

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/339263697>

FMNet: Feature Mining Networks for Brain Tumor Segmentation

Conference Paper · November 2019

DOI: 10.1109/ICTAI.2019.00083

CITATIONS

5

READS

121

7 authors, including:



Fengming Lin

University of Leeds

14 PUBLICATIONS 1,584 CITATIONS

SEE PROFILE



Wali Ullah Khan

University of Luxembourg

195 PUBLICATIONS 3,102 CITATIONS

SEE PROFILE



Enshuai Pang

Shandong University

2 PUBLICATIONS 19 CITATIONS

SEE PROFILE

FMNet: Feature Mining Networks for Brain Tumor Segmentation

Fengming Lin, Ju Liu, Qiang Wu, Xiangmao Kong, Wali Ullah Khan, Wei Shi, Enshuai Pang
School of Information Science and Engineering, Shandong University, Qingdao266237, China.

Emails: 1900258258@qq.com, juliu@sdu.edu.cn, 20034369@qq.com, silencekxm@qq.com, waliullahkhan30@gmail.com, 1836109780@qq.com, 2919028238@qq.com

Abstract—Brain tumor is one of the primary diseases that endanger human life and health. Multi-modality magnetic resonance imaging (MRI) for tumor analysis is one of the key techniques to clinical diagnosis. Experienced experts artificially segment the classical method of brain tumor segmentation according to their anatomical and pathological knowledge, which is time-consuming. In this paper, we propose a deep neural network structure, namely Feature Mining Networks (FMNet), for brain tumor segmentation. The proposed FMNet adopts three innovative structure, including semantic information mining unit (SIMU), macro information mining unit (MIMU), and a feature correction unit (FCU). These three units can enhance the mining of semantic information and spatial information, and further, modify the information in a direction that is conducive to segmentation results. Each unit can bring an improvement in segmentation performance. We evaluate the proposed framework on BraTS2017 and BraTS2018 dataset. The experimental results show that our FMNet performs better than state-of-the-art networks such as fully convolutional networks (FCN), U-Net and so on.

Index Terms—Brain Tumor, Segmentation, Feature Mining Networks, Deep Neural Networks, Deep Learning

I. INTRODUCTION

Brain tumor is one of the primary diseases that endanger human life and health [1]. Multi-modality magnetic resonance imaging (MRI) for tumor analysis is one of the keys to treatment [2]. Segmentation of brain tumors is very popular in the field of diagnosis. Owing to the high complexity of brain tumor MRI, brain tumor segmentation is still a challenging task and need to be further explored. The classical method is artificially segmented by experienced experts according to their anatomical and pathological knowledge, with the help of specific software. However, this method is time-consuming and labor-intensive, and the correction rate of labeling varies from person to case. Over the past few decades, there has been increasing interest in developing computational methods to assist in the analysis of MRI in pathology.

Since the AlexNet won the imagenet large scale visual recognition challenge (ILSVRC) in 2012 [3], deep learning stands out in machine learning algorithms, especially in tasks classification, object detection [4], and semantic segmentation [5] [6]. ResNet introduced residual unit to reduce gradient dispersion and performed better on these tasks [7]. Recently,

The work was supported by the foundation of the key techniques of brain-like computing and application in multimodal information processing under grants No. 2017CXGC1504.

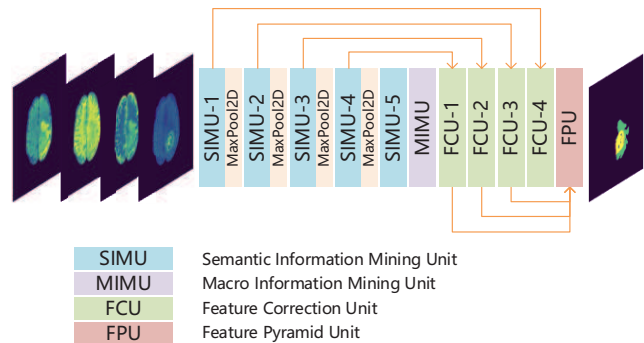


Fig. 1: Structure of FMNets. Consists of four parts: SIMU, MIMU, FCU, and FPU.

image processing based on deep neural networks has become a hot topic for research in academia and industry. The initial segmentation methods based on convolutional neural networks train a classification network by taking patches as input [8]. The category of the patch is treated as the category of the central pixel on the patch. Such a classification network is trained to achieve pixel-level classification, namely segmentation.

In this way, it is difficult to select the size of small color blocks, and there is the problem of repeated calculation of adjacent pixel points. Therefore, the fully convolutional networks (FCN) replaces the full connection layer in the traditional classification network structure. It uses the fully convolutional layer to predict the category of pixel points from the extracted feature graph to complete segmentation tasks. Fully convolutional networks can take arbitrary scale images as input, to solve the problem of repeated calculation of adjacent pixels. Compared with natural images, medical images have fewer data samples and higher complexity, i.e., large gray value range and complex boundary. U-Net performs well in solving the problem of medical image segmentation [9]. Its main structure is to combine images of the same scale in the process of up-sampling and down-sampling, to realize the fusion of low-level features and high-level features. There are many variants of U-Net, such as combinations with modules like ResNet, DenseNet [10] [11] [12].

Methods such as PANet [13] and FPN [14] combine a

large amount of information. The merged information does not necessarily contribute to more accurate segmentation. So we need to correct the feature in a direction that is favorable for segmentation results. The traditional U-Net model only uses a convolution kernel of size 3, which is insufficient for spatial information extraction. Therefore, we need to extract more spatial location information for segmentation. Methods such as V-Net [15] use the structure of ResNet to extract more abundant semantic information, but this is not enough. There are still many details waiting to be mined.

In this paper, we propose a deep neural network structure, namely Feature Mining Networks (FMNet), for brain tumor segmentation. We innovate three network structural units for semantic segmentation. Each unit brings a significant improvement in segmentation performance. The structure of FMNet is shown in Figure 1. The main contributions of this paper is listed below:

- 1) We innovate the Semantic Information Mining Unit (SIMU). Based on feature extraction of the encoder, a branch is added to each block to search for new deep semantic information. This allows for more abundant semantic information.
- 2) We innovate the Macro Information Mining Unit (MIMU). We use multi-scale convolution kernels to scan features of small size, including convolution kernels with large scale. This can be an excellent way to mine spatial location information. After the fusion of this information, we conduct the attention operations for the integrated features.
- 3) We innovate the Feature Correction Unit (FCU). Since the fused information is not necessarily beneficial to the segmentation results, this unit corrects the fused features in a direction conducive to segmentation. The corrections are made using features from the decoder.

II. RELATED WORKS

A. Feature Pyramid Unit (FPU)

Image segmentation method based on convolutional neural network loses the original position information after multiple convolution and downsampling of the data. Upsampling directly can restore resolution but affects segmentation accuracy. The feature pyramid structure solves this problem well, which can effectively use position information and multi-scale information to improve performance of image segmentation. The structure of FPU is shown in Figure 2.

The output of each upsampling region in the decoder is upsampled to the scale of the original image respectively, with the same number of channels. Then add them up to get the fused features. The feature passes through a convolution layer and is segmented after passing the softmax activation function. The fusion process is shown below:

$$Output = \sum_{i=1}^n H(x_i). \quad (1)$$

where *Output* indicates the output of the FPU. $H(x_i)$ represents the *i*th feature from upsample layer. Of course, if

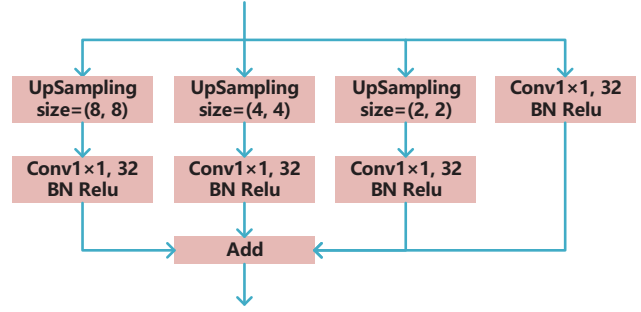


Fig. 2: Feature Pyramid Unit.

the feature is already original scale, it does not need to be upsampled but added directly.

III. METHODS

The proposed FMNet adopts three innovative structure, including FCU, SIMU, and MIMU, respectively. These three units can enhance the mining of semantic information and spatial information, and can further modify the information in a direction that is conducive to segmentation results. In addition, resnet structure and feature pyramid structure are also used in this segmentation network.

A. Semantic Information Mining Unit (SIMU)

Traditional deep learning algorithms usually include encoder and decoder, where the encoder is used to extract detailed information. The decoder fuses these details and downsamples them to get the final segmentation result. At present, most of the segmentation methods use only one path in the encoder part. The segmentation method based on ResNet adds a residual path to each downsampling area of the encoder to supplement information. However, the supplementary information comes from the convolution layer before, which lead to the insufficiency of extracted details. Therefore, semantic information compensation should be carried out in the segmentation process. The structure of SIMU is shown in Figure 3.

The encoder section is composed of multiple downsampling blocks which include a SIMU and a pooling layer. Each SIMU contains two branches. The first is the main feature extraction branch, which is a residual unit. It consists of feature extraction sub-branch and residual sub-branch. The feature extraction sub-branch includes two convolutional layers with the kernel size of 3. The residual sub-branch is a convolutional layer with the kernel size of 1. After each convolutional layer, there is a batch normalization layer and an activation layer. The output of main feature extraction branch is obtained by adding the features of two sub-branches. The second is the supplementary feature extraction branch. It is composed of two residual units. The structure of each residual unit is the same as that of the main information extraction branch, but the number of channels is half of it. Then the features of these two

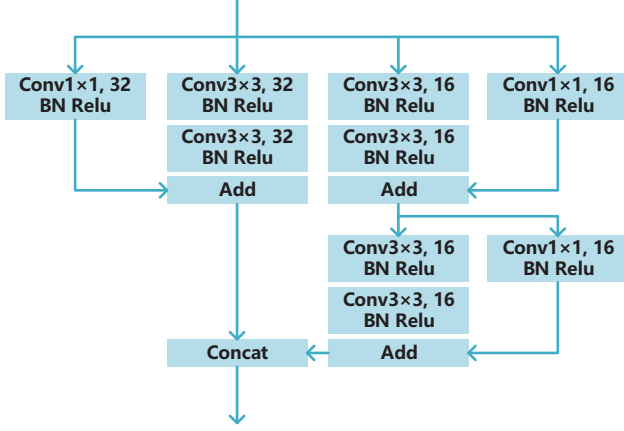


Fig. 3: Structure of SIMU. The first SIMU was taken as example.

branches are concatenated to get the output of SIMU. After the SIMU, a pooling layer is used to change the image scale. The output features of the pooling layer are the input features of the next downsampling block. Finally, the ultimate output of the encoder is obtained through a downsampling block without pooling layer. The process is shown below:

$$E(x_i) = M(x_i) \oplus S(x_i). \quad (2)$$

where $E(x_i)$ is the output of the i th SIMU in encoder. $M(x_i)$ is the output of main feature extraction branch in the i th SIMU. $S(x_i)$ is the output of supplementary feature extraction branch in the i th SIMU. \oplus means concatenate operation.

B. Macro Information Mining Unit (MIMU)

The traditional deep neural network based segmentation algorithm directly connects the encoder to the decoder. In the encoding process, the scale of features gradually decreases while channels increase. In general, the decoder directly upsamples these small scale features. We have found through experiments that the features obtained by the last layer of encoders are essential for the segmentation results. Due to the small-scale nature, we can use a large convolution kernel to scan it and extract more spatial localization information for segmentation. The structure of MIMU is depicted in Figure 4.

The MIMU consists of three branches. The first branch is an ordinary convolution layer, which carries out convolution operation on features from the last layer of the encoder. The second is the macro information mining branch, which uses three convolution kernels with different scales to carry out convolution operations on features, respectively. The scale of these three convolution kernels is 3, 5 and 7. Then add the three features, and finally multiplying added features by output features of the first branch to obtain the output of the second branch. The third is global attention branch. Meanwhile, global average pooling is conducted for the output of the first branch. Then it is multiplied by the output of the second branch.

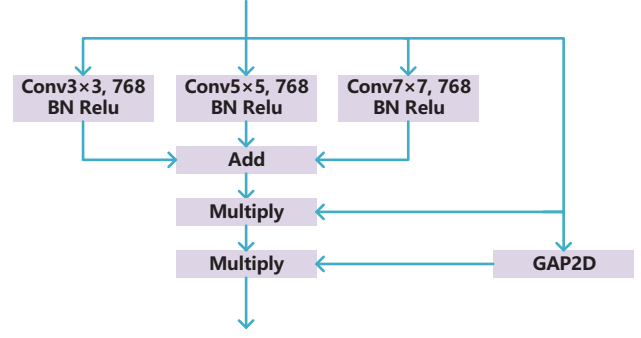


Fig. 4: Structure of MIMU.

Finally, we get the output of the MIMU. The process is shown below:

$$M(x) = E(x_n) \cdot Multi(E(x_n)) \cdot GAP(E(x_n)), \quad (3)$$

$$Multi(x) = Conv3(x) + Conv5(x) + Conv7(x). \quad (4)$$

where $M(x)$ is the output of the MIMU. $E(x_n)$ is the output of the last downsampling area of encoder. $Multi(E(x_n))$ is the output of macro information mining branch which consists of three convolution layers of different scales. $GAP(E(x_n))$ is the output of global attention branch.

C. Feature Correction Unit

The current segmentation methods have focused on feature fusion. These methods fuse all the information from different levels of space and detail. More information is extracted in SIMU and MIMU, and both favorable and unfavorable information for segmentation results are fused. Therefore, it is necessary to correct the features in a direction that favors the segmentation result during the segmentation process. The structure of FCU is shown in Figure 5.

The decoder part is composed of multiple upsampling areas. The number can be set freely and eventually needs to be upsampled to the original scale. The upsampling area includes an upsample layer, a feature correction part, and a residual part. After each convolution layer, local response normalization is adopted to prevent gradient dispersion. The feature correction section consists of two multiplicative layers and one concatenate layer. The first feature multiplication layer multiplies the encoder output by the upsampled output from the decoder, both of which are of the same scale. The second feature multiplication layer performs a self-multiplication operation on upsampled output. Finally, the concatenate layer connects these two features to obtain the output of feature correction part. The correction process is shown below:

$$D(x_i) = (E(x_i) \cdot D(x_{i-1})) \oplus (D(x_{i-1}) \cdot D(x_{i-1})). \quad (5)$$

Where $D(x_i)$ and $E(x_i)$ are the same level output of decoder and encoder, respectively. $D(x_{i-1})$ is the output of previous level in decoder.

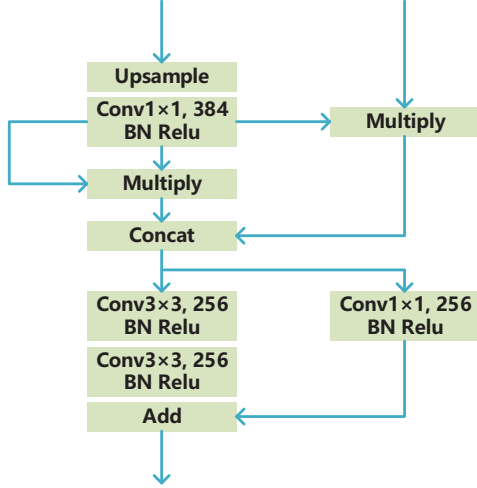


Fig. 5: Structure of FCU. The first FCU was taken as example.

IV. EXPERIMENTS

We have calculated dice of our proposed FMNet method on BraTS2017 and BraT2018 datasets.

A. Dataset

BraTS2018 [16] dataset and BraTS2017 [17] dataset are employed respectively for training and testing. These are standard datasets from MICSAT-BraTS competition, which contain multimodal MRI scans of glioblastoma (GBM/HGG) and lower grade glioma (LGG). Among them, HGG comprises 210 cases, and LGG comprises 75 cases. HGG and LGG are mixed randomly during training. These multimode scans describe native (T1), post-contrast T1-weighted (T1Gd), T2-weighted (T2), and T2 Fluid Attenuated Inversion Recovery (FLAIR) volumes. They are obtained by various scanners from multiple agencies. All imaging datasets are manually segmented by one to four raters according to the same annotation protocol. Experienced neuroradiologists recognize their annotations. Annotations comprise the GD-enhancing tumor (ET-label 4), the peritumoral edema (ED-label 2), and the necrotic and non-enhancing tumor core (NCR/NET-label 1), as described in [18].

B. Evaluation standard

Dice Similarity Coefficient (DSC) was evaluated in the following three segmentation tasks in our experiment: Whole Tumor (WT), Tumor Core (TC), Enhancing Tumor (ET). The segmented result can be divided into four parts, namely True Positive (TP), False Positive (FP), True Negative (TN) and False Negative (FN).

$$DSC = \frac{2TP}{FP + 2TP + FN}. \quad (6)$$

C. Preprocessing

Due to the small number and high complexity of medical images, it is challenging to extract useful features for analysis. Therefore, we need to preprocess these images using multimodal features of MRI. First, images with four modalities were normalized to adjust the gray value distribution. Then these four modalities images were fused into four channels datasets. At the same time, we removed slices from the dataset without any label. After shuffling the order, they were divided into five parts, including four parts for training and one part for testing.

D. Baseline

In a comparative experiment, VGG [19] is a brain tumor segmentation method using VGGNet structure. DUNet [20] is a brain tumor method to deepen U-Net, which is more effectual than the traditional U-Net structure. FCNN [21] is a fully convolutional networks method on brain tumor segmentation. HPUNet [22] adds feature pyramid structure to the conventional U-Net, which improves the segmentation accuracy. For each convolution block in HPU-Net, we added residuals unit, namely baseline. The experiments used a GTX-1080ti graphics card with 11G of graphics memory.

E. Results

TABLE I: Comparison with state-of-the-arts on the testing set of BRATS2017. **Boldface is our method.**

BraTS2017	Dice		
	WT	TC	ET
VGG	0.9058	0.6909	0.6479
DUNet	0.8959	0.7035	0.6275
FCNN	0.9097	0.6511	0.5705
HPUNet	0.9244	0.8001	0.7636
baseline(HPUNet+ResNet)	0.9257	0.8068	0.7683
baseline+FCU	0.9257	0.8223	0.7904
baseline+FCU+MIMU	0.9272	0.8273	0.7964
baseline+FCU+MIMU+SIMU	0.9272	0.8390	0.8003

Table 1 shows the segmentation results of our methods on BraTS2017 dataset. As can be seen from the table, each innovation point contributes to the improvement of dice. FMNet is a two-dimensional model that surpassed BraTS2017 champion in each of the three split tasks. Also, FMNet performed even better than the 3D model in BraTS2017 competition.

Table 2 shows the segmentation effect of our method on BraTS2018 dataset. Our method FMNet is still superior to FCNN, DUNet, and other traditional methods. It is well known that in general, the effect of two-dimensional model is not as good as that of the three-dimensional model. In BraTS2018 competition, the first three winners are 3D models, but our 2D model can still surpass the third place in WT and the second place in ET.

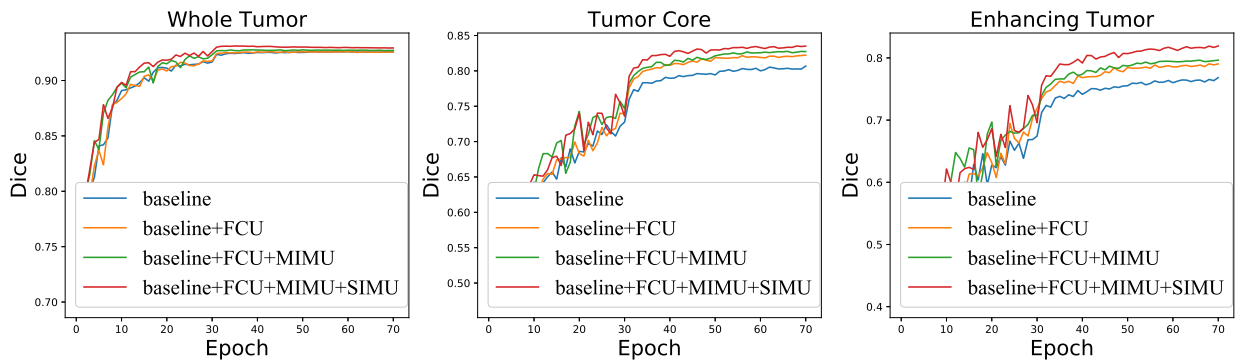


Fig. 6: Dice on BraTS2017.

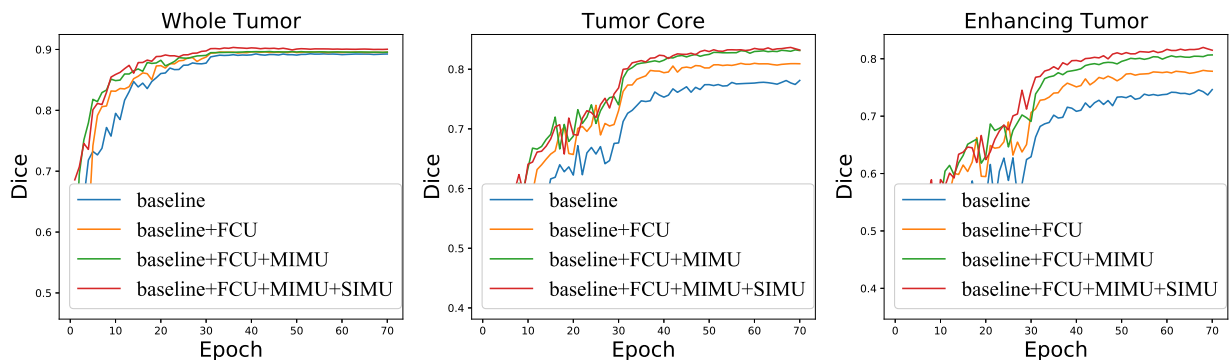


Fig. 7: Dice on BraTS2018.

TABLE II: Comparison with state-of-the-arts on the testing set of BRATS2018. **Boldface is our method.**

BraTS2018	Dice		
	WT	TC	ET
VGG	0.8724	0.6933	0.6598
DUNet	0.8631	0.6884	0.6205
FCNN	0.8687	0.6537	0.5828
HPUNet	0.8930	0.8108	0.7558
baseline(HPUNet+ResNet)	0.8925	0.7811	0.7463
baseline+FCU	0.8954	0.8091	0.7783
baseline+FCU+MIMU	0.8956	0.8313	0.8067
baseline+FCU+MIMU+SIMU	0.9019	0.8374	0.8176

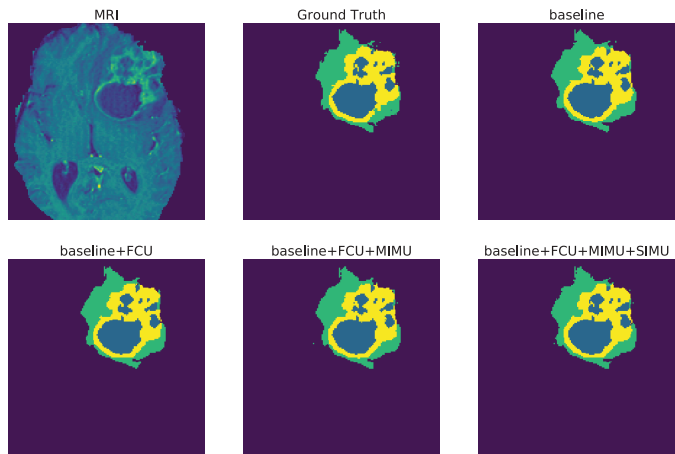


Fig. 8: Visual segmentation results of BraTS2017 sample.

As can be seen in Figure 6 and Figure 7, each method we proposed can improve the segmentation effect. The results of both BraTS2017 and BraTS2018 tend to stabilize after the 30th Epoch. Figure 8 and Figure 9 are visualizations of FMNet segmentation performance. The segmentation results of FMNet is very close to ground truth. It is hard to see the difference between them and ground truth with eyes.

V. CONCLUSION

In this paper, we proposed a deep neural network structure, namely Feature Mining Networks, for brain tumor segmentation. The innovation lies in the following three aspects: The first is Feature Correction Unit, which modifies the information in a direction that is conducive to segmentation results. Secondly, we proposed a Macro Information Mining Unit, which scans small-scale features with multi-scale convolution

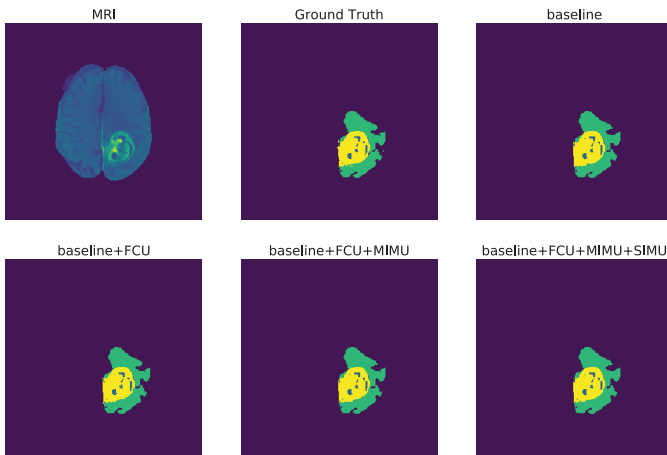


Fig. 9: Visual segmentation results of BraTS2018 sample.

kernels to obtain more spatial location information. The third is the Semantic Information Mining Unit, which uses three branches to dig more in-depth details. Experiments proved that our three innovations could bring the improvement of segmentation accuracy. In the future, we plan to focus on weakly supervised learning and the clinical use of appropriate methods.

REFERENCES

- [1] Stefan Bauer, Roland Wiest, Lutz-P Nolte, and Mauricio Reyes. A survey of mri-based medical image analysis for brain tumor studies. *Physics in Medicine & Biology*, 58(13):R97, 2013.
- [2] Bjoern H Menze, Andras Jakab, Stefan Bauer, Jayashree Kalpathy-Cramer, Keyvan Farahani, Justin Kirby, Yuliya Burren, Nicole Porz, Johannes Slotboom, Roland Wiest, et al. The multimodal brain tumor image segmentation benchmark (brats). *IEEE transactions on medical imaging*, 34(10):1993, 2015.
- [3] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems*, pages 1097–1105, 2012.
- [4] Ross Girshick, Jeff Donahue, Trevor Darrell, and Jitendra Malik. Rich feature hierarchies for accurate object detection and semantic segmentation. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 580–587, 2014.
- [5] Liang-Chieh Chen, George Papandreou, Iasonas Kokkinos, Kevin Murphy, and Alan L Yuille. Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs. *IEEE transactions on pattern analysis and machine intelligence*, 40(4):834–848, 2018.
- [6] Jonathan Long, Evan Shelhamer, and Trevor Darrell. Fully convolutional networks for semantic segmentation. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 3431–3440, 2015.
- [7] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778, 2016.
- [8] Mohammadreza Mostajabi, Payman Yadollahpour, and Gregory Shakhnarovich. Feedforward semantic segmentation with zoom-out features. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 3376–3385, 2015.
- [9] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention*, pages 234–241. Springer, 2015.
- [10] Gao Huang, Zhuang Liu, Laurens Van Der Maaten, and Kilian Q Weinberger. Densely connected convolutional networks. In *CVPR*, volume 1, page 3, 2017.
- [11] Simon Jégou, Michal Drozdal, David Vazquez, Adriana Romero, and Yoshua Bengio. The one hundred layers tiramisu: Fully convolutional densenets for semantic segmentation. In *Computer Vision and Pattern Recognition Workshops (CVPRW), 2017 IEEE Conference on*, pages 1175–1183. IEEE, 2017.
- [12] Mazhar Shaikh, Ganesh Anand, Gagan Acharya, Abhijit Amrutkar, Varghese Alex, and Ganapathy Krishnamurthi. Brain tumor segmentation using dense fully convolutional neural network. In *International MICCAI Brainlesion Workshop*, pages 309–319. Springer, 2017.
- [13] Shu Liu, Lu Qi, Haifang Qin, Jianping Shi, and Jiaya Jia. Path aggregation network for instance segmentation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 8759–8768, 2018.
- [14] Tsung-Yi Lin, Piotr Dollár, Ross B Girshick, Kaiming He, Bharath Hariharan, and Serge J Belongie. Feature pyramid networks for object detection. In *CVPR*, volume 1, page 4, 2017.
- [15] Fausto Milletari, Nassir Navab, and Seyed-Ahmad Ahmadi. V-net: Fully convolutional neural networks for volumetric medical image segmentation. In *3D Vision (3DV), 2016 Fourth International Conference on*, pages 565–571. IEEE, 2016.
- [16] Spyridon Bakas, Hamed Akbari, Aristeidis Sotiras, Michel Bilello, Martin Rozycki, Justin S Kirby, John B Freymann, Keyvan Farahani, and Christos Davatzikos. Advancing the cancer genome atlas glioma mri collections with expert segmentation labels and radiomic features. *Scientific data*, 4:170117, 2017.
- [17] Spyridon Bakas, Hamed Akbari, Aristeidis Sotiras, Michel Bilello, Martin Rozycki, Justin S Kirby, John B Freymann, Keyvan Farahani, and Christos Davatzikos. Advancing the cancer genome atlas glioma mri collections with expert segmentation labels and radiomic features. *Scientific data*, 4:170117, 2017.
- [18] Bjoern H Menze, Andras Jakab, Stefan Bauer, Jayashree Kalpathy-Cramer, Keyvan Farahani, Justin Kirby, Yuliya Burren, Nicole Porz, Johannes Slotboom, Roland Wiest, et al. The multimodal brain tumor image segmentation benchmark (brats). *IEEE transactions on medical imaging*, 34(10):1993–2024, 2015.
- [19] Haocheng Shen, Jianguo Zhang, and Weishi Zheng. Efficient symmetry-driven fully convolutional network for multimodal brain tumor segmentation. In *Image Processing (ICIP), 2017 IEEE International Conference on*, pages 3864–3868. IEEE, 2017.
- [20] Geena Kim. Brain tumor segmentation using deep u-net. In *International MICCAI BraTS Challenge*, pages 154–160, 2017.
- [21] Varghese Alex, Mohammed Safwan, and Ganapathy Krishnamurthi. Brain tumor segmentation from multi modal mr images using fully convolutional neural network. In *International MICCAI BraTS Challenge*, pages 1–8, 2017.
- [22] Xiangmao Kong, Guoxia Sun, Qiang Wu, Ju Liu, and Fengming Lin. Hybrid pyramid u-net model for brain tumor segmentation. In *International Conference on Intelligent Information Processing*, pages 346–355. Springer, 2018.