3차원 뇌 MRI 영상 분할을 위한 밀집 교차 결합 합성 곱 네트워크

판 레 중•, 안 상 일*, 신 지 태°

Cross-Linked Fully Convolution-DenseNet for Volumetric Segmentation of Brain MRI

Trung Le Phan•, Sangil Ahn*, Jitae Shin°

요 약

일반적으로 뇌의 얇은 시트 구조, 강도 비균질성, 강도와 포화 사이의 낮은 대비로 뇌 조직 영상 분할에 어려움이 있다. 위 문제점을 해결하며 자동으로 정확한 뇌 조직 자기공명영상 분할을 위해 이 논문에서 3D FC-DenseNet과 HyperDenseNet의 두 가지 모델의 장점을 기반한 밀집 교차 결합 합성 곱 네트워크를 소개한다. MRBrainS13, iSeg_2019 데이터를 이용한 실험을 통해 교차 결합 합성 곱 네트워크가 기존의 방법보다 정확도 측면에서 더 좋은 성능을 보이는 것을 증명한다.

Key Words : iSeg_2019, Multi-modal imaging, 3D FC-DenseNet, HyperDenseNet, Segmentation.

ABSTRACT

Automatic brain tissue segmentation in brain MRI is one of challenging issues of segmentation task due to thin sheet structure, intensity inhomogeneity and low contrast between intensity and saturation. Addressing the above problems, this paper introduces cross-linked fully convolution-DenseNet(cross-linked FC-DenseNet) based on the advantages of to models, 3D FC-DenseNet and HyperDenseNet, for the purpose of automatically accurate brain tissue MRI segmentation. As a result of experiments using MRBrainS13, iSeg_2019 dataset, the proposed method performs better in terms of accuracy than existing methods.

Ⅰ. Introduction

Different MRIs modalities can be used to brain segment for diagnosis purpose. However, there is a challenging issue due to thin sheet structure, intensity inhomogeneity and low contrast among tissues which can interfere with the correct segmentation. Therefore, developing of accurate segmentation algorithm is important of this research area.

Recently, convolutional Neural Networks (CNNs)[1] have shown to be prospered in several computer vision tasks including the cognition of medical image segmentation. 3D FC-DenseNet[2] can learn local and global features from both high and low resolution on medical image.

In this work, we first modify 3D FC-DenseNet[2]. We add an end-to-end concurrent spatial and channel squeeze & excitation (scSE) followed by every Denseblock that allows it to explore the interdependency between the channels. Secondly, we combine the 3D FC-DenseNet[2] with cross-link[3] which can learn more features from the different layers. The proposed method exploits the multi-level and multi-model features of brain MRIs. There are cross-dense connections not only between in the pairs of layers within the same path, but also between in those across different paths and therefore can capture more feature information[3].

※ 이 논문은 정부(교육부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업임 (NRF-2017R1D1A1B03031752, NRF-2018R1C1B6007462).
• First Author : (ORCID:0000-0002-1831-5065) Department of Electrical and Computer Engineering, Sungkyunkwan University, trungphan@skku.edu, 정회원
° Corresponding Author : (ORCID:0000-0002-2599-3331) Department of Electrical and Computer Engineering, Sungkyunkwan University, jtshin@skku.edu, 정회원
* (ORCID:0000-0002-3632-3974) Department of Electrical and Computer Engineering, Sungkyunkwan University, il2s@skku.edu, 학생회원
논문번호 : 201901-250-C-LU, Received October 24, 2019; Revised December 7, 2019; Accepted December 12, 2019

Ⅱ. Proposed Methods

In Fig. 1. we propose the network which is the combination of three phases: 3D FC-DenseNet concurrent spatial and channel squeeze & excitation (scSE) and HyperDenseNet. The network consists of basic main two paths: contracting and expanding path. The contracting path aims to understand what features are in the object and increase the receptive field. The role of the expanding path is to support more where features are in the object. The model has the capacity to learn both low resolution global features and high resolution local features with cross-link. The initial part of the network has three 3x3x3 convolution layers with stride 1 followed a batch normalization layer (BN) and a ReLU that generate 64 output feature maps. We implement this contracting path with three dense blocks, three scSE blocks, three transition blocks. Each dense block has four BN-ReLU-Conv(1x1x1)-BN- ReLU-Conv(3x3x3) with a growth rate of k = 16. To against the over-fitting after the Conv(3x3x3), we use a dropout layer with a dropout rate of 0.2. After each dense block, we add scSE block to explore the interdependencies between the channels. After scSE block, the transition block includes BN-ReLU-Conv(1x1x1)-BN-ReLU followed by a convolution layer of stride 2 to reduce feature map resolution while preserving the spatial information in limited dataset. Meanwhiles, for recovering the feature resolution, the expanding path has 3 convolution layers. To reuse the feature map information of different size during the expanding path, third convolution layer get outputs from each transition block in contracting path using cross-link path. We resize the feature maps of first transition block and last using max pooling and up-convolution of 2x2x2 by stride 2 to get equal resolution with second transition output and concatenate all of them as input of third convolution layer. Also, second convolution layer use 3 different size of feature map: output of initial part, output of first transition block, output of third convolution layer. Otherwise we use 2 different size of feature map in first convolution layer: output of initial part, output of second convolution layer. This cross-link path allows model to capture the multiple contextual information from different layer's features. Furthermore, it improves gradient flow within a limited dataset. Each convolution layer consists of ReLU-Conv (3 x 3 x 3)-BN. And finally, a classifier consisting of ReLU-Conv (1 x 1 x 1) is target classes.

Ⅲ. Results

Fig. 2 displays segmentation results of the competing methods. The magnified views of the segmentation of the rectangular regions are displayed at the bottom of each image. It shows that

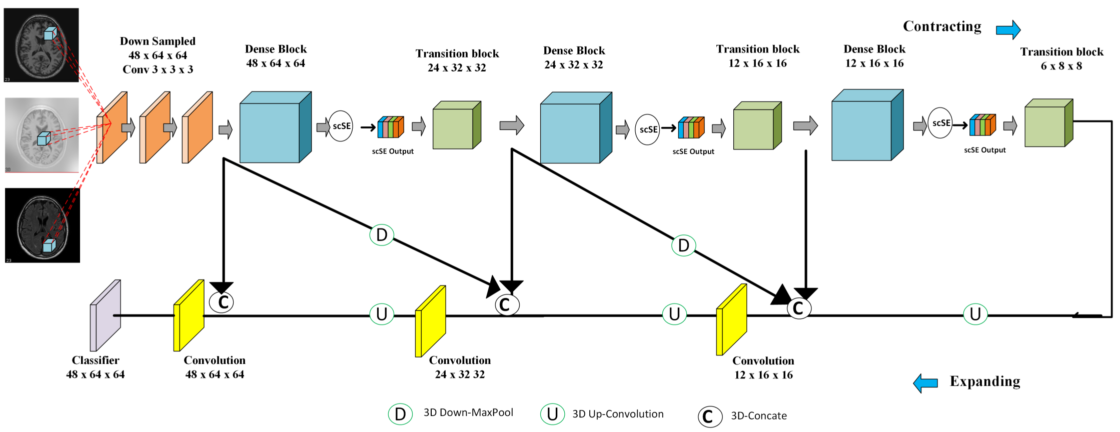

Fig. 1. The proposed cross-linked FC-DenseNet architecture for volumetric segmentation.

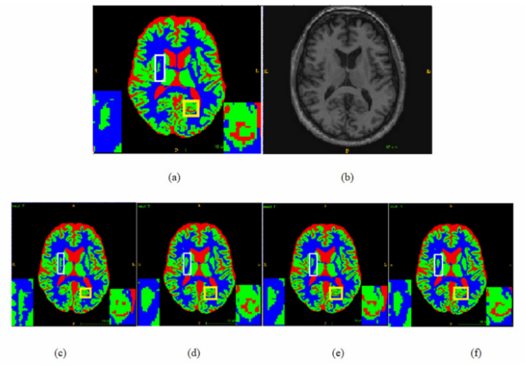

Fig. 2. Results segmentation result of validation data: (a) ground truth (blue, green, red color represent the WM, GM, and CSF respectively), (b)original T1 image, (c) 3D-Unet[1], (d) FC-Dens-eNet[2], (e) FC-DenseNet-scSE, (f) cross-linked FC-DenseNet-scSE.

the proposed method capture the details of the ground-truth labels considerably better than the other methods.

From the Table 1, it's clear that the proposed architecture outperformed the existing architectures. Averaging the DSC accuracy over the brain tissue classes, we note that cross-linked FC-DenseNet-scSE achieved highest accuracy with 87.57%, FC-DenseNet-scSE achieved 87.47%, FC-DenseNet achieved 87.01%, and 3D-Unet achieved 86.88%. The overall score is determined using the sum over all ranks. The magnified views of the segmentation of the rectangular regions are also displayed at the bottom of each image.

In addition, our method achieves good results when applying to the similar iSeg_2019 data set. Our results were in the top of the challenge.

Table 1. Performance on the validation set of the proposed cross-linked FC-DenseNet method and two recent deep learning-based methods: 3D-Unet[1], FC-DenseNet[2] in term of Dice similarity coefficient (DSC:%), Average Distance (AVD:mm).

METHODS	WM		GM		CSF		Avg. DSC
	DSC	AVD	DSC	AVD	DSC	AVD	
3D-Unet	90.7	0.11	87.3	0.16	82.6	0.36	86.88
FC-DenseNet	90.8	0.12	87.6	0.16	82.6	0.33	87.01
FC-DenseNet -scSE	91.3	0.11	88.0	0.15	83.1	**0.25**	87.47
Cross-linked FC-DenseNet -scSE	**91.3**	**0.11**	**88.1**	**0.14**	**83.3**	0.33	**87.57**

Table 2. Results of iSeg-2019 challenge (DSC:%), Average Surface Distance (ASD:mm). Only top 8 teams are show here)

METHODS	WM		GM		CSF	
	DSC	ASD	DSC	ASD	DSC	ASD
QL111111	87.7	0.47	83.5	0.51	83.2	0.55
SmartDSP	86.4	0.52	83.1	0.55	82.9	0.56
FightAutism	86.3	0.52	82.9	0.56	82.6	0.57
Cross-linked FC-DenseNet-scSE (Ours)	86	0.53	82.3	0.57	80.7	0.68
RB	86	0.53	81.4	0.57	80.7	0.7
WorldSeg	85.4	0.54	82.1	0.59	80.2	0.7
PerceptionComputing Lab_HIT	84.4	0.61	81.4	0.61	81	0.63
OxfordIBME	82.8	0.64	77.8	0.62	80.6	0.61

Ⅳ. Conclusion

In this paper, we proposed a fully automatic segmentation method for brain image based on extend and modify 3D FC-DenseNet with adding scSE block and cross-Link. We added an end-to-end scSE followed by every dense-block that allows explores the interdependency between the channels. We put the cross-link after the and to learn features more from many feature maps with difference sizes that is proven through the HyperDense-Net[3]. We further incorporate multi-modality information for accurate brain segmentation. The significant advantage of the proposed network is exhibited by quantitative evaluations on MRBrainS13 and iSeg_2019.

References

[1] Ö. Çiçek, A. Abdulkadir, S. S. Lienkamp, T. Brox, and O. Ronneberger, "3D U-Net: learning dense volumetric segmentation from sparse annotation," *Int. Conf. Medical Image Computing and Computer-assisted Intervention*, Athens, Greece, Oct. 2016.

[2] S. Hashemi, S. P. Prabhu, S. K. Warfield, and A. Gholipour, "Exclusive independent probability estimation using deep 3D fully convolutional DenseNets: Application to

isoIntense infant brain MRI segmentation," in *Proc. Machine Learning Res.*, vol. 102, pp. 260-274, 2019.

[3] D. Jose, et al., "HyperDense-Net: A hyper-densely connected CNN for multi-modal image segmentation," *IEEE Trans. Med. Imaging*, vol. 38, no. 5, pp. 1116-1126, 2018.