공학박사 학위논문

Cuffless Blood Pressure Estimation based on Deep Learning approach: From Personalization to Generalization

딥 러닝 기반 비가압 혈압 추정:
개인화에서 일반화로

2021년 2월

서울대학교 대학원

협동과정 바이오엔지니어링 전공

이 동 석

Cuffless Blood Pressure Estimation based on Deep Learning approach: From Personalization to Generalization

딥 러닝 기반 비가압 혈압 추정: 개인화에서 일반화로

지도교수 박 광 석

이 논문을 공학박사 학위논문으로 제출함

2021년 2월

서울대학교 대학원

협동과정 바이오엔지니어링 전공

이 동 석

이동석의 공학박사 학위논문을 인준함

2020년 12월

위 원 장 최 진 욱 (인)

부 위 원 장 박 광 석 (인)

위 원 Sungwan Kim (인)

위 원 이 태 수 (인)

위 원 임 용 규 (인)

Ph.D. Dissertation

Cuffless Blood Pressure Estimation based on Deep Learning approach: From Personalization to Generalization

February 2021

Interdisciplinary Program in Bioengineering

The Graduate School

Seoul National University

Dongseok Lee

Cuffless Blood Pressure Estimation based on Deep Learning approach: From Personalization to Generalization

Academic advisor **Kwang Suk Park**

Submitting a Ph.D. Dissertation

February 2021

Interdisciplinary Program in Bioengineering
The Graduate School
Seoul National University

Dongseok Lee

Confirming the Ph.D. Dissertation written by Dongseok Lee

December 2020

Chair	**Choi, Jinwook, M.D., Ph.D.**	**(Seal)**
Vice Chair	**Park, Kwang Suk, Ph.D.**	**(Seal)**
Examiner	**Kim, Sungwan, Ph.D.**	**(Seal)**
Examiner	**Lee, Taesoo, Ph.D.**	**(Seal)**
Examiner	**Lim, Yonggyu, Ph.D.**	**(Seal)**

Abstract

Cuffless Blood Pressure Estimation based on Deep Learning approach: From Personalization to Generalization

Dongseok Lee

Interdisciplinary Program in Bioengineering

The Graduate School

Seoul National University

Blood pressure (BP) is one of the vital signs that provide fundamental health information on the heart and cardiovascular system of a patient. An estimated 1.13 billion individuals worldwide have hypertension. It is known as a "silent killer" because it is a risk factor for various diseases. Regular BP monitoring is important for diagnosing hypertension and predicting heart disease. Various cuffless blood

pressure estimation methods using physiological signal such as electrocardiogram and photoplethysmogram have been proposed.

In this study, I developed a cuffless BP estimation algorithm based on the deep learning approach. In addition, ballistocardiogram (BCG) signal was used to improve the performance of the model.

First, I developed a personalized BP estimation model based on the convolutional neural network (CNN) with the attention mechanism. CNN is known to show high performance in image learning. It exhibited the capability to extract features to estimate BP. In addition, the attention mechanism was used in the model to strengthen the importance of the feature in specific timesteps. The proposed model displayed mean absolute error (MAE) values of 4.06 mmHg and 3.33 mmHg for systolic BP and diastolic BP estimation, respectively.

Second, I investigated a generalized BP estimation model without the intermittent BP calibration. The calibration process is required to improve the accuracy of BP estimation. However, the process is laborious for a patient attempting to measure BP. The long short-term memory (LSTM) network model was applied rather than the CNN. This was because the CNN model may include noisy signal, which yields low performance. The model was investigated in terms of reproducibility. A multi-day test was performed using multiple measurement data. The results revealed that the proposed model showed performance higher than that of the personalized model. The MAE values of the proposed model were 5.82 mmHg and 5.24 mmHg for SBP and DBP, respectively.

In this study, I developed a cuffless BP estimation model based on the deep learning approach involving CNN and LSTM. The CNN model with the attention mechanism had the advantage of being applicable on the entire signal without the feature extraction procedure. The results showed that the BP estimation model with BCG signal displayed performance higher than that without the BCG signal. In addition, the LSTM model was investigated to make a generalized BP estimation model. The results showed that the proposed model displayed higher performance in a multi-day test. This study attempted to establish a BP estimation model that is both personalized and generalized. It enables continuous BP monitoring on a daily basis using unobtrusively measured ECG, PPG, and BCG.

Keyword: Cuffless blood pressure estimation, convolutional neural network, attention, long short-term memory network, personalization, generalization

***Student Number*: 2015-31047**

Table of Contents

List of Tables

List of Figures

List of Abbreviation

BP	**Blood pressure**
SBP	**Systolic blood pressure**
DBP	**Diastolic blood pressure**
ABP	**Arterial blood pressure**
ECG	**Electrocardiogram**
PPG	**Photoplethysmogram**
BCG	**Ballistocardiogram**
PTT	**Pulse transit time**
PWV	**Pulse wave velocity**
PVDF	**Polyvinylidene fluoride**
RRI	**R-R interval**
RJI	**R-J interval**
CNN	**Convolutional neural network**
RNN	**Recurrent neural network**
LSTM	**Long short-term memory**
GRU	**Gated recurrent unit**
ReLU	**Rectified Linear Unit**
MLR	**Multiple linear regression**
MAE	**Mean absolute error**
STD	**Standard deviation**
RMSE	**Root-mean-square error**
AAMI	**Association for the Advancement of Medical Instrumentation**
BHS	**British Hypertension Society**

Chapter 1. Introduction

1.1. Introduction to blood pressure measurement

1.1.1. Present blood pressure measurement

Blood pressure (BP) is one of the vital signs that provide fundamental health information on the heart and cardiovascular system of a patient. When the heart beats, the blood-flow from the heart exerts pressure on the blood vessels. The BP varies between systolic BP (SBP), which is an increase in pressure caused by the systolic contraction of the heart, and diastolic BP (DBP), which is the low pressure between instances of SBP. High BP (hypertension) is known as a "silent killer" because it is a risk factor for various diseases such as arrhythmia, heart attack, blindness, and brain stroke if it is not detected early nor treated appropriately. An estimated 1.13 billion people worldwide have hypertension.

The gold standard for BP monitoring is the mercury sphygmomanometer. It is generally used in the physician’s office. However, this method is not continuous and generally requires frequent calibration. In addition, BP measurement in medical circumstances may be incorrect because certain patients exhibit a higher BP than usual in a clinical setting, which is known as the "white-coat effect."

Arterial BP (ABP) is considered as a gold standard for continuous BP monitoring, which is performed in the intensive care unit. The BP is measured directly by an intravascular cannula module inserted in the arterial vessel. Because

it is an invasive method, the measurement of ABP can cause side effects such as bleeding and infection. Moreover, it is difficult to measure ABP in daily life because it requires a clinical setting and professional knowledge.

Because regular BP monitoring is important for diagnosing hypertension and predicting heart diseases, numerous devices have been developed to help patients measure BP at home or during their daily lives. These devices are generally based on the oscillometric method and use an inflatable upper-arm cuff (Figure 1-1) [1, 2]. However, although these methods provide noninvasive measurement in daily life, these do not offer continuous measurement. Wearing the cuff whenever the patient attempts to measure BP is a cumbersome task and can upset the patient. In addition, the significance of beat-to-beat BP analysis has increased. It involves parameters such as the blood pressure variability (BPV) [3].

Figure 1-1. Conventional BP measurement device using cuff.

1.1.2. Cuffless blood pressure measurement

Many researchers have investigated cuffless and continuous BP monitoring methods based on the pulse wave velocity (PWV). The PWV can be measured using physiological signals and can be expressed by the Moens–Korteweg (M–K) equation (1-1) and Hughes equation (1-2) [4]:

$$PWV = \sqrt{\frac{Eh}{\rho d}} \tag{1-1}$$

$$E = E_0 e^{\gamma P} \tag{1-2}$$

where E is the elastic modulus at the BP P; ρ denotes the density of the blood; and h and d are the thickness and radius, respectively, of the blood vessel. Furthermore, E_0 is the elastic modulus at zero BP, and γ is the coefficient of the blood vessel. When P increases, the elastic modulus increases, and thereby, the PWV increases.

PWV is inversely related to the pulse transit time (PTT) as follows:

$$\text{PWV} = \frac{L}{PTT}, \tag{1-3}$$

where L denotes the length from the heart to a certain peripheral site of the body, e.g., the finger. The PTT is the time taken by the pulse to propagate between two locations. The PTT can be calculated between the R-peak of the electrocardiogram (ECG) and characteristic point of the photoplethysmogram (PPG) that is measured at the finger. The ECG and PPG are measured using a noninvasive method and can conveniently be used in long-term monitoring or daily life without a cuff. The PTT

is known to be negatively correlated with BP [5], and various models have been developed to estimate BP with the PTT [6-8]. However, Payne et al. reported that the PTT method is not a reliable marker for BP estimation [9].

Several studies have recommended the ballistocardiogram (BCG) as a substitute for the PPG in the calculation of the PTT. The BCG is a measurement of the forces exerted on the body by the blood ejected from the heart. The BCG can be acquired by force sensors such as accelerometers, load cells, and film sensors including polyvinylidene fluoride (PVDF) sensors. Shin et al. proposed a BP measurement system using ECG and BCG on a weighing scale [10]. The BCG was measured on the weighing scale, and the RJ interval (RJI) between the R-peak of the ECG and J-peak of the BCG was measured. The results demonstrated that the RJI had a negative correlation with BP. BP was estimated using the linear regression method. Lee et al. recommended a BP monitoring chair using two-channel BCGs [11]. Two BCGs were measured at the back of the chair and the cushion on the seat, and the BP was estimated according to the phase difference between the two BCGs.

Although the PTT method showed remarkable performance in cuffless BP estimation, generality issue occurred. The present PTT method is based on personalized fitting. The accuracy of the model can be improved using a personal coefficient because the variables that were regarded as constant (such as the density of the blood, and thickness and diameter of the blood vessel) differ among individuals. However, the calculation of the coefficients for each individual is laborious and time-consuming. In addition, the personalized model requires

intermittent calibration with the cuff for an acceptable accuracy. This is because the arterial dimensions of the subject can vary with time owing to the effect of the autonomic nervous system and the vascular vasomotion.

Meanwhile, the generalized model requires more data than the personalized model does and thereby, more initial training time. In addition, the performance of the model may be lower than that of the personalized model. However, once the generalized model is trained, the training time of the model can be saved, and it can be applied rapidly to the subject in emergency scenarios. Furthermore, the temporal-generality in a subject circumvents the need for cumbersome calibration task.

Table 1-1. Comparison between personalized model and generalized model.

	Personalized model		**Generalized model**
Performance		>	
Initial training time		<	
Additional training time		>	
Intermittent calibration	Yes		No

1.2. Purpose of research

This study aimed to develop a cuffless BP estimation model using the deep learning approach. The deep learning algorithm has shown high performance for the classification problem and nonlinear regression in various areas including BP estimation. The deep learning algorithm can model the nonlinear expression between the physiological signals (ECG, PPG, and BCG) and target BP. In addition, I intended to improve the BP estimation performance using BCG additionally. In this study, two approaches of deep learning were applied according to the characteristic of the network.

Moreover, the research was designed to overcome the limitations of previous studies. Certain studies used a method to artificially vary the BP, such as the Valsalva maneuver, which can affect abnormal BP variation unlike in a real scenario. I intended to verify BP estimation without a specific BP variation scenario.

In addition, a generalized BP estimation model without the intermittent BP calibration was designed based on deep learning. The deep learning approaches are known to be highly effective algorithms for abstraction and generalization. The calibration process is required to improve the accuracy of BP estimation. However, the process is laborious for patients attempting to measure BP. The model applying the deep learning approach was investigated in terms of reproducibility, i.e., its capability to guarantee consistent performance across days.

1.3. Dissertation outline

This dissertation consists of the following chapters:

- Chapter 2 addresses the background information of the method used in this study and prior works related to the study.
- Chapter 3 addresses the waveform-based blood pressure estimation using convolutional neural network and attention mechanism.
- Chapter 4 describes the feature-based blood pressure estimation using bidirectional long short-term memory network.
- Chapter 5 presents the discussions and limitation of the preceding chapters, and Chapter 6 presents the conclusion of the dissertation.

This dissertation is based on the following publications:

- H. Eom, D. Lee, S. Han, Y. S. Hariyani, Y. Lim, I. Sohn, K. Park, and C. Park, “End-to-End Deep Learning Architecture for Continuous Blood Pressure Estimation Using Attention Mechanism,” *Sensors (Basel)*, vol. 20, no. 8, Apr. 2020 [12].
- 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 (Basel),* vol. 21, no. 1, Dec. 2020 [13].

The author of this dissertation contributed to the above studies as follows: conceptualization and design of the experiments; data acquisition, analysis, and interpretation; and draft and revision of the manuscript.

Chapter 2. Background and related studies

2.1. Cuffless blood pressure estimation based on deep learning approaches

Several studies have proposed the use of machine learning (including deep learning) algorithms to perform continuous BP estimation automatically without a cuff. Chan et al. [14] and Kachuee et al. [15] proposed a model for estimating BP based on features extracted from ECG and PPG using conventional machine learning algorithms such as linear regression and AdaBoost. Su et al. proposed a recurrent neural network (RNN)-based BP estimation model using features extracted from ECG and PPG [16]. Kurylyak et al. [17], Lee et al. [11], and Wang et al. [18] also proposed a simple artificial neural network (ANN) model that uses features extracted from only one signal such as PPG or BCG. However, these methods have problems in that the extraction of features is expensive and laborious. In addition, if the signal is noisy, it may be difficult to obtain adequate data to train the neural network.

In recent studies, a few authors attempted to estimate BP using raw signals without feature engineering. Slapničar et al. estimated BP with only raw PPG using ResNet, a deep learning model that have displayed good performance in the field of image classification [19]. Tanveer et al. also used ECG and PPG raw signals and achieved good performance [20]. However, the length of the data used for BP estimation was 16 s, which is excessively long for emergency scenarios.

2.2. Deep learning network

2.2.1. Convolutional neural network (CNN)

Convolutional neural networks (CNNs) have achieved considerable success in various challenging areas by extracting key features from a large amount of data such as images. Recently, CNNs have achieved remarkable performance in signal processing, particularly in biomedical areas [21-24].

The key concept underlying CNN is that it has a convolutional layer rather than a classical perceptron layer. In image data training, the convolutional layer can learn the spatial pattern of the image, whereas the perceptron layer can learn the pattern through the pixels. A CNN generally has multiple convolutional layers. The previous layer concentrates on low-level features, and the feature map is connected to the next layer that can learn high-level features.

2.2.2. Recurrent neural network

A recurrent neural network (RNN) is a network with a loop structure. RNN have displayed higher performance than other networks for processing sequential time-series data. It has been applied for sequential modeling in natural language understanding and video processing. An RNN has a loop in the network, and the output of a timestep is used as a state of the next timestep. A schematic diagram of an RNN is shown in Figure 2-1. The network repeats the procedure in each timestep according to the sequence.

However, the conventional RNN displays the vanishing gradient problem, particularly while handling long time series data [25]. Two representative RNN-based models have been proposed to solve this long-term dependency problem: the long short term memory (LSTM) developed by Hochreiter et al. [26] and the gated recurrent unit (GRU) introduced by Cho et al. [27].

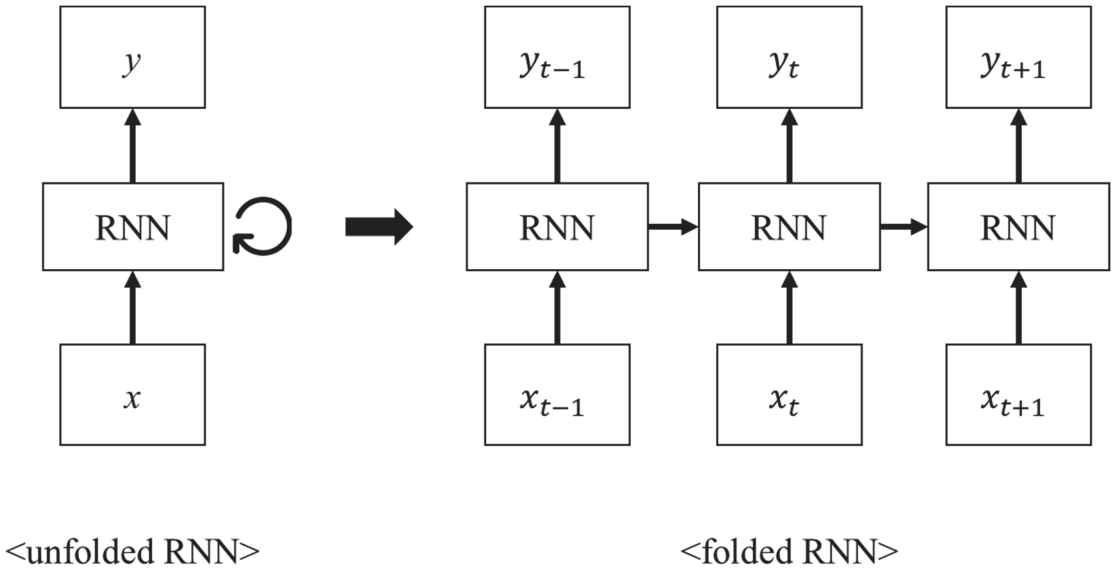

Figure 2-1. Schematic diagram of RNN.

.

2.2.2.1. Long short-term memory network

The long short-term memory (LSTM) network replaces the RNN cell with LSTM cells. The central concept underlying LSTM is that an LSTM cell has three gates: forget gate, input gate, and output gate. The structure of LSTM is illustrated in Figure 2-2a).

The forget gate (f_t) controls the amount of information that would be forgotten using the hidden state and input vector. The input gate (i_t) decides which value would be updated and updates the state of the cell. The output gate (o_t) controls the amount of information that would be outputted. These gates can aid the network in learning long time-series data or eliminating inconsequential data and thereby, learn patterns with a long duration. The gates can be expressed by the following equations:

$$f_t = \sigma\big(W_f \cdot [h_{t-1}, x_t] + b_f\big)$$

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i)$$

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o)$$

$$\tilde{c}_t = \tanh(W_c \cdot [h_{t-1}, x_t] + b_c)$$

$$c_t = f_t * c_{t-1} + i_t * \tilde{c}_t$$

$$h_t = \tanh(c_t) * o_t$$

σ, *W, h, x, b, c,* and *o* represent the sigmoid, weight matrix of each gate, hidden state, input vector, bias of each gate, cell state, and output vector, respectively.

Bidirectional LSTM (Bi-LSTM) is an extension of LSTM in which the input sequence is read forward and backward and both outputs are concatenated. Bi-LSTM is more effective than LSTM because it can learn the pattern in both directions.

2.2.2.2. Gated Recurrent Unit

The gated recurrent unit (GRU) is a variation of LSTM. It has two gates: forget gate and input gate (See Figure 2-2b)). A hidden state h_t replaces c_t and h_t in LSTM, and a reset gate vector z_t controls two gates. When the output of z_t is one, the forget gate opens and input gate closes. When the output of z_t is zero, the forget gate closes and input gate opens.

Similar to bidirectional LSTM, GRU can be used bidirectionally. The bidirectional GRU (Bi-GRU) has two layers: the forward and backward layers. The outputs of both directions are concatenated. The structure of Bi-GRU is illustrated in Figure 2-3.

GRU has less learning parameters than LSTM and requires less training time. However, the expression capability may be lower than that of LSTM. In this study, GRU was used with the CNN (Chapter 3), and LSTM was used independently (Chapter 4).

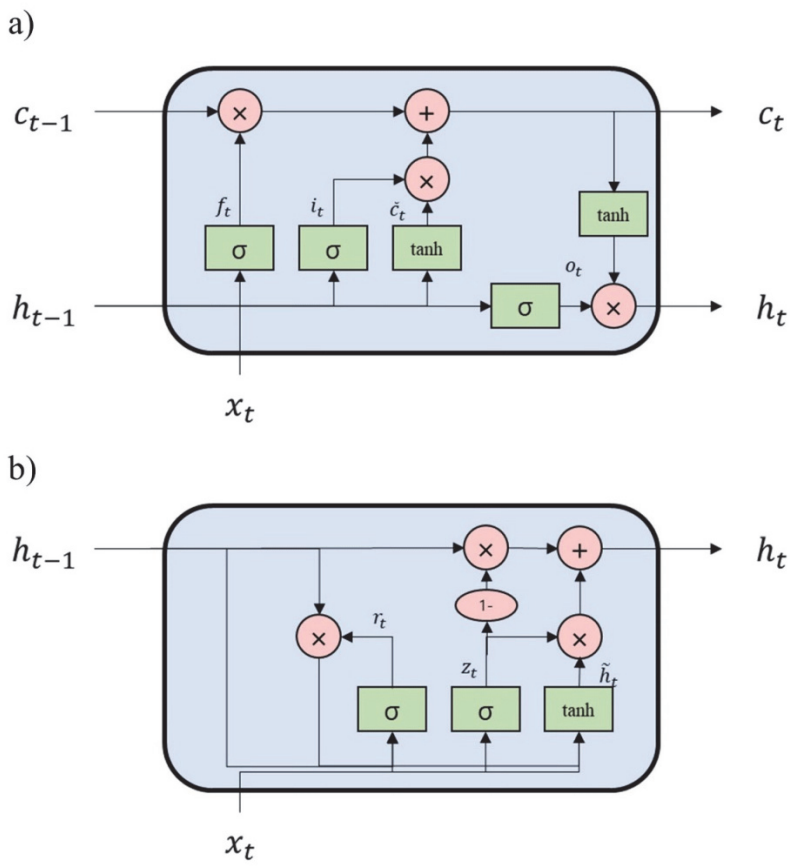

Figure 2-2. a) Internal structure of LSTM cell. b) Internal structure of GRU cell.

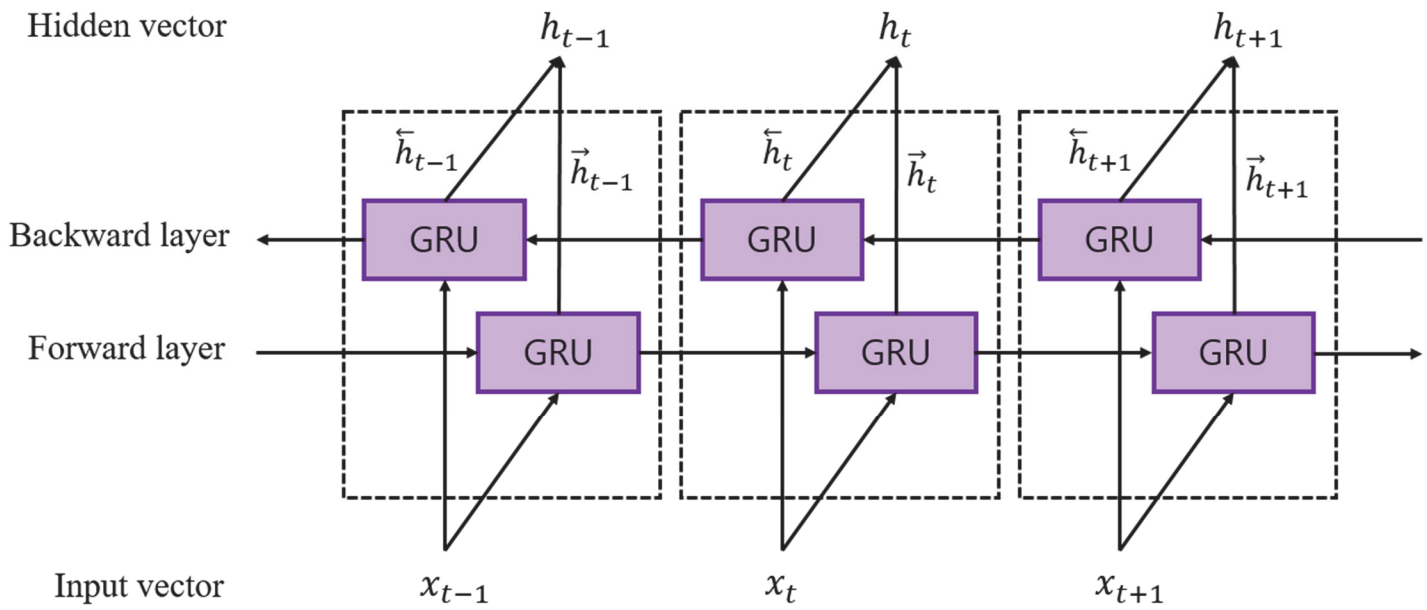

Figure 2-3. Structure of bidirectional GRU.

2.2.3. Attention Mechanism

The attention mechanism has recently been demonstrated to be efficient in sequence modeling tasks such as image captioning, neural machine translation, and signal processing [28-30]. The key concept underlying the attention mechanism is that the model can pay "attention" to the specific timestep that is strongly related to the target output, rather than applying an identical weight in each timestep [31]. Figure 2-4a) shows an example of an attention algorithm for sequence-to-sequence (encoder–decoder network). The attention score is calculated as the dot product of the hidden state in the encoder h_i and that in the decoder s_t. Then, the attention weights a_i are generated using the softmax function, with the sum of weights equated to one. The final attention value (known as context vector) is calculated with the weighted sum of the hidden state values and concatenated to calculate the output vector.

Because sequential feature vectors from Bi-GRU in the proposed model may contribute differently for estimating BP values, I added an attention layer to automatically train how important the feature vectors were in each time step. Larger weights can be assigned to significant information using the attention mechanism, and various methods to apply the attention mechanism have been proposed [32]. In this work, the feed-forward attention model was used [33]. The structure of the attention layer used in this study is illustrated in Figure 2-4b).

Given the Bi-GRU hidden state vector h_i at each timestep $i \in [1, N]$, the importance score s_i was calculated through the score function $score(\cdot)$ as follows:

$$s_i = score(W_s h_i + b)$$

where the equation can be represented as a single-layer perceptron having a trainable weight W_s and a bias b. The score function can be set as an activation function in the neural network. The tanh function was used in the proposed model.

After obtaining each importance score s_i for the hidden state vector h_i, the attention weight a_i was evaluated using the softmax function. It is expressed as

$$a_i = softmax(s_i) = \frac{\exp(s_i)}{\sum_i \exp(s_i)}$$

Finally, the context vector c was obtained by calculating the weighted sum of the attention weight vector and the corresponding hidden state vector, as follows:

$$c = \sum_i^N a_i h_i$$

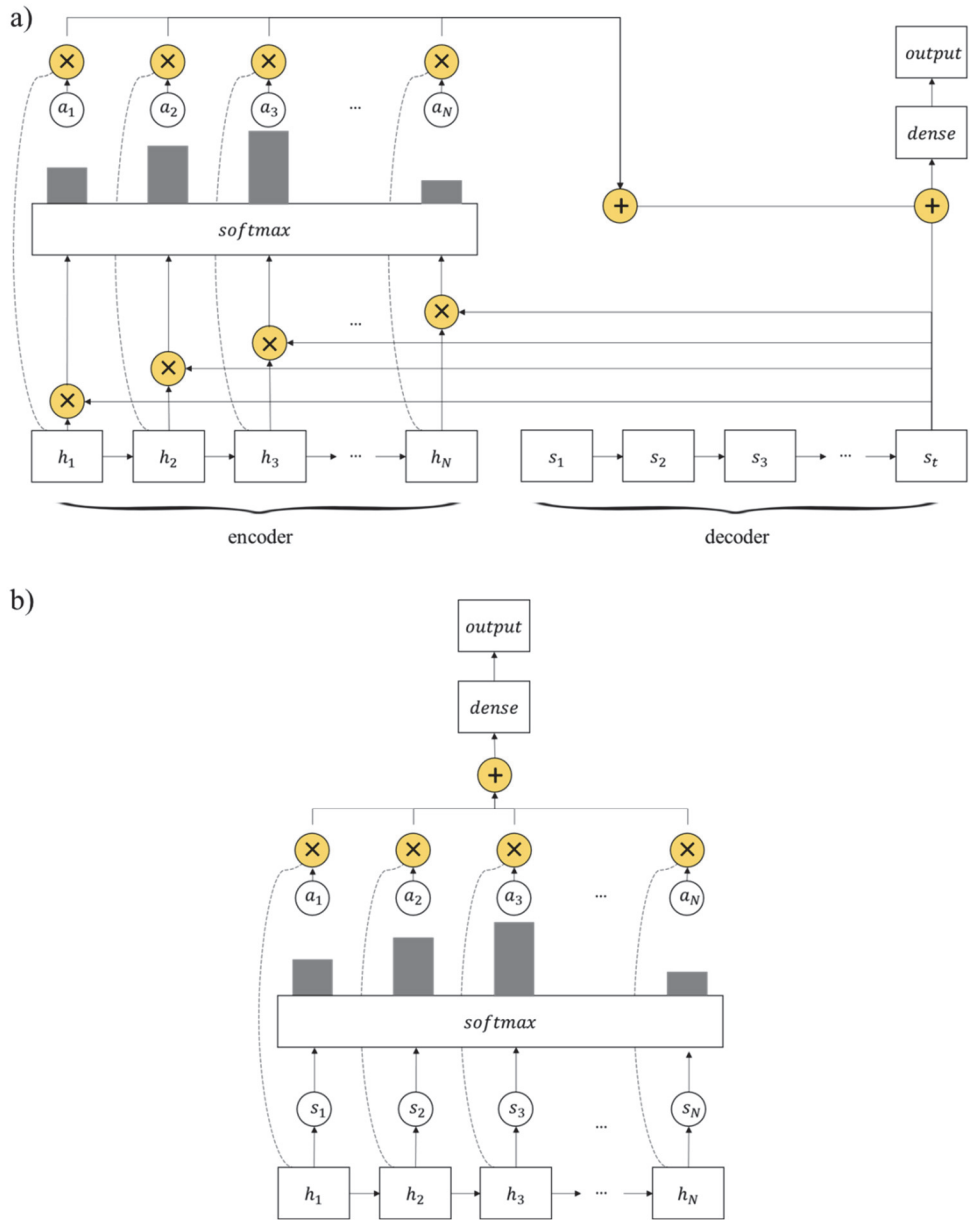

Figure 2-4. a) Structure of dot-product attention in seq2seq model. b) Structure of feed-forward attention.

Chapter 3. Waveform-based BP estimation using CNN and attention mechanism

3.1. Methods

3.1.1. Data Acquisition

A total of 15 subjects (6 men, 9 women, age: 26.2 ± 3.0) were recruited for this study. The subjects had no reported medical conditions. Written informed consent was obtained from the subjects. Furthermore, this study was approved by the Institutional Review Board of Seoul National University Hospital (IRB No. 1801-016-912).

The experimental setup is illustrated in Figure 3-1. Three Ag/AgCl electrodes were attached to the subject's left arm, right arm, and left leg according to Einthoven's triangle. ECG was acquired on Lead II using the electrodes with the BIOPAC ECG100C module, and PPG was measured at the subject's index finger using a commercial module (PSL-iPPG2C) [34]. In addition, the BCG signals were measured using a PVDF film sensor attached to the seat of the chair. The SBP and DBP were also measured simultaneously using a continuous BP monitoring device (Finometer Pro (Finapres Medical Systems, Enschede, The Netherlands)). Finometer Pro is composed of a processing unit and two cuffs (a finger cuff and an arm cuff). The reference BP is measured with the arm cuff at the beginning of the measurement, and the beat-to-beat BP is measured at the finger cuff sensor. All the

data were synchronized and sampled at 1000 Hz with a data acquisition device (BIOPAC MP150 module (BIOPAC Systems Inc., Goleta, CA, USA)).

After the measurement device was attached, the subjects were asked to sit on the armchair. The data were measured for 30 min while the subjects were in the resting condition. The measured BP values are shown in Figure 3-3. The mean and standard deviation (STD) of SBP and DBP were 115.04 ± 14.64 mmHg and 70.01 ± 9.56 mmHg, respectively.

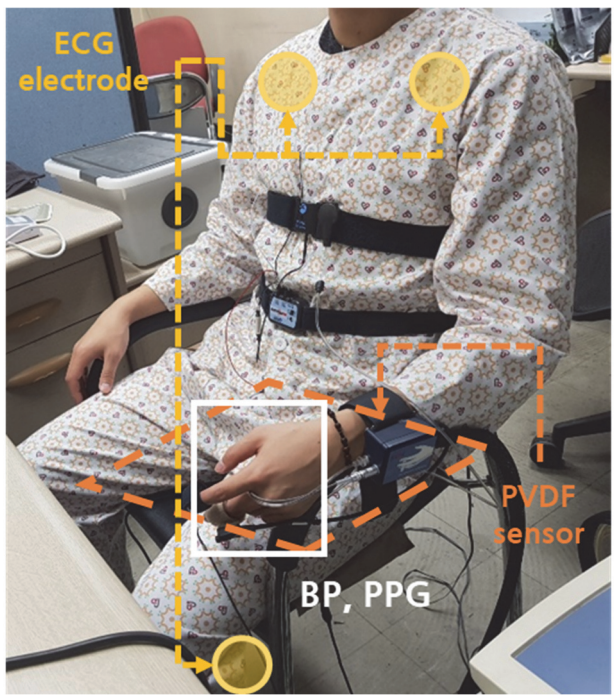

Figure 3-1. Overview of measurement setup.

3.1.2. Data Preprocessing

A second-order Butterworth bandpass filter was applied to the data to remove baseline wandering and power-line noise. The cutoff frequency of each filter applied to each signal is summarized in Table 3-1.

All the combinations of the raw signals (ECG, PPG, and BCG) were used as input for the deep learning model to investigate the effect of each signal on the BP estimation. To ensure information sufficiency, 5 s segments of the signals were used as input to the model with an overlap of 1 ms. The data of the first 20 s and final 5 s were excluded because of the noise in the data acquired. A total of 26,625,000 samples were generated.

Because a large amount of data was used as input, the data were resampled to 125 Hz for efficient learning. The target labels were set as the BP values (SBP, DBP) corresponding to the end of each segment, as shown in Figure 3-2. The BP values that were beyond the range (mean ± 1.96·STD) were considered as outliers and were eliminated. The BP distribution that was used in this study is illustrated in Figure 3-3. The y-axis represents the number of sequences.

Table 3-1. Cutoff frequency of bandpass filter applied to each signal.

Signal	**HPF (Hz)**	**LPF (Hz)**
ECG	0.5	35
BCG	4	15
PPG	0.5	15

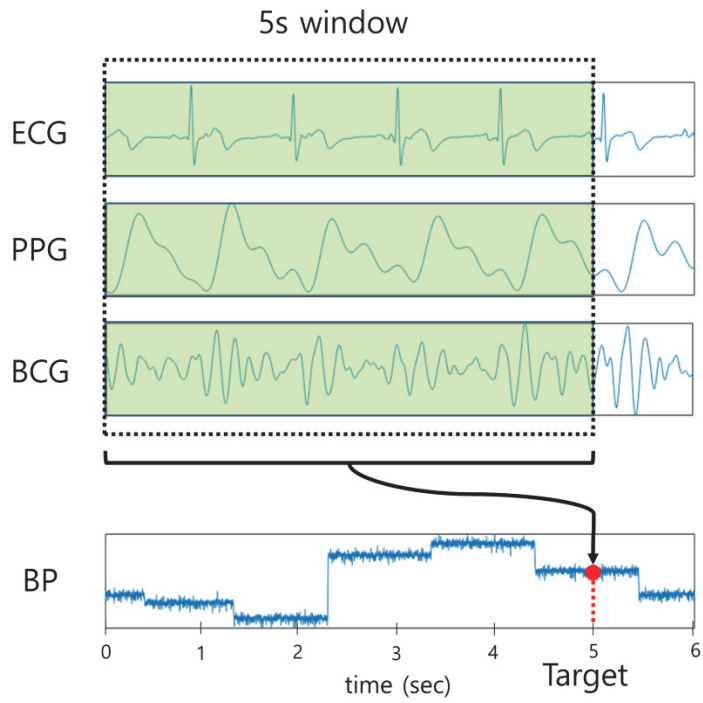

Figure 3-2. Target BP labeling of the deep learning model.

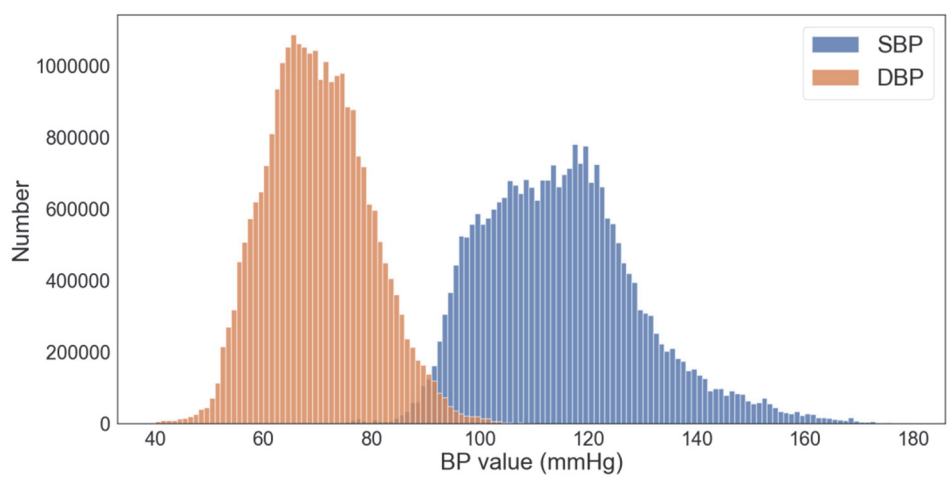

Figure 3-3. BP distribution used in this study.

3.1.3. Deep learning architecture

The proposed end-to-end deep learning network consists of a CNN layer, a Bi-GRU layer, and an attention layer, as shown in Figure 3-4. In the proposed model, the CNN structure was designed by referring to the VGG-16 network structure [35]. The original VGG-16 has 13 convolutional layers and 3 fully connected layers. However, the structure was modified to be applied to the proposed model. The detailed structure of the model is presented in Table 3-2. The fully connected layers in the VGG-16 network were changed to a Bi-GRU layer. In addition, the final three convolutional blocks were excluded to leave an acceptable number of hidden nodes in the sequences to be applied for Bi-GRU and the subsequent attention layer.

A total of 10 convolution layers with a rectified linear unit (ReLU) activation function were used to extract spatial pattern vectors from signals. Each convolution layer was followed by a batch normalization layer to reduce the internal covariate shift. The final layer of each convolution module was set to a max-pooling layer to reduce the length of the inputs. The same padding was applied to each convolution operation. Both kernel and pooling size were set to three, and the output channel size of the convolution layer was scaled up by a factor of 2 from 64 to 512 as it passed through each convolution module.

In the Bi-GRU phase, 64 hidden nodes were set up in each of the forward and backward layers, and 128 features were generated at each timestep. The Bi-GRU layer can encode temporal information between features in the proposed model. In

addition, representative features can be acquired in the Bi-GRU layer by reducing the feature dimensions from 512 to 128. The output of the Bi-GRU hidden state vectors were weighted and summed. Then, the SBP and DBP values were output through a one-layer perceptron. The detailed structure of the proposed model is illustrated in Table 3-2.

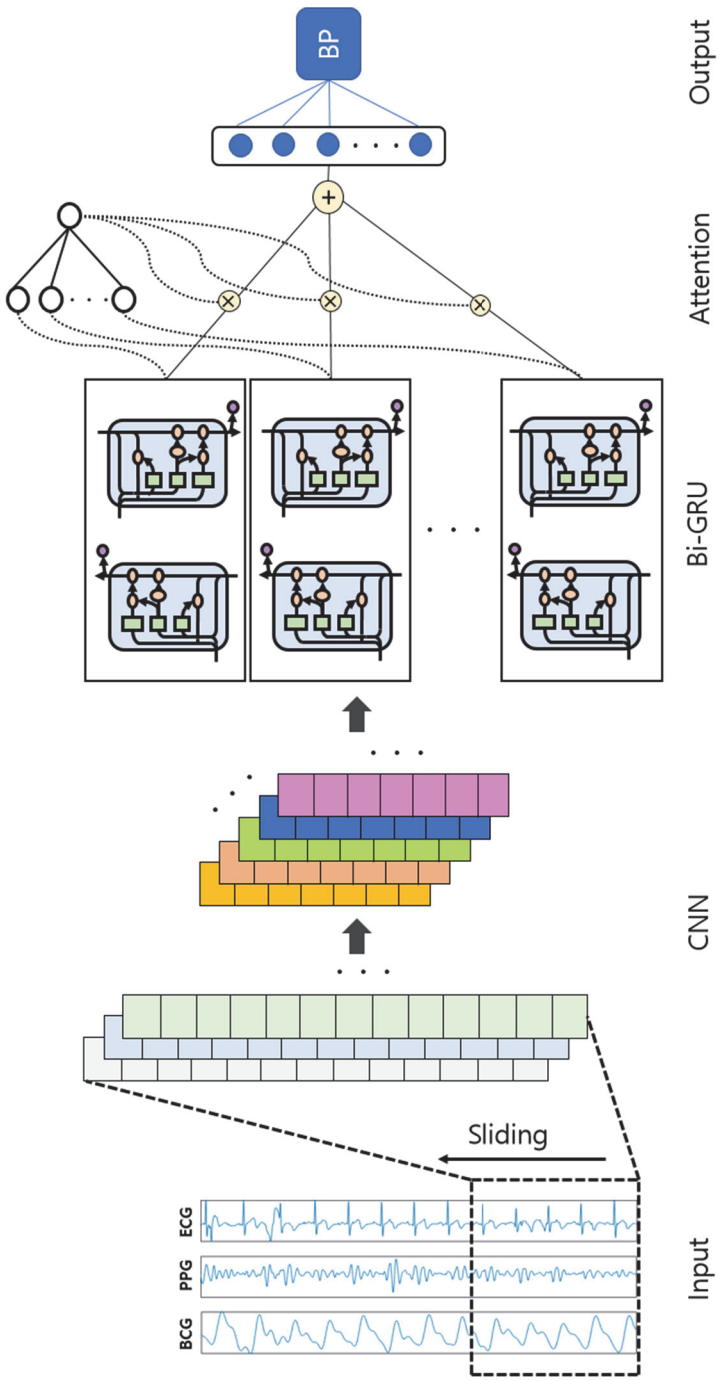

Figure 3-4. Overall structure of proposed network.

Table 3-2. Detailed structure of proposed model.

Network	Layer	Input Shape	Output Shape	Padding	Stride	Kernel
CNN	Conv	625×3	625×64	Same	1	3
	BN+ReLU					
	Conv	625×64	625×64	Same	1	3
	BN+ReLU					
	Maxpool (size = 3)	625×64	209×64	Same	3	-
	Conv	209×64	209×128	Same	1	3
	BN+ReLU					
	Conv	209×128	209×128	Same	1	3
	BN+ReLU					
	Maxpool (size = 3)	209×128	70×128	Same	3	-
	Conv	70×128	70×256	Same	1	3
	BN+ReLU					
	Conv	70×256	70×256	Same	1	3
	BN+ReLU					
	Conv	70×256	70×256	Same	1	3
	BN+ReLU					
	Maxpool (size = 3)	70×256	24×256	Same	3	-
	Conv	24×256	24×512	Same	1	3
	BN+ReLU					
	Conv	24×512	24×512	Same	1	3
	BN+ReLU					
	Conv	24×512	24×512	Same	1	3
	BN+ReLU					
	Maxpool (size = 3)	24×512	8×512	Same	3	-
Bi-GRU	Forward	8×512	8×64		-	
	Backward	8×512	8×64		-	
	Concatenation (Forward + Backward)					
Attention	1-layer perceptron	8×128	128		-	
	Activation tanh					
	Softmax					
	Weighted sum					
Output	1-layer perceptron	128	2		-	

3.1.4. Training setting

Seventy percent of the dataset was used for training, 10% for validation, and 20% for testing. The Adam optimizer [36] was used with a learning rate of 10^{-3} and decay of 10^{-4} to optimize the hyperparameters of the model. The learning rate was set to the optimal value empirically, and the initial weights were selected randomly. The mean squared error (MSE) was used for the loss function. The model was trained with the early stopping method with a patience of 10 in a maximum of 50 epochs. The batch size was set to 512. The Keras deep learning framework with TensorFlow backend and NVIDIA GeForce RTX 2080Ti (NVIDIA corporation, Santa Clara, CA, USA) (with 11 GB VRAM) was used as the computing environment for network training. A sample of convergence curve is illustrated in Figure 3-5.

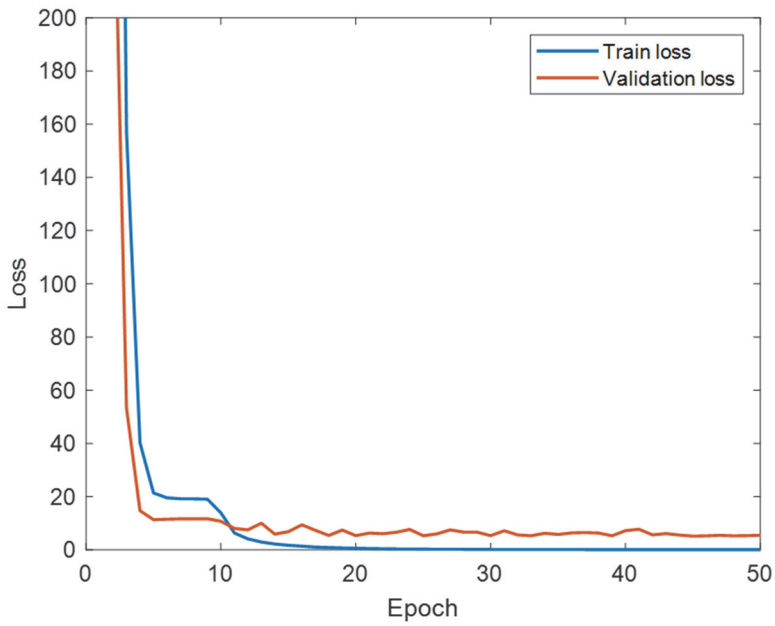

Figure 3-5. Sample convergence curve in 50 epochs.

3.2. Results

Three types of results are presented in this section. First, the BP estimation results are compared for each combination of signals. Second, the performances of the models with and without attention are compared when all the signals were used as inputs. Finally, the performance of the end-to-end deep learning model is compared with that of the multiple linear regression (MLR) model that used interval features from the characteristic point of each signal. The root-mean-square error (RMSE) and mean absolute error (MAE) were used as metrics of the BP estimation accuracy. In addition, the values of the coefficient of determination (R^2) between the reference and estimated BP were calculated for all the results. Furthermore, the Bland–Altman plot [37] was presented to increase the reliability of the results.

3.2.1. Performance comparison by signal combination

A summary of the results from all the combinations of the three signals is shown in Table 3-4. The model with ECG + PPG and ECG + PPG + BCG shows performance higher than that of the other models in terms of BP estimation. In addition, the proposed model with attention displays an error lower than that without attention. The MAE value is 4.06 mmHg and 3.33 mmHg for SBP and DBP, respectively.

The estimation accuracy when a combination of signals (ECG + PPG, ECG + BCG, PPG + BCG, and ECG + PPG + BCG) was used as input was significantly higher than those of the cases where a single signal (ECG, PPG, and BCG) was used as the input. The detailed results are shown in Table 3-3.

Table 3-3. Mean values of RMSE, MAE, and R^2 when the input was a single signal and when it comprised multiple signals.

Input	**SBP (mmHg)**			**DBP (mmHg)**		
	RMSE	MAE	Mean R^2	RMSE	MAE	Mean R^2
Single signal	7.04	5.47	0.24	5.39	4.19	0.21
Multiple signals	6.21	4.78	0.40	4.83	3.76	0.35

Table 3-4. Performance comparison for combinations of input signals for the model without attention and that with attention. The 95%-confidence interval is indicated below the error of the proposed model.

Model	Input	SBP				DBP			
		RMSE (mmHg)	MAE (mmHg)	STD	R^2	RMSE (mmHg)	MAE (mmHg)	STD	R^2
CNN+Bi-GRU	ECG	7.02	5.51	4.66	0.24	5.16	4.06	3.45	0.27
	PPG	6.88	5.34	4.60	0.28	5.73	4.45	4.09	0.14
	BCG	7.24	5.59	5.03	0.20	5.29	4.06	3.71	0.22
	ECG, PPG	5.83	4.46	4.06	0.46	4.74	3.70	3.37	0.38
	ECG, BCG	6.74	5.30	4.60	0.31	4.82	3.74	3.27	0.34
	PPG, BCG	6.44	4.86	4.50	0.36	5.04	3.88	3.62	0.27
	ECG, PPG, BCG	5.87	4.51	4.14	0.48	4.73	3.71	3.39	0.40
CNN+Bi-GRU +Attention (proposed model)	ECG, PPG, BCG	5.42 [1.97, 8.87]	4.06 [1.53,6.59]	4.04	0.52	4.30 [0.94, 7.72]	3.33 [0.61, 6.05]	3.42	0.49

A repeated measurement analysis of variance (ANOVA) test was used to compare the performance with varied combinations as input. The differences between models were significant in both SBP and DBP estimation ($p < 0.01$). In addition, a paired t-test was performed between the results to compare each model. As shown in Table 3-5 and Table 3-6, the proposed model statistically outperformed the other methods regardless of the input, in both SBP and DBP estimation.

Table 3-5. Results of paired t-test between various inputs for SBP estimation.

Inputs	ECG	PPG	BCG	ECG, PPG	ECG, BCG	BCG, PPG	ECG, BCG, PPG	ECG, BCG, PPG (with attention)
ECG		-	-	$p < 0.05$	-	-	$p < 0.05$	$p < 0.05$
PPG			-	$p < 0.05$	-	$p < 0.05$	$p < 0.05$	$p < 0.05$
BCG				$p < 0.05$	-	$p < 0.05$	$p < 0.05$	$p < 0.05$
ECG, PPG					$p < 0.05$	-	-	$p < 0.05$
ECG, BCG						-	$p < 0.05$	$p < 0.05$
BCG, PPG							-	$p < 0.05$
ECG, BCG, PPG								$p < 0.05$

Table 3-6. Results of paired t-test between various inputs for DBP estimation.

Inputs	ECG	PPG	BCG	ECG, PPG	ECG, BCG	BCG, PPG	ECG, BCG, PPG	ECG, BCG, PPG (with attention)
ECG		-	-	-	$p < 0.05$	-	-	$p < 0.05$
PPG			-	$p < 0.05$	$p < 0.05$	$p < 0.05$	$p < 0.05$	$p < 0.05$
BCG				-	-	-	-	$p < 0.05$
ECG, PPG					-	-	-	$p < 0.05$
ECG, BCG						-	-	$p < 0.05$
BCG, PPG							-	$p < 0.05$
ECG, BCG, PPG								$p < 0.05$

3.2.2. Attention Mechanism Performance

As shown in Table 3-4, the model with the attention mechanism showed lower RMSE and MAE values than that without the attention mechanism. An example of BP estimation results is presented in Figure 3-6. The model with the attention mechanism estimated the fluctuation in BP more precisely than that without the attention mechanism. The accuracy of the model can be improved with the attention mechanism by learning more intensively in the timestep that contains relatively important information in the 5 s input data.

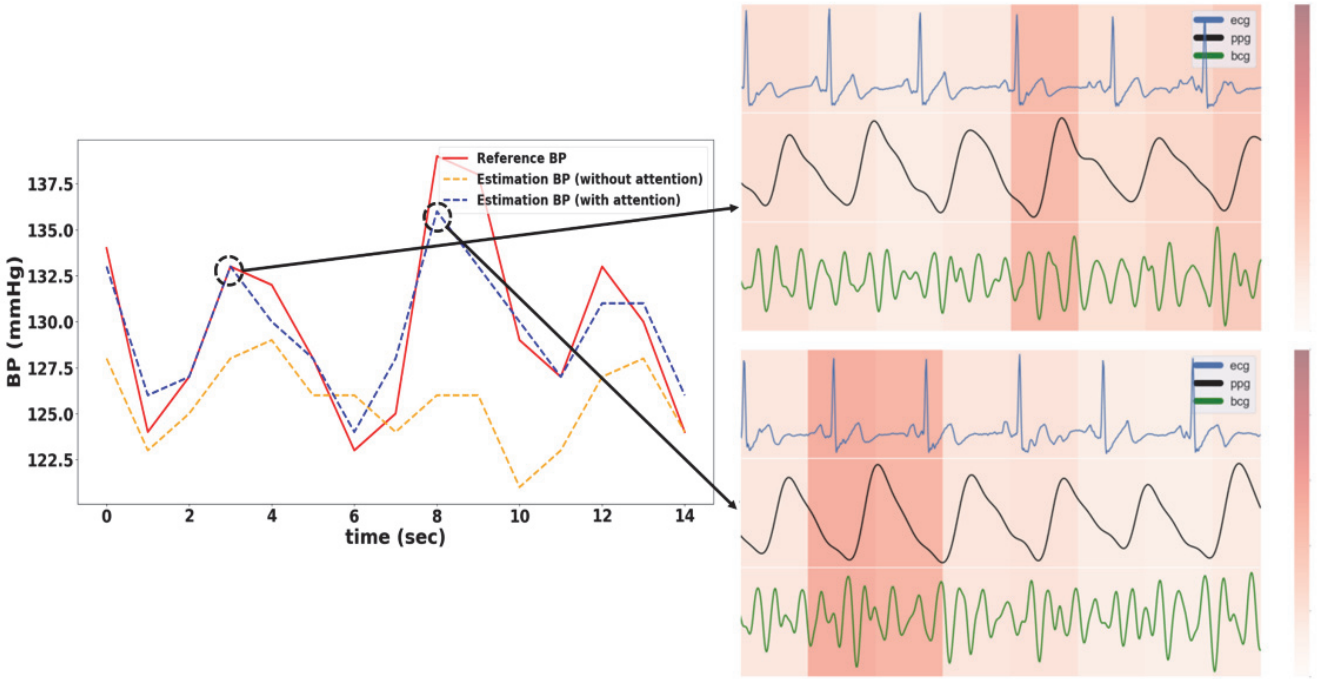

Figure 3-6. Left: Sample of estimated BP with and without attention mechanism; Right: Heatmap of the weights of the attention mechanism at the point where the error was low. The darker color denotes a higher attention weight.

In addition, an investigation of the attention heat map revealed that the attention weight was high at a specific time. The process of generating the attention heat map was as follows: The length of the data was reduced from 625 to 8 by the CNN pooling layer. Thereby, Bi-GRU generated eight hidden state vectors. Accordingly, the original signal could be divided into eight sections. The attention heat map was generated by assigning the attention weight obtained from the attention layer to each section.

The ANOVA test was conducted to interpret the attention weights of each timestep. The results are summarized in Figure 3-7. The differences between timesteps are statistically significant ($p < 0.001$), and the weights from Timestep 2 to Timestep 6 are significantly higher than those for the other timesteps. Timesteps 1–7 includes 648 ms of feature information, and Timestep 8 has 464 ms of feature information. This implies that the data from Timestep 2 to Timestep 6 (from 1.112 s to 3.704 s before the target BP) had information that is significant for BP estimation.

Figure 3-8 shows the Bland–Altman plot between the estimated and reference BP. The limits of agreement (LOA) at 95%-confidence intervals for DBP and SBP are measured as [-9.50, 9.50] and [-11.24, 11.63], respectively. This implies that 95% of the error was within [lower LOA, upper LOA]. Moreover, the mean error values between estimation and reference were 0.03 and 0.20 for DBP and SBP estimation, respectively. This implies that the model displayed negligible bias.

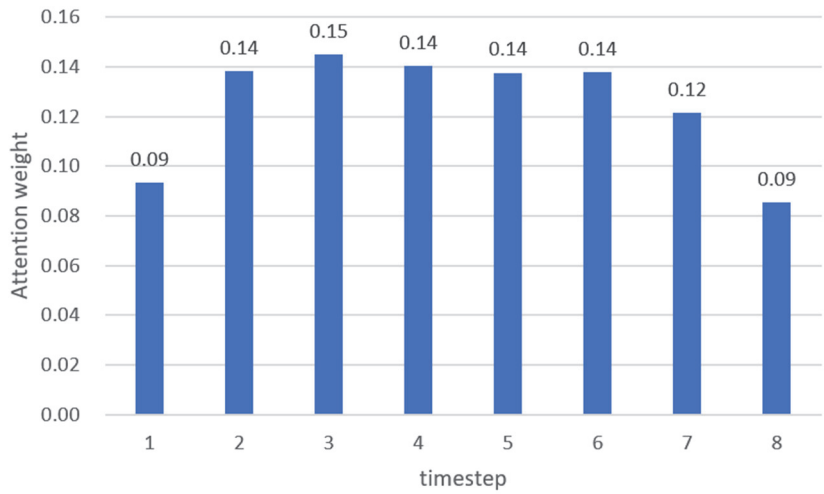

Figure 3-7. Mean attention weight across all datasets for each timestep.

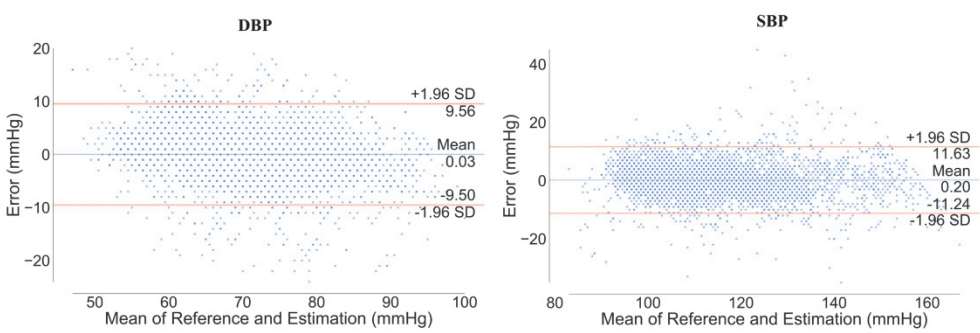

Figure 3-8. Bland–Altman plot of DBP and SBP. The orange line denotes the limit of agreement (LOA), and the blue line denotes the mean of difference (error) between reference BP and estimated BP.

Figure 3-9 shows the BP estimation result from two subjects. Overall, the estimated BP is similar to the reference BP, as shown in Figure 3-9a). However, as shown in Figure 3-9b), the error is high in cases where BP rapidly varies in a short time or if the overall range of BP is wide.

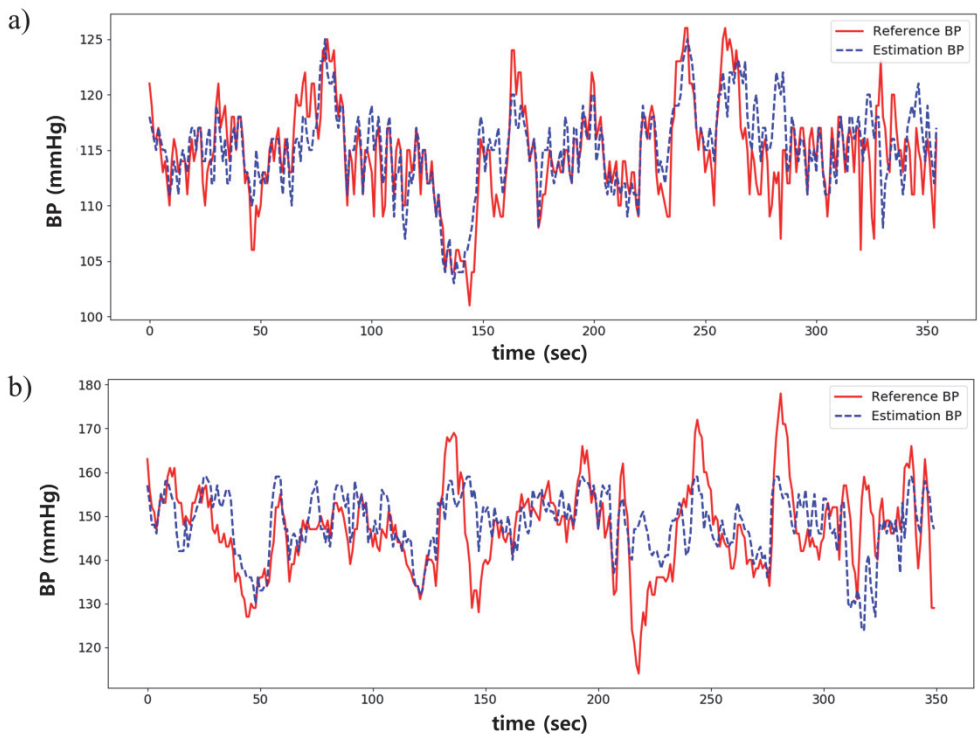

Figure 3-9. Comparison between estimated and reference BP. a) best case; b) worst case.

3.2.3. Comparison with Multiple Linear Regression model

The results of the proposed model were compared with those of the multiple linear regression (MLR) model. The characteristic points from ECG, BCG, and PPG were detected to extract the features for the MLR model. The ECG R-peak and the peak of the first derivative PPG were detected based on Pan and Tompkins's algorithm [38]. The BCG J-peak was detected by identifying the highest peak between 110 ms and 250 ms after each R-peak. False-positive peaks were excluded manually. Then, the R-R interval (RRI), PTT, and RJI were calculated from each cardiac cycle (Figure 3-10a)). In addition, 21 ± 11% of the cardiac cycles were excluded in cases where the peak was not detected because of motion artifacts (Figure 3-10b)). The features were utilized as inputs of the MLR model, and SBP and DBP were used as dependent variables of the MLR model.

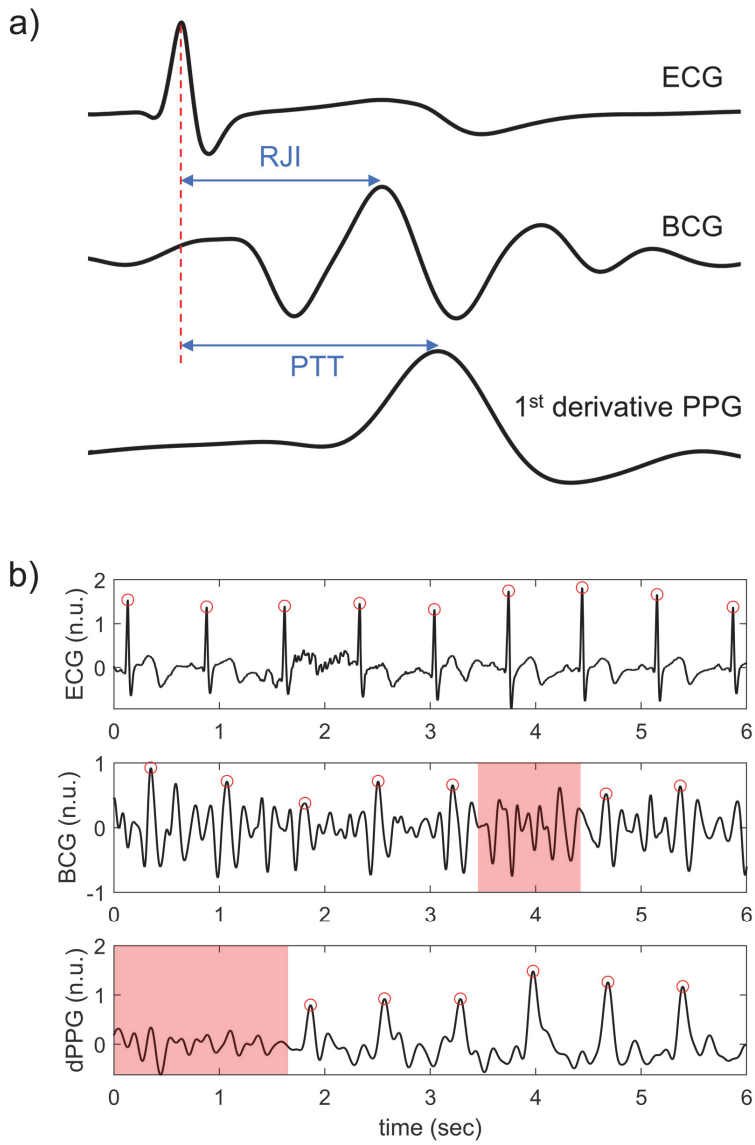

Figure 3-10. a) Example of calculation of PTT and RJI in a cardiac cycle. b) example of excluded peaks. Each red dot denotes a characteristic point, and the red shaded region shows the area where peaks were not detected.

As summarized in Table 3-7, the MLR model showed a lower performance of BP estimation than that for the proposed model. The input parameters of the MLR model were RRI, RJI, and PTT that were calculated from ECG, PPG, and BCG. Although the input signal for the MLR and proposed model were identical, the MAE of the MLR model was statistically higher in both SBP and DBP estimation ($p < 0.05$).

The scatter plots between PTT and SBP of the two cases are presented in Figure 3-11. Although the values of the correlation coefficient (CC) between PTT and SBP were high in both the cases, the R^2 value for the good and inferior cases were 0.29 and 0.05, respectively. The BP variation could not be explained by the interval features alone (such as PTT). However, the proposed model could extract other features as well as the interval feature.

Table 3-7. Comparison between proposed model and MLR model.

	SBP				DBP			
	RMSE (mmHg)	**MAE (mmHg)**	**STD (mmHg)**	**R^2**	**RMSE (mmHg)**	**MAE (mmHg)**	**STD (mmHg)**	**R^2**
Proposed model	5.42	4.06	4.04	0.52	4.30	3.33	3.42	0.49
MLR	6.40	5.19	3.45	0.26	4.75	3.85	2.69	0.22

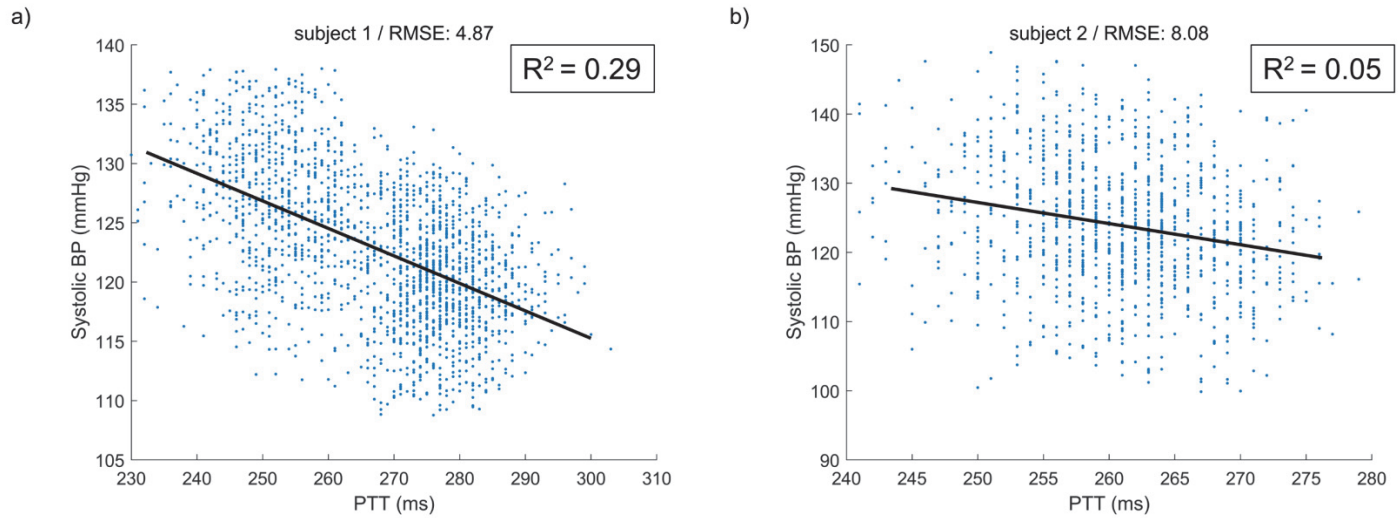

Figure 3-11. Scatter plots between PTT and systolic BP. The black line indicates the fitting line. a) good case; b) inferior case.

3.3. Discussion

3.3.1. Interpretation of result using global standards

Two global standards of BP monitoring are used as performance indicators in most studies: the Association for the Advancement of the Medical Instrumentation (AAMI) standard [39] and the British Hypertension Society (BHS) standard [40]. Table 3-8 and Table 3-9 show the results of using the proposed model with reference to these standards. The study population of this study was 15, which is insufficient because the AAMI and BHS standards require at least 85 subjects. Nonetheless, the results complied with the AAMI and satisfied BHS standards in Grade A for both SBP and DBP.

Table 3-8. Performance comparison with AAMI standard.

		ME (mmHg)	STD (mmHg)
AAMI standard		≤ 5	≤ 8
Proposed model	**SBP**	-0.20	5.83
	DBP	-0.02	4.91

Table 3-9. Performance comparison with BHS standard.

		Cumulative absolute error percentage			Grade
		≤ 5 mmHg	≤ 10 mmHg	≤ 15 mmHg	
BHS standard		60%	85%	95%	A
		50%	75%	90%	B
		40%	65%	85%	C
			Worse than C		D
Proposed model	**SBP**	**73%**	**93%**	**98%**	**A**
	DBP	**80%**	**96%**	**99%**	**A**

3.3.2. Comparison with related works

The performances of the models used in related studies are shown in Table 3-10. Chan et al. proposed a linear regression model based on PTT [14]. Although the study used mean error, the error was high and the data size unspecified. Kachuee et al. [15], Kurylyak et al. [17], Su et al. [16], and Wang et al. [18] also proposed feature-based approaches. The studies recommended various methods to extract the features, which require a substantial amount of time. Meanwhile, the proposed model used only bandpass filtering and did not require a complex feature extraction process. Lee et al. proposed a BP estimation model using only two channels of BCGs [11]. Although BP estimation using only BCG is effective, the evaluation was conducted with a one-time BP measurement rather than a continuous one. Tanveer et al. achieved remarkable performance using raw signals from ECG and PPG [20]. However, the window length to estimate BP was excessively long (16 s), whereas the window length of this study was 5 s. In addition, the authors observed that the performance of their proposed model was highly dependent on an appropriate division of ECG and PPG cycles. This implies that additional work such as peak detection is required and that the signal quality affects the performance of the model. Slapničar et al. also proposed a deep learning model using raw signals of PPG [19]. However, its performance was low compared to those of other studies. In addition, it is difficult to evaluate the model accurately because other evaluation metrics were not presented.

Table 3-10. Performance comparison with related works.

Author	Data size	Model	Input		SBP error (mmHg)	DBP error (mmHg)
			Inputs	Signal		
Chan [14]	Unspecified	Linear regression	Features (PTT)	ECG PPG	ME: 7.49 STD: 8.82	ME: 4.08 STD: 5.62
Kachuee [15]	1,000 subjects 10 min (MIMIC III)	AdaBoost	Features	PPG	MAE: 8.21 STD: 5.45	MAE: 4.31 STD: 5.62
Kurylyak [17]	15,000 heartbeats	Deep learning (ANN)	Features	PPG	ME: 3.80 STD: 3.46	ME: 2.21 STD: 2.09
Lee [11]	30 subjects	Deep learning (ANN)	Features (IPD)	BCG	ME: 0.01 STD: 6.75	ME: 0.05 STD: 5.83
Slapničar [19]	510 subjects 700 hours (MIMIC III)	Deep learning (ResNet)	Raw	PPG	MAE: 9.43	MAE: 6.88
Su [16]	84 subjects 10 min	Deep learning (RNN)	Features	ECG PPG	RMSE: 3.73	RMSE: 2.43
Tanveer [20]	39 subjects (MIMIC I)	Deep learning (ANN+LSTM)	Raw	ECG PPG	RMSE: 1.27 MAE: 0.93	RMSE: 0.73 MAE: 0.52
Wang [18]	58,795 intervals of PPG (MIMIC I)	Deep learning (ANN)	Features	PPG	MAE: 4.02 STD: 2.79	MAE: 2.27 STD: 1.82
This study	15 subjects 30 min	Deep learning (CNN+Bi-GRU +Attention)	Raw	ECG PPG BCG	MAE: 4.06 STD: 4.04	MAE: 3.33 STD: 3.42

Chapter 4. Feature-based generalized BP estimation model based on LSTM Network

4.1. Methods

An overview of the proposed approach is depicted in Figure 4-1. First, measured data was bandpass filtered to remove baseline noise, and the characteristic peaks of the signal were detected. Thereafter, the features were extracted from the characteristic points. The feature sequence vector was used as an input to the Bi-LSTM model, and SBP and DBP were regressed. The methods are demonstrated in the following subsections.

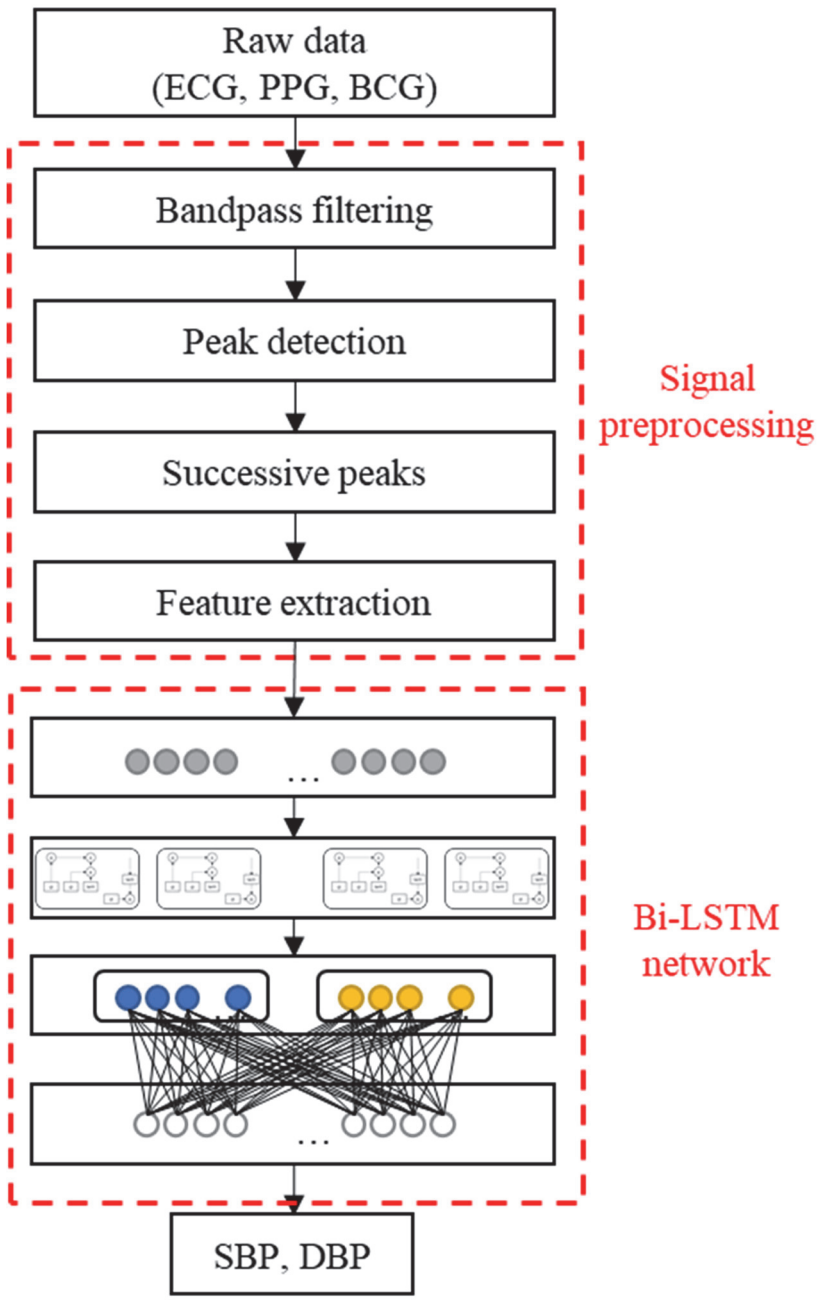

Figure 4-1. Overview of proposed approach.

4.1.1. Data Acquisition

A total of 18 subjects (male: 8, female: 10) with no reported medical conditions were recruited for the experiment. Written informed consent was obtained from the subjects, and the study was approved by the Institutional Review Board of Seoul National University Hospital (IRB No. 1801-016-912).

Several devices were attached to the subject to measure the physiological signals. Three Ag/AgCl electrodes were attached to the subject according to Einthoven's triangle, and the ECG was acquired on Lead II with the BIOPAC ECG100C module. The PPG was measured from the index finger of the subject using a commercial module (PSL-iPPG2C), whereas the BCG signal was measured from the PVDF sensor installed on the chair seat. The reference SBP and DBP were measured with a continuous BP monitoring device (Finometer® PRO, Finapres Medical Systems). Once the devices were attached, the subject was asked to sit on the chair with the PVDF sensor, and the signal was recorded for 30 min. All the data were synchronized and digitized at 1000 Hz using a data acquisition device (BIOPAC MP150). Furthermore, 15 subjects visited again after one to two weeks, and the measurement procedure was repeated with an identical experimental setup. After the measurement was completed, personal information (age, gender, height, weight, body mass index (BMI)) was measured (where required) and recorded. The information is summarized in Table 4-1.

Table 4-1. Summary of personal information of participants.

	Age	**Height (cm)**	**Weight (kg)**	**BMI (kg/m^2)**
Mean	25.7	167.1	60.9	21.5
STD	3.1	9.3	14.9	3.2

4.1.2. Signal preprocessing and feature extraction

A second-order Butterworth filter was applied to the signal to remove baseline wandering and power-line noise (ECG: 0.5 to 35 Hz; BCG: 4 to 15 Hz; PPG: 0.5 to 8 Hz). The characteristic points from the ECG, BCG, and PPG were used to extract the features. First, the R-peak of the ECG was detected using the Pan–Tompkins algorithm. The J-peak of the BCG was detected by identifying the highest peak between 110 and 250 ms after each R-peak. The PPG peak was detected after the signal was differentiated. Once the false-positive peaks had been excluded manually, the features were extracted. Thereafter, the features were standardized with the mean and standard deviation values to be used as input for the neural network model. The features are listed in Table 4-2, and the feature extraction method is depicted in Figure 4-2. The features from 10 cardiac cycles were regarded as a sequence. The number of cardiac cycles was determined empirically. The sequences that included undetected peaks were excluded. The BP value immediately after the final peak in the sequence was used as a reference. The SBP and DBP distributions are illustrated in Figure 4-3. The average values are 111.2 and 67.7 mmHg for SBP and DBP, respectively.

Table 4-2. List of features used as inputs in the proposed model.

Feature	Description
RRI	ECG R peak to R peak interval
PTT	ECG R peak to PPG peak interval
RJI	ECG R peak to BCG J peak interval
IPI	BCG I notch to PPG peak interval
ECGamp	Amplitude of ECG R peak
BCGamp	Amplitude of BCG J peak
PPGamp	Amplitude of PPG peak interval

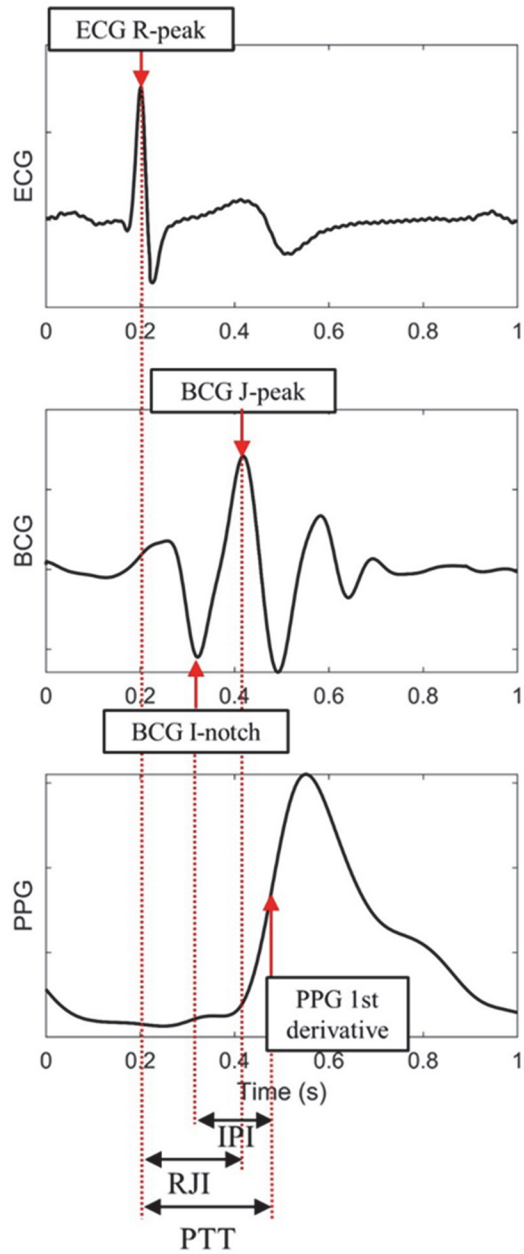

Figure 4-2. Feature extraction from characteristic points of three signals: ECG, BCG, and PPG.

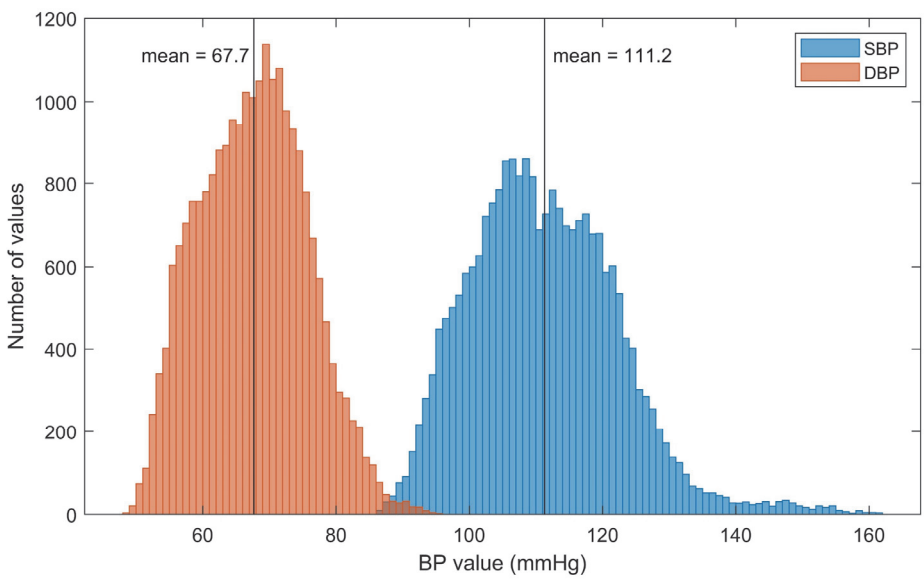

Figure 4-3. BP distributions in the data.

4.1.3. Deep learning architecture

The proposed network architecture is summarized in Figure 4-4. Because 10 cardiac cycles with 7 features were used as the input, the shape of the input layer is 10 × 7. The number of hidden nodes of the Bi-LSTM network is empirically set to 128, and a total of 256 features were generated at each timestep. Tanh was used as the activation function of the LSTM layer. The outputs of the forward and backward LSTM cells were concatenated (10 × 256 neurons) and transformed into a one-dimensional flattened layer (1 × 2560 neurons) for connecting to the fully connected layer after the LSTM layer. In the case of the general model, personal information including gender, age, height, weight, and body mass index (BMI) was included in the first fully connected layer. The second fully connected layer was used for the BP regression. ReLU and a linear activation function were utilized at the first and second fully connected layer, respectively. The number of hidden nodes in the first fully connected layer was set to 64.

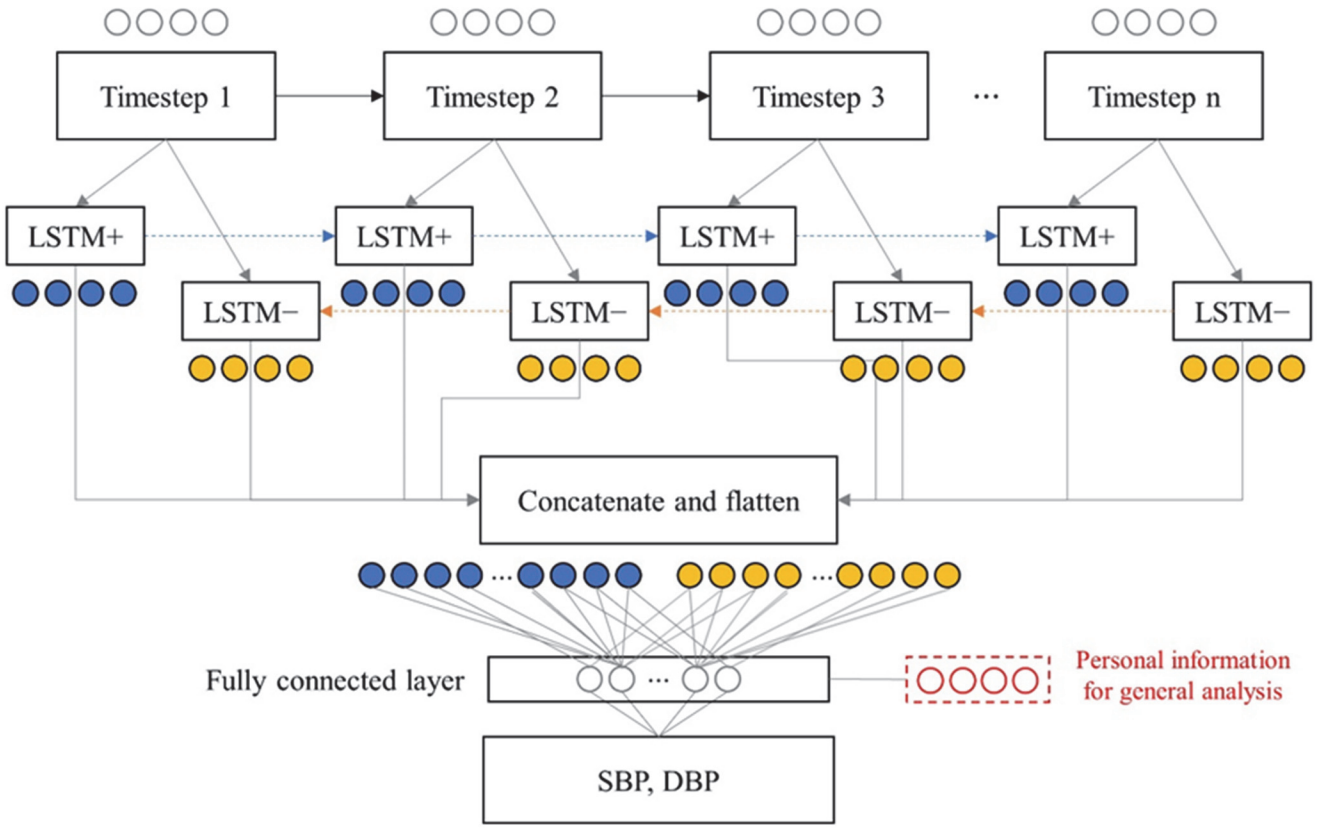

Figure 4-4. Diagram of proposed model architecture.

4.1.4. Experimental setup

The deep learning model was implemented in the Keras framework with a TensorFlow backend. The data were shuffled and randomly selected to train the deep learning model. In total, 60% of the data was used for training, 20% was used for validation, and 20% was used for testing. The Adam optimizer was used to optimize the model with a learning rate of 10^{-3}. The initial value was randomly determined, and the MSE was selected as the loss function. To overcome the overfitting issue, a regularization method was applied with the dropout mask on 10% of the connection in the LSTM layer. The model was trained with the early stopping method with a patience of 10 in a maximum of 100 epochs. The batch size was set to 64.

Following the training process, the test set was used to estimate the SBP and DBP. The model was trained three times with different random initial values, and the regression result was averaged. The CC, MAE, and RMSE between the estimated and reference BPs were calculated to evaluate the performance of the algorithm.

4.2. Results

4.2.1. Feature analysis

The performance of the model was evaluated with different inputs. The features that were used as inputs are described in Table 4-3. As indicated in Table 4-4, the model with all the three signals exhibits performance higher than those of the other models. The MAEs are 2.62 and 2.03 mmHg and CCs are 0.77 and 0.76, for the SBP and DBP estimations, respectively. The difference between the models is statistically significant ($p < 0.01$).

Table 4-3. Feature list of different inputs.

Inputs	Features
ECG	RRI, ECGamp
ECG, BCG	RRI, RJI, ECGamp, BCGamp
ECG, PPG	RRI, PTT, ECGamp, PPGamp
ECG, PPG, BCG	RRI, PTT, RJI, IPI, ECGamp, BCGamp, PPGamp

Table 4-4. Mean values of MAE, RMSE, and CC for different inputs of the personalized model.

Inputs	SBP			DBP		
	MAE	**RMSE**	**CC**	**MAE**	**RMSE**	**CC**
ECG	3.81	4.75	0.50	2.70	3.40	0.51
ECG, BCG	3.50	4.42	0.59	2.51	3.15	0.62
ECG, PPG	2.84	3.57	0.74	2.29	2.88	0.70
ECG, PPG, BCG	**2.62**	**3.36**	**0.77**	**2.03**	**2.57**	**0.76**

4.2.2. General model analysis

Leave-one-subject-out (LOSO) analysis was performed to develop a general model. The data of one subject were removed from the training set and those of the other subjects were used as input to train the model. Moreover, a fine-tuning approach was applied after each training run. The weight in the Bi-LSTM layer was not trained, and the fully connected layer was trained with 20% of the data of the excluded subject.

The results are summarized in Table 4-5. The MAE values of the LOSO model are 10.01 and 5.64 mmHg for the SBP and DBP, respectively. This model exhibited a higher error than that of the personal model ($p < 0.01$). The tuned LOSO model yielded MAE values of 2.56 and 2.06 mmHg for the SBP and DBP, respectively. It exhibited a marginally lower error than that for the personal model in the case of SBP and a higher error in the case of DBP. The difference was not statistically significant.

Table 4-5. Mean values of MAE, RMSE, and CC for each model.

Model	SBP			DBP		
	MAE	**RMSE**	**CC**	**MAE**	**RMSE**	**CC**
Personal	2.62	3.36	0.77	2.03	2.57	0.76
LOSO	10.01	11.26	0.40	5.64	6.52	0.40
Tuned LOSO	2.56	3.25	0.80	2.05	2.61	0.76

A comparison of the personal and tuned LOSO models is illustrated in Figure 4-5. Although the difference is not statistically significant, the tuned LOSO model exhibits higher performance than that of the personal model when the reference BP is exceptionally high or low. Moreover, the tuned LOSO model requires fewer parameters to be trained, thereby the model can be trained with a smaller amount of data. This, in turn, results in lesser time-consumption.

The Bland–Altman plot is illustrated in Figure 4-6. The bias is not significant in any of the three models, and the limits of agreement at a 95%-confidence interval of the tuned LOSO model are [-6.08, 6.26] and [-4.87, 5.00] for SBP and DBP, respectively.

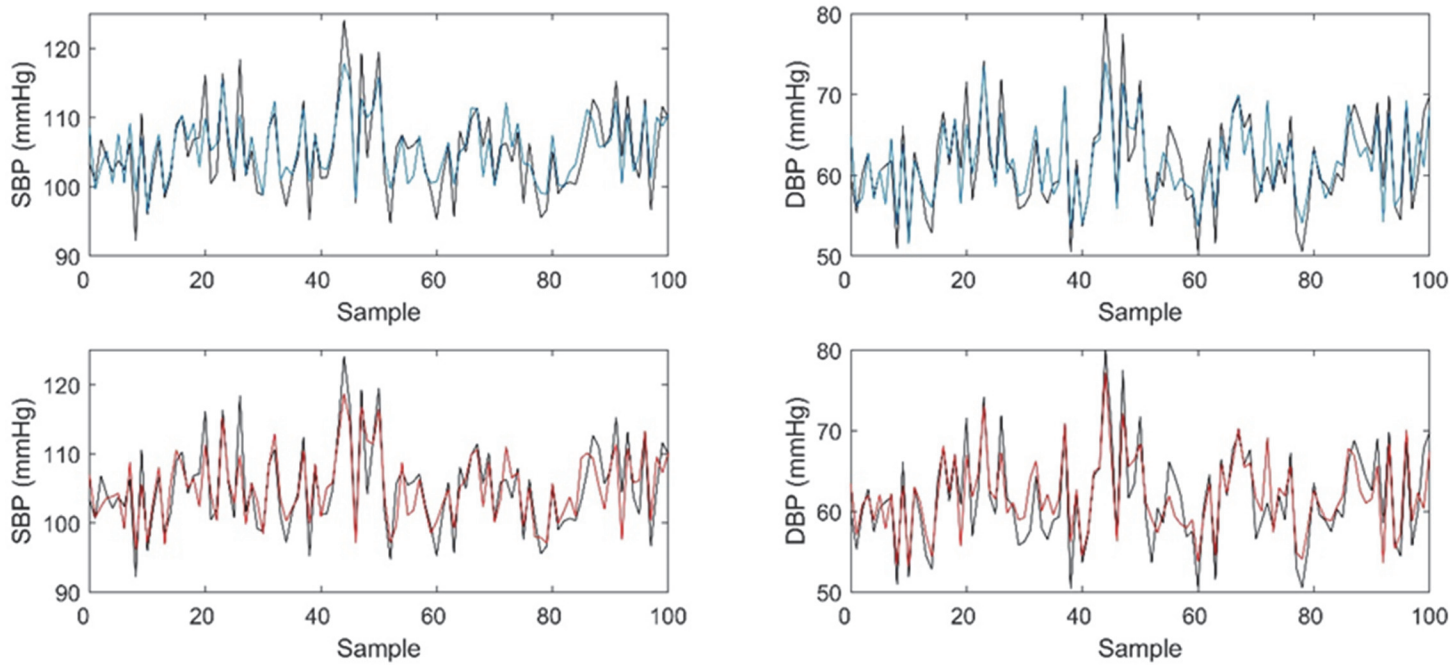

Figure 4-5. Comparison of reference and estimated BP (black line: reference BP; blue line: personal model; red line: tuned LOSO model).

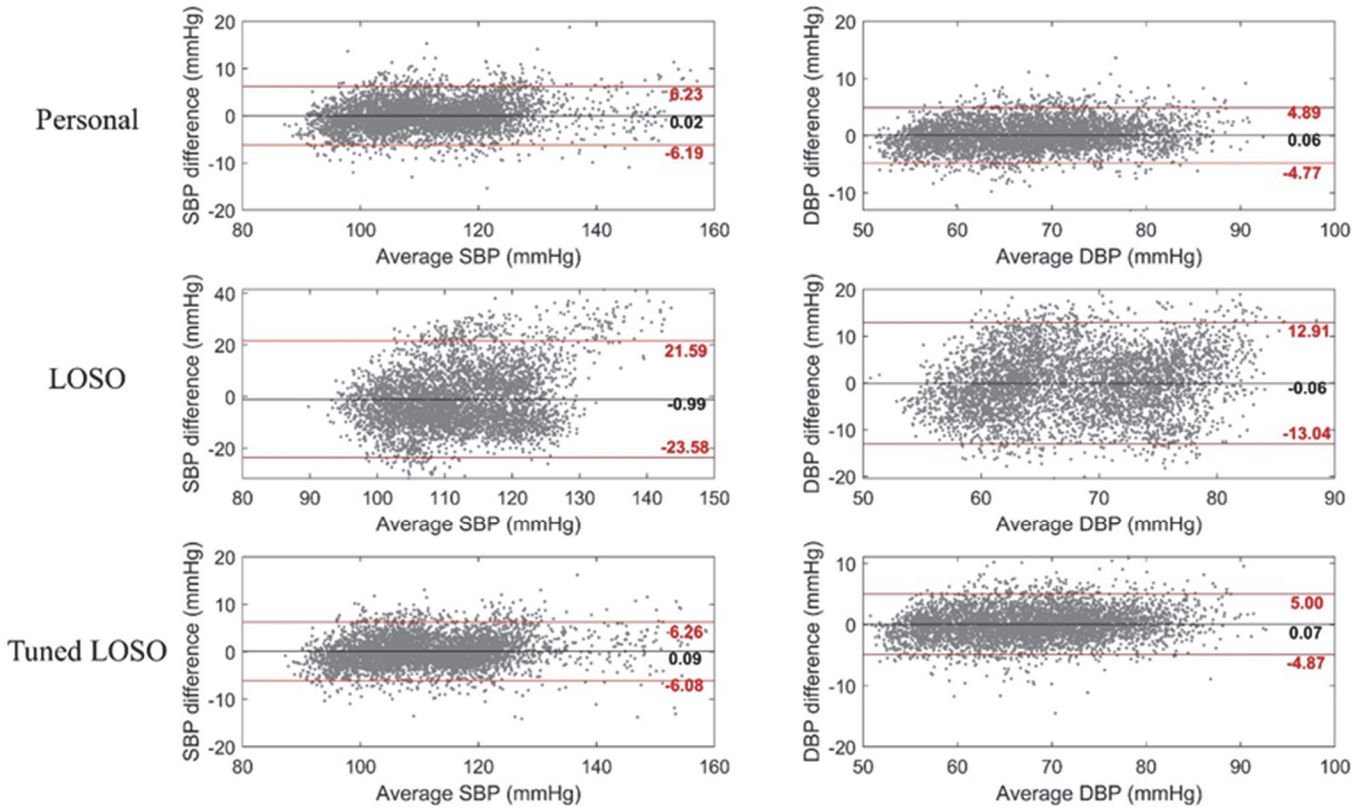

Figure 4-6. Bland–Altman plots of models. Black line: mean value; red line: 95% confidence interval.

4.2.3. Reproducibility analysis

The model reproducibility was investigated to evaluate the model generalization. A multi-day test was performed using second-visit data in addition to a one-day test. The model was trained with the data of the one visit, and the data of the other visit was used as a test set. The results are presented in Table 4-6. The error is higher than that of the test results with only the first visit, for the personal and tuned LOSO models. The MAE values of the tuned LOSO model are 5.82 and 5.24 mmHg for the SBP and DBP estimation, respectively. Although the difference between the personal and tuned LOSO models presented in Section 4.2.2 is not significant, the reproducibility of the tuned LOSO model is higher than that of the personal model ($p < 0.05$). This is because the personal model was overfitted with the one-day condition of the subject.

Table 4-6. Mean MAE, RMSE, and CC values in the multi-day test analysis.

train	test	model	SBP			DBP		
			MAE (mmHg)	**RMSE (mmHg)**	**CC**	**MAE (mmHg)**	**RMSE (mmHg)**	**CC**
visit #1	visit #2	personal	7.12	8.99	0.41	6.22	7.61	0.38
		LOSO	10.23	11.49	0.41	5.94	6.81	0.41
		tuned LOSO	**5.81**	**6.78**	**0.53**	**5.34**	**6.14**	**0.51**
visit #2	visit #1	personal	6.23	7.60	0.45	5.20	6.27	0.41
		LOSO	10.82	11.97	0.45	6.17	7.06	0.41
		tuned LOSO	**5.84**	**6.85**	**0.52**	**5.14**	**5.97**	**0.49**
Total		personal	6.67	8.29	0.43	5.71	6.94	0.40
		LOSO	10.52	11.73	0.43	6.06	6.94	0.41
		tuned LOSO	**5.82**	**6.82**	**0.53**	**5.24**	**6.06**	**0.50**

Scatter plots for the model results are presented in Figure 4-7. The R^2 values are presented therein. The personal model tends to underestimate the BP, and the R^2 values are 0.51 and 0.40 for SBP and DBP, respectively. The R^2 values of the tuned LOSO models are 0.63 and 0.49 for SBP and DBP, respectively. This implies that the model learned more general patterns to estimate the BP than the personal model did, with high reproducibility.

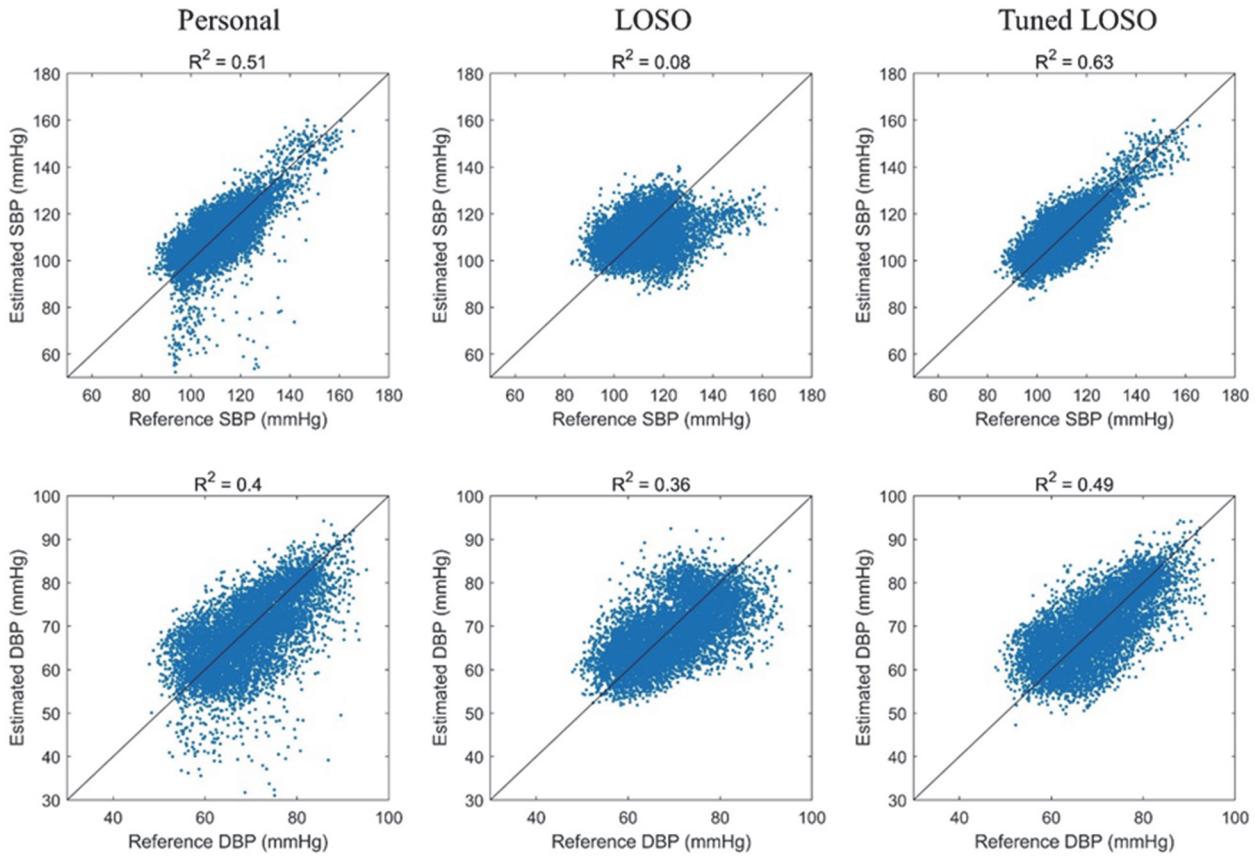

Figure 4-7. Scatter plots of reproducibility analysis.

4.3. Discussion

4.3.1. Evaluation using global standards

The proposed model of one-day and multi-day tests was evaluated using two international standards of BP estimation: the British Hypertension Society (BHS) standard [40] and the Association for the Advancement of Medical Instrumentation (AAMI) standard. The evaluation results are presented in Table 4-7. The BHS standard evaluates the BP estimation device based on the cumulative percentage of absolute errors under thresholds of 5, 10, and 15 mmHg. According to the BHS standard, the proposed model is consistent with Grade A in the one-day test and Grade B in the multi-day test, for SBP and DBP.

The evaluation results using the AAMI standard are described in Table 4-8. The AAMI standard requires mean error (ME) values lower than 5 mmHg and STD values lower than 8 mmHg. According to the AAMI, the number of populations should be at least 85. Although this study did not satisfy the population criterion, both the models satisfied the ME and STD values in the SBP and DBP estimation.

Table 4-7. Performance evaluation using BHS standard.

		Cumulative absolute error percentage			Grade
		≤ 5 mmHg	≤ 10 mmHg	≤ 15 mmHg	
BHS standard		60%	85%	95%	A
		50%	75%	90%	B
		40%	65%	85%	C
			Worse than C		D
Proposed model (one-day test)	**SBP**	**89.3%**	**99.4%**	**100.0%**	**A**
	DBP	**94.7%**	**99.8%**	**100.0%**	**A**
Proposed model (multi-day test)	**SBP**	**51.6%**	**81.4%**	**96.3%**	**B**
	DBP	**56.1%**	**87.9%**	**98.3%**	**B**

Table 4-8. Performance evaluation using AAMI standard.

		ME (mmHg)	STD (mmHg)
AAMI standard		≤ 5	≤ 8
Proposed model (one-day test)	**SBP**	-0.09	3.15
	DBP	-0.07	2.52
Proposed model (multi-day test)	**SBP**	-0.07	7.30
	DBP	-0.17	6.4

4.3.2. Average BP analysis

The model was based on beat-to-beat BP estimation, which implies that the model can express the high frequency variation of the BP. Most of the reference BP values are based on the measurement of the average BP. The average BP analysis is also discussed in this section.

The reference BP was averaged using moving average filter. Two numbers of points (5 and 10) were selected to perform averaging. Figure 4-8 shows a sample of BP with moving average filter. As the number of points of the moving average increases, the high frequency variation of the BP is attenuated.

The results are illustrated in Figure 4-9. The error in average BP estimation is lower than that in beat-to-beat BP estimation. The MAE values are 5.72 and 4.95 mmHg for SBP and DBP, respectively. In addition, the error is smaller when a larger number of points are used to perform average BP estimation. The MAE values are 5.54 and 4.84 mmHg for SBP and DBP estimation, respectively. The difference is not significant for SBP estimation, whereas it is so for DBP estimation.

In many health-care applications, BP is classified using numerical values. According to the guideline for BP classification for adults, BP can be classified into four: normal, prehypertension, Stage 1 hypertension, and Stage 2 hypertension [41]. The performance in terms of hypertension classification was evaluated using the average BP estimation model. The result is shown in Table 4-9.

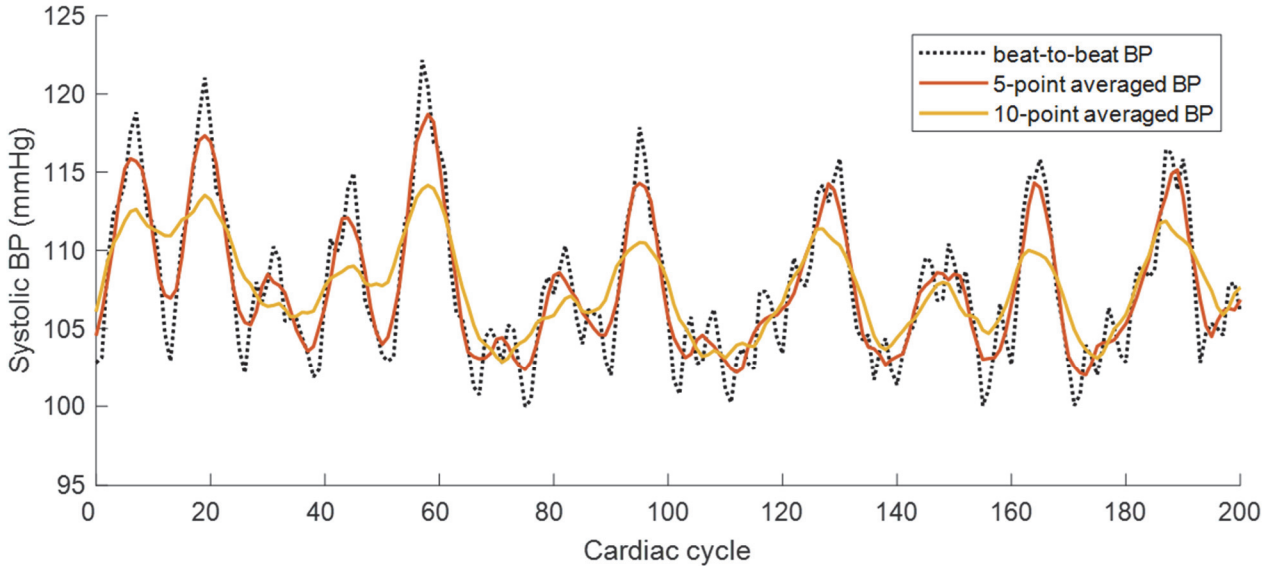

Figure 4-8. Sample of BP with moving average filter.

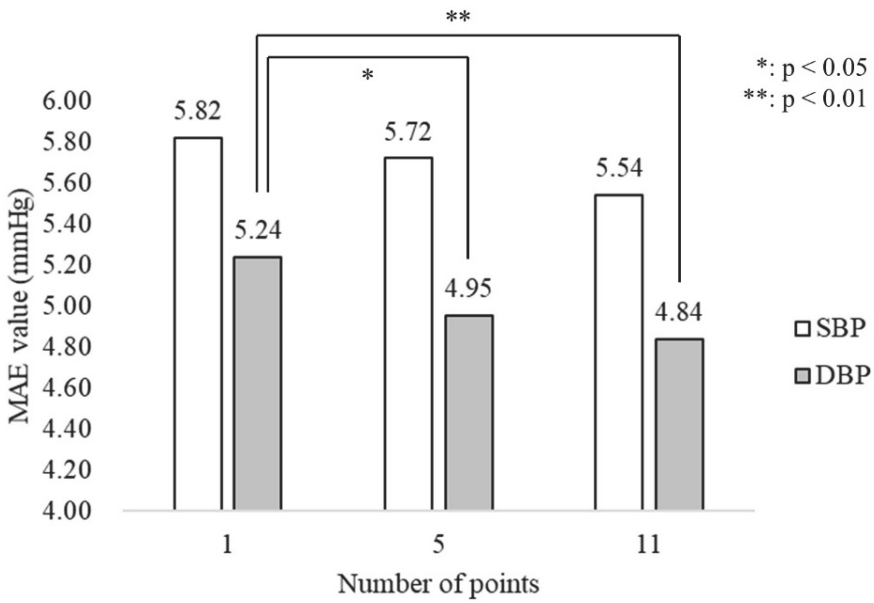

Figure 4-9. Result of average BP analysis.

Table 4-9. Accuracy of hypertension classification for SBP and DBP.

BP Class	SBP		DBP	
	Range (mmHg)	Accuracy	Range (mmHg)	Accuracy
Normal	BP < 120	84%	BP < 80	90%
Prehypertension	120 ≤ BP < 130	82%		
Stage 1 Hypertension	130 ≤ BP < 140	97%	80 ≤ BP < 90	89%
Stage 2 Hypertension	BP ≥ 140	98%	BP ≥ 90	100%
Total		**81%**		**89%**

4.3.3. Comparison with related works

To evaluate the performance of the proposed method, the algorithm was compared with the representative BP estimation methods. Chen et al. recommended the use of an SBP estimation equation based on linear regression with PTT [6]:

$$SBP = SBP_0 - \frac{2}{\gamma PTT_0}(PTT - PTT_0)$$

In addition, Poon et al. investigated an algorithm with the initially calibrated BP as follows [7]:

$$SBP = MBP_0 + \frac{2}{\gamma}\ln\left(\frac{PTT_0}{PTT}\right) + \frac{2}{3}\cdot PP_0 \cdot \left(\frac{PTT_0}{PTT}\right)^2$$

$$DBP = MBP_0 + \frac{2}{\gamma} ln\left(\frac{PTT_0}{PTT}\right) - \frac{1}{3}\cdot PP_0 \cdot \left(\frac{PTT_0}{PTT}\right)^2$$

Meanwhile, Ding et al. recommended the pulse intensity ratio (PIR) as a reflection of the variation in the arterial diameter. In addition, they estimated BP with PIR and PTT [8]:

$$SBP = DBP_0 \cdot \frac{PIR_0}{PIR} + PP_0 \cdot \left(\frac{PTT_0}{PTT}\right)^2$$

$$DBP = DBP_0 \cdot \frac{PIR_0}{PIR}$$

The model was compared with these three methods. Furthermore, a conventional multiple linear regression (MLR) method with the features that were used in the model was compared with proposed method. The comparison result is presented in

Table 4-10. The PIR model shows the lowest error among the previous methods. The difference between the PIR model and MLR model without BCG features is not significant. However, the performance is improved with the BCG features. The MAE values of the MLR model with BCG features are 4.17 mmHg and 3.12 mmHg for SBP and DBP, respectively. In addition, the use of the features of the previous 10 cardiac cycles yielded a large improvement in BP estimation compared with that of the model with only one cardiac cycle. Proposed model shows the lowest error because the previous cycles are utilized as inputs and the deep learning could obtain a nonlinear expression that related the features and the target BP.

The scatter plots of PTT model 2, the PIR model, the MLR model with BCG features, and proposed model are presented in Figure 4-10. As shown in the plots, when the relationship between PTT and BP is not significant, the PTT-based model cannot yield a meaningful fitting equation. In addition, the DBP estimation in the MLR model is improved with the BCG features, compared with that for the PIR model.

Table 4-10. Comparison with other methods.

Model	Equation	MAE for SBP (mmHg)	MAE for DBP (mmHg)
PTT model 1 [6]	$SBP = SBP_0 - \frac{2}{\gamma PTT_0}(PTT - PTT_0)$	4.32	-
PTT model 2 [7]	$SBP = MBP_0 + \frac{2}{\gamma}\ln\left(\frac{PTT_0}{PTT}\right) + \frac{2}{3}\cdot PP_0 \cdot \left(\frac{PTT_0}{PTT}\right)^2$ $DBP = MBP_0 + \frac{2}{\gamma}\ln\left(\frac{PTT_0}{PTT}\right) - \frac{1}{3}\cdot PP_0 \cdot \left(\frac{PTT_0}{PTT}\right)^2$	4.70	3.28
PIR model [8]	$SBP = DBP_0 \cdot \frac{PIR_0}{PIR} + PP_0 \cdot \left(\frac{PTT_0}{PTT}\right)^2$ $DBP = DBP_0 \cdot \frac{PIR_0}{PIR}$	4.47	3.15
MLR model (w/o BCG features)	$SBP = a_1 + b_1 \cdot PTT + c_1 \cdot RRI$ $DBP = a_2 + b_2 \cdot PTT + c_2 \cdot RRI$	4.25	3.16
MLR model (with BCG features)	$SBP = a_1 + b_1 \cdot PTT + c_1 \cdot RRI + d_1 \cdot RJI + e_1 \cdot IPI$ $DBP = a_2 + b_2 \cdot PTT + c_2 \cdot RRI + d_2 \cdot RJI + e_2 \cdot IPI$	4.17	3.12
MLR model (with previous features)	$SBP = a_1 + \sum_N (b_{1i} \cdot PTT_i + c_{1i} \cdot RRI_i + d_{1i} \cdot RJI_i + e_{1i} \cdot IPI_i)$ $DBP = a_2 + \sum_N (b_{2i} \cdot PTT_i + c_{2i} \cdot RRI_i + d_{2i} \cdot RJI_i + e_{2i} \cdot IPI_i)$	3.71	2.65
Proposed model	**LSTM**	**2.62**	**2.03**

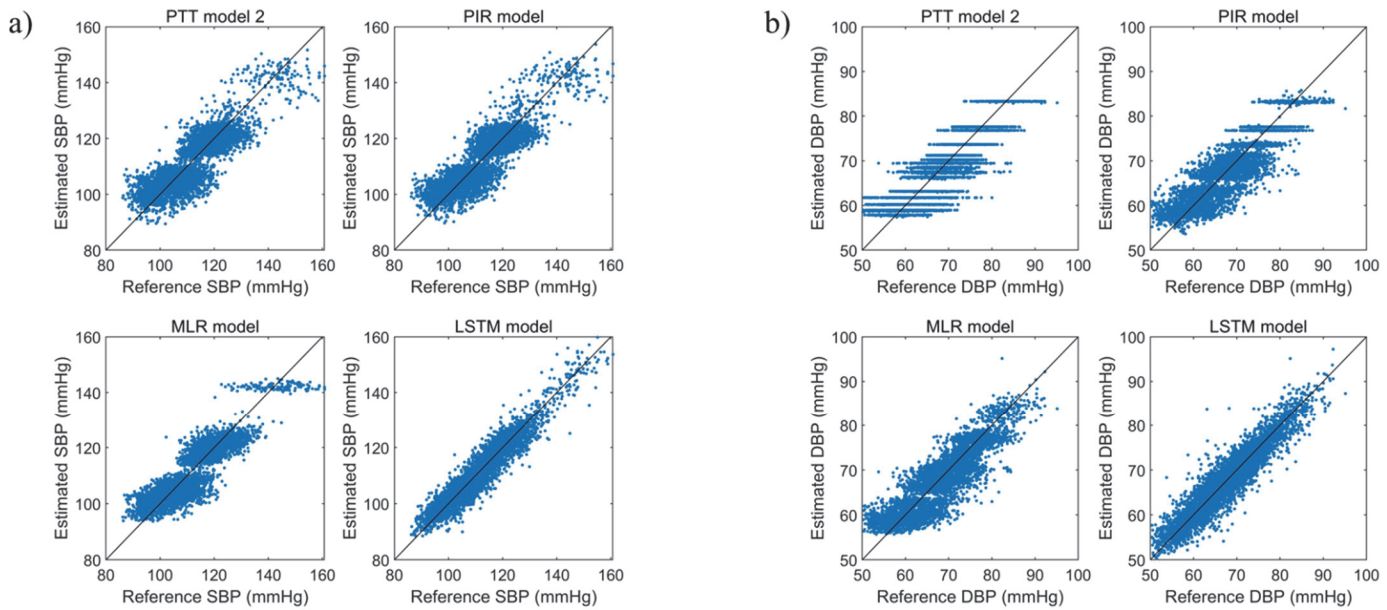

Figure 4-10. Scatter plots of proposed model and the other methods; a) SBP estimation, b) DBP estimation.

In addition, comparisons of the proposed method with related works are summarized in Table 4-11. It was difficult to perform an unbiased comparison with other studies because the datasets used in the studies may have differed significantly. In addition, the validation methods varied. Kachuee et al. [15], Slapničar et al. [19], and Hsu et al. [42] used an online database named “Medical Information Mart for Intensive Care unit (MIMIC)” [43]. This database contains a large number of clinical data including those of ECG, breathing, PPG, and BP. However, the data may not be compatible for normal individuals because these were obtained from patients in intensive care units, and the patients could have been influenced by drugs that could affect BP variation. Kachuee et al. recommended a continuous BP estimation algorithm based on AdaBoost. However, the error was relatively higher than that in other studies [15]. Slapničar et al. implemented a network architecture using ResNet and a spectro-temporal block, and performed LOSO analysis with the data [19]. Hsu et al. [42] and Wu et al. [44] proposed DNN models, in which the error was lower than that in other studies. However, these applied 10-fold cross-validation, which is different from the LOSO analysis. Su et al. proposed a long-term BP prediction model using a Bi-LSTM network [16]. A multi-day analysis was performed on the second and fourth days, and at six months. The MAE values were 5.81 and 5.21 mmHg for SBP and DBP, respectively. Although the error value was lower than that of proposed model, the validation was performed with only a personalized model. Furthermore, The BP

estimation performance was enhanced compared with that of previous work that utilized a CNN and an attention mechanism [12].

Table 4-11. Comparisons with related works.

Author	Dataset	Model	Input	Validation method	SBP error (mmHg)		DBP error (mmHg)	
					MAE	RMSE	MAE	RMSE
Kachuee et al. [15]	N=1000 10 min (MIMIC III)	Adaboost	ECG, PPG features	Personal	8.21		4.31	
				10-fold cross validation	11.17		5.35	
Slapničar et al. [19]	510 subjects 700 hours (MIMIC III)	ResNet	Raw PPG	LOSO	15.41		12.38	
				tuned LOSO	9.43		6.88	
Hsu et al. [42]	N=9000 (MIMIC II)	DNN	PPG features	10-fold cross validation	3.21	4.63	2.23	3.21
Wu et al. [44]	N=85	DNN	ECG, PPG features	10-fold cross validation	3.31	4.60	2.22	3.15
Su et al. [16]	N=84 10 min	bi-LSTM	ECG, PPG features	Personal (one-day)		3.73		2.43
				Personal (multi-day)		5.81		5.21
previous work [12]	N=15 30 min	CNN, bi-GRU, Attention	Raw ECG, PPG, BCG	Personal	4.06	5.42	3.33	4.30
Proposed work	**N=18 30 min**	**bi-LSTM**	**ECG, PPG, BCG features**	**LOSO**	**10.01**	**11.26**	**5.60**	**6.52**
				tuned LOSO	**2.56**	**3.25**	**2.05**	**2.61**
				tuned LOSO (multi-day)	**5.82**	**6.82**	**5.24**	**6.06**

Chapter 5. Discussion

Two types of deep learning networks (waveform-based and feature-based) were discussed in the previous chapters. The comparison of the two networks follows. The waveform-based model discussed in Chapter 3 does not require a feature extraction process. However, the mean training time was long (48 min for a subject) because the CNN has a significantly larger number of parameters. Meanwhile, the feature-based model addressed in Chapter 4 requires feature engineering. However, the model consumed a shorter training time (1 min for a subject), and the performance was better than that of the waveform-based model. This is because the model concentrates on the signal without noise. In addition, the model consumes 20 ms per sequence and therefore, can be applied in real-time BP estimation. However, the coverage was 60 % when the sequence length was 10 because the sequences with undetected peaks were excluded owing to the noise. As shown in Figure 5-1, when the sequence length is one, the MAE values of SBP and DBP are 3.88 mmHg and 3.20 mmHg, respectively. These are marginally higher than those of the waveform-based model. However, the coverage is 81 %.

In addition, the generalized model was discussed in Chapter 4. Although the inter-subject generalized model was designed as the LOSO model, the performance was significantly lower than that of the personalized model. However, the model was improved with a fine-tuning method that uses a smaller number of data. Thereby, it showed performance similar to that of the personalized model.

Moreover, the tuned LOSO model showed performance higher than that of the personalized model in the multi-day test.

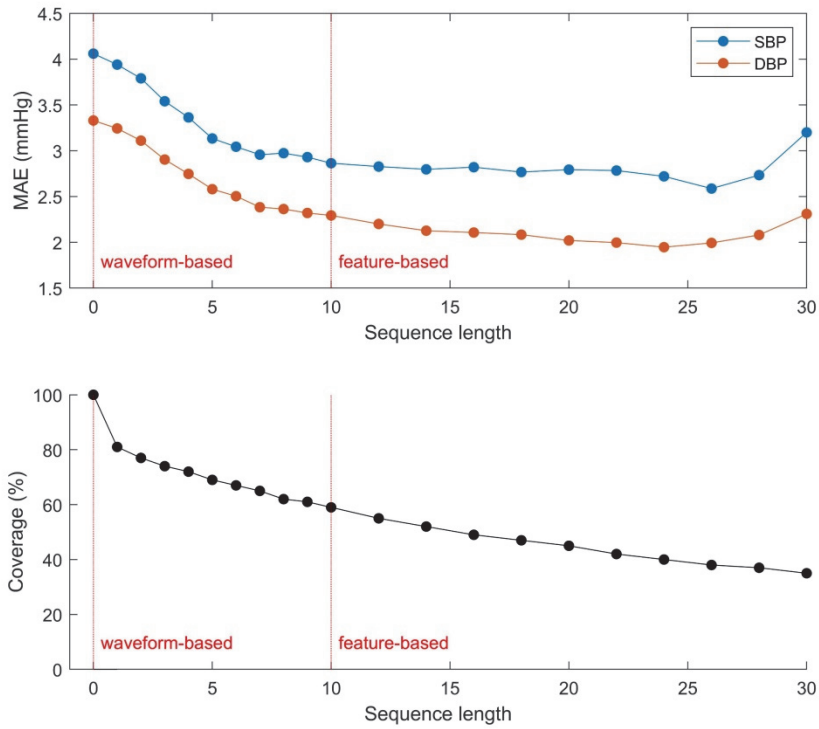

Figure 5-1. MAE values and coverage between different sequence lengths. (Zero sequence length denotes waveform-based model.)

Finally, the limitation of this study is discussed briefly. The data was measured in a short time (30 min). Hence, the performance of the method while predicting long-term BP should be considered. In addition, the data of patients with hypertension were not included in the study. However, approximately 8 % and 2 % of the BP data were in Hypertension stages 1 and 2, respectively, although no subject was diagnosed as a hypertension patient.

Moreover, the number of subjects was small. However, the inputs in the waveform-based model had a length of 5 s. In addition, the total sample size was approximately 27,000,000, which was adequate to evaluate the model. The comparison between feature-based models with different numbers of subjects is presented in Figure 5-2, and the MAE, RMSE, mean CC values are summarized in Table 5-1. The model with 30 subjects was generated with the data of both the visits. This model shows enhanced performance with a larger number of datasets, although it was trained with the same individual as the model with 15 subjects. In further work, the model can be improved with the data of more subjects including patients with hypertension.

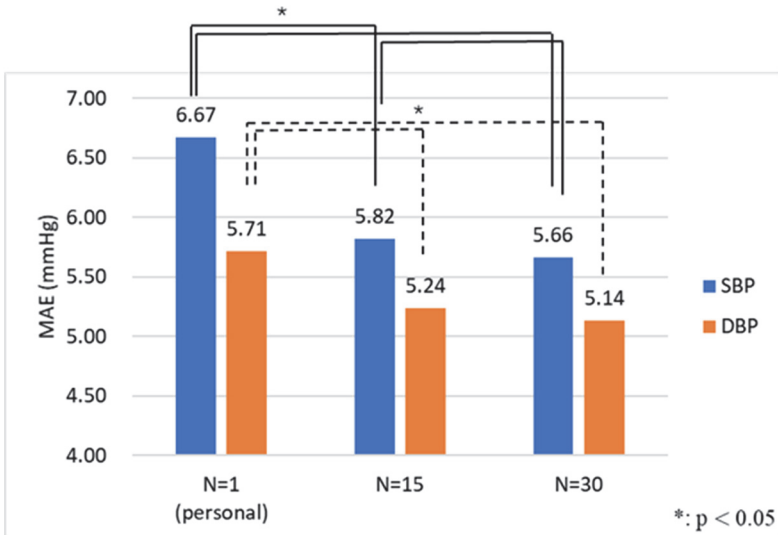

Figure 5-2. Comparison of the model with different numbers of subjects.

Table 5-1. MAE, RMSE, and mean CC values of the models with different numbers of subjects.

	SBP			DBP		
	MAE	**RMSE**	**Mean CC**	**MAE**	**RMSE**	**Mean CC**
N=1 (personal)	6.67	8.29	0.43	5.71	6.94	0.40
N=15	5.82	6.82	0.53	5.24	6.06	0.50
N=30	**5.66**	**6.69**	**0.56**	**5.14**	**5.96**	**0.51**

Chapter 6. Conclusion

In this study, I developed a cuffless BP estimation model based on the deep learning approaches. Two types of deep learning algorithm were applied: waveform-based learning and feature-based algorithm. The results showed that the BP estimation model displayed higher performance with the BCG signal. The waveform-based model was designed with CNN and the attention mechanism. The model had the advantage of being applied on the entire signal without the feature extraction procedure. In addition, the bi-LSTM model was investigated to develop a feature-based BP estimation model. Although the LSTM model could not be applied for noisy signal, it showed performance higher than that of the waveform-based model. In addition, the feature-based model was designed to ensure generality. It was evaluated in terms of reproducibility and showed good performance in a multi-day test.

This study is significant because it attempted to establish both personalized and generalized BP estimation model. The model enables continuous BP monitoring in daily life using unobtrusively measured ECG, PPG, and BCG. The signal used in this study can be measured unobtrusively and therefore, can result in unobtrusive BP monitoring. It can be applied widely with a ubiquitous healthcare platform.

Bibliography

[1] G. Ogedegbe and T. Pickering, "Principles and techniques of blood pressure measurement," *Cardiol Clin,* vol. 28, no. 4, pp. 571-86, Nov 2010.

[2] S. Yoo *et al.*, "Validation of the mobile wireless digital automatic blood pressure monitor using the cuff pressure oscillometric method, for clinical use and self-management, according to international protocols," *Biomed Eng Lett,* vol. 8, no. 4, pp. 399-404, Nov 2018.

[3] D. Wu *et al.*, "Analysis of beat-to-beat blood pressure variability response to the cold pressor test in the offspring of hypertensive and normotensive parents," *Hypertens Res,* vol. 40, no. 6, pp. 581-589, Jun 2017.

[4] X. Ding and Y. T. Zhang, "Pulse transit time technique for cuffless unobtrusive blood pressure measurement: from theory to algorithm," *Biomed Eng Lett,* vol. 9, no. 1, pp. 37-52, Feb 2019.

[5] M. Y. Wong, C. C. Poon, and Y. T. Zhang, "An evaluation of the cuffless blood pressure estimation based on pulse transit time technique: a half year study on normotensive subjects," *Cardiovasc Eng,* vol. 9, no. 1, pp. 32-8, Mar 2009.

[6] W. Chen, T. Kobayashi, S. Ichikawa, Y. Takeuchi, and T. Togawa, "Continuous estimation of systolic blood pressure using the pulse arrival time and intermittent calibration," *Med Biol Eng Comput,* vol. 38, no. 5, pp. 569-74, Sep 2000.

[7] C. Poon and Y. Zhang, "Cuff-less and noninvasive measurements of arterial blood pressure by pulse transit time," in *2005 IEEE engineering in medicine and biology 27th annual conference*, 2006, pp. 5877-5880: IEEE.

[8] X. R. Ding, Y. T. Zhang, J. Liu, W. X. Dai, and H. K. Tsang, "Continuous Cuffless Blood Pressure Estimation Using Pulse Transit Time and Photoplethysmogram Intensity Ratio," *IEEE Trans Biomed Eng,* vol. 63, no. 5, pp. 964-972, May 2016.

[9] R. A. Payne, C. N. Symeonides, D. J. Webb, and S. R. Maxwell, "Pulse transit time measured from the ECG: an unreliable marker of beat-to-beat blood pressure," *J Appl Physiol (1985),* vol. 100, no. 1, pp. 136-41, Jan 2006.

[10] J. H. Shin, K. M. Lee, and K. S. Park, "Non-constrained monitoring of systolic blood pressure on a weighing scale," *Physiol Meas,* vol. 30, no. 7, pp. 679-93, Jul 2009.

[11] K. J. Lee, J. Roh, D. Cho, J. Hyeong, and S. Kim, "A Chair-Based Unconstrained/Nonintrusive Cuffless Blood Pressure Monitoring System Using a Two-Channel Ballistocardiogram," *Sensors (Basel),* vol. 19, no. 3, Jan 31 2019.

[12] H. Eom *et al.*, "End-to-End Deep Learning Architecture for Continuous Blood Pressure Estimation Using Attention Mechanism," *Sensors (Basel),* vol. 20, no. 8, Apr 20 2020.

[13] D. Lee *et al.*, "Beat-to-Beat Continuous Blood Pressure Estimation Using Bidirectional Long Short-Term Memory Network," *Sensors (Basel),* vol. 21, no. 1, Dec 25 2020.

[14] K. Chan, K. Hung, and Y. Zhang, "Noninvasive and cuffless measurements of blood pressure for telemedicine," in *2001 Conference Proceedings of the 23rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2001, vol. 4, pp. 3592-3593: IEEE.

[15] M. Kachuee, M. M. Kiani, H. Mohammadzade, and M. Shabany, "Cuffless Blood Pressure Estimation Algorithms for Continuous Health-Care Monitoring," *IEEE Trans Biomed Eng,* vol. 64, no. 4, pp. 859-869, Apr 2017.

[16] 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)*, 2018, pp. 323-328: IEEE.

[17] Y. Kurylyak, F. Lamonaca, and D. Grimaldi, "A Neural Network-based method for continuous blood pressure estimation from a PPG signal," in *2013 IEEE International instrumentation and measurement technology conference (I2MTC)*, 2013, pp. 280-283: IEEE.

[18] L. Wang, W. Zhou, Y. Xing, and X. Zhou, "A novel neural network model for blood pressure estimation using photoplethesmography without electrocardiogram," *Journal of healthcare engineering,* vol. 2018, 2018.

[19] G. Slapničar, N. Mlakar, and M. Lustrek, "Blood Pressure Estimation from Photoplethysmogram Using a Spectro-Temporal Deep Neural Network," *Sensors (Basel),* vol. 19, no. 15, Aug 4 2019.

[20] M. S. Tanveer and M. K. Hasan, "Cuffless blood pressure estimation from electrocardiogram and photoplethysmogram using waveform based ANN-LSTM network," *Biomedical Signal Processing and Control,* vol. 51, pp. 382-392, 2019.

[21] F. Rundo, S. Conoci, A. Ortis, and S. Battiato, "An Advanced Bio-Inspired PhotoPlethysmoGraphy (PPG) and ECG Pattern Recognition System for Medical Assessment," *Sensors (Basel),* vol. 18, no. 2, Jan 30 2018.

[22] O. Yildirim, P. Plawiak, R. S. Tan, and U. R. Acharya, "Arrhythmia detection using deep convolutional neural network with long duration ECG signals," *Comput Biol Med,* vol. 102, pp. 411-420, Nov 1 2018.

[23] I. Ullah, M. Hussain, and H. Aboalsamh, "An automated system for epilepsy detection using EEG brain signals based on deep learning approach," *Expert Systems with Applications,* vol. 107, pp. 61-71, 2018.

[24] D. Dey, S. Chaudhuri, and S. Munshi, "Obstructive sleep apnoea detection using convolutional neural network based deep learning framework," *Biomed Eng Lett,* vol. 8, no. 1, pp. 95-100, Feb 2018.

[25] Y. Bengio, P. Simard, and P. Frasconi, "Learning long-term dependencies with gradient descent is difficult," *IEEE Trans Neural Netw,* vol. 5, no. 2, pp. 157-66, 1994.

[26] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Comput,* vol. 9, no. 8, pp. 1735-80, Nov 15 1997.

[27] J. Chung, C. Gulcehre, K. Cho, and Y. J. a. p. a. Bengio, "Empirical evaluation of gated recurrent neural networks on sequence modeling," 2014.

[28] K. Xu *et al.*, "Show, attend and tell: Neural image caption generation with visual attention," in *International conference on machine learning*, 2015, pp. 2048-2057.

[29] M.-T. Luong, H. Pham, and C. D. J. a. p. a. Manning, "Effective approaches to attention-based neural machine translation," 2015.

[30] H. Song, D. Rajan, J. J. Thiagarajan, and A. J. a. p. a. Spanias, "Attend and diagnose: Clinical time series analysis using attention models," 2017.

[31] A. Vaswani *et al.*, "Attention is all you need," in *Advances in neural information processing systems*, 2017, pp. 5998-6008.

[32] S. Chaudhari, G. Polatkan, R. Ramanath, and V. J. a. p. a. Mithal, "An attentive survey of attention models," 2019.

[33] C. Raffel and D. P. J. a. p. a. Ellis, "Feed-forward networks with attention can solve some long-term memory problems," 2015.

[34] Physiolab, Busan, Korea. Available online: http://www.physiolab.co.kr

[35] K. Simonyan and A. J. a. p. a. Zisserman, "Very deep convolutional networks for large-scale image recognition," 2014.

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

[37] D. G. Altman and J. M. Bland, "Measurement in Medicine - the Analysis of Method Comparison Studies," (in English), *Journal of the Royal Statistical Society Series D-the Statistician,* vol. 32, no. 3, pp. 307-317, 1983.

[38] J. Pan and W. J. Tompkins, "A real-time QRS detection algorithm," *IEEE Trans Biomed Eng,* vol. 32, no. 3, pp. 230-6, Mar 1985.

[39] A. f. t. A. o. M. I. J. A. A. S. 10-, "American national standards for electronic or automated sphygmomanometers," 1987.

[40] E. O'Brien *et al.*, "The British Hypertension Society protocol for the evaluation of automated and semi-automated blood pressure measuring devices with special reference to ambulatory systems," *J Hypertens,* vol. 8, no. 7, pp. 607-19, Jul 1990.

[41] P. K. Whelton and R. M. Carey, "The 2017 American College of Cardiology/American Heart Association Clinical Practice Guideline for High Blood Pressure in Adults," *JAMA Cardiol,* vol. 3, no. 4, pp. 352-353, Apr 1 2018.

[42] Y. C. Hsu, Y. H. Li, C. C. Chang, and L. N. Harfiya, "Generalized Deep Neural Network Model for Cuffless Blood Pressure Estimation with Photoplethysmogram Signal Only," *Sensors (Basel),* vol. 20, no. 19, Oct 4 2020.

[43] A. E. W. Johnson *et al.*, "MIMIC-III, a freely accessible critical care database," *Scientific Data,* vol. 3, no. 1, 2016.

[44] D. Wu *et al.*, "Continuous cuff-less blood pressure estimation based on combined information using deep learning approach," vol. 8, no. 6, pp. 1290-1299, 2018.

초 록

혈압은 환자의 심장과 심혈관계에 관한 기본적인 건강 정보를 알려주는 중요한 활력 징후 중에 하나이다. 세계적으로 약 11 억 3 천만 명의 사람들이 갖고 있는 고혈압은 다양한 질병들의 위험 요인으로서, “silent killer”로 알려져 있다. 고혈압의 진단과 심혈관계 질병의 예방을 위해서는 규칙적인 혈압 관찰이 중요하다. 일상 생활에서 쉽게 혈압 관찰을 할 수 있도록 심전도와 혈관용적파와 같은 생체 신호를 사용한 다양한 비가압 혈압 추정 방법들이 제안되었다.

본 연구에서는 딥 러닝 방법을 기반으로 한 무구속 혈압 추정 알고리즘이 제안되었다. 또한, 혈압 추정 모델의 성능을 증대시키기 위해 심탄도 신호가 사용되었다.

첫 번째로, 합성곱 신경망을 기반으로 개인화된 혈압 추정 모델을 개발하였다. 합성곱 신경망은 이미지 학습에 좋은 성능을 보이는 것으로 알려져 있으며, 혈압 추정에 필요한 특성들을 추출할 수 있는 능력을 보여주었다. 또한, 시간에 따른 특성들의 중요도를 강화하기 위해 어텐션 메커니즘이 추가되었다. 제안된 모델은 수축기 혈압, 이완기 혈압 추정에 각각 평균절대오차 4.06 mmHg 와 3.33 mmHg 의 값을 보였다.

두 번째로, 중간 혈압 교정 작업을 필요로 하지 않는 일반화된 혈압 추정 모델 또한 개발되었다. 혈압 교정 과정은 혈압 추정 정확도를 높이기 위해 필요하지만, 환자가 혈압을 측정하려 할 때마다 교정하는 것은 힘든 일이 될 수 있다. 합성곱 신경망은 잡음이 포함된 신호까지 입력으로 사용할 수 있지만 그로 인해 성능이 감소할 수 있는 단점이 있어, 합성곱 신경망 대신 장단기 기억 신경망 모델이 적용되었다. 제안된 모델은 재현성의 관점에서 시험되었고, 서로 다른 날 측정한 데이터로 테스트했을 때, 개인화된 모델보다 더 좋은 성능을 보여주는 것을 확인하였다. 제안된 모델의 평균절대오차 값은 수축기 혈압과 이완기 혈압 추정에 각각 5.82 mmHg 와 5.24 mmHg 이었다.

본 연구에서는 합성곱 신경망과 장단기 기억 신경망을 사용한 무구속 혈압 추정 모델이 개발되었다. 어텐션 메커니즘을 더한 합성곱 신경망은 신호 전체에 적용이 가능하고 특성 추출 과정이 필요 없는 장점을 가지고 있다. 혈압 추정 모델은 심탄도 신호를 추가하지 않았을 때보다 추가했을 때 높은 성능을 보여주었다. 또한 장단기 기억 신경망이 일반화된 혈압 추정 모델을 만들기 위해 사용되었다. 실험 결과를 보면, 제안된 모델이 더 일반화된 성능을 보여주는 것을 확인할 수 있었다. 본 연구는 개인화된 혈압 추정 모델과 일반화된 혈압 추정 모델 두 가지

모두를 포함하려고 시도하였다. 이는 일상생활에서 연속적으로 비가압 혈압 추정을 가능하게 하는 데 의의가 있다.

주요어: 비가압 혈압 추정, 컨볼루션 신경망, 어텐션, 장단기 기억 신경망, 개인화, 일반화

학 번: 2015-31047