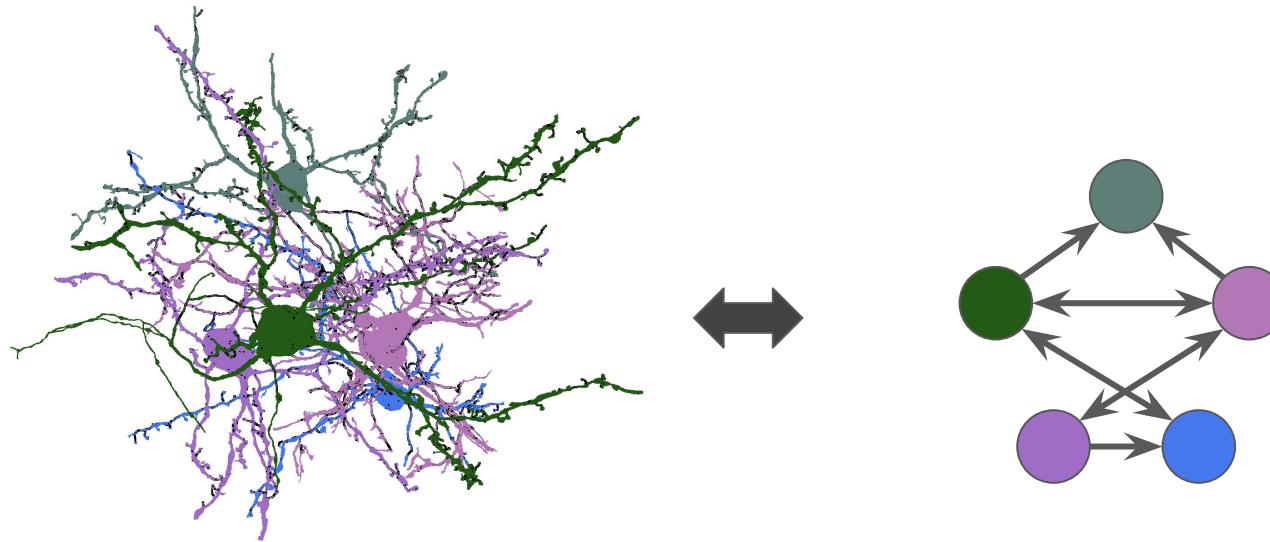


Biologically-Aware Algorithms for Connectomics

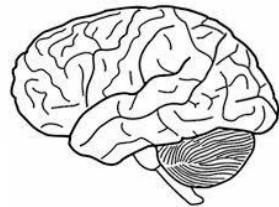


Brian Matejek

Advisor: Hanspeter Pfister
Committee: Michael Mitzenmacher, Todd Zickler

Connectomics

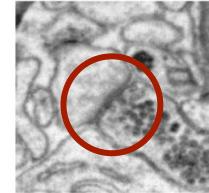
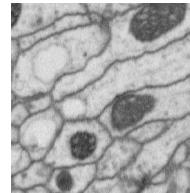
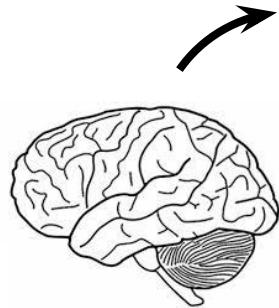
Goal: Extract the wiring diagram from a brain



Connectomics

Goal: Extract the wiring diagram from a brain

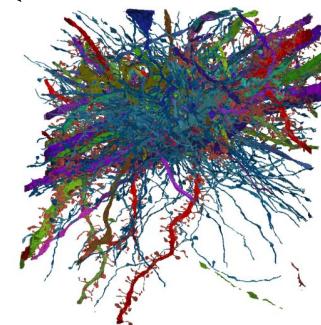
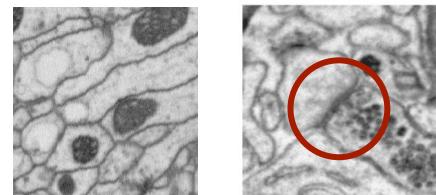
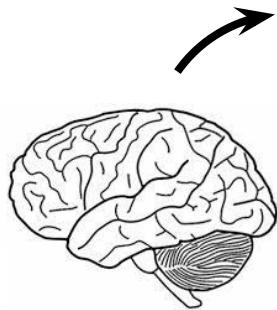
Nanometer-Resolution Imaging



Connectomics

Goal: Extract the wiring diagram from a brain

Nanometer-Resolution Imaging

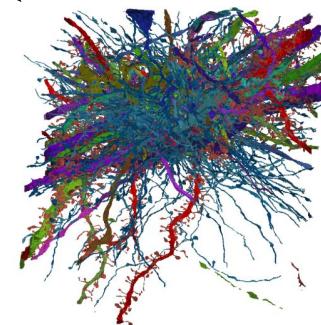
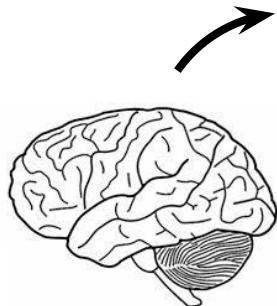


Circuit
Reconstruction

Connectomics

Goal: Extract the wiring diagram from a brain

Nanometer-Resolution Imaging



Circuit
Reconstruction

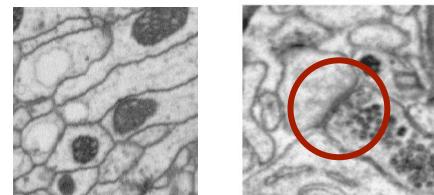


Improved Neural Networks, Medicine, Models

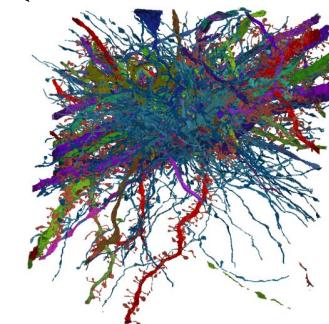
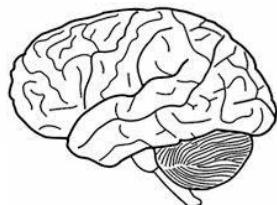
Connectomics

Goal: Extract the wiring diagram from a brain

Nanometer-Resolution Imaging



Better
Understanding



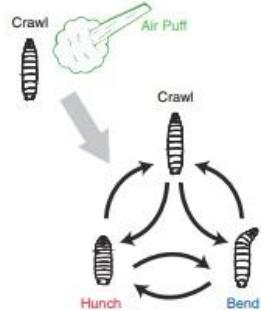
Circuit
Reconstruction



Improved Neural Networks, Medicine, Models

Connectomics

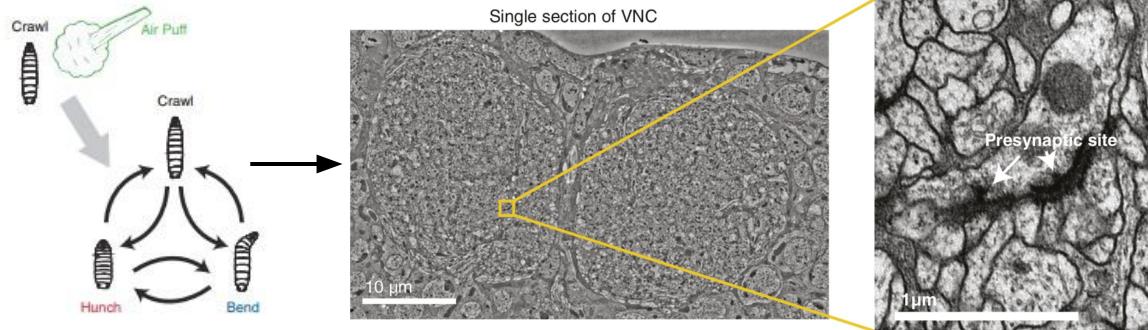
Goal: Extract the wiring diagram from a brain



Behavior

Connectomics

Goal: Extract the wiring diagram from a brain

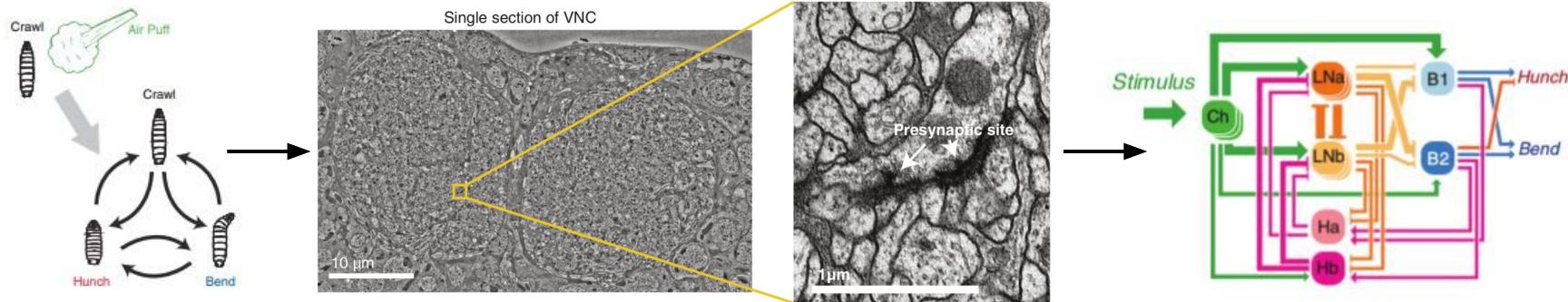


Behavior

Structure

Connectomics

Goal: Extract the wiring diagram from a brain



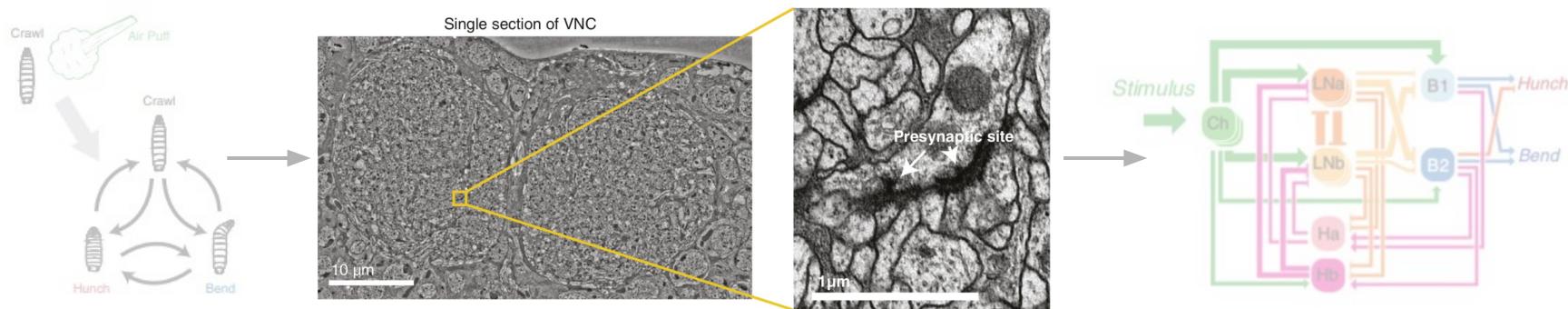
Behavior

Structure

Function

Connectomics

Goal: Extract the wiring diagram from a brain

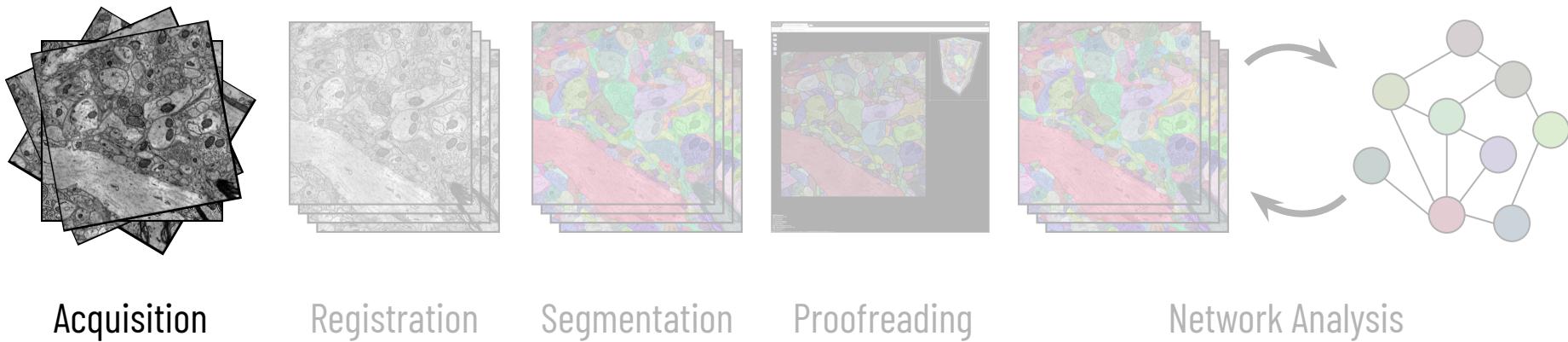


Behavior

Structure

Function

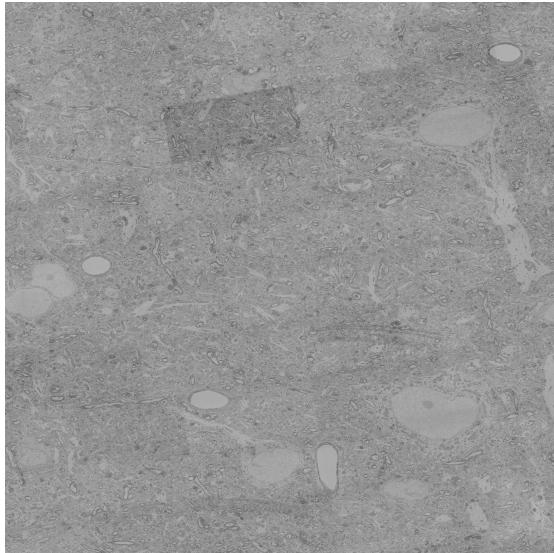
Connectomics Pipeline



Suissa-Peleg *et al.*, Automatic Neural Reconstruction from Petavoxel of Electron Microscopy, *Microscopy and Microanalysis* 2016
Schalek *et al.*, Imaging a 1 mm^3 Volume of Rat Cortex Using a MultiBeam SEM, *Microscopy and Microanalysis* 2016
Xu *et al.*, Enhanced FIB-SEM Systems for Large-Volume 3D Imaging, *biorxiv* 2020
Yin *et al.*, A Petascale Automated Imaging Pipeline for Mapping Neuronal Circuits with High-Throughput Transmission Electron Microscopy, *Nature Communications* 2020

Image Acquisition

Multi-beam electron microscopes collect 1 TB of raw image data every hour

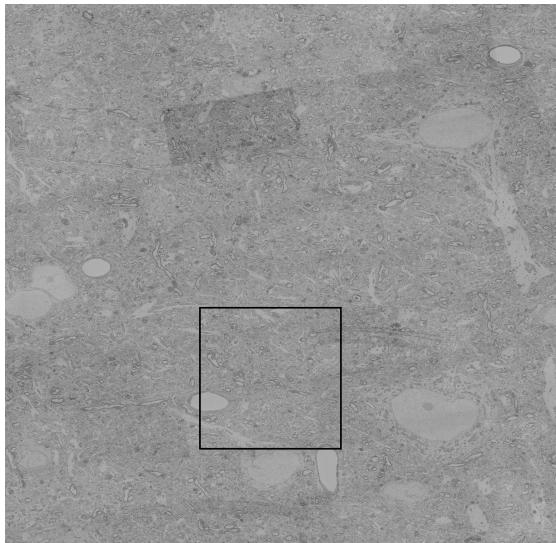


100 μm

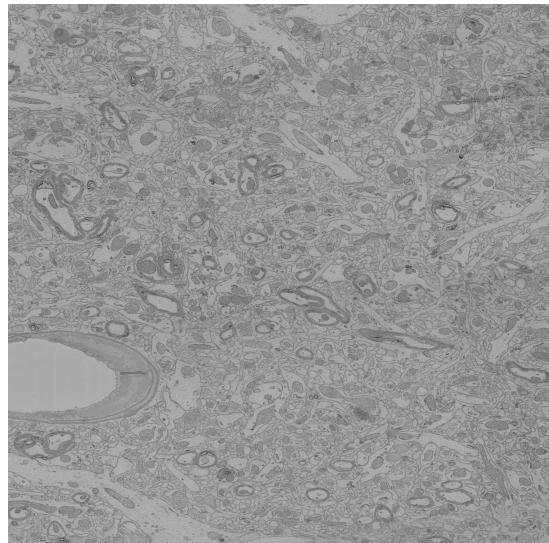
Image Acquisition

Multi-beam electron microscopes collect 1 TB of raw image data every hour

Can image 1 mm^3 of image data (2 PB) in 6 months



100 μm

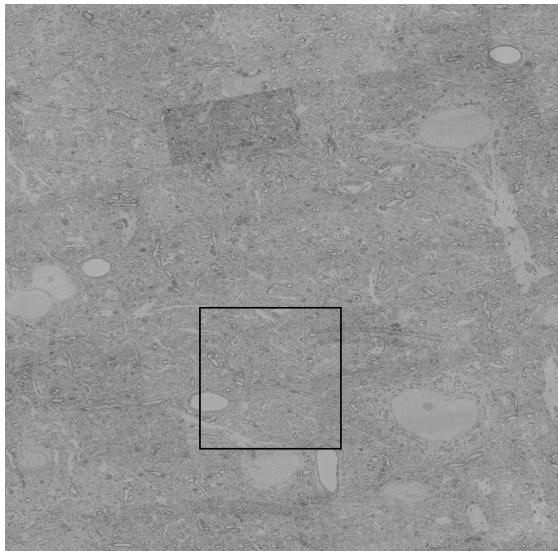


25 μm

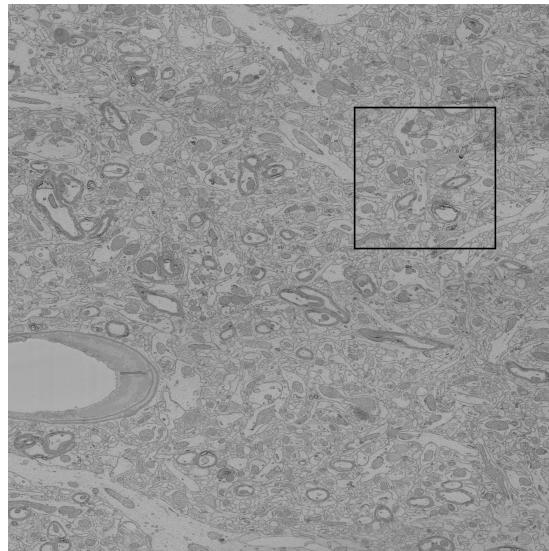
Image Acquisition

Multi-beam electron microscopes collect 1 TB of raw image data every hour

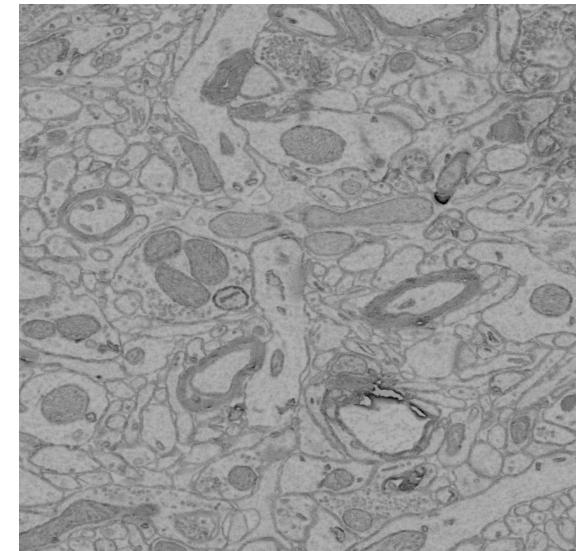
Can image 1 mm³ of image data (2 PB) in 6 months



100 μm

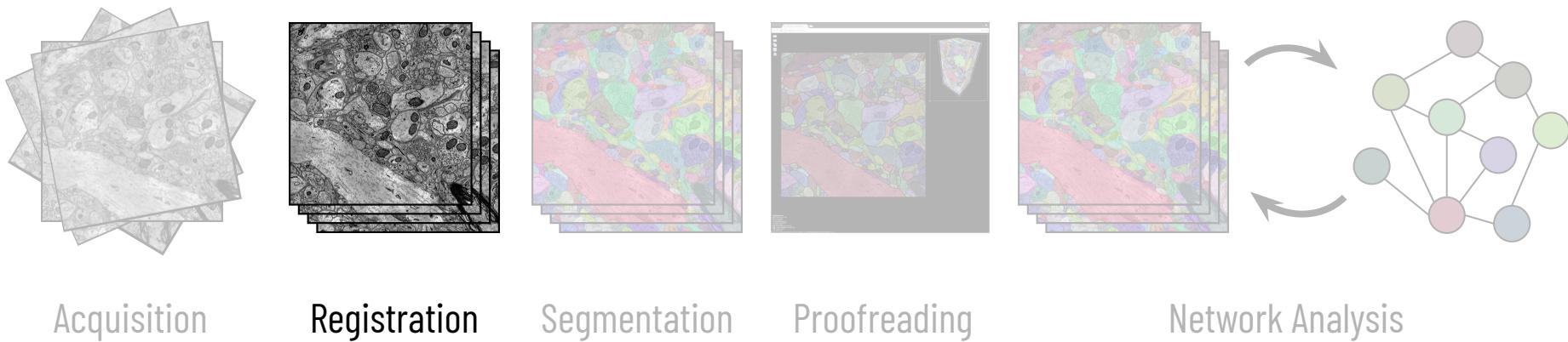


25 μm



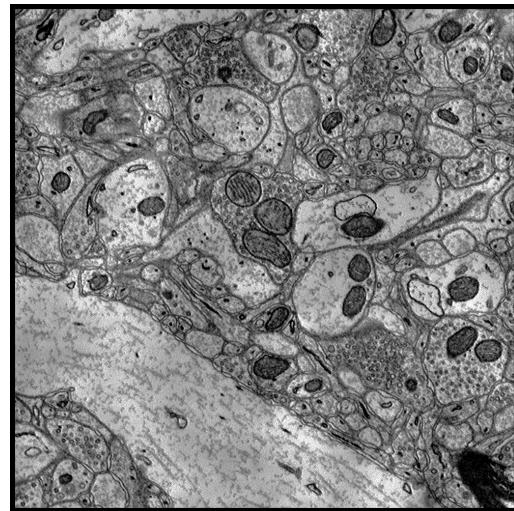
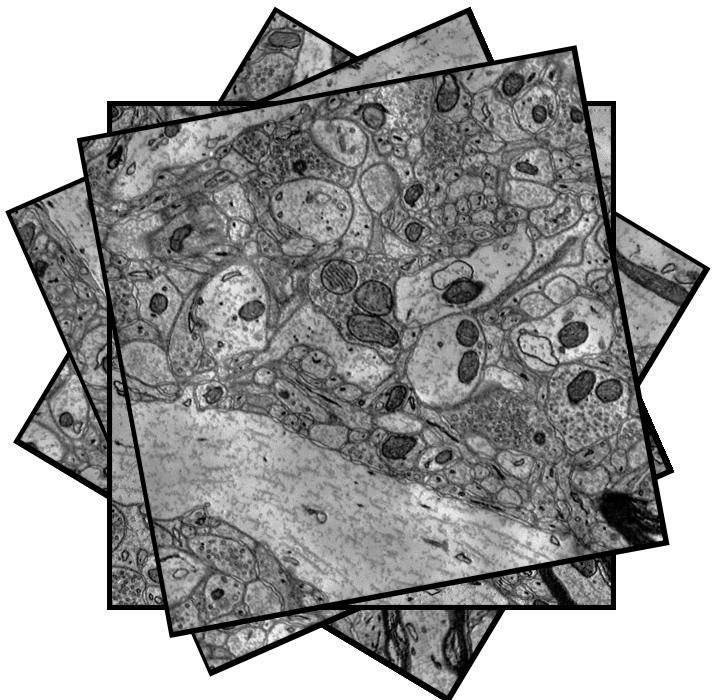
6250 nm

Connectomics Pipeline

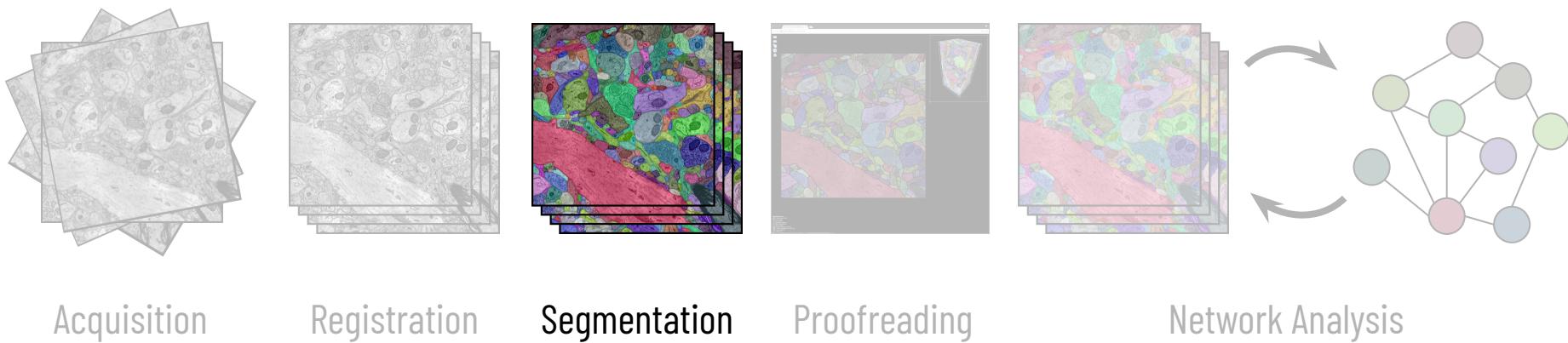


Saalfeld *et al.*, Elastic Volume Reconstruction from Series of Ultra-thin Microscopy Sections, Nature 2012
Khairy *et al.*, Joint Deformable Registration of Large EM Image Volumes: A Matrix Solver Approach, arxiv 2018

Registration



Connectomics Pipeline



Nunez-Iglesias *et al.*, Machine Learning of Hierarchical Clustering to Segment 2D and 3D Images, PLoS ONE 2014

Cicek *et al.*, 3D U-Net: Learning Dense Volumetric Segmentation from Sparse Annotation, MICCAI 2016

Zeng *et al.*, DeepEM3D: Approaching Human-Level Performance on 3D Anisotropic EM Image Segmentation, Bioinformatics 2017

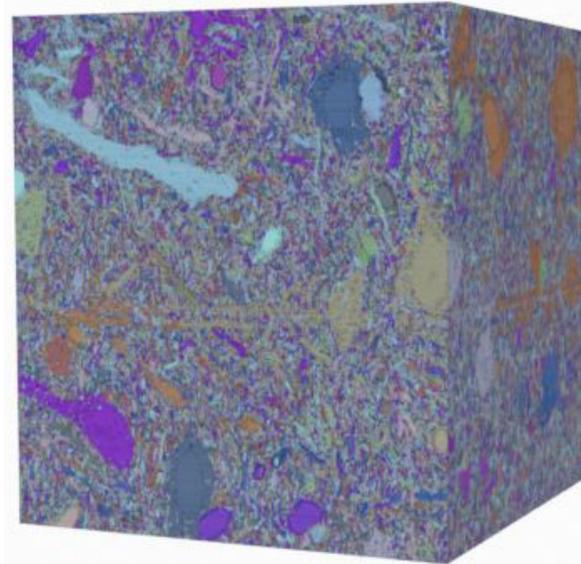
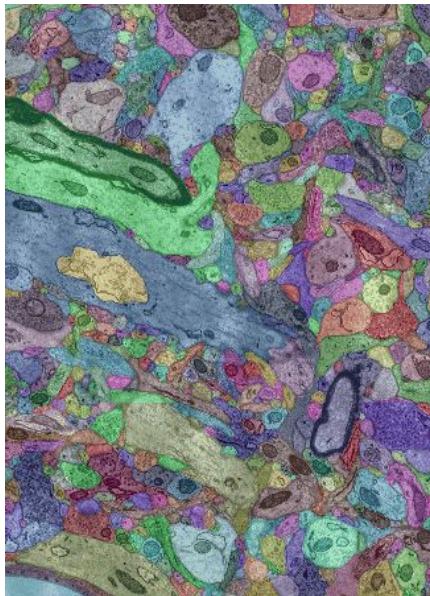
Pape *et al.*, Solving Large Multicut Problems for Connectomics via Domain Decomposition, ICCV 2017

Lee *et al.*, Superhuman Accuracy on the SNEMI3D Connectomics Challenge, arxiv 2017

Januszewski *et al.*, High-Precision Automated Reconstruction of Neurons with Flood-Filling Networks, Nature Methods 2018

Label Volumes

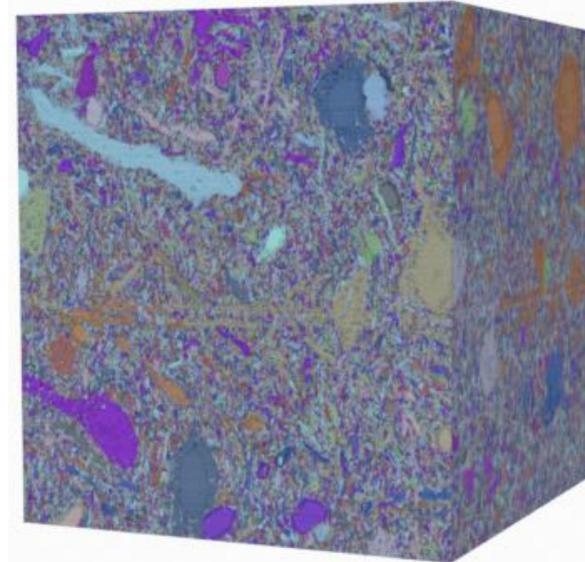
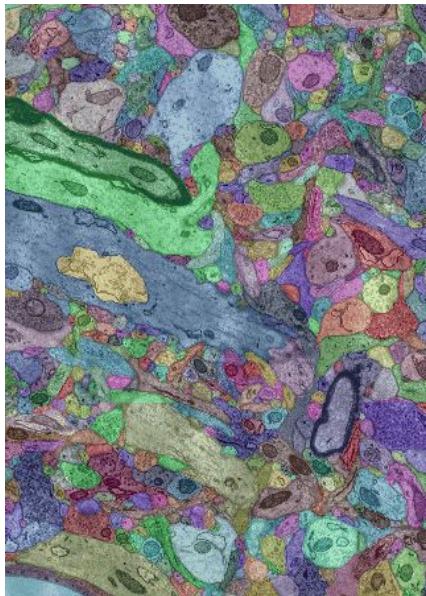
Two voxels have the same label only if they belong to the same neuron



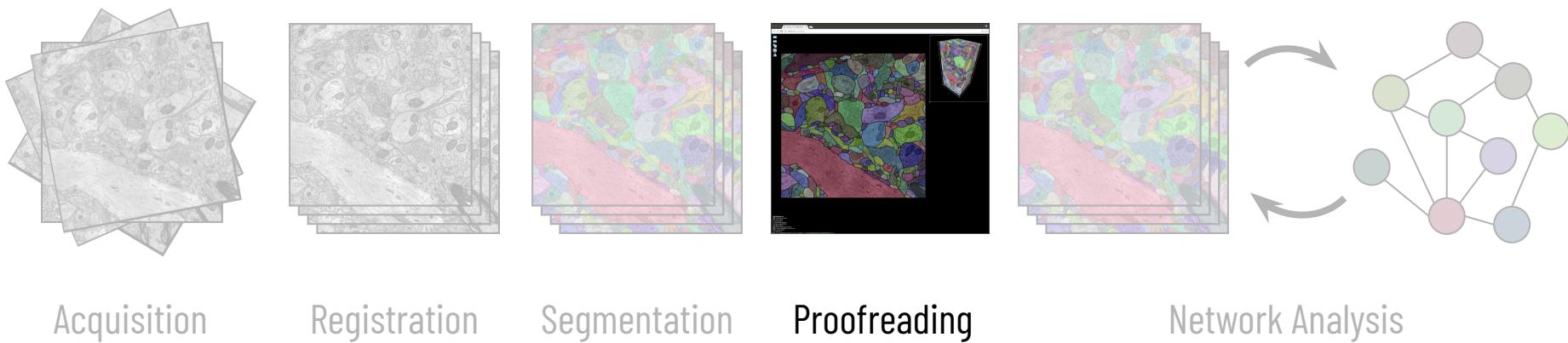
Label Volumes

Two voxels have the same label only if they belong to the same neuron

Typically use 64 bits per voxel to label each segment uniquely



Connectomics Pipeline



Haehn *et al.*, Design and Evaluation of Interactive Proofreading Tools for Connectomics, IEEE VIS 2014

Zung *et al.*, An Error Detection and Correction Framework for Connectomics, NIPS 2017

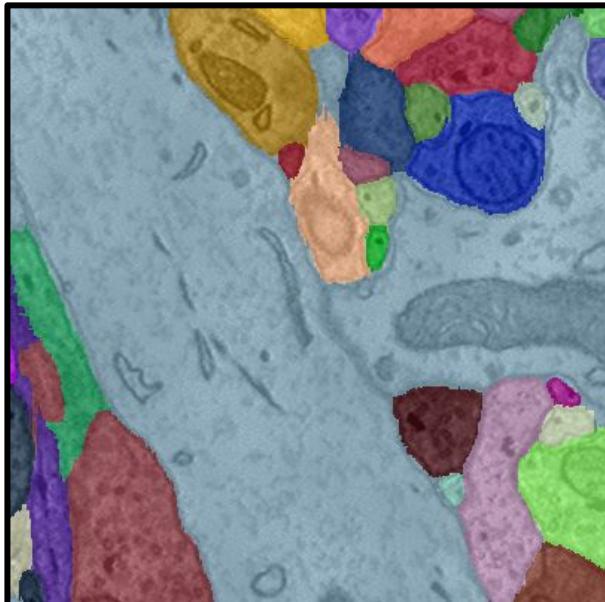
Haehn *et al.*, Guided Proofreading of Automatic Segmentations for Connectomics, CVPR 2018

Dmitriev *et al.*, Efficient Correction for EM Connectomics with Skeletal Representation, BMVC 2018

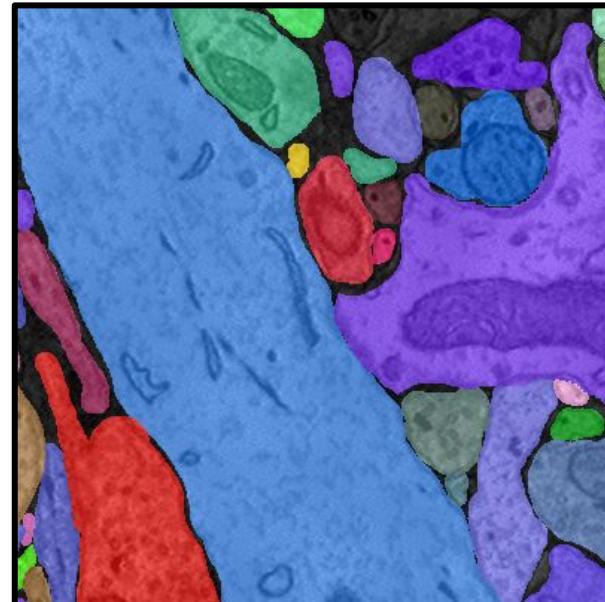
Matejek *et al.*, Biologically-Constrained Graphs for Global Connectomics Reconstruction, CVPR 2019

Merge Errors

Automatic Segmentation

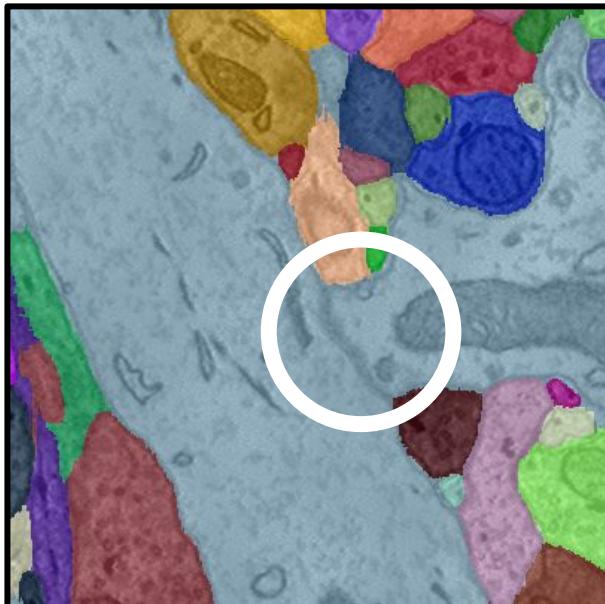


Ground Truth

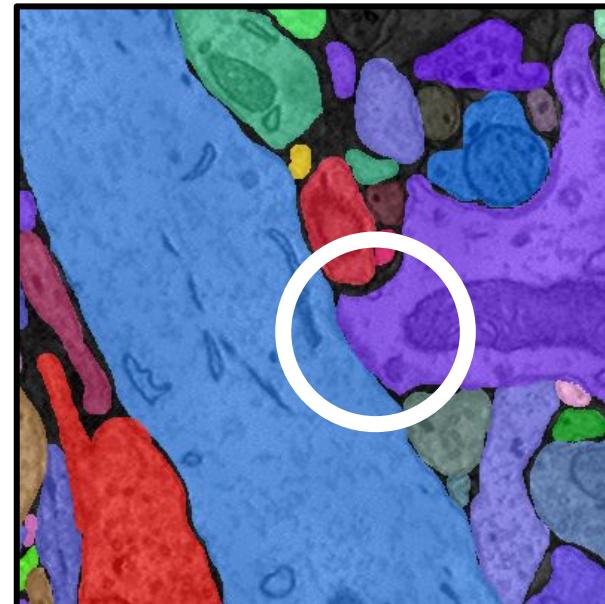


Merge Errors

Automatic Segmentation

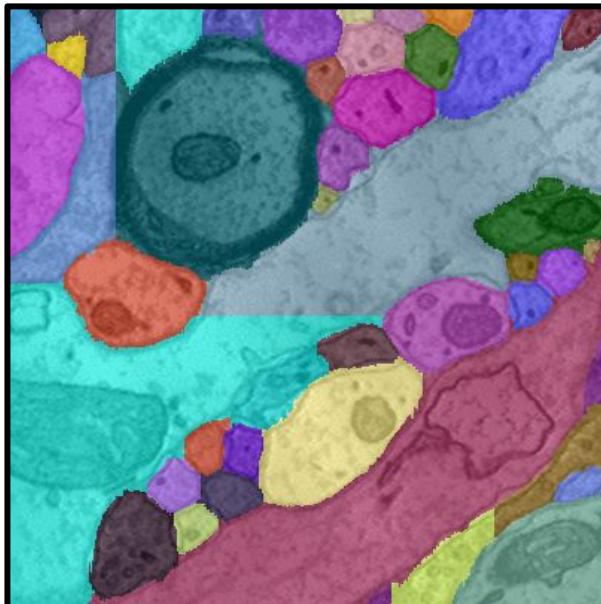


Ground Truth

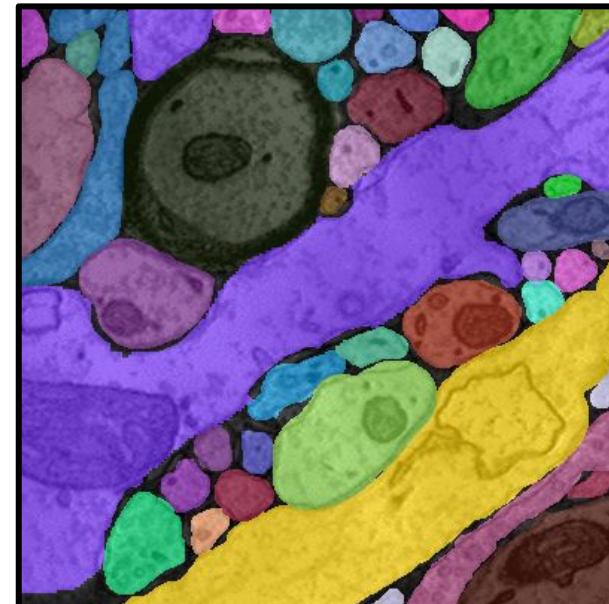


Split Errors

Automatic Segmentation

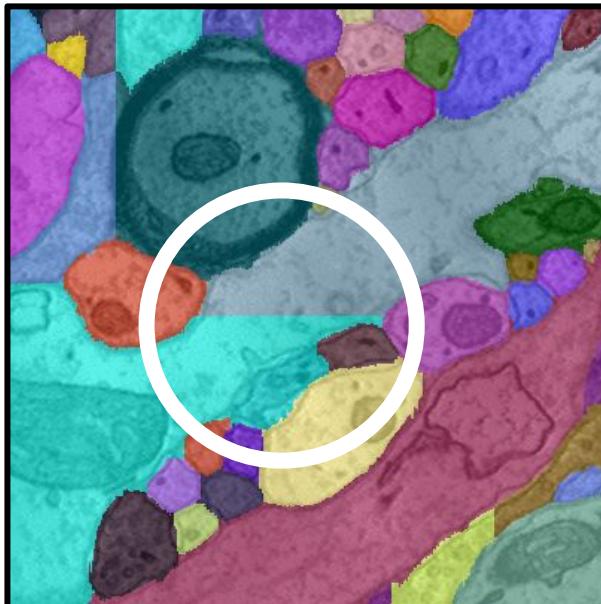


Ground Truth

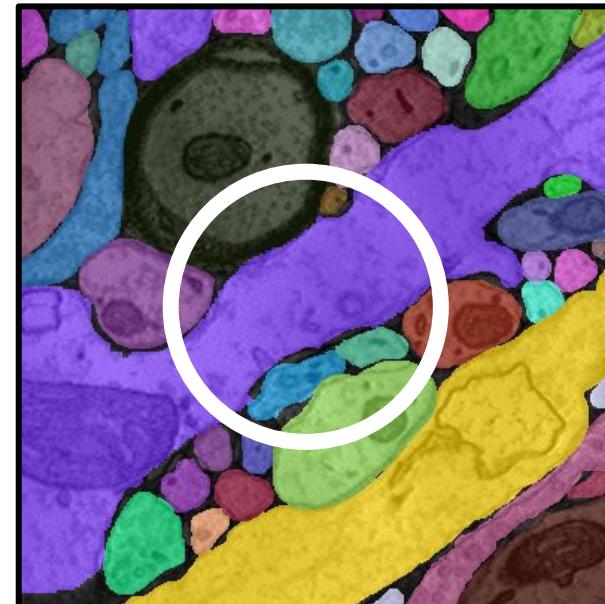


Split Errors

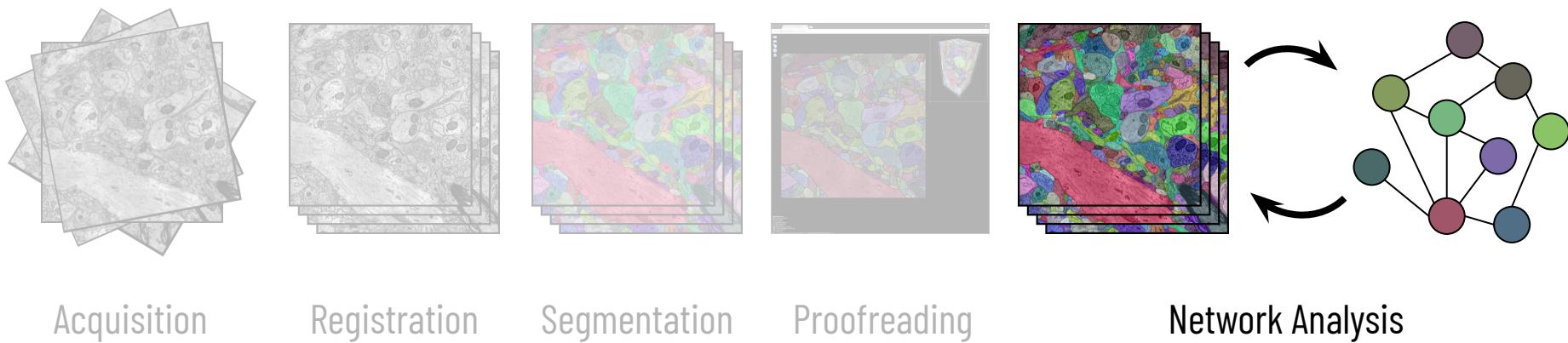
Automatic Segmentation



Ground Truth



Connectomics Pipeline



Sorger *et al.*, neuroMAP - Interactive Graph-Visualization of the Fruit Fly's Neural Circuit, BioVIS 2013

Al-Awami *et al.*, NeuroLines: A Subway Map Metaphor for Visualizing Nanoscale Neuronal Connectivity, IEEE VIS 2014

Haehn *et al.*, Scalable Interactive Visualization for Connectomics, MDPI Informatics 2017

Cook *et al.*, Whole-Animal Connectomes of Both *Caenorhabditis elegans* Sexes, Nature 2019

Scheffer *et al.*, A Connectome and Analysis of the Adult *Drosophila* Central Brain, eLife 2020

Network Analysis

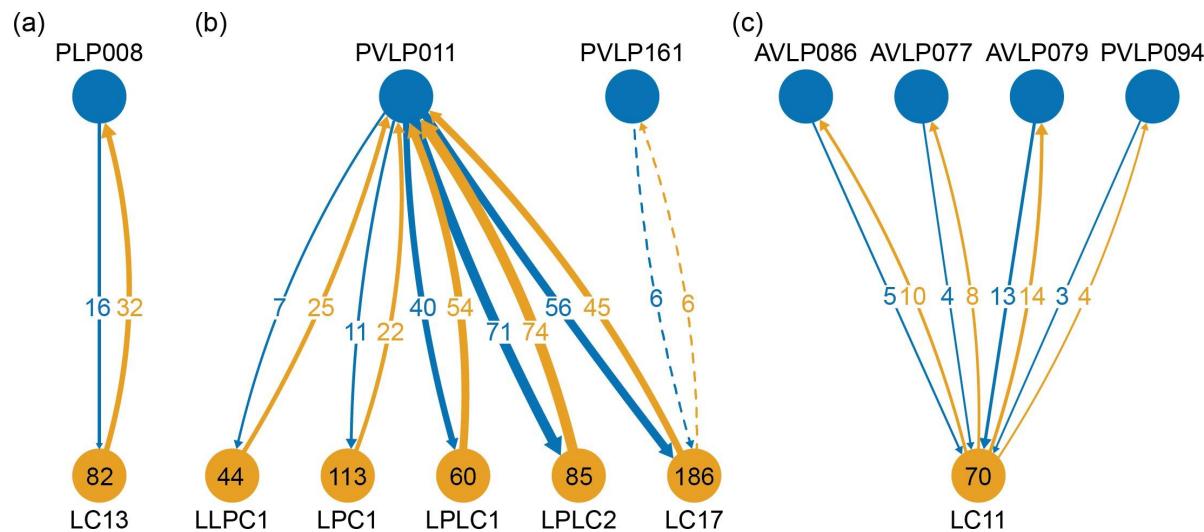
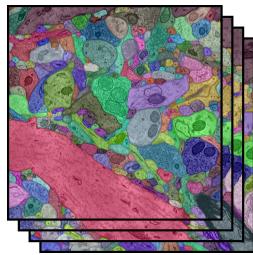
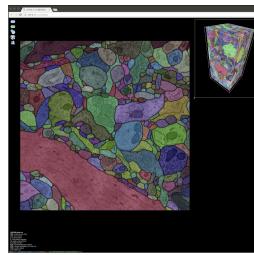


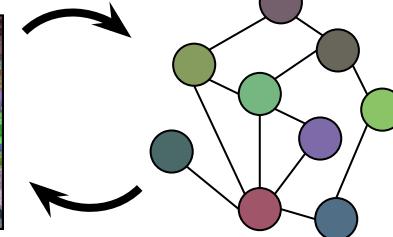
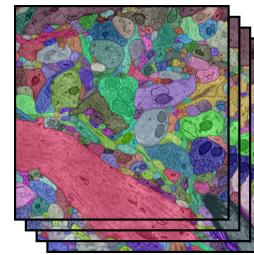
Image to Analysis Pipeline



Segmentation



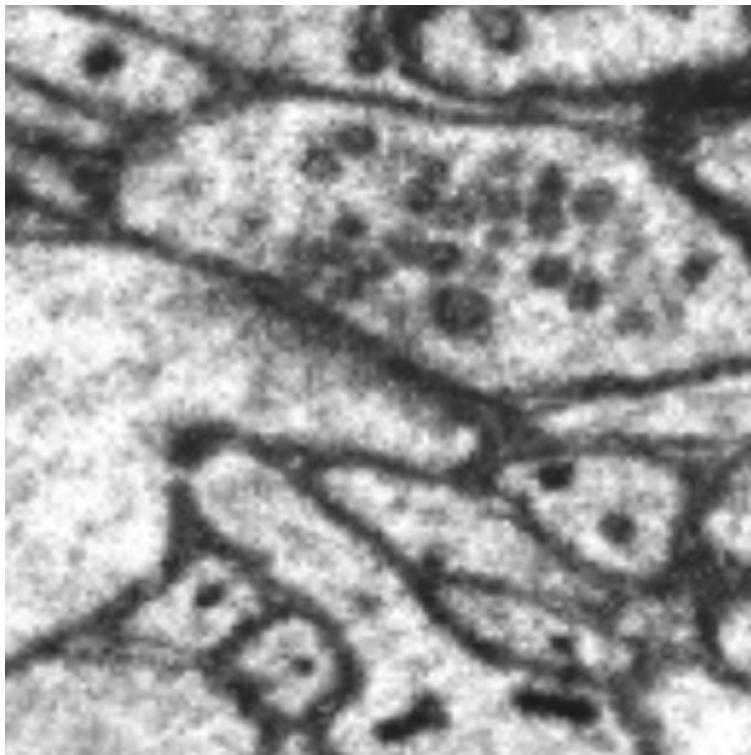
Proofreading



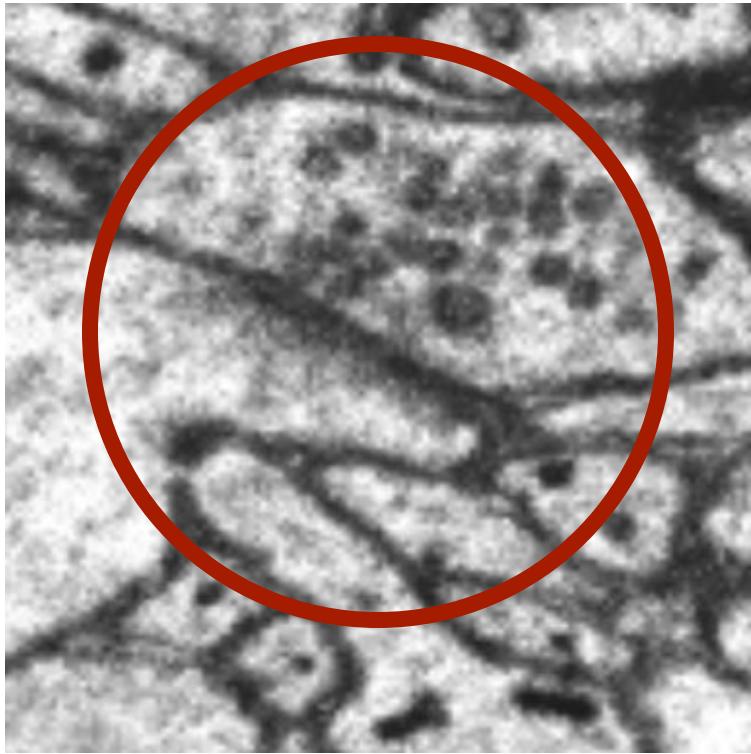
Network Analysis

Scale Challenges of Connectomics

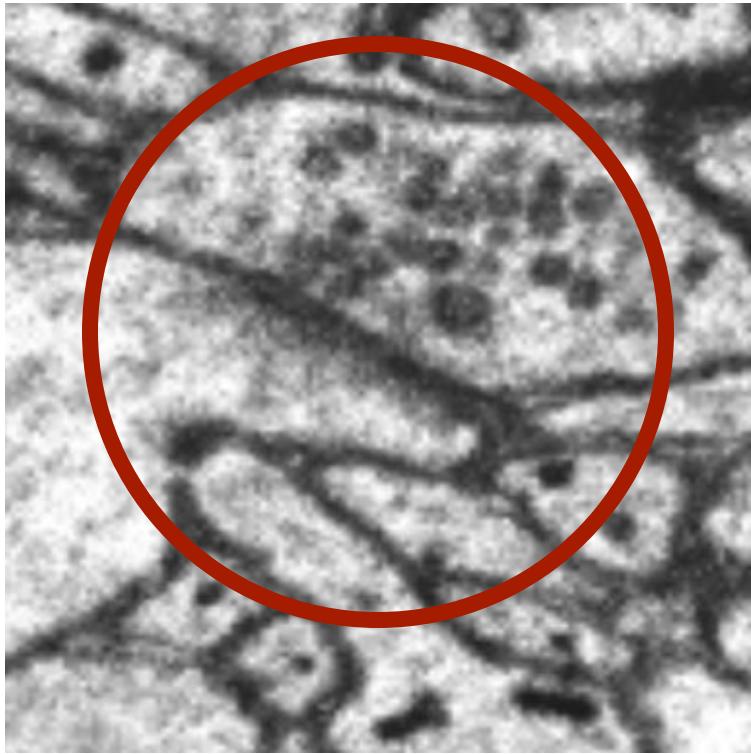
Reconstructing wiring diagrams requires extreme differences in scale



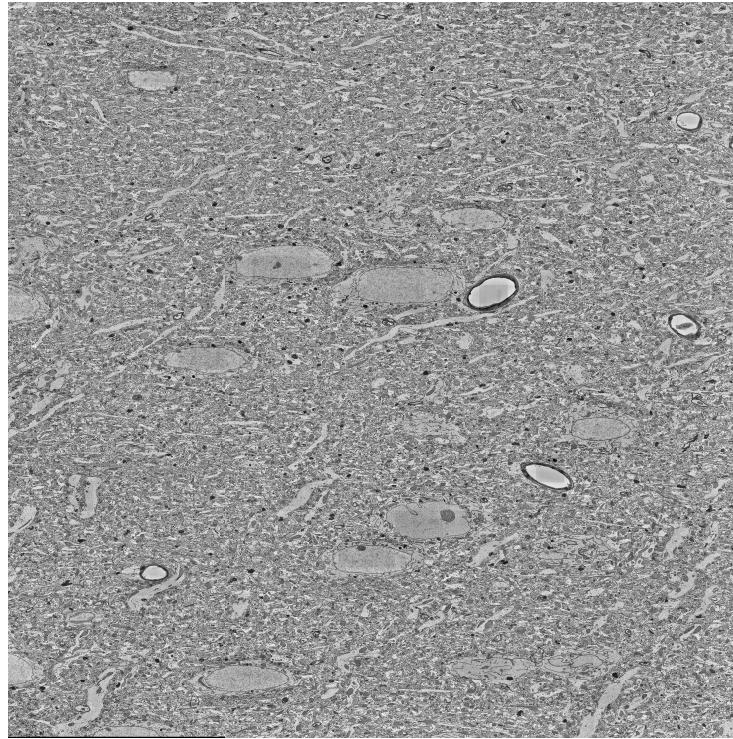
4-10 nanometer resolution



4-10 nanometer resolution

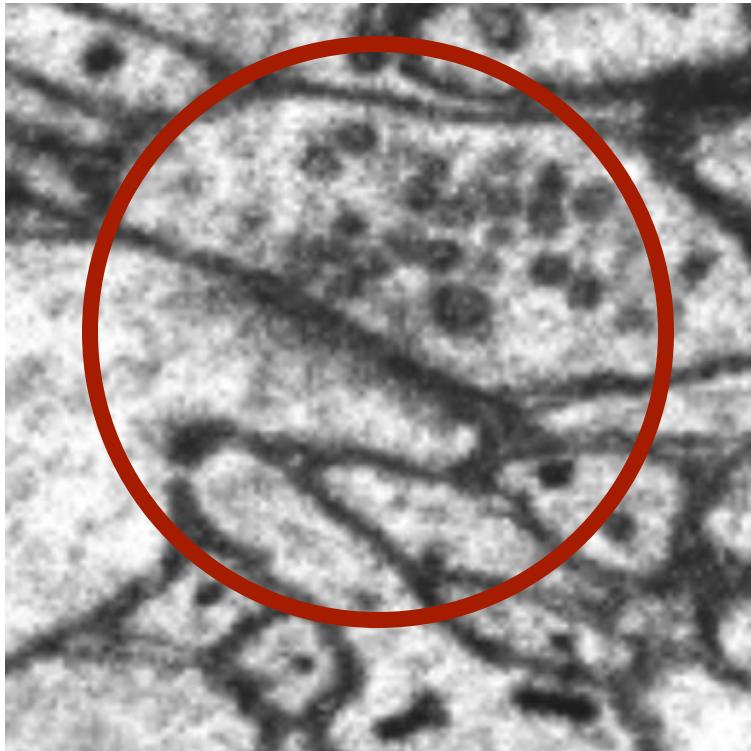


4-10 nanometer resolution

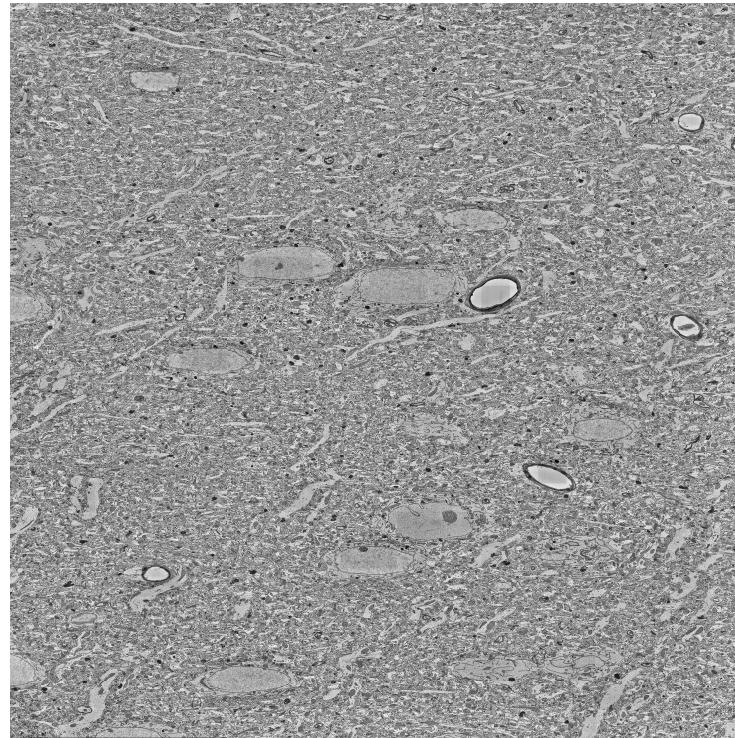


100 micrometers span

Four Orders of Magnitude Difference



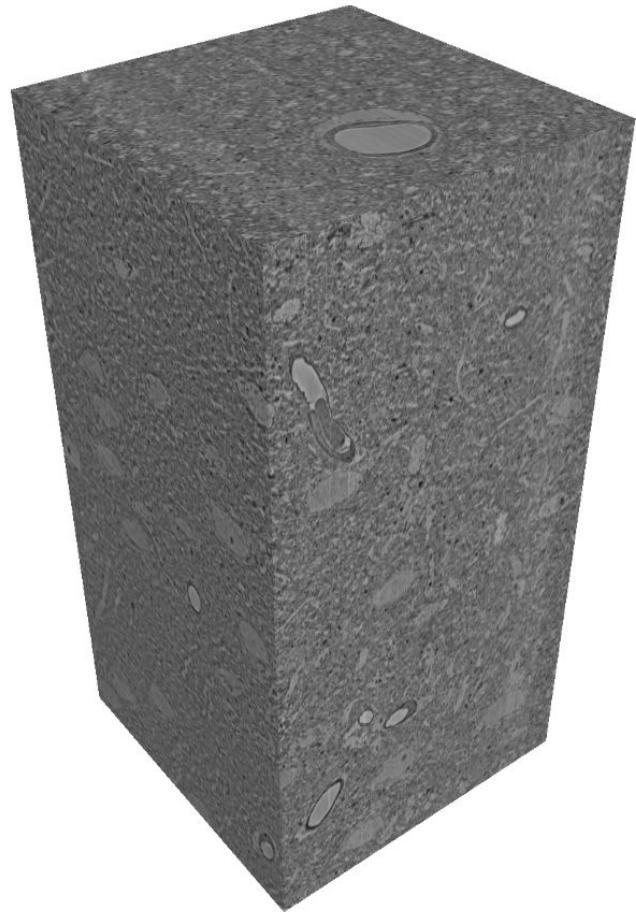
4-10 nanometer resolution



100 micrometers span

Scale Challenges of Connectomics

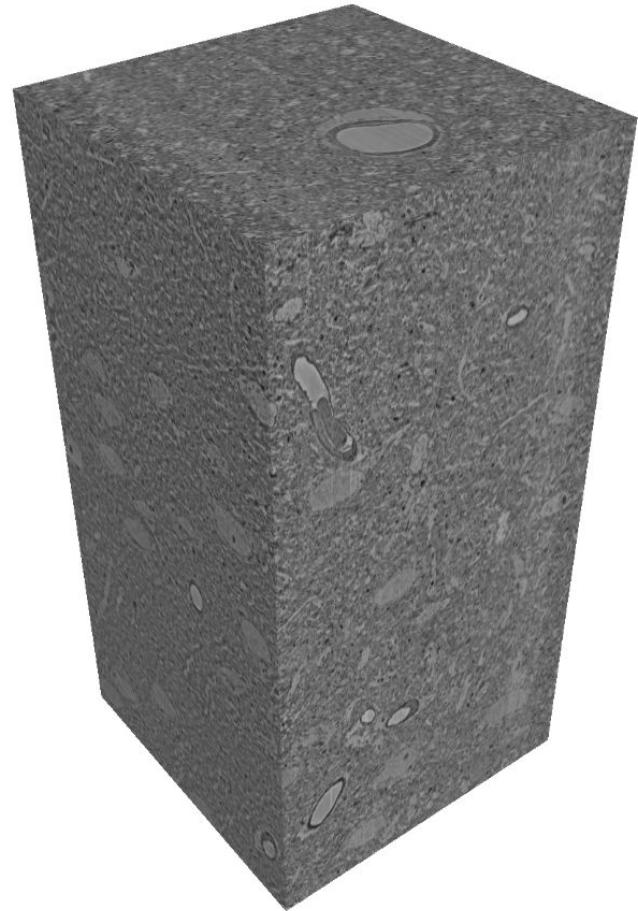
Scales compound with three dimensions



Scale Challenges of Connectomics

Scales compound with three dimensions

Small circuits with ~100-400 neurons require one trillion voxels

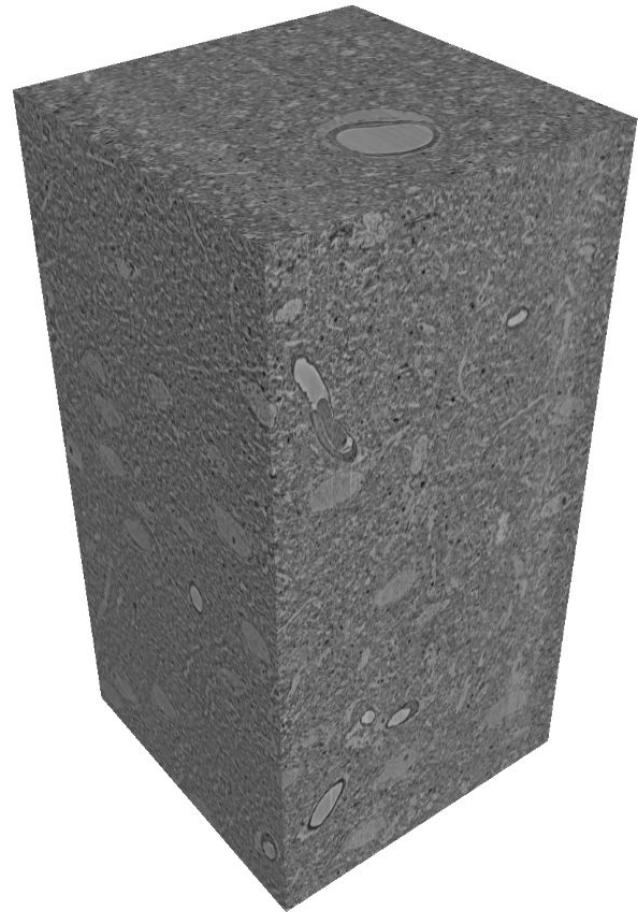


Scale Challenges of Connectomics

Scales compound with three dimensions

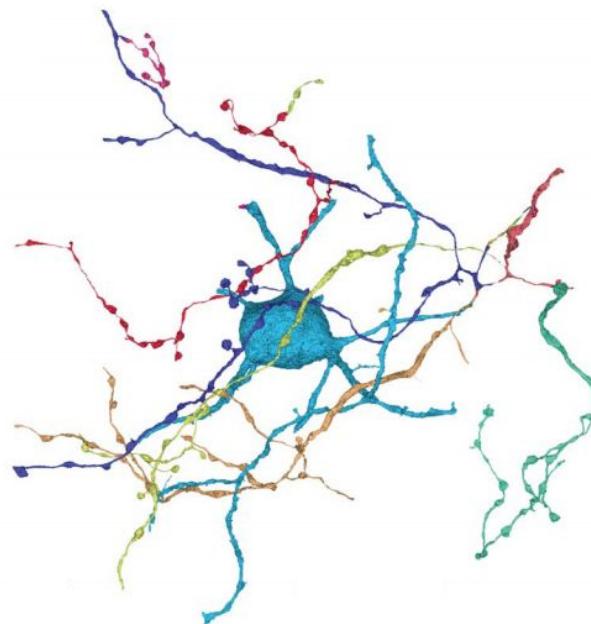
Small circuits with ~100-400 neurons require one trillion voxels

We require automatic solutions for reconstruction and analysis



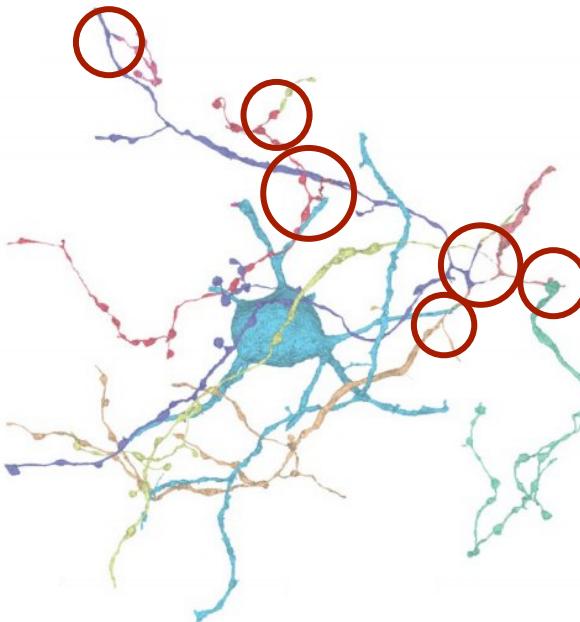
End-to-End Deep Learning

Increasingly, end-to-end trained deep learning models reconstruct the neurons



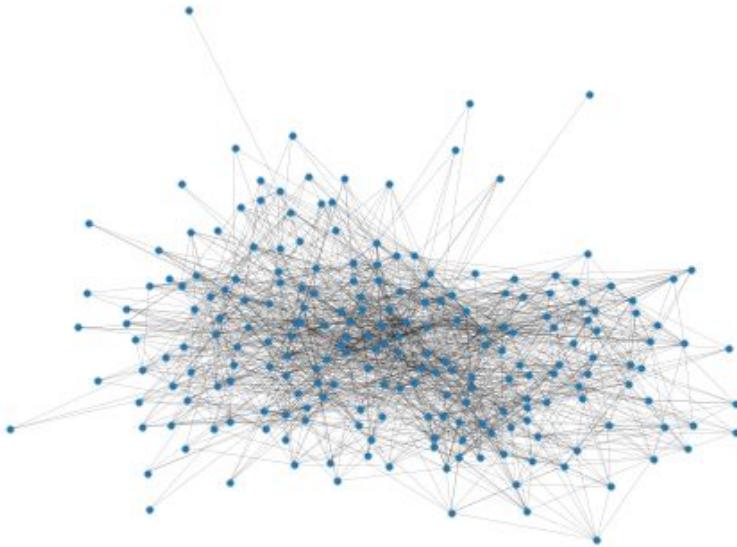
Errors with Automatic Methods

Despite incredible accuracies, these methods make mistakes at such large scales



Dense Graphs

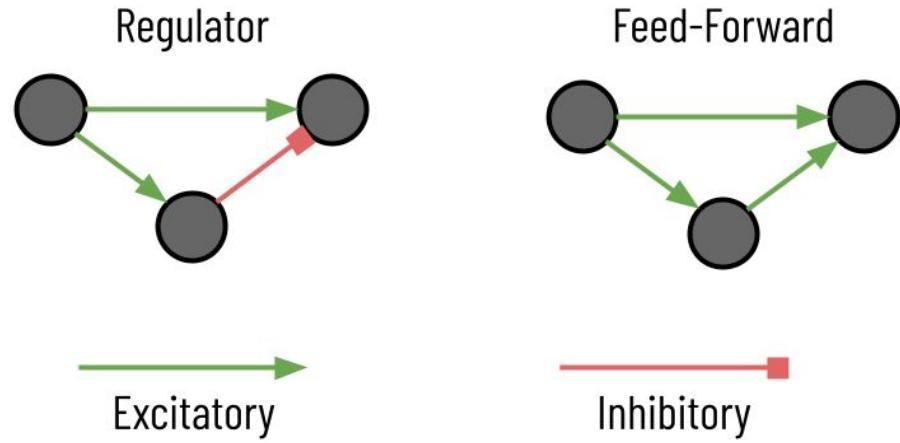
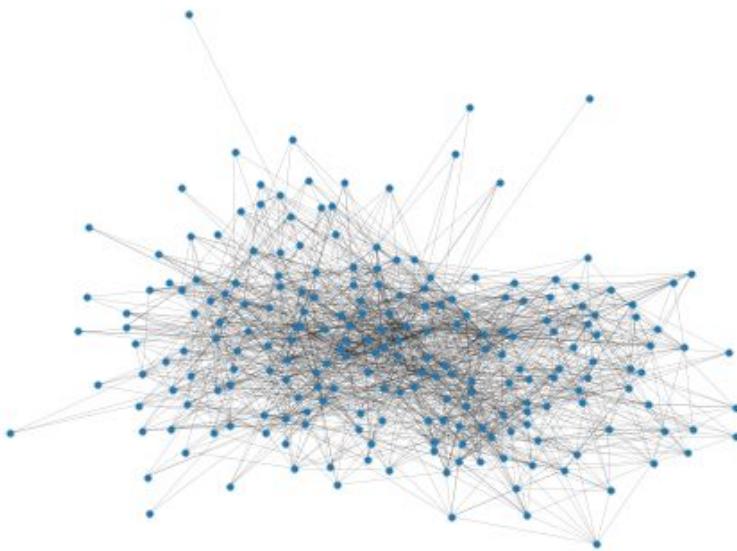
After reconstruction, the extracted wiring diagrams can be quite dense (20-100 connections per neuron)



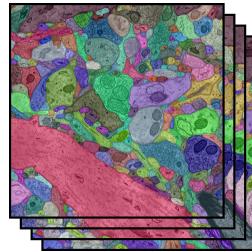
Dense Graphs with Biological Attributes

After reconstruction, the extracted wiring diagrams can be quite dense (20-100 connections per neuron)

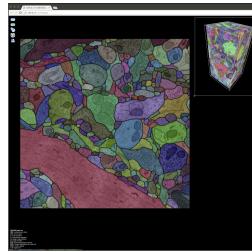
The connections themselves have biological significance such as excitatory/inhibitory synapses



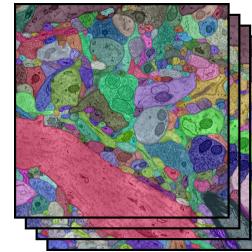
Biologically-Aware Algorithms Along the Connectomics Pipeline



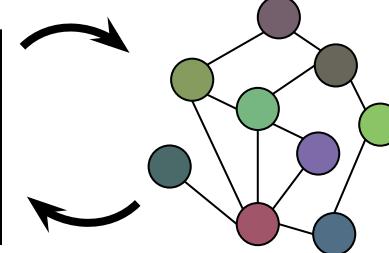
Segmentation



Proofreading

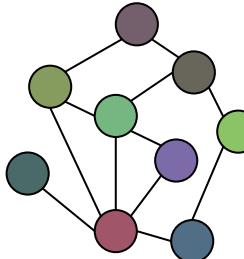
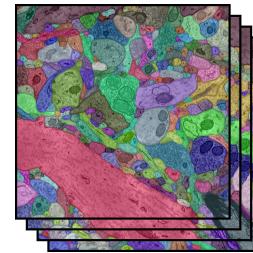
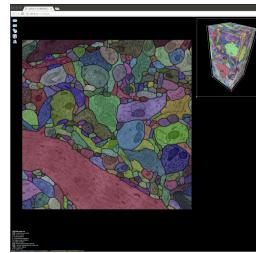
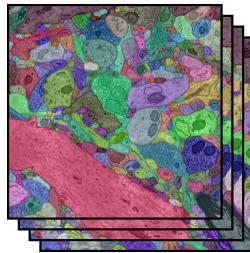


Network Analysis



Biologically-Aware Algorithms Along the Connectomics Pipeline

Compression

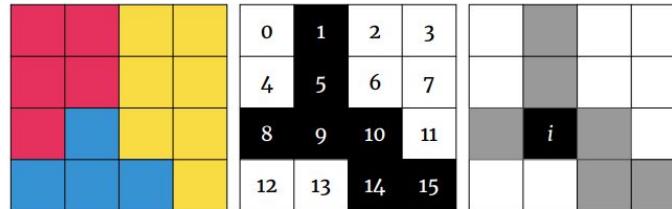
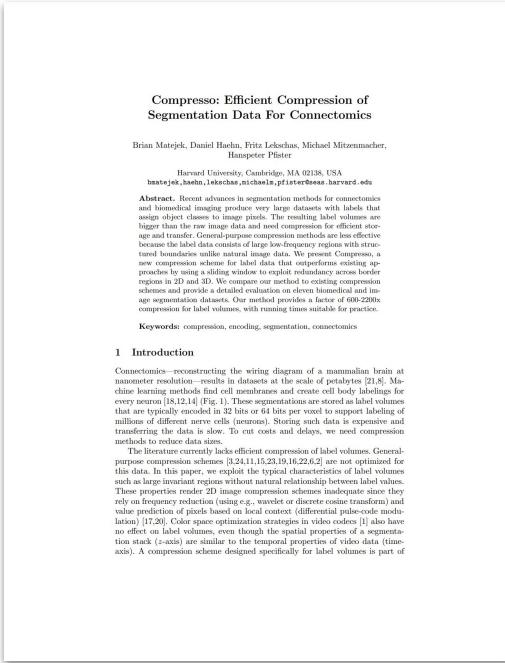


Segmentation

Proofreading

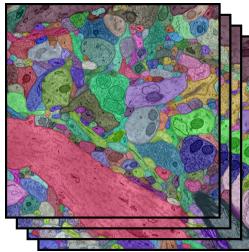
Network Analysis

Compresso: Efficient Compression of Segmentation Data for Connectomics

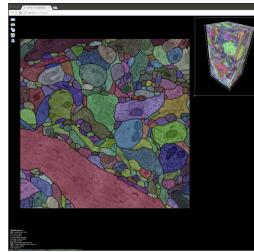


Biologically-Aware Algorithms Along the Connectomics Pipeline

Compression



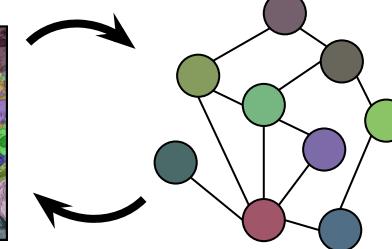
Error Correction



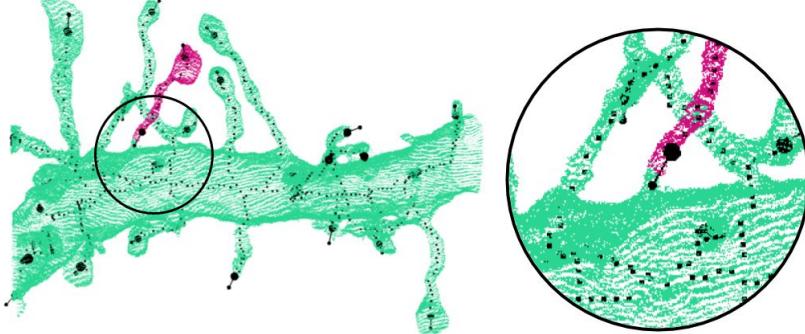
Segmentation

Proofreading

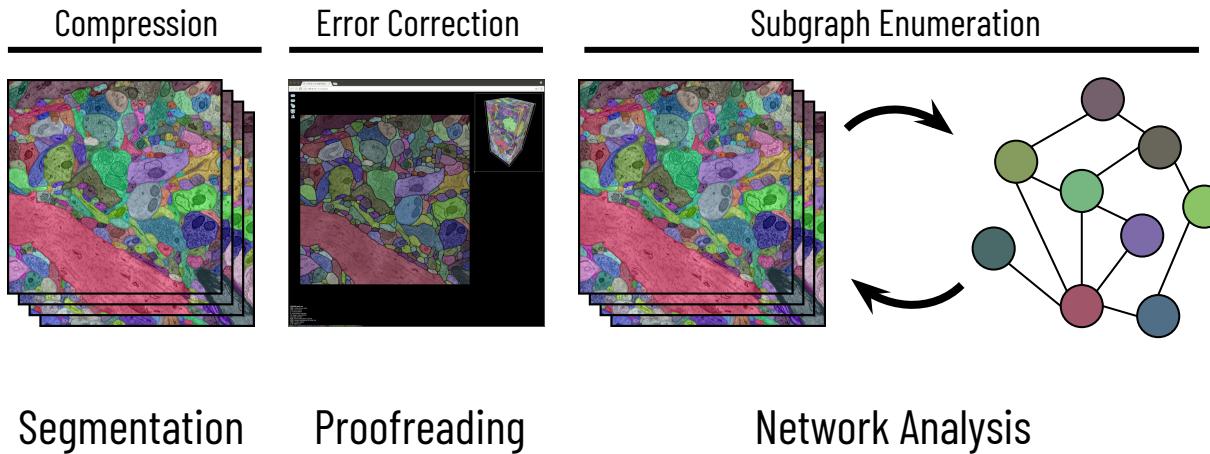
Network Analysis



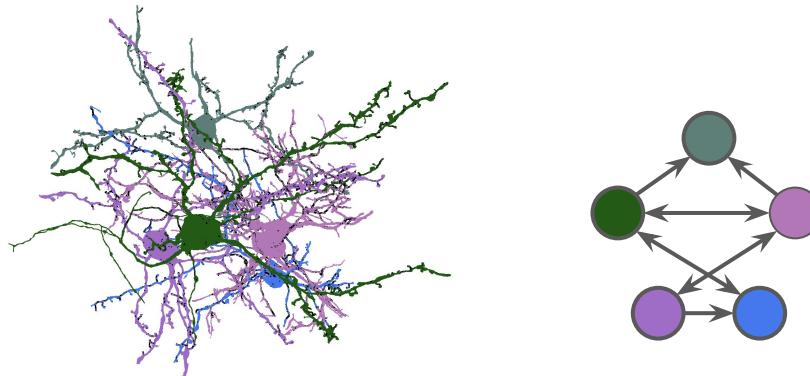
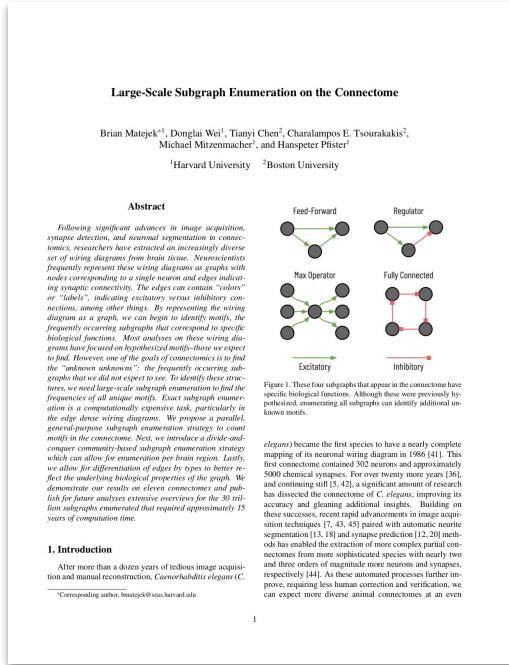
Biologically-Constrained Graphs for Global Connectomics Reconstruction



Biologically-Aware Algorithms Along the Connectomics Pipeline

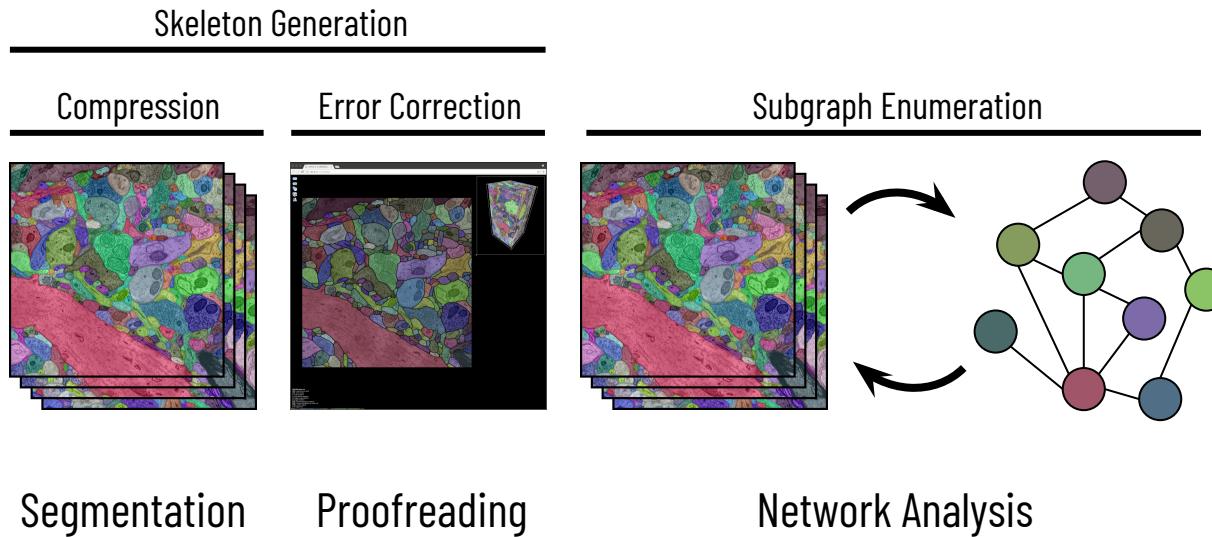


Large-Scale Subgraph Enumeration on the Connectome

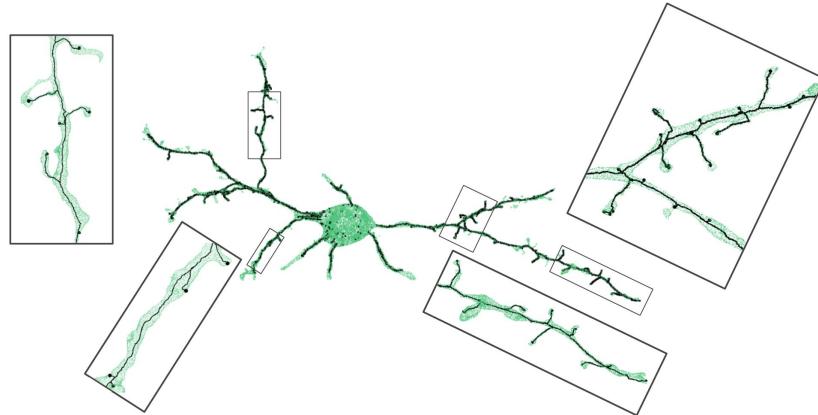


Brian Matejek, Donglai Wei, Tianyi Chen, Charalampos E. Tsourakakis, Michael Mitzenmacher, and Hanspeter Pfister
Preprint

Biologically-Aware Algorithms Along the Connectomics Pipeline



Synapse-Aware Skeleton Generation for Neural Circuits



Synapse-Aware Skeleton Generation for Neural Circuits

Brian Matejek¹, Donglai Wei¹, Xueying Wang², Jinglin Zhao²,
Kálmán Palágyi¹, Hanspeter Pfister¹

¹ John A. Paulson School of Engineering and Applied Sciences
Harvard University, Cambridge, MA, USA

² Center for Brain Science, Department of Molecular Cellular Biology
Harvard University, Cambridge, MA, USA

³ Department of Image Processing and Computer Graphics
University of Szeged, Szeged, Hungary
bmatejek@sees.harvard.edu

Abstract. Reconstructed terabyte and petabyte electron microscopy image volumes are multi-segmented at resolutions fine enough to identify every synaptic connection. After manual or automated reconstruction, neuroscientists want to extract wiring diagrams and connectivity information to analyze the data at a higher level. Despite significant advances in image acquisition, neuron segmentation, and synapse detection tools, the connectivity diagram generation step remains challenging and often do not take into account the wealth of information in the densely reconstructed volumes. We propose a synapse-aware skeleton generation strategy to transform the volumetric volumes into a schematic, amenable yet abstract form on which neuroscientists can perform biological analysis and modeling. Our strategy uses a novel learning-based topological thinning strategy and guarantees a one-to-one correspondence between skeleton endpoints and synapses while simultaneously generating vital geometric features of the neurons and their connections. We demonstrate our results on three large-scale connectomics datasets and compare against current state-of-the-art skeletonization algorithms.

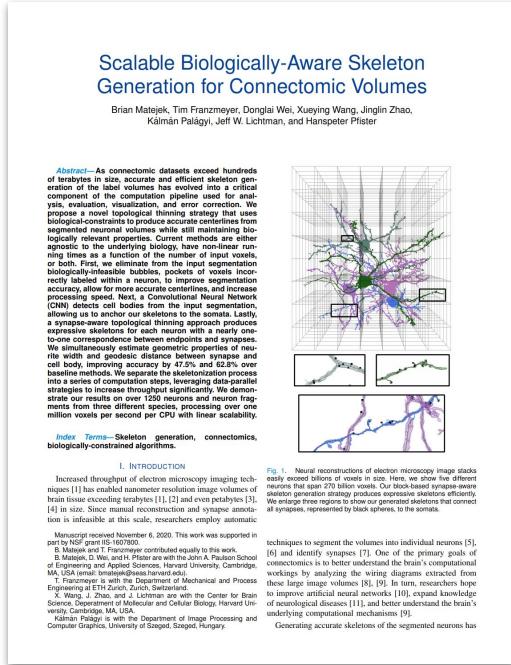
Keywords: neural circuits · connectomics · skeleton generation.

1 Introduction

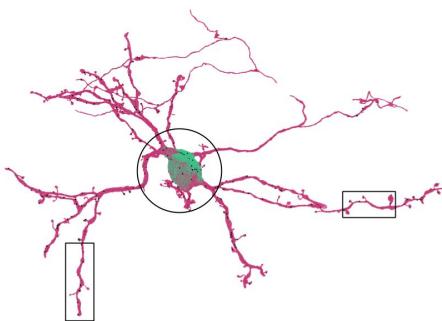
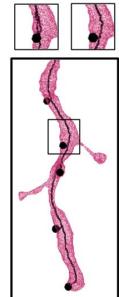
Acquisition techniques [17], automatic segmentation methods [5], and synapse detection strategies [3] in connectomics have all progressed rapidly in the last decade, yielding densely reconstructed volumes at nanometer resolution. These terabyte and petabyte volumes contain hundreds of thousands of interconnected neurons and their synaptic connections. Due to the size of the reconstructed 3D volumes, most analysis of this data occurs at a very coarse level [4].

Little research has focused on generating accurate wiring diagrams from the raw reconstructions. Current approaches [5] directly use an off-the-shelf skeletonization method to reduce these volumes into a series of nodes (neurons) and

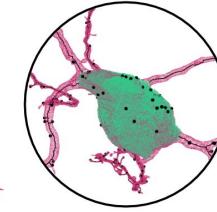
Scalable Biologically-Aware Skeleton Generation for Connectomic Volumes



Bubble Filling



Soma Detection



Synapse Connectivity

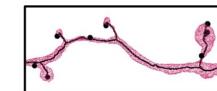


Fig. 1. Neural reconstructions of electron microscopy image stacks can exceed billions of voxels in size. Here, we show five different brain tissue regions. Our novel, biologically-aware skeleton generation strategy produces expressive skeletons efficiently. We enlarge three regions to show our generated skeletons that connect all synapses, represented by black spheres, to the soma.

INTRODUCTION

Increased interest of electron microscopy imaging techniques [1] has enabled connectomic reconstruction of brain tissue exceeding terabytes [1], [2] and even petabytes [3], [4] in size. Since manual reconstruction and synapse annotation is infeasible at this scale, researchers employ automatic, biologically-constrained algorithms.

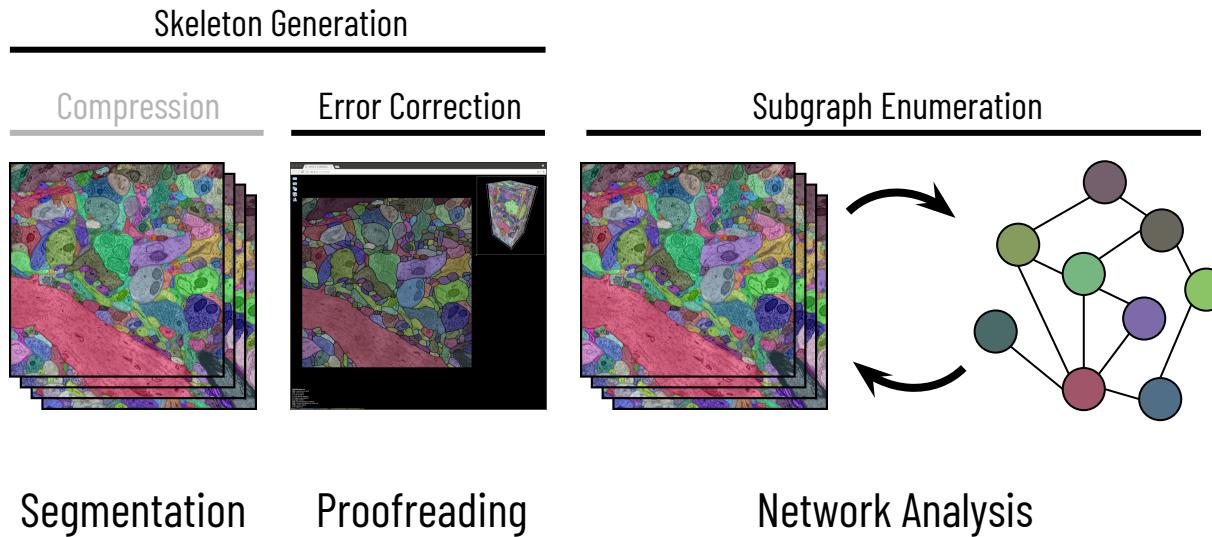
I. INTRODUCTION

Manuscript received November 6, 2020. This work was supported in part by grants from the National Science Foundation (NSF) (B. Matejek and T. Franzmeyer contributed equally to this work). B. Matejek and T. Franzmeyer contributed equally to this work. B. Matejek is with the Center for Brain Science, Department of Engineering and Applied Sciences, Harvard University, Cambridge, MA, USA (email: bmatejek@seas.harvard.edu). T. Franzmeyer is with the Department of Mechanical and Process Engineering at ETH Zurich, Zurich, Switzerland.

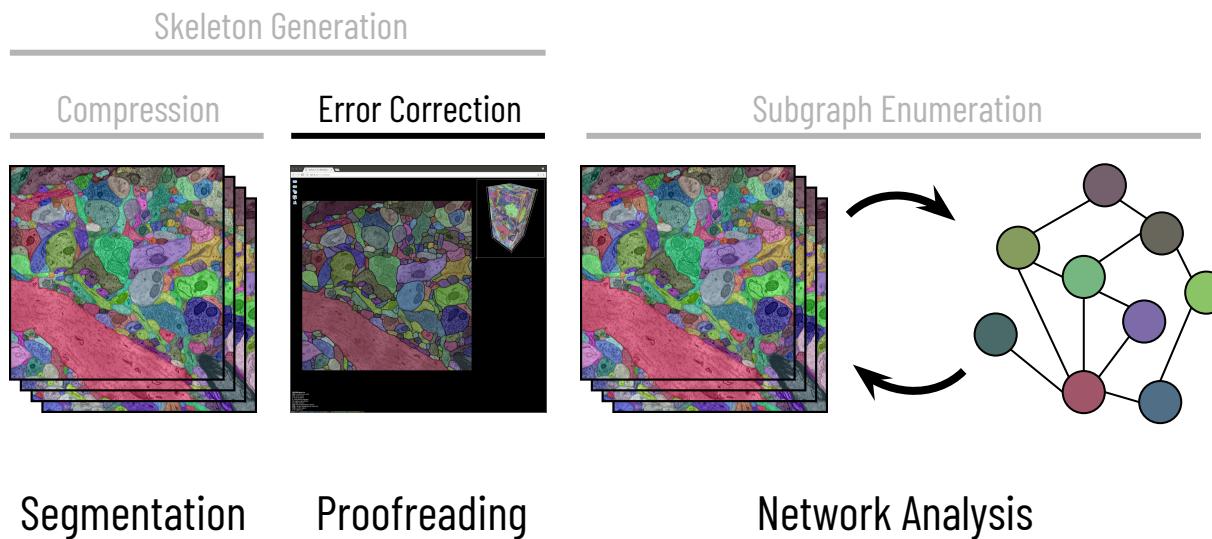
A. Wei, X. Wang, J. Zhao, and H. Pfister are with the Center for Brain Science, Department of Molecular and Cellular Biology, Harvard University, Cambridge, MA, USA (email: xueying.wang@mcg.harvard.edu). Kálmán Palágyi is with the Department of Image Processing and Computer Graphics, University of Szeged, Szeged, Hungary.

Generating accurate skeletons of the segmented neurons has techniques to segment the volumes into individual neurons [5], [6] and identify synapses [7]. One of the primary goals of connectomics is to better understand the brain's computational workings by analyzing the weight diagrams extracted from the neural connections [8], [9]. These findings hope to improve artificial neural networks [10], expand knowledge of neurological diseases [11], and better understand the brain's underlying computational mechanisms [9].

Biologically-Aware Algorithms Along the Connectomics Pipeline



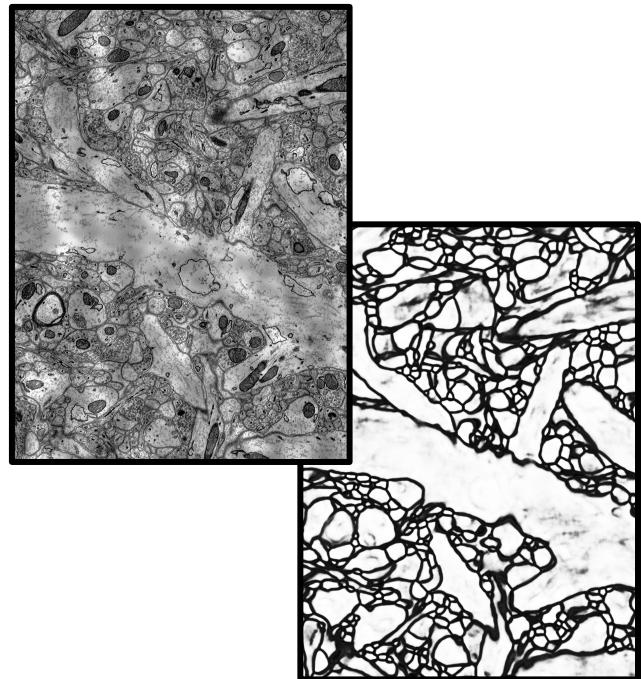
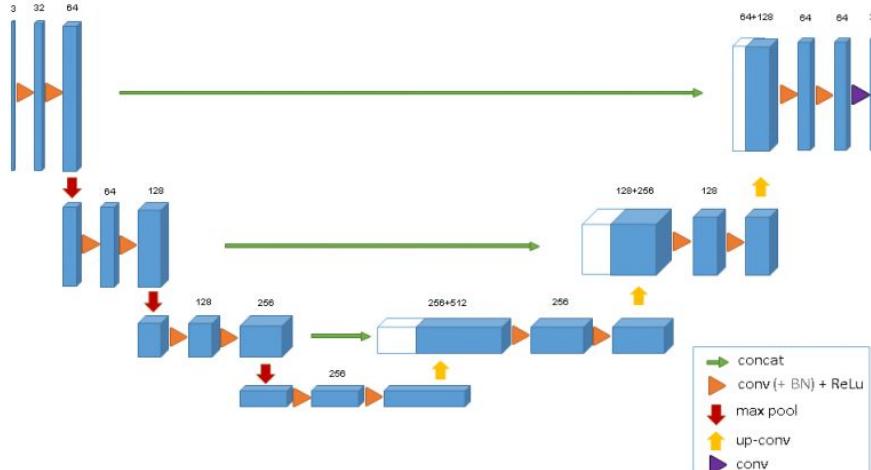
Biologically-Aware Algorithms Along the Connectomics Pipeline



Biologically-Constrained Graphs for Global Connectomics Reconstruction

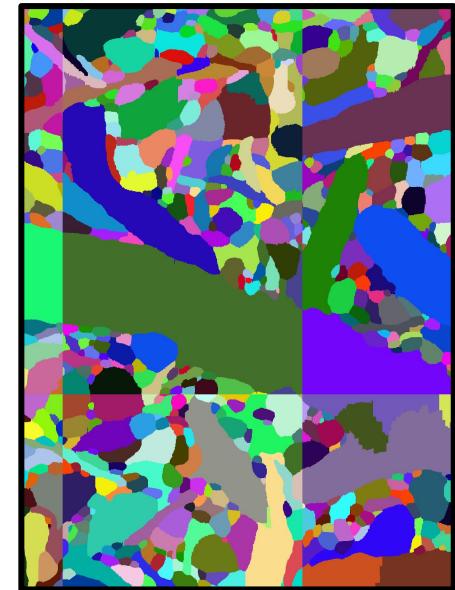
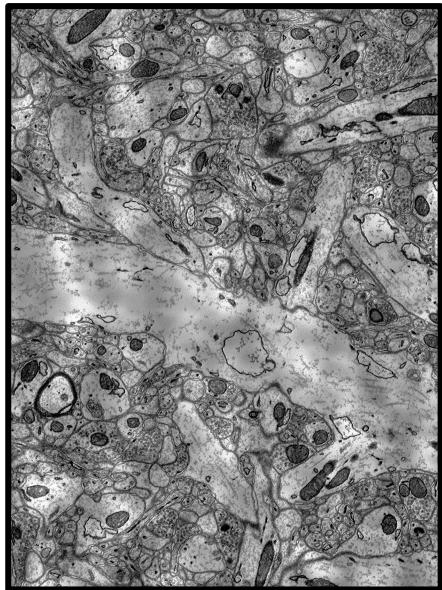
Brian Matejek, Daniel Haehn, Haidong Zhu,
Donglai Wei, Toufiq Parag, and Hanspeter Pfister

Affinity Generation



Ronneberger *et al.*, U-Net: Convolutional Networks for Biomedical Image Segmentation, MICCAI 2015
Cicek *et al.*, 3D U-Net: Learning Dense Volumetric Segmentation from Sparse Annotation, MICCAI 2016

3D Watershed on Affinities



Zlateski et al., Image Segmentation by Size-Dependent Single Linkage Clustering of a Watershed Basin Graph, 2015

Funke et al., A Deep Structured Learning Approach Towards Automating Connectome Reconstruction from 3D Electron Micrographs, 2017

Zeng et al., DeepEM3D: Approaching Human-Level Performance on 3D Anisotropic EM Image Segmentation, Bioinformatics 2017

Agglomeration

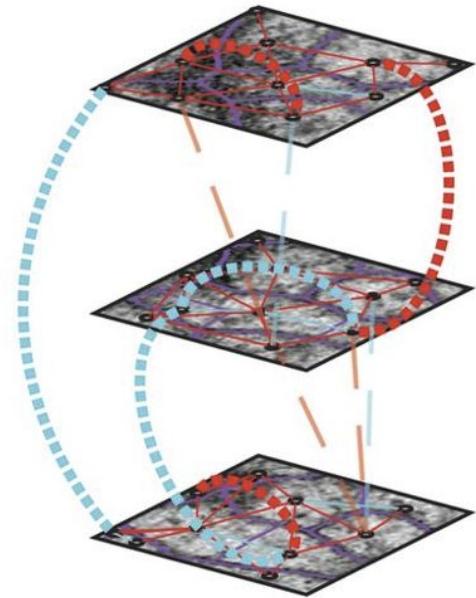
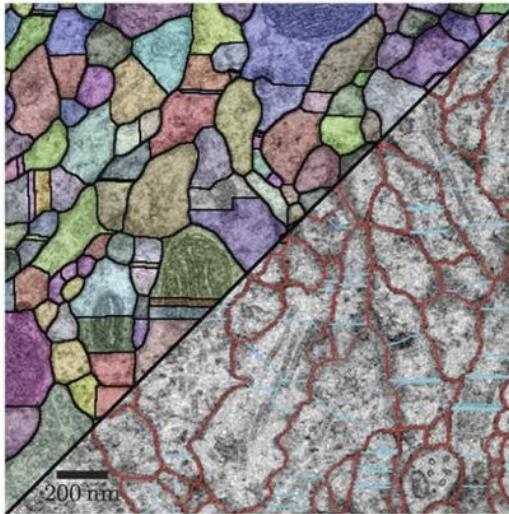
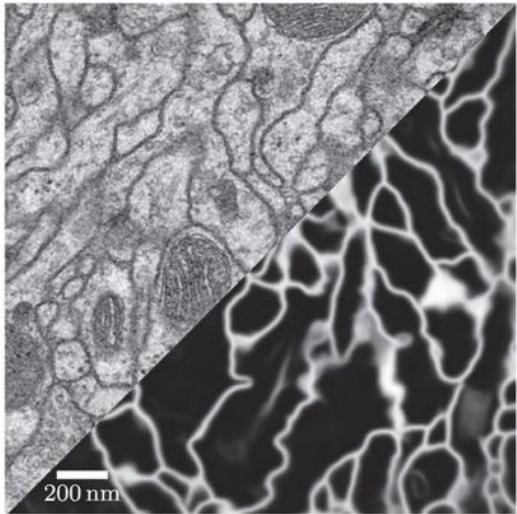


Nunez-Iglesias et al., Machine Learning of Hierarchical Clustering to Segment 2D and 3D Images, PLoS ONE, 2013

Parag et al., A Context-Aware Delayed Agglomeration Framework for Electron Microscopy Segmentation, PLoS ONE 2015

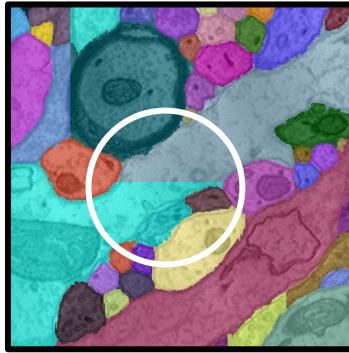
Funke et al., A Deep Structured Learning Approach Towards Automating Connectome Reconstruction from 3D Electron Micrographs, 2017

Lifted Multicuts

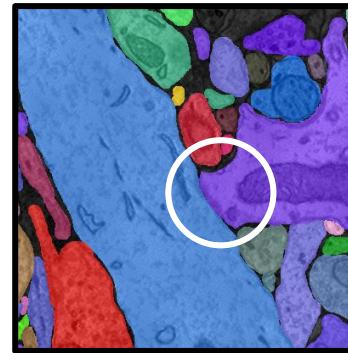
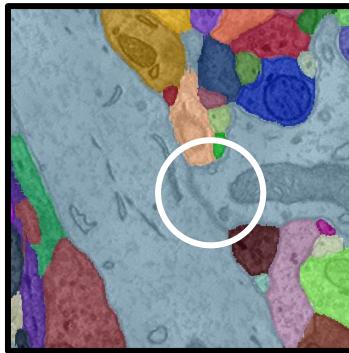
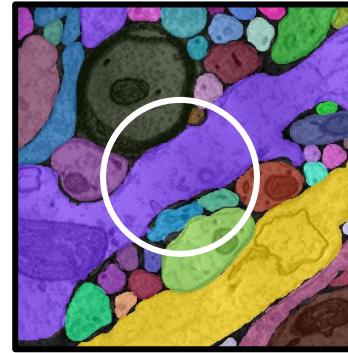


Errors

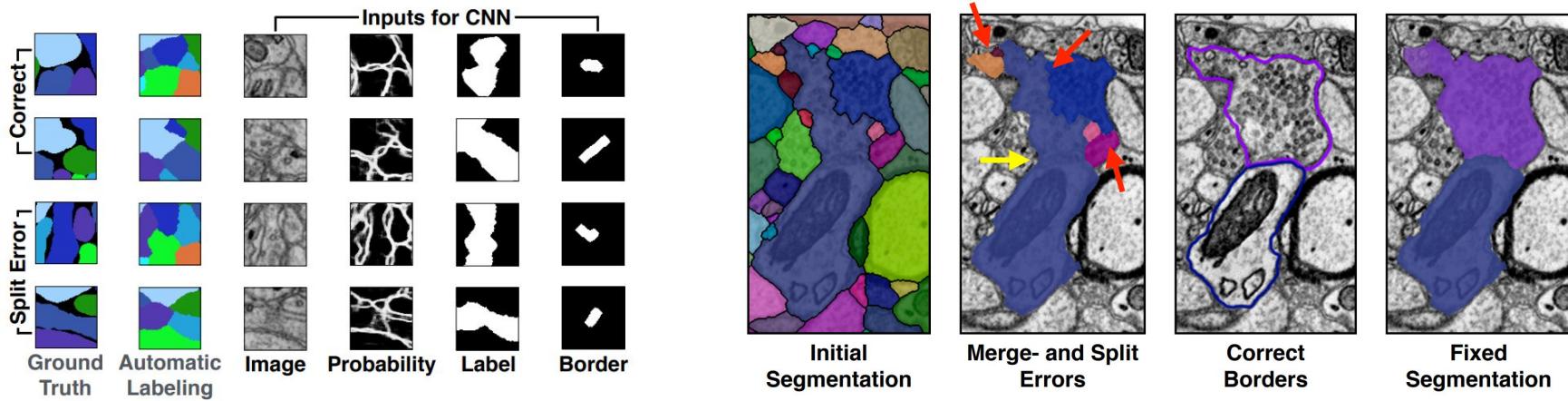
Automatic Segmentation



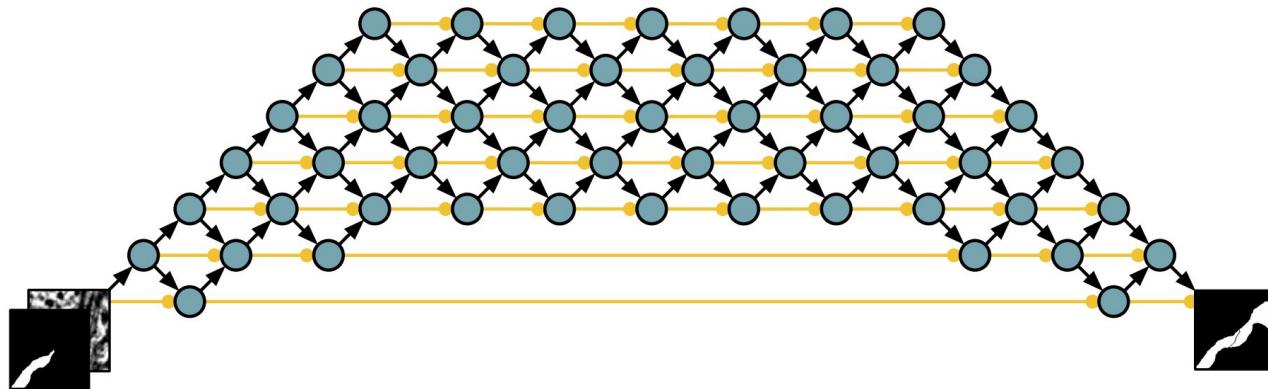
Ground Truth



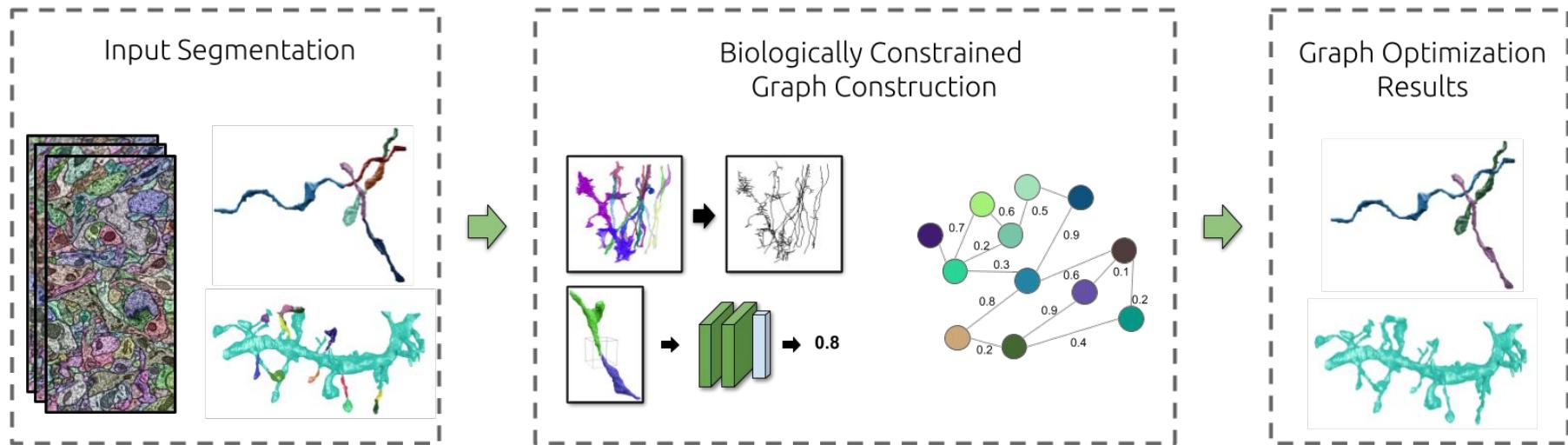
Guided Proofreading



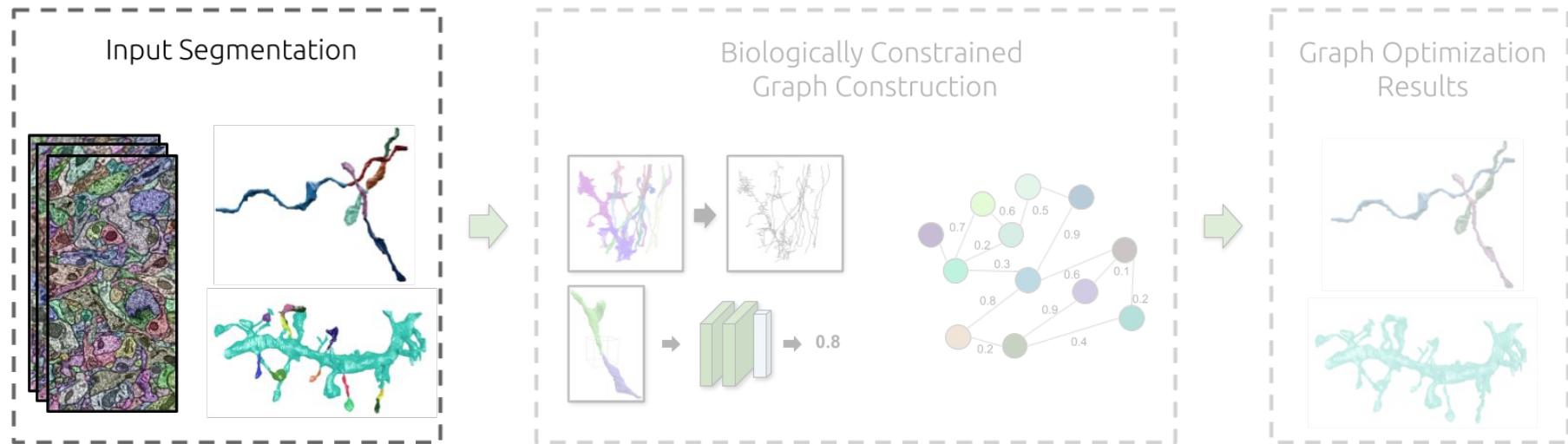
Automatic Proofreading



Proposed Automatic Error Correction

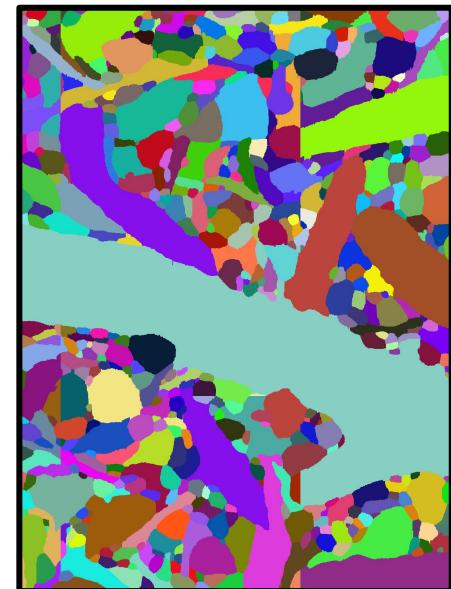
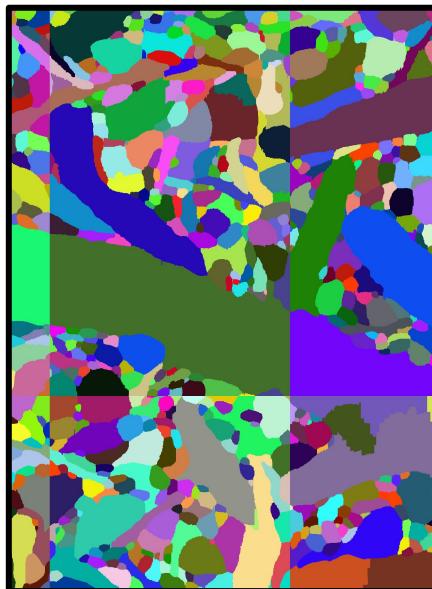


Input



Traditional Two-Stage Frameworks

Existing segmentation strategies typically produce over-segmentations



Affinity Generation

Watershed Transform

Agglomeration

Traditional Two-Stage Frameworks

Existing segmentation strategies typically produce over-segmentations

We use the result from an existing strategy as our input

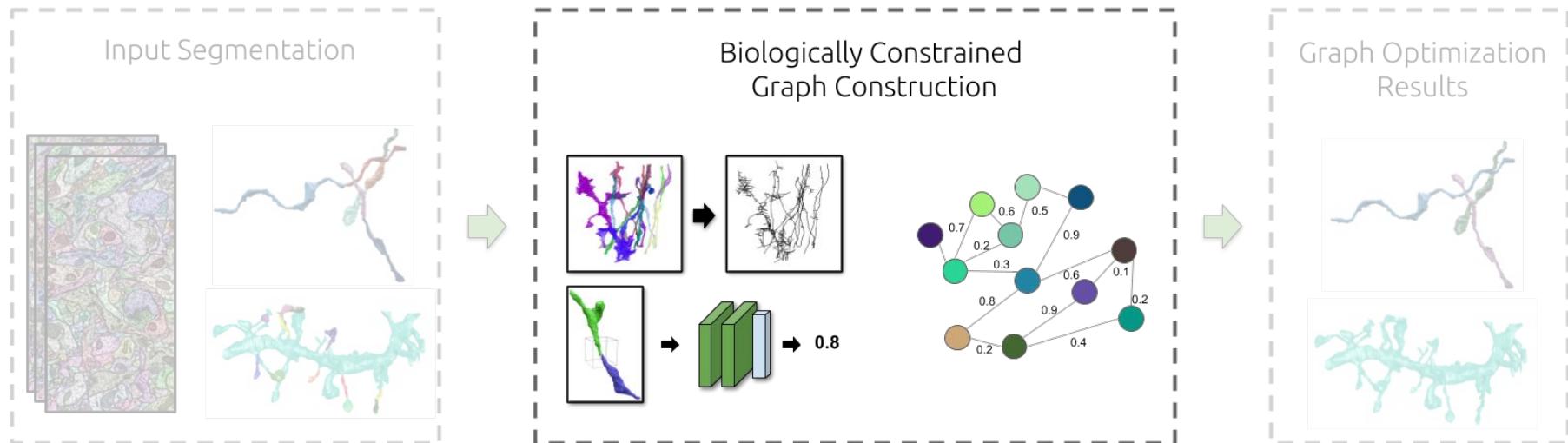
Traditional Two-Stage Frameworks

Existing segmentation strategies typically produce over-segmentations

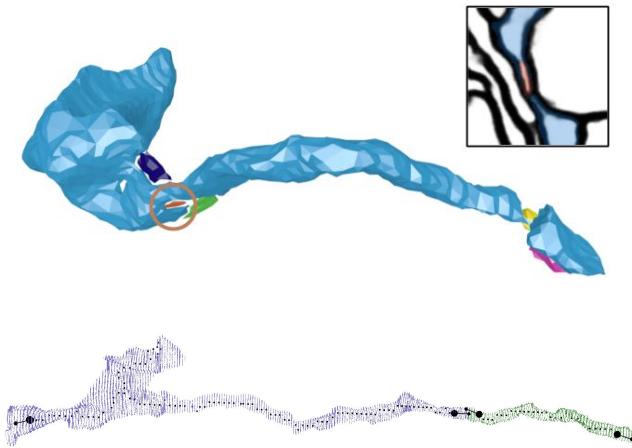
We use the result from an existing strategy as our input

Allows us to leverage larger local context when forming our graph

Goal: Construct a graph with as few nodes and edges as possible

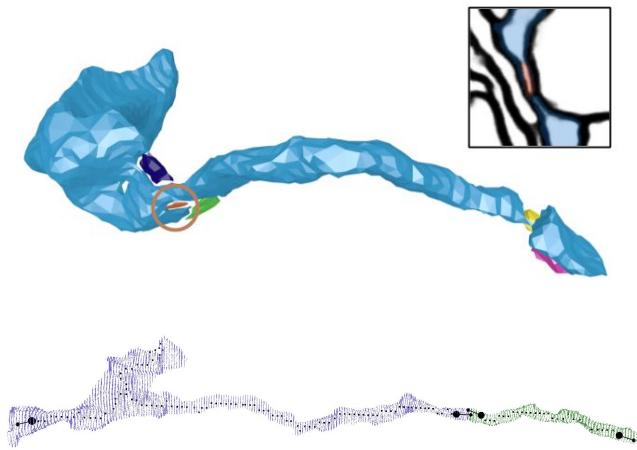


Graph Construction with Biological Constraints

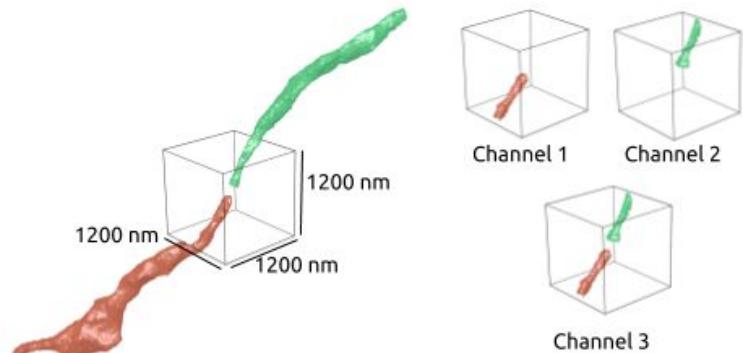


Hand-Designed
Geometric Constraints

Graph Construction with Biological Constraints



Hand-Designed
Geometric Constraints



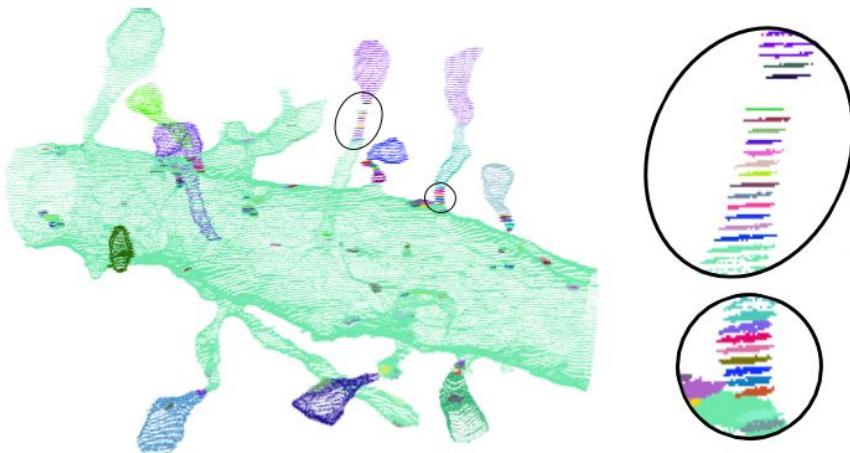
Machine-Learned
Morphologies

Node Generation

Existing segmentation strategies produce a large number of small segments

Node Generation

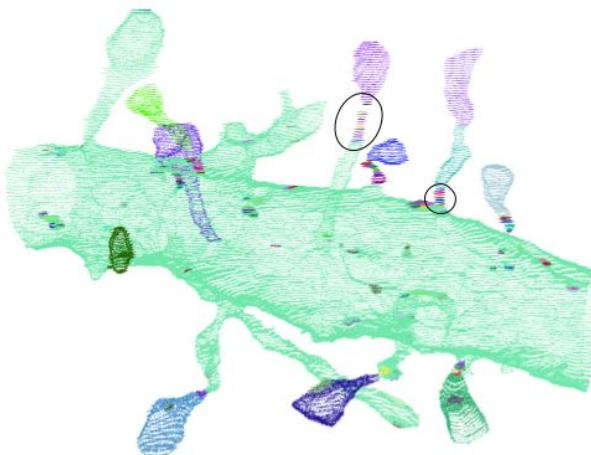
Existing segmentation strategies produce a large number of small segments



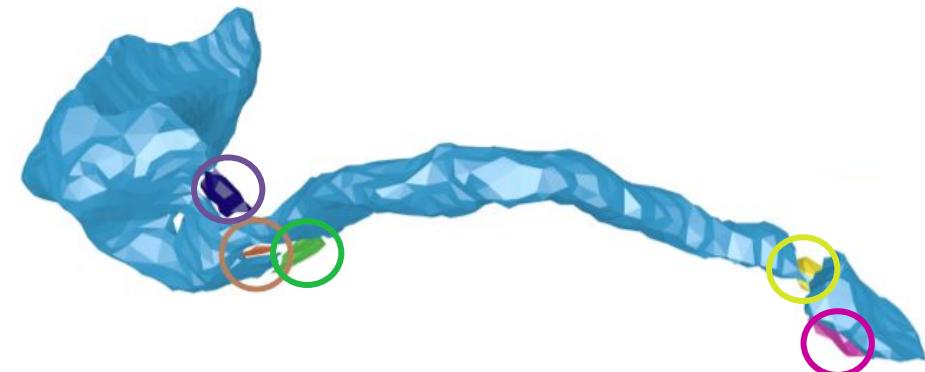
Singleton Slices

Node Generation

Existing segmentation strategies produce a large number of small segments



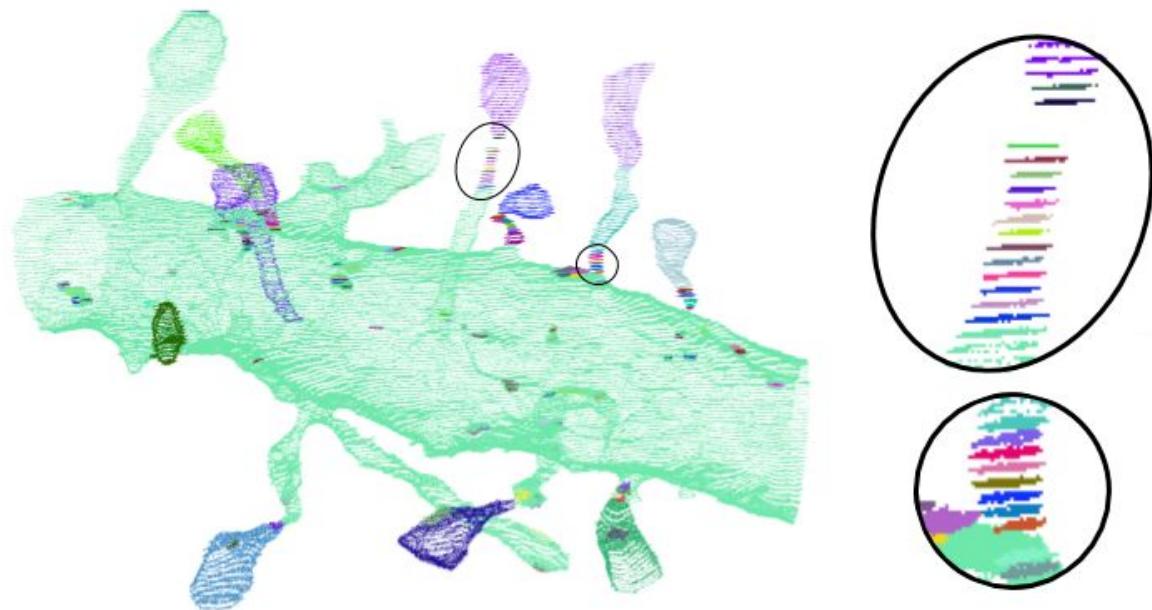
Singleton Slices



5 Small Segments

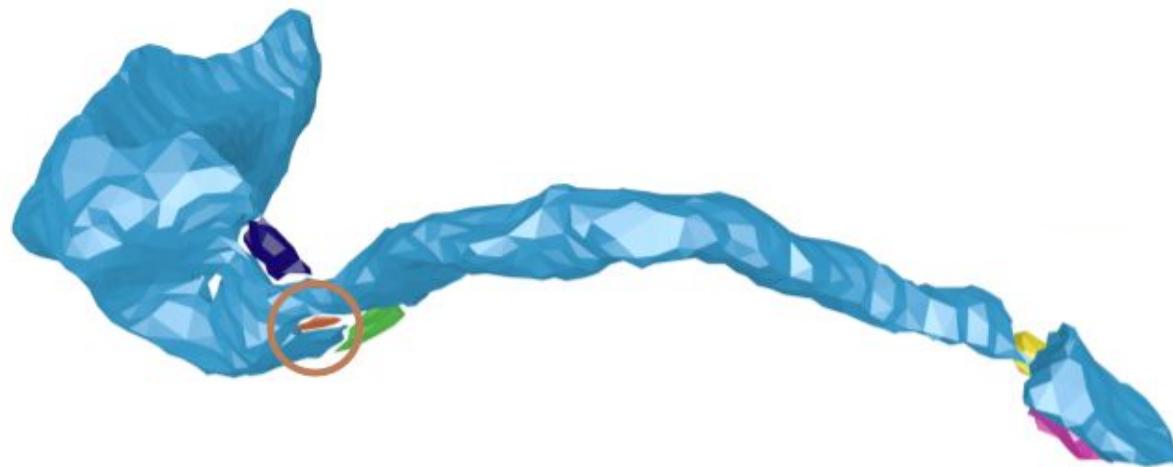
Singleton Removal

Merge adjacent singleton slices that have an Intersection-over-Union above 0.30



Merging Other Small Segments

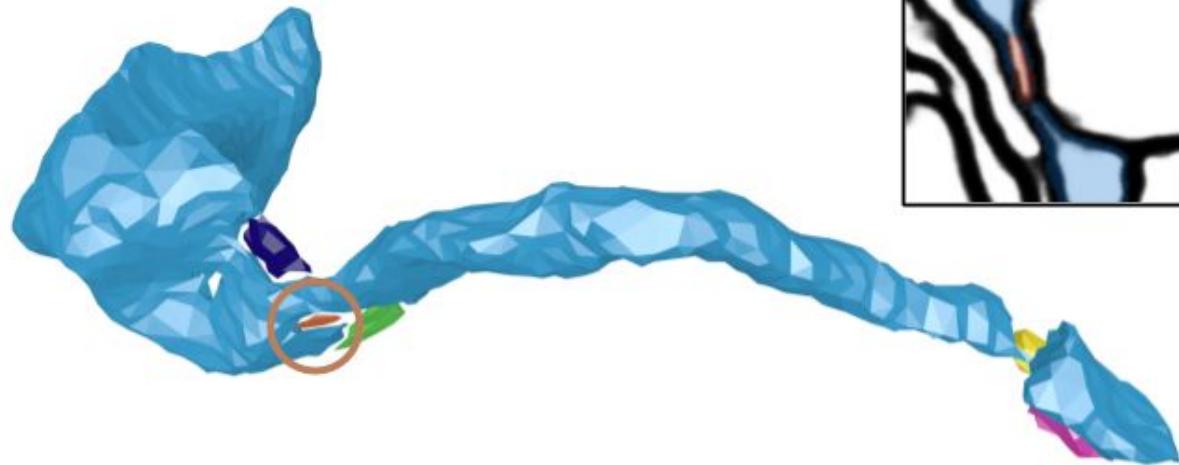
Up to 80% of remaining segments are very small with little shape information



Merging Other Small Segments

Up to 80% of remaining segments are very small with little shape information

These small segments often occur at narrow locations with noisy affinities



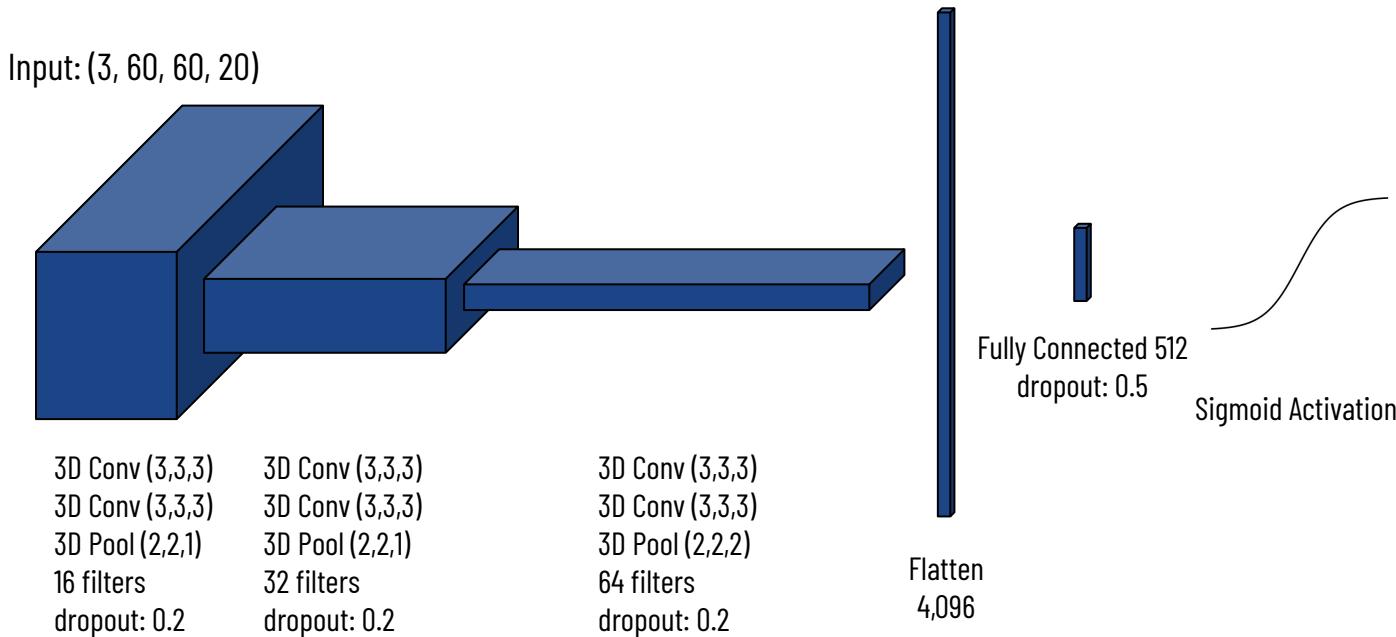
Small Segment Merging

Each small segment is merged with a nearby large segment

Small Segment Merging

Each small segment is merged with a nearby large segment

A 3D CNN predicts the most likely neighbor to belong to the same neuronal process



Edge Generation

Each segment has too many adjacent neighbors to use the adjacency matrix

Edge Generation

Each segment has too many adjacent neighbors to use the adjacency matrix



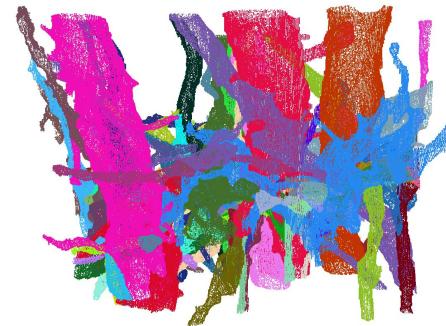
Typical Segment

Edge Generation

Each segment has too many adjacent neighbors to use the adjacency matrix



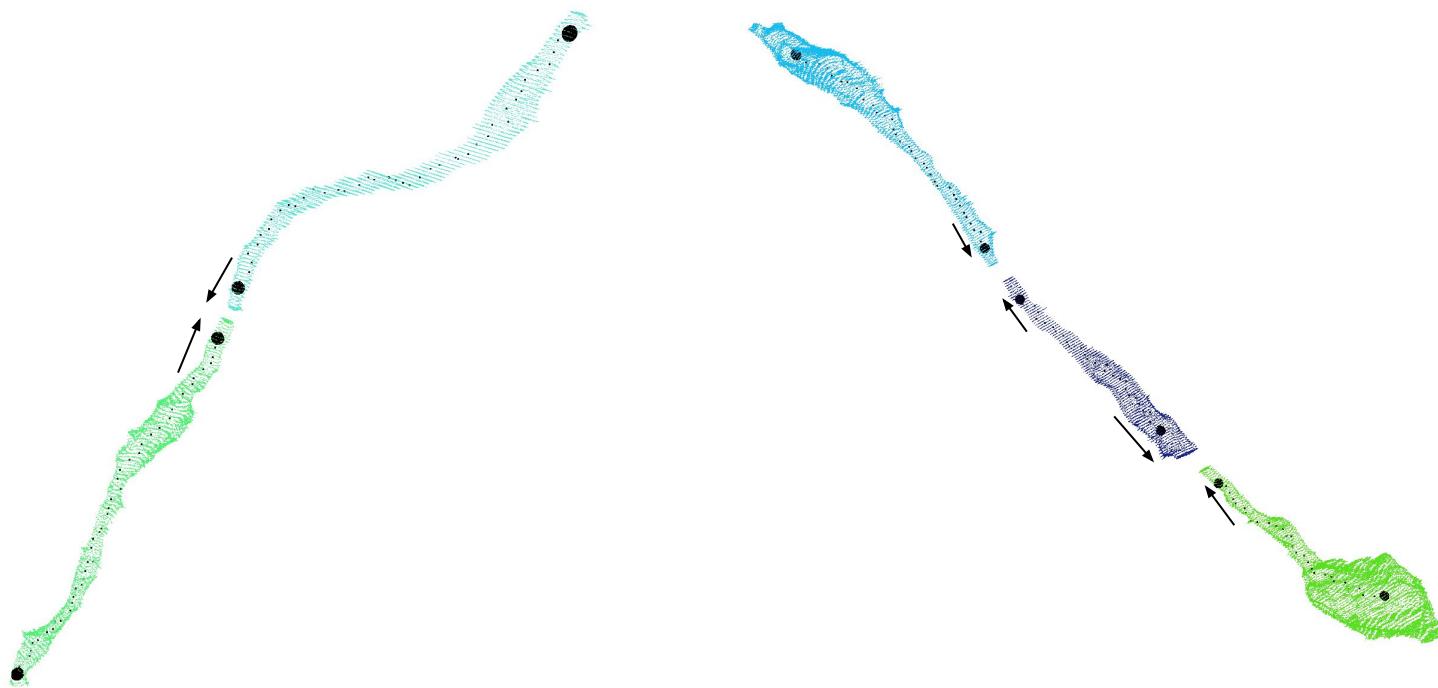
Typical Segment



103 Adjacent Neighbors

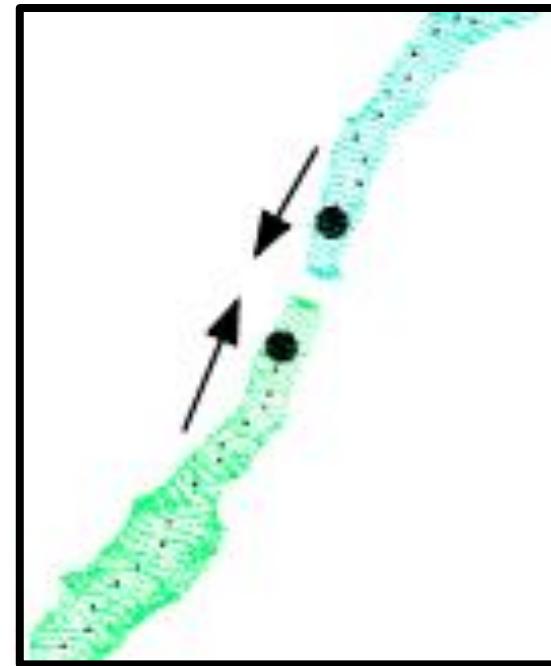
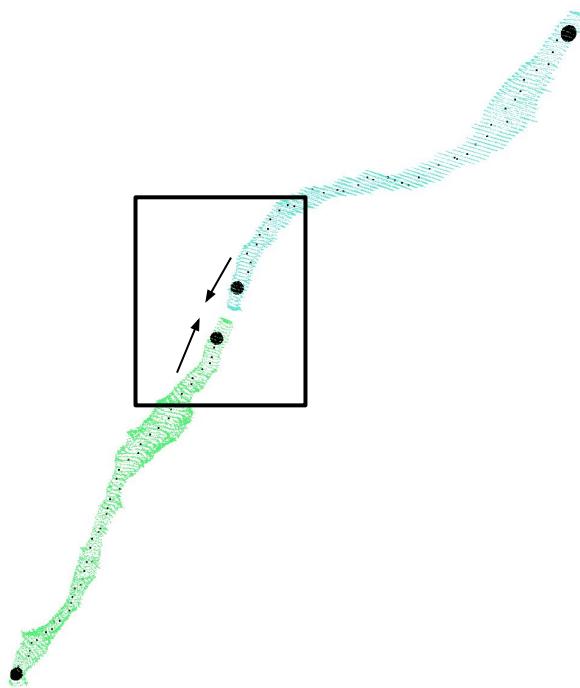
Handcrafted Geometric Constraints

Use directional information to identify potential split error locations



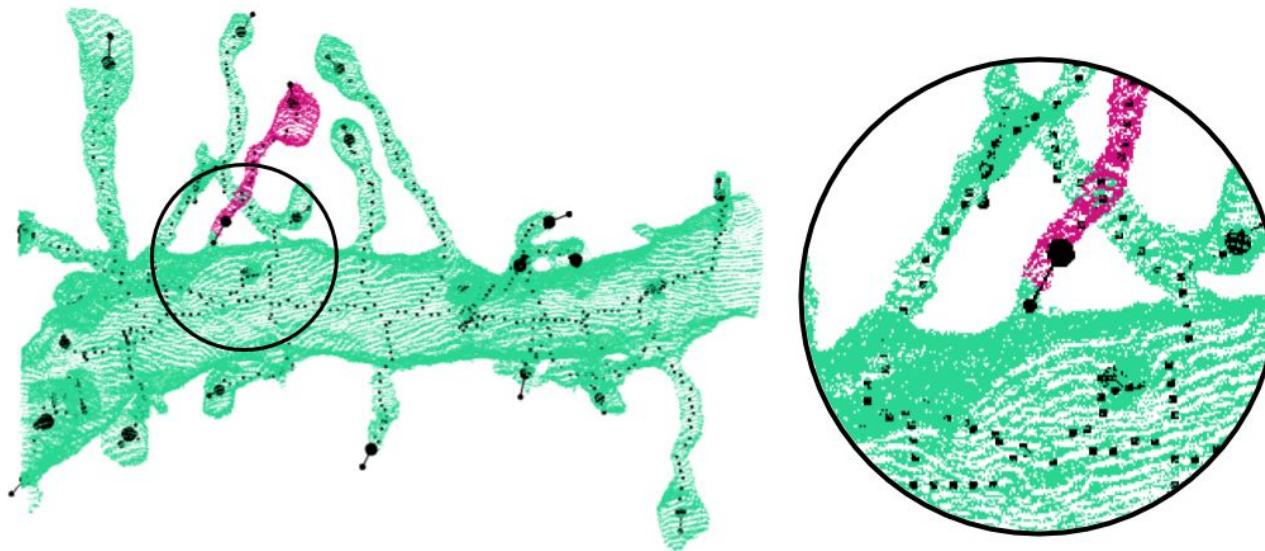
Skeleton Generation

Approximate volume shapes with 1D skeletons and identify potential split errors based on skeletal geometry around the endpoints



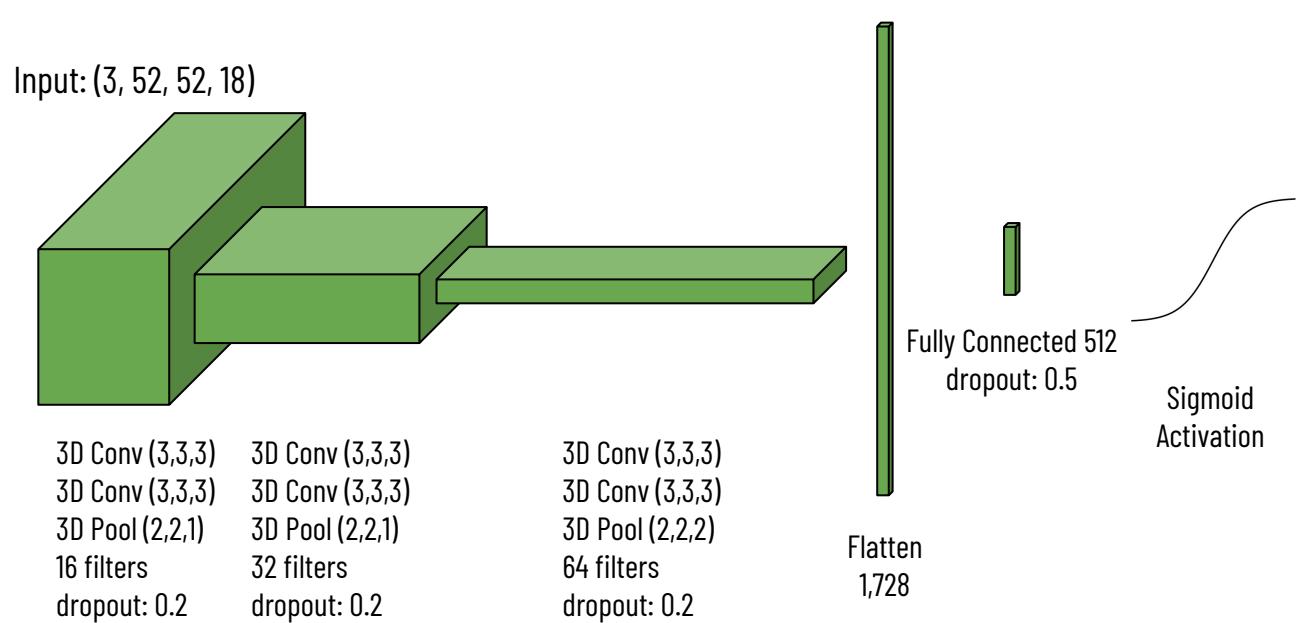
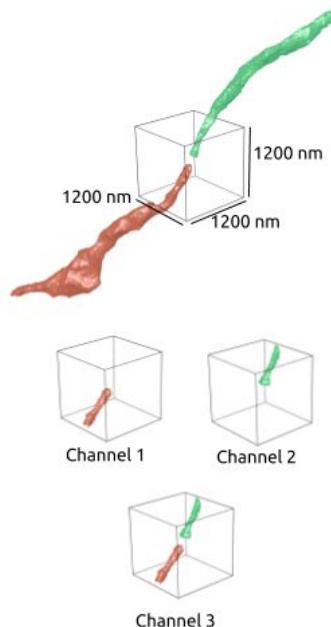
Edge Generation

Two nodes receive an edge in the graph if one of the corresponding skeletons has an endpoint vector towards the other segment



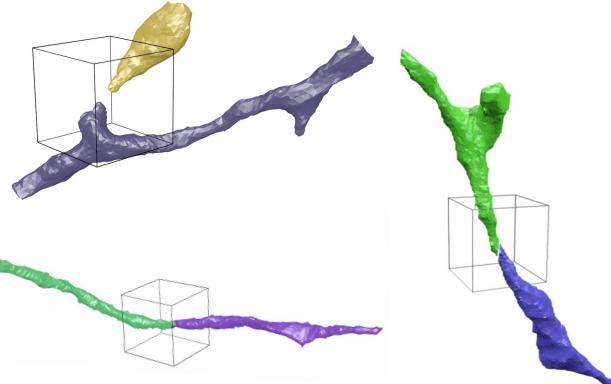
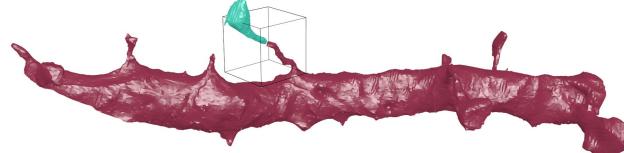
Generating Edge Weights with Machine-learned Morphologies

We train a convolutional neural network to predict if two segments belong to the same neuronal process



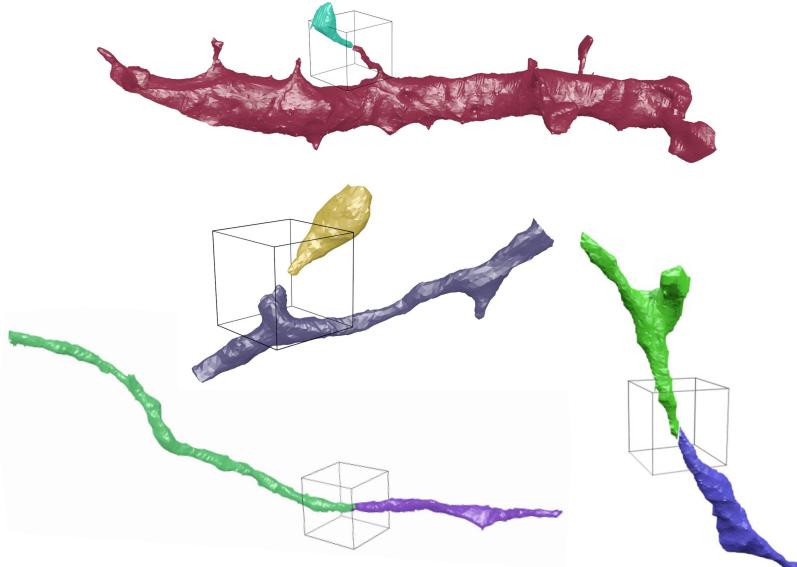
Input Examples

Should Merge

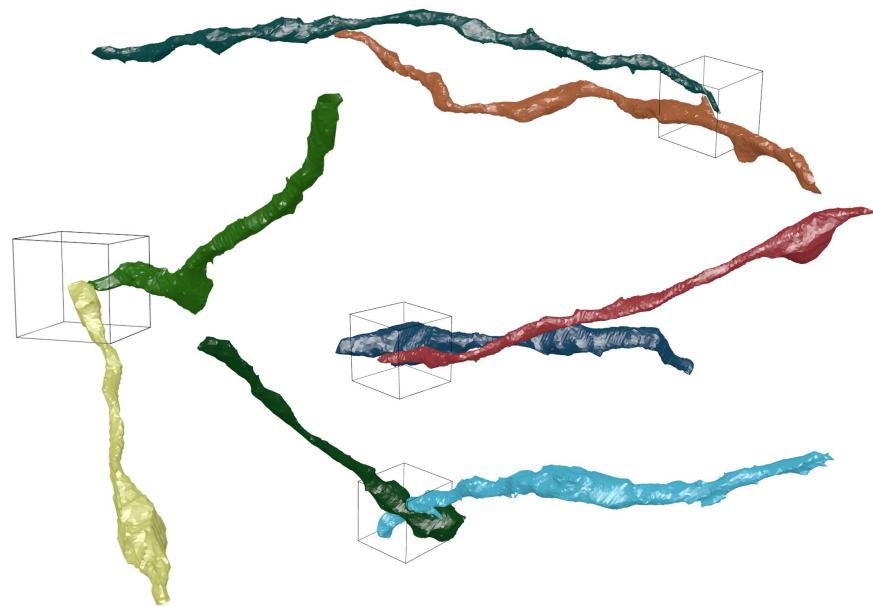


Input Examples

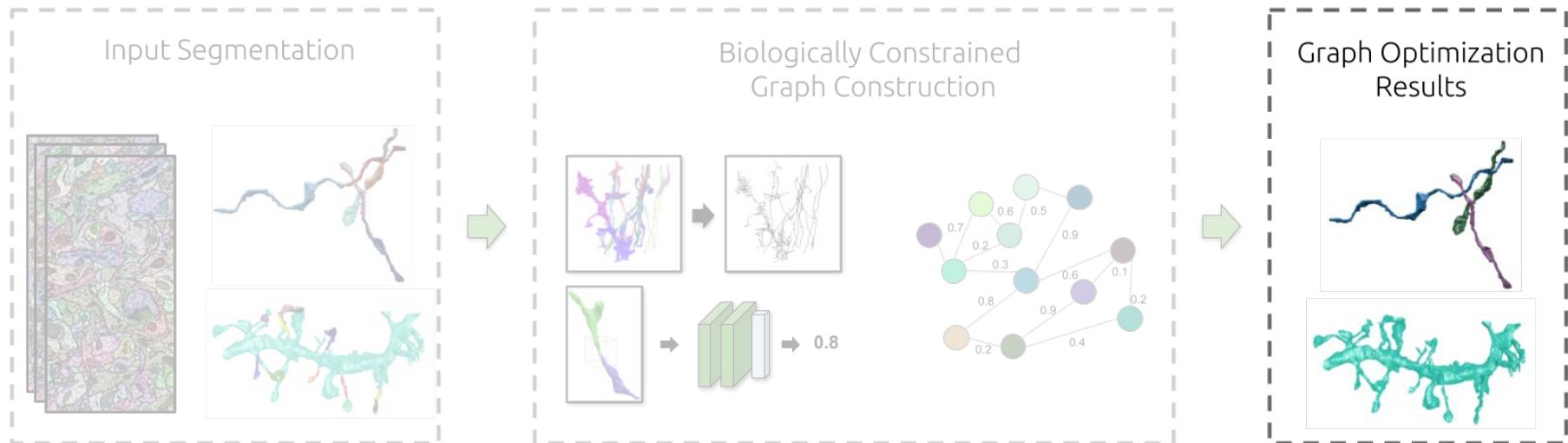
Should Merge



Should Not Merge

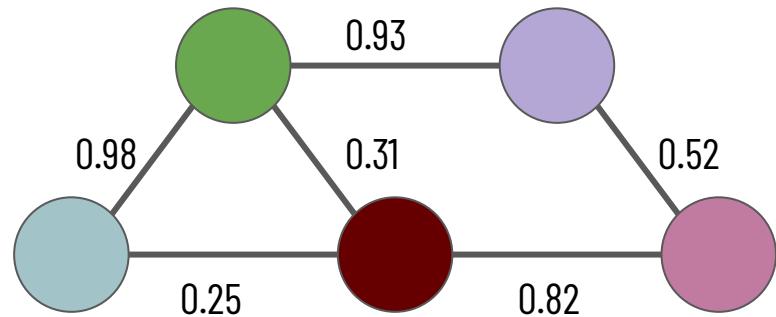


Goal: Partition graph into neuronal processes



Multicut

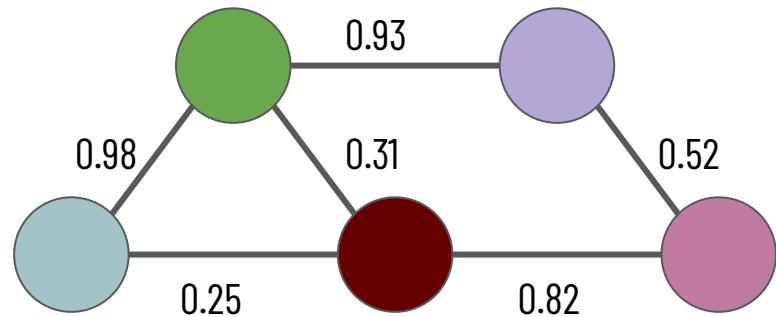
Reformulate the segmentation problem as a multicut graph partitioning one



Multicut

Reformulate the segmentation problem as a multicut graph partitioning one

The final number of segments is not predetermined

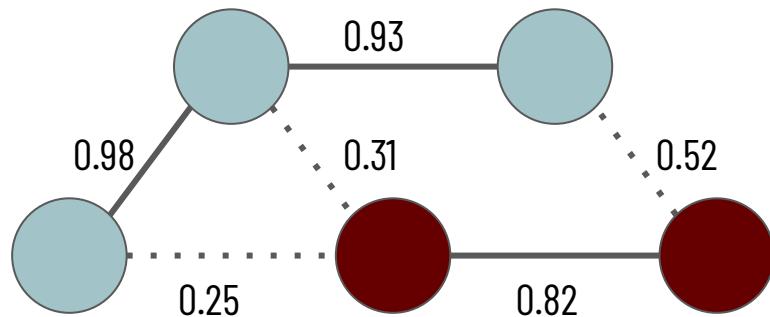


Multicut

Reformulate the segmentation problem as a multicut graph partitioning one

The final number of segments is not predetermined

Guarantees a globally consistent solution

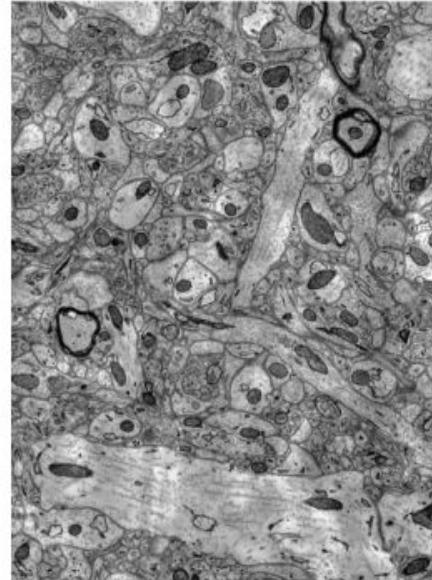
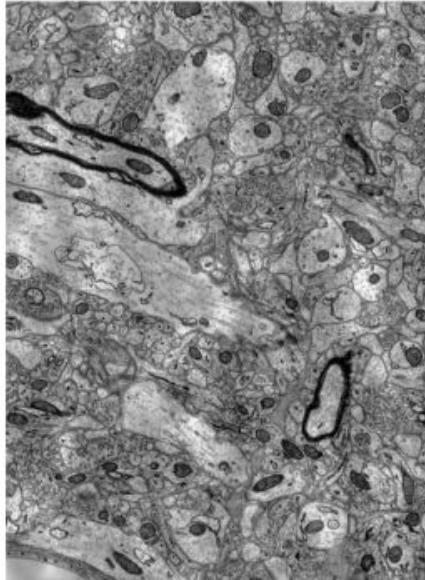


Datasets

Kasthuri

Princeton Neuroscience Institute

SNEMI3D



2 Volumes

$6 \times 6 \times 30 \text{ nm}^3 / \text{vx}$

$1335 \times 1809 \times 338 \text{ vx}$

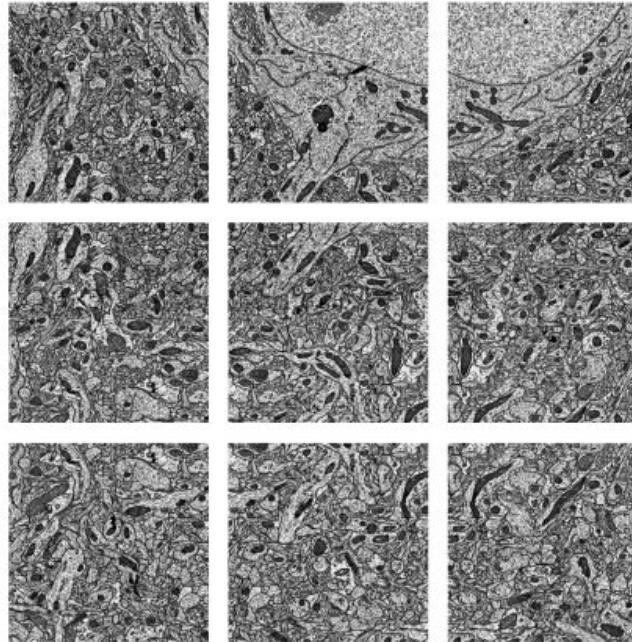
$8.01 \times 10.85 \times 10.14 \mu\text{m}^3$

Datasets

Kasthuri

Princeton Neuroscience Institute

SNEMI3D



9 Volumes

$3.6 \times 3.6 \times 40 \text{ nm}^3 / \text{vx}$

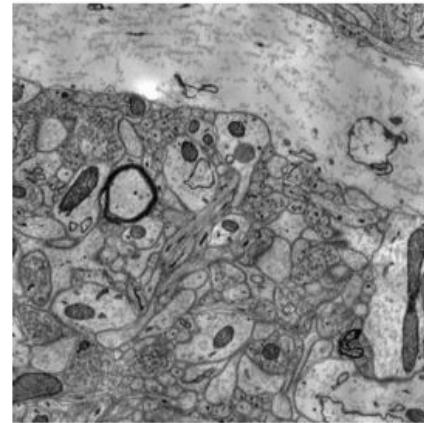
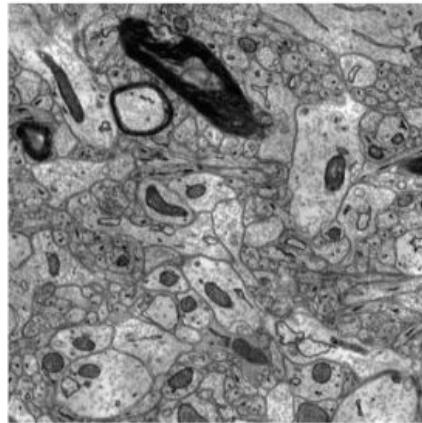
$2048 \times 2048 \times 256 \text{ vx}$

$7.37 \times 7.37 \times 10.24 \mu\text{m}^3$

Datasets

Kasthuri

Princeton Neuroscience Institute



SNEMI3D

2 Volumes

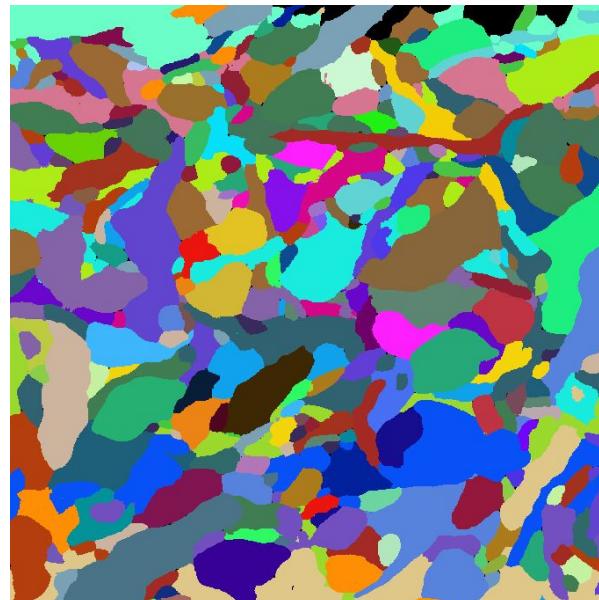
$3 \times 3 \times 30 \text{ nm}^3 / \text{vx}$

$1024 \times 1024 \times 100 \text{ vx}$

$3.07 \times 3.07 \times 3 \mu\text{m}^3$

Input Segmentations

For the two PNI Test datasets, we use zwatershed and mean agglomeration



Input Segmentations

For the Kasthuri and SNEMI3D datasets, we use the waterz agglomeration strategy



Split Variation of Information

Measure of entropy between segmentation and ground truth

Split Variation of Information

Measure of entropy between segmentation and ground truth

VI Split: Increases if two voxels from the same neuron have different labels



Split Variation of Information

Measure of entropy between segmentation and ground truth

VI Split: Increases if two voxels from the same neuron have different labels



VI Merge: Increases if two voxels from different neurons have the same label



Split Variation of Information

Measure of entropy between segmentation and ground truth

VI Split: Increases if two voxels from the same neuron have different labels



VI Merge: Increases if two voxels from different neurons have the same label



Total Variation of Information = VI Split + VI Merge

Variation of Information

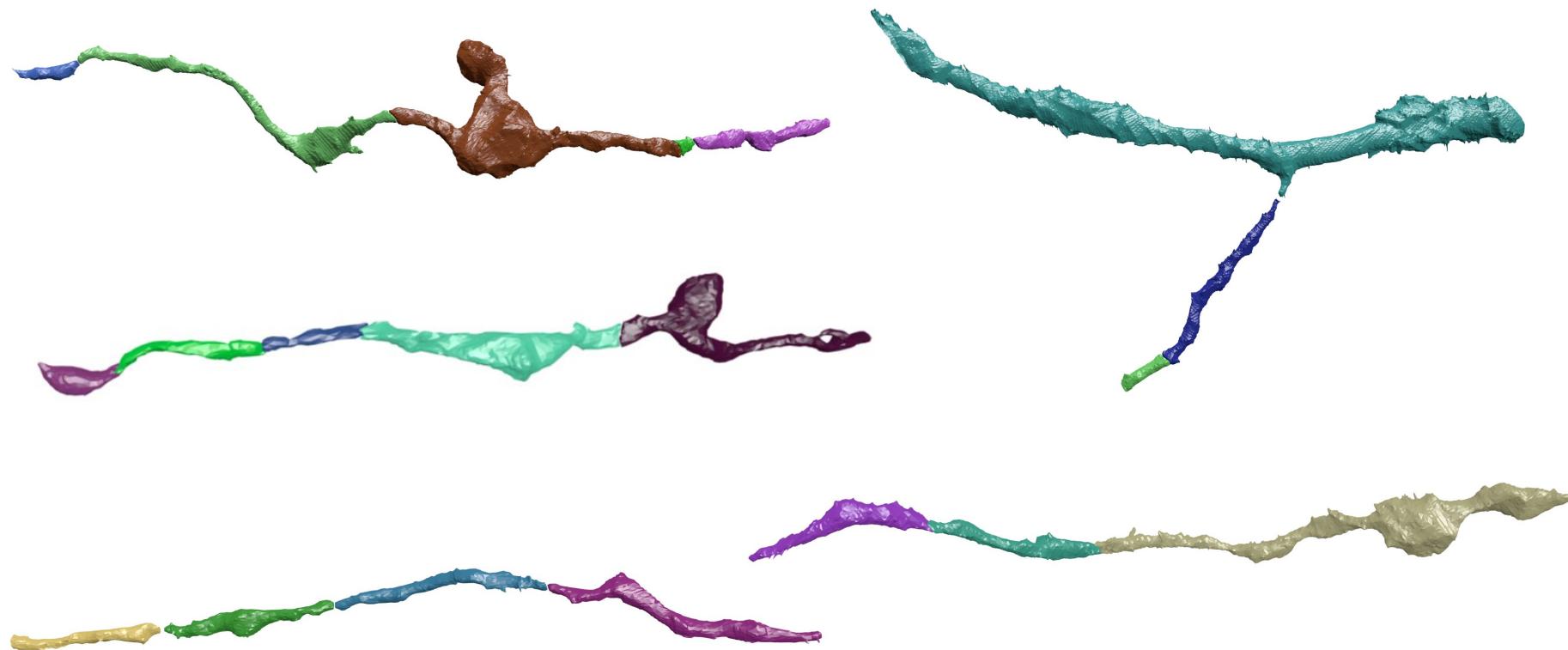
Dataset	Baseline (↓)	Proposed (↓)	Decrease (↑)
PNI Test One	0.491	0.388	-20.9%
PNI Test Two	0.416	0.297	-28.7%
Kasthuri Test	0.965	0.815	-15.6%
SNEMI3D	0.807	0.647	-19.8%

Variation of Information

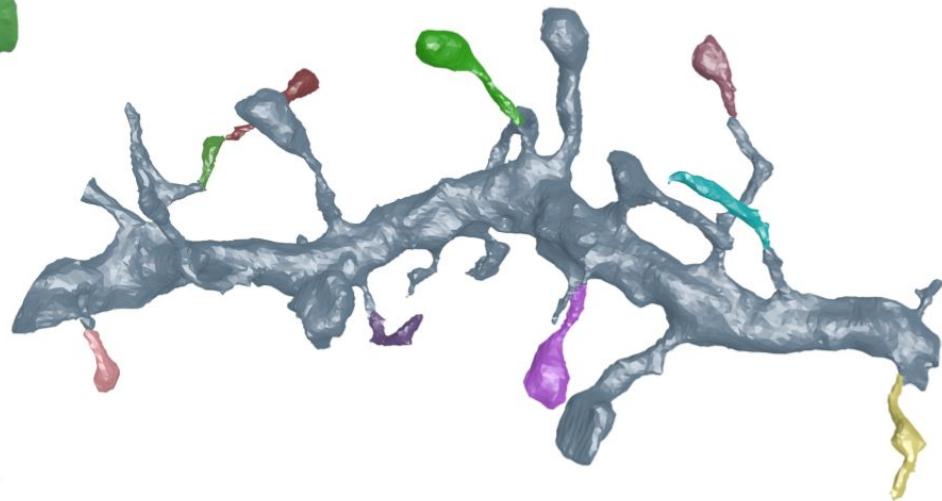
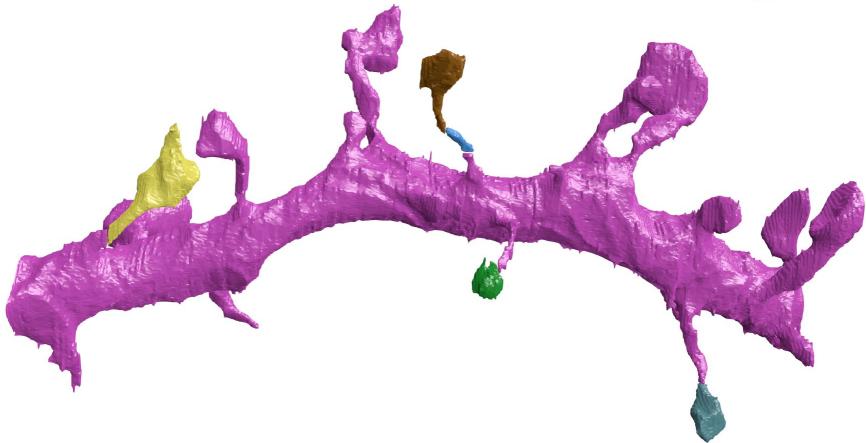
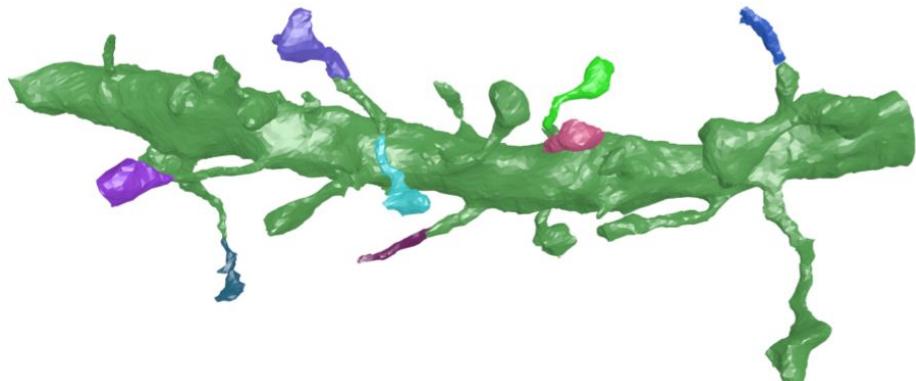
Dataset	Baseline (↓)	Proposed (↓)	Decrease (↑)
PNI Test One	0.491	0.388	-20.9%
PNI Test Two	0.416	0.297	-28.7%
Kasthuri Test	0.965	0.815	-15.6%
SNEMI3D	0.807	0.647	-19.8%

Average decrease of Variation of Information by 21.3% over the four datasets

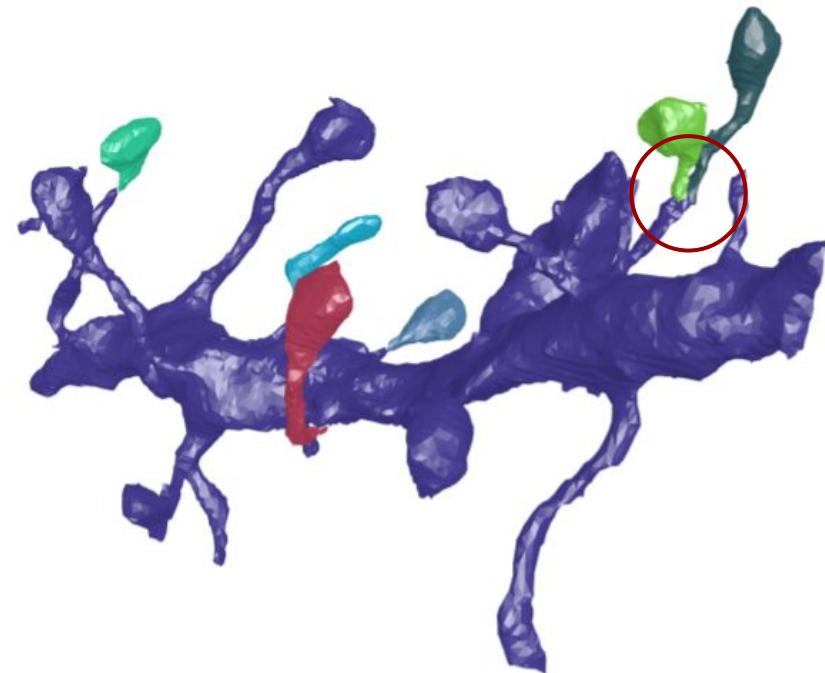
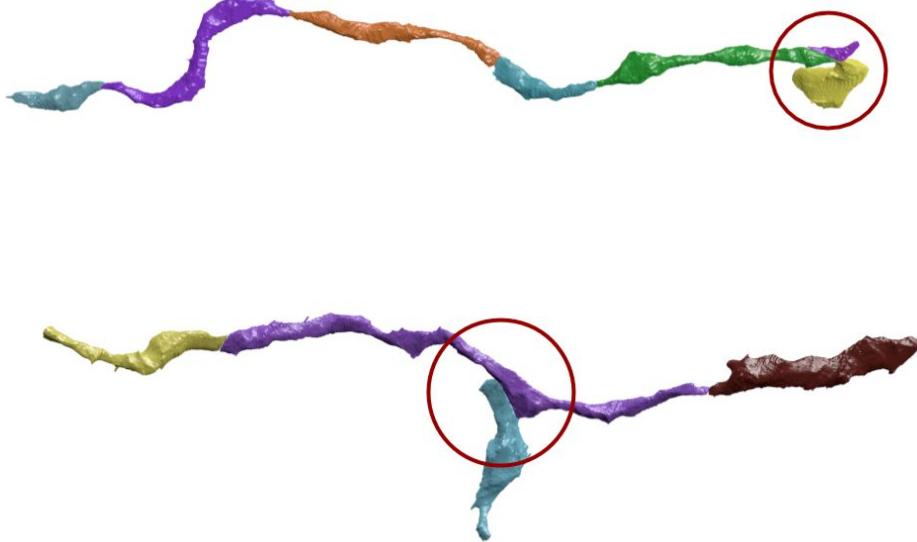
Qualitative Results



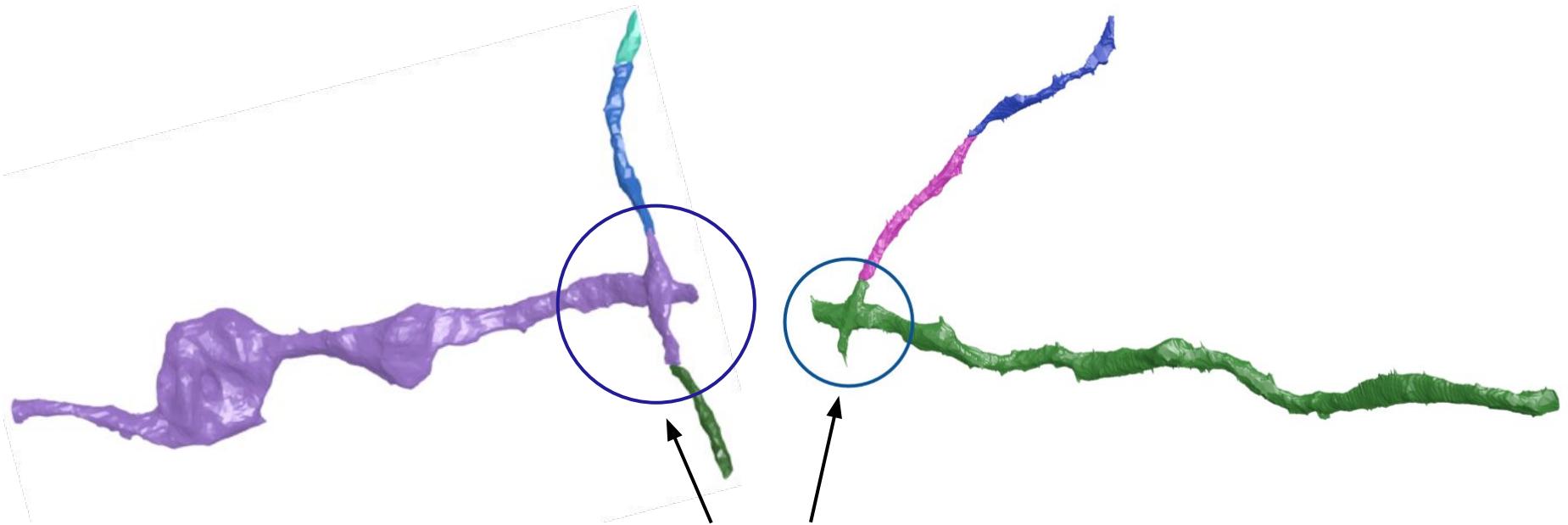
Qualitative Results



Failure Cases



Failure Cases



Errors in Input Segmentation

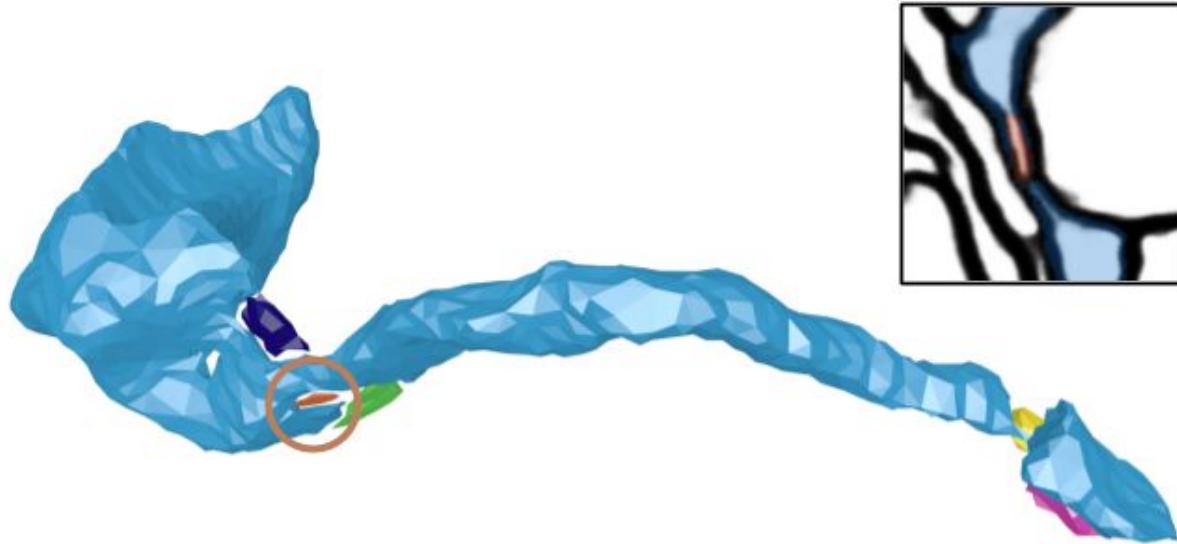
Ablation Studies: Node Generation

Goal: Merge all small segments with a nearby larger segment from the same neuronal process

Ablation Studies: Node Generation

Goal: Merge all small segments with a nearby larger segment from the same neuronal process

Baseline: How many small segments belong to the same neuron as the high affinity large neighbor?



Ablation Studies: Node Generation

Goal: Merge all small segments with a nearby larger segment from the same neuronal process

Baseline: How many small segments belong to the same neuron as the high affinity large neighbor?

Dataset	Baseline (↑)	Proposed (↑)
PNI Test One	305/521 (36.9%)	686/129 (80.2%)
PNI Test Two	185/281 (39.7%)	444/75 (85.5%)
Kasthuri Test	4,514/8,604 (52.5%)	6,623/2,020 (76.6%)

The number of correctly merged small segments versus the number of incorrectly merged segments

Ablation Studies: Edge Generation

Goal: Identify all split errors while minimizing the number of total edges in the graph

Ablation Studies: Edge Generation

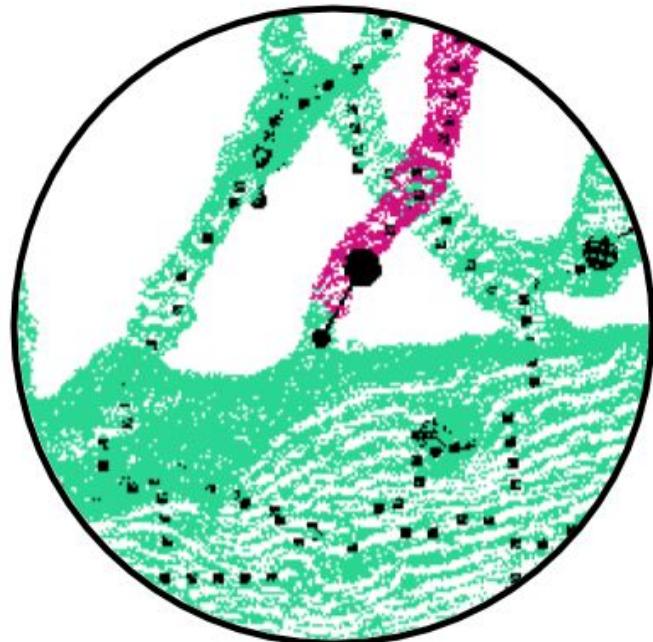
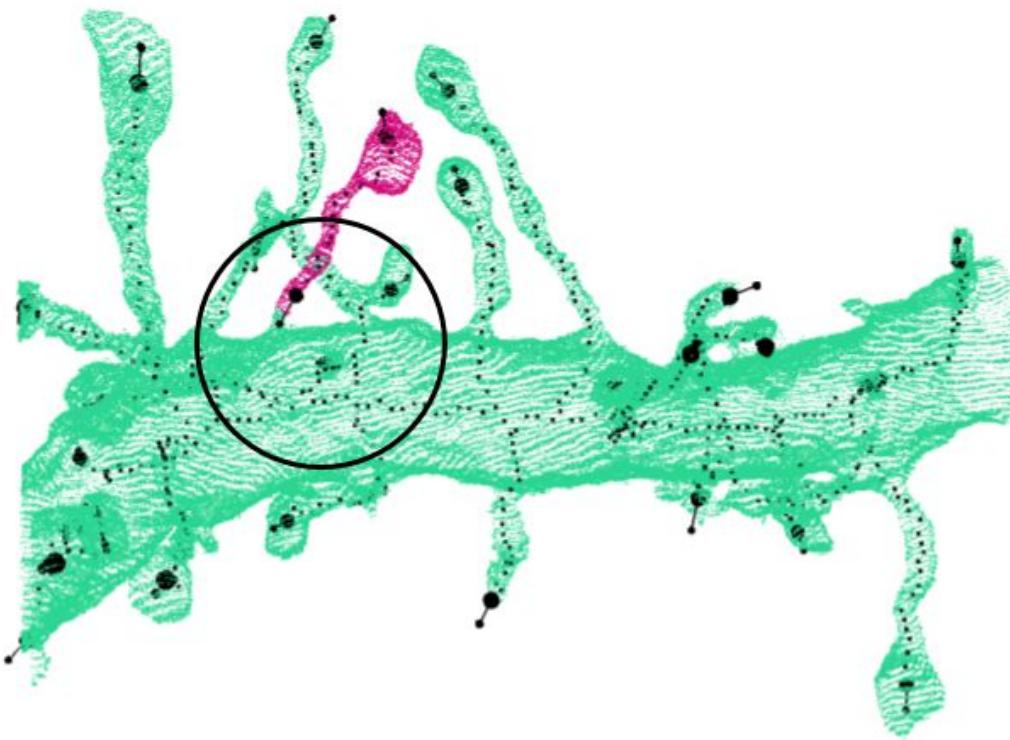
Goal: Identify all split errors while minimizing the number of total edges in the graph

Baseline: How many total edges are there in the adjacency graph?

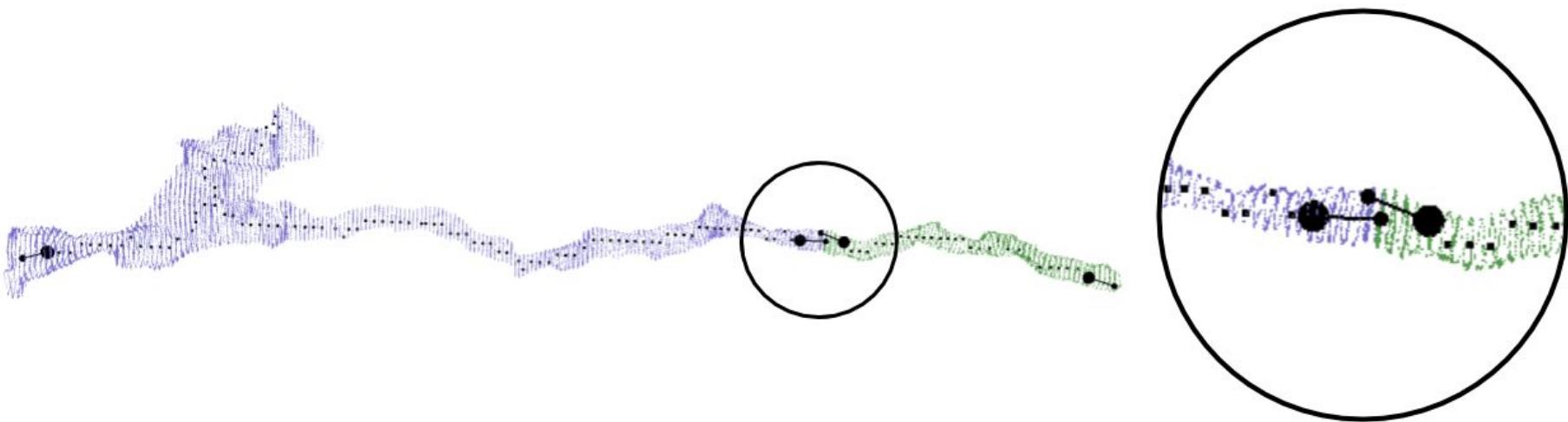
Dataset	Baseline	Proposed	Edge Recall (\uparrow/\downarrow)
PNI Test One	528 / 25,619	417 / 10,074	79.0% / 39.3%
PNI Test Two	460 / 30,388	370 / 11,869	80.4% / 39.1%
Kasthuri Test	1,193 / 43,951	936 / 18,168	78.5% / 41.3%

The number of edges in the graph that correspond to split errors, the total number of edges, and the recall

Ablation Studies: Edge Generation

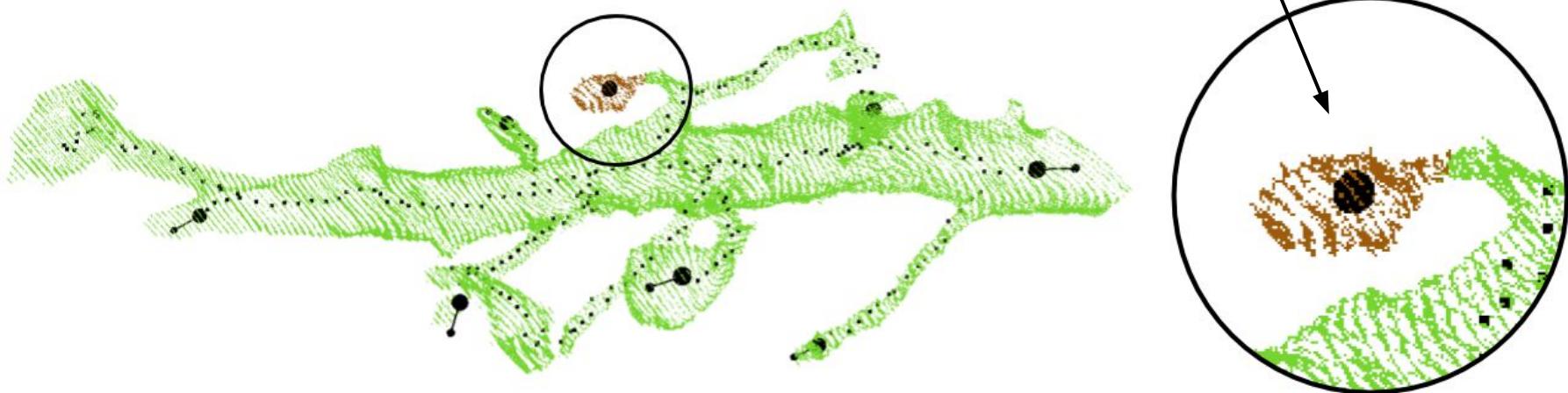


Ablation Studies: Edge Generation

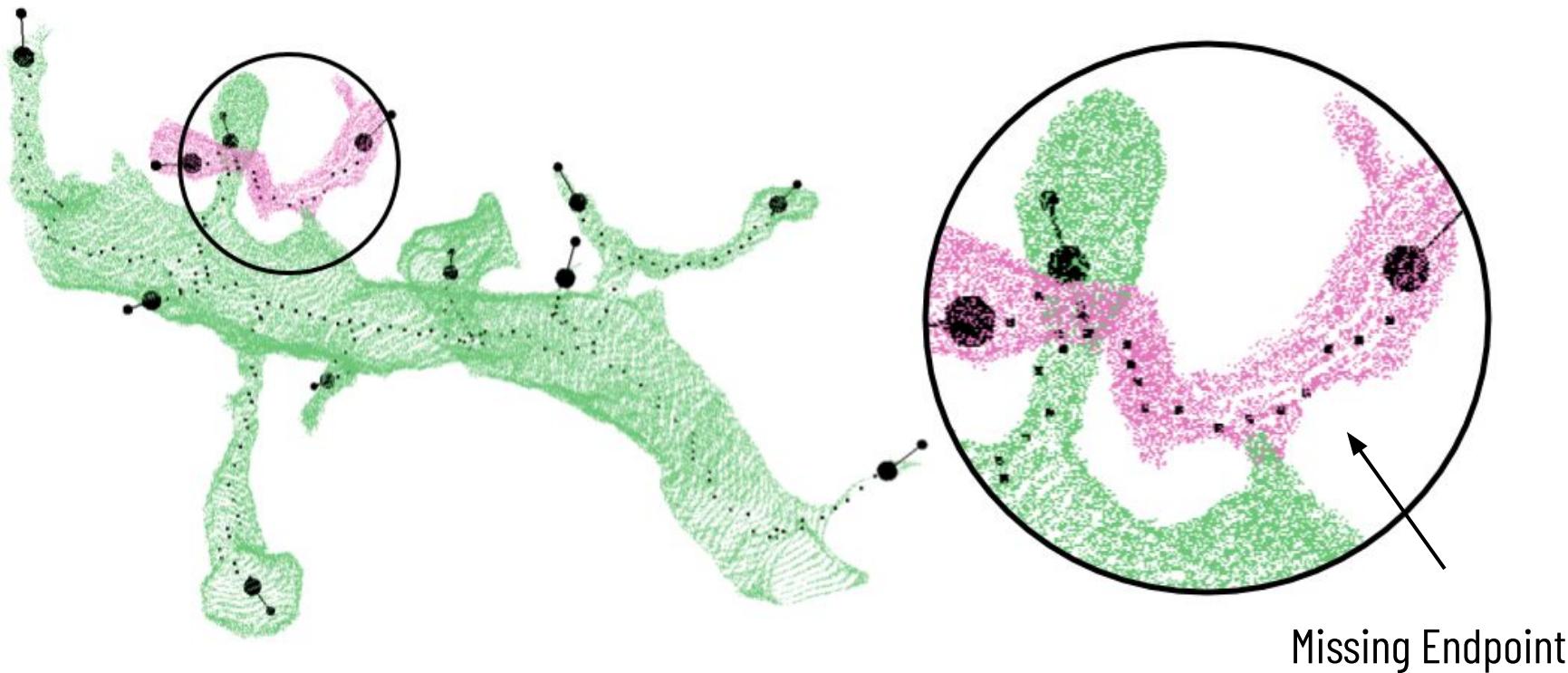


Edge Generation Failure Cases

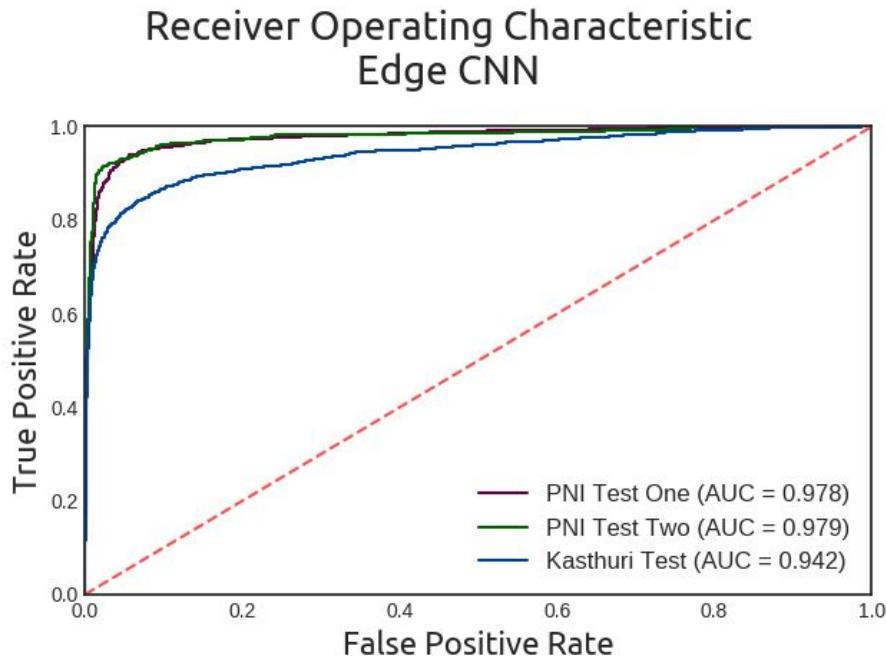
Trivial Skeleton



Edge Generation Failure Cases



Ablation Studies: Edge Weight Assignment



Accuracies:

PNI Test One: 96.4%

PNI Test Two: 97.2%

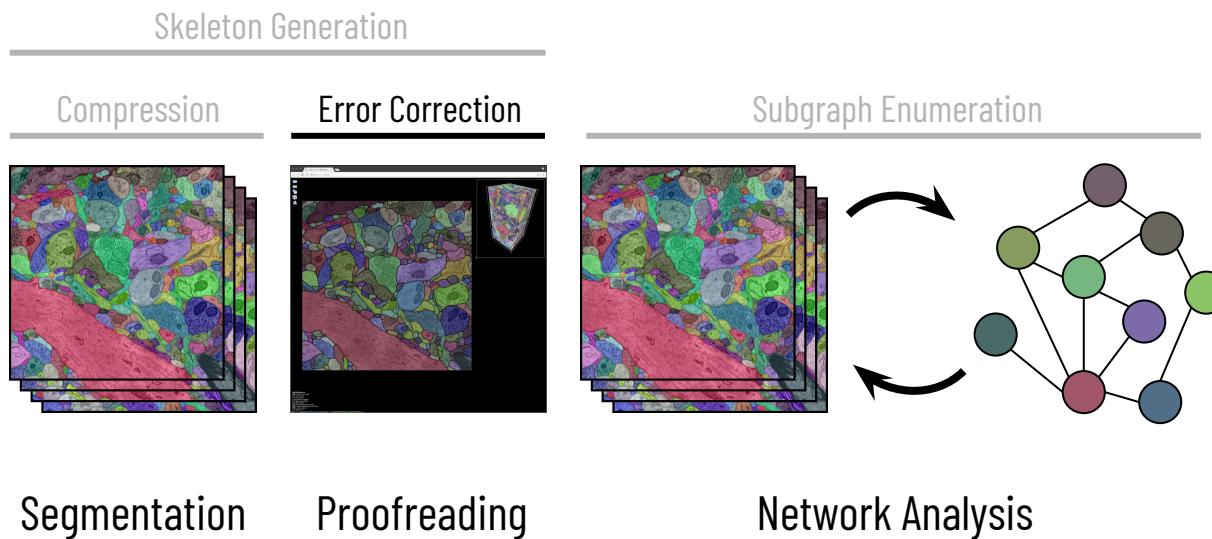
Kasthuri: 93.4%

Running Times

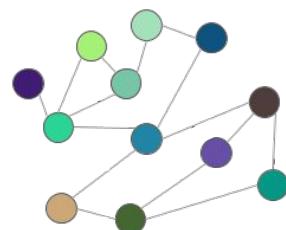
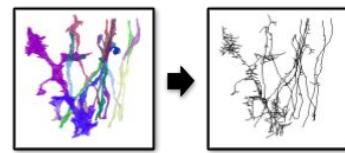
Time to process a gigavoxel dataset

Step	Running Time
Node Generation	281 seconds
Edge Generation	351 seconds
Lifted Multicut	13 seconds
Total	10.75 minutes

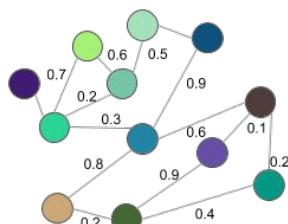
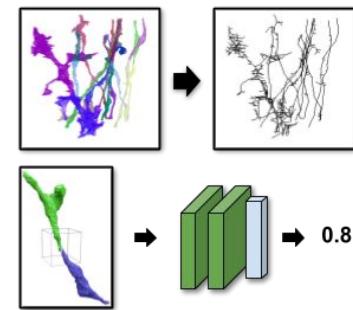
Biologically-Aware Algorithms Along the Connectomics Pipeline



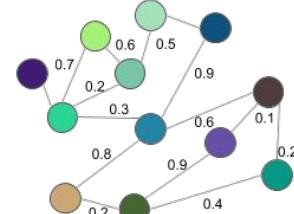
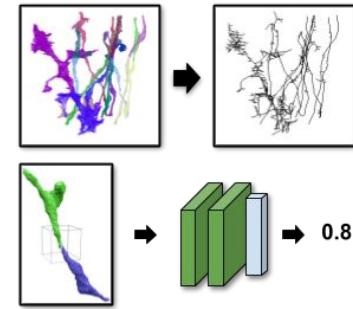
Biologically Constrained
Graph Construction



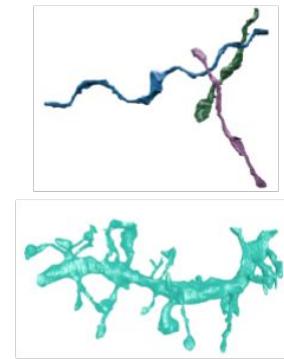
Biologically Constrained Graph Construction



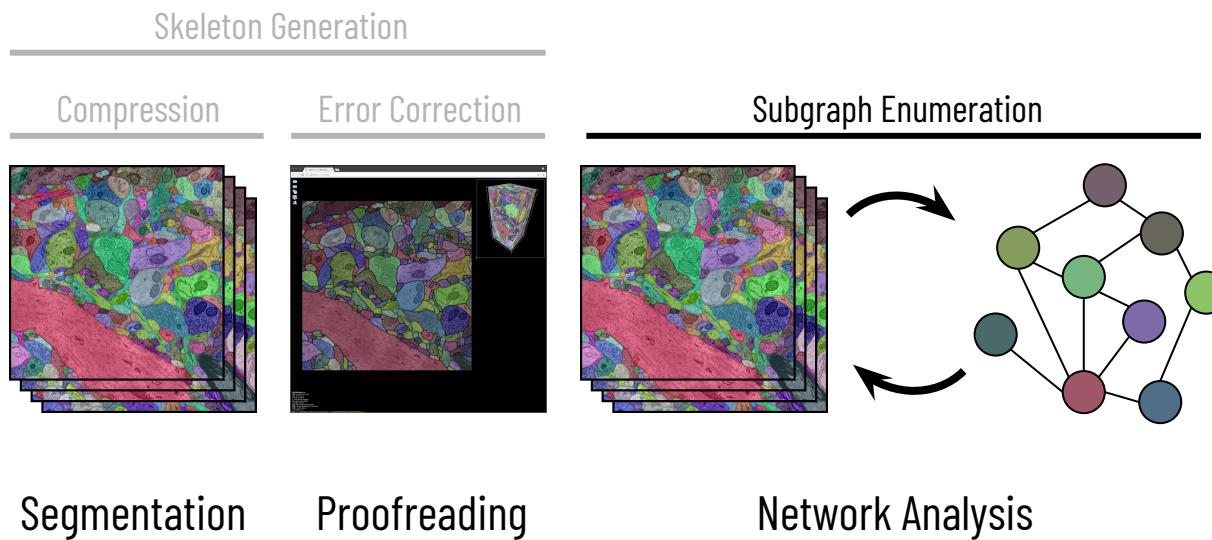
Biologically Constrained Graph Construction



Graph Optimization Results



Biologically-Aware Algorithms Along the Connectomics Pipeline

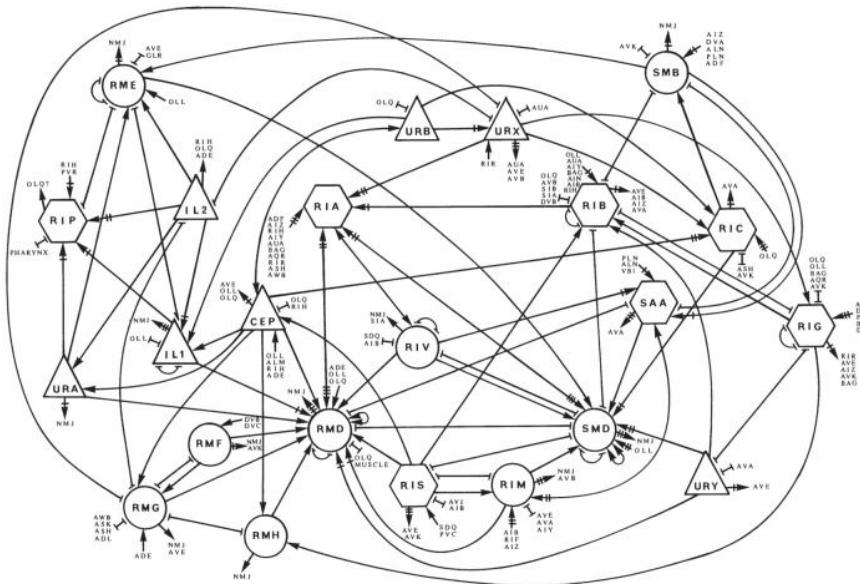


Large-Scale Subgraph Enumeration on the Connectome

Brian Matejek, Donglai Wei, Tianyi Chen, Charalampos E. Tsourakakis,
Michael Mitzenmacher, and Hanspeter Pfister

Extracting the Entire Wiring Diagram

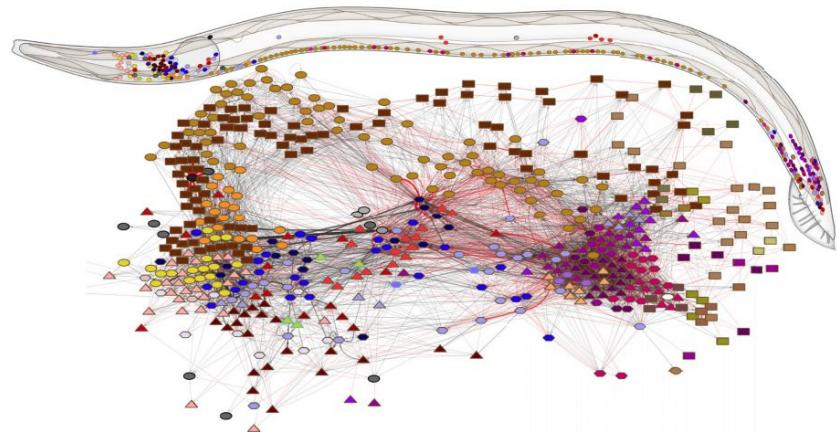
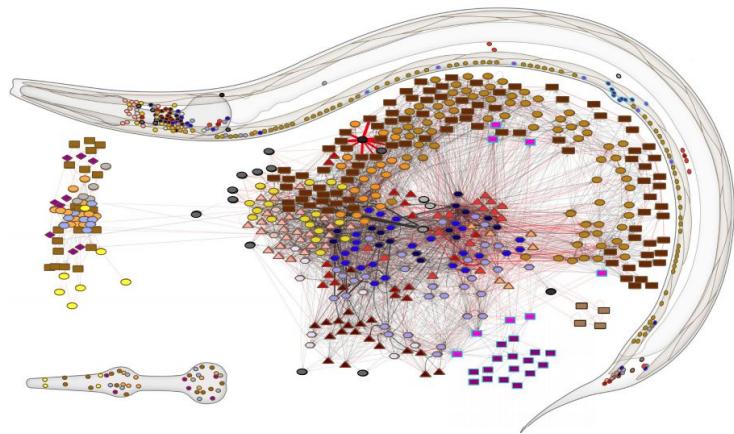
C. elegans, the first nearly complete connectome



Emmons S.W., The Beginning of Connectomics: A Commentary on White et al. (1986) 'The Structure of the Nervous System of the Nematode *Caenorhabditis elegans*.' Philosophical Transactions of the Royal Society of London 2015.
Copyright © The Royal Society

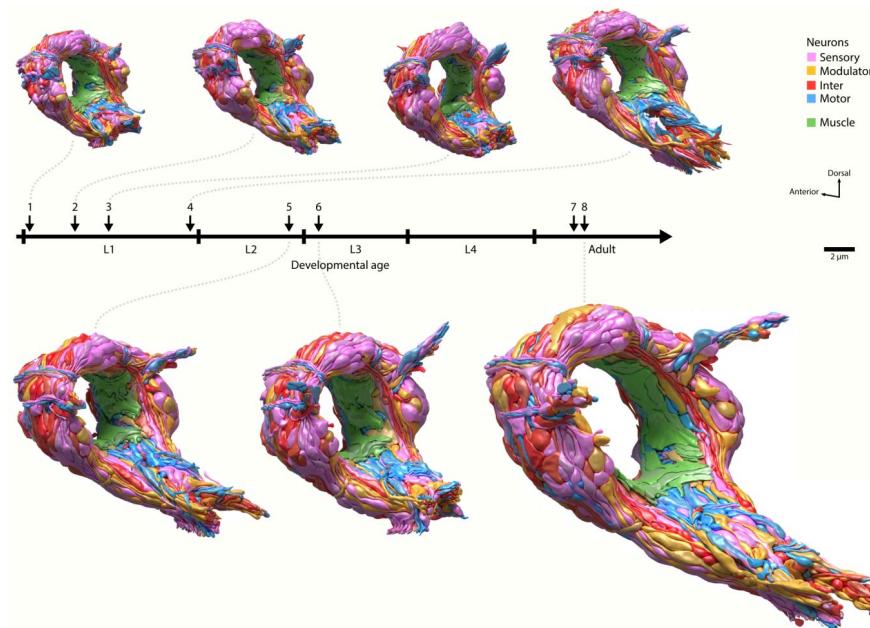
Rapid Expansions in Large-Scale Wiring Diagrams

Improvements in the automatic processes allow for larger and more diverse connectomes



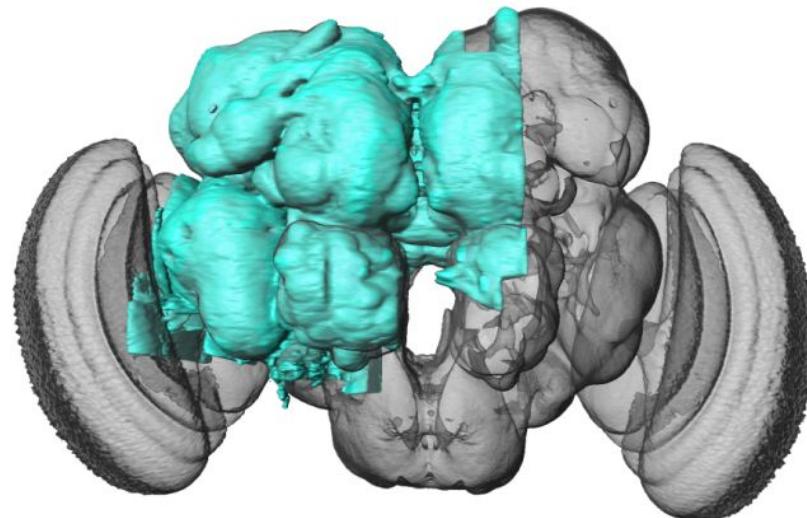
Rapid Expansions in Large-Scale Wiring Diagrams

Improvements in the automatic processes allow for larger and more diverse connectomes



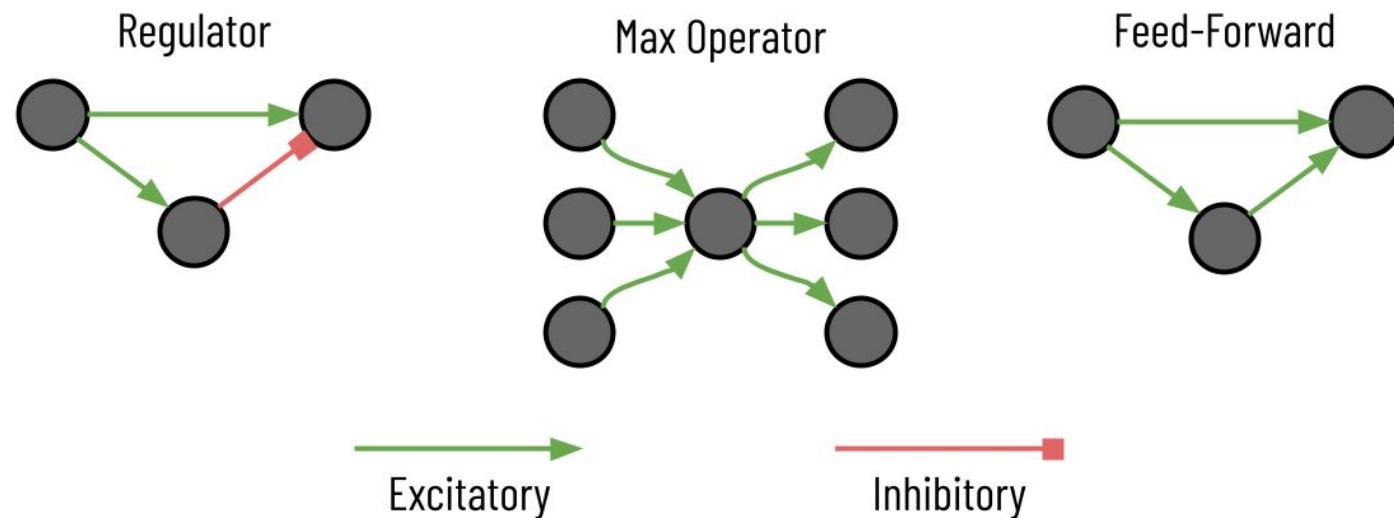
Rapid Expansions in Large-Scale Wiring Diagrams

Improvements in the automatic processes allow for larger and more diverse connectomes



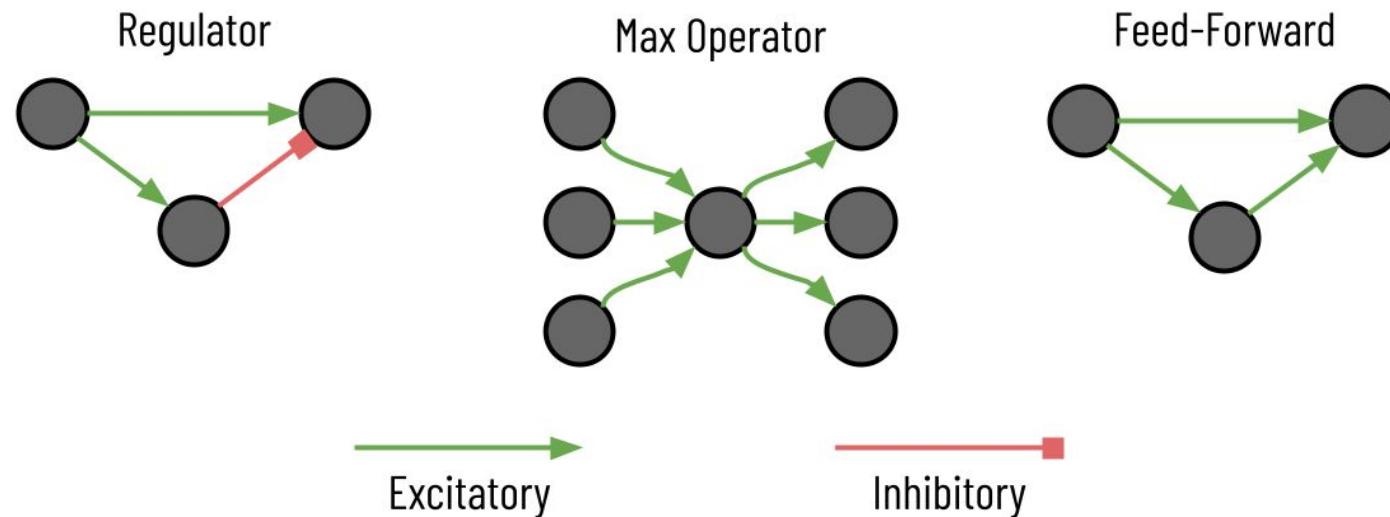
Motif Discovery

Some subgraphs, or motifs, in the wiring diagram correspond to specific computational functions



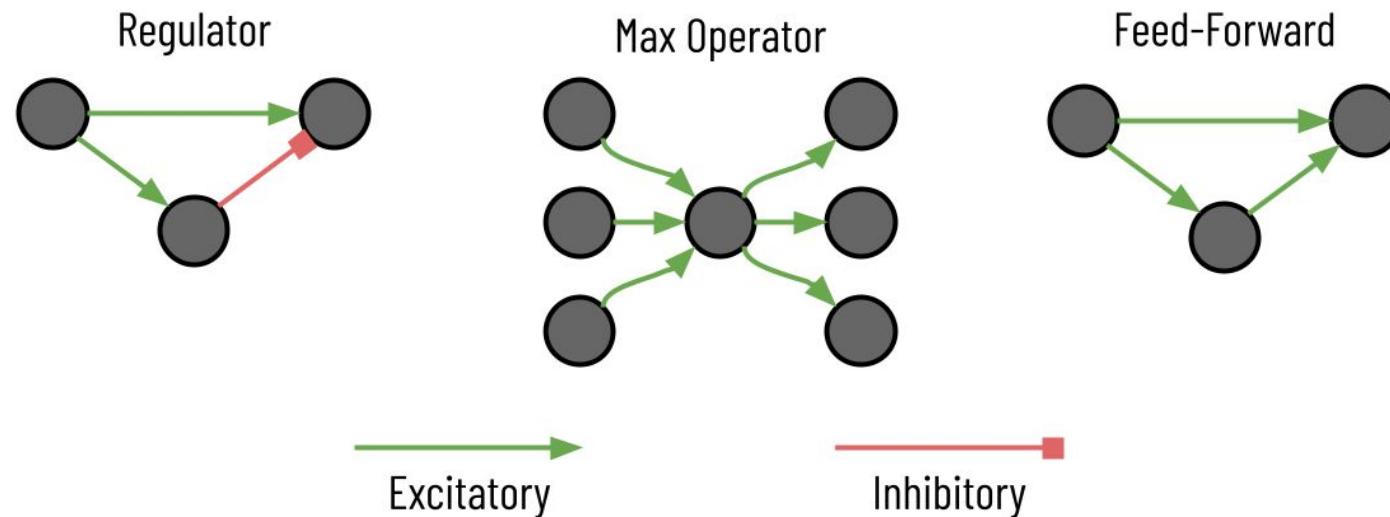
Motif-Centric Subgraph Enumeration

Most current motif analysis on these connectomes looks only for currently expected motifs



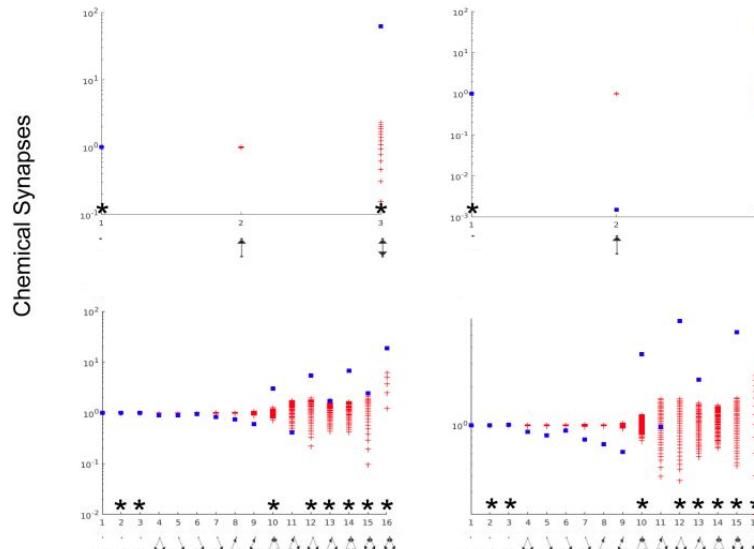
Motif-Centric Subgraph Enumeration

Most current motif analysis on these connectomes looks only for currently expected motifs
However, these strategies cannot identify “unknown unknowns”



Network-Centric Subgraph Enumeration

Analyses that consider all viable subgraphs of size k in a connectome typically restrict k to 2 or 3



Computational Issues

The number of subgraphs to enumerate grows quickly as k increases

Subgraph Size (k)	No. Subgraphs
3	122,933
4	3,753,059
5	124,780,035
6	4,218,875,364
7	141,111,975,907

Results from one connectome with 585 neurons/muscles/end-organs and 7,415 edges

Computational Issues

The number of subgraphs to enumerate grows quickly as k increases

Subgraph Size (k)	No. Subgraphs	
3	122,933	
4	3,753,059	30 x
5	124,780,035	33 x
6	4,218,875,364	34 x
7	141,111,975,907	33 x

Results from one connectome with 585 neurons/muscles/end-organs and 7,415 edges

Computational Issues

The number of subgraphs to enumerate grows quickly as k increases

The complexity to correctly classify a single subgraph increases exponentially with k

Subgraph Size (k)	Subgraphs per second
4	285,136
5	221,352
6	161,812
7	119,991
8	95,244

Averaged over eight connectomes with 225 neurons each (variable edges)

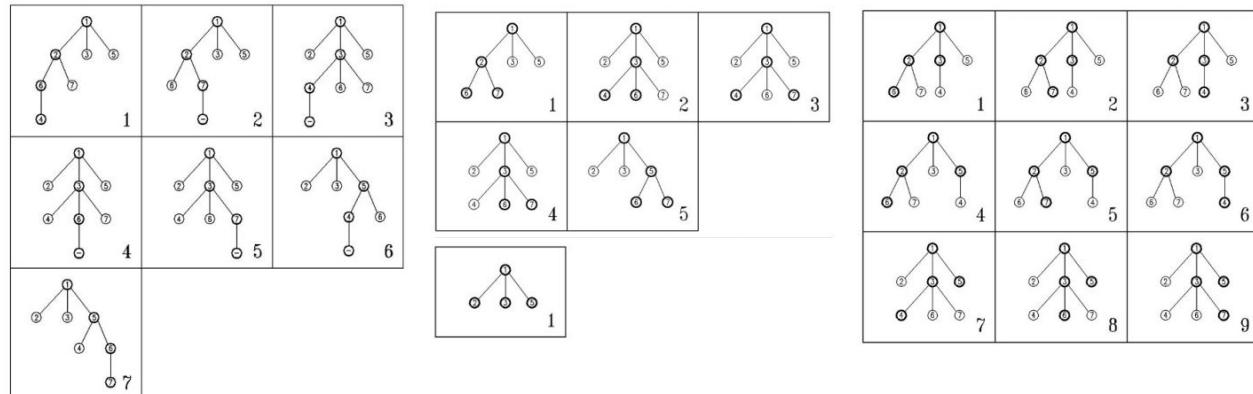
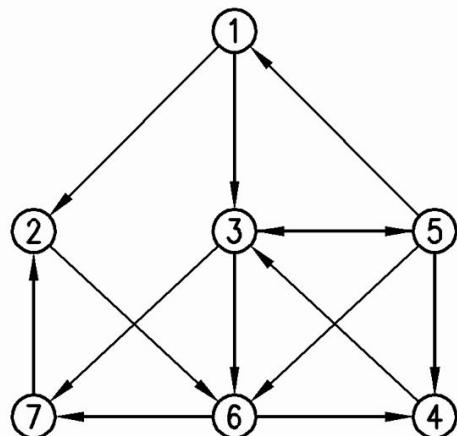
Computational Issues

Subgraph Size (k)	No. Subgraphs	Computation Time
3	126,610,248	9.44 min
4	36,041,949,778	2.15 d
5	12,522,283,314,604	2.77 yr

Results from one connectome with 21,739 neurons and 841,720 edges

Kavosh Subgraph Enumeration

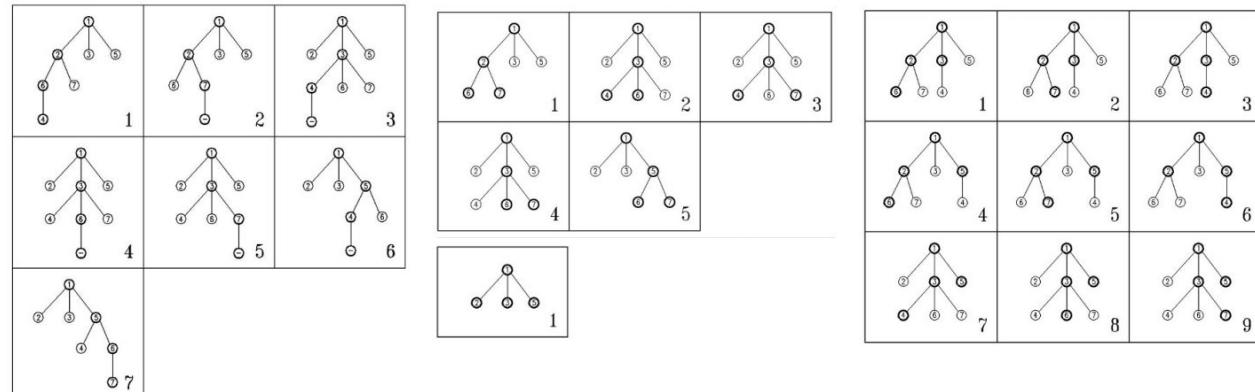
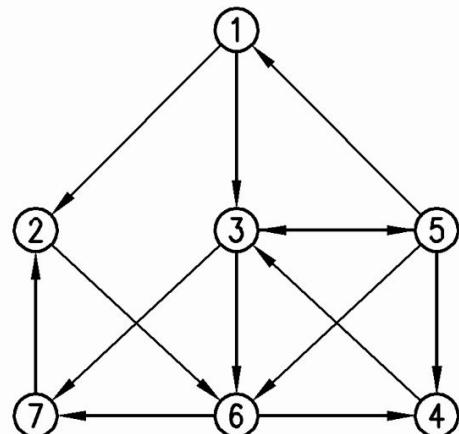
We extend on the Kavosh algorithm, an existing fast method for network-centric enumeration



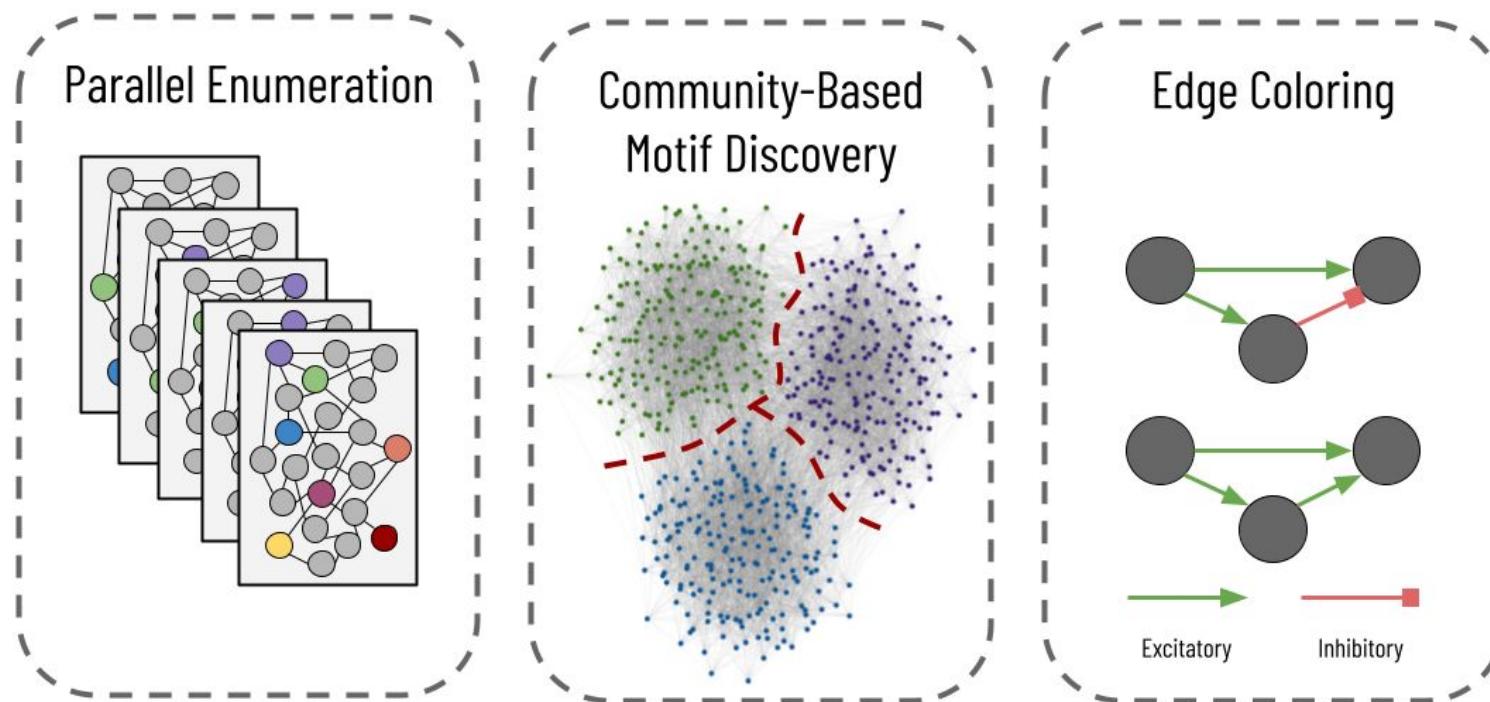
Kavosh Subgraph Enumeration

We extend on the Kavosh algorithm, an existing fast method for subgraph enumeration

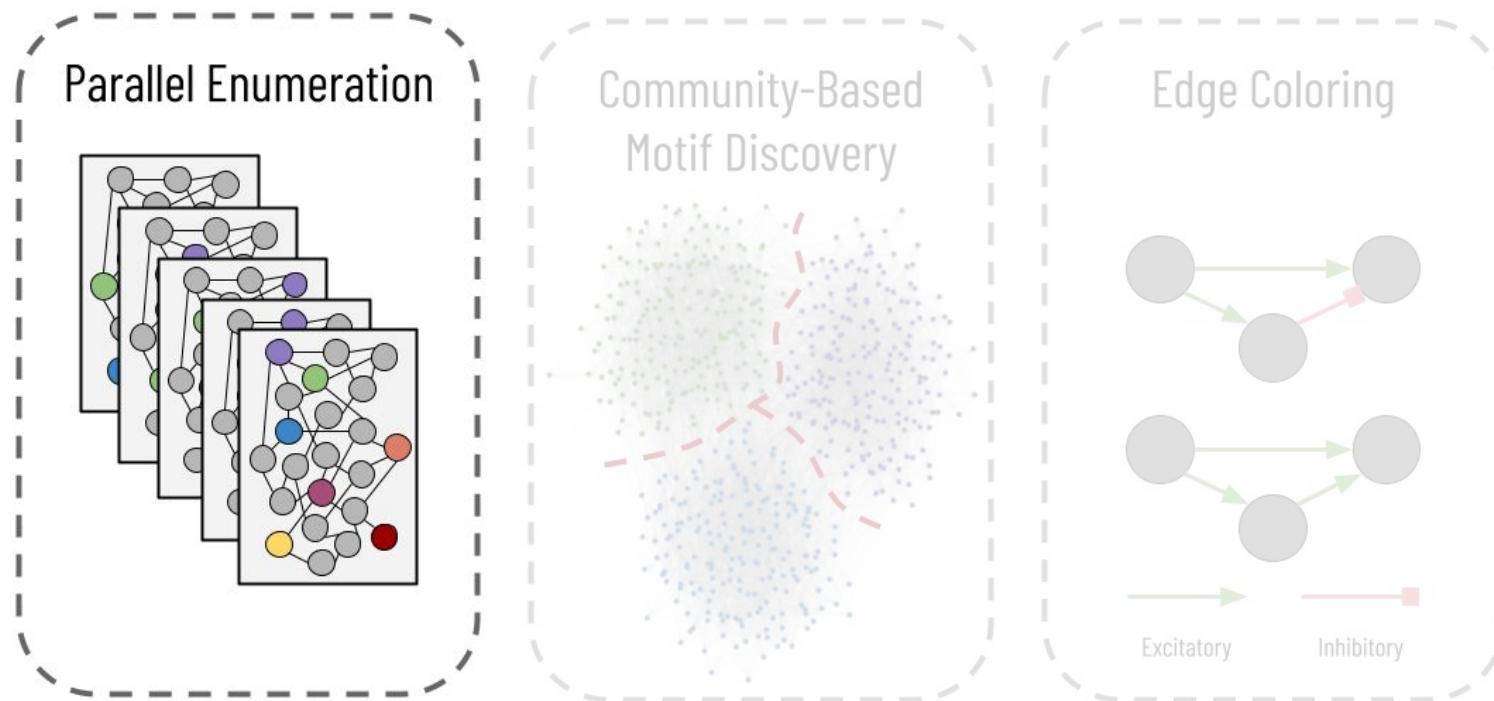
To avoid subgraph duplication, the algorithm considers all subgraphs rooted at a given vertex



Proposed Large-Scale Subgraph Enumeration

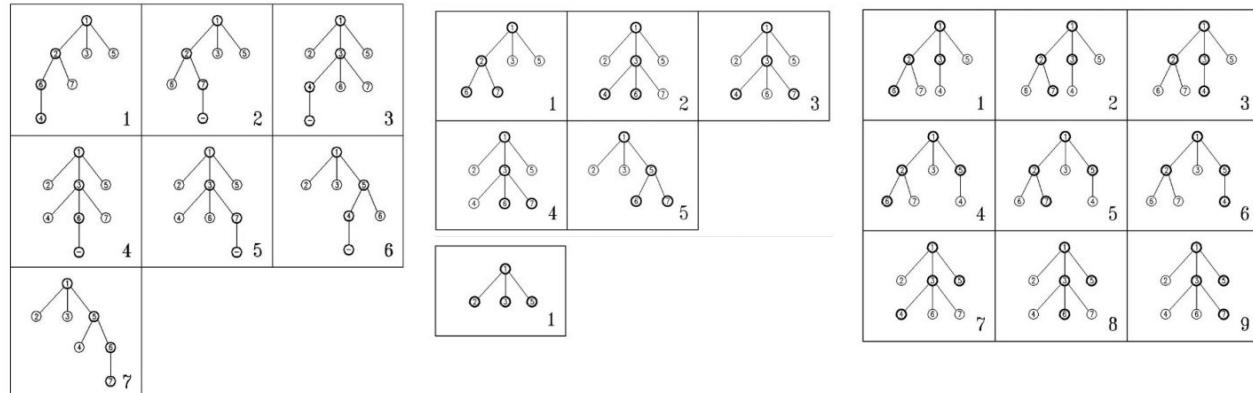
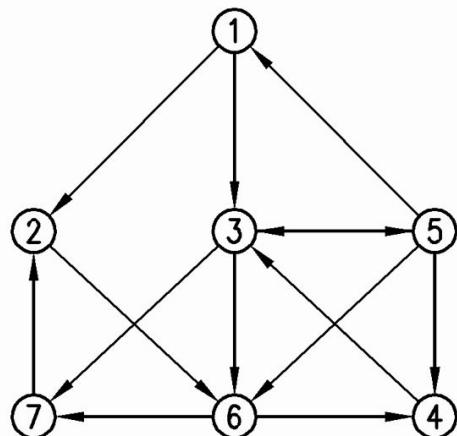


Goal: Distribute Enumeration over a Compute Cluster



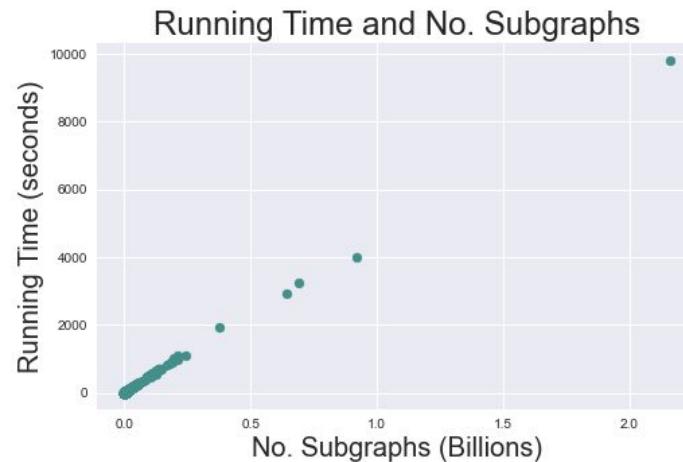
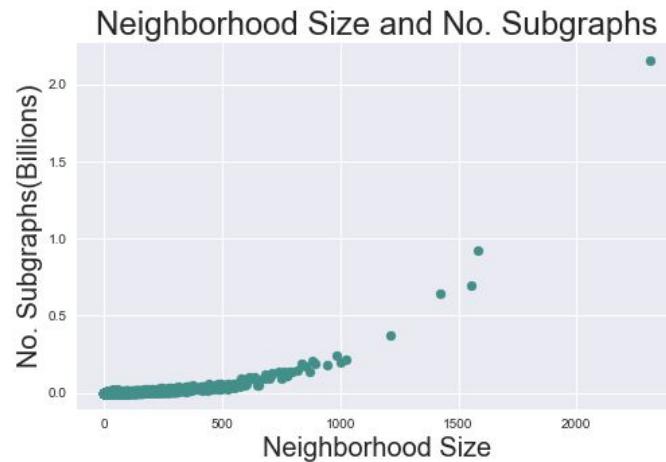
Parallelizing Kavosh

The algorithm itself is easy to parallelize—enumerate the subgraphs rooted at each vertex in parallel



Better Parallelization

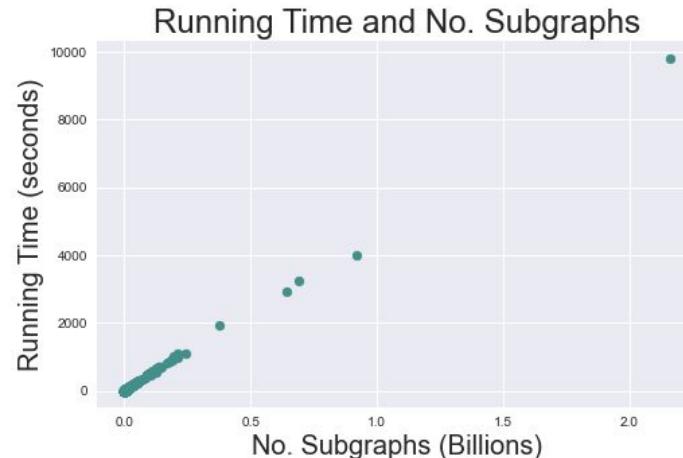
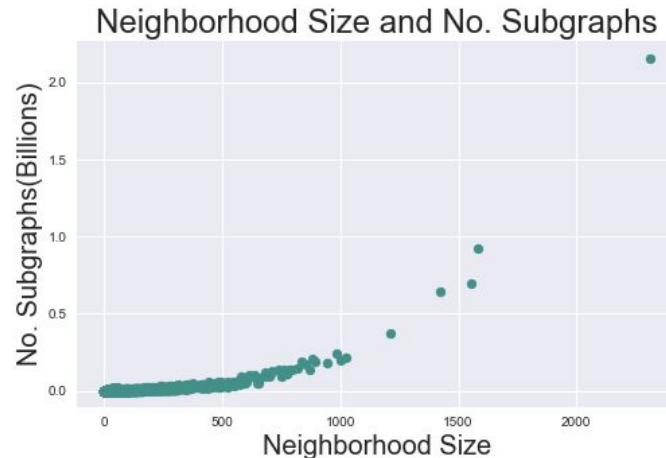
There are significant disparities in the number of subgraphs rooted at a given vertex



Better Parallelization

There are significant disparities in the number of subgraphs rooted at a given vertex

The many large cliques in connectomics can produce vertices with thousands of times more subgraphs



Readjusting the Enumeration Order

We relabel the vertices to minimize the influence of these cliques and reduce the idle CPU time

	Standard Order	Readjusted Order
Mean Time	8.21 sec	8.54 sec
Median Time	0.92 sec	5.80 sec
Maximum Time	9,820.40 sec	74.12 sec
Wall Time	175.20 min	15.78 min
Idle CPU Time	680.44 hr	14.18 hr

Results for $k = 4$ from one connectome with 21,739 neurons and 841,720 synaptic connections

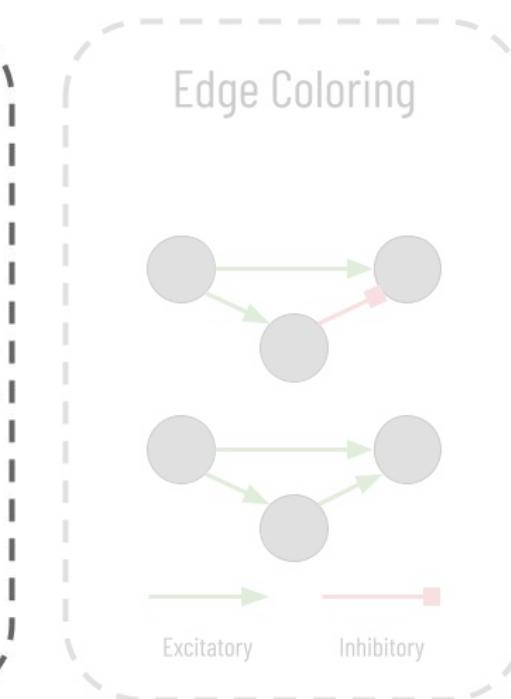
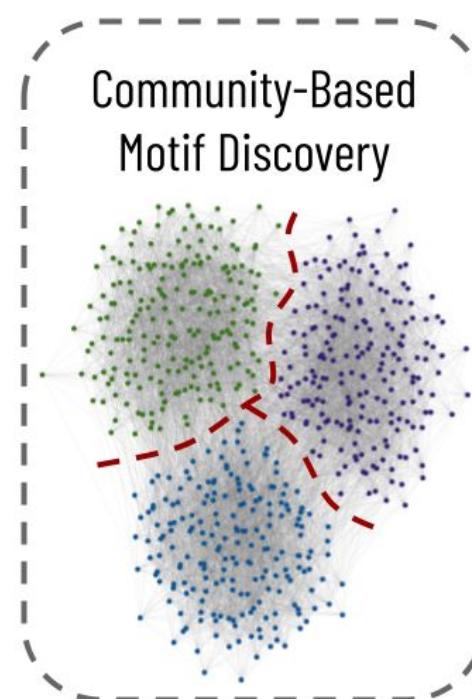
Readjusting the Enumeration Order

We relabel the vertices to minimize the influence of these cliques and reduce the idle CPU time

	Standard Order	Readjusted Order
Mean Time	8.21 sec	8.54 sec
Median Time	0.92 sec	5.80 sec
Maximum Time	9,820.40 sec	74.12 sec
Wall Time	175.20 min	15.78 min
Idle CPU Time	680.44 hr	14.18 hr

Results for $k = 4$ from one connectome with 21,739 neurons and 841,720 synaptic connections

Goal: Divide and Conquer Enumeration for Very Large Datasets



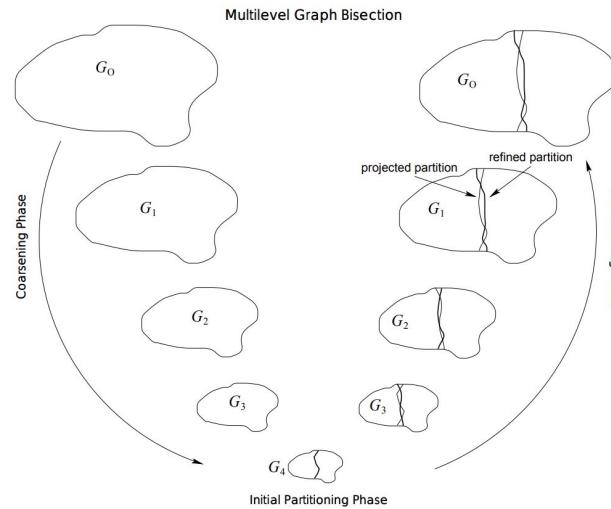
Clustering Before Enumeration

For very large connectomes, parallelization will not be enough as the number of subgraphs grows quickly

Clustering Before Enumeration

For very large connectomes, parallelization will not be enough as the number of subgraphs grows quickly

We use the METIS algorithm to create similarly sized clusters and enumerate subgraphs within clusters

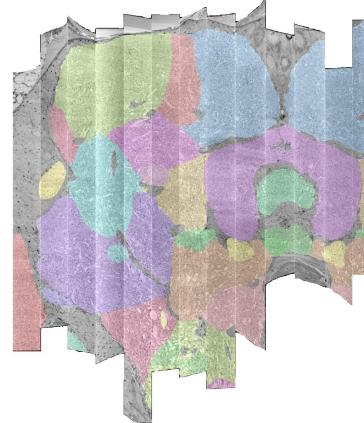


Clustering Before Enumeration

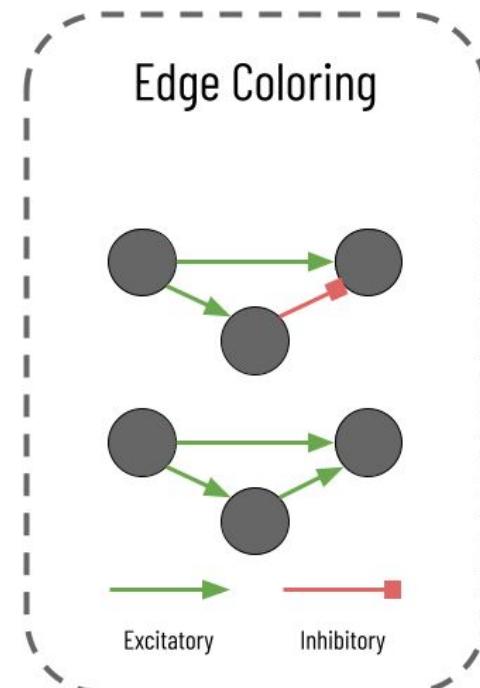
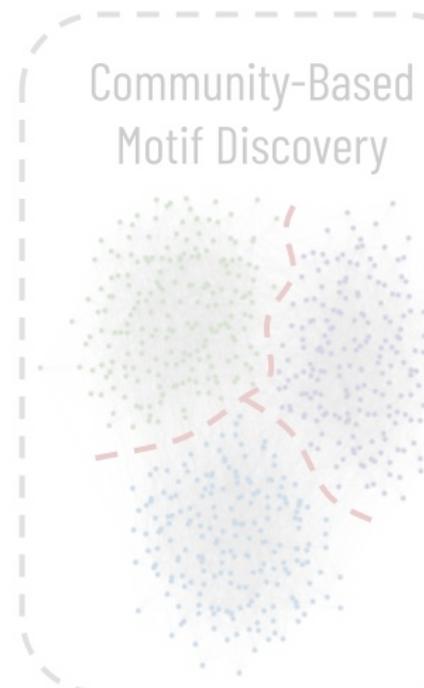
For very large connectomes, parallelization will not be enough as the number of subgraphs grows quickly

We use the METIS algorithm to create similarly sized clusters and enumerate subgraphs within clusters

These clusters could similarly be determined by domain-knowledge of the brain regions

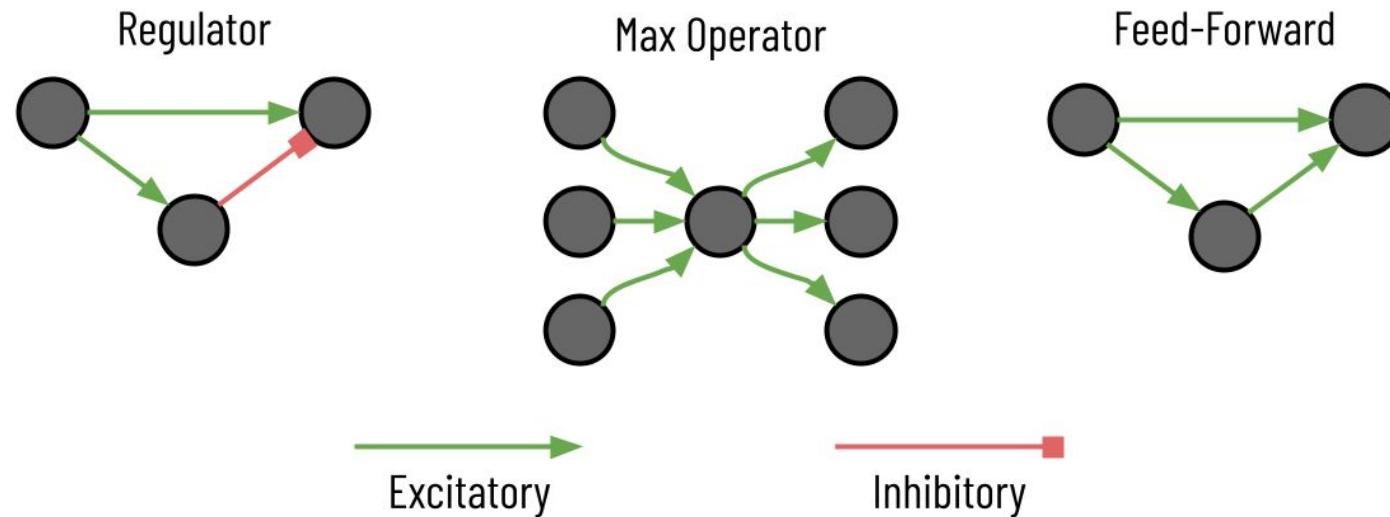


Goal: Augment Graphs with Biologically Relevant Features



Creating Graphs that Better Resemble the Biology

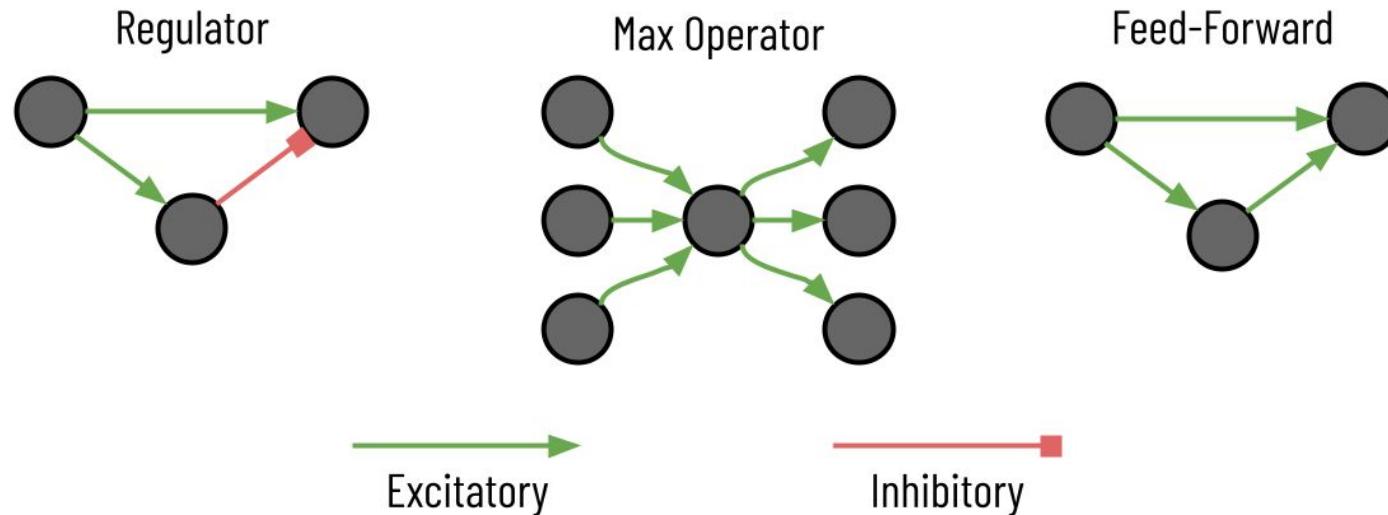
Similar motifs can have variable functionalities depending on the connections



Creating Graphs that Better Resemble the Biology

Similar motifs can have variable functionalities depending on the connections

Edges can be excitatory/inhibitory, or even represent chemical/electrical synapse connections



Datasets

Drosophila

	Age	Sex	Neurons	Edges	Edge Types
	Adult	Female	21,739	841,720	Moderate/Strong

C. elegans Development

C. elegans Sexes

Datasets

	Age	Sex	Neurons	Edges	Edge Types
<i>Drosophila</i>	0 hr	Hermaphrodite	225	775	N/A
	5 hr	Hermaphrodite	225	986	N/A
<i>C. elegans Development</i>	8 hr	Hermaphrodite	225	1,006	N/A
	16 hr	Hermaphrodite	225	1,101	N/A
<i>C. elegans Sexes</i>	23 hr	Hermaphrodite	225	1,504	N/A
	27 hr	Hermaphrodite	225	1,524	N/A
	Adult (50 hr)	Hermaphrodite	225	2,193	N/A
	Adult (50 hr)	Hermaphrodite	225	2,189	N/A

Datasets

Drosophila

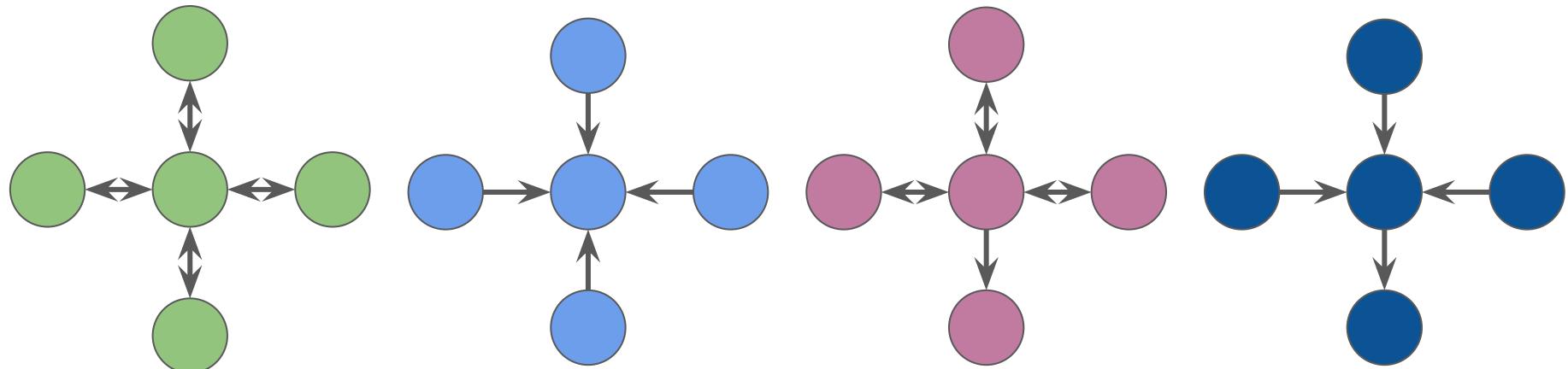
	Age	Sex	Neurons	Edges	Edge Types
	Adult	Hermaphrodite	460	6,469	Chem/Elec/Both
	Adult	Male	585	7,415	Chem/Elec/Both

C. elegans Development

***C. elegans* Sexes**

Common Subgraphs in *Drosophila*

The four most common motifs of size 5 accounts for over 27.5% of all subgraphs



13.87%

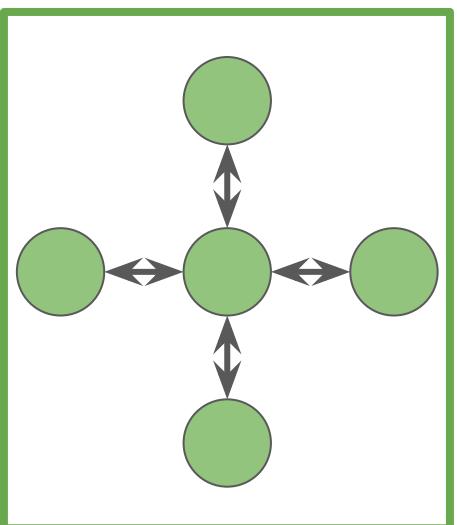
6.55%

3.67%

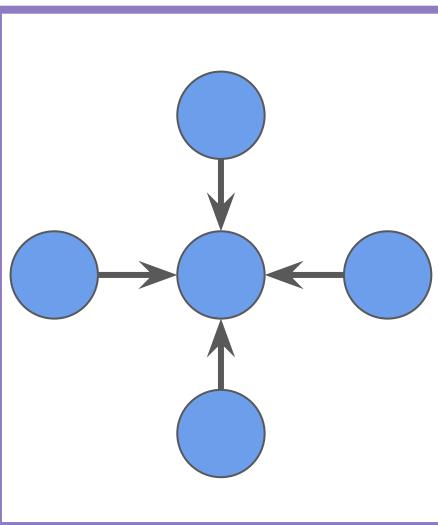
3.45%

Common Subgraphs in *Drosophila*

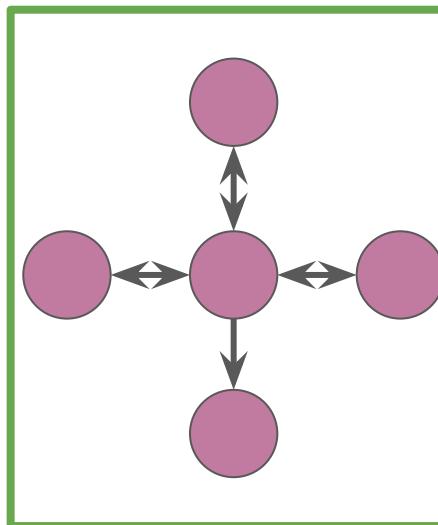
The four most common motifs of size 5 accounts for over 27.5% of all subgraphs



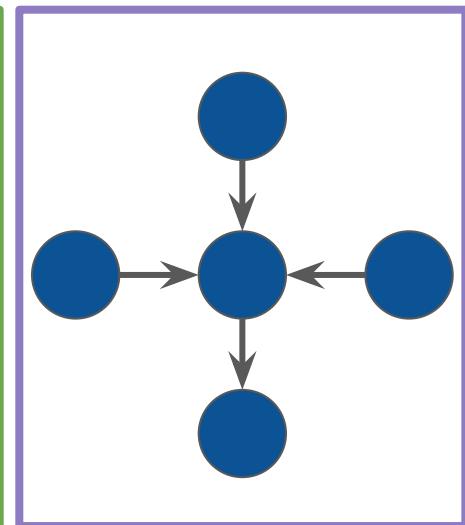
13.87%



6.55%



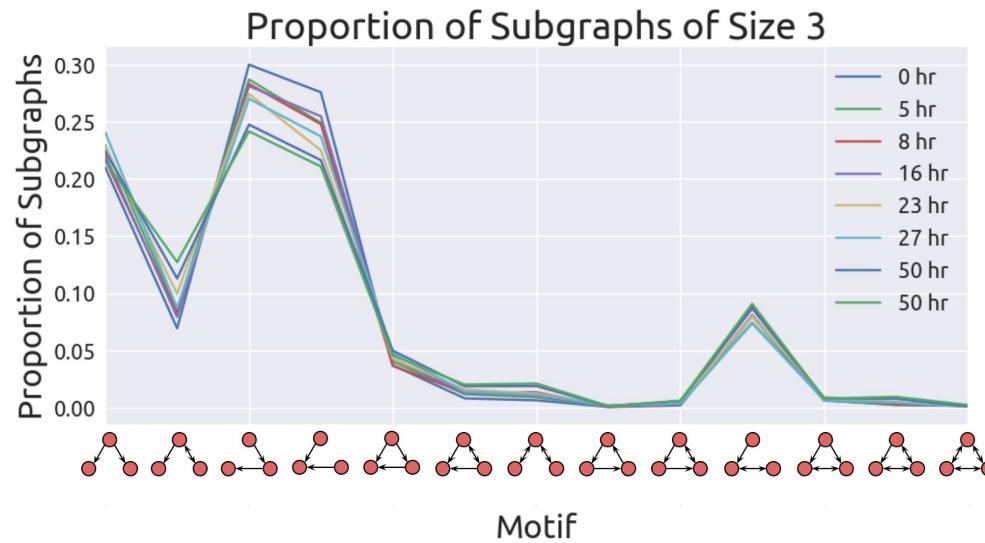
3.67%



3.45%

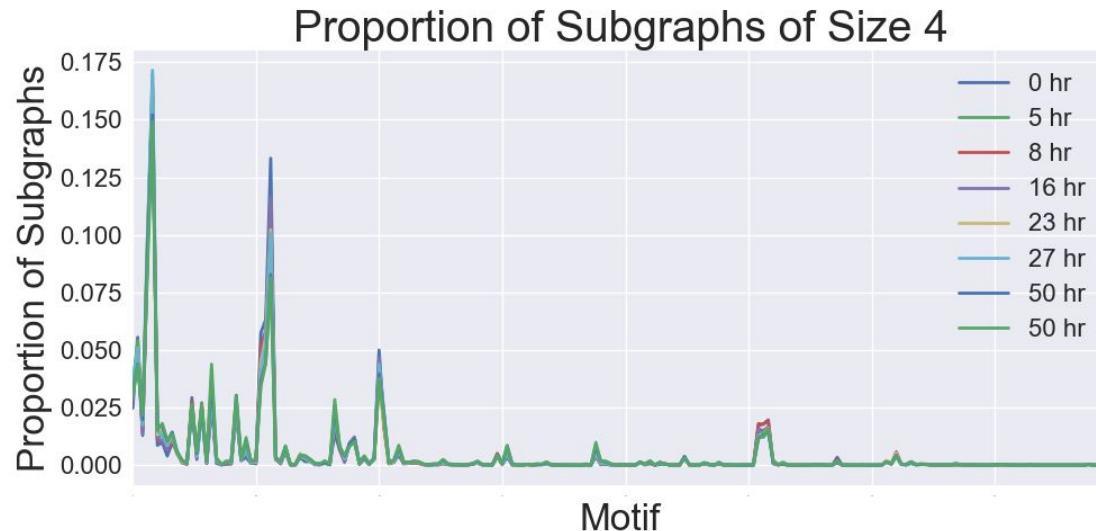
Developmental Growth of *C. elegans*

Amazingly, the relative proportions of specific subgraphs of size 3 and 4 does not differ significantly across the *C. elegans* development



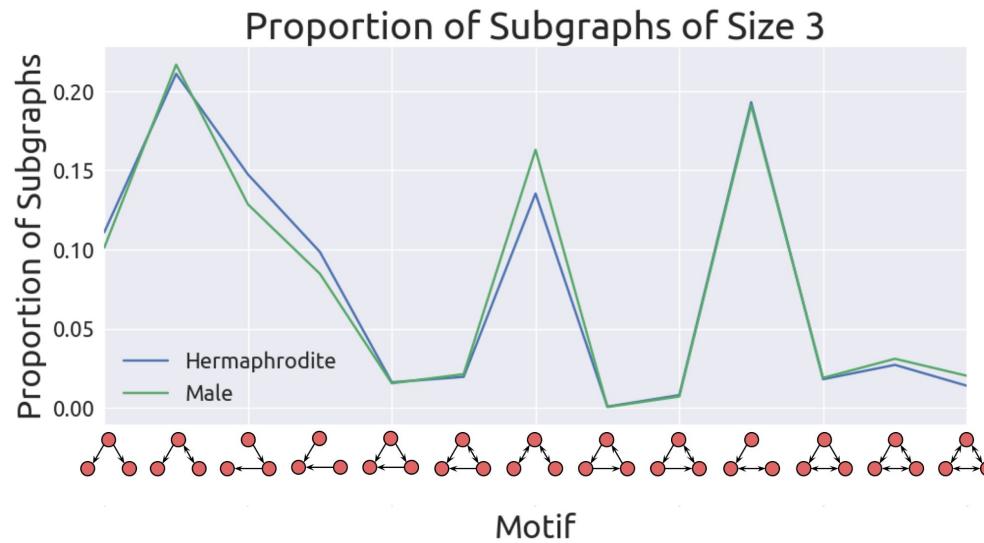
Developmental Growth of *C. elegans*

Amazingly, the relative proportions of specific subgraphs of size 3 and 4 does not differ significantly across the *C. elegans* development



Comparison Between the Two *C. elegans* Sexes

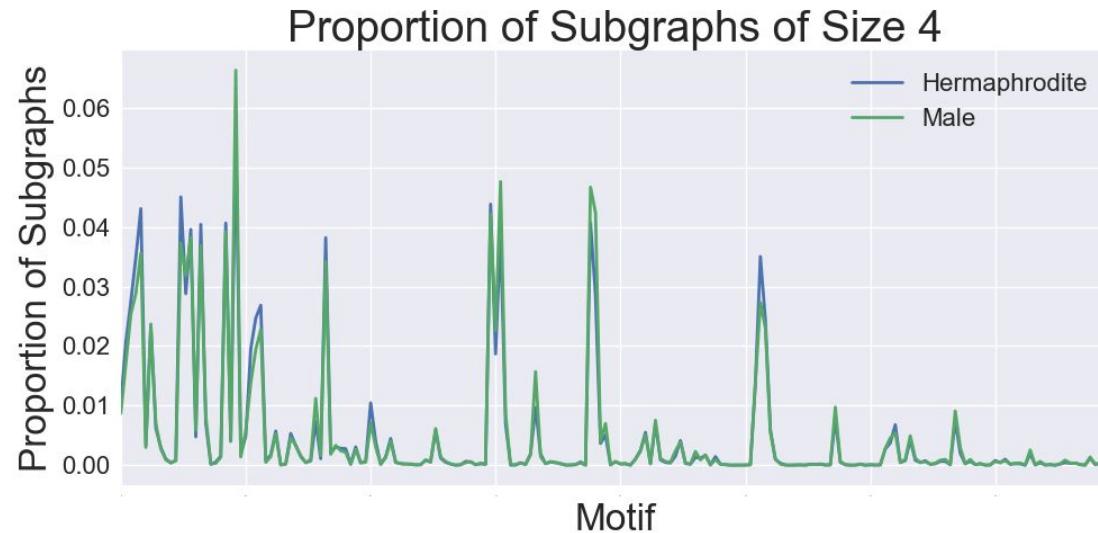
Both the adult sexes of the *C. elegans* samples compared had very similar motif counts for $k = 3$ and 4



These *C. elegans* datasets are not directly comparable to the previous set since they contain end-organs, muscles, and gap junctions

Comparison Between the Two *C. elegans* Sexes

Both the adult sexes of the *C. elegans* samples compared had very similar motif counts for $k = 3$ and 4



These *C. elegans* datasets are not directly comparable to the previous set since they contain end-organs, muscles, and gap junctions

Ablation Studies: Community-Based Motif Discovery

No. Communities	No. Subgraphs (k = 5)	Computation Time
1	12,522,283,314,604	2.77 yr
5	3,645,456,758,599	0.82 yr
10	2,130,560,777,611	0.48 yr
15	1,458,545,460,827	0.31 yr
20	1,014,016,146,436	0.22 yr
25	607,287,828,074	0.13 yr
30	456,303,664,640	0.098 yr

Ablation Studies: Edge Coloring

Dataset	No Edge/Edge Color Computation Times			
	3	4	5	6
<i>Drosophila</i>	566.63 / 1053.48 s	2.15 / 4.39 d	2.77 / 5.66 yr	N/A
<i>C. elegans</i> Herm.	0.37 / 0.77 s	14.87 / 33.40 s	726.04 / 1654.80 s	9.43 / 21.61 hr
<i>C. elegans</i> Male	0.35 / 0.77 s	13.28 / 29.24 s	585.68 / 1313.99 s	7.04 / 16.23 hr

Dataset Publication

We publish summaries of all enumerated subgraphs from the twelve connectomes to enable further analysis from the community

Graph Type	No. Subgraphs	Computation Time
Standard	13,143,518,416,074	2.92 yr
Edge Colored	12,562,956,618,837	5.67 yr
Community-Based	9,362,073,052,898	2.08 yr

Dataset Publication

We publish summaries of all enumerated subgraphs from the eleven connectomes to enable further analysis from the community

Graph Type	No. Subgraphs	Computation Time
Standard	13,143,518,416,074	2.92 yr
Edge Colored	12,562,956,618,837	5.67 yr
Community-Based	9,362,073,052,898	2.08 yr

35 trillion enumerated subgraphs

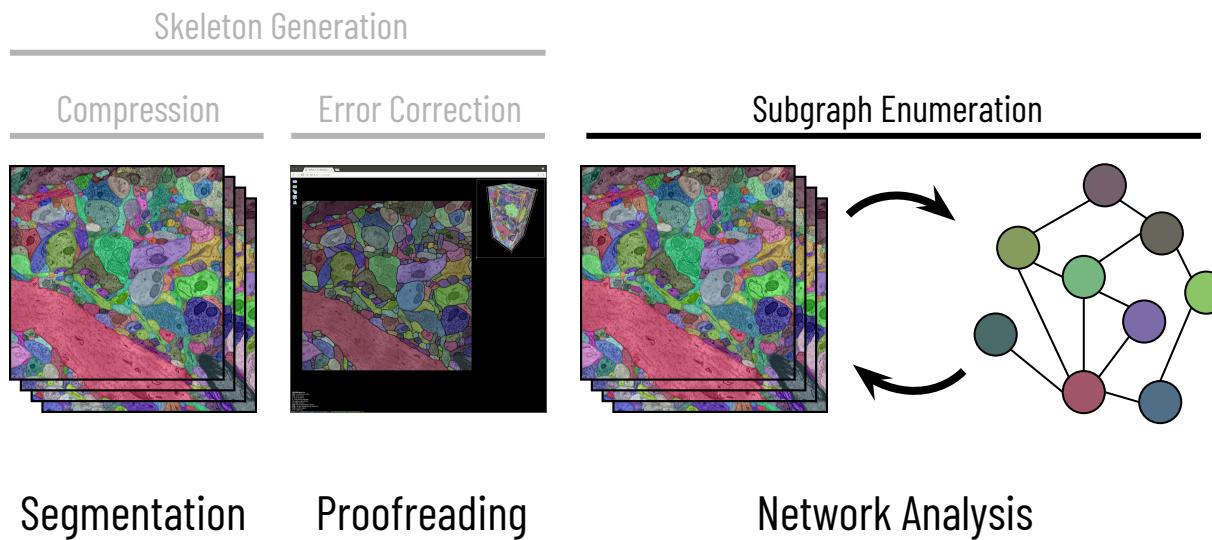
Dataset Publication

We publish summaries of all enumerated subgraphs from the eleven connectomes to enable further analysis from the community

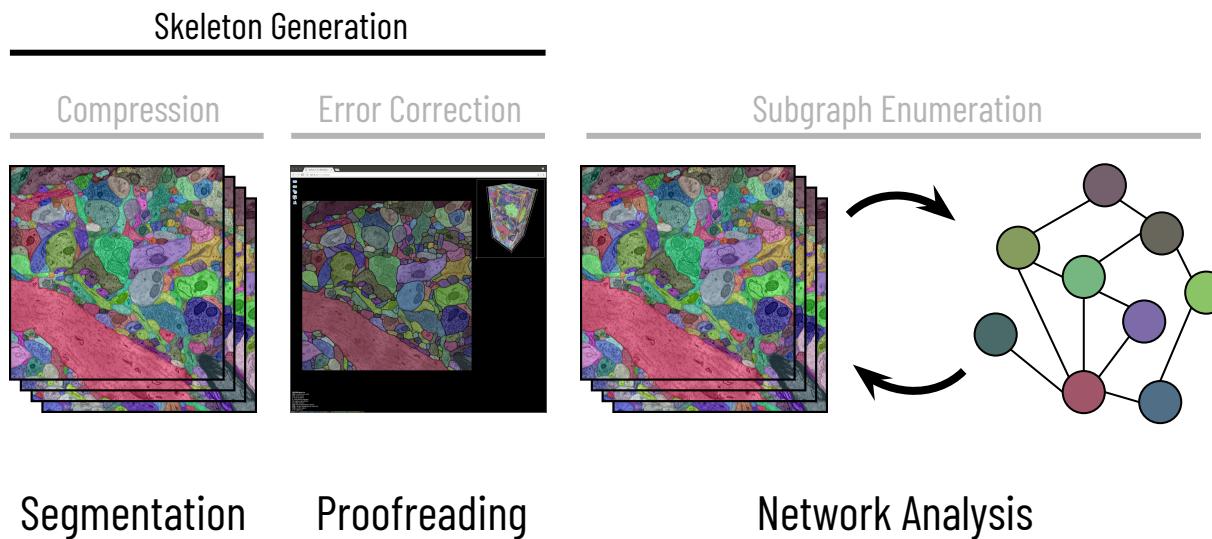
Graph Type	No. Subgraphs	Computation Time
Standard	13,143,518,416,074	2.92 yr
Edge Colored	12,562,956,618,837	5.67 yr
Community-Based	9,362,073,052,898	2.08 yr

35 trillion enumerated subgraphs over 10.67 years of computation time

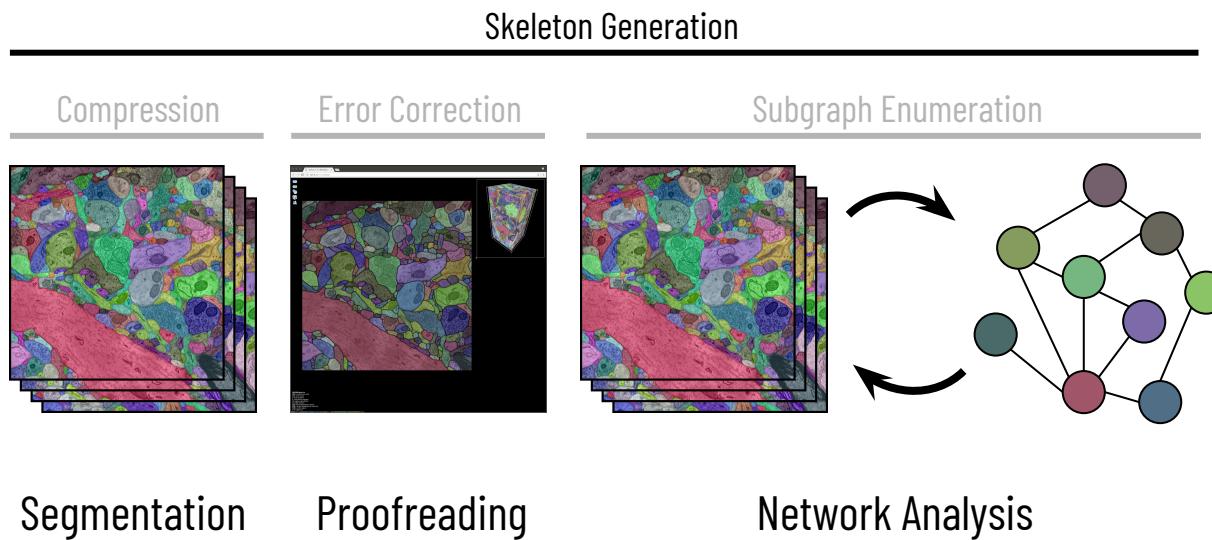
Biologically-Aware Algorithms Along the Connectomics Pipeline



Biologically-Aware Algorithms Along the Connectomics Pipeline



Biologically-Aware Algorithms Along the Connectomics Pipeline



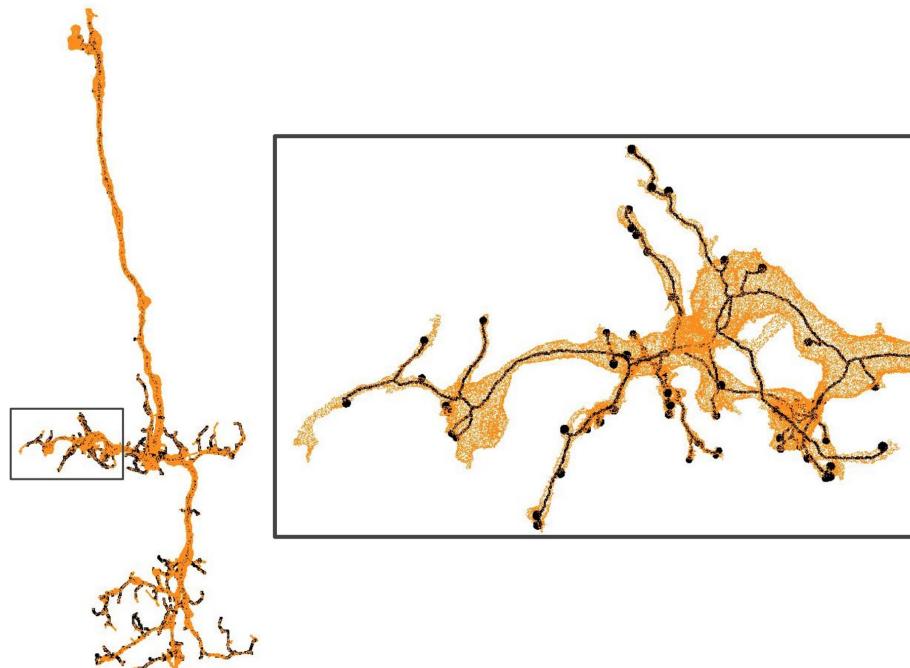
Scalable Biologically-Aware Skeleton Generation for Connectomic Volumes

Brian Matejek, Tim Franzmeyer, Donglai Wei, Xueying Wang, Jinglin Zhao,
Kálmán Palágyi, Jeff W. Lichtman, and Hanspeter Pfister

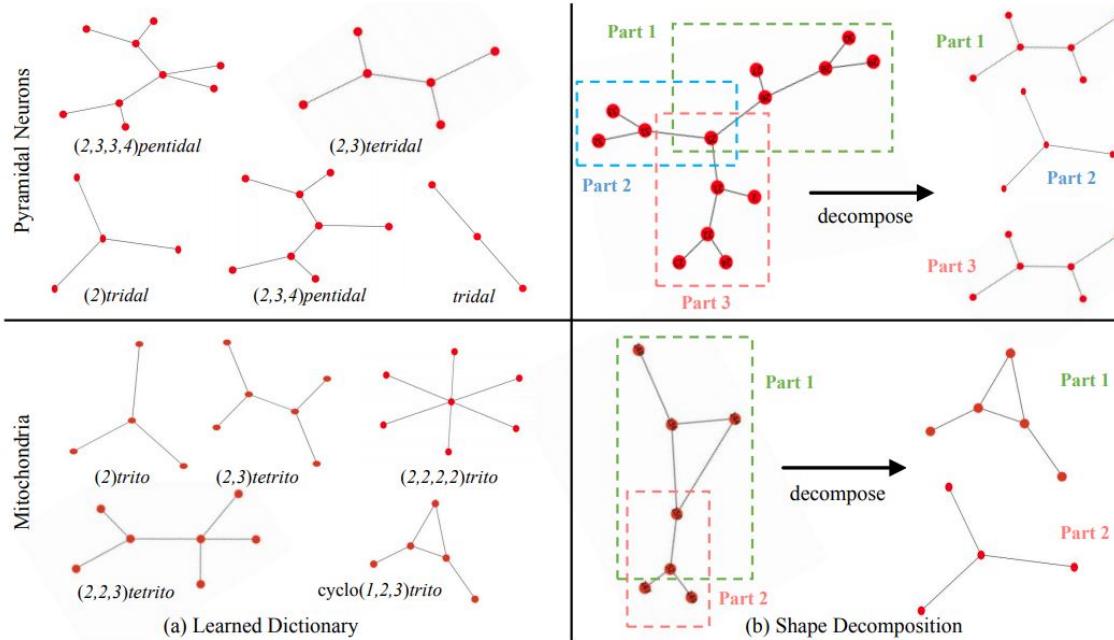
International Conference on Medical Image Computing and Computer Assisted Intervention, 2019
Under Review, 2021

Skeletonized Representations of Label Volumes

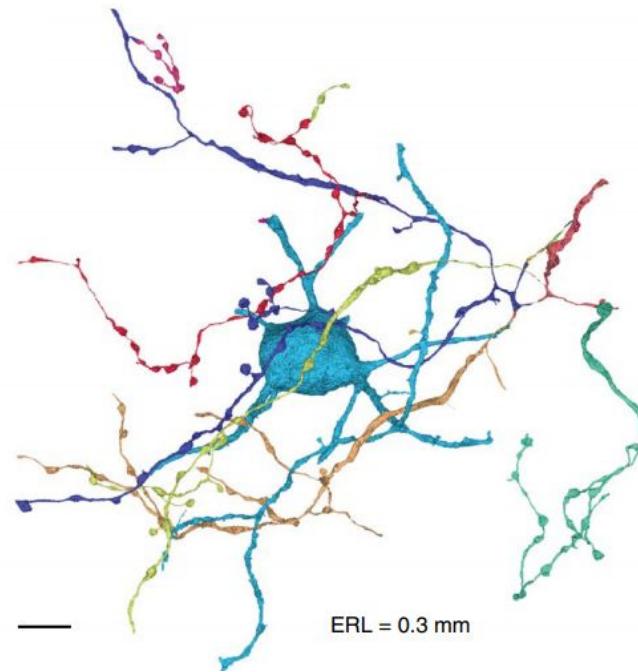
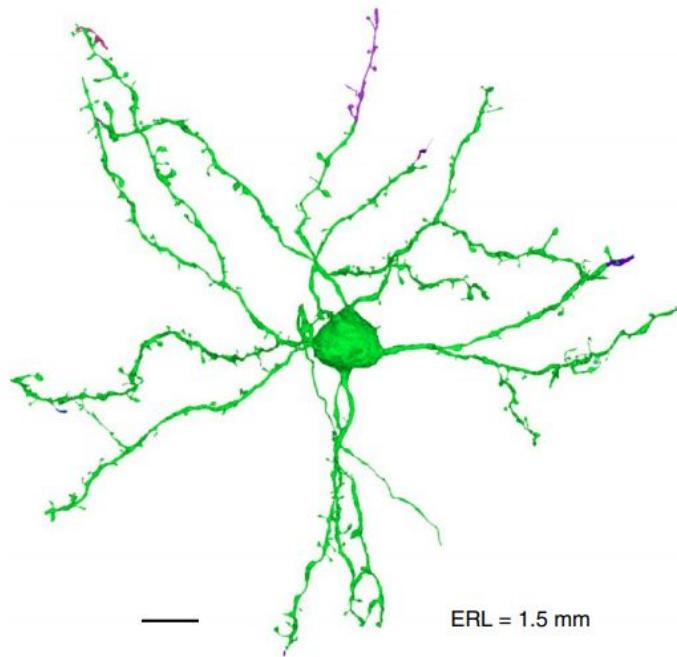
Across the connectomics pipeline, skeletonized representations of the label volumes are incredibly useful



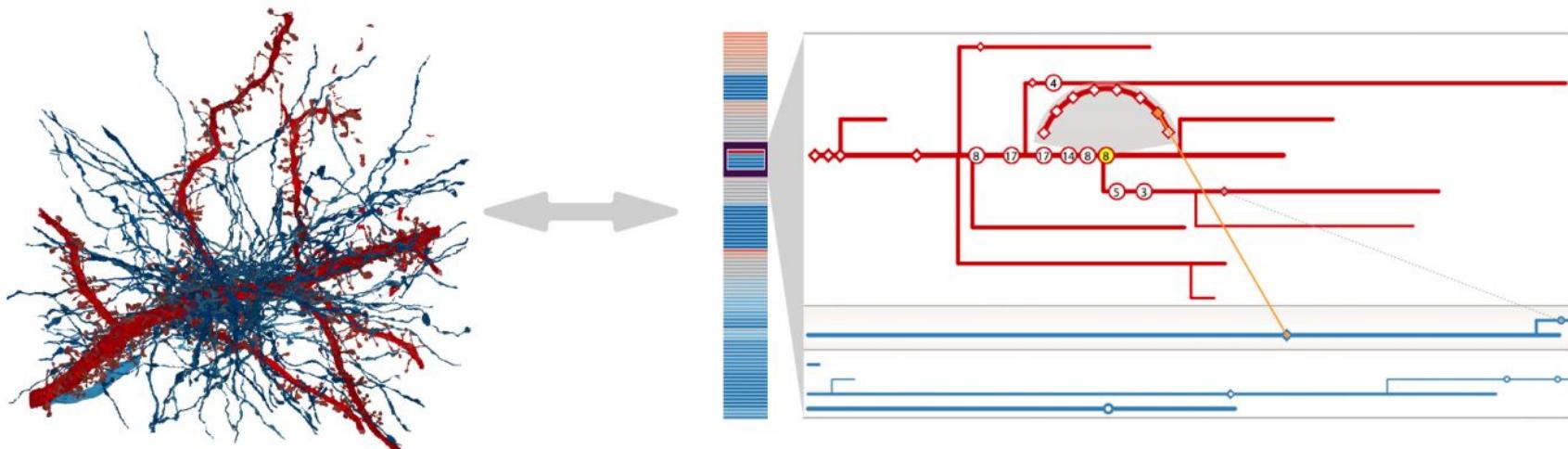
Skeletonization for Analysis



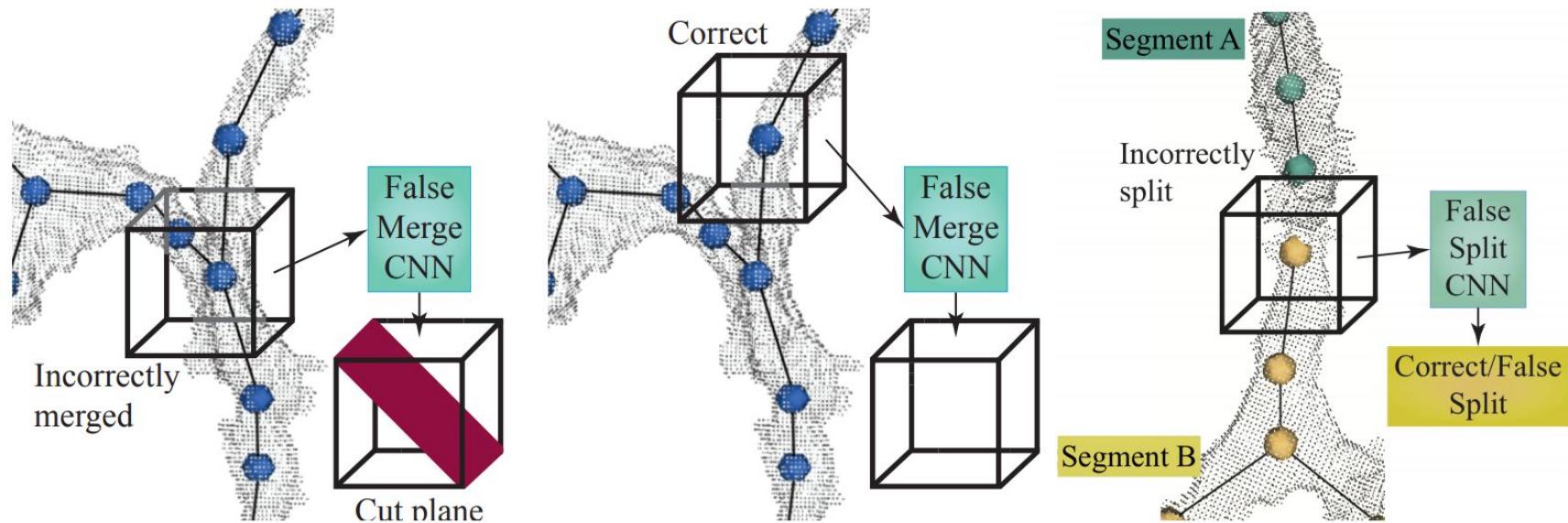
Skeletonization for Segmentation Evaluation



Skeletonization for Visualization



Skeletonization for Error Correction

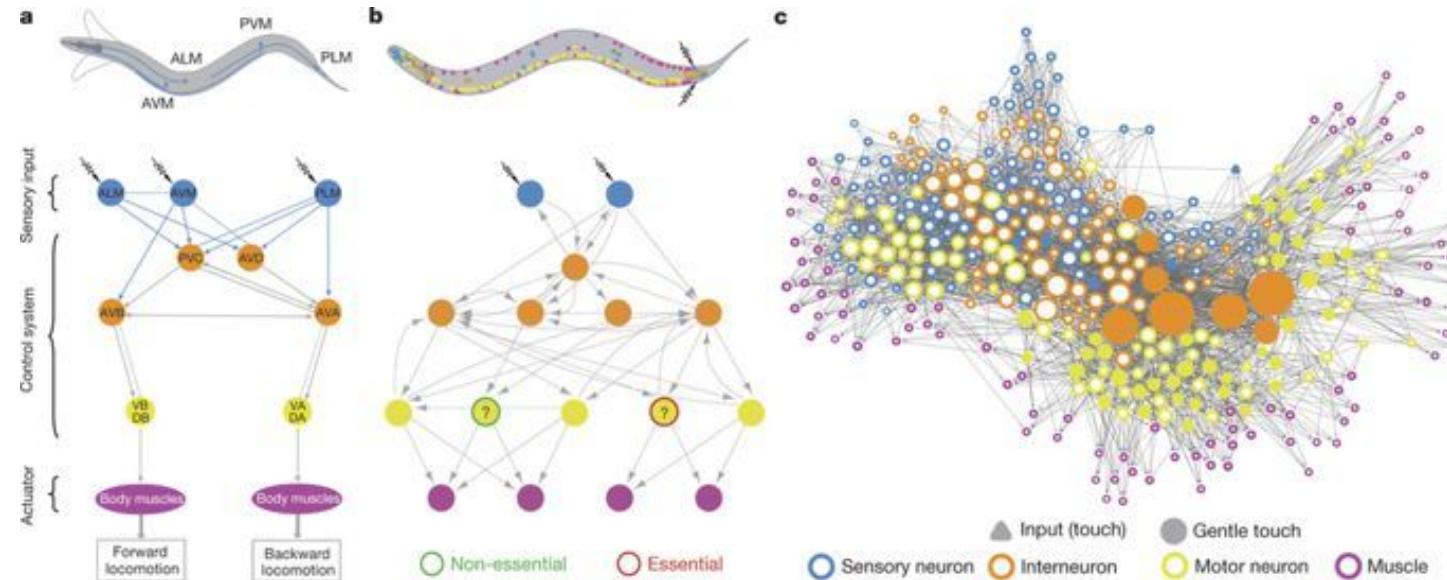


Berning *et al.*, SeqEM: Efficient Image Analysis for High-Resolution Connectomes, Neuron 2015

Dmitriev *et al.*, Efficient Correction for EM Connectomics with Skeletal Representation, BMVC 2018

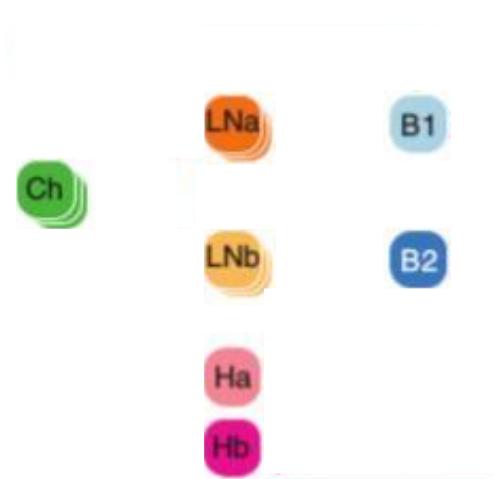
Matejek *et al.*, Biologically-Constrained Graphs for Global Connectomics Reconstruction, CVPR 2018

Improving Wiring Diagrams with Skeletons



Current Graph-Based Wiring Diagram Methodology

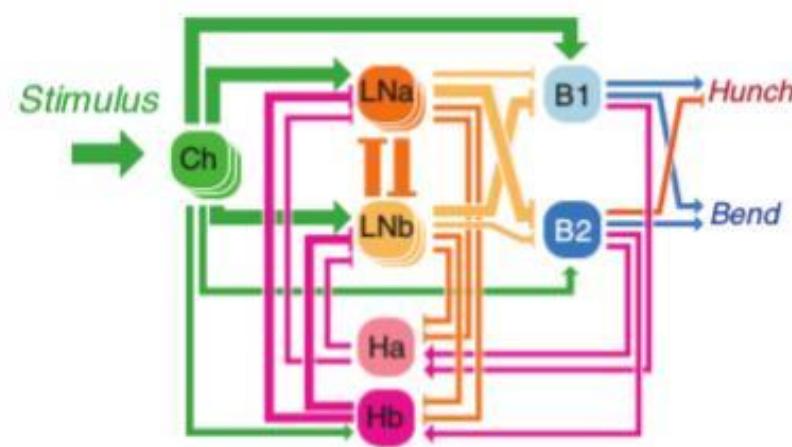
Each node represents one neuron



Current Graph-Based Wiring Diagram Methodology

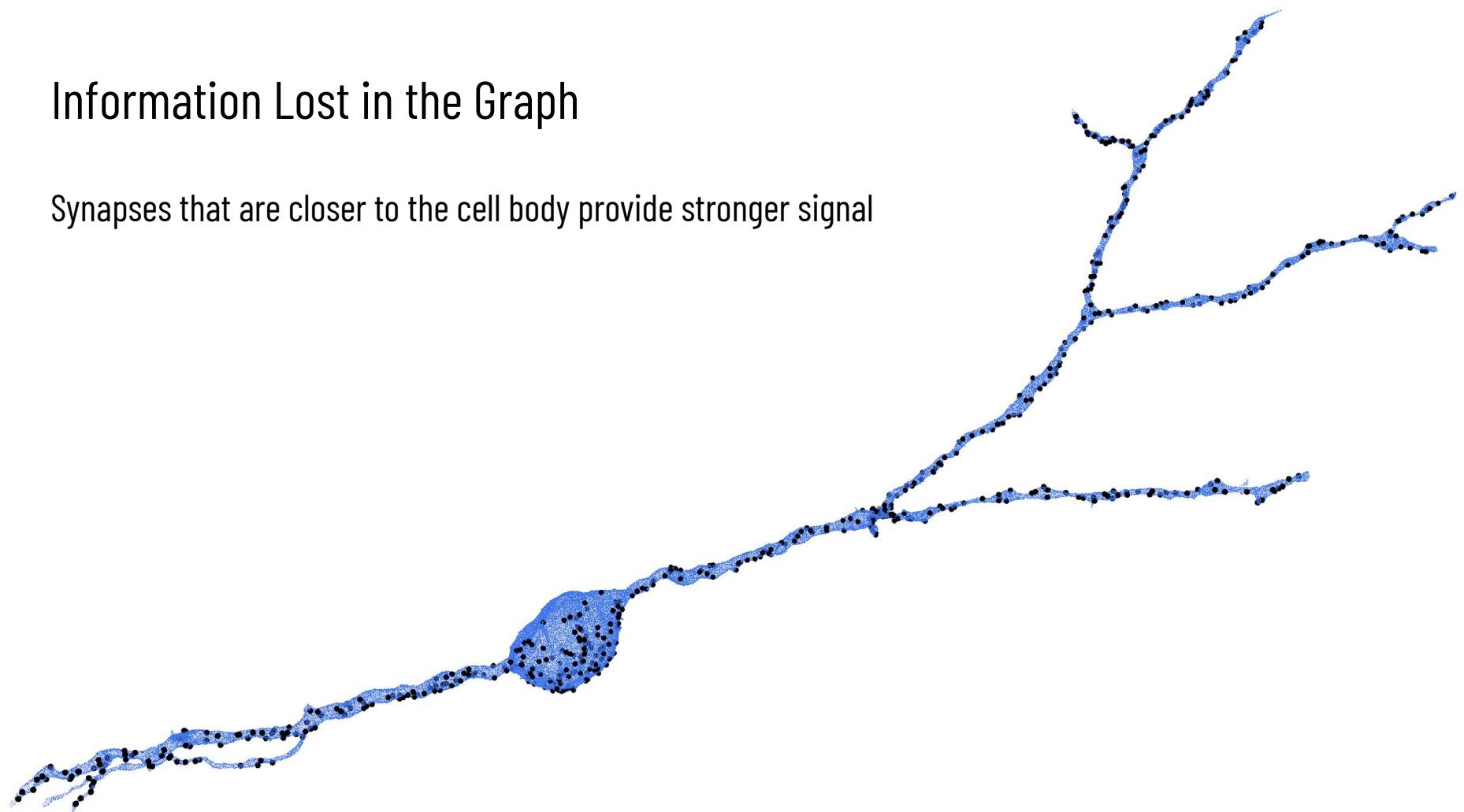
Each node represents one neuron

Weighted edges indicate number of synapses between two neurons



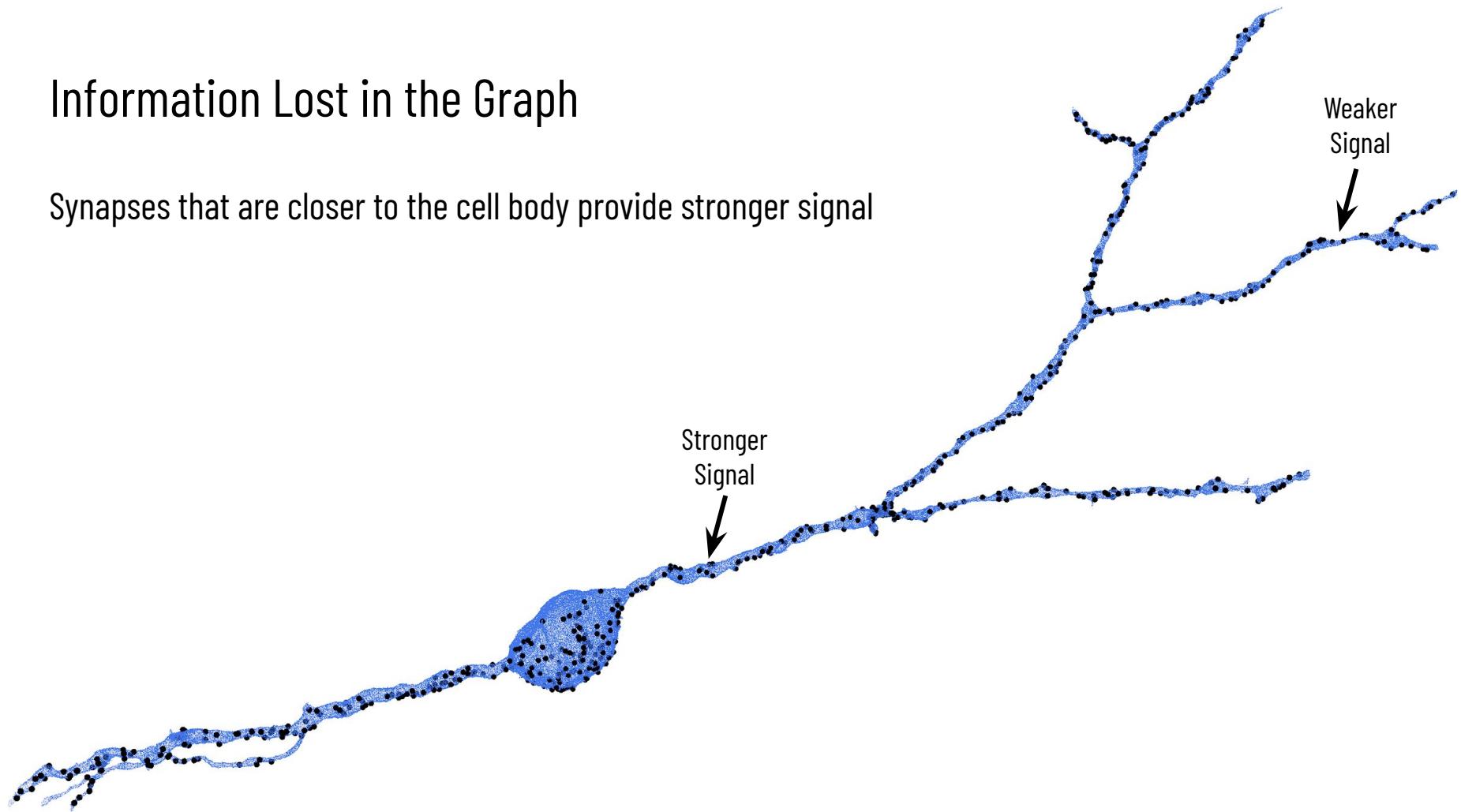
Information Lost in the Graph

Synapses that are closer to the cell body provide stronger signal

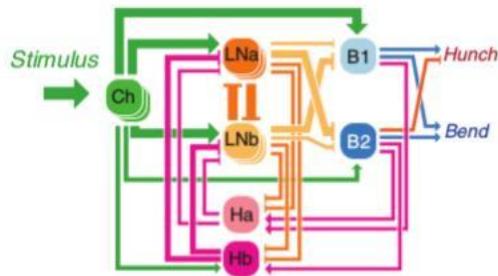


Information Lost in the Graph

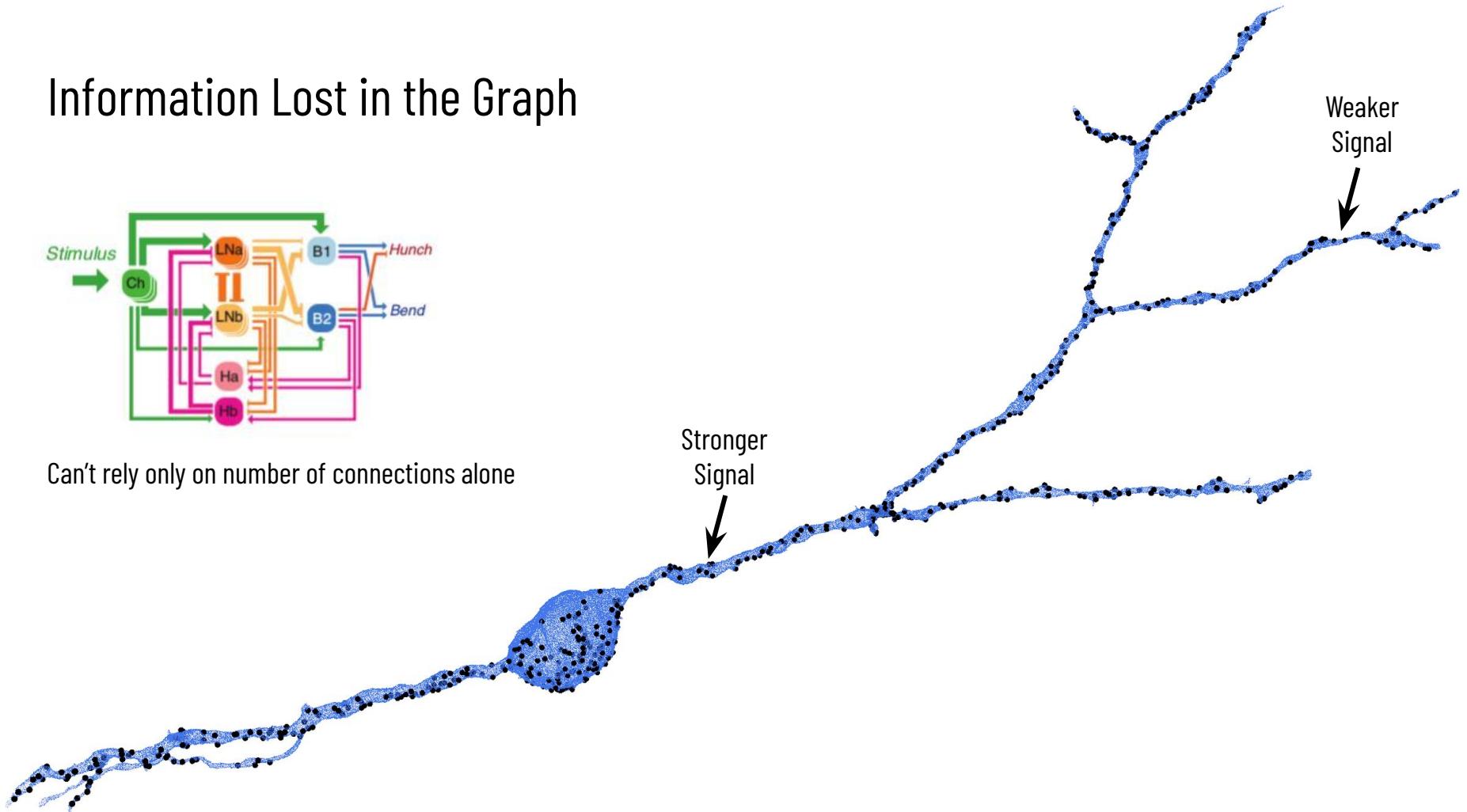
Synapses that are closer to the cell body provide stronger signal



Information Lost in the Graph



Can't rely only on number of connections alone



Information Lost in the Graph

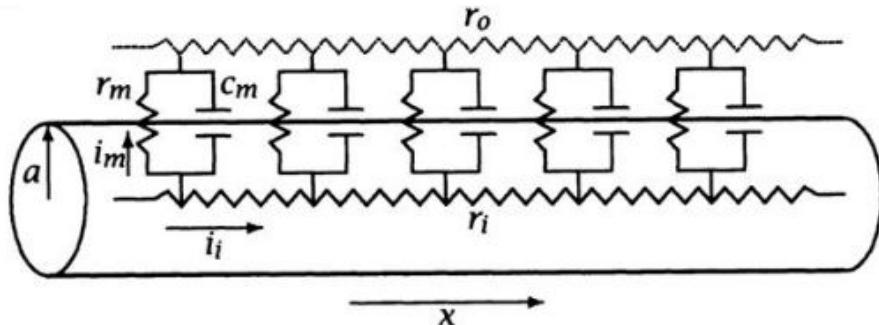
Width of neuron along path from synapse to cell body

Information Lost in the Graph

Width of neuron along path from synapse to cell body

Cable Theory:

$$\frac{1}{r_l} \frac{\partial^2 V}{\partial x^2} = c_m \frac{\partial V}{\partial t} + \frac{V}{r_m}$$

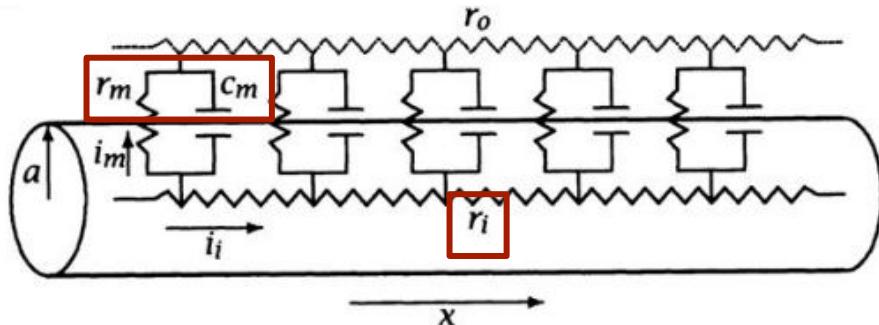


Information Lost in the Graph

Width of neuron along path from synapse to cell body

Cable Theory:

$$\frac{1}{r_l} \frac{\partial^2 V}{\partial x^2} = c_m \frac{\partial V}{\partial t} + \frac{V}{r_m}$$

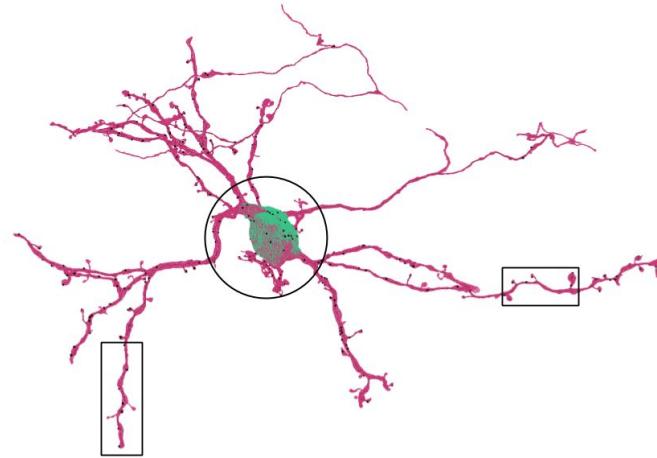
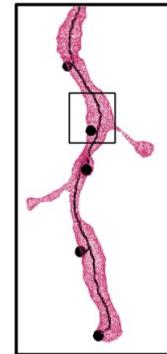
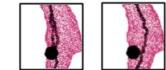


r_l, c_m, r_m are functions of the neurite diameter

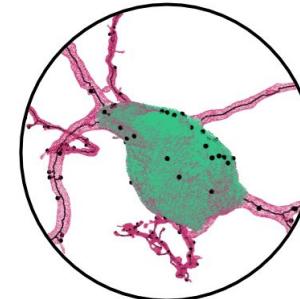
Biologically-Aware Skeleton Generation

Generate skeletons that connect all synapses to the cell body

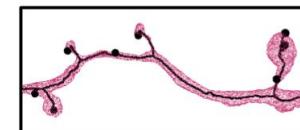
Bubble Filling



Soma Detection



Synapse Connectivity

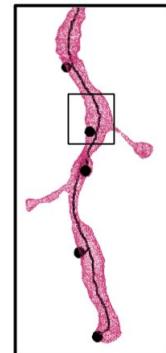
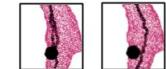


Biologically-Aware Skeleton Generation

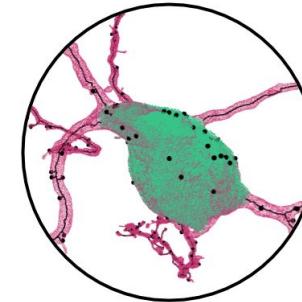
Generate skeletons that connect all synapses to the cell body

Calculate vital geometric statistics needed for evaluating the perceived synaptic strength like geodesic distance from the synapse to the cell body and the width along each neurite

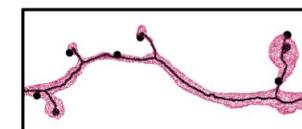
Bubble Filling



Soma Detection

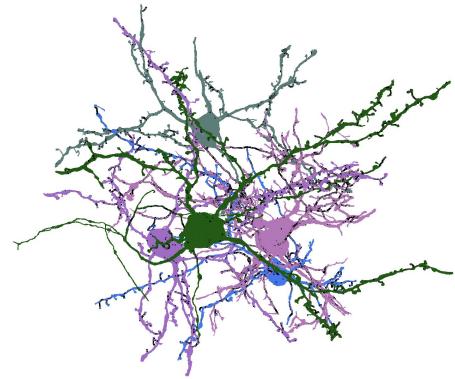


Synapse Connectivity



Block-Based Processing

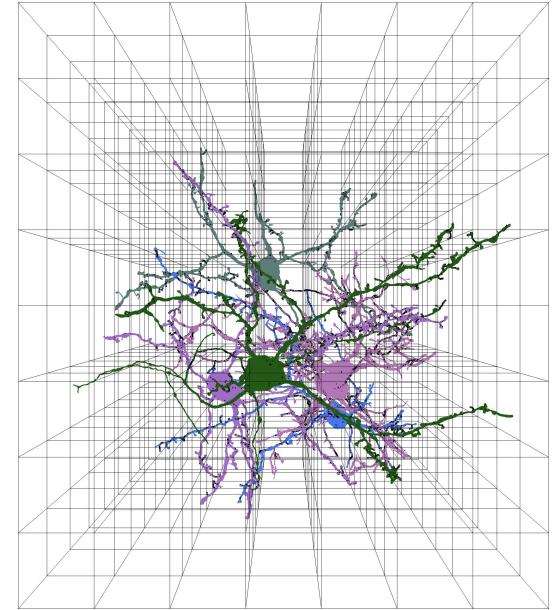
Connectome label volumes now regularly exceed hundreds of gigabytes



Block-Based Processing

Connectome label volumes now regularly exceed hundreds of gigabytes

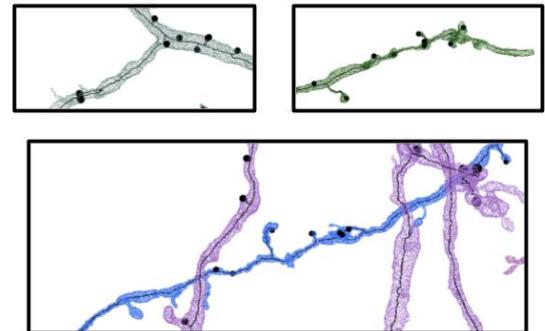
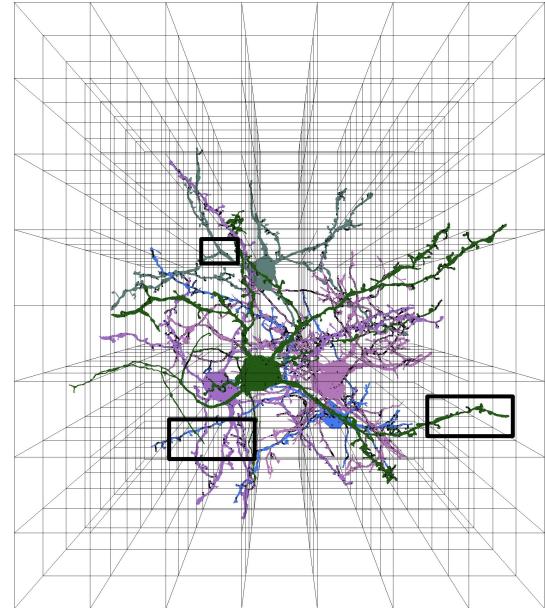
Processing needs to be mostly block-based for widespread adoption



Block-Based Processing

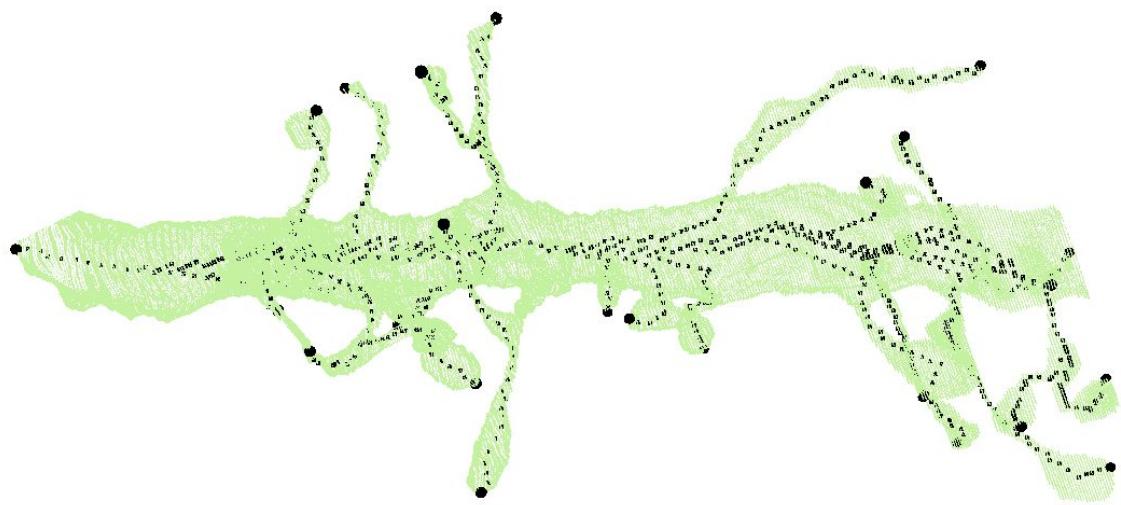
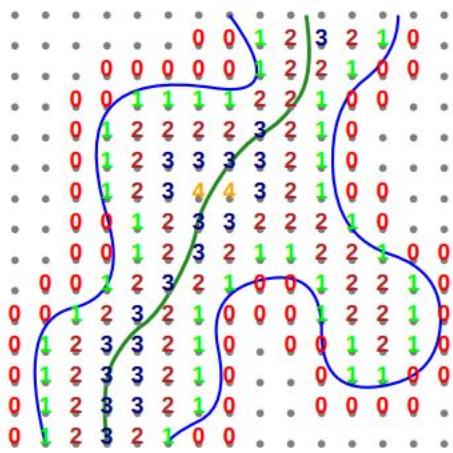
Connectome label volumes now regularly exceed hundreds of gigabytes

Processing needs to be mostly block-based for widespread adoption



Tree-structure Extraction Algorithm for Accurate and Robust Skeletons

TEASER is the predominantly used skeleton generation technique for connectomics



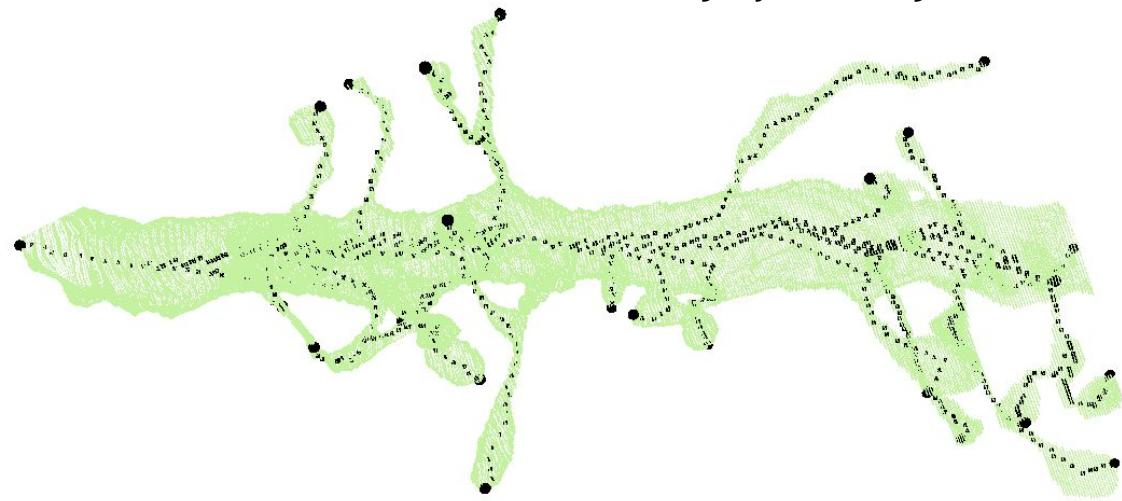
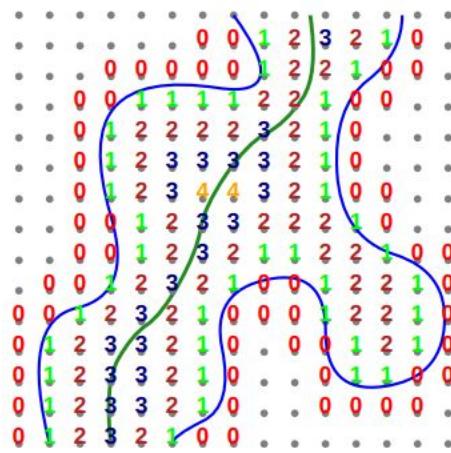
Sato et al., TEASER: Tree-structure Extraction Algorithm for Accurate and Robust Skeletons, PCCGA 2000

Zhao et al., NeuTu: Software for Collaborative, Large-Scale, Segmentation-Based Connectome Reconstruction, Frontiers in Neural Circuits 2018

Tree-structure Extraction Algorithm for Accurate and Robust Skeletons

TEASER is the predominantly used skeleton generation technique for connectomics

The algorithm continually identifies distant voxels to attach to a root voxel using Dijkstra's algorithm

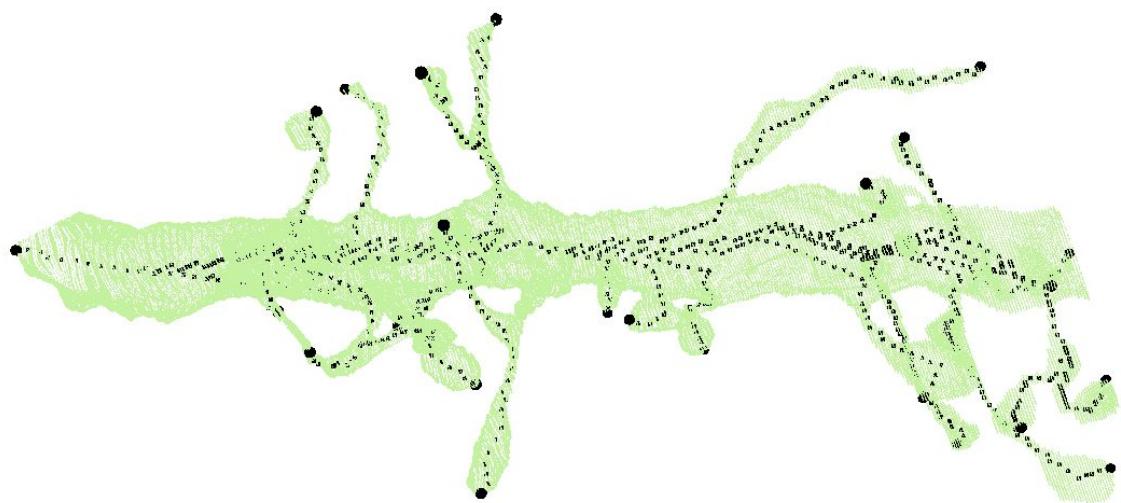
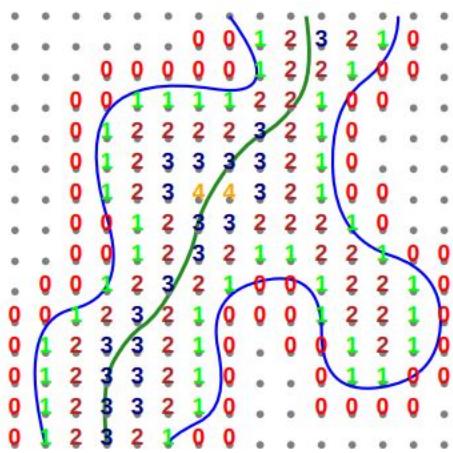


Sato et al., TEASER: Tree-structure Extraction Algorithm for Accurate and Robust Skeletons, PCCGA 2000

Zhao et al., NeuTu: Software for Collaborative, Large-Scale, Segmentation-Based Connectome Reconstruction, Frontiers in Neural Circuits 2018

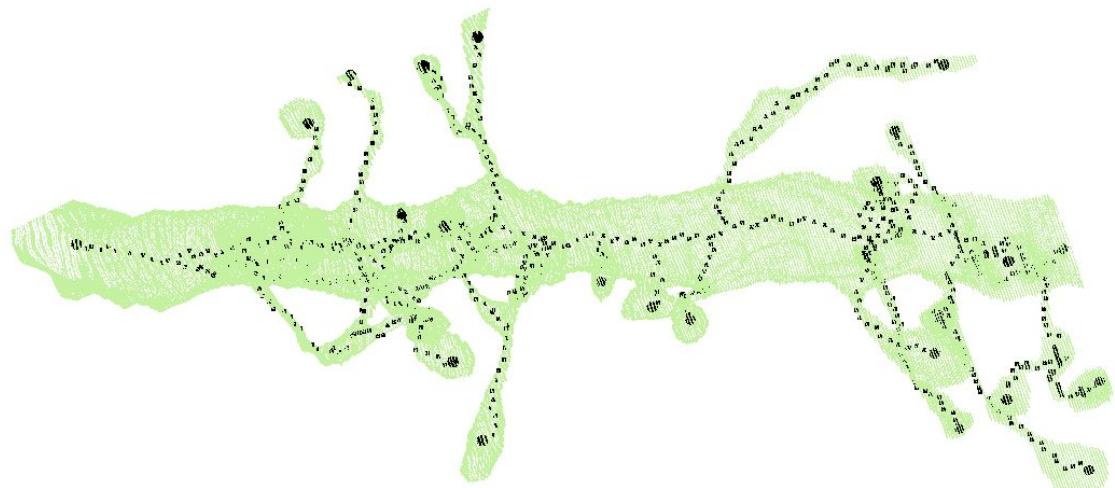
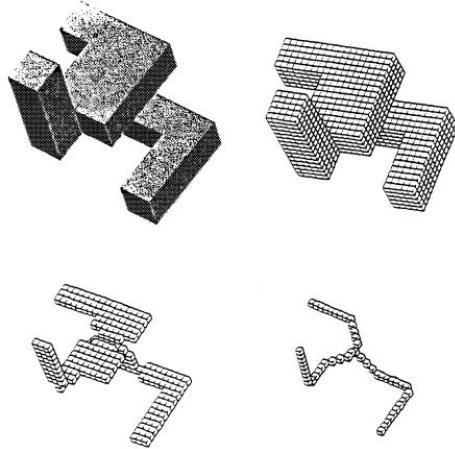
Tree-structure Extraction Algorithm for Accurate and Robust Skeletons

Silversmith *et al.*, have implemented a block-based version of the algorithm for larger connectome volumes



Topological Thinning

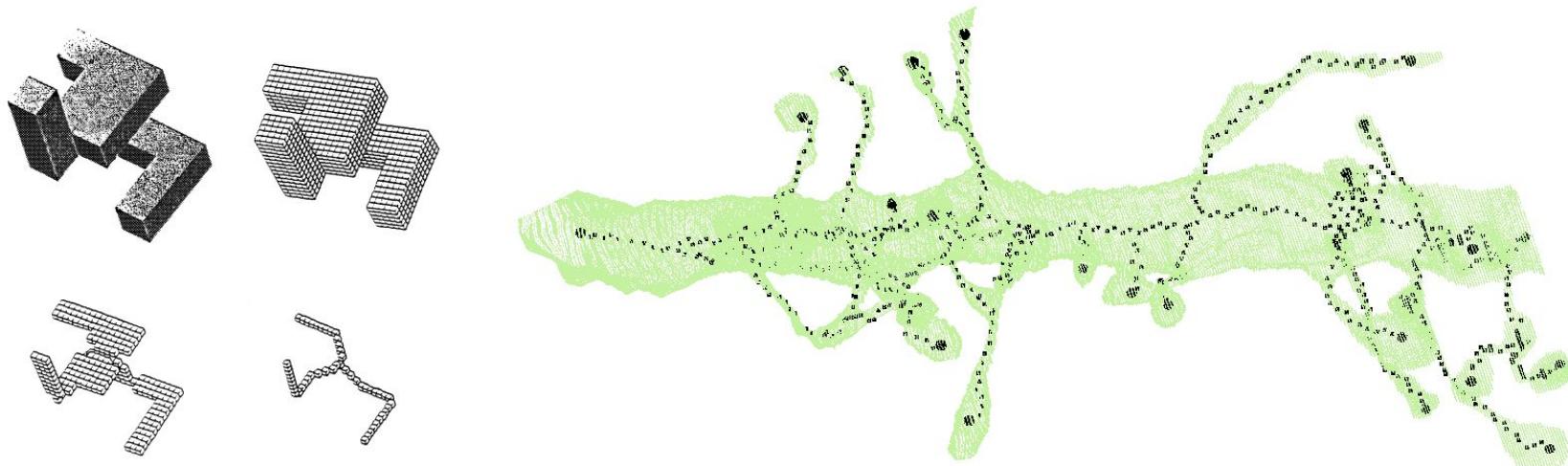
Topological thinning algorithms gradually erode the surface of a volume to a centerline



Topological Thinning

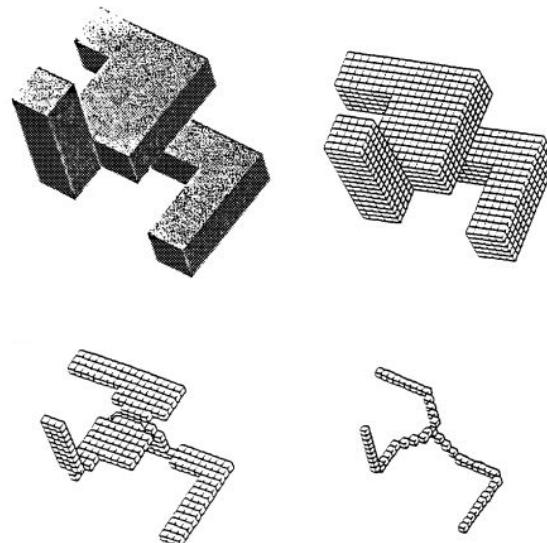
Topological thinning algorithms erode the surface of a volume to a centerline while preserving topology

These algorithms rely only on the immediate local neighborhood around a voxel to determine deletion



Topological Thinning

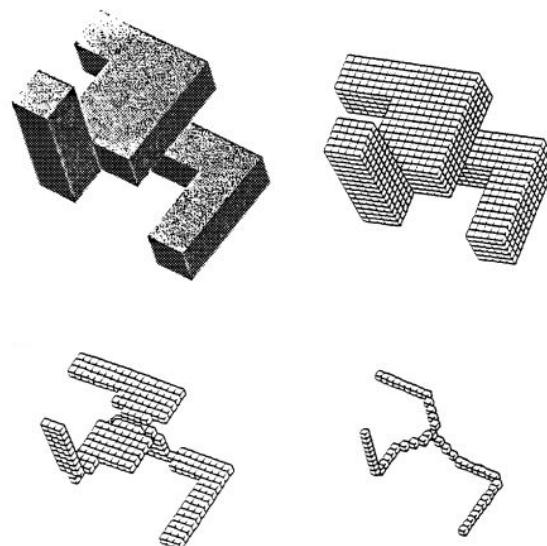
Topological thinning algorithms iteratively consider all voxels on the boundary of a volume for deletion



Topological Thinning

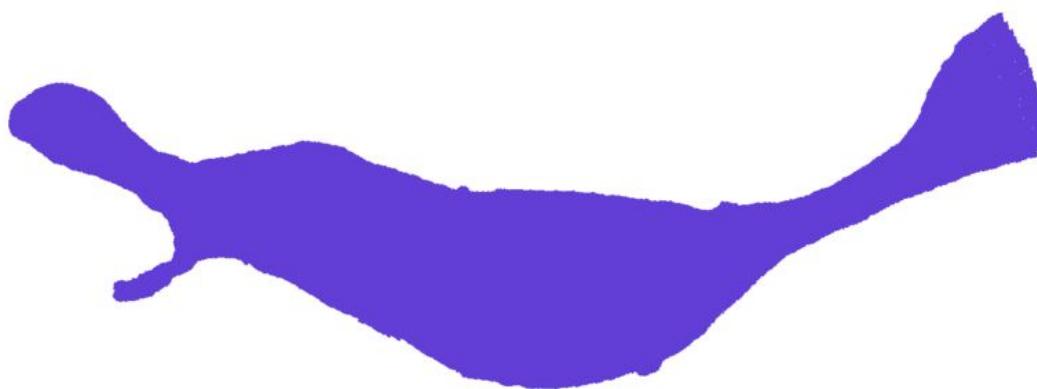
Topological thinning algorithms iteratively consider all voxels on the boundary of a volume for deletion

Points that are considered “Simple” are deleted, i.e., those that do not change the topology of the object



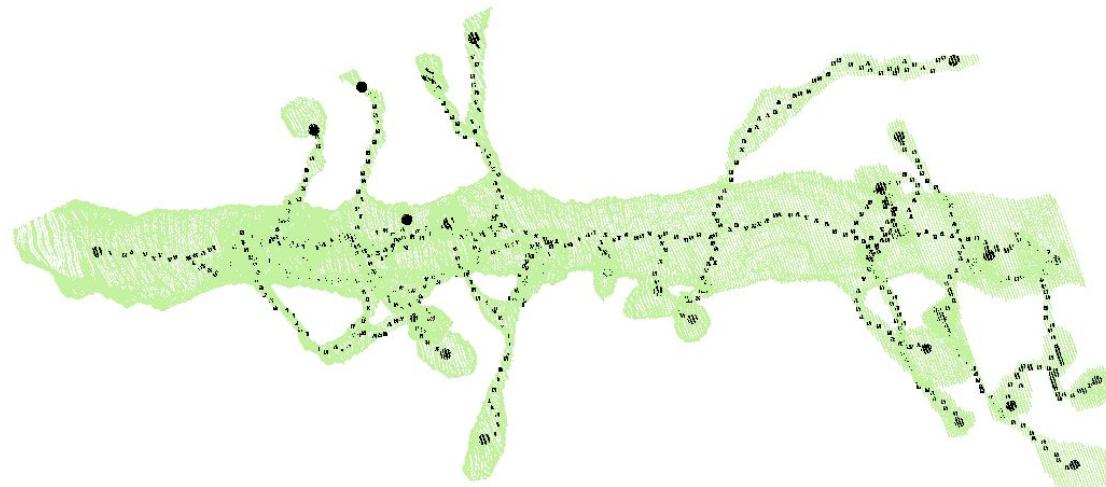
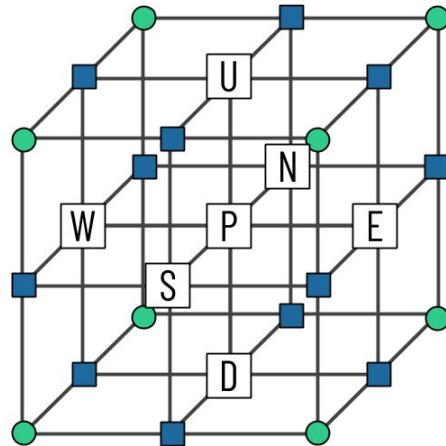
Topological Thinning

However, certain points (e.g., endpoints) are preserved despite being simple

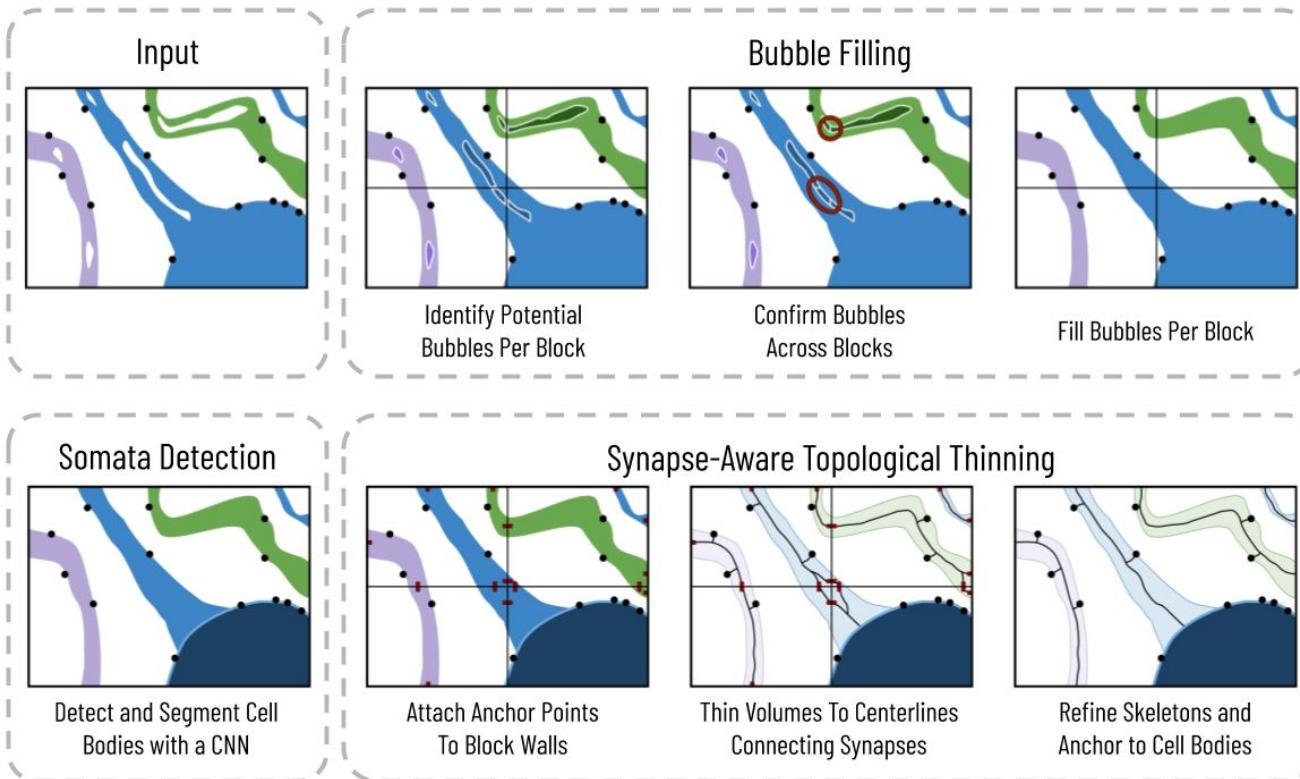


Isthmus Thinning

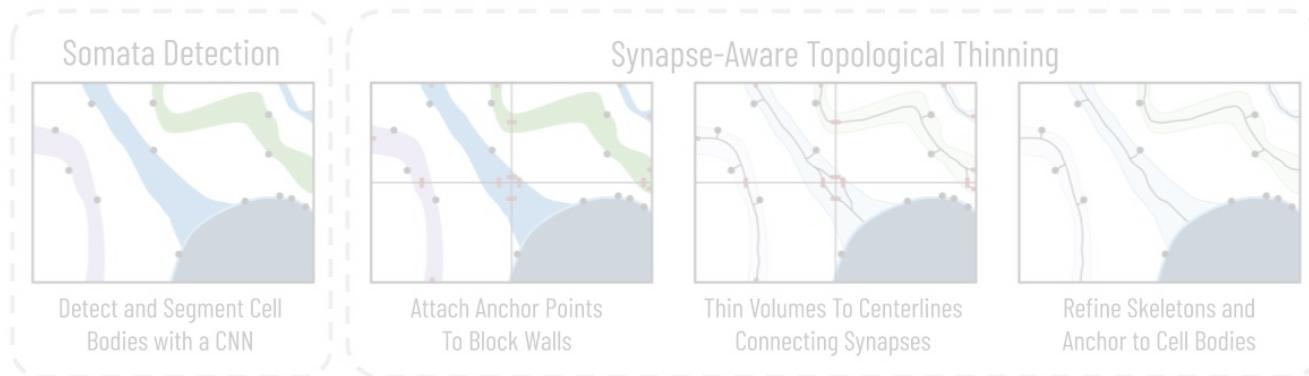
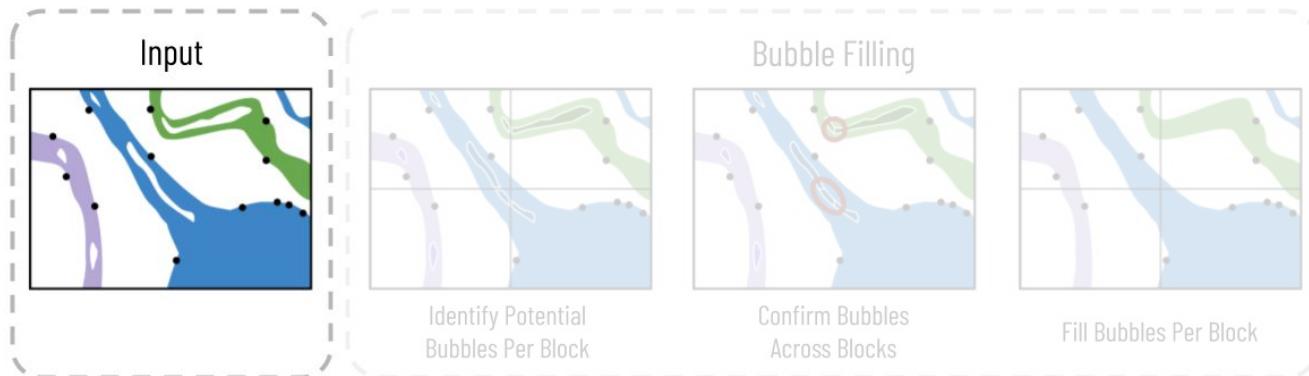
Other extensions on topological thinning define classes of voxels (e.g., isthmuses) that cannot be deleted to both preserve topology and create more expressive skeletons



Proposed Skeleton Generation

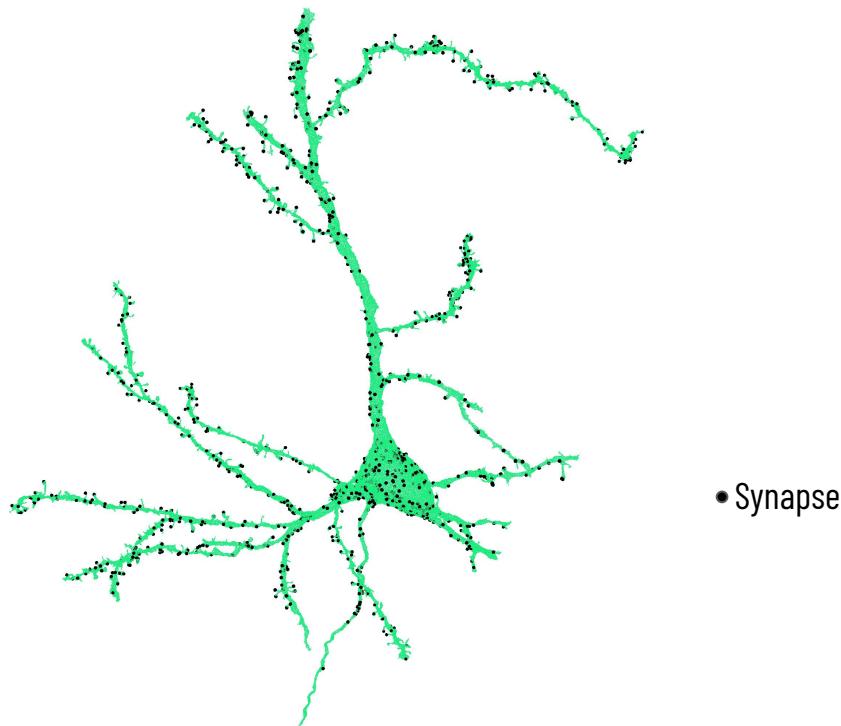


Input



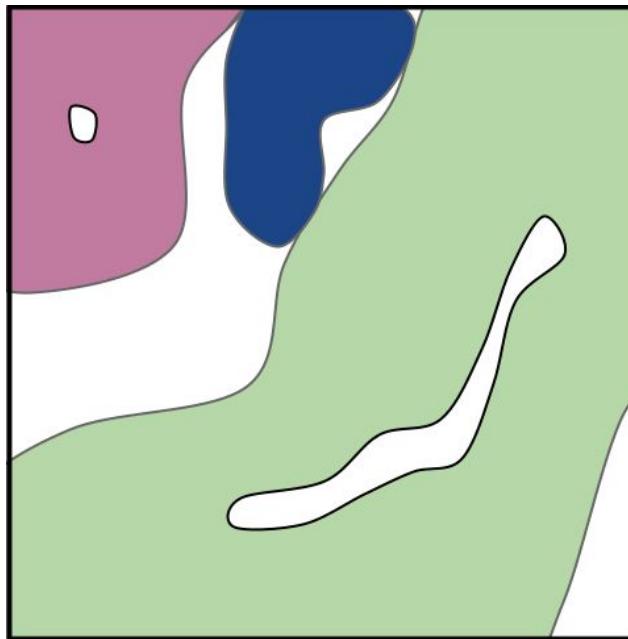
Input Segmentations and Synapses

Our method takes as input a label volume and a corresponding set of synapse locations



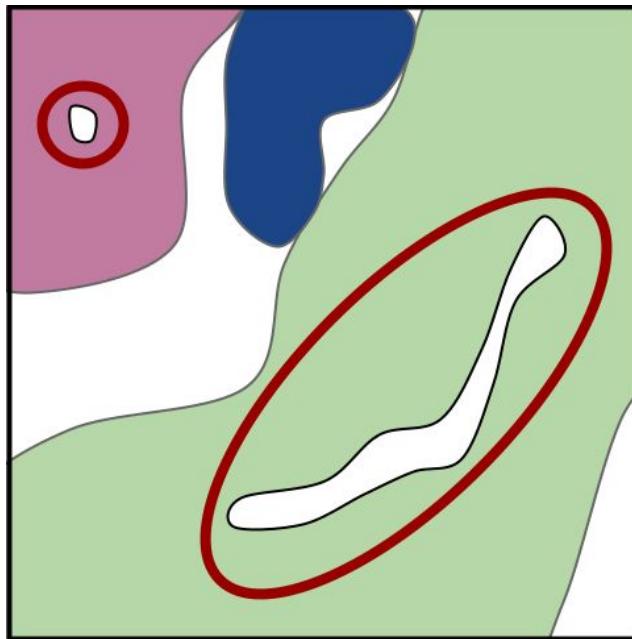
Bubbles in Input Segmentation

Segmentations can contain millions of bubbles



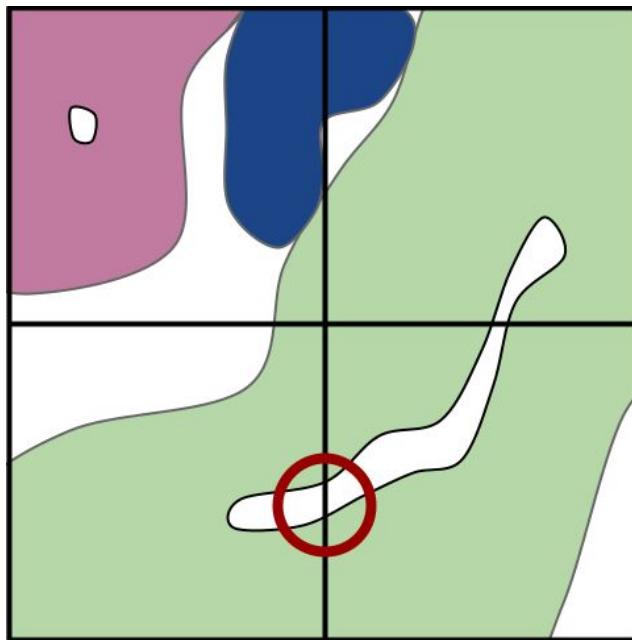
Bubbles in Input Segmentation

Segmentations can contain millions of bubbles, i.e., pockets of mislabeled voxels contained within a label



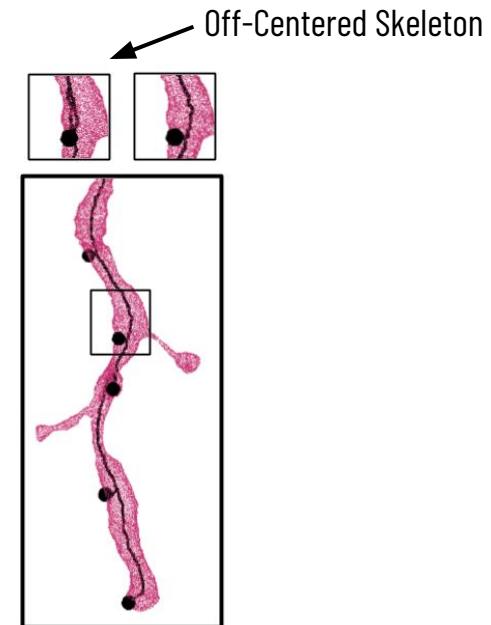
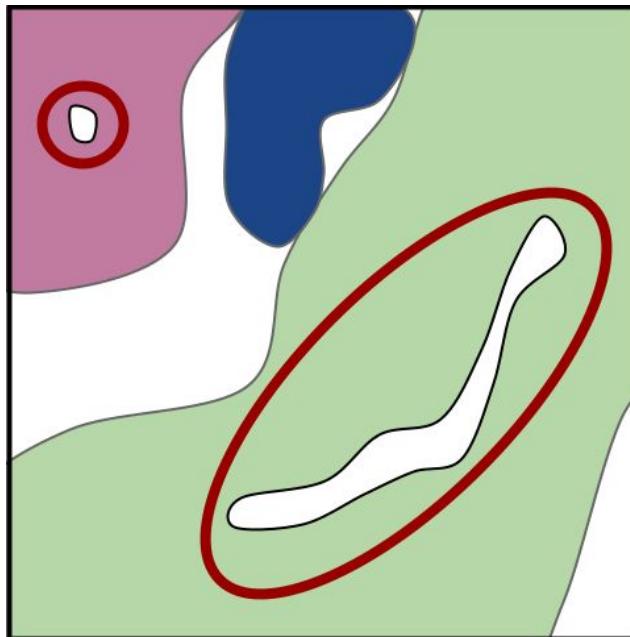
Bubbles in Input Segmentation

These bubbles can span across multiple blocks and therefore we need to consider global information

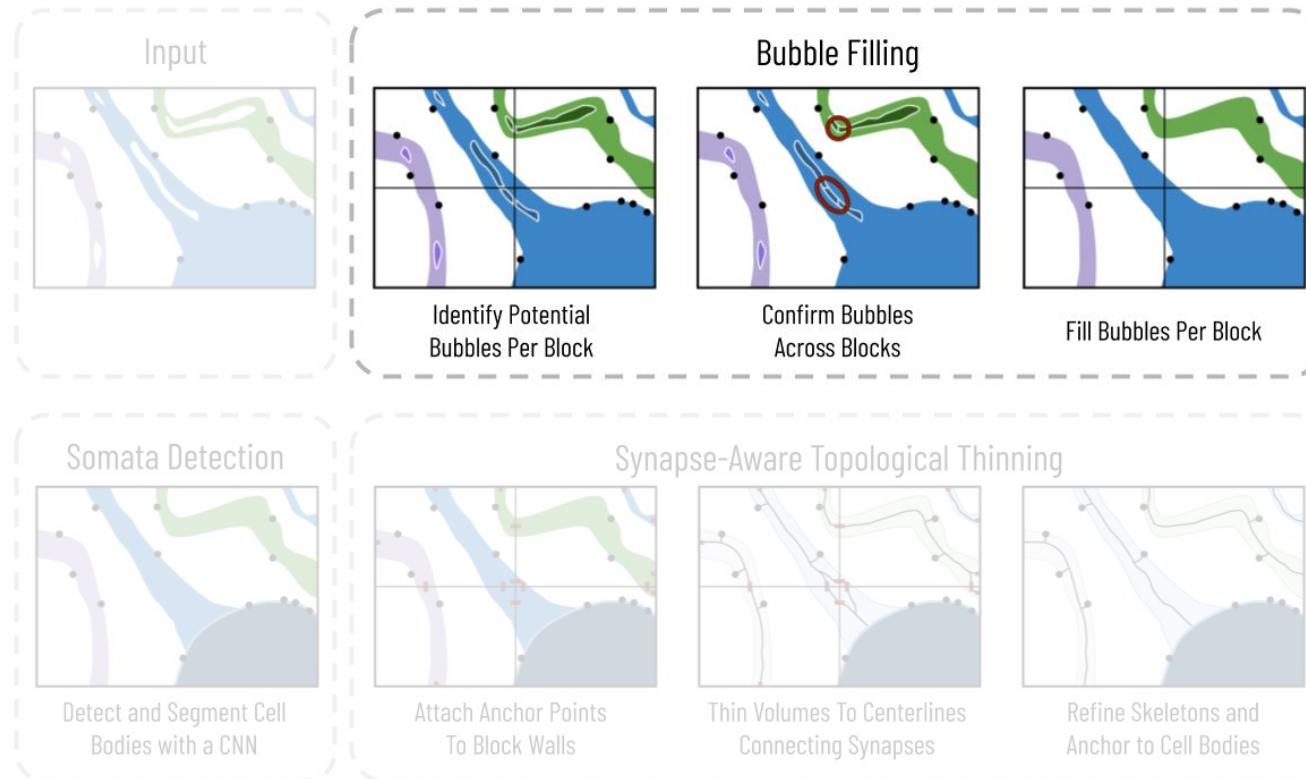


Bubbles in Input Segmentation

These biologically-infeasible bubbles increase thinning runtime and cause errors in width estimation

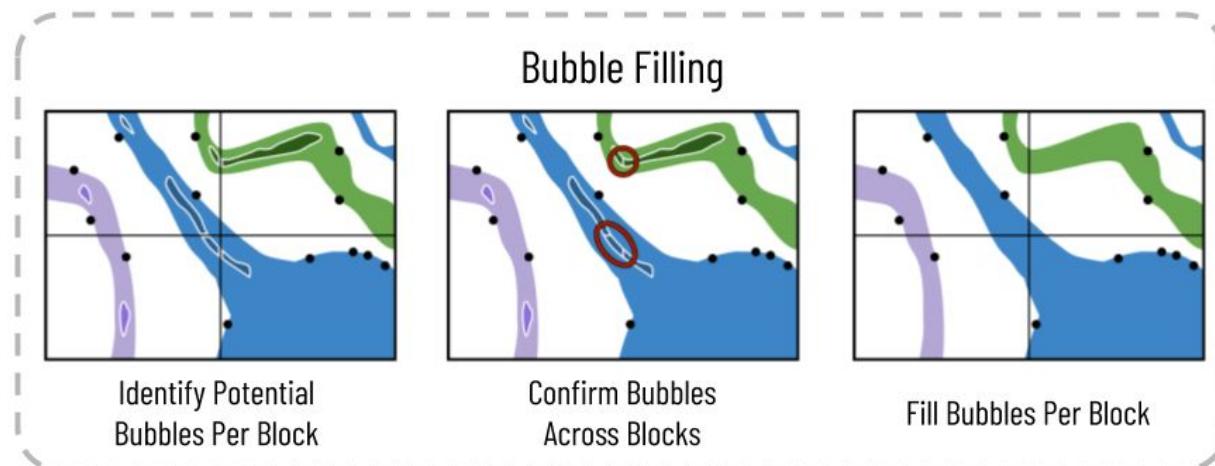


Goal: Fill bubbles in the input segmentation



Parallelizable Bubble Filling

We divide the bubble filling process into two computationally expensive but parallelizable operations and one computationally cheap step that requires global scope



Identify Potential Bubbles Per Block

Our method identifies bubbles per block

5	5	5	8	8	8		8	8	8	-37	-37	-37
5	5	8	8	8	-1		-37	-37	-37	-37	-37	-37
5	-2	8	8	8	8		8	8	8	8	8	-37
5	8	8	-3	-3	-3		-38	-38	8	8	-39	8
-4	8	8	8	8	8		8	8	8	8	-39	8
-4	-4	8	8	8	8		8	8	8	8	8	-40

Identify Potential Bubbles Per Block

Our method identifies bubbles per block, links bubbles across blocks

5	5	5	8	8	8		8	8	8	-37	-37	-37
5	5	8	8	8	-1		-37	-37	-37	-37	-37	-37
5	-2	8	8	8	8		8	8	8	8	8	-37
5	8	8	-3	-3	-3		-38	-38	8	8	-39	8
-4	8	8	8	8	8		8	8	8	8	-39	8
-4	-4	8	8	8	8		8	8	8	8	8	-40

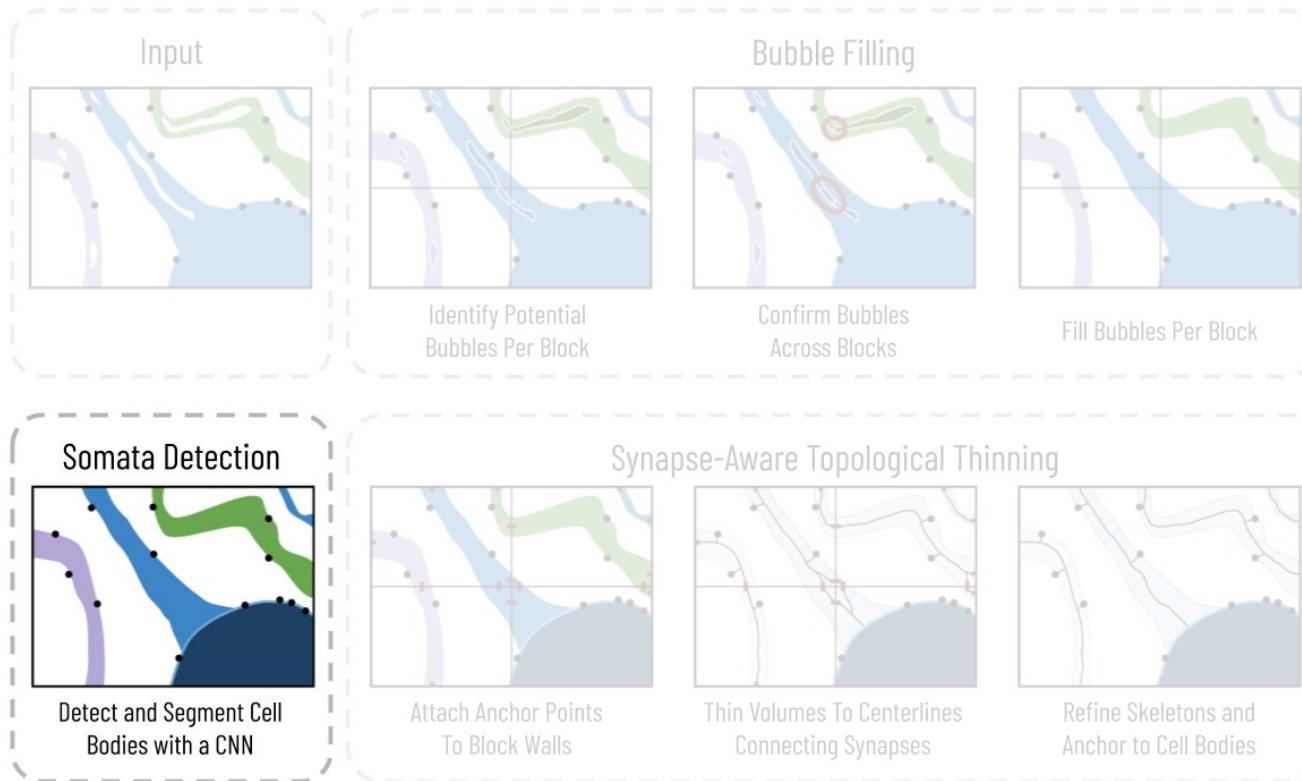
Identify Potential Bubbles Per Block

Our method identifies bubbles per block, links bubbles across blocks, and then fills in the bubbles creating dense neurons

5	5	5	8	8	8
5	5	8	8	8	-1
5	-2	8	8	8	8
5	8	8	8	8	8
-4	8	8	8	8	8
-4	-4	8	8	8	8

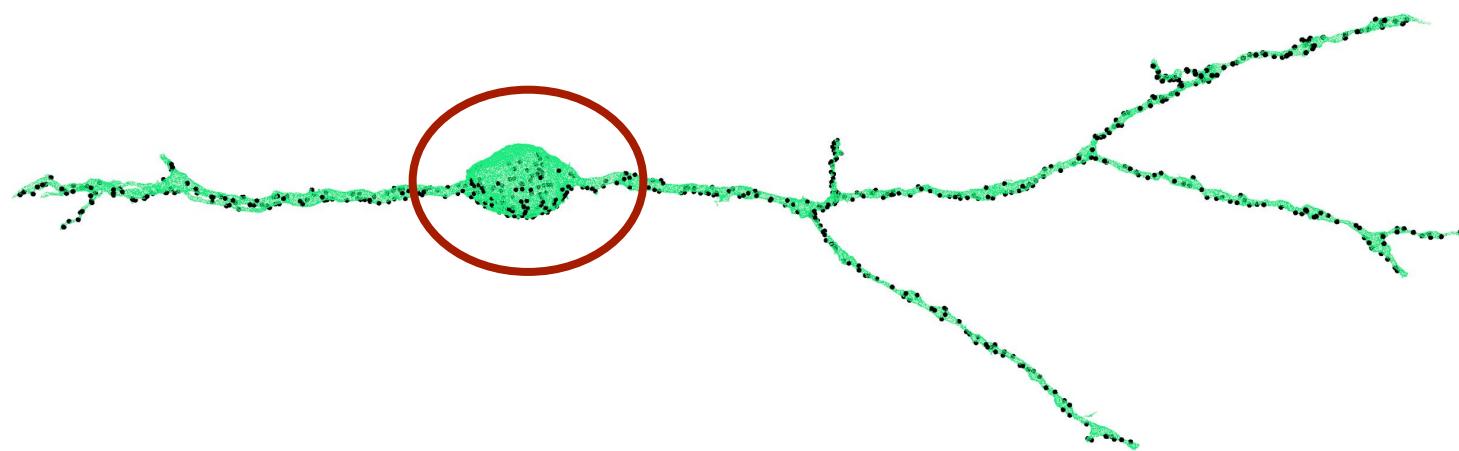
8	8	8	-37	-37	-37
-37	-37	-37	-37	-37	-37
8	8	8	8	8	-37
8	8	8	8	8	8
8	8	8	8	8	8
8	8	8	8	8	-40

Goal: Detect somata to anchor skeletons onto the cell body



Cell Bodies

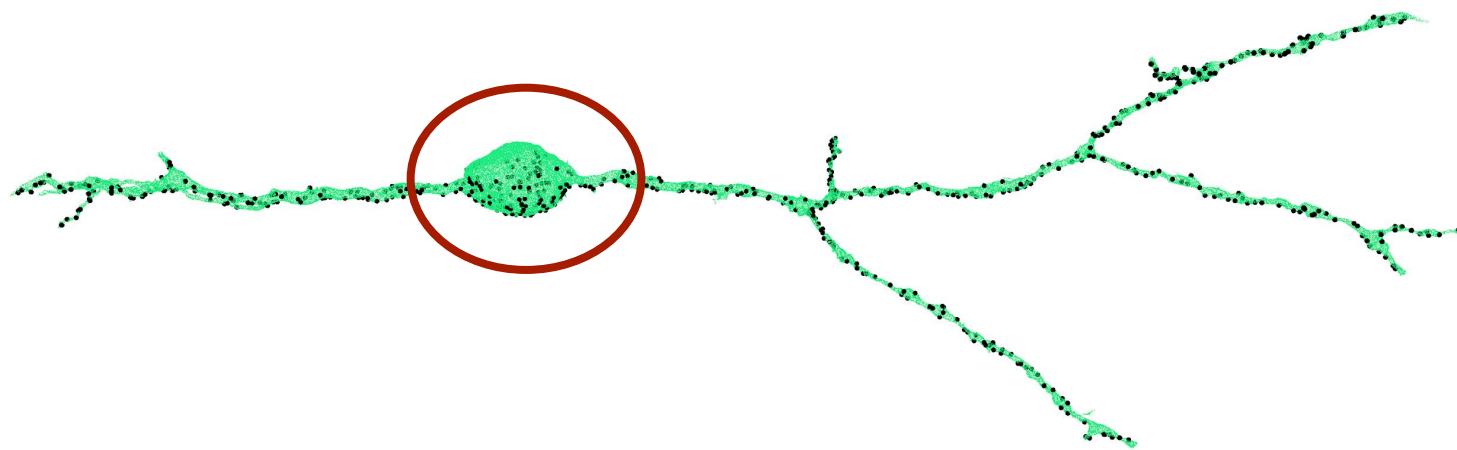
We can significantly reduce the total runtime by masking out the cell bodies before thinning



Cell Bodies

We can significantly reduce the total runtime by masking out the cell bodies before thinning

In some volumes, cell bodies represent up to 65% of the total volume

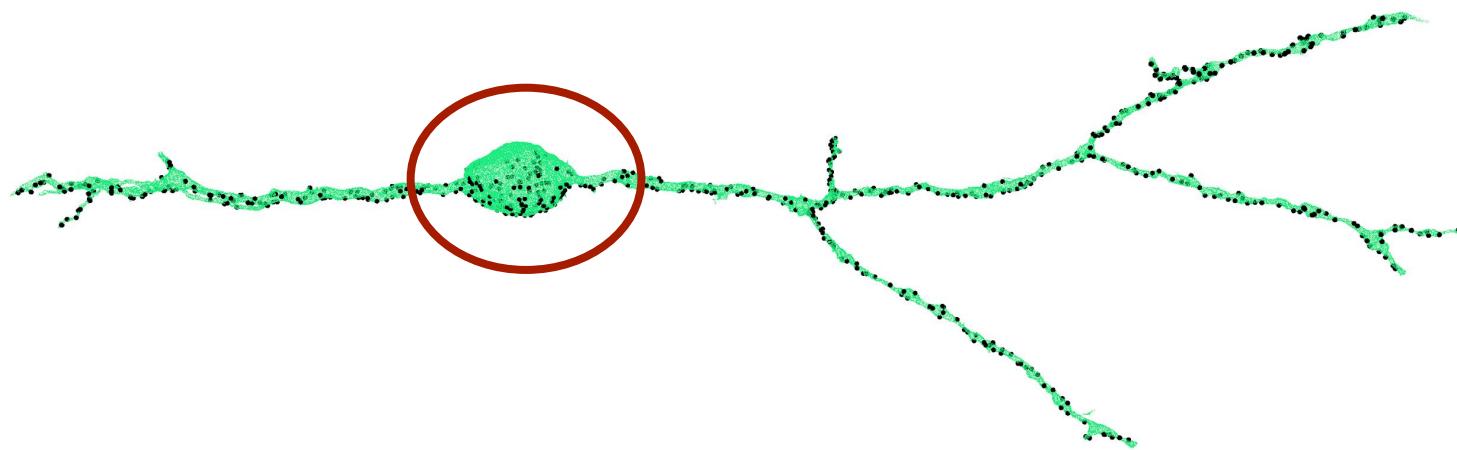


Cell Bodies

We can significantly reduce the total runtime by masking out the cell bodies before thinning

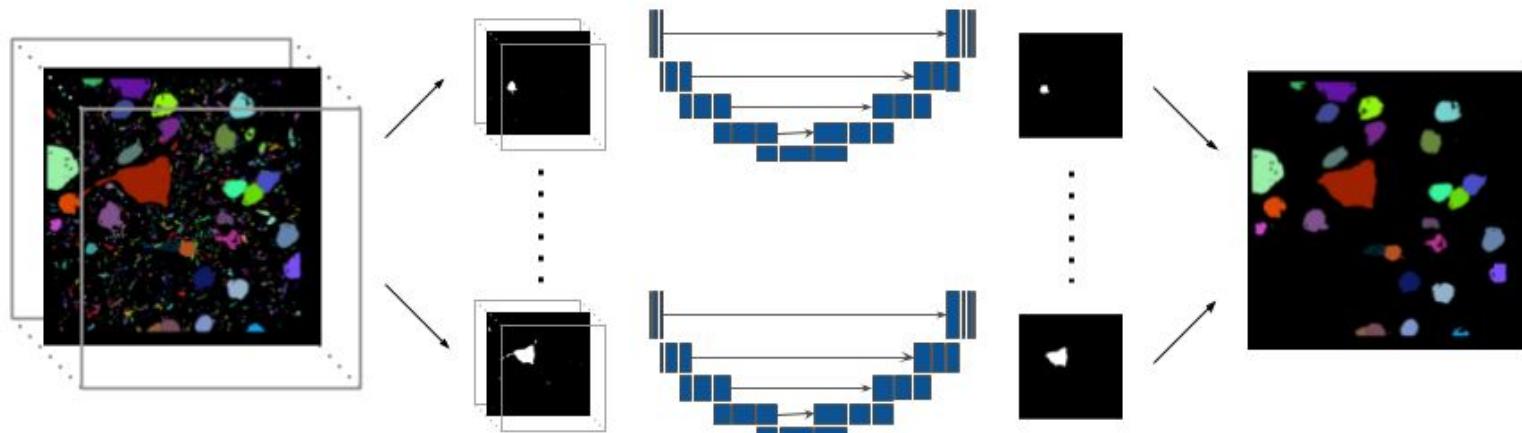
In some volumes, cell bodies represent up to 65% of the total volume

This also enables us to anchor the skeletons on the cell body



Cell Body Detection

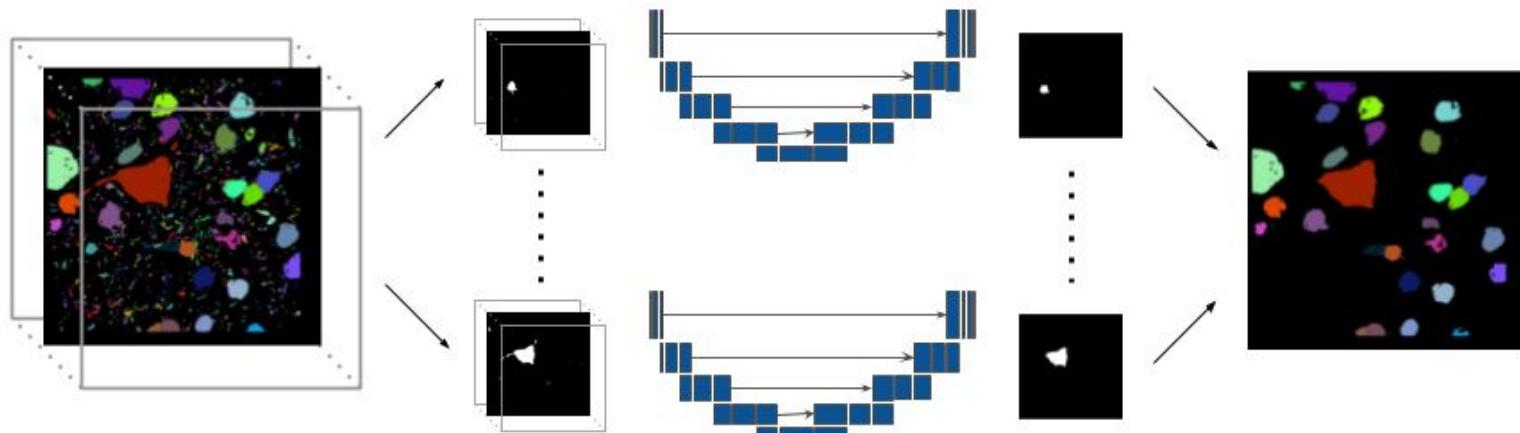
We train a U-Net to identify individual cell bodies in each slice of the label volume



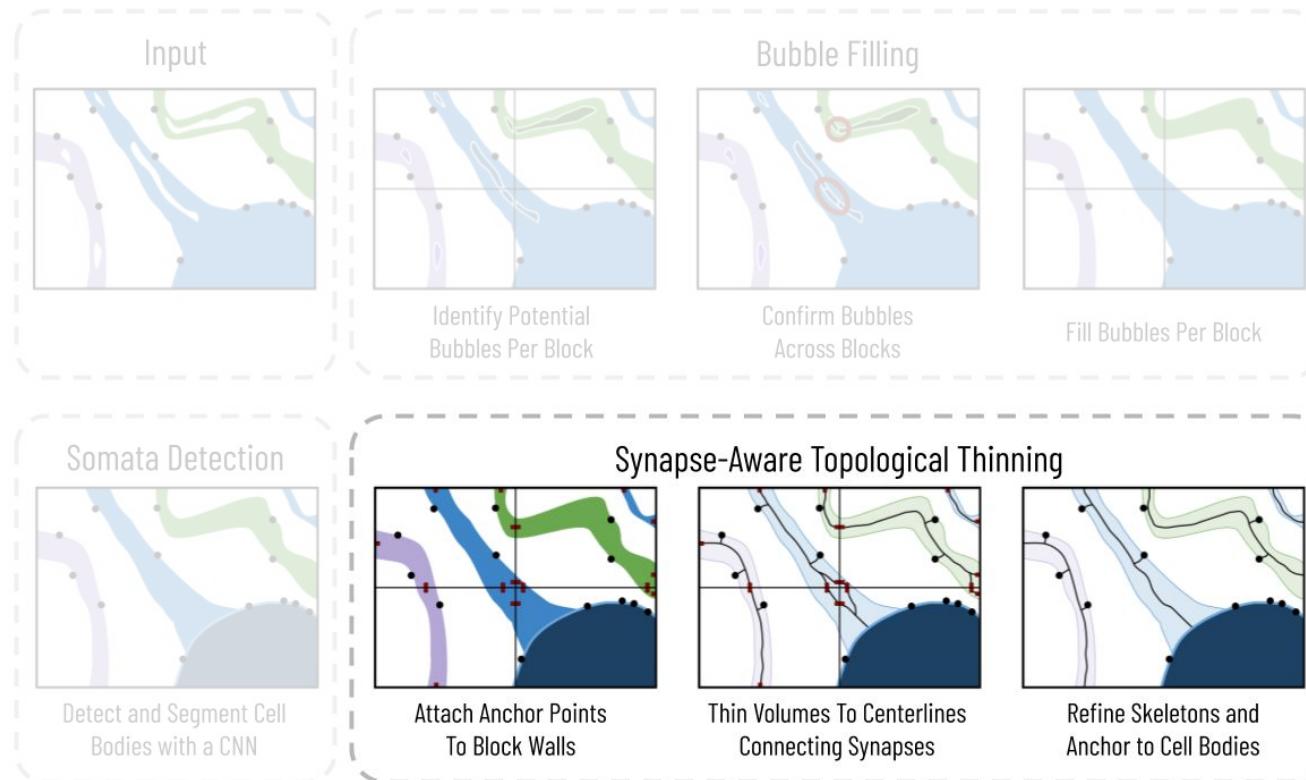
Cell Body Detection

We train a U-Net to identify individual cell bodies in each slice of the label volume

We find that U-Nets trained on one dataset transfer well to others

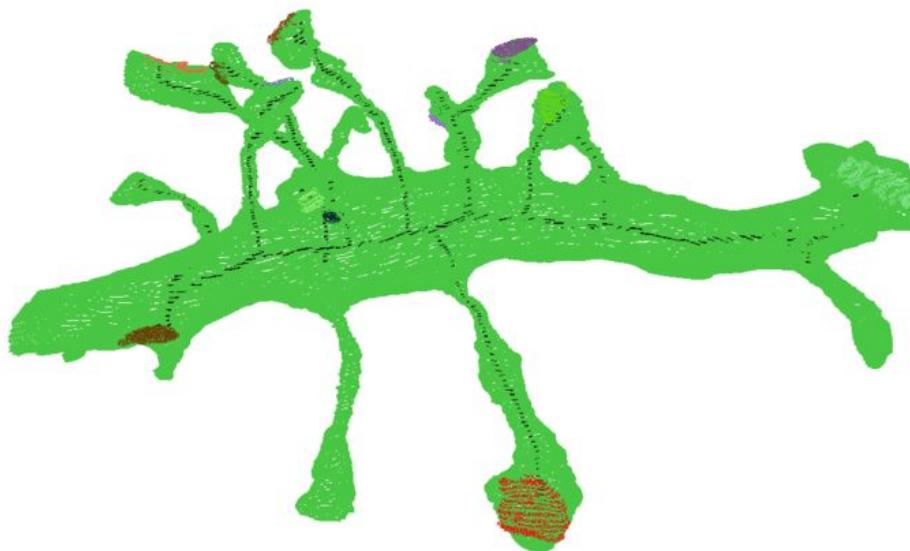


Goal: Extract accurate skeletons that connect all synapses to the cell body



Synapse-Aware Topological Thinning

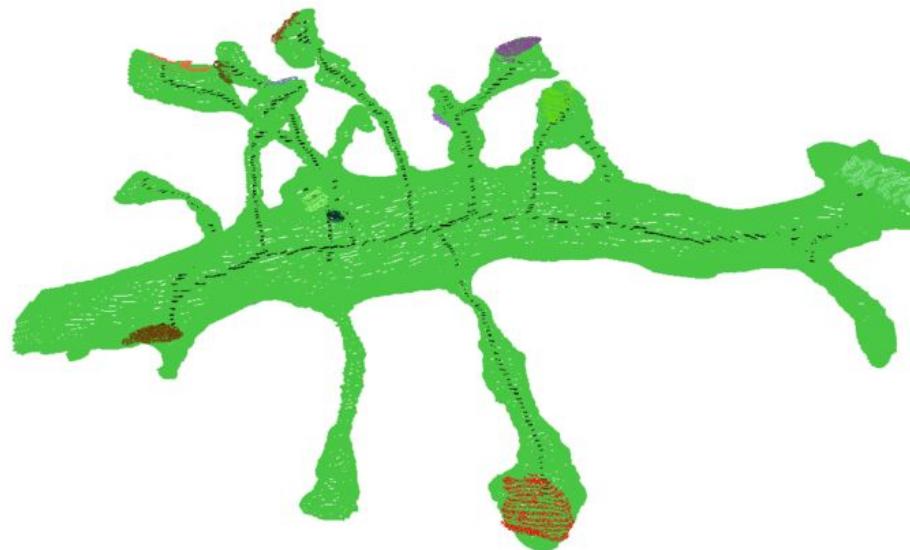
We do not allow the deletion of synapses from the skeleton



Synapse-Aware Topological Thinning

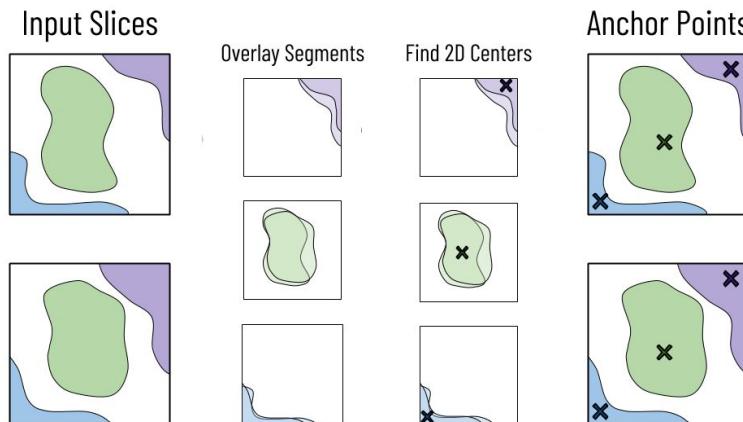
We do not allow the deletion of synapses from the skeleton

This guarantees that every synapse will remain connected in the skeletons



Attach Anchor Points to Block Walls

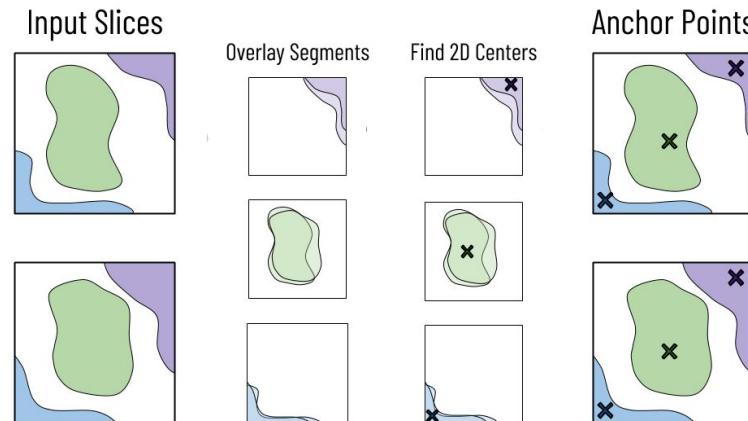
We need to guarantee that the skeletons generated in each block connect to those in neighboring blocks



Attach Anchor Points to Block Walls

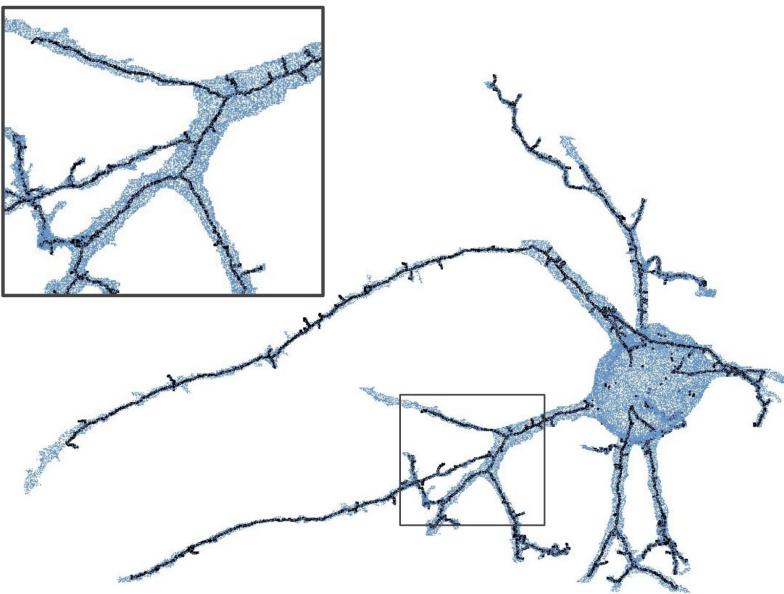
We need to guarantee that the skeletons generated in each block connect to those in neighboring blocks

For each pair of neighboring blocks, we look at the intersections of the adjacent block surfaces



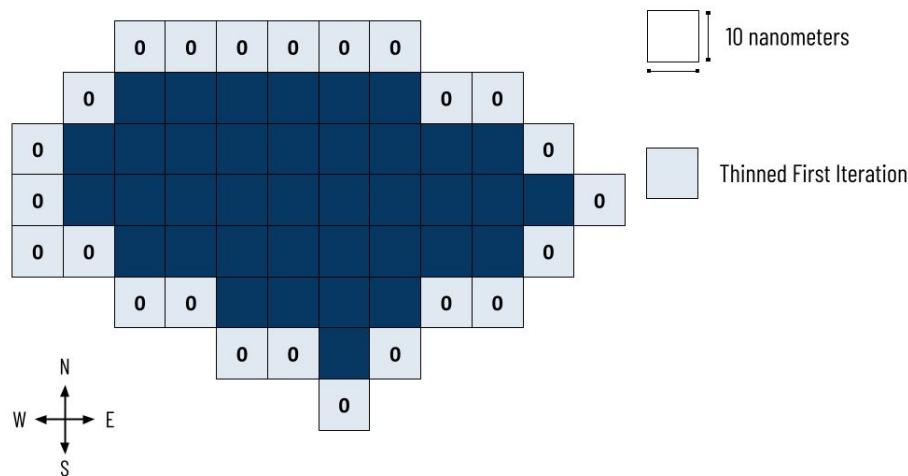
Thin Volumes to Centerlines Connecting Synapses

After determining anchor points, we can thin each segmentation per block



Estimating Neurite Width

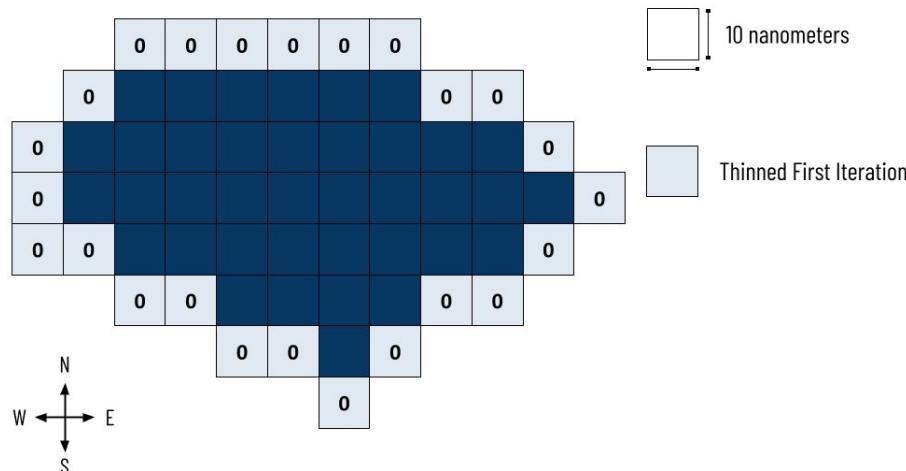
During topological thinning, we estimate the nearest distance from a thinned point to the cell boundary



Estimating Neurite Width

During topological thinning, we estimate the nearest distance from a thinned point to the cell boundary

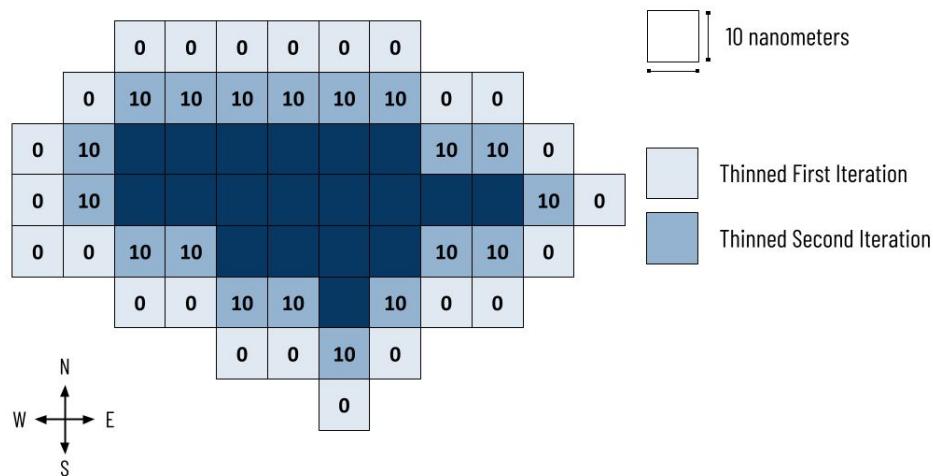
Voxels on the cell boundary are initialized at a distance of 0



Estimating Neurite Width

During topological thinning, we estimate the nearest distance from a thinned point to the cell boundary

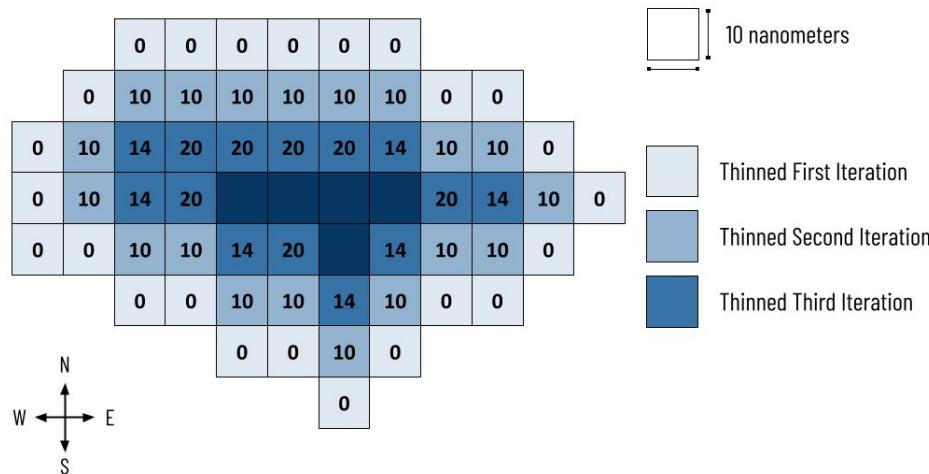
Voxels on the cell boundary are initialized at a distance of 0



Estimating Neurite Width

During topological thinning, we estimate the nearest distance from a thinned point to the cell boundary

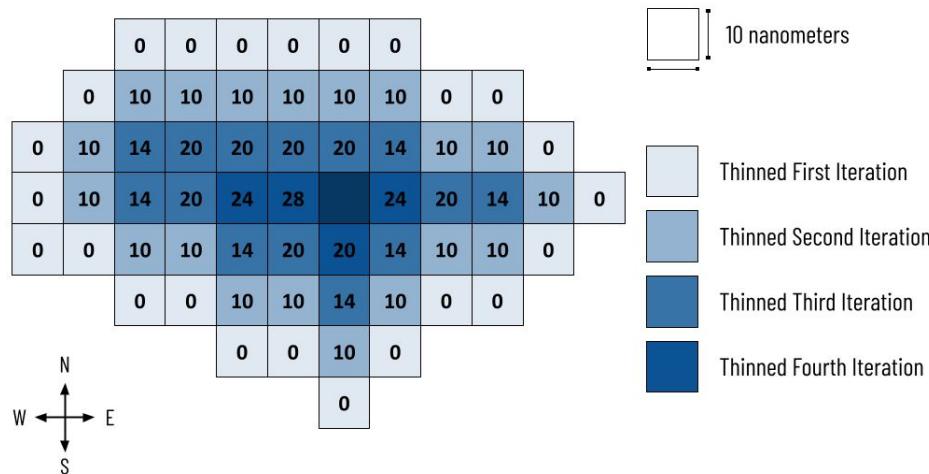
Voxels on the cell boundary are initialized at a distance of 0



Estimating Neurite Width

During topological thinning, we estimate the nearest distance from a thinned point to the cell boundary

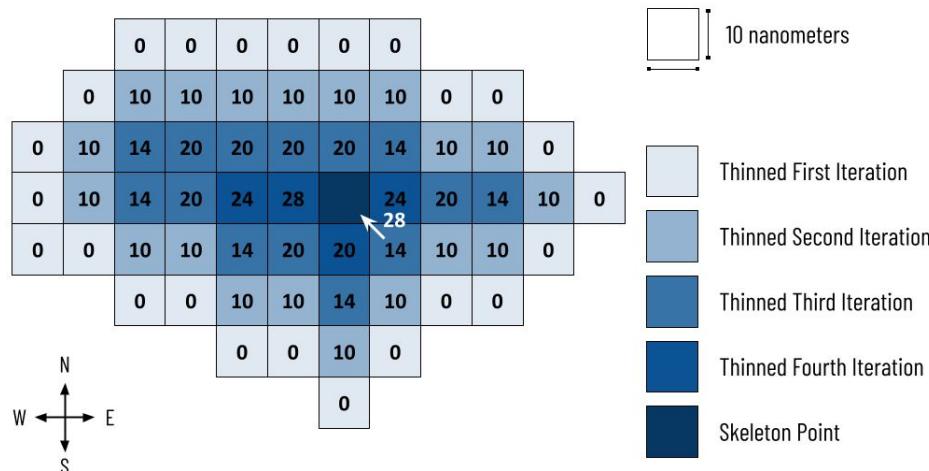
Voxels on the cell boundary are initialized at a distance of 0



Estimating Neurite Width

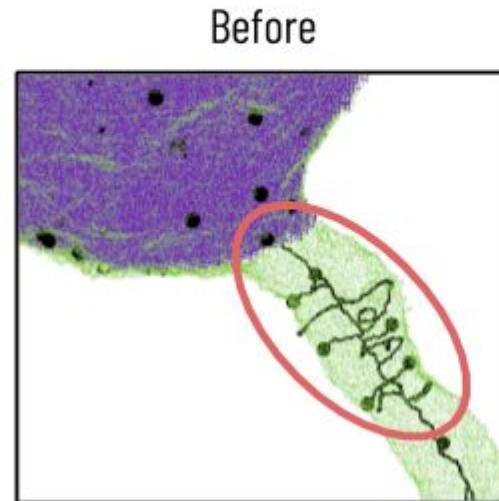
During topological thinning, we estimate the nearest distance from a thinned point to the cell boundary

Voxels on the cell boundary are initialized at a distance of 0



Refine Skeletons and Anchor to Cell Bodies

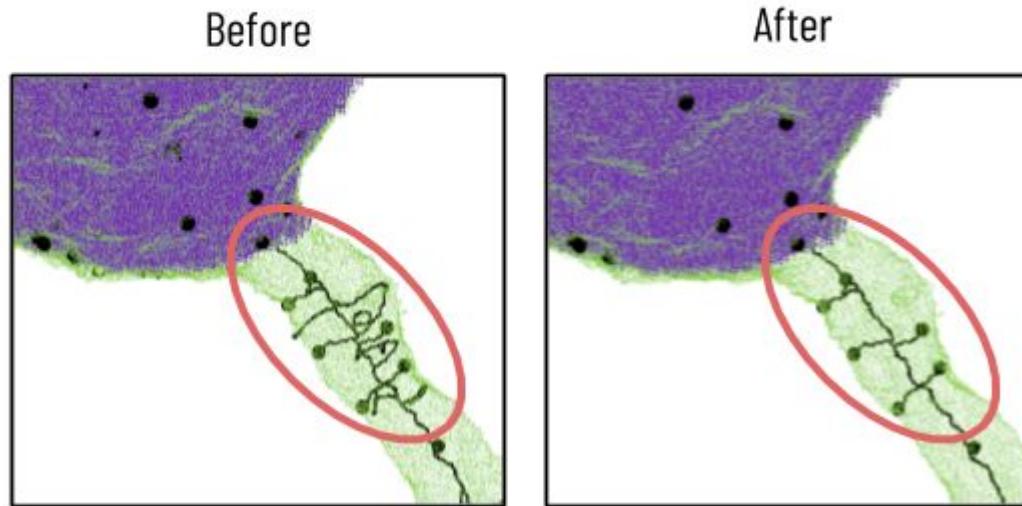
Tunnels through the label volumes can cause loops in the skeletons



Refine Skeletons and Anchor to Cell Bodies

Tunnels through the label volumes can cause loops in the skeletons

We simultaneously remove these loops and calculate the geodesic distance from each synapse to the soma

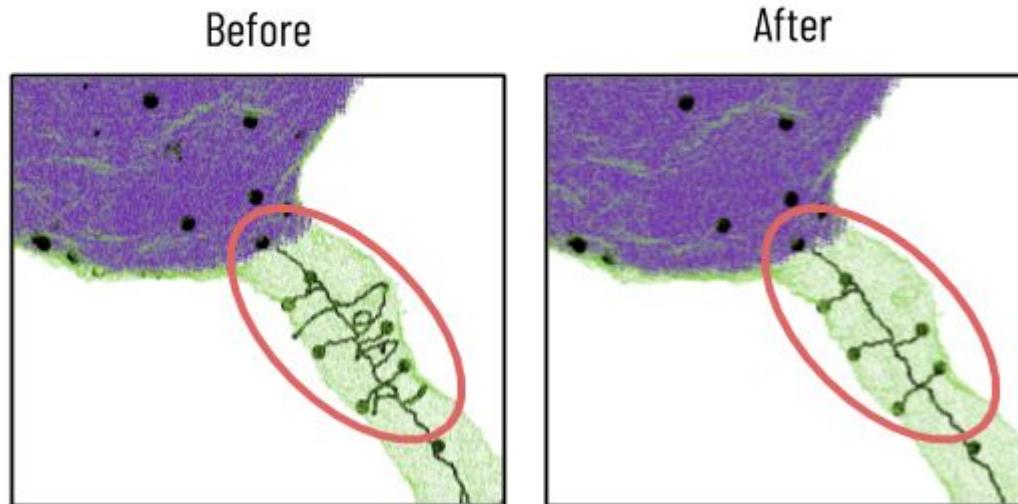


Refine Skeletons and Anchor to Cell Bodies

Tunnels through the label volumes can cause loops in the skeletons

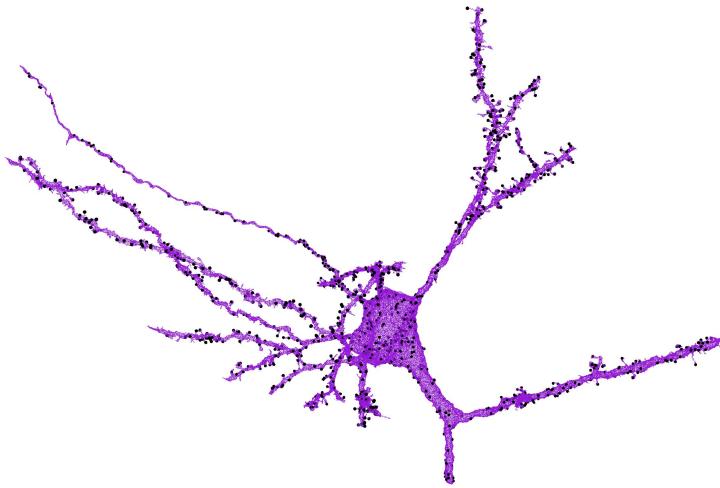
We simultaneously remove these loops and calculate the geodesic distance from each synapse to the soma

Although this step requires global scope, it is incredibly quick since it requires only the skeletons



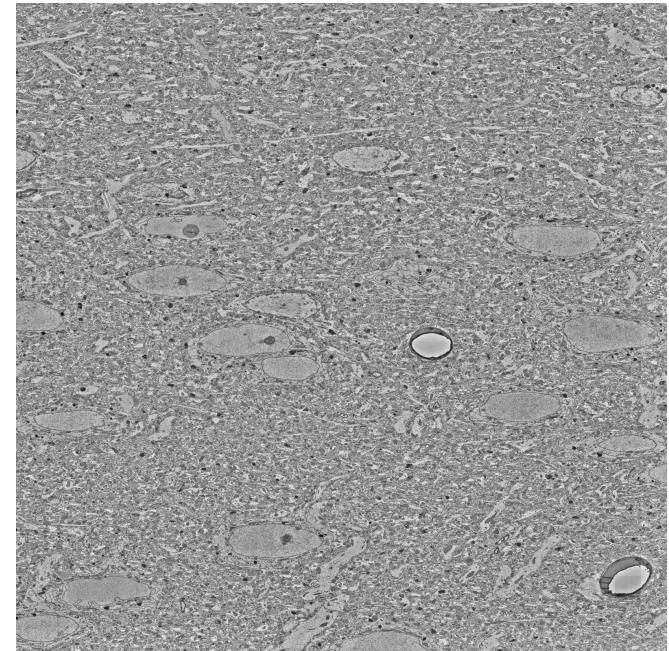
Datasets

JWR



FIB-25

J0126



$106 \times 106 \times 93 \mu\text{m}^3$

$32 \times 64 \times 30 \text{ nm}^3 / \text{vx}$

85 Neurons

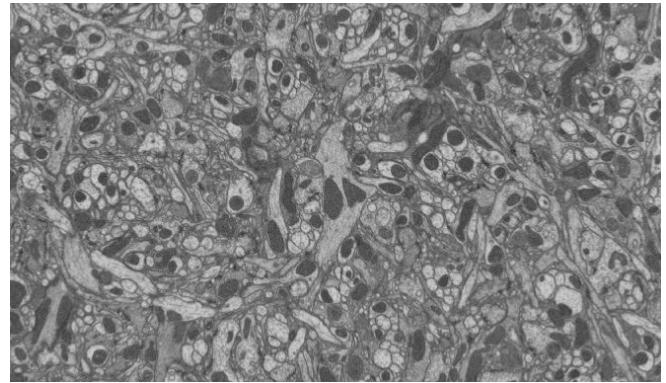
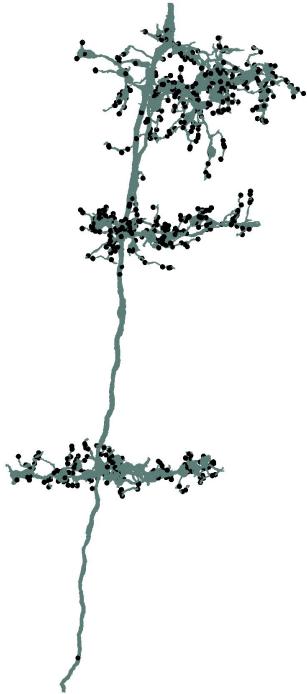
50,334 Synapses

Datasets

JWR

FIB-25

J0126



$36 \times 29 \times 69 \mu\text{m}^3$

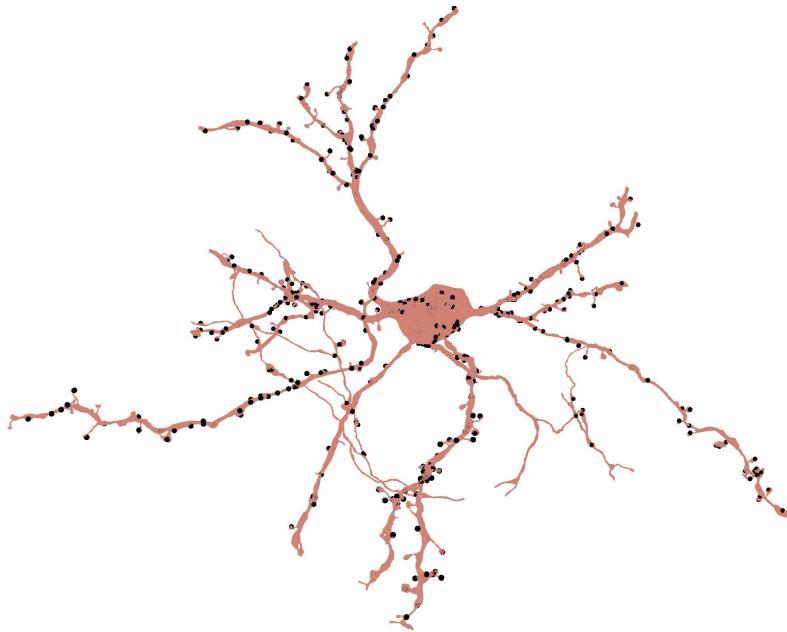
$10 \times 10 \times 10 \text{ nm}^3 / \text{vx}$

763 Neurons

84,157 Synapses

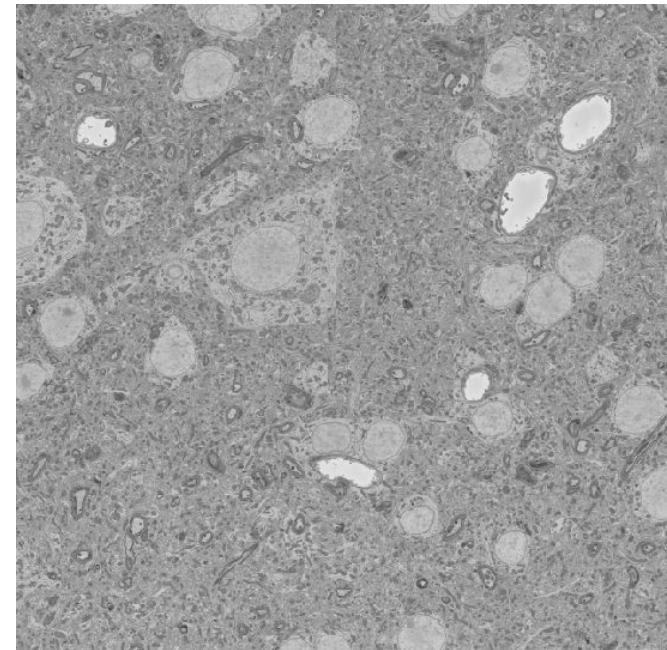
Datasets

JWR



FIB-25

J0126



$96 \times 98 \times 114 \mu\text{m}^3$

$18 \times 18 \times 20 \text{ nm}^3 / \text{vx}$

407 Neurons

91,465 Synapses

Neural Reconstruction Integrity (NRI)

The NRI score measures how well a method preserves the intracellular pathways between synapses

Neural Reconstruction Integrity (NRI)

The NRI score measures how well a method preserves the intracellular pathways between synapses

We adapt the method to measure the correspondence between synapses and endpoints

Neural Reconstruction Integrity (NRI)

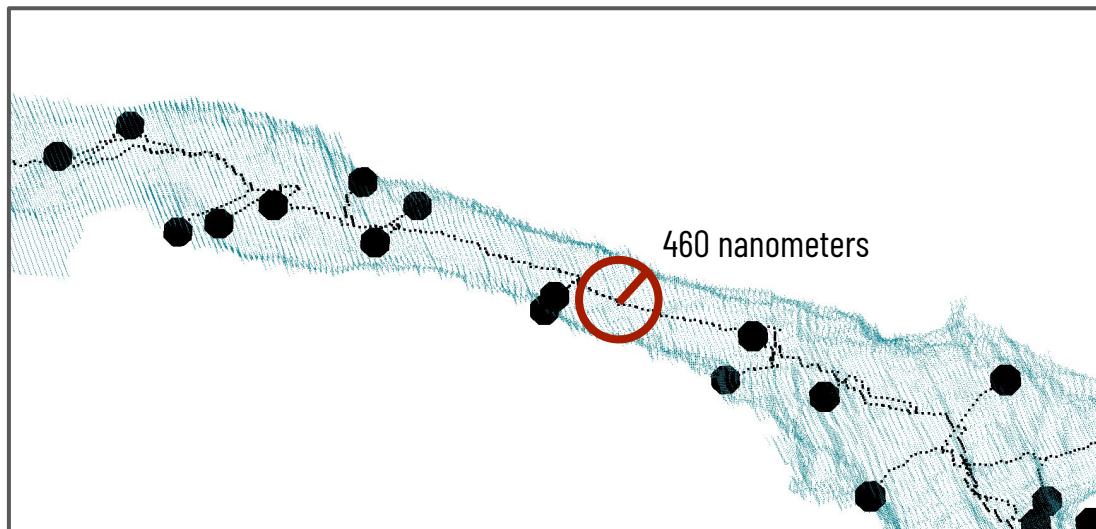
The NRI score measures how well a method preserves the intracellular pathways between synapses

We adapt the method to measure the correspondence between synapses and endpoints

For the baselines, we consider a synapse mapped to an endpoint if they are within 1600 nanometers

Width Estimation

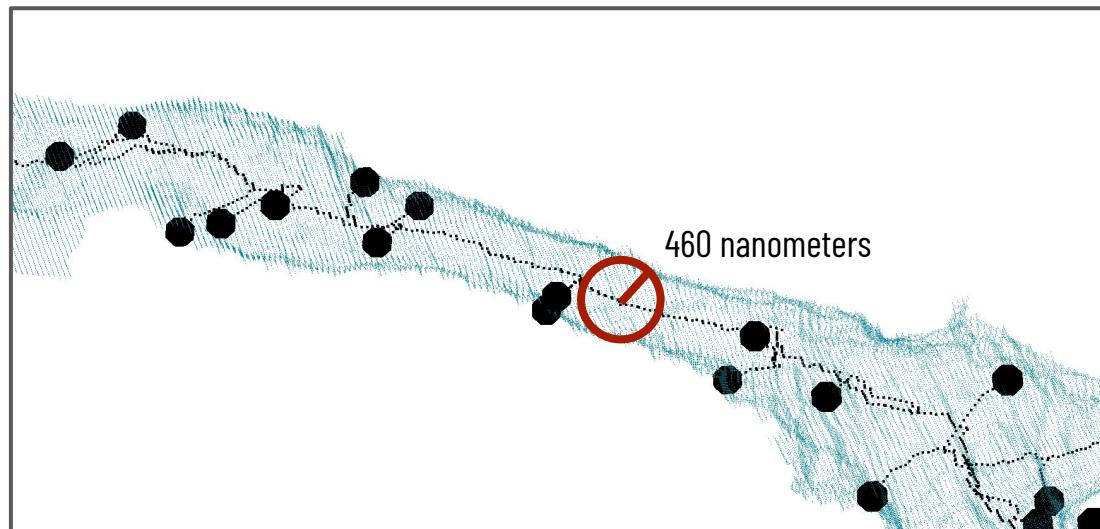
For each point along the skeleton path, we calculate the largest sphere that we could center at that point without leaving the volume



Width Estimation

For each point along the skeleton path, we calculate the largest sphere that we could center at that point without leaving the volume

We use this as an estimate for the cross-sectional width of the neurite at a given location



Skeleton Simplicity

Our final evaluation metric considers skeleton simplicity—the number of points in the skeleton

Skeleton Simplicity

Our final evaluation metric considers skeleton simplicity—the number of points in the skeleton

All else equal, fewer points is better

Quantitative Results

Method	JWR			FIB-25			J0126		
	NRI (↑)	Width (↓)	Points (↓)	NRI (↑)	Width (↓)	Points (↓)	NRI (↑)	Width (↓)	Points (↓)
Proposed	0.9988	40.03 nm	26,752	0.9952	14.42 nm	11,755	0.9997	25.55 nm	25,562
TEASER	0.1011	120.69 nm	18,250	0.2477	19.78 nm	10,216	0.1729	171.33 nm	33,022
Isthmus Thinning	0.2574	N/A	1,645,966	0.3158	N/A	39,873	0.2454	N/A	1,089,923

Quantitative Results

Method	JWR			FIB-25			J0126		
	NRI (↑)	Width (↓)	Points (↓)	NRI (↑)	Width (↓)	Points (↓)	NRI (↑)	Width (↓)	Points (↓)
Proposed	0.9988	40.03 nm	26,752	0.9952	14.42 nm	11,755	0.9997	25.55 nm	25,562
TEASER	0.1011	120.69 nm	18,250	0.2477	19.78 nm	10,216	0.1729	171.33 nm	33,022
Isthmus Thinning	0.2574	N/A	1,645,966	0.3158	N/A	39,873	0.2454	N/A	1,089,923

We improve on NRI score by 288%, 215%, and 307% over the next best method

Quantitative Results

Method	JWR			FIB-25			J0126		
	NRI (↑)	Width (↓)	Points (↓)	NRI (↑)	Width (↓)	Points (↓)	NRI (↑)	Width (↓)	Points (↓)
Proposed	0.9988	40.03 nm	26,752	0.9952	14.42 nm	11,755	0.9997	25.55 nm	25,562
TEASER	0.1011	120.69 nm	18,250	0.2477	19.78 nm	10,216	0.1729	171.33 nm	33,022
Isthmus Thinning	0.2574	N/A	1,645,966	0.3158	N/A	39,873	0.2454	N/A	1,089,923

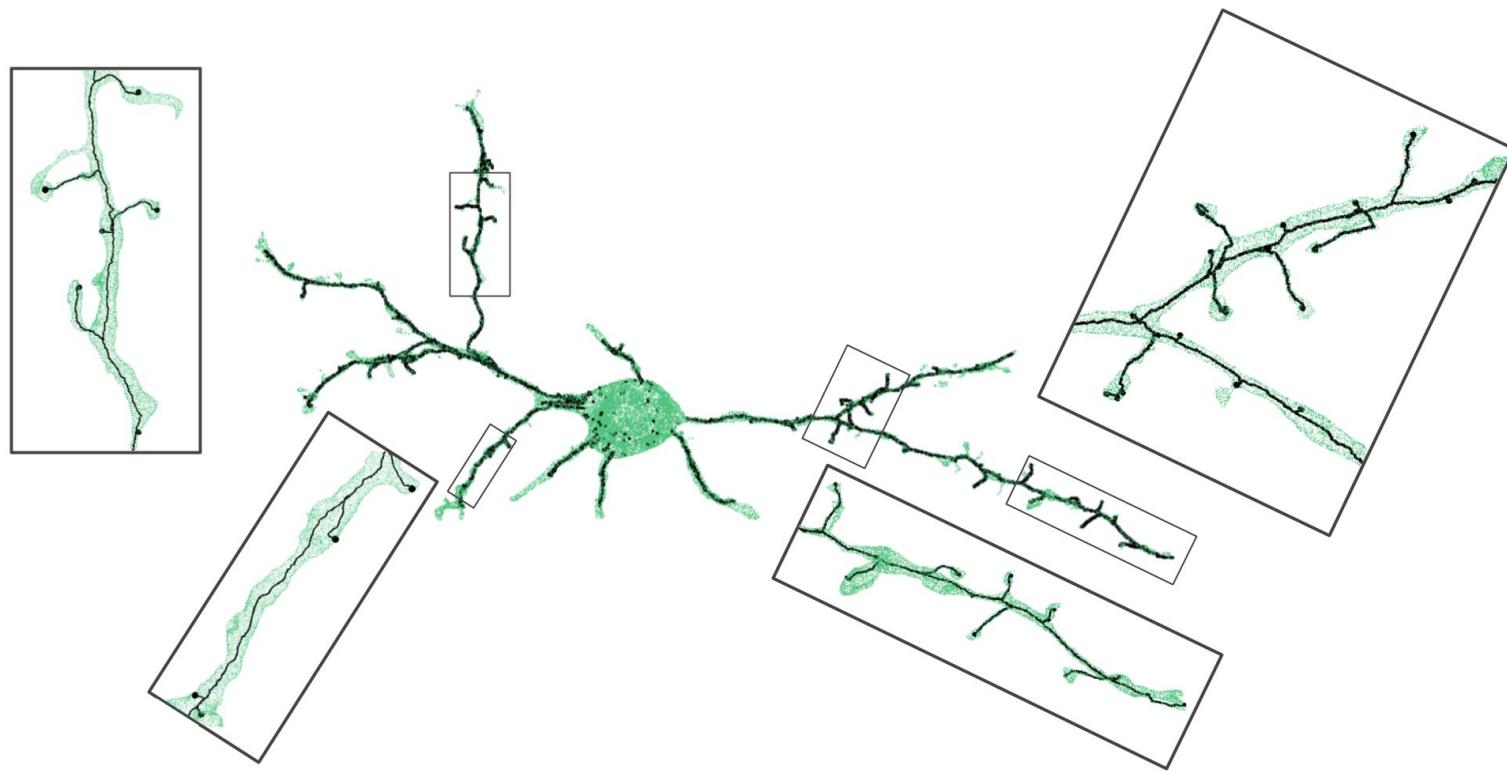
We improve on the width estimates by 67%, 27%, and 85% over the next best method

Quantitative Results

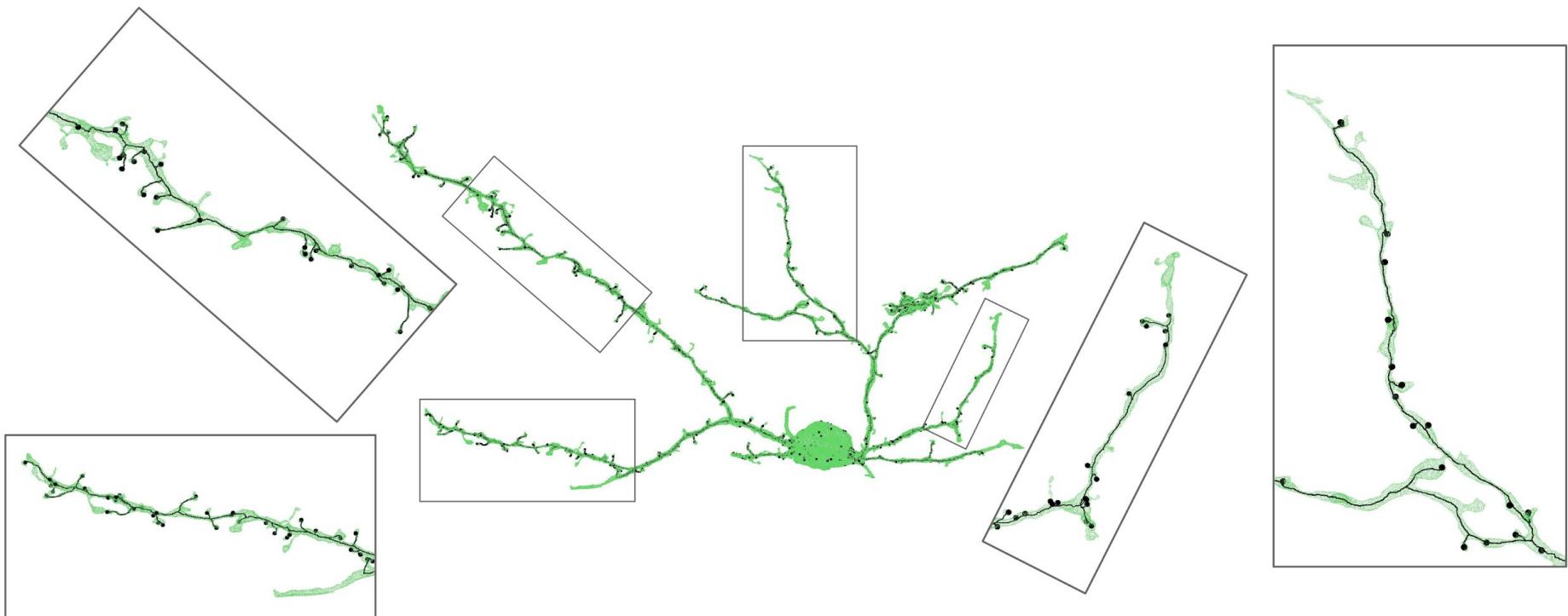
Method	JWR			FIB-25			J0126		
	NRI (↑)	Width (↓)	Points (↓)	NRI (↑)	Width (↓)	Points (↓)	NRI (↑)	Width (↓)	Points (↓)
Proposed	0.9988	40.03 nm	26,752	0.9952	14.42 nm	11,755	0.9997	25.55 nm	25,562
TEASER	0.1011	120.69 nm	18,250	0.2477	19.78 nm	10,216	0.1729	171.33 nm	33,022
Isthmus Thinning	0.2574	N/A	1,645,966	0.3158	N/A	39,873	0.2454	N/A	1,089,923

TEASER has fewer points on two of the three datasets

Qualitative Results



Qualitative Results



Ablation Studies: Bubble Filling

The J0126 reconstruction (automatic with Flood-Filling Networks), had over 24 million bubbles, corresponding to 0.80% of the total volume

Ablation Studies: Bubble Filling

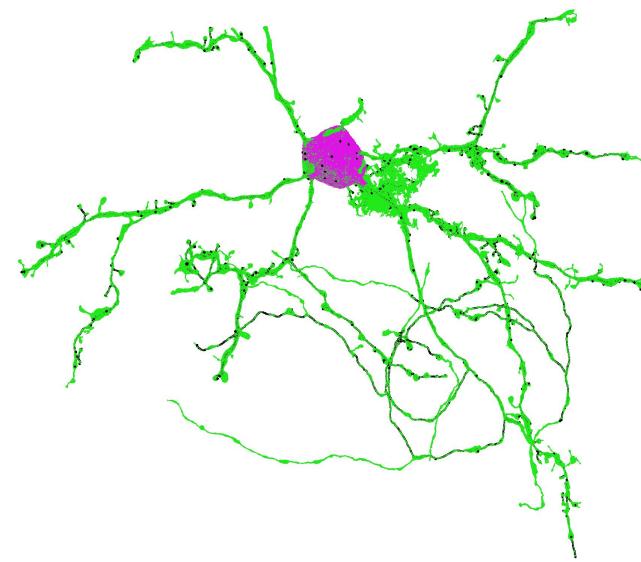
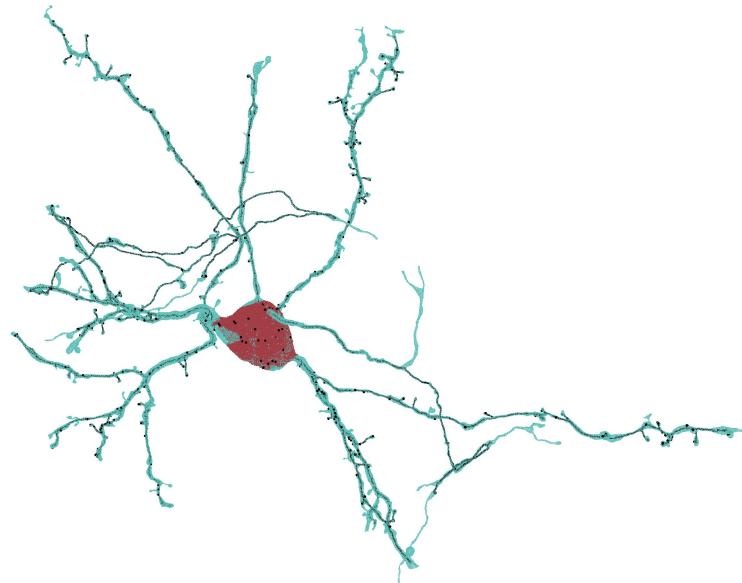
The J0126 reconstruction (automatic with Flood-Filling Networks), had over 24 million bubbles, corresponding to 0.80% of the total volume

Removing these bubbles reduced topological thinning time by 57.16%

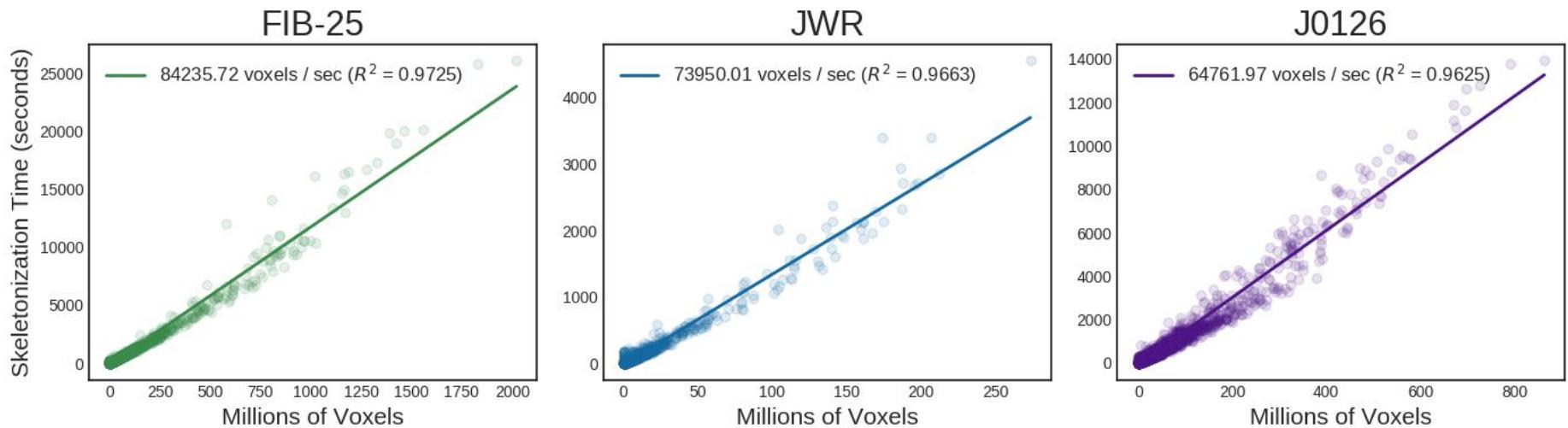
Ablation Studies: Soma Detection

We predict which voxels belong to cell bodies with 99.28% accuracy (TPR: 99.77%, FPR: 0.76%)

Masking out the cell bodies reduces the running time of topological thinning by 49.95% and 60.21% on the JWR and J0126 datasets, respectively



Ablation Studies: Synapse-Aware Topological Thinning



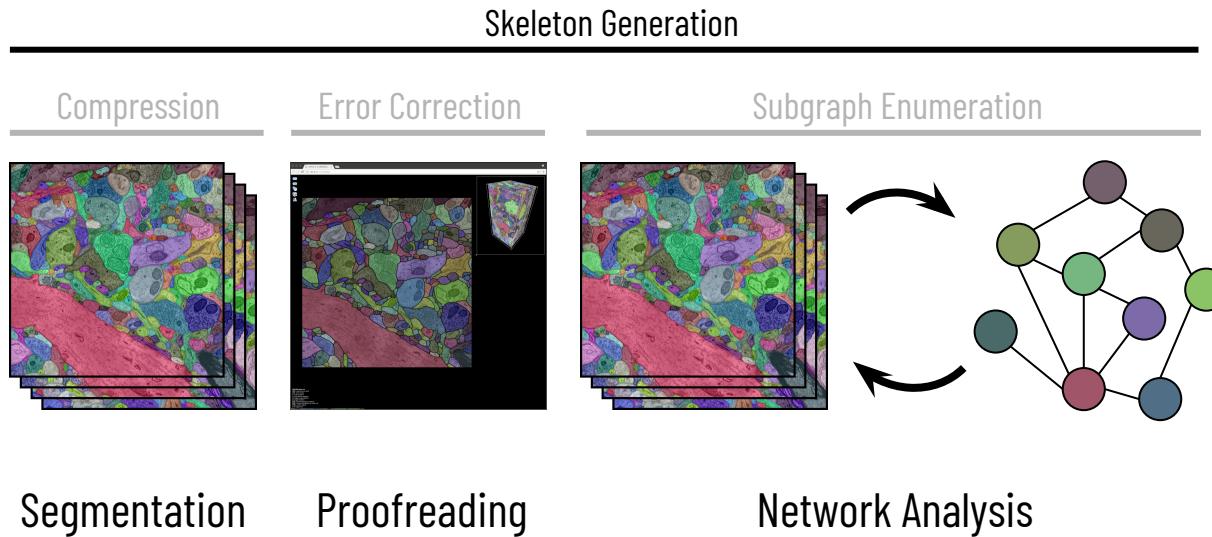
Ablation Studies: Computational Complexity

Method	JWR	FIB-25	J0126
Entire Pipeline	5.19 hr	N/A	45.29 hr
No Bubble Filling	4.03 hr	N/A	78.85 hr
No Soma Detection	10.37 hr	33.54 hr	113.82 hr
Only Thinning	20.56 hr	30.72 hr	479.19 hr

Ablation Studies: Computational Complexity

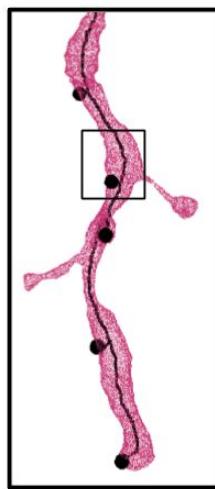
Method	JWR	FIB-25	J0126
Entire Pipeline	5.19 hr	N/A	45.29 hr
No Bubble Filling	4.03 hr	N/A	78.85 hr
No Soma Detection	10.37 hr	33.54 hr	113.82 hr
Only Thinning	20.56 hr	30.72 hr	479.19 hr

Biologically-Aware Algorithms for Connectomics

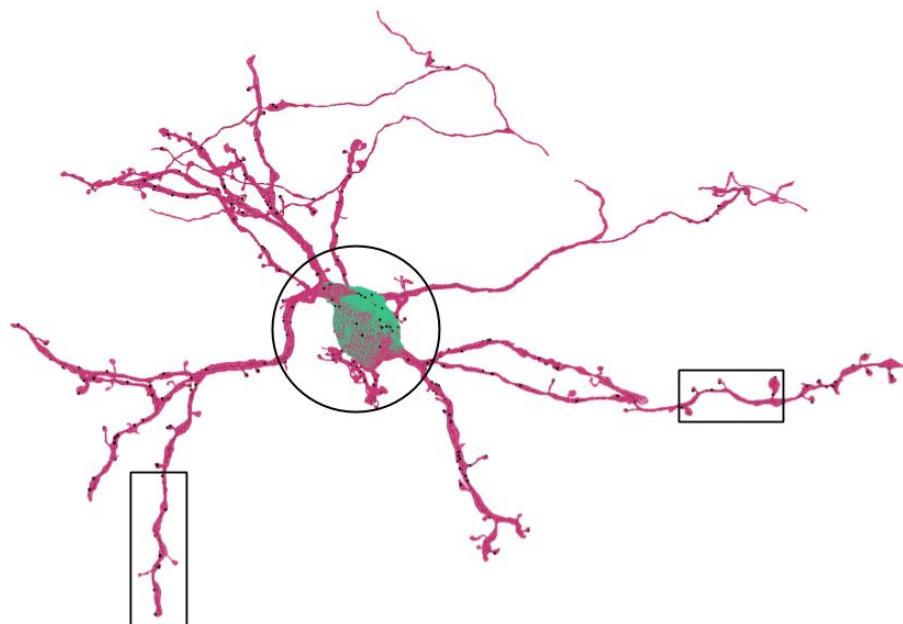
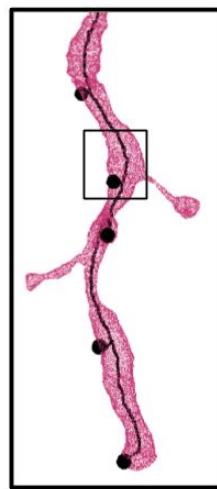




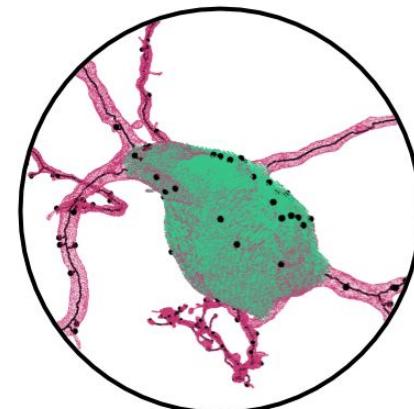
Bubble Filling



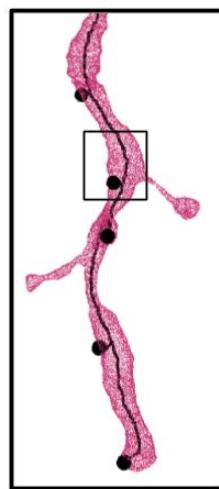
Bubble Filling



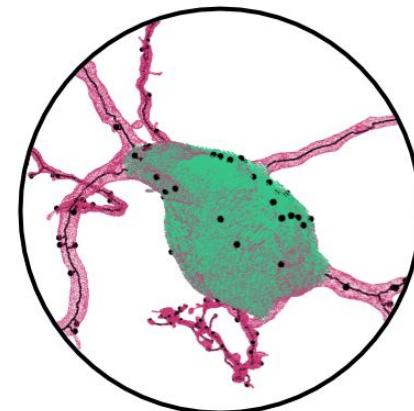
Soma Detection



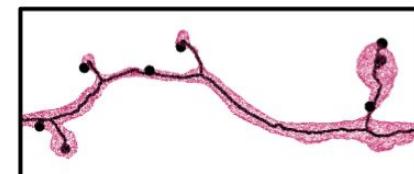
Bubble Filling



Soma Detection



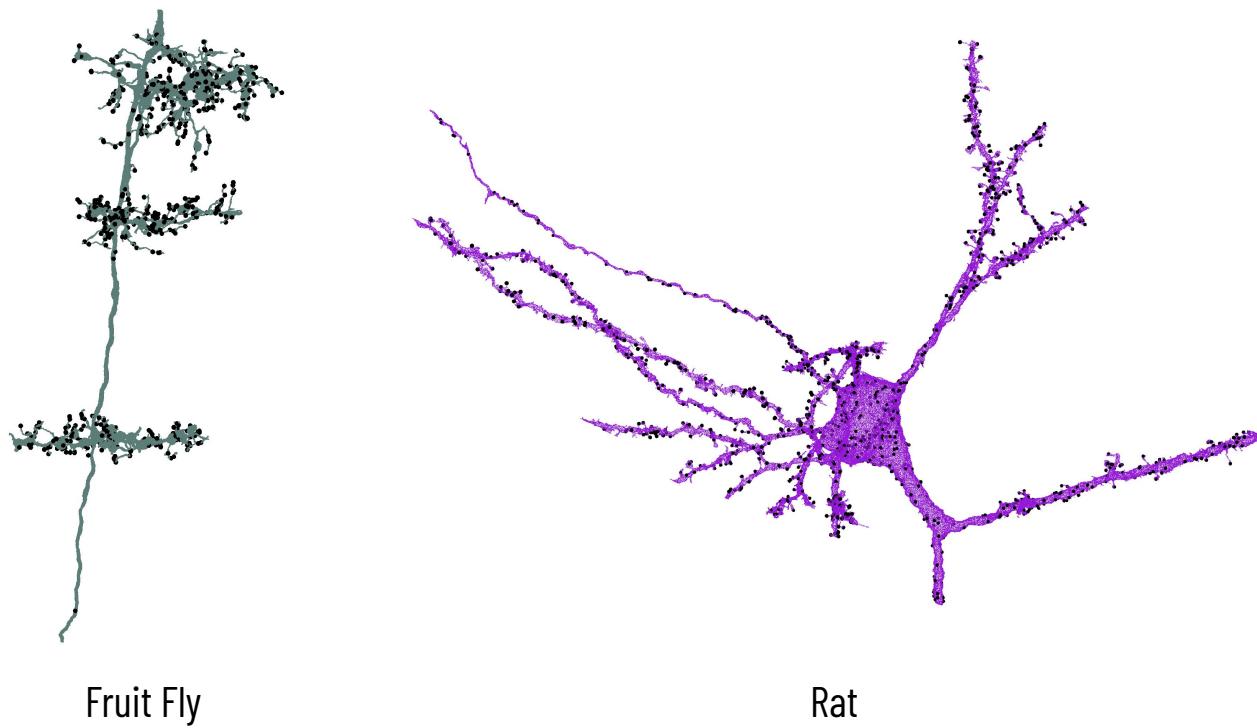
Synapse Connectivity



Future Directions

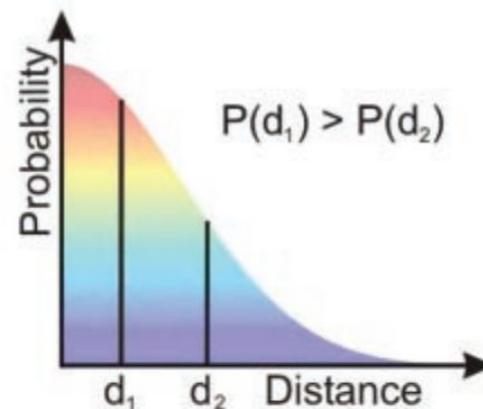
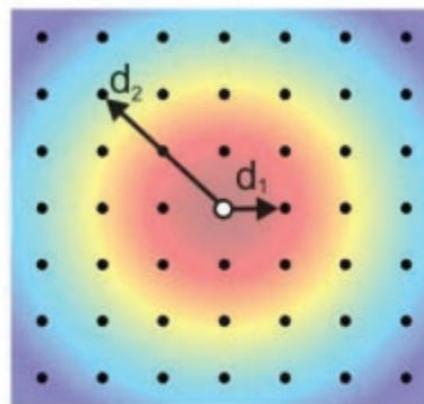
Different Biological-Constraints for Error Correction

Biological-constraints should differ between species as morphologies can look wildly different



Random Graph Generation for Motif Discovery

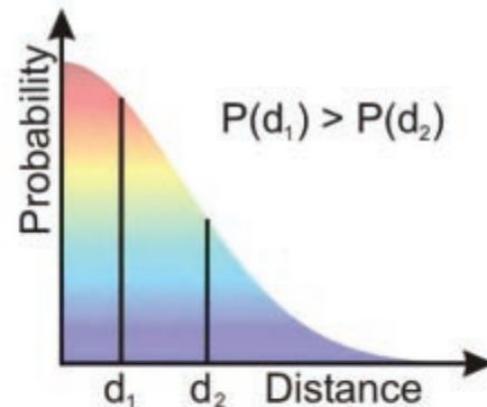
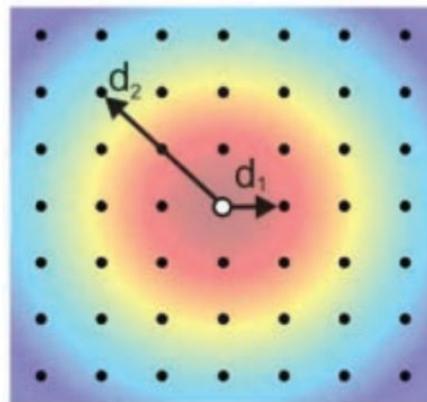
Random graphs allow us to identify subgraphs that appear more frequently than expected



Random Graph Generation for Motif Discovery

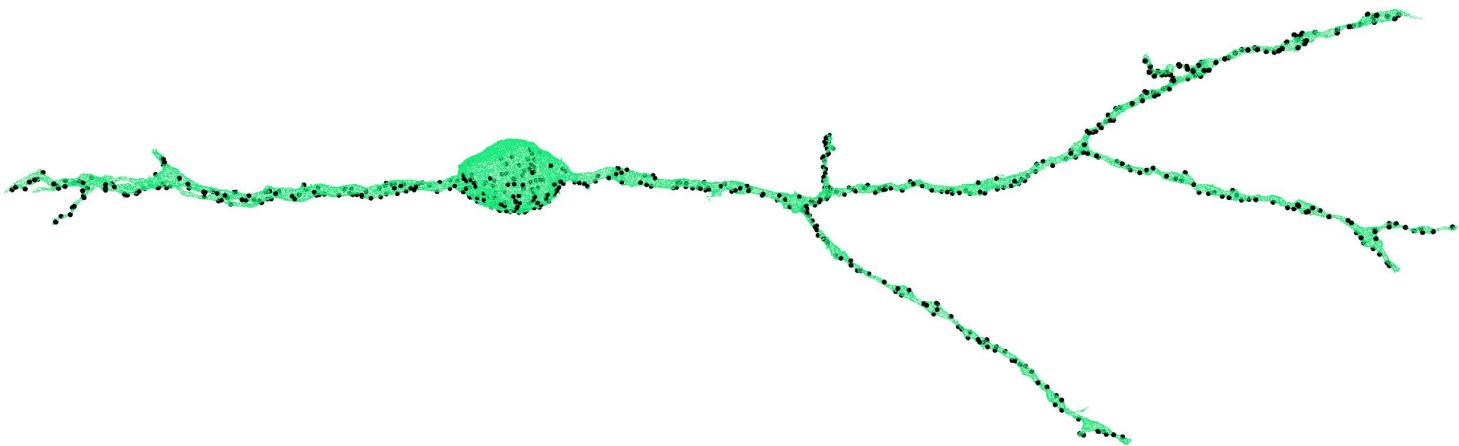
Random graphs allow us to identify subgraphs that appear more frequently than expected

Without using biological priors on our random graphs, we can mistakenly identify “important motifs”



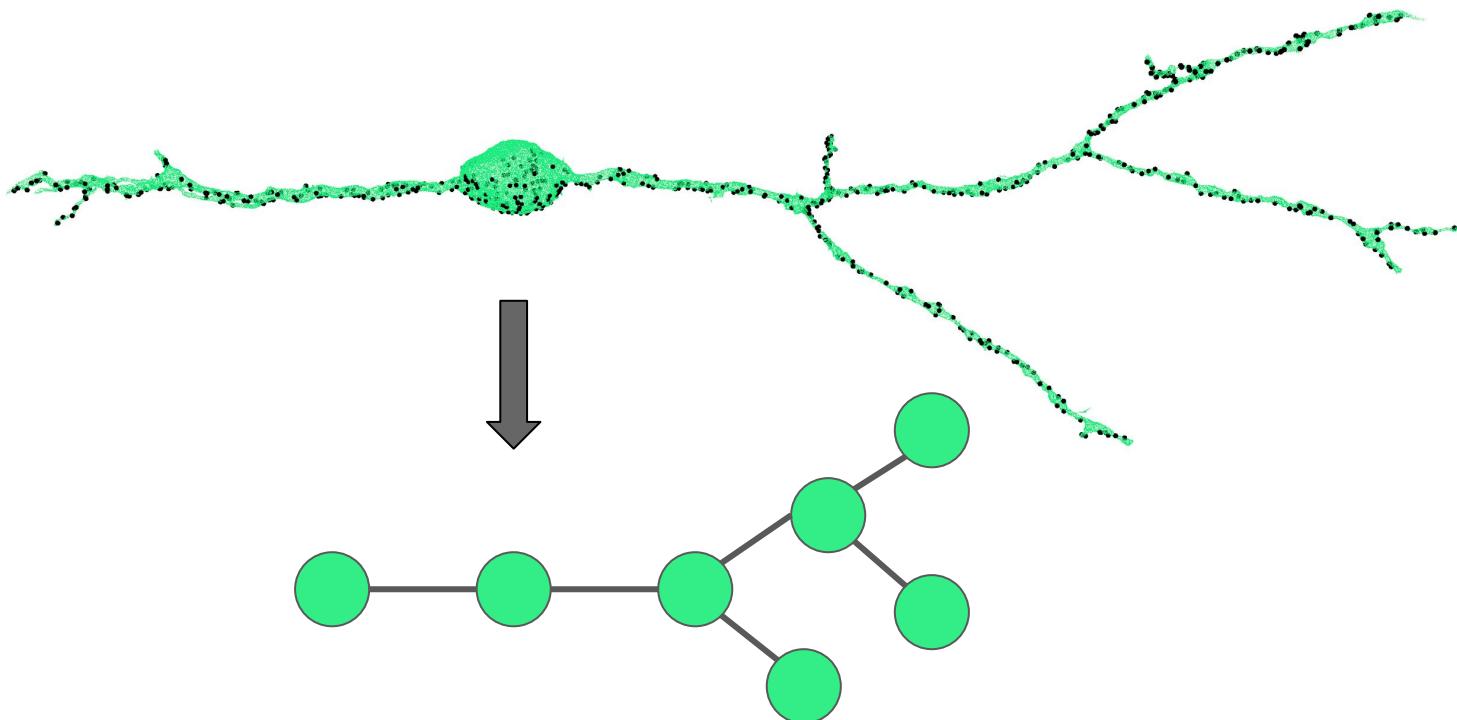
Augmenting the Wiring Diagram with Skeletonization

Current wiring diagrams ignore any interplay between neurites



Augmenting the Wiring Diagram with Skeletonization

Neurites themselves can each become nodes in a wiring diagram where neurons become multiple nodes



Acknowledgements



Hanspeter Pfister

Acknowledgements



Todd Zickler



Michael Mitzenmacher

Acknowledgements



Tianyi Chen



Tim
Franzmeyer



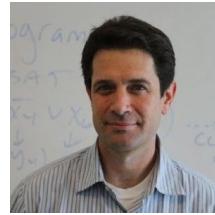
Daniel Haehn



Fritz Lekschas



Jeff W.
Lichtman



Michael
Mitzenmacher



Kálmán Palágyi



Toufiq Parag



Hanspeter
Pfister



Babis
Tsourakakis



Snow Wang



Donglai Wei



Jinglin Zhao



Haidong Zhu

Visual Computing Group



Thank you!

Questions?