Personalized Blood Pressure Estimation using Photoplethysmography and Wavelet Decomposition

Jared Leitner Po-Han Chiang Sujit Dey
Mobile Systems Design Lab, Dept. of Electrical and Computer Engineering, University of California, San Diego
Email: {jjleitne, pochiang, sdey}@ucsd.edu

***Abstract* — Blood pressure (BP) is the most important indicator of cardiovascular diseases. Traditional cuff-based methods for measuring BP require manual intervention and time. These methods may lead to inaccurate measurements and are not practical for continuous BP monitoring, which is crucial for detecting abnormal BP fluctuations. In this study, we propose a personalized machine learning model to estimate BP using his/her previous BPs and the photoplethysmogram (PPG) signal, the simplest and most popular tool for non-invasive diagnosis. To best utilize the information contained in the PPG signal, we propose to apply wavelet decomposition to extract features from the PPG signal. The arterial blood pressure (ABP) time series is processed with an exponentially weighted moving average (EWMA) and a peak detection technique to derive the SBP, DBP, and their corresponding trends. Finally, a random forest model is used to construct a predictive model based on these features. The MIMIC dataset is used for analysis and comparison with other BP estimation methods. Our experimental results demonstrate the proposed approach has smaller estimation error than existing methods, with mean average errors (MAE) for SBP and DBP equal to 3.43 and 1.73, respectively.**

***Index Terms*— Machine Learning; Blood Pressure; photoplethysmogram; Wearables; Wavelets**

I. Introduction

High blood pressure (BP), or hypertension, affects 30% of American adults and contributes to over 410,000 deaths per year [1,2]. This condition has been called "the silent killer," as typically no symptoms are recognized before significant damage has already been done to the heart and arteries [3]. BP is defined as the pressure exerted on the arteries as blood is pumped throughout the body and is measured in millimeters of mercury (mmHg). The primary metrics used to measure BP are systolic (SBP) and diastolic blood pressure (DBP), which are defined as the maximum and minimum blood pressure, respectively, during a pulse.

For accurate diagnosis and treatment of hypertension, regular BP measurement is necessary. According to the American College of Cardiology, increased at home BP monitoring is essential for recognizing inconsistencies in measurements taken in a medical setting [4]. Currently, the predominant device for measuring BP is a mercury sphygmomanometer which involves attaching an inflatable cuff around the upper arm [5]. This process requires significant user effort, which limits the frequency of BP measurements and increases the chance of measurement error. The use of an arterial catheter can provide continuous BP measurement; however, it is highly invasive and impractical for daily life. On the other hand, continuous blood pressure estimation could be incorporated into one's daily routine using data collected non-invasively. One prominent method is utilizing the photoplethysmogram (PPG) sensor, which is available in most wrist wearables. The principle of the PPG sensor is to measure the dilation and constriction of blood vessels. In this study, we propose the use of machine learning techniques to construct accurate and personalized models for BP estimation based on the PPG signal and historical BP values.

A. Related Work

Previous work focused on continuous BP estimation is described in [6-10]. The use of both electrocardiogram (ECG) and PPG signals for BP estimation is studied using neural networks [6] and regression trees [7]. Although the ECG signal provides useful insight into the operation of the circulatory system, the sensor is expensive and currently not available in most wearables. [8] uses both PPG and physiological features (e.g., height, weight, and age) for BP estimation. The inclusion of physiological features improves their model performance; however, their non-personalized approach does not meet the required performance of cuff-based methods. This higher performance when including individual physiological features emphasizes the fact that personal characteristics play an important role in BP estimation. [9,10] propose methods for BP estimation using only the PPG signal. However, the performance is limited since their tree-based models are not personalized.

All of the above studies cannot inherently extract time series information, and more descriptive time series features are required. In addition, none of the above papers consider historical BP information as an additional predictor variable. Since our goal is to produce a practical, cost-effective solution for continuous BP measurement, we only consider the PPG signal as an input to our system. [9,10] extract statistical features from the PPG series on a per cycle basis, which does not fully capture the time series dependence between cycles. In order to best extract features from the PPG time series, our approach utilizes wavelet decomposition [11], which is known for its ability to extract the true underlying variation of a signal. As indicated in [8], individual physiology ranges from person to person. Taking this into consideration, we train personalized models for each patient in our study. Wavelet decomposition creates a large number of features, some of which may be an irrelevant and redundant representation of the PPG signal. To address this

issue, we select the random forest, one of the most popular ensemble learning methods, as our machine learning model due to its ability to perform well with redundant features. Furthermore, if historical BP information using an exponentially weighted moving average (EWMA) is provided as an additional feature, the performance exceeds all previous work for BP estimation based on photoplethysmography.

The rest of the paper is organized as follows. In section II, data retrieval, feature extraction, and the proposed random forest approach to personalized BP estimation are presented. In section III, the performance of our model is evaluated. Finally, we conclude the paper in section IV.

II. Proposed Method

A. *Dataset and BP Estimation Framework*

Data was obtained from the Multiparameter Intelligent Monitoring in Intensive Care (MIMIC) database [12]. This database contains records of over 90 intensive care unit patients, with an average of 40 hours of data per patient. The waveforms collected include ECG, respiration, continuous blood pressure, and PPG signals each sampled at 125 Hz. The arterial blood pressure (ABP) was directly measured from a radial artery using an invasive catheter. A fingertip sensor was used to measure the PPG data. Not all patients have data for all signals, and only patients with sufficient PPG and blood pressure data were considered for this study. As a result, we filtered out 40 patients with sufficient PPG and ABP data.

The block diagram of our proposed framework is shown in Figure 1. The first stage of the preprocessing phase is time series decomposition using wavelet transformation (WT) [11], a signal processing technique used to analyze the PPG signal in both the temporal and spectral domain. In our study, we use the stationary wavelet transform (SWT) to derive additional features with length equal to that of the original PPG signal. Both the mathematical formulation and application of the wavelet transform is described in the following section.

The next step is to extract the SBP and DBP (maximum and minimum ABP) from the raw ABP signal in each pulse. We adopt a peak detection technique [13], which uses a moving average of the ABP as the threshold to determine the indices of the peaks in the ABP series. Once these indices are known, the original dataset for each patient is downsampled to only include the BP and feature values for each peak index. This results in a dataset with SBP values and corresponding PPG and wavelet values. DBP indices are determined by inverting the ABP series to detect the local minimum in each pulse. A separate dataset is constructed for the DBP values with corresponding PPG and wavelet values. Both the SBP and DBP values along with the raw ABP waveform can be seen in Figure 2.

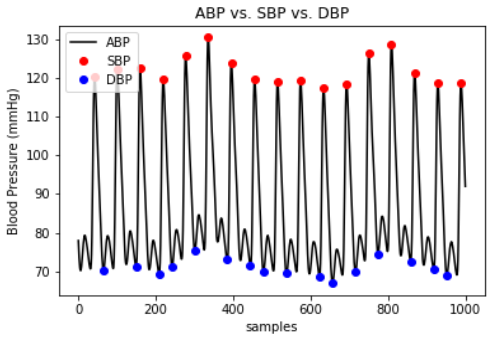

Figure 2. SBP and DBP values vs. raw ABP series

Once the SBP and DBP datasets have been constructed, an additional EWMA feature is included for both SBP and DBP. The rationale is that the EWMA provides a low-variance representation of the SBP and DBP waveforms. In our study, the EWMA series is offset by 10 minutes. Finally, we use a random forest to estimate current SBP and DBP with our derived dataset containing the EWMA and wavelet-decomposed PPG data. Our proposed model uses a combination of SBP and DBP features for every cycle to estimate SBP/DBP values. Since SBP and DBP are strongly correlated, merging the SBP and DBP datasets into a single dataset provides a stronger representation of PPG and historical BP.

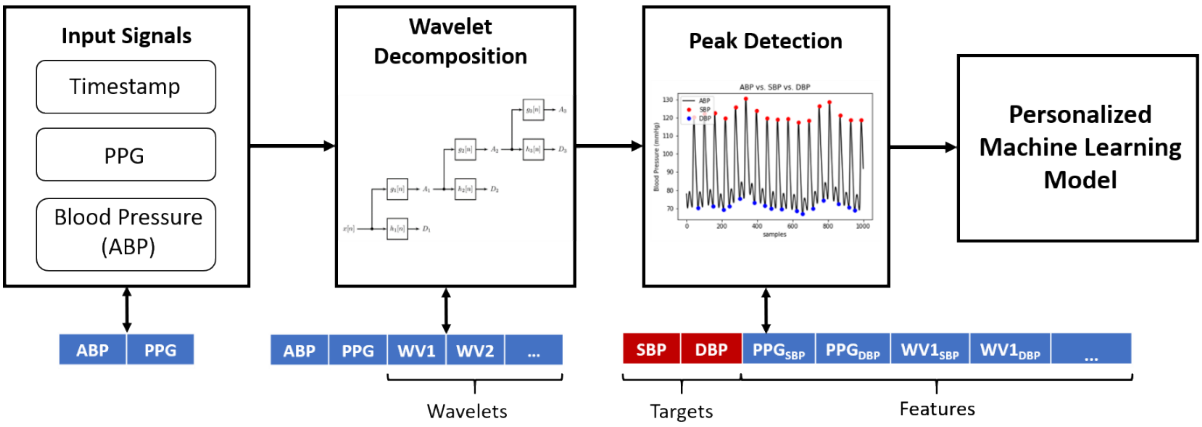

Figure 1. Block Diagram of proposed framework

B. *Wavelet Decomposition*

Wavelet transform is used in order to understand the underlying variation of the PPG signal at multiple levels of resolution. Although the Fourier transform is widely used to extract global frequencies present in a signal, it is not able to provide any resolution in time. That is, it cannot determine at which times specific frequency components are present. Wavelet analysis trades off resolution in frequency for resolution in time by instead projecting the signal onto wavelet functions, as compared to sinusoids in the Fourier case. The mother wavelet function is defined as a short duration wave, and other wavelets are obtained by shifting and changing the frequency of the mother wavelet. The following equation describes a wavelet function parameterized by translation and dilation parameters b and a, where $\Psi(t)$ is the mother wavelet function.

$$\Psi_{a,b}(t) = \frac{1}{\sqrt{a}}\Psi\left(\frac{t-b}{a}\right)$$

The PPG signal is decomposed using these wavelets at different dilations and translations in order to get a high-resolution frequency representation of the signal. The above equation represents the continuous wavelet transform. Figure 3 displays a filter representation of the SWT, which consists of a cascade of low and high pass filters [11]. A_i and D_i correspond to the approximate and detailed coefficients at level i. This cascade of filters allows for a multi-tiered frequency representation of the PPG signal to be extracted. Figure 4 displays a segment of a PPG signal and the corresponding detailed wavelet components at multiple levels.

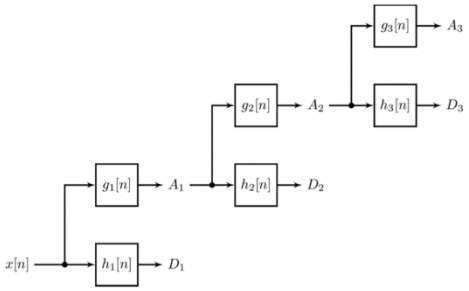

Figure 3. Filter implementation of SWT

C. *Random Forest*

After the preprocessing phase, the SBP and DBP datasets are used to train our machine learning model. Due to the redundant representation of the PPG signal acquired from the stationary wavelet transform, the top performing machine learning algorithm must be robust to potentially irrelevant features. The random forest (RF) model is known to perform well even when using redundant features [14]. In addition, the random forest model only has two primary parameters, making the parameter tuning process far more tractable than a deep learning approach. A model that is not hyper-sensitive

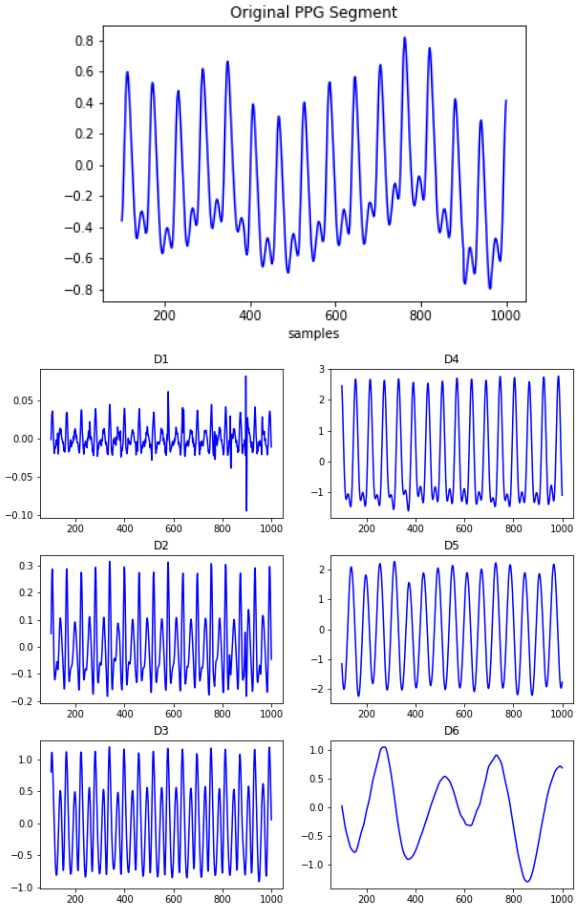

Figure 4. Detailed wavelet coefficients of PPG segment

to its parameters allows it to be more easily applied to new patient data.

The random forest is an ensemble model, which trains multiple decision trees and averages the outputs of each of these trees when making a prediction. Each individual tree learns decision rules from the input PPG data that lead to the greatest reduction in the variance of the target blood pressure values. Every tree consists of 3 types of nodes: 1) root node, where all data is input before any splits, 2) decision nodes, where the data is split according to the feature values, and 3) leaf nodes, where no more splits are performed. Since SBP and DBP are continuous values, these trees are called regression trees. When provided with a new data sample, each decision tree will output the average value of the leaf node in which that sample falls.

Random forests take advantage of the principle that a group of weak learners (decision trees) can form a strong learner. Bootstrap aggregation is the method used for training random forests. Data samples are randomly drawn from the entire dataset with replacement to form a bootstrap dataset that is the same size as the original dataset. Each decision tree is trained using a different bootstrap dataset. This process reduces the variance of the overall model, as the output is averaged over all individual trees. In addition, the bias remains unchanged as each tree is identically distributed and

therefore the expectation of the average of the trees is equal to the expectation of any individual tree. As a result, RF decreases the overall error by reducing the variance while maintaining the same bias.

III. Results and Discussion

In this section, we will describe the experiment settings and compare BP estimation results obtained from our proposed method to previous work. We also discuss feature importance and test our model on future BP prediction.

A. Experiment Setting

We implement and evaluate our machine learning model using the Scikit-learn library [15] and Keras [16] in the python environment on an Intel i5 3.2GHz quad-core and 16GB RAM computer. Root mean square error (RMSE), mean absolute error (MAE), and mean absolute percentage error (MAPE) are calculated and used as our evaluation metrics. Their definitions are as follows:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}\left(P_{pred}^{i} - P_{actual}^{i}\right)^2}{n}}$$

$$MAE = \frac{\sum_{i=1}^{n}\left|P_{pred}^{i} - P_{actual}^{i}\right|}{n}$$

$$MAPE = \frac{nMAE}{\sum_{i=1}^{n}\left|P_{actual}^{i}\right|} \times 100\%$$

We use 5-fold cross-validation to randomly split the SBP and DBP datasets for each individual into the train (80%) and the test (20%) sets 5 times and average the estimation results. Personalized models are built for all 40 patients, each using 10 hours of data. For our RF models, we set the number of trees to 200 and the minimum number of samples required to split an internal node as 2. For comparison, a long short-term memory (LSTM) network is also trained to estimate BP based solely on historical BP. This is carried out in order to determine how our proposed method compares to a deep learning approach that is known to effectively model time series. The LSTM is trained using 64 as the batch size and the Adam optimizer [17]. We also use early stopping and insert dropout layers with dropout rate equal to 0.2 to avoid overfitting. In addition, we compare our results to the RReliefF framework proposed by [9]. This framework selects a subset of the most relevant statistical features they extract from the PPG series to provide to their model. This direct comparison to [9] can be made because the MIMIC dataset was also used in their study. They do not provide results for RMSE or MAPE. We also compare the performance of the variants of our proposed method with different subsets of features. These different feature sets include 1) PPG + Wavelets, 2) PPG + Wavelets + EWMA, and 3) PPG + Wavelets + EWMA + BP-Merge, where the third feature set corresponds to our proposed method described in Section II.

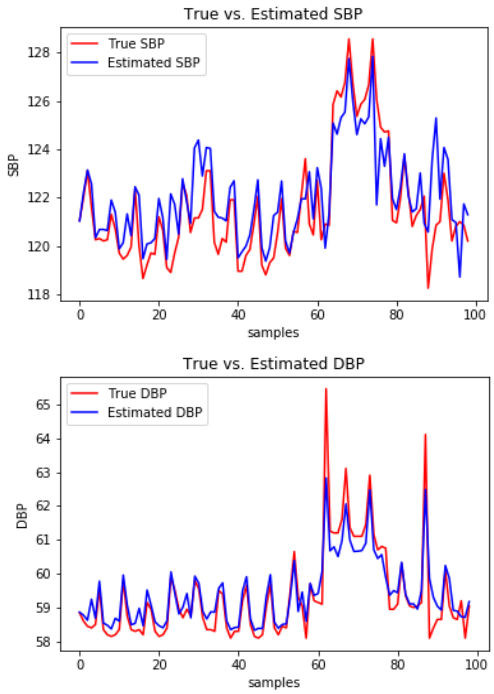

Figure 5. True vs. Estimated SBP/DBP series

B. BP Estimation Results

Table I summarizes the performance for BP estimation using the RF model with different feature sets compared to both the LSTM and RReliefF method [9]. These values are the average over 40 patients. As can be seen, the performance of the RF model when using PPG and wavelet features is significantly better than the LSTM when using historical BP. The MAE for SBP and DBP estimation when using the RF and these features is 4.88 and 2.61, respectively, while that of the LSTM is 10.14 and 4.94. This is a surprising result, as the LSTM is expected to work well with time series data; however, it indicates that the PPG and wavelet features are a robust set of features on their own for BP estimation.

	SBP			DBP		
	MAE	RMSE	MAPE (%)	MAE	RMSE	MAPE (%)
LSTM - Historical BP	10.14	13.13	7.99	4.94	6.86	7.79
RReliefF [9]	4.47	-	-	2.02	-	-
RF - PPG + Wavelets	4.88	7.31	3.87	2.61	4.09	4.20
RF - PPG + Wavelets + EWMA	3.90	6.02	3.12	2.04	3.39	3.30
RF - PPG + Wavelets + EWMA + BP-Merge	3.43	5.36	2.74	1.73	2.97	2.81

TABLE I. Comparison of performance for different methods

The performance of our RF model increases when the EWMA is added in addition to the PPG and wavelet features. The average MAE for SBP and DBP decreases from 4.88 to 3.90 and from 2.61 to 2.04, respectively. Furthermore, our proposed method (PPG + Wavelets + EWMA + BP-Merge) results in the top performance. The MAE for SBP and DBP estimation when using this approach is 3.43 and 1.73, respectively. This satisfies the standard for cuff-based measurement, which is a MAE of 5 for both SBP and DBP [18]. In addition, our performance exceeds that of the RReliefF method, which obtained a MAE of 4.47 and 2.02 for SBP and DBP. We attribute this increase in performance to our use of wavelet decomposition, the inclusion of historical BP as an additional feature, and a combination of SBP and DBP features. Figure 5 displays a segment of the true SBP/DBP series compared to our estimated series for one patient.

The random forest model is also able to obtain a ranking of the feature importance. This provides insight into which features have the largest effect on BP estimation. The following are the 3 most common top feature for estimating SBP when using our combined SBP/DBP features approach: 1.) SBP EWMA, 2.) DBP EWMA, 3.) SBP PPG. This indicates that historical BP information is a significant factor in estimating new BP values. Figure 6 displays the top 5 features and their relative importance for 2 of the 40 patients. Each feature importance score is in between 0 and 1.

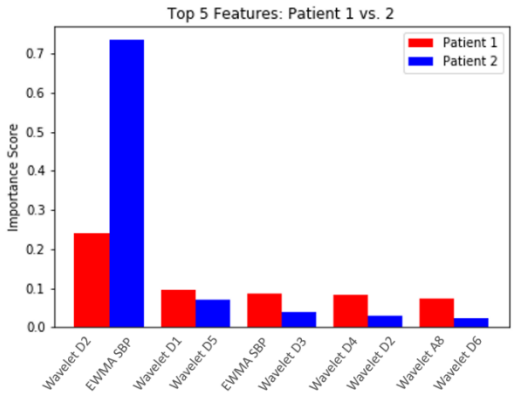

Figure 6. Comparison of top 5 features for two patients

As can be seen, the top features along with their importance score range from patient to patient. This indicates that certain features are more important for specific individuals, which further validates our approach to using personalized models. The top 5 features along with their importance score for patient 1, in order, are: 1) Wavelet D2 (0.241), 2) Wavelet D1 (0.096), 3) EWMA SBP (0.087), 4) Wavelet D4 (0.081), and 5) Wavelet A8 (0.074). Clearly, the detailed wavelet coefficients have a strong effect on BP estimation for this patient. As described earlier, these coefficients correspond to the outputs of the high pass filters in the SWT implementation. Larger wavelet numbers correspond to deeper levels of the filter cascade. For patient 2, the top 5 features and importance scores are: 1) EWMA SBP (0.734), 2) Wavelet D5 (0.071), 3) Wavelet D3 (0.034), 4) Wavelet D2 (0.030), and 5) Wavelet D6 (0.022). The differences in feature importance for patient 1 are far less than those for patient 2, whose top feature clearly dominates with a score of 0.734. The remaining four features for patient 2 also consist of detailed wavelet coefficients, indicating this representation of the PPG series contributes to our high performance.

We also test our model's ability to predict future BP values when using the combined feature set. Five different look-ahead windows are considered and include 5 minutes, 10 minutes, 30 minutes, 1 hour, and 2 hours. As can be seen from Table II, the performance decreases when predicting further into the future but levels off after a certain point. The MAE for predicting SBP 5 minutes into the future is 5.01 as compared to 3.43 for estimating current values. For DBP, these MAEs are 2.37 and 1.73, respectively. This significant increase in error indicates that current PPG information does not provide a strong representation of future BP values, and loses its predictive power relatively quickly. As the look-ahead window increases, the MAE for both SBP and DBP prediction saturates. The difference in performance for SBP prediction 30 minutes and 2 hours into the future is 0.03. For DBP, this difference is 0.002. This demonstrates that after a certain look-ahead window, our model reaches a performance limit. Overall, these results indicate that predicting future BP values is more difficult than estimating current values, and additional derived features, possibly from health and sleep data, are required.

	SBP MAE	DBP MAE
5 min	5.001	2.396
10 min	5.221	2.474
30 min	5.596	2.622
1 hr	5.502	2.547
2 hr	5.626	2.624

TABLE II. Performance for different look-ahead windows

IV. Conclusion

In this paper, we provide a novel machine learning approach to continuous BP estimation. Our preprocessing framework successfully extracts information from the raw PPG and ABP signals. Our use of wavelet decomposition to extract features from the PPG series, in conjunction with using historical BP information in the form of an exponentially weighted moving average, provides a robust set of features for BP estimation. By including the extra historical BP feature, the performance greatly improves, indicating that past BP information should be considered when estimating new BP values. Furthermore, when combing the SBP and DBP feature sets, our top performing model exceeds the cuff-based standard and all other studies conducted on this topic. Provided PPG and BP data from a new patient, our proposed method can be used to construct a personalized BP estimation model that can carry out

continuous BP measurement. Future work involves developing a method for transferring existing models to new patients that do not have enough personal data to form their own models.

REFERENCES

[1] C. Fryar, Y. Ostchega , C. Hales ,G. Zhang and D. Kruszon-Moran, "Hypertension prevalence and control among adults: United States, 2015–2016," *NCHS data brief*, no.289, 2017.

[2] Centers for Disease Control and Prevention, and National Center for Health Statistics. "Underlying cause of death 1999-2013 on CDC WONDER online database, released 2015." *Data are from the multiple cause of death files*, 2013.

[3] The World Health Organization, "A Global Brief on Hypertension," April 2013.

[4] American College of Cardiology, "Guideline for the Prevention, Detection, Evaluation, and Management of High Blood Pressure in Adults," *Journal of the American College of Cardiology*, pp.4-28, 2017.

[5] G. Odedegbe and T. Pickering, "Principles and techniques of blood pressure measurement," Cardiol. Clinics, vol. 28, no. 4, pp. 571–586, 2010.

[6] P. Su, X. Ding, Y. Zhang, J. Liu, F. Miao and N. Zhao, "Long-term blood pressure prediction with deep recurrent neural networks," *2018 IEEE EMBS International Conference on Biomedical & Health Informatics (BHI)*, Las Vegas, NV, 2018, pp. 323-328.

[7] B. Zhang, Z. Wei, J. Ren, Y. Cheng and Z. Zheng, "An Empirical Study on Predicting Blood Pressure Using Classification and Regression Trees," in *IEEE Access*, vol. 6, pp. 21758-21768, 2018.

[8] Siu-Yeung Cho, Stephen P. Morgan, Ricardo Correia, and L. Wen. "Cuff-less blood pressure measurement using fingertip photoplethysmogram signals and physiological characteristics," In *Optics in Health Care and Biomedical Optics VIII*, vol. 10820, p.p. 1082036. International Society for Optics and Photonics, 2018.

[9] G. Slapničar, M. Luštrek, and M. Marinko, "Continuous Blood Pressure Estimation from PPG Signal," *Informatica,* vol. 42, no. 1, 2018.

[10] S. Khalid, J. Zhang, F. Chen, and D. Zheng, "Blood Pressure Estimation Using Photoplethysmography Only: Comparison between Different Machine Learning Approaches," *Journal of Healthcare Engineering*, vol. 2018, 2018.

[11] P. Chiang, S. P. V. Chiluvuri, S. Dey and T. Q. Nguyen, "Forecasting of Solar Photovoltaic System Power Generation Using Wavelet Decomposition and Bias-Compensated Random Forest," *2017 Ninth Annual IEEE Green Technologies Conference (GreenTech)*, Denver, CO, 2017.

[12] GB. Moody, and RG. Mark, "A Database to Support Developmental and Evaluation of Intelligent Intensive Care Monitoring," *Computers in Cardiology,* vol. 23, pp. 657-660, 1996.

[13] P. Gent, H. Farah, N. Nes, and B. Arem,"Heart Rate Analysis for Human Factors: Development and Validation of an Open Source Toolkit for Noisy Naturalistic Heart Rate Data," *HUMMANIST Conference*, Hague, NL, Jun. 2018

[14] L. Breiman, "Random Forests," *Machine Learning*, 45(1):5–32, 2001

[15] "scikit-learn Machine Learning in Python," 2015, http://scikit-learn.org/.

[16] F. Chollet, "Keras," 2015.

[17] D. Kingma and J. Ba, "Adam: A method for stochastic optimization." *arXiv preprint arXiv:1412.6980*, 2014

[18] Association for the Advancement of Medical Instrumentation (AAMI), "American National Standard Manual," *Electronic or Automated Sphygmonanometers*, AASI/AAMI SP 10:2002, 2003.