Pencils down, everyone: When to stop proofreading an electron microscopy volume

Hannah Martinez[1], Jordan K. Matelsky[1], Daniel Xenes[1], Cara Cavanaugh[2], Katharine Merfeld[2], Cody J. Smith[2], Brock A. Wester[1]

Johns Hopkins Applied Physics Laboratory[1], University of Notre Dame[2]

Abstract

The rate-limiting step of large scale connectome generation is manual proofreading, and it's difficult to determine when the dataset has achieved a sufficient level of quality for downstream science.

We present **a fast, inexpensive way to evaluate the quality of connectomes** and other data products at incremental checkpoints during proofreading. We subdivide computation into thousands of small tasks and compute a connectome and contactome in minutes at low cost, then repeat at proofreading checkpoints to **quantify proofreading completeness**.

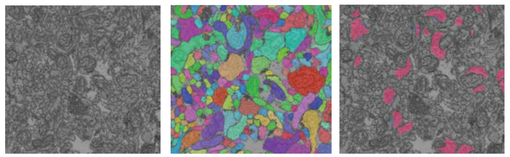
Our method reduces the uncertainty associated with the planning and prioritization of proofreading activities, and enables data owners to accurately forecast and minimize of the amount of proofreading necessary for impactful science.

Background

Connectome generation workflows still produce reconstruction errors such as merged and split cell segments. These must be corrected through automated **[PSTR 418.19 UU21]** or manual proofreading, which is a cost driver. It is critical to understand proofreading progress in order to plan science and budget appropriately.

We share a statistical technique to assess proofreading progress. We validate our approach on a 74 × 74 × 207 µm zebrafish spinal cord SBEM volume [1]. Over 31 months, 20 proofreaders have spent ~5300 hours committing approximately 50 thousand edits.

We replay edit history and generate checkpoints after every 5K proofreading operations. In the zebrafish dataset, this results in 10 timeseries checkpoints, revealing the evolution of the data over proofreading time.



SBEM Image *Segmentation* *Pre/Postsynapses*

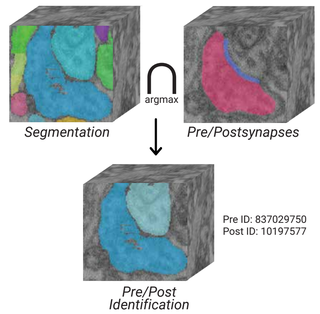
[1] Fabian N (2018). Volume EM Reconstruction of Spinal Cord Reveals Wiring Specificity in Speed-Related Motor Circuits. Cell Reports. https://doi.org/10.1016/j.celrep.2018.05.023.

We thank the NIH for supporting this work under BRAIN Initiative grant R24MH114785, as well as the University of Notre Dame and the Gallagher Family.

Methods

Synaptic Participation Tagging

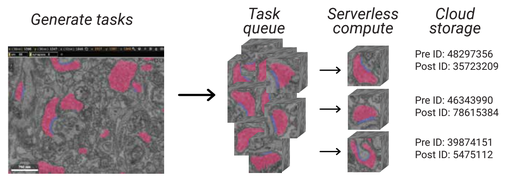
Connectome generation is parallelized at the level of individual synapses (125.6K total). We identify synaptic participants by masking the segmentation IDs with the maximum overlap with pre- and post-synaptic paint.



Connectome & Contactome Generation at Scale

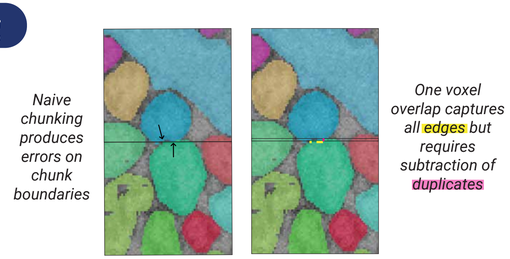
We shard contactome and connectome itemwise tasks, which scales linearly with the image volume: For 125.6K synapses, we incur a cost of $2 in compute, or **$15.92/million synapses.**

Contactome volumes can be parallel-processed at the level of cuboid chunks. These must be stitched to avoid edge-effects.



Contactome Chunking

When processing a cuboid for contactome generation, we must account for edge effects. We add a one voxel overlap on 3 out of 6 faces to ensure all edges are included, then subtract out duplicates.



Invariant Modeling & Forecasting

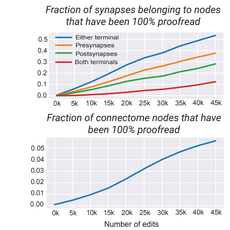
Connectome and contactome graph invariants were fit with a family of strictly monotone models (inverse-power, exponential-decay, and low-order polynomials) via linear least squares. Model selection used Bayesian Information Criterion (BIC) with a slight prior toward convergent forms to reduce model complexity without sacrificing fit, yielding asymptote estimates.

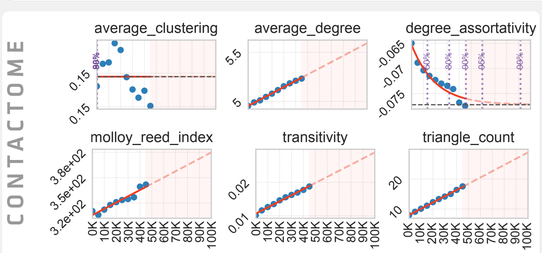
$$\mathrm{BIC} = n \times \ln(\mathrm{SSE} \,/\, n) + k \times \ln(n)$$

Results

Proofreading Progress

Through our checkpointing evaluation it is possible to view proofreading progress statistics at a glance. **Incremental checkpoints reveal predictable trends of connectome synapse reliability.**





Avg Clustering: Fraction of triples that form closed triangles. Reflects local cohesiveness.

Degree Assortativity: Correlation between degrees of connected nodes. Indicates hub/hub or hub/leaf structure.

Degree Entropy: Measures uneven node connectivity.

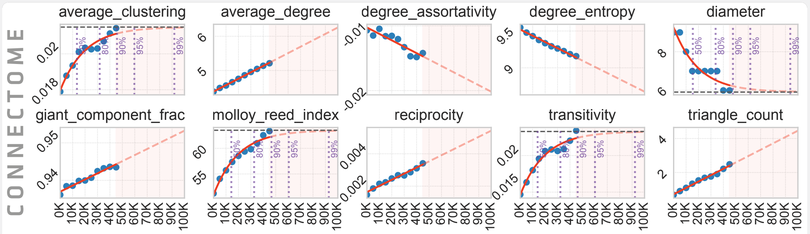
Diameter: Longest shortest path, global reachability.

Giant Comp. Frac.: Fraction of nodes in the largest component.

Molloy Reed: Ratio predicting percolation threshold. Higher values imply well-connected network.

Reciprocity: Fraction of bidirectional edges.

Transitivity: Global clustering coefficient across all triples.



Proofreading dynamics reveal asymptotic convergence of network invariants

These inexpensive graph invariants calculated using the connectome results follow trends that converge on **well-defined asymptotes.**

Invariant	Asymptote	50%	80%	90%	95%	99%
avg_clustering	0.0214889	*13,462*	*31,786*	45,698	59,655	91,846
diameter	5.9147	*15,443*	*33,767*	47,724	61,681	93,872
molloy_reed	63.528	*14,295*	*32,641*	46,598	60,555	92,746
transitivity	0.0232872	*13,957*	*32,281*	46,148	60,105	92,296
		14.3K ± 730	*32.6K ± 729*	**46.5K ± 750**	**60.5K ± 753**	**92.7K ± 753**

When these asymptotes are calculated independently, their future behaviors independently predict extremely similar trajectories of connectome quality.

Our method accurately evaluates current checkpoint quality and forecasts additional proofreading effort required to achieve a desired error ceiling.