End-to-end design for visOCT denoising by speckle reduction scanning and deep learning

Tianyi Ye[1], Jingyu Wang[2], Ji Yi[1,2,†]

1. Department of Biomedical Engineering, Johns Hopkins University, Baltimore USA
2. Department of Ophthalmology, Johns Hopkins University, Baltimore USA
[†]Correspondence: jiyi@jhu.edu

Abstract

Purpose:
Visible light optical coherence tomography (VIS-OCT) is an emerging imaging modality that provides one-micron level axial resolution. Improving VIS-OCT image quality by denoising is an essential step in the overall workflow in VIS-OCT clinical applications because its illumination power is limited for patient safety and comfort. In this study, we provide the first deep learning-based, end-to-end design for visOCT denoising using our speckle reduction scanning dataset.

Methods:
The HD dataset includes 105 retinal B-scans obtained on our 2nd Gen dual-channel VIS-OCT system for 12 subjects. With our speckle reduction scanning protocol, 16 or 32 B-scans at the same location are obtained and averaged to get the noisy-clean pairs. Using the U-Net architecture, we provide both supervised and self-supervised strategies using the noise-to-void method(N2V) to fulfill practical scenarios where the clean images are not available. We split the HD dataset into training, validation, and test sets as (51:17:37). The dataset is augmented 8-fold by three 90° rotations and horizontal flip. For the supervised method, the input images are resized to 512x512, and the batch size is 2. For the self-supervised method, the input images are resized to 512x512 followed by random cropped to 64×64 with a batch size of 128. They are trained for 200 epochs by Adam optimizer. We also test the models on our Raster scan dataset with a 3D volume of the human retina.

Results
Fig1. A-D visualized the denoising performance on HD dataset. We repeat the experiment 8 times and the average PSNR/SSIM for self-supervised and supervised strategies are 27.26(±0.2470)/0.5296(±0.0083) and 30.96((±0.1556)/0.7636((±0.0024), respectively, both of which significantly improved the PSNR/SSIM of noisy images from 22.90/0.2639.

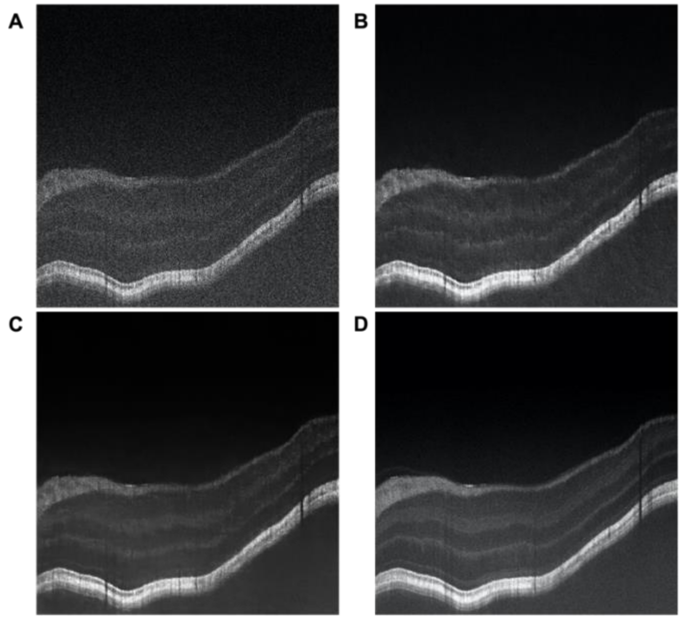

Fig .1. A) A noisy example from the HD testing dataset. B) The denoised result of the self-supervised strategy. C) The denoised result of the supervised strategy. D) The multi B-scans averaged clean image.

Fig. 2 A, B and C qualitatively show the models trained on HD dataset generalize well to Raster scan dataset and denoise the whole 3D volumes. Since we do not have clean ground truth for Raster scan dataset, the quantitative analysis is not available.

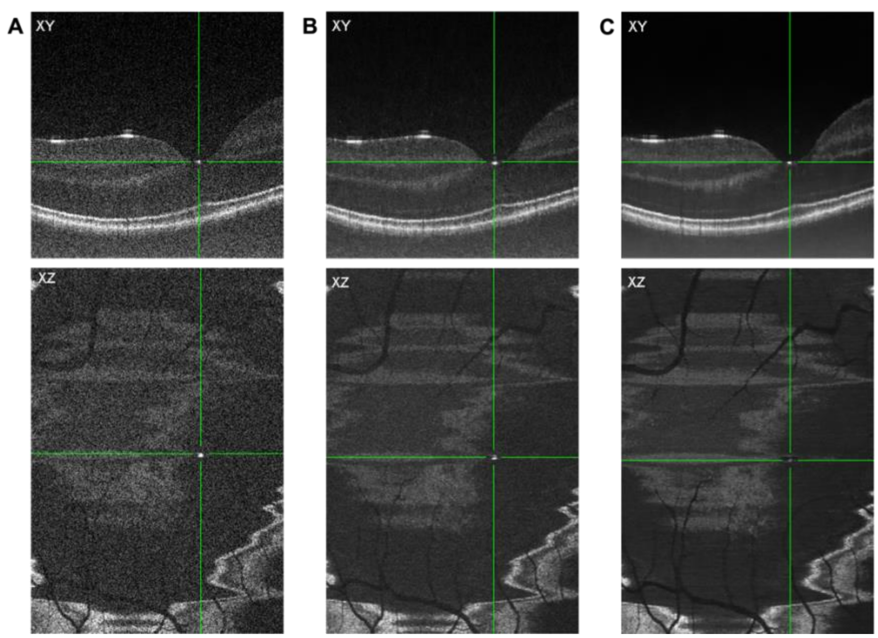

Fig. 2. Top: B-scan, bottom: en face image. Fig. 2A, B and C are the noisy volume, volume denoised by the self-supervised method and volume denoised by the supervised method.

Conclusions:

The proposed end-to-end framework that includes both self-supervised and supervised strategies successfully denoise the HD visOCT images with significantly improved PSNRs and SSIMs. The framework is also robust to denoise the never-seen Raster scan visOCT volume. This study reveals the strength, robustness and flexibility of deep learning-based denoising for visOCT images.