

EMGauss: Continuous Slice-to-3D Reconstruction via Dynamic Gaussian Modeling in Volume Electron Microscopy

Supplementary Material

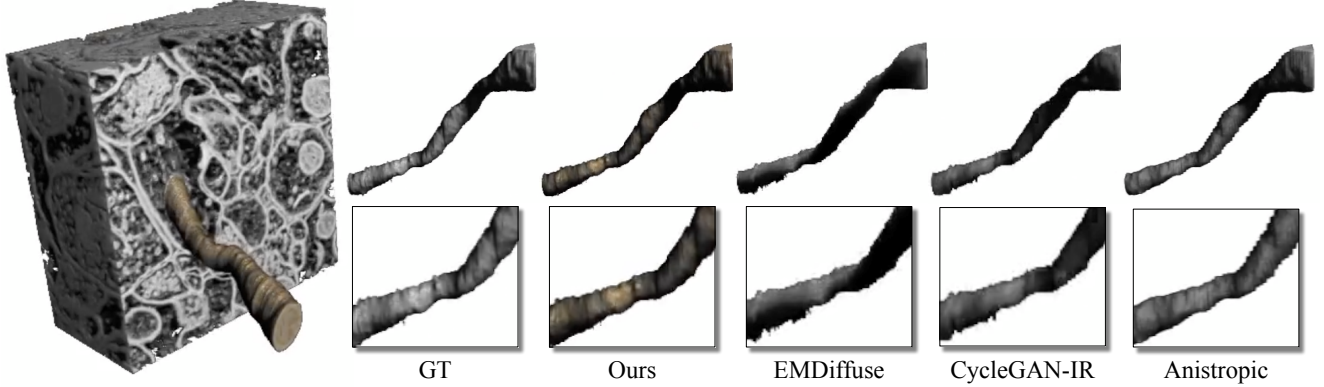


Figure 6. 3D reconstruction results on the interpolated EPFL dataset. We apply SAM2 [5] on the interpolation outputs of different isotropic reconstruction models to generate the mask of a cellular structure across continuous xy slices, and use DragonFly Software for 3D reconstruction. We provide visualizations of the full reconstruction process and comparisons in the *supplementary video*.

A. Segmentation and 3D Reconstruction of Isotropic-reconstructed Volume

Isotropic slice interpolation is not only an image enhancement task but also a prerequisite for accurate 3D reconstruction in volume electron microscopy (vEM). Severe axial anisotropy in typical acquisition pipelines hampers downstream analysis such as segmentation and structural reconstruction. High-fidelity slice interpolation is therefore crucial for restoring geometric continuity and enabling volumetric reconstructions that more faithfully reflect underlying biological structures.

To evaluate the practical utility of our method, we apply the reconstructed isotropic volumes to downstream volumetric analysis. Following the protocol of prior work [3], we conduct video-based segmentation on the interpolated slice sequences and use the resulting masks to drive 3D reconstruction. Specifically, we adopt Segment Anything Model 2 (SAM2) [5] to segment cellular structures across continuous xy slices. With ground-truth intermediate slices available, we compute their segmentation masks and measure the Intersection-over-Union (IoU) between predictions and ground truth to assess segmentation accuracy.

We provide 3D reconstruction visualizations and comparisons in the *supplementary video* (Figure 6). These experiments collectively demonstrate that high-quality anisotropy correction is crucial for downstream 3D vEM analysis, and that our reconstruction pipeline yields more coherent structural continuity and higher morphological fidelity than existing methods.

B. More Ablation Studies

Different anisotropy ratios. To assess the robustness and generality of our method, we further evaluate isotropic reconstruction under synthetic anisotropy factors of $\times 4$ and $\times 8$ on both the EPFL and FIB-25 datasets. The full quantitative results are reported in Table 5 and Table 6. Across all settings, our approach consistently outperforms prior methods, with particularly notable gains under the challenging $\times 8$ anisotropy where preserving structural continuity becomes substantially harder. These results indicate that our deformable 2D Gaussian formulation remains effective even when axial information is highly sparse, demonstrating strong generalization across varying anisotropy levels.

Number of training images within a single reconstruction process. For isotropic reconstruction under an anisotropy factor of $\times 6$, we group every 30 consecutive axial slices as a training bundle (as summarized in Table 2), using only 6 uniformly spaced slices (indices 0, 6, 12, 18, 24, 30) as inputs while treating the remaining slices as interpolation targets. Each bundle is fitted with an independent deformable 2DGS model that captures the local slice-to-slice geometric evolution. To assess the effect of the number of training slices, we conduct a sensitivity analysis. As shown in Table 7, reconstruction quality remains stable across different choices, indicating that our method is robust to the choice of training image count. In practice, we adopt 6 training slices per bundle, which offers a favorable balance between computational cost and memory usage without sacrificing performance.

Table 5. *Results of isotropic reconstruction on EPFL synthetic anisotropic datasets under anisotropy ratios of $\times 4$ and $\times 8$.*
[†]EMDiffuse uses additional data for training.

Method	EPFL					
	$\times 4$			$\times 8$		
	PSNR	SSIM	FSIM	PSNR	SSIM	FSIM
CycleGAN-IR[1]	24.39	0.5657	0.9063	20.62	0.4239	0.8269
EMDiffuse[3] [†]	23.53	0.5649	0.9002	21.82	0.4466	0.8741
Ours	27.03	0.7052	0.9513	25.37	0.5811	0.9041

Table 6. *Results of isotropic reconstruction on FIB-25 synthetic anisotropic datasets under anisotropy ratios of $\times 4$ and $\times 8$.*
[†]EMDiffuse uses additional data for training.

Method	FIB-25					
	$\times 4$			$\times 8$		
	PSNR	SSIM	FSIM	PSNR	SSIM	FSIM
CycleGAN-IR[1]	27.06	0.6704	0.9113	23.76	0.5516	0.8426
EMDiffuse[3] [†]	27.54	0.6789	0.9295	23.58	0.5096	0.8500
Ours	28.15	0.7313	0.9331	26.09	0.5943	0.8974

Table 7. *Sensitivity analysis on the number of training images within one training process.* The results are averaged over the two isotropic datasets. The results indicate that the proposed framework exhibits strong robustness with respect to the choice of the number of training slices.

No. of training images	4	5	6	7
PSNR	26.92	26.82	26.98	26.90
SSIM	0.7120	0.7131	0.7126	0.7133
FSIM	0.9322	0.9340	0.9315	0.9305

C. More Results on Isotropic Reconstruction

Cubic Interpolation yields over-smooth results. Cubic interpolation relies on fixed, analytically defined polynomial kernels that perform local weighted averaging across neighboring slices, yet it lacks any mechanism for capturing the underlying morphological evolution along the axial direction. This formulation inherently imposes strong smoothness priors that bias the reconstruction toward low-frequency components, suppressing high-frequency components that are essential for delineating ultrastructural details. Consequently, although cubic interpolation may attain moderate PSNR, its predictions exhibit pronounced blurring and attenuation of membrane contrast, ultimately limiting their suitability for analyses that require precise delineation of nanoscale structures. As illustrated in Fig. 7, cubic interpolation produces noticeably smoother boundaries and markedly reduced detail in features such as vesicular patterns and thin neurites. These artifacts reflect the fundamental limitations of purely interpolation-based approaches when applied to complex biological volumes. In contrast, our method explicitly models volumetric deformation and structural evolution across slices, enabling it to preserve fine-scale morphology and more faithfully reflect the underlying biological architecture.

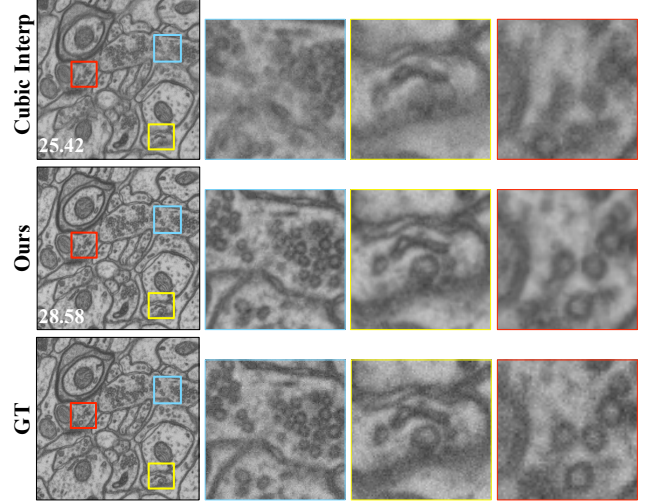


Figure 7. *Comparison of cubic interpolation and our methods on EPFL dataset with an anisotropy ratio of 6.* Cubic interpolation enforces excessive smoothness, leading to blurred membranes and loss of fine ultrastructural details, whereas our method better preserves high-frequency morphology and aligns more closely with the ground truth.

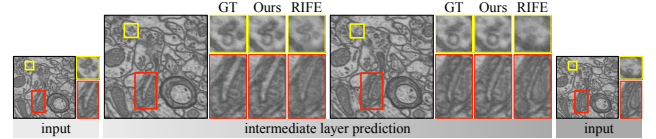


Figure 8. *Comparison with the VFI method RIFE [2].* Given two adjacent input slices, each method predicts the intermediate slice. Our method reconstructs sharper and more anatomically consistent structures, while RIFE produces oversmoothed results (see zoomed regions).

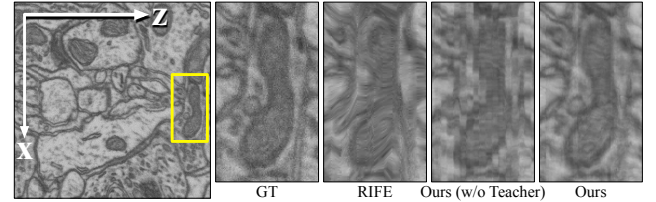


Figure 9. *xz-plane visualization of reconstructed volumes.* Our method produces more coherent axial structures and smoother cross-slice transitions than RIFE [2] and the ablated variant without the teacher module.

Comparison with video frame interpolation (VFI). Conventional video frame interpolation (VFI) methods are typically Eulerian, operating on fixed pixel grids, which often leads to blurred bifurcations and oversmoothing when modeling structural appearance and disappearance across slices. In contrast, EMGauss adopts a Lagrangian representation that lifts pixels to deformable Gaussian primitives with independent trajectories, enabling more natural handling of structural deformation and topological changes.

Moreover, unlike VFI methods that rely on discrete indexing and heavy pretraining (e.g., RIFE [2] trained on Vimeo90K), EMGauss supports continuous slice synthesis, accommodates irregular sampling, and is trained in a self-supervised manner directly on the target volume, making it particularly suitable for data-scarce biomedical scenarios. Empirically, we compare our method with RIFE [2] fine-tuned on our training data across two datasets. As illustrated in Figure 8, our method produces sharper and more anatomically consistent intermediate slices than RIFE, which often yields oversmoothed structures. The xz -plane visualization in Figure 9 further shows that our approach better preserves axial structural continuity across depth.

Interpolation results on FANC datasets. We evaluate the interpolation performance of different methods on a real-captured anisotropic volume EM dataset, FANC dataset [4]. As illustrated in Figure 10, given two adjacent training slices, each method is tasked with reconstructing the isotropic ($\times 10$) intermediate slices. Our approach consistently produces the most anatomically coherent results, exhibiting smooth and biologically plausible morphological transitions across depth. In contrast, CycleGAN-IR [1] exhibits noticeable instability during training. It exhibits contrast shifts and global intensity fluctuations, and it also produces temporal inconsistencies where the predicted slice is discontinuous from the subsequent ground-truth slice. While EMDiffuse [3] can capture scene-level variation, it often explains cross-slice differences through opacity modulation rather than geometric deformation (highlighted by the red boxes). This results in flickering or abrupt appearance and disappearance of structures, rather than biologically meaningful motion. Overall, the visualizations demonstrate that our deformable Gaussian formulation yields substantially more faithful slice interpolation and provides reconstructions that are better aligned with the needs of downstream volumetric analysis.

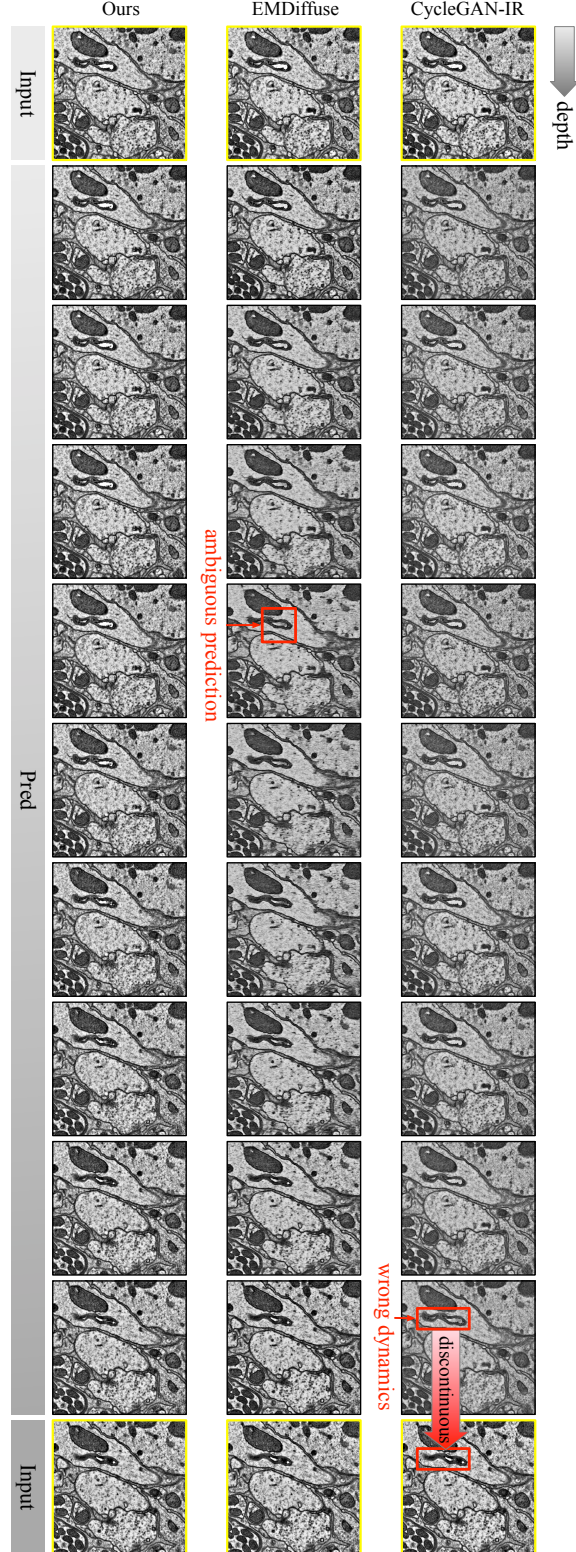


Figure 10. *Isotropic reconstruction sequence on real-captured anisotropic FANC dataset with an anisotropic factor of $10\times$.* The first and last slices (highlighted in yellow) are the ground-truth inputs, and the middle rows show the interpolated slices produced by each method. CycleGAN-IR exhibits temporal inconsistency, where its predicted slice becomes discontinuous with the next slice (the input slice). EMDiffuse suffers from modeling cross-slice variation through opacity modulation rather than geometric deformation. In contrast, our method yields consistent slice-to-slice transitions and preserves anatomically meaningful continuity throughout the sequence.

References

- [1] Shiyu Deng, Xueyang Fu, Zhiwei Xiong, Chang Chen, Dong Liu, Xuejin Chen, Qing Ling, and Feng Wu. Isotropic reconstruction of 3d em images with unsupervised degradation learning. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 163–173. Springer, 2020. [2](#), [3](#)
- [2] Zhewei Huang, Tianyuan Zhang, Wen Heng, Boxin Shi, and Shuchang Zhou. Real-time intermediate flow estimation for video frame interpolation. In *European conference on computer vision*, pages 624–642. Springer, 2022. [2](#), [3](#)
- [3] Chixiang Lu, Kai Chen, Heng Qiu, Xiaojun Chen, Gu Chen, Xiaojuan Qi, and Haibo Jiang. Diffusion-based deep learning method for augmenting ultrastructural imaging and volume electron microscopy. *Nature Communications*, 15(1):4677, 2024. [1](#), [2](#), [3](#)
- [4] Jasper S Phelps, David Grant Colburn Hildebrand, Brett J Graham, Aaron T Kuan, Logan A Thomas, Tri M Nguyen, Julia Buhmann, Anthony W Azevedo, Anne Sustar, Sweta Agrawal, et al. Reconstruction of motor control circuits in adult drosophila using automated transmission electron microscopy. *Cell*, 184(3):759–774, 2021. [3](#)
- [5] Nikhila Ravi, Valentin Gabeur, Yuan-Ting Hu, Ronghang Hu, Chaitanya Ryali, Tengyu Ma, Haitham Khedr, Roman Rädle, Chloe Rolland, Laura Gustafson, Eric Mintun, Junting Pan, Kalyan Vasudev Alwala, Nicolas Carion, Chao-Yuan Wu, Ross Girshick, Piotr Dollár, and Christoph Feichtenhofer. Sam 2: Segment anything in images and videos. *arXiv preprint arXiv:2408.00714*, 2024. [1](#)