International Journal of Medical, Medicine and Health Sciences
ISSN: 2517-9969
Vol:10, No:2, 2016

HRV Analysis Based Arrhythmic Beat Detection Using kNN Classifier

Onder Yakut, Oguzhan Timus, Emine Dogru Bolat

***Abstract*—**Health diseases have a vital significance affecting human being's life and life quality. Sudden death events can be prevented owing to early diagnosis and treatment methods. Electrical signals, taken from the human being's body using non-invasive methods and showing the heart activity is called Electrocardiogram (ECG). The ECG signal is used for following daily activity of the heart by clinicians. Heart Rate Variability (HRV) is a physiological parameter giving the variation between the heart beats. ECG data taken from MITBIH Arrhythmia Database is used in the model employed in this study. The detection of arrhythmic heart beats is aimed utilizing the features extracted from the HRV time domain parameters. The developed model provides a satisfactory performance with ~89% accuracy, 91.7 % sensitivity and 85% specificity rates for the detection of arrhythmic beats.

***Keywords*—**Arrhythmic beat detection, ECG, HRV, kNN classifier.

I. Introduction

ARRHYTHMIAS in cardiovascular system affect regular rhythmic activity of the heart. It is an important issue to detect the arrhythmia beats within a normal heart activity in clinical cardiology. So, a significant amount of studies on automatic arrhythmia detection using machine learning for classification the ECG signal can be seen in literature. Machine learning based decision support systems ease the clinicians to analyze the heart activity. Therefore, various features extracted from the ECG signal and classification methods are proposed in literature.

Xue et al. propose an android based application classifying the arrhythmia using HRV analysis for monitoring the situation of the heart [1]. The present databases for the ECG, different preprocessing techniques for filtering the signal and the performance of Artificial Neural Network (ANN) based classifiers are searched by Jambukia et al. [2]. Benmalek et al. suggest a methodological approach for analyzing cardiac arrhythmias utilizing short time Fourier transform [3]. Sherbakova et al. propose an algorithm designed for telemetry systems and detecting the basic arrhythmias within the ECG signal [4]. Timus designs a reliable pattern recognition system using soft computing algorithms for classifying sleep apneas from the ECG signal at a high accuracy rate [5]. Sarma et al. propose an approach detecting the arrhythmias within the ECG signal employing the wavelet based feature extraction methods and multilayer perceptron (MLP) classifier [6]. Zairi et al. suggest an FPGA based application determining the cardiac arrhythmias [7]. Boonperm et al. develop a Matlab GUI detecting the cardiac arrhythmia types by processing the ECG signal [8]. Balachandran et al. propose a multi-resolution wavelet transform based feature extraction method for increasing the performance of the ECG analysis [9]. Yakut et al. develop an ECG measuring system using e-Health Sensor Shield and obtain the HRV parameter utilizing the Pan-Tompkins QRS detection algorithm [10]. A web-based wireless ECG measuring and recording system is realized by Yakut et al. [11]. This e-health application is ready for use including the specifications of being portable, compact, easy to use, user friendly and real time. It is aimed to support the clinicians while diagnosing, treating and following up the patients.

In this study, R peaks and rhythm annotations are taken from the MITBIH arrhythmia database [12]. 17727 heart beats are obtained from total 8 recordings. 66% of 17727 heart beats is used for training data while the left 6027 beats are utilized for test data. The features are extracted from the time domain HRV parameters and the effective features are selected using the Wrapper based approach. The arrhythmic beats are detected employing the selected features and the k-Nearest Neighbor (kNN) classifier.

II. Materials And Methods

The proposed model in this study is given in Fig. 1. The ECG recordings are provided from MIT-BIH Arrhythmia Database using Matlab program. RR intervals, employed for calculating the HRV parameter, are obtained using the R peaks and rhythm annotation taken from the database. Then the HRV is derived utilizing the obtained RR intervals. Finally, a feature set is prepared calculating the HRV time domain parameters.

The features are extracted from the time domain HRV parameters. The feature selection process is realized using the Wrapper based approach to perform the classification with the most effective features, to save the processing load and time. Finally, the feature vector is obtained using these optimal selected features.

Working with the optimized parameters instead of default parameters affects the performance of the classification positively during the classification process [5]. In this study, nonparametric and supervised learning classifier kNN is used

O. Yakut is with the Information Technologies Department, Kocaeli University, Umuttepe, Izmit, 41380 Turkey (e-mail: onderyakut.41@gmail.com).

O. Timus is with the Inventory Control Center, Turkish Navy, Golcuk, Kocaeli, Turkey (e-mail: oguzhantimus@gmail.com).

E. D. Bolat is with the Biomedical Engineering Department, Faculty of Technology, Kocaeli University, Umuttepe, Izmit, 41380 Turkey (Corresponding author, Phone: +90 262 303 22 39; Fax: +90 262 303 22 03; e-mail: ebolat@gmail.com).

because of its high classification capability. The classification realized in two classes as normal or arrhythmic beats with a satisfactory performance.

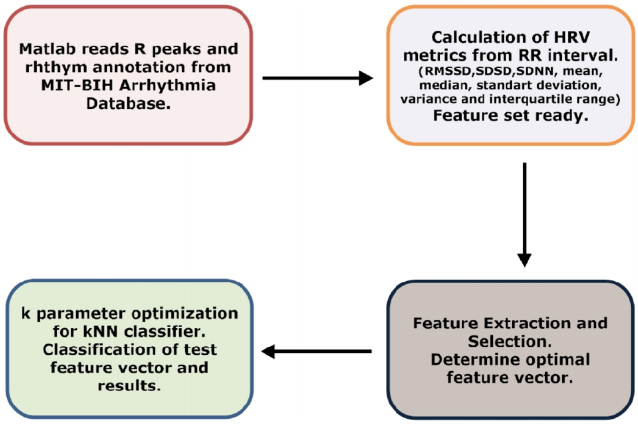

Fig. 1 Block diagram of the proposed model

A. ECG Recordings

ECG recordings of MIT-BIH Arrhythmia Database which is served by a free web access Physionet, are used in this study. This database includes 48 30-minute two channel ECG recordings. These recordings are digitized at a 360 Hz sampling frequency, a ±5 mV amplitude and 11 bit resolution for each channel [12].

TABLE I
ECG BEAT TYPE DISTRIBUTION

The number of subject	103	109	113	118	119	200	202	220
Normal Sinus Rhythm	x	x	x		x	x	x	x
Left Bundle Branch Block Beat		x						
Sinus Arrhytmia			x					
Right Bundle Branch Block Beat				x				
Ventricular Trigeminy					x			
Ventricular Tachycardia						x	x	
Sinus Bradycardia							x	
Atrial Premature Contraction							x	
Atrial Flutter							x	
Atrial Fibrillation							x	
Atrial Couplet								x
Premature Ventricular Contraction		x		x	x	x	x	
Fusion of Ventricular and Normal Beat		x				x	x	
Aberrated Atrial Premature Beat			x				x	
Atrial Premature Beat	x			x		x	x	x

In this study, a feature vector is prepared using ECG recordings of 8 subjects having arrhythmic beat types. The used ECG recordings have a total number of 17727 heart beats. The beat types of the ECG recordings are given in Table I.

B. RR Interval

In this study, R peaks of ECG signals and rhythm annotation are taken from the MIT-BIH Arrhythmia Database and used for feature extraction.

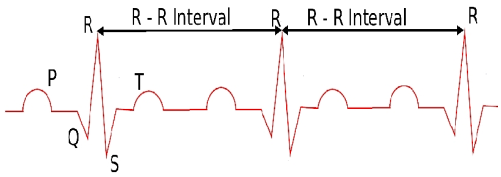

Fig. 2 RR Interval between two consecutive R peaks [10]

As seen in Fig. 2, the interval between the two consecutive R peaks is calculated using (1):

$$RR(i) = R(i+1) - R(i) \tag{1}$$

HRV is the change in the time intervals between adjacent heart beat as (1). HRV also decreases during periods of mental stress. HRV is regulated by the autonomic nervous system. Parasympathetic activity decreases heart rate and increases HRV, whereas sympathetic activity increases heart rate and decreases HRV [13].

In this study, feature calculations using (2) [14] and (3)-(5) [15], based on HRV time domain parameters, are performed. A sliding window including 12 RR interval segments is employed during the feature extraction process.

HRV metrics are given between (2) and (5). In these equations; N, number of RR interval; RR, mean of RR interval;

$$\overline{RR} = \frac{1}{N-1}\sum_{n=1}^{N}(RR_{(n)}) \tag{2}$$

SDNN, standard deviation of all RR intervals;

$$SDNN = \sqrt{\frac{1}{N-1}\sum_{n=1}^{N}[RR_{(n)} - \overline{RR}]^2} \tag{3}$$

RMSSD, the square root of the mean of the sum of the squares of differences between adjacent RR intervals;

$$RMSSD = \sqrt{\frac{1}{N-1}\sum_{n=1}^{N-1}[RR_{(n+1)} - RR_n]^2} \quad (4)$$

SDSD, standard deviation of differences between adjacent RR intervals; $\Delta R_{(i)}$; differences between adjacent RR interval; $\overline{\Delta R}$, mean of $\Delta R_{(i)}$; N, number of $\Delta R_{(i)}$ intervals.

$$\Delta R_{(i)} = RR_{(i+1)} - RR_{(i)}$$

$$SDSD = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N-1}[\Delta R_{(i)} - \overline{\Delta R}]^2} \quad (5)$$

C. Feature Extraction and Selection

In this study, the parameters as RMSSD, SDSD, SDNN, mean, median, standard deviation, variance and interquartile range are extracted from the RR intervals derived from the ECG signal to get the feature vector. This vector is optimized discarding the irrelevant and inefficient features. Wrapper based feature selection algorithm is used for this purpose. All the features are selected and transferred to the feature vector at the end of this process.

D. Classification

kNN is used as a classifier in this study. Nonparametric and supervised learning classifier, kNN has a significant place among classifiers and is commonly preferred in various science and engineering applications. It is resistant to noisy data sets and its efficiency increases with the number of data set. Since the only k parameter is required, kNN is a classifier which is easy to use [5].

The optimal k parameter is defined using different k values with the training data set and the most effective features. Then, this optimal k value is used in the test data.

Total number of heartbeats is 17727 in this study. 66% of this data is used for training while the left 6027 beats are separated for test. The accuracy, sensitivity, and specificity values obtained during optimizing k value are given in Table II and k parameter optimization graph is seen in Fig. 3.

TABLE II
k PARAMETER OPTIMIZATION

k	Accuracy	Sensitivity	Specificity
1	0.80589744	0.847365	0.708904
3	0.82435897	0.865718	0.732044
5	0,82865672	0.866388	0.743397
7	0,83316239	0.869576	0.75007
9	0,83683761	0.874012	0.75333
11	0,8417094	0.877952	0.760587
13	0,84324786	0.879441	0.762497
15	0,84484527	0.88139	0.763881
17	0,84435897	0.881236	0.762768
19	0,84410256	0.881666	0.761423
21	0,8457265	0.882324	0.764819
23	**0,84615385**	**0.883253**	**0.764561**
25	0,8465812	0.882849	0.766337
27	0,84529915	0.881966	0.764205
29	0,84521368	0.882426	0.763273
31	0,84606838	0.884289	0.762772

III. RESULTS

The classification of test data is realized in Weka program using the trained model and 10 fold cross validation technique. The optimal k value of kNN classifier is found and used as 23 in test. At the end of classification process, arrhythmic heart beats are obtained.

The obtained Receiver Operating Curve (ROC) is depicted in Fig. 4. Area Under Curve (AUC) is obtained as 0.9401. AUC shows how much the classification can distinguish the arrhythmic beats. The gradation of AUC is given in Table III [16].

TABLE III
GRADATION OF AUC

0.90 – 1.00	Excellent (A)
0.80 - 0.90	Good (B)
0.70 - 0.80	Fair (C)
0.60 – 0.70	Poor (D)
0.50 – 0.60	Fail (F)

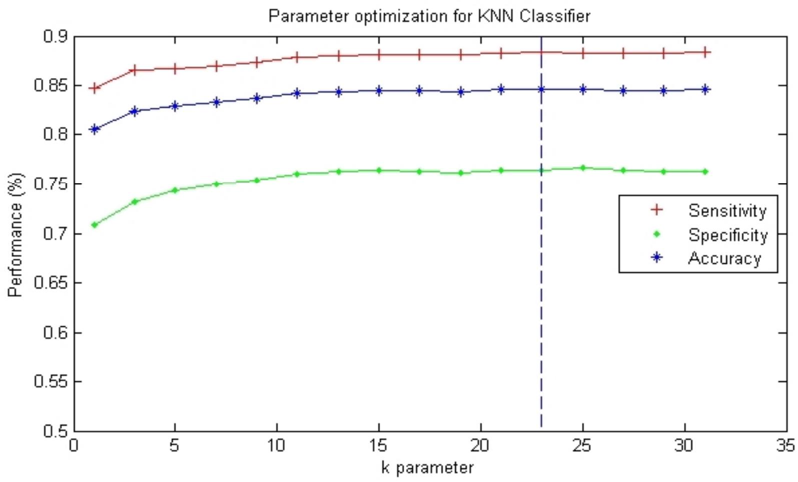

Fig. 3 k parameter optimization for kNN classifier

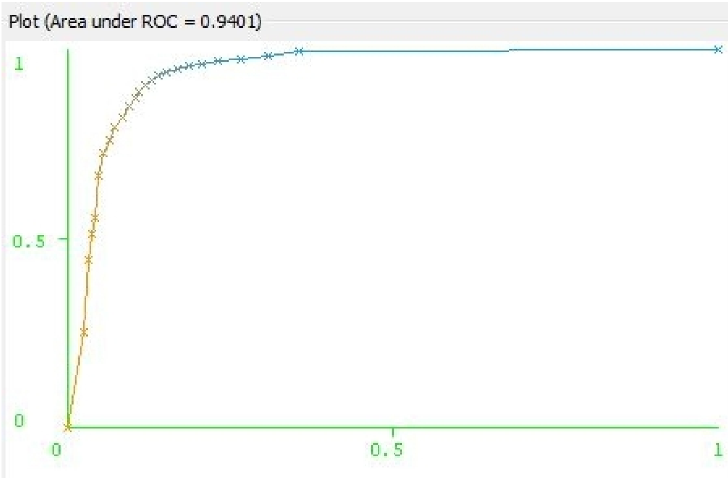

Fig. 4 The obtained ROC curve

TABLE IV
STATISTICAL CRITERION OBTAINED FOR CLASSIFICATION

Accuracy (%)	Sensitivity (%)	Specificity (%)
88.83	91.68	85.17

The statistical criterion results obtained for classification of the arrhythmic beats is shown in Table IV.

IV. CONCLUSION

The developed model classifies the arrhythmic beats at about 89% accuracy rate in the test data. The model has the capability to determine the true positive (sensitivity) arrhythmic beats at a rate of about 91.7% and it can distinguish the true negative (specificity) arrhythmic beats at a rate of about 85%. The obtained AUC as 0.9401 is excellent according to Table III. In this study, the only HRV time domain parameters are used to determine the arrhythmic beats. At the conclusion of the study, it is found out that the used features are simple and appropriate for a rapid diagnosis. This model can be improved to classify the arrhythmic beats and arrhythmia types using HRV time and frequency domain methods in future studies.

ACKNOWLEDGMENT

Kocaeli University Scientific Research Projects Unit supports this study with the project number as BAP 2014/69HDP.

REFERENCES

[1] Xue, S., Chen, X., Fang, Z., Xia, S., “An ECG arrhythmia classification and heart rate variability analysis system based on android platform”, 2nd International Symposium on Future Information and Communication Technologies for Ubiquitous HealthCare, Beijing, China, 2015, pp. 1-5.

[2] Jambukia, S.H., Dabhi, V.K., Prajapati, H.B., “Classification of ECG signals using machine learning techniques: A survey”, International Conference on Advances in Computer Engineering and Applications, Ghaziabad, India, 2015, pp. 714-721.

[3] Benmalek, E., Elmhamdi, J., “Arrhythmia ECG signal analysis using non parametric time-frequency technique”, International Conference on Electrical and Information Technologies, Marrakech, 2015, pp. 281-285.

[4] Sherbakova, T.F., Sherbakov, G.I., Sedov, S.S., “Cardiac arrhythmia analysis used in the system of electrocardiosignal transmission to the supervision center”, International Conference on Antenna Theory and Techniques, Kharkiv, 2015, pp. 1-3.

[5] Timus, O., “Sleep Respiration Disorders Diagnosis and Classification Utilizing Soft Computing Algorithms”, Phd. thesis, Natural and Applied Sciences Eletronics and Computer Education Department, Kocaeli University, Kocaeli, Turkey, 2015.

[6] Sarma, P., Nirmala, S.R., Sarma, K.K., “ECG classification using wavelet subband energy based features”, International Conference on Signal Processing and Integrated Networks, Noida, 2014, pp. 785-790.

[7] Zairi, H., Kedir-Talha, M., Benouar, S., Ait-Amer, A., “Intelligent system for detecting cardiac arrhythmia on FPGA”, 5th International Conference on Information and Communication Systems, Irbid, 2014, pp. 1-5.

[8] Boonperm, P., Supakasemwong, D., Naiyanetr, P., “ECG analyzing program for arrhythmia detection”, 7th Biomedical Engineering International Conference, Fukuoka, 2014, pp. 1-4.

[9] Balachandran, A., Ganesan, M., Sumesh, E.P., “Daubechies algorithm for highly accurate ECG feature extraction”, International Green Computing Communication and Electrical Engineering, Coimbatore, 2014, pp. 1-5.

[10] O. Yakut, S. Solak, E. Dogru Bolat, “Measuring ECG Signal Using e-Health Sensor Platform”, International Conference on Chemistry, Biomedical and Environment Engineering, Antalya, 2014, pp. 71-75.

[11] O. Yakut, S. Solak, E. Dogru Bolat, “Implementation of a Web-Based Wireless ECG Measuring and Recording System”, 17th International Conference on Medical Physics and Medical Sciences, Istanbul, 2015, vol. 9(10), pp. 815-818.

[12] MIT-BIH Arrhythmia Database. Available: http://physionet.org (accessed October, 2015).

[13] http://www.polar.com/us-en/support/Heart_Rate_Variability__HRV_

[14] Wang, H.M., Sheng-Chieh, H., “SDNN/RMSSD as a Surrogate for LF/HF: A Revised Investigation”, Modelling and Simulation in Engineering, vol. 2012, pp. 1-8, June, 2012.

[15] Kampouraki, A., Manis, G., Nikou, C., “Heartbeat Time Series Classification with Support Vector Machines”, IEEE Transactions on Information Technology in Biomedicine, vol. 13, no. 4, pp. 512-518, July, 2009.

[16] http://gim.unmc.edu/dxtests/roc3.htm (accessed October, 2015).