Using PPG and Machine Learning to Measure Blood Pressure

G. S. Cardoso, M. G. Lucas, S. S. Cardoso, J. C. M. Ruzicki
and A. A. S. Junior

Abstract

Measurement of blood pressure is critical for patients with cardiovascular disease. Using a cuff sphygmomanometer with non-invasive registration is currently the most common practice since its fast and does not require an experienced operator. However, it does not allow continuous pressure monitoring and the discomfort of the procedure can discourage consistent use. Photoplethysmography (PPG) is a technique increasingly used for non-invasive, portable devices to monitor arterial oxygen saturation (SpO_2), heart rate and, more recently, for glycemic control. In this study, we evaluate different methods to estimate blood pressure using PPG. Two methods presented are adapted from the literature, while the third is an improvement proposal. Multiple linear regression (MLR), artificial neural network (ANN), support vector machine regression (SVR) and decison tree regression (DTR) using temporal and spectral PPG features are evaluated. Principal component analysis (PCA) is used in order to reduce dimensionality. The MIMIC (Multiparameter Intelligent Monitoring in Intensive Care) database is used to train and evaluate the approaches. Results indicate that the proposal improves diastolic (DBP) and systolic (SBP) blood pressure estimation with mean absolute errors (MAE) of 6.52 ± 5.75 mmHg and 13.19 ± 11.90 mmHg, respectively.

G. S. Cardoso
Communications Department (DECOM/FEEC), UNICAMP, Campinas, Brazil

M. G. Lucas · S. S. Cardoso
Biomedical Engineering Institute (IEB), UFSC, Florianópolis, Brazil

J. C. M. Ruzicki · A. A. S. Junior (✉)
Department of Electrical Engineering, Sul-Rio-Grandense Federal Institute (IFSul), Praça Vinte de Setembro, 455 Pelotas, Brazil

1 Introduction

Primary hypertension is an intermediate risk factor for cardiovascular disease (CVD) and affects more than thirty million people in Brazil [1]. The treatment for this incurable, chronic disease consists in lifestyle changes, use of antihypertensive medications as well as daily blood pressure monitoring [2]. Hypertension is also one of the major comorbities and an important risk factor for COVID-19 mortality [3]. Common methods used for the measurement of arterial blood pressure (BP) are the auscultatory and oscillometric, which estimate maximum (systolic) and minimum (diastolic) BP pressure [2]. However, these methods have some disadvantages. To perform this procedure, artery must be occluded to interrupt the blood flow in the measurement region, causing discomfort to the patient [4,5]. It also precludes online and continuous monitoring, which, when needed, often use more invasive techniques [6].

Due to the limitations of the traditional blood pressure measurement, several non-invasive methods are being researched to provide greater comfort and safety for continuous pressure monitoring [6]. Among the biological signals with potential for continuous and non-invasive pressure monitoring applications, photoplethysmography has shown great promise since PPG signals are non-invasive and require simple optical sensors (see Fig. 1).

Given the widespread use of PPG signals in HR (Heart Rate) and SpO_2 monitoring [7], and, even glicemic control [8], it is also a very robust and well known technology. Thus, several groups have been studying the application of PPG signals to estimate blood pressure. Kachuee et al. [9] demonstrate a technique called pulse transit time (PTT), which uses PPG and electrocardiography (ECG) signals to estimate aver-
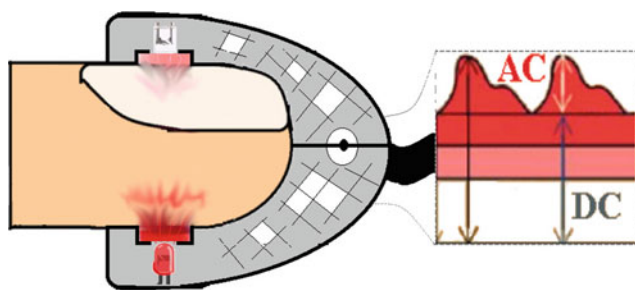
Fig. 1 Photoplethysmography (PPG) technique

age BP within an 7.52 ± 9.54 mmHg mean absolute error (MAE) range. Also, Nabeel et al. [10] shows that one can use two PPG signals acquired from different parts of the body to assess the average BP within an 0.96 ± 6.77 mmHg mean error (ME) range in a method called pulse wave velocity (PWV). Although these methods have achieved interesting results they also add implementation complexity due the use of two sensors and the performance sensitivity to synchronization issues between the inputs. Also, when the method also requires using a signal other than PPG (f.i. ECG is required in the PTT method) this furthers increases the hardware for signal conditioning.

To reduce hardware complexity some authors have proposed algorithms using parameter extracted from a single PPG signal. Methods using features in time [11,12] and frequency [13] domains, or both [14], have been explored. Among these studies, [11,13] have results within the standards proposed for BP measurement [15]. However, in similar studies such as [14,16,17], found error to be above the requirements to meet the standard. There is room, therefore, to subsequent work that helps clarify this divergence.

This paper aims to reproduce and compare the features extraction methods reported by Kurylyak et al. [11] and Xing and Sun [13] and compare the performance of different machine learning algorithms for BP measurement in order to assist in the investigation of these reported error differences. The method reported by Kurylyak et al. [11] uses time domain features of the PPG signal to train an artificial neural network (ANN) to estimate the BP. In [13], an ANN is trained using frequency domain features of the PPG signal. In addition, in this work, an algorithm is proposed to estimate systolic and diastolic pressures using the principal component analysis (PCA) technique with time and frequency domain features and machine learning classification.

This work is organized as follows: Sect. 2 presents the general methodology. Section 3 describes the experiments and results achieved. Finally, in Sect. 4 a general discussion and the main conclusions are presented.

2 Methodology

Tests were performed to evaluate the impact of the following issues on the BP measurement error: signal feature selection and dimensional reduction of those features; signal range used; and, finally, choice of regression algorithms.

2.1 Database

Data was extracted from the database presented in [9], which is a preprocessed and validated version of the MIMIC II [18], consisting of 3363 signal segments of at least 10 min, sampled at 125 Hz with a minimum resolution of 8 bits. PPG signals were captured at the fingertip, while ECG lead II and an invasive blood pressure signal were recorded simultaneously.

A few signals of this database contain movement artifacts. By visual inspection, these sections were identified and the entire section was discarded from the beginning of the movement artifact to the end of the signal. A total of 147 blocks were discarded in this step, which corresponds to 658 minutes of signal.

Finally, the detection of peaks and valleys was made using the algorithm presented in [19], which separates the PPG signal in segments and classifies each one as positive slope or negative slope. When a transition from a positive slope to a negative slope segment occurs, the maximum value in the segment with the latest positive slope is considered to be the peak value in the signal. Similarly, when a transition from a negative to a positive slope occurs, the minimum value in the segment with the latest negative slope is considered to be the valley value of the signal. An 8 points (64 ms window) constant segmentation was used. This algorithm was chosen due to the simplicity of being implemented in embedded devices. The detection of peaks and valleys is important both for extracting the features in the time domain and for determining the cycle used for the features in the frequency domain.

2.2 Extracted Features

Kurylyak et al. [11] used 21 temporal features of PPG signal to train ANNs (see Fig. 2). The high number of features were used because of the variability of the data between different individuals to the same pressure value due to physiological characteristics of each individual.

Xing and Sun [13] used normalized PPG signals by applying a simplified model of volumetric expansion of arteries in order to reduce the influence of the variability of the elastic
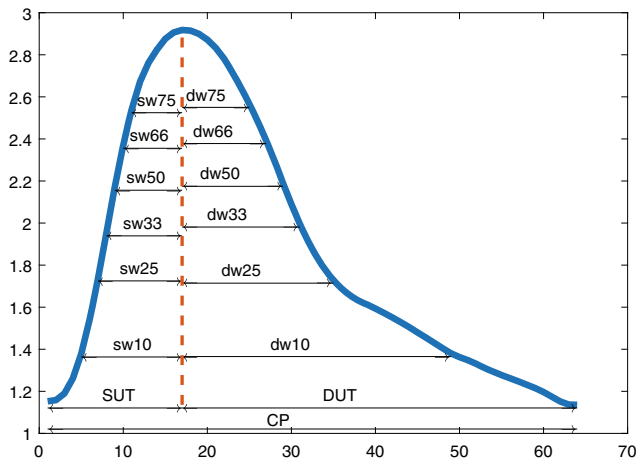
Fig. 2 Extracted features of PPG signal

properties of the arteries between different individuals. After normalization, intervals containing a complete cardiac cycle, 10 % of the previous and 5 % of the following cycle are taken and zeros are added until 150 points are obtained. After that, the fast Fourier transform (FFT) of these intervals were performed and the module/phase information up to 10.8 Hz was used to train an ANN with a hidden layer of 35 neurons and two outputs, one for the SBP and the other for the DBP.

In this work, two configurations to extract features based in [11,13] were used. The first, extracts the 21 time domain features described in [11] while the second applies a normalization and extracts the frequency domain features described in [13]. A normalization different from the described in [13] was used, given by the Eq. 1, which is also very effective according to [20].

$$PPG_{norm} = \frac{PPG_0 - \min(PPG_0)}{\max(PPG_0) - \min(PPG_0)} \qquad (1)$$

2.3   Dimensionality Reduction

To reduce the dimensionality of the extracted features, Pearson correlation coefficient, Relief [21] algorithm and principal component analysis (PCA) are used.

Pearson correlation coefficient and Relief weights were calculated in correlation to each feature and the SBP value. The dimensionality reduction by these two methods was performed considering only the eight characteristics with the highest weights when using Relief or with the highest coefficients when using Pearson.

When using PCA, the components that account for at least 95 % of the variability of the input data were considered.

2.4   Blood Pressure Ranges

Some tests were done training the algorithms using data in the entire pressure range available in the dataset. In the other cases we followed the methodology described in [12], in which the data was divided into three categories according to the SBP values - hypotensive (80–100 mmHg), normotensive (100 to 140 mmHg) and hypertensive (140–180 mmHg). The algorithms were then trained in each of these ranges.

2.5   Regression Algorithms

To estimate systolic and diastolic blood pressures from the features of the PPG signal, four regression algorithms were used: multiple linear regression (MLR), ANNs, support vector machine regression (SVR) and decision tree regression (DTR).

The coefficients for MLR were obtained by the least squares method and the neural network used was a multi-layer perceptron (MLP) with a hidden layer of 35 neurons and two neurons in the output layer (one for SBP and one for DBP). For SVR and DTR algotihms, linear kernel and binary trees were used, respectively.

3   Experiment and Outcomes

The development of the experiment took place through a set of evaluations carried out in the methods described in section II. A total of twenty-four tests were performed, varying different parameters of the method proposed in [11]. Fifteen tests were performed varying the parameters of the method described in [13] and sixteen tests were performed following the method proposed in this article, which combined the time and frequency features of PPG signal The evaluations carried out on these three cases are illustrated in Fig. 3 by the test flow diagrams.

Tables 1 and 2 demonstrate the test results for the methods proposed in [11,13]. The tests were performed according to the test flow diagram described in Fig. 3. In these experiments, tests were performed with different features, different regression algorithms and regressors for different pressure ranges.

All results obtained employ the MAE and standard deviation (SD) of the pressure errors as a performance metric. As already mentioned, for MAE and SD to meet the requirements proposed in [15], they must remain below 5 ± 8 mmHg, respectively.

The best results obtained for the method proposed in [11] for systolic pressure were with the use of ANN, in which a
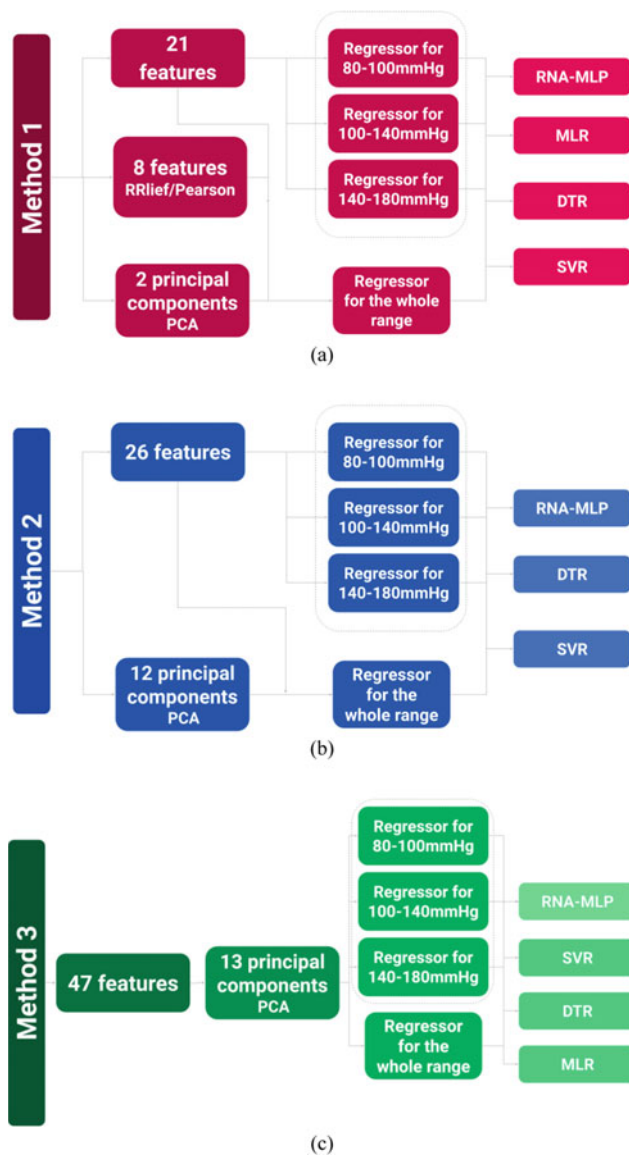
Fig. 3 Test flow diagrams: **a** method 1 corresponding to proposed method in [11]; **b** method 2 corresponding to proposed method in [13] and **c** method 3 corresponding to the proposed method in this work

MAE of 15.58 ± 14.78 mmHg was obtained for SBP and 7.34 ± 7.05 mmHg for DBP. For the method proposed in [13], the best results were 21.97 ± 16.79 mmHg for SBP and 6.97 ± 5.73 for DBP using the SVR regression algorithm, as seen in Table 2. Note that the obtained results for both cases are far from the standards described in [15].

The method proposed in [11] produces the best result for a general test when compared to other regression algorithms. However, the regression algorithms for fractional pressure ranges shown in Table 1 presented superior performance, even in tests with a reduced number of features using the Relief algorithm and also using the principal components given by PCA dimensionality reduction shown in Table 3. Applying

the method proposed in [13], it can be observed that for a test in the general pressure range the ANN regressor is not better than the SVR regression algorithm. In addition, the input feature reduction using PCA to 12 principal components can improve the performance of this method as shown in the Table 1. However, the most promising approach was the one in which the general pressure range was divided into three sub-ranges.

In order to improve the results obtained by the regression algorithms presented in [11,13], a new approach was proposed using the features of both methods with a dimensionality reduction by PCA. In this approach, the parameters were reduced from 47 to 13 principal components. With this approach, a better result was obtained when compared to the tests of the other two methods shown in the Tables 1, 2, 3, and 4, as can be seen in Table 5.

4 Conclusions

In this study, we developed an experiment to reproduce the methods presented in [11,13] and, based on those we propose a new method. A summary of the results obtained in previous works and the results obtained in this study is given in Table 6.

From the results obtained, it was observed that, in general, the proposed method, using the temporal and spectral characteristics of the PPG signal, improves the blood pressure estimation in comparison to the other two methods. In addition, better results were observed in tests in which the algorithms were trained only in certain pressure ranges in comparison to when they were trained in the entire available range. This difference is more evident in the estimation of SBP.

Regarding the reproduced methods, it should be noted that the results obtained in this study agreed with the results obtained in [16,17]. As for the regression algorithms, in general the DTR presented the highest MAE values. The lowest MAE values were obtained using SVR with linear kernel, which indicates a predominantly linear relationship between characteristics and blood pressure values.

Most results are outside the range determined by AAMI, especially those referring to SBP. Despite this, some obtained results for DBP and SBP are within the range of 5 ± 8 mmHg, which demonstrates the feasibility of extracting blood pressure only with the use of PPG signs.

The results indicated that the proposed method could improve the blood pressure estimation by evaluating the temporal and spectral features extracted from the acquired PPG signal. The exploration of additional characteristics of PPG signals and their normalization in order to eliminate differences in these signals due to characteristics such as variability in the elasticity of arteries among individuals should be explored in future works, with the potential to obtain better results.

Table 1 Results for the method proposed in [11] according to the pressure range

Range	Method	DBP (mmHg)	SBP (mmHg)
Full range	MLR	7.57 ± 6.36	16.98 ± 12.53
	DTR	9.68 ± 10.20	19.34 ± 17.84
	RNA	7.34 ± 7.05	15.58 ± 14.78
	SVR	7.19 ± 6.82	16.81 ± 13.48
80–100 mmHg	MLR	6.65 ± 6.16	5.63 ± 3.42
	DTR	7.20 ± 5.67	4.39 ± 4.24
	RNA	5.36 ± 6.29	5.63 ± 3.75
	SVR	$6, 12 \pm 5.94$	5.58 ± 3.22
100–140 mmHg	MLR	6.51 ± 5.90	9.80 ± 6.64
	DTR	8.57 ± 11.65	11.58 ± 9.43
	RNA	6.49 ± 5.78	8.49 ± 6.19
	SVR	5.68 ± 5.73	10.45 ± 6.84
140–180 mmHg	MLR	8.25 ± 7.75	9.17 ± 5.44
	DTR	8.38 ± 8.88	9.73 ± 6.96
	RNA	7.63 ± 6.57	8.78 ± 5.49
	SVR	7.52 ± 6.94	8.82 ± 5.25

Table 2 Results for the method proposed in [13] according to the pressure range

Range	Method	DBP (mmHg)	SBP (mmHg)
Full range	DTR	10.05 ± 9.84	24.89 ± 21.51
	RNA	9.16 ± 7.43	22.08 ± 16.74
	SVR	6.51 ± 5.73	21.97 ± 16.79
80–100 mmHg	DTR	8.46 ± 7.86	6.28 ± 4.61
	RNA	6.68 ± 7.13	5.47 ± 2.87
	SVR	7.61 ± 7.65	6.06 ± 3.50
100–140 mmHg	DTR	10.51 ± 12.20	10.72 ± 8.32
	RNA	7.71 ± 5.85	8.89 ± 6.34
	SVR	4.54 ± 3.48	8.41 ± 6.02
140–180 mmHg	DTR	11.31 ± 11.06	13.16 ± 9.15
	RNA	10.21 ± 10.00	11.09 ± 8.48
	SVR	6.67 ± 5.86	11.44 ± 8.63

Table 3 Results for the method proposed in [11] according to the number of features

# caract.	Method	DBP (mmHg)	SBP (mmHg)
8	MLR	7.80 ± 6.50	19.75 ± 12.88
	DTR	8.67 ± 9.50	20.74 ± 19.32
	RNA	7.45 ± 6.77	14.81 ± 13.23
	SVR	7.00 ± 6.81	19.86 ± 13.13
2 (PCA)	MLR	8.26 ± 6.20	22.68 ± 12.60
	DTR	9.29 ± 8.92	23.54 ± 19.25
	RNA	8.03 ± 6.55	17.87 ± 13.34
	SVR	7.30 ± 6.76	21.49 ± 12.35

Table 4 Results for the method proposed in [13] using 12 principal components

Method	DBP (mmHg)	SBP (mmHg)
DTR	9.81 ± 9.30	24.68 ± 20.80
RNA	7.16 ± 5.81	21.19 ± 17.04
SVR	6.54 ± 5.77	21.99 ± 16.94

Table 5 Results for the proposed method using both time and frequency domains features according to the pressure range

Range	Method	DBP (mmHg)	SBP (mmHg)
Full range	MLR	6.89 ± 5.34	15.25 ± 12.86
	DTR	9.46 ± 8.66	15.82 ± 14.93
	RNA	6.68 ± 5.18	13.58 ± 10.53
	SVR	6.52 ± 5.75	13.19 ± 11.90
80–100 mmHg	MLR	6.56 ± 5.33	5.12 ± 4.13
	DTR	6.79 ± 7.15	4.42 ± 3.71
	RNA	4.85 ± 6.52	2.29 ± 2.52
	SVR	4.95 ± 6.00	4.85 ± 4.60
100–140 mmHg	MLR	7.04 ± 5.22	7.50 ± 6.45
	DTR	9.14 ± 12.87	7.56 ± 7.23
	RNA	5.03 ± 5.87	6.58 ± 5.46
	SVR	4.92 ± 5.51	6.76 ± 6.49
140–180 mmHg	MLR	9.33 ± 7.39	8.60 ± 5.80
	DTR	10.72 ± 9.74	9.63 ± 6.78
	RNA	7.63 ± 8.15	6.18 ± 4.37
	SVR	6.63 ± 7.36	6.29 ± 5.53

Table 6 Comparison of the results of this paper with other works

Work	Signal	Error	SBP (mmHg)	DBP (mmHg)
[9]	ECG+PPG	MAE ± SD	12.38 ± 16.17	6.34 ± 8.45
[10]	PPG	ME ± SD	1.15 ± 7.98	0.86 ± 6.36
[11]	PPG	MAE ± SD	3.80 ± 3.46	2.21 ± 2.09
[12]	PPG	ME ± SD	−0.1 ± 6.5	−0.6 ± 5.2
[13]	PPG	ME ± SD	0.06 ± 7.8	0.01 ± 4.66
[14]	PPG	MAE	9.43	6.88
[20]	PPG	MAE ± SD	3.97 ± 7.99	2.43 ± 3.37
[22]	PPG	MAE	4.47	3.21
[23]	PPG	MAE ± SD	11.64 ± 8.20	7.62 ± 6.78
[16]	PPG	MAE ± SD	8.54 ± 10.9	4.34 ± 5.8
[17]	PPG	MAE ± SD	1.87 ± 7.20	0.79 ± 3.40
Our	PPG	MAE ± SD	13.19 ± 11.90	6.52 ± 5.75

References

1. Malachias MVB, Souza WKSB, Plavnik FL et al (2016) 7th Brazilian guideline of arterial hypertension. Arq Bras Cardiol 107:1–83
2. Brasil (2006) Systemic arterial hypertension for the unified health system. Ministério da Saúde, Brasília
3. Wang B, Li R, Lu Z, Huang Y (2020) Does comorbidity increase the risk of patients with COVID-19: evidence from meta-analysis. Aging (Albany NY) 12:6049–6057
4. Hall JE (2017) Guyton e Hall tratado de fisiologia médica. Elsevier Brasil, Rio de Janeiro
5. Shioya K, Dohi T (2013) Blood pressure measurement device based on the arterial tonometry method with micro triaxial force sensor. In: 2013 Transducers & Eurosensors XXVII: the 17th international conference on solid-state sensors, actuators and microsystems (TRANSDUCERS & EUROSENSORS XXVII), pp 2389–2392

[6] Meidert AS, Saugel B (2018) Techniques for non-invasive monitoring of arterial blood pressure. Front Med 4:231

[7] Webster JG (1997) Design of pulse oximeters. CRC Press, New York

[8] Rosenthal RD, Paynter LN, Mackie LH (1991) Non-invasive measurement of blood glucose 1991. US Patent 5,028,787

[9] Kachuee M, Kiani MM, Mohammadzade H, Shabany M (2015) Cuff-less high-accuracy calibration-free blood pressure estimation using pulse transit time. In: 2015 IEEE international symposium on circuits and systems (ISCAS), pp 1006–1009

[10] Nabeel PM, Jayaraj J, Mohanasankar S (2017) Single-source PPG-based local pulse wave velocity measurement: a potential cuffless blood pressure estimation technique. Physiol Measure 38:2122–2140

[11] Kurylyak Y, Lamonaca F, Grimaldi D (2013) A neural network-based method for continuous blood pressure estimation from a PPG signal. In: 2013 IEEE international instrumentation and measurement technology conference (I2MTC), pp 280–283

[12] Khalid SG, Zhang J, Chen F, Zheng D (2018) Blood pressure estimation using photoplethysmography only: comparison between different machine learning approaches. J Healthcare Eng

[13] Xing X, Sun M (2016) Optical blood pressure estimation with photoplethysmography and FFT-based neural networks. Biomed Opt Express 7:3007–3020

[14] Slapničar G, Mlakar N, Luštrek M (2019) Blood pressure estimation from photoplethysmogram using a spectro-temporal deep neural network. Sensors 19:3420

[15] AAMI (1987) American national standards for electronic or automated sphygmomanometers. In: ANSI/AAMI SP 10-1987

[16] Liu M, Po LM, Fu H (2017) Cuffless blood pressure estimation based on photoplethysmography signal and its second derivative. Int J Comput Theory Eng 9:202–206

[17] Bersano N, Sanson H (2018) Non-invasive blood pressure estimation from photoplethysmography signals using artificial neural networks. In: 2018 20th International conference on advanced communication technology (ICACT), pp 29–35

[18] Goldberger AL, Amaral LAN, Glass L et al (2000) PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals. Circulation 101:e215–e220

[19] Kavsaoğlu AR, Polat K, Bozkurt MR (2016) An innovative peak detection algorithm for photoplethysmography signals: an adaptive segmentation method. Turkish J Electr Eng & Comput Sci 24:1782–1796

[20] Mousavi SS, Firouzmand M, Charmi M, Hemmati M, Moghadam M, Ghorbani Y (2019) Blood pressure estimation from appropriate and inappropriate PPG signals using A whole-based method. Biomed Signal Process Control 47:196–206

[21] Robnik-Šikonja M, Kononenko I (1997) An adaptation of relief for attribute estimation in regression. In: Machine learning: proceedings of the fourteenth international conference (ICML'97), vol 5, pp 296–304

[22] Gaurav A, Maheedhar M, Tiwari VN, Narayanan R (2016) Cuffless PPG based continuous blood pressure monitoring—A smartphone based approach. In: 2016 38th Annual international conference of the IEEE engineering in medicine and biology society (EMBC), pp 607–610

[23] Zhang Y, Feng Z (2017) A SVM method for continuous blood pressure estimation from a PPG signal. In: Proceedings of the 9th international conference on machine learning and computing, New York, NY, USA, pp 128–132