Supplemental Online Content

Fan R, Bowd C, Christopher M, et al. Detecting glaucoma in the Ocular Hypertension Study using deep learning. *JAMA Ophthalmol*. Published online March 17, 2022. doi:10.1001/jamaophthalmol.2022.0244

This supplementary material has been provided by the authors to give readers additional information about their work.

eMethods. Implemented class weights added to loss function to address POAG vs non-POAG end point imbalance

Dataset Preparation

Because the 22 OHTS sites used different fundus cameras, resulting in inherent variability in image quality and resolution, the training of DL models was much more challenging than if photographs came from a single site or camera. To provide consistent model inputs, we first extracted a region centered on the ONH using a semantic segmentation DL model, DeepLabv3+ (with a ResNet-18 backbone network trained for ONH extraction). A square region surrounding the extracted ONH was then automatically cropped for input to the DL model, where the side length of each cropped image is approximately two times larger than the ONH diameter. The cropped fundus images were then resized to 224×224 pixels. A single reviewer manually reviewed each cropped image to ensure it was correctly centered on the ONH (see Supplemental eFigure 1). The OHTS stereophotographs included both simultaneous and sequential photographs – which cover different areas of the optic nerve head. The images were cropped to provide consistent inputs to the deep learning model.

Data Augmentation

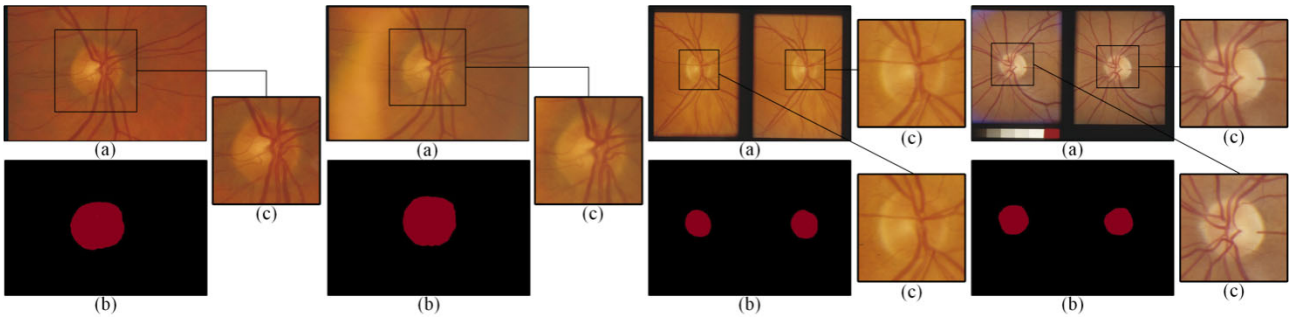
Several data augmentation strategies were applied to increase variation in the training set. To mimic the inclusion of both OD and OS orientations, horizontally mirrored versions of all photographs were added. In addition, we completed horizontal and vertical translation (≤40 pixels) and rotation (≤5°), in which the ONH center of each photograph was randomly perturbed by a small amount to reflect the common situation in which photographs are not well-centered. Each augmented image was assigned the same label (healthy or POAG) as the original input image from which it was derived.

Supplement highlights the conventional cross entropy loss *L* as follows:

$$L = -\frac{n_1}{n_0 + n_1} y \log p - \frac{n_0}{n_0 + n_1} (1 - y) \log(1 - p)$$

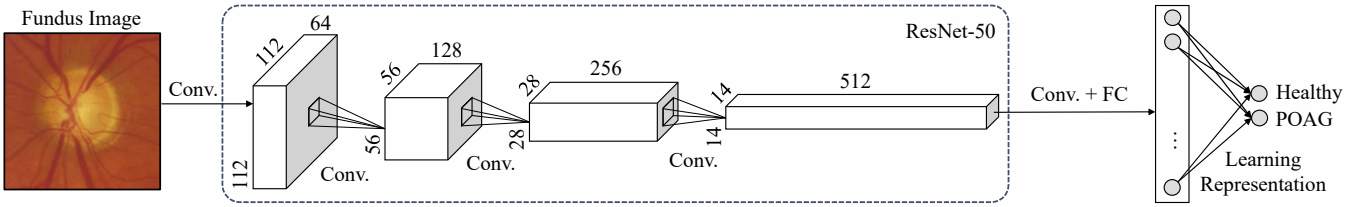
where y denotes the class label ($y = 0$ for healthy images and $y = 1$ for POAG images), p represents POAG prediction probability output by the network, and n_0 and n_1 denote the number of healthy and POAG images, respectively. We utilized the stochastic gradient descent with momentum (SGDM) optimizer too minimize (1), where the learning rate is set to 0.001 and the batch size is set to 30. The DL models we used were initially trained on the ImageNet database. In addition, due to the class imbalance of the OHTS dataset, we selected the best parameters of each DL model based on its achieved F-scores on the validation set, as this metric is better to use for seeking a balance between precision and recall, especially when the class distribution is uneven. Furthermore, we adopt the early stopping mechanism on the validation set to avoid over-fitting, where the tolerance is 5 epochs.

eFigure 1. Examples of optic nerve head detection and fundus photograph cropping



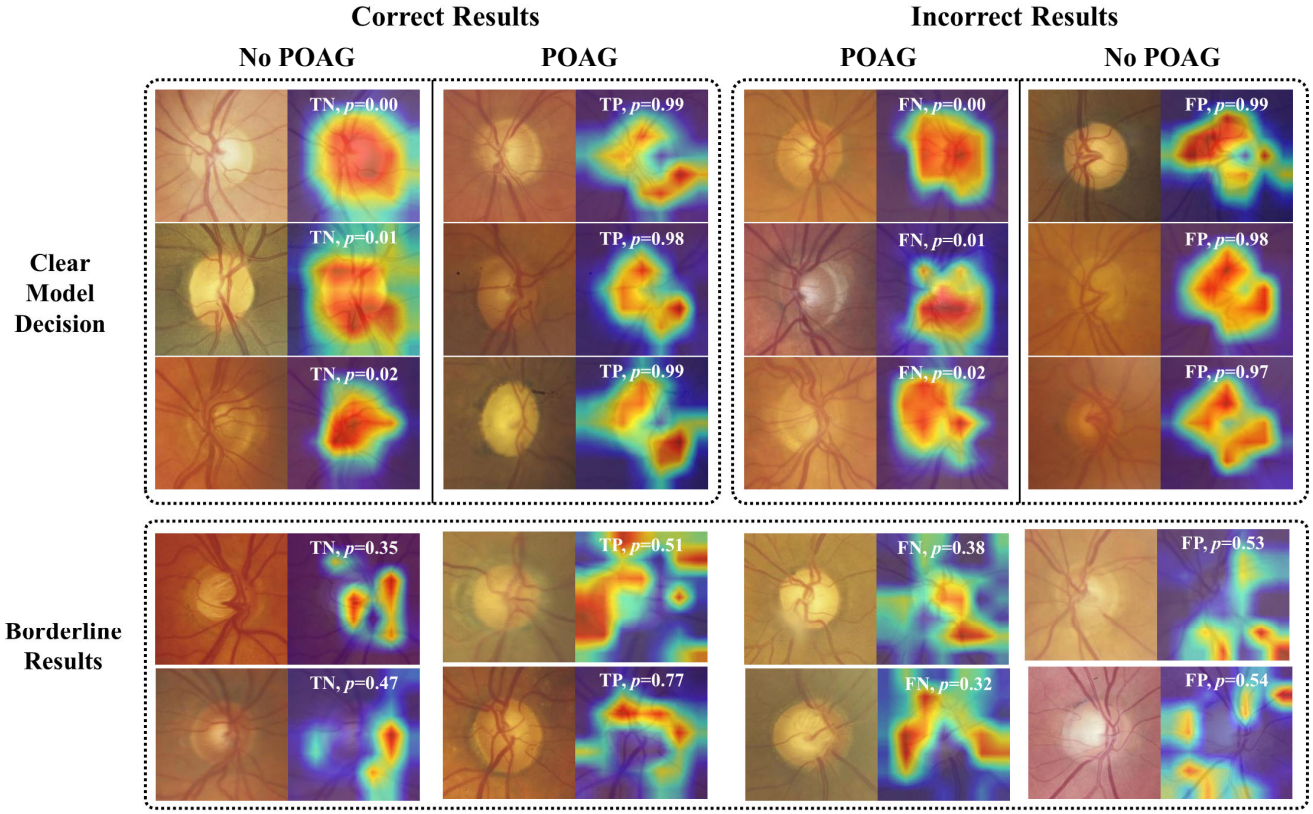
Examples of optic nerve head (ONH) detection and fundus photograph cropping: (a) raw fundus photographs; (b) ONHs (in red) detected by our trained DeepLabv3+ model; (c) cropped regions centered on the ONHs. The Ocular Hypertension Treatment Study (OHTS) dataset contains sequential (left two groups) and stereo (right two groups) fundus photographs.

eFigure 2. ResNet-50 architecture



ResNet-50 architecture, where Conv. represents a convolution layer and FC represents a fully-connected layer. A Softmax layer was added in the last to produce two scales indicating the probability distribution of healthy and primary open angle glaucoma (POAG) classes, respectively.

eFigure 3. Examples of deep learning model visualizations of the optic disc changes attributable to primary open-angle glaucoma (POAG) by the OHTS Endpoint Committee



Examples of deep learning model visualizations of the optic disc changes attributable to primary open-angle glaucoma (POAG) by the OHTS Endpoint Committee. Both clear model decisions (top) and borderline results (bottom) are shown. These results suggest that the region within the ONH had the greatest impact on clear model decisions (*p*, the probability of POAG estimated by the model, between 0.0 and 0.1 or between 0.9 and 1.0). The borderline results (*p* between 0.3 and 0.7) were less consistent with respect to the location on the disc in which the model based its decisions. TP refers to true positive, TN refers to true negative, FP refers to false positive, FN refers to false negative.

eTable 1. Time elapsed between the earliest false-positive photograph for each case and later OHTS Endpoint Committee determination of POAG

Ground Truth Determined by	POAG Detection Modality	Number of false positive (FP) results Participants (Eyes)/Visits	Mean (95% CI) number of years between model first false positive results and OHTS Endpoint Committee detection of primary open angle glaucoma (POAG)
Endpoint Committee	Optic Disc Photograph and/or Visual Field	31 (38) /198	5.2 (4.1, 6.3)
	Optic Disc Photograph	27 (32) /158	4.5 (3.4, 5.6)
	Visual Field	22 (26) /125	4.6 (3.3 , 5.8)

eTable 2. Diagnostic accuracy of deep learning model performance in identifying POAG using only 3 randomly selected visits/images in the test set

Ground Truth Determined by	POAG Detection Modality	POAG (n)	Area Under the Receiver Operating Characteristic Curve (95% CI)		Sensitivity at Specificity of:			
		Participants (Eyes)/Visits	All eyes	Early glaucoma VF MD ≥ -6 dB	80%	85%	90%	95%
Endpoint Committee	Optic Disc Photograph and/or Visual Field	47 (60) /102	0.89 (0.84, 0.94)	0.87 (0.79, 0.92)	0.83	0.81	0.69	0.56
	Optic Disc Photograph	36 (46) / 76	0.91 (0.85, 0.95)	0.91 (0.87, 0.95)	0.86	0.83	0.76	0.57
	Visual Field	31 (35) / 61	0.85 (0.75, 0.92)	0.84 (0.72, 0.93)	0.74	0.70	0.69	0.48
Reading Centers	Optic Disc Photograph	49 (69) / 93	0.89 (0.83, 0.93)	0.89 (0.83, 0.93)	0.85	0.77	0.73	0.55
	Visual Field	42 (49) / 76	0.83 (0.75, 0.89)	0.81 (0.71, 0.89)	0.68	0.63	0.61	0.49