Automatic Arrhythmia Detection Based on Convolutional Neural Networks

Zhong Liu[1, 2], Xinan Wang[1, *], Kuntao Lu[1] and David Su[3]

Abstract: ECG signal is of great importance in the clinical diagnosis of various heart diseases. The abnormal origin or conduction of excitation is the electrophysiological mechanism leading to arrhythmia, but the type and frequency of arrhythmia is an important indicator reflecting the stability of cardiac electrical activity. In clinical practice, arrhythmic signals can be classified according to the origin of excitation, the frequency of excitation, or the transmission of excitation. Traditional heart disease diagnosis depends on doctors, and it is influenced by doctors' professional skills and the department's specialty. ECG signal has the characteristics of weak signal, low frequency, large variation, and easy to be interfered. In this investigation, an ECG anomaly automatic classification system based on the convolutional neural network is proposed. The training sets of the convolutional neural network are ECG beats extracted from the MIT-BIH database as training sets. A 36-layer convolutional neural network (CNN) is trained based on Caffe framework to classify ECG signals automatically. The experimental results show that it can reach or even exceed the level of a senior cardiologist in judging three diseases: FIB, AFL and IVR.

1 Introduction

Arrhythmia is a common clinical cardiac abnormality, not only relates to cardiovascular disease, but also relates to many other diseases and occur in a few healthy people. The type and frequency of arrhythmia are essential indicators of cardiac stability [Schwartz and Menotti (2005)]. Arrhythmia affects the normal synchronous systolic timing of the heart, reduces the efficiency of cardiac pumping, and threatens the safety of human life. Therefore, timely and correct diagnosis of arrhythmia is of great significance [Li, Zheng and Tai (1995)]. ECG is a macroscopic record of depolarization and repolarization of cardiac cells, which objectively reflects the physiological status of various parts of the heart to a certain extent, therefore it is of great significance in clinical medicine. In clinical practice, arrhythmic signals can be classified according to the origin of excitation, the frequency of excitation, or the transmission of excitation. ECG arrhythmia detection

---

[1] Peking University Shenzhen Graduate School, Shenzhen, 518055, China.

[2] School of Electronics Engineering and Computer Science, Peking University, Beijing, 100871, China.

[3] iFutureLab Inc. 955 Alma St., Suite B, CA 94301, Palo Alto.

[*] Corresponding Author: Xinan Wang. Email: anxinwang@pku.edu.cn.

is essential for the early diagnosis of heart disease.

1. Power frequency interference: Human Distributed Capacitance, which is caused by power frequency and magnetic field belong to the interference sink of electrode lead circuit, produces interference confluence, and it has a frequency of 50 Hz (or 60 Hz) power frequency and its amplitude is low, it is showed by the electrocardiogram regularity of fine wavelet ripple. This interference often obscures the small transitions in the original ECG and affects the electrocardiogram diagnosis [Shirbani and Setarehdan (2013)].

2. Baseline drift [Chouhan and Mehta (2007)]: When the body breathes, the organs and tissues in the thoracic cavity will change to a certain extent. When the electrode is not fixed very well, it will affect the amplitude and shape of the ECG waveform recorded on the body surface. Its frequency is generally lower than 1 Hz, showing that the baseline drifts periodically with respiration. It is very difficult to analyze and recognize the waveform of each part of ECG, especially the measurement of ST segment in the waveform, which has important diagnostic value for myocardial ischemia and myocardial infarction.

3. EMG interference: EMG interference is mV class interference caused by human activities and muscle tension, its wide frequency range performs like irregular fine wavelet lines on the electrocardiogram, which conceals the original ECG waveform in the small error, making the electrocardiogram blurred or distorted, and difficult to identify and diagnose [Wang, Xu and He (2015)].

Therefore, it is also very difficult to detect arrhythmia by automation.

At present, SVM [Raj and Ray (2017)] is used to classify arrhythmias, while Kiranyaz et al. [Kiranyaz, Ince and Gabbouj (2016); Kiranyaz, Ince and Hamila (2015)], Acharya [Martis, Acharya and Min (2013); Acharya, Fujita and Lih (2017] and others use convolutional neural network to extract and classify arrhythmias. The classification accuracy of some arrhythmias is good.

In this study, a new depth learning algorithm based on convolutional neural network was proposed to classify arrhythmias, the algorithm effectively solves the above problems and shows good performance in arrhythmia classification.

In this research, the MIT-BIH arrhythmia database were provided by the Massachusetts Institute of Technology in the training. It can be obtained from Physio Net. It is recognized as one of the standard ECG databases in the world and has been widely used. ECG in the MIT-BIH Arrhythmia Database consist of a set of over 4000 long-term Holter recordings obtained by the Beth Israel Hospital Arrhythmia Laboratory. It is collected from 47 clinical patients, there are 48 groups sets of annotated ECG recording data. The 48 records are all slightly over 30 minutes long. The 25 male subjects were from 32 year-old to 89 year-old, and the 22 women were from 23 year-old to 89 year-old. (Records 201 and 202 were from the same male subject.) Sampling rate is 360 Hz and sampling accuracy is 11 bits by 0.1 Hz-100 Hz bandpass filter [https://www.physionet.org/physiobank/database/html/mitdbdir/mitdbdir.htm].

In this investigation, a 36-layer convolutional neural network depth learning algorithm based on Cafe depth learning framework is used to implement ECG classification (as shown in the Fig. 1); MIT-BIH database provided by MIT is used in training; access to data from MIT-BIH database by MATLAB ; using the adaptive filtering to reduce the

noise; QRS wave group detection is realized by wavelet transform; and The ECG rhythm is intercepted by fixed windows (R N or N=60). The authors compared the judgments of cardiologists. The method proposed in this article can reach or even exceed the level of cardiologists in the classification of three abnormalities: AFIB, AFL and IVR.

However, arrhythmia testing usually requires patients to wear a dynamic electrocardiograph and other equipment to record their ECG for a long time, usually 24 hours or more. Such a large amount of data is challenging for doctors to diagnose and analyze in a short time, and there is also the possibility of misjudgment because doctors usually judge the morphological changes of ECG by observation. Therefore, the automatic detection system for arrhythmia detection will be an important and much more effective method. Because different patients have significant differences in ECG, and even the same patient's ECG itself also has more significant differences. In addition, ECG is also susceptible to the following disturbances.
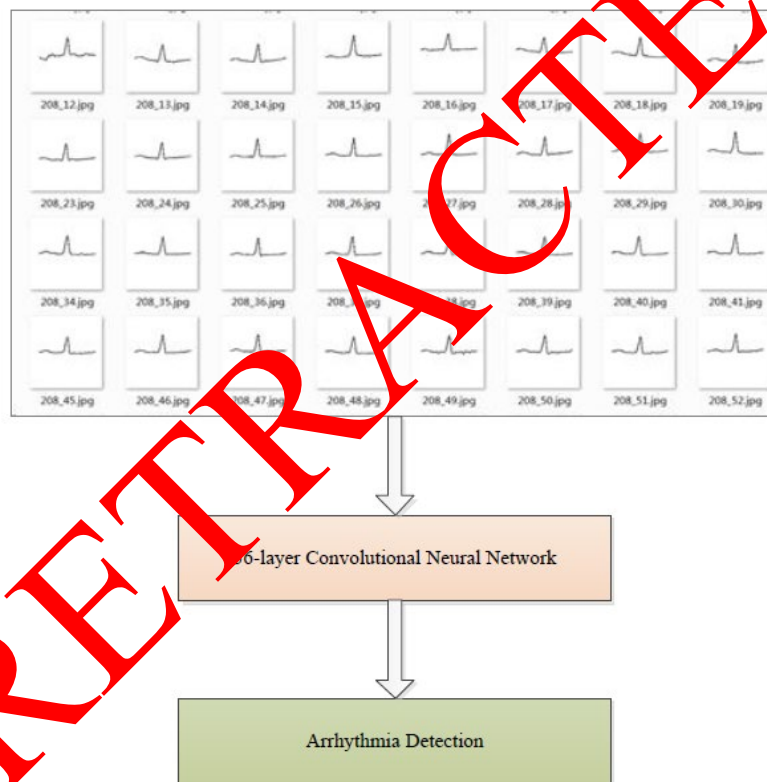

Figure 1: Arrhythmia classification detection based on convolutional neural network

2 Model & data

2.1 Design framework

According to Fig. 2, a 36-layer convolutional neural network depth learning algorithm based on Cafe depth learning framework was proposed in this study through reading MIT-BIH database data, data are preprocessed, data sets are constructed, and classification models are

trained according to the trained model. To further explain, data pre-processing includes using adaptive filtering for noise reduction; data set construction includes constructing training set and test set; training is in the framework of caffe, then using random initial weight, training from scratch; finally, ECG arrhythmias are classified by the trained model.
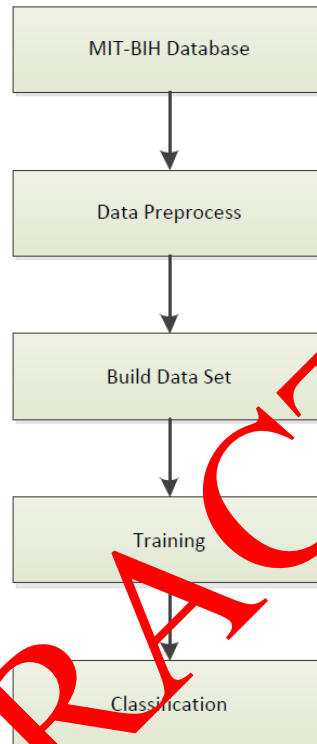

Figure 2: Design of the experimental framework

2.2 Data preprocessing

Firstly, reading the contents of the MIT data record format file and transforming it into JPG picture. MIT-BIH database is read through MATLAB. There are 48 record groups. Each group records two channels, MLII places usually in the first channel, but some others not. For convenience, only the record group of MLII in the first channel is selected.

Due to the performance of cardiac electrical activity on the human body surface, ECG signal is generally weak, with its amplitude of 10_V~5_V, and its frequency of 0.05~100 Hz. It is very vulnerable to the influence of the external environment. In the process of acquisition, amplification and transmission, there will be a lot of interference coupled to the ECG signal, which greatly reduces the signal-to-noise ratio of ECG signal and will have a great influence on the reliability of anomaly detection of the rhythm [Shirbani and Setarehdan (2013)]. These interfering signals mainly include power frequency interference, baseline drift and EMG interference. Therefore, the proposed method combines morphological filtering [Gang (1999); Lulu, Lin and Yuliang (2012)] and

wavelet transform [Sahambi, Tandon and Bhatt (1997); Benzid, Marir and Boussaad (2003); Saritha, Sukanya and Murthy (2008)] to filter low-frequency noise and high-frequency noise. ECG signal preprocessing method is shown in Fig. 3.
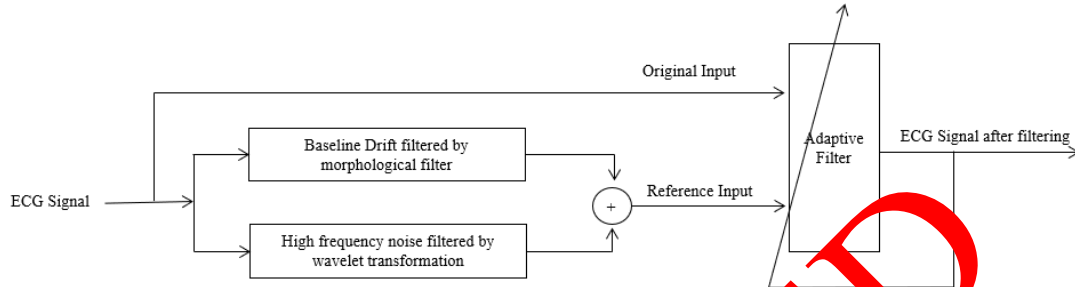
Figure 3: ECG data preprocessing

2.3 Data set construction

2.3.1 Introduction of MIT-BIH database

To save storage space, the MIT-BIH database uses a custom unique file format. In 48 sets of ECG records, each set of records is composed of three files, that are header file (.hea), data file (.dat) and annotation file (.atr). The three files are uniformly numbered. The first 24 ECG records were numbered from 100 to 124, and the last 24 ECG records were numbered from 200 to 234. Some of them were not used. The complete original data file is shown below (Fig. 4).

Figure 4: MIT-BIH data reading

For example, the filenames of the three files recorded by ECG 100 are: 100. hea, 100. dat, and 100. atr. The three file formats are as follows:

1. Header file (.Hea) consists of one or more lines of ASCII codeword. Recorded information includes file name, lead number, sampling rate, data points, gain,

storage format, etc., the header files with "#" at the beginning of the annotation line are generally used to explain the patient's medication situation.

2. Data file (.dat), data stored in binary, two data per three bytes, the length of data is 12 bits. MIT-BIH database data storage formats are Format 8, Format 16 and Format 80, Format 212, Format 310 and so on. The arrhythmia database used in this study utilizes Format212 format uniformly.

3. Annotation file (.atr) is stored in binary form and the diagnostic information of ECG experts is recorded on the corresponding ECG signals. There are two main formats, MIT format and AHA format. The length of each comment of MIT format occupies an even number of bytes; AHA format with 16 bytes of space. MIT format are used in the arrhythmia database in this paper. The MIT format is stored in such a way that the first byte of the first two bytes of each comment unit is the least significant bit, the upper 6 bits of the 16 bits represent the comment type code, and the remaining 10 bits indicate the time at which the comment point occurred or auxiliary information.

In this investigation, we use the MATLAB program provided by the official website to extract R-wave pictures, so as to only select the data with MLII leads, and to exclude 102 and 104 of 48 groups without MLII leads, and 114 of MLII signals in the second channel (easy to error). Finally, the total number of selected signal groups is 45. The specified resolution is 256*256, and one of the R waves of data 100th is shown in Fig. 5:
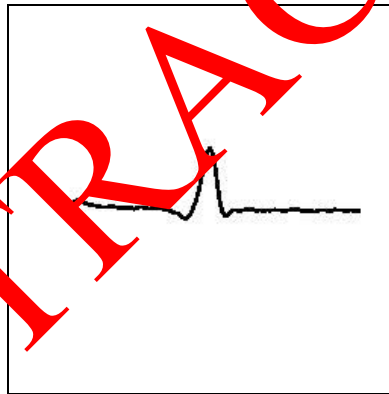

Figure 5: R wave sample plot

2.3.2 Sample graph extraction

The location of the main wave R peak of the QRS wave group is the key step of the ECG signal waveform detection. In this investigation, we locate R wave [Narayana and Rao (2011)] by detecting the modulus maxima of R waves. At the peak of R wave, 60 points (121 data points) are taken from left and right, and a sample image is generated with a size of 256*256. The curve is black and the base image is white. The sample plots extracted through this method are shown in Fig. 6:

Figure 6: extracted sample plots

2.3.3 Construction of training set and test set

Data sets can be categorized into training set train and test set.

2.4 Model

The human brain is a highly complex information processing system composed of a large number of biological neurons through the interconnection. Artificial Neural Network (ANN) is similar to the human brain in two aspects: one for acquiring knowledge from the external environment through the learning process; the other for using the internal neurons (corresponding to the weight of the network) to store the acquired knowledge information. Fig. 7 is the schematic diagram of artificial neural network.
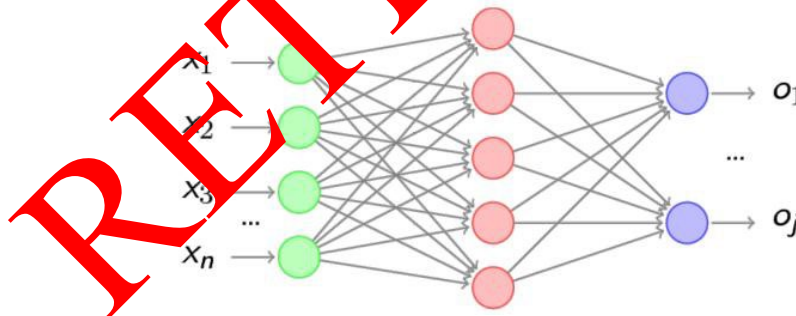
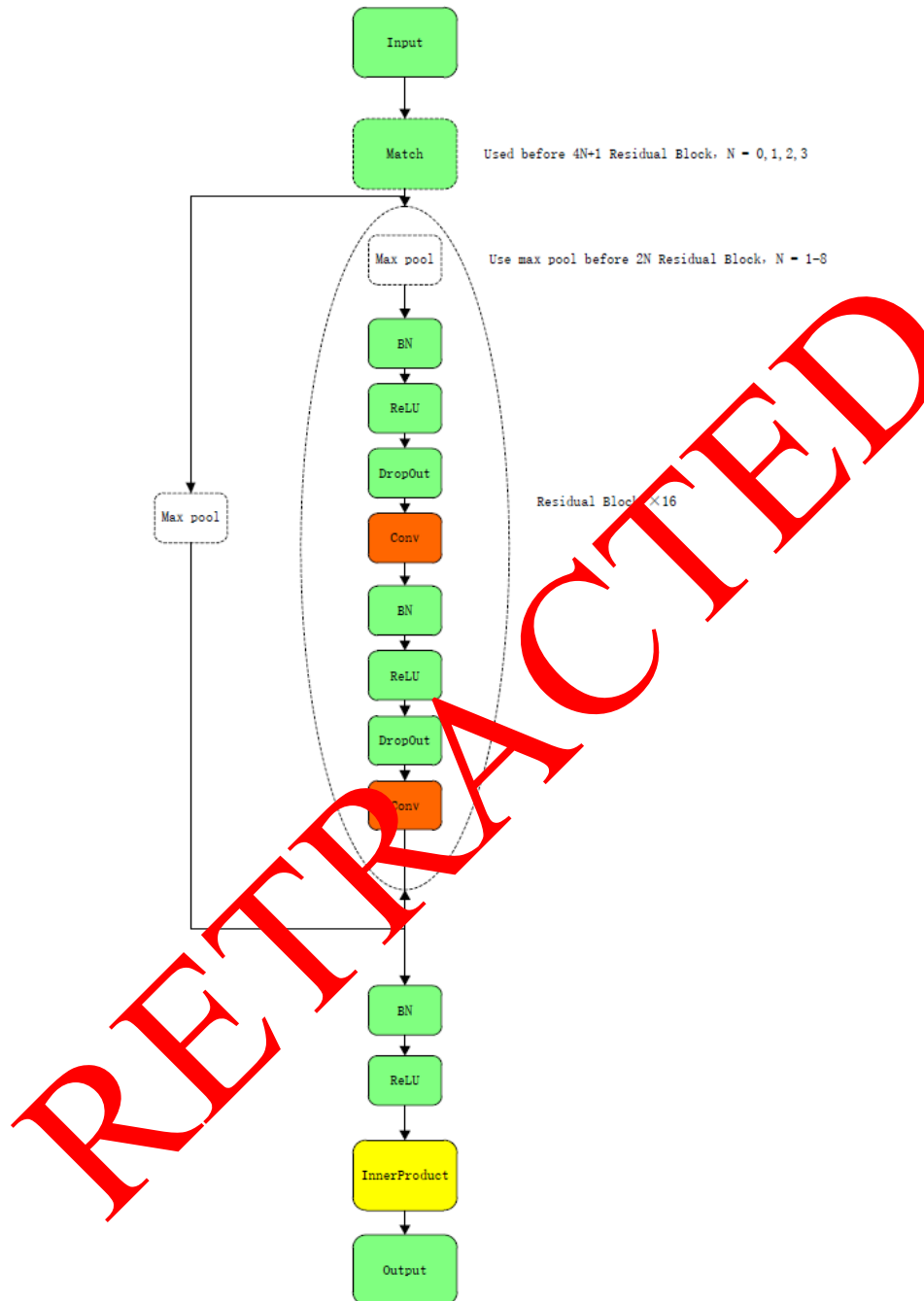

Figure 7: ANN Schematic diagram

Figure 8: Network frame structure

External data is input to the input layer, and the wiring represents the weights. There may be one or more hidden layers in the middle, and finally an output layer. Some artificial neural networks, such as BP neural network, are similar to a black box. Given input x and

output o, the network learns the best weight by itself through forward propagation and backward propagation. This is the learning process. However, the biggest problem of the traditional artificial neural network is that the weight is too much, therefore the model is very large, and the calculation efficiency is very low. In 2006, Geoffrey Hinton, the father of in-depth learning, put forward convolutional neural network, using convolution, pooling and other methods [Lecun, Bengio and Hinton (2015)]. The network weight (i.e., the number of connections between nodes) was effectively reduced, and the network accuracy was greatly improved. It has brought about revolutionary progress in artificial intelligence. Based on the latest achievements of convolutional neural network and Caffe framework, in this investigation, 36-layer convolution neural network is established by using convolutional layer, residual unit (proposed by Kaiming He of Microsoft Research Institute for Asia [He, Zhang and Ren (2016)], which can make neural network deeper), random deactivation (Dropout, which is mainly used to reduce the connection between some units randomly in each iteration and reduce the amount of computation), batch normalization (BN) and ReLU activation function. Fig. 8 is the network frame structure.

According to the network structure, there are 16 residual blocks with two convolutional layers in each cell. The lengths of all convolutional layer filters are 16, and the number of filters is 64k. The initial value of K is 1, which is increased by 1 after passing through four residual units. Two adjacent residual units (even numbers) sample the input sent by odd numbers twice down, so the final sample is 2^8 times down on the initial input. When the residual unit is down-sampled, the corresponding residual unit short-cut is also down-sampled, and the maximum pooling method (Max Pooling) is used. The match is the 1*1 volume layer, which realizes the transformation and matching between different residuals modules. The detailed network frame structure is as shown in Fig. 9 and Fig. 10:
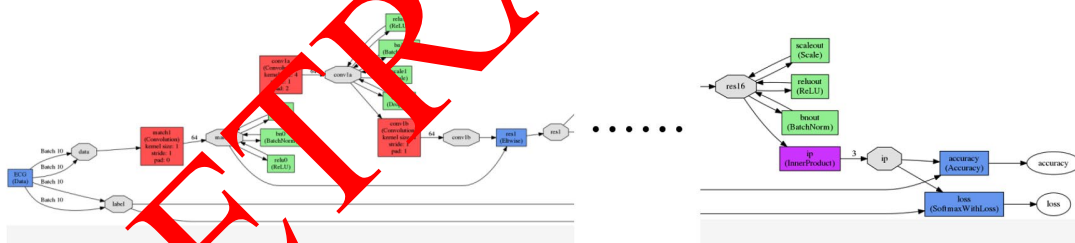

Figure 9: Detailed structure of the convolutional neural network
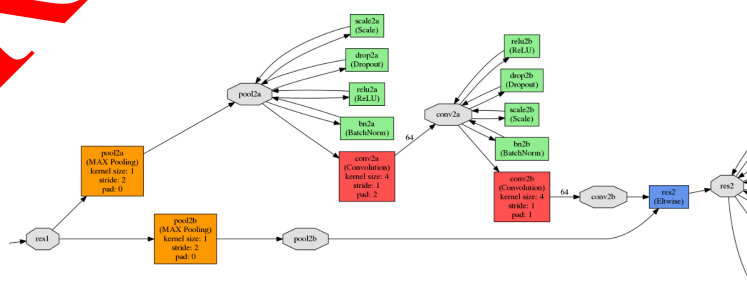

Figure 10: Laminated structure

After the above model is established, the training set based on MIT-BIH database is trained by Adam method [Wilson, Roelofs and Stern (2018)]. The Adam method is the Adaptive Moment Estimation, and it can calculate the adaptive learning rate of each parameter. Similar to the stochastic gradient descent method, it is a way to adjust the model updating weights and deviation parameters. In practical application, the Adam method has a good effect. Compared with other adaptive learning rate algorithms, it has faster convergence speed and more effective learning effect, and it can correct the problems existing in other optimization techniques, such as the disappearance of learning rate, slow convergence or high variance parameter updating resulting in large fluctuation of loss function.

Then the training model is used for classification.

3 Result

The data in the MIT-BIH database is read by MATLAB, and the original signal is obtained as shown in Fig. 11(a). The data obtained from the original signal denoised by adaptive filtering is shown in Fig. 11(b). It is evident that the baseline drift noise at low frequencies and the burr at high frequencies are well suppressed.
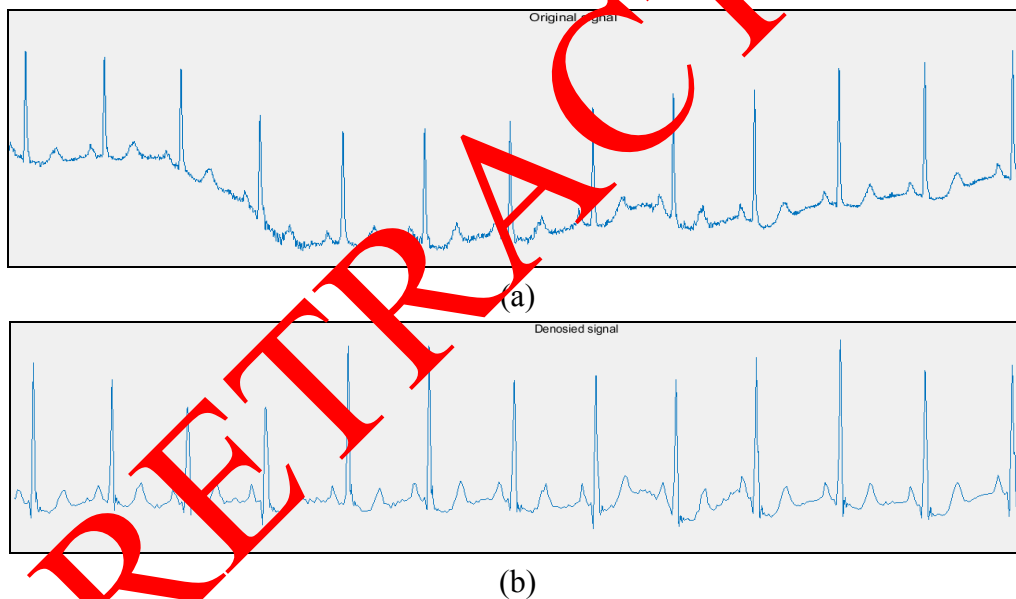

(a)

(b)

Figure 11: contrast before and after ECG signal preprocessing

After model training, the results were compared with those of the cardiologist, as shown in Tab. 1:

Table 1: Comparison of the convolutional neural network test results and cardiology results [Rajpurkar, Hannun and Haghpanahi (2017)]

	Model	Cardiologist
AFIB	**0.562**	0.515
AFL	**0.707**	0.635
BIGEMINY	0.766	**0.837**
IVR	**0.683**	0.632
SINUS	0.714	**0.847**

As can be seen from Tab. 1, the classification results based on convolutional neural network proposed in this study have exceeded the accuracy of human cardiologist in the analysis and judgment of three types of arrhythmias: AFIB, AFL and IVR. The accuracy of BIGEMINY and SINUS is slightly lower than that of a cardiologist.

4 Discussion

Based on the universal problem of cardiovascular disease in the world and the need of screening and diagnosis, the classification method of arrhythmia is studied. With the denoising and classification of ECG signal as the breakthrough point, the adaptive filtering is chosen to denoise and the convolutional neural network is chosen as research methods. Aiming at the research of ECG arrhythmia classification method based on convolutional neural network, a new improved convolutional neural network structure is proposed to improve the accuracy of arrhythmia classification.

The depth learning algorithm based on 36-layer convolutional neural network proposed in this article has good performance in the analysis and judgment of three types of arrhythmias: AFIB, AFL and IVR, but the accuracy of judgment of other types of arrhythmias is still insufficient. The next work will continue to optimize this algorithm.

Further study of QRS wave group detection method to further improve the accuracy of QRS group detection. Because the ECG signal itself is susceptible to interference, it is very difficult to accurately detect the QRS wave group if the ECG signal is superimposed with the interference signal. The accurate detection of QRS wave group is the basis of the subsequent classification methods. Therefore, it is very important to improve the accuracy of QRS wave group detection.

In addition, optimizing the structure and parameters of convolutional neural network and putting forward more applicable network structure are also important aspects to improve the classification accuracy of arrhythmia. In addition, it is vital to apply the proposed method to Holter's actual measurement and analysis, and further verify the reliability of this method by the measured data. The next step is to collect, record and store the ECG signals of patients with arrhythmias in the hospital. The data obtained will be applied to test the method proposed in this study, in order to increase the classification range of arrhythmias, and to strive to further improve the accuracy of classification of other types of arrhythmias.

On the other hand, the proposed method needs a strong computing power to complete, and cannot be directly applied to wearable devices with weak computing power. In order

to realize the practical application of wearable devices, the lightweight optimization of the network is also needed. Applying arrhythmia classification method to portable devices will have a higher value, which will be a very important direction in the future.

References

Acharya, U. R.; Fujita, H.; Lih, O. S. (2017): Application of deep convolutional neural network for automated detection of myocardial infarction using ECG signals. *Information Sciences*, vol. 415, pp. 190-198.

Acharya, U. R.; Fujita, H.; Lih, O. S. (2017): Automated detection of arrhythmias using different intervals of tachycardia ECG segments with convolutional neural network. *Information Sciences*, vol. 405, pp. 81-90.

Benzid, R.; Marir, F.; Boussaad, A. (2003): Fixed percentage of wavelet coefficients to be zeroed for ECG compression. *Electronics Letters*, vol. 39, no. 11, pp. 830-831.

Chouhan, V. S.; Mehta, S. S. (2007): Total removal of baseline drift from ECG signal. *International Conference on Computing: Theory and Applications*, pp. 512-515.

Georganis, A.; Doulgeraki, N.; Asvestas, P. (2017): A real-time ECG signal processing application for arrhythmia detection on portable devices. *Journal of Physics Conference Series*, vol. 931, no. 1, 012004.

Gang, L. (1999): The mathematical morphology filter and it's application in the ECG. *Chinese Journal of Scientific Instrument*, no. 4, pp.

He, K.; Zhang, X.; Ren, S. (2016): Deep residual learning for image recognition. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 770-778.

Kiranyaz, S.; Ince, T.; Gabbouj, M. (2016): Real-time patient-specific ECG classification by 1-d convolutional neural networks. *IEEE Transactions on Biomedical Engineering*, vol. 63, no. 3, pp. 664-675.

Kiranyaz, S.; Ince, T.; Hamila, R. (2015): Convolutional neural networks for patient-specific ECG classification. *37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 2608-2611.

Lulu, Z.; Lin, Y.; Yuliang, Y. (2012): Removal of baseline drift from ECG signals based on morphology and adaptive filter. *Control Conference*, pp. 3750-3753.

Lecun, Y.; Bengio, Y.; Hinton, G. (2015): Deep learning. *Nature*, vol. 521, no. 7553, pp. 436.

Li, C.; Zheng, C.; Tai, C. (1995): Detection of ECG characteristic points using wavelet transforms. *IEEE Transactions on Biomedical Engineering*, vol. 42, no. 1, pp. 21-28.

Martis, R. J.; Acharya, U. R.; Min, L. C. (2013): ECG beat classification using PCA, LDA, ICA and discrete wavelet transform. *Biomedical Signal Processing & Control*, vol. 8, no. 5, pp. 437-448.

Narayana, K. V. L.; Rao, A. B. (2011): Wavelet based QRS detection in ECG using MATLAB. *Innovative Systems Design & Engineering*, vol. 2, no. 7.

Rajpurkar, P.; Hannun, A. Y.; Haghpanahi, M. (2017): Cardiologist-level arrhythmia detection with convolutional neural networks. arXiv:1707.01836v1.

Raj, S.; Ray, K. C. (2017): ECG signal analysis using DCT-based DOST and PSO optimized SVM. *IEEE Transactions on Instrumentation & Measurement*, vol. 66, no. 3, pp. 470-478.

Shirbani, F.; Setarehdan, S. K. (2013): ECG power line interference removal using combination of FFT and adaptive non-linear noise estimator. *Electrical Engineering*, pp. 1-5.

Sahambi, J. S.; Tandon, S. N.; Bhatt, R. K. (1997): Using wavelet transforms for ECG characterization. An on-line digital signal processing system. *IEEE Engineering in Medicine & Biology Magazine*, vol. 16, no. 1, pp. 77-83.

Saritha, C.; Sukanya, V.; Murthy, Y. N. (2008): ECG signal analysis using wavelet transforms. *Bulgarian Journal of Physics*, vol. 35, no. 1, pp. 155-163.

Schwartz, W. R.; Menotti, D. (2016): ECG-based heartbeat classification for arrhythmia detection. *Computer Methods & Programs in Biomedicine*, vol. 127, pp. 144-164.

Wang, X.; Xu, X. J.; He, Q. Y. (2015): A study of ECG signal myoelectricity interference removal method. *Intelligent Computer & Applications*.

Wilson, A. C.; Roelofs, R.; Stern, M. (2018): The marginal value of adaptive gradient methods in machine learning. *Advances in Neural Information Processing Systems*, pp. 4148-4158.