Automatic segmentation of low-grade glioma in MRI image based on UNet++ model

Dan Xu¹, Xidong Zhou, Xuefen Niu and Junwei Wang*  
Department of computer science and Engineering, Northeastern University, Shenyang 110000, China  
*Corresponding author’s e-mail: 13684634211@163.com  
¹1871728@stu.neu.edu.cn

Abstract: Glioma is one of the common brain tumors, and the prognosis of patients with low-grade glioma is relatively good. Therefore, this paper takes the segmentation of low-grade glioma as the research direction, and proposes an automatic segmentation algorithm of low-grade glioma MRI image based on UNet++. Firstly, the sample data in the data set is normalized, and then the sample data is divided into training set, verification set and test set, and data augmentation is performed on the training set; finally, all the data in the training set is used to train the network model. In the training process, in order to alleviate the over-fitting problem of model training, we draw into the dropout after each convolution layer. In order to verify the effectiveness of the proposed model, all intracranial tumor images of patients in the test sample set are selected for segmentation, and the final average Dice coefficient can reach: 89.1%. Compared with the segmentation algorithm based on U-Net, the average Dice coefficient of this algorithm is increased by 3.89%. The experimental results show that this algorithm can achieve better automatic segmentation results for low-level glioma images, provide effective reference for doctors' diagnosis and surgery, and alleviate the problems of low accuracy and low efficiency caused by manual annotation of images only based on personal experience of doctors.

1. Introduction
Medical image data plays a significant role in clinical diagnosis and medical research. In recent years, the medical image processing technology has gradually developed into an important research field, in which the use of computer technology to segment medical images can reduce the measurement error of doctors, reduce the workload of doctors, and improve the efficiency of medical work. It is a research direction with high clinical application value. Medical image segmentation technology designs automatic or semi-automatic segmentation algorithm to segment the interested part of medical image, and makes the segmentation result close to the original structure of the region as much as possible. Automatic and high accuracy intracranial tumor segmentation algorithm is very important for clinical accurate diagnosis and treatment of intracranial tumor. Measuring tumor size with segmented image can assist doctors to make the best diagnosis and treatment plan for patients, and tumor segmentation image can assist doctors to demarcate radiotherapy target area. In the field of medical image segmentation, tumor image segmentation has always been a laborious and time-consuming task for further research.

Brain tumor is one of the most deadly diseases. In 2018, there were about 297000 new cases of brain cancer, accounting for 1.6% of all new cancer cases, and 241000 brain cancer deaths, accounting for 2.5% of all cancer deaths[1]. Gliomas originate from the pathological changes of glial cells, ac-
counting for 40%-50% of intracranial tumors. They are the result of the interaction between congenital genetic risk factors and environmental carcinogenic factors. Gliomas can be divided into high-grade gliomas and low-grade gliomas. High-grade gliomas are highly malignant and are prone to recurrence and metastasis. The median survival time of high-grade gliomas is about 3-4 years, or even shorter. The growth cycle of low-grade glioma is different from that of malignant tumor, but it still has the biological ability to transform into high-grade malignant tumor. Therefore, low-grade glioma is the focus of intracranial tumor segmentation. In order to improve the survival rate of patients, only for clinically common gliomas, accurate diagnosis of early stage of low-grade gliomas can better prolong the life of patients[2][3]. The contrast of MRI images of brain tumors is obvious, which is conducive to accurate segmentation of brain tumors. Efficient, accurate and automatic segmentation of brain tumor is of great significance for its diagnosis, treatment and postoperative tracking. However, in clinical application, it mainly depends on the personal experience of doctors in manually segmenting tumor tissue, and there is still a lack of accurate, repeatable and adaptable brain tumor segmentation method.

2. Related Work
In the field of brain tumor image segmentation, unsupervised segmentation is often used in the early stage. Unsupervised segmentation is a technology of image segmentation in the absence of labeled images. It makes statistical analysis of medical images with the help of computer to solve the problem of pattern recognition. In pace with the high-speed development of computer technology and theory, machine learning has been widely used in various fields. In the field of medical image segmentation, there are a slew of traditional machine learning algorithms for intracranial tumor segmentation, such as support vector machine (SVM), conditional random field (CRF) and random forests (RF). Supervised segmentation uses labeled image data, which can accomplish tasks in different scenes simply by changing the training set.

With the rise of artificial intelligence, convolutional neural network has made a major breakthrough in the field of computer vision, and the research of computer-aided diagnosis of brain tumor based on CNN has also received extensive attention. In 2014, fully convolutional network (FCN) was introduced by J. Long et al. [4]. They used fully convolutional network to segment images end-to-end. In the same year, Ian J. Goodfellow et al. [5] proposed GAN (Generative Adversarial Networks), which made the network no longer rely on data volume and labels. Subsequently, in 2015, O. Ronneberger et al. [6] proposed the most famous CNN architecture U-Net in medical image analysis. The main architecture innovation in U-Net is the combination of the same number of up-sampling layers and down-sampling layers. G. Huang et al. [7] proposed the architecture called DenseNet with dense connection. These deep neural network models have been successfully applied to many computer vision tasks, which vastly promoted the development of MRI brain tumor segmentation research based on deep learning. Among them, the famous U-Net network has achieved good results in medical image segmentation. The architecture consists of a contracting path to capture context and a symmetric expanding path that enables precise localization[6]. The skip connection structure of the encoder in U-Net network links the high-resolution features of the encoder with the up-sampling output of the decoder, and finally obtains more accurate segmentation results. Due to the excellent performance of U-Net in image segmentation, many medical image segmentation networks based on U structure of U-Net are produced. For example, UNet++[8]. UNet++ uses nested dense skip connections to reduce the semantic gap between encoder and decoder. This is in contrast to the concatenation commonly used in U-Net, which directly fast-forward high-resolution feature maps from the encoder to the decoder network, resulting in the fusion of semantically dissimilar feature maps [8]. Although a large number of deep learning models have appeared in the field of medical image segmentation, there are not many deep learning algorithm models for brain tumors. Therefore, based on the deep learning theory, this paper proposes a low-level glioma segmentation algorithm based on the UNet++ model.
3. Methods
The whole process of this algorithm mainly includes three steps: data preprocessing, data augmentation and brain tumor image segmentation based on UNet++ network model.

3.1. Data preprocessing
In the field of image processing, it is usually necessary to standardize the input image. After standardization, the original data can be converted to the same order of magnitude, and then the processed image is fed into the neural network. The common method of data standardization is to normalize the data. Normalization is a process of transforming the image into a fixed standard form through a series of standard processing and transformation. The fixed standard form is invariant to the geometric changes of images, that is, normalization can find out the invariants in the images, so as to distinguish whether these images are the same or belong to the same series. Normalization can facilitate the later data processing and accelerate the convergence speed of the whole network when training the network. In order to display and calculate conveniently, the gray value of the image is normalized to 0-1. Normalization can reduce the interference caused by uneven light in medical images.

3.2. Data augmentation
The network model used in this paper has millions of parameters, so it needs a sufficient proportion of sample data for training. Due to the lack of labeled brain tumor image data, it is difficult to train a deep neural network model. In order to obtain more data, we need to use some random image processing methods to increase the available training samples. A robust CNN framework can be invariant to translation, viewpoint, size, illumination or any combination of the above. In this algorithm, three data augmentation methods are used, including flip (horizontal and vertical flip), rotation (random angle rotation), and translation (moving image along X or Y direction or both). Through these three data augmentation methods, the original data set can be expanded, and the over-fitting of the network can be avoided. Through the transformation of training image, the model with stronger generalization ability can be obtained, which can better adapt to specific application scenarios.

3.3. Network framework
Before UNet++, U-Net has achieved remarkable results in medical image segmentation. U-Net network is most commonly used in medical image processing. It is first used for cell wall segmentation, and then used for pulmonary nodule detection and fundus vascular image extraction. The U-Net network has achieved outstanding results in these problems, and has overcome the problems of fewer data capacity, fuzzy boundary and high gray scale in medical image analysis. The U-Net network framework includes an encoder for processing input MRI images and a decoder for generating output. The encoder decomposes the image into different levels of feature maps, and extracts the coarse-grained features of the feature maps. The decoder restores the feature maps of each layer by up-sampling, and the concatenation cascades the features of each layer of the encoder with the features obtained by the transpose convolution operation in the decoder, thus reducing the loss of accuracy in the feature extraction process.

In this paper, we use the UNet++ network framework, which is a variant of U-Net network. The whole framework includes 15 feature extraction modules, and each module has two convolution layers. Each convolution layer adopts a $3 \times 3$ convolution kernel with step size of 1 for convolution. At the same time, each convolution layer is followed by a batch normalization (BN) layer and a leaky-relu layer to improve the performance of the network. While keeping the basic architecture of U-Net unchanged, UNet++ changes its simple concatenation to dense skip connection. In UNet++, when the high-resolution feature maps from the encoder network are fused with the corresponding semantically-rich feature maps from the decoder network, the model can capture the fine-grained details of brain tumors more effectively. When the feature maps of decoder and encoder network are similar in semantics, the network will learn the segmentation task more easily. In contrast, U-Net quickly forwards the high-resolution feature mapping from the encoder to the decoder, resulting in the fusion of different
semantic feature maps, resulting in a semantic gap between low-level features and high-level features, which fails to achieve the best prediction effect. UNet++ fuses the semantic information between modules in different layers from shallow to deep, and fuses the semantic information of dense skip connection between modules in the same layer, so that the model can fully and effectively capture the detailed features of the target and improve its segmentation performance [9].

The network structure of UNet++ is shown in the Figure 1, with layers 1 to 5 from top to bottom. The size of feature map of the same layer is alike, and each node represents a feature extraction module.

In this network, from top to bottom, the features of L1 encoder are fused with the up-sampled output of L2 layer. Then the fused output is fused with up-sampled output of L3 layer. This cycle continues until there is no corresponding up-sampling module in the next layer.

$$x^{i,j} = \begin{cases} H(x^{i-1,j}) & j = 0 \\ H(\left[ x^{i,k} \ast U(x^{i+1,j-1}) \right]) & j > 0 \end{cases}$$

(1)

The stack of feature maps represented by $x^{i,j}$ is shown in Equation(1) [8]. $x^{i,j}$ in Equation(1) represent the current node, which is the output of feature extraction module. $i$ represents the sequence in which the encoder performs down-sampling, ranging from 1 to 5. $j$ is the module serial number of the same layer, and $j = 0$ represents the encoder feature extraction module. $H(\cdot)$ represents the convolution operation in the feature extraction module. $U(\cdot)$ is the up-sampling operation. $[\cdot]$ denotes the concatenation of features in dense skip-connections.

### 3.4 Loss function and training strategy

In this model, Dice coefficient loss [10] function is used as the loss function of the predicted result for each output, which is defined as Equation (2).

$$L(Y_{GT}, Y_{P}) = \frac{1}{N} \sum_{b=1}^{N} \left( 1 - \frac{2 \times Y_{GT}^b \cap Y_{P}^b}{Y_{GT}^b + Y_{P}^b} \right)$$

(2)
In the Equation (2), \( Y_{GT} \) represents the labeling of brain tumors and \( Y_p \) represents the prediction results. \( N \) is the batch size. \( Y_{GT}^b \) is the segmentation annotation of the B image, and \( Y_p^b \) is the prediction probability of the B image.

In order to improve the generalization ability of the model and prevent the problem of overfitting in the process of model training, this paper uses the regularization method of neural network, dropout. Dropout is an optimization method for artificial neural networks with deep structure. In the learning process, some weights or outputs of hidden layer are returned to zero randomly to reduce the co-dependency between nodes, so as to realize the regularization of neural networks and reduce the structural risk [11]. Dropout can make the prediction result independent of a certain part of neurons, which makes the model more generalized.

4. Results

4.1. Dataset

The data used in the experiment are from the brain MRI segmentation data set published on the machine learning competition platform kaggle (https://www.kaggle.com). The author of the dataset is Mateusz Buda. An example of the dataset is shown in Figure 2.

![Brain tumor images of the patients.](image)

This data set contains brain MRI images and manual flair anomaly segmentation mask. The images are from The Cancer Imaging Archive (TCIA). They correspond to 110 TCGA low-grade gliomas. The data set contains 3929 images, of which 2556 is positive and 1373 is negative, and the image size is 256 * 256. The training set, verification set and test set are divided according to the ratio of 7:2:1, including 2828 training sets, 708 verification sets and 393 test sets.

4.2. Metrics

Dice coefficient [12], a common index in medical images, is used to measure the coincidence degree between the actual segmentation results and the theoretical segmentation results. The evaluation definition is shown in the Equation (3).

\[
    f_{Dice} = \frac{2 \times |Y_{GT} \cap Y_p|}{|Y_{GT}| + |Y_p|}
\]  

(3)

\( Y_{GT} \) is the labeled pixel and \( Y_p \) is the prediction result. Dice coefficient is used to evaluate the model segmentation effect. The higher the dice value, the better the effect of the model.

4.3. Segmentation results

In order to facilitate the construction, training, verification and testing of the segmentation network, this experiment uses the Keras deep learning framework to build the segmentation network, and uses python programming language to code. This experiment was completed on NVIDIA K80 GPU pro-
vided by kaggle (https://www.kaggle.com/) platform. The training time of each epoch was about 2 minutes. It takes about 5 hours to complete each training. Four tumor MRI images were used to show the segmentation accuracy of the algorithm. Figure 3 shows the segmentation results of this algorithm on some tumor images. As can be seen from Figure 3, the results of tumor segmentation using the network framework in this paper are close to the results of labeling.

![Figure 3. The final segmentation results of this algorithm.](image)

In order to verify the effectiveness of our algorithm in low-level glioma segmentation, this paper selects one image segmentation algorithm to compare with our algorithm. The algorithm is U-Net [6] based segmentation model.

| Algorithm  | Dice   |
|------------|--------|
| U-Net      | 0.8521 |
| UNet++     | 0.8910 |

As can be seen from Table 1, compared with the segmentation model based on U-Net, the mean Dice coefficient of this model is improved by 0.0389. Experimental results show that the proposed algorithm can achieve better automatic segmentation effect for low-level glioma images.

5. Conclusion
In this paper, an automatic segmentation model based on UNet++ network framework is proposed for the segmentation of low-grade glioma, and the similarity of segmentation is verified and analyzed by Dice coefficient. In this study, the mean Dice coefficient of the model based on UNet++ network for glioma segmentation reached 89.1% and achieved a better segmentation result. Compared with the segmentation model based on U-Net, this algorithm can achieve higher segmentation accuracy and vastly shorten the training time. However, the segmentation of medical image needs higher accuracy. In the experimental process, it is found that when the tumor block in the low-level glioma image is small or the edge is fuzzy, the existing algorithm model has the situation of missing segmentation or false segmentation, which is difficult to complete segmentation with pinpoint accuracy. In order to solve this problem, we plan to design a more effective automatic segmentation algorithm for low-level glioma by using the 3D data of glioma and the relevant feature information of 3D image of glioma.

Acknowledgements
I sincerely thank my friends and teachers for taking time out of their busy work and study to help me. They inspired me when I met bottlenecks and encouraged me when I encountered setbacks.

References
[1]. Li Q., Bai K. X., Zhao L. and Guan X. (2020) Progresss and challenges of MRI brain tumor image segmentation. Journal of Image and Graphics, 25(3):0419-0431.
[2]. Menze BH., Jakab A., Bauer S., et al. (2015) The multimodal brain tumor image segmentation benchmark (BRATS)[J].IEEE Trans Medical Imaging, 34(10):1993-2024.

[3]. Bakas S., Akbari H., Sotiras A., et al. (2017) Advancing the cancer genome atlas glioma mri collections with expert segmentation labels and radiomic features[J].Sci Data, 4:170117.

[4]. Shelhamer E., Long J., Darrell T. (2014) Fully Convolutional Networks for Semantic Segmentation[J].IEEE Transactions on Pattern Analysis & Machine Intelligence, 39(4):1-1.

[5]. Goodfellow I. J., Pouget-Abadie J., Mirza M., et al. (2014) Generative Adversarial Networks[J].Advances in Neural Information Processing Systems, 3:2672-2680.

[6]. Ronneberger, O., Fischer, P., Brox T. (2015) U-net: Convolutional networks for biomedical image segmentation. In: MICCAI. LNCS, vol. 9351, pp. 234–241. Springer (2015).

[7]. Huang G., Liu Z., Weinberger K. Q., and Van der Maaten L. (2017) Densely connected convolutional networks. In Proceedings of the IEEE conference on computer vision and pattern recognition, volume 1, page 3, 2017.

[8]. Zhou Z. W., Siddiquee M. M. R., Tajbakhsh N., and Liang J. M. (2018) UNet++: A nested U-Net architecture for medical image segmentation. In Deep Learning Medical Image Analysis and Multimodal Learning for clinical Decision support. Cham: Springer:3-11.

[9]. Xing Y. Y., Yang F., Tang Y. J., and Zhang L. Y. (2020) Ultrasound fetal head edge detection using fusion UNet++. Journal of Image and Graphics, 25(02):0366-0377.

[10]. Milletari F., Navab N., Ahmadi S-A. (2016) V-Net: Fully Convolutional Neural Networks for Volumetric Medical Image Segmentation[C].International Conference on 3D Vision; IEEE International Conference on 3D Vision, et al. 2016 Fourth International Conference on 3D Vision: 3DV 2016: 25-28 October 2016, Stanford, California, USA: proceedings. Piscataway, NJ: IEEE, 2016: 565-571.

[11]. Hinton G. E., Srivastava N., Krizhevsky A., et al. (2012) Improving neural networks by preventing co-adaptation of feature detectors[J].Computer ence, 3(4):pages. 212-223.

[12]. Dice L. R. (1945) Measures of the Amount of Ecologic Association Between Species[J]. Ecology, 1945, 26(3).