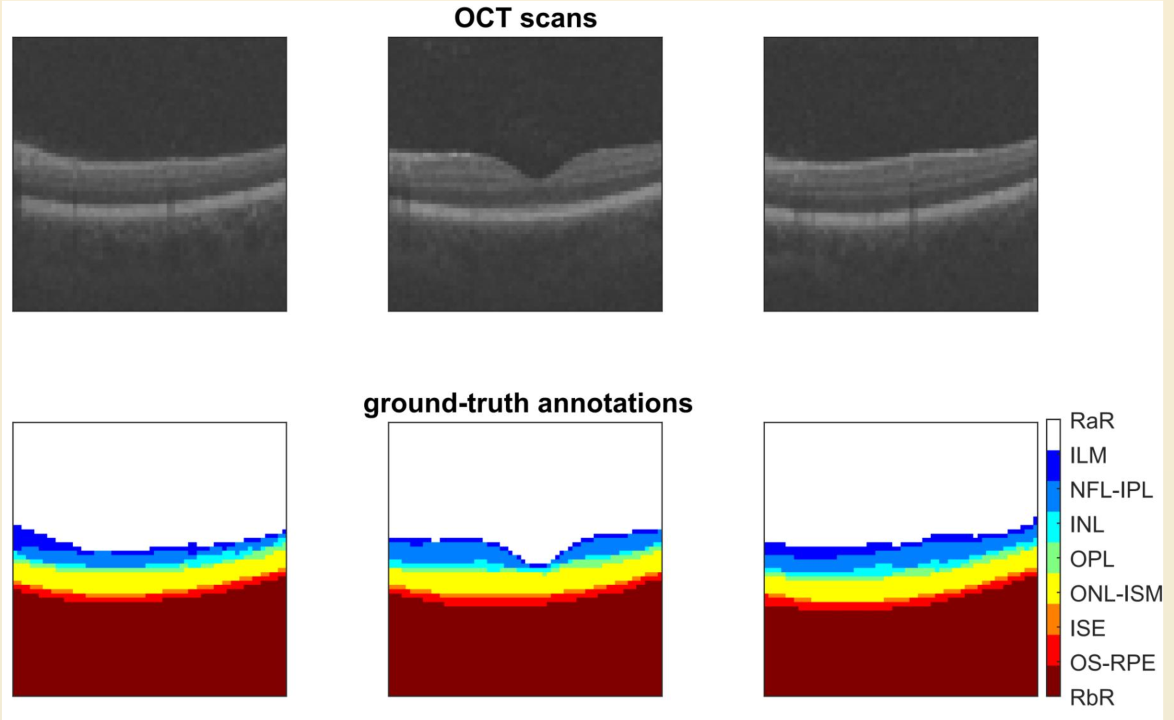
Intra-Retinal Layer Segmentation of Optical Coherence Tomography Using 3D Fully Convolutional Networks

Farkhondeh Kiaee, Hamed Fahimi, and Hossein Rabbani, *Senior Member, IEEE*

Department of Biomedical Engineering, Medical Image and Signal Processing Research Center, School of Advanced Technologies in Medicine,Isfahan University of Medical Sciences, Isfahan, IRAN

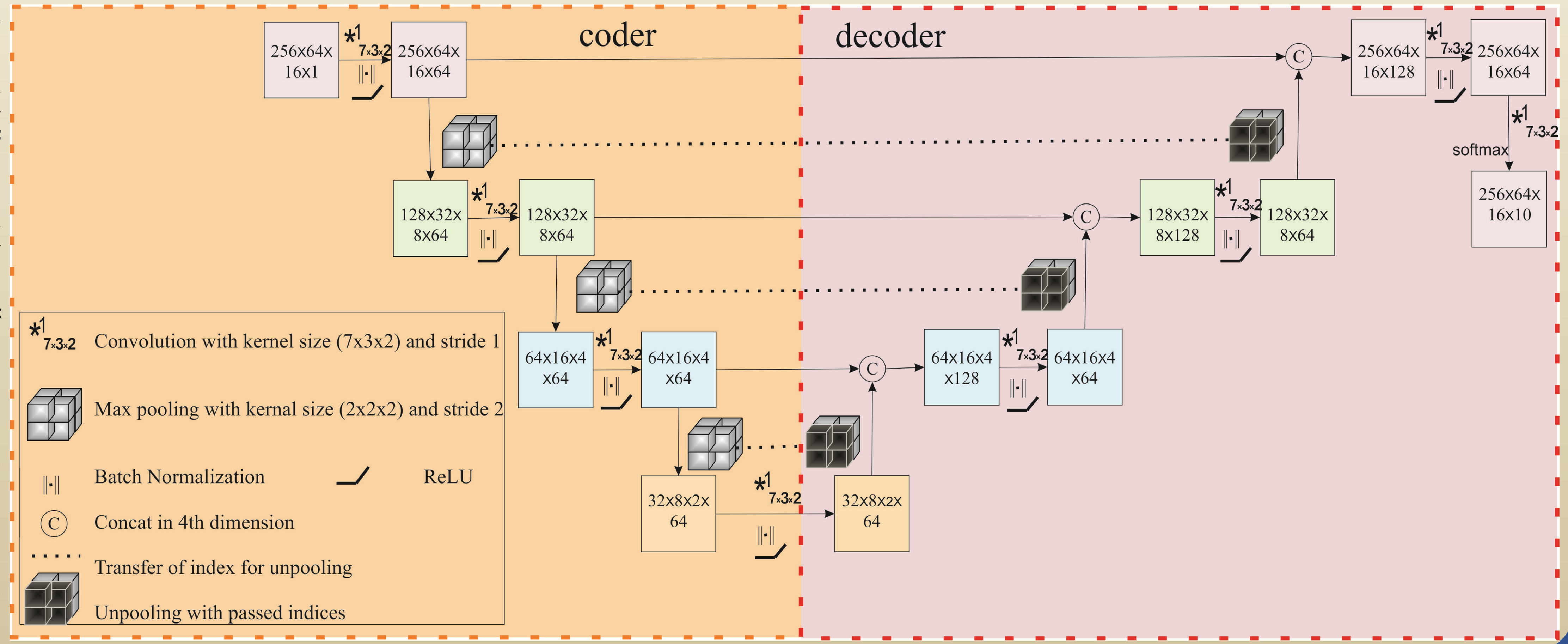
1. Introduction

- The task of segmenting a retinal OCT scan into the constituent retinal layers is one of the enabling steps for many applications of automated retinal OCT analysis.
- Following figure previews 3 frames of 3D OCT images along with ground-truth annotations of 7 retinal layers that we desire to find on volumetric scans.



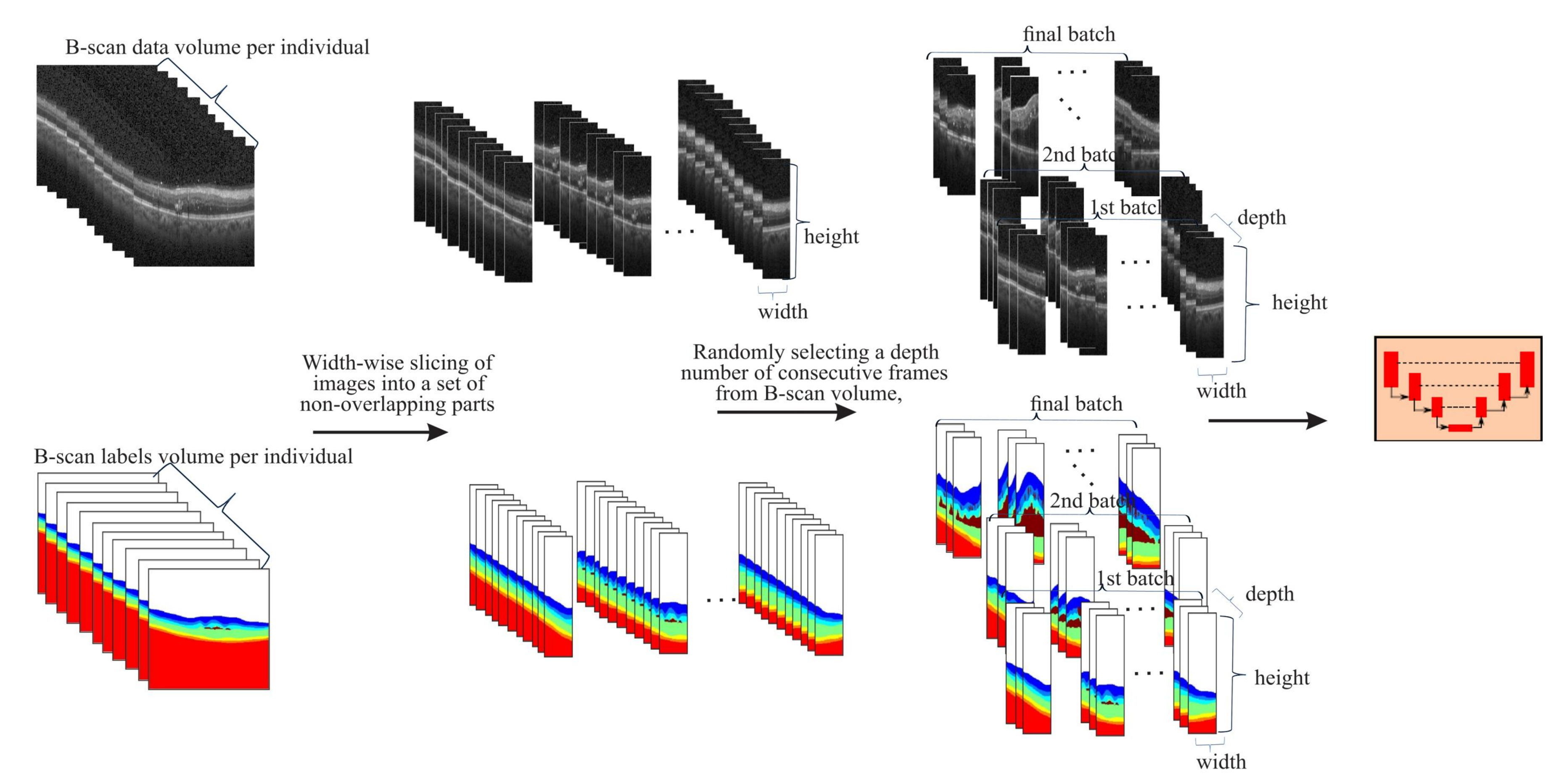
2. Methodology

- We propose an end-to-end 3D fully convolutional architecture for retinal layers segmentation which uses the correlation among nearby frames and makes predictions for the whole OCT volume in one pass.
- The architecture of the proposed 3D deep network is illustrated in following figure. It is constructed by stacking encoder, decoder, & classification blocks. The encoder block learns a hierarchy of shrinking 3D feature maps. Decoder block enlarges the feature maps to the size of original input image for semantic segmentation. Skip layers combine coarse and semantic information with fine & appearance information.



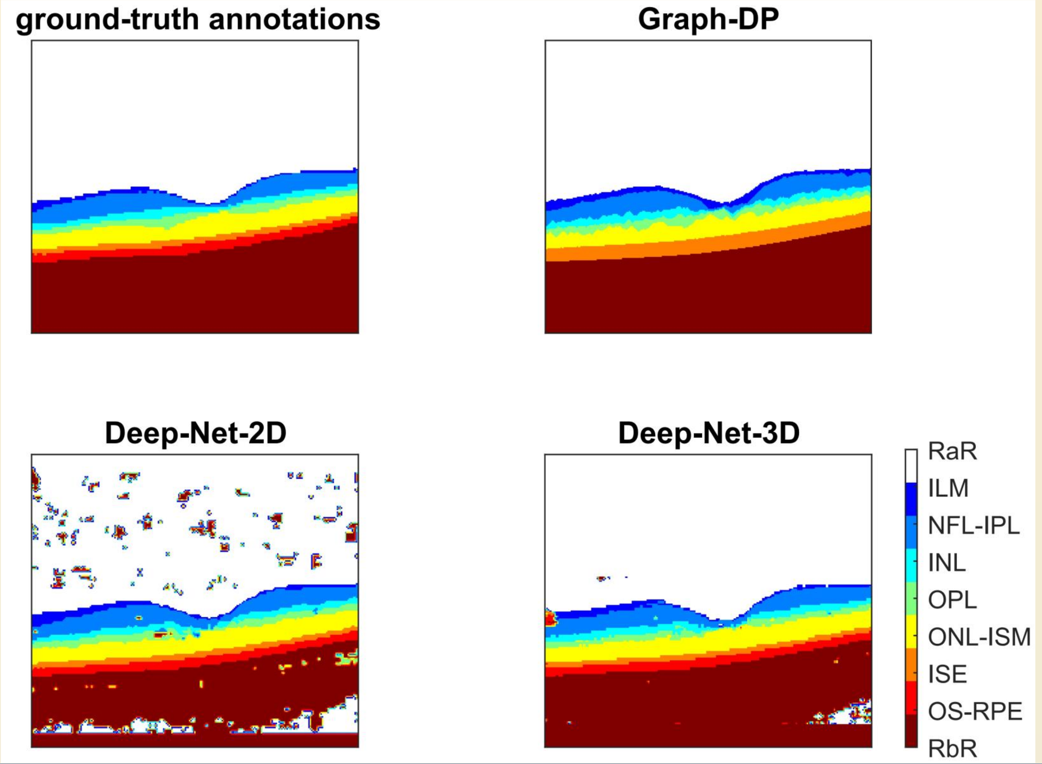
3. Experimental Design

- Training on a big OCT volume of a subject is limited by the RAM of the GPU. In order to make the network fit on a single GPU during training, the OCT volume is sliced width-wise and depth-wise into a determined size of s_{width} =64 and s_{depth}=16, respectively, as depicted in the following figure.
- The proposed framework is evaluated on the Isfahan publicly available dataset from the Ophthalmology Dept. of Feiz Hospital, Isfahan, Iran. The dataset consists of thirteen 3D macular SD OCT images from 13 subjects with size 512× 650×128 (i.e. 128 B-scans per subject). 10 B-scans per subject were randomly selected and annotated for the retinal layers by an expert clinician.
- In the first place, subjects 1-8 in the dataset is considered as the training set and subjects 9-13 is used for the testing phase.



4. Qualitative Results

- The ensemble model is constructed using 8-folded cross-validation within the training dataset. We contrast the ensemble results with the graph-based dynamic programming (Graph-DP) and a 2D deep retinal layer segmentation network (Deep-Net-2D). A qualitative comparison is presented in the following figure.



5. Quantitative Results and Conclusion

- In this paper, a 3D deep learning based end-to-end learning framework was proposed for segmentation of multiple retinal layers in OCT images. The single and ensemble performance of the proposed model evaluated with standard metrics i.e.. dice score and layer contour error (CE) are tabulated in the following table.
- The method outperformed two state-of-the-art retinal layer segmentation i.e. the Deep-Net-2D and Graph-DP by a significant increase of 6% in the Dice metric for OPL and INL layers and consistent improvements across the retinal layers. Despite the strategies used for dealing with the class imbalance, CE values are rather inferior for OPL and INL classes, but still promising for most of the classes.

		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[15]	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[15]	-	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	-