

Microvasculature Segmentation and Intercapillary Area Quantification of the Deep Vascular Complex Using Transfer Learning

Julian Lo¹, Morgan Heisler¹, Vinicius Vanzan², Sonja Karst^{2,3},
Ivana Zadro Matovinović⁴, Sven Lončarić⁴, Eduardo V. Navajas², Mirza Faisal Beg¹, and
Marinko V. Šarunić¹

¹ School of Engineering Science, Simon Fraser University, Burnaby, BC, Canada

² Department of Ophthalmology and Visual Sciences, University of British Columbia, Vancouver, BC, Canada

³ Department of Ophthalmology and Optometry, Medical University of Vienna, Vienna, Austria

⁴ Faculty of Electrical Engineering and Computing, University of Zagreb, Zagreb, Croatia

Correspondence: Marinko V. Šarunić, School of Engineering Science, Simon Fraser University, 8888 University Drive, Burnaby, BC V5A 1S6, Canada. e-mail: msarunic@sfu.ca

Received: September 29, 2019

Accepted: May 8, 2020

Published: July 10, 2020

Keywords: optical coherence tomography; angiography; machine learning; diabetic retinopathy; neural networks

Citation: Lo J, Heisler M, Vanzan V, Karst S, Matovinović IZ, Lončarić S, Navajas EV, Beg MF, Šarunić MV. Microvasculature segmentation and intercapillary area quantification of the deep vascular complex using transfer learning. *Trans Vis Sci Tech.* 2020;9(2):38. <https://doi.org/10.1167/tvst.9.2.38>

Purpose: Optical coherence tomography angiography (OCT-A) permits visualization of the changes to the retinal circulation due to diabetic retinopathy (DR), a microvascular complication of diabetes. We demonstrate accurate segmentation of the vascular morphology for the superficial capillary plexus (SCP) and deep vascular complex (DVC) using a convolutional neural network (CNN) for quantitative analysis.

Methods: The main CNN training dataset consisted of retinal OCT-A with a 6 × 6-mm field of view (FOV), acquired using a Zeiss PlexElite. Multiple-volume acquisition and averaging enhanced the vasculature contrast used for constructing the ground truth for neural network training. We used transfer learning from a CNN trained on smaller FOVs of the SCP acquired using different OCT instruments. Quantitative analysis of perfusion was performed on the resulting automated vasculature segmentations in representative patients with DR.

Results: The automated segmentations of the OCT-A images maintained the distinct morphologies of the SCP and DVC. The network segmented the SCP with an accuracy and Dice index of 0.8599 and 0.8618, respectively, and 0.7986 and 0.8139, respectively, for the DVC. The inter-rater comparisons for the SCP had an accuracy and Dice index of 0.8300 and 0.6700, respectively, and 0.6874 and 0.7416, respectively, for the DVC.

Conclusions: Transfer learning reduces the amount of manually annotated images required while producing high-quality automatic segmentations of the SCP and DVC that exceed inter-rater comparisons. The resulting intercapillary area quantification provides a tool for in-depth clinical analysis of retinal perfusion.

Translational Relevance: Accurate retinal microvasculature segmentation with the CNN results in improved perfusion analysis in diabetic retinopathy.

Introduction

Diabetic retinopathy (DR), a complication of diabetes mellitus, is the most common cause of vision loss among people with diabetes and affects 749,800 Canadians.¹ DR damages the structure of the capillar-

ies in the retina,² leading to widespread areas of retinal ischemia as it progresses. Optical coherence tomography angiography (OCT-A) is a rapidly emerging imaging technology that allows for the retinal microvasculature to be seen volumetrically in micrometer-scale detail.^{3,4} OCT-A has been shown to produce images that closely relate to histology,⁵⁻⁸ and it presents a

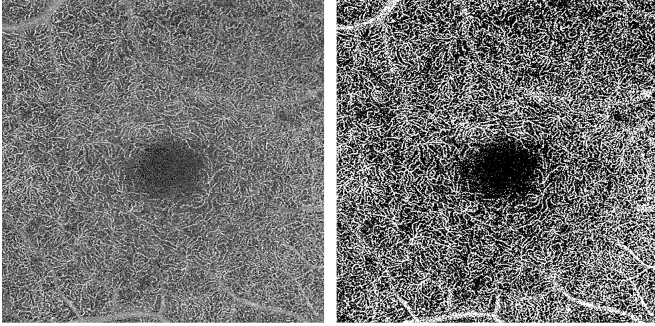


Figure 1. (Left) Original single-frame image of the deep vascular complex. (Right) Image thresholded using Otsu's method.

noninvasive and dye-free alternative with a lower risk of complications⁹ when compared to the current gold standard, fluorescein angiography (FA).

Analysis and quantification of the retinal microvasculature benefit from multi-scale imaging with fields of view (FOVs) ranging from $\sim 2 \times 2$ mm to $\sim 6 \times 6$ mm. At a smaller FOV, the capillaries that comprise the structure of the microvasculature can be individually resolved; with a larger FOV, macroscopic features, including regions of capillary non-perfusion, can be identified. With recent research hypothesizing that early manifestations of DR form in the retinal periphery,¹⁰ improving the vessel segmentations and tools for quantification in wider fields of view for both the superficial capillary plexus (SCP) and deep vascular complex (DVC)¹¹ would be important to clinicians. The DVC resulted from combining the intermediate and deep capillary plexuses due to difficulty in resolving each plexus individually, and it has shown a higher correlation to retinal ischemia in DR.^{12,13}

OCT-A images are information rich, and it is time consuming for a clinician to trace the vessels for detailed analysis. For the cases of highest clinical interest, small changes in the capillaries must be detected. Consequently, developing accurate automated methods of microvasculature segmentation is an essential step toward quantification; however, the efficacy of traditional image-processing algorithms can vary based on artifacts present in the image, most notably from noise. Segmentation of microvasculature in fundusoscopic photographs, as well as by FA, has been examined,¹⁴ but fewer algorithms specifically designed for OCT-A have been developed. Simple thresholding of the OCT-A image intensity has been applied,^{15,16} but such approaches pose numerous drawbacks in their invariance to microvasculature features and performance when applied to lower quality images with a low signal-to-noise ratio (SNR). The disadvantages of thresholding can be seen in Figure 1, where an image of the DVC in a patient with mild diabetic

retinopathy was processed with Otsu's method.¹⁷ This representative example of thresholding demonstrates that vessels do not maintain continuity, and a significant portion of the speckle is erroneously delineated as a vessel.

Methods using vesselness filters have been developed,^{18,19} but they either are similarly limited by the SNR or require manual correction. A tophat filter and optimally oriented flux method for segmenting the vessels²⁰ have been implemented and demonstrated for brain imaging in mice. In addition, some commercial OCT systems also provide segmentation of the vessels but generally face the same issues with images with low SNRs.

Machine learning is a rapidly growing field, showing promising results for numerous ophthalmological applications. Topics in this field that have recently been investigated include retinal layer^{21–23} and capillary plexus²⁴ segmentation, cone photoreceptor identification,^{25,26} macular fluid segmentation,²⁷ geographic atrophy segmentation,²⁸ OCT image categorization,²⁹ diagnosis and referral for retinal disease patients,^{30,31} and synthesis of fundusoscopic images.³² Additionally, recent reports have published online tools to improve the accessibility of machine learning-based retinal layer segmentation through intuitive user interfaces that can be used directly by clinicians.³³ Machine learning algorithms have also been applied toward OCT-A segmentation, with a recent approach (MEDnet)³⁴ applying a convolutional neural network (CNN) to identify and segment avascular areas in widefield images of the SCP. We have also previously published a method of using a CNN to segment 1×1 -mm images of the SCP.³⁵ Machine learning algorithms using CNNs are well suited to addressing the issues of vessel segmentation through a series of trainable filters. These filters allow the segmentation to be sensitive to vessel boundaries; hence, they also have the potential to preserve vessel widths. However, even with the strengths of machine learning, the quality of the OCT-A images will have a significant impact on the results of vessel segmentation and quantification.

In our previous work, we proposed a method to register and average multiple sequentially acquired OCT-A images in order to significantly improve image quality and vessel discernibility.³⁶ Related works in the literature have also investigated averaging of OCT-A images to improve vessel contrast³⁷ and automated biomarker identification algorithms.³⁸ However, these approaches require prolonged imaging sessions, which is not always possible, particularly in a high-volume clinical environment. Therefore, there is greater clinical utility in the development of an algorithm that can accurately segment and subsequently quantify the vasculature and correspond-

ing intercapillary areas (ICAs) from one single-frame OCT-A image. Quantification of individual ICAs has been previously explored,^{15,18} but these studies used 3×3 -mm images of the SCP. Similarly, quantification of the vasculature in the DVC has been explored, but this research did not include ICAs.¹⁹

With many approaches presenting accurate analyses of ICAs in the SCP, the contribution of this report lies in its description of an original and novel method to accurately and automatically segment and quantify these regions in the DVC. We used an approach of transfer learning, referred to as fine-tuning, for the segmentation of retinal microvasculature in single-frame, widefield 6×6 -mm OCT-A images for the purposes of ICA quantification. The developed framework allows for the adaptation of an initial segmentation network to a new dataset with significantly fewer manually graded training examples. We combined the approach of OCT-A averaging to generate high-contrast images of the vascular networks with supervised learning to provide the CNN with accurate ground truth data in order to guide the vessel segmentation even in the case of a single (unaveraged) OCT-A image. The computer-generated segmentations were qualitatively examined by retinal specialists and compared to manual segmentations from a trained rater. The outputs of the automated vessel analysis provide nearly immediately available, quantitative information on the microvasculature and ICAs from a single OCT-A volume and hence can potentially accelerate treatment plans and improve DR prognosis.

Methods

Subject Criteria and Data Preparation

Subject recruitment and imaging took place at the Eye Care Centre of Vancouver General Hospital and at North Shore Eye Associates. The project protocol was approved by the research ethics boards at the University of British Columbia and Vancouver General Hospital, and the experiment was performed in accordance with the tenets of the Declaration of Helsinki. Written informed consent was obtained by all subjects.

Subjects in the control group ($n = 8$) displayed no evidence of retinal or ocular pathology upon examination by an experienced retina specialist (Table 1). Subjects classified as diabetic ($n = 28$) were diagnosed with DR based on the international DR severity scale.³⁹ All subjects were screened for clear ocular media, ability to fixate, and ability to provide informed consent before imaging. In addition, patients with diabetic macular edema were not included in the study.

Table 1. Demographics of the Control Dataset Used in This Study

Gender	<i>n</i>	Mean Age, <i>y</i> (SD)
Male	4	24.5 (3)
Female	4	53.5 (23.1)

Optical Coherence Tomography Instrumentation

The data used for this study were acquired with the Zeiss PlexElite (Carl Zeiss Meditec, Dublin, CA) with software version 1.7.31492. The nominal 6×6 -mm scanning protocol was used, sampling at a 500×500 resolution at a rate of 100,000 A-scans per second at a visual angle of 20.94° . Each B-scan was repeated twice at the same position, and the optical microangiography implemented on the commercial imaging system was used to generate the angiographic information. The A-scan depth was 3 mm with an axial resolution of $6.3 \mu\text{m}$ and a transverse resolution of $20 \mu\text{m}$, as described in the product specifications.

The inner limiting membrane and posterior boundary of the outer plexiform layer were used as the segmentation boundaries for the commercial device, and the inner plexiform layer/inner nuclear layer complex was used as the SCP/DVC boundary. The SCP and DVC were subsequently extracted, with projection artifacts removed, via a built-in software feature in the Zeiss PlexElite and exported at a 1024×1024 resolution. Scans were only included in the study if the system specified signal strength was 8 (out of 10) or higher.

Network Architecture

The network for vessel segmentation used a variation of the U-Net⁴⁰ architecture, which was adapted for two classes: vessel and background. The basic U-Net architecture is shown in Figure 2 and consists of convolutional and pooling layers. The convolutional layers consist of a series of trainable filters, which are correlated across the image and subsequently passed through a rectifier linear unit activation with units capped at 6 (ReLU-6).⁴¹ Each convolutional layer was followed by a batch normalization layer, as well as a dropout layer with a coefficient of 0.5.⁴¹ Pooling layers were inserted to increase the receptive field of the subsequent filters in the convolutional layers, helping with generalization to prevent overfitting.

The U-Net architecture provides high-resolution feature extraction through its structure, which consists of a contracting and expanding path. For each level in the contracting path, high-resolution weights are combined with the upsampled, generalized weights (span connections) in the corresponding level of the

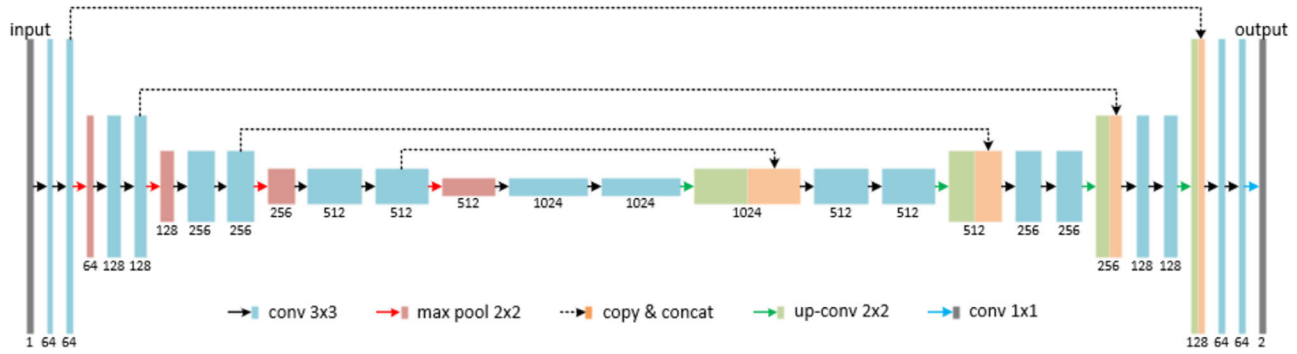


Figure 2. U-net architecture.

expanding path. This allows the network to retain the learned localization information and better segment smaller, more detailed structures in the image. In addition, methods using residual blocks for each convolutional block were investigated; however, such methods resulted in similar or lower performance (data not shown).

Training

Two OCT-A datasets were used for training the network. To construct the initial weights, data were acquired from a previous study.⁴² Briefly, the data consisted of 29 images with a 2 × 2-mm FOV acquired with a prototype swept-source OCT instrument⁴³ and 47 images with a 3 × 3-mm FOV acquired with a commercial spectral-domain OCT instrument. Each OCT-A image was manually segmented using a Surface Pro tablet (Microsoft Corporation, Redmond, WA) and GNU image manipulation program (GIMP) by one trained rater and verified and accepted by two additional trained raters.

To construct the initial weights, the network hyperparameters were optimized using three-fold cross-validation. This resulted in a network trained over 120 epochs using the Adam optimizer, with an initial learning rate of 10⁻⁴ and a custom epsilon value of 10⁻⁵. Evaluation was performed qualitatively on a set of acquired 3 × 3-mm FOV OCT-A images across all devices based on images most recently acquired at the clinic. Segmentation of a single 3 × 3-mm or 2 × 2-mm image using the network took approximately 2 seconds on a Nvidia RTX 2060 graphics card (Nvidia Corporation, Santa Clara, CA), with a possible decrease to 0.3 seconds per image when segmented in larger batches of ~10 images.

The network with the initial weights was subsequently fine-tuned on the 6 × 6-mm FOV images acquired with the PlexElite. This was done in two stages: First, a smaller dataset of 10 single-frame

OCT-A images of each SCP and DVC for which there existed corresponding high-quality averaged images was identified. As described in our previously published study,³⁶ images were registered and averaged based on a template image that was free of microsaccadic motion. This allowed us to use the averaged OCT-A images to construct ground truth labels for each single-frame template image. These labeled ground truth vessel segmentations were subsequently paired with the single-frame template OCT-A images to train the deep neural network to perform segmentations approaching the quality of averaged images, while using only single-frame images (Figure 3).

For the SCP, the automated segmentations of the averaged images generated by the network with initial weights were adequate, as determined by a separate group of trained raters; hence, these results were fed back into the network as additional training examples. However, the automated segmentations on the DVC required additional manual correction due to the lower SNR of the images in this layer and the different morphological features of the vasculature relative to the SCP. Using the initial weights as a guideline for the manual raters would introduce biases that could negatively impact further stages of training; thus, the vasculature of the DVC images was instead first segmented through Otsu’s method.¹⁷ Another masked and trained rater manually corrected the resulting segmentation using a Microsoft Surface Pro tablet and GIMP. All segmentations were reviewed and accepted by two of three other trained raters. A second trained rater segmented three images to obtain inter-rater metrics.

Due to memory limitations when training, each 6 × 6 image was separated into four quadrants, which were saved as four separate images. The same method of augmentation and cross-validation was used. This resulted in the network being trained through the Adam optimizer, using an initial learning rate of 10⁻²

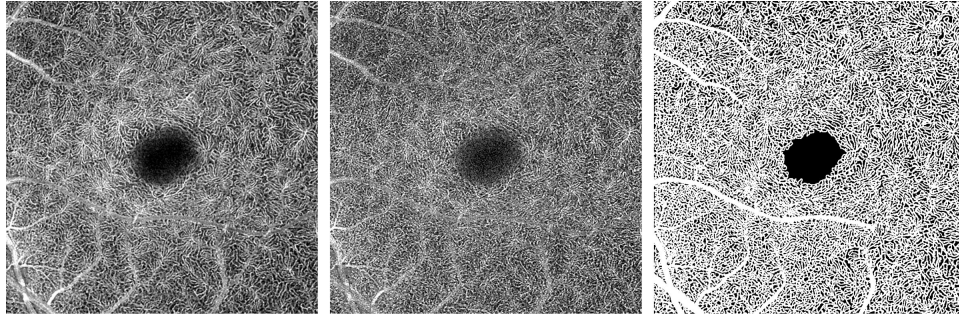


Figure 3. (Left) Ten-frame averaged 6×6 -mm image of the DVC. (Center) Single-frame template image of the same region. (Right) Manually segmented averaged image to be paired with the template image for training.

Table 2. Overview of the Three Datasets Used to Train the Fine-Tuned Network

	Dataset		
	2×2 mm and 3×3 mm	First 6×6 mm	Second 6×6 mm
Training images	SCP: 76 averaged and single-frame	SCP: 10 single-frame DVC: 10 single-frame	SCP: 39 single-frame DVC: 39 single-frame
Ground truth segmentations	SCP: manual	SCP: automated DVC: manual	SCP: automated DVC: automated

and a custom epsilon value of 10^{-2} . The forward inference segmentation of a single 6×6 -mm image using the network took approximately 4 seconds, with a possible decrease to 0.5 seconds per image when segmented in larger batches.

To further reinforce the manually segmented dataset, extensive data augmentation was performed. Each OCT-A image (along with its corresponding manual segmentation) in the training set was rotated 90° three times with no processing. Next, to account for noise, each image was rotated 90° an additional five times with various contrast adjustments, which included contrast-limited adaptive histogram equalization, as well as the built-in `imadjust` function in MATLAB (MathWorks, Natick, MA). To account for motion, each rotated image was also separated into randomly sized strips, which were re-ordered randomly to simulate motion artifacts in the image. The probability maps resulting from the automated segmentations were binarized at a value of 0.5, the default class cutoff in the probability map. After binarization, isolated clusters of fewer than 30 pixels were deemed false positives and removed.

Additional training data were required to improve the performance of the network on the DVC using a single-frame OCT-A image. The intermediate network (after training on the 10 manually segmented images) was able to segment additional 6×6 -mm averaged OCT-A images of the DVC. Subsequently, the next

stage of fine-tuning the CNN involved using this intermediate network to generate additional training data, using manual inspection for quality but not requiring laborious manual segmentation at the scale of capillaries. We applied the intermediate network trained on the manual segmentations to all 50 averaged OCT-A images in our 6×6 -mm dataset. After manual inspection, we identified 39 images for each SCP and DVC with adequate automated segmentations that constituted the new training set. A summary of the training sets used is provided in Table 2.

Similar to what was done in the first stage, the weights were initialized with the original network trained with the first dataset of 2×2 -mm and 3×3 -mm images. The same methods of image augmentation and training were applied, and cross-validation resulted in an initial learning rate of 10^{-2} and a custom epsilon value of 10^{-2} using the Adam optimizer.

Performance Evaluation

To evaluate the automated segmentation performance, a number of metrics were calculated. The number of true positives (TPs), false positives (FPs), false negatives (FNs), and true negatives (TNs) was calculated using pixel-wise comparison between a reference manual segmentation and the thresholded binary output of the neural network. To calculate inter-rater

metrics, these metrics were calculated by comparing one manual segmentation to another by a different rater. In the context of this study, pixels corresponding to vessels and the background comprised the positive and negative classes, respectively. Using the TP, FP, FN, and TN numbers we can calculate the accuracy of the segmentation, as shown in Equation 1:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

Additionally, we can compute the Dice similarity coefficient (DSC), which quantifies the similarity between two segmented images, by measuring the degree of spatial overlap. The DSC values range from 0, indicating no spatial overlap, to 1, indicating complete overlap, and can be calculated by Equation 2:

$$\text{DSC} = \frac{2TP}{2TP + FP + FN} \quad (2)$$

Three methods were tested: using only the initial weights, using a network solely trained on the 6×6 -mm dataset, and using the fine-tuned network trained on both datasets.

Post-Processing of the Automated Vessel Segmentation

The neural network generated segmentation of the vessels in the OCT-A images, but further processing was required for quantitative analysis. Next, the ICAs were identified, as determined by the non-vessel pixels. The largest ICA within a small region in the center of the image was defined as the foveal avascular zone (FAZ). All erroneously segmented pixels within the FAZ were set to a non-vessel classification, with the centroid then used as the center of the Early Treatment of Diabetic Retinopathy Study (ETDRS) grid.

Two metrics were of interest when quantifying ICAs: the area of the region and maximum ischemic point, defined as the point of maximum distance to the nearest vessel within the ICA. The metrics were measured for each ETDRS region and compared to a database of healthy eyes from which the SCP and DVC were extracted. As outlined earlier, eight healthy controls were recruited, resulting in a possible 16 eyes. Of these, we were able to obtain high-quality averaged images for 12 eyes to construct the reference database, and these were also included in the training dataset outlined in Table 2. Each measured ICA was color coded and overlaid on the original image based on the number of standard deviations from the mean. Perifoveal vessel density (for each

ETDRS region) was also calculated as the proportion of measured area occupied by pixels which were classified by the algorithm as a vessel. In addition, the projection artifacts remaining in segmentations of the DVC were excluded for the calculation of vascular metrics, including vessel density and vessel index. This was done through an automated MATLAB post-processing step using image erosion and dilation.

Results

Network Performance Evaluation

Quantitative Segmentation Comparison

Tables 3 and 4 show comparative quantitative results when segmenting the SCP and DVC, respectively. The network trained on the initial dataset of 2×2 -mm and 3×3 -mm images, the network solely trained on the 6×6 -mm dataset, and the network trained with our proposed transfer learning method were labeled as Networks A, B, and C, respectively. The accuracy and Dice index for Network C showed a high similarity between segmentations of the single-frame template images and the averaged images. The inter-rater metrics were conducted only on the manually segmented datasets and are intended to be a representative number illustrating the difficulty of this problem and the variation in the metrics between human raters. Table 5 shows the same networks but evaluated on the original 2×2 -mm and 3×3 -mm dataset.

Qualitative Segmentation Comparison

The fine-tuned network was qualitatively evaluated on data unseen by the CNN during training on control and DR patients. Figure 4 focuses on a peripheral area of the SCP located close to the optic nerve head, where the elongated vascular structure of radial peripapillary capillaries (RPCs) are visible. As shown in Figure 4C2, Network C (the fine-tuned network) preserved the features characteristic of the RPCs and segmented larger vessels more accurately than did Network A (the initial weights). The differences between Network B (trained solely with the 6×6 -mm dataset) and Network C are less pronounced, due to the higher SNR present in images of the SCP.

Figure 5 shows an additional enlarged comparison for an image of the SCP. It can be seen here that Network C was able to accurately segment the areas of ischemia observed in the averaged images and, as also shown in Figure 4, segmented larger vessels more accurately. Figure 6 shows an enlarged comparison of the segmentations results obtained by Network A and Network C when segmenting the elongated, lobular

Table 3. Comparative Quantitative Results of the Segmentation of the SCP Among Three Networks

	Network A	Network B	Network C	Inter-Rater
Accuracy	0.8141	0.8534	0.8599	0.8300
Dice similarity index	0.8060	0.8586	0.8618	0.6700

Network A consisted of only the initial weights, Network B was trained on only the images from the 6×6 -mm dataset, and Network C was the fine-tuned network using our proposed transfer learning method.

Table 4. Comparative Quantitative Results of the Segmentation of the DVC Among Three Networks

	Network A	Network B	Network C	Inter-Rater
Accuracy	0.6934	0.7822	0.7986	0.6874
Dice similarity index	0.6469	0.8065	0.8139	0.7416

Network A consisted of only the initial weights, Network B was trained on only the images from the 6×6 -mm dataset, and Network C was the fine-tuned network using our proposed transfer learning method.

Table 5. Comparative Quantitative Results of the Segmentation of the 2×2 -mm and 3×3 -mm Dataset Among Three Networks

	Network A	Network B	Network C	Inter-rater
Accuracy	0.8677	0.8329	0.8350	0.8300
Dice similarity index	0.8395	0.8059	0.8066	0.6700

Network A consisted of only the initial weights, Network B was trained on only the images from the 6×6 -mm dataset, and Network C was the fine-tuned network using our proposed transfer learning method.

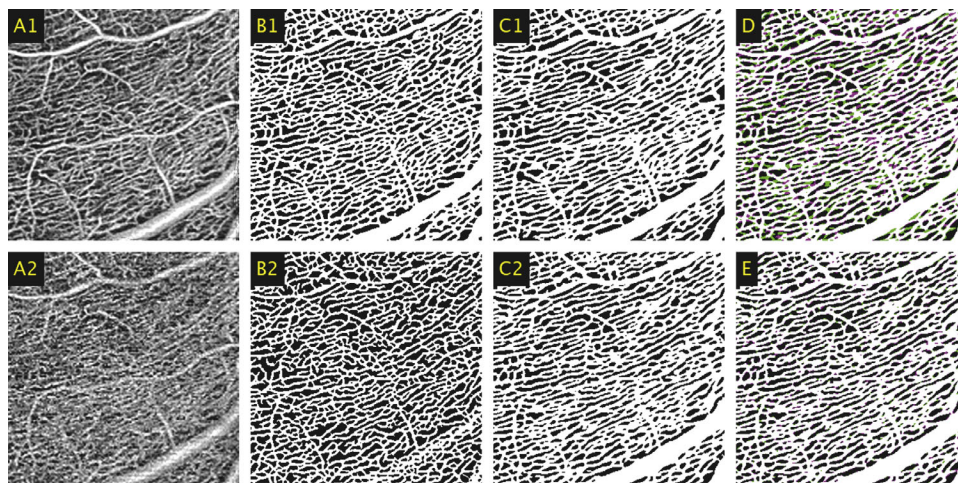


Figure 4. (A1) A 2×2 -mm window of an averaged 6×6 -mm image taken of the SCP. (A2) A 2×2 -mm window of the corresponding region in the equivalent single-frame template image. (B1) Averaged image segmented using the initial weights (Network A). (B2) Single-frame image segmented using Network A. (C1) Averaged image segmented using the fine-tuned network (Network C). (C2) Single-frame image segmented using Network C. (D) Comparison between the automated segmentations between the averaged and single-frame images produced by Network C, represented by *magenta* and *green*, respectively, with *white* representing the union. (E) Comparison between single-frame segmentations between Network B and C, represented by *magenta* and *green*, respectively, with *white* representing the union.

capillary structure of the DVC in a lower quality image. As shown in [Figure 6B2](#), certain clusters of vessels were erroneously treated as noise by Network A, resulting in regions of false negatives. This is characteristic of single-frame OCT-A images; the blurred-out

regions were replaced by a discernible vessel structure when using the averaged OCT-A images. The results presented in [Figure 6C2](#) are representative of the outputs from Network C, which eliminated a portion of these false negatives.

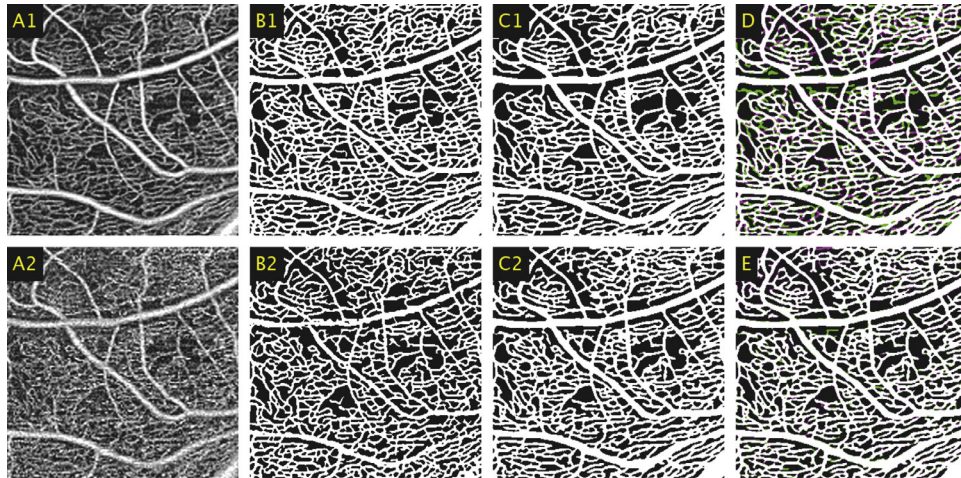


Figure 5. (A1) A 2×2 -mm window of an averaged 6×6 -mm image taken of the SCP. (A2) A 2×2 -mm window of the corresponding region in the equivalent single-frame template image. (B1) Averaged image segmented using the initial weights (Network A). (B2) Single-frame image segmented using Network A. (C1) Averaged image segmented using the fine-tuned network (Network C). (C2) Single-frame image segmented using Network C. (D) Comparison between the automated segmentations between the averaged and single-frame images produced by Network C, represented by *magenta* and *green*, respectively, with *white* representing the union. (E) Comparison between single-frame segmentations between Network B and C, represented by *magenta* and *green*, respectively, with *white* representing the union.

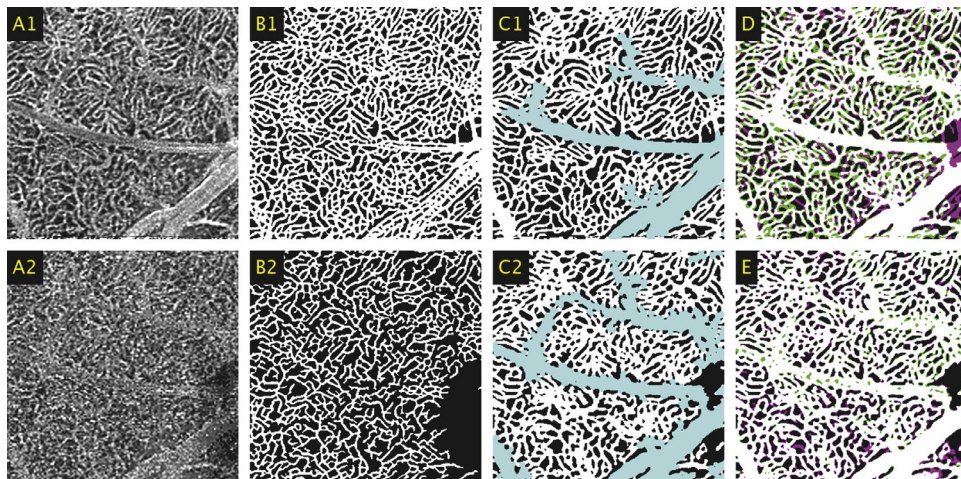


Figure 6. (A1) A 2×2 -mm window of an averaged 6×6 -mm image taken of the DVC. (A2) A 2×2 -mm window of the corresponding region in the equivalent single-frame template image. (B1) Averaged image segmented using the initial weights (Network A). (B2) Single-frame image segmented using Network A. (C1) Averaged image segmented using the fine-tuned network (Network C), with projection artifacts to be excluded highlighted in *cyan*. (C2) Single-frame image segmented using Network C, with projection artifacts to be excluded highlighted in *cyan*. (D) Comparison between the automated segmentations between the averaged and single-frame images produced by Network C, represented by *magenta* and *green*, respectively, with *white* representing the union. (E) Comparison between single-frame segmentations between Network B and C, represented by *magenta* and *green*, respectively, with *white* representing the union.

Residual projection artifacts from the SCP were automatically identified and are highlighted in cyan in Figures 6C1 and 6C2. As outlined earlier, these regions obscured the capillaries underneath and were subsequently excluded as a post-processing step when calculating the vessel density and vessel index metrics. When segmenting with Network A, these projection artifacts were erroneously segmented as additional capillaries. This can also be seen in Figure 6E where

the green spots along the projection artifacts indicate that the segmentation produced by Network B was not as continuous as for Network C. With accurate and continuous projection artifact delineation, these can be more accurately removed in post-processing.

Figure 7 shows an additional enlarged comparison of the results when segmenting the DVC with different versions of the deep neural network. The images segmented by Network C, shown in Figures 7C1

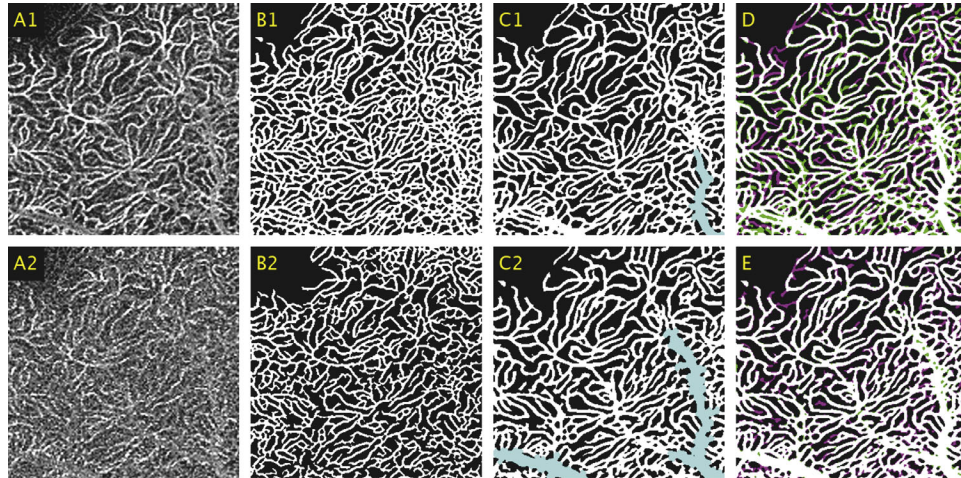


Figure 7. (A1) A 2×2 -mm window of an averaged 6×6 -mm image taken of the DVC. (A2) A 2×2 -mm window of the corresponding region in the equivalent single-frame template image. (B1) Averaged image segmented using the initial weights (Network A). (B2) Single-frame image segmented using Network A. (C1) Averaged image segmented using the fine-tuned network (Network C), with projection artifacts to be excluded highlighted in cyan. (C2) Single-frame image segmented using Network C, with projection artifacts to be excluded highlighted in cyan. (D) Comparison between the automated segmentations between the averaged and single-frame images produced by Network C, represented by magenta and green, respectively, with white representing the union. (E) Comparison between single-frame segmentations between Network B and C, represented by magenta and green, respectively, with white representing the union.

and 7C2, more closely follow the elongated, lobular ICA morphology of the DVC and the results were less prone to over-segmenting noise. This presents a substantial improvement over the images segmented by Network A, as shown in Figures 7B1 and 7-B2, the results of which incorrectly applied the branching structure characteristic of the SCP to the DVC.

Intercapillary Area Evaluation

Figure 8 shows representative images, segmentations, and standard deviation maps for diabetic subjects without DR, with mild/moderate non-proliferative DR, and with severe non-proliferative DR as graded by a retina specialist.

Discussion

Early detection of DR is paramount to ensuring effective treatment and improved patient quality of life. Changes in both the SCP and DVC have been identified as potential early biomarkers of DR. As a result, accurate segmentation and quantification of increasingly widefield images of both the SCP and DVC will provide further insight into the emergence and progression of DR.

We designed a transfer learning-based framework for automated segmentation of the microvasculature in the SCP and DVC, as well as quantification of

the ICAs in 6×6 -mm single-frame OCT-A images. The framework consists of two convolutional neural networks: an initial network trained on 2×2 -mm and 3×3 -mm images and a second network that utilized the pre-existing weights and fine-tuning on a smaller dataset of 6×6 -mm images of both the SCP and DVC. This approach allowed for accurate feature detection despite a limited training set, with results that exceed the intra-rater accuracy. In particular, fine-tuning from an existing set of data provided more robust projection artifact delineation in the DVC, allowing for removal in post-processing when computing vascular metrics. The resulting ICA quantifications allow for a closer investigation into suspected areas of low perfusion but do not expressly define what constitutes such areas.

A prevailing limitation of many machine learning problems is training dataset acquisition. For our study, manually segmenting an individual 6×6 -mm image of the DVC took each rater an average of 4 hours to complete, which can pose a significant challenge for problems requiring larger datasets. Solely training a new network on our limited, manually segmented 6×6 -mm dataset would overfit to the training set, and including this new dataset with the original would result in a heavy data imbalance. The introduction of additional automated segmentations of averaged images greatly increased the size and quality of our training set, from 10 images of each SCP and DVC to 39 images. This allowed for a larger variation in training samples, consequently improving network performance.

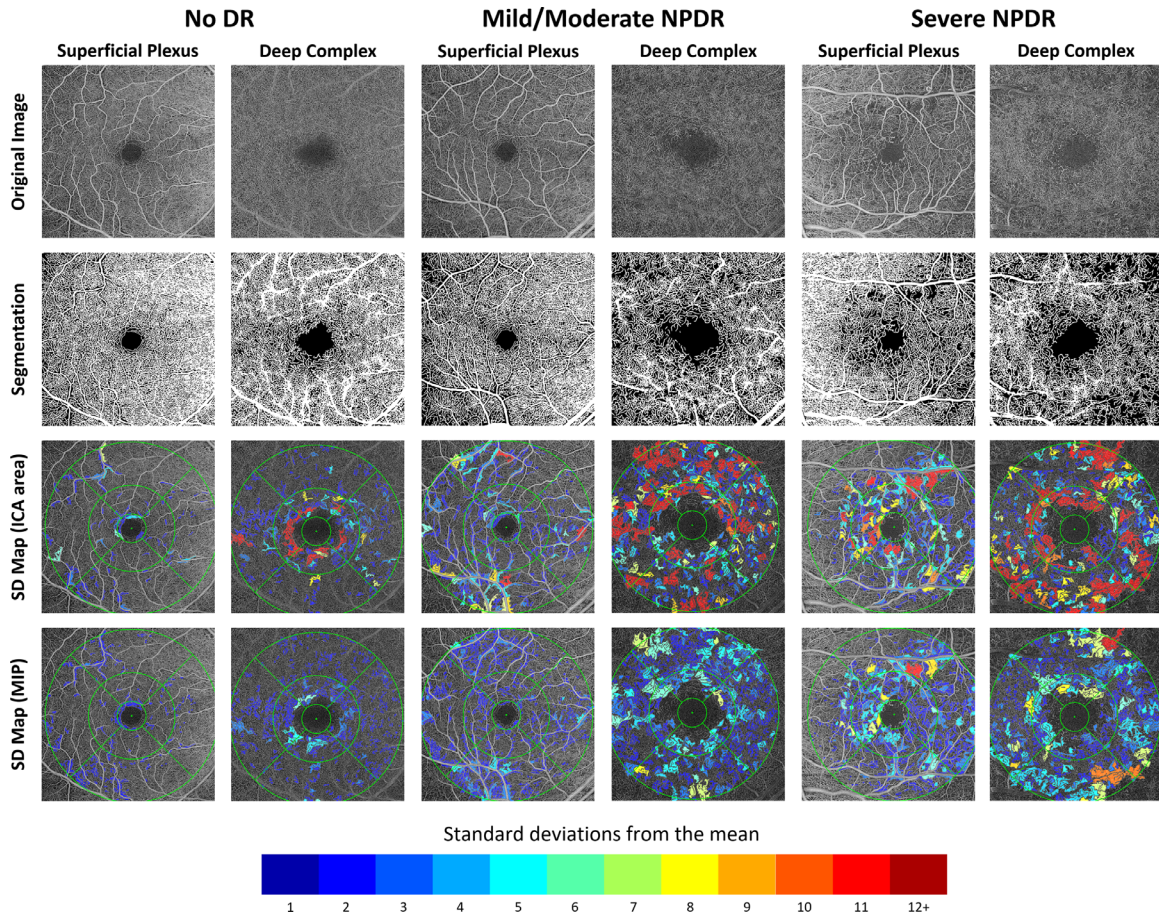


Figure 8. Labeled standard deviation maps for subjects with no diabetic retinopathy, mild/moderate non-proliferative diabetic retinopathy, or severe non-proliferative diabetic retinopathy. Original images of the DVC have been brightened for clarity. All intercapillary areas are labeled based on the number of standard deviations its maximum ischemic point and area exceeded the reference mean.

The impact of the training examples is most evident in the segmentations of the DVC, where the initial weights produced segmentations that differed significantly from the images produced by the fine-tuned network. As seen in Figures 6 and 7, vessels segmented by the initial weights closely resembled the denser morphological characteristics of the SCP. In particular, the ICAs in the DVC followed a lobular pattern, which was reflected more accurately in the segmentations generated by the fine-tuned network.

Another limitation is the quality of the data. Images with a signal strength index of 8 or lower, as well as images with excessive microsaccadic eye motion, were omitted from the study. If there are excessive microsaccadic artifacts, microvascular features begin to blur and can be subsequently classified as noise by the network. This emphasizes the importance of using the averaged 6 × 6-mm images as the ground truth data obtained from manual segmentations because it will be the most anatomically accurate. Our previously published method of averaging and registering

single-frame images based on a template³⁶ allowed for segmentations of averaged images to be paired with single-frame training data, greatly improving the quality of our training samples. Segmentation quality appeared to be independent of location in the image, as automated segmentation accuracy was consistent across the 6 × 6-mm FOV in the absence of additional artifacts.

To summarize, we designed a machine learning framework to accurately segment and quantify the retinal microvasculature in the SVP and the DVC. It produces immediately available segmentations that provide clinicians with a tool for in-depth analysis of ICAs and the level of retinal perfusion. Through this framework, patient care for DR can be adapted individually, improving compliance and patient prognosis. In addition, visualization and quantification of retinal vasculature at a high level of accuracy provide more information about disease activity and therefore may add to individualized patient care.

Acknowledgments

Supported by funding from the following agencies: Diabetes Action Canada, National Sciences and Engineering Research Council of Canada, Canadian Institutes of Health Research, Alzheimer Society Canada, and Michael Smith Foundation for Health Research.

Disclosure: **J. Lo**, None; **M. Heisler**, None; **V. Vanzan**, None; **S. Karst**, None; **I.Z. Matovinović**, None; **S. Lončarić**, None; **E.V. Navajas**, None; **M.F. Beg**, None; **M.V. Šarunić**, Seymour Vision, Inc. (I)

References

1. CNIB Foundation. Blindness in Canada. Available at: <https://cnib.ca/en/sight-loss-info/blindness/blindness-canada?region=bc>. Accessed June 24, 2020.
2. Nentwich MM, Ulbig MW. Diabetic retinopathy - ocular complications of diabetes mellitus. *World J Diabetes*. 2015;6:489–499.
3. Kashani AH, Chen C-L, Gahm JK, et al. Optical coherence tomography angiography: a comprehensive review of current methods and clinical applications. *Prog Retin Eye Res*. 2017;60:66–100.
4. Jia Y, Bailey ST, Hwang TS, et al. Quantitative optical coherence tomography angiography of vascular abnormalities in the living human eye. *Proc Natl Acad Sci USA*. 2015;112:E2395–E2402.
5. An D, Balaratnasingam C, Heisler M, et al. Quantitative comparisons between optical coherence tomography angiography and matched histology in the human eye. *Exp Eye Res*. 2018;170:13–19.
6. Balaratnasingam C, An D, Sakurada Y, et al. Comparisons between histology and optical coherence tomography angiography of the periarterial capillary-free zone. *Am J Ophthalmol*. 2018;189:55–64.
7. Tan PEZ, Balaratnasingam C, Xu J, et al. Quantitative comparison of retinal capillary images derived by speckle variance optical coherence tomography with histology. *Invest Ophthalmol Vis Sci*. 2015;56:3989–3996.
8. Yu D-Y, Cringle SJ, Yu PK, et al. Retinal capillary perfusion: spatial and temporal heterogeneity. *Prog Retin Eye Res*. 2019;70:23–54.
9. Yannuzzi LA, Rohrer KT, Tindel LJ, et al. Fluorescein angiography complication survey. *Ophthalmology*. 1986;93:611–617.
10. Silva PS, Cavallerano JD, Haddad NMN, et al. Peripheral lesions identified on ultrawide field imaging predict increased risk of diabetic retinopathy progression over 4 years. *Ophthalmology*. 2015;122:949–956.
11. Campbell JP, Zhang M, Hwang TS, et al. Detailed vascular anatomy of the human retina by projection-resolved optical coherence tomography angiography. *Sci Rep*. 2017;7:42201.
12. Carnevali A, Sacconi R, Corbelli E, et al. Optical coherence tomography angiography analysis of retinal vascular plexuses and choriocapillaris in patients with type 1 diabetes without diabetic retinopathy. *Acta Diabetol*. 2017;54:695–702.
13. Hasegawa N, Nozaki M, Takase N, Yoshida M, Ogura Y. New insights into microaneurysms in the deep capillary plexus detected by optical coherence tomography angiography in diabetic macular edema. *Invest Ophthalmol Vis Sci*. 2016;57:OCT348–OCT355.
14. Srinidhi CL, Aparna P, Rajan J. Recent advancements in retinal vessel segmentation. *J Med Syst*. 2017;41:70.
15. Krawitz BD, Phillips E, Bavier RD, et al. Parafoveal nonperfusion analysis in diabetic retinopathy using optical coherence tomography angiography. *Transl Vis Sci Technol*. 2018;7:4.
16. Hwang TS, Gao SS, Liu L, et al. Automated quantification of capillary nonperfusion using optical coherence tomography angiography in diabetic retinopathy. *JAMA Ophthalmol*. 2016;134:367–373.
17. Otsu N. A threshold selection method from gray-level histograms. *IEEE Trans Syst Man Cybern*. 1979;9:62–66.
18. Schottenhamml J, Moulton EM, Ploner S, et al. An automatic, intercapillary area-based algorithm for quantifying diabetes-related capillary dropout using optical coherence tomography angiography. *Retina*. 2016;36(Suppl 1):S93–S101.
19. Zhu TP, Li EH, Li JY, et al. Comparison of projection-resolved optical coherence tomography angiography-based metrics for the early detection of retinal microvascular impairments in diabetes mellitus. *Retina*. 2019, doi:10.1097/IAE.0000000000002655.
20. Li A, You J, Du C, Pan Y. Automated segmentation and quantification of OCT angiography for tracking angiogenesis progression. *Biomed Opt Express*. 2017;8:5604–5616.
21. Roy AG, Conjeti S, Karri SPK, et al. ReLayNet: retinal layer and fluid segmentation of macular optical coherence tomography using fully convolutional networks. *Biomed Opt Express*. 2017;8:3627.

22. Maloca PM, Lee AY, de Carvalho ER, et al. Validation of automated artificial intelligence segmentation of optical coherence tomography images. *PLoS One*. 2019;14:e0220063.
23. 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;8:2732–2744.
24. Zhu Q, Xing X, Zhu M, et al. A new approach for the segmentation of three distinct retinal capillary plexuses using optical coherence tomography angiography. *Transl Vis Sci Technol*. 2019;8:57.
25. Heisler M, Ju MJ, Bhalla M, et al. Automated identification of cone photoreceptors in adaptive optics optical coherence tomography images using transfer learning. *Biomed Opt Express*. 2018;9:5353–5367.
26. Cunefare D, Huckenpahler AL, Patterson EJ, Dubra A, Carroll J, Farsiu S. RAC-CNN: multimodal deep learning based automatic detection and classification of rod and cone photoreceptors in adaptive optics scanning light ophthalmoscope images. *Biomed Opt Express*. 2019;10:3815–3832.
27. Schlegl T, Waldstein SM, Bogunovic H, et al. Fully automated detection and quantification of macular fluid in OCT using deep learning. *Ophthalmology*. 2018;125:549–558.
28. Ji Z, Chen Q, Niu S, Leng T, Rubin DL. Beyond retinal layers: a deep voting model for automated geographic atrophy segmentation in SD-OCT images. *Transl Vis Sci Technol*. 2018;7:1.
29. Lu W, Tong Y, Yu Y, Xing Y, Chen C, Shen Y. Deep learning-based automated classification of multi-categorical abnormalities from optical coherence tomography images. *Transl Vis Sci Technol*. 2018;7:41.
30. De Fauw J, Ledsam JR, Romera-Paredes B, et al. Clinically applicable deep learning for diagnosis and referral in retinal disease. *Nat Med*. 2018;24:1342–1350.
31. Kermany DS, Goldbaum M, Cai W, et al. Identifying medical diagnoses and treatable diseases by image-based deep learning. *Cell*. 2018;172:1122–1131.e9.
32. Zhao H, Li H, Maurer-Stroh S, Cheng L. Synthesizing retinal and neuronal images with generative adversarial nets. *Med Image Anal*. 2018;49:14–26.
33. Ometto G, Moghul I, Montesano G, et al. ReLayer: a free, online tool for extracting retinal thickness from cross-platform OCT images. *Transl Vis Sci Technol*. 2019;8:25.
34. 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;9:5147–5158.
35. Prentašić P, Heisler M, Mammo Z, et al. Segmentation of the foveal microvasculature using deep learning networks. *J Biomed Opt*. 2016;21:075008.
36. Heisler M, Lee S, Mammo Z, et al. Strip-based registration of serially acquired optical coherence tomography angiography. *J Biomed Opt*. 2017;22:36007.
37. Uji A, Balasubramanian S, Lei J, et al. Multiple enface image averaging for enhanced optical coherence tomography angiography imaging. *Acta Ophthalmol*. 2018;96:e820–e827.
38. Schmidt TG, Linderman RE, Strampe MR, Chui TYP, Rosen RB, Carroll J. The utility of frame averaging for automated algorithms in analyzing retinal vascular biomarkers in AngioVue OCTA. *Transl Vis Sci Technol*. 2019;8:10.
39. Wilkinson CP, Ferris FL, Klein RE, et al. Proposed international clinical diabetic retinopathy and diabetic macular edema disease severity scales. *Ophthalmology*. 2003;110:1677–1682.
40. Ronneberger O, Fischer P, Brox T. U-Net: convolutional networks for biomedical image segmentation. Available at: <https://arxiv.org/pdf/1505.04597.pdf>. Accessed June 24, 2020.
41. Krizhevsky A. Convolutional deep belief networks on CIFAR-10. Available at: <http://www.cs.utoronto.ca/~kriz/conv-cifar10-aug2010.pdf>. Accessed June 24, 2020.
42. Heisler M, Chan F, Mammo Z, et al. Deep learning vessel segmentation and quantification of the foveal avascular zone using commercial and prototype OCT-A platforms. Available at: <https://arxiv.org/ftp/arxiv/papers/1909/1909.11289.pdf>. Accessed June 24, 2020.
43. Xu J, Han S, Balaratnasingam C, et al. Retinal angiography with real-time speckle variance optical coherence tomography. *Br J Ophthalmol*. 2015;99:1315–1319.