Evaluation of patient electrocardiogram datasets using signal quality indexing

Nazrul Anuar Nayan, Hafifah Ab Hamid
Center for Integrated Systems Engineering and Advanced Technologies
Faculty of Engineering and Built Environment, Universiti Kebangsaan Malaysia, UKM Bangi Selangor, Malaysia

Article Info

Article history:

Received Jul 31, 2018
Revised Jan 7, 2019
Accepted Feb 1, 2019

Keywords:

ECG
Heart rate
Peak detectors
Signal quality indexing
Vital sign

ABSTRACT

Electrocardiogram (ECG) is widely used in the hospital emergency rooms for detecting vital signs, such as heart rate variability and respiratory rate. However, the quality of the ECGs is inconsistent. ECG signals lose information because of noise resulting from motion artifacts. To obtain an accurate information from ECG, signal quality indexing (SQI) is used where acceptable thresholds are set in order to select or eliminate the signals for the subsequent information extraction process. A good evaluation of SQI depends on the R-peak detection quality. Nevertheless, most R-peak detectors in the literature are prone to noise. This paper assessed and compared five peak detectors from different resources. The two best peak detectors were further tested using MIT-BIH arrhythmia database and then used for SQI evaluation. These peak detectors robustly detected the R-peak for signals that include noise. Finally, the overall SQI of three patient datasets, namely, Fantasia, CapnoBase, and MIMIC-II, was conducted by providing the interquartile range (IQR) and median SQI of the signals as the outputs. The evaluation results revealed that the R-peak detectors developed by Clifford and Behar showed accuracies of 98% and 97%, respectively. By introducing SQI and choosing only high-quality ECG signals, more accurate vital sign information will be achieved.

Corresponding Author:

Nazrul Anuar Nayan,
Center for Integrated Systems Engineering and Advanced Technologies,
Faculty of Engineering and Built Environment, Universiti Kebangsaan Malaysia,
UKM Bangi Selangor, Malaysia.
Email: lsntl@ccu.edu.tw

1. INTRODUCTION

Electrocardiogram (ECG) reflects the electrical activity of the heart and contains vast diagnostic information that can guide clinical decision making [1]. ECG is one of the most indispensable tools in medical diagnosis. The ECG and photoplethysmogram (PPG) waveform show similarity in phases when the signals are derived into second derivative signals [2]. ECG and PPG have recently attracted increasing attention due to their specialty in the medical field, particularly in the extraction of vital signs in relation to cardiopathy [3]. Cardiopathy is affecting over 20% people worldwide. Therefore, a universal method for detecting these diseases from ECG signals is highly desired [4]. Each ECG pulse comprises five points, which are known as P, Q, R, S, and T, as shown in Figure 1. It also consists of PR and ST segments, PR and QT intervals, and QRS complex. Automatic analysis of ECG with the use of a computer algorithm is a fundamental task in cardiac monitoring, particularly in long-term monitoring, in which a large amount of data is recorded [5-7]. QRS complexes are typically narrow and tall, resulting in large areas over the curve around these locations. A few algorithms have been developed to detect the onset and duration of QRS complexes [8]. Acquired signals are seldom affected by noise and require for advanced filtering techniques [9]. Accurate

R-peak detection is an important step in ECG analysis, and various methods have been proposed in the past [10].

A signal quality index (SQI) algorithm is used to evaluate the overall signal quality of ECG [11]. For this purpose, the R-peak detector must be capable of robustly detecting the correct R-peak even in a high-noise ECG. From the perspective of signal processing, the R-peak is an important aspect of heart rate variability (HRV) measurement, respiration rate (RR) extraction, and signal quality indexing (SQI) evaluation. Optimal SQI can potentially be used to improve the diagnosis and monitoring of abnormalities, such as hypertension [12].

Multiparameter Intelligent Monitoring in Intensive Care II (MIMIC-II) is a public-access ICU database that stores information regarding numerous patients admitted to intensive care units (ICUs) in the Beth Israel Deaconness Medical Centre in Boston, MA, United States [13, 14]. The original database was split into two parts: the MIMIC-II clinical database and the MIMIC-II waveform database. The clinical database contains the data on the patient charts (heart rate and blood pressure (bp) every hour, blood measurements, etc.). The waveform database contains the raw signals monitored from the patient, such as ECG, photoplethysmogram, and RR.

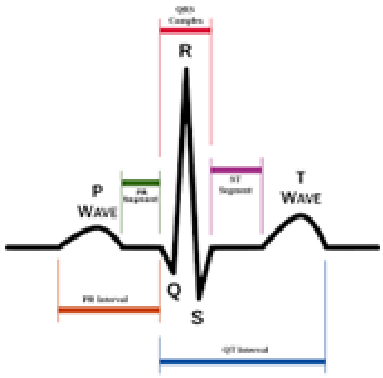

Figure 1. ECG waveform comprises of five points, two segments, two intervals and a complex

Capnobase was an initiative of Dr. Walter Karlen and Dr. Mark Ansermino from the Electrical Computer Engineering in Medicine, Univ. of British Columbia, Canada in 2009. On 11 February 2015, a new revision of this database added the demographic information, such weight, age, and ventilation mode [15]. For Fantasia, 20 young (21–34 years old) and 20 elderly (68–85 years old) rigorously screened healthy subjects underwent 120 min of continuous supine resting and continuous ECG, and respiration signals were collected. In half of each group, the recordings also included an uncalibrated continuous non-invasive blood pressure signal. Each subgroup of subjects included equal numbers of men and women. All subjects remained in a resting state in sinus rhythm while watching the movie Fantasia (Disney, 1940) to help maintain wakefulness. The continuous ECG, respiration, and blood pressure signals were digitized at 250 Hz. Each heartbeat was annotated using an automated arrhythmia detection algorithm, and each beat annotation was verified by visual inspection.

The interquartile range (IQR) was used in this study to measure the variability by dividing a dataset into quartiles, which are denoted by Q1, Q2, and Q3. In this study, a method of examining the signal quality of ECG datasets was investigated. Five R-peak detection algorithms for 12 lead ECG recordings were evaluated using MIMIC-II ECG and MIT-BIH arrhythmia database. Then, two peak detectors that exhibited the best result for R-peak detection were selected to run SQI for the three patient datasets, namely MIMIC-II, Capnobase, and Fantasia.

2. RESEARCH METHOD

Prior to the SQI process, five R-peak detectors were evaluated by applying them to actual ECG datasets. The R-peak detectors used in this study were; Behar's jQRS [16], Qinghua Zhang's rpeak [17], Pan and Tomkins QRSDetector, MATLAB's findpeak, and Clifford's [16] rpeakdetect. First, the detected R-peak on top of ECG was plotted to ensure that it makes sense across the whole record. The two best detectors were used to gauge its accuracy by searching for the R-peak of the MIT-BIH arrhythmia dataset prior to its use in the SQI process. The ECG data used for this evaluation were MIMIC-II and Capnobase. The recorded time for each ECG was 8 min. A 10 s window SQI that move every 1 s after each SQI was applied. Thus, 470 windows were evaluated. Then, the two sets of annotations from the two QRS detectors were transmitted as an input to another algorithm. This algorithm provided an output or score SQI in the range of 0–1. In this study, an SQI of more than 0.9 was categorized as good signals. The median IQR for 30 s non-overlapped window SQI were then obtained to evaluate the qualitative validation of the signal. The median SQI and the IQR were measured for MIMIC-II, Capnobase, and Fantasia as the final output to examine the ECG quality. The methodology is simplified in Figure 2.

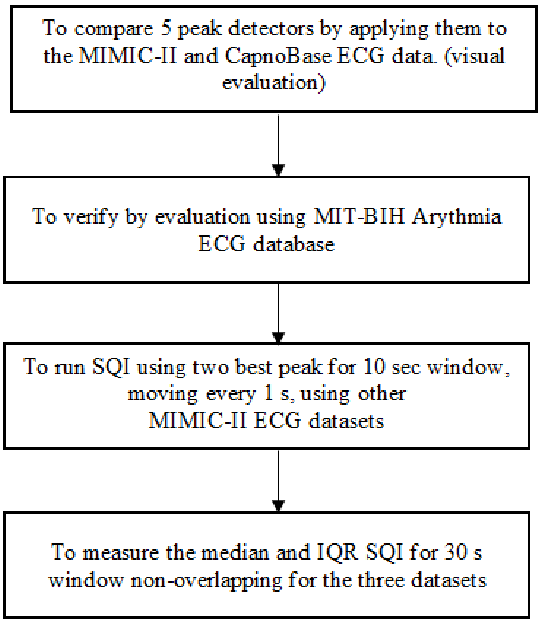

Figure 2. Methodology

3. RESULTS AND ANALYSIS

3.1. Peak detectors evaluation

The five peak detectors were evaluated using an ECG signal at lead II of a female patient aged 50 at the ICU, recorded for the MIMIC-II project, as shown in Figure 3. An 8-min ECG was analyzed from the patient, and the graphs display from 250 s to 300 s of the total ECG. The top graph shows the peak detectors developed by Behar, followed by the original MATLAB peak detector, Zhang's, Pan & Tompkins's, and Clifford's ECG peak detector results on the bottom graph. The analysis suggests that Behar and Clifford's peak detectors are the most reliable and robust. To confirm this finding, we run another evaluation using Capnobase as shown in Figure 4, which confirmed that the R-peaks detected by Behar and Clifford were consistent and robust.

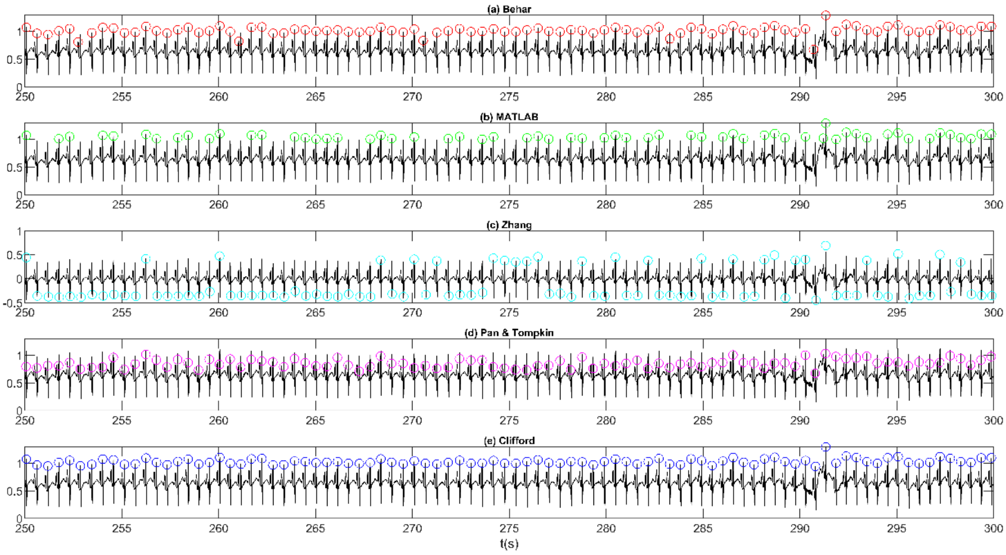

Figure 3. ECG R-peak detector evaluation using MIMIC-II dataset

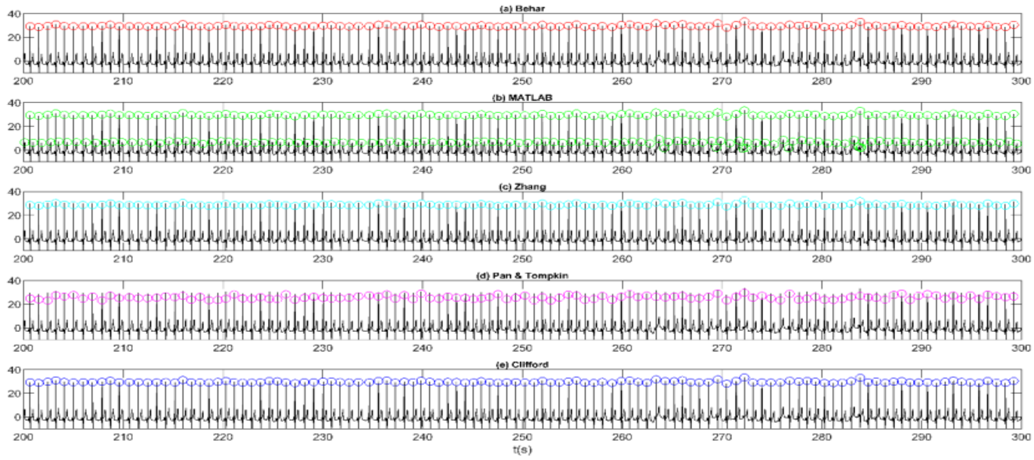

Figure 4. ECG R-peak detector evaluation using Capnobase dataset

3.2. Presicion using MIT-BIH Arrythmia database

Behar and Clifford's peak detectors were used to detect the R-peak of the MIT-BIH arrhythmia database as shown in Figure 5. The result showed that when 60 s ECG data were used, the accuracy was 98% for Behar's and 97% for Clifford's. This data was based on raw and preprocessed ECG signal that include artifacts and noise.

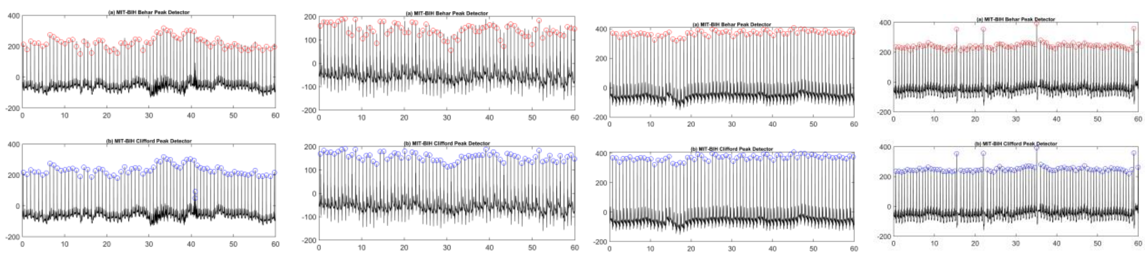

Figure 5. ECG R-peak detector evaluation using MIT-BIH Arrythmia dataset

3.3. SQI from the input of the two selected peak detectors

Figure 6 shows that 655 peaks were detected for both peak detectors during the 6 min of recording. Behar and Clifford's peak detectors were used to confirm the signal quality. Figure 7 shows a very good quality signal detected during 70 s of the recorded data. To confirm this finding, SQI was also performed on a file of the Fantasia dataset, which is shown in Figure 8 and Figure 9 shows a low quality of signals detected during the first 50 s of the recording.

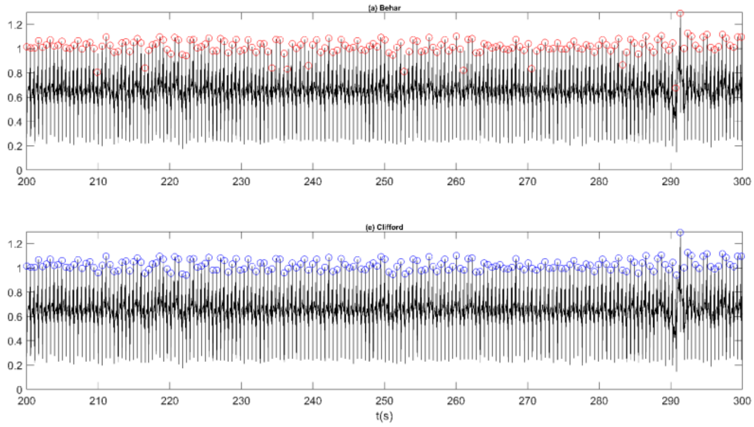

Figure 6. Behar and Clifford peak detectors used for signal quality indexing using MIMIC-II ECG data

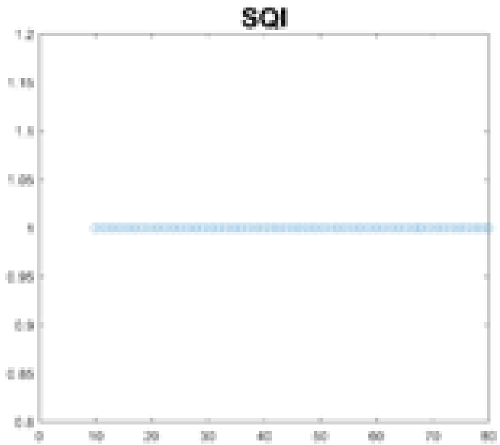

Figure 7. SQI of one MIMIC-II data

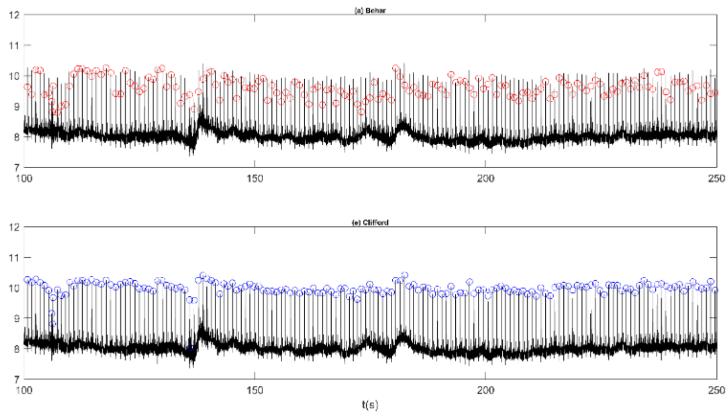

Figure 8. Behar and Clifford peak detectors used for SQI using Fantasia data

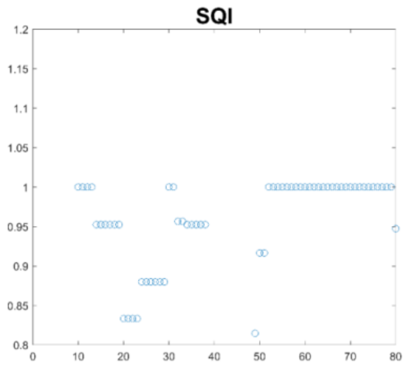

Figure 9. SQI of one Fantasia data

3.4. Median SQI and IQR measurement

Using this technique, all ECG data in three different datasets were evaluated on their SQI. The datasets were Fantasia, MIMIC-II, and Capnobase. The median SQI and IQR of each data were recorded.

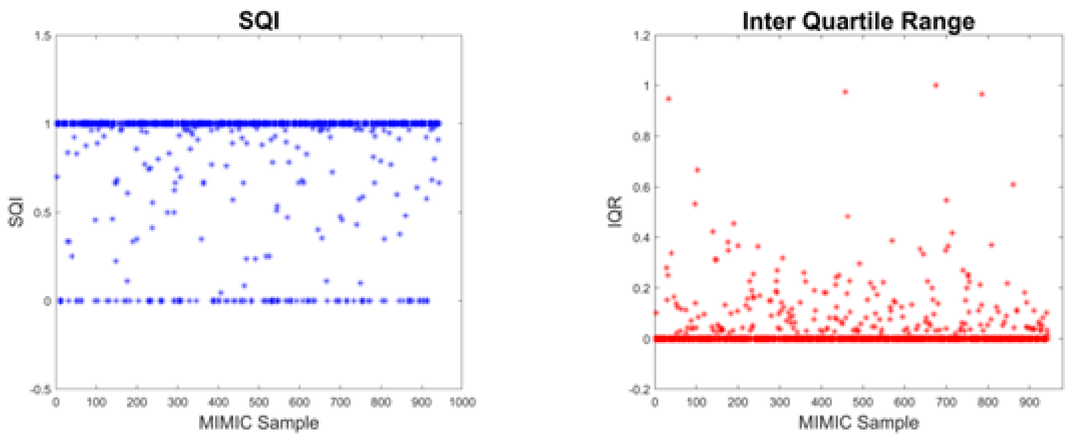

Figure 10. SQI and IQR of all MIMIC-II patients ECG data

As shown in Figure 10, by setting the good data as having 0.9 SQI and above, we categorized 76% out of 944 patient ECG as good. The data also included 93 data that had SQI 0. For the IQR, MIMIC-II had 932 out of 944 that had less than 0.4 IQR. In. Figure 11, Capnobase ECG quality was good, as 40 out of 43 patient data had SQI of 1. This finding was also verified by the IQR graph, in which only one dataset had 0.6 IQR (the largest). Figure 12 shows the result for the Fantasia dataset. The data were relatively good, as 35 out of 38 showed a SQI of 0.9 and above. This result was also verified using IQR, where only six patient ECG data had IQR of more than 0. The results verify that the use of SQI confers biomedical engineers in using the ECG data.

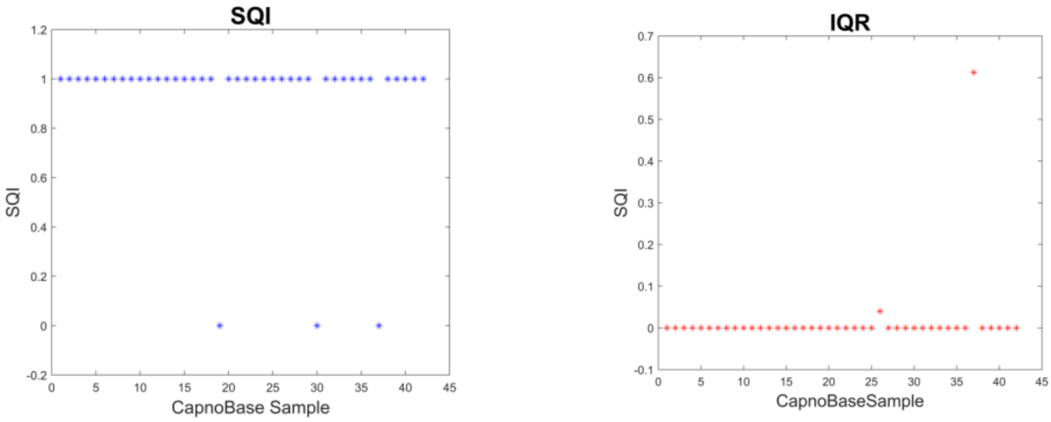

Figure 11. SQI and IQR of all Capnobase patients ECG data

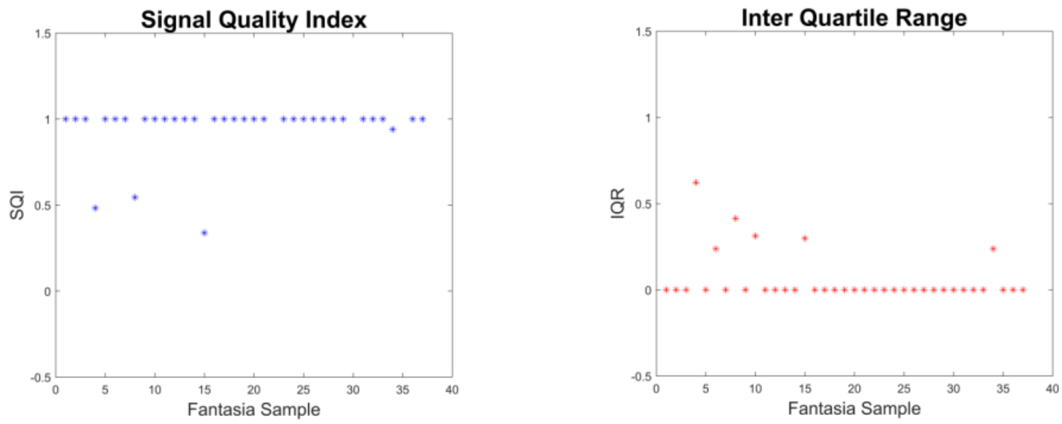

Figure 12. SQI and IQR of all Fantasia patients ECG data

4. CONCLUSION

We evaluated five R-peak detectors in this study. The peak detector algorithms constructed by Behar and Clifford achieved the two best results. These detectors were then used to run SQI. On the basis of the SQI results of each data, the median SQI and the IQR were measured. The proposed method could be useful to evaluate ECG signals prior to being sources for RR and HRV extraction.

REFERENCES

[1] Gothwal H, Kedawat S, Kumar R. Cardiac arrhythmias detection in an ECG beat signal using fast fourier transform and artificial neural network. *J. Biomedical Science and Engineering.* 2011; 4(4): 289-296.

[2] Malek SNH, Chellappan K, Jaafar R. Short Review of Electrocardiogram (ECG) Technique Versus Optical Techniques for Monitoring Vascular Health. *International Conference for Innovation in Biomedical Engineering and Life Sciences. Singapore*. 2015; 56: 222-225.

[3] Sameen AZ, Jaafar R, Zahedi E, Being G. A novel waveform mirroring technique for systolic blood pressure estimation from anacrotic photoplethysmogram. *Journal of Engineering Science and Technology.* 2018; 13(10): 3252-3262.

[4] C. Wei, C. Zhang and M. Wu, "A study on the universal method of EEG and ECG prediction," *2017 10th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI)*, Shanghai, 2017, pp. 1-5.

[5] Maršánová L, Ronzhina M, Smíšek R, Vítek M, Němcová A, Smital L, Nováková M. ECG features and methods for automatic classification of ventricular premature and ischemic heartbeats: A comprehensive experimental study. *Scientific Reports.* 2017; 7(1):11239.

[6] Nayan NA, Risman NS, Jaafar R. A portable respiratory rate estimation system with a passive single-lead electrocardiogram acquisition module. *Technology and Health Care.* 2016; 24(4): 591-597.

[7] Nayan NA, Risman NS, Jaafar R. Breathing rate estimation from a single-lead electrocardiogram acquisition system. *International Journal of Applied Engineering Research*. 2015; 10(17): 38154-38158.

[8] Zong W, Moody GB, Jiang D. A robust open-source algorithm to detect onset and duration of QRS complexes. *Computers in Cardiology.* 2003: 737-740.
[9] Gacek A, Pedrycz W. ECG signal processing, classification and intertretation: A comprehensive framework of computational intelligence. New York: Springer Science & Business Media. 2011.
[10] Liao Y, Na RX, Rayside D. *Accurate ECG R-peak detection for telemedicine.* Humanitarian Technology Conference-(IHTC) IEEE. Canada. 2014: 1-5.
[11] Nayan NA, Jaafar R, Risman NS. Development of respiratory rate estimation technique using electrocardiogram and photoplethysmogram for continuous health monitoring. *Bulletin of Electrical Engineering and Informatics (BEEI).* 2018; 7(3): 487-494.
[12] Elgendi M. Optimal signal quality index for photoplethysmogram signals. *Bioengineering.* 2016; 3(4): 21.
[13] Goldberger AL, Amaral LA, Glass L, Hausdorff JM, Ivanov PC, Mark RG, Mietus JE, Moody GB, Peng CK, Stanley HE. PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals. *Circulation.* 2000; 101(23): 215-220.
[14] Silva I, Moody GB. An open-source toolbox for analysing and processing physionet database in matlab and octave. *Journal of open research software.* 2014; 2(1):27.
[15] Karlen W, Turner M, Cooke E, Dumont G, Ansermino JM. *CapnoBase: Signal database and tools to collect, share and annotate respiratory signals.* Annual Meeting of the Society for Technology in Anesthesia (STA). West Palm Beach. 2010: 25.
[16] J. Behar, J. Oster, Q. Li and G. D. Clifford, "ECG Signal Quality During Arrhythmia and Its Application to False Alarm Reduction," in *IEEE Transactions on Biomedical Engineering*, vol. 60, no. 6, pp. 1660-1666, June 2013.
[17] Zhang Q, Manriquez AI, Medigue C, Papelier Y, Sorine M. An algorithm for robust and eficiente location of T-wave ends in electrocardiograms. *IEEE Transactions on Biomedical Engineering.* 2006; 53(12): 2544-2552.