Segmentation of Brain Gliomas Based on a Multi-modal Multi-scale Double-pathway 3D Residual CNN

Mingyuan Pan¹,², Yonghong Shi¹,²,*, Zhijian Song¹,²,*

¹Digital Medical Research Center, School of Basic Medical Science, Fudan University, Shanghai, China
²Shanghai Key Laboratory of Medical Imaging Computing and Computer Assisted Intervention, Shanghai, China

*Correspondence: Zhijian Song, zjsong@fudan.edu.cn; Yonghong Shi, yonghong.shi@fudan.edu.cn.

Keywords: Brain glioma, Multimodality, Brain tumor segmentation, CNN, Brain tumor

Abstract

Background: The automated segmentation of brain gliomas regions in magnetic resonance (MR) images plays an important role in the early diagnosis, intraoperative navigation, radiotherapy planning and prognosis of brain tumors. It is very challenging to segment gliomas and intratumoral structures since the location, size, shape, edema range and boundary of gliomas are heterogeneous, and multimodal brain gliomas images (such as T1, T2, fluid-attenuated inversion recovery (FLAIR), and T1c images) are collected from multiple radiation centers.

Methods: This paper presents a multimodal, multi-scale, double-pathway, 3D residual convolution neural network (CNN) for automatic gliomas segmentation. First, a robust gray-level normalization method is proposed to solve the multicenter problem, such as very different intensity ranges due to different imaging protocols. Second, a multi-scale, double-pathway network based on DeepMedic toolkit is trained with different combinations of multimodal MR images for gliomas segmentation. Finally, a fully connected conditional random field (CRF) is used as a post-processing strategy to optimize the segmentation results for addressing the isolated segmentations and holes.

Results: Experiments on the Multimodal Brain Tumor Segmentation (BraTS) 2017 and 2019 challenge data show that our methods achieve a good performance in delineating the whole tumor with a Dice coefficient, a sensitivity and a positive predictive value (PPV) of 0.88, 0.89 and 0.88, respectively. Regarding the segmentation of the tumor core and the enhancing area, the sensitivity reached 0.80.

Conclusions: Experiments show that our method can accurately segment gliomas and intratumoral structures from multimodal MR images, and it is of great significance to clinical neurosurgery.
1 Background

Brain tumors, whose annual incidence is 4 per 10000 people, are one of the most lethal cancers in the world. Gliomas account for 40-50% of all brain tumors [1]. Identifying gliomas structures based on precise segmentation of brain magnetic resonance (MR) images not only assists diagnosis but also plays an important role in intraoperative navigation of tumor location, regional planning of targeted radiotherapy, and follow-up observation of patients after surgery [2]. Therefore, developing a reliable, accurate and fast gliomas segmentation method has always been a research hotspot in the field of medical image processing. However, the locations, sizes, shapes, edema ranges and boundaries of gliomas in brain images are heterogeneous. It is very difficult to segment gliomas and the intratumoral structures [3].

Manual and semiautomatic segmentation of gliomas takes several hours for an experienced radiologist to delineate the tumor boundary. The results of manual segmentation largely depend on the subjective judgment of the observer or clinician. It was reported that gliomas segmented by the same or different doctors may differ by up to 20%+15% or 28%+12%, respectively [4].

Based on prior knowledge, the traditional automatic segmentation algorithm preprocesses images, calculates and outputs image features (such as texture primitives, wavelet transforms) to random forest, support vector machine, self-organizing map or K-means classifiers for judgment and classification, and achieve the automatic segmentation of gliomas [5, 6]. These traditional classifiers can process only pre-extracted image features. In recent years, deep learning algorithms based on human visual systems, such as convolutional neural networks (CNNs), can independently acquire certain feature information from images and achieve feature classification and target segmentation [7]. This method greatly saves the time consumed by traditional algorithms in feature extraction.

To achieve brain gliomas segmentation in multimodal brain MR images, some studies divide multimodal 3D MR images into groups of 2D cross-sectional data along the vertical axis and input them into a 2D CNN for training [8]. In essence, this method is still based on a 2D CNN. The complexity of the network structure is relatively low, and the amount of computation is relatively small. The disadvantage of this algorithm is that it ignores the relative relationship of voxels on the vertical axis and loses the interlayer relationship of the image information. Some studies directly build a 3D CNN network to process the 3D structure feature of voxels [9]. This method greatly increases the computational complexity and introduces higher requirements for the hardware performance. To address the heterogeneity of tumors, a two-channel cascaded CNN with multi-scale sample input is an effective method [10, 11] both in 2D or 3D CNN since such architecture combined the detailed features of tumor and its larger context information well. Except of altering the kernel size in each pathway, DeepMedic [12] was ranked top in Multimodal Brain Tumor Segmentation (BraTS) 2016 challenge with an architecture of deep layers and two pathways each dealing with the multi-scale information. Besides, the traditional automatic segmentation methods, for example, a conditional random field (CRF), was proved efficient to be a post-processing step of a deep learning algorithm to improve the network performance, as it can take into account voxels and their surrounding domain information [13]. And with the basic idea of multi-pathway architecture, many studies [14] and even DeepMedic itself has been improved to a three- or more- pathway network to deal with multi-resolution patches. At the same time, there are also other multi-pathway architecture [15, 16] focusing on the information of multimodalities, with a densely-connection between convolutional layers in the one pathway and between the pathways, they had proved the combination of multi-modality MRI in multi-pathway worked well in infant brain segmentation and ischemic stroke lesion segmentation. Recently, Aigün et
al. [17] had explored that a late fusion strategy did perform better than early one as it avoids the early merge of particular features in each modalities.

Inspired by the above mentioned methods, we trained the multimodal multi-scale double-pathway 11-layer residual CNN with a late fusion strategy for automatic segmentation of gliomas in MR images. Since MR imaging scans may be altered by the bias field distortion and suffer from the multicenter issue, it is necessary to vary the intensity of the same tissues consistently across the images. In the preprocessing stage, a robust strategy of denoising and intensity normalization of multimodal MR images is proposed to make the images more comparable. Then, the datasets are randomly divided into a training set, a validation set and a test set. After data enhancement, multiscale training samples of different modalities are sampled and input into two pathways according to our modality combination to enhance the feature information. The soft segmentation results by the CNN may show some isolated, misclassified voxels or small clusters, sometimes in physiological and anatomically unlikely locations. Finally, a post-processing procedure including a fully connected 3D CRF is employed to optimize the soft segmentation probability maps and obtain the final segmentation results.

This paper is organized as follows: the experimental results are summarized in section 2, and the discussion and conclusion are given in sections 3 and 4, the details of our methods are presented in section 5 respectively.

2 Results

2.1 Dataset

The datasets are from BraTS2017[18] and BraTS 2019[19], respectively. BraTS2017 consists of 75 cases with LGGs and 210 cases with HGGs. BraTS 2019 consists of 76 cases with LGGs and 259 cases with HGGs. Each case was scanned with four modalities: T1, T2, FLAIR and T1c. All the images were preprocessed by registration, reinterpolation and skull removal. The necrotic areas, edema areas and enhancement areas of the gliomas were labeled manually.

For BraTS2017, to ensure the accuracy of the experimental results, 10 groups of data were run in each experiment, and all the LGG and HGG data were randomly divided into a training group with 205 cases, a validation group with 40 cases and a test group with 40 cases. The number of LGG cases in the training set was doubled to ensure the balance of LGGs and HGGs in the training samples. Here, the doubled number of LGG cases cannot lead to the imbalance of sample since the study is about classifying tissues into necrotic, edema, and enhancement areas of the gliomas. For BraTS2019, we adopt the same training strategy, with 10 groups of data tested, and 40 cases both in test and validation set.

The study is based on the DeepMedic open source toolkit Keras in the Theano deep learning framework [12]. The experimental environment is a CUDA 7.5 parallel computing structure and cuDNN 7.5 GPU acceleration library. The GPU is an NVIDIA GeForce GTX 1080.

2.2 Network configuration

The total number epochs of the training process was set to 35. Each epoch contains 20 iterations, and during each iteration, 10 sets of sample data are processed in parallel. Each training session consisted of 1000 samples, which consisted of 10 samples from healthy tissue and 10 samples from tumor regions from 50 patients randomly selected from the training set. After each training epoch, 5000 samples were randomly selected from the ROIs of all 40 patients in the validation set with a ratio of healthy tissue to
tumor area of 3:7 to evaluate the performance of the CNN and determine the training condition. After training, 40 case images in the test set were input into the network for segmentation.

The dropout rate of the full connection layer was set to \{0, 0.5, 0.5\}. The activation function of each layer was the PReLU function, and the classifier was the softmax function. The batch size was 10, while the standardized number of samples of each batch was 60 [20]. The activation function values were normalized to prevent output explosion. The initial learning rate is $10^{-4}$ with a decline by half. After three epoch validations, the accuracy remains stable, and the learning rate is halved to improve the learning accuracy.

### 2.3 Evaluation metrics

According to the BraTS evaluation index, we need to evaluate the complete tumor, core region and enhancing region segmented manually ($V_{GT}$) and algorithmically ($V_M$) through the Dice coefficient, positive predictive value (PPV), and sensitivity, respectively. These measures are calculated as follows:

\[
DICE = \frac{2(V_M \cap V_{GT})}{V_M + V_{GT}}, \quad PPV = \frac{V_M \cap V_{GT}}{V_M}, \quad Sensitivity = \frac{V_M \cap V_{GT}}{V_{GT}}.
\]

In particular, the complete tumor includes a necrotic area, edema area and enhancing core; the core region includes a necrotic area and an enhancing core; and the enhancing region includes only the enhancing core.

The Hausdorff distance is also taken into consideration as the maximum distance of a set to the nearest point in another set, in other words, how close the segmentation and the expected output are. Given two point sets $A = \{a_1, a_2, ..., a_n\}$ and $B = \{b_1, b_2, ..., b_m\}$ in Euclidean space, the Hausdorff distance between $A$ and $B$ is defined as

\[
HD(A, B) = \max \left( \max_{a_i \in A} \left( \min_{b_j \in B} d(a_i - b_j) \right), \max_{b_j \in B} \left( \min_{a_i \in A} d(a_i - b_j) \right) \right),
\]

and a 95% Hausdorff distance was chosen to avoid the influence of outliers [21].

### 2.4 Evaluations

#### 2.4.1 Evaluation of the intensity normalization methods

Results in Table 1 shows that our robust intensity normalization performs better than the simple normalized filter in 3D slicer module in all the metrics. As the latter one normalized the image globally, which account for a lot of zero point, and its intensity distribution is still the same as the original images without a histogram matching.
Table 1. Performance comparison of different intensity normalization.

| Normalization | DICE Tumor | Core | Enha. | PPV Tumor | Core | Enha. | Sensitivity Tumor | Core | Enha. | Hausdorff distance Tumor | Core | Enha. |
|---------------|------------|------|-------|-----------|-------|-------|-------------------|-------|-------|--------------------------|-------|-------|
| Robust Filters| 0.87±0.11  | 0.77±0.19 | 0.77±0.16 | 0.86±0.15 | 0.81±0.20 | 0.79±0.17 | 0.91±0.08 | 0.78±0.19 | 0.80±0.20 | 18.59±19.42 | 19.97±21.30 | 14.11±22.22 |
| Simple Filters | 0.80±0.13  | 0.71±0.22 | 0.64±0.25 | 0.85±0.14  | 0.77±0.19 | 0.77±0.23 | 0.79±0.16 | 0.72±0.25 | 0.65±0.30 | 27.56±23.69 | 37.82±27.59 | 30.44±32.83 |

2.4.2 Evaluation of the patch size

As shown in Table 2, based on the constructed network structure, three training sample sizes, 19³, 25³ and 31³, are considered. When shrinking the patch size from 25³ to 19³, for the whole tumor, the Dice coefficient declines from 0.87 to 0.83, the PPV declines from 0.86 to 0.82, and the Hausdorff distance increases. A similar result is obtained when enlarging the patch size from 25³ to 31³. However, when considering the performance of the enhancing core and neurotic core, the larger size of 31³ is much poorer, as its Dice coefficient declines to 0.65 and 0.40 and PPV declines to 0.76 and 0.52. The smaller patch size of 19³ tends to have good accuracy in segmenting the enhancing core with a PPV of 0.84.

The segmentation results are shown in Fig. 1. Compared with the results of a patch size of 25³, the blue neurotic core and the whole tumor were poorly segmented regardless of whether the patch was reduced to a size of 19³ or increased to a size of 31³. Thus, the patch size was set as 25³ in the study.

Table 2. Performance comparison based on the training patches of different sizes.

| Patch size | DICE Tumor | Core | Enha. | PPV Tumor | Core | Enha. | Sensitivity Tumor | Core | Enha. | Hausdorff distance Tumor | Core | Enha. |
|------------|------------|------|-------|-----------|-------|-------|-------------------|-------|-------|--------------------------|-------|-------|
| 19³        | 0.83±0.13  | 0.76±0.23 | 0.77±0.24 | 0.82±0.14  | 0.80±0.22 | 0.84±0.20 | 0.87±0.15 | 0.77±0.23 | 0.76±0.23 | 24.36±22.18 | 17.49±22.78 | 16.22±23.95 |
| 25³        | 0.87±0.11  | 0.77±0.19 | 0.77±0.16 | 0.86±0.15  | 0.81±0.20 | 0.79±0.17 | 0.91±0.08 | 0.78±0.19 | 0.80±0.20 | 18.59±19.42 | 19.97±21.30 | 14.11±22.22 |
| 31³        | 0.83±0.13  | 0.65±0.17 | 0.40±0.21 | 0.83±0.15  | 0.76±0.21 | 0.52±0.26 | 0.85±0.13 | 0.63±0.18 | 0.37±0.21 | 22.89±23.58 | 26.55±28.28 | 20.93±21.44 |

2.4.3 Evaluation of the modality combinations

Based on the modality combinations and training pathways defined in Table 7, Table 3 compares the segmentation performance of different combinations. For the whole tumor segmentation, all four modalities were used in the E_G and C1_G groups, and their Dice coefficients for whole tumor segmentation were all higher than 0.85 in the T_G group with only two modalities. However, the Dice coefficient in the E_G group was better than that in the C1_G group, with an improvement in the enhancing part. Although the sensitivity for detecting tumors in the E_G group is not as good as that for the other two groups, its accuracy is best, with a PPV of 0.86 in whole tumor segmentation and approximately 0.80 in inner structure segmentation. In addition, the Hausdorff distance of the E_G group is also the smallest among all four groups. After exchanging the training pathways from the E_G group into the C2_G group, the Dice and PPV indicators fell catastrophically in both the whole tumor and the core region segmentation tasks. However, when referring to the enhancing part, the accuracy and Hausdorff distance are even better than those of the other groups. The results of size 19³ patch (see Table 3) also show high precision but low sensitivity in the enhancing core and indicate under-segmentation. With a much poorer performance in the Hausdorff distance, we deduce that the C1_G group has a more severe over-segmentation problem with higher sensitivity and lower precision.

Fig. 2 shows that the T_G group was more successful in identifying the blue necrotic areas than the E_G group. However, it was easy to make mistakes in the segmentation of whole tumors, especially in
the peritumoral area, which was characterized by the over-segmentation of necrotic areas. Here, the necrotic areas, enhanced nuclei and edema areas are expressed in blue, red and green, respectively. After using the information from the four modes, the segmentation performance of the C1_G group is improved, but compared with the E_G group, the accuracy of the C1_G group is still insufficient. The results for the C2_G group were worse than those for the C1_G groups. From multimodal combination training, grouping the training of the information from the four modes can indeed increase the segmentation accuracy.

### Table 3. Evaluation of the performance of the modality combinations and training pathways.

| Mode | DICE | PPV | Sensitivity | Hausdorff distance |
|------|------|-----|-------------|-------------------|
|      | Tumor | Core | Enha. | Tumor | Core | Enha. | Tumor | Core | Enha. | Tumor | Core | Enha. |
| T_G  | 0.85±0.14 | 0.75±0.22 | 0.77±0.20 | 0.81±0.18 | 0.78±0.24 | 0.78±0.17 | 0.91±0.09 | 0.78±0.22 | 0.81±0.19 | 27.04±25.45 | 27.92±27.53 | 20.43±24.86 |
| E_G  | 0.87±0.11 | 0.77±0.19 | 0.77±0.16 | 0.86±0.15 | 0.81±0.20 | 0.79±0.17 | 0.91±0.08 | 0.78±0.19 | 0.80±0.20 | 18.59±19.42 | 19.97±21.30 | 14.11±22.22 |
| C1_G | 0.87±0.12 | 0.77±0.20 | 0.74±0.21 | 0.83±0.16 | 0.79±0.20 | 0.71±0.24 | 0.92±0.07 | 0.83±0.21 | 0.85±0.21 | 26.20±26.18 | 25.20±25.71 | 21.76±29.04 |
| C2_G | 0.71±0.28 | 0.71±0.31 | 0.75±0.27 | 0.72±0.30 | 0.75±0.32 | 0.81±0.27 | 0.77±0.30 | 0.70±0.30 | 0.74±0.27 | 29.83±26.18 | 26.64±28.73 | 14.83±22.52 |

### 2.4.4 Evaluating the effectiveness of the post-processing steps

We incorporate the fully connected 3D CRF with several other procedures into the 3D CNN to achieve more structured predictions. As shown in Table 4, after the post-processing step, the Dice coefficient of the whole tumor and enhancing core all increase by two percent, and the PPV improves as well. Combined with the post-processing step, the void area disappears, the adhesion part is separated, and the segmentation result is more consistent with the manual labeling standard. Therefore, the performance of segmentation improves as the Hausdorff distance decrease a lot. One interesting finding is that with a post-processing step, the sensitivity of detecting the whole tumor weakens slightly; this finding is considered a side-effect of CRF addressing the over-segmentation issue. As shown in the comparisons in Fig. 3, our full method (CNN+post) achieves better results where the isolated holes disappear and the merged areas are detached.

### Table 4. Evaluation of the effectiveness of the post-processing

| Mode | DICE | PPV | Sensitivity | Hausdorff distance |
|------|------|-----|-------------|-------------------|
|      | Tumor | Core | Enha. | Tumor | Core | Enha. | Tumor | Core | Enha. | Tumor | Core | Enha. |
| CNN  | 0.87±0.11 | 0.77±0.19 | 0.77±0.16 | 0.86±0.15 | 0.81±0.20 | 0.79±0.17 | 0.91±0.08 | 0.78±0.19 | 0.80±0.20 | 18.59±19.42 | 19.97±21.30 | 14.11±22.22 |
| CNN+post | 0.89±0.09 | 0.78±0.24 | 0.78±0.23 | 0.93±0.07 | 0.83±0.24 | 0.82±0.21 | 0.85±0.13 | 0.79±0.21 | 0.78±0.24 | 13.50±21.06 | 13.30±14.36 | 10.22±11.22 |

### 2.4.5 Comparison with other methods

Table 5 compares our approach with several other methods participated in BraTS 2017 proceedings. Because our method generates a CNN segmentation model by training four MR modality images (FLAIR, T1c, T1 and T2), we list several methods for brain gliomas segmentation based on the four MR modalities with different architectures and networks. And as we had not participated the BraTS 2017 to get our validation results, all the comparison are made on the training data. Here, the specificity is excluded as all of the methods can reach a specificity of almost 0.98-0.99. And several participants has missed some of their metrics. As shown, we can see that though our hausdorff distance need
improvement, our other metrics performs well and even ranked top. Further details would be discussed in next section.

Table 5. Comparison with the segmentation models participated in BraTS 2017.

| Method        | Dice       |   | Sensitivity |   | Hausdorff distance |
|---------------|------------|---|-------------|---|-------------------|
|               | Tumor | Core | Enha.| Tumor | Core | Enha.| Tumor | Core | Enha.  |
| Cata et al. [37] | 0.87  | 0.69 | 0.67 | 0.85  | 0.62 | 0.72 | 6.41  | 9.59 | 7.15   |
| Hu et al. [38]  | 0.79  | 0.77 | 0.72 | 0.79  | 0.76 | 0.77 | -     | -    | -     |
| Li et al. [39]  | 0.88  | 0.87 | 0.78 | 0.89  | 0.87 | 0.78 | -     | -    | -     |
| Castillo et al. [40] | 0.87 | 0.78 | 0.74 | 0.87  | 0.74 | 0.78 | 21.32 | 9.18 | 7.52   |
| CNN+post       | 0.89  | 0.78 | 0.76 | 0.85  | 0.78 | 0.78 | 13.50 | 13.38 | 10.22  |

2.4.6 Segmentation performance on the BraTS 2019

We also tested our methods in BraTS 2019. The results are shown in Table 6. As post-processing procedure is still useful in improvement on performance. Because BraTS 2019 dataset abandons the post-operative data and refines the data categorization, our method obtains a slightly better result by comparing with the one of BraTS 2017 dataset, especially in Hausdorff distance.

Table 6. Performance comparison on the BraTS 2019.

| Method | Dice | PPV | Sensitivity |   | Hausdorff distance |
|--------|------|-----|-------------|---|-------------------|
|        | Tumor | Core | Enha. | Tumor | Core | Enha. | Tumor | Core | Enha.  | Tumor | Core | Enha.  |
| CNN    | 0.88±0.10 | 0.76±0.21 | 0.74±0.23 | 0.86±0.13 | 0.77±0.23 | 0.70±0.20 | 0.92±0.08 | 0.82±0.19 | 0.86±0.22 | 20.68±23.40 | 19.94±22.38 | 17.40±2.13 |
| CNN+post | 0.90±0.09 | 0.78±0.22 | 0.76±0.26 | 0.91±0.10 | 0.83±0.18 | 0.73±0.21 | 0.89±0.11 | 0.81±0.23 | 0.86±0.20 | 13.28±17.42 | 15.29±19.56 | 15.66±22.55 |

3 Discussion

This paper proposed a multi-scale residual CNN with a double pathway to segment gliomas in multi-modality MR images. Our full method, CNN+CRF, includes three necessary steps. First, to build the CNN segmentation model, all the images are first normalized by our proposed histogram matching techniques. Then, the multi-scale, double-pathway residual network is trained by 4 MR image modalities and is applied to generate the rough segmentation results on the test data. Finally, the post-processing step with CRF, refines the segmentation results by adhering isolated areas and removing holes.

Patch sizes and modality combinations were analyzed in the framework. In the training process, the size of the sample patches has certain limitations. When it decreases, the details in the ROI will be reduced, leading to an impairment in the segmentation of large lesions and under-segmentation of smaller parts. In contrast, when the patch becomes larger, it will contain too much irrelevant information that will hinder useful feature extraction.

By comparing different modal combinations in the training methods, it was found that modal fusion can enhance the feature information of the double pathway and improve the network performance to a certain extent. T1 and T2 images can provide information when segmenting the whole tumor. However, when segmenting the enhancing core, T1c images provide more useful information than the other modes. In this study, a double pathway with different resolutions is used to learn complementary
information in each modality. For example, T1c and FLAIR images can display detailed information such as tumor nuclei and edema areas, while T1 and T2 images provide location information of tumors relative to brain structures.

The comparison between the different methods is shown in Table 4. All the methods are designed based on 4 modalities of MR images and tested on the same dataset. It’s worth to mention that all the results are based on the test on validation set, while Cata et al.[22] built a Mask-V-net architecture with two V-net for segmentation. With four modalities as an input in first net and adding a tumor mask as a box into the next one, their methods help to confine the tumor area and obtain a good results in Hausdorff Distance, however the Dice and sensitivity is not promising than us. Hu et al.[23] also built multi Cascaded U-net each training one modality in 2D slices to fuse the feature of multi-modalities, and the results is worse than our segmentation performance, considering it results from the shortage of 2D network in MRI gliomas segmentation. Meanwhile, Li et al.[24] used an adversarial network with two simple 2D CNN to obtain a promising results, especially their Dice and sensitivity of tumor core is much better than us. Though, our whole tumor’s Dice is still the best and the ability of detecting enhancing part is as good as theirs. Castillo et al.[25] also made their work based on DeepMedic, while they add several pathways in it. However, the job is focusing on the multi-resolution information in MRI images and all the modalities are still sampled with early fusion before inputing into the same connection DeepMedic adopted. Compared with us, the performance is very close in several metrics. Our dice and hausdorff distance is better, while their method is more sensitive to lesion. And our further network is considering adding more pathway and make a more complicated connection, such as densely-connection for a more complex and accurate multi-modal feature fusion strategy.

4 Conclusion

This paper presents a multimodal 3D residual CNN algorithm for the segmentation of gliomas and intratumoral structures. Specifically, a multimodal, multiscale, double-pathway, 11-layer 3D residual CNN was established in this study. After normalizing all the images in the BraTS 2017 dataset by histogram matching and network initiation, the data will be divided into a training group, an evaluation group and a test group. This study builds and trains the segmentation network on the training group data, adjusts the parameters in the evaluation group data, generates rough segmentation results on the test data, and uses a CRF to remove holes and merge isolated areas to achieve the final optimization of these segmentation results. The method can make full use of modal image information, such as T1, T2, T1C, FLAIR, to achieve the accurate and effective segmentation of gliomas and their internal structure. To improve the segmentation performance for tumors and their internal structures and ultimately achieve the end-to-end classification of LGGs and HGGs, we will continue to explore the methods of mixed modal combinations and multi-pathway CNNs to extract the structural features of tumors.

5 Methods

5.1 Preprocessing image

MR artifacts are usually caused by the bias field and the tremendous influence of brain tumors on different intensity ranges. Specifically, one problem with combining images from different scanners
and analyzing them with automatic tools is that the qualities of MR images can be dependent on the scanner manufacturer, field strength, MRI protocol, and so on; this issue is called the multicenter problem and especially manifests as a varied range of intensities [26]. Traditional histogram matching techniques cannot achieve proper intensity normalization. Goetz [27] proposed a normalization method for any kind of mode image, which subtracts the whole gray-level mean of the mode image from each image and adjusts the standard deviation of the gray distribution to 1. As gliomas often appears in high-signal areas, which causes deviations in the gray-level mean and variance of the image, the gray-level value of the highest histogram bin is taken as the mean (typically the gray-level distribution of white matter [28]). The K-means clustering algorithm is used to obtain the optimal clustering center of all the images of the same mode, that is, the mean and variance of the distribution. Therefore, this method normalizes all the images by histogram matching. The procedures are as follows:

For each type of mode image of each subject,
① Reduce the noise with the method of SUSAN [29] and correct the bias field using FAST [30].
② Linearly transform the intensity range of the image to the range of [0,255] and calculate its histogram.
③ For an image \( I = \{ I_k \} k = 1,2, ..., N \), where \( N \) is the number of voxels and \( I_k \) is the intensity of the \( k \)-th image. \( \bar{I} \) denotes the gray value of the highest histogram bin, and the robust deviation is \( \sigma = \sqrt{\sum_{k=1}^{N}(I_k - \bar{I})^2/N} \). Then, the image can be normalized to the range of [-1,1] with its own parameters \( \bar{I} \) and \( \sigma \).
④ Using these two parameters, the mean \( \bar{I} \) and variance \( \sigma \), as coordinate axes, the mean and variance distributions of all the images of the same mode can be obtained.
⑤ Using the K-means clustering algorithm to cluster the distribution, we can obtain a set of optimal mean and variance parameters, which will achieve gray matching for all the images.

Fig. 4 shows that the FLAIR images with different low-grade gliomas (LGGs) (a1-a3) or T1c images with different high-grade gliomas (HGGs) (a4-a6) are normalized to the images (b1-b3) and (b4-b6), respectively, with a more consistent grayscale. The results indicate that the robust normalization method overcomes the multicenter issue and makes the MRI scans of different patients comparable. The improvement can also be confirmed by the change in the image intensity histogram before and after the normalization (Fig. 5). Fig. 5b shows that the histograms of all the FLAIR images with LGGs are more compact after normalization. It should be noted that the simple intensity normalization technology in 3D Slicer [31] can also match all images of the same mode to the same intensity space, but obtain the same distribution results similar to original one in Fig. 5a. Therefore, the segmentation performance of such normalized images for subsequent CNN segmentation is slightly worse than ours (shown in Table 1).

### 5.2 Multi-scale double-pathway 3D residual CNN

Gliomas can occur in different parts of the brain, such as the supratentorial infratentorial regions or the brainstem. The shapes of gliomas vary, such as stellate, oligo-branched, ventricular or mixed types, and the sizes of gliomas are very different. To successfully segment tumor tissue, a large number of multi-scale spatial features with context information around the tumor should be considered. Therefore, by incorporating both local and global contextual information into 3D CNN, the study employs parallel convolutional pathways of multi-scale resolution containing multimodal context information of gliomas; this method can effectively segment gliomas, as shown in Fig. 6 [32][33].

The blue pathway (\( Pathway_{Nor} \)) inputs the sample patches of normal resolution for handling the tumor appearance details, while the down-sampled sample patches that record the spatial context information of the sample patches of normal resolution are input into the yellow pathway (\( Pathway_{Low} \)) for learning the high level features, such as locations, in low resolution. Note that the image in the yellow pathway was down-sampled to reduce the computational complexity before training. Moreover, these two pathways are independent of each other, which makes it possible to
extract and study each of their characteristics and information using multi-scale resolution. The patches in the two pathways are all normalized to standard patches with a mean of 0 and standard deviation of 1 and augmented with symmetry, reversal and duplication for convolution processing. Each pathway consists of 8 convolution layers, where the last 6 layers are connected with residual blocks in two adjacent layers to enhance the robustness. After 8-layer convolution, the feature maps from Pathway_{Low} are upsampled to match the ones from Pathway_{nor}. The two groups of features are concatenated and input into the fully connected layer to integrate the trained features and location information. Finally, based on the trained features, the softmax function is used to classify the voxels into four rough segmentation results: necrotic core (red), enhanced area (blue), edema area (green) and normal tissue area (yellow).

**Convolutional layer** The \( l^{th} \ (l \in [1, L]) \) convolutional layer has \( F_l \) feature maps. The value at the position \((x, y, z)\) in the \( i^{th} \) feature map from the \( l^{th} \) layer, \( h_{li}^{xyz} \), is calculated as

\[
h_{li}^{xyz} = f \left( \sum_{k=1}^{F_{l-1}} \frac{y}{2} \sum_{v=-\frac{y}{2}}^{\frac{y}{2}} \frac{w}{2} \sum_{w=-\frac{w}{2}}^{\frac{w}{2}} h_{(l-1)k}^{(x+u)(y+v)(z+w)} W_{lik}^{uvw} + b_{li} \right),
\]

which is the result of convolving each of the feature maps \( h_{(l-1)k} \ (k \in [1, F_{l-1}]) \) from the \((l - 1)^{th} \) layer with a 3D kernel \( W_{lik} \) and with a size of \((U + 1) \times (V + 1) \times (W + 1)\), adding a learned bias, \( b_{li} \), and applying a nonlinearity active function. Here, the kernel size is set to \( 3 \times 3 \times 3 \) to make the architecture deeper since deeper networks have more discriminative power due to the additional nonlinearity and better quality of local optima [34]. In this work, we choose the stride as 1, as larger strides downsample the feature maps and cause inaccurate segmentation. The size of the outputted feature maps is calculated as follows: \( \text{size of the input feature maps} / \text{size of the kernel} + 1 \).

**Fully connected layer** After the convolution, the feature maps from Pathway_{Low} are upsampled to match the feature maps from Pathway_{nor}. Then, the two groups of features are concatenated together and inputted into the following two fully connected layers. The value at the position \((x, y, z)\) in the outputted feature map, \( h_{li}^{xyz} \), is calculated as a result of connecting the corresponding value at the same position from each of the input feature maps \( h_{(l-1)k} \ (k \in [1, F_{l-1}]) \) with respective weight \( W_{lik}^{xyz} \), and adding a bias \( b_{li}^{xyz} \), then applying a nonlinearity active function, as shown in Eq. (2),

\[
h_{li}^{xyz} = f \left( \sum_{k=1}^{F_{l-1}} h_{(l-1)k}^{xyz} W_{lik}^{xyz} + b_{li}^{xyz} \right).
\]

**Activation function** The parametric rectified linear unit (PReLU) [35] is chosen as the nonlinear activation function in the convolution and fully connected layers instead of the rectified linear unit (ReLU) since its coefficient \( a_i \) is self-adaptive to guarantee the output of each neuron during the training process. PReLU is defined as

\[
f(t_i) = \begin{cases} 
    t_i, & t_i > 0 \\
    a_i t_i, & t_i \leq 0
\end{cases}
\]

where \( t_i \) denotes the temporary value of the active function during the training stage. In the case of over-fitting problem, the dropout regularization is adopted to cut off some hidden neurons with a probability of 50% [36].
**Classification layer** A position-wise softmax function is chosen as a classifier,

\[ P_c(h^{xyz}) = \frac{e^{h_{c}^{xyz}}}{\sum_{c=0}^{C-1} e^{h_{c}^{xyz}}}, \]  

(4)

where \( h^{xyz} \) denotes the pixel to be classified, \( h_{c}^{xyz} \) is the value of the pixel in the \( c^{th} \) feature map, \( C = 4 \) is the number of classes, and \( P_c(\cdot) \) means the possibility of pixel at position \( (x, y, z) \) belonging to class \( c \).

**Optimization and residual connection** For the gradient descent method, the networks are trained using the root mean square prop (RMSProp) optimizer, which introduces an attenuation coefficient \( \rho \) to regulate the cumulative square gradient \( r \) and is calculated as

\[ r \leftarrow \rho r + (1 - \rho) \cdot dg^2, \]  

Subj. to \( \Delta \theta = -\frac{l_r}{\sqrt{\delta + r}} \cdot dg; \ \theta \leftarrow \theta + \Delta \theta; \]

where \( l_r \) is the learning rate, \( \theta \) denotes the network parameters, and \( \delta \) is a small constant to ensure that the denominator is not zero. As the network goes deeper, it becomes more difficult to train. One reason for this difficulty is the vanishing or exploding gradient of the activation function during propagation; the other reason is degradation with amplified errors caused by over-fitting. Therefore, a residual module [37] is introduced to achieve an identical mapping. As shown in Fig. 6, the traditional mapping (Fig. 6b) is altered to the residual function \( h_{t+2} - h_t \) between the two layers (Fig. 6c). When its value is 0, it is an identity mapping. Moreover, the residual blocks focus on the very small disturbance in the network, which is relatively more sensitive to the learning of parameters.

**5.3 Dense training on image patches**

**Patch size and balanced sampling** Tumor volume varies greatly in different brain images. For example, Fig. 7 shows the registered FLAIR (Fig. 7a) and T1c (Fig. 7b) images of the same patient, showing the larger green edema area (Fig. 7a) and pink necrotic core (Fig. 7b) near the midsagittal section, respectively, while Fig. 7b also shows the multiple small lesions (indicated by yellow arrow) in T1c. The information in different modalities differs a lot. The green area of edema is pretty large in FLAIR image while small necrotic core in pink is hardly to find in T1c images. This phenomenon leads to two problems. First, what patch size is appropriate? As shown in Fig. 7, both large (upper) and small (lower) patches may contain different amounts of normal tissue (green dots) and tumor tissue (red dots). The determination of the patch size needs to balance whether the tumor boundary contains normal tissue and the degree of inclusion. Large patches tend to contain too much irrelevant normal tissue when delineating the internal structure of the tumor (red cube in Fig. 7), and small patches fail to include many details, which makes it difficult to extract features (blue cube in Fig. 7). The second problem is related to sampling. As shown in Fig. 8a, uniform sampling can easily lead to a significant reduction in the sensitivity of tumor tissue recognition. Therefore, we sampled the tumor tissue (red dot) and the normal tissue (green dot) with equal weight (50%) (Fig. 8b).
**Modality combination** Fig. 9 shows that multimodal MRI images can provide different medical diagnosis information due to different imaging parameters. Edema caused by tumors has a low signal on T1 images, as T1 images can provide only the anatomical information of the brain. In T2 images, the edema area presents a high signal intensity, and it is easy to observe the uneven gray enhancement area, i.e., the relative location of the lesion. To avoid the interference of a high signal of cerebrospinal fluid (CSF) on tumor tissue, FLAIR imaging can suppress the signal of the CSF region, thus allowing a better detection of small hyper intense lesions. Because of the administration of a contrast agent, peritumoral angiography on T1c images is helpful for distinguishing necrotic areas from the enhancing core within a tumor.

Previous methods based on multimodal images usually mix the patches from FLAIR and T1c together to train each pathway simultaneously. This study employs all four modalities, and different groupings are used to train the two pathways. Table 7 shows the four modality combinations: the traditional group (T_G) {FLAIR, T1c}, the experimental group (E_G) {{FLAIR, T1c}, {T1, T2}}, control group 1 (C1_G) {FLAIR, T1c, T1, T2}, and control group 2 (C2_G){{T1, T2}, {FLAIR, T1c}}. Table 7 also shows the methods for training the pathways.

**Table 7.** Modality combinations for training the pathways.

| Pathway | Two modalities | Four modalities |
|---------|----------------|----------------|
| T_G: Traditional | E_G: Experimental | C1_G: Control 1 |
| Pathway_Nur | Pathway_Low | |
| [FLAIR, T1c] | (FLAIR, T1c) | (FLAIR, T1c) |
| (T1c, FLAIR) | (T1, T2) | (T1, T2) |

### 5.4 Image post-processing

The results of CNN segmentation may be affected by noise and local extrema, and some isolated regions or holes appear, which need to be processed. Based on CRFs [38, 39], a 3D fully connected CRF is built to refine the segmentation results. The energy loss function is defined as

\[
E(X, I) = \sum_i \varphi_u(X_i) + \sum_{i<j} \varphi_p(X_i, X_j),
\]  
(6)

where, for voxel \(i\), \(\varphi_u(X_i) = -\ln P(X_i|I)\) and \(P(X_i|I)\) is the output from the CNN. According to Koltun’s observation [40], when referring to the multiclass segmentation of images, \(\varphi_p(X_i, X_j)\) can be defined as a linear combination of Gaussian kernels that is sensitive to the image contrast:

\[
\varphi_p(X_i, X_j) = \mu(X_i, X_j) \left( w^1 \exp \left( \frac{-|p_{l,d} - p_{j,d}|^2}{2\theta_{z,d}^2} - \frac{|l_{i} - j_{i}|^2}{2\theta_{d}^2} \right) + w^2 \exp \left( \frac{-|p_{l,d} - p_{j,d}|^2}{2\theta_{z,d}^2} \right) \right),
\]  
(7)

when \(X_i \neq X_j\), \(\mu(X_i, X_j) = 1\); otherwise, \(\mu(X_i, X_j) = 0\). In Eq. (7), \(p\) and \(I\) denote the position and intensity of pixel \(X\). The first Gaussian function describes the probability distribution of two adjacent pixels and spatially identical labels belonging to the same class by two parameters: \(\theta_{z,d}\) and \(\theta_{d}\). The other Gaussian function uses the parameter \(\theta_{y,d}\) to enforce the smooth classification of pixels with
homogenous labels in a large range to avoid local voids. Finally, the weights $w^1$ and $w^2$ adjust the relative ratio of the two factors.

After the CRF, we also further improve the segmentation results by a clustering method and applying following post-processing procedures according to the clinical experience (Here, noted that all the gray value is normalized to 0-255):

a) All the clusters in the tumor mask with mean gray value of FLAIR image and T2 image are both higher than the 150, are considered image noise and removed from the tumor mask.

b) In general, tumor tissues have high signal in at least one modality. Therefore, for the voxels in the tumor mask, if their gray values are less than 0.95 times of the average intensity of FLAIR image and T2 image, and their gray values on T1C image are less than 150, then these voxels will be excluded from the tumor mask.

c) Fill the holes in the necrotic area, as they are very likely to be necrosis.

d) For those voxels within the enhancing tumor areas whose T1c gray value is less than 110, they will be regarded as classification errors and changed into the ones in necrotic regions.

e) Voxels in clusters whose volume is smaller than 0.2 times the maximum clustering volume are considered to be non-tumorous regions, while those in necrotic clustering are the same.

6 Declarations

Ethics Approval and Consent to Participate

All patients signed an informed consent approved in BraTs dataset.

Consent for Publication

Not applicable.

Availability of Data and Materials

The BraTs data are available in https://ipp.cbica.upenn.edu.

Competing Interests

The authors have declared that no competing interests exist.

Funding

This research was supported by grants from the National Key Research and Development Program of China (2017YFC0110701). This research was also supported by the National Natural Science Foundation of China (grant 60972102).

Authors’ Contributions

MP and YS designed the experiments; MP carried out the experiments; MP, YS and ZS analyzed the experimental results; MP and YS wrote the manuscript; MP, YS and ZS revised the manuscript.
References

1. Bray, F., et al., Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. 2018. 68(6): p. 394-424.

2. İşın, A., C. Direkoğlu, and M. Şah, Review of MRI-based Brain Tumor Image Segmentation Using Deep Learning Methods. Procedia Computer Science, 2016. 102: p. 317-324.

3. Gordillo, N., E. Montseny, and P. Sobrevilla, State of the art survey on MRI brain tumor segmentation. Magnetic Resonance Imaging, 2013. 31(8): p. 1426-1438.

4. Mazzara, G.P., et al., Brain tumor target volume determination for radiation treatment planning through automated MRI segmentation. International Journal of Radiation Oncology Biology Physics, 2004. 59(1): p. 300-312.

5. Tustison, N.J., et al., Optimal Symmetric Multimodal Templates and Concatenated Random Forests for Supervised Brain Tumor Segmentation (Simplified) with ANTsR. 2015. 13(2): p. 209-225.

6. Anitha, V. and S.J.I.C.V. Murugavalli, Brain tumour classification using two-tier classifier with adaptive segmentation technique. 2016. 10(1): p. 9-17.

7. Lecun, Y., et al., Gradient-based learning applied to document recognition. Proceedings of the IEEE, 1998. 86(11): p. 2278-2324.

8. Zikic, D. and I. M. Segmentation of brain tumor segmentation using deep convolutional neural networks. in MICCAI Multimodal Brain Tumor Segmentation Challenges. 2014.

9. Urban, G. Multi-modal brain tumor segmentation using deep convolutional neural networks. in MICCAI Multimodal Brain Tumor Segmentation Challenges. 2014.

10. Havaei, M., et al., Brain tumor segmentation with Deep Neural Networks. Medical Image Analysis, 2017. 35: p. 18-31.

11. Davy, A. Brain tumor segmentation with deep neural networks. in MICCAI Multimodal Brain Tumor Segmentation Challenges. 2014.

12. Kamnitsas, K., et al., Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation. Medical Image Analysis, 2017. 36: p. 61-78.

13. Zhao, X., et al., A deep learning model integrating FCNNs and CRFs for brain tumor segmentation. 2017. 43: p. 98-111.

14. Naíomi, F. Deep Learning Neural Networks for Segmentation of brain Tumor in Multi Modal MRI. in MICCAI Multimodal Brain Tumor Segmentation Challenge 2017.

15. Jose, D., Ismail. Ben , A, and Christian. JD. Dense Multi-path U-net for Ischemic Stroke Lesion Segmentation in Multi Images Modalities. Available from:https://arxiv.org/abs/1810.07003

16. J . Dolz, K . Gopinath, J. Yuan, H. Lombaert, C. Desrosiers, and I. B. Ayed. Hyperdense-net: a hyper-densely connected cnn for multi-modal image segmentation. Available from:https://arxiv.org/abs/1804.02967.

17. M. Augun, Y. H .Sahin, and G. Unal. Multi modal convolutional neural networks for barin tumor segmentation. Available from:https://arxiv.org/abs/1809.06191.

18. Bakas, S., et al., Advancing The Cancer Genome Atlas glioma MRI collections with expert segmentation labels and radiomic features. 2017. 4(170117): p. 170117.

19. S. Bakas, M. Reyes, A. Jakab, S. Bauer, M. Rempfler, A. Crimi, et al., "Identifying the Best Machine Learning Algorithms for Brain Tumor Segmentation, Progression Assessment, and Overall Survival Prediction in the BRATS Challenge", arXiv preprint arXiv:1811.02629 (2018)

20. Ioffe, S. and C. Szegedy. Batch normalization: accelerating deep network training by reducing internal covariate shift. in International Conference on International Conference on Machine Learning. 2015.

21. Soltaninejad, M., et al., Identifying the Best Machine Learning Algorithms for Brain Tumor
Segmentation, Progression Assessment, and Overall Survival Prediction in the BRATS Challenge. 2019; Available from: [https://arxiv.org/abs/1811.02629](https://arxiv.org/abs/1811.02629).

22. Casamitjana, A., et al., Cascaded V-Net Using ROI Masks for Brain Tumor Segmentation. in Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries. 2018. Cham: Springer International Publishing. p 381-391.

23. Yan, H. and Yong, X., 3D Deep Neural Network-Based Brain Tumor Segmentation Using Multimodality Magnetic Resonance Sequences. in Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries. 2018. Cham: Springer International Publishing. p 424-435.

24. Zeju, L., et al., Brain Tumor Segmentation Using an Adversarial Network. in Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries. 2018. Cham: Springer International Publishing. p 123-132.

25. Castillo, L.S., et al., Brain Tumor Segmentation and Parsing on MRIs Using Multiresolution Neural Networks. in Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries. 2018. Cham: Springer International Publishing. p 332-343.

26. Chalavi, S., et al., Quantitative and qualitative assessment of structural magnetic resonance imaging data in a two-center study. BMC Medical Imaging, 2012. 12(1): p. 27.

27. Goetz, M., W. C, and B. F. Extremely randomized trees based brain tumor segmentation. in MICCAI Multimodal Brain Tumor Segmentation Challenges. 2014.

28. Fields, R.D., Neuroscience. Change in the brain's white matter. Science (New York, N.Y.), 2010. 330(6005): p. 768-769.

29. Smith, S.M. and J.M.J.I.J.o.C.V. Brady, SUSAN—A New Approach to Low Level Image Processing. 1997. 23(1): p. 45-78.

30. Zhang, Y., M. Brady, and S. Smith, IEEE Transactions on Medical Imaging. Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm. 2002. 20(1): p. 45-57.23.

31. Fedorov A., Beichel R., Kalpathy-Cramer J., Finet J., Fillion-Robin J-C., Pujol S., Bauer C., Jennings D., Finessy F.M., Sonka M., Buatti J., Aylward S.R., Miller J.V., Pieper S., Kikinis R. 3D Slicer as an Image Computing Platform for the Quantitative Imaging Network. Magn Reson Imaging. 2012 Nov;30(9):1323-41. PMID: 22770690. PMCID: PMC3466397.

32. Hubel, D.H. and T.N.J.o.P. Wiesel, Receptive fields and functional architecture of monkey striate cortex. 1968. 195(1): p. 215-243.

33. Eigen, D., C. Puhrsch, and R. Fergus. Depth Map Prediction from a Single Image using a Multi-Scale Deep Network. in International Conference on Neural Information Processing Systems. 2014.

34. Choromanska, A., et al., The Loss Surface of Multilayer Networks. 2014: p. 192-204.

35. He, K., et al., Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification. 2015.

36. Srivastava, N., et al., Dropout: A Simple Way to Prevent Neural Networks from Overfitting. 2014. 15(1): p. 1929-1958.

37. He, K., et al. Deep Residual Learning for Image Recognition. in 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2016.

38. Sutton, C., A.J.F. McCallum, and T.i.M. Learning, An Introduction to Conditional Random Fields. 2010. 4(4): p. 267-373.

39. Swersky, M.S.K. crfChain: Matlab code for chain-structured conditional random fields with categorical features. 2008; Available from: [http://www.cs.ubc.ca/~schmidtm/Software/crfChain.html](http://www.cs.ubc.ca/~schmidtm/Software/crfChain.html).

40. Krähenbühl, P. and V. Koltun, Efficient Inference in Fully Connected CRFs with Gaussian Edge Potentials. Vol. 24. 2012.
Fig. 1. Comparison of the segmentation results based on training samples of different sizes. The necrotic cores, enhanced cores and edema areas are shown in blue, red and green, respectively.
Fig. 2. Comparison of the segmentation results obtained through training with different modal combinations. FLAIR and T1c views of the same subject and the corresponding manual annotations are displayed in the first three columns. The segmentation results for the T_G, E_G, C1_G and C2_G groups are shown in the last four columns.
Fig. 3. Illustration the segmentation results after post-processing. The necrotic areas, enhanced nuclei and edema areas are expressed in blue, red and green, respectively.

Fig. 4. (a1-a3) are the randomly selected FLAIR images with low-grade gliomas (LGGs), and (a4-a6) are the randomly selected T1c images with high-grade gliomas (HGGs). In addition, (b1-b3) and (b4-b6) are the images corresponding to (a1-a3) and (a4-a6), respectively, after intensity normalization by the proposed method.
Fig. 5. Illustration of the intensity histograms of the 75 FLAIR images with LGGs before (A) and after (B) the normalization.

Fig. 6. Illustration of the framework of the multi-scale, double-pathway 11-layer 3D residual CNN (c). The traditional connection of two convolution layers and the residual connections are shown in (a) and (b), respectively.
Fig. 7. Illustration of patch size. (a) FLAIR images with the green region indicates the edema area, while the registered T1c images in (b) shows the necrotic core in pink. The yellow arrow shows the multiple small lesions; and red and blue cubes stands for the larger or smaller size of sampling patches, respectively.

Fig. 8. (a) Uniform sampling. (b) Weighted sampling.
Fig. 9. The four modalities in the study.