

# Automated Segmentation of Retinal Fluid Volumes From Structural and Angiographic Optical Coherence Tomography Using Deep Learning

Yukun Guo<sup>1</sup>, Tristan T. Hormel<sup>1</sup>, Honglian Xiong<sup>1,2</sup>, Jie Wang<sup>1,3</sup>, Thomas S. Hwang<sup>1</sup>, and Yali Jia<sup>1,3</sup>

<sup>1</sup> Casey Eye Institute, Oregon Health & Science University, Portland, OR, USA

<sup>2</sup> School of Physics and Optoelectronic Engineering, Foshan University, Foshan, Guangdong, China

<sup>3</sup> Department of Biomedical Engineering, Oregon Health & Science University, Portland, OR, USA

**Correspondence:** Yali Jia, Casey Eye Institute, Oregon Health & Science University, Portland, OR 97239, USA. e-mail: [jjaya@ohsu.edu](mailto:jjaya@ohsu.edu)

**Received:** March 25, 2020

**Accepted:** September 7, 2020

**Published:** October 8, 2020

**Keywords:** retinal fluid volume; segmentation; OCT/OCTA; deep learning

**Citation:** Guo Y, Hormel TT, Xiong H, Wang J, Hwang TS, Jia Y. Automated segmentation of retinal fluid volumes from structural and angiographic optical coherence tomography using deep learning. *Trans Vis Sci Tech.* 2020;9(2):54. <https://doi.org/10.1167/tvst.9.2.54>

**Purpose:** We proposed a deep convolutional neural network (CNN), named Retinal Fluid Segmentation Network (ReF-Net), to segment retinal fluid in diabetic macular edema (DME) in optical coherence tomography (OCT) volumes.

**Methods:** The 3- × 3-mm OCT scans were acquired on one eye by a 70-kHz OCT commercial AngioVue system (RTVue-XR; Optovue, Inc., Fremont, CA, USA) from 51 participants in a clinical diabetic retinopathy (DR) study (45 with retinal edema and six healthy controls, age 61.3 ± 10.1 (mean ± SD), 33% female, and all DR cases were diagnosed as severe NPDR or PDR). A CNN with U-Net-like architecture was constructed to detect and segment the retinal fluid. Cross-sectional OCT and angiography (OCTA) scans were used for training and testing ReF-Net. The effect of including OCTA data for retinal fluid segmentation was investigated in this study. Volumetric retinal fluid can be constructed using the output of ReF-Net. Area-under-receiver-operating-characteristic-curve, intersection-over-union (IoU), and F1-score were calculated to evaluate the performance of ReF-Net.

**Results:** ReF-Net shows high accuracy (F1 = 0.864 ± 0.084) in retinal fluid segmentation. The performance can be further improved (F1 = 0.892 ± 0.038) by including information from both OCTA and structural OCT. ReF-Net also shows strong robustness to shadow artifacts. Volumetric retinal fluid can provide more comprehensive information than the two-dimensional (2D) area, whether cross-sectional or en face projections.

**Conclusions:** A deep-learning-based method can accurately segment retinal fluid volumetrically on OCT/OCTA scans with strong robustness to shadow artifacts. OCTA data can improve retinal fluid segmentation. Volumetric representations of retinal fluid are superior to 2D projections.

**Translational Relevance:** Using a deep learning method to segment retinal fluid volumetrically has the potential to improve the diagnostic accuracy of diabetic macular edema by OCT systems.

## Introduction

Diabetic macular edema (DME) is the most common cause of vision loss in diabetic retinopathy (DR).<sup>1</sup> Accurate detection of DME for screening and treatment response is critical in preventing vision loss.<sup>2</sup> Currently, clinicians use structural optical coherence

tomography (OCT) to diagnose DME from retinal thickness maps, central macular thickness (CMT), and qualitative inspection of the raster scans.<sup>3</sup> CMT is an imperfect biomarker for DME, because atrophy caused by cell loss can reduce the thickness in the presence of edema, and the presence of other pathology such as epiretinal membranes can increase the thickness without edema.<sup>4,5</sup> Segmentation and quantification of

retinal fluid cysts provide a more specific biomarker for DME that is useful even when other confounding abnormalities are present.<sup>6</sup>

For a more accurate measurement of retinal fluid cysts, several reading centers have explored the measurement of retinal fluid area on cross-sectional commercially available OCT.<sup>7–9</sup> However, quantification of fluid based on the quantity of fluid present may not be accurately captured by cross-sections at large step intervals. Retinal fluid cysts are inherently three-dimensional, and measurements using CMT and cross-sectional OCT cannot precisely measure volumes, only projected areas. With increasing laser speed, the sampling density in commercially available OCT/OCTA volumes is getting higher, which provides a basis for volumetric measurement of retinal fluid. Previously, we presented a fuzzy level-set method<sup>10</sup> to measure retinal fluid volume in OCT/OCT angiography (OCTA) scans of DME eyes. This approach, however, was vulnerable to the shadow artifacts caused by vitreous floaters and large vessels, as well as pupil vignetting.

Contemporary deep-learning-based methods have shown a great advantage for image segmentation tasks.<sup>11–18</sup> In ophthalmology, researchers have proposed a number of deep neural networks to solve specific problems, such as retinal layer segmentation in OCT,<sup>19–21</sup> retinal vessel segmentation in fundus photography,<sup>22,23</sup> choroidal neovascularization segmentation,<sup>24</sup> high-resolution reconstruction on OCT angiograms,<sup>25</sup> and retinal nonperfusion area segmentation in OCTA.<sup>26–28</sup> Deep-learning-based retinal fluid segmentation on cross-sectional OCT has also been reported by many scholars. Bai et al.<sup>29</sup> use a fully convolutional neural network (CNN) and a fully connected conditional random field method to segment cystoid macular edema. This method can get good segmentation results when the fluid deposits are extensive, but it is not sensitive to small target regions. Schlegl et al.<sup>30</sup> propose an encoder-decoder-based deep learning method to detect and quantify macular fluid in OCT images that achieved high accuracy. To solve the challenges due to speckle noise and imaging artifacts, Girish et al.<sup>31</sup> add denoising and subretinal layer segmentation during preprocessing before feeding the data to a CNN, which improve performance. Denoising also helped their algorithm perform well on data from different instruments. Li et al.<sup>32</sup> apply a three-dimensional (3D) CNN on Spectralis OCT (Heidelberg Engineering Inc., Heidelberg, Germany) scans and achieved high performance, but the sparse sampling density hindered the accurate measurement of fluid volume. Some researchers tried to combine a deep-learning-based method with a

traditional image processing method to get a better segmentation result.<sup>33</sup> However, all of these methods segment retinal fluid from OCT data alone. We hypothesize that OCTA signal could improve segmentation accuracy, because retinal fluid and blood flow are never collocal.

In this study, we propose a new deep CNN, named Retinal Fluid Segmentation Network (ReF-Net), to segment intraretinal and subretinal fluid from simultaneously generated volumetric OCT/OCTA scans. Our network provides three key innovations: first, we include OCTA data in the network input. As part of this work, we characterized the effect of this inclusion on network performance. Second, we provide 3D segmentation results and data representations. Last, our network shows strong, robust performance with different types of shadow artifacts.

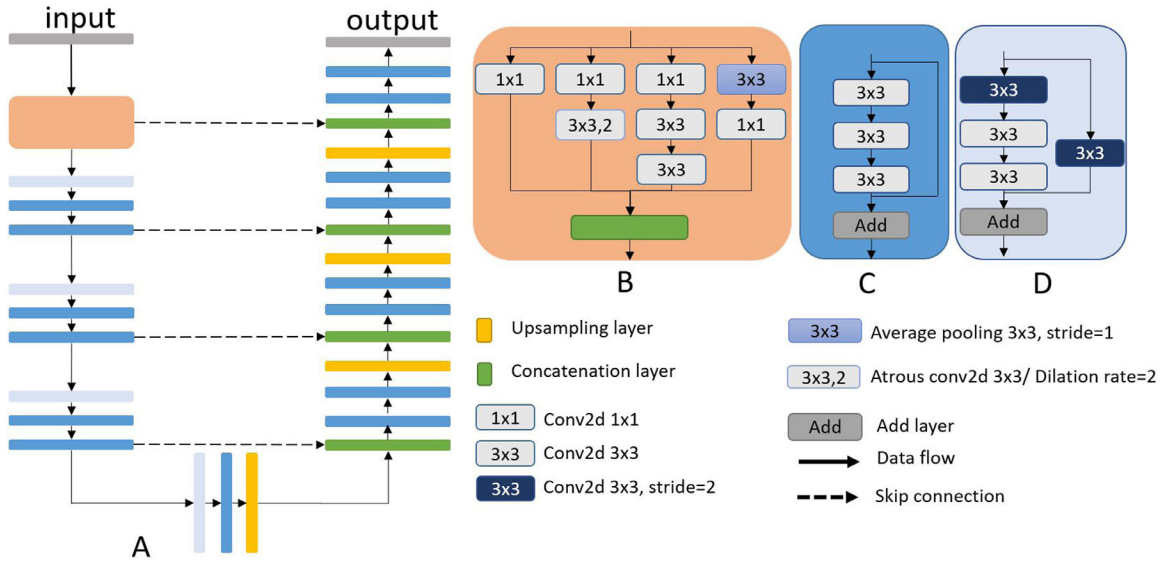
## Methods

### Data Acquisition

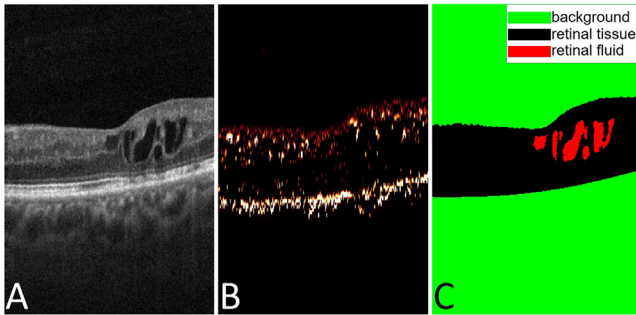
Volumetric OCT data were acquired over the central 3- × 3-mm region using a 70-kHz OCT commercial AngioVue system (RTVue-XR; Optovue, Inc., Fremont, CA, USA) centered at 840 nm with a full-width half-maximum bandwidth of 45 nm. Two repeated B-scans were taken at each 304 raster positions, and each B-scan consists of 304 A-lines. The structural OCT was generated by averaging the two repeated B-scans, and the OCTA was generated by using the split-spectrum amplitude-decorrelation angiography algorithm<sup>34</sup> to compute the decorrelation between the two repeated B-scans, simultaneously. Projection-resolved OCTA (PR-OCTA) removed shadow graphic artifacts from the superficial vasculatures while preserving true flow signal in deeper layers.<sup>35</sup>

### Convolutional Neural Network Architecture

Deep convolutional neural networks (CNNs) are superior to traditional methods for semantic segmentation tasks. U-Net-like CNNs have high adaptability to medical image segmentation due to skip connections that enable feature extraction with minimal loss of resolution.<sup>36</sup> In this study, we adopted U-Net-like architecture and designed ReF-Net to segment retinal fluid. To increase its capability for feature extraction, we applied some modifications to the original U-Net (Fig. 1A). A Multi-scale feature extraction block (Fig. 1B), inspired by Inception,<sup>37</sup> is placed after the input layer. This block can extract multiscale features



**Figure 1.** The architecture of deep convolutional neural network constructed in this study. (A) ReF-Net architecture. (B) Multi-scale block. (C, D) Residual convolutional blocks.



**Figure 2.** Representative OCT/OCTA B-scan showing retinal fluid. (A) OCT B-scan. (B) OCTA B-scan. (C) Ground truth map with three categories, background (green), retinal tissue (black), and retinal fluid area (red).

that enhance the ability of the neural network to detect targets with different sizes.<sup>26,27</sup> We also replaced the regular forward convolutional layers with residual units (Figs. 1C, 1D), borrowed from ResNet,<sup>38</sup> to increase the feature extraction ability of ReF-Net. More details on ReF-Net appear in the Appendix.

## Dataset Preprocessing

The data we used in our experiment contain three types of image: structural OCT B-scans (Fig. 2A), OCTA B-scans (Fig. 2B), and the ground truth map (Fig. 2C). We collected data from a total of 51 eyes (45 with DME and six healthy controls) from a clinical DR study, and each eye has two repeated volumetric scans. To remove excessive speckle noise, each B-scan was enhanced by performing a moving average with

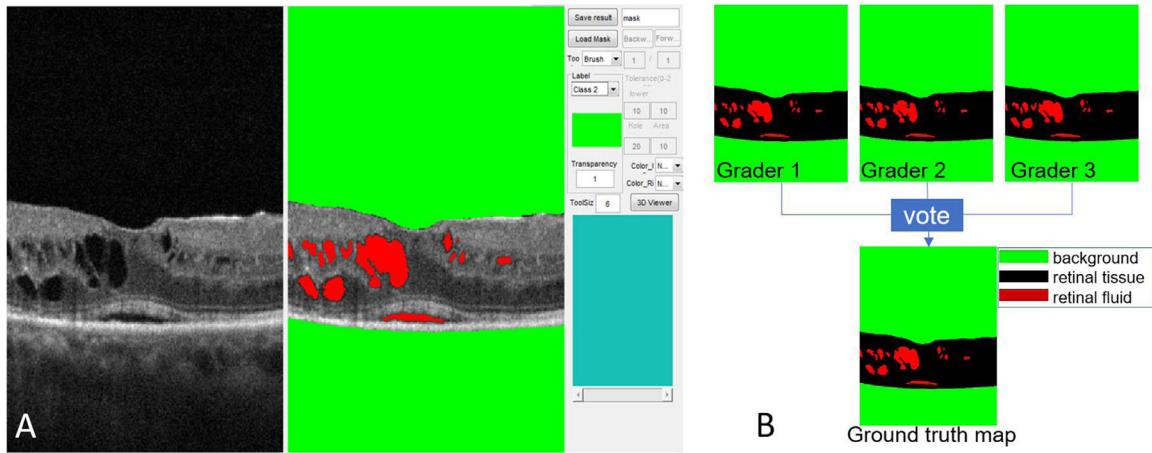
the two adjacent B-scans. In this study, each volumetric scan contained a total of 304 B-scans.

Previously, researchers only used OCT data to segment retinal fluid. Because the fluid region does not contain any vasculature, the simultaneously computed OCTA data may contribute to segmentation accuracy. To verify that OCTA data can indeed improve the segmentation performance, we designed two versions of the CNN in this study, each of which works with different inputs. The input of the first network (ReF-Net-OCT) only contains OCT data, and the other one (ReF-Net-OCTA) contains both OCT and OCTA data. Before feeding OCT and OCTA data to ReF-Net-OCTA, an image fusion operation (Equation 1) was applied to merge these two types of data together:

$$I_{\text{fusion}} = (1 - \beta) \times I_{\text{OCT}} + \beta \times I_{\text{OCTA}}. \quad (1)$$

Here,  $I_{\text{fusion}}$  is the fused data,  $\beta \in [0, 1]$  is a fusion factor, and  $I_{\text{OCT}}$  and  $I_{\text{OCTA}}$  represent the OCT and OCTA data, respectively. To get the optimal parameter value of  $\beta$ , we tested 12 values from 0.05 to 0.60 at intervals of 0.05.

The ground truth map that was used to train ReF-Net contains three categories: background, retinal tissue, and retinal fluid area. To obtain the ground truth, three graders manually delineated retinal fluid area using a customized graphical user interface (Fig. 3A). A guided bidirectional graph search (GBGS) method was used to segment retinal tissue boundaries.<sup>39</sup> We merged three manual grading outputs together by using a pixel-wise voting method to obtain the final ground truth map (Fig. 3B). In rare instances



**Figure 3.** Manual delineation of the ground truth for training. (A) The in-house graphical user interface software. (B) Three graders manually delineated the background (green), retinal tissue (black), and retinal fluid area (red). Pixel-wise voting method to generate the final ground truth map.

when three graders assigned a pixel to three disparate categories (one vote for background, one vote for tissue, one vote for fluid), the graders would reach a consensus through discussion.

## Ref-Net Hyperparameter Settings

### Loss Function

As a multiclass segmentation task, retinal fluid segmentation encounters a serious category imbalance problem because of the huge difference in the area between these three categories. To suppress the effect of the category imbalance, we used a categorical cross-entropy loss combined with weighted Jaccard coefficient loss:<sup>27</sup>

$$L = \sum_{i=1}^N J_i \times w_i, \quad \sum_{i=1}^N w_i = 1, \quad (2)$$

$$J = \left( 1 - \frac{\sum_x y(x) \times \hat{y}(x) + \alpha}{\sum_x (y(x) + \hat{y}(x)) - \sum_x y(x) \times \hat{y}(x) + \alpha} \right) \times \alpha,$$

where  $N$  is the number of categories,  $J_i$  is the Jaccard loss of  $i$ th category, and  $w_i$  is the weight of  $i$ th category associated with Jaccard coefficient  $J_i$ . In our experiment, we associated the three categories (retinal fluid area, retinal tissue, and background) with weights  $w = (0.5, 0.25, 0.25)$ . We set a higher value to the retinal fluid region to make the ReF-Net pay more attention to this category. The  $y$  and  $\hat{y}$  denote the ground truth and output of the ReF-Net, respectively,  $x$  is the position of each pixel in the sample. The  $\alpha$  is a smoothing factor usually set to 100<sup>27</sup> to get a similar gradient change from a similar loss change.

## Optimizer and Training

A modified Adam algorithm, AdamW,<sup>40</sup> was used to train ReF-Net by minimizing the weighted Jaccard coefficient loss. The initial learning rate was set to 0.001. Training batch size was set to 8. We used a global learning decay strategy that reduces the learning rate by 90% when the loss reaches a plateau. Early stopping was used to stop the training phase when the validation loss did not show a decrease over 15 training epochs. The total dataset (51 eyes) was randomly split into a training set (40 eyes) and test set (11 eyes). The training set consisted of 36 DME cases and four healthy controls, the testing set consisted of nine DME cases and two healthy controls. In the hyperparameters tuning step, we split five cases (four DME cases and one healthy control) to form a validation set from the training set. After parameters tuning, we used the whole training set (including the cases from the validation set) to train our model. To increase the cases available for training, we augmented the training data with horizontal flips. We only used horizontal flips because they are anatomically reasonable (a horizontal flip turns a left eye into a right eye, and vice-versa), and just this transformation provided sufficient data for training.

## Results

ReF-Net was implemented in Python 3.7 with Keras (Tensorflow-backend) on a PC with an Intel i7 CPU, two GTX 1080Ti GPUs, and 64GB RAM. In the training phase, each training epoch takes about 13 minutes,

**Table 1.** Agreement (in Voxels) Between Automated Detection and Manual Delineation of Volumetric Retinal Fluid Region (Mean  $\pm$  Standard Deviation)

CNNs	AROC	IoU	F1-Score	P Value <sup>a</sup>
<b>ReF-Net-OCT</b>	0.992 $\pm$ 0.011	0.770 $\pm$ 0.124	0.864 $\pm$ 0.084	–
<b>ReF-Net-OCTA (<math>\beta = 0.05</math>)</b>	0.995 $\pm$ 0.004	0.775 $\pm$ 0.107	0.870 $\pm$ 0.073	0.6541
<b>ReF-Net-OCTA (<math>\beta = 0.10</math>)</b>	0.996 $\pm$ 0.005	0.787 $\pm$ 0.124	0.875 $\pm$ 0.086	0.2841
<b>ReF-Net-OCTA (<math>\beta = 0.15</math>)</b>	0.996 $\pm$ 0.007	0.800 $\pm$ 0.096	0.885 $\pm$ 0.064	0.1063
<b>ReF-Net-OCTA (<math>\beta = 0.20</math>)</b>	<b>0.996 <math>\pm</math> 0.003</b>	<b>0.807 <math>\pm</math> 0.061</b>	<b>0.892 <math>\pm</math> 0.038</b>	0.0725
<b>ReF-Net-OCTA (<math>\beta = 0.25</math>)</b>	0.992 $\pm$ 0.008	0.771 $\pm$ 0.089	0.868 $\pm$ 0.060	0.6134
<b>ReF-Net-OCTA (<math>\beta = 0.30</math>)</b>	0.992 $\pm$ 0.007	0.760 $\pm$ 0.083	0.860 $\pm$ 0.057	0.5321
<b>ReF-Net-OCTA (<math>\beta = 0.35</math>)</b>	0.991 $\pm$ 0.009	0.728 $\pm$ 0.097	0.840 $\pm$ 0.065	0.1220
<b>ReF-Net-OCTA (<math>\beta = 0.40</math>)</b>	0.993 $\pm$ 0.006	0.716 $\pm$ 0.100	0.830 $\pm$ 0.069	0.0728
<b>ReF-Net-OCTA (<math>\beta = 0.45</math>)</b>	0.992 $\pm$ 0.013	0.672 $\pm$ 0.128	0.797 $\pm$ 0.091	0.0043
<b>ReF-Net-OCTA (<math>\beta = 0.50</math>)</b>	0.989 $\pm$ 0.015	0.664 $\pm$ 0.169	0.784 $\pm$ 0.137	0.0037
<b>ReF-Net-OCTA (<math>\beta = 0.55</math>)</b>	0.988 $\pm$ 0.013	0.636 $\pm$ 0.195	0.757 $\pm$ 0.172	0.0042
<b>ReF-Net-OCTA (<math>\beta = 0.60</math>)</b>	0.988 $\pm$ 0.017	0.638 $\pm$ 0.190	0.760 $\pm$ 0.162	0.0026

Bold indicates best performance.

<sup>a</sup>Paired samples *t*-test was used for F1-score between ReF-Net-OCT and ReF-Net-OCTAs.

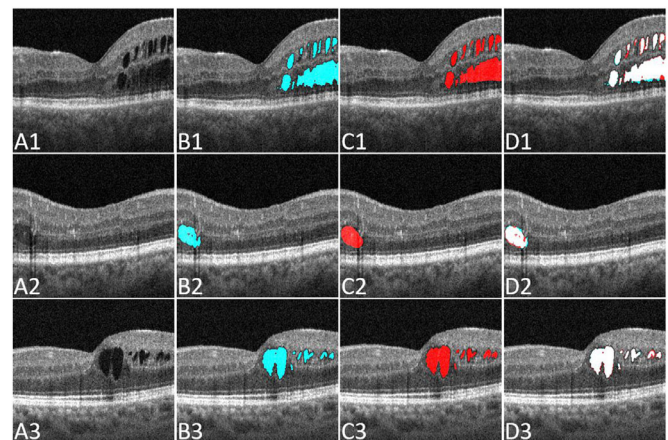
and each model reached the best performance after 55 epochs on average.

### Segmentation accuracy

The performance of each model (Table 1) was measured for the area-under-receiver-operating-characteristic-curve (AROC), intersection-over-union (IoU; also known as Jaccard coefficient), and F1-score, defined as

$$F1 = \frac{2 \times TP}{2 \times TP + FP + FN}, \quad (3)$$

where TP is true positive, FP is false positive, and FN is false negative. ReF-Net-OCT can achieve good results using only the OCT data as the input, which is expected because the retinal fluid region shows extremely low OCT reflectance intensity compared to the healthy retina. For ReF-Net-OCTA, the AROC, IoU and F1-score depend on the  $\beta$  parameter in Equation 1, with each reaching a peak when  $\beta = 0.20$ . The value of  $\beta$  regulates the proportion of information contributed to segmentation from both structural OCT and OCTA. As  $\beta$  increases, so does the proportion of OCTA information used for decision making, while the corresponding proportion of information from OCT decreases. The network achieved its best performance with  $\beta = 0.20$ ; thus we can confirm that the information from the OCT data played the major role in the segmentation, although OCTA still improved performance. Comparing the segmentation results of ReF-Net (ReF-Net-OCTA,  $\beta = 0.20$ ) (Fig. 4B) to the ground truth

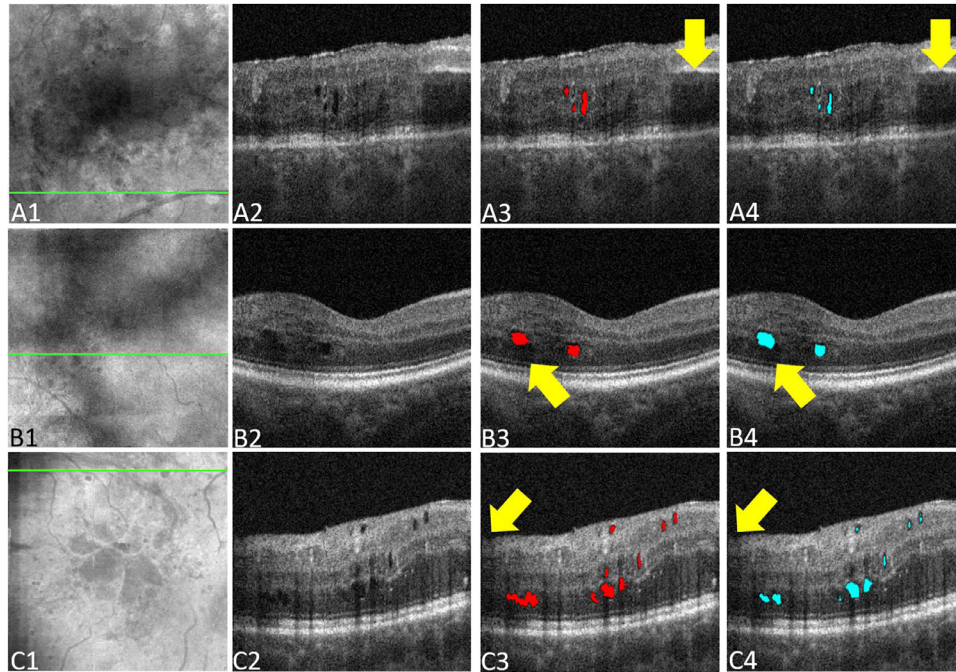


**Figure 4.** Comparison between ReF-Net-OCTA ( $\beta = 0.20$ ) and ground truth on structural OCT B-scans. (A) Structural OCT B-scans. (B) Segmented fluid maps from ReF-Net (blue) and (C) the ground truth maps (red) overlaid on structural cross-sections. (D) Difference map between segmented fluid from ReF-Net and ground truth. White area is the overlap region of two maps. The blue and red in (D) show pixels exclusively in the algorithm output or ground truth, respectively.

maps (Fig. 4 C), the large overlapping areas (Fig. 4D) indicate the high accuracy of ReF-Net.

### Resistance to Shadow Artifacts

Shadow artifacts caused by large vessels, vitreous floaters, and pupil vignetting can reduce the signal reflectance strength in retinal tissue, which reduces contrast in shadow area. To verify the robustness of ReF-Net on shadow artifacts, we applied ReF-Net



**Figure 5.** Automated retinal fluid segmentation results on shadow artifact effected scans. *Yellow arrows* indicate shadow artifacts. (Row A) Example case with large vessel shadow artifacts. (Row B) Example case with vitreous floater shadow artifacts. (Row C) Example case with pupil vignetting shadow artifacts. (Column 1) Reflectance en face images, with the *green line* indicating the position of the B-scan shown in the other columns. (Column 2) Raw cross-sectional scans. (Column 3) Ground truth map (*red*) overlaid on B-scans. (Column 4) ReF-Net (ReF-Net-OCTA,  $\beta = 0.20$ ) outputs (*blue*) overlaid on the B-scans.

(ReF-Net-OCTA,  $\beta = 0.20$ ) on cases with three types of typical shadow artifacts (Fig. 5); ReF-Net could handle all three types. This is likely because ReF-Net does not just rely on contrast information between the fluid and tissue, but also on other geometrical information extracted by convolutional kernels.

### Volumetric Versus 2D Projected Retinal Fluid and Cross-Sectional Quantification

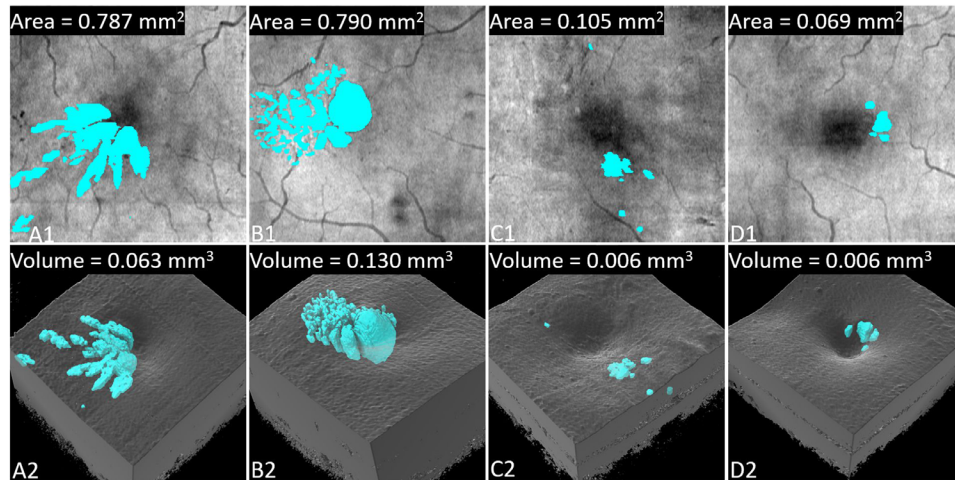
By applying the best model (ReF-Net-OCTA,  $\beta = 0.20$ ) to each frame of OCT/OCTA data, we can construct a 3D segmentation. The total fluid volume in the retina is a prognostic indicator for visual acuity.<sup>41</sup> The average percent difference between the fluid volume predicted by Ref-Net-OCTA with  $\beta = 0.20$  and the ground truth is  $8.71\% \pm 4.48\%$  (mean  $\pm$  SD). Compared to the 2D en face projected retinal fluid regions overlaid on projected OCT reflectance (Figs. 6A1–6D1), the 3D retinal fluid region (Figs. 6A2–6D2) is more intuitive. Furthermore, the 3D volumetric result produces a more meaningful quantification than the 2D projected area, which may not reflect the actual extent of the fluid region. In (Figs. 6A1, 6B1); for example, although the two DME

cases have similar fluid areas in the 2D en face image, the actual fluid volumes differ by a factor of 2 (Figs. 6A2, 6B2). Similarly, cases with very different 2D projected fluid area on en face image (Figs. 6C1, 6D1), may have similar fluid volume (Figs. 6C2, 6D2).

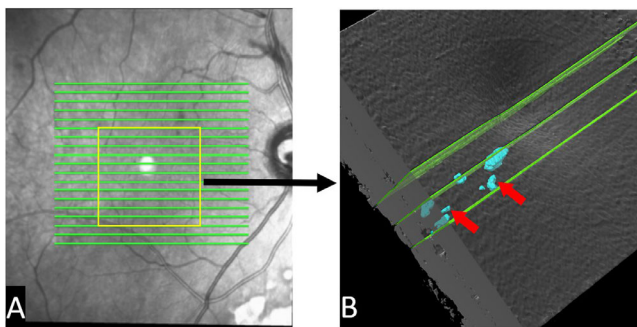
Figure 7 further demonstrates how 2D fluid measurement on cross-sectional raster scans are more likely to miss a substantial portion of retinal fluid because of undersampling. Using densely sampled OCT and OCTA data, ReF-Net can render retinal fluid cysts in 3D (the volumetric of fluid is shown in Fig. 7B, blue) and reveal anatomic changes related to macular edema more completely.

### Recovering DME Diagnosis From a False-Negative Cmt Measurement

Central macular thickness (CMT) is a commonly used biomarker for DME diagnosis. Major clinical trials have used CMT greater than 2 SD from the population mean for inclusion criteria in clinical trials.<sup>42</sup> Because there is a significant population variation in retinal thickness, and specific pathology such as epiretinal membrane or retinal atrophy can cause changes unrelated to macular edema, diagnosis



**Figure 6.** Comparison between 2D projected fluid areas and 3D fluid volumes in DME cases. (A1-D1) 2D structural OCT and retinal fluid projections. (A2-D2) 3D structural OCT and retinal fluid representations. Apparent fluid areas can be similar while volumes are quite different (A, B), and apparent fluid areas can be quite different while volumes are similar (C, D). In such cases, the 2D projection is misleading.



**Figure 7.** A DME case in which a substantial portion of retinal fluid would be missed by under-sampled scans. (A) Infrared photograph with sampling positions (green lines) from a Spectralis OCT (Heidelberg Engineering Inc.) scan. (B) Dense volumetric OCT (RTVue-XR; Optovue, Inc.) with retinal fluid volume (blue). The yellow square in (A) indicates the scanning position in (B). Red arrows indicate the retinal fluid missed by the undersampled scan, which can be detected by our algorithm using the densely-sampled OCT. Green lines indicate the sampling position of Spectralis OCT scan.

of DME solely based on CMT measurements can be unreliable. Figure 8 shows an eye with a CMT of 217  $\mu\text{m}$ , which does not meet the CMT definition of center-involved DME, but it has a known clinical diagnosis of DME with retinal fluid caused by DME.<sup>3</sup> ReF-Net automatically detected a fluid volume of 0.044  $\text{mm}^3$ .

### Longitudinal Study of Retinal Fluid In Oct/Octa Scans

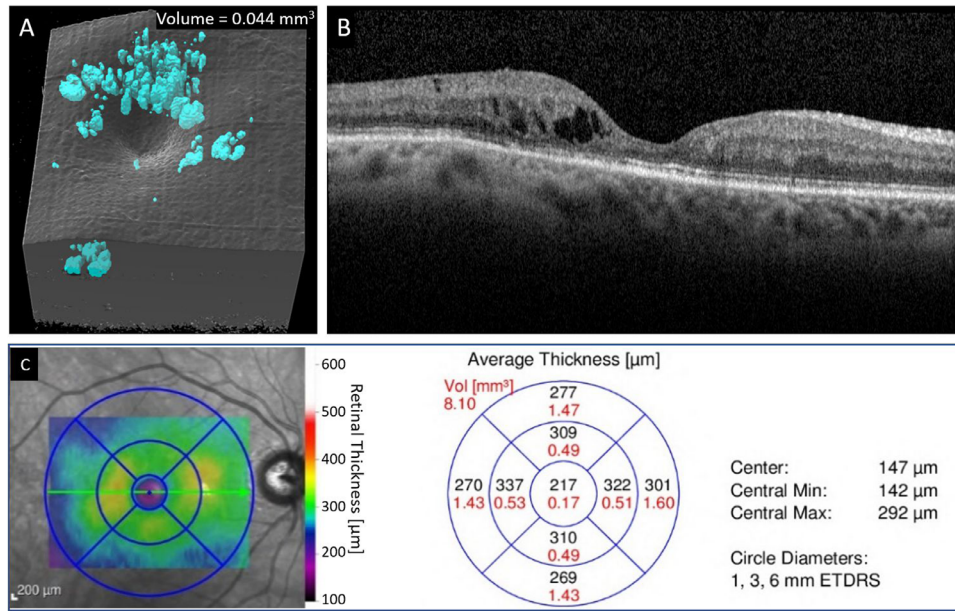
The change in retinal fluid volume is an important indicator of treatment response in DME. Using our algorithm, it is easy to visualize fluid volume changes

longitudinally. To do so, we register the baseline and follow-up scans (Figs. 9A, 9B) using Bruch's membrane and large retinal vessels as a reference for the axial and lateral directions, respectively. After the omnidirectional registration (Fig. 9C), the changes in the shape and size of the fluid (Fig. 9D) can be easily visualized. Furthermore, we can identify the vascular changes caused by retinal fluid accumulation by overlaying the retinal fluid volumes on the angiographic volumes (Figs. 9E, 9F).

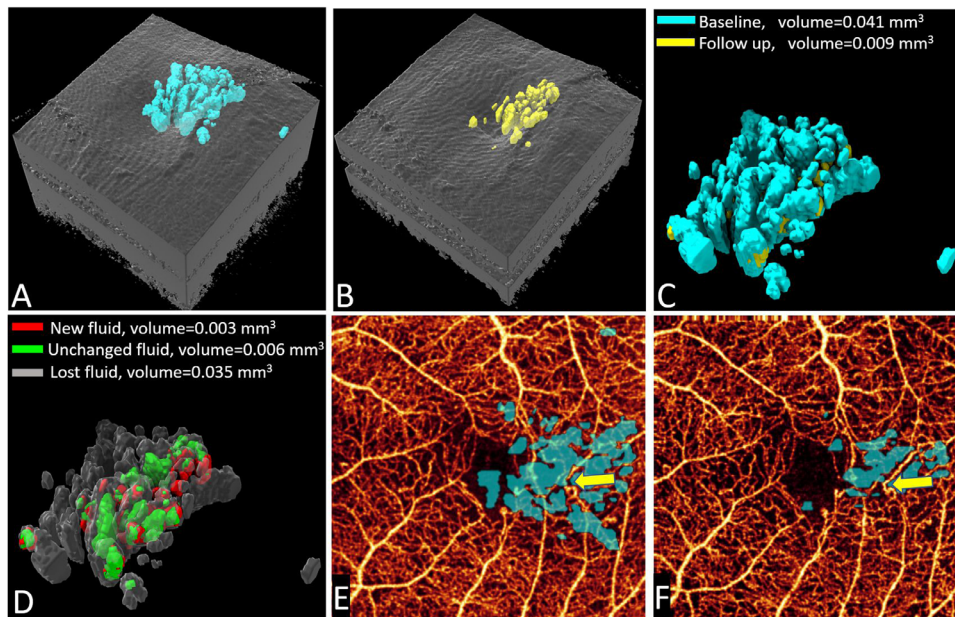
## Discussion

We have presented a deep-learning-based method (ReF-Net) for segmenting and quantifying retinal fluid in 3D using OCT and OCTA volumes. We demonstrated that OCTA data enhances the segmentation task, and the 3D approach provides a more intuitive and complete representation of the anatomic changes in DME than 2D cross-sections or projections.

ReF-Net is a U-Net-like convolutional neural network. We added several useful modifications to enhance its feature extraction capability, such as a multi-scale feature extraction block and residual blocks. We also compared our network's performance to previously published results. For a fair comparison, all hyperparameters in the various networks were set to the same value. ReF-Net achieved the best performance out of the networks we examined (Table 2). The methods proposed by Bai et al.<sup>29</sup> and Schlegel et al.<sup>30</sup> (both of which were based on fully convolu-



**Figure 8.** A diabetic macular edema (DME) case with a false-negative result from central macular thickness (CMT) was automatically detected and measured by ReF-Net. (A) Retinal fluid volume segmented by ReF-Net. (B) Cross-sectional structural OCT. (C) Retinal thickness map and average thickness distribution in early treatment diabetic retinopathy study (ETDRS) grid. The CMT value is 217, which does not meet the definition of DME.



**Figure 9.** Local dynamics of retinal fluid in longitudinal monitoring of a DME eye. (A) Baseline. (B) One year follow-up after the treatment. (C) Registered baseline and follow-up scans. (D) Changes in the retinal fluid region. (E) Baseline retinal fluid area overlaid on an inner retinal OCT angiogram. (F) Follow-up retinal fluid area overlaid on an inner retinal OCT angiogram. The yellow arrow indicated the change of vasculature caused by retinal fluid.

tional neural networks) show lower performance than U-Net-like CNNs (the other three methods in Table 2). This may be because the skip connections could transfer feature information from the initial layers to the deeper layers directly, thereby improving the

CNN’s ability to identify minute details in the target. Comparing to the method of Girish et al.,<sup>31</sup> ReF-Net shows higher accuracy, which is likely because of the multiscale feature extraction block and residual blocks, which made ReF-Net more adaptable to different



**Table 2.** Performance Comparison Between Deep-Learning-Based Methods (Mean  $\pm$  SD)

Methods	F1-Score	IoU	AROC
Bai et al. <sup>29</sup>	0.836 $\pm$ 0.084	0.727 $\pm$ 0.119	0.989
Schlegl et al. <sup>30</sup>	0.842 $\pm$ 0.068	0.733 $\pm$ 0.102	0.990
Girish et al. <sup>31</sup>	0.871 $\pm$ 0.075	0.779 $\pm$ 0.113	0.992
Li et al. <sup>32</sup>	0.862 $\pm$ 0.073	0.764 $\pm$ 0.106	0.991
ReF-Net-OCTA, $\beta = 0.20$	<b>0.892 <math>\pm</math> 0.038</b>	<b>0.807 <math>\pm</math> 0.061</b>	<b>0.996</b>

sizes of edema. Although Li's method<sup>32</sup> reported high accuracy on sparsely sampled OCT volumes, in our dataset, however, its performance did not surpass ReF-Net, which may be because the architecture of Li's method was not optimized for high sampling density OCT volumes.

The main feature useful for retinal fluid segmentation is the low reflectance of fluid regions in structural OCT scans. Additionally, because the fluid region does not contain any flow signal, OCTA data have the potential to improve segmentation performance. In this study, we analyzed the contribution of OCTA data to ReF-Net-OCTA's performance. By merging OCT and OCTA data with different fusion factors, we trained 12 different models based on the ReF-Net-OCTA architecture, each with different segmentation accuracy. Experimental results show that ReF-Net-OCTA achieves the best performance (F1-score  $> 0.89$ ) when the fusion factor equals 0.2 (Table 1). The accuracy of ReF-Net-OCTA is superior to ReF-Net-OCT for this value of  $\beta$ , indicating that OCTA data is helpful to segment fluid correctly. While the inclusion of OCTA data did improve network performance (Table 1), the improvement was not statistically significant ( $\beta = 0.2$ ;  $P = 0.0725$ , paired samples *t*-test). Nonetheless we still believe that inclusion of OCTA data improved performance, given both the higher accuracy demonstrated by the network with OCTA data included, and the clear relationship between the  $\beta$  parameter and network performance. At this and similar values, the  $\beta$  network performance was also most consistent, as indicated by the low measurement error values. Additional benefits of using OCT/OCTA data are as follows: (1) high-sampling-density structural OCT and OCTA can be simultaneously processed from the same scan; (2) they are inherently coregistered, facilitating the study on anatomic and angiographic pathologies; (3) OCTA vasculatures can be used for registering multiple scans during longitudinal studies.

Because the input to ReF-Net is a single B-scan, ReF-Net is compatible with conventional cross-

sectional data. Additionally, ReF-Net also achieves high performance on scan volumes with structural OCT only, which allows us to apply ReF-Net to the scans that only have OCT data at the expense of some accuracy loss. Furthermore, ReF-Net does not require segmented layers for its input. This is a critical advantage in this context, since retinal slab segmentation is especially error prone in the presence of pathological disruptions to normal slab anatomy, of which the presence of retinal fluid is an important example. ReF-Net also shows strong robustness on different types of topical shadow artifacts, including large vessel shadows, vitreous floater shadows, and pupil vignetting shadows. ReF-Net was trained on 3-  $\times$  3-mm central macular scans, but it could also easily be migrated to larger scan patterns or the scans acquired from other OCT instruments using transfer learning.<sup>43</sup>

Although our method can segment retinal fluid with high accuracy, there are some drawbacks that may hinder its application. Retinal fluid can be classified into intraretinal fluid and subretinal fluid according to location, with each category having different diagnostic and prognostic value.<sup>44</sup> Because a CNN is insensitive to the location of the target and the different types of retinal fluid have similar features, we labeled all the fluid regions in one category to help ReF-Net learn consistent features in order to improve its performance. Thus, ReF-Net cannot distinguish these two types of retinal fluid. However, differentiating type could be easily accomplished by an additional retinal layer segmentation step. Alternatively, other unidentified fluid accumulation features that could be discovered with the precise 3D segmentation provided by ReF-Net could possibly serve as superior biomarkers than the intra/subretinal fluid classification. Another limitation concerns the features used for decision making. The most obvious feature indicating retinal fluid is the high contrast between the fluid region and retinal tissue. In future work, we can try to improve the performance of ReF-Net by resolving this drawback.

## Conclusions

In summary, we designed a deep-learning-based method named ReF-Net to segment volumetric retinal fluid on OCT/OCTA scans. By combining the OCT and OCTA data as the input to the network, ReF-Net demonstrated that OCTA data can improve retinal fluid segmentation. Our results indicate volumetric representations of retinal fluid can provide more comprehensive information than 2D either cross-sections or projections.

## Acknowledgments

Supported by grant National Institutes of Health (R01 EY027833, R01 EY024544, P30 EY010572); Unrestricted Departmental Funding Grant and William & Mary Greve Special Scholar Award from Research to Prevent Blindness (New York, NY).

Disclosure: **Y. Guo**, None; **T.T. Hormel**, None; **H. Xiong**, None; **J. Wang**, None; **T.S. Hwang**, None; **Y. Jia**, Optovue (F, P)

## References

- Spaide RF. Retinal vascular cystoid macular edema: review and new theory. *Retina*. 2016;36(10):1823–1842.
- Gonzalez VH, Campbell J, Holekamp NM, et al. Early and long-term responses to anti-vascular endothelial growth factor therapy in diabetic macular edema: analysis of protocol I data. *Am J Ophthalmol*. 2016;172:72–79.
- Browning DJ, Glassman AR, Aiello LP, et al. Optical coherence tomography measurements and analysis methods in optical coherence tomography studies of diabetic macular edema. *Ophthalmology*. 2008;115(8):1366–1372.
- Virgili G, Parravano M, Jr E, Gordon I, Lucen-teforte E. Anti-vascular endothelial growth factor for diabetic macular oedema: a network meta-analysis. *Cochrane Database Syst Rev*. 2018;10(10):CD007419.
- Nunes S, Pereira I, Santos A, Bernardes R, Cunha-Vaz J. Central retinal thickness measured with HD-OCT shows a weak correlation with visual acuity in eyes with CSME. *Br J Ophthalmol*. 2010;94(9):1201–1204.
- Gerendas BS, Prager S, Deak G, et al. Predictive imaging biomarkers relevant for functional and anatomical outcomes during ranibizumab therapy of diabetic macular oedema. *Br J Ophthalmol*. 2018;102(2):195–203.
- Bogunović H, Abramoff MD, Sonka M. Geodesic graph cut based retinal fluid segmentation in optical coherence tomography. *Ophthalm Med Imaging Anal*. 2015:49–56.
- Montuoro A, Waldstein SM, Gerendas BS, Schmidt-Erfurth U, Bogunović H. Joint retinal layer and fluid segmentation in OCT scans of eyes with severe macular edema using unsupervised representation and auto-context. *Biomed Opt Express*. 2017;8(3):1874.
- Tennakoon R, Gostar AK, Hoseinnezhad R, Bab-Hadiashar A. Retinal fluid segmentation in OCT images using adversarial loss based convolutional neural networks. In: *2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)*. 2018:1436–1440.
- Wang J, Zhang M, Pechauer AD, et al. Automated volumetric segmentation of retinal fluid on optical coherence tomography. *Biomed Opt Express*. 2016;7(4):1577–1589.
- Long J, Shelhamer E, Darrell T, et al. Fully convolutional networks for semantic segmentation. *IEEE Trans Pattern Anal Mach Intell*. 2014;39(4):640–651.
- Badrinarayanan V, Kendall A, Cipolla R. Segnet: A deep convolutional encoder-decoder architecture for image segmentation. *arXiv Prepr arXiv151100561*. 2015.
- Yu F, Koltun V. Multi-scale context aggregation by dilated convolutions. *arXiv Prepr arXiv151107122*. November 2015.
- Chen L-CC, Papandreou G, Kokkinos I, Murphy K, Yuille AL. DeepLab: semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs. *IEEE Trans Pattern Anal Mach Intell*. 2018;40(4):834–848.
- Lin G, Milan A, Shen C, Reid I. RefineNet: Multi-Path Refinement Networks for High-Resolution Semantic Segmentation. *arXiv Prepr arXiv161106612*. November 2016.
- Zhao H, Shi J, Qi X, Wang X, Jia J. Pyramid scene parsing network. *arXiv Prepr arXiv161201105*. 2016.
- Jegou S, Drozdal M, Vazquez D, Romero A, Bengio Y. The one hundred layers tiramisu: fully convolutional DenseNets for semantic segmentation. *IEEE Comput Soc Conf Comput Vis Pattern Recognit Work*. 2017;2017-July:1175–1183.
- Chen LC, Zhu Y, Papandreou G, Schroff F, Adam H. Encoder-decoder with atrous separable

- convolution for semantic image segmentation. *Lect Notes Comput Sci (including Subser Lect Notes Artif Intell Lect Notes Bioinformatics)*. 2018;11211 LNCS:833–851.
19. Fang L, Cunefare D, Wang C, Guymer RH, Li S, Farsiu S. Automatic segmentation of nine retinal layer boundaries in OCT images of non-exudative AMD patients using deep learning and graph search. *Biomed Opt Express*. 2017;8(5):2732.
  20. Devalla SK, Renukanand PK, Sreedhar BK, et al. DRUNET: a dilated-residual U-Net deep learning network to segment optic nerve head tissues in optical coherence tomography images. *Biomed Opt Express*. 2018;9(7):3244.
  21. Zang P, Wang J, Hormel TT, Liu L, Huang D, Jia Y. Automated segmentation of peripapillary retinal boundaries in OCT combining a convolutional neural network and a multi-weights graph search. *Biomed Opt Express*. 2019;10(8):4340.
  22. Laibacher T, Weyde T, Jalali S. M2U-Net: Effective and efficient retinal vessel segmentation for resource-constrained environments. *arXiv Prepr arXiv181107738*. 2018.
  23. Alom MZ, Hasan M, Yakopcic C, Taha TM, Asari VK. Recurrent residual convolutional neural network based on u-net (r2u-net) for medical image segmentation. *arXiv Prepr arXiv180206955*. 2018.
  24. Wang J, Hormel TT, Gao L, et al. Automated diagnosis and segmentation of choroidal neovascularization in OCT angiography using deep learning. *Biomed Opt Express*. 2020;11(2):927–944.
  25. Gao M, Guo Y, Hormel TT, Sun J, Hwang TS, Jia Y. Reconstruction of high-resolution 6×6-mm OCT angiograms using deep learning. *Biomed Opt Express*. 2020;11(7):3585.
  26. Guo Y, Camino A, Wang J, Huang D, Hwang TS, Jia Y. MEDnet, a neural network for automated detection of avascular area in OCT angiography. *Biomed Opt Express*. 2018;9(11):5147.
  27. Guo Y, Hormel TT, Xiong H, et al. Development and validation of a deep learning algorithm for distinguishing the nonperfusion area from signal reduction artifacts on OCT angiography. *Biomed Opt Express*. 2019;10(7):3257–3268.
  28. Wang J, Hormel TT, You Q, et al. Robust non-perfusion area detection in three retinal plexuses using convolutional neural network in OCT angiography. *Biomed Opt Express*. 2020;11(1):330–345.
  29. Bai F, Marques MJ, Gibson SJ. Cystoid macular edema segmentation of optical coherence tomography images using fully convolutional neural networks and fully connected CRFs. 2017, <http://arxiv.org/abs/1709.05324>.
  30. Schlegl T, Waldstein SM, Bogunovic H, et al. Fully automated detection and quantification of macular fluid in OCT using deep learning. *Ophthalmology*. 2018;125(4):549–558.
  31. Girish GN, Thakur B, Chowdhury SR, Kothari AR, Rajan J. Segmentation of intra-retinal cysts from optical coherence tomography images using a fully convolutional neural network model. *IEEE J Biomed Heal Informatics*. 2019;23(1):296–304.
  32. Li MX, Yu SQ, Zhang W, et al. Segmentation of retinal fluid based on deep learning: Application of three-dimensional fully convolutional neural networks in optical coherence tomography images. *Int J Ophthalmol*. 2019;12(6):1012–1020.
  33. Lu D, Heisler M, Lee S, et al. Deep-learning based multiclass retinal fluid segmentation and detection in optical coherence tomography images using a fully convolutional neural network. *Med Image Anal*. 2019;54:100–110.
  34. Jia Y, Tan O, Tokayer J, et al. Split-spectrum amplitude-decorrelation angiography with optical coherence tomography. *Opt Express*. 2012;20(4):4710–4725.
  35. Wang J, Zhang M, Hwang TS, et al. Reflectance-based projection-resolved optical coherence tomography angiography [Invited]. *Biomed Opt Express*. 2017;8(3):1536.
  36. 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*. 2015:234–241.
  37. Szegedy C, Liu W, Jia Y, et al. Going deeper with convolutions. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2015:1–9.
  38. 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*. 2016:770–778.
  39. Guo Y, Camino A, Zhang M, et al. Automated segmentation of retinal layer boundaries and capillary plexuses in wide-field optical coherence tomographic angiography. *Biomed Opt Express*. 2018;9(9):4429.
  40. Loshchilov I, Hutter F. Decoupled weight decay regularization. *arXiv preprint arXiv:1711.05101*. 2017.
  41. Pelosini L, Hull CC, Boyce JF, McHugh D, Stanford MR, Marshall J. Optical coherence tomography may be used to predict visual acuity in patients with macular edema. *Investig Ophthalmol Vis Sci*. 2011;52(5):2741–2748.

42. Chalam K V., Bressler SB, Edwards AR, et al. Retinal thickness in people with diabetes and minimal or no diabetic retinopathy: Heidelberg spectralis optical coherence tomography. *Investig Ophthalmol Vis Sci.* 2012;53(13):8154–8161.
43. Yosinski J, Clune J, Bengio Y, Lipson H. How transferable are features in deep neural networks? *Adv Neural Inf Process Syst.* 2014;4(January):3320–3328.
44. Daruich A, Matet A, Moulin A, et al. Mechanisms of macular edema: beyond the surface. *Prog Retin Eye Res.* 2018;63(October 2017):20–68.