



## Ensemble Learning for Detection of Short Episodes of Atrial Fibrillation

Peimankar, Abdolrahman; Puthusserypady, Sadasivan

*Published in:*  
Proceedings of 2018 26th European Signal Processing Conference

*Link to article, DOI:*  
[10.23919/EUSIPCO.2018.8553253](https://doi.org/10.23919/EUSIPCO.2018.8553253)

*Publication date:*  
2018

*Document Version*  
Peer reviewed version

[Link back to DTU Orbit](#)

*Citation (APA):*  
Peimankar, A., & Puthusserypady, S. (2018). Ensemble Learning for Detection of Short Episodes of Atrial Fibrillation. In *Proceedings of 2018 26th European Signal Processing Conference* (pp. 66-70). IEEE.  
<https://doi.org/10.23919/EUSIPCO.2018.8553253>

---

### General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

# Ensemble Learning for Detection of Short Episodes of Atrial Fibrillation

Abdolrahman Peimankar

Department of Electrical Engineering  
Technical University of Denmark  
2800 Kgs. Lyngby, Denmark  
apeima@elektro.dtu.dk

Sadasivan Puthusserypady

Department of Electrical Engineering  
Technical University of Denmark  
2800 Kgs. Lyngby, Denmark  
spu@elektro.dtu.dk

**Abstract**—Early detection of atrial fibrillation (AF) is of great importance to cardiologists in order to help patients suffer from chronic cardiac arrhythmias. This paper proposes a novel algorithm to detect short episodes of atrial fibrillation (AF) using an ensemble framework. Several features are extracted from long term electrocardiogram (ECG) signals based on the heart rate variability (HRV). The most significant subset of features are selected as inputs to the four classifiers. Outputs of these classifiers are then combined for the final detection of the AF episodes. Results from an extensive analysis of the proposed algorithm show high classification accuracy (around 85%) and sensitivity (around 92%) for classifying very short episodes of AF (10 beats per segment, which is approximately 6 seconds). The accuracy and sensitivity of the proposed algorithm are improved significantly to 96.46% and 94%, respectively, for slightly longer episodes (60 beats per segment) of AF. Compared to the state-of-the-art algorithms, the proposed method shows the potential to pave the way to extend to real-time AF detection applications.

**Index Terms**—Electrocardiogram (ECG), Ensemble learning, Atrial fibrillation, Feature selection, Classification.

## I. INTRODUCTION

A significant proportion of worldwide mortality are caused by cardiac diseases and atrial fibrillation (AF) is one of the most common cardiac arrhythmias among elderly population [1], [2]. The demographics of western countries is alarming this health issue [3], [4]. It should be mentioned that the consequences of AF is much more serious than AF itself since it may end up in serious heart failures and strokes [5]. From an economic point of view, treatment of AF can be very expensive as well as challenging [4], [5].

One of the most common ways for physicians to diagnose AF is through visual examination of the electrocardiogram (ECG) recordings. However, it is not always easy and in most cases cumbersome to analyse these big amounts of ECG data. Therefore, it is required to develop analytic software in order to automatically analyse the heart beats and rhythms which helps accelerating the process of detecting AF. Various state-of-the-art algorithms for detecting AF have been introduced in the literature. Zhou et al. used symbolic dynamics and Shannon entropy for automatic detection of AF [6]. Dash et al. proposed a real time AF detection algorithm using heart beat intervals and tested their method on two different databases [7]. In [8], the possibility of AF detection using P wave absence was studied. In a recent study, Andersen et al. investigated AF detection using data driven and HRV

approaches [9]. Asgari et al. proposed an algorithm using statistical modelling of atrial activity and a single classifier to detect AF [10]. An automatic detection of AF using discrete wavelet transform and HRV was also proposed in [11]. The purpose of this work is to develop an intelligent algorithm that will overcome inaccuracies and uncertainties that exist in the state-of-the-art diagnosis methodologies.

Ensemble learning methods are being used for prognostics and decision making in various applications, such as biomedical [12], financial [13], and power systems engineering [14]–[16]. Three main parts of ensemble learning systems are [17]: (i) Sampling from a dataset to make a training set, (ii) Training a group of classifiers, and (iii) Combining the output of classifiers. It has been shown in the literature that using ensemble learning increases the chance of selecting more accurate classifiers by avoiding selection of a single weak classifier [17].

Four different HRV based features have been derived and used as inputs to the four chosen classification algorithms to distinguish the AF episodes from normal sinus rhythm (NSR). The four used classifiers are, the Random Forests (RF) [18], Support Vector Machine (SVM) [19], Adaptive Boosting (AdaBoost) [20], and Group Method of data Handling (GMDH) [21]. Each of these classifiers is trained separately using 5-fold cross validation and then the outputs of all the classifiers are combined using a combination rule called the Dempster-Shafer theory (DST) [22] to enhance the classification accuracy.

## II. MATERIALS AND METHODS

### A. Preprocessing

The recorded ECG signals are filtered to remove noise and baseline wanders. The well-known PhysioNet WFDB Toolbox is also used to detect the  $R$  peaks and extract the RR intervals (RRI) [23], [24]. The signals are then segmented into  $M$  beats ( $R$  peaks) per segment.

### B. Feature Extraction

Four HRV based features are extracted from each RRI segment. They are: (i) The standard deviation of RRI in a segment (SD\_RRI), (ii) mean of RRI in a segment (Mean\_RRI), (iii) Root mean square of successive differences (RMSSD), and (iv) Normalised root mean square of successive differences (nRMSSD) [25]–[27]. SD\_RRI and Mean\_RRI measure

the standard deviation and mean of RRIs in each segment. RMSSD (Eq.(1)) and nRMSSD (Eq.(2)) are statistical time-domain based methods to estimate HRV in a specified intervals of an ECG.

$$RMSSD = \sqrt{\frac{1}{N-1} \sum_{i=1}^{N-1} (RR_{i+1} - RR_i)^2} \quad (1)$$

$$nRMSSD = \frac{RMSSD}{\overline{Mean\_RRI}}, \quad (2)$$

where  $N$  is the number of RRIs in a segment. Due to the irregular heart beats in AF episodes, the values of RMSSD and nRMSSD are usually higher.

### C. Classification

1) *Support Vector Machine(SVM)*: It is one of the popular statistical based learning algorithms which seeks optimum hyper-planes to classify a dataset into different classes or approximate a function. Suppose that we have a dataset of  $N$  inputs and targets as:  $Z = \{(\mathbf{x}_1, t_1), (\mathbf{x}_2, t_2), \dots, (\mathbf{x}_N, t_N)\}$ , where  $\mathbf{x}_n \in \mathbb{R}^m$  and  $t_n \in \mathbb{R}$  are inputs vectors and targets, respectively. The algorithm uses this dataset to approximate the function  $f(\mathbf{x})$ , which maps inputs to targets, as follows [19]:

$$f(\mathbf{x}) \simeq \sum_{n=1}^N (\mathbf{w}^T \mathbf{x}_n + b), \quad (3)$$

where  $\mathbf{w}$  denotes the weights vector and  $b$  is the bias term. The separating hyper-plane can be completely defined by determining  $\mathbf{w}$  and  $b$ . In order to take the advantage of non-linear decision boundaries, a radial basis function (RBF) kernel is used, which is defined as:

$$K(\mathbf{x}_n, \mathbf{x}_m) = \exp\left(\frac{-\|\mathbf{x}_n - \mathbf{x}_m\|^2}{2\sigma^2}\right), \quad (4)$$

where  $\sigma$  is the width of the kernel function which needs to be optimized. There is another hyper-parameter,  $C$ , which controls the mis-classification's penalty to achieve a trade-off between minimising the error and the model complexity. A grid search approach is used to find the optimum values of these hyper-parameters.

2) *Group Method of Data Handling (GMDH)*: This method, which is also known as the polynomial neural network, was introduced by Ivakhnenko in 1971 [28]. Here, the relationship between multiple inputs and outputs of the network can be modelled as:

$$\hat{Y}(\mathbf{u}) = a_0 + \sum_{i=1}^L a_i f_i(\mathbf{u}), \quad (5)$$

where  $\mathbf{u} = (u_1, \dots, u_L)$  is the input feature vector,  $\hat{Y}$  is the output. In Eq.(5),  $a_0$  and  $a_i$  are scalar coefficients,  $f_i$  are the elementary functions, and  $L$  is the number of base function components in the GMDH network.

In the GMDH algorithm, various subsets of Eq.(5), which are called partial-models, are defined. The coefficients ( $a_0, a_i, a_{ij}, a_{ijk}, \dots$  in Eq.(6)) of these partial-models are

determined using least-squares techniques [28]. The core concept of GMDH is to find a model (network) with optimal complexity by gradually increasing the partial-models. In our work, we use one of the most well-known base functions called the Kolmogorov-Gabor polynomial defined in the following equation:

$$\begin{aligned} \hat{Y}(\mathbf{u}) = & a_0 + \sum_{i=1}^L a_i u_i + \sum_{i=1}^L \sum_{j=1}^L a_{ij} u_i u_j \\ & + \sum_{i=1}^L \sum_{j=1}^L \sum_{k=1}^L a_{ijk} u_i u_j u_k + \dots \end{aligned} \quad (6)$$

3) *Random Forests (RF)*: RF is an ensemble learning approach which constructs different decision trees during the training phase and then make a prediction based on the majority voting. It was first introduced by Tin Kam Ho [29] in 1995 and was extended by Leo Breiman in 2001 [30]. Decision trees are invariant and stable models to transformation of feature inputs and adding irrelevant features [31]. In addition, training deep decision trees easily over-fits the training set [31]. In contrast, different deep decision trees are trained on different parts of the training set to alleviate the overfitting problem.

4) *Adaptive Boosting (AdaBoost)*: It is another important ensemble learning algorithm introduced by Yoav Freund and Robert Schapire in 1997 [20]. Compared to RF, AdaBoost is less likely to over-fit but is very sensitive to noisy data. It also uses decision trees as the weak learners and weighs the mis-classification samples in the next iterations where new decision trees are added to the ensemble.

### D. Dempster-Shafer Combination Rule

The Dempster-Shafer theory (DST) is a powerful technique for information fusion and combining information from different sources [22]. This method is also capable of capturing the degree of certainties from different information sources [32].

The three main functions used in DST are: a mass probability function ( $m$ ), a belief function ( $Bel$ ), and a plausibility function ( $Pl$ ) of which  $m$  is the most important in binary classification problems and should meet the following conditions [32]:

$$\begin{aligned} m & : 2^X \rightarrow [0, 1], \\ m(\emptyset) & = 0, \\ \sum_{A \subseteq X} m(A) & = 1, \end{aligned} \quad (7)$$

where  $X$  is the universal set and  $\emptyset$  is the empty set. For our application, the universal set is defined as:  $X = \{\text{NSR}, \text{AF}\}$ .

One advantage of DST is its capability for combining independent evidences (mass probability functions),  $m_1$  and  $m_2$ , in order to produce more informative evidence, which is shown by  $m_1 \oplus m_2$  and is calculated as follows:

$$(m_1 \oplus m_2)(A) = \frac{1}{1 - P} \sum_{B \cap C = A \neq \emptyset} m_1(B) m_2(C), \quad (8)$$

where  $P = \sum_{B \cap C = \emptyset} m_1(B)m_2(C)$ . In the proposed ensemble method (Section III), the outputs of the single classifiers are actually normalized mass functions between 0 and 1 which can be used by DST to compute a combined output.

### III. ENSEMBLE AF DETECTION

The proposed method utilizes the advantages of ensemble learning to detect short episodes of AF in ECG recordings. The flowchart of the method is illustrated in Fig. 1 and the algorithm is described in eight steps as follows:

- 1) *Noise reduction*: The recorded ECG signals are filtered to remove the noise and baseline wanders.
- 2) *R peaks detection*: As mentioned in Section II-A, the R peaks in all ECG signals are detected using the PhysioNet WFDB Toolbox.
- 3) *Segmentation*: In this step, the ECG signals are segmented into  $M$  beats per segment.
- 4) *Create a synthetic dataset*: In this study, the number of NSR episodes in the dataset are much more than the number of AF episodes. Thus, the size of AF class is not sufficient for training the classification algorithms. Therefore, the use of this imbalanced dataset increases the chance of a biased classification, which in turns leads to a higher error rate on the minority class [33]. To improve the training procedure of the classifiers by making the dataset balanced, adaptive synthetic over-sampling technique (ADASYN) is applied to enable the classification algorithms achieving their desirable performance [33]. The ADASYN algorithm consists of three main steps: (i) computing the degree of class imbalance to calculate the number of synthetic samples for the minority class, (ii) calculating the Euclidean distance to find the  $K$  nearest neighbours in a minority class, and (iii) generating the synthetic dataset for the minority class by:  $d_i = x_i + \lambda(x_{k_i} - x_i)$ , where  $x_i$  is a minority class sample,  $x_{k_i}$  is a randomly chosen sample from the determined  $K$  nearest neighbors, and  $\lambda \in [0, 1]$  is a random number.
- 5) *Cross validation*: In this step, the whole dataset is divided into training and validation sets using the 5-fold cross validation.
- 6) *Train all classifiers*: The 5-fold training sets are used to train all the single classifiers to estimate the average error.
- 7) *Evaluate all the trained classifiers*: The trained classifiers are tested on the validation sets to evaluate the performance of each classification algorithms.
- 8) *Combine the outputs of the classifiers*: The outputs of all the classifiers which are posterior probabilities assigned to the two classes (AF & NSR) are combined using DST.

### IV. EXPERIMENTAL VALIDATION

In this study, we have used the Long-Term Atrial Fibrillation Database (LTAfDB) on PhysioNet [23] to verify the performance of the developed algorithm. This database consists of 84 ECG recordings of length between 24 to 25 hours and sampled at 128 Hz. The imbalanced LTAfDB dataset is composed

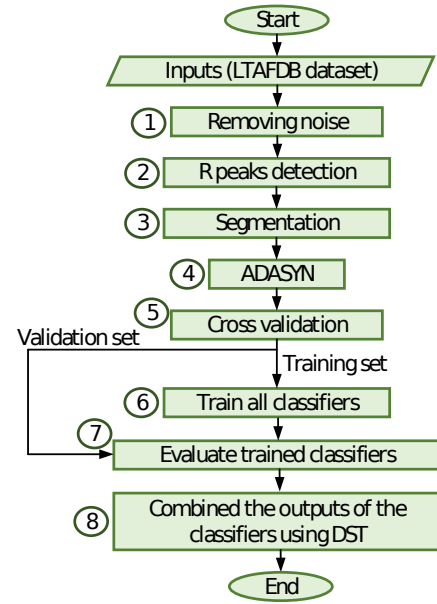


Fig. 1. Flowchart of the proposed algorithm.

of 893618 beats. The number of beats for the two classes are represented in Table I. After creating synthetic data, the number of beats for both classes were distributed equally (Table I). As an example, distribution of the imbalanced and balanced number of beats for 30 segments of signal # 2 for two arbitrary features (Mean\_RRI and SD\_RRI) are plotted in Fig. 2a and Fig. 2b, respectively. This shows how synthetic data are distributed with respect to the two features (Mean\_RRI and SD\_RRI).

TABLE I  
NUMBER OF R PEAKS FOR BALANCED AND IMBALANCED DATASETS.

Classes	NSR	AF
Number of beats (imbalanced)	752473	141145
Number of beats (balanced)	752473	757707

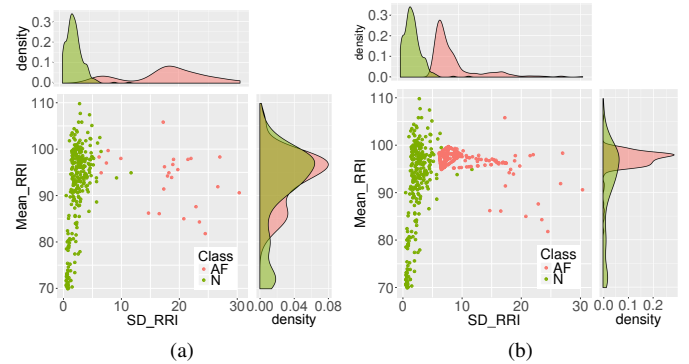


Fig. 2. Scatter plots of R peaks for two arbitrary features: (a) Imbalanced, (b) Synthetically balanced.

The significant features, which are the most efficient ones in separating the AF class from NSR class, are selected as inputs

to the classifiers. Amongst the four extracted features (Section II-B), the three most efficient features are the SD\_RRI, RMSSD, and nRMSSD. Figures 3a and 3b show the scatter plots along with the density curves of four features for signal # 2 in the dataset. It is evident that the density curves for these three features have less overlaps and using these three features leads to a better classification between AF and NSR classes.

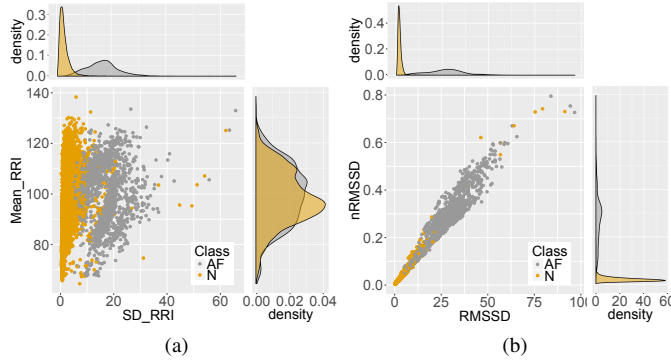


Fig. 3. Scatter plots along with kernel densities for signal # 2 of length 23 hours. Each data point represents an extracted feature for one segment.

In the next step, the selected features are used as inputs to the classifiers. The four classifiers are trained using 5-fold cross validation techniques. The training and validation accuracies for all four classifiers are reported in Table II. As given in this table, RF performs the best on validation sets followed by AdaBoost, SVM, and GMDH.

TABLE II  
5-FOLD CROSS VALIDATION ACCURACY OF INDIVIDUAL CLASSIFICATION ALGORITHMS.

Algorithms	RF	SVM	AdaBoost	GMDH
Training (%)	<b>90.62</b>	86.54	90.14	87.39
Validation (%)	<b>83.86</b>	78.60	83.08	78.50

The outputs of the four classifiers are then combined using DST which improves the overall classification accuracy of the best classifier by more than 1% (83.86 to 84.97%). The confusion matrices on training and validation phases are also shown in Figs. 4a and 4b. In addition, other classification measures of the combined classifiers such as  $F1$ -score, sensitivity ( $Se$ ), and positive predictive value ( $PPV$ ) are reported in Table III. These metrics are defined as follows:

$$Se = \frac{TP}{TP + FN}, \quad (9)$$

$$PPV = \frac{TP}{TP + FP}, \quad (10)$$

$$F = (1 + \beta^2) \frac{PPV \cdot Se}{(\beta^2 \cdot PPV) + Se}, \quad (11)$$

$$Acc = \frac{TP + TN}{TP + FN + FP + TN}. \quad (12)$$

where  $TP$ ,  $FN$ , and  $FP$  are the number of true positive, false negative, and false positive cases, respectively.  $F$ -score is actually a weighted harmonic mean of  $Se$  and  $PPV$ . When

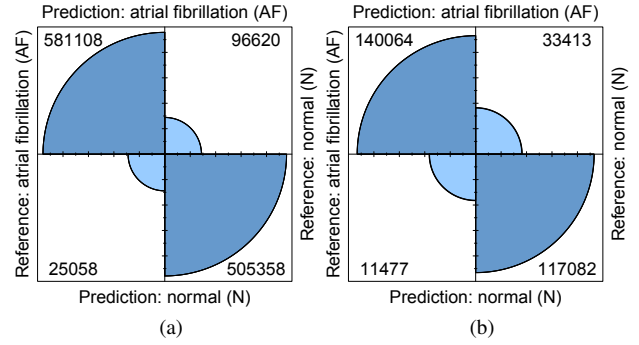


Fig. 4. Confusion matrix. (a) Training. (b) Validation.

TABLE III  
CLASSIFICATION MEASURES OF THE PROPOSED ALGORITHM ON THE TRAINING AND VALIDATION SETS.

Measure	Acc (%)	F1 score	Se (%)	PPV (%)
Training	90	0.90	96.00	86.00
Validation	85	0.86	92.00	81.00

$\beta$  is equal to 1 the measure is called balanced  $F$ -score ( $F1$  score) which takes both  $Se$  and  $PPV$  into account equally. The value of  $F1$  score is between 0 and +1 in which +1 shows the highest classification ability and 0 represents the lowest classification ability.

In addition, the performance of the proposed algorithm was evaluated on the longer segment lengths ( $M = 60, 150$ , and 300 beats per segment) and the corresponding results are given in Table IV. As it is expected, the performance of the proposed algorithm enhances significantly for longer segments. For example, for  $M = 60$ , the  $Acc$ ,  $F1$  score,  $Se$ , and  $PPV$  increase by over 11%, 10%, 2%, and 16%, respectively.

TABLE IV  
COMPARISON OF THE CLASSIFICATION MEASURES FOR DIFFERENT NUMBER OF BEATS PER SEGMENT ON THE VALIDATION SET.

Measure	Acc (%)	F1 score	Se (%)	PPV (%)
$M = 10$	85.00	0.86	92.00	81.00
$M = 60$	96.46	0.96	94.00	97.00
$M = 150$	97.00	0.96	95.00	97.00
$M = 300$	97.86	0.97	96.00	98.00

It should be mentioned that the majority of the state-of-the-art algorithms in the literature considered fairly large segment of beats to detect AF. To the best of our knowledge, this is one of the first studies that investigates detection of AF for a very short segment (10-beat). For example, Dash et al. [7] defined 128-beat segments to train their proposed algorithm. In addition, they excluded AF episodes of less than 64 beats long in the reported results as they stated that their algorithm is not robust enough for very short episodes of AF. In another study, Babaeizadeh et al. [34], which uses RRI and P wave absent as detecting features, achieved a very good performance for segments of longer than 60 beats with  $Se$  and  $PPV$  of 93% and 98%, respectively. It should be noted that there is no

reported results on shorter segments in their study, however, the performance of the proposed algorithm in our paper for  $M=60$  is comparable with them.

## V. CONCLUSION

In this paper, a novel ensemble method has been proposed to detect short episodes of AF in ECG recordings. First, four different features (SD\_RRI, Mean\_RRI, RMSSD, and nRMSSD) were extracted from the derived HRV signals. Then, three of the most significant features, namely the SD\_RRI, RMSSD, and nRMSSD were selected and used them as inputs to the four chosen classification schemes (SVM, GMDH, RF, and AdaBoost). All classification algorithms were trained using 5-fold cross validation. Finally, the outputs of the classifiers which were actually the posterior probabilities assigned to the two classes (AF & NSR) were combined using the DST combination rule to improve the overall accuracy. The very impressive classification results obtained with short segments ( $M = 10$ ) of ECG signals provide us the opportunity to use this algorithm in real-time to detect AF which could eventually be used in house by clinicians to diagnose incipient AF.

## ACKNOWLEDGMENT

This project is supported by the Danish Innovation funded project, REAFEL - Reaching the Frail Elderly patient for optimizing diagnosis of atrial fibrillation (IFD Project No: 6153-00009B). We acknowledge the Innovation Fund Denmark for this support.

## REFERENCES

- [1] Y.-k. Iwasaki, K. Nishida, T. Kato, and S. Nattel, "Atrial fibrillation pathophysiology: implications for management," *Circulation*, vol. 124, no. 20, pp. 2264–2274, 2011.
- [2] V. Markides and R. J. Schilling, "Atrial fibrillation: classification, pathophysiology, mechanisms and drug treatment," *Heart*, vol. 89, no. 8, pp. 939–943, 2003.
- [3] S. Stewart, N. Murphy, A. Walker, A. McGuire, and J. McMurray, "Cost of an emerging epidemic: an economic analysis of atrial fibrillation in the uk," *Heart*, vol. 90, no. 3, pp. 286–292, 2004.
- [4] M. Zoni-Berisso, F. Lercari, T. Carazza, and S. Domenicucci, "Epidemiology of atrial fibrillation: European perspective," *Clinical epidemiology*, vol. 6, p. 213, 2014.
- [5] J.-Y. Le Heuzey, O. Paziand, O. Piot, M. A. Said, X. Copie, T. Lavergne, and L. Guize, "Cost of care distribution in atrial fibrillation patients: the cofac study," *American heart journal*, vol. 147, no. 1, pp. 121–126, 2004.
- [6] X. Zhou, H. Ding, B. Ung, E. Pickwell-MacPherson, and Y. Zhang, "Automatic online detection of atrial fibrillation based on symbolic dynamics and shannon entropy," *Biomedical engineering online*, vol. 13, no. 1, p. 18, 2014.
- [7] S. Dash, K. Chon, S. Lu, and E. Raeder, "Automatic real time detection of atrial fibrillation," *Annals of biomedical engineering*, vol. 37, no. 9, pp. 1701–1709, 2009.
- [8] S. Ladavich and B. Ghoraani, "Rate-independent detection of atrial fibrillation by statistical modeling of atrial activity," *Biomedical Signal Processing and Control*, vol. 18, pp. 274–281, 2015.
- [9] R. S. Andersen, E. S. Poulsen, and S. Puthusserypady, "A novel approach for automatic detection of atrial fibrillation based on inter beat intervals and support vector machine," in *2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, July 2017, pp. 2039–2042.
- [10] S. Asgari, A. Mehrnia, and M. Moussavi, "Automatic detection of atrial fibrillation using stationary wavelet transform and support vector machine," *Computers in biology and medicine*, vol. 60, pp. 132–142, 2015.
- [11] I. H. Bruun, S. M. S. Hissabu, E. S. Poulsen, and S. Puthusserypady, "Automatic atrial fibrillation detection: A novel approach using discrete wavelet transform and heart rate variability," in *2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, July 2017, pp. 3981–3984.
- [12] L. Shi, L. Xi, X. Ma, M. Weng, and X. Hu, "A novel ensemble algorithm for biomedical classification based on ant colony optimization," *Applied Soft Computing*, vol. 11, no. 8, pp. 5674–5683, 2011.
- [13] D. West, S. Dellana, and J. Qian, "Neural network ensemble strategies for financial decision applications," *Computers & operations research*, vol. 32, no. 10, pp. 2543–2559, 2005.
- [14] A. Peimankar, S. J. Weddell, T. Jalal, and A. C. Laphorn, "Multi-objective ensemble forecasting with an application to power transformers," *Applied Soft Computing*, vol. 68, pp. 233–248, 2018.
- [15] A. Peimankar, S. J. Weddell, T. Jalal, and A. C. Laphorn, "Evolutionary multi-objective fault diagnosis of power transformers," *Swarm and Evolutionary Computation*, vol. 36, pp. 62–75, 2017.
- [16] A. Peimankar, S. J. Weddell, T. Jalal, and A. C. Laphorn, "Ensemble classifier selection using multi-objective pso for fault diagnosis of power transformers," in *2016 IEEE Congress on Evolutionary Computation (CEC)*, July 2016, pp. 3622–3629.
- [17] R. Polikar, "Ensemble based systems in decision making," *IEEE Circuits and systems magazine*, vol. 6, no. 3, pp. 21–45, 2006.
- [18] L. Breiman, "Random forests," *Machine Learning*, vol. 45, no. 1, pp. 5–32, Oct 2001.
- [19] C. Cortes and V. Vapnik, "Support-vector networks," *Machine learning*, vol. 20, no. 3, pp. 273–297, 1995.
- [20] Y. Freund and R. E. Schapire, "A decision-theoretic generalization of on-line learning and an application to boosting," *Journal of computer and system sciences*, vol. 55, no. 1, pp. 119–139, 1997.
- [21] A. G. Ivakhnenko, "Polynomial theory of complex systems," *IEEE transactions on Systems, Man, and Cybernetics*, no. 4, pp. 364–378, 1971.
- [22] G. Shafer, *A mathematical theory of evidence*. Princeton university press, 1976, vol. 42.
- [23] A. L. Goldberger, L. A. Amaral, L. Glass, J. M. Hausdorff, P. C. Ivanov, R. G. Mark, J. E. Mietus, G. B. Moody, C.-K. Peng, and H. E. Stanley, "Physiobank, physiotoolkit, and physionet," *Circulation*, vol. 101, no. 23, pp. e215–e220, 2000.
- [24] I. Silva and G. B. Moody, "An open-source toolbox for analysing and processing physionet databases in matlab and octave," *Journal of open research software*, vol. 2, no. 1, 2014.
- [25] V. Gokana, C. Phua, and G. Lissorgues, "Automatic detection of atrial fibrillation using rr interval from ecg signals," in *The 15th International Conference on Biomedical Engineering*. Springer, 2014, pp. 215–218.
- [26] R. Colloca, A. E. Johnson, L. Mainardi, and G. D. Clifford, "A support vector machine approach for reliable detection of atrial fibrillation events," in *Computing in Cardiology Conference (CinC), 2013*. IEEE, 2013, pp. 1047–1050.
- [27] D. D. McManus, J. Lee, O. Maitas, N. Esa, R. Pidikiti, A. Carlucci, J. Harrington, E. Mick, and K. H. Chon, "A novel application for the detection of an irregular pulse using an iphone 4s in patients with atrial fibrillation," *Heart Rhythm*, vol. 10, no. 3, pp. 315–319, 2013.
- [28] A. G. Ivakhnenko, "Polynomial theory of complex systems," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 1, no. 4, pp. 364–378, 1971.
- [29] T. K. Ho, "Random decision forests," in *Proceedings of 3rd International Conference on Document Analysis and Recognition*, vol. 1, Aug 1995, pp. 278–282 vol.1.
- [30] L. Breiman, "Random forests," *Machine learning*, vol. 45, no. 1, pp. 5–32, 2001.
- [31] T. Hastie, R. Tibshirani, and J. Friedman, *The elements of statistical learning*. Springer, 2009.
- [32] L. A. Klein and L. A. Klein, *Sensor and data fusion: a tool for information assessment and decision making*. SPIE press Bellingham eWA WA, 2004, vol. 324.
- [33] H. He, Y. Bai, E. A. Garcia, and S. Li, "Adasyn: Adaptive synthetic sampling approach for imbalanced learning," in *Neural Networks, 2008. IJCNN 2008.(IEEE World Congress on Computational Intelligence). IEEE International Joint Conference on*. IEEE, 2008, pp. 1322–1328.
- [34] S. Babaeizadeh, R. E. Gregg, E. D. Helfenbein, J. M. Lindauer, and S. H. Zhou, "Improvements in atrial fibrillation detection for real-time monitoring," *Journal of electrocardiology*, vol. 42, no. 6, pp. 522–526, 2009.