

OPEN

# A Deep Learning Approach to Denoise Optical Coherence Tomography Images of the Optic Nerve Head

Sripad Krishna Devalla<sup>1</sup>, Giridhar Subramanian<sup>1</sup>, Tan Hung Pham<sup>1,2</sup>, Xiaofei Wang<sup>1,3</sup>, Shamira Perera<sup>2,4</sup>, Tin A. Tun<sup>1,2</sup>, Tin Aung<sup>2,4</sup>, Leopold Schmetterer<sup>2,6,7,8</sup>, Alexandre H. Thiéry<sup>5</sup> & Michaël J. A. Girard<sup>1,2</sup>

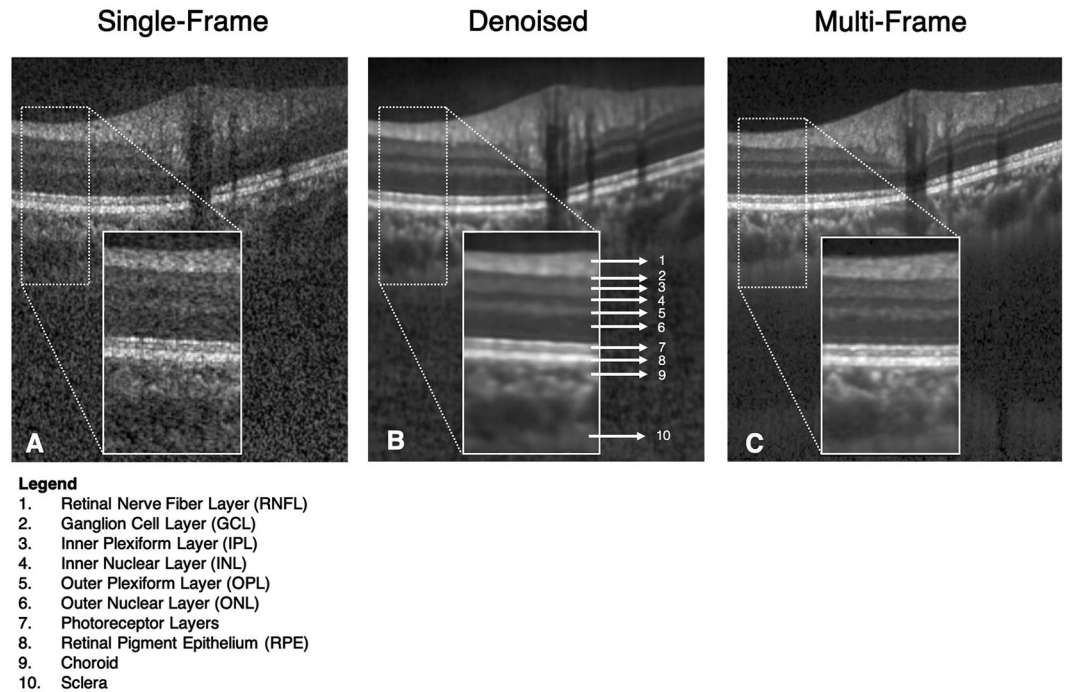
Optical coherence tomography (OCT) has become an established clinical routine for the *in vivo* imaging of the optic nerve head (ONH) tissues, that is crucial in the diagnosis and management of various ocular and neuro-ocular pathologies. However, the presence of speckle noise affects the quality of OCT images and its interpretation. Although recent frame-averaging techniques have shown to enhance OCT image quality, they require longer scanning durations, resulting in patient discomfort. Using a custom deep learning network trained with 2,328 'clean B-scans' (multi-frame B-scans; signal averaged), and their corresponding 'noisy B-scans' (clean B-scans + Gaussian noise), we were able to successfully denoise 1,552 unseen single-frame (without signal averaging) B-scans. The denoised B-scans were qualitatively similar to their corresponding multi-frame B-scans, with enhanced visibility of the ONH tissues. The mean signal to noise ratio (SNR) increased from  $4.02 \pm 0.68$  dB (single-frame) to  $8.14 \pm 1.03$  dB (denoised). For all the ONH tissues, the mean contrast to noise ratio (CNR) increased from  $3.50 \pm 0.56$  (single-frame) to  $7.63 \pm 1.81$  (denoised). The mean structural similarity index (MSSIM) increased from  $0.13 \pm 0.02$  (single frame) to  $0.65 \pm 0.03$  (denoised) when compared with the corresponding multi-frame B-scans. Our deep learning algorithm can denoise a single-frame OCT B-scan of the ONH in under 20 ms, thus offering a framework to obtain superior quality OCT B-scans with reduced scanning times and minimal patient discomfort.

In recent years, optical coherence tomography (OCT) imaging has become a well-established clinical tool for assessing optic nerve head (ONH) tissues, and for monitoring many ocular<sup>1,2</sup> and neuro-ocular pathologies<sup>3</sup>. However, despite several advancements in OCT technology<sup>4</sup>, the quality of B-scans is still hampered by speckle noise<sup>5–11</sup>, low signal strength<sup>12</sup>, blink<sup>12,13</sup> and motion artefacts<sup>12,14</sup>.

Specifically, the granular pattern of speckle noise deteriorates the image contrast, making it difficult to resolve small and low-intensity structures (e.g., sub-retinal layers)<sup>5–7</sup>, thus affecting the clinical interpretation of OCT data. Also, poor image contrast can lead to automated segmentation errors<sup>14–16</sup>, and incorrect tissue thickness estimation<sup>17</sup>, potentially affecting clinical decisions. For instance, segmentation errors for the retinal nerve fiber layer (RNFL) thickness can lead to over/under estimation of glaucoma<sup>18</sup>.

<sup>1</sup>Ophthalmic Engineering & Innovation Laboratory, Department of Biomedical Engineering, Faculty of Engineering, National University of Singapore, Singapore, Singapore. <sup>2</sup>Singapore Eye Research Institute, Singapore National Eye Centre, Singapore, Singapore. <sup>3</sup>Beijing Advanced Innovation Center for Biomedical Engineering, School of Biological Science and Medical Engineering, Beihang University, Beijing, China. <sup>4</sup>Duke-NUS Graduate Medical School, Singapore, Singapore. <sup>5</sup>Department of Statistics and Applied Probability, National University of Singapore, Singapore, Singapore. <sup>6</sup>Nanyang Technological University, Jurong West, Singapore. <sup>7</sup>Department of Clinical Pharmacology, Medical University of Vienna, Vienna, Austria. <sup>8</sup>Center for Medical Physics and Biomedical Engineering, Medical University of Vienna, Vienna, Austria. Sripad Krishna Devalla, Giridhar Subramanian, Alexandre H. Thiéry and Michaël J. A. Girard contributed equally. Correspondence and requests for materials should be addressed to A.H.T. (email: [a.h.thiery@nus.edu.sg](mailto:a.h.thiery@nus.edu.sg)) or M.J.A.G. (email: [mgirard@nus.edu.sg](mailto:mgirard@nus.edu.sg))

## Qualitative Analysis



**Figure 1.** Single-frame (A), denoised (B), and multi-frame (C) B-scans for a healthy subject are shown. The denoised B-scan can be observed to be qualitatively similar to its corresponding multi-frame B-scan. Specifically, the visibility of the retinal layers, and choroid, and lamina cribrosa were prominently improved. Sharp and clear boundaries were also obtained for retinal layers, and the choroid-scleral interface.

Currently, there exist many hardware<sup>19–27</sup> and software schemes<sup>27–29</sup> to denoise OCT B-scans. Hardware approaches offer robust noise suppression through frequency compounding<sup>24–27</sup> and multi-frame averaging (spatial compounding)<sup>19–23</sup>. While multi-frame averaging techniques have shown to enhance image quality and presentation<sup>28,29</sup>, they are sensitive to registration errors<sup>29</sup>, and require longer scanning times<sup>30</sup>. Moreover, elderly patients often face discomfort and strain<sup>31</sup>, when they remain fixated for long durations<sup>31,32</sup>. Software techniques, on the other hand, attempt to denoise through numerical algorithms<sup>5–11</sup> or filtering techniques<sup>33–35</sup>. However, registration errors<sup>36</sup>, computational complexity<sup>5,37–39</sup>, and sensitivity to choice of parameters<sup>40</sup> limit their usage in the clinic.

While deep learning has shown promising segmentation<sup>41–44</sup>, classification<sup>45–47</sup>, and denoising<sup>48–50</sup> applications in the field of medical imaging for modalities such as magnetic resonance imaging (MRI), its application to OCT imaging is still in its infancy<sup>51–62</sup>. Although recent deep learning studies have shown successful segmentation<sup>51–59</sup> and classification applications<sup>60–62</sup> in OCT imaging, to the best of our knowledge no study exists yet to assess the success of denoising OCT B-scans.

In this study, we propose a deep learning approach to denoise OCT B-scans. We aimed to obtain multi-frame quality B-scans (i.e. signal-averaged) from single-frame (without signal averaging) B-scans of the ONH. We hope to offer a denoising framework to obtain superior quality B-scans, with reduced scanning duration and minimal patient discomfort.

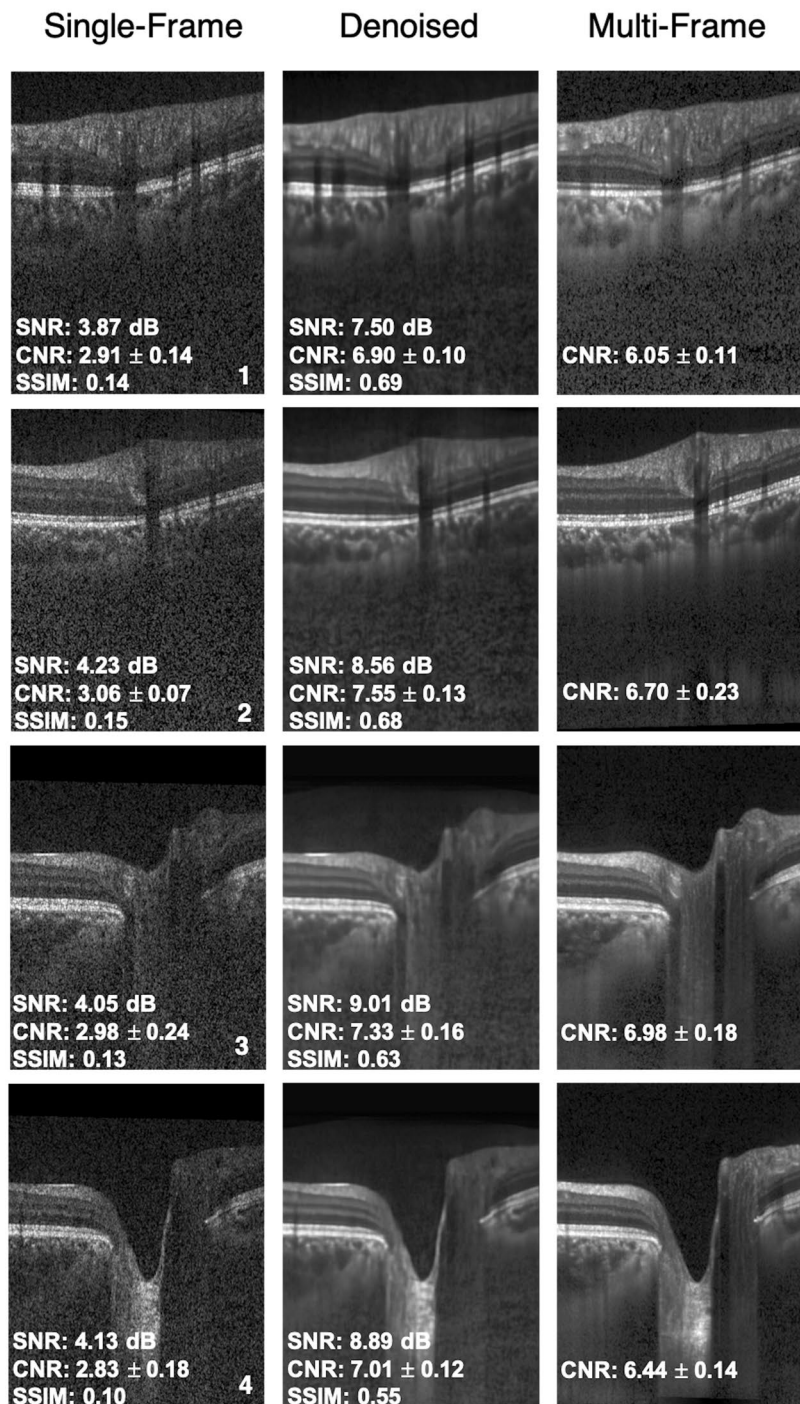
### Results

When trained on 23,280 pairs (2,328 B-scans after extensive data augmentation resulted in 23,280 B-scans) of ‘clean’ (multi-frame) and their corresponding ‘noisy’ B-scans (clean B-scans + Gaussian noise), our deep learning network was able to successfully denoise the unseen single-frame B-scans. An independent test set of 1,552 single-frame B-scans was used to evaluate the denoising performance of the proposed network qualitatively and quantitatively.

**Denoising performance – qualitative analysis.** The single-frame, denoised and multi-frame B-scan for a healthy subject can be found in Fig. 1. In all the cases, the denoised B-scans were qualitatively similar to their corresponding multi-frame B-scans (Fig. 2). Specifically, in all the denoised B-scans, we observed no deep learning induced image artifacts, and the overall visibility of all the ONH tissues were prominently enhanced (Fig. 2; 2<sup>nd</sup> column).

**Denoising performance – quantitative analysis.** When evaluated on the independent test set of 1,552 B-scans, on average, we observed a two-fold increase in SNR upon denoising. Specifically, the mean SNR for the

## Qualitative Analysis



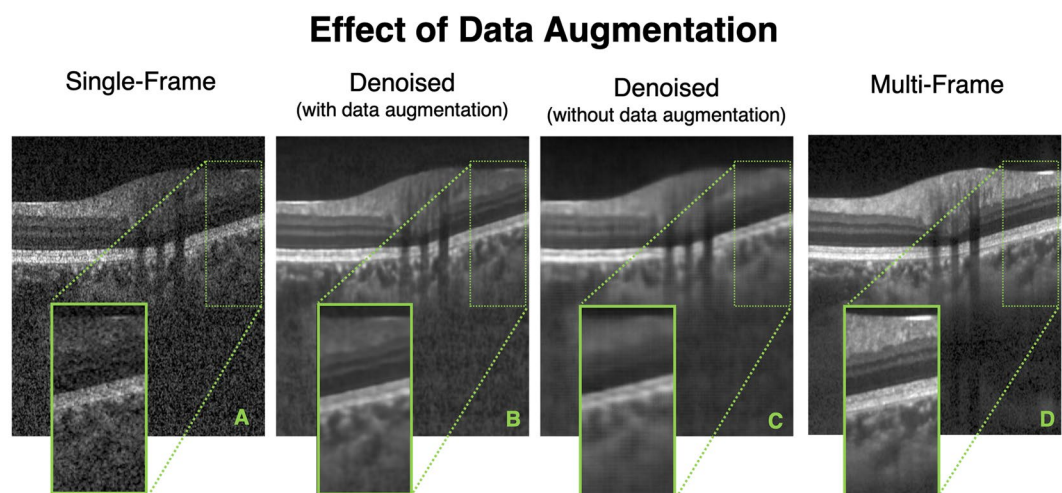
**Figure 2.** Single-frame, denoised and multi-frame B-scans for four healthy subjects (1–4) are shown. The signal to noise ratio (SNR), contrast to noise ratio (CNR; mean of all tissues) and the structural similarity index (SSIM) for the respective B-scans are shown as well. In all cases, the denoised B-scans (2<sup>nd</sup> column) were consistently similar (qualitatively) to their corresponding multi-frame B-scans (3<sup>rd</sup> column).

unseen single-frame/denoised B-scans were:  $4.02 \pm 0.68$  dB/ $8.14 \pm 1.03$  dB, respectively, when computed against their respective multi-frame B-scans. The two-fold reduction in the noise levels resulted in the enhanced overall visibility of the ONH tissues.

In all cases, the multi-frame B-scans always offered a higher CNR compared to their corresponding single-frame B-scans. Further, the denoised B-scans consistently offered a higher CNR compared to the single-frame B-scans, for all tissues (Table 1). Specifically, the mean CNR (mean of all tissues) increased from

Tissue	Single-frame	Denoised	Multi-frame
RNFL	2.97 ± 0.42	7.28 ± 0.63	5.18 ± 0.76
GCL + IPL	3.83 ± 0.43	12.09 ± 4.22	11.62 ± 1.85
All other retinal layers	2.71 ± 0.33	5.61 ± 1.46	4.62 ± 0.86
RPE	5.62 ± 0.72	9.25 ± 2.25	8.10 ± 1.44
Choroid	2.99 ± 0.43	5.99 ± 0.45	5.75 ± 0.63
Sclera	2.42 ± 0.39	6.40 ± 1.68	6.00 ± 0.96
LC	4.02 ± 1.23	6.81 ± 1.99	6.46 ± 1.81

**Table 1.** Mean CNR for all ONH tissues computed for the single-frame, denoised and multi-frame B-scans.



**Figure 3.** Single-frame (A), denoised (with data augmentation) (B), denoised (without data augmentation) (C), and multi-frame (D) B-scans for a healthy subject are shown. The denoised B-scan (B) obtained from a network trained with data augmentation can be observed to be qualitatively similar to its corresponding multi-frame B-scan (D). However, when trained with limited training data (without data augmentation), although the network is able to reduce the speckle noise primarily, the denoised B-scan is over-smoothed (C) with smudged and unclear tissue boundaries.

3.50 ± 0.56 (single-frame) to 7.63 ± 1.81 (denoised). For each tissue, mean CNR in a single-frame, denoised and multi-frame B-scan can be found in Table 1. The increased CNR values for each ONH tissue in the denoised images implied sharper and improved visibility of the tissue boundaries.

Finally, on average, our denoising approach offered a five-fold increase in MSSIM. Specifically, the mean MSSIM for the single-frame/denoised B-scans were: 0.13 ± 0.02/0.65 ± 0.03, when computed against their respective multi-frame B-scans. Thus, the denoised B-scans were five-times structural more similar (compared to the single-frame B-scans) to the multi-frame B-scans.

**Denoising performance – effect of data augmentation.** When trained without data augmentation, in all the test cases, while the network was able to reduce the speckle noise primarily, the denoised B-scans appeared to be over-smoothed (blurred). Specifically, the tissue boundaries (especially for the retinal layers) appeared smudged and unclear (Fig. 3).

While the network trained with the baseline dataset (without data augmentation) offered a slightly higher mean SNR of 10.11 ± 2.13 dB, (vs. with data augmentation: 8.14 ± 1.03 dB), the higher SNR can be attributed to the over-smoothed (blur) denoised B-scans, and not the improved image quality and reliability.

Further, we would like to clarify that we were unable to obtain reliable CNR values for the individual ONH tissues in all the denoised images (network trained without data augmentation) due to smudged and unclear tissue boundaries.

Finally, while the network trained with the baseline dataset resulted in denoised B-scans with nearly two-fold increased mean MSSIM (0.25 ± 0.07; ‘noisy’ B-scans mean MSSIM: 0.13 ± 0.02), the use of data augmentation helped retrieve structural information (mean MSSIM: 0.65 ± 0.03; five-fold increase compared to ‘noisy’ B-scans).

Overall, we observed that a 10-fold increase in the dataset size through extensive data augmentation did improve the overall quality and reliability of the denoised B-scans.

**Denoising performance: clinical reliability.** For all the three structural parameters, there were no significant ( $p > 0.05$ ) differences (means) in the measurements when obtained from the denoised or the multi-frame radial B-scans.

The percentage errors (denoised vs. multi-frame radial B-scans; mean  $\pm$  standard deviation) in the measurements of the p-RNFL, the p-GCCT, and the p-CT were  $3.07 \pm 0.75\%$ ,  $2.95 \pm 1.02\%$ , and  $3.90 \pm 2.85\%$  respectively.

## Discussion

In this study, we present a custom deep learning approach to denoise single-frame OCT B-scans of the ONH. When trained with the ‘clean’ (multi-frame) and the corresponding ‘noisy’ B-scans, our network denoised unseen single-frame B-scans. The proposed network leveraged on the inherent advantages of U-Net, residual learning, and dilated convolutions<sup>58</sup>. Further, the multi-scale hierarchical feature extraction<sup>63</sup> pathway helped the network recover ONH tissue boundaries degraded by speckle noise. Having successfully trained, tested and validated our network on 1,552 single-frame OCT B-scans of the ONH, we observed a consistently higher SNR and CNR for all ONH tissues, and a consistent five-fold increase in MSSIM in all the denoised B-scans. Thus, we may be able to offer a robust deep learning framework to obtain superior quality OCT B-scans with reduced scanning duration and minimal patient discomfort.

Using the proposed network, we obtained denoised B-scans that were qualitatively similar to their corresponding multi-frame B-scans (Figs 1 and 2), owing to the reduction in noise levels. The mean SNR for the denoised B-scans was  $8.14 \pm 1.03$  dB, a two-fold improvement (reduction in noise level) from  $4.02 \pm 0.68$  that was obtained for the single-frame B-scans. Given the significance of the neural (retinal layers)<sup>64–68</sup> and connective tissues (sclera and LC)<sup>69–73</sup>, in ocular pathologies such as glaucoma<sup>2</sup>, and age-related macular degeneration<sup>74</sup>, their enhanced visibility is critical in a clinical setting. Furthermore, reduced noise levels would likely increase the robustness of aligning/registration algorithms used to monitor structural changes over time<sup>17</sup>. This is crucial for the management of multiple ocular pathologies<sup>75,76</sup>.

In denoised B-scans (vs single-frame B-scans), we consistently observed higher CNRs. Our approach enhanced the visibility of small (e.g. RPE and photoreceptors) and low-intensity tissues (e.g. GCL and IPL; Fig. 2: 2<sup>nd</sup> column). For all tissues, the mean CNR increased from  $3.50 \pm 0.56$  (single-frame) to  $7.63 \pm 1.81$  (denoised). Since existing automated segmentation algorithms rely on high contrast, we believe that our approach could potentially reduce the likelihood of segmentation errors that are relatively common in commercial algorithms<sup>14–16,77</sup>. For instance, the incorrect segmentation of the RNFL can lead to inaccurate thickness measurements, leading to under-/over- estimation of the glaucomatous damage<sup>18</sup>. By using the denoising framework as a precursor to automated segmentation/thickness measurement, we could increase the reliability<sup>78</sup> of such clinical tools.

Upon denoising, we observed a five-fold increase in MSSIM (single-frame/denoised:  $0.13 \pm 0.02/0.65 \pm 0.03$ ), when validated against the multi-frame B-scans. The preservation of features and structural information plays an important role in accurately measuring cellular level disruption to determine retinal pathology. For instance, the measurement of the ellipsoid zone (EZ) disruption<sup>79</sup> provides an insight into the photoreceptor structure, that is significant in pathologies such as diabetic retinopathy<sup>80</sup>, macular hole<sup>81</sup>, macular degeneration<sup>82</sup>, and ocular trauma<sup>83</sup>. Existing multi-frame averaging techniques<sup>29</sup> significantly enhance and preserve the integrity of the structural information by suppressing speckle noise<sup>30,39–41</sup>. However, they are limited by a major clinical challenge: the inability of the patients to remain fixated for long scanning times<sup>31,32</sup>, and the resultant discomfort<sup>31</sup>.

In this study, we are proposing a methodology to significantly reduce scanning time while enhancing OCT signal quality. In our healthy subjects, it took on average 3.5 min to capture a ‘clean’ (multi frame) volume, and 25 s for a ‘noisy’ (single frame) volume. Since we can denoise a single B-scan in 20 ms (or 2 s for a volume of 97 B-scans), this means that we can theoretically generate a denoised OCT volume in about 27 seconds (=time of acquisition of the ‘noisy’ volume [25 s] + denoising processing [2 s]). Thus, we may be able to drastically reduce the scanning duration by more than 7 folds, while maintaining superior image quality.

Besides speckle noise, patient dependent factors such as cataract<sup>84–87</sup> and/or lack of tear film in dry eyes can significantly diminish OCT scan quality<sup>12,84–88</sup>. While lubricating eye drops and frequent blinking can instantly improve image quality for patients with corneal drying<sup>88,89</sup>, the detrimental effects of cataract on OCT image quality might be reduced only if cataract surgery is performed<sup>12,84,85</sup>. Moreover, pupillary dilation may be needed especially in subjects with small pupil sizes to obtain acceptable quality B-scans<sup>12,90</sup>, which is highly crucial in the monitoring of glaucoma<sup>90</sup>. Pupillary dilation is also time consuming and may cause patient discomfort<sup>91</sup>. It is plausible that the proposed framework, when extended, could be a solution to the afore-mentioned factors that limit image quality, avoiding the need for any additional clinical procedure.

In this study, several limitations warrant further discussion. First, the proposed network was trained and tested only on B-scans from one device (Spectralis). Every commercial OCT device has its own proprietary algorithm to pre-process the raw OCT data, potentially presenting a noise distribution different from what our network was trained with. Hence, we are unsure of our network’s performance on other devices. Nevertheless, we offer a proof of concept which could be validated by other groups on multiple commercial OCT devices.

Second, we were unable to train our network with a speckle noise model representative of the Spectralis device. Such a model is currently not provided by the manufacturer and would be extremely hard to reverse-engineer because information about all pre- and post-processing done to the OCT signal is also not provided. While there exist a number of OCT denoising studies that assume a Rayleigh<sup>8</sup>/Generalized Gamma distribution to describe speckle noise<sup>39</sup>, we observed that they were ill-suited for our network. From our experiments, the best denoising performance was obtained when our network was trained with a simple Gaussian noise model ( $\mu = 0$ ,  $\sigma = 1$ ). It is possible that a thorough understanding of the raw noise distribution prior to the custom pre-processing on the OCT device could improve the performance of our network. We aim to test this hypothesis with a custom-built OCT system in our future works.

Third, while we have discussed the need for reliable clinical information from poor quality OCT scans, that could be critical for the diagnosis and management of ocular pathology (e.g., glaucoma), we have yet to test the networks' performance on pathological B-scans.

Fourth, we observed that the CNR was higher for the denoised B-scans than for the corresponding multi-frame B-scans. This could be attributed to over-smoothing (or blurring) of tissue textures that was consistently present in the denoised B-scans. We are currently exploring other deep learning techniques to improve the B-scan sharpness that is lost during denoising.

Fifth, we were unable to provide further validation of our algorithm by comparing our outputs to histology data. Such a validation would be extremely difficult, as one would need to first image a human ONH with OCT, process with histology, and register both datasets. Furthermore, while we believe our algorithm is able to restore tissue texture accurately (when comparing denoised B-scans with multi-frame B-scans), an exact validation of our approach is not possible. Long fixation times in obtaining the multi-frame B-scans lead to subtle motion artifacts (eye movements caused by microsaccades or unstable fixation)<sup>92</sup>, displaced optic disc center<sup>93</sup>, and axial misalignment<sup>12</sup>, causing minor registration errors between the single-frame and multi-frame B-scans, thus preventing an exact comparison between the denoised B-scans and the multi-frame B-scans.

Finally, although we observed no significant differences in the measurements of the p-RNFLT, p-GCCT and the p-CT when measured on the denoised or the multi-frame B-scans, it must be noted that the measurements were obtained on a testing cohort of limited (8 subjects) and healthy subjects only. Further studies on larger and diverse (presence of pathology) cohorts are necessary to assert and robustly validate the clinical relevance of the proposed technique.

In conclusion, we have developed a custom deep learning approach to denoise single-frame OCT B-scans. With the proposed network, we were able to denoise a single-frame OCT B-scan in under 20 ms. We hope that the proposed framework could resolve the current trade-off in obtaining reliable and superior quality scans, with reduced scanning times and minimal patient discomfort. Finally, we believe that our approach may be helpful for low-cost OCT devices, whose noisy B-scans may be enhanced by artificial intelligence (as opposed to expensive hardware) to the same quality as in current commercial devices.

## Methods

**Patient recruitment.** A total of 20 healthy subjects were recruited at the Singapore National Eye Centre. All subjects gave written informed consent. This study was approved by the institutional review board of the SingHealth Centralized Institutional Review Board and adhered to the tenets of the Declaration of Helsinki. The inclusion criteria for healthy subjects were: an intraocular pressure (IOP) less than 21 mmHg, and healthy optic nerves with a vertical cup-disc ratio (VCDR) less than or equal to 0.5.

**Optical coherence tomography imaging.** The subjects were seated and imaged under dark room conditions by a single operator (TAT). A spectral-domain OCT (Spectralis, Heidelberg Engineering, Heidelberg, Germany) was used to image both eyes of each subject. Each OCT volume consisted of 97 horizontal B-scans (32- $\mu$ m distance between B-scans; 384 A-scans per B-scan), covering a rectangular area of  $15^\circ \times 10^\circ$  centered on the ONH. For each eye, single-frame (without signal averaging), and multi-frame (75x signal averaging) volume scans were obtained. Enhanced depth imaging (EDI)<sup>94</sup> and eye tracking<sup>92,95</sup> modalities were used during the acquisition. From all the subjects, we obtained a total of 3,880 B-scans for each type of scan (single-frame or multi-frame).

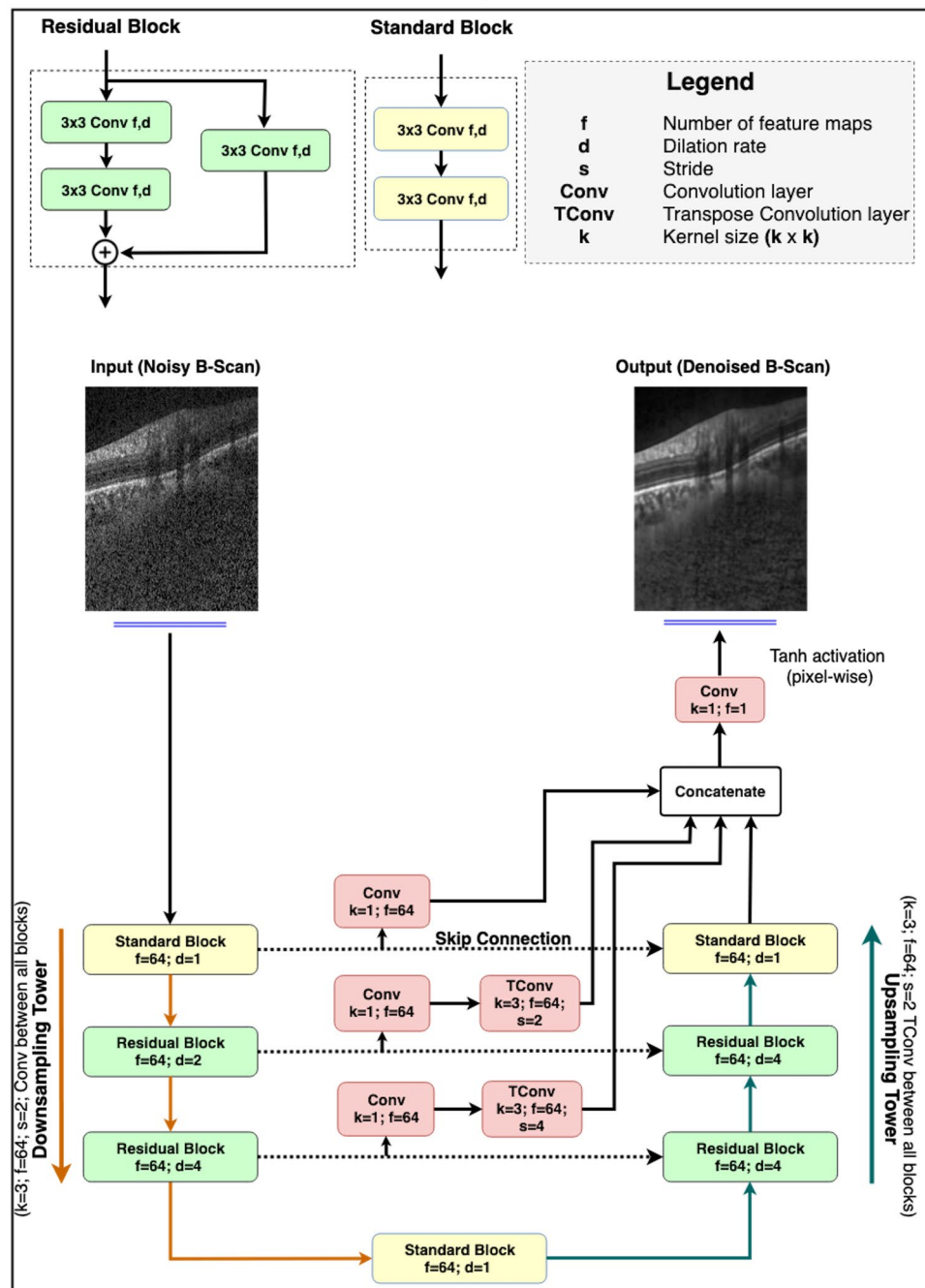
**Volume registration.** The multi-frame volumes were reoriented to align with the single-frame volumes through rigid translation/rotation transformations using 3D software (Amira, version 5.6; FEI). This registration was performed using a voxel-based algorithm that maximized mutual information between two volumes<sup>96</sup>. Registration was essential to quantitatively validate the corresponding regions between the denoised and multi-frame B-scans. Note that Spectralis follow-up mode was not used in this study. Although the follow-up mode allows a new scanning of the same area by identifying previous scan locations, in many cases, it can distort B-scans and thus provide unrealistic tissue structures in the new scan.

**Deep learning based denoising.** In this study, we developed a fully-convolutional neural network, inspired by our earlier DRUNET architecture<sup>58</sup> to denoise single-frame OCT B-scans of the ONH. It leverages on the inherent advantages of U-Net<sup>97</sup>, residual learning<sup>98</sup>, dilated convolutions<sup>99</sup>, and multi-scale hierarchical feature extraction<sup>63</sup> to obtain multi-frame quality B-scans. Briefly, the U-Net and its skip connections helped the network learn both the local (tissue texture) and contextual information (spatial arrangement of tissues). The contextual information was further exploited using dilated convolution filters. Residual connections improved the flow of the gradient information through the network, and multi-scale hierarchical feature extraction helped restore tissue boundaries in the B-scans.

**Network architecture.** The proposed network consisted of two types of feature extraction units: (1) standard block; and (2) residual block. While each block consisted of two subsequent dilated convolutional layers (64 filters; size =  $3 \times 3$ ) for feature extraction, the residual block consisted of an additional  $3 \times 3$  convolutional layer (identity connection)<sup>98</sup> that improved the flow of the gradient information throughout the depth of the network. The use of dilated convolution filters<sup>99</sup> helped the network better understand the contextual information (i.e., the spatial arrangement of tissues), that is crucial for obtaining reliable tissue boundaries.

Inspired by the U-Net<sup>97</sup>, the network was composed of a downsampling and an upsampling tower, connected via skip-connections (Fig. 4). By sequentially halving the dimensionality of the B-scan after each feature extraction unit, the downsampling tower extracted the contextual information (i.e., the spatial arrangement of tissues).

## Deep Learning Architecture

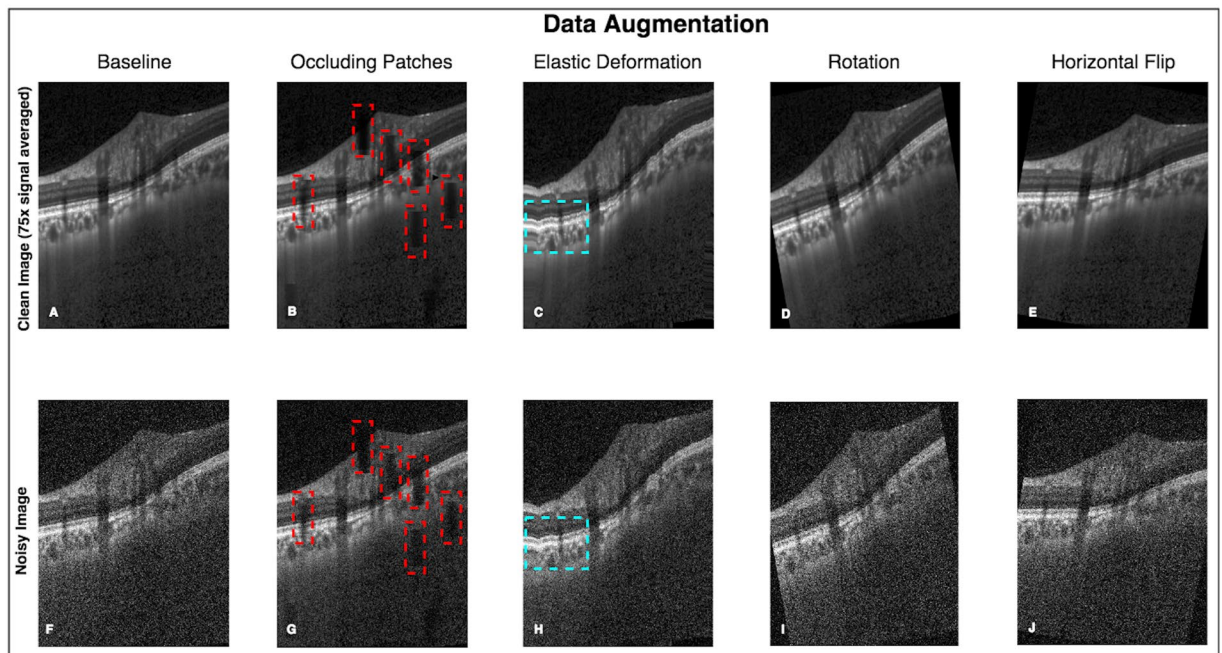


**Figure 4.** The architecture comprised of two towers: (1) A downsampling tower – to capture the contextual information (i.e., spatial arrangement of the tissues), and (2) an upsampling tower – to capture the local information (i.e., tissue texture). Each tower consisted of two blocks: (1) a standard block, and (2) a residual block. The latent space was implemented as a standard block. The multi-scale hierarchical feature extraction unit helped better recover tissue edges eroded by speckle noise. The network consisted of 900 k trainable parameters.

On the other hand, in the process of sequentially restoring the B-scan to its original dimensions, the upsampling tower extracted local information (i.e., tissue texture). Skip connections between both towers helped the network learn the extracted contextual and local information jointly.

In the downsampling tower, an input B-scan (size:  $496 \times 384$ ) was first passed on to a standard block (dilation rate: 1) followed by two residual blocks (dilation rate: 2 and 4, respectively). A convolution layer (64 filters; size =  $3 \times 3$ ; stride = 2) after every block sequentially halved the dimensionality of the B-scan.

A standard block (dilation rate: 1) was then used to transfer the contextual feature maps from the downsampling to the upsampling tower.



**Figure 5.** An exhaustive offline data augmentation was done to circumvent the scarcity of training data. (A–E) represent the original and the data augmented ‘clean’ B-scans (multi-frame). (F–J) Represent the same for the corresponding ‘noisy’ B-scans. The occluding patches (B,G; red boxes) were added to make the network robust in the presence of blood vessel shadows. Elastic deformations (C,H; cyan boxes) were used to make the network invariant to atypical morphologies<sup>104</sup>. A total of 23,280 B-scans of each type (clean/noisy) were generated from 2,328 baseline B-scans.

The upsampling tower consisted of two residual blocks (dilation rate: 4) and a standard block (dilation rate: 1). After each block, a transpose convolution layer (64 filters; size =  $3 \times 3$ ; stride = 2) was used to restore the B-scan sequentially to its original dimension.

The use of multi-scale hierarchical feature extraction<sup>63</sup> improved the recovery of tissue boundaries eroded by speckle noise in the single-frame B-scans. It was implemented by passing the feature maps at each downsampling level through a convolution layer (64 filters; size =  $1 \times 1$ ), followed by a transpose convolution layer (64 filters; size =  $3 \times 3$ ) to restore the original B-scan resolution. The restored maps were then concatenated with the output feature maps from the upsampling tower.

Finally, the concatenated feature maps were fed to the output convolution layer (1 filter; size =  $1 \times 1$ ), followed by pixel-wise hyperbolic tangent (tanh) activation to produce a denoised B-scan.

In both towers, all layers except the last output layer, were activated by an exponential linear unit (ELU)<sup>100</sup> function. In addition, in each residual block, the feature maps were batch normalized<sup>101</sup> and ELU activated before addition.

The proposed network comprised of 900,000 training parameters. The network was trained end-to-end using the Adam optimizer<sup>102</sup>, and we used the mean absolute error as loss function. We trained and tested the proposed network on an NVIDIA GTX 1080 founders edition GPU with CUDA v8.0 and cuDNN v5.1 acceleration. With the given hardware configuration, each single-frame B-scan was denoised under 20 ms.

**Training and testing of the network.** From the dataset of 3,880 B-scans, 2,328 of them (from both eyes of 12 subjects) were used as a part of the training dataset. The training set consisted of ‘clean’ B-scans and their corresponding ‘noisy’ versions. The ‘clean’ B-scans were simply the multi-frame (75x signal averaging) B-scans. The ‘noisy’ B-scans were generated by adding Gaussian noise ( $\mu = 0$ ,  $\sigma = 1$ ) to the respective ‘clean’ B-scans (Fig. 5).

The testing set consisted of 1,552 single-frame B-scans (from both eyes of 8 subjects) to be denoised. We ensured that the scans from the same subject weren’t used in both training and testing sets.

**Data augmentation.** An exhaustive offline data augmentation was done to circumvent the scarcity of training data. We used elastic deformations<sup>58,103</sup>, rotations (clockwise and anti-clockwise;  $10^\circ$ ), occluding patches<sup>58</sup>, and horizontal flipping for both ‘clean’ and ‘noisy’ B-scans. Briefly, elastic deformations were used to produce the combined effects of shearing and stretching in an attempt to make the network invariant to atypical morphologies (as seen in glaucoma<sup>104</sup>). Ten occluding patches of size  $60 \times 20$  pixels were added at random locations to non-linearly reduce (pixel intensities multiplied by a random factor between 0.2 and 0.8) the visibility of the ONH tissues. This was done to make the network invariant to blood vessel shadows that are common in OCT B-scans<sup>105</sup>. Note that a full description of our data augmentation approach can be found in our previous paper<sup>58</sup>.



Using data augmentation, we were able to generate a total of 23,280 ‘clean’ and 23,280 corresponding ‘noisy’ B-scans that were added to the training dataset. An example of data augmentation performed on a single ‘clean’ and corresponding ‘noisy’ B-scan is shown in Fig. 5.

**Denosing performance – qualitative analysis.** All denoised single-frame B-scans were manually reviewed by expert observers (S.K.D. & G.S.) and qualitatively compared against their corresponding multi-frame B-scans to assess the following: 1) presence of deep learning induced image artifacts in the denoised B-scans; and 2) overall visibility of the ONH tissues.

**Denosing performance – quantitative analysis.** The following image quality metrics were used to assess the denosing performance of the proposed algorithm: (1) signal to noise ratio (SNR); (2) contrast to noise ratio (CNR); and (3) mean structural similarity index measure (MSSIM)<sup>106</sup>. These metrics were computed for the single-frame, multi-frame, and denoised B-scans (all from the independent testing set; 1,552 B-scans of each type).

The SNR (expressed in dB) was a measure of signal strength relative to noise. It was defined as:

$$SNR = -10 * \log_{10} \left( \frac{\|f_o - \tilde{f}\|^2}{\|f_o\|^2} \right) \quad (1)$$

where  $f_o$  is the pixel-intensity values of the ‘clean’ (multi-frame) B-scan, and  $\tilde{f}$  is the pixel-intensity values of the B-scan (either the ‘noisy’ [single-frame] or the denoised B-scan) to be compared with  $f_o$ . A high SNR value indicates low noise in the given B-scan with respect to the ‘clean’ B-scan.

The CNR was a measure of contrast difference between different tissue layers. It was defined as:

$$CNR_i = \frac{|\mu_t - \mu_b|}{\sqrt{0.5(\sigma_t^2 + \sigma_b^2)}} \quad (2)$$

where  $\mu_t$  and  $\sigma_t^2$  denoted the mean and variance of pixel intensity for a chosen ROI within the tissue ‘ $i$ ’ in a given B-scan, while  $\mu_b$  and  $\sigma_b^2$  represented the same for the background ROI. The background ROI was chosen as a  $20 \times 384$  (in pixels) region at the top of the image (within the vitreous). A high CNR value suggested enhanced visibility of the given tissue.

The CNR was computed for the following tissues: (1) RNFL; (2) ganglion cell layer + inner plexiform layer (GCL + IPL); (3) all other retinal layers; (4) retinal pigment epithelium (RPE); (5) peripapillary choroid; (6) peripapillary sclera; and (7) lamina cribrosa (LC). Note that the CNR was computed only in the visible portions of the peripapillary sclera and LC. For each tissue, the CNR was computed as the mean of twenty-five ROIs ( $8 \times 8$  pixels each) in a given B-scan. All the ROIs were manually chosen in each tissue by an expert observer (G.S.) using a custom MATLAB (R2015a, MathWorks Inc., Natick, MA) graphical user interface.

The structural similarity index measure (SSIM)<sup>106</sup> was computed to assess the changes in tissue structures (i.e., edges) between the single-frame/denoised B-scans and the corresponding multi-frame B-scans (ground-truth). The SSIM was defined between  $-1$  and  $+1$ , where  $-1$  represented ‘no similarity’, and  $+1$  ‘perfect similarity’. It was defined as:

$$SSIM(x, y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \quad (3)$$

where  $x$  and  $y$  represented the denoised and multi-frame B-scan respectively;  $\mu_x$ ,  $\sigma_x^2$  denoted the mean intensity and standard deviation of the chosen ROI in B-scan  $x$ , while  $\mu_y$ ,  $\sigma_y^2$  represented the same for B-scan  $y$ ;  $\sigma_{xy}$  represented the cross-covariance of the ROIs in B-scans  $x$  and  $y$ .  $C_1$  and  $C_2$  (constants to stabilize the division) were chosen as 6.50 and 58.52, as recommended in a previous study<sup>106</sup>.

The MSSIM was computed as the mean of SSIM from ROIs ( $8 \times 8$  pixels each) across a B-scan (stride = 1; scanned horizontally). It was defined as:

$$MSSIM(X, Y) = \frac{1}{M} \sum_{k=1}^M SSIM(x_k, y_k) \quad (4)$$

Note that the SNR, and MSSIM were computed for an entire B-scan, as opposed to the CNR that was computed for individual tissues.

**Denosing performance: effect of data augmentation.** In an attempt to understand the significance of data augmentation, the entire process of training and testing was performed on two datasets: (1) baseline dataset (without data augmentation; 2,328 pairs of ‘clean’ and corresponding ‘noisy’ B-scans); and (2) data augmented dataset (23,280 pairs of ‘clean’ and corresponding ‘noisy’ B-scans).

**Denosing performance: clinical reliability.** The clinical reliability of the denoised B-scans was assessed by comparing the measurements of 3 clinically relevant ONH structural parameters obtained from the denoised and its corresponding multi-frame radial B-scans.

For each of the eight subjects in the testing set, we obtained 12 radial B-scans passing through the center of the Bruch’s membrane opening (BMO) from both the denoised volumes (single-frame volumes denoised using

the proposed network) and their corresponding multi-frame volumes. A total of 192 pairs (8 subjects; both eyes; 12 B-scans per volume) of radial B-scans (denoised and its corresponding multi-frame B-scans) were obtained.

In all the radial B-scans, the peripapillary retinal nerve fiber layer thickness (p-RNFLT), the peripapillary ganglion cell complex layer thickness (ganglion cell layer + inner plexiform layer; p-GCCT), and the peripapillary choroidal thickness (p-CT) were manually measured by an expert observer (S.K.D.) using Image<sup>107</sup>.

In each radial B-scan, the BMO points were defined as the extreme-tips of the RPE. The BMO reference line was obtained by joining the BMO points.

The p-RNFLT was computed as the distance between the inner limiting membrane and the posterior RNFL boundary taken at 1.7 mm on either side from the center of the BMO reference line.

The p-GCCT was computed as the distance between the posterior RNFL boundary and the posterior inner plexiform layer boundary taken at 1.7 mm on either side from the center of the BMO reference line.

The p-CT was computed as the distance between the posterior RPE boundary and the choroid-sclera interface taken at 1.7 mm on either side from the center of the BMO reference line.

All the three structural parameters were calculated as the average of the measurements taken on either side of the BMO.

For each structural parameter, paired t-test were used to assess the difference (means) in measurements when obtained from the denoised and their corresponding multi-frame radial B-scans.

Finally, we also computed the percentage error (mean) in measurements (between denoised and multi-frame radial B-scans) for each parameter.

### Data Availability

The dataset used for the training, validation, and testing of the proposed deep learning network was obtained from the Singapore National Eye Center and transferred to the National University of Singapore in a de-identified format. The study was approved by the SingHealth Centralized Institutional Review. Due to regulations, SingHealth does not authorize the dataset to be shared publicly.

### Code Availability

The authors wish to license the code to another party and thus are unable to release the entire codebase publicly at this stage. However, all the experiments, network descriptions, and data augmentations are described in sufficient details to enable independent replication with non-proprietary libraries.

### References

- Adhi, M. & Duker, J. S. Optical coherence tomography – current and future applications. *Current opinion in ophthalmology* **24**, 213–221, <https://doi.org/10.1097/ICU.0b013e32835f8bf8> (2013).
- Bussel, I. I., Wollstein, G. & Schuman, J. S. OCT for glaucoma diagnosis, screening and detection of glaucoma progression. *British Journal of Ophthalmology* **98**, ii15 (2014).
- Maldonado, R. S., Mettu, P., El-Dairi, M. & Bhatti, M. T. The application of optical coherence tomography in neurologic diseases. *Neurology: Clinical Practice* **5**, 460–469, <https://doi.org/10.1212/CPJ.0000000000000187> (2015).
- van Velthoven, M. E., Faber, D. J., Verbraak, F. D., van Leeuwen, T. G. & de Smet, M. D. Recent developments in optical coherence tomography for imaging the retina. *Progress in retinal and eye research* **26**, 57–77, <https://doi.org/10.1016/j.preteyeres.2006.10.002> (2007).
- Yongzhao, D., Gangjun, L., Feng, G. & Zhongping, C. Speckle reduction in optical coherence tomography images based on wave atoms. *Journal of Biomedical Optics* **19**, <https://doi.org/10.1117/1.JBO.19.5.056009> (May 2014).
- Bashkansky, M. & Reintjes, J. Statistics and reduction of speckle in optical coherence tomography. *Optics letters* **25**, 545–547, <https://doi.org/10.1364/OL.25.000545> (2000).
- Schmitt, J. M., Xiang, S. H. & Kin Man, Y. Speckle in optical coherence tomography. *Journal of Biomedical Optics* **4** (January 1999).
- Baghaie, A., Yu, Z. & D'Souza, R. M. State-of-the-art in retinal optical coherence tomography image analysis. *Quantitative Imaging in Medicine and Surgery* **5**, 603–617, <https://doi.org/10.3978/j.issn.2223-4292.2015.07.02> (2015).
- Szkulmowski, M. *et al.* Efficient reduction of speckle noise in Optical Coherence Tomography. *Optics express* **20**, 1337–1359, <https://doi.org/10.1364/oe.20.001337> (2012).
- Esmaili, M., Dehnavi, A. M., Rabbani, H. & Hajizadeh, F. Speckle Noise Reduction in Optical Coherence Tomography Using Two-dimensional Curvelet-based Dictionary Learning. *Journal of Medical Signals and Sensors* **7**, 86–91 (2017).
- Jian, Z. *et al.* Speckle attenuation in optical coherence tomography by curvelet shrinkage. *Optics letters* **34**, 1516–1518, <https://doi.org/10.1364/OL.34.001516> (2009).
- Hardin, J. S., Wong, G. T. & Seth, C. N., Chao, D., and Vizzeri, G. Factors Affecting Cirrus-HD OCT Optic Disc Scan Quality: A Review with Case Examples. *Journal of Ophthalmology* **2015** (2015).
- Mansouri, K., Medeiros, F. A., Tatham, A. J., Marchase, N. & Weinreb, R. N. Evaluation of Retinal and Choroidal Thickness by Swept-Source Optical Coherence Tomography: Repeatability and Assessment of Artifacts. *American journal of ophthalmology* **157**, 1022–1032, <https://doi.org/10.1016/j.ajo.2014.02.008> (2014).
- Asrani, S., Essaid, L., Alder, B. D. & Santiago-Turla, C. Artifacts in spectral-domain optical coherence tomography measurements in glaucoma. *JAMA ophthalmology* **132**, 396–402, <https://doi.org/10.1001/jamaophthalmol.2013.7974> (2014).
- Liu, Y. *et al.* Patient characteristics associated with artifacts in Spectralis optical coherence tomography imaging of the retinal nerve fiber layer in glaucoma. *American journal of ophthalmology* **159**, 565–576.e562, <https://doi.org/10.1016/j.ajo.2014.12.006> (2015).
- Kim, K. E., Jeoung, J. W., Park, K. H., Kim, D. M. & Kim, S. H. Diagnostic classification of macular ganglion cell and retinal nerve fiber layer analysis: differentiation of false-positives from glaucoma. *Ophthalmology* **122**, 502–510, <https://doi.org/10.1016/j.ophtha.2014.09.031> (2015).
- Balasubramanian, M., Bowd, C., Vizzeri, G., Weinreb, R. N. & Zangwill, L. M. Effect of image quality on tissue thickness measurements obtained with spectral-domain optical coherence tomography. *Optics express* **17**, 4019–4036 (2009).
- Mansberger, S. L., Menda, S. A., Fortune, B. A., Gardiner, S. K. & Demirel, S. Automated Segmentation Errors When Using Optical Coherence Tomography to Measure Retinal Nerve Fiber Layer Thickness in Glaucoma. *American journal of ophthalmology* **174**, 1–8, <https://doi.org/10.1016/j.ajo.2016.10.020> (2017).
- Iftimia, N., Bouma, B. E. & Tearney, G. J. Speckle reduction in optical coherence tomography by “path length encoded” angular compounding. *J Biomed Opt* **8**, 260–263, <https://doi.org/10.1117/1.1559060> (2003).
- Desjardins, A. E. *et al.* Angle-resolved optical coherence tomography with sequential angular selectivity for speckle reduction. *Optics express* **15**, 6200–6209 (2007).

21. Bajraszewski, T. *et al.* Improved spectral optical coherence tomography using optical frequency comb. *Optics express* **16**, 4163–4176 (2008).
22. Kennedy, B. F., Hillman, T. R., Curatolo, A. & Sampson, D. D. Speckle reduction in optical coherence tomography by strain compounding. *Optics letters* **35**, 2445–2447, <https://doi.org/10.1364/ol.35.002445> (2010).
23. Klein, T., Andre, R., Wieser, W., Pfeiffer, T. & Huber, R. Joint aperture detection for speckle reduction and increased collection efficiency in ophthalmic MHz OCT. *Biomed Opt Express* **4**, 619–634, <https://doi.org/10.1364/boe.4.000619> (2013).
24. Pircher, M., Gotzinger, E., Leitgeb, R., Fercher, A. F. & Hitzinger, C. K. Speckle reduction in optical coherence tomography by frequency compounding. *J Biomed Opt* **8**, 565–569, <https://doi.org/10.1117/1.1578087> (2003).
25. Schmitt, J. M. Array detection for speckle reduction in optical coherence microscopy. *Physics in medicine and biology* **42**, 1427–1439 (1997).
26. Schmitt, J. M. Restoration of Optical Coherence Images of Living Tissue Using the CLEAN Algorithm. *J Biomed Opt* **3**, 66–75, <https://doi.org/10.1117/1.429863> (1998).
27. Schmitt, J. M., Xiang, S. H. & Yung, K. M. Speckle in optical coherence tomography. *J Biomed Opt* **4**, 95–105, <https://doi.org/10.1117/1.429925> (1999).
28. Behar, V., Adam, D. & Friedman, Z. A new method of spatial compounding imaging. *Ultrasonics* **41**, 377–384 (2003).
29. Wu, W., Tan, O., Pappuru, R. R., Duan, H. & Huang, D. Assessment of frame-averaging algorithms in OCT image analysis. *Ophthalmic surgery, lasers & imaging retina* **44**, 168–175, <https://doi.org/10.3928/23258160-20130313-09> (2013).
30. Chen, C.-L. *et al.* Virtual Averaging Making Nonframe-Averaged Optical Coherence Tomography Images Comparable to Frame-Averaged Images. *Translational Vision Science & Technology* **5**, 1, <https://doi.org/10.1167/tvst.5.1.1> (2016).
31. Chitchian, S., Mayer, M. A., Boretzky, A. R., van Kuijk, F. J. & Motamedi, M. Retinal optical coherence tomography image enhancement via shrinkage denoising using double-density dual-tree complex wavelet transform. *Journal of Biomedical Optics* **17**, 116009, <https://doi.org/10.1117/1.JBO.17.11.116009> (2012).
32. R. Daniel Ferguson, D. X. H., Lelia Adelina Paunescu, Siobahn Beaton, Joel S. Schuman Tracking Optical Coherence Tomography *Optics letters* **29** (2004).
33. Ozcan, A., Bilenca, A., Desjardins, A. E., Bouma, B. E. & Tearney, G. J. Speckle reduction in optical coherence tomography images using digital filtering. *Journal of the Optical Society of America. A, Optics, image science, and vision* **24**, 1901–1910 (2007).
34. Bernardes, R. *et al.* Improved adaptive complex diffusion despeckling filter. *Optics express* **18**, 24048–24059, <https://doi.org/10.1364/oe.18.024048> (2010).
35. Wong, A., Mishra, A., Bizheva, K. & Clausi, D. A. General Bayesian estimation for speckle noise reduction in optical coherence tomography retinal imagery. *Optics express* **18**, 8338–8352, <https://doi.org/10.1364/oe.18.008338> (2010).
36. Bian, L., Suo, J., Chen, F. & Dai, Q. Multiframe denoising of high-speed optical coherence tomography data using interframe and intraframe priors. *Journal of Biomedical Optics* **20**, 1–11, 11 (2015).
37. Grzywacz, N. M. *et al.* Statistics of optical coherence tomography data from human retina. *IEEE transactions on medical imaging* **29**, 1224–1237, <https://doi.org/10.1109/tmi.2009.2038375> (2010).
38. Hossein Rabbani, M. S. & Abramoff, M. D. Optical Coherence Tomography Noise Reduction Using Anisotropic Local Bivariate Gaussian Mixture Prior in 3D Complex Wavelet Domain. *International Journal of Biomedical Imaging* **2013** (2013).
39. Li, M., Idoughi, R., Choudhury, B. & Heidrich, W. Statistical model for OCT image denoising. *Biomedical Optics Express* **8**, 3903–3917, <https://doi.org/10.1364/BOE.8.003903> (2017).
40. Mayer, M. A. *et al.* Wavelet denoising of multiframe optical coherence tomography data. *Biomedical Optics Express* **3**, 572–589, <https://doi.org/10.1364/BOE.3.000572> (2012).
41. Agravat, R. R. & Raval, M. S. In *Soft Computing Based Medical Image Analysis* (eds Nilanjan Dey, Amira S. Ashour, Fuqian Shi, & Valentina E. Balas) 183–201 (Academic Press, 2018).
42. Akkus, Z., Galimzianova, A., Hoogi, A., Rubin, D. L. & Erickson, B. J. Deep Learning for Brain MRI Segmentation: State of the Art and Future Directions. *Journal of Digital Imaging* **30**, 449–459, <https://doi.org/10.1007/s10278-017-9983-4> (2017).
43. Cui, Z., Yang, J. & Qiao, Y. In 2016 35th Chinese Control Conference (CCC). 7026–7031.
44. Liu, J. *et al.* Applications of deep learning to MRI images: A survey. *Big Data Mining and Analytics* **1**, 1–18, <https://doi.org/10.26599/BDMA.2018.9020001> (2018).
45. Mohsen, H., El-Dahshan, E.-S. A., El-Horbaty, E.-S. M. & Salem, A.-B. M. Classification using deep learning neural networks for brain tumors. *Future Computing and Informatics Journal* **3**, 68–71, <https://doi.org/10.1016/j.fcij.2017.12.001> (2018).
46. Viktor Wegmayr, S. A. & Buhmann, J. Classification of brain MRI with big data and deep 3D convolutional neural networks. *Proceedings Volume 10575, Medical Imaging 2018: Computer-Aided Diagnosis; 105751S* (2018).
47. Bertrand, H., Perrot, M., Ardon, R. & Bloch, I. Classification of MRI data using Deep Learning and Gaussian Process-based Model Selection. Preprint at, <https://arxiv.org/abs/1701.04355> (2017).
48. Benou, A., Veksler, R., Friedman, A. & Riklin Raviv, T. Ensemble of expert deep neural networks for spatio-temporal denoising of contrast-enhanced MRI sequences. *Medical image analysis* **42**, 145–159, <https://doi.org/10.1016/j.media.2017.07.006> (2017).
49. Jiang, D. *et al.* Denoising of 3D magnetic resonance images with multi-channel residual learning of convolutional neural network. Preprint at, <https://arxiv.org/abs/1712.08726> (2017).
50. Gondara, L. in 2016 IEEE 16th International Conference on Data Mining Workshops (ICDMW). 241–246 (2016).
51. Sui, X. *et al.* Choroid segmentation from Optical Coherence Tomography with graph-edge weights learned from deep convolutional neural networks. *Neurocomputing* **237**, 332–341, <https://doi.org/10.1016/j.neucom.2017.01.023> (2017).
52. Al-Bander, B., Williams, B. M., Al-Tae, M. A., Al-Nuaimy, W. & Zheng, Y. In 2017 10th International Conference on Developments in eSystems Engineering (DeSE). 182–187 (2017).
53. Zhang, Q., Cui, Z., Niu, X., Geng, S. & Qiao, Y. In *Neural Information Processing*. (eds Derong Liu *et al.*) 364–372 (Springer International Publishing).
54. Venhuizen, F. G. *et al.* Robust total retina thickness segmentation in optical coherence tomography images using convolutional neural networks. *Biomedical Optics Express* **8**, 3292–3316, <https://doi.org/10.1364/BOE.8.003292> (2017).
55. Fang, L. *et al.* Automatic segmentation of nine retinal layer boundaries in OCT images of non-exudative AMD patients using deep learning and graph search. *Biomedical Optics Express* **8**, 2732–2744, <https://doi.org/10.1364/BOE.8.002732> (2017).
56. Lu, D. *et al.* Deep-learning based multiclass retinal fluid segmentation and detection in optical coherence tomography images using a fully convolutional neural network. *Medical image analysis* **54**, 100–110, <https://doi.org/10.1016/j.media.2019.02.011> (2019).
57. Roy, A. G. *et al.* ReLayNet: retinal layer and fluid segmentation of macular optical coherence tomography using fully convolutional networks. *Biomedical Optics Express* **8**, 3627–3642, <https://doi.org/10.1364/BOE.8.003627> (2017).
58. Devalla, S. K. *et al.* DRUNET: a dilated-residual U-Net deep learning network to segment optic nerve head tissues in optical coherence tomography images. *Biomedical Optics Express* **9**, 3244–3265, <https://doi.org/10.1364/BOE.9.003244> (2018).
59. Devalla, S. K. *et al.* A Deep Learning Approach to Digitally Stain Optical Coherence Tomography Images of the Optic Nerve Head. *Investigative ophthalmology & visual science* **59**, 63–74, <https://doi.org/10.1167/iovs.17-22617> (2018).
60. Awais, M., Müller, H., Tang, T. B. & Meriaudeau, F. In 2017 IEEE International Conference on Signal and Image Processing Applications (ICSIPA). 489–492 (2017).
61. Prahs, P. *et al.* OCT-based deep learning algorithm for the evaluation of treatment indication with anti-vascular endothelial growth factor medications. *Graefes Archive for Clinical and Experimental Ophthalmology* **256**, 91–98, <https://doi.org/10.1007/s00417-017-3839-y> (2018).

62. Lee, C. S., Baughman, D. M. & Lee, A. Y. Deep Learning is Effective for Classifying Normal versus Age-Related Macular Degeneration OCT Images. *Ophthalmology Retina* **1**, 322–327, <https://doi.org/10.1016/j.oret.2016.12.009> (2017).
63. Liu, Y., Cheng, M. M., Hu, X., Wang, K. & Bai, X. In 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 5872–5881.
64. Al-Mujaini, A., Wali, U. K. & Azeem, S. Optical Coherence Tomography: Clinical Applications in Medical Practice. *Oman Medical Journal* **28**, 86–91, <https://doi.org/10.5001/omj.2013.24> (2013).
65. Bowd, C., Weinreb, R. N., Williams, J. M. & Zangwill, L. M. The retinal nerve fiber layer thickness in ocular hypertensive, normal, and glaucomatous eyes with optical coherence tomography. *Archives of ophthalmology (Chicago, Ill.: 1960)* **118**, 22–26 (2000).
66. McLellan, G. J. & Rasmussen, C. A. Optical Coherence Tomography for the Evaluation of Retinal and Optic Nerve Morphology in Animal Subjects: Practical Considerations. *Veterinary ophthalmology* **15**, 13–28, <https://doi.org/10.1111/j.1463-5224.2012.01045.x> (2012).
67. Miki, A. *et al.* Rates of retinal nerve fiber layer thinning in glaucoma suspect eyes. *Ophthalmology* **121**, 1350–1358, <https://doi.org/10.1016/j.ophtha.2014.01.017> (2014).
68. Ojima, T. *et al.* Measurement of retinal nerve fiber layer thickness and macular volume for glaucoma detection using optical coherence tomography. *Japanese journal of ophthalmology* **51**, 197–203, <https://doi.org/10.1007/s10384-006-0433-y> (2007).
69. Downs, J. C. *et al.* Posterior scleral thickness in perfusion-fixed normal and early-glaucoma monkey eyes. *Investigative ophthalmology & visual science* **42**, 3202–3208 (2001).
70. Lee, K. M. *et al.* Anterior lamina cribrosa insertion in primary open-angle glaucoma patients and healthy subjects. *PLoS One* **9**, e114935, <https://doi.org/10.1371/journal.pone.0114935> (2014).
71. Park, S. C. *et al.* Lamina cribrosa depth in different stages of glaucoma. *Investigative ophthalmology & visual science* **56**, 2059–2064, <https://doi.org/10.1167/iovs.14-15540> (2015).
72. Quigley, H. A. & Addicks, E. M. Regional differences in the structure of the lamina cribrosa and their relation to glaucomatous optic nerve damage. *Archives of ophthalmology (Chicago, Ill.: 1960)* **99**, 137–143 (1981).
73. Quigley, H. A., Addicks, E. M., Green, W. R. & Maumenee, A. E. Optic nerve damage in human glaucoma. II. The site of injury and susceptibility to damage. *Archives of ophthalmology (Chicago, Ill.: 1960)* **99**, 635–649 (1981).
74. Regatieri, C. V., Branchini, L. & Duker, J. S. The Role of Spectral-Domain OCT in the Diagnosis and Management of Neovascular Age-Related Macular Degeneration. *Ophthalmic surgery, lasers & imaging: the official journal of the International Society for Imaging in the Eye* **42**, S56–S66, <https://doi.org/10.3928/15428877-20110627-05> (2011).
75. Health Quality, O. Optical Coherence Tomography for Age-Related Macular Degeneration and Diabetic Macular Edema: An Evidence-Based Analysis. *Ontario Health Technology Assessment Series* **9**, 1–22 (2009).
76. Srinivasan, V. J. *et al.* High-Definition and 3-dimensional Imaging of Macular Pathologies with High-speed Ultrahigh-Resolution Optical Coherence Tomography. *Ophthalmology* **113**, 2054.e2051–2054.2014, <https://doi.org/10.1016/j.ophtha.2006.05.046> (2006).
77. Alshareef, R. A. *et al.* Prevalence and Distribution of Segmentation Errors in Macular Ganglion Cell Analysis of Healthy Eyes Using Cirrus HD-OCT. *PLOS ONE* **11**, e0155319, <https://doi.org/10.1371/journal.pone.0155319> (2016).
78. Stankiewicz, A. *et al.* Denoising methods for improving automatic segmentation in OCT images of human eye. **65**, 71, <https://doi.org/10.1515/bpasts-2017-0009> (2017).
79. Scoles, D. *et al.* Assessing Photoreceptor Structure Associated with Ellipsoid Zone Disruptions Visualized with Optical Coherence Tomography. *Retina (Philadelphia, Pa.)* **36**, 91–103, <https://doi.org/10.1097/IAE.0000000000000618> (2016).
80. Kern, T. S. & Berkowitz, B. A. Photoreceptors in diabetic retinopathy. *Journal of Diabetes Investigation* **6**, 371–380, <https://doi.org/10.1111/jdi.12312> (2015).
81. Baba, T. *et al.* Correlation of visual recovery and presence of photoreceptor inner/outer segment junction in optical coherence images after successful macular hole repair. *Retina* **28**, 453–458, <https://doi.org/10.1097/IAE.0b013e3181571398> (2008).
82. Hayashi, H. *et al.* Association between Foveal Photoreceptor Integrity and Visual Outcome in Neovascular Age-related Macular Degeneration. *American journal of ophthalmology* **148**, 83–89.e81, <https://doi.org/10.1016/j.ajo.2009.01.017> (2009).
83. Flatter, J. A. *et al.* Outer retinal structure after closed-globe blunt ocular trauma. *Retina* **34**, 2133–2146, <https://doi.org/10.1097/iae.0000000000000169> (2014).
84. Bambo, M. P. *et al.* Influence of cataract surgery on repeatability and measurements of spectral domain optical coherence tomography. *British Journal of Ophthalmology* **98**, 52 (2014).
85. Kok, P. H. B. *et al.* The relationship between the optical density of cataract and its influence on retinal nerve fibre layer thickness measured with spectral domain optical coherence tomography. *Acta Ophthalmologica* **91**, 418–424, <https://doi.org/10.1111/j.1755-3768.2012.02514.x> (2012).
86. Mwanza, J. C. *et al.* Effect of cataract and its removal on signal strength and peripapillary retinal nerve fiber layer optical coherence tomography measurements. *Journal of glaucoma* **20**, 37–43, <https://doi.org/10.1097/IJG.0b013e3181cb93b> (2011).
87. Savini, G., Zanini, M. & Barboni, P. Influence of pupil size and cataract on retinal nerve fiber layer thickness measurements by Stratus OCT. *Journal of glaucoma* **15**, 336–340, <https://doi.org/10.1097/01.ijg.0000212244.64584.c2> (2006).
88. Stein, D. M. *et al.* Effect of Corneal Drying on Optical Coherence Tomography. *Ophthalmology* **113**, 985–991, <https://doi.org/10.1016/j.ophtha.2006.02.018> (2006).
89. Nicola G. Ghazi, J. W. M The effect of lubricating eye drops on optical coherence tomography imaging of the retina. *Digital Journal of Ophthalmology* **15** (2009).
90. Smith, M., Frost, A., Graham, C. M. & Shaw, S. Effect of pupillary dilatation on glaucoma assessments using optical coherence tomography. *The British Journal of Ophthalmology* **91**, 1686–1690, <https://doi.org/10.1136/bjo.2006.113134> (2007).
91. Moisseiev, E. *et al.* Pupil dilation using drops vs gel: a comparative study. *Eye* **29**, 815, <https://doi.org/10.1038/eye.2015.47> (2015).
92. Ferguson, R. D., Hammer, D. X., Paunescu, L. A., Beaton, S. & Schuman, J. S. Tracking optical coherence tomography. *Optics letters* **29**, 2139–2141 (2004).
93. Shin, J. W. *et al.* The Effect of Optic Disc Center Displacement on Retinal Nerve Fiber Layer Measurement Determined by Spectral Domain Optical Coherence Tomography. *PLOS ONE* **11**, e0165538, <https://doi.org/10.1371/journal.pone.0165538> (2016).
94. Wong, I. Y., Koizumi, H. & Lai, W. W. Enhanced depth imaging optical coherence tomography. *Ophthalmic surgery, lasers & imaging: the official journal of the International Society for Imaging in the Eye* **42**(Suppl), S75–84, <https://doi.org/10.3928/15428877-20110627-07> (2011).
95. Hammer, D. *et al.* Advanced scanning methods with tracking optical coherence tomography. *Optics express* **13**, 7937–7947 (2005).
96. Wells, W. M. 3rd, Viola, P., Atsumi, H., Nakajima, S. & Kikinis, R. Multi-modal volume registration by maximization of mutual information. *Medical image analysis* **1**, 35–51 (1996).
97. O. Ronneberger, P. F. & Brox, T. U-Net: Convolutional Networks for Biomedical Image Segmentation. *Medical Image Computing and Computer-Assisted Intervention (MICCAI)* **9351**, 234–241 (2015).
98. He, K., Zhang, X., Ren, S. & Sun, J. In 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 770–778 (2016).
99. Yu, F. & Koltun, V. In 4th International Conference on Learning Representations, ICLR 2016 (Poster) (2016).
100. Clevert, D.-A., Unterthiner, T. & Hochreiter, S. In 4th International Conference on Learning Representations, ICLR 2016 (Poster) (Feb 2016).
101. Sergey Ioffe, C. S. In *Proceedings of the 32nd International Conference on Machine Learning - Volume 37* 448–456 (JMLR.org, Lille, France, 2015).

102. Kingma, D. P. & Ba, J. in 3rd International Conference for Learning Representations, ICLR 2015 (Poster) (2015).
103. Simard, P. Y. & John, D. S. C. Platt Best Practices for Convolutional Neural Networks Applied to Visual Document Analysis *Proceedings of the Seventh International Conference on Document Analysis and Recognition (ICDAR 2003)* (2003).
104. Wu, Z., Xu, G., Weinreb, R. N., Yu, M. & Leung, C. K. Optic Nerve Head Deformation in Glaucoma: A Prospective Analysis of Optic Nerve Head Surface and Lamina Cribrosa Surface Displacement. *Ophthalmology* **122**, 1317–1329, <https://doi.org/10.1016/j.ophtha.2015.02.035> (2015).
105. Girard, M. J., Strouthidis, N. G., Ethier, C. R. & Mari, J. M. Shadow removal and contrast enhancement in optical coherence tomography images of the human optic nerve head. *Investigative ophthalmology & visual science* **52**, 7738–7748, <https://doi.org/10.1167/iovs.10-6925> (2011).
106. Zhou, W., Bovik, A. C., Sheikh, H. R. & Simoncelli, E. P. Image quality assessment: from error visibility to structural similarity. *IEEE Transactions on Image Processing* **13**, 600–612, <https://doi.org/10.1109/TIP.2003.819861> (2004).
107. Rueden, C. T. S. *et al.* ImageJ2: ImageJ for the next generation of scientific image data. *BMC Bioinformatics* **18** (2017).

## Acknowledgements

Singapore Ministry of Education Academic Research Funds Tier 1 (R-155-000-168-112 [A.H.T.]; R-397-000-294-114 [M.J.A.G.]); National University of Singapore (NUS) Young Investigator Award Grant (NUSYIA\_FY16\_P16, R-155-000-180-133; [A.H.T.]); Singapore Ministry of Education Academic Research Funds Tier 2 (R-397-000-280-112, R-397-000-308-112 [M.J.A.G.]; MOE2016-T2-2-135 [A.H.T.]); National Medical Research Council (Grant NMRC/STAR/0023/2014 [T.A.]).

## Author Contributions

S.K.D. and G.S. conceived the study, designed the experiments, and wrote the paper; T.H.P. assisted the design of the algorithm; X.W. and T.A.T. assisted in imaging all the subjects, and edited the manuscript; S.P., T.A., L.S., edited the manuscript, A.H.T. and M.J.A.G., supervised the study and edited the manuscript.

## Additional Information

**Competing Interests:** Dr. Michaël J. A. Girard and Dr. Alexandre H. Thiéry are co-founders of Abyss Processing Pte Ltd.

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2019