Appendix S1. Pixel-Level Annotations and Segmentation Model

Annotations

Retinal pathology biomarkers and retinal layers were annotated by experts from the Liverpool Ophthalmology Reading Center on the B-scan level and subjected to internal quality assurance processes (a subselection of B-scan annotations of each grader was adjudicated and reviewed by a senior clinician). Specifically, a sparse selection of B-Scans across a total of 143 volume scans (860 B-scans in total), obtained from Cirrus (Carl Zeiss Meditec, Inc., Dublin, CA) OCT machines were annotated by drawing contours of the intraretinal fluid (IRF; cystoid spaces), subretinal fluid (SRF), pigment epithelial detachment (PED), and subretinal hyperreflective material (SHRM; a morphological feature seen on OCT as hyperreflective material located external to the neurosensory retina but internal to the retinal pigment epithelium [RPE]),[1] along with the following 5 retinal layers: internal limiting membrane, interface between outer plexiform layer and Henle's fiber layer, inner and outer boundary of RPE, and Bruch's membrane. Contours were drawn on the B-scans, stored in raster format, and then converted to label maps (fluids) and elevation maps (layers) of the original image dimension.

Definition of Training and Holdout Sets

The annotated B-scans were divided into training and holdout (validation) sets. There were 860 B-scans corresponding to 143 different patients in the training set, and 99 B-scans corresponding to 17 patients in the holdout set. To ensure that each feature was represented with approximately equal ratios in both sets, a fraction of 10% of the annotated B-scans was randomly designated to the holdout set under the stratifications

of SHRM, IRF, SRF, and PED volume. All the B-scans from a given eye were assigned to either the training or the holdout set to avoid overfitting.

Details on Training and Tuning

The U-Net, a convolutional neural network for biomedical image segmentation,[2] was trained using the training set to recognize the 4 retinal pathology-related features using the annotated volumes as training material, i.e., pixel-level semantic segmentation (Fig S3). To reach the best segmentation results on the training data with a 4-fold cross-validation approach, the U-Net was configured as follows. It processed images at 5 different resolution levels, each using 2 convolution layers with 3 × 3 kernels followed by Rectified Linear Unit[3,4] activations. The number of kernels in each layer was 64×2^l, where l is the resolution level from 0 (finest) to 4 (coarsest). Zero padding was applied to the first level to maintain the original input size of 512 × 512 pixels. The model was trained with the Adam optimizer[5] with a learning rate of 0.001 using batches of 12 images for 60 epochs. Dropout[6] was applied during training with 0.5 probability before and after the convolutions of the coarsest resolution level, and deconvolution (or transposed convolution) for the upsampling/decoder branch.

Performance

Segmentation performance for SHRM, IRF, SRF, and PED on the holdout set was assessed against the corresponding reference annotations and measured via performance metrics (Dice coefficient [Sørensen-Dice similarity coefficients]). Compared with human annotation, our model yielded a mean Dice coefficient (±

standard deviation) of 0.69 (± 0.26) for SHRM, 0.70 (± 0.26) for IRF, 0.67 (± 0.27) for SRF, and 0.73 (± 0.24) for PED. Median values were significantly higher, with 0.76, 0.78, 0.74, and 0.82, respectively, indicating distributions with tails to the left.

References

1. Willoughby AS, Ying GS, Toth CA, et al; Comparison of Age-Related Macular Degeneration Treatments Trials Research Group. Subretinal hyperreflective material in the Comparison of Age-Related Macular Degeneration Treatments Trials. *Ophthalmology.* 2015;122(9):1846–1853.e5.
2. Ronneberger O, Fischer F, Brox T. U-Net: convolutional networks for biomedical image segmentation. In: Navab N, Hornegger J, Wells W, Frangi A, eds. *Medical Image Computing and Computer-Assisted Intervention - MICCAI 2015 Lecture Notes in Computer Science*. Cham, Switzerland: Springer; 2015:234–241.
3. Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks. Presented at: 26th Annual Conference on Neural Information Processing Systems; December 3–6, 2012; Lake Tahoe, NV.
4. Nair V, Hinton GE. Rectified linear units improve restricted Boltzmann machines. Presented at: 27th International Conference on Machine Learning; June 21–24, 2010; Haifa, Israel.
5. Kingma DP, Ba J. Adam: a method for stochastic optimization. Presented at: 3rd International Conference on Learning Representations; May 7–9, 2015; San Diego, CA.

6. Srivastava N, Hinton G, Krizhevsky A, et al. Dropout: a simple way to prevent neural networks from overfitting. *J Mach Learn Res.* 2014;15:1929–1958.