optica

Weakly supervised individual ganglion cell segmentation from adaptive optics OCT images for glaucomatous damage assessment: supplement

SOMAYYEH SOLTANIAN-ZADEH,[1] KAZUHIRO KUROKAWA,[2] ZHUOLIN LIU,[3] FURU ZHANG,[3] OSAMAH SAEEDI,[4] DANIEL X. HAMMER,[3] DONALD T. MILLER,[2] AND SINA FARSIU[1,5,*]

[1]*Department of Biomedical Engineering, Duke University, Durham, North Carolina 27708, USA*
[2]*School of Optometry, Indiana University, Bloomington, Indiana 47405, USA*
[3]*Center for Devices and Radiological Health (CDRH), U.S. Food and Drug Administration, Silver Spring, Maryland 20993, USA*
[4]*Department of Ophthalmology and Visual Sciences, University of Maryland Medical Center, Baltimore, Maryland 21201, USA*
[5]*Department of Ophthalmology, Duke University Medical Center, Durham, North Carolina 27710, USA*
*Corresponding author: sina.farsiu@duke.edu

Supplementary Information for

Weakly-Supervised Individual Ganglion Cell Segmentation from Adaptive Optics OCT Images for Glaucomatous Damage Assessment

SOMAYYEH SOLTANIAN-ZADEH,[1] KAZUHIRO KUROKAWA, [2] ZHUOLIN LIU, [3] FURU ZHANG, [3] OSAMAH SAEEDI, [4] DANIEL X. HAMMER, [3] DONALD T. MILLER,[2] AND SINA FARSIU[1,5,*]

[1]Department of Biomedical Engineering, Duke University, Durham, NC 27708, USA
[2]School of Optometry, Indiana University, Bloomington, IN 47405, USA
[3]Center for Devices and Radiological Health (CDRH), U.S. Food and Drug Administration, Silver Spring, MD 20993, USA
[4]Department of Ophthalmology and Visual Sciences, University of Maryland Medical Center, Baltimore, MD 21201, USA
[5]Department of Ophthalmology, Duke University Medical Center, Durham, NC 27710, USA
*Corresponding author: sina.farsiu@duke.edu

This PDF file includes:

        Supplementary Text
        Supplementary Methods
        Supplementary Figures
        Supplementary Tables
        Supplementary Video Captions
        References

Supplementary Text

1. The Ganglion Cell Layer

In the human retina, GCL somas reach a maximum density and cell stack depth near 3°-4.5°, where these GCs project to the densely packed cone photoreceptors at the foveal center [1, 2]. At increasing eccentricities from this peak, the cell density and GCL thickness monotonically decrease, with GCs eventually arranged in a monolayer around 12°-13° [1, 2]. The GC soma size also varies with retinal eccentricity, with some GC types varying more than others. The two primary subtypes of GCs are the midget and parasol cells; parasol GC somas are generally larger than midget GC somas (e.g., at 12°-13°) but are increasingly similar and smaller in size closer to the fovea. At 3°, the two GC types are effectively identical in size and thus, indistinguishable based on size [3, 4].

2. Ophthalmic Examination and Glaucoma Diagnosis

All participants underwent a complete ophthalmological examination. The exam included the measurement of IOP, a slit lamp examination, dilated fundus examination, determination of axial length with biometry (IOLMaster, Zeiss and Lenstar, Haag Streit), and OCT imaging (Heidelberg Spectralis, Heidelberg, Germany) of the peripapillary retinal NFL and macula. The glaucoma subjects were examined by a glaucoma subspecialist and included Humphrey 24-2 and 10-2 visual field (VF) tests (Humphrey Field Analyzer, Carl Zeiss Meditec Inc.). Full clinical records and OCT imaging data were examined to confirm the diagnosis of glaucoma. Automatic OCT retinal layer segmentation of GCL was confirmed and compared to AO imaging locations accounting for the raphe angle. To determine the total deviation (TD) and pattern deviation (PD) values in the areas corresponding to AO imaged locations, VF 24-2 map was used with accounting for GC displacement [5]. TD and PD values represent the local loss in sensitivity and focal depressed areas compared with age-matched controls, respectively. To remove potential bias in analysis, one subject was omitted from the structure-function study because imaging of this subject was done with a different imaging instrument (Optovue, Fremont, CA, USA), which was different than the predefined protocol used for all other subjects. The clinical OCT and VF measurements are summarized in Table S2. All glaucoma patients included in this work were under treatment to control IOP.

Supplementary Methods

1. Retinal Layer Segmentation

Using graph theory and dynamic programming (GTDP), we sequentially segmented retinal layer boundaries from the contrast-enhanced median B-scan image (denoted as the *input image* in Fig. S1) by limiting the search region using the previously identified layer. The schematics in Supplementary Fig. 1 outline the layer segmentation steps. First, we set the initial search region (R_0) to be the upper half of the image based on the prior knowledge that all layers above the inner nuclear layer (INL) are in this region. Specifically, R_0 is a binary mask with only the upper half set to 1. In a case in which the outer retina was cropped out from the AO-OCT volume during the manual grading process, R_0 was set to 1 for the entire image. We then segmented the vitreous-nerve fiber layer (NFL). Since the inner plexiform layer (IPL)-INL boundary is generally stronger than the ganglion cell layer (GCL)-IPL, we identified the former prior to segmentation of the latter. While the vitreous-NFL is easily discernable due to its strong hyper-reflectivity, IPL-INL and GCL-IPL are not as prominent. To accurately segment these two boundaries, we applied a bilateral filter with a photometric spread of $\sigma_r = 1$ and geometric spread of $\sigma_d = (12, 1.5)$ (horizontal and vertical directions, respectively) to primarily smooth the input image in the horizontal direction while preserving the edges. We then generated a binary mask to narrow the initial search region, R_0, for a more accurate identification of IPL-INL. We removed the region within W μm of the vitreous-NFL boundary (defined in (1)) from R_0 to get the final search region R_1 for IPL-INL. After identifying the IPL-INL with GTDP, we constructed the search region for GCL-IPL (R_2). We defined R_2 to be between S μm below the vitreous-NFL (defined in (2)) and 10 pixels above the IPL-INL. Based on these boundaries, we extracted volumes extending from the vitreous-NFL to the GCL-IPL for further analysis. To avoid missing sparsely scattered somas at the GCL-IPL boundary, we kept $N = 10$ additional C-scans below this boundary. We empirically set the parameters as

$$W = \begin{cases} \frac{3}{4}Y, & \text{Healthy} \\ \frac{3}{8}Y, & \text{Glaucoma} \end{cases}, \tag{1}$$

and

$$S = \begin{cases} \frac{1}{3}Y, & \text{Healthy} \\ \frac{1}{10}Y, & \text{Glaucoma} \end{cases}, \tag{2}$$

where Y was selected to be 94 μm, 70 μm, and 55 μm for 3.75°, 8.5°, and 12°-12.75° recordings, respectively.

2. Ablation Experiments

A. Network Architecture

We investigated the effect of using residual connections in each level of the encoder path and the use of convolutional downsampling instead of maxpooling. As the precision-recall curves in Fig. S4 show, WeakGCSeg, which uses both residual connections and convolutional downsampling, outperformed the other variations, especially for the IU 3.75° data and images from the FDA system (based on the average precision scores; Table S7).

B. Post-processing

To investigate the effectiveness of our post-processing approach for soma segmentation, we conducted the following experiments:

1. Comparison to two other variations. These included simple thresholding of the CNN output (the soma probability map) and using the soma-class feature map (features before the Softmax layer, normalized to the range 0 and 1) for post-processing instead of the probability map. As the results in Fig. S5(A) show, simple thresholding of the probability map yielded soma masks that were significantly smaller than previously reported soma sizes, even across different threshold levels (Fig. S5(B)). Using the feature map or the probability map as the input to our segmentation pipeline yielded similar estimates of soma diameters that were in line with previously reported values. As Fig. S5(B) shows, the estimates with these two maps converged to approximately the same level at high threshold values.

2. Modification of watershed masks with thresholding. In addition to the above experiment, we analyzed the effectiveness of removing voxels from the set of watershed masks in the final thresholding step (parameter *TH* explained in Section 2.B.3 of main text). For this, we compared the diameter estimates calculated from the automatic segmentation masks ($d_{automatic}$) to the soma diameters obtained from the manually segmented cells (d_{manual}; three subjects, 300-340 somas from each subject). The results showed that the initial watershed masks (without thresholding) yielded $d_{manual}-d_{automatic}$ = -0.795 µm (averaged over all subjects and locations), whereas thresholding reduced this difference to -0.005 µm. This shows that watershed alone gave masks that were on average bigger than the underlying somas.

3. Effect of Gaussian smoothing. Finally, we quantified the effect of the Gaussian smoothing step on the estimated soma diameters. For this experiment, we changed the standard deviation (σ) of the Gaussian filter in the lateral plane and compared the estimated diameters to d_{manual}. We split the somas to small (diameter < 15 µm) and large cells (diameter ≥ 15 µm) to separately quantify the effect of smoothing on the diameter estimates. The results in Fig. S5(C) show that larger smoothing factors decrease the difference between manual and automatic diameter estimates in large somas. For small somas, the effect was the opposite. Based on the average absolute diameter differences calculated over all somas, the results indicate that on average, smaller smoothing factors of $\sigma < 1$ yield diameter estimates closer to d_{manual}.

Supplementary Figures
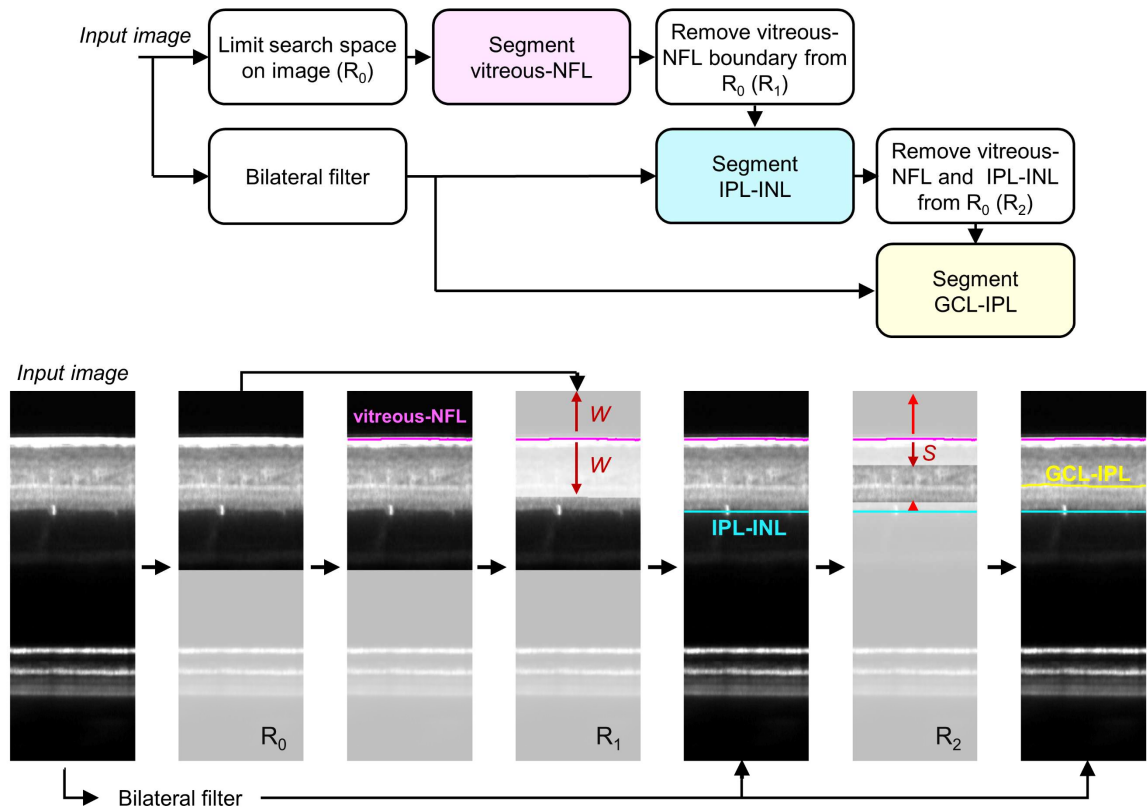
Fig. S1| The pre-processing step involved sequential boundary detection to extract the NFL+GCL from AO-OCT volumes. (*Top*) Overview of the layer segmentation steps applied to the contrast-enhanced median B-scan image, denoted as the *input image*, and *(bottom)* illustrative examples for each step of the pipeline. The less transparent areas in R_0, R_1, and R_2 denote regions excluded from the boundary search space. NFL: nerve fiber layer, IPL: inner plexiform layer, INL: inner nuclear layer. W and S are defined in Eqs. (1)-(2).

(A)

Zoom

(B)

Zoom

(C)

Zoom
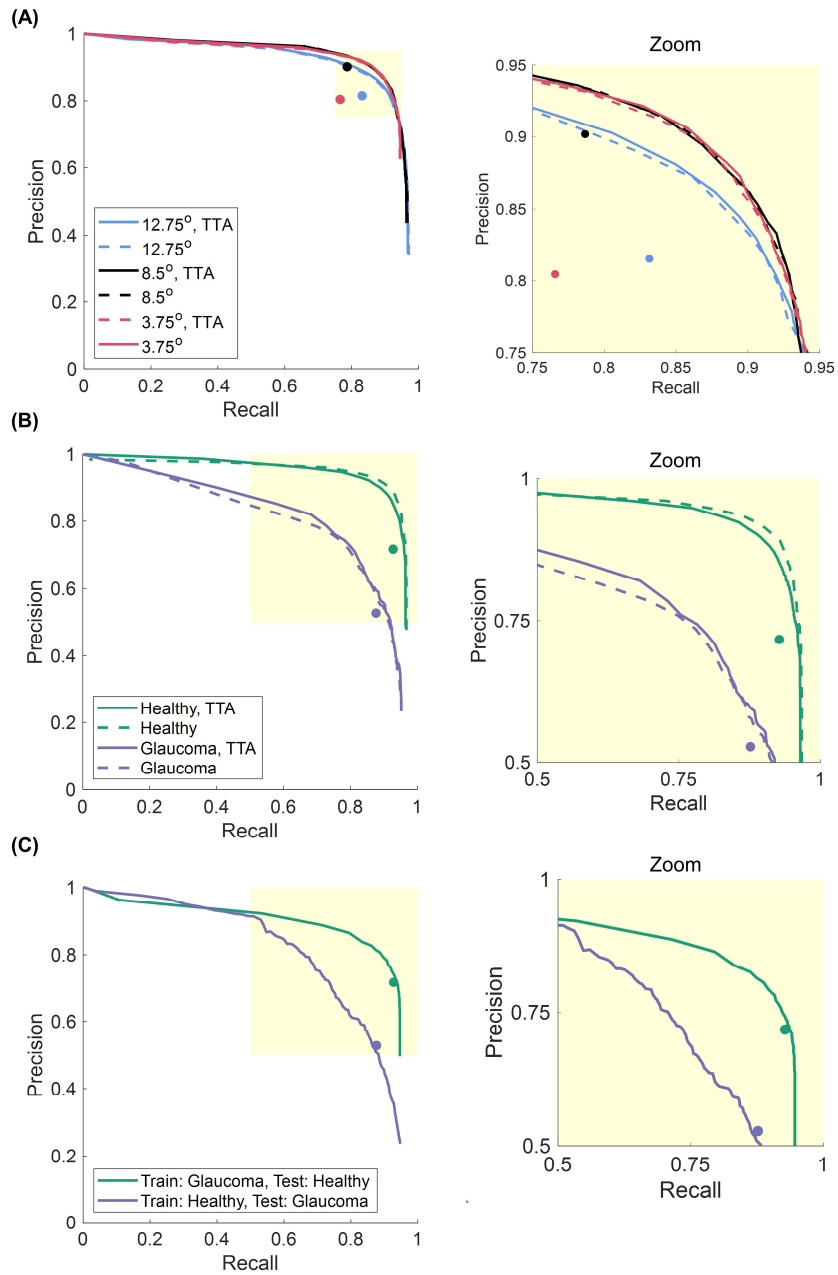
Fig. S2| Overall detection performance of the method compared to expert grader under different conditions. Effect of test-time-augmentation (TTA) on our method's average precision-recall curves for the **(A)** IU and **(B)** FDA datasets. **(C)** Average precision-recall curves of our method trained on one group of subjects and tested on the other. The average of expert graders' performance is shown with circle markers. Precision-recall curves for each trained network were created by changing the probability threshold from 0 to 1. The average of all curves at each threshold value yielded the final average precision-recall curve.

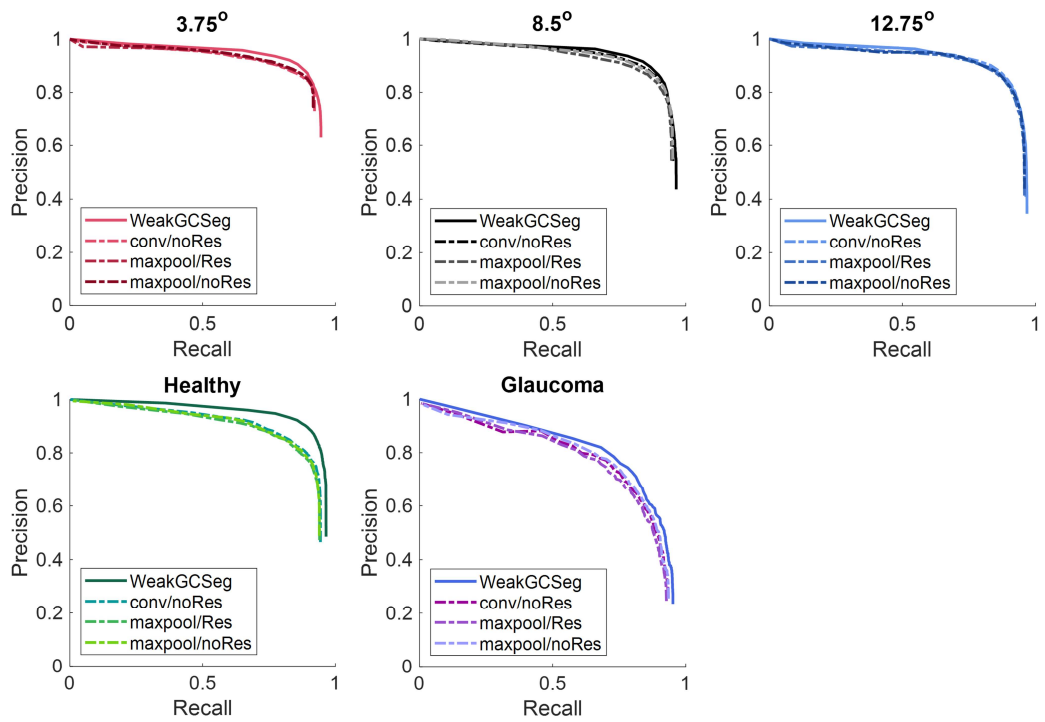**Fig. S3**| Mean precision-recall curves for the ablation experiment testing different variations of WeakGCSeg with convolutional or maxpooling downsampling (conv or maxpool) and with or without the residual connections (Res or noRes).
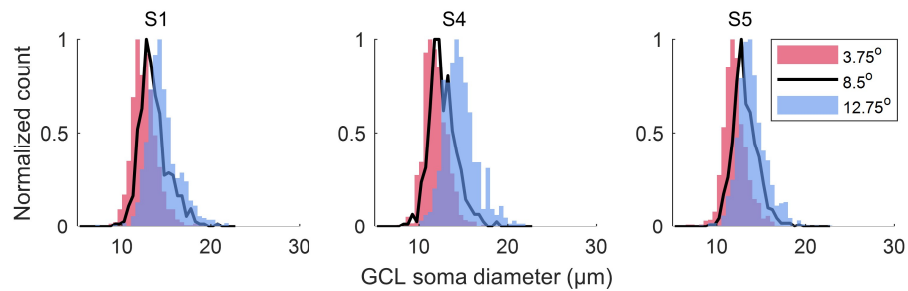

Fig. S4| Soma size distributions automatically determined by WeakGCSeg for three healthy subjects across different retinal locations.
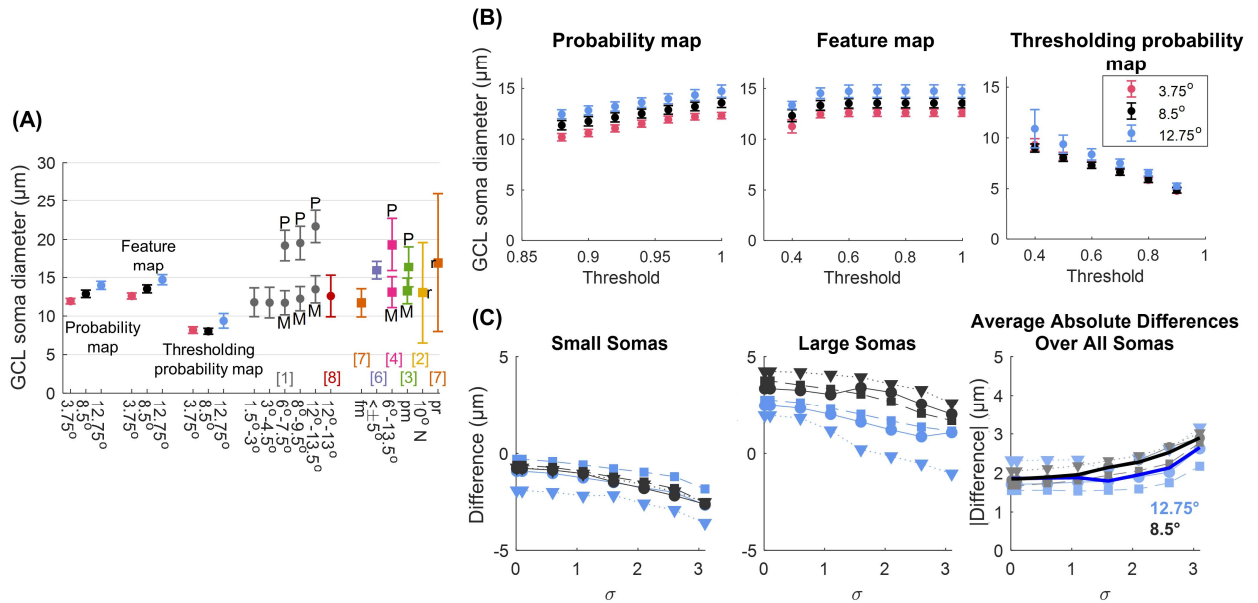
Fig. S5| Results of the ablation study for the post-processing step. (A) Comparison of previously reported GCL soma diameters with the estimated values calculated through direct thresholding the CNN output probability map (threshold = 0.5), post-processing the CNN soma feature map (threshold = 0.6), or post-processing the probability map (threshold = 0.96). **(B)** Effect of threshold on the estimated soma diameters for each segmentation approach. For our post-processing approach, the threshold refers to the value in the 3D watershed step that determines the background pixels. **(C)** Effect of Gaussian smoothing on the difference between soma diameter measurements from manual and automatic segmentation masks (d_{manual} - $d_{automatic}$). We split the somas to two groups based on their diameters (Small somas: diameter < 15 µm, and Large somas: diameter ≥ 15 µm). X-axis represents the standard deviation of the Gaussian filter (σ) in the lateral plane. Each marker shape represents data from a different subject. The thicker lines in the right panel show the average result over all subjects.
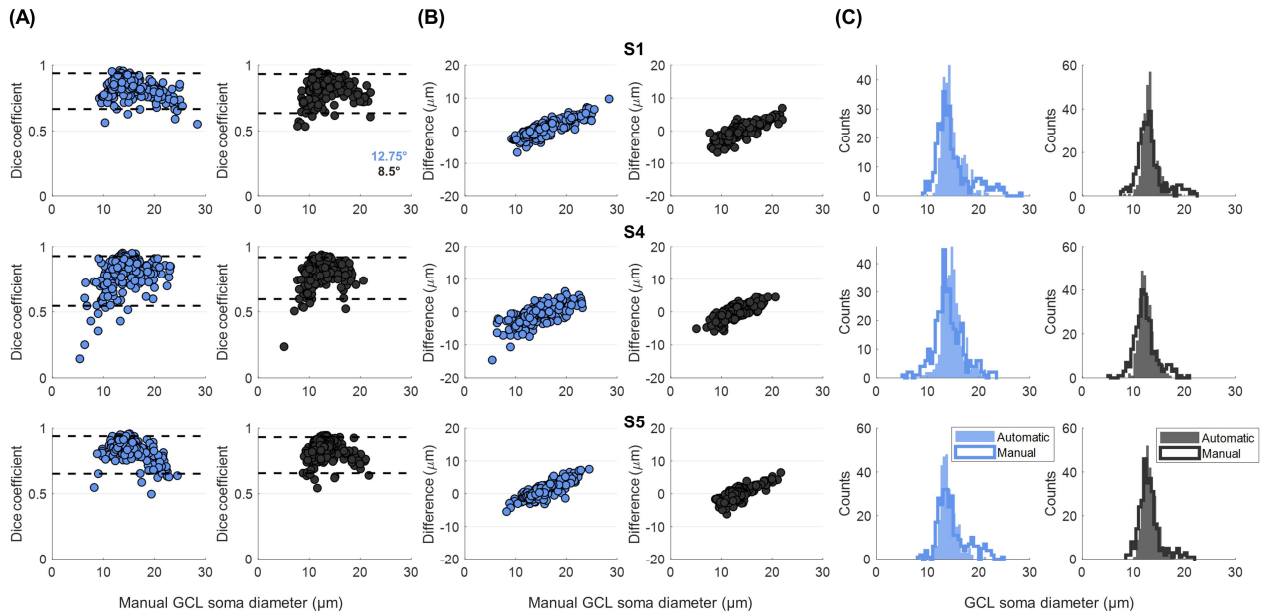
Fig. S6| Comparison between automatic and manual GCL soma segmentations at 8.5° (black) and 12.75° (blue) locations. (A) Dice similarity coefficients between the automatic and manual *en face* segmentation masks. **(B)** The difference in soma diameter measurements defined as d_{manual} - $d_{automatic}$, where d_{manual} and $d_{automatic}$ correspond to diameter estimations using the manual and automatic segmentation masks, respectively. **(C)** Histogram of measured diameters. Markers in **(A-B)** denote individual somas (*n* = 300-343 somas). Dice similarity coefficients and soma diameters were calculated based on the binary masks at the *en face* plane where each soma's manually marked center was located. Dashed black lines in **A** are the 95% data intervals. Rows in panels **A-C** correspond to the results obtained for subjects S1, S4, and S5, respectively.

(A)

(B)
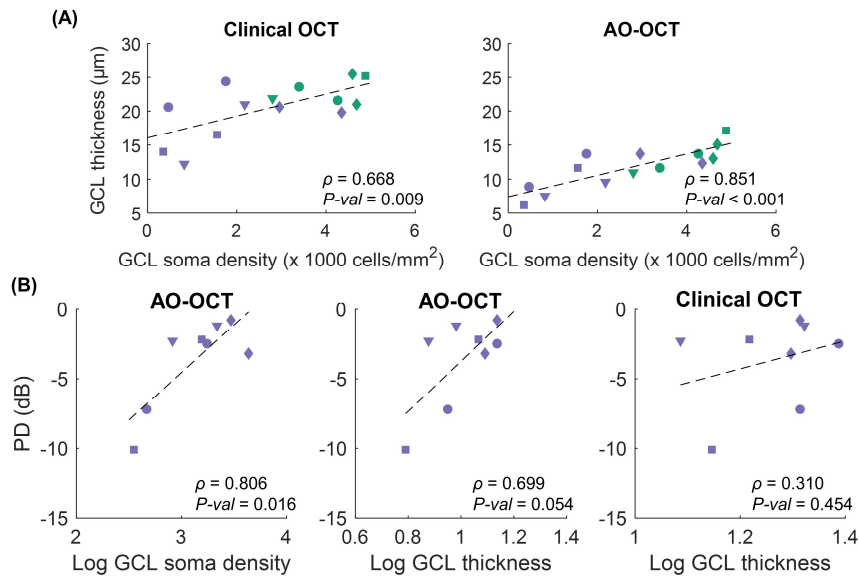
Fig. S7| Structural and functional characteristics of glaucomatous eyes compared to controls. (A) GCL thickness measured with clinical OCT and AO-OCT versus soma densities for 4 healthy and 4 glaucoma subjects. **(B)** PD measurements versus cell densities and GCL thickness values for 4 glaucoma subjects. ρ: Pearson correlation coefficient. Each subject is shown with a different marker shape.

Fig. S8| Mean precision-recall curves for localizing GCL somas using different neural networks.

Supplementary Tables

Table S1. Subject information for the IU dataset and expert grader (Gr) assignments for creating the gold-standard ground truth (GT) and the 2nd Grading markings. AL: axial length.

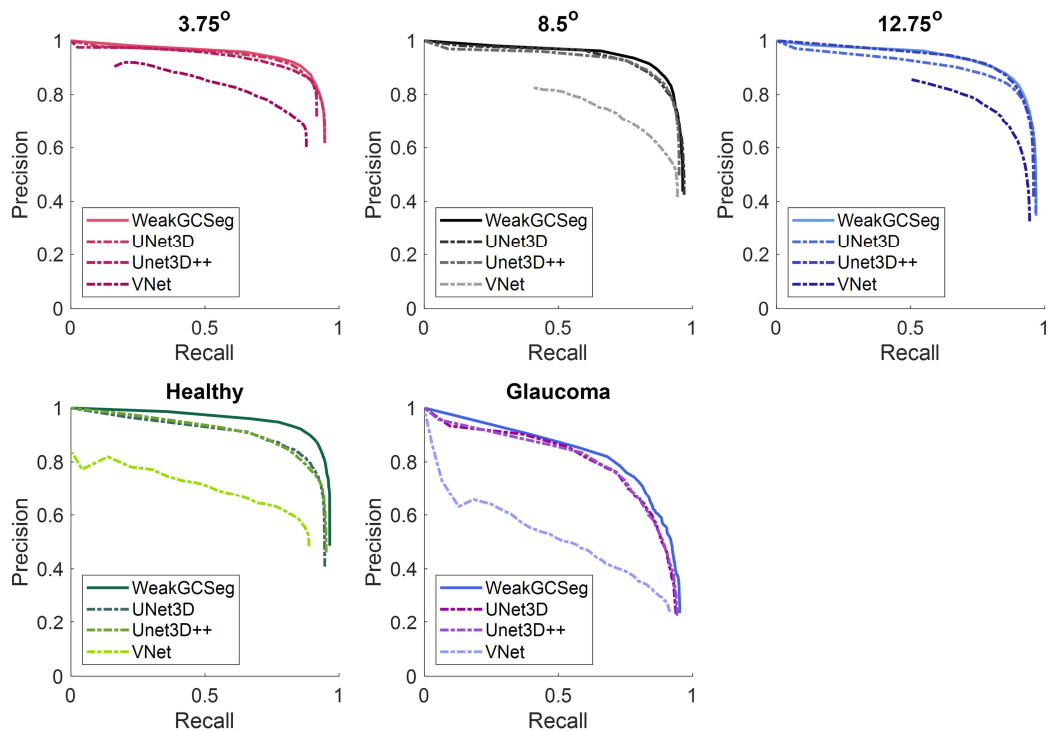| Subject	General			Marking	Retinal Eccentricity		
	Gender	Age	AL [mm]		3.75°	8.5°	12.75°
S1	M	50	25.40	GT	Gr1/Gr2	Gr6/Gr8	Gr1/Gr2
				2nd Grading	Gr7	Gr9	Gr7
S2	M	26	24.04	GT	Gr1/Gr2	Gr8	Gr1/Gr2
				2nd Grading	Gr6	Gr9	Gr6
S3	F	47	23.27	GT	Gr1/Gr2	Gr3/Gr6	Gr1/Gr2
				2nd Grading	Gr6	Gr9	Gr8
S4	M	24	23.98	GT	Gr1/Gr2	Gr6	Gr1/Gr2
				2nd Grading	Gr8	Gr9	Gr8
S5	M	26	25.24	GT	Gr3/Gr6	Gr6	Gr5/Gr6
				2nd Grading	Gr7	Gr9	Gr7
S6	M	31	24.75	GT	Gr5/Gr6	-	Gr7/Gr6
				2nd Grading	Gr8	-	Gr8
S7	M	22	23.63	GT	Gr4/Gr2	-	Gr6/Gr8
				2nd Grading	Gr8	-	Gr7
S8	M	33	26.44	GT	Gr7/Gr8	-	Gr7/Gr8
				2nd Grading	Gr6	-	Gr6

Table S2. Clinical eye exam data for the glaucoma (G) and healthy (H) subjects from the FDA dataset.

Subject	General							OCT Macula Thickness [μm]		Visual Field [dB]	
	Gender	Age	AL [mm]	C:D	IOP [mmHg]	HF Defect	Severity	Location	GCL	TD	PD
S₁₋G	F	62	23.64	0.7	11	I	E	12T2.5S	24.4	-2.4	-2.5
								12T2.5I	20.6	-6.3	-7.2
S₂₋G	F	52	25.98	0.3	14	S	E	12T2.5S	12.2	-1.4	-2.3
								12T2.5I	21.0	-0.6	-1.2
S₃₋G	F	57	24.06	0.6	15	I	E	12T2.5S	20.6	-0.7	-0.8
								12T2.5I	19.8	-2.3	-3.2
S₄₋G	M	54	23.09	0.8	14	I	M	12T2.5S	16.5	-6.2	-2.2
								12T2.5I	14.0	-13.7	-10.1
S₁₋H	M	50	26.08	0.5	14	-	-	12T2.5S	21.6	-	-
								12T2.5I	23.6	-	-
S₂₋H	M	58	23.28	0.4	11	-	-	12T2.5S	21.9	-	-
								12T2.5I	-	-	-
S₃₋H	F	54	24.9	0.5	22	-	-	12T2.5S	21.0	-	-
								12T2.5I	25.5	-	-
S₄₋H	M	67	24.91	0.4	20	-	-	12T2.5S	-	-	-
								12T2.5I	25.2	-	-

Notes: 1. General: AL – axial length, C:D – cup-to-disc ratio, IOP – intraocular pressure, HF –hemifield (S or I), Severity – Early (E), Moderate (M), or Severe (S). All glaucoma subjects were under treatment to control IOP.

2. OCT Macula: values were interpolated from Heidelberg maps, including angle of raphe.

3. Visual Field (24-2 algorithm): TD – total deviation, PD – pattern deviation; values were interpolated from VF TD maps accounting for GC-VF displacement and flipped (S-I) to account for corresponding anatomical region.

Table S3. Number of training and test data along with the number of manually labeled cells for each experiment. For training and test data, the numbers denote subjects:volumes, whereas for the cells the numbers denote the average (std).

Exp.	Imager (train/test)	Subject	Training samples	Training cells	Testing samples	Total testing samples
1	IU/IU	Healthy	7:14	13,345±399	1:2	8:16
2	IU/IU	Healthy	7:14	13,345±399	1:1	5:5
3	FDA/FDA	Healthy	3:4-5	2,296±431	1:1-2	4:6
		Glaucoma	4:8	1,538±282	1:2	5:10
4	IU/FDA	Healthy	8:16	15,251	4:6	4:6
	FDA/IU	Healthy	4:6	3,061	8:16	8:16
	IU+FDA/IU+FDA	Healthy	9:16-17	15,641±830	3:6-5	12:22

Table S4. Effect of intensity normalization and test-time-augmentation on detection performance for the IU dataset. Scores are reported as mean ± standard deviation for F_1 (recall, precision), calculated across $n = 8$ subjects for 3.75° and 12.75° and $n = 5$ subjects for the 8.5° locations.

Intensity normalization	Eccentricity	Test-time-augmentation	
		Yes	No
Identity	3.75°	0.87 ± 0.04 (0.88 ± 0.09, 0.87 ± 0.06)	0.87 ± 0.04 (0.87 ± 0.10, 0.88 ± 0.05)
	8.5°	0.87 ± 0.02 (0.90 ± 0.05, 0.85 ± 0.04)	0.88 ± 0.01 (0.89 ± 0.04, 0.87 ± 0.02)
	12.75°	0.87 ± 0.04 (0.88 ± 0.07, 0.85 ± 0.06)	0.86 ± 0.05 (0.88 ± 0.06, 0.85 ± 0.07)
Whiten	3.75°	0.85 ± 0.06 (0.85 ± 0.14, 0.87 ± 0.07)	0.86 ± 0.05 (0.86 ± 0.13, 0.88 ± 0.05)
	8.5°	0.88 ± 0.02 (0.89 ± 0.03, 0.86 ± 0.02)	0.88 ± 0.01 (0.90 ± 0.04, 0.86 ± 0.02)
	12.75°	0.85 ± 0.06 (0.86 ± 0.08, 0.85 ± 0.08)	0.88 ± 0.06 (0.87 ± 0.06, 0.85 ± 0.08)

Table S5. Average precision scores for each dataset in the ablation study across all cross-validations. For each dataset, the highest score is written in bold. Conv: strided convolutional downsampling, Res: residual connection.

Dataset	WeakGCSeg	Conv/noRes	Maxpool/Res	Maxpool/noRes
IU 3.75°	**0.90 ± 0.06**	0.87 ± 0.06	0.87 ± 0.06	0.87 ± 0.06
IU 8.5°	**0.91 ± 0.02**	0.90 ± 0.03	0.89 ± 0.03	0.90 ± 0.03
IU 12.75°	**0.90 ± 0.05**	**0.90 ± 0.04**	0.88 ± 0.05	0.89 ± 0.05
FDA Healthy	**0.92 ± 0.05**	0.85 ± 0.06	0.86 ± 0.05	0.86 ± 0.05
FDA Glaucoma	**0.77 ± 0.14**	0.75 ± 0.12	0.74 ± 0.13	0.76 ± 0.14

Table S6. Effect of intensity normalization and test-time-augmentation on detection performance for the FDA dataset. Scores are reported as mean ± standard deviation for F_1 (recall, precision), calculated across $n = 6$ volumes from four healthy subjects and $n = 10$ volumes from five glaucoma subjects. Training and optimizations were done independently for the two groups of subjects.

Intensity normalization	Dataset	Test-Time-Augmentation	
		Yes	No
Identity	Healthy	0.76 ± 0.17 (0.76 ± 0.24, 0. 83 ± 0.05)	0.77 ± 0.12 (0.78 ± 0.21, 0.80 ± 0.06)
	Glaucoma	0.64 ± 0.19 (0.59 ± 0.25, 0. 74 ± 0.16)	0.61 ± 0.20 (0.57 ± 0.27, 0. 71 ± 0.20)
Whiten	Healthy	0.84 ± 0.05 (0.90 ± 0.04, 0.78 ± 0.07)	0.83± 0.05 (0.87 ± 0.06, 0.80 ± 0.05)
	Glaucoma	0.75 ± 0.11 (0.75 ± 0.14, 0.78 ± 0.15)	0.73 ± 0.12 (0.78 ± 0.14, 0.72 ± 0.16)

Table S7. Generalizability test between groups of subjects from the FDA dataset. Detection scores are reported as mean ± standard deviation for F_1 (recall, precision) across $n = 6$ volumes from four healthy subjects and $n = 10$ volumes from five glaucoma subjects, respectively. Networks were trained with whitened volumes, and predictions were made with test-time-augmentation. Validation data denote volumes used to determine the post-processing parameter T. There are three types of experiments in this table: (1) training, validation, and test data are from the same group, (2) test data are from a different group than the training and validation data, and (3) training data are from a different group than the validation and test data. The first category encompasses the experiments explained in Section 2.C of main text. In experiments (2), all volumes in the test group (e.g. healthy) were used to evaluate performance, and one volume from the other group (e.g. glaucoma) was randomly selected as the validation data while the rest of volumes were use to train the CNN. In experiments (3), all volumes in the training group (e.g. healthy) were used to train the CNN and volumes from the other group (e.g. glaucoma) were split to test (one subject) and validation (rest of the subjects in the group) data.

Test Data	Training Data	Validation Data	
		Healthy	Glaucoma
Healthy	Healthy	0.84 ± 0.05 (0.90 ± 0.04, 0.78 ± 0.07)	-
	Glaucoma	0.84 ± 0.03 (0.88 ± 0.05, 0.81 ± 0.05)	0.84 ± 0.03 (0.89 ± 0.04, 0.81 ± 0.05)
Glaucoma	Healthy	0.58 ± 0.22 (0.47 ± 0.27, 0.92 ± 0.07)	0.71± 0.11 (0.71 ± 0.18, 0.75 ± 0.11)
	Glaucoma	-	0.75 ± 0.11 (0.75 ± 0.14, 0.78 ± 0.15)

Table S8. Generalizability of trained CNN across different AO-OCT imaging systems with different scan characteristics. Detection scores are reported as mean ± standard deviation for F1 (recall, precision) across $n = 16$ and 6 volumes for the IU (eight subjects) and FDA (four subjects) datasets, respectively. The IU dataset consisted of the 3.75° and 12.75° locations, and the FDA dataset consisted of data from the healthy subjects. Networks were trained with whitened volumes, and predictions were made with test-time-augmentation. Resizing denotes interpolation of the test volumes to the voxel size of the training dataset prior to the network.

		Resizing Test Volumes	
Test	**Train**	**Yes**	**No**
IU	1. IU	-	0.85 ± 0.06 (0.86 ± 0.11, 0.86 ± 0.07)
	2. FDA	0.86 ± 0.05 (0.83 ± 0.08, 0.90 ± 0.05)	0.64 ± 0.16 (0.52 ± 0.17, 0.92 ± 0.04)
	3. Mixed	0.87 ± 0.04 (0.88 ± 0.06, 0.86 ± 0.06)	-
FDA	4. IU	0.75 ± 0.10 (0.93 ± 0.07, 0.63 ± 0.12)	0.55 ± 0.16 (0.43 ± 0.17, 0.83 ± 0.06)
	5. FDA	-	0.84 ± 0.05 (0.90 ± 0.04, 0.78 ± 0.07)
	6. Mixed	-	0.81 ± 0.09 (0.77 ± 0.15, 0.86 ± 0.04)

Supplementary Video Captions

Video S1| Automatically identified and segmented GCL somas at 3.75° temporal to the fovea. *En face* flythrough of a volume from the IU dataset with (*right*) soma detection results compared to the gold-standard manual markings and (*left*) overlay of soma segmentation masks, with each soma represented by a randomly assigned color. Cyan, red, and yellow markers denote true positives, false negatives, and false positives, respectively. Scale bar: 50 μm. The volume corresponds to the data shown in Fig. 3 with the same color for each cell.

Video S2| Automatically identified and segmented GCL somas at 8.5° temporal to the fovea. *En face* flythrough of a volume from the IU dataset with (*right*) soma detection results compared to the gold-standard manual markings and (*left*) overlay of soma segmentation masks, with each soma represented by a randomly assigned color. Cyan, red, and yellow markers denote true positives, false negatives, and false positives, respectively. Scale bar: 50 μm. The volume corresponds to the data shown in Fig. 3 with the same color for each cell.

Video S3| Automatically identified and segmented GCL somas at 12.75° temporal to the fovea. *En face* flythrough of a volume from the IU dataset with (*right*) soma detection results compared to the gold-standard manual markings and (*left*) overlay of soma segmentation masks, with each soma represented by a randomly assigned color. Cyan, red, and yellow markers denote true positives, false negatives, and false positives, respectively. Scale bar: 50 μm. The volume corresponds to the data shown in Fig. 3 with the same color for each cell.

Video S4| Three-dimensional illustration of automatically identified and segmented GCL somas at 3.75° temporal to the fovea. 3D *en face* flythrough of a volume from the IU dataset, with only somas located within 5.6 μm from each plane visualized. An ellipsoid was fit to each soma's segmentation mask and represented by a randomly assigned color. *En face* images are 195×195 μm^2 regions. The volume corresponds to Fig. 3 with the same color for each cell.

References

1. Z. Liu, K. Kurokawa, F. Zhang, J. J. Lee, and D. T. Miller, "Imaging and quantifying ganglion cells and other transparent neurons in the living human retina," Proceedings of the National Academy of Sciences **114**, 12803-12808 (2017).
2. C. A. Curcio and K. A. Allen, "Topography of ganglion cells in human retina," Journal of comparative Neurology **300**, 5-25 (1990).
3. M. Pavlidis, T. Stupp, M. Hummeke, and S. Thanos, "Morphometric examination of human and monkey retinal ganglion cells within the papillomacular area," Retina **26**, 445-453 (2006).
4. R. W. Rodieck, K. Binmoeller, and J. Dineen, "Parasol and midget ganglion cells of the human retina," Journal of Comparative Neurology **233**, 115-132 (1985).
5. A. B. Watson, "A formula for human retinal ganglion cell receptive field density as a function of visual field location," Journal of vision **14**, 15-15 (2014).
6. J. C. Blanks, Y. Torigoe, D. R. Hinton, and R. H. Blanks, "Retinal pathology in Alzheimer's disease. I. Ganglion cell loss in foveal/parafoveal retina," Neurobiology of aging 17, 377-384 (1996).
7. J. Stone and E. Johnston, "The topography of primate retina: A study of the human, bushbaby, and new- and old-world monkeys," J. Comp. Neuro. 196, 205-223 (1981).
8. E. A. Rossi, C. E. Granger, R. Sharma, Q. Yang, K. Saito, C. Schwarz, S. Walters, K. Nozato, J. Zhang, and T. Kawakami, "Imaging individual neurons in the retinal ganglion cell layer of the living eye," Proc Natl Acad Sci 114, 586-591 (2017).