IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received August 7, 2019, accepted September 6, 2019, date of publication September 23, 2019,
date of current version October 21, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2943197

Dual-Input Neural Network Integrating Feature Extraction and Deep Learning for Coronary Artery Disease Detection Using Electrocardiogram and Phonocardiogram

HAN LI[1], XINPEI WANG[1], CHANGCHUN LIU[1], YAN WANG[2], PENG LI[1,3], (Member, IEEE), HONG TANG[4], LIANKE YAO[1], AND HUAN ZHANG[1]

[1]School of Control Science and Engineering, Shandong University, Jinan 250061, China
[2]Health Management Center, Affiliated Hospital of Qingdao University, Qingdao 266000, China
[3]Division of Sleep and Circadian Disorders, Brigham and Womens Hospital, Harvard Medical School, Boston, MA 02115, USA
[4]Department of Biomedical Engineering, Dalian University of Technology, Dalian 116024, China

Corresponding authors:Xinpei Wang (wangxinpei@sdu.edu.cn) and Changchun Liu (g595018526@163.com)

This work was supported in part by the National Natural Science Foundation of China under Grant 61471223, Grant 61601263, and Grant 61501280, and in part by the Shandong Provincial Science and Technology Project under Grant 2016GSF201037.

ABSTRACT Electrocardiogram (ECG) and phonocardiogram (PCG) signals reflect the electrical and mechanical activities of the heart, respectively. Although studies have documented that some abnormalities in ECG and PCG signals are associated with coronary artery disease (CAD), only few researches have combined the two signals for automatic CAD detection. This paper aims to differentiate between CAD and non-CAD groups using simultaneously collected ECG and PCG signals. To entirely exploit the underlying information in these signals, a novel dual-input neural network that integrates the feature extraction and deep learning methods is developed. First, the ECG and PCG features are extracted from multiple domains, and the information gain ratio is used to select important features. On the other hand, the ECG signal and the decomposed PCG signal (at four scales) are concatenated as a five-channel signal. Then, the selected features and the five-channel signal are fed into the proposed network composed of a fully connected model and a deep learning model. The results show that the classification performance of either feature extraction or deep learning is insufficient when using only ECG or PCG signal, and combining the two signals improves the performance. Further, when using the proposed network, the best result is obtained with accuracy, sensitivity, specificity, and G-mean of 95.62%, 98.48%, 89.17%, and 93.69%, respectively. Comparisons with existing studies demonstrate that the proposed network can effectively capture the combined information of ECG and PCG signals for the recognition of CAD.

INDEX TERMS Deep learning, feature extraction, ECG, PCG, coronary artery disease, classification.

I. INTRODUCTION

Coronary artery disease (CAD) is a major type of cardiovascular diseases and a leading cause of death worldwide. It is primarily caused by the accumulation of plaques along the inner walls of coronary arteries, which reduces the blood flow to the myocardium [1], [2]. Under severe conditions, the ruptured plaques can completely occlude the arterial lumen, eventually triggering an acute myocardial infarction. At present, coronary angiography, the gold standard in the clinical diagnosis of CAD, is an invasive technique that requires professional surgical procedures, considerable time, and cost. Thus, it is not attractive as a screening method for general medical conditions. Electrocardiogram (ECG) and phonocardiogram (PCG) signals have valuable information about the electrical and mechanical activities of the heart, respectively. Studies have documented that in the resting ECG signals of CAD patients, symptoms such as T-wave inversion, ST-T abnormalities, left ventricular hypertrophy, and premature ventricular contractions can be observed [3]. In the PCG signals of CAD patients, turbulent flow in narrowed coronary arteries may produce weak murmurs in diastolic heart sounds [4]. Thus, both ECG and PCG are

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

promising tools for CAD screening, with advantages of simple operation, high efficiency, non-invasiveness, and low cost.

With the development of signal processing and machine learning techniques, the automated analysis of ECG and PCG signals has been increasingly investigated for the diagnosis of cardiac diseases. A computer-aided system can capture important information that may be overlooked by the subjective interpretation of the physicians. Conventional methods focus on feature extraction and classification process, and commonly used features are extracted from time [5], frequency [6], nonlinear [7]–[9], and time-frequency [10]–[15] domains. When useful features are obtained, the next step is classification. Numerous classifiers have been employed thus far, for example, k-nearest neighbours [11], Gaussian mixture model [12], artificial neural network [16], [17], support vector machine (SVM) [18], [19], and decision tree [8], [9]. In recent years, deep learning techniques based on the convolutional neural network (CNN) have become very popular. Owing to its strong feature learning capabilities, CNN has gained research interest in its application to the classification of ECG and PCG signals [20]–[24]. CNN is a hierarchical neural network whose convolutional layers alternate with downsampling layers [25]. It primarily imitates the human visual system and can effectively recognize the patterns in visual objects. Compared with conventional methods, CNN can detect hidden patterns from physiological signals without any feature extraction and selection processes.

Researchers have attempted using various ECG or PCG databases for the automatic detection of CAD. Among the studies that use the ECG signal, a majority of the focus is on conventional feature extraction and classification process [8], [12], [15], [16], [18]. Moreover, Acharya *et al.* [20] and Tan *et al.* [21] adopted the CNN model as a tool for feature learning and attained improved performance using the ECG signals of 47 subjects. The previously mentioned studies have reported the effectiveness of ECG for diagnosis of CAD, however, they all use very small datasets (less than or equal to 47 subjects). Considering significant morphological differences in physiological signals among different people, a small amount of data may lack diversity and universality, which may easily lead to insufficient learning ability and poor generalization of the classifier model. Hence, several studies attempted using comparatively larger databases [26]–[28], and among them, Lee *et al.* [26] achieved the highest classification accuracy of 90% using the ECG signals of 193 subjects. The studies using the PCG signal are based on traditional methods, and the obtained results are not satisfactory [29]–[31]. Of these studies, Gauthier *et al.* [31] achieved the highest accuracy of 73.3% using the PCG signals of 30 subjects. In fact, researches have pointed out that using either the ECG or PCG signals to detect CAD proves insufficient because signal abnormalities may be absent in some patients [32], [33].

Additionally, current studies use either conventional methods or CNN technique to differentiate CAD from the non-CAD group. Nevertheless, the deep learning features extracted from the signals focus on temporal and morphological information, which may be inadequate to diagnose the disease as features from other representational domains have also been found to be useful [26], [27], [30]. If multi-domain feature extraction can be combined with deep learning, sufficient information can be provided for the classification task, and better results are likely to be achieved. To the best of our knowledge, there is no literature integrating the two techniques for the identification of CAD at present.

Consider the complementarity between the ECG and PCG signals in CAD detection, this paper uses simultaneously collected ECG and PCG signals from 195 subjects. The combined utilization of ECG and PCG signals can overcome the disadvantages of using only one of the signals and thus provide more evidence for an accurate diagnosis of CAD. To classify the CAD and non-CAD classes, a novel dual-input neural network architecture is developed using an ensemble of feature extraction and deep learning. One input of the network comprises the extracted and selected multi-domain features from ECG (time, frequency, and time-frequency domains) and PCG (time, frequency, entropy, energy, and kurtosis domains) signals. The other input of the network is a five-channel concatenation of the ECG signal and the PCG signal decomposed into four scales. Inside the dual-input neural network, a fully connected model and a deep learning model are used to process the two inputs, and their outputs are consolidated to predict the class label. The proposed network takes advantages of both the traditional method and the deep learning technique. Through the mutual complement between the hand-crafted features and the deep learning features, more underlying patterns within the signals can be recognized to enhance the classification performance. The main contributions of this work are summarized below:

- Combined utilization of simultaneously collected ECG and PCG signals.
- Development of a novel dual-input neural network that integrates conventional feature extraction and deep learning techniques.
- Extraction of multi-domain features from ECG and PCG signals and design of a five-channel signal for the proposed dual-input network.
- Achieving superior classification performance compared with existing studies.

The remainder of this paper is organized as follows. Section II separately describes the data acquisition, signal preprocessing, feature extraction, feature selection, proposed dual-input neural network, and performance evaluation methods. The feature selection result and classification result are presented in Section III, and Section IV discusses the results. Finally, Section V concludes this paper and gives some future extensions.

II. MATERIALS AND METHODS

A. DATA ACQUISITION

The study received full approval from the Institutional Review Board of Shandong Provincial Qianfoshan

TABLE 1. Basic characteristics of CAD patients and non-CAD subjects.

Characteristic	CAD	Non-CAD
Age	62 ± 10	56 ± 7
Male/female	89/46	26/34
Height	166 ± 8	166 ± 8
Weight	71 ± 11	64 ± 9
Body mass index	26 ± 4	23 ± 3
Heart rate	67 ± 8	68 ± 7
Systolic blood pressure	133 ± 16	116 ± 17
Diastolic blood pressure	82 ± 12	72 ± 10

Values are expressed as male/female or mean $\pm$ standard deviation.

Hospital, Jinan, China, and it was carried out under the principles in the Declaration of Helsinki and its following amendments. Before participation, all subjects provided informed consent, and the patients who had undergone percutaneous coronary intervention or coronary artery bypass surgery were excluded. A cardiovascular function detection device (CVFD-II, Huiyironggong Technology Co., Ltd, Jinan, China) was used to record the physiological signals. During the measurement, the subjects were requested to lie in a supine position in a quiet and temperature-controlled room (25 ± 3°C). The standard lead-II ECG and PCG signals were simultaneously recorded for 5 min at a sampling rate of 1 kHz. The inclusion criteria were subjects who underwent coronary angiography.

Subjects with at least one major coronary artery stenosis $\geq$ 50% were categorized into the CAD class and the others into the non-CAD class. A total of 195 subjects enrolled in the data collection, including 135 CAD patients and 60 non-CAD subjects (115 males and 80 females; age range: 42 $\sim$ 86 years). The basic characteristics of all subjects are given in Table 1.

B. SIGNAL PREPROCESSING

First, the ECG signals are denoised using a Butterworth filter with a pass-band of 1 $\sim$ 60 Hz, and the baseline wander is removed. For the PCG signals, a high-pass Butterworth filter (10 Hz) is applied to remove the low-frequency noise and the baseline wander. Then, the power interference (50 Hz) in the ECG and PCG signals is removed. To enlarge the sample size for machine learning, each five-min recording is cropped to 20 signals lasting 15 s. Consequently, a total of 3,900 samples are generated, including 2,700 CAD and 1,200 non-CAD samples. Each signal is regularized with z-score normalization before inputting it into the network.

C. FEATURE EXTRACTION

A total of 81 ECG and 154 PCG features are extracted from multiple domains.

1) ECG FEATURE EXTRACTION

First, the R-peaks of the ECG signals are detected using Afonso's algorithm [34]. Based on the detected R-peaks, the other characteristic points including P-peaks, Q-points, S-points, and T-peaks are identified successively. Next, the ECG features are extracted from the time, frequency, and time-frequency domains. The details of the features are described below.

TABLE 2. Extracted time-domain features during a cardiac cycle of the ECG signal.

Abbreviation	Description
RR_interval	RR interval duration
DRR_interval	The difference of successive RR intervals
PR_interval	The interval duration between P-peak and R-peak
RT_interval	The interval duration between R-peak and T-peak
QS_interval	The interval duration between Q-point and S-point
QT_interval	The interval duration between Q-point and T-peak
R_peak	The amplitude of R-peak
PQ_peak	The amplitude ratio of P-peak to Q-point
PT_peak	The amplitude ratio of P-peak to T-peak
TP_peak	The amplitude ratio of T-peak to P-peak
TQ_peak	The amplitude ratio of T-peak to Q-point
R_signal	The average amplitude of the ECG signal in a 0.7-s window around R-peak that contains the P-QRS-T wave
SP_signal	The average amplitude of the ECG signal between S-point and P-peak that contains the ST-T segment
ST_bias	The depression or elevation degree of ST segment [35]

(1) Time domain (62 features). The time interval and amplitude features likely to be associated with CAD are extracted from each cardiac cycle and listed in Table 2. Then, the maximum, minimum, average, and standard deviation of these features are calculated over all cycles of a signal segment, and the heart rate is derived. Further, the maximum (max_All_signal), minimum (min_All_signal), average (avg_All_signal), variance (var_All_signal), and standard deviation (std_All_signal) of the amplitudes over the entire signal are calculated.

(2) Frequency domain (5 features). Discrete Fourier transform is used to obtain the frequency spectrum f_i ($i = 1, 2, \cdots, N$) of the ECG signals. The average and standard deviation of the spectrum are computed and denoted as *avg_Sp* and *std_Sp*, respectively. On this basis, three- and four-order statistics, i.e., skewness and kurtosis, are calculated separately, which measure the symmetry and "peakedness" degree of the spectrum distribution. Further, the entropy of the spectrum is computed. The three features are expressed as

$$Sp\_skew = \frac{\sum_{i=1}^{N}(f_i - avg\_Sp)^3}{N * std\_Sp^3}, \tag{1}$$

$$Sp\_kur = \frac{\sum_{i=1}^{N}(f_i - avg\_Sp)^4}{N * std\_Sp^4} - 3, \tag{2}$$

$$Sp\_entropy = -\sum_{i=1}^{N} \frac{f_i}{\sum_{j=1}^{N} f_j} * \log \frac{f_i}{\sum_{j=1}^{N} f_j}. \tag{3}$$

(3) Time-frequency domain (14 features). Wavelet and wavelet packet features are extracted from the ECG

signals, respectively. First, in view of the reported correlation between the fragmented QRS complex and CAD [36], the discrete wavelet transform is employed for the time-frequency analysis of the QRS complex. The signal in a 0.3-s window around the R-peak is decomposed into six scales using the db2 mother wavelet. Then, the energy of the approximation coefficients (Ea) at the fourth level is computed as $E = \sum_{i=1}^{L} Ea(i)^2$, where L is the length of Ea. Similarly, the energy of the detail coefficients at six scales is computed and expressed as Eb_k ($k = 1 \sim 6$). On this basis, five wavelet features are obtained and shown below, which represent the energy ratio of the coefficients in different frequency bands. Their average and standard deviation over all cycles are computed.

$$Wd\_ratio1 = \frac{Eb_6}{\sum_{k=1}^{6} Eb_k}, \tag{4}$$

$$Wd\_ratio2 = \frac{Eb_4 + Eb_5}{\sum_{k=1}^{6} Eb_k}, \tag{5}$$

$$Wd\_ratio3 = \frac{Eb_4}{E}, \tag{6}$$

$$Wd\_ratio4 = \frac{Eb_5}{E}, \tag{7}$$

$$Wd\_ratio5 = \frac{Eb_6}{E}. \tag{8}$$

Wavelet packet analysis is an extension of the wavelet transform and offers a rich signal analysis [37]. Using the db4 mother wavelet, the entire ECG signal is decomposed into four scales, and the coefficients of 16 nodes at the fourth level of the wavelet packet tree are reconstructed. Then, the energy of each reconstructed vector is computed as E_k ($k = 1 \sim 16$), and the total energy, wavelet energy entropy, and energy ratios are derived as

$$Wp\_energy = \sum_{k=1}^{16} E_k, \tag{9}$$

$$Wp\_entropy = -\sum_{i=1}^{16} \frac{E_i}{\sum_{k=1}^{16} E_k} * \log \frac{E_i}{\sum_{k=1}^{16} E_k}, \tag{10}$$

$$Wp\_ratio1 = \frac{E_1}{Wp\_energy}, \tag{11}$$

$$Wp\_ratio2 = \frac{E_2}{Wp\_energy}. \tag{12}$$

2) PCG FEATURE EXTRACTION

During each cardiac cycle, the PCG signal is segmented into four states: S1, systole, S2, and diastole, using the algorithm proposed by Springer *et al.* [38]. Then, referring to the literature on the recognition of abnormal PCG signals [39], [40], the time, frequency, entropy, energy, and kurtosis features are extracted from each cardiac cycle, as described below. In addition to the spectrum features, the mean and standard deviation of the other features are computed over all cycles.

TABLE 3. Extracted time-domain features during a cardiac cycle of the PCG signal.

Abbreviation	Description
CC	The cardiac cycle duration
IntS1	The S1 interval duration
IntS2	The S2 interval duration
IntSys	The systole interval duration
IntDia	The diastole interval duration
Ratio_SysCC	The ratio of systole interval to the cardiac cycle duration
Ratio_DiaCC	The ratio of diastole interval to the cardiac cycle duration
Ratio_SysDia	The ratio of systole interval to the diastole interval
Amp_SysS1	The ratio of average amplitude during systole to that during S1
Amp_DiaS2	The ratio of average amplitude during diastole to that during S2

TABLE 4. Extracted frequency-domain features during a cardiac cycle of the PCG signal.

Abbreviation	Description
Spectrum_S1 (20 $\sim$ 130 Hz)	The spectrum value of S1 at frequency of 20 Hz, 30 Hz, $\cdots$, 130 Hz respectively
Spectrum_S2 (20 $\sim$ 130 Hz)	The spectrum value of S2 at frequency of 20 Hz, 30 Hz, $\cdots$, 130 Hz respectively
Spectrum_Sys (20 $\sim$ 300 Hz)	The spectrum value of systole at frequency of 20 Hz, 30 Hz, $\cdots$, 300 Hz respectively
Spectrum_Dia (20 $\sim$ 300 Hz)	The spectrum value of diastole at frequency of 20 Hz, 30 Hz, $\cdots$, 300 Hz respectively
HF_S1	The proportion of high-frequency component in the total spectrum of S1
HF_Sys	The proportion of high-frequency component in the total spectrum of systole
HF_S2	The proportion of high-frequency component in the total spectrum of S2
HF_Dia	The proportion of high-frequency component in the total spectrum of diastole
LF_S1	The proportion of low-frequency component in the total spectrum of S1
LF_Sys	The proportion of low-frequency component in the total spectrum of systole
LF_S2	The proportion of low-frequency component in the total spectrum of S2
LF_Dia	The proportion of low-frequency component in the total spectrum of diastole

(1) Time domain (20 features). The interval durations, duration ratios, and average amplitude ratios are calculated, and the details are given in Table 3.

(2) Frequency domain (98 features). Using the discrete Fourier transform, the spectrum and spectrum ratio features are extracted and listed in Table 4. First, the spectrum values of S1, systole, S2, and diastole states are computed at different frequencies with a 10 Hz interval. Their mean values are then calculated over all cycles. Moreover, the proportions of the high-frequency (above 200 Hz) and the low-frequency (below 50 Hz) components in the spectrums of the four states are obtained separately.

(3) Energy domain (20 features). The energy ratio is investigated between two segmented states, and it is defined as

$$Energy\_ratio = \frac{\sum_{i=1}^{M} s_i^2}{\sum_{i=1}^{N} c_i^2}, \tag{13}$$

TABLE 5. Extracted energy-domain features during a cardiac cycle of the PCG signal.

Abbreviation	Description
Energy_S1ToSys	The energy ratio of S1 to systole
Energy_S1ToDia	The energy ratio of S1 to diastole
Energy_S2ToSys	The energy ratio of S2 to systole
Energy_S2ToDia	The energy ratio of S2 to diastole
Energy_DiaToSys	The energy ratio of diastole to systole
Energy_S1Total	The energy ratio of S1 to the total cardiac cycle
Energy_SysTotal	The energy ratio of systole to the total cardiac cycle
Energy_S2Total	The energy ratio of S2 to the total cardiac cycle
Energy_DiaTotal	The energy ratio of diastole to the total cardiac cycle
Energy_HsTotal	The energy ratio of S1 and S2 to the total cardiac cycle

TABLE 6. Extracted entropy- and kurtosis-domain features during a cardiac cycle of the PCG signal.

Abbreviation	Description
SE_Sys	The sample entropy of systole
rFE_Sys	The refined fuzzy entropy of systole
SE_Dia	The sample entropy of diastole
rFE_Dia	The refined fuzzy entropy of diastole
S1_kur	The kurtosis of S1
Sys_kur	The kurtosis of systole
S2_kur	The kurtosis of S2
Dia_kur	The kurtosis of diastole

where M and N denote the lengths of the time series s_i and c_i, respectively. The details are described in Table 5.

(4) Entropy domain (8 features). Entropy measures the complexity of the time series from the viewpoint of cardiac dynamics. In addition to the commonly used sample entropy [41], a refined fuzzy entropy [42] is applied to the systole and diastole states, which replaces the conventional Gaussian function with the piecewise fuzzy membership function. The details are given in Table 6.

(5) Kurtosis domain (8 features). The kurtosis of the S1, systole, S2, and diastole states are computed and the details are described in Table 6.

D. FEATURE SELECTION

Before feeding the extracted features into the proposed network, IGR [43] is used for feature selection, which helps reduce feature dimension and enhance classification performance. IGR overcomes the drawback of the information gain when dealing with attributes with a large number of distinct values. A feature with a larger IGR value is considered to make a higher contribution to the classification. The computation of IGR is based on the information gain and the split information, as shown below

$$Entropy(S) = \sum_{i \in Classes} -p_i * \log p_i, \tag{14}$$

$$Gain(S, A) = Entropy(S) - \sum_{v \in Values(A)} \frac{|S_v|}{|S|} * Entropy(S_v), \tag{15}$$

$$SplitInfo(S, A) = -\sum_{v \in Values(A)} \frac{|S_v|}{|S|} * \log \frac{|S_v|}{|S|}, \tag{16}$$

$$IGR(S, A) = \frac{Gain(S, A)}{SplitInfo(S, A)}, \tag{17}$$

where p_i is the probability of records belonging in class i, and S_v denotes the collection of records in class label S whose value is v for a feature A.

In this study, a large number of spectrum features at different frequencies are extracted from the PCG signals. In fact, there is a considerable redundancy and correlation among them, and it is difficult to obtain a good performance by directly inputting these values into the learning model. Consequently, in each state of S1, systole, S2, and diastole, only one feature with the largest IGR is selected and a total of four mean spectrum value features are retained. Then, these four features and the remaining 72 PCG features are combined with the 81 ECG features into one set of 157 features. Based on the IGR values calculated from the entire dataset, these features are ranked in a descending order. The number of highly ranked features that should be selected depends on the performance of five-fold cross validation.

E. PROPOSED DUAL-INPUT NEURAL NETWORK

In this paper, we present a new dual-input neural network using an ensemble of feature extraction and deep learning to classify the CAD and non-CAD classes. A graphical description of the network architecture is provided in Fig. 1.

The network consists of a fully connected model and a deep learning model. It has two inputs: the selected multi-domain ECG and PCG features, and a five-channel signal obtained from the ECG and the decomposed PCG signals. A two-layer fully connected model with 32 and 64 neurons is used to classify the selected features, which provides a nonlinear transformation function $f(\cdot)$ that maps the input x to the output y. The transformation is expressed as

$$y = f(wx + b), \tag{18}$$

where w and b represent the weight matrix and bias vector separately. On the other hand, for the PCG signal, the frequencies of S1, S2, and heart murmurs (in systole and diastole) are about 50 $\sim$ 100, 50 $\sim$ 200, and 120 $\sim$ 660 Hz, respectively. In the case of coronary stenosis, there may be an increase in the high-frequency components [4]. Owing to a wide frequency distribution, the PCG signal is decomposed into four scales using the db6 mother wavelet for the time-frequency analysis. Then, the ECG signal and the PCG signal decomposed into four scales are concatenated into a five-channel signal as the input for the deep learning model.

The deep learning model consists of two components: the CNN and a bidirectional gated recurrent unit (GRU) with the attention mechanism. The detailed configurations are listed in Table 7. A leaky rectifier linear unit [45] is used as the activation function for all layers. The CNN is used as an automatic feature extractor, and it comprises ten identical convolutional blocks, each of which has three end-to-end

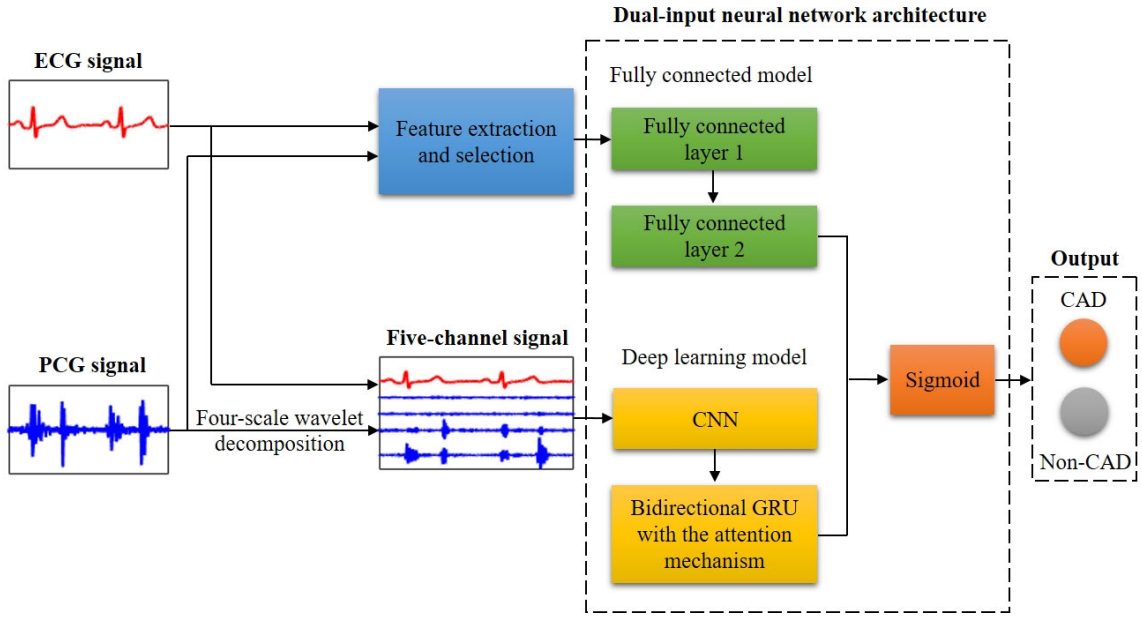


FIGURE 1. Architecture of the dual-input neural network composed of a fully connected model and a deep learning model.

TABLE 7. Configurations of the deep learning model composed of the CNN and a bidirectional GRU with the attention mechanism.

Layers	Type	Output shape	Kernel size	Stride	Dropout
1	Convolutional layer	15000×32	16	1	–
2	Convolutional layer	15000×32	16	1	–
3	Convolutional layer	7500×32	64	2	0.2
⋮					
31	Bidirectional GRU	15×24	–	–	0.2
32	Attention mechanism	24	–	–	–

To avoid repetition, only the configurations of three layers in the first convolutional block are presented. The structure of the remaining convolutional blocks is the same as that of the first one.

convolutional layers. The convolutional layers use convolution instead of multiplication in the fully connected layers, which is defined as

$$y_j = f(\sum_{i \in M} k_{ij} * x_i + b_j) \tag{19}$$

where M denotes the filter size, i is the size of the input feature maps, j is the size of convolution kernels, and k_{ij} is the convolution kernel for the i-th input and the j-th output. The output of the convolutional layers is referred to as feature maps. As seen in Table 7, in each convolutional block, the strides of the first two layers are set as the default parameter 1, and the stride of the third layer is set as 2. The increase in the stride can down-sample the feature maps to minimize overfitting and improve learning speed. Further, the dropout with a probability of 0.2 is used to reduce the generalization error.

A bidirectional GRU with the attention mechanism follows the CNN to process the feature sequences learned by the convolutional layers. The GRU is a variant of the recurrent neural network, suitable for processing the time series [46]. It solves the vanishing gradient problem and preserves long-distance memory. There are two gates, a reset gate r_t and an update gate z_t in the GRU structure, which together control the update of information. The bidirectional GRU merges the representations of a forward and a backward GRU to identify more patterns. At time t, the hidden state $\overrightarrow{h}_t$ of the forward GRU is computed as

$$\overrightarrow{h}_t = (1 - \overrightarrow{z}_t)\overrightarrow{h}_{t-1} + \overrightarrow{z}_t \overrightarrow{\tilde{h}}_t, \tag{20}$$

$$\overrightarrow{z}_t = \sigma(\overrightarrow{W}_z x_t + \overrightarrow{U}_z \overrightarrow{h}_{t-1}), \tag{21}$$

$$\overrightarrow{\tilde{h}}_t = \tanh\left(\overrightarrow{W} x_t + \overrightarrow{U}(\overrightarrow{r}_t \odot \overrightarrow{h}_{t-1})\right), \tag{22}$$

$$\overrightarrow{r}_t = \sigma(\overrightarrow{W}_r x_t + \overrightarrow{U}_r \overrightarrow{h}_{t-1}), \tag{23}$$

where x_t is the sequence vector, $\overrightarrow{h}_{t-1}$ is the previous state, $\overrightarrow{\tilde{h}}_t$ is the candidate state, σ is a logistic sigmoid function, and $\odot$ denotes an element-wise multiplication. Similarly, the hidden state $\overleftarrow{h}_t$ of the backward GRU is computed, and

finally, the hidden state h_t of the bidirectional GRU is

$$h_t = \begin{bmatrix} \overrightarrow{h}_t \\ \overleftarrow{h}_t \end{bmatrix}. \tag{24}$$

The attention mechanism [47] is then added to assign different importance weights to each element of the bidirectional GRU output. The mechanism emphasizes the important elements that can distinguish CAD from non-CAD class, and thus, it helps boost the classification performance. The weight α_t is derived through a softmax function, and the output vector of the attention mechanism is computed by $c = \sum_t \alpha_t h_t$.

At the end of the network, the outputs of the fully connected model and the deep learning model are consolidated, and a sigmoid layer transforms the real values into prediction probability, which is expressed as

$$\sigma_i = \omega x_i + b, \tag{25}$$

$$p_i = \frac{1}{1 + \exp(-\sigma_i)}, \tag{26}$$

where x_i is the i-th output and p_i denotes the output of the nonlinear activation function.

F. CROSS VALIDATION AND PERFORMANCE EVALUATION

Five-fold cross validation is performed in this work. First, the 5 min ECG and PCG recordings of all subjects are divided into five parts by stratified sampling. Subsequently, in each part, every 5 min recording is cut into twenty 15 s segments. Four out of the five parts are used as the training set, and the rest is used as the validation set. This way, the signal segments for the training and validation phases are assured to come from totally different subjects, thereby making the evaluation more realistic. Then, five iterations are conducted, and the final classification result is the average of five cross validations.

In order to evaluate the classification performance, the standard metrics, including accuracy (Acc), sensitivity (Sen), and specificity (Spe), are used in this study. Furthermore, considering the class imbalance in the data, G-mean is used as well, which measures the balanced performance of a learning model between the CAD and non-CAD classes. The equations associated with these metrics are calculated as

$$Acc = \frac{TP + TN}{TP + TN + FP + FN}, \tag{27}$$

$$Sen = \frac{TP}{TP + FN}, \tag{28}$$

$$Spe = \frac{TN}{TN + FP}, \tag{29}$$

$$G - mean = \sqrt{Sen * Spe}, \tag{30}$$

where TP, TN, FP, and FN stand for the number of the true positives, true negatives, false positives, and false negatives, respectively.

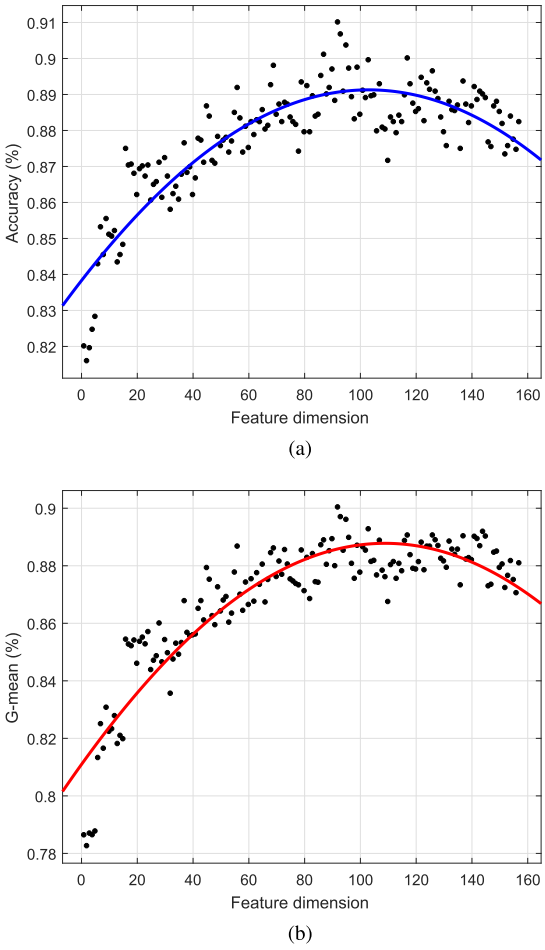


FIGURE 2. Variation in the accuracy and G-mean with an increase in feature dimension: (a) Accuracy; (b) G-mean. (The scatters indicate the true values, and the curve is the fitting of the scatters.)

III. RESULTS

In our experiment, feature extraction was implemented in Matlab R2016b, and machine learning was performed using Python 3.5. The deep learning codes were executed on Keras framework with a Tensorflow backend using an NVIDIA GeForce GTX 1080Ti GPU. The proposed method was performed on a PC with 3.70 GHz Intel Core i7 CPU, 16 GB RAM and a windows 10 operating system. The results of the feature selection, feature classification, deep learning classification, and dual-input network classification are analyzed in the following subsections.

A. RESULTS OF FEATURE SELECTION AND CLASSIFICATION

1) RESULTS OF FEATURE SELECTION

The variations in the overall metrics accuracy and G-mean caused by an increase in the dimension of the input features are shown in Fig. 2, and the fully connected model is used for cross validation. The scatters indicate the true values of the metrics. To clearly observe the change in the trend, a Gaussian function is used for curve fitting the scatters.

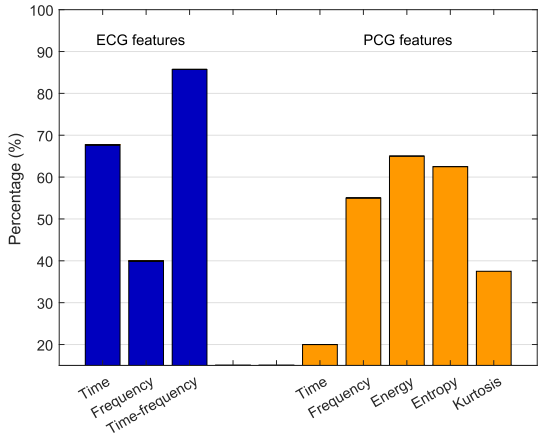


FIGURE 3. Contributions of ECG and PCG features from different domains.

The fitting curves indicate that, with an increase in feature dimension, both the accuracy and G-mean increase first and then decrease, and they reach the maximum value when the feature dimension is 92. Consequently, 92 features from the ECG and PCG signals are selected and used for the classification.

To further determine the feature domains that contribute to the classification, the ratio between the number of selected features and original features is computed for each domain, as shown in Fig. 3. The ECG features from the time-frequency domain contribute the most, followed by the time domain ECG features. Furthermore, the energy, entropy, and frequency domain PCG features also play a positive role in the classification. In contrast with these features, the contribution of the time domain PCG features is relatively small.

According to the IGR values, the top 22 features are listed in Table 8 to illustrate the features that are most useful in the detection of CAD. Among them, there are 11 ECG and 11 PCG features. The wavelet packet and wavelet features of the ECG signals rank in the top five, indicating their importance in the classification. Then, the amplitude ratio, energy, and spectrum features of the PCG signals rank from the sixth to fifteenth, and the energy features account for a majority of them. In addition to the above features, the amplitude ECG features also show their significance in the identification of CAD.

2) CLASSIFICATION PERFORMANCE BEFORE AND AFTER FEATURE SELECTION

Table 9 shows the classification results of the extracted ECG and PCG features and the selected feature set, as well as the significance test results (p-value) between their accuracies. In the classification of the CAD and non-CAD groups, the extracted ECG features perform better than the PCG features, and they achieve an accuracy of 88.41%. Then, by simply combining all ECG and PCG features, the accuracy is improved by 1.87% compared to that of the ECG features. After feature selection, the classification performance is further improved, and an accuracy of 91.0% is obtained.

TABLE 8. The IGR ranking of the selected ECG and PCG features. (max/min/avg/std_abbreviation denotes the maximum/minimum/average/standard deviation of the ECG features, respectively. m/std_abbreviation denotes the the mean/standard deviation of the PCG features.)

No.	Feature	IGR
1	**Wp_entropy**	0.2904
2	**Wp_ratio1**	0.2894
3	**Wp_ratio2**	0.2889
4	**avg_Wd_ratio1**	0.2441
5	**avg_Wd_ratio2**	0.2416
6	m_Amp_SysS1	0.2411
7	sd_Energy_HsTotal	0.2408
8	m_Energy_HsTotal	0.2401
9	m_Energy_S1ToSys	0.2239
10	m_Energy_SysTotal	0.2115
11	sd_Energy_SysTotal	0.2093
12	sd_Amp_SysS1	0.1975
13	sd_Energy_S1Total	0.1961
14	m_Energy_DiaTotal	0.1948
15	m_Spectrum_Sys (40Hz)	0.1934
16	**max_All_signal**	0.1889
17	**min_R_signal**	0.1873
18	m_Energy_S1ToDia	0.1822
19	**avg_R_signal**	0.1804
20	**max_TP_peak**	0.1784
21	**std_Wd_ratio1**	0.1764
22	**std_Wd_ratio2**	0.1764

The bold text indicates the ECG features.

B. CLASSIFICATION RESULTS OF DEEP LEARNING MODEL AND DUAL-INPUT NEURAL NETWORK

First, the ECG, PCG, decomposed PCG, and five-channel signals are fed into the deep learning model, respectively. Among them, the ECG and PCG are both one-channel signals; the PCG signal decomposed into four scales is concatenated into a four-channel signal; the five-channel signal is the concatenation of the ECG signal and the decomposed PCG signal. Then, the selected features and the five-channel signal are fed into the proposed neural network. Table 10 presents the classification results of the deep learning model and the proposed network, as well as the significance test results (p-value) between their accuracies.

In the results of the deep learning model, although the classification performance of the ECG signal is better than that of the PCG signal, the overall results are not satisfactory. By decomposing the PCG signal into four frequency bands, the performance is better than that of the original PCG signal. After the ECG signal and the decomposed PCG signal are concatenated into a five-channel signal, there is an improvement in the performance. On this basis, the classification performance is further boosted by the proposed network using both the selected features and the five-channel signal as inputs. Simultaneously, the accuracy reaches 95.62%, and the sensitivity is as high as 98.48%, which exceeds the performance of either the fully connected model or the deep learning model.

TABLE 9. Classification results of the extracted and selected features using the fully connected model.

Features	Acc (%)	Sen (%)	Spe (%)	G-mean (%)	p-value
ECG features (81)	88.41 ± 4.69	89.63 ± 5.86	85.67 ± 4.02	87.58 ± 4.18	0.36
PCG features (154)	80.23 ± 2.15	85.19 ± 2.84	69.08 ± 6.81	76.58 ± 3.44	0.00
ECG & PCG features (235)	90.28 ± 3.25	91.74 ± 3.76	87.00 ± 4.04	89.31 ± 3.21	0.69
Selected features (92)	91.00 ± 2.23	92.48 ± 3.01	87.67 ± 2.03	90.03 ± 1.88	1.00

[1] The feature dimension is indicated in parenthesis.
[2] The p-value indicates the significance test result between the accuracies of the extracted features and selected features.

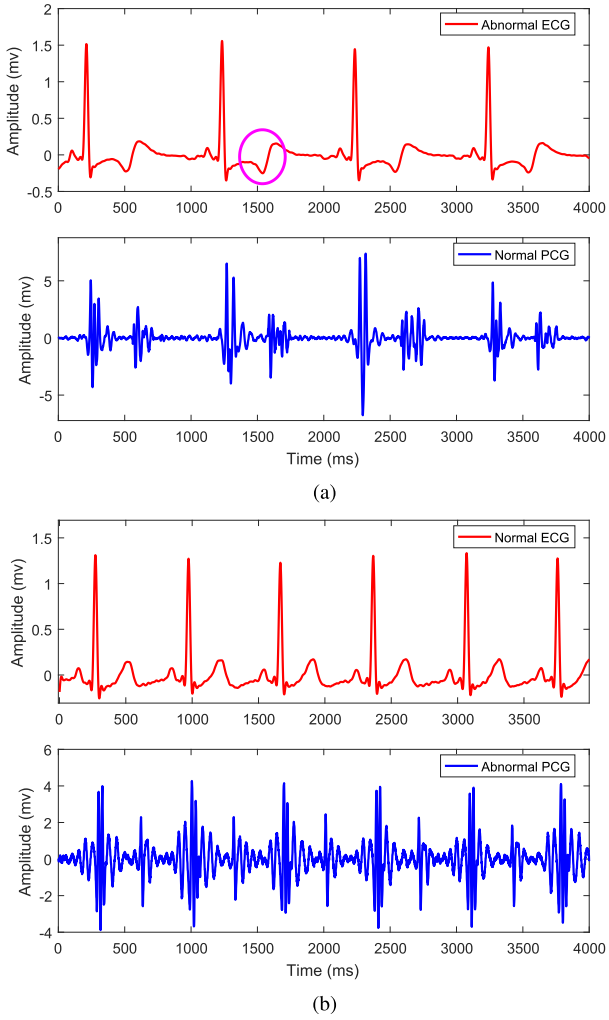


FIGURE 4. Illustration of the ECG and PCG signals of two typical CAD patients: (a) A CAD patient with abnormal ECG and normal PCG; (b) A CAD patient with normal ECG and abnormal PCG.

IV. DISCUSSION

Clinically, it is very difficult to identify CAD using only ECG or PCG signal. Fig. 4 illustrates the ECG and PCG signals of two CAD patients with considerably typical symptoms. As seen in Fig. 4 (a), the ECG signal of the patient shows the abnormality of the T-wave inversion. However, there are no obvious abnormal heart murmurs in the PCG signal. The symptoms of the patient in Fig. 4 (b) are the opposite. The ECG signal appears normal, while there is a significant increase in the heart murmurs of the PCG signal. In general, the clinical conditions of the CAD patients are far more complicated than those described above. In some cases, the variations in the ECG and PCG signals caused by the coronary stenosis are very subtle and inconspicuous. Thus, computer-aided diagnosis using combined ECG and PCG is essential, which provides a more reliable basis for accurately detecting CAD.

Tables 9 and 10 indicate that the classification performance of feature extraction or deep learning when using the combination of the two signals exceeds that when using only ECG or PCG signal. The results imply that the simultaneous use of ECG and PCG signals is more promising for assisting clinical CAD diagnosis. To the best of our knowledge, no similar studies have been applied to the diagnosis of CAD so far. One research on myocardial infraction detection used the combination of ECG, PCG, and clinical data [48]. The results of that study are consistent with those in this study: neither clinical data nor ECG nor PCG alone were sufficient for detecting the disease, and using multimodal features could improve the performance.

In practice, to capture more hidden information from the signals, as many features as possible are extracted. However, in the classification stage, not all these features are useful and there is some redundancy among them. Hence, feature selection is necessary to reduce the feature dimension. As seen in Fig. 3, among the selected ECG features, the wavelet packet and wavelet features from the time-frequency domain are the most important for the classification. Similarly, several studies used the wavelet decomposition [11], [49] and wavelet packet decomposition [50] for the ECG signals, and the extracted features also performed well in identifying CAD. Among the extracted PCG features, the energy, entropy, and frequency domain features indicate their importance in classification. The result is consistent with the study of Tang *et al.* [40], who used multi-domain features for the detection of abnormal PCG recordings.

To better understand the feature learning process of the CNN, the first seven-channel outputs of ten convolutional blocks are analyzed visually when using the ECG signal as the input. As depicted in Fig. 5, the feature maps of the convolutional blocks $1 \sim 6$ appear to approximate the patterns associated with the original ECG signal. Nevertheless, as the convolutional layer deepens to blocks $7 \sim 10$, the feature maps become increasingly abstract and less visually

TABLE 10. Classification results of the deep learning model and the dual-input neural network.

Model	Input	Acc (%)	Sen (%)	Spe (%)	G-mean (%)	*p*-value
Deep learning	ECG signal	84.59 ± 6.45	88.67 ± 4.70	75.42 ± 11.32	81.63 ± 8.21	0.01
Deep learning	PCG signal	71.95 ± 7.16	77.30 ± 11.59	59.92 ± 11.77	67.28 ± 6.54	0.00
Deep learning	Decomposed PCG signal	81.44 ± 3.35	90.15 ± 4.78	61.83 ± 12.50	74.13 ± 6.74	0.00
Deep learning	Five-channel signal	91.82 ± 1.68	93.15 ± 2.32	88.83 ± 7.56	90.84 ± 3.29	0.00
Proposed network	Selected features & five-channel signal	95.62 ± 0.99	98.48 ± 1.22	89.17 ± 3.11	93.69 ± 1.51	1.00

The *p*-value indicates the significance test result between the accuracies of the deep learning model and proposed network.

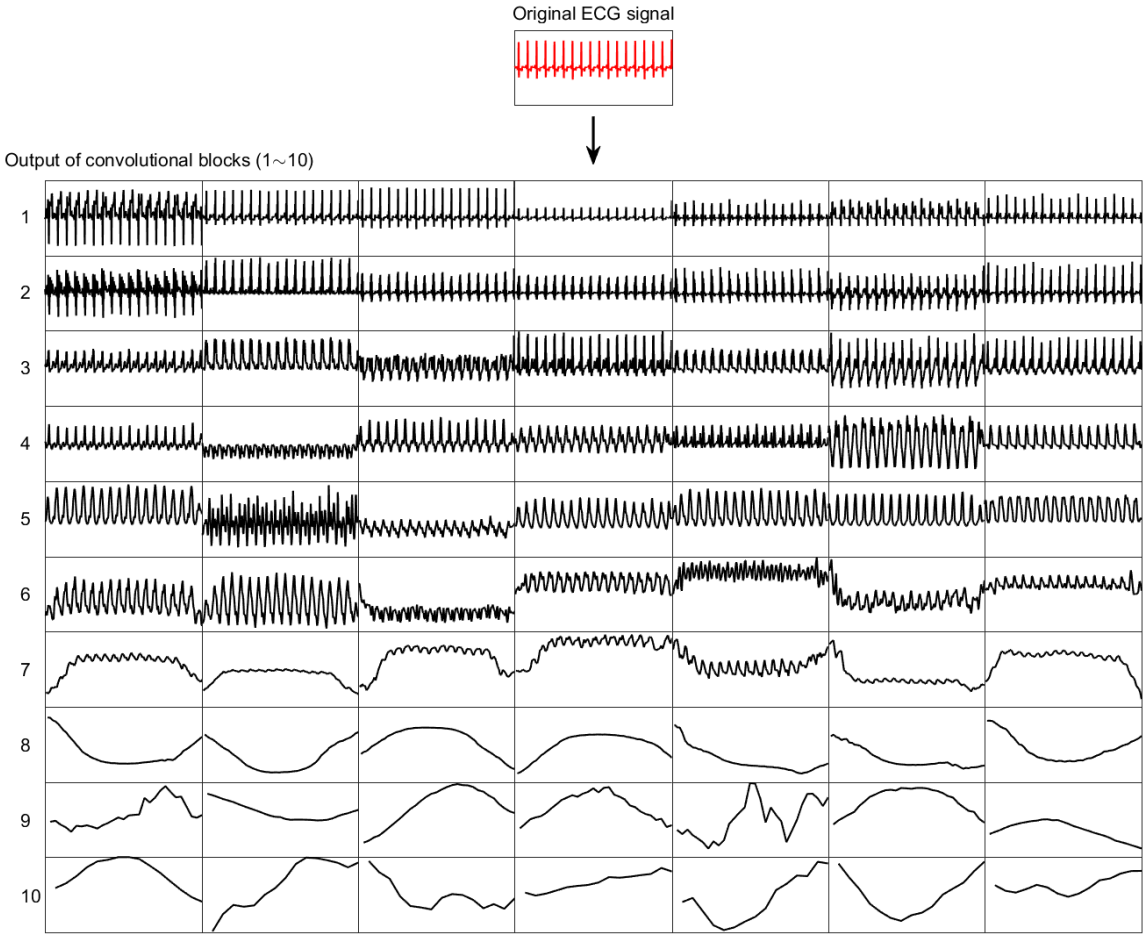


FIGURE 5. Visualization of feature learning process of the CNN using the ECG signal as the input.

interpretable. At this stage, the local patterns are learned and the outputs represent high-level concepts [51]. Higher-level representations carry less information about the visual contents of the original signal and more information related to the category.

Through the above visualization of feature maps, it is worth noting that the features learned by the CNN are primarily temporal patterns constituting structural elements that form discriminative representations for classification [24]. Although the CNN method proves to be effective in the classification of ECG or PCG signals [20]–[23], it may be challenging to learn all the features related to the disease, especially those features from frequency domain, time-frequency domain, entropy domain, and so on, when using the original signals as input. In this study, among the hand-crafted features, the time-frequency domain contributes the most for the ECG signal, and the energy and entropy domain contribute the most for the PCG signal. Furthermore, on comparing Tables 9 and 10, we find that the performance of the ECG or PCG features obtained by the fully connected model is better than that of the ECG or PCG signal obtained by the deep learning model. This implies that the CNN learns temporal and waveform information from the signals, however, it can hardly capture the information from other representational domains that play a very important role in the classification. Consequently, merging the feature extraction with deep learning can provide

TABLE 11. Summary of the existing studies on the diagnosis of CAD using ECG or PCG signals.

Author	Database	Feature & Classifier	Result (%)
Lee et al. [26] (2007)	ECG signal (193 subjects: 99 CAD & 94 normal)	Time, frequency, and nonlinear features SVM, Naïve Bayes, multiple association rules, decision tree	Acc=90.0 – –
Kaveh & Chung [27] (2013)	ECG signal (89 subjects: 43 CAD & 46 normal)	Time and frequency features SVM	Acc=88.0 Sen=84.2 Spe=92.6
Deng et al. [28] (2017)	ECG signal (421 subjects: 347 CAD & 74 normal)	Cardiodynamicsgram Priori threshold method	Acc=84.6 Sen=84.7 Spe=83.7
Makaryus et al. [29] (2013)	PCG signal (161 subjects: 19 CAD & 142 normal)	Microbruit score Logistic regression	Acc=61.5 Sen=89.5 Spe=57.7
Schmidt et al. [30] (2015)	PCG signal (133 subjects: 63 CAD & 70 normal)	Frequency and nonlinear features Quadratic discriminant function	Acc=68.4 Sen=72.0 Spe=65.2
This paper	ECG & PCG signals (195 subjects: 135 CAD & 60 non-CAD)	Time, frequency, and time-frequency features Energy, entropy, and kurtosis features Dual-input neural network	Acc=95.6 Sen=98.5 Spe=89.2

richer and more diverse information for distinguishing the CAD from non-CAD groups. In contrast with traditional methods or CNN, the proposed dual-input neural network has a higher representation ability, and thus, it can extract more underlying features from the signals, thereby allowing better identification of CAD patients.

Table 11 summarizes the existing studies that use a comparatively larger ECG or PCG database for CAD diagnosis. These studies are based on traditional feature extraction and classification methods. Among them, Lee *et al.* [26] used ECG features from the time, frequency, and nonlinear domains and achieved the highest accuracy of 90%. Although studies have reported the association between abnormal heart murmurs and CAD [29], [30], the classification performance of the PCG signal remains inferior to that of the ECG signal. Concerning this study, we used the combination of ECG and PCG signals from more subjects than in most other studies, and we extracted features from six different domains that could represent more potential patterns. A new network architecture was developed to classify the CAD and non-CAD classes, and both the obtained accuracy and sensitivity were found to be higher than those in the other studies. In clinical diagnosis, a high sensitivity means that the patients can be detected more accurately, which is of great practical significance.

In addition to the above studies, several investigations attempted single-channel CNN for the classification of ECG or PCG signals [20], [21], [52]. Similar to the study of Potes *et al.* [52], in this study, the PCG signal was decomposed into four frequency bands for the time-frequency analysis rather than using the original signal directly, which contributed to mining more patterns. However, the novelty of this study was that by concatenating the ECG and four sub-band PCG signals, we designed a five-channel signal as the input of CNN instead of using the single-channel input. Multi-channel input facilitated the simultaneous learning of the ECG and PCG features. In the feature learning phase, the CNN could extract not only the features of each channel signal but also the combined features between different channels. This helped to capture more useful information.

V. CONCLUSION

In this work, a dual-input neural network structure was developed, and the ECG and PCG signals were used in combination for CAD diagnosis. The proposed network integrated the traditional feature extraction and deep learning models to identify more underlying features inside the signals. The experimental results showed that the proposed method achieved a classification accuracy of 95.62% on the ECG and PCG recordings cropped to 15 s, outperforming the use of only the traditional method or deep learning technique. Further, these results exceeded those of existing studies on CAD detection using the ECG or the PCG signal. Our study suggests that the combined utilization of the ECG and PCG signals, as well as the proposed network, is potentially promising for noninvasive CAD detection.

However, it should be mentioned that the proposed method requires a large amount of annotated data for supervised learning, and some of the extracted features are of less use to the classification. Besides, after feature selection, some of remained features may be strongly correlated. Therefore, in future work, signals from more subjects are necessary to test the performance of the proposed method further. Moreover, we will pay attention to exploring more useful features and investigating whether removing highly relevant features can help improve classification performance.

ACKNOWLEDGMENT

The authors would like to thank the Shandong Provincial Qianfoshan Hospital for the full support and thank all the volunteers who participated in the study. The authors also would like to thank the undergraduate Guangyi Chen for his help in programming.

CONFLICT OF INTEREST STATEMENT

The authors declare that there are no conflicts of interest to this work.

REFERENCES

[1] *Global Health Estimates: Deaths by Cause, Age, Sex and Country, Geneva*, World Health Org., Geneva, Switzerland, 2015, pp. 2000–2012.

[2] M. Nikpay, A. Goel, H.-H. Won, L. M. Hall, C. Willenborg, S. Kanoni, D. Saleheen, T. Kyriakou, C. P. Nelson, and J. C. Hopewell, "A comprehensive 1000 Genomes–based genome-wide association meta-analysis of coronary artery disease," *Nature Genet.*, vol. 47, no. 10, pp. 1121–1130, Sep. 2015.

[3] D. De Bacquer, G. De Backer, M. Kornitzer, K. Myny, Z. Doyen, and H. Blackburn, "Prognostic value of ischemic electrocardiographic findings for cardiovascular mortality in men and women," *J. Amer. College Cardiol.*, vol. 32, no. 3, pp. 680–685, 1998.

[4] M. Akay, Y. M. Akay, D. Gauthier, R. G. Paden, W. Pavlicek, F. D. Fortuin, J. P. Sweeney, and R. W. Lee, "Dynamics of diastolic sounds caused by partially occluded coronary arteries," *IEEE Trans. Biomed. Eng.*, vol. 56, no. 2, pp. 513–517, Feb. 2009.

[5] S. Ari, K. Hembram, and G. Saha, "Detection of cardiac abnormality from PCG signal using LMS based least square SVM classifier," *Expert Syst. Appl.*, vol. 37, no. 12, pp. 8019–8026, 2010.

[6] Z. Zhidong, "Noninvasive diagnosis of coronary artery disease based on instantaneous frequency of diastolic murmurs and SVM," in *Proc. 27th Annu. Conf. IEEE Eng. Med. Biol.*, Jan. 2006, pp. 5651–5654.

[7] S. K. Nayak, K. Pal, B. Mohapatra, A. Dey, and A. Bit, "A review on the nonlinear dynamical system analysis of electrocardiogram signal," *J. Healthcare Eng.*, vol. 2018, no. 2, May 2018, Art. no. 6920420.

[8] U. R. Acharya, Y. Hagiwara, J. E. W. Koh, S. L. Oh, J. H. Tan, M. Adam, and R. S. Tan, "Entropies for automated detection of coronary artery disease using ECG signals: A review," *Biocybern. Biomed. Eng.*, vol. 38, no. 2, pp. 373–384, 2018.

[9] M. E. Karar, S. H. El-Khafif, and M. A. El-Brawany, "Automated diagnosis of heart sounds using rule-based classification tree," *J. Med. Syst.*, vol. 41, no. 4, p. 60, Mar. 2017.

[10] S. M. Debbal and F. Bereksi-Reguig, "Time-frequency analysis of the first and the second heartbeat sounds," *Appl. Math. Comput.*, vol. 184, no. 2, pp. 1041–1052, Jan. 2007.

[11] U. R. Acharya, H. Fujita, M. Adam, O. S. Lih, V. K. Sudarshan, T. J. Hong, J. E. Koh, Y. Hagiwara, C. K. Chua, C. K. Poo, and T. R. San, "Automated characterization and classification of coronary artery disease and myocardial infarction by decomposition of ECG signals: A comparative study," *Inf. Sci.*, vol. 377, pp. 17–29, Jan. 2017.

[12] D. Giri, U. R. Acharya, R. J. Martis, S. V. Sree, T.-C. Lim, T. Ahamed, and J. S. Suri, "Automated diagnosis of coronary artery disease affected patients using LDA, PCA, ICA and discrete wavelet transform," *Knowl.-Based Syst.*, vol. 37, pp. 274–282, Jan. 2013.

[13] L. Samantaray, M. Dash, and R. Panda, "A review on time-frequency, time-scale and scale-frequency domain signal analysis," *IETE J. Res.*, vol. 51, no. 4, pp. 287–293, 2005.

[14] M. Wacker and H. Witte, "Time-frequency techniques in biomedical signal analysis: A tutorial review of similarities and differences," *Methods Inf. Med.*, vol. 52, no. 4, pp. 279–296, 2013.

[15] S. Patidar, R. B. Pachori, and U. R. Acharya, "Automated diagnosis of coronary artery disease using tunable-Q wavelet transform applied on heart rate signals," *Knowl.-Based Syst.*, vol. 82, pp. 1–10, Jul. 2015.

[16] S. Dua, D. U. Xian, S. V. Sree, and A. V. I. Thajudin, "Novel classification of coronary artery disease using heart rate variability analysis," *J. Mech. Med. Biol.*, vol. 12, no. 4, 2012, Art. no. 1240017.

[17] M. Karimi, R. Amirfattahi, S. Sadri, and S. A. Marvasti, "Noninvasive detection and classification of coronary artery occlusions using wavelet analysis of heart sounds with neural networks," in *Proc. 3rd IEE Int. Seminar Med. Appl. Signal Process.*, Nov. 2005, pp. 117–120.

[18] A. D. Dolatabadi, S. E. Z. Khadem, and B. M. Asl, "Automated diagnosis of coronary artery disease (CAD) patients using optimized SVM," *Comput. Methods Programs Biomed.*, vol. 138, pp. 117–126, Jan. 2017.

[19] I. Babaoğlu, O. Findik, and M. Bayrak, "Effects of principle component analysis on assessment of coronary artery diseases using support vector machine," *Expert Syst. Appl.*, vol. 37, no. 3, pp. 2182–2185, 2010.

[20] U. R. Acharya, H. Fujita, O. S. Lih, M. Adam, J. H. Tan, and C. K. Chua, "Automated detection of coronary artery disease using different durations of ecg segments with convolutional neural network," *Knowl.-Based Syst.*, vol. 132, pp. 62–71, Sep. 2017.

[21] J. H. Tan, Y. Hagiwara, W. Pang, I. Lim, S. L. Oh, M. Adam, R. S. Tan, M. Chen, and U. R. Acharya, "Application of stacked convolutional and long short-term memory network for accurate identification of CAD ECG signals," *Comput. Biol. Med.*, vol. 94, pp. 19–26, Mar. 2018.

[22] M. S. Wibawa, I. M. D. Maysanjaya, N. K. Novianti, and P. N. Crisnapati, "Abnormal heart rhythm detection based on spectrogram of heart sound using convolutional neural network," in *Proc. 6th Int. Conf. Cyber IT Service Manage.*, Aug. 2018, pp. 1–4.

[23] A. Meintjes, A. Lowe, and M. Legget, "Fundamental heart sound classification using the continuous wavelet transform and convolutional neural networks," in *Proc. 40th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, Jul. 2018, pp. 409–412.

[24] B. Pourbabaee, M. J. Roshtkhari, and K. Khorasani, "Deep convolutional neural networks and learning ECG features for screening paroxysmal atrial fibrillation patients," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 48, no. 12, pp. 2095–2104, Dec. 2018.

[25] S. Kiranyaz, T. Ince, and M. Gabbouj, "Real-time patient-specific ECG classification by 1-D convolutional neural networks," *IEEE Trans. Biomed. Eng.*, vol. 63, no. 3, pp. 664–675, Mar. 2016.

[26] H. G. Lee, K. Y. Noh, and K. H. Ryu, *Mining Biosignal Data: Coronary Artery Disease Diagnosis Using Linear and Nonlinear Features of HRV*. Berlin, Germany: Springer, 2007.

[27] A. Kaveh and W. Chung, "Automated classification of coronary atherosclerosis using single lead ECG," in *Proc. IEEE Conf. Wireless Sensor (ICWISE)*, Dec. 2013, pp. 108–113.

[28] M. Deng, M. Tang, C. Wang, L. Shan, L. Zhang, J. Zhang, W. Wu, and L. Xia, "Cardiodynamicsgram as a new diagnostic tool in coronary artery disease patients with nondiagnostic electrocardiograms," *Amer. J. Cardiol.*, vol. 119, no. 5, pp. 698–704, 2017.

[29] A. N. Makaryus, J. N. Makaryus, A. Figgatt, D. Mulholland, H. Kushner, J. L. Semmlow, J. Mieres, and A. J. Taylor, "Utility of an advanced digital electronic stethoscope in the diagnosis of coronary artery disease compared with coronary computed tomographic angiography," *Amer. J. Cardiol.*, vol. 111, no. 6, pp. 786–792, 2013.

[30] S. Schmidt, C. Holst-Hansen, J. Hansen, E. Toft, and J. Struijk, "Acoustic features for the identification of coronary artery disease," *IEEE Trans. Biomed. Eng.*, vol. 62, no. 11, pp. 2611–2619, Nov. 2015.

[31] D. Gauthier, Y. M. Akay, R. G. Paden, W. Pavlicek, F. D. Fortuin, J. K. Sweeney, R. W. Lee, and M. Akay, "Spectral analysis of heart sounds associated with coronary occlusions," in *Proc. 6th Int. Special Topic Conf. Inf. Technol. Appl. Biomed.*, Nov. 2007, pp. 49–52.

[32] F. Sato, "Symposium on limitation of diagnostic value of electrocardiography and phonocardiography," *Jpn. Circulat. J.*, vol. 30, no. 12, pp. 1537–1541, 1966.

[33] M. H. Crawford, C. A. Mendoza, and R. A. O'Rourke, D. H. White, C. A. Boucher, and J. Gorwit, "Limitations of continuous ambulatory electrocardiogram monitoring for detecting coronary artery disease," *Ann. Internal Med.*, vol. 89, no. 1, pp. 1–5, 1978.

[34] V. X. Afonso, W. J. Tompkins, T. Q. Nguyen, and S. Luo, "ECG beat detection using filter banks," *IEEE Trans. Biomed. Eng.*, vol. 46, no. 2, pp. 192–202, Feb. 1999.

[35] M. Xu, S. Wei, X. Qin, Y. Zhang, and C. Liu, "Rule-based method for morphological classification of ST segment in ECG signals," *J. Med. Biol. Eng.*, vol. 35, no. 6, pp. 816–823, 2015.

[36] J. J. Lee, J. H. Lee, J. W. Jeong, and J. Y. Chung, "Fragmented QRS and abnormal creatine kinase-MB are predictors of coronary artery disease in patients with angina and normal electrocardiographys," *Korean J. Internal Med.*, vol. 32, no. 3, p. 469, 2017.

[37] E. Avci and Z. H. Akpolat, "Speech recognition using a wavelet packet adaptive network based fuzzy inference system," *Expert Syst. Appl.*, vol. 31, no. 3, pp. 495–503, 2006.

[38] D. B. Springer, L. Tarassenko, and G. D. Clifford, "Logistic regression-HSMM-based heart sound segmentation," *IEEE Trans. Biomed. Eng.*, vol. 63, no. 4, pp. 822–832, Apr. 2016.

[39] C. Liu *et al.*, "An open access database for the evaluation of heart sound algorithms," *Physiol. Meas.*, vol. 37, no. 12, p. 2181, 2016.

[40] H. Tang, Z. Dai, Y. Jiang, T. Li, and C. Liu, "PCG Classification Using Multidomain Features and SVM Classifier," *BioMed Res. Int.*, vol. 2018, Jul. 2018, Art. no. 4205027.

[41] J. S. Richman and J. R. Moorman, "Physiological time-series analysis using approximate entropy and sample entropy," *Amer. J. Physiol.-Heart Circulatory Physiol.*, vol. 278, no. 6, pp. H2039–H2049, 2000.

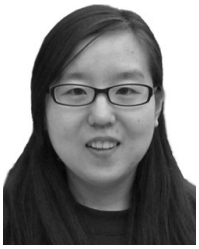
[42] L. Ji, P. Li, K. Li, X. Wang, and C. Liu, "Analysis of short-term heart rate and diastolic period variability using a refined fuzzy entropy method," *Biomed. Eng. Online*, vol. 14, no. 1, p. 64, Jul. 2015.

[43] P. N. Tan, M. Steinbach, and V. Kumar, *Introduction to Data Mining*. London, U.K.: Pearson, 2006.

[44] M. Akay, Y. M. Akay, D. Gauthier, R. G. Paden, W. Pavlicek, F. D. Fortuin, J. P. Sweeney, and R. W. Lee, "Dynamics of diastolic sounds caused by partially occluded coronary arteries," *IEEE Trans. Biomed. Eng.*, vol. 56, no. 2, pp. 513–517, Feb. 2009.
[45] K. He, X. Zhang, S. Ren, and J. Sun, "Delving deep into rectifiers: Surpassing human-level performance on imagenet classification," in *Proc. ICCV*, Dec. 2015, pp. 1026–1034.
[46] J. Chung, C. Gulcehre, K. Cho, and Y. Bengio, "Empirical evaluation of gated recurrent neural networks on sequence modeling," in *Proc. Conf. Neural Inf. Process. Syst.*, 2014, pp. 1–4.
[47] Z. Yang, D. Yang, C. Dyer, X. He, A. Smola, and E. Hovy, "Hierarchical attention networks for document classification," in *Proc. Conf. North Amer. Chapter Assoc. Comput. Linguistics: Hum. Lang. Technol.*, 2016, pp. 1480–1489.
[48] M. Zarrabi, H. Parsaei, R. Boostani, A. Zare, Z. Dorfeshan, K. Zarrabi, and J. Kojuri, "A system for accurately predicting the risk of myocardial infarction using PCG, ECG and clinical features," *Biomed. Eng., Appl., Basis Commun.*, vol. 29, no. 3, 2017, Art. no. 1750023.
[49] M. Kumar, R. B. Pachori, and U. R. Acharya, "Characterization of coronary artery disease using flexible analytic wavelet transform applied on ECG signals," *Biomed. Signal Process. Control*, vol. 31, pp. 301–308, Jan. 2017.
[50] S. L. OH, M. Adam, J. H. Tan, Y. Hagiwara, V. K. Sudarshan, J. E. W. Koh, K. C. Chua, K. P. Chua, R. S. Tan, and E. Y. Ng, "Artery disease from short-term 12 lead electrocardiogram signals by using wavelet packet decomposition and common spatial pattern techniques," *J. Mech. Med. Biol.*, vol. 17, no. 7, 2017, Art. no. 1740007.
[51] F. Chollet, *Deep Learning with Python*. Shelter Island, New York, USA: Manning Publications, 2018.
[52] C. Potes, S. Parvaneh, A. Rahman, and B. Conroy, "Ensemble of feature-based and deep learning-based classifiers for detection of abnormal heart sounds," in *Proc. IEEE Comput. Cardiol. Conf.*, Sep. 2016, pp. 621–624.

HAN LI received the B.S. degree in automation from Shandong University, Jinan, China, in 2014, where she is currently pursuing the Ph.D. degree with the School of Control Science and Engineering.

Her current research interests include the application of computational intelligence in the detection of cardiovascular diseases and biomedical signal processing.



XINPEI WANG received the B.S. and Ph.D. degrees in biomedical engineering from Shandong University, Jinan, China, in 2005 and 2011, respectively, where she is currently a Lecturer with the School of Control Science and Engineering.

Her current research interests include biomedical signal and image processing, biomedical measurements and devices, and machine learning.

CHANGCHUN LIU received the B.S. and M.S. degrees in automation from Shandong Industrial University (now merged into Shandong University), Jinan, China, in 1982 and 1987, respectively. Since 2000, he has been a Professor in biomedical engineering with the School of Control Science and Engineering, Shandong University.

His current research interests include novel solutions for noninvasive detection of cardiovascular function, biomedical signal processing, biomedical measurements and devices, and physiological system modeling.

YAN WANG is currently the Director of the Health Management Center of the Affiliated Hospital of Qingdao University. She is also a member of the Health Management Branch of the Chinese Medical Association and the Vice Director of the Health Management Branch of the Shandong Provincial Medical Association.

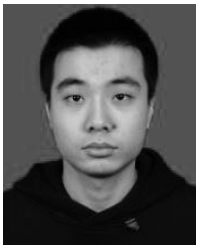
Her current research interests include physical examination project, health promotion program, and health management service.

PENG LI received the B.S. and Ph.D. degrees in biomedical engineering from Shandong University, Jinan, China, in 2009 and 2014, respectively. He held multiple disciplinary postdoctoral training with Shandong University, from 2014 to 2015, and then with the Harvard Medical School, Boston, MA, USA, from 2016 to 2018. He is currently an Instructor in medicine with the Division of Sleep Medicine, Harvard Medical School, and an Associate Physiologist with the Division of Sleep and Circadian Disorders, Brigham and Womens Hospital, Boston.

His current research interests include understanding the nonlinear dynamics in cardiovascular systems, brain activity, and motor activity. In particular, he is recently interested in predicting Alzheimers dementia and cardiovascular diseases noninvasively and the underlying neuro-pathological mechanisms.

HONG TANG received the B.S. degree in mechanical manufacture and automation from the Zhongyuan University of Technology, in 2000, the M.S. degree in biomedical engineering from Jilin University, in 2003, and the Ph.D. degree in signal processing from the Dalian University of Technology, China, in 2006. He was a Research Fellow with the Regional Innovation Center of Yeungnam University, South Korea, from 2006 to 2007. He is currently an Associate Professor with the School of Biomedical Engineering, Dalian University of Technology.

His current research interests include signal processing in heart sounds, mechanism of heart sound generation, relations between heart sounds and heart hemodynamics, and biomedical signal compression.



LIANKE YAO received the B.S. degree in biomedical engineering from the Chongqing University of Technology, Chongqing, China, in 2012. He is currently pursuing the Ph.D. degree with Shandong University.

His current research interests include biomedical signal processing and analysis of QT interval variability.

HUAN ZHANG received the B.S. degree in electronic science and technology from Shanxi University, Taiyuan, China, in 2013. She is currently pursuing the Ph.D. degree with the School of Control Science and Engineering, Shandong University, China.

Her current research interests include biomedical signal processing, machine learning, and early detection of coronary artery disease.