

Cuffless Blood Pressure Estimation Based on both Artificial and Data-driven Features from Plethysmography

Huan Li¹(✉), Yue Wang², and Yumpeng Guo³

College of Software Engineering, Sichuan University, Chengdu 610065, China
lihuan@stu.scu.edu.cn

Abstract. Blood pressure (BP) is an important indicator of individuals' health conditions for the prevention or treatment of cardiovascular disease. However, conventional measurements require inconvenient cuff-based instruments and are not able to detect continuous blood pressure. Advanced methods utilize machine learning to estimate BP by constructing artificial features in plethysmography (PPG) or using an end-to-end deep learning framework to estimate BP directly. Empirical features are limited by current research on cardiovascular disease and are not sufficient to express BP variability, while data-driven approaches neglect expert knowledge and lack interpretability. To address this issue, in this paper we propose a method for continuous BP estimation that extracts both artificial and data-driven features from PPG to take advantage of expert knowledge and deep learning at the same time. Then a deep residual neural network is designed to reduce information redundancy in the gathered features and refine high-level features for BP estimation. The results show that our proposed methods outperforms the compared methods in three commonly used metrics.

Keywords: Blood pressure estimation · Plethysmography · Artificial features · Data-driven features

1 Introduction

As the World Health Organization reports, an increasing number of people today are suffering from cardiovascular diseases[1]. Blood pressure(BP) is one of the most critical physiological factors of human health and is necessary for the accurate diagnosis of hypertension[2]. However, traditional BP measurements are not convenient and cannot detect continuous BP. They can be classified into invasive and noninvasive groups. In hospitals, invasive methods such as arterial puncture measurements are widely used in the intensive care unit (ICU). Although carefully prepared, they still carry a risk to patients not to mention their comfort[8]. For noninvasive methods, cuff-required devices such as sphygmomanometers are the most commonly used. However, the cuff must be inflated during measurement, which causes discomfort and does not allow for continuous BP. In addition, this

method requires advanced training and is therefore not quite friendly to elderly individuals, who need BP monitoring the most[9].

To solve these problems, many studies have been conducted to obtain accurate continuous BP via other noninvasive convenient-detection physiological signals. Pulse waves (PWs) are periodic blood ripples that travel along the arteries produced by heartbeat[3]. It can be easily measured on the finger end or radial artery by the plethysmography (PPG) method[5]. Based on the wave propagation theory of fluids in elastic pipes, two main kinds of BP measurements using PW have been developed[6]. The first group utilizes pulse wave velocity (PWV), which uses multichannel PPG signals to calculate the velocity of the pressure wave propagation in the vessels. Nevertheless, the PWV-based method requires complex calibration procedures for individual physiological parameters which are not practical yet. Another kind of method is based on the analysis of a single PPG signal. Researchers have focused on the morphological features of PPG[10]. They apply several analysis schemes to the original PPG signal or its derivatives to construct better artificial features. However, this kind of empirical feature may be limited by expert knowledge and cannot sufficiently express BP variability. Thus, some deep learning models are also applied to extract features from the PPG signal directly. However, these data-driven estimation methods neglect the current cardiovascular research and therefore have poor medical interpretability.

Inspired by the latest work of Hong[7] and Deng[4], this paper proposes a method for BP estimation based on a single PPG signal. To express more information in features and improve the robustness of estimation, we utilize existing knowledge and a deep learning model to extract empirical and data-driven features at the same time. Specifically, we gather features from two designed branches: one branch constructs morphological features artificially based on current research while the other branch uses an LSTM network to extract data-driven features. Despite the information redundancy in those features, an efficient residual deep learning model is introduced to filter and refine valuable features in an end-to-end model. The experimental results confirm the validity of each module in our method and indicate that it surpasses the compared methods.

The rest of this paper is organized as follows. Section 2 presents the current works on the empirical features or deep learning frameworks. Section 3 describes each module of the proposed method in detail. Section 4 compares our model with other methods, while Section 5 concludes the paper.

2 Related Work

In this section, we briefly introduce some related works that implement continuous BP estimation by extracting artificial features from PPG signals or using deep learning techniques. Kachuee extracts features from PPG and electrocardiograms (ECGs), and collects additional personal information(age, sex, etc.) for calibration, then various conventional machine learning methods are used for es-

timation[11]. Xie extracted time features, area features, and ratio features from the preprocessed PPG signal and then uses a regression model method to estimate BP[12]. Aman Gaurav extracted 46 features from PPG and its derivatives while the target BP is estimated with an artificial neural network (ANN)[9]. The ANN is designed to reduce the information redundancy in the extracted features. These methods merely extract features artificially, which might neglect some implicit information and are not sufficient enough to express BP variability.

In addition to traditional machine learning methods based on artificial features, more end-to-end deep learning techniques are applied to extract features through a data-driven approach. Su built a sequence-to-sequence deep recurrent neural network to directly estimate arterial blood pressure (ABP)[13]. Shimazaki used a deep convolutional neural network (CNN) to estimate BP from PPG and its third and fourth order derivatives, and personal information (age, sex, height, weight, etc.) is also utilized.[14]. Despite the enrichment of the input data, his method still merely utilizes a data-driven feature extraction approach and neglects expert knowledge. Thus, in this paper, we explicitly take advantage of both artificial and data-driven features to capture more information from the PPG signal and make our method more robust.

3 Proposed Model

In general, the model we proposed is a combination of an LSTM branch and an MLP-included feature engineering branch embedded in a deep neural network. These two branches are responsible for extracting artificial features and data-driven features respectively. The final module gathers the extracted features, and refines the information for BP estimation. Since systolic blood pressure (SBP) and diastolic blood pressure (DBP) are both required in diagnosing hypertension, the targets of our model are SBP and DBP. The overview of our model is illustrated in Figure.1.

To obtain the SBP and DBP at time t , the empirical feature extraction branch utilizes the last valid cycle in the PPG signal before t to construct 21 artificially set features and obtains high-level patterns with a simple MLP. The data-driven feature extracting branch uses a single layer LSTM to learn the implicit features from the input PPG sequence of length L before moment t . The feature gathering and multichannel output module concatenates the features extracted by two branches and captures deeper information from the fusion of those features in a deep learning framework. These two branches and feature gathering modules comprise our complete end-to-end model to estimate SBP and DBP simultaneously.

$$SBP = \text{Maximum}(ABP \text{ in the following minute}) \quad (1)$$

$$DBP = \text{Minimum}(ABP \text{ in the following minute}) \quad (2)$$

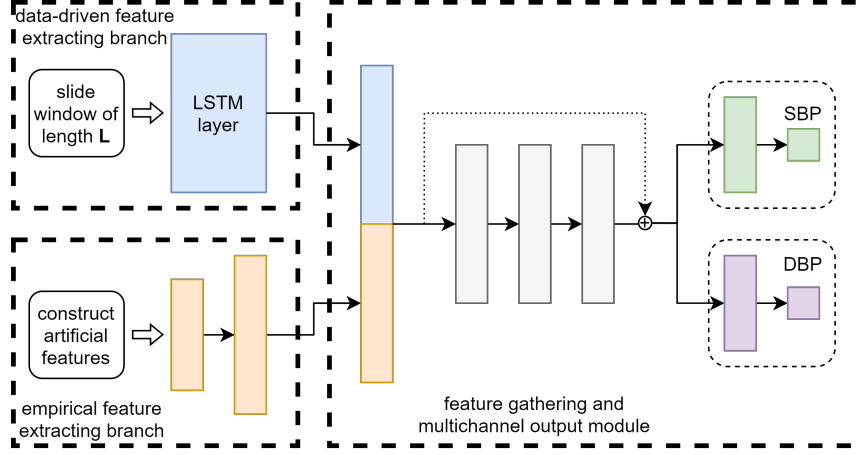


Fig. 1: The architecture of the proposed method

3.1 Empirical Feature Extracting Branch

This branch includes two main processes. We first capture the last complete valid PPG cycle before moment t from the original PPG sequence. Then 21 artificially set features will be constructed from that cycle. Finally, the constructed features are fed into a simple MLP to obtain high-level empirical patterns.

Segmenting the Last Valid Cycle Figure.2 shows an example fluctuation of a normal PPG cycle, which can be broadly divided into systolic and diastolic states. During diastole, the aortic valve closes causing some blood to rush back, which results in a repulse wave in the diastolic part called the dicrotic wave.

Although the pulse wave has the same periodicity as the heartbeat, the actual fluctuations of the PPG signal are rather complex because of the volatile heart rate and external interference. To guarantee the value of artificially constructed features, we need to filter those abnormal cycles by the following rules:

1. The deviation between A and A' should be less than $1/10$ of the main peak amplitude. The heights of A and A' represent the arterial pressure in the heartbeat interval, which should not be mutated. The sudden change is usually caused by external disturbances, for example, the movement of the device or the subject under measurement.
2. The height of both the dicrotic notch and dicrotic crest should be between 25% and 75% of the main peak amplitude. The dicrotic wave happens due to a small amount of refluxed blood after the aortic valve closes during diastole. Therefore, those dicrotic waves that happen too early or too late or too drastically mean that there is some error in the measurement process.

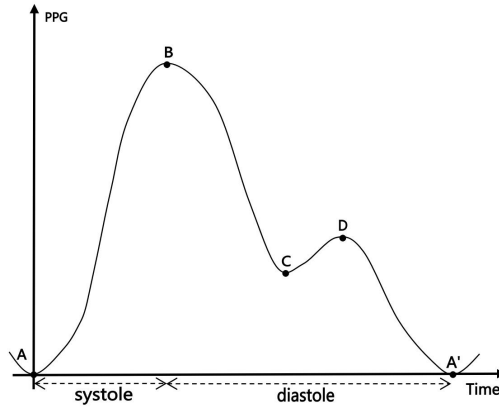


Fig. 2: An example waveform of one PPG cycle. Points A and A' are the beginning and end of one PPG cycle respectively; B is the systolic peak; C is the dicrotic notch and D is the dicrotic peak also called the diastolic peak.

Table 1: Further information of artificially constructed features

category	id	description
magnitude features	1	The height of point B
	2	The height of point C
	3	The height of point A
	4	The height deviation between point E and D
temporal features	5	The total time of the systole
	6	The time interval from A to reach 25% height of systole amplitude
	7	The time interval from A to reach 50% height of systole amplitude
	8	The time interval from A to reach 75% height of systole amplitude
	9	The total time of the diastole
	10	The time interval from B to reach 75% height of diastole amplitude
	11	The time interval from B to reach 50% height of diastole amplitude
	12	The time interval from B to reach 25% height of diastole amplitude
	13	The time interval between B and C
	14	The time interval between C and D
	15	The time interval from C to reach 50% height of dicrotch pulse amplitude
	16	The time interval from D to reach the height of C
	17	The time interval from D to reach 50% height of dicrotch pulse amplitude
area features	18	The integral from A to B
	19	The integral from B to C
	20	The integral from C to D
	21	The integral from D to A'

Artificially Construct Features After analyzing the PPG curve, we design 21 simple features including magnitude features, temporal features, and area features. Further descriptions of these features are shown in Table.1.

Each kind of feature has its physiological significance. The magnitude features have an inner correlation with age and the arterial characteristic information of a subject which greatly affects the SBP and DBP. The temporal and area features correspond to the force and blood volume during ventricular systole and diastole[15].

Then a neural layer is designed to turn the 21 obtained artificial features into a more informative form. We use the typical linear neural layer with the ReLU function to perform primary information fusion among those empirical features. Although further feature gathering will be subsequently carried out, this fusion process is still necessary since it can discover some inner correlations among artificial features without the interference of data-driven features.

3.2 Data-driven Feature Extracting Branch Based on LSTM

To extract data-driven features, long short-term memory (LSTM) [16] is utilized. The excellent capability of LSTM to extract information from long sequences is achieved by the ingenious cooperation of its cell and input gate, output gate, and forget gate[18]. The cell is designed to remember valuable information over arbitrary time intervals, while the three gates regulate the update of the cell state.

With the excellent power, LSTM could effectively capture the implicit features from the input PPG sequence which reflects the softness of blood vessels or the condition of the subject's heart. Since this process is data-driven, LSTM may extract some hidden features that are neglected in the empirical branch.

3.3 Feature Gathering and Multichannel Output Module

This module is designed to gather the features from two previous branches and to estimate both SBP and DBP. The processes of obtaining empirical features and data-driven features are isolated, so that some information redundancy may exist.

To reduce the information redundancy and capture some high-level implicit patterns, the gathered features are fed forward into a simple MLP. However, those additional layers make our model much deeper, which could lead to gradient vanishing or exploding problems during the end-to-end training process. To overcome this issue, we add a residual connection[22], which is a sort of skip-connection that learns residual parameters with reference to the layer inputs. By adding the shortcut, the gathering module can obtain some new complex information without missing existing valuable features.

Although SBP and DBP both reflect the human cardiovascular status which means that the features they depend on are quite similar, the mapping function might be rather different so it is necessary to build an output channel for

each estimation target. After obtaining the condensed features, we build two isolated channels for SBP and DBP estimation. This mode not only reinforces the representation of shared patterns of different targets but also maintains the task-oriented patterns for each task. The evaluation indicator Eq.3 is used to analyze the estimation accuracy.

$$MAE = \frac{1}{n} \sum_{i=1}^n \left| S\hat{B}P_i - SBP_i \right| + \left| D\hat{B}P_i - DBP_i \right| \quad (3)$$

4 Experiments

4.1 Experimental Setup

Dataset and Data Preprocessing In this paper, we use the Multiparameter Intelligent Monitoring in Intensive Care (MIMIC) dataset[19], which is publicly available and commonly utilized in biomedical and health research. This dataset provides various clinical physiological signals of thousands of intensive care unit (ICU) patients over weeks. The signals required in this paper, arterial blood pressure (ABP) and PPG are both sampled at 125Hz.

For our experimentation, we filtered the subjects for lack of ABP or PPG and randomly selected 100 subjects. Then the records of the beginning and end of 3 hours are deleted because of null values and strong fluctuation. The input data consist of two parts, the L-second-long segmentation of PPG before t and the 21 artificial features extracted from the last PPG cycle. To obtain the target output of time t, we pick the 60-second-long segmentation of ABP before t and then analyze its maximum and minimum as SBP and DBP respectively. The proportions of SBP and DBP are shown in Figure.3. To ensure the robustness of the model, we select 8000 samples from each subject.

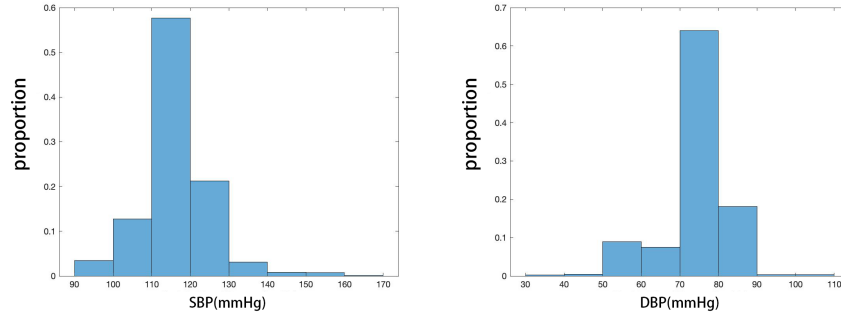


Fig. 3: the proportion of target SBP and DBP of our samples

Basic Training Settings The samples are split into three subsets, training set, validation set, and testing set with a ratio of 6:2:2. The obtained PPG segmentation data and target BP data are not normalized but the artificial feature data are normalized to $[0, 1]$ via the feature-specific scope obtained from the training set.

Performance Metrics We choose three widely used metrics to evaluate the performance of each estimation method. Mean Absolute Percentage Error (MAPE), Mean Absolute Error (MAE), Root Mean Squared Error (RMSE). The definitions follow:

$$MAPE = \frac{100\%}{n} \sum_{i=1}^n \left| \frac{\hat{y}_i - y_i}{y_i} \right| \quad (4)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |\hat{y}_i - y_i| \quad (5)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2} \quad (6)$$

4.2 Competing Methods

The methods chosen in the comparison experiments cover classical machine learning techniques and some state-of-the-art approaches.

These chosen machine learning methods are the most typical schemes and perform well in many aspects, including multilayer perceptron (MLP)[17], convolution neural network (CNN)[20], and support vector machine (SVM)[21]. The CNN only uses the original PPG segmentation, while MLP and SVM merely utilize the normalized artificial features. All competing models have a similar number of parameters as our proposed model. The SVM method does not need the validation set, so it uses both the training set and the validation set for training.

We also compare our method with 3 current related works, including the FFT-based neural network by Xing, X.[23], a calibration-free CNN by Schlesinger, O.[24], and a fully convolutional network using only the PPG signal proposed by Baek, S.[25]

4.3 Results and Analysis

Main Results Table.2 shows the results of our model and competing methods. Comparing the results, we can find that except for the RMSE when estimating the DBP, our proposed model all obtains the best score, which demonstrates that our model surpasses the other methods. Meanwhile, to check whether CNN could have better capability to extract data-driven features, we also build a complete model with CNN. However, its performance is close to that of CNN-only model.

Table 2: Estimation Results of the Different Methods

	Results for SBP			Results for DBP		
	MAPE	MAE	RMSE	MAPE	MAE	RMSE
SVM	5.109%	6.075	8.855	8.103%	5.427	8.080
MLP	3.617%	4.283	7.927	5.032%	3.508	6.074
CNN	3.684%	4.380	7.478	4.796%	3.315	5.945
Xing[23]	3.344%	4.053	7.175	4.872%	3.377	6.185
Schlesinger[24]	3.266%	3.979	7.035	4.784%	3.301	6.044
Baek[25]	3.248%	3.913	6.910	4.759%	3.298	5.917
proposed	3.243%	3.869	6.886	4.755%	3.292	6.059

Ablation Study To verify the contribution of each module in our proposed method, we conduct a clear ablation study. We remove the empirical feature extracting branch, the data-driven feature extracting branch, and the multichannel output mode in the gathering learning module with the same hyperparameters to form three new models: 1) model A utilizes the PPG sequence with only the LSTM branch; 2) model B uses artificial features only; 3) model C uses both empirical and data-driven features but does not use the multichannel output mode, which means there are two models for SBP and DBP estimation respectively. The results are shown in Table.3. In general, the performance of our hybrid-feature model is significantly better than other incomplete models, indicating that each module in our proposed framework contributes to the improvement of measuring accuracy and model robustness.

Table 3: Results of the Ablation Study

	Results for SBP			Results for DBP		
	MAPE	MAE	RMSE	MAPE	MAE	RMSE
A	3.897%	4.601	7.071	5.474%	3.635	6.240
B	3.492%	4.148	7.151	4.990%	3.512	6.161
C	3.573%	4.247	7.082	4.787%	3.363	6.256
complete	3.243%	3.869	6.886	4.755%	3.292	6.059

By comparing models A, B, and the complete model, we can conclude that both empirical features and data-driven features contain valuable information for BP estimation while the latter includes some implicit patterns neglected by the former. As presented in the table, the performance of model B is better than model A, which reveals that the 21 artificially constructed explicit features

express more hidden information than data-driven features. However, the complete model combining two branches of features obtains the best score. It not only proves that the LSTM can learn some implicit patterns neglected in artificial features but also verifies the effectiveness of our gathering module to blend and refine information.

The experimentation also verifies the effect of the multichannel output mode. To test the performance of model C, we have to train two task-oriented models for SBP and DBP, which is less efficient. As the results show, the complete model with multichannel output mode improves the performance for both SBP and DBP estimation, but the change of the former is much more significant compared to the latter. The reason behind this might be that some valuable hidden patterns are easier to learn from the DBP channel while the multichannel mode makes it possible for the SBP channel to eavesdrop on those shared patterns.

Parameter Study There are two the most significant hyperparameters in our proposed model, the input PPG sequence length (L) of the data-driven feature extraction branch based on LSTM, and the output feature size (H) of the two feature extraction branches. Considering both medical practicality, we set 4 groups of experiments to investigate the impact of L on the measuring results, L=10 seconds, L=20 seconds, L=30 seconds, and L=40 seconds, while H is set to 16, 32, 64, and 128. The other parameters remain the same as the proposed method and the experiment results are shown in Figure.4 and Figure.5 respectively.

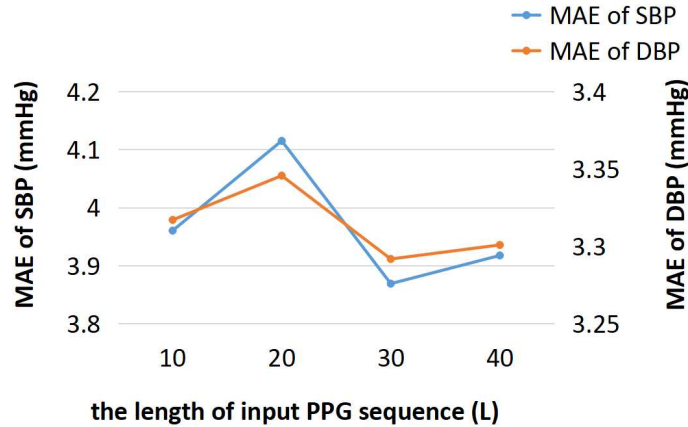


Fig. 4: Model performance with different input lengths of PPG (L)

L determines the input length of the data-driven pattern learning module, which represents the temporal range for LSTM to extract hidden patterns. Intuitively, the longer the signal we feed into the LSTM, the more information

it can learn. However, the results of experiments show that the performance of LSTM is neither linearly related to the input length nor does it simply improve before it decreases because of the underfitting or overfitting. The $L=30$ group has the best performance while the $L=10$ group performs better than the $L=20$ and $L=40$ groups. For this, we speculate that the PPG signals in the range of 10 to 20 seconds ahead or 30 to 40 seconds ahead contain less extra valuable information but some interference factors.

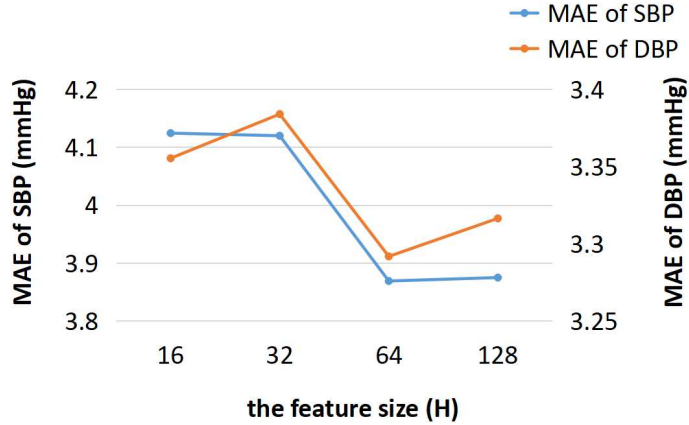


Fig. 5: Model performance with different feature sizes of the middle layer (H)

H is the hyper-parameter determining the capability of both the empirical pattern learning module and data-driven pattern learning module to represent a high level hidden pattern. When the pattern size is too small to pass all useful information to the deeper layers, an excessive size of patterns will not only lead to more parameters but also sparse the pattern and result in poor performance. Via experiments, we have found that the best value of H is 64.

5 Conclusions

In this paper, we proposed a cuffless BP estimation method using both artificial and data-driven features in PPG. The LSTM discovers some implicit features neglected by artificial features, while the empirical features utilize expert knowledge to make the end-to-end model more interpretable. A deep residual network is designed to reduce information redundancy in the gathered features and construct some higher-level features. Meanwhile, the multichannel output mode we applied to estimate SBP and DBP simultaneously can not only learn the task-oriented estimation function efficiently but also reinforce the shared features. We consider introducing some state-of-the-art empirical features and applying efficient noise filtering methods to the original PPG sequence in future works.

Acknowledgments We would like to thank Prof. Yingjie Zhou at Sichuan University who provided us with valuable suggestions and feedback that improved the quality of this manuscript.

References

1. W. H. Organization, *World health statistics 2015*. World Health Organization, 2015.
2. W. WHO, “A global brief on hypertension: silent killer, global public health crisis,” 2013.
3. Van de Vosse, F. N., Stergiopoulos, N.: Pulse wave propagation in the arterial tree. *Annual Review of Fluid Mechanics*, 43, 467-499. (2011)
4. Deng, Y., Zhou, Y., Zhang, Z.: Short-Long Correlation Based Graph Neural Networks for Residential Load Forecasting. In *International Conference on Neural Information Processing* (pp. 428-438). Springer, Cham. (2021)
5. Parreira, V. F., Vieira, D. S., Myrrha, M. A., Pessoa, I. M., Lage, S. M., Britto, R. R.: Optoelectronic plethysmography: a review of the literature. *Brazilian Journal of Physical Therapy*, 16, 439-453. (2012)
6. Chung, E., Chen, G., Alexander, B., Cannesson, M.: Non-invasive continuous blood pressure monitoring: a review of current applications. *Frontiers of medicine*, 7(1), 91-101. (2013)
7. Hong, Y., Zhou, Y., Li, Q., Xu, W., Zheng, X.: A deep learning method for short-term residential load forecasting in smart grid. *IEEE Access*, 8, 55785-55797. (2020)
8. H. Sorvoja and R. Myllyla, “Noninvasive blood pressure measurement methods,” *Molecular and quantum acoustics*, vol. 27, pp. 239-264, 2006.
9. A. Gaurav, M. Maheedhar, V. N. Tiwari, and R. Narayanan, “Cuff-less ppg based continuous blood pressure monitoring a smartphone based approach,” in *2016 38th annual international conference of the IEEE engineering in medicine and biology society (EMBC)*. IEEE, 2016, pp. 607-610.
10. D. Korpas, J. Halek, and L. Doležal, “Parameters describing the pulse wave.” *Physiological research*, vol. 58, no. 4, 2009.
11. 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*, vol. 64, no. 4, pp. 859-869, 2016.
12. Q. Xie, G. Wang, Z. Peng, and Y. Lian, “Machine learning methods for real-time blood pressure measurement based on photoplethysmography,” in *2018 IEEE 23rd International Conference on Digital Signal Processing (DSP)*. IEEE, 2018, pp. 1-5.
13. P. Su, X.-R. Ding, Y.-T. Zhang, J. Liu, F. Miao, and N. Zhao, “Long-term blood pressure prediction with deep recurrent neural networks,” in *2018 IEEE EMBS International conference on biomedical & health informatics (BHI)*. IEEE, 2018, pp. 323-328.
14. S. Shimazaki, H. Kawanaka, H. Ishikawa, K. Inoue, and K. Oguri, “Cuffless blood pressure estimation from only the waveform of photoplethysmography using cnn,” in *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. IEEE, 2019, pp. 5042-5045.
15. Elgendi, M.: On the analysis of fingertip photoplethysmogram signals. *Current cardiology reviews*, 8(1), 14-25. (2012)
16. Hochreiter, S., Schmidhuber, J.: Long short-term memory. *Neural computation*, 9(8), 1735-1780. (1997)

17. Ramchoun, H., Ghanou, Y., Ettaouil, M., Janati Idrissi, M. A.: Multilayer perceptron: Architecture optimization and training. (2016)
18. Zhao, R., Wang, J., Yan, R., Mao, K.: Machine health monitoring with LSTM networks. In 2016 10th international conference on sensing technology (ICST) (pp. 1-6). IEEE. (2016)
19. Johnson, A. E., Pollard, T. J., Shen, L., Lehman, L. W. H., Feng, M., Ghassemi, M., ... Mark, R. G.: MIMIC-III, a freely accessible critical care database. *Scientific data*, 3(1), 1-9. (2016)
20. Li, Z., Liu, F., Yang, W., Peng, S., Zhou, J.: A survey of convolutional neural networks: analysis, applications, and prospects. *IEEE Transactions on Neural Networks and Learning Systems*. (2021)
21. Sapankevych, N. I., Sankar, R.: Time series prediction using support vector machines: a survey. *IEEE computational intelligence magazine*, 4(2), 24-38. (2009)
22. Alaeddine, H., Jihene, M.: Deep Residual Network in Network. *Computational Intelligence and Neuroscience*, (2021).
23. Xing, X., Sun, M.: Optical blood pressure estimation with photoplethysmography and FFT-based neural networks. *Biomedical optics express*, 7(8), 3007-3020. (2016)
24. Schlesinger, O., Vigderhouse, N., Eytan, D., Moshe, Y.: Blood pressure estimation from PPG signals using convolutional neural networks and Siamese network. In ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) (pp. 1135-1139). IEEE. (2020)
25. Baek, S., Jang, J., Yoon, S.: End-to-end blood pressure prediction via fully convolutional networks. *IEEE Access*, 7, 185458-185468. (2019)