Tech Science Press

Classification of Arrhythmia Based on Convolutional Neural Networks and Encoder-Decoder Model

Jian Liu[1,*], Xiaodong Xia[1], Chunyang Han[2], Jiao Hui[3] and Jim Feng[4]

[1]School of Computer & Communication Engineering, University of Science and Technology Beijing, Beijing, 100083, China
[2]Beijing Satellite Navigation Center, Beijing, 100093, China
[3]The Institute of NBC Defense, Chinese PLA Army, Beijing, 102205, China
[4]Amphenol Global Interconnect Systems, San Jose, CA 95131, US
*Corresponding Author: Jian Liu. Email: liujian@ustb.edu.cn

Abstract: As a common and high-risk type of disease, heart disease seriously threatens people's health. At the same time, in the era of the Internet of Thing (IoT), smart medical device has strong practical significance for medical workers and patients because of its ability to assist in the diagnosis of diseases. Therefore, the research of real-time diagnosis and classification algorithms for arrhythmia can help to improve the diagnostic efficiency of diseases. In this paper, we design an automatic arrhythmia classification algorithm model based on Convolutional Neural Network (CNN) and Encoder-Decoder model. The model uses Long Short-Term Memory (LSTM) to consider the influence of time series features on classification results. Simultaneously, it is trained and tested by the MIT-BIH arrhythmia database. Besides, Generative Adversarial Networks (GAN) is adopted as a method of data equalization for solving data imbalance problem. The simulation results show that for the inter-patient arrhythmia classification, the hybrid model combining CNN and Encoder-Decoder model has the best classification accuracy, of which the accuracy can reach 94.05%. Especially, it has a better advantage for the classification effect of supraventricular ectopic beats (class S) and fusion beats (class F).

Keywords: Electroencephalography; convolutional neural network; long short-term memory; encoder-decoder model; generative adversarial network

1 Introduction

In recent years, with improvement of the standard of living, people are paying more and more attention to their own health. Early detection and treatment of the disease puts higher demands on medical workers and corresponding equipment. Among various diseases, heart disease is not only more likely to happen, but also poses a greater threat to human life. As a common examination method, electroencephalography (ECG) can reflect the state of the heart at every moment, which

is an important reference for doctors to diagnose. However, the identification of ECG still requires experienced medical staff to accurately diagnose the pathology. Therefore, the use of intelligent devices and related algorithms to achieve real-time monitoring of the patient's heartbeat state has a strong practical significance, which is a research hotspot for scholars.

The realization of the traditional automatic arrhythmia classification algorithm can be divided into four parts: data acquisition, preprocessing, feature extraction and classification. In the data acquisition part, we can collect the ECG signals through relevant medical diagnostic equipment. However, due to the need to protect the privacy, the open arrhythmia data sets are often used in most research. The data preprocessing part mainly completes the analysis and filtering of the related noise in the ECG signal, which can improve the efficiency of subsequent classification. And the common types of the noise include baseline drift, power frequency interference and emg interference [1]. In the feature extraction part, we mainly complete the wave form positioning and the feature value extraction of the ECG signals. Since directly providing the input data of the classifier, this part has a crucial impact on the classification accuracy of the classifier. The methods, which are commonly used in this part, include morphological and temporal features, wavelet transforms [2], principal component analysis (PCA) [3,4], et al. The classifier is the kernel as well as the most direct technique of the automatic arrhythmia classification algorithm. And the commonly used algorithms include support vector machine (SVM) [5,6], Particle Swarm Optimization (PSO) [7], logistic regression algorithm, neural network and related hybrid algorithms.

In the traditional classifier implementation, the output characteristic value of the former three steps is often taken as the input of the classifier. And the classification model is constructed by relevant algorithm to complete the automatic classification of arrhythmia. Salam et al. [8] use three algorithm, which are discrete wavelet transform (DWT), adaptive least mean square (ALMS) and SVM, to detect ST segments, QRS complexes, and R peaks to diagnose arrhythmia. Albuquerque et al. [9] use the supervised learning techniques including Optimum-Path Forest (OPF) classifier and SVM classifier to detect arrhythmia. The simulation results show that the SVM classifier has better accuracy, while the OPF classifier has an advantage in computing time. Mathews et al. [10] use the restricted Boltzmann machine (RBM) and the deep belief network (DBN) to classify arrhythmia on a single-lead ECG. By selecting the appropriate parameters, RBM and DBN can complete classification at a low sampling rate of 114 Hz. And the high average recognition accuracy of indoor ectopic beat (93.63%) and supraventricular ectopic beat (95.57%) has been achieved. Due to the feature of the traditional methods, the quality of the extracted feature values greatly affects the classification accuracy of the subsequent classifiers. Since the large error of manually extracting eigenvalues has a great negative impact on classification accuracy, the classification effect of traditional classification algorithms is limited. For the image processing, some related algorithms have been well applied in some scenarios [11–14], and deep learning has been widely used in research of image classification and various disease diagnosis fields [15]. So the arrhythmia classification algorithm based on deep learning, which has good feature extraction ability, has become a research hotspot in recent years [16–18].

According to the characteristics of data, deep learning can accomplish two functions of feature extraction and classification, which can avoid complex feature extraction engineering to a certain extent and reduce the impact of manually extracting eigenvalues on classification effects. Ali Isin et al. [19] use the AlexNet as a feature extractor and feed the extracted features into a simple backpropagation neural network to perform the final classification. The highest correct recognition rate obtained by the experiment is up to 98.51%, while the test accuracy is up to 92%. Serkan Kiranyaz et al. [20] propose a patient-specific ECG heartbeat classifier based on the adaptive one-dimensional convolutional neural networks (1-D CNN) model, which has a good classification effect. Rajendra

Acharya et al. [21] develop a 9-layer deep CNN to automatically identify five different categories of heartbeats, and achieve an accuracy of 94.03% and 93.47%, respectively, in the original and noise-free ECG.

However, the above methods do not take advantage of the time series characteristics of the ECG. Owing to Long Short-Term Memory (LSTM) is an improvement on Recurrent Neural Network (RNN), it makes good use of the characteristics of timing and has a good application in solving sequence problems [22]. Many scholars have combined CNN and LSTM to build an arrhythmia classification model. Philip Warrick et al. [23] use a combination of the CNN and LSTM units to classify Atrial fibrillation (AF). And using the ten-fold cross-validation method, the F-measurement value of the algorithm is up to $0.83.10 \pm 0.015$. Shu Lih Oh et al. [24] use CNN and LSTM models to diagnose five kinds of heartbeats, including normal sinus rhythm, left bundle branch block (LBBB), right bundle branch block (RBBB), atrial premature beat (APB) and ventricular premature beats (PVC), and the accuracy, sensitivity and specificity of the simulation results reach about 98.10%, 97.50% and 98.70%, respectively. Although domestic and foreign scholars have achieved certain results in the field of physiological signal research, due to the imbalance, time series and patient specificity of ECG signal data, big differences of the classification accuracy and sensitivity between different categories exist in the inter-patient arrhythmia classification model.

In this paper, we mainly build models based on how to solve these three problems. For the imbalance of data, Generative Adversarial Networks (GAN) can realize the learning of data distribution in the game between generator and discriminator, and it already achieves better results in the field of image data enhancement [25,26]. Therefore, we consider applying the GAN to the data enhancement process of ECG signals in this paper. In addition, CNN and sequence to sequence model have achieved good application results in the fields of image and speech signal processing [27]. ECG signals are essentially a sequence of time and have some similarities with speech signals, so some speech processing algorithms can be used for arrhythmia classification. For the timing characteristics of ECG signal data, we mainly study the network model and architecture of CNN in arrhythmia classification, and use encoder-decoder model to optimize the algorithm in this paper. The contribution of this paper is to apply GAN to data augmentation and construct the arrhythmia classification model based on LSTM. By comparing the effects of arrhythmia classification under different models, the effect of the inter-patient arrhythmia classification model constructed in this paper was verified.

The rest of this paper is organized as follows. The database and the data preprocessing are presented in Section 2. The theoretical description of the arrhythmia classification model constructed in this paper are presented in Section 3. And then the experimental results are presented in Sections 4. Finally, the conclusion is discussed in Section 5.

2 Prepared

2.1 Datasets

Since the patient information has strong privacy, we use the MIT-BIH arrhythmia database as the data set for training and testing in this paper. The database was collected from 48 different patients and 48 heartbeat records were recorded, each approximately lasting for 30 minutes [28]. Due to the complexity of the type of arrhythmia, according to the recommendations of the Advancement of Medical Instrumentation (AAMI), ECG beats can be divided into five types, which are N (normal beats), S (supraventricular ectopic beats), V (ventricular ectopic beats), F (fusion beats) and Q (unclassifiable beats).

As is shown in Fig. 1, the heartbeat waveform mainly consists of P wave, T wave, R peak and QRS wave group. The former four classes of ECG heartbeats have their own waveform characteristics. For the first class N, the R wave peak is upward, the peak value is large, and the waveforms of the P wave and T wave are clear. For the second class S, the P wave appears earlier and has a longer R interval. For the third class V, the waveform has no P wave, but the QRS wave group is wide, and the position is advanced. The fourth class F is a fusion form of S and V. Each type of heartbeat has obvious features, which is beneficial to the corresponding feature extraction for the CNN model. Since the amount of data of class Q differs greatly from other types of heartbeats, in this paper, we only classify the former four classes.
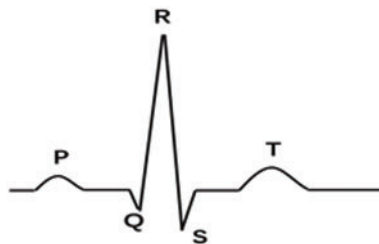

Figure 1: The waveform structure the ECG heartbeat

2.2 Preprocessing

In this section, we mainly complete two aspects of work, including heart beat segmentation and dataset partitioning.

In the heart beat segmentation part, we mainly realize the interception of the individual heartbeat. In the MIT-BIH database, each piece of data is in units of records and contains multiple heart beats. In this paper, we mainly focus our research on the classification of arrhythmia. Therefore, the model is constructed directly based on the heartbeat, which is beneficial to the training of the model by increasing the total data amount.

In the MIT-BIH arrhythmia database, the annotation file contains the manually labeled R-peak position, which is convenient for the researchers to segment the heartbeat. We take this file as a reference, and complete the heart beats segmentation by taking the corresponding data from the left and right. The specific implementation is described as follows. The definition 'R-R interval' is the sample between two adjacent R peak positions. So the samples in this interval can be divided into two parts to obtain a sample of the individual heartbeat. We take the R peak position of each heart beat as the center, and 45% of the samples are collected from the left interphase, while 55% of the samples are collected on the right side to complete the segmentation of the heartbeat, which is shown in Fig. 2.
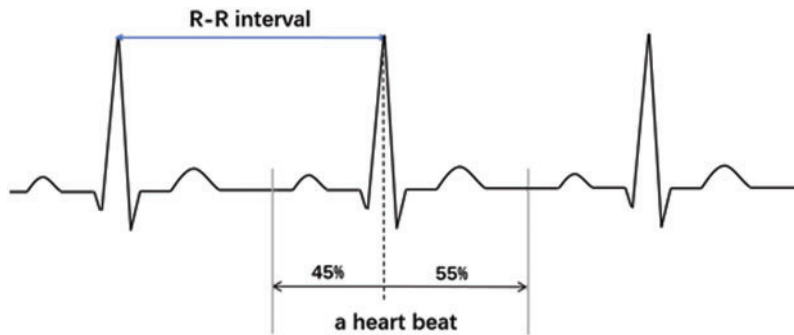
**Figure 2:** The way of heart beats segmentation

In the data set division, we divide the data into training set and testing set based on the inter-patient heartbeats, which can improve the scalability of the classification model. We divide all records into two categories according to the existing proposed data division method [29], one category is used as training set for model training, and the other is used as test set for testing. The specific division is as follows. Dataset 1 (DS1) used as training set contains the following records: 101, 106, 108, 109, 112, 114, 115, 116, 118, 119, 122, 124, 201, 203, 205, 207, 208, 209, 215, 220, 223 and 230. While Dataset 2 (DS2) used as test set contains: 100, 103, 105, 111, 113, 117, 121, 123, 200, 202, 210, 212, 213, 214, 219, 221, 222, 228, 231, 232, 233 and 234. According to this method, the data of some patients is used for model training, and the data of different patients is used for model testing, which provides data guarantee for the scalability of the arrhythmia classification.

In addition, in order to improve the generalization of the algorithm, we use a linear function to normalize each heartbeat, which is beneficial to the training of the model. Finally, a series of $1*64$ heartbeat data is obtained, which is directly used for model construction.

3 Method

In this paper, a hybrid model combined with CNN and encoder-decoder model is designed based on the characteristics of ECG signals, and GAN is used for data enhancement. Using the good feature extraction ability of the CNN and the time series features extraction ability of the LSTM, the classification model is constructed, which is shown in Fig. 3.

The preprocessing part has already been described in Section 2, and the principles of the later parts of the model are described in this section. Since most of the networks used in this paper are constructed based on CNN and LSTM, in this section, we first briefly introduce the basic principles of CNN and LSTM.
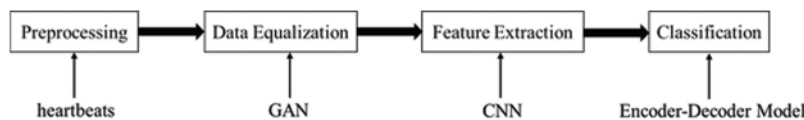
Figure 3: The process structure of classification model

3.1 Convolution Neural Network

CNN is mainly composed of convolution layer, pooling layer and fully connected layer. The feature value is extracted by convolution layer and pooling layer, and finally the classification output result is obtained through the fully connected layer.

The convolutional layer is the core of the CNN, and the feature extraction is mainly performed by the convolution kernel. By convolving the input data with the kernel function, the corresponding feature map is obtained. Different convolution filters correspond to different feature values, and finally the output of the layer is obtained by the activation function. The specific expression result of the convolutional layer is represented by the weight k and the bias b. Taking the output result of the j-th neuron in the l-th convolution layer as an example, the implementation is expressed as follows Eq. (1):

$$y_i^l = f\left(\sum_{i \in N_j} k_{ij}^l * y_i^{l-1} + b_j^l\right) \tag{1}$$

where k_{ij}^l represents the convolution kernel between two neurons, b_j^l is the bias of the current layer, y_i^l is the output of this layer, y_i^{l-1} indicates the input of the current layer, which is also the output of the previous layer, and $f(\cdot)$ represents activation function. The function of the activation function is to simulate the human nervous system, that means it only reacts to the input of some neurons.

The pooling layer, also known as the downsampling layer, is mainly used to reduce data parameters. The pooling operation achieves dimensionality reduction of data and combines low-level local features into higherlevel features. Similarly, taking the j-th neuron in the l-th pooling layer as an example, the output of the pooling layer can be defined as Eq. (2)

$$s_j^l = down(y_j^l) \tag{2}$$

where $down(\cdot)$ indicates the downsampling function and y_j^l is the output of the convolutional layer. In this paper, we use the maximum downsampling as the pooling layer.

The function of the fully connected layer is to integrate the abstract features of the former layer and then send the output values to the classifier for classification. After flattening into one-dimensional, the feature data is directly sent into the fully connected layer, and the mapping between the feature value and the output category is completed. For fully connected layer, a commonly used model is the Multi-Layer Perception (MLP). In addition, in the improvement of the model, the SVM is usually used to replace the MLP to improve the classification effect of the model.

3.2 Long Short-Term Memory

Since the LSTM is an improved network based on RNN, it mainly adds three logic gates, including the forgot gate, the input gate and the output gate. Each hidden state h_{t-1} is determined by the current input x_t and the previous time state h_{t-1} under the action of three gates, which can be explained as follows.

The forgot gate primarily determines how much of the previous moment output is retained as the input to the current state. The specific expression formula can be expressed as Eq. (3)

$$f_t = \sigma(w_f * [h_{t-1}, x_t] + b_f) \tag{3}$$

where w_f and b_f are the weight matrix and the bias of the forgot gate, respectively. The processing of the state and the input is realized by merging the previous time state h_{t-1} with the current time input x_t.

The input gate primarily determines how much of the current time input value is retained as the actual input to the current state, which can be expressed as follows Eqs. (4)–(6).

$$i_t = \sigma(w_i * [h_{t-1}, x_t] + b_c) \tag{4}$$

$$\tilde{C} = \tanh(w_c * [h_{t-1}, x_t] + b_c) \tag{5}$$

$$c_t = f_t * c_{t-1} + i_t * \tilde{C}_t \tag{6}$$

where i_t indicates how much the information of last state is kept and $\tilde{C}_t$ indicates the new information generated by the input.

The output gate mainly determines how much the output state at this moment is reserved as the output of the next-time state, which can be expressed as Eqs. (7) and (8)

$$o_t = \sigma(w_o * [h_{t-1}, x_t] + b_o) \tag{7}$$

$$h_t = o_t * tanh(c_t) \tag{8}$$

Similar to the CNN, the LSTM training method is to propagate the error through the loss function, and calculate the partial derivative of weights to obtain the final classification model.

3.3 Data Equalization

In data equalization part, we mainly solve the problem of large difference in the amount of data between different types of heart beats. According to the data set division method and the AAMI arrhythmia classification method, we can know that the amount of data between the categories are quite different, which make it difficult to train the model. In order to improve the accuracy of model training and achieve a balanced number of heartbeats as much as possible, we use the GAN to realize data equalization.

The GAN model is mainly composed of the generator G and the discriminator D. The generator receives the random noise z, and then generates a certain distribution of fake data, which can be written as $G(z)$. The discriminator mainly completes the identification of the fake data and determines whether it is real or fake. Since the goal is to achieve learning of the distribution from the original data, the network is implemented by a dynamic game of two sub-networks. The main purpose of the generator is to generate fake data close to the original distribution as much as possible, so that the discriminator cannot judge its authenticity. Assuming that the distribution of the true data x is $P_t(x)$, the noise distribution is $P_z(z)$, where $P_t(x)$ is unknown and $P_z(z)$ is known. Ideally, the GAN can finally achieve the distribution of $G(z)$ as close as possible to $P_t(x)$, which means the generator G maps the noise onto the original data. The GAN constructs the loss function by cross entropy loss, which can be written as follows Eq. (9).

$$V(D, G) = E_{x \sim P_t(x)}[ln(D(X)] + E_{z \sim P_z(z)}[ln(1 - D(G(z))] \tag{9}$$

In the training process, we use the gradient descent method to optimize the G and D alternately, to realize the balance in the continuous game.

Since there is no clear definition of the specific implementation of G and D in the GAN, various ways can be formed in the implementation. In this paper, we use the Deep Convolutional Generative

Adversarial Networks (DCGAN) for ECG data, which combining the GAN with CNN. The network is constructed by using the corresponding CNN in the specific implementation of *G* and *D*, respectively.

Since there is no clear definition of the specific implementation of *G* and *D* in the GAN, various ways can be formed in the implementation. In this paper, we use the Deep Convolutional Generative Adversarial Networks (DCGAN) for ECG data, which combining the GAN with CNN. The network is constructed by using the corresponding CNN in the specific implementation of *G* and *D*, respectively.

For the discriminator D, a three-layer one-dimensional convolution operation is used to obtain the convolution feature, and the corresponding output is obtained through the logistic regression layer. For the generator G, the model uses the full connection, upsampling and convolution operations for the input noise vector to learn the original data distribution. The composition of the model is shown in the Fig. 4.
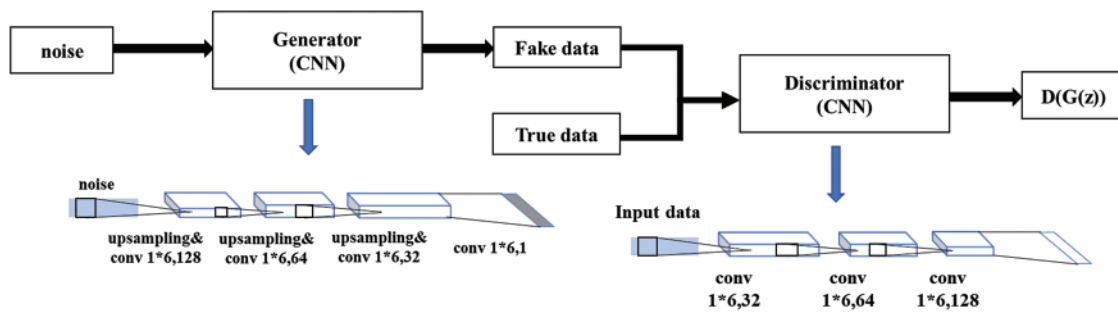
Figure 4: The structure of the DCGAN model

In this paper, we set the generator updates once when the discriminator updates 5 times. The DCGAN model is constructed and trained for the class S, class V, and class F heartbeat data, respectively. And using the trained model, we can achieve the generation data, of which the distribution is as similar as the original data to realize the relative balance of the classification data. The data distribution of DS1 after using DCGAN isshown in Tab. 1.

Table 1: The distribution of DS1 after using DCGAN

Class	Original data	Enhanced data
N	45866	45866
S	944	44991
V	3788	44992
F	415	44989

3.4 Feature Extraction

As a commonly used deep learning algorithm, CNN mainly realizes feature extraction and classification of target data. Therefore, the difference in waveform between different heartbeat categories provides data guarantee for feature extraction.

Through the previous design and simulation, we design a simple CNN model in this paper. The model only contains four layers of convolution and pooling operations, simplifying the complexity of the model without affecting the effect of classification. The CNN structure used in this paper is shown

in Tab. 2. We mainly use the feature extraction ability of CNN to complete the feature extraction function of the classification model.

Table 2: The model structure and the parameters of the CNN network

| Layer function	Kernel	Stride	Number of kernels
Convolution1	$1 * 3$	1	32
Maxpooling1	$1 * 2$	2	32
Convolution2	$1 * 3$	1	64
Maxpooling2	$1 * 2$	2	64
Convolution3	$1 * 3$	1	128
Maxpooling3	$1 * 2$	2	128
Convolution4	$1 * 3$	1	256
Maxpooling4	$1 * 2$	2	256

3.5 Classifier

The encoder-decoder model is a commonly framework structure which is widely used in the solution of the sequence to sequence problem. The ECG signal is essentially a time-based sequence signal, which is similar to the speech signal. So the encoder-decoder model can be used as a classification model of the ECG signal due to its good practicability to the time series. In this paper, for the feature information extracted by CNN, a classifier based on encoder-decoder model is designed, and its time series features are further considered. The specific principle and implementation of the model are described below.

The encoder-decoder model consists of three parts, including the encoder, the semantic vector and the decoder. The encoder mainly completes encoding of the input information, and generalizes the information into a memory mode as a semantic vector. The decoder takes the semantic vector as the initial input state and completes the semantic transformation through the corresponding decoding algorithm. The specific implementation of the encoder and decoder are flexible. The optional models include CNN, RNN, Bi-directional Recurrent Neural Network (BiRNN) and LSTM, etc. There is no uniform specification for the encoder and decoder model algorithms, so a variety of encoder-decoder models can be constructed by different combinations of algorithms.

In this paper, we combine the CNN with the encoder-decoder model based on LSTM to construct the classification model. The original input signal is passed through the CNN model, and the feature extraction is performed to generate the corresponding feature variable as the input of the encoder-decoder model. And the classification of the arrhythmia is realized by the LSTM model. The structure of this model is shown in the Fig. 5. The relevant principles of the input and output models of each part are described as follows.

Encoder: The LSTM model structure is used. The input is the training result of the fully connected layer in the CNN model, which means the feature vector. And the output is the semantic vector representation of the corresponding target value, which is used to initialize the decoder input.

C: Represents a semantic vector. The encoder encodes any length of sequence information into a fixed length of context information vector as input to the decoder.
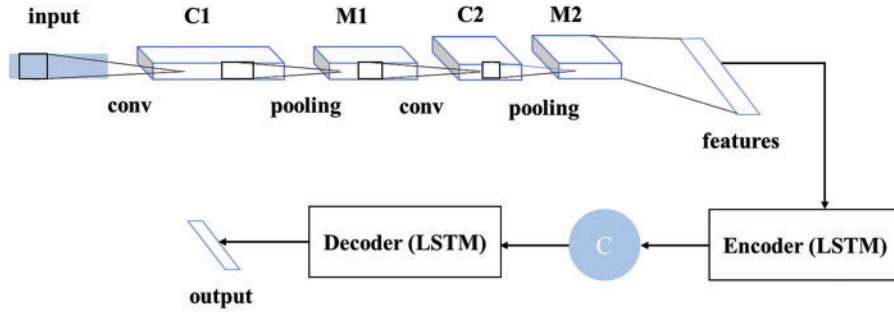
Figure 5: The structure of CNN with the encoder-decoder model

Decoder: The LSTM model structure is used. The input is the semantic vector output representation of the encoder, and the output is the vector corresponding to the target. Then, it is converted to the probability value by the sofmax function, and the different types of the arrhythmia is generated one by one.

4 Experiments and Results

In this paper, the ECG signals are classified abnormally by a hybrid model based on encoder-decoder model. At the same time, in order to compare the classification effect of the model, three different algorithms are used to classify ECG arrhythmia. The three algorithms include one-dimensional CNN model (1-D CNN), a combination of CNN and SVM (CNN+SVM), and the encoder-decoder model combined with CNN model (CNN+ED). We use the MIT-BIH arrhythmia database as the input data set. Besides, in order to measure the effect of data equalization, 1-D CNN is used as the classifier to verify its effect based on the results, in a condition of whether using data equalization or not. After applying the preprocessing technique mentioned above, the data are sent into the models for training and testing.

In the evaluation index, we measure the simulation results of four types of arrhythmia in four standard metrics based on the confusion matrix, which includes classification accuracy (ACC), sensitivity (TPR), specificity (TNR) and positive predictivity (PPV). At the same time, we use F1-score to evaluate classification effects of different types. The respective definitions of these five metrics adopting true positive (TP), true negative (TN), false positive (FP) and false negative (FN) are expressed as follows Eqs. (10)–(14).

$$ACC = \frac{TP + TN}{TP + TN + FP + FN} \tag{10}$$

$$TPR = \frac{TP}{TP + FN} \tag{11}$$

$$TNR = \frac{TN}{TN + FP} \tag{12}$$

$$PPV = \frac{TP}{TP + FP} \tag{13}$$

$$F1 = \frac{2PPV * TPR}{PPV + TPR} \tag{14}$$

Firstly, Tab. 3 reflects classification effect under the condition that whether GAN is used as the data equalization method in the arrhythmia classification model. The data in the table are the average of the four classification results.

Table 3: The results of average classification effects from four indicators using data equalization or not

Method	ACC	TPR	TNR	PPV
Not use data equalization	93.53	50.68	89.76	42.79
Use data equalization	94.45	53.23	90.41	46.56

The results in the Tab. 3 shows that the TPR and PPV indicator using data balance has increased significantly. Since the TPR can reflect the classification effect under the condition of data imbalance, it can be known that using GAN as the data equalization method is beneficial to improve the classification efficiency of the model. Therefore, data equalization is used for the comparison of different classification models.

Through the training and test process of the model, the simulation results of the three models are shown in Tab. 4, where the accuracy index is used to measure the integral test results of the model.

By comparing the average results of various metrics corresponding to the three models, it can be known that the basic CNN network model has better average accuracy and specificity, both of which are about 90%. While 1-D CNN model has worse average sensitivity and predictivity. The CNN+SVM model has improved the integral classification accuracy, but for individual indicators, the improvement effect is not obvious. The CNN+ED model has a greater degree of improvement in all the metrics, which can reach more than 70%. The average accuracy and specificity are improved by about 3% and 6%, respectively, while the average sensitivity and predictivity increase the most, which can reach up to 20%.

Table 4: The results of average classification effects from four indicators using three different models

Method	ACC	TPR	TNR	PPV
1-D CNN	94.45	53.23	90.41	46.56
CNN+SVM	96.20	48.77	88.72	50.35
CNN+ED	97.03	78.64	96.46	71.56

Tab. 5 shows the comparison between the traditional method SMOTE and the data enhancement method in this paper under the CNN + ED model. The experiment shows that GAN is better as the data enhancement method. The performance of precision PPV and sensitivity TPR is improved by about 20%, and the specificity is improved by 3%. The data enhancement method in this paper improves the accuracy of classification.

**Table 5:** Comparison with traditional data enhancement methods

Method	ACC	TPR	TNR	PPV
SMOTE	97.53	59.06	92.21	66.22
DCGAN	97.03	<u>78.64</u>	<u>96.46</u>	<u>71.56</u>

Tab. 6 reflects the classification indicators of four types of ECG signals under the CNN+ED model. It can be known that when classifying the arrhythmia, the classification effect of the class N and class V are better, while the class S and class F are worse. By comparing the dataset distribution, we can know that the dataset obtained by GAN has more uncertainty which may lead to a worse classification effect.

Table 6: The results of four classes classification effects using CNN+ED model

CLASS	ACC	TPR	TNR	PPV
N	94.22	94.62	91.00	98.85
S	94.83	85.73	95.18	40.6
V	99.85	97.79	99.99	99.87
F	99.20	36.41	99.68	46.94
Ave	97.03	78.64	96.46	71.56
Accuracy		94.05		

Besides, Fig. 6 shows the classification effect of different algorithms in different classification categories by the index of F1-score. By comparing the classification effects of the three models, the results of the CNN+ED model for the class S and class F are both significantly improved, while the classification effects of the other two categories have not improved much.
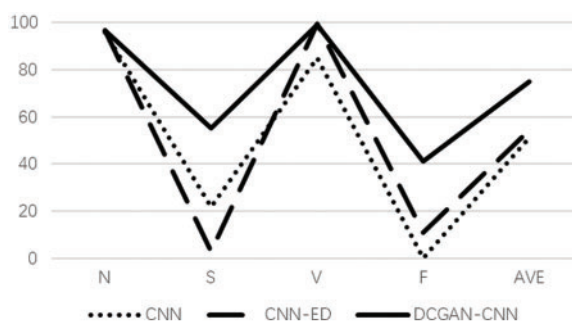
Figure 6: F1 values of four different types of arrhythmia under three models

Tab. 7 shows the comparison of the proposed method with other inter-patient ECG classification methods. It can be seen from the table that the network structure and data enhancement method used in this paper have good classification results.

Table 7: Comparison of this method with other methods

	Method	ACC
A Mustaqeem l [5]	SVM	92.07
S. Raj [7]	PSO+SVM	89.10
U. R. Acharya [19]	1D-CNN	93.47
The proposed method	CNN-ED	93.80
The proposed method	DCGAN-CNN-ED	94.05

5 Citations

In this paper, a hybrid model combined with CNN and encoder-decoder model is designed for the classification of arrhythmia, and GAN method is used as data equalization method. The inter-patient heartbeat data processing results are used to verify the classification effect. The simulation results show that the classification model constructed in this paper has a good classification effect, especially in the class S and class F. And the accuracy of this model is as high as 94.05%. The CNN models used in this paper are all four layers. And the accuracy can be improved by combining them with other learning models, which can avoid complex convolution operations to some extent. In a word, in this paper we had basically completed the classification of arrhythmia under the premise of automatically extracting the characteristic parameters, which is conducive to the auxiliary treatment of heart disease.

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

[1] S. Satheeskumaran and M. Sabrigiriraj, "VLSI implementation of a new LMS-based algorithm for noise removal in ECG signal," *International Journal of Electronics*, vol. 103, no. 6, pp. 975–984, 2016.

[2] J. Park, M. Z. A. Bhuiyan, M. Kang, J. Son and K. Kang, "Nearest neighbor search with locally weighted linear regression for heartbeat classification," *Soft Computing*, vol. 22, no. 4, pp. 1225–1236, 2018.

[3] M. Abu-Alhaija and N. M. Turab, "Automated learning of EDG streaming data through machine learning internet of things," *Intelligent Automation & Soft Computing*, vol. 32, no. 1, pp. 45–53, 2022.

[4] T. Bekıryazici and H. Gürkan, "ECG Compression method based on convolutional autoencoder and discrete wavelet transform," in *2020 28th Signal Processing and Communications Applications Conf. (SIU)*, Gaziantep, Turkey, pp. 1–4, 2020.

[5] A. Mustaqeem, S. M. Anwar and M. Majid, "Multiclass classification of cardiac arrhythmia using improved feature selection and SVM invariants," *Computational and Mathematical Methods in Medicine*, vol. 2018, no. 1, pp. 1–10, 2018.

[6] Y. Liu, L. Dong, B. Zhang, Y. Xin and L. Geng, "Real time ECG classification system based on DWT and SVM," in *2020 IEEE Int. Conf. on Integrated Circuits, Technologies and Applications (ICTA)*, Nanjing, China, pp. 155–156, 2020.

[7] S. Raj, K. C. Ray and O. Shankar, "Cardiac arrhythmia beat classification using dost and pso tuned svm," *Computer Methods and Programs in Biomedicine*, vol. 136, pp. 163–177, 2016.

[8]   K. A. Salam and G. Srilakshmi, "An algorithm for ecg analysis of arrhythmia detection," in *2015 IEEE Int. Conf. on Electrical, Computer and Communication Technologies (ICECCT)*, Coimbatore, India, pp. 1–6, 2015.

[9]   V. Albuquerque, T. M. Nunes, D. R. Pereira, E. Luz, D. Menotti *et al.,* "Robust automated cardiac arrhythmia detection in ECG beat signals," *Neural Computing and Applications*, vol. 29, no. 3, pp. 679–693, 2018.

[10] S. N. Mathews, C. Kambhamettu and K. E. Barner, "A novel application of deep learning for single-lead ECG classification," *Computers in Biology and Medicine*, vol. 99, no. 7, pp. 53–62, 2017.

[11] M. B. Nejad and M. E. Shiri, "A new enhanced learning approach to automatic image classification based on salp swarm algorithm," *Computer Systems Science and Engineering*, vol. 34, no. 2, pp. 91–100, 2019.

[12] Z. Xia, N. N. Xiong, A. V. Vasilakos and X. Sun, "Epcbir: An efficient and privacy-preserving content-based image retrieval scheme in cloud computing," *Information Sciences*, vol. 387, no. 1, pp. 195–204, 2017.

[13] Weize Sun, Shaowu Chen, Lei Huang, Hing Cheung So and Min Xie, "Deep convolutional neural network compression via coupled tensor decomposition," *IEEE Journal of Selected Topics in Signal Processing*, vol. 15, no. 3, pp. 603–616, 2020.

[14] Junhao Huang, Weize Sun and Lei Huang, "Deep neural networks compression learning based on multiobjective evolutionary algorithms," *Neurocomputing*, vol. 378, no. 6, pp. 260–269, 2020.

[15] J. Xu and W. Chen, "Convolutional neural network-based identity recognition using ECG at different water temperatures during bathing," *Computers, Materials & Continua*, vol. 71, no. 1, pp. 1807–1819, 2022.

[16] Y. Yang, J. Duan, H. Yu, Z. Gao and X. Qiu, "An image classification method based on deep neural network with energy model," *Computer Modeling in Engineering & Sciences*, vol. 117, no. 3, pp. 555–575, 2018.

[17] A. Noorwali, A. Yengui, K. Ammous and A. Ammous, "Design and realization of non invasive fetal ECG monitoring system," *Intelligent Automation & Soft Computing*, vol. 32, no. 1, pp. 455–466, 2022.

[18] Q. Cui, S. McIntosh and H. Sun, "Identifying materials of photographic images and photorealistic computer generated graphics based on deep cnns," *Computers, Materials & Continua*, vol. 55, no. 2, pp. 229–241, 2018.

[19] A. Isin and S. Ozdalili, "Cardiac arrhythmia detection using deep learning," *Procedia Computer Science*, vol. 120, pp. 268–275, 2017.

[20] S. Kiranyaz, T. Ince and M. Gabbouj, "Real-time patient-specific ecg classification by 1d convolutional neural networks," *IEEE Transactions on Biomedical Engineering*, vol. 63, no. 3, pp. 664–675, 2015.

[21] U. R. Acharya, S. L. Oh, Y. Hagiwara, J. H. Tan, M. Adam *et al.,* "A deep convolutional neural network model to classify heartbeats," *Computers in Biology and Medicine*, vol. 89, no. 1, pp. 389–396, 2017.

[22] Z. Qu, B. Cao, X. Wang, F. Li, P. Xu *et al.,* "Feedback lstm network based on attention for image description generator," *Computers, Materials & Continua*, vol. 59, no. 2, pp. 575–589, 2019.

[23] P. Warrick and M. N. Homsi, "Cardiac arrhythmia detection from ecg combining convolutional and long short-term memory networks," in *Proc. CinC*, France, pp. 1–4, 2017.

[24] S. Oh, E. Y. Ng, R. San and U. R. Acharya, "Automated diagnosis of arrhythmia using combination of CNN and LSTM techniques with variable length heart beats," *Computers in Biology and Medicine*, vol. 102, no. 4, pp. 278–287, 2018.

[25] I. J. Goodfellow, J. Pouget-Abadie, M. Mirza, X. Bing, D. Warde-Farley *et al.,* "Generative adversarial nets," in *Proc. Int. Conf. on Neural Information Processing Systems*, Montreal, Canada, 2014.

[26] W. Fang, F. Zhang, V. S. Sheng and Y. Ding, "A method for improving cnn-based image recognition using dc-gan," *Computers, Materials & Continua*, vol. 57, no. 1, pp. 167–178, 2018.

[27] S. Mousavi and F. Afghah, "Inter-and intra-patient ecg heartbeat classification for arrhythmia detection: A sequence to sequence deep learning approach," in *Proc. ICASSP*, Brighton, UK, pp. 1308–1312, 2019.

[28] Z. F. M. Apandi, R. Ikeura and Hayakawa, "Arrhythmia detection using mit-bih dataset: A review," in *Proc ICASSDA*, Kuching, Malaysia, pp. 1–5, 2018.

[29] P. D. Chazal, M. O'Dwyer and R. B. Reilly, "Automatic classification of heartbeats using ECG morphology and heartbeat interval features," *IEEE Transactions on Biomedical Engineering*, vol. 51, no. 7, pp. 1196–1206, 2004.