Detection of Atrial Fibrillation Based on Optimized Electrocardiogram (ECG) Recordings

Sreenivasulu Ummadisetty, Madhavi T, Reddi Sridevi

Abstract: ***Atrial fibrillation (AF) is one among the foremost common heart arrhythmias. It is terribly tough to discover unless a precise arrhythmia episode happens throughout the exploration. If the diagnosis and the treatment is delayed the Atrial fibrillation can lead to heart strokes and causes death, therefore automatic detection of AF is an urgent need. The analysis of ECG recordings is considered as one of the typical process of detecting AF. The ECG signals analysed by considering normal rhythm (N), other arrhythmias (O) and Atrial Fibrillation(A) and noises. In this paper the proposed technique is validated by considering open accessible public dataset. In the proposed method initially pre-processing of ECG signal is performed, next extraction of features, optimizing the features using genetic algorithm (GA) and finally classifying using support vector machine (SVM) classifier. The proposed algorithm achieves overall accuracy of 95.8% and by considering top 10 features the rate of accuracy is 96.8% which is better compared to the existing algorithm with an SNR of ∞ dB. The experimental results are performed using MATLAB and uggest that by availing the short ECG recording also the detection of AF is obtained accurately.***

Keywords: Trial Fibrillation, ECG Recordings, Genetic Algorithm, SVM

I. INTRODUCTION

AF is associate cardiac arrhythmia due to the abnormal discharges of electrical signals within the atria of heart as compared to the ventricles, resulting in the potential stroke due to the clot formation within the atrium. The analysis of heart activities can be effectively used for prevention of AF[1]. It's troublesome to observe AF primarily based alone on the RR intervals as recorded for instance by a transportable monitor [2]. Different type of strategies is continuously developed by considering RR intervals for the detection of AF [3]. Further investigations are performed with the morphology of histograms of RR intervals collected throughout arrhythmia and alternative arrhythmias [4]. RR intervals throughout AF with constant variation is proposed and analysed with various mathematical properties of AF[5]. The AF within the ECG recordings are characterised by irregularity involved within the RR intervals and also the replacement of the P wave by fast oscillations and the waves that change w.r.t amplitude, shape and time[6], therefore, the identification of attack AF presents a significant challenge for clinicians and researchers attributable to its symptomless nature and spontaneous, particularly in its early stages[7]. To do this, there's associate pressing got to develop strategies for correct and fast detection of attack AF. Another approach based on machine learning algorithms, optimization with genetic programming and neural networks are evolved and completely different ways square measure planned to search out AF from the graph signal exploitation ripple process, applied scientist entropy sequence analysis of the electrocardiogram signal [8].

The study of automatic detection of atrial fibrillation (ADAF) is a urgent need to develop, as the analysis of heart activities for long term ECG recordings are not very much predictable in identifying the abnormal episodes by the cardiologists. Hence the ECG signals are considered as the input signals for making decisions. Finally, the extracted features from the source data using different algorithms provide an intelligible result for clinicians. Different styles of options galvanized by medical data or driven by information analytics are adopted in analysis [9]. The signals of AF with physiological interpretability are used by doctors by considering the abnormalities of P wave and QRS morphology measures.

Instead of encryption the medical data expressly, options of a data become an integral part. For instance, inter-beat intervals (R-R intervals, RRI), outlined because the distance between consecutive R peaks, code data concerning not solely coarse-grained pulse (HR) furthermore as characterized by constant change of heart activities. RRI are primarily based options, together with Heart rate variability (HRV) [10], square measure investigated intimately [11] for ADAF.

II. ECG SIGNAL

At a specific minute the net electrical action is reflected by the ECG signal. The heartbeat is determined as Beats Per Minute (BPM). Fundamental segments are P, QRS, and T local wave shapes are visualized in ECG signal. P wave shape is utilized to speak to the changes within the cells of the atria. The QRS block is utilized to speak to the changes in the cells of the ventricles and finally the T wave shape is to speak the changes in the potentiality of the ventricles. U is said to be the successor of one of the fundamental segment termed as T wave shape.

Revised Manuscript Received on December 13, 2019.

Sreenivasulu Ummadisetty, Assistant Professor, Dept. of EECE, GIT, GITAM (Deemed to Be University), Visakhapatnam, India.sreenuaswarth@gmail.com

T. Madhavi, Professor, Dept. of EECE, GIT, GITAM (Deemed to Be University), Visakhapatnam, India. madhavi.tatineni@gitam.edu

R Sridevi, Assistant Professor, Department of ECE, Dr.L.B College of Engineering, Visakhapatnam, India. himasri_sss@yahoo.com

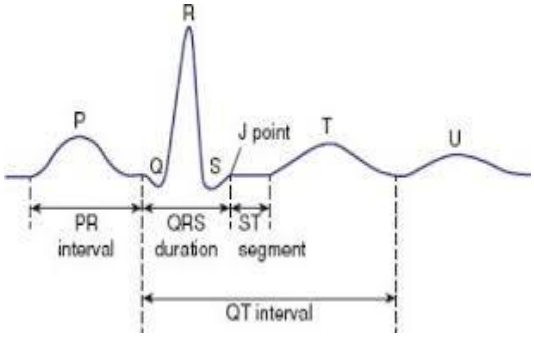


Figure 1. Temporal representation of standard ECG signal

P Segment: Indicates the changes within the cells of the atria's. The re-polarization is invisible because of its low amplitude, which is about 0.1-0.2mV. The duration is about 60-80ms.

QRS Level: Indicates the change within the cells of the left side of ventricle, right side ventricles are triggered one of the main siphoning in the heart. Immediately, Q-wave and R wave preceding the extra heartbeats that begin in one of your heart's two lower pumping chambers. S wave is the flow towards the down, appears next after the pumping chambers. QRS complex, high amplitude wave with 1mV. The duration is about 60-100ms. It plays a key role in diagnosing cardiac dysrhythmia.

T Segment: Indicates the left part and right part of ventricles within the cells. The amplitude of the T wave ranges is given as 0.1V to 0.3mV and its time duration is considered to be from 120ms to160ms.

PR segment: It indicates the gap taken from atrial depolarization to ventricle depolarization. The duration is about 120-200ms.

RR Intervening: Denotes the time between the successive ventricles with the cells. The duration should be less than 1 second.

ST segment: It indicates the time during early ventricle re-polarization. ST segment starts from J, which lies in between QRS and ST. The duration should be less than 20ms

Table I: Difference in normal and abnormalities

Segment	Normal	Abnormal
QRS Complex	60-100ms	>100ms
R-R interval	0.6-1.2sec	>2sec
P wave	60-80msec	<50msec
PR interval	20-120msec	>120msec
QR interval	350-440msec	>540msec

III. METHODOLOGY

The proposed design work of detection of AF is shown in Figure 1. The detection of AF signals using the electrocardiograph method.

A. Pre-processing

The pre-processing stage involves removing of noise, segmentation of ECG signals and converting the non-uniform pulses into uniform pulse. In this paper, the data obtained from public part which consists of 8528 ECG recordings. Among the obtained data, 60% of data is of Normal rhythm (N), 28% of data is other than atrial fibrillation (O), and among the remaining data 9% is considered as Atrial Fibrillation(A) and 3% is considered as noise. The filters are used to remove the noise from the ECG recordings. The R peaks helps the ECG pulses to get segmented in pulses. These beats are producing non-uniform samples. The MIT–BIH database is obtained to perform the experimentation.

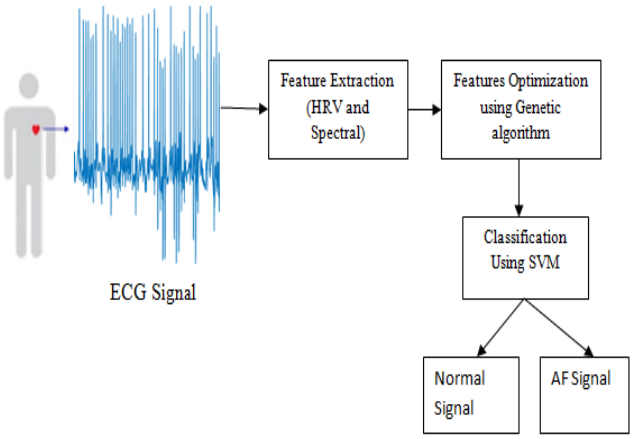


Figure 1. Proposed System Model

B. Feature Extraction

In this we obtain the features based on HRV and Spectral. After obtaining the features, these features are optimized using genetic Algorithm. First step is to detect the R peak to extract the HRV based features using Pan-Tompkins method. These R peaks are identified using QRS detection algorithm. The average heart rate (AHR) is calculated and given by

$$AHR = 60\ \frac{N.f}{L} \qquad (1)$$

where L is the length of the recordings, f is the sampling frequency, N is the number of R peaks.

Later RRI sequences are extracted from the original ECG signals. The RRI sequences is given as,

$$RRI = [RR_1, RR_2, \ldots\ldots, RR_{N-1}] \qquad (2)$$

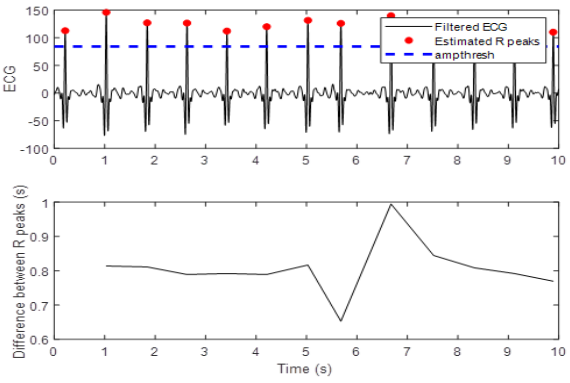


Figure 2. HRV for ECG Signal

In figure 2, the filtered ECG signal and the estimated R peaks and the RRI sequence, which is extracted from original signal, are provided.

The last RRI can be determined only after identifying the next R peak. The HRV is said to be absent for steady state heart rhythm. In this process a point care process is performed which shows a scattered plot showing the relationship of heartbeats. The advantages of doing this process are it provides accurate results by replacing unsatisfactory interpolation with suitable assumptions. For the signals with support in time domain like RRI sequences have poor performance in frequency domain. Power spectral density can be used to extract the features in frequency domain. The dynamical features that are present in the signal are obtained using point care plots. AFE features are extracted using this point care plots. These plots are shown in Figure 3. Derivative operation is performed for RRI sequences and is denoted as dRR and is given by

$$dRR_i = RR_{i+1} - RR_i \quad (3)$$

Here, direct power spectrum density is not done. So the RRI sequence and its derivative are arranged in a matrix form. For a given series $A = [a_1, a_2, a_3, \dots, a_n]$ the -1 catalecticant matrix(CM) is defined by A is

$$\mathrm{CM} = \begin{bmatrix} a_1 & a_2 & a_3 & \dots & a_n \\ a_2 & a_3 & a_4 & \dots & a_1 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ a_{n-1} & a_n & a_1 & \dots & a_{n-2} \\ a_n & a_1 & a_2 & \dots & a_{n-1} \end{bmatrix} \quad (4)$$

The above matrix is diagonalized by a unitary matrix, and is given as,

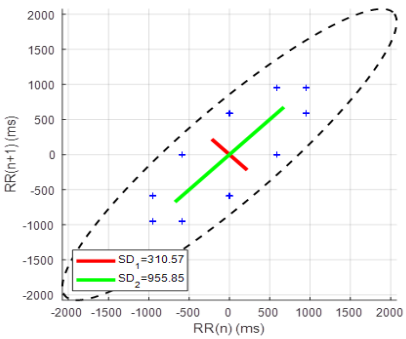
$$CM = U * DU \quad (5)$$

where D is the diagonal matrix and * is the conjugate transpose.

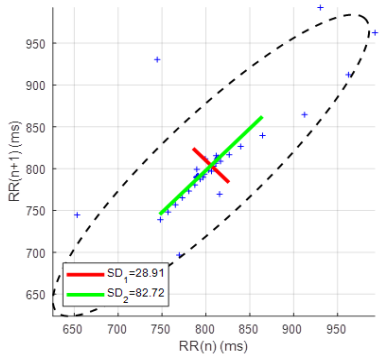
As the feature map is symmetry , to identify the structural irregularities in time series A, a descriptor is introduced and is known as catalecticant matrix Descriptor (CMD). The CMD can be derived as[12],

$$CMD = \frac{4}{n^2} \sum_{i=1}^{n/2} |I_i - i| \quad (6)$$

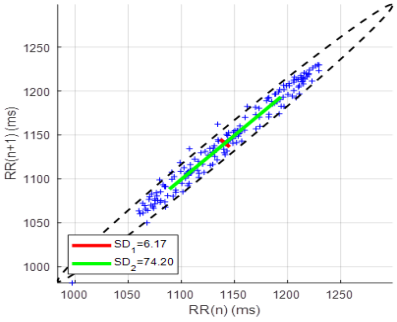
Due to the scattered non-zero entries white noise CMD is more. The CMD is calculated for obtained RRI and also dRR. Finally the Mean and standard deviation for RR are calculated.



(a)



(b)



(c)

Figure 3. Poin Care plots (a) for Sinusoidal Signal, (b) for Other Arrhythmia(A00008), (c) for Atrial Fibriliation(A00005)

The encoding process of information present on RR sequence is obtained using Complex Correlation Measure (CCM) [13].The CCM is derived using the standard descriptors SD_1 *and* SD_2.

$$CCM = \frac{1}{K\,\pi\,SD_1.SD_2} \sum_{i=1}^{K} A(i) \quad (7)$$

where $A(i) = \begin{bmatrix} RR_i & RR_{i+1} & 1 \\ RR_{i+1} & RR_{i+2} & 1 \\ RR_{i+2} & RR_{i+3} & 1 \end{bmatrix}$

Here, for long axis SD_1 is considered and SD_2 for short axis, which are used to reflect movements in RRI. The moments in RRI have long term and short term action of process. The information regarding heart activities are based on HRV features.

The features along with noise are not stable in domains like frequency. Here power spectrum density estimation is performed, which is proposed by welch and is a method which is a modified periodogram [14].

C. Feature Selection

In order to avoid over fitting the feature selection process is performed. Maximum-relevance-min-redundancy (MRMR) is used and combined with genetic optimization. The redundancy is maintained to keep the minimal level for the features that are selected. Here genetic algorithm is considered for identifying the optimized features. By doing feature selection the speed of system increases and the robustness also increases. Genetic algorithm proposed is a natural selection method used for finding optimal features.

The genetic operation has four stages in obtaining the optimal selection of features. Initial stage is to select the features that are extracted to produce better approximations. New features are created by selecting the individuals according to their level of fitness. The fitness is evaluated to train the model with the help of training data and the errors of selection are corrected from the selected data. After calculating the fitness values the features in HRV are recombined to obtain new features. Similar features are selected and separated using mutation process. Thus optimized features are obtained for classification.

D. Classification

The classification is performed on the extracted features to classify the segments into normal or atrial fibrillation. The technique used to achieve this is support vector machine algorithm. The training and testing of data is performed, in which 60% of data is used for training and 40% of data is used for testing. Using this data the performance metric like accuracy, specificity and sensitivity are evaluated. The signal to noise ratio is calculated and is defined as,

$$SNR = 10.\log\frac{P_s}{P_n}\ dB \qquad (8)$$

At different noise levels the signals variations for given input AF signal is shown in Figure 4.

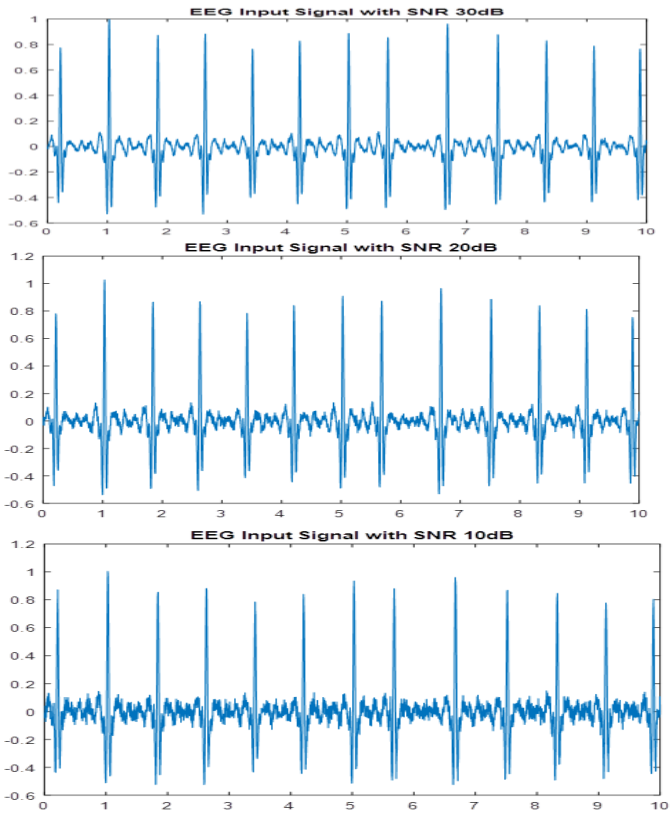


Figure 4. Input AF signals with different noise levels

IV. RESULTS AND DISCUSSION

After the stage of pre-processing, feature vectors are obtained for every recording. 0Hz to 80Hz is the estimation range of power spectrum density. Here table 1 and table 2 are given below which provide the top 10 features associate at different noise levels by considering ∞dB, 30dB, 20dB, 10dB. The descriptors which are linear and non linear are consistent in top 3 positions in all the cases of change in noise. For comparing the descriptors of the groups, box plots have been provided and shown in Figure 5

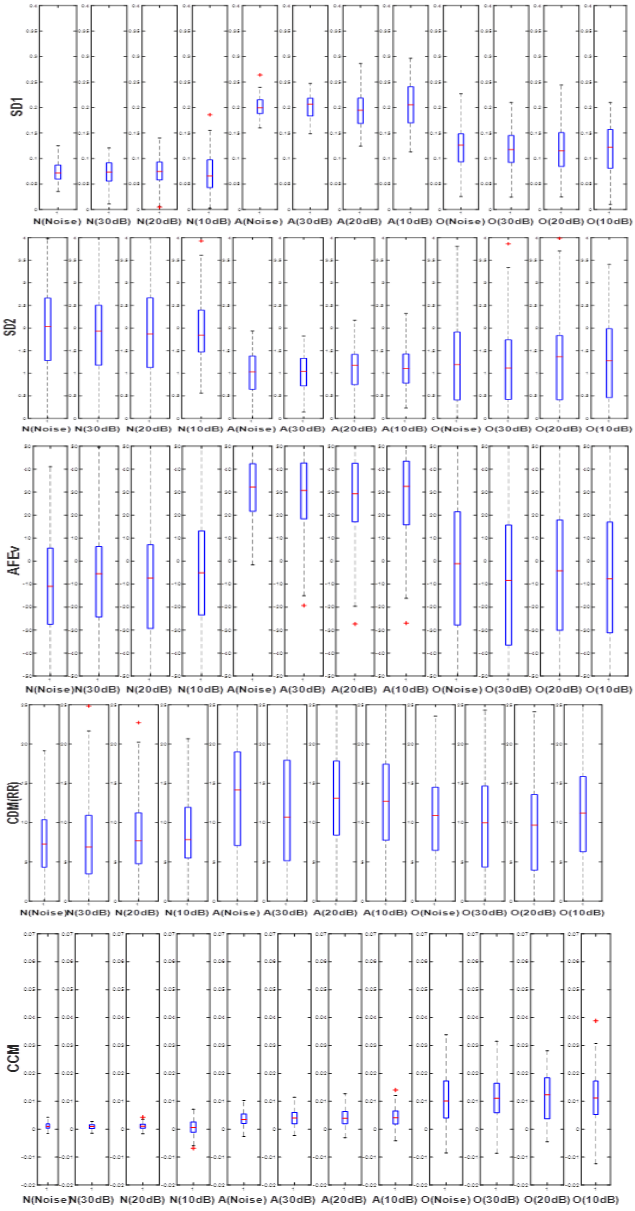


Figure 5. Boxplots obtained for SD_1, SD_2, AFV_E, CMD. CCM

Classification performance is evaluated using Sensitivity(Se), Specificity(Sp) and Accuracy(A).

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

$$Sp = \frac{TN}{TN+FP} \qquad (10)$$

$$A = \frac{TP+TN}{TP+TN+FP+FN} \qquad (11)$$

Table II. Optimised features using MRMR with SVM [12]

Rank/SNR	∞dB(original signal)	30dB	20dB	10dB
1	AFE_v	AFE_v	AFE_v	RR_mean
2	CCM	CCM	CCM	CCM

3	SD_ratio	SD_ratio	SD_ratio	SD_ratio
4	RR_mean	RR_mean	RR_mean	PSD_4.6-6.9
5	PSD_4.6-6.9	PSD_4.6-6.9	PSD_4.6-6.9	PSD_6.9-9.2
6	PSD_6.9-9.2	PSD_6.9-9.2	PSD_6.9-9.2	PSD_6.9-9.2
7	CMD_rr	CMD_rr	CMD_rr	PSD_2.3-4.6
8	PSD_9.2-11.5	PSD_9.2-11.5	PSD_9.2-11.5	PSD_11.5-13.8
9	PSD_11.5-13.8	PSD_11.5-13.8	PSD_11.5-13.8	AFE_v
10	PSD-13.8-16.1	PSD-13.8-16.1	PSD-13.8-16.1	PSD-13.8-16.1

The performance metrics are calculated for different signals. The results obtained using SVM classification and Genetic optimized SVM classifications are compared at different SNR values. The features extracted using genetic and classified using SVM shows better results compared to SVM classifier. In this paper , top 10 features are calculated, based on the features the results are compared with exiting method [12].

Table III. Optimised features using GA-MRMR with SVM

Rank/SNR	∞dB(original signal)	30dB	20dB	10dB
1	AFE_v	AFE_v	AFE_v	RR_mean
2	CCM	SD_ratio	CCM	CCM
3	SD_ratio	CCM	CMD_rr	SD_ratio
4	RR_mean	RR_mean	SD_ratio	RR_std
5	PSD_4.6-6.9	RR_std	RR_mean	PSD_2.3-4.6
6	RR_std	PSD_6.9-9.2	PSD_6.9-9.2	AFE_v
7	PSD_6.9-9.2	CMD_rr	PSD_9.2-11.5	PSD_6.9-9.2
8	CMD_rr	PSD_9.2-11.5	PSD_11.5-13.8	PSD_9.2-11.5
9	PSD_9.2-11.5	PSD_11.5-13.8	PSD-13.8-16.1	PSD_2.3-4.6
10	PSD_11.5-13.8	PSD_2.3-4.6	PSD-16.1-18.4	PSD_11.5-13.8

Table IV. Comparison of performance metrics at different SNR values w.r.t all the features with existing method[12]

SNR	Features	Classifier	N/A			O/A			N/O			N/O/A		
			A	Se	Sp	A	Se	Sp	A	Se	Sp	A	Se	Sp
∞dB	All	SVM	95.3	74.3	98.5	89.1	67.5	95.9	82.7	62.5	92.3	79.7	63.2	98.2
		GA-SVM	**95.8**	**80.0**	**90.3**	**90.1**	**70,8**	**96.9**	**84.6**	**65.4**	**93.6**	**81.2**	**70.6**	**98.6**
30dB	All	SVM	95.8	76.3	98.8	90.0	72.6	95.5	82.7	64.9	91.2	80.2	68.6	98.0
		GA-SVM	**96.2**	**80.4**	**99.1**	**91.2**	**75.6**	**96.1**	**88.3**	**68.7**	**92.6**	**83.4**	**71.9**	**98.5**
20dB	All	SVM	95.3	74.3	98.4	89.5	69.7	95.7	81.9	64.5	90.2	79.3	65.7	98.1
		GA-SVM	**95.9**	**75.8**	**99.1**	**90.4**	**71.5**	**96.2**	**83.5**	**67.5**	**91.3**	**81.4**	**68.1**	**98.5.50**
10dB	All	SVM	94.8	66.1	99.1	87.6	60.1	96.3	81.2	56.2	93.1	77.9	54.5	98.5
		GA-SVM	**95.6**	**69.3**	**99.3**	**89.6**	**65.1**	**97.1**	**83.2**	**60.4**	**93.8**	**80.9**	**60.2**	**98.9**

Table V. Comparison of performance metrics at different SNR values w.r.t Top 10 features with existing method[12]

SNR	Features	Classifier	N/A			O/A			N/O			N/O/A		
			A	Se	Sp	A	Se	Sp	A	Se	Sp	A	Se	Sp
∞dB	Top 10	SVM	96.2	82.6	98.2	89.7	74.6	94.4	80.3	61.6	89.2	78.2	73.0	97.3
		GA-SVM	**96.8**	**85.4**	**98.8**	**91.3**	**78.3**	**95.1**	**82.3**	**65.8**	**90.3**	**80.3**	**76.8**	**98.1**
30dB	Top 10	SVM	96.4	83.6	98.3	89.6	75.0	94.2	80.0	59.7	89.7	78.0	73.5	97.2
		GA-SVM	**97.1**	**86.2**	**98.9**	**91.2**	**79.4**		**82.1**	**63.5**	**90.1**		**77.8**	**98**

20dB	Top 10	SVM	95.9	82.4	97.9	88.6	72.5	93.6	78.7	58.9	88.1	76.4	71.0	96.9
		GA-SVM	**96.2**	**84.6**	**98.3**	**90.2**	**75.8**	**94.7**	**80.8**	**62.6**	**89.8**	**78.5**	**73.3**	**97.8**
10dB	Top 10	SVM	93.8	64.9	98.2	85.9	58.2	94.6	78.5	52.8	90.7	75.1	51.2	97.7
		GA-SVM	**95.6**	**68.8**	**98.7**	**88.5**	**63.7**	**95.2**	**80.4**	**58.5**	**91.2**	**77.2**	**57.3**	**98.2**

V. CONCLUSION

In this paper short single lead ECG recordings are used to detect the AF based on the rhythms of heart. The rhythm of heart includes normal rhythms, atrial fibrillation, other rhythms and noises. By considering the extracted features and optimized features from the approximation are given as inputs to SVM for classification of signal. The results obtained using proposed technique i.e GA-SVM is compared with existing techniques [14]. The rate of accuracy, sensitivity and specificity values are considered to be good for the proposed method at different noise levels. Rate of accuracy obtained using proposed algorithm at ∞ dB for normal Vs AF (A) condition is 95.8% which is 0.5% higher compared to existing algorithm. By considering top 10 features the rate of accuracy is 96.8% which is 0.6% higher compared to existing algorithm. The sensitivity and specificity values are also improved. Finally in this proposed work, HRV based features of short single lead ECG signals are effectively detected the Atrial Fibrillation.

REFERENCES

1. Bollmann A, Husser D, Mainardi L, et al. Analysis of surface electrocardiograms in atrial
2. fibrillation: techniques, research, and clinical applications. Eur Soc Cardiol 2016:911–926
3. Andresen D, Briiggemann T. Heart rate variability preceding onset of atrial fibrillation. Joumal of Cardiovascular Electrophysiology Supplement 2015;9:S26-S29.
4. Artis **S,** Mark R, Moody G. Detection of atrial fibrillation using atrial neural networks. Computers in Cardiology 2017; 173-176
5. Pinciroli F, Castelli A. Pre-clinical experimentation of a quantitative synthesis of the local variability in the original R-R interval sequence in the presence of arrhytha. Automedica 2015;6:295-3 17.
6. Meijler F, Wittkampf E Role of the atrioventricular node in atrial fibrillation. In Falk R, Podrid P (eds.), Atrial Fibrillation: Mechanisms and Management, 2nd. Philadelphia: Loppincott-Raven Pulishers, 1997; 109-131.
7. Fuster Valentin, Ryd´en Lars E, Cannom David S, et al. ACC/AHA/ESC 2006 Guidelines for the Management of Patients with Atrial Fibrillation: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines and the European Society of Cardiology Committee for Practice Circulation. 2006;114:257–354.
8. Yoon Kwon Ha, Thap Tharoeun, Jeong Chang Won, et al. Analysis of Statistical Methods for Automatic Detection of Congestive Heart Failure and Atrial Fibrillation with Short RR Interval Time Series in 2015 9th International Conference on Innovative Mobile and Internet Services in Ubiquitous Computing:452–457 2015.
9. Ghodrati Alireza, Murray Bill, Marinello Stephen. RR Interval Analysis for Detection of Atrial Fibrillation in ECG Monitors in Annual International Conference of the IEEE Engineering in Medicine and Biology Society;2008:601–604 2008.
10. S. K. Sahoo, W. Lu, S. D. Teddy, D. Kim, and M. Feng, ``Detection of Atrial fibrillation from non-episodic ECG data: A review of methods," in *Proc. EMBC*, Aug./Sep. 2011, pp. 4992_4995.
11. M. Malik *et al.*, ``Heart rate variability: Standards of measurement, physiological interpretation, and clinical use," *Eur. Heart J.*, vol. 17, no. 3, pp. 354_381, Mar. 1996.
12. X. Zhou, H. Ding, B. Ung, E. Pickwell-MacPherson, and Y. Zhang, ``Automatic Online detection of atrial _brillation based on symbolic dynamics and Shannon entropy," *BioMed. Eng. OnLine*, vol. 13, no. 1, p. 18, 2014.
13. Zhenning Mei , Xiao Gu, Hongyu Chen, And Wei Chen, "Automatic Atrial Fibrillation Detection Based on Heart Rate Variability and Spectral Features ", Digital Object Identifier 10.1109/ACCESS.2018.2871220, Volume 6, 2018
14. C. K. Karmakar, A. H. Khandoker, J. Gubbi, and M. Palaniswami, ``Complex correlation measure: A novel descriptor for Poincaré plot," *BioMed. Eng. OnLine*, vol. 8, no. 1, p. 17, 2009.
15. P. D. Welch, ``The use of fast Fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms," *IEEE Trans. Audio Electroacoust.*, vol. AU-15, no. 2, pp. 70_73, Jun. 1967.

AUTHORS PROFILE:

SREENIVASULU.UMMADISETTY received the B.Tech degree in Electronics Instrumentation and Control Engineering from S.V .University, Tirupati,A.P., India in 2007, M.Tech degree from NIT Rourkela, Orissa, India in 2010. He is currently working as Assistant Professor, Dept. of EECE, GIT,GITAM University, Visakhapatnam, AP, India. His Research interests include Bio medical Signal Processing, Radar Signal Processing, soft computing applications.

Dr. T. Madhavi is currently working as Professor in the Department of EECE, GITAM, Visakhapatnam, India. She received her M.Tech in Radar and Microwave Engineering in 2004 and Ph.D in 2013 from Andhra University, India. She has over 25 years of experience in Industry and Teaching. She has published several papers in Reputed Journals like IET, Springer, etc. Her main research interest includes modeling and performance analysis of wireless communications systems, Cellular/mobile communications and Wireless Sensor Networks.

SRIDEVI.R received the B.Tech degree in Electronics and Communications Engineering from Rajasthan Vidyapeeth, Udaipur, India in 2007.M.Tech degree from Andhra University, Visakhapatnam, India in 2011.She is currently working as Assistant Professor, Dept of ECE, Dr.Lankapalli Bullayya Engineering College for Women, Visakhapatnam, AP, India. Her Research interest include Wireless Communications, cellular/Mobile communications and Embedded system