A Deep Learning Approach to Predict Blood Pressure from PPG Signals

Ali Tazarv*, Marco Levorato†

⋆ EECS Department, † Dept. of Computer Science

University of California, Irvine

Abstract— **Blood Pressure (BP) is one of the four primary vital signs indicating the status of the body's vital (life-sustaining) functions. BP is difficult to continuously monitor using a sphygmomanometer (i.e. a blood pressure cuff), especially in everyday-setting. However, other health signals which can be easily and continuously acquired, such as photoplethysmography (PPG), show some similarities with the Aortic Pressure waveform. Based on these similarities, in recent years several methods were proposed to predict BP from the PPG signal. Building on these results, we propose an advanced personalized data-driven approach that uses a three-layer deep neural network to estimate BP based on PPG signals. Different from previous work, the proposed model analyzes the PPG signal in time-domain and automatically extracts the most critical features for this specific application, then uses a variation of recurrent neural networks (RNN) called Long-Short-Term-Memory (LSTM) to map the extracted features to the BP value associated with that time window. Experimental results on two separate standard hospital datasets, yielded absolute errors mean and absolute error standard deviation for systolic and diastolic BP values outperforming prior works.**

I. INTRODUCTION

Hypertension, defined as systolic blood pressure (SBP) larger than 140mmHg or diastolic blood pressure (DBP) larger than 90mmHg[1] [1], is estimated to have caused 9.4 million annual deaths globally, 17% of the total death in 2012 and 7% of total disability-adjusted life years (DALYs) [2]. If left uncontrolled, hypertension causes stroke, myocardial infarction (MI), cardiac failure, dementia, renal failure, and even blindness. In adults, hypertension after diabetes is the second reason to increase the risk of cardiovascular disease (CVD) and several types of cancer, as well as multiple non-fatal diseases. Hypertension has been increasing in recent years. By 2030, 40.5% of the US population is projected to have some form of CVD [3]. While people with the risk of hypertension need to measure their blood pressure frequently, conventional cuff-based BP measurement devices are expensive and inconvenient for continuous monitoring. Thus the development of alternative methods is necessary.

Blood Pressure (BP) – commonly measured in mmHg – is a quasi-periodic signal in sync with an individual's heartbeats. The upper peak in each period is called the Systolic Blood Pressure or SBP, and the lower bound in each period is called Diastolic Blood Pressure or DBP (Figure 1). While blood pressure is difficult to monitor continuously in a non-clinical setting, Photoplethysmography (PPG) is a non-invasive optical method that measures a related signal: blood volume temporal variations in the vessels and tissues. PPG signals are obtained from pulse oximeters, emitting visible light (LED) on the skin and measuring the micro-variations in the transmitted, or reflected light intensity (photo-diode). PPG sensors are small in size and low cost to build, and they already exist in most newer wearables (e.g. smartwatches, activity trackers, and smart rings).

In recent years, there has been an extensive body of research studying similarities and correlation of PPG signals and Aortic Pressure waveforms, as well as the possibility of estimating SBP and DBP based on PPG signals. Since both signals are originating from the same source (the individual's heartbeats), they are highly correlated. However, since Aortic and PPG are generally measured from different parts of the body (e.g. arm and wrist) using different devices, they are typically out of phase. Figure 1 is based on experimental data after time-shift alignment (since PPG signal does not have a unit, in the figure it is scaled for easier readability). Some methods such as the one in [4] are proposed to automatically detect and compensate for this phase difference.

Due to convenience of use and also low cost of PPG sensors, estimating blood pressure metrics (SBP and DBP) from PPG signals is of great interest. The common approach for this is to first extract a set of predefined features from PPG, and then use some regression models to estimate BP from those features. A shortcoming of this approach is that since these features are predefined and generic, some of the important information inside the raw PPG signals (including patterns specific to each subject) might be lost, which if captured, could improve the accuracy of BP estimation. Other than that, these methods do not utilize characterizations of PPG as a time-series signal. In this paper, we propose the use of a new framework to address these shortcomings.

Our proposed approach makes the following advancements with respect to prior works in the literature: (*i*) The proposed method utilizes a Convolutional Neural Network (CNN) layer, which "learns" the most informative features of the
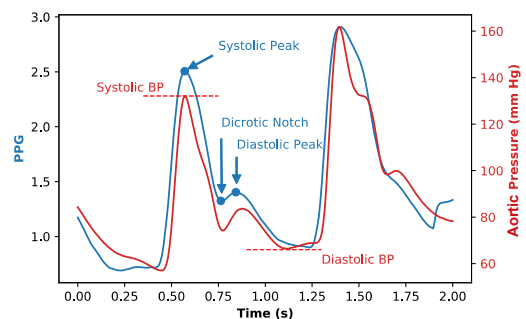
[1]for a patient who is not undergoing drug treatment for hypertension.

Fig. 1: Aortic Pressure and PPG raw signals

PPG signal for this application (BP estimation) in a nonlinear and efficient way, as part of an optimization problem; (*ii*) We adopt a Long-Short Term Memory (LSTM) model to capture long-term temporal inter-dependencies in a seamless and automated procedure.

We train and test the model for each subject separate from others. We test the proposed framework on two standard hospital datasets. Our proposed model on 20 randomly picked subjects from MIMIC-II dataset [5] gives prediction error ($MAE \pm SDAE$) of 3.70±3.07 mmHg and 2.02±1.76 mmHg, for SBP and DBP values respectively. Also on UQVSD dataset, prediction errors ($MAE \pm SDAE$) are 3.70±3.07 mmHg and 2.02±1.76 mmHg, for SBP and DBP respectively. To the best of our knowledge, these results on similar datasets outperform other methods in the literature.

The rest of the paper is organized as follows. Section II overviews recent work in this area. Section III describes our proposed framework and methodology. Section IV introduces the two real-world datasets we used. In Section V we present the results and compare those with state-of-the-art methods. Section VI concludes the paper.

II. RELATED WORK

Estimation of Blood Pressure from PPG signal is studied in prior works and several methods are proposed for that. In [6], [7] and [8], the authors propose methods based on Pulse Transit Time (PTT). PTT is defined as the time required for the blood pressure wave to travel from the source (heart) to the wrist (where the PPG signal is recorded, and it is measured from the time shift between the Electrocardiography (ECG) and PPG signals. In [9], authors consider every sample point in one cycle as a separate feature and use Principal Component Analysis (PCA) to reduce the number of features, and then apply regression algorithms on the set of reduced features. In [10], authors propose a list of some potential features of the PPG signals (e.g. area under the curve, time length of certain points on the signal in one cycle, etc.), and then use an autoencoder to reduce the list of features and at the end estimate BP based on the reduced set of features using a feed-forward neural network.

All these methods start by extracting a set of features from PPG signals and then use those features to estimate blood pressure. One shortcoming of this approach is that since these features are predefined, they may fail to capture all the details in the signal that might be useful for a specific task such as BP estimation. Some of these details might also be different from one subject to another. Even in [9] that the authors start with the whole signal, the dimensionality reduction procedure is not optimized for the BP estimation task, as PCA is a generic data compression method. Additionally, none of these approaches exploit possible characterizations of the PPG as a time-series signal.

III. METHODOLOGY

Different from the common approach described in Section II, our proposed method for each subject automatically extracts a set of features that are optimal for this application,
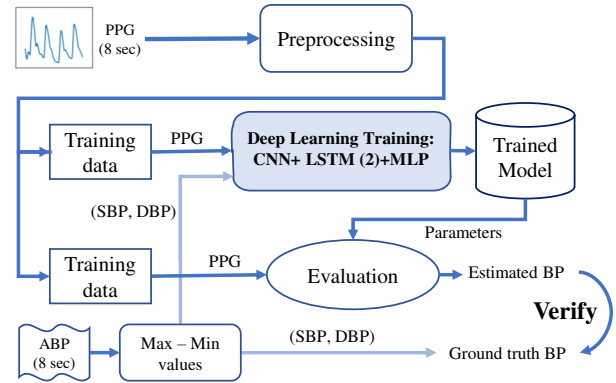

Fig. 2: Overview of the proposed method to estimate BP from single channel PPG signal

as part of the training process. It also captures and utilizes temporal dependencies in the data using an advanced neural architecture. Figure 2 presents the block diagram of the method we propose to estimate SBP and DBP using PPG signal. The method is composed of the following blocks and components: (*A*) Signal Pre-processing; (*B*) Machine Learning, composed of an automatic feature extraction module (CNN) and a time series analysis module (LSTM); and (*C*) Model Evaluation block. We explain the first two components here and describe the Evaluation block later in Section V.

A. Pre-processing

Since both Aortic Pressure and PPG originate from heartbeats, we do not expect to see high-frequency elements in their waveforms. Thus, higher frequency elements are likely due to noise. Additionally, since the PPG is a relative signal and its DC offset (mean value) does not have an interpretation of our interest, we eliminate the mean value of the signal (the zero frequency element). There are many popular methods to eliminate a range of frequencies from a signal. Herein, we use a traditional Fourier Transform (FFT)-based approach. In the PPG signal, we set a band-pass filter eliminating the frequencies outside the range $0.1 - 8$ Hz. The resulting waveform still contains all the information necessary for our estimation, while being affected by smaller noise energy. In preprocessing the Aortic Pressure signal, we apply a low-pass filter with a cut-off frequency of 5Hz to the raw signal to eliminate sharp peaks in the signal. After these steps, we split both the PPG and ABP into windows of 8 s with a step of 2s, resulting in 6s overlaps between two consecutive windows. We further scale the PPG signals in each window to zero mean and unit variance. Each window of the PPG signal is equivalent to a vector of length $8f_s$, which f_s is the sampling frequency. We use this vector as the model input. In the ABP signal, the maximum and minimum values in each window are interpreted as the SBP and DBP respectively for that interval (Figure 1).

B. Deep Learning

Deep Neural Networks primarily including Convolutional Neural Network (CNN), Recurrent Neural Network, and

Multi-Layer-Perceptron (MLP), facilitate task-adapted feature representation of the data. CNN, which was originally designed to capture the variability of 2D data, in [11] is shown to perform well on 1D data and outperform other approaches. CNN consists of an initial layer of convolutional filters – boxes with certain sizes sF, and certain sets of trainable weights which slide over the input data – followed by an activation function (e.g. sigmoid), a pooling layer with pool size sP, and after that a batch normalization layer [12].

RNN is an effective tool for the analysis of time-series data since it learns contextual patterns in the input from previous time steps. Unlike feed-forward networks, RNN uses some internal states to process temporal sequences of the input. However, learning to store information over extended time intervals via recurrent back-propagation may necessitate an excessive training time, due to the insufficient decaying error gradient back-flow (vanishing gradient) discussed in [13].

Long-Short-Term-Memory (LSTM) is a variation of RNN and solves this issue by using memory blocks, for which the trainable "forget gate", "input gate" and "output gate" control which parts of the data are worth saving and which parts are not. This idea solves the issue of vanishing gradient mentioned above, and the result is a strong tool to learn and track long-term inter-dependencies in the input data (e.g. time series data) [13]. The combination of CNN with a subsequent LSTM has been shown to perform well on PPG signals. In [14] authors use a similar architecture for estimating heart-rate and Bio-metric ID using PPG.

The model is trained in Python environment (version 3.6.3) and implemented using Keras 2.2.4 [15], with Tensorflow 1.3.1 backend. The CNN layer starts with a 1-D filtering operation with filter size 15 (sF), followed by a Rectified Linear Unit (RELU) [16] as activation function, a batch-normalization layer [12], a max-pooling layer, and finally a dropout layer [17] with $dropout\_rate = 0.1$. In the max-pooling layer the pooling size sP is set to 4, which means the layer takes the maximum out of every 4 consecutive values in a filter (non-overlapping) and discards the rest. Unlike some other common deep network layers (e.g. MLP), the Max-pooling layer does not have trainable weights. Dropout layers are probabilistic masks that in every gradient update (mini-batch) block a portion of the nodes. In our implementation, the $dropout\_rate = 0.1$ means that in every gradient update, there is a 10% chance for each node in the layer to be dropped out (or masked). These two-layers (Max-pooling and Dropout) are well-known methods used to avoid overfitting.

Our LSTM network is composed of two identical LSTM module in series. Each of them has 64 units (nU), with *tanh* as the activation function for the hidden state data and output data, and *hard-sigmoid* as the recurrent activation functions for the forget, input, output gates [13].

The model was trained using Adam optimizer [18]. The batch size was set to 20 to balance the training time vs noisiness of gradient updates trade-off. The hyperparameters mentioned above and also the number of layers for each network (CNN and LSTM), were optimized with grid search. Changing the hyperparameters (e.g. sF, nU, sP, etc.) or the
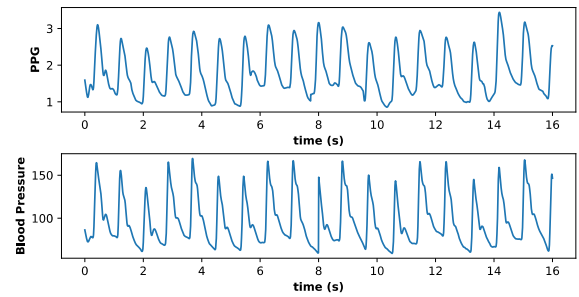

Fig. 3: Scaled raw PPG and ABP signals in the same time interval, from MIMIC-II dataset.

number of layers did not result in an improved final output.

IV. DATASET

To evaluate the accuracy and efficiency of the proposed method, we used two publicly accessible real-life datasets.

MIMIC V.3 2015: Multi-parameter Intelligent Monitoring in Intensive Care, provided by [5]. The MIMIC-II dataset includes healthcare information and signals of thousands of patients at hospitals between the years 2001 and 2012. An example of PPG and corresponding ABP signals from this dataset is given in Figure 3 (the PPG signal is scaled for readability). These signals were originally recorded at a sampling frequency f_s=125 Hz, with a minimum accuracy of 8 bits. We randomly selected data corresponding to 20 different subjects -five minutes long each- from this dataset to test our model.

The University of Queensland Vital Signs Dataset: This dataset covers a wide range of BP values, recorded from 32 surgical patients in duration ranging from 13 minutes to 5 hours, at the Royal Adelaide Hospital, Australia [19]. This dataset is recorded at a sampling frequency $f_s = 100Hz$. We used 49 measurements (10 minutes each) from this dataset.

The majority of studies on this topic have used parts of one of these two datasets. However, details of how portions of these datasets were selected for experiments are not revealed. We acknowledge that the data we used for our experiment, might not exactly match the data used in similar studies, but in order to conduct a fair comparison, we picked a portion of each dataset randomly.

For both datasets, we re-sampled the PPG and ABP signals at 20 Hz and re-adjusted the timings [4]. The distribution of BP in MIMIC-II data (the portion we used in this experiment) is presented in Figure 4. After these steps, the resulting PPG signals and ABP values are ready to be used for training and evaluation of our method.

V. EXPERIMENTAL RESULTS

The preprocessed PPG signals and BP values from both datasets (Section IV) are further normalized to zero mean and unit variance before being used as the input to the network. The most important nonlinear features of the input data are extracted through a CNN network, and then these features are given to a two-layer LSTM network to do the time series analysis. It is important to note that unlike conventional methods, feature extraction is performed as part

TABLE I: Performance comparison with prior works

Work	Dataset	SBP		DBP	
		MAE	SD	MAE	SD
[20]	MIMIC-II	4.47	6.85	3.21	4.72
[21]	MIMIC-II	8.54	-	4.34	-
[9]	MIMIC-II	3.97	7.99	2.43	3.37
[22]	MIMIC-II	3.80	3.46	2.21	2.09
Our work	**MIMIC-II**	**3.70**	**3.07**	**2.02**	**1.76**
[23]	UQVSD	11.64	8.20	7.62	6.78
[24]	UQVSD	4.77	7.68	3.67	5.69
Our work	**UQVSD**	**3.91**	**4.78**	**1.99**	**2.45**

Fig. 5: Prediction Error for SBP (left), and DBP (right)

of the training process; the trainable weights in the CNN are optimized to extract the features that are most informative toward the estimation of the BP values. We perform *leave-one-window-out* validations. Hence, one 8s time window is kept separate, and training is done on the rest of the data. Then, the trained model is evaluated on the test sample. Also for training, we do not use three time-windows on each side of the test set, to account for the 6s overlap between two consecutive time windows. With this strategy, the training data and test data will be completely separate from one another. We are training and testing the model for each subject independent from the others. The performance of the network is evaluated using the following three metrics:

Absolute Error (AE): for each test sample, AE is defined as $AE_i = |BP_{T_i} - BP_{E_i}|$, in which BP_{T_i} is the true or target blood pressure, and BP_{E_i} is the estimated blood pressure when the window number i is held out as the test sample. Consequently the mean value of AE_i and the standard deviation over the `number_of_samples` iterations, and over the set of subjects are calculated and reported in the Table I. The mean and standard-deviation of AE would be a rough estimate of the error value on predictions, and the robustness and reliability of the method respectively. A model is selected as a strong predictor if it produces small MAE and $SDAE$ values. The method proposed in this paper, to the best of our knowledge, outperforms other frameworks and techniques that are available until this time.

BHS: As a second metric, we consider a standard commonly used for blood pressure measurements accuracy, provided by the British Hypertension Society (BHS) [25], which grades the measurement accuracy into three groups (A-C). This metric measures the fraction of measurements (or estimations) which are within a certain range – 5, 10
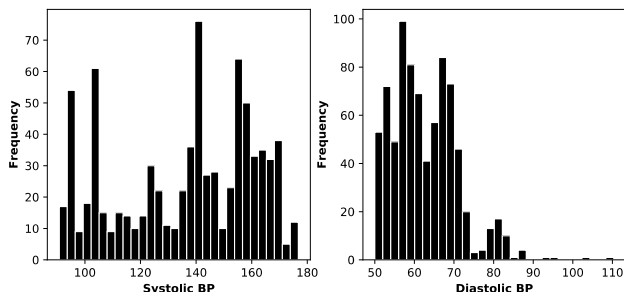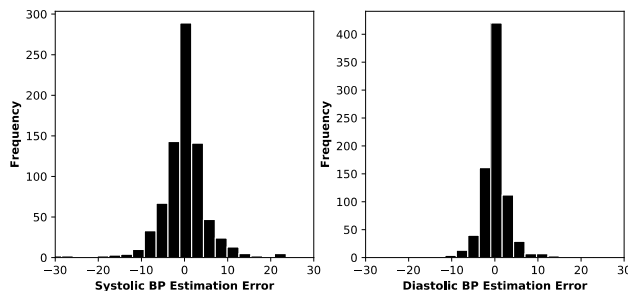
and 15 mmHg – of the target values for unobserved data. Based on this metric, our proposed model (to the best of our knowledge) is the only method that gets Grade *A* for both SBP and DBP (in predictions based on PPG only). Detailed results are given in Table II. For comparison, authors in [9] report grade *A* for DBP and below *C* for SBP estimation on MIMIC-II dataset. Also authors in [24] report grade *A* for DBP, and *B* for SBP estimation on UQVSD dataset.

AAMI: As the third metric, we consider a different standard provided by the US Association for the Advancement of Medical Instrumentation (AAMI) [26]. Based on this Standard, a measurement algorithm is valid if the ME (Mean Error) of the measurements is below 5 mmHg and the SD of Errors is smaller than 8 mmHg. The results are presented in Table III. Our model passes the AAMI standard criteria for both SBP and DBP on both dataset. For comparison, our proposed model produces better error values (lower mean and σ) compared to [24] on UQVSD data. On MIMIC-II dataset we get lower error σ but higher error mean compared to [9].

The distribution of the prediction error for one subject from the MIMIC-II dataset is depicted in Figure 5. Based on the histograms, prediction errors are mostly small. However, in some rare cases (less than 1% of times) the estimation error for SBP goes beyond the 20 mmHg range from the target value. For DBP, that never happened in our experiments.

The Bland-Altman plot for the predictions on the same

TABLE II: BHS Standard [25] vs. our results

(a) BHS Standard Minimum Requirements

	Cumulative Percentage Error		
	≤ 5 mmHg	≤ 10 mmHg	≤ 15 mmHg
Grade A	60%	85%	95%
Grade B	50%	75%	90%
Grade C	40%	65%	85%

(b) Our results on the two datasets

Dataset	Signal	Cumulative Percentage Error (Grade)		
		≤ 5 mmHg	≤ 10 mmHg	≤ 15 mmHg
MIMIC-II	SBP	77% (**A**)	92% (**A**)	96% (**A**)
	DBP	93% (**A**)	97% (**A**)	99% (**A**)
UQVSD	SBP	75% (**A**)	92% (**A**)	96% (**A**)
	DBP	92% (**A**)	98% (**A**)	99% (**A**)

Fig. 4: Distribution of Systolic BP and Diastolic BP

TABLE III: Comparison of results with AAMI Standard [26]

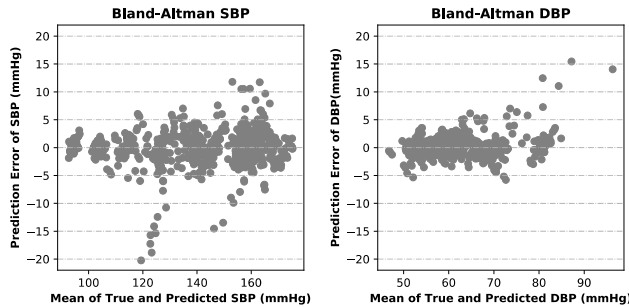|		Mean Error	σ of Error
AAMI Criteria:		≤ 5	≤ 8
MIMIC-II	SBP	0.21	6.27
	DBP	0.24	3.40
UQVSD	SBP	0.52	6.16
	DBP	0.20	3.15

Fig. 6: Bland-Altman plot for SBP and DBP

subject is presented in Figure 6. Predictions on lower values of BP – left side of each plot - are more reliable (smaller error). Note that these samples correspond to when the subject is likely to be less active (potentially at rest), at which the PPG signal is expected to be less noisy and more reliable.

VI. CONCLUSIONS AND FUTURE WORK

Hypertension or high blood pressure is closely related to cardiovascular disease. While continuous and portable BP monitoring is of high importance, conventional cuff-based BP measurement devices are primarily confined to clinical settings. The method proposed in this paper is a cuff-less, noninvasive and feasible method which estimates Blood Pressure with high accuracy based on a single channel PPG signal. PPG sensors are convenient to use, simple and low cost, and are integrated in most of newer wearable devices. We use Convolutional Neural Network to extract the optimal set of features from the input, and then use a LSTM network to capture the temporal correlations in the extracted features. Results of the proposed method outperform previous works that use a single channel PPG signal as input. Importantly, the proposed method ranks as *A* in the BHS Standard for both SBP and DBP, complies with the AAMI Standard, and outperforms similar works in terms of prediction error (as well as BHS criteria).

Importantly, the proposed method is feasible for practical applications in commercial products. The processing power required to execute the model (and more importantly, to keep it updated with the new incoming data) might be too intensive to be executed on the edge layer in an IoT architecture. However, the raw PPG data can be transferred to the cloud, and the prediction output can be sent back to the user through a smart-watch or smart-phone, all in a real-time scenario. The development of such architecture is left to future work.

REFERENCES

[1] E. J. Benjamin *et al.*, "Heart disease and stroke statistics-2018 update: A report from the american heart association," *Circulation*, 2018.

[2] W. H. Organization *et al.*, "Global status report on noncommunicable diseases 2014," tech. rep., World Health Organization, 2014.

[3] P. A. Heidenreich *et al.*, "Forecasting the future of cardiovascular disease in the united states," *Circulation*, vol. 123, no. 8, 2011.

[4] X. Xing and M. Sun, "Optical blood pressure estimation with photoplethysmography and fft-based neural networks," *Biomedical optics express*, vol. 7, no. 8, pp. 3007–3020, 2016.

[5] A. L. Goldberger *et al.*, "Physiobank, physiotoolkit, and physionet," *Circulation*, vol. 101, no. 23, pp. e215–e220, 2000.

[6] M. Kachuee *et al.*, "Cuff-less high-accuracy calibration-free blood pressure estimation using pulse transit time," in *2015 IEEE International Symposium on Circuits and Systems (ISCAS)*, May 2015.

[7] J. Lass *et al.*, "Continuous blood pressure monitoring during exercise using pulse wave transit time measurement," in *The 26th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, vol. 1, pp. 2239–2242, IEEE, 2004.

[8] F. Miao *et al.*, "A novel continuous blood pressure estimation approach based on data mining techniques," *IEEE Journal of Biomedical and Health Informatics*, vol. PP, pp. 1–1, 04 2017.

[9] S. S. Mousavi *et al.*, "Blood pressure estimation from appropriate and inappropriate ppg signals using a whole-based method," *Biomedical Signal Processing and Control*, vol. 47, pp. 196 – 206, 2019.

[10] S. Shimazaki *et al.*, "Features extraction for cuffless blood pressure estimation by autoencoder from photoplethysmography," in *2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pp. 2857–2860, IEEE, 2018.

[11] Y. Lecun *et al.*, "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, vol. 86, pp. 2278–2324, Nov 1998.

[12] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," *CoRR*, 2015.

[13] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997.

[14] D. Biswas *et al.*, "Cornet: Deep learning framework for ppg-based heart rate estimation and biometric identification in ambulant environment," *IEEE Transactions on Biomedical Circuits and Systems*, vol. 13, pp. 282–291, April 2019.

[15] F. Chollet *et al.*, "Keras." https://keras.io, 2015.

[16] V. Nair and G. E. Hinton, "Rectified linear units improve restricted boltzmann machines," ICML'10, (USA), Omnipress, 2010.

[17] N. Srivastava *et al.*, "Dropout: A simple way to prevent neural networks from overfitting," *JMLR*, 2014.

[18] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *CoRR*, vol. abs/1412.6980, 2014.

[19] D. Liu *et al.*, "University of queensland vital signs dataset: Development of an accessible repository of anesthesia patient monitoring data for research," *Anesthesia & Analgesia*, vol. 114, no. 3, 2012.

[20] A. Gaurav *et al.*, "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)*, pp. 607–610, Aug 2016.

[21] M. Liu, L. Po, and H. Fu, "Cuffless blood pressure estimation based on photoplethysmography signal and its second derivative," *International Journal of Computer Theory and Engineering*, vol. 9, 01 2017.

[22] Y. Kurylyak *et al.*, "A neural network-based method for continuous blood pressure estimation from a ppg...," in *IEEE International Instrumentation and Measurement Technology Conference (I2MTC)*, 2013.

[23] Y. Zhang and Z. Feng, "A svm method for continuous blood pressure estimation from a ppg signal," ICMLC 2017, ACM, 2017.

[24] K. Duan *et al.*, "A feature exploration methodology for learning based cuffless blood pressure measurement...," EMBC 2016.

[25] E. O'Brien *et al.*, "Blood pressure measuring devices: recommendations of the european society of hypertension," *BMJ*, 2001.

[26] A. for the Advancement of Medical Instrumentation *et al.*, "American national standard. manual, electronic or automated sphygmomanometers," *ANSI/AAMI SP10-2002/A1*, 2003.