

Received August 30, 2020, accepted October 2, 2020, date of publication October 9, 2020, date of current version October 21, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3029842

Evaluating Quality of Photoplethymographic Signal on Wearable Forehead Pulse Oximeter With Supervised Classification Approaches

SHING-HONG LIU¹, (Member, IEEE), HAN-CHI LIU², WENXI CHEN³, (Member, IEEE), AND TAN-HSU TAN¹, (Senior Member, IEEE)

¹Department of Computer Science and Information Engineering, Chaoyang University of Technology, Taichung City 413310, Taiwan, R.O.C.

²Department of Electrical Engineering, National Taipei University of Technology, Taipei City 10608, Taiwan, R.O.C.

³Biomedical Information Engineering Laboratory, The University of Aizu, Aizuwakamatsu 965-8580, Japan

Corresponding author: Tan-Hsu Tan (tthan@ntut.edu.tw)

This work was supported in part by the Ministry of Science and Technology under Grant MOST 109-2221-E-324-002-MY2.

ABSTRACT Pulse oximeter is a common and important instrument in medical clinic, which uses the photoplethysmography to measure oxygen saturation ratio (SpO_2). However, the photoplethysmographic (PPG) signal is easily corrupted by the motion artifact when SpO_2 is measured in a dynamic scenario. Moreover, the probe of most pulse oximeters available in the market is finger-type clip, which is only suitable in the static scenario. This study aims to develop a wearable forehead pulse oximeter which could be used in a dynamic scenario, and uses two classification approaches, support vector machine (SVM) and convolutional neural network (CNN), for evaluating the qualities of its PPG signals. The higher the quality of PPG signals, the higher the accuracy of SpO_2 values. The SVM classified the SQI of PPG signal by twelve statistic features calculated from 7-second PPG segment. In the CNN approach, the PPG signals converted to an image was used as the input. The VGG-19 model, was then used to evaluate the SQI of PPG segments. Twenty subjects were recruited to perform the static and dynamic experiments. For the dynamic experiment, subjects ran on a treadmill with three different speeds, 3 km/hour, 5 km/hour and 7 km/hour. Experimental results indicated that the accuracies of SVM and CNN for SQI classifications were $89.9 \pm 1.18\%$ (mean \pm deviation %) and $88.7 \pm 1.54\%$, respectively. Then, the dynamic data were used to test the two classification models which were trained by the static data. The accuracies of SVM and CNN for SQI classifications were $89.5 \pm 3.87\%$ and $86.2 \pm 4.28\%$, respectively. The error ratios of SpO_2 in the case of static condition before and after the SQI classification with the SVM were respectively $5.6 \pm 6.6\%$ and $1.9 \pm 1.1\%$. The results suggested that the performance of SVM was better than CNN.

INDEX TERMS Pulse oximeter, oxygen saturation ratio, photoplethysmography, SVM, CNN.

I. INTRODUCTION

Oxygen saturation ratio (SpO_2) is an important parameter in clinical monitor, which can be measured by the finger-type pulse oximeter. Various devices for SpO_2 measurement have been developed in the past decades. With the advancement of technology, development of an accurate and ease-to-use device for SpO_2 measurement has attracted much attention in recent years [1]. In the SpO_2 measurement, two different specific wavelengths of light, red light and infrared (IR) light, are employed. A photodiode receives the light signals that are

transmitted through or reflected from the tissue. This noninvasive optical technique is the photoplethysmography (PPG). PPG-oriented sensors have two main types: transmission mode and reflection mode [2]. The finger-type pulse oximeter on the market uses transmission sensor, which probe usually is placed on peripheral parts of the body, such as fingertip or toe. The reflective sensor can be applied to other parts of body, such as forehead or wrist. Therefore, the reflective PPG sensor is more suitable for the long-term and wearable measurement of physiological signals [3].

The PPG signal consists of a tiny pulsating signal, or the alternating current (AC) signal, and a large baseline signal, or the direct current (DC) signal. The AC signal comes from

The associate editor coordinating the review of this manuscript and approving it for publication was Wenbing Zhao¹.

the change of arterial blood flow following heartbeats. The DC signal is mainly due to the light absorption of body tissues, such as skin, muscles, bones and venous blood. The SpO_2 is calculated according to the ratio of the AC and DC components of the PPG signals measured by the red and infrared (IR) lights, respectively [4].

Pulse oximeters have been used at the variety conditions, like as anesthesia, labor and delivery, emergency treatment, and cardiovascular or respiratory dysfunction. During exercise of patients with cardiopulmonary limitations, the pulse oximeter is useful in assessing the severity of disease, monitoring the effectiveness of therapeutic interventions, and determining the need for supplemental oxygen [5]. For the training of athletes, pulse oximeter could be used to assess the degree of hypoxemia [6]. Recently, the comparison between finger and forehead SpO_2 measurements has been presented [7], [8]. According to Sugino's study [7], the forehead pulse oximeter could detect the varying of hypoxia with faster speed and higher accuracy than the finger-type pulse oximeter. The reason is that shaking of a finger is large, and the corruption of PPG signal with the finger-type probe is severe during dynamic measurement. The forehead-type probe can be tightly tied to the forehead, and its PPG signal is not easily corrupted by the motion artifact. The study of Blaylock also showed that the pulse signal measured from the finger was susceptible to low temperature and low perfusion, which led to more loss of pulses than that from the forehead [8]. Thus, the accuracy of the SpO_2 measurement would be reduced. However, the commercial forehead pulse oximeter is suit to the static measurement, and needs the medical personnel to place the sensor probe at the proper position. Therefore, designing a wearable forehead pulse oximeter that could be easily used in the exercise training or home care will be a benefit for athletes or users with limited cardiopulmonary function.

The interference of motion artifacts and ambient light are two factors that cause the distortion of PPG signal, and thus decrease the accuracy of SpO_2 measurement [9]. The problem of ambient light could be overcome by the structural design of the sensor probe. But, the overcoming methods for the problem of motion artifacts are to detect the signal qualities of corrupted PPG segments or pulses, and delete them. The waveform of PPG pulse bears the regular morphological characteristics [10], [11]. According to these characteristics of PPG pulses, like as main peak, dicrotic notch, pulse width, and amplitude, many studies used the rule-based methods to determine the quality of each PPG pulse. Then, the signal quality index (SQI) represents the corrupted degree of the PPG pulse. Liu *et al.* have employed the fuzzy rules to determine the SQI of PPG pulses [12], and Fischer *et al.* have applied the characteristics of PPG waveform and the decision tree to classify its SQI [13]. Li *et al.* have used the Bayesian hypothesis testing method to analyze the SQI [14]. Some studies used the machine learning approaches to classify the SQI. Pereira *et al.* [15] and Chong *et al.* [16] used the support vector machine (SVM) to determine the SQI of PPG segment.

Liu *et al.* used the fuzzy neural network to evaluate the SQI of PPG pulse [17]. Liu *et al.* also used the deep learning approaches to classify the SQI of PPG pulse with the same data measured by the previous study [18]. They found that the performance of the latest deep learning methods was better than the traditional machine learning.

In this study, the goal was to develop a wearable forehead pulse oximeter to measure the SpO_2 under the static or dynamic scenarios. In order to improve the accuracy of SpO_2 , we evaluated the SQI of PPG signals, and used the PPG signals with the high SQI to calculate SpO_2 . For users easily wearing the pulse oximeter, we designed an array-type probe worn at the forehead. After signing the informed consent agreement, twenty subjects participated in this experiment which included the static and dynamic scenarios. The SQI of each 7-sec PPG segment was annotated as high or low according to morphology features of PPG pulses in this segment. Our SQI evaluation algorithms used two supervised learning approaches, SVM and CNN. The former approach used the statistical parameters of PPG segments as the input variables to train the model, and then the trained SVM was tested by the unknown PPG data. The latter approach used a 2-D CNN model to determine the SQI by converting each 7-sec PPG signal segment of the red light and IR light into a corresponding image. We used the 11-fold cross validation to test the performances of two supervised learning approaches. Moreover, because the static data of forehead pulse oximeter was easily measured, we explored whether the static data could be used to classify the SQI of dynamic data. Thus, we also used the static data to train two models, and then determining the SQI of dynamic data by the trained models. The results showed that the performance of SVM approach was better than the CNN approach.

II. HARDWARE IMPLEMENTATION AND SIGNAL PRE-PROCESSING

A forehead pulse oximeter was designed to measure the SpO_2 in static and dynamic scenarios, which had an array-type probe. The PPG signals of red and IR lights on this pulse oximeter were processed and used to calculate the beat-to-beat SpO_2 .

A. HARDWARE IMPLEMENTATION

The array-type probe consists of three MAX30102 PPG sensors (Maxim Integrated TM, San Jose, CA, USA). Each MAX30102 contains two light-emitting diodes with wavelengths of 660 nm (red light) and 880 nm (IR light), respectively. Moreover, a photo diode, an 18-bit analog-to-digital converter, and low-noise electronics with ambient light rejection are also included in this PPG sensor. Communication between sensors and microcontroller is accomplished via the standard I²C-compatible interface. The MCU (MSP430F5438A) is a 16-bit microcontroller (Texas Instruments TM, Dallas, TX, USA), which reconciles the workflow of the MAX30102, acquisition of PPG signals, digital signal processing to determine the pulse amplitude, and transmits

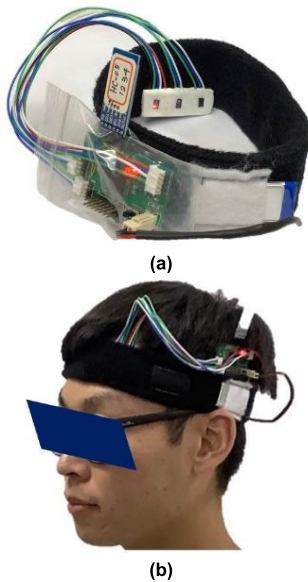


FIGURE 1. (a) The real photo of our designed forehead pulse oximeter, (b) The forehead pulse oximeter is worn on the position.

the data to a notebook PC by a Bluetooth v3.0 module using the UART protocol. The sampling rate was 100 Hz. Figure 1(a) shows the real photo of our designed forehead pulse oximeter, and Fig.1(b) shows when it is worn on the forehead position. The average DC components (digital code) of two light sources within one second must be the range from 1.8×10^6 to 1.9×10^6 according to the Maxim Integrated Products technique report [19]. Thus, if the average DC components for the red light and IR light were lower or higher than the range, the system would automatically adjust the power of LED light to maintain SpO_2 measurement for each subject at the same condition.

B. SIGNAL PRE-PROCESSING

The PPG signal comprises a very large DC component (PPG baseline signal) and a very small AC component (PPG pulse signal), and is always coupled with some noises due to motion artifact, ambient light and 60 Hz power line interference. Because the power spectrum of the PPG pulse signal is usually below 5 Hz [20], [21], the proposed system applied a fourth-order Butterworth low-pass filter with a cut-off frequency of 5 Hz to filter out high-frequency noise, and smooth the PPG pulse signal. A second-order Butterworth high-pass filter with a cut-off frequency of 0.3 Hz was used to filter out the PPG baseline signal. The PPG pulses were detected from the filtered PPG pulse signal by the differential method [22]. The differential PPG (dPPG) pulse has a main peak during the systolic duration which is significantly larger than the other peaks. Then, the systolic phase of the PPG pulse could be found by seeking the maximum value of the dPPG pulse. Figure 2 shows the PPG and dPPG pulse signals. According to the references [23], [24], the systolic duration is about 0.49 seconds. The maximum point (circle symbol) of each dPPG pulse is marked at the synchronous PPG pulse

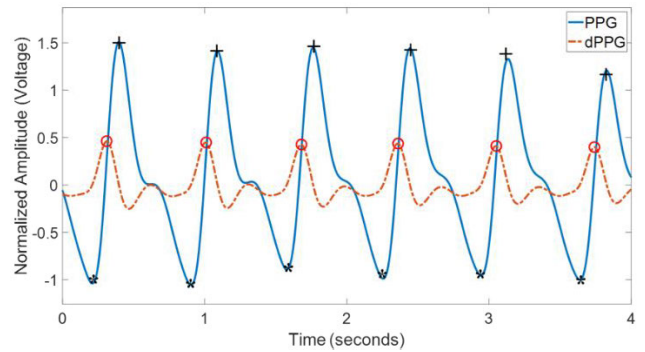


FIGURE 2. PPG signal (solid blue line) and dPPG signal (dashed red line).

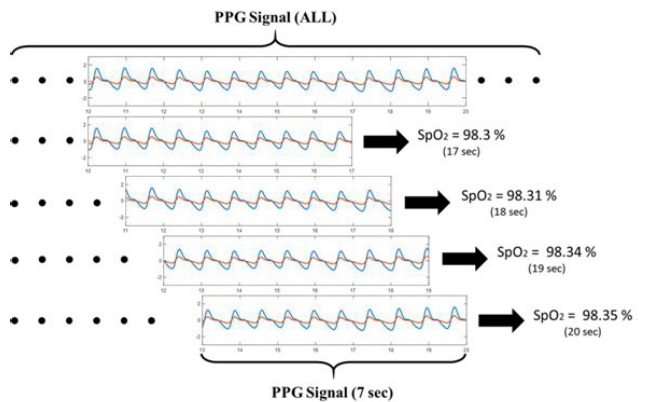


FIGURE 3. The PPG pulse signals, red light (red) and IR light (blue), are segmented by seven seconds, and the window shift is one second. The average SpO_2 in the segment represents the SpO_2 value of the segment.

as the center position of systolic phase. The valley (star symbol) of the PPG pulse is found backward to the minimum value within 25 points from the center point. The peak (cross symbol) of the PPG pulse is found forward to the maximum value within 25 points from the center point.

C. ESTIMATION OF OXYGEN SATURATION

The SpO_2 value is calculated based on the Maxim Integrated Products technique report [19]. The AC component is the amplitude of PPG pulse signal. The DC component is the average value of the PPG baseline signal within one second. Then the ratio R is obtained by AC and DC components of the red and IR lights by Eq. (1). Then, the SpO_2 value is calculated using Eq. (2), which is updated with one second. The PPG pulse signals, red light (red) and IR light (blue), are segmented by seven seconds, and the window is shifted by one second, as shown in Fig. 3. The SpO_2 of the segment is the average value of all PPG pulses.

$$R = \frac{AC_{Red}/DC_{Red}}{AC_{IR}/DC_{IR}} \quad (1)$$

$$SpO_2 = 104 - 17R. \quad (2)$$

III. SQI ANNOTATION OF PPG SEGMENTS

Figure 4 shows the flowchart of SQI annotation for the PPG segments. Four features are extracted from the PPG pulse.

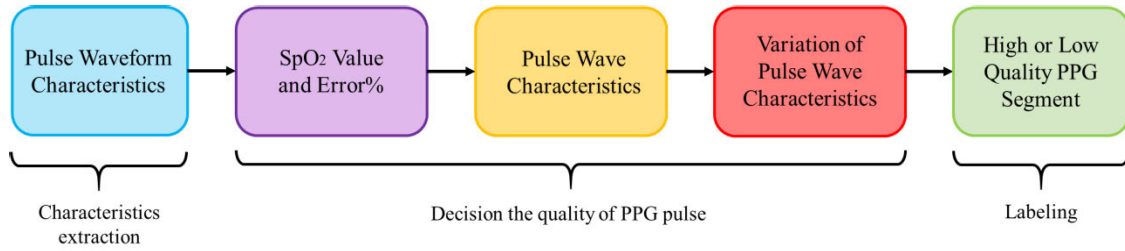


FIGURE 4. The flowchart of SQI annotation for the PPG segments.

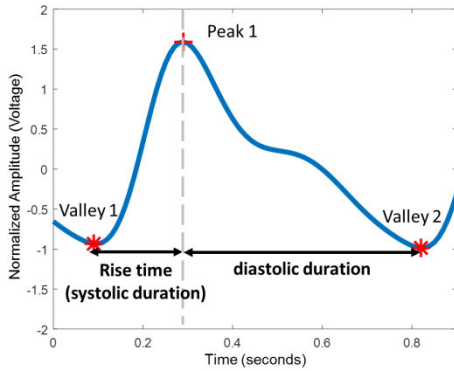


FIGURE 5. The characteristics of PPG wave.

Then, we annotate the SQI of PPG segment as high or low according to the SpO₂ value and its error ratio, the ranges and varieties of four characteristics.

A. CHARACTERISTICS OF PPG PULSE

There are four characteristics in the PPG pulse, as shown in Fig. 5, including the pulse wave amplitude (PWA), pulse wave duration (PWD), rise time (RT), and systolic-to-diastolic duration ratio (SDR). They are defined as follow,

$$PWA_{left} = V_{peak1} - V_{valley1}, \quad (3)$$

$$PWA_{right} = V_{peak1} - V_{valley2}, \quad (4)$$

$$PWD = T_{valley2} - T_{valley1}, \quad (5)$$

$$RT = T_{peak1} - T_{valley1}, \quad (6)$$

$$SDR = \frac{T_{peak1} - T_{valley1}}{T_{valley2} - T_{peak1}}, \quad (7)$$

where V_{peak1} is the voltage of peak1, $V_{valley1}$ is the voltage of valley1, T_{peak1} is the time of peak1, $T_{valley1}$ is the time of valley1, and $T_{valley2}$ is the time of valley2.

B. QUALITY OF PPG PULSE

There were three stages to determine the quality of the pulses within the PPG segment. In the first stage, the forehead SpO₂ measured by our apparatus was larger than 100%, or the error percentage (Error %) between the forehead SpO₂ and finger SpO₂ was larger than 5%. The finger SpO₂ measured by the commercial Rossmax SA310 (Rossmax, Taipei, Taiwan) was

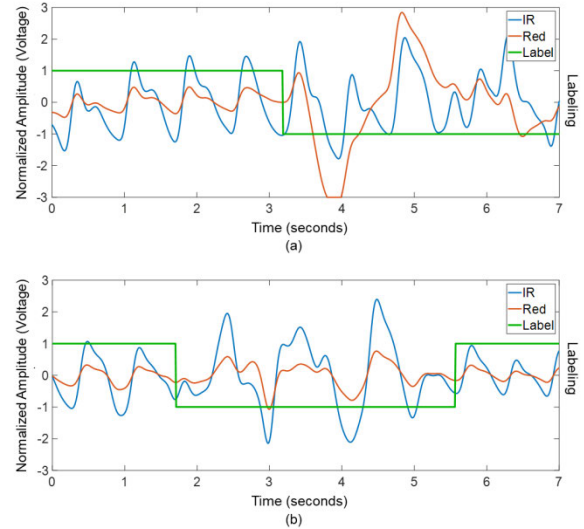


FIGURE 6. The PPG pulses are labeled according to the wave characteristics. The green line represents the pulse quality, -1 and 1 represent the low and high qualities. The red and blue lines are the PPG signals for the red light and IR light, respectively.

used as the reference data. The error percentage is defined as,

$$Error\% = \frac{|forehead\ SpO_2 - finger\ SpO_2|}{forehead\ SpO_2}. \quad (8)$$

If the pulse fits any of the above two conditions, the quality of the pulse is annotated low.

In the second stage, the following six conditions including the contour and four characteristics of pulse wave were used to determine the quality of pulse according to previous study [13]. If the pulse fits any condition, the quality of the pulse is annotated low. The six conditions are explained as following.

- (1) The pulse wave has the chopping on the peak or valley.
- (2) The absolute peak or valley value for red light is larger than the IR light.
- (3) The RT is smaller than 0.08 seconds, or larger than 0.49 seconds.
- (4) The SDR is larger than 1.1.
- (5) The $\min\{PWA_{right} / PWA_{left}, PWA_{left} / PWA_{right}\}$ is smaller than 0.4.
- (6) The PWD is larger than 2 seconds, or smaller than 0.3 seconds.

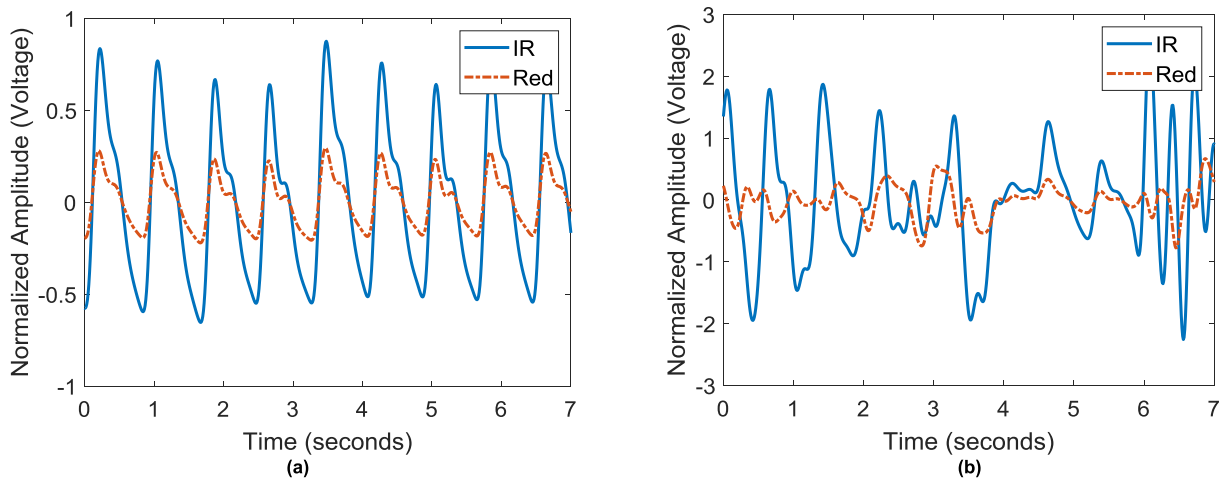


FIGURE 7. (a) shows the segment is annotated as high SQI, and (b) shows the segment is annotated as low SQI.

In the third stage, we considered the three-feature varieties of the neighbor pulses. There were three conditions. If the pulse fits any condition, the quality of the pulse is annotated low. The three conditions are explained as following, n represents the n th pulse.

(1) The ratio of $RT(n-1)$ and $RT(n)$ is out of the range between 33% and 300%.

(2) The ratio of $PWD(n-1)$ and $PWD(n)$ is out of the range between 33% and 300%.

(3) The ratio of $PWA(n-1)$ and $PWA(n)$ is out of the range between 50% and 200%.

Figure 6(a) shows the labeling results (green line) of all PPG pulses in one segment (red line for the red light, and blue line for the IR light). One pulse wave has the chopping on the valley for the red light. Moreover, the PPG pulses for the red light are more severely corrupted than the IR light. These pulses are all labeled as low quality (-1 value). Otherwise, the pulses are labeled as high quality (1 value). Figure 6(b) shows the PPG signals for the red light and IR light have the synchronous corrupted segment from 1.8 seconds to 5.7 seconds. This segment is annotated as low SQI.

C. SQI ANNOTATION OF PPG SEGMENT

When there were over one-fifth pulses labeled as low quality in the segment, this segment was annotated as low SQI. Figure 7(a) shows the segment annotated as high SQI, and Fig. 7(b) shows the segment annotated as low SQI.

IV. FEATURES EXTRACTION AND CLASSIFICATION

The supervised classification approaches used SVM and CNN to classify the SQI of the PPG segment being high or low. In the SVM approach, twelve statistic features extracted from the PPG signals for the red and IR lights in one segment were used to classify the types of SQI. In the CNN approach, the PPG pulse signals for the red and IR lights in one segment were transferred to a corresponding image which size was

300×300 pixels, as shown in Fig. 7. These images were fed into the 2-D CNN to classify the SQI.

A. FEATURES EXTRACTION

In one segment, there were two PPG signals, red light and IR light. We defined six statistic features for each PPG signal, the standard deviation of PWD (SD_{PWD}), the standard deviation of PWA (SD_{PWA}), the standard deviation of SDR (SD_{SDR}), the standard deviation of peak-to-peak interval (SD_{PPI}), the standard deviation of PPG pulse signal (SD_{pulse}), and the standard deviation of PPG baseline signal ($SD_{baseline}$).

$$SD_{PWD} = \sqrt{\frac{1}{N} \sum_{i=1}^N (PWD_i - \overline{PWD})^2}, \quad (9)$$

where N is the number of pulses in the segment, \overline{PWD} is the PWD mean of all pulses.

$$SD_{PWA} = \sqrt{\frac{1}{N} \sum_{i=1}^N (PWA_i - \overline{PWA})^2}, \quad (10)$$

where \overline{PWA} is the PWA mean of all pulses.

$$SD_{SDR} = \sqrt{\frac{1}{N} \sum_{i=1}^N (SDR_i - \overline{SDR})^2}, \quad (11)$$

where \overline{SDR} is the SDR mean of all pulses.

$$SD_{PPI} = \sqrt{\frac{1}{N} \sum_{i=1}^N (PPI_i - \overline{PPI})^2}, \quad (12)$$

where PPI is the peak-to-peak interval, \overline{PPI} is the PPI mean of all pulses.

$$SD_{pulse} = \sqrt{\frac{1}{M} \sum_{i=1}^M (AC_i - \overline{AC})^2}, \quad (13)$$

where M is the sampling number in the segment, $M = 700$, \overline{AC} is the mean of PPG pulse signal.

$$SD_{baseline} = \sqrt{\frac{1}{M} \sum_{i=1}^M (DC_i - \overline{DC})^2}, \quad (14)$$

where \overline{DC} is the mean of PPG baseline signal.

TABLE 1. Fundamental information about the VGG-19 layers and associated parameters of the network architecture.

Type	Filter Size	Channel Number	Input Size
Conv1	3 × 3	64	300 × 300 × 3
	3 × 3	64	300 × 300 × 64
Max pool	3 × 3	-	300 × 300 × 64
	3 × 3	128	150 × 150 × 64
Conv2	3 × 3	128	150 × 150 × 128
Max pool	3 × 3	-	150 × 150 × 128
	3 × 3	256	75 × 75 × 128
	3 × 3	256	75 × 75 × 256
Conv3	3 × 3	256	75 × 75 × 256
	3 × 3	256	75 × 75 × 256
Max pool	3 × 3	-	75 × 75 × 256
	3 × 3	512	37 × 37 × 256
	3 × 3	512	37 × 37 × 512
Conv4	3 × 3	512	37 × 37 × 512
	3 × 3	512	37 × 37 × 512
Max pool	3 × 3	-	37 × 37 × 512
	3 × 3	512	18 × 18 × 512
	3 × 3	512	18 × 18 × 512
Conv5	3 × 3	512	18 × 18 × 512
	3 × 3	512	18 × 18 × 512
Max pool	3 × 3	-	18 × 18 × 512
Flatten	-	1	9 × 9 × 512
Fc	-	1	41472
Out	-	2	1024

B. SUPPORT VECTOR MACHINE

SVM is a machine learning approach that has been widely used for solving supervised classification problems due to its generalization ability [25]. In essence, SVM classifies the maximum margin for the training data with a separating hyperplane that can be formulated as a quadratic optimization problem in feature space. The subsets of patterns closest to the decision boundary are called support vectors. Considering a linearly separable dataset $\{\bar{X}_i, D_i\}$, where \bar{X}_i is the input pattern for the i th example and D_i is the corresponding desired output (1, or -1), a hyperplane is found as the decision surface. It can be written as follows:

$$W^T \bar{X}_i + b \geq 0 \text{ then, } D_i = 1 \text{ } W^T \bar{X}_i + b < 0 \text{ then, } D_i = -1, \tag{15}$$

where W is the coefficients' vector of the hyperplane function. Cortes *et al.* proposed the soft-margin for SVM [26]. A slack variable (ξ_i) was added to get the margin which could tolerate some data points with the mistaken classification. Thus, the margin between the hyperplane and the nearest point is maximized and is considered a quadratic optimization problem:

$$\text{minimize } \frac{1}{2} \|W\|^2 + C \sum_{i=1}^i \xi_i, \quad \xi_i \geq 0. \tag{16}$$

$$\text{subject to } D_i(W^T \bar{X}_i + b) - 1 + \xi_i \geq 0. \tag{17}$$

where C is the cost parameter. When C is larger, the affection of the data with the mistaken classification on the Eq. (16) is larger. When W and b are rescaled, the point nearest to the hyperplane has a distance of $1/\|W\|$ [27]. Using Lagrange multipliers and the Kuhn-Tucker theorem, the solution is

given by:

$$W = \sum D_i \alpha_i \bar{X}_i. \tag{18}$$

Only a small fraction of α_i coefficient is nonzero. The corresponding pairs of $\{\bar{X}_i, D_i\}$ are known as support vectors and they define the decision boundary. All other input patterns multiplied with α_i values are rendered irrelevant. The hyperplane decision function for the input pattern vector can be written as follows:

$$f(\bar{X}) = \text{sign}(\sum D_i \alpha_i (\bar{X}_i^T \bar{X}_i) + b). \tag{19}$$

Via replacing the inner product ($\bar{X}_i^T \bar{X}_i$) with the kernel function $K(x_i, x_j)$, the input patterns are mapped to a higher dimensional space [27]. In this higher dimension space, a separating hyperplane is constructed to maximize the margin. In this study, the linear function and Gaussian radial basis function were used to construct the kernel function as follows:

$$K(x_i, x_j) = x_i^T x_j, \tag{20}$$

$$K(x_i, x_j) = \exp(-\gamma \|x_i - x_j\|^2). \tag{21}$$

C. CONVOLUTION NEURAL NETWORK

Because the number of samples is too few and the characteristics of patterns do not have big difference, we chose a 2-D CNN with the multiple layers. We build a 2-D CNN based on the pretrained CNN architecture with 19 layers (VGG-19) [28]. In the output layer, we replaced the 1000 fully-connection with the softmax activation by a 2 fully-connection. VGG-19 model was pretrained for object detection task on the ImageNet dataset [29]. The detailed description of VGG-19 model is shown in Table 1. All of the filters in VGG-19 are of 3 × 3 in size. The down-sampling

TABLE 2. The number of high and low SQI for twenty subjects in the static and dynamic experiments. (N is the number of PPG pulses).

Subject	SQI for Static		SQI for 3 km/hour		SQI for 5 km/hour		SQI for 7 km/hour	
	High (N)	Low (N)	High (N)	Low (N)	High (N)	Low (N)	High (N)	Low (N)
1	135	39	145	29	123	51	56	118
2	109	65	146	28	151	23	42	132
3	50	124	174	0	140	34	75	99
4	77	97	82	92	31	143	52	122
5	90	84	169	5	157	17	77	97
6	85	89	150	24	173	1	66	108
7	76	98	Non	Non	Non	Non	Non	Non
8	72	102	125	49	134	40	51	123
9	68	106	148	26	87	87	53	121
10	119	55	139	35	97	77	79	95
11	61	113	Non	Non	Non	Non	Non	Non
12	78	96	156	18	158	16	55	119
13	66	108	137	37	149	25	66	108
14	69	105	174	0	167	7	60	114
15	54	120	154	20	106	68	93	81
16	72	102	37	137	47	127	39	135
17	69	105	174	0	141	33	62	112
18	86	88	147	27	126	48	48	126
19	72	102	151	23	165	9	57	117
20	88	86	122	52	76	98	35	139
Total	1596	1884	2530	602	2228	904	1066	2066

is performed directly by maximum pooling layers that have a stride of 2, batch size is 16, learning rate is 0.00001, and batch normalization is performed right after each convolution and before ReLU activation. Two fully connected layers have sizes of 1024.

V. EXPERIMENTS

A series of static and dynamic experiments was conducted in which a total of 20 subjects (12 males and 8 females) were invited to participate in the experiments. All the subjects were healthy adults and aged 20 to 25 years old (21.7 ± 2.3 years, mean \pm standard deviation), weight was between 45 and 87 Kg (61.4 ± 11.5 Kg), height was between 151 and 185 cm (169.8 ± 9.6 cm) and heart rate was between 58.5 and 95.9 BPM (77.6 ± 11.0 beats/minute (BPM)). They did not have past histories of related major injury. This experiment was approved by the Research Ethics Committee of China Medical University & Hospital (No. CMUH107-REC1-167), Taichung, Taiwan.

A commercial medical pulse oximeter (Rossmax® SA310, Rossmax, Taiwan) with the finger clip was used to measure the beat-to-beat SpO₂ and was considered as the reference value in the study. The Rossmax® SA310 has the Bluetooth function to transmit the measured SpO₂ values to a computer. Thus, the SpO₂ measured by our designed forehead pulse oximeter could synchronize with the SpO₂ measured by the Rossmax® SA310. In the static experiment, subjects were asked to wear the forehead pulse oximeter on their heads, and

take the finger clip of the Rossmax SA310 on the index finger of right hand, as shown in Fig. 8. Subjects would be measured three minutes. In the first thirty seconds, they were requested to keep their body still, and to slightly move their heads left and right in the next thirty seconds. These activities would be repeated three times. When the subjects slightly move their heads, the PPG signals measured by forehead probe would be corrupted by the motion artifact. But, the PPG signals measured by finger-type probe were stable. Thus, some PPG segments for the forehead probe would be annotated as low SQI. Table 2 shows the number of high and low SQI samples for all subjects in the static experiment. The total sample has 3480. There are 1596 samples belonging to high SQI, and 1884 samples belonging to low SQI.

In the dynamic experiment, it was difficult to keep the fingers still when the subject was running on the treadmill. Then, the Rossmax® SA310 pulse oximeter could not measure the SpO₂. Thus, subjects only wore the forehead pulse oximeter. When PPG pulse signals were in stable status, subjects ran on the treadmill. In order to observe the corrupted influences of the PPG pulse signals for the different exercise intensities, the running speeds had three types, 3 km/h (walking), 5 km/h (brisk walking) and 7 km/h (jogging), respectively. At each speed, the data were recorded for 3 minutes. Table 2 shows the number of high and low SQI samples for all subjects. The total sample has 9396. There are 5824 samples belonging to high SQI, and 3572 samples belonging to low SQI. Subjects 7 and 11 did not perform the dynamic experiment.

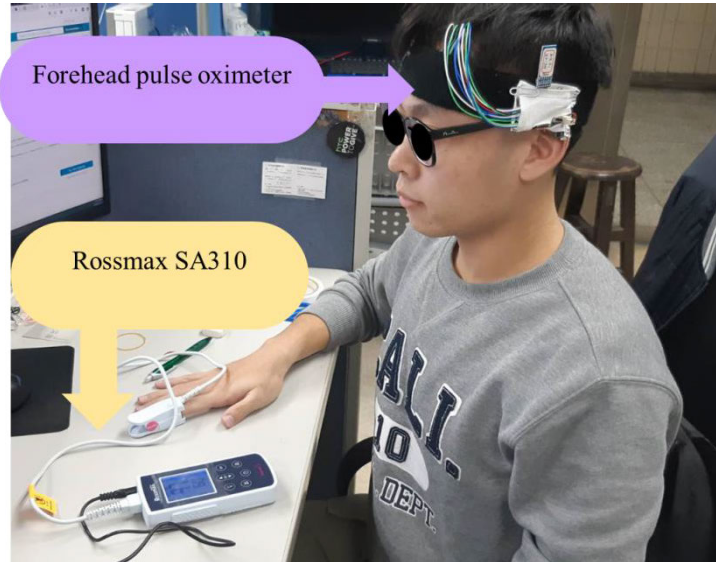


FIGURE 8. The real photo for the static experiment. Subject wears the forehead pulse oximeter and his index finger takes the finger clip of Rossmax® SA310.

TABLE 3. The training and testing sample numbers of high and low SQI for each fold.

Subjects	Training Data		Testing Data	
	High SQI (N)	Low SQI (N)	High SQI (N)	Low SQI (N)
1	3803	2635	3617	2821
2	3405	2511	4015	2945
3	3404	2512	4016	2944
4	3383	2533	4037	2923
5	3611	2305	3809	3151
6	3525	2391	3895	3065
7	3246	2670	4174	2786
8	3616	2822	3804	2634
9	3641	2797	3779	2659
10	3730	2708	3690	2748
11	3617	2821	3803	2635

According to our proposed method, a PPG segment was considered true positive (TP) as its quality level was correctly identified, false positive (FP) as its quality level was incorrectly identified, true negative (TN) as its quality level was correctly rejected, and false negative (FN) as its quality level was incorrectly rejected. Here, the performance of the proposed method was evaluated using accuracy, $(TP+TN)/(TP+FP+FN+TN)$, precision, $TP/(TP+FP)$, sensitivity, $TP/(TP+FN)$, and specificity, $TN/(FN+TN)$.

VI. RESULTS

In order to evaluate the performance of SVM and CNN approaches, we designed two analytic methods. First, we used the 11-fold cross validation to train and test the two approaches. We organized the eleven data sets according the subject's serial number. In the first set, the data of subject 1 to subject 10 was used as the training data, and the data of subject 11 to subject 20 was used as the testing data. The data included the static and dynamic data. Thus, the number of training and test data was 6438, separately. In the second set, the data of subject 2 to subject 11 was used as the training data, and the data of subject 12 to subject 20 and subject

1 was used as the testing data. According to this composed method, in the eleventh set, the training data was the data of subject 11 to subject 20, and the testing data was the data of subject 1 to subject 10. Second, we used the static data from all subjects as the training data, the dynamic data of all subjects as the testing data. Thus, the sample numbers of training and testing data were 3480 and 9396, respectively.

We used the first set to search the optimal values of parameters of kernel functions for SVM by the grid-search method. The range of C parameter was from 2^{-8} to 2^8 , and γ parameter was from 2^{-15} to 2^{15} . The step was 2 times. Figure 9 (a) shows the distribution of SVM performance with the linear kernel function. The highest accuracy is 91.2 % when the C is 64. Figure 9 (b) shows the distribution of SVM performance with the Gaussian radial basis kernel function. The highest accuracy is 92.1% when the C is 256 and γ is 0.125. Thus, for the training and testing of SVM, we chose these values.

A. 11-FOLD VALIDATION

Table 3 shows the training and testing sample numbers of high and low SQI for each fold. In the general analysis,

TABLE 4. The performance of SVM with linear kernel function for the 11-fold validation.

Data Set	Accuracy (%)	Specificity (%)	Sensitivity (%)	Precision (%)
1	91.2	88.3	93.4	91.1
2	90.4	86.2	93.5	90.2
3	89.7	84.6	93.4	89.2
4	88.5	84.1	91.6	88.8
5	87.7	87.4	87.9	89.4
6	88.5	85.3	91.0	88.7
7	89.3	86.9	91.0	91.2
8	89.2	86.2	91.2	90.5
9	90.0	86.4	92.6	90.6
10	90.5	86.6	93.3	90.3
11	89.8	87.9	91.2	91.5
Mean±SD	89.5 ± 0.98	86.4 ± 1.24	91.8 ± 1.61	90.2 ± 0.94

TABLE 5. The performance of SVM with Gaussian radial basis kernel function for the 11-fold validation.

Data Set	Accuracy (%)	Specificity (%)	Sensitivity (%)	Precision (%)
1	92.1	88.9	94.5	91.6
2	90.9	87.0	93.7	90.8
3	90.4	85.3	94.2	89.7
4	87.5	83.4	90.5	88.3
5	88.3	88.4	88.2	90.2
6	89.1	87.9	90.0	90.4
7	90.0	88.9	90.7	92.4
8	90.0	87.8	91.5	91.6
9	90.3	87.1	92.5	91.0
10	90.2	86.4	93.0	90.2
11	90.3	89.0	91.1	92.3
Mean±SD	89.9 ± 1.18	87.3 ± 1.64	91.8 ± 1.87	90.8 ± 1.15

TABLE 6. The performance of VGG-19 model for the 11-fold validation.

Data Set	Accuracy (%)	Specificity (%)	Sensitivity (%)	Precision (%)
1	92.1	93.6	90.9	94.8
2	90.5	85.5	94.3	89.8
3	89.3	90.5	88.4	92.7
4	87.7	89.5	86.4	91.9
5	86.7	88.8	85.1	90.2
6	88.3	83.7	92.0	87.7
7	87.9	94.3	83.8	95.8
8	87.3	90.3	85.2	92.7
9	89.1	89.3	89.0	92.2
10	88.1	89.4	87.1	91.7
11	89.1	92.3	86.8	94.2
Mean±SD	88.7 ± 1.54	89.7 ± 3.16	88.1 ± 3.23	92.2 ± 2.34

the sample numbers of high SQI are more than those of low SQI. Table 4 shows the performance of SVM with the linear kernel function for the 11-fold validation, the accuracy is $89.5 \pm 0.988\%$, sensitivity is $91.8 \pm 1.61\%$, specificity is $86.4 \pm 1.24\%$, and precision is $90.2 \pm 0.94\%$. Table 5 shows the performance of SVM with the Gaussian radial basis kernel function for the 11-fold validation, the accuracy is $89.9 \pm 1.18\%$, sensitivity is $91.8 \pm 1.87\%$, specificity is $87.3 \pm 1.64\%$, and precision is $90.8 \pm 1.15\%$.

For the VGG-19 model training, in order to reduce the difference between the sample numbers of high and low SQI, we extended few sample number to the more sample number for each fold. In the first fold, the sample number of high quality is 3803 and low SQI is 2635. Thus, the sample number of low SQI is extended to 3803. The epoch was 50 times. Table 6 shows the performance of VGG-19 model, the accuracy is $88.7 \pm 1.54\%$, sensitivity is $88.1 \pm 3.23\%$, specificity is $89.7 \pm 3.16\%$, and precision is $92.2 \pm 2.34\%$. According

to Tables 4 and 5, the performance and stability of SVM are better than the VGG-19 model.

B. CLASSIFICATION OF DYNAMIC DATA USING STATIC DATA-TRAINED MODELS

The second evaluation used the static data-trained models to classify the dynamic data. In Table 2, the total training number is 3480. There are 1596 samples belonging to high SQI, and 1884 samples belonging to low SQI. The total testing number is 9396. There are 5824 samples belonging to high SQI, and 3572 samples belonging to low SQI. Table 7 shows the testing performances of SVM with the linear kernel function at three speeds, 3 km/h, 5 km/h, and 7 km/h. The average accuracy is $89.1 \pm 3.53\%$, sensitivity is $90.8 \pm 4.67\%$, specificity is $77.3 \pm 18.87\%$, and precision is $90.9 \pm 0.35\%$. Table 8 shows the testing performances of SVM with the Gaussian radial basis kernel function at three speeds, 3 km/h, 5 km/h, and 7 km/h. The average

TABLE 7. The testing performances of SVM with linear kernel function at three speeds, 3 km/h, 5 km/h, and 7 km/h.

	Accuracy (%)	Specificity (%)	Sensitivity (%)	Precision (%)
3 km/h	88.8	58.0	96.1	90.6
5 km/h	85.8	78.2	88.9	91.0
7 km/h	92.9	95.7	87.3	91.3
Mean±SD	89.1 ± 3.53	77.3 ± 18.87	90.8 ± 4.67	90.9 ± 0.35

TABLE 8. The testing performances of SVM with Gaussian radial basis kernel function at three speeds, 3 km/h, 5 km/h, and 7 km/h.

	Accuracy (%)	Specificity (%)	Sensitivity (%)	Precision (%)
3 km/h	89.7	63.6	95.9	91.7
5 km/h	85.5	77.9	88.6	90.8
7 km/h	93.2	96.3	87.2	92.5
Mean±SD	89.5 ± 3.87	79.3 ± 16.39	90.6 ± 4.67	91.7 ± 0.85

TABLE 9. The testing performances of VGG-19 model at three speeds, 3 km/h, 5 km/h, and 7 km/h.

	Accuracy (%)	Specificity (%)	Sensitivity (%)	Precision (%)
3 km/h	90.2	63.3	96.6	91.7
5 km/h	86.7	62.6	96.4	86.4
7 km/h	81.7	75.0	94.7	66.1
Mean±SD	86.2 ± 4.28	67.0 ± 6.95	95.9 ± 1.08	81.4 ± 13.51

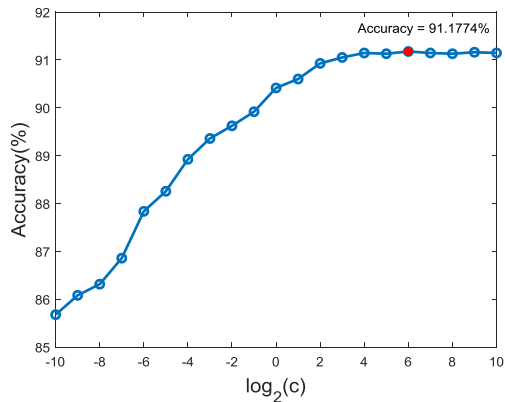
TABLE 10. The SpO₂ values were measured by the forehead pulse oximeter at static condition, 3 km/h, 5 km/h, and 7 km/h before and after the SQI classification with the SVM. (Sub. is the abbreviation of subject, E is the abbreviation of Error %, M is the abbreviation of mean).

Sub.	Before Classification					After Classification				
	Static	3 km/h	5 km/h	7 km/h	Static	3 km/h	5 km/h	7 km/h		
	SpO ₂ (%), M±SD	E (%)	SpO ₂ (%), M±SD		SpO ₂ (%)	E (%)	SpO ₂ (%), M±SD			
1	97.6 ± 0.4	0.82	99.5 ± 0.4	99.1 ± 0.3	98.2 ± 0.8	97.6 ± 0.2	0.82	99.4 ± 0.4	99.1 ± 0.2	99.1 ± 0.6
2	97.8 ± 1.2	1.48	96.4 ± 0.4	96.5 ± 0.4	96.6 ± 0.6	98.0 ± 0.6	1.36	96.4 ± 0.2	96.6 ± 0.3	96.2 ± 0.3
3	92.8 ± 31.9	9.07	98.1 ± 0.3	97.9 ± 0.8	97.7 ± 0.6	98.9 ± 0.3	3.05	98.1 ± 0.3	98.2 ± 0.7	98.0 ± 0.3
4	98.6 ± 1.1	1.68	98.4 ± 0.5	95.5 ± 11.7	97.7 ± 1.1	98.8 ± 0.4	1.60	98.6 ± 0.2	98.5 ± 0.7	98.7 ± 0.2
5	94.9 ± 3.7	2.55	98.3 ± 0.2	98.2 ± 0.2	97.0 ± 1.2	97.1 ± 1.7	0.81	98.3 ± 0.2	98.2 ± 0.2	97.5 ± 0.9
6	99.3 ± 0.1	3.98	98.0 ± 0.3	98.0 ± 0.4	97.9 ± 0.6	99.3 ± 0.1	4.13	97.9 ± 0.2	98.0 ± 0.4	98.3 ± 0.3
7	97.6 ± 1.9	1.76				98.8 ± 0.2	1.76			
8	78.2 ± 86.8	20.63	98.5 ± 0.4	98.6 ± 0.3	93.4 ± 2.6	97.3 ± 0.6	1.74	98.6 ± 0.4	98.6 ± 0.3	99.0 ± 0.2
9	98.3 ± 2.3	3.08	99.1 ± 0.5	99 ± 0.3	97.8 ± 1.3	99.0 ± 0.2	2.71	99.1 ± 0.5	99.1 ± 0.2	99.4 ± 0.4
10	98.1 ± 0.5	2.64	97.8 ± 0.3	97.9 ± 0.4	97.9 ± 0.3	98.3 ± 0.3	2.78	97.8 ± 0.3	97.9 ± 0.3	97.9 ± 0.2
11	93.4 ± 16.9	5.74				98.4 ± 0.4	0.91			
12	97.2 ± 1.4	1.08	98.1 ± 0.3	98.1 ± 0.3	98.2 ± 0.7	98.0 ± 0.4	0.69	98.1 ± 0.3	98.1 ± 0.3	98.1 ± 0.2
13	97.7 ± 3.7	2.45	98.8 ± 0.3	98.9 ± 0.3	98.8 ± 0.4	98.9 ± 0.6	2.29	98.9 ± 0.2	99.0 ± 0.2	99.1 ± 0.2
14	96.3 ± 11.3	4.95	96.5 ± 0.3	96.9 ± 0.6	96.9 ± 0.7	99.4 ± 0.3	3.37	96.5 ± 0.3	97.0 ± 0.3	97.0 ± 0.4
15	78.7 ± 29.9	20.51	97.5 ± 3.2	98.1 ± 0.7	98.0 ± 0.5	98.6 ± 0.4	1.27	98.1 ± 0.3	98.2 ± 0.3	98.1 ± 0.4
16	90.4 ± 28.8	10.22	96.5 ± 4.0	97.5 ± 1.2	97.0 ± 1.7	99.7 ± 0.4	3.67	98.1 ± 0.5	97.9 ± 0.3	98.5 ± 0.4
17	95.4 ± 5.4	4.06	97.1 ± 0.1	97.2 ± 0.2	97.0 ± 1.0	98.6 ± 5.4	3.00	97.1 ± 0.1	97.2 ± 0.2	97.4 ± 0.2
18	80.0 ± 86.3	18.83	97.9 ± 0.3	97.6 ± 0.8	97.0 ± 0.9	98.0 ± 0.6	0.80	97.9 ± 0.2	97.8 ± 0.2	97.9 ± 0.2
19	97.2 ± 0.4	1.07	94.8 ± 0.2	94.8 ± 0.2	93.7 ± 6.1	97.3 ± 0.2	1.02	94.8 ± 0.2	94.8 ± 0.2	95.1 ± 0.3
20	97.0 ± 0.3	0.47	99.2 ± 0.4	99.8 ± 0.7	98.8 ± 1.2	97.1 ± 0.2	0.40	99.2 ± 0.4	99.5 ± 0.5	99.6 ± 0.9
M±	93.8 ±	5.6 ±	97.8 ±	97.8 ±	97.2 ±	98.4 ±	1.9 ±	97.9 ±	98.0 ±	98.0 ±
SD	15.7	6.6	0.7	1.1	2.5	0.7	1.1	0.3	0.3	0.4

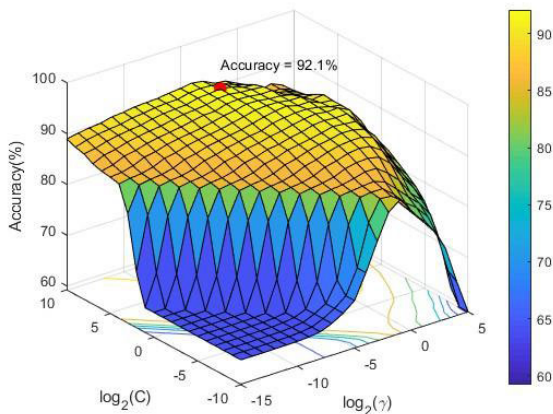
accuracy is $89.5 \pm 3.87\%$, sensitivity is $90.6 \pm 4.67\%$, specificity is $79.3 \pm 16.39\%$, and precision is $91.7 \pm 0.85\%$. The data under 7 km/h has the highest accuracy, 93.2%, and the data under 5 km/h has the lowest accuracy, 85.5%. Table 9 shows the testing performances of VGG-19 model at three speeds. The data under 3 km/h has the highest accuracy, 90.2%, and the data under 7 km/h has the lowest accuracy, 81.7%. The average accuracy is $86.2 \pm 4.28\%$, sensitivity is $95.9 \pm 1.08\%$, specificity is $67.0 \pm 6.95\%$, and precision is $81.4 \pm 13.51\%$. The overall assessment is that the SVM has the better performance than VGG-19 model.

C. IMPROVEMENT OF SpO₂ CALCULATION

Table 10 shows the change of SpO₂ measured by the forehead pulse oximeter at static condition, 3 km/h, 5 km/h, and 7 km/h before and after the SQI classification with the SVM using Gaussian radial basis kernel function. We only used the PPG segments with the high SQI to calculate the SpO₂ after the SQI classification. Because subjects were measured by the forehead and finger-type pulse oximeters only in the static experiment, there was the Error %. We could find that the Error % reduces from 5.6 ± 6.6 to 1.9 ± 1.1 after the SQI classification. Figure 10(a) shows the Bland-Altman plot



(a)

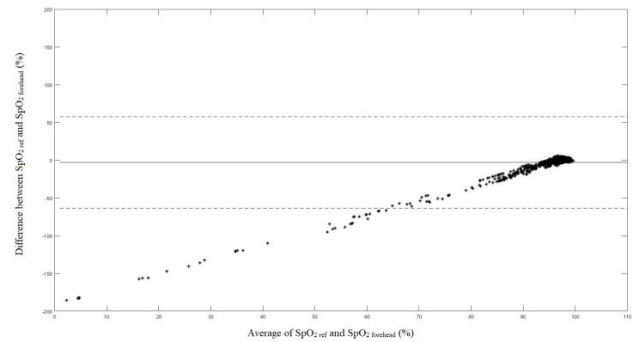


(b)

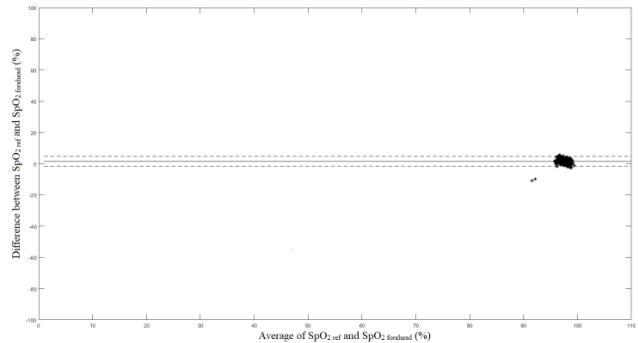
FIGURE 9. (a) The distribution of SVM performance with the linear kernel function using the parameter, C. the highest accuracy is 91.2%. (b) The distribution of SVM performance with the Gaussian radial basis kernel function using the parameters, C and γ . The highest accuracy is 92.1%.

displaying the difference between the reference SpO₂ and forehead SpO₂ before the SQI classification. PPG signals in some segments were seriously corrupted. Thus, these SpO₂ would be very large negative values. Figure 10(b) shows the Bland-Altman plot displaying the difference between the reference SpO₂ and forehead SpO₂ after the SQI classification.

The mean and standard deviation of the differences were $-2.9 \pm 15.8\%$ and $1.6 \pm 1.5\%$ before and after SQI classification, respectively. There are 50 points and 46 points before and after the classification out of the 95% limits of agreement. In the dynamic condition, the standard deviation of SpO₂ measured by forehead pulse oximeter at the three speeds all reduce from 0.7, 1.1 and 2.5 to 0.3, 0.3 and 0.4 after the SQI classification. Box plot is used to display the distribution of SpO₂ before and after the SQI classification for the three speeds, as shown in Figure 11. We used the t-test to analyze the SpO₂ values whether they have the significant difference before and after the SQI classification for the different speeds. In the cases of 3 km/h and 7 km/h, SpO₂ values all had the significant differences. But, in the case of 5 km/h, SpO₂ values did not show the significant difference. Thus, accuracy and stability of SpO₂ could be improved by the SQI classification.



(a)



(b)

FIGURE 10. (a) Bland-Altman plot displaying the difference between the reference SpO₂ and forehead SpO₂ before the SQI classification. (b) Bland-Altman plot displaying the difference between the reference SpO₂ and forehead SpO₂ after the SQI classification.

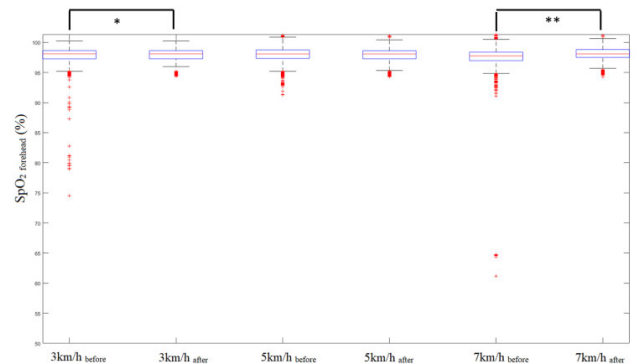


FIGURE 11. Box plot displaying the distribution of SpO₂ before and after the SQI classification for the three speeds, 3 km/h, 5 km/h and 7 km/h. (* represents $p < 0.05$, ** represents $p < 0.001$).

VII. DISCUSSION

PPG signal has been widely used to measure many physiological parameters, like as pulse rate [30], blood oxygen saturation [1], blood pressure [31], respiration rate [32], and left ventricular ejection time (LVET) [33]. But, the information of PPG wave used by these physiological parameters is different. The heart rate measurement used the PPI of PPG signal. Thus, the accuracy of heart rate would not be reduced when the PPG wave has the clear main peak even if the PPG signal is corrupted by motion artifact. The SQI classification for the heart rate measurement would not contribute higher accuracy when the morphology of PPG pulse is seriously

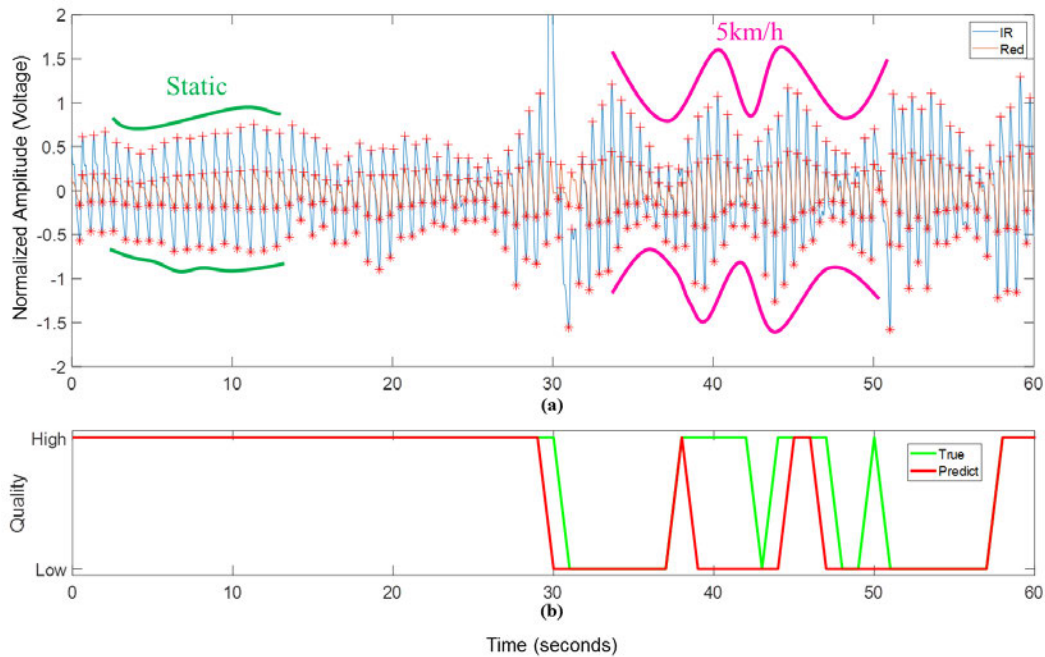


FIGURE 12. (a) The PPG pulse signals for the red light (red line) and IR light (blue line) under the static condition and 5 km/h condition. (b) The annotated result (green line) and classified result (red line).

distorted [34], [35]. Because the LVET is defined in the DPPG signal as the time interval between the first zero crossing point and the minimum point, when the PPG wave has a slight distortion at the rise time duration, the accuracy of LVET would be reduced. Thus, the SQI decision has to be peak by peak. Liu *et al.* compared the performances of the SVM and CNN approaches [17], [18]. They found that the CNN approach could detect slighter distortion for the PPG wave than the SVM approach. In the SpO₂ measurement, there are two PPG signals, red light and IR light, which baseline and pulse components all are used to calculate SpO₂ value. Thus, the SQI decision of PPG signals has to consider the four signals, PPG pulse signal and PPG baseline signal for the red and IR lights. Sukor *et al.* used decision-tree classifier to classify the SQI of PPG signals with peak by peak which sensitivity, specificity and accuracy were $89 \pm 10\%$, $77 \pm 19\%$ and $83 \pm 11\%$ [36]. Chong *et al.* used SVM to classify the SQI with the PPG segment which accuracy was 94.4% for the laboratory control [16]. Subjects only were requested to move their body, and walk or climb stairs. Thus, the density of motion artifact could be controlled in their study.

In our study, subjects were running on the treadmill for the different speeds. The intensity and density of the motion artifact could not be controlled. The accuracies of the SVM and VGG-19 model were $89.9 \pm 1.18\%$ and $88.7 \pm 1.54\%$ when the classified data included the static and dynamic data. Even if we used the static data to estimate the SQI of dynamic data, the accuracies of the SVM and VGG-19 model also had $89.5 \pm 3.87\%$ and $86.2 \pm 4.28\%$. It appears that our results were not as competitive as Chong's study. however, our

dataset had a large amount of data with low SQI. In Table 2, the amount of data with the low SQI increases from 602, 904 to 2066 when the running speed increases from 3 km/h, 5 km/h to 7 km/h. In Table 8, the accuracy for the 7 km/h data is 93.2% which is the best than other speeds, because the PPG signals were seriously corrupted by the motion artifact. The SVM could easily classify the PPG segments with high and low SQI. While in the case of 5 km/h, the accuracy is the worst because the PPG signals were moderately corrupted by the motion artifact. Thus, SpO₂ values before and after SQI classification did not have the significant difference in this speed. Figure 12(a) shows the PPG pulse signal during the static and 5 km/h conditions. Figure 12(b) shows the annotated result (green line) and classified result (red line). In the static duration, the accuracy is 100%. But, in the 5 km/h duration, both changes of PWA and SD of pulses are large.

Classifying the SQI of PPG segments appertains to a generalized decision. The annotations of the PPG segments would directly affect the accuracy of classification. In this study, we used SpO₂ value, the characteristics of PPG pulse, and the characteristic change to annotate the SQI. The features of PPG segment were the standard deviation of these characteristics. Thus, the characteristics of PPG pulses with tiny error would make this PPG segment annotated incorrectly. Figure 13(a) shows a PPG segment which SQI is annotated as low but it is classified as high in the static experiment. The first row shows the SpO₂ values measured by the Rossmax® SA310, the second row shows the SpO₂ values measured by the forehead pulse oximeter. In this PPG segment, the first to fourth pulses are annotated as low quality because the RTs of these pulses are over the desired range.

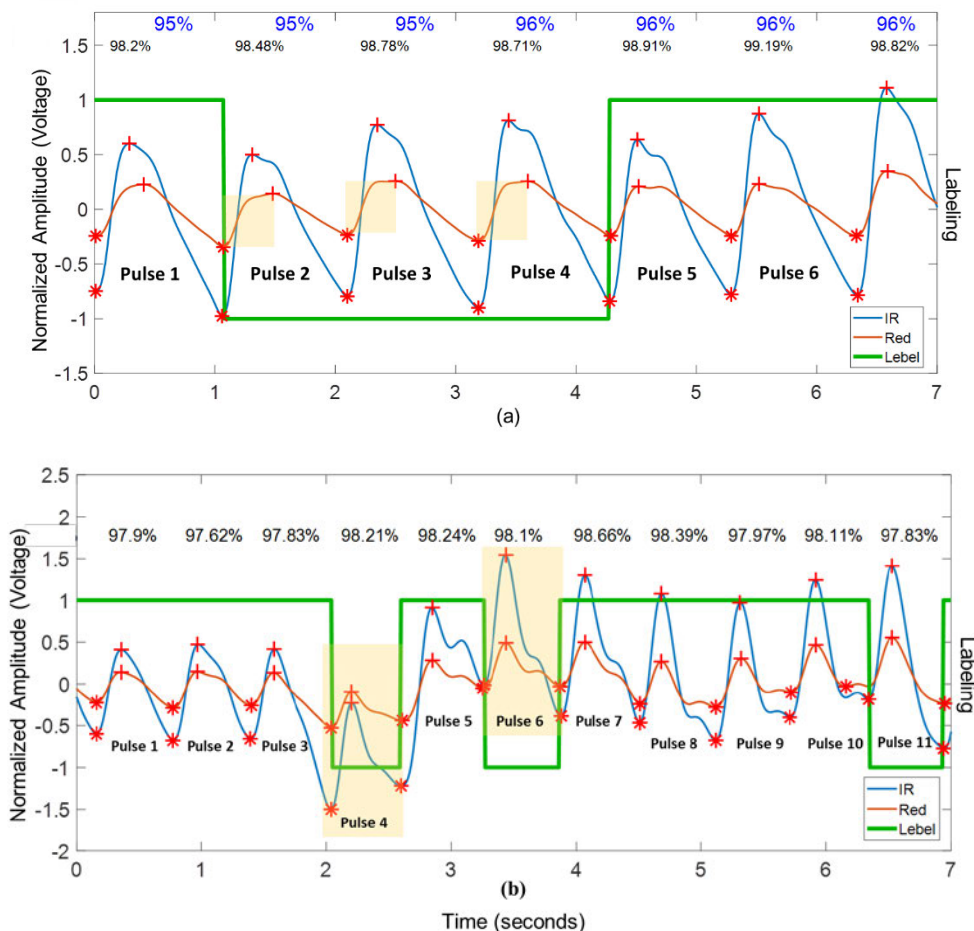


FIGURE 13. Two examples for the PPG segments with the wrong annotation because characteristics of PPG pulses with tiny error.

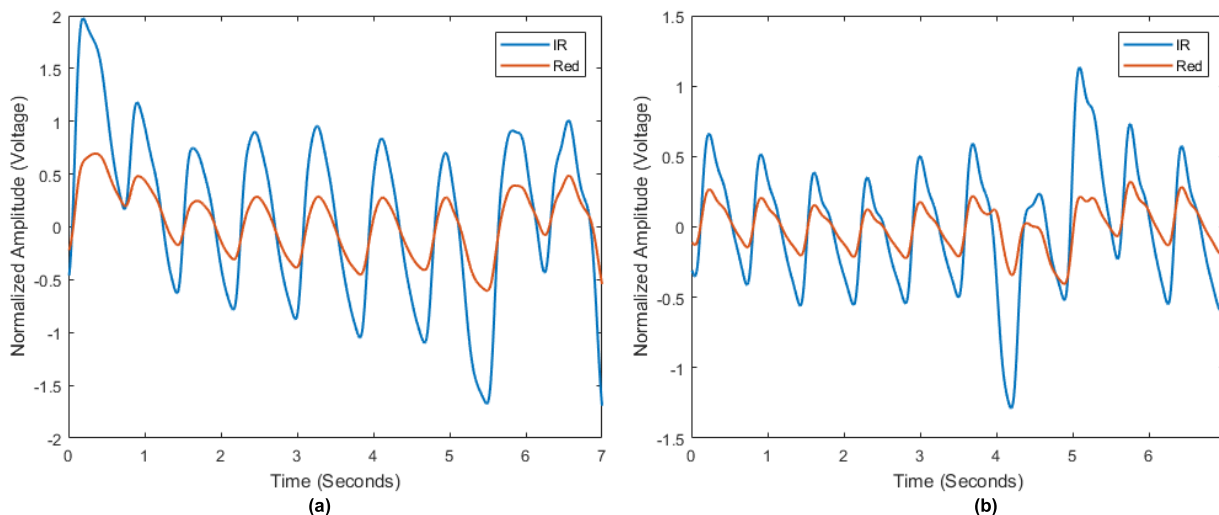


FIGURE 14. Two PPG segments are classified as high SQI by SVM, but VGG-19 model classified them as low SQI.

But the PPG pulse signals are very stable. Figure 13(b) shows the PPG pulse signals in the dynamic experiment. Those SpO₂ values were measured by the forehead pulse oximeter.

The fourth, sixth and eleventh pulses are annotated as low quality according to “The absolute peak or valley value for red light is larger than the IR light”. Thus, this PPG segment

is annotated as low SQI. But, the standard deviations of PWA and PWD are small. This segment was classified as high SQI.

In this study, the performance of VGG-19 model was not better than the SVM. This result had the opposite view to the previous studies of Liu *et al.* [17], [18]. The reason was considered that the SQI classification of PPG segment pertains to a generalized decision. The VGG-19 model could detect the tiny difference of the patterns. Figure 14(a) shows that a PPG segment is classified as high SQI by SVM, but as low SQI by VGG-19. We could find that the PPG pulse signals have some baseline drift in the beginning of PPG segment. Figure 14(b) also shows the same problem. The PPG pulse signals have sort of baseline drift between 4 and 5 seconds. This PPG segment was classified by VGG-19 model as low SQI.

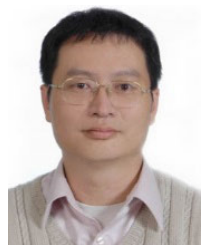
VIII. CONCLUSION

In this study, a forehead pulse oximeter was implemented which could measure SpO₂ in static and dynamic scenarios. The array-type sensor was able to find the best measurement position. Because the forehead pulse oximeter would be used in dynamic scenarios, the SQIs of PPG signals have to be evaluated to increase the accuracy of SpO₂. The SVM and VGG-19 model were used to classify the SQI. Although the performance of SVM was slightly better than VGG-19 model, advantage of using the VGG-19 model is that we did not need to explore the characteristics of PPG pulses in preprocessing step. Noticeably, the main limitation of the study was that the number of the PPG segments was not large enough. If more PPG segments are used in the training process of the VGG-19 model, better results will be expected. Moreover, we did not compare the performance of wrist pulse oximeter with the forehead pulse oximeter in the dynamic scenario. In the future, this issue is worthy to explore.

REFERENCES

- [1] T. Tamura, "Current progress of photoplethysmography and SPO₂ for health monitoring," *Biomed. Eng. Lett.*, vol. 9, no. 1, pp. 21–36, Feb. 2019.
- [2] J. Wannenburg and R. Malekian, "Body sensor network for mobile health monitoring, a diagnosis and anticipating system," *IEEE Sensors J.*, vol. 15, no. 12, pp. 6839–6852, Dec. 2015.
- [3] A. Bideaux, T. Schwendemann, and S. Hey, "Evaluation of design parameters for a reflection based long-term pulse oximetry sensor," in *Proc. 5th EAI Int. Conf. Wireless Mobile Commun. Healthcare, Transforming Healthcare Through Innov. Mobile Wireless Technol.*, 2015, pp. 190–193.
- [4] S. Fouzas, K. N. Priftis, and M. B. Anthracopoulos, "Pulse oximetry in pediatric practice," *Pediatrics*, vol. 128, no. 4, pp. 740–752, Oct. 2011.
- [5] L. J. Mengelkoch, D. Martin, and J. Lawler, "A review of the principles of pulse oximetry and accuracy of pulse oximeter estimates during exercise," *Phys. Therapy*, vol. 74, no. 1, pp. 40–49, Jan. 1994.
- [6] J. H. Williams, S. K. Powers, and M. K. Stuart, "Hemoglobin desaturation in highly trained athletes during heavy exercise," *Med. Sci. Sports Exerc.*, vol. 18, p. 168–173, Apr. 1986.
- [7] S. Sugino, N. Kanaya, M. Mizuuchi, M. Nakayama, and A. Namiki, "Forehead is as sensitive as finger pulse oximetry during general anesthesia," *Can. J. Anesthesia/J. canadien d'Anesthésie*, vol. 51, no. 5, pp. 432–436, May 2004.
- [8] V. Blaylock, M. Brinkman, S. Carver, P. McLain, S. Matteson, P. Newland, R. Pettit, C. Schulman, W. vanSchijndel, and S. Watson, "Comparison of finger and forehead oximetry sensors in postanesthesia care patients," *J. PeriAnesthesia Nursing*, vol. 23, no. 6, pp. 379–386, Dec. 2008.
- [9] M. T. Petterson, V. L. Begnoche, and J. M. Graybeal, "The effect of motion on pulse oximetry and its clinical significance," *Anesthetic Analgesia*, vol. 105, no. 6, pp. 78–84, 2007.
- [10] J. Allen and A. Murray, "Similarity in bilateral photoplethysmographic peripheral pulse wave characteristics at the ears, thumbs and toes," *Physiol. Meas.*, vol. 21, no. 3, pp. 369–377, Aug. 2000.
- [11] J. Allen, "Photoplethysmography and its application in clinical physiological measurement," *Physiol. Meas.*, vol. 28, no. 3, pp. R1–R39, Mar. 2007.
- [12] S.-H. Liu, K.-M. Chang, and T.-H. Fu, "Heart rate extraction from photoplethysmogram on fuzzy logic discriminator," *Eng. Appl. Artif. Intell.*, vol. 23, no. 6, pp. 968–977, Sep. 2010.
- [13] C. Fischer, B. Dömer, T. Wibmer, and T. Penzel, "An algorithm for real-time pulse waveform segmentation and artifact detection in photoplethysmograms," *IEEE J. Biomed. Health Informat.*, vol. 21, no. 2, pp. 372–381, Mar. 2017.
- [14] K. Li, S. Warren, and B. Natarajan, "Onboard tagging for real-time quality assessment of photoplethysmograms acquired by a wireless reflectance pulse oximeter," *IEEE Trans. Biomed. Circuits Syst.*, vol. 6, no. 1, pp. 54–63, Feb. 2012.
- [15] T. Pereira, K. Gadhomi, M. Ma, X. Liu, R. Xiao, R. A. Colorado, K. J. Keenan, K. Meisel, and X. Hu, "A supervised approach to robust photoplethysmography quality assessment," *IEEE J. Biomed. Health Informat.*, vol. 24, no. 3, pp. 649–657, Mar. 2020.
- [16] J. W. Chong, D. K. Dao, S. M. A. Salehizadeh, D. D. Mcmanus, C. E. Darling, K. H. Chon, and Y. Mendelson, "Photoplethysmograph signal reconstruction based on a novel hybrid motion artifact detection-reduction Approach. Part I: Motion and noise artifact detection," *Ann. Biomed. Eng.*, vol. 42, no. 11, pp. 2238–2250, Nov. 2014.
- [17] S.-H. Liu, J.-J. Wang, W. Chen, K.-L. Pan, and C.-H. Su, "Classification of photoplethysmographic signal quality with fuzzy neural network for improvement of stroke volume measurement," *Appl. Sci.*, vol. 10, no. 4, p. 1476, Feb. 2020.
- [18] S.-H. Liu, R.-X. Li, J.-J. Wang, W. Chen, and C.-H. Su, "Classification of photoplethysmographic signal quality with deep convolution neural networks for accurate measurement of cardiac stroke volume," *Appl. Sci.*, vol. 10, no. 13, p. 4612, Jul. 2020.
- [19] Maxim Integrated Products, Inc. *Recommended Configurations and Operating Profiles for MAX30101/MAX30102 EV Kits*. Accessed: 2018. [Online]. Available: <https://www.maximintegrated.com/en/design/technical-documents/userguides-and-manuals/6/6409.html>
- [20] J. M. Cho, Y. K. Sung, K. W. Shin, D. J. Jung, Y. S. Kim, and N. H. Kim, "A preliminary study on photoplethysmogram (PPG) signal analysis for reduction of motion artifact in frequency domain," in *Proc. IEEE-EMBS Conf. Biomed. Eng. Sci.*, Dec. 2012, pp. 28–33.
- [21] M. Okada, S. Kimura, and M. Okada, "Estimation of arterial pulse wave velocities in the frequency domain: Method and clinical considerations," *Med. Biol. Eng. Comput.*, vol. 24, no. 3, pp. 255–260, May 1986.
- [22] A. Zhao, P. Wang, and Y. Duan, "Peak detection of pulse signal based on dynamic difference threshold," *J. Jilin Univ.*, no. 3, pp. 847–853, 2014.
- [23] A. M. Weissler, W. S. Harris, and C. D. Schoenfeld, "Systolic time intervals in heart failure in man," *Circulation*, vol. 37, no. 2, pp. 149–159, Feb. 1968.
- [24] R. P. Lewis, S. E. Rittogers, W. F. Froester, and H. Boudoulas, "A critical review of the systolic time intervals," *Circulation*, vol. 56, no. 2, pp. 146–158, 1977.
- [25] V. Vapnik, *Statistical Learning Theory*. New York, NY, USA: Wiley, 1998.
- [26] C. Cortes and V. Vapnik, "Support-vector network," *Mach. Learn.*, vol. 20, no. 3, pp. 273–297, 1995.
- [27] C. J. C. Burges, "A tutorial on support vector machines for pattern recognition," *Data Mining Knowl. Discovery*, vol. 2, pp. 955–971, Jun. 1998.
- [28] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," 2014, *arXiv:1409.1556*. [Online]. Available: <http://arxiv.org/abs/1409.1556>
- [29] *Keras Applications*. Accessed: May 14, 2020. [Online]. Available: <https://keras.io/applications/>
- [30] S. Majumder, T. Mondal, and M. J. Deen, "Wearable sensors for remote health monitoring," *Sensors*, vol. 17, pp. 130–174, Jan. 2017.
- [31] G. Sharwood-Smith, J. Bruce, and G. Drummond, "Assessment of pulse transit time to indicate cardiovascular changes during obstetric spinal anaesthesia," *Br. J. Anaesth.*, vol. 96, pp. 100–105, Jan. 2006.
- [32] K. H. Chon, S. Dash, and K. Ju, "Estimation of respiratory rate from photoplethysmogram data using time-frequency spectral estimation," *IEEE Trans. Biomed. Eng.*, vol. 56, no. 8, pp. 2054–2063, Aug. 2009.

- [33] S.-H. Liu, J.-J. Wang, C.-H. Su, and D.-C. Cheng, "Improvement of left ventricular ejection time measurement in the impedance cardiography combined with the reflection photoplethysmography," *Sensors*, vol. 18, no. 9, p. 3036, Sep. 2018.
- [34] A. Johnson, J. Behar, F. Andreotti, G. D. Clifford, and J. Oster, "Multi-modal heart beat detection using signal quality indices," *Physiol. Meas.*, vol. 36, no. 8, p. 1665, 2015.
- [35] G. B. Papini, P. Fonseca, L. M. Eerikäinen, S. Overeem, J. W. M. Bergmans, and R. Vullings, "Sinus or not: A new beat detection algorithm based on a pulse morphology quality index to extract normal sinus rhythm beats from wrist-worn photoplethysmography recordings," *Physiol. Meas.*, vol. 39, no. 11, Nov. 2018, Art. no. 115007.
- [36] J. A. Sukor, S. J. Redmond, and N. H. Lovell, "Signal quality measures for pulse oximetry through waveform morphology analysis," *Physiol. Meas.*, vol. 32, no. 3, pp. 369–384, Mar. 2011.



SHING-HONG LIU (Member, IEEE) received the B.S. degree in electronic engineering from Feng Jia University, Taichung City, Taiwan, in 1990, the M.S. degree in biomedical engineering from National Cheng Kung University, Tainan, Taiwan, in 1992, and the Ph.D. degree from the Department of Electrical and Control Engineering, National Chiao-Tung University, Hsinchu, Taiwan, in 2002. Since 1994, he has been a Lecturer with the Department of Biomedical Engineering, Yuanpei University, Hsinchu. He was an Associate Professor from 2002 to 2008. He is currently a Professor with the Department of Computer Science and Information Engineering, Chaoyang University of Technology. His current research interests include digital signal processing, fuzzy control, and designing biomedical instrument.



HAN-CHI LIU was born in Taichung City, Taiwan, in 1996. He received the B.S. degree in electrical engineering from the National Chin-Yi University of Technology, Taichung City, in 2018, and the M.S. degree in electrical engineering from the National Taipei University of Technology, Taipei City, Taiwan, in 2020. His research interests include digital signal processing, microcontroller design, and biomedical instrument design.



WENXI CHEN (Member, IEEE) received the B.S. and M.S. degrees in biomedical electronics from Zhejiang University, Hangzhou, China, in 1983 and 1986, respectively, and the Ph.D. degree in biomedical instrumentation from the Institute of Biomaterials and Bioengineering, Tokyo Medical and Dental University, Japan, in 2001. He is currently a Professor with the Biomedical Information Engineering Laboratory (BIE Lab), and the Head of the Information Systems Division, The University of Aizu, Aizuwakamatsu, Japan. His current research interests include two main aspects, development of various modalities to monitor diverse vital signs under daily life environment, and comprehensive interpretation of multifarious long-term data over multiple time scales to reveal statistical links and causalities among diseases and health condition.



TAN-HSU TAN (Senior Member, IEEE) received the B.S. degree in electrical engineering from the National Taiwan Institute of Technology, in 1983, the M.S. degree in electrical engineering from National Tsing Hua University, in 1988, and the Ph.D. degree in electronics engineering from National Chiao Tung University, Hsinchu, Taiwan, in 1998. Since 1988, he has been with the Department of Electrical Engineering, National Taipei University of Technology, Taipei, Taiwan, where he is currently a Professor. His research interests include wireless communications, telecare, machine learning, and optimization algorithms.

...