Self-supervised Denoising and Bulk Motion Artifact Removal of 3D Optical Coherence Tomography Angiography of Awake Brain

Zhenghong Li, Jiaxiang Ren, Zhilin Zou, Kalyan Garigapati, Congwu Du, Yingtian Pan, and Haibin Ling

Stony Brook University, USA
{zhenghli,hling}@cs.stonybrook.edu

Abstract. Denoising of 3D Optical Coherence Tomography Angiography (OCTA) for awake brain microvasculature is challenging. An OCTA volume is scanned slice by slice, with each slice (named B-scan) derived from dynamic changes in successively acquired OCT images. A B-scan of an awake brain often suffers from complex noise and Bulk Motion Artifacts (BMA), severely degrading image quality. Also, acquiring clean B-scans for training is difficult. Fortunately, we observe that, the slice-wise imaging procedure makes the noises mostly independent across B-scans, while preserving the continuity of vessel (including capillaries) signals across B-scans. Thus inspired, we propose a novel blind-slice self-supervised learning method to denoise 3D brain OCTA volumes slice by slice. For each B-scan slice, named center B-scan, we mask it entirely black and train the network to recover the original center B-scan using its neighboring B-scans. To enhance the BMA removal performance, we adaptively select only BMA-free center B-scans for model training. We further propose two novel refinement methods: (1) a non-local block to enhance vessel continuity and (2) a weighted loss to improve vascular contrast. To the best of our knowledge, this is the first self-supervised 3D OCTA denoising method that effectively reduces both complex noise and BMA while preserving capillary signals in brain OCTA volumes. Code is available at https://github.com/ZhenghLi/SOAD.

Keywords: Optical Coherence Tomography Angiography · Self-supervised Volume Denoising · Bulk Motion Artifacts Removal

1 Introduction

Background. Optical Coherence Tomography (OCT) is a rapidly developing tomography imaging technique for biological diagnoses [27]. OCT Angiography (OCTA) volume is scanned slice by slice, with each slice (B-scan) derived from dynamic changes in successively acquired cross-sectional OCT images at the same slow-axis position [9]. Despite an increasing amount of research carried out in retinal OCT/OCTA [16,8,28,12,20], less focus has been directed towards 3D OCTA of brain microvasculature, especially in an awake brain. Due to challenging imaging conditions such as heterogeneous tissue and deeper penetration and

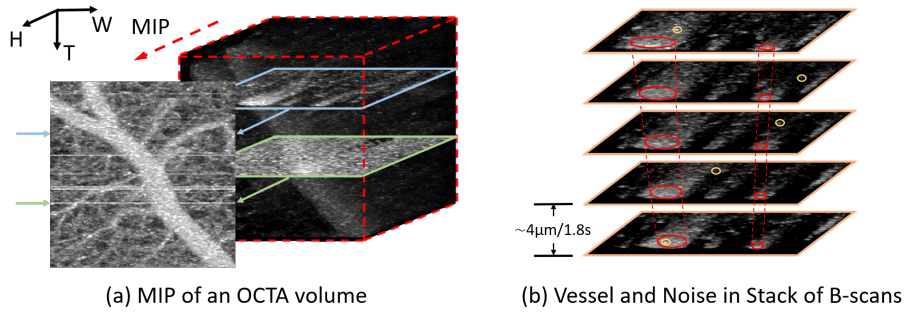

Fig. 1. (a) Maximum Intensity Projection (MIP) of a raw OCTA volume of the brain cortex of an awake mouse. The blue slice is a *normal* B-scan consisting of unknown true signal, shadow, speckle noise, and background noise. The green slice is a *corrupted* B-scan affected by severe bulk motion artifacts, and it is projected as a bright stripe in the MIP image. **(b)** The slice-wise imaging process causes a relatively large spatial and time gap between OCTA B-scans, making the noise independent and discontinuous across B-scans (yellow circles), while vessel signals remain continuous (red tubes).

possible motions of subjects, an OCTA B-scan of awake brains often suffers from noise and motion artifacts, severely degrading the imaging quality. In this work, we employ Maximum Intensity Projection (MIP), a common processing against shadow effect, to visualize vessels (Fig. 1(a)). Our goal is to enhance MIP images by denoising and removing bulk motion artifacts (BMA) from the volumes.

Besides the lack of groundtruth, denoising OCTA images of the brain cortex presents two main challenges. **(1)** Complex noise, including speckle noise due to light scattering and motion artifacts caused by the respiration or movement of awake subjects, can be non-zero mean and correlated within a B-scan. In the paper, we denote background noise as the noise caused by micro motions or background tissue scattering (Fig. 1(a) blue) and motion artifacts as the severe BMA (Fig. 1(a) green) caused by the movement of awake subjects. Also, we call the B-scans unaffected by BMA as normal B-scans and those affected by BMA as corrupted ones. **(2)** Capillaries often have weak signals and consist of only a few pixels in an OCTA B-scan due to their small size and slow blood flow. Hence, capillary signals are likely mistaken for noise and removed during denoising.

Unfortunately, most existing methods cannot effectively solve the above challenges. Traditional OCT/OCTA denoising methods [3,17,6,28,32] are tailored to retinal images/volumes. They may struggle with the noise characteristics of OCTA of awake brains and be confused by capillary signals and noise. Learning-based methods [24,21,18,10] often require dense scanning and registration for generating pseudo-ground truth B-scans for supervision, yet OCTA images of awake brains remain noisy even after this processing. Unsupervised OCT denoising [8] or BMA removal methods [25] may require extra regions of interest (ROI) labeling or prior knowledge for BMA synthesis on MIP images, making them less practical for brain OCTA volumes. Generic self-supervised methods

based on blind-spot networks [13,2,14,29,31,34] have shown promise in biomedical domain [33], but they assume noise to be 2D independently distributed with zero means, which is not the case in OCTA volumes of awake brains.

Motivation. Although noise is corrected within a B-scan, due to the slice-wise imaging property of OCTA volumes, it shows inter-B-scan discontinuity and approximate independence, while vessel signals are still continuous scross B-scans (Fig. 1 (b)). This property suggests us to model the inter-B-scan 3D continuity to separate noise and signal.

Contribution. Inspired by above observation, we propose an effective *Self-supervised 3D OCTA Denoising* method (SOAD) based on a blind-slice masking strategy to simultaneously remove speckle noise, background noise, and BMA. Specifically, SOAD denoises an OCTA volume slice by slice. For each B-scan slice, we sample its neighboring B-scans to form a sub-volume as the input. The center B-scan of the sub-volume is the slice to be denoised. We mask it entirely black and train a network to recover the original center B-scan. Our blind-slice strategy employs the difference in the inter-B-scan continuity to adaptively separate vessel signal and noise, so it can effectively reduce complex noise while preserving capillaries. To further enhance the BMA removal performance, we propose an empirical criterion that utilizes the intensity-increasing effect of BMA to effectively exclude BMA B-scans from contaminating the training process. In this way, the network will only learn to restore normal B-scans and thus remove BMA even for several consecutive corrupted B-scans. It is also noteworthy that our method does not need additional data synthesis and ROI labeling.

Besides, since most vessel signals share similar representation in a volume, inspired by the non-local mean technique [4] for image denoising, we further propose use a non-local block [30] to implicitly realize the non-local mean operation on deep features. Moreover, we introduce a novel weighted loss function that adaptively assigns higher weights to vessel regions to improve the contrast. In summary, our main contributions include: **(1)** the first 3D self-supervised learning framework for OCTA volume denoising, **(2)** a novel blind-slice masking strategy that forces the model to learn 3D features to remove noise and BMA while preserving capillary signals, **(3)** a non-local block that leverages the similarity shared by vessel signals to enhance the vascular features and improve the denoising results, and **(4)** a novel weighted loss function that adaptively assigns higher weights to vessel regions to enhance the results. With these contributions, our method show promising results in the evaluation on real world data in comparison with state-of-the-arts.

2 Self-supervised OCTA Denoising (SOAD)

2.1 Noise and BMA Model

A 3D OCTA volume of awake brains of N B-scans can be represented by $\boldsymbol{Y} = \{\boldsymbol{y}_1, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_N\}$, where a 2D B-scan $\boldsymbol{y}_i$ can be approximately modeled as:

$$\boldsymbol{y}_i \approx f(s(\boldsymbol{x}_i), \boldsymbol{\nu}_i) + \boldsymbol{b}_i + \boldsymbol{\mu}_i \tag{1}$$

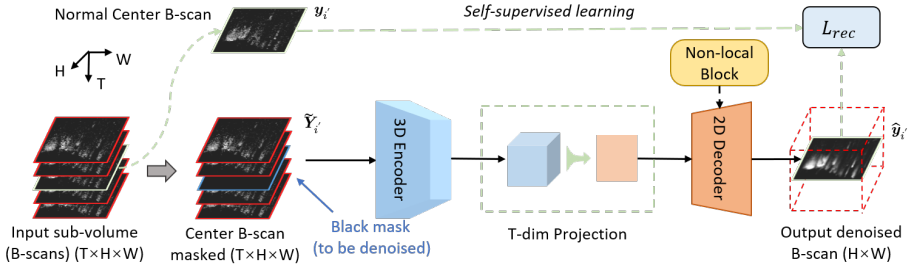

Fig. 2. Illustration of our Self-supervised OCTA Denoising (SOAD). Five input B-scans are shown for illustration (seven are actually used), and the center B-scan is masked and to be denoised. The 3D encoder is modified from the V-Net [22] encoder. The encoded multi-scale 3D feature maps are projected into 2D before inputting to the decoder. The projection is realized by 3D Convs whose T-dimensions are the same as the encoded features. Then a 2D decoder, modified from U-Net [26], restores the center B-scan from the projected features. A non-local block is inserted after the deepest decoder block to refine the feature. The loss function is the weighted recovery loss defined in the text.

where $\boldsymbol{x}$ is the groundtruth signal unavailable through current OCTA systems, $s(\cdot)$ is the shadow function caused by light scattering in the blood, $\boldsymbol{\nu}$ is the foreground speckle noise component caused mainly by scattering of red blood cells and tissues, $f(\cdot)$ is the foreground imaging function, $\boldsymbol{b}$ is the background noise caused by micro motions or tissue scattering, and $\boldsymbol{\mu}$ is the BMA. We model $\boldsymbol{b}$ as additive noise following [24] and $\boldsymbol{\mu}$ also as an additive component due to its intensity-increasing effect.

For OCTA volumes of awake brains, a key observation is the rough independence and discontinuity of the noise across B-scans. Due to the slice-wise OCTA imaging process, the $\boldsymbol{\nu}_i$ and $\boldsymbol{b}_i$ are independent and discontinuous w.r.t $\boldsymbol{\nu}_{j|j\neq i}$ and $\boldsymbol{b}_{j|j\neq i}$. Therefore, instead of exactly mathematically model the noise in 2D, which is still an open question, we propose to utilize the difference in inter-B-scan 3D continuity between noise and vessel signal to separate them.

For motion artifact, since BMA $\boldsymbol{\mu}_i$ appears only in a few $\boldsymbol{y}_i$ and drastically increases the overall intensity of $\boldsymbol{y}_i$, we can easily detect such B-scans and then prevent them from participating the self-supervised learning. Thus, although $\boldsymbol{\mu}$ may emerge continuously, such corruption only spans up to only a few B-scans due to the time gap between B-scans. Intuitively, such BMA effects can be eliminated by exploring neighbor B-scans as will be described later.

Note that our final purpose is to suppress the noise of the MIP visualization of OCTA volumes. For this purpose, we do not need to specifically deal with $\boldsymbol{s}$, since it will be suppressed by $\boldsymbol{x}$ during MIP, i.e., $\text{MIP}(s(\boldsymbol{x})) \approx \text{MIP}(\boldsymbol{x})$.

2.2 Blind-slice Self-supervised OCTA Denoising and BMA Removal

The pipeline of our blind-slice self-supervised method SOAD is shown in Fig. 2. The OCTA volume is denoised slice by slice. For each B-scan $\boldsymbol{y}_i$, i.e., the center B-scan of the input sub-volume composed of $\boldsymbol{y}_i$ and its neighboring B-scans, the proposed system can be formulated as $\hat{\boldsymbol{y}}_i = F_\theta(\tilde{\boldsymbol{Y}}_i)$, where $F_\theta(\cdot)$ denotes the network with parameters in θ, $\tilde{\boldsymbol{Y}}_i = \{\boldsymbol{y}_{i-t}, \ldots, \boldsymbol{y}_{i-1}, \mathcal{M}(\boldsymbol{y}_i), \boldsymbol{y}_{i+1}, \ldots, \boldsymbol{y}_{i+t}\}$, and $T = 2t + 1$ is the window size. $\mathcal{M}(\boldsymbol{y}_i)$ denotes the center B-scan $\boldsymbol{y}_i$ of the input that is masked black, meaning that $\tilde{\boldsymbol{Y}}_i$ contains no direct information of $\boldsymbol{y}_i$.

Because using corrupted B-scans for supervision will introduce bias, to enhance BMA removal, we aim to only use normal B-scans $\boldsymbol{y}_{i'}$ for supervision, i.e., $\boldsymbol{y}_{i'} = \nu(\boldsymbol{x}_{i'} + \boldsymbol{s}_{i'}) + \boldsymbol{b}_{i'}$, where i' denotes the index of normal B-scans and $\tilde{\boldsymbol{Y}}_{i'}$ denotes the input for training whose center B-scan is $\boldsymbol{y}_{i'}$. Note that $\boldsymbol{\mu}_i$ appears only in several $\boldsymbol{y}_i$ and it can drastically increase the mean intensity of $\boldsymbol{y}_i$ (denoted as $\bar{y}_i$). Thus, we exclude the sub-volumes whose center B-scans' mean intensities are significantly above the average over the B-scan sequences. To avoid manual annotations for training, we empirically treat $\bar{y}_i < \text{mean}(\bar{\boldsymbol{y}}) + 1.5 * \text{std}(\bar{\boldsymbol{y}})$ as normal B-scans $\boldsymbol{y}_{i'}$, where $\bar{\boldsymbol{y}}$ is a sequence of all $\bar{y}_i$ in a volume.

Since background noise $\boldsymbol{b}_{i'}$ is independent of $\boldsymbol{y}_{j|j\neq i'}$, $\boldsymbol{b}_{i'}$ is independent of $\tilde{\boldsymbol{Y}}_{i'}$. Because $\boldsymbol{b}_{i'}$ is sparse and discontinuous across different B-scans, we assume that $\hat{\boldsymbol{y}}_{i'}$ cannot fit $\boldsymbol{b}_{i'}$ given the input $\tilde{\boldsymbol{Y}}_{i'}$, and $\hat{\boldsymbol{y}}_{i'}$ will converge to 0 in the background regions, which is a local minimum of the squared ℓ_2 loss function. Similarly, for the foreground speckle noise $\boldsymbol{\nu}_{i'}$, we also assume that $\hat{\boldsymbol{y}}_{i'}$ cannot fit $\boldsymbol{\nu}_{i'}$ but learns to smooth the foreground to approach a local minimum of the squared ℓ_2 loss function. Since vessels are continuous, $\hat{\boldsymbol{y}}_{i'}$ will finally only converge to $s(\boldsymbol{x}_{i'})$.

2.3 Refinement

Non-local Block. Vessel signals in OCTA volumes share similar appearances and continuities, indicating that applying non-local mean [4] could refine denoising. However, directly applying this as a post-processing step is computationally expensive. Instead, we use a non-local block (NL) [30] in the decoder. NL is originally inspired by non-local mean [4] and can adaptively assign higher weights to similar features to enhance vessel features and improve the results.

Weighted Loss Function. Because the vessels are sparse in the volumes, if we directly use the common squared ℓ_2 loss as the recovery loss, the outputs of the network will be dim and low in contrast. To improve the contrast of the output volume, we propose a weighted loss (WL) function:

$$\mathcal{L}_{\text{rec}} = \frac{1}{N'} \sum_{i'} \|\boldsymbol{w}_{i'} * (\hat{\boldsymbol{y}}_{i'} - \boldsymbol{y}_{i'})\|_2^2 \tag{2}$$

where N' is the number of normal B-scans $\boldsymbol{y}_{i'}$ in a batch, $\boldsymbol{w}_{i'} = \alpha \cdot \text{sg}(\hat{\boldsymbol{y}}_{i'}^{\gamma}) + \boldsymbol{y}_{i'}^{\gamma} + 0.5$, α is the weight parameter, γ the curve parameter, and $\text{sg}(\cdot)$ means stop-gradient. The bias term 0.5 is to prevent the background weights from zero. This WL function adaptively assigns higher weights to vessel regions that are usually bright in denoised B-scans $\hat{\boldsymbol{y}}_{i'}$. We include raw B-scans $\boldsymbol{y}_{i'}$ in $\boldsymbol{w}_{i'}$ to help the convergence because $\hat{\boldsymbol{y}}_{i'}$ is not stable at the beginning of the training.

3 Experiment

3.1 Experiment Details

OCTA Dataset of Awake Mouse Cortex. Seven OCTA volumes in the brain cortex of awake mice are collected. Five of them are originally of $500 \times H \times 1000$ ($T \times H \times W$) voxels, and two are of $625 \times H \times 1000$, where H ranges from 400 – 512. The physical sizes of the cropped volumes are around $2 \times 0.7 \times 2.4mm^3$ for $T = 500$ B-scans or $2.5 \times 0.7 \times 2.4mm^3$ for $T = 625$ B-scans. About 3.6% B-scans in the dataset are corrupted B-scans.

Implementation Details. Our encoder and decoder follow that in [1]. We downsample the T-dimension of the features at the 1st and 3rd stages. The input sub-volume is of size $7 \times 128 \times 128$, and we sample the original volume by a stride of $1 \times 64 \times 64$. The model is trained by Adam [11] for 15 epochs with 128 batch size. The initial learning rate is $1e-3$, and decays by 0.1 at the 4th, 8th, and 12th epochs. For WL, we set $\alpha = 100$ and $\gamma = 1/3$. We first train and evaluate SOAD based on leave-one-out validation and denote it as SOAD (L). Then, inspired by [23,15] that self-supervised denoising method can also be trained on the validation or test dataset to get better result, we also directly train and evaluate SOAD on all seven volumes denoted as SOAD (S).

We compare SOAD with two classical denoisers (BM3D [5] and BM4D [19]), three OCT/OCTA volume denoisers (Median-TVD [28], Shearlet3D [32] and Pipeline [10]), and two 3D self-supervised methods (Noise2Self [2] and UDVD [29]). The deep learning methods [10,2,29] are evaluated in the leave-one-out protocol.

Evaluation Metrics We separately evaluate the results on normal and corrupted B-scans. Since there is no groundtruth, we employ two no-reference ROI-based metrics [7,6]. The first one, contrast-to-noise ratio (CNR), evaluates the contrast between foreground ROIs and background. The second one, mean-to-standard-deviation ratio (MSR), evaluates the homogeneity of foreground ROIs. Specifically, we randomly sample 30 normal B-scans from each volume and label 9 ROIs for each normal B-scan. For corrupted B-scans, we sample 8 corrupted B-scans from each volume and label 6 ROIs for each. We select 5 large noisy background regions for each B-scan and concatenate them as a whole background.

Since ROI-based metrics can be biased in small regions, we further use vessel segmentation on MIP images for evaluation. Specifically, we use the MIP image of each result volume to train a standard U-Net [26], supervised by manually labeled vessel masks. We use leave-one-out validation and Dice scores to evaluate the result. To further evaluate results on capillaries, we pick the masks of small vessels (capillaries and some small arterioles and venules) and report additional Dice scores over them, denoted as Dice_s.

3.2 Experimental Results

Quantitative Results. The **normal** B-scan denoising results are presented in Table 1 (left). Our method SOAD (L) clearly outperforms others, and can be further improved via self-supervised training on the validation set, *i.e.*, SOAD

Table 1. Comparison of denoising and BMA removal results

	Normal B-scans				Corrupted B-scans			
	CNR	MSR	Dice	Dice_s	CNR	MSR	Dice	Dice_s
Raw	3.76	3.32	76.96	45.81	2.39	3.60	70.56	36.15
BM3D [5]	12.81	10.87	76.74	41.22	7.05	14.51	72.36	33.89
BM4D [19]	12.02	9.99	82.19	54.60	5.63	11.92	76.21	42.58
Median-TVD [28]	12.62	10.54	74.30	34.35	8.21	14.17	71.10	29.96
Shearlet3D [32]	9.59	8.62	77.63	44.46	8.16	9.75	75.21	39.20
Pipeline [10]	6.41	6.27	77.14	45.63	5.73	5.97	76.74	38.82
Noise2Self [2]	4.12	3.60	77.67	47.08	2.61	3.91	71.30	37.17
UDVD [29]	4.02	3.52	77.54	46.60	2.54	3.82	71.66	37.37
SOAD (L)	14.01	14.69	87.71	69.69	14.50	16.84	86.93	64.65
SOAD (S)	**14.50**	**15.28**	**88.18**	**70.52**	**14.92**	**17.63**	**87.64**	**65.91**

Table 2. Ablation study on the Non-local Block and the Weighted Loss

		Normal B-scans				Corrupted B-scans			
NL	WL	CNR	MSR	Dice	Dice_s	CNR	MSR	Dice	Dice_s
		13.37	12.80	87.42	68.96	14.37	15.01	87.38	64.67
✓		13.96	14.11	87.95	70.06	14.45	16.32	87.22	65.39
	✓	14.21	14.11	87.71	69.67	**14.92**	16.03	87.48	65.25
✓	✓	**14.50**	**15.28**	**88.18**	**70.52**	**14.92**	**17.63**	**87.64**	**65.91**

(S). SOAD produces not only good ROI-background contrast but also good capillary signal preservation. Among other methods, only BM4D can have both effective denoising and segmentation results. Note that the blind-spot strategy-based denoising methods, Noise2Self and UDVD, do not work well as they may restore the noise using neighboring noisy pixels in the same B-scan.

The **corrupted** B-scan denoising and BMA removal results are presented in Table 1 (right). Again, SOAD consistently outperforms other methods, which shows the effectiveness of SOAD on both denoising and BMA removal.

Qualitative Results. Fig. 3 compares the MIP images of different methods. For **normal** B-scan denoising, the raw image is noisy with low contrast, making vessels unclear or even invisible in the MIP image. Results show that our SOAD can effectively remove the noise, with large vessels smoothed and capillaries recovered. Other methods, e.g., BM3D and Median-TVD, remove background noise but deteriorate capillary signals as well. BM4D preserves some capillary signals, but suffers from relatively low smoothness and continuity.

For **corrupted** B-scan denoising, our SOAD removes most BMA and eliminates stripes in the resultant MIP images. Shearlet3D also removes BMA, but its result is low in contrast and smoothness. Pipeline does not effectively recover the information in the corrupted B-scans because there can be big misalignment across the corresponding raw OCT B-scans. The other methods do not work for BMA removal because BMA can be densely spread in corrupted OCTA B-scans.

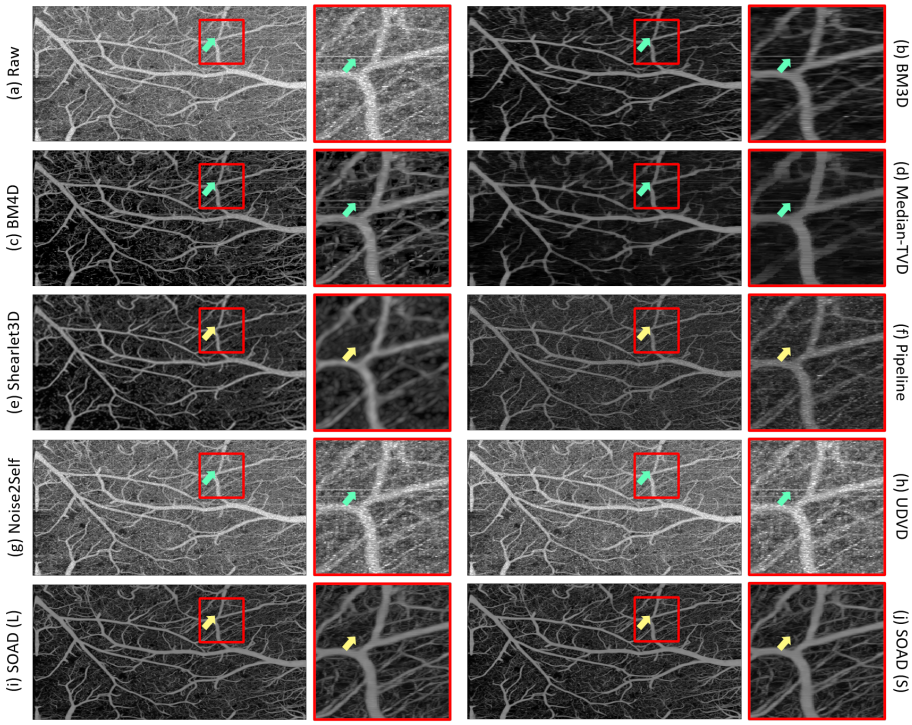

Fig. 3. Comparison of MIP results. Regions of interest are zoomed in. Stripes are marked by green arrows, and removed stripes are marked by yellow arrows.

Ablation Study. Table 2 compares SOAD with different refinement options. It shows that both refinement methods effectively improve the performance. Specifically, NL mainly improves the vessel structures by aggregating more non-local information to enhance vessel features, and WL mainly improves the contrast by making the model pay more attention to the foreground in the supervision. It is worth mentioning that, in rare cases, there can be consecutive corrupted B-scans. NL sometimes aggregates the neighbor BMA information for the inference, and the BMA in the center B-scan may not be perfectly removed. Such situation can be improved when incorporating WL in training.

Besides the refinement methods, we also present ablation study of normal B-scan selection threshold in the supplementary. The result shows that without normal B-scan selection, the overall performance on corrupted B-scans will drop.

4 Conclusion

We propose an effective blind-slice self-supervised learning method, named SOAD, for denoising and BMA removal of OCTA volumes of awake brains. SOAD denoises an OCTA volume slice by slice, by masking the center B-scans of input

sub-volumes and training the network to recover them. The network uses a 3D-to-2D encoder-decoder architecture to aggregate 3D information to recover the center B-scan. A non-local block and a weighted loss function are further designed to enhance the performance. Experiments on an awake brain OCTA dataset show that SOAD effectively removes speckle noise, background noise, and bulk motion artifacts while restoring capillaries. Despite the good performance, because noise distribution in awake brain OCTA is very complex without a closed-form mathematical model, this work does not present rigorous proof of the self-supervised learning procedure. Future work will be carried out on this.

Acknowledgments. This work was partially supported by NIH grants 1R21DA057699 (HL/YP/CD), 1RF1DA048808 (YP/CD) and 2R01DA029718 (CD/YP), and partially supported by NSF grants 2006665 (HL) and 2128350 (HL).

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. https://github.com/Hsuxu/Magic-VNet
2. Batson, J., Royer, L.: Noise2self: Blind denoising by self-supervision. In: International Conference on Machine Learning. pp. 524–533. PMLR (2019)
3. 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**(3), 036006 (2015)
4. Buades, A., Coll, B., Morel, J.M.: A non-local algorithm for image denoising. In: 2005 IEEE computer society conference on computer vision and pattern recognition (CVPR'05). vol. 2, pp. 60–65. Ieee (2005)
5. Dabov, K., Foi, A., Katkovnik, V., Egiazarian, K.: Image denoising by sparse 3-d transform-domain collaborative filtering. TIP **16**(8), 2080–2095 (2007)
6. Daneshmand, P.G., Mehridehnavi, A., Rabbani, H.: Reconstruction of optical coherence tomography images using mixed low rank approximation and second order tensor based total variation method. TMI **40**(3), 865–878 (2020)
7. Fang, L., Li, S., Nie, Q., Izatt, J.A., Toth, C.A., Farsiu, S.: Sparsity based denoising of spectral domain optical coherence tomography images. Biomedical optics express **3**(5), 927–942 (2012)
8. Guo, A., Fang, L., Qi, M., Li, S.: Unsupervised denoising of optical coherence tomography images with nonlocal-generative adversarial network. IEEE Transactions on Instrumentation and Measurement (2020)
9. Hossbach, J., Husvogt, L., Kraus, M.F., Fujimoto, J.G., Maier, A.K.: Deep oct angiography image generation for motion artifact suppression. In: Bildverarbeitung für die Medizin 2020, pp. 248–253. Springer (2020)
10. Jiang, Z., Huang, Z., Qiu, B., Meng, X., You, Y., Liu, X., Liu, G., Zhou, C., Yang, K., Maier, A., et al.: Comparative study of deep learning models for optical coherence tomography angiography. Biomedical optics express **11**(3), 1580–1597 (2020)
11. Kingma, D.P., Ba, J.: Adam: A method for stochastic optimization. In: ICLR (Poster) (2015)

12. Koch, V., Holmberg, O., Spitzer, H., Schiefelbein, J., Asani, B., Hafner, M., Theis, F.J.: Noise transfer for unsupervised domain adaptation of retinal oct images. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 699–708. Springer (2022)
13. Krull, A., Buchholz, T.O., Jug, F.: Noise2void-learning denoising from single noisy images. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp. 2129–2137 (2019)
14. Laine, S., Karras, T., Lehtinen, J., Aila, T.: High-quality self-supervised deep image denoising. Advances in Neural Information Processing Systems **32** (2019)
15. Lee, W., Son, S., Lee, K.M.: Ap-bsn: Self-supervised denoising for real-world images via asymmetric pd and blind-spot network. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 17725–17734 (2022)
16. Li, M., Zhang, Y., Ji, Z., Xie, K., Yuan, S., Liu, Q., Chen, Q.: Ipn-v2 and octa-500: Methodology and dataset for retinal image segmentation. arXiv preprint arXiv:2012.07261 (2020)
17. Li, M., Idoughi, R., Choudhury, B., Heidrich, W.: Statistical model for oct image denoising. Biomedical optics express **8**(9), 3903–3917 (2017)
18. Liu, X., Huang, Z., Wang, Z., Wen, C., Jiang, Z., Yu, Z., Liu, J., Liu, G., Huang, X., Maier, A., et al.: A deep learning based pipeline for optical coherence tomography angiography. Journal of Biophotonics **12**(10), e201900008 (2019)
19. Maggioni, M., Katkovnik, V., Egiazarian, K., Foi, A.: Nonlocal transform-domain filter for volumetric data denoising and reconstruction. TIP **22**(1), 119–133 (2012)
20. Mahapatra, D., Bozorgtabar, B., Shao, L.: Pathological retinal region segmentation from oct images using geometric relation based augmentation. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 9611–9620 (2020)
21. Mehdizadeh, M., MacNish, C., Xiao, D., Alonso-Caneiro, D., Kugelman, J., Bennamoun, M.: Deep feature loss to denoise oct images using deep neural networks. Journal of Biomedical Optics **26**(4), 046003 (2021)
22. Milletari, F., Navab, N., Ahmadi, S.A.: V-net: Fully convolutional neural networks for volumetric medical image segmentation. In: Fourth international conference on 3D vision (3DV) (2016)
23. Neshatavar, R., Yavartanoo, M., Son, S., Lee, K.M.: Cvf-sid: Cyclic multi-variate function for self-supervised image denoising by disentangling noise from image. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 17583–17591 (2022)
24. Qiu, B., Huang, Z., Liu, X., Meng, X., You, Y., Liu, G., Yang, K., Maier, A., Ren, Q., Lu, Y.: Noise reduction in optical coherence tomography images using a deep neural network with perceptually-sensitive loss function. Biomedical Optics Express **11**(2), 817–830 (2020)
25. Ren, J., Park, K., Pan, Y., Ling, H.: Self-supervised bulk motion artifact removal in optical coherence tomography angiography. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 20617–20625 (2022)
26. 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. pp. 234–241. Springer (2015)
27. Schmitt, J.M.: Optical coherence tomography (oct): a review. IEEE Journal of selected topics in quantum electronics **5**(4), 1205–1215 (1999)
28. Shamouilian, M., Selesnick, I.: Total variation denoising for optical coherence tomography. In: IEEE Signal Processing in Medicine and Biology Symposium (SPMB) (2019)

29. Sheth, D.Y., Mohan, S., Vincent, J.L., Manzorro, R., Crozier, P.A., Khapra, M.M., Simoncelli, E.P., Fernandez-Granda, C.: Unsupervised deep video denoising. In: Proceedings of the IEEE/CVF International Conference on Computer Vision. pp. 1759–1768 (2021)
30. Wang, X., Girshick, R., Gupta, A., He, K.: Non-local neural networks. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 7794–7803 (2018)
31. Wang, Z., Liu, J., Li, G., Han, H.: Blind2unblind: Self-supervised image denoising with visible blind spots. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 2027–2036 (2022)
32. Yang, J., Hu, Y., Fang, L., Cheng, J., Liu, J.: Universal digital filtering for denoising volumetric retinal oct and oct angiography in 3d shearlet domain. Optics Letters **45**(3), 694–697 (2020)
33. Yu, X., Ge, C., Li, M., Yuan, M., Liu, L., Mo, J., Shum, P.P., Chen, J.: Self-supervised blind2unblind deep learning scheme for oct speckle reductions. Biomedical Optics Express **14**(6), 2773–2795 (2023)
34. Zhang, D., Zhou, F.: Self-supervised image denoising for real-world images with context-aware transformer. IEEE Access **11**, 14340–14349 (2023)