# A LSTM-Based Realtime Signal Quality Assessment for Photoplethysmogram and Remote Photoplethysmogram

Haoyuan Gao, Xiaopei Wu, Chenyun Shi, Qing Gao, Jidong Geng
School of Computer Science and Technology, Anhui University
Anhui University, Hefei, China
gaohaoyuan98@outlook.com, wxp2001@ahu.edu.cn, cyshi98@qq.com,
gqing_sun@163.com, 769904039@qq.com

## Abstract

*Monitoring physiological parameters is very important to access individuals' health status. Recent years, remote photoplethysmogram (rPPG) captured from human face by consumer-level cameras is used to estimate heart rate (HR). However, remote sensing signals are more easily affected by motion artifacts and environmental noise, which make the evaluation results unreliable. In this paper, we propose a long-short term memory network (LSTM) to assess the quality of the PPG(rPPG) signals in real time. This algorithm can also seek out the high quality segments from the ultra-long signals quickly. First, we labeled the PPG data by the combination of three traditional methods. Then, a LSTM network was trained to distinguish between clean signals and noisy signals in the PPG database. Finally, the network from the PPG data was verified in the rPPG data. The results of the experiments show that our method can get the signal quality index in real time, and the high quality fragments extracted by our method indirectly increase the accuracy of HR evaluation.*

## 1. Introduction

Photoplethysmogram(PPG) [2, 35] is often obtained by a pulse oximeter which illuminates the skin surface and measures changes in light absorption. It is a simple and low-cost technique which can monitor physiological parameters, such as blood pressure [21], heart and respiration rate [24], HR variability [18] and other changes of vital signs.

Recent years, remote PPG (rPPG) [4, 37] captured from human face by consumer-level cameras is used to estimate vital signs [3, 10, 29]. The rPPG is also referred as image Photoplethysmogram (iPPG) and non-contact Photoplethysmogram. The method can provide comfortable physiological assessment [28, 40]. It is a distinct convenience to get the rPPG with smart phones, laptops, or other high-definition cameras. However, the rPPG is easier contaminated by external light sources [12, 15] and motion artifacts [6, 36] than the PPG. It makes vital signs assessment by the rPPG more difficult. So identifying the high quality signals from the large amount of rPPG data is significant. The computational complexity and accuracy of the localization algorithm are of great importance.

It has been well investigated that the PPG measurements are significant in clinic. But there is still a lack of studies in identifying the high quality signals which also known as signal quality assessment (SQA) or signal quality index (SQI). We first introduce SQA methods in another cardiac monitoring signal, electrocardiogram (ECG) [32]. ECG is an earlier and more sophisticated cardiac monitoring signal. In this paper, ECG and PPG features for SQA are grouped into two categories, fiducial features and nonfiducial features.

The fiducial features include waveform morphological features of R-peak [27], QRS-complex [17, 30] and T-wave, and heartbeat interval features such as RR interval [16]. The nonfiducial features include frequency-domain, time-domain, statistical and information theoretic features. Signal-to-noise ratio [19], correlation coefficient [20], entropy [13], KL divergence [26], skewness and kurtosis [14] are common characteristics for quality assessment. The features extracted from the signals can be used as input features of machine learning to predict signal quality scores[5, 23, 25, 39].

The PPG signal is consistent with the ECG signal, and they both reflect the periodic changes of heartbeat. Signal quality assessment methods [27, 41] are shared between them. Not only that, signal quality evaluation of PPG signals also requires the following characteristics. Fiducial features include R-peak, RR interval [38]and perfusion [7, 34], nonfiducial features include signal power spectrum [42], entropy [33], skewness and kurtosis [7, 34] are common characteristics for quality assessment.

The above methods obtain new features by preprocessing a section of the signal. The quality of that signal is then evaluated by the new features. In real-time systems, the preprocessing features need to be calculated again after additional sampling points are acquired. The computation of statistical features and frequency domain features need enough sampling points. It makes the space complexity and time complexity of the algorithm unsatisfactory.

In this paper, we propose a LSTM-based realtime signal quality assessment (LSTM-SQA). It is an end-to-end algorithm and the algorithm can detect high-quality signal quickly. The memory cells can save the characteristics of the preceding signal and predict quality score for the current time signal.

For the rPPG signals, there are almost no signals with the same high-quality as PPG. We used the PPG data to solve the label imbalance problem of the rPPG data and shared the network parameters from the PPG to assess the quality of the rPPG. It is transfer learning that the PPG data are in the source domain and the rPPG data are in the target domain.

To verify the results, the mean scores of the sampling points are considered as the score of the whole signal. The mean values are compared with the scores given by the traditional methods and errors in heart rate assessment. We also visually display the results through pictures.

There are two contributions of this paper. Firstly, we applied LSTM to the PPG signal and propose a new realtime SQA algorithm. Secondly, we used domain adaptive method to expand the rPPG data and improve the stability of the model by sharing parameters.

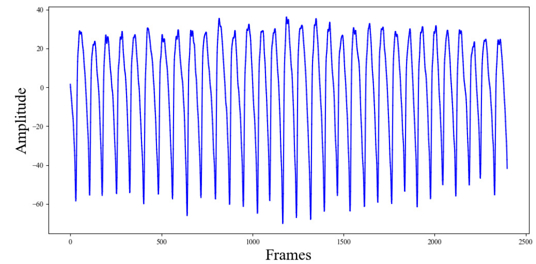
## 2. Methods

### 2.1. Introduction to the PPG databases

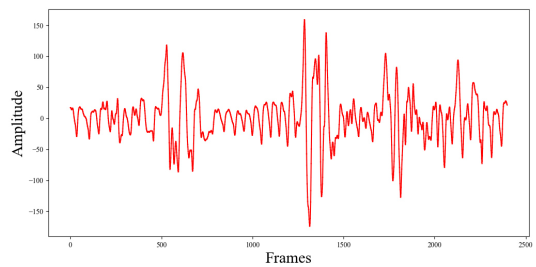
We use two PPG public data sets and two rPPG data sets collected by ourselves to verify our method.

#### 2.1.1 The PPG-DaLiA data set

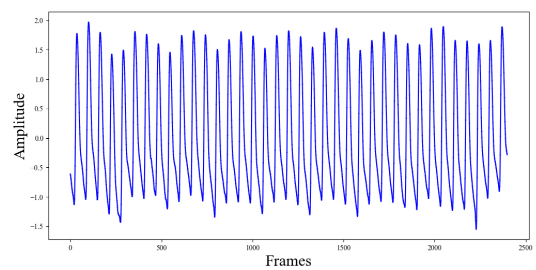
The PPG-DaLiA data set [31] contains heart rate, PPG, ECG and ACC signals. PPG data collection uses a mobile device worn by the wrist and ECG data collection uses a mobile device worn by the chest. The sampling rate of the PPG signal is 64 Hz. Heart rate was evaluated by ECG with pseudo peaks removed manually. The noise includes motion artifacts produced by walking up and down stairs, playing foosball, cycling, driving, eating lunch, walking and working. Fig.1(a) and Fig.1(b) show the signal fragments we processed. Red color represents low-quality signals and blue color represents high-quality signals.




(a) The high quality signal in PPG-DaLiA data set.




(b) The low quality signal in PPG-DaLiA data set.




(c) The signal in Cuff-Less blood pressure estimation data set.

Figure 1. Examples of PPG data sets.

#### 2.1.2 The Cuff-Less blood pressure estimation data set

Cuff-Less blood pressure estimation data set [11] contains ECG, PPG and arterial blood pressure (ABP) waveform signals recorded by patient monitors at various hospitals. Fig.1(c) shows the signal fragments we processed. It is the opposite of the PPG-DaLiA data set in value, so that the image is flipped up and down. The signals are sampled at the frequency of 125 Hz. With machine learning algorithms, the signals can estimate blood pressure. It can be seen intuitively that the signal quality of this database is higher than that of PPG-DaLiA dataset. We labeled the original signals as high quality signals. We used this data set to expand the high-quality signals, so the distribution of high and low quality signals is balanced.

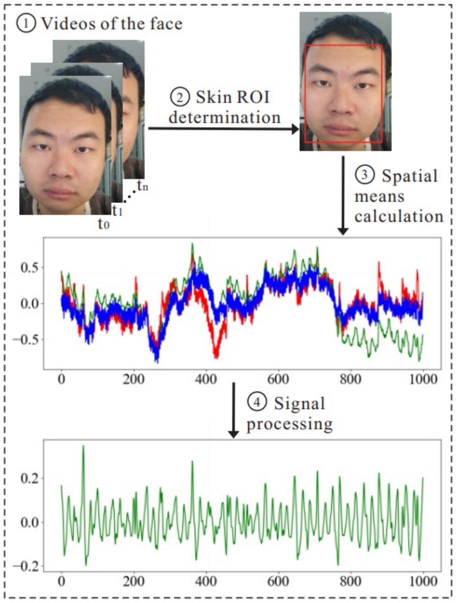



Figure 2. Basic framework of rPPG measurement.

### 2.1.3 The rPPG data sets

We built a facial image video dataset captured by the camera of nokia's mobile phone, and a supervised data set containing facial videos collected by microsoft lifecam studio and heart rates collected by CMS50E. The CMS50E is a PPG signal collector worn on human's finger. The frame rate of the videos is 30 fps, which also means the sampling rate of the signals is 30 Hz. Based on relevant rPPG studies in the literature, the corresponding basic framework can be summarized and described in Fig.2. First, video of faces was collected via a smartphone or other camera. To make sure the signals are reliable, the subjects were asked to remain still and we collect data in an environment with sufficient light. Such then, region of interesting (ROI) areas were obtained by manual annotation. We selected facial skin as ROI from the first frame of the video. The RGB color channels means were calculated from the ROI. As shown in the middle of the Fig.2, the red, green and blue channels provide three sets of original signals. Finally, signal processing methods (band-pass filtering algorithms, blind source separation algorithms or other denoise methods) were applied to derive the cleaner component. In this paper, we selected the green channel and carry out simple filtering processing as the original signal of rPPG. The processed signal is shown in the lower part of the Fig.2.

## 2.2. Long Short Term Memory Network

As would be expected, LSTM has been applied to a variety of practical problems. LSTM is a good solution for requiring the use of long-range contextual information [9], such as music generation [43], speech recognition [1] and reinforcement learning [8]. LSTM is local in space and time. Its computational complexity per time step is $O(1)$.

The architectures of the LSTM block can be seen in Fig.3. It features block input, three gates (a input gate $\boldsymbol{i}_t$, a forget gate $\boldsymbol{f}_t$ and a output gate $\boldsymbol{o}_t$), cells $\boldsymbol{c}_t$. The input gate controls which new information flows into the cells, the forget gate controls which information remains in the cells and the output gate controls which information in the cell is used to output in the LSTM block. The vector formulas for a LSTM block forward pass can be written as:

$$\boldsymbol{f}_t = \sigma(\boldsymbol{W_f}\boldsymbol{x}_t + \boldsymbol{R_f}\boldsymbol{h}_{t-1} + \boldsymbol{b_f}) \tag{1}$$

$$\boldsymbol{i}_t = \sigma(\boldsymbol{W_i}\boldsymbol{x}_t + \boldsymbol{R_i}\boldsymbol{i}_{t-1} + \boldsymbol{b_i}) \tag{2}$$

$$\boldsymbol{o}_t = \sigma(\boldsymbol{W_o}\boldsymbol{x}_t + \boldsymbol{R_o}\boldsymbol{h}_{t-1} + \boldsymbol{b_o}) \tag{3}$$

$$\tilde{\boldsymbol{c}}_t = \tanh(\boldsymbol{W_c}\boldsymbol{x}_t + \boldsymbol{R_c}\boldsymbol{h}_{t-1} + \boldsymbol{b_c}) \tag{4}$$

$$\boldsymbol{c}_t = \boldsymbol{f}_t \odot \boldsymbol{c}_{t-1} + \boldsymbol{i}_t \odot \tilde{\boldsymbol{c}}_t \tag{5}$$

$$\boldsymbol{h}_t = \boldsymbol{o}_t \odot \tanh(\boldsymbol{c}_t) \tag{6}$$

where the vector $\boldsymbol{x}_t \in \mathbb{R}^d$ is input vector to the LSTM block. The vector $\boldsymbol{h}_t \in \mathbb{R}^h$ is hidden state vector. It also known as output vector and recurrent vector of the LSTM block. The vector $\boldsymbol{c}_t \in \mathbb{R}^h$ is cell state vector and the vector $\tilde{\boldsymbol{c}}_t \in \mathbb{R}^h$ is cell input activation vector. The matrices $\boldsymbol{W_f} \in \mathbb{R}^{h\times d}$, $\boldsymbol{W_i} \in \mathbb{R}^{h\times d}$ and $\boldsymbol{W_o} \in \mathbb{R}^{h\times d}$ are weight matrices of input vector for the forget gate $\boldsymbol{f}_t$, the input gate $\boldsymbol{i}_t$ and the output gate $\boldsymbol{o}_t$ respectively. The matrices $\boldsymbol{R_f} \in \mathbb{R}^{h\times h}$, $\boldsymbol{R_i} \in \mathbb{R}^{h\times h}$ and $\boldsymbol{R_o} \in \mathbb{R}^{h\times h}$ are weight matrices of hidden state vector for the forget gate, the input gate and the output gate respectively. The vector $\boldsymbol{b_f} \in \mathbb{R}^h$, $\boldsymbol{b_i} \in \mathbb{R}^h$ and $\boldsymbol{b_o} \in \mathbb{R}^h$ are bias vector for the forget gate, the input gate and the output gate respectively. We set $\boldsymbol{c}_0 = \boldsymbol{0}$ and $\boldsymbol{h}_0 = \boldsymbol{0}$. The operator $\odot$ denotes the Hadamard product. The matric $\boldsymbol{W_c} \in \mathbb{R}^{h\times d}$ is weight matrices of input vector and the matric $\boldsymbol{R_c} \in \mathbb{R}^{h\times h}$ is weight matrices of hidden state vector for cell input activation vector. The functions $\sigma()$ and $\tanh()$ are nonlinear activation functions. The logistic sigmoid $\sigma()$ is used as gate activation function and the hyperbolic tangent $\tanh()$ is used as the LSTM block input and output activation function.

From Fig.3 and Formulate.5, we can see that the current time cell vector $\boldsymbol{c}_t$ is the sum of the last time cell vector $\boldsymbol{c}_{t-1}$ passing through the forget gate $\boldsymbol{f}_t$ and the cell input activation vector $\tilde{\boldsymbol{c}}_t$ passing through the input gate $\boldsymbol{i}_t$. Finally, we can get output vector $\boldsymbol{h}_t$ after cell vector $\boldsymbol{c}_t$ passing through the output gate $\boldsymbol{o}_t$ .

## 2.3. Traditional quality index of PPG(rPPG)

In the traditional SQA methods, characteristics are extracted by preprocessing the signal waveforms. In this section, $\mu_s$ and $\sigma_s$ are the mean and standard deviation of the

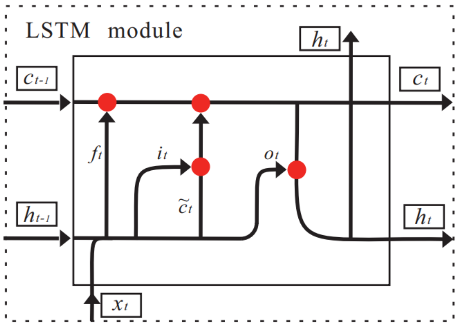



Figure 3. Internal structure of LSTM block.

signal $s$ respectively. $N$ is the number of sampling points and $\hat{s}$ is the filtered PPG signal. As follow show:

Perfusion($P_{SQA}$):The perfusion is the gold standard for assessing PPG signal quality [7]. The perfusion index is the ratio of difference of pulse blood flow to average blood flow in capillary bed. In other words, this is the difference of the light reflected by the blood pulse when the finger is illuminated, which can be defined as follows:

$$P_{SQA} = [(\hat{s}_{\max} - \hat{s}_{\min})/|\mu_s|] \times 100 \tag{7}$$

where $\hat{s}_{\max}$ and $\hat{s}_{\min}$ are the maximum and minimum values of the signal respectively.

Skewness( $S_{SQA}$):The skewness is the third standardized moment. It is a measure of the symmetry of a probability distribution, which is defined as:

$$S_{SQA} = \frac{1}{N}\sum_{i=1}^{N}\left[(s_i - \mu_s)/\sigma_s\right]^3 \tag{8}$$

Kurtosis( $K_{SQA}$ ):The kurtosis is the fourth moment of the distribution. It is a statistical measure used to describe the distribution of observed data around the mean, which is defined as:

$$K_{SQA} = \frac{1}{N}\sum_{i=1}^{N}\left[(s_i - \mu_s)/\sigma_s\right]^4 \tag{9}$$

Signal-to-noise ratio( $N_{SQA}$ ):The SNR measures the ratio of the energy in a desired signal to the energy in background noise. In this paper, the ratio of signal variance to the noise variance is used, as follows:

$$N_{SQA} = \sigma^2_{signal}/\sigma^2_{noise} \tag{10}$$

Template matching ( $TM_{SQA}$ ): The template matching approaches have been used for signal quality assessment of the PPG and ECG in [27]. The approximately periodic of heartbeat activity means that the more periodic the PPG signal, the higher the quality. The template matching quantizes regularity of signals in a segment, which is an indicator of reliability. The template matching of PPG signal can define as:

$$TM_{SQA} = \frac{1}{M}\sum_{r=1}^{M}\rho(\bar{s}_r, s_r) \tag{11}$$

where $s_r$ is a segment in PPG signal. The center of $s_r$ is R-peak/PPG-pulse peak and the width of $s_r$ is a fixed value. The variable $r$ means the $r{-}th$ peak and $M$ is the number of peaks. $\bar{s}_r$ is the mean of $s_r$ . The function $\rho()$ is Pearson correlation coefficient.

Standard deviation of peak-to-peak time interval( $STDT_{SQA}$ ): The $STDT_{SQA}$ [38] is also based on the pseudo-periodicity of the heartbeat. The time interval between adjacent peaks are smooth with no mutation:

$$STDT_{SQA} = \sqrt{\frac{1}{(M-2)}\sum_{r=2}^{M-1}(T_r - T_{r-1})^2} \tag{12}$$

where $T_r = t_{p_{r+1}} - t_{p_r}$ is the time interval between the $r{-}th$ successive peak $p_r$ and the $(r+1){-}th$ successive peak $p_{r+1}$. Since the time interval between neighboring peaks is pseudo-periodic for clean parts, it is expected that $STDT_{SQA}$ for high quality one to be smaller than low quality one.

The methods above use statistical information such as mean value and standard deviation. It means the signals should be long enough for evaluating statistics. It is not appropriate to position of high-quality segments from extra-long signals.

### 2.4. LSTM-SQA

Unlike the above methods, we propose a signal quality assessment method based on long short term memory network (LSTM-SQA). Our method is an end to end algorithm used the sampling points of the signal directly. This method does not have to calculate statistical characteristics or other features. It is a real-time algorithm and the LSTM network can process time sequence with time complexity of $O(1)$.

The detail of algorithm is shown in Fig.4. In the preprocessing stage, we built a small window to split the signal. The stride $d_1$ means we move the window $d_1$ sampling points at one time step. The width of the window $d_2$ is the length put into the network every moment. We set $d_1 = 1$, $d_2 = 5$ and the label of center point as the window's label. So the network predicts the quality index of a sampling point.

As depicted in Fig.4, the net contains three units. When new sampling points flow in, the network will update the information of hidden cells, and then determine the quality of the current sampling points. At the beginning of LSTM, enough sample points are necessary for updating the memory cells. After this stage, the network can automatically

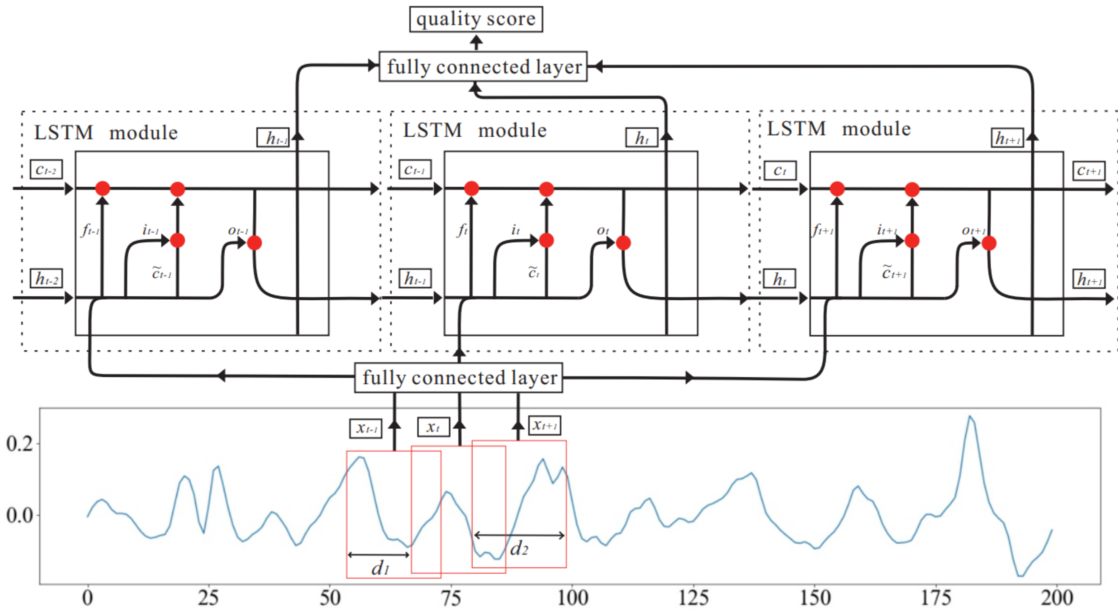



Figure 4. The LSTM net for PPG (rPPG) SQA. In each time step, the network reads the sampling points in the red box and gives the data quality of the current sampling points. The net contains three units: the first is fully connected layer for input; the intermediate unit is a standard LSTM block; the last is a fully connected layer for output with sigmoid activation function.

discard unimportant information and update valuable information to ensure the accuracy of the current period of evaluation. The whole process can output quality score in terms of sampling points, but scores are reliable only when the memory cells record enough information.

With sigmoid function, its output dimension is 1 and the range of network output is [0,1]. It is a binary classification model. Its output $p_o$ is the probability of been high quality, while $1 - p_o$ is the probability of been low quality. When the probability of high quality was greater than that of low quality, we considered the input signal as high quality. On the other hand, we also regarded the output of the network as continuous quality scores.

## 3. Experiment

### 3.1. Labelling of the PPG signals

Firstly, LSTM belongs to supervised machine learning, which requires data to be labeled. Traditional methods can only estimate the quality of long signals, but real time methods need higher resolution tag information. Therefore, we have improved some traditional methods so that they can evaluate the quality of signals in 1 second or even less, but the method still needs nearby signal information.

We evaluated the signal quality based on $P_{SQA}$, $TM_{SQA}$ and $STDT_{SQA}$. First, the signals were divided into smaller units by the average time interval of R-peak. The average of perfusion and $\bar{s}_r$ in $TM_{SQA}$ method have been calculated in this process. For the unit signals, the difference between the individual perfusion and the average perfusion could be obtained. We calculated the correlation between the unit signals and the average template. We also judged whether the unit signals lacks the R-peak. On this basis, the unit signals were divided into good or bad.

### 3.2. Generalization performance in motion artifacts noise

In this section, we verified the generalization of the network through cross-subject task. There are 15 subjects in the PPG-DaLiA dataset and we numbered them from 1 to 15. Every time we trained the network, we extracted the data of two subjects in turn as the test set, and the others were used to train the network. We judged whether the signal is good or bad according to the probability of network output. The network only has three layers, in which the length of hidden cells in RNN layer is 8. The results of the experiments are as shown in Table 1.

From the training accuracy, it can be seen that the network can stably converge. The average value of the training accuracy is 81.34%. From the accuracy of the test set, it can be seen that the average accuracy of the test set is 79.73%, which proves that the network is learnable and has good generalization in cross-subject experiments.

### 3.3. Correlation analysis between LSTM score and other signal quality assessment characteristics

In this paper, we used three traditional signal quality assessment methods as artificial labels. We regarded the prob-

| Subjects' number in Test Set | Train Set Accuracy | Test Set Accuracy |
|---|---|---|
| 1,2 | 81.45% | 80.02% |
| 3,4 | 80.02% | 82.00% |
| 5,6 | 82.62% | 75.22% |
| 7,8 | 81.75% | 80.96% |
| 9,10 | 83.18% | 82.99% |
| 11,12 | 79.34% | 79.56% |
| 13,14 | 81.04% | 77.33% |
| average | 81.34% | 79.73% |

Table 1. Accuracy of LSTM network in cross-subjects experiment.

ability values given by the network as the quality scores. We directly used the output of network to analyze the correlation with three other signal quality index parameters, including the SNR, the skewness and the kurtosis.

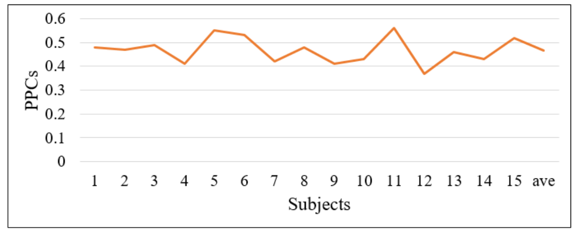
Since the network scores each sampling point, the SNR, the skewness and the kurtosis require an amount of sampling points to get statistical information, we need to average the scores of the sampling points from LSTM within a section of signal which has enough information to evaluate other parameters.

In this section, we cut the signal into segments with a length of 8 seconds. Correlation coefficients were calculated for each subject's average scores from the network and the three parameters. As showed in Fig.5, the abscissa represents the subject number, the last digit is the average value, and the ordinate represents the Pearson correlation coefficients (PCCs).

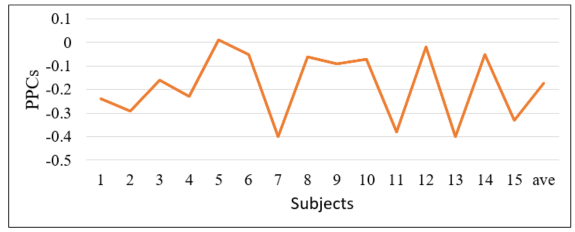
It can be seen that the quality scores given by the network have a strong correlation with the SNR, a weak negative correlation with the skewness, and no correlation with the kurtosis. The correlation between the quality scores and the SNR is the strongest, and the average correlation coefficient is 0.47. Subject 11 had the strongest correlation, and the correlation coefficient is 0.56. The quality scores from network have negative correlation with the skewness. Subject 13 has the strongest negative correlation, and the correlation coefficient is -0.39. However, the quality scores have almost no correlation with the kurtosis, and the average correlation coefficient is closed to 0.

### 3.4. Correlation analysis between heart rate accuracy and network score in PPG

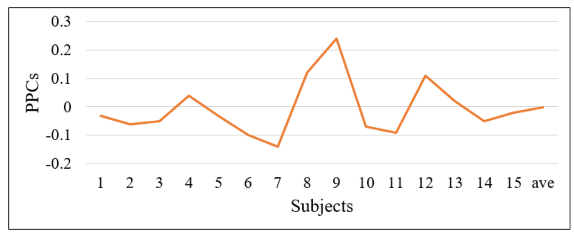
We have used the PPG signals in the PPG-DaLiA dataset to evaluate the heart rate. The database provides reliable heart rate obtained from the ECG signals. The Fast Fourier Transform analysis [22] was used for heart rate measurement in the PPG signals. The PPG signals contain a lot of motion noise, which makes the calculation results obviously different from the artificial labels. We calculated the absolute errors between the artificial labels and the FFT method. Then we compared them with the scores given by the LSTM-SQA.




(a) Correlation between LSTM and signal-to-noise ratio.




(b) Correlation between LSTM and skewness.

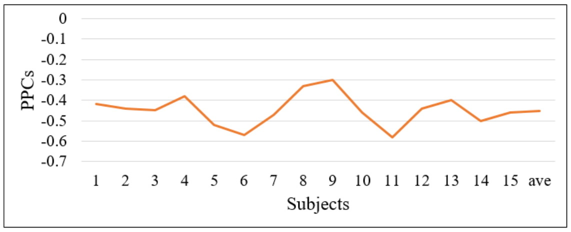



(c) Correlation between LSTM and kurtosis.

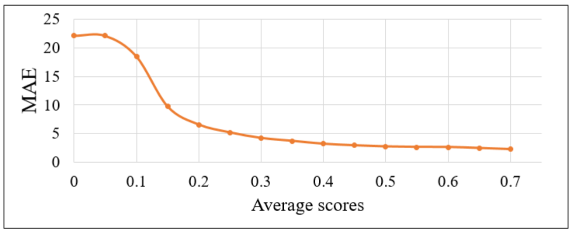
Figure 5. Correlation analysis.

As showed in Fig.6(a), the abscissa represents the subject number, the last digit is the average value, and the ordinate represents the PCCs. In Fig.6(b) and Fig.6(c), abscissa represents the average of network score, and ordinate represents the mean absolute error (MAE) between evaluation results by FFT method and artificial labels. It can be observed in Fig.6(a) that the signal quality score is negatively correlated with the mean absolute error of the HR evaluation. The correlation coefficient of subject 11 is -0.58 and the correlation is largest. Subject 9 have the lowest correlation, with the average correlation coefficient is -0.45.

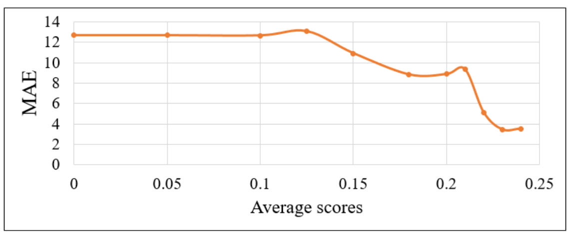
We selected signals with the quality score greater than a certain value and calculated the mean of absolute error for heart rate evaluation. As showed in Fig.6(b), the mean absolute error of all data on heart rate evaluation is 22.13. As low scores signals are excluded, the error became lower and lower. When the scores given by the network greater than 0.3, the average error is lower than 5. In other words, the error of heart rate evaluation is less than 5 beats per minute. The MAE of heart rate evaluation is 2.75 when the scores are greater than 0.5. In general, the higher the score given by the network, the more accurate the heart rate




(a) Correlation between LSTM and MAE.




(b) The influence of LSTM on MAE in PPG.




(c) The influence of LSTM on MAE in rPPG.

Figure 6. Mean absolute error of heart rate analysis.

assessment.

### 3.5. LSTM-SQA in rPPG

It is difficult to extract high quality rPPG signals even from high quality video. The signal quality of rPPG is lower than that of mobile PPG. This creates an imbalance between low quality data and high quality data which is a challenge for machine learning. However, the signal quality index criteria for rPPG signals are consistent with PPG. We enhance the PPG data and transfer to data domain of the rPPG. PPG-DaLiA dataset contain signals with motion noise, while Cuff-Less blood pressure estimation dataset contain a large amount of high quality signals.

We mixed the PGG datesets and the rPPG datasets to solve the class imbalance problem at first. We flipped the data, whitened signals, aligned the sampling rate and the amplitude, so that the feature distributions of the three data fields (include PPG-DaLiA dataset, Cuff-Less dataset and rPPG datasets) were similar. The sampling frequency of all signals is 64Hz by quadratic interpolation. Next, a LSTM network for PPG signals quality index has trained. Finally, the rPPG signals were used to fine tune the network parameters.

With the same method in the Section 3.4, we compared the error of heart rate assessment with the scores of the network. Because the signal quality of rPPG is far lower than that of PPG, the maximum value of average score is only 0.25. As shown in Fig.6(c), LSTM-SQA still conforms to the rule that the higher the scores, the smaller the error of heart rate evaluation. The rPPG signals is filtered in advance, and the average error of heart rate evaluation is 12.67. When the score is greater than 0.23, the average error of heart rate evaluation is less than 4.

We extracted the rPPG signal from the video of three women and three men. The green channel signals were filtered to get the original signal of 150s. The length of the extracted optimal signal segments was set to 10 seconds in advance. Each sample point was scored by LSTM. The sampling point scores in the 10s segment were summed to obtain the optimal segment of one rPPG signal. Some experimental results are shown in Fig.7.

It could be observed that the fragments with higher quality are selected by LSTM. In the segments, the beating cycle of the heart is obvious, and there are no obvious mutations in the heartbeat interval and perfusion. Heart rate can be estimated with less error. However, compared with high-quality PPG signals, there are still noise, some pseudo-peaks and mutation of extremum, which are also the limitation of rPPG without special denoising algorithm.

## 4. Conclusion and future work

We have used LSTM to assess the signal quality in real time. From the experimental results, using LSTM for SQA can enhance the credibility and accuracy of heart rate estimation. The quality scores of LSTM are directly and significantly correlated with the accuracy of heart rate. It works effectively for both heart rate estimation of mobile devices and heart rate estimation of rPPG, helping to eliminate the hard-to-evaluate signals and giving more accurate results.

However, mobile devices and rPPG still have limitations in that even the best fragments are not good as the data collected by a professional device. This requires a denoising algorithm to remove the noise, or train a network to regenerate high-quality signals. We have trained a mature discriminator and its gradient can be propagated back to the sampling points. In the future, we will introduce the generative adversarial network into it, and strive to enhance the quality of mobile data and rPPG data to the level of professional equipment. Such high-quality signals obtained by non-contact methods can extract more vital signs.

## References


[1] Navdeep Jaitly Alex Graves and Abdel rahman Mohamed. HYBRID SPEECH RECOGNITION WITH DEEP BIDIRECTIONAL LSTM University of Toronto Department of Computer Science 6 King ' s

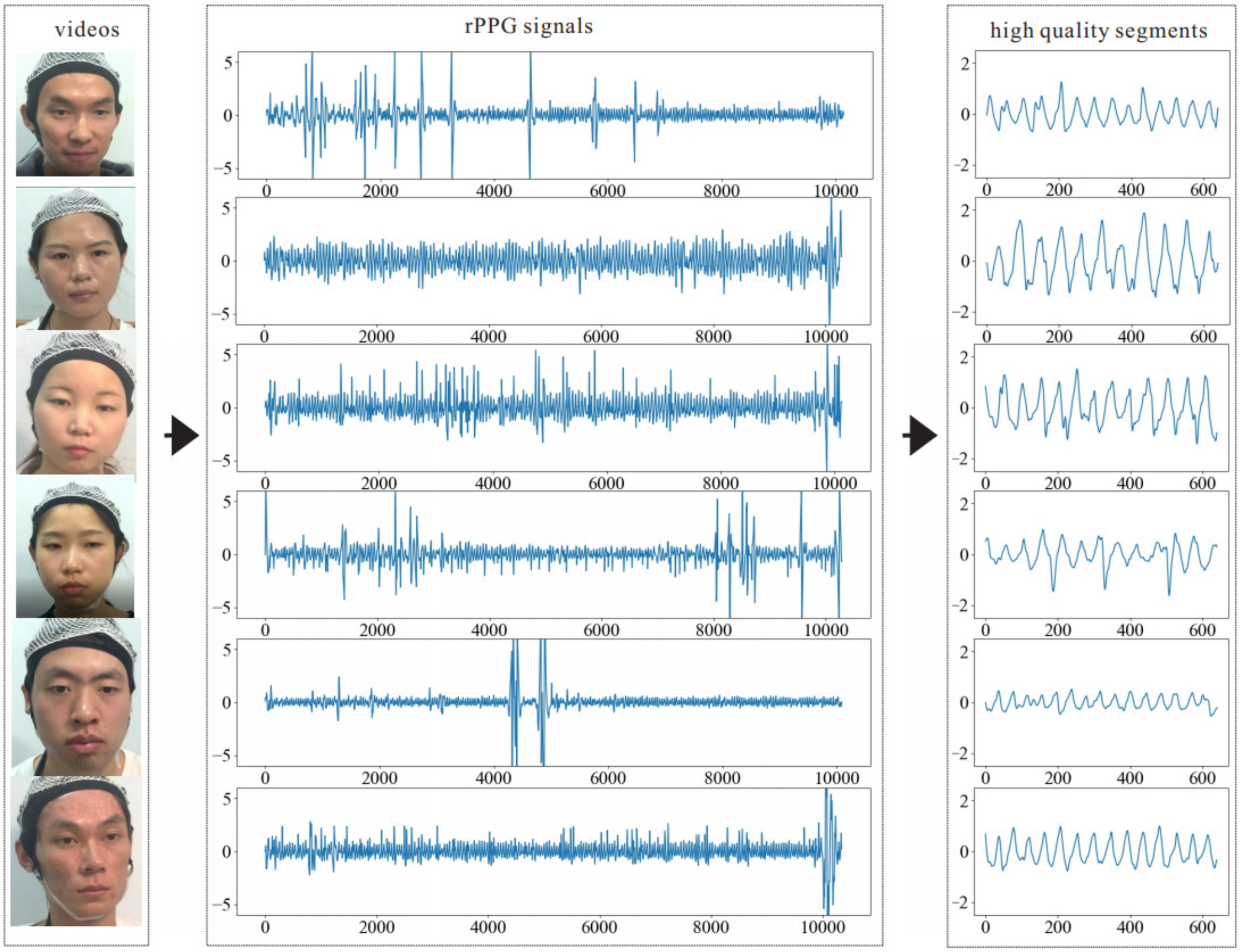



Figure 7. Display of quality assessment results in rPPG

College Rd . Toronto , M5S 3G4 , Canada. pages 273–278, 2013. 3

[2] John Allen. Photoplethysmography and its application in clinical physiological measurement. *Physiological Measurement*, 28(3), 2007. 1

[3] Ufuk Bal. Non-contact estimation of heart rate and oxygen saturation using ambient light. *Biomedical Optics Express*, 6(1):86, 2015. 1

[4] Xun Chen, Juan Cheng, Rencheng Song, Yu Liu, Rabab Ward, and Z. Jane Wang. Video-Based Heart Rate Measurement: Recent Advances and Future Prospects. *IEEE Transactions on Instrumentation and Measurement*, 68(10):3600–3615, 2019. 1

[5] G. D. Clifford, J. Behar, Q. Li, and I. Rezek. Signal quality indices and data fusion for determining clinical acceptability of electrocardiograms. *Physiological Measurement*, 33(9):1419–1433, 2012. 1

[6] G. De Haan and A. Van Leest. Improved motion robustness of remote-PPG by using the blood volume pulse signature. *Physiological Measurement*, 35(9):1913–1926, 2014. 1

[7] Mohamed Elgendi. Optimal signal quality index for photoplethysmogram signals. *Bioengineering*, 3(4), 2016. 1, 4

[8] Matthew Hausknecht and Peter Stone. Deep recurrent q-learning for partially observable MDPs. *AAAI Fall Symposium - Technical Report*, FS-15-06:29–37, 2015. 3

[9] Sepp Hochreiter and Jürgen Schmidhuber. Long Short-Term Memory. *Neural Computation*, 9(8):1735–1780, 1997. 3

[10] In Cheol Jeong and Joseph Finkelstein. Introducing Contactless Blood Pressure Assessment Using a High Speed Video Camera. *Journal of Medical Systems*, 40(4):1–10, 2016. 1

[11] Mohamad Kachuee, Mohammad Mahdi Kiani, Hoda Mohammadzade, and Mahdi Shabany. Cuff-less high-accuracy calibration-free blood pressure estimation

using pulse transit time. *Proceedings - IEEE International Symposium on Circuits and Systems*, 2015-July(2):1006–1009, 2015. 2

[12] Antony Lam and Yoshinori Kuno. Robust heart rate measurement from video using select random patches. *Proceedings of the IEEE International Conference on Computer Vision*, 2015 International Conference on Computer Vision, ICCV 2015:3640–3648, 2015. 1

[13] Jinseok Lee, David D. McManus, Sneh Merchant, and Ki H. Chon. Automatic motion and noise artifact detection in holter ECG data using empirical mode decomposition and statistical approaches. *IEEE Transactions on Biomedical Engineering*, 59(6):1499–1506, 2012. 1

[14] Qiao Li and G Clifford. Signal quality indices and data fusion for determining acceptability of electrocardiograms collected in noisy ambulatory environments. *Computing in Cardiology*, 38, 2011. 1

[15] Kuan Yi Lin, Duan Yu Chen, and Wen Jiin Tsai. Face-Based Heart Rate Signal Decomposition and Evaluation Using Multiple Linear Regression. *IEEE Sensors Journal*, 16(5):1351–1360, 2016. 1

[16] Chengyu Liu, Peng Li, Lina Zhao, Feifei Liu, and Ruxiang Wang. Real-time signal quality assessment for ECGs collected using mobile phones. *Computing in Cardiology*, 38(June 2014):357–360, 2011. 1

[17] Chengyu Liu, Xiangyu Zhang, Lina Zhao, Feifei Liu, Xingwen Chen, Yingjia Yao, and Jianqing Li. Signal quality assessment and lightweight QRS detection for wearable ECG smartvest system. *IEEE Internet of Things Journal*, 6(2):1363–1374, 2019. 1

[18] G. Lu, F. Yang, J. A. Taylor, and J. F. Stein. A comparison of photoplethysmography and ECG recording to analyse heart rate variability in healthy subjects. *Journal of Medical Engineering and Technology*, 33(8):634–641, 2009. 1

[19] Steffen A. Mann and Reinhold Orglmeister. A flexible PCA-based ECG-reconstruction algorithm with confidence estimation for ECG during exercise. *Computing in Cardiology*, 41(January):33–36, 2014. 1

[20] F. J. Martinez-Tabares, J. Espinosa-Oviedo, and G. Castellanos-Dominguez. Improvement of ECG signal quality measurement using correlation and diversity-based approaches. *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*, pages 4295–4298, 2012. 1

[21] Devin B Mccombie, Andrew T Reisner, and H Harry Asada. Multi-Channel Blind Identification of Local Arterial Dynamics . *New York*, pages 3521–3524, 2006. 1

[22] N. H. Mohd Sani, W. Mansor, Khuan Y. Lee, N. Ahmad Zainudin, and S. A. Mahrim. Determination of heart rate from photoplethysmogram using Fast Fourier Transform. *2015 International Conference on BioSignal Analysis, Processing and Systems, ICBAPS 2015*, pages 168–170, 2015. 6

[23] Eduardo Morgado, Felipe Alonso-Atienza, Ricardo Santiago-Mozos, Óscar Barquero-Pérez, Ikaro Silva, Javier Ramos, and Roger Mark. Quality estimation of the electrocardiogram using cross-correlation among leads. *BioMedical Engineering Online*, 14(1), 2015. 1

[24] Kazuki Nakajima, T. Tamura, and H. Miike. Monitoring of heart and respiratory rates by photoplethysmography using a digital filtering technique. *Medical Engineering and Physics*, 18(5):365–372, 1996. 1

[25] H. Naseri and M. R. Homaeinezhad. Electrocardiogram signal quality assessment using an artificially reconstructed target lead, 2015. 1

[26] Sungho Oh. *A new quality measure in electrocardiogram signal*. PhD thesis, University of Florida, 2004. 1

[27] Christina Orphanidou, Timothy Bonnici, Peter Charlton, David Clifton, David Vallance, and Lionel Tarassenko. Signal-quality indices for the electrocardiogram and photoplethysmogram: Derivation and applications to wireless monitoring. *IEEE Journal of Biomedical and Health Informatics*, 19(3):832–838, 2015. 1, 4

[28] Ming-Zher Poh, Daniel J. McDuff, and Rosalind W. Picard. Non-contact, automated cardiac pulse measurements using video imaging and blind source separation. *Optics Express*, 18(10):10762, 2010. 1

[29] A P Prathosh, Pragathi Praveena, Lalit K Mestha, and Sanjay Bharadwaj. Selective Ensemble Aggregation. 65(11):2902–2916, 2017. 1

[30] P. X. Quesnel, A. D.C. Chan, and H. Yang. Real-time biosignal quality analysis of ambulatory ECG for detection of myocardial ischemia. *MeMeA 2013 - IEEE International Symposium on Medical Measurements and Applications, Proceedings*, pages 1–5, 2013. 1

[31] Attila Reiss, Ina Indlekofer, Philip Schmidt, and Kristof Van Laerhoven. Deep PPG: Large-scale heart rate estimation with convolutional neural networks. *Sensors (Switzerland)*, 19(14):1–27, 2019. 2

[32] Udit Satija, Barathram Ramkumar, and M. Sabarimalai Manikandan. A Review of Signal Processing Techniques for Electrocardiogram Signal Quality Assessment. *IEEE Reviews in Biomedical Engineering*, 11:36–52, 2018. 1

[33] Nandakumar Selvaraj, Yitzhak Mendelson, Kirk H. Shelley, David G. Silverman, and Ki H. Chon. Statisti-

cal approach for the detection of motion/noise artifacts in Photoplethysmogram. *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*, pages 4972–4975, 2011. 1

[34] Hyeon Seok Seok, Sangjin Han, Junyung Park, Donggeun Roh, and Hangsik Shin. Photoplethysmographic pulse quality assessment methods based on similarity analysis. *Proceedings - 2018 Joint 10th International Conference on Soft Computing and Intelligent Systems and 19th International Symposium on Advanced Intelligent Systems, SCIS-ISIS 2018*, pages 350–353, 2018. 1

[35] Kirk Shelley and Stacey Shelley. Pulse Oximeter Waveform: Photoelectric Plethysmography. *Clinical Monitoring: Practical Applications for Anesthesia and Critical Care*, (May):420–423, 2001. 1

[36] Yu Sun, Hu Sijung, Vicente Azorin-Peris, Jonathon A Chambers, Zhu Yisheng, and Stephen E Greenwald. Motion-compensated noncontact imaging photoplethysmography to monitor cardiorespiratory status during exercise. *Journal of biomedical optics*, 16(7):077010, 2011. 1

[37] Yu Sun and Nitish Thakor. Photoplethysmography Revisited: From Contact to Noncontact, from Point to Imaging. *IEEE Transactions on Biomedical Engineering*, 63(3):463–477, 2016. 1

[38] Fatemehsadat Tabei, Rajnish Kumar, Tra Nguyen Phan, David D. McManus, and Jo Woon Chong. A Novel Personalized Motion and Noise Artifact (MNA) Detection Method for Smartphone Photoplethysmograph (PPG) Signals. *IEEE Access*, 6:60498–60512, 2018. 1, 4

[39] Diana P. Tobon and Tiago H. Falk. Online ECG quality assessment for context-aware wireless body area networks. *Canadian Conference on Electrical and Computer Engineering*, 2015-June(June):587–592, 2015. 1

[40] Yonggang Yana, Xiang Maa, Lifeng Yaoa, and Jianfei Ouyanga. Noncontact measurement of heart rate using facial video illuminated under natural light and signal weighted analysis. *Bio-Medical Materials and Engineering*, 26:S903–S909, 2015. 1

[41] Chenggang Yu, Zhenqiu Liu, Thomas McKenna, Andrew T. Reisner, and Jaques Reifman. A Method for Automatic Identification of Reliable Heart Rates Calculated from ECG and PPG Waveforms. *Journal of the American Medical Informatics Association*, 13(3):309–320, 2006. 1

[42] Yue Zhang and Junjun Pan. Assessment of photoplethysmogram signal quality based on frequency domain and time series parameters. *Proceedings - 2017 10th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics, CISP-BMEI 2017*, 2018-January:1–5, 2018. 1

[43] Kun Zhao, Siqi Li, Juanjuan Cai, Hui Wang, and Jingling Wang. An emotional symbolic music generation system based on LSTM networks. *Proceedings of 2019 IEEE 3rd Information Technology, Networking, Electronic and Automation Control Conference, ITNEC 2019*, (Itnec):2039–2043, 2019. 3