IJCSE International Journal of Computer Sciences and Engineering Open Access
Research Paper Vol.-6, Issue-8, Aug 2018 E-ISSN: 2347-2693

Implementation of K-Nearest Neighbor (KNN) algorithm for detection of QRS Complexes

P. Mathur[1*], V.S. Chouhan [2]

[1]Research Scholar, JNU Jodhpur, India
[2]Dept. of ECE, MBM Engineering College, Jodhpur, India

Corresponding Author: peeyush14_mathur@yahoo.com

Available online at: www.ijcseonline.org

Accepted: 18/Aug/2018, Published: 31/Aug/2018

Abstract**—** In this paper, K-Nearest Neighbor (KNN) algorithm as a classifier is implemented with slope as feature for detection of QRS-complex in ECG, the detection rate of 99.32% is achieved. The proposed algorithm is evaluated on standard databases CSE dataset-3.

Keywords**—**K-NN Alogorithm, QRS detection.

I. Introduction

The impact of digital signal processing techniques promoted revolutionary advances in many fields of application e.g. biomedical engineering, speech communication, data communication, nuclear science and many others. Electrocardiogram (ECG) is one of the most important electrical signals in the field of medical science which has a great need to be processed before further analysis. Arrhythmias or abnormalities of the heart rhythm can be detected using electrocardiograms (ECGs) that record the electrical activity of the heart. Cardiac arrhythmias occur when disturbances are caused in the normal electrical events related to the basic process of automaticity, conduction and triggering mechanisms of the heart.

A standard ECG record is the best test for diagnosing all arrhythmias, whether of ventricular or supraventricular origin. An ECG tracing is a series of waves that represent the electrical events of the various chambers and conduction pathways within the heart.

In this paper we have implemented K-Nearest Neighbor (KNN) algorithm as a classifier is implemented with slope as feature for detection of QRS-complex in ECG. The paper is organized as initially a brief theory about k-NN classifier is discussed followed by the implementation of proposed algorithm for identification of QRS complexes. Than the results for various subjects are shown, followed by the conclusion at the last.

II. KNN Classifier

In pattern recognition, the k-nearest neighbors algorithm (k-NN) [1] is a non-parametric method used for classification and regression [2]. In both cases, the input consists of the k closest training examples in the feature space. The output depends on whether k-NN is used for classification or regression:

- In k-NN classification, the output is a class membership. An object is classified by a majority vote of its neighbors, with the object being assigned to the class most common among its k nearest neighbors (k is a positive integer, typically small). If k = 1, then the object is simply assigned to the class of that single nearest neighbour [3].
- In k-NN regression, the output is the property value for the object. This value is the average of the values of its k nearest neighbors.

k-NN is a type of instance-based learning, or lazy learning, where the function is only approximated locally and all computation is deferred until classification. The k-NN algorithm is among the simplest of all machine learning algorithms.

Both for classification and regression, a useful technique can be to assign weight to the contributions of the neighbors, so that the nearer neighbors contribute more to the average than the more distant ones. For example, a common weighting scheme consists in giving each neighbor a weight of 1/d, where d is the distance to the neighbor.

The neighbors are taken from a set of objects for which the class (for k-NN classification) or the object property value (for k-NN regression) is known. This can be thought of as the training set for the algorithm, though no explicit training step is required.

A peculiarity of the k-NN algorithm is that it is sensitive to the local structure of the data. The algorithm is not to be confused with k-means, another popular machine learning technique.

Algorithm

A case is classified by a majority vote of its neighbors, with the case being assigned to the class most common amongst its K nearest neighbors measured by a distance function. If K = 1, then the case is simply assigned to the class of its nearest neighbor.

Distance functions	
Euclidean	$\sqrt{\sum_{i=1}^{k}(x_i - y_i)^2}$
Manhattan	$\sum_{i=1}^{k}\lvert x_i - y_i\rvert$
Minkowsi	$\left(\sum_{i=1}^{k}(x_i - y_i)^q\right)^{1/q}$

III. Identification of QRS-Complexes

The input to the algorithm for QRS detection is 12 lead simultaneously recorded ECG vectors.

Step 1: A raw digital single lead ECG signal of a patient is acquired.

Step 2: Raw ECG signal is often contaminated by disturbances such as power line interference and baseline wander. The finite impulse response (FIR) notch filter proposed by Van Alste and Schilder [4] is used to remove baseline wander. The adaptive filter to remove base line wander is a special case of notch filter, with notch at zero frequency (or dc). This filter has a "zero" at dc and consequently creates a notch with a bandwidth of (μ/π)*fs, where fs is the sampling frequency of the signal and μ is the convergence parameter. Frequencies in the range 0-0.5Hz are removed to reduce the base line drift. The filter proposed by Furno and Tompkins [5] is used to remove 50Hz power line interference.

Step 3: The absolute value of slope at every sampling instant is calculated to enhance the signal in the region of QRS complex. These slope values are then normalized. The slope is used as an important discriminating feature because slope of the signal is much more in the QRS region than in the non-QRS-region.

Step 4: K-NN algorithm is applied in order to cluster the QRS and non-QRS regions. An output of 1 is marked or labeled if the sample belongs to the QRS cluster and 0 for the non-QRS cluster.

Step 5: After clustering has been done using K-NN algorithm, the slope curve is scanned. The membership of slope, at a given sampling instant, is found means whether it belongs to QRS or non-QRS domain and accordingly it is labeled as 1 or 0 respectively. It is observed that continuous train of 1's and 0's is obtained in the QRS and non-QRS region respectively.

IV. Results

Fig. 1 display QRS-detection in record MO1_004. All the wave components are normal in this case. Hence all the QRS-complexes have been correctly identified by the algorithm.

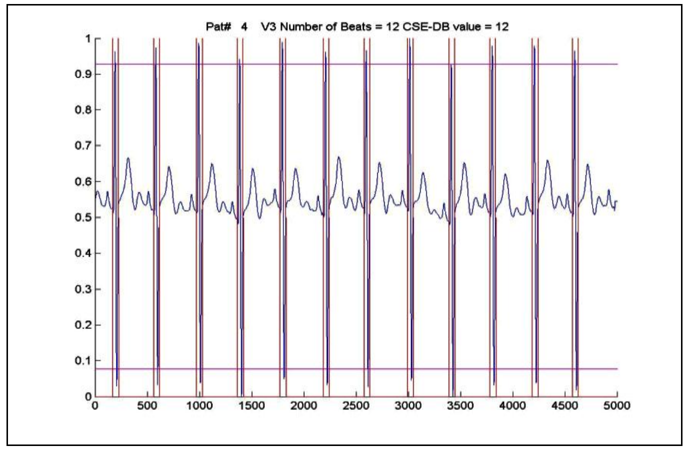

Fig. 1 Detection of QRS-complexes of record MO1_004

Fig. 2 shows QRS-detection in lead V3 of record MO1_007. T-waves are peaky in this case. Even though the T-wave are of higher amplitude compared to the amplitude of the QRS-complexes, amplitude of their slopes and further the squared slope signal is very low hence these peaky T-waves are correctly not detected as QRS-complexes by the algorithm.

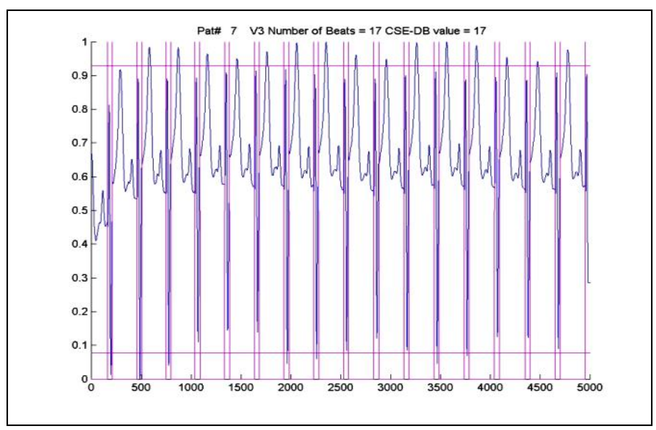

Fig. 2 Detection of QRS-complexes of record MO1_007

Fig. 3 illustrates QRS-detection of record MO1_103. Morphology of some of the QRS-complexes and T-waves is different in this record. The KNN based algorithm correctly detects all the QRS-complexes of different morphologies.

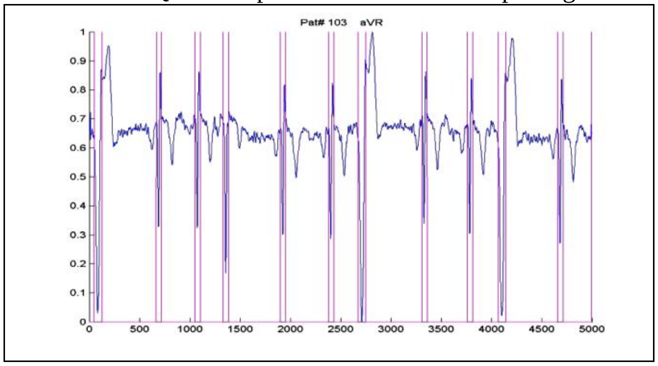

Fig. 3 Detection of QRS-complexes of record MO1_103

V. CONCLUSION

A successful attempt has been made in the present research work for the detection of QRS-complexes in complexes in simultaneously recorded 12-lead ECG signal using KNN Algorithm using slope as feature, the detection rate of 99.32 % is acheived. The algorithm has been developed and tested against the data-set 3 of CSE multi-lead measurement library [6].

REFERENCES

[1] https://en.wikipedia.org/wiki/K-nearest_neighbors_algorithm

[2] N. S. Altman, "An introduction to kernel and nearest-neighbor nonparametric regression". The American Statistician, Vol. 46, Issue 3, pp. 175–185, 1992.

[3] P. A. Jaskowiak, R. J. G. B. Campello, "Comparing Correlation Coefficients as Dissimilarity Measures for Cancer Classification in Gene Expression Data". http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.208.993. Brazilian Symposium on Bioinformatics (BSB 2011). pp. 1–8. Retrieved 16 October 2014.

[4] J.A. Alste Van and T.S. Schilder, "Removal of baseline wander and power line interference from the ECG by an efficient FIR filter with a reduced number of taps," IEEE Transactions on Biomedical Engineering, Vol. 32, Issue 12, pp.1052-1059, 1985.

[5] G.S. Furno and W.J. Tompkins, "A learning filter for removing noise interference," IEEE Transactions on Biomedical Engineering, Vol. 30, issue 4, pp. 234-235, 1983.

[6] Willems J.L., "The CSE Multilead atlas measurement results data set 3," Common Standards for Quantitative Electrocardiography, Leuven, Belgium, 1988

Authors Profile

Peeyush Mathur received M.Tech degree in Computer Science and Engineering from GGSIPU, Delhi. Presently he is Ph.D. scholar at Jodhpur National University,Jodhpur. His area of interest is Soft Computing, Computer Networks, Software Engineering.

Dr. V.S. Chouhan received B.E. degree in Electronics & Communication Engineering, M.E. degree in Digital Communication Engineering and Ph.D. in Bio-medical Signal Processing from J. N.V. University, Jodhpur (India). Presently he is Professor in the Department of Electronics & Communication Engineering at M.B.M. Engineering College, J.N.V. University, Jodhpur. His research interests include fields of Biomedical Signal Processing, Soft Computing and Digital Communications.

Mr C H Lin pursed Bachelor of Science and Master of Science from University of New York, USA in year 2009. He is currently pursuing Ph.D. and currently working as Assistant Professor in Department of Telecommunication, University of New York, USA since 2012. He is a member of IEEE & IEEE computer society since 2013, a life member of the ISROSET since 2013 and ACM since 2011. He has published more than 20 research papers in reputed international journals including Thomson Reuters (SCI & Web of Science) and conferences including IEEE and it's also available online. His main research work focuses on Cryptography Algorithms, Network Security, Cloud Security and Privacy, Big Data Analytics, Data Mining, IoT and Computational Intelligence based education. He has 5 years of teaching experience and 4 years of Research Experience.

