

Study and Literature Survey for Classification of Cardiac Arrhythmia With Respect to ECG and HRV Signal

Mr. Vyankatesh S. Thorat¹, Dr. Suresh D. Shirbahadurkar², Mrs. Vaishali V. Thorat³

Student of M.E of D.Y. Patil School of Engg, Ambi, Pune, India¹

Principal and Professor of D. Y. Patil College of Engg, Ambi, Pune, India²

Assistant Prof and P.G. Co-ordinator of D.Y. Patil School of Engg, Ambi Pune, India³

Abstract: Consistent or periodical heart rhythm disorders may result cardiac arrhythmias. presence of recurring arrhythmic events (also known as cardiac dysrhythmia or irregular heartbeats), as well as erroneous beat detection due to low signal quality, significantly affects estimation of both time and frequency domain indices of heart rate variability (HRV). A reliable, real-time classification and correction of ECG-derived heartbeats is a necessary prerequisite for an accurate online monitoring of HRV and cardiovascular control. In this, Heart Rate Variability (HRV) signals are analyzed and various features including time domain, frequency domain and nonlinear parameters are extracted. The additional nonlinear features are extracted from electrocardiogram (ECG) signals. These features are helpful in classifying cardiac Arrhythmias. In this, we are going to use genetic programming which is applied to classify heart Arrhythmias using both HRV and ECG features. Genetic programming selects effective features, and then finds the most suitable trees to distinguish between different types of Arrhythmia. By considering the variety of extracted parameters from ECG and HRV signals, genetic programming can be used precisely to differentiate various arrhythmias. The performance of used algorithm is evaluated on MIT-BIH Database. Here we are going to see seven different types of arrhythmia classes which includes normal beat, left bundle branch block beat, right bundle branch beat, premature ventricular contraction, fusion of ventricular and normal beat, atrial premature contraction and paced beat are classified with an accuracy of 98.75%, 98.93% , 99.10%, 99.46%, 99.82%, 99.46% and 99.82% respectively. In this paper, we are going to classify arrhythmias by using genetic algorithm.

Index Terms: Arrhythmia, Electrocardiogram (ECG), Heart Rate Variability (HRV), Genetic Programming (GP), Feature Selection.

I. INTRODUCTION

Regarding to classification of cardiac arrhythmia, any disturbance in the heart rhythm can be cause danger to person. Although cardiac arrhythmia is one of the leading causes of death, it can be treated if detected on time. Heart is a muscular organ which is responsible to pump oxygenated blood throughout blood vessels by rhythmic contractions. Heart Arrhythmia can cause too slow or too fast performance of the heart. To detect it, ECG and HRV signals are widely used. ECG signal records electrical performance of the heart. It contains a lot of important information related to the condition of the heart and one of the most important tools in detecting heart diseases. A typical ECG signal consists of the P-wave, QRS complex, and T-waves. The P wave is the result of slow moving depolarization of the atria. QRS complex which is made of Q, R and S waves shows ventricular depolarization. The T wave represents repolarization of the ventricles, and is longer in duration than depolarization. HRV signal describes the variations between consecutive heartbeats. It is a non invasive evaluation method of the nervous system which controls cardiovascular system and is a measurement of the interaction between sympathetic and parasympathetic activity in autonomic nervous system.

HRV signal is a non-stationary signal and its changes can be interpreted as a current or upcoming disease. Several methods for automatic detection and classification of cardiac arrhythmias have been proposed in literature, including: artificial immune recognition system with fuzzy weighted, threshold-crossing intervals, neural networks, fuzzy neural networks, fuzzy equivalence relations, Bayesian classifiers, support vector machines, wavelet transforms, combined wavelet transformation and radial basis neural networks, fuzzy logic combined with the Markov models and the rule based algorithms. Some papers used techniques which are based on ECG segment. The various features of the ECG signal including the morphological features are extracted and used for classification of the cardiac arrhythmias. This is a time consuming procedure and the results are very sensitive to the amount of noise. An alternative approach would be to extract the HRV signal from the ECG signal first by recording the R-R time intervals and then processing the HRV signal instead.

This is a more robust method since the R-R time intervals are less affected by noise. One drawback of the proposed HRV-based algorithm is that some of the arrhythmia types

such as the left bundle branch block and the right bundle branch block beats cannot be detected using only the heart rate variability (HRV) features. In this, a new arrhythmia classification algorithm is proposed which is able to effectively classify seven types of arrhythmia. These arrhythmias are namely the normal beat (NB), left bundle branch block beat (LBBB), right bundle branch beat (RBBB), premature ventricular contraction (PVC), fusion of ventricular and normal beat (FUSION), atrial premature contraction (APC) and paced beat (PACE). In this, various features from both ECG and HRV signal are extracted and given to a genetic programming to produce the suitable solution trees to distinguish between different types of arrhythmia. From the various identified features the proposed method selects the effective ones and categorizes the seven classes of heart arrhythmia highly precisely. The main objective of the proposed work is: genetic programming is applied to classify heart arrhythmias using both HRV and ECG features. Genetic programming selects effective features, and then finds the most suitable trees to distinguish between different types of arrhythmia.

II.LITERATURE REVIEW

In this paper, the potential of this method for clinical uses and real-time detection was examined using 180 electrocardiogram records [60 atrial fibrillation, 60 ventricular fibrillation, and 60 ventricular tachycardia]. The proposed algorithm has achieved high accuracy (more than 97%) and is computationally fast in detection. [6] In this paper, new classification algorithm of the ECG beats, applying the fuzzy hybrid neural network and the features drawn from the higher order statistics has been proposed in the paper. The cumulants of the second, third, and fourth orders have been used for the feature selection. The hybrid fuzzy neural network applied in the solution consists of the fuzzy self-organizing subnetwork connected in cascade with the multilayer perceptron, working as the final classifier. The c-means and Gustafson-Kessel algorithms for the self-organization of the neural network have been applied. The results of experiments of recognition of different types of beats on the basis of the ECG waveforms have confirmed good efficiency of the proposed solution. The investigations show that the method may find practical application in the recognition and classification of different type heart beats.[7] In this paper, reliable, real-time classification and correction of ECG-derived heartbeats is a necessary prerequisite for an accurate online monitoring of HRV and cardiovascular control. We have developed a novel point-process-based method for real time R-R interval error detection and correction. Given an R-wave event, we assume that the length of the next R-R interval follows a physiologically motivated, time-varying inverse Gaussian probability distribution. We then devise an instantaneous automated detection and correction procedure for erroneous and arrhythmic beats by using the information on the probability of occurrence of the observed beat provided by the model. We test our algorithm over two datasets from the PhysioNet archive. The Fantasia normal

rhythm database is artificially corrupted with known erroneous beats to test both the detection procedure and correction procedure. The benchmark MIT-BIH Arrhythmia database is further considered to test the detection procedure of real arrhythmic events and compare it with results from previously published algorithms. Our automated algorithm represents an improvement over previous procedures, with best specificity for the detection of correct beats, as well as highest sensitivity to missed and extra beats, artificially misplaced beats, and for real arrhythmic events. [1] In this following paper investigates four novel methods for noncontact measurement of heart rate (HR) and consequently its derivate HR variability, an important marker of autonomic activity proven to be predictive of likelihood of future health related events. Feasibility study of basic principles is focused on measurements of signal-to-noise ratio with respect to the distance between the subject and HR sensor/apparatus. [2] In this paper, analyzing the Holter recordings collected at baseline during the European Myocardial Infarction Amiodarone Trial (EMIAT), we evaluate the possibility of using , the slope of the power spectrum of heart rate variability signals (HRV) in the vicinity of $f = 0$, for postinfarction risk stratification.[3] In this, study proposes the use of pulse rate variability (PRV) extracted from the PPG signal instead of HRV. Polysomnographic registers from 21 children (aged 4.47 ± 2.04 years) were studied. The subject classification based on DA Pevents and PRV analysis obtained an accuracy of 86.67% which represents an improvement of 6.67% with respect to the HRV analysis. These results suggest that PRV can be used in apnea detectors based on DAP events, to discriminate apneic from nonapneic events avoiding the need for ECG recordings.[4] In this paper, the potential of the authors' method for clinical uses and real-time detection was examined using human surface ECGs and intra cardiac electrograms (EGMs). The method achieved high sensitivity and specificity (>0.98) in discrimination of supraventricular rhythms from ventricular ones. The authors also present a hardware implementation of the algorithm on a commercial single-chip CPU.[5]

III. METHODOLOGY

A) Block diagram

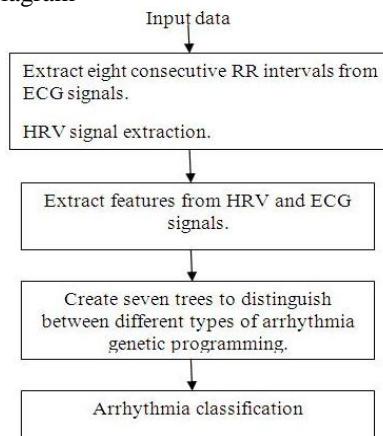


Fig 1. Block diagram

In this, the proposed method is tested using the MIT-BIH arrhythmia database. This database contains 48 ECG records. Each record is approximately 30 minutes long and it includes 109000 R-R intervals with sampling frequency of 360 Hz. Each beat has been annotated independently by two cardiologists. Their annotations are compared, consensus on disagreements was obtained, and the reference annotation files were prepared. In most records, the upper signal is a modified limb lead II (ML II) and the lower signal is a modified lead V1 (VI). All of the ECG records are chosen from lead II and include 8 sequential R-R intervals. Extracted records of this database include all seven classes of heart beat. After dividing the ECG signals into 8 sequential RR intervals, HRV signal is extracted from calculating the time intervals between every two consequential R-wave in an ECG signal (R-wave is located in the maximum absolute value of the signal within the time window).

IV. FEATURE EXTRACTION

A) Extraction of features from HRV signal

The behavior of HRV signal includes both linear and nonlinear behavior combination. The combination of these features is considered. These feature includes time domain, frequency domain and nonlinear parameters. Each ECG signal is divided into 8 consequential R-R intervals and each segment of HRV(heart rate variability) signal includes time distance between every two consequential R-wave in ECG signals.

a) Time domain features

Time-domain parameters of HRV(heart rate variability) are the easiest as these are based on common statistical methods. In this paper, seven commonly used time domain features are as follows:

The mean value of the eight R-R intervals within each segment (Mean).

- The root mean square successive difference of the eight R-R intervals in each segment (RMSSD).
- The standard deviation of the 8 R-R intervals within each segment (SDNN).
- The standard deviation of differences between the adjacent R-R intervals within each segment (SDSD).
- The number of successive difference of intervals which differ by more than 50, 10 and 5 ms,

respectively, divided by 8, the total number of the RR Interval within each segment (pNN50, pNN10, pNN5).

b) Frequency domain features

LF/HF: Although time domain features are important to classify arrhythmias, they are not capable of distinction of sympathetic and parasympathetic content of the HRV signal. For this purpose, HRV signal is transformed into frequency domain and the ratio of spectral power in lower bound (0.04-0.15Hz) to spectral power in upper bound (0.15- 0.5Hz) is to be calculated. The lower bound frequency power is related to controlling temperature and cardiovascular mechanism and the upper frequency is related to the cardiac vagal activity.

c) Nonlinear features

The nonlinear properties of HRV can be analyzed using such as follow measures:

ApEn: Approximate entropy is use to measure the complexity or irregularity of the signal. Large values of ApEn indicates high irregularity and smaller values of ApEn implies higher regularity. The ApEn is computed as follows. For each segment in HRV signal with length N, u_j is defined as follow:

$$u_j = (RR_j, \dots, RR_{j+m-1}), j=1, \dots, N-m+1, \dots \quad (1)$$

where m is called the embedding dimension and N is the number of measured RR intervals. The distance between these vectors is defined as the maximum absolute difference between the corresponding elements. For each u_j the relative number of vectors u_k for which $d(u_j, u_k) \leq r$ is calculated by Eqs. (2) and (3)

$$d(u_j, u_k) = \max_{1 \leq i \leq m} |RR_{j+i} - RR_{k+i}| \quad |n = 0, \dots, m-1\}, \quad (2)$$

$$C_j^m(r) = \frac{|\{u_k | d(u_j, u_k) \leq r\}|}{N-m+1} \quad (3)$$

Due to the normalization, the value of $C_j^m(r)$ is always smaller or equal to 1. Afterward, the mean of natural logarithm of each $C_j^m(r)$ over j is taken to yield:

$$\varnothing^m(r) = \frac{1}{N-m+1} \sum_{j=1}^{N-m+1} \ln C_j^m(r). \quad (4)$$

Finally approximate entropy calculated by Eq. (5) $ApEn(m, r, N) = \varnothing^m(r) - \varnothing^{m+1}(r)$

$$ApEn(m, r, N) = \varnothing^m(r) - \varnothing^{m+1}(r) \quad (5)$$

For calculating ApEn for each HRV segment, the value of m and r are chosen as $m = 2$ and $r = 0.2SDNN$.

SpEn: Spectral entropy evaluates the HRV signal complexity in frequency domain [31]. Shannon Channel entropy estimates HRV entropy as

$$H = - \sum_f P_f \log(P_f), \quad (6)$$

where P_f is the value of the probability density function (PDF) of the process at frequency f.

V. FLOWCHART

Genetic programming is use to classify arrhythmias. Genetic programming was first introduced by Koza using tree representation. A member in a genetic programming has tree structure and as a result two types of genes, called functions and terminals are defined. In fact, the intermediate nodes act as functions and the leaves act as terminals. In this method, extracted features and some random numbers are used as terminals of the tree. The initialization of population is random and each tree is

evaluated by fitness function. Different genetic operators like crossover and mutation are used to make new trees and the new generation is generated. As shown in Fig. 2, this procedure is continued until the ending requirements are met. The procedure is explained in details in the following sub section.

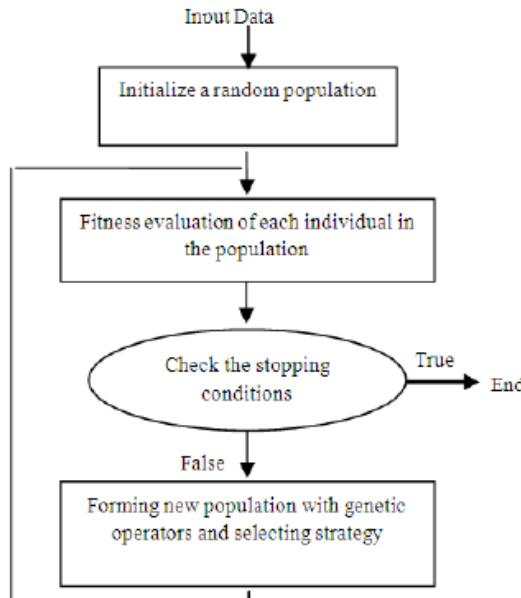


Fig.2 Schematically overview of Genetic Programming

A).Create an initial population:

The initial population is created by ramped half-and-half method. In this method, both Grow and Full methods are used to produce the initial population. For each depth level of trees considered, half of the individuals are initialized using the full method, and the other half using the grow method. The population of trees resulting from this initialization method is very diverse with balanced and unbalanced trees of several different depths. The individual's terminals are selected from extracted features and from set of random numbers. In this, we are using the following operators as function set: add, subtract, times, divide, sin, cos and if-then-else structures. While calculating X1/X2, if X2 is zero, X1 is returned as the response. The if-then-else structure is defined as follows:

$$\text{myif}(a, b, c) = \begin{cases} c & a < 0 \\ b & \text{otherwise} \end{cases}$$

When the initial population is created, each individual's fitness is measured. The value of each individual reflects the accuracy of output of genetic programming with respect to ideal output.

B. Generating new population

To generate a new population, some processes should be repeated until the number of new population is made equal to number of old population.

C. Genetic operators

After defining the fitness function, genetic operators should also be defined in order to improve the population.

Common genetic operators are like-reproduction, crossover and mutation. In reproduction, 10% of population in each generation is transmitted to the new population without any changes. The procedure of choosing individuals to transmit to the new population depends on selecting strategy.

In crossover, random nodes are chosen from both parent trees, and the sub trees starting at these two nodes are swapped resulting in two offspring. There is no bias in choosing internal or terminal nodes as the crossing sites. In tree mutation, a random node is chosen from the parent tree and the subtree starting by a randomly generated tree. This new random tree is created with the Grow initialization method and obeys the size/depth restrictions imposed on the trees created for the initial generation.

D. Selection strategy

A genetic programming without a suitable selection strategy is not different than random search. Some common approaches for selecting suitable parents are fitness proportionate selection, Greedy selection and Tournament selection. In this paper Lexicographic Parsimony Pressure Tournament strategy is chosen for selecting parents. In this strategy, like the Tournament selection, a random number of individuals are selected from the population and the best of them is determined. The main difference is that if two individuals are equally fit, the tree with fewer nodes is chosen as the best. This technique has shown to effectively control bloat in different types of problems.

VI. EXPECTED RESULTS

It is expected to classify arrhythmia of Normal beat (NB), Left Bundle Branch Block beat (LBBB), Right Bundle Branch Block beat (RBBB), Premature Ventricular Contraction (PVC), fusion of ventricular and normal beat (FUSION), Atrial Premature Contraction (APC) and paced beat (PACE), genetic programming produce the appropriate tree to distinguish any arrhythmia from others. For avoiding bias, the number of training data of each specific arrhythmia type has about the same size of the summation of other classes. The four commonly used measures of sensitivity, specificity, positive predictivity, and accuracy are derived for the proposed algorithm.

VII. CONCLUSION

This system used for classification of cardiac arrhythmia, each ECG signals divide to 8 consequential R-R intervals and HRV signal are extracted from it. The 12 features are extracted from HRV signals and 3 nonlinear features are extracted from ECG signals. To classify cardiac arrhythmias, these features are used as terminal in genetic programming and the algorithm selects affective features to distinguish each arrhythmia from others. Due to short processing time and relatively high accuracy of the used method, it can be used as a real-time arrhythmia classification system.

ACKNOWLEDGMENT

We are using this opportunity to express our gratitude to everyone who supported us for writing this paper. We are thankful for their guidance and invaluable advice during this work. We are sincerely grateful to them for sharing their truthful and illuminating views on a number of issues related to this paper.

I am also appreciative and genuinely thankful to our H.O.D. **Prof. Santosh G. Bari** and I am also thankful to our beloved Principal **Dr. V. N. Nitnaware** sir for providing us the basic infrastructure and other amenities



Mrs. Vaishali V. Thorat B.E , M.TECH (VLSI). Asst Prof and P.G. Co-ordinator of D.Y .Patil School of Engg, Ambi Pune.

REFERENCES

- [1] Luca Citi, Emery N. Brown, and Riccardo Barbieri, "A Real-Time Automated Point-Process Methodfor the Detection and Correction of Erroneous and Ectopic Heartbeats", IEEE TRANSACTIONS ON BIOMEDICAL ENGINEERING, VOL. 59, NO. 10, OCTOBER 2012.
- [2] Jure Kranjec, Samo Beguš, Janko Drnovšek, and Gregor Geršak, "Novel Methods for Noncontact Heart Rate Measurement: A Feasibility Study", IEEE TRANSACTIONS ON INSTRUMENTATION AND MEASUREMENT, VOL. 63, NO. 4, APRIL 2014.
- [3] Roberto Sassi, Katerina Hnatkova, Marek Malik, and Maria Gabriella Signorini, "HRV Scaling Exponent Identifies Postinfarction Patients Who Might Benefit From Prophylactic Treatment With Amiodarone", IEEE TRANSACTIONS ON BIOMEDICAL ENGINEERING, VOL. 53, NO. 1, JANUARY 2006.
- [4] Jes'us L'azaro, Eduardo Gil, Jos'e Mar'a Vergara, and Pablo Laguna, "Pulse Rate Variability Analysis for Discrimination of Sleep-Apnea-Related Decreases in the Amplitude Fluctuations of Pulse Photoplethysmographic Signal in Children", IEEE JOURNAL OF BIOMEDICAL AND HEALTH INFORMATICS, VOL. 18, NO.1, JANUARY 2014.
- [5] K. Minami, H. Nakajima, and T. Toyoshima, "Real time discrimination of ventricular tachyarrhythmia with Fourier transform neural network" IEEE Trans. Biomedical Engineering, vol. 46, NO. 2, 179-185, 1999.
- [6] Y. Wang, Y.S. Zhu, NV. Thakor and YH. Xu, "A short-time multifractal approach for arrhythmia detection based on fuzzy neural network" IEEE Trans. Biomedical Engineering, vol. 48, NO. 9, pp. 989-995, 2001.
- [7] S. Osowski and TH. Linh, "ECG beat recognition using fuzzy hybrid neural network". IEEE Trans. Biomedical Engineering, vol. 48, pp.1265-271, 2001.
- [8] OT. Inan, L. Giovangrandi and G. Kovacs, "Robust neural-networkbased classification of premature ventricular contractions using wavelet transform and timing interval features" IEEE Trans. Biomedical Engineering, vol. 53, NO. 12, pp.2507-515, 2006.

BIOGRAPHIES



Mr. Vyankatesh S. Thorat, Student of M.E of D.Y. Patil School of Engg, Ambi, Pune, Maharashtra, India



Dr. Suresh D. Shirbahadurkar [B.E (EC), M.E (EC), Ph.D(EC)] Principal of D.Y. Patil College of Engg. Ambi Pune. Maharashtra, India