

Supplementary Materials: Serial section microscopy image inpainting guided by axial optical flow

Anonymous Authors

1 EXPERIMENTS

1.1 Dataset and metric

To comprehensively estimate the network performance by evaluating its contribution to segmentation work and designing stability experiment, we introduce another **EPFL Dataset**¹. It provides an image stack with 1065 serial sections of size 2048×1536. This dataset is collected from the CA1 hippocampus region of a mouse brain. The dataset is acquired by FIB-SEM and the resolution of each voxel is approximately 5 × 5 × 5 nm.

The common segmentation map assessment metrics consist of VOI (Variation of Information) [3] and ARAND (Adapted Rand Error) [4]. VOI calculates the sum of two conditional entropies of the predicted and the ground truth segmentation maps, while ARAND measures the similarity between the predicted segmentation maps and the ground truth segmentation maps based on the cluster to which randomly selected pixels belong.

1.2 Further Comparison on Segmentation Results

To demonstrate the rationality of FlowInpaint’s inpainted results, we further perform neural segmentation tasks on them. Here, we adopt a random walker-based image segmentation method [2] to generate the segmentation maps of the inpainted results. The quantitative and visual comparison results of inpainted images in segmentation tasks on the CREMI dataset are presented in Table 1 and Figure 1. From these segmentation maps, we can tell that the generated textures that violate the trend of structural changes and blurry content that affects image feature extraction would lead to imprecise neural segmentation results. However, our method has already addressed the aforementioned issues during the scheme design. Thus FlowInpaint’s results can well enhance the accuracy of subsequent work to some extent.

| Method | Metric | | | | | |
|-----------------|--------|--------|--------|--------|--------|--------|
| | CREMIA | | CREMIB | | CREMIC | |
| | VOI↓ | ARAND↓ | VOI↓ | ARAND↓ | VOI↓ | ARAND↓ |
| PEN-NET | 1.887 | 0.468 | 2.779 | 0.563 | 2.131 | 0.511 |
| AOT-GAN | 1.569 | 0.413 | 2.473 | 0.559 | 2.033 | 0.504 |
| CCPGAN | 1.175 | 0.218 | 1.292 | 0.369 | 1.183 | 0.355 |
| SSF-Restoration | 1.098 | 0.214 | 1.174 | 0.361 | 1.171 | 0.312 |
| Ours | 1.073 | 0.194 | 1.168 | 0.357 | 1.164 | 0.305 |

Table 1: Inpainting results on neuron segmentation task. ↑ means higher is better. ↓ means lower is better.

1.3 Stability Analysis

1.3.1 Stability on Contrast. In real-world scenarios, low contrast is often introduced during the imaging process and affects the recognition of structure information. We simulate the neighboring images I_{i-1}^{con} under different contrast ratios by adjusting the brightness (B) and the haze (H) level of the original image I_{i-1} as follows:

$$I_{i-1}^{con} = I_{i-1} \times (1 - H) + 255 \times B \times H. \quad (1)$$

A lower brightness (B) value corresponds to a darker image, whereas a higher haze (H) level value indicates lower clarity. We set the brightness (B) between 0.6~1 and the haze (H) between 0.4~0.8 based on the contrast variation in the real world, and Table 2 summarizes the quantitative results. The change in contrast does not alter the relative intensity relationship between pixels, thus having little impact on optical flow estimation. Actually, FlowInpaint achieves nearly the same high-quality restoration results in different combinations of brightness and haze, which can also be derived from Figure 2.

| H(haze) | Metric | B(brightness) | | | | |
|---------|--------|---------------|--------|--------|--------|--------|
| | | 0.6 | 0.7 | 0.8 | 0.9 | 1.0 |
| 0.2 | PSNR↑ | 29.992 | 30.108 | 30.186 | 30.327 | 30.748 |
| | SSIM↑ | 0.911 | 0.916 | 0.920 | 0.925 | 0.930 |
| | FSIM↑ | 0.946 | 0.947 | 0.949 | 0.950 | 0.950 |
| | FID↓ | 8.111 | 8.030 | 7.971 | 7.859 | 7.806 |
| 0.3 | PSNR↑ | 29.887 | 30.062 | 30.140 | 30.202 | 30.299 |
| | SSIM↑ | 0.908 | 0.913 | 0.917 | 0.922 | 0.925 |
| | FSIM↑ | 0.946 | 0.947 | 0.948 | 0.948 | 0.948 |
| | FID↓ | 8.159 | 7.999 | 7.915 | 7.859 | 7.742 |
| 0.4 | PSNR↑ | 29.779 | 29.943 | 30.032 | 30.152 | 30.201 |
| | SSIM↑ | 0.904 | 0.910 | 0.914 | 0.918 | 0.921 |
| | FSIM↑ | 0.945 | 0.947 | 0.948 | 0.948 | 0.948 |
| | FID↓ | 8.190 | 8.116 | 7.968 | 7.892 | 7.825 |
| 0.5 | PSNR↑ | 29.716 | 29.847 | 29.903 | 29.971 | 30.040 |
| | SSIM↑ | 0.902 | 0.906 | 0.913 | 0.915 | 0.915 |
| | FSIM↑ | 0.945 | 0.945 | 0.946 | 0.947 | 0.948 |
| | FID↓ | 8.278 | 8.182 | 8.038 | 7.958 | 7.906 |
| 0.6 | PSNR↑ | 29.546 | 29.642 | 29.793 | 29.588 | 29.622 |
| | SSIM↑ | 0.898 | 0.902 | 0.905 | 0.908 | 0.911 |
| | FSIM↑ | 0.944 | 0.945 | 0.946 | 0.946 | 0.946 |
| | FID↓ | 8.411 | 8.265 | 8.183 | 8.073 | 8.003 |

Table 2: Stability Studies on contrast. ↑ means higher is better. ↓ means lower is better.

¹Dataset available at <https://www.epfl.ch/labs/cvlab/data/data-em/>

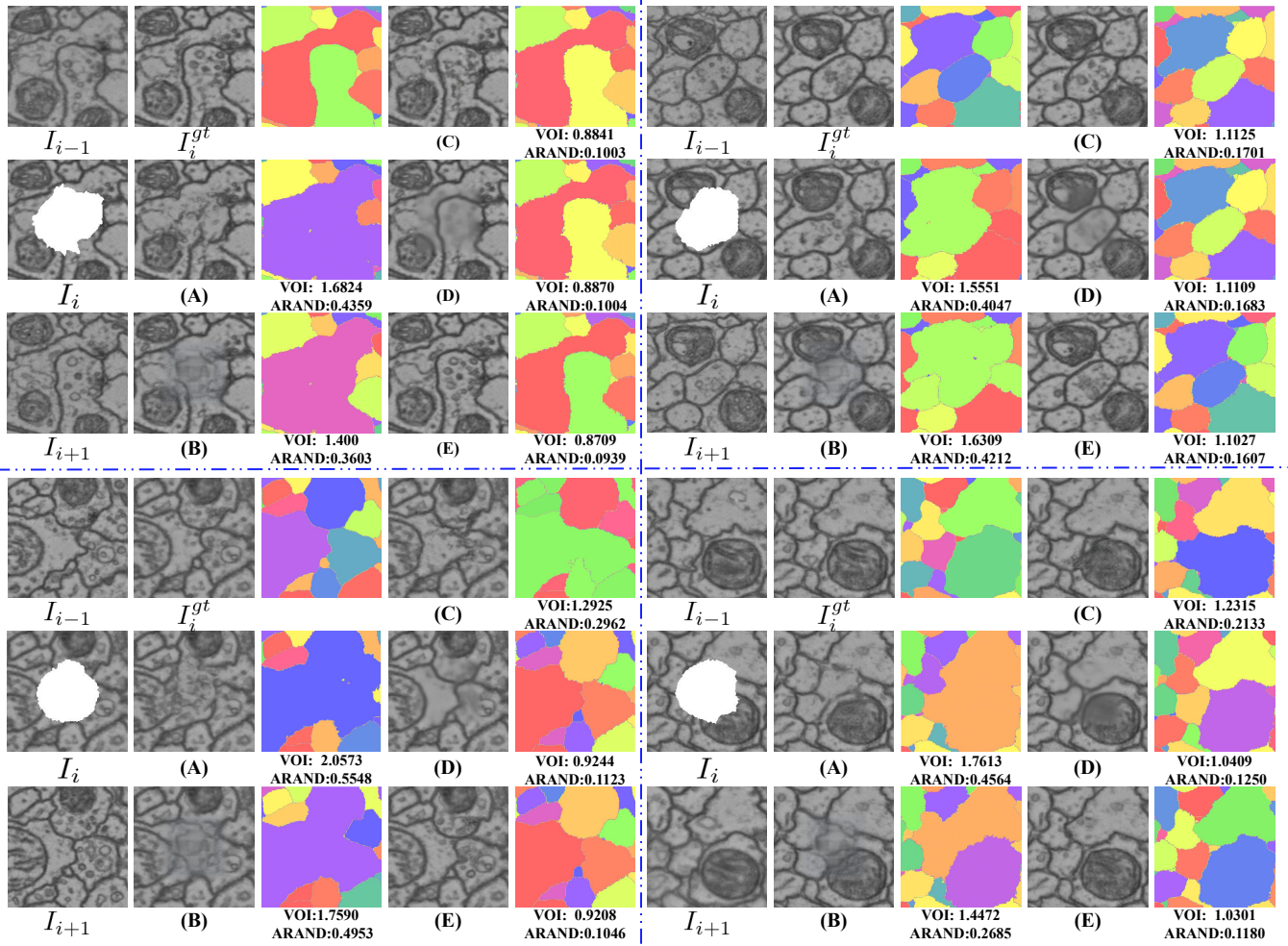


Figure 1: The neuron segmentation maps of inpainted results. (A) PEN-Net, (B) AOT-GAN, (C) CCPGAN, (D) SSF-Restoration, and (E) Ours.

1.3.2 Stability on Axial Resolution. The thickness of tissue sections determines the resolution along the sectioning direction (axial resolution). Typically, the common axial resolution of serial sections ranges from 30 to 70 nm, and the appearance of FIB-SIM allows for the improvement of resolution to 5~10 nm [1]. To test the stability of FlowInpaint under different axial resolutions, we simulate the axial resolution by adjusting the section interval among the EPFL image stacks, whose raw section interval is 5nm. We train the FlowInpaint on the data of $5 \times 5 \times 50 \text{ nm}^3$ resolution, and apply it on all testing data of $5 \times 5 \times 20 \sim 70 \text{ nm}^3$ resolution, which is shown in Table 3. With the reduction of axial resolution, the transformation of biological structures along the axial direction becomes larger and more difficult to estimate. When the axial resolution ranges from 20 to 60 nm, FlowInpaint generates acceptable recovered results.

REFERENCES

- [1] TD Ambegoda, Julien NP Martel, J Adamcik, M Cook, and Richard HR Hahnloser. 2017. Estimation of Z-thickness and XY-anisotropy of electron microscopy images using Gaussian processes. *Journal of Neuroinformatics and Neuroimaging* 2, 2

- (2017), 15–22.
- [2] Lorenzo Cerrone, Alexander Zeilmann, and Fred A Hamprecht. 2019. End-to-end learned random walker for seeded image segmentation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 12559–12568.
- [3] Marina Meilă. 2003. Comparing clusterings by the variation of information. In *Learning Theory and Kernel Machines: 16th Annual Conference on Learning Theory and 7th Kernel Workshop, COLT/Kernel 2003, Washington, DC, USA, August 24–27, 2003. Proceedings*. Springer, 173–187.
- [4] Ranjith Unnikrishnan, Caroline Pantofaru, and Martial Hebert. 2007. Toward objective evaluation of image segmentation algorithms. *IEEE transactions on pattern analysis and machine intelligence* 29, 6 (2007), 929–944.

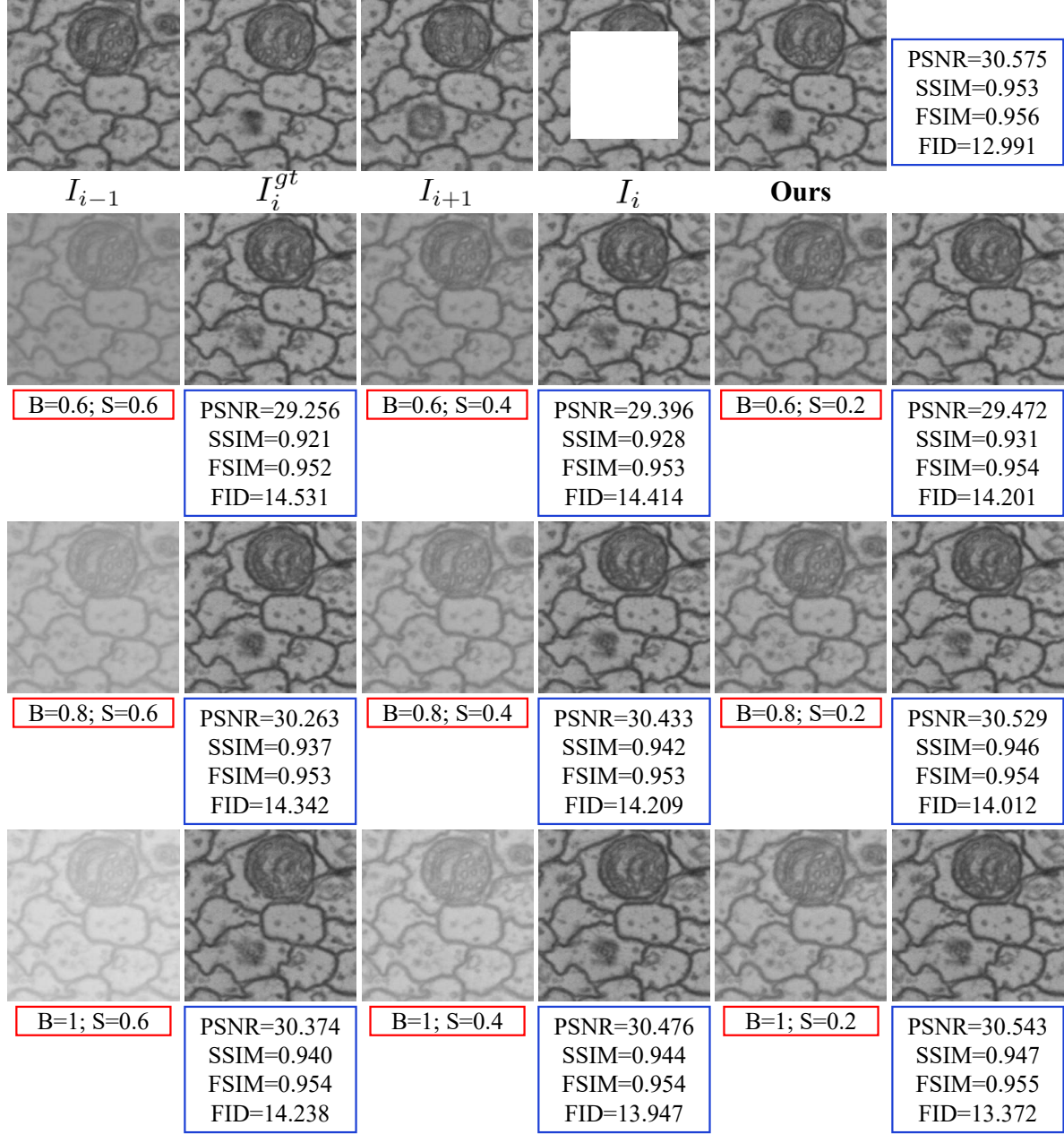


Figure 2: The example under different contrast ratios, B represents brightness and S represents the degree of spray.

| Resolution(nm ³) | 5 × 5 × 20 | 5 × 5 × 30 | 5 × 5 × 40 | 5 × 5 × 50 | 5 × 5 × 60 | 5 × 5 × 70 |
|------------------------------|------------|------------|------------|------------|------------|------------|
| SSIM↑ | 0.886 | 0.867 | 0.850 | 0.832 | 0.804 | 0.778 |
| PSNR↑ | 28.404 | 27.747 | 27.248 | 26.762 | 26.086 | 25.556 |
| FSIM↑ | 0.941 | 0.938 | 0.935 | 0.932 | 0.929 | 0.927 |
| FID↓ | 10.321 | 10.884 | 11.085 | 12.555 | 13.567 | 16.684 |

Table 3: Stability Studies on axial resolution (nm³). ↑ means higher is better. ↓ means lower is better.