Learning to Restore ssTEM Images from Deformation and Corruption

Anonymous ECCV submission

Paper ID 7

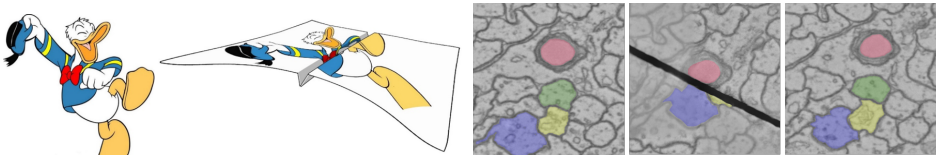
Abstract. Serial section transmission electron microscopy (ssTEM) plays an important role in biological research. Due to the imperfect sample preparation, however, ssTEM images suffer from inevitable artifacts that pose huge challenges for the subsequent analysis and visualization. In this paper, we propose a novel strategy for modeling the main type of degradation, *i.e.*, Support Film Folds (SFF), by characterizing this degradation process as a combination of content deformation and corruption. Relying on that, we then synthesize a sufficient amount of paired samples (degraded/groundtruth), which enables the training of a tailored deep restoration network. To the best of our knowledge, this is the first learning-based framework for ssTEM image restoration. Experiments on both synthetic and real test data demonstrate the superior performance of our proposed method over existing solutions, in terms of both image restoration quality and neuron segmentation accuracy.

Keywords: Degradation modeling · image restoration · deep learning.

1 Introduction

The development of electron microscopy (EM) at synapse resolution has greatly promoted the study of neuron morphology and connectomics, which is essential for understanding the working principle of intelligence. Among different EM imaging techniques [2,7,11,16,30,33], serial section transmission electron microscopy (ssTEM) has a clear advantage in the imaging speed and resolution [9,11,33]. It is thus widely adopted in analyzing the connectivity in volumetric samples of brain tissue by imaging many thin sections in sequence. Recently, relying on the ssTEM technique, researchers have imaged the first complete EM volume of the brain of adult *Drosophila* melanogaster [33], which is regarded as a milestone in brain science.

Due to the imperfect sample preparation, however, ssTEM images suffer from inevitable artifacts. Among these artifacts, one type of degradation is most common and has a significant influence on the image quality, *i.e.*, Support Film Folds (SFF, accounting for 3.2% of samples) [9]. SFF is caused by imperfect cutting in a complicated way. Fig. 1(a) visualizes this process with an intuitive example, where a 2D cartoon image is folded along a certain line and content around this line disappears in the deformed image. Fig. 1(b) shows three consecutive images with the middle one suffering from SFF degradation. As can be seen, the



(a) (b)

Fig. 1. (a) A cartoon image as an intuitive example to demonstrate SFF degradation. (b) Three consecutive ssTEM images where the middle one is with SFF artifacts.

folding effect not only results in content corruption (in a form of dark line), but also introduces severe deformation of surrounding regions which thus have large misalignments with the adjacent images.

The quality of ssTEM images is greatly reduced with the aforementioned SFF degradation, which severely hinders the subsequent tasks, such as alignment [26] and segmentation [10,17]. For example, in the neuron segmentation task, one image with artifacts could interrupt many neuron structures, resulting in erroneous reconstruction (see Fig. 7 for exemplar cases). Li *et al.* [17] also indicated that neuron segments are interrupted at multiple sections due to SFF degradation. However, it is challenging to restore ssTEM images from SFF degradation. Compared with degradations that generally occur on 2D natural images, SFF exhibits drastically different characteristics, making classic restoration methods difficult to apply. On the other hand, since there is no corresponding groundtruth for the degraded ssTEM image, it is also difficult to directly leverage on the power of deep learning based restoration. To the best of our knowledge, ssTEM image restoration from SFF degradation still remains an open problem.

To fill this gap, we propose to model SFF degradation as a combination of content deformation and corruption. Backed up by a statistic analysis from abundant real samples, we design an algorithm to synthesize degraded images from artifacts-free ones, where the latter are adopted as the corresponding groundtruth. Leveraging on the above obtained image pairs for training, we then propose a tailored deep learning framework, which consists of three modules, *i.e.*, interpolation, unfolding, and fusion. Specifically, the interpolation module utilizes the adjacent images to obtain an interpolated image, which is then used as a reference for the unfolding and fusion modules to address content deformation and corruption, respectively.

As demonstrated by comprehensive experimental results, our method significantly outperforms existing solutions for SFF restoration in ssTEM images, both quantitatively and qualitatively. With the generalizability of the degradation modeling and the scalability of the restoration network, the proposed method could play an essential role in bridging the gap between ssTEM images acquisition and subsequent analysis tasks, which thus facilitates the research of neuron morphology and connectomics.

The main contributions of this work are summarized as follows:

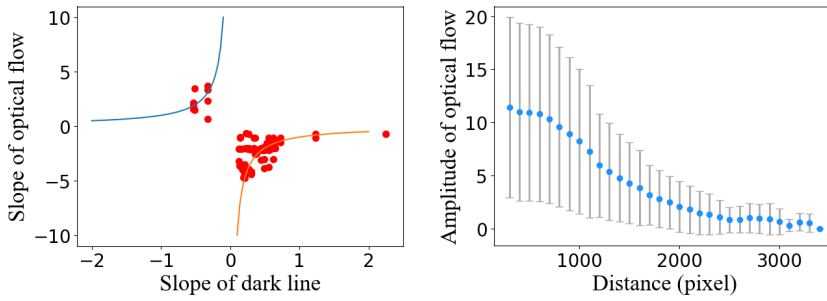
- We propose the first learning-based framework for ssTEM image restoration from SFF artifacts.
- We conduct a comprehensive analysis on the statistics of SFF. Based on the degradation modeling derived from the analysis, we propose a synthesis algorithm to generate degraded/groundtruth image pairs for training the deep restoration network.
- Experiments on both synthetic and real test data demonstrate the advantage of our proposed method, in terms of both image restoration quality and neuron segmentation accuracy.

2 Related Work

EM image restoration. Mainstream EM imaging techniques can be divided into two groups: transmission EM (TEM) [8,11,29,33] (with ssTEM as a representative) and scanning EM (SEM) [4,7,16,30]. These two types of techniques have different imaging advantages, and the acquired EM images suffer from different artifacts due to the flaws of respective imaging principles. Previously, EM image restoration is largely investigated within the scope of denoising [27], where the target degradation is relatively simple. Recently, a few works are reported to restore other types of artifacts in EM images. Khalilian-Gourtani *et al.* [14] proposed a method to detect and correct the striping artifacts in SEM images by solving a variational formulation of the image reconstruction problem. Maraghechi *et al.* [19] corrected three dominant types of SEM artifacts, *i.e.*, spatial distortion, drift distortion, and scan line shifts under an integrated digital image correlation framework. Meanwhile, Minh-Quan *et al.* [20] presented an asymmetrically cyclic adversarial network to remove the blob-like artifacts alongside Gaussian noise-like corruption in SEM images and background noise in TEM images. However, the dominant type of degradation in ssTEM images are seldom investigated, since SFF is much more complicated and challenging compared with the above mentioned ones.

Image inpainting. Restoration from SFF degradation is kind of similar to the image inpainting task which aims to fill in the corrupted areas in an image. Recently, a number of deep learning based approaches are proposed to accomplish this task, which generate promising results in certain challenging conditions [18,21,25,31,32]. However, directly applying existing inpainting methods would not work for SFF restoration. Take Fig. 1(b) for example, only filling in the dark line cannot fully recover the corrupted content which actually has a larger area than the dark line. Even the dark line is repaired in this way, the resultant image would have large misalignments with the adjacent images.

Substitution/Interpolation. In the neuron segmentation task [17], Li *et al.* completely discarded the degraded image and replaced it with a copy of the adjacent image. This substitution strategy reduces interrupted neurons to a certain extent, but may still lead to inaccurate reconstruction. Another straightforward solution for SFF restoration would be video frame interpolation. However, it is difficult to recover sufficient details since the content of adjacent images is



(a) (b)

Fig. 2. Statistical analysis of SFF degradation from 50 real samples. (a) illustrates the relationship between the slope of optical flow and the slope of dark line. Red points represent the slope values on real samples. The curve is the inverse function: $y = -1/x$. Within the error range, this curve can fit these red points well, which indicates that the main orientation of deformation is roughly perpendicular to the dark line. (b) illustrates the relationship between the amplitude of optical flow and the radial distance of dark line. The horizontal axis is the radial distance from each point in the image to the dark line. Blue points represent the mean of amplitude and gray lines represent variance. It can be seen that the degree of deformation gradually attenuates away from the corruption.

less similar in ssTEM compared to that in common video due to low imaging resolution along the axial direction. Although SFF artifacts are accompanied by content deformation, there is still useful information in the degraded image that should not be discarded along with artifacts. To this end, we propose to synthesize degraded/groundtruth image pairs for training a restoration network, which enables the usage of information in the degraded image.

3 SFF Modeling

3.1 Statistical Analysis

In order to have a better understanding on the characteristics of SFF degradation, we collect 50 images with SFF artifacts from the Full Adult Fly Brain (FAFB) data [33]. Based on these samples, we statistically analyze the relationship between content deformation and dark line corruption, which is important for SFF simulation. We adopt optical flow to describe the deformation. A straightforward solution is to extract optical flow by using one of the adjacent images as reference. However, such a solution will suffer from intrinsic estimation error, due to low imaging resolution along the axial direction in ssTEM. Therefore, we first use the video frame interpolation method [23] to obtain an interpolation result from the adjacent images, and then extract optical flow between the degraded image and the interpolated one through a block-matching method [34]. The obtained results are then utilized for statistical analysis.

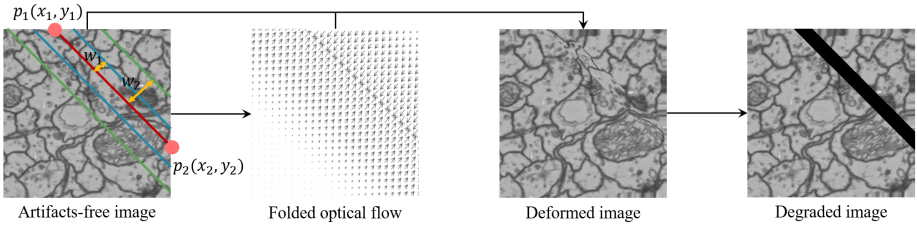


Fig. 3. Pipeline to simulate the SFF degradation process. Firstly, we generate the optical flow by using parameters randomly selected in fixed ranges, including two endpoints $p_1(x_1, y_1)$ and $p_2(x_2, y_2)$, the width of dark line w_1, and the width of actually corrupted region w_2. Then, based on the optical flow, we warp the artifacts-free image to obtain the deformed image. Finally, we add a straight black line on top of the deformed image to get the degraded result.

We take the dark line as the boundary and divide the degraded image into two parts for analysis, since the orientations of optical flows in the two parts are opposite. In order to alleviate the influence of estimation error, we only consider the main orientation, *i.e.*, the major component in the orientation histogram. We adopt the concept of slope in mathematics to describe the main orientation in a 2D coordinate system. We observe that the main orientations of the dark line and the extracted optical flow are roughly perpendicular to each other. In other words, the product of their slopes should be close to -1. Analysis on the 50 samples verify this observation, as shown in Fig. 2(a). In addition to the orientation information, we also analyze the relationship between the amplitude of optical flow and the radial distance to the dark line. As shown in Fig. 2(b), we find that the degree of deformation gradually attenuates away from the dark line. The above two observations on orientation and amplitude of content deformation serve as the basis of SFF simulation.

3.2 SFF Simulation

Realistic artifacts are generated during the acquisition process of ssTEM images, yet we cannot get the groundtruth corresponding to these degraded samples. It is thus difficult to directly train a deep restoration network which generally requires a large amount of paired samples (degraded/groundtruth). To address this issue, we propose an effective modeling strategy considering the aforementioned characteristics of SFF degradation, relying on which we can then synthesize a sufficient amount of paired samples to train a deep restoration network.

As demonstrated in Fig. 1(b), SFF degradation consists of two kinds of artifacts: corruption in the dark line and deformation of surrounding regions. Ideally, if we can unfold the dark line in the correct way, the deformation issue can be addressed and then filling up the corrupted content would not be difficult with the assistance of adjacent images. The challenge is that, since the folding effect during sample preparation is highly non-rigid, the deformation could vary in dif-

Algorithm 1: Folded optical flow synthesis.

Input: I: Artifacts-free image
$p_1(x_1, y_1)$ and $p_2(x_2, y_2)$: Two endpoints on two different boundaries of image I
w_1: Width of dark line
w_2: Width of actually corrupted region ($w_2 \geq w_1$)
α: Amplitude decay factor of optical flow
H: Height of image I
W: Width of image I

$p_1(x_1, y_1)$ and $p_2(x_2, y_2)$ determine the linear function of dark line l:
$y = k_l * x + b_l, k_l = (y_2 - y_1)/(x_2 - x_1), b_l = y_1 - k * x_1$.
The slope of optical flow: $k_f = -1/k_l$.
for $i \in [1, H]$ **do**
  for $j \in [1, W]$ **do**
    The distance from point (i, j) to the line l:
    $d(i, j) = |(k_l * i - j + b_l)/\sqrt{k_l^2 + 1}|$
    The amplitude of optical flow:
    $A(i, j) = \alpha * d(i, j) + b_f, b_f = (w_2 - w_1) - \alpha * w_2$
    if $A(i, j) < 0$ **then**
      $A(i, j) = 0$
    end
    $F(i, j, 1) = A(i, j) * \cos(\arctan(k_f))$
    $F(i, j, 2) = A(i, j) * \sin(\arctan(k_f))$
  endfor
endfor

Output: $F[H, W, 2]$: Folded optical flow.

ferent regions and the corrupted content is unpredictable. Therefore, accurately modeling SFF degradation in an analytical way is difficult. Instead, we simulate SFF degradation using the following strategy as shown in Fig. 3.

First, we collect a number of artifacts-free images from FAFB. According to the characteristics analyzed in Sec. 3.1, we produce simulated optical flows to deform the artifacts-free images (see Algorithm 1 for detailed implementation). Specifically, two endpoints $p_1(x_1, y_1)$ and $p_2(x_2, y_2)$ are randomly generated on two different boundaries of the image, which determine the slope of the dark line and thus the slope of optical flow (since the two are orthogonal). For the amplitude of the optical flow, we randomly assign the width of the dark line (w_1) and the width of actually corrupted region ($w_2, w_2 \geq w_1$). For simplicity, we only consider a linear decay relationship between the amplitude of the optical flow and the radial distance to the dark line.

Based on the above synthesized optical flows, the artifacts-free images are deformed correspondingly and the content falling in the corrupted region with width w_2 is masked by a black line with width w_1, as shown in Fig. 3. In practice, however, the realistic SFF artifacts may not satisfy the ideal corruption in the simulation process, *i.e.*, with a straight line shape and zero intensity values. To

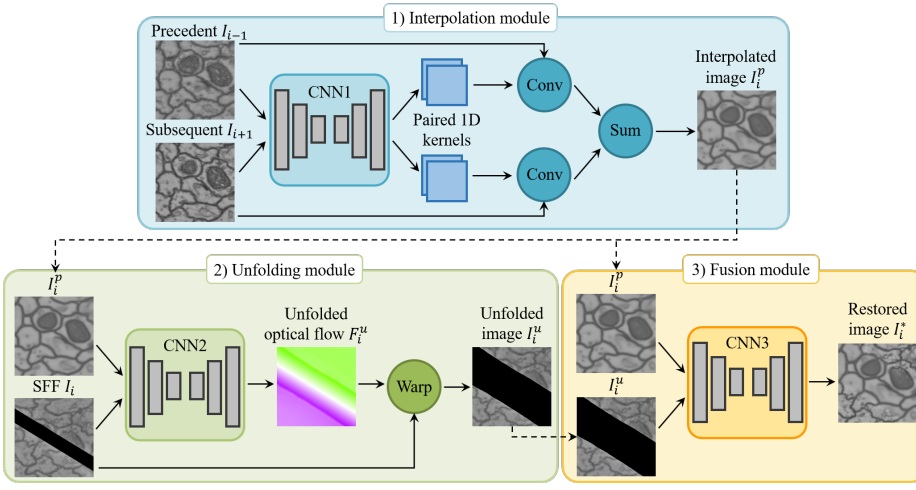


Fig. 4. The proposed restoration framework consists of three modules: interpolation, unfolding and fusion. Interpolation module takes the two adjacent images as input to obtain the interpolated result, which is used as the reference image in the next two modules. Unfolding module aims to address the content deformation in the degraded image. Fusion module is designed to fill up the corrupted content in the unfolded image. Detailed specifications of each module can be found in the supplementary material.

address this issue, we propose to cover a straight black line (with a slightly larger width) on top of the realistic dark line for regularization during the inference stage, which contributes to the restoration performance (see ablation study in Sec. 5.5). In addition to the optical flow for simulating the deformation process, we also generate the inverse optical flow for the unfolding purpose, which serves as the groundtruth in the unfolding module, as detailed in Sec. 4.2.

4 SFF Restoration

4.1 Artifacts Detection

In practice, automatic detection of SFF artifacts is highly desired, especially when handling a vast amount of samples from the whole brain. In this work, we adopt a simple yet effective line detector, *i.e.*, Hough transformation [6], to achieve automatic detection of the dark line corruption caused by SFF degradation [9,26]. Specifically, we perform Hough transformation in the binarized version of each image. For the binarization process, we traverse different threshold settings and find a suitable one in a validation dataset. Hough transformation can also localize the dark line, which enables regularization of corruption for a better restoration, as mentioned in Sec. 3.2.

4.2 Restoration Framework

The proposed restoration framework is composed of three modules, *i.e.*, interpolation, unfolding, and fusion, as shown in Fig. 4. On the one hand, the interpolation result is adopted as the reference for the optical-flow based unfolding, which addresses the deformation caused by SFF degradation. On the other hand, the interpolated image is fused with the unfolded image, which fills up the corrupted content in the dark line. For the implementation of each module, advanced network design may lead to a better performance. Yet, as the first attempt of learning-based restoration, we focus more on the realization of functions.

Interpolation module. Our interpolation module is built upon the kernel prediction network (KPN) [22], which is originally designed for the video frame interpolation task. Here, we view two adjacent images (I_{i-1} and I_{i+1}) as input and predict the corresponding per-pixel kernels to interpolate I_i^p in between. Specifically, the predicted kernels are applied to input images in a convolutional manner. Following the implementation of [23], we adopt two separable 1D kernels as a replacement of 2D kernel for efficiency. Details of the backbone structure can be found in the supplementary material.

Unfolding module. In this module, we adopt optical flow to represent the per-pixel position correspondence between a pair of input images. Based on the estimated optical flow, the deformation in the degraded image I_i can be addressed by a warping operation, as shown in Fig. 4. We term this alignment process as unfolding. We adopt a residual variant of U-Net [28] to implement the flow estimation network, where the degraded image I_i and the reference image I_i^p from the interpolation module form the input pair. Different from classical networks for optical flow estimation, our unfolding module can directly estimate optical flow at full image resolution, which eliminates the checkerboard effects caused by subsequent up-sampling. More specifically, the network is composed of four down-sampling and four up-sampling blocks. Each block contains three regular convolutional layers stacking with a skip connection.

Fusion module. Intuitively, to fill up the corrupted content in the unfolded image I_i^u from the unfolding module, one can directly crop the corresponding part in the interpolated image I_i^p and stitch it with the rest part of I_i^u. However, direct stitching is not optimal since the unfolded image may not be exactly aligned with the interpolated one. On the other hand, in addition to the content deformation and corruption, SFF degradation usually decreases the contrast of image. The above two issues call for an independent module to serve as an advanced solution for image fusion. In our implementation, a simplified version of U-Net [28] is adopted here, which addresses the potential misalignment and the low contrast issues for a better restoration performance (as detailed in Sec. 5.5).

Loss function. We train the above three modules independently and cascade them in the order of interpolation, unfolding, and fusion. In the proposed restoration framework, we adopt the L_1 distance between the input and the groundtruth as the loss function for each module. Specifically, given the simulated input with SFF artifacts, its artifacts-free version is adopted to supervise

the interpolation and fusion modules, and the simulated inverse optical flow is adopted to supervise the unfolding module.

5 Experiments

5.1 Data Preparation

In the training and validation phase, all data we use is from FAFB, the first ssTEM volume of a complete adult drosophila brain imaged at 4×4 nm resolution and sectioned at 40 nm thickness [33]. There are a total of 7062 sections in FAFB, and the original resolution of each section is 286720×155648 which is partitioned into 8192×8192 images, resulting in 40 TB data in storage. The SFF artifacts frequently occur in FAFB, which severely hinder the analysis and visualization on this valuable data. We select a central cube out of the raw volume as our main experimental data (approximately 150 GB) to demonstrate the effectiveness of the proposed method, yet the results could generalize to other portions of the volume. We further partition the selected cube into 512×512 images for easy manipulation. From this cube, we randomly select 4000 artifacts-free samples to generate the training data. Each sample contains three consecutive images, the middle image is used for the simulation of SFF degradation. In addition, we also select 100 samples as validation data for hyper-parameter tuning.

In the test phase, we use the public data from the CREMI challenge in 2016 [3] that aims to facilitate the neuron reconstruction in ssTEM images. The CREMI dataset is also for adult drosophila brain and with the same imaging resolution as FAFB. It consists of three subsets corresponding to different neuron types, each containing 125 images for training and 125 images for testing. Since the training images have manually obtained segmentation labels, we adopt them as the test images in our task to quantitatively evaluate the proposed method in terms of both image restoration quality and neuron segmentation accuracy. Specifically, we select a few samples out of an image bunch in each subset for SFF simulation, and the restored images are evaluated against the original ones for calculating quantitative metrics. Besides the synthetic test data used above, we also select a few number of real samples with SFF artifacts from FAFB (no overlap with the training and validation sample) for test, where only qualitative results are reported.

5.2 Implementation Details

During the statistical analysis of SFF degradation (in Sec. 3.1), we adopt a block matching method to generate a sparse optical flow between the degraded image and the interpolation result. The block size is set to 71×71 with an overlap of 41 pixels, and the maximum search range is 21 pixels. The estimated sparse optical flow is then upsampled to a dense one at the original resolution. During SFF simulation (in Sec. 3.2), there are three parameters, *e.g.*, the width of dark line w_1, the width of corrupted region w_2 and the decay factor of optical flow

Table 1. Quantitative comparisons of restoration results on synthetic data from CREMI dataset, in terms of both image restoration quality and neuron segmentation accuracy. For each subset, 25 out of 125 images are selected to synthesize SFF artifacts.

CREMI	Metric \ Method	SFF	Sub. [17]	PC [18]	PC-unfold	Interp [23]	Ours
A	PSNR ↑	15.57	19.57	16.35	18.01	22.70	**26.20**
	SSIM ↑	0.5615	0.4353	0.5665	0.7517	0.6595	**0.8261**
	FID ↓	229.71	33.30	36.83	62.44	144.69	**27.80**
	VOI ↓	2.6780	1.2505	2.6440	0.9507	0.8967	**0.7833**
	ARAND ↓	0.4971	0.2881	0.4783	0.1442	0.1518	**0.0968**
B	PSNR ↑	15.11	18.16	16.43	18.17	22.22	**26.81**
	SSIM ↑	0.5842	0.3586	0.6161	0.7532	0.6041	**0.8202**
	FID ↓	260.27	50.54	41.78	53.57	175.26	**38.17**
	VOI ↓	4.0629	3.8864	3.7147	3.4817	3.1898	**3.0957**
	ARAND ↓	0.5806	0.5855	0.3931	0.4355	0.3680	**0.3517**
C	PSNR ↑	14.52	17.74	15.09	16.97	21.96	**25.74**
	SSIM ↑	0.4988	0.3066	0.5037	0.7257	0.5766	**0.7957**
	FID ↓	335.38	42.70	44.60	73.33	168.13	**42.26**
	VOI ↓	4.5572	3.6755	4.5789	3.4882	3.1606	**3.0825**
	ARAND ↓	0.4244	0.4280	0.4308	0.3309	0.2835	**0.2789**
Inference time (s)		*	*	0.1235	0.5130	0.0075	0.4251

α. To increase the diversity of simulated samples, we set the three parameters to be within a dynamic range. The ranges of w_1, w_2 and α are $(5, 30)$, $(w_1, 80)$ and $(-0.1, -0.0001)$, respectively. During each iteration of the training process, we randomly generate a set of parameters to simulate SFF degradation on the training samples.

We adopt the same training setting for the three modules in our restoration framework. The resolution of input image to the network is set as 256, which is obtained by random cropping from the training samples. We train these networks using Adam [15] with $\beta_1 = 0.9$, $\beta_2 = 0.999$, a learning rate of 0.0001, and a batch size of 32 on four *NVIDIA Titan Xp* GPUs. We perform random rotation and flip in the training phase for data augmentation. In addition, we perform random contrast and brightness adjustment to address the contrast and brightness variation around the dark line. We adopt the early termination strategy to obtain an optimal model on the validation data. Each module requires about one day for training. In the inference phase, it takes $0.42s$ to process an image with 2048×2048 resolution.

5.3 Restoration on Synthetic Data

To quantitatively evaluate the proposed method, we first conduct experiments on three subsets from CREMI (termed as A, B, and C), where the degraded images

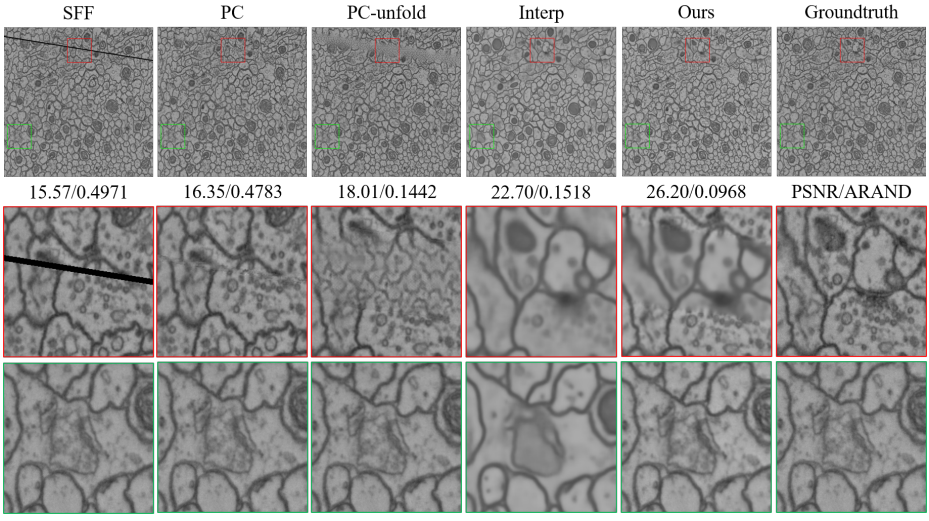


Fig. 5. Visual comparison of restoration results on *synthetic* test data. More visual comparison results on the synthetic test data can be found in the supplementary material.

(SFF) are synthesized from the artifacts-free ones. For each subset, we select 25 out of 125 samples for simulation. Four representative restoration strategies are adopted for comparison, including: (a) the substitution strategy (Sub.) which directly replaces the degraded image with the previous adjacent image, (b) the inpainting strategy which directly repairs the dark line, without consideration of the surrounding deformation. We adopt the state-of-the-art image inpainting method, *i.e.*, Partial Convolution (PC) [18], as a representative. We fine-tune the PC model pretrained on natural images with our simulated training samples. (c) Inpainting with unfolding (PC-unfold), where the PC model is fine-tuned with the unfolded images obtained from our proposed unfolding module. (d) The interpolation strategy (Interp) [23] which completely discards the degraded image and generates the result from two adjacent images.

As shown in Table 1, the quantitative results on the three subsets demonstrate that our proposed method outperforms the other restoration strategies by a large margin, in terms of image restoration quality. Besides two widely used fidelity metrics PSNR and SSIM, we also adopt a perceptual index for evaluation, *i.e.*, Fréchet Inception Distance (FID) [12], which is generated by computing the feature distance between the restored and groundtruth images. This perceptual index validates the superiority of the proposed method again.

As can be observed from the visual results in Fig. 5 (red box), the plain inpainting method PC only fills in the dark line but cannot address the surrounding deformation. Although the deformation can be largely corrected by PC-unfold with our unfolding module, it still cannot repair the complex neuron structure in the corrupted region only depending on the degraded image itself. On the

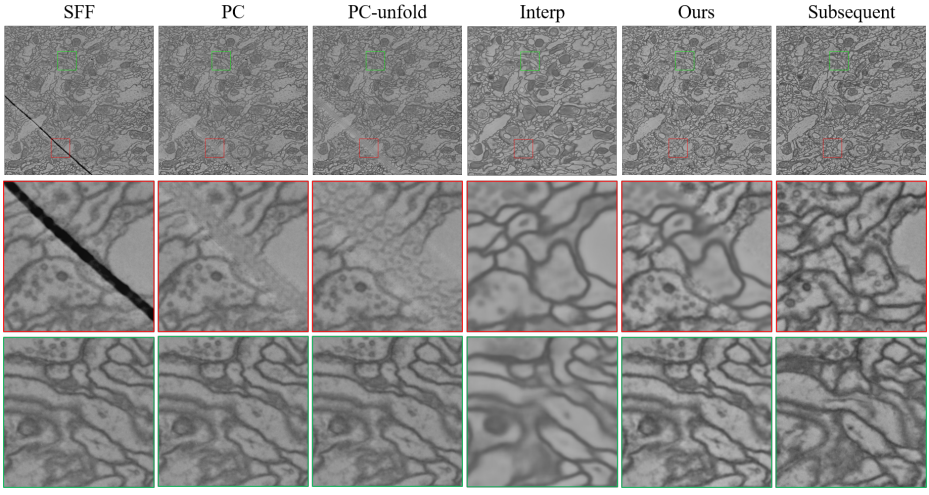


Fig. 6. Visual comparison of restoration results on *real* test data. More visual comparisons can be found in the supplementary material. Note that, due to the lack of groundtruth, we include the subsequent image to the degraded one as a reference here.

other hand, as can be seen from Fig. 5 (green box), although the interpolation result gets rid of misalignment with adjacent images, it loses fine details in the original degraded images away from the dark line. In contrast, these details are well preserved in our result, while both the corrupted and deformed regions are recovered.

To validate the effectiveness of our proposed restoration method in the subsequent analysis tasks, we also conduct segmentation experiments on the above restored images. To this end, we utilize a state-of-the-art neuron segmentation method [10], in which a 3D U-Net architecture is used to predict the affinity maps of ssTEM images and the final segmentation results are obtained after the operations of seeded watershed and agglomeration. We use two common metrics for evaluation of segmentation accuracy: adapted Rand error (ARAND) [1] and variation of information (VOI) [24]. As shown in Table 1, quantitative segmentation results demonstrate that our restoration method preserves more details that are useful for subsequent analysis tasks compared with other restoration strategies.

5.4 Restoration on Real Data

To evaluate the generalization capability of the proposed method, we further conduct experiments on real degraded images with SFF artifacts from FAFB. As shown in Fig. 6, our proposed method generalizes well to real data. Compared with baseline methods, the superior perceptual quality of recovered images demonstrates the advantage of our proposed method. In addition to perceptual quality, we also conduct segmentation experiments on the recovered images. Due

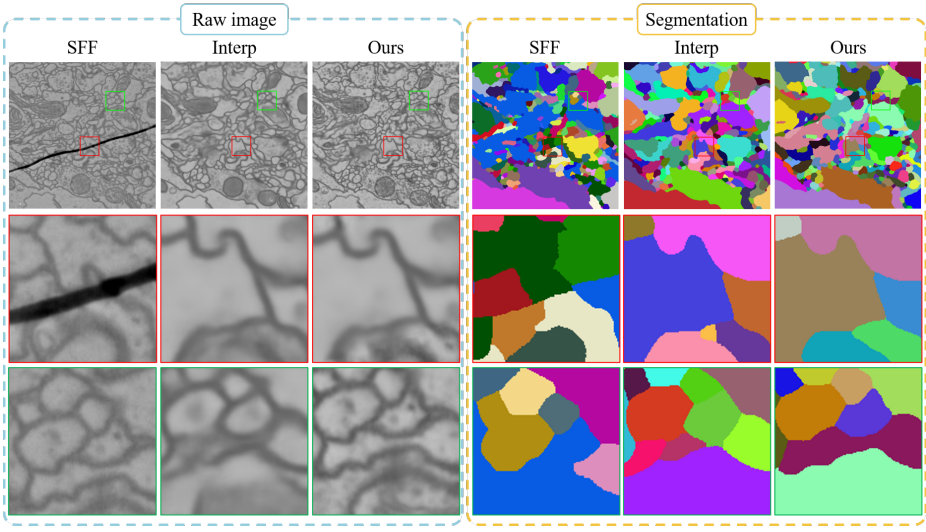


Fig. 7. Exemplar segmentation results on real data with SFF degradation. We adopt a representative segmentation method [10] for evaluation. Each pseudo color represents one neuron. Red box denotes a region near the dark line with severe deformation. Green box denotes a region away from the dark line with fine details.

Table 2. Ablation results for each module in the proposed restoration framework.

Interpolation	Unfolding	Fusion	PSNR	SSIM	FID	VOI	Rand
✓			22.70	0.6595	144.69	0.8967	0.1518
✓	✓		18.30	0.7393	115.55	0.9416	0.1430
✓		✓	24.68	0.7767	80.54	0.8301	0.1254
	✓	✓	25.94	0.8058	39.96	0.8534	0.1259
✓	✓	✓	**26.20**	**0.8261**	**27.80**	**0.7833**	**0.0968**

to the lack of groundtruth segmentation labels, we provide qualitative results on one exemplar case to demonstrate the superiority of our proposed method. As shown in Fig. 7 (red box), the interpolation method and our proposed method both obtain good segmentation results near the dark line. Nevertheless, our method preserves more details away from the dark line, which avoids merge errors introduced by interpolation, as shown in Fig. 7 (green box).

5.5 Ablation Study

Function of module. To verify the function of each module in our restoration framework, we conduct ablation experiments as shown in Table 2. Without the fusion module (Interpolation + Unfolding), the direct stitching cannot ensure the continuity of neuron structure. On the other hand, the issue of low contrast is left unsolved. Therefore, the PSNR score is obviously lower than other settings.

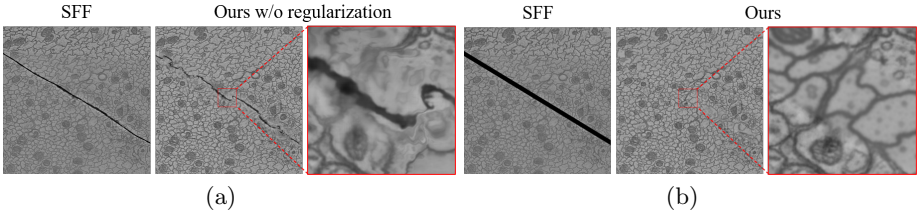


Fig. 8. A visual example to illustrate the effect of regularization of corruption. By covering a straight black line (b) on top of the realistic dark line in the original input (a), a significant improvement of performance is achieved.

Without the unfolding module (Interpolation + Fusion), the useful information in the surrounding area of the dark line cannot be utilized for fusion due to the deformation error. Without the interpolation module (Unfolding + Fusion), the corrupted content cannot be well recovered.

Regularization of corruption. As described in Sec. 3.2, based on the localization results of Hough transformation, the realistic dark line is automatically covered by a straight black line in the test phase. This regularization process is essential to the restoration performance, which bridges the gap between the synthetic corruption and the realistic one. We demonstrate the effectiveness of this regularization process in Fig. 8.

6 Conclusion

In this paper, we present the first learning-based restoration framework to address content deformation and corruption in ssTEM images, relying on effective modeling and simulation of SFF degradation. Evaluated on both synthetic and real test data, we demonstrate the advantage of our proposed method in terms of both image restoration quality and neuron segmentation accuracy. We believe the proposed method could benefit the future research of neuron morphology and connectomics using ssTEM images.

References

1. Arganda-Carreras, I., Turaga, S.C., Berger, D.R., Cireşan, D., Giusti, A., Gambardella, L.M., Schmidhuber, J., Laptev, D., Dwivedi, S., Buhmann, J.M., et al.: Crowdsourcing the creation of image segmentation algorithms for connectomics. Frontiers in neuroanatomy **9**, 142 (2015)
2. Bock, D.D., Lee, W.C.A., Kerlin, A.M., Andermann, M.L., Hood, G., Wetzel, A.W., Yurgenson, S., Soucy, E.R., Kim, H.S., Reid, R.C.: Network anatomy and in vivo physiology of visual cortical neurons. Nature **471**(7337), 177 (2011)

3. CREMI: Miccal challenge on circuit reconstruction from electron microscopy images. https://cremi.org/ (2016)
4. Denk, W., Horstmann, H.: Serial block-face scanning electron microscopy to reconstruct three-dimensional tissue nanostructure. PLoS biology **2**(11), e329 (2004)
5. Dosovitskiy, A., Fischer, P., Ilg, E., Hausser, P., Hazirbas, C., Golkov, V., Van Der Smagt, P., Cremers, D., Brox, T.: Flownet: Learning optical flow with convolutional networks. In: Proceedings of the IEEE international conference on computer vision. pp. 2758–2766 (2015)
6. Duda, R.O., Hart, P.E.: Use of the hough transformation to detect lines and curves in pictures. Communications of the ACM **15**(1), 11–15 (1972)
7. Eberle, A., Mikula, S., Schalek, R., Lichtman, J., Tate, M.K., Zeidler, D.: High-resolution, high-throughput imaging with a multibeam scanning electron microscope. Journal of microscopy **259**(2), 114–120 (2015)
8. Feist, A., Echternkamp, K.E., Schauss, J., Yalunin, S.V., Schäfer, S., Ropers, C.: Quantum coherent optical phase modulation in an ultrafast transmission electron microscope. Nature **521**(7551), 200 (2015)
9. Funke, J.: Automatic Neuron Reconstruction from Anisotropic Electron Microscopy Volumes. Ph.D. thesis, ETH Zurich (2014)
10. Funke, J., Tschopp, F., Grisaitis, W., Sheridan, A., Singh, C., Saalfeld, S., Turaga, S.C.: Large scale image segmentation with structured loss based deep learning for connectome reconstruction. IEEE transactions on pattern analysis and machine intelligence **41**(7), 1669–1680 (2018)
11. Harris, K.M., Perry, E., Bourne, J., Feinberg, M., Ostroff, L., Hurlburt, J.: Uniform serial sectioning for transmission electron microscopy. Journal of Neuroscience **26**(47), 12101–12103 (2006)
12. Heusel, M., Ramsauer, H., Unterthiner, T., Nessler, B., Hochreiter, S.: Gans trained by a two time-scale update rule converge to a local nash equilibrium. In: Advances in neural information processing systems. pp. 6626–6637 (2017)
13. Ilg, E., Mayer, N., Saikia, T., Keuper, M., Dosovitskiy, A., Brox, T.: Flownet 2.0: Evolution of optical flow estimation with deep networks. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 2462–2470 (2017)
14. Khalilian-Gourtani, A., Tepper, M., Minden, V., Chklovskii, D.B.: Strip the stripes: Artifact detection and removal for scanning electron microscopy imaging. In: ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). pp. 1060–1064. IEEE (2019)
15. Kingma, D.P., Ba, J.: Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980 (2014)
16. Knott, G., Marchman, H., Wall, D., Lich, B.: Serial section scanning electron microscopy of adult brain tissue using focused ion beam milling. Journal of Neuroscience **28**(12), 2959–2964 (2008)
17. Li, P.H., Lindsey, L.F., Januszewski, M., Zheng, Z., Bates, A.S., Taisz, I., Tyka, M., Nichols, M., Li, F., Perlman, E., et al.: Automated reconstruction of a serial-section em drosophila brain with flood-filling networks and local realignment. bioRxiv p. 605634 (2019)
18. Liu, G., Reda, F.A., Shih, K.J., Wang, T.C., Tao, A., Catanzaro, B.: Image inpainting for irregular holes using partial convolutions. In: Proceedings of the European Conference on Computer Vision (ECCV). pp. 85–100 (2018)
19. Maraghechi, S., Hoefnagels, J., Peerlings, R., Rokoš, O., Geers, M.: Correction of scanning electron microscope imaging artifacts in a novel digital image correlation framework. Experimental mechanics **59**(4), 489–516 (2019)

20. Minh Quan, T., Grant Colburn Hildebrand, D., Lee, K., Thomas, L.A., Kuan, A.T., Allen Lee, W.C., Jeong, W.K.: Removing imaging artifacts in electron microscopy using an asymmetrically cyclic adversarial network without paired training data. In: Proceedings of the IEEE International Conference on Computer Vision Workshops. pp. 0–0 (2019)
21. Nazeri, K., Ng, E., Joseph, T., Qureshi, F., Ebrahimi, M.: Edgeconnect: Structure guided image inpainting using edge prediction. In: Proceedings of the IEEE International Conference on Computer Vision Workshops. pp. 0–0 (2019)
22. Niklaus, S., Mai, L., Liu, F.: Video frame interpolation via adaptive convolution. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. pp. 670–679 (2017)
23. Niklaus, S., Mai, L., Liu, F.: Video frame interpolation via adaptive separable convolution. In: Proceedings of the IEEE International Conference on Computer Vision. pp. 261–270 (2017)
24. Nunez-Iglesias, J., Ryan Kennedy, T.P., Shi, J., Chklovskii, D.B.: Machine learning of hierarchical clustering to segment 2d and 3d images. PloS one **8**(8) (2013)
25. Pathak, D., Krahenbuhl, P., Donahue, J., Darrell, T., Efros, A.A.: Context encoders: Feature learning by inpainting. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 2536–2544 (2016)
26. Popovych, S., Bae, J.A., Seung, H.S.: Caesar: Segment-wise alignment method for solving discontinuous deformations. In: 2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI). pp. 1214–1218. IEEE (2020)
27. Roels, J., Aelterman, J., Luong, H., Lippens, S., Pižurica, A., Saeys, Y., Philips, W.: An overview of state-of-the-art image restoration in electron microscopy. Journal of microscopy **271**(3), 239–254 (2018)
28. 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)
29. Schorb, M., Haberbosch, I., Hagen, W.J., Schwab, Y., Mastronarde, D.N.: Software tools for automated transmission electron microscopy. Nature methods **16**(6), 471–477 (2019)
30. Tapia, J.C., Kasthuri, N., Hayworth, K.J., Schalek, R., Lichtman, J.W., Smith, S.J., Buchanan, J.: High-contrast en bloc staining of neuronal tissue for field emission scanning electron microscopy. Nature protocols **7**(2), 193 (2012)
31. Yang, C., Lu, X., Lin, Z., Shechtman, E., Wang, O., Li, H.: High-resolution image inpainting using multi-scale neural patch synthesis. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. pp. 6721–6729 (2017)
32. Zeng, Y., Fu, J., Chao, H., Guo, B.: Learning pyramid-context encoder network for high-quality image inpainting. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. pp. 1486–1494 (2019)
33. Zheng, Z., Lauritzen, J.S., Perlman, E., Robinson, C.G., Nichols, M., Milkie, D., Torrens, O., Price, J., Fisher, C.B., Sharifi, N., et al.: A complete electron microscopy volume of the brain of adult drosophila melanogaster. Cell **174**(3), 730–743 (2018)
34. Zhu, S., Ma, K.K.: A new diamond search algorithm for fast block-matching motion estimation. IEEE transactions on Image Processing **9**(2), 287–290 (2000)