A Deep Learning Approach to Digitally Stain Optical Coherence Tomography Images of the Optic Nerve Head

Sripad Krishna Devalla,[1] Khai Sing Chin,[2] Jean-Martial Mari,[3] Tin A. Tun,[4] Nicholas G. Strouthidis,[4,5,6] Tin Aung,[4,7] Alexandre H. Thiéry,[2] and Michaël J. A. Girard[1,4]

[1]Ophthalmic Engineering and Innovation Laboratory, Department of Biomedical Engineering, Faculty of Engineering, National University of Singapore, Singapore
[2]Department of Statistics and Applied Probability, National University of Singapore, Singapore
[3]GePaSud, Université de la Polynésie Française, Tahiti, French Polynesia
[4]Singapore Eye Research Institute, Singapore National Eye Centre, Singapore
[5]NIHR Biomedical Research Centre at Moorfields Eye Hospital NHS Foundation Trust and UCL Institute of Ophthalmology, London, United Kingdom
[6]Discipline of Clinical Ophthalmology and Eye Health, University of Sydney, Sydney, New South Wales, Australia
[7]Yong Loo Lin School of Medicine, National University of Singapore, Singapore

Correspondence: Michaël J. A. Girard, Ophthalmic Engineering and Innovation Laboratory, Department of Biomedical Engineering, National University of Singapore, Engineering Block 4, #04-8, 4 Engineering Drive 3, Singapore 117583; mgirard@nus.edu.sg.
Alexandre H. Thiéry, S-16, 06-113, Department of Statistics and Applied Probability, National University of Singapore, 6 Science Drive 2, Singapore-17546; a.h.thiery@nus.edu.sg.

AHT and MJAG contributed equally to the work presented here and should therefore be regarded as equivalent authors.

Submitted: July 13, 2017
Accepted: November 20, 2017

Citation: Devalla SK, Chin KS, Mari J-M, et al. A deep learning approach to digitally stain optical coherence tomography images of the optic nerve head. *Invest Ophthalmol Vis Sci*. 2018;59:63–74. https://doi.org/10.1167/iovs.17-22617

Purpose. To develop a deep learning approach to digitally stain optical coherence tomography (OCT) images of the optic nerve head (ONH).

Methods. A horizontal B-scan was acquired through the center of the ONH using OCT (Spectralis) for one eye of each of 100 subjects (40 healthy and 60 glaucoma). All images were enhanced using adaptive compensation. A custom deep learning network was then designed and trained with the compensated images to digitally stain (i.e., highlight) six tissue layers of the ONH. The accuracy of our algorithm was assessed (against manual segmentations) using the dice coefficient, sensitivity, specificity, intersection over union (IU), and accuracy. We studied the effect of compensation, number of training images, and performance comparison between glaucoma and healthy subjects.

Results. For images it had not yet assessed, our algorithm was able to digitally stain the retinal nerve fiber layer + prelamina, the RPE, all other retinal layers, the choroid, and the peripapillary sclera and lamina cribrosa. For all tissues, the dice coefficient, sensitivity, specificity, IU, and accuracy (mean) were 0.84 ± 0.03, 0.92 ± 0.03, 0.99 ± 0.00, 0.89 ± 0.03, and 0.94 ± 0.02, respectively. Our algorithm performed significantly better when compensated images were used for training ($P < 0.001$). Besides offering a good reliability, digital staining also performed well on OCT images of both glaucoma and healthy individuals.

Conclusions. Our deep learning algorithm can simultaneously stain the neural and connective tissues of the ONH, offering a framework to automatically measure multiple key structural parameters of the ONH that may be critical to improve glaucoma management.

Keywords: glaucoma, artificial intelligence, deep learning, optic nerve head, optical coherence tomography, digital staining, adaptive compensation

In glaucoma, the optic nerve head (ONH) exhibits complex structural changes, including, but not limited to, thinning of the retinal nerve fiber layer (RNFL)[1]; changes in choroidal thickness,[2,3] minimum rim width,[4] and lamina cribrosa (LC) depth[5]; and scleral canal expansion and bowing.[6,7] If all these structural parameters (and their changes) could be measured automatically with optical coherence tomography (OCT), it could considerably assist clinicians in their day-to-day management of glaucoma.

For OCT research, manual segmentation has remained the gold standard to extract structural information of the ONH, and this is especially true for deeper connective tissues.[8,9] However, manual segmentation is time consuming, prone to bias, and unsuitable in a clinical setting.[10,11] Although several techniques have been proposed to automatically segment some (but not all) ONH tissues in OCT images,[10,12–20] each tissue currently requires its own processing algorithm. This lack of a "universal" approach may limit the clinical translation and appeal for these algorithms.

Furthermore, the quality of automated segmentations/delineations largely depends on that of the OCT images. Poor deep-tissue visibility and shadow artifacts[21] in OCT images as a result of light attenuation make the development of robust segmentation tools difficult. With the advent of swept-source OCT,[22] enhanced depth imaging,[23–25] and compensation technology,[26] the quality of OCT images has been improved, opening the door to new possibilities. Recently, our group has developed a postprocessing technique that, when combined with compensation, could digitally stain (highlight) neural and connective tissues in OCT images of the ONH. However, this approach remains limited, as it cannot identify each ONH tissue separately, and in some cases requires manual inputs.[27]

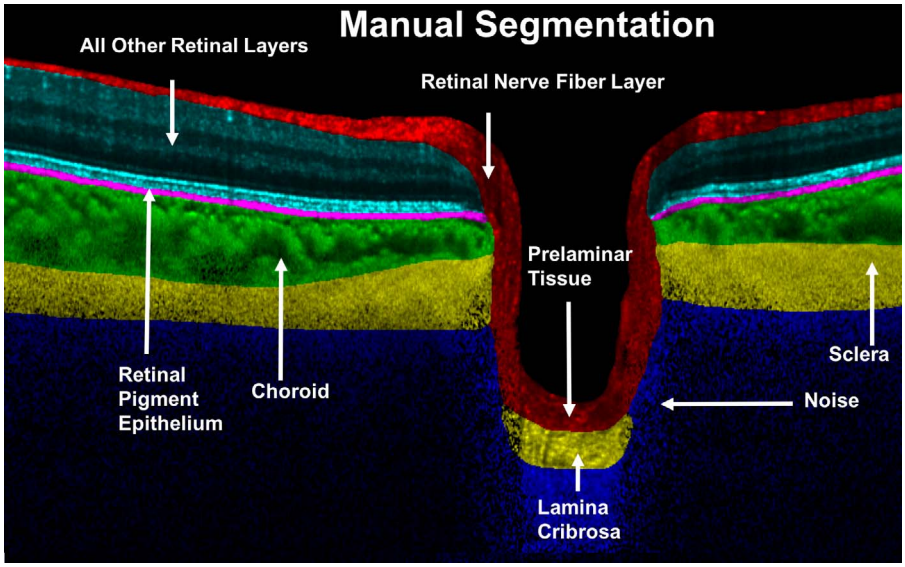

FIGURE 1. Manual segmentation of a compensated OCT image of the ONH. The RNFL and prelamina are shown in *red*, the RPE in *pink*, all other retinal layers in *cyan*, the choroid in *green*, the peripapillary sclera and LC in *yellow*, and noise in *blue*.

In this study, we aimed to develop a custom deep learning algorithm to automatically and simultaneously stain six important neural and connective tissue structures in OCT images of the ONH. We hope to offer a framework to automatically extract key structural information that has remained difficult to obtain in OCT scans of the ONH.

METHODS

Patient Recruitment

A total of 100 subjects were recruited at the Singapore National Eye Centre. All subjects gave written informed consent, and the study adhered to the tenets of the Declaration of Helsinki and was approved by the institutional review board of the hospital. The subject population consisted of 40 healthy controls, 41 subjects with primary open-angle glaucoma (POAG), and 19 patients with primary angle closure glaucoma (PACG). Inclusion criteria for healthy controls were as follows: IOP $\leq$ 21 mm Hg, healthy optic nerves with vertical cup-disc ratio (CDR) less than or equal to 0.5, and normal visual fields. POAG was defined as glaucomatous optic neuropathy (GON; characterized as loss of neuroretinal rim with vertical CDR $>$ 0.7 and/or focal notching with nerve fiber layer defect attributable to glaucoma and/or asymmetry of CDR between eyes $>$ 0.2) with repeatable glaucomatous visual field defects. PACG was defined as the presence of GON with compatible visual field loss, in association with a closed anterior chamber angle and/or peripheral anterior synechiae in at least one eye. A closed anterior chamber angle was defined as the posterior trabecular meshwork not being visible in at least 180° of anterior chamber angle.

OCT Imaging

OCT imaging was performed on seated subjects under dark room conditions after dilation with tropicamide 1% solution. Images were acquired by a single operator (TAT). The diagnosis was masked with the right ONH being imaged in all the subjects, unless the inclusion criteria were met only in the left eye, in which case the left eye was imaged. A horizontal B-scan (0°) of 8.9 mm (composed of 768 A-scans) was acquired through the center of the ONH for all the subjects using spectral-domain OCT (Spectralis; Heidelberg Engineering, Heidelberg, Germany). Data averaging was set to 48 and enhanced depth imaging was used for all scans.

Correction of Light Attenuation Using Adaptive Compensation

To remove the effects of light attenuation from OCT images, all B-scans were postprocessed using adaptive compensation (AC).[26] For OCT images of the ONH, AC has been shown to remove blood vessel shadows, improve tissue contrast, and increase the visibility of several features of the ONH.[28,29] For all B-scans, we used a threshold exponent of 12 (to limit noise overamplification at high depth), and a contrast exponent of 2 (to improve overall image contrast).[23]

Manual Segmentation of OCT Images

We performed manual segmentation of all compensated OCT images to train our digital staining algorithm to identify and highlight tissues, and to validate the accuracy of our approach. Specifically, each compensated OCT image were manually segmented by two expert observers (SD and KC) using Amira (version 5.4; FEI, Hillsboro, OR, USA) to identify the following classes: (1) the RNFL and the prelamina (in red, Fig. 1); (2) the RPE (in pink); (3) all other retinal layers (in cyan); (4) the choroid (in green); and (5) the peripapillary sclera and the LC (in yellow). Noise (below the peripapillary sclera and LC) was color-coded in blue. Note that in most cases, a full-thickness segmentation of the peripapillary sclera and of the LC was not possible due to limited visibility.[29] Therefore, we segmented only the visible portions of the sclera/LC as detected from the compensated OCT signal, and no effort was made to capture their accurate thickness. The manual segmentation assigned a label (defined between 1 and 6) to each pixel of each OCT image to indicate the tissue class.

Investigative Ophthalmology & Visual Science

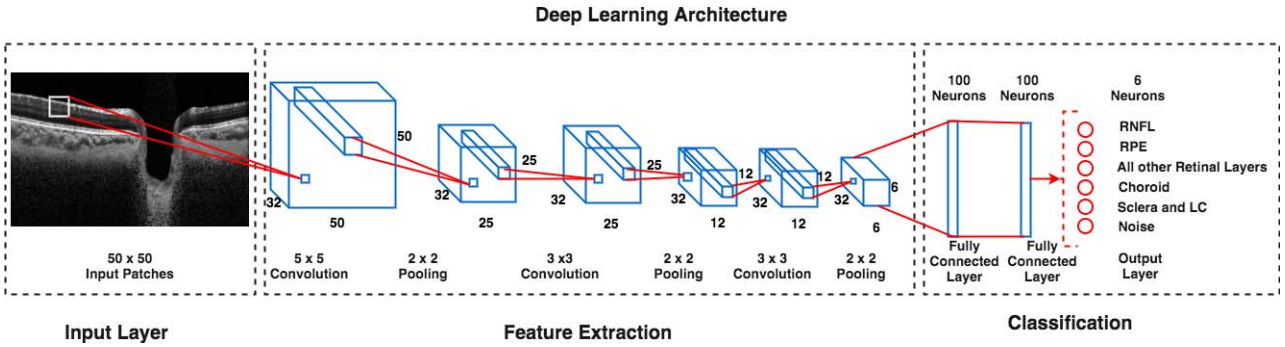

FIGURE 2. The deep learning architecture is an eight-layer CNN composed of three convolution layers, three max-pooling (subsampling) layers, and two fully connected layers.

Digital Staining of the ONH Using Deep Learning

In this study, we developed a custom deep learning approach to automatically stain ONH tissues in OCT images of the ONH. In recent years, deep learning has been extensively used in the field of medical imaging for diagnosis and segmentation applications. Although the concept of deep learning for medical image segmentation is not new for other imaging modalities such as magnetic resonance imaging,[30,31] its application to OCT is still relatively recent.[31–34] Although a number a studies have shown the successful segmentation of retinal[31,33,34] and macular edema[32] in OCT images using deep learning, to the best of our knowledge, no studies have been able to simultaneously stain neural and connective tissues in OCT images of the ONH. To this end, we used a two-dimensional convolution neural network (CNN) that was trained with manually segmented OCT images to recognize the most representative features of each tissue (present in small image patches). When our CNN was presented with an OCT image it had not yet seen, it was able to identify patches in the new image with features that "matched" those from the training set; each patch was then assigned a color code in its center (or probability to belong to a given tissue; one color per tissue between 1 and 6), to generate a digitally stained image of the ONH.

Network Architecture

For this study, we used an eight-layer CNN that was composed of three convolution layers, three max-pooling layers, and two fully connected layers (see detailed architecture in Fig. 2). Overlapping patches (quantity: approximately 2700; size: 50 × 50 pixels; stride length: 1 pixel) were extracted from each OCT image (size: 497 × 768 pixels) and fed as a single channel gray-scale image to the input layer.Each of the three convolution layers extracted 32 feature maps with filters of size 5 × 5, 3 × 3, and 3 × 3 pixels, respectively. The feature maps from the hidden layers (all layers except the input and output layers) were activated using a rectified linear unit (ReLu) function. Two fully connected layers with 100 neurons each were used to connect all the activations in the previous layers and funnel their excitations to the output layer. The output layer had six neurons, each corresponding to one class of tissue (i.e., RNFL, RPE, all other retinal layers, choroid, peripapillary sclera + LC, and noise). A softmax activation function was then applied to the output layer to obtain the class-wise probabilities (to belong to a given class) for each patch. For simplicity, every patch was assigned the class label with the highest probability in its center.

The proposed CNN was composed of 130,000 trainable parameters and was able to learn the initially unknown weights and biases during training using a standard cross-entropy loss function and an ADAM gradient descent optimization algorithm (learning rate: 0.001).[35] To reduce overfitting, a dropout of 35% was used in the last layer before the softmax activation. The loss function was scaled during the training using class weights for each output class of tissue to circumvent the fact that tissues covering large areas in OCT images (e.g., RNFL + prelamina) were represented by more patches. Specifically, the class weights assigned to each class of tissue were inversely proportional to the number of patches representing it in the training set (i.e., the more patches representing a particular class of tissue, the lesser its weight). Due to the limited size of our dataset, we performed online data augmentation (as is common in machine learning) by rotating (10°; clockwise and counterclockwise), flipping (horizontally), and translating (5 pixels; vertically and horizontally) our patches. The proposed CNN learned the specific features for each class of tissue in batches of 50 patches over 100 epochs (iterations). Note that in each epoch, all the patches in a batch underwent data augmentation before the start of the training. The CNN was developed using the Python programming language (Python Software foundation, https://www.python.org/) and implemented using the Keras[36] framework for deep learning with Tensorflow[37] as backend. We trained and tested the proposed CNN on an NVIDIA GTX 1080 (Nvidia, CA, USA) founder's edition GPU with CUDA v8.0 (Nvidia) and cuDNN v5.1 (Nvidia) acceleration.

Hyperparameter Tuning

The proposed architecture design was finalized purely on a heuristic basis after performing several experiments with varying kernel sizes, number of kernels, and depth. Due to the scarcity of available segmented OCT images, we have used the whole dataset for fine-tuning the architecture of the network. The regularization techniques of data augmentation and dropout layer were manually designed after observing the results over multiple experiments (for training set of size 10 images). With low (5%–10%) or no dropout, we observed that the network was overfitting. Thus, we kept increasing the dropout (35%) until we achieved a network that does not overfit. For training the network, we used an ADAM optimizer with a learning rate of 0.001.

Training and Testing of Our CNN

Once a robust network architecture was found and then fixed until the end of the study, we carried out several set of cross-

validation experiments for assessing the accuracy of the segmentation outputs and estimating the impact of the number of training examples on the accuracy of the network. In each of the cross-validation experiments, an equal number of compensated glaucoma and healthy OCT images were used in each of the training sets. We carried out experiments with training datasets of size 10, 20, 30, and 40 B-scans. In each one of these experiments, the accuracy was evaluated on the remaining set of images that were not part of the training sets. To study the effect of compensation on the accuracy of the segmentation, a similar set of experiments were performed with uncompensated images. We would like to emphasize one more time, though, that the entire dataset of images was initially leveraged for fine-tuning the architecture of the network. Consequently, there indeed is a small leakage of the test sets of images in the training procedures.

Digital Staining Performance: Qualitative Assessment

All digitally stained images were reviewed manually by the expert observers (SD and KC) for all training sets (with compensated and uncompensated images) and compared (qualitatively) with their corresponding manual segmentations.

Digital Staining Performance: Quantitative Assessment

To estimate the accuracy of digital staining in identifying individual ONH tissues in OCT images during the testing process, the following metrics were computed for each tissue and for each entire image: the dice coefficient, sensitivity, specificity, intersection over union (IU), and accuracy.

It is important to emphasize that these metrics could not be directly applied to the peripapillary sclera and LC, as their through-thickness visibility varied considerably across images. Instead, staining of the sclera and LC was assessed qualitatively.

The dice coefficient is a standard measure of similarity between two shapes and was used to assess the "overlap" between manual segmentation and digital staining. The dice coefficient is typically defined between 0 and 1, where 1 represents a perfect overlap and 0 no overlap. The dice coefficient DC_i was calculated for each tissue i (1: RNFL and prelamina, 2: RPE, 3: all other retinal layers, and 4: choroid), and for each B-scan in each testing set. It was defined as follows:

$$DC_i = \frac{2 * |DS_i \cap MS_i|}{|DS_i| + |MS_i|}, \quad (1)$$

where DS_i is the set of pixels representing the ONH tissue i in the digitally stained B-scan, and MS_i is that in the corresponding manually segmented B-scan.

Specificity, also defined as true negative rate (TNR) can assess the false predictions made by digital staining, and was calculated for each ONH tissue i and for each B-scan in each testing set as follows:

$$S.p_i = \frac{|\overline{DS}_i \cap \overline{MS}_i|}{|\overline{MS}_i|}, \quad (2)$$

where $\overline{DS}_i$ is the set of all pixels not belonging to tissue i in the digitally stained B-scan and $\overline{MS}_i$ is that in the corresponding manually segmented B-scan.

Sensitivity, also defined as the true positive rate (TPR), can assess the ability of digital staining to accurately stain a given ONH tissue, and was calculated for each ONH tissue i and for each B-scan in each testing set. It was defined as in Equation 3:

$$S.n_i = \frac{|DS_i \cap MS_i|}{|MS_i|} \quad (3)$$

We also computed two other metrics, namely the IU and the accuracy, both of which can assess the area of overlap between digital staining and manual segmentation. They are defined as follows:

$$IU = \frac{TPR}{TPR + FPR + FNR} \quad (4)$$

$$Accuracy = \frac{TPR + TNR}{TPR + TNR + FPR + FNR}, \quad (5)$$

where FNR is the false negative rate (FNR = 1 − TPR) and FPR the false positive rate (FPR = 1 − TNR).

Specificity, Sensitivity, IU, and accuracy were reported on a scale of 0 to 1.

Effect of Training Set Size on Digital Staining Accuracy

We used a 1-way ANOVA to assess differences in dice coefficients, sensitivities, and specificities (mean) for a given tissue across training set sizes (data were pooled for a given training set size). The test was performed in MATLAB (R2015a; MathWorks, Inc., Natick, MA, USA) and statistical significance was set at $P < 0.05$.

Digital Staining Reliability

We assessed the reliability of digital staining using the manual segmentations from the two expert observers. For this experiment, two CNNs were trained: one with the manual segmentation from the first observer, and the other with the manual segmentation from the second observer. Note that 10 images were used for training for each CNN. Dice coefficients (averaged for all tissues) were then calculated for the four following cases:

A. Manual segmentation from the first observer versus digital staining trained with the first observer.
B. Manual segmentation from the second observer versus digital staining trained with the first observer.
C. Manual segmentation from the first observer versus digital staining trained with the second observer.
D. Manual segmentation from the second observer versus digital staining trained with the second observer.

Paired t-tests were then used to assess the differences in dice coefficients between cases A and B; and between cases C and D. In addition, we aimed to understand differences in manual segmentations between the two expert observers by calculating the dice coefficient for the following case:

E. Manual segmentation from the first observer versus manual segmentation from the second observer.

Performance Comparison Between Glaucoma and Healthy Subjects

We used unpaired Student's t-test to quantitatively compare the performance of digital staining when testing was performed either on healthy or glaucoma OCT images. Specifically, for a training set of size 10, we used unpaired t-tests to assess the

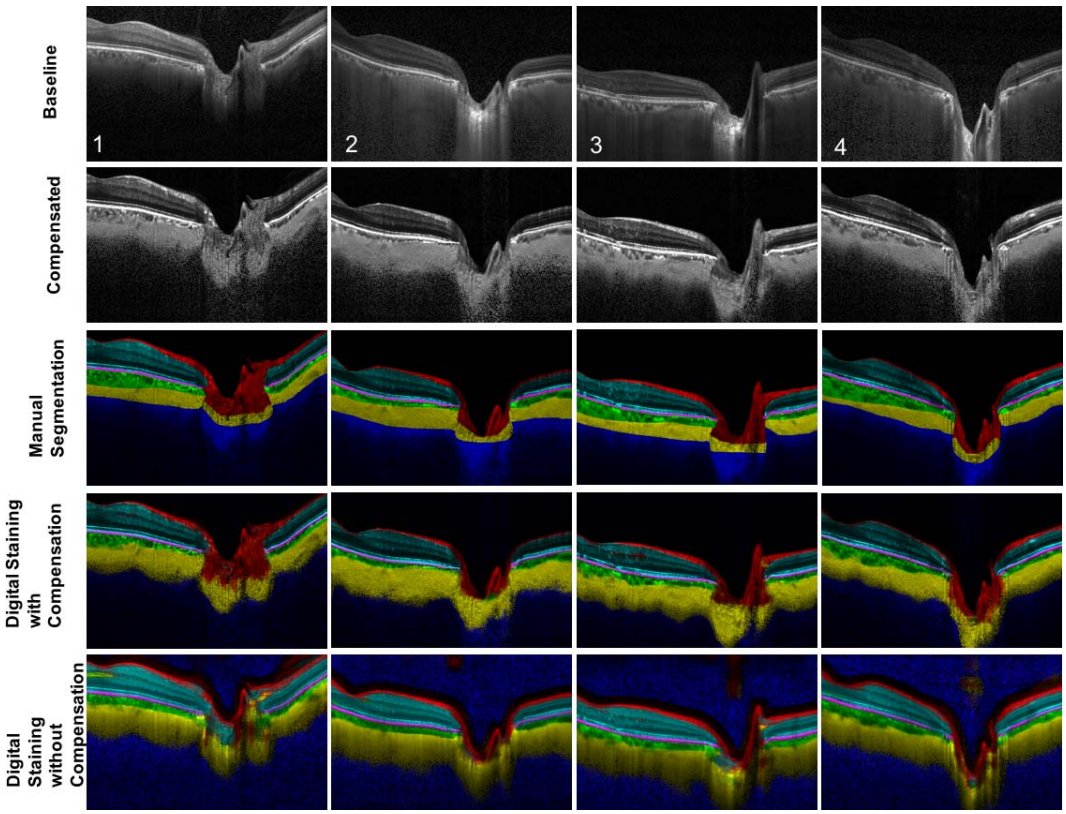

FIGURE 3. Baseline (*first row*), compensated (*second row*), manually segmented (*third row*), digitally stained images (trained on 10 compensated images; *fourth row*), and digitally stained images (trained on 10 baselined images; *fifth row*) for four selected subjects (1: Healthy, 2 and 3: POAG, 4: PACG).

differences in the mean values of dice coefficients, specificities and sensitivities, IUs, and accuracies for each tissue (for a training set of size 10, data were pooled across all datasets separately for glaucoma and healthy images).

Effect of Compensation on Digital Staining Accuracy

We used paired *t*-tests to assess whether our digital staining algorithm exhibited improved performance when trained with compensated images (as opposed to uncompensated images). Specifically, for a training set size of 10, we used *t*-tests to assess the differences in the mean values of dice coefficients, sensitivities, and specificities for each tissue (data were pooled for the training set of size 10).

RESULTS

Qualitative Analysis

Baseline, compensated, manually segmented, and digitally stained images (with training performed on 10 compensated or 10 baseline images) for four selected subjects (1: Healthy, 2 and 3: POAG, 4: PACG) can be found in Figure 3.

When compensated images were used for training (Fig. 3, fourth row), we found that our digital staining algorithm was able to simultaneously highlight the RNFL + prelamina (in red), the RPE (in pink), all other retinal layers (in cyan), the choroid (in green), the sclera + LC (in yellow), and noise (in blue). Digitally stained images were similar to those obtained from manual segmentation, and the results were consistent across all subjects and for all testing sets. Overall, the anterior LC was well captured, but our algorithm had a tendency to always identify LC insertions into the sclera that were not always present in the manual segmentations (e.g., subject 4). Small errors were sometimes observed. For instance, a small portion of the central retinal trunk was identified as choroidal tissue (green) in subject 2. Interestingly, although we provided a "smooth" delineation of the choroid-scleral interface, our algorithm had a tendency to follow the "undulations" of choroidal vessels.

When baseline images were instead used for training (Fig. 3, fifth row), more errors were observed. For instance, parts of the retina and prelamina were identified as scleral tissue (yellow) in subject 1.

Quantitative Analysis

Across all tests (with training performed on compensated images), we found that the average dice coefficient was 0.82 ± 0.05 for the RNFL + prelamina, 0.84 ± 0.02 for the RPE, 0.86 ± 0.03 for all other retina layers, and 0.85 ± 0.02 for the choroid. Sensitivity and specificity were high for all tissues: 0.89 ± 0.04 and 0.99 ± 0.00 for the RNFL + prelamina, 0.90 ± 0.03 and 0.99 ± 0.00 for the RPE, 0.98 ± 0.02 and 0.99 ± 0.00 for all other retina layers, and 0.91 ± 0.02 and 0.99 ± 0.00 for the choroid, respectively. The IU and accuracy were also relatively high for all tissues: 0.87 ± 0.06 and 0.93 ± 0.02 for the RNFL + prelamina, 0.86 ± 0.04 and 0.93 ± 0.02 for the RPE, 0.97 ± 0.02 and 0.98 ± 0.01 for all other retina layers, and 0.87 ± 0.03 and 0.93 ± 0.01 for the choroid, respectively.

For a given training set size, results were highly consistent across all tests (see Fig. 4 showing dice coefficients,

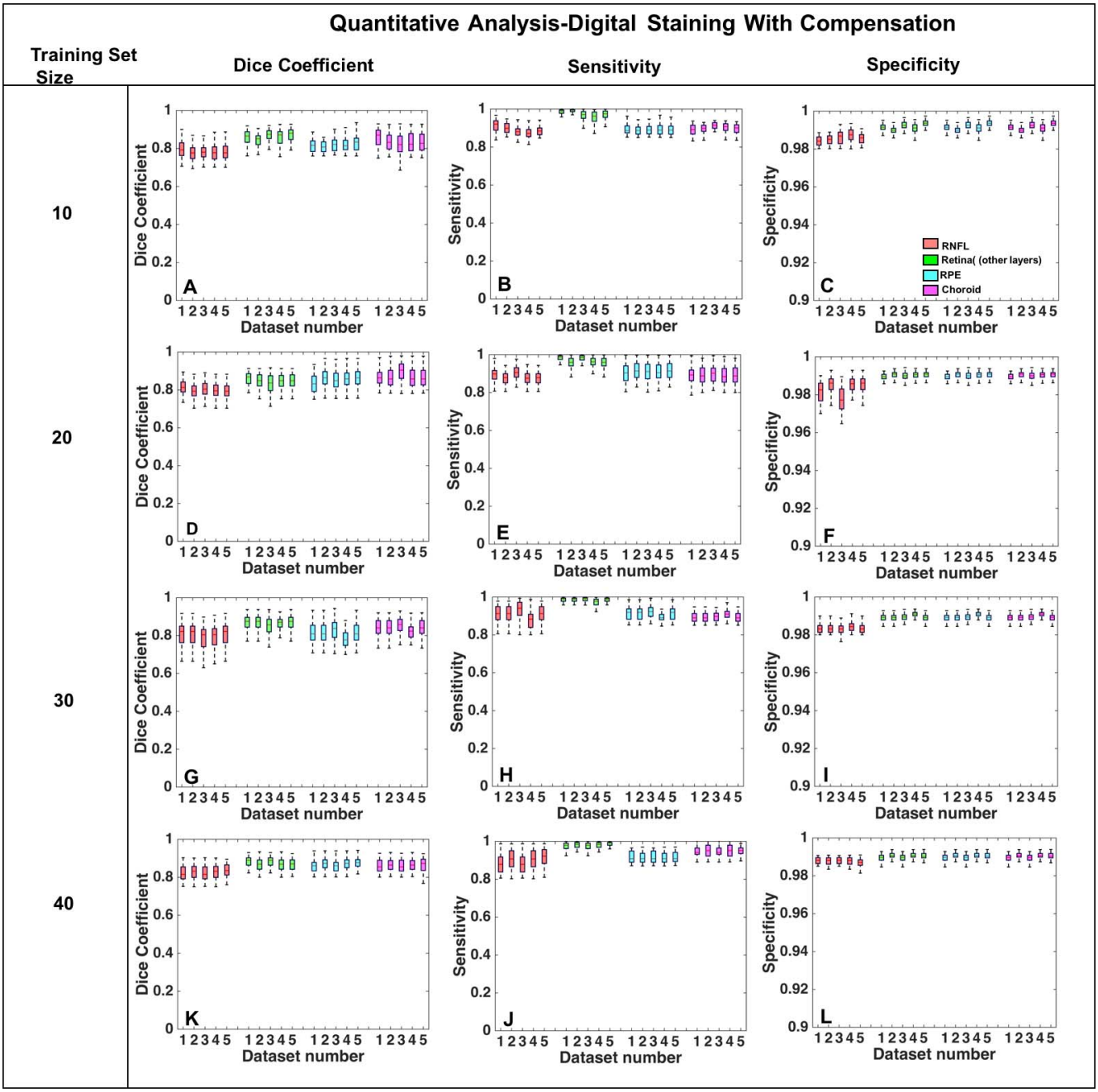

FIGURE 4. (**A–C**) Dice coefficients, sensitivities, and specificities represented as *box plots* for each tissue (RNFL in *red*, all other retina layers in *green*, RPE in *blue*, and the choroid in *magenta*) for each of five testing sets per tissue (training images per set: 10; testing images per set: 90). (**D–F**) Same as above except that 20 images were used for each training set, and 80 images for each testing set. (**G–I**) Thirty images were used for each training set, and 70 images for each testing set. (**K–L**) Forty images were used for each training set, and 60 images for each testing set.

sensitivities, and specificities for each tissue, five tests per tissue, training set size: 10).

Dice coefficients, sensitivities, specificities, IU, and accuracy (mean ± SD) for all training set sizes and all tissues are listed in the Table. For all tissues but RPE, we found that the training set size had no significant impact on the dice coefficient and on sensitivity ($P > 0.05$ for all cases). However, increasing the training set size from 10 to 40 significantly improved the dice coefficient for RPE ($P < 0.001$) from 0.81 ± 0.04 to 0.87 ± 0.03. Finally, the training set size had a significant impact on specificity for all tissues ($P < 0.001$ for all cases); however, we noted that specificity values were always higher than 0.98 for all cases.

No significant differences ($P > 0.05$ for all cases) were observed in the dice coefficients, specificities, sensitivities, IUs, and accuracies (means) between glaucoma and healthy OCT images (see Fig. 5; showing dice coefficient, sensitivity, specificity, IU, and accuracy for each tissue and a training set of size 10; data were pooled across all datasets).

There was no significant difference ($P > 0.05$) in dice coefficient when digital staining was compared against the manual segmentations from both expert observers irrespective of whose manual segmentation was used for training (A versus

Table. Mean Dice Coefficient, Sensitivity, Specificity, IU, and Accuracy for All Tissues Across All Datasets for All Sizes of Testing Set When Evaluated With Respective Training Sets

Training Set Size, Metric	Mean ± SD			
Tissue	10	20	30	40
Dice coefficient				
RNFL	0.824 ± 0.051	0.831 ± 0.04	0.819 ± 0.061	0.821 ± 0.038
Retinal layers (all others)	0.858 ± 0.028	0.856 ± 0.024	0.861 ± 0.022	0.872 ± 0.031
RPE	0.812 ± 0.040	0.828 ± 0.001	0.853 ± 0.045	0.867 ± 0.030
Choroid	0.836 ± 0.013	0.839 ± 0.021	0.854 ± 0.031	0.862 ± 0.035
Sensitivity				
RNFL	0.889 ± 0.030	0.891 ± 0.060	0.899 ± 0.014	0.897 ± 0.048
Retinal layers (all others)	0.969 ± 0.029	0.973 ± 0.021	0.978 ± 0.019	0.981 ± 0.009
RPE	0.890 ± 0.029	0.889 ± 0.034	0.901 ± 0.021	0.915 ± 0.031
Choroid	0.899 ± 0.024	0.898 ± 0.022	0.887 ± 0.006	0.889 ± 0.029
Specificity				
RNFL	0.981 ± 0.001	0.988 ± 0.005	0.986 ± 0.001	0.989 ± 0.007
Retinal layers (all others)	0.991 ± 0.001	0.984 ± 0.002	0.988 ± 0.000	0.991 ± 0.003
RPE	0.991 ± 0.002	0.989 ± 0.002	0.989 ± 0.002	0.992 ± 0.011
Choroid	0.990 ± 0.001	0.989 ± 0.002	0.991 ± 0.001	0.993 ± 0.002
IU				
RNFL	0.873 ± 0.043	0.859 ± 0.080	0.873 ± 0.078	0.868 ± 0.062
Retinal layers (all others)	0.966 ± 0.028	0.970 ± 0.022	0.968 ± 0.023	0.967 ± 0.022
RPE	0.851 ± 0.035	0.848 ± 0.051	0.890 ± 0.044	0.861 ± 0.048
Choroid	0.859 ± 0.041	0.855 ± 0.031	0.869 ± 0.033	0.879 ± 0.041
Accuracy				
RNFL	0.935 ± 0.022	0.926 ± 0.041	0.927 ± 0.040	0.932 ± 0.001
Retinal layers (all others)	0.982 ± 0.014	0.984 ± 0.011	0.984 ± 0.011	0.983 ± 0.023
RPE	0.925 ± 0.017	0.923 ± 0.025	0.941 ± 0.020	0.930 ± 0.011
Choroid	0.929 ± 0.021	0.926 ± 0.017	0.939 ± 0.017	0.929 ± 0.021

B, and C versus D in Fig. 6). Thus, digital staining was deemed reliable.

Overall, we found that digital staining performed significantly better when compensated images where used for training (as opposed to baseline or noncompensated images). Specifically, dice coefficients, sensitivities, and specificities were always significantly higher when our algorithm was trained with compensated images (versus baseline images; $P < 0.001$ for all cases; Fig. 7).

Discussion

In this study, we have developed a custom deep learning algorithm to digitally stain tissues in OCT images of the ONH. Our algorithm was tested and validated using OCT images from 100 subjects, and was found to exhibit relatively good performance across all tissues for both glaucoma and healthy images. Digital staining also was found to be significantly reliable when validated against the manual segmentations from the two expert observers. Furthermore, there was a good agreement in the manual segmentations between the two observers (averaged dice coefficient for all tissues = 0.89). We also found that our algorithm performed significantly better when it was trained with enhanced (i.e., compensated) OCT images. Our method is attractive because it identifies all tissue layers simultaneously, and because it is "universal," in the sense that the exact same processing approach is being used to identify each tissue layer individually. Our work could have applications for the clinical management of glaucoma using OCT. To the best of our knowledge, no deep learning techniques have yet been proposed for OCT images of the ONH.

In this study, we found that our digital staining algorithm was able to simultaneously isolate the RNFL + prelamina, the RPE, all other retina layers, the choroid, and the peripapillary sclera + LC. Our results were consistent across all subjects and we obtained relatively good agreements with respect to manual segmentations (for all tissues, averaged dice coefficients varied between 0.82 and 0.86, averaged sensitivities between 0.89 and 0.97, averaged specificities were always higher than 0.98, averaged IU varied between 0.87 and 0.96, and averaged accuracy varied between 0.93 and 0.98). We believe our work offers a framework to automatically measure structural parameters of the ONH. Typically, ONH tissues exhibit complex three-dimensional (3D) structural changes during the development and progression of glaucoma, including, but not limited to, changes in RNFL thickness and minimum rim width,[38] changes in LC depth,[5] changes in LC curvature,[39] changes in LC global shape index,[40] changes in choroidal thickness,[2,3] peripapillary atrophy,[41,42] scleral canal expansion,[6] migration of the LC insertion sites,[43,44] LC focal defects,[45,46] and scleral bowing.[7] On digital staining there might exist imperfections and tissue discontinuities due to false predictions that can be eliminated using simple nearest neighbors–based post processing technique. Following such a postprocessing approach, one can attempt to automatically extract several important structural parameters of the ONH. For instance, choroidal thickness and RNFL thickness maps could be derived from our choroid and RNFL + prelamina staining (simply as the number of colored pixels in each A-scan times a physical scale factor). To compute Bruch's membrane opening (BMO)–minimum rim width, one could first isolate the two BMO points (from our RPE staining) and compute the minimum distance from the inner limiting membrane (anterior boundary that can be obtained from our RNFL + prelamina staining).[47] One also could compute the amount of scleral bending in the scleral flange, LC curvature (nasotemporal and inferosuperior), and LC global shape index from our peripapillary sclera + LC staining.[40,48] Finally, LC depth could be

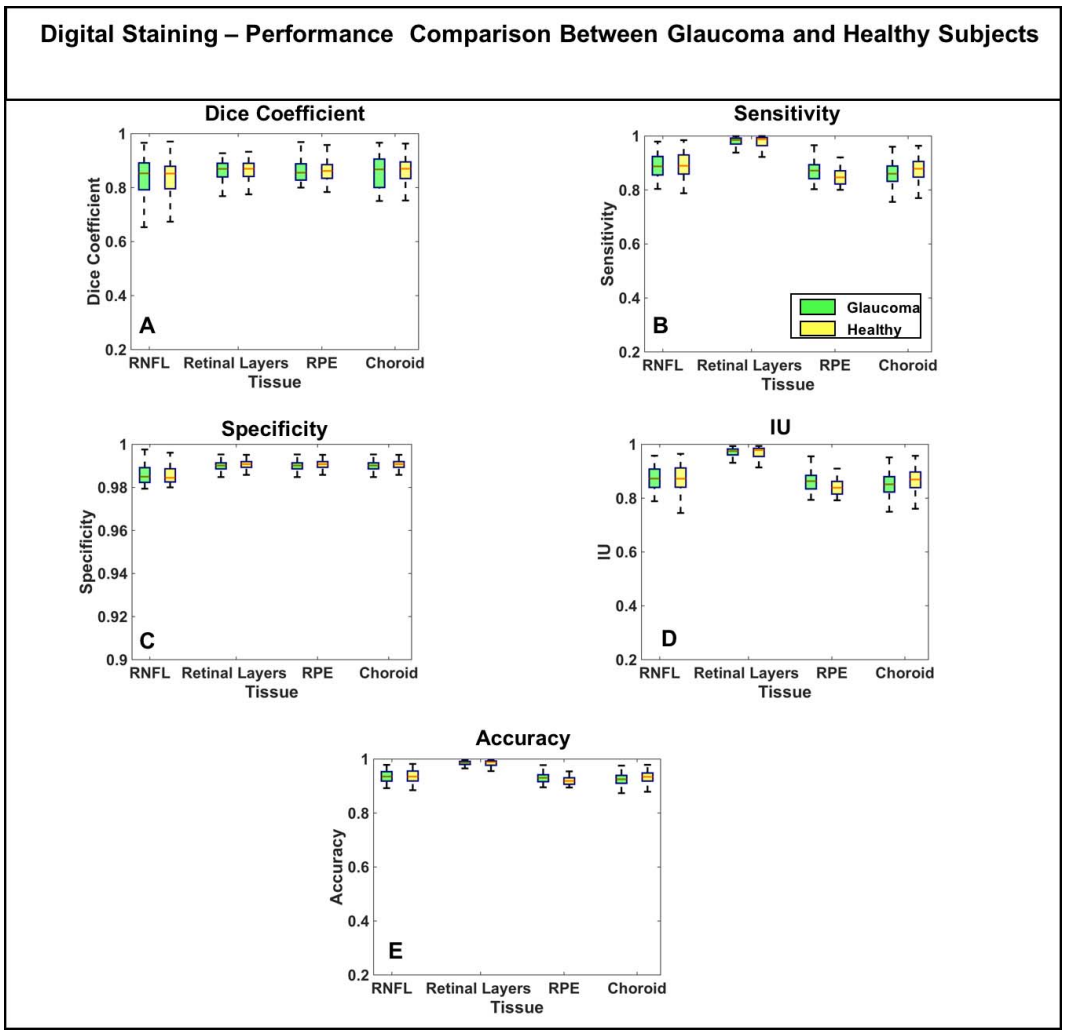

Figure 5. (**A–E**) Dice coefficient, sensitivity, specificity, IU, and accuracy represented as *box plots* for each tissue (RNFL, all other retina layers, RPE, and the choroid) for a training set of size 10 (data were pooled across all datasets) for glaucoma (*green*) and healthy OCT images (*yellow*).

automatically computed from our RPE (to identify BMO) and LC staining.[49] We believe these parameters may prove critical for glaucoma diagnosis, management, and risk profiling. Our approach thus may be of high importance for glaucoma diagnosis, management, and risk profiling.

Although accurate automated segmentation tools exist for retinal layers,[50,51] for the choroid,[12,13,18,52–54] or the choroid-scleral interface,[54] and to a lesser extent, for the LC,[55–57] there exists no "universal" tool to isolate both connective and neural tissues simultaneously. Each tissue currently requires its own specific algorithm, each of which may be computationally expensive.[53] It is interesting to note that a number of existing automated segmentation tools are prone to segmentation errors in images with pathology (e.g., AMD, optic disc edema).[58] On the contrary, our algorithm exhibited the same performance when tested on healthy or glaucoma images (while the training was performed on a mix of healthy and glaucoma images). This could have significant advantages for clinical translation of our tools; however, more work should be carried out to understand whether this would remain true for all glaucoma severities, and for ONHs with specific characteristics, such as peripapillary atrophy. Our current digital staining approach highlights all tissues simultaneously with the exact same deep learning backbone and requires only a few seconds of processing time for each image on a standard GPU card. Note that our group is currently developing a real-time digital staining solution to make it more attractive for glaucoma clinics.

We found that the quality of digital staining was relatively poor when our deep learning network was trained with baseline (uncompensated) images (Fig. 5). On average (all tissues) the dice coefficient was 0.56 ± 0.06 (versus 0.84 ± 0.03 when training with compensated images), sensitivity 0.65 ± 0.261 (versus 0.92 ± 0.03), and specificity 0.93 ± 0.02 (versus 0.99 ± 0.00). This is not surprising, as baseline images (versus compensated images) typically exhibit lower intra- and interlayer contrasts, low visibility at high depth, and strong blood vessel shadow artifacts.[26] Our work illustrates that adaptive compensation[26] may be a necessary first step toward a simple solution to automatically segment the ONH tissues.

Interestingly, we found that increasing the size of our training set (from 10 to 40 images) did not significantly improve digital staining accuracy, except for the RPE. This result may appear counterintuitive. However, contrary to most deep learning applications, our situation is intrinsically low dimensional: most OCT scans of the ONH are fundamentally similar to each other (e.g., the sclera is always posterior to the choroid). Furthermore, we would like to emphasize that, on

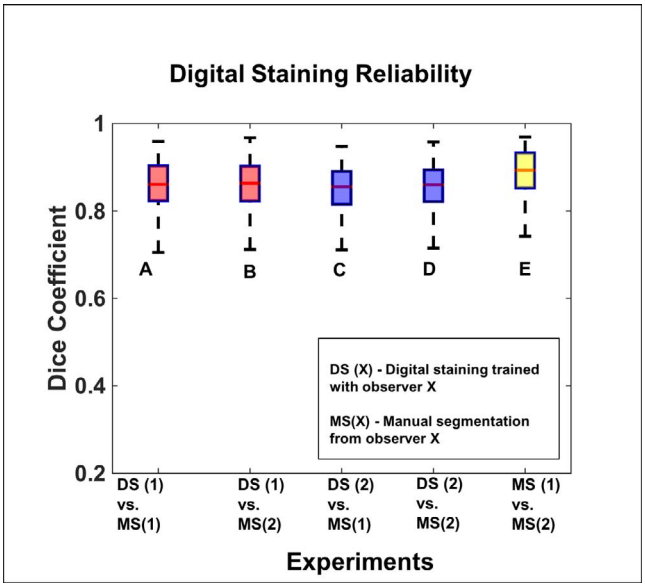

FIGURE 6. Digital staining reliability. Two CNN models were trained: one with the manual segmentation from the first observer, and the other with the manual segmentation from the second observer. Note that 10 images were used for training for each CNN. Dice coefficients (averaged for all tissues) were then calculated for the five following cases: (**A**) manual segmentation from the first observer versus digital staining trained with the first observer; (**B**) manual segmentation from the second observer versus digital staining trained with the first observer; (**C**) manual segmentation from the first observer versus digital staining trained with the second observer; (**D**) manual segmentation from the second observer versus digital staining trained with the second observer; (**E**) manual segmentation from the first observer versus manual segmentation from the second observer.

average, approximately 2700 patches of size 50×50 pixels for a stride of 1 pixel were extracted from each image. Thus, we effectively trained our deep learning network with a large amount of data (for 10 images, 10 images $\times$ 2700 patches = 27,000 patches in total). And because each ONH tissue exhibits similar texture at different locations, a patch that, for example, represents the nasal part of the choroid also can be used to identify the temporal part of the choroid during testing. These reasons may potentially explain why our results were accurate (while circumventing overfitting) even when training was performed on a small number of subjects (10). This can be seen as a strong advantage, as obtaining access to images is one the largest limiting factors for deep learning applications. However, we believe that more validations are required, and we aim to study much larger populations in the near future.

Finally, to the best of our knowledge, there are very few traditional image processing and machine learning approaches that have been proposed to simultaneously segment connective and neural tissues in OCT images of the ONH. In our previous study,[59] we proposed a histogram-based algorithm to digitally stain connective and neural tissues; such an approach was limited, as it was not able to fully separate different connective tissue layers (e.g., sclera from choroid). It also was found that techniques such as K-means clustering were not able to accurately dissociate neural from connective tissues. In all, we believe that a deep learning approach may be the preferred technique for OCT images of the ONH.

Several limitations in our study warrant further discussion. First, our algorithm was trained with OCT images from a single device (Spectralis), and it is currently unknown if our approach could be directly applied to images captured with other OCT devices. However, one could consider retraining the network for each device separately. We are currently exploring such an approach.

Second, given the limitation in dataset size (100 images), we did use the whole dataset for fine-tuning the global architecture of the network. Thus, there was an overall mixing of training/testing sets across all the experiments. Given a slightly larger dataset, our future works can definitely have an exclusive testing set. Nevertheless, we offer here a proof of principle of ONH digital staining that also could be used by other groups for further validation.

Third, we were unable to provide an additional validation of our algorithm by comparing our stained images with those obtained from histology. This is extremely difficult to achieve, as one would need to image a human ONH with OCT, process it with histology, and register both datasets. Note that the broad understanding of OCT ONH anatomy to histology has been based on a single comparison with a normal monkey eye scanned in vivo at an IOP of 10 mm Hg and then perfusion fixed at time of euthanasia at the same IOP.[60] The tissue classification derived from our algorithm matches the expected relationships observed in this canonical work. At the time of writing, there have been no published experiments matching human ONH histology to OCT images. Although the absence of this work prevents an absolute validation of our technique, the same shortcoming necessarily applies to every other in vivo investigation of deep OCT imaging of the human ONH, many publications of which predate even the publication of the comparison with the monkey ONH.

Fourth, in some subjects, we observed false predictions for a few pixels in the LC. This shortcoming could potentially be addressed with the use of a deeper network, a more advanced neural network architecture, a 3D CNN, or a simple postprocessing approach to filter tissue discontinuities following the digital stain step. Further work is required to explore all these options.

Fifth, in the patch-based approach, overlapping patches result in multiple convolutions on similar sets of pixels, which are a waste of computational memory and time. Recently developed architectures[61–64] for other biomedical imaging applications have circumvented these issues, which could be explored for OCT images of the ONH.

Sixth, in this study, the number of neurons in the two fully connected layers was less than the number of features in the last pooling layer. However, please note that we used a completely heuristic approach to identify the optimum number of neurons required in each of the fully connected layers after several experiments. When we increased the number of neurons (in the range of 500–1000) in the fully connected layers, the model became computationally expensive and was overfitting. When we used a smaller number of neurons, the accuracy was compromised. Thus, after several experiments, we have heuristically concluded that, for digital staining using the proposed architecture, we required two fully connected layers, each with 100 neurons to provide a good accuracy without overfitting.

Seventh, in some subjects, the LC insertions into the sclera highlighted by the algorithm that were not visible to the expert observer during manual segmentations makes it unclear if our algorithm was able to identify faint signals that resembled LC insertions, or whether it was introducing artifacts. A 3D validation may be required to address this phenomenon.

In conclusion, we have developed a custom deep learning algorithm to digitally stain nervous and connective tissues in OCT images of the ONH. Because these tissues exhibit significant structural changes in glaucoma, digital staining may be of interest in the clinical management of glaucoma.

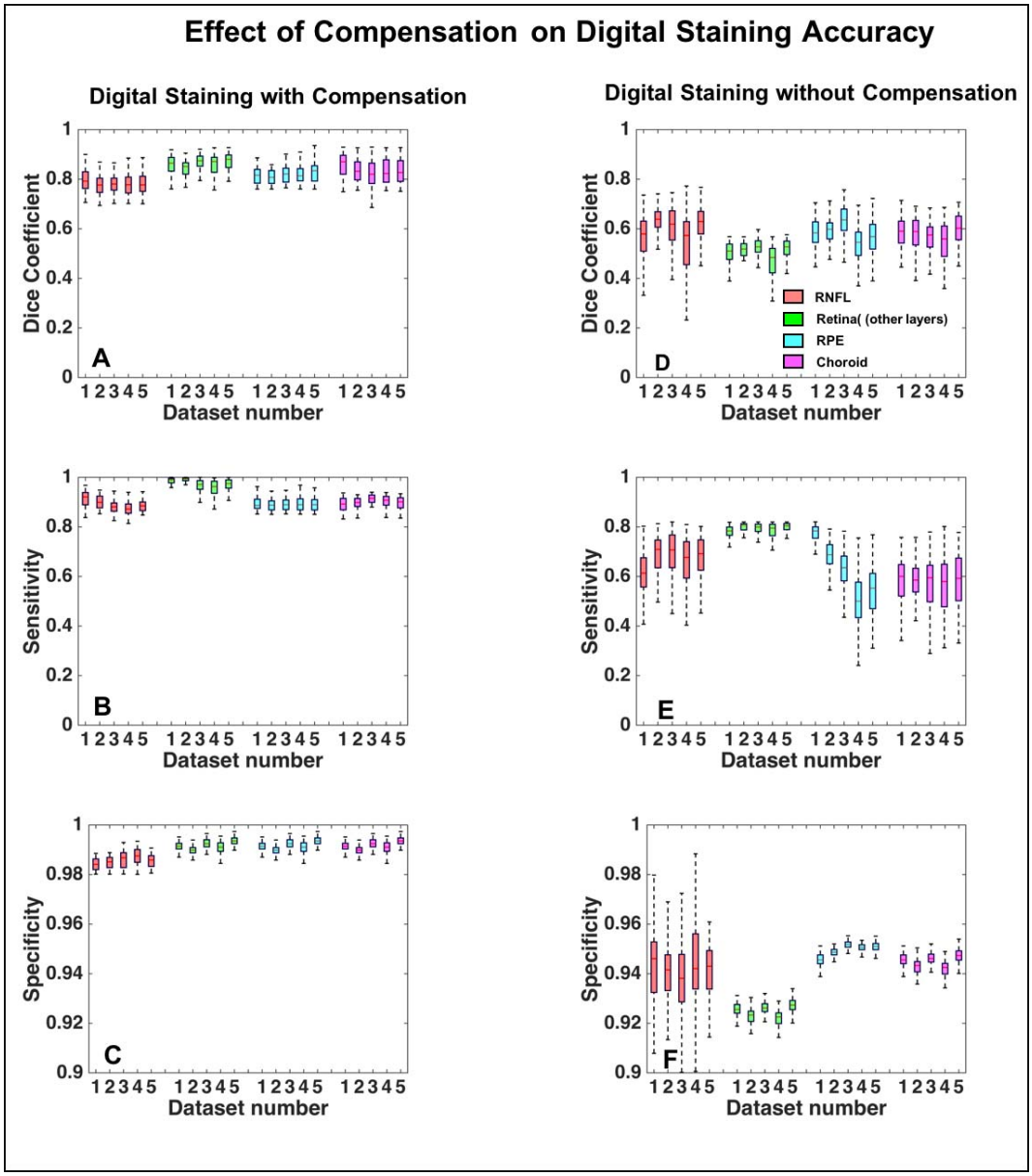

FIGURE 7. (**A–C**) Dice coefficients, sensitivities, and specificities when our algorithm was trained on compensated images (10 images per training set). (**D–F**) Dice coefficients, sensitivities, and specificities when our algorithm was trained on baseline (uncompensated) images (10 images per training set).

Acknowledgments

Supported by the Singapore Ministry of Education Academic Research Funds Tier 1 (R-155-000-168-112 [AHT]), a National University of Singapore Young Investigator Award Grant (NUSYIA_FY13_P03; R-397-000-174-133 [MJAG]), and by the National Medical Research Council (Grant NMRC/STAR/0023/2014 [TA]).

Disclosure: **S.K. Devalla**, None; **K.S. Chin**, None; **J.-M. Mari**, None; **T.A. Tun**, None; **N.G. Strouthidis**, None; **T. Aung**, None; **A.H. Thiéry**, None; **M.J.A. Girard**, None

References

1. Bowd C, Weinreb RN, Williams JM, Zangwill LM. The retinal nerve fiber layer thickness in ocular hypertensive, normal, and glaucomatous eyes with optical coherence tomography. *Arch Ophthalmol*. 2000;118:22–26.
2. Hosseini H, Nilforushan N, Moghimi S, et al. Peripapillary and macular choroidal thickness in glaucoma. *J Ophthalmic Vis Res*. 2014;9:154–161.
3. Zhang C, Tatham AJ, Medeiros FA, Zangwill LM, Yang Z, Weinreb RN. Assessment of choroidal thickness in healthy and glaucomatous eyes using swept source optical coherence tomography. *PLoS One*. 2014;9:e109683.
4. Gardiner SK, Boey PY, Yang H, Fortune B, Burgoyne CF, Demirel S. Structural measurements for monitoring change in glaucoma: comparing retinal nerve fiber layer thickness with minimum rim width and area. *Invest Ophthalmol Vis Sci*. 2015;56:6886–6891.
5. Sawada Y, Hangai M, Murata K, Ishikawa M, Yoshitomi T. Lamina Cribrosa depth variation measured by spectral-domain optical coherence tomography within and between four glaucomatous optic disc phenotypes. *Invest Ophthalmol Vis Sci*. 2015;56:5777–5784.

6. Sigal IA, Yang H, Roberts MD, et al. IOP-induced lamina cribrosa deformation and scleral canal expansion: independent or related? *Invest Ophthalmol Vis Sci*. 2011;52:9023–9032.
7. Kupersmith MJ, Sibony P, Mandel G, Durbin M, Kardon RH. Optical coherence tomography of the swollen optic nerve head: deformation of the peripapillary retinal pigment epithelium layer in papilledema. *Invest Ophthalmol Vis Sci*. 2011;52:6558–6564.
8. Han JC, Choi DY, Kwun YK, Suh W, Kee C. Evaluation of lamina cribrosa thickness and depth in ocular hypertension. *Jpn J Ophthalmol*. 2016;60:14–19.
9. Girard MJ, Beotra MR, Chin KS, et al. In vivo 3-dimensional strain mapping of the optic nerve head following intraocular pressure lowering by trabeculectomy. *Ophthalmology*. 2016; 123:1190–1200.
10. Lang A, Carass A, Hauser M, et al. Retinal layer segmentation of macular OCT images using boundary classification. *Biomed Opt Express*. 2013;4:1133–1152.
11. Gaiarsa VM, Cardenas D, Furuie SS. Approaches to segment stent area from intravascular optical coherence tomography. *Res Biomed Eng*. 2017;33:11–20.
12. Zhang L, Lee K, Niemeijer M, Mullins RF, Sonka M, Abràmoff MD. Automated segmentation of the choroid from clinical SD-OCT. *Invest Ophthalmol Vis Sci*. 2012;53:7510–7519.
13. Tian J, Marziliano P, Baskaran M, Tun TA, Aung T. Automatic segmentation of the choroid in enhanced depth imaging optical coherence tomography images. *Biomed Opt Express*. 2013;4:397–411.
14. Naz S, Ahmed A, Akram MU, Khan SA. Automated segmentation of RPE layer for the detection of age macular degeneration using OCT images. Paper presented at the 2016 Sixth International Conference on Image Processing Theory, Tools and Applications (IPTA). Available at: http://ieeexplore.ieee.org/document/7821033/. Accessed June 22, 2017.
15. Mayer MA, Hornegger J, Mardin CY, Tornow RP. Retinal nerve fiber layer segmentation on FD-OCT scans of normal subjects and glaucoma patients. *Biomed Opt Express*. 2010;1:1358–1383.
16. Mansberger SL, Menda SA, Fortune BA, Gardiner SK, Demirel S. Automated segmentation errors when using optical coherence tomography to measure retinal nerve fiber layer thickness in glaucoma. *Am J Ophthalmol*. 2017;174:1–8.
17. Koozekanani D, Boyer K, Roberts C. Retinal thickness measurements from optical coherence tomography using a Markov boundary model. *IEEE Trans Med Imaging*. 2001;20: 900–916.
18. Kajić V, Esmaeelpour M, Glittenberg C, et al. Automated three-dimensional choroidal vessel segmentation of 3D 1060 nm OCT retinal data. *Biomed Opt Express*. 2013;4:134–150.
19. Cabrera Fernandez D, Salinas HM, Puliafito CA. Automated detection of retinal layer structures on optical coherence tomography images. *Opt Express*. 2005;13:10200–10216.
20. Al-Diri B, Hunter A, Steel D. An active contour model for segmenting and measuring retinal vessels. *IEEE Trans Med Imaging*. 2009;28:1488–1497.
21. Girard MJ, Strouthidis NG, Mari JM, Ethier CR. Shadow removal, contrast and connective tissue visibility enhancement in optical coherence tomography images of the human optic nerve head. *Invest Ophthalmol Vis Sci*. 2011;52:1333–1333.
22. Lavinsky F, Lavinsky D. Novel perspectives on swept-source optical coherence tomography. *Int J Retina Vitreous*. 2016;2: 25.
23. Shields CL, Kaliki S, Rojanaporn D, Ferenczy SR, Shields JA. Enhanced depth imaging optical coherence tomography of small choroidal melanoma: comparison with choroidal nevus. *Arch Ophthalmol*. 2012;130:850–856.
24. Adhi M, Duker JS. Optical coherence tomography—current and future applications. *Curr Opin Ophthalmol*. 2013;24: 213–221.
25. Spaide RF, Koizumi H, Pozzoni MC. Enhanced depth imaging spectral-domain optical coherence tomography. *Am J Ophthalmol*. 2008;146:496–500.
26. Strouthidis N, Mari JM, Park SC, Girard M. Enhancement of lamina cribrosa visibility in optical coherence tomography images using adaptive compensation. *Invest Ophthalmol Vis Sci*. 2013;54:2149–2149.
27. Mari JM, Aung T, Cheng CY, Strouthidis NG, Girard MJ. A digital staining algorithm for optical coherence tomography images of the optic nerve head. *Trans Vis Sci Tech*. 2017;6(1): 8.
28. Mari JM, Strouthidis NG, Park SC, Girard MJA. Enhancement of lamina cribrosa visibility in optical coherence tomography images using adaptive compensation: improving lamina cribrosa visibility in OCT images. *Invest Ophthalmol Vis Sci*. 2013;54:2238–2247.
29. Girard MJA, Tun TA, Husain R, et al. Lamina cribrosa visibility using optical coherence tomography: comparison of devices and effects of image enhancement techniques: lamina cribrosa visibility in OCT. *Invest Ophthalmol Vis Sci*. 2015; 56:865–874.
30. Havaei M, Davy A, Warde-Farley D, et al. Brain tumor segmentation with deep neural networks. *Med Image Anal*. 2017;35:18–31.
31. Hariharan B, Arbeláez P, Girshick R, Malik J. Simultaneous detection and segmentation. In: Fleet D, Pajdla T, Schiele B, Tuytelaars T, eds. In: *Computer Vision – ECCV 2014: 13th European Conference, Zurich, Switzerland, September 6–12, 2014, Proceedings, Part VII*. Cham: Springer International Publishing; 2014:297–312.
32. Lee CS, Tyring AJ, Deruyter NP, Wu Y, Rokem A, Lee AY. Deep-learning based, automated segmentation of macular edema in optical coherence tomography. Available at: https://www.biorxiv.org/content/early/2017/05/09/135640. Accessed June 22, 2017.
33. Roy AG, Sailesh Conjeti, Karri SPK, et al. ReLayNet: retinal layer and fluid segmentation of macular optical coherence tomography using fully convolutional network. Available at: https://arxiv.org/abs/1704.02161. Accessed June 22, 2017.
34. Venhuizen FG, van Ginneken B, Liefers B, et al. Robust total retina thickness segmentation in optical coherence tomography images using convolutional neural networks. *Biomed Opt Express* 2017;8:3292–3316.
35. Kingma D, Ba J. Adam: a method for stochastic optimization. Available at: https://arxiv.org/abs/1412.6980. Accessed June 22, 2017.
36. GitHub, Inc. fchollet/keras. Available at: https://github.com/fchollet/keras. Accessed June 22, 2017.
37. Martín Abadi AA, Agarwal A, Barham P, et al. TensorFlow: large-scale machine learning on heterogeneous systems. Available at: http://tensorflow.org/. Accessed June 22, 2017.
38. Gardiner SK, Boey PY, Yang H, Fortune B, Burgoyne CF, Demirel S. Structural measurements for monitoring change in glaucoma: comparing retinal nerve fiber layer thickness with minimum rim width and area. *Invest Ophthalmol Vis Sci*. 2015;56:6886–6891.
39. Kim YW, Jeoung JW, Kim DW, et al. Clinical assessment of lamina cribrosa curvature in eyes with primary open-angle glaucoma. *PLoS One*. 2016;11:e0150260.
40. Thakku SG, Tham YC, Baskaran M, et al. A global shape index to characterize anterior lamina cribrosa morphology and its determinants in healthy Indian eyes. *Invest Ophthalmol Vis Sci*. 2015;56:3604–3614.

41. Manjunath V, Shah H, Fujimoto JG, Duker JS. Analysis of peripapillary atrophy using spectral domain optical coherence tomography. *Ophthalmology*. 2011;118:531–536.
42. Hayashi K, Tomidokoro A, Lee KYC, et al. Spectral-domain optical coherence tomography of β-zone peripapillary atrophy: influence of myopiaand glaucoma. *Invest Ophthalmol Vis Sci*. 2012;53:1499–1505.
43. Lee EJ, Kim T-W, Kim M, Girard MJA, Mari JM, Weinreb RN. Recent structural alteration of the peripheral lamina cribrosa near the location of disc hemorrhage in glaucoma. Recent alteration of LC and disc hemorrhage. *Invest Ophthalmol Vis Sci*. 2014;55:2805–2815.
44. Yang H, Williams G, Downs JC, et al. Posterior (outward) migration of the lamina cribrosa and early cupping in monkey experimental glaucoma. *Invest Ophthalmol Vis Sci*. 2011;52:7109–7121.
45. You JY, Park SC, Su D, Teng CC, Liebmann JM, Ritch R. Focal lamina cribrosa defects associated with glaucomatous rim thinning and acquired pits. *JAMA Ophthalmol*. 2013;131:314–320.
46. Kiumehr S, Park S, Dorairaj S, et al. In vivo evaluation of focal lamina cribrosa defects in glaucoma. *Arch Ophthalmol*. 2012;130:552–559.
47. Reis AS, O'Leary N, Yang H, et al. Influence of clinically invisible, but optical coherence tomography detected, optic disc margin anatomy on neuroretinal rim evaluation. *Invest Ophthalmol Vis Sci*. 2012;53:1852–1860.
48. Lee SH, Yu D-A, Kim T-W, Lee EJ, Girard MJA, Mari JM. Reduction of the lamina cribrosa curvature after trabeculectomy in glaucoma. *Invest Ophthalmol Vis Sci*. 2016;57:5006–5014.
49. Fazio MA, Grytz R, Morris JS, Bruno L, Girkin CA, Downs JC. Human scleral structural stiffness increases more rapidly with age in donors of African descent compared to donors of European descent. *Invest Ophthalmol Vis Sci*. 2014;55:7189–7198.
50. Garvin MK, Abramoff MD, Kardon R, Russell SR, Wu X, Sonka M. Intraretinal layer segmentation of macular optical coherence tomography images using optimal 3-D graph search. *IEEE Trans Med Imaging*. 2008;27:1495–1505.
51. Fang L, Cunefare D, Wang C, Guymer RH, Li S, Farsiu S. Automatic segmentation of nine retinal layer boundaries in OCT images of non-exudative AMD patients using deep learning and graph search. *Biomed Opt Express*. 2017;8:2732.
52. Song YJ, Kim YK, Jeoung JW, Park KH. Assessment of open-angle glaucoma peripapillary and macular choroidal thickness using swept-source optical coherence tomography (SS-OCT). *PLoS One*. 2016;11:e0157333.
53. Alonso-Caneiro D, Read SA, Collins MJ. Automatic segmentation of choroidal thickness in optical coherence tomography. *Biomed Opt Express*. 2013;4:2795–2812.
54. Duan L, Yamanari M, Yasuno Y. Automated segmentation of chorio-scleral interface using polarization sensitive optical coherence tomography. *Invest Ophthalmol Vis Sci*. 2012;53:4088–4088.
55. Nadler Z, Wang B, Wollstein G, et al. Automated lamina cribrosa microstructural segmentation in optical coherence tomography scans of healthy and glaucomatous eyes. *Biomed Opt Express*. 2013;4:2596–2608.
56. Tan MH, Ong SH, Thakku SG, Cheng C-Y, Aung T, Girard M. Automatic feature extraction of optical coherence tomography for lamina cribrosa detection. *J Image Graph*. 2015;3:102–106.
57. Campbell IC, Coudrillier B, Mensah J, Abel RL, Ethier CR. Automated segmentation of the lamina cribrosa using Frangi's filter: a novel approach for rapid identification of tissue volume fraction and beam orientation in a trabeculated structure in the eye. *J R Soc Interface*. 2015;12:20141009.
58. Chen JJ, Kardon RH. Avoiding clinical misinterpretation and artifacts of optical coherence tomography analysis of the optic nerve, retinal nerve fiber layer, and ganglion cell layer. *J Neuroophthalmol*. 2016;36:417–438.
59. Mari J-M, Aung T, Cheng C-Y, Strouthidis NG, Girard MJA. A digital staining algorithm for optical coherence tomography images of the optic nerve head. *Trans Vis Sci Tech*. 2017;6(1):8.
60. Strouthidis NG, Grimm J, Williams GA, Cull GA, Wilson DJ, Burgoyne CF. A comparison of optic nerve head morphology viewed by spectral domain optical coherence tomography and by serial histology. *Invest Ophthalmol Vis Sci*. 2010;51:1464–1474.
61. Shelhamer E, Long J, Darrell T. Fully convolutional networks for semantic segmentation. *IEEE Trans Pattern Anal Mach Intell*. 2017;39:640–651.
62. Brox ORaPFaT. U-Net: convolutional networks for biomedical image segmentation. *Med Image Comput Comput Assist Interv*. 2015;9351:234–241.
63. Zhang W, Li R, Deng H, et al. Deep convolutional neural networks for multi-modality isointense infant brain image segmentation. *NeuroImage* 2015;108:214–224.
64. Li R, Zhang W, Suk H-I, et al. Deep learning based imaging data completion for improved brain disease diagnosis. *Med Image Comput Comput Assist Interv*. 2014;17:305–312.

Investigative Ophthalmology & Visual Science