3D Convolutional Networks for Brain Tumor Segmentation

Adrià Casamitjana, Santi Puch, Asier Aduriz, Elisa Sayrol, and Verónica Vilaplana*

Signal Theory and Communications Department, Universitat Politècnica de Catalunya. BarcelonaTech, Spain
{adria.casamitjana,elisa.sayrol,veronica.vilaplana}@upc.edu

Abstract. This paper presents our work on applying 3D Convolutional Networks for brain tumor segmentation for the BRATS challenge. We are currently experimenting with different 3D fully convolutional architectures. We present preliminary results using these architectures and outline our future steps and experiments, which involve hyperparameter optimization, comparison of the models' performance and implementation of a post-processing stage to eliminate false positive predictions.

1 Introduction

The problem of automatic brain tumor segmentation has attracted considerable attention during the past years due to its high clinical relevance and its challenging nature [1]. Methods are usually categorized in two broad groups: generative models, which rely on prior knowledge about the appearance and distribution of different tissue types and discriminative models, which directly learn the relationship between image features and segmentation labels.

Within the second group, in the last two years there has been an increasing use of deep learning methods (and specifically convolutional neural networks CNN) to tackle the problem, motivated by the state of the art performance of deep learning models in several computer vision tasks. As opposed to classical discriminative models based on hand-crafted features, deep learning models learn a hierarchy of increasingly complex features directly from data, by applying several layers of trainable filters and optional pooling operations. Most of these methods do not completely exploit the available volumetric information but use two-dimensional CNNs, processing 2D slices independently or using three orthogonal 2D patches to incorporate contextual information (see [1, 3, and references therein]). A fully 3D approach is proposed in [2], consisting on a 3D CNN that produces soft segmentation maps, followed by a fully connected 3D CRF that imposes generalization constraints and obtains the final labels.

* This work has been partially supported by the project BIGGRAPH-TEC2013-43935-R, financed by the Spanish Ministerio de Economía y Competitividad and the European Regional Development Fund (ERDF)

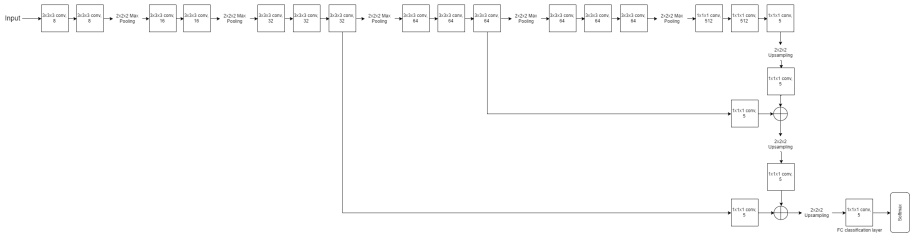

Fig. 1. 3DNet_1

In this work we explore three different 3D CNN architectures to solve the brain tumor segmentation problem. We report preliminary results using the BRATS 2015 Training dataset.

2 Methods

We propose two fully convolutional 3D CNN architectures inspired in two well known 2D models used for generic image segmentation. We also train a third model which is a variant of the two-pathway DeepMedic net proposed in [2].

The first model, 3DNet_1, is a 3D fully convolutional network based on the VGG architecture [8], with skip connections that combine coarse, high layer information with fine, low layer information. The configuration of the net is illustrated in Figure 1. Given the characteristic large amount of parameters of 3D networks, a reduction in the number and dimensions of the filters with respect to its 2D analog was necessary in order to ensure that the model could be trained with the available resources.

The second model, 3DNet_2, is the 3D version of the network proposed in [4]. It is based on the architecture presented in [5], where on top of a VGG net (a contracting path) there is a multilayer deconvolution network (an expansive path). There are connections between corresponding layers in the contracting and expanding paths. The model is illustrated in Figure 2.

The third architecture, 3DNet_3, is a modification of DeepMedic network [2] and is illustrated in Figure 3. The aim of using two paths is gathering both low and high resolution features from the input segment. In a different approach than that of literature, where it is usual to employ different input sizes for each path, we feed segments of equal dimensions to our network. This way, we get coarser features by using larger receptive fields in one path (by means of max-pooling layers) and finer features by using smaller receptive fields (combination of convolutional layers).

3 Implementation details

Preprocessing: We normalize the data within each input volume by subtracting the volume's mean and dividing by the volume's standard deviation.

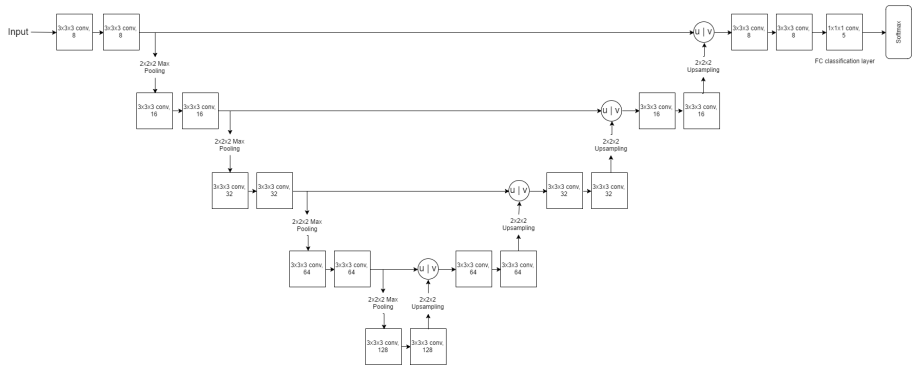

Fig. 2. 3DNet_2

Dataset sampling and class imbalance: Due to memory constraints we do not scan the whole volume in one forward pass but divide it into multiple segments of size 64^3. We adopt the training scheme proposed in [2]: training batches are formed by extracting segments from the training images with 50% probability of being centered on a background or foreground voxel. This scheme has the advantage of automatically mitigating the class imbalance problem. In order to further alleviate this problem we weight the cross-entropy loss function taking into account the class distributions as proposed in [7].

Training: All models use ReLu activations and batch normalization. They are trained using the Adam optimizer, with elastic net regularization and using He initialization for the weights. We use a small batch size of 10 due to memory constraints.

4 Results

Preliminary results obtained with the BRATS 2015 Training dataset are presented in Table 1, using 60% of the data for training and the remaining 40% for test. Recall values are promising for Edema and Enhancing Core classes, but poor for the Necrotic core and Non-enhancing core classes, despite the strategies used for dealing with the class imbalance.

	Accuracy	Dice score			Recall				
		Whole	Core	Active	1-Nec	2-Edm	3-NEnh	4-Enh	0-Else
3DNet_1	99.69	89.64	76.87	63.12	44.71	74.09	28.40	66.94	99.95
3DNet_2	99.71	91.59	69.90	73.89	41.10	84.16	32.35	73.38	99.93
3DNet_3	99.71	91.74	83.61	76.82	51.29	77.50	37.61	87.29	99.95

Table 1. Results for BRATS 2015 Training dataset

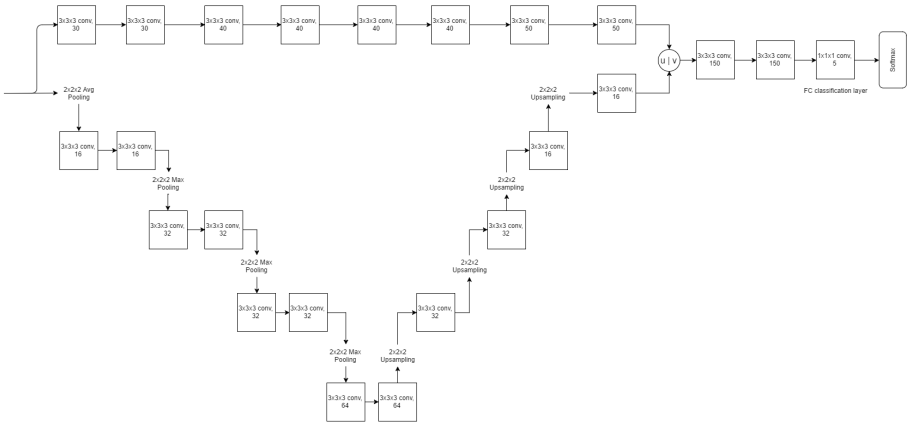

Fig. 3. 3DNet_3

5 Conclusions and future work

In this work we explore the use of 3D CNNs for brain tumor segmentation. The three models are fully connected, being capable of dense-inference, that is making predictions for whole volumes in one pass. In addition, the use of trainable upsampling layers increases the effective batch size without an increase in memory or computational cost.

Future work for the final submission will include: implementation of more elaborated strategies to tackle the class imbalance problem, hyperparameter optimization in order to increase the performance of the models, analysis and comparison of the three architectures and implementation of a post-processing stage to eliminate false positives.

References

1. Menze, B.H., Jakab, A., et al.: The Multimodal Brain Tumor Image Segmentation Benchmark (BRATS). Medical Imaging, IEEE Trans. on 34(10), 1993–2024 (2015)
2. Kamnitsas, K., Ledig, et al.: Efficient Multi-Scale 3D CNN with fully connected CRF for Accurate Brain Lesion Segmentation, arXiv:1603.05959v2 (2016)
3. Havaei, M., Davy, A., et al.: Brain Tumor Segmentation with Deep Neural Networks. arXiv:1505.03540v3 (2016)
4. Ronneberger, O., Fischer, P., Brox, T.: U-Net: Convolutional Networks for Biomedical Image Segmentation. arXiv:1505.04597v1 (2015)
5. Noh, H., Hong, S., Han B.: Learning Deconvolution Network for Semantic Segmentation. ICCV, Santiago, Chile (2015)
6. Long, J., Shelharmer, E., Darrel, T.: Fully Convolutional Networks for Semantic Segmentation. CVPR, Boston, USA (2015)
7. Badrinarayanan, V., Kendall, A., Cipolla, R.: SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation. arXiv:1511.00561 (2015)
8. Simonyan, K., Zisserman, A.: Very Deep Convolutional Networks for Large-Scale Image Recognition. arXiv:1409.1556v6 (2015)