



Predicting blood pressure from face videos using face diagnosis theory and deep neural networks technique



Weiyi Xing^a, Yinni Shi^b, Chaoyong Wu^a, Yiqiao Wang^{c,*}, Xu Wang^{d,**}

^a School of Traditional Chinese Medicine, Beijing University of Chinese Medicine, Beijing, 102488, China

^b School of Information, North China University of Technology, Beijing, 100144, China

^c School of Management, Beijing University of Chinese Medicine, Beijing, 102488, China

^d School of Life Science, Beijing University of Chinese Medicine, Beijing, 102488, China

ARTICLE INFO

Keywords:

Blood pressure prediction
Face diagnosis
Deep neural networks

ABSTRACT

Hypertension is a major cause of cardiovascular diseases. Accurate and convenient measurement of blood pressure are necessary for the detection, treatment, and control of hypertension. In recent years, face video based non-contact blood pressure prediction is a promising research topic. Interestingly, face diagnosis has been an important part of traditional Chinese medicine (TCM) for thousands of years. TCM practitioners observe some typical regions of the face to determine the health status of the Zang Fu organs (i.e., heart). However, the effectiveness of face diagnosis theory in conjunction with computer vision analysis techniques to predict blood pressure is unclear. We proposed an artificial intelligence framework for predicting blood pressure using deep convolutional neural networks in this study. First, we extracted pulse wave signals through 652 facial videos. Then, we trained and compared nine artificial neural networks and chose the best performed prediction model, with an overall true predict rate of 90%. We also investigated the impact of face reflex regions selection on blood pressure prediction model, and the five face regions outperformed. Our high effectiveness and stability framework may provide an objective and convenient computer-aided blood pressure prediction method for hypertension screening and disease prevention.

1. Introduction

Hypertension is closely related to the development of cardiovascular diseases (CVDs) and imposes substantial public health burden [1–3]. In China, the number of people with hypertension has reached 245 million [4]. Accurate and convenient measurement of blood pressure are necessary for the detection, treatment, and control of hypertension. The traditional blood pressure estimation methods are usually contact-based and may induce discomfort. Therefore, a more convenient non-contact blood pressure monitoring method is an urgent requirement.

Vision is the most important source of information in humans and may provide opportunity to record blood pressure related changes from a distance. Interestingly, visual inspection has been the essential diagnostic method in traditional Chinese medicine (TCM) for thousands of years. TCM practitioners observe some typical regions of the face to determine the health status of the Zang Fu organs (heart, liver, spleen,

lung, and kidney) and to diagnose diseases at an early stage [5,6]. The convergence of this traditional face diagnosis theory, digital camera photography, image processing, and artificial intelligence (AI) techniques may contribute to accurately measure blood pressure comfortably and conveniently.

In 2016, Jeong and Finkelstein presented a contactless blood pressure estimation using video camera [7]. Specifically, they extracted image-based photoplethysmography (PPG), calculated pulse transit time between face and right palm, and then estimate blood pressure. PPG is an optical technique for detecting blood volume changes with each heart beat and can be utilized to extract blood pressure values [8]. As a variant of image-based PPG, transdermal optical imaging (TOI) [9] can measure hemoglobin concentrations changes from raw digital camera images representing facial blood flow fluctuations and permit continuous blood pressure monitoring using multilayer perceptron algorithm. These non-contact methods provide an easy and comfortable replacement for

* Corresponding author.

** Corresponding author.

E-mail addresses: xwy_1216@163.com (W. Xing), Yinnishi0105@163.com (Y. Shi), wucy@bucm.edu.cn (C. Wu), yqwang@bucm.edu.cn (Y. Wang), wangx@bucm.edu.cn (X. Wang).

current cuff-based devices. In recent years, AI techniques (e.g., deep learning) play a more and more important role in predicting blood pressure from contact-based physiological signals [10–16]. However, face video and deep neural network blood pressure prediction technique have rarely been jointly studied, the effectiveness of the TCM face diagnosis theory in the paradigm also remains to be verified.

In the present study, we selected five face regions of interest (ROIs) based on Zang Fu organs in face diagnosis and extracted pulse wave signals from facial videos. Then, we constructed a variety of deep artificial neural networks for training and investigating the performances of blood pressure prediction models.

2. Materials and methods

2.1. Dataset construction and preprocessing

The multi-scene vital sign dataset for physiological signal recognition created by our research group in 2021 was used in the present study [17]. This dataset contains 107 samples (37 males and 70 females). All participants reported no history of hypertension, but 2 participants' blood pressure were larger than 140/90 mmHg during the collection. Due to incomplete information and difficulty in format conversion on some of the data, 652 facial videos were finally included in our study. Specifically, samples 1–74 had 8 facial videos under open light source (4 different states: resting, chatting, deep breathing, and after-exercise; 2 cameras), whereas samples 75–107 contained 2 facial videos captured using 2 cameras under resting state in the standard environment. The two cameras, including an industrial camera (Medavision, MVSUA134 GC/MT) and a web camera (Aoni, C27Pro full HD Video), recorded facial videos simultaneously. The industrial camera has a CMOS sensor, a maximum resolution of 1280×960 , a maximum frame rate of 15 fps, and a color space of RGB. The collection of the data was approved by the Ethics Committee of Beijing University of Chinese Medicine (No. 2019BJZYYLL0101) and written informed consent was obtained from all participants.

First, the 68-point face monitoring model, which is built on the Dlib database, was used to track and annotate facial feature points on video images frame by frame, thereby selecting five regions based on TCM face diagnosis (heart, liver, spleen, lung, and kidney) [18] were primarily chosen as the ROIs (Fig. 1-A) for deriving pulse wave signals in RGB color space. Then, We divide the subject's video into frames, find the key points in each frame to target the region of interest, and take the mean value of the green channel for each frame. We selected a single green channel value for processing for two reasons. On the one hand, a

person's skin color is not sensitive to the green channel, and the green channel values of facial images of people with different skin colors are relatively small; on the other hand, transdermal optical imaging (TOI) explains that the green channel in an image is highly responsive to changes in blood flow [9]. Third, there is too much noise in the image, especially from the acquisition equipment and external environment. Therefore, we extract the spatial average values of pixels in each ROI of the image, and then average these spatial average values again among the five ROIs to form a one-dimensional signal, thereby avoiding "acquisition errors". In addition, a whole face image contains excessive amounts of irrelevant information, such as eyebrows, eyes, lips, etc., and the approach in this paper does not concentrate on the region of interest, which is the region of interest.

2.2. Network architecture

In the present study, the models were developed with the following experimental settings: Pytorch 1.10.1, Python 3.9, Ubuntu 18.04 operating system, NVIDIA GeForce RTX 2080Ti GPU, and CUDA 11.3. This study focused on the effectiveness and comparison of deep neural network models with different network structures, and selected best model structure for predicting blood pressure.

2.3. Model training and testing

The model was trained using processed sample data, and its depth is changed to get a better depth range before training and comparing models with various structures to get the best model structure. The network operation flow is shown in Fig. 2.

This article uses a series of models to predict blood pressure based on the characteristics of medical datasets. Firstly, we selected typical and less complex convolutional neural networks such as ResNet50, ResNet18, GoogleNet, and VGG16 for extensive comparative experiments to determine the optimal model structure. In addition, since the input data we studied is a one-dimensional signal, we updated the prediction network by replacing the original 2D convolution with 1D convolution, and modified the size and step size of the pooling layer to adapt to the input. This enables classic networks like ResNet to process 1D signals and accurately predict blood pressure values. Secondly, we mainly avoid overfitting of network training by adding discard layer, hyperparameter optimization and model structure adjustment. Finally, we selected the VGG16 network with the best prediction performance among the nine models, reduced its network layer and reset some convolutional channels to obtain the prediction model L-VGG (lightweight

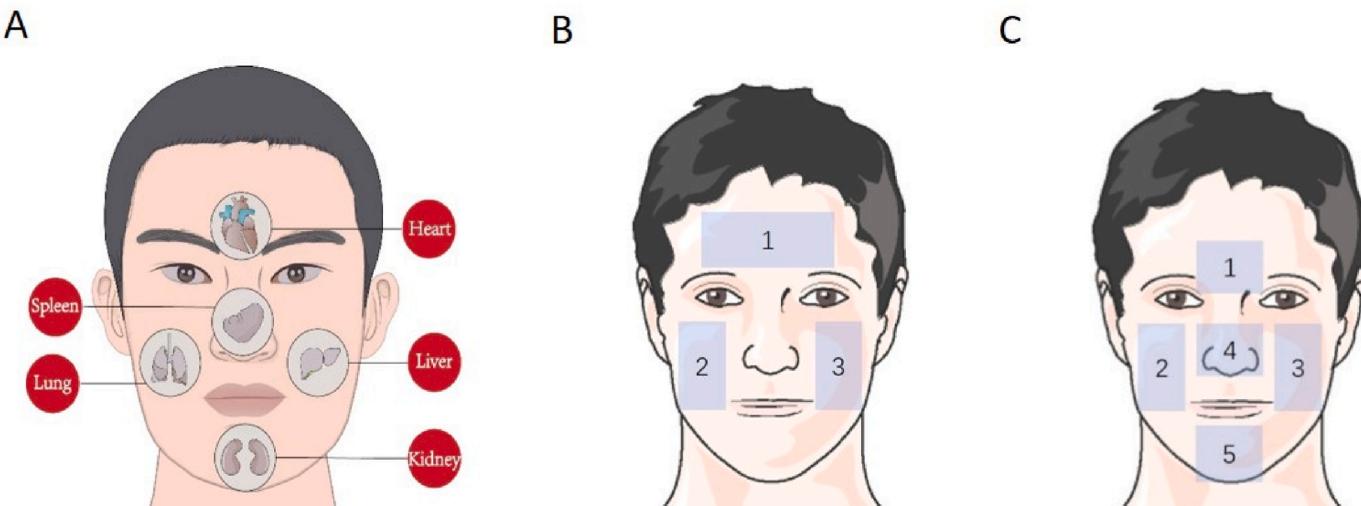


Fig. 1. Diagram of the different facial regions of interest (ROIs). (A) The five facial reflex zones, (B) Three facial ROIs, (C) Five facial ROIs.

Video image input

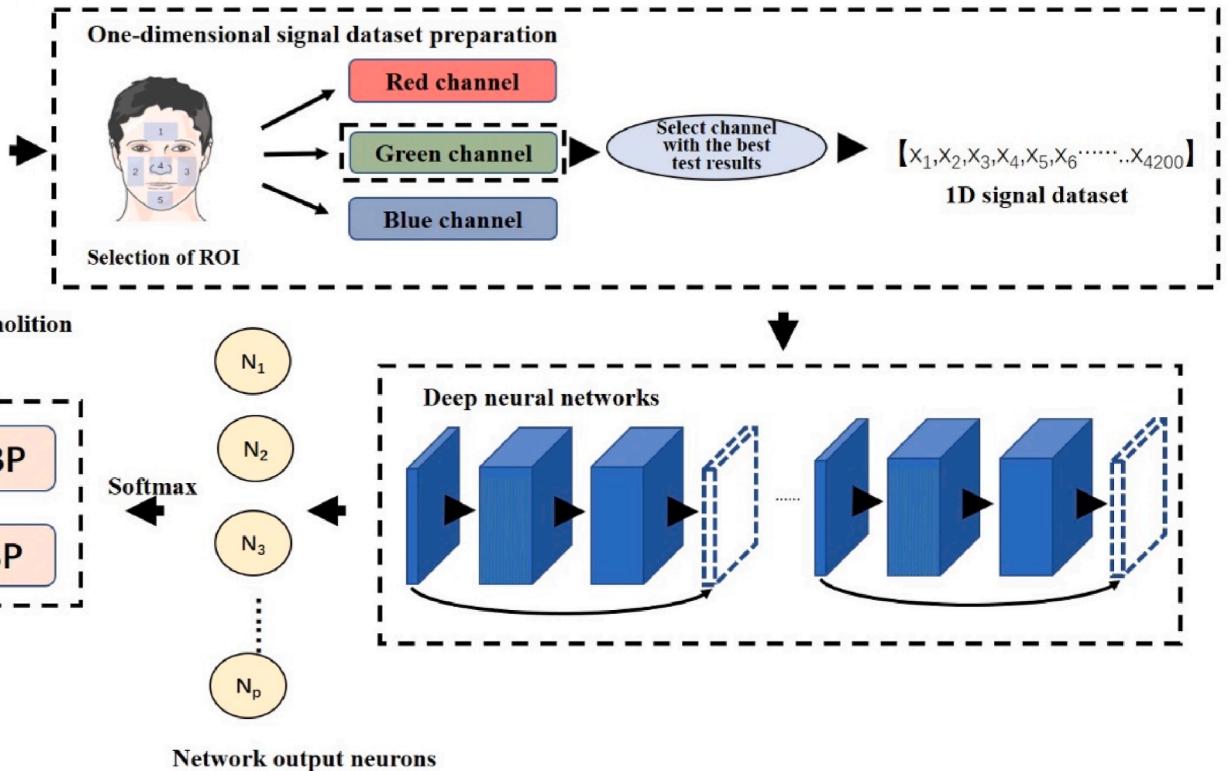


Fig. 2. Diagram of the network operation flow
(DBP: Diastolic Blood Pressure, SBP: Systolic Blood Pressure, N: Abbreviations for neurons).

VGG16). Experiments have shown that this method improves prediction accuracy while reducing model complexity. The specific framework of the nine models was shown in Fig. 3.

To examine the contribution of the five regions of interest on the face to predicted blood pressure, we first converted the one-dimensional signal dataset of the face, which consisted of the mean green channel pixel values of the five regions, to a two-dimensional signal dataset, as shown below.

$$X = [(a_1, b_1, c_1, d_1, e_1), (a_2, b_2, c_2, d_2, e_2), \dots, (a_n, b_n, c_n, d_n, e_n)]_n \quad (2.1)$$

In this two-dimensional signal, a, b, c, d, e signify the pixel average of the green channel in the five regions of the face (heart, spleen, kidney, lung, and liver), and n denotes the number of frames in a single subject's face after frame segmentation in the video.

We input this two-dimensional signal into the network and multiply it with a 5×1 matrix through the first linear layer. This 5×1 matrix is automatically learned and updated by the network during training and parameter updating, and the value of this matrix is determined at the end of training, which is the weight generated when the two-dimensional signal is converted into a one-dimensional signal. The weights are given as follows.

$$W = \begin{bmatrix} w_1 \\ w_2 \\ w_3 \\ w_4 \\ w_5 \end{bmatrix} \quad (2.2)$$

The greater the value of w , the more important the facial region corresponding to that value is considered by the network during training, i.e. the greater the contribution of that facial region to the

prediction result of the facial 1D signal.

The weights were processed and fed into a Softmax function so that the weight values corresponding to each of the five facial regions were transformed into a probability distribution ranging from (0, 1) with all elements adding up to 1. The Softmax function is shown below.

$$\text{Softmax}(z_i) = \frac{e^{z_i}}{\sum_{c=1}^C e^{z_c}} \quad (2.3)$$

Where z_i is the output value of the i th node and C is the number of output nodes, i.e. the number of weights.

2.4. Model evaluation statistics

The prediction accuracy of blood pressure prediction was assessed by the mean absolute error (MAE), root mean squared error (RMSE), mean absolute percentage error (MAPE), Pearson correlation coefficient (PCC, denoted as r), and coefficient of determination R^2 . These are the equations.

$$MAE = \frac{1}{N} \sum_{i=1}^N |y_i - \hat{y}_i| \quad (2.4)$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2} \quad (2.5)$$

$$MAPE = 100\% \times \frac{\sum_{i=1}^N \left| \frac{y_i - \hat{y}_i}{y_i} \right|}{n} \quad (2.6)$$

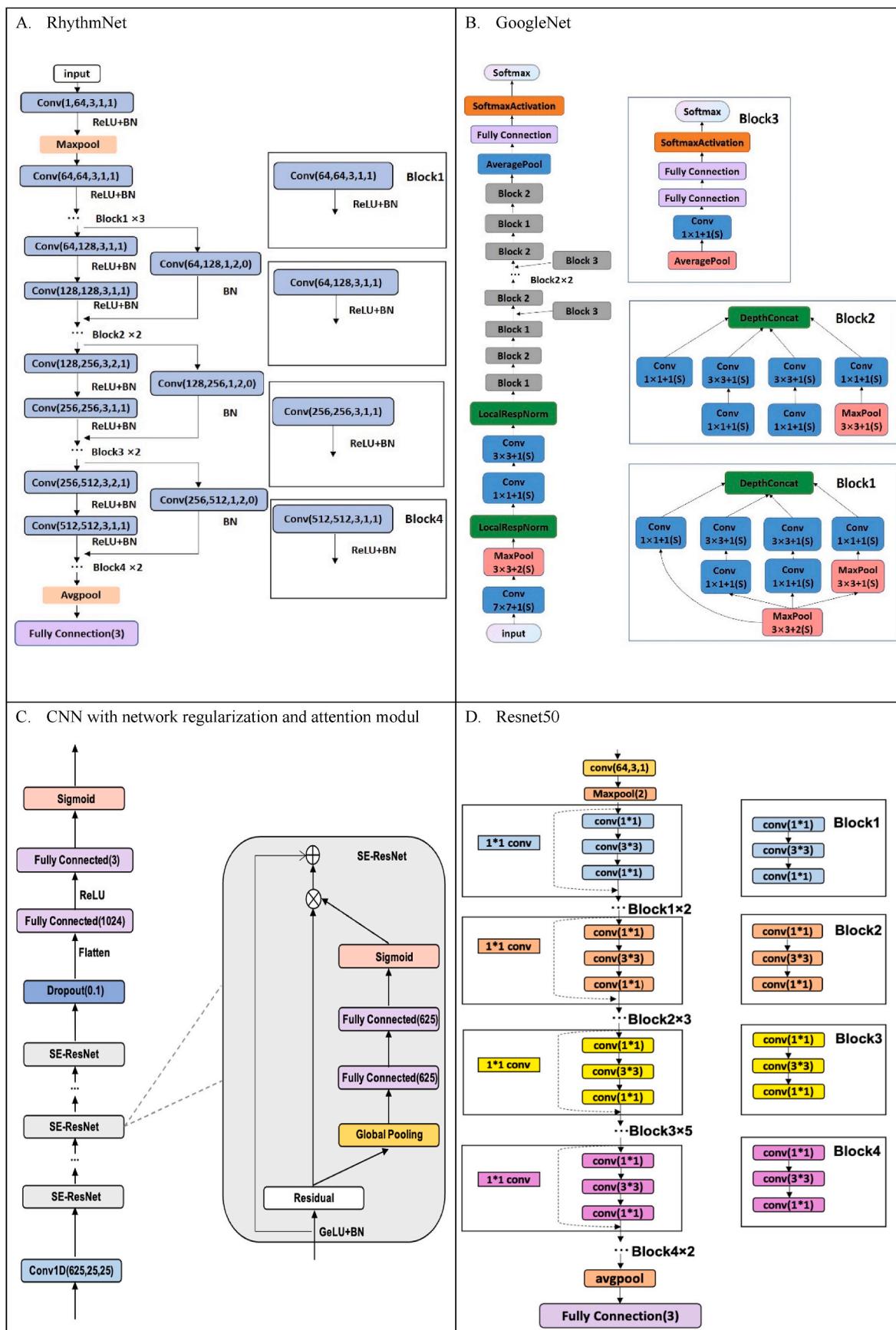


Fig. 3. Diagram of specific framework of nine models.

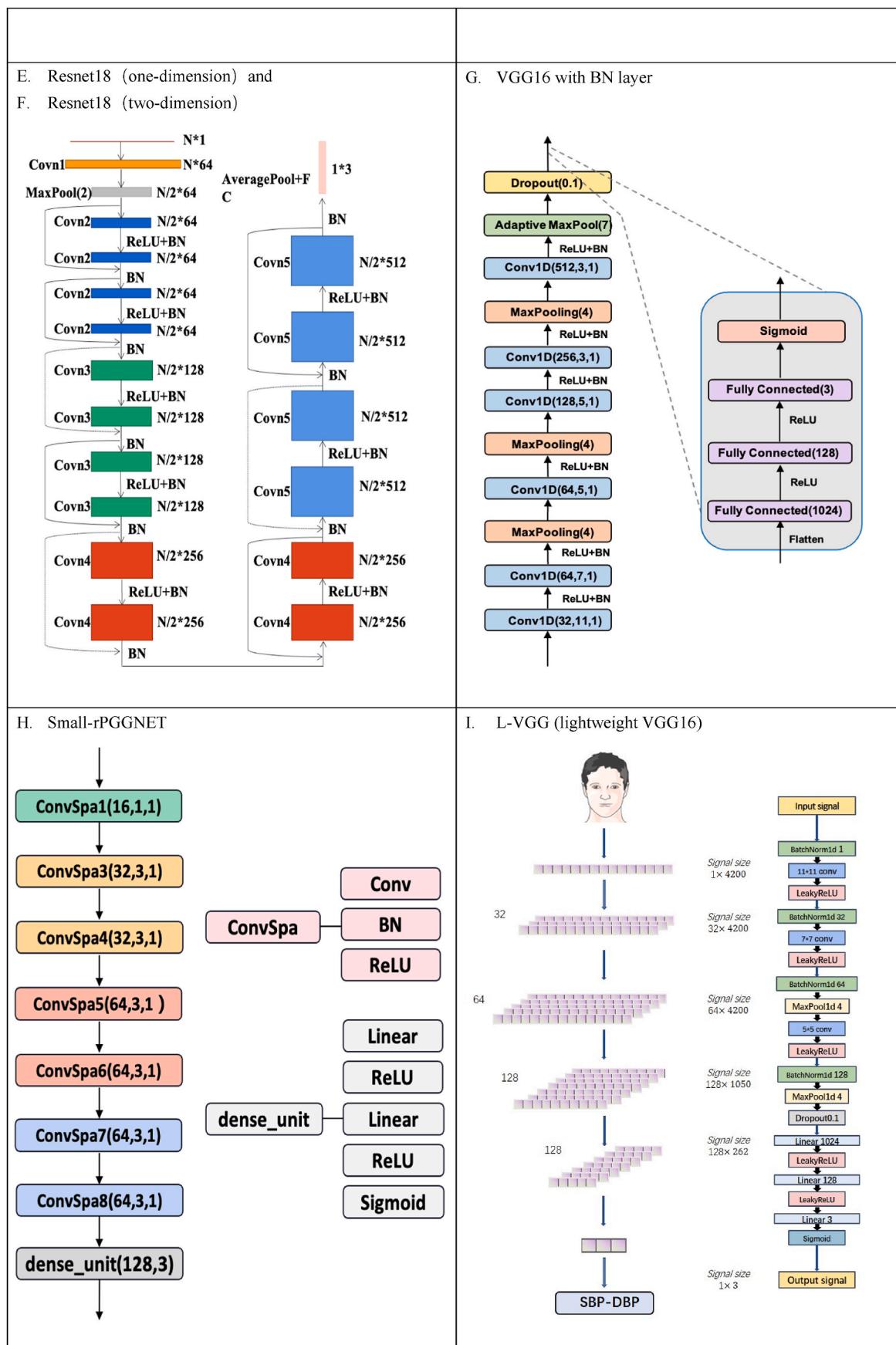


Fig. 3. (continued).

$$PCC = \frac{\sum_{i=1}^N (\hat{y}_i - \bar{\hat{y}})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^N (\hat{y}_i - \bar{\hat{y}})^2} \sqrt{\sum_{i=1}^N (y_i - \bar{y})^2}} \quad (2.7)$$

$$R^2 = 1 - \frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{\sum_{i=1}^N (y_i - \bar{y})^2} \in [0, 1] \quad (2.8)$$

$$TPR = n/N \quad (2.9)$$

In the equation above, N is the number of samples, \hat{y}_i , y_i are the predicted and true blood pressure values of the i-th sample respectively, $\bar{\hat{y}}$, \bar{y} are the mean of the predicted and true values respectively, and n is the number of samples where the difference between the predicted and true values is less than 5 in absolute value. The mean blood pressure value (test value - reference value) must be 5 mmHg or lower, and for systolic blood pressure (SBP) and diastolic blood pressure (DBP), standard deviation (SD) must be 8 mmHg or below, under the AAMI/ESH/ISO General Standard (ISO 81060-2: 2018) [19].

3. Results

3.1. Model construction

Our study focused on the prediction of blood pressure based on pulse wave signals, with a variety of artificial neural networks chosen for training and improvement. The findings were in line with those reported by SCHRUMPF F [20], who used MAE as the primary evaluation metric. As shown in Table 1, the lightweight VGG16 (L-VGG) model, which featured continuous iterations of convolutional, pooling, and fully connected layers, ultimately reduced the training loss and obtained accurate prediction values after comparative analysis of the nine models. The model predicted blood pressure prediction models with MAE values below 2.5, which complied with the AAMI/ESH/ISO general standard (ISO 81060-2: 2018), and had excellent metrics for RMSE, MAPE, PCC, SD and R^2 , it was chosen as the best model for blood pressure prediction.

Fig. 3(I) displayed visualization of the L-VGG model structure.

When we ran tests with the test set, Fig. 4 shows the actual values of the subject's blood pressure on the blue line and the predicted values of the best prediction model on the red line. Qualitatively, it is evident that the model accurately forecasts, diastolic blood pressure, and systolic blood pressure, with a small error between the predicted and actual values.

3.2. Model cross-validation

K-fold cross-validation is a method used to assess the stability of a

model with low bias [21]. Ten-fold cross-validation is used in this study. The basic steps of ten-fold cross-validation are as follows: (i) divide the original number into 10 subsets with the most balanced sample size possible; (ii) use the first subset as the test set and the second-9th subsets as the training set; (iii) train the model using the training set and calculate the results of each evaluation indicator in the test set; (iv) repeat steps 2 and 3 using the second-10th subsets as the test set in turn; (v) calculate the average of each evaluation index as the final result. Tables 2–3 are the result of ten-fold cross-validation.

The cross-validation experiments showed that the optimal model (lightweight VGG16) was relatively accurate in predicting blood pressure, rather than the results being accurate by chance. The MAE of the second group was 1.39 and TPR was also greater than 95%. In addition, an intra-group correlation coefficient (ICC) close to one was calculated, indicating good measurement and evaluation reliability.

3.3. Comparative analysis of regions of interest

Depending on the degree of relevance to the rPPG, three regions of interest—the forehead (ROI 1), the right cheek (ROI 2), and the left cheek (ROI 3)—with relatively wide areas and uniform colors were initially chosen as ROI [22], as shown in Fig. 1-B. Face diagnosis theory guided the selection of five points of interest [23]. The following areas are of interest: heart (forehead area near the eyebrows, ROI 1), lung (right cheek, ROI 2), liver (left cheek, ROI 3), spleen (nose, ROI 4), and kidney (chin, ROI 5), as shown in Fig. 1-C.

Fig. 5 depicts the anticipated outcomes, and it can be seen that the data results for all five parts of the face (Fig. 5-A) performed better than those for the three regions of the face (Fig. 5-B). It is maybe that the five facial regions are the most important for this study, with additional comparison analysis to come in different scenarios.

3.4. Contribution analysis of sub-region on face

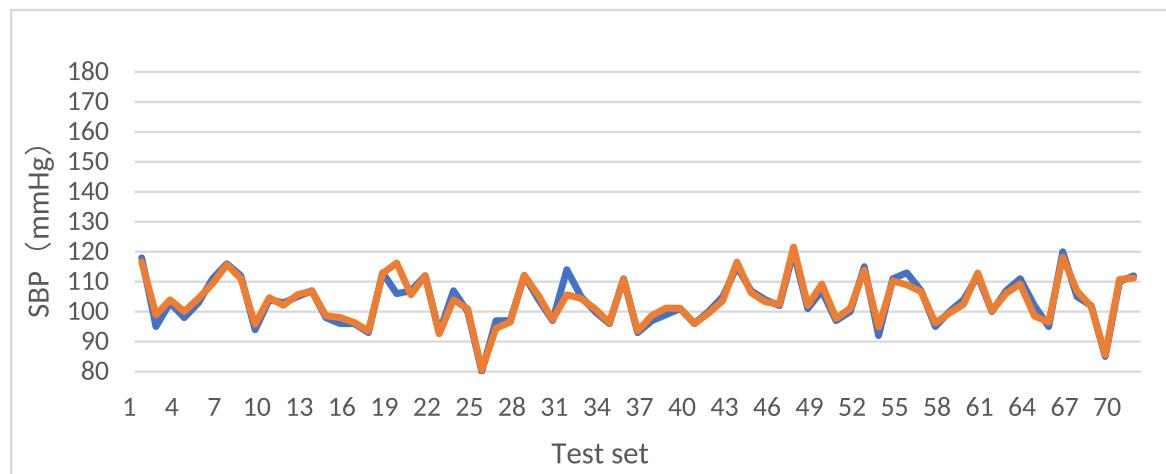
Each sub-region of the face, including the heart (forehead area near the eyebrows, ROI), liver (left cheek, ROI), lung (right cheek, ROI), spleen (nose, ROI), and kidney (chin, ROI). Fifteen sets of experiments were conducted processing in conjunction with equations (2.1)–(2.3), the results of the contribution of each facial region to the predicted outcome are shown in Table 4 and Table 5. It was found that in ten out of fifteen experiments, the top three weights were all kidney, liver and lung.

4. Discussion

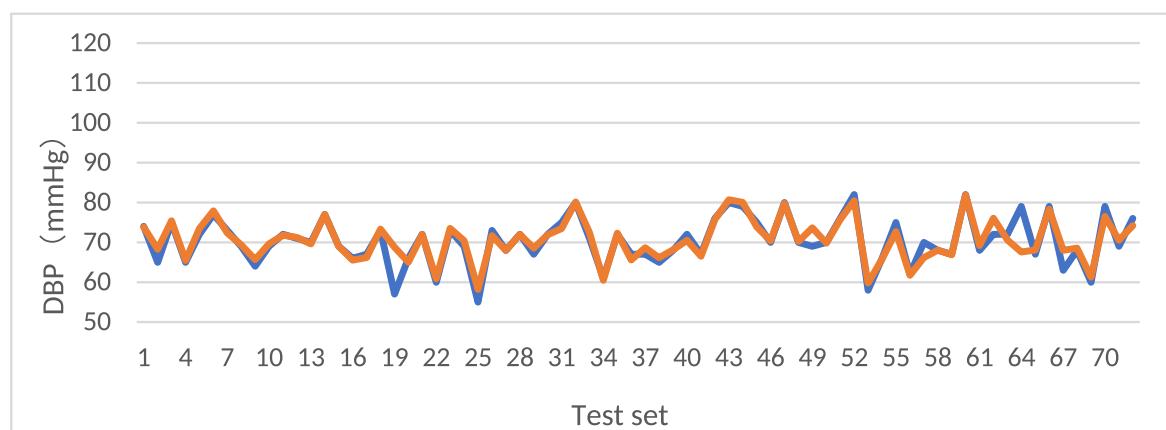
Based on the concepts of facial diagnosis in Chinese medicine and the

Table 1
Summary of blood pressure predicted by different artificial neural network models.

Model	Label	MAE	RMSE	MAPE	SD	PCC	R ²	TPR
RhythmNet	SBP	14.07	16.50	14.86	10.77	0.03	-0.68	0.00
	DBP	10.26	12.08	12.96	8.55	0.03	-0.04	0.00
GoogleNet	SBP	16.06	19.92	18.84	9.31	-0.05	-0.73	0.00
	DBP	14.83	17.09	16.05	10.69	0.01	-0.70	0.00
CNN with network regularization and attention module	SBP	7.52	9.03	7.16	5.04	0.23	0.03	0.26
	DBP	6.19	8.17	8.27	5.38	0.08	-0.12	0.54
Resnet50	SBP	6.33	8.76	5.89	6.11	0.44	0.09	0.54
	DBP	5.02	7.13	6.73	5.11	0.47	0.15	0.60
Resnet18 (one-dimension)	SBP	5.74	8.27	5.42	5.96	0.49	0.19	0.60
	DBP	5.34	7.28	7.31	4.99	0.52	0.11	0.57
Resnet18 (two-dimension)	SBP	6.27	8.61	5.84	5.94	0.43	0.12	0.50
	DBP	4.64	5.99	6.28	3.82	0.65	0.40	0.65
VGG16 with BN layer	SBP	5.01	7.29	4.70	5.43	0.63	0.37	0.64
	DBP	4.47	5.82	6.04	4.04	0.67	0.43	0.67
Small-rPGGNET	SBP	5.57	7.33	5.23	4.79	0.62	0.36	0.53
	DBP	4.53	6.51	6.15	4.70	0.56	0.29	0.67
L-VGG (lightweight VGG16)	SBP	2.44	4.21	2.43	3.44	0.89	0.79	0.88
	DBP	2.01	3.47	2.81	2.94	0.88	0.79	0.90

**Fig. 4 A.** Graph of true vs. predicted systolic blood pressure fluctuations

(The blue line represents the actual value of the subject's blood pressure, while the red line represents the predicted value of the best predictive model, SBP: Systolic Blood Pressure) .

**Fig. 4 B.** Graph of true vs. predicted diastolic blood pressure fluctuations

(The blue line represents the actual value of the subject's blood pressure, while the red line represents the predicted value of the best predictive model, DBP: Diastolic Blood Pressure) .

Table 2
SBP evaluation indicators of ten-fold cross-validation.

Test set	MAE	RMSE	MAPE	R ²	SD	PCC	ICC	TPR
Fold-1	1.76	2.81	2.81	0.92	2.21	0.96	0.99	0.90
Fold-2	1.39	2.23	1.34	0.92	1.75	0.96	0.99	0.97
Fold-3	2.51	5.74	2.21	0.73	5.20	0.86	0.92	0.92
Fold-4	2.54	4.23	2.40	0.85	3.40	0.93	0.97	0.90
Fold-5	2.32	4.12	2.29	0.79	3.42	0.89	0.95	0.89
Fold-6	1.86	3.08	1.80	0.86	2.48	0.93	0.97	0.90
Fold-7	5.09	7.99	4.94	0.46	6.20	0.68	0.79	0.67
Fold-8	2.69	4.87	2.51	0.71	4.09	0.84	0.92	0.85
Fold-9	2.43	4.27	2.22	0.76	3.54	0.88	0.93	0.82
Fold-10	1.86	2.78	1.78	0.93	2.09	0.96	0.99	0.94
Average	2.44	4.21	2.43	0.79	3.44	0.89	0.94	0.88

principles of pulse wave, we suggest a framework for predicting blood pressure. We choose the five reflex regions of heart, liver, spleen, lung and kidney as our regions of interest and perform image processing on facial videos to extract a one-dimensional signal. A well-known artificial neural network model is then fed the one-dimensional signal for training. The L-VGG is presented as the best prediction model after comparing the training outcomes of nine models, which has an overall accuracy up to 90% on the dataset.

Table 3
DBP evaluation indicators of ten-fold cross-validation.

Test set	MAE	RMSE	MAPE	R ²	SD	PCC	ICC	TPR
Fold-1	1.64	2.56	2.43	0.90	1.98	0.95	0.98	0.92
Fold-2	1.28	2.41	1.89	0.93	2.06	0.91	0.96	0.96
Fold-3	1.76	2.84	2.52	0.91	2.25	0.90	0.96	0.94
Fold-4	2.11	4.05	2.83	0.71	3.49	0.85	0.91	0.92
Fold-5	1.84	3.12	2.77	0.78	2.53	0.90	0.95	0.89
Fold-6	1.53	2.50	2.20	0.90	1.99	0.95	0.99	0.93
Fold-7	3.47	6.57	4.45	0.44	5.61	0.67	0.77	0.85
Fold-8	2.18	3.79	2.96	0.79	3.12	0.89	0.95	0.88
Fold-9	2.44	4.03	3.34	0.71	3.23	0.85	0.91	0.83
Fold-10	1.81	2.78	2.69	0.88	2.12	0.94	0.98	0.90
Average	2.01	3.47	2.81	0.79	2.84	0.88	0.94	0.90

Our study might offer new approaches and concepts for hypertension early detection. Delay in diagnosis has a substantial impact on the course of hypertension disease and increases the risk of cardiovascular disease (CVD) and all-cause mortality. How to implement early hypertension monitoring, more accurately identify potential patients, and eliminate the threat of the disease is one of the most important public health issues of the day. The majority of earlier studies relied on machine learning techniques [24,25], which require additional pre-processing steps to

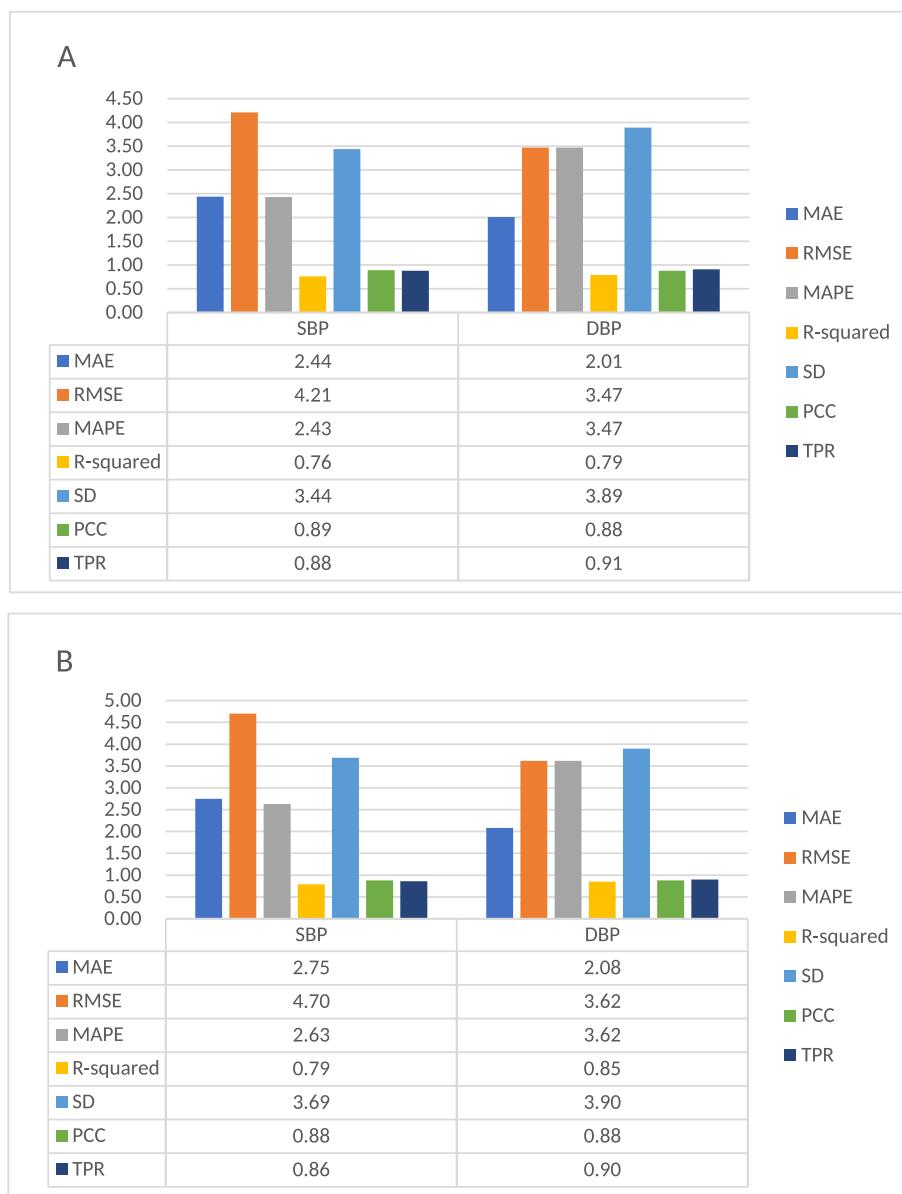


Fig. 5. Histogram of a comparison of five facial regions and three facial regions for prediction results
(MAE: Mean Absolute Error, RMSE: Root Mean Square Error, MAPE: Mean Absolute Percentage Error, R-squared: Coefficient of determination, SD: Standard Deviation, PCC: Pearson Correlation Coefficient, TPR: True Positive Rate, DBP: Diastolic Blood Pressure, SBP: Systolic Blood Pressure).

Table 4
Weights of sub-region contribution.

Experimental group	Heart	Spleen	Kidney	Lung	Liver	Sort by weight
1	0.1351	0.4104	0.1295	0.3877	0.2572	spleen-lung-liver-heart-kidney
2	-0.0641	-0.0553	0.1429	0.1293	0.2605	liver-kidney-lung-spleen-heart
3	-0.0132	0.2585	0.2832	0.1255	-0.0619	kidney-spleen-lung-heart-liver
4	-0.1000	-0.2964	-0.0439	0.1808	0.3836	liver-lung-kidney-heart-spleen
5	-0.1193	-0.1300	0.2471	0.2010	-0.0414	kidney-lung-liver-heart-spleen
6	-0.3264	0.0047	0.3224	0.3712	0.3614	lung-liver-kidney-spleen-heart
7	0.0836	0.2576	0.0812	0.0903	-0.0297	spleen-lung-heart-kidney-liver
8	-0.2205	-0.4097	-0.2799	0.0852	0.3429	liver-lung-heart-kidney-spleen
9	-0.1011	-0.0929	0.5107	-0.0709	0.4735	kidney-liver-lung-spleen-heart
10	-0.0379	-0.0793	0.2831	0.2715	0.2193	kidney-lung-liver-heart-spleen
11	0.2243	-0.1119	0.2718	-0.029	0.3113	kidney-liver-heart-lung-spleen
12	0.1496	0.1422	0.3165	0.4264	0.3694	lung-kidney-liver-heart-spleen
13	0.2024	-0.0991	0.2064	0.2317	0.3365	liver-lung-kidney-heart-spleen
14	0.1338	0.4452	-0.2913	0.1823	-0.1654	spleen-lung-heart-liver-kidney
15	-0.2283	-0.2983	0.3264	0.3918	0.1843	lung-kidney-liver-heart-spleen

Note: In nine out of fifteen experiments, the top three weights were all kidney, liver and lung.

Table 5
Probability weights of sub-region contribution.

Experimental group	Heart	Spleen	Kidney	Lung	Liver	Sort by weight
1	0.1799	0.1372	0.2123	0.2665	0.2041	lung-kidney-liver-heart-spleen
2	0.1532	0.1795	0.2278	0.2397	0.1998	lung-kidney-liver-spleen-heart
3	0.2260	0.2148	0.2512	0.1526	0.1554	kidney-spleen-heart
4	0.2090	0.1548	0.2067	0.2191	0.2104	lung-liver-kidney-heart-spleen
5	0.2406	0.1877	0.1286	0.2145	0.2286	heart-liver-lung-spleen
6	0.1745	0.1795	0.2059	0.2370	0.2031	lung-kidney-liver-spleen
7	0.1618	0.2974	0.2016	0.1901	0.1491	spleen-kidney-lung-heart
8	0.1509	0.1485	0.1863	0.2892	0.2251	lung-liver-kidney-heart
9	0.1706	0.1925	0.2047	0.2303	0.2019	spleen-lung-kidney-liver
10	0.2195	0.2309	0.2068	0.1779	0.1649	heart-kidney-lung
11	0.2118	0.1845	0.2362	0.2029	0.1646	kidney-heart-lung
12	0.1993	0.1810	0.2025	0.2012	0.2160	spleen-liver-kidney-lung
13	0.1634	0.1575	0.2151	0.2426	0.2214	heart-spleen-lung-liver
14	0.1587	0.1452	0.2558	0.2374	0.2029	kidney-lung-heart
15	0.1814	0.1472	0.2285	0.2331	0.2098	spleen-lung-kidney-liver

Note: In ten out of fifteen experiments, the top three weights were all kidney, liver and lung.

extract features based on experience. In this study, a deep neural network model was used to predict blood pressure data.

This method facilitates end-to-end blood pressure prediction, eliminates the need for human feature selection, and makes it simpler to install the system in clinical settings while enhancing operability. In the course of the study, we trained the nine networks that are used for training the most frequently and, based on comparisons of experimental results, we deduced improvements to the algorithm, ultimately proposing the best model. After conducting manual network training repeatedly and for a very long time, we found that feature extraction happens instead of failure and poor training results when the network structure is complex and has many layers. Because of its minimal structural complexity, quick computation speed, and resistance to overfitting, the CNN prediction model is easy to incorporate and use. The prediction is more precise thanks to an enhancement to the standard CNN model. The prediction error is controlled within 3 mmHg and the best model can be achieved within 50 epochs, which is much better than the AAMI/ESH/ISO standard (ISO 81060-2: 2018). The model has fewer network layers and rapid training convergence. The model's outstanding reliability has also been confirmed. In this study, the conversion of each input video duration into a one-dimensional signal took only 1 min, i.e. someone was in front of the screen for 1 min, and the blood pressure and heart rate values could be accurately predicted, which is a technique with practical value and future application possibilities.

Yet, there are several important topics for future research. First, artificial neural network prediction models are more accurate than conventional techniques but produce less explanatory predictions. This is a common feature of all deep learning models [26]. When we examined the contribution of the separate area of reflex regions, we discovered that the reflex zones of kidney, liver, and lungs were generally more contributing. This conclusion is supported by Chinese medical theory to a certain extent, but further comparative trials are needed in the future. Secondly, our dataset sample needs to be extended further in order to provide a more thorough comparison of the predictive impacts between populations with and without hypertension. Thirdly, the present model is only for blood pressure prediction. However, the most lethality of hypertension is its complications, such as stroke and heart diseases. In the future, it is very needed to investigated whether the proposed method can predict these cardiovascular events and other critical diseases.

CRediT authorship contribution statement

Weiyi Xing: Methodology, Formal analysis, Software, Writing original draft & editing. Yinni Shi: Investigation, Data curation, Software. Chaoyong Wu: Data collection, Data curation. Yiqiao Wang: Conceptualization, Supervision, Writing-review & editing. Xu Wang: Methodology, Formal analysis, Investigation, Writing-review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

This work is supported by the Key Research Program of the Chinese Academy of Sciences (ZDRW-ZS-2021-1-2), and the National Natural Science Foundation of China (No. 82004222, 82174224).

References

- [1] R. Philip, T. Beaney, N. Appealbaum, C.R. Gonzalvez, C. Koldweij, A.K. Golestaneh, N. Poulter, J.M. Clarke, Variation in hypertension clinical practice guidelines: a

- global comparison, *BMC Med.* 19 (2021) 117. <https://doi:10.1186/s12916-02-101963-0>.
- [2] W.L. Shi, J.X. Fan, K. Bai, R. Song, X.Q. Hu, K. He, The effectiveness of hypertension prevention and treatment and the influencing factors of hypertension control rate in Henan, *Mod. Prev. Med.* (2022) 49.
- [3] K.T. Mills, A. Stefanescu, J. He, The global epidemiology of hypertension, *Nat. Rev. Nephrol.* 16 (2020) 223–237. <https://doi:10.1038/s41581-019-0244-2>.
- [4] Committee on Revision of Chinese guidelines on prevention and control of hypertension in China. 2018 Chinese guidelines for the management of hypertension, *Chin. J. Cardiovasc. Med.* 19 (2019) 1–44. <https://doi:10.3969/j.issn.1672-5301.2019.03.001>.
- [5] L. Zhu, F. Liu, F. Ding, Advances in objectification of facial visualization, *J. Guizhou Univ. Traditional Chin. Med.* 42 (2020) 6. <https://doi:10.16588/j.cnki.issn2096-8426.2020.04.019>.
- [6] F. Lin, Y. Tan, P. Zhou, P. Luo, Survey on facial complexion system for clinical diagnosis in tradition Chinese medicine, *Chin. J. Stereology Image Analysis* 24 (2019) 225–240. <https://doi:cnki:sun:ztxs.0.2019-03-0070>.
- [7] I. Jeong, J. Finkelstein, Introducing contactless blood pressure assessment using a high speed video camera, *J. Med. Syst.* 40 (2016) 77. <https://doi:10.1007/s10916-016-0439-z>.
- [8] P.M. Nabeel, V.R. Kiran, J. Joseph, V.V. Abhidev, M. Sivaprakasam, Local pulse wave velocity: theory, methods, advancements, and clinical applications, *IEEE Rev. Biomed. Eng.* 13 (2020) 74–112. <https://doi:10.1109/RBME.2019.2931587>.
- [9] H. Luo, D. Yang, A. Barszczylk, N. Vempala, J. Wei, S.J. Wu, P.P. Zheng, G. Fu, K. Lee, Z.P. Feng, Smartphone-based blood pressure measurement using transdermal optical imaging technology, *Circ.-Cardiovasc. Imag.* 12 (2019), e008857. <https://doi:10.1161/CIRCIMAGING.119.008857>.
- [10] T.H. Wu, K.H. Pang, W.Y. Kwong, Predicting Systolic Blood Pressure Using Machine Learning, *the 7th International Conference on Information and Automation for Sustainability*, IEEE, 2015, pp. 1–6.
- [11] S.N.A. Ismail, N.A. Nayan, R. Jaafar, Z. May, Recent advances in non-invasive blood pressure monitoring and prediction using a machine learning approach, *Sensors* 22 (2022) 6195. <https://doi:10.3390/s22166195>.
- [12] S. Baek, J. Jang, S. Yoon, End-to-end blood pressure prediction via fully convolutional networks, *IEEE Access* 7 (2019) 185458–185468. <https://doi:10.1109/ACCESS.2019.2960844>.
- [13] S. Lee, J.H. Chang, Deep belief networks ensemble for blood pressure estimation, *IEEE Access* 5 (2017) 9962–9972. <https://doi:10.1109/ACCESS.2017.2701800>.
- [14] P. Su, X. Ding, Y. Zhang, J. Liu, F. Miao, N. Zhao, Long-term blood pressure prediction with deep recurrent neural networks, in: *2018 IEEE EMBS International Conference on Biomedical & Health Informatics (BHI)*, IEEE, 2018, pp. 323–328.
- [15] S.Y. Liu, Research on Blood Pressure Prediction Model of Hypertensive Patients Based on Deep Learning, Qingdao University of Science and Technology, 2020.
- [16] Z.J. Tan, Y.C. Xu, The study of blood pressure measurement system based on PPG and machine learning, *Med. Equip.* 33 (2020) 26–28.
- [17] C.Y. Wu, Research on Heart Rate and Blood Pressure Monitoring and Pulse Condition Feature Recognition Based on Video Images, Beijing University of Chinese Medicine, 2022.
- [18] Q. Guan, Y. Xu, S. Yang, Y.C. Guo, F.F. Li, Research on the relationship between the characteristics of traditional Chinese medicine facial diagnosis and disease, *CJTICMP* 37 (2022) 902–905. <https://doi:cnki:sun:bxyy.0.2022-02-077>.
- [19] G.S. Stergiou, B. Alpert, S. Mieke, R. Asmar, N. Atkins, S. Eckert, G. Frick, B. Friedman, T. Graßl, T. Ichikawa, J.P. Ioannidis, P. Lacy, R. McManus, A. Murray, M. Myers, P. Palatinis, G. Parati, D. Quinn, J. Sarkis, A. Shennan, T. Usuda, J. G. Wang, C.O. Wu, E. O'Brien, A universal standard for the validation of blood pressure measuring devices: association for the advancement of medical instrumentation/european society of hypertension/international organization for standardization (aami/esh/iso) collaboration statement, *Hypertension* 71 (2018) 368–374. <https://doi:10.1161/HYPERTENSIONAHA.117.10237>.
- [20] F. Schrumpf, P. Frenzel, C. Aust, G. Osterhoff, M. Fuchs, Assessment of non-invasive blood pressure prediction from ppg and rppg signals using deep learning, *Sensors* 21 (2021) 6022. <https://doi:10.3390/s21186022>.
- [21] Z. Lin, J. Lai, X. Chen, L. Cao, J. Wang, Curriculum reinforcement learning based on K-fold cross validation, *Entropy* 24 (2022) 1787. <https://doi:10.3390/e24121787>.
- [22] Z.T. Yu, W. Peng, X.B. Li, X.P. Hong, G.Y. Zhao, Remote heart rate measurement from highly compressed facial videos an end-to-end deep learning solution with video enhancement, in: *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, 2019, pp. 151–160.
- [23] S.Q. Jiang, A Study of the Principle of Face Diagnosis in the Divisions of the Huangdi Neijing, Hunan University of Chinese Medicine, 2020.
- [24] Y. Chen, Research on Blood Pressure Measurement and Prediction Based on Machine Learning, University of Science and Technology of China, 2020.
- [25] L.S. Cai, Research of Heart Rate Measurement Based on Multi-Channel PPG Sensor Signals, Zhejiang Normal University, 2017.
- [26] C. Voyant, G. Notton, S. Kalogirou, M.L. Nivet, C. Paoli, F. Motte, A. Fouilloy, Machine learning methods for solar radiation forecasting: a review, *Renew. Energy* 105 (2017) 569–582. <https://doi:10.1016/j.renene.2016.12.095>.