APPLICATIONS OF
MODELLING AND SIMULATION
http://arqiipubl.com/ams eISSN 2600-8084 VOL 10, 2026, 1-10

A SHAP-Explainable Framework for Blood Pressure Prediction Based on PPG and ECG Signal Analysis

C. B. Tai[1], S. L. Loh[1,*] and A. Huong[2]

[1]Centre for Robotics and Industrial Automation, Faculty of Electrical Technology and Engineering, Universiti Teknikal Malaysia Melaka, Melaka, Malaysia
[2]Faculty of Electrical and Electronic Engineering, Universiti Tun Hussein Onn, Johor, Malaysia

[*]Corresponding author: slloh@utem.edu.my

Submitted 04 August 2025; Revised 06 November 2025; Accepted 07 November 2025; Available online 02 January 2026.

Abstract: Continuous blood pressure (BP) monitoring is crucial in managing hypertension, but current techniques are risky and uncomfortable. Extensive research has been conducted to explore the prediction of BP using features extracted from the photoplethysmogram (PPG) and the electrocardiogram (ECG). However, most of these generated features are not significantly linked with blood pressure and frequently lack a rigorous scientific explanation. This paper identifies a set of features clinically relevant to the BP prediction for predicting systolic and diastolic BP using machine learning; the process is further enhanced with Shapley Additive Explanations (SHAP) for optimal feature selection. The feature set extracted from PPG and ECG signals, and patients' demographic data are used as the input for Support Vector Regression (SVR) and Random Forest algorithms for BP prediction. In this research, it was found that Random Forest is superior to SVR. The findings from the experiment combining ECG and demographic features revealed a decrease in mean error rate by 24.13% and 81.50%, respectively, for systolic and diastolic BP prediction compared to the analysis using only PPG. SHAP-based feature selection reduced the feature set by 50% while achieving the necessary predictive capabilities. These results offer essential insights for developing a valid and noninvasive method for BP monitoring, assisting medical teams conducting clinical research on hypertension and cardiovascular illness.

Keywords: Blood pressure; Electrocardiogram; Photoplethysmogram; Random Forest; SHapley Additive exPlanations (SHAP).

1. INTRODUCTION

Reduced or uncontrolled blood pressure (BP) level, i.e., hypertension, is a major risk factor for hypertensive and ischemic heart disease. In severe circumstances, it can lead to dire health consequences, including death. According to the World Health Organization (WHO), more than 1.28 billion adults worldwide suffer from this disease [1]. In Malaysia, cardiovascular diseases, including hypertensive heart disease, claimed the second leading cause of death in 2023. Therefore, continuous BP monitoring, especially for postoperative care and managing patients with significant comorbidities, is essential to prevent health deterioration [2].

The traditional methods for measuring BP can be divided into invasive and non-invasive. Despite the extensive use of the invasive approach as the "gold standard" due to its high diagnostic accuracy and sensitivity, this method requires arterial puncture for its measurement. The procedure may pose risks to the patients, such as wound infection at the puncture site, excessive bleeding, and nerve damage [3]. Non-invasive methods, typically performed using cuff-based sphygmomanometers, offer greater patient safety and rapid measurement, although they provide only indirect BP readings. However, repeated and prolonged cuff inflation can cause significant patient discomfort, skin bruising, or inflammatory reactions. Most of these measurements are intermittent and non-continuous.

Today, technology has made it possible to estimate BP values based on photoplethysmogram (PPG) and electrocardiogram (ECG) data [4]. Since BP is a significant biomarker often used to reveal one's cardiovascular health by assessing its systolic (SBP) and diastolic (DBP) values, PPG and ECG are gaining significant attention and becoming alternative approaches for predicting these BP variables using either machine learning or deep learning techniques. Numerous studies have identified various ECG and PPG features with strong predictive value for estimating BP. Nonetheless, few studies have been carried out to document and compare the relationship between the different features of PPG and ECG, and blood pressure prediction accuracy using machine learning methods. In addition, there is no consensus on the optimal combination of these clinical features for BP prediction.

The main contribution of this paper is identifying a set of features that are clinically relevant to BP, which come from

PPG, ECG, and demographics categories. Features with high correlation (correlation > 0.9) were selected, but those with low mutual information (MI) were removed. The process is followed by feature selection using Shapley Additive Explanations (SHAP) to reduce the total features used in the model. The BP prediction models are then tested on different combinations of PPG, ECG, and individuals' demographic features, using Support Vector Regression (SVR) and Random Forest algorithms. This study uses Kaggle datasets of 500 subjects for investigation and MATLAB/Python software for simulation analysis. The goal is to establish a reliable and efficient continuous BP monitoring system suitable for healthcare and wearable technology.

In subsequent sections, Section 2 reviews the past related research, and Section 3 covers the feature extraction method and model selection. Section 4 discusses the experimental results, and the paper concludes in Section 5 with overall findings and future research directions.

2. LITERATURE REVIEW

2.1 Introduction

PPG and ECG are important pathological parameters containing rich information about blood pressure, which relates directly to a patient's cardiac condition. Numerous features can be extracted from the PPG and ECG signals to measure blood pressure and diagnose cardiac diseases. The commonly used features from PPG for BP estimation include PPG_K, also known as PPG K value, which reflects total peripheral resistance, arterial wall elasticity, and blood viscosity [5]-[7], Photoplethysmogram intensity ratio (PIR) that quantifies the ratio of PPG peak to bottom intensity [5], [7], [8], and aging index [9]-[11]. Besides, features from Acceleration Photoplethysmogram (APG) [9], [10], [12], PPG morphology and dynamic features [5], [6], [9], [10], [12]-[14], Large Artery Stiffness Index (LASI) [4], [8], [9], [15], Argument Index (AI) [4], [5], [8], [15], Kurtosis [10] and many more, are the common used features from PPG for BP estimation. The relevant features from ECG are Hjorth mobility and complexity [10], [16], [17], and heart rate [4], [6]-[9], [15], Pulse Arrival Time (PAT) peak [4], [5], [8], [12], [15], PAT max slope [4], [6]-[8], [15] and PAT onset [4], [6], [8], [15] are among the shared features of PPG and ECG for BP prediction.

The existing methods of BP prediction using PPG and ECG signals can be separated into deep learning and traditional machine learning approaches. The differences in datasets can lead to different experimental results even with the same model; thus, preprocessing techniques, feature selection methods, hyperparameter settings, and validation methods used in the past can vary depending on the data used. Some of the important approaches that offer unique advantages in capturing important features relevant to BP detection and their specific challenges are discussed in the following subsections.

2.2 Long Short-Term Memory (LSTM)

Long Short-Term Memory (LSTM) is a temporal recurrent neural network (RNN) that is widely adopted in sequential data tasks, including BP prediction. A study in [18] used Bidirectional LSTM (Bi-LSTM) with ballistocardiogram signal and features from PPG and ECG for the same purpose. The corresponding work reported a mean absolute error (MAE) and RMSE of 5.82 mmHg and 6.82 mmHg, respectively, for SBP, 5.24 mmHg and 6.06 mmHg for DBP in multi-day tests. Meanwhile, the study in [19] used Bi-LSTM as the first layer, followed by residual layers, to predict BP levels based on PPG and ECG features. The reported an MAE, standard deviation (SD), and RMSE of 6.726 mmHg, 14.505 mmHg, and 8.051 mmHg, respectively, for SBP, while 2.516 mmHg, 6.442 mmHg, and 3.998 mmHg for DBP. Other studies worth highlighting include [20], which employed an LSTM network integrated with an autoencoder to convert raw PPG signals into a continuous ABP signal. The result demonstrated strong performance in BP prediction with MAE, SD, and RMSE of 4.05 mmHg, 4.60 mmHg, and 5.25 mmHg for SBP, and 2.41 mmHg, 3.11 mmHg, and 3.17 mmHg for DBP. Although LSTM models are effective for temporal analysis, there are issues with interpretability, computational cost, data availability, and model resilience when using them to predict blood pressure.

2.3 Hybrid Convolution Neural Network with Long Short-Term Memory (CNN+LSTM)

Some studies in the literature combined Convolution Neural Network (CNN) with LSTM to form a hybrid network (CNN+LSTM) for BP level prediction. CNN offers the advantage of learning relevant features across different scales and uses a weight-sharing mechanism to help minimize memory usage compared to fully connected networks [21]. In [21], two CNNs were used to extract morphological features from PPG signals for initial SBP and DBP predictions. This was followed by two-layer stacked LSTMs to capture temporal dependencies and enhance predictions by incorporating the dynamic relationship between SBP and DBP. The study achieved ME±SD and MAE of +1.91±5.55 mmHg and 3.97 mmHg for SBP and 0.67±2.84 mmHg and 2.10 mmHg for DBP. In a separate study [22], PPG and ECG signals were used as the input to CNN+LSTM for BP prediction. The study conducted separate predictions for SBP and DBP, achieving MAE and SD values of 4.41 mmHg and 6.11 mmHg for SBP, while 2.91 mmHg and 4.23 mmHg for DBP. In [23], the study employed a CNN+LSTM architecture to capture morphological and temporal features from PPG and ECG for the simultaneous prediction of SBP and DBP. The research achieved ME±SD and MAE values of -0.02±1.6 mmHg and 1.2 mmHg for SBP, and 0.2±1.3 mmHg and 1.0 mmHg for DBP, indicating high prediction accuracy. However, the authors pointed out that the duration of data recordings may influence the accuracy of LSTM models, and model generalization is still a problem to be considered. In [24], the research examined different deep learning frameworks, incorporating Residual Network (ResNet) and WaveNet as CNNs, along with LSTM as an RNN for BP prediction. Combining ResNet with LSTM produced the best results, with an MAE and RMSE of 4.118 mmHg and 5.682 mmHg for SBP, and 2.228 mmHg and 2.986 mmHg for DBP. However, this model is also the most computationally expensive; it suffers from memory complexity when handling long-term dependencies in PPG signals.

2.4 Support Vector Regression (SVR)

One of the traditional machine learning algorithms used by many researchers for predicting BP is SVR. SVR is a type of Support Vector Machine (SVM) used for regression tasks. This method finds a line hyperplane that best fits the data points in a higher dimension [13]. This algorithm maps input features to a higher-dimensional space through a kernel function [12]. In [13], the study used Linear Regression (LR), Artificial Neural Networks (ANN), and SVR for BP prediction using a PPG signal. The study reported that the SVR outperforms other machine learning algorithms with MAE±SD of 13.57±3.23 mmHg and 8.30±1.88 mmHg for SBP and DBP estimation, respectively. This observation is supported by [4], who extracted features from PPG and ECG signals in their analysis and found that SVR outperformed Regularized Linear Regression (RLR) and ANN with MAE and SD of 12.38 mmHg and 16.17 mmHg for SBP, and 6.34 mmHg and 8.45 mmHg for DBP. Although most studies in the past used online and public datasets, some used real-world, self-collected data in their analysis. In [12], SVR performed better than Random Forest, AdaBoost, and ANN when real-world, self-collected PPG and ECG signals were used, with MAE and RMSE of 6.97 mmHg and 8.15 mmHg for SBP estimation. Meanwhile, the study in [25] proposed a hybrid model, i.e., the SVR model was used as the last output layer of the CNN model for BP prediction. The result showed a good performance with MAE±SD and RMSE of 1.23±2.45 mmHg and 1.89 mmHg for SBP, while 3.08±5.67 mmHg and 3.91 mmHg for DBP. In general, deep learning outperforms traditional machine learning. SVR often needs to be combined with strong feature engineering for more robust and scalable performance.

2.5 Random Forest

Random Forest is another traditional machine learning algorithm used by many researchers for the same problem. Random forest utilizes ensemble learning techniques, combining multiple decision trees to make predictions through random sampling with replacement [6]. The study in [26] used Random Forest for BP prediction based on PPG and ECG features. The RMSE was reported as 13.01 mmHg and 12.89 mmHg for SBP and DBP, respectively. In some studies, Random Forest outperformed other traditional machine learning methods, including SVR. Additionally, Random Forest also performed better in [8] compared to other traditional machine learning models, i.e., LR, ridge regression, SVR, and AdaBoost. After implementing GA for feature optimization, the Random Forest model achieved MAE and RMSE of 9.54 mmHg and 13.83 mmHg for SBP, and 5.48 mmHg and 6.80 mmHg for DBP. In [27], the study used the tree-based pipeline optimization tool (TPOT), an automated machine learning tool, in their analysis. Random Forest gave the best result for SBP estimation with an MAE and MSE of 6.52 mmHg and 7.48 mmHg. Another work in [28] used SVR for comparison with Random Forest. In their study, GA was used to optimize the hyperparameters of SVR, while grid search was employed for Random Forest. The study suggested that Random Forest performed better than SVR under the same conditions, with an MAE of 4.45 mmHg and 3.95 mmHg for SBP and DBP, respectively. In [29], MLR, SVR, and Random Forest were used to predict BP by combining morphological features of the ECG signal with Photoacoustic Tomography (PAT). From the results obtained, Random Forest achieved an MAE±SD of 6.12±9.52 mmHg for SBP and 4.02±6.58 mmHg for DBP. Random Forest may produce good results, but the limitation is the model's reliance on handcrafted features in BP prediction using biomedical signals.

BP prediction models often involve deep learning or machine learning techniques. Deep learning extracted features automatically while machine learning involves the physiological feature extraction to improve accuracy. Deep learning is well-known with its powerful automated modelling but with large data size. Machine learning has the substantial practical benefit over deep learning where the features used are known, which benefit to medical teams in clinical study on BP and cardiovascular diseases, and it needs much smaller size of data. Hence, machine learning is chosen over deep learning in this project.

3. METHODOLOGY

This paper proposed an optimized framework for predicting BP by leveraging morphological and dynamic features extracted from PPG and ECG signals. The framework integrates SHAP for feature selection and employs traditional machine learning models for BP prediction. The process encompasses signal preprocessing, feature extraction, and model evaluation, yielding separate SBP and DBP estimation models. The method begins with signal importation, where PPG and ECG signals are imported for analysis. These waveforms undergo preprocessing to eliminate signal noise before important features are extracted for further analysis. The features were used for model development;in this stage, two experiments were conducted: the first identifies the most effective machine learning model for BP prediction, while the second tests various feature combinations to assess their predictive power. The performance of the developed models was tested for SBP and DBP estimation, and evaluated based on the selected feature combinations. The detailed information on these key activities is elaborated in the following subsections.

3.1 Data Preprocessing

A large cleaned dataset called "PulseDB" has been published [30]. This dataset consists of 5245454 high-quality 10-second segments of ECG, PPG, and ABP signal from 5361 subjects retrieved from the MIMIC-III waveform database matched subset and the VitalDB database. One of the advantages of this dataset is that it includes subjects' identification and demographic information, allowing for the evaluation of the generalizability of models to unseen data and avoiding data leakage that could lead to overly optimistic results. Signals acquired from the VitalDB database were downsampled from 500 Hz to 125 Hz to have a consistent sampling rate with the signals from the MIMIC-III matched subset.

The study utilizes the "PulseDB" dataset, which comprises over 5 million high-quality 10-second segments of ECG, PPG, and ABP signals collected from more than 5,300 subjects. This dataset includes demographic information, enhancing the model's generalizability. In this study, 10 signal segments and gathered data on the subjects' age, gender, weight, height, and body mass index (BMI) from 500 participants are chosen, resulting in a total of 5000 data used in this project. The data is split into 60% for the training set to train the model, 20% for the validation set for hyperparameter tuning, and the remaining 20% for the test set to verify the model's performance on unseen data.

The PPG and ECG signals were resampled from 500 Hz to 125 Hz to reduce data size and computational burdens. To minimize noise in the signals, the PPG signal is filtered using an eighth-order Chebyshev-II filter with cutoff frequencies of 0.5-8 Hz, while the ECG signal is filtered with an 8^{th}-order Butterworth filter with cutoff frequencies of 0.5-40 Hz. The amplitudes of both signals are normalized to a range of 0 to 1. Despite initial preprocessing, some corrupted signal segments remain. Signals that fail to yield usable features are excluded. Additionally, SBP and DBP values are constrained to clinically relevant ranges: SBP between 80-180 mmHg and DBP between 50-130 mmHg.

The segmentation process begins with identifying the R peaks of the ECG signal using MATLAB's 'findpeaks' function. The first and last R peaks define the segment boundaries. These criteria ensure the extracted signals maintain the correct order, including checks on the number of systolic peaks in the PPG relative to the ECG R peaks and heart rate consistency through the point detection using MATLAB code.

3.2 Feature Extraction

SBP values are extracted from the ABP signal by referencing the maximum and minimum values between the first and last R peaks of the ECG signal, while the DBP value is derived by referencing the minimum value between the first R peak and the last ABP systolic peak. The averages of these values are computed to represent the final SBP and DBP for each segment.

Fiducial points representing the cardiac events in the PPG signal are critical for feature extraction. A total of 15 fiducial points are identified, including onset points, systolic peaks, and dicrotic notches. Besides, the first and second derivatives of the PPG signal, i.e., Velocity Plethysmogram (VPG) and Acceleration Plethysmogram (APG), are computed to aid in feature extraction. Figure 1 shows a PPG signal, its fiducial points, and their derivatives.

There are three points in the first PPG derivative signal important for the VPG estimation: w, y, and z, as shown in Figure 1 (center). These points can be detected using the 'findpeak' function. w refers to the maximum peak of VPG between the first onset and systolic peak of the PPG, y is the minimum point of the VPG, and z is the sub-peak of VPG after the y. Meanwhile, five points, i.e., a, b, c, d, e, and f, in the second derivatives of PPG (shown at the bottom of Figure 1) are used for APG estimation. Similarly, VPG prediction used the 'findpeaks' function to automatically identify points a, b, e, and f. Point a is an early systolic positive peak, detected by locating the maximum peak, while b is an early systolic negative peak, detected by locating the trough of the APG after a. Point e is an early diastolic positive wave, detected by locating the maximum peak after b, point f is a diastolic negative wave, detected by locating the trough after e. Points e and f correspond to the dicrotic notch and the diastolic peak of the PPG signal, respectively. The ECG signal requires the detection of five key waves: P, Q, R, S, and T, as illustrated in Figure 2. The R peak is a reference point for identifying the other waves, which are detected systematically.

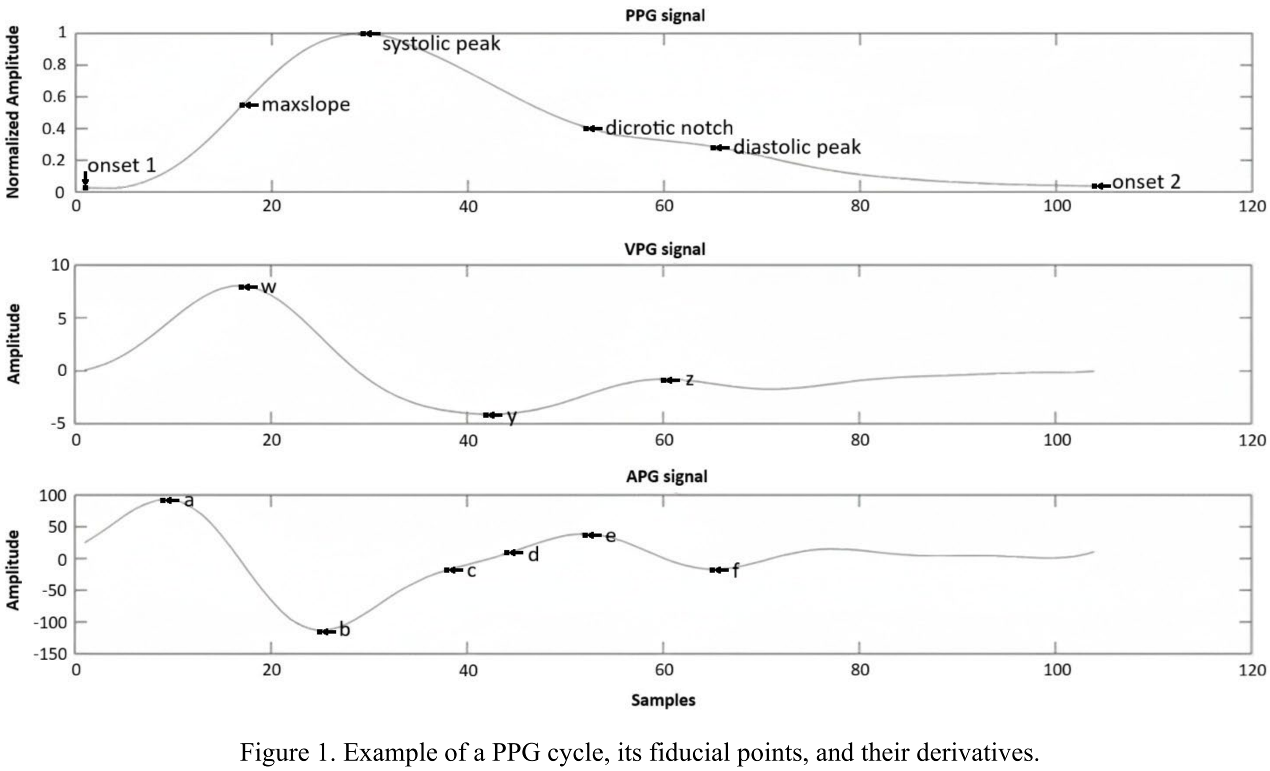


Figure 1. Example of a PPG cycle, its fiducial points, and their derivatives.

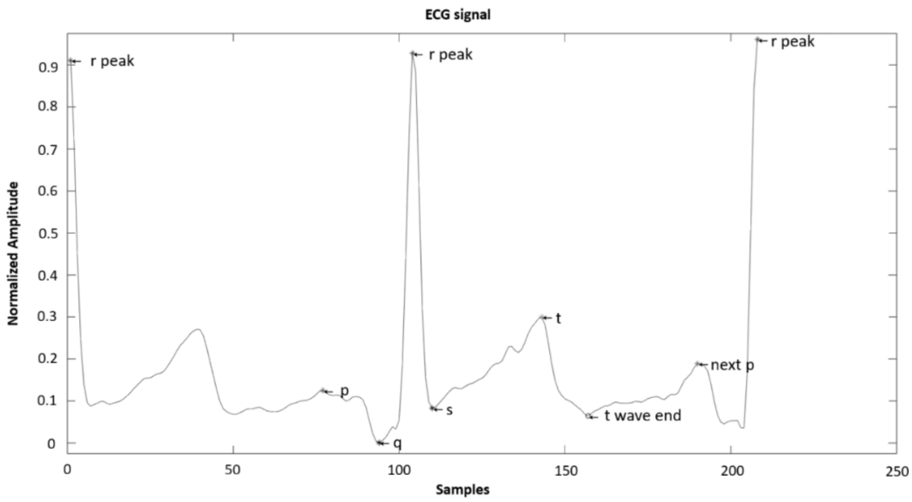


Figure 2. An example of an ECG signal and its key features.

3.3 Feature Selection and Model Development

A total of 78 features is selected and considered in the study, where 68 features are extracted from the PPG and ECG signals, and 10 demographic features are extracted from the subject. Demographic features are included to enhance model performance. All features are categorized in Table 1.

Table 1. Breakdown of features considered in the study.

Source of Feature	Number of Features	Examples of Feature
PPG	45	- Photoplethysmography Intensity Ratio, PIR - PPG characteristic value, PPG_K
ECG	20	- Heart rate - Hjorth complexity
PPG + ECG	3	- Pulse Arrival Time (PAT) peak - Pulse Arrival Time (PAT) maxslope
Demographic feature	10	- BMI - Age

SHAP was used for feature selection. SHAP is a game theoretical approach that explains the output of machine learning models [31]. The fundamental concept behind SHAP is to compute Shapley values which represent the influence of a feature on a model's prediction. Shapley values are computed by determining the average marginal contribution of a feature value across all possible combinations of feature sets, and each feature was ranked based on its contribution to the prediction performance. A correlation analysis was conducted to remove redundant features, and mutual information was used to discard features with low importance. The final feature set was determined using backward elimination, with the best model chosen based on the lowest Root Mean Square Error (RMSE).

Two machine learning models were compared: Random Forest and Support Vector Regression (SVR). Random Forest, which combines multiple decision trees, was chosen for its ability to handle non-linear relationships and reduce variance. Hyperparameters were tuned using Optuna, an open-source framework for automatic optimization.

In this study, two experiments were conducted. The first experiment compared SVR and Random Forest models for BP prediction. The second experiment tested various feature combinations to evaluate their impact on prediction accuracy. These models were evaluated using Pearson correlation coefficient (r), Mean Error (ME), Standard Deviation (SD), and Cumulative Percentage Error (CPE) expressed as:

$$r = \frac{\sum_{i=1}^{n}(y_i-\bar{y})(x_i-\bar{x})}{\sqrt{\sum_{i=1}^{n}(y_i-\bar{y})^2}\sqrt{\sum_{i=1}^{n}(x_i-\bar{x})^2}} \tag{1}$$

where y_i is the ith actual value, x_i is the i-th predicted value, n is the number of observations, $\bar{y}$ is the average of actual values, and $\bar{x}$ is the average of predicted values.

$$ME = (1/n)\sum_{i=1}^{n}(y_i - x_i) \tag{2}$$

$$SD = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(z_i - \bar{z})^2} \tag{3}$$

where z_i represents the i-th ME and $\bar{z}$ is the mean value.

$$CPE_{\leq T} = \left(\frac{Number\ of\ readings\ with\ |y_i - x_i| \leq T}{Total\ number\ of\ readings}\right) \times 100\% \quad (4)$$

where T is the error threshold (5, 10, 15 mmHg).

The model's performance was assessed using the BHS grading criteria (percentage of predictions within 5, 10, and 15 mmHg of reference), and also AAMI/ANSI/ISO 81060-2 criteria (≤ 5 mmHg mean bias and ≤ 8 mmHg standard deviation). Although the VitalDB dataset was not collected according to the formal AAMI standard and BHS protocol, the criteria were applied as a benchmark for comparative performance assessment.

4. RESULTS AND DISCUSSIONS

In this study, 78 features were extracted from PPG and ECG signals, along with SBP and DBP reference values derived from ABP signals, and demographic data of 500 subjects. The following sections discuss the findings from two key experiments that evaluated the efficacy of different machine-learning models and feature combinations for BP prediction.

4.1 Experiment 1: Model Selection

Two models, SVR and Random Forest, were optimized using the Optuna framework. The hyperparameters for each model were fine-tuned based on the validation set. Table 2 shows the SBP and DBP prediction results using SVR and Random Forest, respectively. For SBP prediction, Random Forest outperformed SVR, showing a lower ME and better overall performance across different evaluation metrics. The Random Forest model achieved an ME ± SD of 0.9673 ± 14.0735 mmHg for SBP and 0.041 ± 9.3297 mmHg for DBP, slightly better than the SVR model (as indicated by the bold font in the table). This result led to selecting Random Forest as the model for further experimentation.

Table 2. Performance of SVR and Random Forest in systolic (SBP) and diastolic blood pressure (DBP) prediction using full feature set.

Calculated prediction error	SBP prediction		DBP prediction	
	SVR	Random Forest	SVR	Random Forest
ME± SD (mmHg)	1.8092 ± 14.6449	**0.9673 ± 14.0735**	1.1894± 9.7163	**0.041 ± 9.3297**
r	0.5008	**0.5146**	**0.4142**	0.3890
$CPE_{\leq 5}$ (%)	27.3	**29.0**	**41.2**	40.1
$CPE_{\leq 10}$ (%)	52.5	**53.3**	70.0	**74.6**
$CPE_{\leq 15}$ (%)	70.8	**72.7**	88.2	**90.0**

4.2 Experiment 2: Manual Feature Selection and Analysis

Since Table 2 showed that Random Forest outperformed the SVR, the subsequent experiment used this model to investigate different feature categories manually chosen and combined to assess their significance on the prediction accuracy. Four different feature categories combinations were evaluated, i.e., (1) using all features (PPG, ECG, and demographic data), (2) PPG and ECG features without patients' demographics information, (3) PPG and patients' demographics, and (4) PPG alone.

Tables 3 and 4 show the feature analysis results for SBP and DBP, respectively. For SBP prediction, the combination of all features (PPG, ECG, and demographics) yielded the best results with an ME of 0.9673 mmHg, while using only PPG features resulted in the highest error at 1.275 mmHg. For DBP, the trend was similar, with the full feature set producing the lowest ME of 0.041 mmHg and the PPG-only approach showing the highest error. These findings indicate that including ECG and demographic features enhances the models' prediction accuracy.

Table 3. Random Forest performance in systolic blood pressure (SBP) prediction using different feature categories.

ME ± SD in SBP prediction (mmHg)	**Features categories**			
	PPG+ECG +Demographic	PPG+ECG	PPG +Demographic	PPG
	0.9673 ± 14.0735	1.0798 ± 14.0859	1.0409 ±14.7398	1.275 ± 14.7802
r	**0.5146**	0.5118	0.4382	0.4362
$CPE_{\leq 5}$ (%)	29.0	**29.4**	26.1	27.1
$CPE_{\leq 10}$ (%)	53.3	**53.7**	49.8	52.2
$CPE_{\leq 15}$ (%)	72.7	**73.4**	70.5	69.7

Table 4. Random Forest performance in diastolic blood pressure (DBP) prediction using different feature categories.

DBP	PPG + ECG + Demographic	PPG + ECG	PPG + Demographic	PPG
ME $\pm$ SD (mmHg)	**0.041 $\pm$ 9.3297**	0.1227 $\pm$ 9.3486	0.1936 $\pm$ 9.7038	0.2216 $\pm$ 9.8246
r	**0.3890**	0.3843	0.3037	0.2771
$CPE_{\leq 5}$ (%)	**40.1**	**40.1**	38.3	39.7
$CPE_{\leq 10}$ (%)	**74.6**	74.3	71.3	71.1
$CPE_{\leq 15}$ (%)	90.0	90.0	**90.1**	88.3

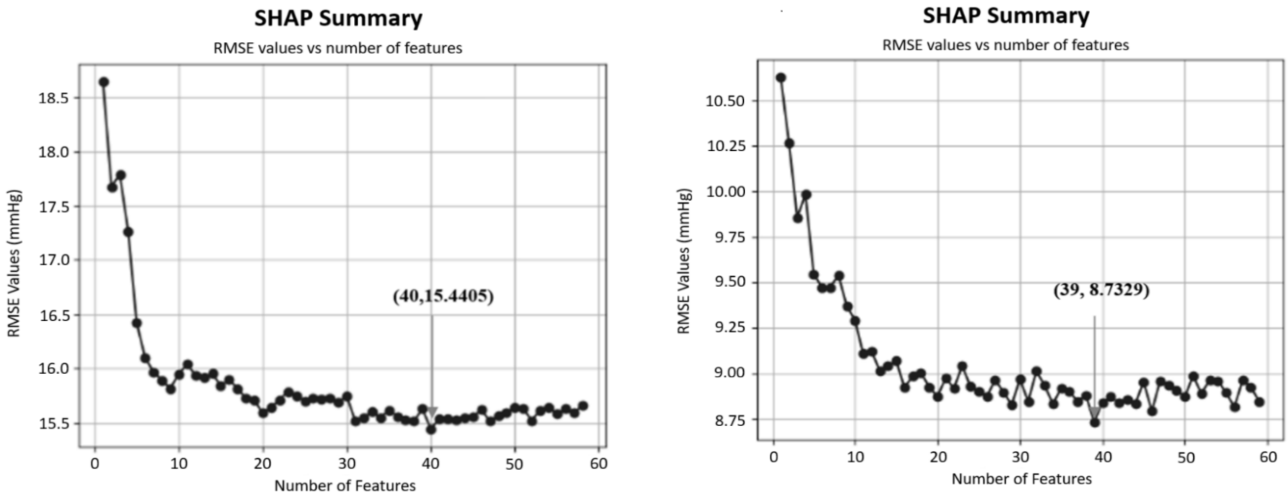
4.3 Automatic Feature Selection Using SHAP

To further optimize the model, SHAP was employed to rank the importance of the considered features. Features with high correlation (correlation > 0.9) were selected, but those with low mutual information (MI) were removed, reducing the feature set to 58 for SBP and 59 for DBP. The remaining features were ranked by their mean absolute SHAP values, and iterative feature removal was performed to identify the optimal feature combination.

Figure 3 shows the RMSE of the Random Forest model across different numbers of features used for (a) SBP and (b) DBP predictions. Based on this figure, the lowest RMSE (15.4405 mmHg) in SBP prediction was achieved with the top-ranked 40 features, whereas top-ranked 39 features were used for DBP, producing an RMSE of 8.7329 mmHg. According to SHAP, $Angle_{ed}$ and PT are the features that contribute the most in predicting SBP and DBP, respectively. $Angle_{ed}$ is computed by determining the slope from point e to the point d in the APG, as shown in Figure 1, while PT is computed by measuring the time difference between P and T in the ECG signal, as shown in Figure 2.

4.4 Comparison of Optimized Feature Combinations

The models trained with the optimized feature sets were compared with those using the full feature sets, and the results are tabulated in Table 5. For SBP, the optimized model with 40 features slightly improved ME from 0.9673 mmHg to 0.9464 mmHg while reducing the feature count by approximately 48.72%. For DBP, the optimized model with 39 features maintained similar accuracy with a slight ME increase, demonstrating the effectiveness of the feature selection process in reducing model complexity without sacrificing performance.

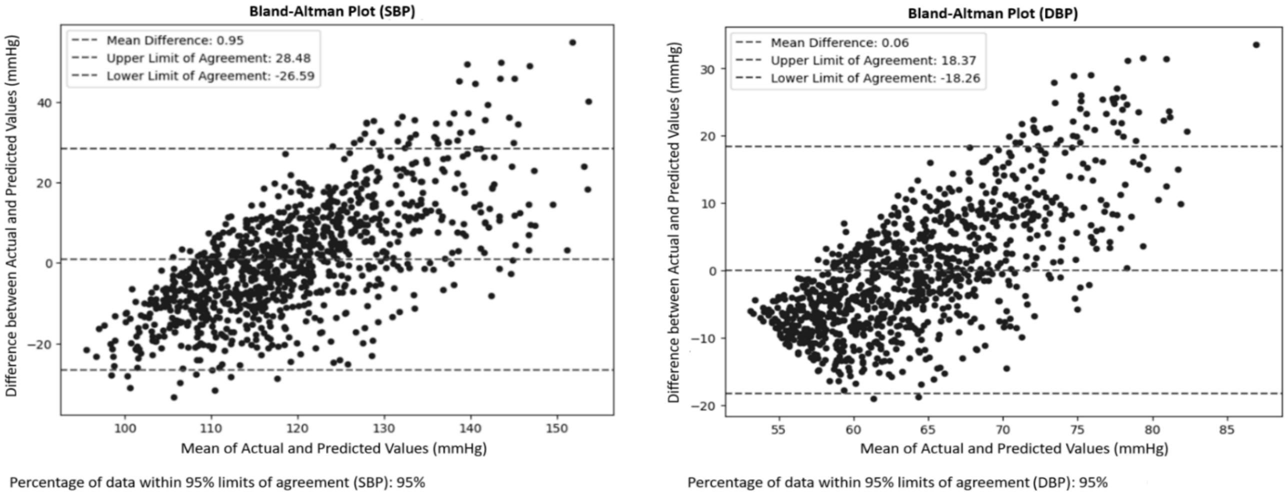


(a) (b)

Figure 3. RMSE of the Random Forest model across different numbers of features used for (a) systolic blood pressure (SBP) and (b) diastolic blood pressure (DBP) prediction.

Table 5. Comparison between full and selected features. All predictions are made using Random Forest.

	SBP		DBP	
	PPG+ECG+ Demographic	Feature selection using SHAP	PPG+ECG+ Demographic	Feature selection using SHAP
Features count	78	40	78	39
ME$\pm$ SD (mmHg)	0.9673 $\pm$ 14.0735	**0.9464 ± 14.0502**	**0.041$\pm$ 9.3297**	0.0574 ± 9.3436
r	0.5146	**0.5161**	**0.3890**	0.3868
$CPE_{\leq 5}$ (%)	**29.0**	28.6	**40.1**	39.6
$CPE_{\leq 10}$ (%)	53.3	**53.6**	74.6	**74.8**
$CPE_{\leq 15}$ (%)	72.7	**74.2**	90.0	**90.4**



(a) (b)

Figure 4. Bland-Altman plot for the estimated (a) systolic blood pressure (SBP) and (b) diastolic blood pressure (DBP).

Figure 4 shows the Bland-Altman plot for SBP and DBP, respectively. Referring to Figure 4, the mean differences are 0.95 mmHg and 0.06 mmHg for SBP and DBP, respectively. These values differ insignificantly from 0, indicating the nonexistence of fixed bias for both models. The differences are more widely spread at lower and higher mean values, indicating variability in prediction accuracy across the range of values. Several points lie outside the 95% limits of agreement, suggesting some outliers. However, for both models, 95% of the data is falls within these limits, implying overall good agreement between the actual and predicted values.

The models were evaluated against the Association for the Advancement of Medical Instrumentation (AAMI) Standard and British Hypertension Society (BHS) Protocol, widely used benchmarks for BP measurement devices. The SBP model had an ME of 0.9464 mmHg, and the DBP model had an ME of 0.0574 mmHg, both within the acceptable range of the AAMI standard. However, the SD for both SBP (14.0502 mmHg) and DBP (9.3436 mmHg) exceeded the AAMI threshold of 8 mmHg. According to the BHS protocol, both models performed below a grade C, with cumulative percentage errors of 28.6% and 39.6% for SBP and DBP, respectively, within the $CPE_{\leq 5}$ category.

Recent research has highlighted that data leakage is a common issue overlooked in many BP studies, leading to overly optimistic results [32-34]. When measures are taken to prevent data leakage, ensuring that records from the same subject do not appear in both the training and testing sets, many deep learning and traditional machine learning models fail to fulfill the AAMI and BHS criteria.

Furthermore, in [32], the study also addressed the issue of data leakage in their results. XGBoost and CatBoost were utilized to predict BP based on features extracted from PPG signal. The findings indicated that CatBoost demonstrated superior performance, achieving a MAE and ME±SD of 5.368 mmHg and 0.050±7.837 mmHg for SBP, and 2.521 mmHg and 0.022±3.767 mmHg for DBP. However, these results were obtained when considering data leakage. When the researchers minimized the data leakage, the MAE and ME±SD increased to 18.209 mmHg and -1.230±22.368 mmHg for SBP, and 7.524 mmHg and -0.257±9.784 mmHg for DBP.

Hence, the proposed models were further compared to this similar study in [32]. The Random Forest model with SHAP-based feature selection outperformed the referenced study's CatBoost model, as shown in Table 6, particularly in cumulative percentage errors under different thresholds. This highlights the effectiveness of the proposed method in providing accurate BP predictions despite not meeting all the stringent criteria of the AAMI and BHS standards.

Table 6. A comparison of the prediction performance between the proposed method and the previous study.

Criteria	**Dias et al. [32]**	**Proposed method**
Dataset (subject/signal segment)	UCI dataset (- / 50182)	VitalDB (500 / 5000)
Method	CatBoost	Random Forest+ SHAP feature selection
Number of features used (SBP/DBP)	133 / 133	39 / 40
ME±SD (SBP/DBP)	-1.23±22.368 mmHg / -0.257±9.784mmHg	**0.9464 ± 14.0502** mmHg / **0.0574 ± 9.3436** mmHg
r (SBP/DBP)	0.218 / 0.306	**0.5146 / 0.3890**
$CPE_{\leq 5}$ (SBP/DBP) (%)	16.2 / **40.7**	**28.6** / 39.6
$CPE_{\leq 10}$ (SBP/DBP) (%)	32.3 / 73.2	**53.6 / 74.8**
$CPE_{\leq 15}$ (SBP/DBP) (%)	47.6 / 90.1	**74.2 / 90.4**

Overall, the Random Forest model demonstrated superior performance compared to SVR, and the integration of SHAP significantly optimized feature selection. Reducing the feature count improved the model's computational efficiency and reduced the risk of overfitting. Although the models did not fully meet the AAMI and BHS standards, they exhibited promising results, particularly in real-time applications or scenarios with limited computational resources.

5. CONCLUSION

The features extracted from PPG and ECG signals, which are clinically relevant to BP, were used in SVR and Random Forest for blood pressure prediction. The Random Forest algorithm demonstrated superior performance compared to SVR in predicting SBP and DBP. Meanwhile, the feature analysis results revealed that the prediction models' performance varies with the features considered in the experiments. Including ECG signals and demographic features reduces the calculated ME by approximately 24.13% for SBP and 81.50% for DBP, compared to using only PPG signals. Optimized feature selection using SHAP enhanced predictive accuracy even with significantly fewer features, i.e., a reduction of 48.72% for SBP and 50% for DBP. The model evaluation based on AAMI standards and BHS protocols indicated that even though the model could not fulfill the criteria, it is consistent with findings from recent papers.

The results demonstrate that the extracted features are associated with BP and suitable for application in the BP prediction model. The inclusion of ECG signals shows potential for enhancing BP predictions. Nonetheless, creating a high-performance BP prediction model that fulfills the AAMI standard and BHS protocol remains challenging due to variations in these signals among individuals and the complexity of physiological factors. In terms of feature analysis, it is advisable to incorporate more features prior to feature selection to thoroughly explore high-potential features for predicting BP. Besides that, exploring features that do not rely on the detection of fiducial points should be emphasized to enhance the model's applicability to real-life conditions. Further exploration into integrating additional signals, such as ballistocardiogram signals, could further be explored to enhance prediction accuracy.

ACKNOWLEDGEMENT AND FUNDING

The authors wish to express their gratitude to the Center for Robotics and Industrial Automation (CeRIA), the Center for Research and Innovation Management (CRIM), Faculty of Electrical Technology and Engineering (FTKE), and Universiti Teknikal Malaysia Melaka (UTeM) for supporting the research and publication

DECLARATION OF CONFLICTING INTERESTS

The authors declare no potential conflicts of interest with respect to the research and publication of this article.

REFERENCES

[1] World Health Organization, Hypertension. https://www.who.int/news-room/fact-sheets/detail/hypertension, 2023 (accessed 21.07.2025).

[2] Statista, Malaysia: most common causes of death 2021. https://www.statista.com/statistics/1343021/malaysia-most-common-causes-of-death/ (accessed 21.07.2025).

[3] B. Alexander, M. Cannesson and T. J. Quill, Blood Pressure Monitoring, *Anesthesia Equipment: Principles and Applications*, 2013, 273-282.

[4] M. Kachuee, M. M. Kiani, H. Mohammadzade and M. Shabany, Cuff-less high-accuracy calibration-free blood pressure estimation using pulse transit time, *Proceedings of IEEE International Symposium on Circuits and Systems*, Lisbon, Portugal, 2015, 1006-1009.

[5] S. Chen, Z. Ji, H. Wu and Y. Xu, A non-invasive continuous blood pressure estimation approach based on machine learning, *Sensors*, 19(11), 2019, 2585.

[6] G. Ma, J. Zhang, J. Liu, L. Wang and Y. Yu, A multi-parameter fusion method for cuffless continuous blood pressure estimation based on electrocardiogram and photoplethysmogram, *Micromachines*, 14(4), 2023, 804.

[7] F. Miao, N. Fu, Y. -T. Zhang, X. -R. Ding, X. Hong, Q. He and Y. Li, A novel continuous blood pressure estimation approach based on data mining techniques, *IEEE Journal of Biomedical and Health Informatics*, 21(6), 2017, 1730-1740.

[8] G. Thambiraj, U. Gandhi, U. Mangalanathan, V. J. M. Jose and M. Anand, Investigation on the effect of Womersley number, ECG and PPG features for cuffless blood pressure estimation using machine learning, *Biomedical Signal Processing and Control*, 60, 2020, 101942.

[9] J. Liu, S. Hu, Z. Xiao, Q. Hu, D. Wang and C. Yang, A novel interpretable feature set optimization method in blood pressure estimation using photoplethysmography signals, *Biomedical Signal Processing and Control*, 86, 2023, 105184.

[10] E. Finnegan, S. Davidson, M. Harford, P. Watkinson, L. Tarassenko and M. Villarroel, Features from the photoplethysmogram and the electrocardiogram for estimating changes in blood pressure, *Scientific Reports*, 13(1), 2023, 1-20.

[11] M. Liu, L. M. Po and H. Fu, Cuffless blood pressure estimation based on photoplethysmography signal and its second derivative, *International Journal of Computer Theory and Engineering*, 9(3), 2017, 202.

[12] M. K. F. Wong, H. Hei, S. Z. Lim and E. Y. -K. Ng, Applied machine learning for blood pressure estimation using a small, real-world electrocardiogram and photoplethysmogram dataset, *Mathematical Biosciences and Engineering*, 20(1), 2022, 975-997.

[13] S. Kılıçkaya, A. Güner and B. Dal, Comparison of Different machine learning techniques for the cuffless estimation of blood pressure using PPG signals, *Proceedings of 2020 International Congress on Human-Computer Interaction, Optimization and Robotic Applications (HORA)*, Ankara, Turkey, 2020, 1-6.
[14] X. -R. Ding, B. P. Yan, Y. -T. Zhang, J. Liu, P. Su and N. Zhao, Feature exploration for knowledge-guided and data-driven approach based cuffless blood pressure measurement, *arXiv*, 2019, 1-4.
[15] M. Kachuee, M. M. Kiani, H. Mohammadzade and M. Shabany, Cuffless blood pressure estimation algorithms for continuous health-care monitoring, *IEEE Transactions on Biomedical Engineering*, 64(4), 2017, 859-869.
[16] M. Simjanoska, M. Gjoreski, M. Gams and A. Madevska Bogdanova, Non-invasive blood pressure estimation from ECG using machine learning techniques, *Sensors*, 18(4), 2018, 1160.
[17] S. Yang, W. S. W. Zaki, S. P. Morgan, S-Y. Cho, R. Correia, L. Wen and Y. Zhang, Blood pressure estimation from photoplethysmogram and electrocardiogram signals using machine learning, *Nottingham ePrints*, University of Nottingham, China, 2018.
[18] D. Lee, H. Kwon, D. Son, H. Eom, C. Park, Y. Lim, C. Seo and K. Park, Beat-to-beat continuous blood pressure estimation using bidirectional long short-term memory network, *Sensors*, 21(1), 2020, 96.
[19] Y. -H. Li, L. N. Harfiya, K. Purwandari and Y.-D. Lin, Real-time cuffless continuous blood pressure estimation using deep learning model, *Sensors*, 20(19) 2020, 5606.
[20] L. N. Harfiya, C. -C. Chang and Y. -H. Li, Continuous Blood pressure estimation using exclusively photopletysmography by LSTM-based signal-to-signal translation, *Sensors*, 21(9), 2021, 2952.
[21] J. Esmaelpoor, M. H. Moradi and A. Kadkhodamohammadi, A multistage deep neural network model for blood pressure estimation using photoplethysmogram signals, *Computers in Biology and Medicine*, 120, 2020, 103719.
[22] S. Baker, W. Xiang and I. Atkinson, A hybrid neural network for continuous and non-invasive estimation of blood pressure from raw electrocardiogram and photoplethysmogram waveforms, *Computer Methods and Programs in Biomedicine*, 207, 2021, 106191.
[23] D. U. Jeong and K. M. Lim, Combined deep CNN–LSTM network-based multitasking learning architecture for noninvasive continuous blood pressure estimation using difference in ECG-PPG features, *Scientific Reports*, 11(1), 2021, 1-8.
[24] A. Paviglianiti, V. Randazzo, S. Villata, G. Cirrincione and E. Pasero, A comparison of deep learning techniques for arterial blood pressure prediction, *Cognitive Computation*, 14(5), 2021, 1689-1710.
[25] S. Rastegar, H. Gholam Hosseini and A. Lowe, Hybrid CNN-SVR blood pressure estimation model using ECG and PPG signals, *Sensors*, 23(3), 2023, 1259.
[26] A. Tiloca, G. Pagana and D. Demarchi, A random tree based algorithm for blood pressure estimation, *Proceedings of 2020 IEEE MTT-S International Microwave Biomedical Conference*, Toulouse, France, 2020, 1-4.
[27] S. M. Fati, A. Muneer, N. A. Akbar and S. M. Taib, A continuous cuffless blood pressure estimation using tree-based pipeline optimization tool, *Symmetry*, 13(4), 2021, 686.
[28] X. Chen, S. Yu, Y. Zhang, F. Chu and B. Sun, Machine learning method for continuous noninvasive blood pressure detection based on random forest, *IEEE Access*, 9, 2021, 34112-34118.
[29] H. Wang, Random forest based blood pressure prediction model from ECG And PPG Signal, *Proceedings of 12th International Conference on Bioscience, Biochemistry and Bioinformatics*, Tokyo, Japan, 2022, 1-6.
[30] W. Wang, P. Mohseni, K. L. Kilgore and L. Najafizadeh, PulseDB: A large, cleaned dataset based on MIMIC-III and VitalDB for benchmarking cuff-less blood pressure estimation methods, *Frontiers in Digital Health*, 4, 2023, 1-16.
[31] V. Fleischhauer, A. Feldheiser and S. Zaunseder, Beat-to-Beat blood pressure estimation by photoplethysmography and its interpretation, *Sensors*, 22(18), 2022, 7037.
[32] F. M. Dias, T. B. S Costa, D. A. C Cardenas, M. A. F Toledo, J. E. Krieger and M. A. Gutierrez, A machine learning approach to predict arterial blood pressure from photoplethysmography signal, *Proceedings of Computing in Cardiology*, Tampere, Finland, 2022, 1-4.
[33] T. B. D. S. Costa, F. M. Dias, D. A. C. Cardenas, M. A. F. D. Toledo, D. M. D. Lima and J. E. Krieger, Blood pressure estimation from photoplethysmography by considering intra- and inter-subject variabilities: Guidelines for a fair assessment, *IEEE Access*, 11, 2023, 57934-57950.
[34] S. González, W. -T. Hsieh and T. P. -C. Chen, A benchmark for machine-learning based non-invasive blood pressure estimation using photoplethysmogram, *Scientific Data*, 10(149), 2023, 1-16.