Open-Source Deep Learning-Based Automatic Segmentation of Mouse Schlemm's Canal in Optical Coherence Tomography Images

Kevin C. Choy[1], **Guorong Li**[2], **W. Daniel Stamer**[2,1], **Sina Farsiu**[1,2]

[1]Department of Biomedical Engineering, Duke University, Durham, North Carolina, United States

[2]Department of Ophthalmology, Duke University, Durham, North Carolina, United States

Abstract

The purpose of this study was to develop an automatic deep learning-based approach and corresponding free, open-source software to perform segmentation of the Schlemm's canal (SC) lumen in optical coherence tomography (OCT) scans of living mouse eyes. A novel convolutional neural network (CNN) for semantic segmentation grounded in a U-Net architecture was developed by incorporating a late fusion scheme, multi-scale input image pyramid, dilated residual convolution blocks, and attention-gating. 163 pairs of intensity and speckle variance (SV) OCT B-scans acquired from 32 living mouse eyes were used for training, validation, and testing of this CNN model for segmentation of the SC lumen. The proposed model achieved a mean Dice Similarity Coefficient (DSC) of 0.694 ± 0.256 and median DSC of 0.791, while manual segmentation performed by a second expert grader achieved a mean and median DSC of 0.713 ± 0.209 and 0.763, respectively. This work presents the first automatic method for segmentation of the SC lumen in OCT images of living mouse eyes. The performance of the proposed model is comparable to the performance of a second human grader. Open-source automatic software for segmentation of the SC lumen is expected to accelerate experiments for studying treatment efficacy of new drugs affecting intraocular pressure and related diseases such as glaucoma, which present as changes in the SC area.

Keywords

Schlemm's canal; glaucoma; optical coherence tomography; deep learning; image segmentation

Declarations of interest: none

1 Introduction

Glaucoma, a leading cause of irreversible blindness worldwide, is characterized by impaired outflow facility, which results in high intraocular pressure (IOP) (Brubaker, 2003). The secretion of aqueous humor into the eye and regulation of its outflow are physiologically important processes for IOP homeostasis and proper functioning of anterior eye tissues. Schlemm's canal (SC) is a structure in the anterior segment of the eye and a part of the conventional outflow pathway which is responsible for the drainage of the majority of aqueous humor. Abnormal outflow resistance at the SC and trabecular meshwork (TM) interface, a critical region in regulating outflow and IOP, results in elevated IOP, a major risk factor in the progression of glaucoma (Gordon et al., 2002). Reduced cross-sectional area of SC has been observed at elevated IOPs (Kagemann et al., 2014). Further, SC in glaucomatous eyes has a significantly smaller cross-sectional area compared to healthy eyes (Allingham et al., 1996; Hong et al., 2013; Imamoglu et al., 2016; Wang et al., 2012). In addition, changes in SC dimensions in living eyes have been used to predict TM biomechanical properties and outflow function (Li et al., 2019; Li et al., 2021; Wang et al., 2017). Therefore, the size of the SC lumen can provide useful insight into the function and health of the conventional outflow pathway and serve as a biomarker for elevated IOP.

Spectral-domain (SD) optical coherence tomography (OCT) has been demonstrated to be an effective tool for visualizing changes of the SC lumen in vivo (Kagemann et al., 2010; Sarunic et al., 2008; Usui et al., 2011). In addition, OCT-angiography techniques such as speckle variance (SV) OCT can be used to visualize SC due to the presence of blood cells in the SC lumen (Li et al., 2014a). This method calculates the interframe variance of intensity between structural images, utilizing differences in the time-varying properties between fluids and non-moving tissue to create contrast (Mariampillai et al., 2008). OCT has been used to study the properties of SC in human and animal models (Ang et al., 2018). Xin et al. (2016) utilized OCT to measure dynamic pressure-dependent changes of SC in ex vivo human eyes. Huang et al. (2017) developed a three-dimensional model of circumferential aqueous humor outflow and an automated SC detection algorithm using SD-OCT in a living human eye. Daniel et al. (2018) measured SC cross-sectional area to study morphologic changes in aqueous outflow structures during accommodation in humans. Previous studies have also used OCT to monitor anatomical changes to SC due to canaloplasty and phacoemulsification (Fuest et al., 2016; Paulaviciute-Baikstiene et al., 2016; Zhao et al., 2016).

Previous work monitoring the SC lumen using in vivo OCT imaging in mouse eyes is of particular interest. Mice have several advantages for studying conventional outflow function: First, the conventional outflow pathway of mice is similar to that of humans in terms of anatomy, physiology, and pharmacology (Aihara et al., 2003; Boussommier-Calleja et al., 2012; Lei et al., 2011; Li et al., 2014b; Millar et al., 2011). Second, mouse eyes can be cannulated to control aqueous humor content (e.g., drug concentration or contrast agent) or IOP level while imaging (Li et al., 2014a). Third, in terms of imaging advantages, mice have a thin nonpigmented sclera, large SC, and a small eye (enabling 360° visualization of outflow). Fourth, since aging is a major risk factor for glaucoma, the short lifespan of mice enables the study of outflow function in normal, aged, and diseased eyes. Li et al. (2019) used a mouse model to investigate the effects of rho-kinase inhibitors on the

perfusion of outflow tissues by monitoring the size of SC in OCT images. Additionally, Li et al. (2016) monitored the effects of ocular corticosteroid treatment, which can lead to steroid-induced glaucoma, on the cross-sectional area of the SC lumen in mice. Furthermore, Zhang et al. (2020) utilized visible-light OCT for visualization of the entire SC and limbal microvascular network in mice. Although analysis of SC can provide useful insight on intraocular hypertension and glaucoma, image segmentation is a current bottleneck. Manual segmentation of the SC lumen is time-consuming and also susceptible to inter-observer variability. There have been several attempts at automating the segmentation of SC in human eyes. Huang et al. (2017) used a Bayesian Ridge method on confocal laser scanning ophthalmoscopic images to approximate the location of SC as a starting point for a fuzzy hidden Markov Chain segmentation approach on OCT images. Tom et al. (2015) used a region-growing method for segmentation of SC in OCT images. Wang et al. (2020) used K-means, fuzzy C-means, and the level set method for segmentation of SC in ultrasound biomicroscopy (UBM) images. Yao et al. (2021) performed full circumferential imaging of the SC using swept-source OCT and semi-automatic segmentation using active contours. However, these algorithms which are based on classic machine learning techniques still require manual intervention, or are not evaluated on an adequate independent dataset, or do not provide delineation of the SC boundary. Also, none of these algorithms are directly applicable for the segmentation of the SC lumen in OCT images of living mouse eyes.

Recently, deep convolutional neural networks (CNNs) have achieved remarkable improvements in a wide range of computer vision tasks, including image classification, object detection, and segmentation (Long et al., 2015). Although deep learning-based methods have been widely utilized in many areas of ophthalmic image analysis (Cabrera DeBuc and Arthur, 2019; Devalla et al., 2018a; Fang et al., 2017a; Guo et al., 2020; Loo et al., 2020; Venhuizen et al., 2017; Xiao et al., 2017), they have yet to be applied for the segmentation of the SC lumen, which is a challenging task due to the variable size of the SC and low contrast with surrounding tissues.

In this paper, we present LF-DRAG-UNet (Late Fusion Dilated Residual Attention-Gated U-Net), a deep learning-based approach for segmentation of the SC lumen in OCT images of living mouse eyes, which is trained end-to-end on a dataset of paired OCT and SV-OCT images. This proposed model uses separate encoders to learn features for each input image and utilizes an input image pyramid, dilated residual convolutions, and attention gating to enhance segmentation performance. To facilitate future studies on the assessment of outflow facility, we made our dataset and algorithms freely available online as an open-source software package.

2   Materials and supplies

2.1   Animals

All data used in this study were imaged at the Duke Eye Center. The animals were handled in accordance with approved Institutional Animal Care and Use Committee of Duke University protocol (A001-19-01) and in compliance with the Association for Research in Vision and Ophthalmology (ARVO) Statement for the Use of Animals in Ophthalmic and Vision Research. C57BL/6 (C57) and 129s mice were obtained from the Jackson

Laboratory. $Loxl1^{+/-}$ mice on mixed 129s/ C57BL/6 background were generously provided by Profs. Tiansen Li and Janey Wiggs (Massachusetts Eye and Ear Infirmary/National Institutes of Health) and inbred to generate $Loxl1^{+/+}$, $Loxl1^{+/-}$, and $Loxl1^{-/-}$ littermates. Generation of $Loxl1^{-/-}$ mice has been previously described in detail (Liu et al., 2004). The mice were bred/housed in clear cages and kept in housing rooms at 21°C with a 12-hour light-dark cycle. Mice were between two and twelve months of age during experiments. Mice were anesthetized using intraperitoneal injection of ketamine (100 mg/kg) and xylazine (10 mg/kg) and maintained by injection of ketamine (60 mg/kg) every 20 minutes if necessary. During experiments, mice were kept at room temperature because (1) a temperature-controlled heating pad could not be placed on the custom-built mouse platform due to limited space, and (2) OCT experiments were end point measurements, usually lasting about 30 minutes, and mice were euthanized after imaging.

2.2 OCT imaging

Mice imaging data was sourced from our previous experiments in (Li et al., 2019; Li et al., 2020). In vivo imaging of mouse eyes was performed using an Enisu R2200 SD-OCT system (Bioptigen Inc., Research Triangle Park, NC) following our previously established methodology (Li et al., 2014a; Li et al., 2019; Li et al., 2020). Briefly, mice were anesthetized then placed on a custom-made OCT imaging mount, equipped with an onboard micromanipulator which was used to insert a glass cannula into the anterior chamber to control and measure IOP. The OCT probe was aimed at the inferior lateral limbus and the image was centered and focused at SC.

While collecting images at the same region of SC, eyes were subjected to a series of IOP steps and imaged sequentially at each pressure (Li et al., 2019; Li et al., 2020). In a subset of mice, the IOP was returned to the initial pressure of 10 mmHg for an additional measurement, resulting in no more than two samples collected at each IOP, for each eye. OCT scans were collected from 32 living mouse eyes, with 5 – 7 samples obtained per eye. These samples were attained from IOP values of 10, 12, 15, 17, 20, 22, and 25 mmHg. IOPs that far exceeded this were not included since the SC lumen of normal mice is almost entirely collapsed at 20 mmHg. For each sample, two hundred consecutive B-scans (each with 1,000 A scans spanning 0.5 mm in lateral length) were captured at the same imaging location. Image registration was performed using ImageJ (National Institutes of Health, Bethesda, MD, USA) with the StackReg plugin using a rigid body transformation (Thevenaz et al., 1998).

2.3 Dataset

For each sample, the average intensity OCT image was obtained by taking the pixel-wise mean across all B-scans to create a high signal-to-noise ratio image. The SV-OCT image was obtained by taking the pixel-wise variance over five consecutive B-scans and then computing the pixel-wise average of the resulting interframe variance calculations. Images were cropped to remove border artifacts from image registration. OCT image quality was assessed using the Quality Index (*QI*) introduced by Stein et al. (2006), and images with a *QI* < 50 were excluded from the dataset since the SC in these images was not discernable,

and therefore could not be accurately segmented even by experienced graders. Out of 167 images collected, 163 (97.6 %) images met this criterion and were included in the dataset.

Manual segmentation of SC in OCT images was performed by an expert reader using a freehand annotation tool provided by the SchlemmSeg software (Li et al., 2016). The grader performed segmentation based on the intensity and SV-OCT images. Speckle variance was monitored to provide a complementary source of information to help identify the SC boundary. In an OCT image, the SC typically appears as a dark oblong closed contour. In contrast, the SC in an SV-OCT image is brighter compared to the surrounding area, due to increased speckling generated by the flow of blood cells through the SC. Also, the grader observed the two hundred B-scan frames as a video to aid in the accurate identification of the boundaries of the SC lumen. Additionally, manual segmentation was performed by a second grader to assess inter-observer variability and to obtain a benchmark for human-level performance for segmentation.

In total, the dataset used in this study was comprised of 163 pairs of intensity OCT and SV-OCT images. Figure 1 displays sample images from one eye in the dataset. The dataset was partitioned into independent training, validation, and test sets using six-fold cross-validation as discussed in Section 3.3. The following data augmentation steps were performed during training. Images were first resized so that pixels were isomorphic in x- and y-directions. Images were then transformed by applying random horizontal flips, random translations in the x- and y-directions up to 20% of the image size, random scalings from 75% to 125% of the image size, and random rotations within a range of 30 degrees. Images were then cropped to 512 by 512 pixels. The pixel values for each input image were linearly scaled by subtracting the mean and dividing by the standard deviation and normalized to [0, 1].

3   Detailed methods

3.1   Network architecture

In this work, we propose LF-DRAG-UNet, a custom CNN architecture for the segmentation of SC using intensity and SV-OCT images. The structure of the proposed network is illustrated in Figure 2. Our model was based on U-Net (Ronneberger et al., 2015), a popular architecture for semantic segmentation, in which each pixel in an image is assigned a probability of belonging to a particular class: SC or background. A standard U-Net is composed of symmetric encoder (contractive) and decoder (expansive) paths. The encoder is made up of four encoder blocks, in which each block consists of two successive series of a convolution layer, batch normalization layer (Ioffe and Szegedy, 2015), and rectified linear unit (ReLU) activation, followed by a max pooling layer. Similarly, the decoder is made up of four decoder blocks, whereby the input to each decoder block is upsampled using a transposed convolution layer and concatenated to features passed through a skip connection from the corresponding encoder block. This is followed by two successive series of a convolution layer, batch normalization layer, and ReLU activation. In the proposed LF-DRAG-UNet, several modifications were made to the standard U-Net by incorporating dilated residual convolutions, an asymmetrical encoder-decoder structure, a multi-scale input image pyramid, and attention gating.

Residual learning with dilated convolutions was used, which has been shown to improve segmentation performance in previous works (Apostolopoulos et al., 2017; Chen et al., 2018; Devalla et al., 2018b; Diakogiannis et al., 2020). The standard encoder block found in U-Net was modified using parallel atrous convolutions with dilation rates of $d = 1$, $d = 3$, and $d = 5$ to increase the receptive field of each layer with little overhead in computational cost (Chen et al., 2018). The outputs of the atrous convolution branches were added to the input of the block for residual learning, which has been demonstrated to improve gradient flow through the network (He et al., 2016a; Zhang et al., 2018). Batch normalization and activations were applied in the pre-activation design (He et al., 2016b). The structure of this dilated residual convolution block is illustrated in Figure 3.

In addition, the proposed network possessed an asymmetrical structure, consisting of two encoders and one decoder. In order to learn meaningful representations for both the intensity and SV-OCT images, a separate encoder path was used for each input, and their features were fused downstream in the network. This type of late fusion technique is common in tasks involving multiple inputs (Cunefare et al., 2018; Hazirbas et al., 2017; Karpathy et al., 2014). An input image pyramid was also used, whereby sequentially downsampled inputs were injected into the encoders to capture features at multiple scales.

In addition, attention gating was used to highlight salient regions of the image. Attention gating involves taking an element-wise product of the feature maps with learned attention coefficients (Oktay et al., 2018). This has the effect of pruning regions of the image with less useful information and has been shown to be effective for the segmentation of targets with high variability in size. A gating signal, g, obtained from a coarser resolution, was used to determine salient regions in the image. For a feature map x^l at layer l, the attention coefficient α^l was computed with the following equations:

$$q_{att}^l = \psi^\top\left(\sigma_1\left(W_x^\top x_i^l + W_g^\top g_i^l + b_{xg}\right)\right) + b_\psi,$$

$$\alpha^l = \sigma_2\left(q_{att}^l\left(x^l, g; \theta_{att}\right)\right),$$

where $\sigma_1(x)$ is a ReLU nonlinearity, $\sigma_2(x)$ is a sigmoid normalization, and θ_{att} is a set of parameters consisting of linear transformations and bias terms: ψ, W_x, W_g, b_{xg}, and b_ψ.

Figure 4 illustrates the structure of the late fusion attention gate used to combine attention maps learned using features from the dual encoders. The attention coefficients, which were optimized to recognize salient regions in each input, were combined by taking a weighted element-wise sum of the attention coefficients from both encoder branches as follows:

$$\hat{\alpha}_i^l = \frac{1}{w_1 w_2}\left(w_1 \alpha_{1,\,i}^l + w_2 \alpha_{2,\,i}^l\right),$$

where $\hat{\alpha}_i^l$ is the final attention coefficient for pixel i, and w_1 and w_2 are the weights applied to each attention map. Finally, deep supervision, which forces intermediate feature maps

to be semantically discriminative at every scale, was used to encourage attention gates at different scales to be able to influence the responses to a large range of foreground content (Lee et al., 2015).

3.2 Implementation and training

Deep learning models were implemented in Python 3.7 with Pytorch (version 1.7.1) and trained using NVIDIA RTX 2080 Ti GPUs on a Linux (Ubuntu) machine (Paszke et al., 2019). Network weights were randomly initialized using He initialization (He et al., 2015). Optimization was performed using stochastic gradient descent with Nesterov acceleration with an initial learning rate of 0.01 and momentum of 0.9 (Sutskever et al., 2013). L_2 weight regularization was used with a factor of 0.0001. The learning rate was reduced using a cosine annealing schedule with warm restarts to encourage the optimizer to jump out of local minima and reach a flatter minimum (Loshchilov and Hutter, 2017).

Networks were trained using a Dice loss objective function, which is defined as

$$L_{Dice} = 1 - \frac{2\sum_i p_i g_i}{\sum_i p_i^2 + \sum_i g_i^2},$$

where $p_i \in p$ are pixels in the predicted segmentation p and $g_i \in G$ are pixels in the ground truth G (Milletari et al., 2016). Dice loss has been shown to be robust to class imbalance, which makes it suitable for the segmentation of SC, since SC occupies a small area relative to the size of the background in an OCT image (Milletari et al., 2016).

3.3 Cross-validation

Six-fold cross-validation was used to train and evaluate the networks using all available data while avoiding selection bias. Four folds contained B-scans from five eyes and two folds contained B-scans from six eyes. Four folds were used for training, one fold was used for validation, and the remaining fold was left out for testing. The partition was done so that if B-scans of a particular eye were in one set, no B-scans from the same eye appeared in a different set, thereby ensuring that training, validation, and test sets were independent.

All networks were trained from scratch for 200 epochs with a batch size of 2. The network corresponding to the best validation loss achieved during the last 20 epochs was used to evaluate performance on the test set. This process was repeated six times, cycling the partitions so that each eye was included in the test set a single time. All code for the method proposed in this paper, along with the annotated SC dataset, are available at https://github.com/kevinchoy/oct-schlemm-seg.

3.4 Testing and evaluation

The output of each network is a 2D probability map in which each pixel value (ranging from 0 to 1) represents the predicted probability of that pixel being classified as SC. A fixed threshold of 0.5 was applied on the output to obtain a binary mask, such that pixels where the network predicted a probability greater than or equal to 0.5 were classified as SC, while pixels with probabilities less than 0.5 were classified as background. The following

post-processing steps were performed on the binary mask to obtain the final segmentation mask. First, a morphological dilation operation was performed with a circular structuring element with a radius of 8 pixels to connect nearby discontinuous segments, since gaps between these segments were typically only a few pixels wide, while erroneously labeled false positive pixels were generally further away (Soille, 2013). Second, since the SC is a single continuous structure, all but the largest connected component was discarded. Third, a morphological erosion was performed to restore the original size of the mask. Last, a binary hole filling operation was used to obtain the final segmentation.

The performance of each network was evaluated using Dice Similarity Coefficient (DSC), also known as Sørensen–Dice coefficient or F1 score (Dice, 1945). DSC measures the overlap between two binary masks and is commonly used to evaluate the performance of segmentation algorithms. The ideal case of perfect overlap between a predicted mask and target mask results in a DSC of 1, whereas no overlap results in a DSC of 0. In addition to DSC, performance was also evaluated using precision, and recall (also known as sensitivity). These metrics are given by the following equations:

$$DSC = \frac{2\sum_i p_i g_i}{\sum_i p_i^2 + \sum_i g_i^2} = \frac{2TP}{2TP + FP + FN},$$

$$Precision = \frac{TP}{TP + FP},$$

$$Recall = \frac{TP}{TP + FN},$$

where true positive (TP) refers to the number of pixels in the output mask that were correctly predicted as part of SC, false positive (FP) refers to the number of pixels incorrectly labeled as part of SC but were part of the background, and false negative (FN) refers to the number of pixels that were incorrectly classified as part of the background but belonged to SC.

3.5   Deep learning-based segmentation is comparable to segmentation by a human grader

We trained our proposed LF-DRAG-UNet model along with a standard U-Net on the dataset of intensity and SV-OCT images of living mouse eyes. The outputs of all trained networks were corrected using the post-processing steps described in Section 3.4 to obtain final segmentation masks. Figure 5 displays an illustrative example, which shows that LF-DRAG-UNet obtained segmentations that were qualitatively more accurate compared to the baseline model. We also evaluated the performance of both U-Net and our modified network when trained using OCT images, SV-OCT images, and both OCT and SV-OCT images. When training was conducted using either OCT or SV-OCT images alone, only one encoder was necessary. We refer to the single-encoder version of the proposed network as DRAG-UNet (Dilated Residual Attention-Gated U-Net) since the architecture of this modified network did not have a late-fusion component. Evaluation of deep learning models and manual

segmentation performed by a second human grader are summarized in Table 1. Comparison between proposed and baseline methods was made using the Wilcoxon signed-rank test. LF-DRAG-UNet surpassed the baseline U-Net trained using both OCT and SV-OCT images in mean and median performance on all metrics except median precision. The differences in DSC, precision, and recall were statistically significant ($p < 0.05$).

Both the standard U-Net and DRAG-UNet trained using only SV-OCT images performed significantly worse than their counterparts trained using only intensity OCT images and those trained using both input images. The U-Net and DRAG-UNet trained using only OCT images performed similarly to the corresponding models trained using both input images. The U-Net trained using only OCT images had a higher DSC and recall than the U-Net trained on both inputs. However, these differences were not statistically significant. Furthermore, we compare the performance of the proposed model with that of a second human grader. LF-DRAG-UNet recorded a lower mean DSC and precision than the human grader but achieved comparable mean recall. In addition, LF-DRAG-UNet also achieved higher median scores for each metric compared to the expert human grader. Differences between the proposed model and human grader were not statistically significant.

3.5.1 Discussion

The TM/SC plays a major role in regulating IOP, and is especially important in the study of glaucoma, as ocular hypertension is a major risk factor for this disease (Allingham et al., 1996). Prior research monitored the size of SC to evaluate the effectiveness of glaucoma drugs and treatment (Li et al., 2014a; Li et al., 2019; Li et al., 2021; Li et al., 2016). Image segmentation is a crucial step in many ophthalmic applications and is required for the analysis of the SC cross-sectional area. In previous studies, the SC region was annotated by manual segmentation, which is both time-consuming and subject to inter-observer variability (Li et al., 2014a; Li et al., 2019; Li et al., 2016). Automatic segmentation of SC is expected to accelerate experiments for studying treatment efficacy of new drugs affecting IOP and facilitate the study and development of glaucoma treatments.

In this study, an automatic deep learning-based approach was used for semantic segmentation of SC. Our proposed CNN model, LF-DRAG-UNet, obtained superior segmentation performance compared to the baseline U-Net. A late fusion scheme was used, whereby, instead of concatenating the intensity OCT and SV-OCT images as a two-channel input to the network, separate encoders were trained for each input, which was better able to combine information from both sources. In addition, an input image pyramid and dilated residual convolution blocks were used to improve the model's ability to use global information. Furthermore, attention gating was used to prune less important information. Figure 6 displays an example heatmap of the attention coefficients for the final attention gate in the network, which shows that the model was able to focus on the salient region around SC and suppressed activations in the background.

3.5.2 OCT and SV-OCT comparison

Table 1 provides a summary of the performance of both standard and modified U-Net models trained using OCT images, SV-OCT images, and both input images. Figure 7 displays illustrative example segmentations using these models. In the majority of cases, models that used both intensity and SV-OCT images performed better than those that used only one type of input image. The models

which used only intensity images performed similarly to the corresponding models which used both inputs, while models that used only SV-OCT images performed significantly worse than their counterparts. This indicates that the networks relied heavily on the OCT images over the SV-OCT images. The benefit of using both input images was especially apparent when SC was difficult to see in the intensity images but clearly visible in the SV-OCT image. However, models that used both input images would occasionally miss the SC lumen when SC was visible in the OCT image but difficult to see in the SV-OCT image. While SV-OCT provided an auxiliary source of information, which could aid in segmentation when the OCT image was poor, it could also confuse the models when the SV-OCT image itself was poor. However, typically, when the SC boundary was visible in the OCT image, it was also visible in the SV-OCT image.

4 Potential pitfalls and troubleshooting

Our method generally performed well on samples that had clear boundaries and high contrast in both intensity and SV-OCT images. The SC lumen in these images was clearly visible and easy to identify. In discussing the limitations of our method, we highlight the cases in which our algorithm did not correctly segment the SC lumen. Figure 8 displays example cases where the network performed poorly. Several factors make the segmentation of SC challenging. Low contrast between SC and surrounding tissue makes it difficult to find the boundaries of SC. In addition, the presence of superficial blood vessels with high scattering properties and associated shadowing can obscure the SC in the SV-OCT image. Furthermore, SC shrinks and can even collapse at high enough pressure in healthy eyes, making segmentation more difficult for eyes at elevated IOP (Li et al., 2016). In these cases, our deep learning-based method may not be able to successfully annotate the SC lumen, and the resulting segmentations may require manual correction. These manually corrected datasets can be used to retrain the algorithm and reduce the automatic segmentation algorithm's error in future instances (Loo et al., 2018).

We used averaging to improve the signal-to-noise ratio of OCT images. In applications where repeated imaging and averaging is not possible, efficient single or multi-frame denoising algorithms can be utilized to reduce the impact of speckle noise on segmentation accuracy (Fang et al., 2017b).

Figure 9 shows the performance of the baseline U-Net and LF-DRAG-UNet on OCT scans of mouse eyes at different IOP. The proposed network was better able to segment images at higher pressure, but both the baseline and proposed networks performed worse with samples obtained at increased IOP. Although lower performance on small targets may be attributed to the nature of pixel overlap metrics such as DSC, segmentation of SC at high IOP remains a challenge and a potential area of improvement. However, these factors are also difficulties that are present during manual segmentation by a human grader, and automatic segmentation of SC by the proposed deep network was not found to be significantly different than that of manual segmentation.

5 Concluding remarks

In this paper, we proposed the first fully automatic approach for segmentation of the SC lumen that is on-par with human-level performance. This study demonstrates that deep learning-based methods can be used to automatically segment the SC lumen in OCT images. Our method can aid in the study of SC by removing the need for time-intensive manual segmentation. This would make it easier to study the effects of drugs and treatment affecting the conventional outflow pathway, IOP, and related diseases such as glaucoma, which present as changes in SC area.

Acknowledgments

Abbreviations:

SC	Schlemm's Canal
TM	trabecular meshwork
IOP	intraocular pressure
OCT	optical coherence tomography
SV-OCT	speckle variance optical coherence tomography
QI	Quality Index
DSC	Dice Similarity Coefficient
CNN	convolutional neural network
ReLU	rectified linear unit

References

Aihara M, Lindsey JD, Weinreb RN, 2003. Experimental mouse ocular hypertension: establishment of the model. Invest Ophthalmol Vis Sci 44, 4314–4320. 10.1167/iovs.03-0137. [PubMed: 14507875]

Allingham RR, de Kater AW, Ethier CR, 1996. Schlemm's canal and primary open angle glaucoma: correlation between Schlemm's canal dimensions and outflow facility. Exp Eye Res 62, 101–109. 10.1006/exer.1996.0012. [PubMed: 8674505]

Ang M, Baskaran M, Werkmeister RM, Chua J, Schmidl D, Aranha Dos Santos V, Garhofer G, Mehta JS, Schmetterer L, 2018. Anterior segment optical coherence tomography. Prog Retin Eye Res 66, 132–156. 10.1016/j.preteyeres.2018.04.002. [PubMed: 29635068]

Apostolopoulos S, De Zanet S, Ciller C, Wolf S, Sznitman R, 2017. Pathological OCT Retinal Layer Segmentation Using Branch Residual U-Shape Networks, Medical Image Computing and Computer Assisted Intervention – MICCAI 2017, pp. 294–301. 10.1007/978-3-319-66179-7_34.

Boussommier-Calleja A, Bertrand J, Woodward DF, Ethier CR, Stamer WD, Overby DR, 2012. Pharmacologic manipulation of conventional outflow facility in ex vivo mouse eyes. Invest Ophthalmol Vis Sci 53, 5838–5845. 10.1167/iovs.12-9923. [PubMed: 22807298]

Brubaker RF, 2003. Targeting outflow facility in glaucoma management. Surv Ophthalmol 48 Suppl 1, S17–20. 10.1016/s0039-6257(03)00003-1.

Cabrera DeBuc D, Arthur E, 2019. Recent Developments of Retinal Image Analysis in Alzheimer's Disease and Potential AI Applications, Computer Vision – ACCV 2018 Workshops, pp. 261–275. 10.1007/978-3-030-21074-8_21.

Chen LC, Papandreou G, Kokkinos I, Murphy K, Yuille AL, 2018. DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs. IEEE Trans Pattern Anal Mach Intell 40, 834–848. 10.1109/TPAMI.2017.2699184. [PubMed: 28463186]

Cunefare D, Langlo CS, Patterson EJ, Blau S, Dubra A, Carroll J, Farsiu S, 2018. Deep learning based detection of cone photoreceptors with multimodal adaptive optics scanning light ophthalmoscope images of achromatopsia. Biomed Opt Express 9, 3740–3756. 10.1364/BOE.9.003740. [PubMed: 30338152]

Daniel MC, Dubis AM, Quartilho A, Al-Hayouti H, Khaw SPT, Theodorou M, Dahlmann-Noor A, 2018. Dynamic Changes in Schlemm Canal and Iridocorneal Angle Morphology During Accommodation in Children With Healthy Eyes: A Cross-Sectional Cohort Study. Invest Ophthalmol Vis Sci 59, 3497–3502. 10.1167/iovs.17-23189. [PubMed: 30025073]

Devalla SK, Chin KS, Mari JM, Tun TA, Strouthidis NG, Aung T, Thiery AH, Girard MJA, 2018a. A Deep Learning Approach to Digitally Stain Optical Coherence Tomography Images of the Optic Nerve Head. Invest Ophthalmol Vis Sci 59, 63–74. 10.1167/iovs.17-22617. [PubMed: 29313052]

Devalla SK, Renukanand PK, Sreedhar BK, Subramanian G, Zhang L, Perera S, Mari JM, Chin KS, Tun TA, Strouthidis NG, Aung T, Thiery AH, Girard MJA, 2018b. DRUNET: a dilated-residual U-Net deep learning network to segment optic nerve head tissues in optical coherence tomography images. Biomed Opt Express 9, 3244–3265. 10.1364/BOE.9.003244. [PubMed: 29984096]

Diakogiannis FI, Waldner F, Caccetta P, Wu C, 2020. ResUNet-a: A deep learning framework for semantic segmentation of remotely sensed data. ISPRS Journal of Photogrammetry and Remote Sensing 162, 94–114. 10.1016/j.isprsjprs.2020.01.013.

Dice LR, 1945. Measures of the amount of ecologic association between species. Ecology 26, 297–302. 10.2307/1932409.

Fang L, Cunefare D, Wang C, Guymer RH, Li S, Farsiu S, 2017a. Automatic segmentation of nine retinal layer boundaries in OCT images of non-exudative AMD patients using deep learning and graph search. Biomed Opt Express 8, 2732–2744. 10.1364/BOE.8.002732. [PubMed: 28663902]

Fang L, Li S, Cunefare D, Farsiu S, 2017b. Segmentation Based Sparse Reconstruction of Optical Coherence Tomography Images. IEEE Trans Med Imaging 36, 407–421. 10.1109/TMI.2016.2611503. [PubMed: 27662673]

Fuest M, Kuerten D, Koch E, Becker J, Hirsch T, Walter P, Plange N, 2016. Evaluation of early anatomical changes following canaloplasty with anterior segment spectral-domain optical coherence tomography and ultrasound biomicroscopy. Acta Ophthalmol 94, e287–292. 10.1111/aos.12917. [PubMed: 26648049]

Gordon MO, Beiser JA, Brandt JD, Heuer DK, Higginbotham EJ, Johnson CA, Keltner JL, Miller JP, Parrish RK 2nd, Wilson MR, Kass MA, 2002. The Ocular Hypertension Treatment Study: baseline factors that predict the onset of primary open-angle glaucoma. Arch Ophthalmol 120, 714–720; discussion 829–730. 10.1001/archopht.120.6.714. [PubMed: 12049575]

Guo Y, Hormel TT, Xiong H, Wang J, Hwang TS, Jia Y, 2020. Automated Segmentation of Retinal Fluid Volumes From Structural and Angiographic Optical Coherence Tomography Using Deep Learning. Translational Vision Science & Technology 9. 10.1167/tvst.9.2.54.

Hazirbas C, Ma L, Domokos C, Cremers D, 2017. FuseNet: Incorporating Depth into Semantic Segmentation via Fusion-Based CNN Architecture, Computer Vision – ACCV 2016, pp. 213–228. 10.1007/978-3-319-54181-5_14.

He K, Zhang X, Ren S, Sun J, 2015. Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification, 2015 IEEE International Conference on Computer Vision (ICCV), pp. 1026–1034. 10.1109/iccv.2015.123.

He K, Zhang X, Ren S, Sun J, 2016a. Deep Residual Learning for Image Recognition, 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 770–778. 10.1109/cvpr.2016.90.

He K, Zhang X, Ren S, Sun J, 2016b. Identity Mappings in Deep Residual Networks, Computer Vision – ECCV 2016, pp. 630–645. 10.1007/978-3-319-46493-0_38.

Hong J, Xu J, Wei A, Wen W, Chen J, Yu X, Sun X, 2013. Spectral-domain optical coherence tomographic assessment of Schlemm's canal in Chinese subjects with primary open-angle glaucoma. Ophthalmology 120, 709–715. 10.1016/j.ophtha.2012.10.008. [PubMed: 23352198]

Huang AS, Belghith A, Dastiridou A, Chopra V, Zangwill LM, Weinreb RN, 2017. Automated circumferential construction of first-order aqueous humor outflow pathways using spectral-domain optical coherence tomography. J Biomed Opt 22, 66010. 10.1117/1.JBO.22.6.066010. [PubMed: 28617922]

Imamoglu S, Sevim MS, Alpogan O, Ercalik NY, Kumral ET, Pekel G, Bardak H, 2016. In vivo biometric evaluation of Schlemm's canal with spectral-domain optical coherence tomography in pseuduexfoliation glaucoma. Acta Ophthalmol 94, e688–e692. 10.1111/aos.13080. [PubMed: 27167320]

Ioffe S, Szegedy C, 2015. Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift, in: Francis B, David B (Eds.), Proceedings of the 32nd International Conference on Machine Learning. PMLR, Proceedings of Machine Learning Research, pp. 448–456.

Kagemann L, Wang B, Wollstein G, Ishikawa H, Nevins JE, Nadler Z, Sigal IA, Bilonick RA, Schuman JS, 2014. IOP elevation reduces Schlemm's canal cross-sectional area. Invest Ophthalmol Vis Sci 55, 1805–1809. 10.1167/iovs.13-13264. [PubMed: 24526436]

Kagemann L, Wollstein G, Ishikawa H, Bilonick RA, Brennen PM, Folio LS, Gabriele ML, Schuman JS, 2010. Identification and assessment of Schlemm's canal by spectral-domain optical coherence tomography. Invest Ophthalmol Vis Sci 51, 4054–4059. 10.1167/iovs.09-4559. [PubMed: 20237244]

Karpathy A, Toderici G, Shetty S, Leung T, Sukthankar R, Fei-Fei L, 2014. Large-Scale Video Classification with Convolutional Neural Networks, 2014 IEEE Conference on Computer Vision and Pattern Recognition, pp. 1725–1732. 10.1109/cvpr.2014.223.

Lee C-Y, Xie S, Gallagher P, Zhang Z, Tu Z, 2015. Deeply-Supervised Nets, Proceedings of the Eighteenth International Conference on Artificial Intelligence and Statistics. PMLR, pp. 562–570.

Lei Y, Overby DR, Boussommier-Calleja A, Stamer WD, Ethier CR, 2011. Outflow physiology of the mouse eye: pressure dependence and washout. Invest Ophthalmol Vis Sci 52, 1865–1871. 10.1167/iovs.10-6019. [PubMed: 21169533]

Li G, Farsiu S, Chiu SJ, Gonzalez P, Lutjen-Drecoll E, Overby DR, Stamer WD, 2014a. Pilocarpine-induced dilation of Schlemm's canal and prevention of lumen collapse at elevated intraocular pressures in living mice visualized by OCT. Invest Ophthalmol Vis Sci 55, 3737–3746. 10.1167/iovs.13-13700. [PubMed: 24595384]

Li G, Farsiu S, Qiu J, Dixon A, Song C, McKinnon SJ, Yuan F, Gonzalez P, Stamer WD, 2014b. Disease progression in iridocorneal angle tissues of BMP2-induced ocular hypertensive mice with optical coherence tomography. Mol Vis 20, 1695–1709. [PubMed: 25558173]

Li G, Lee C, Agrahari V, Wang K, Navarro I, Sherwood JM, Crews K, Farsiu S, Gonzalez P, Lin CW, Mitra AK, Ethier CR, Stamer WD, 2019. In vivo measurement of trabecular meshwork stiffness in a corticosteroid-induced ocular hypertensive mouse model. Proc Natl Acad Sci U S A 116, 1714–1722. 10.1073/pnas.1814889116. [PubMed: 30651311]

Li G, Lee C, Read AT, Wang K, Ha J, Kuhn M, Navarro I, Cui J, Young K, Gorijavolu R, Sulchek T, Kopczynski C, Farsiu S, Samples J, Challa P, Ethier CR, Stamer WD, 2021. Anti-fibrotic activity of a rho-kinase inhibitor restores outflow function and intraocular pressure homeostasis. Elife 10. 10.7554/eLife.60831.

Li G, Mukherjee D, Navarro I, Ashpole NE, Sherwood JM, Chang J, Overby DR, Yuan F, Gonzalez P, Kopczynski CC, Farsiu S, Stamer WD, 2016. Visualization of conventional outflow tissue responses to netarsudil in living mouse eyes. Eur J Pharmacol 787, 20–31. 10.1016/j.ejphar.2016.04.002. [PubMed: 27085895]

Li G, Schmitt H, Johnson WM, Lee C, Navarro I, Cui J, Fleming T, Gomez-Caraballo M, Elliott MH, Sherwood JM, 2020. Integral role for lysyl oxidase-like-1 in conventional outflow tissue function and behavior. The FASEB Journal 34, 10762–10777. 10.1096/fj.202000702rr. [PubMed: 32623782]

Liu X, Zhao Y, Gao J, Pawlyk B, Starcher B, Spencer JA, Yanagisawa H, Zuo J, Li T, 2004. Elastic fiber homeostasis requires lysyl oxidase–like 1 protein. Nature genetics 36, 178–182. 10.1038/ng1297. [PubMed: 14745449]

Long J, Shelhamer E, Darrell T, 2015. Fully convolutional networks for semantic segmentation, 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 3431–3440. 10.1109/cvpr.2015.7298965.

Loo J, Clemons TE, Chew EY, Friedlander M, Jaffe GJ, Farsiu S, 2020. Beyond Performance Metrics: Automatic Deep Learning Retinal OCT Analysis Reproduces Clinical Trial Outcome. Ophthalmology 127, 793–801. 10.1016/j.ophtha.2019.12.015. [PubMed: 32019699]

Loo J, Fang L, Cunefare D, Jaffe GJ, Farsiu S, 2018. Deep longitudinal transfer learning-based automatic segmentation of photoreceptor ellipsoid zone defects on optical coherence tomography images of macular telangiectasia type 2. Biomedical Optics Express 9, 2681–2698. 10.1364/BOE.9.002681. [PubMed: 30258683]

Loshchilov I, Hutter F, 2017. SGDR: Stochastic Gradient Descent with Warm Restarts, International Conference on Learning Representations (ICLR), Toulon, France.

Mariampillai A, Standish BA, Moriyama EH, Khurana M, Munce NR, Leung MK, Jiang J, Cable A, Wilson BC, Vitkin IA, Yang VX, 2008. Speckle variance detection of microvasculature using swept-source optical coherence tomography. Opt Lett 33, 1530–1532. 10.1364/ol.33.001530. [PubMed: 18594688]

Millar JC, Clark AF, Pang IH, 2011. Assessment of aqueous humor dynamics in the mouse by a novel method of constant-flow infusion. Invest Ophthalmol Vis Sci 52, 685–694. 10.1167/iovs.10-6069. [PubMed: 20861483]

Milletari F, Navab N, Ahmadi S-A, 2016. V-Net: Fully Convolutional Neural Networks for Volumetric Medical Image Segmentation, 2016 Fourth International Conference on 3D Vision (3DV), pp. 565–571. 10.1109/3dv.2016.79.

Oktay O, Schlemper J, Folgoc LL, Lee M, Heinrich M, Misawa K, Mori K, McDonagh S, Hammerla NY, Kainz B, 2018. Attention U-Net: Learning where to look for the pancreas, pp. 1–10.

Paszke A, Gross S, Massa F, Lerer A, Bradbury J, Chanan G, Killeen T, Lin Z, Gimelshein N, Antiga L, 2019. PyTorch: an imperative style, high-performance deep learning library, pp. 8026–8037.

Paulaviciute-Baikstiene D, Vaiciuliene R, Jasinskas V, Januleviciene I, 2016. Evaluation of Outflow Structures In Vivo after the Phacocanaloplasty. J Ophthalmol 2016, 4519846. 10.1155/2016/4519846. [PubMed: 27516898]

Ronneberger O, Fischer P, Brox T, 2015. U-Net: Convolutional Networks for Biomedical Image Segmentation, Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015, pp. 234–241. 10.1007/978-3-319-24574-4_28.

Sarunic MV, Asrani S, Izatt JA, 2008. Imaging the ocular anterior segment with real-time, full-range Fourier-domain optical coherence tomography. Arch Ophthalmol 126, 537–542. 10.1001/archopht.126.4.537. [PubMed: 18413525]

Soille P, 2013. Morphological image analysis: principles and applications. Springer Science & Business Media, Berlin, Germany.

Stein DM, Ishikawa H, Hariprasad R, Wollstein G, Noecker RJ, Fujimoto JG, Schuman JS, 2006. A new quality assessment parameter for optical coherence tomography. Br J Ophthalmol 90, 186–190. 10.1136/bjo.2004.059824. [PubMed: 16424531]

Sutskever I, Martens J, Dahl G, Hinton G, 2013. On the importance of initialization and momentum in deep learning. PMLR, pp. 1139–1147.

Thevenaz P, Ruttimann UE, Unser M, 1998. A pyramid approach to subpixel registration based on intensity. IEEE Trans Image Process 7, 27–41. 10.1109/83.650848. [PubMed: 18267377]

Tom M, Ramakrishnan V, van Oterendorp C, Deserno TM, 2015. Automated detection of Schlemm's canal in spectral-domain optical coherence tomography. SPIE, pp. 744–752. 10.1117/12.2082513.

Usui T, Tomidokoro A, Mishima K, Mataki N, Mayama C, Honda N, Amano S, Araie M, 2011. Identification of Schlemm's canal and its surrounding tissues by anterior segment fourier domain optical coherence tomography. Invest Ophthalmol Vis Sci 52, 6934–6939. 10.1167/iovs.10-7009. [PubMed: 21757587]

Venhuizen FG, van Ginneken B, Liefers B, van Grinsven M, Fauser S, Hoyng C, Theelen T, Sanchez CI, 2017. Robust total retina thickness segmentation in optical coherence tomography images using convolutional neural networks. Biomed Opt Express 8, 3292–3316. 10.1364/BOE.8.003292. [PubMed: 28717568]

Wang F, Shi G, Li X, Lu J, Ding Z, Sun X, Jiang C, Zhang Y, 2012. Comparison of Schlemm's canal's biological parameters in primary open-angle glaucoma and normal human eyes with swept source optical. J Biomed Opt 17, 116008. 10.1117/1.JBO.17.11.116008. [PubMed: 23117803]

Wang K, Johnstone MA, Xin C, Song S, Padilla S, Vranka JA, Acott TS, Zhou K, Schwaner SA, Wang RK, Sulchek T, Ethier CR, 2017. Estimating Human Trabecular Meshwork Stiffness by Numerical Modeling and Advanced OCT Imaging. Invest Ophthalmol Vis Sci 58, 4809–4817. 10.1167/iovs.17-22175. [PubMed: 28973327]

Wang X, Zhai Y, Liu X, Zhu W, Gao J, 2020. Level-Set Method for Image Analysis of Schlemm's Canal and Trabecular Meshwork. Transl Vis Sci Technol 9, 7. 10.1167/tvst.9.10.7.

Xiao S, Bucher F, Wu Y, Rokem A, Lee CS, Marra KV, Fallon R, Diaz-Aguilar S, Aguilar E, Friedlander M, Lee AY, 2017. Fully automated, deep learning segmentation of oxygen-induced retinopathy images. JCI Insight 2. 10.1172/jci.insight.97585.

Xin C, Johnstone M, Wang N, Wang RK, 2016. OCT Study of Mechanical Properties Associated with Trabecular Meshwork and Collector Channel Motion in Human Eyes. PLoS One 11, e0162048. 10.1371/journal.pone.0162048. [PubMed: 27598990]

Yao X, Tan B, Ho Y, Liu X, Wong D, Chua J, Wong TT, Perera S, Ang M, Werkmeister RM, Schmetterer L, 2021. Full circumferential morphological analysis of Schlemm's canal in human eyes using megahertz swept source OCT. Biomedical Optics Express 12, 3865–3877. 10.1364/BOE.426218. [PubMed: 34457385]

Zhang X, Beckmann L, Miller DA, Shao G, Cai Z, Sun C, Sheibani N, Liu X, Schuman J, Johnson M, Kume T, Zhang HF, 2020. In Vivo Imaging of Schlemm's Canal and Limbal Vascular Network in Mouse Using Visible-Light OCT. Invest Ophthalmol Vis Sci 61, 23. 10.1167/iovs.61.2.23.

Zhang Z, Liu Q, Wang Y, 2018. Road Extraction by Deep Residual U-Net. IEEE Geoscience and Remote Sensing Letters 15, 749–753. 10.1109/lgrs.2018.2802944.

Zhao Z, Zhu X, He W, Jiang C, Lu Y, 2016. Schlemm's Canal Expansion After Uncomplicated Phacoemulsification Surgery: An Optical Coherence Tomography Study. Invest Ophthalmol Vis Sci 57, 6507–6512. 10.1167/iovs.16-20583. [PubMed: 27918824]
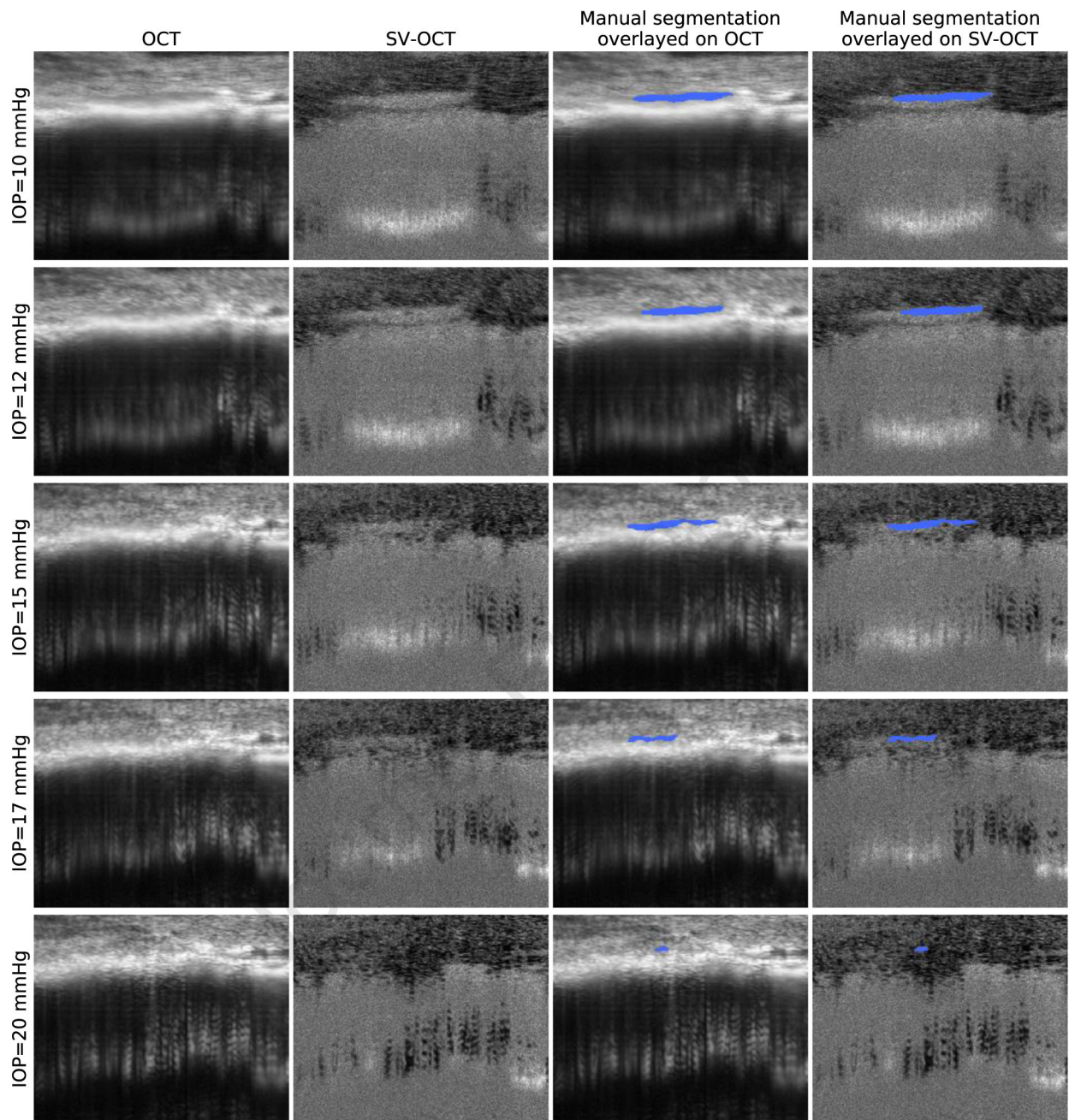
Figure 1.
Visualization and manual segmentation of SC in OCT images of a living mouse eye. Each row represents OCT imaging data and corresponding overlayed ground truth annotations acquired on the same eye at a different IOP level (from top to bottom: 10, 12, 15, 17, 20 mmHg). Each column from left to right represents the intensity OCT image, the SV-OCT image, the ground truth overlayed on the intensity OCT image, and the ground truth overlayed on the SV-OCT image.
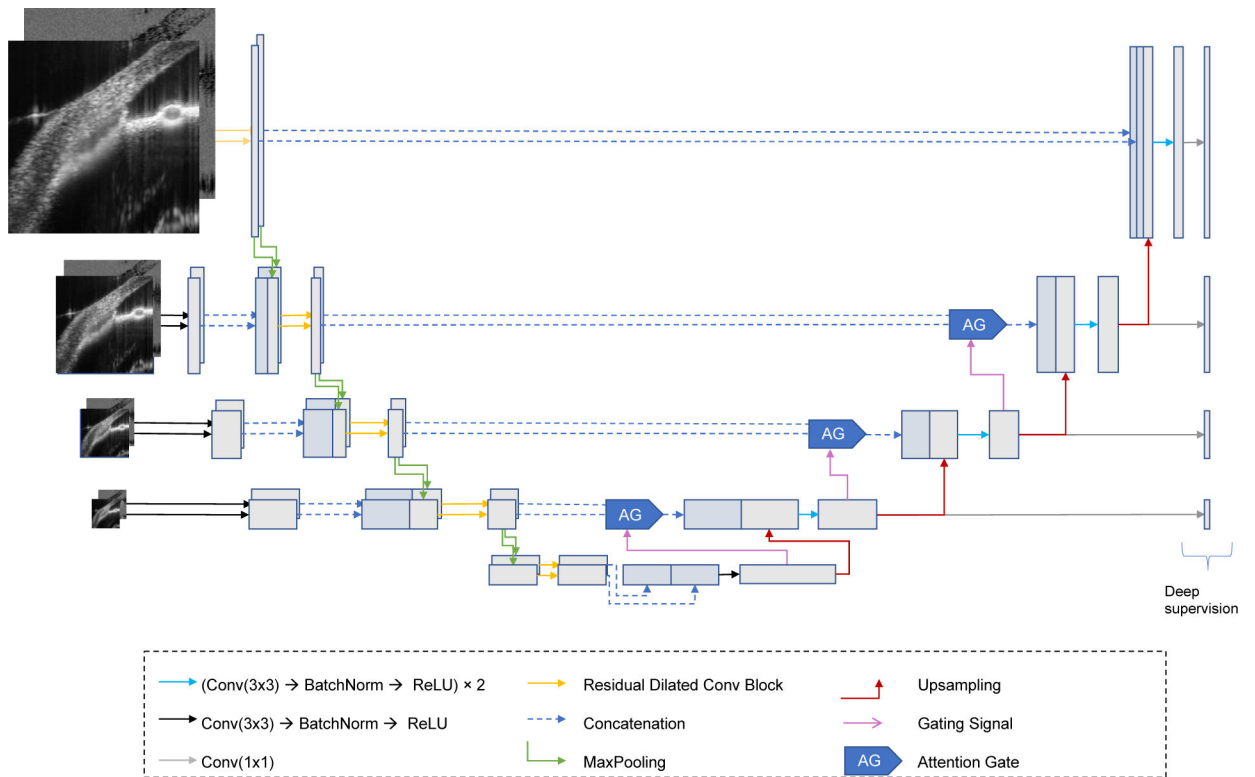
Figure 2.
Proposed CNN Model: LF-DRAG-UNet. The network is trained on OCT and SV-OCT images using separate encoder branches that are fused downstream. A multi-scale input image pyramid and dilated residual convolutions are used to provide more global information. Attention gating is used to prune irrelevant regions of the feature maps.
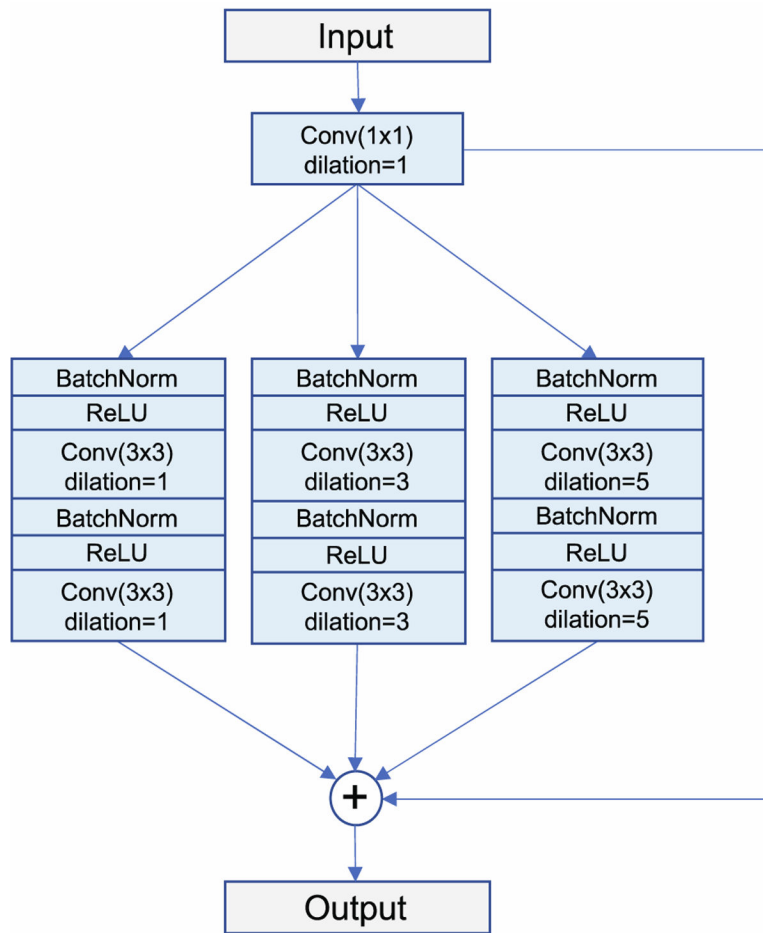
Figure 3.
Structure of the dilated residual convolution block, the encoder sub-unit in LF-DRAG-UNet. Three parallel branches containing atrous convolutions with dilation rates of $d = 1$, $d = 3$, and $d = 5$ are used to expand the receptive field of the convolution layers. The outputs of all branches are added to the input for residual learning.
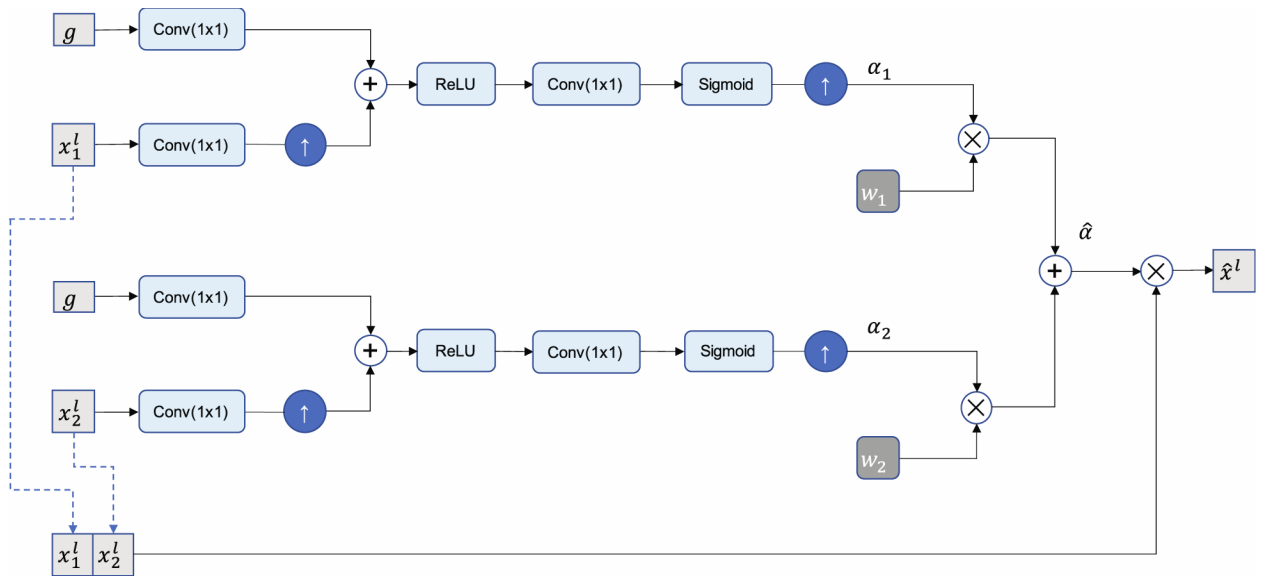
Figure 4.
Diagram of the late fusion attention gate. Feature maps x^l are scaled by attention coefficients $\hat{\alpha}$, which are obtained by taking the weighted sum of attention coefficients learned for OCT and SV-OCT inputs. A gating signal, g, obtained from a coarser resolution, is used to determine salient regions and prune less relevant features.
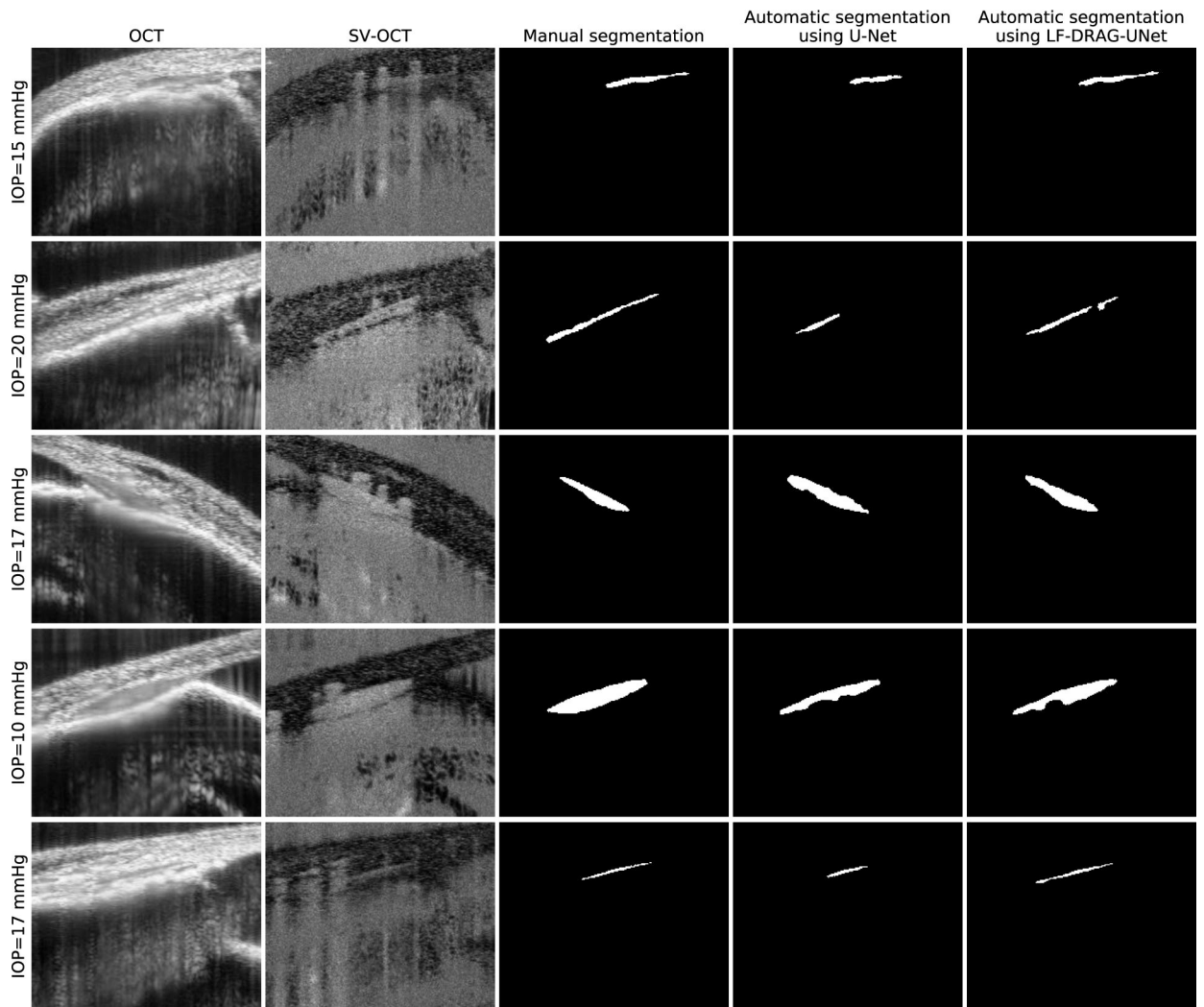
Figure 5.
Illustrative example showing the superior performance of the proposed LF-DRAG-UNet compared to baseline U-Net. Columns from left to right depict the intensity OCT image, SV-OCT image, ground truth obtained by manual segmentation, segmentation obtained by the baseline U-Net, and segmentation obtained by LF-DRAG-UNet.
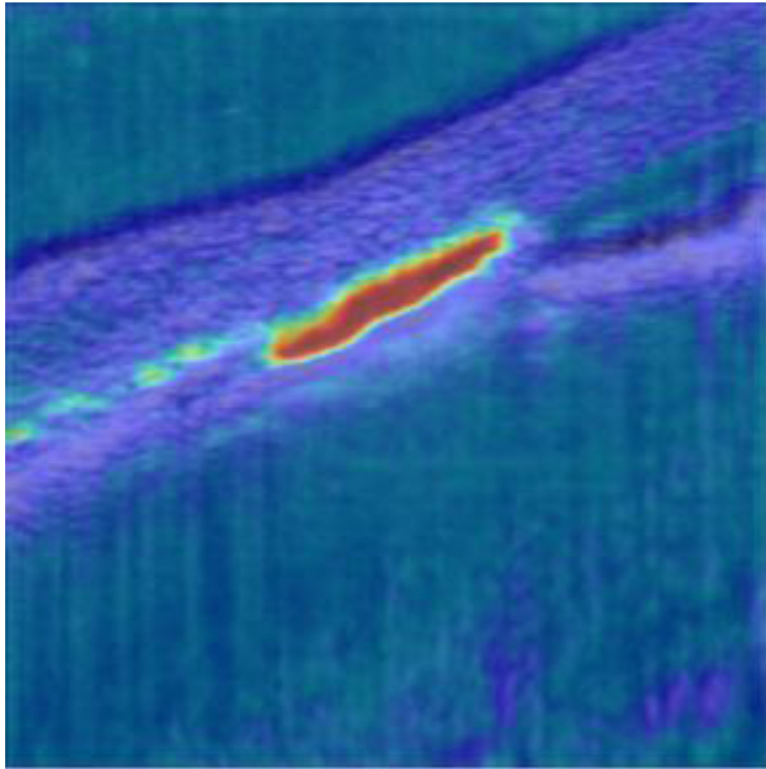
Figure 6.
Attention coefficient heatmap for an OCT image of a mouse eye at $IOP = 12\ mmHg$. The attention gate was able to focus in on the region of the image corresponding to the SC lumen and suppressed regions in the background.
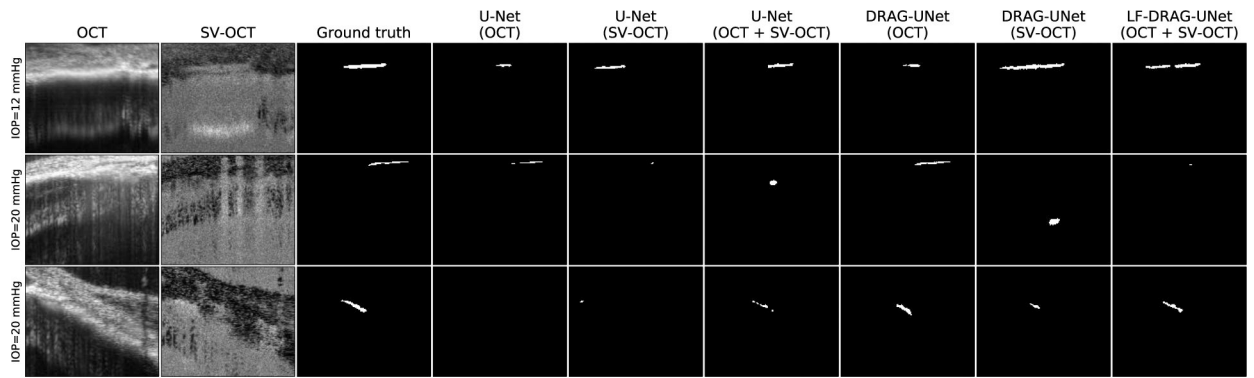
Figure 7.
Comparison of segmentations obtained by proposed and alternative networks trained using intensity OCT images, SV-OCT images, and both input images. Row 1 depicts that when SC is difficult to see in the intensity image but visible in the SV-OCT image, networks using SV-OCT images performed well. Row 2 depicts that when SC is not visible in the SV-OCT image, networks using SV-OCT images performed worse than those using only intensity images. Row 3 depicts when OCT and SV-OCT images provided complementary information, networks that used both inputs had superior results.
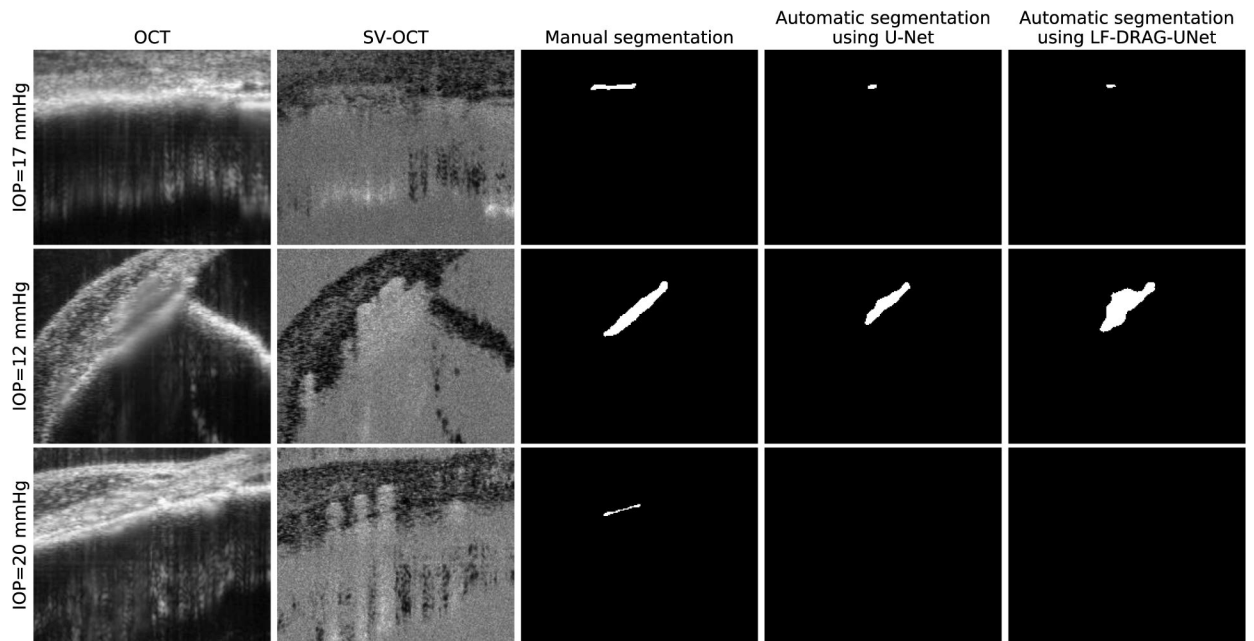
Figure 8.
Limitations of CNNs for automatic SC segmentation. Columns from left to right represent the intensity OCT image, SV-OCT image, ground truth obtained by manual segmentation, segmentation obtained by U-Net, and segmentation obtained by the proposed LF-DRAG-UNet model. Low contrast, shadowing due to superficial blood vessels, and small SC size can lead to incorrect segmentations.
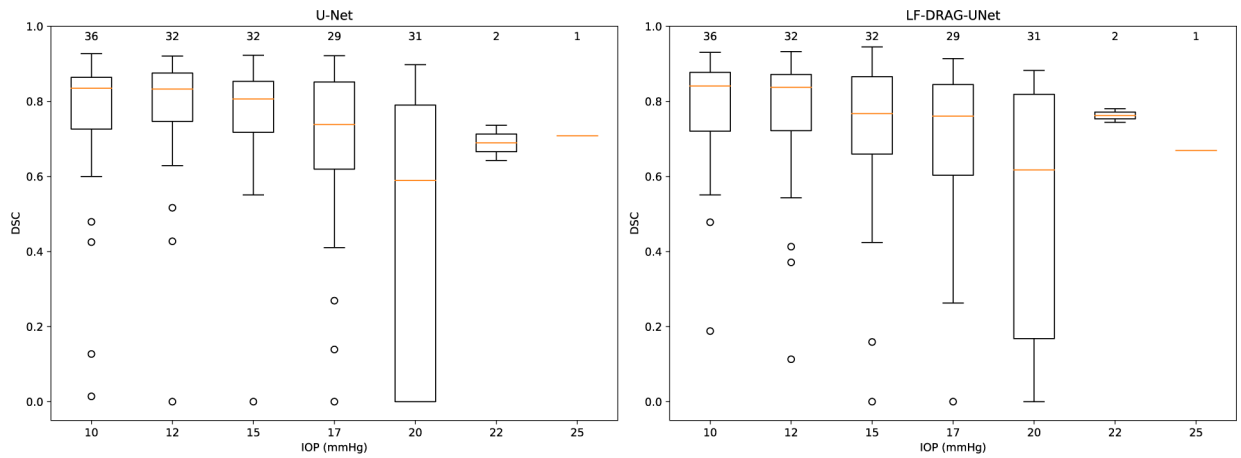
Figure 9.

Boxplots showing DSC for samples at different IOP. Left plot shows results of the baseline U-Net; right plot shows results for the proposed LF-DRAG-UNet. The following summary statistics are visualized: lower and upper quartile (range of the box); median (horizontal orange line); $1.5 \times$ the interquartile range below and above the lower and upper quartiles, respectively (whiskers). The number above each box denotes the number of samples at the specified IOP. DSC scores were lower for samples at higher IOPs which had SCs with relatively smaller cross-sectional areas. The proposed model was better able to segment samples at higher IOP.

Table 1

Average (mean ± SD) and median segmentation performance of an expert human grader, proposed model, and alternative models evaluated using DSC, precision, and recall.

Models	DSC		Precision		Recall	
	Mean ± SD	Median	Mean ± SD	Median	Mean ± SD	Median
Human grader	0.713 ± 0.209	0.763	0.731 ± 0.212	0.772	0.737 ± 0.250	0.829
U-Net (OCT)	0.677 ± 0.270	0.782	0.688 ± 0.294	0.782	0.710 ± 0.294	0.829
U-Net (SV-OCT)	0.598 ± 0.300	0.724	0.658 ± 0.305	0.775	0.615 ± 0.331	0.763
U-Net (**baseline**) (OCT & SV-OCT)	0.671 ± 0.282	0.774	0.711 ± 0.299	0.830	0.700 ± 0.314	0.832
DRAG-UNet (OCT)	0.686 ± 0.254	0.784	0.726 ± 0.266	0.813	0.719 ± 0.281	0.821
DRAG-UNet (SV-OCT)	0.613 ± 0.294	0.735	0.649 ± 0.306	0.740	0.655 ± 0.339	0.799
LF-DRAG-UNet (**our method**) (OCT & SV-OCT)	0.694 ± 0.256	0.791	0.722 ± 0.264	0.812	0.743 ± 0.294	0.849