



Heartbeat classification using deep residual convolutional neural network from 2-lead electrocardiogram

Zhi Li ^{a,b}, Dengshi Zhou ^a, Li Wan ^c, Jian Li ^{a,*}, Wenfeng Mou ^a

^a College of Electronic and Information Engineering, Sichuan University, Chengdu 610065, China

^b Key Laboratory of Wireless Power Transmission of Ministry of Education, Sichuan University, 610065, China

^c Department of Gynecology and Obstetrics, West China Second Hospital, Sichuan University, No.20 Renminnan Road, Chengdu 610041, PR China

ARTICLE INFO

Keywords:

Heartbeat classification
Arrhythmia
Deep learning
Residual convolutional neural network
ECG signals
2-Lead

ABSTRACT

Background: The electrocardiogram (ECG) has been widely used in the diagnosis of heart disease such as arrhythmia due to its simplicity and non-invasive nature. Arrhythmia can be classified into many types, including life-threatening and non-life-threatening. Accurate detection of arrhythmic types can effectively prevent heart disease and reduce mortality.

Methods: In this study, a novel deep learning method for classification of cardiac arrhythmia according to deep residual network (ResNet) is presented. We developed a 31-layer one-dimensional (1D) residual convolutional neural network. The algorithm includes four residual blocks, each of which consists of three 1D convolution layers, three batch normalization (BP) layers, three rectified linear unit (ReLU) layers, and an “identity shortcut connections” structure. In addition, we propose to use 2-lead ECG signals in combination with deep learning methods to automatically identify five different types of heartbeats.

Results: We have obtained an average accuracy, sensitivity and positive predictivity of 99.06%, 93.21% and 96.76% respectively for single-lead ECG heartbeats. In the 2-lead datasets, the results show that the deep ResNet model has high classification performance, achieving an accuracy of 99.38%, sensitivity of 94.54%, and specificity of 98.14%.

Conclusion: The proposed method can be used as an adjunct tool to assist clinicians in their diagnosis.

© 2019 Elsevier Inc. All rights reserved.

Introduction

Electrocardiogram (ECG) is a comprehensive potential activity produced by cardiomyocytes during cardiac contraction and relaxation. ECG is mainly composed of QRS complex, P wave, T wave, etc. It is usually possible to determine whether a patient has arrhythmia or other heart disease through an ECG waveform, which includes lots of pathological information. Arrhythmia is the abnormal rhythm of the heartbeat. During arrhythmia, the heartbeat is too fast, too slow or irregular [1]. However, such diagnostics require professionally trained health professionals to visually assess irregular ECGs, which is often a time consuming and subjective process in some cases [2]. Due to the large differences in morphology, it is difficult to accurately classify arrhythmia beats. Therefore, computer-aided diagnostic (CAD) techniques have been used to automatically identify arrhythmia classifications over the past few decades [3].

The ECG heartbeat classification can be divided into four parts, namely, pre-processing, heartbeat segmentation, feature extraction or feature selection, and classification. In each of the above sections, a

number of algorithms are proposed [4–12,30,31]. However, the parts associated with classification (feature extraction and classification learning 4algorithms) still have room for exploration and improvement [25]. The segmented ECG signal has many sampling points, and researchers use many traditional machine learning algorithms for feature extraction and dimensionality reduction. Some statistical methods, such as principal component analysis (PCA) [5,31], higher order statistic (HOS) [6], independent component analysis (ICA) [7,30] and linear discriminant analysis (LDA) [8], were used for feature extraction of ECG signals. Elhaj et al. [9] combines multiple feature extraction (HOS, ICA and PCA) methods and use SVM and NN classifiers to achieve 98.91% accuracy in arrhythmia classification. Most studies have shown that wavelet transform has a good performance for ECG signal feature extraction, because wavelet transform can extract both frequency domain and time domain information [10,11]. Yildirim et al. [12] used WT method to classify five types of beats and achieved 97.29% accuracy.

In recent years, many classification methods based on deep learning have been proposed in the field of ECG signal research. Unlike traditional methods, deep learning methods provide a structure in which feature extraction and classification are performed together, rather than using hand-crafted features. In fact, the method based on deep learning has the ability to learn useful features from the input signals.

* Corresponding author.

E-mail address: lijianz@scu.edu.cn (J. Li).

Long-term short-term memory (LSTM) is widely used in time series signal analysis, such as classification of ECG signals [3,35]. In addition to LSTM, convolutional neural network (CNN) play an important role in deep learning and are widely used in various fields. CNN has achieved great success in the field of image processing and computer vision [14,15]. CNN has been applied to morphological analysis of physiological signals due to its unique ability to capture position and displacement invariance [16–18]. Rajpurkar et al. [19] used a 34-layer CNN model to detect arrhythmia heartbeats in a total of 64,121 records collected from 29,163 patients, and eventually achieved 80.0% precision and 78.4% sensitivity. Considering that CNN is less sensitive to noise, Acharya et al. [1] applied a 9-layer CNN to achieve an accuracy of 94.03% and 93.47% in the classification of arrhythmia in original and denoising heartbeats respectively.

With the deepening of the number of network layers, the learning ability of CNN will be enhanced. However, simply stacking the number of network layers cannot improve the accuracy. When the network layers are very deep, the main obstacle to training a deep network model is the problem of vanishing/exploding gradients [20]. This problem has been largely solved by normalized initialization and intermediate normalized methods [21,22]. Even with these methods, the training of deep neural network still has the phenomenon that the accuracy decreases with the increase of network depth. ResNet provides a solution for deep network training [23]. The network structure does not add a lot of parameters and computation, but it can solve the problem of training accuracy degradation. In this work, a novel approach using 19-layer, 25-layer and 31-layer deep residual convolutional neural network is proposed to identify the five heartbeat types. Compared with the traditional CNN, we use a deeper network structure and improve the accuracy of ECG classification. In addition, we used single-lead and 2-lead ECG signals for experimental comparison.

Database

In this study, the data are obtained from the MIT-BIH arrhythmia database [13], which includes many types of arrhythmia. The database contains 48 records from 47 subjects, each record including half-hour ambulatory ECG signals (noted as lead A and lead B). In addition to 102, 104 and 114 records, the first lead of the other 45 records is modified-lead II (MLII). The lead V1 contains the most records in the second lead. The details of the lead have been recorded in Table 1. We chose the lead MLII and lead V1 as a 2-lead datasets with a total of 40 records.

According to standards advocated by Advanced Medical Instruments (AAMI) [24], the fifteen types of heartbeats in Table 2 are re-divided into five classes. These five classes are defined by AAMI as N, S, V, F and Q. More five classes of partitioning details are shown in Table 3. We used a total of 94, 013 heartbeats for training and testing. In this study, we randomly select 20% of the total data set for testing, and the remaining data uses 80% as the training set and 20% as the verification set. The data distribution details for training and testing are shown in the Fig. 1.

Methods

Removal of noise

The raw ECG signals contains a lot of noise, such as power frequency interference, baseline drift, high frequency noises caused by muscle

contraction and electrode movement, etc. [4]. In order to remove these noises while preserving the characteristics of the ECG signal and reducing the computational overhead [25], we use the method based on wavelet transform [26].

Heartbeat segmentation

Each heartbeat is independently labeled by 2 or more cardiologists and ultimately agreed. The segmentation of ECG signals is based on manual labeling of R wave peak position or main local extremum [4]. In order to retain most of the heartbeat information, we choose 99 sampling points before the marked R peak position and 200 sampling points after it. The sampling rate of ECG signals in the MIT-BIH arrhythmia database is 360 Hz, so 300 sampling points are equivalent to 0.83 s. Each segmented heartbeat contains most of the pathological information. Five types of heartbeats containing 2-lead are shown in Fig. 2, with 30 ECG waveforms for each lead in each type of heartbeat.

Convolutional neural network (CNN)

CNN has great potential to automatically extract important features from 2D input data such as images [24]. However, many biomedical signals are 1D signals related with time series, such as ECG and electroencephalogram (EEG) signals, so we use 1D convolution to process the signal. The basic unit of CNN is composed of input layer, convolution layer, activation function and output layer which can be expressed as:

$$y = f(Wx + b) \quad (1)$$

where x is the input; y is the output; f is the ReLU function; W is the convolution matrix; b is the bias.

Batch normalization (BN)

In the process of deep network training, the input distribution of each layer varies with the parameters of the previous layer. This phenomenon is called internal covariant shift, which is usually considered as an important factor affecting the speed of network training. In order to reduce the internal covariate shift caused by the progressive transformation, Ioffe et al. [28] proposed batch normalization, which effectively solves the problem by introducing normalization and shift steps in each nonlinear transformation. In the training process, the training data needs to be divided into basic data units, namely mini-batch, and then batch normalization. Considering a mini-batch of size m , the steps and formulas for batch normalization are as follows:

$$\mu = \frac{1}{m} \sum_{i=1}^m x_i \quad (2)$$

$$\sigma^2 = \frac{1}{m} \sum_{i=1}^m (x_i - \mu)^2 \quad (3)$$

$$BN_{y,\beta}(x_i) = y \left(\frac{x_i - \mu}{\sqrt{\sigma^2 + \epsilon}} \right) + \beta \quad (4)$$

where x_i represents entries in the current mini-batch, and σ^2 and μ are the standard deviation and mean of the mini-batch, respectively. In

Table 1
Distribution of lead types and record numbers in MIT-BIH arrhythmia database.

Lead A	V5	V5	MLII	MLII	MLII	MLII																													
Lead B	V2	MLII	V2	V4	V5	V1																													
Records	102	114	103	124	100	101	105	106	107	108	109	111	112	113	115	116	118	119	121	122	200	201	202	203	205	207	208	209	210	212	213	214	215	217	219
	104		117		123	220	221	222	223	228	230	231	232	233	234																				

Table 2

Fifteen arrhythmia heartbeats in the MIT-BIH database.

Heartbeat type	Annotation	Heartbeat type	Annotation
Normal beat	N	Supraventricular premature beat	S
Left bundle branch block beat	L	Nodal (junctional) escape beat	j
Right bundle branch block beat	R	Fusion of paced and normal beat	f
Atrial premature beat	A	Ventricular escape beat	E
Premature ventricular contraction	V	Nodal (junctional) premature beat	J
Aberrated atrial premature beat	a	Atrial escape beat	e
Paced beat	P	Unclassifiable beat	Q
Fusion of ventricular and normal beat	F	Total	15

formula (4), ε is a constant added to the mini-batch variance for numerical stability, and γ and β are trained parameters.

Residual network

The depth of the deep learning network is an important factor affecting the final classification and recognition results, so the usual idea is to make the design of the neural network as deep as possible. In fact, after building a certain depth, increasing the depth will degrade the performance of the deep learning network. One of the reasons is that the deeper the network design, the more obvious the problem of vanishing/exploding gradients, which makes the network training more difficult [27]. In contrast, the residual network is an improved deep learning algorithm for CNN, which avoids these problems by using “shortcut connections” that skip multiple network layers [23].

One-dimensional convolution is usually used to process one-dimensional signals. The 1D residual block is shown in Fig. 3(a), and the relationship between the input and output of the residual block is expressed as

$$I_j = \text{ReLU}(F(I_{j-1}, \mathbf{w}_j)) + I_{j-1} \quad (5)$$

where I_{j-1} is the input vectors of the current residual block and I_j is the

output vectors of the current residual block. Each residual structure is activated by a ReLU function. The \mathbf{w}_j is the weight parameter of the j th layer of the ResNet block that needs to be learned. The function $F(I_{j-1}, \mathbf{w}_j)$ contains 1D convolutional layer, BN layer and ReLU layer.

Since the dimensions of I_{j-1} and F may be different, we use linear projection \mathbf{W}_s to match the dimensions of both. The residual structure shown in the Fig. 3(b) is defined as:

$$I_j = \text{ReLU}(F(I_{j-1}, \mathbf{w}_j) + \mathbf{W}_s I_{j-1}) \quad (6)$$

In the shortcut connections branch, we uses feature maps from a 1D convolutional and BP layer as residual information to supplement the residual block features.

The architecture

We propose a novel deep ResNet to effectively learn the features of ECG signals. In the network structure of this model, the input layer has two channels because of the data of two leads. After the input layer, there is a 1D convolution layer (the size of kernels, the strides and the number of filters are 5, 2 and 32 respectively), a BP layer, a ReLU layer and a max-pooling layer (the pooling size and the strides are 3 and 2 respectively). As shown in Fig. 4, there are 4 residual blocks in the hidden layer, each of which has a similar structure. Residual block 1 structure is shown in Fig. 3(b), and other residual block structure parameters and network parameters are shown in Table 4. The convolution kernel size of the first convolution layer in each residual block is 1×1 , the second is 3×1 , and the third is 1×1 . The advantage of using 1×1 convolution kernel is to greatly reduce the number of parameters and computational consumption. After the last residual block, we apply an average pooling layer followed by a full connection layer. The last layer is the softmax layer, which classifies five classes of arrhythmic heartbeats.

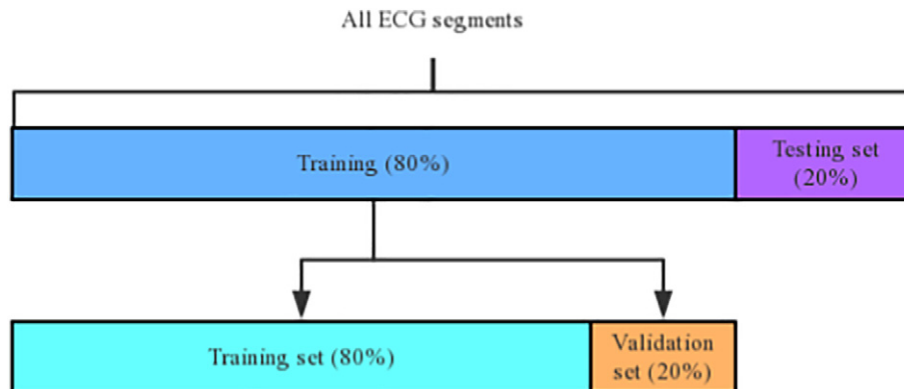
Results

This work runs on a deep learning framework of Keras with Tensorflow as the backend. The workstation used consists of 8 GB GPU

Table 3

The five classes of division schemes recommended by AAMI.

AAMI class	MIT-BIH classes annotations	Group description	Number of heartbeat
N	L, N, R, e, j	Normal beats	79,508
S	A, J, S, a	Supraventricular ectopic beats	2696
V	E, V	Ventricular ectopic beats	7129
F	F	Fusion beats	793
Q	P, Q, f	Unclassified beats	3887
Total	15	5	94,013

**Fig. 1.** The data distribution details for training and testing.

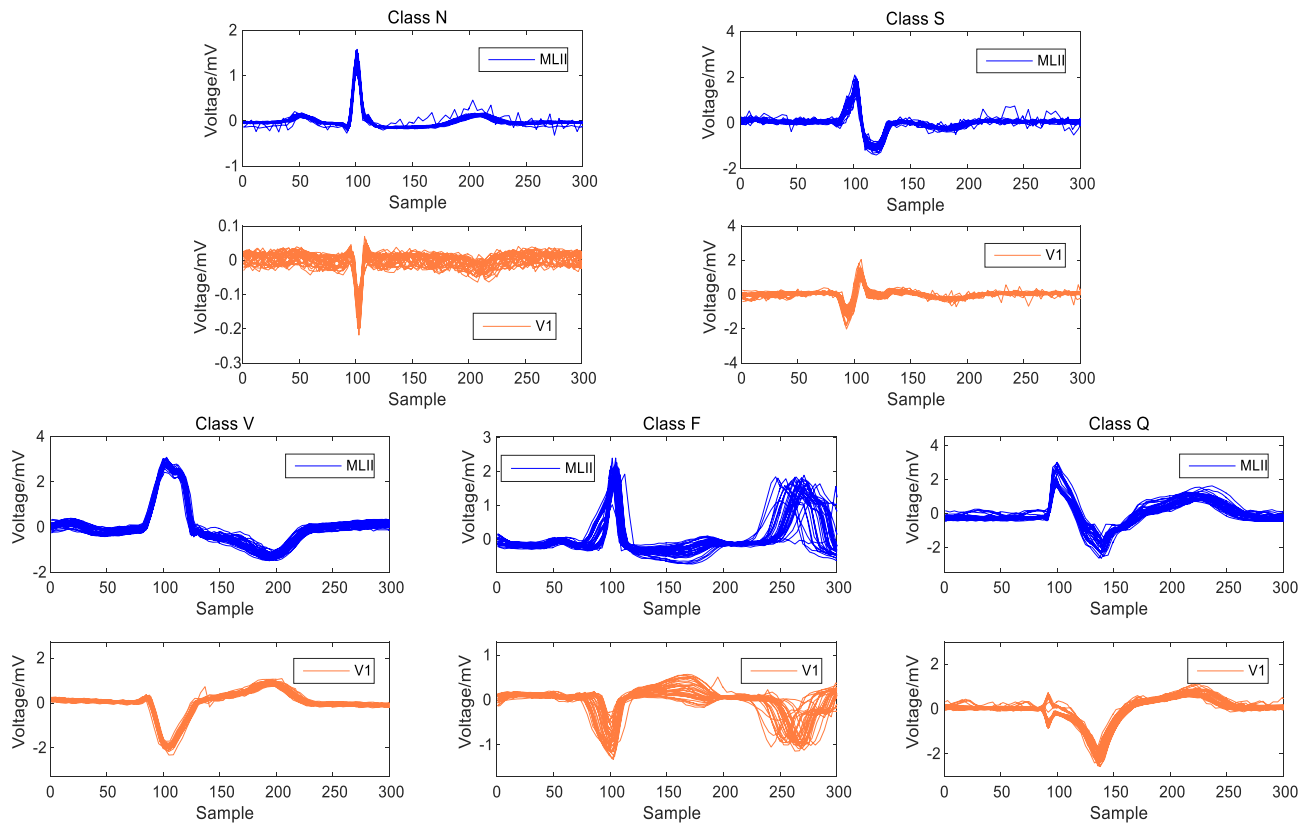


Fig. 2. Five types of heartbeats in MLII lead and V1 lead.

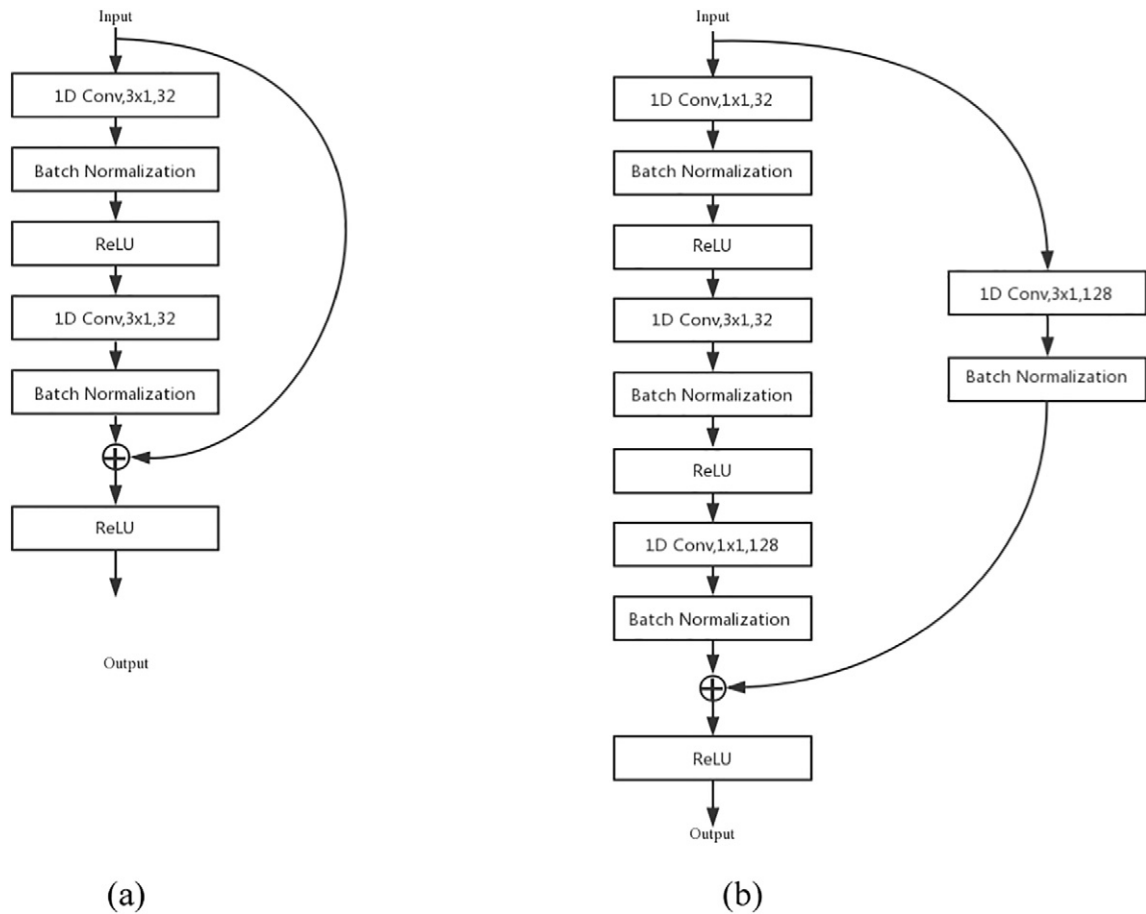


Fig. 3. ResNet architectures. (a) 1D convolutional residual block. (b) The proposed 1D convolutional residual block with linear projection.

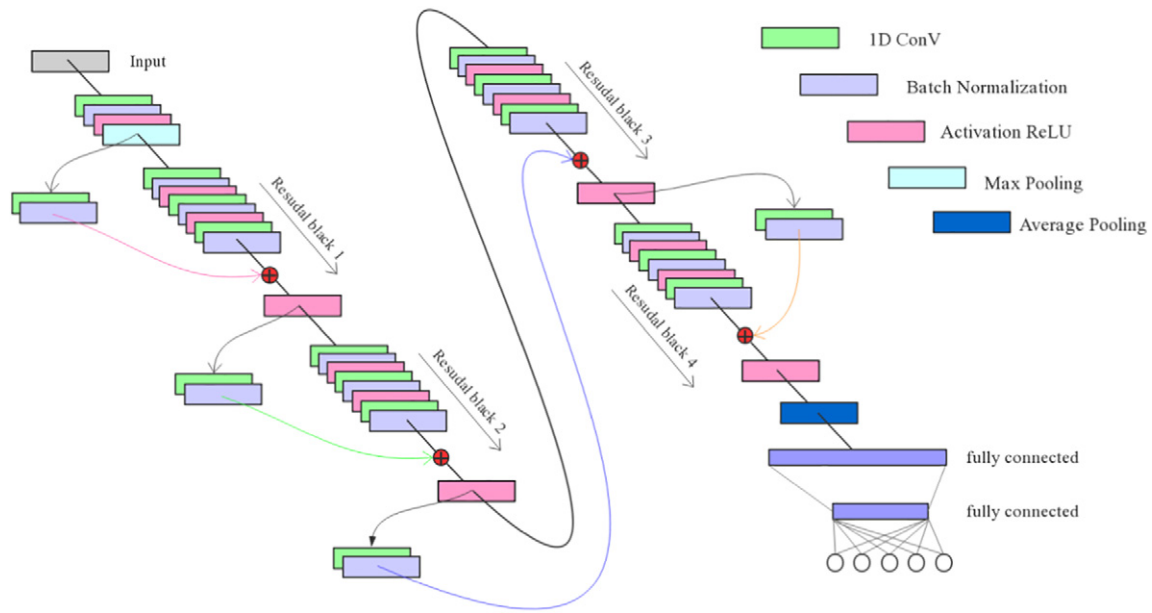


Fig. 4. The architecture for the proposed ResNet model.

(NVIDIA GeForce GTX-1070), Intel i7-4790 processor (3.60 GHz) and 8 GB RAM. Usually the ResNet model is deeper than the traditional CNN model. Due to the different performance of different layers of networks, we designed a variety of different layers of ResNet architectures for training. As shown in Table 4, we proposed a 31-layer ResNet model, and we also designed a 19-layer ResNet model with the first two residual blocks and a 25-layer ResNet model with the first three residual blocks.

In order to evaluate the performance of single-lead and 2-lead on the different deep learning models, we used a single-lead datasets with only MLI lead and a 2-lead datasets with MLI lead and V1 lead. In our paper, the batch size of each deep learning model was 100. The validation

results of single-lead are shown in Fig. 5(a). In the 20th epoch, validation datasets were correctly identified at 99.17%, 99.15% and 99.18% respectively in the proposed ResNet-19, ResNet-25 and ResNet-31 model.

The 2-lead recognition performance of the ResNet-25 and ResNet-31 model was 98.66% and 98.53% respectively at the end of the first epoch in Fig. 5(b). It can be observed that the validation accuracy of the proposed three ResNet models starts to plateau at the end of 15 epochs. Overall, the validation curve of ResNet-31 model performs better than that of ResNet-19 model and ResNet-25 model.

In Tables 5 and 6, confusion matrix of ResNet-31 model test results in single-lead and 2-lead are presented respectively. The ResNet-31 model has high performance for each type of heartbeats, and the use of 2-lead

Table 4
Details of the proposed ResNet structure.

Layers	Type	Activation function	Pool size	No. of filters	Kernel size	Stride	Output shapes
0	Input	—	—	—	—	—	300×2
1	1D convolution	ReLU	—	32	5×1	2	148×32
2	Batch normalization	—	—	—	—	—	148×32
3	1D max-pooling	—	3	—	5×1	2	74×32
4–9	1D convolution in residual blocks 1	ReLU	—	32, 32, 128	1×1 , 3×1 , 1×1	1, 1, 1	74×32 , 74×32 , 74×128
10–15	1D convolution in shortcut connection 1D convolution in residual blocks 2	ReLU ReLU	— —	128 64, 64, 256	3×1 1×1 , 3×1 , 1×1	2, 2, 1, 1	74×128 37×64 , 37×64 , 37×256
16–21	1D convolution in shortcut connection 1D convolution in residual blocks 3	ReLU ReLU	— —	256 128, 128, 512	3×1 1×1 , 3×1 , 1×1	2, 2, 1, 1	37×256 19×128 , 19×128 , 19×512
22–27	1D convolution in shortcut connection 1D convolution in residual blocks 4	ReLU ReLU	— —	512 256, 256, 1024	3×1 1×1 , 3×1 , 1×1	2, 2, 1, 1	19×512 10×256 , 10×256 , 10×1024
28	1D convolution in shortcut connection	ReLU	—	1024	3×1	2	10×1024
29	1D average-pooling	—	3	—	—	—	1×1024
30	Flatten	—	—	—	—	—	1024
31	Fully-connected	tanh	—	—	—	—	40
32	Fully-connected	softmax	—	—	—	—	5

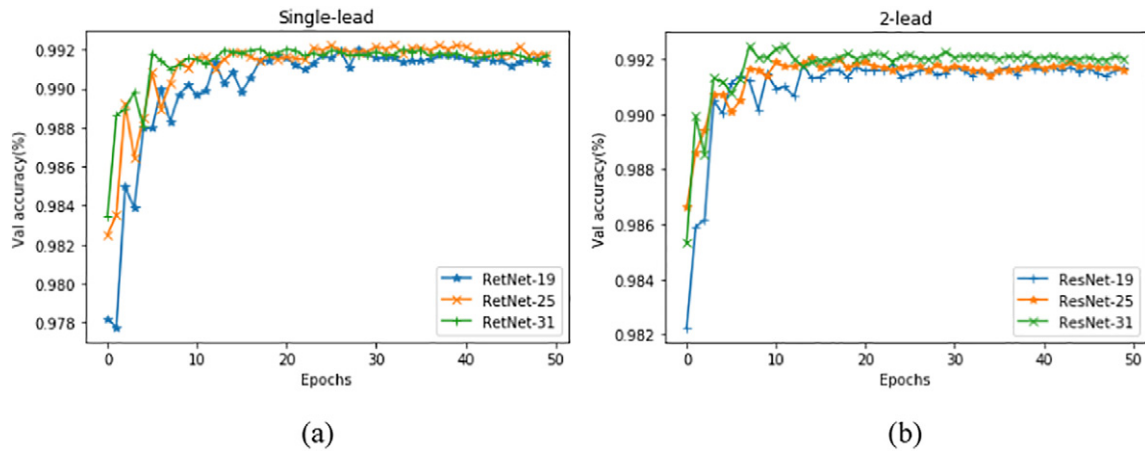


Fig. 5. Validation results of the three proposed ResNet models. (a) Validation results for single-lead datasets. (b) Validation results for 2-lead datasets.

further improves classification accuracy. Due to the imbalance of datasets, the classification performance of different types of heartbeat varies greatly.

The heartbeats datasets were fed into the 9-layer CNN model [1], 11-layer CNN model [37] and the proposed deep ResNet model for training and testing. Table 7 represents the comprehensive performance comparisons of different architecture in different lead datasets. Table 7 shows that ResNet-31 model achieved the highest recognition performance but took the longest average time per epoch.

Discussion

The MIT-BIH arrhythmia database has been used in many studies on the classification of arrhythmias for many years. Table 8 summarizes the comparison of the proposed methods with others studies in the past decade. Li et al. [32] proposed multi-domain feature extraction method, which employs kernel-independent component analysis (KICA) and discrete wavelet transform (DWT) to extract non-linear and frequency domain features of ECG signals, respectively. The system obtained a 98.8% recognition performance. In other study [29], they used the entropy of the wavelet packet decomposition (WPE) coefficients and the Random forest (RF) classifier to achieve 94.61% accuracy. Elhaj et al. [9] proposed a neural network and support vector machine classification method showed a 98.91% accuracy. Sahoo et al. [33] used a QRS complex features detection algorithm based on the multi-resolution wavelet transform to detect four types of ECG beats. In their study, they used 109,494 heartbeats from the MIT-BIH arrhythmia database and reached 98.39% performance.

Deep learning has become increasingly popular in recent years and has been applied in arrhythmia detection. Kiranyaz et al. [34] and Acharya et al. [1] used CNN to identify arrhythmias in single-lead data sets (MLII leads), and achieved 99% and 94.03% accuracy, respectively. Since the ECG signals are time series signals, Yildirim

et al. [35] used bidirectional and unidirectional LSTM to achieve an accuracy of 99.39% in 7737 heartbeats. In another work of Yildirim et al. [36], they used convolutional auto-encoder (CAE) LSTM to obtain 99.23% and 99.11% accuracy of raw data and coded features data, respectively. Oh et al. [3] combined CNN and LSTM to detect arrhythmia and obtained 98.1% accuracy rate. They used 16,499 heartbeat segments, each containing 1000 sampling points, about 3 or 4 heartbeats.

Generally speaking, if the number of neural network model layers is deep, the model may have the problem of vanishing gradients, thus make it difficult to improve the training results. Through the deeper network structure and residual connection, the proposed deep ResNet model has better test results than the traditional CNN [1,37] model. As shown in Table 8, most of the current studies only used MLII lead data from the MIT-BIH arrhythmia database, and some studies did not provide lead information. In our study, single-lead and 2-lead datasets were fed into the ResNet-31 model for training and testing, and achieved 99.06% and 99.38% accuracy, respectively.

The advantages of our work are as follows:

- The feature extraction process is not required, and the pre-processed ECG signals are trained and classified directly.
- Compared with the traditional convolutional neural network, this paper designs a deeper residual convolutional neural network model to classify ECG heartbeats. The model can still improve the accuracy of the network with the increase of network depth.
- The 2-lead ECG signal is used for each heartbeat to further improve classification accuracy.

On the other hand, due to the deeper layers of the Deep ResNet model, the average training time per epoch is longer. In addition, it is inevitable that the training of 2-lead is more time-consuming than that of single-lead.

Table 5

A confusion matrix of single-lead in ResNet-31.

		Predicted					+P (%)	Se (%)	Acc (%)
		N	S	V	F	Q			
Original	N	15,952	24	19	5	2	99.33	99.69	99.06
	S	56	475	1	0	0	95.0	89.29	
	V	24	1	1310	5	0	97.69	97.76	
	F	21	0	10	127	0	92.03	80.38	
	Q	6	0	1	1	763	99.74	98.96	

Se: Sensitivity, +P: Positive predictivity, Acc: accuracy.

Table 6

A confusion matrix of 2-lead in ResNet-31.

		Predicted					+P (%)	Se (%)	Acc (%)
		N	S	V	F	Q			
Original	N	15,974	18	8	2	0	99.54	99.83	99.38
	S	39	491	2	0	0	96.27	92.29	
	V	13	0	1324	3	0	98.59	98.81	
	F	18	1	9	130	0	96.30	82.28	
	Q	4	0	0	0	767	1.0	99.48	

Se: Sensitivity, +P: Positive predictivity, Acc: accuracy.

Table 7

Comparison of comprehensive performance between CNN model and proposed ResNet model.

Architecture	Lead	Acc (%)	Se (%)	+P (%)	Average time (sec) per epoch
CNN-9 [1]	Single-lead	96.01	67.38	76.86	8.90
	2-Lead	96.84	75.20	91.29	9.32
CNN-11 [37]	Single-lead	96.89	77.85	92.69	9.24
	2-Lead	97.50	82.84	92.74	9.51
ResNet-19	Single-lead	99.03	91.88	96.50	18.55
	2-Lead	99.27	94.25	97.37	19.70
ResNet-25	Single-lead	98.88	92.19	95.04	25.11
	2-Lead	99.31	94.14	98.17	25.72
ResNet-31	Single-lead	99.06	93.21	96.76	28.34
	2-Lead	99.38	94.54	98.14	34.25

Se: Sensitivity, +P: Positive predictivity, Acc: accuracy.

Conclusion

The automatic diagnosis technology of arrhythmia has important research value and can reduce the harm of cardiovascular diseases. In this article, a novel deep learning algorithm is proposed to automatically recognize five different ECG signals. This algorithm has higher classification performance than the traditional machine learning algorithm and other deep learning algorithms in the classification of arrhythmia. In the future studies, we will try to combine deep ResNet model with other deep learning methods to improve the classification performance of arrhythmia heartbeats.

Acknowledgements

This work was supported by the Key Research and Development Project of Science & Technology Department of Sichuan Province [Grant No. 2019YFG0192].

Table 8

Performance comparison of the proposed methods with others state-of-the-art studies.

Study	Years	Leads	# of beats	Approach	Performance
Martis et al. [30]	2013	–	110,094	• ICA, DWT • PNN	Acc: 99.28% Se: 97.97% +P: 99.21%
Martis et al. [31]	2013	–	34,989	• Bispectrum, PCA • Least Square-SVM	Acc: 93.48% Se: 99.27%
Li et al. [32]	2016	Signal lead	1800	• DWT, kernel ICA and PCA • SVM	Acc: 98.8% Se: 98.50% +P: 98.91%
Li et al. [29]	2016	–	100,688	• WPE, RR • RF	Acc: 94.61%
Elhaj et al. [9]	2016	–	110,094	• DWT, PCA, HOS, ICA • SVM-RBF	Acc: 98.91% Se: 98.91%
Sahoo et al. [33]	2017	MLII	109,494	• DWT • SVM	Acc: 98.39% Se: 99.87% +P: 99.69%
Kiranyaz et al. [34]	2016	MLII	83,648	• End-to-end • CNN	Acc: 99.00% Se: 93.90%
Acharya et al. [1]	2017	MLII	109,449	• End-to-end • CNN	Acc: 94.03% Se: 96.71%
Yildirim et al. [35]	2018	MLII	7376	• End-to-end • DBLSTM	Acc: 99.39%
Yildirim et al. [36]	2019	MLII	100,022	• End-to-end • CAE-LSTM	Acc: 99.23%
Oh et al. [3]	2018	MLII	16,499	• End-to-end • LSTM and CNN	Acc: 98.10% Se: 97.50% +P: 98.69%
This study	2019	MLII	94,013	• End-to-end • 31-layer deep ResNet	Acc: 99.06% Se: 93.21% +P: 96.76%
This study	2019	MLII and V1	94,013	• End-to-end • 31-layer deep ResNet	Acc: 99.38% Se: 94.54% +P: 98.14%

Acc: Accuracy, Se: Sensitivity, +P: Positive predictivity; LDA: Linear Discriminant Analysis; HOS: High Order Statistics; PCA: Principal Component Analysis; SVM-RBF: Support Vector Machine (Radial Basis Function); ICA: Independent Component Analysis; DWT: Discrete wavelet transform; PNN: Probabilistic Neural Network; WPE: Wavelet Packet Entropy; RR: RR intervals; RF: Random Forest; CNN: Convolutional Neural Network; LSTM: Long Short-Term Memory; DBLSTM: Deep Bidirectional LSTM; ResNet: Residual Network.

References

- [1] Acharya UR, Oh SL, Hagiwara Y, et al. A deep convolutional neural network model to classify heartbeats [J]. *Comput Biol Med* 2017;89.
- [2] Osowski S, Hoai LT, Markiewicz T. Support vector machine-based expert system for reliable heartbeat recognition [J]. *IEEE Transactions on Biomedical Engineering* 2004;51(4):582–9.
- [3] Lih OS, Ng EYK, San TR, et al. Automated diagnosis of arrhythmia using combination of CNN and LSTM techniques with variable length heart beats[J]. *Computers in Biology and Medicine* 2018;102:278–87 (S0010482518301446-).
- [4] Chen S, Hua W, Li Z, et al. Heartbeat classification using projected and dynamic features of ECG signal [J]. *Biomedical Signal Processing & Control* 2017;31:165–73.
- [5] Kim J, Shin HS, Shin K, et al. Robust algorithm for arrhythmia classification in ECG using extreme learning machine [J]. *Biomed Eng Online* 2009;8(1):31.
- [6] Afkhami RG, Azarnia G, Tinati MA. Cardiac arrhythmia classification using statistical and mixture modeling features of ECG signals [J]. *Pattern Recognition Letters* 2016;70:45–51.
- [7] Yu SN, Chou KT. Selection of significant independent components for ECG beat classification [J]. *Expert Systems with Applications* 2009;36(2):2088–96.
- [8] Wang JS, Chiang WC, Hsu YL, et al. ECG arrhythmia classification using a probabilistic neural network with a feature reduction method[J]. *Neurocomputing* 2013;116:38–45.
- [9] Elhaj FA, Salim N, Harris AR, et al. Arrhythmia recognition and classification using combined linear and nonlinear features of ECG signals[J]. *Computer Methods & Programs in Biomedicine* 2016;127(C):52–63.
- [10] Lin CH, Du YC, Chen T. Adaptive wavelet network for multiple cardiac arrhythmias recognition[J]. *Expert Systems with Applications* 2008;34(4):2601–11.
- [11] Güler İnan, Übeyli Elif Derya. ECG beat classifier designed by combined neural network model[J]. *Pattern Recognition* 2005;38(2):199–208.
- [12] Yildirim O. Ecg beat detection and classification system using wavelet transform and online sequential elm[J]. *Journal of Mechanics in Medicine and Biology* 2019;19:1.
- [13] MIT-BIH arrhythmias database [online]. (Available) <https://www.physionet.org/physiobank/database/mitdb/>.
- [14] Kahali S, Adhikari SK, Sing JK. Convolution of 3D Gaussian surfaces for volumetric intensity inhomogeneity estimation and correction in 3D brain MR image data [J]. *IET Computer Vision* 2018;12(3):288–97.
- [15] Li Q, Peng Q, Yan C. Multiple VLAD encoding of CNNs for image classification [J] ; 2018.
- [16] Tan JH, Hagiwara Y, Pang W, et al. Application of stacked convolutional and long short-term memory network for accurate identification of CAD ECG signals [J]. *Computers in Biology and Medicine* 2018;94:19–26 (S0010482517304201).

- [17] Zhang J, Wu Y. A new method for automatic sleep stage classification[J]. IEEE Transactions on Biomedical Circuits and Systems 2017;1–14.
- [18] Zhang Q, Zhou D, Xuan Z. HeartID: a multiresolution convolutional neural network for ECG-based biometric human identification in smart health applications[J]. IEEE Access 2017(99) 1–1. (PP).
- [19] Rajpurkar P, Hannun AY, Haghighpanahi M, et al. Cardiologist-level arrhythmia detection with convolutional neural networks [J]; 2017 (Rajpurkar P, Hannun A Y, Haghighpanahi M, et al. Cardiologist-Level Arrhythmia Detection with Convolutional Neural Networks[J]. 2017).
- [20] Bengio Y. Learning long-term dependencies with gradient descent is difficult[J]. IEEE Trans Neural Netw 1994;5.
- [21] He K, Zhang X, Ren S, et al. Delving deep into rectifiers: surpassing human-level performance on ImageNet classification[J]; 2015.
- [22] Saxe AM, McClelland JL, Ganguli S. Exact solutions to the nonlinear dynamics of learning in deep linear neural networks[J]; 2013.
- [23] He K, Zhang X, Ren S, et al. Deep residual learning for image recognition[C]//2016 IEEE conference on computer vision and pattern recognition (CVPR). IEEE Computer Society; 2016.
- [24] Testing and, reporting performance results of cardiac rhythm and ST segment measurement algorithms, ANAI/AAMI EC57:1998 standard. Assoc.Adv.Med.Instrum; 1998.
- [25] Luz EJ, Schwartz WR, Cámara-Chávez G, et al. ECG-based heartbeat classification for arrhythmia detection: a survey [J]. Comput Methods Programs Biomed 2016;127 (C):144–64.
- [26] Seena V, Yomas J. A review on feature extraction and denoising of ECG signal using wavelet transform[C]//. International conference on devices, circuits and systems. IEEE; 2014. p. 1–6.
- [27] Bengio Y, Simard P, Frasconi P. The problem of learning long-term dependencies in recurrent networks[C]//. IEEE international conference on neural networks. IEEE; 1993.
- [28] Ioffe S, Szegedy C. Batch normalization: accelerating deep network training by reducing internal covariate shift [J]; 2015.
- [29] Li T, Zhou M. ECG classification using wavelet packet entropy and random forests [J]. Entropy 2016;18(8):285.
- [30] Martis RJ, Acharya UR, Min LC. ECG beat classification using PCA, LDA, ICA and discrete wavelet transform [J]. Biomedical Signal Processing & Control 2013;8(5): 437–48.
- [31] Martis RJ, Acharya UR, Mandana KM, et al. Cardiac decision making using higher order spectra[J]. Biomedical Signal Processing & Control 2013;8(2):193–203.
- [32] Li H, Yuan D, Wang Y, et al. Arrhythmia classification based on multi-domain feature extraction for an ECG recognition system:[J]. Sensors 2016;16(10):1744.
- [33] Sahoo S, Kanungo B, Behera S, et al. Multiresolution wavelet transform based feature extraction and ECG classification to detect cardiac abnormalities[J]. Measurement 2017(108):55–66 (accepted).
- [34] Kiranyaz S, Ince T, Gabbouj M. Real-time patient-specific ECG classification by 1D convolutional neural networks [J]. IEEE Trans Biomed Eng 2015;63(3):664–75.
- [35] Yildirim O. A novel wavelet sequences based on deep bidirectional LSTM network model for ECG signal classification. [J]. Computers in Biology & Medicine 2018;96:189.
- [36] Yildirim O, baloglu U, Tan R, et al. A new approach for arrhythmia classification using deep coded features and LSTM networks [J]. Comput Methods Programs Biomed 2019;176:121–33.
- [37] Acharya UR, Fujita H, Lih OS, et al. Automated detection of arrhythmias using different intervals of tachycardia ECG segments with convolutional neural network[J]. Information Sciences 2017;405:81–90.