

DeshadowGAN: A Deep Learning Approach to Remove Shadows from Optical Coherence Tomography Images

Haris Cheong¹, Sripad Krishna Devalla¹, Tan Hung Pham¹, Liang Zhang¹,
Tin Aung Tun², Xiaofei Wang³, Shamira Perera^{2,4}, Leopold Schmetterer^{2,4-7},
Tin Aung^{1,2}, Craig Boote^{1,8,9}, Alexandre Thiery⁵, and Michaël J. A. Girard^{1,2}

¹ Ophthalmic Engineering and Innovation Laboratory, Department of Biomedical Engineering, Faculty of Engineering, National University of Singapore, Singapore

² Singapore Eye Research Institute, Singapore National Eye Centre, Singapore

³ Beijing Advanced Innovation Center for Biomedical Engineering, School of Biological Science and Medical Engineering, Beihang University, Beijing, China

⁴ Ophthalmology Department, Duke-NUS Medical School, Singapore

⁵ Department of Statistics and Applied Probability, National University of Singapore, Singapore

⁶ Lee Kong Chian School of Medicine, Nanyang Technological University, Singapore

⁷ Department of Clinical Pharmacology, Medical University of Vienna, Austria

⁸ School of Optometry & Vision Sciences, Cardiff University, UK

⁹ Newcastle Research & Innovation Institute, Singapore

Correspondence: Michaël J. A. Girard, Ophthalmic Engineering and Innovation Laboratory, Department of Biomedical Engineering, Faculty of Engineering, National University of Singapore, 4 Engineering Dr 3, Block E4 #04-08, Singapore 117583. e-mail: mgirard@nus.edu.sg

Received: September 30, 2019

Accepted: December 13, 2019

Published: April 15, 2020

Keywords: glaucoma; generative adversarial network; deep learning; shadow removal

Citation: Cheong H, Devalla SK, Pham TH, Zhang L, Tun TA, Wang X, Perera S, Schmetterer L, Aung T, Boote C, Thiery A, Girard MJA. DeshadowGAN: a deep learning approach to remove shadows from optical coherence tomography images. *Trans Vis Sci Tech.* 2020;9(2):23, <https://doi.org/10.1167/tvst.9.2.23>

Purpose: To remove blood vessel shadows from optical coherence tomography (OCT) images of the optic nerve head (ONH).

Methods: Volume scans consisting of 97 horizontal B-scans were acquired through the center of the ONH using a commercial OCT device for both eyes of 13 subjects. A custom generative adversarial network (named DeshadowGAN) was designed and trained with 2328 B-scans in order to remove blood vessel shadows in unseen B-scans. Image quality was assessed qualitatively (for artifacts) and quantitatively using the intralayer contrast—a measure of shadow visibility ranging from 0 (shadow-free) to 1 (strong shadow). This was computed in the retinal nerve fiber layer (RNFL), the inner plexiform layer (IPL), the photoreceptor (PR) layer, and the retinal pigment epithelium (RPE) layer. The performance of DeshadowGAN was also compared with that of compensation, the standard for shadow removal.

Results: DeshadowGAN decreased the intralayer contrast in all tissue layers. On average, the intralayer contrast decreased by $33.7 \pm 6.81\%$, $28.8 \pm 10.4\%$, $35.9 \pm 13.0\%$, and $43.0 \pm 19.5\%$ for the RNFL, IPL, PR layer, and RPE layer, respectively, indicating successful shadow removal across all depths. Output images were also free from artifacts commonly observed with compensation.

Conclusions: DeshadowGAN significantly corrected blood vessel shadows in OCT images of the ONH. Our algorithm may be considered as a preprocessing step to improve the performance of a wide range of algorithms including those currently being used for OCT segmentation, denoising, and classification.

Translational Relevance: DeshadowGAN could be integrated to existing OCT devices to improve the diagnosis and prognosis of ocular pathologies.

Introduction

Glaucoma is the leading cause of irreversible blindness and occurs due to the death of retinal ganglion

cells within the optic nerve head (ONH).¹ In its most common form, there are no symptoms, making regular diagnostic tests crucial for early detection and treatment.² Recent research suggests that glaucoma eyes have a unique biomechanical and structural profile that

may allow us to differentiate them from non-glaucoma eyes.³

To better understand how glaucoma affects the structure and biomechanics of the eye, optical coherence tomography (OCT) has been shown to be a promising tool.⁴ It uses low-coherence light to capture micrometer resolution and three-dimensional images, allowing in vivo visualization of a patient's retinal layers⁵; however, because light does not travel through blood, information from locations beneath blood vessels is significantly decreased.⁶ This causes artifacts known as retinal shadows. These artifacts may introduce errors in retinal nerve fiber layer (RNFL) thickness measurements, which has clinical implications for the management of glaucoma where changes in RNFL thickness must be monitored accurately over time.⁷ These shadows may also occlude deep structures such as the lamina cribrosa (LC), the main site of axonal loss in glaucoma.⁸ Other studies have identified retinal shadows as being a challenge for their study of retinal layers. Mujat et al.⁹ found that retinal shadows generate “holes” within the posterior boundary of the RNFL, and they strived to probe behind these blood vessels by identifying the index of OCT A-scans so as to compensate for the shadows. Consequently, it may be crucial to develop algorithms to replenish the information lost within these shadows.

Other studies had since found compensatory methods to combat information loss within shadows. Fabritius et al.¹⁰ described a compensatory method to reduce the effects of vessel artifacts on interpretation of the retinal pigment epithelium (RPE) layer. Mari et al.¹¹ also improved the quality of OCT images through compensation methods by correcting the effects of light attenuation and by better estimating the optical properties (e.g., reflectivity) of the tissues. These predictions are, however, only estimations or are based on simple optical models that may result in secondary artifacts being produced such as inverted shadows.

Artificial intelligence techniques have been applied extensively to shadow removal algorithms for normal images with varying levels of success.^{12,13} In 2014, Goodfellow et al.¹⁴ introduced faux image generation using generative adversarial networks (GANs). This technique paved the way for GANs to be applied for other purposes, such as shadow removal,¹² shadow detection,^{15,16} and unwanted artifact removal.¹⁷ In this study, we aimed to test whether a custom GAN (DeshadowGAN) could automatically detect and remove shadows according to a predicted “shadow score” in order to improve the quality of OCT images of the ONH.

Materials and Methods

Patient Recruitment

A total of 13 healthy subjects (average age, 28 years) were recruited at the Singapore National Eye Centre. All subjects gave written informed consent. This study adhered to the tenets of the Declaration of Helsinki and was approved by the institutional review board of the hospital. The inclusion criteria for healthy subjects were an intraocular pressure (IOP) less than 21 mmHg and healthy optic nerves with a vertical cup-to-disc ratio ≤ 0.5 .

Optical Coherence Tomography Imaging

Recruited subjects were seated and imaged in dark room conditions by a single operator. A standard spectral domain OCT system (Spectralis; Heidelberg Engineering, Heidelberg, Germany) was used to image both eyes of each subject. We obtained 97 horizontal B-scans (32- μm distance between B-scans; 384 A-scans per B-scan) from a rectangular area $15^\circ \times 10^\circ$ centered on the ONH. We obtained 75 times signal-averaged images from multiframe volumes. In total, our training set consisted of 2328 multiframe baseline B-scans from 24 three-dimensional (3D) volumes. Our test set consisted of 291 multiframe baseline B-scans from three 3D volumes.

DeshadowGAN: Overall Description

Our algorithm was comprised of two networks competing with one another. The first network was referred to as the shadow detection network, and it predicted which pixels would be considered as shadowed pixels. The second network was referred to as the shadow removal network, and it aimed to remove shadows from OCT images such that the first network (shadow detection network) could no longer identify shadowed pixels. Briefly, we first trained the shadow detection network five times on baseline images with their corresponding manually segmented shadow mask as the ground truth. Binary segmentation masks (size 496×384) were manually created for all 2 328 B-scans using ImageJ software (National Institutes of Health, Bethesda, MD)¹⁸ by one observer (HC); shadowed pixels were labeled as 1 and shadow-free pixels as 0. Next, we trained the shadow removal network once by passing the baseline as input and using the predicted binary masks as part of the loss function. Finally, we trained the shadow detection network another five times with the output from the shadow removal network and another five times with the manually

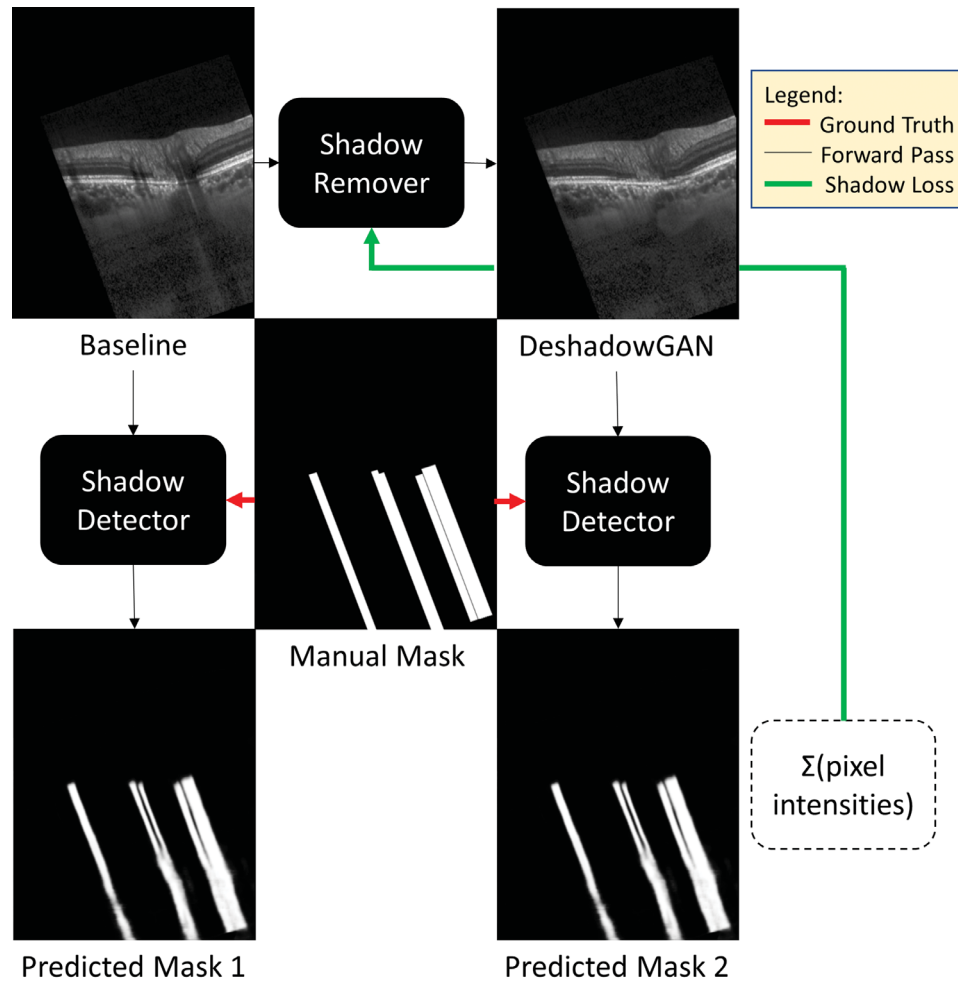


Figure 1. Overall algorithm training diagram.

segmented binary masks as ground truth (Fig. 1). More details about the two networks can be found below.

Shadow Detection Network

A neural network inspired by the U-Net architecture¹⁹ (Fig. 2) was trained with a simple binary cross entropy loss²⁰ using the handcrafted segmentation masks as ground truth. This network had a sigmoid layer as its final activation, making it a per-pixel binary classifier. It was then trained with original images concatenated with the output from a shadow removal network, using the manually segmented masks as ground truth.

The shadow detection network first performed two convolutions with kernel size 3 and stride 1, followed by a ReLU activation²¹ after each convolution. Then, images were downsampled using a 2×2 maxpool operation, halving the size of the height and width of the feature maps. This occurred four times, with the number of feature maps at each smaller size increasing from 1 to 64, 128, 256, and 512, respectively.

The shadow detection network was comprised of two towers. A downsampling tower at each stage sequentially halved the dimensions of the baseline image (size 512×512) via maxpooling to capture the contextual information (i.e., spatial arrangement of tissues), and an upsampling tower sequentially restored it back to its original resolution to capture the local information (i.e., tissue texture).²² A transposed convolution was performed four times in the upsampling tower for the predicted segmentation masks to be size 512×512 , before passing to a sigmoid activation function for compression of each pixel to a value between 0 and 1.

Shadow Removal Network

The shadow removal network was inspired from the deep video portraits reported by Kim et al.²³ A schematic of the architecture is shown in Figure 3. Baseline images were inputted into the network and passed through a downsampling segment and an upsampling segment (colored in yellow and blue,

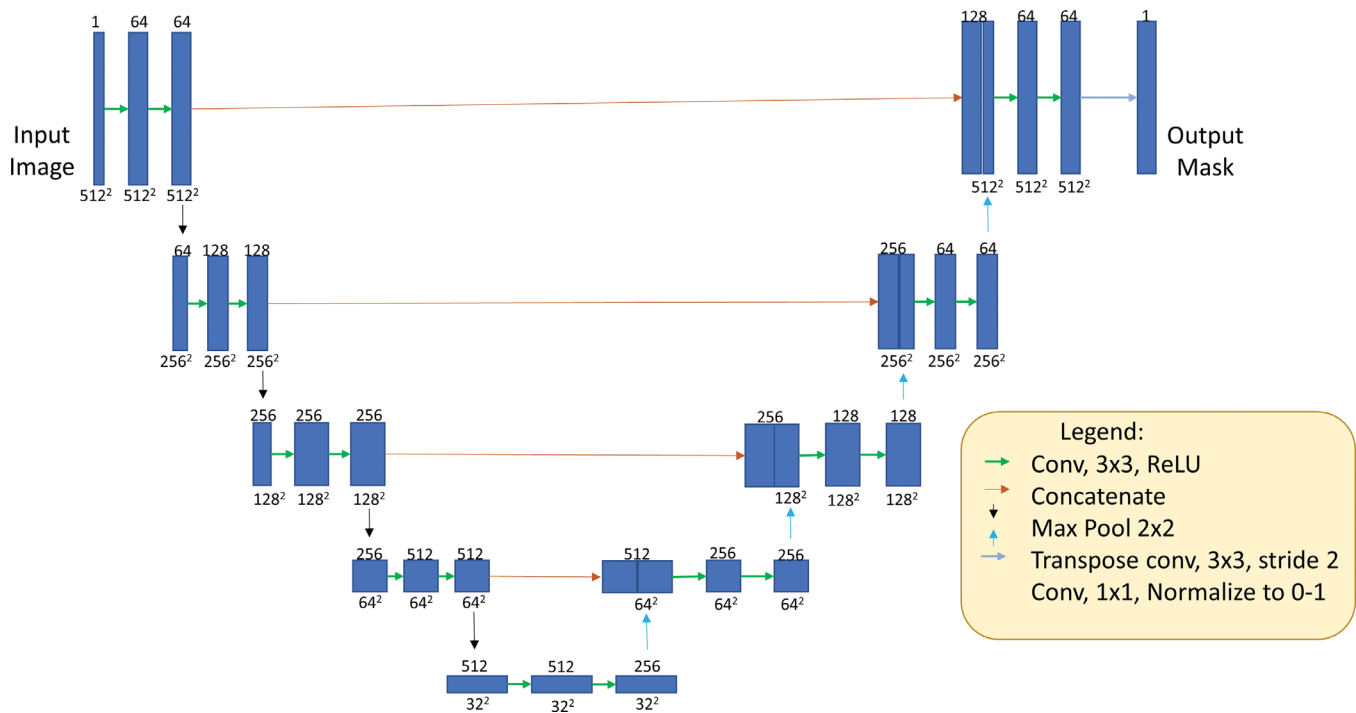


Figure 2. Shadow detection network architecture. Numbers on top of each rectangle represent the number of feature maps, and numbers below each rectangle represent the feature map size. The network consists of 13.4M parameters, occupying 648 MiB of RAM on a single Nvidia GTX 1080 Ti.

respectively) (Fig. 3b). The downsampling segment allowed the network to understand contextual information, and the upsampling segment increased the resolution of the output. Features from both segments were combined to produce a more precise output in the successive convolution layer.²⁴

The shadow removal network consisted of eight downsampling modules and eight upsampling modules (Fig. 3b). The first encoding layer and the last decoding layer did not employ batch normalization. We included a dropout of 0.5 only in the first three upsampling layers. The network had 55.7M parameters and occupied 820 MiB of RAM on a Nvidia GTX 1080 Ti (Santa Clara, CA). Each module in the downsampling module consisted of a convolution layer (stride 2, kernel size 4×4) followed by a batch normalization and a leaky ReLU activation function. Every downsampling module reduced the feature map size by half, enabling the network to derive contextual information and encode its input into an increasing number of filters. The maximum number of filters plateaued at 512 for four times and then moved on to the decoding segment.

Each decoding segment consisted of three other submodules: Up and two Refine submodules. The Up submodule consisted of a transpose convolution (stride 2, size 4×4) followed by a batch normalization and a ReLU activation function. Every Up submodule allowed the network to improve its decoding efficiency

from encoded information from the input. The Refine submodule consisted of a convolution (stride 1, size 3×3) followed by a batch normalization and a 0.5 dropout. We repeated this process until a 512×512 feature map was obtained and we reduced the number of feature maps produced in the last layer to one to mimic input images. Finally, we applied a pixel-wise sigmoid activation to compress all activations from the decoding segment to values between 0 and 1.

Image Augmentation

An image augmentation network was created using Pytorch²⁵ to perform on-demand image augmentation during training. Our data augmentation consisted of random transformations, including horizontal flipping, image rotations (angle between -40° and 40°), XY translations (-20% to 20% of the image size), image scaling (scaling factor between 0.8 and 1.2), and image shear (shear angle between -20° and 20°). All images were then resized to 512×512 pixels.

Weighted-Mixture Loss Function for the Shadow Removal Network

An adversarial shadow removal network was simultaneously trained using a custom loss function (to be minimized during training) that reduced the appearance of shadows in output images. This custom loss

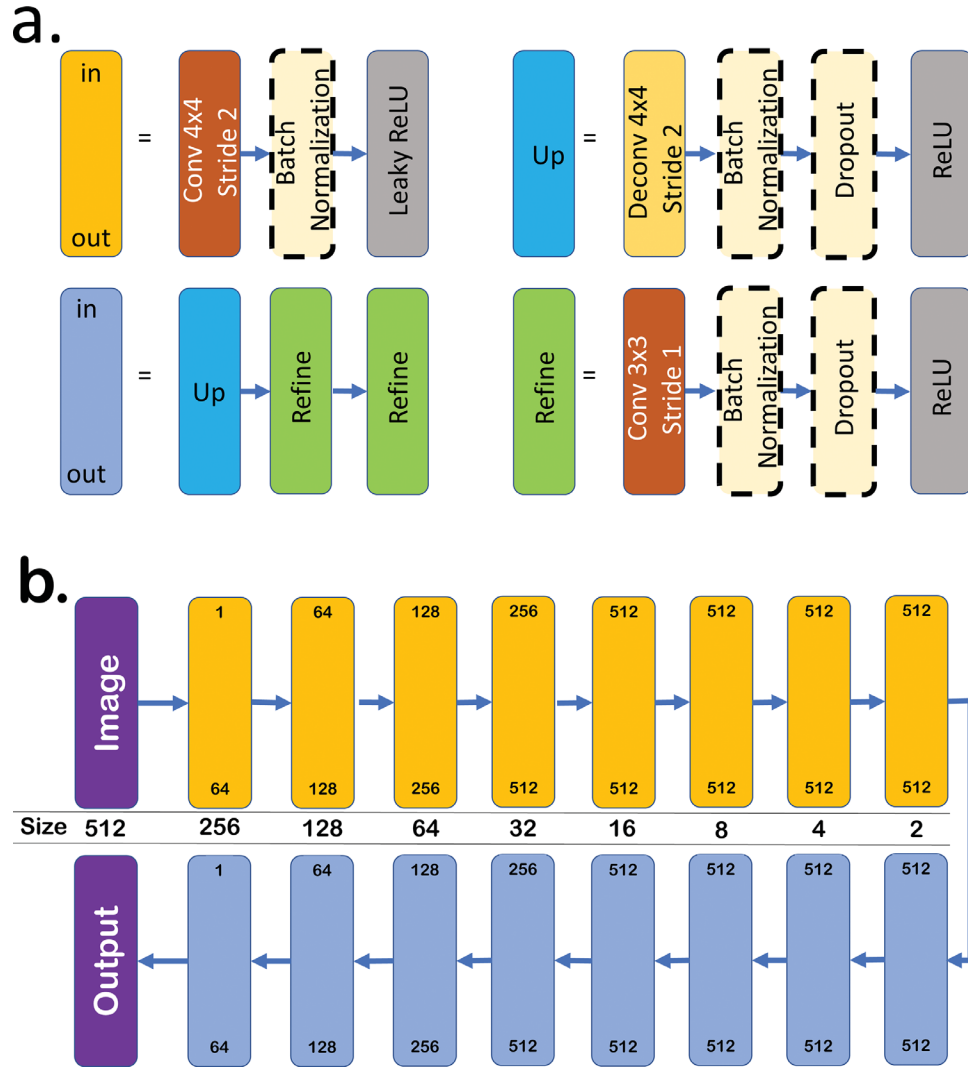


Figure 3. All arrows represent a forward pass of the output from one layer to the input of the next layer. Each box represents a module (a set of layers). The size of our input image is 512×512 . (a) Definitions of the layers in downsampling and upsampling modules within the shadow removal network. Dotted boundaries indicate that the module is present only within some layers. In and out values at the top and bottom of each rectangle represent the number of feature maps being input and output from that module, respectively. (b) The size row indicates the size of the output of each module (rectangles above and below it).

function was used to restore structural information under blood vessel shadows while maintaining structural information in all other areas. It consisted of a tuned weighted combination of four losses: content, style, total variation, and shadow losses, as briefly explained below.

Content Loss

We used the content loss to ensure that all non-shadowed regions of a given image remained the same after shadow correction. To do so, we compared high-level image feature representations (using a pretrained convolutional neural network known as ResNet-152) from a given baseline image, B , with those from

its deshadowed output, D . Note that the content loss function has been used in Style Transfer²⁶ and has been shown to maintain fine image details and edges.

To calculate the content loss, we first segmented all shadows from the baseline image with our shadow detection network. All shadows were then masked (pixel intensity values equal to zero) in order to generate two images: B_{masked} (baseline image with masked shadows) and D_{masked} (deshadowed image with masked shadows) as shown in Figure 4.

B_{masked} and D_{masked} were then passed to the ResNet-152 network,²⁷ itself trained using the ImageNet dataset. The content loss was then calculated for each

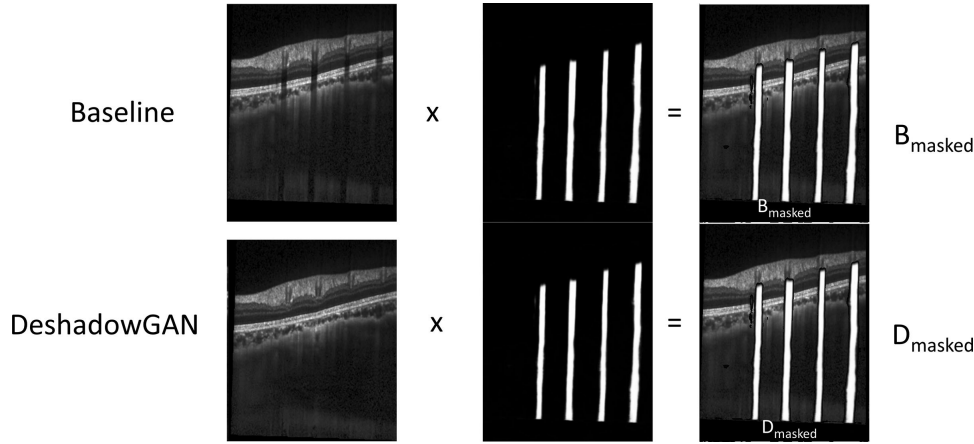


Figure 4. Masking of baseline and deshadowed images during content loss and style loss calculations. Predicted shadow mask for the baseline image is used to mask both the baseline and deshadowed image.

image pair as an Euclidean norm as follows:

$$\begin{aligned} \mathcal{L}_{content}(B_{masked}, D_{masked}) \\ = \sum_{i=9,33,141,150} \frac{1}{C_i H_i W_i} \|P_i(B_{masked}) - P_i(D_{masked})\|^2 \end{aligned} \quad (1)$$

where $P_i(x)$ is a feature map that contains the activations of the i th convolutional layer of the ResNet-152 network for a given image x ; C_i , H_i , and W_i represent the channel number, height, and width of the feature map $P_i(x)$, respectively. The convolutional layers mentioned in Equation 1 (i.e., 9, 33, 141, and 150) were selected because they are the final convolutional layers in the ResNet-152 network before downsampling.

Style Loss

On top of the content loss, we also used the style loss²⁸ to ensure that the image style (texture) remained the same in the non-shadowed regions after shadow correction. We computed the Gram matrix of an image to find a representation of its style. The style loss was then computed for each image pair (B_{masked} , D_{masked}) and was defined as the Euclidean norm between the Gram matrices of B_{masked} and D_{masked} :

$$\begin{aligned} \mathcal{L}_{style}(B_{masked}, D_{masked}) \\ = \sum_{i=9,33,141,150} \|G_i(B_{masked}) - G_i(D_{masked})\|^2 \end{aligned} \quad (2)$$

where G_i is a $C_i \times C_i$ matrix defined as

$$G_i(x) = P_i(x)_{C_i, W_i, H_i} \times P_i(x)_{H_i, W_i, C_i} \quad (3)$$

Total Variation Loss

We used the total variation loss to prevent checker-board artifacts from appearing in deshadowed images. It was defined as the sum of the differences between neighboring pixels in a given deshadowed image, D :

$$\mathcal{L}_{TV}(D) = \frac{1}{n} \sum_{i,j} |D_{i+1,j} - D_{i,j}| + |D_{i,j+1} - D_{i,j}| \quad (4)$$

where n is the total number of pixels in the deshadowed image, and i and j are the row and column numbers, respectively.

Shadow Loss

The shadow loss was defined to ensure that shadows were properly removed so that they become undetectable to the shadow detection network. When a given image D had been deshadowed, it was passed to the shadow detection network to produce a predicted shadow mask, M_D (with pixel intensities equal to 1). All pixel intensities in the shadow mask were summed, and this sum was defined as the shadow loss function.

Total Loss

The total loss for the shadow removal network was defined as

$$\begin{aligned} \mathcal{L}_{total} = w_1 * content\ loss + w_2 * style\ loss \\ + w_3 * shadow\ loss + w_4 * total\ variation\ loss \end{aligned} \quad (5)$$

where w_1 , w_2 , w_3 , and w_4 are weights given the following values: 100, 0.1, 100, and 1e-5, respectively. Note that all weights were tuned manually through an iterative approach. First, we found that with a value of $w_1 = 100$,

we generated images with no content loss (i.e., no structural changes), but not without the presence of checkerboard artifacts. These artifacts could be removed when choosing $w_2 = 0.1$ and $w_4 = 1e-5$.²⁹ Finally, we increased the value of w_3 until the shadow loss became the largest component in the total loss function (so that the focus remained on removing shadows) and until shadow removal was deemed qualitatively acceptable for smaller width shadows. This was optimum when $w_3 = 100$.

Training Parameters

All training and testing were performed on a Nvidia GTX 1080 Ti with CUDA 10.1 and cuDNN v7.6.0 acceleration. Using these hardware specifications, each image took an average of 10.3 ms to be deshadowed. The total training time was 7 days using the Adam optimizer³⁰ at a learning rate of 1×10^{-5} and a batch size of 2. A learning rate decay was implemented to halve learning rates every 10 epochs. We stopped the training when no observable improvements in output images could be observed.

Shadow Removal Metrics

Intralayer Contrast

We used the intralayer contrast to assess the performance of our algorithm in removing shadows. The intralayer contrast was defined as

$$\text{Intralayer Contrast} = \left| \frac{I_1 - I_2}{I_1 + I_2} \right| \quad (6)$$

where I_1 is the mean pixel intensity from five manually selected regions of interest (size 5×5 pixels) that are shadow free in a given retinal layer, and I_2 is that from five neighboring shadowed regions of the same tissue layer. The intralayer contrast varied between 0 and 1, where values close to 0 indicate the absence of blood vessel shadows and values close to 1 indicate a strong presence of blood vessel shadows.

We computed the intralayer contrast for multiple tissue layers of the ONH region—namely, the RNFL, photoreceptor (PR) layer, inner plexiform layer (IPL), and retinal pigment epithelium (RPE) layer—before and after application of our deshadowing algorithm. The intralayer contrast was computed on an independent test set consisting of 291 images. Results were reported in the form of mean \pm standard deviation (SD).

Comparison with Adaptive Compensation

To evaluate the effectiveness of our deshadowing algorithm, we compared images deshadowed using DeshadowGAN with images enhanced with adaptive compensation,^{6,31} the gold-standard for correcting OCT shadows. For adaptive compensation, we used contrast, decompression, compression, and threshold exponents of 1, 4, 4, and 6, respectively. Intralayer contrasts were also computed for all compensated images (same regions as those used for the baseline images).

Validation Using a Test Scenario with Known Ground Truth

We investigated whether our deshadowing algorithm was capable of restoring information below blood vessel shadows without introducing unwanted artifacts. To do so, we required ground truth images without blood vessel shadows, but such images are not easy to obtain in vivo. An alternative is to add artificial shadows to a given baseline image and assess whether our algorithm can remove them without introducing artifacts. Accordingly, we created exponential decay maps on images to simulate the effect of light attenuation and trained DeshadowGAN with such images. A shadow can be simply simulated as

$$\text{ShadowPixel}_{ij} = \text{BaselinePixel}_{ij} \times e^{-\alpha i} \quad (7)$$

where i is the row number and α indicates the rate of decay. We used the same training and testing image sets, except that two artificial shadows (random width between 1 and 100 pixels; random α between 100 and 300) were randomly added to each baseline image. DeshadowGAN was retrained with the exact same aforementioned procedure, including the manual segmentation of the artificial shadows. Note also that, during training, the DeshadowGAN algorithm did not have access to the ground truth baseline images without shadows. After deshadowing, the presence of artifacts was assessed qualitatively.

Quantification of Test Outcome Using the Peak Signal-to-Noise Ratio

To further quantify the performance of DeshadowGAN, we compared the peak signal-to-noise ratio (PSNR) values when artificial shadows were (1) more pronounced (i.e., higher exponential decay of the OCT signal), and (2) wider. To investigate the effects of shadow width on the PSNR, the exponential decay was fixed at 0.005. Artificial shadows were created with widths of 240, 600, 960, and 1440 μm in four

separate experiments on a test set composed of 291 images. These width values represent those of true retinal shadows. The PSNR was then calculated on deshadowed images as

$$PSNR = 20 \times \log_{10} MAX_I - 10 \times \log_{10} MSE \quad (8)$$

where MSE refers to the mean squared error between D (deshadowed image) and T (ground truth image without artificial shadows), given by

$$MSE(D, T) = \sum ||D - T||^2 \quad (9)$$

To investigate the effects of exponential decay on the PSNR, we created artificial shadows with varying exponential decay (i.e., 0.00333, 0.00400, 0.00500, and 0.00667) on the same test set composed of 291 images. Here, a higher exponential decay indicates a stronger shadow. These exponential decay values were derived from true shadows present in ONH images. We also used a distribution of shadow widths (69.0 ± 25.0 pixels). When the shadows had been created then corrected with DeshadowGAN, the PSNR was calculated on deshadowed images.

Results

DeshadowGAN Decreased the Intralayer Contrast

After application of our algorithm, blood vessel shadows from unseen images were successfully identified and corrected from each retinal layer as observed quantitatively and qualitatively. On average, we observed improvements in intralayer contrast of $33.7 \pm 6.81\%$, $28.8 \pm 10.4\%$, $35.9 \pm 13.0\%$, and $43.0 \pm 19.5\%$ for the RNFL, IPL, PR layer, and RPE layer respectively. This can be qualitatively observed in a B-scan of the peripapillary tissues shown in Figure 5.

Comparison with Adaptive Compensation

DeshadowGAN was able to correct shadows without affecting the contrast of anterior layers, without adding noise and without creating artifacts (Fig. 6). In addition, DeshadowGAN had better shadow removal capabilities than compensation as layer depth increased. This can be observed from the box plot in Figure 7, where the 25th and 75th percentiles of the intralayer contrast for DeshadowGAN gradually increased against those of compensation from the RNFL to the RPE layer.

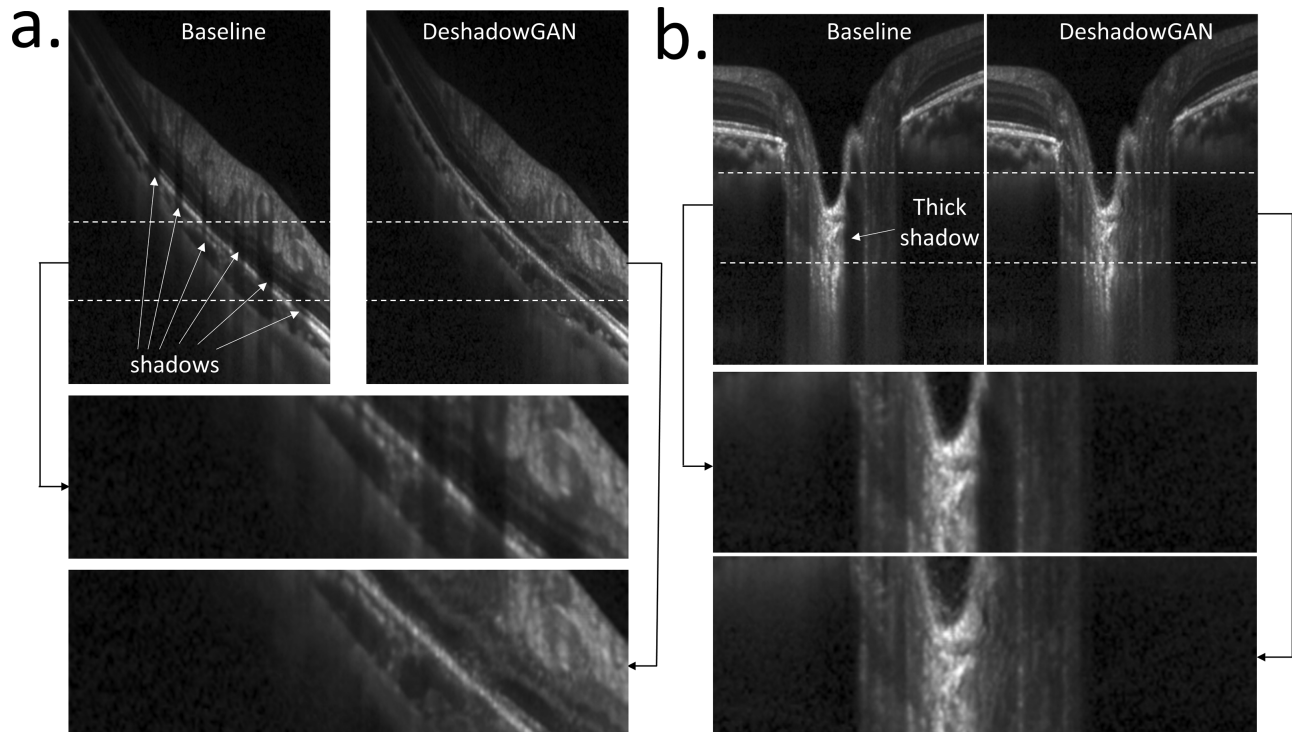


Figure 5. Images of retinal layers before and after deshadowing of (a) areas away from the optic disc and (b) areas around the optic disc.

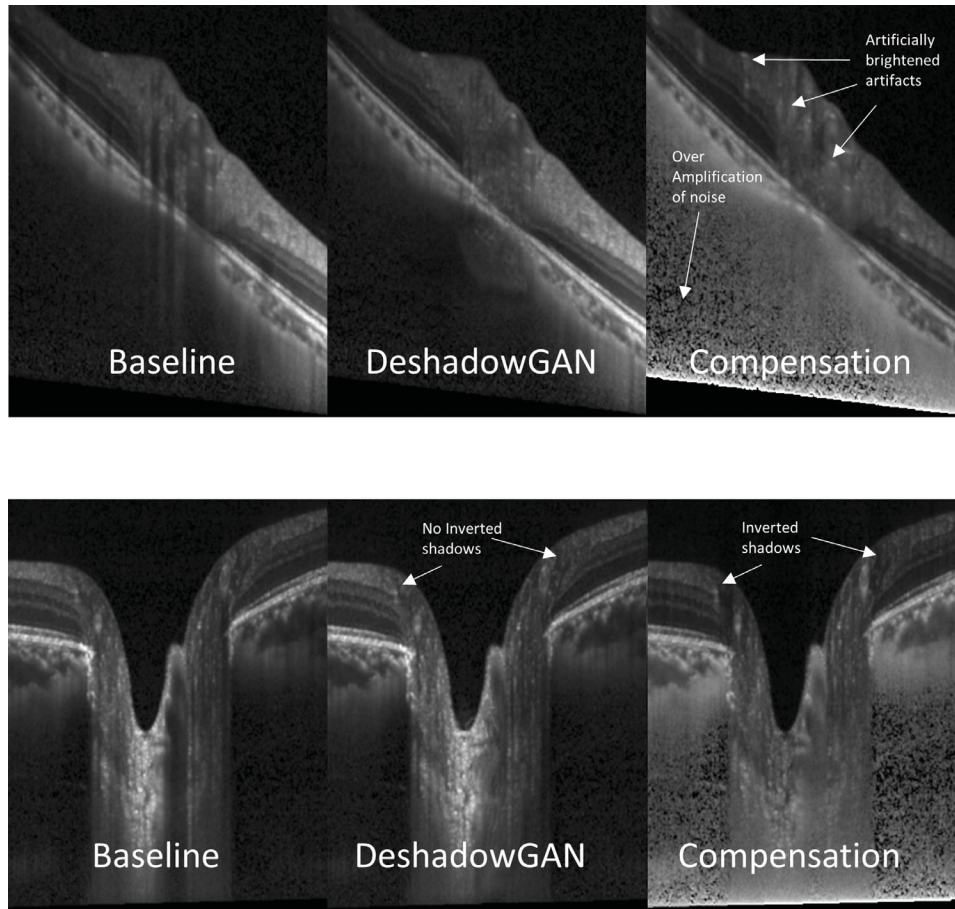


Figure 6. Compensation artifacts comparison with DeshadowGAN. (Top right) Artificially brightened artifacts and overamplification of noise in the compensated image. (Bottom right) Inverted shadows in compensated images.

Shadow removal was also qualitatively corroborated by observation of the flattened lateral pixel intensities (across shadows) for the PR layer, RPE layer, and RNFL before and after shadow removal (Fig. 8, right column). DeshadowGAN recovered the shadows to a larger extent as compared to compensation. Furthermore, we observed that compensation did not have an increase in shadow information but rather a decrease in non-shadow intensities in shallow layers, as non-shadow pixel intensities were found to be up to 50% lower after compensation.

Proof of Principle: DeshadowGAN Did Not Create Artifacts

Qualitative analysis of our results showed that no artificial anatomical information was created within deshadowed images. This can be qualitatively observed in Figure 9, where both genuine retinal shadows were retained, albeit not as clearly defined as compared to the ground truth (baseline images in this case).

Effect of Artificial Shadow Width and Exponential Decay on the PSNR

Overall, we observed that DeshadowGAN was more sensitive to shadow width as compared to shadow contrast (higher decay = higher contrast). This can be qualitatively observed in the boxplots in Figure 10.

Discussion

In this study, we proposed a novel deep learning algorithm (DeshadowGAN) with a weighted-mixture loss function to remove retinal blood vessel shadows in OCT images of the ONH. When trained with baseline OCT images and manually created binary shadow masks, DeshadowGAN improved tissue visibility under shadows at all depth, regardless of shadow width. DeshadowGAN may be considered as a preprocessing step to improve the performance of a wide range of algorithms, including those currently being

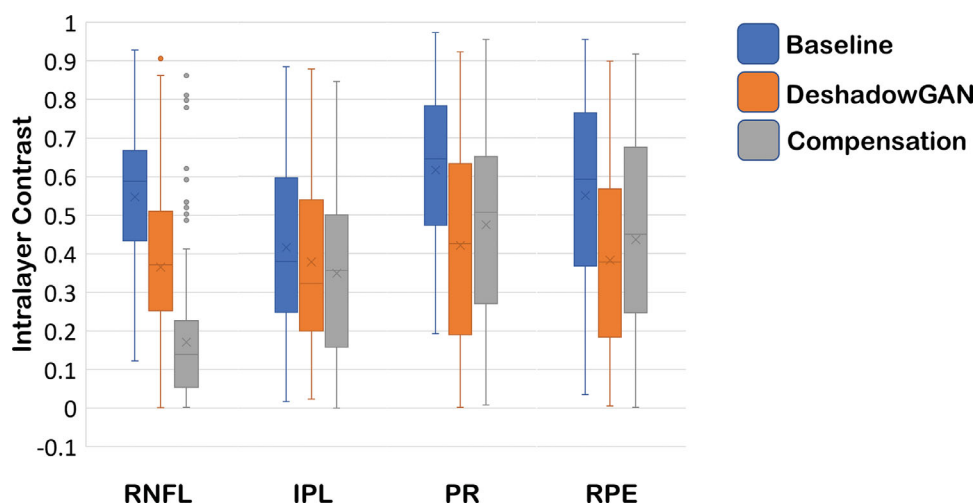


Figure 7. Intralayer contrast comparison among baseline, deshadowed, and compensated images. When compared with compensation, DeshadowGAN tends to perform better in deeper layers.

used for OCT image segmentation, denoising, and classification.

Having successfully trained, validated, and tested our algorithm with a total of 2619 baseline OCT images, we found that DeshadowGAN can be applied to new images not previously seen by the network in order to correct shadows. Furthermore, for new images, DeshadowGAN does not require any segmentation, delineation, or identification of shadows by the user. Our results confirmed consistently higher intralayer contrasts, flatter layer-wise pixel intensity profiles across shadows, and the absence of many artifacts commonly found in compensated images. Thus, we may be able to provide a robust deep learning framework to consistently remove retinal blood vessel shadows of varying sizes and intensities.

In addition, DeshadowGAN was able to successfully eliminate the deleterious effects of light attenuation affecting the visibility of retinal layers and deeper tissues such as the LC. DeshadowGAN helped substantially recover the visibility of the anterior lamina cribrosa boundary, where sensitive pathophysiologic deformation could signal the onset of early glaucoma.^{32–34} Deep collagenous tissues such as the LC and adjacent peripapillary sclera are the main load-bearing tissues of the eye in the ONH region,³⁵ and it has been reported that biomechanical and morphological changes in these tissues may serve as risk factors for glaucoma.^{36–38} The robustness of the OCT-based measurements performed on these tissues could be substantially improved after application of our proposed algorithm.

Corrected images with DeshadowGAN did not exhibit the strong artifacts that are often observed with adaptive compensation, such as inverted shadows,

hyperreflective spots, noise overamplification at high depth (see examples in Fig. 6), and hyporeflective retinal layers. For this latter case, we found that compensation can indeed reduce tissue brightness in the anterior retinal layers (while enhancing deeper connective tissue layers) by up to 50%. Brightness is typically not affected with DeshadowGAN. We also believe that compensation artifacts could cause issues for automated segmentation algorithms that rely on the presence of homogeneous pixel intensity values within the same layer.^{39–41} Because DeshadowGAN generates significantly fewer artifacts, it has the potential to be used as an artificial intelligence preprocessing step for many automated OCT applications in ophthalmology, such as, but not limited to, segmentation, denoising, signal averaging, and disease classification.^{42–46}

In addition, we observed that the performance of DeshadowGAN was slightly affected when dealing with wider shadows with stronger contrast. This was assessed with artificial cases and using the PSNR as a measure of performance. We may consider taking these parameters into account within DeshadowGAN in future work in order to improve performance.

As a first proof of principle, we also found that DeshadowGAN did not create anatomically inaccurate information under shadows and maintained all other image regions true to their original quality. This showed that our algorithm does not introduce or obscure any information within shadowed regions during deshadowing; however, this was only confirmed with artificial data by simply adding fake shadows (simulated as an exponential decay). It will also be extremely important to validate this in pathological cases to ensure that DeshadowGAN does not obscure true pathology from

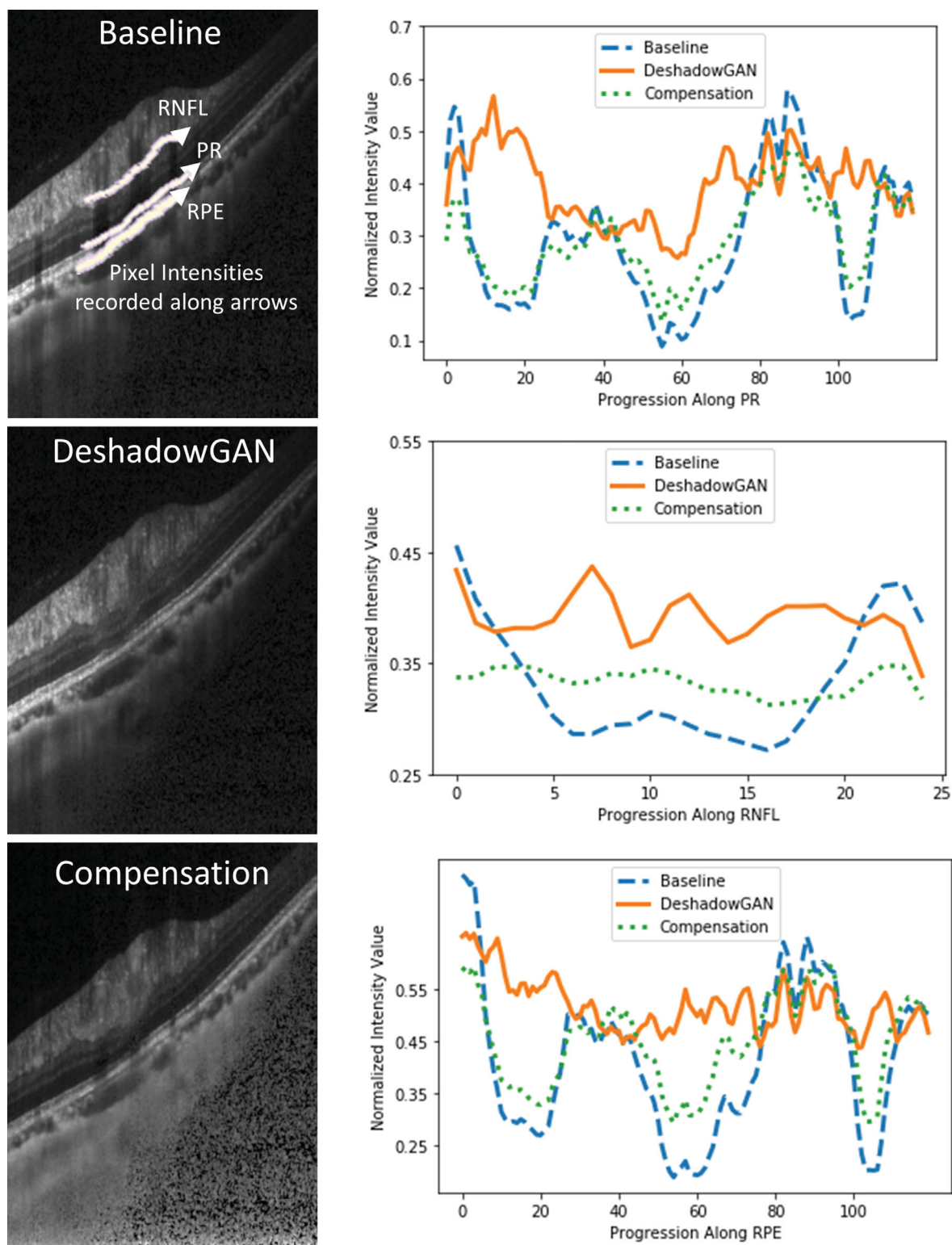


Figure 8. Layer-wise lateral pixel intensities across the PR layer, RPE layer, and RNFL. The direction of progression is along the arrow at the bottom of each image.

the shadowed region. If one wanted to confirm such results with ex- or in-vivo data, one would need to image the exact same tissue region with and without the presence of blood flow. Such experiments would

be extremely complex to perform, especially in humans in vivo, even if blood is flushed with saline temporarily (as is done with intravascular OCT). However, we understand that such validations may be necessary for

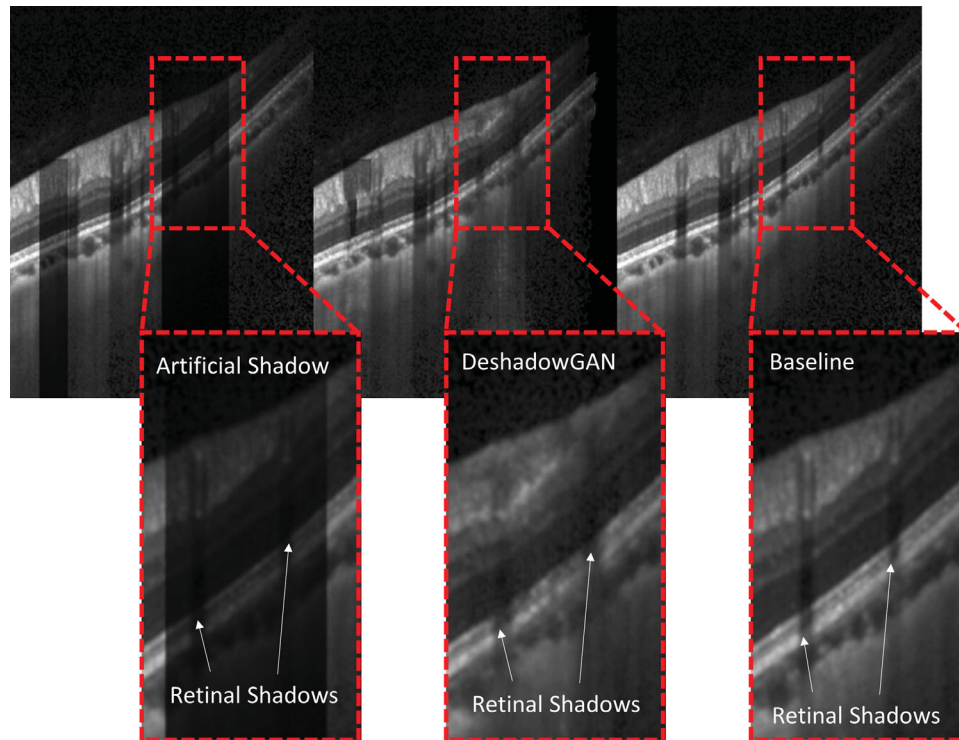


Figure 9. Artificial shadow removal experiment results. From left, the baseline with an artificial shadow, a deshadowed image from DeshadowGAN, and a baseline image without an artificial shadow.

full clinical acceptance of this methodology. From our point of view, it would also be imperative to further confirm that DeshadowGAN would not interfere with another AI algorithm aimed at improving diagnosis

or prognosis. On the other hand, it is also very possible that DeshadowGAN may increase the diagnosis or prognosis performance of other algorithms, and we hope to test such hypotheses in detail in the future.

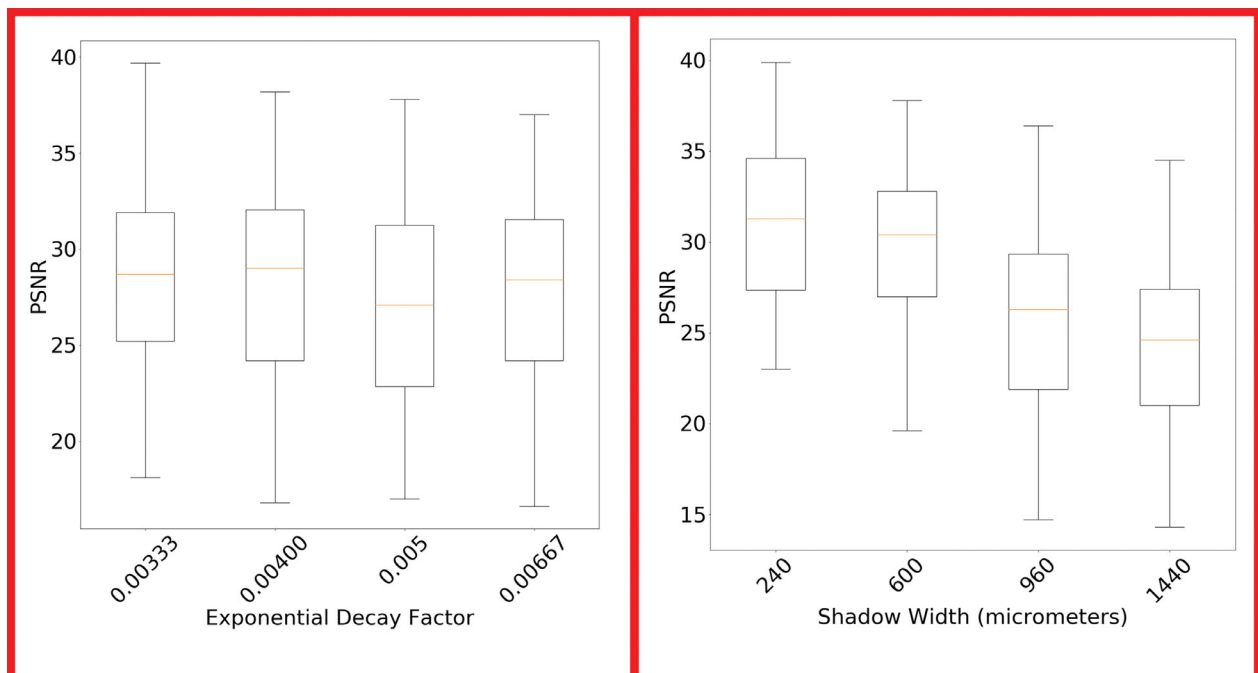


Figure 10. (Left) PSNR values as exponential decay values increased from 0.00333 to 0.00667. (Right) PSNR values as shadow width increased from 240 μm to 1440 μm .

Several limitations of this work warrant further discussion. Although DeshadowGAN has performed relatively well on baseline OCT images from healthy eyes, we cannot confirm that its performance will remain the same for eyes with pathological conditions such as glaucoma. This is because deep learning approaches respond unpredictably when the input is very different from its training images,^{47,48} and pathological training sets may be required. Furthermore, DeshadowGAN was trained on high-quality multi-frame OCT images from a single Spectralis OCT device. It is unknown if the algorithm would be able to perform as effectively if applied to OCT images obtained from other OCT devices or OCT images from the same device but with significantly less or no signal averaging. Similarly, each scenario may require a separate training set. We aim to perform further tests to assess all possible scenarios.

In conclusion, we have proposed a novel algorithm to correct blood vessel shadows in OCT images. Such an algorithm can be considered as a preprocessing step to improve the performance of a wide range of algorithms, including those currently being used for OCT image segmentation, denoising, and classification.

Acknowledgments

Supported by Singapore Ministry of Education Academic Research Funds Tier 1 (R-155-000-168-112 to AT; R-397-000-294-114 to MJAG); National University of Singapore Young Investigator Award Grants (NUSYIA FY16 P16, R-155-000-180-133 to AT; NUSYIA FY13 P03, R-397-000-174-133 to MJAG); Singapore Ministry of Education Academic Research Funds Tier 2 (R-397-000-280-112, R-397-000-308-112 to MJAG); and National Medical Research Council Grant NMRC/STAR/0023/2014 (TA).

Disclosure: **H. Cheong**, None; **S.K. Devalla**, None; **T.H. Pham**, None; **L. Zhang**, None; **T.A. Tun**, None; **X. Wang**, None; **S. Perera**, None; **L. Schmetterer**, None; **T. Aung**, None; **C. Boote**, None; **A. Thierry**, Abyss Processing (I); **M.J.A. Girard**, Abyss Processing (I)

References

1. Rezaie T, Child A, Hitchings R, et al. Adult-onset primary open-angle glaucoma caused by mutations in optineurin. *Science*. 2002;295:1077–1079.
2. Posner A, Schlossman A. Syndrome of unilateral recurrent attacks of glaucoma with cyclitic symptoms. *Arch Ophthalmol*. 1948;39:517–535.
3. Coudrillier B, Tian J, Alexander S, et al. Biomechanics of the human posterior sclera: age- and glaucoma-related changes measured using inflation testing. *Invest Ophthalmol Vis Sci*. 2012;53:1714–1728.
4. Garvin MK, Abrahmoff D, Wu X, Russell SR, Burns TL, Sonka M. Automated 3-D intraretinal layer segmentation of macular spectral-domain optical coherence tomography images. *IEEE Trans Med Imaging*. 2009;28:1436–1447.
5. Hartl I, Li XD, Chudoba C, et al. Ultrahigh-resolution optical coherence tomography using continuum generation in an air-silica microstructure optical fiber. *Opt Lett*. 2001;26:608–610.
6. Girard MJ, Strouthidis NG, Ethier CR, Mari JM, et al. Shadow removal and contrast enhancement in optical coherence tomography images of the human optic nerve head. *Invest Ophthalmol Vis Sci*. 2011;52:7738–7748.
7. Leung CK, Cheung CY, Weinreb RN, et al. Retinal nerve fiber layer imaging with spectral-domain optical coherence tomography: pattern of RNFL defects in glaucoma. *Ophthalmology*. 2010;117:2337–2344.
8. Tan MH, Ong SH, Thakku SG, et al. Automatic feature extraction of optical coherence tomography for lamina cribrosa detection. *J Image Graph*. 2015;3:102–106.
9. Mujat M, Chan R, Cense B, et al. Retinal nerve fiber layer thickness map determined from optical coherence tomography images. *Opt Express*. 2005;13:9480–9491.
10. Fabritius T, Makita S, Hong Y, Myllylä R, Yasuno Y. Automated retinal shadow compensation of optical coherence tomography images. *J Biomed Opt*. 2009;14:010503.
11. Mari JM, Strouthidis NG, Park SC, Girard MJ. Enhancement of lamina cribrosa visibility in optical coherence tomography images using adaptive compensation. *Invest Ophthalmol Vis Sci*. 2013;54:2238–2247.
12. Wang J, Li X, Yang J. Stacked conditional generative adversarial networks for jointly learning shadow detection and shadow removal. In: Proceedings of the IEEE International Conference on Computer Vision (ICCV). Piscataway, NJ: Institute of Electrical and Electronics Engineers. 2018;1788–1797.
13. Qu L, Tian J, He S, et al. DeshadowNet: a multi-context embedding deep network for shadow removal. In: Proceedings of the IEEE International

- Conference on Computer Vision (ICCV). Piscataway, NJ: Institute of Electrical and Electronics Engineers. 2017;2308–2316.
14. Goodfellow I, Pouget-Abadie J, Mirza M, et al. Generative adversarial nets. In: Ghahramani Z, Welling M, Cortes C, Lawrence ND, Weinberger KQ, eds. *Advances in Neural Information Processing Systems* 27. Red Hook, NY: Curran. 2015;2672–2680.
 15. Nguyen V, Vicente TFY, Zhao M, et al. Shadow detection with conditional generative adversarial networks. In: *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*. Piscataway, NJ: Institute of Electrical and Electronics Engineers. 2017;4520–4528.
 16. Le H, Vicente TFY, Nguyen V, et al. A+D net: training a shadow detector with adversarial shadow attenuation. In: *Proceedings of the European Conference on Computer Vision (ECCV)*. Berlin, Germany: Springer. 2018;662–678.
 17. Zhang H, Sindagi V, Patel VM. Image de-raining using a conditional generative adversarial network. *IEEE Trans Circuits Syst Video Technol*. 2019, doi: [10.1109/TCSVT.2019.2920407](https://doi.org/10.1109/TCSVT.2019.2920407).
 18. Schindelin J, Arganda-Carreras I, Frise E, et al. Fiji: an open-source platform for biological-image analysis. *Nat Methods*. 2012;9:676–682.
 19. Falk T, Mai D, Bensch R, et al. U-Net: deep learning for cell counting, detection, and morphometry. *Nat Methods*. 2019;16:67–70.
 20. Mannor S, Peleg D, Rubinstein R. The cross entropy method for classification. In: *Proceedings of the 22nd International Conference on Machine Learning*. New York, NY: Association for Computing Machinery. 2005;561–568.
 21. Zeiler MD, Ranzato M, Monga R, et al. On rectified linear units for speech processing. In: *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. Piscataway, NJ: Institute of Electrical and Electronics Engineers. 2013;3517–3521.
 22. Devalla SK, Subramanian G, Pham TH, et al. A deep learning approach to denoise optical coherence tomography images of the optic nerve head. *Sci Rep*. 2019;9:14454.
 23. Kim H, Garrido P, Tewari A, et al. Deep video portraits. *ACM Trans Graph*. 2018;37:163.
 24. Ronneberger O, Fischer P, Brox T. U-Net: convolutional networks for biomedical image segmentation. In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*. New York, NY: Springer. 2015;234–241.
 25. Paszke A, Gross S, Chintala S, et al. Automatic differentiation in PyTorch. In: *Proceedings of the NIPS 2017 Autodiff Workshop*. 2017;1–4.
 26. Johnson J, Alahi A, Fei-Fei L. Perceptual losses for real-time style transfer and super-resolution. In: *Proceedings of the European Conference on Computer Vision (ECCV)*. Berlin, Germany: Springer. 2016;1–18.
 27. Szegedy C, Ioffe S, Vanhoucke V, Alemi AA. Inception-v4, Inception-ResNet and the impact of residual connections on learning. In: *AAAI'17: Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence*. 2017;4278–4284.
 28. Johnson J, Alahi A, Fei-Fei L. Perceptual losses for real-time style transfer and super-resolution. In: *Proceedings of the European Conference on Computer Vision (ECCV)*. Berlin, Germany: Springer. 2016;1–18.
 29. Liu G, Reda FA, Shih KJ, et al. Image inpainting for irregular holes using partial convolutions. In: *Proceedings of the European Conference on Computer Vision (ECCV)*. Berlin, Germany: Springer. 2018;89–105.
 30. Kingma DP, Ba J. Adam: a method for stochastic optimization. In: *International Conference on Learning Representations*. 2015;1–13.
 31. Girard MJ, Ang M, Chung CW, et al. Enhancement of corneal visibility in optical coherence tomography images using corneal adaptive compensation. *Transl Vis Sci Technol*. 2015;4:3.
 32. Bellezza AJ, Rintalan CJ, Thompson HW, Downs JC, Hart RT, Burgoyne CF. Deformation of the lamina cribrosa and anterior scleral canal wall in early experimental glaucoma. *Invest Ophthalmol Vis Sci*. 2003;44:623–637.
 33. Jonas JB, Berenshtein E, Holbach L. Anatomic relationship between lamina cribrosa, intraocular space, and cerebrospinal fluid space. *Invest Ophthalmol Vis Sci*. 2003;44:5189–5195.
 34. Quigley HA, Hohman RM, Addicks EM, Massof RW, Green WR. Morphologic changes in the lamina cribrosa correlated with neural loss in open-angle glaucoma. *Am J Ophthalmol*. 1983;95:673–691.
 35. Girard MJ, Strouthidis NG, Ethier CR, Mari JM. Shadow removal and contrast enhancement in optical coherence tomography images of the human optic nerve head. *Invest Ophthalmol Vis Sci*. 2011;52:7738–7748.
 36. Burgoyne CF, Downs JC, Bellezza AJ, Suh JK, Hart RT. The optic nerve head as a biomechanical structure: a new paradigm for understanding the role of IOP-related stress and strain in the patho-

- physiology of glaucomatous optic nerve head damage. *Prog Retin Eye Res.* 2005;24:39–73.
37. Sigal IA. Interactions between geometry and mechanical properties on the optic nerve head. *Invest Ophthalmol Vis Sci.* 2009;50:2785–2795.
 38. Ethier CR. Scleral biomechanics and glaucoma—a connection? *Can J Ophthalmol.* 2006;41:14.
 39. Mirsharif Q, Tajeripour F. Investigating image enhancement methods for better classification of retinal blood vessels into arteries and veins. In: *The 16th CSI International Symposium on Artificial Intelligence and Signal Processing (AISP 2012)*. Piscataway, NJ: Institute of Electrical and Electronics Engineers. 2012;591–597.
 40. Van de Voorde T, De Genst W, Canters F. Improving pixel-based VHR land-cover classifications of urban areas with post-classification techniques. *Photogramm Eng Remote Sensing.* 2007;73:1017–1027.
 41. Shahtahmassebi A, Yang N, Wang K, Moore N, Shen Z. Review of shadow detection and deshadowing methods in remote sensing. *Chinese Geogr Sci.* 2013;23:403–420.
 42. Zijdenbos AP, Dawant BM. Brain segmentation and white matter lesion detection in MR images. *Crit Rev Biomed Eng.* 1994;22:401–465.
 43. Giani A, Cigada M, Esmaili DD, et al. Artifacts in automatic retinal segmentation using different optical coherence tomography instruments. *Retina.* 2010;30:607–616.
 44. Lee C, Zhang YT. Reduction of motion artifacts from photoplethysmographic recordings using a wavelet denoising approach. In: *IEEE EMBS Asian-Pacific Conference on Biomedical Engineering, 2003*. Piscataway, NJ: Institute of Electrical and Electronics Engineers. 2003;194–195.
 45. Dong C, Deng Y, Loy CC, Tang X. Compression artifacts reduction by a deep convolutional network. In: *2015 IEEE International Conference on Computer Vision (ICCV)*. Piscataway, NJ: Institute of Electrical and Electronics Engineers. 2015;576–584.
 46. Gjestebj L, Yang Q, Xi Y, Zhou Y, Zhang J, Wang G. Deep learning methods to guide CT image reconstruction and reduce metal artifacts. In: Flohr TG, Lo JY, Schmidt TG, eds. *Medical Imaging 2017: Physics of Medical Imaging*. Bellingham, WA: International Society for Optics and Photonics. 2017;1–7.
 47. Shwartz-Ziv R, Tishby N. This paper was done with the support of the Intel Collaborative Research institute for Computational Intelligence (ICRI-CI) and is part of the “Why & When Deep Learning works: looking inside Deep Learning” ICRI-CI paper bundle.. arXiv preprint arXiv:1703.00810, 2017.
 48. Glorot X, Bengio Y. Understanding the difficulty of training deep feedforward neural networks. *J Mach Learn Res.* 2010;9:249–256.