

# Automatic Segmentation of Retinal Capillaries in Adaptive Optics Scanning Laser Ophthalmoscope Perfusion Images Using a Convolutional Neural Network

Gwen Musial<sup>1</sup>, Hope M. Queener<sup>2</sup>, Suman Adhikari<sup>2</sup>, Hanieh Mirhajianmoghadam<sup>2</sup>, Alexander W. Schill<sup>1,2</sup>, Nimesh B. Patel<sup>2</sup>, and Jason Porter<sup>1,2</sup>

<sup>1</sup> Department of Biomedical Engineering, University of Houston, Houston, TX, USA

<sup>2</sup> College of Optometry, University of Houston, Houston, TX, USA

**Correspondence:** Gwen Musial, Department of Biomedical Engineering, University of Houston, 3517 Cullen Blvd, Room 2027, Houston, TX 77204-5060, USA. e-mail: [gmusial@uh.edu](mailto:gmusial@uh.edu)

**Received:** October 23, 2019

**Accepted:** June 2, 2020

**Published:** July 23, 2020

**Keywords:** machine learning; convolutional neural network; adaptive optics; vascular segmentation; capillary imaging

**Citation:** Musial G, Queener HM, Adhikari S, Mirhajianmoghadam H, Schill AW, Patel NB, Porter J. Automatic segmentation of retinal capillaries in adaptive optics scanning laser ophthalmoscope perfusion images using a convolutional neural network. *Trans Vis Sci Tech.* 2020;9(2):43. <https://doi.org/10.1167/tvst.9.2.43>

**Purpose:** Adaptive optics scanning laser ophthalmoscope (AOSLO) capillary perfusion images can possess large variations in contrast, intensity, and background signal, thereby limiting the use of global or adaptive thresholding techniques for automatic segmentation. We sought to develop an automated approach to segment perfused capillaries in AOSLO images.

**Methods:** 12,979 image patches were extracted from manually segmented AOSLO montages from 14 eyes and used to train a convolutional neural network (CNN) that classified pixels as capillaries, large vessels, background, or image canvas. 1764 patches were extracted from AOSLO montages of four separate subjects, and were segmented manually by two raters (ground truth) and automatically by the CNN, an Otsu's approach, and a Frangi approach. A modified Dice coefficient was created to account for slight spatial differences between the same manually and CNN-segmented capillaries.

**Results:** CNN capillary segmentation had an accuracy (0.94), a Dice coefficient (0.67), and a modified Dice coefficient (0.90) that were significantly higher than other automated approaches ( $P < 0.05$ ). There were no significant differences in capillary density and mean segment length between manual ground-truth and CNN segmentations ( $P > 0.05$ ).

**Conclusions:** Close agreement between the CNN and manual segmentations enables robust and objective quantification of perfused capillary metrics. The developed CNN is time and computationally efficient, and distinguishes capillaries from areas containing diffuse background signal and larger underlying vessels.

**Translational Relevance:** This automatic segmentation algorithm greatly increases the efficiency of quantifying AOSLO capillary perfusion images.

## Introduction

Changes in vascular structure and perfusion are known to contribute to several systemic, retinal, and optic nerve head pathologies. For example, cross-sectional studies have reported radial peripapillary capillary dropout in eyes of patients with primary open-angle glaucoma<sup>1-3</sup> and Alzheimer's disease,<sup>4,5</sup>

while changes in capillary morphology have been observed near the fovea in diabetic patients.<sup>6,7</sup> With the advent of optical coherence tomography angiography (OCTA)<sup>8</sup> and adaptive optics scanning laser ophthalmoscope (AOSLO) imaging techniques,<sup>9-11</sup> it is now possible to non-invasively image perfused vasculature, including the smallest of capillaries within the retina and surrounding the optic nerve head. Quantification of perfused retinal capillaries may prove valuable

for developing sensitive biomarkers to earlier diagnose and monitor the progression of ocular and systemic diseases.<sup>4,5,12–14</sup>

The quantification of capillary parameters typically requires a segmented, binary image. Although manual segmentation of the capillaries in grayscale AOSLO or OCTA images can be performed, such a procedure is largely subjective and requires many hours for completion by a skilled observer. Therefore, automated methods are desirable to facilitate objective and time-efficient image analyses. Different global and adaptive thresholding algorithms have been explored for automated segmentation of vasculature in fundus images and, more recently, in OCTA images, including Otsu's method<sup>15–17</sup> and multi-scale vesselness filters (e.g., Frangi filter).<sup>22–25</sup> However, translation of these thresholding algorithms to grayscale AOSLO perfusion images for the purposes of automatically segmenting retinal vasculature has proven challenging, primarily due to the large variations in contrast, brightness, and background signal that can typically manifest in AOSLO perfusion images. Machine learning techniques, such as convolutional neural networks (CNNs), have been developed for fundus<sup>26–28</sup> and OCTA<sup>29</sup> images. However, there is a lack of development in comparable techniques for AOSLO images.

The purpose of this work was to design an automated algorithm that can segment the radial peripapillary capillary network in grayscale AOSLO perfusion images with high repeatability. We developed a CNN that distinguishes capillaries from major vasculature and background signal. The network was trained on grayscale AOSLO perfusion images acquired in healthy human eyes, in healthy non-human primate eyes, and in non-human primate eyes with laser-induced experimental glaucoma. The algorithm's performance was evaluated by computing its accuracy and Dice similarity coefficient, with manually marked images serving as the ground truth, and was compared with the same measures obtained using traditional segmentation techniques. Capillary metrics were also calculated for images following manual, CNN, and traditional segmentations and were compared among the approaches.

## Methods

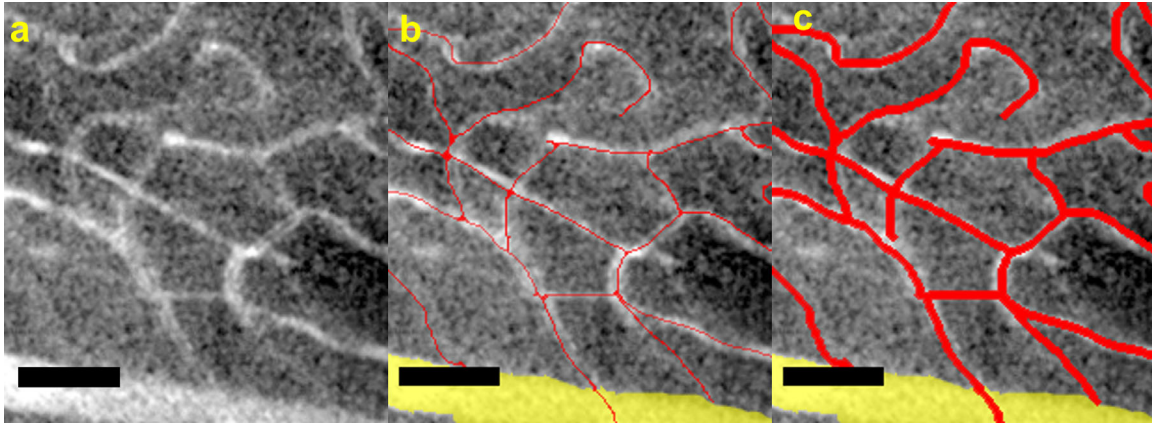
All human subject research procedures were approved by the University of Houston's institutional review board and adhered to the tenets of the Declaration of Helsinki. Informed consent was

obtained from each human subject prior to performing any experimental procedures. All animal care experimental procedures were approved by the University of Houston's Institutional Animal Care and Use Committee and adhered to the ARVO Statement for the Use of Animals in Ophthalmic and Vision Research.

## Adaptive Optics Scanning Laser Ophthalmoscope Imaging

The pupil of each subject was dilated prior to imaging using 2.5% phenylephrine and 1% tropicamide. Pupils were then centered on the optical axis of the AOSLO using a bite bar attached to a three-dimensional translation stage. Human subjects were instructed to fixate on a laser pointer projected onto a fixation target. The pointer was moved on the target until a portion of the subject's optic nerve head (ONH) was within the field of view of the AOSLO. Non-human primate (NHP) subjects (rhesus monkeys, *Macaca mulatta*) were anesthetized with 20 to 25 mg/kg ketamine and 0.8 to 0.9 mg/kg xylazine to minimize eye movements during imaging.<sup>30</sup> The head of each monkey was positioned using a head mount attached to a three-dimensional translation stage and was steered using the tip, tilt, and rotation capabilities of the head mount until the monkey's ONH was within the field of view of the AOSLO. Monkey eyelids were held open using a lid speculum. Imaging was performed while monkeys wore a rigid gas-permeable contact lens, which was used to prevent corneal dehydration and to correct for any inherent spherical refractive errors.<sup>31</sup>

En face reflectance videos of the most superficial retinal nerve fiber layer (RNFL) axon bundles were acquired using a confocal AOSLO imaging channel (200- $\mu\text{m}$  pinhole diameter) over a 2° field of view at a rate of 25 Hz using a superluminescent diode (SLD) light source (S-Series Broadlighter; Superlum, Carrigtwohill, Ireland) with a center wavelength of 840 nm (full width at half maximum = 50 nm). The power of the SLD at the corneal plane was 150  $\mu\text{W}$ , a value that was more than 10 times below the maximum permissible exposure for an imaging duration of 1 hour.<sup>32</sup> When imaging the most superficial retina near the ONH, the confocal imaging channel yielded high-contrast images of the RNFL axon bundles (which directly backscatter light) with very limited visualization of capillary structure; therefore, non-confocal, split-detector<sup>33</sup> AOSLO videos of blood flow perfusion were collected simultaneously with the confocal videos at the same retinal location and depth. The split-detector channel emphasizes local changes in the index of refraction.<sup>17</sup>



**Figure 1.** Dilution of a single pixel manual trace with a uniform structuring element closely matches the diameter of perfused retinal capillaries in AOSLO images. (a) Original grayscale image of perfused retinal capillaries near a larger vessel. (b) Single-pixel manual trace through the center of perfused capillaries (red) and the manual tracing of a larger vessel (yellow). (c) Uniform dilation of the manual trace with a disk ( $r = 5$  pixels) closely matches the diameter of perfused capillaries. Scale bar: 50  $\mu\text{m}$ .

Consequently, split-detector videos were used to calculate perfusion images at the retinal plane of focus.<sup>34</sup>

Using DeMotion, a cross-correlation program based on CUDA (NVIDIA Corporation, Santa Clara, CA),<sup>35</sup> individual frames from AOSLO confocal videos were subdivided into strips and registered with respect to a pre-selected reference frame to remove eye motion and create a stabilized video. The same offsets were then applied to identical strips from each frame in the corresponding split-detector video to generate stabilized split-detector videos. Perfusion images of the radial peripapillary capillaries were generated following an approach similar to that of Chui et al.<sup>11</sup> After compensation for intra- and interframe eye motion, videos were normalized to the maximum pixel value within the entire video. To reduce noise, this result was median filtered using a  $3 \times 3$ -pixel kernel. To limit the influence of slower (less than 0.5 Hz) tissue reflectance changes, the 150-to-250-frame stabilized video was divided into 25-frame intervals. The standard error of each pixel was computed over each 25-frame interval, and a frame containing the standard error of each pixel (or a standard-error frame) was generated for each interval. After applying a  $3 \times 3$ -pixel median filter to each standard error frame, all frames were averaged and subsequently normalized by the maximum value. Histogram stretching was applied so that the lower and upper 1% of the histogram were set to 0 and 255, respectively, in the resulting perfusion image. Multiple images were taken along the ONH rim and manually stitched together (Adobe Photoshop, Adobe Systems, San Jose, CA) to generate a larger perfusion montage (Fig. 3a).

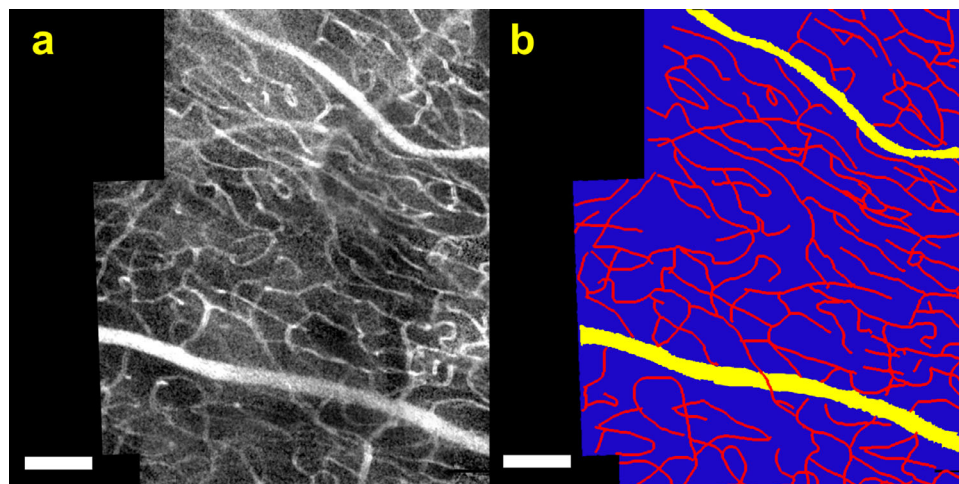
## Manual Segmentation of Perfused Capillaries

Manual tracing of perfused retinal capillaries was performed in Adobe Photoshop using a Wacom tablet (Wacom, Saitama, Japan). Perfused capillaries, defined as bright tube-like structures that were no wider than 20 pixels across ( $\sim 20 \mu\text{m}$  in width), were manually traced through the center of the capillary using a 1-pixel line (Fig. 1b). The single-pixel-diameter trace was dilated using a uniform structuring element (disk) with a radius of 5 pixels to mimic the average diameter of the capillaries in the perfusion images (Fig. 1c). Major vasculature was manually traced with a variable pencil size larger than 20 pixels (Fig. 1).

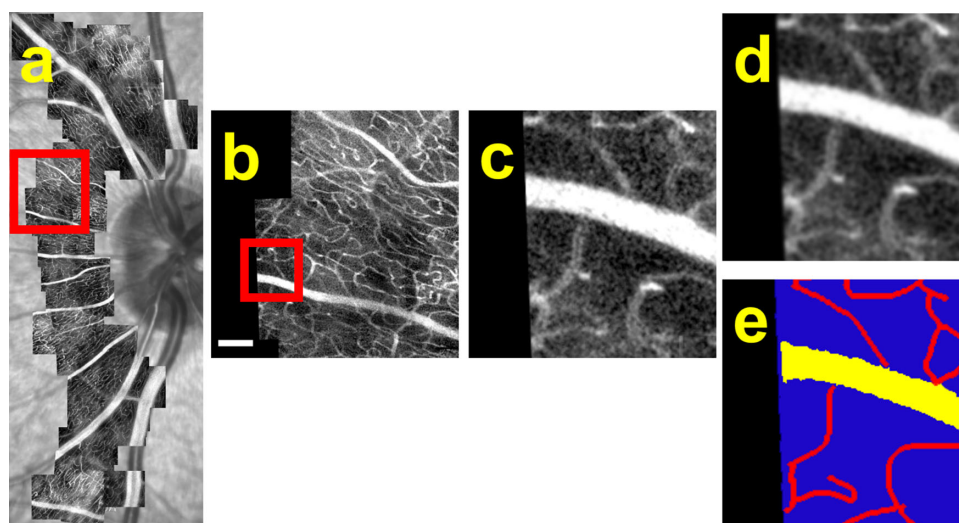
The manual marking resulted in images with pixels belonging to one of four ground-truth classes (Fig. 2b)—capillary (red), large vessel (yellow), background (blue), and image canvas (black). Canvas pixels result from the non-uniform perimeter of the montage and the padding required to achieve uniform rectangular images for processing by the neural network. Canvas pixels have an intensity of 0, whereas AOSLO background signal pixels have variable intensity and can be any grayscale value. One rater (rater A) segmented all 185 AOSLO images used for training, and two raters (rater A and rater B) segmented 14 images used for testing CNN performance.

## Subjects and Training Dataset

The training dataset consisted of six montages from four healthy adult human eyes (mean age,



**Figure 2.** Training dataset was manually segmented into four classes. (a) Original grayscale image cropped from a larger AOSLO montage. (b) Multi-class image showing the four classes of ground truth: capillaries (*red*), large vessel (*yellow*), background (*blue*), and image canvas (*black*). Scale bars: 200  $\mu$ m.



**Figure 3.** AOSLO perfusion montage from a single subject was subdivided into patches for input into the convolutional neural network. (a) Original grayscale AOSLO montage overlaid on the corresponding SLO image from the same eye. (b) Example of one of the smaller image regions ( $768 \times 768$  pixels) cropped from the larger montage in (a) denoted by the *red square*. Scale bar: 200  $\mu$ m. (c) Example of a  $128 \times 128$ -pixel patch extracted from the smaller image region in (b) denoted by the *red square*. The 185-image dataset generated 12,979 such patches. (d) Output generated after applying a Gaussian filter with a kernel size of 3 to the patch in (c). (e) Representation of (c) according to different pixel classes (*red*, capillary; *yellow*, large vessel; *blue*, background; *black*, image canvas), which serves as the ground truth for CNN training.

28.4  $\pm$  1.1 years) with no ocular pathology and a best-corrected visual acuity of 20/20 or better, as well as 11 montages from six healthy NHP eyes and six montages from four NHP eyes with laser-induced experimental glaucoma.<sup>36</sup> Due to variability in the amount of retina imaged within a given imaging session, some eyes in the training dataset were imaged more than once over a time span of at least 1 week. After manually marking perfused vasculature, smaller

image regions were extracted from each AOSLO perfusion montage and were cropped and/or padded to a uniform  $768 \times 768$  pixels for computation, resulting in 185 multi-class images (Fig. 3). Each  $768 \times 768$ -pixel image was subdivided into 50% overlapping patches of  $128 \times 128$  pixels. We required that no more than half of the pixels within a patch belong to the image canvas class in order to decrease computation time. This process yielded 12,979 patches, 75% of which

**Table 1.** Network Layer Architecture

Layer	Type	Maps	Kernel Size
0	Input	1 map of $128 \times 128$ neurons	—
1	Convolutional	32 maps of $128 \times 128$ neurons	$25 \times 25$
2	Dropout	20% dropout rate	—
3	Batch normalization	—	—
4	Max pooling	—	$4 \times 4$
5	Convolutional	64 maps of $64 \times 64$ neurons	$15 \times 15$
6	Dropout	20% dropout rate	—
7	Batch normalization	—	—
8	Max pooling	—	$4 \times 4$
9	Convolution	128 maps of $32 \times 32$ neurons	$10 \times 10$
10	Upsampling	—	$4 \times 4$
11	Batch normalization	—	—
12	Convolution	64 maps of $64 \times 64$ neurons	$15 \times 15$
13	Dropout	20% dropout rate	—
14	Upsampling	—	$4 \times 4$
15	Batch normalization	—	—
16	Convolution	32 maps of $128 \times 128$ neurons	$25 \times 25$
17	Dropout	20% dropout rate	—
18	Batch normalization	—	—
19	Convolution	4 maps of $128 \times 128$ neurons	$1 \times 1$
20	Fully connected	softmax activation function	—

were randomly selected as the training dataset and the remaining 25% of which comprised the validation set.

### Convolutional Neural Network Architecture

The CNN that we developed to segment capillaries from AOSLO perfusion images was built in Python and based on U-Net, an open-source CNN initially used to segment cells in microscopy images.<sup>37</sup> The software and sample data described in this work are available on GitHub (<https://github.com/porter-lab-software/AOVesselCNN>). Two key steps in the U-Net architecture are (1) contraction for optimizing learned content and (2) a symmetric expansion for precision localization. The network described in this work was based on CNNs previously used to segment vasculature in fundus images<sup>26</sup> and OCTA images<sup>29</sup> and was subsequently altered to optimize the model for automatically segmenting AOSLO perfusion images.

The detailed architecture for our novel CNN is shown in Table 1. The general pattern of including repeating groups of convolution, dropout, batch normalization, and pooling layers is a common feature of CNNs.<sup>38</sup> The CNN begins with a convolutional layer, which convolves an input, or image, with a filter of a specified kernel size. The convolutional response of the filter with the input is passed to the next layer.

Dropout layers within the network prevent overfitting of network units to the training data by randomly removing units from the CNN. For this CNN, max-pooling layers were used to decrease computational demand and to increase the robustness of the network against small image distortions.<sup>39</sup> Max pooling takes the maximum value from a convolutional layer over a specified kernel and passes this response to the next layer. Batch normalization prevents overfitting and decreases training time by reducing internal covariate shift through normalization of the mean and variance statistics of the network units.<sup>40</sup> In the second half of the U-Net architecture structure, upsampling is used in place of pooling to connect the coarse outputs from the pooled convolutional layers back to the pixel segmentation.<sup>41</sup> The final fully connected layer uses a softmax activation function<sup>42</sup> to provide probability maps for each class (capillary, large vessel, background, image canvas), which can then be converted to binary maps using a global threshold determined by Otsu's method<sup>19</sup> to produce the final segmentation. For the purpose of computing capillary metrics, Otsu's method was applied only to the capillary class to segment those pixels that were capillaries from those that were not capillaries.

Our CNN contains important alterations from the base U-Net structure. First, we have expanded on

previous CNN designs that have classified pixels into one of only two categories (i.e., vessel and non-vessel) by developing a 4-pixel type classification system (i.e., capillary, large vessel, background, and image canvas classes). Second, the size of the convolutional filter kernel was changed from a fixed  $3 \times 3$  pixels to a varying size of  $25 \times 25$ ,  $15 \times 15$ , and  $10 \times 10$  pixels to accommodate for large vessels (defined to be  $>20$  pixels in diameter) and capillaries (that typically range from 7 to 14 pixels in diameter). In addition, a weighting function was implemented using TensorFlow,<sup>43</sup> as the proportion of background and canvas pixels in the training set was much greater than the number of pixels classified as capillaries or large vessels. The weighting function more heavily weights capillary and large vessel pixels in inverse proportion to their percent representation in training set patches. Using TensorFlow, a weighted categorical cross-entropy loss function was implemented for the network training and adjusted to incorporate multiple pixel classes.

### Convolutional Neural Network: Training and Testing

The network was trained for 250 epochs with a batch size of 64 on a NVIDIA Volta graphics processing unit at the University of Houston Research Computing Data Core (formerly the Core Facility for Advanced Computing and Data Science) using the aforementioned training set. The testing dataset was collected in a separate group of eyes and consisted of 1764 patches that were extracted from 14 images ( $768 \times 768$  pixels) acquired in one healthy human eye, two healthy NHP eyes, and one NHP eye with laser-induced experimental glaucoma. Test images were manually segmented by two raters (rater A and rater B). In addition, we generated a combined testing set to serve as a manual ground truth for evaluating CNN performance. The combined dataset was the union of the manual segmentations separately performed by rater A and rater B and was skeletonized and re-dilated using a uniform 5-pixel radius to better mimic true capillary diameters.

### Traditional Segmentation Algorithms

In addition to performing manual and CNN segmentations, test images of perfused vasculature were automatically segmented using two traditional techniques: an Otsu's approach and a Frangi approach.<sup>19,25</sup> In both approaches, a custom program (MATLAB; The MathWorks, Inc., Natick, MA) was developed to first apply a Gaussian filter ( $\sigma = 3$ ) to grayscale AOSLO images. In the Otsu's approach,

Otsu's method was subsequently applied to local regions of each Gaussian filtered image using a kernel size of  $49 \times 49$  pixels to determine an intensity threshold value and generate a binary image. The kernel size was automatically determined by the "imbinarize" function in MATLAB to be 1/16th of the image dimension ( $768 \times 768$  pixels). In the Frangi approach, a Frangi filter ( $\sigma = 1$ ) was applied to each Gaussian filtered perfusion image, with the output being converted to a binary image using a global Otsu's method. The binary images generated using each traditional approach were quantified using the same method as for the images that were segmented manually or automatically using the CNN.

### Performance Evaluation: Accuracy

Capillary segmentation performance was evaluated by computing the accuracy of the CNN (and traditional) segmentation techniques with respect to the combined testing set for the capillary class. The numbers of pixels classified as true-positive (TP), false-positive (FP), false-negative (FN), and true-negative (TN) were calculated using a pixel-wise comparison between the manual ground-truth segmentation and the segmentation from the CNN. Pixels marked as the capillary class by both the manual ground truth and the CNN were labeled TP, and those that were not marked by both the manual ground truth and the CNN were labeled TN. Pixels marked as the capillary class by the CNN, but not the manual ground truth, were labeled as FP, and those marked as the capillary class by the manual ground truth but not the CNN were labeled FN. The accuracy<sup>44</sup> was then calculated as

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

### Performance Evaluation: Dice and Modified Dice

Computing accuracy alone may not fully capture how well the CNN segmentation matches the manual ground-truth segmentation. For example, having a high percentage of background pixels in a testing set could inflate those pixels labeled TN and artificially yield a high accuracy. Sensitivity and the Dice coefficient of the segmentation could be more robust tools for evaluating a segmentation which contains a high proportion of background (or negative) class pixels, as is common in AOSLO perfusion images.<sup>45</sup> The sensitivity<sup>46</sup> of the segmentation, also called the true-positive ratio, is calculated using only the pixels



**Figure 4.** A modified Dice coefficient accounts for slight spatial differences between manual ground-truth and CNN segmentations of the same capillary. (a) Example in which manual (green) and CNN (red) segmentations of the same capillary have little overlap (yellow), yielding a very low Dice coefficient of 0.14. (b) To account for slight spatial differences in segmentations of the same capillary on measures of performance, the boundaries of the ground-truth segmentation were expanded outward by 5 pixels (green lines). (c) The boundaries of the CNN segmentation were expanded outward by 5 pixels (red lines). (d) Image showing the upper and lower boundaries of the intersection of the two expansions from (b) and (c). The pixels within these boundaries are considered to be overlapping pixels, or true positives. (e) The new overlapping region (yellow) was used to calculate the modified Dice coefficient (with a value of 0.93 for this image).

that are marked as positive, or capillary class, in the manual ground-truth segmentation image. Because sensitivity does not account for times when the CNN segmentation erroneously marks capillaries where the manual ground-truth segmentation does not (FPs), we chose to use the Dice coefficient as our performance metric. The Dice coefficient<sup>47,48</sup> uses the incorrectly marked pixels (FPs) as part of the evaluation and is computed as

$$\text{Dice coefficient} = \frac{2*TP}{2*TP + FP + FN} \quad (2)$$

The Dice coefficient has the potential to be artificially low if the segmentations being compared do not perfectly overlap when marking the same capillary (Fig. 4a). To account for slight spatial differences between segmentations of the same marked capillaries, we developed a modified Dice coefficient to include segmentations that were separated by less than the average capillary radius (5 pixels) as TPs. The boundaries of the automatic segmentation and the manual ground-truth segmentation were expanded outward by 5 pixels (Fig. 4b,c). Pixels were re-classified as TP if the expanded CNN segmentation overlapped the original ground-truth segmentation or if the expanded ground-truth segmentation overlapped the original CNN segmentation (Figs. 4d, 4e). The modified Dice coefficient was then computed with the reclassified pixels using Equation 2.

## Capillary Metrics

A custom MATLAB program was used to calculate capillary density and mean segment length for all segmentation techniques. Capillary density has been a commonly reported metric used to quantify perfused capillaries from OCTA images,<sup>49,50</sup> and is calculated from binary images as the ratio of pixels identified as a capillary class to the total number of image pixels. Mean segment length (MSL) has been used as a metric to assess the continuity of capillary segmentation algorithms, with increased mean segment lengths correlating to improved segmentation.<sup>51</sup> When computing MSL, the MATLAB function “bwmorph” was used to thin the binary image. The function then identified endpoints and branchpoints of the capillary segments and used these points to define the continuous pixels making up a segment. A length was computed for each segment in the image, and the MSL is the average of these lengths.

## Statistics

We assessed whether significant differences in capillary metrics existed between different raters or between different segmentation techniques using a one-way ANOVA followed by a Tukey–Kramer<sup>52</sup> post hoc test. Determinations of whether significant differences existed in the accuracy and performance of the different segmentation techniques were also performed using a one-way ANOVA followed by a Tukey–Kramer<sup>52</sup> post hoc test. Values of  $P < 0.05$  were considered

to be statistically significant (SigmaPlot 13.0; Systat Software, Inc., San Jose, CA).

## Results

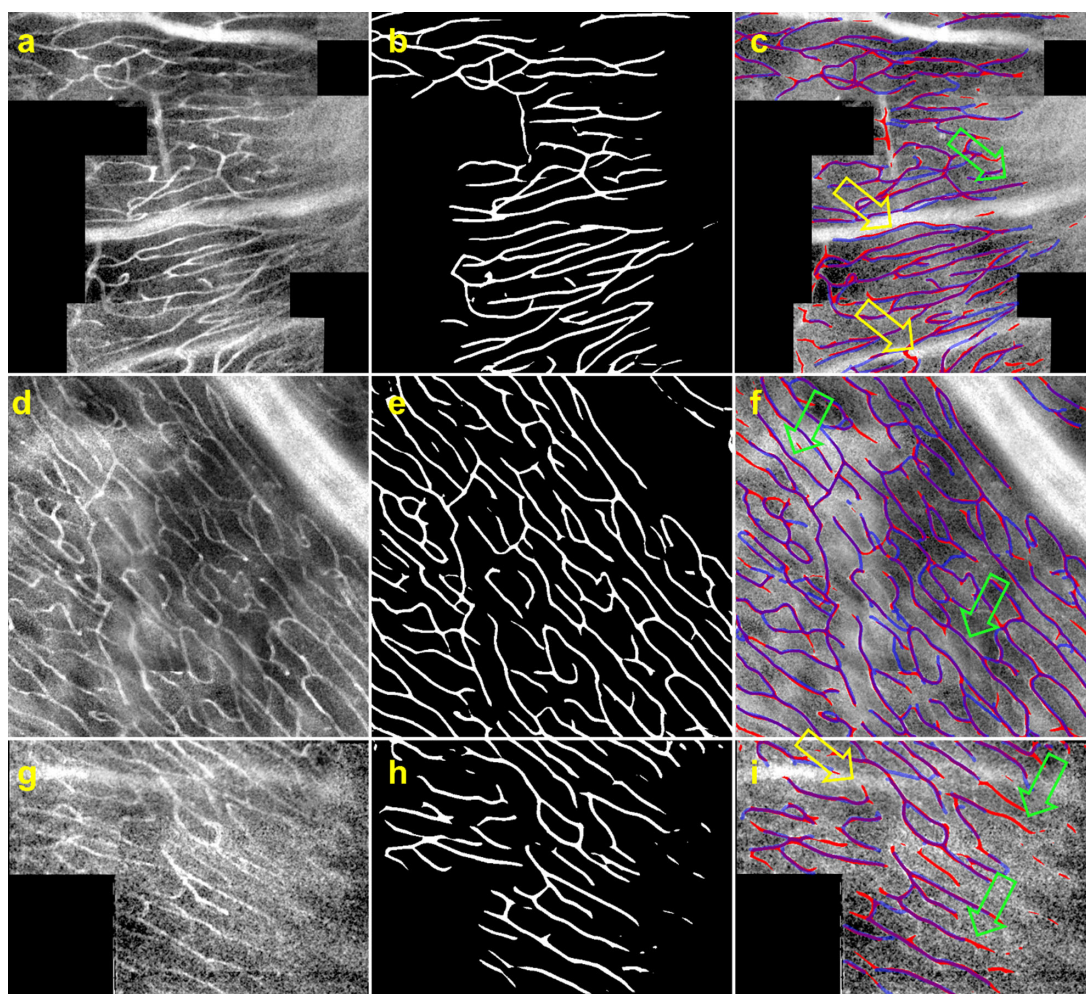
Table 2 shows the levels of agreement in capillary metrics quantified following manual segmentations of 14 testing images performed by rater A, rater B, and the union of raters A and B (i.e., ground truth). There were no significant differences in measurements of capillary density or MSL across raters ( $P > 0.05$ ). The Dice and modified Dice coefficients between the two manual raters were  $0.63 \pm 0.06$  and  $0.90 \pm 0.04$ , respectively.

Training the CNN for 250 epochs with 64 batches per epoch on the full training dataset took 3 hours

**Table 2.** Mean  $\pm$  SD Values of Capillary Density and MSL Computed Following Manual Segmentation of 14 Test Images

	Density (%)	MSL (pixels)
Rater A	$14.8 \pm 2.7$	$72 \pm 12$
Rater B	$14.3 \pm 3.0$	$65 \pm 10$
Union of raters A and B (ground truth)	$13.9 \pm 3.1$	$64 \pm 8$

for completion, and segmenting the 14 images in the testing set took 2 minutes and 30 seconds on an NVIDIA Volta GPU. Comparisons of manual and CNN segmentations of representative test images are shown in Figure 5. In general, the CNN can automatically segment capillaries throughout the entire image,



**Figure 5.** The CNN algorithm successfully segments perfused retinal capillaries in areas with high background signal and over larger vessels. Original grayscale image regions containing perfused capillaries and larger vasculature from the eyes of (a) a healthy human subject, (d) a healthy non-human primate, and (g) a non-human primate with laser-induced experimental glaucoma. Image regions were taken from larger AOSLO montages as described in Figure 3. (b, e, h) Binary CNN segmentations of perfused capillaries from the corresponding grayscale images in (a), (d), and (g). (c, f, i) Original grayscale images from (a), (d), and (g) showing CNN segmentations from (b), (e), and (h) in red and manual ground-truth segmentations in blue. The CNN is capable of segmenting capillaries in regions with high background signal (green arrows) and regions where capillaries traverse larger vessels (yellow arrows).



**Table 3.** Comparison of Performance and Capillary Metrics Generated Following Various Manual and Automated Segmentation Techniques

	Performance Metrics (Ground Truth = Union of Raters A and B)			Capillary Metrics	
	Accuracy	Dice Coefficient	Modified Dice Coefficient	Density (%)	MSL (pixels)
Ground truth	—	—	—	13.9 ± 3.1	64 ± 8
Otsu's approach	0.79 ± 0.04	0.42 ± 0.10	0.65 ± 0.11	34.7 ± 5.5 <sup>b</sup>	28 ± 4 <sup>c</sup>
Frangi approach	0.90 ± 0.01	0.50 ± 0.12	0.82 ± 0.10	14.1 ± 5.5	31 ± 8 <sup>c</sup>
CNN	0.94 ± 0.02 <sup>a</sup>	0.67 ± 0.07 <sup>a</sup>	0.90 ± 0.05 <sup>a</sup>	11.9 ± 3.7	62 ± 6

<sup>a</sup>Indicates a significant difference relative to other automated segmentation approaches ( $P < 0.05$ ).

<sup>b</sup>Indicates a significant difference relative to all other segmentation approaches ( $P < 0.001$ ).

<sup>c</sup>Indicates a significant difference relative to ground-truth segmentation (raters A and B) ( $P < 0.05$ ).

including regions with high background signal and locations where capillaries traverse larger vessels. The accuracy of the CNN in correctly classifying a pixel as being a capillary (relative to the ground-truth manual marking) was 0.94 across all testing images (Table 3). The accuracies obtained for the large vessel, background, and canvas classes were 0.96, 0.90, and 0.99, respectively.

To compare our newly developed CNN to traditional segmentation algorithms, test images were also segmented using an Otsu's approach and a Frangi approach. Figure 6 illustrates a representative grayscale perfusion image with the corresponding manual ground-truth, traditional, and CNN segmentation techniques. The result obtained following Otsu's approach (Fig. 6c) highlights the oversegmentation of background features and the inability for the algorithm to segment capillaries only. Capillaries segmented with this algorithm have uneven edges and non-uniform diameters. The Frangi approach (Fig. 6d) yielded capillaries with regular diameters and smooth edges, was able to separately classify capillaries, and did not segment background signal to the extent that was observed following Otsu's approach. However, it is common to see gaps in segmentation along the length of the capillary using this approach, which tends to produce smaller values of mean segment length. Qualitatively, the CNN segmentation (Fig. 6e) closely resembles the manual ground-truth segmentation. Capillaries tend to have uniform diameters and smooth edges and to be continuously segmented along their length with this method.

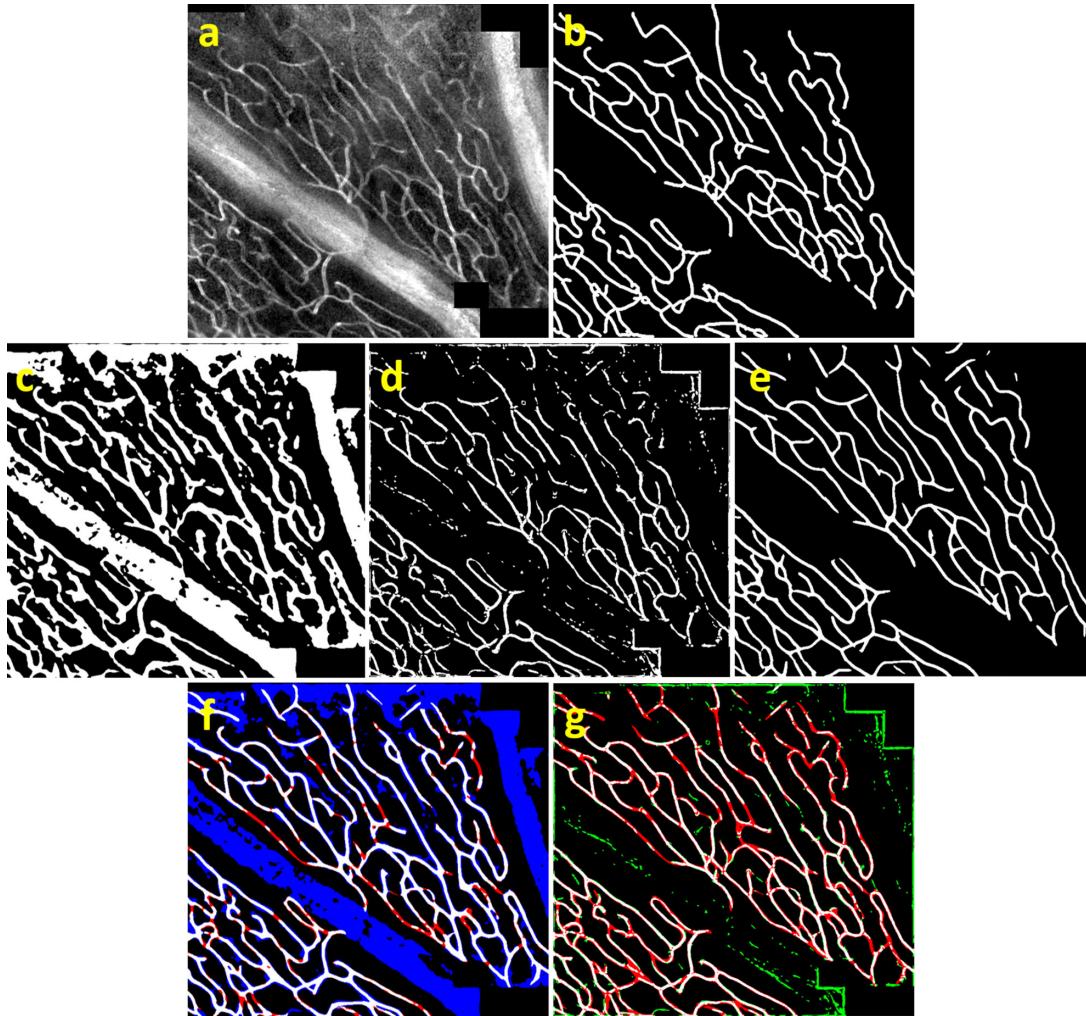
We compared performance and capillary metrics obtained after conducting manual segmentation, Otsu's approach, the Frangi approach, and CNN segmentations on the test images (Table 3). Significant differences were found between automated segmentation approaches (relative to manual ground-truth segmentations) for accuracy and the Dice and

modified Dice coefficients ( $P < 0.05$ ). Post hoc analysis determined there were significant differences in the values of accuracy, Dice coefficient, and modified Dice coefficient between the segmentations generated by the CNN and Otsu's approach and those generated by the CNN and Frangi approach ( $P < 0.05$ ).

Capillary metrics computed after segmenting test images using our newly developed CNN and traditional automated techniques were compared with those calculated using the manual ground truth. Capillary density was significantly higher when calculated using Otsu's approach compared to all other segmentation approaches ( $P < 0.001$ ). This result was expected, as Otsu's method does not distinguish between capillary and large vasculature. (As shown in Supplementary Table S1, Otsu's method still yielded significantly higher densities and lower mean segment lengths relative to manual segmentations when segmenting both capillaries and large vessels.) No significant differences in capillary density were found between manual ground-truth, Frangi, and CNN segmentations ( $P > 0.05$ ). MSL was significantly different between segmentation approaches ( $P < 0.001$ ), with post hoc analysis identifying significant differences in MSL between Otsu's approach and manual ground truth and between the Frangi approach and manual ground truth ( $P < 0.05$ ). No significant difference was found between manual ground truth and the CNN segmentation ( $P > 0.05$ ).

## Discussion

The main purpose of this study was to develop a method capable of automatically and accurately segmenting perfused retinal capillaries in AOSLO images. We designed a CNN that incorporated the novel use of four classes in the training set to enhance



**Figure 6.** CNN segmentation outperforms traditional segmentation techniques. (a) Original grayscale image containing perfused capillaries and larger vasculature. (b) Manual ground-truth image generated by the manual segmentation of perfused capillaries only. (c) Resultant image generated after applying Otsu's approach to the original grayscale image in (a). In addition to segmenting capillaries, this approach also segments major vasculature and background. (d) Application of the Frangi approach to the original grayscale image in (a) resulted in a binary image that excluded major vasculature but tended to leave small gaps along the length of individual capillaries, thereby reducing the metric of mean segment length. (e) CNN segmentation tends to exclude major vasculature and background signal while also maintaining continuity along the segmented capillaries. (f) Color-coded overlay of the CNN segmentation (*red*) from (e) with the segmentation following Otsu's approach (*blue*) from (c), where areas of common segmentation are shown in *white*. (g) Color-coded overlay of the CNN segmentation (*red*) from (e) with the segmentation following the Frangi approach (*green*) from (d), where areas of common segmentation are shown in *white*.

the segmentation and quantification of perfused radial peripapillary capillaries relative to traditional segmentation techniques, particularly in areas with high background signal and large perfused vasculature. There were no statistical differences between metrics of capillary density and MSL calculated from segmentations performed by two independent manual raters or the union of the two raters. Capillary metrics quantified from the CNN segmentation were not statistically different from those quantified following manual ground-truth segmentation. Overall, the CNN dramatically decreased processing time while providing a more

objective approach to the segmentation of perfused retinal capillaries.

Our newly developed CNN outperforms two traditional segmentation techniques (Otsu's and Frangi approaches) and closely mimics manual segmentations. Although commonly used to segment perfused capillaries in OCTA images,<sup>20,21</sup> the Otsu's approach employed in this study does not distinguish smaller diameter capillaries from larger diameter, major vasculature. In addition, Otsu's approach is a pixel intensity-based method for segmentation that can result in the inclusion of large regions of background signal, all of

which can result in artificially high values for capillary density. The Frangi approach for automated segmentation yielded capillary densities that were similar to those in manual ground-truth segmentations, largely because Frangi filters were originally developed to segment bright tube-shaped objects and can be tuned to include specific spatial frequencies (such as excluding lower spatial frequencies that are more characteristic of larger vessels).<sup>25</sup> Despite their similar capillary density and relatively high performance values, the Frangi approach also resulted in segmentations that included gaps along the length of individual capillaries, yielding low values of mean segment length relative to manual ground-truth segmentations. The newly developed CNN more accurately separated capillaries from larger vasculature and was more continuous in its segmentation of individual capillaries. Consequently, the CNN yielded high performance metrics and generated values of capillary density and MSL that were not significantly different from the manual ground-truth segmented values.

The separation of the training dataset into four different classes (capillaries, large vessel, image background, image canvas) is an exciting improvement on existing CNNs that have been used for vascular segmentation. Previous CNNs developed for and applied to fundus images have not differentiated between smaller or larger vessels,<sup>26–28</sup> likely due to the fact that fundus images do not possess the resolution necessary to visualize capillaries. A more recent CNN designed to segment capillaries in OCTA images was trained on images centered on the fovea (surrounding the foveal avascular zone),<sup>29</sup> where smaller vasculature is more prominent. In contrast, the peripapillary retina surrounding the optic nerve head contains significant major vasculature that ideally needs to be differentiated from smaller vasculature to better understand capillary perfusion. Hence, we trained our CNN on images acquired near the optic nerve head where the superficial retina contains a large range of vessel diameters (from major arterioles and venules to the smallest of radial peripapillary capillaries) using classes that successfully separated perfused capillaries from larger vasculature in our test images.

A variety of kernel sizes were explored for the pooling (and corresponding upsampling) layers and convolutional layers when developing the CNN. Our use of variable convolutional layer kernel sizes was motivated by a previously reported CNN that used a  $4 \times 4$  kernel size for the pooling layers but variable convolutional layer kernel sizes to enhance the ability of the algorithm to segment vasculature with different diameters in OCTA images.<sup>25</sup> During development, it

was found that the values of capillary metrics output by our CNN depended on the type of convolutional layer kernel size (variable versus fixed) and the kernel size of the pooling layers. As shown in Supplementary Figures S1d and S1e, the use of a  $4 \times 4$  kernel size for the pooling layer (relative to a  $2 \times 2$  kernel size) provided a segmentation with increased smoothing, uniformity in vessel width, and connectivity of capillaries, regardless of the type of convolutional layer kernel size. A  $2 \times 2$  kernel size for the pooling layer yielded higher resolution segmentations that were also more fragmented and dependent on the type of convolutional layer kernel size (Supplementary Figures S1b, S1c). These qualitative observations are supported quantitatively in Supplementary Table S2. Capillary metrics were similar to values obtained from manual ground-truth segmentations when using a  $4 \times 4$  kernel size for the pooling layers and a variable or a fixed  $3 \times 3$  kernel size for the convolutional layers (see Supplementary Figures S1d and S1e, respectively). However, when using a  $2 \times 2$  kernel size for the pooling layers, a fixed,  $3 \times 3$  convolutional layer kernel size (Supplementary Figure S1b) yielded values of MSL that were significantly different from ground truth, whereas those obtained using the variable convolutional layer kernel size (Supplementary Figure S1c) were not. The publicly available code on GitHub is configured for variable convolutional layer kernel sizes, as these maintained the mean segment length regardless of the pooling kernel sizes we explored (Supplementary Table S2).

The newly developed CNN objectively segments perfusion images and reduces potential variability in capillary metrics that could result from differences in manual markings performed by different raters. We found no statistical difference in the values of mean capillary density or mean segment length that were calculated between the manual segmentations performed by two independent raters on the testing images. Therefore, the markings performed by a single rater (rater A) were used as ground truth for the training dataset. Rather than choosing the markings made by rater A or rater B as the ground truth for the testing dataset, we adopted a more restrictive criterion—namely, the union of the manual segmentations performed by the two raters. In addition, the performance of the CNN segmentation is directly dependent on the ground-truth segmentation method used in the training dataset. An alternative approach to the one used in this work to generate ground-truth segmentations for training could be to have multiple raters segment images and use only the common segmentation as the ground-truth training set. However, such an approach could be time inefficient

and may result in capillary segmentations that are more fragmented, potentially hindering subsequent analyses seeking to quantify mean segment length on test images of perfused capillaries.

When combined with high-performance computing resources, the CNN detailed in this work decreased the time required to segment one complete montage (as illustrated in Fig. 3a), from 6 hours when performed manually by an expert rater to approximately 2 to 3 minutes. In general, high-performance computing resources can be leveraged to increase the efficiency and speed with which the network can be trained; however, less computational power is required when the trained network weights have been determined, allowing researchers to efficiently use the trained CNN on most computing systems.

Additional enhancements to the described CNN could be considered to increase its general performance and utility across a wide range of conditions, particularly for applications in which accurate segmentation and quantification of perfused capillaries are important for drawing conclusions about disease pathology. For example, even though our CNN was trained on a dataset that included images from healthy eyes and eyes with varying severities of laser-induced experimental glaucoma, the training dataset could be refined by adding images taken in eyes with other diseases (such as diabetic retinopathy). A potential benefit to including more diseased eyes is that image quality tends to be worse in pathological cases than in healthy eyes. Therefore, the CNN segmentation may benefit from including a larger number of images that are challenging to segment in the training set. Also, our training dataset consisted of images acquired near the optic nerve head, where both capillaries and major vessels are abundant. AOSLO perfusion images acquired near the fovea, where capillaries are very abundant, could increase the number of capillaries in the training dataset (relative to the amount of major vasculature) and potentially improve segmentation performance. In addition, future work could be conducted to better understand the performance of this algorithm when applied to capillary perfusion images acquired from different AOSLO systems and laboratories.

In conclusion, the architecture for the CNN presented in this work can be used to accurately, objectively, and quickly segment perfused capillaries from high-resolution images obtained using an AOSLO. The close agreement between the CNN and manual segmentation enables the robust quantification of perfused capillary metrics. This automated segmentation technique could be applied to evaluate changes in capillary metrics over time in retinal and optic nerve head diseases.

## Acknowledgments

The authors thank David Cunefare and Sina Farsiou for their helpful discussions in developing the CNN. The authors acknowledge the use of the Maxwell/Opuntia/Sabine Clusters and the advanced support from the Core Facility for Advanced Computing and Data Science at the University of Houston.

Supported by the donors of the National Glaucoma Research Grant (G2018061), a program of the Bright-Focus Foundation; by a National Institutes of Health Grant (P30 EY007551); and by the University of Houston College of Optometry. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Disclosure: **G. Musial**, None; **H.M. Queener**, None; **S. Adhikari**, None; **H. Mirhajianmoghadam**, None; **A.W. Schill**, None; **N.B. Patel**, None; **J. Porter**, None

## References

1. Geyman LS, Garg RA, Suwan Y, et al. Peripapillary perfused capillary density in primary open-angle glaucoma across disease stage: an optical coherence tomography angiography study. *Br J Ophthalmol*. 2017;101:1261–1268.
2. Sripsema NK, Garcia PM, Bavier RD, et al. Optical coherence tomography angiography analysis of perfused peripapillary capillaries in primary open-angle glaucoma and normal-tension glaucoma. *Invest Ophthalmol Vis Sci*. 2016;57:OCT611–OCT620.
3. Kim SB, Lee EJ, Han JC, Kee C. Comparison of peripapillary vessel density between preperimetric and perimetric glaucoma evaluated by OCT-angiography. *PLoS One*. 2017;12:e0184297.
4. Zhang YS, Zhou N, Knoll BM, et al. Parafoveal vessel loss and correlation between peripapillary vessel density and cognitive performance in amnesic mild cognitive impairment and early Alzheimer's disease on optical coherence tomography angiography. *PLoS One*. 2019;14:e0214685.
5. Yoon SP, Grewal DS, Thompson AC, et al. Retinal microvascular and neurodegenerative changes in Alzheimer's disease and mild cognitive impairment compared with control participants. *Ophthalmol Retin*. 2019;3:489–499.
6. Ishibazawa A, Nagaoka T, Takahashi A, et al. Optical coherence tomography angiography in

- diabetic retinopathy: a prospective pilot study. *Am J Ophthalmol*. 2015;160:35–44.e1.
7. Nesper PL, Roberts PK, Onishi AC, et al. Quantifying microvascular abnormalities with increasing severity of diabetic retinopathy using optical coherence tomography angiography. *Invest Ophthalmol Vis Sci*. 2017;58:BIO307–BIO315.
  8. Jia Y, Hornegger J, Tan O, et al. Split-spectrum amplitude-decorrelation angiography with optical coherence tomography. *Opt Express*. 2012;20:4710–4725.
  9. Chui TYP, Mo S, Krawitz B, et al. Human retinal microvascular imaging using adaptive optics scanning light ophthalmoscopy. *Int J Retin Vitre*. 2016;2:11.
  10. Scoles D, Gray DC, Hunter JJ, et al. In-vivo imaging of retinal nerve fiber layer vasculature: imaging histology comparison. *BMC Ophthalmol*. 2009;9:1–9.
  11. Chui TYP, VanNasdale DA, Burns SA. The use of forward scatter to improve retinal vascular imaging with an adaptive optics scanning laser ophthalmoscope. *Biomed Opt Express*. 2012;3:2537–2549.
  12. Tekin K, Inanc M, Kurnaz E, et al. Quantitative evaluation of early retinal changes in children with type 1 diabetes mellitus without retinopathy. *Clin Exp Optom*. 2018;101:680–685.
  13. Tang PH, Jauregui R, Tsang SH, Bassuk AG, Mahajan VB. Optical coherence tomography angiography of RPGR-associated retinitis pigmentosa suggests foveal avascular zone is a biomarker for vision loss. *Ophthalmic Surg Lasers Imaging Retin*. 2019;50:e44–e48.
  14. Hui F, Nguyen CTO, He Z, et al. Retinal and cortical blood flow dynamics following systemic blood-neural barrier disruption. *Front Neurosci*. 2017;11:568.
  15. Fercher AF, Drexler W, Hitzenberger CK, Lasser T. Optical coherence tomography—principles and applications. *Reports Prog Phys*. 2003;66:239–303.
  16. Roorda A, Romero-Borja F, Donnelly WJ, III, Queener H, Hebert TJ, Campbell MCW. Adaptive optics scanning laser ophthalmoscopy. *Opt Express*. 2002;10:405–412.
  17. Burns SA, Elsner AE, Sapoznik KA, Warner RL, Gast TJ. Adaptive optics imaging of the human retina. *Prog Retin Eye Res*. 2019;68:1–30.
  18. Optovue. Angiovue. Available at: <https://www.optovue.com/international/products/angiovue>. Accessed July 3, 2020.
  19. Otsu N. A threshold selection method from gray-level histograms. *IEEE Trans Syst Man Cybern*. 1979;20:62–66.
  20. Rabiolo A, Gelormini F, Sacconi R, et al. Comparison of methods to quantify macular and peripapillary vessel density in optical coherence tomography angiography. *PLoS One*. 2018;13:e0205773.
  21. Yao C, Chen HJ. Automated retinal blood vessels segmentation based on simplified PCNN and fast 2D-Otsu algorithm. *J Cent South Univ Technol*. 2009;16:640–646.
  22. Mochi T, Anegondi N, Girish M, Jayadev C, Roy AS. Quantitative comparison between optical coherence tomography angiography and fundus fluorescein angiography images: effect of vessel enhancement. *Ophthalmic Surg Lasers Imaging Retin*. 2018;49:e175–e181.
  23. Tan B, Wong A, Bizheva K. Enhancement of morphological and vascular features in OCT images using a modified Bayesian residual transform. *Biomed Opt Express*. 2018;9:2394–2406.
  24. Jothi A, Jayaram S. Blood vessel detection in fundus images using Frangi filter technique. In: Pani-grahi B, Trivedi M, Mishra K, Tiwari S, Singh P, eds. *Advances in Intelligent Systems and Computing*. Vol. 670. Singapore: Springer; 2019:49–57.
  25. Frangi AF, Niessen WJ, Vincken KL, Viergever MA. Multiscale vessel enhancement filtering. In: Wells WM, Colchester A, Delp S, eds. *Medical Image Computing and Computer-Assisted Intervention*. Vol. 1496. Berlin: Springer; 1998:130–137.
  26. Xiancheng W, Wei L, Bingyi M, et al. Retina blood vessel segmentation using a U-Net based convolutional neural network. *Procedia Comput Sci*. 2018;00:1–11.
  27. Biswas R, Vasani A, Roy SS. Dilated deep neural network for segmentation of retinal blood vessels in fundus images. *Iran J Sci Technol Trans Electr Eng*. 2020;44:505–518.
  28. Samuel PM, Veeramalai T. Multilevel and multiscale deep neural network for retinal blood vessel segmentation. *Symmetry (Basel)*. 2019;11:946.
  29. Prentašić P, Navajas E, Lončarić S, et al. Segmentation of the foveal microvasculature using deep learning networks. *J Biomed Opt*. 2016;21:075008.
  30. Frishman LJ, Shen FF, Du L, et al. The scotopic electroretinogram of macaque after retinal ganglion cell loss from experimental glaucoma. *Invest Ophthalmol Vis Sci*. 1996;37:125–141.
  31. Ivers KM, Li C, Patel N, et al. Reproducibility of measuring lamina cribrosa pore geometry in human and nonhuman primates with in vivo adaptive optics imaging. *Invest Ophthalmol Vis Sci*. 2011;52:5473–5480.
  32. Delori FC, Webb RH, Sliney DH. Maximum permissible exposures for ocular safety (ANSI 2000),

- with emphasis on ophthalmic devices. *J Opt Soc Am A*. 2007;24:1250–1265.
33. Scoles D, Sulai YN, Langlo CS, et al. In vivo imaging of human cone photoreceptor inner segments. *Invest Ophthalmol Vis Sci*. 2014;55:4244–4251.
  34. Sulai YN, Scoles D, Harvey Z, Dubra A. Visualization of retinal vascular structure and perfusion with a nonconfocal adaptive optics scanning light ophthalmoscope. *J Opt Soc Am A Opt Image Sci Vis*. 2014;31:569–79.
  35. Dubra A, Harvey Z. Registration of 2D images from fast scanning ophthalmic instruments. In: Fischer B, Dawant BM, Lorenz C, eds. *Biomedical Image Registration*. Berlin: Springer; 2010:60–71.
  36. Ivers KM, Sredar N, Patel NB, et al. In vivo changes in lamina cribrosa microarchitecture and optic nerve head structure in early experimental glaucoma. *PLoS One*. 2015;10:e0134223.
  37. Ronneberger O, Fischer P, Brox T. U-Net: convolutional networks for biomedical image segmentation. Available at: <https://arxiv.org/pdf/1505.04597.pdf>. Accessed July 3, 2020.
  38. Schmidhuber J. Deep learning in neural networks: an overview. *Neural Networks*. 2015;61:85–117.
  39. Jarrett K, Kavukcuoglu K, Ranzato M, LeCun Y. What is the best multi-stage architecture for object recognition? In: *2009 IEEE 12th International Conference on Computer Vision*. Los Alamitos, CA: Institute of Electrical and Electronics Engineers; 2009:2146–2153.
  40. Ioffe S, Szegedy C. Batch normalization: accelerating deep network training by reducing internal covariate shift. *Proc Mach Learn Res*. 2015;37:448–456.
  41. Shelhamer E, Long J, Darrell T. Fully convolutional networks for semantic segmentation. *IEEE Trans Pattern Anal Mach Intell*. 2017;39:640–651.
  42. Bishop CM. *Pattern Recognition and Machine Learning*. New York: Springer, 2006.
  43. TensorFlow. TensorFlow: an end-to-end open source machine learning platform. Available at: <http://www.tensorflow.org>. Accessed July 3, 2020.
  44. Rand WM. Objective criteria for the evaluation of clustering methods. *J Am Stat Assoc*. 1971;66:846–850.
  45. Taha AA, Hanbury A. Metrics for evaluating 3D medical image segmentation: analysis, selection, and tool. *BMC Med Imaging*. 2015;15:29.
  46. Altman DG, Bland JM. Statistics Notes: Diagnostic tests 1: sensitivity and specificity. *BMJ*. 1994;308:1552.
  47. Sørensen TJ. A method of establishing groups of equal amplitudes in plant sociology based on similarity of species content and its application to analyses of the vegetation on Danish commons. *K Dansk Vidensk Selsk Skr*. 1948;5:1–34.
  48. Dice LR. Measures of the amount of ecologic association between species. *Ecology*. 1945;26:297–302.
  49. Manalastas PIC, Zangwill LM, Saunders LJ, et al. Reproducibility of optical coherence tomography angiography macular and optic nerve head vascular density in glaucoma and healthy eyes. *J Glaucoma*. 2017;26:851–859.
  50. Al-Sheikh M, Tepelus TC, Nazikyan T, Sadda SVR. Repeatability of automated vessel density measurements using optical coherence tomography angiography. *Br J Ophthalmol*. 2017;101:449–452.
  51. Mo S, Phillips E, Krawitz BD, et al. Visualization of radial peripapillary capillaries using optical coherence tomography angiography: The effect of image averaging. *PLoS One*. 2017;12:1–17.
  52. Dubitzky W, Wolkenhauer O, Cho K-H, Yokota H, eds. *Encyclopedia of Systems Biology*. New York: Springer; 2013:2304.