A classification method for judging the depth of chest compression based on CNN

Zhao Liang\textsuperscript{1*}, Bao Yu\textsuperscript{2}, Ye Ruidong\textsuperscript{2}, Zhang Aijuan\textsuperscript{2} and Zhang Yu\textsuperscript{2}

\textsuperscript{1}Correspondence: zhaol@cumt.edu.cn
\textsuperscript{2}School of mine, China University of Mining and Technology, Daxue road, 221116 Xuzhou, China

Full list of author information is available at the end of the article

Abstract
The displacement can be calculated based on the integrated value of the acceleration signal waveform obtained by the acceleration sensor or gyroscope. However, this method is not effective in accurate measurement. Although some studies have improved the method of calculating accurate distance values by overcoming the effects of sensor noise or integration delay, the evaluation is still affected by sensor accuracy and environment. However, there are some special displacements, such as the chest compression. The displacement is a reciprocating motion and will return to the starting point again. Therefore, the acceleration waveform changes have obvious characteristics in the two stages from moving to the equilibrium position and returning to the starting point. Therefore, we propose an embedded classification method based on one-dimensional Convolutional Neural Network (CNN), which directly learns from the data of chest compressions and performs the signal formed by the Classification, distinguish the signal waveform under the standard pressing distance, so as to replace the calculation of distance measurement, and is not affected by factors such as pressure occlusion and electromagnetic wave interference, and has certain practical value on site. We tagged 937 presses and collected data from the simulator. The experiment evaluates the proposed network structure, and compares the classification results of the sample data with several CNN networks and Support Vector Machines with different structures in the literature. The results show that with sufficient training, the proposed one dimension CNN method can achieve an accuracy rate of more than 95\%, and balances the accuracy rate and the hardware requirements.

Keywords: Convolutional neural network (CNN); Chest compression classification; Accelerometer sensor application; Cardiopulmonary resuscitation

1 Introduction
Nowadays, more and more sensor devices are used in medical devices to assist doctors in treatment. As a necessary means of first aid measures, chest compressions (CC) have precise requirements for the depth of compressions in first aid. As shown in the latest cardiopulmonary resuscitation (CPR) manual, an adequate CC depth (CCD) is associated with a higher survival rate. The recommended CCD is around 5 cm, while avoiding excessive (greater than 6 cm) or insufficient (lower than 5 cm) CCDs during CPR [1]. However, it is difficult to maintain a high-quality CCD due to potential tension, fatigue, and fear of injury to patients during CPR. To improve the quality of CPR training, it is necessary to create a portable training device to track the rescue process and provide feedback to correct non-standard CC movements.

Some old CPR auxiliary equipment are depends on pressure that have been re-
ported not reliable [2]. At present, various auxiliary equipment have been developed for measuring CCD [3]. The main technologies of these devices are to monitor the process of CC and many methods have been proposed to support it. One of the key problems is to obtain the accuracy displacement of CC. If it is solved, the rescuers or trainees would know whether theirs CCs are valid. Methods for accurately measuring depth include ultrasonic/Impulse Radio-Ultra Wideband (IR-UWB) ranging, laser ranging, infrared ranging, accelerometer sensor, etc.

However, when CPR is performed on a bed, the obtained CCD result may be larger than the actual one due to the additional acceleration effect of the mattress movement [4]. The sloping floor or bed is also affects the accelerometer sensor results [5]. In addition, most current precision equipment is bulky, inconvenient to carry, or expensive, such as [3]. Therefore, one trend of CPR equipment is to obtain higher accuracy under different conditions, and the other trend is that it is cheap, simple and light, and easy to use outdoors. These characteristics of outdoor CPR equipment require the installation of miniature sensors, therefore, acceleration sensors or gyroscopes are very suitable for the needs. Since it is difficult to hold a smartphone when CC is being executed, installing the sensor in Smartwatch may be a good choice.

Analyzing previous studies, we found that most of these studies focused on the accuracy of distance calculation. If we change the way of deciding, we think more about the type of resuscitation compression results, and classify according to whether the compression sampling results are good or bad, and distinguish the compression results that meet the conditions, we can replace the judgment method of the compression distance accuracy. According to common sense, CC is a typical regular reciprocating motion, which is easy to operate in the experiment to obtain acceleration data without too much error, because the sample value of each cycle will be zero again. This provides an opportunity for the practical value of monitoring equipment in the field of emergency cardiopulmonary resuscitation. In term of that, what we should do is to summarize the evaluation results and obtain their relevant classifications. According to the accurate measurement of the length and amplitude characteristics of the waveform formed by the acceleration measurement values, the classification method has the ability to judge the state satisfaction. Thus, we can choose a method for the sampled data classification to solve this problem. This is our motivation.

In this paper, the micro-low-power acceleration sensor is used to trace the movement action of CC. First, the noise is filtered, and then the pulse waveform formed by the pressing is segmented, and then the characteristics of the single pulse waveform are extracted and filtered to find the waveform that meets the filtering condition. Finally, the characteristics of sampling data are analyzed when the object moves at a short distance, and the purpose of accurate classification is realized according to signal features. Our main contributions is to change the method of CCD measurement calculation operation into the method of a waveform evaluation operation. Then, we adapted the maximum alignment of training data according to the feature of CC samples. Finally, we compared the error between our method and traditional accurate measurement and compared the running times of different classification methods and chooses one suitable method for portable devices.
This paper is organized as follows. Section 2 introduces related work. Section 3 introduces the collection and experimental methods of training data, and preprocesses the data to adapt to the training processing of convolutional networks. Section 4 uses LeNet to establish a one-dimensional CNN network model, and discusses the improvement of one-dimensional CNN to improve its performance and precision. In Section 5, we evaluate the size of different filters through some experiments, and discuss and compare the different structure of convolutional networks. The last part introduces the conclusions and opinions of future work.

2 Related Work

In the ranging sensor device, the displacement sensor itself is bulky and difficult to install, and cannot meet the requirements of miniaturization of components, portability, and low power consumption. So as the IR-UWB and laser devices. Ultrasonic / IR-UWB ranging method is based on time of arrival. The distance calculation is to use wireless signals to transmit to the distance and the time it takes in the radar process. Time difference of arrival can provide higher accuracy, but its algorithm is not superior to the quadratic integration method [6]. When the signal is measured, nothing should be placed in front of the Ultrasonic / IR-UWB generator. Sometimes, clothes can make the results inaccurate. Although a radar detecting method based on the deep learning technique demonstrated in [7] can recognize breathing and a fall, It will also be disturbed by clothing during accuracy measurement. Regarding the accelerometer sensor, some studies have shown that the quadratic integration method for micrometer with acceleration sensor for CC depth (CCD) measurement is effective[8] [9].Because accelerometer sensors are cheap and lightweight, in micrometers, the most commonly used compression depth measurement depends on the acceleration sensor.

Most methods for CCD computing based on accelerometer sensors use the quadratic integration method. In the opinion of most studies, the quadratic integration method is prone to error accumulation and the drift caused by numerical integration has an unpredictable trend. Therefore, many error compensation methods based on the integral drift assumption usually fail in practice. The author of [10] believes that there is noise during the sampling process of the accelerometer, which will be propagated into the calculation of the integrated velocity and position. Therefore, they reduced the noise through the Kalman filter. It seems to cut off the effect of noise. However, the time interval they chose is 10ms, which is a bit long for the CC sampling frequency. Moreover, their analysis tool is Matlab, which is difficult to embed in portable devices. The authors of [11] proposed to use flexible pressure sensors in real time, but the device must be connected to a computer to calculate the results.

Physiological signals can also be recognized using artificial intelligence (AI) classification algorithms. AI classifier can effectively learn body monitoring features while at the same time reducing the redundant information of the signal and improving classification efficiency. Deep learning that is one of AI methods help us obtain a higher accuracy on these classification. Latest scientific research referring to deep learning in physiological 1D signal data such as electromyogram (EMG), electrocardiogram (ECG), electroencephalogram (EEG), and electrooculogram (EOG) shows
its high potential [12]. Hou integrates a long short-term memory (LSTM)-based auto-encoder (AE) network with support vector machine (SVM) for ECG arrhythmias classification. In the model, the LSTM-based AE network extract ECG signal features, and the SVM classifier is applied for classifying different ECG arrhythmias signals [13]. The result shows the proposed method has more than 99% accuracy. The author of [14] proposed a methods for the detection of pathological voice from healthy speech based on glottal source information. Two combination method are used, including a combination of convolutional neural network (CNN) and multi-layer perceptron (MLP), and a combination of CNN and LSTM network. Their results are better than that gotten by the best traditional pipeline systems. The other researches shows the same conclusion[15]. The authors of [16] make a single high-density surface electromyography (HD-sEMG) dry electrodes device by a matrix of sensor nodes. A 3-layer CNN with a majority vote on 5 successive inferences is used to recognize 8 hand postures and the accuracy reaches 98.15%. Based on the similar method, Maachi et al [17] placed 18 sensors on patients foot and each signal is sent to the one dimension Convenet of deep neural networks (DNN). Important clinical spatio-temporal gait features, such as swing phase, stance phase and stride time, can be derived from vertical ground reaction force signals to distinguish the Parkinson features. These outstanding works using AI show that it is more effective to use classification methods to judge medical monitoring signals than calculating methods.

3 Methods

3.1 Data Preparation methods

Before established the model, as shown in Figure 1, some preparations including data sampling and data preprocess have to be finished.

3.1.1 Training data obtainment

In order to obtain the normal and abnormal chest compression signal data for the experiment, a chest compression simulation system was designed to collect the acceleration data during the chest compression. The entire compression data acquisition system, as shown in Figure 1 can be divided into four parts: the simulation device, the control module, the measurement module and the sensor module. The simulation device or manikin can simulate different pressure waveforms generated by the compression simulation process. The control module is responsible for data analysis, processing and saving. All the compression data will be recorded in a secure digital memory (SD) card. The measurement module is to obtain the computed distance value for comparing with the results gained by classification methods. The sensor module is responsible for compression information monitoring and measuring when the compression is performed. The sensor module is responsible for the acceleration data collection of the CC process. During the sampling process, the hardware is the low power triaxial acceleration ADXL345 for simulation and bm250 for smart watch.

3.1.2 Data preprocess

After collecting samples, the data should be preprocessed for filtering the noise in sampling data. This article simulates the actual situation of CC during the
data collection process. Each set of data is the acceleration curve collected during continuous compression. However, the data collected by time point is affected by electromagnetic and voltage, and the original data contains a lot of white noise. And each group of experiments includes several compressions that create several waveform pulse signals. It is necessary to find the pulse generated by the pressing from the original waveform, remove the interference signals, and distinguish the obtained waveform to obtain the corresponding acceleration changes that are corresponding to each compression. Therefore, data preprocessing such as noise reduction filtering, pulse recognition, waveform segmentation and fitting is required, as shown in Figure 1.

3.1.3 Denoising filtering

As shown in [18], this is accordance with other mentioned researches. According to the influence of electromagnetic and voltage at the sampling time, the data directly collected by the experimental equipment contains a lot of white noise, and there are many burrs in the original pressing waveform. If it is not reduced, it will seriously interfere with the analysis and processing of the compression data. The filtering method has been tested. Certainly, the effect of single-dimensional Kalman filter is better. However, its operation time and complexity are high, which makes the operation delay on a microcontroller. In the case of small losses, the median filtering and low-pass amplitude-controlled filtering are adapted in this paper.

By analyzing the raw data, the low-pass amplitude-controlled filtering is first used to filter the unsuitable white noise and hammering anomalies before performing waveform analysis. After passing through the low-pass filter, there is a significant correlation between adjacent sampling points, and the narrower the filtering bandwidth, the stronger the correlation. Then, using median filtering through sliding
window. This has the advantage of suppressing periodic interference and has high smoothness.

3.1.4 **CC pulse recognition**

In this case, the gradient and amplitude feature calculation method are used to identify the pulse, and then the sliding window is used to distinguish the waveform of each pressing pulse. Then, waveform verification is necessary. The width of the waveform limits the time consumption of CC and the height limits the strength of CC. The three gradient values of the waveform will obtain a normalized sensing data of CC. The specific process is as follows:

Step1: Establish two sliding windows. One named sliding window A stores the sampled value, and the other called sliding window B is to store the filtered result;

Step 2: Do threshold monitoring on the sliding window B, and determine the threshold value when the static gravity value shows an \( \alpha \) percent change;

Step 3: Cut the wave in the window B. To select the optimal point from window B or A as the starting point and terminate on the pulse end point, which physical meaning is a compression process. In the cutting wave, the hand speed is zero when the compression reaches the lowest point position;

Step 4: Recognize the CC pulse according to three restrictions.

Then a set of standardized waveforms will be obtained. According to this method, pulse identification and waveform segmentation were sequentially performed on 18 sets of experimental data, and finally 975 CC waveforms were obtained.

3.2 **Solution based on 1D CNN Method**

3.2.1 **One dimension LeNet5 CNN model**

In order to do CC results classifying, one dimension convolutional neural networks, a deep learning method, is used for less computational cost than RNN. Two dimension CNNs (2D-CNNs) are now widely used in the machine vision community as a latest technology for many image and video recognition problems. 1D-CNN will perform convolutional calculation on the one dimension signal [19]. The input of our 1D-CNN is an array representing the CPR waveform, which is denoted as \( X \). The network is designed to learn a set of parameters to map the input to the prediction \( Y \) according to a convolutional feature given by Eq. (1):

\[
Y = F_n(X_n|\Theta_n) = h(WX_n + b), \Theta_n = [W, b] 
\]  

(1)

\( X_n \) is a one-dimensional input matrix of \( N \) feature maps, \( W \) is a set of \( N \) one-dimensional kernels that used to extract a set of features from the input values, \( b \) is a bias vector, and \( h \) is an activation function.

For simplicity, we define two classes of the result of CCD, 0 is error and 1 right. Thus, the problem becomes a binary classification problem. Then, the crossentropy is defined in Eq. (2):

\[
L_H = -\frac{1}{N} \sum_{k=1}^{N} [\tilde{y}_k log y_k - (1 - \tilde{y}_k log(1 - y_k)]
\]  

(2)

where \( \tilde{y} \) is the label value of a sample.

The number of parameters of a CNN is directly related to the computational cost.
of training as well as for the need of a large amount of training data. Especially, the recognition model of CCD classification will be used for devices with low computational ability. Therefore, we try to find a simple neural net structure for training. One-dimension signals are usually used in many monitoring cases, and pytorch or tensorflow have a special function for it. We still choose 2D-CNNs to modify for this case because we hope to extend the sensors type in future. We first consider the LeNet-5 net because of its lower cost. The LeNet-5 architecture consists of two sets of convolutional and average pooling layers, followed by a flattening convolutional layer, then two fully-connected (FC) layers and finally a softmax classifier [20]. However, LeNet-5 is for 2D image recognition of handwriting. We set one dimension as 1 so that the model is suitable for one dimension cases, and the model is shown in 2.

The architecture shown in 2 is made of 3 convolution layers with filter length of 5 interlaced with two max-pooling layers that are not drawn, followed by two fully connected layers and an output layer. The stride set one step. The ReLU activation function \( h(x) = \max(x, 0) \) is used for all layers, except for the output layer where a softmax classifier was used to output the posterior probability of each class. Finally, the fully connected layer with 128 neurons and the output layer are used for fault detection and classification. The three convolution layers have 32, 64, and 768 neurons respectively.

The 1D raw CC data of accelerometer sensors are preprocessed before being input into the 1D-CNN classifier for learning. The data preprocessing mentioned above will normalized the data and distinguish each CC plus as a wave signal for 1D-CNN inputting.

3.2.2 Data feature analysis and labelling

In the actual emergency cardiac compression process, it is necessary to continuously press multiple times in a short period of time. It is inevitable that there are
problems such as obstruction and jitter during distance measurement. Moreover, in most cases, it is difficult to achieve high-precision millimeter-level measurement, therefore, it is difficult to obtain a large number of reliability data labels. In order to solve the problem of accurate training set labels, we use high-speed cameras and ultrasonic rangefinder to correct them. With high speed camera recordings, readings can be taken directly on the range scale. Ultrasonic ranging can reach an accuracy 1mm when measuring short distances. An ultrasonic rangefinder is installed inside the emergency care simulator to measure the compression distance. Because of that, ultrasonic rangefinder is only used as a correcting tool.

Figure 3 Data alignment of normalized waveform data. The waveform of the left figure is aligned by the start points and the right is by the maximums point.

After the data preprocessing, as another feature of the sampling, the input data set for CNN has missing values due to the inconsistencies at the endpoints, which can be seen from Figure 3. We adapt the model for different wavelengths and sampling rates in two ways: (1) change the neutral node sizes to accommodate the number of the wave inputs; (2) padding or segmenting each piece to make it fit for the input dimensions of the network. The first one implies modifying the size of the first convolutional layers as well as the number and the dimension of model. Then, we consider the inputs alignment. However, after we align the inputs, the data of the starting point alignment is not easy to observe the discrete condition of the wave of compression signals, then all the waves are converted from the origin alignment to the maximum alignment. In order to let the input data have the same size and prevent the occurrence of missing values, each curve takes the maximum point as the boundary and takes 50 points on the left and 20 points on the right to form a data set. These waves are drawn in the same picture for convenience display, as shown in Figure 3. Now, the length of waves have the same points. Otherwise, zero value is filled.

A preliminary analysis of the data curve containing the label and the maximum alignment graph shows that the wave similarity is high when they are denoting valid CCP. When the gap between the compression distance exceeds 5mm, there is a large jitter in the middle of the compression signals. If the interval of two compressions is too long, the waveform display are abnormal compared with the standard one. There are also obvious correct waves, which are concentrated on the two extreme points of the wave set that are too smooth or too prominent. Therefore, the variance at the two extreme points can be used to measure the degree of discreteness of the data, which may be another important character.
3.2.3 Improvement of 1D-LeNet5 model

After training, the testing result get a good result in our previous simulation [18]. However, the accuracy rate reduced when we used the model on emergency care simulator. We hope to obtain a better model by improving the original one.

The structure of the 1D-CNNs was composed of the basic structure of a convolution layer, the sampling layer, and fully-connected layer. Different CNN structures have different effects on signals. According to these layer types, we try three ways to improve the original model. Firstly, according to [21, 22], a typical observation is that deeper networks offer better performance. Considering the running environment of the CPR program, we add one or two hidden layers that may be feasible with a low cost. If it is proved to be effective, we try to add three or more hidden layers and testing whether the device with low ability could respond in time. Secondly, a large filter size of CNN means large receptive fields obtained by this layer [19]. However, some researches have shown that stacking more little filter may achieve the same goal with computational benefits. If stacking two stacked layers, we should insert nonlinearities between them, which increases the representational power of the CNN network and subsequently, leads to an accuracy increase. We will compare the two methods. Thirdly, if there is large receptive fields in the first convolutional layers, it is assumed that the first layer should have a more global view of the wave signal. Moreover, the electronic noise is non-stationary, i.e. the frequency or spectral contents of the noise is stochastic and like a pulse. Therefore, shorter filters do not provide a general view on the spectral contents of the signal and easy absorbed the noise signal to make indecision. we enlarge the filter of the first layer and compare it with others. In order to avoid the over-fitting, batch normalization should be done after the activation function of each convolution layer.

4 Experiments and Results

In this section, we perform three experiments to compare some parameters of the one-dimensional model. The first one is to compare filters of different sizes, the second one is to compare the different numbers of the network layer and find which is better for low level devices. The last one is to compare the performance of different 1D-CNN methods on the data set. For the sampled data, the network outputs the results of the two classifications through the fully connected layer, which is used to determine whether the sampled data meets the CC criteria. We use the parameters of accuracy (Accuracy, Acc) and F-Score [23] to evaluate the performance of the network classification results. The specific definitions of these parameters are as follows (3):

\[
ACC = \frac{(TP + TN)}{(TP + TN + FP + FN)} \quad (3)
\]

\[
F - Score = 2 \times \frac{TP \times TP}{TP + FP + TP + FN} \quad (4)
\]

Among them, TP is true normal data or actual normal data classified as normal data, and TN is true abnormal data, which represents correct data classified as abnormal data. FP is false normal data, which means abnormal data is classified as normal data, FN is false abnormal data that means actually normal data is
classified as abnormal data. In the statistical analysis of binary classification, the F-score indicates a test’s accuracy. The F-score is the harmonic mean of the precision and recall, where an F-score reaches its optimal value at 1 and worst at 0 [24].

4.1 comparison of different filter size

We have designed several sets of contrast experiments to test the relationship between the feature extraction and the filter size in this study. The experimental results are summarized in Table 1. Table 1 shows that the deep features extracted by these filters with different size exhibit positive and negative recognition rates under different learning loop times, thereby indicating that the extracted features can filter out correct CCD well and adapt to the characteristics of the networks favorably. However, firstly, the results also shows that the model 1d-Lenet5 is easy to fall into local solution with small filter size. when we enlarge the filter size, the Lenet5 can jump out the local point. secondly, the wide filter size is better than the small one on short learning time and recognition accuracy. Sometimes the small size filter net fall into local solution if the initial filter is created with random value.(this should shown in some pics that wide compares to small).

<table>
<thead>
<tr>
<th>1st convolutional filter size</th>
<th>Second filter shape</th>
<th>Iteration</th>
<th>Max/Min/Avg accuracy rate(%)</th>
<th>Max/min/Avg F-score(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>5 1 × 5</td>
<td>500</td>
<td>95/65.83/82.62</td>
<td>95.16/73.55/88.488</td>
<td></td>
</tr>
<tr>
<td>5 1 × 5</td>
<td>1000</td>
<td>92.5/65/87.64</td>
<td>92.91/73.08/88.534</td>
<td></td>
</tr>
<tr>
<td>5 1 × 5</td>
<td>1500</td>
<td>95.83/65.83/88.166</td>
<td>95.93/73.55/89.918</td>
<td></td>
</tr>
<tr>
<td>5 1 × 5</td>
<td>2000</td>
<td>95/91.67/93.334</td>
<td>95.16/91.53/93.466</td>
<td></td>
</tr>
<tr>
<td>7 1 × 5</td>
<td>500</td>
<td>93.33/90.83/0.8262</td>
<td>93.65/90.91/92.27</td>
<td></td>
</tr>
<tr>
<td>21 1 × 5</td>
<td>500</td>
<td>95/93.33/94.166</td>
<td>95.16/93.44/94.208</td>
<td></td>
</tr>
</tbody>
</table>

4.2 comparison of different numbers of CNN layer

To obtain enhanced results, we use the proposed 1D-CNNs and then add some new hidden layers on the original neural networks to form a 6 or 7 layer CNNs. Then, in the CNN structure, the convolution and maxpooling layers are added and the convolution and pool layers alternately appear. The last layers are fully connected to obtain the output. We compare the performance of the network under different layers and hope to find the proper numbers of CNN for embedded system.

The experimental results are presented in Table 2. The results indicate that the increase in the convolution layer has increased the learning time and improved the recognition accuracy of the signal for the samples in this study. However, we could not add too much convolution layers because the CP recognition should be finished in a very short time that is no more than 600ms. A computation of the 1D-LeNet5 consumes about 150ms on our 240MHz embedded device. we compared the running time of one sample recognition in the figure 4. The time is obtained based on the same CPU time. Thereafter, from the figure 4, we can see 1D-AlexNet is not proper for the needs, which means the number of convolution layers is up to 5.

4.3 comparison of different methods

We will compare some other CNN models here, such as Alexnet and Knn. Table 3 shows the average classification accuracy achieved by multiple one-dimensional
Table 2 Performance comparison of different numbers of 1D convolution layer

<table>
<thead>
<tr>
<th>Filter shape</th>
<th>Iteration</th>
<th>Max/Min/Avg accuracy rate(%)</th>
<th>Max/min/Avg F-score(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 × 7.5</td>
<td>500</td>
<td>95/91.67/93.33</td>
<td>95.16/91.53/93.466</td>
</tr>
<tr>
<td>1 × 7,5,3</td>
<td>500</td>
<td>93.33/90.83/0.8262</td>
<td>93.65/90.91/92.27</td>
</tr>
<tr>
<td>1 × 11,7,5,3</td>
<td>300</td>
<td>95/93.33/94.166</td>
<td>95.16/93.44/94.208</td>
</tr>
<tr>
<td>1D-LeNet</td>
<td>2000</td>
<td>95/93.33/94.166</td>
<td>95.16/93.44/94.208</td>
</tr>
<tr>
<td>1D-AlexNet</td>
<td>300</td>
<td>95/93.33/94.166</td>
<td>95.16/93.44/94.208</td>
</tr>
</tbody>
</table>

CNNs and the results achieved by other other state-of-the-art methods described in the literatures. In the worst case involving falling into local optimum, the average accuracy of our proposed one-dimensional CNN with 4 convolution layers is 85.67%. The other of our proposed methods with 5 convolution layers is better, the average accuracy is 94%, and the deviation is only 0.26% in 10 folds.

The proposed chest compression CNN with 4 or 5 convolution layers (CPCNN4 or CPCNN5), the DAENet [25], the GammatoneNet [19] are 1D-CNN, which learn the representation directly from the signals. Heart sounds were extracted by the denoising autoencoder (DAE) algorithm as the input feature of 1D-CNN of DAENet. The periods of heart sound signal is very similar with the compression signal in waveform, and GammatoneNet use 2D and 1D representations of the audio signal as input and have a good performance on 1D signals. Therefore, the two method may be suitable our project.

In Table 3, we list the network structure and filter size of all methods. The proposed method is not structurally more complicated than the same accuracy methods, and has fewer parameters than the latest methods described in the literatures, which means that our method is more suitable for low devices in terms of computational complexity. In the figure 5, we compare the ACC and F-score after the experiments. Our method performs better than the general classification method and integration method in the collected samples, and has the feasibility to be embedded into low devices.
### Table 3 Performance comparison of different numbers of 1D convolution layer

<table>
<thead>
<tr>
<th>Methods</th>
<th>Convolution Layers</th>
<th>Filter shape</th>
</tr>
</thead>
<tbody>
<tr>
<td>Integration</td>
<td>-</td>
<td>Median filter</td>
</tr>
<tr>
<td>Logistic</td>
<td>-</td>
<td>Sigmoid</td>
</tr>
<tr>
<td>SVM</td>
<td>-</td>
<td>RBF</td>
</tr>
<tr>
<td>1D-AlexNet</td>
<td>32<em>64</em>128<em>256</em>1024</td>
<td>1 × 11.5,3,3</td>
</tr>
<tr>
<td>GammatoneNet 19</td>
<td>16<em>16</em>32<em>64</em>128</td>
<td>1 × 16,8,4,2</td>
</tr>
<tr>
<td>CPCNN4</td>
<td>32<em>64</em>128*1152</td>
<td>1 × 7.5,3</td>
</tr>
<tr>
<td>CPCNN5</td>
<td>32<em>64</em>128<em>196</em>896</td>
<td>1 × 11.7,5,3</td>
</tr>
</tbody>
</table>

**Figure 5** Precision comparing in different layers and filters

### 5 Discussion

The experimental results show that the features extracted by the proposed system have obvious discrimination, and the features are more obvious in the amplitude and time series of the sampling signals, which improves the classification. Because the length of the waveform limits the size of the filter, when the filter exceeds half of the waveform length, the learning result is not good on testing. Although a large size of convolution kernel increases the amount of calculation, it is easy to obtain features when the number of convolution layers is small. In Table 2, the numbers of the convolution kernel sizes we selected are relatively prime, and the effect of the prime size filter is better than the case with the filters of multiple convolution kernel size. In Table 3 and our previous study, we can confirm that this 1D-CNN method is more accurate than quadratic integral calculation distance.

With deep networks, as the number of layers increases, the amount of calculation increases slightly, and the time for the duration classification and identification is extended. In some cases, the computing ability of the embedded device may be unsatisfied. In the learning process of our experiments, GammatoneNet and AlexNet and CPCNN5 methods did not fall into local optimum situation. The network architecture consists of 4 or 5 convolutional layers, depending on the processing capabilities of the device. This shows that the probability of these methods falling into local optimum are small. From the perspective of learning time, the CPCNN5 method may support direct learning on the device, which will be tried in later research.
6 Conclusion

This paper proposes a method based on 1D-CNN network classification to replace the CCD measurement and its rationality judgment. By comparing different network performances, the method was evaluated on a data set of 937 compression samples. Through the experimental analysis, the following conclusions can be obtained:

1. The proposed classification method performs better than the quadratic integral measurement with 9.4% higher ACC, and also performs better than SVM and other methods.

2. The modified one-dimensional architecture network is better than the 1D-LeNet method in accuracy and better in calculation time than the 1D-ALexNet method, and can be applied to low-capacity devices.

In summary, the model proposed in this study can effectively improve the recognition accuracy of regular reciprocating motion measurement, and very suitable for portable measurement equipment of regular reciprocating motion, like cardiac compression orthosis. In future work, we will verify the feasibility of this combination and whether it can bring better performance in monitoring data detection. In addition, whether it can be applied to lower-end devices, it is necessary to further investigate to find out how to reduce the hardware requirements of such problems, and may further improve the performance of 1D-CNN.

Competing interests

The authors declare that they have no competing interests.

Abbreviations


Author’s contributions

Liang and Bao and A. Zhang conceptualized and designed the study and write the paper. Y. Zhang, Bao and Ye write the program and analyzed the data. All authors have read and approved the final version of the manuscript to be published.

Acknowledgement

This work was supported by the Project of the National Natural Science Foundation of China (no. 51204185) and the State Key Development Program for Basic Research of Xuzhou City (No. KC17073).

Authors’ information

Zhao liang is currently a Ph.D. candidate in School of mine, China University of Mining and Technology. Since 1998, she has been working as a management staff and in 2008 she received a M.D. in Econometrics from China University of Mining and Technology. Yu Bao is currently Associate Professor in Department of Computer Science and Information Technology, China University of Mining and Technology. He received the PhD degree from Tongji University in 2011, the MEng and BEng degree from China University of mining and Technology. He stayed at the Colorado School of Mines, U.S., in 2016 as a visiting scholar. His research interests include artificial intelligence and distributed system, specifically, he is interested in optimization computing of IoT and security and privacy in wireless networks.

Availability of data and material

The training data supporting the conclusions of this article are not available because it refers to subjects’ personal privacy and patent. We will protect the privacy and patent right of subjects to the maximum extent possible. The model of the subjects are public and attached.

Ethics approval and consent to participate

The rights and interests of the data shown in the paper are public and there is no potential risk to the subjects.

CONSENT FOR PUBLICATION

All the data shown in the paper are not involved personal privacy.
References

Author details

1 School of mine, China University of Mining and Technology, Daxue road, 221116 Xuzhou, China.
2 Department of Computer Science and Technology, China University of Mining and Technology, Daxue, road, 221116 Xuzhou, China.