Retinal layers segmentation using Fully Convolutional Network in OCT images

Avi Ben-Cohen[1⋆], Dean Mark[1⋆], Ilya Kovler[1], Dinah Zur[2], Adiel Barak[2], Matias Iglicki[3], and Ron Soferman[1]

[1]RSIP Vision
[2]Division of Ophthalmology, Tel Aviv Sourasky Medical Center, Sackler Faculty of Medicine, Tel Aviv University, Tel Aviv, Israel
[3]Ophthalmology Department, Hospital de Clnicas Jos de San Martn, Buenos Aires, University of Buenos Aires, Argentina

Abstract. In this work we explore a fully convolutional network (FCN) for the task of retinal layers segmentation on retinal optical coherence tomography (OCT) images. FCN has proven to be a powerful tool for many medical applications in recent years. The FCN is trained to segment the different layers based on manually annotated data. At test time, the probability maps for the different layers are extracted from each test case. Next, Sobel edge filter along with graph search methods are used to detect each layers boundary. The proposed method was validated on 24 patients with different pathologies including epiretinal membrane (ERM) and Diabetic Macular Edema (DME), and shows promising results compared to the state of the art.

Keywords: OCT, Deep learning, Fully convolutional network

1 Introduction

Optical coherence tomography (OCT) has been widely used in a variety of medical imaging applications [9]. Today, it is an indispensable tool for the diagnosis and follow up of macular and neurodegenerative diseases, such as age-related macular degeneration, diabetic macular edema, glaucoma and many more. Exact measurement of different layer thickness and accurate comparison over time is an important tool for the evaluation of disease progression. Manual segmentation of these layers can be very time consuming and inconsistent for different experts. This highlights the need for computerized analysis to assist clinicians in the segmentation of the retinal layers.

Several OCT retinal segmentation algorithms have been proposed in recent years. Graph based methods have been used in many of the works for extraction of the boundaries between the different layers [4,10]. Other works formulated this problem as a classification problem using several machine learning techniques. Features extracted from the different layers or layer boundaries were used for training a classifier [12].

⋆ These authors contributed equally.

In recent years, deep learning has become a dominant research topic in numerous fields. Specifically, Convolutional Neural Networks (CNN) [13] obtained outstanding performance on different tasks including many medical applications [3, 7, 16]. Deep CNNs are fully data-driven and can retrieve hierarchical features automatically by building high-level features from low-level ones, thus obviating the need to manually customize hand-crafted features. CNN has been recently used for automatic retinal layer boundaries segmentation and showed promising results [6]. The CNN was trained using patches taken out of the relevant boundary pixels.

In this paper we used a fully convolutional architecture [14] for retinal layers segmentation in OCT images. Fully convolutional networks (FCN) can take input of arbitrary size and produce correspondingly-sized output with efficient learning. Unlike patch based methods, the loss function using this architecture is computed over the entire image segmentation result pixel-wise and includes more global context. Processing entire images instead of patches removes the need to select representative patches, eliminates redundant calculations where patches overlap, and therefore scales up more efficiently with image resolution. Moreover, skip layers are used to combine information across different scales. The output of this method is the different layers probability map. We next used Dijkstra's algorithm [5] to find each layer boundaries based on the network's output. The purpose of this study was to explore the segmentation performance on four retinal layers which were more difficult to handle using previous mentioned techniques.

2 Fully Convolutional Network Architecture

In the following we provide details on the network architecture as well as the the training process.

2.1 U-Net

The network architecture is U-Net based [16]. Let $C_{k,s}$ denote a Convolution-ReLU layer and $CB_{k,s}$ denote a Convolution-ReLU-BatchNorm with k filters of size $s \times s$, $CD_{k,s}$ denotes a Convolution-ReLU-Dropout layer and $CBD_{k,s}$ denotes a Convolution-ReLU-BatchNorm-Dropout with a dropout rate of 50%. $CBP_{k,s}$ denotes a Convolution-ReLU-BatchNorm-Maxpool with pooling size of 2×2, and U_f denotes an Upsampling (deconvolution) layer by factor f.

The "U-Net" encoder:
$C_{64,3} \rightarrow CBP_{64,3} \rightarrow C_{128,3} \rightarrow CBP_{128,3} \rightarrow C_{256,3} \rightarrow CBP_{256,3} \rightarrow C_{512,3} \rightarrow CBP_{512,3} \rightarrow C_{1024,3} \rightarrow CB_{1024,3}$.

The "U-Net" decoder:
$U_2 \rightarrow C_{512,3} \rightarrow CB_{512,3} \rightarrow U_2 \rightarrow C_{256,3} \rightarrow CBD_{256,3} \rightarrow U_2 \rightarrow C_{128,3} \rightarrow CBD_{128,3} \rightarrow U_2 \rightarrow CD_{64,3} \rightarrow C_{64,3} \rightarrow C_{4,1}$.

The "U-net" includes skip connections between every second layer i in the encoder and layer $n-i-1$ in the decoder, where n is the total number of layers. The skip connections concatenate activations from layer i to layer $n-i-1$.

2.2 Training

Input images and their corresponding segmentation maps are used to train the network using Adam optimizer [11] with GPU acceleration. We concentrated on four retinal layer complexes: RNFL, GCL+IPL, INL, and OPL. Figure 1 shows the mentioned layers. The regions surrounding these layers were ignored during the training process. All images were aligned using the RPE boundary segmentation and Bilateral filter [17] was used to reduce noise while preserving strong edges in the image.

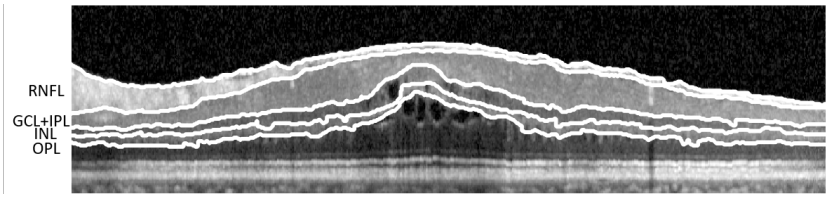

Fig. 1. OCT image with the four retinal layers used for training and testing: RNFL, GCL+IPL, INL, and OPL. The RPE boundary segmentation was used to align the images.

Cross entropy loss function was used following a pixel-wise soft-max over the network final output. The learning rate was chosen to be 0.001 for the first 10 epochs and 0.0001 for the last 10 epochs (total of 20 epochs). The weight decay was chosen to be 0.0005 and the momentum parameter was 0.9.

2.3 Data Augmentation

Data augmentation is essential to teach the network the desired invariance and robustness properties. We generate different scales from 0.8 to 1.2 as layers size can change significantly. Moreover, gamma correction was applied with values ranging from 0.8 to 1.2. The values are sampled using uniform distribution. For each image in our dataset, four augmentations were created during training time.

3 Dijkstra Based Boundary Segmentation

The output from the trained FCN model is a per-pixel and per-class probability map. For each class probability map we extract the boundary using Sobel filter

[1]. Sobel filter can be used to distinguish between the upper and the lower boundaries. We make use of the fact that we know the order of the layers to retrieve each layer boundary one at a time, ignoring regions that are not relevant after finding a layer boundary. We use Dijkstra to find the shortest path that represents each specific layer boundary. Figure 2 shows an illustration of this process.

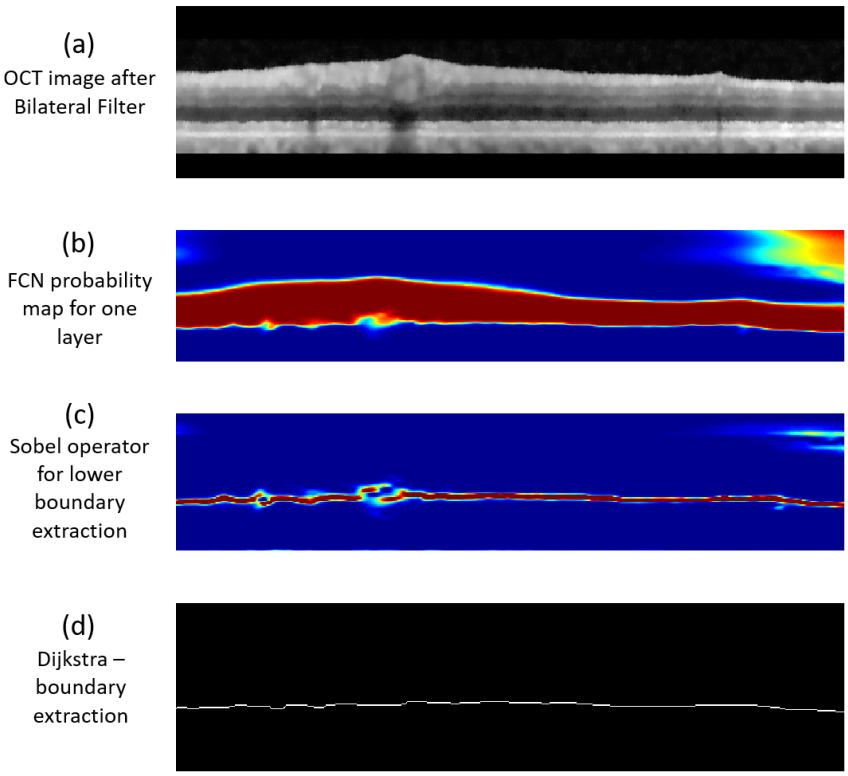

Fig. 2. The proposed framework for one retinal layer. (a) An OCT image after RPE layer alignment and bilateral filter; (b) The FCN probability map extracted for one layer; (c) Sobel operator is used to extract the lower boundary of the layer; (d) Dijkstra algorithm is used to achieve the final boundary.

4 Experiments

4.1 Data

Our dataset included 24 OCT B-scans from 24 different patients with 734 images collected from the retina clinic at the Tel Aviv Sourasky Medical Center. The study was approved by the Ethics committee review board of the Tel Aviv Sourasky Medical Center and conducted in adherence with the Declaration of Helsinki. Six of the patients were diagnosed as having epiretinal membrane (ERM), six patients were diagnosed with Diabetic Macular Edema (DME), while the remaining 12 patients were healthy. After pupil dilation, all patients underwent spectral-domain OCT (Spectralis; Heidelberg Engineering, Heidelberg, Germany). Images were acquired using horizontal raster pattern scans, which were obtained via a 30 x 25 scan field, composed of 31 horizontal B-scans, spaced at 238 μm. Each B-scan was set to average 9 frames each. The layers were initially segmented using the OCTSEG software [15], manual corrections were made by an expert ophthalmologist. Since our dataset was relatively small, we used 3-fold cross validation to test the proposed method's performance.

Table 1. Thickness Differences (mean and standard deviation in pixels) and DICE Index between manual grading and automated grading using U-Net, patch-based CNN, and OCTExplorer.

Retinal Layer	U-Net			Patch-based CNN			OCTExplorer		
	Mean Diff.	STD	DICE	Mean Diff.	STD	DICE	Mean Diff.	STD	DICE
RNFL	1.12	1.39	0.95	1.36	1.66	0.94	3.65	2.62	0.85
GCL+IPL	1.85	1.87	0.92	2.34	2.05	0.89	2.43	2.11	0.78
INL	1.85	1.66	0.85	2.13	1.68	0.81	1.91	1.70	0.62
OPL	1.60	1.60	0.86	1.94	1.76	0.83	3.19	2.17	0.84

4.2 Segmentation- Comparative Evaluation

We compared the FCN to the patch based method. We designed a patch-based CNN similar to the one introduced by Fang et al [6] to segment the four retinal layers. In this method the CNN is trained to provide the boundaries probability maps and Dijkstra is used to recover each layer boundaries. Another comparison was made to the publicly available OCTExplorer software [2]. The mean thickness difference (in pixels) between the automated and manual segmentation was computed for the four layers of interest. Next, the absolute value of the differences was used, and its mean and standard deviation was calculated across all cases. Moreover, DICE index was used for each layer to evaluate the layer segmentation performance. Table 1 presents the results for each layer using the FCN, the patch-based CNN, and OCTExplorer. The U-Net architecture

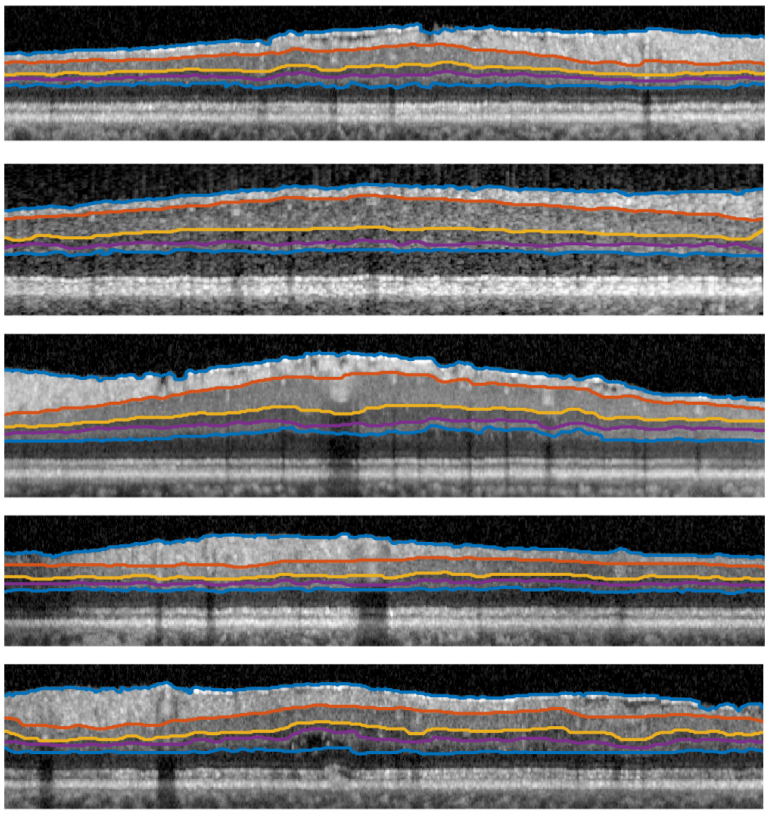

Fig. 3. Example results of the proposed method.

provided better results in all layers in both thickness measures and DICE index. Figure 3 shows examples of our proposed method results using the U-Net architecture.

5 Conclusions

Our objective was to explore two deep learning based methods for the segmentation of the retinal layers. The U-Net based fully convolutional model used in our work provided better results in all layers compared to the patch-based CNN. This is a preliminary work, additional experiments should be conducted using a larger dataset including additional pathologies, and exploring different fully convolutional architectures. Furthermore, our model can be easily extended to segment additional retinal layers, this extension can be explored in future experiments.

While finishing this manuscript we became aware of the preprint [8] which has some overlap with our results.

References

1. Abdou, I. E., and Pratt, W. K. (1979). Quantitative design and evaluation of enhancement/thresholding edge detectors. Proceedings of the IEEE, 67(5), 753-763, (1979).
2. Abramoff, M. D., Garvin, M. K., and Sonka, M.: Retinal imaging and image analysis. IEEE reviews in biomedical engineering, 3, 169-208, (2010).
3. Ben-Cohen, A., Diamant, I., Klang, E., Amitai, M., and Greenspan, H.: Fully Convolutional Network for Liver Segmentation and Lesions Detection. In International Workshop on Large-Scale Annotation of Biomedical Data and Expert Label Synthesis. pp. 77-85, (2016).
4. Chiu, S. J., Li, X. T., Nicholas, P., Toth, C. A., Izatt, J. A., and Farsiu, S.: Automatic segmentation of seven retinal layers in SDOCT images congruent with expert manual segmentation. Optics express, 18(18), 19413-19428,(2010).
5. Dijkstra, E. W.: A note on two problems in connexion with graphs. Numerische mathematik, 1(1), 269-271, (1959).
   Choroid Segmenta
6. Fang, L., Cunefare, D., Wang, C., Guymer, R. H., Li, S., and Farsiu, S.: Automatic segmentation of nine retinal layer boundaries in OCT images of non-exudative AMD patients using deep learning and graph search. Biomedical Optics Express, 8(5), 2732-2744, (2017).
7. Setio, A.A., Ciompi, F., Litjens, G., Gerke, P., Jacobs, C., van Riel, S., Wille, M.W., Naqibullah, M., Sanchez, C. and van Ginneken, B.: Pulmonary nodule detection in CT images: false positive reduction using multi-view convolutional networks. IEEE Transactions on Medical Imaging, pre-print, (2016).
8. Guha Roy, A., Conjeti, S., Phani Krishna Karri, S., Sheet, D., Katouzian, A., Wachinger, C., and Navab, N.: ReLayNet: Retinal Layer and Fluid Segmentation of Macular Optical Coherence Tomography using Fully Convolutional Network. arXiv:1704.02161, (2017).

9. Huang, D., Swanson, E. A., Lin, C. P., Schuman, J. S., Stinson, W. G., Chang, W. and Fujimoto, J. G.: Optical coherence tomography. Science (New York, NY), 254(5035), 1178, (1991).
10. Keller, B., Cunefare, D., Grewal, D. S., Mahmoud, T. H., Izatt, J. A., and Farsiu, S.: Length-adaptive graph search for automatic segmentation of pathological features in optical coherence tomography images. Journal of Biomedical Optics, 21(7), 076015-076015,(2016).
11. Kingma, D., and Ba, J.: Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980, (2014).
12. Lang, A., Carass, A., Hauser, M., Sotirchos, E. S., Calabresi, P. A., Ying, H. S., and Prince, J. L.: Retinal layer segmentation of macular OCT images using boundary classification. Biomedical optics express, 4(7), 1133-1152,(2013).
13. LeCun, Y., Bottou, L., Bengio, Y. and Haffner, P.: Gradient-Based Learning Applied to Document Recognition. Proceedings of the IEEE, 86, 2278-2324, (1998).
14. Long, J., Shelhamer, E. and Darrell, T.: Fully convolutional networks for semantic segmentation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 3431-3440),(2015).
15. Mayer, M. A., Hornegger, J., Mardin, C. Y., and Tornow, R. P.: Retinal Layer Segmentation on OCT-Volume Scans of Normal and Glaucomatous Eyes. Investigative Ophthalmology and Visual Science, 52(14), 3669-3669, (2011).
16. Ronneberger, O., Fischer, P., and Brox, T.: U-net: Convolutional networks for biomedical image segmentation. Springer. In International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 234-241 , (2015).
17. Tomasi, C., and Manduchi, R.: Bilateral filtering for gray and color images. IEEE. Sixth International Conference In Computer Vision, .pp. 839-846, (1998).