Automated detection of atrial fibrillation using ECG signals

Under the guidance of: Dr Shivnarayan Patidar
16ECE1008: Harshvardhan Paithane
16ECE1012: Mahima Tendulkar
16ECE1028: Sukkhada Joshii

National Institute of Technology Goa
Department of Electronics and Communication Engineering

Contents

1. Motivation
2. Objective
3. Algorithm
4. Results
5. Conclusion
6. Future Work
7. References

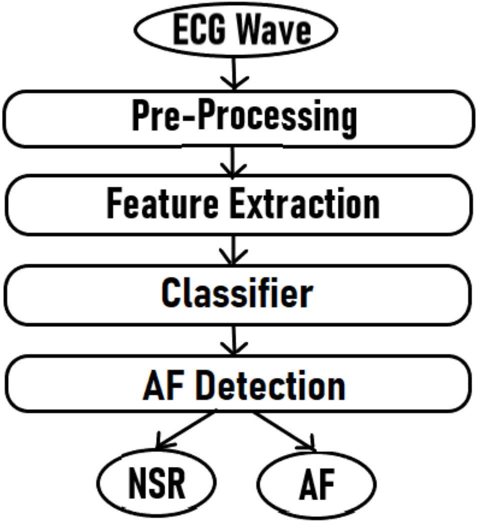
Motivation

- Possibility of avoiding severe heart related ailments by detecting Atrial fibrillation.
- To enhance the efficiency of AF detection and alleviate the workload of cardiologists
- To make life easier for heart patients and trying to improve their conditions by using the limited knowledge gained in our engineering years

Objective

- To implement the algorithm for detection of AF using wavelet packet transform and correlation function. It is an implementation of existing work[1].
- To perform the analysis by using different supervised binary classifiers and compare the results.
- To improve the generalization ability of the algorithm by training it over multiple datasets.

Algorithm for detection



Dataset

Total three datasets are used to implement the model. MIT-BIH AF Dataset (DS1) [2], PhysioNet challenge 2017 Dataset (DS2) [3], PhysioNet challenge 2020 Dataset (DS3) [3]. The algorithm is trained, validated and tested on each of them.

	Records	**No of Lead**	**Sampling Rate**
DS1	23	2	250 Hz
DS2	5925	1	300 Hz
DS3	1894	12	500 Hz

Table 1: Dataset specifications

We have used single lead from each dataset to implement the model.

AF vs NSR

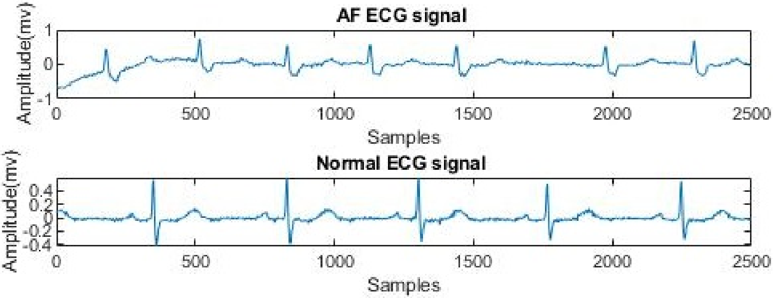


Figure 1: Plot of AF ECG signal (top) vs Normal ECG signal (bottom). DS1 record #4015.

Pre-Processing

- Denoising
  - 50 Hz Notch filter.
  - 0.3-45 Hz Bandpass filter.
- Segmentation
  - Each ECG episode is divided into segments for every 10 seconds.
  - Each segment is equivalent to window without overlapping.

ECG Signal after Pre-Processing

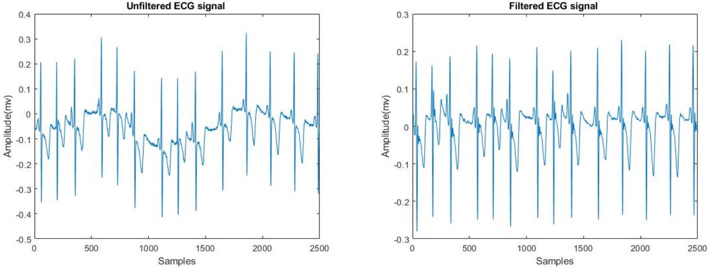


Figure 2: ECG Signal before denoising and after denoising. (DS1 record #4015)

We have filtered out Baseline wander, 50Hz powerline interference, and electrode motion noise using the mentioned filters.

Wavelet Packet Transform

WPT decomposes the signal into weighted sums of wavelet base function at different scales.

- Ability of WPT is to efficiently characterize local feature in signals.
- Output of signals decomposed by WPT in different sub-bands is 'wavelet coefficients'.
- It is a sophisticated method for signal spectral analysis and can be combined with proper statistical strategy to design the time-frequency features.

Wavelet Packet Transform

The wavelet coefficients can be computed as follows:

$$d_l^{2p}(t) = \sum_{n \in Z} h(n) d_{l-1}^{p}(2t - n)$$

$$d_l^{2p+1}(t) = \sum_{n \in Z} g(n) d_{l-1}^{p}(2t - n)$$

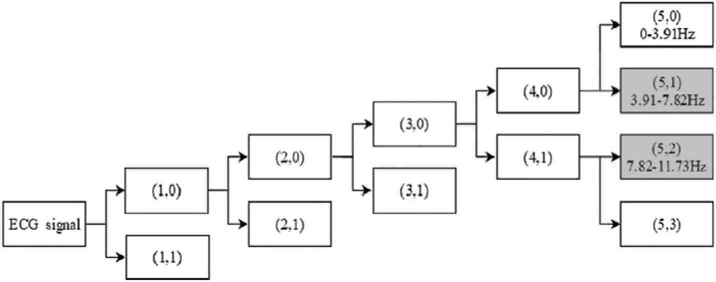


Figure 3: A part of the tree-structure of a 5-level wavelet packet transform.

ECG Signal after WPT

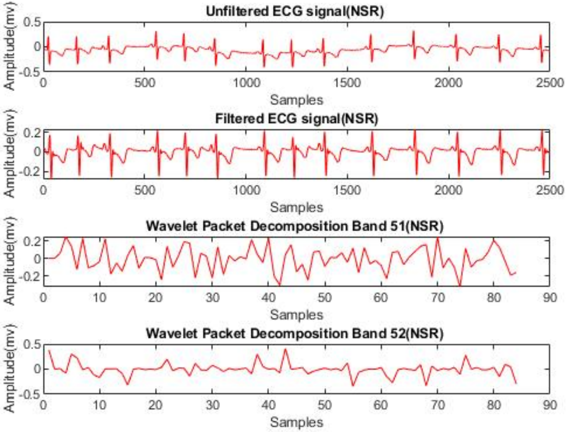


Figure 4: Plot of the sub-bands $Band_{51}$ and $Band_{52}$ of WPT decomposed NSR signal. (DS1 record #4015)

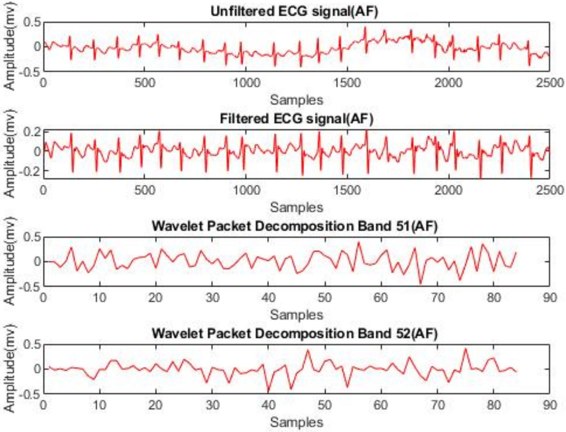


Figure 5: Plot of the sub-bands $Band_{51}$ and $Band_{52}$ of WPT decomposed AF signal. (DS1 record #4015)

Feature Extraction steps

- WPT to decompose filtered ECG segment and obtain wavelet coefficients from selected sub-bands.
- Divide coefficients series into n equal segments.
- Use these segments to compute normalized correlation matrix.

$$R_{\overline{d}} = \begin{bmatrix} B_{1,1} & \cdots & B_{1,1+\tau} & \cdots & B_{1,n} \\ \vdots & & \vdots & & \vdots \\ B_{i,1} & \cdots & B_{i,1+\tau} & \cdots & B_{i,n} \\ \vdots & & \vdots & & \vdots \\ B_{n,1} & \cdots & B_{n,1+\tau} & \cdots & B_{n,n} \end{bmatrix}$$

Feature Extraction steps

- Calculate features from the normalized matrix obtained in the previous step as follows:

$$W_B = \sum_{i=1}^{n} B_{i,i+\tau} n_{i,i+\tau}$$

$$H_B = -\sum_{i=1}^{n} p_{i,i+\tau} \log p_{i,i+\tau}$$

where, $W_B =$ Weighted sum. $H_B =$ Information entropy.

- Assemble features as feature set for the classifier.

Feature Set

Correlation among wavelet coefficient series in ECG signals affected from disorder of atrial activity will be decreased.

- Hence, now we have the feature set ready for next stage ie. classification.
  - Weighted sum (W_B)
  - Information entropy (H_B)

Feature set

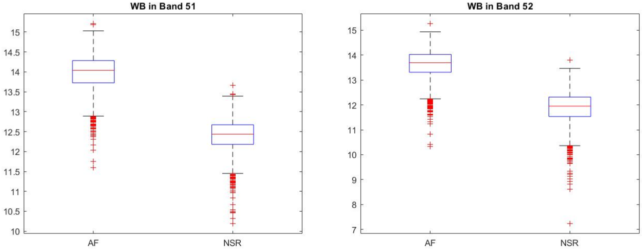
Kruskal–Wallis test is conducted to prove that the extracted features are significant. The obtained p-values are given in the Table below.

Features	MIT-BIH AF	PNC 2017	PNC 2020
W_B in $Band_{51}$	7.4424e-55	4.6259e-223	1.8753e-31
W_B in $Band_{52}$	9.1575e-60	1.5020e-135	6.4447e-33
H_B in $Band_{51}$	9.3999e-08	6.5043e-145	2.7211e-19
H_B in $Band_{52}$	2.9049e-20	5.9010e-140	5.8739e-15

Table 2: p-value of the constructed features for respective dataset.

Box-plot

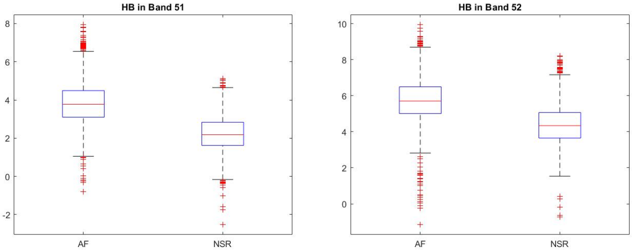
The box-plots of features show the superior classification ability in the selected sub-bands.



(a) p-value= 7.4424e-55 (b) p-value= 9.1575e-60

Figure 6: The example of box-plot for W_B in (a)$Band_{51}$ (b)$Band_{52}$ (Dataset DS1)

Box-plot



(a) p-value= 9.3999e-08 (b) p-value= 2.9049e-20

Figure 7: The example of box-plot for H_B in (a)$Band_{51}$ (b)$Band_{52}$ (Dataset DS1)

These box-plots suggest that the constructed features can be considered as important for AF detection.

Classification

- **Different classifiers such as ANN, KNN, SVM are used to implement the model.**
- For every dataset results are obtained using all three classifiers.
- In this work, the classifiers used perform Binary classification by segregating the ECG data into two classes, ie. AFib or NSR.

Classification

SVM

- A Support vector machine (SVM) constructs an optimal hyperplane as a decision surface such that the margin of separation between the two classes in the data is maximized.
- Polynomial type of SVM's like Quadratic and Cubic are used.

KNN

- A Nearest neighbor search locates the k-nearest neighbors or all neighbors within a specified distance to query data points, based on the specified distance metric.
- Weighted and Cosine type of KNN are used.

Classification

ANN

Feed forward three layered Artificial Neural Network structure is constructed for classification.

- Input layer(4 Nodes)
- Hidden layer(10 Nodes)
- Output layer(2 Nodes)

The activation function of the hidden layer and output layer is set as Sigmoid function and Softmax function, respectively.

Classification

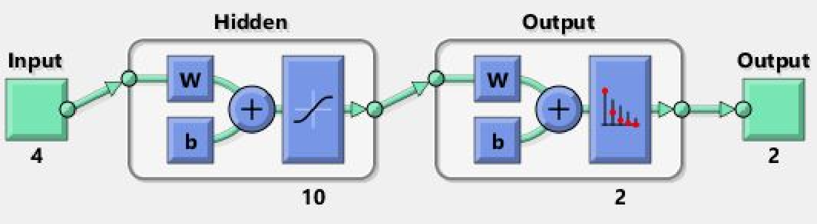


Figure 8: ANN topology in proposed scheme.

The dataset used is divided into 80% for training, 10% for validation and 10% for testing.

10-Fold cross validation

Ten-fold cross-validation is implemented to assure the credible classification performance of classifiers, which enhances the generalization ability of the model.

Three standard metrics are employed. [5]

- Accuracy (ACC)
- Sensitivity (SEN)
- Specificity (SPE)

The F1-score is also calculated using the evaluation parameters given above.

Binary classification using DS1

	ANN	KNN	SVM
ACC (%)	100	99.44	99.44
SEN (%)	100	99.44	100
SPE (%)	100	99.44	98.90
F1 Score	1.000	0.994	0.994

Table 3: Performance evaluation parameters' comparison of the algorithm using MIT-BIH AF dataset with different classifiers.

Binary classification using DS2

	ANN	KNN	SVM
ACC (%)	99.28	99.17	99.27
SEN (%)	98.35	99.33	99.46
SPE (%)	99.54	98.58	98.59
F1 Score	0.983	0.994	0.995

Table 4: Performance evaluation parameters' comparison of the algorithm using PhysioNet Challenge 2017 dataset with different classifiers.

Binary classification using DS3

	ANN	KNN	SVM
ACC (%)	99.26	98.96	99.15
SEN (%)	99.08	98.50	98.88
SPE (%)	99.44	99.43	99.43
F1 Score	0.992	0.989	0.991

Table 5: Performance evaluation parameters' comparison of the algorithm using PhysioNet Challenge 2020 dataset with different classifiers.

Binary classification (All datasets combined)

We combined all three datasets and used it to implement the method and the results obtained are given in table below.

	ANN	KNN	SVM
ACC (%)	98.03	97.78	98.50
SEN (%)	95.05	97.66	98.43
SPE (%)	99.40	98.04	98.39
F1 Score	0.968	0.983	0.988

Table 6: Performance evaluation parameters' comparison of the algorithm using all three datasets combined with different classifiers.

Binary classification (All datasets combined)

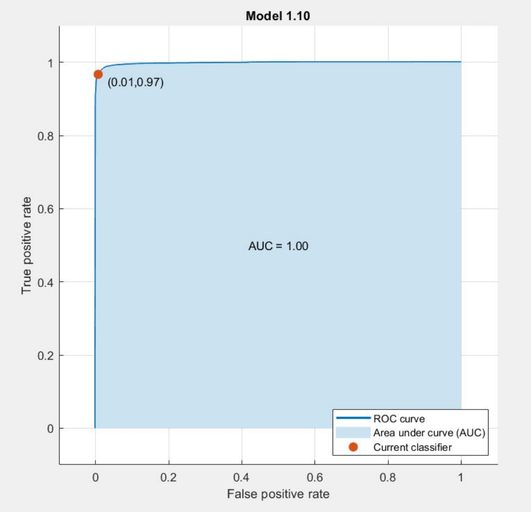


Figure 9: ROC curve for the SVM classifier.

Parameter Analysis

	DS1	DS2	DS3	All Combined
ACC	99.44	99.27	99.15	98.50
SEN	100	99.46	98.88	98.43
SPE	98.90	98.59	99.43	98.39
F1 Score	0.994	0.995	0.991	0.988

Table 7: Evaluation parameter comparison of the algorithm using different dataset with SVM classifier.

Conclusion

- The algorithm for detection of AF using Wavelet packet Transform and Correlation function was implemented successfully.
- The generalization ability of the algorithm was improved by implementing it on a larger and more diverse dataset. An accuracy of 98.5% was achieved using the SVM classification method.
- It was observed that while ANN gave a better accuracy when the datasets were trained separately, SVM classification was more effective for the generalized approach.

Future Work

Software:

- Diagnosis of other types of arrhythmias and heart valve related diseases.
- Using Deep neural networks such as a recurrent neural network to improve the classification.
- Improving the accuracy of the current machine learning model used.
- Using other wavelet transforms instead of WPT to implement proposed model.

Hardware:

- Exploring implementation of low cost device for detection of AF.

References I

[1] Jibin Wanga, Ping Wangb, Suping Wanga, "Automated detection of atrial fibrillation in ECG signals based on wavelet packet transform and correlation function of random process", *Biomedical Signal Processing and Control*, Volume 55, 101662, August 2019.

[2] MIT-BIH AF dataset, *https://physionet.org/content/afdb/1.0.0/*.

[3] PhysioNet challenge 2017/2020 dataset, *https://physionet.org/about/challenge/*.

[4] Velayudhan, Aswathy, and Soniya Peter, "Noise analysis and different denoising techniques of ECG signal-a survey", *IOSR Journal of Electronics and Communication Engineering (IOSR-JECE)*, pp 40-44, eISSN-2278, 2016.

[5] Parikh R, Mathai A, Parikh S, Chandra Sekhar G, Thomas R, "Understanding and using Sensitivity, Specificity and Predictive values", Indian J Ophthalmol, 56(1), pp 45-50, June 2008.

[6] S. Patidar, A. Sharma and N. Garg, "Automated detection of atrial fibrillation using Fourier-bessel expansion and teager energy operator from electrocardiogram signals", *2017 Computing in Cardiology (CinC)*, Rennes, 349-105, pp. 1-4, 2017.

Thank You