A Multi-Domain Feature Framework for Bone-conducted Phonocardiogram Signal Quality Assessment

Biao Wu[1], Yumin Li[1], Li Ling[1], Junjie Pan[2], Chaohong Liu[2], Huan Li[2], Chenghao Sui[2], Yanan Zhou[2], Wenkai Wang[2], Jianqing Li[1], Chengyu Liu[1]
[1]School of Instrument Science and Engineering, Southeast University, Nanjing, China
[2]Goertek Technology Co., Qingdao, China

Abstract

Bone-conducted phonocardiogram (BCPCG) recordings acquired in practical environments are often contaminated by motion artifacts and muscle tremor, making signal quality assessment essential for reliable downstream analysis. This study proposes a multi-domain handcrafted feature framework for BCPCG quality assessment based on raw signals and their envelope representations. A 25-dimensional feature set was designed to characterize signal quality from waveform morphology, energy variation, spectral structure, envelope peak organization, and rhythm consistency. Experiments on 24,581 labeled BCPCG segments were conducted under a three-class setting using five-fold stratified cross-validation. Among six evaluated machine-learning classifiers, LightGBM achieved the best overall performance, with an ACC of 0.914 and an AUC of 0.967. The proposed framework provides an effective and compact solution for BCPCG quality assessment.

Keywords: Bone-conducted phonocardiogram; signal quality assessment; multi-domain features

1. Introduction

Cardiac auscultation is a simple and cost-effective tool for cardiovascular assessment. Phonocardiogram (PCG) signals contain important information related to cardiac mechanical activity and therefore have potential for early screening and long-term wearable monitoring [1,2]. However, PCG recordings acquired in real-world settings are often degraded by motion artifacts, ambient noise, muscle interference, and sensor displacement, which can impair downstream analyses such as PCG segmentation, inter-beat interval estimation, and abnormality detection [3,4,6,8]. Therefore, signal quality assessment (SQA) is a necessary preprocessing step for robust PCG analysis.

Existing studies on PCG signal quality assessment can generally be divided into three categories. The first category includes traditional signal quality indices (SQIs), which use statistical descriptors to measure signal quality. Although these methods are efficient and easy to implement, their reliance on limited single-domain descriptors often leads to insufficient discrimination under complex noisy conditions [4]. The second category includes deep-learning-based approaches, which have shown promising results in PCG analysis and abnormality recognition [3,8]. These methods usually require larger annotated datasets and higher computational cost, which may limit their applicability in lightweight wearable device. The third category comprises machine-learning-based methods using handcrafted features by integrating information from temporal, spectral, and morphological domains [5,7]. For example, Das et al. introduced autocorrelation and spectral descriptors to capture periodic patterns in PCG signals [5].

The aforementioned SQA methods are not universally applicable, as they are constrained by device form factor and no unified quality assessment standard has been established across acquisition settings. In contrast, bone-conducted phonocardiogram (BCPCG) offers a promising solution for daily and wearable PCG monitoring, as it can naturally mitigate environmental noise interference while reducing the dependence on complex circuitry [10]. However, because of its acquisition mechanism, BCPCG signals are still susceptible to interference from muscle tremor and body motion, making quality assessment more challenging. To address this problem, Ling et al. proposed a multi-domain feature-based framework and demonstrated the effectiveness of feature fusion for PCG quality assessment [7]. Nevertheless, its performance and generalizability for BCPCG recordings remain limited.

This study proposes a multi-domain feature framework for BCPCG quality assessment, leveraging complementary information from both raw BCPCG signals and their envelope representations.

2. Materials and methods

2.1. Data Acquisition and Preprocessing

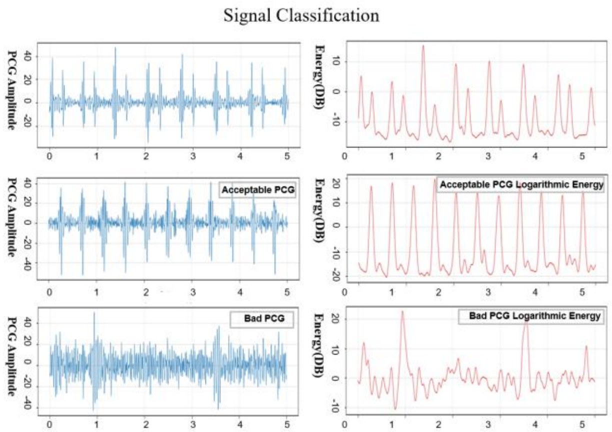

Figure 1. Signal classification. The left column shows normalized BCPCG waveforms, while the right column shows the corresponding logarithmic energy curves.

The dataset consists of 24,581 BCPCG segments obtained from 750 participants for quality assessment. It was developed based on the dataset framework of Ling et al.[7] and further expanded through additional data collection in this study. Each labeled segment was linked to its original recording using a recording identifier and segment index.

The BCPCG channel was band-pass filtered using a fourth-order Butterworth filter (20-80 Hz). Each recording was segmented using a 5s sliding window with a 2s step.

2.2. Data Annotation

All BCPCG segments were annotated by two trained annotators with professional knowledge in phonocardio-graphy. As illustrated in Fig. 1, segments with clear cardiac-cycle morphology, stable amplitude distribution, and regular energy variation were labeled as **Good**. Segments with partial noise contamination but still identifiable cardiac activity were labeled as **Medium**, whereas segments dominated by noise or lacking recognizable periodic structure were labeled as **Poor**.

2.3. Feature Extraction

A compact feature set was extracted from both the raw BCPCG signal and its envelope representation (PCGenv). The retained features describe waveform morphology, energy distribution, spectral structure, transient disturbance, and envelope rhythm organization. In this paper, these features are divided into five groups, labeled G1 to G5 respectively.

1) Temporal Morphology and Time Statistics (G1)

Basic waveform morphology and global amplitude features were characterized using statistical descriptors computed from the raw BCPCG signal, including zero-crossings, skewness, kurtosis, mean value, standard deviation, signal energy, and root mean square (RMS).

2) Log-Energy and Amplitude Counts (G2)

To characterize the energy distribution and local variation of BCPCG signals, several energy-related features were extracted based on the log-energy sequence

$$L(n) = \log(x_n^2 + \epsilon)$$

where ϵ is a small constant for numerical stability. The mean log-energy and the standard deviation of its first-order difference were used to describe the global energy level and local energy variation, including LogE_Diff_Std, LogE_Mean, AmpCount_GreaterThan0.8, AmpCount_ LessThan0.4 and AmpCount_LessThan0.1.

3) Frequency-Domain (G3)

Frequency-domain characteristics were summarized using spectral centroid and spectral entropy. The spectral centroid is defined as:

$$SC = \frac{\sum_f f|X(f)|}{\sum_f |X(f)|}$$

where $X(f)$ denotes the magnitude spectrum. This feature indicates the center of spectral energy distribution. Spectral entropy measures the dispersion of energy across frequency components, including spectral centroid, spectral entropy and spike fraction.

4) Envelope Peak Structure(G4)

Structural features were extracted from the envelope signal obtained using the Hilbert transform. After detecting envelope peaks, several descriptors were computed. These include the energy ratio between peak and non-peak regions.

$$ratio_{pm} = \frac{\sum_{n \in \Omega_{peak}} e_n^2}{\sum_{n \in \Omega_{nonpeak}} e_n^2}$$

where e_n measures the prominence of PCG events relative to background activity. Features describe the variability of peak heights and the variability of inter-peak intervals, characterizing the structural relationship between adjacent intervals, including Peak-NonPeak Ratio, PeakHeightCV, PeakIntervalCV, S1/S2 AlternationRatio and S1/S2 interval separation.

Table 1. Feature Groups and Dimensionality

Feature Group	Signal	Dim.
Time statistics + Temporal Morphology	BCPCG	7
Log-energy + Amplitude counts	BCPCG	5
Spectral features + Spike fraction	BCPCG	3
Peak structure + Alternation descriptors	PCGenv	5
Rhythm consistency	PCGenv	5
Total	BCPCG +PCGenv	25

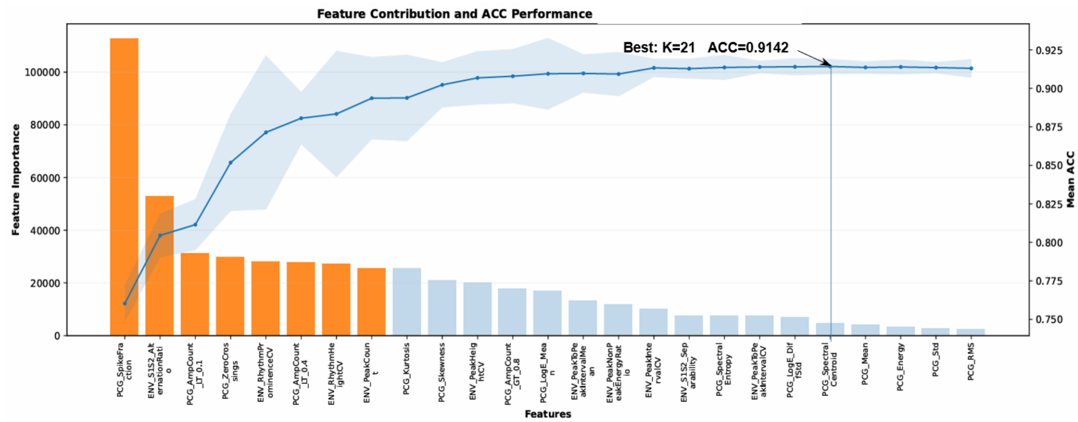

Figure 2. Feature Contribution and ACC Performance

5) Envelope Rhythm Consistency(G5)

Rhythm-level characteristics were derived from the sequence of detected envelope peaks. These include the number of peaks and the mean interval between adjacent peaks.

$$iv_{mean} = \frac{1}{M-1}\sum_{i=1}^{M-1} \Delta_i$$

where i denotes the interval between consecutive peaks and M is the number of detected peaks. Additional descriptors quantify the variability of peak, including Peak Count, PeakToPeak_IntervalMean, PeakToPeak_ IntervalCV, RhythmHeightCV and RhythmProminence.

2.4. Feature Selection

Feature selection was performed with LightGBM to reduce redundancy and rank feature contribution. Gain-based feature importance was averaged across five-fold stratified cross-validation to obtain a stable feature ranking.

As shown in Fig.2, the classification accuracy increased rapidly as top-ranked features were included and gradually stabilized thereafter. The best performance was achieved when the top 21 features were retained, reaching the best accuracy.

2.5. Classification Pipeline

As shown in Fig.3, the selected 21 features were used as inputs to the subsequent classification models for the BCPCG quality assessment task. We standardized features with StandardScaler and applied SMOTE to each cross-validation training fold to perform data enhancement.

Different machine-learning classifiers were used to evaluate the effectiveness of the proposed feature representation, including Support Vector Machine (SVM), Random Forest (RF), Logistic Regression (LR), K-Nearest Neighbor (KNN), XGBoost and LightGBM.

2.6. Evaluation Metrics and Statistical Analysis

To comprehensively evaluate the proposed three-class PCG quality assessment framework, five metrics were adopted: accuracy (ACC), area under the receiver operating characteristic curve (AUC), macro-averaged sensitivity (Se), macro-averaged specificity (Sp), and macro-averaged F1-score (Macro-F1) [10]. And all results were obtained under five-fold stratified cross-validation, in which the class distribution was preserved in each fold.

3. Results and Discussion

3.1. Classification Performance

All models were tested under the same preprocessing procedure and five-fold stratified cross-validation framework. The results are listed in Table 2.

Among them, LightGBM obtained the best classification results, with ACC, AUC, Se, Sp, and Macro-F1 values of 0.914, 0.967, 0.851, 0.926, and 0.850.

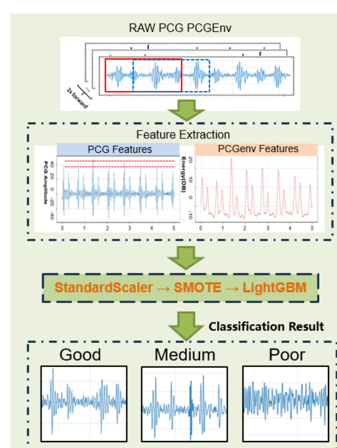

Figure 3. Overall workflow of the proposed BCPCG quality assessment framework.

Table 2. Classification performance of different models.

Method	ACC	AUC	Se	Sp	Macro-F1
LR	0.885	0.944	0.785	0.907	0.772
KNN	0.895	0.946	0.798	0.914	0.786
SVM	0.909	0.952	0.840	0.928	0.836
RF	0.905	0.950	0.829	0.925	0.826
XGBoost	0.912	0.959	0.847	0.929	0.844
LightGBM	0.914	0.967	0.851	0.926	0.850

The results indicate that the proposed multi-domain feature representation provides stable and reliable discriminative capability for three-class BCPCG quality assessment.

3.2. Discussion

To further investigate the contribution of different feature components, an ablation study was conducted by removing one feature group at a time from the full model. The corresponding results are presented in Table 3.

Table 3. Ablation Study Results.

Feature Set	ACC	Se	Sp	Macro-F1
Full	0.914	0.851	0.926	0.850
Minus-G1*	0.907	0.830	0.918	0.831
Minus-G2	0.904	0.831	0.920	0.823
Minus-G3	0.909	0.845	0.922	0.842
Minus-G4	0.910	0.847	0.923	0.841
Minus-G5	0.910	0.848	0.924	0.840

G1*: Feature group one(G1) without mean, energy, std and RMS

As shown in Table 3, the complete feature set achieved the best overall performance, with an accuracy of 0.914 and a Macro-F1 score of 0.850. When any feature group was excluded, the classification performance showed a consistent decline across all evaluation metrics. This pattern suggests that the selected features describe different aspects of BCPCG signal quality and jointly form a more complete representation.

4. Conclusion

This study proposed a multi-domain handcrafted feature framework for three-class BCPCG quality assessment using raw BCPCG signals and their envelope representations. Experimental results on 24,581 segments showed that the proposed representation achieved effective and stable classification performance, with LightGBM providing the best results. Feature selection results suggested that a compact subset of 21 features was sufficient to maintain strong performance. The ablation study further confirmed the complementary value of different feature groups, indicating that the proposed method is a practical and interpretable solution for BCPCG signal quality assessment.

References

[1] McConnell ME. Pediatric Heart Sounds. Berlin: Springer Science & Business Media; 2008.

[2] Wen YN, Lee APW, Fang F, Jin CN, Yu CM. Beyond auscultation: acoustic cardiography in clinical practice. Int J Cardiol. 2014;172(3):548-560.

[3] Zabihi M, Rad AB, Kiranyaz S, Gabbouj M, Katsaggelos AK. Heart sound anomaly and quality detection using ensemble of neural networks without segmentation. In: 2016 Computing in Cardiology Conference (CinC); 2016. p. 613-616.

[4] Springer DB, Brennan T, Ntusi N, et al. Automated signal quality assessment of mobile phone-recorded heart sound signals. J Med Eng Technol. 2016;40(7-8):342-355.

[5] Das D, Banerjee R, Choudhury AD, et al. Novel features from autocorrelation and spectrum to classify phonocardiogram quality. In: 2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC); 2017. p. 4516-4520.

[6] Springer DB, Tarassenko L, Clifford GD. Logistic regression-HSMM-based heart sound segmentation. IEEE Trans Biomed Eng. 2015;63(4):822-832.

[7] Ling L, Li Y, Lv Y, et al. A phonocardiogram quality assessment method using multi-domain feature. In: 2024 IEEE SENSORS; 2024. p. 1-4.

[8] Rubin J, Abreu R, Ganguli A, et al. Recognizing abnormal heart sounds using deep learning. arXiv. 2017;1707.04642.

[9] Branco P, Torgo L, Ribeiro RP. A survey of predictive modeling on imbalanced domains. ACM Comput Surv. 2016;49(2):1-50.

[10] Li Y, Ling L, Xiao Z et al. Using bone conduction technology for phonocardiogram measurement based on the watch-type device: a feasibility study. IEEE Sens J. 2025;25(17):33779–33791.

Acknowledgments

This research was funded by the National Key Research and Development Program of China (2023YFC3603600).

Address for correspondence:

Chengyu Liu.

School of Instrument Science and Engineering, Southeast University, Nanjing, China.

Email: chengyu@seu.edu.cn