Real-time OCT image denoising using a self-fusion neural network

JOSE J. RICO-JIMENEZ,[1] ⬤ DEWEI HU,[2] ERIC M. TANG,[1] ⬤ IPEK OGUZ,[2] AND YUANKAI K. TAO[1],*

[1]*Department of Biomedical Engineering, Vanderbilt University, Nashville, TN 37232, USA*
[2]*Department of Electrical Engineering and Computer Science, Vanderbilt University, Nashville, TN 37235 USA, USA*
*yuankai.tao@vanderbilt.edu

Abstract: Optical coherence tomography (OCT) has become the gold standard for ophthalmic diagnostic imaging. However, clinical OCT image-quality is highly variable and limited visualization can introduce errors in the quantitative analysis of anatomic and pathologic features-of-interest. Frame-averaging is a standard method for improving image-quality, however, frame-averaging in the presence of bulk-motion can degrade lateral resolution and prolongs total acquisition time. We recently introduced a method called self-fusion, which reduces speckle noise and enhances OCT signal-to-noise ratio (SNR) by using similarity between from adjacent frames and is more robust to motion-artifacts than frame-averaging. However, since self-fusion is based on deformable registration, it is computationally expensive. In this study a convolutional neural network was implemented to offset the computational overhead of self-fusion and perform OCT denoising in real-time. The self-fusion network was pretrained to fuse 3 frames to achieve near video-rate frame-rates. Our results showed a clear gain in peak SNR in the self-fused images over both the raw and frame-averaged OCT B-scans. This approach delivers a fast and robust OCT denoising alternative to frame-averaging without the need for repeated image acquisition. Real-time self-fusion image enhancement will enable improved localization of OCT field-of-view relative to features-of-interest and improved sensitivity for anatomic features of disease.

1. Introduction

Optical coherence tomography (OCT) has become ubiquitous in ophthalmic diagnostic imaging over the last three decades [1,2]. However, clinical OCT image-quality is highly variable and often degraded by inherent speckle noise [3,4], bulk-motion artifacts [5–7], and ocular opacities/pathologies [8,9]. Poor image-quality can limit visualization and introduce errors in quantitative analysis of anatomic and pathologic features-of-interest. Averaging of multiple repeated frames acquired at the same or neighboring locations is often used to increase signal-to-noise ratio (SNR) [10–12]. However, frame-averaging in the presence of bulk-motion can degrade lateral resolution and prolongs total acquisition time, which can make clinical imaging challenging or impossible in certain patient populations.

Many computational techniques for improving ophthalmic OCT image-quality have been previously described, including compressed sensing, filtering, and model-based methods [13]. One compressed sensing approach creates a sparse representation dictionary of high-SNR images that is then applied to denoise neighboring low-SNR B-scans [14]. However, this method requires a non-uniform scan pattern to slowly capture high-SNR B-scans to create a sparse representation dictionary, which limits its robustness in clinical applications. A different dictionary-based approach obviates the need for high-SNR B-scans and frame-averaging by utilizing K-SVD dictionary learning and curvelet transform for denoising OCT [15]. Various well-known image denoising filters, such as Block Matching 3-D and Enhanced Sigma Filters combined with the Wavelet Multiframe algorithm, have also been evaluated for OCT denoising [16]. However,

these aforementioned methods are computationally expensive, particularly when combined with wavelet-based compounding algorithms. Patch-based approaches, such as the spatially constrained Gaussian mixture model [17] and non-local weighted group low-rank representation [18], have also been proposed for OCT image-enhancement but have similar computational overhead and are, thus, unsuitable for real-time imaging applications.

Deep-learning based denoising has gained popularity in medical imaging [19–21] and OCT [22,23] applications. In addition to producing highly accurate results, these methods also overcome the computational burden of traditional methods and enable real-time processing once a model is trained. Convolutional Neural Networks (CNNs) for OCT denoising have included a variety of network architectures such as GANS [11,24–28], MIFCN [29], DeSpecNet [30], DnCNN [10,31], Noise2Noise [32], GCDS [33], U-NET [34–37], and intensity fusion [38,39]. The critical barrier to translating deep-learning methods to clinical OCT is the lack of an ideal reference image to be used as the ground-truth. To overcome this limitation, some CNN-based methods use averaged frames as the ground-truth to mimic frame-averaging image-quality. Recently, a ground-truth free CNN method was demonstrated that uses one raw image as the input and a second raw image as the ground-truth [40]. Here, the image-enhancement benefit was sacrificed in favor of faster processing rates. This trade-off highlights the limitations of current-generation deep-learning based OCT denoising methods, which require a large number of repeated input images to compute the ground-truth and a small number of input images to achieve real-time processing rates but can result in artifactual blurring (Table 1).

Table 1. Summary of current-generation neural network architectures for OCT denoising.

Architecture	Num. repeated images for ground truth	Short description	Ref.
U-net	128	U-net	[34]
SM-GAN	100	Speckle-modulated GAN	[11]
DRUNET	75	Dilated-Residual U-Net	[35,36]
DnCNN	50	Feed-forward CNN with a perceptually-sensitive loss function	[10]
SiameseGAN	40	GAN with a siamese twin network	[27]
GCDS	40	Gated convolution–deconvolution structure	[33]
DeSpecNet	20	Residue-learning-base deconvolution	[30]
Caps-cGAN	20	Capsule conditional GAN	[24]
cGAN	20	Edge-sensitive cGAN	[25]
BRUNet	9	Branch Residual U-shape Network	[37]
CNN-MSE, **CNN-WGAN**	6	MSE and Wasserstein GANs	[28]
SRResNet	1	Super-resolution residual network	[40]

Oguz et al. [41] recently demonstrated robust OCT image-enhancement using self-fusion. This method is based on multi-atlas label fusion [42], which exploits the similarity between adjacent B-scans. Self-fusion does not require repeat-frame acquisition, is edge preserving, enhances retinal layers, and significantly reduces speckle noise. The main limitation of self-fusion is its computationally complexity due to the required deformable image-registration and similarity computations, which precludes real-time OCT applications. Our inability to access image-enhanced OCT images in real-time significantly limits the utility of self-fusion for evaluating dynamic retinal changes. Similarly, because real-time image aiming and guidance is performed using noisy raw OCT cross-sections, it is challenging to accurately evaluate image focus, whether the field-of-view sufficiently covers features-of-interest, and image quality after self-fusion, thus reducing yield of usable clinical datasets. In this study, we overcome processing limitations of

self-fusion by developing a CNN that uses denoised self-fusion images as the ground-truth [41]. This approach combines the robustness of self-fusion denoising and the high processing-speed of neural networks. Here, we demonstrate integration and translation of optimized data acquisition and processing for real-time self-fusion image-enhancement of ophthalmic OCT at ~ 22 fps for an image size of 512×512 pixels. While potentially more prone to artifacts from implementation of a neural network, our proposed approach enables real-time denoising of raw OCT images, which can directly be used as an indicator of image quality following artifact-free offline self-fusion processing of the acquired data. Similar strategies have been demonstrated in OCT angiography applications to provide previews of volumetric vascular projection maps in real-time [43,44]. This real-time denoising technology can also enhance diagnostic utility in applications that require immediate feedback or intervention, such as during OCT-guided therapeutics or surgery [45,46].

2.   Materials and methods

2.1.   OCT system

All images were acquired with a handheld OCT system previously reported in [47,48] with a 200 kHz, 1060-nm center wavelength swept-source laser (Axsun) optically buffered to 400 KHz. The OCT signal was detected using a 1.6 GHz balanced photodiode (APD481AC, Thorlabs) and discretized with a 12-bit dual-channel 4 GS/s waveform digitizer board (ATS-9373, AlazarTech). The OCT sample-arm beam was scanned using a galvanometer pair and relayed using a 2x demagnifying telescope to a 2 mm diameter spot at the pupil. All human imaging data were acquired under a protocol approved the Vanderbilt University Institutional Review Board.

2.2.   Dataset

Volumetric OCT datasets of healthy human retina centered on the fovea and optic nerve head (ONH) were acquired to train and test the self-fusion neural network. OCT optical power incident on the pupil was attenuated (1-2 mW) to simulate different SNR levels. Each volume contained 2500 raw B-scans (500 sets of 5-repeated frames). The repeated frames were averaged and the resulting 500-frame volume was self-fused with a radius of 3 frames (3 adjacent images before and 3 after the current frame for total of 7 images) to achieve high quality ground-truth images to train the neural network. The supplemental Fig. 1 shows the effect of self-fusion with different radii. Sets of 3 raw and non-averaged B-scans (radius of 1) were used as self-fusion neural network inputs to obtain one denoised image. Figure 1 shows examples of raw and corresponding self-fused OCT B-scans used as ground-truth images.

Additional OCT images from external datasets were added as test images. The first set of images was taken from the dataset used to test the Sparsity Based Simultaneous Denoising and Interpolation (SBSDI) method [49]. Two sets of images acquired with different OCT systems (Cirrus: Zeiss Meditec; T-1000 and T-2000: Topcon) were taken from the Retinal OCT Fluid Detection and Segmentation Benchmark and Challenge (RETOUCH) dataset [50].

2.3.   Network architecture and training

The network was designed and implemented in PyTorch based on the multi-scale U-Net architecture proposed by Devalla et al. [35]. The model was trained on 9 ONH volumes with various SNR and validated on 3 fovea volumes to avoid information leakage (Fig. 2(A)). Additionally, 3 ONH and 3 fovea volumes were used as test data. The self-fusion neural network was trained on a RTX 2080 Ti 11GB GPU (NVIDIA) until the loss-function reached the plateau (30 epochs). Parameters in the network were optimized using the Adam optimization algorithm with a starting learning rate of 1e-3 and decay factor of 0.8 for every epoch. Batches of 3 adjacent registered OCT B-scans (radius of 1) were used as inputs to train the network. Here, the
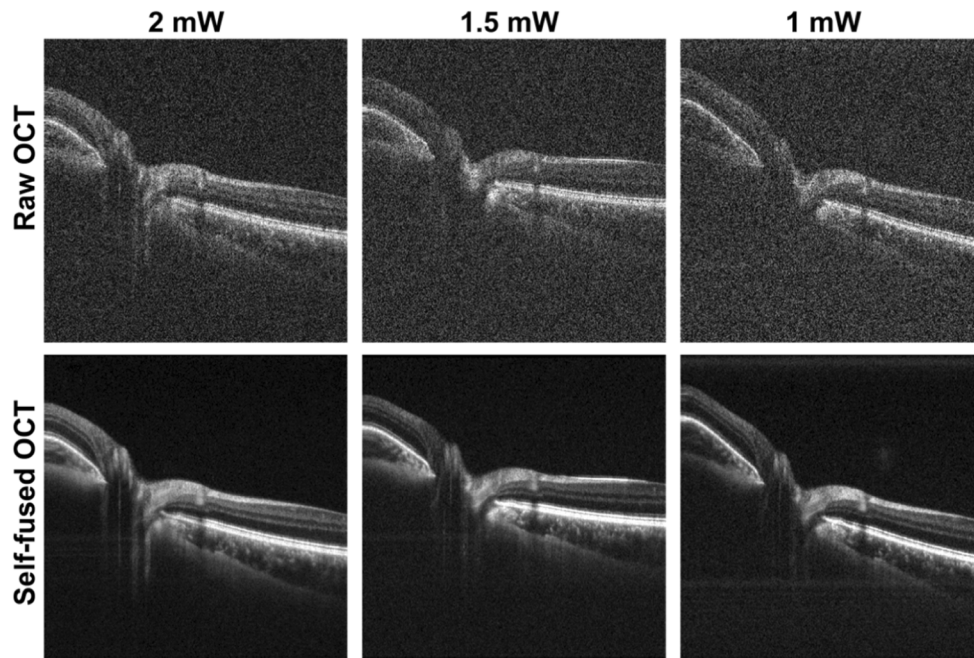
Fig. 1. Representative raw and self-fused OCT B-scans used for training for different optical power levels incident on the pupil.

central frame was denoised based on information from the neighboring slices. The number of requisite input B-scans is kept low to achieve video-rate self-fusion processing. Similarly, while deformable image-registration is ideal for denoising and, in this case, used to get the ground truth self-fused images (Fig. 1), a discrete Fourier transform (DFT) based rigid image-registration [51] was adopted as the motion correction strategy to minimize computational overhead. The computational cost of the original self-fusion method using rigid DFT registration or deformable registration (Symmetric Normalization - SyN) ANTsPy [52] is compared in Table 2. The computer used for this test had an 11th Gen IntelCore i7-11700 @ 2.50GHz×16 CPU.

Table 2. Average offline processing time of self-fusion using rigid DFT registration and deformable registration.

	3-Frame Registration	3-Frame Self-fusion with registration
Rigid DFT Registration	0.004 s.	0.654 s. (1.52 fps)
ANTsPy Deformable Registration	1.714 s.	2.392 s. (0.42 fps)

2.4. Real-time implementation

Custom C++ software was used to acquire and process OCT images. TorchScript was used to create a serializable version of the self-fusion neural network model that could be used in C++ via LibTorch (C++ analog of PyTorch). The model was loaded in a LibTorch-based module and executed to denoise and display self-fusion denoised OCT images (Fig. 2(B)). The OCT acquisition and processing software consists of a main thread that controls the graphical user interface and image visualization, and two sub-threads running asynchronously to control data acquisition (DAQ) and processing. The DAQ module acquired 16-bit integer raw OCT
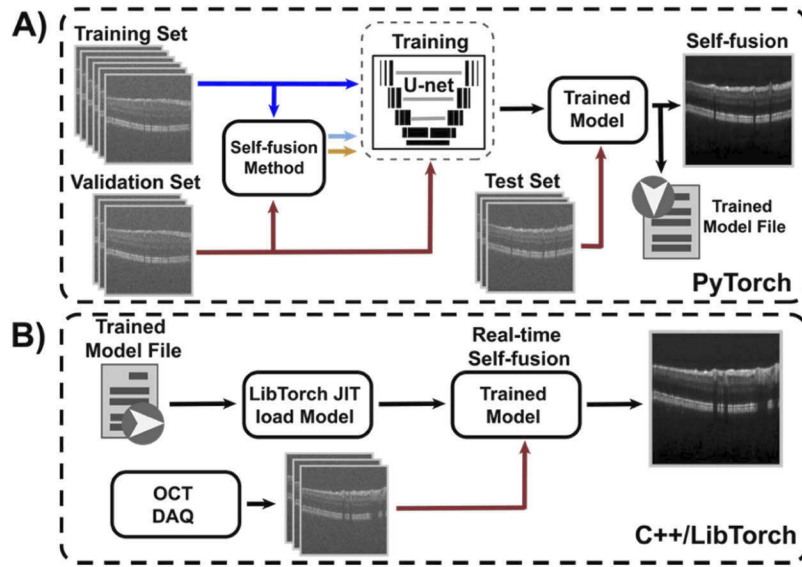
Fig. 2. Self-fusion neural network. A) Training pipeline. B) Real-time self-fusion with the pre-trained model integrated with OCT data acquisition.

data with 2560 pixels/A-lines and 512 A-lines/B-scan. Sets of three images were acquired and copies of these images were loaded into 32-bit float LibTorch-GPU tensors. LibTorch-based GPU-accelerated OCT processing pipeline includes: background subtraction, Hanning spectral windowing, dispersion-compensation, Fourier transform, logarithmic compression, and image cropping (Fig. 3). The resulting images were then intensity-normalized, motion-corrected using DFT fast registration, denoised using the pretrained self-fusion neural network, and finally contrast-adjusted using the 1st and 99th percentiles of the image data as lower and upper intensity limits respectively.

2.5. Quantitative evaluation

The two most commonly used quantitative metrics for assessment of noise reduction were adopted to evaluate the performance of the self-fusion neural network:

Peak-signal-to-noise ratio (PSNR):

$$NR = 10\,Log_{10}\left(\frac{max(I_f)^2}{\sigma_b^2}\right) \qquad (1)$$

where $max(I_f)$ denotes the maximum foreground intensity and σ_b denotes the standard deviation of the background.

Contrast-to-noise ratio (CNR):

$$CNR = |\mu_f - \mu_b|\,/\sqrt{0.5(\sigma_f^2 - \sigma_b^2)} \qquad (2)$$

where μ_f and μ_b are the mean of the foreground and background, and σ_f and σ_b. are the standard deviation of the foreground and background, respectively.
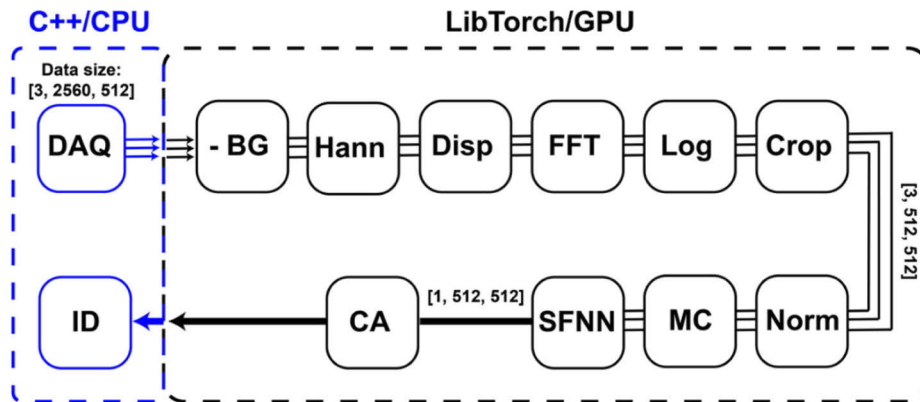
Fig. 3. Real-time self-fusion neural network processing pipeline. Three OCT B-scans were acquired and loaded into LibTorch-GPU tensors. The LibTorch-based OCT processing that included background subtraction (-BG), Hanning spectral window (Hann), dispersion-compensation (Disp), Fourier transform (FFT), logarithmic compression (Log), and image cropping (Crop). The input images for self-fusion neural network were then normalized (Norm), motion-corrected (MC), processed with the pretrained self-fusion neural network (SFNN), and contrast-adjusted (CA) before being displayed (ID).

3. Results

Figure 4 depicts a representative set of images of the fovea processed with the self-fusion neural network. All images were contrast-adjusted using the same percentiles.
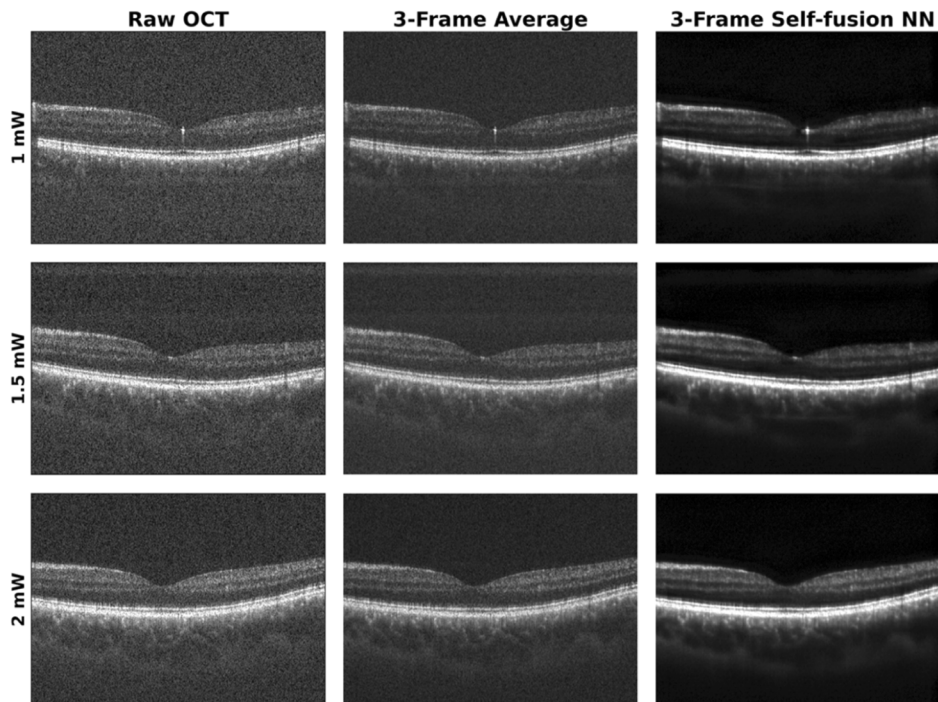

Fig. 4. Representative raw, average, and self-fusion neural network denoised OCT images of the fovea taken using different laser powers on the pupil.

The CNR and PSNR were calculated for the raw, 3-frame average, and 3-frame self-fusion neural network denoised OCT B-scans. The quantitative comparison of frame-averaging and self-fusion with respect to the original raw OCT image is summarized in Fig. 5. Experimental results showed CNR improved by ~50% and ~100% for frame-averaging and self-fusion over raw OCT B-scans, respectively. Likewise, self-fusion outperformed frame-averaging by improving PSNR by ~90% and ~20%, respectively. The contrast of retinal layers and vessels was improved, which facilitates the identification of anatomical and potentially pathological features.
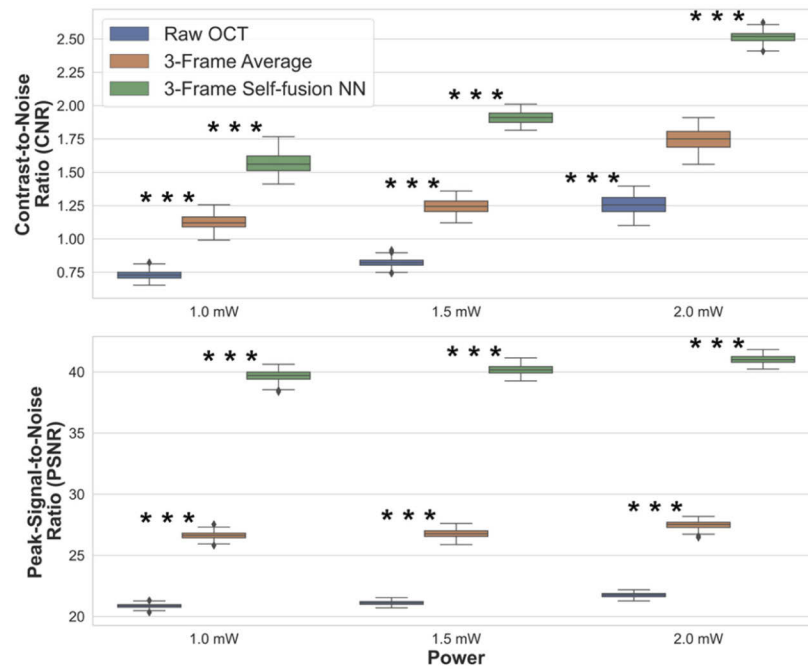

Fig. 5. Quantitative evaluation for OCT images acquired with different laser power on the pupil (*** $P \leq 0.001$).

Self-fusion neural network processing time was also compared to off-line processing (Table 3). The average of individual processing time and frame-rate for the OCT processing, DFT registration, and self-fusion neural network was quantified using 100 frames an OCT testing dataset. The CPU and GPU used for benchmarking were a XEON E5-2630 v4 2.2 GHz (Intel) and GeForce RTX 2080 Ti (NVIDIA), respectively. The results showed that self-fusion neural network can achieve near video-rate performance at ~22 fps (Visualization 1 in supplemental material).

Table 3. Average processing time and frame-rate for OCT processing and self-fusion neural network denoising on CPU and GPU.

	3-Frame OCT Processing	3-Frame DFT Registration	3-Frame Self-fusion NN	3-Frame SFNN Total
Python/CPU	0.048 s	0.060 s	1.630 s	1.741 s (0.58 fps)
Python/GPU	0.009 s	0.004 s	0.055 s	0.068 s (15.85 fps)
C++/GPU	**0.008 s**	**0.003 s**	**0.033 s**	**0.044 s (22.49 fps)**

The total processing time was calculated as the sum of the three most critical processing blocks (OCT processing, DFT registration, and self-fusion NN) plus trivial memory-to-memory data transfer time. For real-time processing, DFT registration was selected over deformable registration to reduce computation time. The results reported in Tables 2 and 3 demonstrate the image processing advantage of the self-fusion neural network over original self-fusion method. As expected, the image quality of self-fused images increases with the use of deformable registration at the expense of increased processing time.

The self-fused images from the external datasets are illustrated in Fig. 6. A set of three raw OCT images from each dataset was registered and self-fused with the neural network. Improvements in image quality and reduction in speckle noise in the SBSDI images self-fused with the neural network show enriched visualization of blood vessels and retinal layers in the fovea. The images taken with the Cirrus system show the retina with improved visualization of macular edema. The images acquired with the Topcon system show abnormal RPE, where drusenoid deposits are clearly seen in the denoised image as a smooth dome-shaped elevation. In all cases, the self-fusion neural network outperformed averaging in terms of CNR and PSNR enhancing retinal features. Therefore, the images self-fused with the neural network provided enhanced
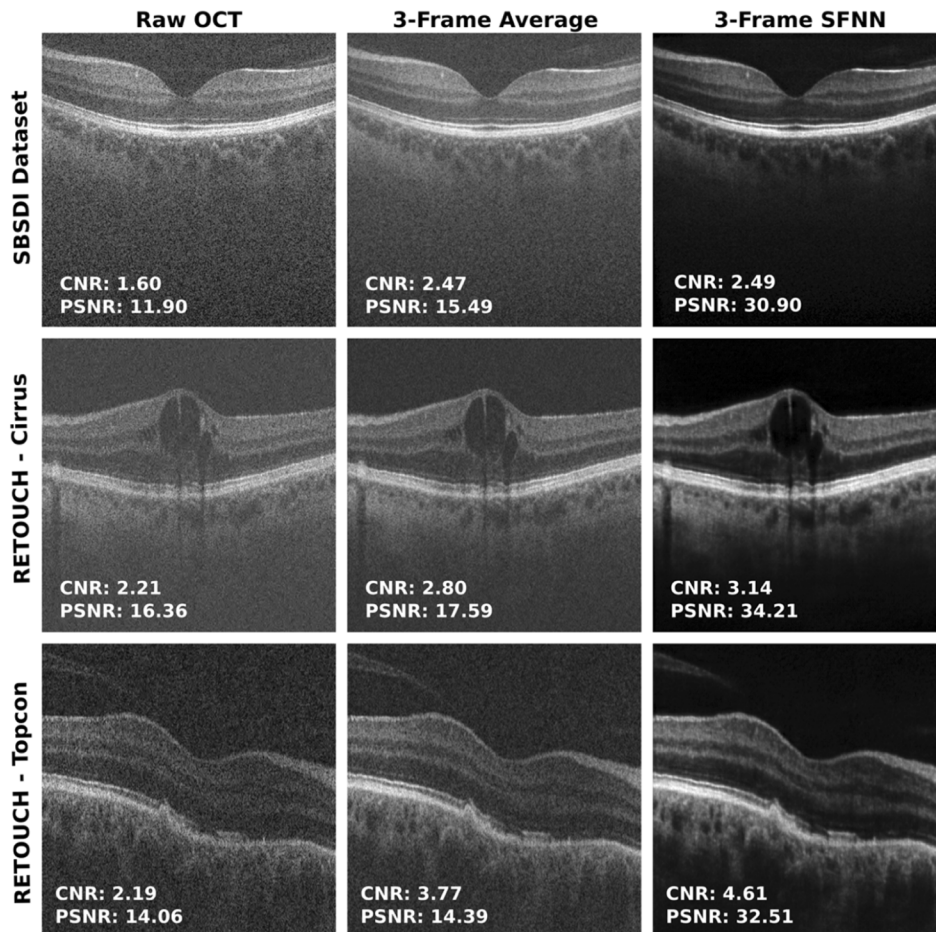

Fig. 6. Raw, average, and self-fusion neural network denoised OCT images of external datasets. Images from the SBSDI dataset (first row) and RETOUCH dataset (Cirrus: middle row; Topcon: bottom row).

visualization of diagnostic-relevant pathological features in images with different image-quality and limited visualization.

4. Discussion

Noise and poor image-quality can limit accurate identification and quantitation of pathological features on ophthalmic OCT. While robust denoising methods are well-established, these are limited to off-line implementations due to long processing times and high computational complexity. Real-time OCT image-enhancement is critical for clinical imaging to ensure patient data is of sufficient quality to perform structural and functional diagnostics in post-processing. These benefits are even more critical for OCT-guided applications, such as in ophthalmic surgery where image-quality is degraded by ocular opacities [53–55].

Deep-learning methods have shown potential for real-time image denoising. However, existing methods have shown a tradeoff between preserving structural details and reducing noise that can result in over-smoothing and loss of resolution. More importantly, deep-learning based methods require robust training data of ocular pathologies to avoid inclusion of unwanted artifacts. In this study, we implemented real-time OCT image-enhancement at near video-rates based on self-fusion. Self-fusion is more robust to motion-artifacts as compared to frame-averaging and overcomes the need for extensive training by using similarity between adjacent OCT B-scans to improve image-quality. These benefits were confirmed experimentally with our video-rate self-fusion implementation, and we show significant advantages in CNR and PSNR over frame-averaging.

We demonstrated the ability of the self-fusion neural network to denoise OCT images not only from our research grade systems but also from external datasets acquired with different commercial OCT technology. Although the neural network was trained on images from healthy human retina, the denoised external OCT images present relevant pathological features that were enhanced with the neural network such as vascularization, layer detachment, macular holes, and drusenoid deposits.

While the proposed self-fusion neural network outputs suffer from a slight image-smoothing effect produced by convolution and rigid registration when compared to self-fusion, better generalization of the data by using more images, more robust network architectures, data augmentation, a larger training database, and more images as input channels for the neural network may help to preserve features [35]. In addition, the use of more powerful GPUs will enable increasing the number of input images, which can reduce smoothing artifacts without sacrificing processing speed. The proposed method may also be directly applied to OCT variants such as OCT angiography, Doppler OCT, OCT elastography, and polarized-sensitive OCT to improve image-quality and diagnostic utility.

Our results showed a significant improvement in CNR and PSNR in the self-fused B-scans over the frame-averaged and raw B-scans, where reduced speckle noise and improved contrast benefit identification of anatomical features such as retinal layers, vessels, and potential pathologic features. The proposed approach delivers a fast and robust OCT denoising alternative to frame-averaging without the need for multiple repeated image acquisition at the same location. While we expect a few image artifacts from our neural-network implementation, conventional offline self-fusion may be directly applied to corresponding datasets that require quantitative analyses or precision diagnostic feature extraction in post-processing. Real-time self-fusion image enhancement will enable improved localization of OCT field-of-view relative to features-of-interest and improved sensitivity for anatomic features.

Disclosures. The authors declare no conflicts of interest.

Data availability. Data underlying the results presented in this paper are not publicly available at this time but may be obtained from the corresponding author upon reasonable request.

Supplemental document. See Supplement 1 for supporting content.

References

1. D. Thomas and G. Duguid, "Optical coherence tomography: a review of the principles and contemporary uses in retinal investigation," Eye **18**(6), 561–570 (2004).
2. M. Adhi and J. S. Duker, "Optical coherence tomography: current and future applications," Current Opinion in Ophthalmology **24**(3), 213–221 (2013).
3. B. Karamata, K. Hassler, M. Laubscher, and T. Lasser, "Speckle statistics in optical coherence tomography," JOSA A **22**(4), 593–596(2005).
4. J. M. Schmitt, S. H. Xiang, and K. M. Yung, "Speckle in optical coherence tomography," J. Biomed. Opt. **4**(1), 95 (1999).
5. T. T. Hormel, D. Huang, and Y. Jia, "Artifacts and artifact removal in optical coherence tomographic angiography," Quant. Imaging Med. Surg. **11**(3), 1120–1133 (2020).
6. R. F. Spaide, J. G. Fujimoto, N. K. Waheed, R. F. Spaide, J. G. Fujimoto, and N. K. Waheed, "Image artifacts in optical coherence angiography," Journal Articles Donald and Barbara Zucker School of Medicine Academic Works **35**(11), 2163–2180 (2015).
7. B. E. Bouma, G. J. Tearney, J. F. de Boer, and S. H. Yun, "Motion artifacts in optical coherence tomography with frequency-domain ranging," Opt. Express **12**(13), 2977–2998 (2004).
8. N. Venkateswaran, A. Galor, J. Wang, and C. L. Karp, "Optical coherence tomography for ocular surface and corneal diseases: a review," Eye and Vis. **5**(1), 13 (2018).
9. R. K. Murthy, S. Haji, K. Sambhav, S. Grover, and K. V. Chalam, "Clinical applications of spectral domain optical coherence tomography in retinal diseases," Biomedical Journal **39**(2), 107–120 (2016).
10. B. Qiu, Z. Huang, X. Liu, X. Meng, Y. You, G. Liu, K. Yang, A. Maier, Q. Ren, and Y. Lu, "Noise reduction in optical coherence tomography images using a deep neural network with perceptually-sensitive loss function," Biomed. Opt. Express **11**(2), 817 (2020).
11. Z. Dong, G. Liu, G. Ni, J. Jerwick, L. Duan, and C. Zhou, "Optical coherence tomography image denoising using a generative adversarial network with speckle modulation," J. Biophotonics **13**(4), e201960135 (2020).
12. W. Wu, O. Tan, R. R. Pappuru, H. Duan, and D. Huang, "Assessment of frame-averaging algorithms in OCT image analysis," Ophthalmic Surg. Lasers Imaging Retina **44**(2), 168–175 (2013).
13. A. Baghaie, Z. Yu, and R. M. D'Souza, "State-of-the-art in retinal optical coherence tomography image analysis," Quantitative Imaging in Medicine and Surgery **5**(4), 603 (2015).
14. L. Fang, S. Li, Q. Nie, J. A. Izatt, C. A. Toth, and S. Farsiu, "Sparsity based denoising of spectral domain optical coherence tomography images," Biomed. Opt. Express **3**(5), 927 (2012).
15. M. Esmaeili, A. M. Dehnavi, H. Rabbani, and F. Hajizadeh, "Speckle noise reduction in optical coherence tomography using two-dimensional curvelet-based dictionary learning," J. Med. Signals Sens **7**(2), 86–91 (2017).
16. J. J. Gomez-Valverde, J. E. Ortuno, P. Guerra, B. Hermann, B. Zabihian, J. L. Rubio-Guivernau, A. Santos, W. Drexler, and M. J. Ledesma-Carbayo, "Evaluation of speckle reduction with denoising filtering in optical coherence tomography for dermatology," in *Proceedings - International Symposium on Biomedical Imaging* (IEEE Computer Society, 2015), 2015-July, pp. 494–497.
17. Z. Amini and H. Rabbani, "Optical coherence tomography image denoising using Gaussianization transform," J. Biomed. Opt. **22**(08), 1 (2017).
18. C. Tang, L. Cao, J. Chen, and X. Zheng, "Speckle noise reduction for optical coherence tomography images via non-local weighted group low-rank representation," Laser Phys. Lett. **14**(5), 056002 (2017).
19. I. Rizwan, I. Haque, and J. Neubert, "Deep learning approaches to biomedical image segmentation," Informatics in Medicine Unlocked **18**, 100297 (2020).
20. J. Schaefferkoetter, J. Yan, C. Ortega, A. Sertic, E. Lechtman, Y. Eshet, U. Metser, and P. Veit-Haibach, "Convolutional neural networks for improving image quality with noisy PET data," EJNMMI Res **10**(1), 105 (2020).
21. A. S. Lundervold and A. Lundervold, "An overview of deep learning in medical imaging focusing on MRI," Zeitschrift fur Medizinische Physik **29**(2), 102–127 (2019).
22. R. T. Yanagihara, C. S. Lee, D. S. W. Ting, and A. Y. Lee, "Methodological challenges of deep learning in optical coherence tomography for retinal diseases: a review," Trans. Vis. Sci. Tech. **9**(2), 11 (2020).
23. K. Alsaih, M. Z. Yusoff, T. B. Tang, I. Faye, and F. Mériaudeau, "Deep learning architectures analysis for age-related macular degeneration segmentation on optical coherence tomography scans," Computer Methods and Programs in Biomedicine **195**, 105566 (2020).
24. M. Wang, W. Zhu, K. Yu, Z. Chen, F. Shi, Y. Zhou, Y. Ma, Y. Peng, D. Bao, S. Feng, L. Ye, D. Xiang, X. Chen, and S. Member, "Semi-supervised capsule cGAN for speckle noise reduction in retinal OCT images," IEEE Trans. Med. Imaging **40**(4), 1168–1183 (2021).
25. Y. Ma, X. Chen, W. Zhu, X. Cheng, D. Xiang, and F. Shi, "Speckle noise reduction in optical coherence tomography images based on edge-sensitive cGAN," Biomed. Opt. Express **9**(11), 5129 (2018).
26. A. Guo, L. Fang, M. Qi, and S. Li, "Unsupervised denoising of optical coherence tomography images with nonlocal-generative adversarial network," IEEE Trans. Instrum. Meas. **70**, 1–12 (2021).

27. N. A. Kande, R. Dakhane, A. Dukkipati, P. Kumar Yalavarthy, and S. Member, "SiameseGAN: a generative model for denoising of spectral domain optical coherence tomography images," IEEE Trans. Med. Imaging **40**(1), 180–192 (2021).

28. K. J. Halupka, B. J. Antony, M. H. Lee, K. A. Lucy, R. S. Rai, H. Ishikawa, G. Wollstein, J. S. Schuman, and R. Garnavi, "Retinal optical coherence tomography image enhancement via deep learning," Biomed. Opt. Express **9**(12), 6205–6221 (2018).

29. A. Abbasi, A. Monadjemi, L. Fang, H. Rabbani, and Y. Zhang, "Three-dimensional optical coherence tomography image denoising through multi-input fully-convolutional networks," Comput. Biol. Med. **108**, 1–8 (2019).

30. F. Shi, N. Cai, Y. Gu, D. Hu, Y. Ma, Y. Chen, and X. Chen, "DeSpecNet: a CNN-based method for speckle reduction in retinal optical coherence tomography images," Phys. Med. Biol. **64**(17), 175010 (2019).

31. N. Gour and P. Khanna, "Speckle denoising in optical coherence tomography images using residual deep convolutional neural network," Multimed Tools Appl **79**(21-22), 15679–15695 (2020).

32. G. Gisbert, N. Dey, H. Ishikawa, J. Schuman, J. Fishbaugh, and G. Gerig, "Self-supervised denoising via diffeomorphic template estimation: application to optical coherence tomography," Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) 12069 LNCS, 72–82 (2020).

33. S. N. Menon, V. B. Vineeth Reddy, A. Yeshwanth, B. N. Anoop, and J. Rajan, "A novel deep learning approach for the removal of speckle noise from optical coherence tomography images using gated convolution–deconvolution structure," in *Advances in Intelligent Systems and Computing* (Springer, 2020), 1024, pp. 115–126.

34. Z. Mao, A. Miki, S. Mei, Y. Dong, K. Maruyama, R. Kawasaki, S. Usui, K. Matsushita, K. Nishida, and K. Chan, "Deep learning based noise reduction method for automatic 3D segmentation of the anterior of lamina cribrosa in optical coherence tomography volumetric scans," Biomed. Opt. Express **10**(11), 5832 (2019).

35. S. K. Devalla, G. Subramanian, T. H. Pham, X. Wang, S. Perera, T. A. Tun, T. Aung, L. Schmetterer, A. H. Thiéry, and M. J. A. Girard, "A Deep Learning Approach to Denoise Optical Coherence Tomography Images of the Optic Nerve Head," Sci. Rep. **9**(1), 14454–13 (2019).

36. S. K. Devalla, P. K. Renukanand, B. K. Sreedhar, S. Perera, J.-M. Mari, K. S. Chin, T. A. Tun, N. G. Strouthidis, T. Aung, A. H. Thiéry, and M. J. A. Girard, "DRUNET: a dilated-residual u-net deep learning network to digitally stain optic nerve head tissues in optical coherence tomography images," Biomed. Opt. Express **9**, 3244–3265 (2018).

37. S. Apostolopoulos, J. Salas, J. L. P. Ordóñez, S. S. Tan, C. Ciller, A. Ebneter, M. Zinkernagel, R. Sznitman, S. Wolf, S. de Zanet, and M. R. Munk, "Automatically enhanced OCT scans of the retina: a proof of concept study," Sci. Rep. **10**(1), 7819 (2020).

38. D. Hu, C. Cui, H. Li, K. E. Larson, Y. K. Tao, and I. Oguz, "LIFE: a generalizable autodidactic pipeline for 3D OCT-A vessel segmentation," in Medical Image Computing and Computer Assisted Intervention – MICCAI 2021, Lecture Notes in Computer Science, vol 12901 (2021).

39. D. Hu, J. D. Malone, Y. Atay, Y. K. Tao, and I. Oguz, "Retinal OCT denoising with pseudo-multimodal fusion network," Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) 12069 LNCS, 125–135 (2020).

40. Y. Huang, N. Zhang, and Q. Hao, "Real-time noise reduction based on ground truth free deep learning for optical coherence tomography," Biomed. Opt. Express **12**(4), 2027 (2021).

41. I. Oguz, J. Malone, Y. Atay, and Y. K. Tao, "Self-fusion for OCT noise reduction," in *Medical Imaging 2020: Image Processing*, B. A. Landman and I. Išgum, eds. (SPIE, 2020), 11313, p. 11.

42. H. Wang, J. W. Suh, S. R. Das, J. B. Pluta, C. Craige, and P. A. Yushkevich, "Multi-atlas segmentation with joint label fusion," IEEE Trans. Pattern Anal. Mach. Intell. **35**(3), 611–623 (2013).

43. J. Xu, K. Wong, Y. Jian, and M. V. Sarunic, "Real-time acquisition and display of flow contrast using speckle variance optical coherence tomography in a graphics processing unit," J. Biomed. Opt. **19**(02), 1 (2014).

44. A. Camino, D. Huang, J. C. Morrison, S. Pi, T. T. Hormel, W. Cepurna, X. Wei, and Y. Jia, "Real-time cross-sectional and en face OCT angiography guiding high-quality scan acquisition," Opt. Lett. **44**(6), 1431–1434 (2019).

45. M. J. Ringel, E. M. Tang, and Y. K. Tao, "Advances in multimodal imaging in ophthalmology," Ophthalmol. Eye Dis. **13**, 251584142110024 (2021).

46. M. T. El-Haddad and Y. K. Tao, "Advances in intraoperative optical coherence tomography for surgical guidance," Curr. Opin. Biomed. Eng. **3**, 37–48 (2017).

47. M. T. El-Haddad, I. Bozic, and Y. K. Tao, "Spectrally encoded coherence tomography and reflectometry: simultaneous en face and cross-sectional imaging at 2 gigapixels per second," J. Biophotonics **11**(4), e201700268 (2018).

48. J. D. Malone, M. T. El-Haddad, S. S. Yerramreddy, I. Oguz, and Y. K. K. Tao, "Handheld spectrally encoded coherence tomography and reflectometry for motion-corrected ophthalmic optical coherence tomography and optical coherence tomography angiography," Neurophotonics **6**(04), 1 (2019).

49. L. Fang, S. Li, R. P. McNabb, Q. Nie, A. N. Kuo, C. A. Toth, J. A. Izatt, and S. Farsiu, "Fast acquisition and reconstruction of optical coherence tomography images via sparse representation," IEEE Trans. Med. Imaging **32**(11), 2034–2049 (2013).

50. H. Bogunovic, F. Venhuizen, S. Klimscha, S. Apostolopoulos, A. Bab-Hadiashar, U. Bagci, M. F. Beg, L. Bekalo, Q. Chen, C. Ciller, K. Gopinath, A. K. Gostar, K. Jeon, Z. Ji, S. H. Kang, D. D. Koozekanani, D. Lu, D. Morley, K. K. Parhi, H. S. Park, A. Rashno, M. Sarunic, S. Shaikh, J. Sivaswamy, R. Tennakoon, S. Yadav, S. de Zanet, S. M. Waldstein, B. S. Gerendas, C. Klaver, C. I. Sanchez, and U. Schmidt-Erfurth, "RETOUCH: the retinal OCT fluid detection and segmentation benchmark and challenge," IEEE Trans. Med. Imaging **38**(8), 1858–1874 (2019).

51. M. Guizar-Sicairos, S. T. Thurman, and J. R. Fienup, "Efficient subpixel image registration algorithms," Opt. Lett. **33**(2), 156–158 (2008).

52. N. J. Tustison, Y. Yang, and M. Salerno, "Advanced normalization tools for cardiac motion correction," Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) **8896**, 3–12 (2014).

53. P. Hahn, J. Migacz, R. O'Connell, R. S. Maldonado, J. A. Izatt, and C. A. Toth, "The use of optical coherence tomography in intraoperative ophthalmic imaging," Ophthalmic Surg Lasers Imaging **42**(4), S85 (2011).

54. O. M. Carrasco-Zevallos, B. Keller, C. Viehland, L. Shen, G. Waterman, B. Todorich, C. Shieh, P. Hahn, S. Farsiu, A. N. Kuo, C. A. Toth, and J. A. Izatt, "Live volumetric (4D) visualization and guidance of in vivo human ophthalmic surgery with intraoperative optical coherence tomography," Sci. Rep. **6**(1), 31689–16 (2016).

55. J. P. Ehlers, "Intraoperative optical coherence tomography: past, present, and future," Eye **30**(2), 193–201 (2016).