Study on MRI Brain Tumor Recognition and Segmentation Based on an Improved U-Net Network Model

Hailan Yu, Gengrun Yao, Hang Xu and Ruoshan Xiong

School of Information Technology and Engineering, Guangzhou College of Commerce, Guangzhou, Guangdong, China
Email: 360724201@qq.com

Abstract. Current malignant tumor is one of the main causes of death in Chinese residents, and the incidence of brain tumor in the population is higher. With the continuous development of modern imaging technology, especially MRI imaging technology, it provides a favorable help for doctors to assess brain tumors and choose treatment options. The accuracy of brain tumor segmentation is the key for doctors to diagnose patients' medical condition information. Brain tumor's structure is complex, shape is changeable and grayscale isn’t uniform, showing considerable differences in different patients. Artificial segmentation of MRI brain tumor images is time-consuming, and it is often affected by subjective differences. For the above reasons, researchers are working to develop an MRI brain tumor recognition segmentation method. This paper presents a recognition and segmentation method for MRI brain tumors, based on an improved U-Net network model. The proposed method avoids too large or too small learning rate leading to too fast or too slow convergence, and adds the BatchNormalization module to normalize the input of the input activation function of the network, thus realizing the training length, robustness, and accurately and completely dividing the brain tumor image. This paper performs detection experiments on the Kaggle+LGG Datasets, with the average classification accuracy of 99.8309%, the average IoU is 0.9935, and the average Dice coefficient is 0.9956.

Keywords. MRI brain tumor image; U-Net; recognition and segmentation; BatchNormalization.

1. Introduction

With the MRI (Magnetic Resonance Imaging) Rapid development in technology, medical imaging has gradually become an important basis for medical diagnosis and disease treatment, extensive used in clinical medicine, in the current medical imaging medical diagnosis, current brain tumor image segmentation methods are divided into three categories: artificial segmentation, semi-automatic segmentation and automatic segmentation [1], there is a certain rate of error. However, the accuracy of the results on medical imaging evaluation greatly affects the treatment plan of surgery, so how to reduce the error rate of medical image processing analysis is a big problem that is urgently waiting to be solved.

In recent years, some tumor image segmentation techniques have also emerged in the field of image segmentation, and image recognition segmentation also began to transfer from artificial to automatic. The method of automatic image segmentation is mainly in the traditional method of segmentation, such as edge-based segmentation method, area growth method [2], etc. The current automatic segmentation is fundamentally different from other segmentation methods, such as the use
of neural networks for image segmentation. Neural network is not limited to a single image, it is
summarized on the basis of the image datasets, the segmentation process is constantly iteratively
corrected, and finally get the overall information of such images [3].

The contrast between the tumor area and the normal area can be shown with a more obvious
distinction. The biggest challenge faced in the process of brain tumor division both here and abroad is
how to ensure the accuracy of brain tumor images in the process of segmentation. The algorithm for
image segmentation, such as the threshold segmentation and the region segmentation which are based
on the traditional segmentation method [4], are more traditional in dealing with brain tumor
segmentation. However, FCM algorithm is time-consuming and various defects make it difficult to
guarantee the efficiency and quality when processing brain tumor images. According to the
shortcomings of the above algorithm in brain tumor image detection, in this paper, a brain tumor
image segmentation algorithm based on improved U-Net network is proposed, which realizes the
training length, strong robustness, and accurately and completely divides the brain tumor image. By
applying image recognition and deep learning techniques, U-Net network's brain tumor image
segmentation algorithm can help doctors locate diseases, analyze conditions, assist in diagnosis, and
improve their productivity without artificial intervention, which has significance clinical guiding for
diagnosis and treatment.

2. MRI Brain Tumor Segmentation

2.1. MRI

In imaging examination for brain tumor evaluation, MRI can achieve variable image contrast,
specifically by using different pulse series and changing imaging parameters corresponding to
longitudinal relaxation time and transverse relaxation time. Longitudinal relaxation time weighted and
transverse relaxation time weighted imaging signal intensity is related to the characteristics of specific
tissue [5], and enables enhanced tissue resolution to determine the anatomy more accurately; therefore,
MRI is widely used for tumor and cardiovascular detection and diagnosis [6]. Because the images
obtained by MRI have clear and fine characteristics, doctors can improve the diagnosis efficiency
during the diagnosis process. MRI conducts multiple and multiplane imaging of various parts of the
human body, which can more objectively and specifically display the pathogen location and adjacent
tissues in the human body. When scanning these images, MRI mostly adopts the mode of flat scanning
and sequence scanning to enhance the lesion positioning. Multiple sequences obtained through
scanning and highlight the different characteristics of the brain and tumor, and avoids the skull
artifacts on CT, especially the posterior skull fossa, enabling MRI to provide more accurate specific
information related to the brain tumor.

2.2. Medical Image Segmentation

There are multiple image modalities in medical imaging, reflecting images of the physiological and
physical properties of the human body in the 2D and 3D regions, such as CT, PET, ultrasound
imaging, MRI and others. This paper is primarily based on marking segmentation of parts of MRI
images with special markings, and extracting relevant features, to provide accurate and reliable data
support for doctors in clinical diagnosis and treatment and pathological research.

Medical image segmentation is a process of dividing the medical images into several unrelated
connected regions according to the brightness, shape, position and other characteristics of the medical
image. The goal of the segmentation is to study the anatomy, identify the location of the tumor and the
abnormal tissue, measure the volume size, and observe the growth of the tumor [7]. Target
segmentation from medical image is a more difficult task, because most of the medical image is
complex, lack of obvious linear characteristics. Therefore, for medical image data, it must be clear that
the purpose and effect of segmentation is to show the removal of lesions or organ extraction at first,
and then to do the image segmentation. The accuracy of the split results is influenced by some volume
effect, grayscale unevenness, artifacts, proximity of grayscale between different soft tissues [8], etc.
So using deep learning tools to optimize these image segmentation algorithms is an urgent development technology.

3. Traditional U-Net Network Model

The traditional U-Net [9] model uses Encoder-Decoder approach to apply full convolutional neural networks in the biomedical image segmentation domain [10], first published in MICCAI in 2015, in the processing of natural images today, more and more researchers begin to follow and use the U-type structures. Since the Type U structure earlier used multi-scale features for semantic segmentation tasks and object detection tasks [11], U structure plays a enlightening role in the semantic segmentation algorithm, and U-Net, as one of the ancestors, its structure has two greatest features, Type U structure and Skip-connection.

The U-Net network model includes two symmetric sections. And the entire network is bounded by the middle and it can be divided into left and right parts. The network on the left is identical to the ordinary convolutional network, which is the contraction path. And the contraction path consists of 4 blocks (Block), each block utilizes Max Pooling downsampling and 3*3 convolution module to extract features of different scales, grasp the relationship between the pixels. The entire network did not carry out full connection operation, through Encoder sampling 4 times, a total of 16 times down sampling, after the completion of the downsampling feature map channel number doubled, the feature map size changed to 32*32. While the right half is the Expansive path. It also consists of 4 blocks, some of the networks are basically symmetric to the left, using upper sampling and 3*3 convolutional module to recover the scale and fuse previous features, sampling it 4 times on the decoder. Finally, the obtained semantic feature diagram is gradually restored to the original resolution, which is to achieve the output image segmentation purposes.

4. Improved U-Net Network Model

Compared to the traditional U-Net network model, the improved U-Net network model used in the convolutional blocks in the network joins the BatchNormalization module, normalizing the input activation function of the network. As shown in figure 1, improved U-Net network model appears. It solves the impact of offset and increase of input data, accelerates the training speed and increases the learning rate, even the small learning rate can have a fast learning rate, as well as reduce or ignore the parameter selection of dropout layer and L2 regularization items. Because the parameters are normalized after the BatchNormalization module, the neurons not too much influence by the activation function, the distribution becomes obvious. After an activation function, neurons will automatically weaken or remove some neurons, and there is no need to dropout them. In addition, since L2 regularization is normalized for each training, there are large parameter changes due to different data distribution rarely occur. The increasing parameters can disrupt the training datasets and prevent training offset.

Figure 1. Improved U-Net network model diagram.
5. Analysis of the Experimental Results

5.1. Preprocessing

The data used in this paper comes from the Kaggle Open Source Brain Tumor Datasets, which contains 3929 images of brain tumor Tif format and 3929 Label label images. In this paper, we divide the training datasets, including 10% and 90% of the training datasets, and into network training and network validation datasets in a ratio of 8:2. Data augmentation is performed by picture rotation, reflection transformation, random deformation, image flipping, picture scaling, and image stretching.

5.2. Parameter Setting

U-Net network input image is 256*256*3, network convolutional core size is 3*3, the activation function is set to RELU, convolution mode is Same convolution to keep the image input and output size; the pooling layer selects the maximum value pooling, its step length is set to 2; every two convolutional BatchNormalization to accelerate model convergence. However, in the network anticonvolution phase step is 2, the convolutional core is 3*3, the convolution mode for same convolution maintains the same image input and output size, preserves the residual connection with the previous network layer while the anticonvolution, and prevents the gradient from vanishing [12].

5.2.1. Dice Coefficient

Dice coefficient is a computational measure set similarity measure function, usually used to compute the similarity of two samples with a value range of [0, 1], used to predict segmentation excellence of the model in medical image segmentation. So it is counting all the pixels of a category for Loss as a whole. Because Dice Loss directly takes the segmentation effect evaluation index as Loss to supervise the network, and ignores a large number of background pixels, solving the problem of positive and negative sample imbalance, the convergence rate is fast.

This paper uses Dice coefficient as a method to evaluate the segmentation effect and as a measure to measure the gap between the segmentation result and the label as a loss function. As shown in equation (2), it is pixel level, with the value range between 0 to 1, the real target appears in an area X, the model predicts the target is Y, then the closer the dice coefficient results are to 1 showing that the more significant the model effect.

\[
s = \frac{2|X \cap Y|}{|X| + |Y|}
\]  

(1)

5.2.2. Learning Rate

Too small a learning rate can cause convergence to slow down, but too large a learning rate can lead to divergence, and a changing learning rate scheduling system can provide faster convergence than a fixed learning rate. Traditional experience suggests that large learning rates should not be used, but may lead to “Hyperconvergence” and Regularizing Effect [13]. The adaptive learning rate used in this paper, combined with the cosine annealing learning rate idea, learning rate drops rapidly and rises rapidly, and repeats this process, allowing the model to escape the current local best and find other local best advantages to achieve higher correctness as shown in equation (2).

\[
f(x) = \begin{cases} 
(LR_{\text{MAX}} - LA_{\text{START}}) / LR_{\text{RAMPUP}}_{\text{EPOCHS}} \times x + LR_{\text{START}} , x < LR_{\text{RAMPUP}}_{\text{EPOCHS}} \\
LR_{\text{MAX}}, x < LR_{\text{RAMPUP}}_{\text{EPOCHS}} + LR_{\text{SUSTAIN}}_{\text{EPOCHS}} \\
(LR_{\text{MAX}} - LR_{\text{MIN}}) \times LR_{\text{EXP}}_{\text{DECAY}}^{(x - LR_{\text{RAMPUP}}_{\text{EPOCHS}} - LR_{\text{SUSTAIN}}_{\text{EPOCHS}} - LR_{\text{MIN}})} 
\end{cases}
\]  

(2)

where f(x) is the learning rate, x is the number of iterations, LR_MAX is the maximum value of learning rate decay interval; LR_MIN is the minimum value of learning rate decay interval; LA_START is the initial value of learning rate; LR_RAMPUP_EPOCHS is the maximum number of iterations; LR_SUSTAIN_EPOCHS is the number of iterations.
The adaptive learning rate is set so that the gradient descent can be more accurate and reduce the training length, thus realizing related operations such as image segmentation processing accurately and completely.

5.2.3. Datasets. This paper selects the LGG Segmentation Datasets, which includes brain MRI and manual FLAIR exception segmentation templates. Images came from The Cancer Imaging Archive (TCIA). They correspond to 110 patients collected by The Cancer Genome Atlas (TCGA) in low-grade glioma, with at least Fluid Attenuated Inversion Recovery (FLAIR) sequences and genomic cluster data available. There are 3 channels per image. There were 1013 sequences, namely before and after FLAIR, (in channel order). Nine post-contrast sequences were missing and six pre-contrast sequence were absent. The missing sequence is replaced with the FLAIR sequence, making all the images three-channel. The mask is a binary single-channel image. They split the FLAIR anomalies in the FLAIR sequence (applicable in all cases).

5.3. Results Analysis
After 100 Epochs training for the model, the traditional U-Net training parameters diagram are shown in figure 2, and the improved network training parameters diagram are shown in figure 3.

![Figure 2. Traditional U-Net training parameters diagram.](image)

![Figure 3. Improved network training parameters diagram.](image)
Figure 4 shows the traditional network training learning rate. This paper uses the adaptive learning rate and the cosine annealing. The results are shown in the learning rate diagram of cosine annealing training diagram in figure 5. By improving the learning rate, we escape the local minimum and find the path to the global minimum.

![Figure 4. Traditional network training learning rate diagram.](image)

Figure 5. Learning rate diagram of cosine annealing training diagram.

Figure 6 shows the model results test, the red area, the green area are the marked brain tumor site and no tumor in the brain. The accuracy of the final classification results reached 99.81%, intersection ratio reached 0.963 and DICE coefficient reached 0.952. Among them, the predicted segmentation results of the figure in the first row, the second row left figure and the left figure in the third row were highly compatible with the real part of the brain tumor in the actual situation and with a high accuracy. While predicting the normal brain image, there was no dividing area, and the results showed that there was no tumor in the brain.

![Figure 6. Model results test.](image)
6. Conclusion
In this paper, the algorithm framework of traditional U-Net network is improved with BatchNormalization algorithm module to normalize the input of the input activation function of the network, greatly improves the training speed and increases the learning rate, and it has greater accuracy for the recognition and segmentation of brain tumors.

In future work, we will continue to improve the image processing, such as separating the non-brain tissue in the image to make the data clearer, with more accurate segmentation and better results. In addition, expanding the identification of the segmentation object, which is the tumor recognition of other parts of the human body, improving the identification and segmentation system are to achieve more human medical image segmentation system.

Acknowledgments
This work was supported by Guangdong University Science Research Project (No.2019KTSCX256), the Study of Technology (No.2020KQNCX105), and the Intelligent Science and Technology Team (No.2019XJJXTD01).

References
[1] Ai L M, Li T D, Liao F Y and Shi K Z 2020 Attention U-Net-based MRI segmentation Laser and Optoelectronics Advances 673 (14) 279-286.
[2] Xiao D W, Zhang S L, Wang C, et al. 2018 A MRI brain tumor image segmentation method based on an optimized U-net network model CN109087318A.
[3] Levine M D and Shaheen S I 1981 A modular computer vision system for picture segmentation and interpretation *IEEE Transactions on Pattern Analysis and Machine Intelligence* 3 (5) 540-557.

[4] Shenton M E, Dickey C C, Frumin M, et al. 2001 A review of MRI findings in schizophrenia *Schizophrenia Research* 49 (1-2) 1-52.

[5] Liu Y and Chen S 2017 Review of medical image segmentation method *Electronic Science Technology* 30 (8) 169-172.

[6] Jiang G P, Qin W J, Zhou S J, et al. 2015 Medical image segmentation and its development status *Journal of Computer* 38 (6) 1222-1242.

[7] Liu L, et al. 2019 Fractional wavelet scattering network and applications *IEEE Transactions on Biomedical Engineering* 66 (2) 553-563.

[8] Xu H X 2018 Value of CT Diagnosis and differential diagnosis of cranirarian tumor and magnetic resonance imaging *Journal of Practical Medical Technology* 25 (11) 1216-1217.

[9] Abdulkadir A, Ronneberger O, Wolf R C, et al. 2012 Functional and structural MRI biomarkers to detect pre-clinical neurodegeneration *Current Alzheimer research* 10 (2).

[10] Ronneberger O, Fischer P and Brox T 2015 U-Net: Convolutional networks for biomedical image segmentation *International Conference on Medical Image Computing and Computer-Assisted Intervention* (Berlin: Springer).

[11] Garcia-Garcia A, Orts-Escolano S, Oprea S, et al. 2017 A review on deep learning techniques applied to semantic segmentation arXiv:1704.06857.

[12] Li Q F and Mou S X 2021 Study on Brain Tumor Segmentation Methods Based on MRI Images (Nanjing University).

[13] Kong X M 2019 *Research and Implementation of Brain Tumor Segmentation Algorithm Based on Deep Learning* (Shandong University).