Research on ventricular segmentation based on deep learning

Xuefeng Peng[1, *]

[1] Electronic information, Hebei University of Technology, Tianjin, 300401, China

[*] Corresponding author:jiaminghui22@mails.ucas.ac.cn

Abstract. Pattern recognition algorithms have achieved great success in medical image analysis, especially in terms of lesion area segmentation in magnetic resonance images. Different from the labor-intensive manual segmentation and diagnosis, image segmentation technology based on deep learning has made breakthroughs in the accuracy and speed of lesion area recognition in recent years. Focusing on the task of left ventricle segmentation from magnetic resonance images, this paper utilizes the powerful feature representation ability of the convolutional neural network to complete accurate and automatic left ventricle segmentation. For example, the left ventricle is segmented from the heart image based on MRI(Magnetic Resonance Imaging) through UNet, which is based on Convolutional neural network Meanwhile, the evaluation indicators and the process of data set preprocessing are explained. On this basis, the influence of changing the parameters of basic UNet is further studied, and the experimental analysis is carried out from three aspects: changing the batch size, changing the number of initial extracted features and changing the number of network layers. Finally, Dice coefficients of 0.9331, 0.9316 and 0.8416 were obtained on the test set, verification set and training set respectively, and good style results were obtained. At last, the further improvement was explained.

Keyword: UNet, deep learning, ventricular segmentation, magnetic resonance images.

P. Kar et al. (eds.), *Proceedings of the 2023 International Conference on Image, Algorithms and Artificial Intelligence (ICIAAI 2023)*, Advances in Computer Science Research 108,
https://doi.org/10.2991/978-94-6463-300-9_7

1 Introduction

Since cardiovascular disease happens very quickly, it is difficult to take effective measures to deal with it, resulting in a very high case fatality rate. Nowadays, cardiovascular patients are getting younger and more than 90 percent of the cases come from nine common factors [1]. Therefore, how to effectively prevent and timely detect cardiovascular diseases has attracted a lot of research attention.

The left ventricular has been demonstrated as one of the important components when evaluating cardiac function and condition [2]. In this case, accurate segmentation of the left ventricular region is of great significance for the diagnosis and treatment of cardiovascular diseases. During the initial determination of cardiovascular disease, the left ventricular region generally depends on manual segmentation by medical personnel, which is time-consuming and laborious. With the rapid development of pattern recognition technology, especially image segmentation technology, automatically segmenting the left ventricular region from magnetic resonance images based on computer vision algorithms shows a promising future. According to the differences in design ideas of segmentation algorithms, existing left ventricular segmentation methods can mainly be divided into traditional methods and deep learning-based methods. Traditional methods are mainly based on image features, deformable models, and map, which can get good segmentation effects under certain environments but have limitations more or less.

In recent years, thanks to the strong feature representation ability of convolutional neural networks, deep learning is gradually applied to medical image processing tasks [3]. Emad et al. achieved the localization of the left ventricle by some operations of the convolutional neural network [4]. However, it fails to separate the left ventricle. In the same year, Long et al. achieved left ventricular segmentation in cardiac MRI with a deep learning method for the first time [5]. However, due to differences in heart size or shape among different individuals, there are still many challenges in the practical application of left ventricular segmentation algorithms, mainly due to: (1) blood flow causing artifacts in the image; (2) The interference of other tissues in the image, such as liver and fat, on the left ventricular region; (3) The left ventricular area is relatively small.

To address the above issues, this article proposes a network that is more suitable for left ventricle image segmentation based on UNet. UNet has a very simple structure, which can be easily implemented on a computer. In addition, UNet can achieve better segmentation performance when there are fewer datasets since the magnetic resonance images of left ventricle are difficult to obtain. Through the modification of the network, qualified segmentation results are obtained in the heart dataset of the medical segmentation of decathlon. This paper is mainly composed of six parts. In Section 2, the related work involved in this paper is introduced. In Section 3 and 4, the details and results of proposed method are introduced, respectively. Finally, the future development of left ventricular segmentation research was discussed and the contributions of the entire article were summarized.

2 Related work

2.1 Convolutional Neural Network

CNN(Convolutional Neural Network) is a feed-forward neural network with multiple convolution computations and a deep level, and it is one of the main representatives of deep learning [6]. In recent times, there have been many areas of progress in deep learning, so it is very necessary to use CNN to optimize ventricular segmentation technology.

2.2 Filling method

Because when the picture goes through the convolution layer, the image size is getting smaller and smaller, which will affect the effect of feature recognition. To this end, different army filling methods are selected depending on the purpose. The main filling methods include Valid Padding, Same Padding, Full Padding, whose difference is that the size of the image obtained after filling is reduced, equal and increased respectively. A special method is Arbitrary Padding, while it is generally not used because it is difficult to operate. This article mainly uses the Same Padding. The input and output images are the same size, so the results of ventricular segmentation can be compared better.

2.3 Optimization algorithm

In the training process of convolutional neural network, many problems will be met, which will affect the training of the whole network. Corresponding measures should be taken to solve these problems, and the use of optimization algorithm will help the network to better train [7]. By using the optimization algorithms, the training time will be greatly reduced and the optimal solution can be found more quickly. The main methods include stochastic gradient descent optimization algorithm and momentum stochastic gradient descent optimization algorithm. AdaGrad algorithm and RMSProp algorithm have a dynamic adjustment of the learning rate, whose difference lies in the formula of cumulative historical gradient square. On this basis, Adam algorithm can even achieve different learning rates for different parameters. Based on this, this paper selects Adam optimization algorithm for training optimization.

3 Method

3.1 Full Convolutional Network

The ends of traditional convolutional neural networks are composed of some fully connected layers. Although images of any size can be accepted according to reason,the existence of these fully connected layers will cause the image to lose spatial information. Therefore, the relatively original convolutional neural networks do not perform well in image segmentation. Fully Convolutional Networks (FCN) is a type of network proposed by Long's team [5], which has played a huge role in the field of image segmentation. It replaced the previously fully connected layers with convolutional layers with only a portion of the connections and the same weight, which not only reduces the information used by the neural network but also restores the image to its original size,it further preserves the spatial information of the original image. any size of images can send into the fully connected convolutional neural network ,then achieve image segmentation from end to end at the pixel level [8].

3.2 UNet

The UNet is a modified convolutional neural network developed by Olaf Ronneberger's team in 2015, whose basic schematic diagram is shown in Figure 1. The UNet can mainly be divided into two parts: the downsampling part on the left, which is the contraction path, and the upsampling part, which is the extension path. The contraction path on the left is mainly a general convolutional network, where each layer in the contraction path is composed of two layers of a 3×3 convolutional kernel, and it also has a 2×2 downsampling. After this convolution operation, the network is activated through the ReLU activation function, and then the information is pooled to the maximum value to complete the image compression process. It is worth noting that after each pooling is completed, the number of convolutional kernels in the next round will double to ensure the number of image features, which is the downsampling part of the network. After each downsampling, The area of the feature map has been cut in half, but the amount of features obtained by the network doubles. The expansion path on the right side of the network corresponds one-to-one to the contraction path on the left, and each layer in the expansion path is composed of two layers of a 3×3 convolution kernel, it also has a 2×2 convolution kernel in the deconvolution layer. UNet combines the output in the contraction path with the results of the previous layer in the expansion path, and uses this as the input at this level in the expansion path. After each upsampling process, the area of the obtained feature map will double, while the amount of features obtained will be reduced to half of the previous one. Since left ventricular segmentation is a yes or no problem, the output layer is generally activated through the Sigmod activation function, and then can get whether the different points are left ventricles and the corresponding probabilities.

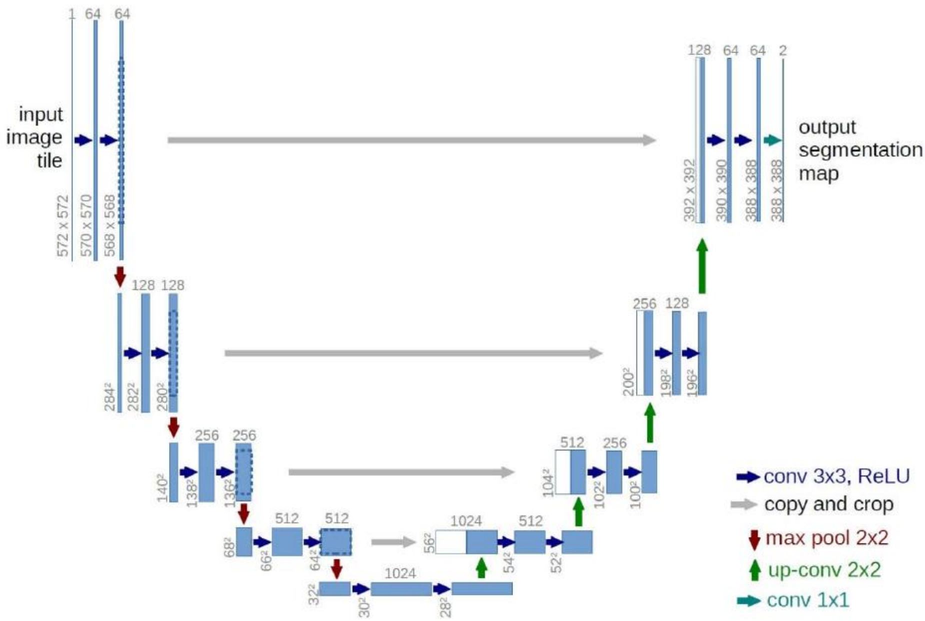

Fig. 1. UNet architecture [9]

3.3 Loss function and strategy

Due to the fact that left ventricular segmentation is a binary classification method, this article uses Dice loss and BCE loss. Dice loss ignores most of the background of cardiac MRI, thus minimizing the interference caused by background factors on loss values, making it more suitable for use in the field of medical image segmentation. BCE loss is a commonly used binary loss. Due to the common BCE Loss will magnify the effect of background, it's not very good for segmenting the left ventricle [10]. Therefore, Dice Loss (DL) is used as the training loss function in this project, which can reduce the influence of other background loss values.

This experiment used a learning rate with an initial learning rate of 10-2 to schedule the program. If the Dice loss obtained does not decrease for five consecutive cycles, the learning rate will be reduced to half of the previous one. By using this learning rate scheduler, the learning rate can be adjusted in a timely manner based on the training results. This operation can greatly improve the quality and efficiency of training and achieve better left ventricular segmentation results.

4 Experiment

4.1 Evaluation index

This paper mainly gets evaluations of resulting segmentation by accuracy and Dice Similarity Coefficient (DSC). Accuracy is the most common and basic evaluation index. DSC is a parameter used to judge how well the actual forecast results match the real situation. It could be any number from 0 to 1. The closer it gets to 1, the better the network's pixel classification.

4.2 Data processing

This data set is based on MRI of the heart in the Medical Segmentation Decathlon challenge. The cardiac MRI in this data set were stored using the NIfTI-2 data format. Each data file is an image of a patient's heart during breathing. It is a multi-channel 3D array in which each channel represents a cross-sectional view of the axial, coronal, and sagittal planes. For each file, a corresponding label file exists that contains information about the position of the left ventricle in the image. The dataset consists of 180 annotated MR Images of the heart, each 320 by 320 pixels in size. The specific example is shown in Figure 2. The following details the processing means of the data set of this subject.

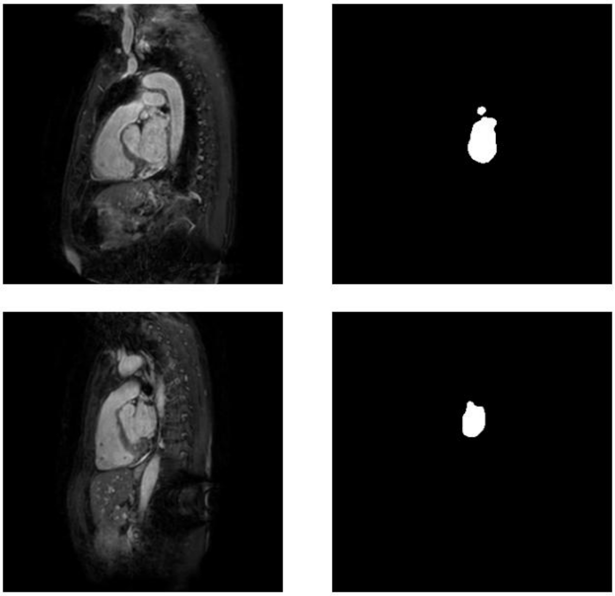

Fig. 2. Data sample(self-draw)

Image denoising. Due to some inherent noise in the original image, the image will be affected, so the removal of outliers should be carried out. The measures to be taken are to replace the pixels in the position with obvious anomalies with the pixels closest to them and with normal values, so as to achieve the effect of image denoising.

Slices were sliced and normalized. Because the format used to store raw data is in 3D format, the original image should be sliced according to a certain step size and then displayed as 2D image. At the same time, due to the uneven gray distribution of cardiac MRI, It need to carry out gray normalization processing for the image, so as to facilitate the subsequent image processing [11].

Image format conversion. Since the image of the original data set is in NII format, which is difficult to operate, the image format should be converted to a format that is easier to operate.

According to the above processing content, the program is written according to the characteristics of the corresponding data set. Use this program to read the data set and process it into a format Pytorch recognizes.

4.3 Experiment settings

The network of this project was completed based on the Python programming language and PyTorch deep learning framework. The entire experiment was conducted on the Win10 system, with the main configurations being an Inter i7-7700HQ @ 2,80GHz CPU and an NVIDIA GeForce GTX 1050 graphics card with 2GB memory, as well as 16 (8+8) GB of 2400MHz memory and 128GB of the solid-state hard drive.

4.4 Performance analysis

The UNet mentioned above is applied to the left ventricular segmentation of cardiac MRI. Firstly,the data of cardiac MR Image is loaded and the image is preprocessed for better image training. Then,the image after processing is sent into UNet,after training,the test image is sent into the model.Then the test result is gotten from the trained model.

The impact of changing parameters on the network. In the basic UNet network after the completion of the left ventricular segmentation, this topic is also on this basis to further study the impact of changing parameters on the segmentation results, mainly by modifying the basic parameters of three aspects to explore, including the batch size, the number of initial extraction features, the number of network layers.

Effect of changing batch size. Batch size means that when training the whole network, a batch will be sent for training each time. Due to various limitations, the pictures of each batch training are limited, and the change of batch size will affect the accuracy of the whole network training. When choosing a relatively small batch size, randomness will be increased, but the weight will be updated more frequently. Choosing a larger batch size will reduce randomness, but at the same time, it will slow down the frequency of weight updating. This topic observed the effect of segmentation by changing different batch sizes, which is to compare the segmentation performance of the network by dice coefficient of segmentation. The specific parameter changes are shown in Table 1, and the results are shown in Figure 3. It can be seen from the results that the best effect of left ventricular segmentation was obtained when the batch size was 4 in our experiment.

Table 1. Change batch size experiment parameters

Variable	Variable value
Batch size	1,2,4,8,16
Initial filter depth	64
Number of network layers	4

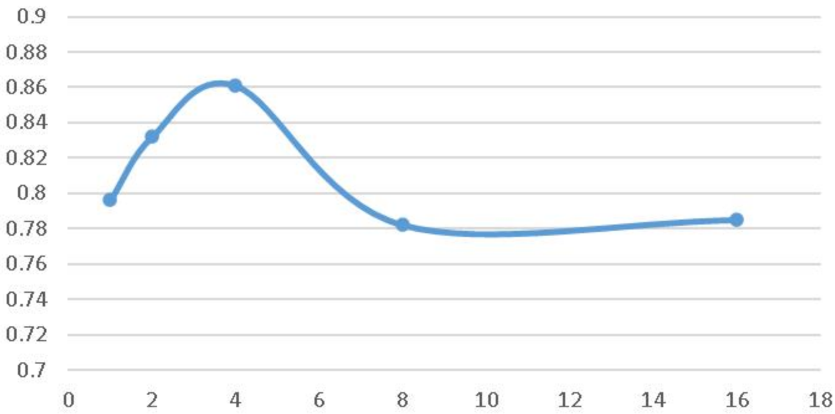

Fig. 3. Effect of changing batch size(self-draw)

Effect of changing the number of features initially extracted. The initial filtering depth refers to the feature number to be obtained in the first convolution. The initial filtering depth of basic UNet is 64. Now change this value, and then observe the impact on left ventricular segmentation. The specific parameter changes are shown in Table 2, and the results are shown in Figure 4. It can be seen that increasing the filter depth can indeed improve the dice coefficient of image segmentation, but at the same time, one problem can not be ignored, that is, the larger the initial filter depth is, the larger the size of the entire network model will also increase, and the training speed will be greatly reduced, so it is necessary to find a proper balance between the accuracy and the size of the network model. For this project, combined with multiple considerations, the initial filtering depth of 16 was selected as the final training parameter.

Table 2. Change the experimental parameters of the number of initial extracted features

Variable	Variable value
Batch size	4
Initial filter depth	8,16,32,64,128
Number of network layers	4

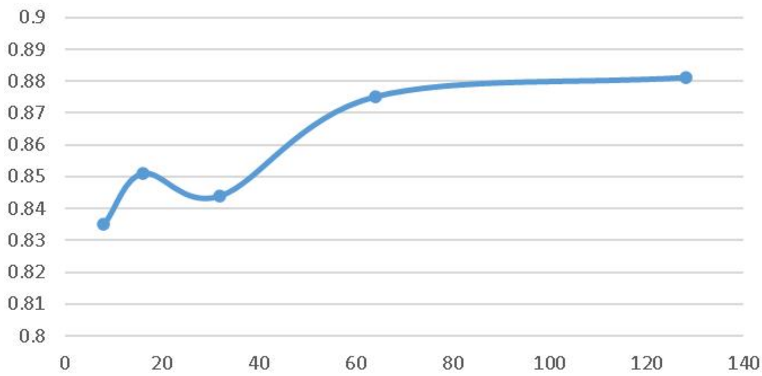

Fig. 4. Effect of changing Number of features initially extracted(self-draw)

*Effect of changing the number of network layers.*In the basic UNet, according to the content mentioned before, each network has 4 contraction paths and 4 expansion paths, that is to say, each UNet is composed of 4 layers.Since this project can achieve the function of changing the number of UNet layers, the influence of increasing the number of network layers on the effect of left ventricular segmentation will be studied next. The specific parameter changes can be seen in Table 3, the obtained results can be seen in Figure 5. It can be shown that with the network model increases with the number of layers, the effect of left ventricular segmentation becomes worse. This is because the experimental configuration of this experiment is not enough to support the training of deeper network, so the experiment still adopts the original UNet four-layer network for training segmentation.To sum up, this experiment obtains the best segmentation results when the batch size is 4, the initial filtering depth is 16, and the number of network layers is 4. These results are obtained when the epoch is 100.

Table 3. Change the experimental parameters of network layers

Variable	Variable value
Batch size	4
Initial filter depth	16
Number of network layers	4,5,6,7,8

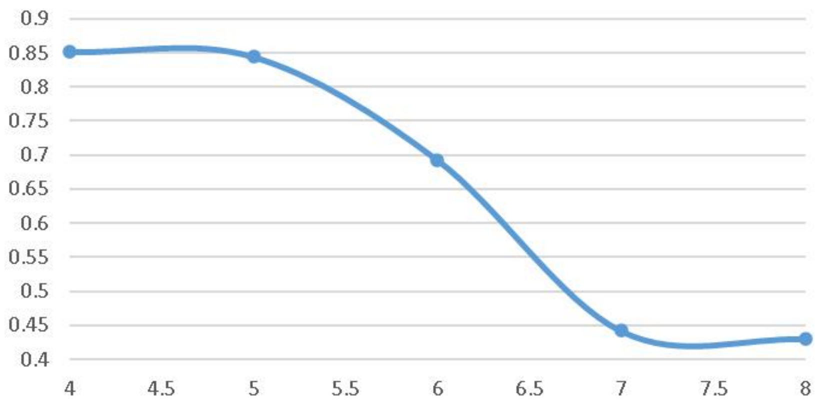

Fig. 5. The effect of changing the number of network layers(self-draw)

Result presentation. Figure 6 shows the changes of loss and coefficient in the training process. It can be seen that with the progress of training, parameters are based on convergence, indicating that our experimental method is effective. Figure 7 shows a specific example of UNet based left ventricular segmentation results. It can be seen that the effect of UNet based left ventricle segmentation is relatively good and stable, and the contour of left ventricle can be well divided in the experiment.

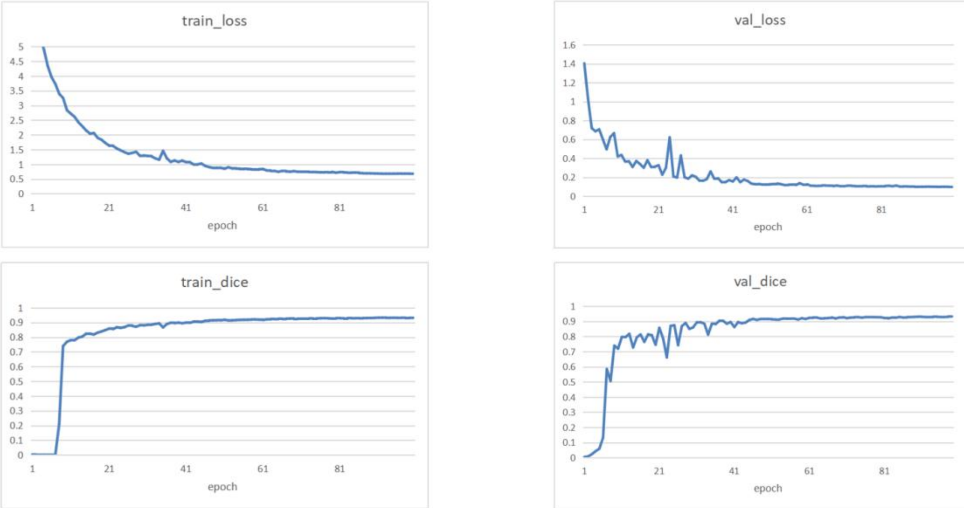

Fig. 6. UNet training process(self-draw)

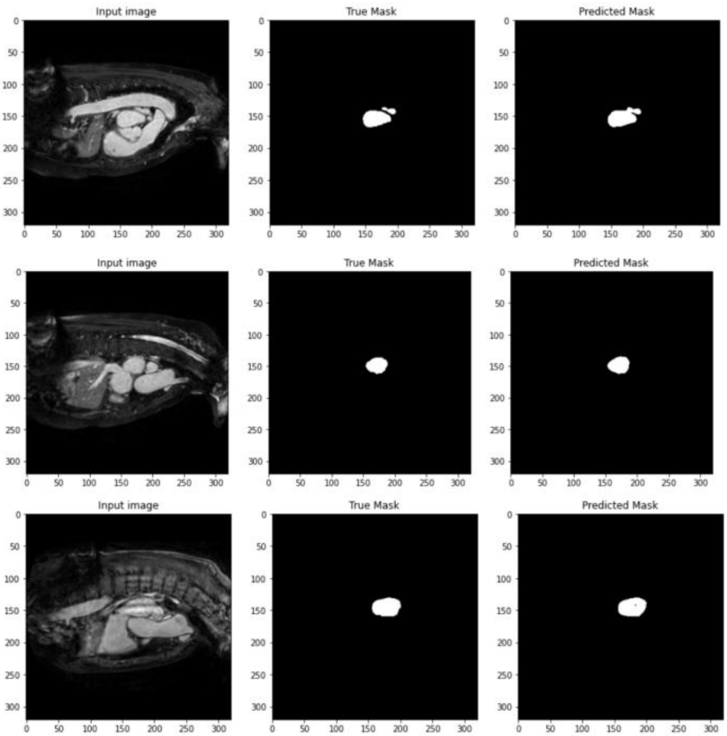

Fig. 7. Segmentation example(self-draw)

5 Discussion

With the emergence of more and more science and technology along with the progress of The Times, the treatment of people's diseases is becoming more and more advanced. MRI imaging technology allows patients no longer to be exposed to electromagnetic radiation, and has a clear image, good distinction effect for different tissues, and clear distinction between myocardium and blood pool. Based on these advantages, this technology has become the standard for the evaluation of cardiac function. This project takes UNet network as the starting point and then realizes the segmentation of the left ventricle in MR Image of the heart. The specific research content is as follows:

(1)Put the data into the model training based on the basic UNet, complete the left ventricular segmentation of the heart MR Image, and obtain the corresponding evaluation indicators.

(2)On the basis of the original UNet network, the appropriate parameter selection in this subject was studied from the three aspects of changing the batch size, changing the initial feature extraction quantity and changing the number of network layers, and finally the network model with better performance in the experimental environment of this subject was obtained.

After these experimental results were obtained, the subject was analyzed and the left ventricle segmentation results obtained by the network were analyzed. It was found that UNet performed well in the left ventricle segmentation with a small sample size. However, due to the influence of experimental environment and other factors, the segmentation results were still not particularly ideal, which was also the place to be further improved. There is still room for the following improvements:

(1)The original data in the data set is small, and there is a certain risk of overfitting.

(2)Although the segmentation results obtained from the model in this article are not bad, there are still many areas for improvement.After that, the structure of UNet network can be improved on the basis of this paper, and it can be combined with other segmentation methods to improve the segmentation effect of left ventricle.

(3)Improve the universality of the model so that it can be used for other 3D medical images, such as liver images.

6 Conclusion

This paper uses UNet network to achieve left ventricular segmentation of cardiac MR Images, and explains the evaluation indicators and the process of data set preprocessing. Afterwards, the data was placed into the network for training and the model was obtained. Then, it was segmented into the left ventricle and preliminary segmentation results were obtained. On this basis, the influence of changing parameters on the basic UNet is further studied, and the experimental analysis is carried out from three aspects: changing the batch size, changing the initial filtering depth and changing the number of network layers.It was found that changing the batch size may improve the results, but due to device limitations, being too large may actually have a negative effect; Although changing the number of initially extracted features may result in better results, the experimental time also greatly increases and the efficiency is not high; Finally, changing the number of layers in the network reveals that the training effect has even deteriorated, indicating that the more layers the network has, the better the effect may not necessarily be.Finally, after experimental analysis, it was found that the most suitable environment for this experiment is when the batch size is 4, the initial feature extraction is 16, and the number of network layers is 4. the Dice coefficient of 0.8416 was obtained in the test set. It is found that UNet has more advantages than traditional convolutional neural network in the field of medical image segmentation with small samples. In this paper, the deep learning method is used to solve the problem of medical left ventricular segmentation to a certain extent, and the parameter adjustment is combined with the experimental equipment to get better segmentation results. However, due to the limitations of the experimental environment, the results obtained are still not the best. In the future,I hope to conduct experiments under better hardware conditions to obtain better experimental results and more accurately segment the left ventricle.

References

1. Lopez E O, Ballard B D, Jan A. Cardiovascular disease[M]//StatPearls [Internet]. StatPearls Publishing, (2022).
2. Shoaib M A, Chuah J H, Ali R, et al. An Overview of Deep Learning Methods for Left Ventricle Segmentation[J]. Computational intelligence and neuroscience, (2023).
3. Esteva A, Chou K, Yeung S, et al. Deep learning-enabled medical computer vision[J].

NPJ digital medicine, 4(1): 5, (2021).

4. Emad O, Yassine I A, Fahmy A S. Automatic localization of the left ventricle in cardiac MRI images using deep learning[C]//2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). IEEE, 683-686, (2015).
5. Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation[C]. Proceedings of the IEEE conference on computer vision and pattern recognition. 3431-3440, (2015).
6. Khan, A., Sohail, A., Zahoora, U. et al. A survey of the recent architectures of deep convolutional neural networks. Artif Intell Rev 53, 5455–5516 (2020).
7. Cong, S., Zhou, Y. A review of convolutional neural network architectures and their optimizations. Artif Intell Rev 56, 1905–1969 (2023).
8. Tang P , Yang P , Nie D ,et al.Unified medical image segmentation by learning from uncertainty in an end-to-end manner[J].Knowledge-based systems, Apr.6: 241, (2022).
9. Ronneberger O, Fischer P, Brox T. U-Net: Convolutional networks for biomedical image segmentation[C]. International Conference on Medical image computing and computer-assisted intervention. Springer, Cham, 234-241, (2015).
10. S. Jadon, "A survey of loss functions for semantic segmentation," 2020 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB), Via del Mar, Chile, pp. 1-7, doi: 10.1109/CIBCB48159.2020.9277638, (2020).
11. Zhu H C, Tong D, et al. Temporally downsampled cerebral CT perfusion image restoration using deep residual learning[J]. International Journal Of Computer Assisted Radiology And Surgery. 15(2):193-201, (2020).