Research article

SGEResU-Net for brain tumor segmentation

Dongwei Liu[1,2]**, Ning Sheng**[3]**, Tao He**[1,2]**, Wei Wang**[1]**, Jianxia Zhang**[4] **and Jianxin Zhang**[1,2,5,*]

[1] School of Computer Science and Engineering, Dalian Minzu University, Dalian 116600, China

[2] Institute of Machine Intelligence and Biocomputing, Dalian Minzu University, Dalian 116600, China

[3] Key Lab of Advanced Design and Intelligent Computing (Ministry of Education), Dalian University, Dalian 116622, China

[4] School of Intelligent Engineering, Henan Institute of Technology, Xinxiang 453003, China

[5] SEAC Key Laboratory of Big Data Applied Technology, Dalian Minzu University, Dalian 116600, China

* **Correspondence:** Email: jxzhang0411@163.com; Tel: +8641187557005.

Abstract: The precise segmentation of tumor regions plays a pivotal role in the diagnosis and treatment of brain tumors. However, due to the variable location, size, and shape of brain tumors, the automatic segmentation of brain tumors is a relatively challenging application. Recently, U-Net related methods, which largely improve the segmentation accuracy of brain tumors, have become the mainstream of this task. Following merits of the 3D U-Net architecture, this work constructs a novel 3D U-Net model called SGEResU-Net to segment brain tumors. SGEResU-Net simultaneously embeds residual blocks and spatial group-wise enhance (SGE) attention blocks into a single 3D U-Net architecture, in which SGE attention blocks are employed to enhance the feature learning of semantic regions and reduce possible noise and interference with almost no extra parameters. Besides, the self-ensemble module is also utilized to improve the segmentation accuracy of brain tumors. Evaluation experiments on the Brain Tumor Segmentation (BraTS) Challenge 2020 and 2021 benchmarks demonstrate the effectiveness of the proposed SGEResU-Net for this medical application. Moreover, it achieves DSC values of 83.31, 91.64 and 86.85%, as well as Hausdorff distances (95%) of 19.278, 5.945 and 7.567 for the enhancing tumor, whole tumor, and tumor core on BraTS 2021 dataset, respectively.

1. Introduction

Brain tumors are abnormal cells that grow in the human brain, and they can be divided into benign and malignant tumors [1]. Nowadays, the incidence of malignant tumors is relatively high, posing a considerable threat to human health. Glioma is the most common malignant tumor, and it mainly includes high-grade glioma (HGG) and low-grade glioma (LGG). Magnetic resonance imaging (MRI) can produce high-quality brain images without damage and skull artifacts, which provides comprehensive information for the diagnosis and treatment of brain tumors [2]. In addition, accurate brain tumor segmentation can not only provide valuable information such as morphology, size, and location of the tumor but also help clinicians improve the diagnosis of brain tumors. Therefore, the automatic segmentation of brain tumors on MRI is of great significance to patient treatment and has been widely studied for a long time.

Recently, deep learning methods have demonstrated superiority in the brain tumor segmentation task compared to the traditional methods. Especially, the Multimodal Brain Tumor Segmentation Challenge (BraTS), jointly held with the International Association for Medical Image Computing and Computer Assisted Intervention since 2012, has greatly promoted the progress of deep learning-related brain tumor segmentation. The earlier deep learning methods for brain tumor segmentation mainly utilize the idea of patch classification in small-scale images based on convolutional neural networks. However, these methods have limitations in maintaining image spatial continuity, which also occupies large storage space and leads to low efficiency. Motivated by the fully convolutional network, Ronneberger et al. [3] present a novel symmetric 2D U-Net model which contains down-sampling layers for feature encoding, up-sampling layers for feature recovery, and skip connections to integrate up-sampling and down-sampling features. Due to its concise and efficient architecture, U-Net largely improves the segmentation performance for medical images and quickly becomes the mainstream in brain tumor segmentation tasks. Then, to discover the potential of the model, researchers combine U-Net with various other advanced methods, such as residual module [4], attention module [5] and ensemble module [6], to produce more robust brain tumor segmentation models. Though 2D U-Net models require low memory consumption, they have limitations in containing spatial context information of 3D MRI images. Therefore, 3D U-Net models are developed to capture more spatial context information, which also achieves higher accuracy on this medical task. Similar to 2D U-Net models, a great deal of 3D modules, including residual module, dense connection block, and attention module [7–9], are integrated to gain higher segmentation performance. Among them, Myronenko [10] utilizes an asymmetrical residual U-Net, in which most of the trainable parameters of the model resided in the encoder and win the first place in the BraTS 2018 challenge. Zhao et al. [11] utilize an 3D U-Net model with dense blocks along with three tricks including data processing methods, semi-supervised learning and optimizing processes, win the second place in the BraTS 2019 challenge. Besides, Jiang et al. [12] achieve the first place using the two-stage cascaded asymmetric 3D residual U-Net. Moreover, Isensee et al. [13] place the first in the BraTS 2020 challenge based on the 3D nnU-Net [14] model. The successful application of the above method is enough to prove that the transformation based on the 3D U-Net network has a huge advantage in brain tumor segmentation.

In this work, following merits of the 3D U-Net architecture, we try to construct a novel 3D U-Net model called SGEResU-Net to segment brain tumors. SGEResU-Net embeds residual modules and spatial group-wise enhance (SGE) attention modules into a single 3D U-Net architecture, where SGE

attention modules are adopted to enhance the feature learning of semantic regions and reduce possible noise and interference with almost no extra parameters. In addition, motivated by the effectiveness of the self-ensemble module [11], we also integrate this module into the network to achieve satisfied segmentation accuracy. The overall architecture of the SGEResU-Net can be shown in Figure 1. We evaluate the SGEResU-Net model using the Brain Tumor Segmentation (BraTS) Challenge 2020 and 2021 datasets, and the experiment results on the BraTS 2021 demonstrate that it can achieve DSC values of 83.31, 91.64 and 86.85%, as well as Hausdorff distances (95%) of 19.278, 5.945 and 7.567 for the enhancing tumor, whole tumor, and tumor core, respectively.
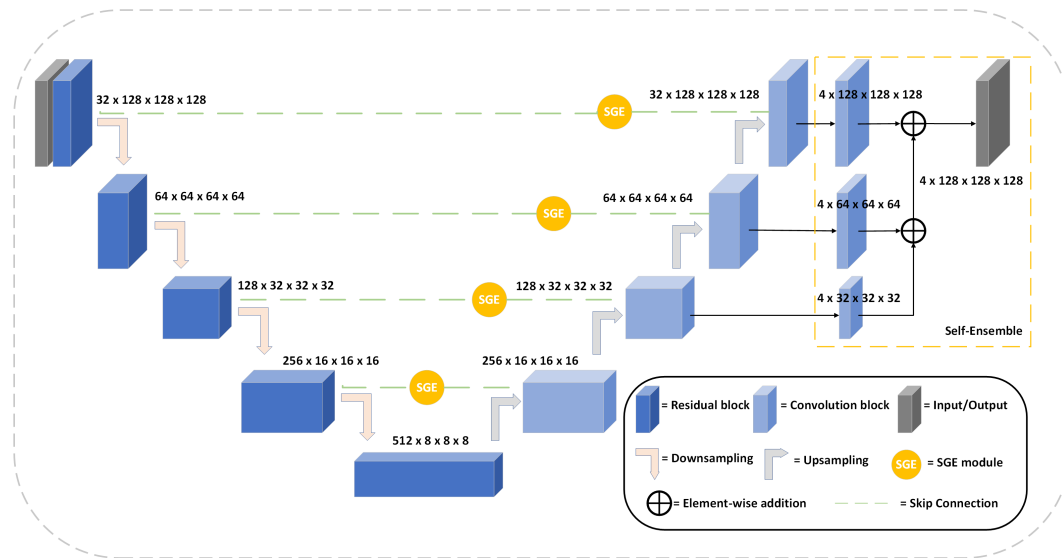

Figure 1. The overall architecture of SGEResU-Net. SGEResU-Net is an end-to-end network that integrates spatial group-wise enhance (SGE) attention and residual modules into a single 3D U-Net model, in which a series of SGE modules are adopted to replace skip connection operations for enhancing semantic feature representation and suppressing possible noise. Besides, the self-ensemble module is utilized to improve the brain tumor segmentation performance by combining prediction results from different scales.

2. Methods

In this section, we mainly give a detailed description of the spatial group-wise enhance attention residual U-Net (SGEResU-Net) model for the brain tumor segmentation. We first introduce the overall architecture of SGEResU-Net and then describe the utilized 3D SGE, 3D residual and the self-ensemble modules. Finally, the loss function employed is given.

2.1. Network architecture

The SGEResU-Net model is a trainable 3D brain tumor segmentation network that simultaneously embeds 3D spatial group-wise enhance attention module and residual module into a single U-Net

architecture, as shown in Figure 1. The size of the input data for this network is $4 \times 128 \times 128 \times 128$, which means that the image size is $128 \times 128 \times 128$, and the channel number is 4. The same as the U-Net model, the core of SGEResU-Net consists of an encoder path, a decoder path and skip connection units. The encoder path includes four residual blocks and a bottom convolutional layer to capture high-level features of the context semantic information, while its decoder path utilizes four normal convolutional blocks to recover features. Besides, to provide the effective information for brain tumor segmentation, the spatial group-wise enhance (SGE) attention module is added as the horizontal connection, which may enhance the feature learning of semantic regions and reduce possible noise and interference with almost no extra parameters, and it will also gain the possible accuracy improvement. Finally, an ensemble module given by Zhao et al. [11] is adopted in this network, aiming to combine prediction results on different scales. Using the typical encoder-decoder structure, SGEResU-Net can well conserve the high-level context information in deep layers by fusing shallow features and deep features. To better describe the SGEResU-Net model, we will introduce the details of the utilized 3D SGE, 3D residual and the self-ensemble modules in the following context of this section.
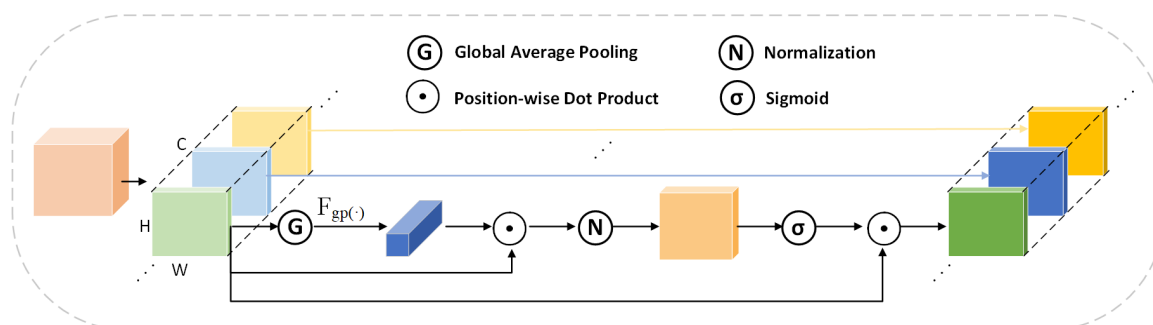

Figure 2. 3D spatial group-wise enhance (SGE) attention module. SGE groups the feature along the channel dimension of deep features to form several sub-features and then utilizes the attention mechanism guided by similarities between the global and local feature representation inside each semantic feature group.

2.1.1. SGE module

Motivated by the idea of grouping features and attention mechanism, Li et al. [15] propose the spatial group-wise enhance (SGE) attention module for visual tasks. SGE module aims to change the importance of semantic features in each semantic group by using an attention factor to reduce the influence of similar patterns and noisy background. More specifically, Li et al. think that CNNs generate the feature representation of complex objects by collecting hierarchical and different parts of semantic sub-features, however, the activation of these sub-features is often spatially affected by similar patterns and noisy backgrounds, which leads to uncorrected localization and classification. To resolve these problems, SGE groups the feature along the channel dimension of the deep feature map to form several sub-features, and then utilizes the attention mechanism guided by similarities between the global and local feature representation inside each semantic feature group. The SGE model not only enhances the feature learning of semantic regions but also compresses noise and interference with

almost no extra parameters. Therefore, in this work, we embed the SGE module at each horizontal connection position of the segmentation network. Besides, as the MRI brain tumor image is in 3D format, we extend the original 2D SGE attention module into the 3D module to fit for the input data, and the overall architecture of 3D SGE can be demonstrated in Figure 2. In this way, not only the shallow feature information can be better extracted, but also the deep semantic information can be well combined to achieve a better segmentation result.

2.1.2. Residual module

As we all know, the residual module can effectively avoid the problem of vanishing or exploding gradients during model training, which also accelerates the convergence of the network [16]. In this work, SGEResU-Net utilizes a variety of residual blocks (Resblocks) to capture semantic features of brain tumor images during the down-sampling process, and its basic architecture can be illustrated in Figure 3. Each residual block consists of three convolutional layers, three GroupNorm layers, and two Relu layers. We do not replace all the conventional convolutions with residual modules like the traditional one, but only replace the second conventional convolution of each layer. It is worth noting that the residual module is not employed during the up-sampling process, which means that we adopt the normal convolution block (Convblock) to decode the encoding feature.
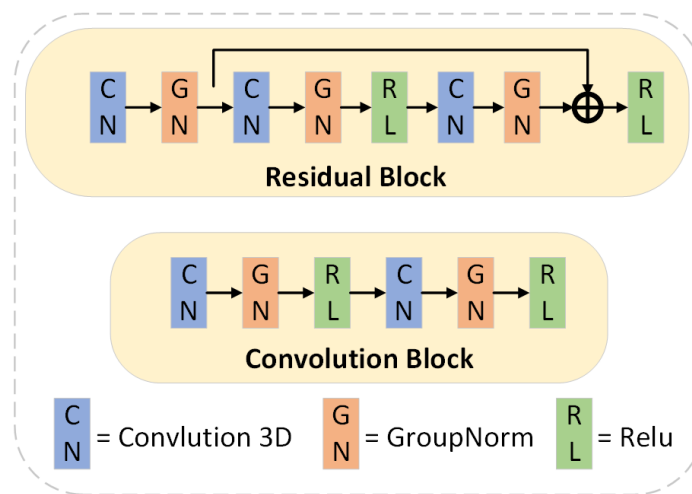

Figure 3. The overall architectures of Resblock (top) and Convblock (bottom).

2.1.3. Self-ensemble module

We employ the self-ensemble module given by Zhao et al. [11] to assist the segmentation of brain tumors. Fusing prediction results of different layers to generate the final result is a common strategy to improve the network performance in various computer vision tasks. Inspired by this idea, Zhao et al. also utilize the self-ensemble module for brain tumor segmentation, which joins predictions at each scale to obtain the final segmentation result. Following this network, we also embed the self-ensemble module into the last three layers of SGEResU-Net as shown in Figure 1. Specifically, a prediction

result is firstly generated in the third-to-last layer, and then up-sampled to the size of the feature map of the second-to-last layer. Then, combined with the prediction generated by the second-to-last layer, it will be up-sampled to the last layer. Finally, it fuses the prediction result of the last layer to generate the final prediction result.

2.2. Loss function

The brain tumor segmentation task suffers from a serious class imbalance problem. To address this problem, we adopt a combination of Dice loss and Cross-Entropy loss as the loss function for our network. The combined loss function Loss can be defined as follows:

$$Loss = \alpha L_{DC} + (1 - \alpha)L_{CE} \tag{2.1}$$

where α is the balance parameter varying from 0 to 1.

The Dice loss is a commonly used loss function for the medical image segmentation, which can make the model effectively learn those samples with unbalanced categories. Meanwhile, the cross-entropy loss is a good solution to the problem of multi-task imbalance, which decreases the difference between the training samples and evaluation metrics. In this way, we combine these two losses to solve the problem of category imbalance to a certain extent. The Dice loss and the cross-entropy loss can be expressed as follows:

$$L_{DC} = 1 - \frac{2\sum_{i=0}^{N} y_i \hat{y}_i}{\sum_{i=0}^{N}(y_i + \hat{y}_i)} \tag{2.2}$$

$$L_{CE} = -\sum_{i=0}^{N} y_i log \hat{y}_i \tag{2.3}$$

where N indicates the total number of classes, y_i is the one-hot encoding (0 or 1) for class i, and $\hat{y}_i$ denotes the correct predicted probability for the class.

3. Experimental datasets and settings

In this section, we first introduce the BraTS 2020 and 2021 datasets utilized to evaluate the SGEResU-Net model, followed by the brief description of the pre-processing method. After that, the evaluation metrics and experiment settings are briefly described.

3.1. Datasets and pre-processing

We adopt two recently released MRI brain tumor benchmarks of BraTS 2020 and BraTS 2021 to evaluate the proposed SGEResU-Net model. The BraTS 2020 dataset contains a training dataset of 369 glioma patients, a validation dataset consisting of 125 patient cases, and a test dataset including 166 samples, respectively. Compared with the BraTS 2020 dataset, the BraTS 2021 dataset largely increases brain tumor patient cases from 660 to 2,000. More specifically, it includes the training dataset of 1,251 cases, the validation dataset of 219 cases, and the test dataset of 530 cases, respectively [17–22]. For the two benchmarks, ground truths of training datasets are provided by the BraTS organizers. However, labels of validation and test datasets are unavailable to the public. In order to ensure the fairness and accuracy of the experimental results, the evaluation results of validation and test datasets

can be obtained only via the BraTS online website. Besides, each patient case includes four modalities describing with Flair, T1, T1ce and T2, and the size of each image is $240 \times 240 \times 155$ pixels. The basic labels include four types, which are called healthy parts (label 0), necrotic and non-enhancing tumors (label 1), edema around the tumor (label 2) and GD enhancing tumors (label 4). The various sub-regions considered for the segmentation evaluation are the whole tumor (the combined area of labels 1, 2 and 4), the core tumor (the combined area of labels 1, 4), and the enhancing tumor (the area of label 4). Figure 4 shows some typical images of a patient case in the BraTS 2021 training dataset.
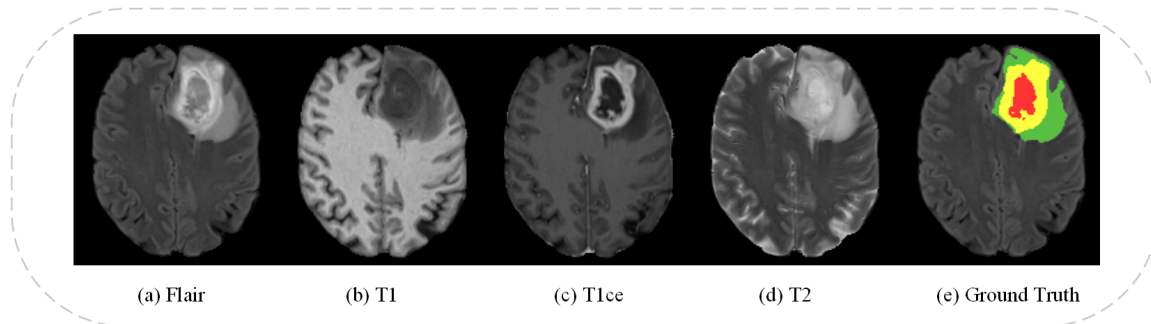

(a) Flair   (b) T1   (c) T1ce   (d) T2   (e) Ground Truth

Figure 4. Typical MRI brain images in the BraTS 2021 training dataset. From left to right: (a) Flair, (b) T1, (c) T1ce, (d) T2 and (e) Ground Truth. Each color represents a tumor class, i.e., red: necrotic and non-enhancing; yellow: enhancing tumor; green: edema.

Since the BraTS 2020 and 2021 datasets are collected in multiple institutions with different scanners and protocols, which makes the intensity values not standardized. Thus, we utilize the z-score method to standardize non-standardized brain tumors, which applies average and standard deviation to process each image. The detailed calculation formula is as follows:

$$z' = \frac{z - \mu}{\delta} \tag{3.1}$$

where z is the input image, z' is the normalized image. μ and δ denote the mean value and standard deviation of the input image, respectively. Besides, as the brain tumor image contains a lot of useless background information, all images are cropped into $128 \times 128 \times 128$ pixels as input. In addition, a variety of data transformation strategies such as random rotation, random scaling, random elastic deformation, random flip and random intensity changes are utilized to preprocess the data.

3.2. Evaluation metrics and settings

In this paper, four commonly used metrics including Dice score/ Dice similarity coefficient (DSC), Hausdorff distance 95%, Sensitivity and Specificity are adopted to evaluate the segmentation network, and they are defined as follows:

$$Dice\ score = \frac{2TP}{FP + 2TP + FN} \tag{3.2}$$

$$Sensitivity = \frac{TP}{TP + FN}, \quad Specificity = \frac{TN}{TN + FP} \tag{3.3}$$

where TP, FP, TN and FN denote the number of false negative, true negative, true positive and false positive voxels, respectively. The DSC value, as a metric of ensemble similarity, is generally used to calculate the similarity of two samples. The value range is 0–1, and values close to 1 indicate more similar contours. Sensitivity and Specificity measure voxel-wise overlap between the predicted results and the ground-truth.

$$HD(T, P) = max\left\{sup_{t \in T}\inf_{p \in P}d\left(t, p\right), sup_{p \in P}\inf_{t \in T}d\left(t, p\right)\right\} \qquad (3.4)$$

where t and p represent the points on the the ground-truth regions T and the predicted regions P, respectively. $d\left(t, p\right)$ is the function that computes the distance between points t and p. Hausdorff95 is a metric of using Hausdorff distance to measure the 95% quantile of the surface distance. Since Dice score and Hausdorff95 are the overall evaluation metrics for the entire BraTS challenges, we adopt them as the key metric for evaluation consistently across all the challenges.

In our experiments, we utilize the Adam optimizer with an initial learning rate of 0.001. Momentum is 0.95, the weight decay is $1e^{-5}$, and the batch size is set to 4. Our model is implemented using the PyTorch deep learning framework on an NVIDIA GeForce GTX 3090 GPU with 24 GB of memory.

4. Results and discussion

To verify the effectiveness of the embedded SGE and Residual modules, we firstly conduct basic experiments on the BraTS 2020 dataset, and then compare the SGEResU-Net model with the representative brain tumor segmentation networks. Finally, we focus on experiments performed on the BraTS 2021 dataset to evaluate SGEResU-Net, which contains more brain tumor samples and may provide more stable results.

4.1. Compared experiment results on the BraTS 2020 benchmark

Comparisons with the baseline on the BraTS 2020 training dataset. Firstly, we perform compared experiment with the baseline model of 3D U-Net on the BraTS 2020 training dataset to evaluate SGEResU-Net. We divide the 2020 training dataset into two data sets with a ratio of 8:2 for training and validation, and the five-fold cross-validation strategy is adopted to achieve more stable results. The compared results are listed in Table 1.

Table 1. DSC results on the BraTS 2020 training dataset.

Method	Enhancing	Whole	Core
3D U-Net (baseline)	0.7851	0.9010	0.8297
SGEResU-Net	0.7940	0.9048	0.8522

As can be seen in Table 1, our SGEResU-Net can achieve the better segmentation performance than its baseline model, and it obtains DSC of 0.7940, 0.9048 and 0.852 on the enhancing tumor, whole tumor and core tumor segmentation, respectively. It is superior to the U-Net with 0.89, 0.38 and 2.25% accuracy improvement on the enhancing tumor, whole tumor and core tumor, respectively. This experiment demonstrates the effectiveness of SGEResU-Net in the segmentation of brain tumors to a certain extent.

Table 2. Ablation DSC values on the BraTS 2020 validation dataset.

Methods	DSC			Hausdorff95		
	Enhancing	Whole	Core	Enhancing	Whole	Core
U-Net	0.7686	0.8938	0.7992	32.56	7.70	12.11
ResU-Net	0.7761	0.8979	0.8254	35.07	7.21	11.42
SGEU-Net	0.7748	0.8991	0.8104	30.17	7.39	12.06
SGEResU-Net	0.7742	0.9026	0.8306	29.80	6.95	11.60

Ablation experiments on the BraTS 2020 validation dataset. Then, we perform ablation experiments to further test the effects of SGE module and Residual module on the BraTS 2020 validation dataset. Here, we utilize the whole BraTs 2020 training dataset of 369 MRI images to train segmentation models, and then submit 125 predicted results of the validation dataset via BraTS online website to achieve the final DSC and Hausdorff95 values. The ablation results on the BraTS 2020 validation dataset are reported in Table 2.

As shown in Table 2, the baseline model of U-Net achieves DSC values of 0.7686, 0.8938 and 0.7992 on enhancing tumor, whole tumor, and core tumor segmentation, respectively. After embedding Residual blocks, it accordingly gains 0.75, 0.41 and 2.62% accuracy improvements on these three regions, which demonstrates the effects of Residual block on brain tumor segmentation task. When it comes to the SGE module, we incorporate it into four skip connections for capturing long-range context features of brain tumor images. As shown in the table, SGEU-Net also gains better segmentation results than its baseline, and its DSC values of 0.7748, 0.8991 and 0.8104 on enhancing tumor, whole tumor, and core tumor segmentation, respectively. Finally, we simultaneously embed Residual blocks and SGE modules into a single 3D U-Net architecture, and construct the final segmentation model of SGEResU-Net. And it achieves DSC values of 0.7742, 0.9026 and 0.8306 on enhancing tumor, whole tumor, and core tumor segmentation, respectively. Compared to its baseline, it accordingly gains 0.56, 0.88 and 3.14% accuracy improvements on these three regions, which further demonstrates the effects of integrating Residual blocks and SGE modules on brain tumor segmentation task. Therefore, ablation experiment results well prove that the effectiveness of SGEResU-Net for the brain tumor segmentation task.

Comparisons with the state-of-the-arts on the BraTS 2020 validation dataset. Additionally, we also perform compared experiments with some state-of-the-art models on the BraTS 2020 validation dataset, whose results are listed in Table 3. In this table, SGEResU-Net achieves DSC values of 77.4, 90.3 and 83.1% on the enhancing tumor, whole tumor, and tumor core, respectively. Especially, our SGEResU-Net gets the optimal result on the whole tumor segmentation, and it outperforms by 1.3% accuracy gains over the second highest result of 0.890 [26, 27] on the whole tumor segmentation, which further demonstrates that our proposed model is effective for the segmentation of brain tumors. On the enhancing tumor segmentation, Wang et al. [26] obtain the optimal DSC value of 0.785 by introducing the self-attention mechanism into the segmentation model of deep learning. Compared with Sundaresan et al. [27], Zhang et al. [30] and Huang et al. [29], the DSC values obtained by our method are 0.4, 0.4 and 7.4% higher on enhancing tumor, respectively. Meanwhile, Zhang et al. [30] obtain the optimal DSC values of 0.839 on core tumor, and our SGEResU-Net is slightly lower than it on core tumor segmentation. They simultaneously integrate asymmetric convolution

Table 3. Compared results with the state-of-the-art methods on the BraTS 2020 validation dataset.

Methods	DSC			Hausdorff95		
	Enhancing	Whole	Core	Enhancing	Whole	Core
Tang et al. [23]	0.698	0.889	0.784	34.3	4.5	10.1
Cheng et al. [24]	0.780	0.894	0.814	24.4	7.1	12.7
Zhang W et al. [25]	0.702	0.883	0.739	38.6	7.0	30.2
Wang et al. [26]	0.785	0.890	0.814	16.7	6.5	10.5
Sundaresan et al. [27]	0.770	0.890	0.770	29.4	4.4	15.3
Fang et al. [28]	0.670	0.870	0.769	50.8	9.4	12.5
Huang et al. [29]	0.700	0.860	0.772	39.1	6.7	15.1
Zhang J et al. [30]	0.770	0.896	0.839	32.4	7.7	11.7
SGEResU-Net (ours)	0.774	0.903	0.831	29.8	7.0	11.6

blocks and an expectation-maximization attention module into the DMF-Net architecture to capture more powerful deep features for the brain tumor segmentation, which may lead to higher DSC values on core tumor segmentation. Note that SGEResU-Net outperforms by 1.7% accuracy gains over the second highest result of 0.814 [24, 26] on core tumor segmentation. When it comes to Hausdorff95 results, SGEResU-Net gains Hausdorff95 distances are 29.8, 7.0 and 11.6 on the enhancing tumor, whole tumor, and tumor core, respectively. Compared with latest counterpart models, SGEResU-Net is only inferior to Sundaresan et al. [27] on enhancing tumor, and Wang et al. [26] on whole tumor and core tumor. It consistently has a better segmentation performance than models given by Fang et al. [28] and Zhang et al. [30]. Nevertheless, SGEResU-Net still ranks the third and fourth places on the core tumor and enhancing tumor segmentation considering the Hausdorff95 metric. In general, the compared experiment results demonstrate the competitive performance of SGEResU-Net and again prove its effectiveness for segmenting brain tumors.

4.2. Compared experiment results on the BraTS 2021 benchmark

Comparisons with the baseline on the BraTS 2021 training dataset. In this experiment, we utilize 1251 labeled cases from BraTS 2021 training dataset to randomly split the data for training and validation with a ratio of 8:2, and the five-fold cross-validation is also adopted. The experiment results are reported in Table 4. As shown in Table 4, SGEResU-Net achieves DSC values of 0.8790,

Table 4. DSC results on the BraTS 2021 training dataset.

Methods	Enhancing	Whole	Core
3D U-Net (baseline)	0.8744	0.9324	0.9162
SGEResU-Net	0.8790	0.9365	0.9219

0.9365 and 0.9219 on enhancing, whole and core tumor segmentation, respectively. The average DSC on three segmentation models is 0.9125. These results well show the effectiveness of SGEResU-Net for brain tumor segmentation. In addition, we also provide some visualization results of SGEResU-
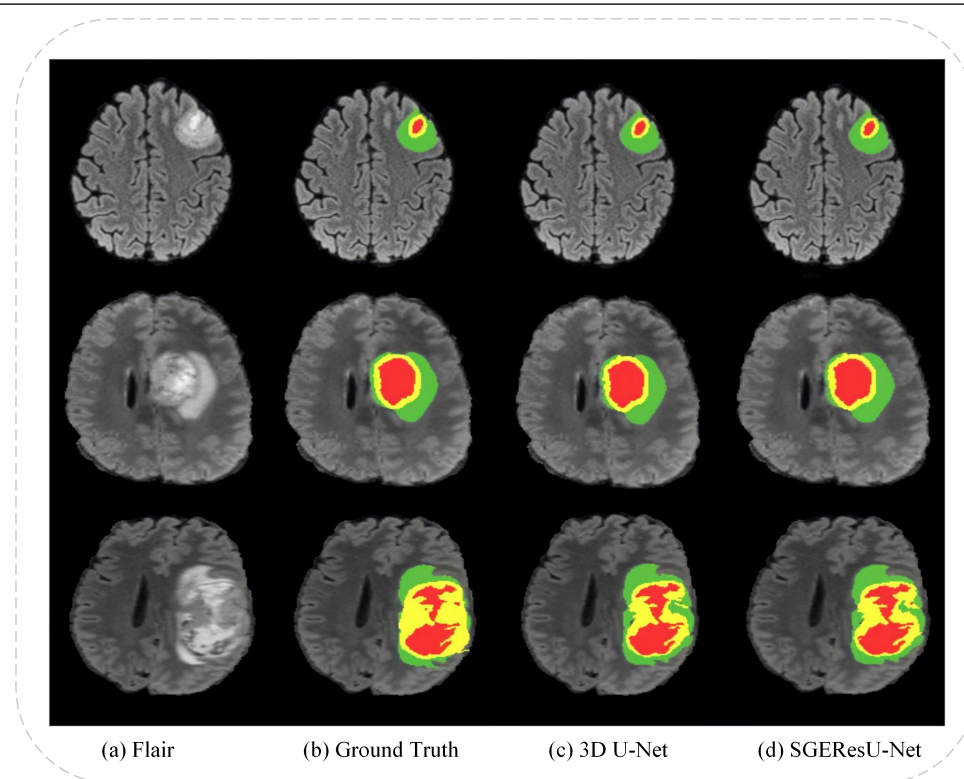
Figure 5. Example segmentation results on the BraTS 2021 Training Dataset. From left to right: (a) Flair, (b) Ground Truth, (c) 3D U-Net and (d) SGEResU-Net results overlaid on Flair image. Each color represents a tumor class, i.e., red: necrosis and non-enhancing; yellow: enhancing tumor; green: edema.

Net in Figure 5. Different colors represent the different types of tumors, i.e., the red regions are necrosis and non-enhancing, the yellow regions represent enhancing tumor, the green regions indicate edema. Meanwhile, images from left to right are Flair, Ground Truth, 3D U-Net and SGEResU-Net segmentation results overlaid on flair image, respectively. As shown in this figure, SGEResU-Net can well segment the enhancing, whole and core tumor regions.

Experiment results on the BraTS 2021 validation dataset. In this experiment, we leverage the whole BraTS 2021 training dataset of 1251 MRI images to train the brain tumor segmentation model. Then, it is used to segmentation tumor images in the BraTS 2021 validation dataset, whose results will be uploaded to the official website platform to obtain the final validation dataset results. We tabulate the final evaluation results of SGEResU-Net on the BraTS 2021 validation dataset in Table 5. Quantitatively, SGEResU-Net achieves DSC values of 83.31, 91.64 and 86.85%, as well as Hausdorff distances (95%) of 19.278, 5.945 and 7.567 for the enhancing tumor, whole tumor, and tumor core, respectively. These results further demonstrate the effectiveness of SGEResU-Net for brain tumor segmentation. In addition, this table also lists the mean, standard deviation, median, and 25th and 75th percentile of each metric. Finally, Figure 6 shows the box plot of DSC values obtained on the BraTS 2021 validation dataset. These results to a certain prove the effects of our SGEResU-Net model on this medical image segmentation application.

Table 5. Evaluation results of SGEResU-Net on the BraTS 2021 validation dataset.

Metric	Tumor type	Mean	StdDev	Median	25quantile	75quantile
	Enhancing	0.8331	0.2332	0.8963	0.8343	0.9439
DSC	Whole	0.9164	0.0981	0.9414	0.8985	0.9651
	Core	0.8685	0.1879	0.9353	0.8599	0.9634
	Enhancing	0.8296	0.2236	0.9021	0.8116	0.9520
Sensitivity	Whole	0.9292	0.1036	0.9630	0.9183	0.9833
	Core	0.8609	0.1834	0.9224	0.8451	0.9650
	Enhancing	0.9998	0.0004	0.9999	0.9998	0.9999
Specificity	Whole	0.9992	0.0008	0.9995	0.9989	0.9997
	Core	0.9997	0.0006	0.9999	0.9998	0.9999
	Enhancing	19.278	77.467	1.414	1.000	2.343
Hausdorff95	Whole	5.945	25.991	2.236	1.414	4.123
	Core	7.567	35.750	2.000	1.000	4.062
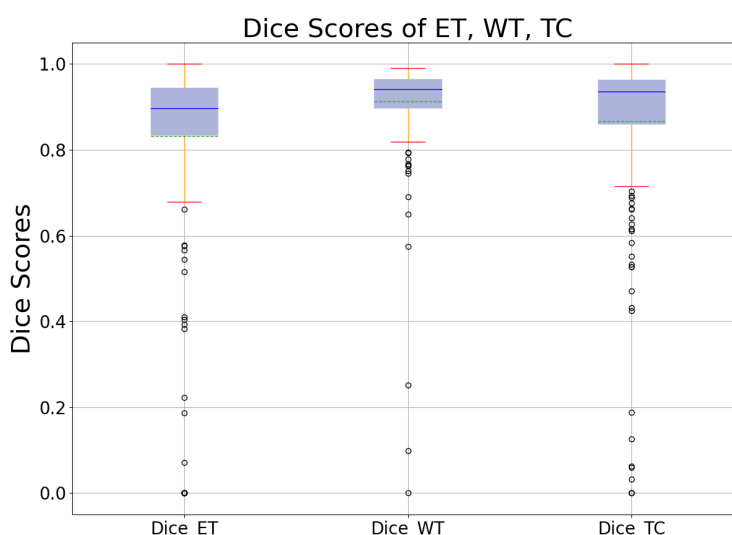
Figure 6. Box plot of DSC scores achieved on the BraTS 2021 validation dataset.

5. Conclusions

In this work, we propose a novel 3D U-Net model called SGEResU-Net to segment brain tumors. SGEResU-Net replaces the horizontally linked part of the baseline network with an improved 3D SGE module to adpat the complex feature distribution of brain tumor images. The 3D SGE module can learn sub-features and suppress noise in a targeted manner for each group, and is also a lightweight module that can adapt brain tumor segmentation models without stress. Experiment results on the Brain Tumor Segmentation Challenge 2020 and 2021 benchmarks demonstrate its effectiveness for the MRI brain tumor segmentation task. In the future, we will evaluate SGEResU-Net on other typical medical image

segmentation applications. Additionally, we will also exploit where the SGE module is more beneficial for brain tumor segmentation performance in 3D U-Net networks.

Acknowledgments

Conflict of interest

The authors have no conflict of interest.

References

1. J. Liu, M. Li, J. Wang, F. Wu, Y. Pan, A survey of MRI-based brain tumor segmentation methods, *Tsinghua Sci. Technol.*, **19** (2014), 578–595. https://doi.org/1007-0214-19-6-578

2. S. Bauer, R. Wiest, L. P. Nolte, M. Reyes, A survey of MRI-based medical image analysis for brain tumor studies, *Phys. Med. Biol.*, **58** (2013), R97. https://doi.org/10.1088/0031-9155/58/13/R97

3. O. Ronneberger, P. Fischer, T. Brox, U-Net: convolutional networks for biomedical image segmentation, in *Medical Image Computing and Computer-Assisted Intervention - MICCAI 2015-18th International Conference Munich*, Lecture Notes in Computer Science, Springer, (2015), 234–241. https://doi.org/10.1007/978-3-319-24574-4_28

4. A. Kermi, I. Mahmoudi, M. T. Khadir, Deep convolutional neural networks using U-Net for automatic brain tumor segmentation in multimodal MRI volumes, in *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries - 4th International Workshop*, Springer, (2019), 37–48. https://doi.org/10.1007/978-3-030-11726-9_4

5. J. X. Zhang, Z. K. Jiang, J. Dong, Y. Q. Hou, B, Liu, Attention gate ResU-Net for automatic MRI brain tumor segmentation, *IEEE Access*, **8** (2020), 58533–58545. https://doi.org/10.1109/ACCESS.2020.2983075

6. A. Albiol, A. Albiol, F. Albiol, Extending 2D deep learning architectures to 3D image segmentation problems, in *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries - 4th International Workshop*, Springer, (2019), 73–82. https://doi.org/10.1007/978-3-030-11726-9_7

7. H. Jia, W. Cai, H. Huang, Y. Xia, H2NF-Net for brain tumor segmentation using multimodal MR imaging: 2nd place solution to BraTS Challenge 2020 Segmentation Task, in *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries - 6th International Workshop*, Springer, (2021), 58–68. https://doi.org/10.1007/978-3-030-72087-2_6

8. X. Zhang, W. Jian, K. Cheng, 3D dense U-nets for brain tumor segmentation, in *Pre-Conference Proceedings of the 7th MICCAI BraTS Challenge*, (2018), 562–570.

9. P. Liu, Q. Dou, Q. Wang , P. A. Heng, An encoder-decoder neural network with 3D squeeze-and-excitation and deep supervision for brain tumor segmentation, *IEEE Access*, **8** (2020), 34029–34037. https://doi.org/10.1109/ACCESS.2020.2973707

10. A. Myronenko, 3D MRI brain tumor segmentation using autoencoder regularization, in *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries - 4th International Workshop*, Springer, (2019), 311–320. https://doi.org/10.1007/978-3-030-11726-9_28

11. Y. Zhao, Y. Zhang, C. Liu, Bag of tricks for 3D MRI brain tumor segmentation, in *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries - 5th International Workshop*, Springer, (2020), 210–220. https://doi.org/10.1007/978-3-030-46640-4_20

12. Z. Jiang, C. Ding, M. Liu, Two-Stage cascaded U-Net: 1st place solution to BraTS challenge 2019 segmentation task, in *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries - 5th International Workshop*, Springer, (2019), 231–241. https://doi.org/10.1007/978-3-030-46640-4_22

13. F. Isensee, P. F. Jager, P. M. Full, P. Vollmuth, K. H. Maier-Hein, nnU-Net for brain tumor segmentation. in *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries - 6th International Workshop*, Springer, (2021), 118–132. https://doi.org/10.1007/978-3-030-72087-2_11

14. F. Isensee, P. Kickingereder, W. Wick, M. Bendszus, K. H. Maier-Hein, No new-net, in *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries - 4th International Workshop*, Springer, (2019), 234–244. https://doi.org/10.1007/978-3-030-11726-9_21

15. X. Li, X. L. Hu, J. Yang Li, Spatial group-wise enhance: Improving semantic feature learning in convolutional networks, preprint, arXiv: 1905.09646.

16. K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, in *2016 IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2016*, IEEE Computer Society, (2016), 770–778. https://doi.org/10.1109/CVPR.2016.90

17. U. Baid, S. Ghodasara, S. Mohan, M. Bilello, E. Calabrese, E. Colak, et al., The RSNA-ASNR-MICCAI BraTS 2021 Benchmark on brain tumor segmentation and radiogenomic classification, preprint, arXiv: 2107.02314.

18. B. H. Menze, A. Jakab, S. Bauer, J. Kalpathy-Cramer, K. Farahani, J. Kirby, et al., The multimodal brain tumor image segmentation benchmark (BraTS), *IEEE Trans. Med. Imaging*, **34** (2015), 1993–2024. https://doi.org/10.1109/TMI.2014.2377694

19. S. Bakas, H. Akbari, A. Sotiras, M. Bilello, M. Rozycki, J. S. Kirby, et al, Advancing the cancer genome atlas glioma MRI collections with expert segmentation labels and radiomic features, *Sci. Data*, **4** (2017), 170117. https://doi.org/10.1038/sdata.2017.117

20. S. Bakas, H. Akbari, A. Sotiras, M. Bilello, M. Rozycki, J. S. Kirby, et al, Segmentation labels and radiomic features for the preoperative scans of the TCGAGBM collection, *Cancer Imaging Arch.*, 2017. https://doi.org/10.7937/K9/TCIA.2017.KLXWJJ1Q

21. S. Bakas, H. Akbari, A. Sotiras, M. Bilello, M. Rozycki, J. S. Kirby, et al, Segmentation labels and radiomic features for the preoperative scans of the TCGALGG collection, *Cancer Imaging Arch.*, 2017. https://doi.org/10.7937/K9/TCIA.2017.GJQ7R0EF

22. 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, preprint, arXiv: 1811.02629.

23. J. Tang, T. Li, H. Shu, H. Zhu, Variational-Autoencoder regularized 3D MultiResUNet for the BraTS 2020 brain tumor segmentation, in *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries - 6th International Workshop*, Springer, (2021), 431–440. https://doi.org/10.1007/978-3-030-72087-2_38

24. K. Cheng, C. Hu, P. Yin, et al. Glioma sub-region segmentation on Multi-parameter MRI with label dropout, in *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries - 6th International Workshop*, Springer, (2021), 420–430. https://doi.org/10.1007/978-3-030-72087-2_37

25. W. B. Zhang, G. Yang, H. Huang, W. J. Yang, X. M. Xu, Y. K. Liu, et al., ME-Net: Multi-encoder net framework for brain tumor segmentation. *Int. J. Imag. Syst. Tech.*, **31** (2021), 1834–1848. https://doi.org/10.1002/ima.22571

26. W. Wang, C. Chen, M. Ding, H. Yu, S. Zha, J. Li, TransBTS: Multimodal brain tumor segmentation using transformer, in *Medical Image Computing and Computer Assisted Intervention – MICCAI 2021, 24th International Conference*, Springer, (2021), 109–119. https://doi.org/10.1007/978-3-030-87193-2_11

27. V. Sundaresan, L. Griffanti, M. Jenkinson, Brain tumour segmentation using a triplanar ensemble of U-Nets on MR images, in *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries - 6th International Workshop*, Springer, (2021), 340–353. https://doi.org/10.1007/978-3-030-72084-1_31

28. Y. Fang, H. Huang, W. J. Yang, X. M. Xu, W. W. Jiang, X. B. Lai, Nonlocal convolutional block attention module VNet for gliomas automatic segmentation, *Int. J. Imag. Syst. Tech.*, **32** (2022), 528–543. https://doi.org/10.1002/ima.22639

29. H. Huang, G. Yang, W. B. Zhang, X. M. Xu, W. J. Yang, W. W. Jiang, et al., A deep multi-task learning framework for brain tumor segmentation, *Front Oncol.*, **11** (2021), 690244. https://doi.org/10.3389/fonc.2021.690244

30. J. X. Zhang , Z. K. Jiang , D. W. Liu , Q. L. Sun, Y. Q. Hou, B. Liu, 3D asymmetric expectation-maximization attention network for brain tumor segmentation. *NMR Biomd.*, (2021), e4657. https://doi.org/10.1002/nbm.4657