A squeeze M-SegNet architecture for segmentation of brain tissues on MRI

Nagaraj Yamanakkanavar, Bumshik Lee*
광주광역시 남구 필문대로 309, 조선대학교 IT융합대학 정보통신공학과
nagarajpy@chosun.ac.kr, bslee@chosun.ac.kr (corresponding author)

요 약

Accurate segmentation of brain tissue in a magnetic resonance imaging (MRI) is an important biomarker of medical imaging. Hence, segmentation methods are in research focus and various methods are presented in the literature. In this paper, we proposed the novel M-SegNet architecture with fire module (squeeze and expand layers) for segmentation of brain MRI. The proposed model utilizes long-skip connections, as well as squeeze and expand convolutional layers from the fire module to segment brain MRI. The M-SegNet architecture consists of a multi-scale deep network at the encoder side, deep supervision at the decoder side, and the architecture uses pooling indices along with skip-connections from the encoder to the decoder layer. The multi-scale side input layers are used to support deep layers for extracting the discriminative information and the decoder side provides deep supervision to reduce the gradient problem. The skip-connections are used to pass features from the encoder to the decoder path to recover the spatial information lost during down-sampling and pooling indices helps for faster convergence of the model. Besides, the proposed method results in fewer parameters generation with efficient memory use and, hence is faster to train in comparison with conventional methods. The proposed method was evaluated against widely used segmentation methods on publically available datasets. Experimental results show that the proposed method can segment brain MRI more accurately as compared with several state-of-the-art methods.

Ⅰ. 서 론

Magnetic resonance imaging (MRI) is preferably used for structural analysis of the brain because it gives high spatial resolution and contrast for soft tissues. Generally, it is known that MRI is associated with fewer health risks compared to other modalities like computed tomography and positron emission tomography [1]. Over the years, enormous development has been made in accessing brain injuries and exploring brain anatomy with the aid of MRI [2]. The artifacts in the neurological structure of the brain are the main challenging factor in the process of brain tissue segmentation. At present, medical MRI segmentation mainly focuses on the segmentation of gray matter (GM), white matter (WM), and cerebrospinal fluid (CSF) brain tissues. The manual segmentation by specialists is time-consuming, often prone to human error, and impractical for large datasets. Therefore, developing accurate methods for automated brain-tissue segmentation has become an active research area. To segment the object rapidly and accurately, various machine learning approaches have been developed in the literature such as, [3] based on the clustering algorithm, [4] based on level set methods and [5] based on pattern recognition approaches. However, the performances of these approaches are often hindered due to the complex brain structure, low soft-tissue contrast, non-uniform intensity, partial volume effect, and noise in brain MRI. Recently, deep learning approaches have been made possible to obtain accurate and fast tissue segmentation in medical images. In particular, convolutional neural networks (CNNs) have shown high performance in various applications such as handwritten digit recognition, object detection, and semantic segmentation [6]. Deep learning-based methods enable self-learning of features, thus do not depend on hand-crafted features extraction, while traditional machine learning-based techniques usually extract features, such as the Gaussian or Haar-like kernels. However, the need for a large amount of data for training is one of the main limitations of deep learning-based methods. In particular, collecting such a large volume of labeled training data in the field of medical image analysis is quite challenging [7]. Furthermore, the CNN's have some drawbacks in segmentation applications because, during the segmentation process, reconstruction should be performed using vectors, meaning, one not only needs to convert a feature map into vectors but also needs to reconstruct brain images from vectors. . To overcome these problems, we propose a squeeze M-SegNet architecture for automatic segmentation of brain MRI scans.

Ⅱ. 본론

The conventional methods such as SegNet [8], U-Net [9], and M-net [10] have shortcomings even if they provide good segmentation performances. In order to solve the problems and enhance the performance of the segmentation, we proposed a hybrid architecture of SegNet and M-net, which is, namely, the M-SegNet that captures the best of both the models by using pooling indices from the SegNet and M-net architecture, thus providing multi-scale information through side-paths and skip-connections for better segmentation performance. To achieve better segmentation accuracy, we employ long-skip connections, as well as squeeze and expand convolutional layers from the fire module to segment brain MRI. Furthermore, apart from accuracy, computation complexity is also an important metric to evaluate the overall performance of the network. The computation time is directly proportional to the learnable parameters generated during model training. The lesser is the parameters required, the faster to train the model.

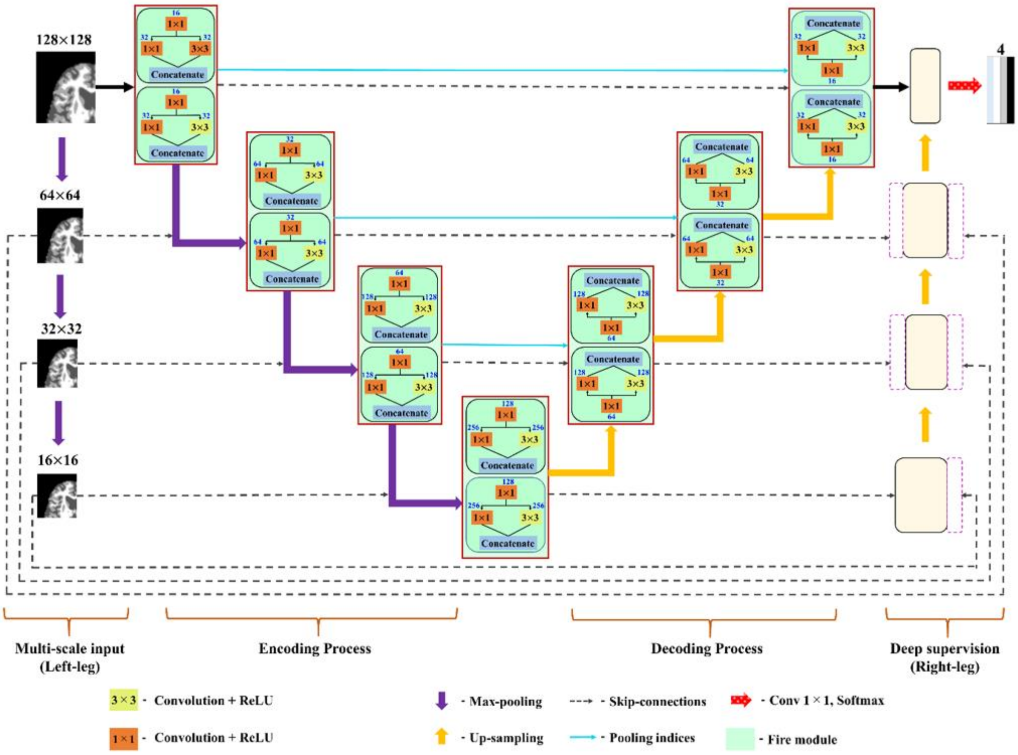

Fig 1. The schematic representation of the proposed method.

Hence, we adopted the fire module to reduce memory requirement and computational complexity. Fig. 1 shows the overall architecture of the proposed model. As shown in Fig. 1, it has an encoder-decoder based structure, with two side paths such as multi-scale input (left-leg) and deep supervision (right-leg). In the encoding path, each level has a cascade of fire modules with convolutional sizes of 1×1 and 3×3. The output of the fire module is down-sampled using 2×2 max-pooling operation with a stride of two. The max-pooling reduces the dimension of the image and helps to capture fine details of the feature map. In the left-leg of the architecture, input image is down-sampled using 2×2 max-pooling operation with a stride of two and concatenated to corresponding encoder layer through side-skip connections. The multi-scale inputs are used to support deep layers for extracting the discriminative information. The feature obtained by each fire module in the encoder layer is transferred to the corresponding decoding layer with a skip connections and pooling indices. The skip connections are used to pass features from the encoder to the decoder path to retrieve spatial information lost during down-sampling. Furthermore, the proposed model has faster convergence because pooling indices are passed to the deconvolution layers. The pooling indices and skip connections used in the proposed architecture which are indicated by blue and dotted gray color arrows as shown in Fig. 1. Each decoding layer consists of two consecutive fire modules. The max-pooling operation at decoder side is replaced by un-pooling layers, which up-samples the input feature maps, without any additional learnable parameters. The long-skip connections are used to stabilize the gradient updates in the network and also it will improve the optimization convergence speed. The deep supervision (right-leg) used to reduce the gradient problem and hence improve the optimization convergence speed. The classification layer consists of 1×1 convolutional layer with a softmax activation function that outputs a reconstructed image. The softmax layers predict four output classes such as GM, WM, CSF, and background. To measure the loss of the proposed model, we used cross-entropy loss.

The proposed method experiments with a publically available open access series of imaging studies (OASIS) dataset [12] for segmentation of brain tissues on MRI. In our, experiment, 50 subjects were selected and the first 30 subjects were used for model training, and the remaining 20 subjects were used to test the data. The proposed method was evaluated against SegNet [8], U-net [9], and M-net [10] architectures. The experiments were performed using the Keras framework on an Nvidia 3090Ti GPU. During training, we set the number of epochs as 10, a validation split of 0.2, the learning rate of 0.001, and used stochastic gradient descent with a high momentum rate of 0.99. To objectively evaluate the performances of the methods, we used the Dice similarity coefficient (DSC) [13] and Jaccard index (JI) [14] to compare segmentation outputs against the ground truths. The DSC and JI used to determine the extent of overlap between a given ground truth segmentation map and predicted segmentation map is defined as (1) and (2),

$$DSC(s,s') = \frac{2|s \cap s'|}{|s|+|s'|} \quad (1)$$

$$JI(s,s') = \frac{|s \cap s'|}{|s \cup s'|} \quad (2)$$

where the term $\cap$ denotes the intersection of the ground truth and predicted segmentation map, and |.| represents the cardinality of the set.

Table 1 Experimental results of the proposed method in comparison with the existing methods using OASIS dataset.

Model	WM		GM		CSF	
	DSC	JI	DSC	JI	DSC	JI
SegNet [8]	0.89	0.83	0.86	0.80	0.85	0.80
U-net [9]	0.93	0.89	0.93	0.88	0.90	0.85
M-net [10]	0.94	0.83	0.93	0.88	0.90	0.85
Proposed method	**0.96**	**0.93**	**0.96**	**0.91**	**0.94**	**0.89**

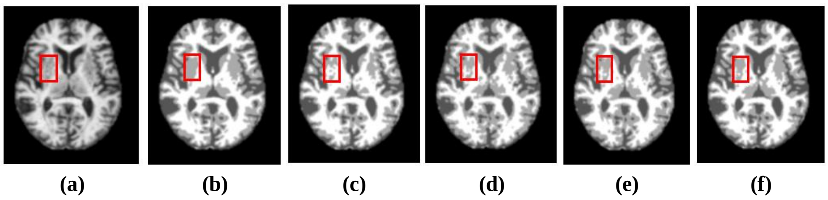

Fig 1. Segmented results of existing methods and proposed method: (a) Original input image, (b) ground truth segmentation map, (c) Predicted segmentation map of SegNet (d) segmentation map of U-net (e) Predicted segmentation map of M-net (f) Predicted segmentation map of proposed method respectively.

Ⅲ. 결론

In this paper, we presented the squeeze M-SegNet architecture that can achieve better performance compared to conventional methods for the segmentation of brain MRI. In the proposed architecture, the multi-scale side input layer helps in extracting discriminative information and deep supervision in the output side is used to speed up the convergence of the model through long-skip connections. Furthermore, the use of fire modules in the encoder and decoder paths of the proposed architecture reduces the learnable parameter and form a memory-efficient segmentation model. Finally, our method demonstrates significant improvement in terms of popular metrics, such as the DSC and JI, for the segmentation of brain MRI scans into CSF, GM, and WM regions, exhibiting average DSC and JI values of 0.95 and 0.90.

ACKNOWLEDGMENT

This work was supported by the National Research Foundation of Korea (NRF) grant funded by Korea government (No. NRF-2019R1I1A3A01058959).

참 고 문 헌

[1] Miriam, R.; Malin, B.; Jan, A.; Patrik, B.; Fredrik, B.; Jorgen, R.; Ingrid, L.; Lennart, B.; Richard, P.; Martin, R.; et al. PET/MRI and PET/CT hybrid imaging of rectal cancer–description and initial observations from the RECTOPET (Rectal Cancer trial on PET/MRI/CT) study. Cancer Imaging 2019, 19, 52.

[2] J. Bernal, K. Kushibar, M. Cabezas, S. Valverde, A. Oliver and X. Lladó, "Quantitative Analysis of Patch-Based Fully Convolutional Neural Networks for Tissue Segmentation on Brain Magnetic Resonance Imaging," IEEE Access, vol. 7, pp. 89986-90002, 2019.

[3] Nagaraj Y., Asha C.S. and A. V. Narasimhadhan, "Assessment of speckle denoising in ultrasound carotid images using least square Bayesian estimation approach," 2016 IEEE Region 10 Conference (TENCON), pp. 1001-1004, 2016.

[4] P. Moeskops, M. A. Viergever, A. M. Mendrik, L. S. De Vries, M. J. N. L. Benders, and I. Isgum, "Automatic Segmentation of MR Brain Images with a Convolutional Neural Network," IEEE Trans. Med. Imaging, 2016.

[5] C. Wachinger, M. Reuter, and T. Klein, "Deepnat: Deep convolutional neural network for segmenting neuroanatomy". NeuroImage, vol.170, pp.434-445, 2018.

[6] Yamanakkanavar N, Choi JY, Lee B. MRI Segmentation and Classification of Human Brain Using Deep Learning for Diagnosis of Alzheimer's Disease: A Survey. Sensors. 20(11), 3243,2020.

[7] W. Jie, Y. Xia and Y. Zhang, "M3Net: A multi-model, multi-size, and multi-view deep neural network for brain magnetic resonance image segmentation", Pattern Recognition, vol. 91, pp.366-378, 2019.

[8] V. Badrinarayanan, A. Kendall, and R. Cipolla, "SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation," IEEE Trans. Pattern Anal. Mach. Intell., 2017.

[9] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," Medical Image Computing and Computer Assisted Intervention-MICCAI 2015, Springer, pp. 234–241, 2015.

[10] V. Adiga., J. Sivaswamy. "FPD-M-net: Fingerprint Image Denoising and Inpainting Using M-net Based Convolutional Neural Networks", Inpainting and Denoising Challenges, Springer, 2019.

[11] N. Yamanakkanavar and B. Lee, "Using a Patch-Wise M-Net Convolutional Neural Network for Tissue Segmentation in Brain MRI Images," in IEEE Access, vol. 8, pp. 120946-120958, 2020.

[12] H. Tracy and L. Randy, "Open Access Series of Imaging Studies (OASIS): Cross-Sectional MRI Data in Young, Middle Aged, Nondemented, and Demented Older Adults". Journal of Cognitive Neuroscience, vol.19, pp.1498-1507, 2007.

[13] L. R. Dice, "Measures of the amount of ecologic association between species," Ecology, vol. 26, no. 3, pp. 297–302, 1945.

[14] P. Jaccard, "The distribution of the flora in the alpine zone.1," New Phytologist, vol. 11, no. 2, pp. 37-50, Feb. 1912.