Segmentation of the foveal microvasculature using deep learning networks

Pavle Prentašić
Morgan Heisler
Zaid Mammo
Sieun Lee
Andrew Merkur
Eduardo Navajas
Mirza Faisal Beg
Marinko Šarunić
Sven Lončarić

SPIE.

Segmentation of the foveal microvasculature using deep learning networks

Pavle Prentašić,[a,*,†] Morgan Heisler,[b,†] Zaid Mammo,[c] Sieun Lee,[b] Andrew Merkur,[c] Eduardo Navajas,[c] Mirza Faisal Beg,[b] Marinko Šarunić,[b] and Sven Lončarić[a]
[a]University of Zagreb, Faculty of Electrical Engineering and Computing, Unska ul. 3, Zagreb 10000, Croatia
[b]Simon Fraser University, Department of Engineering Science, 8888 University Drive, Burnaby, British Columbia V5A1S6, Canada
[c]University of British Columbia, Department of Ophthalmology and Visual Science, Eye Care Center, 2550 Willow Street, Vancouver, British Columbia V5Z 3N9, Canada

Abstract. Accurate segmentation of the retinal microvasculature is a critical step in the quantitative analysis of the retinal circulation, which can be an important marker in evaluating the severity of retinal diseases. As manual segmentation remains the gold standard for segmentation of optical coherence tomography angiography (OCT-A) images, we present a method for automating the segmentation of OCT-A images using deep neural networks (DNNs). Eighty OCT-A images of the foveal region in 12 eyes from 6 healthy volunteers were acquired using a prototype OCT-A system and subsequently manually segmented. The automated segmentation of the blood vessels in the OCT-A images was then performed by classifying each pixel into vessel or nonvessel class using deep convolutional neural networks. When the automated results were compared against the manual segmentation results, a maximum mean accuracy of 0.83 was obtained. When the automated results were compared with inter and intrarater accuracies, the automated results were shown to be comparable to the human raters suggesting that segmentation using DNNs is comparable to a second manual rater. As manually segmenting the retinal microvasculature is a tedious task, having a reliable automated output such as automated segmentation by DNNs, is an important step in creating an automated output. © 2016 Society of Photo-Optical Instrumentation Engineers (SPIE) [DOI: 10.1117/1.JBO.21.7.075008]

1 Introduction

The human retinal circulation is composed of complex capillary networks that are responsible for satisfying the high metabolic requirements of the multiple neuronal populations within the retina.[1] Retinal vascular diseases, such as diabetic retinopathy and vascular occlusions, contribute significantly to the burden of visual impairment worldwide.[2] Fluorescein angiography (FA) has been considered the gold standard in the evaluation and diagnosis of retinal vascular diseases. Despite its widespread use, this technique is limited by the background choroidal flush from resolving the fine structural details of the multiple layers of retinal capillaries.[3] In addition, FA requires the administration of intravenous contrast dye, which carries a small risk of significant adverse events.[4] Optical coherence tomography angiography (OCT-A) is a new imaging technology that allows noninvasive, dye-free visualization of the retinal circulation.[5] We have implemented a speckle-variance technique for OCT-A as a noninvasive imaging modality that uses the change in the speckle pattern due to red blood cell movement in sequentially acquired OCT images; the corresponding intensity variance in the structural images is used to identify the retinal microvasculature. Using OCT-A, we have been able to show comparable quantitative and qualitative characteristics of the peripapillary,[6–8] foveal,[9] and perifoveal[10] images to cadaveric histological representation.

Macular capillary density is correlated to retinal thickness and visual functioning in patients with diabetic retinopathy.[11] Hence, accurate serial quantification of the retinal microcirculation is a useful marker in evaluating the severity of retinal vascular diseases. Following OCT-A image acquisition, accurate segmentation of the retinal microvasculature is a critical step in the quantitative analysis of the retinal circulation. Retinal vessel segmentation has been demonstrated in multiple medical imaging modalities[12,13] and is well documented in the literature. However, as the vasculature detail and appearance are different for each modality, optimal segmentation approaches may differ between modalities. For vessel segmentation in OCT-A images, only a limited body of work has been conducted.

Automated approaches of segmenting retinal vessels using OCT-A data are becoming more prevalent, yet manual segmentation remains the gold standard. Manual segmentation of the retinal blood vessels in OCT angiography images is a time-consuming and tedious task, which requires training. Reliable automated segmentation of these vessels is paramount for automated microvasculature quantification. The simplest automated approach, adaptive thresholding, has been used[14] but is limited in its sensitivity to the selection of a suitable threshold as well as its insensitivity to the shape and morphology of the microvasculature. One group has skeletonized the OCT-A images of retinal vessels in order to obtain retinal vasculature perfusion density maps[15] but this approach is still insensitive to the various widths of the vessels. Lastly, another group implemented automated blood vessel segmentation using a hybrid Hessian/intensity-based method while imaging wound healing in a

*Address all correspondence to: Pavle Prentašić, E-mail: pavle.prentasic@fer.hr

†Pavle Prentašić and Morgan Heisler are the cofirst authors. They have contributed equally to this work.

mouse ear (pinna) with OCT-A.[16] Although an accuracy of 0.94 was obtained when comparing the automated result to manual segmentations of human retinal fundus images, validation of the technique in human OCT-A retinal images still needs to be done.

This paper presents a new method for automated segmentation of blood vessels in retinal OCT-A images using deep neural networks (DNN). DNNs have shown promising results in solving a variety of problems, such as object recognition in images,[17,18] speech recognition,[19] semantic segmentation of images,[20,21] handwritten character classification recognition,[22] and text analysis.[23]

The main contribution of this paper is to demonstrate the high effectiveness of the deep learning approach to the segmentation of blood vessels in OCT-A images. The automated segmentation results on the images acquired from a clinical prototype OCT-A system were compared with the manual segmentations from two separate trained raters and discussed.

2 Methods

2.1 Ethics Statement

All subject recruitment and imaging took place at the Eye Care Centre of Vancouver General Hospital. 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.

2.2 Speckle Variance Optical Coherence Tomography Imaging

Speckle variance OCT images of the foveal region in 12 eyes from 6 healthy volunteers aged 36.8 ± 7.1 years were acquired using a graphics processing unit-accelerated OCT-A clinical prototype.[24] In total, 80 images were acquired. Briefly, the OCT system uses a 1060-nm swept source (Axsun Inc.) with 100-kHz A-scan rate and a full-width half-maximum bandwidth of 61.5 nm, which corresponds to a coherence length of $\sim 6 \ \mu m$ in tissue. For the speckle variance calculation, three repeat acquisitions were obtained at each B-scan location. The scan area was sampled in a $300 \times 300 (\times 3)$ grid with a $\sim 1 \times 1$ mm field of view in 3.15 s. Images were acquired either directly superiorly, nasally, inferiorly, or temporally from the foveal avascular zone. Processing of the OCT intensity image data and en face visualization of the retinal microvasculature was performed in real time using our open source code.[25,26]

2.3 Manual Segmentation

For comparison, two raters segmented OCT-A images using a Wacom Intuos 4 tablet and GNU image manipulation program. For the cross-validation and training, Rater A segmented all 80 OCT-A images. For the repeatability analysis, 10 images were used and segmented by both rater A and rater B. Rater A segmented each image twice for intrarater agreement, while Rater B segmented each image once for interrater agreement.

2.4 Deep Neural Network Architecture

The automated segmentation of the blood vessels in the OCT-A images was performed by classifying each pixel into either the vessel or the nonvessel class using deep convolutional neural networks. Convolutional and max pooling layers are used as hierarchical feature extractors, which map raw pixel intensities into a feature vector, which is then classified using fully connected layers.

The convolutional layers in our algorithm are made of a sequence of square filters, which perform a two-dimensional convolution with the input image. To calculate the output of each map, convolutional responses are summed and passed through a nonlinear activation function. The nonlinear activation function used in this paper is a rectifying linear unit.

Max pooling layers generate their output by taking the maximum activation over nonoverlapping square regions. These layers do not have adjustable parameters and their size is fixed. By taking the maximum value of the activation function, the most prominent features are selected from the input image.

After six stages of varied convolutional and max pooling layers, a dropout layer was inserted, which can prevent network over-fitting and provide a way of combining an exponentially increasing number of different neural networks in an efficient manner.[27] Then, two fully connected layers are used to classify the feature vector generated by the previous layers. The final fully connected layer contains two neurons where one neuron represents the vessel and other the nonvessel class. The network architecture used in this paper is very similar to the network architecture first used in Ref. 20. An overview is presented in Table 1 and graphically in Fig. 1.

2.5 Network Training Methods

To train our network, original OCT-A images and the corresponding manual segmentations were used as inputs. Each training example consists of a 61×61 pixel square window around the training pixel. Missing pixels in windows at the image border were set to zero. To have a balanced training set, an equal number of vessel and nonvessel pixels were extracted from each image. If the number of vessel pixels was larger than the number of nonvessel pixels in an image, then all nonvessel pixels were selected for the training set and an equal number of vessel pixels were randomly selected from the pool of vessel pixels. Similarly, if the number of nonvessel pixels was larger than

Table 1 Network layers architecture.

Layer	Type	Maps and size	Kernel size
0	Input	1 map of 61×61 neurons	
1	Convolutional	32 maps of 56×56 neurons	6×6
2	Max pooling	32 maps of 28×28 neurons	2×2
3	Convolutional	32 maps of 24×24 neurons	5×5
4	Max pooling	32 maps of 12×12 neurons	2×2
5	Convolutional	32 maps of 9×9 neurons	4×4
6	Max pooling	32 maps of 5×5 neurons	2×2
7	Dropout		
8	Fully connected	150 neurons	
9	Fully connected	2 neurons	

the number of vessel pixels in an image, then all vessel pixels were selected for the training set and an equal number of non-vessel pixels were randomly selected from the pool of nonvessel pixels.

2.6 Network Segmentation Methods

The trained network was then used to segment the original OCT-A images. First, a square window of the same size used for the training purposes was extracted around each pixel of the test images. A forward pass using all test image pixels was performed using the trained network, and each pixel was assigned a grayscale value, with higher values representing higher confidence of the pixel being a vessel pixel. These pixel values were aggregated into the output grayscale images, and median filtering with a small 3×3 window was performed in order to decrease the noise level in the image.

2.7 Cross-Validation Methods

Three-fold cross-validation on all images manually segmented by Rater A was performed. All 80 original images were randomly divided into three sets. Images from two of the sets were used to train the network, and images from the remaining set were used to test the network. This procedure was repeated three times with a different test set each time. Each set of the cross-validation was evaluated on a separate computer in order to decrease the total training and testing time. Each computer had a recent generation NVIDIA graphics card, which decreased computation time. The Caffe deep learning toolkit[28] was used to efficiently use the processing power of the graphics card for computation of convolutional neural network parameters. Using parallel processing, all three sets were used to train the proposed neural network in approximately 30 h. Segmentation of a single image using the trained network took ~2 min.

3 Results

3.1 Performance Evaluation

The segmentation performance was evaluated by pixel-wise comparison of the manually segmented images and the thresholded binary output of the neural network using varying thresholds. The number of true positives (TP), false positives (FP),

false negatives (FN), and true negatives (TN) were calculated using pixel-wise comparison between a reference manual segmentation and a target, which was either another manual segmentation, or the output of our automated method. In our context, a pixel is considered as TP if it is marked as a blood vessel in both the reference manual segmentation and in the target. A pixel is considered as FN if it is marked as blood vessel in the manual segmentation but missed by the target. A pixel is considered as FP if it is marked as vessel by our method but it is not marked as blood vessel in the target. A pixel is considered as TN if it is not marked as blood vessel in both the manual segmentation and in the target. Using the TP, FP, FN, and TN numbers we can calculate the accuracy: $(\mathrm{TP} + \mathrm{TN}) / (\mathrm{TP} + \mathrm{TN} + \mathrm{FP} + \mathrm{FN})$, sensitivity: $\mathrm{TP}/(\mathrm{TP} + \mathrm{FN})$, specificity: $\mathrm{TN}/(\mathrm{TN} + \mathrm{FP})$ and positive predictive value (PPV): $\mathrm{TP}/(\mathrm{TP} + \mathrm{FP})$ of the segmentation.

Using the PPV and sensitivity we can calculate the F_1 measure using Eq. (1).

$$F_1 = \frac{2 \cdot \text{Sensitivity} \cdot \text{PPV}}{\text{Sensitivity} + \text{PPV}}. \tag{1}$$

All of these measures can be calculated on individual images but can also be calculated for the whole dataset. In Fig. 2, the dotted blue line shows the accuracy for all pixels in the dataset against the threshold value used to binarize the output of the network. The accuracy of blood vessel detection increases from the threshold value at 0, peaks at 0.83 with threshold value of 0.78, and then begins to decline. It is important to note that similar results are obtained in a wide range of thresholds, which indicates that the performance is not sensitive to the threshold chosen.

In Fig. 3, the accuracy for each image was calculated and averaged over all images. One standard deviation below the mean values is marked with a green dotted line and one standard deviation above the mean values is marked with a blue dotted line. Qualitatively, the deviation of accuracies is reasonably small for different thresholds, with the maximum mean accuracy of 0.83 ± 0.02 at the threshold value of 0.76, signifying that the performance of the method is consistent over the whole dataset. The accuracy of the deeper capillary network [inner nuclear layer (INL) to outer plexiform layer] is 0.8247 while the accuracy of the superficial capillary networks (inner limiting membrane to INL) is 0.8389; the lower accuracy of the deeper layers
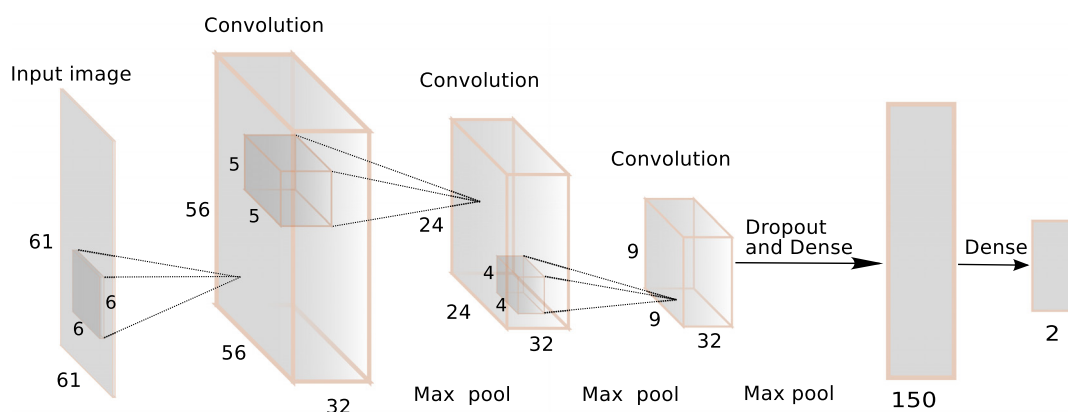

Fig. 1 Graphical representation of the network structure. Input image is a 61×61 patch cut at each image point. Two output neurons represent the probabilities of blood vessels and background at the central pixel in the input image.
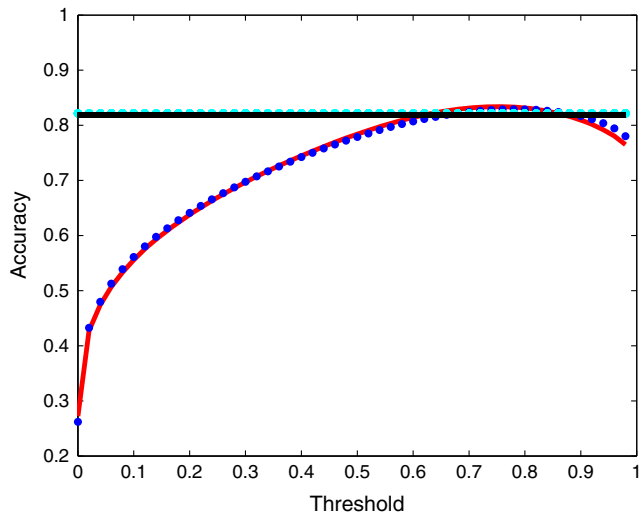
Fig. 2 Accuracy of the segmentation using the DNN. The blue dotted line is the accuracy for all pixels in the dataset and the red line is the accuracy using only the images used for assessing the intrarater and interrater accuracies. The black and cyan lines are the corresponding, intrarater and interrater accuracies, respectively. As the accuracy for several thresholds are above that of the intrarater and interrater accuracies, we can say that the performance is not sensitive to the chosen threshold.
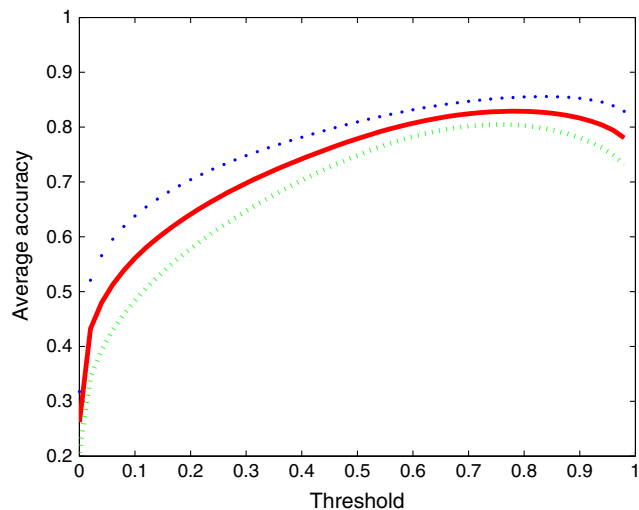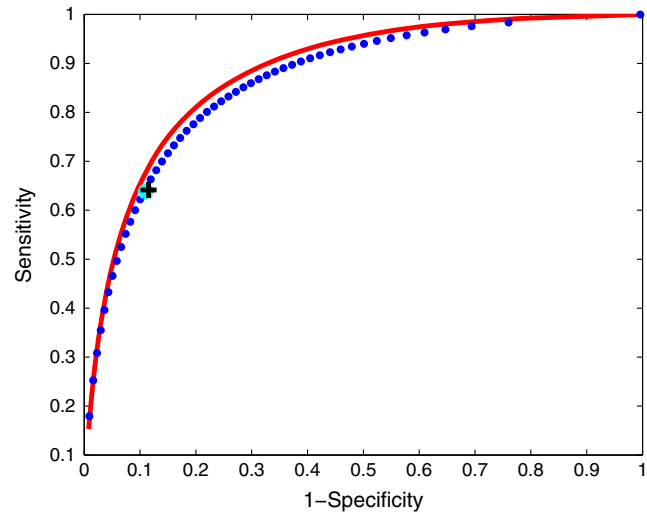

Fig. 4 ROC curves of the segmentation using the DNN. The blue dotted line is the ROC curve for all pixels in the dataset and the red line is the ROC curve for the images used for assessing the intrarater and interrater accuracies. The black cross and cyan dots are the corresponding, intrarater and interrater points, respectively.
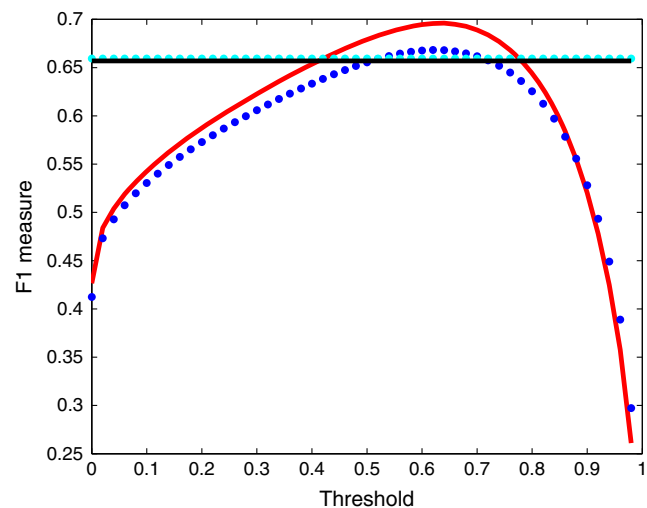

Fig. 3 Mean accuracy of the segmentation using the DNN. The red line is the mean accuracy of the segmentation for all possible threshold values, and the blue dotted line and green dashed line are one standard deviation above and below the mean accuracy, respectively. The small deviation of accuracies signifies a consistent performance of the whole dataset.

Fig. 5 F_1 measure of the segmentation using the DNN. The blue dotted line is the F_1 measure for all pixels in the dataset and the red line is the F_1 measure of the images used for assessing the intrarater and interrater accuracies. The straight black and dotted cyan lines are the corresponding, intrarater and interrater F_1 measures, respectively. The results from the automated DNN method are better than the manual segmentation results for a large range of thresholds, again showing the performance is not sensitive to the threshold chosen.

3.2 Intrarater and Interrater Agreement

As described in Sec. 2.3 among the 80 images segmented by Rater A, 10 images were additionally segmented a second time by Rater A, and also by Rater B for assessing the intra- and interrater agreement. For convenience, we used the accuracy measures discussed above and the original segmentation by Rater A (Rater A1) as the ground-truth in order to assess its agreement with (1) the repeat segmentation of Rater A (Rater A2), (2) Rater B, and (3) the network. The machine segmentation accuracy results of (3) were obtained as part of the threefold validation in Sec. 3.1. The results are shown in Fig. 2 in dotted

is likely due to projection artifact from the superficial vascular layers.

Using the sensitivity and specificity measurements over the range of thresholds we can plot the receiver operator characteristic (ROC) for our method, as shown in Fig. 4 with blue dots. The sensitivity and specificity were calculated using all pixels from the dataset.

In Fig. 5, the F_1 measure was calculated for the machine output using all pixels from the dataset and shown with blue dots.

cyan, solid black, and solid red lines, respectively. The intra- and interrater accuracies for the manual raters are plotted as lines because they are independent of the threshold used for the machine based segmentation. From the Fig. 2, the intrarater, interrater, and machine-rater accuracies are comparable, suggesting that the automated segmentation is comparable to that of a human rater. As it was expected, the accuracy of the repeated segmentation is better than the accuracy of the second rater but the difference is small.

In Fig. 4, the ROC curve of the automated segmentation is compared with Rater A1 (solid red line). In the same figure, the cyan star represents the sensitivity and specificity pair for Rater A2 compared with Rater A1 and the black cross represents the sensitivity and specificity pair for Rater B compared with Rater A1. The ROC curve was created by plotting the sensitivity against the false-positive rate (1-specificity) at various thresholds to depict relative trade-offs between true positives and false positives. A completely random result would be represented by a diagonal line. As seen in Fig. 4, the results from the automated DNN method are better than the manual segmentation results and well above the random result.

In Fig. 5, we can see the F_1 measure curve for the machine output marked with solid red curve, and the F_1 measures for Rater A2 (dotted straight cyan line) and for Rater B (solid straight black line). The F_1 measure depicts the trade-off between precision and recall with each variable weighted equally. As such a higher F_1-measure has a better balance between precision and recall. As seen in Fig. 5, there is a wide range of thresholds in which the balance between precision and recall is higher than the manual raters.

3.3 Capillary Density

Capillary density (CD) is a clinical measure of quantifying retinal capillaries present in the OCT-A images. After segmentation of the vessels, CD can be calculated as the number of pixels in the segmented areas. Using the same 10 images from Sec. 2.3, we obtained the CD values from the segmentations by Rater A1, Rater A2, Rater B, and the network, and calculated the mean capillary density in order to evaluate the intrarater, interrater, and machine-to-rater repeatability of the CD measures. The result is presented in Table 2.

A paired-samples t-test was conducted to compare the capillary density of manual and automated segmentations. There was no significant difference in the scores for either of the manual raters or the machine.

Table 2 Mean capillary density comparison.

	Mean ($N = 10$)	Standard deviation	Standard error mean	p-value
Rater A1	0.2710	0.0399	0.0133	
Rater A2	0.2530	0.0350	0.0117	0.1758
Rater B	0.2583	0.0723	0.0241	0.6187
Machine (threshold = 0.70)	0.2718	0.0342	0.0012	0.9144

4 Discussion

The problem of blood vessel segmentation in OCT-A images is challenging due to the low contrast and high noise levels in OCT-A images. We have presented a deep convolutional neural network-based segmentation method and validation using 80 foveal OCT-A images. In the cross-validation in Sec. 3.1, the accuracy percentage of the trained network fell in range of 80% to 83%. From the results, we conclude that the machine based segmentation was comparable to the manual segmentation by a human rater.

In the intra- and interrater comparison in Sec. 3.2, we found similar degrees of agreement for the repeated segmentations by a single rater, and segmentations from two different raters, showing substantial intra- and interrater variability in the manual
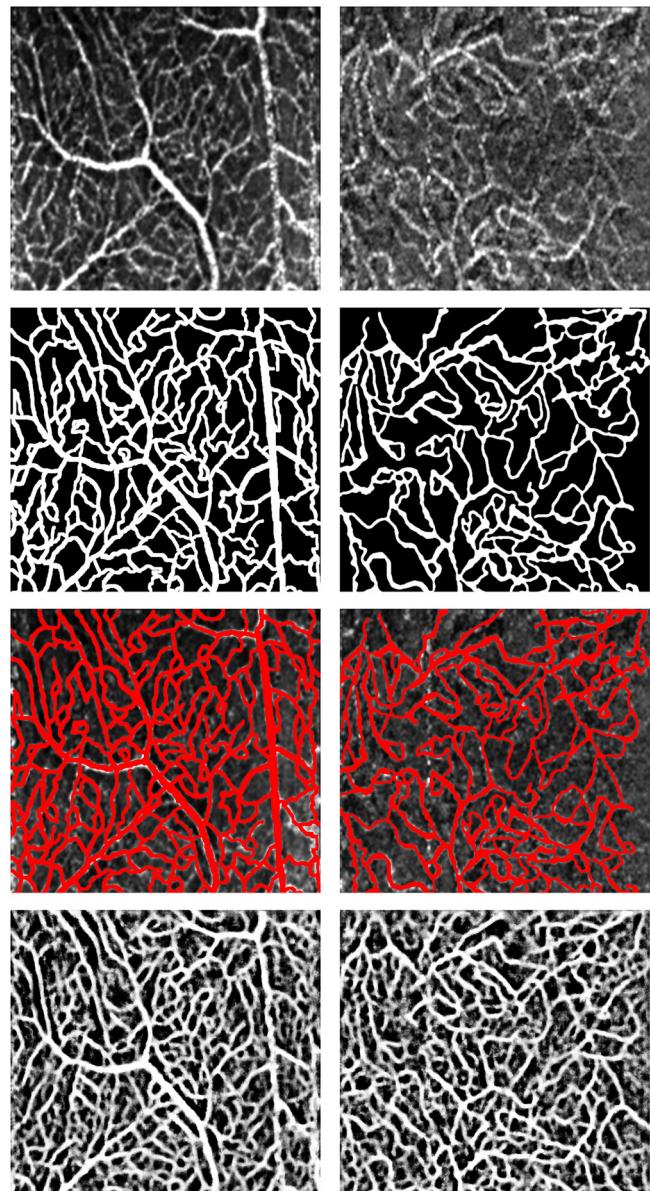

Fig. 6 Examples of OCT-A retinal images acquired with our system (top row), manual segmentations of the vessels (second row), original images with manual segmentations superimposed (third row), and outputs of the proposed DNN method (bottom row). Images in the left column represent an example of a typical dataset and images in the right column represent an example of a low quality dataset.

segmentation. This suggests that the trained network may perform as well as a new human rater. Given the amount of time (~20 to 25 min) required for a human rater to perform the segmentation manually versus 2 min for the automated method, this represents a tool that could be useful in the clinical environment to save valuable human time and present results to the clinician in a shorter interval.

In addition to comparison with manual segmentation, the validity and merit of automated segmentation of medical images can be assessed by deriving clinical parameters such as capillary density. This approach is particularly appropriate if the quality of the derived parameters can be measured, e.g., by the correlation to other relevant clinical features, and if the quality of the manual segmentation ground truth is not reliable. In Sec. 3.3, capillary density was calculated for the manual and machine segmentations. A paired-samples t-test was conducted to compare the capillary density of manual and automated segmentations. There was no significant difference in the scores for either of the manual raters or the machine.

As the performance of a machine learning based approach is closely linked to the quality of the training data, using high quality data is important. However, the performance of a human rater, the ground-truth for training the network, is limited due to the difficulty in delineating the capillaries of some data sets. This was mainly due to poor contrast, vertical motion artifacts, and high noise levels. In Fig. 6, we can see an example of a poor dataset, with an accuracy of 77.12% and an example of a typical dataset with accuracy of 81.16%. We have observed performance variability in the vessel thickness due to the field of view and have chosen to train each field of view separately to take this into account. The dataset in this paper only contains images from one field of view (1 × 1 mm). The automated algorithm does segment the larger vessels (arterioles and venuoles) with a higher degree of certainty than the smaller vessels (capillaries).

This problem could be potentially mitigated by producing ground-truth data that is measurably better than data from a single expert by using images segmented by two or more trained volunteers as the input to the learning procedure. In this case, multiple segmentations of each image would be combined to select regions that are high in agreement by the raters, and the combined image would be then used for the learning procedure. A drawback to this approach would be the human labor cost of several trained raters segmenting a sufficiently large number of images for training purposes. Also, increasing the *enface* image quality in the acquisition stage would increase the quality of the manual rater accuracy and repeatability. This in turn can reduce the noise level in the ground truth data and make this method more robust.

5 Conclusion

Segmentation of the retinal microvasculature is an important step in quantification of retinal images for clinical purposes. For OCT-A, a new method for retinal vasculature visualization, automated segmentation of the retinal vasculature remains a relatively unexplored area. Through comparisons of results from the DNN method and manual raters, the accuracy of our method is found to be comparable to a manual rater. For clinical applications, this is an important step in creating an automated segmentation usable for clinical analysis.

References

1. G. Chan et al., "Quantitative morphometry of perifoveal capillary networks in the human retinaperifoveal capillary networks," *Invest. Ophthalmol. Visual Sci.* **53**(9), 5502 (2012).
2. A. M. Joussen et al., *Retinal Vascular Disease*, Springer, Berlin Heidelberg (2007).
3. K. R. Mendis et al., "Correlation of histologic and clinical images to determine the diagnostic value of fluorescein angiography for studying retinal capillary detail," *Invest. Ophthalmol. Visual Sci.* **51**(11), 5864 (2010).
4. L. A. Yannuzzi et al., "Fluorescein angiography complication survey," *Ophthalmology* **93**(5), 611–617 (1986).
5. A. Zhang et al., "Methods and algorithms for optical coherence tomography-based angiography: a review and comparison," *J. Biomed. Opt.* **20**(10), 100901 (2015).
6. M. S. Mahmud et al., "Review of speckle and phase variance optical coherence tomography to visualize microvascular networks," *J. Biomed. Opt.* **18**(5), 050901 (2013).
7. P. K. Yu et al., "Label-free density measurements of radial peripapillary capillaries in the human retina," *PLoS One* **10**(8), e0135151 (2015).
8. P. E. Z. Tan et al., "Quantitative comparison of retinal capillary images derived by speckle variance optical coherence tomography with histology," *Invest. Ophthalmol. Visual Sci.* **56**(6), 3989–96 (2015).
9. Z. Mammo et al., "Quantitative noninvasive angiography of the fovea centralis using speckle variance optical coherence tomography speckle variance optical coherence tomography of macula," *Invest. Ophthalmol. Visual Sci.* **56**(9), 5074 (2015).
10. G. Chan et al., "In vivo optical imaging of human retinal capillary networks using speckle variance optical coherence tomography with quantitative clinico-histological correlation," *Microvasc. Res.* **100**, 32–39 (2015).
11. K. Sakata et al., "Relationship of macular microcirculation and retinal thickness with visual acuity in diabetic macular edema," *Ophthalmology* **114**(11), 2061–2069 (2007).
12. F. Kirbas and C. Quek, "A review of vessel extraction techniques and algorithms," *ACM Comput. Surv.* **36**(2), 81–121 (2004).
13. Z. Hu et al., "Multimodal retinal vessel segmentation from spectral-domain optical coherence tomography and fundus photography," *IEEE Trans. Med. Imaging* **31**(10), 1900–1911 (2012).
14. T. S. Hwang et al., "Automated quantification of capillary nonperfusion using optical coherence tomography angiography in diabetic retinopathy," *JAMA Ophthalmol.* **134**(4), 367–373 (2016).
15. S. A. Agemy et al., "Retinal vascular perfusion density mapping using optical coherence tomography angiogrphy in normals and diabetic retinopathy patients," *Retina* **35**(11), 2353–2363 (2015).
16. S. Yousefi, T. Liu, and R. K. Wang, "Segmentation and quantification of blood vessels for OCT-based micro-angiograms using hybrid shape/intensity compounding," *Microvasc. Res.* **97**, 37–46 (2015).
17. A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," in *Advances in Neural Information Processing Systems*, pp. 1–9 (2012).
18. K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," 1–13, 2014, arXiv:1409.1556.
19. L. Deng, G. Hinton, and B. Kingsbury, "New types of deep neural network learning for speech recognition and related applications: an overview," in *2013 IEEE Int. Conf. on Acoustics, Speech and Signal Processing*, pp. 8599–8603 (2013).

20. D. Ciresan et al., "Deep neural networks segment neuronal membranes in electron microscopy images," in *Advances in Neural Information Processing Systems (NIPS)*, pp. 1–9 (2012).
21. C. Farabet et al., "Learning hierarchical features for scene labeling," *IEEE Trans. Pattern Anal. Mach. Intell.* **35**(8), 1915–1929 (2013).
22. C. N. dos Santos and M. Gatti, "Deep convolutional neural networks for sentiment analysis of short texts," in *Proc. of COLING 2014, the 25th Int. Conf. on Computational Linguistics: Technical Papers*, pp. 69–78 (2014).
23. D. C. Cirean et al., "Convolutional neural network committees for hand-written character classification," in *Proc. of the Int. Conf. on Document Analysis and Recognition (ICDAR '11)*, pp. 1135–1139 (2011).
24. J. Xu et al., "Retinal angiography with real-time speckle variance optical coherence tomography," *Br. J. Ophthalmol.* **99**(10), 1315–1319 (2015).
25. J. Xu et al., "Real-time acquisition and display of flow contrast using speckle variance optical coherence tomography in a graphics processing unit," *J. Biomed. Opt.* **19**(2), 026001 (2014).
26. J. Xu et al., "GPU open source code with svOCT implementation," Secondary GPU Open Source Code with svOCT Implementation, http://borg.ensc.sfu.ca/research/svoct-gpu-code.html (12 August 2014)

27. N. Srivastava et al., "Dropout: a simple way to prevent neural networks from overfitting," *J. Mach. Learn. Res.* **15**, 1929–1958 (2014).
28. Y. Jia et al., "Caffe: convolutional architecture for fast feature embedding," in *ACM Int. Conf. on Multimedia*, pp. 675–678 (2014).

Pavle Prentašić is a PhD candidate at the Faculty of Electrical Engineering and Computing, University of Zagreb. He received his BS and MEng degrees in computer science from the University of Zagreb in 2010 and 2012, respectively. His current research interests include computer vision, machine learning, and biomedical image processing and analysis. He is a member of IEEE.

Morgan Heisler is a MASc student in the Faculty of Applied Sciences at Simon Fraser University, Canada. She received her BASc (Hons.) from Simon Fraser University in 2015 and her current research interests include optical coherence tomography and biomedical image processing and analysis.

Biographies for the other authors are not available.