Detection of Premature Ventricular Contractions using 12-lead Dynamic ECG based on Squeeze-Excitation Residual Network

Duan Li[1], Tingting Sun[2], Yibai Xue[3], Yilin Xie[4], Xiaolei Chen[5], Jiaofen Nan[6]
School of Computer and Communication Engineering, Zhengzhou University of Light Industry, Zhengzhou, Henan, China[1,2,4,5,6]
Department of Cardiothoracic Surgery, The Fifth Affiliated Hospital of Zhengzhou University, Zhengzhou, Henan, China[3]

Abstract*—Premature ventricular contraction (PVC) is a very common arrhythmia that can originate in any part of the ventricle and is one of the important causes of sudden cardiac death. Timely and rapid detection of PVC on dynamic electrocardiogram (ECG) recording for patients with cardiovascular diseases is of great significance for clinical diagnosis. Furthermore, it can facilitate the planning and execution of radiofrequency ablation. But the dynamic ECGs can be easily contaminated by various noises and its morphological characteristics show significant variations for different patients. Though the deep learning methods achieved outstanding performance in ECG automatic recognition, there are still some limitations, such as overfitting, gradient disappearance or gradient explosion in deep networks. Therefore, a residual module is constructed using the squeeze-excitation method to alleviate the problems. A 20-layer squeeze-extraction residual network (SE-ResNet) containing multiple squeeze-extraction modules was designed for real-time PVC detection on 12-lead dynamic ECG. The algorithm was evaluated using the dynamic 12-lead ECGs in INCART database (168,379 heartbeats in total). The experimental results show that the test accuracy of the method proposed in this paper is 98.71%, and the specificity and sensitivity of PVC are 99.12% and 99.59%, respectively. Under the same dataset and experimental platform, the average recognition accuracy of our proposed method is increased by 0.73%, 1.55%, 2.9% and 1.65% compared with the results obtained by CNN, Inception, AlexNet and deep multilayer perceptron, respectively. The proposed scheme provides a new method for real-time detection of PVC on dynamic 12-lead ECGs. The experiment results show that the proposed method outperforms state-of-the-art methods, and has good potential for clinical applications.

Keywords*—*Dynamic ECG; squeeze-excitation; residual network; premature ventricular contraction

I. Introduction

Premature ventricular contractions (PVCs) are the most common type of arrhythmia and, under certain conditions, can lead to life-threatening heart disease. Electrocardiogram (ECG) is not only a noninvasive and economical tool for routine cardiovascular examination, but also an essential monitoring device in surgical procedures and intensive care units. It is more clinically significant for the diagnosis of PVC. However, it is time-consuming and arduous for cardiologists to analyze many long-term dynamic ECG. Therefore, the automatic detection of PVC on body surface dynamic ECGs can not only improve cardiology workflow efficiency and timely prevent cardiac diseases such as arrhythmia, but also accurately locate the occurrence time and source localization of ventricular premature beats, and then guiding the surgical process such as radiofrequency ablation, etc.

Dynamic ECG is susceptible to various background noises, such as power-line interference, inotropic noise, baseline drift and motion artifact, and its morphological characteristics show significant variations for different patients and under different temporal and physical conditions. Even experienced specialists cannot accurately determine the type of arrhythmias. Machine learning methods such as deep neural networks can more accurately detect arrhythmias such as PVC, and have shown good clinical applications [1].

Deep learning has experienced great breakthroughs in the past decade in many fields, such as image recognition and natural language processing. With the popularity of deep learning and outstanding performance in other fields, researches use deep learning to monitor arrhythmias such as PVCs [2-5]. They transform ECG as one-dimensional time-series signals or two-dimensional signals such as multi-lead ECG beats or time-frequency images as the input of the convolution neural network (CNN). Then conduct layer-by-layer feature extraction and classification. In 2017, Acharya et al.[6] proposed a 9-layer deep CNN to discriminate 5 types of ECG heartbeats and achieved 94.03% accuracy using MIT-BIH arrhythmia database. In 2018, Yildirim et al. [7] designed a new 1D convolutional neural network model (1D-CNN) to recognize 17 different types of long-time dynamic ECG signals. Using the MIT-BIH database, they achieved an overall accuracy of 91.33% for the 17 type arrhythmias. In 2019, Andersen et al. [8] proposed a convolution combined cyclic convolution model, which could search for atrial fibrillation heartbeats from 24-hour dynamic ECG signals, and achieved good results. In 2020, Ullah et al. [9] proposed a two-dimensional (2-D) CNN model to recognize eight types of ECG signals. The model was evaluated on the MIT-BIH dataset and the classification accuracy reached 99.11%. With the deepening of deep network structure, the accuracy of the neural network model will decrease. That is the degradation of neural network. To overcome the problem, deep CNN model with residual structure is developed for ECG arrhythmias detection, which improves classification accuracy [10]. In 2019, Brito et al. [11] proposed a deep learning model based on ResNet architecture. They conducted experiments using MIT-

BIH arrhythmia database and achieved an accuracy of more than 90%. In 2020, Li et al. [12] classified arrhythmias based on deep residual network. The experiments applied to the MIT-BIH arrhythmia database and showed high classification performance with an accuracy of 99.38%. Deep learning has made some achievements in the classification and recognition of PVCs [13, 14]. It has good performance on small evaluation samples and static ECGs, but the accuracy decreases for clinical dynamic 12-lead large data especially for non-equilibrium dataset. Therefore, the evolutionary model and method is crucial to improve the efficiency and effects of PVC detection which will promote the further clinical applications.

Although there are many studies on arrhythmia detection in the literature, there are still various problems such as difficult convergence of deep networks, training cost, and computational complexity. Furthermore, in the literature, most models are trained on relatively clean open-source ECG datasets such as the MIT-BIH database. In this study, considering the advantages and disadvantages of existing technologies, a squeeze-excitation module is constructed which embedded in a residual structure to improve the convergence of the deep network. It aims to improve the non-linear fitting ability of the deep network by reconstructing the hyperplane parameters through the squeeze-excitation operation. The network model uses a feature rescaling strategy, where the importance of each feature channel is automatically obtained by learning, and then the useful features are promoted and the less useful features are suppressed according to its importance. The network model can fully consider the weight of each lead and main wave of ECG signals, and provide a new idea for deep feature extraction of arrhythmia heartbeats. Based on the SE-ResNet model, the performance of the model was evaluated by 168379 12-lead heartbeats from the St Petersburg INCART 12-Lead (INCART) arrhythmia dynamic ECG database. The effectiveness of this method is evaluated by experiments.

The remainder of this paper is organized as follows. Section II described the related work and methods for this study. In Section III, a novel SE-ResNet for detection of premature ventricular beats was implementation. Experimental results are also described. Section IV is discussion and Section V concludes the paper.

II. Method

A. Convolutional Neural Network

CNN is a deep learning algorithm based on artificial neural network structure, trained by a gradient-based optimization algorithm. In contrast to traditional machine learning algorithms, CNN architecture does not need to manually extract features from raw data. Feature extraction and classification are embedded in the architecture, so robust features can be automatically identified from the input data [15]. In general, a CNN consists of multiple back-to-back layers connected in a feedforward manner. As shown in Fig. 1, in the CNN architecture, the main layers are including convolutional layers, pooling layers and a fully-connected layer. Convolutional and pooling layers are responsible for feature extraction, while fully-connected layer is responsible for classification.

B. Squeeze-excitation Residual Network

Bioelectric signals are characterized by individual variability, strong interference, and multi-lead characteristics. Individual variability is reflected in the ECG morphology of different patients with the same disease, and even the difference and translation of characteristic wave directions. In addition, the same patient will also have certain differences in different times and environments. Different leads of the ECG signal reflect the potential transformation of cardiac activity in different parts of the body. The waveforms corresponding to each lead has great difference, and each lead is relatively independent. CNN improves performance by deepening the network structure as much as possible. However, with the increase of CNN depth, namely, the number of network layer increases, the performance of the model tends to saturate and even decline rapidly, which makes the training of deep networks more difficult.

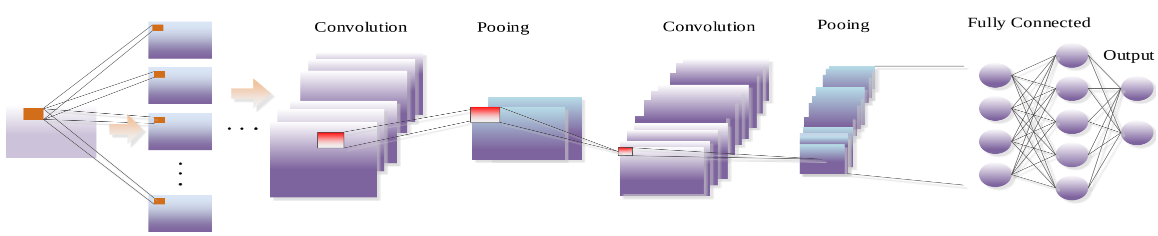

Fig. 1. CNN Architecture.

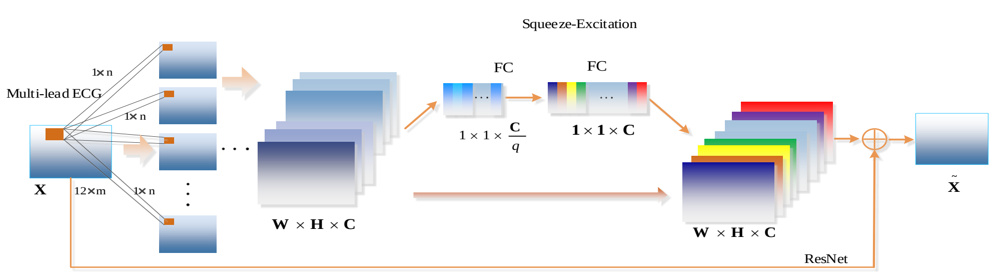

Fig. 2. SE-ResNet Architecture.

Aiming at the above problems, this paper adopts the SE-ResNet model, a deep network architecture with stronger nonlinear fitting ability. The model constructs a squeeze-excitation module which embedded in a residual structure. The network model adopts a feature recalculation strategy, namely, the importance of each feature channel is automatically obtained by learning, and then the useful features are improved and the features that are not useful for the current task are suppressed according to the importance. As shown in Fig. 2, in the squeeze layer, for inputs X = [x_1,x_2, ⋯ ,x_c], where $x_c \in R^{H\times W}$, the simplest aggregation technique, global averaging, is used to generate channel statistics. Formally, the statistic $z \in R^c$ is generated by reducing X by reducing its spatial dimension $H \times W$, and the c-th element of z is calculated by the following equation (1):

$$z_c = \frac{1}{H \times W}\sum_{i=1}^{H}\sum_{i=1}^{W} x_c(i,j) \tag{1}$$

Among them, the H and W represent the height and width of the feature map, respectively, and C represents the number of feature map channels.

In the excitation layer, two fully-connected layers (FC) are used to achieve channel scaling with a reduction rate of q. The dimension of feature data changes from 1 × C to 1 × C / q, and then playback to 1 × C. Finally, the sigmoid activation function is used to scale the data back to the previous data dimension. Since we want to ensure that multiple channels are allowed to be emphasized, a simple gating mechanism with sigmoid activation in equation (2) is employed:

$$s = \sigma(g(z,W)) = \sigma(W_2\sigma(W_1 z)) \tag{2}$$

where σ is the sigmoid function, $W_1 \in R^{\frac{c}{q}\times C}$ and $W_2 \in R^{C\times\frac{c}{q}}$. Here q is a scaling parameter. The final output of the block is obtained by rescaling X with the activations s, as shown in Equation (3) below:

$$\tilde{x}_c = s_c \times x_c \tag{3}$$

Where $\tilde{X}$ = [$\tilde{x}_1$, $\tilde{x}_2$,…, $\tilde{x}_c$] and the channel multiplication between the index quantity sc and the feature map $x_c \in R^{H\times W}$.

Reconstructing hyperplane parameters by squeeze-excitation operation can alleviate the problems of the difficulty tuning of deep network and the nonlinear fitting ability of deep network. The architecture can effectively avoid the effect of traditional ECG signal feature extraction on the subsequent classification accuracy.

III. Detection of Premature Ventricular Contractions based on SE-ResNet

A. Data Sources and Evaluation Metrics

In this experiment, the 12-lead dynamic ECGs from the open INCART Arrhythmia database were used for evaluating the algorithm. The INCART Dynamic Arrhythmia ECG database contains 75 records, sampled at 275 Hz. Each record is about half an hour long and has 12 leads. The original ECG data were collected from patients who were examined for coronary artery disease, and most of them had premature ventricular contractions [16].

Since V6 lead is missing in 102 ECG record, V3 lead is missing in 103 record, and V4 lead is missing in 158 record, the above three ECG records were deleted in consideration of the lead consistency. All heartbeats from the remaining 72 records were used in this experiment. According to our statistics, the data in INCART database included 168379 heartbeats. In order to test the recognition effect of premature ventricular contractions, all cardiac heartbeats were divided into three types: normal heartbeats (N), premature ventricular contractions (V) and other heartbeats (O). The number of normal heartbeats was 143,260, the number of premature ventricular contractions was 19,640 and the number of other heartbeats (premature atrial contractions, supraventricular premature heartbeats and right bundle branch block, etc.) was 5479.

In this study, three metrics were used to evaluate the performance of the proposed classification method: accuracy (Acc), sensitivity (Se) and specificity (Sp), which were defined in formulas (4), (5) and (6) respectively. The calculations are made based on the statistical results of multiple experiments.

$$Acc = \frac{TP+TN}{TP+TN+FP+FN} \tag{4}$$

$$Se = \frac{TP}{TP+FN} \tag{5}$$

$$Sp = \frac{TN}{TN+FP} \tag{6}$$

Among them, TP, FP, TN and FN are true positive, false positive, true negative and false negative, respectively [17, 18].

B. Pre-processing

In contrast, dynamic ECG signals have stronger interference from noise such as motion artefacts, due to factors such as poor skin-to-electrode contact, the effects of breathing and poor contact with the power lines of electrical equipment, et al. Therefore the signal characteristics are chaotic, non-linear and multi-channel. Noises in the ECG signal distort some of its morphological characteristics, which make diagnosis more difficult. So reasonable filtering is very important for subsequent recognition. The useful part of the frequency in the ECG signal mainly ranges from 1 to 40 Hz, and the interest signal is easily submerged in the background noises. The main noise sources include baseline drift, power frequency interference, motion artifact and myoelectric interference [19]. In addition, the low-frequency part of ECG signal contains indicators of malignant arrhythmias such as S-T segment abnormalities, and the high frequency part reflects the amplitude information of main complex wave. In order to keep the morphological characteristics of ECG signal as completely as possible, a wavelet adaptive threshold filtering method is proposed. The algorithm includes three steps: wavelet decomposition, adaptive threshold de-noising and reconstruction. The selection of threshold is adaptive to the signal; therefore, the inherent morphological characteristics of the ECG signal are preserved as much as possible.

The INCART dynamic ECG database contains long-term ECG records with complete annotation of heartbeats. Before applying the model, segmentation of the heartbeat was performed, dividing the long ECG record into heartbeat segments that represent different types, namely, N, V and O. Each heartbeat segment was extracted by selecting a window of 300 samples around the R-peak, which formed by taking 92 points in front of the R-peak and 137 points behind the R-peak, respectively. As each heartbeat segment consists of 230 samples, and the ECG signals are 12 leads, so the size of each beat sample is 12 * 230.

C. SE-ResNet Network Modeling

The architecture of the SE-ResNet network pay more attention to the weights of each lead, and can fully extract the morphological features of the multi-lead ECG waveform. Since the number and complexity of network layers will have a great impact on the training results, we designed SE-ResNet models with different layers and structures, and performed several cross-validation experiments and comparisons. Fig. 3 is the experimental results of the accuracy for each epoch training using the SE-ResNet networks with 8, 12, 16, 20, and 24 layers. As shown in Fig. 3, the 20-layer SE-ResNet achieved better training and testing results, and the 24-layer network performance was comparable to the 20-layer network. Considering computation efficiency and the real-time implementation, especially performed on the embedded processor, we select a 20-layer network for PVC recognition in the following experiments.

In order to optimize the model, we selected the Sgd optimizer, with an initial learning rate of 0.05. To improve the performance of the neural networks, the negative log-likelihood loss (NLLLoss) function and cross-entropy loss (CrossEntropyLoss) function were compared, and the results were shown in Fig. 4. The figure represents the convergence performance for 30 epoch training iterations using different loss functions. As is shown in Fig. 4, using the CrossEntropy loss function, the neural network has more robust stability and better convergence of the training process. Therefore, the CrossEntropyLoss function was selected as the loss function in the following experiments.

In this work, the model architecture of the 20-layer SE-ResNet was designed, as is shown in Fig. 5. The main difference between the proposed network and the original SE-ResNet on ImageNet is that the proposed network uses 1D convolutions instead of 2D convolutions. The input ECG heartbeat is a sample of 12*230 size, and after a convolutional layer with a kernel size of 1*3, followed 9 residual blocks. The residual blocks of the network have 16, 32, and 64 channels respectively. Among them, the first three residual blocks have 16 channels, the second three residual blocks have 32 channels and the remainder three residual blocks has 64 channels.

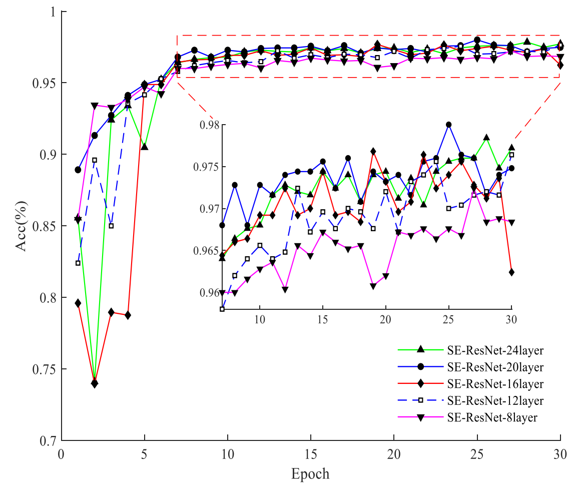

Fig. 3. ECG Recognition Results of different Network Architecture.

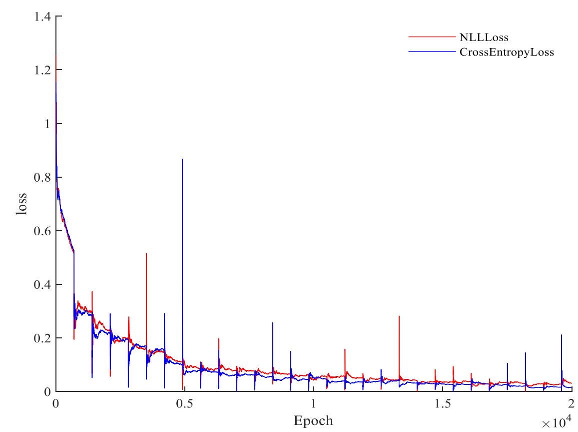

Fig. 4. Comparison of Convergence in Training Process with different Loss Functions.

Each residual block is embedded with squeeze-excitation modules, so a total of 3+ 3+3=9 residual blocks are designed in Fig. 5. In order to accurately extract information from each channel of the ECG heartbeat signal, each residual (short-connect connection) is embedded with a squeeze-excitation module that automatically captures the weight information of each feature channel of the ECG heartbeats in a learning manner, effectively enhancing the features of the useful channels and suppressing information that is not sufficiently useful for the current classification recognition task. Take the first three residual blocks (short-connect) of 16 channels as an example, as shown in Fig. 5, in order to make the best use of the contextual ECG feature information of each channel, channel-level statistics are generated by global averaging pooling. The excitation layer is implemented using two fully-connected layers (FC) for channel scaling, the reduction rate is taken as 4, the dimension of the feature data is changed from 1* 16 to 1* 4 and then replayed as 1* 16. Finally, use the sigmoid activation function to rescale the data back to the dimensions before squeezing. It is equivalent to map the data associated with the input to a set of channel weights, so that the channel features are not limited to the local perceptual field of the convolutional network. Therefore, the context information can be easily understood, and different weights are assigned to the channels. Residual architecture can improve the parameter adjustment ability of the network, namely, the optimization ability. As shown in Fig. 5, there are three 16-channel SE-ResNet modules in total, and each module has two layers of convolution. There are three 32-channel SE-ResNet modules in total, and each module has two layers of convolution. Similarly, there are three 64-channel SE-ResNet modules in total, each with two layers of convolution. Through 19 layers of network transmission, deep feature extraction is completed. Finally, the feature is fed into a fully-connected layer, so the network model has a total of 20 parameter layers.

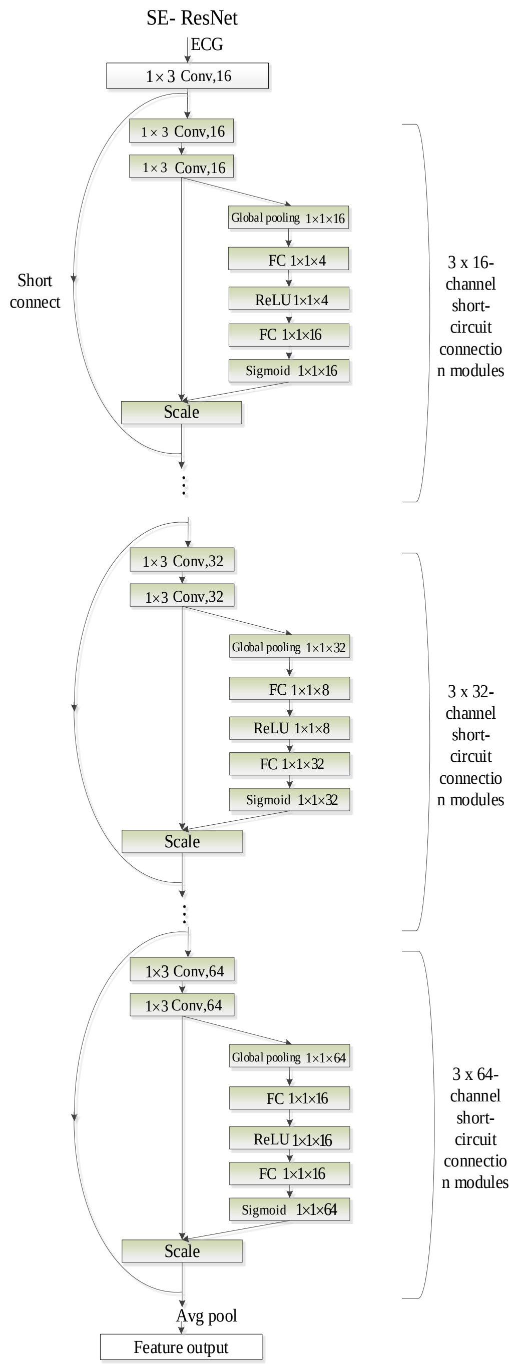

Fig. 5. SE-ResNet Network Model.

D. Analysis of Experimental Results

In order to evaluate the recognition effect of the proposed SE-ResNet model for PVCs recognition based on the INCART dynamic ECG database, this experiment employed a wavelet self-adapting filter to preprocess all 72 lead-consistent ECG records, and segmented 168379 heartbeats according to the R-peak position. We randomly selected 3000 normal heartbeats, 3000 PVCs and 1000 other type heartbeats, respectively from the total 168,379 heartbeats which include 143260 normal heartbeats, 19640 premature ventricular contraction and 5479 other heartbeats. So the total training samples is 7000. Similarly, 1000, 1000 and 500 heartbeats were randomly selected from the remainder data, which were used for cross verification. The remainder 158,879 heartbeats were used for the final evaluation. The classification results were expected to be affected by the heartbeat sample size, and the ratio of training to testing sample size. In general, the lower the ratio of training samples to total samples, the greater the generalization ability of the classifiers. In order to evaluate the network's performance and avoid the occasionality of random sampling, we carried out three random sampling experiments, and the confusion matrix was calculated from the average value of the three experiment results. Then, CNN, Inception, multi-layer perceptron (MLP) and Alexnet of the same complexity were designed. In these experiments, the parameters of the networks were regulated to be best fitted in the classification task. The five networks were trained 30 epochs using the same training samples, and the test results were also compared. The confusion matrix and evaluation metrics are shown in Table I.

Since the experiments were performed three times, three trained models were obtained with exactly the same training dataset. When testing the 12-lead ECG signal, the predicted probabilities of the three models were averaged and finally got the final classification results. The confusion matrix and three different statistical indices, such as Sensitivity (Se), Specificity (Sp), and Accuracy (Acc) were summarized in Table I. As is shown in Table I, it can be seen that the algorithm designed in this paper achieved an accuracy of 98.71% for the detection of the arrhythmias, and the sensitivity (Se) and specificity (Sp) for PVC recognition are 99.12% and 99.59%, respectively. The experimental results showed that the accuracy of the designed SE-ResNet algorithm was improved by 0.73%, 1.55%, 1.65% and 2.9% over CNN, Inception, MLP and AlexNet networks, respectively. The PVC sensitivity (Se) value of the SE-ResNet test was 98.71%, and the sensitivities (Se) of the CNN, Inception, MLP, and Alexnet tests were 92.17%, 97.74%, 94.62, and 95.97%, respectively. It can be seen that the sensitivity of PVC has been significantly improved.

E. Compare Results with other Methods

The PVCs detection results of our proposed method were compared with the recent published research results. These published results on the same dataset are shown in Table II. The detection results were expected to be affected by the heartbeat sample size, the number of classes for classification, and the ratio of training to testing sample size. In general, the lower the ratio of training samples to total samples, the greater the generalization ability of the classifiers, and the fewer the types for classification, the higher the recognition accuracy. In one study, Al Rahhal et al. [20] proposed an electrocardiogram

(ECG) technique based on multi-lead signals and a deep learning architecture. Automatic identification of ECG signals was performed using INCART Arrhythmia Database, which automatically recognized three types: normal heartbeats, PVC and other heartbeats. The overall classification accuracy reached 98.6%, but the sensitivity (Se) of PVC was 91.4%. In a study, Allami [21] used three morphological features and seven statistical features, and also employed the artificial neural network (ANN) classifier for PVC and non-PVC ECG heartbeats recognition. Using 75 ECG records, the classification accuracy and the sensitivity (Se) of PVC they achieved was 95.8% and 93.9%, respectively. In another, Malek et al. [22] developed an improved template matching technique for PVCs and normal heartbeats detecting, by analyzing the maximum value and the correlation coefficients of the maximum and minimum value. The classification accuracy rate was 97.91%, and the sensitivity (Se) and specificity (Sp) of PVC detection were 91.14% and 98.82%, respectively. In general, the proposed SE-ResNet residual network has achieved superior performance on the PVCs recognition experiments using the 12-lead dynamic ECGs. It demonstrates great clinical application prospects.

F. Discussion

Dynamic 12-lead ECG is the gold standard in the detection of arrhythmias. Multi-lead dynamic ECG has strong background noise; there are correlations of ECG leads. Different arrhythmias have corresponding lead characteristics, and the main wave morphological characteristics of some leads are more distinguishable for the specific arrhythmia. Squeeze-excitation network introduces attention mechanism that the importance of each feature channel is automatically obtained by learning. That the useful features are promoted and the less useful features are suppressed according to its importance. The network model can fully consider the weight of each lead and main wave of ECG signals. Therefore, the model can fully extract the morphological features of multi-lead and its main waves, and improve the robustness and generalization ability of the network.

TABLE I. COMPARISON OF PVC RECOGNITION RESULTS ON INCART DATABASE

Method	**Confusion Matrix**				**Evaluation**		
		N	***V***	***T***	***Se***	***Sp***	***Acc***
SE-ResNet	N	137682	496	1082	98.87%	98.22%	98.71%
	V	97	15503	40	99.12%	99.59%	
	T	252	87	3640	91.48%	99.28%	
CNN	N	138079	759	422	99.15%	91.70%	97.98%
	V	862	14416	362	92.17%	99.44%	
	T	766	44	3169	79.64%	99.49%	
Inception	N	135780	1332	2148	97.50%	96.41%	97.16%
	V	142	15287	211	97.74%	98.99%	
	T	563	120	3296	82.83%	98.48%	
MLP	N	136246	1255	1759	97.84%	95.65%	97.06%
	V	204	14798	638	94.62%	99.00%	
	T	650	171	3158	79.37%	98.45%	
Alexnet	N	134939	4202	119	96.90%	88.64%	95.81%
	V	624	15010	6	95.97%	96.99%	
	T	1605	103	2271	57.07%	99.92%	

TABLE II. COMPARISON WITH OTHER METHODS PROPOSED IN THE LITERATURES

Method	**Classification**	**Se**	**Sp**	**Acc**
Al Rahhal et al.[20]	3	91.4%	*	98.6%
Allami[21]	2	93.9%	*	95.8%
Malek et al.[22]	2	91.14%	98.82%	97.91%
Algoritm of this paper	3	99.12%	99.59%	98.71%

In this work, we analysis the characteristics of the long-time dynamic 12-lead ECGs and introduce the squeeze-excitation ResNet model to the real-time PVCs recognition on 12-lead dynamic ECGs. Which is to overcome the and fully extract the multi-lead and multi-dominant complexes. Reconstructing hyperplane parameters by squeeze-excitation operation can alleviate the problems of the difficulty tuning of deep network and the nonlinear fitting ability of deep network. In addition, we considered the influence of SE-ResNet models with different layers and structures on the training results, and through experiments, a 20-layer network was selected as the recognition model. At the same time, the results of two different loss functions under the 20-layer network model are also compared. Finally, the CrossEntropyLoss function was selected as the loss function. The test accuracy of the method proposed in this paper is 98.71%. Under the same dataset and experimental platform, the recognition accuracy of this method is improved compared with CNN, IncexNet and deep multilayer perceptron. The proposed SE-ResNet residual network has achieved superior performance on the PVCs recognition experiments using the 12-lead dynamic ECGs. It demonstrates great clinical application prospects. Then, the PVCs detection results of our proposed method were compared with the recent published research results. Experimental results show that our method has better precision and accuracy than previous studies. Which perform Demonstrate its practical application potential in the medical field.

Although the research results of this paper has achieved good performance, there are still some challenges and study values. First, Bioelectric signals are characterized by individual variability, strong interference, and multi-lead characteristics. The same patient will also have certain differences in different times and environments. Real-time identification of multiple arrhythmias using clinical big data is a challenge work and is of great value in clinical diagnosis. Therefore, we will extend the proposed method to multiple types of arrhythmias on long-term dynamic ECGs, which is of great significance for collecting more clinical ECG data from different patients under different conditions. This also puts forward higher requirements on the robustness and generalization ability of the recognition network. Second, because of the fast and slow changes in heart rhythm, it is not always desirable to use a fixed beat length. It is necessary to study adaptive beat size segmentation to meet different. Third, to achieve higher accuracy, many studies focus on the deep learning trend of making networks deeper and more complex. However, many real-world studies must be performed on computationally limited platforms. We need to consider the computational speed and the computational complexity of the model, as well as its accuracy. Finally, as catheter ablation is an effective therapy for treatment of symptomatic PVCs. And it is important to estimate the targeted anatomic ablation site that prior to the procedure. Based on this study, the localization of the site of origin of a PVCs using 12-lead ECGs is still an interesting and challenge work. Therefore, the further study of us will focus on the location of PVCs. It is important for the planning and execution of the electrophysiological procedure for the catheter ablation and has great clinical application values.

IV. Conclusion

In this study, a 12-lead dynamic ECG PVCs recognition algorithm based on squeeze-excitation residual network is proposed. The squeeze-excitation module is constructed and embedded in the residual structure to improve the performance of the deep network. The hyperplane parameters are reconstructed by squeezing-excitation operations to improve the nonlinear fitting ability of deep networks. A SE-ResNet model based on 20 layers is designed, which overcomes the degradation problem caused by the increase of the network layers when the deep neural network approximates the identity mapping, and ensures the smooth convergence of the network. Experiments of PVCs recognition was performed using 168,379 heartbeats from the INCART dynamic 12-lead ECG database. In the same experimental samples, several popular deep neural network algorithms were compared. The experimental results show that the proposed method effectively improves the overall PVC recognition accuracy on the INCART 12-lead dynamic ECGs, as well as the sensitivity and the specificity have achieved. In the future, we intend to improve the performance of this work with further advanced deep learning techniques. Additional datasets will be added to test the performance of the model to further verify the robustness of the method used. Furthermore, we will further study the localization of the site of origin of a PVC, which is important for the planning and execution of the electrophysiological procedure for the catheter ablation.

Acknowledgment

This work was financially supported by Henan Provincial Science and Technology Research Project (222102210219); Zhengzhou University of Light Industry Doctoral Fund (2018BSJJ031); Henan Provincial University Key Research Project (20A510014).

References

[1] Z. Zhao, X. Wang, Z. Cai, J. Li and C. Liu, "PVC Recognition for Wearable ECGs Using Modified Frequency Slice Wavelet Transform and Convolutional Neural Network," 2019 Computing in Cardiology (CinC), 2019, pp. 1-4, doi: 10.23919/CinC49843.2019.9005872.

[2] S. Hong , Y. Zhou, J. Shang, C. Xiao and J. Sun, "Opportunities and challenges of deep learning methods for electrocardiogram data: a systematic review," Comput Biol Med, vol. 122, pp. 103801, 2020, doi: 10.1016/j.compbiomed.2020.103801.

[3] J. Huang, B. Chen, B. Yao and W. He, "ECG Arrhythmia Classification Using STFT-Based Spectrogram and Convolutional Neural Network," in IEEE Access, vol. 7, pp. 92871-92880, 2019, doi: 10.1109/ACCESS.2019.2928017.

[4] M. Sraitih , Y. Jabrane and A. Hajjam El Hassani, "An Automated System for ECG Arrhythmia Detection Using Machine Learning Techniques," Journal of Clinical Medicine, vol. 10, no. 22, pp. 5450, 2021, doi: 10.3390/jcm10225450.

[5] S. Kiranyaz, O. Avci; O. Abdeljaber, T. Ince, M. Gabbouj and D. J. Inman, "1D convolutional neural networks and applications: A survey," Mech Syst Signal Process, vol. 151, pp. 107398, 2021, doi: 10.1016/j.ymssp.2020.107398.

[6] U. R. Acharya , S. L. Oh , Y. Hagiwara, J. H. Tan, M. Adam, A. Gertych and R. S. Tan, "A deep convolutional neural network model to classify heartbeats," Computers in Biology and Medicine, vol. 89, pp. 389-396, 2017, doi: 10.1016/j.compbiomed.2017.08.022.

[7] O. Yildirim, P. Plawiak , R. S. Tan and U. R. Acharya, "Arrhythmia detection using deep convolutional neural network with long duration ECG signals," Computers in Biology and Medicine, vol. 102, pp. 411-420, 2018, doi: 10.1016/j.compbiomed.2018.09.009.

[8] R. S. Andersen, A. Peimankar and S. Puthusserypady, "A deep learning approach for real-time detection of atrial fibrillation," Expert Systems With Applications, vol. 115, pp. 465-473, 2019, doi: 10.1016/j.eswa.2018.08.011.

[9] A. Ullah, S. M. Anwar and R. M. Mehmood, "Classification of Arrhythmia by Using Deep Learning with 2-D ECG Spectral Image Representation," Remote Sensing, vol. 12, no. 10 pp. 1685, 2020, doi: 10.3390/rs12101685.

[10] P. Rajpurkar , A. Y. Hannun , M. Haghpanahi, C. Bourn and A. Y. Ng, "Cardiologist-Level Arrhythmia Detection with Convolutional Neural Networks," ArXiv, 2017.

[11] C. Brito, A. Machado and A. Sousa, "Electrocardiogram beat-classification based on a ResNet network," Stud Health Technol Inform, vol. 264, pp. 55-59, 2019, doi: 10.3233/SHTI190182.

[12] Z. Li , D. S. Zhou , L. Wan and W. Mou, "Heartbeat classification using deep residual convolutional neural network from 2-lead electrocardiogram," Journal of Electrocardiology, vol. 58, pp. 105-112, 2020, doi: 10.1016/j.jelectrocard.2019.11.046.

[13] A. Y. Hannun, P. Rajpurkar, M. Haghpanahi, G. H. Tison, C. Bourn, M. P. Turakhia and A. Y. Ng , "Cardiologist-level arrhythmia detection and classification in ambulatory electrocardiograms using a deep neural network". Nat Med, vol. 25, pp. 65-69, 2019, doi: 10.1038/s41591-018-0268-3.

[14] B. Hou, J. Yang, P. Wang and R. Yan, "LSTM-Based Auto-Encoder Model for ECG Arrhythmias Classification," in IEEE Transactions on Instrumentation and Measurement, vol. 69, no. 4, pp. 1232-1240, April 2020, doi: 10.1109/TIM.2019.2910342.

[15] Ö. Yıldırım, P. Pławiak, R. S. Tan and U. R. Acharya, "Arrhythmia detection using deep convolutional neural network with long duration ECG signals," Comput Biol Med, vol. 102, pp. 411-420, 2018, doi: 10.1016/j.compbiomed.2018.09.009.

[16] A. L. Goldberger , L. A. Amaral , L. Glass , J. M. Hausdorff, P. C. h. Ivanov, R. G. Mark, J. E. Mietus, G. B. Moody, C. K. Peng and H. E. Stanley, "Physiobank, physiotoolkit, and physionet components of a new research resource for complex physiologic signals," Journal of the American Heart Association, vol. 101, no. 23, pp. 215-220, 2000.

[17] D. Li, R. Shi, N. Yao, F. Zhu and K. Wang. "Real-Time Patient-Specific ECG Arrhythmia Detection by Quantum Genetic Algorithm of Least Squares Twin SVM." Journal of Beijing Institute of Technology, vol. 29, no. 1, pp. 29-37, 2020, doi: 10.15918/j.jbit1004-0579.18156.

[18] L. P. Jin, J. Dong, "Deep Learning Algorithm for Clinical ECG Analysis," Chinese Science: Information Science, vol. 45, no. 3, pp. 398-416, 2015.

[19] T. Tuncer , S. Dogan , P. Plawiak and U. Rajendra Acharya, "Automated arrhythmia detection using novel hexadecimal local pattern and multilevel wavelet transform with ECG signals," Knowledge-Based Systems, vol. 186, pp. 104932, 2019, doi: 10.1016/j.knosys. 2019.104923.

[20] M. M. A. Rahhal, N. A. Ajlan, Y. Bazi, H. A. Hichri and T. Rabczuk, "Automatic Premature Ventricular Contractions Detection for Multi-Lead Electrocardiogram Signal," 2018 IEEE International Conference on Electro/Information Technology (EIT), 2018, pp. 0169-0173, doi: 10.1109/EIT.2018.8500197.

[21] R. Allami, "Premature ventricular contraction analysis for real-time patient monitoring," Biomedical Signal Processing and Control, vol. 47, pp. 358-365, 2019, doi: 10.1016/j.bspc.2018.08.040.

[22] A. S. Malek, A. Elnahrawy, H. Anwar and M. Naeem, "Automated detection of premature ventricular contraction in ECG signals using enhanced template matching algorithm," Biomedical Physics & Engineering Express, vol. 6, pp. 15-24, 2020, doi: 10.1088/2057-1976/ab6995.