

Tractography-Driven Synthetic Data Generation for Fiber Bundle Segmentation in Tracer Histology

Kyriaki-Margarita Bintsi¹, Sparsh Makharia², Yaël Balbastre³, Joselyn Romero Avila⁴, Julia F. Lehman⁵, Suzanne N. Haber^{5,6}, and Anastasia Yendiki¹

¹ Athinoula A. Martinos Center for Biomedical Imaging, Massachusetts General Hospital and Harvard Medical School, Charlestown, MA, United States
kbintsi@mgh.harvard.edu

² Krea University, Sri City, Andhra Pradesh, India

³ Department of Experimental Psychology, University College London, London, United Kingdom

⁴ Universidad Nacional Mayor de San Marcos, Lima, Peru

⁵ Department of Pharmacology and Physiology, University of Rochester School of Medicine, Rochester, NY, United States

⁶ McLean Hospital, Belmont, MA, United States

Abstract. Diffusion MRI (dMRI) tractography enables non-invasive reconstruction of white-matter pathways, but its accuracy is fundamentally limited by indirect, low-resolution measurements of axonal organization. Tracer injection studies in non-human primates provide a gold standard for validating dMRI tractography. This, however, requires time-consuming manual annotation of fiber bundles in histology sections. We propose a synthetic-data augmented framework for automated fiber bundle segmentation in macaque tracer histology. Our approach uses ex vivo dMRI tractography as a generative prior to synthesize 2D image patches for training. This provides us with sufficiently realistic foreground texture, which we compose with backgrounds from blockface photos and diversify via domain randomization. A 2D U-Net is trained on mixed real and synthetic patches. Experiments on held-out brains demonstrate improved generalization across brains and fiber bundle densities compared to training with real data only. Training with synthetic data only leads to poor performance, underscoring the need for real supervision. Overall, our approach achieves performance comparable to the state-of-the-art while requiring 3x less manually annotated data.

Keywords: Synthetic data generation · domain randomization · fiber bundle segmentation.

1 Introduction

Diffusion MRI (dMRI) tractography enables non-invasive reconstruction of white matter pathways in health and disease [3,26]. However, it uses indirect measurements of axonal orientations based on water diffusion and has millimeter-scale resolution, orders of magnitude coarser than individual axons. Anatomic

tracer studies in non-human primates (NHPs) offer a gold standard for validating tractography [20,27], as they allow direct visualization of the complete trajectory of axonal projections at single axon resolution [21,11]. Yet reconstructing fiber pathways from tracer histology requires labor-intensive manual annotation, severely limiting the amount of tracer data available for validation studies.

Machine learning can automate this process, but progress has been constrained by labeled data scarcity and variability across tracers, injection sites, and brains. Methods for axon segmentation in electron microscopy [23,15] or fluorescence microscopy in rodents and marmosets [24,7,25] do not generalize to macaque tracer data, which are particularly valuable for translational studies due to known human-macaque homologies [4]. Initial efforts to deploy deep learning for segmenting fiber bundles in macaque tracer data [22] demonstrated feasibility but relied on *ad hoc* post-processing steps to reduce false positives. Recent work employed semi-supervised pre-training to improve accuracy and generalizability to sparse bundles, while obviating the need for post-processing [2], but still required manual annotations from multiple brains.

Synthetic training data, i.e., images generated from ground-truth segmentation labels, offer a promising strategy for reducing the burden of manually labeling real data. Domain randomization, i.e., systematically varying geometry, appearance, and noise during synthesis, has proven effective in encouraging models to learn structural patterns robust to dataset-specific artifacts, and has found applications across biomedical imaging domains [16,9,5]. In this approach, a network trained on synthetic data can learn to segment real data, not because the synthetic data are realistic, but because they are very diverse. To date, this strategy has not been explored for segmenting tracer histology.

Here, we present a synthetic-data augmented framework for fiber bundle segmentation in macaque tracer histology. Our synthesis framework aims to replicate the key feature of real tracer data, where a tracer labels axons projecting from a single injection site in the gray matter, as they leave the injection site, split into bundles of varying densities, and follow different paths through the white matter to reach their destinations. We propose to leverage dMRI tractography as a geometric prior to synthesize image patches with this appearance. Importantly, we do not need the end-to-end trajectory of the tractography streamlines to be anatomically correct, as we only use them to generate image patches. We compose the streamlines with backgrounds from blockface photos, diversified via domain randomization. Combined with limited real annotated sections, this achieves strong cross-brain generalization with substantially reduced annotation effort. This paper has the following contributions: (1) A tractography-driven synthesis pipeline with controllable domain randomization. (2) A mixed real+synthetic training strategy matching state-of-the-art [2] performance with $3\times$ less annotated data. (3) Ablations on the amount of real data, the synthetic-to-real data ratio, and randomization strength, showing that even a single annotated section yields competitive performance. The code is publicly available at <https://github.com/lincbrain/synthetic-tracer-histology-segmentation>.

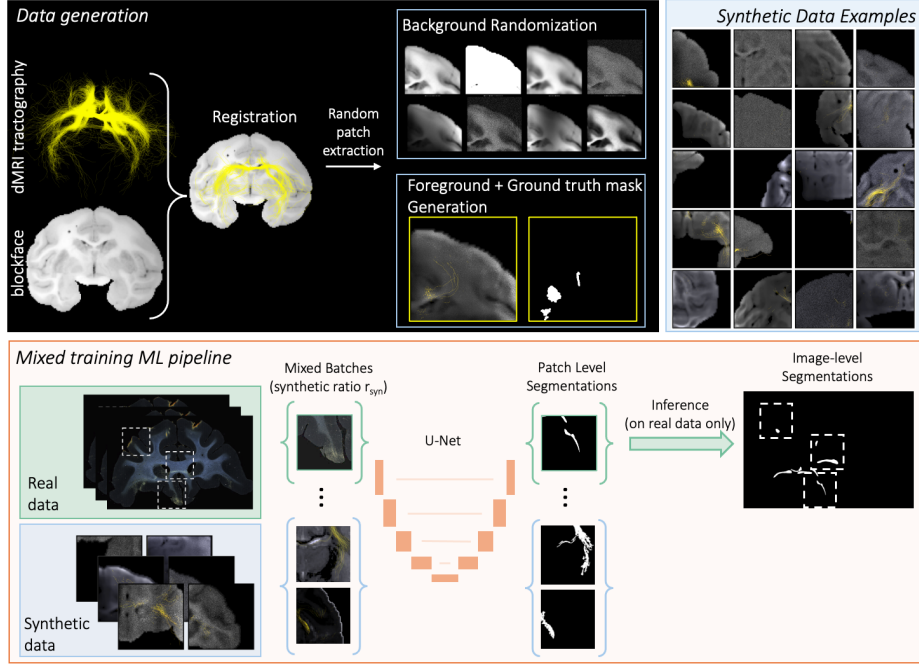


Fig. 1. Overview of the proposed pipeline. **Top:** Ex vivo dMRI is registered to a blockface image from the same brain, and tractography streamlines from a cortical “injection site” are mapped into blockface space. We use subsets of these streamlines with varying fiber densities as the foreground, and compose them with randomized backgrounds to produce synthetic image-mask pairs. **Bottom:** Real histology patches and synthetic patches are mixed in each mini-batch with synthetic ratio r_{syn} to train a 2D U-Net. At inference, only real images are processed, and overlapping patch predictions are merged into full-section bundle segmentations.

2 Methods

2.1 Synthetic Data Generation

We design an on-the-fly synthesis pipeline that generates realistic image-mask pairs from a single ex vivo macaque brain (M0) for which both whole-brain dMRI and co-registered blockface photographs are available. The dMRI data were acquired on a 4.7T preclinical scanner using a 3D echo-planar sequence (0.5 mm isotropic, 515 gradient directions, $b_{\text{max}} = 40,000 \text{ s/mm}^2$), comprising three b-shells ($b = 4,000, 8,000, 12,000 \text{ s/mm}^2$) with 64, 64, and 128 gradient directions respectively, following the protocol of [12]. We used multi-shell multi-tissue constrained spherical deconvolution [6] implemented in MRtrix3 and seeded probabilistic tractography (iFOD2 algorithm, maximum bending angle 30° , minimum FOD amplitude 0.1, all other parameters at default) 10,000 times per voxel, repeated for 96 voxels distributed throughout the cortex to mimic different “in-

jection sites”, yielding anatomically diverse fiber configurations that effectively decouple the synthetic training distribution from any single brain’s anatomy.

After sectioning, blockface photographs were captured at 16–21 μm in-plane resolution and the stacked volume was registered deformably to the $b=0$ dMRI volume using ANTs [1]. Let $B \in \mathbb{R}^{W \times D \times H}$ and $\mathcal{S} = \{s_1, \dots, s_N\}$ denote the blockface volume and tractography streamlines transformed into blockface space, respectively. Each synthetic sample $(\hat{I}, \hat{M})$ is constructed as follows.

Background. To bridge the domain gap between blockface photography and dark-field microscopy, we sample a slice index $y^* \sim \mathcal{U}\{1, \dots, D\}$ and extract the coronal section $B_{y^*}(x, z) = B(x, y^*, z)$. We apply mild CLAHE, a power-law dark-field transform with $\gamma \sim \mathcal{U}(0.8, 1.2)$ restricted to the brain mask, and a dark-field colormap [18], to emulate the appearance of a brain section. To further bridge the domain gap, the image is processed with augmentation presets [5] sampled with weighted probability, spanning optical transforms, parametric noise models, smooth multiplicative bias fields, and Lab color perturbations. Finally, we sample a random $P \times P$ crop ($P = 512$).

Foreground (synthetic tracer signal). Given the slice plane y^* , we select streamlines originating from the current “injection site” that pass within distance τ of the plane:

$$\mathcal{S}_{y^*} = \{s_i \in \mathcal{S} : \min_{(p_x, p_y, p_z) \in s_i} |p_y - y^*| \leq \tau\}, \quad (1)$$

where $(p_x, p_y, p_z) \in s_i$ denotes a 3D point along streamline s_i , p_y its y -coordinate (coronal axis), and τ is an adaptive thickness parameter scaled with output resolution. Points on $\mathcal{S}_{y^*}$ are densified via linear interpolation at 0.2-voxel spacing and projected onto the coronal plane. We then sample a fraction $\rho_{\max} \sim \mathcal{U}(0.02, 0.05)$ of streamlines, from the densely seeded tractography streamlines, and subsample further by $r_{\text{viz}} \sim \mathcal{U}(0.5, 0.7)$ for rendering. Each streamline is rendered as an anti-aliased curve with stochastic linewidth $w \sim \mathcal{U}(0.5, 1.5)$ px, opacity fading linearly from $\alpha = 0.4$ at y^* to 0 at distance δ , and per-streamline color jitter in the orange–yellow range to mimic tracer appearance. Optionally ($p_{\text{inj}} = 0.3$), injection-site artifacts are added by rendering short curved streamlets radiating from a random point with saturated orange hues. The background crop and rendered tracer signal are composited to form the synthetic image $\hat{I}$.

Ground-truth mask. These binary masks mimic the appearance of fiber bundle annotations. They are generated from a dense streamline set, so that $\hat{M}$ captures the full spatial extent of each bundle even when the synthesized tracer signal is sparse. We accumulate a density map $D(x, z)$ by rasterizing line segments between consecutive projected streamline points (Bresenham). We then Gaussian-smooth ($\sigma=2.0$ px) and threshold using an absolute count threshold $\theta=0.08$:

$$\hat{M}_{\text{raw}}(x, z) = \mathbf{1} \left[(G_\sigma * D)(x, z) \geq \theta \right]. \quad (2)$$

Post-processing applies per-component dilation within a disk (to expand bundles while preserving separation) and small-object removal. This decoupling, sparse

fibers in $\hat{I}$ but dense bundles in $\hat{M}$, encourages the model to detect bundles even when only partially visible, reflecting real tracer histology.

On-the-fly generation and domain randomization. Samples are synthesized dynamically during training: each mini-batch randomly selects a seed voxel, slice index, and patch location, rendering the corresponding sample through the full pipeline and providing large diversity without storage overhead (~ 4 s/patch on CPU). Synthesis parameters were selected empirically via visual inspection and prior knowledge of tracer appearance. Beyond the in-pipeline transforms, we apply a second augmentation stage during training that includes spatial transforms (flips, rotations $\pm 20^\circ$, elastic deformations), appearance transforms (color jitter, random gamma, hue-saturation shifts), and stochastic degradations (Gaussian/ISO noise, mild blur).

2.2 Mixed Training Pipeline

An overview of the pipeline is shown in Figure 1. Each training batch combines real and synthetic samples at ratio r_{syn} . Real patches use foreground-aware sampling (half centered on fiber pixels, half at random locations) to address class imbalance. At inference, overlapping 1024×1024 patches are aggregated by averaging to produce full-section probability maps, followed by connected component filtering to remove small spurious detections smaller than 800 pixels..

3 Experiments

Datasets. We use high-resolution coronal histological sections from four in-house macaque datasets (M1–M4) with bidirectional tracer injections at multiple cortical sites. Sections ($50 \mu\text{m}$ thick) were digitized at $0.4 \mu\text{m}$ in-plane resolution, with every 24th section processed to visualize a specific tracer (1.2 mm effective inter-slice spacing). This sampling strategy follows previous tracer studies in non-human primates [13,10]. Expert neuroanatomists annotated fiber bundles (i.e. groups of axons traveling in close proximity from injection sites to terminal fields) into three density classes: dense, moderate, and sparse. Approximately 30 sections per brain were annotated, yielding ~ 90 section-level evaluation samples across the three held-out brains (M2–M4), each downsampled $16\times$ to a $\sim 4\text{k} \times 4\text{k}$ pixel field of view at $6.4 \mu\text{m}$ effective resolution. We train using annotations from a single brain (M1). Generalization is evaluated on three held-out brains: M2 and M3 (partially used to train the SOTA, precluding direct comparison on these brains) and M4 (fully held-out, enabling comparison with all methods).

Evaluation Metrics. Standard overlap metrics (Dice, IoU) are ill-suited to this task: severe class imbalance deflates their values, and the exact boundaries of each bundle are less important than its location and extent. We therefore adopt the component-based evaluation of [22]: connected component matching yields *True Positive Rate (TPR)* per density class (dense/moderate/sparse), *False Discovery Rate* $\text{FDR} = \text{FP}/(\text{TP} + \text{FP})$, and *average TP/FP per section*.

Baselines. SAM2 [17] and MedSAM2 [28] are evaluated as zero-shot baselines, prompted automatically from the white-matter region. Both perform poorly on tracer histology, occasionally detecting dense bundles but mostly missing moderate and sparse bundles, and FDR close to 1.0, underscoring the need for task-specific training. We further compare against: (i) **Sundaresan et al.** [22], the initial self-supervised deep-learning method for this task, which relies on cross-section post-processing; (ii) **Bintsi et al.** [2], the current state-of-the-art (SOTA), which uses semi-supervised pre-training and requires annotations from three brains (M1–M3); and (iii) a **real-only** baseline using the pipeline from [2], trained on only one brain (M1) instead of three, to isolate the effect of synthetic data from the effect of additional annotated brains.

Implementation Details. We use a 2D U-Net [19] with nine encoder–decoder levels, each comprising two 3×3 convolutions with instance normalization and LeakyReLU. Training uses AdamW [14] with learning rate 10^{-4} , cosine annealing with linear warmup, batch size 8, mixed-precision (FP16), and a combined BCE–Dice loss for 150 epochs. Each batch mixes real and synthetic data at ratio $r_{\text{syn}} = 0.7$ (default). Real patches (1024×1024) are sampled with 50% centered on fiber pixels to address class imbalance; synthetic patches (512×512) are generated on-the-fly and resized to match. Real data receives conservative augmentation (flips, small rotations), while synthetic data undergoes extensive domain randomization (Sec. 2.1). Training runs on a single NVIDIA RTX 6000 GPU.

4 Results

Main comparison. Table 1 compares all methods on the held-out test brain M4. Our method outperforms the SOTA [2] in overall detection (TP_{avg} 3.24 vs. 3.00) and achieves similar sparse-bundle sensitivity (0.79 vs. 0.78), while being trained only on one brain instead of three. It substantially outperforms Sundaresan et al. [22] (TP_{avg} 3.24 vs. 1.76) and improves sparse bundle detection over the real-only baseline (0.79 vs. 0.52), at the cost of a higher FDR (0.36 vs. 0.18).

Figure 2 shows qualitative results on two sections from M4, comparing the SOTA [2] (left) with our method (right). Both methods recover dense and moderate bundles well, but the SOTA baseline often misses sparse bundles that our model detects, in line with the higher sparse-bundle TPR in Table 1. These additional detections come at the cost of slightly more FPs.

Cross-brain generalization. Since we train on M1 only, we can additionally evaluate on M2 and M3, which the SOTA [2] used for training. Table 2 shows that our mixed real+synthetic training (Proposed) generalizes well across held-out brains, improving over [22] especially on moderate and sparse bundles (e.g., on M2, $\text{TPR}_{\text{M/S}}$ 0.95/0.64 vs. 0.85/0.48). Compared to the real-only baseline, the proposed model substantially increases TPR_{S} (from 0.35 to 0.64 on M2 and from 0.38 to 0.62 on M3), at the cost of higher FP_{avg} and FDR on M2 (0.40 vs. 0.11), reflecting a systematic shift toward higher recall under cross-brain shift. Notably, on M3 the FDR is reduced relative to the real-only baseline (0.37 vs.

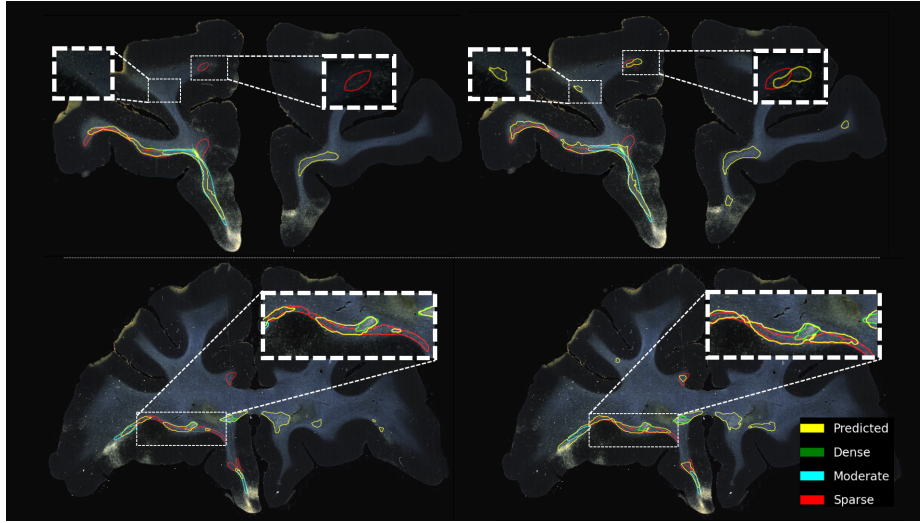


Fig. 2. Visual comparison between the **state-of-the-art method [2]** (left) and **our approach (right)** on two histological sections from the held-out brain M4. Ground-truth bundles are outlined by density class (dense: green, moderate: cyan, sparse: red); predictions are shown in yellow. Insets highlight regions where the baseline often misses sparse bundles that our model detects, consistent with the higher sparse-bundle TPR in Table 1, at the cost of more false positives. Annotations are provided only within the ipsilateral hemisphere; predictions outside this region are excluded from evaluation.

0.47), suggesting that the elevated FDR on M2 reflects a challenging domain shift rather than a systematic failure mode.

Ablations studies. *Real data amount.* Table 3 (left) reports mean performance across held-out brains M2–M4. A single annotated section paired with synthetic data already achieves competitive sensitivity ($\text{TPR}_{D/M/S}$ 0.88/0.80/0.57) with the lowest FDR (0.32); using a full annotated brain further improves detection, especially for sparse bundles (0.69 vs. 0.57), at the cost of a higher FDR (0.38). Performance is non-monotonic so 5 sections do not consistently outperform 1, indicating high variance in the low-shot regime.

Synthetic ratio. Table 3 (right) varies the proportion of synthetic images per batch, averaging over M2–M4. Synthetic-only training fails (mean TPR_S 0.26, FDR 0.94), confirming that real supervision is essential. Compared to the real-only setting ($r_{\text{syn}}=0\%$), mixed real+synthetic training improves dense and moderate TPR, maintains sparse recall, and reduces FDR, achieving a better overall precision–recall trade-off. This holds across a broad range of synthetic ratios ($r_{\text{syn}} \in \{30\%, 50\%, 70\%\}$), indicating robustness to this hyperparameter.

Domain randomization. Across M2–M4 we also compare our default domain randomization to a less aggressive variant. Default DR consistently achieves a better precision–recall trade-off: on all brains it maintains comparable TPR while

Table 1. Performance comparison on the held-out test brain M4. We report true positive rates (TPR) for dense (D), moderate (M), and sparse (S) bundles, average true/false positives per section (TP_{avg} , FP_{avg}), and false discovery rate (FDR).

Method	Train data (# brains)	TPR $\uparrow$			TP_{avg} FP_{avg}		FDR
		D	M	S	$\uparrow$	$\downarrow$	
Sundaresan et al. [22]	1	0.85	0.75	0.41	1.76	2.91	0.62
Real-only [2]	1	0.88	0.75	0.52	2.88	0.64	0.18
Bintsi et al. [2]	3	0.97	0.90	0.78	3.00	1.36	0.30
Bintsi et al. (+ pretrain) [2]	3+10	0.97	0.83	0.62	3.10	0.64	0.20
Proposed	1+synth	0.97	0.89	0.79	3.24	2.06	0.36

Table 2. Cross-brain generalization on M2 and M3 (methods trainable on M1 only; the SOTA [2] cannot be evaluated here as it trains on M1–M3).

Brain	Method	TPR $\uparrow$			TP_{avg} FP_{avg}		FDR
		D	M	S	$\uparrow$	$\downarrow$	
M2	Sundaresan et al. [22]	0.92	0.85	0.48	2.06	3.23	0.61
	Real-only [2]	0.65	0.77	0.35	1.60	0.23	0.11
	Proposed	0.98	0.95	0.64	2.73	3.10	0.40
M3	Sundaresan et al. [22]	0.96	0.76	0.27	0.89	3.63	0.79
	Real-only [2]	0.86	0.81	0.38	1.86	2.00	0.47
	Proposed	0.95	0.95	0.62	3.93	3.06	0.37

reducing FDR (e.g., 0.40 vs. 0.58 on M2 and 0.36 vs. 0.49 on M4), suggesting that stronger augmentation acts as an effective regularizer across brains.

5 Discussion and Conclusion

We have presented a mixed real+synthetic training strategy for fiber bundle segmentation in tracer histology that leverages dMRI tractography as a generative prior. Our results demonstrate that synthetic augmentation effectively complements limited real supervision: performance comparable to the SOTA [2] is achieved with annotations from a single brain (vs. three), and even a single annotated section yields competitive detection when paired with synthetic data. Crucially, synthetic data alone is insufficient. Real supervision remains essential to anchor the learned representations but at intermediate mixing ratios, synthesis acts as a regularizer that improves precision relative to real-only training.

The primary limitation of our approach is elevated FPs on some held-out brains, which we attribute to domain shifts in background appearance, tracer characteristics, and injection patterns not fully captured by the current synthesis pipeline; notably, on M3 FPs increase while FDR *decreases* relative to the real-only baseline, indicating that TP gains outweigh FP increases. We note that this does not reflect circularity from the tractography prior: our model learns only *local* axon-bundle texture from thin histological sections, not long-range connectivity, so the dominant failure mode of tractography does not propagate into the

Table 3. Ablation studies (averaging over M2–M4). Proposed denoted with *.

Training data (# of sections)	TPR (D/M/S) ↑	FDR ↓	r_{syn}	TPR (D/M/S)↑	FDR↓
1	0.88/0.80/0.57	0.32	0%	0.93/0.91/0.69	0.44
5	0.80/0.71/0.53	0.33	30%	0.96/0.89/0.60	0.37
30 (1 brain)*	0.97/0.93/0.69	0.38	50%	0.94/0.90/0.64	0.42
			70%*	0.97/0.93/0.69	0.38
			100%	0.64/0.36/0.26	0.94

segmentation task. Mixed training with real annotated histology further grounds the model in biologically valid tracer appearance. Increasing training diversity through additional real brains or improved synthesis would further improve generalization. The non-monotonic behavior in the low-shot regime also suggests that which sections are annotated matters as much as how many, motivating active selection strategies. Future work will focus on improved synthesis to better capture tracer-dependent intensity profiles and staining artifacts via generative models [8], and uncertainty-aware filtering to reduce FPs.

Ultimately, by reducing dependence on annotated data through synthesis, this work enables reliable bundle segmentation not only for similar protocols but also for unseen tracers and tissue preparations, thus making a key step toward high-throughput mesoscale connectivity mapping.

Acknowledgments. This work was supported by the center for Large-scale Imaging of Neural Circuits (LINC), an NIH BRAIN Initiative Connectivity across Scales (CONNECTS) comprehensive center (UM1-NS132358). Additional support was provided by the National Institute for Mental Health (R01-MH045573, P50-MH106435) and the National Institute for Neurological Disorders and Stroke (R01-NS119911, R01-NS127353).

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

References

1. Avants, B.B., Tustison, N., Song, G., et al.: Advanced normalization tools (ants). *Insight j* **2**(365), 1–35 (2009)
2. Bintsi, K.M., Balbastre, Y., Wu, J., Lehman, J.F., Haber, S.N., Yendiki, A.: Fully automated segmentation of fiber bundles in anatomic tracing data. In: *International Workshop on Computational Diffusion MRI*. pp. 81–92. Springer (2025)
3. Catani, M.: Diffusion tensor magnetic resonance imaging tractography in cognitive disorders. *Current opinion in neurology* **19**(6), 599–606 (2006)
4. Chatterjee, H.J., Ho, S.Y., Barnes, I., Groves, C.: Estimating the phylogeny and divergence times of primates using a supermatrix approach. *BMC evolutionary biology* **9**, 1–19 (2009)
5. Chollet, E., Balbastre, Y., Mauri, C., Magnain, C., Fischl, B., Wang, H.: Neurovascular segmentation in soct with deep learning and synthetic training data. *arXiv preprint arXiv:2407.01419* (2024)

6. Dhollander, T., Zanin, J., Nayagam, B.A., Rance, G., Connelly, A.: Feasibility and benefits of 3-tissue constrained spherical deconvolution for studying the brains of babies. In: Proceedings of the 26th annual meeting of the International Society of Magnetic Resonance in Medicine. vol. 3077 (2018)
7. Friedmann, D., Pun, A., Adams, E.L., Lui, J.H., Kebschull, J.M., Grutzner, S.M., Castagnola, C., Tessier-Lavigne, M., Luo, L.: Mapping mesoscale axonal projections in the mouse brain using a 3d convolutional network. Proceedings of the National Academy of Sciences **117**(20), 11068–11075 (2020)
8. Friedrich, P., Frisch, Y., Cattin, P.C.: Deep generative models for 3d medical image synthesis. In: Generative Machine Learning Models in Medical Image Computing, pp. 255–278. Springer (2024)
9. Gopinath, K., Hoopes, A., Alexander, D.C., Arnold, S.E., Balbastre, Y., Billot, B., Casamitjana, A., Cheng, Y., Chua, R.Y.Z., Edlow, B.L., et al.: Synthetic data in generalizable, learning-based neuroimaging. Imaging Neuroscience **2**, imag-2 (2024)
10. Haynes, W.I., Haber, S.N.: The organization of prefrontal-subthalamic inputs in primates provides an anatomical substrate for both functional specificity and integration: implications for basal ganglia models and deep brain stimulation. Journal of Neuroscience **33**(11), 4804–4814 (2013)
11. Jbabdi, S., Lehman, J.F., Haber, S.N., Behrens, T.E.: Human and monkey ventral prefrontal fibers use the same organizational principles to reach their targets: tracing versus tractography. Journal of Neuroscience **33**(7), 3190–3201 (2013)
12. Jones, R., Grisot, G., Augustinack, J., Magnain, C., Boas, D.A., Fischl, B., Wang, H., Yendiki, A.: Insight into the fundamental trade-offs of diffusion mri from polarization-sensitive optical coherence tomography in ex vivo human brain. Neuroimage **214**, 116704 (2020)
13. Lehman, J.F., Greenberg, B.D., McIntyre, C.C., Rasmussen, S.A., Haber, S.N.: Rules ventral prefrontal cortical axons use to reach their targets: implications for diffusion tensor imaging tractography and deep brain stimulation for psychiatric illness. Journal of Neuroscience **31**(28), 10392–10402 (2011)
14. Loshchilov, I., Hutter, F., et al.: Fixing weight decay regularization in adam. arXiv preprint arXiv:1711.05101 **5**(5), 5 (2017)
15. Naito, T., Nagashima, Y., Taira, K., Uchio, N., Tsuji, S., Shimizu, J.: Identification and segmentation of myelinated nerve fibers in a cross-sectional optical microscopic image using a deep learning model. Journal of neuroscience methods **291**, 141–149 (2017)
16. Pezoulas, V.C., Zaridis, D.I., Mylona, E., Androutsos, C., Apostolidis, K., Tachos, N.S., Fotiadis, D.I.: Synthetic data generation methods in healthcare: A review on open-source tools and methods. Computational and structural biotechnology journal **23**, 2892–2910 (2024)
17. Ravi, N., Gabeur, V., Hu, Y.T., Hu, R., Ryali, C., Ma, T., Khedr, H., Rädle, R., Rolland, C., Gustafson, L., et al.: Sam 2: Segment anything in images and videos. arXiv preprint arXiv:2408.00714 (2024)
18. Reza, A.M.: Realization of the contrast limited adaptive histogram equalization (clahe) for real-time image enhancement. Journal of VLSI signal processing systems for signal, image and video technology **38**(1), 35–44 (2004)
19. Ronneberger, O., Fischer, P., Brox, T.: U-net: Convolutional networks for biomedical image segmentation. In: Medical image computing and computer-assisted intervention–MICCAI 2015: 18th international conference, Munich, Germany, October 5–9, 2015, proceedings, part III 18. pp. 234–241. Springer (2015)

20. Safadi, Z., Grisot, G., Jbabdi, S., Behrens, T.E., Heilbronner, S.R., McLaughlin, N.C., Mandeville, J., Versace, A., Phillips, M.L., Lehman, J.F., et al.: Functional segmentation of the anterior limb of the internal capsule: linking white matter abnormalities to specific connections. *Journal of Neuroscience* **38**(8), 2106–2117 (2018)
21. Schilling, K.G., Nath, V., Hansen, C., Parvathaneni, P., Blaber, J., Gao, Y., Neher, P., Aydogan, D.B., Shi, Y., Ocampo-Pineda, M., et al.: Limits to anatomical accuracy of diffusion tractography using modern approaches. *Neuroimage* **185**, 1–11 (2019)
22. Sundaresan, V., Lehman, J.F., Maffei, C., Haber, S.N., Yendiki, A.: Self-supervised segmentation and characterization of fiber bundles in anatomic tracing data. *Imaging Neuroscience* **3**, imag_a_00514 (2025)
23. Wei, D., Lee, K., Li, H., Lu, R., Bae, J.A., Liu, Z., Zhang, L., dos Santos, M., Lin, Z., Uram, T., et al.: Axonem dataset: 3d axon instance segmentation of brain cortical regions. In: *Medical Image Computing and Computer Assisted Intervention—MICCAI 2021: 24th International Conference, Strasbourg, France, September 27–October 1, 2021, Proceedings, Part I* 24. pp. 175–185. Springer (2021)
24. Winnubst, J., Bas, E., Ferreira, T.A., Wu, Z., Economo, M.N., Edson, P., Arthur, B.J., Bruns, C., Rokicki, K., Schauder, D., et al.: Reconstruction of 1,000 projection neurons reveals new cell types and organization of long-range connectivity in the mouse brain. *Cell* **179**(1), 268–281 (2019)
25. Yan, M., Yu, W., Lv, Q., Lv, Q., Bo, T., Chen, X., Liu, Y., Zhan, Y., Yan, S., Shen, X., et al.: Mapping brain-wide excitatory projectome of primate prefrontal cortex at submicron resolution and comparison with diffusion tractography. *Elife* **11**, e72534 (2022)
26. Yang, J.Y.M., Yeh, C.H., Poupon, C., Calamante, F.: Diffusion mri tractography for neurosurgery: the basics, current state, technical reliability and challenges. *Physics in Medicine & Biology* **66**(15), 15TR01 (2021)
27. Yendiki, A., Aggarwal, M., Axer, M., Howard, A.F., van Walsum, A.M.v.C., Haber, S.N.: Post mortem mapping of connectional anatomy for the validation of diffusion mri. *Neuroimage* **256**, 119146 (2022)
28. Zhu, J., Hamdi, A., Qi, Y., Jin, Y., Wu, J.: Medical sam 2: Segment medical images as video via segment anything model 2. *arXiv preprint arXiv:2408.00874* (2024)