



Wearable Electrocardiogram Quality Assessment Using Wavelet Scattering and LSTM

Feifei Liu¹, Shengxiang Xia^{1*}, Shoushui Wei², Lei Chen³, Yonglian Ren¹, Xiaofei Ren⁴, Zheng Xu¹, Sen Ai¹ and Chengyu Liu^{5*}

¹School of Science, Shandong Jianzhu University, Jinan, China, ²School of Control Science and Engineering, Shandong University, Jinan, China, ³School of Science and Technology, Shandong University of Traditional Chinese Medicine, Jinan, China, ⁴School of Information and Electrical Engineering, Shandong Jianzhu University, Jinan, China, ⁵School of Instrument Science and Engineering, Southeast University, Nanjing, China

OPEN ACCESS

Edited by:

Jichao Zhao,
The University of Auckland,
New Zealand

Reviewed by:

Lakhan Dev Sharma,
VIT-AP University, India
Óscar Barquero-Pérez,
Rey Juan Carlos University, Spain

*Correspondence:

Shengxiang Xia
xias@sdjzu.edu.cn
Chengyu Liu
chengyu@seu.edu.cn

Specialty section:

This article was submitted to
Computational Physiology and
Medicine,
a section of the journal
Frontiers in Physiology

Received: 27 March 2022

Accepted: 30 May 2022

Published: 30 June 2022

Citation:

Liu F, Xia S, Wei S, Chen L, Ren Y,
Ren X, Xu Z, Ai S and Liu C (2022)
Wearable Electrocardiogram Quality
Assessment Using Wavelet Scattering
and LSTM.
Front. Physiol. 13:905447.
doi: 10.3389/fphys.2022.905447

As the fast development of wearable devices and Internet of things technologies, real-time monitoring of ECG signals is quite critical for cardiovascular diseases. However, dynamic ECG signals recorded in free-living conditions suffered from extremely serious noise pollution. Presently, most algorithms for ECG signal evaluation were designed to divide signals into acceptable and unacceptable. Such classifications were not enough for real-time cardiovascular disease monitoring. In the study, a wearable ECG quality database with 50,085 recordings was built, including A/B/C (or high quality/medium quality/low quality) three quality grades (A: high quality signals can be used for CVD detection; B: slight contaminated signals can be used for heart rate extracting; C: heavily polluted signals need to be abandoned). A new SQA classification method based on a three-layer wavelet scattering network and transfer learning LSTM was proposed in this study, which can extract more systematic and comprehensive characteristics by analyzing the signals thoroughly and deeply. Experimental results ($mACC = 98.56\%$, $mF_1 = 98.55\%$, $Se_A = 97.90\%$, $Se_B = 98.16\%$, $Se_C = 99.60\%$, $+P_A = 98.52\%$, $+P_B = 97.60\%$, $+P_C = 99.54\%$, $F_{1A} = 98.20\%$, $F_{1B} = 97.90\%$, $F_{1C} = 99.60\%$) and real data validations proved that this proposed method showed the high accuracy, robustness, and computationally efficiency. It has the ability to evaluate the long-term dynamic ECG signal quality. It is advantageous to promoting cardiovascular disease monitoring by removing contaminating signals and selecting high-quality signal segments for further analysis.

Keywords: dynamic electrocardiogram, signal-quality assessment, wavelet scattering, signal-quality index, long short-term memory network

INTRODUCTION

Cardiovascular diseases (CVDs) are the most common non-communicable diseases globally, responsible for an estimated 17.8 million deaths in 2017, accounting for 31% of all global deaths, of which more than three quarters were in low income and middle-income countries (Liu et al., 2018; Roth et al., 2018). Therefore, early continuous monitoring and prevention for CVDs are very urgent. The recent commercial availability of wearable devices and Internet of things (IoT) technologies with cardiovascular disease detection capabilities has revolutionized the diagnosis and management of these common medical issues, as it has placed the power of arrhythmia detection into

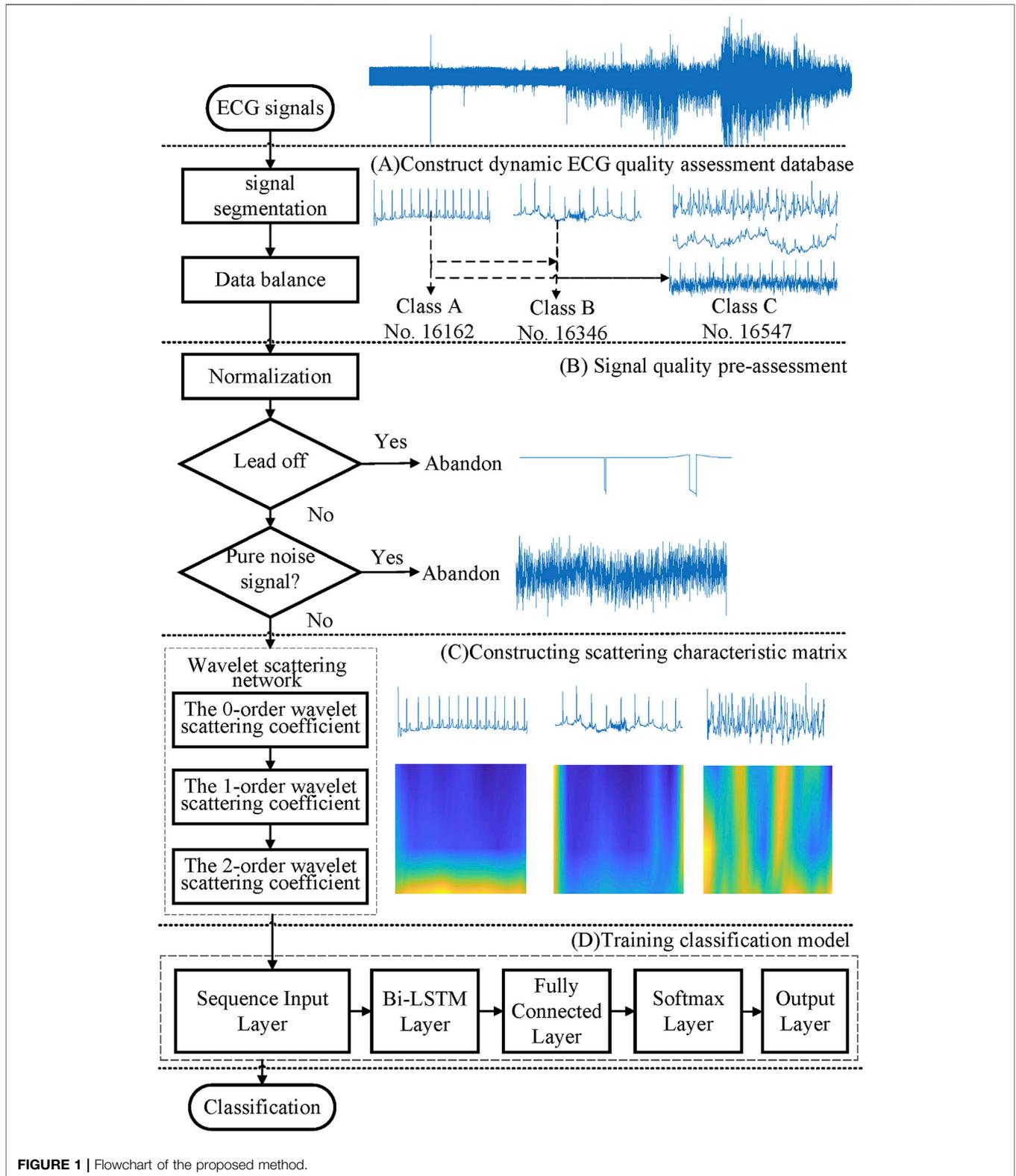


FIGURE 1 | Flowchart of the proposed method.

the hands of the patient (Liu et al., 2019b). However, the dynamic long-term ECG signals suffer from extremely serious noise pollution due to the dynamic long-term unsupervised free-

living monitoring environment (Huerta et al., 2019). A recent study of 100 patients undergoing cardioversion for atrial fibrillation showed that 34% of wearable devices' ECG

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

Score	Description for signal quality scoring
1	ECGs have a clear QRS complex and T wave. Baseline wander does not influence the identification for QRS.
0.75	Transient high amplitude impulse exists, but not more than three episodes. The majority of QRS complexes can be visually clearly identified.
0.5	Both large baseline wander and transient high amplitude impulse exist. It is challenging to visually clearly identify the QRS complexes in a 2–3 s time window.
0.25	More serious larger noises exist, such as strong Gaussian noise and signal saturation and others. In these noise episodes, it is impossible to identify the QRS complex. But at least 4–5 s continuous identifiable heart beats are visible.
0	Strong noises occupy in the more than 5 s episode. It is very hard to identify the heart beat for the most signal.

recordings were categorized as “unclassified” by the device algorithm due to unclear reasons or baseline artifacts and low amplitude recordings (Kaptoge et al., 2019). Poor electrocardiographic signal quality can result in misinterpretation and inappropriate results, hazard the correct diagnosis information (Andrea et al., 2018), increases the risk of false alerts (Liu et al., 2011), which may lead to unnecessary medical referrals and testing (Ip, 2019), and increase the workload of physicians (Zhao and Zhang, 2018). Consequently, it is quite urgent to evaluate the quality of wearable dynamic ECG signals, to eliminate signals with serious noise pollution, to distinguish between clean signals that can be used for disease diagnosis and mildly contaminated signals that can only be used for heart rate extraction, which can effectively reduce false alarm and avoid interference with CVD diagnosis (Xu et al., 2021).

The quality evaluation of wearable dynamic ECG signals has aroused the researchers’ extensive attention (Satija et al., 2018; Liu et al., 2019a; Huerta et al., 2019; Liu F. et al., 2020). As early as in 2011, the PhysioNet Cardiology Challenge addressed the issue of developing an efficient algorithm being able to run in real-time on a mobile phone, which can be able to indicate within a few seconds, while the patient is still present, if the ECG is of adequate quality for interpretation, or if another recording should be made (Silva et al., 2011). From then on, many wearable ECG signal-quality assessment (SQA) methods have been developed, and a variety of signal-quality indexes (SQI) have been explored based on the extraction of statistical, morphological, nonlinear, or time-frequency domain features etc. from the signals (Smital et al., 2020). For instance, Li et al. (2008) proposed a bSQI index based on the principle that different R-wave detectors should be nearly the same for clean ECG signals, while they should have different results for ECG signals polluted by noises, and got a good grade in the 2011 PhysioNet/CinC Challenge (Clifford and Moody, 2012). Based on this index, Liu et al. (2018) proposed the generalized bSQI index, generalized the two QRS detector-based bSQI to multiple QRS detector-based bSQI, and mainly studied the effects of type and number of R wave detectors on signal-quality assessment performances. Smital et al. (2020) proposed continuous signal-to-noise ratio curve using the time-frequency domain approach, including the Wavelet Wiener Filtering method and short-time Fourier transform frequency approach, to estimate real-time quality assessment of long-term ECG signals recorded by wearables in free-living conditions. He et al. (2020)

proposed a fuzzy comprehensive evaluation algorithm based on characteristics of ECG waveform and each band, to comprehensively evaluate the quality of ECG signals.

However, existing SQA methods highly demand robust methods for accurate and reliable detection and measurement of morphological and RR-interval features from noise-free and noisy ECG signals. Although the ECG morphology feature-based methods have shown promising results in noise-free ECG recordings, accuracy and robustness of QRS complex detection and waveform delineation methods are significantly degraded in the presence of severe muscle artifacts and other external noise (Satija et al., 2017). Also, most SQA methods graded the dynamic ECG signal quality into two groups: acceptable versus unacceptable (or good versus bad). In fact, in some wearable ECG signals only R wave could be detected, other waves such as P or ST were drowned out by the noise (Xu et al., 2021). These signals cannot be used for some CVD detection, but they also cannot be abandoned as heart rate information can be obtained. Therefore, these signals could not be simply divided into acceptable or unacceptable. In this study, a wearable ECG quality database with 50,085 recordings was built, which included A/B/C (or high quality/medium quality/low quality) three quality grades (A: high-quality signals can be used for CVD detection; B: slightly contaminated signals can be used for heart rate extracting; C: heavily polluted signals need to be abandoned). The research has revealed that traditional indexes merely based on morphological, nonlinear, or time-frequency domain features did not perform well on this database, as class B signals were easily confused with class A signals. It is essential to extract more systematic and comprehensive characteristics by analyzing the signals thoroughly and deeply.

The wavelet scattering algorithm, proposed by Mallat (2012), Bruna and Mallat (2013), and Anden and Mallat (2014) using the deep convolutional network architecture, iterated over wavelet convolution, nonlinear modulus, and averaging operators to compute higher-order scattering coefficients, which can build the translation invariant, stable and informative signal representation. The wavelet transform method provided stability under the action of small diffeomorphism, while the nonlinear operation and the integration over time give translation invariance (Tang et al., 2015). Cascading wavelet transforms allowed the recovery of high frequencies lost when averaging the absolute values of coefficients of previous wavelet transforms (Destouet et al., 2021). These preprocessing methods provided an

in-depth analysis of signals. First-order scattering coefficients characterize persistent phenomena such as tendency and envelope, while second-order scattering coefficients characterize transient phenomena such as shock signals and amplitude modulation (Anden and Mallat, 2014). The wavelet scattering method has been widely used for acoustic scene classification (Li et al., 2019), speech recognition (Fousek et al., 2015; Joy et al., 2020), and heart sound classification (Mei et al., 2021), which yielded efficient representations for audio processing. However, wavelet scattering currently was seldom used in ECG analysis and application. Sepúlveda et al. (2021) extracted features of the signal at different time scales using the wavelet scattering algorithm for emotion recognition. Also, Liu Z. et al. (2020) employed wavelet scattering transform for ECG beat classification.

In this study, in order to address the classification issue of A/B/C three quality levels wearable ECG signals, a new SQA classification model was proposed based on a three-layer wavelet scattering network and transfer learning long short-term memory (LSTM) method. As the result shows, it performed very well on the quality assessment of wearable dynamic ECGs.

MATERIALS AND METHODS

Figure 1 displays the flowchart of the proposed method. It first established a wearable ECG quality database with 50,085 recordings from two public databases. Then, a quality pre-assessment was established, to delete the lead-fall signals and pure noise, and to avoid the adverse impact of invalid samples on the training models. Also then, the scattering characteristic matrix was extracted by applying a three-layer wavelet scattering network. Finally, a bi-directional long short-term memory (Bi-LSTM) network was employed to train the classification model.

Database

A total of 50,085 recordings of wearable ECGs were used in this study, which were from the Brno University of Technology ECG Quality Database (BUTQDB) (Nemcova et al., 2020) and the 2011 PhysioNet/CinC Challenge (Goldberger et al., 2000; Silva et al., 2011). In the Brno University of Technology ECG Quality Database, the data comprise 18 long-term recordings of single-lead ECGs, collected from 15 subjects (9 females, six males) aged between 21 and 83 years. The signals are longer than 24 h which were detected using the Bittium Faros 180 device (mobile ECG recorder) under free-living conditions. All patients on the datasets did not have any diagnostics. The database contains signal-quality labels for partly data provided by three ECG experts, as well as the consensus of these experts, who grouped the signals into three quality classes.

Class A (high quality): all significant waveforms (P\QRS\ST\T waves) are clearly visible and the onsets and offsets of these waveforms can be detected reliably. The recording with no obvious noise can be used for the diagnosis of cardiovascular disease.

Class B (medium quality): the noise level is increased and significant points in the ECG are unclear (for example, PR interval and/or QRS duration cannot be measured reliably), but QRS complexes are clearly visible and the signal enables reliable QRS detection. Heart rate can be measured correctly.

Class C (low quality): QRS complexes cannot be detected reliably and the signal is unsuitable for any analysis. Heart rate cannot be measured correctly. These signals will interfere with the diagnosis of the cardiovascular disease and need to be removed.

In this study, the annotated recordings and segments have been divided into many fragments of unequal length based on the signal-quality labels provided by ECG experts. Each fragment has an independent label. Also, we segmented the annotated fragments into 10-s fragments with no overlap. Also, a sample of 10-s is the input data to the classification model. The number of class A is 11,708, class B is 7,860, and class C is only 657. It was obvious that data distribution was unbalanced. As we know, imbalanced classes will greatly reduce the generalization ability of the classification model (Clifford et al., 2012). Balancing the database classes can overcome this problem. In this study, we balanced the dataset by expanding the class C data using two ways: one is importing same class data from other databases, and the other is adding noise to clean data.

A total of 1,000 recordings of standard 12-lead ECGs were provided by the 2011 PhysioNet/CinC Challenge (Silva et al., 2011). In 1,000 12-lead ECGs, 773 were labeled as “acceptable,” 225 were “unacceptable,” and two were “intermediate.” Each signal had a length of 10 s. All patients on the datasets did not have any diagnostics. In Liu et al. (2018), every single channel of ECGs had been scored and re-labeled by five researchers, and a total of 9,941 acceptable and a total of 2,059 unacceptable 10-s ECG segments were obtained. In this study, based on the scores in the Liu et al. (2018) and Liu et al. (2019b), we annotated all the leads (10 s segments) individually. For every single channel of ECGs, five scores $S_i, i = 1 \dots 5$ were given by five researchers as presented in **Table 1**. Also, the average score $\bar{S} (\bar{S} = \frac{1}{5} \sum_{i=1}^5 S_i)$ was used as a threshold. The signals were re-labeled as “class A” if it was higher than 0.75, as “class B” if it was higher than 0.25 and lower than 0.75. Otherwise, the signal was labeled as “class C.” We obtained a total of 4,455 “class A,” a total of 5,486 “class B,” and a total of 2,059 “class C” 10-s ECG segments.

If all signals from these two databases were used together simply, the number of class A would be 16,163, class B would be 13,346, and class C would be only 2,716. It was obvious that data distribution was also extremely unbalanced. In this way, class B 7860 signals from the Brno University of Technology ECG Quality Database were employed to expand the class C data by adding noise from the PhysioNet noise stress test database (NSTDB) (Moody et al., 1984). Also, 10,000 recordings chosen randomly from class A were also used to expand the class B database (3,000) and class C database (7,000) by adding noise from NSTDB, for class B; the signal-to-noise ratio (SNR) was equal to 10db, for class C was -10db. In the NSTDB database, three types of noise were existing, record *bw* contains baseline wander noise, record *em* contains electrode motion artifact with a significant amount of baseline wander and muscle noise as well, and record *ma* contains mainly muscle noise (Clifford et al.,

TABLE 2 | Dynamic ECG quality assessment database composition.

Quality class	# Record		Source	Sampling frequency (Hz)	Record length (second)
A	16,163	11,708	BUTQDB	1,000	10
		4,455	2011 PhysioNet/CinC	500	10
B	16,346	7,860	BUTQDB	1,000	10
		5,486	2011 PhysioNet/CinC	500	10
C	17,576	3,000	Class A signal set randomly + NSTDB	—	10
		657	BUTQDB	1,000	10
		2,059	2011 PhysioNet/CinC	500	10
		7,860	Class B signal form BUTQDB + NSTDB	—	10
Total	50,085	7,000	Class A signal set randomly + NSTDB	—	10
		—	250	10	—

2012). Because the baseline wander (*bw*) has little effect on signal quality, Gaussian noises were added to this type of noise to generate new noisy records *gbw*. **Table 2** shows the details of dynamic ECG quality assessment database composition.

Signal-Quality Pre-assessment

Signals from different databases need to be preprocessed. First, each ECG signal was down-sampled to 250 Hz. Then, the min-max standardization method was used to map the original ECG signal data to [0–1]. Lead fall is very common in wearable dynamic ECG signals. Lead fall detection was an important way to decrease data storage costs and computing overhead for wearable devices. **Figure 2** shows several typical cases of lead-fall signals. In this study, if one ECG signal was present with a constant voltage of more than 80% of the recording, it was defined as lead falling. Occasionally, detached electrodes were adhered to clothing and received Gaussian noise signals. If the signal was pure noise, it was needed to be eliminated. Based on the spectrum range of ECG signal, that is, 0–40 Hz, if the ratio of power spectrum energy of the signal in the range of 0–40 Hz to the total energy is less than 30%, it indicates that the main component of the signal is not the ECG signal but the noise signal, which can be directly discarded. By pre-assessment, 1,029 lead off and pure noise signals were eliminated, which belong to the class C. The calculation formula is as follows:

$$MPSQI = \frac{\int_{0Hz}^{40Hz} p(f)df}{\int_{0Hz}^{500Hz} p(f)df} \quad (1)$$

$$PNSQI = \begin{cases} 1 & \text{pure noise if } MPSQI < 30\% \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

Wavelet Scattering Analysis

The wavelet scattering network has the characteristics of translation invariance, deformation stability, and high frequency preservation (Bruna and Mallat, 2013), and it is very sensitive to the deformation of wearable dynamic ECG signals. In this study, the scale function ϕ_I and Morlet wavelet function ψ_λ were employed to construct a three-layer wavelet scattering network. Through this network, ECG signals generate scattering coefficients of order 0, 1, and 2, which can cover the

whole frequency domain of the signal. The network constructing steps are as follows:

- 1) ECG signal $X(t)$ was convoluted with the scale function ϕ_I to obtain the 0-order wavelet scattering coefficient S_0 .

$$S_0 X(t) = X(t) * \phi_I \quad (3)$$

- 2) ECG signal $X(t)$ were convoluted with the first-order wavelet functions, $\psi_{\lambda_{1,i}}$, and the first-order scattering propagation operators $U_{\lambda_{1,i}}$ were generated by nonlinear modulus operation.

$$U_{\lambda_{1,i}} = |X(t) * \psi_{\lambda_{1,i}}|, \quad i = 1 \cdots n. \quad (4)$$

- 3) The first-order wavelet scattering coefficients $S_{1,i}$ are obtained by the convolution of propagators $U_{\lambda_{1,i}}$ and scaling function ϕ_I .

$$S_{1,i} X(t) = |X(t) * \psi_{\lambda_{1,i}}| * \phi_I, \quad i = 1 \cdots n. \quad (5)$$

- 4) The first-order scattering propagator $U_{\lambda_{1,i}}$ were convoluted with the second-order wavelet functions $\psi_{\lambda_{2,j}}$, and the second-order scattering propagators $U_{\lambda_{2,ij}}$ were generated by the nonlinear modulus operation.

$$U_{\lambda_{2,ij}} = ||X(t) * \psi_{\lambda_{1,i}}| * \psi_{\lambda_{2,j}}|, \quad i = 1 \cdots n, j = 1 \cdots m. \quad (6)$$

- 5) The second-order wavelet scattering coefficients $S_{2,ij}$ are obtained by the convolution of the second-order scattering propagators $U_{\lambda_{2,ij}}$ and scaling function ϕ_I .

$$S_{2,ij} X(t) = ||X(t) * \psi_{\lambda_{1,i}}| * \psi_{\lambda_{2,j}}| * \phi_I, \quad i = 1 \cdots n, j = 1 \cdots m. \quad (7)$$

The scattering network can contain more than three layers, but in practice, energy is dissipated with each iteration. Therefore, in this study, three layers were employed. The zero-order wavelet scattering coefficient S_0 mainly average the input ECG signal. The first-order wavelet scattering coefficient S_1 captures details lost in the first step, similar to the scale-invariant feature transformation function. The second-order wavelet scattering coefficient S_2 provides supplementary information that improves classification. The scattering characteristic matrix is composed of these three-layer scattering coefficients S_0, S_1, S_2 . **Figure 3**

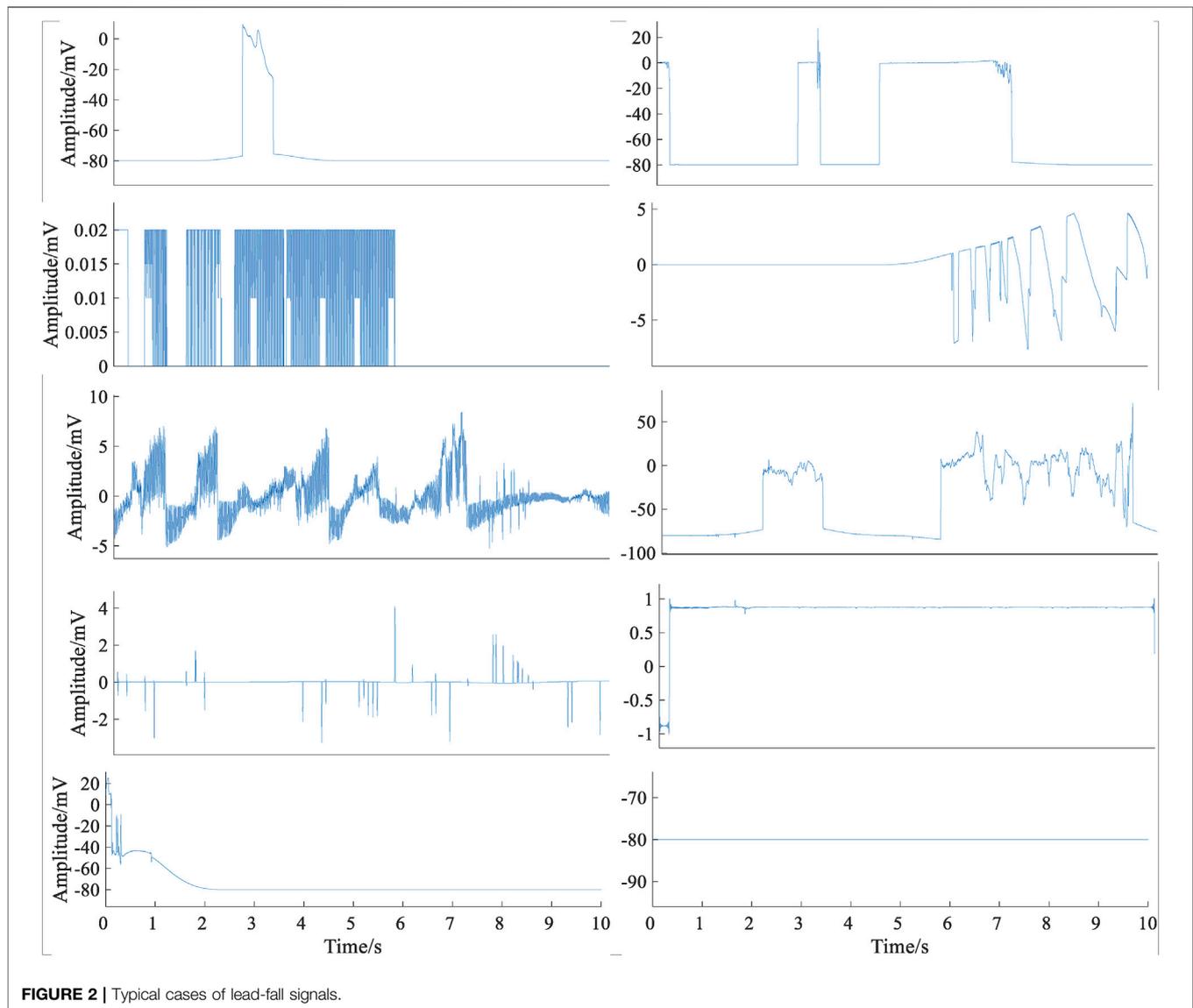


FIGURE 2 | Typical cases of lead-fall signals.

displayed the three-layer wavelet scattering network, and three classes of signals.

Figure 4(A) shows that time-domain plots of the scale function ϕ_I and Morlet wavelet function ψ_λ were employed in the study. Invariance scale I in the scale function needs to be confirmed based on the length of data and sampling frequency. A total of 41 first-order wavelet functions and 7 second-order wavelet functions were used to build this wavelet scattering network, as shown in Figures 4 (C) and (D). Also, Figure 4 (B) described the Littlewood–Paley sums for these scattering filter banks. Wavelet scattering networks could automatically extract feature extraction and could also reduce the signal dimension. The scattering characteristic matrix with dimension 81 – by – 20 was generated by this wavelet scattering network for one ECG signal with a length of 2,500 samples. For scattering coefficients of order 0, an input signal was first averaged using the scale function, which was the

first matrix 1×20 . For scattering coefficients of order 1, performing a continuous wavelet on the input signal yield a set of scalogram coefficients. Also, a modulus was applied to these coefficients and then the outputs were filtered with the wavelet low-pass filter yielding a set of order-1 scattering coefficients. It was the second matrix 41×20 . For scattering coefficients of order 2, the same process was applied to the scalogram coefficients to obtain the third matrix 39×20 . These three matrixes formed a scattering characteristic matrix with a dimension of 81×20 . The columns (20) can be considered as the time dimension and 81 can be considered as the scale dimension. But this time dimension was after processed by average operation. Also, this scale dimension was also after processed by nonlinear modulus and averaging operators. It was different from the time-frequency map generated by the wavelet transform. A long short-term memory (LSTM) classifier with ADAM solver was used for classification.

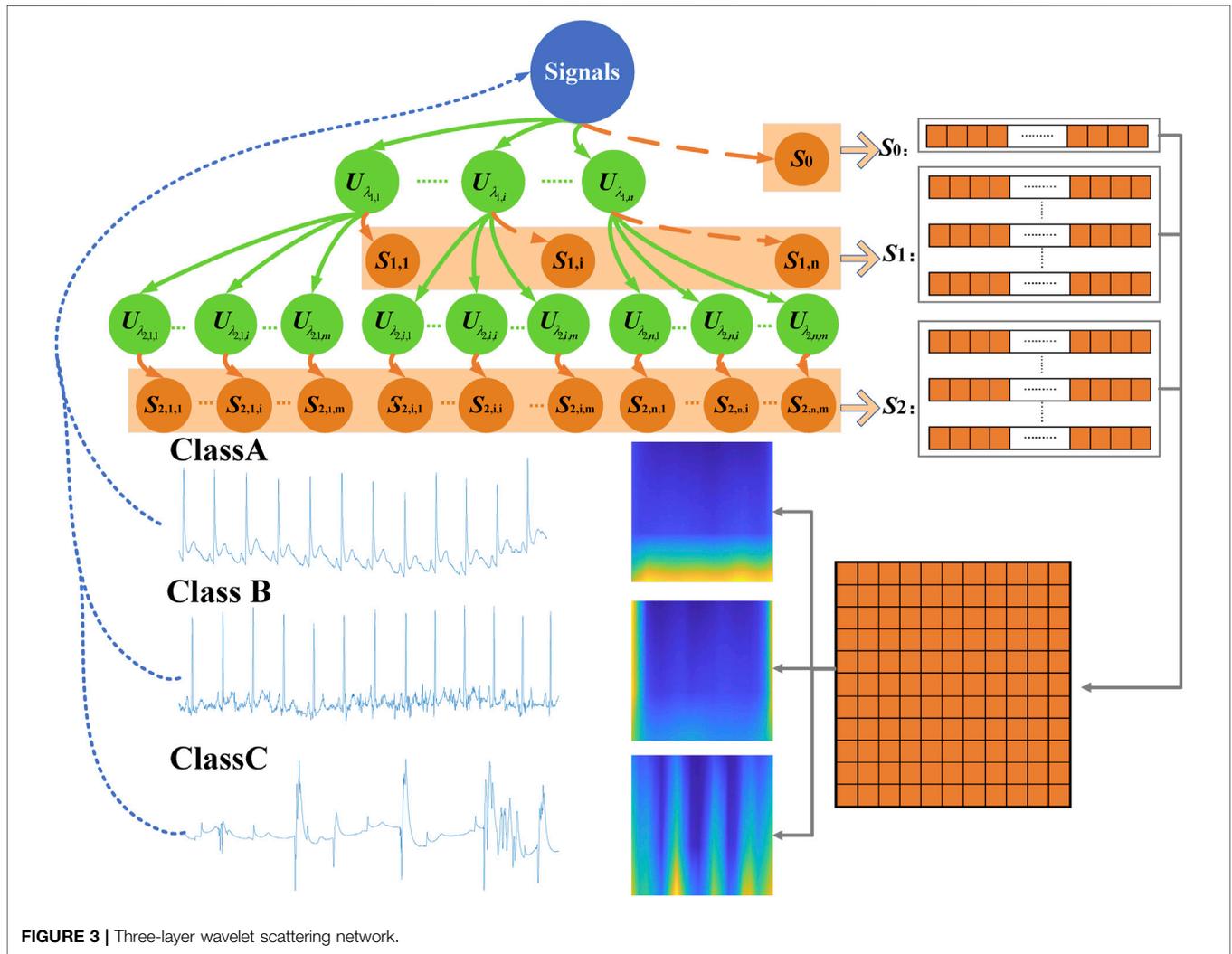


FIGURE 3 | Three-layer wavelet scattering network.

EVALUATION METHOD

The evaluation indexes used in this study are sensitivity (Se), precision ($+P$), comprehensive index F_1 measure for A\B\C three quality grades, and modified accuracy ($mACC$). Se is the proportion of a certain class that has correctly predicted the total number of all real classes in the test dataset, including Se_A, Se_B, Se_C . $+P$ is the proportion of the certain class that has correctly predicted the total number of predicted to be this class in the test set, including $+P_A, +P_B, +P_C$. F_1 measure includes F_{1A}, F_{1B}, F_{1C} for class A/B/C, respectively. mF_1 is an average value of these three indexes.

$$F_{1A} = \frac{2 \times TN_A}{N_A + T_A}, F_{1B} = \frac{2 \times TN_B}{N_B + T_B}, F_{1C} = \frac{2 \times TN_C}{N_C + T_C} \quad (8)$$

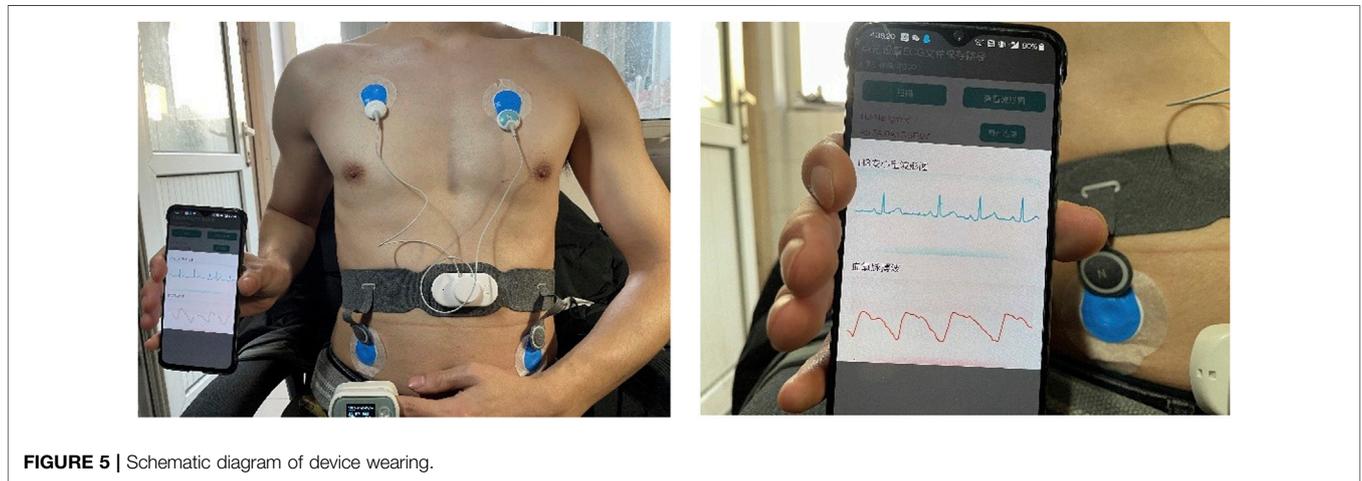
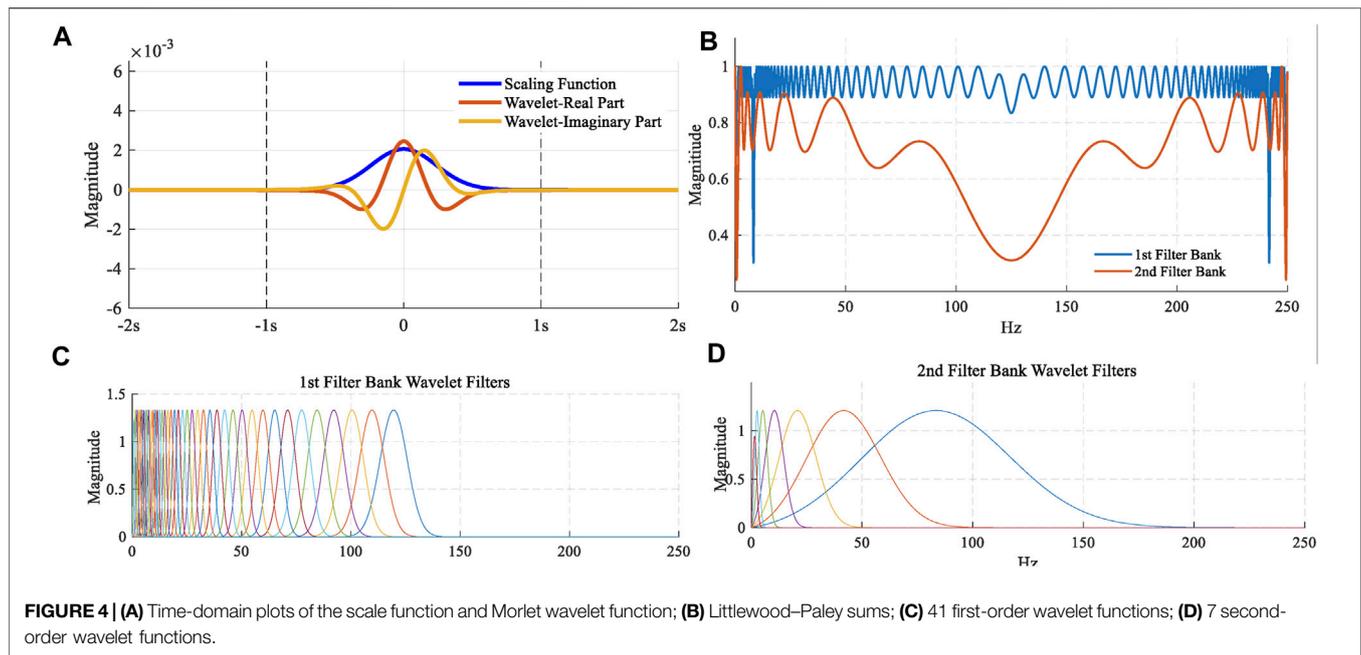
$$mF_1 = \frac{1}{3} (F_{1A} + F_{1B} + F_{1C}), \quad (9)$$

$$mACC = \frac{TN_A + TN_B + TN_C}{N \text{ (the number of all the samples in the test set)}} \quad (10)$$

where $TN_A, TN_B,$ and TN_C are the number of signals accurately predicted as classes A, B, and C, respectively. $N_A, N_B,$ and N_C are the number of all real class A, B, and C signals in the test set, respectively. $T_A, T_B,$ and T_C are the number of all predicted to be class A, B, and C signals in the test set, respectively.

Classification Model

The total number of signals in the database was 50,085. By pre-assessment, 1,029 lead-off and pure noise signals were eliminated. Remaining 49,056 signals were used to study the classification performance of the wavelet scattering network. In this study, a bi-directional long short-term memory (Bi-LSTM) network with the adaptive moment estimation (ADAM) solver was employed to train the classification model. The maximum number of epochs was 1,000. To reduce the amount of padding in the mini-batches, choose a mini-batch size of 490. The 10-fold cross-validation was employed to evaluate the classification performance of the model. All the segments were randomly divided into 10 groups. Also, the number of signals for each fold was 4,905.



Real-Time Validation

For real-time SQA performance analysis, the reliability of the whole wavelet scattering network method on real wearable dynamic ECG signals was tested. A Lenovo H3 dynamic ECG device was used for this experiment. **Figure 5** shows the schematic diagram of device wearing. There were 60 subjects (24 females, 36 males) aged between 19 and 24 years in this experiment. For each subject, a 10-min duration of dynamic ECG signals was recorded under different physical activity conditions. The subjects wearing Lenovo H3 dynamic ECG devices were asked to perform different activities for 10-min duration, including sitting, walking, jogging, sitting, running, sitting, jumping, and sitting. In order to eliminate the interaction between two different physical activities, the subject was asked to sit and rest after strenuous exercise, such as running and jumping. The continuous wearable ECG signals collected by the

Lenovo H3 device were transmitted to the phone via Bluetooth. The ECG signal was segmented with a frame length of 10 s and a hop-size of 1 s. The proposed wavelet scattering network SQA classifier evaluated the quality of the whole signal. The scattering characteristic matrix of the segment signals was generated by the proposed three-layer wavelet scattering network. A classification model was trained by all the signals (a total of 49,056) in the constructed database. By this classifier, the segmented signal was classified into different quality levels.

RESULT

Tables 3 shows the confusion matrices of the classification results for an independent test set, and **table 4** displays 11 evaluation indexes ($mACC$, mF_1 , Se_A , Se_B , Se_C , $+P_A$, $+P_B$, $+P_C$, F_{1A} , F_{1B} , F_{1C})

TABLE 3 | Confusion matrices of the classification results for an independent test set.

Confusion matrix	Pred			Sensitivity (Se)	
	A	B	C		
Actual	A	1,578	37	1	97.65%
	B	23	1,597	8	98.10%
	C	2	8	1,645	99.40%
Precision (+P)	98.44%	97.26%	99.46%	98.39%	

of the average classification results for 10 cross-validation. As it turns out, the classification performance of the wavelet scattering network classifier was very great. The mean values of 11 evaluation indexes were nearly all greater than 98% ($mACC = 98.56\%$, $mF_1 = 98.55\%$, $Se_A = 97.90\%$, $Se_B = 98.16\%$, $Se_C = 99.60\%$, $+P_A = 98.52\%$, $+P_B = 97.60\%$, $+P_C = 99.54\%$, $F_{1A} = 98.20\%$, $F_{1B} = 97.90\%$, $F_{1C} = 99.60\%$), only 97.90% for Se_A and F_{1B} , 97.6% for $+P_B$. Particularly for class C, Se_C , $+P_C$, F_{1C} were all greater than 99%.

For classification performance comparisons, 27 typical SQA methods were selected (picaSQI (Li et al., 2014), tSQI (Liu F. et al., 2020), kSQI (Clifford et al., 2012), ELZ_compl_SQI (Zhang et al., 2014), sSQI (Clifford et al., 2012), bSQI_4 (Liu et al., 2018), DisEn_SQI (Li et al., 2015), pSQI (Li et al., 2008), bsSQI (Li et al., 2014), iSQI (Liu et al., 2019b), basSQI (Li et al., 2014), ApEn_SQI (Pincus et al., 1991), bSQI_2 (Li et al., 2008), HpSQI (Liu F. et al., 2020), FuzzyEn_SQI (Di Marco et al., 2012), SampEn_SQI (Chen et al., 2009), SDN_SQI (Everss-Villalba et al., 2017), eSQI (Li et al., 2014), MSQI (Tobon Vallejo et al., 2014), MpSQI, LpSQI (Liu F. et al., 2020), rsdSQI (Li et al., 2014), purSQI (Nemati et al., 2010), pcaSQI (Behar et al., 2013), PLI_SQI (Everss-Villalba et al., 2017), LZ_compl_SQI (Zhang et al., 2016), and hfSQI (Li et al., 2014)). These SQA methods were mainly based on the SQI indexes extracted from the time domain, frequency domain features, QRS waves, nonlinear characteristic, and others. The support vector machine (SVM) classifier was employed to train the classification model. As shown in **Table 4**, the classification accuracy of multi SQIs for 10-fold cross-validation was 85.33%.

Considering the overfitting influence of deep learning, cross-database validation was carried out to verify the generalization ability of this proposed method. All signals from BUTQDB were used as training data, while all signals from the 2011 PhysioNet/CinC were used as testing data, and vice versa, all signals from the 2011 PhysioNet/CinC were used as training data, while all signals from BUTQDB were used as testing data. **Table 4** also displays the results of cross-database validation. For these two classifiers, classification accuracies were all greater than 80%. It was not as good as 10-fold cross-validation. The classification accuracy of multi SQIs for cross-database validation was about 75%.

DISCUSSION

In this study, we proposed a new SQA classification method based on a three-layer wavelet scattering network and built a wearable

ECG quality database with 50,085 recordings for A/B/C three quality levels. The proposed SQA classifier had an excellent performance on this database ($mACC = 98.56\%$, $mF_1 = 98.55\%$) for 10 cross-validation after all signals mixing. Particularly for class C signals, the proposed approach worked very well and the evaluation indexes were all greater than 99%. For class A and B signals, the results were slightly worse, but all greater than 97%. The wavelet scattering network used the deep convolution network architecture, but filter parameters were predefined. In this study, only the influence of the invariance scale was considered. Meanwhile, for performance comparisons, 27 typical SQA methods were selected to test the performance of this new database. Considering the overfitting influence of deep learning, cross-database validation and real-time validation were also carried out. The classification performance of cross-database validation was also admissible ($mACC$, $mF_1 \geq 80\%$).

Influence of Invariance Scale

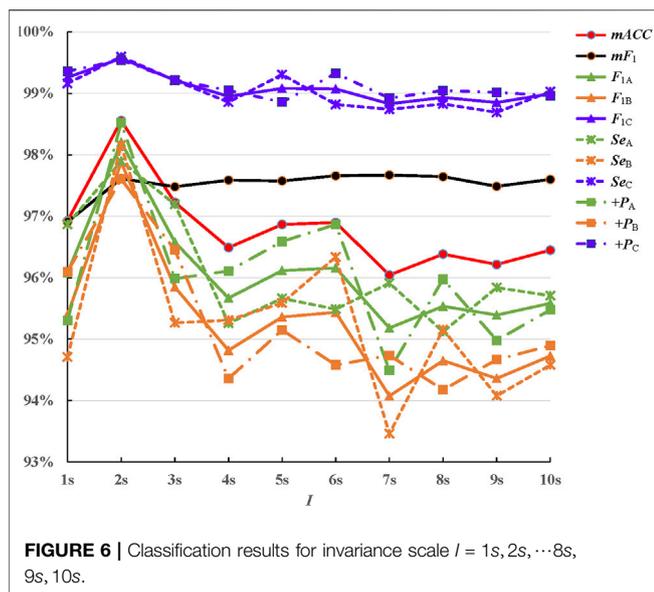
In this study, the proposed three-layer wavelet scattering network was a deep learning framework which could extract complementary compact information automatically. The wavelet scattering network used the deep convolutional network architecture iterates over wavelet convolution, nonlinear modulus, and averaging (pooling) operators to compute higher-order scattering coefficients, which build translation invariant, stable, and informative signal representations. But the filters of the wavelet scattering network were predefined Morlet wavelets (Bruna and Mallat, 2013), which did not need to be learned from data. The Morlet wavelets were a localized waveform, having a better frequency resolution and stability to deformations, which could impose the separation of the different quality signals. The nonlinear modulus propagator recombines real and imaginary parts of complex wavelet coefficients, which could keep the low frequency averaging and obtain the translation invariant representation. Although the modulus operator removed the complex phase and lost information about the high frequencies, it kept the temporal variation of the multiscale envelopes. Also, the high frequencies information lost by the pooling can be recovered as wavelet coefficients in the next layers as the wavelet transform was a redundant representation. High order scattering coefficients could characterize transient phenomena of the different noises from free living. To recover this high-frequency information, a new wavelet transform was implemented to the signal in the next layers before the nonlinear modulus and pooling were performed.

The invariance scale was also termed as the interval of time-shift invariance, which was defined by the size of the time averaging window. The influence of this parameter was also considered in this study. Also, the scattering coefficients were computed at scales $I = 1s, 2s, \dots, 8s, 9s, 10s$. **Figure 6** displayed the classification results. As shown in **Figure 6**, the invariance scale I changing had less influence on the accuracy. All 11 evaluating indexes were above 93%. But obviously, when the invariance scale was set to be 2s, the classification performance was best. The scale I controlled the amount of translation invariance. When it was too small, noises produced by gross movements, such as severe drifting baselines, would miss some. When it was too large, the

TABLE 4 | Classified results of 27 SQL-based SVM classifiers and proposed method.

SQA method		$Se_A/\%$	$Se_B/\%$	$Se_C/\%$	$+P_A/\%$	$+P_B/\%$	$+P_C/\%$	$F_{1A}/\%$	$F_{1B}/\%$	$F_{1C}/\%$	$mF_1/\%$	$mAcc/\%$
27 SQLs + SVM	10-fold cross-validation	87.79 ± 0.97	72.75 ± 1.47	96.06 ± 0.51	77.9 ± 1.09	83.17 ± 0.76	96.2 ± 0.46	82.52 ± 0.69	77.6 ± 0.69	96.1 ± 0.53	85.4 ± 0.44	85.33 ± 0.34
	Cross database validation, BUTQDB signals as training data, 2011 PhysioNet/CinC signals as testing data	61.06 ± 0.45	74.26 ± 0.61	90.74 ± 0.39	65.46 ± 0.87	68.47 ± 0.44	94.38 ± 0.56	63.18 ± 1.02	71.25 ± 0.98	92.53 ± 0.65	75.65 ± 0.33	75.06 ± 0.25
	Cross database validation, 2011 PhysioNet/CinC signals as training data, BUTQDB signals as testing data	63.48 ± 0.69	65.75 ± 0.94	98.03 ± 0.26	75.22 ± 0.47	62.06 ± 0.68	90.62 ± 0.67	68.85 ± 0.81	63.85 ± 0.78	94.18 ± 0.31	75.63 ± 0.84	77.65 ± 0.41
Wavelet scattering + LSTM	10-fold cross validation	97.90 ± 0.54	98.16 ± 0.58	99.60 ± 0.39	98.52 ± 0.81	97.60 ± 0.94	99.54 ± 0.16	98.20 ± 0.85	97.90 ± 0.84	99.60 ± 0.29	98.55 ± 0.40	98.56 ± 0.39
	Cross database validation, BUTQDB signals as training data, 2011 PhysioNet/CinC signals as testing data	79.26 ± 1.50	83.32 ± 0.84	92.33 ± 0.67	80.38 ± 0.77	82.17 ± 1.39	92.82 ± 0.92	79.81 ± 1.01	82.74 ± 1.39	92.58 ± 0.87	85.04 ± 1.48	85.32 ± 0.99
	Cross database validation, 2011 PhysioNet/CinC signals as training data, BUTQDB signals as testing data	79.93 ± 1.03	75.68 ± 0.96	91.59 ± 0.84	84.44 ± 1.01	72.43 ± 0.99	90.11 ± 0.73	82.12 ± 0.88	74.02 ± 1.27	90.85 ± 0.65	82.32 ± 1.23	82.73 ± 1.11

The bold values were the results of the proposed method.



convolution would lose partly high frequency information. In this study, 2s was the best choice for the invariance scale. The classification performances showed small differences in the changing of the invariance scale. The variability within each class A/B/C was not due to translation, but due to time-domain deformations and spectrum noise.

Comparing With Other SQA Methods

At present, most studies about the signal-quality assessment divided the ECG signals into acceptable and unacceptable. There are fewer public databases with three quality levels. For performance comparisons, 27 typical SQA methods were selected to test the performance of this new database. These methods mostly had better performance on the database including two classes of ECG signals. Multiple SQI feature-based classifiers were lower than the proposed novel classification method. In order to analyze the performances of these SQIs better, Figure 7 displays the distribution of these SQIs on the A/B/C quality levels signals. Green, orange, and blue dots represent class A/B/C signals, respectively. The SQIs, which are only based on the QRS waves, such as bSQI-4 and bSQI-2, were defined by the comparison of four or two QRS wave detectors on a single-lead signal. They had good performance on the database with two classes of ECG signals ($mACC > 93\%$) (Liu et al., 2018), but class A signals are mixed up with class B signals, as shown in Figure 7. It is because that QRS wave of class B signals also could be detected accurately. The performance of tSQI and picaSQI was slightly better. These two SQIs were computed not only based on the QRS wave but also based on morphology consistency and nonlinear characteristic (Li et al., 2014; Liu F. et al., 2020). The tSQI was defined as the morphology consistency of any two ECG beats within a fixed time window (Li et al., 2014), and the picaSQI was defined as a periodicity measure of the ECG waveform nonlinear characteristic (Liu F. et al., 2020). For other SQIs based on time and frequency domain features, the distribution ranges for A/B/C quality level signals had a large overlap region.

The class B signals were mostly contaminated by the noise with high frequency and low amplitude, which can make partly class B signals detected to be class A or C.

It should be noted that all these 27 SQIs that we selected were unlikely to be the optimal indexes. We tried to pick as many quality metrics as possible, but it is impossible to pick all of them. Meanwhile, because some SQIs were published in a theoretical way without the executable program, and some literature works lacked detailed necessary preprocessing operations, some SQIs were coded by us. Thus, the classing results in this study could be different from those in the other studies, but the differences are unimportant.

In this study, all the programs were implemented using MATLAB 2020a. Table 5 illustrates the mean time costs and standard deviation values of the 12 SQIs and the proposed method by analyzing 49,056 10-s ECG segments in the database. As shown in Table 5, the proposed method was the most time-efficient compared with 12 SQIs. Also, 18.25 ms was not a long-time cost for 10-s ECG segments.

Real-Time Validation and Cross-Database Validation Analysis

For real-time SQA performance analysis, the reliability of the whole wavelet scattering networks method on real wearable ECG signals was tested. Figure 8 displays two segments of 10-min duration of dynamical ECG signals, physical activities, and evaluation results. As it turns out, under the sitting and walking conditions, the quality of the ECG signal was very good and all signals are assessed as class A, which can be used for the cardiovascular disease diagnosis. In the jogging condition, some signals were contaminated by weak artifacts and assessed as class B. But they could not affect the R wave identification, which can be used for the heart rate measure. In the running and jumping conditions, most of the signals were contaminated by seriously large noises caused by violent motion and assessed as class C. These signals will interfere with the diagnosis of the cardiovascular disease and need to be removed. Also, the proposed SQA method could identify the changing of the position. During the changing stage, there were some fluctuations in the signal. These signals were assessed as class B. The evaluation results show that the proposed wavelet scattering network SQA classifier framework has capability to assess wearable dynamic ECG signal quality.

In this study, cross-database testing was also carried out to verify the generalization ability of this proposed method. As the results show, classification accuracy was greater than 80%. The performance of cross-database validation was not good as 10-fold cross-validation. But for class C, Se_C , $+P_C$, F_{1C} were all greater than 90%. The reason for this phenomenon is the great difference between these two databases. The signals from BUTQDB were single-lead ECGs with three quality classes, while the signals from the 2011 PhysioNet/CinC Challenge were 12-lead ECGs with two quality classes. Although we, based on the scores in Liu et al. (2018), annotated all the leads individually, there are still differences in the re-annotating. The morphological characteristics of class C signals are obvious, and the evaluation of experts is relatively consistent. However, the difference in



FIGURE 7 | Distribution of 27 SQIs on the A/B/C quality levels signals. Green, orange, and blue dots represent class A/B/C signals, respectively.

morphological characteristics between class A and class B signals is not particularly obvious, the evaluations between different experts are different. If all the signals from the 2011 PhysioNet/CinC

Challenge were re-annotated based on the criteria of BUTQDB strictly, the performance of cross-database validation will be better. However, it will need more time cost.

TABLE 5 | Mean time costs and standard deviation values of 12 SQIs and proposed method.

Method	Mean time/ms	SD/ms
picaSQI (Li et al., 2014)	4.04	0.41
tSQI (Liu et al., 2020a)	1.20	0.20
kSQI (Clifford et al., 2012)	1.14	0.13
ELZ_comp_SQI (Zhang et al., 2014)	31.74	6.55
sSQI (Clifford et al., 2012)	1.05	0.23
bSQI_4 (Liu et al., 2018)	6.83	0.65
DisEn_SQI (Li et al., 2015)	25.28	8.31
pSQI (Li et al., 2008)	2.15	0.25
bsSQI (Li et al., 2014)	2.01	0.30
iSQI (Liu et al., 2019b)	1.08	0.26
basSQI (Li et al., 2014)	3.31	0.43
ApEn_SQI (Pincus et al., 1991)	30.41	3.60
Proposed method	18.25	1.38

The bold values were the time cost and standard deviation of the proposed method.

In this study, the approach adding noise to clean data was used to balance the database classes. Also, those clear recordings from classes A and B used to upsample class C, were put back into the original classes. It needs to be considered if this synthetic noise is going to play a role in the classification results. If one clear recording from class A is in the training set and it was also corrupted by noises used to upsample the class C in the test set. If this recording includes some information on the test set, this information will tend to judge this recording as class A. However, it was a contaminated recording and labeled as class C. Therefore, if there is some information generated by balancing the dataset on the test set, classification accuracy will be reduced. However, the classification performance of this method was very good. Therefore, the approach used to balance the dataset did not influence the classification results. The testing of real signals without synthetic addition was also carried out to consider the influence of this balancing data approach. A new database was built with 5,687 recordings without any synthetic noise added,

class A: 2,000, class B: 2,000, and class C: 1,687. We used 30% of the data for testing and the remaining 70% of the data for training. The values of 11 evaluation indexes were nearly all greater than 90%, $mACC = 94.05\%$, $mF_1 = 94.03\%$, $Se_A = 93.35\%$, $Se_B = 91.25\%$, $Se_C = 97.55\%$, $+P_A = 94.44\%$, $+P_B = 92.73\%$, $+P_C = 94.94\%$, $F_{1A} = 93.89\%$, $F_{1B} = 91.98\%$, and $F_{1C} = 96.23\%$. The reduction in the data volume reduced the accuracy of the model, which was also acceptable. The approach of adding noise to clean data to balance the database classes was also used in Clifford et al. (2012).

Most notably, there is no unified evaluation criterion to determine the quality levels of wearable ECG at present. Different databases provide different evaluation methods. For example, the data in the 2011 PhysioNet/CinC Challenge are 12-lead recordings, having a length of 10 s 3–18 annotators marking each signal, and each record was assigned to one of the three groups (acceptable 773, indeterminate 2, and unacceptable 225) based on the average score. Some studies considered that the label of “acceptable” or “unacceptable” was for the whole 12 channels, not for the single channel. Therefore, they re-labeled each channel (Clifford et al., 2012; Liu et al., 2018), and balanced the classes by adding noise to some of the clean data. However, in the BUTQDB database, 18 single-lead signals longer than 24 h were recorded using the Bittium Faros 180 device. The parts of signals were selected to be grouped into three quality levels based on the labels annotated by three experts. Also, some studies constructed a manually annotated gold standard, collected and annotated ECG recordings by themselves (Redmond et al., 2012; Satija et al., 2017; Liu et al., 2019b; Smital et al., 2020). Different classification standards and annotating methods could have great influence on the SQA performance.

Limitations and Prospects

Wearable electrocardiogram quality assessment is quite crucial for cardiovascular disease prevention and diagnosis. It is also an

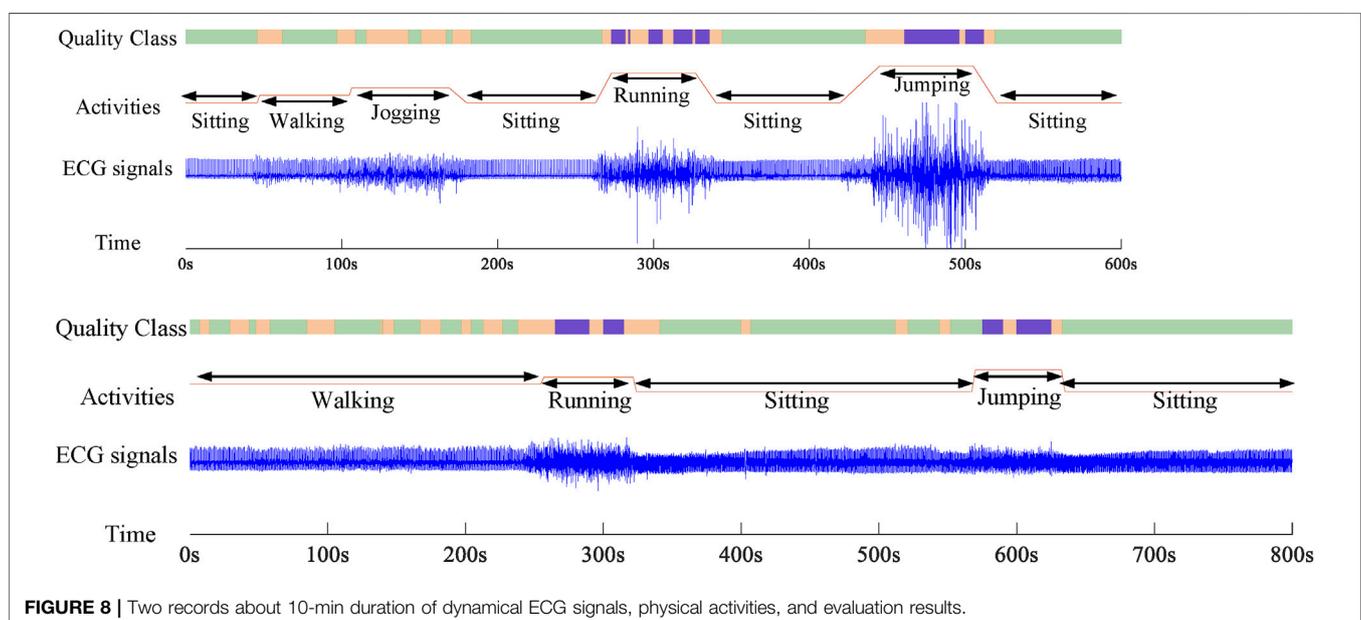


FIGURE 8 | Two records about 10-min duration of dynamical ECG signals, physical activities, and evaluation results.

important issue for wearable device development. Although the proposed new method had great performance on the quality assessment, it was not very well for cross-database validation. The main reason is the difference between annotation methods and classification grades. For future work, uniform and standardized evaluation criterion is quite crucial for the wearable ECG quality assessment.

CONCLUSION

This study aimed to provide a method to classify wearable dynamic ECG signals into three grades: high quality (A), medium quality (B), and low quality (C). A new SQA classification method based on a three-layer wavelet scattering network and transfer learning LSTM was proposed, and a wearable ECG quality database with 50,085 recordings for three quality grades was built. In order to avoid the adverse impact of invalid samples on the training models, the quality pre-assessment was used to delete the lead-fall signals and pure noise. A three-layer wavelet scattering network was performed on the selected 10-s-long signal segments, which can extract more systematic and comprehensive characteristics by analyzing the signals thoroughly and deeply. The Bi-LSTM network with ADAM solver was employed to train the classification model. The 11 evaluating indexes ($mACC$, mF_1 , Se_A , Se_B , Se_C , $+P_A$, $+P_B$, $+P_C$, F_{1A} , F_{1B} , F_{1C}) were 98.56%, 98.55%, 98.52%, 97.60%, 99.54%, 98.20%, 97.90%, 99.60%, 97.90%, 98.16%, and 99.60%, respectively, suggesting that the proposed method can effectively separate three quality grades of wearable ECG signals. For efficacy validation, this method was applied on the real-world data collected using the Lenovo H3 dynamic ECG device. This method had the ability to detect noise signals produced by vigorous activities. With the high computational efficiency, it will have a good application on wearable ECG devices, including removing contaminating signals and selecting high-quality signal segments for CVD diagnosis and analysis. This study verified the feasibility of applying the wavelet scattering network model to wearable

REFERENCES

- Anden, J., and Mallat, S. (2014). Deep Scattering Spectrum. *J IEEE Trans. Signal Process. A Publ. IEEE Signal Process. Soc.* 62 (16), 4114–4128.
- Andrea, N., Radovan, S., Lucie, M., Luká, S., and Martin, V. J. B. R. I. (2018). A Comparative Analysis of Methods for Evaluation of ECG Signal Quality after Compression. *Biomed. Res. Int.* 2018, 1868519. doi:10.1155/2018/1868519
- Behar, J., Oster, J., Qiao Li, Q., and Clifford, G. D. (2013). ECG Signal Quality during Arrhythmia and its Application to False Alarm Reduction. *IEEE Trans. Biomed. Eng.* 60 (6), 1660–1666. doi:10.1109/tbme.2013.2240452
- Bruna, J., and Mallat, S. (2013). Invariant Scattering Convolution Networks. *IEEE Trans. Pattern Anal. Mach. Intell.* 35 (8), 1872–1886. doi:10.1109/tpami.2012.230
- Chen, W., Zhuang, J., Yu, W., and Wang, Z. (2009). Measuring Complexity Using FuzzyEn, ApEn, and SampEn. *Med. Eng. Phys.* 31 (1), 61–68. doi:10.1016/j.medengphys.2008.04.005
- Clifford, G. D., Behar, J., Li, Q., and Rezek, I. (2012). Signal Quality Indices and Data Fusion for Determining Clinical Acceptability of Electrocardiograms. *Physiol. Meas.* 33 (9), 1419–1433. doi:10.1088/0967-3334/33/9/1419

ECG signal-quality assessment. Also, the general framework of this classification method proposed in this study was sufficiently flexible to be used in any given situation.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusion of this article will be made available by the authors, without undue reservation.

AUTHOR CONTRIBUTIONS

FL, CL, and SX designed the research. FL designed the algorithm and analyzed the data. SW, LC, and YR contributed to signal relabeling. XR, ZX, and SA, contributed to collecting and sorting data. FL wrote the manuscript. CL contributed to revising it critically for the important intellectual content. All authors have read and approved of the version to be published.

FUNDING

This work was supported by the National Natural Science Foundation of China (grant numbers 61901114, 82072014, 81871444, 62171123), Shandong Province Key Basic Research program (grant numbers ZR2020ZD25), the National Key R&D Program of China (grant numbers 2019YFE0113800), and the Shandong Province Natural Science Foundation (grant numbers ZR2020MF028).

ACKNOWLEDGMENTS

The authors would like to thank the cardiology team at the Department of Biomedical Engineering, Brno University of Technology and the 2011 PhysioNet/CinC Challenge for providing the ECG Quality Database.

- Clifford, G. D., and Moody, G. B. (2012). Signal Quality in Cardiorespiratory Monitoring. *Physiol. Meas.* 33 (9), 1–6. doi:10.1088/0967-3334/33/9/e01
- Destouet, G., Dumas, C., Frassati, A., and Perrier, V. (2021). *Wavelet Scattering Transform and Ensemble Methods for Side-Channel Analysis*. Cham: Springer International Publishing.
- Di Marco, L. Y., Duan, W., Bojarnejad, M., Zheng, D., King, S., Murray, A., et al. (2012). Evaluation of an Algorithm Based on Single-Condition Decision Rules for Binary Classification of 12-lead Ambulatory ECG Recording Quality. *Physiol. Meas.* 33 (9), 1435–1448. doi:10.1088/0967-3334/33/9/1435
- Everss-Villalba, E., Melgarejo-Meseguer, F., Blanco-Velasco, M., Gimeno-Blanes, F., Sala-Pla, S., Rojo-Álvarez, J., et al. (2017). Noise Maps for Quantitative and Clinical Severity towards Long-Term ECG Monitoring. *Sensors* 17 (11), 2448. doi:10.3390/s17112448
- Fousek, P., Goel, V., Kingsbury, B., Marcheret, E., Maymon, S., Nahamoo, D., et al. (2015). Deep Scattering Spectrum in Acoustic Modeling for Speech Recognition.
- Goldberger, A. L., Amaral, L. A. N., Glass, L., Hausdorff, J. M., Ivanov, P. C., Mark, R. G., et al. (2000). PhysioBank, PhysioToolkit, and PhysioNet. *Circulation* 101 (23), e215–e220. doi:10.1161/01.CIR.101.23.e215

- He, J., Liu, D., and Chen, X. (2020). Wearable Exercise Electrocardiograph Signal Quality Assessment Based on Fuzzy Comprehensive Evaluation Algorithm. *Comput. Commun.* 151, 86–97. doi:10.1016/j.comcom.2019.12.051
- Huerta, A., Martínez-Rodrigo, A., González, V., Quesada, A., Rieta, J. J., and Alcaraz, R. (2019). “Quality Assessment of Very Long-Term ECG Recordings Using a Convolutional Neural Network,” in 2019 E-Health and Bioengineering Conference (EHB), Iasi, Romania, 21–23 November 2019.
- Ip, J. E. (2019). Wearable Devices for Cardiac Rhythm Diagnosis and Management. *JAMA* 321 (4), 337–338. doi:10.1001/jama.2018.20437
- Joy, N. M., Oglic, D., Cvetkovic, Z., Bell, P., and Renals, S. (2020). “Deep Scattering Power Spectrum Features for Robust Speech Recognition,” in Interspeech 2020, Shanghai, China, October 25–29, 2020. doi:10.21437/interspeech.2020-2656
- Kaptoke, S., Pennells, L., De Bacquer, D., Cooney, M. T., Kavousi, M., Stevens, G., et al. (2019). World Health Organization Cardiovascular Disease Risk Charts: Revised Models to Estimate Risk in 21 Global Regions. *Lancet Glob. Health* 7 (10), e1332–e1345. doi:10.1016/S2214-109X(19)30318-3
- Li, P., Liu, C., Li, K., Zheng, D., Liu, C., Hou, Y., et al. (2015). Assessing the Complexity of Short-Term Heartbeat Interval Series by Distribution Entropy. *Med. Biol. Eng. Comput.* 53 (1), 77–87. doi:10.1007/s11517-014-1216-0
- Li, Q., Mark, R. G., and Clifford, G. D. (2008). Robust Heart Rate Estimation from Multiple Asynchronous Noisy Sources Using Signal Quality Indices and a Kalman Filter. *Physiol. Meas.* 29 (1), 15–32. doi:10.1088/0967-3334/29/1/002
- Li, Q., Rajagopalan, C., and Clifford, G. D. (2014). A Machine Learning Approach to Multi-Level ECG Signal Quality Classification. *Comput. Methods Programs Biomed.* 117 (3), 435–447. doi:10.1016/j.cmpb.2014.09.002
- Li, Z., Hou, Y., Xie, X., Li, S., and Liu, W. (2019). “Multi-level Attention Model with Deep Scattering Spectrum for Acoustic Scene Classification,” in 2019 IEEE International Conference on Multimedia & Expo Workshops (ICMEW), Shanghai, China, 08–12 July 2019.
- Liu, C., Li, P., Zhao, L., and Liu, F. (2011). Real-time Signal Quality Assessment for ECGs Collected Using Mobile Phones. *Computing in Cardiology. Hangzhou, China*, 38, 357–360.
- Liu, C., Yang, M., Di, J., Xing, Y., Li, Y., and Li, J. (2019a). Wearable ECG: History, Key Technologies and Future Challenges. *Chin. J. Biomed. Eng.* 38 (6), 641–652. doi:10.3969/j.issn.0258-8021.2019.06.001
- Liu, C., Zhang, X., Zhao, L., Liu, F., Chen, X., Yao, Y., et al. (2019b). Signal Quality Assessment and Lightweight QRS Detection for Wearable ECG SmartVest System. *IEEE Internet Things J.* 6, 1363–1374. doi:10.1109/JIOT.2018.2844090
- Liu, F., Liu, C., Zhao, L., Jiang, X., Zhang, Z., Li, J., et al. (2018). Dynamic ECG Signal Quality Evaluation Based on the Generalized bSQI Index. *IEEE Access* 6, 41892–41902. doi:10.1109/ACCESS.2018.2860056
- Liu, F., Wei, S., Lin, F., Jiang, X., and Liu, C. (2020). “An Overview of Signal Quality Indices on Dynamic ECG Signal Quality Assessment,” in *Feature Engineering and Computational Intelligence in ECG Monitoring*. Editors C. Liu. and J. Li (Singapore: Springer), 33–54. doi:10.1007/978-981-15-3824-7_3
- Liu, Z., Yao, G., Zhang, Q., Zhang, J., and Zeng, X. (2020). Wavelet Scattering Transform for ECG Beat Classification. *J Comput. Math. Methods Med.* 2020 (7), 3215681. doi:10.1155/2020/3215681
- Mallat, S. (2012). Group Invariant Scattering. *Comm. Pure Appl. Math.* 65 (10), 1331–1398. doi:10.1002/cpa.21413
- Mei, N., Wang, H., Zhang, Y., Liu, F., Jiang, X., and Wei, S. (2021). Classification of Heart Sounds Based on Quality Assessment and Wavelet Scattering Transform. *Comput. Biol. Med.* 137, 104814. doi:10.1016/j.compbiomed.2021.104814
- Moody, G. B., Muldrow, W. K., and Mark, R. G. (1984). Noise Stress Test for Arrhythmia Detectors. *J Comput. Cardiol.* 11, 381–384.
- Nemati, S., Malhotra, A., and Clifford, D. C. (2010). Data Fusion for Improved Respiration Rate Estimation. *EURASIP J. Adv. Signal Process* 2010, 926305. doi:10.1155/2010/926305
- Nemcova, A., Smisek, R., Opravilová, K., Vitek, M., Smital, L., and Maršánová, L. (2020). Brno University of Technology ECG Quality Database (BUT QDB). *PhysioNet*. doi:10.13026/ka4-qw24
- Pincus, S. M., Gladstone, I. M., and Ehrenkrantz, R. A. (1991). A Regularity Statistic for Medical Data Analysis. *J. Clin. Monit. Comput.* 7 (4), 335–345. doi:10.1007/bf01619355
- Redmond, S. J., Xie, Y., Chang, D., Basilakis, J., and Lovell, N. H. (2012). Electrocardiogram Signal Quality Measures for Unsupervised Telehealth Environments. *Physiol. Meas.* 33 (9), 1517–1533. doi:10.1088/0967-3334/33/9/1517
- Roth, G. A., Abate, D., Abate, K. H., Abay, S. M., Abbafati, C., Abbasi, N., et al. (2018). Global, Regional, and National Age-sex-specific Mortality for 282 Causes of Death in 195 Countries and Territories, 1980–2017: a Systematic Analysis for the Global Burden of Disease Study 2017. *Lancet* 392, 1736–1788. doi:10.1016/S0140-6736(18)32203-7
- Satija, U., Ramkumar, B., and Manikandan, M. S. (2018). A Review of Signal Processing Techniques for Electrocardiogram Signal Quality Assessment. *IEEE Rev. Biomed. Eng.* 11, 36–52. doi:10.1109/RBME.2018.2810957
- Satija, U., Ramkumar, B., and Sabarimalai Manikandan, M. (2017). Real-Time Signal Quality-Aware ECG Telemetry System for IoT-Based Health Care Monitoring. *IEEE Internet Things J.* 4 (3), 815–823. doi:10.1109/jiot.2017.2670022
- Sepúlveda, A., Castillo, F., Palma, C., and Rodríguez-Fernández, M. (2021). Emotion Recognition from ECG Signals Using Wavelet Scattering and Machine Learning. *Appl. Sci.* 11 (11), 4945.
- Silva, I., Moody, G. B., and Celi, L. (2011). “Improving the Quality of ECGs Collected Using Mobile Phones: The PhysioNet/Computing in Cardiology Challenge 2011,” in 2011 Computing in Cardiology, Hangzhou, China, September 18–21, 2011, 273–276.
- Smital, L., Haider, C. R., Vitek, M., Leinveber, P., Jurak, P., Nemcova, A., et al. (2020). Real-Time Quality Assessment of Long-Term ECG Signals Recorded by Wearables in Free-Living Conditions. *IEEE Trans. Biomed. Eng.* 67 (10), 2721–2734. doi:10.1109/TBME.2020.2969719
- Tang, Y. Y., Lu, Y., and Yuan, H. (2015). Hyperspectral Image Classification Based on Three-Dimensional Scattering Wavelet Transform. *IEEE Trans. Geosci. Remote Sens.* 53 (5), 2467–2480. doi:10.1109/tgrs.2014.2360672
- Tobon Vallejo, D., Falk, T., and Maier, M. (2014). MS-QI: A Modulation Spectrum-Based ECG Quality Index for Telehealth Applications. *Ieee. Trans. Biomed. Eng.* 63 (8), 1613–1622. doi:10.1109/TBME.2014.2355135
- Xu, H., Yan, W., Lan, K., Ma, C., Wu, D., Wu, A., et al. (2021). Assessing Electrocardiogram and Respiratory Signal Quality of a Wearable Device (SensEcho): Semisupervised Machine Learning-Based Validation Study. *JMIR Mhealth Uhealth* 9 (8), e25415. doi:10.2196/25415
- Zhang, Y.-t., Liu, C.-y., Wei, S.-s., Wei, C.-z., and Liu, F.-f. (2014). ECG Quality Assessment Based on a Kernel Support Vector Machine and Genetic Algorithm with a Feature Matrix. *J. Zhejiang Univ. - Sci. C* 15 (7), 564–573. doi:10.1631/jzus.c1300264
- Zhang, Y., Wei, S., Liu, H., Zhao, L., and Liu, C. (2016). A Novel Encoding Lempel-Ziv Complexity Algorithm for Quantifying the Irregularity of Physiological Time Series. *Comput. Methods Programs Biomed.* 133, 7–15. doi:10.1016/j.cmpb.2016.05.010
- Zhao, Z., and Zhang, Y. (2018). SQI Quality Evaluation Mechanism of Single-Lead ECG Signal Based on Simple Heuristic Fusion and Fuzzy Comprehensive Evaluation. *Front. Physiol.* 9. doi:10.3389/fphys.2018.00727

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher’s Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors, and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Liu, Xia, Wei, Chen, Ren, Ren, Xu, Ai and Liu. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.