



## Research Article

<https://doi.org/10.1631/jzus.B2000381>



# 3D brain glioma segmentation in MRI through integrating multiple densely connected 2D convolutional neural networks

Xiaobing ZHANG<sup>1</sup>, Yin HU<sup>1</sup>, Wen CHEN<sup>1</sup>, Gang HUANG<sup>2</sup>, Shengdong NIE<sup>1✉</sup>

<sup>1</sup>School of Medical Instrument and Food Engineering, University of Shanghai for Science and Technology, Shanghai 200093, China

<sup>2</sup>School of Health, Shanghai University of Medicine & Health Sciences, Shanghai 201318, China

**Abstract:** To overcome the computational burden of processing three-dimensional (3D) medical scans and the lack of spatial information in two-dimensional (2D) medical scans, a novel segmentation method was proposed that integrates the segmentation results of three densely connected 2D convolutional neural networks (2D-CNNs). In order to combine the low-level features and high-level features, we added densely connected blocks in the network structure design so that the low-level features will not be missed as the network layer increases during the learning process. Further, in order to resolve the problems of the blurred boundary of the glioma edema area, we superimposed and fused the T2-weighted fluid-attenuated inversion recovery (FLAIR) modal image and the T2-weighted (T2) modal image to enhance the edema section. For the loss function of network training, we improved the cross-entropy loss function to effectively avoid network over-fitting. On the Multimodal Brain Tumor Image Segmentation Challenge (BraTS) datasets, our method achieves dice similarity coefficient values of 0.84, 0.82, and 0.83 on the BraTS2018 training; 0.82, 0.85, and 0.83 on the BraTS2018 validation; and 0.81, 0.78, and 0.83 on the BraTS2013 testing in terms of whole tumors, tumor cores, and enhancing cores, respectively. Experimental results showed that the proposed method achieved promising accuracy and fast processing, demonstrating good potential for clinical medicine.

**Key words:** Glioma; Magnetic resonance imaging (MRI); Segmentation; Dense block; 2D convolutional neural networks (2D-CNNs)

## 1 Introduction

Glioma is one of the most common and aggressive types of primary brain tumor, with a low survival rate (Zhuge et al., 2017). The World Health Organization (WHO) has divided brain gliomas into four grades according to their severity: Grade 1 and Grade 2 are mildly dangerous and known as low-grade glioma (LGG), slowly advancing tumors, while Grade 3 and Grade 4 are high-grade glioma (HGG) tumors with high malignancy (Mohan and Subashini, 2018). Accurate quantification of tumor size can be considered as a method of efficacy evaluation. Precise segmentation also provides a basis for developing a radiotherapy plan and surgical strategy. Therefore, accurate and reliable

segmentation of brain glioma is a goal of great clinical significance. Up to now, clinical segmentation of brain glioma is generally carried out manually. In addition, the manual segmentation of brain glioma is time-consuming and laborious; its success also relies on the doctor's clinical experience and is subject to subjective factors (Mengqiao et al., 2017). Hence, research on accurate automatic segmentation methods of brain glioma has been a long-time goal of medical image processing. However, gliomas may appear in any position of the brain, with varied shape, appearance, and size, making it a challenge to segment gliomas automatically and accurately (Udupa and Vishwakarma, 2016).

Magnetic resonance imaging (MRI) is a widely used, non-intrusive imaging modality, because it gives sensitive tissue contrast. It is a unique imaging technique that is especially suitable for human brain tumors. Thus, MRI is the most commonly used image for brain tumor detection and segmentation. Multiple sequences, including T2-weighted fluid-attenuated inversion recovery (FLAIR), T1-weighted (T1), T1-weighted

✉ Shengdong NIE, nsd4647@163.com

✉ Xiaobing ZHANG, <https://orcid.org/0000-0001-9344-2630>  
Shengdong NIE, <https://orcid.org/0000-0001-7825-4455>

Received July 15, 2020; Revision accepted Dec. 2, 2020;  
Crosschecked Apr. 12, 2021

© Zhejiang University Press 2021

contrast-enhanced (T1c), and T2-weighted (T2) sequences, were assessed to jointly diagnose and segment brain gliomas. Notably, the intensity inhomogeneity of MRI data further increases the difficulty of automatic segmentation of brain gliomas.

Recent brain glioma research has mainly analyzed lesion tissue in images generated by image processing, pattern recognition, and artificial intelligence. The existing automatic and semi-automatic brain tumor segmentation methods can be broadly categorized as generative model-based or discriminative model-based methods (Goetz et al., 2015). The generative models need prior information, such as tumor shape and tumor appearance, but have the advantage of fast convergence during training. The discriminative models include random forests, support vector machine, and conditional random field (CRF). Segmentation methods based on the discriminant model can directly learn the characteristics of tumors without prior knowledge. In the classification recognition task, the accuracy of methods based on discriminant model is higher than that based on the generation model (Zhao and Jia, 2015). More recently, traditional methods and deep learning methods have been widely used in the segmentation of gliomas. For example, Li et al. (2018) proposed a unified glioma segmentation algorithm that combines spatial fuzzy c-mean clustering, region growth, and post-processing. Islam et al. (2020) proposed an efficient multilevel segmentation method that combines optimal threshold and watershed segmentation technique followed by a morphological operation to separate the tumor. Currently, segmentation methods based on convolutional neural networks (CNNs) have achieved good results in brain glioma segmentation tasks. Both two-dimensional (2D)-CNNs and three-dimensional (3D)-CNNs have been adopted for segmentation of brain gliomas. Zikic et al. (2014) developed a brain tumor segmentation method based on 2D-CNNs in which they built a five-layer CNN segmentation model with four channels of  $19 \times 19$  2D image patches as the input. The four channels were image patches of T1, T2, T1c, and FLAIR. Havaei et al. (2017) proposed a deep learning model with two pathways of CNNs, including a convolution pathway and a fully connected pathway. Dvořák and Menze (2016) modeled the multi-class brain tumor segmentation task as three binary segmentation sub-tasks; each sub-task was solved using CNNs, with each task segmenting a

sub-region of a brain glioma. Most previous brain tumor segmentation tasks based on 2D-CNNs use image patches (local areas in MRI images) to train the models, transform the segmentation tasks into classification tasks, and sort the image patches into (1) edema region, (2) necrosis/non-enhancing tumor, and (3) enhancing tumor. Then the classification results of each image patch are used to mark its central voxel for tumor segmentation. In addition, Krishna et al. (2019) used an end-to-end 2D U-Net network to segment glioma. The methods of brain glioma segmentation based on 2D-CNNs, however, which use 2D MRI image patches as input, cannot make full use of the neighborhood and spatial features of each pixel. To make full use of the 3D information of brain glioma MRI, it is more effective to directly use the 3D data of brain glioma MRI for segmentation. Urban et al. (2014) first applied the 3D-CNN model to segmentation of gliomas; they proposed a 3D-CNN segmentation model with inputs of  $9 \times 9 \times 9$  in size and used four modal MRI data to achieve 3D segmentation of a brain glioma. Chen et al. (2018) developed a densely connected 3D-CNNs with an input image size of  $38 \times 38 \times 38$ . The output is not a classification probability value of individual pixels, but a central image body of  $12 \times 12 \times 12$  in size. This output method greatly improves the brain glioma segmentation speed. The dense connection method of the convolutional layer in the segmentation model solves the problem of gradient disappearance during the network training process to a certain extent, and improves the segmentation accuracy of the brain glioma. Guo et al. (2019) proposed a cascade of global context CNNs; their network is a modification of the 3D U-Net, consisting of residual connection, group normalization, and deep supervision. Baid et al. (2020), Pan et al. (2020), and Zhou et al. (2020) designed different 3D segmentation models of gliomas, based on 3D-CNNs; each model can segment various radiologically identifiable sub-regions such as edema region, enhancing tumor, and necrosis. The brain glioma segmentation model based on 3D-CNNs takes 3D MRI image patches as input. Even if the 3D information of MRI data is used fully, it also increases the network parameters and cost of calculation, taking up more memory. Though many algorithms have been proposed over the decades to achieve this task, most have shortcomings that limit their utility in routine clinical practice.

To address the above challenges in glioma segmentation, we have proposed a fully automatic brain tumor segmentation model with densely connected architecture based on CNNs. The aim of this study is to design a fully automated brain tumor segmentation algorithm that will accurately segment the tumors and act as an assistive tool for radiologists for exact tumor quantification. The main contributions are as follows:

(1) The boundary of brain glioma edema is ambiguous and difficult to segment. FLAIR and T2 superimposed fusion was used as one of the preprocessing steps to enhance the edema region of brain gliomas, so that the boundary of the brain glioma edema area was clear and easy to segment.

(2) For the network structure design, 2D dense connection blocks were added; these blocks consist of a series of convolutional layers. Each of the layers accepts the feature of all of the layers in front of it as an input. In addition, in the process of feature extraction, we also extracted features in stages; both low-level features and high-level features can be used, so that the network did not miss any information as the number of network layers increased.

(3) We defined a new loss function that makes model training easier and avoids network over-fitting.

(4) In order to make full use of the neighborhood and spatial characteristics of pixel points, different views of the multi-modal MRI data were used as the training dataset. Three densely connected 2D-CNNs segmentation models with the same structure were

trained. Then the segmentation results of three views were fused and post-processed, and the brain glioma was 3D-segmented by multiple 2D-CNNs.

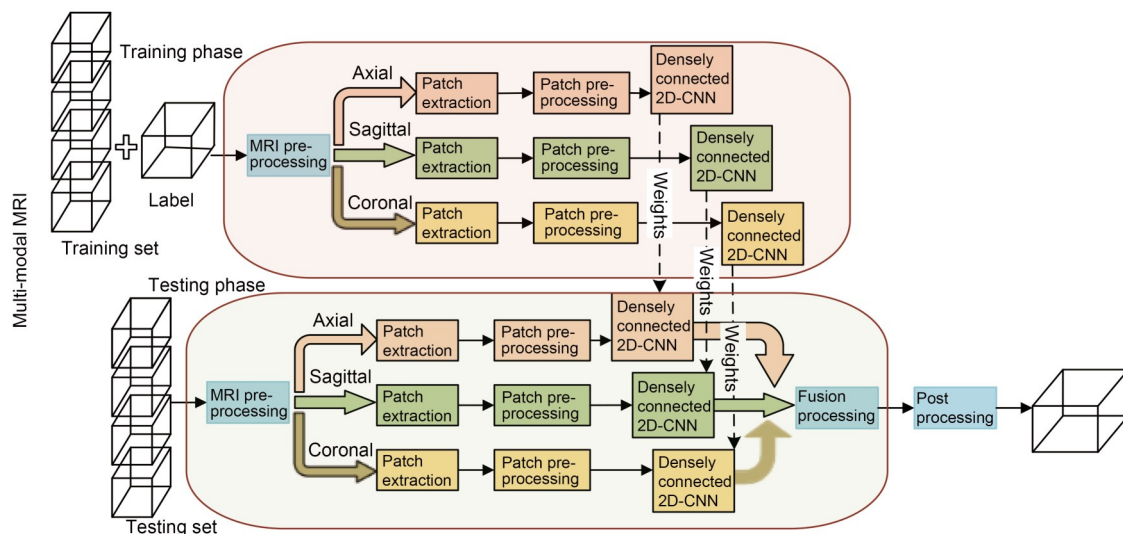
## 2 Methods and experiments

### 2.1 Methods

The proposed densely connected 2D-CNN model for brain tumor segmentation is illustrated in Fig. 1. This model includes the following steps: (1) multi-modal MRI pre-processing; (2) training three densely connected 2D-CNN segmentation models; (3) combining the segmentation result of three views using fusion strategy; (4) post-processing the fused image.

#### 2.1.1 Pre-processing

The MRI images extracted from volumetric data have bias field distortion and intensity inhomogeneity due to the principles of MRI, technical limitations, and various factors in the image acquisition process. These limitations cause the intensity of the same tissues in different patients to vary across the image. Kamnitsas et al. (2017) applied N4ITK bias correction to T1 and T1c MRI sequences as a step in pre-processing, which resulted in removal of the intensity gradient of each scan. Zhao et al. (2018) adopted a robust intensity normalization method to make MRI scans of different patients comparable; this technique effectively reduced the non-uniform intensity of the MRI image. The



**Fig. 1** Flowchart of the proposed method. MRI: magnetic resonance imaging; 2D-CNN: two-dimensional convolutional neural network.

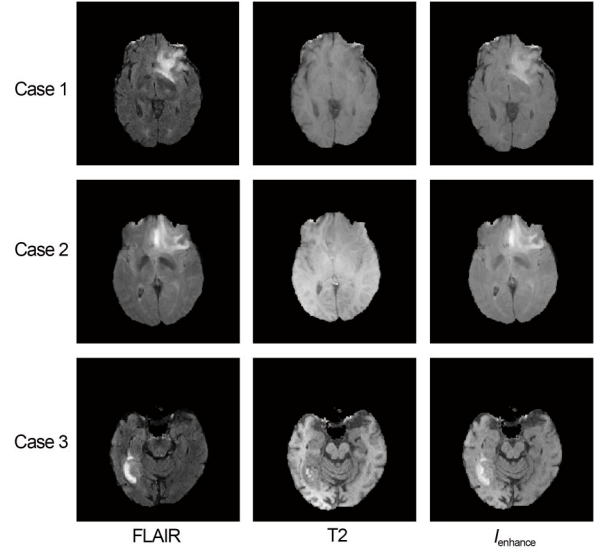
blurred edge of the brain glioma and unclear boundary of the edema region increase the difficulty of segmentation. The T2 sequence of an MRI is the generated MR signal when the transverse magnetic vector attenuates and disappears; it highlights the difference in T2 transverse relaxation between tissues and is suitable for observing diseased tissue. The T2 sequence shows fluid-filled tissues as bright signals. In FLAIR, brain ventricles show low signals, and tissue lesions show bright signals. To avoid the influence of bright ventricular signals, due to the high fluid content in gliomas, on segmentation of glioma, the two modalities of T2 and FLAIR can be combined to segment the glioma as a whole. Therefore, in order to segment the edema portion of the glioma more completely, a new pre-processing method was presented based on the characteristics of brain glioma. The first step in our pre-processing is algebraic superposition fusion of the two modal images of FLAIR and T2 to get the enhanced image of edema ( $I_{\text{enhance}}$ ), which is given by

$$I_{\text{enhance}} = (\text{FLAIR} + \text{T2}) / (\text{FLAIR} \times \text{T2}). \quad (1)$$

Some examples after pre-processing are indicated in Fig. 2. The normalization of zero mean and unit variances of FLAIR,  $I_{\text{enhance}}$ , T1c, and T1 data is regarded as the second step of our pre-processing. Such processing preserves the original information of the image to the greatest extent. Finally, patches are normalized with respect to mean and variance.

### 2.1.2 Classification

The structure of our proposed densely connected 2D-CNNs is shown in Fig. 3. The proposed architecture takes patches of multiple modalities as input and predicts the class of center pixel in respective patches. Simonyan and Zisserman (2015) found that when a network model has the same performance, the number of parameters in the network can be reduced and take up less memory by using a smaller convolution kernel. Inspired by their research, a small convolution kernel of  $3 \times 3$  is used in our network structure. VggNet and ResNet are commonly used medical image segmentation models (Simonyan and Zisserman, 2015; He et al., 2016). These models can only learn higher-level features from the characteristics of the previous layer. Low-level features cannot be fully utilized, and important information may be ignored. DenseNet was proposed



**Fig. 2** Three cases from the 2018 global multi-modal brain tumor segmentation challenge (BraTS2018) training datasets (<https://www.med.upenn.edu/sbia>), showing representative improvements when using the fusion of two modal images of T2-weighted fluid-attenuated inversion recovery (FLAIR) and T2 approach.  $I_{\text{enhance}}$ : enhanced image of edema.

to overcome the problem of low feature utilization (Huang et al., 2017). The greatest advantage of this program is that each layer accepts the features of all the previous layers as input and reuses them. So, we added two dense connection blocks when designing the network structure. A densely connected block is formulated as:

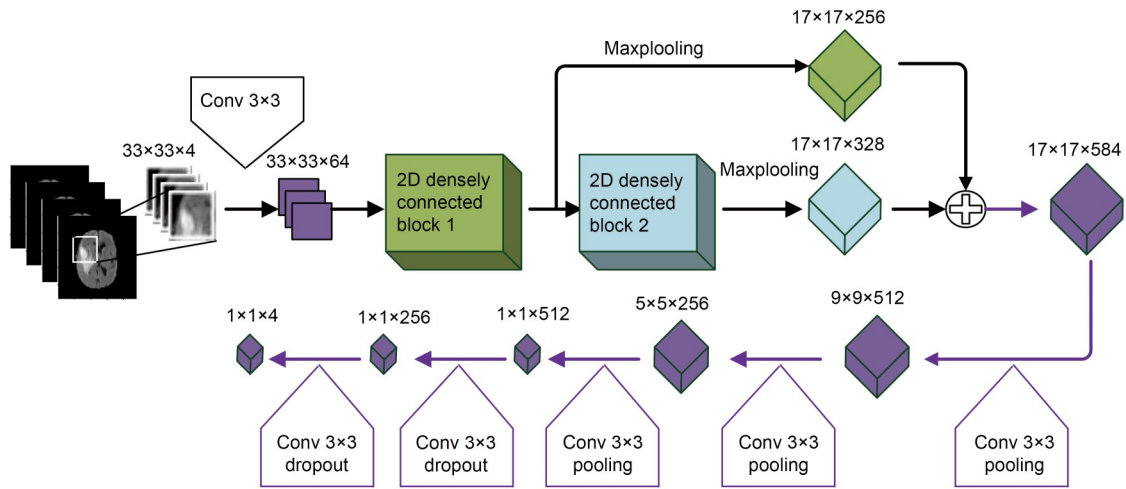
$$y(m+1) = f([y(0), y(1), \dots, y(m)]), \quad (2)$$

$$f(y(n)) = W * \delta(B(y(n))), n \in \{0, 1, \dots, m\}, \quad (3)$$

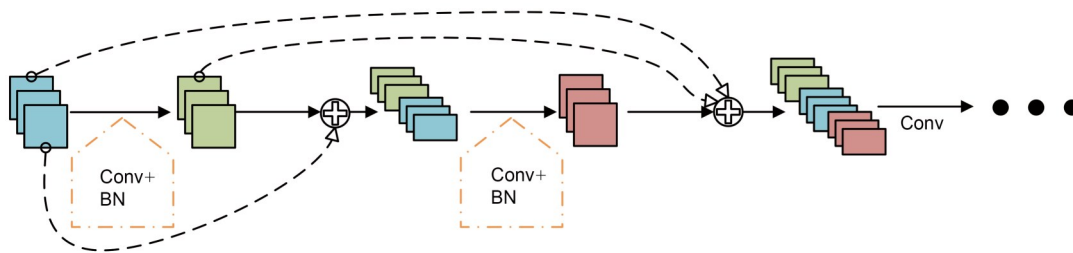
where  $W$  is the weight matrix,  $*$  denotes convolution operation,  $B(\cdot)$  represents batch normalization,  $\delta(\cdot)$  denotes rectified linear unit for activation (Nair and Hinton, 2010), and  $[y(0), y(1), \dots, y(m)]$  represents the concatenations of all outputs of previous layers before the layer of  $m+1$ . The design is illustrated in Fig. 4.

The input size of our network is  $33 \times 33 \times 4$ . Four channels are four modal image patches. We used 64 kernels with each kernel size of  $3 \times 3$  for the initial convolution; then two densely connected blocks are utilized respectively. The proposed architectures of the densely connected block 1 and the densely connected block 2 are presented in Tables 1 and 2, respectively.

In addition, the feature map extracted by the densely connected block 1 has two functions: (1) it is



**Fig. 3** Network structure of our densely connected 2D-CNNs. CNNs: convolutional neural networks; Conv: convolutional layer.



**Fig. 4** Structure of densely connected 2D-CNN feature extractor. BN: batch normalization; CNN: convolutional neural networks; Conv: convolutional layer.

**Table 1** Architecture of the 2D densely connected block 1

Layer	Filter size	Stride	Input	Output
1	3×3	1×1	64×33×33	24×33×33
2	3×3	1×1	88×33×33	24×33×33
3	3×3	1×1	112×33×33	24×33×33
4	3×3	1×1	136×33×33	24×33×33
5	3×3	1×1	160×33×33	24×33×33
6	3×3	1×1	184×33×33	24×33×33
7	3×3	1×1	208×33×33	24×33×33
8	3×3	1×1	232×33×33	24×33×33

**Table 2** Architecture of the 2D densely connected block 2

Layer	Filter size	Stride	Input	Output
1	3×3	1×1	256×33×33	12×33×33
2	3×3	1×1	268×33×33	12×33×33
3	3×3	1×1	280×33×33	12×33×33
4	3×3	1×1	292×33×33	12×33×33
5	3×3	1×1	304×33×33	12×33×33
6	3×3	1×1	316×33×33	12×33×33

directly used as low-level features; (2) it is used as the input of densely connected block 2 to extract high-level features. Low-level features and high-level features

were concatenated. Passing through a series of convolutional and pooling layers, feature maps were then fed into a four-class soft-max classifier (e.g., background, edema region, necrosis/non-enhancing tumor, and enhancing tumor).

### 2.1.3 Fusing segmentation results obtained in axial, coronal, and sagittal views

Three densely connected 2D-CNN segmentation models were trained with training datasets of axial, coronal, and sagittal views. In the test, three segmentation models were used to segment the 3D MRI data of patients. Three views of segmentation results are obtained. Next, a fusion strategy is used to fuse the three segmentation results.  $r_a$ ,  $r_c$ , and  $r_s$  denote the segmentation results of one voxel gotten in axial, coronal, and sagittal views, respectively.  $r$  denotes the segmentation result after fusion, and 0, 1, 2, and 4 denote a voxel labeled as healthy tissue, necrosis/non-enhancing core, edema, and enhancing core, respectively. The fused segmentation result is obtained by following

strategies: (1) if  $r_a=r_c=r_s$ , let  $r=r_a$ ; (2) if any two of  $r_a$ ,  $r_c$ , and  $r_s$  are equal, let  $r$  equal to the equal value (e.g., if  $r_a=r_c$ , let  $r=r_a=r_c$ ); (3) if two or more than two of  $r_a$ ,  $r_c$ , and  $r_s$  are greater than 0, let  $r$  equal to 2, and otherwise, let  $r$  equal to 0.

#### 2.1.4 Post-processing

To further improve the brain tumor segmentation performance, fully connected CRFs are used as our post-processing method. This method can be combined with the relationship among all the pixels in the original image to process the classification results obtained by depth learning, optimize the rough and uncertainty marks in the classification image, correct the fine misclassification area, and obtain a more detailed segmentation boundary.

## 2.2 Experiments

For the implementation of our models, Keras (<https://keras.io>) and Tensorflow (<https://tensorflow.google.cn>) were used as backends. Keras is an open-source deep learning library in high level; it runs on the top of Tensorflow which benefits from a massive parallel architecture such as graphics processing unit (GPU) to optimize the deep learning models.

#### 2.2.1 Materials

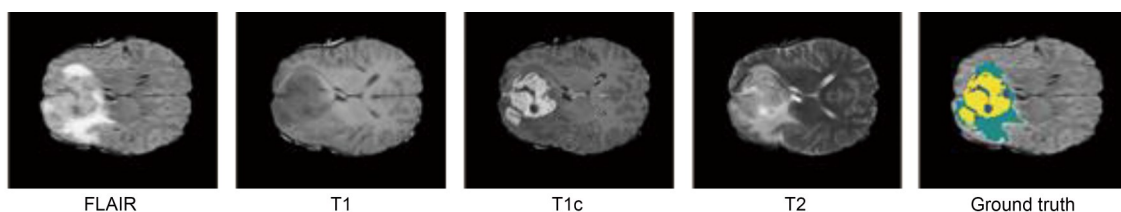
We use datasets provided by the 2018 global multi-modal brain tumor segmentation challenge (BraTS2018) to train and test our segmentation model (Menze et al., 2015; Bakas et al., 2017). The training dataset included 210 HGG cases and 75 LGG cases. Brain images of each patient come with four MRI sequences (i.e., FLAIR, T1, T1c, and T2) and the ground truth labels are manually determined by experts. In Fig. 5, the edema region, necrosis/non-enhancing tumor, and enhancing tumor are indicated in green, yellow, and blue, respectively.

The validation dataset consists of 66 cases with unknown grades. Because brain images of each patient include four MRI sequences (i.e., FLAIR, T1, T1c, and T2) without ground truth labels, we need to upload our segmentation results for online evaluation. To further test the generalization ability of our designed segmentation model, we also used the BraTS2013 training datasets (including 20 HGG cases and 10 LGG cases) for testing (Kistler et al., 2013).

#### 2.2.2 Implementation details

We use 80% of the BraTS2018 training dataset for the training, and the remaining 20% is our test datasets. In the training dataset, four sequences of each patient, including FLAIR, T1, T1c, and T2, were sliced, and then axial, coronal, and sagittal 2D patches were extracted from each slice. Forty patches were extracted from each slice. The number of extracted patches for different classes are equal, effectively alleviating the problem of data imbalance. In our experiments, the training image patches and their labels of the axial, coronal, and sagittal views are sent to the densely connected 2D-CNN model. The back propagation (BP) algorithm and the stochastic batch gradient descent (SGD) algorithm are used to supervise the loss function and to optimize the network parameters. Finally, three optimized densely connected 2D-CNN models were obtained.

When training the axial densely connected 2D-CNNs, axial patches are fed into densely connected 2D-CNNs. The input size is  $33 \times 33 \times 4$ , and patches in these four channels are extracted from pre-processed FLAIR, T1c,  $I_{\text{enhance}}$ , and T1, respectively. The training batch size is set to 64, and the initial learning rate is set to 0.0001. The learning rate is divided by 10 after each 30 epochs. The convolution weight is regularized



**Fig. 5** Four MRI modalities of a patient with glioblastomas disease and its ground truth. The edema region, necrosis/non-enhancing tumor, and enhancing tumor are indicated in green, yellow, and blue, respectively. MRI: magnetic resonance imaging; FLAIR: T2-weighted fluid-attenuated inversion recovery; T1: T1-weighted; T1c: T1-weighted contrast-enhanced; T2: T2-weighted (Note: for interpretation of the references to color in this figure legend, the reader is referred to the web version of this article).

by L1, and the regularization coefficient is 0.0001. First, the cross entropy loss function is used as the loss layer function for the experiment. The training loss and the validation loss changes with the training epoch are shown in Fig. 6a. Fig. 6a shows that with the increase of the epoch, the problem of over-fitting appeared in the network. In order to avoid network over-fitting, a new loss function is proposed. The loss function in 2D-CNNs model is given by

$$\text{Loss} = -(1 - \varepsilon) \sum_i q_i \ln p_i - \varepsilon \sum_{i=1}^n \frac{1}{n} q_i \ln p_i, \quad (4)$$

where  $p_i$  is the predicted distribution,  $q_i$  is the true distribution,  $n$  represents the number of categories, and  $\varepsilon$  is a constant. Our loss function consists of two parts, the first part is the most commonly used cross entropy in classification studies and the second part is the uniform distribution function. For example, if the output is  $[z_1, z_2, z_3, z_4]$  and the target is  $[1, 0, 0, 0]$ , the cross entropy loss function is

$$\text{Loss} = -\ln(e^{z_1}/Z), \quad Z = e^{z_1} + e^{z_2} + e^{z_3} + e^{z_4}. \quad (5)$$

As long as  $z_1$  is already the maximum value of  $[z_1, z_2, z_3, z_4]$ , we can always then “intensify” by

increasing the training parameters, so that  $z_1, z_2, z_3,$  and  $z_4$  increase by a large enough ratio (equivalently, increase the vector  $[z_1, z_2, z_3, z_4]$  modulus length) and  $e^{z_1}/Z$  is close enough to 1 (equivalently, loss is close enough to 0). As long as the model length is blindly increased, the loss can be reduced. This is the source of overconfidence in softmax. In order to prevent over-confidence, one solution is to use a little effort to fit the uniform distribution rather than simply fitting the one-hot distribution. From Eq. (4) in our study, it can be concluded that Eq. (5) is equivalent to

$$\text{Loss} = -(1 - \varepsilon) \ln(e^{z_1}/Z) - \varepsilon \sum_j \frac{1}{4} \ln(e^{z_j}/Z), \quad (6)$$

$$Z = e^{z_1} + e^{z_2} + e^{z_3} + e^{z_4}.$$

In this way, blindly increasing the ratio to make  $e^{z_1}/Z$  close to 1 is no longer the optimal solution, thus alleviating softmax overconfidence and preventing over-fitting.

In BraTS2018, we verify the superiority of the improved loss function. The results of the training process of axial densely connected 2D-CNNs are shown in Fig. 6. Fig. 6a demonstrates the loss and the accuracy of the network training process when the

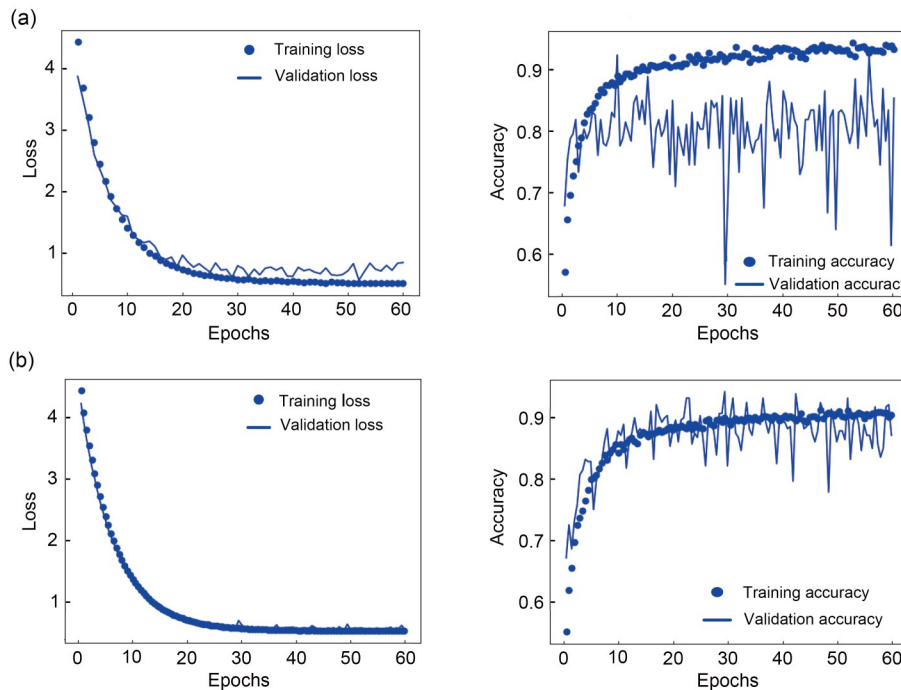


Fig. 6 Results of the training process of axial densely connected 2D-CNNs. (a) Cross-entropy loss function; (b) New loss function. CNNs: convolutional neural networks.

cross-entropy loss function is used, and Fig. 6b shows the loss and the accuracy when using the improved cross entropy loss function. It can be seen from Fig. 6 that the improved loss function increases the segmentation accuracy of the network from 80% to 86%. Throughout the entire training process, there is no over-fitting or under-fitting. The loss is continuously decreasing and the network is continuously optimized. Fig. 6 fully illustrates the necessity and effectiveness of adding a uniform distribution to the basis of the cross-entropy loss function. When the training is completed, the axial densely connected 2D-CNN segmentation model is obtained. Meanwhile, the coronal densely connected 2D-CNNs and the sagittal densely connected 2D-CNNs are gained by using the same training parameters.

### 2.2.3 Evaluation parameters

We evaluated the experimental results based on three metrics: dice score, sensitivity, and specificity. These metrics are calculated as follows:

$$\text{Dice} = 2TP / (2TP + FP + FN), \quad (7)$$

$$\text{Sensitivity} = TP / (TP + FN), \quad (8)$$

$$\text{Specificity} = TN / (TN + FP), \quad (9)$$

where TP, TN, FP, and FN are “true positive,” “true negative,” “false positive,” and “false negative” predictions, respectively. The dice score measures the overlap area of the predicted lesion region and the ground truth. Sensitivity is the measure of tumor pixels that have been correctly classified. Specificity is the measure of normal regions that have been correctly classified. Specifically, the complete region includes the enhancing core, edema, non-enhancing core, and necrosis components; the core region includes the enhancing core, non-enhancing core, and necrosis; the enhancing region only includes the enhancing core.

## 3 Results

In the test phase, the segmentation task is transformed into a classification task, and each image patch is classified to realize segmentation. First, in order to make every pixel of the image predictable, we fill 0 around the original image to expand the image size from  $240 \times 240$  to  $273 \times 273$ . Next, the  $33 \times 33$  size image patches are captured by column in turn, and sent to

the segmentation model to obtain the classification results of the central pixel pixels. Finally, the segmentation result of the slice was obtained by combining the classification results of each pixel by columns.

### 3.1 Primary experiments tested on the BraTS2018 training datasets

We use 20% (57 cases) of the data in the BraTS2018 for our preliminary test set. Table 3 shows the evaluation results of axial, coronal, and sagittal densely connected 2D-CNNs, fusion processing, and fusion with post-processing. Table 3 also lists the segmentation results of gliomas with and without pre-processing. The segmentation results of axial, coronal, and sagittal views obtained with pre-processing are superior to those obtained without pre-processing. The dice coefficient, sensitivity, and specificity of the whole tumor area were increased by 5.3%, 5.1%, and 3.4%, respectively. The evaluation metrics of tumor core area and tumor enhancement area are also slightly improved, indicating the effectiveness of the pre-processing step for the edema area of brain glioma. The scores in Table 3 also indicate that the fusion of segmentation results obtained from different views markedly improves segmentation accuracy. Without pre-processing, the fusion process will increase the dice of the complete region, core region, and enhancing region by 2.7%, 6.7%, and 1.3%, respectively; if pre-processing is performed, the fusion process will increase the dice by 8.0%, 9.7%, and 5.2%, respectively. The segmentation model of a single view is prone to over-segmentation and under-segmentation due to the lack of spatial location information of the glioma. The segmentation results of multiple views are merged through the majority voting strategy to improve the accuracy of glioma segmentation, especially of the entire tumor. For some obvious false positives that appear in the segmentation results of a certain view, such as small nodules, such false positives can be directly reduced through majority voting strategy fusion processing. Furthermore, post-processing slightly improved the segmentation results. Fig. 7 shows that the segmentation results of fusing three views can remove some obvious false positives, which appear in one of the three results and do not appear in the other two results. From the two sample segmentations, we can clearly see that when the glioma slice has only one tumor type, it can also be accurately segmented.



### 3.2 Segmentation performance on BraTS2018 validation datasets

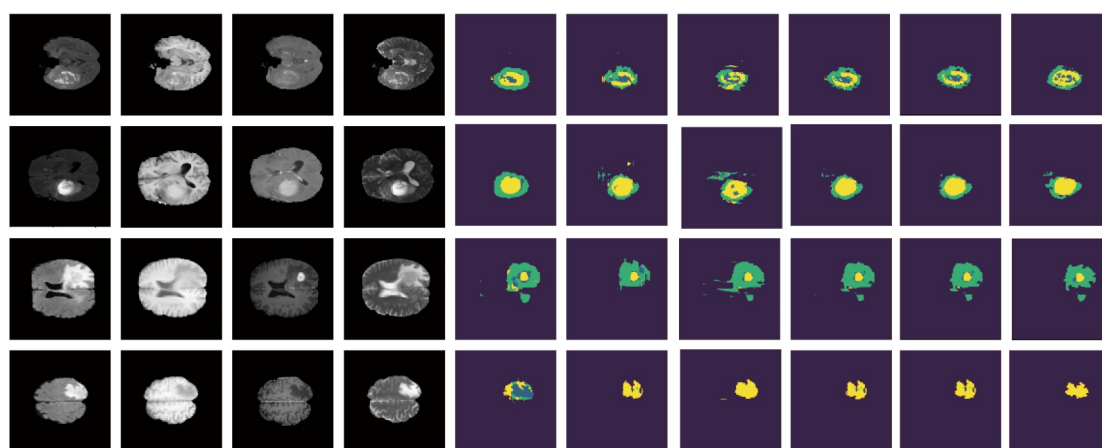
The evaluation scores of our method for the BraTS2018 validation dataset are provided by BraTS2018 organizer (<https://ipp.cbica.upenn.edu>). The evaluation results of our method are listed in Table 4, along with the segmentation results of other methods participating in the BraTS2018 challenge. McKinley et al. (2018) introduced a new family of classifiers based on DeepSCAN architecture, in which densely connected

blocks of dilated convolutions are embedded in a shallow U-Net-style structure of down/up sampling and skip connections. Zhou et al. (2018) designed multiple deep architectures of varied structures for learning contextual and attentive information, and then integrated the predictions of these models to obtain more robust segmentation results. The proposed method has great advantages in sensitivity, with sensitivity scores of 0.8441, 0.9511, and 0.9228 for the enhancing tumor, whole tumor, and tumor core, respectively (Table 4).

**Table 3 Average performance of our system on the 57 glioma cases examined**

Method	Dice			Sensitivity			Specificity		
	Comp.	Core	Enh.	Comp.	Core	Enh.	Comp.	Core	Enh.
Without pre-processing									
Axial	0.7630	0.6931	0.7879	0.8443	0.8391	0.8710	0.9276	0.9569	0.9745
Coronal	0.7521	0.7248	0.7649	0.8376	0.8535	0.8539	0.9147	0.9567	0.9689
Sagittal	0.7434	0.7570	0.7850	0.8567	0.8620	0.8628	0.9248	0.9429	0.9730
Mean	0.7528	0.7249	0.7792	0.8462	0.8516	0.8625	0.9223	0.9521	0.9721
+Fusion	0.7804	0.7923	0.7932	0.9028	0.8798	0.8932	0.9429	0.9632	0.9876
+Post	0.7806	0.7930	0.7931	0.9031	0.8812	0.8932	0.9432	0.9634	0.9845
With pre-processing									
Axial	0.8103	0.7047	0.8142	0.9172	0.8778	0.9084	0.9585	0.9937	0.9946
Coronal	0.8055	0.7566	0.7888	0.8740	0.8948	0.8813	0.9679	0.9820	0.9948
Sagittal	0.8112	0.7950	0.8070	0.9024	0.9004	0.8919	0.9651	0.9651	0.9964
Mean	0.8090	0.7521	0.8033	0.8978	0.8910	0.8938	0.9638	0.9802	0.9952
+Fusion	0.8337	0.8227	0.8319	0.9532	0.9261	0.9128	0.9783	0.9972	0.9974
+Post	0.8393	0.8231	0.8327	0.9532	0.9261	0.9129	0.9784	0.9972	0.9974

Comp.: complete; Enh.: enhancing.



**Fig. 7** Two segmentation examples demonstrate the effectiveness of integrating multiple densely connected 2D-CNNs. The first and second rows show segmentation results of the 50th and 75th slices of the axial view of patient Brats18\_TCIA06\_211\_1, respectively. The third and fourth rows show the segmentation results of the 95th and 125th slices of the axial view of patient Brats18\_2013\_18\_1, respectively. From left to right: FLAIR, T1, T1c, T2, segmentation results of axial densely connected 2D-CNNs, segmentation results of coronal 2D densely connected CNNs, sagittal densely connected 2D-CNNs, fusion, fusion+post-process, and the ground truth labels. In the segmentation results, the edema region, necrosis/non-enhancing tumor, and enhancing tumor are indicated in green, yellow, and blue, respectively. FLAIR: T2-weighted fluid-attenuated inversion recovery; T1: T1-weighted; T1c: T1-weighted contrast-enhanced; T2: T2-weighted; CNNs: convolutional neural networks (Note: for interpretation of the references to color in this figure legend, the reader is referred to the web version of this article).

Baid et al. (2020) have designed a novel 3D U-Net architecture that segments various radiologically identifiable subregions such as edema, enhancing tumor, and necrosis; they proposed a weighted patch extraction scheme from the tumor border regions to address the problem of class imbalance between tumor and non-tumorous patches. Fig. 8 shows the dice scores of complete regions obtained with our methods on the BraTS2018 validation dataset, which reached an average of 0.81. Fig. 9 shows the mean, median, and standard deviation of the segmentation results of the BraTS2018 validation dataset. The median and mean values of each evaluation metrics were highly consistent for enhancing tumor, whole tumor, and tumor core, and the standard deviation was small, indicating that the proposed segmentation method has high stability (Fig. 9). Fig. 10 shows the segmentation results of seven patients.

### 3.3 Segmentation performance of BraTS2013 datasets

The BraTS2013 training dataset including 20 cases of HGG and 10 cases of LGG was used for testing. For evaluating the effectiveness of the

proposed model, a comparative analysis is presented in Table 5. Hussain et al. (2017) developed a brain tumor segmentation method based on the cascaded deep CNNs. Li et al. (2016) proposed a probabilistic model that combines sparse representation and Markov random field to classify tumor pixels. Table 5 demonstrates the superior performance of our proposed model in terms of dice score and sensitivity.

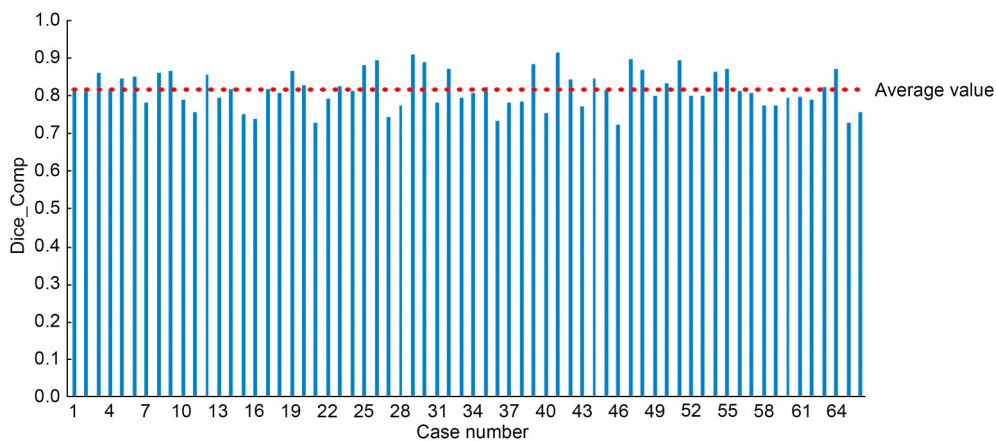
## 4 Discussion and conclusions

Brain tumor segmentation plays an important role in diagnostic procedures in brain glioma, radiation therapy, clinical surgical planning, and assessment of benign and malignant conditions. Not only is clinical diagnosis facilitated, but the survival chances of a brain glioma patient also greatly increase when segmentation accuracy is high. In this study, 3D brain images were segmented by integrating the segmentation results of multiple densely connected 2D-CNNs, which were trained to segment brain images from axial, coronal, and sagittal views. Segmentation begins at a pre-processing stage consisting of bias field correction,

**Table 4 Average sensitivity scores of the BraTS2018 validation dataset (66 cases)**

Study	Dice			Sensitivity			Specificity		
	Comp.	Core	Enh.	Comp.	Core	Enh.	Comp.	Core	Enh.
This study	0.8165	0.8488	0.8249	0.9511	0.9228	0.8441	0.9961	0.9823	0.9948
McKinley et al. (2018)	0.9007	0.8473	0.7924	0.9107	0.8359	0.8290	0.9937	0.9912	0.9980
Zhou et al. (2018)	0.9094	0.8650	0.8135	0.9142	0.8682	0.8134	0.9941	0.9968	0.9983
Baid et al. (2020)	0.8780	0.8267	0.7480						

BraTS2018: 2018 global multi-modal brain tumor segmentation challenge; Comp.: complete; Enh.: enhancing.



**Fig. 8 Bar plots for the dice score of complete tumor on the validation dataset. Dice\_Comp means the dice score of complete tumor.**

edema enhancement, and patch normalization. Next, the same numbers of image patches for each class were used as training image patches to avoid data imbalance problem. Convolutional layers with small kernels 3×3 in size to allow deeper architectures build the densely connected CNNs. Three segmentation models were trained using 2D image patches to obtain axial, coronal, and sagittal views. The three models are integrated to segment brain glioma using fusion strategy.

Based on the characteristics of brain glioma itself, a new pre-processing method was proposed to acquire clear brain glioma edema, FLAIR, and T2 modal

images. Fig. 2 indicates that the proposed pre-processing method clarifies the edema boundary. The segmentation results for BraTS2013 shown in Table 5 also demonstrate that our method segments the complete tumor very well. Compared with other methods, our approach is more accurate for segmentation of the edema tumor.

To overcome the limitations of existing segmentation models based on CNNs, 2D densely connected blocks are added to our network structure. The effect of densely connected blocks is to reuse the features of all the previous layers, each of which accepts the features of all the layers in front of it as input. Because the image features are extracted in stages, both

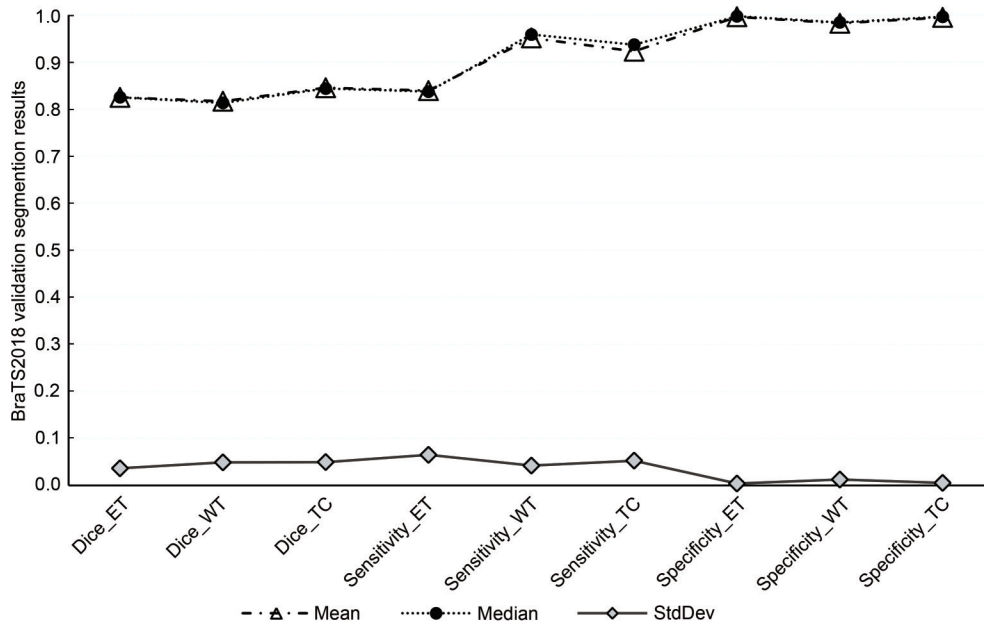


Fig. 9 Mean, median, and standard deviation for the segmentation results of the BraTS2018 validation dataset. BraTS2018: 2018 global multi-modal brain tumor segmentation challenge; StdDev: standard deviation; ET: enhancing tumor; WT: whole tumor; TC: tumor core.

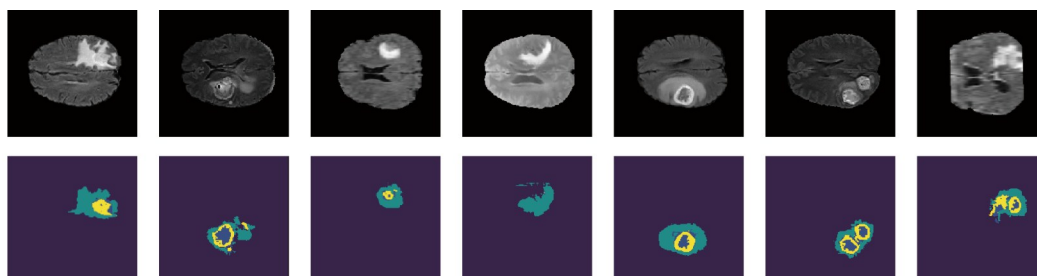


Fig. 10 Examples of segmentation results for the BraTS2018 validation dataset. The first row shows the 90th slice of FLAIR. The second row shows the segmentation results: green, edema; yellow, non-enhancing/necrosis; blue, enhancing tumor. From left to right: Brats18\_CBICA\_AAM\_1, Brats18\_CBICA\_BHN\_1, Brats18\_TCIA03\_604\_1, Brats18\_MDA\_922\_1, Brats18\_WashU\_S041\_1, Brats18\_CBICA\_AP-M\_1, and Brats18\_TCIA07\_600\_1. BraTS2018: 2018 global multi-modal brain tumor segmentation challenge; FLAIR: T2-weighted fluid-attenuated inversion recovery (Note: for interpretation of the references to color in this figure legend, the reader is referred to the web version of this article).

**Table 5 Average segmentation results on BrsTS2013 training dataset**

Method	Dice			Specificity			Sensitivity		
	Comp.	Core	Enh.	Comp.	Core	Enh.	Comp.	Core	Enh.
Proposed method	0.81	0.78	0.80	0.84	0.71	0.72	0.90	0.76	0.79
Hussain et al. (2017)	0.80	0.67	0.85				0.82	0.63	0.83
Li et al. (2016)	0.78	0.52	0.52				0.85	0.54	0.58

BraTS2018: 2018 global multi-modal brain tumor segmentation challenge; Comp.: complete; Enh.: enhancing.

low-level features and high-level features can be used. Therefore, the network does not ignore information as the number of network layers increases during the training process. Moreover, we defined a new loss function which facilitates model training. The training loss values shown in Fig. 6 indicate that our model has good learning ability and does not result in over-fitting.

We evaluated the proposed method in BraTS2013 and BraTS2018 glioma datasets. The experimental results shown in Table 3 indicate that the fusion result improves the segmentation accuracy by removing some obvious false positive results. The HGG tissue was obvious, but the boundary between edema tumor and core tumor in LGG was fuzzy. Fig. 8 shows that the segmentation results of each patient are slightly different, with a standard deviation is 0.079. However, the average dice score of 0.82 is equal to the dice scores of 57 patients in the training set, indicating that the algorithm developed in this paper has high stability. Tables 3 and 5 also indicate that the segmentation results of BraTS2018 were better than those of BraTS2013. These notable differences suggest a reason why the generalization ability of our model needs to be further improved. Accordingly, the next important goal of our research is to explore an automatic segmentation model that can adapt to the large differences in data by adjusting the input size of the network and changing the network structure to obtain a more stable segmentation model. Currently, adversarial networks are outperforming state of the art methods for semantic segmentation in several computer vision tasks. This approach might be further investigated as a means of improving segmentation in medical images.

In this paper, we have presented a fully automatic brain tumor segmentation method based on three densely connected 2D-CNNs. We considered different training schemes with variable loss functions, data pre-processing methods, and fusion strategy. Our enhanced brain glioma segmentation method, analyzed three testing datasets, and achieved high accuracies of

0.86, 0.82, and 0.81 dice scores, respectively, over the entire tumor region. However, despite application of multiple densely connected 2D-CNNs, the 3D information of MRI data were still not obtained. In addition, the post-processing method is relatively simple. Our ongoing work aims to change the network structure and post-processing methods to make full use of the 3D information of MRI in brain glioma to further improve tumor segmentation performance.

### Acknowledgments

This work was supported by the National Natural Science Foundation of China (No. 81830052), the Shanghai Natural Science Foundation of China (No. 20ZR1438300), and the Shanghai Science and Technology Support Project (No. 18441900500), China.

### Author contributions

Shengdong NIE contributed to the conception of the study. Xiaobing ZHANG performed the experiment and wrote the manuscript. Yin HU and Wen CHEN contributed significantly to analysis and manuscript preparation. Gang HUANG helped perform the analysis with constructive discussions. All authors have read and approved the final manuscript and, therefore, have full access to all the data in the study and take responsibility for the integrity and security of the data.

### Compliance with ethics guidelines

Xiaobing ZHANG, Yin HU, Wen CHEN, Gang HUANG, and Shengdong NIE declare that they have no conflict of interest.

All procedures followed were in accordance with the ethical standards of the responsible committee on human experimentation (institutional and national) and with the Helsinki Declaration of 1975, as revised in 2008 (5). Informed consent was obtained from all patients for being included in the study.

### References

- Baid U, Talbar S, Rane S, et al., 2020. A novel approach for fully automatic intra-tumor segmentation with 3D U-Net architecture for gliomas. *Front Comput Neurosci*, 14:10. <https://doi.org/10.3389/fncom.2020.00010>

- Bakas S, Akbari H, Sotiras A, et al., 2017. Advancing The Cancer Genome Atlas glioma MRI collections with expert segmentation labels and radiomic features. *Sci Data*, 4:170117.  
<https://doi.org/10.1038/sdata.2017.117>
- Chen LL, Wu Y, DSouza AM, et al., 2018. MRI tumor segmentation with densely connected 3D CNN. Proceedings Volume 10574, Medical Imaging 2018: Image Processing. SPIE Medical Imaging, 2018, Houston, Texas, USA, 105741F.  
<https://doi.org/10.1117/12.2293394>
- Dvořák P, Menze B, 2016. Local structure prediction with convolutional neural networks for multimodal brain tumor segmentation. *In: Menze B, Langs G, Montillo A, et al. (Eds.), Medical Computer Vision: Algorithms for Big Data*. Springer, Cham, p.59-71.  
[https://doi.org/10.1007/978-3-319-42016-5\\_6](https://doi.org/10.1007/978-3-319-42016-5_6)
- Goetz M, Weber C, Binczyk F, et al., 2015. DALSA: domain adaptation for supervised learning from sparsely annotated MR images. *IEEE Trans Med Imaging*, 35(1):184-196.  
<https://doi.org/10.1109/tmi.2015.2463078>
- Guo D, Wang L, Song T, et al., 2019. Cascaded global context convolutional neural network for brain tumor segmentation. *In: Crimi A, Bakas S (Eds.), Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries. BrainLes 2019. Lecture Notes in Computer Science*, Vol. 11992. Springer, Cham, p.315-326.  
[https://doi.org/10.1007/978-3-030-46640-4\\_30](https://doi.org/10.1007/978-3-030-46640-4_30)
- Havaei M, Davy A, Warde-Farley D, et al., 2017. Brain tumor segmentation with Deep Neural Networks. *Med Image Anal*, 35:18-31.  
<https://doi.org/10.1016/j.media.2016.05.004>
- He KM, Zhang XY, Ren SQ, et al., 2016. Deep residual learning for image recognition. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, p.770-778.  
<https://doi.org/10.1109/CVPR.2016.90>
- Huang G, Liu Z, van der Maaten L, et al., 2017. Densely connected convolutional networks. 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, p.2261-2269.  
<https://doi.org/10.1109/CVPR.2017.243>
- Hussain S, Anwar SM, Majid M, 2017. Brain tumor segmentation using cascaded deep convolutional neural network. 2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Jeju, Republic of Korea, p.1998-2001.  
<https://doi.org/10.1109/EMBC.2017.8037243>
- Islam R, Imran S, Ashikuzzaman M, et al., 2020. Detection and classification of brain tumor based on multilevel segmentation with convolutional neural network. *J Biomed Sci Eng*, 13(4):45-53.  
<https://doi.org/10.4236/jbise.2020.134004>
- Kamnitsas K, Ledig C, Newcombe VFJ, et al., 2017. Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation. *Med Image Anal*, 36:61-78.  
<https://doi.org/10.1016/j.media.2016.10.004>
- Kistler M, Bonaretti S, Pfahrer M, et al., 2013. The virtual skeleton database: an open access repository for biomedical research and collaboration. *J Med Internet Res*, 15(11):e245.  
<https://doi.org/10.2196/jmir.2930>
- Krishna N, Khalander MR, Shetty N, et al., 2019. Segmentation and detection of glioma using deep learning. *In: Chiplunkar N, Fukao T (Eds.), Advances in Artificial Intelligence and Data Engineering. Advances in Intelligent Systems and Computing*, Vol. 1133. Springer, Singapore, p.109-120.  
[https://doi.org/10.1007/978-981-15-3514-7\\_10](https://doi.org/10.1007/978-981-15-3514-7_10)
- Li QN, Gao ZF, Wang QY, et al., 2018. Glioma segmentation with a unified algorithm in multimodal MRI images. *IEEE Access*, 6:9543-9553.  
<https://doi.org/10.1109/ACCESS.2018.2807698>
- Li YH, Jia FC, Qin J, 2016. Brain tumor segmentation from multimodal magnetic resonance images via sparse representation. *Artif Intell Med*, 73:1-13.  
<https://doi.org/10.1016/j.artmed.2016.08.004>
- McKinley R, Meier R, Wiest R, 2018. Ensembles of densely-connected CNNs with label-uncertainty for brain tumor segmentation. *In: Crimi A, Bakas S, Kuijf H, et al. (Eds.), Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries. BrainLes 2018. Lecture Notes in Computer Science*, Vol. 11384. Springer, Cham, p.456-465.  
[https://doi.org/10.1007/978-3-030-11726-9\\_40](https://doi.org/10.1007/978-3-030-11726-9_40)
- Mengqiao W, Jie Y, Yi C, et al., 2017. The multimodal brain tumor image segmentation based on convolutional neural networks. 2017 2nd IEEE International Conference on Computational Intelligence and Applications (ICCIA). Springer, Cham, p.336-339.  
<https://doi.org/10.1109/CIAPP.2017.8167234>
- Menze BH, Jakab A, Bauer S, et al., 2015. The multimodal brain tumor image segmentation benchmark (BRATS). *IEEE Trans Med Imaging*, 34(10):1993-2024.  
<https://doi.org/10.1109/TMI.2014.2377694>
- Mohan G, Subashini MM, 2018. MRI based medical image analysis: survey on brain tumor grade classification. *Biomed Signal Proc Control*, 39:139-161.  
<https://doi.org/10.1016/j.bspc.2017.07.007>
- Nair V, Hinton GE, 2010. Rectified linear units improve restricted Boltzmann machines. Proceedings of the 27th International Conference on Machine Learning (ICML-10), Haifa, Israel, p.807-814.
- Pan MY, Shi YH, Song ZJ, 2020. Segmentation of gliomas based on a double-pathway residual convolution neural network using multi-modality information. *J Med Imaging Health Informatics*, 10(11):2784-2794.  
<https://doi.org/10.1166/jmihi.2020.3216>
- Simonyan K, Zisserman A, 2015. Very deep convolutional networks for large-scale image recognition. Proceedings of the 3rd International Conference on Learning Representations. Springer, Cham, p.6.  
<https://arxiv.org/abs/1409.1556>
- Udapa P, Vishwakarma S, 2016. A survey of MRI segmentation techniques for brain tumor studies. *Bonfring Int J Adv Image Proc*, 6(3):22-27.  
<https://doi.org/10.9756/bijaiip.10467>

- Urban G, Bendszus M, Hamprecht FA, et al., 2014. Multimodal brain tumor segmentation using deep convolutional neural networks. Proceedings MICCAI-BRATS, Springer, Cham, p.31-35.  
<https://doi.org/10.1117/12.2557599>
- Zhao LY, Jia KB, 2015. Deep feature learning with discrimination mechanism for brain tumor segmentation and diagnosis. 2015 International Conference on Intelligent Information Hiding and Multimedia Signal Processing (IIH-MSP). Adelaide, SA, Australia, p.306-309.  
<https://doi.org/10.1109/IIH-MSP.2015.41>
- Zhao XM, Wu YH, Song GD, et al., 2018. A deep learning model integrating FCNNs and CRFs for brain tumor segmentation. *Med Image Anal*, 43:98-111.  
<https://doi.org/10.1016/j.media.2017.10.002>
- Zhou CH, Chen SC, Ding CX, et al., 2018. Learning contextual and attentive information for brain tumor segmentation. In: Crimi A, Bakas S, Kuijf H, et al. (Eds.), Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries. BrainLes 2018. Lecture Notes in Computer Science, Vol. 11384, Springer, Cham, p.497-507.  
[https://doi.org/10.1007/978-3-030-11726-9\\_44](https://doi.org/10.1007/978-3-030-11726-9_44)
- Zhou ZX, He ZS, Shi MF, et al., 2020. 3D dense connectivity network with atrous convolutional feature pyramid for brain tumor segmentation in magnetic resonance imaging of human heads. *Comput Biol Med*, 121:103766.  
<https://doi.org/10.1016/j.compbiomed.2020.103766>
- Zhuge Y, Krauze AV, Ning H, et al., 2017. Brain tumor segmentation using holistically nested neural networks in MRI images. *Med Phys*, 44(10):5234-5243.  
<https://doi.org/10.1002/mp.12481>
- Zikic D, Ioannou Y, Brown M, et al., 2014. Segmentation of brain tumor tissues with convolutional neural networks. Proceedings MICCAI Workshop on Multimodal Brain Tumor Segmentation Challenge. Springer, Cham, p.36-39.