International Journal on Applications in Information and Communication Engineering
Volume 1: Issue 7: July 2015, pp 11-13. www.aetsjournal.com ISSN (Online) : 2394-6237

ECG Abnormality Detection Using Time Series Motif

P.Vinoth, S.Ganesh

Abstract— Automatic detection of electrocardiogram (ECG) signals is very important for clinical diagnosis of heart disease.P-QRS-T wave representing the cardiac depolarization and re-polarization, recorded at the body surface. Analysis of ECG signal is proposed based on three step system Signal pre-processing, Feature Extraction, Classification.The suggested method can accurately classify and discriminate normal (Normal) and abnormal heartbeats. Abnormal heartbeats include left & Right bundle branch block (LBBB, RBBB),atrial premature contractions (APC) and premature ventricular contractions (PVC). Simulation results show that classification accuracy of 95.79% for the first dataset & an overall accuracy of detection of 95.18% are achieved over eight files from MIT/BIH arrhythmia database.

Keywords: ECG signals, preprocessing, feature Extraction, fuzzification and classification

I. INTRODUCTION

The analysis of ECG has been extensively used for diagnosing many cardiac diseases. A typical system for such an goal can be divided into three subsequent categories (preprocessing, feature extraction, and classification) in which N (normal), L (left bundle branch block), R (right bundle branch block), A (atrial premature contraction) and V (premature ventricular contraction) illustrate heartbeat. In the preprocessing step, frequencies of the ECG signal related to artifacts is detected and attenuated.

II. SIGNAL PREPROCESSING

De-noising is an important issue for the analysis of signals. Most troublesome noise sources are electrical activity of muscles To eliminate such noise, an advanced signal processing technique using discrete wavelet transform (DWT) denoising technique should be use

III. FEATURE EXTRACTION

The ECG signal are pre-processed from noise for better accuracy and further discredited it to symbolic characters for efficient mining. The noise artifacts that generally affects ECG signals is baseline wandering. Normally it appears from respiration and lies between 0.15 and 0.3Hz. The baseline wander of ECG signal is eliminated using moving average window which diminishes the irregularities in beats To perform pre-processing of univariate Time-series signals to discrete symbolic sequence, a Symbolic Aggregate approximation (SAX) is used. SAX allows a Time-series signals of arbitrary length to be reduced to a string (symbolic representation) with alphabets arbitrary of length, where typically the alphabet set size is also an arbitrary integer value.

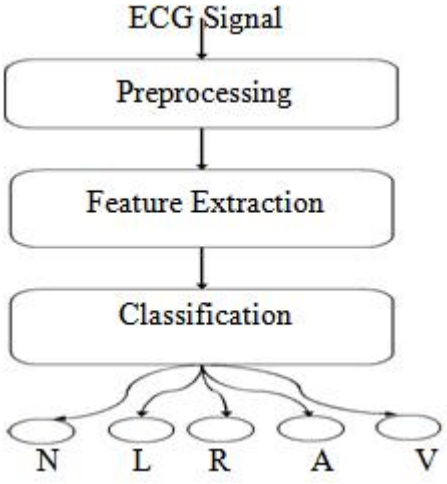

Figure1: Block Diagram Of ECG System

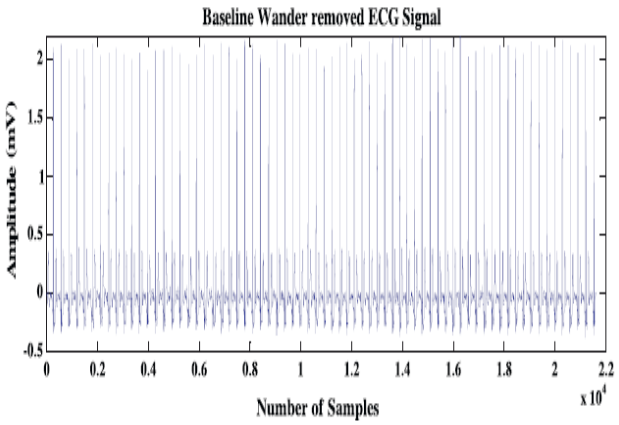

Figure 2 : Base Line Wandering

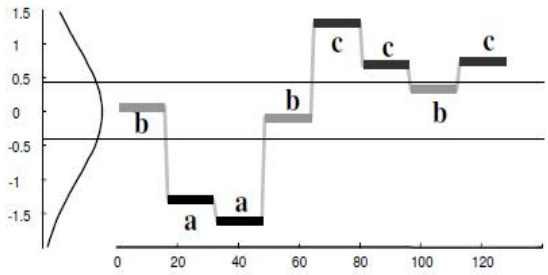

Figure 3: Removal & SAX

P.Vinoth, PG student, Fatima Michael College of Engineering and Technology. (Email: gstarvinoth92@gmail.com)
S.Ganesh, Assistant professor, Fatima Michael College of Engineering and Technology. (Email: s.ganeshvel@gmail.com)

IV. FUZZIFICATION(OPTIMIZATION)

Fuzzification is a process which performs machine learning techniques methodology. The heartbeats are observed by ECG machine which convert the beats into a waveform which cannot be understood by human interface. Hence this technique involves learning machine language and show the normal & abnormal level.

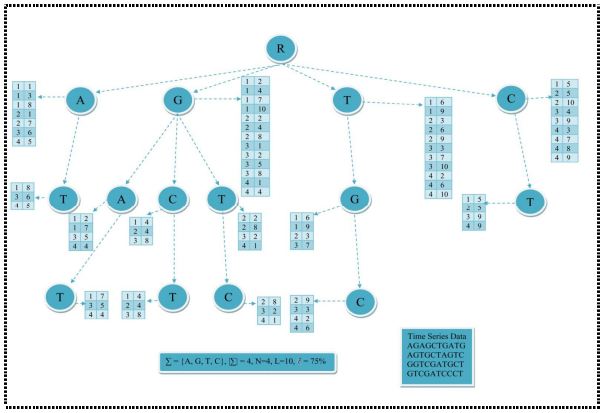

Figure4: Optimization(Machine Learning)

V.CLASSIFICATION

This is a final module which shows output values detection of normal and abnormal heart beat levels. This classifies the output values of normal heart beat rates as well as abnormal heart beat rates. Also classifies the time series motif values.

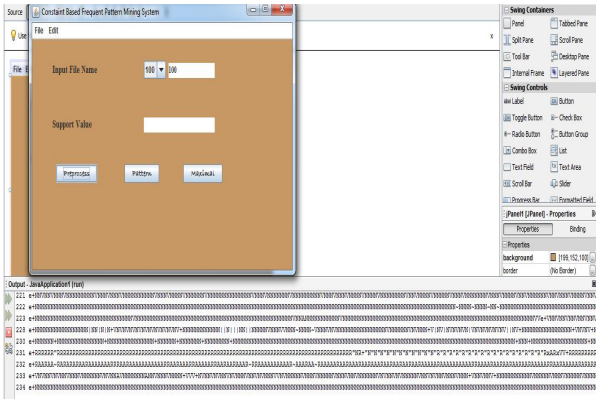

Figure 5: Time Series Motif System

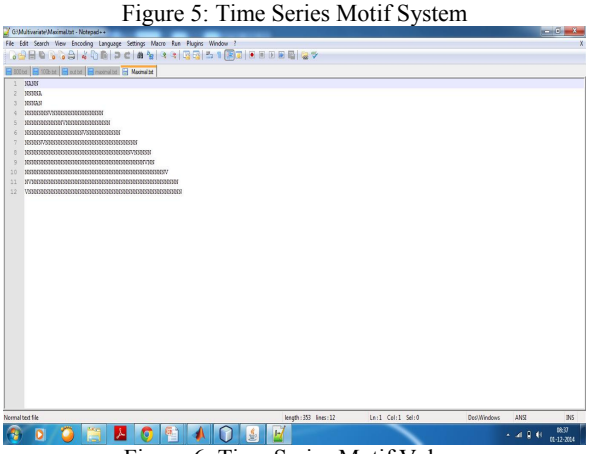

Figure 6: Time Series Motif Values

The Classification System also shows an Feature Vector For Five Classes Such As

- Normal Heart Beats
- Abnormal Heart Beats
- Left Bundle Branch Block (LBBB)
- Right Bundle Branch Block(RBBB)
- Arterial Premature Contraction(APC)
- Premature Ventricular Contraction(PVC)

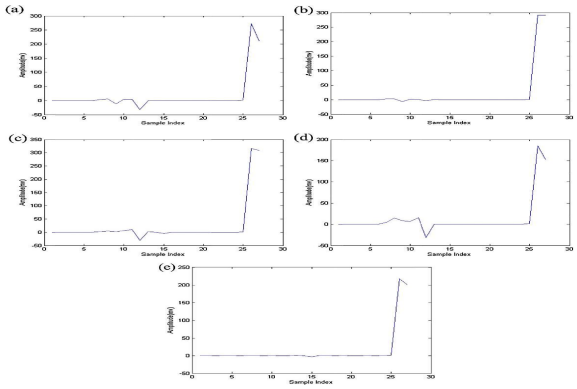

Figure 7: Normal Beats, LBBB, RBBB, PVC, APC

VI. EXPECTED RESULTS

Experiments have been performed random 4000 beats (800 from each beat type) are chosen from all the 18,299 beats. 75 beats from each class are used for train and the others are used to test the classifiers. For the second experiment, all of 18,299 beats are used to obtain the overall performance of classifier. However, the size of training samples for each class did not varied. The final experiment shows that the pro-posed method is able to generalize, because the all training set was 375 samples (approximately 2% of all beats in experiment).We found that the recognition rates of pro-posed method were significantly higher than the other combining techniques at the 99% confidence level for arrhythmia classification in ECG.

VII. CONCLUSION

In this work, we have discussed motif discovery and anomaly detection for time series data. Experiments also show that our approach can efficiently find motifs and anomalies in a large dataset. Our work provides a general solution motif discovery and anomaly detection on any time series data. In the future, we will apply this work in some specific domains, such as Data mining

REFERENCES

[1] H.M. Rai , A.Trivedi , S.Shukla, ECG signal processing for abnormalities detection using multi-resolution wavelet transform and Artificial Neural Network classifier, Measurement 46 (2013) 3238–3246.

[2] P.C. Chang, J.J. Lin, J.C. Hsieh, J. Weng, Myocardial infarction classification with multi-lead ECG using hidden Markov models and Gaussian mixture models ,Appl. Soft Compute. 12 (2012) 3165–3175.

[3] Z. Zidelmal, A. Amirou, D. Ould-Abdeslamb, J. Merckle, ECG beat classification using a cost sensitive classifier, Compute. Methods Programs Biomed. 111 (2013)570–577.

[4] C. Bu, Y. Gao, K. Mathee , G. Narasimhan, X. Wang, and N. Xu,(2002) ,"Mining Protein Sequences for Motifs," J. Computational Biology, vol. 9, no. 5, pp. 707-720.

[5] Lee Yeuan-kuen et al, "An Advanced Least-Significant- Bit Embedding Scheme for Steganographic Encoding", 5414/2009: pp. 349-360, 2009.

[6] L. Marsan and M.-F. Sagot, (2000), "Algorithms for Extracting Structured Motifs Using a Suffix Tree with Application to Promoter and Regulatory Site Consensus Identification," J. Computational Biology, vol. 7, nos. 3/4, pp. 345-36

[7] M. J. Zaki, (Jan/Feb 2001) ,"SPADE: An Efficient Algorithm for Mining Frequent Sequences", in Machine Learning Journal, special issue on Unsupervised Learning (Doug Fisher, ed.),pages 31-60, Vol. 42 Nos.

[8] J. Wang and J. Han (2004), "BIDE: Efficient Mining of Frequent Closed Sequences," Proc. 20th IEEE Int'l Conf. Data Eng. (ICDE), pp. 79-90.

[9] J. Han, L.V.S. Lakshmanan, T.N.Raymond, (1999), "Constraint-Based Multidimensional Data Mining," IEEE Computer, 32(8), 46-50.

[10] JigneshM.Patel,(August 2011), "FLAME: Efficient and Accurate Discovery of Patterns in Sequence Data Sets,"P.No 1154 – 1168, vol 23