Ensembles of Multiple Scales, Losses and Models for Segmentation of Brain Area

Xuhua Ren[1], Dinggang Shen[2] and Qian Wang[1]

[1] Institute for Medical Imaging Technology, School of Biomedical Engineering, Shanghai Jiao Tong University, Shanghai 200030, China
[2] Department of Radiology and BRIC, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599
renxuhua@sjtu.edu.cn

A variety of CNN architectures has shown good results in recent literature. Such architectural choices are very important influence for results. In this work, we utilized three types of CNN architectures: DeepMedic [1], Cascaded Anisotropic CNN [2] and Cascaded Anisotropic Unet [3].

1.1 DeepMedic

Model description: The first architecture we employ is DeepMedic, it is the 11-layers deep, multi-scale 3D CNN for medical image segmentation. The architecture consists of two parallel convolutional pathways that process the input at multiple scales to achieve a large receptive field for the final classification while keeping the computational cost low. Inspired by Residual, the use of residual connection is adopted. We only utilized the default parameter for DeepMedic version during these Challenge.

Training details: The wider variant is trained on larger inputs of width 34 and 22 for the two scales respectively. They are trained with cross-entropy loss, with all hyper-parameters adopted from the original configuration.

1.2 Cascaded Anisotropic Convolutional Neural Networks

Model description: They use three networks to sequentially segment sub-region of brain tumor, and each of these networks deals with a binary segmentation problem. The first network (WNet) segments the whole tumor from four modalities 3D image patches. Then a bounding box of the whole tumor is obtained. The cropped region would be used in the second network (TNet) for enhancing tumor (ET) and non-enhancing tumor (NET) which has the same architecture with WNet. Finally, the third network called ENet would separate ET and NET. Rely on the three networks, they could solve the BRATS segmentation task. Binary segmentation problems and take advantage of the hierarchical structure of tumor sub regions to reduce false positives. And, it uses dilated convolution, residual connection and multi-scale prediction to improve segmentation performance. They also propose to fuse the output of CNNs in three orthogonal views (i.e., WNet, TNet and ENet for three views: axial, sagittal, and coronal) for more robust segmentation of brain tumor.

Training details: Our networks were implemented in Tensorflow [4] using NiftyNet [5]. We used Adaptive Moment Estimation (Adam) for training. Batch size is 5, and maximal iteration 20k.

1.3 U-Net

Model description: We employ one versions of the U-Net architecture in our ensemble models. In this version we use residual block to increase model complexity, where residual skip connections [6] are implemented via summations of the signals in the up-sampling part of the network, all layers use batch normalization [7], RELU [8] and zero-padding. We also use three networks to hierarchically and sequentially segment substructures of brain tumor, and each of these networks deals with a binary segmentation problem, as same as Cascaded Anisotropic Convolutional Neural Networks.

Training details: Our networks were implemented in Tensorflow using NiftyNet. We used Adaptive Moment Estimation (Adam) for training. Batch size is 5, and maximal iteration 20k.

1.4 Hybrid Loss

The **Dice** similarity coefficient (DSC) measures the amount of agreement between two image regions. It is widely used as a metric to evaluate the segmentation performance with the given ground truth in medical images. The DSC is defined in (1), we utilize | | to indicate the number of foreground voxels in the ground truth and segmentation images.

$$DSC = \frac{2|S \cap R|}{|S|+|R|}$$

where S is the segmentation result and R is the corresponding ground truth label. This function however is not differentiable and hence cannot directly be used as a loss function in deep learning. Hence, continuous versions of the Dice score have been proposed that allow differentiation and can be used as loss in optimization schemes based on stochastic gradient descent:

$$LD = -\frac{2\sum_{i}^{N} s_i r_i}{\sum_{i}^{N} s_i + \sum_{i}^{N} r_i}$$

We also address to add a modulating factor $(1-p_t)^{\gamma}$ to the cross-entropy loss, with tunable focusing parameter $\gamma \geq 0$. The focal loss [9] definition as:

$$FL = -1(1-p_t)^{\gamma} log(p_t)$$

Finally, we propose a **hybrid loss** HL which add dice loss and focal loss together.

$$HL = LD + FL$$

1.5 Ensemble

The above models are all trained completely separately. At testing time, each model segments individually an unseen image and outputs its class-confidence maps. We utilize three level ensemble strategies: input image sizes, CNN models and losses. So, we have 13 number of models in the finally. Considering each of Cascade CNN has three direction CNNs, we have 40 single model CNNs for the eight-sub-region segmentation task.

Table 1. Ensemble strategies

Item	DeepMedic	Cascade CNN			Cascade Unet		
Input	34x34x34 & 22x22x22	WNET	TNET	ENET	WNET	TNET	ENET
		144x144x19	72x72x19	64x64x19	144x144x19	72x72x19	64x64x19
		144x144x29	72x72x29	64x64x29	144x144x29	72x72x29	64x64x29
Loss	Entropy	Dice		Hybrid	Dice		Hybrid

1.6 Postprocess

Finally, the above models are all trained completely separately. At testing time, each model segments individually an unseen image and outputs its class-confidence maps. Then we ensembled all the models results into one result with some post processing: get rid of tiny subregions; get rid of sub regions which has large eccentricity.

2 References

1. Kamnitsas, K., Ferrante, E., Parisot, S., Ledig, C., Nori, A.V., Criminisi, A., Rueckert, D., Glocker, B.: DeepMedic for brain tumor segmentation. In: International Workshop on Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries, pp. 138-149. Springer, (Year)
2. Wang, G., Li, W., Ourselin, S., Vercauteren, T.: Automatic brain tumor segmentation using cascaded anisotropic convolutional neural networks. In: International MICCAI Brainlesion Workshop, pp. 178-190. Springer, (Year)
3. Ronneberger, O., Fischer, P., Brox, T.: U-net: Convolutional networks for biomedical image segmentation. In: International Conference on Medical image computing and computer-assisted intervention, pp. 234-241. Springer, (Year)
4. Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., Corrado, G.S., Davis, A., Dean, J., Devin, M.: Tensorflow: Large-scale machine learning on heterogeneous distributed systems. arXiv preprint arXiv:1603.04467 (2016)
5. Gibson, E., Li, W., Sudre, C., Fidon, L., Shakir, D.I., Wang, G., Eaton-Rosen, Z., Gray, R., Doel, T., Hu, Y.: NiftyNet: a deep-learning platform for medical imaging. Computer methods and programs in biomedicine 158, 113-122 (2018)
6. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 770-778. (Year)

7. Ioffe, S., Szegedy, C.: Batch normalization: Accelerating deep network training by reducing internal covariate shift. arXiv preprint arXiv:1502.03167 (2015)
8. Nair, V., Hinton, G.E.: Rectified linear units improve restricted boltzmann machines. In: Proceedings of the 27th international conference on machine learning (ICML-10), pp. 807-814. (Year)
9. Lin, T.-Y., Goyal, P., Girshick, R., He, K., Dollár, P.: Focal loss for dense object detection. arXiv preprint arXiv:1708.02002 (2017)