Sensors and Materials, Vol. 34, No. 11 (2022) 4167–4184
MYU Tokyo

S & M 3104

Contactless Blood Pressure Measurement by AI Robot

Shu-Yin Chiang* and Yi-Feng Chen

Department of Information and Telecommunications Engineering, Ming Chuan University,
No. 5 De Ming Rd., Gui Shan District, Taoyuan City 333, Taiwan

(Received July 25, 2022; accepted October 18, 2022)

Keywords: non-contact measurement, PPG, blood pressure, heart rate, deep learning, arrhythmia

In this study, an AI robot uses a camera and computer to perform face recognition and uses non-contact image physiological signal measurement technology to predict heartbeat and blood pressure. The predicted heartbeat and blood pressure are displayed on the robot tablet and are transmitted to a cloud database to assist healthcare management. In this study, we use RGB images to extract facial features and points of interest of the face and palm. The changes in vasoconstriction at these points of interest reflect the relationship between the absorption of light by blood, heartbeat, and blood pressure. Using the photoplethysmography (PPG) signal of the green channel in the RGB image through a convolutional neural network (CNN), deep learning technology can predict heartbeat and blood pressure values and even determine whether a subject has arrhythmia. Our results demonstrate that the predicted heart rate and blood pressure errors are 2.6 and 1.7%, respectively. The AI companion robot in this study can obtain the subject's physical information by a non-contact method, reducing anxiety and the cost of labor in medical care.

1. Introduction

Many traditional medical measurement methods require significant labor. With the current development of AI technology, measurements can now be obtained from training images using AI. The measurement of heart rate and blood pressure using non-contact methods is a new technology in the development of AI. One notable advantage of non-contact measurement is it reduces the anxiety of subjects when measuring blood pressure; anxiety frequently increases the measured value, especially because of the so-called white coat hypertension or white coat effect.[1,2] Some people experience increased blood pressure when they see medical staff, and some become anxious whenever their blood pressure is measured, leading to increased blood pressure. When people feel stressed, they become more nervous and their heart beats faster, leading to incorrect blood pressure measurements. Therefore, non-contact measurement allows subjects to have their blood pressure measured easily without generating stress. Moreover, measurements with robots can reduce labor costs in healthcare.

Many studies on non-contact measurement have dealt with variations in subjects' facial movement, light intensity, and detection distance.[3] Verkruysse *et al.*[4] proposed the use of a

*Corresponding author: e-mail: sychiang@mail.mcu.edu.tw
https://doi.org/10.18494/SAM4036

ISSN 0914-4935

digital camera under an indoor ambient light source to measure the pulse signal through photoplethysmography (PPG). Poh *et al.*[5] proposed the use of independent component analysis (ICA) to decompose a facial region of interest (ROI) signal into three independent signals according to the RGB signal in color images. When using the green signal G in an RGB image as the PPG signal to estimate the heart rate, the ICA method is used to render the waveform of the signal to make it clearer and reduce noise to obtain better results. Hsieh *et al.*[6] proposed an adaptive ROI selection method that is robust against the natural movement of the human head. Lewandowska *et al.*[7] used principal component analysis combined with ICA to reduce the signal processing time and the amount of calculation required to measure the heart rate. Monkaresi *et al.*[8] used a common camera instead of a high-resolution camera to estimate the heartbeat using linear regression and k-nearest neighbors (KNN) machine learning and showed that a common camera coupled with a machine learning method can obtain the heart rate. Li *et al.*[9] used discriminative response map fitting to find 66 feature points of the human face, and they extracted nine feature points to generate a mask of the ROI. They used a model of the illumination variations in the ROI, divided the pulse signal into segments, and discarded segments contaminated by sudden non-rigid movements to obtain the heartbeat by normalizing the value of the signal.

In research on non-contact blood pressure measurement, the flow velocity of blood vessels is mainly calculated through PPG, and the blood pressure is deduced from parameters such as the outer diameter of the blood vessel wall, the thickness of the blood vessel, and the blood density. The waveform is measured at different positions as blood flows through the same artery. The waveform changes according to the difference in the pulse transit time (PTT); thus, the relative blood pressure can generally be obtained by dividing the transit time of the pulse wave between two points by the length of the blood vessel between the two points. Lewington[10] used the Moens–Korteweg (MK) equation and Hughes equation to express the relationship between pulse wave velocity (PWV) and blood pressure. Ma *et al.*[11] measured the two ends of the arm and used the time difference of the signals to calculate the relevant parameters. Fan *et al.*[12] proposed that multiple records of systolic blood pressure (SBP), diastolic blood pressure (DBP), and PWV of a test subject can be used to generate a linear regression equation to predict their SBP and DBP. Jeong and Finkelstein[13] used a high-speed camera and light source to measure the image-PTT (iPTT) signals of the face and palm, then used the relationship between the actual test subject's blood pressure and the image-PPG (iPPG) linear regression equation after obtaining the PWV parameters, allowing the blood pressure of a human body to be read simply from images. Deep learning methods have been used to improve the performance of conventional contactless methods for heart rate measurement, with PhysNet generating the best measurement outcome among the deep learning methods.[14] Even though research on non-contact physiological measurement using facial video has made considerable progress in the past few years, there are still some limitations with this technology, such as the lack of challenging datasets, the long time required for the estimation process, and the non-portability of the system.[15] Liu *et al.*[16] proposed and demonstrated the effectiveness of personalized video-based cardiac measurement for non-contact pulse and heart rate monitoring, where only 18 s of the

video was used for customization.

Two main contributions are described in this paper: the first is a portable AI service robot and the second is a method that requires only 6 s of video to predict the subject's heart rate and blood pressure. The rest of the paper is structured as follows. The hardware structure of the robot is introduced in Sect. 2, and the materials and methods are described in Sect. 3. The experimental results and conclusion are presented in Sects. 4 and 5, respectively.

2. Hardware

The AI robot designed for this study is shown in Fig. 1. It is a service robot with a height of 160 cm. The specifications of the computer inside the robot are an Intel® Core i7-9750H CPU, 8 GB RAM, and a GeForce GTX 1650 graphics card. An Intel RealSense SR300 camera is set up on the robot head. The resolution of the experimental images is 640 × 480, and 60 images can be taken per second, i.e., 60 frames per second (FPS) or 60 Hz. The depth range of the camera is from 0.2 to 1.5 m. A Surface Go tablet computer is placed on the robot's body to display information on the human heartbeat, blood pressure, and any abnormalities in the heart rhythm predicted through the algorithm. The actual blood pressure and heart rate measurement in the experiment is performed using an OMRON HEM-7121 arm sphygmomanometer. The research is carried out using the Robot Operating System (ROS) with OpenCV3.4.2 and Dlib19.21.0.

3. Methods

In this study, the following methods are used to achieve fixed-location non-contact measurement of the heart rate and blood pressure and to determine whether the heart rate is abnormal. The non-contact heartbeat and blood pressure measurement system is implemented using RGB images captured by the camera in the AI robot. The overall structure of the measurement system is shown in Fig. 2. First, a video of about 6 s is recorded, wherein images include the correct positions of the face and hand through face detection (facial landmarks) and

Fig. 1. (Color online) AI robot.

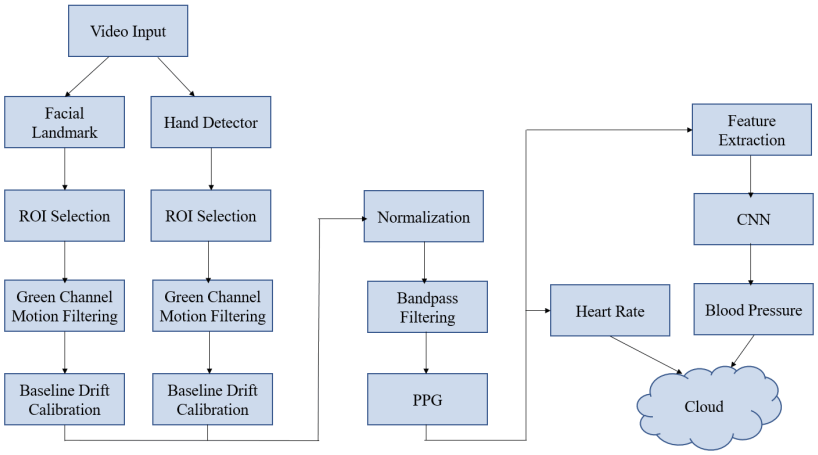
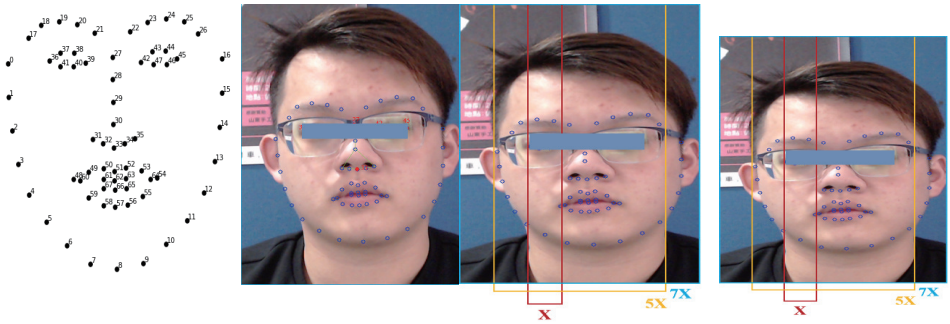


Fig. 2. (Color online) Flow chart of non-contact heart rate and blood pressure measurement system.

hand detection (hand detector). Then the optical flow method is used for motion correction, and the ROIs of the face and hand in the image are captured frame by frame as the output. Second, the ROIs of the hand and face obtained in the first part are pre-processed, the average value of the green channel of the RGB image is obtained, and then the average green values of the hand and face are scaled to a similar range through normalization, so as to facilitate the subsequent PPG waveform calculation. After calculating the normalized values, bandpass filtering is used to filter out the noise and obtain the PPG waveform in the frequency range of the heartbeat. Then, the delay time between the PPG of the hand and the face is combined with deep learning to predict the blood pressure. The average heart rate is calculated using the time difference between the PPG peaks. Finally, the third part uses convolutional neural network (CNN) models and pre-trained models for training. The training dataset uses the public MIMIC dataset, with the PPG waveforms of abnormal signals in the acquired dataset deleted to reduce their effect on the acquired characteristics of an irregular heart rate. The characteristics of arrhythmia are obtained during training and, finally, the trained model is incorporated into the model to judge the arrhythmia.

3.1 Extraction and tracking of facial feature points

After inputting the face image, the facial feature points (facial landmarks)[4] obtained by Dlib training are used to obtain 66 facial features, as shown in Fig. 3(a), and palm detection (blood pressure measurement) is performed. The locations of 68 (x, y) coordinates indicate the positions of facial structures on the face image obtained by Dlib facial landmark detector. To increase the speed of the calculation of the heart rate, the width of the eyes and nose [Fig. 3(b)] is calculated using the characteristics of the human face. The width of the face is about five times the width of the eyes, and the height of the face is about three times the height of the bridge of the nose, which is used as the basis for the search box. Therefore, the width of the right eye (subtraction of feature points 36 and 39) is calculated as Eq. (1), the width of the left eye (subtraction of feature points 42 and 45) is calculated as Eq. (2), and then the average value is obtained as Eq. (3). The



(a) (b) (c) (d)

Fig. 3. (Color online) Facial feature point extraction and tracking.

height of the bridge of the nose is calculated (subtraction of feature points 33 and 27) as Eq. (4). The width of the search box is set as seven times the width of an eye [Fig. 3(c)], and the height is set as five times the height of the bridge of the nose [Fig. 3(d)].

$$W_{RE} = \sqrt{(x_{36} - x_{39})^2 + (y_{36} - y_{39})^2} \tag{1}$$

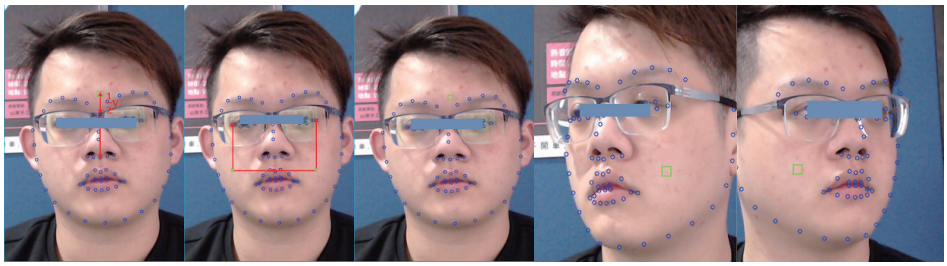
$$W_{LE} = \sqrt{(x_{42} - x_{45})^2 + (y_{42} - y_{45})^2} \tag{2}$$

$$W_{EYE} = (W_{LE} + W_{RE}) / 2 \tag{3}$$

$$W_{NOZE} = \sqrt{(x_{27} - x_{33})^2 + (y_{27} - y_{33})^2} \tag{4}$$

3.2 ROI calculation

The ROI of the study takes the junction between the bridge of the nose and the inner canthus of the two eyes as the center of the face. Therefore, when calibrating the position of the face search window, this junction is used as the center of the entire window. In this study, to extend the results to the mobile service robot for measurement and to consider the offset of the human face, the positions of the ROIs are set as the forehead (ROI_{FH}), the right cheek (ROI_{RF}), and the left cheek (ROI_{LF}). When the image is taken in front of the face, the ROI position of the forehead is set at the height of one-half of the bridge of the nose above the bridge of the nose [as shown in Fig. 4(a)], when on the side, the left and right sides and the junction below the outer canthus of both eyes, as shown in Fig. 4(b). To change the ROI position, the right cheekbone (facial feature point 2) and the nose tip (facial feature point 30) are subtracted to obtain the right face W_{RF} [Eq. (5)], as shown in Fig. 4(d); the left cheekbone (facial feature point 14) is subtracted from the nose tip (facial feature point 30) to obtain the left face WLF, as shown in Eq. (6). The ROIs of 30 × 30 pixels in the images of the left and right cheeks and the forehead of the human face are captured for testing using Eq. (7).



(a) (b) (c) (d) (e)

Fig. 4. (Color online) Selection and calculation of face ROIs. (a) Head ROI, (b) face ROI, (c) forehead ROI, (d) left Face fOI, and (e) right face ROI.

$$W_{RF} = \sqrt{(x_2 - x_{30})^2 + (y_2 - y_{30})^2} \tag{5}$$

$$W_{LF} = \sqrt{(x_{14} - x_{30})^2 + (y_{14} - y_{30})^2} \tag{6}$$

$$ROI = \begin{cases} ROI_{RF} & W_{RF} > 3W_{EYE} \\ ROI_{LF} & W_{LF} > 3W_{EYE} \\ ROI_{FH} & \text{otherwise} \end{cases} \tag{7}$$

Since blood pressure measurement requires measurement signals between two points, hand images are added. The hand detection in this study uses a model trained on TensorFlow[17] through a single shot detector (SSD) neural network to identify the hand. When detecting the hand, the hand is placed naturally on the table, palm facing up, with the hand and the head at the same level. When the hand is detected, the ROI is used to frame the palm and find the position of the palm center. The location of the ROI sampling point is shown in Fig. 5. The captured hand image is expanded by 30 × 30 pixels from the center point to obtain the ROI.

3.3 Signal processing and normalization

The ROI is divided into the B, G, and R channels using the split function. Since blood appears red because it absorbs green light and reflects red light, the absorption of green light changes significantly when blood flows through blood vessels;[4] therefore, G (green) is discussed separately for the channel waveform and channel signal. The ROI pixels for the signal of the G (green) channel are averaged using Eq. (8) to obtain I_m, as shown in Fig. 6. Also, according to the time series, a waveform composed of average pixel values is obtained, as shown in Fig. 7.

$$I_m = \sum_{i=1}^{30}\sum_{j=1}^{30} I(x_i, y_j) / (ROI) \tag{8}$$

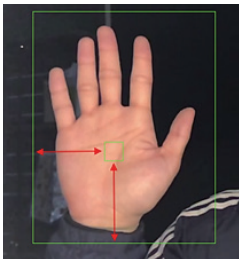


Fig. 5. (Color online) Palm ROI.

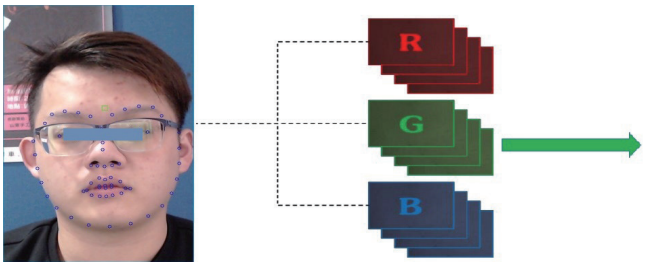


Fig. 6. (Color online) Using the signal of the G (green) channel.

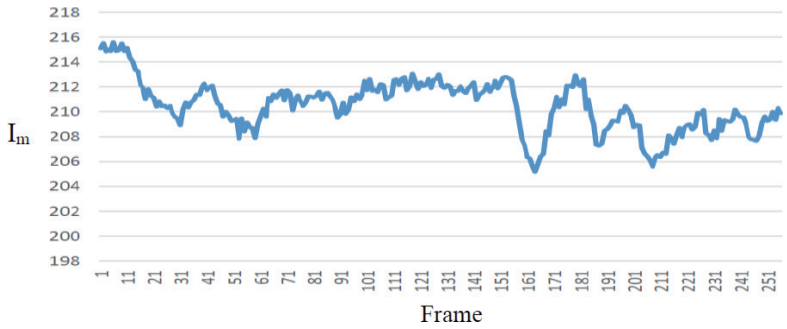


Fig. 7. (Color online) Waveform of average values of ROI.

To verify that the green component of the RGB image captures blood flow changes, we separate the RGB image into three components, R, G, and B images, as shown in Fig. 8. The left column of Fig. 8 shows the RGB image and the second column shows the three components, i.e., R, G, and B images. Next, we apply the proposed method to obtain the average PPG signals of the ROI of both the hand and face. The green channel of the PPG signal has a clear and regularly undulating rhythm. Thus, it is used in the study to predict the heart rate and blood pressure. To reduce the noise and jitter of the images, optical flow estimation is used to find pixelwise motion between consecutive images. We assume that $I(x, y, t)$ is the intensity at t, then the variation between two image frames is Δx, Δy, and Δt. Thus, Eq. (9) is satisfied for a constant intensity between two frames. From the Taylor expansion, Eq. (10) is satisfied, and $R(x, y, t)$ is close to zero. From simultaneous equations (9) and (10), we arrive at Eq. (11), which we rewrite as Eq. (12).

$$I(x,y,t) = I(x+\Delta x, y+\Delta y, t+\Delta t) \tag{9}$$

$$I(x+\Delta x, y+\Delta y, t+\Delta t) = I(x,y,t) + \frac{\partial I}{\partial x}\delta x + \frac{\partial I}{\partial y}\delta y + \frac{\partial I}{\partial t}\delta t + R(x,y,t) \tag{10}$$

$$\frac{\partial I}{\partial x}\Delta x + \frac{\partial I}{\partial y}\Delta y + \frac{\partial I}{\partial t}\Delta t = 0 \tag{11}$$

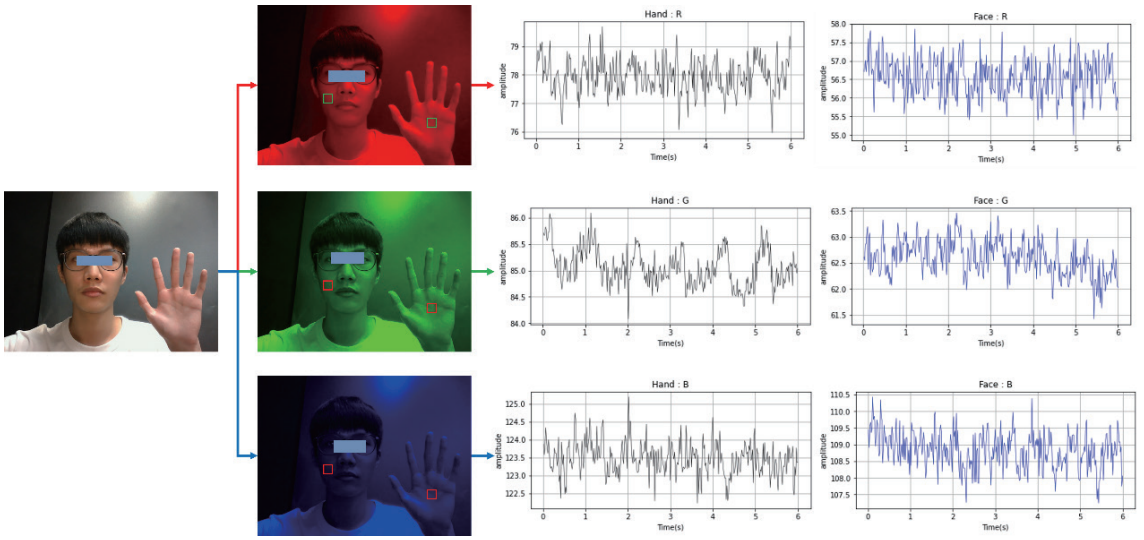


Fig. 8. (Color online) RGB and three components, i.e., R, G, and B images.

$$\frac{\partial I}{\partial x}\frac{\Delta x}{\Delta t}+\frac{\partial I}{\partial y}\frac{\Delta y}{\Delta t}+\frac{\partial I}{\partial t}\frac{\Delta t}{\Delta t}=0 \tag{12}$$

Writing $\Delta x/\Delta t$ and $\Delta y/\Delta t$ as the derivatives of the pixel along the x and y directions and V_x and V_y as the velocity components along the x and y directions, respectively, Eq. (12) can be simplified to Eqs. (13) and (14). Finally, using a set value with points q_1 to q_n to solve V_x and V_y gives the optical flow corresponding to the Lucas–Kanade optical flow method[18] in Eq. (15).

$$\frac{\partial I}{\partial x}V_x+\frac{\partial I}{\partial y}V_y+\frac{\partial I}{\partial t}=0 \tag{13}$$

$$I_xV_x+I_yV_y=-I_t \tag{14}$$

$$\begin{aligned} I_x(q_1)V_x+I_y(q_1)V_y&=-I_t(q_1)\\ I_x(q_2)V_x+I_y(q_2)V_y&=-I_t(q_2)\\ &\vdots\\ I_x(q_n)V_x+I_y(q_n)V_y&=-I_t(q_n) \end{aligned} \tag{15}$$

To make the results visible before and after correction by optical flow, we use Fig. 9 to explain the results of correction by optical flow. In Fig. 9, the left and right sides are the signals obtained before and after optical flow correction, respectively. Figures 9(a) and 9(c) show the PPG signals obtained from hand and face ROIs without correction, respectively. Thus, the results of bandpass filtering in Fig. 9(e) are less regular waveforms. Figures 9(b) and 9(d) show the PPG

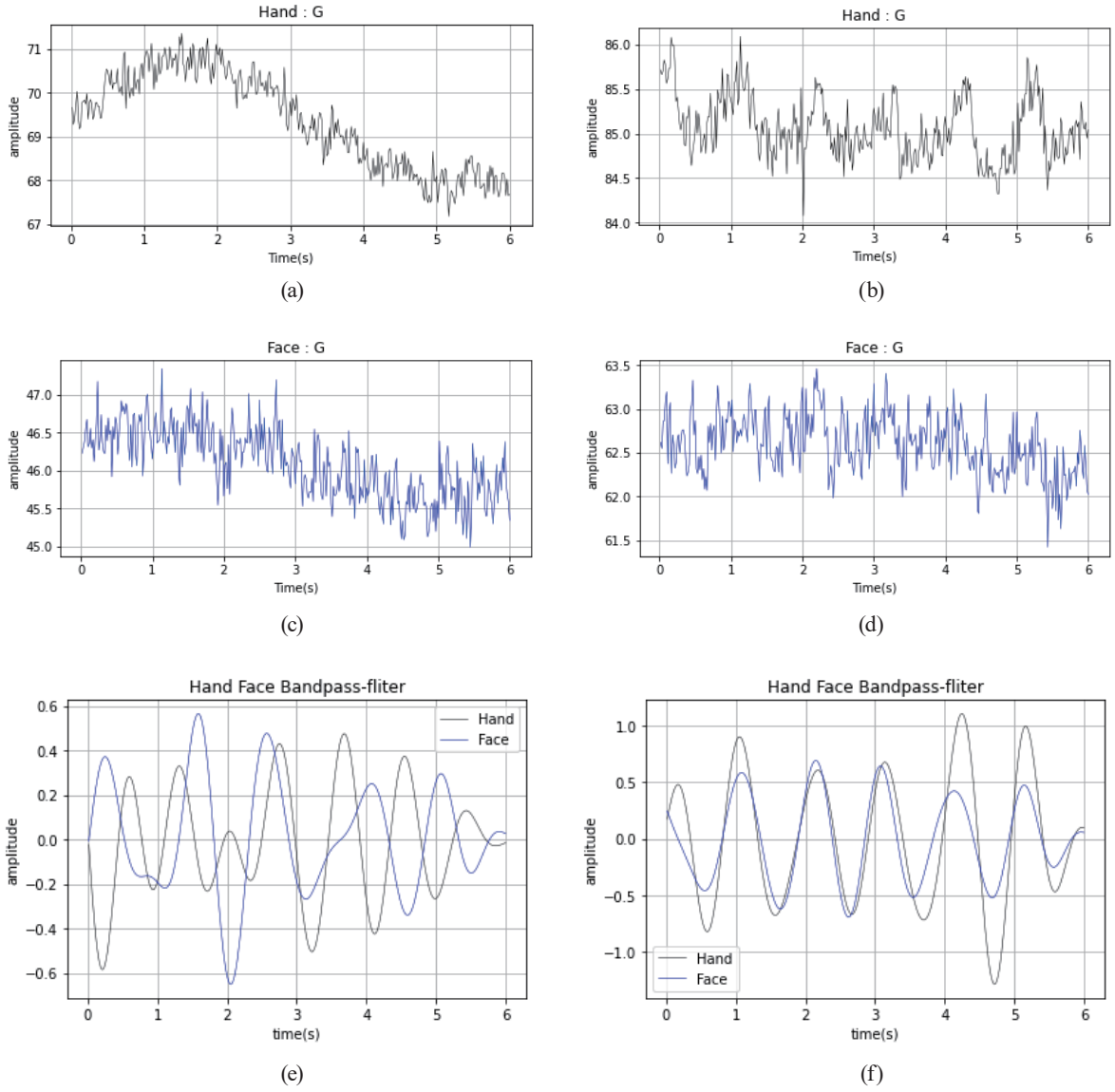


Fig. 9. (Color online) Waveforms before and after optical flow correction.

signals obtained from hand and face ROIs after the optical flow correction, respectively. We find that the corrected PPG signals obtained by bandpass filtering in Fig. 9(f) are more regular waveforms. Hence, it can demonstrate the difference resulting from optical flow correction.

Optical flow methods can eliminate the effect of instantaneous changes in the light of the external environment. However, the movement of the subject's body back and forth may change the value of the green light component, resulting in the displacement of the visually presented waveform. We use baseline drift calibration to detrend the displacement component. We divide the 6 s video into three parts of 2 s each to detrend each component using a linear relationship. When the data baseline is corrected, we normalize the data of the hand and face to facilitate the subsequent calculation of the blood pressure and heartbeat, which makes it easier to observe the relationship between the PPG waveforms of two different parts, as shown in Fig. 10.

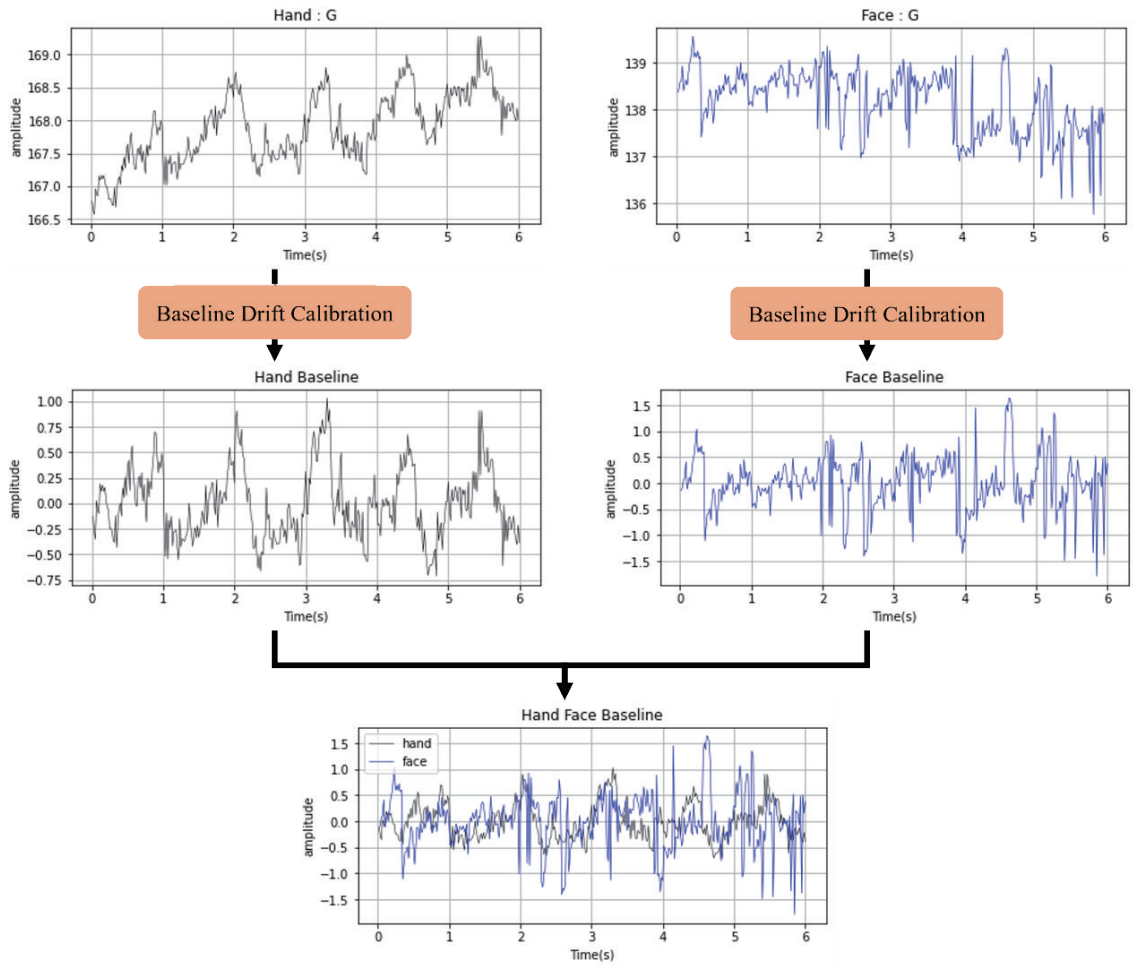


Fig. 10. (Color online) Normalization and baseline calibration of hand and face PPG waveforms.

3.4 Bandpass filtering and heart rate calculation

The obtained time domain signal $x[n]$ is converted into a frequency domain signal $y[k]$, as shown in Eq. (16), and then the frequency peak of the image is analyzed. The index in the frequency domain signal, k in Eq. (16), i.e., the frequency length after converting from the original time length n to the frequency, is used for systematic analysis and heart rate calculation. Since the heartbeat is about 40–100 beats per minute (about 0.7–4 Hz) for a human not exercising, a bandpass filter with frequency 0.7–4 Hz is used. The results are shown in Fig. 11.

$$y[k] = \sum_{n=0}^{N-1} e^{\frac{2\pi}{N}nk} x[n] \quad k = 0,1,\cdots,N-1 \tag{16}$$

To predict the heart rate, we use the interval between the peaks of the waveform (pulse wave interval, PWI) shown in Fig. 12. The peak times for the face and hand are calculated, the average

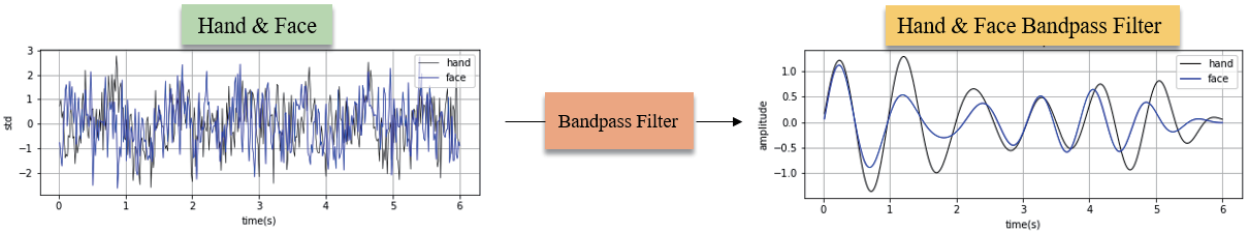


Fig. 11. (Color online) Bandpass filtering results of hand and face PPG waveforms.

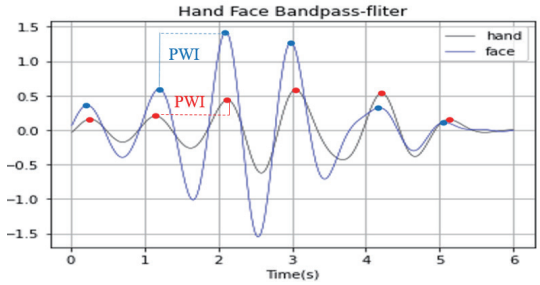


Fig. 12. (Color online) Peaks of the face and hand signal waveforms.

value, PWI_m, of PWI is obtained using Eq. (17), and the reciprocal of the average value multiplied by 60 is used as the heart rate [Eq. (18)].

$$PWI_m = \sum_{i=1}^{n} PWI_i / n \tag{17}$$

$$HR_m = 60 / PWI_m \tag{18}$$

3.5 Deep learning to predict blood pressure

After obtaining the PPG signals of the hand and face, three physiological characteristics are calculated, namely, the interval between the peaks, called PWI, PTT, and PWV as shown in Fig. 13. These three values are used as the input of the training data of the CNN for training, and the output is the SBP and DBP values. This CNN model uses 1D three-stroke features as the input of three neurons. The output is composed of two dense layers (Dense); each dense layer has 64 neuron outputs and finally outputs a neuron for prediction. The model is matched with the ReLU activation function, wherein the optimizer uses the Root Mean Square Prop (RMSprop) optimizer, and the loss function uses the mean square error (MSE). The architecture of the CNN is shown in Fig. 14.

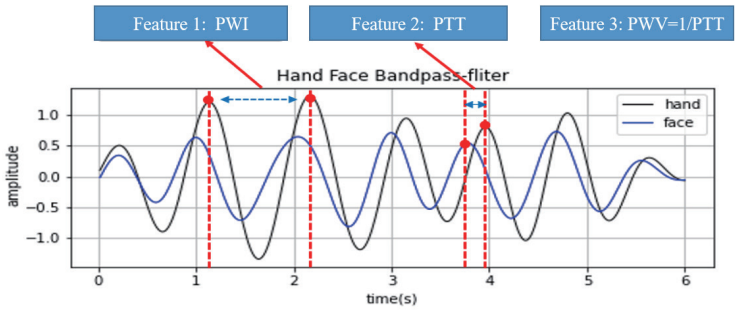


Fig. 13. (Color online) Features of deep learning to predict blood pressure.

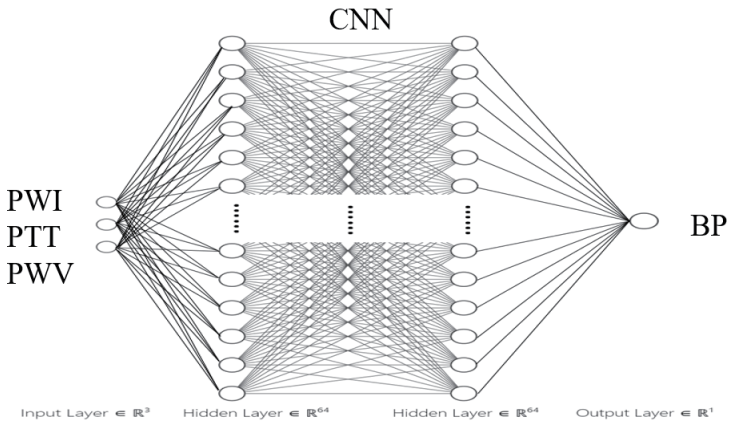


Fig. 14. (Color online) CNN architecture diagram for deep learning to predict blood pressure.

3.6 Prediction of arrhythmia

The arrhythmia study is conducted using MIMIC data within the PhysioNet dataset, which comprises physiological data records from more than 90 intensive care unit (ICU) patients. The PPG signals in the MIMIC dataset are trained. Figure 15 shows the PPG signals of normal and abnormal heart rates. Next, we introduce PPG signals of different subjects into the 1D CNN model for training and identification and arrange the CNN architecture as shown in Fig. 16.

4. Results and Discussion

After only one year of research, we achieved the non-contact prediction of heartbeat, blood pressure, and irregular heartbeat through a fixed-point robot using visual images. The instrument used to measure blood pressure and heart rate in this study is an OMRON HEM-7121 arm blood pressure monitor, with accuracy of pressure measurement of ±3 mmHg and accuracy of pulse measurement of ±5%. The distance between the robot and the human is fixed at 30–40 cm, as shown in Fig. 17. To obtain the approximate distance between the user and the robot, the depth information of an RGBD camera is used. Images of the human face and hand are captured

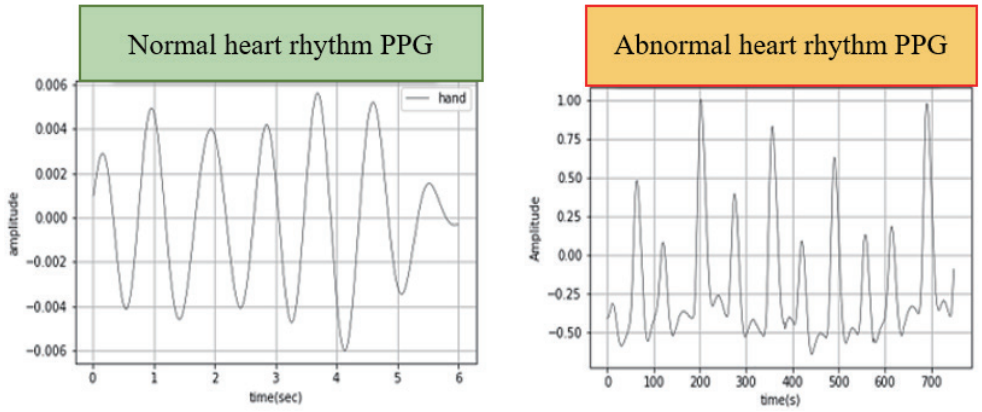


Fig. 15. (Color online) PPG waveforms of normal and abnormal heart rates in the dataset.

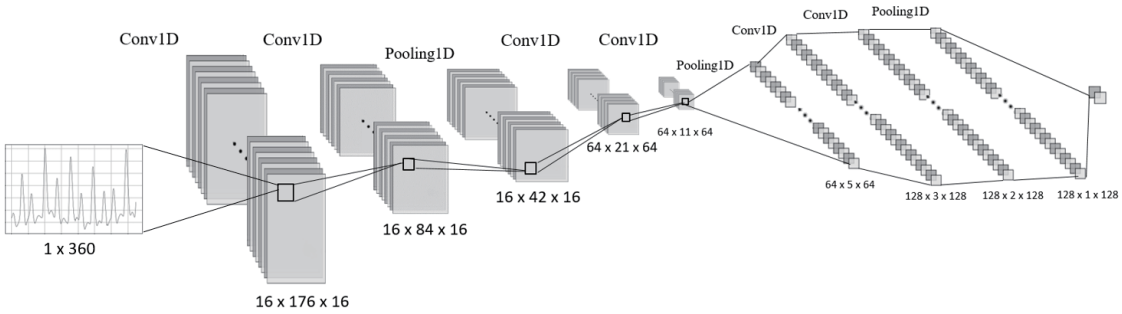


Fig. 16. (Color online) Structure of the CNN model for identifying arrhythmia.

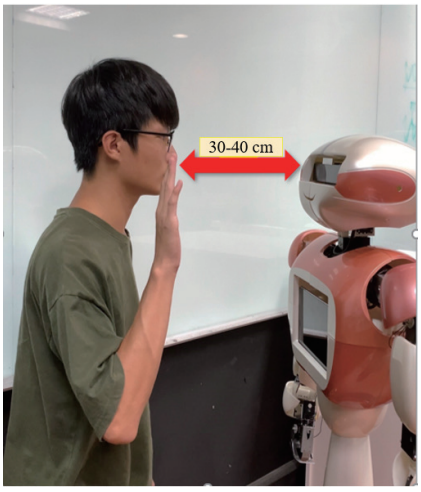


Fig. 17. (Color online) Distance between robot and human subject.

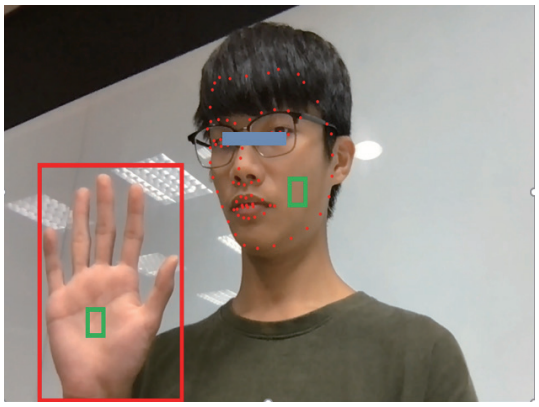


Fig. 18. (Color online) ROIs of face and palm.

through the camera on the robot for 6 s, and the ROIs of the hand and face are calculated as shown in Fig. 18. After adding and averaging the green channel signal in the RGB image in the

ROI and passing the signal through the motion filter, it becomes the PPG waveform of the face and hand. Since the human hand sometimes moves, when there is movement, a baseline drift correction is performed on the PPG signal; the 6 s signal is divided into three 2 s segments and then corrected. Finally, the results corrected for baseline drift are used.

4.1 Prediction of heart rate

As shown in Sect. 3.2, a smooth waveform can be obtained after normalizing the signal and bandpass filtering. Using Eqs. (17) and (18), we can predict the heart rate. Several experiments were conducted to verify the correctness of the research results. The heart rates predicted for the face and hand of 10 subjects are shown in Tables 1 and 2, respectively. From Table 1, it is found that the mean error percentage of the heart rate predicted from the face is 4.65% with a standard deviation of 1.80, while from Table 2, the mean error percentage of the heart rate predicted from the hand is 2.60% with a standard deviation of 1.07. Therefore, it is found that the PPG signal of the face leads to a larger error, which is due to the changes in the breathing and face. The ROI of

Table 1
Predicted heart rate from face.

Subject	PPG heart rate	Measured heart rate	Error (%)
1	70	68	2.94
2	72	68	5.88
3	70	65	7.69
4	66	63	4.76
5	68	67	1.49
6	70	66	6.06
7	76	73	4.11
8	63	66	4.55
9	73	69	5.80
10	64	62	3.23
Mean error (%)			4.65
Standard deviation			1.80

Table 2
Predicted heart rate from hand.

Subject	PPG heart rate	Measured heart rate	Error (%)
1	67	69	2.90
2	64	63	1.59
3	60	58	3.45
4	73	71	2.82
5	60	61	1.64
6	62	65	4.62
7	78	77	1.30
8	68	67	1.49
9	66	64	3.13
10	67	65	3.08
Mean error (%)			2.60
Standard deviation			1.07

the palm is less likely to change, resulting in a smaller error. To verify that the proposed method can provide an acceptable result for different people, five people are chosen to perform 10 measurements to predict their heart rate. The results are summarized in Table 3. The average percentage errors of the PPG heart rate from the hand and face are 3.74 and 5.33% with standard deviations of 1.85 and 3.09, respectively. This indicates that the results of the non-contact measurement for each subject are acceptable.

4.2 Prediction of blood pressure

Next, the characteristic values of the face and hand signals are calculated as extracted features, which are the intervals between waveform peaks or troughs [PWI, PTT, and pulse wave velocity (PWV)]. In this experiment, 200 samples are used, 90% for training and 10% for testing. During training, the validation_split function is used on 20% of the training samples for verification data. We set the training period as 100 times and the batch size as five times. The three characteristic values of the PPG signal are input to the CNN to predict the SBP and DBP values through machine learning. The results are shown in Tables 4 and 5, respectively. From Table 4, the mean error range of SBP predicted for 15 subjects is 1.69% with a standard deviation of 1.03, while in Table 5, the mean error range of DBP predicted for 15 subjects is 1.32% with a standard deviation of 1.29. Therefore, it is found that blood pressure can be accurately predicted through deep learning using images. Our results are compared with those calculated using Fan's[12] linear equation, as shown in Tables 4 and 5 for the predicted SBP and DBP, respectively. From Table 4, the mean error range of SBP predicted for 15 subjects is 2.23% with a standard deviation of 0.85, while in Table 5, the mean error range of DBP predicted for 15 subjects is 3.94% with a standard deviation of 2.31. Therefore, it is found that our deep learning method can predict blood pressure more accurately than Fan's method.

To demonstrate the repeatability and deviation of the non-contact measurement for each subject, the SBP and DBP for five subjects for 10 measurements are predicted. The results are summarized in Table 6. The mean percentage errors for SBP and DBP are 1.24% and 2.52% with standard deviations of 0.81 and 1.82, respectively. Hence, we can conclude that the proposed methods can correctly predict users' blood pressure.

Table 3
Predicted heart rate for different subjects for 10 measurements.

Subject	Measured mean heart rate	PPG heart rate from hand	PPG heart rate from face	Mean error of heart rate from hand (%)	Mean error of heart rate from face (%)	Standard deviation of error from hand	Standard deviation of error from face
Subject 1	62.2	62.0	61.7	3.21	5.36	1.28	2.18
Subject 2	60.9	61.9	62.9	3.64	5.69	1.90	4.29
Subject 3	64.8	63.4	62.4	3.37	5.52	1.91	3.15
Subject 4	65.5	63.9	62.7	4.29	5.16	1.96	2.10
Subject 5	67.9	66.0	64.8	4.19	5.90	2.20	3.73
Mean				3.74	5.53	1.85	3.09

Table 4
Predicted SBP.

Subject	Measured SBP	Our SBP	Our error (%)	Fan's SBP[12]	Fan's error (%)[12]
1	130	130	0.00	133	2.31
2	134	131	2.24	136	1.49
3	130	130	0.00	134	3.08
4	136	138	1.47	133	2.21
5	138	137	0.72	134	2.90
6	139	136	2.16	133	4.32
7	129	131	1.55	131	1.55
8	138	135	2.17	134	2.90
9	134	130	2.99	136	1.49
10	129	132	2.33	132	2.33
11	133	134	0.75	135	1.50
12	133	131	1.50	135	1.50
13	134	132	1.49	132	1.49
14	135	132	2.22	137	1.48
15	135	130	3.70	131	2.96
Mean error (%)			1.69		2.23
Standard deviation			1.03		0.85

Table 5
Predicted DBP.

Subject	Measured DBP	Our DBP	Our error (%)	Fan's DBP[12]	Fan's error (%)[12]
1	66	66	0.00	65	1.52
2	59	60	1.69	63	6.78
3	66	66	0.00	65	1.52
4	62	62	0.00	64	3.23
5	60	62	3.33	64	6.67
6	60	62	3.33	64	6.67
7	62	62	0.00	64	3.23
8	61	60	1.64	64	4.92
9	64	63	1.56	65	1.56
10	62	62	0.00	64	3.23
11	60	62	3.33	64	6.67
12	63	62	1.59	65	3.17
13	59	60	1.69	63	6.78
14	62	62	0.00	64	3.23
15	64	63	1.56	64	0.00
Mean error (%)			1.32		3.94
Standard deviation			1.29		2.31

Table 6
Predicted SBP and DBP for different subjects for 10 measurements.

Subject	Measured mean SBP	PPG mean SBP	Measured mean DBP	PPG mean SBP	Mean error of SBP (%)	Mean error of DBP (%)	Standard deviation of SBP error	Standard deviation of DBP error
Subject 1	129.5	130.8	62.7	63.3	1.32	1.99	0.74	2.55
Subject 2	128.6	129.4	59.2	61.3	1.09	3.64	0.55	2.33
Subject 3	131.5	131.2	62.1	61.4	0.99	2.07	0.71	1.28
Subject 4	135.4	134.1	64.9	64.1	0.95	2.16	0.68	1.48
Subject 5	139.8	137.4	68.8	66.9	1.83	2.73	1.38	1.42
Mean					1.24	2.52	0.81	1.82

Fig. 19. (Color online) AI robot displaying the measurement results on its panel.

To inform the user of their heart rate and blood pressure, the obtained heart rate and blood pressure are displayed on the robot panel, as shown in Fig. 19. The panel displays the subject's predicted heart rate, blood pressure value, and information such as whether the heart beat is regular. The panel of the AI robot in Fig. 19 indicates that the heart rate is 68 bpm and SBP and DBP are 136 and 63, respectively. The heart beat is normal with no arrhythmia.

5. Conclusions

We propose an AI robot with non-contact measurement technology to predict a subject's heart rate and blood pressure and to determine whether the heart beat is abnormal. The AI robot uses its panel to ask the subject to stand 30 to 40 cm in front of it with one hand lifted for about 6 s, then the robot displays the subject's heart rate and blood pressure and whether they are experiencing arrhythmia. The results indicate that the prediction error is within 2.5%. For further study, the AI robot will be trained to automatically trace a subject's location to perform dynamic contactless measurement.

Acknowledgments

This work was supported by the Ministry of Science and Technology of Taiwan, R.O.C., under grant number MOST 109-2221-E-130-014.

References

1 W. Gerin, G. Ogedegbe, J. E. Schwartz, W. F. Chaplin, T. Goyal, L. Clemow, and T. G. Pickering: J. Hypertens. **24** (2006) 67.
2 E. D. Manios, E. A. Koroboki, G. K. Tsivgoulis, K. M. Spengos, I. K. Spiliopoulou, F. G. Brodie, and N. A. Zakopoulos: Am. J. Hypertens. **21** (2008) 153.
3 R. Huang, W. Su, S. Zhang, and W. Qin: J. Comput. Commun. **7** (2019) 17.

4 W. Verkruysse, L. O. Svaasand, and J. S. Nelson: Optics Express **16** (2008) 21434.
5 M. Z. Poh, D. J. McDuff, and R. W. Picard: IEEE Trans. Biomed. Eng. **58** (2010) 7.
6 C. H. Hsieh, C. M. Kuo, and K. Tseng: Image **2** (2012) 16.
7 M. Lewandowska, J. Rumiński, T. Kocejko, and J. Nowak: Federated Conf. Computer Science and Information Cystems (IEEE, 2011) 405–410.
8 H. Monkaresi, R. A. Calvo, and H. Yan: IEEE J. Biomed. Health Inf. **18** (2013) 1153.
9 X. Li, J. Chen, G. Zhao, M. Pietikainen: Proc. IEEE Conf. Computer Vision and Pattern Recognition (2014) 4264-4271.
10 S. Lewington: Lancet **360** (2002) 1903.
11 Y. Ma, J. Choi, A. Hourlier-Fargette, Y. Xue, H. U. Chung, J. Y. Lee, and Y. Huang: Proc. National Acad.Sci. **115** (2018) 11144–11149.
12 X. Fan, Q. Ye, X. Yang, S. D. Choudhury: J. Ambient Intell. Hum. Comput. **11** (2020) 4329.
13 I. C. Jeong and J. Finkelstein: J. Med. Syst. **40** (2016) 1.
14 A. Ni, A. Azarang, and N. Kehtarnavaz: Sensors **21** (2021) 3719.
15 H. Wang, Y. Zhou, and A. El Saddik: Comput. Electr. Eng. **95** (2021) 107392.
16 X. Liu, Z. Jiang, J. Fromm, X. Xu, S. Patel, and D. McDuff: Proc. Conf. Health, Inference, and Learning (2021) 154–163.
17 V. Dibia: GitHub Repository (2017).
18 N. Sharmin and R. Brad: Sensors **12** (2012) 12694.

About the Authors

Shu-Yin Chiang received her M.S. and Ph.D. degrees from the University of Michigan, Ann Arbor, in 1994 and 1999, respectively, both in electrical engineering. She is currently dean of the School of Information Technology and a professor in the Department of Information and Telecommunications Engineering, Ming Chuan University, Taiwan, R.O.C. Her research interests include intelligent control, robot applications, and wireless sensor networks. (sychiang@mail.mcu.edu.tw)

Yi-Feng Chen received his B.S. and M.S. degrees from Ming Chuan University, Taiwan, R.O.C., in 2020 and 2021, respectively, both in information and telecommunications engineering. His research interests include AI and robot applications.