A Novel Method for Extracting High-Quality RR Intervals from Noisy Single-Lead ECG Signals

Shan Xue[1], Leirong Tian[1], Zhilin Gao[1], and Xingran Cui[1,2](✉)

[1] Key Laboratory of Child Development and Learning Science, Ministry of Education, School of Biological Science and Medical Engineering, Southeast University, Nanjing, China
cuixr@seu.edu.cn

[2] Institute of Biomedical Devices (Suzhou), Southeast University, Suzhou, China

Abstract. In previous studies, plenty of high-accuracy R-peak detection methods were performed on electrocardiogram (ECG) signal analysis. However, these excellent results were usually obtained from some standard and common databases. When applying these detectors on ECG signals collected in daily life and ordinary experiments, or acquired from wearable single-lead ECG devices, the R peak detection accuracies were usually unsatisfying. Due to the influence of data-acquiring environment and devices, the collected ECG signals were often noisy. Each R-peak detection method has its own advantages and may be superior in a certain kind of ECGs. In this study, we proposed a method combining seven R-peak detection methods to get high-quality RR Intervals (RRIs) from noisy ECGs. This new method included two steps, 1) obtain preliminary R-peak annotations through combining seven R-peak detection methods, and 2) calculate the quality score of each R-peak annotations detected in 1) according to the ECG waveform features including kurtosis, skewness and the frequency band power ratio, then exclude the wrong annotations based on the quality scores. The proposed method was evaluated on two databases: MIT-BIH Arrhythmia database and the CPSC2019 training set. The R peak detection average accuracies on these two databases were 98.89% and 55.47% respectively. The results showed that the method proposed in this paper performed better than the seven common R-peak detection methods, especially in noisy ECG signals.

Keywords: Electrocardiogram (ECG) · R-peak detection · High-quality RR intervals · Noisy ECG · Combining methods

1 Introduction

Electrocardiogram (ECG) is one of the most important physiological signals. One of the commonly used methods for analyzing ECG is the Heart Rate Variability (HRV) analysis. HRV refers to the dynamical changes of the difference of heart rate cycle. HRV analysis can reflect the activity, balance and related pathological state of autonomic nervous system. In recent years, there are many studies on HRV [1, 2].

Y. Chen et al. (Eds.): BICT 2020, LNICST 329, pp. 68–79, 2020.
https://doi.org/10.1007/978-3-030-57115-3_6

HRV is obtained by RR intervals. In order to get accurate HRV, it is necessary to detect R peaks accurately. In previous studies, many R-peak detection methods have been proposed. Pan et al. proposed a method for detecting R peaks achieving 99.56% in MIT-BIH Arrhythmia database [3]. It is a benchmark in the R peak detection field [4]. A knowledge-based and robust QRS detection method was proposed by Mohamed Elgendi [5], and it can be easily implemented in a digital filter design. Valtino et al. presented a multirate processing algorithm incorporating FB's for ECG beat detection [6, 7]. MTEO algorithm and the statistical analysis approach were adopted in [8] for QRS complexes detection. Dohare et al. proposed a method for QRS detection that applying sixth power of signal to enhance the peaks of ECG [9]. Matteo et al. showed a QRS detection method especially designed for noisy applications followed by a parameters space reduction operated by the KL transform modified on a "user-fit" basis [10].

However, the quality of ECG collected in real daily life is often worse than that collected in clinical or laboratory conditions. During noisy ECG analysis, the accuracy of R peak detection method is difficult to reach the ideal values in the published papers. Therefore, in order to achieve a higher accuracy of R peak detection, this paper proposed a new method to synthesize the advantages of multiple annotators. Then, in order to get higher quality RRIs, we also extracted the ECG waveform features to find out the wrong annotations and miss-detections.

2 Data and Methods

2.1 Data

The proposed method was evaluated using two public databases. The first database is the MIT-BIH Arrhythmia Database (https://physionet.org/content/mitdb/1.0.0/) [11]. This database contained 48 records collected in the Beth Israel Hospital Arrhythmia Laboratory. Each of the 48 records is over 30 min and sampled at 360 Hz. Most of these records were obtained from inpatients, including different shapes of QRS complexes and many kinds of noises. This database contains approximately 109,000 beat labels, which were labeled and checked by experts. Because this database has enough data samples and accurate labels, it is often used as a standard database to evaluate R-peak detection methods.

The second database was downloaded from CPSC 2019: Challenging QRS Detection and Heart Rate Estimation from Single-Lead ECG Recordings (http://2019.icbeb.org/Challenge.html) [12]. The training data consists of 2,000 single-lead ECG recordings collected from patients with cardiovascular disease, each of the recordings last for 10 s. The sample rate is 500 Hz. This database contains noisy ECG episodes and signals with different arrhythmia patterns. Therefore, the CPSC 2019 database was used to evaluate the performance of the proposed method on noisy ECG data.

2.2 Overview of the Proposed Method

Each R-peak detection method has its own advantages, disadvantages and specific scope of application. However, ECG signals were affected by various kinds of noises. The performance of a single method on an ECG database might not be satisfactory. Therefore, the aim of this paper is to obtain more accurate R-peak detection results by combining seven common R-peak detection results, which is making use of detection methods' advantages to obtain high-quality RRI data from noisy ECG signals.

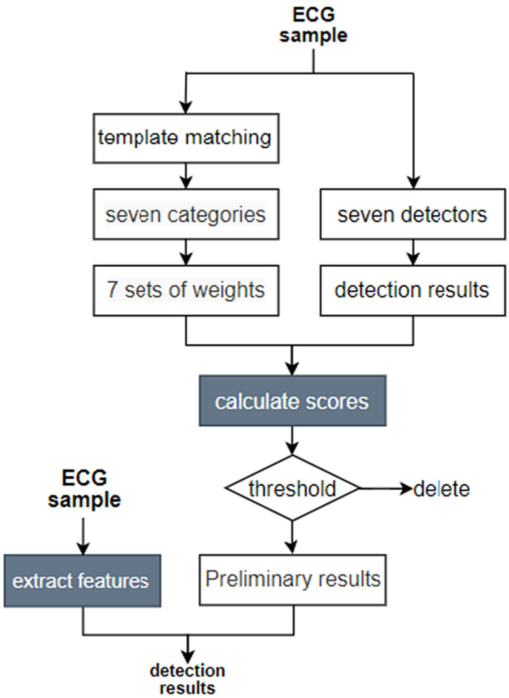

Fig. 1. Block diagram of proposed method

A block diagram of the proposed method was given in Fig. 1, consisting of five steps:

1) Detect R peaks of the ECG samples with seven detection methods. Seven sets of initial detection results were obtained.
2) Categorize the ECG sample by template matching and select weights by the category. There were seven categories in total that corresponding to seven sets of weights. A set of weights included seven weights corresponding to seven detection methods. For a specific kind of ECG signal, the method with better performance was assigned more weight. The worse performance of a detection method, the smaller weight was assigned, and the weight was even set to be negative, accordingly. A weight represented the average performance of a detection method on a specific kind of ECG signals. The weights were obtained through a method mentioned below. Some fine adjustments of weights were made in the algorithm.
3) Calculate the "scores" of each R-peak annotation in seven sets of detection results in 1) based on the weights in 2). A high score indicated that the annotation might be correct. The method of calculating the scores was described in detail in Sect. 2.4.

Then, the scores were filtered by a threshold, and the annotations with low scores were deleted. After this step, the preliminary R-peak detection results were obtained.
4) According to the preliminary results of 3), cut the ECG sample into multiple sections.
5) Extract waveform features from each section in 4). Based on these features, the quality scores of each section were calculated. The purpose of calculating the quality scores was to further eliminate wrong R-peak annotations. The method of calculating the quality scores was described in detail in Sect. 2.4.

The final detection results were obtained through the above steps.

The proposed method was developed using the MATLAB software environment. The programs run in MATLAB R2018a on Intel(R) Core(TM) i7-6700 k CPU @ 4.00 GHz processor, Lenovo, Beijing, China.

2.3 The Seven R-Peak Detection Methods

This method was based on seven popular R-peak detection methods. These methods have been widely used in many studies, especially in single-lead ECG analysis.

Optimized knowledge-based algorithm (OKB)

OKB is a knowledge-based and robust algorithm for detecting QRS complexes in ECG signals. It used the prior knowledge of the ECG features to improve the detection accuracy [7].

Pan-Tompkins

Pan-Tompkins is also an algorithm widely used in QRS complexes detection. It detected QRS complexes by calculating the slope, amplitude, and width. In addition, it proposed that using a digital bandpass filter to reduce the false detections [3].

Jqrs

Jqrs method is an adaptation of the Pan-Tompkins algorithm. It used search back technology and 200 ms refractory blanking to improve the detection results [3, 13].

Nqrs

Nqrs is based on the filter bank (FB) which decomposes the ECG into sub-bands with uniform frequency bandwidths. This is a real-time algorithm since its beat detection is minimal [6, 7].

MTEO

MTEO method can detect R, Q, S, T and P wave. It based on a modified version of the multiresolution Teager energy operator (MTEO). The unsupervised clustering of action potentials is achieved by applying a combination of label and template matching techniques [8].

Phase Space QRS

This method employs the area under the non-linear phase space reconstruction of the ECG recording in order to identify the QRS complexes [14].

R detection by wavelet

In this method, in order to remove noise and baseline drift, the signal is first passed through a filter, and then the wavelet is used to detect the R wave [15].

2.4 Details of the Novel R Peak Detection Method

The proposed new method is presented as the following four steps:

Obtain the initial R-peak detection results and select weights

The input ECG signal was first analyzed by seven R-peak detection methods mentioned above to obtain seven sets of detection results. Then through template matching described below, ECG signals were classified into seven categories.

The length of each template was set to 10 s. The cross-correlation coefficients between input ECG signals and templates were calculated respectively, and the input signals were classified into the category with the largest cross-correlation coefficient.

Then, select the weights corresponding to the category that input signal belonged.

Calculate the "score" for each R-peak annotation

The seven sets of R detection results were merged into a vector, and the scores of the initial R-peak annotations obtained in step (1) were calculated one by one.

First, to calculate the score of R-peak annotation j, the score was calculated based on the weight of the detection method that marked annotation j.

Then, find out whether there were other R-peak annotations marked by other six methods in the left or right range (i.e., within 50 ms around the center of annotation j) of annotation j. There would be some deviation between the results of several detection methods, so when many close annotations were marked within this range, it is likely that there was a correct R-peak annotation in this range. When other annotations within the range were more concentrated on annotation j, it is more likely that j was a correct R-peak annotation. Therefore, the score of annotation j was inversely related to the distances from it to other annotations within the range. The score of R-peak annotation j was defined in expression (1).

$$score_j = weight_j \sum_i \frac{fs \cdot weight_i}{100 \cdot |annotation_j - annotation_i|} \tag{1}$$

where $weight_j$ is the weight corresponding to the detection method that marked $annotation_j$, $annotation_i$ is the annotation within the range, fs is the sampling rate of ECG signals.

Finally, a threshold was used to filter out the R-peak annotations with high scores.

$$threshold = 0.8 \cdot mean \tag{2}$$

where mean is the average score of the ten R-peak annotations before the annotation being calculated. The parameter was adjusted according to the average accuracy scores.

Method for getting weights

The procedure was presented in Fig. 2. Each of the seven categories has ten ECG recordings, and 70 recordings composed the data set.

1) For each category of ECG signal, first initialize seven weights corresponding to seven methods.
2) Randomly select one of the seven weights for the operation of $+/-0.5$. The weights before operation were called "previous weights"; the weights after operation were called "current weights".

3) Obtain the detection results based on "current weights", and calculate the average accuracy of R-peak detection of ten recordings. If the accuracy did not change, "previous weight" were updated to "current weights", then continue to perform 2); if the accuracy increased, the "previous weights" were updated to "current weights", and the same operation was performed on the weights in the next iteration instead of randomly selecting operation; if the accuracy decreased, the previous weights were not updated, then continue to perform 1).
4) This process was repeated 100 times.

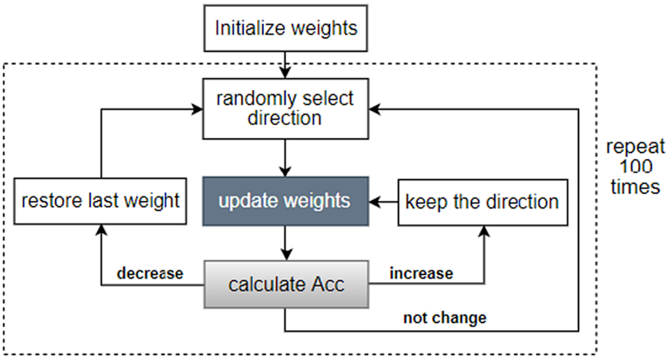

Fig. 2. Block diagram of getting weights

Extracting ECG waveform features

Although the accuracy of R-peak detection was improved, the accuracy might be still unsatisfying, so it is necessary to find the wrong detections or missed detections to get high-quality R-R intervals (RRIs).

In the previous steps, we synthesized seven commonly-used R-peak detection methods and got the preliminary R-peak results. However, in some cases, all the seven methods made the same wrong detection or some R peaks were miss-detected.

Therefore, to improve the detection accuracy, we extracted some features of each heartbeat's waveform to calculate the quality score of a heartbeat. The heartbeats were divided by the detection results in the previous steps. The features extracted were kurtosis, skewness and frequency band power ratio.

In order to eliminate the influence of outliers, the features' values were limited. If the kurtosis value was greater than 8, the kurtosis value was 8. The skewness value was generally small, so special treatment is not carried out for the skewness. The frequency band power ratio was the power ratio of 5–15 Hz to the power ratio of 5–50 Hz, and then calculated the square of the power ratio, which was to expand the difference. If the square value was less than 0.5, then took 0.5.

The poor quality signals include some abnormal signals caused by electrode falling off and other reasons. By calculating the variance of signals, we can judge and punish the scores.

After the eigenvalues were obtained, the quality score of a heartbeat interval was calculated as follows:

$$quality\ score = \begin{cases} \frac{kur \cdot 4.9}{p} - 5 \cdot e^{-ske \cdot 0.8} + \frac{40}{var}, ifske < -0.5 \\ \frac{kur \cdot 4.9}{p} - 5 \cdot e^{1+ske} + \frac{40}{var}, if -0.5 \le ske < 0 \\ \frac{kur \cdot 4.9}{p} + ske \cdot 5 + \frac{40}{var}, others \end{cases} \quad (3)$$

where kur is kurtosis, p is frequency band power ratio, ske is skewness, var is the variance of signals. The parameters were obtained by fitting the eigenvalues to the quality scores.

The R peak quality scores in ECG recordings were shown in Fig. 3.

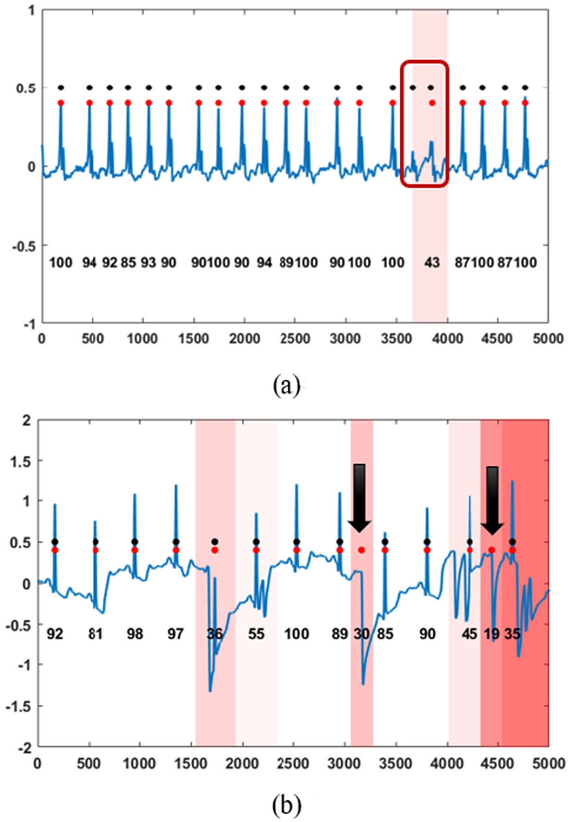

Fig. 3. Representative result of the R peak quality score. (a) Scores of recording 52 from CPSC2019 training set. Red dots represent the R detection results of the proposed method, black dots represent the reference label of R-peak annotations. The black number below the R peaks were the corresponding signal quality scores. The parts in red box was miss-detected. Score of this part was lower than the others. (b)Scores of recording 55 from CPSC2019 training set. In this record, there are two wrong marks (indicated by the black arrow). The scores of these two places are lower than others. Additionally, in CPSC2019, the evaluation method refines the search edge for answers within [0.5, 9.5]s, eliminating the misjudgment of scores while the predictive R peak location is in [0.5 (0.5 + 0.075)]s or [(9.5 − 0.075), 9.5]s. (Color figure online)

3 Results

3.1 Results on the MIT-BIH Arrhythmia Database

The proposed method was firstly evaluated on the MIT-BIH Arrhythmia Database. Results are shown in Table 1.

Sensitivity (Se), prediction (+P) and average accuracy were used to evaluate the performance of the method on the database. The accuracy of the method proposed in this paper is higher than that of using each of the seven detection methods separately. It should be noted that the accuracies of some methods are not as high as that presented in the published papers [3, 5, 8, 14, 15], which may be due to different parameter settings or different data preprocessing.

Table 1. Performance of the proposed R peak detection method on the MIT-BIH arrhythmia database.

Methods	Se (%)	+P (%)	Average accuracy
OKB	97.96%	97.75%	96.15%
Jqrs	99.23%	98.91%	98.19%
Nqrs	98.76%	99.10%	97.89%
MTEO_qrs	99.07%	99.06%	98.26%
Pan-Tompkin	99.04%	98.94%	98.04%
Phase Space QRS	90.44%	87.89%	83.25%
Wavelet	83.29%	89.47%	80.85%
Proposed method	**99.54%**	**99.33%**	**98.89%**

3.2 Results on the CPSC 2019 Challenging Data

To check the effectiveness of the proposed method in noisy ECG signals, the evaluation was done using the CPSC 2019 challenging data. The performances of various methods are shown in Table 2.

The true positive (TP) is the number of correct annotations detected, the false positive (FP) is the number of wrong R-peak annotations, the false negative (FN) is the number of missed R peak. In order to compare the results with competitors, we used the same rules for calculating accuracy as used in CPSC 2019 website [5]. The rules are:

1) Complete matching scores one point;
2) A false positive (FP) detection scores 0.7 points;
3) A false negative (FN) detection scores 0.3 points, since from a clinical perspective, missed diagnosis is more serious than misdiagnosis, thus penalize FN detection here;
4) Other situations score 0;
5) The final accuracy is the average score of all recordings.

Comparing to performances of the submitted competition algorithms (in MATLAB group) reported in the challenge, the proposed method achieved the accuracy of 55.47%, apparently exceeding the second place algorithm (accuracy 40.73%), and very close to the best algorithm performance (accuracy 57.32%). On the other hand, when analyzing this noisy ECG database, all the seven commonly-used methods performed worse than the proposed method (see Table 2). Therefore, this novel method can improve the R-peak detection accuracy.

Table 2. Performances of various methods on CPSC 2019 training data.

Methods	TP	FN	FP	Accuracy
OKB	24976	2063	2895	48.71%
Jqrs	22934	4105	2323	45.20%
Nqrs	18910	8129	1752	6.43%
MTEO_qrs	22304	4711	2157	38.68%
Pan-Tompkin	22028	5011	3491	34.94%
Phase Space QRS	21155	5884	7134	28.56%
Wavelet	24160	2879	4445	41.82%
Proposed method	**24813**	**2226**	**1819**	**55.47%**

The rules of calculating accuracy used in CPSC 2019 were much more rigorous than the rules used in MIT-BIH Arrhythmia database, and the penalty for wrong detection was greater. This was also one of the reasons for that the accuracies of CSPC 2019 data are much lower than on MIT-BIH Arrhythmia data.

Figure 4 shows the representative results of record 00243 in the CPSC2019, comparing the results of this method with those of seven commonly-used methods. In FP, FN and accuracy, the results of the method proposed were better than those of the seven methods respectively.

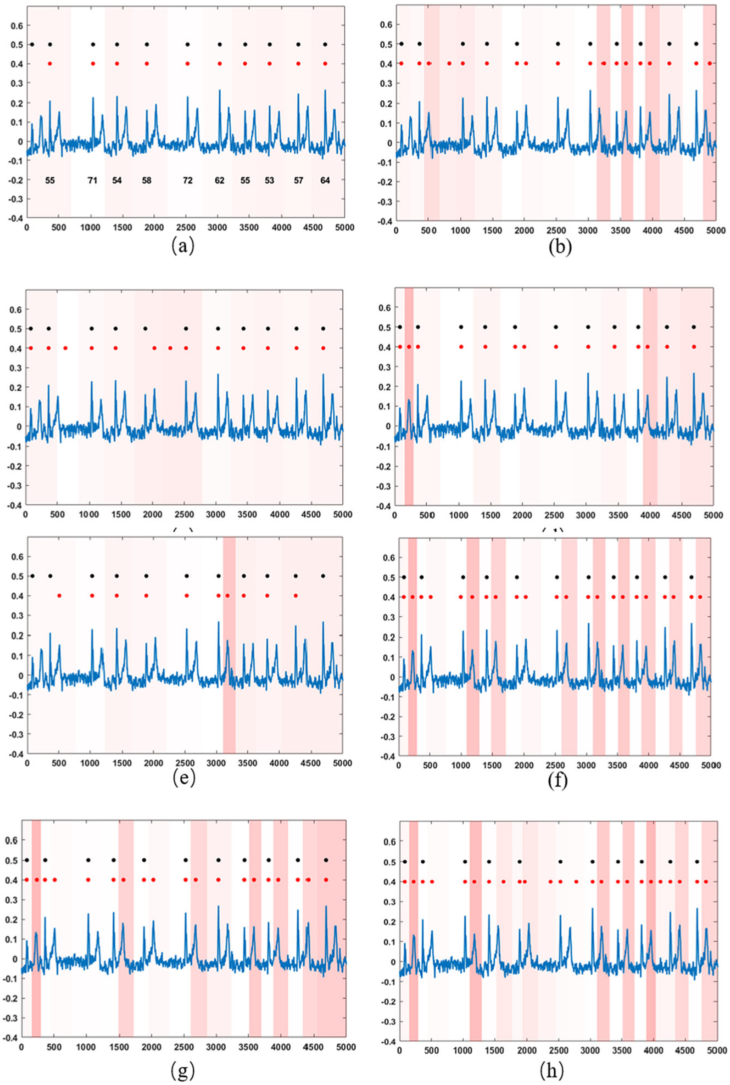

Fig. 4. Representative results of record 00243 among different R-peak detection methods: (a) the proposed method: FP = 0, FN = 0, accuracy score = 1; (b) OKB method: FP = 3, FN = 0, accuracy score = 0; (c) Jqrs method: FP = 3, FN = 1, accuracy score = 0; (d) Pan-Tompkin method: FP = 3, FN = 0, accuracy score = 0; (f) Nqrs method: FP = 2, FN = 2, accuracy score = 0; (g) Phase Space QRS method: FP = 8, FN = 0, accuracy score = 0; (h) wavelet method: FP = 13, FN = 0, accuracy score = 0. Additionally, the accuracy scores here are based on the rules in the in CPSC 2019 website [5], and when the number of FN and FP exceeds 2, the accuracy score is 0.

4 Conclusions

In this study, we proposed a novel method combining seven commonly-used R-peak detection methods for getting high-quality R-R intervals (RRI) from noisy ECG recordings. The detection results were obtained by assigning a weight to each method, and then the results were filtered by signal quality scores based on every heartbeat interval's waveform, which can be used to correct detection errors and obtain high-quality RRI data.

Based on the evaluation results of two public databases, the method proposed in this paper effectively improved the accuracy of R-peak recognition, and can take the advantages of different methods in different types of ECGs, and more accurate RRIs can be extracted from noisy ECG signals. Besides, the following-up R peak quality score evaluation enhanced the quality of RRIs.

As for several detection methods used in this study, we may add more available methods and do more comparison to find the best combination in further study.

Acknowledgment. This research was funded by National Natural Science Foundation of China, grant number 61807007; National Key Research and Development Program of China, grant number 2018YFC2001100; Fundamental Research Funds for the Central Universities of China, grant number 2242019K40042.

References

1. Li, G., Chung, W.Y.: Detection of driver drowsiness using wavelet analysis of heart rate variability and a support vector machine classifier. Sensors **13**(12), 16494–16511 (2013)
2. Atri, R., Mohebbi, M.: Obstructive sleep apnea detection using spectrum and bispectrum analysis of single-lead ECG signal. Physiol. Meas. **36**(9), 1963 (2015)
3. Pan, J., Tompkins, W.J.: A real-time QRS detection algorithm. IEEE Trans. Biomed. Eng. **3**, 230–236 (1985)
4. D'Aloia, M., Longo, A., Rizzi, M.: Noisy ECG signal analysis for automatic peak detection. Information **10**(2), 35 (2019)
5. Elgendi, M.: Fast QRS detection with an optimized knowledge-based method: evaluation on 11 standard ECG databases. PLoS ONE **8**(9), e73557 (2013)
6. Afonso, V.X., et al.: ECG beat detection using filter banks. IEEE Trans. Biomed. Eng. **46**(2), 192–202 (1999)
7. Oppenheim, A.V.: Discrete-Time Signal Processing. Pearson Education, India (1999)
8. Sedghamiz, H., Santonocito, D.: Unsupervised detection and classification of motor unit action potentials in intramuscular electromyography signals. In: 2015 e-Health and Bioengineering Conference (EHB). IEEE (2015)
9. Dohare, A., Kumar, V., Kumar, R.: An efficient new method for the detection of QRS in electrocardiogram. Comput. Electr. Eng. **40**(5), 1717–1730 (2014)
10. Paoletti, M., Marchesi, C.: Discovering dangerous patterns in long-term ambulatory ECG recordings using a fast QRS detection algorithm and explorative data analysis. Comput. Methods. Programs Biomed. **82**(1), 20–30 (2006)
11. Moody, G.B., Mark, R.G.: The impact of the MIT-BIH arrhythmia database. IEEE Eng. Med. Biol. Mag. **20**(3), 45–50 (2001)

12. Gao, H., et al.: An open-access ECG database for algorithm evaluation of QRS detection and heart rate estimation. J. Med. Imag. Health Inform. **9**(9), 1853–1858 (2019)
13. Behar, J., et al.: A comparison of single channel fetal ECG extraction methods. Ann. Biomed. Eng. **42**(6), 1340–1353 (2014)
14. Lee, J.W., et al.: A real time QRS detection using delay-coordinate mapping for the microcontroller implementation. Ann. Biomed. Eng. **30**(9), 1140–1151 (2002)
15. Bsoul, A.A.R., et al.: Detection of P, QRS, and T components of ECG using wavelet transformation. In: 2009 ICME International Conference on Complex Medical Engineering. IEEE (2009)