IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received June 13, 2018, accepted July 22, 2018, date of publication July 26, 2018, date of current version August 20, 2018.

Digital Object Identifier 10.1109/ACCESS.2018.2860056

Dynamic ECG Signal Quality Evaluation Based on the Generalized bSQI Index

FEIFEI LIU[1,2], CHENGYU LIU[1], (Member, IEEE), LINA ZHAO[1,3], XINGE JIANG[3], ZHIMIN ZHANG[3], JIANQING LI[4], (Member, IEEE), SHOUSHUI WEI[3], AND YUAN ZHANG[5], (Senior Member, IEEE)

[1]State Key Laboratory of Bioelectronics, Jiangsu Key Laboratory of Remote Measurement and Control, School of Instrument Science and Engineering, Southeast University, Nanjing 210096, China
[2]Shandong Zhong Yang Software Limited Company, Jinan 250101, China
[3]School of Control Science and Engineering, Shandong University, Jinan 17923, China
[4]School of Biomedical Engineering and Informatics, Nanjing Medical University, Nanjing 211166, China
[5]Shandong Provincial Key Laboratory of Network Based Intelligent Computing, University of Jinan, Jinan 250022, China

Corresponding authors: Chengyu Liu (chengyu@seu.edu.cn) and Shoushui Wei (sswei@sdu.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 61571113, Grant 61572271, and Grant 61572231, in part by the Key Research and Development Programs of Jiangsu Province under Grant BE2017735, in part by the China Postdoctoral Science Foundation under Grant 2017M612280, in part by the Shandong Province Key Research and Development Plan under Grant 2018GSF118133, and in part by the Shandong Provincial Natural Science Foundation in China under Grant ZR2014EEM003.

ABSTRACT With the fast development of wearable electrocardiogram (ECG) monitoring, real-time and dynamic signal quality assessment (SQA) become an imperious demand. Thus, many signal quality indices (SQIs) have been developed in the past several years. bSQI is a typical SQI defined from two common QRS complex detectors ('*ep_limited*' and '*wqrs*'). However, in actual application, bSQI heavily relies on the QRS complex detectors used. Therefore, if using different combination of QRS detectors can improve the performance of SQA needs to be explored. In this paper, we utilized up to ten QRS detectors to re-define bSQI from the combination of any two QRS detectors to test which combination outputs the highest performance. Then, we generalized the two QRS detector-based bSQI to multiple QRS detector-based bSQI (i.e., GbSQI), to systematically test the effects of type and number of QRS detectors on SQA performances. The results showed that for the single GbSQI feature-based classifier, the combination of six QRS detectors reported the highest classification accuracy with a mAcc of 94.03%. For the multiple GbSQI feature-based classifier, the combination of four QRS detectors showed the best classification accuracy with a mAcc of 94.76%. As a conclusion, we recommended using *U3*, *UNSW*, *DOM*, and *OKB* detectors for calculating GbSQI for the wearable ECG monitoring application.

INDEX TERMS Electrocardiogram (ECG), signal quality assessment (SQA), signal quality index (SQI), QRS detection.

I. INTRODUCTION

Recently, the fast-developing wearable and Internet of things (IoT) technologies significantly promote the progress in ambulatory electrocardiogram (ECG) monitoring [1], which is an essential useful tool for the early detection of cardiovascular diseases (CVDs) [2]. However, dynamic ECGs suffer from the problem of poor signal quality due to the bad electrode contact or incorrect electrode positioning due to unsupervised operators in a remote condition [3]. Poor signal quality impedes the reliable manual or automated measurement, hazards the correct diagnosis information [4], increases the risk of false alerts [5], and increases the workload of physicians [6], which may limit the mHealth application for rural populations [4]. Therefore, automated signal quality assessment (SQA) method is important to be established, and can remind the user to re-take recordings when signal quality is low [7], to reject the unavailable ECGs to avoid the network congestion [8], and to provide the reliable signals for CVDs scanning [9].

A variety of SQA methods have been explored, including time-domain, frequency-domain, joint time-frequency, self-correlation, cross-correlation, entropy methods, etc. In 1989, Moody and Mark developed a Karhunen–Loeve transform to estimate ECG noise [10]. In 1996, Allen and Murray proposed quantitative indexes, i.e., the frequency content in six different bandwidths and the time length exceeded a preset threshold [6]. In 2008, Redmond *et al.* [11] developed a signal artifact masking algorithm for automatically marking ECGs with sections of obvious artifact, including three feature masks: a rail contact mask, a high-frequency mask and a low power mask, which was the first report about the ECG quality measures in unsupervised telecare environments. Based on this signal artifact masking algorithm, Redmond *et al.* [12] proposed a SQA method using a Parzen window supervised statistical classifier model in 2012. In 2008, Li *et al.* [13] proposed four signal quality indexes: (1) bSQI: comparison of two beat detectors on a single ECG lead, (2) iSQI: comparison of the same beat detector on different ECG leads, (2) kSQI: evaluation of the kurtosis (randomness) of ECG episode, and (3) sSQI: calculating the spectral distribution of ECG episode within a certain physiological frequency band. Using these indexes, Clifford *et al.* [14] achieved good results in the 2011 PhysioNet/CinC challenge. The 2011 Physionet/CinC Challenge addressed the issue of developing an efficient algorithm being able to run in real-time on a mobile phone, which can provide useful feedback to a layperson in the process of acquiring a diagnostically useful ECG recording [8]. Since then, many SQA methods have been proposed [2], including six wave features (flat baseline, saturation, baseline drift, low amplitude, high amplitude and steep slope) identified by Marco *et al.* [15], and a signal quality matrix used by Xia *et al.* [9]. In addition, recently Zhang *et al.* [16] developed a novel encoding Lempel–Ziv complexity algorithm for quantifying the irregularity of ECGs.

The developed signal quality indexes (SQIs) fall into two categories. The first category of SQIs were directly calculated from ECG waveform characteristics, such as signal amplitude, spectrum, standard deviation, mean square error and signal noise ratio (SNR). The second category of SQIs were calculated by analyzing the feature characteristics extracted from ECGs. bSQI was a typical representative. It was based on the principle that QRS detections from different algorithms should be nearly the same for good quality signals, while they should be different if signal quality was poor. Since different QRS detectors were sensitive and specific to different types of noises [17], the comparison of how accurately multiple QRS detectors isolate each event (such as a beat or a noise artifact) provides an estimation of noise level. bSQI has been widely used in several years, as the essential SQI feature for single/multiple channel ECGs signal quality determination [18] or as feature for support vector machine (SVM)-based ECG SQA [19].

However, it is clear that bSQI heavily relies on the used QRS detectors. bSQI itself uses two common QRS detectors: '*ep-limited*' [20] and '*wqrs*' [21]. If one QRS detector misses one or more beats (due to low QRS amplitudes) or registers extra beats (due to artifact or high amplitude T waves), bSQI will fail to give a good signal quality estimation [13]. Accordingly, how robust will the widely used bSQI be on the poor signal quality ECGs? If using other QRS detectors, or combining multiple QRS detectors, will improve the performance of bSQI? These questions should be clarified. Thus, in this study, the contributions are from two aspects. First, we utilized up to ten QRS detectors to re-calculate bSQI from the combination of any two and obtained the best combination of QRS detectors for defining two QRS detector-based bSQI. Then, we generalized the two QRS detector-based bSQI to multiple QRS detector-based bSQI (i.e., GbSQI), to compare the effects of the type and number of QRS detectors on the SQA performances of GbSQI.

II. METHOD

A. DATABASE AND RE-LABELLING

A total of 1,000 recordings of standard 12-lead ECGs from the 2011 PhysioNet/CinC Challenge were used [8], which were collected by the Sana Project [22] and were provided freely via PhysioNet [23]. ECGs were sampled at 500 Hz with 16-bit resolution and were filtered with full diagnostic bandwidth (0.05 through 100 Hz). Each signal had a length of 10 s. In 1000 ECGs, 773 were labeled as 'acceptable,' 225 were 'unacceptable' and 2 were 'intermediate'. However, the labeling for 'acceptable' or 'unacceptable' was for the whole 12 channels, not for the single channel, making the evaluation of bSQI on single ECG channel impossible. For example, many 'acceptable' ECGs have a channel with total noises or being even a flat line. Thus, we re-labeled each channel of ECGs and obtained a total of 12,000 10-s ECG data segments.

TABLE 1. Five signal quality levels for the 10-s ECG segments.

Level	Score	Description for signal quality scoring
1	1	The recording with no obvious noise, QRS complex and T wave is clear.
2	0.75	Noise or artefact may be existing at the beginning or end, but it does not affect the identification for the basic waves of most ECG signal.
3	0.5	The lager noise may continue for about 2-3 s, and it could influence the signal interpretation in this episode. But about 6-7s signals can be identified.
4	0.25	More serious lager noises exist, such as strong gaussian noise and signal saturation et. al. In these noise episode, it is impossible to identify the QRS complex. But at least 4-5 s continuous identifiable heart beats are visible.
5	0	Strong noises occupy in the more than 5 s episode. It is very hard to identify the heart beat for the most signal.

According to the literatures [7], [8], we used five signal quality levels and scored each ECG segment as Table 1. Typical examples in each signal quality level were shown in Fig.1. Five researchers re-labeled all the ECG recordings. For each 10-s ECG segment, five scores were obtained. Then the 10-s ECG segment was labeled as 'acceptable' if $\bar{S}(\bar{S} = \frac{1}{5}\sum_{i=1}^{5} S_i)$ was higher than a threshold of 0.25. Otherwise, it was labeled as 'unacceptable,' resulting in a total of 9,941 acceptable and a total of 2,059 unacceptable 10-s ECG segments.

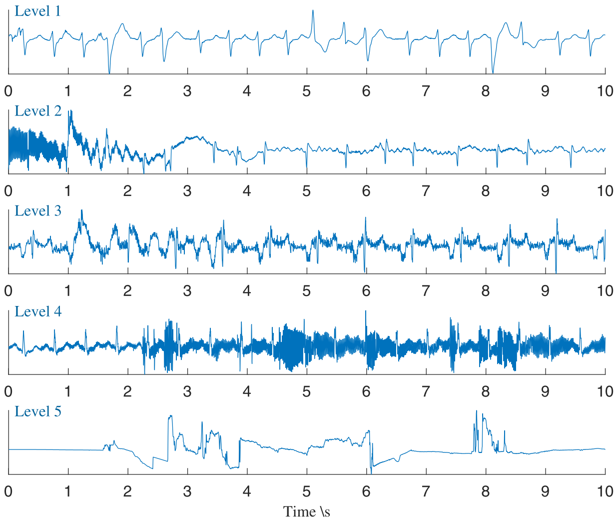


FIGURE 1. Typical 10-s ECG segments from each signal quality level. ECG segment from Level 1 is relatively clean, and those from Levels 2-5 have the decreased signal quality.

B. ORIGINAL bSQI

The original bSQI proposed by Li *et al.* [13] was based on the comparison of two beat detectors on a single lead. Two well documented open-source QRS detectors were used: '*ep_limited*' was based on digital filtering and integration [24] and '*wqrs*' was based on a length transform after filtering [21]. bSQI for a w seconds signal was defined to be the ratio of beats detected synchronously (within an interval, γ) by both detectors to all the detected beats (by either detector) within the window [25]. w is set to be 10 s and γ is set to be 150 ms. bSQI for the kth beat was defined as:

$$bSQI(k) = \frac{N\_matched(k, w)}{N\_ep\_limited(k, w)+N\_wqrs(k, w)-N\_Matched(k, w)} \quad (1)$$

where, $N\_matched(k, w)$ is the beat number agreed upon (within $\gamma = 150$ ms), $N\_ep\_limited(k, w)$ is the beat number detected by '*ep_limited*' and $N\_wqrs(k, w)$ is the beat number detected by '*wqrs*'. Therefore, bSQI ranges between 0 and 1 inclusively. For N beats, there are N windows set to be $w = 10$ s long, centered $\pm$ 5 s around the kth beat. In this study, the signal from the 2011 PhysioNet/CinC Challenge had a length of 10 s merely, so there was only one window for each signal. In this way, only one bSQI value was calculated for each signal.

C. GENERALIZED bSQI (GbSQI)

Thanks to the quick development for studies on QRS detector in the past several decades. A large number of QRS detectors have been developed and have been open-sourced. In this study, we tested up to ten QRS detectors, which detailed in Table 2.

TABLE 2. The QRS detectors with low computational complexity used in this study.

No.	QRS detector
1	*ep_limited* [24]
2	*RS-slope* [26]
3	*Six-power* [27]
4	*Finite state machine (FSM)* [28]
5	*U3 transform (U3)* [29]
6	*Difference operation (DOM)* [30]
7	*'jqrs'* [31]
8	*Optimized knowledge-based (OKB)* [32]
9	*UNSW* [33]
10	*'wqrs'* [21]

There are a large number of QRS detectors have been proposed in the past several years. To analyze all of them would be impractical. Dynamic ECG signal mainly came from wearable or mobile devices. These devices, with the limitations in terms of memory and processor capability, have very high requirements for computational efficiency. So computational efficiency becomes the first criterion for QRS detectors selection. Besides, the high accuracy is an essential basis for the QRS detectors. In fact, the current detectors have high accurate for the ECG signal with high signal quality, but most do not have good performances for the low-quality signals. As is known to all, it is not always to write the right program according to the description of some papers. So, the performability was also a key point for QRS detectors selecting. Therefore, according to these three criteria (algorithm efficiency, detection accuracy and performability), this study selected ten QRS detectors from dozens of papers. Of course, there are many other good detection algorithms with low computation complexity, high detection accuracy and good operability. Because of the limited time and our viewpoints, only these ten QRS detection algorithms were selected in this study.

The generalized bSQI (GbSQI) for the kth beat within a w seconds signal window is defined as (2), as shown at the bottom of this page, where, a total of n QRS detectors were used for calculating GbSQI, and $N\_matched(k, w)$ is the beat number that all algorithms agreed upon (within $\gamma = 150$ ms), $N\_method_n(k, w)$ is the beat number detected by the nth method. GbSQI therefore also ranges between 0 and 1 inclusively. For N beats, there are N windows set to be $w = 10$ s long, centered $\pm$ 5 s around the kth beat. In this study, the signal from the 2011 PhysioNet/CinC Challenge

$$GbSQI(k, w) = \frac{(n-1) \times N\_matched(k, w)}{N\_Method_1(k, w) + N\_method_2(k, w) + \cdots N\_method_n(k, w) - N\_matched(k, w)} \quad (2)$$

had a length of 10 s merely, so there was only one window for each signal. In this way, only one GbSQI value was calculated for each signal.

In this study, seven patterns of GbSQI (bSQI-2, bSQI-3, bSQI-4, bSQI-5, bSQI-6, bSQI-7, bSQI-8) were tested (see Table 3). Take bSQI-2 as an example, any two QRS detectors were chosen randomly from the ten algorithms, and 45 combinations were generated.

TABLE 3. Seven patterns of GBSQI.

Pattern	# used QRS detectors	# combinations
bSQI-2	2	$C_{10}^{2} = 45$
bSQI-3	3	$C_{10}^{3} = 120$
bSQI-4	4	$C_{10}^{4} = 210$
bSQI-5	5	$C_{10}^{5} = 252$
bSQI-6	6	$C_{10}^{6} = 210$
bSQI-7	7	$C_{10}^{7} = 120$
bSQI-8	8	$C_{10}^{8} = 45$

D. EVALUATION METHODS

Sensitivity (Se), specificity (Sp) and modified accuracy (mAcc) were defined in (3-5), based on the number of signals correctly classified as 'unacceptable' (TP), the number of signals falsely classified as 'unacceptable' (FP), the number of signals correctly classified as 'acceptable' (TN) and the number of signals falsely classified as 'acceptable' (FN) [34].

$$\mathrm{Se} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}} \times 100\% \tag{2}$$

$$\mathrm{Sp} = \frac{\mathrm{TN}}{\mathrm{TN} + \mathrm{FP}} \times 100\% \tag{3}$$

$$\mathrm{mAcc} = \frac{\mathrm{Se} + \mathrm{Sp}}{2} \times 100\% \tag{4}$$

We tested the performances from both single GbSQI feature-based classifier and multiple GbSQI feature-based classifier. For single GbSQI feature-based classifier, for each GbSQI feature, the threshold was firstly optimized and then was used for classifying. For multiple GbSQI feature-based classifier, the selected (sorted based on the performance of single GbSQI feature) features were input to an SVM classier for training a classification model. The Gaussian kernel was used in SVM. C and γ were optimized using a grid search method with the search range over C (from 0.5 to 724) and γ (from 4 to 32). For each test, a 10-fold cross validation was used. Figure 2 demonstrates the evaluation process for bSQI-2 pattern.

III. RESULTS

Figure 3 shows the total classification performances (mAcc, Se, Sp) for both single GbSQI feature-based classifiers and multiple GbSQI feature-based classifiers. In order to analyze and compare the performances of each classifier better, first

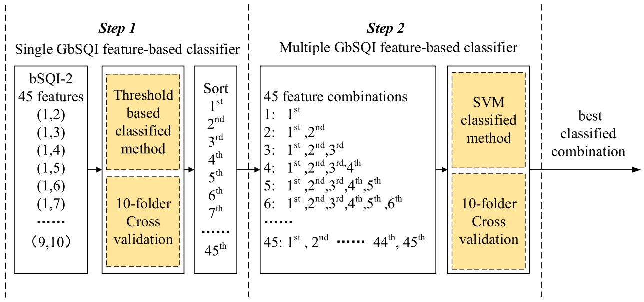


FIGURE 2. Demonstration of the evaluation process for the pattern of bSQI-2.

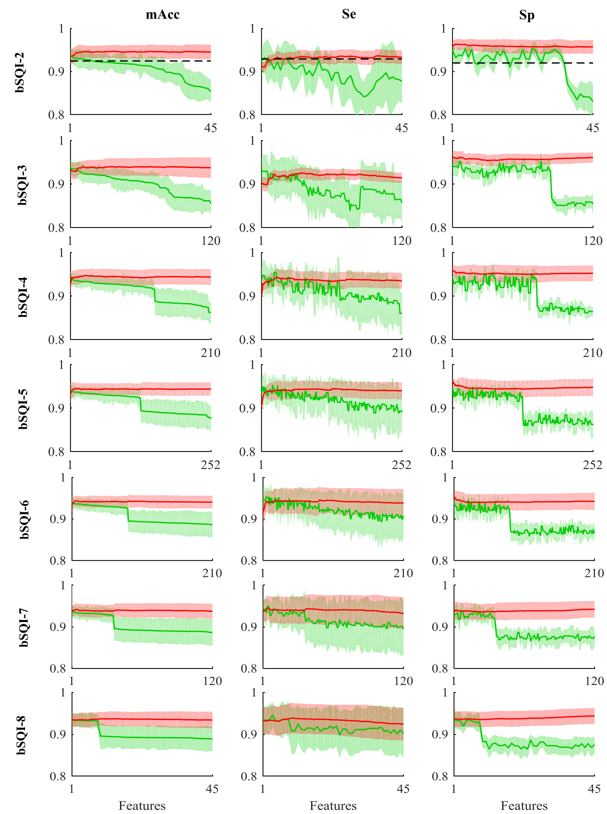


FIGURE 3. Classification performances (mAcc, Se, Sp) for both single GbSQI feature-based classifier and multiple GbSQI feature-based classifier. Green and red lines represent the performances of single GbSQI feature-based classifier and multiple GbSQI feature-based classifier, respectively. The shadow ranges represent the corresponding standard deviations (SDs). Horizontal axis represents the sorted GbSQI features by their mAcc values for single GbSQI feature-based classifier or the numbers of multiple GbSQI features used for multiple GbSQI feature-based classifier. Black dotted lines represent the results from the original bSQI-2 method using '*ep_limited*' and '*wqrs*' detectors as reference. Because the original bSQI only belonged to bSQI-2, other patterns did not have this reference values. Therefore, the black dotted lines were only in the plots of bSQI-2.

we sored the results mAcc from high to low, then we presented the results at 11 key points: the best mAcc, the mAcc corresponding to the top 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90% points respectively, as well as the worst one. Tables 4 and 5 reported the quantitative results respectively. Figure 4 shows the line chart, solid lines representing the

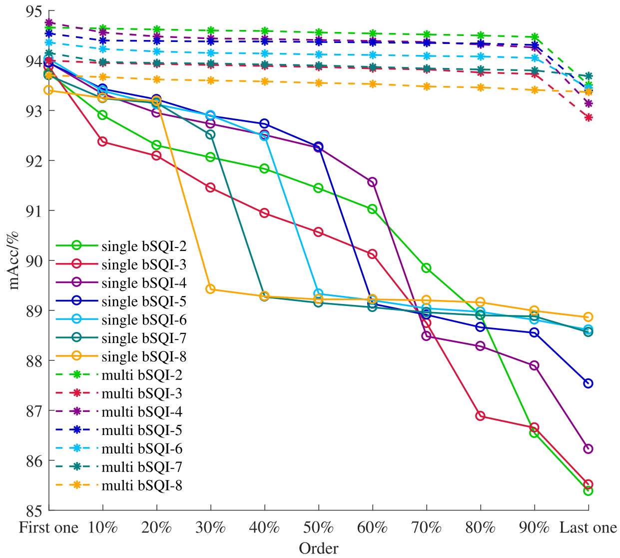


FIGURE 4. The line chart of the classified results (mAcc) of single-feature (solid lines) and multi-features combination (dotted lines).

classified results (mAcc) of single feature and dotted line representing the multi-feature combination.

The original bSQI proposed by Li et al. was based on the QRS detectors combination (1 '*ep_limited*' and 10 '*wqrs*' algorithms). In this paper. this classifier output the result mAcc of 91.81%. The original bSQI belonged to bSQI-2 in this study. 45 bSQI-2 classifiers were developed based on the ten QRS detectors. There were 18 bSQI-2 classifiers whose performances greater than 91.81%, while 26 bSQI-2 classifiers were less than 91.81%, i.e. 40% bSQI-2 classifiers perform better than original bSQI. Besides, the best result has reached 93.74% for bSQI-2 with an increasement about 2%. These results showed that generalizing bSQI had a great impact on bSQI-2. Not only bSQI-2 pattern, bSQI-3 to bSQI-6 also showed the better performance, as shown in Table 4. However, the increase in classified performance was limited. The best result was only increased by 0.29% from bSQI-6. Besides, the performances of bSQI-7 and bSQI-8 were gradually declining.

As shown in Tables 4 and 5 and Fig. 4, the performances of the single GbSQI feature-based classifiers and the multiple GbSQI feature-based classifiers had a certain difference. For the single GbSQI feature-based classifiers, in the same pattern, the performances of the classifiers with different QRS detector had been remarkably different. The maximum difference was 8.36% from bSQI-2 with the highest mAcc of 93.74% and lowest mAcc of 85.38%. However, this difference was not obvious in the multiple GbSQI feature-based classifiers. The maximum was only 1.62% from bSQI-4.

Besides, the single GbSQI feature-based classifiers and the multiple GbSQI feature-based classifiers also had the same feature. The classified performances of each pattern had both very little difference. For the single GbSQI feature-based classifiers, the difference was only 0.57% (94.03% of bSQI-6 and 93.40% of bSQI-8). While in the

TABLE 4. The classified results of single-feature for seven different patterns.

Order	Single feature mAcc /%						
	bSQI -2	bSQI -3	bSQI -4	bSQI -5	bSQI -6	bSQI -7	bSQI -8
First one	93.74	93.87	93.96	93.95	**94.03**	93.70	93.40
10%	92.90	92.37	93.33	93.43	93.39	93.24	93.25
20%	92.30	92.09	92.95	93.22	93.12	93.15	93.19
30%	92.06	91.45	92.73	92.89	92.90	92.51	89.42
40%	91.83	90.94	92.51	92.73	92.48	89.27	89.28
50%	91.44	90.56	92.25	92.27	89.33	89.15	89.22
60%	91.02	90.12	91.56	89.13	89.20	89.06	89.22
70%	89.84	88.74	88.48	88.91	89.04	88.96	89.20
80%	88.90	86.88	88.28	88.66	88.97	88.90	89.16
90%	86.54	86.65	87.89	88.55	88.81	88.88	88.99
Last one	85.38	85.51	86.22	87.53	88.61	88.56	88.86

TABLE 5. The classified results of multi-feature combination for seven different patterns.

Order	Multi features mAcc /%						
	bSQI-2	bSQI-3	bSQI-4	bSQI-5	bSQI-6	bSQI-7	bSQI-8
First one	94.66	94.00	**94.76**	94.54	94.36	94.15	93.70
10%	94.64	93.95	94.56	94.40	94.23	93.97	93.67
20%	94.62	93.93	94.48	94.39	94.18	93.95	93.62
30%	94.60	93.91	94.44	94.38	94.15	93.94	93.60
40%	94.59	93.89	94.43	94.38	94.14	93.92	93.58
50%	94.56	93.87	94.41	94.37	94.12	93.90	93.55
60%	94.54	93.84	94.39	94.36	94.11	93.87	93.53
70%	94.52	93.82	94.37	94.35	94.09	93.84	93.48
80%	94.50	93.76	94.32	94.34	94.08	93.82	93.46
90%	94.47	93.73	94.26	94.31	94.05	93.80	93.41
Last one	93.51	92.86	93.14	93.40	93.45	93.69	93.37

TABLE 6. Top three single GbSQI feature-based classifiers in each pattern.

Pattern	Methods combination		
	First	Second	Third
bSQI-2	**(5,9)**	**(3,5)**	(3,8)
bSQI-3	(5,9,10)	**(3,5,6)**	**(3,5,7)**
bSQI-4	**(5,6,8,9)**	**(5,6,7,9)**	**(3,5,6,9)**
bSQI-5	(4,5,6,8,9)	(4,5,6,7,8)	(5,6,7,8,9)
bSQI-6	(4,5,6,7,8,9)	(7,5,6, 8,9,10)	(4,5,7,8,9,10)
bSQI-7	(1, 5,6,7,8,9,10)	(4,5,6,7,8,9,10)	(3, 5,6,7,8,9,10)
bSQI-8	(1,3,4,5,6,7,9,10)	(1,3,4,5,6,7,8,10)	(1, 3,4,5,7,8,9,10)

multiple GbSQI feature-based classifiers, this difference was only 0.94% (94.76 % of bSQI-6 and 93.70% of bSQI-8).

Table 6 shows the top three single GbSQI feature-based classifiers in each pattern. Figure 5 counts the occurrence

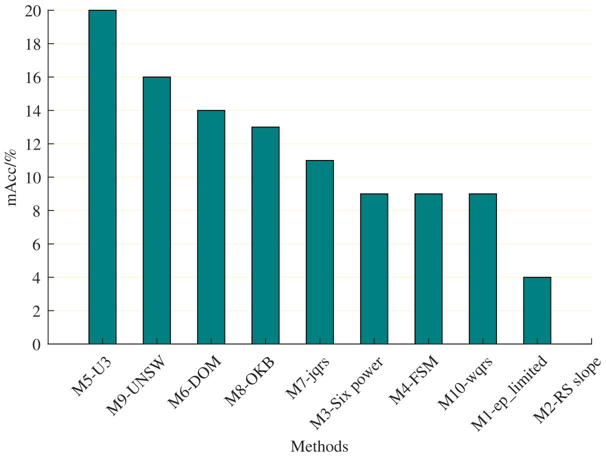


FIGURE 5. The occurrence frequency of each method in the top three single GbSQI feature-based classifiers in each pattern.

frequency for each QRS detector appeared in Table 6. Detector 5 (*U3*) has the highest frequency, and detector 9 (*UNSW*) have the second highest frequency, while detector 2 (*RS-slope*) has the lowest frequency. Therefore, it is recommended to select *U3*, *UNSW*, *DOM* and *OKB* detectors to establish the GbSQI.

TABLE 7. Top three single GbSQI feature-based classifiers in each pattern.

Method	Mean time /ms	SD /ms
1	3.3	0.37
2	11.8	0.60
3	18.7	3.8
4	0.27	0.06
5	2.3	0.21
6	2.5	0.19
7	3.5	0.23
8	1.8	0.14
9	19.7	0.79
10	123	10.27

In this study, all of the tests were implemented in MATLAB 2017a (The MathWorks, Inc., Natick, MA, USA) on Intel TM i5 CPU 3.30 GHz. Table 7 illustrates the mean time costs and SD values of ten QRS detectors by analyzing 12,000 10-s ECG segments in the 2011 PhysioNet/CinC Challenge database. All these ten detectors reported high calculation efficiency ($<$20 ms) except '*wqrs*' (mean 123 ms and SD 10.27 ms). Detector 4 FSM had the highest calculation efficiency (mean 0.27 ms and SD 0.06 ms).

IV. DISCUSSION

In this study, we generalized the two QRS detector-based bSQI to multiple QRS detector-based bSQI (GbSQI), to systematically test the effects of the type and number of QRS detectors on the SQA performance. Ten QRS detectors were selected to establish different GbSQI patterns, and seven patterns of GbSQI were tested. The classification performances of single GbSQI feature-based classifier and multiple GbSQI feature-based classifier were also systematically analyzed.

For the single GbSQI feature-based classifier, from bSQI-2 to bSQI-8, the best classification results of each pattern increased initially and then decreased. In addition, the best classification results were very similar for all the seven patterns. The biggest difference was only 0.57%. bSQI-6 output the highest results with mAcc of 94.03%, while bSQI-8 reported the lowest mAcc of 93.40%. From Table 6, we can find that the top single GbSQI feature-based classifiers of each pattern almost all included QRS detectors 5 and 9 (*U3* and *UNSW* algorithms). These two detectors also had the highest occurrence frequency (20 and 16 respectively) in the top three combinations of QRS detection methods, as shown in the Fig. 5. However, comparing vertical classified results in the Table 4, we found that the performances of the single GbSQI feature-based classifiers with different QRS detector had remarkably difference in the same pattern. Figure 4 clearly showed the same situation. All these phenomena highlighted the importance of the QRS detectors for the single GbSQI feature-based classifier. In addition, overmuch QRS detectors could not improve the classification performance, on the contrary, it would weaken the sensitivity, so reduce the performance and also increase the computational cost.

For the multiple GbSQI feature-based classifier, the classification results were slightly better than the single GbSQI feature-based classifier. Comparing with the single GbSQI feature-based classifier, the situation was different. On the one hand, the best classification results of each pattern had difference. bSQI-4 pattern showed the best performance. This pattern was more flexible and could get the specificity of multi QRS detectors by multiple features. On the other hand, in the same pattern, the performance of the classifier increased initially and then held steady as the features increasing. The classification performance was not proportional to features number. Taking bSQI-2 as an example, the classifier with 37 features reported the highest mAcc of 94.66%, while the classifier with 45 features reported the mAcc of 94.52%. Overmuch features could not improve the classification performance, on the contrary, it would weaken the sensitivity, so reduce the performance and also increase the computational cost.

No matter single GbSQI feature-based classifier or the multiple GbSQI feature-based classifier, the classification results all demonstrated the importance of QRS detectors selection. Two aspects should be considered for the QRS selection: one is the own character of the QRS detection algorithm, the other is the relationship between the selected detectors. Different QRS detection algorithms were sensitive to different types of noise [17]. First, the selected detectors should have high sensitivity and specific in QRS detection. Second, the selected detectors should be complementary.

A. INFLUENCE OF THE DETECTION PERFORMANCE OF QRS DETECTORS

In order to further analyze the detection performances of these ten detectors, we also systematically evaluated these

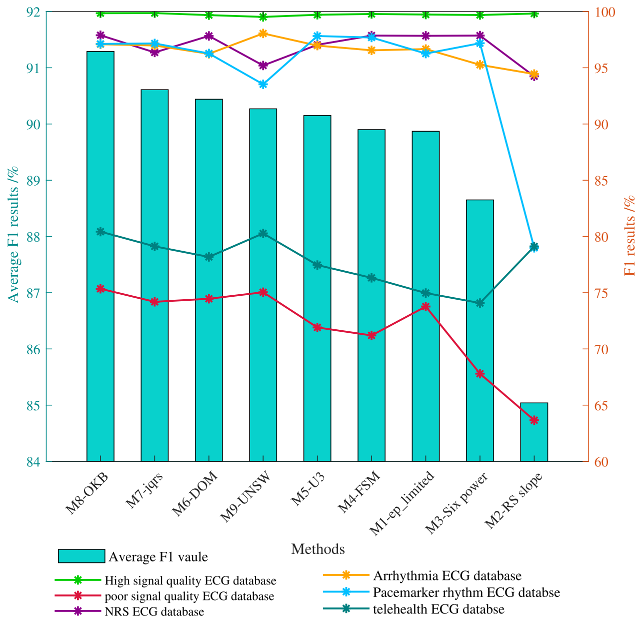


FIGURE 6. Detection results of these QRS detectors on six different ECG databases.

TABLE 8. The list of six ECG databases.

Database	# records	Record length (min)	Sample frequency (Hz)	Source
High quality ECGs	100	10	250	2014 PhysioNet/CinC Challenge training set https://physionet.org/challenge/2014/
Low quality ECGs	100	10	360	2014 PhysioNet/CinC Challenge augmented training set https://physionet.org/challenge/2014/
NRS ECG database	18	120	500	MIT-BIH NSR database https://physionet.org/physiobank/database/nsrdb/
Arrhythmia ECG database	44	30	360	MIT-BIH Arrhythmia database https://www.physionet.org/physiobank/database/mitdb/
Paced rhythm ECGs	4	30	360	MIT-BIH Arrhythmia database https://www.physionet.org/physiobank/database/mitdb/
Telehealth ECGs database	250	0.5	500	Harvard Dataverse TELE database https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/QTG0EP

detectors on six ECG databases [35]. Figure 6 shows the detection results, including the line graph for six databases and histogram for the average values. (‘*wqrs*’ detector was from the WFDB Software Package of PhysioNet [23]. The main role of this method was to compare the performance of GbSQI to the original bSQI. Due to this detector had special requirements on the data format, this paper did not test its performance on these six databases.). Table 8 shows the details of these six databases.

As shown in the Fig. 6, it should be noted that for the clean clinical ECG signals including normal subjects and arrhythmia patients, most QRS detectors had the similar performances and higher detection accuracies, whereas, all these algorithms did not do well for the poor signal quality ECG signals with high noise level and the performances were also with lager difference. The average F1 values of these six databases could further demonstrate the performances of these detectors. The top five detectors were *U3, DOM, ‘jqrs,’ OKB,* and *UNSW*. As shown in the Fig. 5, these five detectors were also the top five with the highest occurrence frequency in the top three single GbSQI feature-based classifiers in each pattern. Meanwhile, it was clear that the *RS-slope* detector had the worst detection performance, it also with the lowest frequency. In the Fig. 3, it was obvious that there were all rapid declines in the results mAcc of each pattern. The reason for these rapid declines was the *RS-slope* detector. For each pattern, when method 2 *RS-slope* detector appeared in the combination of the classifier, the classified results declined rapidly. Due to the length of space, we cannot show the details of all patterns. Table 9 shows the detailed classified results of single-feature for bSQI-2 patterns. From this table, we can find that the combinations with method 2 *RS-slope* all

TABLE 9. The detailed classified results of single-feature for bSQI-2 patterns.

Order	Combination of QRS detectors	Sp/(%)	Se/(%)	mAcc/(%)
1	(5,9)	95.00	92.47	93.74
2	(3,5)	93.23	93.25	93.24
3	(3,8)	93.87	92.23	93.05
4	(6,9)	93.63	92.18	92.91
5	(5,8)	95.23	90.58	92.90
6	(8,10)	93.35	92.18	92.77
7	(8,9)	95.10	90.34	92.72
8	(3,7)	91.98	92.91	92.45
9	(1,3)	91.59	93.01	92.30
10	(1,6)	92.72	91.75	92.23
11	(3,6)	93.52	90.82	92.17
12	(6,10)	93.79	90.53	92.16
13	(5,10)	92.95	91.26	92.10
14	(3,10)	92.58	91.55	92.06
15	(6,8)	94.61	89.32	91.96
16	(1,9)	93.66	90.19	91.93
17	(7,9)	94.83	88.93	91.88
18	(1,8)	93.22	90.44	91.83
19	(1,10)	91.05	92.57	91.81
20	(5,6)	93.11	90.33	91.72
21	(3,9)	92.88	90.53	91.71
22	(9,10)	93.00	89.99	91.50
23	(4,8)	94.98	87.91	91.44
24	(7,10)	92.97	89.70	91.34
25	(1,5)	92.83	89.80	91.31
26	(6,7)	95.33	87.23	91.28
27	(7,8)	94.95	87.08	91.02
28	(1,7)	92.98	88.44	90.71
29	(4,10)	92.11	89.03	90.57
30	(4,6)	94.41	86.65	90.53
31	(3,4)	94.20	85.48	89.84
32	(4,5)	94.62	85.04	89.83
33	(4,9)	95.10	84.17	89.64
34	(4,7)	94.58	84.36	89.47
35	(1,4)	93.73	85.19	89.46
36	(5,7)	91.83	85.96	88.90
37	(2,10)	87.18	87.72	87.45
38	(2,7)	85.84	88.25	87.04
39	(2,6)	84.77	88.88	86.83
40	(1,2)	83.45	89.71	86.58
41	(2,8)	84.83	88.25	86.54
42	(2,5)	84.13	88.78	86.46
43	(2,4)	84.16	88.30	86.23
44	(2,3)	84.04	88.15	86.09
45	(2,9)	83.05	87.71	85.38

had the worst classified results, marked by green. The result mAcc declined rapidly since the method 2 *RS-slope* detector appeared in the combination.

Obviously, the detection performance of the detectors was the most crucial point for the GbSQI classifier. If the selected detector had poor detection performance, the classified performance of the constructed classifier would be influenced greatly. Therefore, before selecting, it was necessary to evaluate the detector's performance. It should be noted that the databases used for evaluated was not only considering the clean ECG database but also the dynamic database with low signal quality. Since the most current QRS detectors had great performance on the clean ECG databases. Employing multi databases would be a good chose to test the performances of the QRS detectors.

B. INFLUENCE OF THE RELATIONSHIP BETWEEN THE SELECTED DETECTORS

The second key point was the relationship between the selected detectors for the GbSQI classifier. Although a large number of QRS detectors have been exited, almost all QRS detectors existed were based on the QRS complexes feature [36]. To track down the origin of most QRS detectors further, four origins were found: power-based, amplitude-based, slope-based, and curve length-based [37]. And these four origins were from the remarkable feature of QRS complexes: high power, high amplitude, steep slope, and long curve length. Figure 7 showed the origins and details of all these ten detectors. The details were divided into four aspects: filtering, extract feature method, threshold setting, and post-processing. Since ECG signals were easily disturbed by noise, the filtering was the first step frequently. Most detectors employed band-pass or low-pass filtering with different frequency range. And median filter was used to remove baseline drift in some detectors. The second step was to extract feature values. Derivative, squaring, and integration were the most common methods to enhance QRS waves. Some special algorithms were also employed, such as sixth power, U3 and length transform. All these feature extracting methods were to enhance the four remarkable features of QRS complexes. The third step was threshold setting. Some detectors employed the simple method to calculate the threshold. For example, one single fixed threshold was designed to be the 98% max value of the integrated signal in '*jqrs*' method. Whereas, in some detectors, the threshold setting was a comprehensive process. For example, in the *ep_limited* method, two sets of adaptive thresholds were employed to detect the QRS peaks in both filtered and integrated energy signals. Moreover, each set included double thresholds. The threshold I adaptively adjusted based upon the detected QRS signal and noise peak levels. The QRS/noise peak levels were designed to be the median value of the eight most-recent beats. The threshold II was set to 30% of the threshold I. As the final step, post-processing was mainly used to find the missed R peak and to eliminate the possibility of a false detection. If the current RR interval was too large, search back

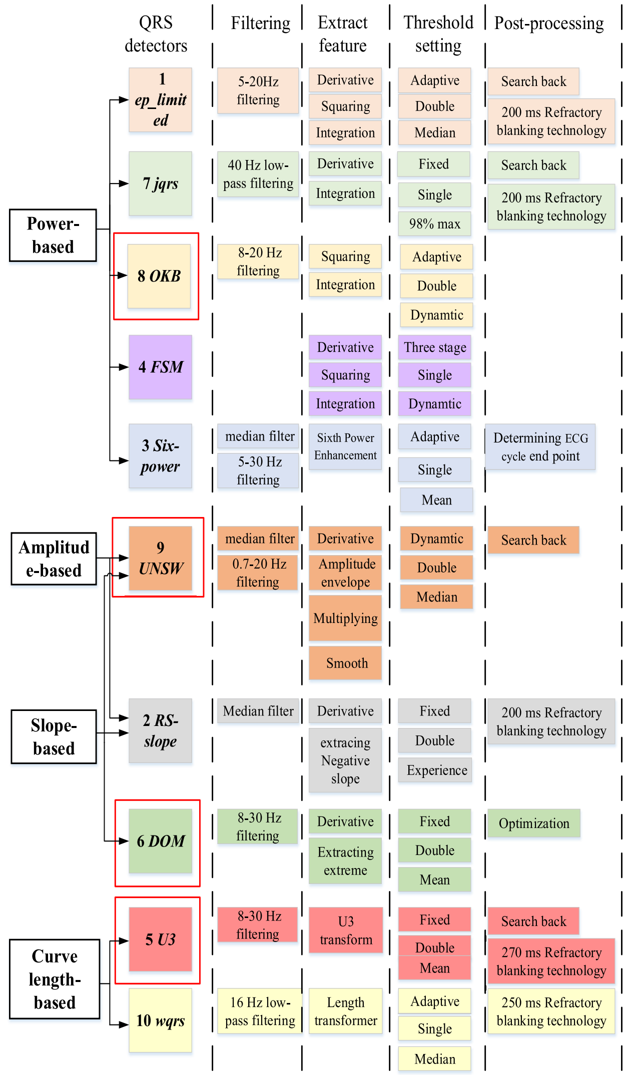


FIGURE 7. Relationship analysis for the selected ten QRS detectors.

process would start to find the missed R peak. And if the current RR interval was too small, refractory blanking technology would be used to eliminate the possibility of a false detection. Due to space limitations, this paper only simply introduced these QRS detectors. For more information, please pay attention to the corresponding references

In Table 6, it was obvious that there were seven combinations (marked in bold) with the detectors having different origins from the nine combinations in the top three single GbSQI feature-based classifiers, i.e., the proportion was closed to 80%. (Here, we only discussed the pattern bSQI-2, bSQI-3, and bSQI-4, since the patterns from bSQ-5 to bSQI-8 must include the detectors from the same origin.)

Taking the best single GbSQI feature-based classifier of bSQI-4 pattern as an example, *U3*, *UNSW*, *DOM*, and *OKB* detectors were selected. These four detectors were also with the top four highest frequency. The top four QRS detectors were emphasized by the red border. It was clear that these four detectors were from different origins. *U3* was based on the

QRS curve-length length; *UNSW* was based on the enhancing the QRS amplitude; *DOM* was based on the extreme point generated by QRS slope changing; *OKB* was based on the enhancing QRS power through integration. Therefore, we suggested that it was better to select the QRS detectors from different origin. In addition, accurate threshold setting and post-processing were also necessary to be considered. *DOM* was good at high-frequency noise by positive and negative thresholds. *UNSW* could remove spurious peaks by smoothing the feature. *U3 transform* could deal with baseline drift better. *OKB* method could reject the influence of the sudden amplitude change by double moving integration.

However, it was not an absolute issue. From table 8, it could be found that some combinations with the same origin could also show the well classified results, such as combination (3,8). Although detectors 3 and 8 were from the same origin, they employed very different way to extract feature. Whereas, the combinations (1,4), (4,7) and (1,7) showed the bad results. These three detectors had not only the same origin but also the same way to extract the feature.

From above analysis, it was obvious that the QRS detectors selecting was very important issue for the GbSQI classifier. The detection performance was the first problems to be considered. The evaluated results were better from multi databases including the dynamic, multi-noise signals. The relationship was second issue to be considered. It was better to select the detectors with many differences in every aspect. In this way, the selected detectors could be complementary on the noise resistance. It should be noted that the second issue was on the basis of the first problem.

Although the performance of the multiple GbSQI feature-based classifier was slightly better than that of the single GbSQI feature-based classifier, computational efficiency and expense also need to be considered, especially for the mobile or wearable device. Clearly, the multiple GbSQI feature-based classifier needs more time cost and internal storage, since it was based on more features than the single GbSQI feature-based classifier. Therefore, it was not worth pursuing a small increase in accuracy with a large computational cost. In this way, the single GbSQI feature-based classifier would be a better choice than the multiple GbSQI feature-based classifier. In the single GbSQI feature-based classifier, bSQI-4 and bSQI-6 showed the best performances. Researchers can choose the classifier based on their own needs.

There was also a special phenomenon that the performance of classifiers with the even number detectors were better than those with odd number detectors. No matter single GbSQI feature-based classifier or the multiple GbSQI feature-based classifier, the performance of bSQI-2 was better than that of bSQI-3, bSQI-4 was better than bSQI-5, and bSQI-6 was greater than bSQI-7, as shown in the Tables 4 and 5. Accordingly, the number of the selected QRS detectors was recommend to be an even number, such as 2, 4 or 6.

It should be noted that, the bSQI-2 classifier with the QRS detectors combination (1 '*ep_limited*' and 10 '*wqrs*' algorithms) was tested on the same ECG database in the work [13], and it gave the best result of 89.9%. In this paper, this classifier output the result mAcc of 91.81%. There are two major reasons for this. First, the definition of the accuracy was different between our work and the reference. Second, since we relabeled the data, the labeling methods were different.

A final important note is that the ten fast and effective QRS detectors we chose are unlikely to be the optimal algorithms. Many detectors could be selected for different contexts, equipment, diagnostic outcomes, or patient populations (particularly for patients with many arrhythmias). The general framework we have described in this paper is sufficiently flexible to allow the use of an arbitrary number of QRS detectors, selecting those that are most appropriate for a given situation. It should be noted that this study only analyzed the performance of GbSQI in the QRS detectors selecting and pattern establishing. Indeed, there are many other signal quality assessment indicators suggested in [2], [34], and [15]. If the optimal GbSQI was cooperated with other indexes, the detecting results must be enhanced. Regarding the feature selection, the literature described a method about how to determine the most relevant group of features [38]. We consider this as our future work.

V. CONCLUSION

In this study, we generalized the two QRS detectors-based bSQI to multiple QRS detectors-based bSQI (GbSQI), to compare the effects of the type and number of QRS detectors on the SQA performances of GbSQI. The general framework and the GbSQI definition described in this paper were sufficiently flexible to allow the use of an arbitrary number of QRS detectors, selecting those that are most appropriate for a given situation. The results presented here indicate that the detectors' detection performance was the first key point for detectors selection and relationship between the selected detectors was second issue. From the results, we recommended using *U3*, *UNSW*, *DOM*, and *OKB* detectors for calculating GbSQI. In conclusion, we have systematically analyzed the performances of GbSQI based classifier. These results and conclusions could offer reference for reasonably employing GbSQI.

AKNOWLEDGMENT

The authors acknowledge the support from the Southeast-Lenovo Wearable Heart-Sleep-Emotion Intelligent monitoring Lab.

CONFLICT OF INTEREST STATEMENT

The authors declare that there are no conflicts of interest to this work.

REFERENCES

[1] L. Johannesen and L. Galeotti, "Automatic ECG quality scoring methodology: Mimicking human annotators," *Physiol. Meas.*, vol. 33, no. 9, p. 1479, Sep. 2012, doi: 10.1088/0967-3334/33/9/1479.

[2] G. D. Clifford and G. B. Moody, "Signal quality in cardiorespiratory monitoring," *Physiol. Meas.*, vol. 33, no. 9, pp. 1–6, 2012, doi: 10.1088/0967-3334/33/9/e01.

[3] A. C. Maan, E. W. van Zwet, S. Man, M. J. Schalij, C. A. Swenne, and S. M. M. Oliveira-Martens, "Assessment of signal quality and electrode placement in ECGs using a reconstruction matrix," in *Proc. Comput. Cardiol.*, Hangzhou, China, 2011, pp. 289–292.

[4] I. Jekova, V. Krasteva, I. Christov, and R. Abächerli, "Threshold-based system for noise detection in multilead ECG recordings," *Physiol. Meas.*, vol. 33, no. 9, pp. 1463–1477, Sep. 2012, doi: 10.1088/0967-3334/33/9/1463.

[5] C. Liu, P. Li, L. Zhao, R. Wang, and F. Liu, "Real-time signal quality assessment for ECGs collected using mobile phones," in *Proc. Comput. Cardiol.*, Hangzhou, China, 2011, pp. 357–360.

[6] J. Allen and A. Murray, "Assessing ECG signal quality on a coronary care unit," *Physiol. Meas.*, vol. 17, no. 4, pp. 249–258, 1996.

[7] G. D. Clifford, J. Behar, Q. Li, and I. Rezek, "Signal quality indices and data fusion for determining clinical acceptability of electrocardiograms," *Physiol. Meas.*, vol. 33, no. 9, pp. 1419–1433, Aug. 2012, doi: 10.1088/0967-3334/33/9/1419.

[8] I. Silva, G. B. Moody, and L. Celi, "Improving the quality of ECGs collected using mobile phones: The PhysioNet/computing in cardiology challenge 2011," in *Proc. Comput. Cardiol.*, Hangzhou, China, 2011, pp. 273–276.

[9] H. Xia, G. A. Garcia, J. Bains, D. C. Wortham, and X. Zhao, "Matrix of regularity for improving the quality of ECGs," *Physiol. Meas.*, vol. 33, no. 9, pp. 1535–1548, Sep. 2012, doi: 10.1088/0967-3334/33/9/1535.

[10] G. B. Moody and R. G. Mark, "QRS morphology representation and noise estimation using the Karhunen–Loeve transform," in *Proc. Comput. Cardiol.*, Jerusalem, Israel, 1989, pp. 269–272.

[11] S. J. Redmond, N. H. Lovell, J. Basilakis, and B. G. Celler, "ECG quality measures in telecare monitoring," in *Proc. 30th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, Vancouver, BC, Canada, Aug. 2008, pp. 2869–2872.

[12] S. J. Redmond, Y. Xie, D. Chang, J. Basilakis, and N. H. Lovell, "Electrocardiogram signal quality measures for unsupervised telehealth environments," *Physiol. Meas.*, vol. 33, no. 9, pp. 1517–1533, 2012.

[13] Q. Li, R. G. Mark, and G. D. Clifford, "Robust heart rate estimation from multiple asynchronous noisy sources using signal quality indices and a Kalman filter," *Physiol. Meas.*, vol. 29, no. 1, pp. 15–32, 2008, doi: 10.1088/0967-3334/29/1/002.

[14] G. D. Clifford, D. Lopez, Q. Li, and I. Rezek, "Signal quality indices and data fusion for determining acceptability of electrocardiograms collected in noisy ambulatory environments," in *Proc. Comput. Cardiol.*, Hangzhou, China, 2011, pp. 285–288.

[15] L. Y. Di Marco *et al.*, "Evaluation of an algorithm based on single-condition decision rules for binary classification of 12-lead ambulatory ECG recording quality," *Physiol. Meas.*, vol. 33, no. 9, pp. 1435–1448, 2012, doi: 10.1088/0967-3334/33/9/1435.

[16] Y. Zhang, S. Wei, H. Liu, L. Zhao, and C. Liu, "A novel encoding Lempel–Ziv complexity algorithm for quantifying the irregularity of physiological time series," *Comput. Methods Programs Biomed.*, vol. 133, no. 8, pp. 7–15, 2016, doi: 10.1016/j.cmpb.2016.05.010.

[17] G. M. Friesen, T. C. Jannett, M. A. Jadallah, S. L. Yates, S. R. Quint, and H. T. Nagle, "A comparison of the noise sensitivity of nine QRS detection algorithms," *IEEE Trans. Biomed. Eng.*, vol. 37, no. 1, pp. 85–98, Jan. 1990, doi: 10.1109/10.43620.

[18] Q. Li and G. D. Clifford, "Dynamic time warping and machine learning for signal quality assessment of pulsatile signals," *Physiol. Meas.*, vol. 33, no. 9, p. 1491, Sep. 2012, doi: 10.1088/0967-3334/33/9/1491.

[19] Y. Zhang, C.-Y. Liu, S.-S. Wei, C.-Z. Wei, and F.-F. Liu, "ECG quality assessment based on a kernel support vector machine and genetic algorithm with a feature matrix," *J. Zhejiang Univ. Sci. C*, vol. 15, no. 7, pp. 564–573, 2014, doi: 10.1631/jzus.C1300264.

[20] *EP Limited*. Accessed: Sep. 8, 2017. [Online]. Available: http://www.eplimited.com/

[21] W. Zong, G. B. Moody, and D. Jiang, "A robust open-source algorithm to detect onset and duration of QRS complexes," in *Proc. Comput. Cardiol.*, Thessaloniki, Greece, 2003, pp. 737–740.

[22] L. A. Celi, L. Sarmenta, J. Rotberg, A. Marcelo, and G. Clifford, "Mobile care (Moca) for remote diagnosis and screening," *J. Health Inform. Develop. Countries*, vol. 3, no. 1, pp. 17–21, 2009.

[23] A. L. Goldberger *et al.*, "PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals," *Circulation*, vol. 101, no. 23, pp. e215–e220, 2000, doi: 10.1161/01.CIR.101.23.e215.

[24] P. S. Hamilton and W. J. Tompkins, "Quantitative investigation of QRS detection rules using the MIT/BIH arrhythmia database," *IEEE Trans. Biomed. Eng.*, vol. BME-12, no. 12, pp. 1157–1165, Dec. 1986, doi: 10.1109/TBME.1986.325695.

[25] M. B. Oefinger, "Monitoring transient repolarization segment morphology deviations in mouse ECG," Ph.D. dissertation, Dept. Elect. Eng. Comput. Sci., Massachusetts Inst. Technol., Boston, MA, USA, 2006.

[26] P. Podziemski and J. Gieraltowski, "Fetal heart rate discovery: Algorithm for detection of fetal heart rate from noisy, noninvasive fetal ECG recordings," in *Proc. Comput. Cardiol.*, Zaragoza, Spain, 2013, pp. 333–336.

[27] A. K. Dohare, V. Kumar, and R. Kumar, "An efficient new method for the detection of QRS in electrocardiogram," *Comput. Elect. Eng.*, vol. 40, no. 5, pp. 1717–1730, 2013, doi: 10.1016/j.compeleceng.2013.11.004.

[28] R. Gutiérrez-Rivas, J. J. García, W. P. Marnane, and Á. Hernández, "Novel real-time low-complexity QRS complex detector based on adaptive thresholding," *IEEE Sensors J.*, vol. 15, no. 10, pp. 6036–6043, Oct. 2015, doi: 10.1109/JSEN.2015.2450773.

[29] M. Paoletti and C. Marchesi, "Discovering dangerous patterns in long-term ambulatory ECG recordings using a fast QRS detection algorithm and explorative data analysis," *Comput. Methods Programs Biomed.*, vol. 82, no. 1, pp. 20–30, 2006, doi: 10.1016/j.cmpb.2006.01.005.

[30] Y. C. Yeh and W. J. Wang, "QRS complexes detection for ECG signal: The difference operation method," *Comput. Methods Programs Biomed.*, vol. 91, no. 3, pp. 245–254, Sep. 2008, doi: 10.1016/j.cmpb.2008.04.006.

[31] J. Behar, J. Oster, and G. D. Clifford, "Non-invasive FECG extraction from a set of abdominal sensors," in *Proc. Comput. Cardiol.*, Zaragoza, Spain, 2013, pp. 297–300.

[32] M. Elgendi, "Fast QRS detection with an optimized knowledge-based method: Evaluation on 11 standard ECG databases," *PLoS ONE*, vol. 8, no. 9, p. e73557, 2013, doi: 10.1371/journal.pone.0073557.

[33] H. Khamis, R. Weiss, Y. Xie, C.-W. Chen, N. H. Lovell, and S. J. Redmond, "QRS detection algorithm for telehealth electrocardiogram recordings," *IEEE Trans. Biomed. Eng.*, vol. 63, no. 7, pp. 1377–1388, Jul. 2016, doi: 10.1109/TBME.2016.2549060.

[34] D. Hayn, B. Jammerbund, and G. Schreier, "QRS detection based ECG quality assessment," *Physiol. Meas.*, vol. 33, no. 9, pp. 1449–1461, Sep. 2012, doi: 10.1088/0967-3334/33/9/1449.

[35] F. Liu *et al.*, "Performance analysis of ten common QRS detectors on different ECG application cases," *J. Healthcare Eng.*, vol. 2018, May 2018, Art. no. 9050812, doi: 10.1155/2018/9050812.

[36] B.-U. Köhler, C. Hennig, and R. Orglmeister, "The principles of software QRS detection," *IEEE Eng. Med. Biol. Mag.*, vol. 21, no. 1, pp. 42–57, Jan./Feb. 2002, doi: 10.1109/51.993193.

[37] M. Elgendi, B. Eskofier, S. Dokos, and D. Abbott, "Revisiting QRS detection methodologies for portable, wearable, battery-operated, and wireless ECG systems," *PLoS ONE*, vol. 9, no. 1, p. e84018, 2014, doi: 10.1371/journal.pone.0084018.

[38] V. Monasterio, F. Burgess, and G. D. Clifford, "Robust classification of neonatal apnoea-related desaturations," *Physiol. Meas.*, vol. 33, no. 9, pp. 1503–1516, Sep. 2012, doi: 10.1088/0967-3334/33/9/1503.

FEIFEI LIU received the B.S. and Ph.D. degrees in biomedical engineering from Shandong University, China, in 2010 and 2015, respectively. He has completed the Post-Doctoral trainings at Shandong University from 2015 to 2017.

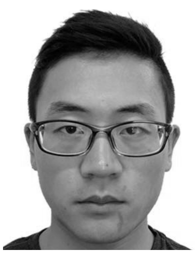
She is currently a Post-Doctoral with Southeast University, Nanjing, China. She has published eight original journal/conference papers. Her research topics include mHealth and intelligent monitoring, machine learning and big data processing for physiological signals, early detection of CADs, and device development for CADs.

CHENGYU LIU (M'14) received the B.S. and Ph.D. degrees in biomedical engineering from Shandong University, China, in 2005 and 2010, respectively. He has completed the Post-Doctoral trainings at Shandong University from 2010 to 2013, Newcastle University, U.K., from 2013 to 2014, and Emory University, USA from 2015 to 2017.

He is currently the Director and a Professor with the Southeast-Lenovo Wearable Heart-Sleep-Emotion Intelligent Monitoring Lab, School of Instrument Science and Engineering, Southeast University, Nanjing, China. He was the PI on over 10 awarded grants attracting a total of over one million. He has published over 130 journal/conference papers, eight chapters in books, and 15 invention patents. His research topics include mHealth and intelligent monitoring, machine learning and big data processing for physiological signals, early detection of CADs, device development for CADs, and sleep and emotion monitoring. He is currently a Federation Journal Committee Member of the International Federation for Medical and Biological Engineering.

LINA ZHAO received the master's degree in control science and control theory from Shandong University, Jinan, China, in 2008, where she is currently pursuing the Ph.D. degree in biomedical sciences and mainly involving in big data processing for physiological signals.

XINGE JIANG received the master's degree in control science and control theory from Shandong University, Jinan, China, in 2006, where she is currently pursuing the Ph.D. degree in biomedical sciences and mainly involving in pulse signals for cardiovascular disease and sleep apnea syndrome.



ZHIMIN ZHANG received the B.E. degree in biomedical engineering from Shandong University, Jinan, China, in 2014, where he is currently pursuing the Ph.D. degree. He is currently a Research Student with the School of Information Technology and Electrical Engineering, The University of Queensland, where he is mainly involving in ECG, EEG signals for cardiovascular disease and sleep stages classification using compressed sensing, machine learning, and entropy analysis techniques.

JIANQING LI received the B.S. and M.S. degrees in automatic technology and the Ph.D. degree in measurement technology and instruments from the School of Instrument Science and Engineering, Southeast Univesity, Nanjing, China, in 1986 and 1992, respectively. He has been with Southeast University since 1986, where he is currently a Professor with the School of Instrument Science and Engineering. He is currently a Professor with the School of Basic Medical Sciences, Nanjing Medical University. His research topics include mHealth and intelligent monitoring, machine learning and big data processing for physiological signals, early detection of CADs, device development for CADs, and sleep and emotion monitoring.

SHOUSHUI WEI received the M.S. degree from the Nanjing University of Aeronautics and Astronautics, Nanjing, China, in 1992. He is currently a Professor with the Institute of Biomedical Engineering, Shandong University, China. His research topics include mHealth and intelligent monitoring, machine learning and big data processing for physiological signals, early detection of CADs, device development for CADs, and sleep and emotion monitoring.

YUAN ZHANG (SM'14) received the M.S. degree in communication systems and the Ph.D. degree in control theory & engineering from Shandong University, China, in 2003 and 2012, respectively. He is currently an Associate Professor with the University of Jinan, China.

He was a Visiting Professor with the Computer Science Department, Georgia State University, USA, in 2014. He has published over 50 peer-reviewed papers in international journals and conference proceedings, one book chapters, and six patents in the areas of mHealth and biomedical big data analytics. His research interests are wearable sensing for smart health, machine learning for auxiliary diagnosis, and biomedical big data analytics. His research has been extensively supported by the Natural Science Foundation of China, the China Postdoctoral Science Foundation, and the Natural Science Foundation of Shandong Province with total grant funding over 1.4 million RMB. He is a Senior Member of ACM. He has served on the technical program committee for numerous international conferences. He is an Associate Editor for IEEE Access. He has served as a Leading Guest Editor for six special issues of the IEEE, Elsevier, Springer, and InderScience publications, including the IEEE Internet of Things Journal special issue on Wearable Sensor Based Big Data Analysis for Smart Health and the IEEE Journal of Biomedical and Health Informatics special issue on Pervasive Sensing and Machine Learning for Mental Health.