

Article

Feasibility and Evaluation of Effective Morphological Feature Representation for Contactless Blood Pressure Estimation Using Remote Photoplethysmography

Hyojin Jo ^{1,†} , Dogyun Park ^{1,†} , Seok Young Lee ^{2,*}  and Nam Kyu Kwon ^{3,*} 

¹ Department of Robotics and AI Convergence, Yeungnam University, Gyeongsan-si 38541, Republic of Korea; hynn0254@yu.ac.kr (H.J.); parkdk4634@yu.ac.kr (D.P.)

² Department of Artificial Intelligence and Information Technology, Sejong University, Seoul 05006, Republic of Korea

³ Department of Electronic Engineering, Yeungnam University, Gyeongsan-si 38541, Republic of Korea

* Correspondence: lsy@sejong.ac.kr (S.Y.L.); namkyu@yu.ac.kr (N.K.K.); Tel.: +82-53-810-3095 (N.K.K.)

† These authors contributed equally to this work.

Abstract

This study evaluates the feasibility of morphological feature-based contactless blood pressure (BP) estimation using remote photoplethysmography (rPPG). A total of 52 morphological features are extracted from each cardiac cycle of rPPG and its first and second derivatives and arranges as temporal sequences for a long short-term memory (LSTM)-based model. In a calibration-based evaluation on a 16-subject public dataset, the 52-feature model achieved mean absolute errors (MAEs) of 11.45 mmHg for systolic BP (SBP) and 8.15 mmHg for diastolic BP (DBP), outperforming direct rPPG waveform input under the same training protocol. Model comparisons also showed that estimation performance varied with the architecture and model capacity. To identify an effective feature representation, recursive feature elimination (RFE) and principal component analysis (PCA) were compared across multiple dimensions. At 24 dimensions, PCA reduced MAE by 1.59 mmHg for SBP and 1.09 mmHg for DBP relative to RFE, while the 15-dimensional PCA representation was selected as the final compact representation considering both performance and feature dimension. The proposed approach was further evaluated against measurement-free and measurement-based reference baselines, while leave-one-subject-out (LOSO) validation was used to assess subject-independent generalization. LOSO showed substantial performance variation across held-out subjects, indicating limited generalizability to unseen subjects. Despite this limitation, the overall results support the feasibility of compact morphological rPPG feature representations for contactless BP estimation and highlight the importance of appropriate feature-space and model selection.



Academic Editors: Cheng-Ying Chung, Yu-Feng Chung and Chang-Wei Hsieh

Received: 8 July 2026

Revised: 26 August 2026

Accepted: 26 August 2026

Published: 27 August 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

Keywords: contactless measurement; blood pressure; remote photoplethysmography; morphological features; deep learning; dimensionality reduction

1. Introduction

Blood pressure (BP) is the pressure that blood pumped from the heart pushes against the walls of blood vessels, providing physiological information about cardiac function, vascular status, organ perfusion, and hemodynamics [1]. Hypertension, in particular, is a major risk factor for cardiovascular diseases such as heart disease and stroke, and is also considered an indicator for preventing deaths from all causes. Recently, the prevalence of

hypertension has been rising globally due to an aging population and increased exposure to lifestyle risk factors like unhealthy diets and lack of physical activity [2]. To prevent this, continuous and uninterrupted BP monitoring technology is required, driving research into various BP measurement methods.

The most accurate method for measuring BP is the invasive approach, which involves inserting a catheter into the skin to directly detect the pressure of blood flow. This catheterization allows for continuous BP measurement and has the advantage of being highly suitable for use as the clinical gold standard [3]. However, this method carries serious risks such as bleeding, making it difficult to use without the assistance of a medical professional [4,5]. In contrast, cuff-based oscillometry is simple to use and easily accessible in daily life. This method uses an automatic BP monitor for measurement and is currently the most widely used in both clinical and home settings. However, it requires direct pressure on the skin, which can be uncomfortable for the user, and each measurement takes at least about three minutes [6]. Therefore, oscillometry has the limitation of being unsuitable for continuous or real-time monitoring. To overcome the unsuitability of conventional direct arterial measurement methods for continuous monitoring, contact-based approaches that utilize biosignals such as photoplethysmography (PPG) and electrocardiogram (ECG) have been proposed [7].

ECG captures the heart's electrical activity using electrodes attached to the skin, while PPG detects pulsatile changes in blood volume by measuring variations in light absorption or reflection from peripheral tissue, such as the fingertip, earlobe, and toe. PPG is typically acquired using an optical sensor consisting of a light-emitting diode and a photodetector. It is widely incorporated into compact, non-invasive pulse oximeters for monitoring pulse rate and estimating peripheral oxygen saturation. Combining them yields the pulse arrival time (PAT), which is the interval between the R-peak of ECG and the systolic peak of PPG, and is equal to the sum of the pre-ejection period (PEP) and pulse transit time (PTT). Increased arterial stiffness accelerates pulse wave velocity; for a fixed path length, PTT shortens and BP rises, so BP and PTT are generally in an inverse relationship [8]. Conventional approaches leverage this characteristic to estimate BP via regression [9,10]. However, obtaining accurate PTT requires separating PEP, demanding extra sensors such as impedance cardiography [11]. For this reason, many studies employ PAT without PEP calibration as an indirect inverse marker of BP [12,13]. But this may cause errors, as the embedded PEP can vary independently of BP [14]. To address this, PPG-only approaches have been reported, as in [15]. This approach calculates PTT by measuring the foot-to-foot or peak-to-peak time difference between PPG signals from two different peripheral sites (e.g., finger–foot) without using ECG [16,17].

However, conventional contact-based BP estimation using PTT can suffer from synchronization issues because it requires at least two sensors. To overcome this, single-site PPG approaches with machine learning [18,19] have been explored as an alternative to dual-site measurements. In practice, PPG is generally preferred over ECG due to easier acquisition and greater convenience [20]. When PPG is used alone, the input features of the network usually comprise the raw waveform, its derivatives, time–frequency spectrograms, and handcrafted waveform features [21–23]. Among these, feature extraction from the PPG waveform has been studied extensively, and diverse morphological and statistical descriptors have been derived from the original signal and its derivatives. Such handcrafted features are not only useful for BP estimation but are also linked to physiologically meaningful waveform characteristics, thereby providing a physiologically interpretable feature representation [24]. In addition, a key consideration is identifying a compact set of informative features that can achieve high performance while preventing an increase in the false detection rate [25]. This is closely related to dimensionality reduction, which

is an essential preprocessing step prior to machine learning. It removes unnecessary or redundant information, improves learning performance, and enhances the interpretability of the results [26]. For example, El-Hajj and Kyriacou [27] extracted 52 PPG features and reduced them to 24 using recursive feature selection (RFE). Nevertheless, wearing a sensor can cause discomfort, may be unsuitable for individuals with skin sensitivity, and raises infection risks. These limitations motivate research on contactless BP estimation as an alternative [28,29].

The contactless method is based on remote PPG (rPPG), which shares the hemodynamic basis of PPG. Unlike contact sensing, rPPG constructs RGB signals from facial skin ROIs in video and then extracts the pulse via preprocessing algorithms. This enables tracking of blood flow in the skin and monitoring various vital signs [30]. Building on this signal, rPPG-based BP estimation has been studied primarily through two approaches. First, the PTT-based approach computes delays between different ROIs (e.g., forehead–cheek) or wavelength bands (e.g., red–green) within a face video, then uses them in regression or deep learning models [31,32]. Second, a direct approach uses facial video or the rPPG waveform as input to a deep learning model to estimate BP [33,34]. There has been active work on explicitly learning temporal information using either spatio-temporal representations (e.g., STMap [35]) or sequence models such as long short-term memory (LSTM) and transformers [36,37].

In parallel, studies have begun to port established morphological features for PPG to rPPG under the same definitions, and estimate BP using machine learning [38–40]. Approaches that combine features are considered to exhibit superior estimation performance compared to those using only demographic information [41]. Furthermore, comparative studies with contact-based features have suggested that rPPG also holds potential for remotely capturing BP-related properties [42]. However, most feature-based rPPG studies for BP estimation do not model the temporal structure of the signal, and they mainly report feasibility with traditional machine learning such as support vector regression (SVR) or simple multi-layer perceptron (MLP) models. Furthermore, while dimensionality reduction is essential when using a large number of features, traditional feature selection-based methods may not sufficiently preserve the information in the defined feature space. Therefore, it is important to evaluate not only whether morphological rPPG features are useful for BP estimation, but also how these features should be represented and processed effectively.

This paper presents a feasibility evaluation of effective morphological feature representation for contactless BP estimation using rPPG. We extract 52 morphological features from each rPPG cardiac cycle and its first and second derivatives, arrange them as temporal sequences, and estimate systolic BP (SBP) and diastolic BP (DBP) using an LSTM-based model. To evaluate how representation and model design affect estimation, we compare the 52-feature input with a matched resampled rPPG waveform, conventional machine learning models and recurrent ablations, and feature representations based on principal component analysis (PCA) and RFE across multiple dimensions. We also use measurement-free global-mean and per-subject-mean baselines and a measurement-based last-known-BP baseline to distinguish estimation beyond average BP levels and temporal persistence. The main analysis uses a blocked calibration-based split with no temporal overlap, while leave-one-subject-out (LOSO) validation is additionally used to assess subject-independent generalization. Overall, the study examines the feasibility of structured morphological rPPG features while identifying practical choices in model configuration, feature-space size, and performance evaluation.

2. Data Preprocessing

2.1. Dataset

The public rPPG data used in this study contain recordings from 17 subjects collected at Leipzig University Hospital [36]. A 2-h video was recorded using a 32 fps industrial USB camera for patients scheduled for surgery, while BP labels were recorded at one-minute intervals via a bedside monitor. The synchronized rPPG signals were extracted from the forehead and cheek region and processed using the Plane-Orthogonal-to-Skin (POS) algorithm [43]. Each released recording was segmented into windows based on heart rate, with each window defined to contain seven cardiac cycles and resampled to a length of 875 points. The BP distribution histogram in the dataset is shown in Figure 1, where the mean and standard deviation of SBP and DBP are 143.40 ± 14.97 mmHg and 65.68 ± 11.30 mmHg, respectively.

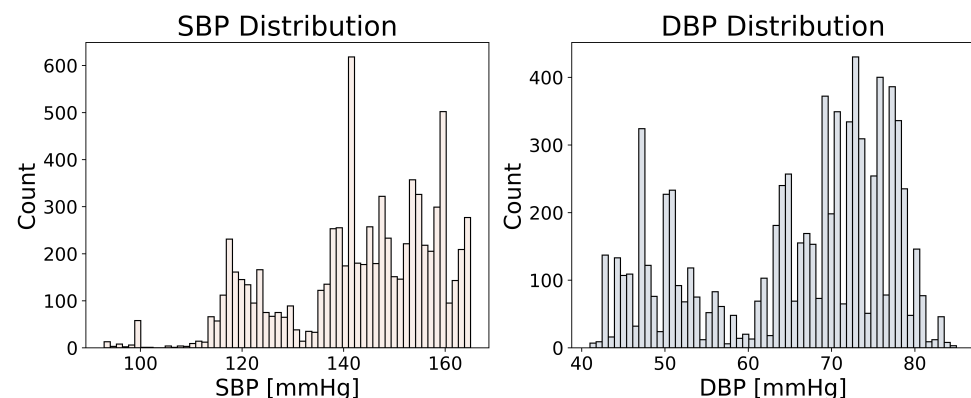


Figure 1. Histogram of BP labels (**left:** SBP, **right:** DBP). The *x*-axis represents BP values and the *y*-axis indicates counts.

2.2. rPPG Morphological Feature Extraction

For each cycle of the rPPG signal, 52 BP-related features derived from PPG morphology are extracted. Prior to feature extraction, min-max normalization is applied to each 7 s rPPG segment, after which individual cycles are delineated. Cycles are excluded when any required reference point or derived feature cannot be reliably identified. Using this strict exclusion criterion, approximately 75.7% of the 111,868 initially extracted cycles were excluded, resulting in 27,135 valid cycles.

The reference points used for feature extraction include the starting point (*foot*), end point (*valley*), and systolic peak (*SP*) of the rPPG waveform; the first peak and valley of $rPPG'$ ($a1$, $b1$) [27]; and the first positive peak, first negative valley, and subsequent re-increasing peak of $rPPG''$ (a , b , c). Here, a denotes the early systolic positive peak, b denotes the early systolic negative peak, and c denotes the late systolic re-increasing peak.

The conventionally used DW_p/SW_p provides BP-related information by capturing the temporal asymmetry between systole and diastole. In this study, p is set to 10, 25, 33, 50, 66, and 75. Measured at the same $p\%$ amplitude level, these widths indirectly reflect changes in wave-reflection timing and arterial stiffness [44]. The value was calculated by defining the 10% amplitude level based on the lower-amplitude point between the foot and valley. However, if the defined 10% amplitude level was lower than all amplitude values on the opposite side, such that no corresponding crossing point could be identified, the minimum-amplitude point on the opposite side was used as the 10% amplitude point. As shown in Figure 2a, systolic time (*ST*) and diastolic time (*DT*) denote the rising and falling phases of the rPPG waveform, respectively. Likewise, systolic area (*SA*) and diastolic area (*DA*) are defined as the signal areas within the corresponding intervals and are strongly associated with the strength of pulse wave reflection [45]. In addition, pulse interval (*PI*)

represents the duration of one cycle and is closely related to heart rate variability [46]. Together, these features reflect the dynamic differences between systole and diastole within a single pulse cycle.

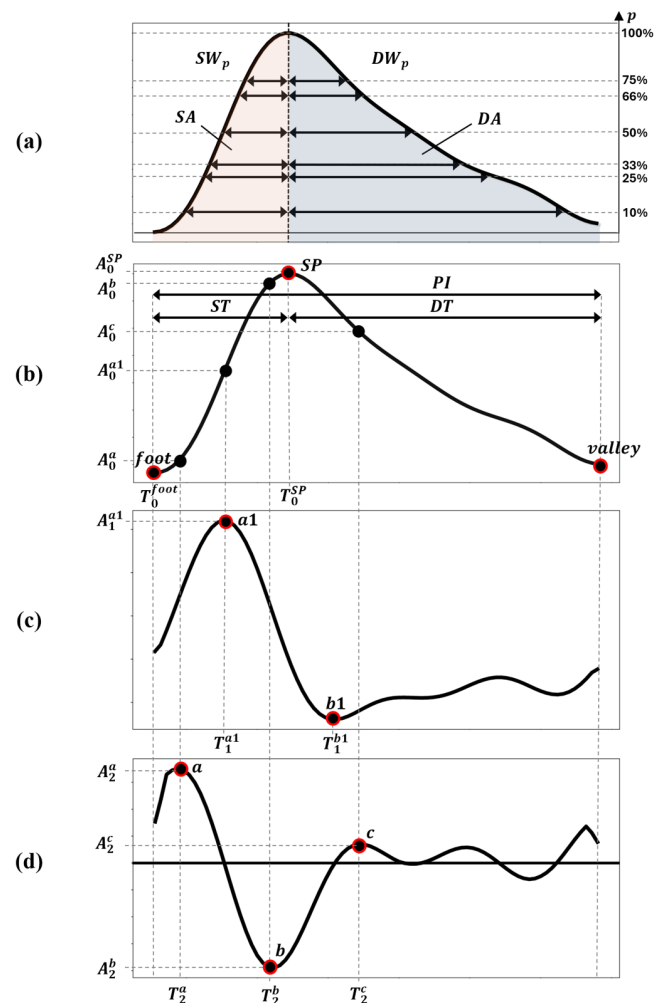


Figure 2. Visualizations of the features (In all subfigures, the x -axis denotes time and the y -axis represents the amplitude of the signal): (a) DW_p/SW_p values and systolic/diastolic areas (right y -axis indicates the p level of the systolic peak amplitude), (b) rPPG signal, (c) rPPG' signal, (d) rPPG'' signal. The black points outlined in red indicate the reference points used for feature extraction.

The remaining features are defined as derived features combining the time and amplitude coordinates of the reference points. The time and amplitude coordinates are denoted by T_o^x and A_o^x , respectively, where $o \in 0, 1, 2$ indicates rPPG, rPPG', and rPPG'', and $x \in foot, SP, a1, b1, a, b, c$ denotes the reference point used for computation. These coordinate-based features reflect not only the absolute position of the pulse wave but also the relative time difference, amplitude ratio, slope, and morphological changes between them. In particular, derivative-based reference points enable more stable localization of turning points in the waveform even when inflection points in the original signal are unclear, and second derivative-based features have repeatedly been reported as significant indicators in BP estimation and vascular aging assessment [47]. Detailed feature definitions are summarized in Table 1, and the reference points and coordinates can be visually confirmed in Figure 2b–d.

Table 1. Summary of Extracted Features.

Source Signal	No. of Features	Feature Definition
rPPG	1	A_0^{SP}
	18	$DW_p, DW_p + SW_p, DW_p/SW_p$
	3	$SA, DA, DA/SA$
	4	$ST, DT, ST/PI, DT/PI$
	2	$PI, PI/T_0^{SP}$
	1	A_0^{SP}/A_0^{foot}
	1	Heart rate = $60/(T_0^{SP} - T_{0,prev}^{SP})$
rPPG'	1	A_1^{a1}
	2	$T_1^{a1} - T_0^{foot}, T_1^{b1} - T_1^{a1}$
rPPG''	4	$A_2^a, A_2^b, A_2^b/A_2^a, A_2^c/A_2^a$
	3	$T_2^a - T_0^{foot}, T_2^b - T_2^a, T_2^c - T_2^b$
	1	$A_2^a - A_2^b$
rPPG, rPPG'	2	$A_0^{a1}, A_0^{a1}/T_1^{a1}$
	2	$T_0^{SP} - T_1^{a1}, (T_0^{SP} - T_1^{a1})/PI$
	2	$T_1^{a1}/PI, T_1^{b1}/PI$
rPPG, rPPG''	2	$A_0^b/A_0^a, A_0^c/A_0^a$
	3	$(T_2^a - T_0^{foot})/PI, (T_2^b - T_2^a)/PI, (T_2^c - T_2^b)/PI$

2.3. Dimensionality Reduction

RFE and PCA reduce dimensionality in different ways. RFE is a supervised feature-selection method that retains a subset of the original features according to feature importance estimated by a specified estimator. In contrast, PCA is a feature-extraction method that transforms the original correlated features into orthogonal principal components ordered by the amount of variance they explain. For the 52 morphological features used in this study, PCA can reduce redundancy among correlated features while preserving information distributed across multiple features in a lower-dimensional representation. Because each principal component is a linear combination of the original features, however, PCA provides less direct physiological interpretability than RFE.

This study adopts PCA as the primary dimensionality reduction strategy. To establish a comparison baseline, a set of 24 features was selected using RFE following [27], and its performance was compared with that of the 24-dimensional representation obtained by PCA. In addition, to further examine the effect of dimensionality on BP estimation performance, both PCA and RFE were evaluated at 5, 10, 15, 20, 30, and 40 dimensions. For these additional RFE experiments, SBP and DBP were jointly considered during feature selection, with a random forest used as the estimator and the elimination step set to 1.

The training-set feature vectors are stacked row-wise to form the high-dimensional data matrix $D \in \mathbb{R}^{n \times m}$, where n represents the number of training feature vectors and m denotes the feature dimension ($m = 52$). To perform PCA, the mean of each feature is first subtracted. The PCA transformation learned from this training matrix is subsequently applied to the validation and test feature vectors, as described in Section 3.2. Let μ_j be the mean of the j th feature, i.e., the j th column. Then, the entry of the resulting matrix $\bar{D} \in \mathbb{R}^{n \times m}$ can be expressed as follows [48]:

$$\bar{d}_{ij} = d_{ij} - \mu_j. \quad (1)$$

The covariance matrix is given by [49]

$$\Sigma = \frac{1}{n} \bar{D}^T \bar{D} \quad (2)$$

which admits the eigendecomposition [50]

$$\Sigma = Q \Lambda Q^T, \quad (3)$$

where $Q = [\mathbf{q}_1, \mathbf{q}_2, \dots, \mathbf{q}_m] \in \mathbb{R}^{m \times m}$ contains orthonormal eigenvectors of Σ satisfying $Q^T Q = I$, and

$$\Lambda = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & \cdots & \lambda_m \end{bmatrix} \quad (4)$$

with eigenvalues ordered as $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_m \geq 0$. This decomposition can be equivalently written as

$$\Sigma = \sum_{k=1}^m \lambda_k \mathbf{q}_k \mathbf{q}_k^T. \quad (5)$$

Given a target reduced dimension $\bar{m}$ ($\bar{m} \leq m$), the data are projected onto the first $\bar{m}$ principal components as [51]

$$Z = \bar{D} Q_{\bar{m}}, \quad (6)$$

where $Q_{\bar{m}} = [\mathbf{q}_1, \mathbf{q}_2, \dots, \mathbf{q}_{\bar{m}}] \in \mathbb{R}^{m \times \bar{m}}$ is the lower dimensional representation of the feature matrix obtained by PCA. The cumulative explained variance ratio (CEV) up to the $\bar{m}$ th component is defined as [48]

$$\text{CEV}(\bar{m}) = \frac{\sum_{k=1}^{\bar{m}} \lambda_k}{\sum_{k=1}^m \lambda_k}. \quad (7)$$

In this study, PCA-based dimensionality reduction was evaluated using 5, 10, 15, 20, 24, 30, and 40 principal components to assess performance across different feature dimensions and to facilitate direct comparison with the RFE-based approach. Among these settings, 15 principal components achieved a CEV of approximately 0.99.

To further examine how each original morphological feature is represented in the PCA-reduced space, a feature contribution analysis was performed using the eigenvalues and eigenvectors obtained from PCA. Let $q_{k,j}$ denote the j -th element of the eigenvector $\mathbf{q}_k$ corresponding to the k -th principal component. The coefficient $q_{k,j}$ can take either a positive or negative value depending on the direction of the corresponding principal-component axis. Therefore, $q_{k,j}^2$ was used to remove the directional effect of the sign and to prevent positive and negative coefficients from canceling each other when contributions are accumulated across multiple principal components. Weighting by λ_k further accounts for the amount of variance represented by each principal component.

Accordingly, the contribution of the j -th original feature across the retained principal components was defined as

$$C_j = \sum_{k=1}^{\bar{m}} \lambda_k q_{k,j}^2, \quad j = 1, \dots, m, \quad (8)$$

where m denotes the number of original features and $\bar{m}$ denotes the number of retained principal components. Since a small number of contribution values were substantially larger than the others, making the remaining contributions difficult to distinguish, a common logarithmic transformation was applied as follows:

$$\tilde{C}_j = \log_{10}(C_j). \quad (9)$$

The transformed values allow the relative representation of the original morphological features within the retained PCA space to be compared. However, these contribution values should be interpreted as indicating how strongly each original feature is represented in the variance preserved by the retained principal components, rather than as a direct measure of its importance for BP estimation.

3. Deep Learning Methodology

This section describes the training procedure using the LSTM-based architecture. To provide a controlled comparison between the feature and signal representations, both inputs were configured to have the same input size and were evaluated using the same network architecture and training protocol. The overall methodology consists of model architecture design, input data construction, and training details.

3.1. Model Architecture

Since BP is estimated using feature vectors arranged over time, this study adopted an LSTM-based model. The model consists of one bidirectional LSTM (BiLSTM), three unidirectional LSTMs (UniLSTMs), one attention layer, and two dense layers. The BiLSTM is used to capture temporal dependencies in both forward and backward directions, allowing the model to incorporate contextual information from preceding and subsequent feature vectors within each sequence. The subsequent UniLSTMs further refine the sequential representation by progressively extracting temporal patterns relevant to BP estimation. The BiLSTM has 256 units per direction (512 total), while the subsequent UniLSTMs have 256, 128, and 64 units, respectively. A dropout rate of 0.3 is applied after each recurrent layer to reduce overfitting. The attention layer is then applied to emphasize temporally informative representations while assigning relatively lower weights to less relevant time steps, enabling the model to focus on the portions of the sequence that contribute more strongly to BP estimation. The resulting representation is passed through two dense layers to estimate the final SBP and DBP values. For the 52-dimensional feature input, the model contains 1,669,026 trainable parameters. A linear activation function is applied to both dense layers. The model architecture is shown in Figure 3.

Furthermore, to evaluate the effectiveness of the proposed model architecture, SVR, MLP, and recurrent ablation models were used as comparison models. The SVR model used a radial basis function kernel, with C , γ , and ϵ set to 10, 0.0005, and 0.01, respectively. The MLP consisted of five hidden layers with 512, 256, 128, 64, and 32 nodes. ReLU activation was applied to the hidden layers, while the output layer simultaneously estimates SBP and DBP using a linear activation function. For the recurrent ablation analysis, three variants were evaluated: a model without the attention layer, a model without the BiLSTM layer, and a model without both the attention and BiLSTM components. These variants were used to examine the effects of bidirectional temporal processing and attention within the recurrent architecture.

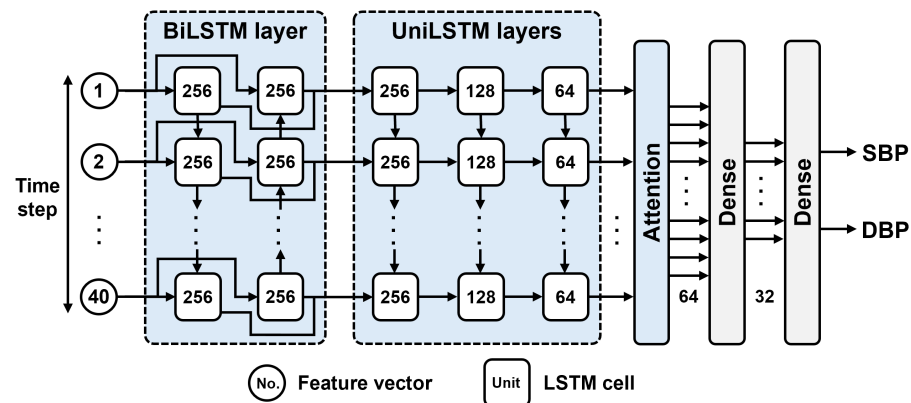


Figure 3. LSTM-based BP estimation model architecture. The model is composed of one BiLSTM layer, three UniLSTM layers, and one attention layer. It takes a sequence of rPPG-derived feature vectors as input and produces SBP and DBP. The input sequence has an analysis length (time step) used when estimating an SBP/DBP pair. LSTM cells share the same parameters and learn temporal patterns by repeatedly updating the hidden state, which has a dimension corresponding to the number of units, for each feature vector.

3.2. Input Data Processing

First, the data are partitioned into training, validation, and test sets. The removal of cycles with feature loss, as described in Section 2.2, leads to variation in the number of feature vectors (i.e., sequence length) across subjects. This variability may lead to overfitting to specific individuals. For example, the number of valid cycles varied from 1125 in one subject to 2422 in another. To mitigate this issue, a calibration-based learning setting is adopted, and each subject's sequence is evenly divided into 10 segments, which are then assigned to each set. For subject s , the sequence length is denoted by l_s . Then, the first nine segments have length $\lfloor l_s / 10 \rfloor$, and the final segment includes any remaining vectors after division. The segments are assigned to training, validation, and test sets in chronological order at an 8:1:1 ratio, ensuring that no temporal overlap exists across the sets, so that the model is evaluated on time ranges not observed during training. This split strategy required at least 400 valid cycles per subject. Subjects who did not satisfy this requirement were excluded, resulting in the exclusion of one subject from the training process. Consequently, 16 subjects were used in the final analysis. In addition to this calibration-based evaluation, LOSO validation was performed by evaluating each subject as the held-out subject in turn, in order to examine subject-independent generalization.

Second, min-max normalization is applied to reduce scale differences across features. The normalization is applied independently to each of the 52 features, regardless of the subject. It is performed only on the training set, and the resulting minimum and maximum values are consistently used to normalize the validation and test sets. Meanwhile, when using the PCA-reduced representation Z from Section 2.3, Z is used directly without additional normalization. Similarly, PCA is fitted only on the training set, and the learned transformation is applied to the validation and test sets.

Finally, the model input time series is constructed. At each training step, a sequence of 40 consecutive feature vectors is used as the input, and this sequence length is defined as the time step. If each feature vector has dimension m , the resulting input size is $40 \times m$. The sequence length of 40 cycles was selected to provide the recurrent model with temporal context across multiple consecutive cardiac cycles. The time step represents the temporal window of cardiac activity used to estimate a single BP pair. The reference SBP and DBP assigned to each input window were calculated by averaging the BP labels corresponding to the 40 cardiac cycles included in that window. The input window is then shifted by a

predefined stride to read the next sample. In this study, the stride is set to 20, resulting in a 50% overlap between adjacent windows. This process is applied separately to each data split after partitioning, ensuring that no temporal information leaks between the training, validation, and test sets. These steps are illustrated in Figure 4.

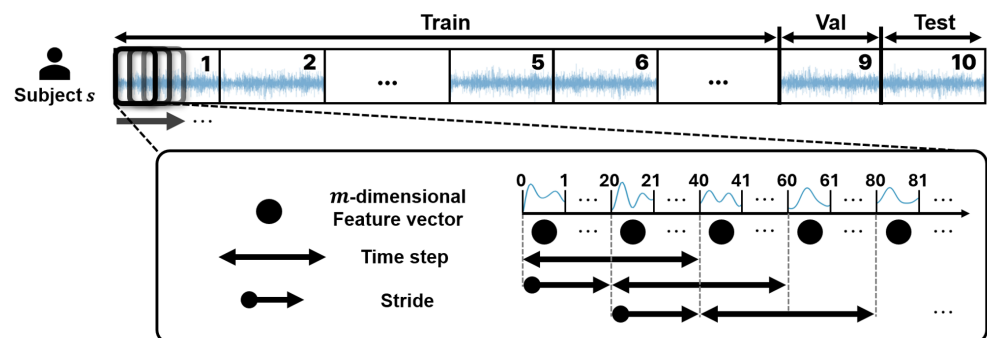


Figure 4. Input data processing illustration: The dataset is split into training, validation, and test sets at an 8:1:1 ratio, followed by input data configuration based on the time step (segment size required to estimate one BP pair) and stride (distance to the next segment).

In addition, to validate the effectiveness of the feature-based approach, a model using the rPPG signal as input was trained. To ensure a consistent input size, each pulse-cycle segment was resampled to a fixed length of 52 samples and treated as a 52-dimensional input vector, analogous to the 52-dimensional feature representation. The resampled signals were then arranged into sequences of 40 time steps with a stride of 20, following the same input configuration used for the feature-based approach. These sequences were fed into the same network architecture and trained using the same training protocol.

3.3. Training Details

The experiments were implemented in TensorFlow. The training process was conducted for 300 epochs with a batch size of 8, a learning rate of 0.0001, and the Adam optimizer. Mean squared error was employed as the cost function. Early stopping was applied when the validation loss did not improve for 20 consecutive epochs, while overfitting was defined as an increase of more than 50% in the validation loss relative to the previous epoch. Both feature-based and signal-based learning, including the MLP model, were conducted under the same conditions as described above. To reduce the effect of random initialization and data shuffling, the calibration-based model comparisons and dimensionality-reduction experiments were trained 20 times using different random seeds, and the reported quantitative results represent the average performance across these 20 runs. For LOSO validation, each held-out-subject fold was repeated five times, and the reported subject-wise metric values represent the mean across the five runs. Model training was performed on RTX A4500 GPUs.

4. Experimental Results

This section reports the quantitative performance of the proposed rPPG-based BP estimation framework under the calibration-based split and LOSO validation, together with comparisons of input representations, model variants, and dimensionality-reduction settings. All experiments used a time step of 40 and a stride of 20. Performance was evaluated using mean absolute error (MAE), root mean squared error (RMSE), standard deviation (SD) of the estimation error, mean difference (MD), and Pearson correlation coefficient (R); MAE, RMSE, SD, and MD are reported in mmHg. For repeated training experiments, values are reported as the mean $\pm$ run-to-run SD unless otherwise noted. The

SD metric itself denotes the sample-wise SD of the estimation errors. The model architecture comparison reports the 20-run mean values together with the corresponding run-to-run SDs and model capacity. For the regression and Bland–Altman visualizations, a representative calibration-based run was selected by identifying the run closest to the 20-run mean for each of the ten evaluation metrics. The run identified most frequently across these metrics was selected for Figure 5. In the case of a tie, the run showing a more balanced number of closest matches across SBP and DBP was chosen. For Figure 6, the 24-dimensional PCA result obtained using the same random seed as the representative run in Figure 5 was used for comparison.

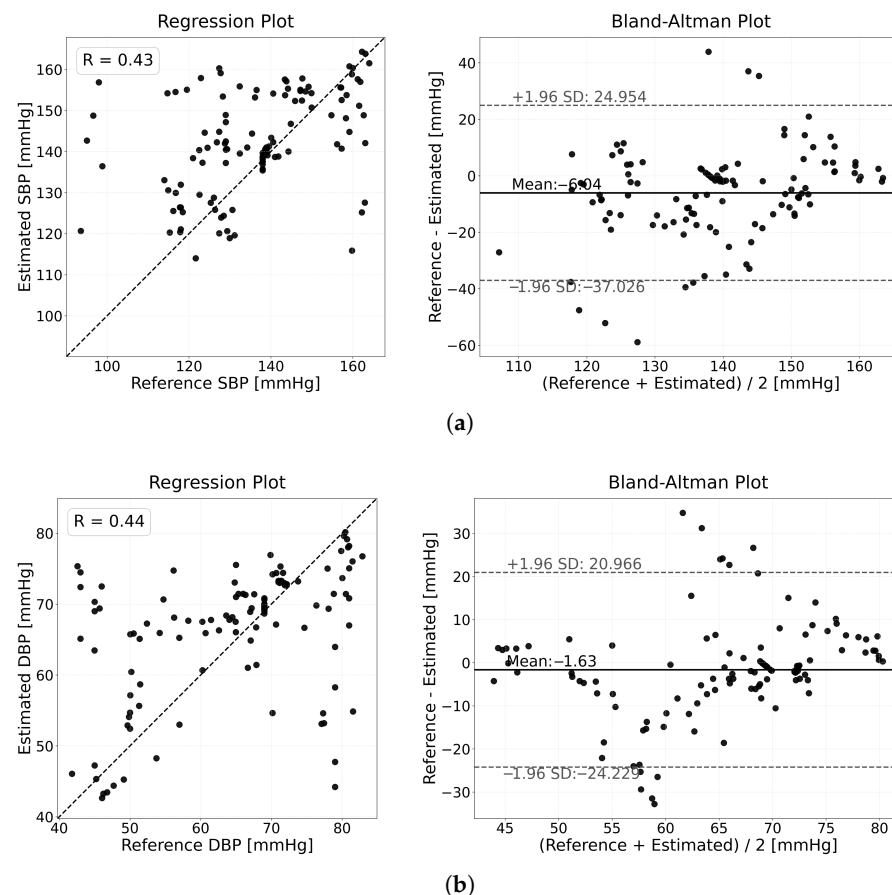


Figure 5. Regression and Bland–Altman plots for the representative run using the 52-dimensional feature representation: (a) SBP and (b) DBP. For SBP, MAE, RMSE, SD, and MD were 11.66, 16.92, 15.81, and -6.04 mmHg, respectively; for DBP, they were 7.93, 11.64, 11.53, and -1.63 mmHg, respectively.

Three non-learned baselines were evaluated using the same window-level reference targets as the proposed model, but they were included for different purposes. The global-mean and per-subject-mean baselines were used as measurement-free baselines that do not rely on a specific recent BP measurement for estimation. The global-mean baseline assigns the mean BP of all training windows to every test window, whereas the per-subject-mean baseline assigns each subject's mean training BP to that subject's test windows. These baselines were included to determine whether the proposed model performs better than simply reproducing the overall or subject-specific average BP level. In contrast, the last-known-BP baseline was treated separately as a measurement-based baseline because it directly uses a specific previously measured BP value from the same subject. For each subject, the final BP label in the training split was retained and used as a constant estimate for all subsequent test windows. This baseline was included to assess whether the proposed model can capture BP changes beyond the temporal persistence represented by the most

recent available BP measurement. Although the last-known-BP baseline does not represent a complete long-term estimation strategy because a fixed previous value cannot reflect subsequent BP changes, it provides a strong reference for evaluating temporal BP tracking. All three baselines generate constant estimates within their corresponding evaluation sets; therefore, R was not reported.

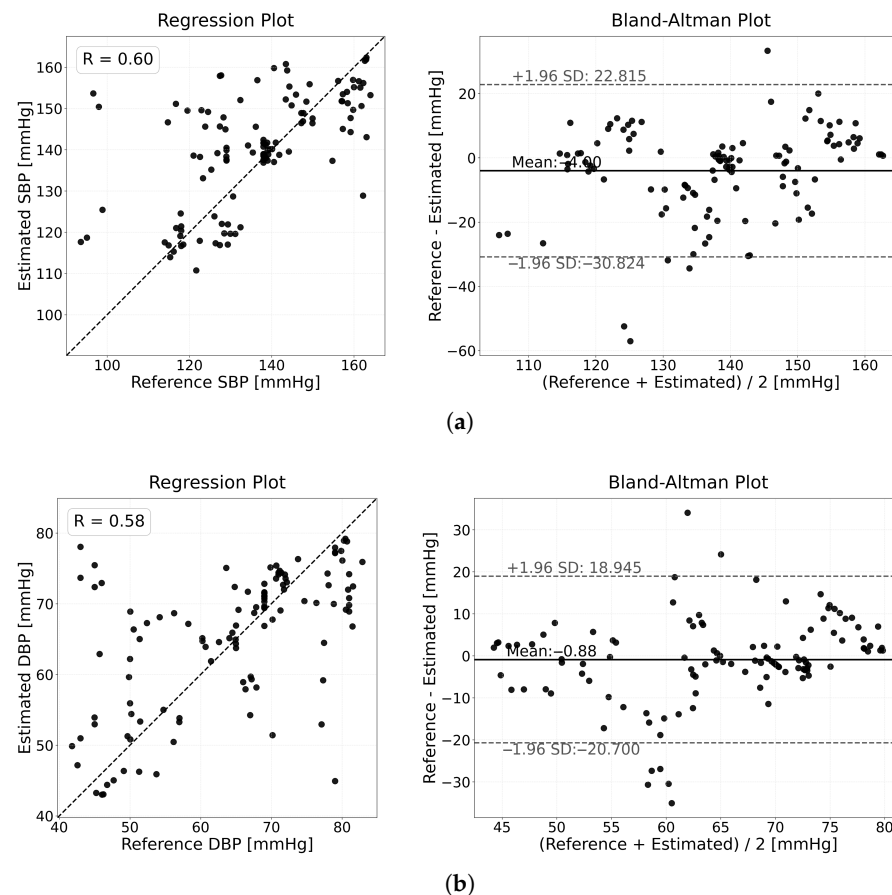


Figure 6. Regression and Bland–Altman plots for the representative run using the 24-dimensional PCA representation: (a) SBP and (b) DBP. For SBP, MAE, RMSE, SD, and MD were 9.72, 14.26, 13.68, and -4.00 mmHg, respectively; for DBP, they were 6.90, 10.15, 10.11, and -0.88 mmHg, respectively.

As shown in Table 2, the global-mean baseline yielded MAEs of 14.28 mmHg for SBP and 9.99 mmHg for DBP. The per-subject-mean baseline in Table 3 showed substantial inter-subject variation; for example, Subjects 1 and 2 had SBP MAEs of 24.10 and 45.98 mmHg despite relatively small within-test error SDs of 2.51 and 1.91 mmHg. This pattern indicates that large errors can arise from a shift in BP level between the training and test periods even when the test-period BP itself varies little. Since the per-subject-mean estimation is constant, its error SD is equivalent to the SD of the corresponding test reference BP and therefore provides a direct indication of within-subject BP variability. The SBP error SD ranged from 0.00 to 12.64 mmHg and the DBP error SD from 0.00 to 5.51 mmHg. Subject 15 contained only one test window, while several other subject–BP combinations had very limited test variation. These cases limit the extent to which within-subject BP tracking can be assessed from correlation or MAE alone.

The last-known-BP baseline in Table 4 further illustrates the effect of temporal persistence. Its performance varied markedly across subjects, and DBP errors were zero for Subjects 10 and 16 because each of these test sets contained only one distinct window-level DBP label and that value matched the retained last-known BP. Thus, very low errors can occur without rPPG-derived information when BP remains unchanged across the evaluated

test windows, underscoring the need for both persistence baselines and explicit reporting of within-subject BP variation.

Table 2. Performance of the global-mean baseline for SBP and DBP estimation.

SBP				DBP			
MAE	RMSE	SD	MD	MAE	RMSE	SD	MD
14.28	17.56	16.44	−6.19	9.99	12.01	11.97	−1.01

Table 3. Performance of the per-subject-mean baseline for SBP and DBP estimation.

Subject No.	Test Windows	SBP				DBP			
		MAE	RMSE	SD	MD	MAE	RMSE	SD	MD
1	11	24.10	24.23	2.51	−24.10	2.47	2.89	1.50	2.47
2	5	45.98	46.02	1.91	−45.98	22.81	22.82	0.64	−22.81
3	3	7.14	7.19	0.78	−7.14	5.85	5.87	0.50	5.85
4	4	5.02	5.02	0.22	5.02	15.51	15.51	0.08	−15.51
5	4	15.76	15.83	1.51	15.76	11.96	11.97	0.31	11.96
6	8	19.15	20.01	5.80	−19.15	11.75	12.08	2.79	−11.75
7	4	1.65	1.73	0.53	1.65	0.45	0.49	0.19	−0.45
8	14	11.47	14.37	12.64	−6.84	7.22	9.08	5.51	−7.22
9	6	15.89	15.97	1.56	15.89	15.21	15.24	1.01	15.21
10	13	4.24	4.25	0.29	−4.24	3.51	3.51	0.00	3.51
11	13	7.02	10.13	9.27	4.09	8.95	9.87	4.16	8.95
12	4	18.80	18.90	1.90	18.80	3.27	3.72	2.58	−2.68
13	8	13.73	13.74	0.54	−13.73	13.81	14.00	2.33	13.81
14	13	15.03	15.26	2.65	−15.03	19.02	19.11	1.82	−19.02
15	1	18.79	18.79	0.00	18.79	4.68	4.68	0.00	4.68
16	4	20.31	20.32	0.32	20.31	13.51	13.51	0.00	13.51

Table 4. Performance of the last-known-BP baseline for SBP and DBP estimation.

Subject No.	SBP				DBP			
	MAE	RMSE	SD	MD	MAE	RMSE	SD	MD
1	4.30	4.98	2.51	4.30	1.96	2.47	1.50	1.96
2	2.58	3.20	1.91	−2.58	0.32	0.56	0.46	−0.32
3	1.74	1.91	0.78	−1.74	0.66	0.83	0.50	−0.66
4	0.42	0.48	0.22	0.42	0.75	0.75	0.08	0.75
5	4.16	4.43	1.51	4.16	1.45	1.48	0.31	1.45
6	9.25	10.92	5.80	9.25	4.80	5.55	2.79	4.80
7	0.53	0.59	0.53	−0.27	1.96	1.97	0.19	−1.96
8	12.91	17.23	12.64	−11.71	5.73	7.38	5.51	−4.91
9	1.23	1.58	1.56	−0.22	0.76	1.01	1.01	−0.06
10	0.17	0.34	0.29	0.17	0.00	0.00	0.00	0.00
11	6.82	10.78	9.27	5.50	3.08	4.82	4.16	2.44
12	2.94	3.50	1.90	2.94	3.05	3.99	2.58	3.05
13	2.68	2.73	0.54	2.68	1.70	2.89	2.33	−1.70
14	6.62	7.14	2.65	−6.62	6.53	6.78	1.82	−6.53
15	1.20	1.20	0.00	1.20	1.16	1.16	0.00	1.16
16	0.72	0.79	0.32	0.72	0.00	0.00	0.00	0.00

4.1. Effectiveness of Feature-Based Estimation

Tables 5 and 6 compare the 52-dimensional rPPG-derived feature representation with the resampled rPPG signal input under the same sequential configuration and training

protocol. The feature input achieved MAE, RMSE, error SD, MD, and R values of 11.45, 16.78, 15.48, -6.39 , and 0.44 for SBP and 8.15 , 11.90 , 11.74 , -1.81 , and 0.40 for DBP, respectively. In comparison, the signal input yielded MAEs of 15.91 mmHg for SBP and 14.89 mmHg for DBP, together with higher RMSE and error SD and lower R for both BP types.

Table 5. SBP estimation performance using the 52-dimensional rPPG-derived feature vectors and the resampled rPPG signal input.

Input	MAE	RMSE	SD	MD	R
Feature	11.45 ± 1.07	16.78 ± 1.12	15.48 ± 0.94	-6.39 ± 1.31	0.44 ± 0.08
Signal	15.91 ± 0.38	21.33 ± 0.49	21.18 ± 0.45	-2.26 ± 1.09	0.27 ± 0.03

Table 6. DBP estimation performance using the 52-dimensional rPPG-derived feature vectors and the resampled rPPG signal input.

Input	MAE	RMSE	SD	MD	R
Feature	8.15 ± 0.81	11.90 ± 0.99	11.74 ± 0.93	-1.81 ± 0.76	0.40 ± 0.12
Signal	14.89 ± 0.41	17.28 ± 0.35	16.64 ± 0.27	4.60 ± 0.78	-0.19 ± 0.04

Relative to the global-mean baseline in Table 2, the 52-feature model reduced MAE by 2.83 mmHg for SBP and 1.84 mmHg for DBP, RMSE by 0.78 and 0.11 mmHg, and error SD by 0.96 and 0.23 mmHg, respectively. The absolute MD was slightly larger for the learned feature model in both BP types, indicating that its improvement was primarily reflected in absolute and dispersion-based error measures rather than reduced signed bias.

Table 7 summarizes the percentages of estimates with absolute errors within 5 , 10 , and 15 mmHg. According to the BHS grading criteria, Grades A, B, and C require at least $60\%/85\%/95\%$, $50\%/75\%/90\%$, and $40\%/65\%/85\%$ of measurements to fall within $5/10/15$ mmHg, respectively. Across the 20 runs, the proposed model achieved mean percentages of 40.43% , 58.39% , and 73.61% for SBP and 51.57% , 72.87% , and 82.17% for DBP, with run-to-run SDs of 5.93 , 4.26 , and 4.91 percentage points for SBP and 4.00 , 3.93 , and 3.17 percentage points for DBP across the three thresholds. Neither BP type satisfied all three thresholds required for BHS Grade C. For contextual comparison with the AAMI/ESH/ISO numerical accuracy criterion, the MD between the estimated and reference BP should be within ± 5 mmHg, with an error SD no greater than 8 mmHg. The 52-feature model had an MD of -6.39 mmHg and an error SD of 15.48 mmHg for SBP and an MD of -1.81 mmHg and an error SD of 11.74 mmHg for DBP. Thus, the SBP result exceeded both numerical limits, whereas DBP satisfied the mean-error limit but exceeded the SD limit. These criteria are provided only as clinical context because the present study does not satisfy the full subject-number, reference-measurement, and protocol requirements for formal device validation.

Table 7. Comparison of the error percentages of the proposed 52-feature model with the BHS criteria listed for Grades A–C.

		<5 mmHg	<10 mmHg	<15 mmHg
Ours	SBP	$40.43 \pm 5.93\%$	$58.39 \pm 4.26\%$	$73.61 \pm 4.91\%$
Ours	DBP	$51.57 \pm 4.00\%$	$72.87 \pm 3.93\%$	$82.17 \pm 3.17\%$
BHS	Grade A	60%	85%	95%
BHS	Grade B	50%	75%	90%
BHS	Grade C	40%	65%	85%

The subject-wise calibration-based results are presented in Tables 8 and 9. For SBP, mean MAE ranged from 1.64 mmHg in Subject 10 to 44.56 mmHg in Subject 2; for DBP,

it ranged from 1.10 mmHg in Subject 10 to 25.38 mmHg in Subject 2. Compared with the per-subject-mean baseline in Table 3, the 52-feature model achieved lower MAE for 10 of 16 subjects for SBP and 9 of 16 subjects for DBP. Run-to-run variability also differed across subjects, with the largest MAE SDs observed for Subject 2 in SBP (6.27 mmHg) and Subject 16 in DBP (5.25 mmHg).

Table 8. Per-subject SBP estimation performance of the LSTM-based model using the 52-dimensional rPPG-derived feature representation.

Subject No.	MAE	RMSE	SD	MD	R
1	9.41 ± 2.91	11.06 ± 2.94	6.74 ± 1.56	−8.58 ± 3.12	−0.25 ± 0.24
2	44.56 ± 6.27	45.82 ± 5.89	10.10 ± 2.86	−44.56 ± 6.27	0.60 ± 0.41
3	10.80 ± 2.24	11.19 ± 2.38	2.66 ± 1.46	−10.80 ± 2.24	0.52 ± 0.56
4	5.31 ± 2.01	5.84 ± 2.06	2.65 ± 0.96	−4.95 ± 2.45	−0.15 ± 0.33
5	8.91 ± 2.19	10.20 ± 1.72	4.83 ± 1.45	8.75 ± 2.25	0.54 ± 0.29
6	9.42 ± 2.24	11.52 ± 3.11	11.07 ± 2.47	−1.92 ± 3.23	−0.18 ± 0.27
7	8.49 ± 1.91	9.52 ± 1.84	4.13 ± 1.60	−8.39 ± 2.00	−0.45 ± 0.51
8	18.70 ± 1.61	23.52 ± 1.39	14.46 ± 0.92	−18.50 ± 1.88	−0.65 ± 0.21
9	3.94 ± 1.43	4.64 ± 1.56	2.37 ± 1.00	3.90 ± 1.46	0.12 ± 0.43
10	1.64 ± 0.95	2.06 ± 1.07	1.31 ± 0.69	−1.20 ± 1.34	0.40 ± 0.25
11	6.47 ± 2.31	11.35 ± 2.89	10.40 ± 2.84	3.35 ± 3.18	0.31 ± 0.31
12	3.73 ± 1.84	4.48 ± 2.05	2.99 ± 0.93	3.00 ± 2.36	−0.15 ± 0.59
13	17.48 ± 2.65	18.15 ± 2.60	4.55 ± 1.67	−17.48 ± 2.65	−0.69 ± 0.22
14	11.51 ± 2.41	13.83 ± 2.13	10.70 ± 1.02	−8.50 ± 2.89	−0.51 ± 0.27
15	5.18 ± 2.41	5.18 ± 2.41	0.00 ± 0.00	4.67 ± 3.34	–
16	23.01 ± 6.13	26.13 ± 4.96	11.86 ± 4.01	22.58 ± 6.49	0.67 ± 0.14

Table 9. Per-subject DBP estimation performance of the LSTM-based model using the 52-dimensional rPPG-derived feature representation.

Subject No.	MAE	RMSE	SD	MD	R
1	6.65 ± 1.51	7.93 ± 1.65	6.73 ± 1.41	3.15 ± 2.97	−0.42 ± 0.20
2	25.38 ± 4.89	27.51 ± 4.17	9.86 ± 3.20	−25.38 ± 4.89	0.73 ± 0.41
3	3.37 ± 0.93	3.62 ± 0.83	1.20 ± 0.45	−3.37 ± 0.93	−0.58 ± 0.56
4	4.37 ± 1.88	4.64 ± 1.88	1.60 ± 0.58	−4.22 ± 2.11	0.36 ± 0.48
5	17.71 ± 3.68	18.99 ± 3.31	6.11 ± 2.82	17.70 ± 3.68	0.44 ± 0.56
6	7.22 ± 2.12	9.34 ± 2.43	7.69 ± 1.57	−4.83 ± 2.91	0.27 ± 0.21
7	8.76 ± 1.56	8.98 ± 1.41	1.73 ± 0.83	−8.73 ± 1.63	−0.11 ± 0.62
8	8.39 ± 0.79	10.16 ± 0.70	5.76 ± 0.50	−8.34 ± 0.87	−0.19 ± 0.50
9	3.02 ± 0.93	3.96 ± 1.33	2.49 ± 1.25	2.98 ± 0.94	0.08 ± 0.35
10	1.10 ± 0.47	1.22 ± 0.44	0.53 ± 0.09	−1.02 ± 0.60	–
11	4.26 ± 0.79	7.09 ± 1.61	6.70 ± 1.49	2.08 ± 1.18	−0.23 ± 0.33
12	4.90 ± 0.98	5.23 ± 0.99	4.32 ± 1.00	−2.70 ± 1.23	−0.33 ± 0.34
13	7.97 ± 1.04	8.34 ± 1.03	2.40 ± 0.57	7.97 ± 1.04	0.56 ± 0.28
14	12.77 ± 3.79	16.42 ± 3.22	12.00 ± 1.04	−10.95 ± 3.90	−0.42 ± 0.31
15	3.37 ± 2.25	3.37 ± 2.25	0.00 ± 0.00	1.16 ± 3.95	–
16	21.87 ± 5.25	24.82 ± 3.77	10.71 ± 3.58	21.83 ± 5.26	–

Table 10 reports the number of distinct 40-cycle window-level BP labels in each subject’s chronological training, validation, and test split. For SBP, the counts ranged from 21 to 95 in training, 1 to 14 in validation, and 1 to 14 in testing; the corresponding DBP ranges were 20–85, 1–11, and 1–14. Subject 2 had only one distinct SBP validation label, Subject 15 had only one distinct test label for both SBP and DBP, and Subjects 10 and 16 each had only one distinct DBP test label. These limited-label cases help explain why correlation is undefined or weakly informative for some subject-specific evaluations and why the measurement-based baseline can occasionally obtain near-zero error.

Table 10. Number of distinct window-level reference BP labels in each subject’s training, validation, and test sets.

Subject No.	SBP			DBP		
	Train	Validation	Test	Train	Validation	Test
1	73	11	10	59	6	11
2	47	1	5	44	4	4
3	28	2	3	27	2	3
4	43	4	4	40	4	3
5	37	4	4	28	4	4
6	68	8	8	58	8	8
7	42	4	4	40	4	3
8	95	14	14	85	11	14
9	56	6	6	47	6	6
10	68	13	5	50	10	1
11	82	6	12	64	11	12
12	29	3	4	27	3	4
13	54	5	5	58	7	5
14	93	11	13	63	8	12
15	21	1	1	20	1	1
16	35	4	4	28	4	1

To further assess whether the pooled correlation reflected within-subject BP tracking, the reference and estimated BP values were mean-centered within each subject before calculating the pooled correlation. Across the 20 independent runs, the subject-mean-centered R was -0.038 ± 0.101 for SBP and -0.077 ± 0.069 for DBP. The near-zero centered correlations indicate that within-subject BP tracking could not be clearly demonstrated in the present dataset. This result should be interpreted in the context of the limited within-subject BP variability observed in several test sets. Notably, the number of distinct reference BP labels may overstate the effective BP variability because decimal-level differences were counted as separate labels; thus, multiple distinct labels could still correspond to a relatively narrow BP range. Consequently, the restricted dynamic range within individual subjects may have limited the ability of the correlation analysis to capture short-term BP tracking performance.

Figure 5 presents the regression and Bland–Altman plots for the representative run using the 52-dimensional feature representation. The regression plots show positive correlations of $R = 0.43$ for SBP and $R = 0.44$ for DBP. In the Bland–Altman analysis, the bias (mean difference) was -6.04 mmHg for SBP and -1.63 mmHg for DBP, with 95% limits of agreement (LoA) ranging from -37.03 to 24.95 mmHg and from -24.23 to 20.97 mmHg, respectively. Because the differences were calculated as reference minus estimated BP, the negative bias indicates a tendency toward overestimation. The relatively wide LoA, particularly for SBP, indicate considerable variation in the estimation errors across individual test samples.

Table 11 summarizes the subject-wise SBP results under LOSO validation. Mean MAE ranged from 7.27 mmHg for Subject 15 to 39.42 mmHg for Subject 16, with a mean subject-wise MAE of 15.88 mmHg. Subjects 14 and 16 also showed large absolute MD values of 34.71 and 39.42 mmHg. R ranged from -0.63 to 0.37 , with negative values in 9 of 16 held-out subjects. In contrast to this large between-subject spread, the repeated-run SD of MAE within each LOSO fold did not exceed 2.09 mmHg, indicating that the dominant variation was across held-out subjects rather than across repeated training runs.

Table 11. Per-subject SBP estimation performance under LOSO validation using the 52-dimensional rPPG-derived feature representation.

Subject No.	MAE	RMSE	SD	MD	R
1	11.83 ± 0.97	13.99 ± 1.11	10.03 ± 0.77	9.75 ± 1.03	−0.56 ± 0.15
2	23.46 ± 2.09	28.58 ± 2.80	16.90 ± 2.16	−23.02 ± 2.10	−0.26 ± 0.10
3	16.88 ± 1.62	18.49 ± 1.76	15.89 ± 1.03	9.15 ± 3.04	0.34 ± 0.15
4	7.50 ± 0.29	9.30 ± 0.35	9.12 ± 0.33	−1.76 ± 0.46	−0.11 ± 0.03
5	9.94 ± 1.17	11.13 ± 1.02	5.04 ± 0.29	9.90 ± 1.18	0.37 ± 0.13
6	12.08 ± 0.71	15.41 ± 0.96	12.46 ± 0.59	−8.98 ± 1.62	0.17 ± 0.10
7	18.21 ± 1.09	22.21 ± 0.70	15.06 ± 0.80	16.29 ± 1.07	−0.49 ± 0.11
8	11.24 ± 1.04	13.68 ± 0.85	8.01 ± 0.24	−11.06 ± 1.25	0.02 ± 0.12
9	9.28 ± 1.57	12.25 ± 1.34	11.23 ± 0.84	4.66 ± 1.92	−0.09 ± 0.06
10	11.28 ± 1.64	13.74 ± 1.57	13.05 ± 1.71	−3.99 ± 1.61	−0.54 ± 0.08
11	18.33 ± 1.38	19.60 ± 1.44	13.72 ± 0.97	−13.97 ± 1.35	−0.63 ± 0.07
12	9.48 ± 0.80	13.13 ± 0.56	9.31 ± 0.76	9.19 ± 1.02	−0.18 ± 0.04
13	12.09 ± 0.11	15.00 ± 0.76	14.76 ± 0.84	2.61 ± 0.52	0.26 ± 0.05
14	35.73 ± 0.85	37.31 ± 0.70	13.67 ± 0.68	−34.71 ± 0.84	−0.31 ± 0.03
15	7.27 ± 1.33	9.16 ± 1.88	7.30 ± 1.50	5.46 ± 1.51	0.27 ± 0.14
16	39.42 ± 1.31	39.50 ± 1.30	2.43 ± 0.64	39.42 ± 1.31	0.01 ± 0.10

Table 12 shows a similarly broad range of DBP performance under LOSO validation. Mean MAE ranged from 4.83 mmHg for Subject 15 to 32.10 mmHg for Subject 14, with a mean subject-wise MAE of 12.82 mmHg. Subjects 14 and 16 had MAEs above 30 mmHg and absolute MD values of 31.93 and 31.20 mmHg, respectively; R ranged from −0.49 to 0.55, with negative values in 10 of 16 held-out subjects. The maximum repeated-run SD of MAE within a fold was 2.18 mmHg, again much smaller than the between-subject error range and supporting the conclusion that unseen-subject identity was the larger source of performance variation.

Table 12. Per-subject DBP estimation performance under LOSO validation using the 52-dimensional rPPG-derived feature representation.

Subject No.	MAE	RMSE	SD	MD	R
1	17.74 ± 1.13	18.21 ± 1.09	4.27 ± 0.65	17.69 ± 1.21	−0.13 ± 0.05
2	15.07 ± 1.83	17.43 ± 1.88	9.93 ± 1.43	−14.24 ± 2.17	−0.10 ± 0.16
3	15.52 ± 2.12	18.42 ± 1.83	10.18 ± 0.31	15.30 ± 2.27	0.26 ± 0.19
4	17.24 ± 1.01	18.42 ± 0.96	6.47 ± 0.34	−17.24 ± 1.01	0.20 ± 0.04
5	9.13 ± 0.94	12.02 ± 0.97	8.30 ± 0.70	8.68 ± 1.03	−0.42 ± 0.02
6	16.30 ± 1.08	18.69 ± 1.02	9.53 ± 0.12	−16.06 ± 1.21	−0.05 ± 0.10
7	7.83 ± 0.40	10.74 ± 0.68	10.16 ± 0.60	3.46 ± 0.74	−0.33 ± 0.04
8	6.46 ± 2.18	7.53 ± 2.12	6.36 ± 0.52	−2.70 ± 3.93	−0.10 ± 0.10
9	6.11 ± 0.87	7.84 ± 1.03	4.87 ± 0.95	6.10 ± 0.88	0.09 ± 0.06
10	6.23 ± 1.13	10.23 ± 1.42	9.35 ± 1.10	3.97 ± 1.58	−0.22 ± 0.05
11	5.96 ± 0.76	8.32 ± 0.53	7.41 ± 0.51	3.09 ± 2.47	−0.35 ± 0.06
12	7.24 ± 0.96	7.68 ± 0.87	3.90 ± 0.28	−6.60 ± 0.98	0.19 ± 0.06
13	6.11 ± 0.67	8.28 ± 1.71	8.11 ± 1.52	1.47 ± 1.20	0.06 ± 0.16
14	32.10 ± 0.69	32.72 ± 0.71	7.14 ± 0.33	−31.93 ± 0.71	−0.49 ± 0.03
15	4.83 ± 0.42	5.93 ± 0.46	5.69 ± 0.59	−0.93 ± 1.51	0.55 ± 0.10
16	31.20 ± 0.99	31.26 ± 0.98	1.92 ± 0.41	31.20 ± 0.99	−0.15 ± 0.26

4.2. Effectiveness of LSTM-Based Estimation

Tables 13 and 14 compare the full recurrent model with MLP, SVR, and three recurrent ablations. To examine how the individual recurrent components affect BP estimation performance, the full BiLSTM–UniLSTM–attention architecture was retained as the reference configuration, and the effects of removing attention, BiLSTM, or both components were

evaluated comparatively. Model capacity is reported using the number of trainable parameters for all neural network-based models, whereas SVR capacity is represented by the number of support vectors because its fitted model is determined by the subset of training samples retained as support vectors. The tables report the 20-run mean performance values together with the corresponding run-to-run SDs.

Table 13. Comparison of model architectures for SBP estimation using rPPG-derived features.

Model	Capacity	MAE	RMSE	SD	MD	R
Full LSTM (proposed)	1,669,026	11.45 ± 1.07	16.78 ± 1.12	15.48 ± 0.94	−6.39 ± 1.31	0.44 ± 0.08
MLP	1,243,170	12.73 ± 1.16	17.21 ± 1.01	16.37 ± 0.77	4.89 ± 2.23	0.24 ± 0.06
SVR	1123 SV	13.40 ± 0.00	16.69 ± 0.00	15.79 ± 0.00	−5.41 ± 0.00	0.33 ± 0.00
LSTM (w/o attention)	1,668,962	11.31 ± 0.81	16.59 ± 1.17	15.53 ± 1.11	−5.78 ± 0.97	0.45 ± 0.08
LSTM (w/o BiLSTM)	565,154	11.55 ± 0.73	16.22 ± 0.98	14.83 ± 0.89	−6.51 ± 0.96	0.50 ± 0.06
LSTM (w/o attention, BiLSTM)	565,090	11.72 ± 0.92	16.79 ± 1.24	15.52 ± 1.10	−6.37 ± 0.99	0.44 ± 0.08

Table 14. Comparison of model architectures for DBP estimation using rPPG-derived features.

Model	Capacity	MAE	RMSE	SD	MD	R
Full LSTM (proposed)	1,669,026	8.15 ± 0.81	11.90 ± 0.99	11.74 ± 0.93	−1.81 ± 0.76	0.40 ± 0.12
MLP	1,243,170	9.93 ± 0.70	12.26 ± 0.56	11.12 ± 0.32	5.04 ± 1.26	0.38 ± 0.05
SVR	1117 SV	8.70 ± 0.00	11.18 ± 0.00	11.16 ± 0.00	−0.59 ± 0.00	0.40 ± 0.00
LSTM (w/o attention)	1,668,962	7.99 ± 0.61	11.70 ± 0.91	11.57 ± 0.86	−1.50 ± 0.95	0.42 ± 0.10
LSTM (w/o BiLSTM)	565,154	8.24 ± 0.61	11.68 ± 0.76	11.53 ± 0.76	−1.66 ± 0.75	0.42 ± 0.08
LSTM (w/o attention, BiLSTM)	565,090	8.30 ± 0.68	11.87 ± 0.92	11.66 ± 0.89	−2.02 ± 1.00	0.40 ± 0.10

For SBP, the full recurrent model achieved an MAE of 11.45 mmHg and R of 0.44, compared with MAEs of 12.73 and 13.40 mmHg for MLP and SVR. Removing attention slightly reduced MAE to 11.31 mmHg, while removing the BiLSTM increased MAE to 11.55 mmHg but improved RMSE, error SD, and R. For DBP, removing attention likewise produced a slightly lower MAE, whereas removing the BiLSTM increased MAE to 8.23 mmHg; removing both components yielded the highest recurrent-ablation MAE for both SBP and DBP. Removing the attention layer slightly improved performance while having little effect on the number of parameters, whereas removing the BiLSTM reduced the recurrent-model capacity from approximately 1.67M to 565k parameters. No single component consistently improved all evaluation metrics. However, the BiLSTM accounted for most of the recurrent model's parameter count, and removing both the BiLSTM and attention components generally increased MAE.

4.3. Comparison with Previous Camera-Based BP Estimation Studies

Table 15 summarizes previously reported camera-based BP estimation results. Schrumppf et al. [36] used the same publicly released rPPG dataset analyzed in the present study. They first trained AlexNet and ResNet using contact PPG data from 5000 subjects and subsequently transferred the models to the rPPG data. Fourteen subjects were included in their final rPPG analysis, and their calibration-based results are listed here. Kawasaki et al. [39] collected an independent dataset from 10 healthy participants in their early 20 s and evaluated green channel, NIR channel, and combined green and NIR rPPG features using SVR with LOSO validation. Since the present study derives rPPG from RGB video, the green channel result was selected for comparison. Rong and Li [38] collected data

from 191 volunteers, of whom 189 were included after screening, and extracted 26 features from green channel IPPG signals.

Table 15. Comparison with previously reported camera-based BP estimation studies. All error metrics are reported in mmHg.

Study	Subjects	Input/Setting	Model	SBP		DBP	
				MAE	SD	MAE	SD
Schrumpf et al. [36]	5000 (PPG)	rPPG transfer learning from contact PPG;	AlexNet	15.20	9.11	8.52	4.92
	14 (rPPG)	calibration-based	ResNet	12.51	12.61	8.30	9.84
Kawasaki et al. [39]	10	Green channel rPPG; LOSO	SVR	11.31	–	6.49	–
Rong and Li [38]	189	Green channel rPPG; 7:3 split	SVR	9.97	3.35	7.59	2.58

Kawasaki et al. used data from 10 healthy participants in their early 20 s, with mean SBP/DBP values of 111.2/70.2 mmHg at rest and 129.0/75.6 mmHg after exercise. Rong and Li used a substantially larger dataset, including 189 analyzed participants, with SBP ranging from 92 to 148 mmHg and DBP from 51 to 102 mmHg. Among the previous studies listed in Table 15, the SVR model reported by Rong and Li showed the lowest MAE for both SBP and DBP. These results suggest that the amount of available data and the diversity of the BP distribution can have a substantial influence on estimation performance.

Compared with the 52-feature results in Tables 5 and 6, the present study achieved lower MAE than the AlexNet and ResNet results reported by Schrumpf et al. for both BP types, despite their models being pretrained using contact PPG data from the large MIMIC-III cohort before being transferred to the same rPPG dataset used in the present study. This result suggests that explicitly extracted morphological features may provide an effective representation of BP-related information in rPPG, particularly when the amount of available rPPG data is limited. The SBP MAE of the present study was also similar to the green channel result reported by Kawasaki et al., whereas its DBP MAE was higher. Overall, the comparison indicates that both dataset characteristics and input representation are important factors affecting BP estimation performance.

4.4. Dimensionality Reduction Analysis

The effect of feature dimensionality was evaluated using PCA and RFE at 5, 10, 15, 20, 24, 30, and 40 dimensions under the same 20-run protocol. Tables 16 and 17 show the PCA results. For SBP, 24 dimensions provided the best overall combination of metrics, with MAE 10.70 ± 0.60 mmHg, RMSE 15.24 ± 0.88 mmHg, error SD 14.13 ± 0.63 mmHg, MD -5.61 ± 1.25 mmHg, and R 0.58 ± 0.03 . For DBP, the minimum mean MAE of 7.55 mmHg occurred at 15, 24, and 40 dimensions, while 24 dimensions gave the lowest RMSE (10.70 ± 0.58 mmHg) and error SD (10.61 ± 0.57 mmHg). Across PCA settings, the run-to-run SD of MAE remained at or below 0.82 mmHg for SBP and 0.53 mmHg for DBP, indicating relatively stable dimensionality comparisons across seeds.

Table 16. SBP estimation performance across different PCA-reduced feature dimensions.

Dim.	MAE	RMSE	SD	MD	R
5	11.06 ± 0.69	15.65 ± 0.90	14.62 ± 0.68	-5.46 ± 1.23	0.55 ± 0.04
10	10.92 ± 0.62	15.62 ± 0.78	14.49 ± 0.53	-5.72 ± 1.28	0.56 ± 0.03
15	11.09 ± 0.82	15.72 ± 1.03	14.40 ± 0.77	-6.25 ± 1.15	0.57 ± 0.04
20	11.23 ± 0.54	15.94 ± 0.71	14.64 ± 0.56	-6.22 ± 1.19	0.56 ± 0.03
24	10.70 ± 0.60	15.24 ± 0.88	14.13 ± 0.63	-5.61 ± 1.25	0.58 ± 0.03
30	10.98 ± 0.75	15.67 ± 1.01	14.48 ± 0.57	-5.85 ± 1.61	0.56 ± 0.03
40	10.88 ± 0.53	15.52 ± 0.75	14.29 ± 0.74	-6.00 ± 0.87	0.58 ± 0.04

Table 17. DBP estimation performance across different PCA-reduced feature dimensions.

Dim.	MAE	RMSE	SD	MD	R
5	8.02 ± 0.40	11.28 ± 0.55	11.18 ± 0.54	-1.22 ± 0.85	0.48 ± 0.05
10	7.63 ± 0.41	10.89 ± 0.55	10.80 ± 0.53	-1.22 ± 0.67	0.51 ± 0.05
15	7.55 ± 0.53	10.76 ± 0.73	10.63 ± 0.70	-1.51 ± 0.73	0.53 ± 0.06
20	7.66 ± 0.36	11.01 ± 0.52	10.86 ± 0.51	-1.62 ± 0.79	0.51 ± 0.05
24	7.55 ± 0.41	10.70 ± 0.58	10.61 ± 0.57	-1.22 ± 0.71	0.53 ± 0.05
30	7.58 ± 0.36	10.84 ± 0.54	10.73 ± 0.49	-1.39 ± 0.70	0.52 ± 0.05
40	7.55 ± 0.42	10.76 ± 0.58	10.65 ± 0.58	-1.38 ± 0.79	0.53 ± 0.05

The RFE results in Tables 18 and 19 showed a different dimensional dependence. For SBP, the lowest mean MAE was obtained at 40 dimensions (11.10 ± 0.79 mmHg), while the highest R occurred at 10 dimensions (0.54 ± 0.06); the 24-dimensional setting produced the highest MAE (12.29 ± 0.82 mmHg) and lowest R (0.39 ± 0.06). For DBP, the 5-dimensional setting yielded the lowest MAE (7.57 ± 0.34 mmHg) and RMSE (10.93 ± 0.43 mmHg) and the highest R (0.51 ± 0.04). The run-to-run SD of MAE remained below 1 mmHg across all RFE settings, suggesting that the estimation performance was reasonably consistent across seeds.

Table 18. SBP estimation performance across different RFE-selected feature dimensions.

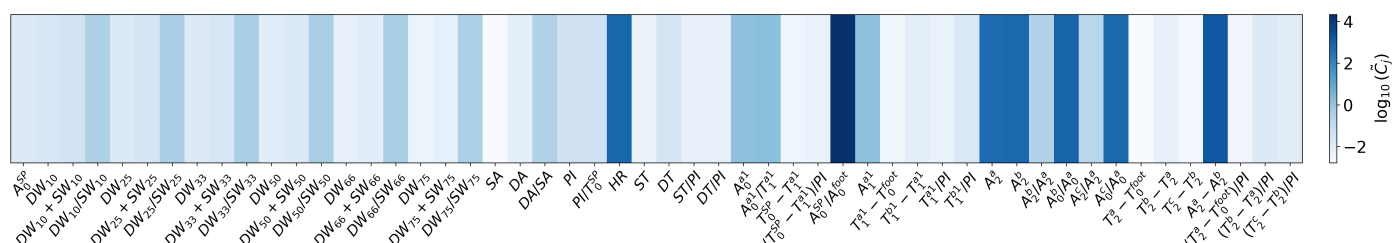
Dim.	MAE	RMSE	SD	MD	R
5	11.50 ± 0.43	16.71 ± 0.55	15.49 ± 0.42	-6.21 ± 0.85	0.47 ± 0.03
10	11.49 ± 0.66	16.30 ± 0.88	14.40 ± 0.88	-7.57 ± 1.06	0.54 ± 0.06
15	11.49 ± 0.79	16.70 ± 0.72	14.87 ± 0.57	-7.56 ± 0.84	0.49 ± 0.06
20	11.65 ± 0.80	16.90 ± 0.85	15.09 ± 0.76	-7.57 ± 1.00	0.48 ± 0.06
24	12.29 ± 0.82	17.43 ± 0.89	16.00 ± 0.78	-6.81 ± 1.21	0.39 ± 0.06
30	11.12 ± 0.98	16.38 ± 1.21	14.96 ± 1.12	-6.60 ± 1.12	0.50 ± 0.08
40	11.10 ± 0.79	16.10 ± 1.30	14.94 ± 1.24	-5.97 ± 0.76	0.49 ± 0.10

Table 19. DBP estimation performance across different RFE-selected feature dimensions.

Dim.	MAE	RMSE	SD	MD	R
5	7.57 ± 0.34	10.93 ± 0.43	10.78 ± 0.40	-1.68 ± 0.64	0.51 ± 0.04
10	7.76 ± 0.65	11.00 ± 0.91	10.76 ± 0.83	-2.14 ± 0.91	0.50 ± 0.08
15	7.82 ± 0.52	11.26 ± 0.61	10.91 ± 0.54	-2.54 ± 1.19	0.48 ± 0.08
20	8.02 ± 0.41	11.47 ± 0.59	11.23 ± 0.50	-2.14 ± 0.96	0.45 ± 0.07
24	8.64 ± 0.55	12.47 ± 0.61	12.21 ± 0.57	-2.38 ± 0.95	0.36 ± 0.07
30	7.82 ± 0.58	11.18 ± 0.70	10.98 ± 0.73	-1.66 ± 1.31	0.48 ± 0.08
40	7.87 ± 0.54	11.38 ± 0.88	11.23 ± 0.86	-1.74 ± 0.64	0.46 ± 0.10

At the directly matched 24-dimensional setting, PCA outperformed RFE across all reported mean metrics, reducing MAE from 12.29 to 10.70 mmHg for SBP and from 8.64 to 7.55 mmHg for DBP. However, the relative advantage depended on BP type and dimensionality; for example, at five dimensions RFE achieved a lower DBP MAE than PCA. Overall, PCA showed more consistently favorable performance, particularly for SBP. Although the 24-dimensional PCA representation produced the strongest numerical performance for several metrics, the 15-dimensional representation retained competitive SBP and DBP performance with a substantially smaller input dimension and was therefore selected as the final reduced representation. Interestingly, 15 dimensions also corresponded to the first point at which the CEV reached approximately 0.99, suggesting the possibility of using a CEV-based threshold as a quantitative criterion for selecting the reduced dimension.

Additionally, to examine how strongly each original feature is represented in the retained PCA space, we visualized the feature contributions described in Section 2.3. Figure 7 shows that the original morphological features were not represented equally within the retained PCA space. Among the 52 features, A_0^{SP} / A_0^{foot} exhibited the largest contribution, while relatively high values were also observed for $A_2^a - A_2^b$, A_0^b / A_0^a , heart rate, A_2^b , A_0^c / A_0^a , and A_2^a . These results suggest that amplitude-related and derivative-based morphological characteristics were more strongly represented in the variance preserved by the retained principal components, whereas several width-based and temporal-ratio features showed comparatively lower contributions. These contribution values indicate how strongly each original feature is represented in the PCA-preserved variance of the present dataset and do not directly indicate its importance for BP estimation. The contribution pattern may also change when a different dataset or population is used, as PCA is determined by the variance and covariance structure of the input data.



5. Discussion

The first major finding is that the structured morphological feature representation was more effective than directly using the resampled rPPG waveform under the same sequential architecture and training protocol. The present estimation accuracy did not satisfy the BHS

Grade C or AAMI/ESH/ISO numerical criteria; however, these criteria are included only as clinical references because the present study does not meet the requirements for formal BP device validation. Nevertheless, the results demonstrated the feasibility of morphological feature-based BP estimation. This finding was further supported by comparison with previous camera-based BP studies, where the present feature-based approach showed favorable performance relative to the AlexNet and ResNet models of Schrumppf et al., despite their use of large-scale contact-PPG pretraining before transfer to the same rPPG dataset. The baseline analyses provided a more cautious interpretation: weak subject-mean-centered correlations and the strong measurement-based last-known-BP reference indicate that reliable dynamic within-subject tracking was not achieved. Even so, the proposed model outperformed the global-mean baseline and the per-subject-mean baseline for a majority of subjects, suggesting that the estimated BP was not determined solely by overall or subject-specific average BP levels. These findings support the feasibility of extracting additional BP-related information from morphological rPPG features.

The model comparisons further showed that the effectiveness of the morphological feature representation depends on what model is used. The recurrent deep-learning models generally showed more favorable estimation performance than the conventional MLP and SVR models, supporting the use of sequential modeling for the cycle-level feature sequences. However, the full recurrent architecture was not uniformly superior. Removing the attention layer slightly improved several metrics, whereas removing the BiLSTM produced mixed changes while substantially reducing the number of model parameters. Removing both components generally degraded MAE compared with the other recurrent configurations. These findings indicate that increasing architectural complexity does not necessarily improve BP estimation and that model selection should consider both estimation performance and model capacity. In this study, the comparison of conventional models and recurrent ablations therefore served not only to justify the use of sequential learning, but also to examine which architectural choices are appropriate for the proposed feature representation.

The dimensionality analysis provided another important finding regarding the construction of an effective morphological feature representation. PCA and RFE showed different behavior across BP types and feature dimensions, indicating that the choice of reduced representation should be evaluated rather than assumed. In the present study, PCA provided more consistently favorable results, and the 15-dimensional representation was selected as a compact input while maintaining competitive estimation performance. The fact that this dimension also corresponded to the first CEV value of approximately 0.99 suggests that a variance-based criterion may be useful when selecting an effective feature dimension. Overall, these findings indicate that retaining all original features is not necessarily beneficial and that an appropriate feature-space size should be considered when using high-dimensional morphological descriptors. The feature-contribution analysis further showed how the original features were represented within the retained PCA space, although these contributions are dataset-dependent and do not directly represent their importance for BP estimation.

The baseline analysis also suggests a useful direction for evaluating feature-based BP estimation methods. The global-mean and per-subject-mean baselines were used as measurement-free measurement-based references to determine whether a model performs beyond the overall BP distribution or the average BP level of an individual. In contrast, the last-known-BP baseline was treated as a measurement-based reference because it directly requires a specific previously measured BP value from the same subject. These references therefore evaluate different aspects of the estimation problem. Measurement-based baselines help determine whether a model is doing more than reproducing an average BP level,

whereas the last-known-BP reference evaluates whether the model can provide information beyond the temporal persistence of a recent BP measurement. The strong performance of the latter is understandable because BP often changes gradually over short periods; however, a fixed previous measurement cannot reflect a subsequent increase or decrease until a new BP measurement becomes available. A signal-based estimator is ultimately intended to capture such changes directly from physiological information. The present model did not demonstrate this capability sufficiently, but this comparison shows why measurement-free and measurement-based references should be distinguished rather than interpreted as equivalent baselines. Together with conventional absolute accuracy criteria such as BHS and AAMI/ESH/ISO, reporting these simple data-dependent references may provide a more informative basis for interpreting BP estimation performance across different datasets and experimental protocols.

Several limitations should be considered when interpreting these findings. The number of subjects was limited, and the LOSO results showed that performance varied considerably across unseen subjects, indicating that subject-independent generalization remains substantially more difficult than calibration-based estimation. More importantly, the dataset contained limited diversity in within-subject BP changes. Although the overall BP range was not extremely narrow, several subjects exhibited only small temporal BP variations, providing relatively few examples from which the model could learn substantial increases or decreases in BP. This limitation may also help explain the strong performance of the last-known-BP reference. The comparison with previous studies further supports the importance of dataset characteristics: Rong and Li obtained favorable performance using a relatively simple SVR model but trained and evaluated it on a considerably larger cohort with a broad BP distribution. Although differences in experimental protocols prevent a direct attribution of the performance difference to dataset size alone, this comparison suggests that the amount and diversity of BP data can be at least as important as model complexity. In addition, the strict morphological-feature extraction procedure excluded a large proportion of cardiac cycles and may have biased the usable data toward cleaner pulse morphologies. Future work should therefore focus on larger datasets containing broader and more diverse within-subject BP changes, more robust morphological-feature extraction with reduced data loss, and evaluation against both measurement-free and measurement-based references. Such evaluations will be necessary to determine whether morphological rPPG features can support reliable subject-independent estimation and, more importantly, capture meaningful BP changes over time.

6. Conclusions

This study evaluated a sequential deep learning framework for contactless BP estimation using 52 morphological features extracted from rPPG, rPPG', and rPPG''. Under the temporally separated calibration-based split, the unreduced feature model achieved mean MAEs of 11.45 mmHg for SBP and 8.15 mmHg for DBP and outperformed direct rPPG signal input across MAE, RMSE, error SD, and correlation. It also improved MAE relative to the global-mean baseline, although improvements over the per-subject-mean baseline were not consistent across all subjects.

Across the dimensionality-reduction experiments, the 24-dimensional PCA representation produced the strongest overall numerical performance for several metrics, with MAEs of 10.70 mmHg for SBP and 7.55 mmHg for DBP. The 15-dimensional PCA representation retained approximately 99% of the cumulative explained variance and competitive performance for both BP targets while using fewer dimensions, and was therefore selected as the final compact representation. The recurrent ablation results were metric-dependent, while LOSO validation showed broad subject-wise error ranges and large biases in several

held-out subjects. Overall, the results support PCA-compressed handcrafted rPPG features as a structured representation for contactless BP estimation, while larger datasets with greater within-subject BP variation are required to establish robust generalization and dynamic BP tracking.

Author Contributions: Conceptualization, H.J., D.P., and N.K.K.; methodology, H.J., D.P., and N.K.K.; software, H.J. and D.P.; validation, H.J., D.P., and N.K.K.; formal analysis, H.J.; investigation, H.J.; resources, D.P.; data curation, H.J. and D.P.; writing—original draft preparation, H.J.; writing—review and editing, H.J., S.Y.L. and N.K.K.; visualization, H.J.; supervision, S.Y.L. and N.K.K.; project administration, N.K.K.; funding acquisition, H.J. and N.K.K. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the National Research Foundation of Korea (NRF) grant funded by the Korean government (MSIT) (No. RS-2023-00219725). Also, this research was supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (No. RS-2025-25431392).

Data Availability Statement: The rPPG dataset analyzed in this study is openly available in the public GitHub repository associated with the original study at <https://github.com/Fabian-Sc85/non-invasive-bp-estimation-using-deep-learning>, accessed on 20 July 2026.

Acknowledgments: During the preparation of this manuscript, the authors used ChatGPT-5.5 (OpenAI) primarily for English-language editing, including grammar correction and stylistic refinement, and to a limited extent for assistance in interpreting and discussing the study results. The authors have reviewed and edited the output and take full responsibility for the content of this publication.

Conflicts of Interest: The authors declare no conflicts of interest.

References

- Chen, G.; Zou, L.; Ji, Z. A review: Blood pressure monitoring based on PPG and circadian rhythm. *Apl Bioeng.* **2024**, *8*, 31501. [[CrossRef](#)] [[PubMed](#)]
- Mills, K.T.; Stefanescu, A.; He, J. The global epidemiology of hypertension. *Nat. Rev. Nephrol.* **2020**, *16*, 223–237. [[CrossRef](#)] [[PubMed](#)]
- Mukkamala, R.; Stergiou, G.S.; Avolio, A.P. Cuffless blood pressure measurement. *Annu. Rev. Biomed. Eng.* **2022**, *24*, 203–230. [[CrossRef](#)] [[PubMed](#)]
- Froehler, M.T.; Chitale, R.; Magarik, J.A.; Fusco, M.R. Comparison of a pressure-sensing sheath and radial arterial line for intraoperative blood pressure monitoring in neurointerventional procedures. *J. Neurointerv. Surg.* **2018**, *10*, 784–787. [[CrossRef](#)] [[PubMed](#)]
- Nuttall, G.; Burckhardt, J.; Hadley, A.; Kane, S.; Kor, D.; Marienau, M.S.; Schroeder, D.R.; Handlogten, K.; Wilson, G.; Oliver, W.C. Surgical and patient risk factors for severe arterial line complications in adults. *Anesthesiology* **2016**, *124*, 590–597. [[CrossRef](#)] [[PubMed](#)]
- Hill, B.L.; Rakocz, N.; Rudas, Á.; Chiang, J.N.; Wang, S.; Hofer, I.; Cannesson, M.; Halperin, E. Imputation of the continuous arterial line blood pressure waveform from non-invasive measurements using deep learning. *Sci. Rep.* **2021**, *11*, 15755. [[CrossRef](#)] [[PubMed](#)]
- Ding, X.; Yan, B.P.; Zhang, Y.T.; Liu, J.; Zhao, N.; Tsang, H.K. Pulse transit time based continuous cuffless blood pressure estimation: A new extension and a comprehensive evaluation. *Sci. Rep.* **2017**, *7*, 11554. [[CrossRef](#)] [[PubMed](#)]
- Zhang, Y.L.; Zheng, Y.Y.; Ma, Z.C.; Sun, Y.N. Radial pulse transit time is an index of arterial stiffness. *Hypertens. Res.* **2011**, *34*, 884–887. [[CrossRef](#)] [[PubMed](#)]
- Esmaili, A.; Kachuee, M.; Shabany, M. Nonlinear cuffless blood pressure estimation of healthy subjects using pulse transit time and arrival time. *IEEE Trans. Instrum. Meas.* **2017**, *66*, 3299–3308. [[CrossRef](#)]
- Fung, P.; Dumont, G.; Ries, C.; Mott, C.; Ansermino, M. Continuous noninvasive blood pressure measurement by pulse transit time. In *Proceedings of the 26th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*; IEEE: Piscataway, NJ, USA, 2004; Volume 1, pp. 738–741.
- Martin, S.L.O.; Carek, A.M.; Kim, C.S.; Ashouri, H.; Inan, O.T.; Hahn, J.O.; Mukkamala, R. Weighing scale-based pulse transit time is a superior marker of blood pressure than conventional pulse arrival time. *Sci. Rep.* **2016**, *6*, 39273. [[CrossRef](#)] [[PubMed](#)]

12. Gesche, H.; Grosskurth, D.; Küchler, G.; Patzak, A. Continuous blood pressure measurement by using the pulse transit time: Comparison to a cuff-based method. *Eur. J. Appl. Physiol.* **2012**, *112*, 309–315. [[CrossRef](#)] [[PubMed](#)]
13. Finnegan, E.; Davidson, S.; Harford, M.; Jorge, J.; Watkinson, P.; Young, D.; Tarassenko, L.; Villarroel, M. Pulse arrival time as a surrogate of blood pressure. *Sci. Rep.* **2021**, *11*, 22767. [[CrossRef](#)] [[PubMed](#)]
14. Mukkamala, R.; Hahn, J.O.; Inan, O.T.; Mestha, L.K.; Kim, C.S.; Toreyin, H.; Kyal, S. Toward ubiquitous blood pressure monitoring via pulse transit time: theory and practice. *IEEE Trans. Biomed. Eng.* **2015**, *62*, 1879–1901. [[CrossRef](#)] [[PubMed](#)]
15. Nitzan, M.; Khanokh, B.; Slovik, Y. The difference in pulse transit time to the toe and finger measured by photoplethysmography. *Physiol. Meas.* **2001**, *23*, 85. [[CrossRef](#)] [[PubMed](#)]
16. Elgendi, M.; Fletcher, R.; Liang, Y.; Howard, N.; Lovell, N.H.; Abbott, D.; Lim, K.; Ward, R. The use of photoplethysmography for assessing hypertension. *Npj Digit. Med.* **2019**, *2*, 60. [[CrossRef](#)] [[PubMed](#)]
17. Nabeel, P.; Jayaraj, J.; Mohanasankar, S. Single-source PPG-based local pulse wave velocity measurement: A potential cuffless blood pressure estimation technique. *Physiol. Meas.* **2017**, *38*, 2122. [[CrossRef](#)] [[PubMed](#)]
18. Teng, X.; Zhang, Y. Continuous and noninvasive estimation of arterial blood pressure using a photoplethysmographic approach. In *Proceedings of the 25th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (IEEE Cat. No. 03CH37439)*; IEEE: Piscataway, NJ, USA, 2003; Volume 4, pp. 3153–3156.
19. El-Hajj, C.; Kyriacou, P.A. A review of machine learning techniques in photoplethysmography for the non-invasive cuff-less measurement of blood pressure. *Biomed. Signal Process. Control* **2020**, *58*, 101870. [[CrossRef](#)]
20. Allen, J. Photoplethysmography and its application in clinical physiological measurement. *Physiol. Meas.* **2007**, *28*, R1. [[CrossRef](#)] [[PubMed](#)]
21. González, S.; Hsieh, W.T.; Chen, T.P.C. A benchmark for machine-learning based non-invasive blood pressure estimation using photoplethysmogram. *Sci. Data* **2023**, *10*, 149. [[CrossRef](#)] [[PubMed](#)]
22. Gupta, S.; Singh, A.; Sharma, A.; Tripathy, R.K. Higher order derivative-based integrated model for cuff-less blood pressure estimation and stratification using PPG signals. *IEEE Sens. J.* **2022**, *22*, 22030–22039. [[CrossRef](#)]
23. Pankaj, Kumar, A.; Kumar, M.; Komaragiri, R. Optimized deep neural network models for blood pressure classification using Fourier analysis-based time–frequency spectrogram of photoplethysmography signal. *Biomed. Eng. Lett.* **2023**, *13*, 739–750. [[CrossRef](#)] [[PubMed](#)]
24. Martinez-Ríos, E.; Montesinos, L.; Alfaro-Ponce, M.; Pecchia, L. A review of machine learning in hypertension detection and blood pressure estimation based on clinical and physiological data. *Biomed. Signal Process. Control* **2021**, *68*, 102813. [[CrossRef](#)]
25. Park, J.; Seok, H.S.; Kim, S.S.; Shin, H. Photoplethysmogram analysis and applications: An integrative review. *Front. Physiol.* **2022**, *12*, 808451. [[CrossRef](#)] [[PubMed](#)]
26. Khalid, S.; Khalil, T.; Nasreen, S. A survey of feature selection and feature extraction techniques in machine learning. In *Proceedings of the 2014 Science and Information Conference*; IEEE: Piscataway, NJ, USA, 2014; pp. 372–378.
27. El-Hajj, C.; Kyriacou, P.A. Cuffless blood pressure estimation from PPG signals and its derivatives using deep learning models. *Biomed. Signal Process. Control* **2021**, *70*, 102984. [[CrossRef](#)]
28. Suresh Kumar, S.; Dashtipour, K.; Abbasi, Q.H.; Imran, M.A.; Ahmad, W. A review on wearable and contactless sensing for COVID-19 with policy challenges. *Front. Commun. Netw.* **2021**, *2*, 636293. [[CrossRef](#)]
29. Zhou, L.; Guess, M.; Kim, K.R.; Yeo, W.H. Skin-interfacing wearable biosensors for smart health monitoring of infants and neonates. *Commun. Mater.* **2024**, *5*, 72. [[CrossRef](#)] [[PubMed](#)]
30. De Haan, G.; Jeanne, V. Robust pulse rate from chrominance-based rPPG. *IEEE Trans. Biomed. Eng.* **2013**, *60*, 2878–2886. [[CrossRef](#)] [[PubMed](#)]
31. Mohan, J.; Kumar, M.L.; Bhattacharya, S.; Saranya, A.S. Non-Contact Blood Pressure Estimation Using Camera-Based rPPG Analysis. In *Proceedings of the 2024 15th International Conference on Computing Communication and Networking Technologies (ICCCNT)*; IEEE: Piscataway, NJ, USA, 2024; pp. 1–5.
32. Huang, Y.; Huang, D.; Huang, J.; Wang, G.; Pan, L.; Lu, H.; He, M.; Wang, W. Camera-based blood pressure monitoring based on multi-site and multi-wavelength pulse transit time features. *IEEE Trans. Instrum. Meas.* **2024**, *73*, 1–14. [[CrossRef](#)]
33. Xing, W.; Shi, Y.; Wu, C.; Wang, Y.; Wang, X. Predicting blood pressure from face videos using face diagnosis theory and deep neural networks technique. *Comput. Biol. Med.* **2023**, *164*, 107112. [[CrossRef](#)] [[PubMed](#)]
34. Hou, Z.; Huang, D.; Huang, Y.; Huang, J.; Zhao, N.; Pan, L.; Lu, H.; Shan, C.; Wang, W. Exploiting Multi-wavelength Morphological Features of Camera-PPG for Blood Pressure Estimation. *IEEE Trans. Instrum. Meas.* **2025**, *74*, 1–12. [[CrossRef](#)]
35. Chen, Y.; Zhuang, J.; Li, B.; Zhang, Y.; Zheng, X. Remote blood pressure estimation via the spatiotemporal mapping of facial videos. *Sensors* **2023**, *23*, 2963. [[CrossRef](#)] [[PubMed](#)]
36. Schrupf, F.; Frenzel, P.; Aust, C.; Osterhoff, G.; Fuchs, M. Assessment of deep learning based blood pressure prediction from PPG and rPPG signals. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Virtual, 19–25 June 2021; pp. 3820–3830.

37. Manullang, M.C.T.; Lin, Y.H.; Chou, N.K. A Transformer Based Network for Estimating Blood Pressure Using Facial Videos. *IEEE Sens. J.* **2024**, *25*, 1969–1977. [[CrossRef](#)]
38. Rong, M.; Li, K. A blood pressure prediction method based on imaging photoplethysmography in combination with machine learning. *Biomed. Signal Process. Control* **2021**, *64*, 102328. [[CrossRef](#)]
39. Kawasaki, S.; Kinefuchi, M.; Hino, Y.; Kobayashi, A.; Kawahito, S.; Takahashi, M.; Tsumura, N. Performance evaluation of blood pressure estimation using an RGB-NIR camera. *Artif. Life Robot.* **2025**, *30*, 484–492. [[CrossRef](#)]
40. van Putten, L.D.; Bamford, K.E. Improving systolic blood pressure prediction from remote photoplethysmography using a stacked ensemble regressor. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Vancouver, BC, Canada, 18–22 June 2023; pp. 5957–5964.
41. Finnegan, E.; Davidson, S.; Harford, M.; Watkinson, P.; Tarassenko, L.; Villarroel, M. Features from the photoplethysmogram and the electrocardiogram for estimating changes in blood pressure. *Sci. Rep.* **2023**, *13*, 986. [[CrossRef](#)] [[PubMed](#)]
42. Djeldjli, D.; Bousefsaf, F.; Maaoui, C.; Bereksi-Reguig, F.; Pruski, A. Remote estimation of pulse wave features related to arterial stiffness and blood pressure using a camera. *Biomed. Signal Process. Control* **2021**, *64*, 102242. [[CrossRef](#)]
43. Wang, W.; Den Brinker, A.C.; Stuijk, S.; De Haan, G. Algorithmic principles of remote PPG. *IEEE Trans. Biomed. Eng.* **2016**, *64*, 1479–1491. [[CrossRef](#)] [[PubMed](#)]
44. Maqsood, S.; Xu, S.; Springer, M.; Mohawesh, R. A benchmark study of machine learning for analysis of signal feature extraction techniques for blood pressure estimation using photoplethysmography (PPG). *IEEE Access* **2021**, *9*, 138817–138833. [[CrossRef](#)]
45. Wang, L.; Pickwell-MacPherson, E.; Liang, Y.; Zhang, Y.T. Noninvasive cardiac output estimation using a novel photoplethysmogram index. In *Proceedings of the 2009 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*; IEEE: Piscataway, NJ, USA, 2009; pp. 1746–1749.
46. Lu, S.; Zhao, H.; Ju, K.; Shin, K.; Lee, M.; Shelley, K.; Chon, K.H. Can photoplethysmography variability serve as an alternative approach to obtain heart rate variability information? *J. Clin. Monit. Comput.* **2008**, *22*, 23–29. [[CrossRef](#)] [[PubMed](#)]
47. Suboh, M.Z.; Jaafar, R.; Nayan, N.A.; Harun, N.H.; Mohamad, M.S.F. Analysis on four derivative waveforms of photoplethysmogram (PPG) for fiducial point detection. *Front. Public Health* **2022**, *10*, 920946. [[CrossRef](#)] [[PubMed](#)]
48. Jolliffe, I.T.; Cadima, J. Principal Component Analysis: A Review and Recent Developments. *Philos. Trans. R. Soc. A Math. Phys. Eng. Sci.* **2016**, *374*, 20150202. [[CrossRef](#)] [[PubMed](#)]
49. Abdi, H.; Williams, L.J. Principal Component Analysis. *Wires Comput. Stat.* **2010**, *2*, 433–459. [[CrossRef](#)]
50. Hotelling, H. Analysis of a Complex of Statistical Variables into Principal Components. *J. Educ. Psychol.* **1933**, *24*, 417–441. [[CrossRef](#)]
51. Pearson, K. LIII. On Lines and Planes of Closest Fit to Systems of Points in Space. *London Edinburgh Dublin Philos. Mag. J. Sci.* **1901**, *2*, 559–572. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.