Computer Science and Information Systems 17(2):445–458 https://doi.org/10.2298/CSIS191229011D

Study of cardiac arrhythmia classification based on convolutional neural network

Yonghui Dai[1], Bo Xu[1], Siyu Yan[2,*], and Jing Xu[2]

[1] Management School, Shanghai University of International Business and Economics, Shanghai 201620, China
daiyonghui@suibe.edu.cn, brianxubo@163.com

[2] School of Information Management and Engineering, Shanghai University of Finance and Economics, Shanghai 200433, China
ysy18956@163.com, 1053517221@qq.com

Abstract. Cardiovascular disease is one of the diseases threatening the human health, and its diagnosis has always been a research hotspot in the medical field. In particular, the diagnosis technology based on ECG (electrocardiogram) signal as an effective method for studying cardiovascular diseases has attracted many scholars' attention. In this paper, Convolutional Neural Network (CNN) is used to study the feature classification of three kinds of ECG signals, which including sinus rhythm (SR), Ventricular Tachycardia (VT) and Ventricular Fibrillation (VF). Specifically, different convolution layer structures and different time intervals are used for ECG signal classification, such as the division of 2-layer and 4-layer convolution layers, the setting of four time periods (1s, 2s, 3s, 10s), etc. by performing the above classification conditions, the best classification results are obtained. The contribution of this paper is mainly in two aspects. On the one hand, the convolution neural network is used to classify the arrhythmia data, and different classification effects are obtained by setting different convolution layers. On the other hand, according to the data characteristics of three kinds of ECG signals, different time periods are designed to optimize the classification performance. The research results provide a reference for the classification of ECG signals and contribute to the research of cardiovascular diseases.

Keywords: Cardiac arrhythmia classification, convolutional neural network, ECG signals, segment size.

1. Introduction

In recent years, with the improvement of information technology and medical technology, the diagnosis technology of some diseases has been improved, which provides a better guarantee for people's health. Among the major diseases that threaten human health, cardiovascular disease has always been at the forefront of easily-occurring diseases and it is extremely threatening. SCD (Sudden Cardiac Death) as a typical cardiovascular disease, it causes many people to face life threatening or even lose their lives. Its cause is closely related to VT (Ventricular Tachycardia) and VF (Ventricular Fibrillation). For example, it is about 390,000 people die of malignant arrhythmias each year in the United States, and it is about 540,000 people die of sudden cardiac death in China each year [19]. When VT signal appears in heart, doctors should observe and diagnose it in time. Otherwise, VT is easy to induce VF. If the patient does not defibrillate in time when VF occurs, it may cause damage to cardiac tissue and function and even death of myocardial tissue [1]. Therefore, the diagnosis of arrhythmia has always been a research hotspot in the medical community. Because the electrocardiogram (ECG) can clearly reflect the health of the heart. When the waveform or value of the ECG signal changes abnormally, it indicates that the body is not very healthy [3]. Therefore, Electrocardiogram-based automatic analysis and diagnosis technology has been adopted by many hospitals as an effective method for determining heart disease. At present, ECG arrhythmia signal classification and recognition are regarded as an important basis for the diagnosis of cardiovascular disease, and this method greatly reduces the workload of doctors, which has important practical value for the timely diagnosis and treatment of heart disease.

Because of the advantages of computers in data calculation, processing, and analysis, many computer applications and analysis tools are used in the medical field. Computer-aided diagnosis is used as an important way of medical diagnosis. For example, medical image recognition based on computer image processing, disease prediction based on computer algorithm, etc. ECG signal reflects objective indicators of people's heart function, and it is an important reference basis for the judgment and diagnosis of cardiovascular and cerebrovascular diseases. The ECG signal has a huge amount of data, and these data are non-linear data, which makes it difficult to use. In order to find rules from these data, it needs a series of data processing to realize. Considering the good adaptability of convolution neural network in nonlinear aspect, this paper uses convolution neural network to study the classification of ECG signal, and studies three kinds of arrhythmia classification of sinus rhythm, ventricular tachycardia and ventricular fibrillation.

This paper is organized as follows. In section 2, related work of cardiac arrhythmia classification is introduced. In section 3, the methodology of convolution neural network technology is shown, which including convolution neural network structure and model for cardiac arrhythmia classification. In section 4, experiment and analysis are illustrated. In section 5, the conclusion of this paper is described.

2. Literature review

In the research of arrhythmia, scholars mainly focus on the preprocessing and classification of ECG data. For example, one-dimensional convolutional neural network was used to study arrhythmia detection, and a six-layer convolutional neural network structure was adopted, which including the number of convolutional layers, pooling layers, and forward fully connected layers are two [8]. The overall accuracy rate of this detection method reached 97%, but the sensitivity of recognition of supraventricular ectopic heartbeat was 60.3%, and the positive predictive value was 63.5%. Guo (2017) et al studied simple convolution neural network for image classification, different parameters were adjusted to achieve good image classification, and they analyzed for different optimization algorithms for the influence of learning rate and optimal parameters on image classification [5]. Thereafter, Some scholars studied five types of the arrhythmias and proposed a new method [17]. In their study, wavelet transform and mean filter were utilized to remove noise in the data preprocessing stage, and the label output by the convolutional neural network was combined with the db6 wavelet coefficient in the feature extraction phase. The functions were normalized before entering libsvm. ECG signals were applied to represent the heartbeat, and The experimental results show that the new method they mentioned is better than the existing method, which is a suitable approach to increase the accuracy of classification. Qiu (2018) et al studied five types of arrhythmia, and they put forward a new method for data pre-processing. In order to decrease the noise, mean filter and wavelet transform were applied, and 250-point signal of ECG were selected to appear the heartbeat, and an accuracy of 98.79% was achieved after the data preprocessing [16].

3. Methodology

3.1. Convolutional Neural Network

Nowadays, deep learning has been successfully applied in various fields. Especially the convolutional neural network, it has been widely used in image recognition and speech recognition, such as image classification, face recognition, speech emotion recognition, video analysis as an important content of deep learning [6][18]. Convolutional neural networks are made up of many neurons connected to each other. After each neuron accepts a linear combination of inputs, it begins with a simple linear weighting, and then adds a nonlinear activation function to each neuron for nonlinear transformation and output, it has a very good performance in terms of picture recognition and classification. [11]. Therefore, if there is a picture of a car, the thing that CNN needs to do is to extract the features on the picture, perform feature recognition, and determine which category and which specific thing the object in the picture belongs to through the established model. The connection between every two neurons represents a weight value, called a weight. Different weights and activation functions result in different outputs of the neural network [21].

Compared with convolution neural network, there are still some problems in full connection neural network [2] [7]. For example, If the image is expanded into vector information, it will lose some spatial information. It has too much parameters, and its efficiency will be affected, which increased the difficulty of training. At the same time, a large number of parameters will also lead to network over-fitting [15][22]. Relatively speaking, convolutional neural networks can solve the above problems well.

The structure of CNN includes three kind of layers, which including input layer, hidden layer and output layer [10]. Every different kinds of neural network have different hidden layers. In this structure, convolutional layer, pooling layer, relu layer and fully connected layer are included by the hidden layers of CNN. The structure is shown as in Fig. 1.

In the CNN structure, there are four main contents, the first content is padding [**?**]. Because the original image will become smaller after being convolved by the filter, the edge count of the image in the convolution is less. At this time, the edge information is easy to lose. In order to solve this problem, we can use the padding method. Before each convolution, we fill in a blank space around the image, so that the image is as large as the original after convolution, and the original edge is also calculated more times [9]. The second content is the step size. In general, the default step size is 1, but in fact, step size can be set to other values. The third content is pooling, which is to extract the main features of a certain area, and reduce the number of parameters to prevent the model from overfitting. When the merge is executed, there is both the maximum pooling and the average value

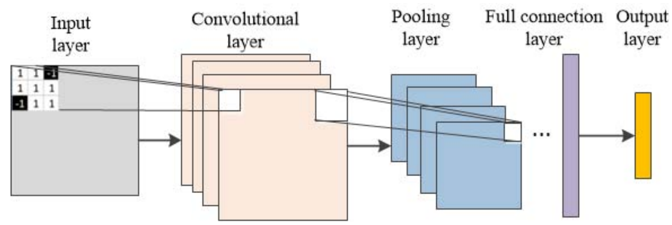

Fig. 1. The structure of CNN

of the area, which is the average pooling. The fourth content is the convolution of multi-channel pictures, color images, usually three channels of RGB, so the input data has three dimensions: (length, width, channel). For example, an image with 28*28 RGB, and the dimension is (28, 28, 3). If the input image is three-dimensional such as (8, 8, 3), and the dimension of the filter is (3, 3, 3) and the last dimension will keep the same dimension as the dimension of input channel. At this time, the convolution is that the elements of the three channels are multiplied and summed.

Therefore, for example, if there is a picture of a car, the thing that CNN needs to do is to extract the features on the picture, perform feature recognition, and determine which category and which specific thing the object in the picture belongs to through the established model.

3.2. The CNN structure for cardiac arrhythmia classification

Because there are many types of arrhythmia, and each type has its own characteristics, which makes the diagnosis of arrhythmia is difficult. For example, the arrhythmia caused by a single irregular heartbeat is called morphological arrhythmia. The arrhythmia caused by a group of irregular heartbeat is called rhythmic arrhythmia. If the data of arrhythmia is detected manually, the whole detection process is not only very tedious and time-consuming, but also human errors may occur in the analysis and detection process. In order to solve the above problems, support vector machine, genetic algorithm, artificial neural network have been used in the detection and recognition of arrhythmia and have achieved good results. In our study, the CNN method was adopted to the classification method of cardiac arrhythmia. In order to get a satisfactory model, the parameters of convolutional neural network model are tried many times. An example of a CNN model's parameters is shown as in Fig. 2.

It can be seen from Fig. 2 that after defining the sequential model, the convolution layer conv1d is added, which is the layer convoluting one-dimensional data in the keras framework. The number of convolution cores is 64, and the number of filters is 1. In the input data, the length of one-dimensional array is 500, and the activation function is relu(). The use of relu() function can accelerate convergence, increase the sparsity of the network, and make the extracted data more representative. The pooling layer uses max-pooling1d, which reduces the dimension of data and prevents over coupling in training. In order to prevent the gradient of neural network from disappearing in the process of back propagation, the batch normalization layer is added to the BN layer to force each input value back to the standard normal distribution of 0 to 1 by some standardized means. The

Layer (type)	Output Shape	Param #
conv1d_27 (Conv1D)	(None, 1, 64)	32064
max_pooling1d_21 (MaxPooling	(None, 1, 4)	0
batch_normalization_18 (Batc	(None, 1, 4)	16
flatten_14 (Flatten)	(None, 4)	0
dense_36 (Dense)	(None, 16)	80
dropout_18 (Dropout)	(None, 16)	0
dense_37 (Dense)	(None, 3)	51

Total params: 32,211
Trainable params: 32,203
Non-trainable params: 8

Fig. 2. Example of the structure parameters of CNN

flatten layer can process the input multidimensional data in one dimension. In the above model, two full connection layers and one dropout layer are added to prevent over fitting during training. The confidence rate of this model is about 92.5%. Its graph converges quickly. However, due to the existence of BN layer and flatten layer, the accuracy will be virtual high. It has the phenomenon of over fitting.

In order to get the best results, during the construction of the convolutional neural network, the subsequent layers are tested one by one. In addition, the number of convolution pooling layers is also changed, its value changed from two to four to verify which is the better result. The structure parameters of the modified convolution neural network are shown in Fig. 3 [20].

Layer (type)	Output Shape	Param #
conv1d_1 (Conv1D)	(None, 291, 100)	1100
max_pooling1d_1 (MaxPooling1	(None, 145, 100)	0
conv1d_2 (Conv1D)	(None, 136, 160)	160160
global_average_pooling1d_1 (	(None, 160)	0
dropout_1 (Dropout)	(None, 160)	0
dense_1 (Dense)	(None, 3)	483

Total params: 161,743
Trainable params: 161,743
Non-trainable params: 0

Fig. 3. The structure parameters of the modified CNN

Convolutional layer: The layer is defined by function Conv1D(), as a one-dimensional convolutional layer, Conv1D has a good recognition performance on sequence data. The number of convolution kernels were defined as 64, the step sizes are 10, and the padding variable is default values of 'same'. In this paper, the activation function of the convolution layer uses the relu() function, because the activation function can make the relationship between input and output close to any function, rather than a simple linear relationship. Moreover, the activation function can accelerate the convergence of neural network, increase the sparsity of neural network, and make the data extracted by deep neural network more representative. The relu function is shown as follows.

$$f(x) = x^+ = max(0, x), \tag{1}$$

Relu function is a piecewise linear function, which changes all negative values to 0, while positive values remain unchanged. Pooling layer: The first layer is defined by function maxpooling1D(), the second pooling layer is defined by global-average-pooling1d. The maximum pooling is used in the paper, which is the most common method. It means that the maximum value will be the selected in a field. The pooling kernel size is 2. The step size is also default. Pooling is actually a process of subsample. Pooling can also improve the fault tolerance of neural networks and prevent overfitting.

Dropout layer: This layer is defined by function dropout. In order to prevent overfitting, this article refers to the research of previous scholars, and set up a dropout layer to discard 30% of the training data, and then put the remaining 70% of the training data into the second dense.

Connected layer and dropout layer: This layer is defined by function Dense(). In this paper, there is only one dense layer. The final output is three categories, which are SR, VT and VF. The overall network structure is shown as in Fig. 4.

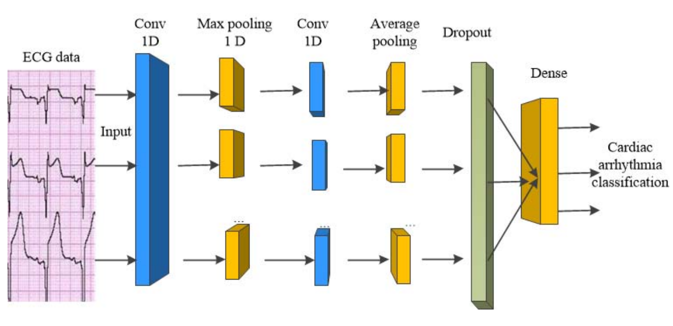

Fig. 4. The CNN structure for Cardiac arrhythmia classification

3.3. Confusion Matrix

Confusion matrix is defined by function confusion_matrix(). After the neural network training is completed, the confusion matrix was added to evaluate the accuracy of the neural network [4]. In this paper, the confusion matrix is a matrix of 3 rows and 3 columns.

Confusion matrix can reflect the accuracy of data classification from different sides. In addition, the confusion matrix can be used to compare the classification results with the average value of the trained neural network, and to compare the calculation to obtain the sensitivity and reliability.

4. Experiment and Analysis

Data pre-processing The data were taken from Physiobank and the extended American Heart Association Database (AHADB). The databases from Physiobank include the European ST-T Database (EDB), the Creighton University Ventricular Tachyarrhythmia Database (CUDB), the MIT-BIH Arrhythmia Database (MITDB), and the MIT-BIH Malignant Ventricular Arrhythmia Database (VFDB). From these databases only records containing examples of VT or VF were used, which is 98 records in total. All records were resampled to 100 Hz sampling rate and processed by a 0.5 Hz high pass FIR filter for removal of baseline wander. The ECG records were further normalized to unit sample variance [13].

Each piece of data that is not originally the same length, segment the data as the same length, and the excess data after division is discarded. In order to explore the influence of the division of different seconds on the result, data is divided into different lengths, which included 100, 200, 300, 1000 (i.e. 1s, 2s, 3s,10s).

At the same time, in order to ensure the same training and testing for each type of data, we separately segment the segmentation data again, 80% of them are used as training data and 20% of then are used as testing data. And in order to ensure the sample balance, VF is the smallest number of samples, therefore, the large number of samples, SR and VT should be down-sampling, and the same amount of data as the VF is selected from SR and VT [12]. Finally, based on data segmentation, the data should be normalized. In this paper, gradient descent iteration method is used when training the model, the feature scaling function can ensure that our different data features are in a similar range, so that the model can converge more quickly during the gradient descent.

In the method of data normalization, the unit variance method is applied [14], and the specific formula is as follows.

$$x[n] = \frac{x[n]}{\sqrt{\frac{\sum_0^n |x[n]|^2}{N}}} \tag{2}$$

After normalizing the data using the unit variance method, the absolute values of all the data are in the interval [0, 1], and put the processed data into the input layer for training, which means the data preprocessing phase ends.

4.1. Data balance method

Class imbalances in data are often encountered when using machine learning algorithms, also known as class skew. For example, if you need to predict whether a patient has a rare disease, however, the percentage of positives in historical data may be only 1%. In this case, it is difficult to learn to get a good classifier, and in this case, the conclusion is usually a big mistake. Some scholars came up with some data balancing method to deal with this situation.

The main methods to deal with data imbalance include that expanding the data set, sampling the data set, generate artificial data samples, different classification algorithms, etc. This paper uses the method of sampling data set to achieve data balance. Sampling includes under-sampling and over-sampling. Under-sampling is to reduce the number of groups with more data, and over-sampling is to increase the number of groups with less data to achieve the balance results. However, the former is usually less used in practical applications, in case of avoiding the data lose. For this reason, new data generation based on existing data is widely adopted, such as synthetic sampling, which is implemented by modifying the data distribution method. In this paper, according to the one-dimensional signal given by the data set and after drawing the time-domain image, the data of SR is much more than that of VT and VF. Moreover, because there is huge difference in data length, data preprocessing is also needed, which is processing samples with unbalanced data. In the data preprocessing stage, the down-sampling was used to balance data, and confusion matrix was utilized for detecting result. Down-sampling is a specific method to resolve sample imbalance, which is to extract the same amount of data from three types of data to ensure sample balance. Confusion matrices are a way to reflect sample results than single precision more closely.

4.2. Experiments and results

In order to classify the cardiac arrhythmia data, this paper tries to use two kinds of CNN architecture to find the best structure and get the best precision results. The first structure includes four convolutional and pooling layers, dropout layer and dense layer. The other includes two convolutional and pooling layers, dropout layer and dense layer. When building the neural network, it is necessary to adjust the neural network itself to obtain the best training effect, which is mainly achieved by controlling different training parameters. For example, Table 1 is the sample data training result of adjusting the different matrix parameter.

Table 1. The result of different matrix parameter

Time	Matrix	SR acc (%)	VF acc (%)	VT acc (%)	Total acc (%)
1s	5000	74.9	28.5	70.0	57.8
1s	10000	83.4	38.7	72.1	64.7
1s	15000	74.3	39.3	68.7	60.7
1s	30000	75.4	41.1	72.1	62.9
Remarks: Batch Size:157, Number of hidden unit:20, Learning rate:0.01					

It can be seen from Table 1 that the learning effect of the whole model is the best when the Matrix is 10,000. While keeping the Matrix, changing the learning Rate to explore the impact of different learning rates on the model. For the adjustment scheme, the increase 5% each time. If the accuracy is lower than the previous one, we will change the range of change to 50%. The results of test are shown as Table 2.

Table 2. The result of different Learning Rate

Time	Learning Rate	SR acc (%)	VF acc (%)	VT acc (%)	Total acc (%)
1s	0.0125	72.3	48.3	82.4	67.6
1s	0.0100	83.4	38.4	72.1	64.6
1s	0.0075	82.6	37.8	71.4	63.9
1s	0.015	\	\	\	\
1s	0.005	\	\	\	\
Remarks: Batch Size:157, Number of hidden unit:20, Matrix:10000					

From the previous parameter adjustment test, the parameters most suitable for the current CNN model are obtained. Next, we try to change the size of the input data to get the best results. The parameters are set to keep the Matrix at 10000 and the learning rate at 0.01. The test is carried out by changing the size of input data to 100, 200, 300, 400. The result is shown as Table 3.

Table 3. The result of different Learning Rate

Time	SR acc (%)	VF acc (%)	VT acc (%)	Total acc (%)
1s	83.4	38.7	72.1	64.7
2s	75.5	36.1	67.6	59.7
3s	62.9	38.6	65.6	55.7
10s	45.5	35.6	71.2	50.8
Average	66.8	37.3	69.1	57.7
Remarks: Batch Size:157, Number of hidden unit:20, Matrix:10000				

At the first time, the first structure includes four convolutional pooling layers, dropout layer and dense layer. Another one includes two convolutional pooling layers, dropout layer and dense layer. Table 4 is the accuracy of first structure including stacking four convolutional max pooling layers.

Table 4. The result of epoch 50 with four convolutional pooling layers

Time	SR acc (%)	VF acc (%)	VT acc (%)	Total acc (%)
1s	70.8	65.7	32.4	56.3
2s	67.7	63.2	30.6	53.8
3s	65.4	56.8	29.7	50.6
10s	30.9	28.3	10.8	23.3
Average	58.7	53.5	25.9	46.1
Remarks: Batch Size:128, epoch: 50				

Obviously, it shows that the average accuracy is very low, the classification performance is not good enough. The accuracy is around from 20% to 60%, which is not a perfect result. However, when the segment size is 1s, the sensitivity is much higher than the segment size is 10s. For example, the SR sensitivity dropped from 70.8% to 30.9%.

Therefore, it is obvious that the segment size will affect the accuracy of the final result. When the segment size is small, the accuracy will perform better. When the neural network was constructed, it is necessary to adjust the neural network itself to obtain the best training effect. By controlling different training parameters, we can obtain the accuracy of each type under different training levels, and then choose the optimal one. The final result is show as Table 5.

Table 5. The result of different epoch

Time	Epoch	SR acc (%)	VF acc (%)	VT acc (%)	Total acc (%)
1s	25	85.3	60.4	82.2	75.9
1s	50	90.7	81.7	92.0	88.1
1s	100	91.7	78.1	92.4	87.4
1s	300	86.6	71.4	79.3	79.1
Average		88.6	72.9	86.5	82.6
Remarks: Batch Size:128					

By adjusting the comparison of the number of epochs. When the epoch is 25, the model didn¡¯t converge apparently. As the increase of the epoch, the model gradually converged, the total accuracy also grows up, until the accuracy has reached up the best result, which is 88.1%. When epoch is 300, the model is overfitting, therefore, the accuracy is 79.1% at this time, which is lower than the former one. Next, keep the batch size is 128, epoch is 50, and change the segment size. Then change the segment size from 1s to 3s, and 10s. The result of epoch 50 is show as the Table 6.

Table 6. The result of epoch 50 with two convolutional pooling layers

Time	SR acc (%)	VF acc (%)	VT acc (%)	Total acc (%)
1s	90.7	81.7	92.0	88.1
2s	92.2	77.3	93.0	87.5
3s	90.1	82.0	89.7	87.3
10s	45.5	35.6	71.2	50.8
Remarks: Batch Size:128, epoch: 50				

When the segment size was changed, and the data was put into the neural network for training. According to the different data length from 1s to 3s, the accuracy changed from 88.1% to 87.3%, and it doesn¡¯t have a dramatic drop. So, the segment size was set to 10s, the sensitivity had a great decrease. So, it can be concluded that the segment size will affect the model training effect. The model performance is better when the segment size is smaller. As far as the results of training model are concerned, the accuracy of identifying SR and VT are higher than those of VF. It can be seen from the ECG signal of these three in time domain, the signal of SR and VT are more regular and more representative, while VF are more chaotic. This is the reason that the sensitivity of VF is not as good as those two, because it is not easy to extract features and learn for neural networks. In addition, in this paper, two different kind of CNN architectures were established, one of the structures is two convolutional-pooling layer model, another is four convolutional-pooling layers. Compared to Table 3 and Table 5, the average accuracy in table 3 is approximately 50%, which is apparently lower than Table 5. Therefore, the second structure performed better than first structure. It is because the amount of data is limited, if I put the limited data into a complex model, it may easier lead to over-fitting. So, the model was changed into the present model, using the convolutional neural network to implement it, and the accuracy and sensitivity of the result is more stable than former model. The result was shown as in Fig. 5.

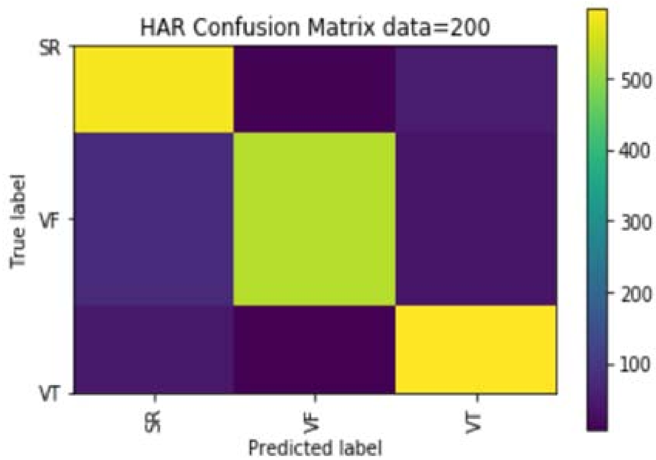

Fig. 5. The confusion matrix

5. Conclusions

In order to study cardiac arrhythmia classification, two CNN structures are used in our study, one is the four-convolutional pooling layer architecture, the other is the two-convolutional pooling layer architecture. From the results of Table 3 and Table 4, CNN with two convolutional pooling layers is better, its average recognition rate is 57.7%, which is higher than 46.1% of the four convolutional pooling layers. In addition, different parameters of the CNN are tested, including the parameters of matrix, learning rate, epoch and segment size. According to the test results in Table 1, when the matrix is 10000, the sensitivity of SR reaches 83.4%, and the total accuracy is 64.7%, which is much higher than the sensitivity when the matrix is 5000 or 15000. Therefore, when the matrix is 10000, the performance of the model is better. According to the test results in Table 2, when the learning rate is 0.0125, the sensitivity of SR is the highest, 83.4%, and the total accuracy is the highest, 67.6%. However, when the learning rate is 0.015 or 0.005, the model does not converge and cannot be classified.

The classification of arrhythmias has always been a research hotspot in the medical field. Before the application of CNN, the classification of heart rate data is a difficult issue. There are two main reasons, one of them is the amount of data to be processed is too large, resulting in high cost and low efficiency. In addition, it is difficult to retain the original features in the process of digitization, resulting in low accuracy of data processing. In order to deal with the issues, two convolutional pooling layers are used to classify the ECG signals, and four time periods (1s, 2s, 3s, 10s) are set and compared to get the best classification results. The research of this paper provides a reference for the classification of ECG signals and is helpful for the research of cardiovascular diseases. In the future research, more sample parameters and continuous optimization of convolutional layer are worth further exploration.

Acknowledgements. This work is supported by the project of Shanghai Philosophy and Social Sciences Plan (No. 2018BGL023).

References

1. Acharya, U.R., Fujita, H., Lih, O.S., Hagiwara, Y., Tan, J.H., Adam, M.: Automated detection of arrhythmias using different intervals of tachycardia ecg segments with convolutional neural network. Information sciences 405, 81–90 (2017)
2. Aljarah, I., Faris, H., Mirjalili, S.: Optimizing connection weights in neural networks using the whale optimization algorithm. Soft Computing 22(1), 1–15 (2016)
3. Benitez, D., Gaydecki, P.A., Zaidi, A., Fitzpatrick, A.P.: The use of the hilbert transform in. ecg signal analysis. Computers in biology and medicine 31(5), 399–406 (2001)
4. Deng, S., Tian, Y., Hu, X., Wei, P., Qin, M.: Application of new advanced cnn structure with adaptive thresholds to color edge detection. Communications in Nonlinear Science and Numerical Simulation 17(5), 1637–1648 (2012)
5. Guo, T.M., Dong, J.W., Li, H.J., Gao, Y.X.: Simple convolutional neural network on image classification. In: IEEE 2nd International Conference on Big Data Analysis (ICBDA). pp. 721–724 (2017)
6. Hong, S., Kwak, S., Han, B.: Weakly supervised learning with deep convolutional neural networks for semantic segmentation: Understanding semantic layout of images with minimum human supervision. IEEE Signal Processing Magazine 34(6), 39–49 (2017)
7. Kapanova, K.G., Dimov, I., Sellier, J.: A genetic approach to automatic neural network architecture optimization. Neural Computing and Applications 39, 1481¨C1492 (2016)
8. Kiranyaz, S., Ince, T., Gabbouj, M.: Real-time patient-specific ecg classification by 1-d convolutional neural networks. IEEE Transactions on Biomedical Engineering 63(3), 664–675 (2016)
9. Limonova, E., Sheshkus, A., Ivanova, A., Nikolaev, D.: Convolutional neural network structure transformations for complexity reduction and speed improvement. Pattern Recognition & Image Analysis 28(1), 24–33 (2018)
10. Liu, H., Li, C., Xu, L.: Dimension reduction and classification of hyperspectral images based on neural network sensitivity analysis and multi-instance learning. Computer Science and Information Systems 16(2), 443–468 (2019)
11. Meng, H.H., Zhang, Y.: Classification of electrocardiogram signals with deep belief networks. In: IEEE International Conference on Computational Science and Engineering (CSE). pp. 7–12 (2015)
12. Mesin, L.: Heartbeat monitoring from adaptively down-sampled electrocardiogram. Computers in Biology and Medicine 84, 217–225 (2017)

13. Mohan, N., Sachin, K.S., Poornachandran, P., Soman, K.: Modified variational mode decomposition for power line interference removal in ecg signals. International Journal of Electrical and Computer Engineering 6(1), 151–159 (2016)
14. Montminy, D.P., Baldwin, R.O., Temple, M.A., Laspe, E.D.: Improving cross-device attacks using zero-mean unit-variance normalization. Journal of cryptographic engineering 3(2), 99–110 (2013)
15. Piotrowski, A.P., Napiorkowski, J.J.: A comparison of methods to avoid overfitting in neural networks training in the case of catchment runoff modelling. Journal of Hydrology Amsterdam 476(1), 97–111 (2013)
16. Qiu, L., Li, W., Cai, W., Zhang, M., Zhu, W., Wang, L.: Heartbeat classification using convolution neural network and wavelet transform to extract features. In: IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computing, Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People and Smart City Innovation. pp. 139–143 (2018)
17. Salem, M., Taheri, S., Yuan, J.: Ecg arrhythmia classification using transfer. learning from 2-dimensional deep cnn features. In: IEEE Biomedical Circuits and Systems Conference (BioCAS). pp. 1–4 (2018)
18. Vikramjit, M., Ganesh, S., Hosung, N., Carol, E.W., Elliot, S., Mark, T.: Hybrid convolutional neural networks for articulatory and acoustic information based speech recognition. Speech Communication 89, 103–112 (2017)
19. Wei, C., Jia, Z.H., Yuan, G.Q., Wu, Y.L.: Integration and regulation- new strategy of arrhythmia drug intervention from ”anti¡ªheart rhythm” to ”regulate heart rhythm”. Chinese Journal of Cardiac Arrhythmias 18(1), 76–78 (2014)
20. Yan, S.Y. (ed.): Cardiac arrhythmia classification on convolution neural network. King’s College London (2019)
21. Yu, S.N., Chen, Y.H.: Noise-tolerant electrocardiogram beat classification based on higher order statistics of subband components. Artificial intelligence in medicine 46(2), 165–178 (2009)
22. Zhou, T., Hou, M.Z., Liu, C.H.: Forecasting stock index with multi-objective optimization model based on optimized neural network architecture avoiding overfitting. Computer Science and Information Systems 15(1), 211–236 (2018)

Yonghui Dai is currently a lecturer at the Management School, Shanghai University of International Business and Economics, China. He received his Ph.D. in Management Science and Engineering from Shanghai University of Finance and Economics, China in 2016. His current research interests include management information systems, affective computing and artificial intelligence. His works have appeared in international journals more than thirty papers. Contact him at daiyonghui@suibe.edu.cn.

Bo Xu is a professor at the Management School, Shanghai University of International Business and Economics, China. He received his Ph.D. in Management Science from Fudan University, China in 1997. His interests include strategic management and educational technology. Contact him at brianxubo@163.com.

Siyu Yan is the corresponding author of this paper, she is a Ph.D. candidate at School of Information Management and Engineering at Shanghai University of Finance and Economics. She received her MS in King’s College London. Her research interests include data mining and big data analytics. Contact her at ysy18956@163.com.

Jing Xu is a Ph.D candidate at School of Information Management and Engineering at Shanghai University of Finance and Economics. She received her MA. in School of Business at Shanghai University of Finance and Economics, China, in 2013. Her research interests include personalized recommendation. E-mail: 1053517221@qq.com.

Received: December 29, 2019; Accepted: May 5, 2020.