UDC 004.021:004.75

HEART RATE ESTIMATION FROM PHOTOPLETHYSMOGRAM AND ACCLERATION SMARTPHONE DATA BASED ON CONVOLUTIONAL NEURALNETWORK AND LONG SHORT TIME MEMORY

Gaowei Qiu
Postgraduate student in the Academic Department of Information and Communication Technology
qgw1314@163.com

Gaowei Qiu
Graduated from Nanchang Aviation University, China. Areas of scientific interest include the development of neural network algorithms and software engineering.

Abstract. The wearable reflective photocapacitance plethysmograph (PPG) sensor can be integrated into the watch or strap to provide instantaneous heart rate (HRs), causing minimal inconvenience to users. However, the existence of motion artifacts (MAs) leads to inaccurate heart rate estimation. In order to solve this problem, I propose a new deep learning neural network to ensure accurate estimation of HR in high-intensity exercise. Methods: I propose a deep neural network based on multi-class and non-uniform multi-label classification for HR estimation. It includes two convolution layers, two short and long term memory (LSTM) layers, one connection layer and three full connection layers including softma,also includes a concatenation layer. The average absolute error of the algorithm for all training data sets and test data sets is less than 1.5 bpm, including 1.09 bpm for training data sets and 1.46 bpm for test data sets. The proposed algorithm is superior to the most advanced methods in accurate estimation of heart rate.
Keywords: PPG, LSTM, Convolution,Concatenation,Heart rate and motion artifacts.

Introduction.

Reflective photoplethysmography (PPG) sensor measures the intensity change of skin reflected light and provides PPG signal representing the change of arterial blood volume between systolic and diastolic phases of the heart cycle. The reason why the sensor is concerned is that it can be installed in a watch or wrist strap to measure and monitor instantaneous heart rate (HRs), minimizing the inconvenience to users. However, the sensor is sensitive to motion artifacts (MAs), which come from the pressure and motion exerted on the wrist of the PPG sensor. Motion artifacts will eventually lead to inaccurate heart rate estimation. A few years ago, Zhang and others shared the data set of acceleration and PPG signals measured simultaneously in [1] motion, which prompted people to study the use of acceleration signals to eliminate motion artifacts in PPG sensors. However, despite the efforts and progress of algorithms, these methods do not always provide accurate results.

In this paper, the purpose is to measure heart rate more accurately,so I propose a new deep neural network based on multi-class and non-uniform multi-label classification for human heart rate estimation. In the proposed algorithm, I consider two power spectra from the input layer of PPG and acceleration signals. In addition, I use the acceleration signal strength in the input layer. I assume that the acceleration signal strength can provide information about the recent HR changes: high intensity represents intense exercise, which may change HR. The proposed algorithm includes two convolution layers, two LSTM layers, one connection layer and three full connection layers (including one softmax layer). In the proposed algorithm, the power spectrum from PPG and acceleration signals is fed to two convolutions to provide MA cancellation in the PPG power spectrum. The output is flattened and connected to a fully connected layer, which is then connected to the acceleration signal strength. Then, the output is fed to two LSTM layers, and then to other full-connection layers including softmax. The LSTM layer uses minimum outliers to track HR tracking. In this algorithm, I also propose a new scheme to modify the real HR value to Gaussian distribution to evaluate the loss value. The performance of the proposed algorithm is evaluated by comparing with the results of [1] - [18] previously reported.

Based on PPG, ACC, and acceleration intensity Datasets.

I used two data sets to evaluate the proposed HR estimation algorithm - the IEEE Signal Processing Cup (ISPC) 2015 data set (n=23) and the direct measurement data set obtained by the developed equipment (n=48). These two sets of data include multi-channel PPG signal and three-axis accelerometer signal, which are measured by the equipment on the wrist during high-intensity physical exercise. At the same time, ECG signals ire measured in the chest as real HR reference. The data includes 12-minute three-channel PPG signal and 50-Hz sampling three-axis acceleration signal. The dataset is divided into two groups, BAMI-I and BAMI-II. During every 4 minutes of running and walking, the subjects walked or ran with a treadmill stick in the last two minutes. I designed this link to reflect the heart rehabilitation exercise of patients with heart disease who have poor exercise ability. They usually walk or run with a treadmill stick. There are 17 males and 6 females, with an average age of 22.0±1.7 years. The whole exercise process is also carried out on the treadmill. For these two data sets, the reference real heart rate was measured by ECG data recorded simultaneously by 24-h the ambulatory ECG monitor (SEER Light, GE Medical, Milwaukee, WI, USA).

In this study, I selected the ISPC data set (n=23) as the training data, because it includes various movements, such as waking up, running, resting, jumping, push up, shaking hands, stretching, pushing and boxing. I also added the own dataset BAMI-I (n=25) to the training data to increase the size of the training data, and used another dataset named BAMI-II (n=23) to test the training algorithm. For all data sets, the ecg-based HRs are calculated using an 8-second window, with a 2-second offset (6-second overlap) to obtain the HRs every 2 seconds. In the whole study, the same window length (8 seconds) and shift (2 seconds) are used to evaluate the performance of the proposed algorithm relative to the existing algorithm [1] - [17].

Description of deep neural network algorithm based on ACC and PPG.

Figure 1 (a) summarizes the architecture of the proposed algorithm. It consists of eight layers: two convolutional layers, two LSTM layers, one tandem layer and three full connection layers, including a softmax layer. More specifically, a 2D convolution layer, a one-dimensional convolution layer, a flat layer, a splice layer, a fully connected layer, two LSTM layers, and a fully connected layer folloid by a softmax.

For the input layer, the power spectrum from PPG signal and acceleration signal is divided into two (size 2 × 222): the top signal is from PPG (Ps (i)), and the bottom signal is from acceleration (Pa (i)). Note that the power spectrum is based on each 8-second window and then shifted for 2 seconds (6-second overlap). Input layer input to 32 2×37 cores, two-dimensional convolution with a step of 4, and then perform the nonlinear function leaky corrected linear unit (ReLU) and 1×2 max pooling, in steps of 2. The resulting size is 1×28×32.The characteristic diagram of 32 shows the intermediate PPG power spectrum after MA cancellation. Send characteristic map into 64 1×5-core, one-dimensional convolution with a side of 1, and then carry out leaky ReLU and 1 with a side of 2×2 max pooling, get other sizes of 1×14×64, showing the PPG power spectrum of MA elimination. The resulting feature map is spread evenly to 896 nodes, which are fully connected to 512 nodes with leaking ReLU.

Subsequently, the acceleration intensity I^{a} (I) is connected with 512 activations. I assume that the acceleration intensity represents how the current HR value will change in the near future. For example, high acceleration intensity means high intensity exercise, which may increase HR. 513 connected nodes are sent to two LSTM layers. These two LSTM layers act as HR tracking algorithms by considering the local HR tracking mode. Due to the existence of LSTM layer, even if the signal to noise ratio (SNR) of PPG signal is extremely low, the dominant frequency corresponding to HR will not deviate seriously in the continuous window. The first layer LSTM contains 512 nodes, each node has 6 time steps, and all nodes are connected to the second layer LSTM. Note that six time steps provide the best accuracy. The numerical analysis of the effect of time step will be introduced. The second LSTM layer also includes six time steps, each of which provides an output with a length of 222. Then, only the output from the last time step is fed to the full connection layer, which is connected to 222 nodes with leaking ReLU. In the frequency range of 0.6-3.3 Hz, the frequency resolution is 0.012 Hz, and the result of the final PPG power spectrum after MA cancellation is 222 activations. The activation is input to the softmax layer, which provides the final probability of the real HR value.

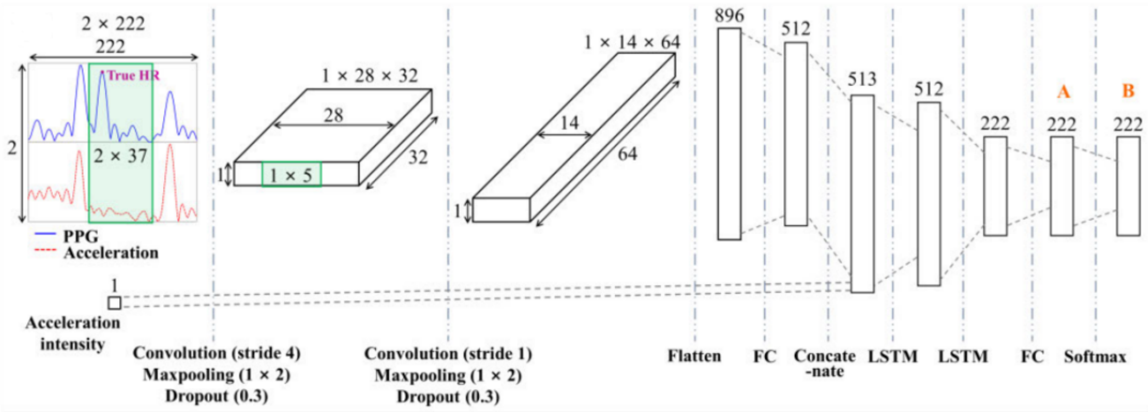

Figure 1. The architecture of the proposed algorithm; 2D convolution layer, 1D convolution layer, a full connection layer,a concatenation layer two LSTM layers and a full connection layer are structured in sequence

To avoid over-fitting, I applied dropout to two convolution layers and two LSTM layers. For convolution layer, the exit rate is set to 0.3. For LSTM layer, the drop rate of input linear transformation is 0.3, and the drop rate of loop state is 0.2. Given a set of parameters of the algorithm, softmax provides the probability corresponding to each HR value subdivided in 222 frequency boxes: from 0.6 to 3.3 Hz, with an interval of 1/222. Therefore, with the real HR value, I can consider the multi-class classification problem, which calculates the cost through multi-class cross entropy ξ Value is

(1)

where and are the true HR probability and the predicted HR probability, respectively, for the frequency bin at the window. In the multiclass cross-entropy, has the value of one only when the frequency bin corresponds to a true HR. Otherwise, has the value of 0. Then, the formulation can be simplified as

(2)

where $\tilde{y}^{i}_{c=true}$ is the predicted HR probability in the frequency bin covering the true HR value.

However, this method has a disadvantage that the frequency bin covering the real HR value may not accurately represent the real HR.Please note that the real HR is obtained from the ECG data sampled at 50 Hz in the BAMI I and II data sets and from the ECG data sampled at 125 Hz in the ISPC data set, while the frequency box is obtained from the PPG signal sampled at 25 Hz. To solve this problem, I will pay ξ Multiply the normalized Gaussian function, where the Gaussian distribution centered on the real HR is normalized to get the maximum value. The revised cost function is

(3)

where represents real HR. Through modification, I can alleviate the problem of non-overlapping HR based on ECG and PPG. In this study, we chose the standard deviation $\sigma = 3$.

Experimental Result.

Based on the proposed algorithm, I found that the AAE and ARE values of the training data set (n=48) ire 1.09 bpm and 0.92% respectively, and the AAE and ARE values of the test data set (n=23) ire 1.46 bpm and 1.23% respectively. Table 1 summarizes the performance. Specifically, the AAE and ARE values of the ISPC dataset are 0.76 bpm and 0.66% respectively, and the AAE and ARE values of the BAMI-I dataset are 1.39 bpm and 1.17% respectively.

$$AAE = \frac{1}{N}\sum_{i=1}^{N}\left|y_c^i - \tilde{y}_c^i\right| \quad (4)$$

$$ARE = \frac{\sum_{i=1}^{N}\left|y_c^i - \tilde{y}_c^i\right|}{\sum_{i=1}^{N}\left|\bar{y}_c^i - \tilde{y}_c^i\right|} \quad (5)$$

Table 1. Performance summary with each dataset: ISPC and BAMI-I ire used as training set, and BAMI-II was used as test set

		AAE(bpm)		ARE(%)	
Training set	ISPC(n=23)	0.76	1.09	0.66	0.92
	BAMI-I(n=25)	1.39		1.17	
Test set	BAMI-II(n=23)	1.46		1.23	

Figure 2 shows the HR trajectory based on the proposed algorithm and acceleration intensity estimation; Figure 2 (a) and (b) show the results of the ISPC dataset, and Figure 2 (c) and (d) show the results of the BAMI-I and BAMI-II dataset. On the top panel of the figure.

The estimated HR trace results are compared with the results considering only the dominant power spectrum of PPG.In addition, it can be seen from the bottom panel of Figure 2 that the acceleration intensity is related to the increase of HR. In particular, Figure 2 (a) and (b) show the results of subjects 14 and 17. These two sets of data are considered to be the most challenging because the measured PPG signal is seriously damaged by MAs, resulting in a very low signal-to-noise ratio.

In fact, for topic 14 in the dataset, the reported AAEs are 6.63 bpm [1], 8.07 bpm [2], 7.29 bpm [3], 4.89 bpm [6], 9.59 bpm [7], 1.60 bpm [17], 12.12 bpm [4], 7.91 bpm [13] and 7.66 bpm [8]. For subject 17, the reported AAEs ire 7.82 bpm [1], 7.01 bpm [2], 3.01 bpm [3], 3.05 bpm [6], 3.01 bpm [7], 2.04 bpm [17], 3.31 bpm [4], 2.44 bpm [13] and 2.77 bpm [8]. On the other hand, the algorithm I proposed provides very accurate estimates of the whole segmented HR for two subjects (AAE: 0.87 bpm for subject 14 and 1.49 bpm for subject 17). Figure 2 (c) and (d) also show the results when the signal-to-noise ratio is very low. For Topic 1 in the BAMI-I dataset, when only the dominant power spectrum of PPG is used, the AAE is 46.64 bpm, and the AAE of WFPV is 14.27 bpm. On the other hand, the results show that AAE is 1.45 bpm.

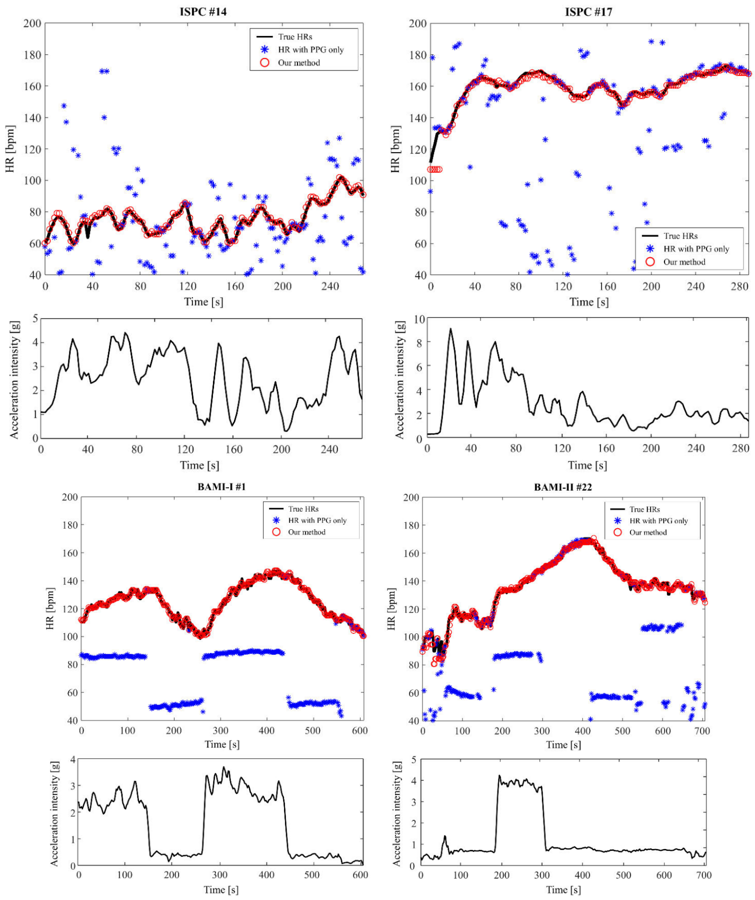

*Figure 2.*Estimated HR trace based on the proposed algorithm (Tops) and acceleration intensity (Bottoms); (a) Subject 14 (ISPC), (b) Subject 17 (ISPC),(c) Subject 1 (BAMI-I) and (d) Subject 22 (BAMI-II).

Conclusion.

The proposed deep learning algorithm estimates HR based on the power spectrum of PPG and acceleration signals and cascaded acceleration strengths. The algorithm consists of a two-dimensional convolutional layer, a one-dimensional convolutional layer, and a fully connected layer for MA cancellation. The algorithm also includes a connection layer, two LSTM layers, a fully connected layer, and a softmax layer for human resource tracking and estimation. The AAE of the algorithm for the training dataset and the test dataset is 1.09 bpm and 1.46 bpm, respectively, which exceeds the results of the most advanced methods currently available and allows for more accurate measurement of heart rate.

However, the most important aspect of future research will involve the implementation of energy saving when implementing the proposed algorithm in real time on wearable devices. Deep learning is undoubtedly to provide good performance, but the implementation of wearable devices faces many challenges, because they require low-power algorithms due to limited computing power.

References

[1] Z. Zhang, Z. Pi, B. Liu. TROIKA: A general framework for heart rate monitoring using wrist-type photoplethysmographic signals during intensive physical exercise, IEEE Trans. Biomed. Eng. Feb. 2015. Vol. 62, no. 2, pp.

[2] Z. Zhang. Photoplethysmography-based heart rate monitoring in physical activities via joint sparse spectrum reconstruction, IEEE Trans. Biomed.Eng. Aug. 2015. Vol. 62, no. 8, pp.

[3] S. Salehizadeh, D. Dao, J. Bolkhovsky, C. Cho, Y . Mendelson, K. Chon. A novel time-varying spectral filtering algorithm for reconstruction of motion artifact corrupted heart rate signals during intense physical activities using a iarable photoplethysmogram sensor, Sensors, 2016. Vol. 16, no. 1, p. 10.

[4] S. Fallet, J.-M. V esin. Robust heart rate estimation using wristtype photoplethysmographic signals during physical exercise: An approach based on adaptive filtering, Physiol. Meas. Feb. 2017. Vol. 38, no. 2, pp.

[5] M. Bolthesaz Mashhadi, E. Asadi, M. Eskandari, S. Kiani, F. Marvasti. Heart rate tracking using wrist-type photoplethysmographic (PPG) signals during physical exercise with simultaneous accelerometry, IEEE Signal Process. Lett. Feb. 2016. Vol. 23, no. 2, pp.

[6] B. Sun, Z. Zhang. Photoplethysmography-based heart rate monitoring using asymmetric least squares spectrum subtraction and Bayesian decision theory, IEEE Sensors J. Dec. 2015. Vol. 15, no. 12, pp.

[7] A. Temko. Accurate heart rate monitoring during physical exercises using PPG, IEEE Trans. Biomed. Eng. Sep. 2017. Vol. 64, no. 9, pp.

[8] M. A. Motin, C. K. Karmakar, M. Palaniswami. PPG derived heart rate estimation during intensive physical exercise, IEEE Access. 2019. Vol. 7, pp. 56062–56069.

[9] E. Khan, F. Al Hossain, S. Z. Uddin, S. K. Alam, M. K. Hasan. A robust heart rate monitoring scheme using photoplethysmographic signals corrupted by intense motion artifacts, IEEE Trans. Biomed. Eng. Mar. 2016. Vol. 63, no. 3, pp.

[10] H. Lee, H. Chung, H. Ko, J. Lee. wearable multichannel photoplethysmography framework for heart rate monitoring during intensive exercise, IEEE Sensors J. Apr. 2018. Vol. 18, no. 7, pp.

[11] V . Nathan, R. Jafari. Particle filtering and sensor fusion for robust heart rate monitoring using wearable sensors, IEEE J. Biomed. Health Informat. Nov. 2018. Vol. 22, no. 6, pp.

[12] Y. Fujita, M. Hiromoto, T. Sato. PARHELIA: Particle filter-based heart rate estimation from photoplethysmographic signals during physical exercise, IEEE Trans. Biomed. Eng. Jan. 2018. Vol. 65, no. 1, pp.

[13] A. Galli, C. Narduzzi, G. Giorgi. Measuring heart rate during physical exercise by subspace decomposition and Kalman smoothing, IEEE Trans.Instrum. Meas. May 2018. Vol. 67, no. 5, pp.

[14] H. Chung, H. Lee, J. Lee. State-dependent Gaussian kernel-based power spectrum modification for accurate instantaneous heart rate estimation, PLoS ONE, Vol. 14, no. Art. no. e0215014.

[15] H. Chung, H. Lee, J. Lee. Finite state machine framework for instantaneous heart rate validation using iarable photoplethysmography during intensive exercise, IEEE J. Biomed. Health Informat. Jul. 2019. Vol. 23, no. 4, pp.

[16] J. Lee, H. Chung, H. Lee. Multi-mode particle filtering methods for heart rate estimation from iarable photoplethysmography, IEEE Trans.Biomed. Eng. Oct. 2019. Vol. 66, no. 10, pp.

[17] D. Biswas, L. Everson. M. Liu, M. Panwar, B.-E. V erhoef, S. Patki, C. H. Kim, A. Acharyya, C. V an Hoof, M. Konijnenburg, and N. V an Helleputte, CorNET: Deep learning framework for PPG-based heart rate estimation and biometric identification in ambulant environment, IEEE Trans. Biomed. Circuits Syst. Apr. 2019. Vol. 13, no. 2, pp.

[18] M. B. Mashhadi, M. Farhadi, M. Essalat, F. Marvasti. Low complexity heart rate measurement from iarable wrist-type photoplethysmographic sensors robust to motion artifacts, in Proc. IEEE Int.Conf. Acoust., Speech Signal Process. (ICASSP), Calgary, AB, Canada, Apr. 2018, pp. 921–924.

ОЦЕНКА ЧАСТОТЫ СЕРДЕЧНЫХ СОКРАЩЕНИЙ, ОСНОВАННАЯ НА ОБЪЕМНОЙ ФОТОГРАММЕТРИИ СВЕРТОЧНЫХ НЕЙРОННЫХ СЕТЕЙ И КРАТКОВРЕМЕННОЙ ПАМЯТИ И ДАННЫХ УСКОРЕНИЯ СМАРТФОНОВ

Цю Гаовей

Аспирантура, Институт
информационно - коммуникационных
технологий

Департамент информационно-коммуникационных технологий Академии информационной безопасности
E-mail: qgw1314@163.com

Аннотация. Носимый отражательный фотоконденсаторный объемный граф (PPG) датчик может быть интегрирован в часы или ремешок, чтобы обеспечить мгновенный сердечный ритм (HR), минимизируя неудобства для пользователя. Однако наличие спортивных псевдотеней (MA) приводит к неточным оценкам сердечного ритма. Чтобы решить эту проблему, я предложил новую нейронную сеть глубокого обучения для обеспечения точной оценки HR в высокоинтенсивных упражнениях. Метод: Для оценки HR была предложена глубокая нейронная сеть, основанная на множественной и неоднородной классификации с несколькими метками. Она включает в себя два слоя намотки, два слоя краткосрочной и долгосрочной памяти (LSTM), соединительный слой и три полных соединения, включая софтму, а также каскадный слой. Средняя абсолютная погрешность алгоритма для всех наборов данных обучения и наборов тестовых данных составляет менее 1,5 bpm, из которых набор данных обучения составляет 1,09 bpm, а набор тестовых данных - 1,46 bpm. Предложенные алгоритмы превосходят самые современные методы точной оценки частоты сердечных сокращений.

Ключевые слова: PPG, LSTM, свертки, каскады, пульс и движущиеся псевдотени