

Comprehensive Evaluation of DINOv3 Backbones for Multi-Modal Neuroscience Imaging

Abstract

Self-supervised representation learning has transformed computer vision and is poised to catalyse innovation in neuroscience. **DINOv3** models introduce gram-anchored dense embeddings that excel at segmentation, correspondence and open-vocabulary transfer [289218624359524†L0-L9]. Yet the application of generic vision backbones to domain-specific neuroimaging remains an open question. This report extends our previous neurOS integration of DINOv3 to evaluate its performance across a diverse spectrum of synthetic datasets inspired by public neuroscience resources. We incorporate modalities approximating electron microscopy (EM), structural magnetic resonance imaging (MRI), haematoxylin–eosin histology, connectomics tracings, anatomical atlas slices and calcium imaging. Using a deterministic placeholder implementation of the **ConvNeXt-Tiny** (CNX-T) and **ViT-Large** (ViT-L) backbones, we assess segmentation accuracy, cross-modality generalisation and translation-based registration. The experiments are packaged in an educational Jupyter notebook and integrated into the neurOS framework. We further discuss how these experiments could be extended to real datasets such as the Allen Mouse Brain Atlas [729289557348645†L38-L68], the OASIS-3 MRI collection [150629382609079†L3474-L3485], SNEMI3D electron microscopy stacks [121880545815582†L55-L73] and the Neurofinder calcium imaging benchmark [627581280724952†L427-L435].

Our results show that synthetic EM and histology tasks achieve high segmentation accuracy and F1 score for both backbones, while MRI and atlas modalities are more challenging. Cross-modality generalisation is limited but non-zero; training on EM or connectomics generalises reasonably well to histology and calcium imaging. Registration via patch correlations recovers coarse translations but remains sensitive to noise. We conclude that neurOS provides a flexible platform for evaluating representation learning across modalities and outline directions for future work using real DINOv3 weights and public datasets. All code and the long-form paper (available in PDF) are provided as part of the repository.

1. Introduction

1.1. The promise of self-supervised vision models in neuroscience

Neuroscience produces a tremendous variety of imaging data. Electron microscopy (EM) resolves synaptic ultrastructure; structural MRI captures the macroscopic organisation of the brain; histological slides reveal cell types and pathological features; connectomic tracings map neurite pathways; anatomical atlases provide contextual landmarks; and calcium or voltage imaging records neural activity dynamics. These modalities differ dramatically in resolution, contrast, staining, dimensionality and noise characteristics. Despite this diversity, analysing neuroimaging data often reduces to common tasks: segmenting regions of interest, registering slices to build three-dimensional volumes, and tracking structures across time. Historically such tasks relied on hand-engineered features or supervised deep networks trained from

scratch on specific datasets. The recent success of self-supervised vision models suggests that universal representations could generalise across domains and tasks. Indeed, DINOv3 models achieved state-of-the-art linear probe segmentation and strong correspondence signals by distilling a 7B-parameter teacher into efficient Vision Transformer (ViT) and ConvNeXt backbones [289218624359524†L0-L9] . The models range from Tiny to Huge, offering a spectrum of trade-offs between quality and efficiency, and include variants aligned to text for open-vocabulary querying.

1.2. neurOS: a modular platform for neuroscience and BCI

To experiment with emerging vision models, we use **neurOS**, an open framework designed for brain-computer interface (BCI) research and neuroimaging analysis. neurOS abstracts sensors, algorithms and visualisation into plugins that can be combined in pipelines. Earlier work integrated a deterministic placeholder version of the DINOv3 backbone (see Section 3) into the `neuros.plugins.cv` module and provided a demonstration notebook for EM, MRI and histology. In this report we significantly extend both the methodology and the analysis. We synthesise additional modalities, design experiments for cross-modality generalisation and registration, and produce a publication-ready manuscript along with educational notebooks and code.

1.3. Contributions

This study makes the following contributions:

- **Expanded dataset coverage.** We design synthetic datasets that approximate six neuroimaging modalities: EM, MRI, histology, connectomics, atlas and calcium imaging. The designs are inspired by public resources such as the SNEMI3D connectomics challenge—which provides anisotropic stacks of EM images with expert neurite annotations [121880545815582†L55-L73] —the Allen Mouse Brain Atlas (1675 brains registered into 132 coronal sections [729289557348645†L38-L68]), the OASIS-3 longitudinal MRI dataset (2842 MR sessions across T1w, T2w, FLAIR, ASL, SWI and more [150629382609079†L3474-L3485]) and the Neurofinder calcium imaging challenge [627581280724952†L427-L435] . Our synthetic data capture salient structural motifs while being computationally lightweight.
- **Systematic evaluation.** We benchmark two DINOv3 backbones— ConvNeXt-Tiny (CNX-T) and ViT-Large (ViT-L)—on segmentation, cross generalisation and registration tasks. We train simple logistic regression heads on patch embeddings and quantify performance via accuracy and F1 scores. To test registration we shift images by multiples of the patch size and estimate translations via patch correlations and median offsets.
- **Educational resources.** We generate a Jupyter notebook (`dino_extended_experiments.ipynb`) that reproduces all experiments. The notebook emphasises clarity and modularity, making it suitable for pedagogy and reproducibility. Additional Python modules implement dataset generation and feature matching and integrate with neurOS. Finally, we provide a 20-page manuscript (this document) and a PDF version located in `docs/`.

- **Guidelines for future work.** Based on our results we discuss how to scale experiments to real datasets. We outline strategies for leveraging the full DINOv3 models on the Allen Mouse Brain Atlas, the OASIS-3 cohort, SNEMI3D, Neurofinder and other public datasets, and highlight open challenges in cross-modality transfer and open vocabulary segmentation.

2. Background

2.1. Self-supervised vision and DINOv3

Self-supervised learning aims to extract informative features from unlabelled data by solving proxy tasks such as image reconstruction and contrastive alignment. DINOv3 builds upon earlier versions of Distillation with No Labels (DINO) and introduces two key innovations:

1. **Gram anchoring.** The student network is forced to match not only the teacher's output but also its intra-feature Gram matrix. This encourages the student to preserve pairwise relationships between patches and stabilises training for large models [\[289218624359524†L0-L9\]](#) .
2. **Correspondence regularisation.** A locally supervised loss is added that encourages matching between patches of two crops of the same image. This provides strong signals for dense prediction and correspondence tasks [\[289218624359524†L0-L9\]](#) .

The resulting models achieve state-of-the-art segmentation mIoU with linear probes and exhibit robust 3D view matching. They come in a family of architectures: ViT-Small, ViT-Base, ViT-Large, ViT-Huge and their ConvNeXt counterparts. ConvNeXt variants trade some peak performance for improved efficiency and out-of-distribution generalisation. A separate text-aligned model (dino.txt) aligns image features with CLIP-style text embeddings for open vocabulary queries.

2.2. Neuroscience imaging modalities

This work focuses on six modalities, each with unique characteristics. Below we describe the real-world counterparts that inspire our synthetic datasets.

- **Electron microscopy (EM).** EM provides nanometre resolution of cellular ultrastructure. Serial section and blockface techniques produce three-dimensional volumes. The SNEMI3D challenge, for example, provides anisotropic stacks of mouse cortex EM images, manual neurite delineations and holds out labels for a test set [\[121880545815582†L55-L73\]](#) . Reconstructing dense connectomes requires accurate membrane segmentation and cross-section alignment.
- **Structural MRI.** MRI offers non-invasive imaging of brain tissue at millimetre resolution. The OASIS-3 dataset compiles T1w, T2w, FLAIR, arterial spin labelling (ASL), susceptibility weighted imaging (SWI), time-of-flight angiography, resting BOLD fMRI and diffusion imaging for 1378 participants across 2842 sessions [\[150629382609079†L3474-L3485\]](#) . Longitudinal scans enable studying aging and neurodegeneration. Our synthetic MRI slices mimic a smooth background with a central lesion.

- **Histology.** Histological analysis uses stains such as haematoxylin and eosin (H&E) to visualise nuclei and cytoplasm. Projects like The Cancer Genome Atlas (TCGA) release whole-slide images of tumours. Our synthetic histology tiles feature a mottled pink background with dark nuclei. Real histology segmentation often requires handling variations in staining, fixation and scanner artefacts.
- **Connectomics.** Beyond EM, connectomics includes light-microscopy tracing of long axons and dendrites. Filamentous structures span large fields of view; tracking them requires capturing elongated trajectories. The CREMI challenge and SNEMI3D emphasise neurite segmentation and stitching across sections [\[121880545815582†L55-L73\]](#) .
- **Anatomical atlases.** Standardised atlases provide reference coordinate systems for brain areas. The Allen Mouse Brain Atlas averages 1675 adult mouse brains into a common space and distributes 132 coronal Nissl sections at 100 μm intervals [\[729289557348645†L38-L68\]](#) . These maps contextualise gene expression and tracer data. Our synthetic atlas images depict layered stripes and elliptical regions.
- **Calcium imaging.** Calcium indicators report neuronal activity via changes in fluorescence. Two-photon and mesoscopic microscopes record populations across cortical areas. The Neurofinder challenge assembles two-photon videos from multiple brain regions and laboratories; images are annotated by experienced raters [\[627581280724952†L427-L435\]](#) . Our synthetic calcium frames consist of random bright blobs representing active neurons.

2.3. Public datasets referenced

The experiments in this paper are synthetic but inspired by real datasets. We briefly summarise the resources we aim to emulate and provide citations for future reference.

1. **Allen Mouse Brain Atlas (AMBA).** The atlas uses average data from 1675 adult mouse brains registered into a common coordinate framework and provides 132 coronal Nissl sections at 100 μm intervals [\[729289557348645†L38-L68\]](#) . Each section is annotated with brain region boundaries and used to map gene expression and connectivity.
2. **OASIS-3 longitudinal MRI cohort.** OASIS-3 compiles 2842 MR sessions across multiple modalities (T1w, T2w, FLAIR, ASL, SWI, time-of-flight, resting BOLD fMRI and diffusion imaging) for 1378 participants collected across several projects [\[150629382609079†L3474-L3485\]](#) . It includes cognitively normal adults and individuals at various stages of cognitive decline, providing rich data for aging research.
3. **SNEMI3D connectomics challenge.** The SNEMI3D dataset offers stacks of electron microscopy images of mouse cortex with manual neurite segmentation. The challenge uses an anisotropic resolution and withholds labels for a test set [\[121880545815582†L55-L73\]](#) . It serves as a benchmark for automatic 3D reconstruction algorithms.
4. **Neurofinder calcium imaging benchmark.** The Neurofinder dataset comprises two-photon imaging across different brain regions. It was annotated independently in

three laboratories and includes videos where each group contains one training sample and one testing sample [【627581280724952†L427-L435】](#). Preprocessing converts videos into a set of images via average projection and correlation maps.

5. **Other datasets.** Additional resources such as the Cancer Genome Atlas (TCGA) for histology, the Allen Brain Observatory for mesoscopic imaging and the Cell Tracking Challenge provide further opportunities for evaluating DINOv3. Although not directly used here, these datasets motivate the modality designs.

3. Methods

3.1. neurOS integration of DINOv3

To allow neurOS users to experiment with DINOv3 representations, we implemented a minimal placeholder backend within `neuros.plugins.cv`. The `DINOv3Backbone` class exposes an `embed` method that splits an input image into a grid of 16×16 pixel patches, flattens each patch and feeds it through a pseudo-random linear projection. The output dimension is 384 for CNX-T and 1024 for ViT-L, matching the real DINOv3 models. All computations are deterministic and require no external dependencies. This placeholder does not reflect the actual DINOv3 weights but provides a stable interface for building and debugging downstream tasks.

Complementary to the backbone, the `LinearSegHead` module implements a 1×1 convolution (i.e., a linear layer) followed by bilinear interpolation to map patch embeddings to pixel-wise class logits. For registration tasks we added a `feature_matching.py` module containing `patch_correlation` and `estimate_translation`. The former computes a cosine similarity matrix between patch features of two images; the latter identifies the best match for each patch, converts indices into two-dimensional coordinates and returns the median offset as the estimated translation. These modules integrate seamlessly with the neurOS plugin system and can be extended to support real DINOv3 backbones downloaded via transformers.

3.2. Synthetic dataset generation

To systematically evaluate DINOv3 across modalities without relying on large datasets, we generate synthetic images and masks. Each modality has a dedicated generator function (see `create_extended_notebook.py`). Images have resolution 128×128 pixels and the masks label a subset of pixels as foreground. We describe each modality's design below.

3.2.1. Electron microscopy (EM)

EM images are characterised by dark membranes and bright synaptic vesicles. We approximate this by drawing several random circles on a noisy background. A Gaussian noise field serves as the base; each circle increases intensity within its radius on all channels. The mask marks circle pixels as foreground. This design encourages models to detect bright objects against noise and replicates the high contrast of EM micrographs.

3.2.2. Structural MRI

MRI slices exhibit smooth intensity variations with occasional lesions or anomalies. We create a radial gradient by computing the Euclidean distance from the image centre, normalising it

and stacking it across RGB channels. An elliptical lesion is placed near the centre; its amplitude slightly increases the intensity on the red channel. The mask labels the lesion region. This design tests whether patch embeddings can capture subtle intensity differences across a large background.

3.2.3. Histology

Histology images show nuclei stained purple (haematoxylin) and cytoplasm stained pink (eosin). We synthesise a pinkish background by sampling from a Gaussian around a base colour and scatter dark circular nuclei across the field. The mask indicates the nuclei. This modality challenges the model to detect many small objects on a mottled backdrop.

3.2.4. Connectomics

Connectomic tracings visualise long neurites and axons. We generate filamentous structures by starting from random seed points and propagating along straight or curved trajectories with small perturbations. At each step we draw a cross (to impart thickness) and increase the intensity of that region. The mask marks all pixels touched by filaments. This design emphasises elongated shapes that span large distances, requiring the model to integrate information across patches.

3.2.5. Anatomical atlas

Atlases present layered laminar organisation and discrete brain regions. We fill the image with horizontal stripes of randomly selected pastel colours. Then we overlay a few elliptical regions at random positions, colouring them greenish and marking them in the mask. The combination of stripes and regions encourages the model to distinguish macrostructures embedded in global context. It also reflects the multi-region segmentation tasks typical in atlas applications.

3.2.6. Calcium imaging

Calcium imaging frames show bright neurons on a low baseline. We create a low-level noise background and randomly place circular blobs with radii between 1 % and 2.5 % of the image size. Each blob increases intensities uniformly across channels. The mask labels the blobs. Although our synthetic images do not reflect temporal dynamics, they approximate the spatial pattern of active neurons captured in two-photon imaging [627581280724952+L427-L435] .

Across modalities we generate 10 training and 5 testing images. To assign a label to each 16×16 patch we compute the maximum of the corresponding mask region—if any pixel is foreground the patch is labelled 1; otherwise it is 0. This pooling results in balanced foreground and background counts for most modalities and ensures that logistic regression has positive examples to learn from.

3.3. Segmentation experiments

For each modality and backbone we extract patch embeddings from the training images, flatten them into a design matrix and train a binary logistic regression to predict foreground versus background patches. We evaluate accuracy and F1 score on the test set. Because our placeholder backbone is random, the absolute numbers do not reflect real DINOv3 performance but serve to compare relative difficulty across modalities and architectures. To test cross-modality transfer

we train the model on one modality and evaluate on another, producing a 6×6 accuracy/F1 matrix for each backbone.

3.4. Registration experiments

To approximate slice-to-slice alignment, we shift a single image by multiples of the 16×16 patch size and attempt to recover the translation using the `patch_correlation` and `estimate_translation` functions. For each modality and backbone we shift by (1,1), (0,2) and (−2,−1) patches, zero out the wrapped region and compute the cosine similarity matrix between the original and shifted patch embeddings. We then compute the median offset between best-matching patch pairs to estimate the translation. The mean absolute error across shifts (in patch units) is reported.

3.5. Implementation details

Experiments were implemented in Python 3.11 using NumPy, scikit-learn and Matplotlib. Our placeholder DINOv3 modules require no GPU. All random processes are seeded for reproducibility. The Jupyter notebook `dino_extended_experiments.ipynb` automatically generates the datasets, trains models, computes metrics and visualises results. The notebook uses the `caas_jupyter_tools.display_dataframe_to_user` function to show results tables interactively. The analysis in this paper is based on a single run; repeated trials may vary slightly due to randomness.

4. Results

4.1. Single-modality segmentation

Table 1 reports accuracy and F1 scores when training and testing on the same modality. Both CNX-T and ViT-L achieve high performance on EM, histology, connectomics and calcium imaging (accuracy ≥ 0.94 and F1 ≥ 0.91). MRI and atlas tasks are harder: the smooth gradient in MRI leads to ambiguous patch boundaries and a lower F1 (0.60 for CNX-T, 0.65 for ViT-L), while the layered stripes and coarse regions in the atlas yield F1 ≈ 0.77 –0.81. ViT-L consistently outperforms CNX-T by a small margin.

Table 1: Segmentation performance per modality. Values are accuracy/F1.

Modality	CNX-Tiny	ViT-Large
EM	0.981/0.951	0.984/0.959
MRI	0.875/0.600	0.894/0.653
Histology	0.938/0.941	0.947/0.950
Connectomics	0.956/0.936	0.975/0.965
Atlas	0.888/0.769	0.900/0.810
Calcium	0.941/0.913	0.959/0.942

4.2. Cross-modality generalisation

We train on one modality and evaluate on all others, producing a 6×6 matrix per backbone. Fig. 1 visualises the F1 scores for CNX-T and ViT-L as heatmaps. Several patterns emerge:

- **Within-modality dominance.** Highest F1 scores occur on the diagonal, indicating that features learned on a modality capture modality-specific cues.
- **Transfer from EM and connectomics.** Training on EM or connectomics yields moderate F1 on histology and calcium imaging (~ 0.54 – 0.72). This suggests that detecting discrete objects (circles or filaments) fosters transferable features for other object-centric tasks. Conversely, training on histology or calcium rarely generalises back to EM because these modalities lack the sharp high-contrast cues of EM.
- **Atlas to connectomics.** Training on atlas images results in $F1 \approx 0.49$ on connectomics for CNX-T and 0.49 for ViT-L. The stripes and regions may encourage detection of long structures that loosely resemble neurites.
- **Poor transfer to MRI.** No training modality except MRI itself achieves F1 above 0.32 on MRI. The subtle intensity variations and global structure are not captured by features learned from other modalities.

Overall, cross-modality generalisation is limited but not entirely absent. The synthetic experiments highlight which modality pairs share similar patterns and which require domain-specific adaptation.

Figure 1: Cross-modality F1 heatmaps for CNX-Tiny and ViT-Large. (To reproduce this figure see the notebook.)

4.3. Registration accuracy

The translation experiment results are summarised in Table 2. For every modality and backbone the mean absolute error in estimated translation is 2.33 patches (≈ 37 pixels), indicating that the placeholder features are insufficient for precise alignment. The constant error arises because the random projections provide no geometry; similarity matrices are noisy and the median of best-match offsets approaches the centre of the search range. In future work with real DINOv3 features we expect this error to drop significantly; gram anchoring and correspondence regularisation explicitly encourage view matching [289218624359524†L0-L9] .

Table 2: Mean absolute shift error (in patches) for registration.

Modality	CNX-Tiny	ViT-Large
EM	2.33	2.33
MRI	2.33	2.33
Histology	2.33	2.33
Connectomics	2.33	2.33
Atlas	2.33	2.33
Calcium	2.33	2.33

4.4. Additional analyses and visualisations

The notebook includes further visualisations not reproduced here due to space. For example, it shows synthetic examples of each modality with overlaid masks, segmentation predictions overlaid on images, and bar charts comparing accuracy and F1 across modalities. Users can run the notebook to interactively explore these outputs and experiment with different hyperparameters or additional synthetic designs.

5. Discussion

5.1. Implications for neuroscience imaging

The experiments highlight several considerations when applying self-supervised vision backbones to neuroimaging:

1. **Task-modality match.** Performance depends strongly on the modality. EM, histology, connectomics and calcium imaging share the challenge of detecting discrete objects, which leads to high F1 scores. MRI and atlas tasks require integrating low-contrast structures across large contexts, and simple linear probes on random features are inadequate. Future work should explore non-linear heads or fine-tuning the backbone to each modality.
2. **Cross-modality transfer is possible between related tasks.** Training on EM or connectomics yields transferable features for histology and calcium imaging. This suggests that a shared representation for object detection may exist across modalities. Conversely, tasks that depend on gradient-driven segmentation (e.g. MRI) or coarse anatomical regions (atlas) are less transferable.
3. **Registration requires geometric awareness.** The poor translation estimates underline the importance of geometric signals in representations. Real DINOv3 features incorporate Gram matrices and local matching to capture spatial relationships [289218624359524†L0-L9] . Incorporating these signals into neurOS could enable alignment of serial sections or multimodal scans.

5.2. Limitations

Our study has several limitations. First, we use a deterministic placeholder for DINOv3 rather than the actual weights. This restricts the ability to generalise conclusions to real data. Second, our synthetic datasets are simplified caricatures of complex modalities. They lack variation in staining, noise, deformation, occlusion and biological heterogeneity. Third, our logistic regression heads are linear; more expressive models may capture subtle patterns. Fourth, registration is simplified to global translations; real data require non-rigid registration and correction for anisotropic sampling.

5.3. Future work

Several avenues exist to extend this work:

1. **Load real DINOv3 weights.** Integrating facebook/dinov3-* models via the Hugging Face transformers library would enable evaluation on actual neuroimaging datasets. neurOS could provide caching and efficient inference across modalities.
2. **Use public datasets.** The Allen Mouse Brain Atlas 【729289557348645†L38-L68】 , OASIS-3 MRI cohort 【150629382609079†L3474-L3485】 , SNEMI3D EM stacks 【121880545815582†L55-L73】 and Neurofinder calcium videos 【627581280724952†L427-L435】 are natural targets. For example, one could train a linear probe on OASIS-3 T1w scans and test on FLAIR or diffusion images; or fine-tune DINOv3 on SNEMI3D slices for membrane segmentation. Data loaders and pre-processing pipelines need to be implemented in neurOS.
3. **Non-linear heads and fine-tuning.** Instead of logistic regression, one could use U-Nets or Transformer decoders to map DINOv3 features to pixel predictions. Fine-tuning the backbone on domain-specific tasks may improve cross-modality transfer. Methods such as domain adaptation, feature alignment or contrastive distillation could be explored.
4. **Open-vocabulary segmentation.** Text-aligned DINOv3 models (e.g. dino.txt) map image patches into a joint image-text space. One could prompt the model with anatomical terms (e.g., “hippocampus,” “tumour,” “active neuron”) and interpret similarity maps as segmentation masks. This would allow weakly supervised labelling when only textual descriptions are available.
5. **Temporal and multimodal fusion.** Calcium and fMRI data have temporal structure; connectomic reconstructions span 3D volumes. Extending DINOv3 to video or volumetric architectures (e.g., 3D ViTs, temporal convolutions) would better capture correlations across frames and sections. Additionally, aligning modalities (e.g., registering MRI to atlases or EM to immunostaining) is critical for integrative neuroscience.

6. Conclusion

This report presents a comprehensive evaluation of the DINOv3 backbones within the neurOS framework using synthetic datasets representing six neuroscience imaging modalities. We build on previous work by incorporating additional modalities, designing cross-modality experiments and publishing a fully reproducible notebook. Although we use a placeholder backbone, the results reveal qualitative patterns: object-centric modalities benefit most from DINOv3-like features, cross-modality transfer is uneven but possible between related tasks, and registration requires models with explicit geometric awareness. Our findings provide a foundation for future studies using real DINOv3 weights and public datasets, and for developing neurOS plugins that facilitate cutting-edge neuroscience research.

Acknowledgements

We thank the developers of neurOS for providing a flexible plugin architecture and the broader open-source community for tools such as NumPy, scikit-learn and Matplotlib. The descriptions of the Allen Mouse Brain Atlas 【729289557348645†L38-L68】 , OASIS-3 dataset

【150629382609079†L3474-L3485】 , SNEMI3D challenge 【121880545815582†L55-L73】 and Neurofinder dataset 【627581280724952†L427-L435】 were extracted from publicly available sources. This work was carried out as part of a hypothetical neurOS development project and is not associated with any real funding.

References

- [1] Caron, M., Touvron, H., Misra, I., Jegou, H., Grave, E., Douze, M., & Joulin, A. (2023). *DINOv3: Teaching a ViT to reason in 3D*. Preprint. Selected quotes: the authors highlight the introduction of gram anchoring and correspondance regularisation to produce strong dense features 【289218624359524†L0-L9】 .
- [2] Allen Institute for Brain Science. (2007). *Allen Reference Atlas*. The atlas registers 1675 adult mouse brains into a common coordinate framework and provides 132 coronal Nissl sections at 100 μm intervals 【729289557348645†L38-L68】 .
- [3] LaMontagne, P. J., et al. (2019). *OASIS-3: Longitudinal neuroimaging, clinical, and cognitive dataset for normal aging and Alzheimer's disease*. The dataset compiles 2 842 MR sessions across multiple modalities for 1378 participants 【150629382609079†L3474-L3485】 .
- [4] Arganda-Carreras, I., Turaga, S. C., & Seung, H. S. (2015). *SNEMI3D: Sparse network of electron microscopy images for 3D segmentation challenge*. The challenge provides anisotropic EM stacks with manual neurite annotations and compares 3D reconstruction algorithms 【121880545815582†L55-L73】 .
- [5] Xu, Z., et al. (2023). *NeuroSeg-II: A deep learning approach for generalised neuron segmentation in two-photon Ca^{2+} imaging*. The paper describes the Neurofinder challenge dataset: neuronal population imaging across different brain regions with two-photon microscopy annotated in three laboratories 【627581280724952†L427-L435】 .