Supplementary material

Shan He[1] and Miodrag Bolić[1]

[1]School of Electrical Engineering and Computer Science, University of Ottawa

May 29, 2024

S1 MLP model for example shown in Figure 1(a)

A simple multilayer perceptron (MLP) model (hidden layer 1-4: 32-128-64-32 nodes, ReLU activation; output layer: 2 nodes, linear activation) was implemented to estimate SBP and DBP using 20 features extracted from ECG and PPG waveforms. A total of 147 subsets were selected from UCI cuffless BP estimation dataset[5], 117 subsets (∼80%) were used for training and 30 subsets (∼20%) were used for testing (no overlap between training and testing set). The mean and standard deviation of absolute difference (MAD±SDAD) of the testing set is 9.00 ± 7.31 and 7.05 ± 6.60 mmHg for SBP and DBP, respectively. The MAD±SDAD of a 543 seconds' testing subset shown in Figure 1(a) is 3.77 ± 2.51 mmHg for SBP. However, the model is unable to track one acute BP change occurred between 325 and 375 seconds.

S2 BP changes caused by activities

Figure S1 demonstrates several procedures used by the wearable physiological and blood pressure measurements during activities of daily living dataset[1] (from now on referred as **DailA BP**) for short-term BP changes. Figure S1(a) illustrates the BP waveform of a subject performing multiple hand grippings for one minute and then recovered. the BP decreases when the force applied to the hand grip increases, and there is a rapid BP drop when hand griping is stopped (e.g., a 32 mmHg decrease of SBP from 60s to 65s) and then BP gradually recovers. Figure S1(b) illustrates the BP steadily increases when a participant walks for one minute (SBP increases from 128 mmHg to 152 mmHg), study[2] also reveals that SBP increases significantly and proportionally to workload increase during exercise test in healthy adults. Figure S1(c) illustrates BP changes responding to Valsalva Maneuver (VM), VM can be divided into four phases based on hemodynamic changes. The BP will transiently increase in Phase 1 and then steadily decrease in Phase 2. In Phase 3, BP will further decreases and increases and overshoot in Phase 4, and gradually recover to normal. Comparing to exercises which steadily change BP, VM can significantly change BP within a very short time, the mean and standard deviation of SBP changes ($\max(\mathrm{SBP}_{\mathrm{VM}}) - \min(\mathrm{SBP}_{\mathrm{VM}})$, e.g., 174-85=89mmHg in Figure S1(c)) is 115 ± 21 mmHg concluded from 5 subjects' 19 VM segments in **DailA BP** dataset. Figure S1(d) illustrates BP decreases when a subject changes posture from sitting to standing and then recovers to normal.

S3 Pearson correlation coefficient (PCC) and Concordance correlation coefficient (CCC)

The relation between CCC (ρ_c) and PCC (ρ) can be represented as:

$$\rho_c = C_{\mathrm{b}} \cdot \rho \tag{S1}$$

where $C_{\mathrm{b}} = \frac{2}{v+\frac{1}{v}+u^2}$, $v = \frac{\sigma_1}{\sigma_2}$ is a scale shift factor and $u = \frac{\mu_1-\mu_2}{\sqrt{\sigma_1\sigma_2}}$ is a location shift factor relative to the scale. σ_{12} is the covariance of two variables, $\mathrm{BP}_{\mathrm{est}}$ and $\mathrm{BP}_{\mathrm{ref}}$. σ_1 and σ_2 are the standard deviations of $\mathrm{BP}_{\mathrm{est}}$ and $\mathrm{BP}_{\mathrm{ref}}$, respectively. μ_1 and μ_2 are the mean values of $\mathrm{BP}_{\mathrm{est}}$ and $\mathrm{BP}_{\mathrm{ref}}$, respectively.

An experiment was implement to demonstrate that CCC is more robust in detecting scale and location shifts than PCC. BP_{ref} is a reference measure of systolic blood pressure. The BP_{est} in Figure (a) to (d) to represent (1) location shift ($\mathrm{BP}_{\mathrm{est}} = \mathrm{BP}_{\mathrm{ref}} - 2.5$), (b) scale shift ($\mathrm{BP}_{\mathrm{est}} = 1.05 * \mathrm{BP}_{\mathrm{ref}}$), (c) location and scale shift ($\mathrm{BP}_{\mathrm{est}} = 1.05 * \mathrm{BP}_{\mathrm{ref}} - 5$), and (d) no concordance ($\mathrm{BP}_{\mathrm{est}} = \mathrm{mean}(\mathrm{BP}_{\mathrm{ref}})$), as shown in Figure S2.

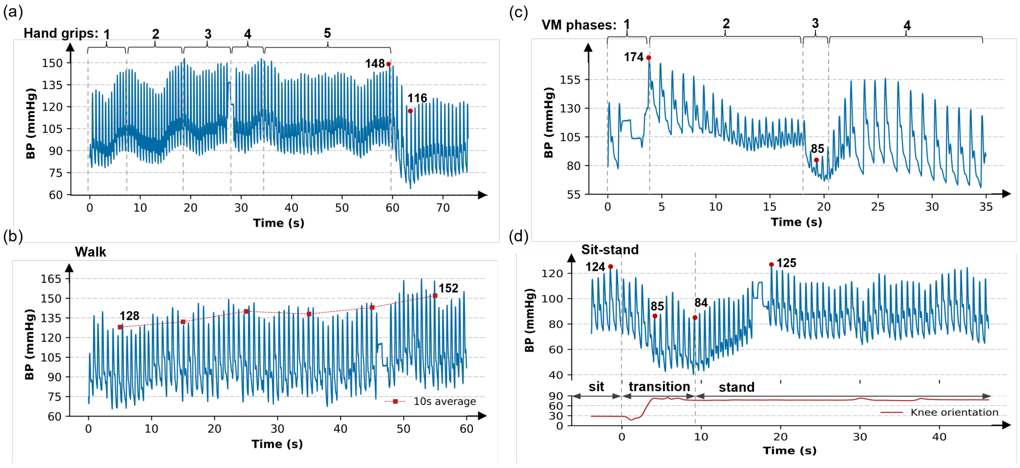

Figure S1: (a) BP waveform when hand grips were performed by squeezing a rolled towel at maximum voluntary contraction (0-60s) and at rest (after 60 s), (b) BP waveform when the subject was walking for 1 minute, (c) BP waveform when Valsalva maneuver was conducted, the BP reached maximal during Phase 1, (d) an example of BP changes extracted from **DailA BP** when a subject changed posture from sitting to standing, the lower track is the orientation of subject's knee (unit: degrees), 0° indicates thighs are parallel to ground (sitting) 90° indicates thighs are perpendicular to ground (standing).

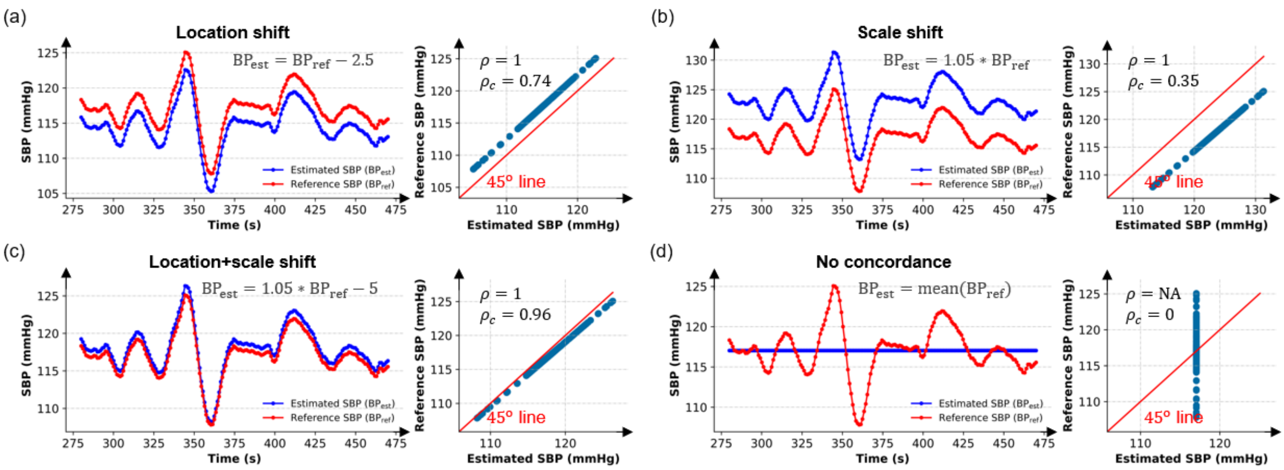

Figure S2: Demonstration of concordance correlation coefficient (ρ_c) and Pearson correlation coefficient (ρ) for different scenarios, (a) location shift, (b) scale shift, (c) location and scale shift, (d) no concordance.

S4 Algorithm: segment distance, trend and compisite similarity metrics and temporally normalized metrics

The pseudocode of the proposed segment distance, trend and composite similarity metrics, as well as their temporally normalized metrics is demonstrated in Algorithm S1.

S5 Re-implementation of SOTA models

Two SOTA cuffless BP estimation models were selected and re-implemented using **DailA BP**[1]. Their results were used to demonstrate and compare the performance of conventional metrics and proposed metrics for BP pattern tracking.

Different training strategies were applied for ApproximateNet and DeepRNN-4L models with different training strategies and due to the fact that The model diagrams and parameters are demonstrated in Figure S3 and experiment details are described below:

- **Data preparation**: For each subject, an average of 2.3 hours of continuous ECG, PPG and BP waveforms were extracted from DailA BP. The sampling rate is 64 Hz and 2^{nd} order Butterworth bandpass filters with pass band of 0.5-30 Hz and 0.5-10 Hz were applied to ECG and PPG signals for noise removal respectively.

Algorithm S1 Segment distance, trend and composite similarity metrics and temporally normalized metrics

Input: two sequences for comparison x and y, $\text{length}(x) = \text{length}(y)$
Output: $\text{TNCSM}(x,y) \in [-1, 1]$

changepoint_x = segmentation(x) ▷ segment x and extract change points, according to tendency
changepoint_y = segmentation(y) ▷ segment y and extract change points, according to tendency
changepoint_xy = sort([0, unique(changepoint_x ∪ changepoint_y), length(x) − 1]) ▷ ESN process
changepoint_process(changepoint_xy) ▷ process changepoints such that each segment has enough samples for PLR
$T_D = \text{length}(x)$ ▷ duration of the sequence
TN_DS = 0, TN_TS = 0, TN_CS = 0
for $i \leftarrow 1$ to length(changepoint_xy) − 1 **do**
  t_i = changepoint_$xy(i) + 1$ ▷ for the first segment, t_i = changepoint_$xy(i)$
  t_{i+1} = changepoint_$xy(i + 1)$
  $x_i = x(t_i : t_{i+1})$
  $y_i = y(t_i : t_{i+1})$
  x_{L_i} = least_squares(x_i)
  y_{L_i} = least_squares(y_i) ▷ approximate a straight line for each segment using least squares method
  $\text{DS}(x_i, y_i)$ = distance_similarity(x_i, y_i) ▷ evaluate segment distance similarity
  $\text{TS}(x_i, y_i)$ = trend_similarity(x_{L_i}, y_{L_i}) ▷ evaluate segment trend similarity
  $\text{CS}(x_i, y_i) = w_1 \cdot \text{DS}(x_i, y_i) + w_2 \cdot \text{TS}(x_i, y_i)$ ▷ calculate composite similarity using weighted sum
  $t_{d_i} = t_{i+1} - t_i$ ▷ duration of i-th segment
  TN_DS += $\frac{1}{T_D} \cdot \text{DS}(x_i, y_i) \cdot t_{d_i}$ ▷ compute sequence distance similarity with temporal normalization
  TN_TS += $\frac{1}{T_D} \cdot \text{TS}(x_i, y_i) \cdot t_{d_i}$ ▷ compute sequence trend similarity with temporal normalization
  TN_CS += $\frac{1}{T_D} \cdot CS(x_i, y_i) \cdot t_{d_i}$ ▷ compute sequence composite similarity with temporal normalization
end for

- **ApproximateNet**: each subject's continuous PPG and BP waveforms were divided into a number of 16 seconds' (fs=64 Hz, $16 \times 64 = 1024$ samples, same to the segment length of the original work[4]) segments. Five-fold cross validation was applied for model training (5 folds as DaliA BP includes 5 subjects). For each iteration, the model was trained on four subjects and estimate BP for the left out subject. Then, the first three segments of reference BP waveform were used to calibrate the estimated BP waveforms, $\text{BP}_{est} = \text{BP}_{est} - (\text{mean}(\text{BP}_{est_{1-3}}) - \text{mean}(\text{BP}_{ref_{1-3}}))$. The SBP and DBP values were then detected from the calibrated BP waveforms and then resampled to uniformly sampled time series sequences (fs=1 Hz). The MAD±SDAD is 9.88±8.65 and 6.50±5.02 mmHg for SBP and DBP respectively.

- **DeepRNN**: as the data of different subjects are independent, a personalized model was trained for each subject. Seven features (demonstated in Section S6) were extracted from the first 25 minutes of ECG and PPG waveforms, as well as corresponding SBP and DBP values were used for model training, then predict SBP and DBP with the rest data. Then the delay between estimation and reference which is caused by the lagged training set was compensated, and the estimated SBP and DBP values were resampled to uniformly sampled time series sequences (fs=1 Hz) for performance evaluation. The MAD±SDAD is 4.14±3.29 and 4.15±3.30 mmHg for SBP and DBP respectively.

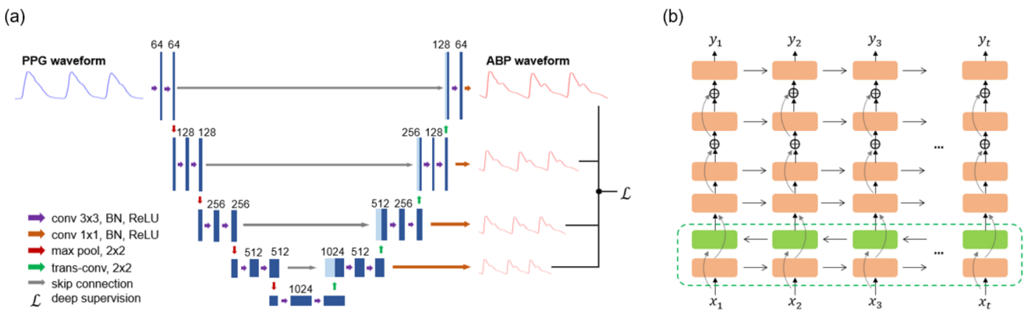

Figure S3: (a) Diagram of ApproximateNet[4], which is a deeply supervised one-dimensional U-Net to approximate ABP waveforms from preprocessed PPG waveforms; (b) Diagram of DeepRNN-4L[3], each rectangle box is an LSTM cell (unit = 128), and the sequence length of each training sample is set to 32. The bottom green dashed block is a bidirectional LSTM layer (unit = 128); there is a dropout layer (rate = 0.2) applied after each bidirectional and forward LSTM layer; residual connections are also applied between LSTM layers. The last layer (unit = 2, activation=linear) was added for SBP and DBP prediction.

S6 Features used for DeepRNN-4L

Seven features extracted from simultaneous ECG and PPG waveforms for BP estimation using DeepRNN-4L model is demonstrated in Figure S4

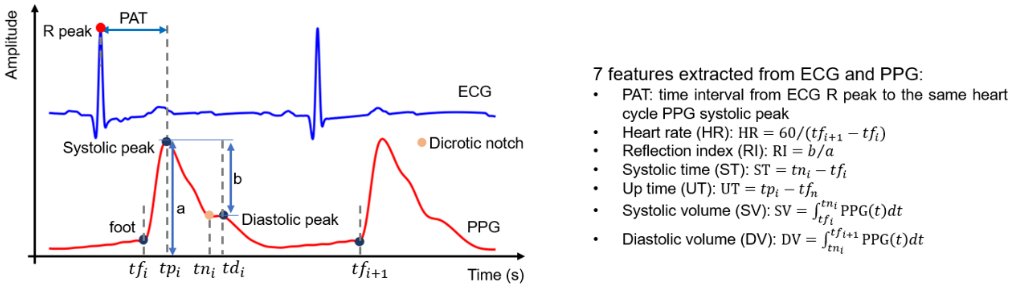

Figure S4: Illustration of 7 ECG and PPG features used for SBP and DBP estimation proposed in[3]. tf_i, tp_i, tn_i and td_i indicate the time index of foot point, systolic peak, dicrotic notch and diastolic peak of i^{th} PPG pulse, respectively.

S7 Benchmark and diagnose estimation algorithms

The proposed methods and metrics can be used to benchmark and diagnose the BP estimation algorithms. For example, two SOTA models, ApproximateNet[4] and DeepRNN-4L[3], were re-implemented with a same dataset **DailA BP**[1]. Therefore, the performance of these two estimation algorithms can be compared in terms of the similarities of distance, trend and the composite of both. For example, the percentages[1] of segments with different MAD levels of two estimation models are shown in Figure S5.

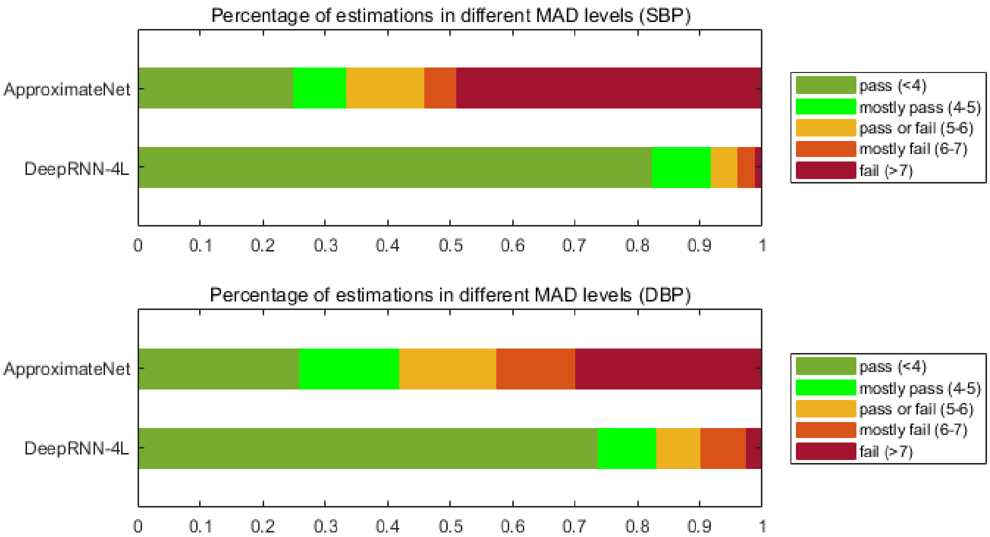

Figure S5: Percentage of segments with different MAD levels for SBP and DBP estimation using ApproximateNet and DeepRNN-4L.

It can be observed that the majority of both SBP and DBP estimations obtained using DeepRNN-4L can be considered as "Grade A" of IEEE Standard (SBP: $> 80\%$, DBP: $> 70\%$). However, only a small proportion of the estimations using ApproximateNet can be considered as "Grade A" of IEEE Standard (SBP: $\approx 25\%$, DBP: $\approx 25\%$) and even $\approx 50\%$ of SBP estimations using ApproximateNet dissatisfy IEEE Standard (MAD$\geq$7 mmHg).

The percentages of segments with correct trend tracking ($TS(x_i, y_i) \geq 0$) for SBP and DBP estimations using ApproximateNet and DeepRNN-4L are shown in Figure S6. It can be observed that DeepRNN-4L performs better than ApproximateNet in terms of trend tracking for both SBP and DBP.

The percentages of segments with MAD$\leq$ 5 mmHg and correct trend tracking for SBP and DBP estimation using ApproximatNet and DeepRNN-4L are shown in FigureS7. It can be observed that DeepRNN-4L performs significantly better than ApproximateNet in terms of both distance and trend. Approximate 78% and 58% of SBP and DBP estimations with MAD$\leq$ 5 mmHg and correct trend tracking ($TS(x_i, y_i) \geq 0$) using DeepRNN-4L. On the contrary, only around 16% and 25% of SBP and DBP estimations using ApproximateNet with MAD$\leq$ 5 mmHg and correct trend tracking.

[1]the percentage is obtained by calculating the ratio between the duration of all segments meet the requirement and the duration of the sequence.

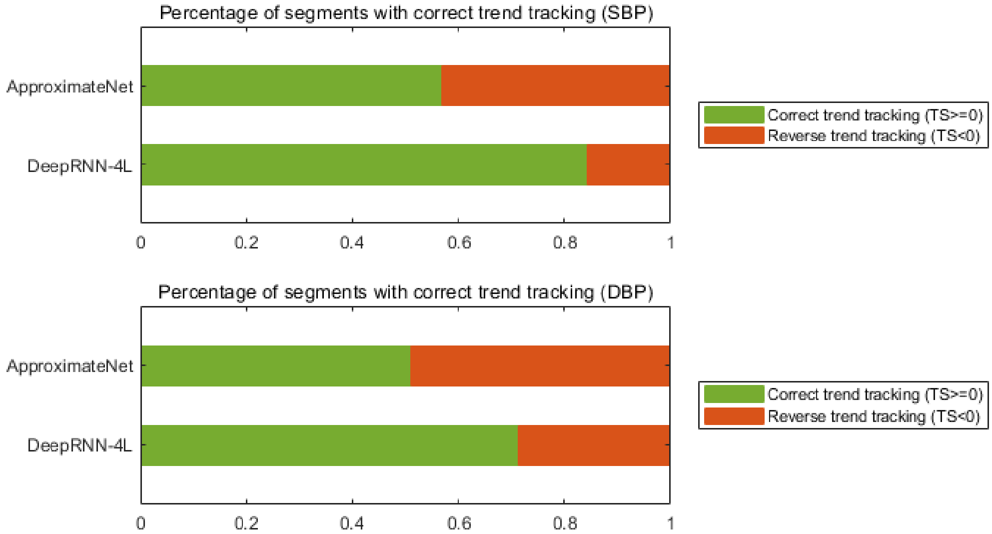

Figure S6: Percentage of segments with correct BP trend tracking for SBP and DBP using ApproximateNet and DeepRNN-4L.

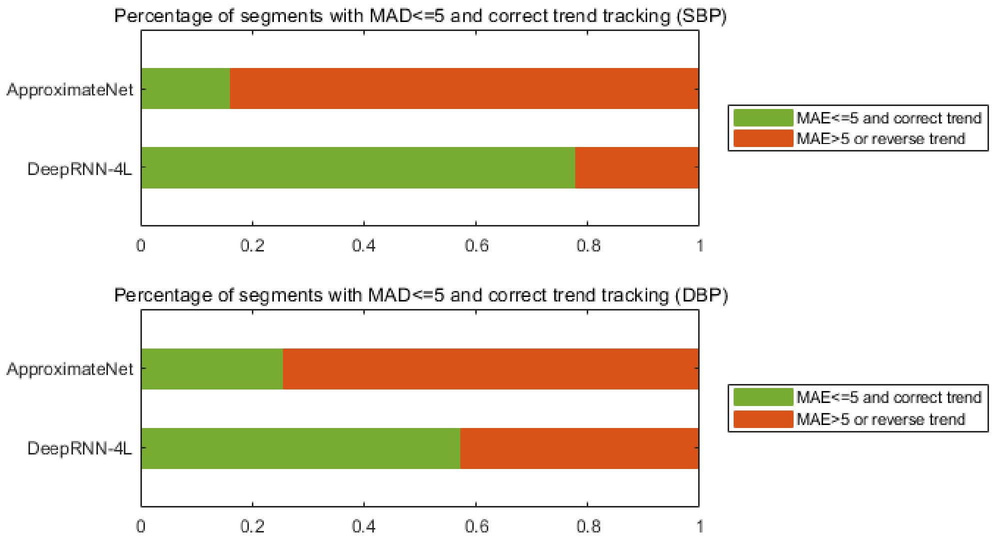

Figure S7: Percentage of segments with MAD$\leq$5 mmHg and correct BP trend tracking for SBP and DBP using ApproximateNet and DeepRNN-4L.

With the examples shown in this section, the proposed methods and metrics can be used to benchmark different estimation algorithms using a same dataset. Also, an estimation algorithm can diagnose and improve its performance by analyzing the similarity metrics proposed in this paper. For example, the DeepRNN-4L already performed well in BP trend tracking, and can just focus on minimizing the estimation error (e.g., MAD). However for ApproximateNet, it needs to improve in terms of both distance and trend tracking accuracy.

References

[1] Landry, C., et al. "Wearable Physiological and Blood Pressure Measurements During Activities of Daily Living." *IEEE Dataport* (2021).

[2] Wielemborek-Musial, Katarzyna, et al. "Blood pressure response to submaximal exercise test in adults." *BioMed research international 2016* (2016).

[3] Su, Peng, et al. "Long-term blood pressure prediction with deep recurrent neural networks." *2018 IEEE EMBS International conference on biomedical & health informatics (BHI).* IEEE, 2018.

[4] Ibtehaz, Nabil, et al. "PPG2ABP: Translating photoplethysmogram (PPG) signals to arterial blood pressure (ABP) waveforms." *Bioengineering* 9.11 (2022): 692.

[5] Kuchuee, Mohamad, et al. "Cuff-Less Blood Pressure Estimation." UCI Machine Learning Repository. (2015) https://doi.org/10.24432/C5B602

[6] Saeed, Mohammed, et al. "Multiparameter Intelligent Monitoring in Intensive Care II (MIMIC-II): a public-access intensive care unit database." *Critical care medicine* 39.5 (2011): 952.