Received: 22 March 2021 | Revised: 18 July 2021 | Accepted: 19 July 2021

DOI: 10.1002/jbio.202100097

RESEARCH ARTICLE

JOURNAL OF BIOPHOTONICS

Deep-learning-based motion correction in optical coherence tomography angiography

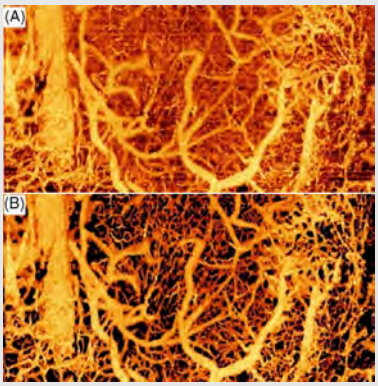
Ang Li | **Congwu Du** | **Yingtian Pan***

Department of Biomedical Engineering, Stony Brook University, Stony Brook, New York, USA

***Correspondence**
Yingtian Pan, Department of Biomedical Engineering, Stony Brook University, Stony Brook, NY 11794, USA.
Email: yingtian.pan@stonybrook.edu

Funding information
National Institutes of Health, Grant/Award Numbers: RF1DA048808, R01DA029718

Abstract



Optical coherence tomography angiography (OCTA) is a widely applied tool to image microvascular networks with high spatial resolution and sensitivity. Due to limited imaging speed, the artifacts caused by tissue motion can severely compromise visualization of the microvascular networks and quantification of OCTA images. In this article, we propose a deep-learning-based framework to effectively correct motion artifacts and retrieve microvascular architectures. This method comprised two deep neural networks in which the first subnet was applied to distinguish motion corrupted B-scan images from a volumetric dataset. Based on the classification results, the artifacts could be removed from the en face maximum-intensity-projection (MIP) OCTA image. To restore the disturbed vasculature induced by artifact removal, the second subnet, an inpainting neural network, was utilized to reconnect the broken vascular networks. We applied the method to postprocess OCTA images of the microvascular networks in mouse cortex in vivo. Both image comparison and quantitative analysis show that the proposed method can significantly improve OCTA image by efficiently recovering microvasculature from the overwhelming motion artifacts.

KEYWORDS
deep neural networks, microvascular network, motion correction, OCTA

1 | INTRODUCTION

Optical coherence tomography angiography (OCTA) is a functional extension of OCT that has been applied to various preclinical and clinical diagnoses, including brain functional imaging [1–3], dermatology [4] and ophthalmology [5, 6]. To differentiate vasculature from the surrounding tissue, OCTA utilizes an endogenous contrast

Abbreviations: ClNet, classification neural network; CNN, convolutional neural network; ComNet, completion subnet; MIP, maximum-intensity-projection; OCTA, optical coherence tomography angiography; SegNet, segmentation subnet; TrainD, training dataset; TestD, testing dataset; ValidD, validation dataset.

mechanism that originates from speckle decorrelation induced by the flowing red blood cells. Specifically, OCTA is obtained by extracting the variance within multiple repeated B-scans at a cross-sectional location by using traditional intensity-variance based methods [7] and more recently deep-learning based methods [8]. The decorrelated pixels, corresponding to vasculature, show significantly higher interframe variance than the pixels of static tissue. However, due to limited raster scanning speed and other imperfections, OCTA is susceptible to motion artifacts induced by various tissue motions such as respiration, heartbeat and mechanical jittering. The motion artifacts during consecutive B-scans result in intensity fluctuations and thus generate strip-like artifacts in the en face OCTA MIP image. Such motion artifacts seriously degrade image quality, for example, reducing image sensitivity and causing distortion in vascular topology.

Efforts to correct motion artifacts in an OCTA image can be categorized in the hardware-based or the software-based approaches. The hardware methods typically require the integration of additional imaging modality to evaluate tissue micromotion and guide the OCTA rescan to counteract the artifacts [9]. Although such techniques have proven effective, the associated system implementation and image acquisition are highly complex and cost-ineffective, which restricts them from widespread adaptations. The software-based methods circumvent the need for hardware modification and thus exhibit high competitiveness for various preclinical or clinical applications. Most of the post-processing methods operate in a manner of image registration, in which strip artifacts on MIP images were replaced by combining information from the repeated OCTA acquisition of the same tissue [10, 11]. Although the initial results were encouraging, such registration-based methods required considerably repeated image acquisition and extended imaging time, thus reducing the practicability for many clinical diagnoses. To solve the problem, we reported an automatic method utilizing tensor voting to reconnect the broken vascular segments and thus enable efficient correction of motion artifacts based on a single OCTA image [12]. However, the effectiveness of this method may deteriorate when severe motion artifacts occur.

To tackle these challenges, here we propose a fully automated motion correction method that employs deep neural networks. The proposed method contains two steps. First, a classification neural network (ClNet) was applied to the acquired OCTA image dataset to differentiate the motion corrupted B-scans from the clean counterparts and remove them from the image cube. However, removal of these motion-corrupted images resulted in disruption of the en face OCTA projection image, including inclusion of multiple blank image gaps. Then, a segmentation network was applied to detect the broken vasculature and an auxiliary inpainting network was used to restore the vascular connections. Results of both visual inspection and quantification show that the deep-learning method is highly effective for correcting the overwhelming artifacts on OCTA images caused by severe and consecutive tissue motion, which is promising for significantly enhancing OCTA for preclinical and clinical diagnosis and evaluation of disease prognosis and treatment effects.

2 | METHODS

2.1 | System setup and data acquisition

An ultrahigh-resolution optical coherence tomography setup (μOCT) was used to acquire in-vivo OCTA images. An ultra-broadband light source ($\lambda = 1{,}310$ nm, $\Delta\lambda_{FWHM} \approx 220$ nm) was applied to illuminate the μOCT system, yielding an axial resolution of 2.5 μm. The beam exiting from sample arm was collimated to 5 mm, steered transversely by a galvo scanner (VM500, General Scanning) and then focused by an achromatic lens (f16mm/NA0.25) onto mouse cortex. The lateral resolution, determined by the effective NA of the achromat was ~5 μm. The interference fringes spectrally encoding the depth profile (A-scan) were detected by a fast linescan InGaAs CMOS camera (2048-pixels; GL2048, Sensors Unlimited) interfaced with an image workstation via a camera link card (PCIe-1433, NI).

In this study, C57/B6 mice (male, 12–14 weeks of age, Jackson Lab) were used, and a chronic cranial window was implanted on the surface of each animal's sensorimotor cortex prior to in vivo imaging. A total of 26 animals were imaged and 22 images were used in the study. For OCTA image acquisition, the animal was anesthetized with inhalational isoflurane and the animal head was mounted on a custom stereotaxic frame to minimize motion artifacts, during which eight consecutive B-scans (x-z plane) at each cross section (y-axis) were acquired at an A-line rate of 27 kHz to derive a cross-sectional OCTA image. The physiology of the animal was monitored, including electrocardiography, respiration rate, and body temperature, and so on. The depth of anesthesia of the animal was adjusted by ramping the isoflurane concentration (e.g., from 2.4% down to 1% or lower) to acquire motion-free and artifacts-corrupted OCTA images, respectively.

2.2 | Deep learning-based artifacts correction pipeline

Removing motion artifacts is a crucial preprocessing step to enhance OCTA image fidelity. The patterns of motion artifacts may share similar features with vasculature. For example, it is difficult to separate the strip artifacts from vessels in an en face MIP OCTA image. A widely used strategy is to select image lines with average intensity or vesselness over a global threshold. However, some clean lines across large branch vessels may be mistakenly selected, whereas artifacts that are not across the entire image but rather unevenly spaced cannot be efficiently detected. To improve the robustness to detect artifacts, a convolutional neural network (ClNet) was first applied for the classification of clean and motion corrupted B-scan images. As shown in Figure 1A, the ClNet is mainly based on the residual CNN [13]. Seven blocks, each containing a 3×3 convolutional layer and a batch normalization layer, were applied to extract the image features. Skip connections with 1×1 convolutional layer used to unify filter number were constructed to facilitate the training. Downsampling was performed by convolutional layers that have a stride of 2 to reduce memory usage and computational time. To train the ClNet, 600 B-scan images were manually selected as training dataset and additional 200 images were selected as testing dataset. All images were resized to 512×512 before feeding into the network after data normalization as shown in Figure 1C. We used Adam with a mini-batch size of 64 and 10^{-4} learning rate. On the testing dataset, the network achieved a 98.5% accuracy of the classification.

After artifacts detection, artificial B-scan images in a volumetric dataset were discarded, resulting in stripe-like gaps (empty spaces) on the projected OCTA image. These gaps need to be refilled to restore the connectivity of the vascular network. This can be implemented by the deep learning based inpainting, which can synthesize alternative contents by learning adaptive image features for different semantics [14]. However, directly applying inpainting network on the broken OCTA image may lead to suboptimal performance on vascular reconnection. This is because the typical inpainting reconstruction ensures the appearance consistency of the generated regions with ground truth, but it does not penalize topological mistakes. To facilitate learning of vascular connectivity, a two-stage neural network was utilized, as shown in Figure 1B. In stage 1, a segmentation subnet (SegNet) was performed on a broken OCTA image to separate vasculature from the background. In stage 2, an image completion subnet (ComNet) was applied to the probability map from the above Segnet to predict the missing vessel segments. With vascular segmentation, the ComNet can pay special attention to learn topological constraints of the vessels and focus on the structure recovery inside the unknown region. The SegNet is based on a U-net structure with dense connection blocks, in which direct connections are built from each convolutional layer to others for better feature propagation and fusion [15]. The max pooling and deconvolution layers are utilized, respectively, in expansive and contracting paths for downsampling and upsampling. The ComNet follows an encoder-decoder structure, which stacks gated convolution layers to handle the gaps with various width in a broken OCTA image [16]. In the ComNet, the feature maps are downsampled by two stride-2 gated convolutional layers to 1/4 of the input size and then recovered to its full size by deconvolutional gated layers before the outputs. The SegNet and ComNet were trained jointly using Dice loss functions:

$$l_{\text{seg}} = l_{\text{dice1}} + l_{\text{dice2}} \tag{1}$$

$$l_{\text{dice}} = 1 - \frac{y\hat{y} + \epsilon}{y + \hat{y} + \epsilon} - \frac{(1-y)(1-\hat{y}) + \epsilon}{2 - y - \hat{y} + \epsilon} \tag{2}$$

where y and $\hat{y}$ represent ground truth and predicted probability map, respectively, and ϵ is a smoothing term. During the training, OCTA images with missing areas were used as the inputs. The corrupted vessel masks and their uncorrupted counterparts were applied as ground truth to train SegNet and ComNet, respectively.

A total of 16 MIP OCTA images of 750×512 pixels (corresponding to $2.5 \times 2\ \text{mm}^2$) were manually labeled and split into the training dataset (TrainD, n = 10) and the validation dataset (ValiD, n = 6). The datasets that were constructed on the OCTA images free of motion artifacts were used to simplify the procedure for manual segmentation. After data normalization, image patches of 128×128 pixels were extracted and fed into the networks as shown in Figure 1C. During the training, a fraction of image lines with a predetermined occupation rate (α) from the images patches and the corresponding vessel masks in each training iteration were randomly removed to simulate motion artifacts. The Adam optimizer was used with a mini-batch size of 16 and 10^{-4} learning rate. In addition, 6 MIP OCTA images corrupted by severe motion artifacts were labeled as testing dataset (TestD, n = 6) to evaluate the performance of the proposed method. The neural networks were implemented on the Keras platform (https://keras.io/) using NVIDIA GPUs (RTX 2070).

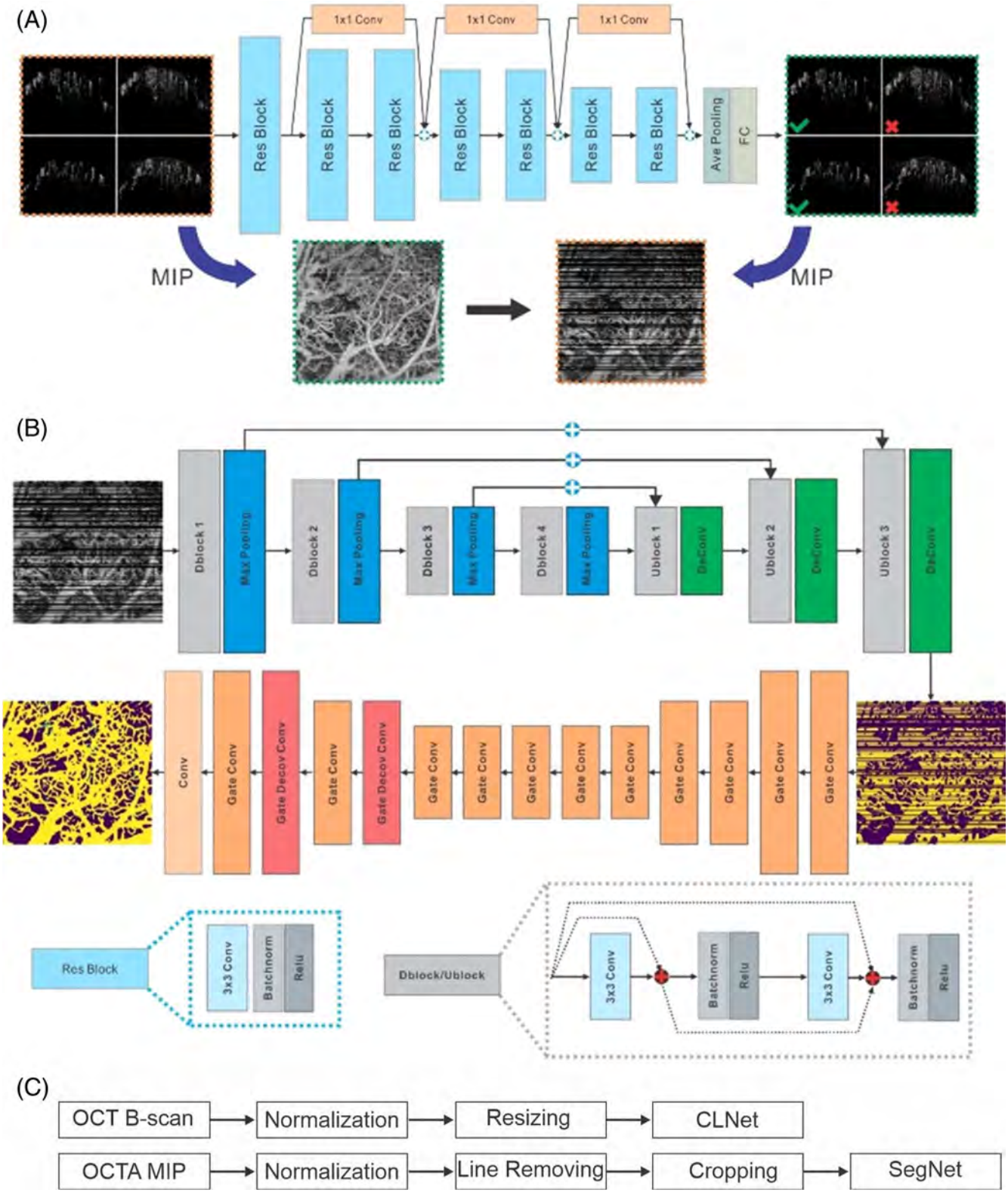


FIGURE 1 A, The classification network (CLNet) was applied to differentiate motion corrupted B-scan images from clean ones. The B-scan images with artifacts were then removed from the MIP OCTA image. B, To fill the gaps in the MIP OCTA image, two-stage networks containing the SegNet and ComNet were used, in which SegNet segmented the broken vessels and ComNet restored the vascular connectivity. C, Flowchart to illustrate the data preprocessing steps

3 | RESULTS AND DISCUSSION

Figure 2 shows an exemplary OCTA image (Figure 2A) that contained corrupted frames induced by motion artifacts, their vessel masks (Figure 2B) derived by our deep-learning method, and the corresponding correction images (Figure 2C) after applying the vessel masks. Artifact reduction is evident and vessel segmentation was successfully achieved by the proposed framework. The corrected vessel masks (Figure 2B) show that the microvascular network can be effectively distinguished from the background without noticeable strip artifacts presented. The resultant artifacts-corrected image (Figure 2C) shows considerable signal-to-noise improvement in comparison with the original image (Figure 2A). Microvasculature buried under overwhelming motion artifacts in the original images can be clearly resolved and appear to be continuous and natural within the vascular trees.

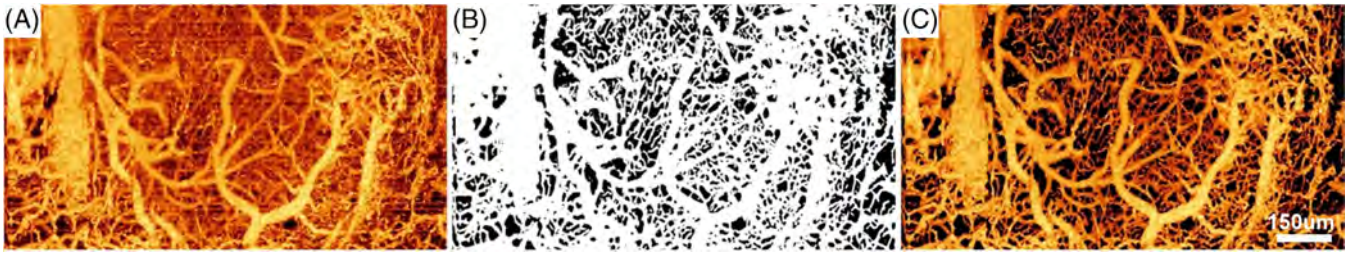


FIGURE 2 Deep-learning-based correction of motion artifacts in OCTA images of 3D microvascular network in mouse sensorimotor cortex. A, two maximum-intensity-projection (MIP) OCTA images corrupted by severe motion artifacts; B, vessel masks of (A) using the proposed deep-learning method; C, correction of motion artifacts in A by applying vessel masks in B

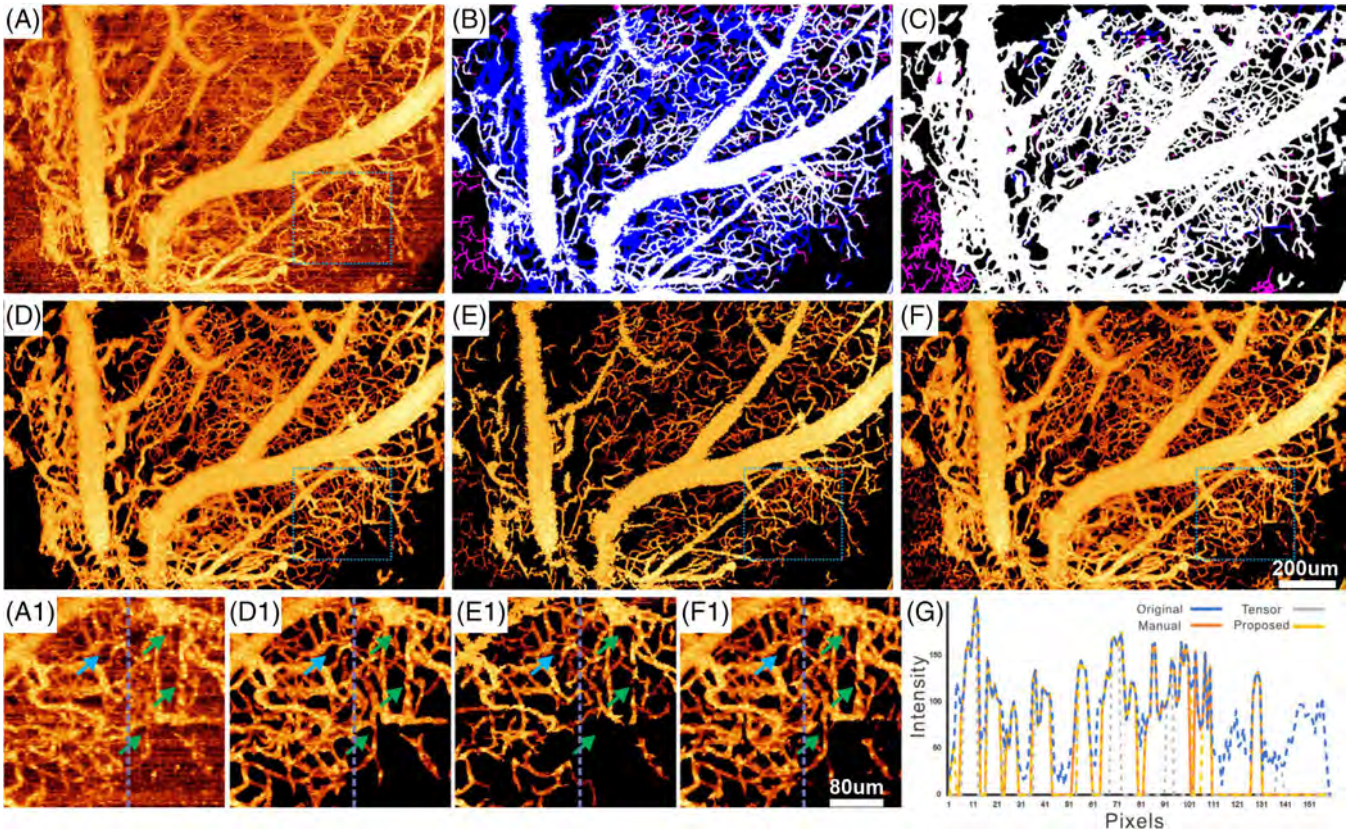


FIGURE 3 A, Original MIP OCTA image with overwhelming artifacts. B,C, vessel masks derived from tensor-voting based method and the proposed deep-learning method overlayed with the ground truth, in which the false positive and false negative pixels are highlighted in purple-color and blue-color regions, respectively. D-F, MIP OCTA images after artifact correction of the ground truth, the tensor-voting method, and the proposed method. A1, D1-F1, zoom-in images of the corresponding dash boxes. G, the intensity profiles across the dashed blue lines in panels A-F

To demonstrate the efficacy of the proposed method for motion-artifact reduction, we compared it with our prior tensor voting based method [12]. The vessel-correction masks from tensor-voting method and the proposed method are overlayed on the ground truth mask respectively to highlight the segmentation accuracy, as shown in Figure 3B,C. Tensor-voting method failed to identify quite some vascular networks due to low image contrast and significant errors caused by the overwhelming motion artifacts, although the majority of the artifacts were removed. In comparison, the proposed method showed significantly improved performance on vessel segmentation, for example, much better vascular connectivity and microvasculature detection. The artifacts-corrected images (Figure 3D-F) and the corresponding zoom-in images (Figure 3D1-F1) further proves that the proposed method is more effective in vessel preservation and removal of artificial stripes. As highlighted by green arrows in Figure 3E1, the tensor-voting method missed considerable numbers of vasculature segments, whereas artifacts were not thoroughly eliminated (blue arrow). On contrary, the proposed method shows significantly better fidelity, providing similar microvascular turnouts as those of the ground truth. The intensity profiles (Figure 3G) across a selected dashed green line in the zoom-in images Figure 3A1-E1 indicates that the proposed method (yellow trace) dramatically suppressed artifacts with high preservation of signals. As a result, the detailed capillary flow network was clearly resolved. A quantitative analysis based on TestD further demonstrated that the proposed method outperformed the tensor-voting method, which significantly increased the Dice score from 0.776 to 0.890.

We further conducted an ablation study and examined the outcome of only using segmentation network to distinguish vasculatures from artifacts. To do so, we trained the SegNet independently without ClNet and ComNet on TrainD and compared it with the proposed framework on TestD. Figure 4A,B shows a pair of exemplary maximum-intensity-projection OCTA images before and after correction using manual segmentation. In the case of applying single SegNet, Figure 4C shows partial differentiation of vascular architecture from artifacts, indicating that it still suffered from undesired artifacts as highlighted by red arrows. In contrast, the proposed framework enhanced artifact removal (Figure 4D). The zoom-in images (Figure 4 A1-D1) revealed that the single SegNet was unable to segment the microvasculature as accurately as the proposed framework, which is evidenced by considerable unconnected vascular segments (green arrows). The presence of strip artifacts led to poor performance of SegNet. The Dice score graph in

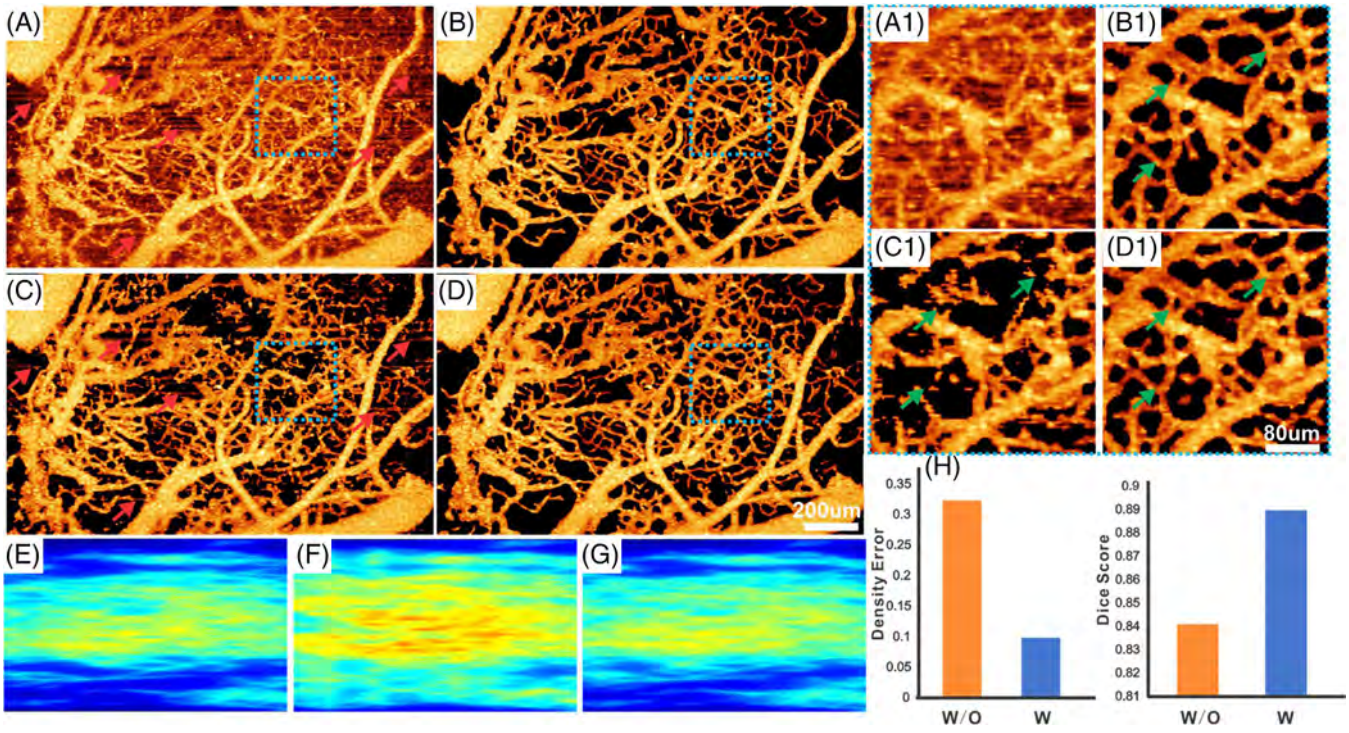


FIGURE 4 A-D are original projected OCTA image, artifacts corrected image with manual segmentation, artifacts corrected image of single SegNet and the proposed framework. A1-D1 are zoom-in images corresponding to dash boxes on A-D. E-G are vessel density maps of ground truth, single SegNet and the proposed framework. H is the quantifications of vessel density errors (left) and Dice score (right)

Figure 4H shows the quantitative improvement of the proposed framework (0.890) over single SegNet (0.841). In addition to the effects on visualization, the error of vessel segmentation caused by single SegNet may significantly bias the quantitative analysis of vascular networks. To evaluate this, we quantified the vessel density on the masks generated by single SegNet and the proposed method. As shown in the density maps, single SegNet caused serious overestimation of vessel density due to the uncorrected strip artifacts (Figure 4F), whereas the proposed method provided far more unbiased assessment (Figure 4G) vs the ground truth (Figure 4E). The error estimation of vessel density in Figure 4H shows the proposed method generated much more accurate quantification than single SegNet (0.098 vs 0.323).

The influence of severity of the motion-induced artifact, represented by the artifact percentage (AP = artifact line #/total line #), was simulated to further evaluate the efficacy of the deep-learning method for restoring the connectivity of the broken vasculature. In the ValiD group, the artifact gaps were created by randomly removing lines (i.e., cross sections) with different AP levels ranging from 0 to 70%. Then, the broken OCTA images were restored by the two-stage reconstruction neural network. Figure 5 shows the simulation results based on an OCTA image of a mouse sensorimotor cortex during week 3 after viral injection (the dark hole is the injection spot showing disruption to the local vascular network). The top panels are the original input OCTA image (Figure 5A: AP = 0%) and those with simulated moderate (Figure 5B: 30%), medium (Figure 5C: 50%) and severe (Figure 5D: 70%) motion artifacts. The mid panels (Figure 5A1-D1) are the corresponding vesselness masks derived by our deep-learning method and the panels (Figure 5 A2-D2) are their zoom-in views of the dashed boxes to highlight the degraded restoration of the capillary networks. Interestingly, a comparison among panels (Figure 5A1-D1) shows that deep learning was able to seamlessly fill up the gaps for moderate (30%, 50%) and even severe (70%) artifacts except that few stripes were present as shown by arrows. This indicates that the efficacy of deep learning might be limited to fully restore large gaps, leading to more vascular disconnections if motion-induced artifacts are substantial. It is noteworthy that higher motion artifacts resulted in more errors of the restored capillary networks. Indeed, as highlighted by the dashed circles in panels Figure 5A2-D2, the number of missing junctions is significantly higher in severely corrupted image Figure 5D2 than the moderately corrupted cases Figure 5B2,C2. This trend is indicated in the quantitative analysis Figure 5E, in which the Dice score decreased with the AP increase, especially when AP increased to 70%.

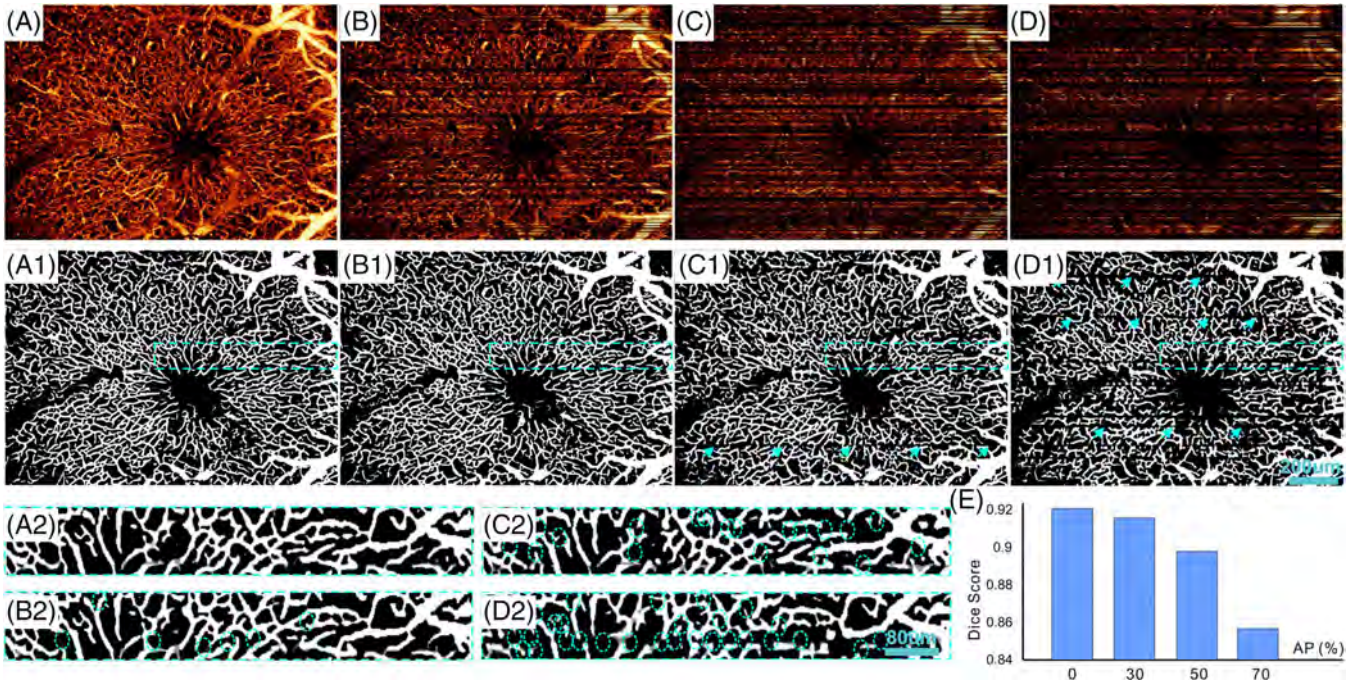


FIGURE 5 Simulation results of the efficacy of deep learning to restore moderate to severe motion artifacts. A-D, OCTA images of original, moderate (AP = 30%), medium (50%) and severe (70%) motion artifacts; A1-D1: binary vesselness maps recovered by our deep-learning method, arrows show the unrestored stripes; A2-D2, zoom-in images corresponding to dash boxes in panels A1-D1, dashed circles show the missing capillary junctions; E, the influence of artifact severity on vascular connection, Dice scores dropped from 0.92 for AP≤50% to 0.85 for AP = 70%

4 | CONCLUSIONS

It is known that OCTA imaging often suffers from severe motion-induced artifacts in various awake scenarios, resulting in suboptimal vascular visualization and biased quantification. These include OCTA imaging in preclinical applications such as brain functional studies and in clinical diagnoses of diseases ophthalmology and dermatology. To tackle the challenge, we reported a deep learning framework for efficient correction of motion artifacts and presented in vivo rodent brain imaging results to demonstrate that this method is well suited for removing severe motion artifacts based on a single OCTA image.

As shown in Figure 1A, the classification network, CLNet, characterized the vascular morphology on a cross-sectional OCTA image (B-scan) and differentiated the motion corrupted B-scans from the clean B-scan dataset, thus allowing more efficient detection of partial or mild strip artifacts than the widely used threshold-based method. Following the artifacts removal, a two-stage neural network was designed to restore the broken vascular image as illustrated in Figure 1B. Specifically, the first subnet distinguished the vascular morphology (vesselness) from the background; then, the second subnet facilitated a robust and accurate restoration of vascular connectivity. The experimental results in Figure 2 demonstrate that the proposed method enabled efficient recovery of the microvascular network overwhelmed by severe motion artifacts. To further examine the efficacy of the proposed deep-learning method, we compared the outcome of this method with that of the previously reported tensor voting method. The results in Figure 3 show that the new deep-learning method significantly outperformed the tensor voting method on motion-artifact removal and vasculature reconnection. We also demonstrated in Figure 4 that the proposed framework with artifacts identification and image inpainting was able to achieve better performance than direct separation of vessels and strip artifacts. It is noteworthy that a potential limitation of this method was that the neural network was unable to fully fill up large gaps induced by several consecutive motion artifacts (strips). The simulation results in Figure 5 indicate that the performance of the proposed method for vesselness restoration was high and stable when moderate (30%) and medium (50%) motion artifacts were present. However, unfilled stripes (highlighted by arrows) and gaps across the microvascular network (dashed circles) started to appear when severe (70%) motion artifacts occurred. The probable reason is that the deep-learning networks could not capture enough contextual information within a large gap to make a reasonable prediction for the missing vascular segments. This problem may be solved by increasing the depth of reconstruction neural network to enlarge the receptive field. Another potential limitation of the proposed method is that it cannot effectively correct the geometric distortion of vascular network caused by severe tissue motion along slow axis. The reconstruction network in the method has the potential to reconnect vasculature with structural distortion by learning features of normal vasculature from motion-free OCTA images. However, the CLNet applied in our study cannot explicitly detect the distortion. To overcome this limitation, an additional deep neural network may be introduced as an extra step to detect the image areas corresponding to the geometric distortion. Another limitation of the proposed method is the small testing dataset. We will increase the number of samples in future studies to further prove the generalization the proposed method.

In conclusion, we reported a deep-learning-based image processing method for motion artifact correction and demonstrated its great potential for efficiently minimizing motion artifacts in OCTA images in both preclinical and clinical applications.

ACKNOWLEDGMENTS

The authors thank Kichon Park for in vivo animal preparations and Yigi Yan for labeling motion-free images. This research was supported by National Institutes of Health grants R01DA029718 (Congwu Du, Yingtian Pan) and RF1DA048808 (Yingtian Pan, Congwu Du).

CONFLICT OF INTEREST

The authors declare no potential conflict of interest.

AUTHOR CONTRIBUTIONS

Ang Li carried out image and data analyses; All authors contributed significantly to data interpretation, discussing the results, and writing the manuscript.

DATA AVAILABILITY STATEMENT

Research data are not shared.

ORCID

Ang Li https://orcid.org/0000-0003-4252-9933

REFERENCES

[1] C. P. Allen, K. Park, A. Li, N. D. Volkow, G. F. Koob, Y. Pan, X. T. Hu, C. Du, *Addict. Biol.* **2019**, *24*, 485.

[2] J. You, C. Pan, K. Park, A. Li, C. Du, *J. Biophotonics* **2020**, *13*, e201960091.

[3] A. Li, J. You, C. Du, Y. Pan, *Biomed. Opt. Express* **2017**, *8*, 5604.

[4] Z. Chen, M. Liu, M. Minneman, L. Ginner, E. Hoover, H. Sattmann, M. Bonesi, W. Drexler, R. A. Leitgeb, *Biomed. Opt. Express* **2016**, *7*, 3032.

[5] A. Y. Kim, Z. Chu, A. Shahidzadeh, R. K. Wang, C. A. Puliafito, A. H. J. I. O. Kashani, *Science* **2016**, *57*, OCT362.
[6] Y. Jia, S. T. Bailey, T. S. Hwang, S. M. McClintic, S. S. Gao, M. E. Pennesi, C. J. Flaxel, A. K. Lauer, D. J. Wilson, J. Hornegger, J. G. Fujimoto, H. G. Huang, *Proc. Natl. Acad. Sci. USA* **2015**, *112*, E2395.
[7] M. S. Mahmud, D. W. Cadotte, B. Vuong, C. Sun, T. W. Luk, A. Mariampillai, V. X. Yang, *J. Biomed. Opt.* **2013**, *18*, 050901.
[8] Z. Jiang, Z. Huang, B. Qiu, X. Meng, Y. You, X. Liu, G. Liu, C. Zhou, K. Yang, A. Maier, Q. Ren, Y. Lu, *Biomed. Opt. Express* **2020**, *11*, 1580.
[9] Q. Zhang, Y. Huang, T. Zhang, S. Kubach, L. An, M. Laron, U. Sharma, R. K. Wang, J. Biomed. Opt. *J. Biomed. Opt.* **2015**, *20*, 066008.
[10] D. W. Wei, A. J. Deegan, R. K. Wang, *J. Biomed. Opt.* **2017**, *22*, 066013.
[11] A. Camino, M. Zhang, S. S. Gao, T. S. Hwang, U. Sharma, D. J. Wilson, D. Huang, Y. Jia, *Biomed. Opt. Express* **2016**, *7*, 3905.
[12] A. Li, G. Zeng, C. Du, H. Zhang, Y. Pan, *Appl. Phys. Lett.* **2018**, *113*, 101102.
[13] K. He, X. Zhang, S. Ren, J. Sun, in 2016 IEEE Conf. Comp. Vision Pattern Recogn., Las Vegas, NV., **2016**, pp.770–778.
[14] J. Xie, L. Xu, E. Chen, *Adv Neural Inform Process Syst* **2012**, *25*, 341.
[15] O. Ronneberger, P. Fischer, T. Brox, in Med. Image Comp. Comp-Assist. Intervent., Munich, Germany, **2015**, pp.234–241.
[16] J. Yu, Z. Lin, J. Yang, X. Shen, X. Lu, T. S. Huang, in IEEE Int. Conf. Com. Vision, Seoul, Korea, **2019**, pp.4471–4480.

How to cite this article: A. Li, C. Du, Y. Pan, *J. Biophotonics* **2021**, *14*(12), e202100097. https://doi.org/10.1002/jbio.202100097