Deep Learning-based Signal Quality Assessment in Wearable ECG Monitoring

Caiyun Ma[1], Zhongyu Wang[1], Lina Zhao[1], Xi Long[2], Rik Vullings[2],

Ronald M Aarts[2], Jianqing Li[1], Chengyu Liu[1*]

[1]State Key Laboratory of Bioelectronics, School of Instrument Science and Engineering, Southeast University, Nanjing, China
[2]Biomedical Diagnostics Lab, Department of Electrical Engineering, Eindhoven University of Technology, Eindhoven, The Netherlands

Abstract

Wearable electrocardiogram (ECG) monitoring is an effective method of screening for occult arrhythmia. However, signals from the wearable ECG monitoring device are often disturbed by various artifacts and noises originating from daily activities and which can significantly affect peak detection and ECG morphological feature extraction, leading to frequent false alarms for arrhythmia detection. Therefore, it is crucial to exclude ECG fragments with poor signal quality. In this study, we developed three xResNet-based ECG signal quality assessment models, trained on the Brno University of Technology ECG Quality Database. The first model can distinguish between ECG data in which the PQRST waves or only QRS complexes are visible from data in which these cannot be used for further analysis with a sensitivity (Se) of 98.87% and specificity (Sp) of 99.83%. The second model detects ECGs with visible PQRST waves with a Se of 97.15% and Sp of 95.95%. The third model classifies ECGs into data with PQRST visible, with only QRS visible, or unsuitable for analysis and achieves an accuracy (Acc) of 96.62%, 93.66%, and 98.97%, respectively. The results indicate that the proposed models can accurately evaluate the ECG signal quality during wearable monitoring, meeting the analysis requirements for arrhythmia.

1. Introduction

ECG signals are widely used in clinical diagnosis and treatment, such as cardiovascular disease diagnosis, arrhythmia identification, and sleep apnea detection. These applications generally require accurate detection of ECG signal feature points, as well as accurate measurement of ECG waveform morphology characteristics (such as amplitude, duration, polarity, and shape.) and interval characteristics (such as PR interval, and QT interval.) [1, 2]. During the collection of long-term ECGs using Holter or wearable dynamic monitoring, a significant amount of noise is generated due to the subject's autonomous activity, which brings a great challenge to ECG analysis [3].

In literature, there is mainly a binary ECG signal quality approch: acceptable and unacceptable, where collected ECG segments are divided into two categories based on the complexity of noise in the ECG segments. It is the most widely used signal quality classification "standard", especially promoted by the Physionet/Computing in Cardiology (CinC) 2011 [4], and it has been utilized in many related studies.

Many signal quality indicators (SQIs) have been developed, including wave feature analysis and QRS complex analysis. In our previous work [5], the performance of quality assessment algorithms based on wave features, including time-domain, frequency-domain, and nonlinear features were analyzed. It detects ECGs with visible PQRST waves with a sensitivity of 92.12% and specificity of 92.19%. The Acc of data with PQRST visible, with only QRS visible, or unsuitable for analysis are 90.74 %, 89.72%, and 97.60%. In 2008, Li *et al.* [6] proposed the bSQI signal quality index, which evaluates signal quality by comparing the QRS complexes detected by two different detection algorithms on a single ECG lead. Liu *et al.* [7] extended the bSQI based on multiple QRS detectors to improve signal quality assessment performance. Additionally, many methods have been developed based on the morphological and interval features of the QRS complex, but these features differ significantly between arrhythmic and normal ECGs. Such QRS-based signal quality assessment methods also require accurate and reliable QRS localization, which remains challenging in dynamic environments [8].

Some scholars have also used deep learning algorithms for ECG quality assessment. Zhou *et al*. [9] used one-dimensional convolutional neural networks (CNN) to identify low-quality ECG signals. Huerta *et al*. [10] transformed ECG signals into time-frequency graphs using continuous wavelet transform (CWT) and designed a CNN to classify ECG signal quality. In the above studies, deep learning models focused on the binary classification of acceptable and unacceptable signal quality.

Arguably, the classification of which ECG signal has good quality and which has poor quality also depends on the application at hand. Requirements on signal quality for disease screening are typically less strict than the requirements that are needed for disease diagnosis. In this study we therefore develop an ensemble of three signal quality assessment models that can be employed based on the desired application.The proposed models are all based on a xResNet34 [11] architecture and as application we aim at detecting the quality of ECG signals in wearable arrhythmia monitoring. For developing and evaluating the algorithm, the Brno University of Technology ECG Quality Database (BUT QDB) is used, which provides Class A, B, and C mobile ECGs. In class A, all waves in the ECG (i.e. P-wave, QRS-complex, T-wave) are clearly visible. In class B, only the QRS-complexes are clearly visible. In class C, the signal is deemed unsuitable for further analysis.

The three proposed models were trained to perform three tasks: (1) identifying Class A and B signals, crucial for arrhythmia screening, (2) detecting Class A signals for accurate disease diagnosis, and (3) differentiating between Class A, B, and C signals.

2. Methods

2.1. Database

The BUT QDB is an ECG quality evaluation database created by the cardiology team at the Department of Biomedical Engineering, Brno University of Technology [12]. The database consists of 18 long-term recordings of single-lead ECGs collected using a mobile ECG with a sampling frequency of 1,000 Hz. Class A, B and C were fully annotated in terms of ECG signal quality. In this study, the data was segmented into sliding windows of 10 s with a 5-s overlap. Further details regarding the data are provided in Table 1.

Table 1. The details of 10-s data

Class	A	B	C	Total
Sample size	33,279	20,663	10,510	64,452

2.2. Processing

To facilitate the processing of ECG signals, this work employed minimum-maximum normalization to standardize the ECG segments to the range (0, 1). To reduce the computational complexity, the signals were downsampled to 200 Hz.

2.3. xResNet34

As mentioned before, the proposed models are based on xResNet34, This is a neural network architecture that consists of a series of Residual blocks and some relatively basic building blocks. The structure of Residual block [11] is illustrated in (a) of Figure 1, and it includes two convolutional layer (Conv), Each Conv's output was rescaled by batch normalization (BN), followed by a rectified linear activation unit (ReLU). Skip connections were utilized.

The architecture of xResNet [11] is shown in part (b) of Figure 1, and consists of an Input stem and 4 Stages. In this paper, the xResNet34 network was used, where the Input stem includes 3 Convs, Stage 1 contains 2 Residual blocks, Stage 2 contains 3 Residual blocks, stage 3 contains 5 Residual blocks, and Stage 4 contains 2 Residual blocks and 1 down sampling layer. The size of the convolutional kernel and the stride (s) are indicated in Figure 1.

In this work, CrossEntropyLoss is used for classifier loss function. The neural network is trained using the stochastic gradient descent (SGD) optimizer with the a learning rate that is transformed during the iterative process, according to the following formula:

$$\mu_p = \frac{\mu_0}{(1+\alpha \cdot p)^{\beta}}$$

Where, μ_0 is the initial learning rate, p represents the relative value of the iteration process, α and β are hyperparameters. $\alpha = 10, \beta = 0.75$, μ_0 is set at 0.01. The model was trained for 30 epochs to identify Class A and B signals. It was trained for 60 for detecting Class A signal and for 100 epoches for differentiating between Class A, B, and C signals.

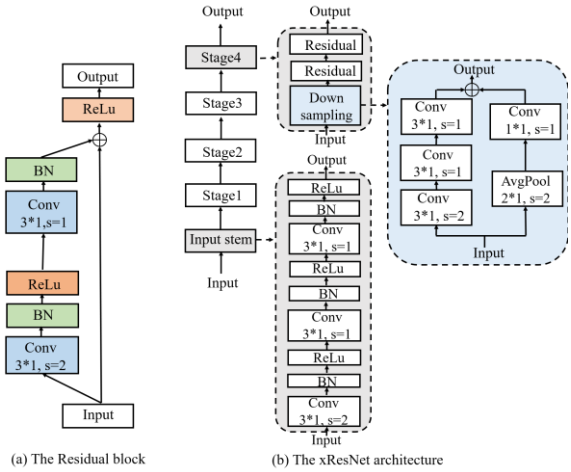

Fig 1. The architecture of the model based on xResNet.

2.4. Evaluation methods

For classification of two categories, Sensitivity (*Se*), Specificity (*Sp*), Accuracy (*Acc*), and Measure of Accuracy (*Macc*) are used as evaluation indicators. According to the label being positive or negative, four indexes were used: true positive (*TP*), true negative (*TN*), false positive (*FP*), and false negative (*FN*). The evaluation indicators are

defined based on these indexes as:

$$Se=TP/(TP+FN),$$
$$Sp=TN/(TN+FP),$$
$$Acc=(TP+TN)/(TP+FN+TN+FP),$$
$$Macc=(Se+Sp)/2.$$

For three-class classification, *Aacc*, *Bacc* and *Cacc* are the proportion of the total number of correct labels predicted for Class A, B and C respectively. *Maccz* is the proportion of the sum of predicted correct labels A, B and C to the total:

$$Maccz=(Aacc+Bacc+Cacc)/3.$$

A 5-fold cross validation scheme was used inthis study for the the evaluation of the proposed method. Additional, the models [5] based on Decision Tree and Features are used to compared.

3. Results

3.1. Results from three models

Table 2. The result from model for arrhythmia scanning

Fold	*Macc* (%)	*Acc* (%)	*Se* (%)	*Sp* (%)
1	99.59	99.64	99.75	99.38
2	99.44	99.74	99.95	98.62
3	99.35	99.68	99.93	98.43
4	99.4	99.74	99.94	98.71
5	99.48	99.66	99.79	99.00
Mean	99.46	99.69	99.87	98.83
SD	0.078	0.046	0.094	0.371

The model for arrhythmias scanning treats class A and B as positive and class C as negative. The results from the 5-fold cross-validation are shown in Table 2. Calculating the average performance over all the folds, gives *Macc*, *Acc*, *Se*, *Sp* of 99.46%, 99.69%, 99.87%, 98.83%, respectively.

Table 3. The result from model for arrhythmia diagnosis

Fold	*Macc* (%)	*Acc* (%)	*Se* (%)	*Sp* (%)
1	96.88	96.35	98.44	95.85
2	96.69	96.70	97.25	96.12
3	96.14	96.14	96.45	95.83
4	96.3	96.31	96.83	95.76
5	96.48	96.49	96.78	96.18
Mean	96.50	96.40	97.15	95.95
SD	0.296	0.210	0.775	0.189

When training the model for arrhythmia diagnosis, we consider class A signals as positive, and class B and C signals as negative. The results are shown in Table 3, with an average performance: *Macc* of 96.50%, *Acc* of 96.40%, *Se* of 97.15%, *Sp* of 95.95%.

Table 4. The result of the three-class classification

Fold	*Maccz* (%)	*Accz* (%)	*Aacc* (%)	*Bacc* (%)	*Cacc* (%)
1	96.43	96.01	96.60	93.44	99.24
2	96.40	96.03	96.71	93.44	99.05
3	96.43	96.09	96.69	93.78	98.81
4	96.36	96.07	97.02	93.15	98.91
5	96.47	96.01	96.06	94.50	98.86
Mean	96.42	96.04	96.62	93.66	98.97
SD	0.042	0.036	0.349	0.519	0.174

In this work, we also trained a three-class classification model on class A, B, and C signals, and the results are shown in Table 4, with average *Maccz, Accz, Aacc*, *Bacc*, *Cacc* 96.42, 96.04, 96.62, 93.66, 98.86, respectively.

3.2 Comparision with model based on Decision Tree+Features

Figure 2 shows the comparison of results between the model based on xResNet34 and Decision Tree+Features [5]. In the model for arrhythmia scanning, the *Macc* increased by 0.57%. For arrhythmia diagnosis, the *Macc* increased by 4.34%. For three-class classification, the Maccz increased by 3.74%.

4. Discussion and conclusion

In this study, we trained three ECG signal quality assessment models based on xResNet34, Compared with Decision Tree+Features, the results indicate that the performance of the quality evaluation scanning algorithm based on xResNet34 is superior.

The quality evaluation scanning algorithm based on xResNet34 is an end-to-end algorithm without human intervention or manual design of intermediate steps. The results domonstrated that the proposed method can provide for reliable method for signal quality scanning for application during arrhythmia ananlysis in wearable ECG monitoring.

Acknowledgement

This work was supported in part by the National Natural Science Foundation of China (62171123, 62071241, 62201144 and 62211530112), the Natural Science Foundation of Jiangsu Province (BK20192004), Postgraduate Research & Practice Innovation Program of Jiangsu Province (KYCX21_0089) and the China

Scholarship Council.

References

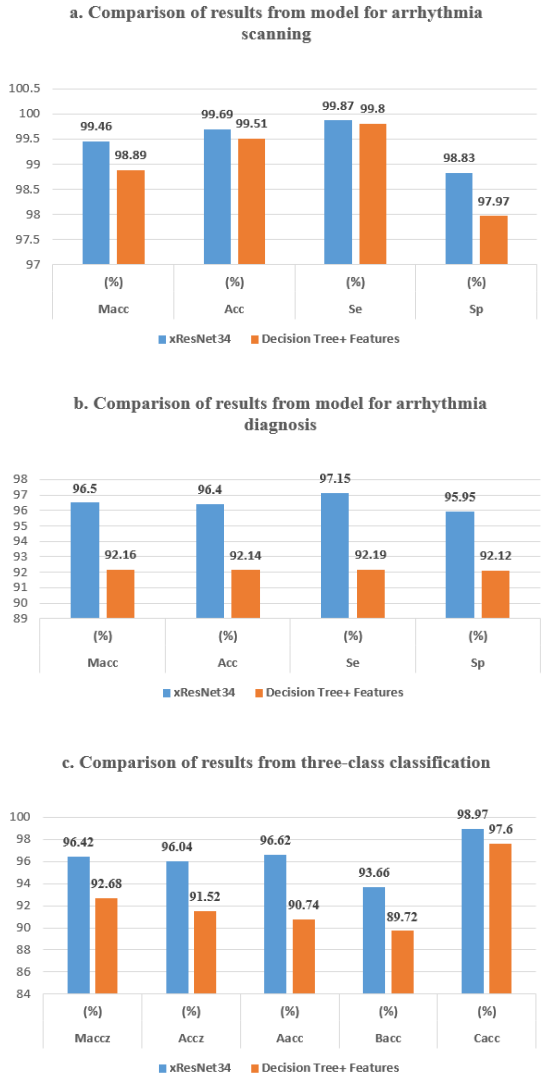

Fig 2 The comparison of results between xResNet34 and Decision Tree+Features.

[1] Qiao, Li, Gari, D. Clifford, "Signal quality and data fusion for false alarm reduction in the intensive care unit," *Journal of Electrocardiology*, vol. 45, no. 6, pp. 596-603, 2012, DOI：10.1016/j.jelectrocard.2012.07.015.

[2] M. Abdelazez, P. Quesnel, A. Chan, and H. Yang, "Signal Quality Analysis of Ambulatory Electrocardiograms to Gate False Myocardial Ischemia Alarms," *IEEE Transactions on Bio-Medical Engineering*, pp. 1318-1325, 2016, DOI：10.1109/TBME.2016.2602283.

[3] S. J. Kang, S. Y. Lee, H. I. Cho, and H. Park, "ECG Authentication System Design Based on Signal Analysis in Mobile and Wearable Devices," *IEEE Signal Processing Letters*, pp. 805-808, 2016, DOI：10.1109/LSP.2016.2531996.

[4] 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 *2011 Computing in Cardiology*, 2012.

[5] C. Ma, Z. Wang, M. Yang, J. Li, and C. Liu, "Decision Tree-based Model for Signal Quality Scanning in Wearable ECG." *in 2022 Computing in Cardiology*, 2022.

[6] 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," *Physiological Measurement*, vol. 29, no. 1, p. 15, 2008, DOI：10.1088/0967-3334/29/1/002.

[7] F. Liu *et al.*, "Dynamic ECG Signal Quality Evaluation based on the Generalized bSQI Index," *IEEE Access*, vol. 6, pp. 41892-41902, 2018, DOI：10.1109/ACCESS.2018.2860056.

[8] U. Satija, B. Ramkumar, and M. S. Manikandan, "A Review of Signal Processing Techniques for Electrocardiogram Signal Quality Assessment," *IEEE Reviews in Biomedical Engineering*, pp. 1-1, 2018, DOI：10.1109/RBME.2018.2810957.

[9] X. Zhou, X. Zhu, K. Nakamura, and N. Mahito, "ECG Quality Assessment Using 1D-Convolutional Neural Network," in *2018 14th IEEE International Conference on Signal Processing (ICSP)*, 2018.

[10] A. Huerta, A. Martínez-Rodrigo, V. González, A. Quesada, J. J. Rieta, and R. Alcaraz, "Quality Assessment of Very Long-Term ECG Recordings Using a Convolutional Neural Network," in *2019 E-Health and Bioengineering Conference (EHB)*, 2020.

[11] N. Strodthoff, P. Wagner, T. Schaeffter, and W. Samek, "Deep learning for ECG analysis: Benchmarks and insights from PTB-XL," *IEEE Journal of Biomedical and Health Informatics*, vol. 25, no. 5, pp. 1519-1528, 2020, DOI：10.1109/JBHI.2020.3022989.

[12] A. Nemcova, Smisek, R., Opravilová, K., Vitek, M., Smital, L., & Maršánová, L., "Brno University of Technology ECG Quality Database (BUT QDB) (version 1.0.0)." *PhysioNet*, 2020.

Address for correspondence:

Chengyu Liu
Sipailou 2, Nanjing, 210096, Jiangsu Province, P.R.China
Southeast University
E-mail: chengyu@seu.edu.cn