

Research Article

A Deep-Learning Model with Learnable Group Convolution and Deep Supervision for Brain Tumor Segmentation

Hengxin Liu,¹ Qiang Li,¹ and I-Chi Wang ²

¹School of Microelectronics, Tianjin University, Tianjin 300072, China

²Department of Industrial Education and Technology, National Changhua University of Education, Changhua 50007, Taiwan

Correspondence should be addressed to I-Chi Wang; perfect_651208@yahoo.com.tw

Received 6 November 2020; Revised 26 December 2020; Accepted 28 January 2021; Published 10 February 2021

Academic Editor: Kim-Hua Tan

Copyright © 2021 Hengxin Liu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The segmentation of brain tumors in medical images is a crucial step of clinical treatment. Manual segmentation is time consuming and labor intensive, and existing automatic segmentation methods suffer from issues such as numerous parameters and low precision. To resolve these issues, this study proposes a learnable group convolution-based segmentation method that replaces convolution in the feature extraction stage with learnable group convolution, thereby reducing the number of convolutional network parameters and enhancing communication between convolution groups. To improve utilization of the feature maps, we added a skip connection structure between learnable group convolution modules, which increased segmentation precision. We used deep supervision to combine output images in the network output stage to reduce overfitting and enhance the recognition capabilities of the network. We tested the proposed algorithm model using the open BraTS 2018 dataset. The experiment results revealed that the proposed model is superior to 3D U-Net and DMFNet and has better segmentation results for tumor cores than No New-Net and NVDLMED, the winning methods in the BraTS 2018 challenge. The segmentation precision of the proposed method with regard to whole tumors, enhancing tumors, and tumor cores was 90.25%, 80.36%, and 86.20%. Furthermore, the proposed method uses fewer parameters and a less complex model.

1. Introduction

Early diagnosis is crucial for the surgical treatment of brain tumors. This has been aided by recent advances in medical imaging technology. Magnetic resonance imaging (MRI) technology can display brain tissue information in great detail and is widely used for the diagnosis of brain tumors. Four types of MRI modes are used: T1 weighted, T2 weighted, postcontrast T1 weighted, and fluid-attenuated inversion recovery (FLAIR). Each of these reflects different aspects of brain tissue. T1-weighted scans highlight tumor contours, T2-weighted scans show distinct tumor regions, and FLAIR scans can distinguish edema from cerebrospinal fluid.

The accurate segmentation of brain tumors in medical images is a critical step before treatment. Manual segmentation is time consuming and labor intensive, and as a result, efficient and accurate automatic segmentation

methods have become a popular research topic in recent years. Brain tumor segmentation methods can generally be divided into three categories: manual segmentation, semi-automatic segmentation, and fully-automatic segmentation. The semiautomatic and fully-automatic methods can be further divided into two categories: unsupervised segmentation and supervised segmentation [1]. Depending on the segmentation principle, unsupervised segmentation includes threshold-based segmentation [2–4], region-based segmentation [5–9], graphic-element classification-based segmentation [10–13], and model-based segmentation [14, 15].

The disadvantages of unsupervised methods are that they require a confirmed number of segmentation regions in advance and the MRI images must first undergo intensity nonuniformity correction and skull stripping. Supervised methods are based on graphic-element classification, including conventional machine learning and convolutional

neural networks (CNNs). Segmentation methods using conventional machine learning include support vector machines [16–22], conditional random fields (CRFs) [23, 24], and random forests (RFs) [25, 26]. In conventional machine learning methods, the features must be manually selected, in which boundary and tumor region details can be easily overlooked.

CNN-based methods include CNN models, fully convolutional neural network (FCNN) models, and U-Net models. CNN models include the CNN structure with small kernels proposed by Pereira [27] and the cascade CNN model proposed by Havaii [28]. FCNN-based models include the residual module-containing FCNN model structure presented by Chen [29] and the model structure integrating FCNNs and CRFs proposed by Zhao [30]. Cicek [31–33] examined 3D convolution operations, upgraded U-Net from 2D to 3D, and proposed a 3D U-Net for the segmentation of 3D medical images. Models based on U-Net include the 3D U-Net structure used by Sherman [34], in which residual structures were added between convolutions in the same layer. Nuechterlein and Mehta [35] developed 3D-ESPNet, which applies the pointwise convolution of semantic segmentation to medical image processing to reduce the number of network parameters; however, the resulting precision of segmentation is lower. Kao et al. [36] employed an ensemble comprising seven 3D U-Nets with different parameters and training strategies for brain tumor image segmentation; the higher number of models resulted in longer training time. In the BraTS 2018 challenge, Isensee et al. [37] made minor structural modifications to a 3D U-Net and obtained No New-Net. With additional training data and a simple postprocessing technique, this approach won the second place in the challenge. Myronenko [38] proposed an encoder-decoder architecture network, NVDLMED, added another decoder pathway to recover input images, and imposed additional constraints. This approach won the first place in the BraTS 2018 challenge.

CNN-based methods all involve a large amount of computation and highly complex models and room for improvement in segmentation precision remains. To reduce the number of segmentation network parameters, Chen et al. [39] proposed a dilated multifiber network (DMFNet), which replaces regular convolutions with group convolutions, greatly reducing the number of parameters while maintaining the precision of the segmentation network. Group convolutions were first proposed for Alexnet [40] and were later successfully applied in ResNeXt [41]; they are currently popular in network design. However, in standard group convolutions, each group processes information independently, and there is no communication between groups, which limits their feature representation capabilities. Zhang et al. [42] presented a dynamic group convolution, which can learn different numbers of convolution groups in training data and improves the information flow between groups, thereby achieving better performance than regular group convolutions.

Skip connections can accelerate network convergence and increase the precision of the segmentation network. Deep convolutional neural networks exhibit better

performance than shallow networks, but the gradient vanishing problem may apply. Residual connections were, thus, introduced to ResNet [43] to solve this degenerative issue. DenseNet [44] presented densely connected layers with more shortcut connections and used the cascade strategy to combine the feature maps of the first few layers. Residual connections and dense connections all use information from the previous convolutional layers and are added to networks in the form of skip connections. DenseNet achieves better performance, but as the number of input channels increases, the network consumes more memory.

The idea underlying deep supervision is to directly supervise the hidden layer rather than just pay attention to the output layer. In GoogLeNet [45], supervising the two hidden layers of a 22-layer network achieves better effects. Dou et al. [46] applied deep supervision to segment 3D liver CT scans. After the features of the lower and middle layers were deconvoluted in the convolutional network, they were then combined with the output layer, reducing training and verification errors and granting the network better convergence. Chen et al. [47] utilized three classifiers to classify features in intermediate layers. The outputs of the classifiers serve as moderators during training, and the network combines multilevel contextual information for deep supervision, thereby enhancing the recognition capabilities of the network.

Regarding the problems of too many parameters and low segmentation precision in conventional CNNs, we modified DMFNet, a group convolution network with a smaller number of parameters. To enhance the communication between groups in DMFNet, we replaced the regular group convolution with learnable group convolution. To make full use of the feature maps, we added a skip connection structure between the learnable group convolution modules and introduced deep supervision to merge output images in the network output stage, thereby enhancing the segmentation precision of the network. We refer to this lightweight brain tumor image segmentation algorithm with learnable grouping as DLSDNet.

2. Lightweight Brain Tumor Image Segmentation Network with Learnable Grouping

We modified DMFNet for our segmentation network, replacing the regular group convolution in the feature extraction stage of DMFNet with learnable group convolution. We also added a skip connection and introduced deep supervision to the network output stage to merge the network outputs.

2.1. DMFNet. DMFNet comprises lightweight 3D convolutional neural networks. The network structure is similar to that of U-Net. It divides complex neural networks into lightweight networks or fiber sets, replaces regular convolution with group convolution, and uses a multiplexer to exchange information.

The multiplexer consists of two $1 \times 1 \times 1$ convolutions to promote information flow between fibers [39]. To expand the receptive field and capture the multiscale 3D spatial correlations of the brain tumors, we added dilated group convolution to the fiber units in the encoder stage. The feature extraction stage involves multiple dilated multifiber units (DMFunits). In the output stage, we used a regular group convolution multifiber unit (MFunit), as shown in Figure 1.

2.2. Learnable Grouping. To further enhance communication for group convolution and the feature extraction capabilities of convolutions, we replaced the group convolution in DMFNet with learnable group convolution (LGConv).

We suppose a convolutional feature map is $F \in \mathbb{R}^{N \times C^{in} \times H \times W}$, where N , C , H , and W , respectively, denote the number of samples, number of channels, and the height and width of the channels in the small batches. If a convolution with $k \times k$ kernels is applied to F , the output feature map is $O \in \mathbb{R}^{N \times C^{out} \times H \times W}$, where each unit output is $o_{ij} \in \mathbb{R}^{N \times C^{out}}$. Learnable group convolution (LGConv) can be defined as follows:

$$o_{ij} = \sum_{m=0}^{k-1} \sum_{n=0}^{k-1} f_{(i+m)(j+n)} (U \odot \omega_{mn}), \quad (1)$$

where $i \in \{1, \dots, H\}$, $j \in \{1, \dots, W\}$; $f_{(i+m)(j+n)} \in \mathbb{R}^{N \times C^{in}}$ represents the hidden unit F of the input feature map, $\omega_{mn} \in \mathbb{R}^{C^{in} \times C^{out}}$ denotes the convolution weight, and \odot indicates the element-wise product.

LGConv is an expansion of group convolution. It can use the binary relation matrix $U \in \{0, 1\}^{C^{in} \times C^{out}}$ to learn group principles. Many convolutions can be regarded as unique forms of learnable grouping.

Let $U = 1$, which gives $1 \odot \omega_{mn} = \omega_{mn}$ and represents a regular convolution, as shown in Figure 2(a). Let $U = I$, where I is a unit matrix. $I \odot \omega_{mn}$ then becomes a matrix where the diagonal elements are 1 and the nondiagonal elements are 0, as shown in Figure 2(b), indicating that each channel is independent. Thus, LGConv is a depth-wise separable convolution [48]. If U is a binary block diagonal matrix as shown in Figure 2(c), then $U \odot \omega_{mn}$ divides the channels into groups. If U is a constant matrix where all of the diagonal elements are 1, then LGConv represents regular group convolution in which adjacent channels are grouped together. If U is any given binary matrix as shown in Figure 2(d), it will result in unstructured convolution. Thus, appropriately constructing binary relation matrix U can produce various convolutional operations.

To reduce the complexity of U , we decompose it into K submatrices:

$$\{U_k | U_k \in \{0, 1\}^{C_k^{in} \times C_k^{out}}, \forall C_k^{in} < C^{in}, \forall C_k^{out} < C^{out}\}. \quad (2)$$

The shape of submatrix U_k is $C_k^{in} \times C_k^{out}$, where $C_k^{in} < C^{in}$ and $C_k^{out} < C^{out}$. We define U as

$$U = U_1 \otimes U_2 \otimes \dots \otimes U_K, \quad (3)$$

where \otimes indicates the Kronecker product. Thus, $\prod_{k=1}^K C_k^{in} = C^{in}$ and $\prod_{k=1}^K C_k^{out} = C^{out}$. Through a series of Kronecker products, we can decompose matrix U into a set of submatrices [49].

To construct each submatrix U_k , let $C^{in} = C^{out}$. This is a general setting in ResNet and ResNeXt. To reduce the parameters in the convolutional operations, we use a single binary variable to express U_k :

$$\begin{cases} U_k = g_k 1 + (1 - g_k) I, & \forall g_k \in g, \\ g = \text{sign}(\tilde{g}), \end{cases} \quad (4)$$

where 1 denotes a 2×2 constant matrix where the elements are 1, I represents a 2×2 unit matrix, g_k is the k th component, $\tilde{g} \in \mathbb{R}^K$ is a learnable gate vector of continuous values, $g \in \{0, 1\}^K$ is a binary gate vector output from \tilde{g} , and $\text{sign}(\cdot)$ denotes a sign function:

$$\text{sign}(x) = \begin{cases} 0, & x < 0, \\ 1, & x \geq 0. \end{cases} \quad (5)$$

We can combine equation (4) with equation (3) as follows:

$$U = (g_1 1 + (1 - g_1) I) \otimes \dots \otimes (g_K 1 + (1 - g_K) I). \quad (6)$$

With such a structural relationship, the parameter in need of optimization becomes \tilde{g} , so the number of parameters in U reduces from $C^{in} \cdot C^{out}$ to $\log_2 C^{in}$. U is a unit matrix where all of the diagonal elements are 1. When $K = 3$ and $g_1 = 1, g_2 = 1, g_3 = 0$, equation (6) becomes $1 \otimes 1 \otimes I$, which is an 8×8 matrix with two groups, as shown in Figure 3(a). When $g_1 = 0, g_2 = 1, g_3 = 0$, equation (6) becomes $I \otimes 1 \otimes I$, which is an 8×8 matrix with four groups, as shown in Figure 3(b). It is shown from that mentioned above that LGConv can group nonadjacent channels. In this way, only three continuous parameters \tilde{g}_1, \tilde{g}_2 , and \tilde{g}_3 are needed to generate g_1, g_2 , and g_3 and learn the original large 8×8 matrix in which 64 parameters require training.

This study replaced the regular group convolution in the network with learnable group convolution to enhance the agility of the network and increase the precision of the segmentation network. The network units following the replacement are as shown in Figure 4.

2.3. Skip Connection Unit. This study proposed a novel skip connection unit to extract early feature maps and enhance feature reuse. Features are transferred between key layers so that the radiation range of early features expands to deeper levels, thereby enhancing the global integration of information flow. The neural network mapping function of the novel skip connection can be expressed as follows:

$$x_n = F_n([x_{n_i}, x_{n_{ii}}, \dots, x_{n_{i \dots ii}}]), \quad (7)$$

where $F_n(\cdot)$ is the nonlinear transform after each level and n indicates level n . The output of layer n expressed as x_n . $[x_{n_i}, x_{n_{ii}}, \dots, x_{n_{i \dots ii}}]$ refers to the cascade of feature maps generated by the selected layers $n_i, n_{ii}, \dots, n_{i \dots ii}$.

Downsampling of the feature maps in the upper layers is first conducted, and then, cascading is used to merge the feature maps with the posterior layer features. Each input includes the features selected from the first layer of the current block and the last layer of the previous block. The structural schematics with the newly added skip connection are displayed in Figure 5.

2.4. Deep Supervision. Deeper networks encode higher-level features. In the training of deep neural networks, deep supervision helps to reduce overfitting, extract more meaningful features, promote network convergence, and solve the problem of vanishing gradients [46, 50]. By adopting deep supervision in every stage of the decoder, the outputs of each intermediate stage can be used for supervision. Via upsampling, the output of each decoder can be adjusted to have the same dimensions as the final output segmentation map. The outputs of these intermediate stages are merged into the final output segmentation map, and then, softmax is used to derive the probability map. Losses can be calculated using ground truths and softmax outputs. In this way, the intermediate stages and the final output will implicitly contain the loss and gradient backpropagation, and the outputs of the intermediate stages will also gradually approach the ground truths. Figure 5 presents the structure of the network following the inclusion of deep supervision. We refer to this network model (which includes LGconv, a skip connection, and deep supervision) as DLSDNet.

$$\begin{aligned} \text{Parameters} &= (k_h \times k_w \times k_d \times C_{\text{in}} + 1) \times C_{\text{out}}, \\ \text{FLOPs} &= [(C_{\text{in}} \times k_h \times k_w \times k_d) + (C_{\text{in}} \times k_h \times k_w \times k_d - 1) + 1] \times C_{\text{out}} \times h \times w \times d. \end{aligned} \quad (8)$$

In the formula given above, k_h , k_w , and k_d denote the height, width, and depth of the convolution kernel, C_{in} and C_{out} are the number of input and output channels, and h , w , and d denote the height, width, and depth of the input data.

For the loss function, we adopted generalized Dice loss (GDL), which was proposed to cope with data imbalance issues. Using Dice loss is disadvantageous for the detection of small targets, so GDL combines multiple Dice classes and uses a weight to quantify and weight the segmentation results:

$$\text{GDL} = 1 - 2 \frac{\sum_{l=1}^L w_l \sum_{n=1}^N p_{\text{ln}} g_{\text{ln}}}{\sum_{l=1}^L w_l \sum_{n=1}^N (p_{\text{ln}} + g_{\text{ln}})}. \quad (9)$$

In the formula given above, g_{ln} denotes the ground truth of voxel n in class l , p_{ln} is the corresponding predicted value,

3. Experiments and Result Analysis

3.1. Dataset and Evaluation Indices. In our experiment, we employed the BraTS 2018 dataset [51, 52], which contains multimodal MRI scans from multiple institutions and has served as the official dataset in a challenge. This dataset comprises four types of MRI sequences: T1 weighted, T2 weighted, postcontrast T1 weighted, and FLAIR. The dimensions of the data are $240 \times 240 \times 155$. The dataset contains a training set and a validation set. The training set provides 285 sets of data for training with ground truth, while the testing set contains 66 sets of data with no ground truth. The objective of the BraTS 2018 challenge was to segment the data images into background, necrotic and nonenhancing tumors, edemas, and enhancing tumors. The researchers had to submit their validation results to an online evaluation platform to validate the effectiveness of their algorithms.

Segmentation accuracy was gauged using the Dice similarity coefficient and the Hausdorff distance. The former indicates the degree of similarity between the experimental segmentation results and the ground truth, with a higher value indicating greater segmentation precision. The latter calculates the maximum distance between the contours of the segmentation results and the ground truths to indicate the segmentation quality of the tumor boundaries, and a smaller absolute value represents better segmentation performance.

The number of model parameters (Parameters) represents the computer memory consumed by the model, and the amount of calculation represents the computing running time of the model, expressed in floating point numbers per second. The calculation is as follows:

N is the total number of voxels, L is the total number of classes, and w_l is the weight of each class:

$$w_l = \frac{1}{(\sum_{n=1}^N g_{\text{ln}})^2}. \quad (10)$$

3.2. Experiment Environment and Preprocessing. We used the deep-learning platform Pytorch to achieve the proposed model. We employed two NVIDIA GeForce 2080Ti GPU for 500 epochs of training. While training the model, we used the Adam optimizer with a self-adjusting learning rate. The initial learning rate was 0.0001. The L2 norm was used to normalize the model, and the weight decay rate was 10-5.

Due to the different imaging equipment and protocols, data artifacts were present within the MRI images [53], so we

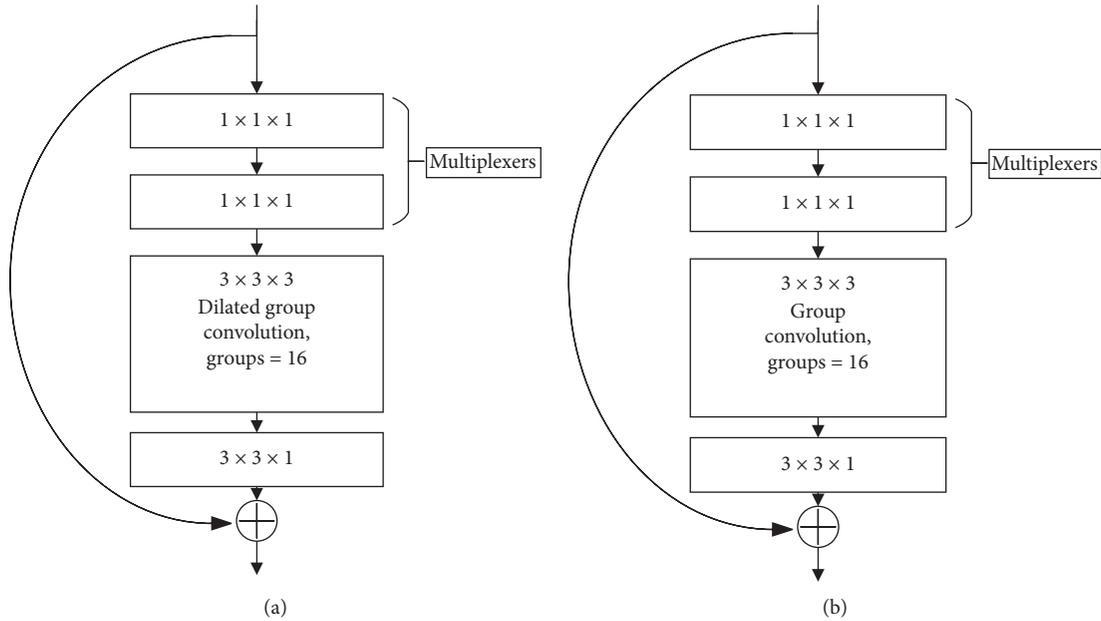


FIGURE 1: Convolution unit. (a) Input DMFunit. (b) Output MFunit.

used the N4 bias field correction algorithm to correct the bias in T1-weighted, T2-weighted, and postcontrast T1-weighted modes.

Due to GPU memory limitations, we expanded the data using the following techniques: random cropping of the original images to $128 \times 128 \times 128$, random mirroring of the axial, coronal, and sagittal views with a probability of 0.5, random rotation of the images between angles $[-10^\circ, 10^\circ]$, random changes to intensities between $[-0.1, 0.1]$, and scaling of images between $[0.9, 1.1]$.

3.3. Implementation of Experiment. Our model uses residual models as building blocks. The overall structure is similar to that of an encoder-decoder. The inputs are four channels of data corresponding to four modes of MRI data. During the feature encoding stage, the residual units of learnable groups are used, and the modified skip connection enhances the multiscale representation capabilities. During the decoding stage, the high-resolution features of cascading the encoder are used to supplement lost information. Upsampling is performed using trilinear interpolation. After each convolution block, Batch Normalization and Rectified Linear Unit (ReLU) activation are executed.

At the encoder end of the network, the skip connection was modified for information links between stages. Before encoding, max pooling is used for the downsampling of the high-level features in the encoder to match the scale on lower

levels; in other words, max pooling is applied to the output of the previous level. Previous features can be accessed directly for the inputs of each stage to enhance feature reuse.

In the decoding stage, the outputs of each stage can guide the final segmentation results, adjusting the outputs of each decoder so that they have the same dimensions as the final output segmentation map, thereby producing three different outputs that are combined and then subjected to the softmax operation to derive the final segmentation map.

3.4. Experiment Results and Analysis. To verify the effectiveness of the proposed network, we trained and validated the proposed network and the original DMFNet using the same training set and validation set. Table 1 presents the experiment results of the original DMFNet model, a network using learnable group convolution (DMFNet + LG), and the proposed DLSDNet (DMFNet + LG + Skip + DS). Performance in terms of brain tumor image segmentation was compared using the Dice score and the Hausdorff distance. Wt, Et, and Tc denote whole tumors, enhancing tumors, and tumor cores, respectively.

A comparison of the first and second rows in Table 1 show that using learnable grouping can improve the Dice score by 1.2% for Tc, thereby indicating that learnable grouping facilitates the detection of small targets. A comparison of the second and third rows in Table 1 show that the addition of the skip connection significantly improves the

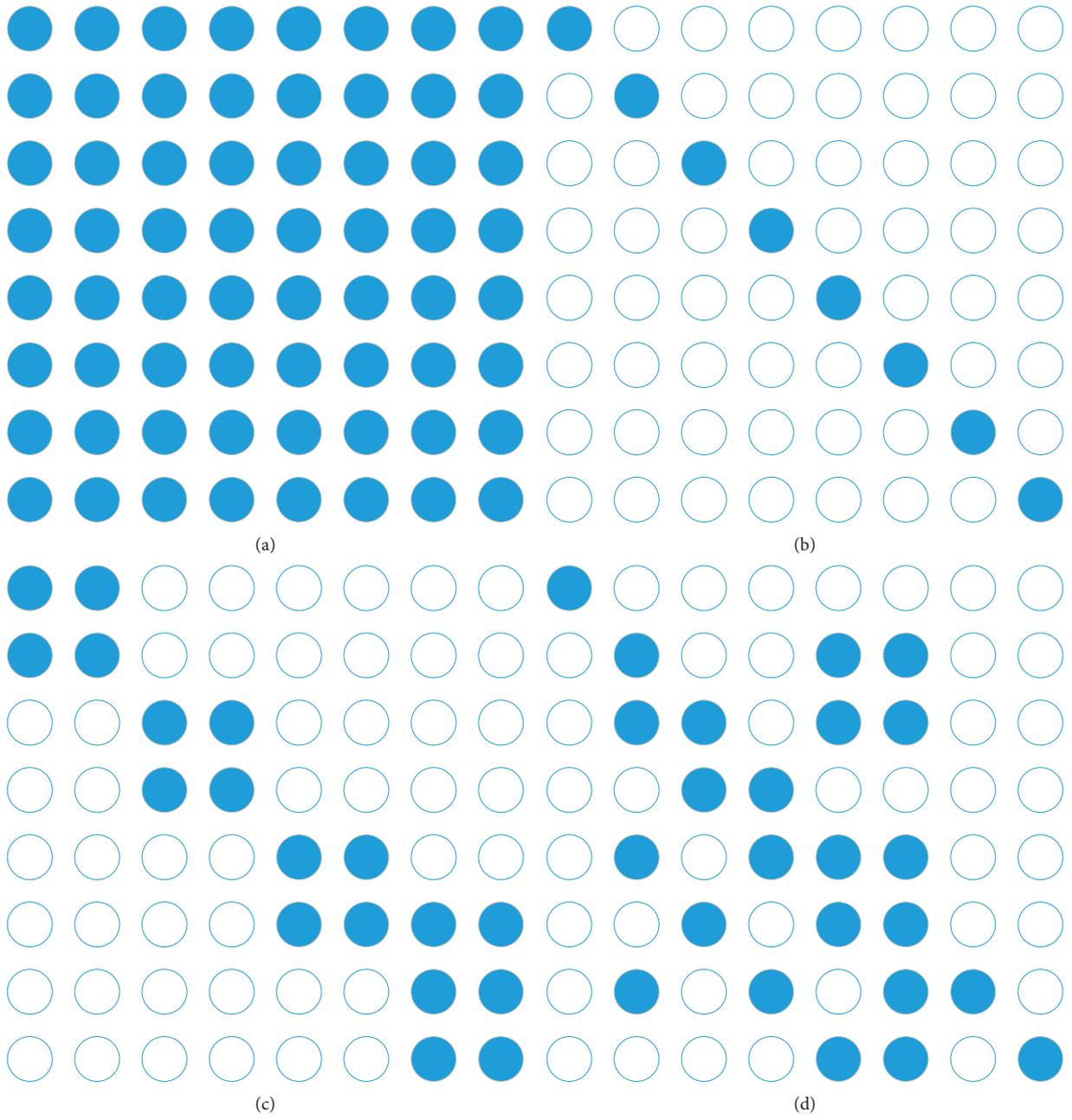


FIGURE 2: Structures of relation matrix U .

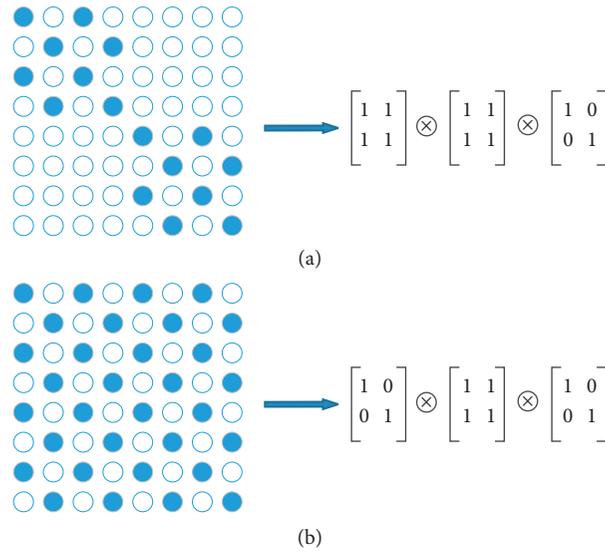


FIGURE 3: Typical relational matrices and associated Kronecker products.

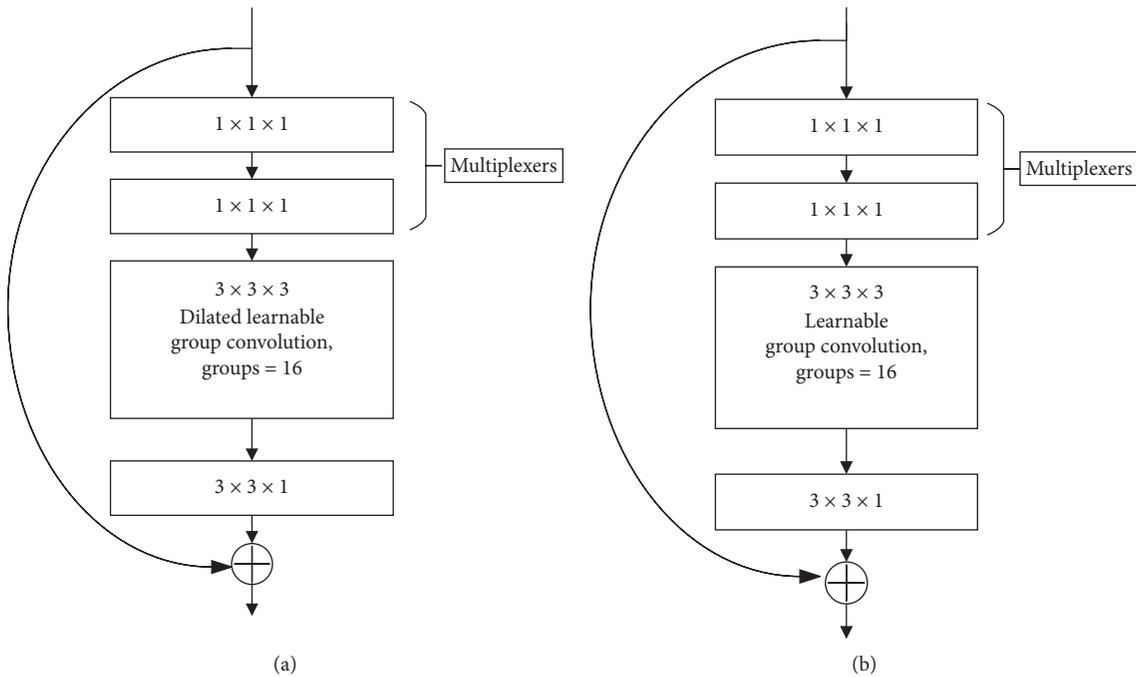


FIGURE 4: Locations of learnable groups in network units. (a) Input LG_DMFnunit. (b) Output LG_MFnunit.

Dice score by 0.19% for Et, 0.23% for Wt, and 0.98% for Tc, thereby indicating that the skip connection enables thorough utilization of multiscale information. Adding deep supervision facilitated the extraction of more features, as well as better supervision of the training process. DLSDNet based on DMFNet improves the Hausdorff by 0.54 mm for Wt and 0.7 mm for Tc.

Table 2 compares the proposed network model and typical brain tumor image segmentation networks. As

shown in Table 2, the proposed network model exhibits better segmentation performance than 3D U-Net [31] with improving the Dice score by 4.4% for Et, 1.72% for Wt, and 14.43% for Tc and improving the Hausdorff by 3.26 mm for Et, 12.63 mm for Wt, and 5.88 mm for Tc. Compared with another similar lightweight network model, 3D-ESPNet [35], DLSDNet improved the Dice score by 6.66% for Et, 1.95% for Wt, and 4.8% for Tc. Compared to the methods that won the first and second place in the BraTS 2018

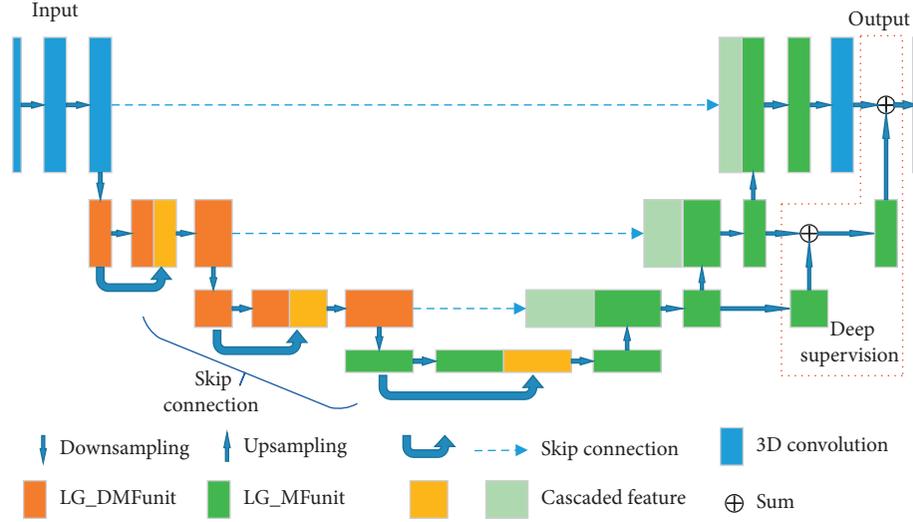


FIGURE 5: Structural schematics of DMFNet with skip connection and deep supervision.

TABLE 1: Experiment results of models applied to the BraTS 2018 validation set.

Model	Parameters (M)	FLOPs(G)	Dice (%)			Hausdorff (mm)		
			Et	Wt	Tc	Et	Wt	Tc
DMFNet [33]	3.84	23.60	79.79	89.90	84.01	2.77	5.01	6.44
DMFNet + LG	4.75	33.78	80.17	90.02	85.22	2.80	4.62	5.73
DLSDNet	5.23	36.14	80.36	90.25	86.20	2.78	4.47	5.74

TABLE 2: Comparison of the proposed network model and typical methods.

Model	Parameters (M)	FLOPs (G)	Dice (%)			Hausdorff (mm)		
			Et	Wt	Tc	Et	Wt	Tc
DLSDNet	5.23	36.14	80.36	90.25	86.20	2.78	4.47	5.74
3D U-Net [27]	16.21	1669.53	75.96	88.53	71.77	6.04	17.10	11.62
3D-ESPNet [29]	3.63	76.51	73.70	88.30	81.40	—	—	—
Kao et al. [30]	9.45	203.96	78.75	90.47	81.35	3.81	4.32	7.56
No New-Net [31]	10.36	202.25	81.01	90.83	85.44	2.41	4.27	6.52
NVDLMED [32]	40.06	1495.53	81.73	90.68	86.02	3.82	4.52	6.85

challenge (No New-Net and NVDLMED) [37, 38], DLSDNet achieved the best Dice score and Hausdorff on Tc with 86.20% and 5.74 mm, respectively. The proposed network model DLSDNet involved a smaller number of network model parameters, was less complex, and occupied fewer resources due to fewer FLOPs, as calculated in [39].

The visualized segmentation results are as shown in Figure 6. As can be seen, the original DMFNet can already roughly segment the contours of the tumor region. However,

in some details, such as smaller tumor core targets, the segmentation performance is poorer. Figure 6(b) displays the results using the network model with learnable grouping. The segmentation performance of this model with regard to tumor cores was superior to that of the original DMFNet. Figure 6(c) shows the segmentation results using the network model with a skip connection and deep supervision. The segmentation performance of this model with regard to whole tumors, enhancing tumors, and tumor cores

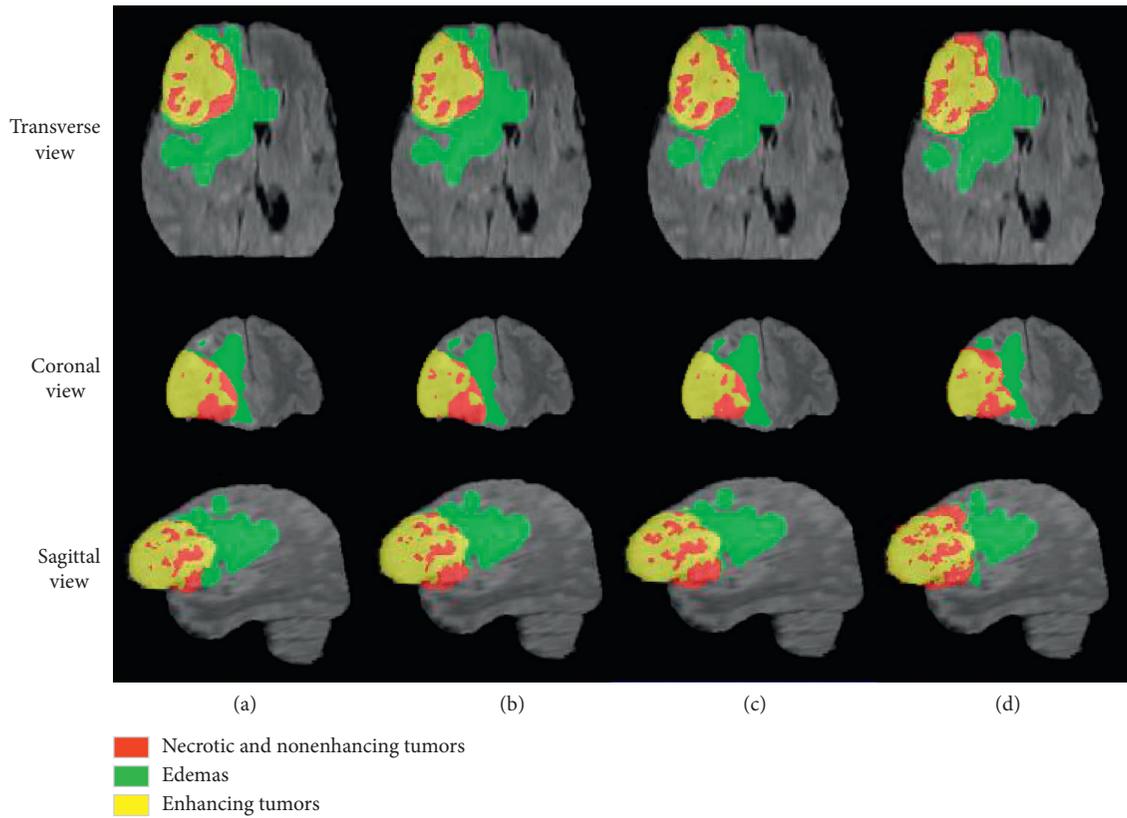


FIGURE 6: Visualized segmentation results. (a) DMFNet. (b) DLNet. (c) DLSDNet. (d) Ground truth.

improved further, and the results were even closer to the ground truth. As can be seen, the proposed DLSDNet model is better at segmenting small targets.

4. Conclusions

Brain tumors vary significantly in intensity and are irregular in shape. This study modified DMFNet to use fewer parameters and introduced LGConv to the feature extraction stage so that it can flexibly choose the number of groups based on dataset and network characteristics. This facilitates adaptation to more complex data features and has wider applicability. We added a skip connection between LGConv blocks to enable thorough utilization of multiscale information and to enhance feature reuse. We added deep supervision to the network output stage to merge the outputs of different stages and to reconstruct outputs with the same dimensions for the extraction of more distinctive features and the enhancement of segmentation accuracy. Experiments using the BraTS 2018 dataset revealed that the proposed model is superior to networks with conventional U-Net structures and has greater precision than other lightweight brain tumor image segmentation methods. The segmentation precision of the proposed network with regard to whole tumors, enhancing tumors, and tumor cores was 90.25%, 80.36%, and 86.20%. Compared with the methods that won the first place in the BraTS 2018 challenge (NVDLMED), the Parameters and FLOPs are reduced by 7 and 40 times, respectively. The proposed network, thus, has

significantly greater precision in the segmentation of enhancing tumors and tumor cores than the original DMFNet and also offers strong competition against the methods that won the first and second place in the BraTS 2018 challenge. Furthermore, the proposed method uses fewer parameters and is a less complex model.

Data Availability

The data used to support the findings of this study can be obtained from the corresponding author on request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported by the National Natural Science Foundation of China under Grant No. 61471263 and the Natural Science Foundation of Tianjin, China, under Grant 16J CZDJ C31100.

References

- [1] Q. Li, K. X. Bai, L. Zhao, and X. Guan, "Progress and challenges of MRI brain tumor image segmentation," *Journal of Image and Graphics*, vol. 25, no. 3, pp. 0419–0431, 2020.
- [2] P. Gibbs, D. L. Buckley, S. J. Blackband, and A. Horsman, "Tumour volume determination from MR images by

- morphological segmentation,” *Physics in Medicine and Biology*, vol. 41, no. 11, pp. 2437–2446, 1996.
- [3] Y. C. Sung, K. S. Han, C. J. Song, S.-M. Noh, and J.-W. Park, “Threshold estimation for region segmentation on MR image of brain having the partial volume artifact,” in *Proceedings of the 2000 5th International Conference on Signal Processing*, pp. 1000–1009, IEEE, Beijing, China, August 2000.
 - [4] A. Stadlbauer, E. Moser, S. Gruber et al., “Improved delineation of brain tumors: an automated method for segmentation based on pathologic changes of 1H-MRSI metabolites in gliomas,” *Neuroimage*, vol. 23, no. 2, pp. 454–461, 2004.
 - [5] R. A. Jan, K. H. A. Horst, K. A. Jan, M. G. Lentschig, and H.-O. Peitgen, “Multispectral brain tumor segmentation based on histogram model adaptation,” in *Proceedings of SPIE-2007 The International Society for Optical Engineering*, vol. 6514, San Diego, CA, USA, 2007.
 - [6] W. Deng, W. Xiao, H. Deng, and J. Liu, “MRI brain tumor segmentation with region growing method based on the gradients and variances along and inside of the boundary curve,” in *Proceedings of the 2010 3rd International Conference on Biomedical Engineering and Informatics*, pp. 393–396, IEEE, Yantai, China, October 2010.
 - [7] T. Węgliński and A. Fabijańska, “Brain tumor segmentation from mri data sets using region growing approach,” in *Proceedings of the Viith International Conference on Perspective Technologies & Methods in Mems Design*, Polyana, Ukraine, May 2011.
 - [8] M. M. J. Letteboer, O. F. Olsen, E. B. Dam, P. W. A. Willems, M. A. Viergever, and W. J. Niessen, “Segmentation of tumors in magnetic resonance brain images using an interactive multiscale watershed algorithm,” *Academic Radiology*, vol. 11, no. 10, pp. 1125–1138, 2004.
 - [9] C. C. Benson, V. L. Lajish, and K. Rajamani, “Brain tumor extraction from mri brain images using marker based watershed algorithm,” in *Proceedings of the 2015 International Conference on Advances in Computing*, Kochi, India, August 2015.
 - [10] Y. F. Tong, Q. Li, X. Guan et al., “An improved multi-modal brain tumor segmentation hybrid algorithm,” *Journal of Single Processing*, vol. 34, no. 3, pp. 340–346, 2018.
 - [11] L. Ren, Q. Li, X. Guan et al., “Three-dimensional segmentation of brain tumors in magnetic resonance imaging based on improved continuous max-flow,” *Laser & Optoelectronics Progress*, vol. 55, no. 11, pp. 215–223, 2018.
 - [12] J. Nie, Z. Xue, T. Liu et al., “Automated brain tumor segmentation using spatial accuracy-weighted hidden markov random field,” *Computerized Medical Imaging and Graphics*, vol. 33, no. 6, pp. 431–441, 2009.
 - [13] S. Bauer, L. P. Nolte, and M. Reyes, “Segmentation of brain tumor images based on at-las-registration combined with a Markov-Random-Field lesion growth model,” in *Proceedings of the IEEE International Symposium on Biomedical Imaging: from Nano to Macro*, pp. 2018–2021, IEEE, Chicago, IL, USA, March 2011.
 - [14] T. Shen, X. Huang, H. Li, E. Kim, S. Zhang, and J. Huang, “A 3D Laplacian-driven parametric deformable model,” in *Proceedings of the IEEE International Conference on Computer Vision*, Barcelona, Spain, November 2011.
 - [15] S. Ho, E. Bullitt, and G. Gerig, “Level-set evolution with region competition: auto-matic 3-d segmentation of brain tumors,” in *Proceedings of the 2002 16th International Conference on Pattern Recognition*, pp. 532–535, Quebec City, Quebec, Canada, August 2002.
 - [16] S. Bauer, L.-P. Nolte, and M. Reyes, “Fully automatic segmentation of brain tumor images using support vector machine classification in combination with hierarchical conditional random field regularization,” *Lecture Notes in Computer Science*, vol. 14, no. 3, pp. 354–361, 2011.
 - [17] A. Kharrat, K. Gasmi, M. B. Messaoud, N. Benamrane, and M. Abid, “A hybrid approach for automatic classification of brain MRI using genetic algorithm and support vector machine,” *Leonardo Journal of Sciences*, vol. 17, no. 1, pp. 71–82, 2010.
 - [18] L.-L. Li, J. Sun, C.-H. Wang, Y.-T. Zhou, and K.-P. Lin, “Enhanced Gaussian process mixture model for short-term electric load forecasting,” *Information Sciences*, vol. 477, pp. 386–398, 2019.
 - [19] K.-P. Lin, H.-F. Chang, T.-L. Chen, Y.-M. Lu, and C.-H. Wang, “Intuitionistic fuzzy C-regression by using least squares support vector regression,” *Expert Systems with Applications*, vol. 64, pp. 296–304, 2016.
 - [20] M. Soltaninejad, X. Ye, G. Yang, N. Allinson, and T. Lambrou, “Brain tumour grading in different MRI protocols using SVM on statistical features,” in *Proceedings of the MIUA 2014*, pp. 259–264, Egham, UK, July 2014.
 - [21] C.-H. Wang, K.-P. Lin, Y.-M. Lu, and C.-F. Wu, “Deep belief network with seasonal decomposition for solar power output forecasting,” *International Journal of Reliability, Quality and Safety Engineering*, vol. 26, no. 6, Article ID 1950029, 2019.
 - [22] Y. Zhou, X. Zhao, K.-P. Lin, C.-H. Wang, and L. Li, “A Gaussian process mixture model-based hard-cut iterative learning algorithm for air quality prediction,” *Applied Soft Computing*, vol. 85, Article ID 105789, 2019.
 - [23] T. L. Jones, T. J. Byrnes, G. Yang et al., “Brain tumor classification using the diffusion tensor image segmentation (D-SEG) technique,” *Neuro-Oncology*, vol. 17, no. 3, pp. 466–476, 2014.
 - [24] W. Wu, A. Y. C. Chen, L. Zhao, and J. J. Corso, “Brain tumor detection and segmentation in a CRF (conditional random fields) framework with pixel-pairwise affinity and superpixel-level features,” *International Journal of Computer Assisted Radiology and Surgery*, vol. 9, no. 2, pp. 241–253, 2014.
 - [25] A. Pinto, S. Pereira, H. Dinis, C. A. Silva, and D. M. L. D. Rasteiro, “Random decision forests for automatic brain tumor segmentation on multi-modal MRI images,” in *Proceedings of the 2015 IEEE 4th Portuguese Meeting on Bioengineering*, pp. 1–5, IEEE, Porto, Portugal, February 2015.
 - [26] M. Goetz, C. Weber, F. Binczyk et al., “DALSA: domain adaptation for supervised learning from sparsely annotated MR images,” *IEEE Transactions on Medical Imaging*, vol. 35, no. 1, pp. 184–196, 2016.
 - [27] S. Pereira, A. Pinto, V. Alves, and C. A. Silva, “Brain tumor segmentation using convolutional neural networks in MRI images,” *IEEE Transactions on Medical Imaging*, vol. 35, no. 5, pp. 1240–1251, 2016.
 - [28] M. Havaei, A. Davy, D. Warde-Farley et al., “Brain tumor segmentation with deep neural networks,” *Medical Image Analysis*, vol. 35, pp. 18–31, 2017.
 - [29] H. Chen, Q. Dou, L. Yu, J. Qin, and P.-A. Heng, “VoxResNet: deep voxelwise residual networks for brain segmentation from 3D MR images,” *NeuroImage*, vol. 170, pp. 446–455, 2018.
 - [30] X. Zhao, Y. Wu, G. Song, Z. Li, Y. Zhang, and Y. Fan, “A deep learning model integrating FCNNs and CRFs for brain tumor segmentation,” *Medical Image Analysis*, vol. 43, pp. 98–111, 2017.
 - [31] Ö. Çiçek, A. Abdulkadir, S. Lienkamp, T. Brox, and O. Ronneberger, “3D U-Net: learning dense volumetric

- segmentation from sparse annotation,” in *Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 424–432, Athens, Greece, October 2016.
- [32] K. Bai, Q. Li, and C. H. Wang, “Integrating improved U-Net and continuous maximum flow algorithm for 3D brain tumor image segmentation,” *Journal of Imaging Science and Technology*, vol. 64, no. 4, Article ID 040412, 2020.
- [33] L. Zhao, Q. Li, C. H. Wang, and Y. C. Liao, “3D brain tumor image segmentation integrating cascaded anisotropic fully convolutional neural network and hybrid level set method,” *Journal of Imaging Science and Technology*, vol. 64, no. 4, Article ID 040411, 2020.
- [34] R. Sherman, “A volumetric convolutional neural network for brain tumor segmentation,” 2018, <http://arxiv.org/abs/1811.02654v1>.
- [35] N. Nuechterlein and S. Mehta, “3D-ESpNet with pyramidal refinement for volumetric brain tumor image segmentation,” in *Proceedings of the 2018 4th International MICCAI Brainlesion Workshop*, pp. 245–253, Granada, Spain, September 2018.
- [36] P.-Y. Kao, T. Ngo, A. Zhang, and J. Chen, “Brain tumor segmentation and tractographic feature extraction from structural mr images for overall survival prediction,” *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*, pp. 128–141, Springer, Cham, Switzerland, 2018.
- [37] F. Isensee, P. Kickingereder, W. Wick, and M. Bendsuz, “No new-net,” *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*, pp. 234–244, Springer, Cham, Switzerland, 2018.
- [38] A. Myronenko, “3D MRI brain tumor segmentation using autoencoder regularization,” *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*, pp. 311–320, Springer, Cham, Switzerland, 2018.
- [39] C. Chen, X. Liu, M. Ding, Z. Junfeng, and J. Li, “3D dilated multi-fiber network for real-time brain tumor segmentation in mri,” *Medical Image Computing and Computer-Assisted Intervention*, Springer, Cham, Switzerland, pp. 184–192, 2019.
- [40] A. Krizhevsky, I. Sutskever, and G. E. Hinton, “Imagenet classification with deep convolutional neural networks,” *Advances in Neural Information Processing Systems*, vol. 25, no. 2, pp. 1097–1105, 2012.
- [41] S. Xie, R. Girshick, P. Dollar, Z. Tu, and K. He, “Aggregated residual transformations for deep neural networks,” in *Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1492–1500, Honolulu, USA, July 2017.
- [42] Z. Zhang, J. Li, W. Shao et al., “Differentiable learning-to-group channels via groupable convolutional neural networks,” in *Proceedings of the 2019 IEEE International Conference on Computer Vision*, pp. 3542–3551, Seoul, Korea, October 2019.
- [43] K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 770–778, Las Vegas, NV, USA, June 2016.
- [44] G. Huang, Z. Liu, L. V. D. Maaten, and K. Q. Weinberger, “Densely connected convolutional networks,” in *Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 4700–4708, July 2017.
- [45] C. Szegedy, W. Liu, Y. Jia et al., “Going deeper with convolutions,” in *Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1–9, Boston, MA, USA, June 2015.
- [46] Q. Dou, L. Yu, H. Chen et al., “3D deeply supervised network for automated segmentation of volumetric medical images,” *Medical Image Analysis*, vol. 41, pp. 40–54, 2017.
- [47] H. Chen, X. Qi, J. Z. Cheng, and P. A. Heng, “Deep contextual networks for neuronal structure segmentation,” in *Proceedings of the 2016 AAAI Conference on Artificial Intelligence*, pp. 1167–1173, Phoenix, AZ, USA, February 2016.
- [48] A. G. Howard, M. Zhu, B. Chen et al., “Mobilenets: efficient convolutional neural networks for mobile vision applications,” 2017, <http://arxiv.org/abs/1704.04861>.
- [49] K. Batselier and W. Ngai, “A constructive arbitrary-degree Kronecker product decomposition of tensors,” *Numerical Linear Algebra with Applications*, vol. 24, no. 3, Article ID e2097, 2017.
- [50] C.-Y. Lee, S. Xie, P. Gallagher, Z. Zhang, and Z. Tu, “Deeply-supervised nets,” in *Proceedings of the Eighteenth International Conference on Artificial Intelligence and Statistics*, pp. 562–570, San Diego, CA, USA, May 2015.
- [51] B. H. Menze, A. Jakab, S. Bauer et al., “The multimodal brain tumor image segmentation benchmark (BRATS),” *IEEE Transactions on Medical Imaging*, vol. 34, no. 10, pp. 1993–2024, 2015.
- [52] S. Bakas, H. Akbari, A. Sotiras et al., “Advancing the Cancer Genome Atlas glioma MRI collections with expert segmentation labels and radiomic features,” *Scientific Data*, vol. 4, p. 170117, 2017.
- [53] G. Collewet, M. Strzelecki, and F. Mariette, “Influence of MRI acquisition protocols and image intensity normalization methods on texture classification,” *Magnetic Resonance Imaging*, vol. 22, no. 1, pp. 81–91, 2004.