International Journal of Computational Intelligence Systems and Applications. Volume 8, Issue 1, (2021) 23-31
http://www.cayley-nielson.com/ijcisa/

Research On An Innovative Image Segmentation Method Of Brain Tumor

Xujing Zhang, Mengxin Li, Tianhui Zhang, Kexue Hu, Kefu Li, Chaoyu Han, Xinghua Wang
School of Information and Control Engineering, Shenyang Jianzhu University, 110168, Shenyang, China
Corresponding author: Xujing Zhang
E-mail: 1198808489@qq.com, limengxin@sjzu.edu.cn, 1005601143@qq.com, 1513823267@qq.com, 1317201638@qq.com, 2692896880@qq.com, 1836907365@qq.com

Abstract

Aiming at the disadvantage that the general network model depends too much on the number of data and labels, this paper proposes a brain tumor segmentation method using 2D-pix2pix on the basis of Generative Adversarial Networks. Pix2pix is a conditional generation countermeasure network method for image migration. Its advantage is that when the image is input, the original image is added as a constraint to avoid the problems such as free generated image and unstable network. For the generated model, the residual module is used to replace the original depth convolution module to avoid the disappearance of gradient in the process of depth convolution.

Keywords: pix2pix, U-net, Image Segmentation, Brain Tumor.

1. Introduction

Brain tumors are highly heterogeneous showing irregularities in gray value, shape, position and size on MRI images. Therefore when extracting the features of brain tumor images, it to be fully extract the boundary information and global semantic information of tumors. At the same time, it to be fully choose the extracted tumor feature information. Because the brain tumor images are different from other brain tissues, which are evenly distributed in the whole brain cavity, the feature information of useless areas needs to be discarded to speed up the model training speed.

The generator of 2D-pix2pix can effectively coordinate the global semantic information by using the global average pooling ability of U-net network. Due to the large image size, the network depth is large, and the gradient is easy to disappear and explode in the process of model training. In this paper, the residual module is used to replace the original convolution module to solve gradient vanishing problem and gradient exploding problem in model training. The discriminator adopts the Patch-GAN network structure commonly used in the image migration network, and $N \times N$ uses the matrix to distinguish the whole image. No matter how large the generated image is, it will be divided into multiple receptive fields of fixed size for judgment, so as to reduce the input of the discriminator, reduce the amount of calculation and speed up the training speed.

2. Network structure design

2.1. Generator network structure

The generator adopts U-net network structure and uses the symmetry of U-net network. In the process of image migration, the image of tumor area will not change, that is the structure of input image and the underlying structure of output image are one-to-one corresponding.

The network structure of the generator is shown in Figure 1,

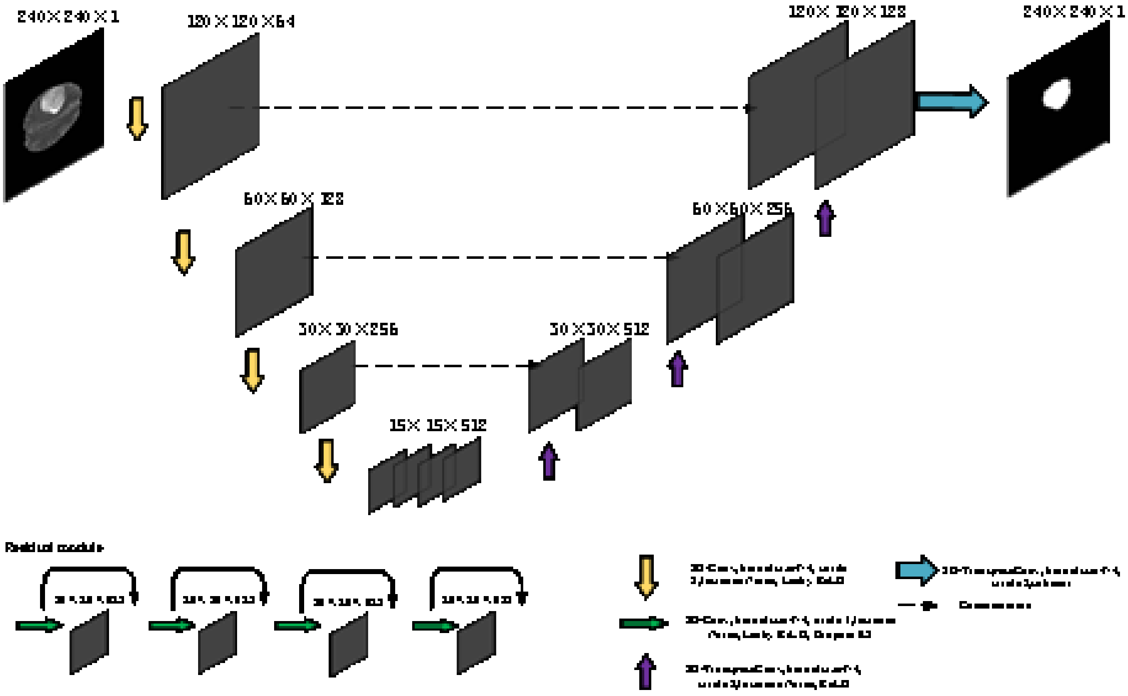

Figure 1. Generator network structure

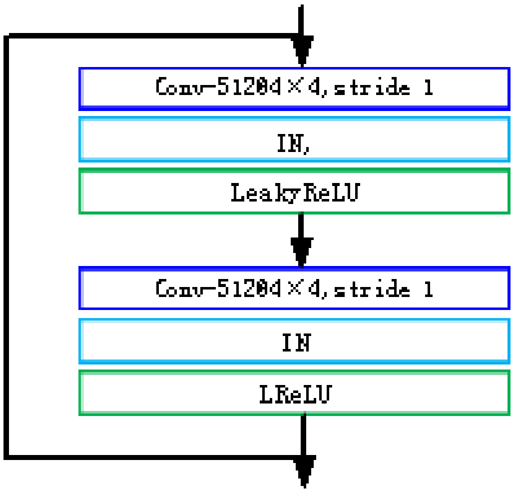

Figure 2. Residual module structure

Specific model architecture details:

Input: a T2 modal 240×240 image of MRI with channel 1;

Encoder: four down sampling blocks. Each uses 2D convolution 4×4 with kernel size.The step is 2, and the instance norm (in) and leaky ReLU activation functions are used. The number of filters used in the first 2D convolution is 64, and the number becomes twice in each downsampling;

Bottleneck (residual module): four residual blocks, each of which uses 2D convolution with kernel size of 4, step size of 1, followed by instance normalized in and leaky ReLU activation functions, and the normalized output of each convolution is connected with the previous residual block;

Decoder: three up sampling blocks. Each up sampling block adopts 2D transpose convolution with core size of 4 and step of 2. Using instance normalization and ReLU activation function, each 2D convolution input is connected with its own encoder output layer;

Output: a 2D transpose convolution uses four filters with convolution kernel size of 4 and step size of 2, and uses softmax activation function to output a single channel 240×240 image with image size of as segmentation prediction.

During the experiment, the U-net network structure adopted by the generator uses the residual module to replace the original convolution module due to the gradient dispersion / explosion and over fitting caused by deep convolution. The residual module structure is shown in Figure 2,

At the same time, in the convolution process, a large number of data training leads to the increase of training time. Therefore, dropout regularization is added to randomly shield some neurons in the training process, so as to improve the generalization ability of the model, and it is not easy to over fit the training data. The regularization probability is 0.2.

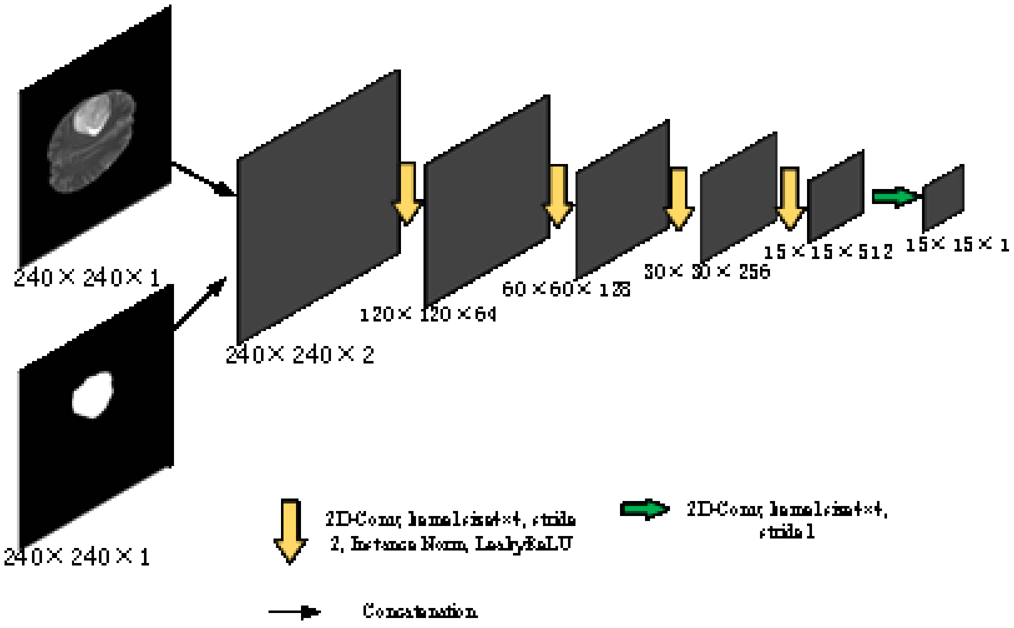

Figure 3. Network structure of 2D-pix2pix discriminator

2.2. Discriminator network structure

The discriminator adopts the Patch-GAN network structure to replace the network of the discriminator with a full convolution neural network. After the image passes through various convolution layers, it will not be input into the full connection layer or activation function, but uses convolution to map the N×N input into a matrix, which is equivalent to the last evaluation value in the original GAN to evaluate the generated image of the generator.N×N Each point in the matrix (true or false) represents a small area evaluation value in the original image, which is the application of receptive field. The original GAN used a value to measure the whole image. Now N×N the matrix is used to evaluate the whole image. No matter how large the generated image is, it is divided into multiple fixed size patches and input into the D discriminator for judgment, D so as to reduce the input, reduce the amount of calculation and speed up the training speed.

At the same time, the generator g itself is fully convoluted and has no limit on the image scale, while the discriminator D processes the image according to the matrix patch and has no limit on the image size. Thus, the entire 2D-pix2pix framework has no limit on the image size. It increases the scalability of the framework. The discriminator network structure is shown in Figure 3,

The network model structure proposed is as follows:

Input: a single channel original brain tumor image and its real binary mask segmentation image, or input its original brain tumor image and the binary segmentation prediction image generated by the generator;

Encoder: the same structure as the generator;

Output: a 2D convolution uses a filter with a convolution kernel size of 4, a step of 1 and a boundary fill of 0. The output 15×15 is a single channel image, which is used as the standard for the generator to generate segmentation prediction results.

In the specific experiment, the coefficients of all ReLU activation functions with leakage in the generator and discriminator are 0.3, so as to reduce the consumption of GPU.

Modified linear unit (ReLU) is the most commonly used activation function in neural networks. Neurons are activated only when the input exceeds the threshold. However, when the input is positive, the derivative is not zero, allowing gradient based learning.This activation function can speed up the calculation, but when the input is negative, the learning speed of ReLU will slow down, or even make the neurons directly invalid. Therefore, when the input is less than zero, the gradient is zero, resulting in the weight can not be updated, and will remain silent in the rest of the training process. In order to solve the disadvantage of ReLU function, a leaky ReLU function is introduced into the negative half region of ReLU function.The leaky ReLU function is a variation of the classical (and widely used) ReLU activation function, whose output has a small slope to the negative input. Since the derivative is always non-zero, this can reduce the emergence of silent neurons, allow gradient based learning (although it will be very slow), and solve the problem that neurons do not learn after ReLU function enters the negative interval.

A convolution layer is used in each module of the encoder and decoder. At the same time, the step of convolution is 2, so that the encoder or decoder can directly down sample or up sample the input, and the sampling multiple is 2.

3. Loss function

2D-pix2pix network is composed of two structures: generation model and discrimination model. Therefore, the loss function is also divided into two parts: generator loss () $L_G$and discriminator loss (). L_D Discriminator loss formula (1):

$$L_D = L_2[D(x,y),1] + L_2[D(x,\hat{y}),0] \tag{1}$$

x Represents the original image, y represents the real segmented image, and $L_2[D(x,y),1]$ represents the discriminator x error between the original image and the y real segmented image with tensor $\hat{y}$1; Represents the segmentation prediction image generated by the $L_2[D(x,\hat{y}),0]$ generator, and represents the discrimination error between the original image and $\hat{y}$ the generated segmentation image with tensor 0.Generator loss function formula (4):

$$L_G = L_2[D(x,y),1] + \alpha GDL(y,\hat{y}) \tag{2}$$

Among them, the image gradient difference loss (GDL) function is used. The GDL function improves the ordinary dice loss function, which produces dice fluctuation due to the prediction error of small target areas in the image, resulting in the problem of unstable training. GDL function is the error loss, gradient weight $y$coefficient and rounding between the real segmented image $\hat{y}$and the predicted segmented image generated $\alpha \to \infty$by the generator $\alpha \geq 0$(at that time, $\alpha = 0$the loss of the generator is only the loss of the unsupervised part of the discriminator, at that time, the whole network can be regarded as a U-net network).

4. Model training

4.1. Adam optimization algorithm

Adam optimization algorithm was proposed by Kingma and Lei BA in 2014. It combines the advantages of adaptive gradient optimization and rmsprop optimization algorithms to comprehensively calculate the first-order moment estimation (i.e. the mean value of the gradient) and the second-order moment estimation (i.e. the decentralized variance of the gradient) of the gradient, so as to calculate the update step size.

Specific formula:

$$g_t = \nabla_\theta J(\theta_{t-1}) \tag{8}$$

$$m_t = \beta_1 m_{t-1} + (1-\beta_1) g_t \tag{9}$$

$$v_t = \beta_2 v_{t-1} + (1-\beta_2) g_t^2 \tag{10}$$

$$\widehat{m}_t = \frac{m_t}{(1-\beta_1^t)} \tag{11}$$

$$\hat{v}_t = \frac{m_t}{(1-\beta_2^t)} \tag{12}$$

$$\theta_t = \theta_{t-1} - \frac{\alpha \cdot \widehat{m}_t}{(\sqrt{\hat{v}_t}+\varepsilon)} \tag{13}$$

Among them, the learning $\alpha = 0.0002$ rate ε is 10^{-8} taken $\beta_1 = 0.8$ as,

β_1(control the weight distribution), (control $\beta_1 = 0.999$ the influence of the square of the gradient before).

Adam optimization algorithm has several significant advantages:

A has the advantages of simple implementation, efficient calculation and less memory requirements

The update of B parameter is not affected by the scaling transformation of gradient

The c-hyperparameter is well interpretable and usually requires no adjustment or little fine-tuning

D the updated step size can be limited to the approximate range, and the learning rate can be adjusted automatically

E applicable to unstable objective function

F is suitable for large-scale data and parameter optimization problems

Based on the advantages of Adam optimization algorithm, the generator and discriminator of the model in this paper adopt Adam optimizer.

5. Dropout algorithm

During the experiment, the U-net network structure used by the generator is prone to over fitting during training. At the same time, a large number of data training leads to the increase of training time in the convolution process. Therefore, the dropout layer is added to alleviate the possible over fitting problem.

Dropout algorithm was proposed by Hinton in 2012. For complex feedforward neural networks, it is easy to cause over fitting when they are trained in small data sets. In the process of training neural network, each training batch can significantly reduce the over fitting phenomenon by ignoring the general feature detection (making the general hidden layer node value 0). In order to reduce the interaction between hidden layer nodes. In short, in the process of forward propagation, let the activation value of a neuron stop working with a certain probability without relying on some local features, so as to make the generalization ability of the model stronger. The regularization probability of this experiment is 0.2. Figure 4 shows the dropout application process:

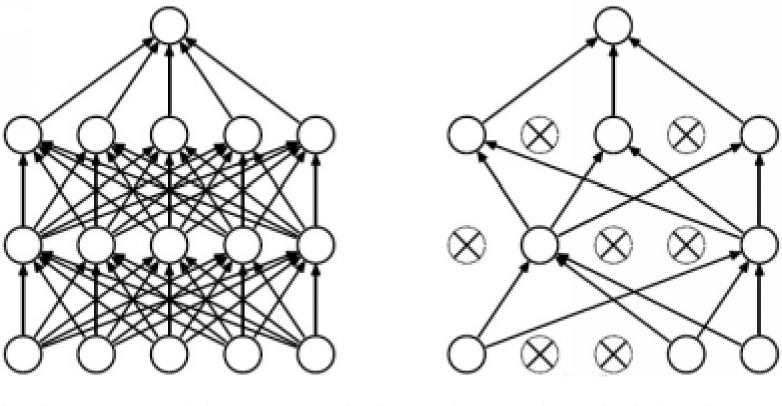

Figure 4. Application process of dropout (standard neural network on the left and neural network after applying dropout regularization on the right)

6. Experimental results and comparative analysis

6.1. Experimental results

The experimental environment in this paper is implemented by Python 3.8 and python 1.8.1. The model is trained and verified on 240×240×1 the images of 160 and 40 subjects respectively. The batch size of each batch is 81000 rounds of training.The computer is equipped with 16GB ram and 10GB memory NVIDIA geforce RTX 3080 graphics card.

The generated brain tumor segmentation image is shown in the figure. The generated segmentation image has similar shape features to the manual segmentation mask and similar texture features to the real MRI image. From figure 3.5, the manual segmentation mask image has similar underlying structure and shape features to the generated segmentation image. The generated results show the effectiveness of brain tumor image segmentation method based on 2D-pix2pix.

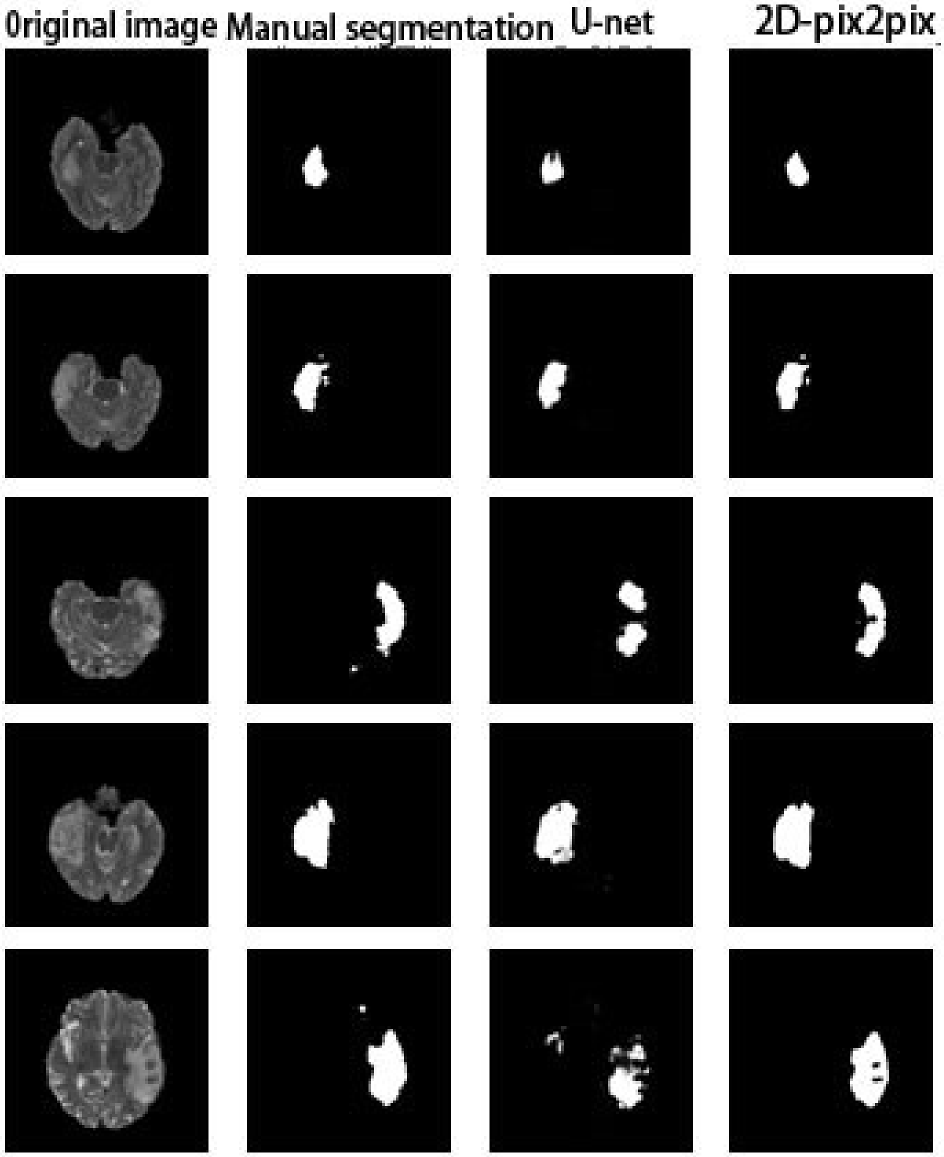

Figure 5. Brain tumor segmentation results of 2D-pix2pix

The experimental results of 2D-pix2pix brain tumor image segmentation method are compared with ordinary U-net segmentation results and manual segmentation results. It can be clearly compared from the figure that this method is used to show obvious effects in both tumor internal segmentation and tumor edge segmentation.However, the segmentation of small positive samples still needs to be improved.

6.2. Comparative analysis of experimental results

In order to verify the effectiveness of this method, the experimental results are compared under the same data set. Dice and hd95 evaluation indexes are used to compare other methods to segment complete tumor (WT). The results are as follows:

Table 1. Comparison of complete tumor segmentation performance results

method	Dice(%)	HD95(mm)
Superpix U-net[1]	89.27	4.97
Zhou[2]	90.52	4.39
Xu[3]	90.70	5.39
ResNet[4]	91.21	3.88
Vox2Vox[5]	91.81	5.02
The method presented	91.93	4.37

As can be seen from Table 3.1, compared with other methods, the surper-pixel U-net tumor segmentation method proposed by Hu et al. [1] has the lowest segmentation accuracy, but has a good effect on tumor edge segmentation; The single channel multi task segmentation method adopted by Zhou et al. [2] improves the dice index by 1.25% and the edge segmentation accuracy; The effect of DCan proposed by Xu et al. [3] on tumor internal segmentation is significantly improved, but the edge segmentation accuracy is reduced and there is over segmentation; The brain tumor image segmentation method based on residual network proposed by Zhou et al. [4] significantly improves the accuracy of image edge segmentation; The Vox2Vox method proposed by Marco DC et al. [5] is a GAN network for brain tumor image segmentation. Although it has little effect on the segmentation edge, it has greatly improved the segmentation accuracy of the complete tumor region. Compared with other methods, the improved pix2pix network not only improves the segmentation accuracy of tumor, but also can obtain segmentation results similar to the real segmentation label, and the segmented tumor boundary is smoother.

Using the confrontation principle of Generative Adversarial Networks, a self-monitoring mechanism is added in the image training process, so as to generate brain tumor images with higher accuracy and more accurate segmentation. In the generator network, the residual module is used to replace part of the convolution module to prevent the depth from being too deep. The image instance is normalized by instance norm to prevent the loss of image information during migration. The experiment compares fcn8, dense-U-net, Vox2Vox and other methods. The dice score of this method in the complete tumor area is 89.82%. The results show that the segmentation network structure proposed in this paper is more appropriate and complete, and effectively solves the problems of large amount of data labels, long training time and low segmentation accuracy in brain tumor segmentation.

7. Conclusions

This paper mainly introduces the network structure of 2D-pix2pix brain tumor segmentation method. In order to solve the gradient disappearance problem caused by deep convolution, the residual module is added to the generator to replace the original convolution, which can improve the coding and decoding ability of the model. The image gradient error function is used to replace the original L1 loss function to improve the segmentation accuracy of the model. The model is simulated and trained. By analyzing the experimental results and comparing the results of other methods, 2D-pix2pix brain tumor segmentation method can segment brain tumor images effectively, and the accuracy of tumor internal segmentation and tumor boundary segmentation is significantly enhanced.

References

[1] Hu Chun-yan, Si Ming-ming, CHEN Wei. A multi-modal brain tumor image segmentation method based on superpixels and improved U-net[J/OL].Small microcomputer system：1-9.2021

[2] Zhou C, Ding C, Wang, et al. One-Pass Multi-Task Networks With Cross-Task Guided Attention for Brain Tumor Segmentation, in IEEE Trans on Image Processing. 2020

[3] Xu H, Xie H, Liu Y, et al. Deep Cascaded Attention Networks for Multi-task Brain Tumor Segmentation [C]. 2019

[4] Zhou Xin-yu. Research on brain tumor segmentation based on three-dimensional residual convolutional neural network [D]. Xiangtan University, 2020.

[5] Marco D.C, David A, Anders E, Vox2Vox: 3D-GAN for Brain Tumor Segmentation. 2020

[6] Pereira S, Pinto A, Alves V, etal, Conditional random fields as recurrent neural networks[C]//Proceedings of the IEEE international conference on computervision.2015:1529-1537.

[7] Mohammad Havaei, Axel Davy, David Warde-Farley,Antoine Biard,Aaron Courville, Yoshua Bengio, Chris Pal, Pierre-Marc Jodoin,Hugo Larochelle. Brain tumor segmentation with Deep Neural Networks[J]. Medical Image Analysis,2017,35

[8] P. L. Chithra, G. Dheepa. Di-phase midway convolution and deconvolution network for brain tumor segmentation in MRI images[J]. International Journal of Imaging Systems and Technology,2020,30(3).

[9] Long J, Shelhamer E, Darrell, T. Fully convolutional networks for Semantic seg mentation[C]//Computer Vision and Pattern Recognition.IEEE,2015:3431-3440.