Computers in Biology and Medicine 120 (2020) 103733

Contents lists available at ScienceDirect

Computers in Biology and Medicine

journal homepage: http://www.elsevier.com/locate/compbiomed

A fusion framework based on multi-domain features and deep learning features of phonocardiogram for coronary artery disease detection

Han Li [a], Xinpei Wang [a,**], Changchun Liu [a,*], Qiang Zeng [b], Yansong Zheng [b], Xi Chu [c], Lianke Yao [a], Jikuo Wang [a], Yu Jiao [a], Chandan Karmakar [d]

[a] School of Control Science and Engineering, Shandong University, Jinan, 250061, China
[b] Health Management Institute, Chinese PLA General Hospital, Beijing, 100853, China
[c] Health Management Center, Xuanwu Hospital, Capital Medical University, Beijing, China
[d] School of Information Technology, Deakin University, Burwood, VIC3125, Australia

ARTICLE INFO

Keywords:
Feature fusion
Multi-domain features
Deep learning
Phonocardiogram
Coronary artery disease

ABSTRACT

Phonocardiogram (PCG) signals reflect the mechanical activity of the heart. Previous studies have reported that PCG signals contain heart murmurs caused by coronary artery disease (CAD). However, the murmurs caused by CAD are very weak and rarely heard by the human ear. In this paper, a novel feature fusion framework is proposed to provide a comprehensive basis for CAD diagnosis. A dataset containing PCG signals of 175 subjects was collected and used. A total of 110 features were extracted from multiple domains, and then reduced and selected. Images obtained by Mel-frequency cepstral coefficients were used as the input for the convolutional neural network for feature learning. Then, the selected features and the deep learning features were fused and fed into a multilayer perceptron for classification. The proposed feature fusion method achieved better classification performance than multi-domain features or deep learning features alone, with accuracy, sensitivity, and specificity of 90.43%, 93.67%, and 83.36%, respectively. A comparison with existing studies demonstrated that the proposed method was a promising noninvasive screening tool for CAD under general medical conditions.

1. Introduction

Coronary artery disease (CAD) is one of the leading causes of death worldwide. It is caused by the deposition of plaque along the inner wall of coronary arteries and may lead to sudden cardiac death in severe cases. Currently, coronary angiography, the gold standard for CAD diagnosis, is invasive and expensive, and thereby only applicable to specialized clinical departments [1]. Phonocardiogram (PCG) auscultation is a low-cost and convenient technique widely used for cardiac disease screening [2]. Earlier studies have reported that coronary stenosis can produce weak heart murmurs in PCG signals [3,4]. However, these murmurs are rarely audible, and thus there is an urgent need to develop automated PCG analysis methods for CAD detection [5].

Conventional methods for the recognition of abnormal PCG signals involve two steps: feature extraction and classification. Extracting appropriate features contributes to capturing abnormal cardiac states. The extracted features in current literature are usually from the time [6,7], frequency [8,9], time-frequency [10–12], and nonlinear [13,14] domains. With the extracted features, the next step is PCG classification. To distinguish between pathological and normal PCG signals, a variety of classifiers have been applied, such as artificial neural network [6], support vector machine [7,11,12,14], random forest [8], Naïve Bayes [9], k-nearest neighbor [9], and tree-based approaches [10,13]. In recent years, inspired by the success of deep learning in the fields of emotion recognition, activity recognition, and image recognition [15–17], the convolutional neural network (CNN) has begun to be applied to the classification of PCG signals [18–24]. The CNN is a hierarchical feature extraction algorithm that replaces manual feature extraction and eliminates the need for feature engineering [21].

Among the current studies on PCG classification, most have been based on the publicly available databases PASCAL [6,11] and PhysioNet/CinC Challenge [7,8,10,14,20–24] for the recognition of abnormal PCG signals. Moreover, several studies have used datasets collected for the identification of valvular heart disease [9,12,13]. In contrast, the number of studies on the detection of CAD is relatively small [25–28]. These studies focused on traditional feature extraction and classification

* Corresponding author.
** Corresponding author.
E-mail addresses: j595018526@163.com (H. Li), wangxinpei@sdu.edu.cn (X. Wang), changchunliu@sdu.edu.cn (C. Liu).

https://doi.org/10.1016/j.compbiomed.2020.103733
Received 8 January 2020; Received in revised form 23 March 2020; Accepted 24 March 2020
Available online 30 March 2020

methods but the results obtained are unsatisfactory. Among them, Samanta et al. [28] obtained the highest accuracy of 82.57% using features extracted from four-channel PCG signals. In terms of deep learning-based PCG classification, few studies have investigated the detection of CAD. In practical applications, manually extracted features can only reflect the variations in PCG signals caused by CAD from certain perspectives. If hand-crafted features can be combined with deep learning features that have strong representation capabilities, more comprehensive disease-related information can be captured.

Therefore, this paper presents a novel framework to fuse extracted multi-domain features with the deep learning features of PCG for the detection of CAD. In the fusion framework, features extracted from multiple domains are reduced and selected.On the other hand, the Mel-frequency cepstral coefficients (MFCCs) extracted from PCG signals are converted into images and used as the input to the CNN for feature learning. Then, the multi-domain features and deep learning features are fused and fed into a multilayer perceptron (MLP) classifier to obtain the prediction probability. The main contributions of this study are as follows.

(1) A new framework is proposed to fuse the extracted multi-domain features with the deep learning features of PCG for the detection of CAD. The feature fusion framework combines the advantages of both traditional methods and deep learning techniques. Therefore, more underlying features are extracted and a more comprehensive feature representation is provided for identifying the variations in PCG signals caused by CAD.
(2) Compared with existing studies, more types of features (six-domain hand-crafted features and MFCCs deep learning features) are extracted to reflect richer information related to heart states.
(3) The proposed method is validated on a PCG dataset of 175 subjects collected from the hospital. The proposed method achieves an accuracy of 90.43% in detection of CAD, which is very competitive compared with the results of existing studies. The results demonstrate that the proposed method is very promising as a non-invasive and available screening tool of CAD, especially under general medical conditions. The computation time and required resources indicate that the proposed computer-assisted method has the potential for real-world applications.

The remainder of this paper is organized as follows. Section 3 describes the data acquisition and preprocessing, multi-domain feature extraction, feature reduction and selection methods, MFCCs, proposed feature fusion method, and performance evaluation strategy. The experimental results are presented in Section 4, Section 5 discusses the results, and Section 6 concludes this work and provides prospective extensions.

2. Mechanism of heart murmurs caused by CAD

As early as 1967, it was reported that heart murmurs occurred during the diastole of PCG because of coronary stenosis [29]. The studies of Akay et al. [4,30] showed that turbulent blood flow was generated when the occlusion rate of the coronary artery was between 25 and 95%. The turbulence excites the vascular system to produce acoustic resonance, namely, heart murmurs. The weak heart murmurs formed by the turbulence contain useful information about coronary stenosis and are distributed at high frequencies. However, the auditory component produced by coronary artery occlusion is usually weakened by intermediate tissues and obscured by loud valve sounds. During the diastolic period without valve sounds, the myocardium exerts minimal pressure on the coronary arteries and the blood flow in the coronary arteries is the greatest. At this time, the pathological murmurs produced are the loudest and enhance the possibility of detecting this weak component. The diastolic heart murmurs are the clearest between the third and fourth ribs of the left margin auscultation area. Our research is centered on the PCG signal in this area.

3. Materials and methods

3.1. Data acquisition and preprocessing

From November 2017 to September 2018, 175 subjects were recruited consecutively in the Shandong Provincial Qianfoshan Hospital. All subjects signed an informed consent form before participation. The research protocol was approved by the clinical ethics committee of the hospital and conducted in accordance with the principles of the Helsinki Declaration and its subsequent amendments. No attempts were made to control the medicine administration of the subjects. The inclusion criterion was subjects scheduled to undergo a coronary angiography within two days. Two types of subjects were excluded from the study: (a) subjects who had previously undergone percutaneous coronary intervention or coronary artery bypass surgery, and (b) subjects who has valvular heart disease as verified by echocardiography.

Each subject was required to lie in a supine position for 15 min in a quiet and temperature-controlled room (25 ± 3°C) between 2 and 6 p.m. prior to the measurement. Then, a piezoelectric sensor was placed in the third intercostal space on the left sternal border of the subject, and a 5-min PCG signal was recorded at a sampling rate of 1 kHz using a cardiovascular function detection device (CVFD-II, Huiyironggong Technology Co., Ltd, Jinan, China). In addition, basic demographic data including age, sex, height, weight, body mass index, heart rate, and blood pressure were also recorded. A total of 120 CAD and 55 non-CAD subjects participated in the study. According to the results of coronary angiography, subjects with at least one major coronary artery branch (left anterior descending, left circumflex, or right coronary artery) $\geq$ 50% stenosis were categorized as CAD, and the others as non-CAD. The basic characteristics of the subjects are given in Table 1.

The raw PCG signals were denoised by a high-pass Butterworth filter (25 Hz), and the power-line interference (50 Hz) was eliminated by an IIR notch filter. The z-score method was used to normalize the signals. Each 5-min recording was cropped into 20 15-s signal segments, and 3,500 samples were generated, including 2,400 CAD and 1,100 non-CAD samples. The PCG signals of a CAD and a non-CAD subject after preprocessing are shown in Fig. 1. Compared with the non-CAD subject, increased heart murmurs in the diastolic period can be observed in the CAD subject's PCG signal.

3.2. Multi-domain feature extraction

Using Springer's algorithm [31], each cardiac cycle of the PCG signal is segmented into four states: S1, systole, S2, and diastole. On this basis, 110 features are extracted from multi-domains. Except for the spectrum features of the PCG signal and the wavelet features, the mean and standard deviation of the features extracted from the four states are calculated over all cycles of a 15-s signal.

Table 1
Demographics of subjects.

Characteristic	CAD	Non-CAD
Age	61 ± 10	56 ± 7
Male/female	81/39	25/30
Height	166 ± 8	166 ± 8
Weight	71 ± 10	63 ± 10
Body mass index	26 ± 3	23 ± 3
Heart rate	67 ± 8	68 ± 7
Systolic blood pressure	133 ± 16	115 ± 16
Diastolic blood pressure	83 ± 12	72 ± 10

Values are represented as male/female or mean ± standard deviation.

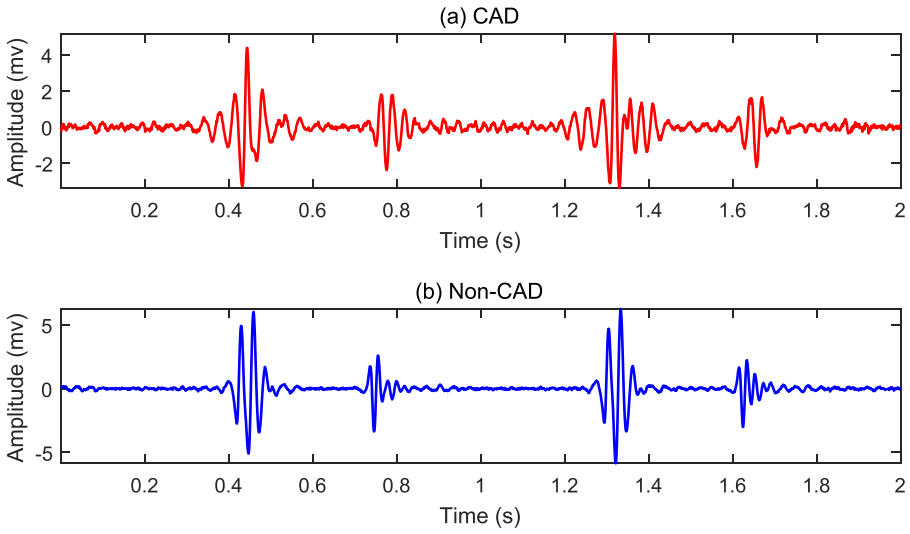

Fig. 1. PCG signals of a CAD and a non-CAD subject.

(1) Time features (20 features). According to the position information of S1, systole, S2, and diastole, the time interval and amplitude features are extracted from each cardiac cycle [32], and the details are given in Table 2.

(2) Energy features (20 features). The energy of the time series x_i is defined as

$$Energy = \sum_{i=1}^{N} x_i^2, \tag{1}$$

where i denotes the discrete time index and N is the length of x_i. The energy ratio between two different states is computed and the details of the features extracted from each cardiac cycle are given in Table 3.

(3) Higher-order statistics features (16 features). The third- and fourth-order statistics, namely, skewness and kurtosis, respectively, are computed for the four states in each cardiac cycle [33]. Skewness measures the asymmetry degree of statistical distribution, and kurtosis represents the peak height of the probability density distribution curve at the average. The two statistics are computed as

$$skewness = \frac{\sum_{i=1}^{N}(x_i - m)^3}{N \times sd^3}, \tag{2}$$

$$kurtosis = \frac{\sum_{i=1}^{N}(x_i - m)^4}{N \times sd^4} - 3, \tag{3}$$

where m and sd denote the mean and standard deviation of the time series x_i, respectively.

Table 2
Time features extracted from each cardiac cycle.

Feature	Description
IntCC	Interval duration of cardiac cycle
IntS1	Interval duration of S1
IntSys	Interval duration of Systole
IntS2	Interval duration of S2
IntDia	Interval duration of Diastole
Ratio_SysCC	Duration ratio of systole to cardiac cycle
Ratio_DiaCC	Duration ratio of diastole to cardiac cycle
Ratio_SysDia	Duration ratio of systole to diastole
Amp_SysS1	Average amplitude ratio of systole to S1
Amp_DiaS2	Average amplitude ratio of diastole to S2

Table 3
Energy features extracted from each cardiac cycle.

Feature	Description
Energy_S1ToSys	Energy ratio of S1 to systole
Energy_S1ToDia	Energy ratio of S1 to diastole
Energy_S2ToSys	Energy ratio of S2 to systole
Energy_S2ToDia	Energy ratio of S2 to diastole
Energy_DiaToSys	Energy ratio of diastole to systole
Energy_S1ToCC	Energy ratio of S1 to cardiac cycle
Energy_SysToCC	Energy ratio of systole to cardiac cycle
Energy_S2ToCC	Energy ratio of S2 to cardiac cycle
Energy_DiaToCC	Energy ratio of diastole to cardiac cycle
Energy_HsToCC	Energy ratio of S1 and S2 to cardiac cycle

(4) Spectrum features of the four states and the PCG signal (24 features). The discrete Fourier transform is employed to obtain the spectra of the four states in each cardiac cycle and the entire 15-s PCG signal. In the spectra of the segmented four states, the proportions of high-frequency (above 200 Hz) and low-frequency (below 50 Hz) components are calculated. Moreover, eight features are extracted from the spectrum of the entire 15-s signal as shown in Table 4. Among them, the entropy of a sequence e_i of length L is calculated as

$$Entropy = -\sum_{i=1}^{L} \frac{e_i}{\sum_{j=1}^{L} e_j} \times \log \frac{e_i}{\sum_{j=1}^{L} e_j}. \tag{4a}$$

(5) Wavelet features (22 features). Wavelet transform is a popular time-frequency analysis method [34]. The db6 mother wavelet is

Table 4
Spectrum features extracted from each 15-s PCG signal.

Feature	Description
m_spectrum	Mean of spectrum
sd_spectrum	Standard deviation of spectrum
ske_spectrum	Skewness of spectrum
kur_spectrum	kurtosis of spectrum
HF_ratio	Proportion of high-frequency components (above 200 Hz) in spectrum
LF_ratio	Proportion of low-frequency components (below 50 Hz) in spectrum
HF_LF_ratio	Ratio of high-frequency to low-frequency components
Entropy_spectrum	Entropy of the spectrum energy

used to decompose each 15-s PCG signal into four levels, and then nine wavelet decomposition features are extracted, including total energy and the energy ratio of wavelet coefficients in different frequency bands. To analyze the high-frequency components of the PCG signal in more detail, three-level wavelet packet decomposition is applied [35]. Eight reconstructed coefficient vectors at the third level of the wavelet packet tree are computed, and 13 wavelet packet decomposition features are extracted, including total energy, the energy ratio, and energy entropy of the reconstructed coefficients. Among them, the energy and energy entropy features are calculated using Equations (1) and (3), respectively. The details of the extracted wavelet features are given in Table 5.

(6) Entropy features (8 features). To measure the nonlinear dynamics of systolic and diastolic time series in each cardiac cycle, sample entropy and refined fuzzy entropy are applied. Sample entropy is a statistical measure developed to quantify the regularity of the time series [36]. Refined fuzzy entropy was proposed by replacing the Gaussian function in the traditional fuzzy entropy measure with a piecewise fuzzy membership function, with good stability and robustness against noise [37].

3.3. Feature reduction and selection

Principal component analysis (PCA) is a frequently-used linear dimensionality reduction technique [38]. It transforms the high-dimensional data into a low-dimensional feature space (typically referred to as principal components) while preserving the maximum variance of the original data. The orthogonal principal components contain no redundant information. The transformation process involves computing the eigenvectors and eigenvalues based on the covariance matrix of the original data, then arranging the eigenvectors in descending order of the eigenvalues, and finally projecting the original data onto the direction of the sorted eigenvectors to obtain the transformed features.

Table 5
Wavelet features extracted from each 15-s PCG signal.

Feature	Description
Extracted wavelet decomposition features:	
Wd_total	Energy sum of wavelet coefficients in decomposed five frequency bands
Wd_ratio1	Energy of wavelet approximation coefficients in 0 – 31.25 Hz
Wd_ratio2	Energy of wavelet detail coefficients in 31.25 – 62.5 Hz
Wd_ratio3	Energy of wavelet detail coefficients in 62.5 – 125 Hz
Wd_ratio4	Energy of wavelet detail coefficients in 125 – 250 Hz
Wd_ratio5	Energy of wavelet detail coefficients in 250 – 500 Hz
Wd_low_ratio	Energy sum of wavelet coefficients below 62.5 Hz
Wd_mid_ratio	Energy sum of wavelet coefficients in 62.5 – 250 Hz
Wd_high_ratio	Energy sum of wavelet coefficients above 250 Hz
Extracted wavelet packet decomposition features:	
Wp_total	Energy sum of reconstructed coefficients in decomposed eight frequency bands
Wp_ratio1	Energy of reconstructed coefficients in 0 – 62.5 Hz
Wp_ratio2	Energy of reconstructed coefficients in 62.5 – 125 Hz
Wp_ratio3	Energy of reconstructed coefficients in 125 – 187.5 Hz
Wp_ratio4	Energy of reconstructed coefficients in 187.5 – 250 Hz
Wp_ratio5	Energy of reconstructed coefficients in 250 – 312.5 Hz
Wp_ratio6	Energy of reconstructed coefficients in 312.5 – 375 Hz
Wp_ratio7	Energy of reconstructed coefficients in 375 – 437.5 Hz
Wp_ratio8	Energy of reconstructed coefficients in 437.5 – 500 Hz
Wp_low_ratio	Energy sum of reconstructed coefficients below 62.5 Hz
Wp_mid_ratio	Energy sum of reconstructed coefficients 62.5 – 250 Hz
Wp_high_ratio	Energy sum of reconstructed coefficients above 250 Hz
Wp_entropy	Energy entropy of reconstructed coefficients in decomposed eight frequency bands

Three feature selection methods are used to determine the least number of principal components that contribute to classification. The Chi-square test is a nonparametric statistical technique used to determine whether the distribution of an observed frequency differs from that of a theoretically expected frequency. The Gini index measures the impurity of a feature, and the correlation method represents the degree of association between a feature and a class label. These three methods are calculated as

$$\chi^2 = \sum \frac{(O-E)^2}{E}, \tag{4b}$$

$$Gini = 1 - \sum_{i \in Classes} [p(i|A)]^2, \tag{5}$$

$$Correlation = \frac{\sum_{i=1}^{n}(A_i - \overline{A}) \times (C_i - \overline{C})}{(n-1) \times \sigma(A) \times \sigma(C)}, \tag{6}$$

where O and E denote the observed frequency and expected frequency, $p(i|A)$ indicates the probability that records in feature A belong to class i; $\overline{A}$ and $\sigma(A)$ are the mean and standard deviation of the feature A, respectively; and $\overline{C}$ and $\sigma(C)$ are the mean and standard deviation of the label C, respectively. The details of feature selection are further explained in Section 4.1.

3.4. MFCCs

In recent literature, MFCCs have been used to transform PCG signals into images, enabling CNN to visually recognize local patterns [21,23,24]. MFCCs are a time-frequency representation approach widely used in speech recognition involving three extraction steps [39]. First, the audio signal is divided into short frames and a discrete Fourier transform is performed on each frame. Then, the triangular band-pass filters at the Mel-frequency scale are used to smooth the spectrum, and the logarithm energies of the spectrum in the filterbank are calculated. Finally, the discrete cosine transform of the log filterbank energies is performed to obtain the 12-dimensional MFCCs as:

$$C_n = \sum_{m=1}^{M} D_m \cos\left[\frac{(m-0.5)\pi n}{M}\right], \tag{8}$$

where M denotes the number of filters, D_m denotes the output of the m-th filter, and C_n denotes the n-th coefficient. In addition to the MFCCs, 12 first-order and 12 s-order differential coefficients are also calculated to describe the trajectories of MFCCs over time, i.e., the dynamic characteristics.

The MFCCs images of the PCG signals from a CAD and a non-CAD subject are shown in Fig. 2. The rhythm of S1 and S2 seems less obvious compared to that of the non-CAD subject because of the increased heart murmurs in the CAD subject's PCG signal.

3.5. Fusion framework of multi-domain features and deep learning features

A fusion framework based on extracted multi-domain features and deep learning features of PCG is proposed for CAD detection, as shown in Fig. 3.

First, multi-domain features are extracted from the preprocessed PCG signals and normalized by the z-score method. Subsequently, PCA is used to reduce the feature dimension, and the principal components contributing to classification are selected using three methods (i.e., Chi-square test, Gini index, and correlation). On the other hand, the MFCCs are extracted from the PCG signals and converted to RGB images for deep learning. The RGB images contribute to improving the classification performance by using the filters estimated by CNN. The end-to-end

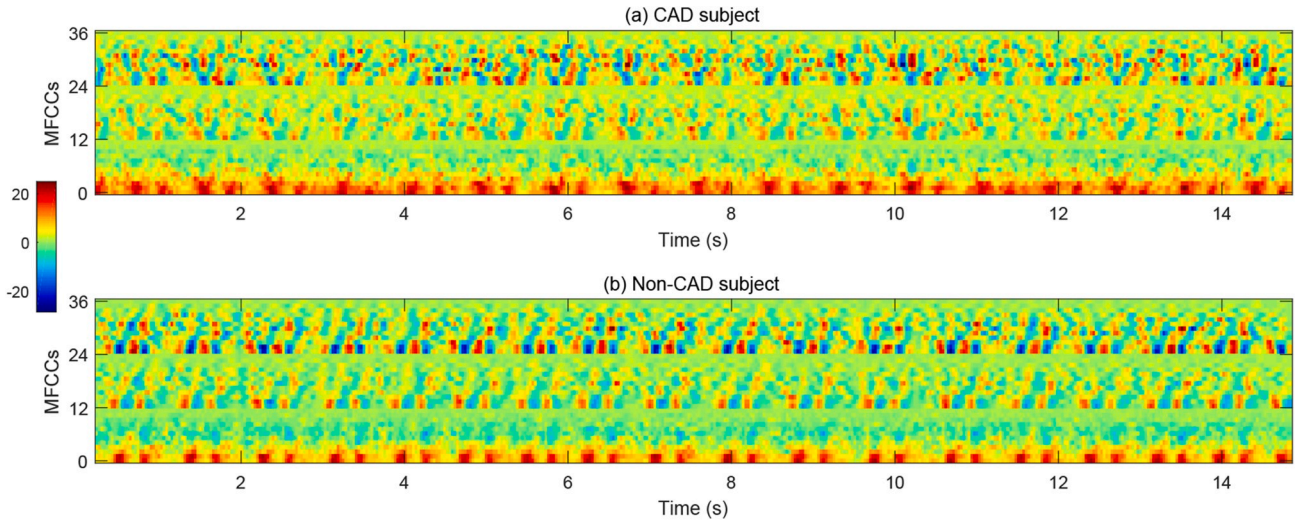

Fig. 2. MFCCs images of the PCG signals from a CAD and a non-CAD subject. The scales 1–12, 13–24, and 25–36 on the y-axis indicate the cepstrum coefficients, first-order differential coefficients, and second-order differential coefficients, respectively.

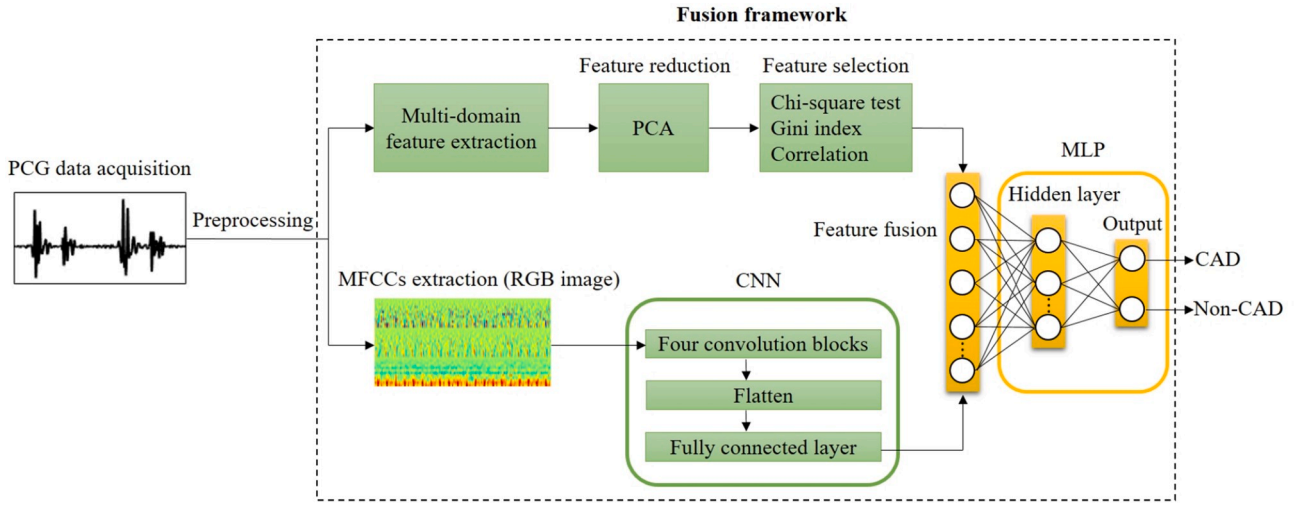

Fig. 3. The fusion framework based on multi-domain features and deep learning features of PCG for CAD detection.

CNN used to learn features from the MFCCs images is composed of four identical convolution blocks and a fully connected layer. Different model parameters are investigated to achieve an almost optimal CNN structure. The optimal parameter combination is obtained by minimizing the binary cross-entropy loss function. The detailed configurations of the CNN are listed in Table 6.

Inside the CNN, the convolution blocks can be regarded as automatic feature extractors, each consisting of three convolutional layers. The stride of the first two layers has default value of 1, and the stride of the third layer is set to 2 to reduce the feature space. In the convolutional layer, the convolution operation is performed through multiple filters to obtain the feature maps, expressed as

$$y_j = f\left(\sum_{i \in M} k_{ij} * x_i + b_j\right) \tag{9}$$

where M denotes the filter size, i denotes the size of the input feature maps, j denotes the size of the convolution kernels, and k_{ij} denotes the convolution kernel for the i-th input and the j-th output. Subsequently, the features extracted by convolution blocks are flattened, and a fully connected layer with 128 neurons is used to map the high-dimensional deep learning features into low-dimensional features for classification,

Table 6
Detailed CNN parameter configurations.

Layer no.	Layer type	Output shape	Kernel size	Stride
0	Input	36 × 280 × 3	–	–
1	Convolution	36 × 280 × 32	3 × 3	1
2	Convolution	36 × 280 × 32	3 × 3	1
3	Convolution	18 × 140 × 32	3 × 3	2
4	Convolution	18 × 140 × 32	3 × 3	1
5	Convolution	18 × 140 × 32	3 × 3	1
6	Convolution	9 × 70 × 32	3 × 3	2
7	Convolution	9 × 70 × 32	3 × 3	1
8	Convolution	9 × 70 × 32	3 × 3	1
9	Convolution	5 × 35 × 32	3 × 3	2
10	Convolution	5 × 35 × 32	3 × 3	1
11	Convolution	5 × 35 × 32	3 × 3	1
12	Convolution	3 × 18 × 32	3 × 3	2
13	Flatten	1728 × 1	–	–
14	Full connection	128 × 1	–	–

which can be expressed as

$$z = W^T x + b, \tag{10}$$

where x, W, and b denote the input, weight, and bias, respectively. Then, the transformed deep learning features are fused with the selected principal components, and these feature values are converted into a two-class prediction probability (i.e., CAD or non-CAD) using an MLP classifier that contains only one hidden layer with 64 neurons.

Batch normalization is used in each convolution block, and a dropout with a probability of 0.2 is added after the fully connected layer and the hidden layer of MLP to improve the learning speed and reduce the risk of overfitting. The leaky rectifier linear unit [40] is applied as the activation function of layers $1-12$ and 14. The adaptive moment estimation [41] is utilized as the optimizer. An initial learning rate of 0.0001 is used, and the batch size is set to 64.

3.6. Performance evaluation

A five-fold cross-validation is conducted in this study. First, the 5-min PCG recordings of all subjects are divided into five subsets by stratified sampling, and the recordings in each subset are cropped into 20 15-s segments. Four of the five subsets are used as the training set, and the fifth is used as the validation set. Then, five iterations are performed, and the final classification result is the average of the cross-validations.

Sensitivity (Sen), specificity (Spe), and accuracy (Acc) are common measures for assessing classification performance. Moreover, to evaluate the performance of imbalanced data, the G-mean is also used to consider the balanced performance between two classes. These evaluation measures are computed as

$$Acc = \frac{TP + TN}{TP + TN + FN + FP}, \tag{11}$$

$$Sen = \frac{TP}{TP + FN}, \tag{12}$$

$$Spe = \frac{TN}{TN + FP}, \tag{13}$$

$$G - mean = \sqrt{Sen \times Spe}, \tag{14}$$

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

4. Results

This section presents the classification results of the selected multi-domain features, deep learning features, and fusion features. In addition, the results obtained are compared with the existing studies on CAD detection using PCG signals.

4.1. Results before and after feature reduction and selection

For the extracted multi-domain features, the PCA is first used to reduce the feature dimension. Then, the Chi-square test, Gini index, and correlation method are applied separately to determine the optimal principal component subset for the subsequent feature fusion. The G-mean is used to evaluate the overall classification performance of the imbalanced data. The classification performances using MLP before and after feature reduction and selection are given in Table 7.

As shown in Table 7, by using the PCA for feature reduction, 110 original features are transformed into 44 principal components and the classification performance is improved. After feature selection, the specificity and G-mean increase, but the sensitivity decreases slightly for the three methods. The accuracy decreases when the Chi-square test and Gini index are used for feature selection, and only increases when the correlation method is used. Among the three feature selection methods, the correlation method performs best, achieving an accuracy of 85.03% with 29 principal components.

Table 7
Classification performance before and after feature reduction and selection.

Method	No. Of features	Acc (%)	Sen (%)	Spe (%)	G-mean (%)
Before feature selection:					
Original features	110	82.69 ± 4.79	89.21 ± 5.39	68.45 ± 12.10	77.78 ± 6.74
PCA	44	84.51 ± 3.54	91.00 ± 3.72	70.36 ± 11.28	79.71 ± 5.92
After feature selection using three methods:					
Chi-square test	38	84.23 ± 4.44	90.04 ± 3.39	71.55 ± 10.99	80.03 ± 6.50
Gini index	27	84.11 ± 4.83	89.71 ± 3.88	71.91 ± 10.65	80.12 ± 6.60
Correlation	29	**85.03 ± 4.53**	**90.50 ± 3.94**	**73.09 ± 11.27**	**81.07 ± 6.60**

The best performance is marked with bold.

The variations of G-mean with an increase in the number of principal components when using three feature selection methods are shown in Fig. 4. The trend of G-mean for the three methods are consistent on the whole. In the beginning, the value of G-mean increases rapidly as the number of principal components increases. Then, when the number of principal components reaches 10 or more, the growth rate of G-mean decreases and the maximum value ia obtained at this stage. The optimal number of principal components is 38, 27, and 29 for the Chi-square test, Gini index, and correlation method, respectively. After that, although the number of components continues to increase, the value of G-mean decreases slightly.

According to the Mann-Whitney U test, a statistical analysis is performed between the CAD and non-CAD subjects for the principal components selected by the correlation method. As shown in Fig. 5, the mean values of the selected principal components of the CAD subjects are smaller than those of the non-CAD subjects. All principal components have a significant difference between the CAD and non-CAD subjects, indicating that the selected components are very useful in the classification. The standard deviations of the top 14 features (first row of Fig. 4) that are more relevant to the class label are very large, whereas the standard deviations of the bottom 15 features (second row of Fig. 4) that are less relevant to the class label are small. The principal components with larger correlations have a more significant differences between the CAD and non-CAD subjects. For the statistical analysis of the top 25 features, $p < 0.01$, while for the statistical analysis of the bottom 4 features, $p < 0.05$.

4.2. Results of MFCCs deep learning features and fusion features

In the application of MFCCs, two adjustable parameters exist: the size of each frame and the overlap between adjacent frames. In this work, different combinations are adopted for parameter optimization, and their G-mean values are shown in Fig. 6. The G-mean is the largest when the frame size is 128 ms and the overlap is 25%. Therefore, this parameter combination is used in this study as it is considered to be optimal.

The CNN + MLP model is first applied to the MFCCs images to obtain the classification result. Then, the principal components selected by the Chi-square test, Gini index, and correlation method are fused with the MFCCs deep learning features. The classification results of the MFCCs deep learning featured and the fusion features are presented in Table 8.

As shown in Table 8, an accuracy of 88.09% is achieved by using only the MFCCs deep learning features for classification, exceeding that of the selected principal components in Table 7. After feature fusion, the

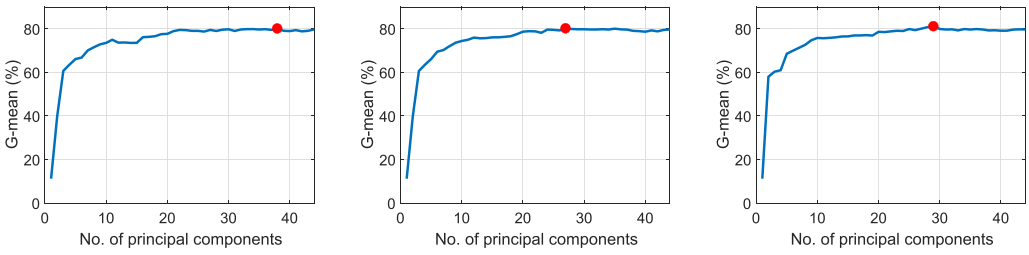

Fig. 4. Variation of G-mean with an increase in the number of principal components using (a) Chi-square test, (b) Gini index, and (c) correlation for feature selection. (Red marker indicates the maximum of G-mean value.). (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

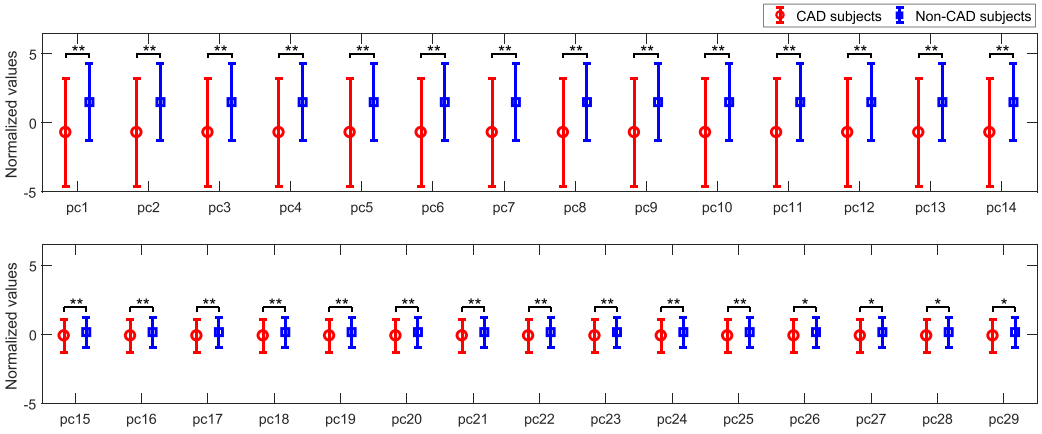

Fig. 5. Error-bar plots of principal components selected by the correlation method for the CAD and non-CAD subjects. Statistically significant differences $p < 0.05$ and $p < 0.01$ are marked with * and ** respectively. The principal components are shown in descending order according to their correlation with the class label.

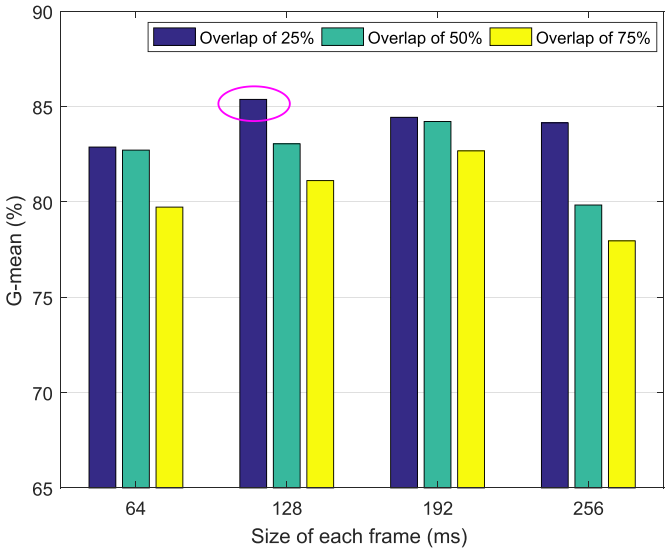

Fig. 6. G-mean of different combinations of frame size and overlap.

accuracy, sensitivity, and G-mean are all improved for the three feature selection methods. However, the specificity decreases when the Chi-square test and Gini index are used, and only increases when the correlation method is used. The best result is obtained by fusing the principal components selected by the correlation method with the MFCCs deep learning features. Compared with the performance of MFCCs deep learning features, the accuracy increases from 88.09% to 90.43%, and the G-mean is improved by 2.81%.

Table 8
Classification performance of the MFCCs deep learning featured and the fusion features.

Features	Acc (%)	Sen (%)	Spe (%)	G-mean (%)
MFCCs deep learning features	88.09 ± 2.92	91.67 ± 3.13	80.27 ± 13.20	85.38 ± 6.18
Fusion features 1	89.31 ± 3.19	93.50 ± 2.81	80.18 ± 11.56	86.30 ± 5.83
Fusion features 2	89.00 ± 3.49	93.17 ± 3.21	79.91 ± 10.29	86.08 ± 5.34
Fusion features 3	**90.43 ± 3.79**	**93.67 ± 3.02**	**83.36 ± 10.07**	**88.19 ± 5.53**

[1] The best performance is marked with bold.
[2] Fusion features 1, 2, and 3 indicate the fusion of principal components (selected by Chi-square test, Gini index, and correlation respectively) and MFCCs deep learning features.

4.3. Comparison with existing studies

Previous studies that have used PCG signals for the diagnosis of CAD are listed in Table 9. The studies all used traditional methods including feature extraction and classification processes. Samanta et al. [28] used PCG signals of 66 subjects and obtained the highest accuracy of 82.6% using an artificial neural network. Makaryus et al. [25] used a relatively larger database containing PCG signals from 161 subjects but only obtained an accuracy of 61.5%. Compared with these studies, we collected

Table 9
Summary of previous studies on the diagnosis of CAD using PCG signals.

Author	Database	Feature type and classifier	Result (%)
Makaryus et al. [25] (2013)	161 subjects (19 CAD, 142 normal)	Microbruit (in frequency range of 400 – 2,700 Hz) Logistic regression	Acc = 61.5 Sen = 89.5 Spe = 57.7
Schmidt et al. [26] (2015)	133 subjects (63 CAD, 70 normal)	Frequency and nonlinear features Multivariate classifier	Acc = 68.4 Sen = 72.0 Spe = 65.2
Banerjee et al. [27] (2017)	25 subjects (10 CAD, 15 normal)	Time and frequency features Support vector machine	Acc = 80.0 Sen = 60.0 Spe = 93.3
Samanta et al. [28] (2019)	66 subjects (29 CAD, 37 normal)	Time and frequency features Artificial neural network	Acc = 82.6 Sen = 85.6 Spe = 79.6
This study	175 subjects (120 CAD, 55 non-CAD)	Extracted six-domain features MFCCs deep learning features Feature fusion framework	**Acc = 90.4** **Sen = 93.7** **Spe = 83.4**

The best result is marked with bold according to accuracy and sensitivity.

PCG signals from 175 subjects. Seven types of features were extracted to mine more potential information, including time, energy, high-order statistics, spectrum, wavelet, entropy, and MFCCs deep learning features. A new feature fusion framework was proposed by combining traditional methods and deep learning techniques. Compared with previous studies, this study achieved higher accuracy and sensitivity, demonstrating the effectiveness of the proposed method.

5. Discussion

In practical applications, one of the main tasks of CAD detection using PCG signals is to extract as many as possible useful features associated with the disease. To this end, this paper proposes a novel feature fusion framework that can take full advantage of both conventional methods and deep learning techniques to achieve more sufficient information mining from PCG signals for CAD diagnosis. Comparing the results of Tables 7 and 8, an accuracy of 88.09% is obtained by using MFCCs images and the CNN, exceeding that of hand-crafted features. This implies that, compared with traditional methods, deep learning techniques have more potential to capture variations caused by the disease. However, deep learning features extracted from MFCCs images focus only on time-frequency characteristics, while the traditionally extracted features can reflect characteristics with multiple types and provide supplementary disease-related information. Therefore, as shown in Table 8, when hand-crafted features and deep learning features are fused for classification, the performance is further enhanced and the accuracy is improved to 90.43%. Fusion features carry more comprehensive and reliable information to differentiate between CAD and non-CAD subjects.

In this study, there are redundancy and correlation among the 110 features extracted from multiple domains. Too many features can interfere with the model learning process and lead to a weak generalization capability. In general, the number of features should be reduced before the classification step. However, directly using the feature selection method to determine the optimal feature subset results in information loss. Therefore, a feature reduction method is needed to replace high-dimensional features with low-dimensional features while retaining useful information. PCA is an effective and commonly used dimensionality reduction method [42–44] and is applied in this study. As shown in Table 7, the classification performance is improved by using PCA to decorrelate the original features. Furthermore, three methods are used to quantitatively evaluate the importance of each principal component in the classification from the perspectives of statistical analysis, impurity, and correlation. The results in Table 7 show that the correlation-based feature selection method performns best. Consistent with this study, in the study of Homsi et al. [8], the correlation-based feature selection was shown to be effective in determining the optimal subset of relevant PCG features.

In terms of time-frequency analysis, the short-time Fourier transform is a frequently used method often combined with the CNN for PCG classification [45–47]. However, the study of Wu et al. [21] pointed out that using MFCCs image as input to the CNN achieved a higher accuracy than the short-time Fourier transform. MFCCs add Mel-frequency and cepstrum analysis to the short-time Fourier transform [39]. Mel-wrapped spectral space is linear for low-frequency components but logarithmic for high-frequency components. This allows MFCCs to operate according to changes in the critical bandwidth of the human ear, thus revealing the acoustic characteristics of audio signals. Moreover, cepstrum analysis can decorrelate the filterbank energies for the purpose of information compression, thereby highlighting those aspects that make a significant contribution to disease detection. The above advantages of MFCCs enable the CNN to extract more useful potential features from images and obtain better performance than the hand-crafted extracted features in this study.

In this study, the feature extraction codes were executed by MATLAB R2016b, and the deep learning codes were compiled in the Keras framework with TensorFlow backend using the Python 3.5 tool. The entire experiment was implemented on a PC with a 3.30 GHz Intel Core i3 CPU, 6 GB RAM, and Windows 10 operating system. From the perspective of computation time and required resources, our proposed feature fusion framework is not very time- or resource-consuming. It takes approximately 9 s to extract 110 features from a 15-s PCG sample on a CPU. Using the CNN to perform a five-fold cross-validation of 3,500 samples takes approximately 9 min on an NVIDIA GeForce GTX 1080Ti GPU. The training time of the feature fusion network for each fold is approximately 2.7 s per epoch, indicating that the proposed method has potential for real-time applications.

Among the studies on automatic PCG classification, the majority have used the PhysioNet/Cinc Challenge 2016 dataset for the recognition of abnormal PCG signals. In this study, the proposed method is also applied to the dataset, and the results obtained are compared with the state-of-the-art studies shown in Table 10. Of the studies based on traditional methods [48–50], Bobillo [48] achieved the highest overall score of 91.9% using time-frequency features as input to the k-nearest neighbor. As for the deep learning-based studies [51–54], Noman [51] and Rubin [52] combined MFCCs and the CNN for PCG classification. Potes [53] used extracted features and decomposed PCG signals to train an AdaBoost and a CNN model, respectively, and integrated the two for classification. Tschannen [54] combined deep features learned by a 1-D CNN with extracted time and frequency features for classification. In contrast to these studies, we extracted different types of features to reflect richer information related to heart states. Considering that a 2-D CNN with MFCCs image as input obtained a higher overall score than that of the 1-D CNN in Noman's study [51], the multi-domain features extracted in this study were fused with deep features learned by a 2-D CNN. The classification results show that the proposed method achieves a higher overall score of 92.4% compared to other methods, demonstrating its effectiveness in detecting abnormal PCG signals.

However, there are still several deficiencies in our study. Recently, Gharehbaghi [55] presented a deep time-growing neural network

Table 10
Comparison of studies on the recognition of abnormal PCG signals using the PhysioNet/CinC Challenge dataset.

Author	Classifier model	Model input	Result (%)
Bobillo[48]	k-nearest neighbor	Time-frequency features	M_{acc} = 91.9 Sen = 94.6 Spe = 89.3
Zabihi[49]	Ensemble of neural networks	Time, frequency features Time-frequency features	M_{acc} = 91.5 Sen = 94.2 Spe = 88.8
Kay[50]	DropConnect neural network	Time, time-frequency features Complexity features	M_{acc} = 84.1 Sen = 84.8 Spe = 83.3
Noman[51]	2-D CNN	MFCCs image	M_{acc} = 88.8 Sen = 86.1 Spe = 91.6
Rubin[52]	2-D CNN	MFCCs image	M_{acc} = 84.8 Sen = 76.5 Spe = 93.1
Potes[53]	Ensemble of AdaBoost and 1-D CNN	Time and time-frequency features Decomposed PCG signal	M_{acc} = 85.0 Sen = 88.0 Spe = 82.0
Tschannen [54]	Feature fusion based on 1-D CNN	Time, frequency features PCG signal	M_{acc} = 87.0 Sen = 90.8 Spe = 83.2
This study	Feature fusion based on 2-D CNN	Time, frequency features Time-frequency, energy features Statistics, entropy features MFCCs image	**M_{acc} = 92.4 Sen = 95.7 Spe = 89.1**

[1] The overall score M_{acc} is an evaluation criteria used in the PhysioNet/CinC Challenge 2016 to measure the balanced performance between the normal and abnormal classes.
[2] The highest overall score M_{acc} is marked with bold.

(DTGNN) that employed a multi-scale learning structure to classify the cyclic time series and effectively retained the dynamic contents in the time series. The DTGNN method was applied to a PCG database and the results showed that it significantly improved the detection accuracy of systolic heart murmurs caused by valvular and septal lesions. Compared with the DTGNN method, our proposed method paid more attention to overcoming the deficiency of deep learning features by fusing them with extracted multi-domain features that provided more information for disease detection. However, in our method, we rarely considered the dynamic contents of the time series as pointed out by Gharehbaghi. If we can develop a proper method to fuse original hand-crafted features and deep learning features with dynamic characteristics, the classification performance may be further enhanced. Moreover, this study only used single-channel PCG signals for CAD detection. Recent studies have shown that using multi-channel PCG signals acquired from different auscultation sites on the chest can provide more information for CAD diagnosis and enhance the classification performance [28].

6. Conclusion

This paper proposed a feature fusion framework that combined hand-crafted features and the deep learning features of PCG to detect CAD. Multi-domain features were extracted from the preprocessed PCG signals, and then reduced and selected. The selected features were fused with the deep learning features extracted from MFCCs images by the CNN, and used as the input for an MLP classifier. The results showed that the fusion features performed better than the hand-crafted and deep learning features, and achieved accuracy, sensitivity, and specificity of 90.43%, 93.67%, and 83.36%, respectively. In comparison with existing studies, the proposed method proved to be very promising for noninvasive CAD detection. In future work, we plan to fuse the features reflecting dynamic contents and utilize multi-channel PCG signals to further improve the detection accuracy of CAD.

Declaration of competing interest

We declare that we have no conflicts of interest.

Acknowledgements

This work was supported by the National Natural Science Foundation of China (Nos. 61471223, 61601263, 61501280) and the Shandong Provincial Jinan Science and Technology Project (Nos. 201816082, 201817001).

References

[1] S.E. Nissen, Limitations of computed tomography coronary angiography, J. Am. Coll. Cardiol. 52 (25) (2008) 2145–2147.
[2] I.R. Hanna, M.E. Silverman, A history of cardiac auscultation and some of its contributors, Am. J. Cardiol. 90 (3) (2002) 259–267.
[3] V. Padmanabhan, J.L. Semmlow, Dynamical analysis of diastolic heart sounds associated with coronary artery disease, Ann. Biomed. Eng. 22 (3) (1994) 264–271.
[4] M. Akay, Harmonic decomposition of diastolic heart sounds associated with coronary artery disease, Signal Process. 41 (1) (1995) 79–90.
[5] J. Semmlow, K. Rahalkar, Acoustic detection of coronary artery disease, Annu. Rev. Biomed. Eng. 9 (2007) 449–469.
[6] G. Eslamizadeh, R. Barati, Heart murmur detection based on wavelet transformation and a synergy between artificial neural network and modified neighbor annealing methods, Artif. Intell. Med. 78 (2017) 23–40.
[7] B.M. Whitaker, P.B. Suresha, C. Liu, G.D. Clifford, D.V. Anderson, Combining sparse coding and time-domain features for heart sound classification, Physiol. Meas. 38 (8) (2017) 1701.
[8] M.N. Homsi, P. Warrick, Ensemble methods with outliers for phonocardiogram classification, Physiol. Meas. 38 (8) (2017) 1631.
[9] A. Yadav, A. Singh, M.K. Dutta, C.M. Travieso, Machine learning-based classification of cardiac diseases from PCG recorded heart sounds, Neural Comput. Appl. (2019) 1–14.
[10] P. Langley, A. Murray, Heart sound classification from unsegmented phonocardiograms, Physiol. Meas. 38 (8) (2017) 1658.
[11] W. Zhang, J. Han, S. Deng, Heart sound classification based on scaled spectrogram and tensor decomposition, Expert Syst. Appl. 84 (2017) 220–231.
[12] F. Safara, S. Doraisamy, A. Azman, A. Jantan, A.R.A. Ramaiah, Multi-level basis selection of wavelet packet decomposition tree for heart sound classification, Comput. Biol. Med. 43 (10) (2013) 1407–1414.
[13] M.E. Karar, S.H. El-Khafif, M.A. El-Brawany, Automated diagnosis of heart sounds using rule-based classification tree, J. Med. Syst. 41 (4) (2017) 60.
[14] J. Li, L. Ke, Q. Du, Classification of heart sounds based on the wavelet fractal and twin support vector machine, Entropy 21 (5) (2019) 472.
[15] M.M. Hassan, M.G.R. Alam, M.Z. Uddin, S. Huda, A. Almogren, G. Fortino, Human emotion recognition using deep belief network architecture, Inf. Fusion 51 (2019) 10–18.
[16] M.Z. Uddin, M.M. Hassan, Activity recognition for cognitive assistance using body sensors data and deep convolutional neural network, IEEE Sensor. J. 19 (19) (2019) 8413–8419.
[17] C. Pan, B. Ni, Y. Xu, X. Yang, Recognition oriented facial image quality assessment via deep convolutional neural network, in: Proceedings of the International Conference on Internet Multimedia Computing and Service, 2016, pp. 160–163.
[18] L. Chen, J. Ren, Y. Hao, X. Hu, The diagnosis for the extrasystole heart sound signals based on the deep learning, J. Med. Image. Health Infomatics 8 (5) (2018) 959–968.

[19] A. Gupta, G. Tang, S. Suresh, Heartfit: an Accurate Platform for Heart Murmur Diagnosis Utilizing Deep Learning, arXiv preprint arXiv:1907.11649.
[20] S.A. Singh, S. Majumder, M. Mishra, Classification of short unsegmented heart sound based on deep learning, in: IEEE International Instrumentation and Measurement Technology Conference (I2MTC), IEEE, 2019, pp. 1–6.
[21] J.M.-T. Wu, M.-H. Tsai, Y.Z. Huang, S.H. Islam, M.M. Hassan, A. Alelaiwi, G. Fortino, Applying an ensemble convolutional neural network with savitzky–golay filter to construct a phonocardiogram prediction model, Appl. Soft Comput. 78 (2019) 29–40.
[22] J.P. Dominguez-Morales, A.F. Jimenez-Fernandez, M.J. Dominguez-Morales, G. Jimenez-Moreno, Deep neural networks for the recognition and classification of heart murmurs using neuromorphic auditory sensors, IEEE Trans. Biomed. Circuit Sys. 12 (1) (2017) 24–34.
[23] W. Han, Z. Yang, J. Lu, S. Xie, Supervised threshold-based heart sound classification algorithm, Physiol. Meas. 39 (11) (2018), 115011.
[24] D.M. Nogueira, C.A. Ferreira, E.F. Gomes, A.M. Jorge, Classifying heart sounds using images of motifs, MFCC and temporal features, J. Med. Syst. 43 (6) (2019) 168.
[25] A.N. Makaryus, J.N. Makaryus, A. Figgatt, D. Mulholland, H. Kushner, J. L. Semmlow, J. Mieres, A.J. Taylor, Utility of an advanced digital electronic stethoscope in the diagnosis of coronary artery disease compared with coronary computed tomographic angiography, Am. J. Cardiol. 111 (6) (2013) 786–792.
[26] S.E. Schmidt, C. Holst-Hansen, J. Hansen, E. Toft, J.J. Struijk, Acoustic features for the identification of coronary artery disease, IEEE (Inst. Electr. Electron. Eng.) Trans. Biomed. Eng. 62 (11) (2015) 2611–2619.
[27] R. Banerjee, A.D. Choudhury, S. Datta, A. Pal, K.M. Mandana, Noninvasive detection of coronary artery disease using PCG and PPG, in: eHealth 360°, Springer International Publishing, 2017, pp. 241–252.
[28] P. Samanta, A. Pathak, K. Mandana, G. Saha, Classification of coronary artery diseased and normal subjects using multi-channel phonocardiogram signal, Biocybernetics Biomed. Eng. 39 (2) (2019) 426–443.
[29] W. Dock, S. Zoneraich, A diastolic murmur arising in a stenosed coronary artery, Am. J. Med. 42 (4) (1967) 617–619.
[30] M. Akay, Y.M. Akay, D. Gauthier, R.G. Paden, W. Pavlicek, F.D. Fortuin, J. P. Sweeney, R.W. Lee, Dynamics of diastolic sounds caused by partially occluded coronary arteries, IEEE (Inst. Electr. Electron. Eng.) Trans. Biomed. Eng. 56 (2) (2008) 513–517.
[31] D.B. Springer, L. Tarassenko, G.D. Clifford, Logistic regression-hsmm-based heart sound segmentation, IEEE (Inst. Electr. Electron. Eng.) Trans. Biomed. Eng. 63 (4) (2015) 822–832.
[32] C. Liu, D. Springer, Q. Li, B. Moody, R.A. Juan, F.J. Chorro, F. Castells, J.M. Roig, I. Silva, A.E. Johnson, et al., An open access database for the evaluation of heart sound algorithms, Physiol. Meas. 37 (12) (2016) 2181–2213.
[33] H. Tang, Z. Dai, Y. Jiang, T. Li, C. Liu, PCG classification using multidomain features and SVM classifier, BioMed Res. Int. 2018 (2018) 1–14.
[34] S.G. Mallat, A theory for multiresolution signal decomposition: the wavelet representation, IEEE Trans. Pattern Anal. Mach. Intell. 11 (7) (1989) 674–693.
[35] L.H. Cherif, S. Debbal, F. Bereksi-Reguig, Choice of the wavelet analyzing in the phonocardiogram signal analysis using the discrete and the packet wavelet transform, Expert Syst. Appl. 37 (2) (2010) 913–918.
[36] J.S. Richman, J.R. Moorman, Physiological time-series analysis using approximate entropy and sample entropy, Am. J. Physiol. Heart Circ. Physiol. 278 (6) (2000) H2039–H2049.
[37] L. Ji, P. Li, K. Li, X. Wang, C. Liu, Analysis of short-term heart rate and diastolic period variability using a refined fuzzy entropy method, Biomed. Eng. Online 14 (1) (2015) 64.
[38] L.I. Smith, A Tutorial on Principal Components Analysis, Cornell University, USA, 2002.
[39] S. Davis, P. Mermelstein, Comparison of parametric representations for monosyllabic word recognition in continuously spoken sentences, IEEE Trans. Acoust. Speech Signal Process. 28 (4) (1980) 357–366.
[40] K. He, X. Zhang, S. Ren, J. Sun, Delving deep into rectifiers: surpassing human-level performance on imagenet classification, in: Proceedings of the IEEE International Conference on Computer Vision, 2015, pp. 1026–1034.
[41] A. Kingmad, A Method for Stochastic Optimization, Arxivepreprint, 2015, p. 6980, arXiv 1412.
[42] E. Kay, A. Agarwal, Dropconnected neural networks trained on time-frequency and inter-beat features for classifying heart sounds, Physiol. Meas. 38 (8) (2017) 1645.
[43] U.R. Acharya, V.K. Sudarshan, J.E. Koh, R.J. Martis, J.H. Tan, S.L. Oh, A. Muhammad, Y. Hagiwara, M.R.K. Mookiah, K.P. Chua, et al., Application of higher-order spectra for the characterization of coronary artery disease using electrocardiogram signals, Biomed. Signal Process Contr. 31 (2017) 31–43.
[44] A.D. Dolatabadi, S.E.Z. Khadem, B.M. Asl, Automated diagnosis of coronary artery disease (CAD) patients using optimized SVM, Comput. Methods Progr. Biomed. 138 (2017) 117–126.
[45] W. Zhang, J. Han, S. Deng, Abnormal heart sound detection using temporal quasi-periodic features and long short-term memory without segmentation, Biomed. Signal Process Contr. 53 (2019) 101560.
[46] W. Zhang, J. Han, Towards heart sound classification without segmentation using convolutional neural network, in: Computing in Cardiology (CinC), IEEE, 2017, pp. 1–4.
[47] T. Nilanon, J. Yao, J. Hao, S. Purushotham, Y. Liu, Normal/abnormal heart sound recordings classification using convolutional neural network, in: Computing in Cardiology Conference (CinC), IEEE, 2016, pp. 585–588.
[48] I.J.D. Bobillo, A tensor approach to heart sound classification, in: Computing in Cardiology Conference (CinC), IEEE, 2016, pp. 629–632.
[49] M. Zabihi, A.B. Rad, S. Kiranyaz, M. Gabbouj, A.K. Katsaggelos, Heart sound anomaly and quality detection using ensemble of neural networks without segmentation, in: Computing in Cardiology Conference (CinC), IEEE, 2016, pp. 613–616.
[50] E. Kay, A. Agarwal, Dropconnected neural network trained with diverse features for classifying heart sounds, in: Computing in Cardiology Conference (CinC), IEEE, 2016, pp. 617–620.
[51] F. Noman, C.-M. Ting, S.-H. Salleh, H. Ombao, Short-segment heart sound classification using an ensemble of deep convolutional neural networks, in: IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), IEEE, 2019, pp. 1318–1322.
[52] J. Rubin, R. Abreu, A. Ganguli, S. Nelaturi, I. Matei, K. Sricharan, Classifying heart sound recordings using deep convolutional neural networks and mel-frequency cepstral coefficients, in: Computing in Cardiology Conference (CinC), IEEE, 2016, pp. 813–816.
[53] C. Potes, S. Parvaneh, A. Rahman, B. Conroy, Ensemble of feature-based and deep learning-based classifiers for detection of abnormal heart sounds, in: Computing in Cardiology Conference (CinC), IEEE, 2016, pp. 621–624.
[54] M. Tschannen, T. Kramer, G. Marti, M. Heinzmann, T. Wiatowski, Heart sound classification using deep structured features, in: Computing in Cardiology Conference (CinC), IEEE, 2016, pp. 565–568.
[55] A. Gharehbaghi, M. Lindén, A deep machine learning method for classifying cyclic time series of biological signals using time-growing neural network, IEEE Trans. Neural Netw. Learn. Sys. 29 (9) (2018) 4102–4115.