Three Dimensional Fully Convolutional Networks for Segmentation of Optical Coherence Tomography Images in Neurodegenerative Disease

F. Kiaee[1], H. Fahimi[1], R. Kafieh[1], A. U. Brandt[2] , H. Rabbani[1]

1 School of Advanced Technologies in Medicine
Medical Image and Signal Processing Research Center
Isfahan University of Medical Sciences, Isfahan, IRAN
f.kiaee@gmail.com, hamed.fahimi1@gmail.com, rkafieh@gmail.com

2 NeuroCure Clinical Research Centery
Charité - Universitätsmedizin Berlin, Berlin, German
alexander.brandt@charite.de

Abstract

Optical coherence tomography (OCT) is an important method for visualization and quantification of intra-retinal layers. OCT-derived measures of retinal layers support investigating the role of afferent visual pathway degeneration in neurodegenerative diseases like multiple sclerosis (MS). Therefore, accurate, robust and repeatable segmentation of retinal layers is of interest in such applications. In this paper, a novel 3D fully convolutional deep architecture is proposed for automated segmentation of retinal layers. For this purpose, 3D convolutions explore spatial and inter-frame dimensions to extract features. The proposed network uses a set of convolution and subsampling layers in an alternating fashion to learn a hierarchy of shrinking 3D feature maps (encoder stage). The encoder is then followed by multiple convolution and upsampling blocks enlarging the feature map to the size of original input image for semantic segmentation (decoder stage). The proposed framework was validated on 3D OCT scans of healthy subjects captured by a Topcon 3D OCT device. We contrast the ensemble results with the Deep-Net-2D and Graph-DP methods and observe a significant increase of 6% in the Dice metric for two layers and consistent improvements across the retinal layers. Despite the strategies used for dealing with the class imbalance, contour error values are rather inferior for two layers, but still promising for most of the classes. The results are promising for further application of the approach in neurodegenerative diseases.

1 Introduction

Optical coherence tomography (OCT) is able to provide high resolution visualization and quantification of intra-retinal layers by low coherence interferometry. In neurodegenerative diseases like multiple sclerosis (MS), OCT is potent to provide information both during initial diagnosis and monitoring of the disease (1). In this context, segmentation of intra-retinal layers with high accuracy and repeatablity will be quintessential to allow clinical usability. Several previous works have been proposing different approaches for reaching this goal. Conventional segmentation methods are based on a models as prior constraints like (2). In segmentation methods based on artificial intelligence, representative features are extracted to train a classifier, e.g. support vector machines or neural networks, to localize the boundaries. Deep learning methods and i.e. the convolutional neural

network (CNN) , one of its most established realizations, have recently gained increasing interest from various research fields (3). Segmentation of OCT retinal images using deep learning is a new field in medical imaging communities.

Semantic segmentation using deep learning methods is already investigated in methods like (fully convolutional networks (FCN) fine-tuned for segmentation), (encoder-decoder based architecture) , and (U-Net).Recent works in retinal layer segmentation combine the probabilistic predictions of CNN model with a graph search method and a pure end-to-end CNN framework is introduced in (4).

However, the 3D intrinsic information in OCT is ignored in most of the mentioned deep learning based methods despite of being very popular in graph-based approaches. In this paper we propose a 3D deep learning based end-to-end learning framework for segmentation of multiple retinal layers in OCT images. Experimental results on normal datasets show high accuracy and reliability, which makes it a potent candidate to be used in clinical applications in the context of neurodegenerative disorders.

2   Proposed segmentation method

Let $\mathcal{V} = \{(n, m, l), n = 1, \ldots, N, m = 1, \ldots, M, l = 1, \ldots, L\}$ be a 3D volume of locations/voxels, on which observed OCT B-scans and their segmentations, are defined. A segmentation solution is a partition of $\mathcal{V}$ into K exhaustive and mutually exclusive regions. In the sequel, it will be convenient to represent each partition by a label b in the label set $\mathcal{B} = \{b_0, \ldots, b_{K-1}\}$, and treat the OCT segmentation problem as a $K = 9$ class classification. The tissue classes include 7 retinal layers, RaR and RbR regions as upper and lower non of interest regions.The layers have the following anatomical correspondence: The lnner Limiting Membrance (ILM), Nerve Fiber Layer to Inner Plexiform Layer (NFL-IFL), Inner Nuclear Layer (INL), Outer Plexiform Layer (OPL), Outer Nuclear Layer to Inner Segment Myeloid (ONL-ISM), Inner Segment Ellipsoid (ISE) and Outer Segment to Retinal Pigment Epithelium (OS-RPE).

The proposed 3D deep network is constructed by stacking the encoder, decoder, and classification blocks. The encoder block learns a hierarchy of shrinking 3D feature maps. The decoder block enlarges the feature maps to the size of original input image for semantic segmentation. The skip architecture of the network leverage the short-cuts across encoder and decoder to connect the intermittent feature maps with the same dimension from encoder to their corresponding in decoder through concatenation layers. These skip layers combine coarse and semantic information with fine and appearance information. The convolution kernels for all the encoder blocks are 3 dimensional with cubic size $7 \times 3 \times 2$ to ensure that the receptive field at the last encoder block covers the whole retinal region. Finally, a convolutional layer with $1 \times 1 \times 1$ kernels is employed to map its input channels (feature maps) to 9 channels corresponding to 9 classes. At the end, a softmax layer estimates the probability of a voxel belonging to each of the 9 classes.

Cross entropy can be used as a measure of dissimilarity between the predicted probability of voxel i to belong to class b, p_i^b and one-hot encoded true label q_i^b. The cost function is then computed by taking the average of all cross-entropies in the OCT volume. In the experiments, the underrepresented retinal classes are rescaled by an empirically selected weight $w = 5$.

3   Experimental Results

The proposed framework was evaluated on the Isfahan publicly available data set from the Ophthalmology Dept. of Feiz Hospital, Isfahan, Iran. The data set consists of 13 normal 3D macular (spectral domain) SD OCT images from 13 subjects with normal eyes. Ten B-scans per subject were randomly selected and annotated for the retinal layers by an expert clinician. The performance of the proposed 3D method is evaluated against two state-of-the-art retinal OCT layer segmentation algorithms, one using graph-based dynamic programming (Graph-DP)(5) and the other a 2D deep retinal layer segmentation network (Deep-Net-2D)(4).

All the networks were run up to convergence. The experiments were run on a workstation with Intel Xeon CPU, one 11 GB Nvidia GTX 1080 Ti GPU and 32 GB RAM.

Table 1: Results of 8-Fold cross validation on 8 patients and ensemble performance of 8 models on rest two patients.

|		RaR	ILM	NFL-IPL	INL	OPL	ONL-ISM	ISE	OS-RPE	RbR
DICE	Deep-Net-3D (ensemble)	0.95	0.90	0.92	0.83	0.84	0.92	0.91	0.90	0.96
	Deep-Net-3D (single)	0.92	0.85	0.90	0.78	0.79	0.89	0.88	0.86	0.93
	Deep-Net-2D(4)	0.88	0.83	0.88	0.75	0.79	0.83	0.81	0.83	0.91
	Graph-DP(5)	-	0.83	0.86	0.71	0.74	0.84	-	-	-
CE	Deep-Net-3D (ensemble)	-	1.08	1.41	1.48	1.66	1.84	0.98	1.01	-
	Deep-Net-3D (single)	-	1.11	1.51	1.56	1.68	1.91	1.03	1.14	-
	Deep-Net-2D(4)	-	1.12	1.59	1.71	1.73	1.91	1.07	1.21	-
	Graph-DP(5)	-	1.14	1.62	1.68	1.72	1.95	-	1.27	-
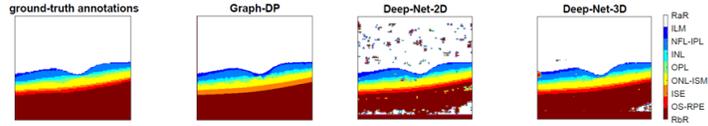
Figure 1: Performance visualization of the methods.

Subjects 1-8 are considered as the training set, and subjects 9-13 are used for the testing phase. A qualitative comparison of the proposed 3D method in contrast with the two comparative methods is presented in Fig. 1.

4   Conclusion

The presented results demonstrated the potential of the proposed method by superior performance on current methods, which makes it a good candidate for accurate segmentation of macular 3D OCTs in neurodegenerative cases. Reliability analysis and further validation will be now performed in a larger data set also containing images from other devices.

References

[1] Axel Petzold, Laura J Balcer, Peter A Calabresi, Fiona Costello, Teresa C Frohman, Elliot M Frohman, Elena H Martinez-Lapiscina, Ari J Green, Randy Kardon, Olivier Outteryck, Friedemann Paul, Sven Schippling, Patrik Vermersch, Pablo Villoslada, Lisanne J Balk, and ERN-EYE IMSVISUAL, "Retinal layer segmentation in multiple sclerosis: a systematic review and meta-analysis.," *The Lancet. Neurology*, vol. 16, no. 10, pp. 797–812, oct 2017.

[2] Raheleh Kafieh, Hossein Rabbani, Michael D Abramoff, and Milan Sonka, "Intra-retinal layer segmentation of 3d optical coherence tomography using coarse grained diffusion map," *Medical image analysis*, vol. 17, no. 8, pp. 907–928, 2013.

[3] Farkhondeh Kiaee, Christian Gagné, and Mahdieh Abbasi, "Alternating direction method of multipliers for sparse convolutional neural networks," *arXiv preprint arXiv:1611.01590*, 2016.

[4] Abhijit Guha Roy, Sailesh Conjeti, Sri Phani Krishna Karri, Debdoot Sheet, Amin Katouzian, Christian Wachinger, and Nassir Navab, "Relaynet: Retinal layer and fluid segmentation of macular optical coherence tomography using fully convolutional network," *arXiv preprint arXiv:1704.02161*, 2017.

[5] Stephanie J Chiu, Xiao T Li, Peter Nicholas, Cynthia A Toth, Joseph A Izatt, and Sina Farsiu, "Automatic segmentation of seven retinal layers in sdoct images congruent with expert manual segmentation," *Optics express*, vol. 18, no. 18, pp. 19413–19428, 2010.