Computers in Biology and Medicine 113 (2019) 103392

Contents lists available at ScienceDirect

Computers in Biology and Medicine

journal homepage: www.elsevier.com/locate/compbiomed

SVR ensemble-based continuous blood pressure prediction using multi-channel photoplethysmogram

Mark Wong Kei Fong, E.Y.K Ng*, Kenneth Er Zi Jian, Tan Jen Hong

School of Mechanical and Aerospace Engineering, Nanyang Technological University, 639798, Singapore

ARTICLE INFO

Keywords:
Blood pressure
Cuff-less
Ensemble learning
Photoplethysmogram
Support vector regression

ABSTRACT

In this paper, a continuous non-occluding blood pressure (BP) prediction method is proposed using multiple photoplethysmogram (PPG) signals. In the new method, BP is predicted by a committee machine or ensemble learning framework comprising multiple support vector regression (SVR) machines. The existing methods for continuous BP prediction rely on a single calibration model obtained from a single arterial segment. Our ensemble framework is the first BP estimation method which uses multiple SVR models for calibration from multiple arterial segments. This permits reducing of the mean prediction error and the risk of overfitting associated with a single model. Each SVR in the ensemble is trained on a comprehensive feature set that is constructed from a distinct PPG segment. The feature set includes pulse morphological parameters such as systolic pulse amplitude and area under the curve, heart rate variability (HRV) frequency, time domain parameters and the pulse wave velocity (PWV). Empirical evaluation using 40 volunteers with no serious health conditions shows that the proposed method is more reliable for estimating both the systolic and diastolic BP than similar methods employing a single calibration model under identical settings. Moreover, the combined output is found to be more stable than the output of any of the constituent models in the ensemble for both the systolic and diastolic cases.

1. Introduction

Hypertension or High blood pressure (High-BP) is a common problem in adults, the condition increases the risk of stroke and myocardial infarction, among other diseases [1–5]. Apart from indicating potential serious health problems, an adult's blood pressure measurement contains a significant amount of information about the individual's physical attributes and helps physicians to plan a healthy diet and exercise regime for him/her. Therefore, it is advised to have blood pressure checked on a regular basis to avoid unnecessary ailments. Auscultatory and oscillometric methods are the most commonly used non-invasive BP measurement techniques [2–4]. Both methods involve constriction of arterial blood flow in the arm using an inflatable sphygmomanometer brachial cuff. In the auscultatory method, air is pumped to inflate the brachial cuff in the upper arm till the brachial artery is absolutely occluded. The cuff is then gradually deflated till the first Korotkoff sound is heard using the stethoscope. This is the sound heard when blood just starts to flow through the artery. The pressure recorded at this point is considered as the systolic pressure. Finally, the cuff pressure is further released until no sound can be heard. The pressure at this point corresponds to the diastolic pressure. For the oscillometric method, the approach differs with the use of a calibrated electronic pressure transducer to detect the pressure oscillations within the cuff during constricted blood flow instead of using the stethoscope. Although the oscillometric method has become more prevalent with caregivers and professionals in recent years due to convenience, it is less accurate than the auscultatory method in general. As these methods are non-invasive, simple, and safe, they are well supported and have been accepted as standard methods for both clinical as well as out-of-clinical BP measurements. However, note that both auscultatory and oscillometric methods measure BP by occluding the artery so the measurement protocol is not convenient for rapid measurements.

Although the brachial cuff-based approach is currently the clinically accepted standard for BP monitoring, it is not appropriate for certain situations where continuous blood pressure monitoring is desired [4–7]. For example, cuff-based blood pressure devices cannot be used for observing sleep patterns without disturbing the patient since the pressure inflation of the brachial or wrist cuff leads to an awakening arousal going along with an increase in the systemic BP [6]. On the other hand, an intra-arterial cannula or artery implanted pressure sensor can measure BP continuously; however, it must be administered by skilled phlebotomists or physicians in an acute or critical care

* Corresponding author.
E-mail addresses: mark_wong@sp.edu.sg (M.W. Kei Fong), mykng@ntu.edu.sg (E.Y.K. Ng), Kenneth_er@sp.edu.sg (K. Er Zi Jian), isstjh@nus.edu.sg (T.J. Hong).

https://doi.org/10.1016/j.compbiomed.2019.103392
Received 14 May 2019; Received in revised form 16 August 2019; Accepted 17 August 2019
Available online 19 August 2019
0010-4825/

hospital setting, moreover the execution of the cannula requires following long and strict protocols to prevent complications [8]. In order to maintain a daily exercise regime and healthy lifestyle, an out-of-clinic continuous BP measuring device is desired which avoids occlusion of arteries associated with cuff-based methods and tedious protocols associated with arterial cannula methods.

In the last few years, various continuous and non-occluding blood pressure monitoring methods have been proposed [8–14]. A common approach is based on pulse transit time (PTT), and/or pulse wave velocity (PWV) [6,8–13]. PTT is defined as the time taken by the arterial blood pulse wave to travel between two arterial sites of measurement within the same cardiac cycle of a single systole-diastole event. There is academic consensus that the PTT value has been closely correlated to arterial stiffness of an individual and can be utilized as a surrogate for blood pressure estimation [11,12]. Notable BP estimation models based on PTT varies from linear to logarithmic to inverse of squares in literature [6,9,10]. PTT is usually determined by calculating the time difference between the QRS-complex or R-wave of the recorded electrocardiogram (ECG) signal and the characteristic point such as the pulse foot or peak on the peripheral pulse of a photoplethysmogram (PPG) waveform [6].

However, ECG may not be convenient in various situations because it is an electrical based measurement and the subject is asked to remove jewellery and metallic objects, but metallic and dental implants have also been proven to affect an ECG signal adversely. Also, the measurement may require consumables such as Ag/AgCl (silver-chloride) electrodes. While PPG, an optically obtained plethysmogram, does not suffer from such shortcomings. Therefore, some PTT based BP measurement methods have been proposed using just the PPG sensors instead of a combination of ECG and PPG sensors [13–15].

Recent studies have shown that decreased heart rate variability (HRV) is also related to hypertension [16–19] i.e. increase in blood pressure. HRV parameters, both time domain and frequency domain [20], are usually derived by calculating R-R interval from the ECG signal [21]. The HRV obtained from the R-R interval of an ECG signal is strongly correlated to the pulse rate variability (PRV) from PPG [22–24]. Consequently, PRV parameters obtained from the PPG signal contain information related to BP changes. It was pointed out as a future work in Ref. [13] that PTT combined with HRV may result in BP estimation improvement. Moreover, it was observed in Ref. [14] that a particular PPG parameter i.e. pulse amplitude is also a strong surrogate for BP estimation. In fact, some works suggest that [25,26] pulse amplitude (PA) is a more accurate surrogate than PTT and the combination of both PA and PTT generally obtains the best results. Another recent study [27] shows that the PPG pulse parameters described in Ref. [28], including PA, can be applied to accurate BP estimation using a machine learning method. In Ref. [29], PTT is used with heart rate without any other HRV parameter. In Ref. [30], HRV parameters from PPG are adapted to predict BP using machine learning methods. A comprehensive feature set which combines PTT, HRV parameters and PPG pulse parameters has not been used for BP prediction. At this stage, it is important to point out that all the methods discussed above use a single calibration model such as a linear regression method for estimating BP from the surrogate parameters.

There are some drawbacks associated with single regression model in calibration such as unstable predictions and risk of overfitting [30]. In machine learning community, committee machine or an ensemble of learning models, with each model trained on different feature set or different subset of data, has been widely used for accurate and stable predictions [31–33]. An ensemble learning framework consists of multiple weak learners. Each learner makes different errors in prediction given that the weak learners are trained on different parts of data. The performance improvement using an ensemble framework in this application thus depends on the diversity of the weak constituent learners. Therefore, ensemble learning framework comprising diverse regression models can be employed for reliable calibration. The advantage of using an ensemble-based BP calibration is twofold. Firstly, ensemble calibration model has more robust performance than a single calibration model since different models make different errors and the overall error is reduced. Secondly, it reduces the effect of overfitting by calibrating with features from multiple arterial segments, during the same cardiac cycle, rather than just relying on a single arterial segment.

With regards to the mentioned considerations for this paper, we then propose an ensemble calibration framework for the goal of achieving robust continuous BP estimation using a multi-channel PPG device. In the new method, four PPG observation points are considered which constitutes six arterial segments. Therefore, the ensemble framework consists of six regression models, with each model trained using features obtained from a distinct arterial segment. We have used support vector or SVR [34] as the base regression model. The feature set consists of PTT, HRV, time and frequency domain values, and pulse morphological parameters including PA. The final ensemble output is more robust than that obtained from any of the constituent SVR model by reducing the risk of overfitting associated with an individual regression model.

From existing literature with regards to BP estimation and to the best of our knowledge scope, this paper proposes the first ensemble learning or multi-regression-based framework for PPG-only BP calibration. PPG features from multiple arterial segments are used for training multiple calibration models rather than just relying on the features from single segment. Moreover, it is a first method that takes into account comprehensive feature set including HRV parameters, PTT and PPG morphological parameters.

This paper is organized into five sections; Section two first discusses the theory of SVR model and then provides the underlying motivation behind employing ensemble learning approach for calibration. Section three explains the new BP estimation method in detail. Section four reports the experimental results and discussions and finally, Section five concludes the paper.

2. Theory on SVR and ensemble learning

For this section, we briefly examine the support vector machine (SVM) algorithm and then review the ensemble learning or committee machine approach in prediction, pointing out the theoretical justification of its superiority over individual learners.

2.1. The support vector machine (SVM) algorithm

The SVM algorithm works for both classifications as well as for regression applications. The regression or the function approximation version of SVM is referred to as support vector machine-regression (SVR) [34]. SVR is considered as a maximum-margin algorithm. Consider a training dataset $\{x_i, y_i\}, i = 1, ..., N, x_i \in R^d$ is references as a d-dimensional data point and y_i is the specified target outcome. $w \in R^d$ is a normal vector to the hyperplane. The goal in SVR is to solve for the function $f(x)$ that has at most ε deviation from the obtained targets of y_i for all the training data recorded that is as flat as possible and at the same time. Another way of explaining is that errors can be ignored if they are less than ε, but the algorithm will not accept any deviation larger than is specified.

The optimization problem in SVR for linear case is given below described in Ref. [34], we begin by describing the linear function in the form:

$$f(x) = \langle w, x \rangle + b \quad \text{with} \quad w \in , b \in R \tag{1}$$

To achieve a small w can be achieved by minimizing the normal where $||w||^2 = \langle w, w \rangle$, the above equation can then be written in a convex optimization as such:

$$\text{Minimize} \ \frac{1}{2}||w||^2 + C\sum_{i=1}^{N}(\varepsilon_i + \varepsilon_i^*) \tag{2}$$

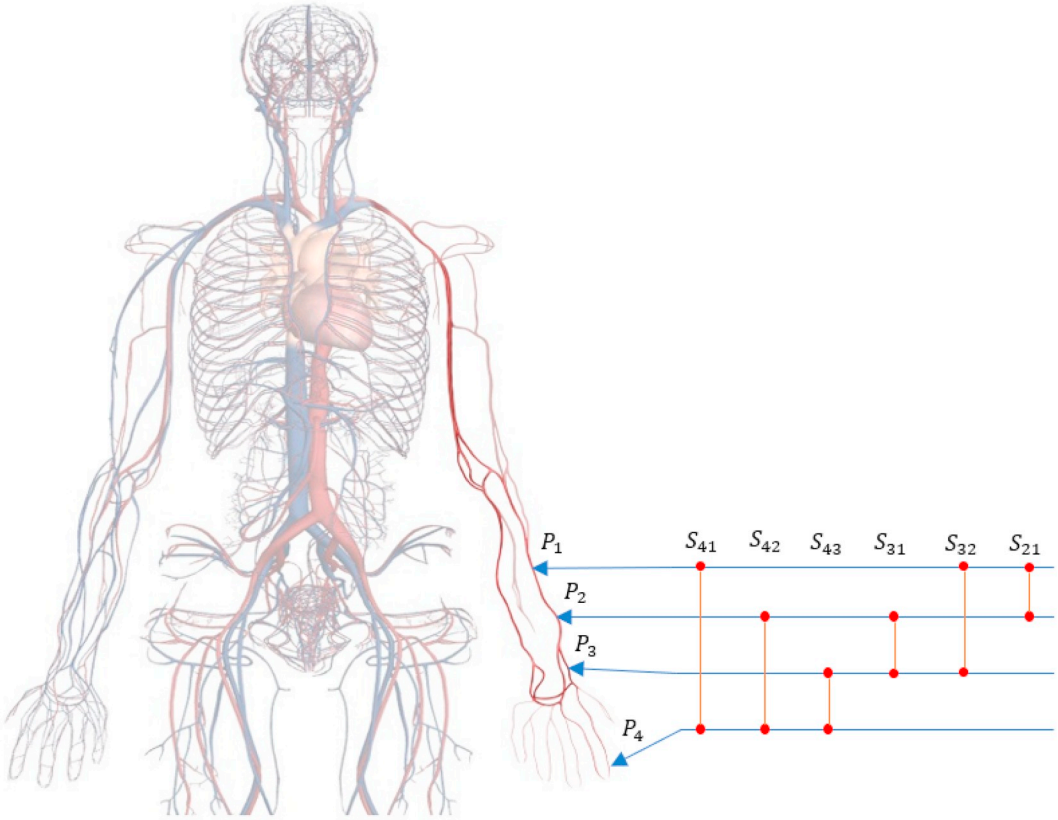

Fig. 1. Four pulse observation points and six corresponding arterial segments.

$$\text{Subject to } \begin{cases} y_i - w, x_i - b \le \varepsilon + \varepsilon_i \\ w, x_i + b - y_i \le \varepsilon + \varepsilon_i^* \\ \varepsilon_i, \varepsilon_i^* \ge 0 \end{cases}$$
$$\text{with } , \; w, x_i \in R^d , \; b \in R \tag{3}$$

The above assumption for equation (3) states that $f(x)$ approximates all pairs of (w, x_i) with ε precision. Here ε_i is a positive slack variable corresponding to the ith data point and C is the specified criteria chosen by the user to balance the tradeoff between the flatness of the function and the training errors.

The above optimization problem can be extended to a more general case e.g. in the non-linear SVM case, where the decision function is the non-linear function of the input training data. The data is then mapped to a higher dimension space using a mapping function $\phi(x)$ where it becomes linearly separable. The exact knowledge of the mapping function is not required. Instead, a kernel function $k(x_i, x_j) = \phi(x_i)\cdot\phi(x_j)$ is used to calculate the inner product of vectors $\phi(x_i)$ and $\phi(x_j)$. The output of the SVM algorithm is given by

$$f(x) = \sum_{i=1}^{N} (\alpha_i - \alpha_i^*) k(x_i, x) + b \tag{4}$$

where α_i, α_i^* are Lagrange multipliers. This is known as a *Support Vector expansion*, where $f(x)$ can be completely described as a linear combination of the training patterns x_i. This simplifies the overall function's representation depending only on the number of SVs described [34], allowing the function to be independent of the input space's dimension.

2.2. Ensemble learning

In various clinical scenarios, it is a common practice to seek opinions from multiple doctors who are experts in the field. The final decision, for a particular treatment, is thus generally made by consulting a committee of experts and combining their opinions in some way. In the context of machine learning, committee machine or ensemble learning systems play a similar role [31–33]. Ensemble learning systems generally provide a more reliable prediction than by any constituent learning model. Ensemble learning systems combine several weak hypotheses to formulate a stronger theory and the effectiveness and successfulness of ensemble learning depends heavily on the diversity of the constituent predictors which then make up the ensemble. It is however not possible to rectify a mistake in prediction if all the models make the same error. The diversity means that each model makes different errors on different instances. Diversity can be achieved by using different model parameters, different training subsets, or different feature sets etc. By combining the output of diverse models through an algebraic expression such as calculating the mean value, the total error can be reduced since various errors of the models are averaged out.

Since a future goal to is to enable such computation on a wearable where processing power is limited, in this paper, we use linear SVR as a constituent model in the ensemble as the dataset here also doesn't appear to have a complex spread and linear SVM is sufficient for efficiency (computational as well as predictive) performance reasons.

The bias-variance-covariance decomposition theory is the basis of theoretical justification for superior performance of ensemble over its constituent predictors [31]. Consider the error equation in (2) which is composed of the average bias, plus a term involving the average variance and another term involving the average covariance. The possibility of a negative covariance may result in the decreased of an expected loss of the said ensemble. Hence, low correlated constituent predictors increase the performance of the ensemble, justifying the importance of diversity among different models.

$$E[\bar{f} - y]^2 = bias^2 + \frac{1}{m}var + \left(1 - \frac{1}{m}\right)covar$$
$$bias = \frac{1}{m}\sum_i (E[f_i] - y)$$
$$var = \frac{1}{m}\sum_i E[f_i - E[f_i]]^2$$
$$covar = \frac{1}{m(m-1)}\sum_i \sum_{j \neq i} E[f_i - E[f_i]](f_j - E[f_j]) \quad (5)$$

where m represents the ensemble size, y refers to the target and f_i is the predicted output from the ith model.

3. Materials and methods

We now propose a novel continuous BP estimation method using a multi-channel PPG device. As can be seen from Fig. 1, PPG signals are collected from four observation points i.e. P_1 to P_4 hence there are six corresponding arterial segments i.e. S_{41}, S_{42}, S_{43}, S_{31}, S_{32} and S_{21}. Instead of relying on a single regression model for calibration, an ensemble of SVR models is used with each SVR trained with features obtained from a different segment. The feature set of a particular arterial segment includes HRV parameters, PTT and PPG morphological features obtained from that segment. The differences in the feature set introduces the diversity in the ensemble where results are better generalization and reduces the risk of overfitting associated with existing regression models for calibration. In general, the final model is more robust than any of the constituent model in the ensemble. Fig. 2 shows the current proposed SVR ensemble-based BP estimation framework.

In Section 3.1, we describe the hardware setup and the signal processing steps required for feature extraction. Section 3.2 discusses PPG waveform features such as pulse amplitude (PA), area under the curve (AUC) and augmentation index (AI). Section 3.3 includes the relevant HRV time and frequency domain parameters. Section 3.4 provides the procedure of obtaining PTT values from multiple arterial segments. Finally, Section 3.5 describes the ensemble method to obtain the final prediction.

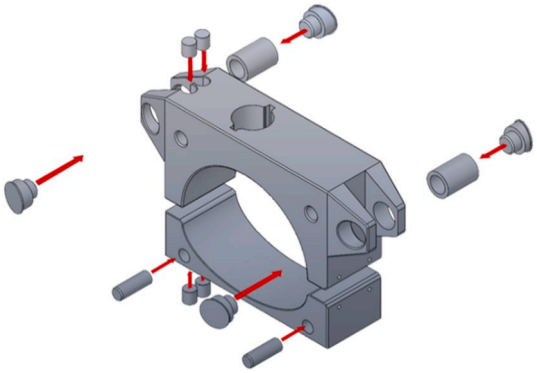

Fig. 3a. Flexible Armbrace Exploded view.

3.1. Hardware setup and signal processing

To simultaneously acquire PPG signals from multiple points on the forearm, we built an armbrace comprising an assembly of clamps (see Fig. 3a and b) and pistons (see Fig. 3c). When clamped down, the brace will be secured with the help of magnets embedded on the ends of both the top and bottom pieces (see Fig. 3d).

The armbrace is very flexible in that the distance between the PPG sensors can be changed (see Fig. 4a) and the heights of the pistons can be adjusted (see Fig. 4b). In this way, one model is used for all the subjects without any noticeable discomfort. The ADS1299 is an analog front-end (AFE) for bio-potential measurements from Texas Instruments, which is a multichannel, low-noise and simultaneous-sampling analog-to-digital converter (ADC) that we used to record multiple PPG signals simultaneously. ADS1299 interface board has a 24-bit sigma-delta (ΔΣ) ADC with an in-built programmable gain amplifier (PGA) that was useful to calibrate analog signal channel gain. Figs. 5 and 6 show the measurement procedure and the block diagram of the

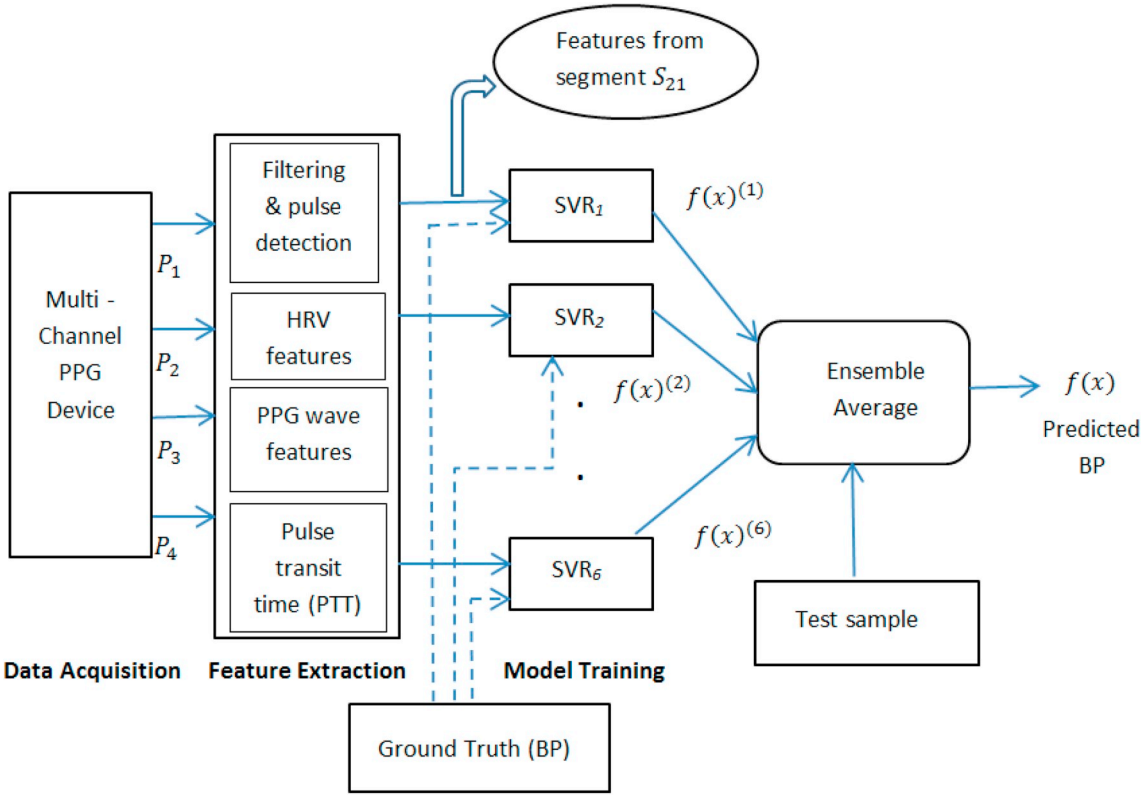

Fig. 2. SVR ensemble framework for BP estimation.

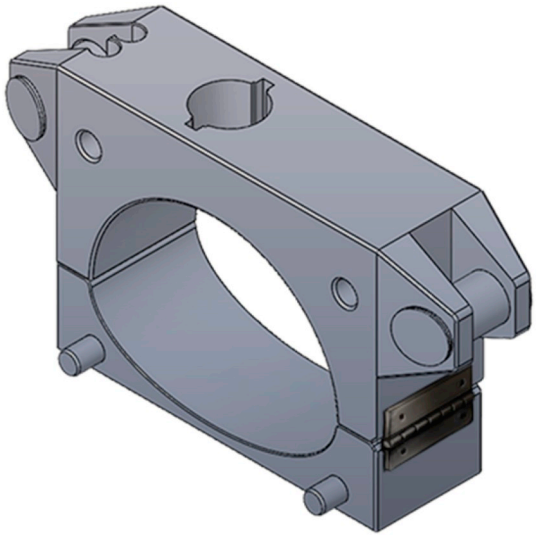

Fig. 3b. Flexible armbrace assembled.

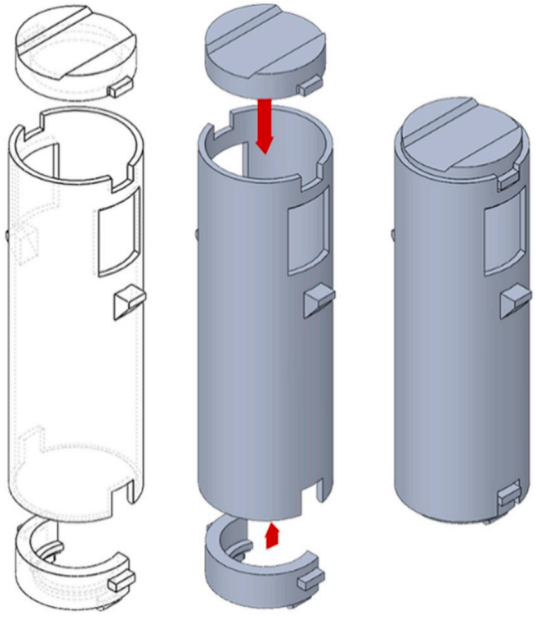

Fig. 3c. Flexible armbrace piston assembly.

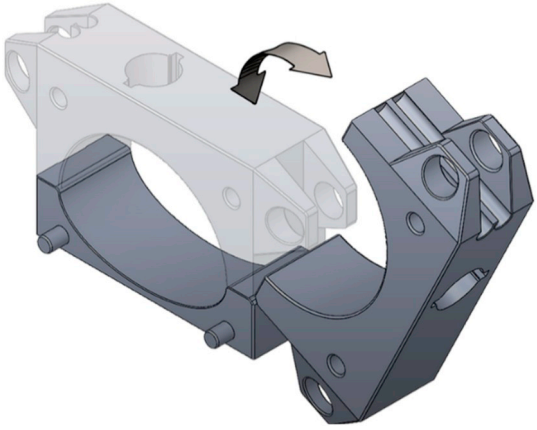

Fig. 3d. Flexible armbrace clamp mechanism.

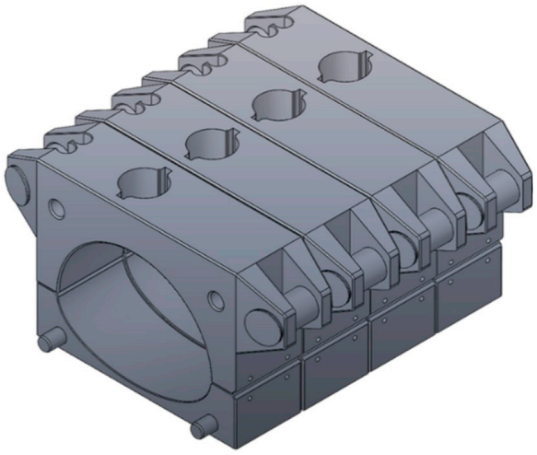

Fig. 4a. Flexible armbrace combined.

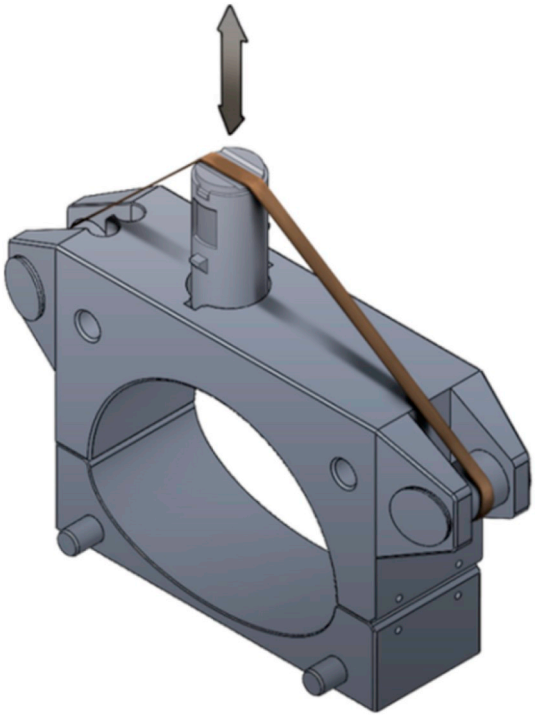

Fig. 4b. Piston Height adjustment.

simultaneous four-channel PPG signal acquisition setup, respectively. All the PPG sensors, attached at the end of each piston (see Fig. 6), are customized in our lab to be used in the armbrace.

The acquired raw signals from the four pulse points are noisy. Therefore, the signals are first passed through Savitzky-Golay smoothing filter with polynomial order 5 and frame size 31. A Finite Impulse Response (FIR) band-bass filter was then used to filter out high-frequency and alternating current (AC) noise from each channel separately. Then, we detect the local peaks in the PPG waveforms where the local peak in this context is the sample point which is greater than its two neighbours. Consequently, if the local peak is greater than a pre-set threshold, it is considered as a valid PPG peak. Later, the corresponding dichroic notch and pulse foot are also identified. A pulse synchronization algorithm is also implemented in MatLab so that peaks from the four channels, corresponding to the same cardiac cycle, are aligned.

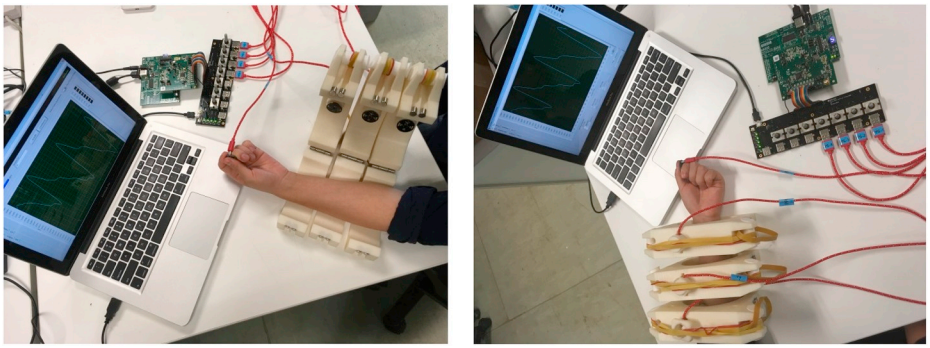

Fig. 5. Four-channel simultaneous PPG data acquisition setup comprising the flexible armbrace with customized PPG sensors and ADS1299 Analog to Digital Convertor.

3.2. PPG pulse features

It has been suggested in various works [14,26–29] that BP has dependence on pulse morphology such as Pulse amplitude, width, area under the curve (AUC) etc. In this work, we calculate systolic pulse amplitude, total area under the PPG curve and augmentation index from each of the four pulse observation points. These parameters are defined below:

Pulse amplitude (PA)

Pulse amplitude, also referred to as the systolic amplitude, is the height of the pulse from foot to systolic peak. In Fig. 7a, X represents the pulse amplitude.

$$PA = X$$

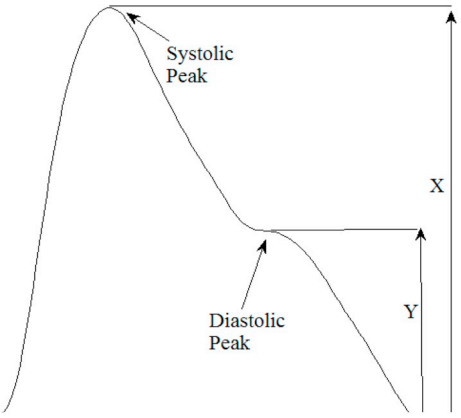

Fig. 7a. Systolic peak amplitude X and diastolic peak amplitudes Y.

Area under the curve (AUC)

In order to calculate the AUC of a single cardiac cycle, PPG curve is obtained by summing the areas A1 and A2 in Fig. 7b, where A1 and A2 are calculated using Trapezoidal numerical integration.

$$AUC = A1 + A2$$

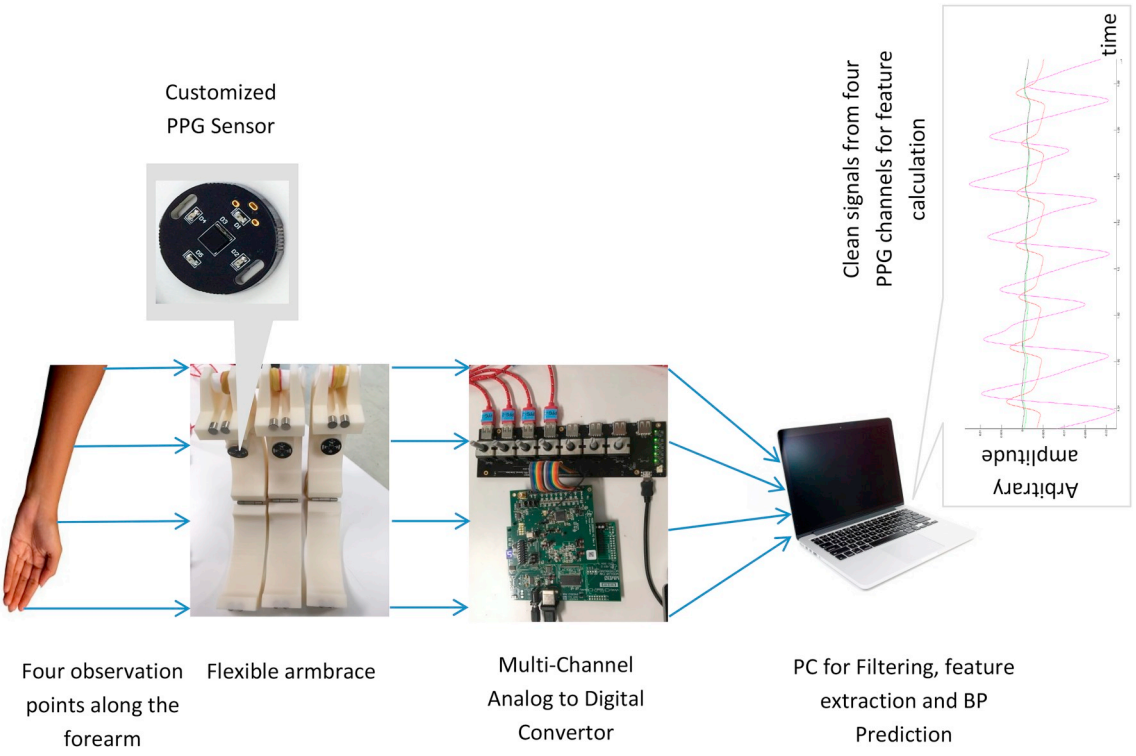

Fig. 6. Block diagram for four-channel PPG data acquisition and processing.

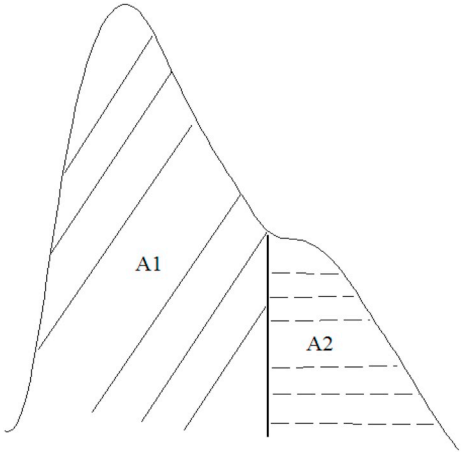

Fig. 7b. A1 and A2 are the areas under the whole PPG wave.

Augmentation index (AI)

The augmentation index (AI) is defined as the ratio of the amplitude of a diastolic peak of the PPG curve to the amplitude the systolic peak of the same waveform [35].

$$AI = Y/X$$

Note that the pulse morphology is different at different pulse points, hence, the corresponding PA, AUC and AI are obtained separately.

3.3. HRV parameters from PPG

It has been suggested in the literature [13,16,18] that heart rate variability (HRV) parameters have some correlation to BP changes. HRV parameters are generally acquired from the R-R intervals from an ECG signal. In Ref. [36], a comparison was done between the HRV by using the Pulse interval in the PPG signal with the HRV of the R-R intervals in a ECG signal and the subsequent results verified that HRV in PPG and ECG signals are firmly correlated. This thus suggests that PPG waveforms could then be obtained as a substitute measurement of ECG for HRV.

In this paper, we hence use numerous time and frequency domain HRV measures with PPG pulse-pulse intervals. For calculating frequency domain parameters, we use autoregressive (AR) model with the bands as follows: Very low frequency (VLF) power in the frequency range of 0.04 Hz, Low frequency (LF) power in the frequency range 0.04–0.15 Hz, and high frequency (HF) power in the frequency range 0.15–0.40 Hz [37]. Kubios software is used to perform HRV analysis [38]. These parameters include mean heart rate, average width of the signal's R-R interval, the standard deviation (SD) of all R-R intervals, root mean square of differences between adjacent R-R intervals and the ratio of LF power to HF power within the signal measure.

3.4. PTT or pulse wave velocity (PWV) using multiple PPG channels

Traditionally, the PTT is obtained by calculating the time differences between each R-wave or QRS-complex of a measured ECG signal and the representative point on the peripheral pulse by PPG such as pulse peak or pulse foot of the systole [39]. In this paper, we calculate PTT using only PPG sensors [13–15]. As shown in Fig. 8, four PPG measurement points i.e. P_1 to P_4 are considered and the six PTT values are determined i.e. $PTT_{1,2}$, $PTT_{1,3}$, $PTT_{1,4}$, $PTT_{2,3}$, $PTT_{2,4}$ and $PTT_{3,4}$.

We then divide the PTT values with the duration of the respective heartbeat segment, this refers to the separation distance between the two corresponding PPG points to get PWV values for the six segments. Note that for simplicity, we assume that the distance between any two PPG points is linear. The distance between any two adjacent PPG points

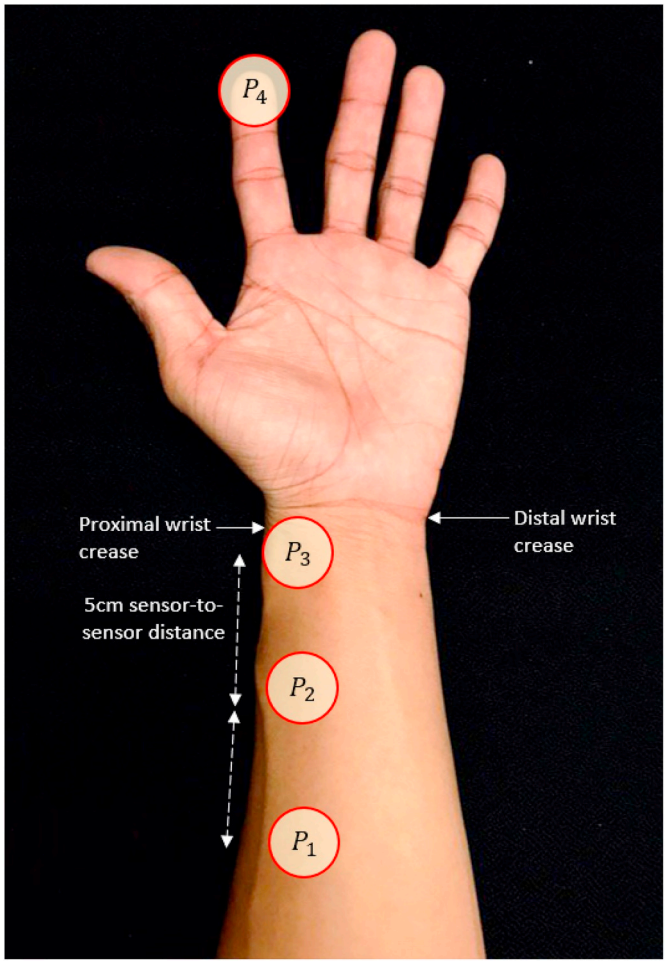

Fig. 8. The above figure represents the positions of the sensors on a volunteer's left arm, positions were selected based on closest proximity to the radial artery P_3 where the sensor assembly could reliably obtain PPG signals and P_1 and P_2 were placed 5 cm as accurately as possible away from P_3. P_4 is a reference finger PPG placed on the index finger typically practiced in today's clinical settings for pulse oximetry.

is always equal and the study fixed the points to 5 cm.

Moreover, we choose pulse peaks to calculate PTT since it was pointed out in Ref. [8] that PTT calculated using pulse peak generally achieves lowest error by means of clear systolic morphology with two point upward signal slope and downward signal slope comparisons in BP prediction.

Whilst the distance between the sensors of different individual's forearm between P_1, P_2 and P_3 could be different, the pulse arrival time should hypothetically arrive with closely between P_1, P_2 and P_3 to offset any difference between the distance selected. Since blood pressure should not change with each heartbeat, the larger the distance between each sensor the greater the PTT differences in each sensor should be expected.

This distance was fixed at 5 cm between each PPG sensor for the purpose of the study, but the brace design is flexible and could be shifted for future study, moreover the size of the brace assembly currently limits the minimum distance that the sensors can be separated.

3.5. SVR ensemble based method for BP prediction

Our ensemble learning method then uses multiple machine learning algorithms to acquire improved predictive performances than could be otherwise obtained from any of the single constituent learning algorithms. In our ensemble learning framework, BP is predicted based on

Table 1
List of extracted features from segment.S_{ij}

No.	Feature	Description
1	X_i	Systolic Amplitude from point P_i Arbitrary amplitude from received PPG light intensity
2	AUC_i	Area under the curve from point P_i
3	AI_i	Augmentation Index from point P_i
4	X_j	Systolic Amplitude from point P_j
5	AUC_j	Area under the curve from point P_j
6	AI_j	Augmentation Index from point P_j
7	aHR (beats/minute)	Average Heart rate
8	aRR (sec)	The average width of the RR interval
9	STD (sec)	The Standard deviation of all RR intervals
10	RMSSD (sec)	Root mean square of differences between adjacent RR intervals
11	LF/HF	Ratio of LF Low-Frequency (0.04–0.15 Hz) power to HF High-Frequency (0.15–0.40 Hz) power
12	PWV_{ij} (mm/second)	Pulse Wave Velocity from segment S_{ij}

Table 2
Systolic blood pressure estimation.

	$ESVR_{MultiPPG}$	OMRON (mmHg)		$ESVR_{MultiPPG}$	OMRON (mmHg)
subject1	113.76	118	subject21	115.73	121
subject2	104.20	88	subject22	111.81	110
subject3	121.09	125	subject23	114.61	109
subject4	134.66	139	subject24	114.02	126
subject5	98.82	99	subject25	110.45	121
subject6	131.45	134	subject26	123.09	130
subject7	107.60	89	subject27	106.66	105
subject8	138.40	140	subject28	105.98	95
subject9	119.68	97	subject29	107.19	112
subject10	122.46	114	subject30	118.65	122
subject11	122.67	117	subject31	111.47	105
subject12	100.17	82	subject32	100.46	104
subject13	130.21	138	subject33	115.13	113
subject14	111.31	104	subject34	113.63	112
subject15	121.01	136	subject35	118.92	117
subject16	128.45	139	subject36	128.16	134
subject17	113.95	127	subject37	111.17	105
subject18	107.27	104	subject38	123.43	132
subject19	112.75	117	subject39	128.25	133
subject20	116.69	126	subject40	112.86	102

Table 3
Diastolic blood pressure estimation.

	$ESVR_{MultiPPG}$	OMRON (mmHg)		$ESVR_{MultiPPG}$	OMRON (mmHg)
subject1	71.20	70	subject21	73.89	77
subject2	69.29	64	subject22	74.06	65
subject3	76.96	80	subject23	73.29	76
subject4	79.73	87	subject24	76.48	82
subject5	65.49	60	subject25	71.18	75
subject6	86.94	86	subject26	73.06	83
subject7	70.72	61	subject27	70.42	70
subject8	82.38	99	subject28	68.07	66
subject9	77.31	65	subject29	69.31	73
subject10	76.22	70	subject30	74.83	75
subject11	77.77	77	subject31	69.06	71
subject12	65.11	64	subject32	68.17	74
subject13	81.96	78	subject33	70.17	67
subject14	70.26	66	subject34	73.80	64
subject15	77.42	81	subject35	74.73	60
subject16	78.80	81	subject36	85.00	80
subject17	73.18	86	subject37	71.69	71
subject18	70.36	63	subject38	77.01	85
subject19	73.61	74	subject39	77.02	80
subject20	71.83	75	subject40	71.06	71

the results of multiple regression models trained on cardiovascular features from multiple segments, with each model making different errors. The ensemble of multiple calibration models, trained on different feature sets, is diverse and makes different errors. Thus, multiple models provide stability and add robustness as compared to a single channel regression model. Features from Table 1 are used to train support vector regression (SVR) [34] model. For each of the six segments, a separate SVR model is trained using its respective feature set. Let $f(x)^{(m)}$ represents output from the mth SVR, where $m = 1, \ldots, 6$ and it is obtained using (1). The final output $f(x)$ is the average of the six SVRs outputs, as expressed below:

$$f(x) = \frac{1}{6}\sum_{m=1}^{6} f(x)^{(m)} \tag{6}$$

The weights are averaged out as each SVR output is a ratio of two points (see Fig. 1), and as such each point P1 – P4 will have three comparison points between the arrival measurement point and the point of difference. The overall distances average out in each SVR.

Table 1 lists the comprehensive feature set used in this paper. Note that pulse amplitude *X*, area under the PPG curve *(AUC)* and augmentation index *(AI)* are used twice in the feature set, one for each PPG observation point in the segment. HRV parameters are obtained using the PPG signals acquired from P4 since for most subjects the cleanest signal is acquired from the fingertip.

Whilst the PPG waveform can be further differentiated into various frequency and sub-domain features to investigate deeper physiological aspects of the PPG waveform with respect to an individual, for purposes of this study the morphological differences are ignored and focus is on pulse transit time (PTT) for blood pressure estimation, the features defined in Table 1 and illustrated in Fig. 9 have been proven to be sufficient for the purposes of this study.

4. Experiments

The experiment section is divided in five sub-sections as follows: Section 4.1 gives the general information about experimental setup and the software packages employed. Section 4.2 discusses the implementation of closely related competing methods. Section 4.3 reports the results obtained by the newly proposed method as well as the competing methods. Section 4.4 performs statistical testing to detect any significant differences among competing methods. Finally, Section 4.5 provides discussion on our findings and results.

4.1. Set up

In order to evaluate the performance of our proposed ensemble method, we acquired data from 40 healthy volunteers. The data collection was approved by the ethical committee of Singapore Polytechnic (IRB Protocol No: 201511-01) and informed consent was obtained from the volunteers. There are 23 male volunteers and 17 female volunteers. Ages of volunteers are in the range of 19–58 ($\bar{x} = 35.2$, $\sigma = 10.436$). The PPG data is acquired for 2 min from all the four channels per subject. Traditional blood pressure measurements from the volunteers were acquired with an Omron™ HEM-7322 automatic cuff-based BP monitor for calibration.

The BP was measured before and after the PPG signals were taken on the **same arm to ensure** pulse-travel time consistency with respect to anatomical distance of the forearm to the heart across all the subjects. The BP cannot be taken during the PPG measurement as the Omron BP cuff occludes the blood vessels and the PPG signals are lost during occlusion. The subjects are all at rest when the BP and PPG signals were taken, so no change was observed to the BP before and after the PPG signals were taken.

Process was as such:

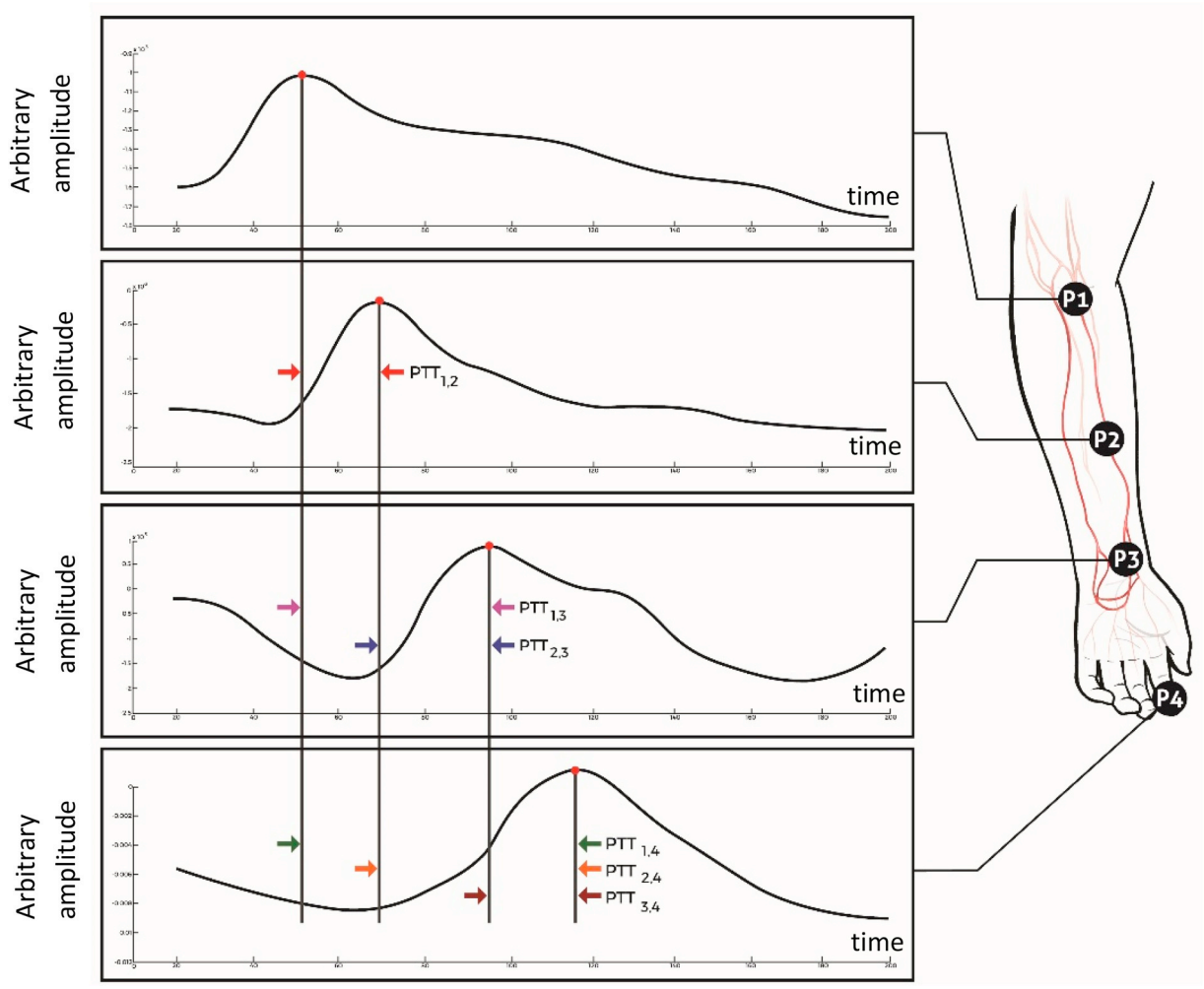

Fig. 9. Six Pulse Transit Time (PTT) for the six arterial segments.

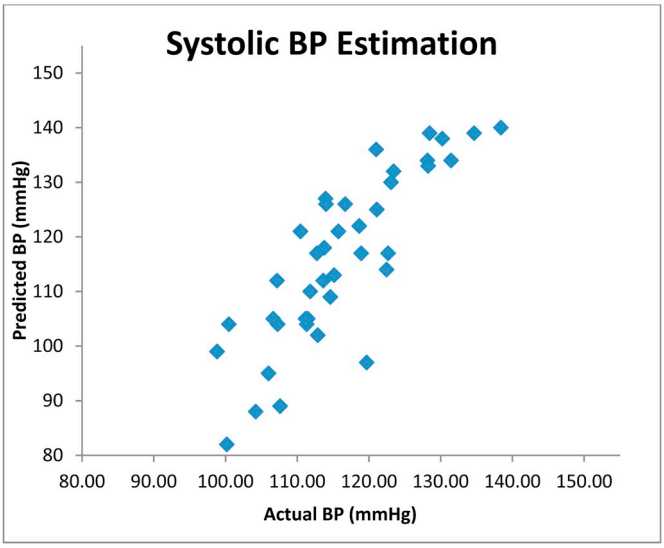

Fig. 10. Actual vs Predicted BP (Systolic).

1. Take BP measurement from Omron HEM-7322 monitor
2. Acquire multi-channel PPG
3. Take BP measurement from Omron HEM-7322 monitor
4. BP reading 1 should be equal to BP reading 2

Table 4
Mean errors in BP predictions.

METHOD	Mean Error (standard deviation)	
	Systolic	Diastolic
$ESVR_{MultiPPG}$	**7.29(5.3)**	**5.01(4.1)**
SVR_{PWV}	9.40(6.6)	5.36(4.2)
SVR_{PWV+PA}	9.38(6.7)	5.30(4.5)
SVR_{PWV+HR}	9.36(6.7)	5.29(4.4)

Traditional BP cannot be taken during PPG measurement as occlusion will suppress PPG signals. BP was taken immediately before and after the PPG data acquisition time.

Signal processing and computation were then performed with MATLAB 2016 on a desktop computer powered by an Intel 2.4 GHz Xeon processor with 6 GB of RAM memory. The SVR based methods are implemented using libsvm library [40].

4.2. Competing methods

Various PPG based BP monitoring methods have been proposed in the past. Since our method is closely related to the method in Ref. [13], we implement this method carefully by ignoring the non-valid PWV associated with measurement artefacts due to the physiological and anatomical limitations. Hence, PWV is constrained within 4 m/s to 20 m/s as in Ref. [41]. Within 2 min of recording, PWV is calculated

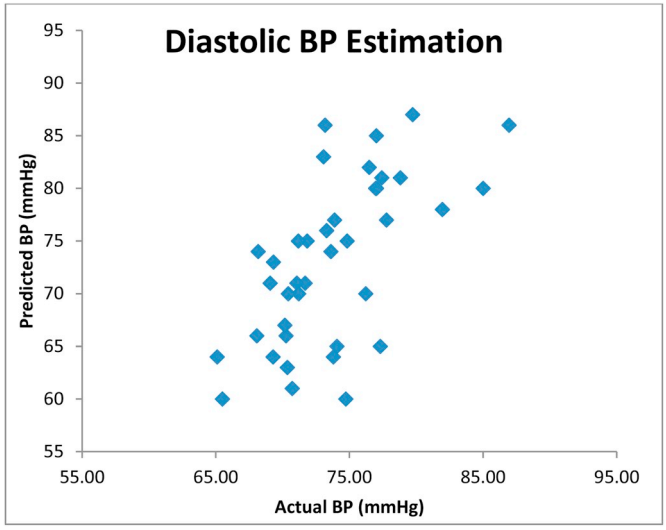

Fig. 11. Actual vs Predicted BP (Diastolic).

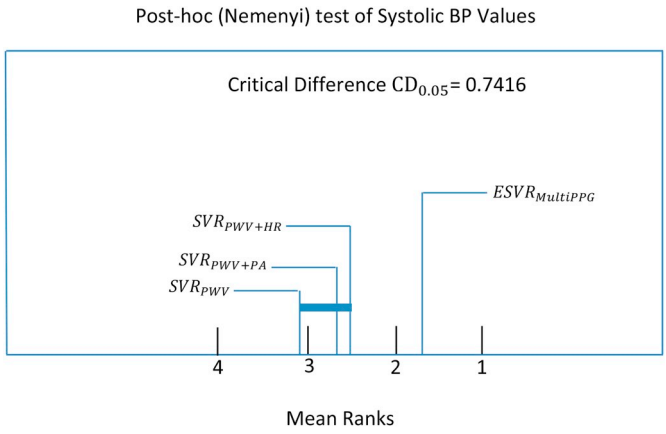

Fig. 12. Visualization of Nemenyi post-hoc test. Horizontal axis represents the mean ranks for $ESVR_{MultiPPG}$, SVR_{PWV}, SVR_{PWV+PA}, and SVR_{PWV+HR}. The SVR_{PWV}, SVR_{PWV+PA}, and SVR_{PWV+HR} are statistically comparable at $p = 0.05$ so they are clustered together.

from a region where correlation values are similar to those reported in Ref. [13]. Note that the dependence of PTT or PWV has already been established extensively [39] and the objective of this research study aims to illustrate effectiveness of ensemble learning frameworks for calibration using multiple PPG segments over single calibration models. Therefore, all the competing methods use identical PWV values and the performance gain or loss among different methods is due to other features or calibration modes. For consistency, we use linear SVR for calibration just as in our proposed method. Even though it was suggested in Refs. [13,14] that a combination of PTT with HRV or PA may produce better result, only PTT based model was used which is referred to asSVR_{PWV} in this paper. Therefore, we use PWV with heart rate (HR) and pulse amplitude (PA) also and termed theseSVR_{PWV+HR}, and SVR_{PWV+PA} respectively.

4.3. Results

Our proposed method is mentioned as $ESVR_{MultiPPG}$ in this section for convenience and we adopted Leave-one-out cross-validation (LOOCV) procedures to train and test the calibration models. For example, the first 39 subjects' data were initially used to train the model and the last subject's data was then applied as the test data. This technique was then repeated 40 times, each time with a different subject's test data. Tables 2 and 3 reports the predicted systolic and diastolic BP values correspondingly by $ESVR_{MultiPPG}$ and also compares to the values recorded by Omron™ HEM-7322 cuff-based BP monitor. The predicted results are also shown as scattered plots represented in Figs. 10 and 11. Table 4 illustrates the mean error and standard deviation (SD) over all the 40 predictions by the four methods in comparison. The best C value for $ESVR_{MultiPPG}$ is found to be 100, while for SVR_{PWV}, SVR_{PWV+PA}, and SVR_{PWV+HR} have value of 10 each.

4.4. Statistical significance testing

In this sub-section, we employed Friedman test [42] to confirm whether the superiority of $ESVR_{MultiPPG}$ is statistically significant when compared to the other methods. We also apply post-hoc analysis to determine if a significant difference is detected by Friedman test, since more than two methods are compared.

For each subject, the prediction method with the lowest absolute error was ranked with increasing order, where the subject with the lowest absolute error ranked as 1, the second lowest subject error ranked as 2, and so on and so forth. If the methods produce the same results, then an average rank is assigned to each of the data sets. For each method, the mean rank over all the 40 subjects is then calculated and obtained with the Matlab function 'Friedman' for the hypothesis testing.

A null hypothesis means that all the methods are comparable, and any differences found among their average ranks are random errors and is thus rejected if the subsequent p-value obtained is found to be lower than a specific confidence level α. For systolic condition, the obtained p-value for all mean errors is 5.2074e-05 therefore a null hypothesis is rejected at $\alpha = 0.05$. For us to analyze the differences among $ESVR_{MultiPPG}$, SVR_{PWV}, SVR_{PWV+PA}, and SVR_{PWV+HR}, a post-hoc Nemenyi test [43] is implemented for pair-wise performance tests. The average ranks of $ESVR_{MultiPPG}$, SVR_{PWV}, SVR_{PWV+PA}, and SVR_{PWV+HR} are 1.725, 3.1125, 2.65 and 2.5125 respectively. If the ranked differences between a pair of prediction methods larger than the critical difference at a certain confidence level, the two methods are considered statistically different. A critical difference of 0.7416 at $p = 0.05$ was reported [42]. The ranked differences from our $ESVR_{MultiPPG}$ and all the other methods were found to be greater than the critical difference thus implies that the $ESVR_{MultiPPG}$ is statistically different while the rank differences among all the remaining 3 methods are smaller than critical difference hence they are statistically alike. Fig. 12 visualizes the Nemenyi post-hoc test at $p = 0.05$, where the mean ranks over the 40 subjects are shown in descending rank order on the x-axis. In this case, statistically comparable methods, SVR_{PWV}, SVR_{PWV+PA}, and SVR_{PWV+HR}, are then grouped with horizontal bar [42].

For the diastolic condition, the p-value obtained for all mean error values is 0.3273 thus the null hypothesis is not rejected at $\alpha = 0.05$.

4.5. Discussion on results

As can be seen from Table 4, $ESVR_{MultiPPG}$ outperformed all the other three methods in both systolic and diastolic BP estimation. SVR_{PWV+PA} and SVR_{PWV+HR} which use the combination of PWV-PA and PWV-HR respectively obtained slightly better performance than PWV alone, with SVR_{PWV+HR} generally a better choice. Moreover, the performance gain obtained by $ESVR_{MultiPPG}$ is contributed to the ensemble framework of weak SVR models. $ESVR_{MultiPPG}$ produces better result than the strongest SVR model in the ensemble as well as SVR_{PWV}, SVR_{PWV+PA}, and SVR_{PWV+HR}. The statistical significant test suggests that $ESVR_{MultiPPG}$ is statistically superior to SVR_{PWV}, SVR_{PWV+PA}, and SVR_{PWV+HR} for systolic BP prediction. The other three methods are statistically similar. For diastolic BP, no statistical difference is observed among the methods on the 40 subjects participated.

This work is a proof of concept to show the advantage of using ensemble learning based regression model over single regression model

for BP calibration. It is observed that the ensemble framework employing our in-house built multi-channel PPG device achieves statistically superior performance than single model-based BP calibration.

This ensemble framework will later be translated into a wearable device. Even though the mean error in systolic and diastolic BP estimation is currently higher than the AMI standard, which is less than 5 mmHg, we foresee better accuracy with the availability of more volunteers. Since it is a non-personalized empirical calibration framework, the accuracy will only increase with more data.

The results are encouraging and supports the hypothesis for the feasibility of non-occluding continuous optical-only BP measurements. For future work, there is motivation to improve the specificity and sensitivity of this technique to meet clinical standards for blood pressure measurements. A larger dataset from a greater number of subjects applied to the newly proposed ensemble approach and additional metrics such as personalized calibration to each individual under measure would certainly improve the specificity of the predictive model developed.

5. Conclusions and future work

In conclusion, we have successfully implemented, developed and proposed a new ensemble learning framework to predict blood pressure using a customized multi-channel photoplethysmogram (PPG) device. Pulse waves received at multiple observation points, corresponding to the same cardiac cycle, help create different feature sets. Multiple support vector machines for regression (SVR) models are then developed in the ensemble, with each model trained on a distinct feature set. The combined output of the ensemble is more stable and also reduces the risk of overfitting associated with a single calibration model. The feature set is comprehensive in that it includes various surrogate markers for BP including pulse amplitude, pulse area under the curve, augmentation index, HRV's time and frequency domain parameters and pulse transit time. The proposed framework is completely optical, non-occluding and measures BP continuously.

This is a preliminary study to show the effectiveness of using ensemble learning based calibration model trained on features from multiple arterial segments over single calibration model trained on features from a single segment. We are currently working towards translating the in-house built multi-channel PPG device, along with the proposed ensemble calibration framework, into a wearable device. We are extremely appreciative and thankful to the 40 volunteers who participated in this research study and with the availability of more volunteers, the proposed framework is expected to achieve higher prediction accuracy since it constitutes non-personalized empirical calibration models. Moreover, the calibration was done using a consumer OMRON device.

The objective of this research was to determine if optical-only measurements (by avoiding ECG reference) could be used to predict BP as a proof-of-concept, the advanced machine learning could be future work to improve BP prediction outcomes and the experiment involving healthy individuals only was to verify that the hypothesis is valid.

Now that the physical concept has been verified on healthy individuals, the authors will further improve on the algorithms in future work by using more advanced machine-learning techniques such as Long Short Term Memory Networks (LSTM-NN) neural networks that are increasingly being considered in studies for blood pressure prediction and processing on mobile devices which is a resource-constraint approach on current mobile devices. But as computational power rapidly improves on future generations of devices, it is expected that such a future solution will be of great benefit in the detection and management of general blood pressure conditions when sensors are developed into long-term wearables instead of discrete bench-top meters, where the future goal of long-term wearable measurements will aid in early detection of cardiovascular anomalies to prevent life-threatening occurrences.

With this encouraging result, the authors are planning future work involving a greater volunteer sample size and collection of data from diseased groups of patients (hypertension etc.) with our healthcare collaborators (SingHealth) and the implementation of more sophisticated machine-learning techniques such as LSTNM-NN to reduce the mean error and to improve predictive outcome of continuous, non-occluding BP measurements.

Conflicts of interest

None declared.

Authors' contributions

Mark Wong, Tan Jen Hong planned and implemented the experiments. Mark Wong and Kenneth Er developed both hardware and firmware for the electronics. Mark, Jen Hong and Eddie Ng performed signal processing, data collection and analysis and also drafted and edited the manuscript. Mark Wong managed the scheduling of human participants and data collection protocol. All authors approved the final manuscript.

References

[1] I. Hajjar, J.M. Kotchen, T.A. Kotchen, Hypertension: trends in prevalence, incidence, and control, Annu. Rev. Public Health 27 (2006) 465–490.
[2] P.A. Iaizzo, Handbook of cardiac anatomy, physiology, and devices, J. Card. Surg. 21 (3) (2006) 327-327.
[3] H. Smulyan, M.E. Safar, Blood pressure measurement: retrospective and prospective views, Am. J. Hypertens. 24 (6) (2011) 628–634.
[4] G. Ogedegbe, T. Pickering, Principles and techniques of blood pressure measurement, Cardiol. Clin. 28 (4) (2010) 571–586.
[5] N.R. Campbell, A. Chockalingam, J.G. Fodor, D.W. McKay, Accurate, reproducible measurement of blood pressure, Can. Med. Assoc. J. 143 (1) (1990) 19.
[6] H. Gesche, D. Grosskurth, G. Küchler, A. Patzak, Continuous blood pressure measurement by using the pulse transit time: comparison to a cuff-based method, Eur. J. Appl. Physiol. 112 (1) (2012) 309–315.
[7] S. Terry, J.S. Eckerle, R.D. Kornbluh, T. Low, C.M. Ablow, Silicon pressure transducer arrays for blood-pressure measurement, Sens. Actuators A Phys. 23 (1–3) (1990) 1070–1079.
[8] Y. Choi, Q. Zhang, S. Ko, Non-invasive cuffless blood pressure estimation using pulse transit time and Hilbert–Huang transform, Comput. Electr. Eng. 39 (1) (2013) 103–111.
[9] M.A. Hassan, M.Y. Mashor, A.M. Saad, S. Mohamed, Non-invasive continuous blood pressure monitoring based on PWTT, J. Adv. Comput. Sci. Technol. Res. 1 (2) (2011) 63–73.
[10] A. Arza, J. Lázaro, E. Gil, P. Laguna, J. Aguiló, R. Bailon, Pulse transit time and pulse width as potential measure for estimating beat-to-beat systolic and diastolic blood pressure, IEEE Comput. Cardiol. Conf. (2013) 887–890.
[11] X.F. Teng, Y.T. Zhang, Continuous and non-invasive estimation of arterial blood pressure using photoplethysmographic approach, IEEE Engineering in Medicine and Biology Society Conference, 4 2003, pp. 3153–3156.
[12] P. Fung, G. Dumont, C. Ries, C. Mott, M. Ansermino, Continuous non-invasive blood pressure measurement by pulse transit time, IEEE Engineering in Medicine and Biology Society Conference, 1 2004, pp. 738–741.
[13] C. Douniama, C.U. Sauter, R. Couronne, Blood pressure tracking capabilities of pulse transit times in different arterial segments: a clinical evaluation, IEEE Computers in Cardiology Conference, 2009, pp. 201–204.
[14] C. Douniama, C.U. Sauter, R. Couronne, Acquisition of Parameters for Non-invasive Continuous Blood Pressure Estimation–Review of the Literature and Clinical Trial, World Congress on Medical Physics and Biomedical Engineering, 2009, pp. 2151–2154.
[15] D.B. McCombie, A.T. Reisner, H.H. Asada, Motion based adaptive calibration of pulse transit time measurements to arterial blood pressure for an autonomous, wearable blood pressure monitor, IEEE Engineering in Medicine and Biology Society Conference, 2008, pp. 989–992.
[16] E.B. Schroeder, D. Liao, L.E. Chambless, R.J. Prineas, G.W. Evans, G. Heiss, Hypertension, blood pressure, and heart rate variability the Atherosclerosis Risk in Communities (ARIC) study, Hypertension 42 (6) (2003) 1106–1111.
[17] G.L. Xie, J.H. WANG, Y. Zhou, H. Xu, J.H. Sun, S.R. Yang, Association of high blood pressure with heart rate variability in children, Iran. J. Paediatr. 23 (1) (2013) 37.
[18] N. Natarajan, A.K. Balakrishnan, K. Ukkirapandian, A study on analysis of Heart Rate Variability in hypertensive individuals, Int. J. Biomed. Adv. Res. 5 (2) (2014) 109–111.
[19] R. Virtanen, A. Jula, T. Kuusela, H. Helenius, L.M. Voipio-Pulkki, Reduced heart rate variability in hypertension: associations with lifestyle factors and plasma renin activity, J. Hum. Hypertens. 17 (3) (2003) 171–179.
[20] Task Force of the European Society of Cardiology the North American Society of

Pacing Electrophysiology, Heart rate variability: standards of measurement, physiological interpretation, and clinical use, Circulation 93 (1996) 1043–1065.
[21] N. Liu, Z.X. Koh, E.C. Chua, L.M. Tan, Z. Lin, B. Mirza, M.E. Ong, Risk scoring for prediction of acute cardiac complications from imbalanced clinical data, IEEE J. Biomed. Health Inf. 18 (6) (2014) 1894–1902.
[22] A. Schäfer, 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. 166 (1) (2013) 15–29.
[23] X.F. Teng, Y.T. Zhang, Study on the peak interval variability of photoplethysmogtaphic signals, IEEE EMBS Asian-Pacific Conference on Biomedical Engineering, 2003, pp. 140–141.
[24] P.S. McKinley, P.A. Shapiro, E. Bagiella, M.M. Myers, R.E. De Meersman, I. Grant, R.P. Sloan, Deriving heart period variability from blood pressure waveforms, J. Appl. Physiol. 95 (4) (2003) 1431–1438.
[25] E.C. Chua, S.J. Redmond, G. McDarby, C. Heneghan, Towards using photo-plethysmogram amplitude to measure blood pressure during sleep, Ann. Biomed. Eng. 38 (3) (2010) 945–954.
[26] C.P. Chua, C. Heneghan, Continuous blood pressure monitoring using ECG and finger photoplethysmogram, IEEE Engineering in Medicine and Biology Society Conference, 2006, pp. 5117–5120.
[27] A. Visvanathan, A. Sinha, A. Pal, Estimation of blood pressure levels from reflective Photoplethysmograph using smart phones, IEEE Bioinformatics and Bioengineering (BIBE) Conference, 2013, pp. 1–5.
[28] M. Elgendi, On the analysis of fingertip photoplethysmogram signal, Curr. Cardiol. Rev. 8 (1) (2012) 14–25.
[29] R. Wang, W. Jia, Z.H. Mao, R.J. Sclabassi, M. Sun, Cuff-free blood pressure estimation using pulse transit time and heart rate, IEEE International Conference on Signal Processing (ICSP), 2014, pp. 115–118.
[30] E. Monte-Moreno, Non-invasive estimate of blood glucose and blood pressure from a photoplethysmograph by means of machine learning techniques, Artif. Intell. Med. 53 (2) (2011) 127–138.
[31] Y. Ren, L. Zhang, P.N. Suganthan, Ensemble classification and regression-recent developments, applications and future directions [review article], IEEE Comput. Intell. Mag. 11 (1) (2016) 41–53.
[32] B. Mirza, Z. Lin, N. Liu, Ensemble of subset online sequential extreme learning machine for class imbalance and concept drift, Neurocomputing 149 (2015) 316–329.
[33] H. Chen, C. Tan, Z. Lin, T. Wu, The diagnostics of diabetes mellitus based on ensemble modelling and hair/urine element level analysis, Comput. Biol. Med. 50 (2014) 70–75.
[34] A.J. Smola, B. Schölkopf, A tutorial on support vector regression, Stat. Comput. 14 (3) (2004) 199–222.
[35] K. Takazawa, N. Tanaka, M. Fujita, O. Matsuoka, T. Saiki, M. Aikawa, S. Tamura, C. Ibukiyama, Assessment of vasoactive agents and vascular aging by the second derivative of photoplethysmogram waveform, Hypertension 32 (2) (1998) 365–370.
[36] S. Lu, H. Zaho, K. Ju, K. Shin, M. Lee, K. Shelly, Can photoplethysmography variability serve as an alternative approach to obtain heart rate variability information? J. Clin. Monit. Comput. 22 (1) (2008) 23–29.
[37] T. Liu, Z. Lin, M.E.H. Ong, Z.X. Koh, P.P. Pek, Y.K. Yeo, B.S. Oh, A.F.W. Ho, N. Liu, Manifold ranking based scoring system with its application to cardiac arrest prediction: a retrospective study in emergency department patients, Comput. Biol. Med. 67 (2015) 74–82.
[38] M.P. Tarvainen, J.P. Niskanen, J.A. Lipponen, P.O. Ranta-Aho, P.A. Karjalainen, Kubios HRV–heart rate variability analysis software, Comput. Methods Progr. Biomed. 113 (1) (2014) 210–220.
[39] R. Mukkamala, J.O. Hahn, O.T. Inan, L.K. Mestha, C.S. Kim, H. Töreyin, S. Kyal, Toward ubiquitous blood pressure monitoring via pulse transit time: theory and practice, IEEE (Inst. Electr. Electron. Eng.) Trans. Biomed. Eng. 62 (8) (2015) 1879–1901.
[40] C.C. Chang, C.J. Lin, LIBSVM: a library for support vector machines, ACM Trans. Intell. Syst. Technol. 2 (3) (2011) 27 [Online]. Available http://www.csie.ntu.edu.tw/~cjlin/libsvm.
[41] S. Hey, A. Gharbi, B. von Haaren, K. Walter, N. König, S. Löffler, Continuous non-invasive pulse transit time measurement for psycho-physiological stress monitoring, IEEE International Conference on eHealth, Telemedicine, and Social Medicine (eTELEMED'09), 2009, pp. 113–116.
[42] J. Demšar, Statistical comparisons of classifiers over multiple data sets, J. Mach. Learn. Res. 7 (2006) 1–30.
[43] P. Nemenyi, Dissertation Abstracts International, Distribution-free Multiple Comparisons (Doctoral Dissertation, Princeton University, 1963), (1963), p. 1233 2 vol. 25.