





Article

A Classification and Prediction Hybrid Model Construction with the IQPSO-SVM Algorithm for Atrial Fibrillation Arrhythmia

Liang-Hung Wang ¹, Ze-Hong Yan ¹, Yi-Ting Yang ¹, Jun-Ying Chen ¹, Tao Yang ¹, I-Chun Kuo ^{1,*}, Patricia Angela R. Abu ², Pao-Cheng Huang ^{3,*}, Chiung-An Chen ^{4,*} and Shih-Lun Chen ⁵

- ¹ Department of Microelectronics, College of Physics and Information Engineering, Fuzhou University, Fuzhou 350108, China; eetommy@fzu.edu.cn (L.-H.W.); N191120015@fzu.edu.cn (Z.-H.Y.); N181127067@fzu.edu.cn (Y.-T.Y.); N181120016@fzu.edu.cn (J.-Y.C.); yangtao_fzu@fzu.edu.cn (T.Y.)
- ² Department of Information Systems and Computer Science, Ateneo de Manila University, Quezon 1108, Philippines; pabu@ateneo.edu
- ³ College of Computer and Information Sciences, Fujian Agriculture and Forestry University, Fuzhou 350002, China
- ⁴ Department of Electrical Engineering, Ming Chi University of Technology, New Taipei 243303, Taiwan
- ⁵ Department of Electronic Engineering, Chung Yuan Christian University, Taoyuan 320314, Taiwan; chrischen@cycu.edu.tw
- * Correspondence: t20106@fzu.edu.cn (I.-C.K.); paulhuang@fafu.edu.cn (P.-C.H.); joannechen@mail.mcut.edu.tw (C.-A.C.); Tel.: +86-150-0592-2083 (I.-C.K.); +86-150-0508-3450 (P.-C.H.); +886-2-2908-9899 (C.-A.C.)



Citation: Wang, L.-H.; Yan, Z.-H.; Yang, Y.-T.; Chen, J.-Y.; Yang, T.; Kuo, I.-C.; Abu, P.A.R.; Huang, P.-C.; Chen, C.-A.; Chen, S.-L. A Classification and Prediction Hybrid Model Construction with the IQPSO-SVM Algorithm for Atrial Fibrillation Arrhythmia. *Sensors* **2021**, *21*, 5222. <https://doi.org/10.3390/s21155222>

Academic Editor: Paweł Pławiak

Received: 31 March 2021

Accepted: 6 July 2021

Published: 1 August 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Abstract: Atrial fibrillation (AF) is the most common cardiovascular disease (CVD), and most existing algorithms are usually designed for the diagnosis (i.e., feature classification) or prediction of AF. Artificial intelligence (AI) algorithms integrate the diagnosis of AF electrocardiogram (ECG) and predict the possibility that AF will occur in the future. In this paper, we utilized the MIT-BIH AF Database (AFDB), which is composed of data from normal people and patients with AF and onset characteristics, and the AFPDB database (i.e., PAF Prediction Challenge Database), which consists of data from patients with Paroxysmal AF (PAF; the records contain the ECG preceding an episode of PAF), and subjects who do not have documented AF. We extracted the respective characteristics of the databases and used them in modeling diagnosis and prediction. In the aspect of model construction, we regarded diagnosis and prediction as two classification problems, adopted the traditional support vector machine (SVM) algorithm, and combined them. The improved quantum particle swarm optimization support vector machine (IQPSO-SVM) algorithm was used to speed the training time. During the verification process, the clinical FZU-FPH database created by Fuzhou University and Fujian Provincial Hospital was used for hybrid model testing. The data were obtained from the Holter monitor of the hospital and encrypted. We proposed an algorithm for transforming the PDF ECG waveform images of hospital examination reports into digital data. For the diagnosis model and prediction model trained using the training set of the AFDB and AFPDB databases, the sensitivity, specificity, and accuracy measures were 99.2% and 99.2%, 99.2% and 93.3%, and 91.7% and 92.5% for the test set of the AFDB and AFPDB databases, respectively. Moreover, the sensitivity, specificity, and accuracy were 94.2%, 79.7%, and 87.0%, respectively, when tested using the FZU-FPH database with 138 samples of the ECG composed of two labels. The composite classification and prediction model using a new water-fall ensemble method had a total accuracy of approximately 91% for the test set of the FZU-FPH database with 80 samples with 120 segments of ECG with three labels.

Keywords: ECG signal; atrial fibrillation; support vector machine; image-to-data; prediction

1. Introduction

Cardiovascular disease is the leading cause of death worldwide [1], and the mortality rate from this type of disease continues to increase. Atrial fibrillation (AF) is one of the

most common cardiovascular diseases. Clinically, AF can be divided into paroxysmal AF (PAF), persistent AF, and permanent AF according to the time of onset [2]. Before onset, PAF has no significant difference in its ECG and normal ECG, and the time of onset is short and is thus constantly misdiagnosed. Therefore, many researchers have explored various algorithms for analyzing AF.

In 1983, Mood et al. [3] used the Markov process model to calculate the average error value of the prediction array for the RR interval and identify the AF signal of the ECG data. In 2006, Petrucci et al. pointed out that when AF occurs, the RR interval in the ECG has asymmetric distribution characteristics [4]. In 2018, Tang et al. [5] proposed a hybrid Taguchi-genetic algorithm that uses the Gaussian function to decompose the P wave in the ECG signal to improve the performance of the disease classifier. Nurmaini et al. [6] combined discrete wavelet transform and one-dimensional CNN, and the accuracy, sensitivity, and specificity of its classifier reached 99.17%, 98.90%, and 99.17%, respectively.

In the prediction problem, the most important part is to find some features that can predict the occurrence of subsequent phenomena. However, no single feature is found in the prediction of AF with high prediction accuracy [7]. The characteristic of P waves is one of the features. In 2000, Andrikopoulos et al. [8] found that the characteristics of P waves in patients with AF were significantly higher than those in the control group. The P wave variability was defined and became a classic parameter for studying AF with a basis on the characteristics of atrial activity. Lepage et al. [9] cut the P waves after the waveform recognition and combined wavelet transform with the Marco Polo model to improve the specificity and sensitivity.

Overall, the algorithms for AF analysis can be divided into traditional machine learning and deep learning methods. Deep learning algorithms can carry out end-to-end learning, and their capacities depend on their structures. In 2016, Abdul-Kadir et al. [10] constructed an AF classification model with a neural network containing three fully interconnected layers where the inputs were extracted ECG features. The accuracy of the model was 95.3% in the MIT-BIH database. In the 2017 PhysioNet/Computing in Cardiology Challenge, Rubin et al. [11] trained a convolutional neural network (CNN) with a Densely Connected block and utilized 8528 ECG records as inputs. It ranked third among 80 teams in the competition, reporting an F1 score of 0.82. Attia et al. [12] developed a CNN AF prediction model with 10 residual blocks. The data were obtained from 180,922 patients with AF for over 23 years. The prediction accuracy of the model reached 83.3% for the test set. Wang [13] proposed an automatic AF detection method based on an 11-layer neural network. The first eight layers were composed of a CNN, and the last three layers were composed of an improved Elman neural network. The data were obtained from NSR, arrhythmia, and AF data sets from MIT-BIH. The accuracy of the model classification was 97.4%.

Deep learning methods exhibit higher accuracy in the classification of heart diseases than traditional machine learning methods [14–17], but they have the disadvantages of depending on large data sets, which are difficult to obtain in the field of medical data analysis and have poor interpretability. Deep learning methods require digital databases collected by advanced ECG equipment. The currently available ECG network database has a limited amount of data, and ECG data collected by many hospitals for decades cannot be effectively used for artificial intelligence (AI) heart disease diagnosis due to format concern problems. In contrast, the traditional method is not subject to these restrictions, and the performance of the model can also reach a considerable level [18–23].

The SVM has been recognized as an efficient tool for classifying data in a hyper-plane [24] and is widely used in ECG analysis. Li et al. [25] constructed two special intuitive grid maps and used the RR interval as the main feature for classifying AF and normal ECG with linear SVM. The sensitivity, specificity, and accuracy of the model were 95.3%, 96.3%, and 95.9%, respectively. Czabanski et al. [26] designed an LSVM classifier with sixteen input features extracted from HRV signals, which achieved the training of a few samples with an excellent classification performance; the classifier had an accuracy rate

of 98.86%. Nuryani et al. [27] tested different SVM structures, used radial basis function (RBF), and examined two ECG features as SVM inputs (the average of RR intervals and standard deviation of the RR intervals). The final sensitivity, specificity, and accuracy were 95.81%, 98.44%, and 97.50%, respectively. In general, SVM has a better performance compared to other traditional machine learning methods [28,29] and can be generalized well even with small samples [30]. The related works are listed in Table 1.

Table 1. Summary of related work in recent years.

Works	Year	Model	Features	Methods
Abdul-Kadir et al. [10]	2016	ANN	ECG features	Deep learning method
Rubin et al. [11]	2018	CNN	ECG records	Deep learning method
Attia et al. [12]	2019	CNN	ECG records	Deep learning method
Wang [13]	2020	CNN/ENN	ECG records	Deep learning method
Cai et al. [14]	2020	DDNN	ECG records	Deep learning method
Cao et al. [15]	2020	ResNet	ECG records	Deep learning method
Li et al. [25]	2017	SVM	RR interval	Traditional machine learning method
Czabanski et al. [26]	2020	LSVM	HRV	Traditional machine learning method
Nuryani et al. [27]	2015	SVM	RR interval	Traditional machine learning method
Nuryani et al. [29]	2017	SVM	RR interval	Traditional machine learning method

In this study, the proposed model was trained and tested on two databases, the MIT-BIH AF database (AFDB) [31] and PAF Prediction Challenge Database (AFPDB) [32,33], and we constructed a two-stage classified-prediction model of AF. The FZU-FPH database was established by Fuzhou University laboratory and Fujian Provincial Hospital [34,35], which was used to test the performance of the model for clinical verification. The analysis of AF was regarded as a classification problem, thus selecting the SVM as the classifier. To test the clinical applicability of the model, a method for converting the PDF ECG image of the hospital into digital data is proposed that can expand the database for model training and testing.

Contributions: This research has various contributions in the domain of intelligence analysis of atrial fibrillation.

1. The proposed classification and prediction hybrid model uses a new waterfall ensemble method. The first step executed the heartbeat classification, classifying normal and AF categories, and then PAF prediction was performed in which the hybrid model was more suitable for clinical applications, with improved accuracy, and realized an intelligent analysis of AF.
2. Second, the research focused on classified features and introduced wavelet changes in the features of AF.
3. Third, research was conducted on improved quantum particle swarm optimization SVM (IQPSO-SVM) training acceleration and used cross-validation to compare the improved acceleration algorithm with the traditional Grid-Search SVM optimization method.
4. Lastly, the proposed algorithm was converted the clinical ECG to digital data, which were acquired from the Holter monitor in the hospital; Thus, the data of the AF database can be greatly increased. The model was validated on the FZU-FPH database, which further reflects the clinical generalization ability of the model.

The paper is organized as follows: Section 2 introduces the materials and methods. Section 3 presents the architecture of the model and explains it in detail in terms of pre-processing, feature extraction, model construction, and clinical data set processing method. The performance evaluation results, corresponding comparison with other published studies, and discussion are shown in Sections 4 and 5. Finally, Section 6 concludes the paper.

2. Materials and Methods

The ECG signals of AF were mainly obtained from three databases. The first database was the MIT-BIH AF database (AFDB). This database contains long-term ECG records of 25 patients with AF. Each record has a duration of 10 h and contains two-channel signals, and each record is sampled at a rate of 250 points per second with a 12-bit resolution at ± 10 mV. Its label includes the following: AFIB (AF), AFL (atrial flutter), J (AV nodules), and N (all other rhythms). The database contains the AF ECGs of patients with AF, which were different from the normal ECGs of normal people. The database is indicated as an AF disease if it has an AF label in the 5 min ECG signal. The database includes 1380 AF and 1140 N labels. Therefore, the database was used in training the classification model of AF, that is, in determining whether a segment of ECG is an AF ECG.

The second database is the PAF Prediction Challenge Database (AFPDB) from physionet. This dual-channel ECG database, which was extracted from the Holter monitor, was created in the 2001 Computer in Cardiology Challenge. The database was divided into training and test sets. The record set started with the p tag coming from patients who had PAF but did not have episodes at the time of recording. The record set started with the n tag from healthy people who had no record of PAF at any time. Each data segment is a 5 min recording. Therefore, 50 normal people and 50 PAF patients with their 30 min data have 300 N and 300 PAF labels, respectively. The database was used in distinguishing the normal ECG records of normal people and patients with PAF and training a predictive model of PAF.

The third database is the FZU-FPH database that was jointly established by Fuzhou University laboratory and Fujian Provincial Hospital. The ECG acquisition equipment used was the DMS300-4A with a sampling frequency of 3000, and the multi-lead data were collected from subjects aged 18–70 from Fujian Province with a 4096-Hz Holter monitor. Each record indicates the gender, age, acquisition time, heartbeat label, and diagnosis conclusion of a patient. The database was used for clinical verification of the model.

3. Establishment of the Composite Model

As shown in Figure 1, the overall ECG analysis algorithm for AF included three blocks: atrial diagnosis model, atrial predict model, and clinical verification part. The first two blocks are composed of three stages: pre-processing, feature extraction, and model training. The data were divided into 5 min sections. In the diagnosis model: after pre-processing, which filters out the interference and R-wave detection, only the time-domain feature was extracted because the difference between normal ECG and AF ECG was significant, as shown in Figure 2a,b, respectively, and their scales is shown in Figure 2c. In the predicted model, given that the normal ECG of patients with AF was similar to normal ECG, wave segmentation to extract features, PR interval, P-wave amplitude, and six frequency-domain features was performed. After the feature set was extracted, normalization, standardization, and regularization were applied. The training set of the AFDB database and AFPDB database was used to train a diagnosis model and prediction model, respectively. Then, the SVM model with an improved quantum–particle–swarm optimization algorithm was adopted. In the clinical verification part, a method for converting clinical images to digital data was used. After the feature set processing, the obtained feature file was used as the clinical verification data of the final model for testing the performance.

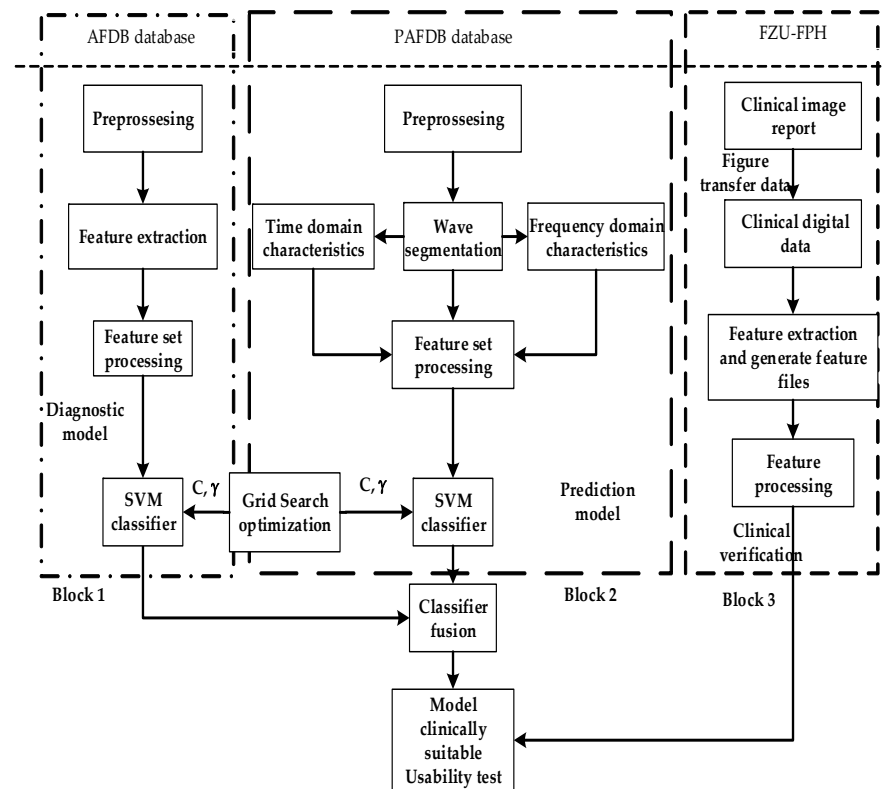


Figure 1. Overall implementation of the ECG analysis algorithm with atrial fibrillation.

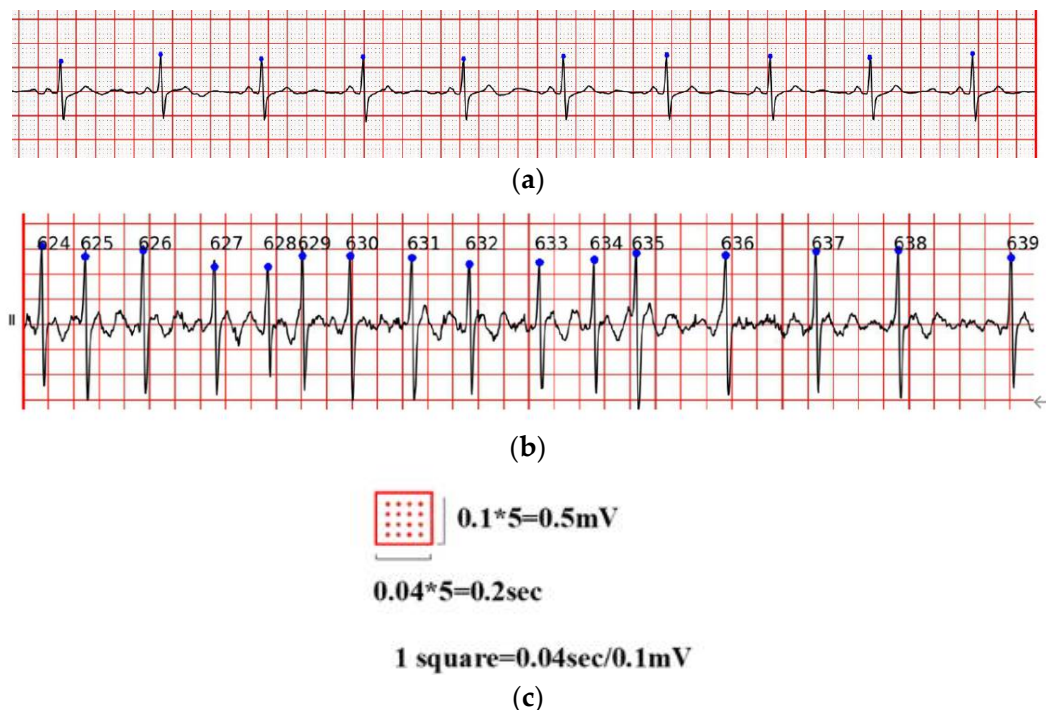


Figure 2. (a) Normal ECG, (b) atrial fibrillation ECG; (c) is the scales of (a,b).

3.1. Pre-Processing

Two steps were used to pre-process the ECG signal: (1) Eliminating the noise interferences to obtain the clear ECG signal, and (2) locating the R wave, QRS complex, and P wave position to analyze the ECG signal information and obtain the ECG features.

The noise interference of the ECG signal has three parts, namely, power frequency noise, baseline drift, and electromechanical (EMG) interference. The low-pass filter was used to filter the T wave interference, power frequency interference, and higher frequency noise of the ECG signal. The high-pass filter was effectively employed to eliminate the EMG interference and baseline drift. The slope of the R wave was amplified by the proposed differentiator to distinguish the QRS complexes. A nonlinear square function was used to square the signal to enhance the slope of the frequency response and reduce the error caused by the T wave. After squaring, the signal was integrated with a window integrator, thus obtaining the slope information of the R wave and the QRS complex. For the R wave location, the improved Pan–Tompkins [36] algorithm was used to eliminate the misdiagnosis point and compensate for the missed detection point. The pre-processing process diagram and flow-chart of the improved R wave location algorithm are shown in Figures 3 and 4, respectively.

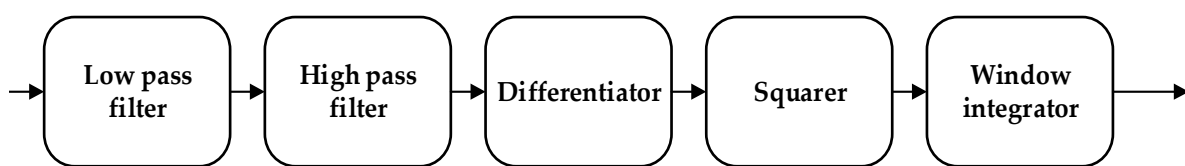


Figure 3. Data pre-processing process.

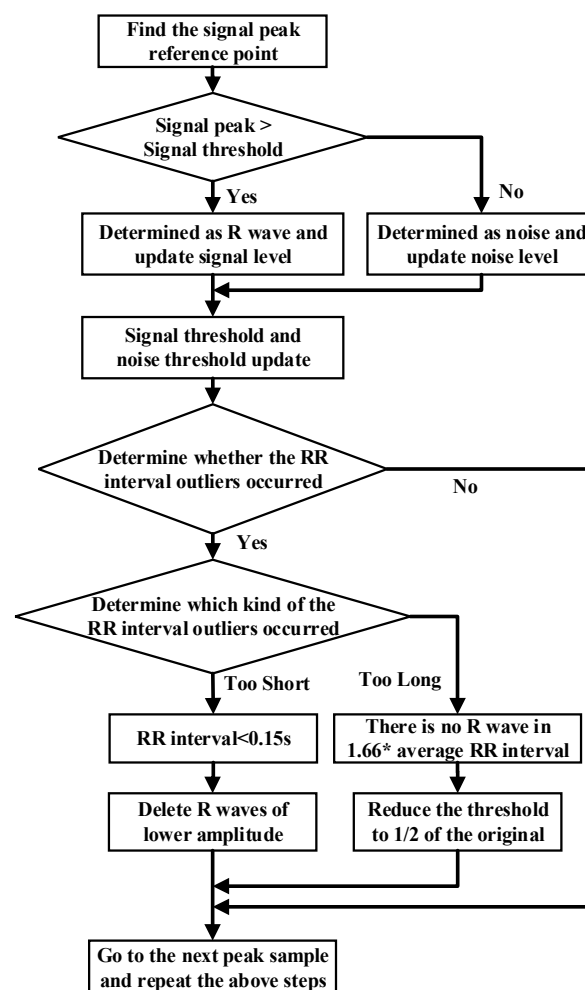


Figure 4. Flow chart of R threshold detection with the dynamic threshold algorithm.

3.2. Feature Extraction

Differences were found between the ECG characteristics of AF and normal conditions:

1. The P wave disappeared and was replaced by AF waves with different shapes, spacing, and amplitudes. AF waves were generally obvious in leads V1 and II. The frequency was 350–600 times/min, and the amplitude was generally 0.05–0.50 mV.
2. The QRS complex morphology is usually normal, but when a ventricular bundle branch block or other diseases occur, the QRS complex widens and the morphology appears abnormal.

When AF occurs, the RR interval (RR Interval, RRI) of the patient becomes irregular. By using the irregularity of RRI in extracting relevant parameters, the diagnosis of AF can be improved. In the related research of atrial activity analysis, the relevant characteristics of P waves are mainly used to represent the changes in atrial activity characteristics, such as P wave variability, which is the square of the standard deviation of the P wave interval. However, owing to low amplitude and high frequency and to the replacement of P waves by irregular f waves, which are confused with noise during the detection process, the extraction of P wave intervals is subject to interference. The accuracy of the algorithm for AF detection with atrial activity analysis is not ideal. The data in the AFDB database were collected by Holter limb leads. Compared with the ECG acquired with standard limb leads, one uses the Mason–Likar lead system and the other uses the Einthoven–Wilson lead system. The lead connection position, acquisition equipment parameters, lead wires, and electrodes are different, which causes the P wave shape and ST segment to be different in the ECG signal acquired by the Holter monitor compared with the standard ECG. Therefore, on the basis of the data in the AFDB database as the training set for training the model and evaluating clinical Holter data, features related to the P wave interval were not used for the purpose of offsetting the poor generalization ability of the model in clinical data diagnosis. The relevant features of the RR interval in the ECG were extracted (such as the average value of the RR interval and the standard deviation of the RR interval) for the construction of a classification model of AF, as the ventricular rate was absolutely uneven. In general, the duration of the RR interval in an ECG is stable at about 0.8 s. However, when AF occurs, irregular excitement appears in various parts of the atrial wall. The specific manifestation is that the RR interval changes are unstable and irregular.

After observing all the data in the database, it was noticed that PAF in many cases appeared after premature beats because RRI would change when premature beats occurred. Therefore, in the selection of the first set of features, RR interval-related features were used, which were consistent with the features used in the diagnosis model. In addition to inter-beat characteristics, the prediction model used intra-beat features, such as PR interval and P-wave amplitude, to compare P-wave time domain differences between patients with AF and those with normal sinus rhythm. According to the pathogenesis and pathophysiology of AF, AF is a rapid stimulation of many focal points produced in the atrium. The frequency of these stimuli in the ECG was greater than the P wave frequency and lower than the QRS compound frequency. Therefore, wavelet transformation was performed on the P band of ECG, and the wavelet coefficients generated after transformation were used as the frequency domain features in the selection of the second set of features. The detail coefficient represented the impact of these rapid stimuli on the atrium. The characteristics of the two AF algorithms are shown in Table 2.

Table 2. Predicted characteristics of atrial fibrillation.

Algorithm Class	Feature Group	Characterization
Atrial fibrillation diagnosis algorithm	Time domain characteristics	RR_std, RR_mean, RR_max, RR_rms, RR_cha, RR_len
Atrial fibrillation prediction algorithm	Time domain characteristics	RR_std, RR_mean, RR_max, RR_rms, RR_cha, RR_len, PR, Pamp
	Frequency domain characteristics	ca1,ca2,ca3, cd1,cd2,cd3

In the selection of time domain features, a series of features of ECG based on the inter-beat interval, intra-beat interval, and amplitude features were extracted, and the first set of discriminant features were generated for the prediction of AF.

Interval features between beats: The interval between heartbeats is the interval between two consecutive heartbeats. The peak point of the R wave was used as the reference point in the conducted series of studies on the relevant characteristics of the RR interval. Interval features within beats: Intracardiac interval is the interval between the post-heartbeat reference point and front reference point. The PR interval required for PAF prediction was based on the peak point of the P wave and the peak point of the R wave, which is the interval between the starting point of the P wave and the starting point of the QRS complex of a given heartbeat. Its characteristic description is shown in Equation (1), where i is the current heartbeat and on represents the starting point of the wave.

$$PR = |QRS(i)_{on} - P(i)_{on}|, \quad (1)$$

Amplitude characteristics: The P wave amplitude was selected as the PAF prediction feature. Equation (2) is the characteristic description, where $ampP$ represents the amplitude of the highest point of P wave amplitude and $ampP_{on}$ represents the amplitude of the starting point of P wave.

$$ampP_{on}P = |ampP - ampP_{on}|, \quad (2)$$

In the selection of frequency domain features, the R wave peak point detected on the signal sample was used as the reference point, and the interval feature was extracted according to the heartbeat. The feature was based on the regional morphology, and the P wave region was extracted using wavelet transform frequency domain characteristics.

A bior5.5 wavelet basis function was used in extracting the P wave features. The biorthogonal (bior Nr. Nd) wavelet is a compactly supported biorthogonal wavelet function, which solves the contradiction between the linear phase and orthogonality requirements. The basic steps of using wavelets for ECG signal processing are as follows:

1. Choose an appropriate wavelet basis function.
2. Decompose the ECG signal into corresponding layers.
3. Keep the coefficients decomposed from a specific layer for the next step of processing.

Given that the frequency span of noise was large, the different dimensions of wavelet transform in processing the signal were used. The process is equivalent to the filtering of signal with different frequencies. Signal characteristics for minimizing the degree of noise interference were obtained.

In general, the wavelet basis function was selected according to the four most obvious characteristics of the wavelet, which are orthogonality, tight support, symmetry, and regularity. Orthogonality is conducive to signal reconstruction. Symmetry reduces the tendency of a signal to be distorted. The tight support length affects the quality of the local characteristics of a signal. The shorter the length, the better the local characteristics. Regularity mainly affects the signal reconstruction smoothing effect. In actual use, the wavelet basis function can be appropriately selected according to the characteristics of

the signal and the required results. A high signal-to-noise ratio (SNR) and correlation coefficient of bior5.5 [37] produce the best wavelet basis function used in ECG signal processing.

The bior5.5 function was used in performing a 5-layer wavelet transform on the n47 ECG signal in the AFPDB database. The result is shown in Figure 5. The uppermost sub-picture in the figure is the original waveform diagram of ECG No. n47, and A1–A5 are the waveform diagrams reconstructed by the approximation coefficients of different layers of ECG No. n47 after wavelet transformation. D1–D5 are the different layers of ECG No. n47 after wavelet transformation. The waveform image was reconstructed using the detail coefficient. With regard to the detailed parameters of the decomposition of the first and second layers, baseline drift and EMG interference, the *R* wave signal of the third layer and the *P* wave signal of the fourth layer were the most obvious parameters. Therefore, bior5.5 was used as the wavelet basis for wavelet transformation on the P-band, and the approximate coefficients and detail coefficients of the fourth layer were extracted after the transformation as frequency domain features.

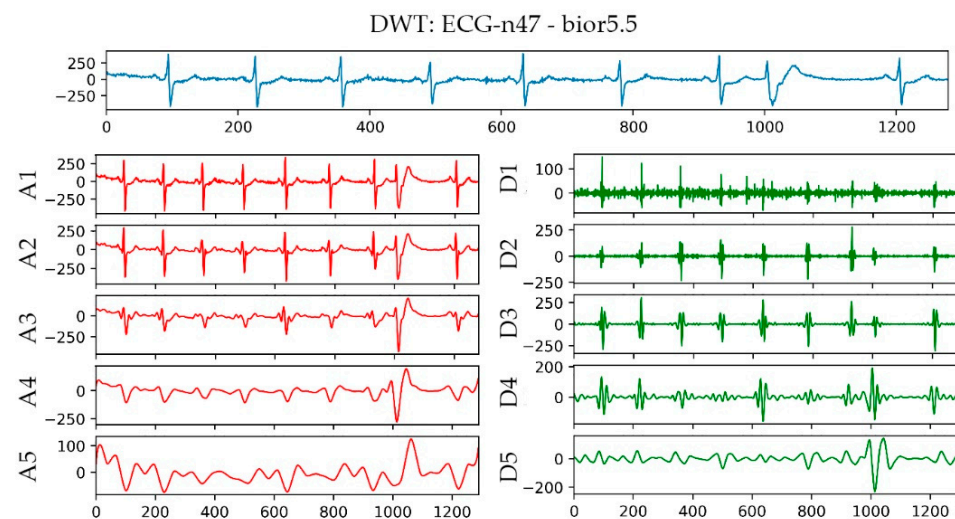


Figure 5. Wavelet transform result graph of No. n47 ECG.

The Decision Tree Classification Algorithm was used to evaluate the importance of features. The comparison in Table 3 lists the importance of the features. The amplitude of the P wave was the most relevant feature for patients with atrial fibrillation.

Table 3. Importance of selected features.

Features	Impact
Pamp	0.268804
cd3	0.220801
RR_std	0.107348
ca3	0.087449
RR_cha	0.081577
PR	0.065737
RR_max	0.058773
ca2	0.027738
cd1	0.021885
RR_rms	0.018379
cd2	0.018140
RR_len	0.012802
ca1	0.010556

3.3. Construction and Verification of the Atrial Fibrillation Model

3.3.1. Construction of the Atrial Fibrillation Model

SVM was used in constructing the AF classifier and PAF predictor. SVM is a two-class classification model and its basis is a linear classifier that maximizes the separation in a feature space for the separation of hyperplanes. The classifier can identify an optimal hyperplane and maximize the interval between two different types of data. In this study, the decision function $f(x)$ was used to represent the optimal hyperplane learned from the training set and was used in evaluating the classification of the prediction set. The discriminant decision function is expressed as shown in Equation (3):

$$f(x) = \text{sign}\left(\sum_{i=1}^n \vec{\alpha}_i^* y_i K(\vec{x}, \vec{x}_i) + b^*\right), \quad (3)$$

In the above equation, $\vec{\alpha}^*$ is the Lagrangian multiplier that satisfies the optimal solution of the SVM classification problem, and b^* is the shortest distance to the optimal hyperplane. SVM with radial basis function was used as the kernel function in classifying and predicting ECG signals. The discriminant decision function is provided as Equation (4):

$$f(x) = \text{sign}\left(\sum_{i=1}^n \vec{\alpha}_i^* y_i (\exp(-\gamma \|\vec{x}_i - \vec{x}\|^2)) + b^*\right), \quad (4)$$

When a data set is linearly inseparable, SVM uses the kernel function to calculate and recognize ECG with the different characteristics of the input ECG signal for the classification of AF ECG and sinus ECG and differentiation between the non-onset ECG of PAF and normal ECG.

3.3.2. Improved Quantum-Particle-Swarm-Optimization Support Vector Machine

The Particle-Swarm-Optimization algorithm is essential to swarm intelligence and is mainly used in solving global optimization problems. The algorithm mainly imitates the behavior of bird and fish swarms to enable particles to find global optimal solutions. However, it has two main disadvantages: first, owing to the lack of randomness in the transformation of its particle position, it is easy to fall into the trap of local optimization. Second, it needs to set additional parameters, which will have a certain impact on the optimization of parameters in a composite model. In view of these two shortcomings, an Improved Quantum-Particle-Swarm-Optimization SVM (IQPSO-SVM) was introduced, which can effectively improve the efficiency of SVM parameter optimization without reducing the accuracy of the model. In the classical physical world, the trajectory of a particle can be described by a velocity vector and a position vector. However, in the quantum world, according to the Heisenberg uncertainty principle, the position and velocity of a particle cannot be determined at the same time. If particles during particle swarm optimization have quantum behavior, they will run in different ways. First, the state of a particle is described by the wave function rather than the previous position and velocity. Through the Monte Carlo method, the iterative equation of particle motion can be obtained.

$$\begin{cases} x_i(t+1) = p + \lambda \cdot |Mbest_i - x_i(t)| \cdot \ln(1/u), & \text{if } k \geq 0.5 \\ x_i(t+1) = p - \lambda \cdot |Mbest_i - x_i(t)| \cdot \ln(1/u), & \text{if } k < 0.5, \end{cases} \quad (5)$$

where λ is the shrinkage-expansion coefficient, also known as the inertia weight or innovation coefficient, u and k are values generated by using the uniform probability distribution

function in the range of [0,1], and M_{best} is the average optimal position of all particles, which is obtained by calculating the mean of c with its Equation is as follows:

$$M_{best} = \frac{1}{N} \sum_{i=1}^N p_{besti}, \quad (6)$$

Inertia weight λ is an important parameter that affects the convergence of IQPSO and is usually considered a fixed value. We improved it in order that λ can change dynamically with the algorithm.

$$\lambda = \lambda_{\max} - \frac{\lambda_{\max} - \lambda_{\min}}{t_{\max}} \times t, \quad (7)$$

where λ_{\max} and λ_{\min} are the upper and lower bounds of inertia weight and t is the number of iterations. In general, the result is better when λ decreases from 1 to 0.5.

In the part of verifying IQPSO, we performed threefold cross-validation to optimize the penalty coefficient C of the SVM and the kernel function parameter gamma with the grid optimization method. The process result is shown in Figure 6.

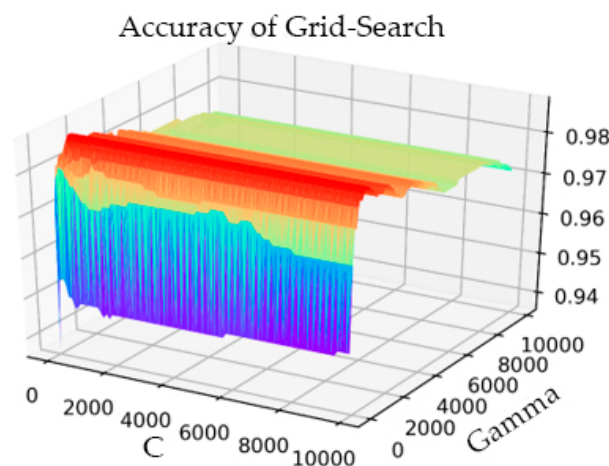


Figure 6. Grid Search parameter optimization process diagram.

The setting range of C was 1–10000, the step was 100, and the gamma range was the same as that of C . After the transformation of C and gamma, the accuracy of the model prediction changed. When the parameter C was 101 and gamma was 801, the accuracy rate (Acc) was at its highest, having a value of 0.9873. The simulation time was 1142.914 s. In the training process of the AF classification model, for the same training data, the IQPSO-SVM algorithm designed in this study was used to determine the optimal parameters. Given that the algorithm introduced the average best position M_{best} , the ability of the particles to cooperate with each other was enhanced. Moreover, the randomness of the quantum space made it difficult for the algorithm to enter local convergence, and the convergence speed of the model was significantly faster than the existing particle swarm optimization (PSO) and SVM using the grid search. As shown in Figure 7, when the accuracy rate reached 0.9875 after iteration, the recorded processing time was 124.4241 s, which was close to only 1/10 of the processing time of the Grid Search method. The values of the relevant parameters are as follows: C , the range of gamma was 0.001–10000, the number of particle swarms was 20, the number of iterations was 30, and the innovation parameter was 0.6.

Experiments were conducted for the cross-validation of different fold numbers, 3-fold, 5-fold, and 10-fold cross-validation, and comparison of the performance of Grid Search and IQPSO-SVM. The detailed classification accuracy and model tuning time-consuming experimental results are listed in Table 4. The results are represented by the average and deviation. In the case of cross-validation with different fold numbers, the IQPSO algorithm took less time to train a better model. When 10-fold cross-validation was performed, the recorded processing time was 1/20 of the process time of the Grid Search.

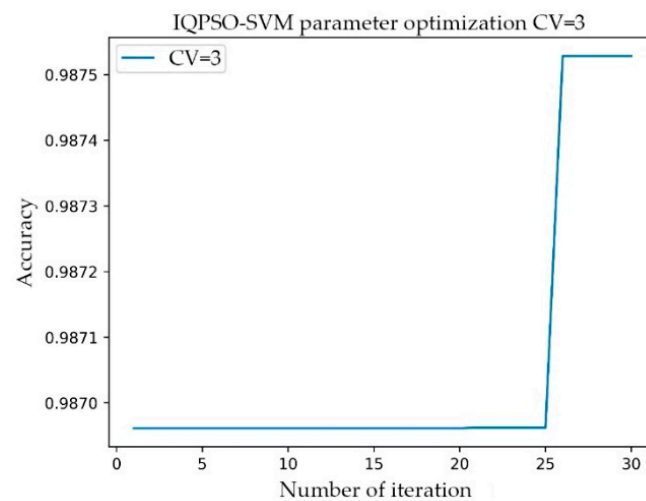


Figure 7. Relationship between the number of iterations and accuracy of IQPSO-SVM.

Table 4. Grid Search and IQPSO-SVM performance comparison.

Cross-Validation	Grid Search		IQPSO-SVM	
	Accuracy	Time(s)	Accuracy	Time(s)
3-fold	0.9836 ± 0.0037	1137.865 ± 5.312	0.9860 ± 0.0018	125.126 ± 1.012
5-fold	0.9810 ± 0.0011	1558.365 ± 6.352	0.9847 ± 0.0009	156.344 ± 2.895
10-fold	0.9835 ± 0.0033	7550.963 ± 5.813	0.9879 ± 0.0010	363.591 ± 3.552

3.3.3. Clinical Image Data to Digital Data

To test the clinical applicability of the model, part of the Holter report from the FZU-FPH database was used. The report was a 5 min two-lead image in a 6-page PDF file. Through a series of PDF to data operations, the report was converted into clinical data. The flow chart illustrating the conversion of clinical PDF images to digital data is shown in Figure 8.

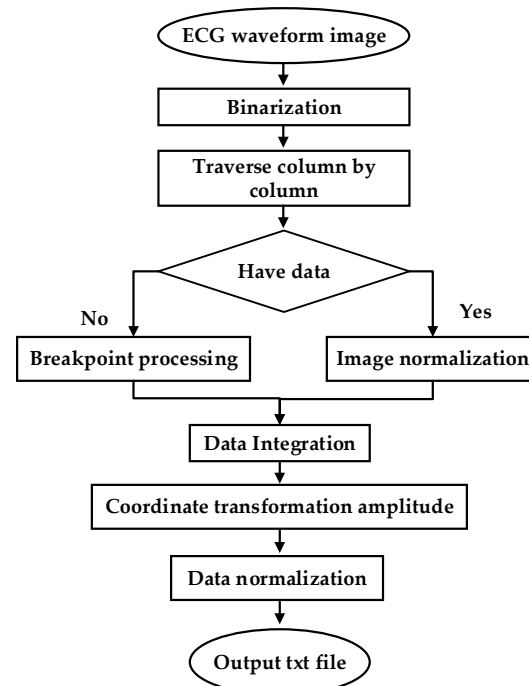


Figure 8. ECG waveform image to data.

ECG is often difficult to share because of its encrypted feature. Its main form is a paper ECG report, and the main storage format of ECG is a PDF file. In this study, we used the PDF2IMG interface of C++ in image conversion and comprehensively considered the distortion of the ECG image. Finally, the normalized image resolution was determined to be 6614×9354 ; three channels were obtained, and the format used was PNG. For the accurate extraction of the ECG waveform information, background and foreground curves should be separated. This process produces a good digital effect. Based on the OTSU method [38], a process that uses an improved OTSU's algorithm for separation was proposed. Gamma transformation was performed on the image before the OTSU. This process greatly improved the success rate of curve extraction. In the separation of the curve from the grid, the effective data in this part were eliminated for the generation of breakpoints. This process was necessary because the effective ECG curve and grid intersected. The method involving traversing column by column to add points was adopted. Figure 9a shows the extracted image, and Figure 9b shows the original Holter image.

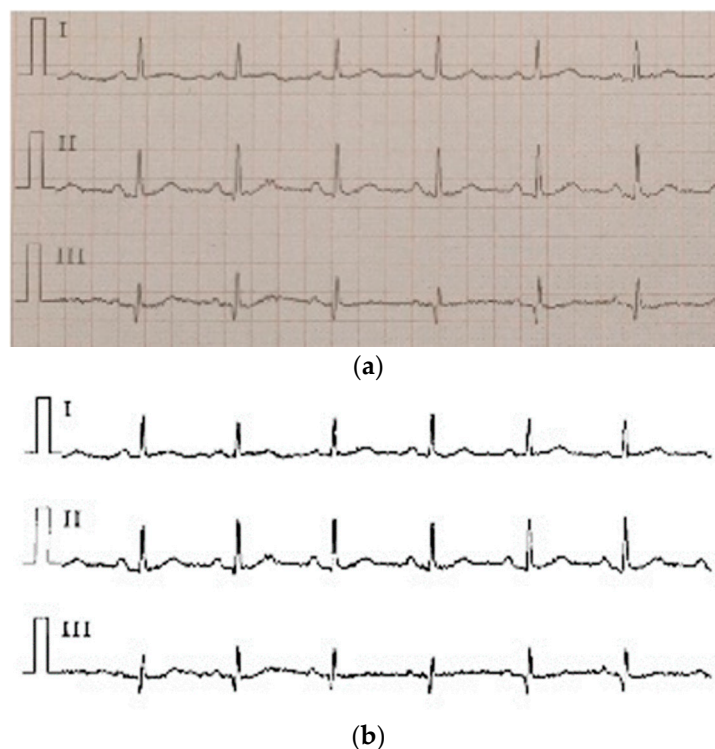


Figure 9. (a) the extracted image, (b) the original Holter image.

4. Experimental Results and Performance Analysis

The data of the MIT-BIH AF database were used in training the AF classification model. Before training, the experimental data were divided into two groups: a training set and a test set. The training and test sets contained a normal sinus ECG and AF ECG with a balanced amount of data. When SVM is used in training a model, the usual process involves the following procedures: standardization of sample data and samples mapped with a kernel function (the kernel function mostly uses RBF and is linear). When the sample is linearly separable, the linear effect is better than RBF. Given that the research sample in this study was a nonlinear sample, the RBF Gaussian was used as the kernel function, and cross-validation and the Grid Search grid optimization algorithm were used in adjusting the hyperparameters. Finally, the optimal parameters were used in training for the generation of the final model, and the model was tested.

In terms of the performance evaluation of the SVM classification model of AF, accuracy (*ACC*), specificity (*SP*), sensitivity (*SE*), and precision (*P*) were used as the evaluation criteria. The calculation formula is shown in Equations (8)–(10):

$$Acc = \frac{TP + TN}{TP + TN + FP + FN} \quad (8)$$

$$SP = \frac{TN}{TN + FP} \quad (9)$$

$$SE = Recall(R) = \frac{TP}{TP + FN} \quad (10)$$

where *TP* represents the proportion of positive samples classified as positive samples, *TN* represents the negative samples classified as negative samples, *FP* represents the proportion of negative samples classified as positive samples, and *FN* is the proportion of positive samples classified as negative samples. Accuracy is the proportion of samples with correct predictions relative to all samples. Specificity represents the proportion of all negative examples that are correctly classified and measures the ability of the classifier to recognize negative examples. Sensitivity represents the proportion of all positive examples that are correctly classified and measures the ability of the classifier to recognize positive examples.

In the AFDB database, the 10 h ECG records of 21 patients were cut down to 5 min records, and 2520 5 min ECG signal segments were obtained. There were 15 patients with 1800 segments in the training set and six patients with 720 segments in the test set. As long as the AF label appeared in the 5 min ECG signal, it was marked as AF disease. The training set included 1020 AF episodes and 780 N episodes. The test set included 360 AF episodes and 360 N episodes. The diagnosis model was trained based on the 1800 segments from the training set and tested with 720 segments from the test set. The classification results obtained are shown in Table 5, in which the x axes and y axes indicate the predicted value and actual value, respectively, and are listed in Tables 6–8.

Table 5. Classification results of the diagnosis model in the AFDB test set.

	AF	N
AF	357	3
N	3	357

Table 6. Classification results of the prediction model in the AFPDB test set.

	PAF	N
PAF	57	3
N	4	56

Table 7. Classification results of the prediction models in the FZU-FPH test set.

	PAF	N
PAF	65	4
N	8	61

Table 8. Classification results of the fusion model in the FZU-FPH test set.

	N	PAF	AF
N	35	5	0
PAF	2	37	1
AF	0	1	39

In the AFPDB database, the 30 min data of the normal sinus ECG data of 50 normal people and 50 patients with paroxysmal atrial fibrillation before the onset of the disease were cut down to obtain 600 5 min ECG signal segments. The 5 min ECG signals of normal sinus ECG data were labelled as the N category and others were labelled as PAF. There were 80 people (i.e., 40 normal people and 40 patients) with 480 segments and 20 people (i.e., 10 normal people and 10 patients) with 120 segments that were used as the training set and test set, respectively. Thus, the training set of the prediction model included 240 PAF episodes and 240 N episodes, and the test set included 60 PAF episodes and 60 N episodes. The classification results are shown in Table 6. Since the input signal of the prediction model is a normal ECG signal, the difficulty of its classification is greatly increased. In order to increase the number of test samples and verify the performance of the prediction model, 138 5 min segments from the FZU-FPH database (i.e., 69 ECG signals of normal people and 69 paroxysmal atrial fibrillation patients) were employed to test the performance of the prediction model. The result is listed in Table 7.

For the intelligent analysis of AF by the model, the classification model and predictive model were integrated through the waterfall hybrid method, and these two models were hybridized in a series. The model identification after hybridization is shown in Figure 10. In order to test the performance of the hybrid model, an extra 80 samples with 120 segments were selected from the FZU-FPH, which are different from the 138 samples mentioned above. The ECG signals at the onset of atrial fibrillation, 20 to 30 min before the onset of atrial fibrillation, and normal ECG signals were labelled as AF, PAF, and N diseases, respectively. The hybrid model first classifies whether the signal is the ECG at the onset of atrial fibrillation or not, namely, AF and N categories. In the N category, the predictive model of atrial fibrillation plays a role, identifying whether it is a normal ECG of patients with paroxysmal atrial fibrillation and classifying the N category into N (i.e., normal person) and PAF (i.e., paroxysmal atrial fibrillation patients).

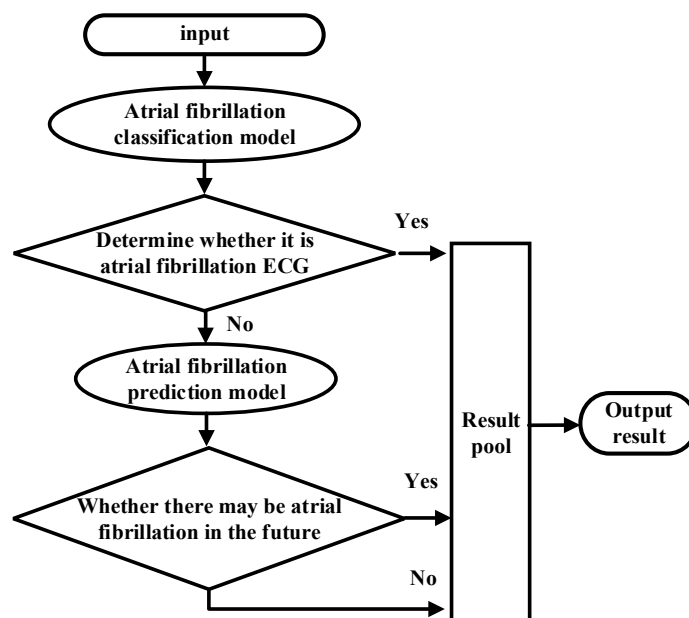


Figure 10. Flow chart of the waterfall fusion algorithm.

A hybrid model was used to classify and predict the N, AF, and PAF categories after the classification and prediction models were integrated into a hybrid model. The classification results between N and AF categories were significantly different, and the accuracy of the model for N and AF classification was 100%. However, the morphology and features of PAF were between N and AF, and the classification results of PAF were easy to misclassify to the N or AF category. The impact of misclassification of PAF into N was much greater than that of N into PAF. The latter can be cancelled by subsequent

medical monitoring. The classification results of fusion models are shown in Table 8, and the performance of different models in different test sets is shown in Table 9.

Table 9. Performance of different models in different test sets.

Performance of Atrial Fibrillation Classification Models in the AFDB Test Set.				
		SE	SP	ACC
AFDB Test Set	AF	99.2%	99.2%	99.2%
AFPDB Test Set	PAF	93.3%	91.7%	92.5%
FZU-FPH Clinical Test Set of Two Labels	PAF	94.2%	79.7%	87.0%
FZU-FPH Clinical Test Set of Three Labels	PAF	92.5%	90%	90.8%
	AF	97.5%	98.8%	98.3%

The calculation methods for the accuracy, sensitivity, and specificity of the three labels are described in Table 9, and the confusion matrix is shown in Table 10. Take N samples as an example, the true positives of N samples are actually the 5 min segments with N labels, and the model recognizes exact N labels, $TP_N = N_N$. The false positives of the N -type samples indicate that they are not actually the N -type labels, but are predicted to be N -type labels by the model, $FP_N = PAF_N + AF_N$. The true negatives of the N -type samples mean that the actual values are not the N -type labels and the model is not mistakenly classified into the N -type labels, $TN_N = PAF_{PAF} + PAF_{AF} + AF_{PAF} + AF_{AF}$. The false negatives of the N -type samples are the actual value of the N -type labels, but they are predicted to be another label, $FN_N = N_{PAF} + N_{AF}$. The classification accuracy of N samples is $N_{ACC} = (TP_N + TN_N) / (TP_N + TN_N + FP_N + FN_N)$. The accuracy of AF labels and PAF labels can be calculated in the same way, and the sensitivity and specificity can be achieved after recognizing the corresponding TP, TN, FP, and FN correctly.

Table 10. Confusion matrix-related definition.

	N	PAF	AF
N	N_N	N_{PAF}	N_{AF}
PAF	PAF_N	PAF_{PAF}	PAF_{AF}
AF	AF_N	AF_{PAF}	AF_{AF}

5. Discussion

A method for automatically detecting AF episodes was described in this study. The method provides a novel means of extracting features from different databases with SVM. The performance of the proposed method was evaluated using the AFDB, AFPDB, and FZU-FPH databases. Fed with 14 features describing time and frequency domains, the SVM classifier had a sensitivity of 99.2%, specificity of 99.2%, and accuracy of 99.2% for the diagnosis model in the AFDB database test set. A sensitivity of 93.3%, specificity of 91.7%, and accuracy of 92.5% were achieved for the prediction model used in the AFPDB database test set. The particle swarm optimization algorithm considerably improved the simulation time and showed good performance. This research combined the diagnosis and prediction of AF and combined the AI algorithm with AF analysis. The training set and test set of the above-mentioned classification and prediction hybrid model were 5 min ECG signals, which were cut from 30 min ECG signals, and were used to train and test the hybrid model. In order to verify the clinical generalization ability of the model, the FZU-FPH database was used.

Table 11 shows the comparison between the performances of the AF-classified model and the previous research models. As indicated in Table 11, the previous works were tested in the AFDB database except literature sources [8,9]. The proposed method demonstrated excellent discrimination with sensitivity, specificity, and accuracy all exceeding 99% for the AFDB test set. Compared with the performance of the previous traditional algorithm,

the performance in the test set of the AFDB database improved. For the use of time-domain-related features, Tateno et al. [37], Li et al. [25], Andersen et al. [18], and Kumar et al. [21] used RRI or some time-domain relative features. The above four literature sources [18,21,25,37] all used the AFDB database as their training material and test set. In 2019, Czabanski et al. [26] used HR features based on RRI and extracted the features of HRV combined with LSVM; the sensitivity, specificity, and accuracy were 98.94%, 98.39%, and 98.66%, respectively. The training heartbeats only played a little part in all heartbeats, but some test set came from the same subject in the training set. For features that use P waves, when deep learning methods are not used, the model performance is greatly reduced. The performance of the model increases with the number of combined features. In [6], the author combined discrete wavelet transform and a one-dimensional convolutional neural network. The training and test material came from four different databases. The classification of the model was excellent but had the disadvantages mentioned before, that is, the amount of data required is relatively large. The classification model proposed in this study can achieve an acceptable accuracy while the training data are insufficient. Through the selection of signal processing algorithms, optimization of a series of algorithms in the processing process and the features selected for research was conducted.

Table 11. Performance comparison of atrial fibrillation classification models.

Works	Characteristic	Methods/Database	SE	SP	ACC
This work	RRI	SVM/AFDB	99.2%	99.2%	99.2%
Tateno et al. [37]	RRI	Coefficient of Variation/AFDB	94.4%	97.2%	—
Li et al. [25]	RRI	LSVM/AFDB	95.9%	95.3%	96.3%
Kumar et al. [21]	ECG features	Random forest/AFDB	95.8%	97.8%	96.8%
Andersen et al. [18]	RRI ECG features	SVM/AFDB	96.81%	96.20%	96.45%
Czabanski et al. [26]	HR features	LSVM/AFDB	98.94%	98.39%	98.66%
Andrikopoulos et al. [8]	P wave	Statistical methods/Self-built database	88%	75%	—
Lepage et al. [9]	P wave	Markov Models/Self-built database	70%	65%	—
Nurmaini et al. [6]	RRI P wave	CNN DWT/AFDB	99.91%	99.91%	99.98%

The sensitivity, specificity, and accuracy of the PAF prediction model for the divided test set of the AFPDB database were 93.3%, 91.7%, and 92.5%, respectively. Table 12 shows the comparison between the performance of the PAF prediction model and the previous research model. All studies used AFPDB databases as training and testing material. In [39–41], a 30 min ECG signal was used for training and for extracting relevant features. Dozens of features are related to HRV, and different modes of feature training cause differences in model performance. In [42,43], 5 min ECG signals were used for model training. The short ECG sequence contained limited information. The importance of feature selection increases with the number of requirements for the classification performance of a model. In [44,45], both studies extracted HRV relative features as the model input, and showed an excellent performance. However, the number of subjects in their training set and test set was insufficient. A verification step was added in the subjects in the FZU-FPH database, in which the subjects and data of the test set were from 69 ECG signals of normal people and 69 paroxysmal atrial fibrillation patients in hospitals. In [12], the database of Team A was twenty years of data collected with partner hospitals. The amount of data was large, and the overall accuracy rate was 83.3%. The database demonstrated strong clinical generalization capabilities. The sensitivity, specificity, and accuracy in the FZU-

FPH database were 94.2%, 79.7%, and 87.0%, which has a better sensitivity compared to literature [12].

Table 12. Performance comparison of predictive models for paroxysmal atrial fibrillation.

Works	Signal Length (min)	Feature/Methods/Database	SE	SP	ACC
This work (AFPDB)	5	P wave RRI/SVM/AFPDB	93.3%	91.7%	92.5%
This work (FZU-FPH)	5	P wave RRI/SVM/FZU-FPH	94.2%	79.7%	87.0%
Costin et al. [40]	30	HRV and PACs /Statistical analysis/AFPDB	89.3%	89.4%	89.4%
Mohebbi et al. [41]	30	HRV/SVM/AFPDB	96.2%	93.1%	94.5%
Boon et al. [39]	30	HRV/SVM/AFPDB	81.1%	79.3%	80.2%
Xin et al. [43]	5	HRV multi-scale wavelet entropy/SVM/AFPDB	92.18%	94.88%	89.48%
Boon et al. [42]	5	HRV/SVM/AFPDB	86.8%	88.7%	87.7%
Elias et al. [44]	5	HRV/ME/AFPDB	100%	95.55%	98.21%
Parsi et al. [45]	5	HRV/SVM/AFPDB	98.8%	96.7%	97.7%
Attia et al. [12]	—	... /CNN/Self-built database	82.3%	83.4%	83.3%

To evaluate the clinical generalization ability of the model, we used the model in clinical trials in which the clinical test set data were obtained from the PDF classification reports including 69 segments of PAF and 69 segments of normal control data. The data came from 69 paroxysmal atrial fibrillation and 69 normal subjects in the FZU-FPH database. Each data point had a duration of 5 min. The prediction model had better sensitivity, less misjudgment, and low specificity in predicting PAF, and normal ECG was easily misdiagnosed as sinus ECG of PAF. However, the clinical impact of this misjudgment was smaller than that of sinus ECG in patients with PAF. Therefore, the model has a certain reference and application value in the clinic.

6. Conclusions

Despite serious medical consequences, AF is still an underestimated clinical and classification problem. The analysis of AF should be intelligent and reduce the burden on doctors. The normal classification of the ECG of patients with AF is a significant research topic since it can provide timely warning for patients with PAF. In this study, starting from the positioning of the reference point of the ECG II-lead signal, the improved dynamic threshold algorithm was used for the accurate extraction of the R wave. On this basis, the relevant features that include six time-domain features were extracted, as well as the model for the automatic classification model of AF. After a series of investigations on the precursors of PAF, the AF prediction from eight time-domain, and six frequency-domain features, we used an improved quantum particle swarm optimization (IQPSO) algorithm combined with SVM to construct a model for the effective prediction of AF. Finally, the classification and prediction hybrid model was established based on the SVM algorithm using a new waterfall ensemble method. This classifier first diagnosed AF on the input ECG data and then identified AF ECG, and non-AF ECG. The sensitivity, specificity, and accuracy were recorded as 99.2%, 99.2%, and 99.2%, respectively, in the AFDB test set. The model analyzed the factors of primary AF, identified the ECG and normal ECG during the onset period of paroxysmal fibrillation and provided early warning for hidden patients. The sensitivity, specificity, and accuracy were 93.3%, 91.7%, and 92.5%, respectively, in the AFPDB test set and 94.2%, 79.7%, and 87.0%, respectively, in the FZU-FPH clinical test set. The classification accuracy of the hybrid model after the three labels of data were used as an input was approximately 91%. The limitation of this research is focused on the feature of

atrial fibrillation without comparison with other classifiers. Future improvements include an end-to-end approach and reducing the manual intervention.

Author Contributions: Conceptualization, L.-H.W., Z.-H.Y., T.Y., C.-A.C. and P.-C.H.; Data curation, L.-H.W. and Z.-H.Y.; Formal analysis, L.-H.W. and Y.-T.Y.; Funding acquisition, L.-H.W., C.-A.C. and P.-C.H.; Investigation, Z.-H.Y. and S.-L.C.; Methodology, L.-H.W., Z.-H.Y., Y.-T.Y., J.-Y.C. and T.Y.; Project administration, L.-H.W.; Resources, L.-H.W., J.-Y.C., I.-C.K. and P.-C.H.; Software, Z.-H.Y., Y.-T.Y. and P.A.R.A.; Supervision, I.-C.K. and C.-A.C.; Validation, L.-H.W. and Z.-H.Y.; Writing—original draft, L.-H.W., Z.-H.Y., P.A.R.A. and P.-C.H.; Writing—review & editing, L.-H.W., P.A.R.A., C.-A.C. and P.-C.H. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported in part by the National Natural Science Foundation of China, the Major Project of Ministry of Science and Technology, and the Health and Education Project of Fujian Province, under Grant Nos. 61971140, 2020IM010200, 2019-WJ-18. And the Ministry of Science and Technology (MOST), Taiwan, under grant numbers of MOST-108-2622-E-033-012-CC2, MOST-109-2622-E-131-001-CC3, MOST-109-2221-E-131-025, MOST 110-2823-8-033-002.

Conflicts of Interest: The authors declare no conflict of interest.

References

- Alonso, A.; Almuwaqqat, Z.; Chamberlain, A. Mortality in atrial fibrillation. Is it changing? *Trends Cardiovasc. Med.* **2020**. [[CrossRef](#)]
- Margulescu, A.D.; Mont, L. Persistent atrial fibrillation vs paroxysmal atrial fibrillation: Differences in management. *Expert Rev. Cardiovasc. Ther.* **2017**, *15*, 601–618. [[CrossRef](#)] [[PubMed](#)]
- Moody, G. A new method for detecting atrial fibrillation using RR intervals. *Comput. Cardiol.* **1983**, *10*, 227–230.
- Petrucci, E.; Balian, V.; Filippini, G.; Mainardi, L. The use of sequential RR distributions to detect atrial fibrillation episodes in very long term ECG monitoring. In Proceedings of the 2006 Computers in Cardiology, Valencia, Spain, 17–20 September 2006; pp. 945–948.
- Tang, W.-H.; Chang, Y.-J.; Chen, Y.J.; Ho, W.-H. Genetic algorithm with Gaussian function for optimal P-wave morphology in electrocardiography for atrial fibrillation patients. *Comput. Electr. Eng.* **2018**, *67*, 52–57. [[CrossRef](#)]
- Nurmaini, S.; Tondas, A.E.; Darmawahyuni, A.; Rachmatullah, M.N.; Umi Partan, R.; Firdaus, F.; Tutuko, B.; Pratiwi, F.; Juliano, A.H.; Khoirani, R. Robust detection of atrial fibrillation from short-term electrocardiogram using convolutional neural networks. *Future Gener. Comput. Syst.* **2020**, *113*, 304–317. [[CrossRef](#)]
- Kennedy, J.; Eberhart, R. Particle swarm optimization. In Proceedings of the ICNN'95-International Conference on Neural Networks, Perth, WA, Australia, 27 November–1 December 1995; pp. 1942–1948. [[CrossRef](#)]
- Andrikopoulos, G.K.; Dilaveris, P.E.; Richter, D.J.; Gialafos, E.J.; Synetos, A.G.; Gialafos, J.E. Increased variance of P wave duration on the electrocardiogram distinguishes patients with idiopathic paroxysmal atrial fibrillation. *Pacing Clin. Electrophysiol.* **2000**, *23*, 1127–1132. [[CrossRef](#)] [[PubMed](#)]
- Lepage, R.; Boucher, J.-M.; Blanc, J.-J.; Cornilly, J.-C. ECG segmentation and P-wave feature extraction: Application to patients prone to atrial fibrillation. In Proceedings of the 2001 Conference Proceedings of the 23rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Istanbul, Turkey, 25–28 October 2001; pp. 298–301. [[CrossRef](#)]
- Abdul-Kadir, N.A.; Safri, N.M.; Othman, M.A. Dynamic ECG features for atrial fibrillation recognition. *Comput. Methods Programs Biomed.* **2016**, *136*, 143–150. [[CrossRef](#)] [[PubMed](#)]
- Rubin, J.; Parvaneh, S.; Rahman, A.; Conroy, B.; Babaeizadeh, S. Densely connected convolutional networks for detection of atrial fibrillation from short single-lead ECG recordings. *J. Electrocardiol.* **2018**, *51*, S18–S21. [[CrossRef](#)]
- Attia, Z.I.; Noseworthy, P.A.; Lopez-Jimenez, F.; Asirvatham, S.J.; Deshmukh, A.J.; Gersh, B.J.; Carter, R.E.; Yao, X.; Rabinstein, A.A.; Erickson, B.J. An artificial intelligence-enabled ECG algorithm for the identification of patients with atrial fibrillation during sinus rhythm: A retrospective analysis of outcome prediction. *Lancet* **2019**, *394*, 861–867. [[CrossRef](#)]
- Wang, J. A deep learning approach for atrial fibrillation signals classification based on convolutional and modified Elman neural network. *Future Gener. Comput. Syst.* **2020**, *102*, 670–679. [[CrossRef](#)]
- Cai, W.; Chen, Y.; Guo, J.; Han, B.; Shi, Y.; Ji, L.; Wang, J.; Zhang, G.; Luo, J. Accurate detection of atrial fibrillation from 12-lead ECG using deep neural network. *Comput. Biol. Med.* **2020**, *116*, 103378. [[CrossRef](#)]
- Cao, X.-C.; Yao, B.; Chen, B.-Q. Atrial fibrillation detection using an improved multi-Scale decomposition enhanced residual convolutional neural network. *IEEE Access* **2019**, *7*, 89152–89161. [[CrossRef](#)]
- Erdenebayar, U.; Kim, H.; Park, J.-U.; Kang, D.; Lee, K.-J. Automatic prediction of atrial fibrillation based on convolutional neural network using a short-term normal electrocardiogram signal. *J. Korean Med. Sci.* **2019**, *34*, 7. [[CrossRef](#)]
- Oh, S.L.; Ng, E.Y.; Tan, S.R.; Acharya, U.R. Automated diagnosis of arrhythmia using combination of CNN and LSTM techniques with variable length heart beats. *Comput. Biol. Med.* **2018**, *102*, 278–287. [[CrossRef](#)] [[PubMed](#)]

18. Andersen, R.S.; Poulsen, E.S.; Puthusserypady, S. A novel approach for automatic detection of Atrial Fibrillation based on Inter Beat Intervals and Support Vector Machine. In Proceedings of the 2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Jeju Island, Korea, 11–15 July 2017; pp. 2039–2042. [[CrossRef](#)]
19. Gliner, V.; Yaniv, Y. An SVM approach for identifying atrial fibrillation. *Physiol. Meas.* **2018**, *39*, 094007. [[CrossRef](#)] [[PubMed](#)]
20. Kennedy, A.; Finlay, D.D.; Guldenring, D.; Bond, R.R.; Moran, K.; McLaughlin, J. Automated detection of atrial fibrillation using RR intervals and multivariate-based classification. *J. Electrocardiol.* **2016**, *49*, 871–876. [[CrossRef](#)] [[PubMed](#)]
21. Kumar, M.; Pachori, R.B.; Acharya, U. Automated diagnosis of atrial fibrillation ECG signals using entropy features extracted from flexible analytic wavelet transform. *Biocybern. Biomed. Eng.* **2018**, *38*, 564–573. [[CrossRef](#)]
22. Pierleoni, P.; Belli, A.; Gentili, A.; Incipini, L.; Palma, L.; Valenti, S.; Raggiunto, S. A eHealth System for Atrial Fibrillation Monitoring. In Proceedings of the Italian Forum of Ambient Assisted Living, Lecce, Italy, 2–4 July 2018; pp. 229–241.
23. Wrobel, J.; Horoba, K.; Matonia, A.; Kupka, T.; Henzel, N.; Sobotnicka, E. Optimizing the automated detection of atrial fibrillation episodes in long-term recording instrumentation. In Proceedings of the 2018 25th International Conference “Mixed Design of Integrated Circuits and System” (MIXDES), Gdynia, Poland, 21–23 June 2018; pp. 460–464.
24. Cortes, C.; Vapnik, V. Support-vector networks. *Mach. Learn.* **1995**, *20*, 273–297. [[CrossRef](#)]
25. Li, Y.; Tang, X.; Wang, A.; Tang, H. Probability density distribution of delta RR intervals: A novel method for the detection of atrial fibrillation. *Australas. Phys. Eng. Sci. Med.* **2017**, *40*, 707–716. [[CrossRef](#)]
26. Czabanski, R.; Horoba, K.; Wrobel, J.; Matonia, A.; Martinek, R.; Kupka, T.; Jezewski, M.; Kahankova, R.; Jezewski, J.; Leski, J.M. Detection of atrial fibrillation episodes in long-term heart rhythm signals using a support vector machine. *Sensors* **2020**, *20*, 765. [[CrossRef](#)]
27. Nuryani, N.; Harjito, B.; Yahya, I.; Lestari, A. Atrial fibrillation detection using support vector machine. In Proceedings of the Joint International Conference on Electric Vehicular Technology and Industrial, Mechanical, Electrical and Chemical Engineering (ICEVT & IMECE), Surakarta, Indonesia, 4–5 November 2015; pp. 215–218. [[CrossRef](#)]
28. Meyer, D.; Leisch, F.; Hornik, K.J.N. The support vector machine under test. *Neurocomputing* **2003**, *55*, 169–186. [[CrossRef](#)]
29. Nuryani, N.; Harjito, B.; Yahya, I.; Solikhah, M.; Chai, R.; Lestari, A. Atrial fibrillation detection using support vector machine and electrocardiographic descriptive statistics. *Int. J. Biomed. Eng. Technol.* **2017**, *24*, 225–236. [[CrossRef](#)]
30. Duin, R.P. Classifiers in almost empty spaces. In Proceedings of the Proceedings 15th International Conference on Pattern Recognition. ICPR-2000, Barcelona, Spain, 3–7 September 2000; pp. 1–7. [[CrossRef](#)]
31. Moody, G.B.; Mark, R.G. The impact of the MIT-BIH Arrhythmia Database. *IEEE Eng. Med. Biol. Mag.* **2001**, *20*, 45–50. [[CrossRef](#)]
32. Goldberger, A.; Amaral, L.A.; Glass, L.; Hausdorff, J.; Ivanov, P.C.; Mark, R.G.; Mietus, J.E.; Moody, G.B.; Peng, C.K.; Stanley, H.E. *Components of a New Research Resource for Complex Physiologic Signals*; American Heart Association: Dallas, TX, USA, 2000; Volume 101, Issue 23.
33. Moody, G.; Goldberger, A.L.; McClennen, S.; Swiryn, S.P. Predicting the onset of paroxysmal atrial fibrillation: The Computers in Cardiology Challenge 2001. In Proceedings of the Computers in Cardiology, Rotterdam, The Netherlands, 23–26 September 2001; Volume 28 (Cat. No. 01CH37287), pp. 113–116. [[CrossRef](#)]
34. Fan, M.-H.; Guan, M.-H.; Chen, Q.-C.; Wang, L.-H. Three-lead ECG detection system based on an analog front-end circuit ADS1293. In Proceedings of the 2017 IEEE International Conference on Consumer Electronics-Taiwan (ICCE-TW), Taipei, Taiwan, 12–14 June 2017; pp. 107–108. [[CrossRef](#)]
35. Wang, L.-H.; Chen, T.-Y.; Lin, K.-H.; Fang, Q.; Lee, S.-Y. Implementation of a wireless ECG acquisition SoC for IEEE 802.15. 4 (ZigBee) applications. *IEEE J. Biomed. Health Inform.* **2014**, *19*, 247–255. [[CrossRef](#)]
36. Pan, J.; Tompkins, W.J. A real-time QRS detection algorithm. *IEEE Trans. Biomed. Eng.* **1985**, 230–236. [[CrossRef](#)]
37. Tatenko, K.; Glass, L. Automatic detection of atrial fibrillation using the coefficient of variation and density histograms of RR and Δ RR intervals. *Med. Biol. Eng. Comput.* **2001**, *39*, 664–671. [[CrossRef](#)]
38. Otsu, N. A threshold selection method from gray-level histograms. *IEEE Trans. Syst. Man Cybern.* **1979**, *9*, 62–66. [[CrossRef](#)]
39. Boon, K.H.; Khalil-Hani, M.; Malarvili, M.; Sia, C.W. Paroxysmal atrial fibrillation prediction method with shorter HRV sequences. *Comput. Methods Programs Biomed.* **2016**, *134*, 187–196. [[CrossRef](#)]
40. Costin, H.; Rotariu, C.; Păsărică, A. Atrial fibrillation onset prediction using variability of ECG signals. In Proceedings of the 2013 8th International Symposium on Advanced Topics in Electrical Engineering (ATEE), Bucharest, Romania, 23–25 May 2013; pp. 1–4. [[CrossRef](#)]
41. Mohebbi, M.; Ghassemian, H. 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.* **2012**, *105*, 40–49. [[CrossRef](#)]
42. Boon, K.H.; Khalil-Hani, M.; Malarvili, M. Paroxysmal atrial fibrillation prediction based on HRV analysis and non-dominated sorting genetic algorithm III. *Comput. Methods Programs Biomed.* **2018**, *153*, 171–184. [[CrossRef](#)] [[PubMed](#)]
43. Xin, Y.; Zhao, Y. Paroxysmal atrial fibrillation recognition based on multi-scale wavelet alpha-entropy. *Biomed. Eng. Online* **2017**, *16*, 121. [[CrossRef](#)]
44. Ebrahimzadeh, E.; Maede, K.; Joulani, M.; Shahraki, R.S.; Fayaz, F.; Ahmadi, F. Prediction of paroxysmal Atrial Fibrillation: A machine learning based approach using combined feature vector and mixture of expert classification on HRV signal. *Comput. Methods Programs Biomed.* **2018**, *165*, 53–67. [[CrossRef](#)]
45. Parsi, A.; Glavin, M.; Jones, E.; Byrne, D. Prediction of Paroxysmal Atrial Fibrillation using New Heart Rate Variability Features. *Comput. Biol. Med.* **2021**, *133*, 104367. [[CrossRef](#)] [[PubMed](#)]