Deep Recurrent Neural Network for Extracting Pulse Rate Variability from Photoplethysmography During Strenuous Physical Exercise

1st Ke Xu, 2nd Xinyu Jiang, 3rd Haoran Ren
Center for Intelligent Medical Electronics (CIME), School of Information Science and Technology
Fudan University
Shanghai, China
{kexu18, jiangxy18,hrren17}@fudan.edu.cn

4th Xiangyu Liu
School of Art Design and Media
East China University of Science and Technology
Shanghai, China
Y12170017@mail.ecust.edu.cn

5th Wei Chen
Center for Intelligent Medical Electronics (CIME), School of Information Science and Technology
Fudan University
Shanghai, China
w_chen@fudan.edu.cn

Abstract*—Pulse rate variability (PRV) extracted from photoplethysmography (PPG) signal is a promising surrogate for heart rate variability (HRV) and has shown its great potential in diagnosing cardiac dysfunctions and autonomic nervous system diseases. However, the accurate extraction of PRV during strenuous physical exercise faces enormous challenges due to PPG's extreme vulnerability to motion artifacts. In this work, we introduce a deep recurrent neural network (RNN) based on bidirectional Long-Short Term Memory Network (biLSTM) for accurate PPG cardiac period segmentation. After that, three important indexes for PRV are calculated, which are peak intervals, pulse intervals, and instantaneous heart rates (IHR). Comparison results with state-of-the-art methods on a dataset including 48 subjects show the promising performance of the proposed algorithm in PRV indexes estimation and recovery. To our best knowledge, this is the first time a deep learning-based algorithm been involved for extraction of PRV from seriously corrupted PPG signals.

***Keywords*—heart rate variability (HRV), pulse rate variability (PRV), photoplethysmography (PPG), motion artifacts, physical exercise**

I. Introduction

Heart rate variability (HRV) is the continuous fluctuation of period length between cardiac cycles. Because HRV is a promising reflection of the many physiological factors modulating the normal rhythm of the heart [1], for the last 20 years it has been used for the diagnose of cardiovascular diseases such as myocardial infarction and cardiac arrhythmia, as well as for autonomic nervous system dysfunctions. HRV has also shown strong correlations with blood pressure and renal failure [1]. Besides, HRV is an important index for sleep staging and diagnosis of sleep disorder [2].

The gold standard for HRV measurement is the R-R interval of Electrocardiograph (ECG) signals. However, the inconvenience of leads placement for ECG has largely impeded it from long-term and wide-spread use.

Photoplethysmography (PPG), on the other hand, is a popular non-invasive physiological measurement utilizing either transmission-type or reflection-type detectors in which photons with specified wavelengths are injected into body parts, and photoelectric receivers placed on the same or the other side of these parts receive transmitted or reflected photons. PPG describes the volume fluctuation in local vascular networks which is mainly caused by cardiac activity, thus includes systolic and diastolic phase in the waveform.

Each peak of the PPG waveform corresponds to a specific R peak in simultaneously recorded ECG, and the peak-to-peak interval changes of PPG waveform, which are called pulse rate variability (PRV), have strong relationship with HRV. Previous researches have proven PRV a promising surrogate for HRV under stationary [3] and non-stationary [4] situations, in short-term and long-term recording [5]. Since PPG has more convenience in measurement setups and has been widely involved in wearable devices such as sport watches, it is expected to be a powerful tool in detecting the occurrence of cardiovascular diseases during daily life and under physical exercise.

However, the accurate extraction of PRV during strenuous physical exercise faces enormous challenges due to PPG's extreme vulnerability to motion artifacts [6]. PPG detectors, which are places on extremes of body such as fingertips and earlobes, can easily be affect by motion induced changes on contact surface and hemodynamic turbulence in local vessels.

Recent researches have focused on the accurate recovery of average heart rate (usually during a period of 8-10s) for contaminated PPG signal, the representatives of which are TROKIA [7], JOSS [8] and SpaMA [9]. Although these studies achieve considerable success in average heart rate estimation from frequency domain, most of them do not pay much attention on the recovery of PPG cardiac period, which is the basics of PRV extraction. Therefore, the HR estimated by these studies can only represent the average level of cardiac activity and cannot be used for PRV evaluation.

To solve the problem above, a new idea is proposed in this research that, instead of recovering the average heart rate, we manage to recover the cardiac phases, which is contaminated by motion artifacts and unable to be recognized manually, and then extract the peaks and feet of waveform from the phases in order to gain accurate information of PRV. To be specific, after pre-processing, a deep recurrent neural network (RNN) based on bidirectional Long-Short Term Memory Network (biLSTM) is used for point-to-point cardiac phases classification of corrupted PPG waveform, dividing which into systolic and diastolic phases. Then peaks and feet of waveform are extracted at the tuning points of these phases and further used for PRV evaluation. Comparison results with state-of-the-art methods on a dataset including 48 subjects show the promising performance of the proposed algorithm in PRV indexes estimation and recovery. To our best knowledge, this is the first time a deep learning-based algorithm been

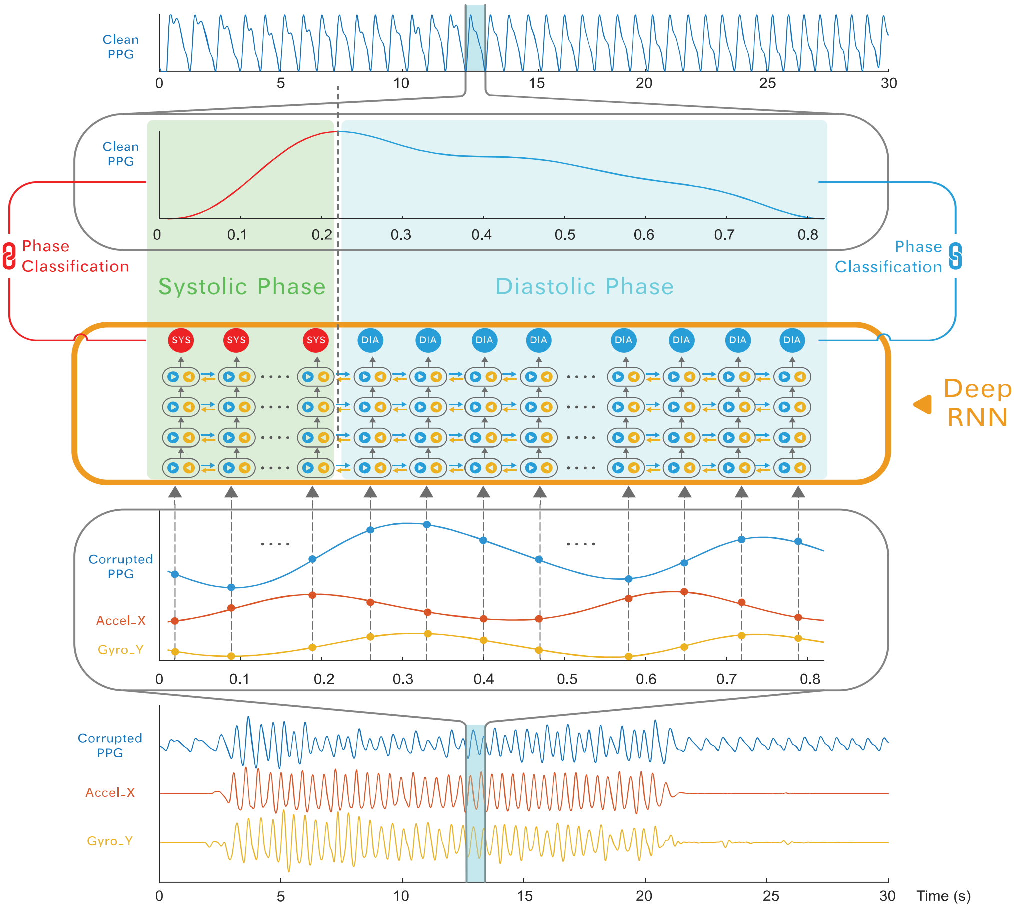

Fig. 1. Algorithm structure of the proposed work. At each sample point, a feature vector combining corrupted PPG, *Accel_X*, and *Gyro_Y* signal is used as the input of the four-layer network. The network with continuous inputs learns from the input and inverted input sequences of feature vectors, and further extracts cardiac phase information layer by layer. The output of the network is expected to be the categories of sample points, either systolic phase or diastolic phase, consistent with the left-hand clean PPG reference.

involved for extraction of PRV from seriously corrupted PPG signals.

This paper is organized as follows: Section II gives descriptions of materials and the algorithm involved in this paper. Section III presents results of performance evaluation and gives some discussion. Section IV concludes this paper.

II. Methodology

This section provides detailed information about the dataset and the algorithm utilized in this study.

A. *Dataset*

Data are acquired simultaneously from left and right hand index fingertips using special designed PPG detectors which are integrated with 6-axis IMU sensors to record movement information. 48 healthy subjects are invited to this experiment (age range 18-65, 20 male and 28 female, white, yellow, and dark skin colors) with written consent. The experiment procedures follow strictly with the principles of the Declaration of Helsinki.

For each subject, an experiment of 31 min is carried out (for subject No. 1-10 each, ten experiments were carried out), with left hand keep static during the experiment and right hand performs movement of bending index finger and moving along horizontal finger-axial, in different frequencies and amplitudes. Although in the free living scenarios, direction and degree of freedom can vary a lot, these two types of movement are selected because they have the strongest effects on PPG observation. Data are sampled at 100 Samples/s. The dataset and detailed description can be found on our website[1].

In this study, the data collected from subject No. 1 to No. 10 are used as training set, and the data from subject No. 11 to No. 48 are used for performance evaluation. Independent testing set is used to evaluate the cross-subject robustness.

B. *Pre-processing*

The corrupted infrared (IR) PPG from right hand as well as IMU sensor data is first de-trended and filtered by 6 Hz-cutoff low pass FIR filter, and then centralized and variance normalized. The X-axis of accelerometer (*Accel_X*) and Y-axis of gyroscope (*Gyro_Y*) from IMU are selected as noise related reference because they are the strongest related ones to experiments defined motion patterns (*Accel_X* is along the horizontal finger-axis, and *Gyro_Y* is related to movement of bending finger).

An envelope filter is also involved to normalize the amplitude of PPG waveform in order to minimize the effect of respiration, in which

$$x_{EF}(t) = \frac{x(t)-e_{lower}(t)}{e_{upper}(t)-e_{lower}(t)} - 0.5, \quad (1)$$

where $x_{EF}(t)$ is the filtered signal, $e_{upper}(t)$ and $e_{lower}(t)$, represent the upper and lower envelopes of unfiltered data

[1] https://github.com/WillionXU/CIME-PPG-dataset-2018/

TABLE I
TRAINING PARAMETERS FOR DEEP RNN

Parameter	Details
Input Vector	$[\, P_{IR} \quad A_X \quad G_Y \,]^T$
Output Vector	$Cate_P$
biLSTM Size	[100 50 50 20]
biLSTM State Activation Function	tanh
biLSTM Gate Activation Function	sigmoid
Output Layer	Soft Max
Optimizer	Adam
Max Epoch	200

P_{IR}, A_X, and G_Y represent IR PPG, X-Axis accelerometer and Y-Axis gyroscope signals respectively. $Cate_P$ indicates the predicted category of this sample point ("systolic" or "diastolic").

$x(t)$, and 0.5 bias used to keep signal within the range of [-0.5, 0.5].

The clean PPG waveforms from left hand are used to extract the category of each sample point as gold standard, which belongs to systolic phase or diastolic phase, based on the method introduced in [10] and checked manually to ensure the correctness of the category markers.

C. Deep Recurrent Neural Network

Long-short Term Memory Network has been widely involved in time-sequence related issues because of its memory characteristics and ability to predict sample points based on previous long-term and short-term information. Bidirectional LSTM, which is an improved version of LSTM, inputs the feature sequences forward and backward, thus is able to gain the information behind and ahead the specific sample point. This makes biLSTM more effective for analyzing complex sequences and extracting abstract information from them. Using deep network will also contribute to the understanding of complex sequences.

In this study, a deep RNN based on bidirectional LSTM is used for the point-to-point classification of PPG sample points. As it is illustrated in Fig. 1, at each sample point, a feature vector combining corrupted PPG, *Accel_X*, and *Gyro_Y* signal is used as the input of the four-layer network. The network with continuous inputs learns from the input and inverted input sequences of feature vectors, and further extracts cardiac phase information layer by layer. The decreasing number of neurons among layers allows more abstract knowledge gained by deeper layers.

The output of the network, with the help of a Soft Max function as output layer, is expected to be the categories of sample points, either systolic phase or diastolic phase, consistent with the left-hand clean PPG reference. Therefore, the network aims to make a precision classification for every sample point of corrupted PPG waveform, even during the period when cardiac phases are unable to be recognized manually from contaminated PPG waveform, as it is demonstrated in Fig. 1.

To train this subject independent network, data from subject No. 1 to No. 10 are used, whose total length is about 52 hours. Data from the rest 38 subjects are used for testing, during which a sequence length of 5 seconds is used for input minibatch in order to satisfy the requirement of real-time scenario. Detailed training parameters can be found in Table I.

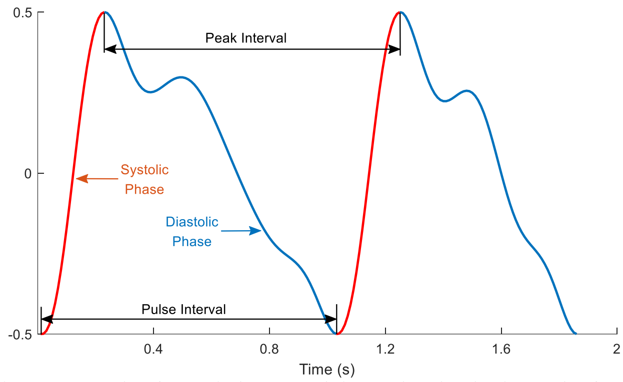

Fig. 2. Example of PRV indexes. Peak interval and pulse interval refer to the time interval between adjacent peaks and feet of waveform.

III. RESULTS AND DISCUSSION

A. PRV Indexes

Using the cardiac phases predicted by the network, peaks and feet of PPG waveform can be extracted at the tuning points of phases. To compare the performance of the proposed algorithm, a pulse registration is first applied to match reference pulses and the pulses predicted by the network one by one, during which a reference pulse and a predicted pulse are paired if they cover at least 50% the pulse length of each other, which indicates the predicted peaks in time are matched with the ones on reference signal. Next, the precision and sensitivity of the proposed algorithm can be calculated as

$$Precision = \frac{TP}{TP+FP} \times 100\% \tag{2}$$

$$Sensitivity = \frac{TP}{TP+FN} \times 100\% \tag{3}$$

in which *TP* is the number of paired pulses, *FP* is the number of unpaired (unmatched) pulses among the network predicted ones (fake predicted pulses), and *FN* refers to the number of unpaired pulsed among the reference ones (pulses fail to predict).

In addition, the peak and pulse interval as well as instantaneous heart rate are calculated and used for performance evaluation. Peak interval and pulse interval, as demonstrated in Fig. 2, refer to the time interval between adjacent peaks and feet of waveform. Instantaneous heart rate is the reciprocal of pulse interval. Both Mean Absolute Error (MAE) and Rooted Mean Square Error (RMSE) are calculated for these three indexes, in which

$$MAE = \frac{1}{N}\sum_{i=1}^{N}\left|Ind_{ref}(i) - Ind_{pred}(i)\right| \tag{4}$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}[Ind_{ref}(i) - Ind_{pred}(i)]^2} \tag{5}$$

where N is the number of the index, $Ind_{ref}(i)$ and $Ind_{pred}(i)$ represent the i^{th} component of reference and predicted index.

B. Comparison Results

Three state-of-the-art methods are selected in this study for comparison, which are based on second-order Volterra Filter [11], Singular Spectrum Analysis (SSA) [12], and RLS adaptive filter plus SSA [13]. These methods are selected because they have the procedures of waveform reconstruction, thus can be used for extracting PRV. The performance indexes are calculated for every test subject and averaged, shown as mean ± std.

TABLE II

PERFORMANCE EVALUATION: COMPARE WITH STATE-OF-THE-ART METHODS (38 SUBJECTS AVERAGE)

PERFORMANCE INDEX (MEAN ± STD)		SECOND-ORDER VOLTERRA FILTER [11]	SSA [12]	RLS FILTER + SSA [13]	RNN - 2 LAYERS	RNN - 3 LAYERS	RNN - 4 LAYERS
Sensitivity (%)		89.3 ± 6.49	91.1 ± 5.83	87.2 ± 6.30	99.3 ± 1.36	**99.4 ± 1.17**	99.3± 1.84
Precision (%)		92.3 ± 2.51	91.6 ± 3.18	89.4 ± 2.95	98.6 ± 1.70	98.9 ± 1.35	**99.1 ± 1.11**
Peak Interval (ms)	*MAE*	97.6 ± 22.7	82.6 ± 24.3	110 ± 22.3	26.5 ± 16.2	**24.2 ± 13.2**	24.2 ± 13.7
	RMSE	179 ± 23.9	164 ± 28.6	198 ± 24.0	59.6 ± 32.1	**54.0 ± 26.5**	56.3 ± 29.1
Pulse Interval (ms)	*MAE*	79.2 ± 14.9	64.8 ± 16.3	90.1 ± 15.2	20.6 ± 11.8	18.7 ± 9.20	**18.2 ± 9.80**
	RMSE	161 ± 18.7	141 ± 24.6	174 ± 19.4	52.5 ± 27.8	45.6 ± 21.9	**44.8 ± 22.8**
Instantaneous HR (Beats/min)	*MAE*	8.68 ± 1.73	7.28 ± 2.29	9.75 ± 1.66	2.20 ± 1.33	2.03 ± 1.20	**1.91 ± 1.28**
	RMSE	26.5 ± 11.0	20.9 ± 5.26	26.5 ± 7.55	6.84 ± 3.70	5.87 ± 3.88	**5.10 ± 3.03**

The comparison results between state-of-the-art methods and the proposed algorithm (with different network layers) are illustrated in Table II. It is clear that deep RNN, even with only 2 layers, outperforms all the state-of-the-art methods in the accurate detection and predict of pulse occurrence, showing extreme high sensitivity and precision ratio. The proposed algorithm also has small standard deviation, indicating good consistency among subjects.

The MAE and RMSE of peak and pulse intervals further prove deep RNN's advantage in fiducial point detection and cardiac interval estimation, which is certainly beneficial for accurate HRV evaluation during exercise. In contrast, the MAE for interval estimation introduced by other methods can be up to 90ms, which is of little clinical meaning because the error is larger than the fluctuation range of cardiac intervals (usually smaller than 80ms [3]).

The instantaneous HR, different from average HR, illustrates the beat-to-beat changes of cardiac frequencies, thus may change rapidly compared with average HR. Performances of state-of-the-art methods on instantaneous heart rate estimation are not as satisfied as those on average heart rate shown in previous studies, partly because these methods have not focused on the recovery of cardiac period and only extracted HR from the frequency domain. On the other hand, the proposed RNN shows remarkable performance in IHR recovery with MAE smaller than 2 Beats/min.

Increasing the depth of network from 1 layer to 3 layers will result in better performance, whereas limited improvements on performance may show on networks deeper than 4 layers, partly because of the small sample size.

IV. CONCLUSION

In this work, a deep RNN based on bidirectional Long-Short Term Memory Network (biLSTM) is proposed for accurate PPG cardiac period segmentation and Pulse Rate Variability (PRV) estimation under strenuous physical exercise where PPG waveforms are contaminated by strong motion artifacts. Comparison results with state-of-the-art methods on a dataset including 48 subjects show the promising performance of the proposed algorithm in PRV indexes estimation and recovery. To our best knowledge, this is the first time a deep learning-based algorithm been involved for extraction of PRV from seriously corrupted PPG signals.

REFERENCES

[1] U. Rajendra Acharya, K. Paul Joseph, N. Kannathal, C. M. Lim, and J. S. Suri, "Heart rate variability: a review," *Med. Biol. Eng. Comput.*, vol. 44, no. 12, pp. 1031–1051, Dec. 2006.

[2] K. L. Dodds, C. B. Miller, S. D. Kyle, N. S. Marshall, and C. J. Gordon, "Heart rate variability in insomnia patients: A critical review of the literature," *Sleep Med. Rev.*, vol. 33, pp. 88–100, Jun. 2017.

[3] S. Lu *et al.*, "Can Photoplethysmography Variability Serve as an Alternative Approach to Obtain Heart Rate Variability Information?," *J. Clin. Monit. Comput.*, vol. 22, no. 1, pp. 23–29, Jan. 2008.

[4] E. Gil, M. Orini, R. Bailón, J. M. Vergara, L. Mainardi, and P. Laguna, "Photoplethysmography pulse rate variability as a surrogate measurement of heart rate variability during non-stationary conditions," *Physiol. Meas.*, vol. 31, no. 9, pp. 1271–1290, Aug. 2010.

[5] B. Vescio, M. Salsone, A. Gambardella, and A. Quattrone, "Comparison between Electrocardiographic and Earlobe Pulse Photoplethysmographic Detection for Evaluating Heart Rate Variability in Healthy Subjects in Short- and Long-Term Recordings," *Sensors*, vol. 18, no. 3, p. 844, Mar. 2018.

[6] A. Schäfer and J. Vagedes, "How accurate is pulse rate variability as an estimate of heart rate variability?: A review on studies comparing photoplethysmographic technology with an electrocardiogram," *Int. J. Cardiol.*, vol. 166, no. 1, pp. 15–29, Jun. 2013.

[7] Z. Zhang, Z. Pi, and B. Liu, "TROIKA: A General Framework for Heart Rate Monitoring Using Wrist-Type Photoplethysmographic Signals During Intensive Physical Exercise," *IEEE Trans. Biomed. Eng.*, vol. 62, no. 2, pp. 522–531, Feb. 2015.

[8] Z. Zhang, "Photoplethysmography-Based Heart Rate Monitoring in Physical Activities via Joint Sparse Spectrum Reconstruction," *IEEE Trans. Biomed. Eng.*, vol. 62, no. 8, pp. 1902–1910, Aug. 2015.

[9] S. M. A. Salehizadeh, D. Dao, J. Bolkhovsky, C. Cho, Y. Mendelson, and K. H. Chon, "A Novel Time-Varying Spectral Filtering Algorithm for Reconstruction of Motion Artifact Corrupted Heart Rate Signals During Intense Physical Activities Using a Wearable Photoplethysmogram Sensor," *Sensors*, vol. 16, no. 1, Dec. 2015.

[10] P. Ghosal and R. Gupta, "Classification of photoplethysmogram signal using self organizing map," in *2015 IEEE International Conference on Research in Computational Intelligence and Communication Networks (ICRCICN)*, 2015, pp. 114–118.

[11] S. Kim, S. Im, and T. Park, "Characterization of Quadratic Nonlinearity between Motion Artifact and Acceleration Data and its Application to Heartbeat Rate Estimation," *Sensors*, vol. 17, no. 8, p. 1872, Aug. 2017.

[12] Y. Ye, W. He, Y. Cheng, W. Huang, and Z. Zhang, "A Robust Random Forest-Based Approach for Heart Rate Monitoring Using Photoplethysmography Signal Contaminated by Intense Motion Artifacts," *Sensors*, vol. 17, no. 2, p. 385, Feb. 2017.

[13] Y. Ye, Y. Cheng, W. He, M. Hou, and Z. Zhang, "Combining Nonlinear Adaptive Filtering and Signal Decomposition for Motion Artifact Removal in Wearable Photoplethysmography," *IEEE Sens. J.*, vol. 16, no. 19, pp. 7133–7141, Oct. 2016.