

Mitochondria Classifier by Pyramidal Neuron Compartment

using machine learning

long mito in dendrites

medium mito in soma
and high density

small short mito in axons

20 μ m

nm³ x 82000, y 55568, z 545

Share

x 2cell_segmentation_v185

x 3synapses

x 4mitochondria

x 5nuclei

+

cell_segmentation_v185

Rendering Annotations Trans

▼ 2D Visualization

Opacity (on)

Opacity (off)

Saturation

Resolution (slice)

▼ 3D Visualization

Resolution (mesh)

Opacity (3d)

Load layer meshes (requires refr)

Load layer skeletons (requires re

▼ Segment Selection

Ignore segment interactions

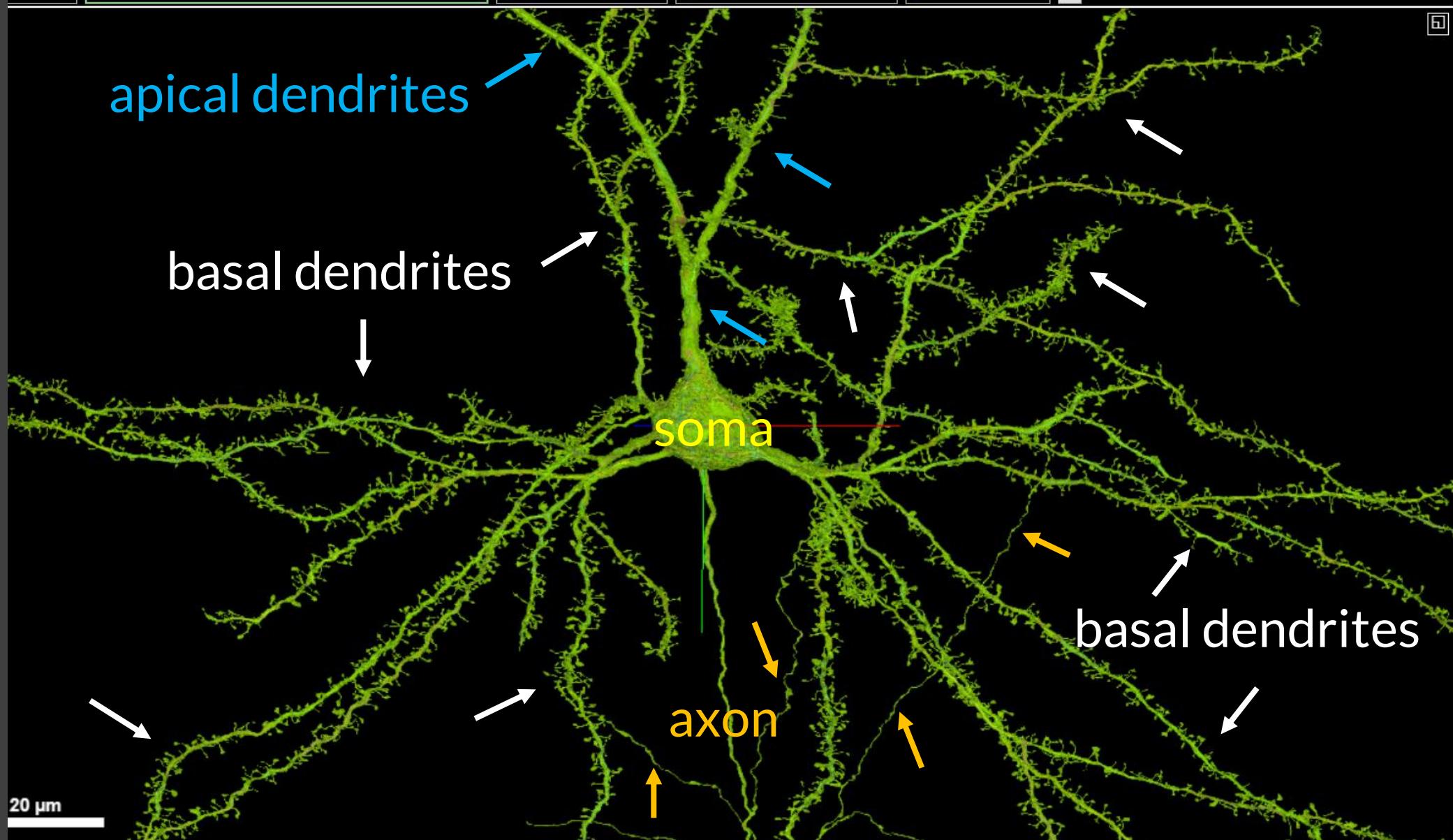
Hide segment ID 0

+ []

x [] [] [] []

648518346349538053 []

▼ Omni Segment Info



Metrics from
Turner et al.
Cell 2022

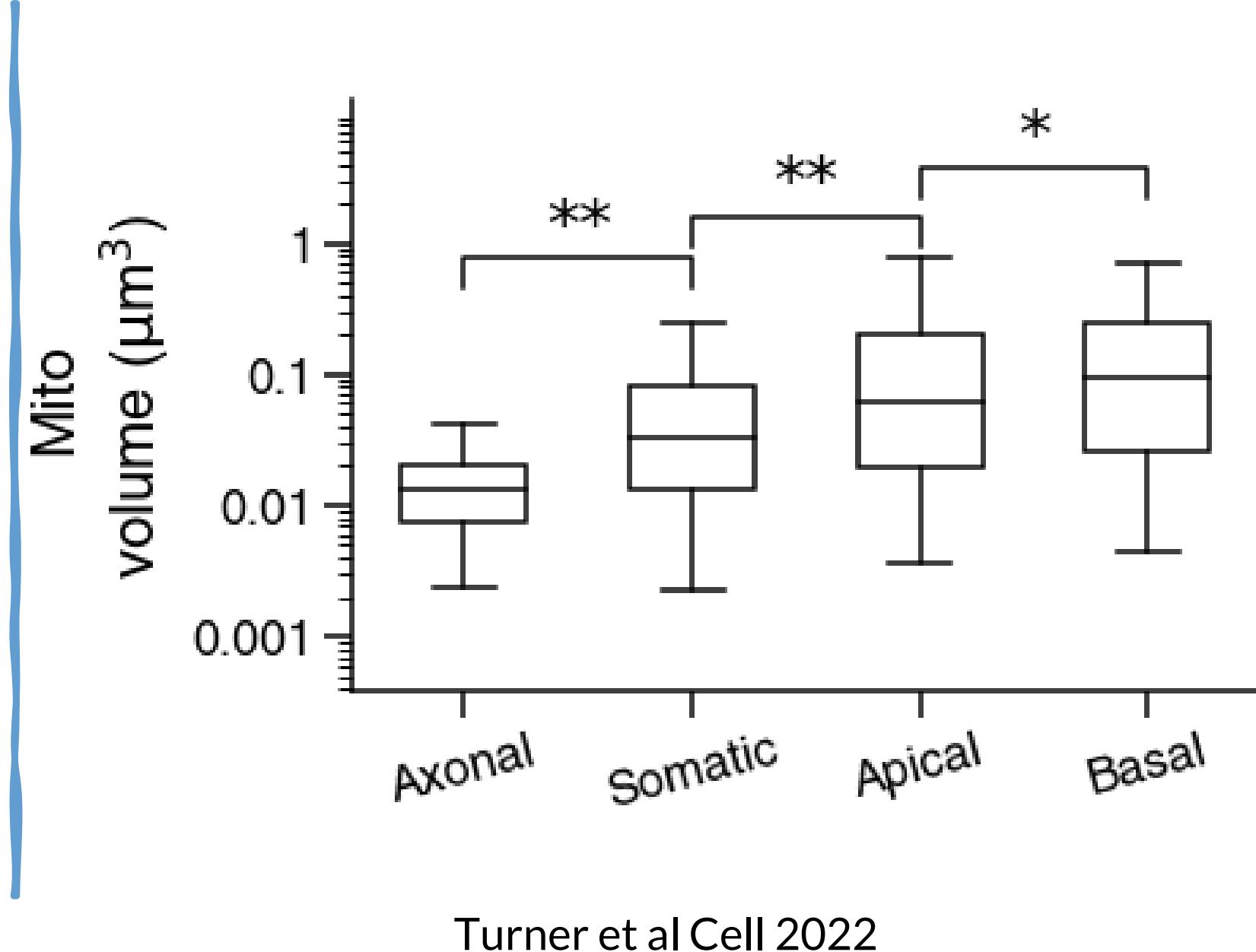
Mito voxels

Mito surface area

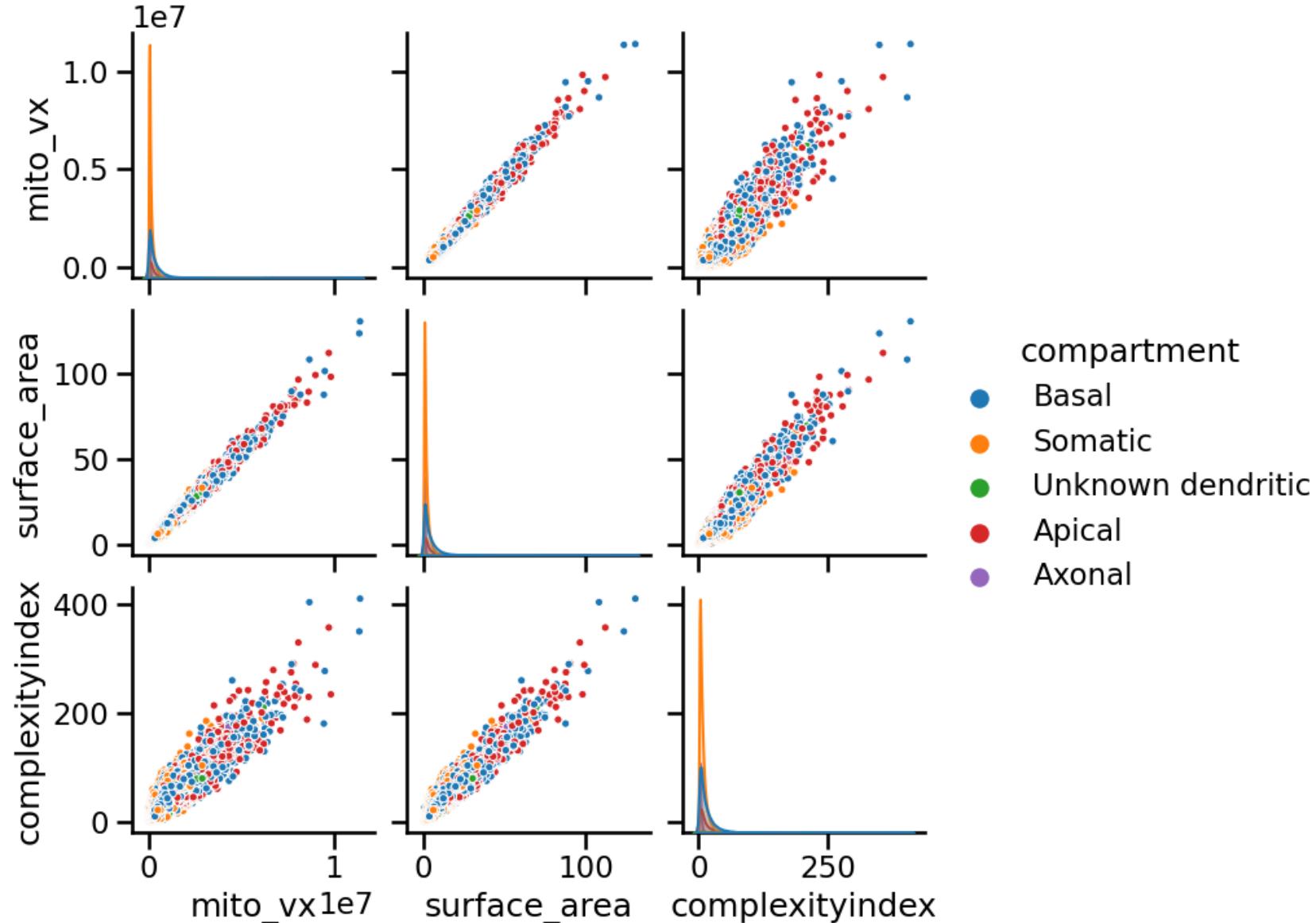
Mito complexity index

Algorithm used mesh skeleton of cell
membrane to assign compartments

Mito volume correlates with compartment



Most metrics are proxies for mito_vx



Will the reverse work??

(mitochondria characteristics used to assign compartment)

The Turner algorithm used cell skeleton characteristics to determine compartments and then assigned each mitochondria accordingly (compartment → mito)

Can characteristics of the mitochondria be used to determine compartment? (mito → compartment)

Use the existing compartment assignments by Turner as a starting point (even though there are assignment errors present in the dataframe)

Train a classifier to analyze mitochondrial metrics and make a compartment assignment accordingly

Metrics calculated from mito df

Distance vector from soma centroid to mito centroid

Boolean if mito centroid is within X voxels (e.g., 1500) of soma centroid

Mito bounding box distance vector (bbox beginning – bbox end)

Boolean if mito centroid is above soma centroid (1) or below soma (0)

Boolean if mito centroid is within a 3d cone of 45° above the soma (toward pia)

nm³ x 82000, y 55568, z 545

Share

x 2cell_segmentation_v185

x 3synapses

x 4mitochondria

x 5nuclei

+

cell_segmentation_v185

Rendering Annotations Trans

▼ 2D Visualization

Opacity (on)

Opacity (off)

Saturation

Resolution (slice)

▼ 3D Visualization

Resolution (mesh)

Opacity (3d)

Load layer meshes (requires refr)

Load layer skeletons (requires re)

▼ Segment Selection

Ignore segment interactions

Hide segment ID 0

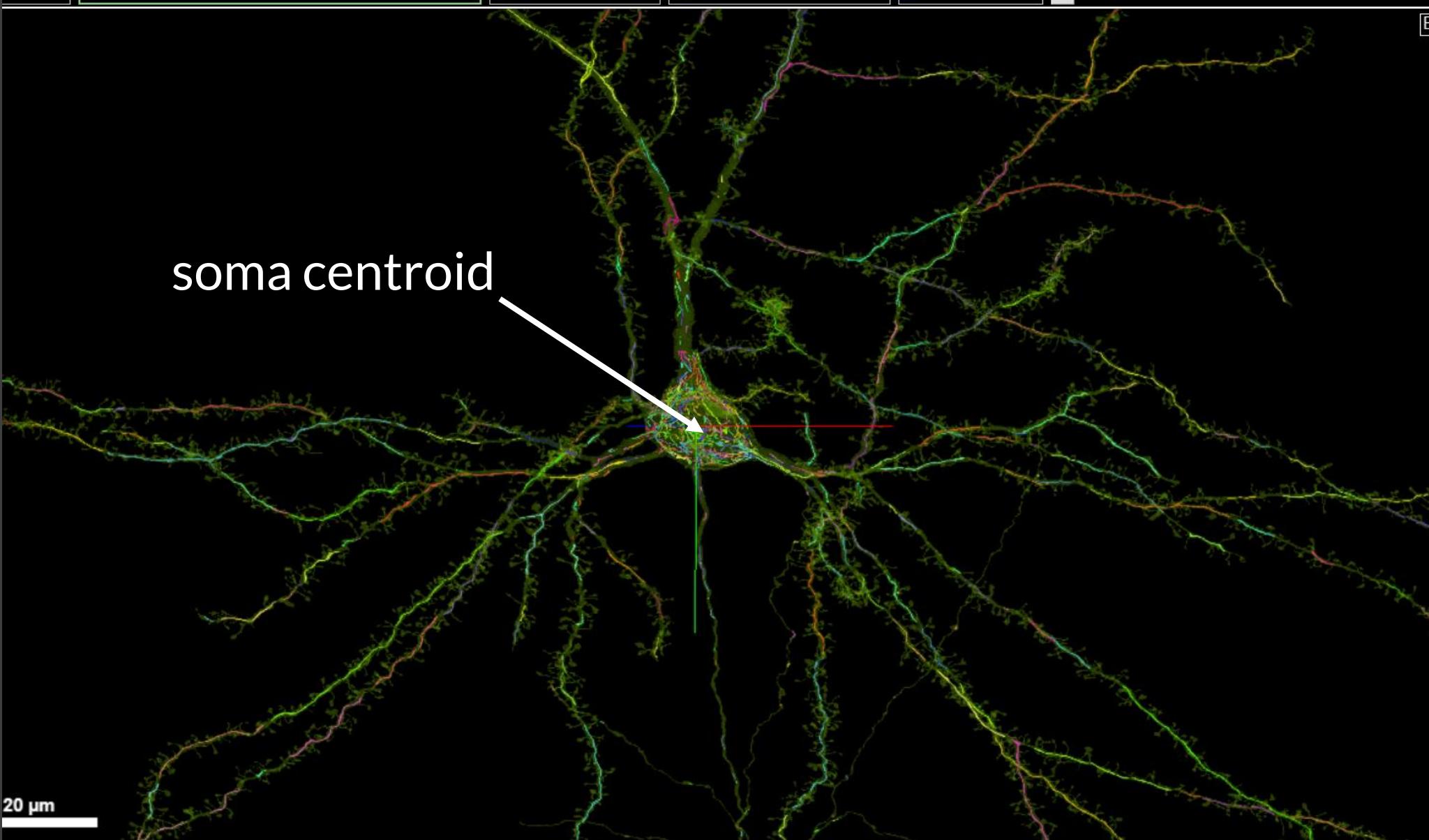
+ []

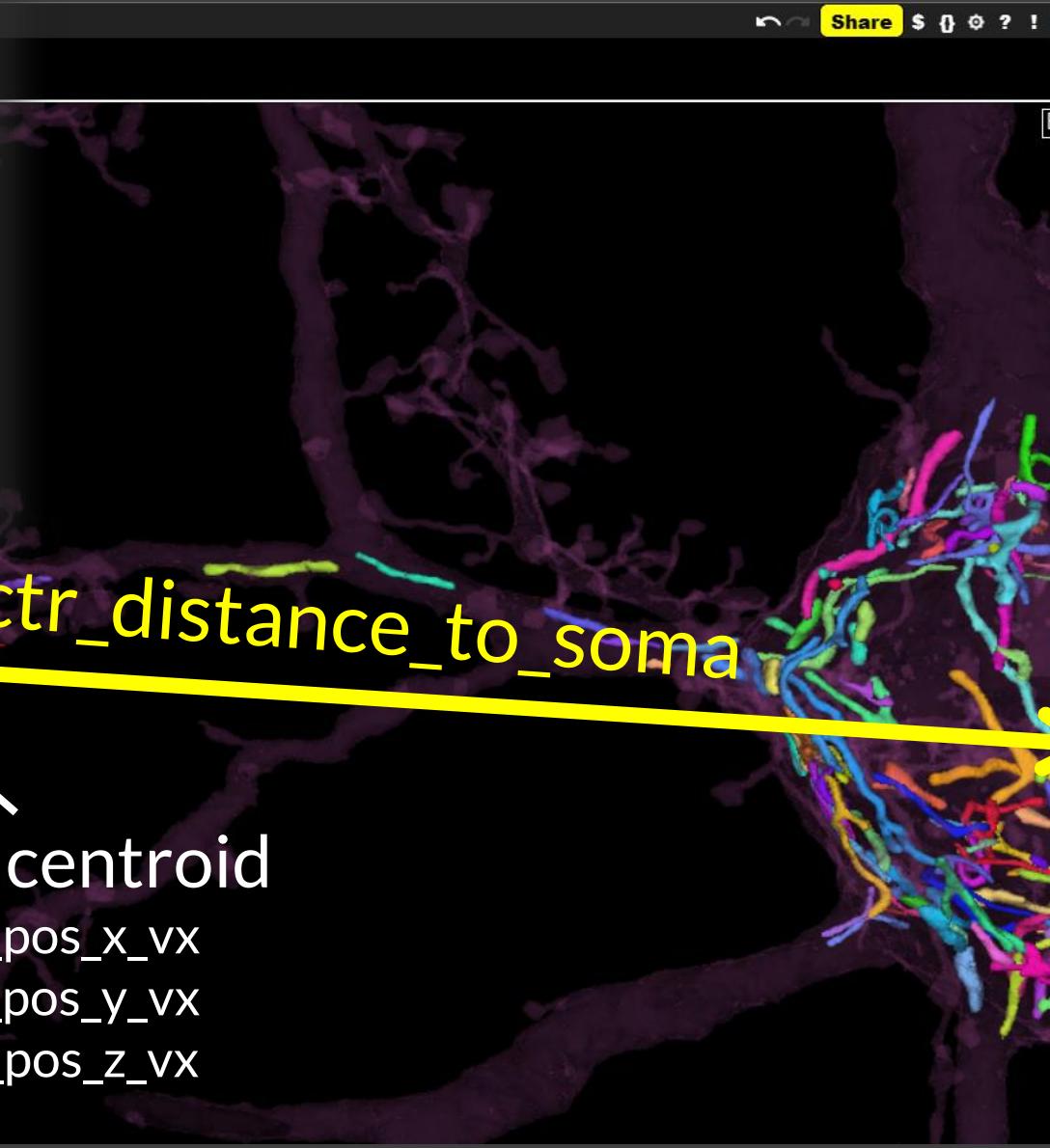
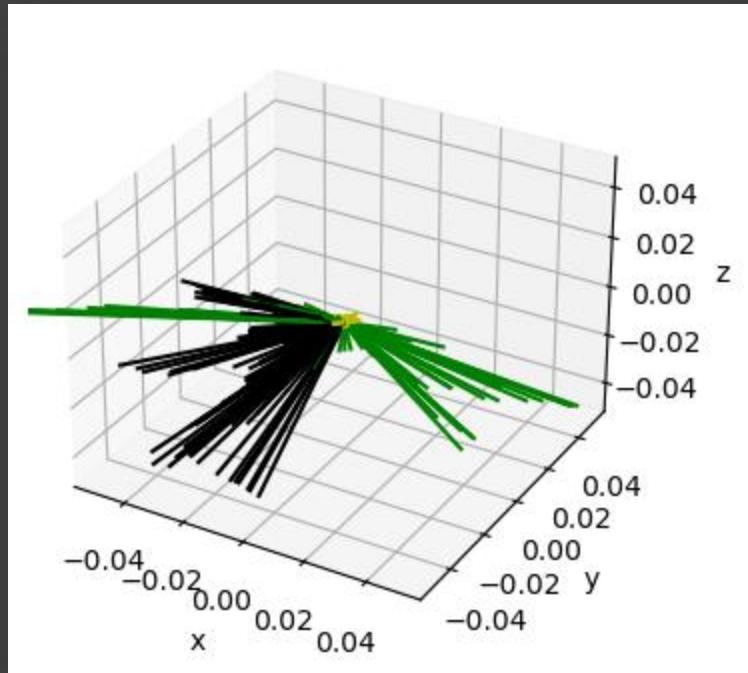
x [] [] [] []

648518346349538053 []

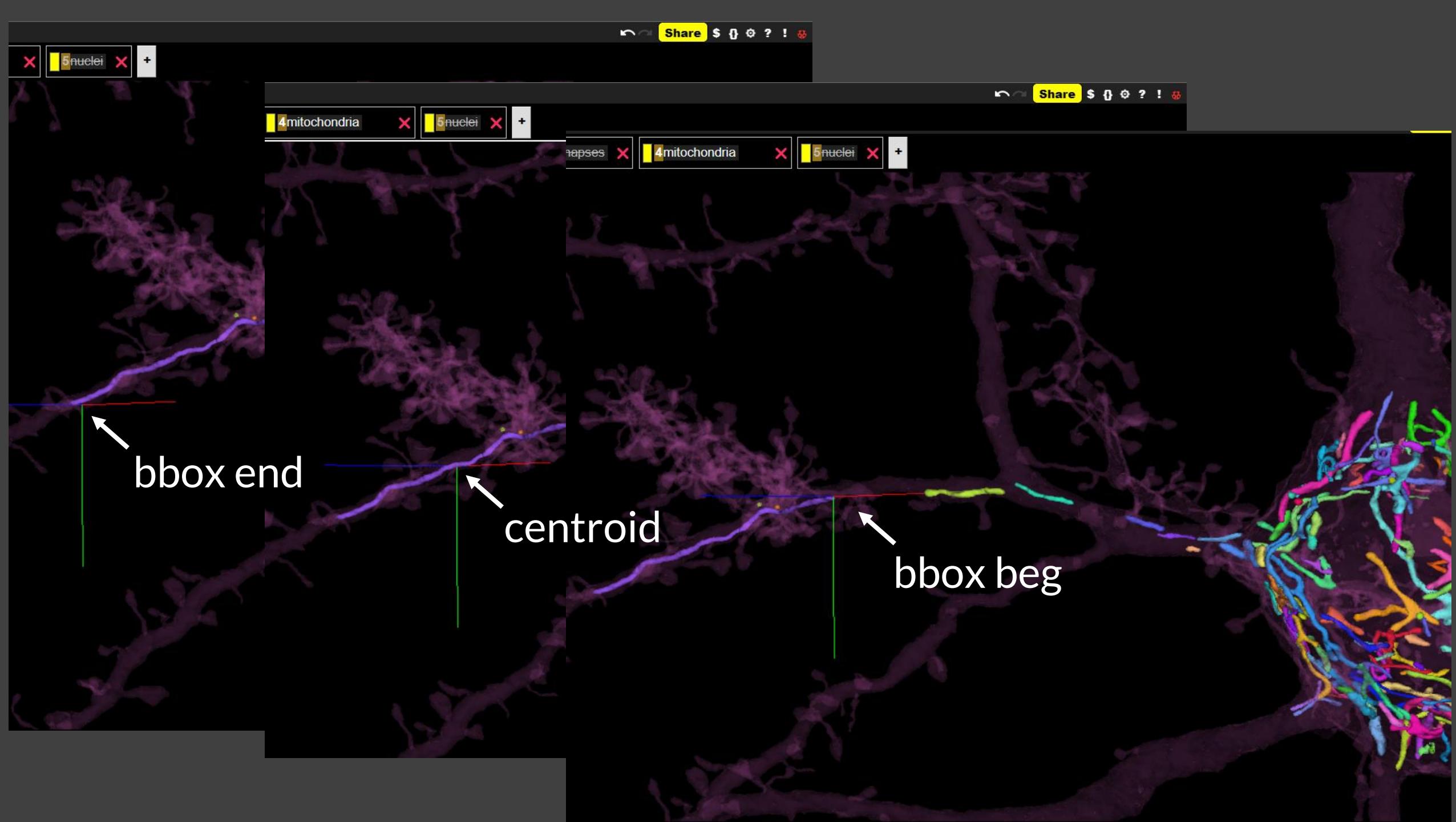
▼ Omni Segment Info

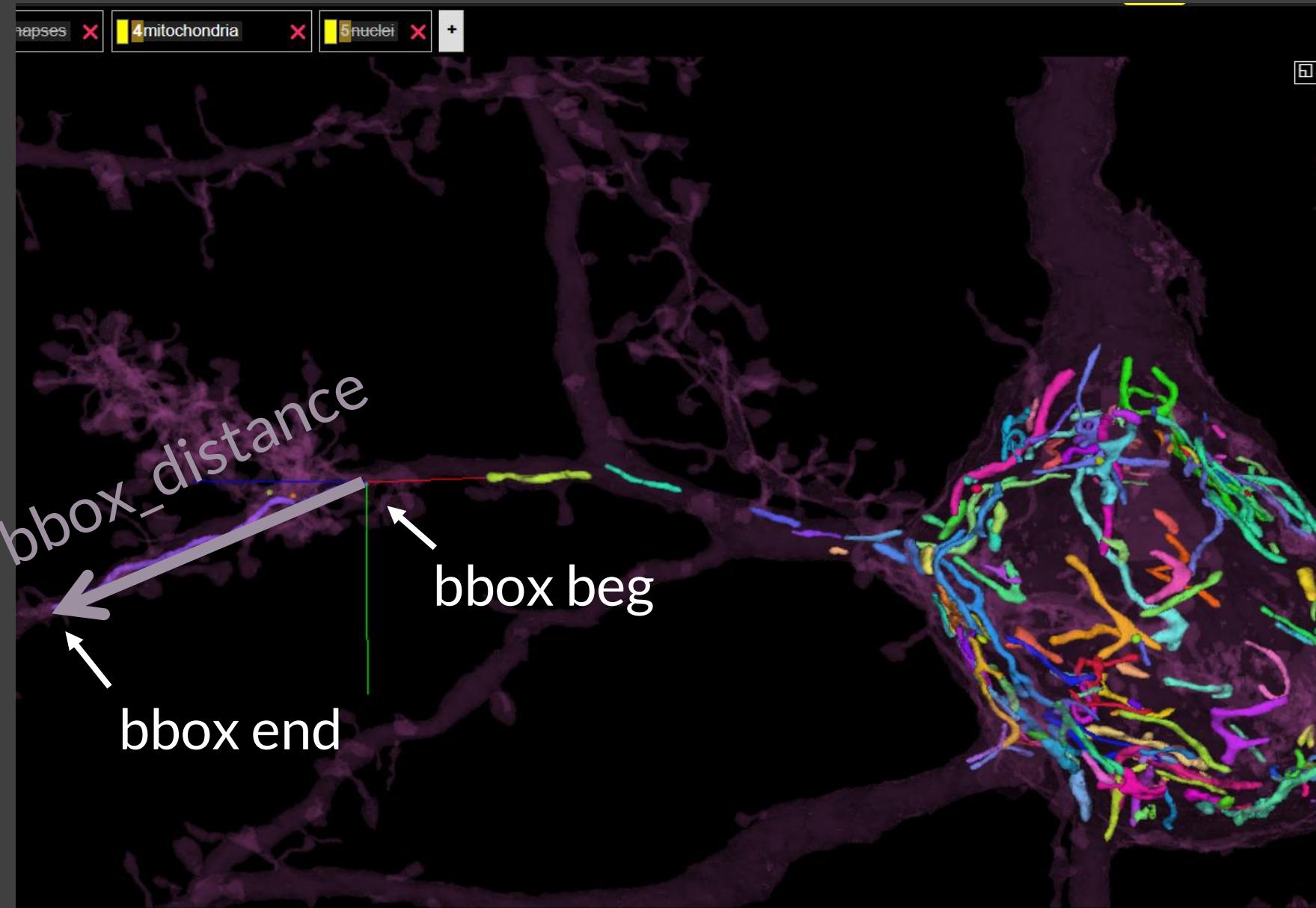
soma centroid





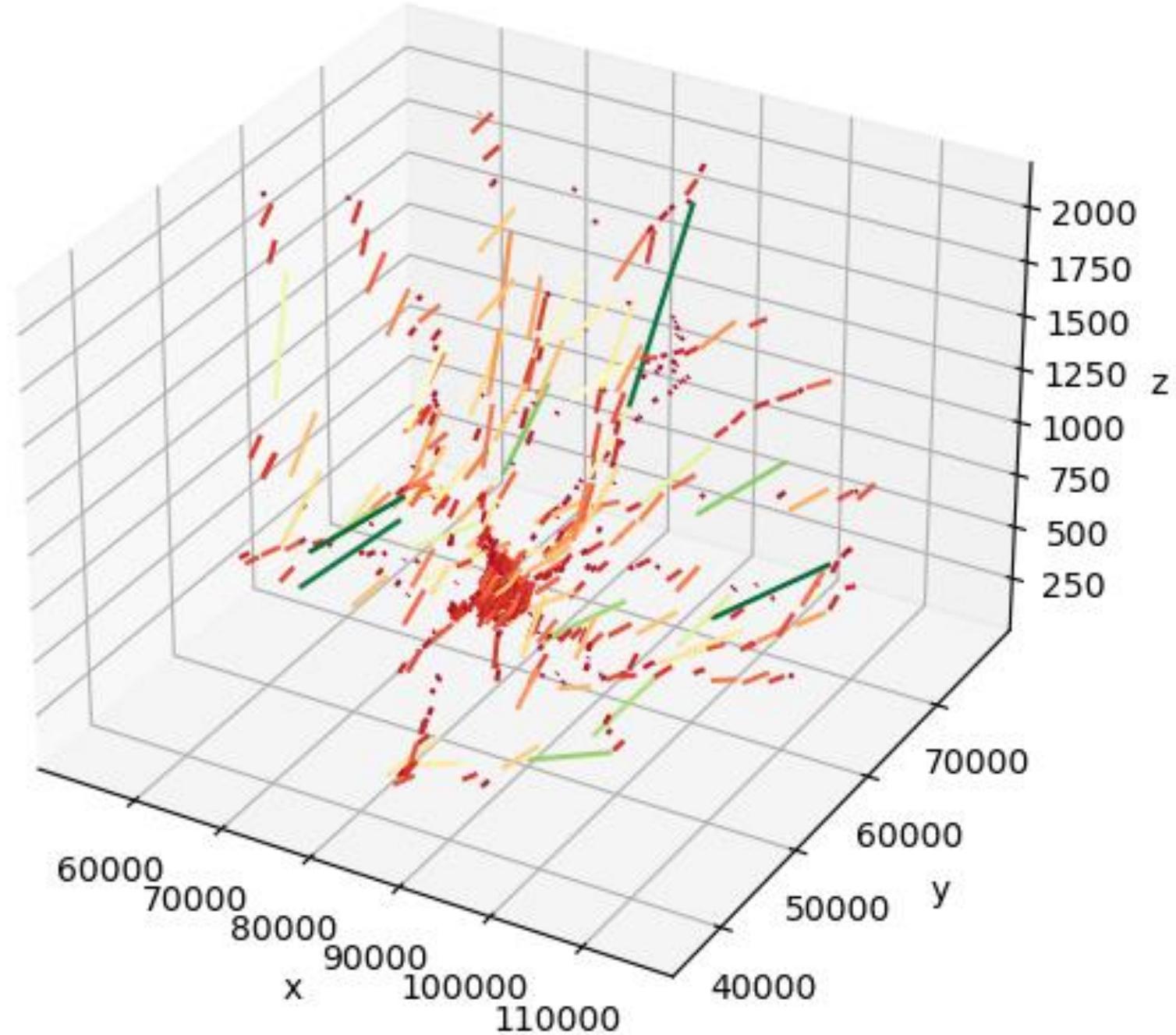
```
np.sqrt((ctr_pos_x_vx - centroid_x)**2 + (ctr_pos_y_vx - centroid_y)**2 + (ctr_pos_z_vx - centroid_z)**2)
```



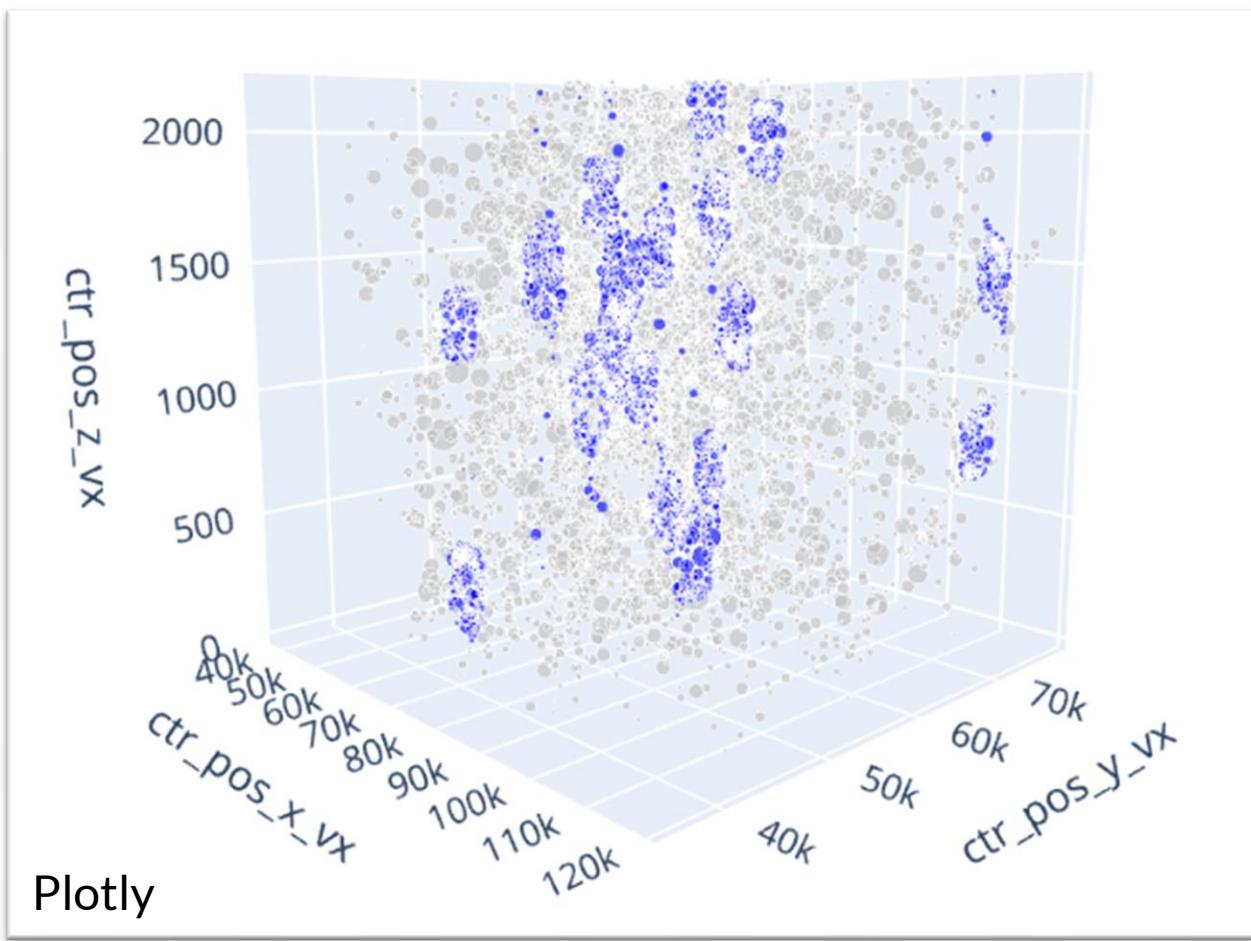


```
np.sqrt((bbox_beg_x_vx - bbox_end_x_vx)**2 + (bbox_beg_y_vx - bbox_end_y_vx)**2 + (bbox_beg_z_vx - bbox_end_z_vx)**2)
```

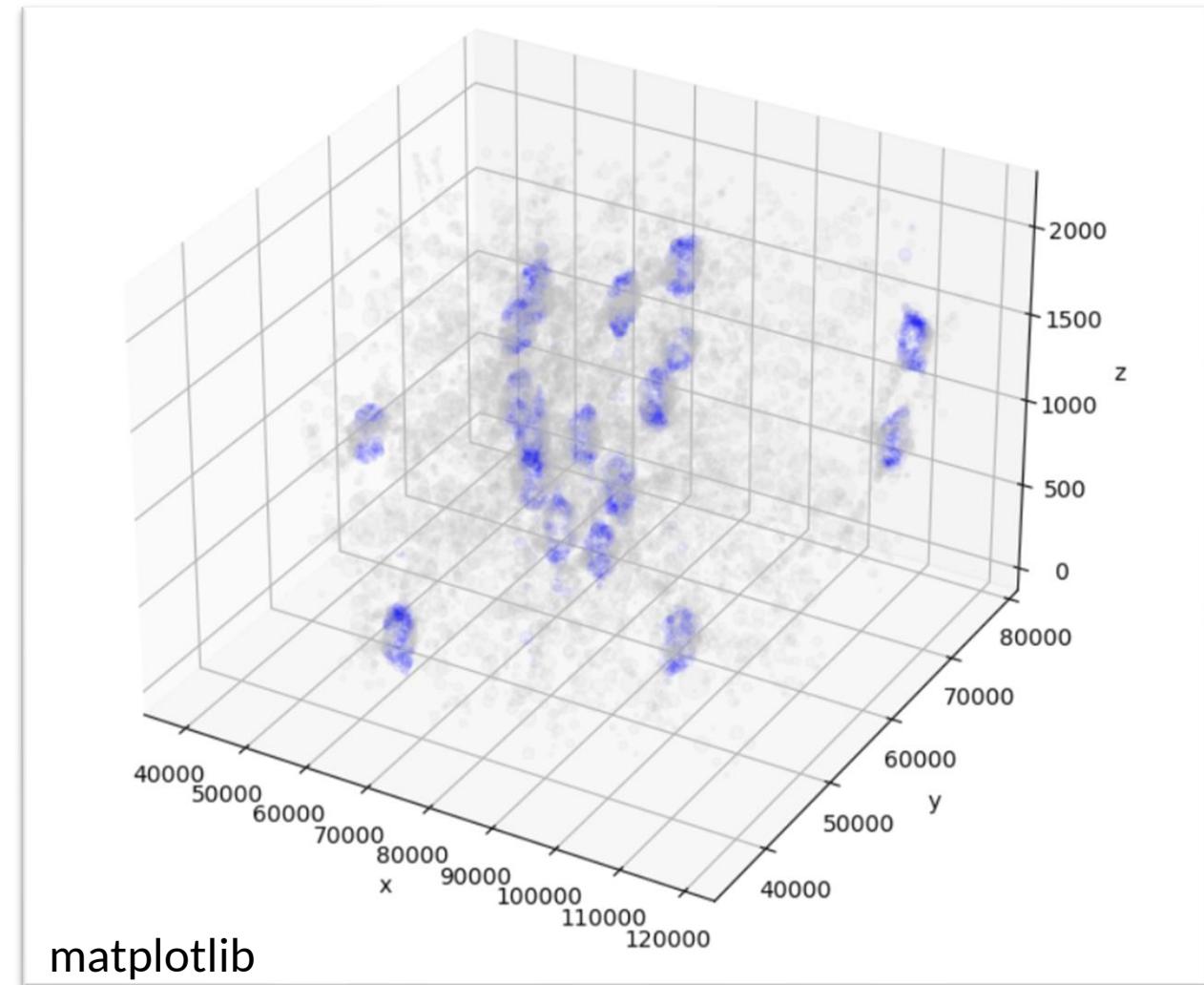
Quiver plot of bbox distance



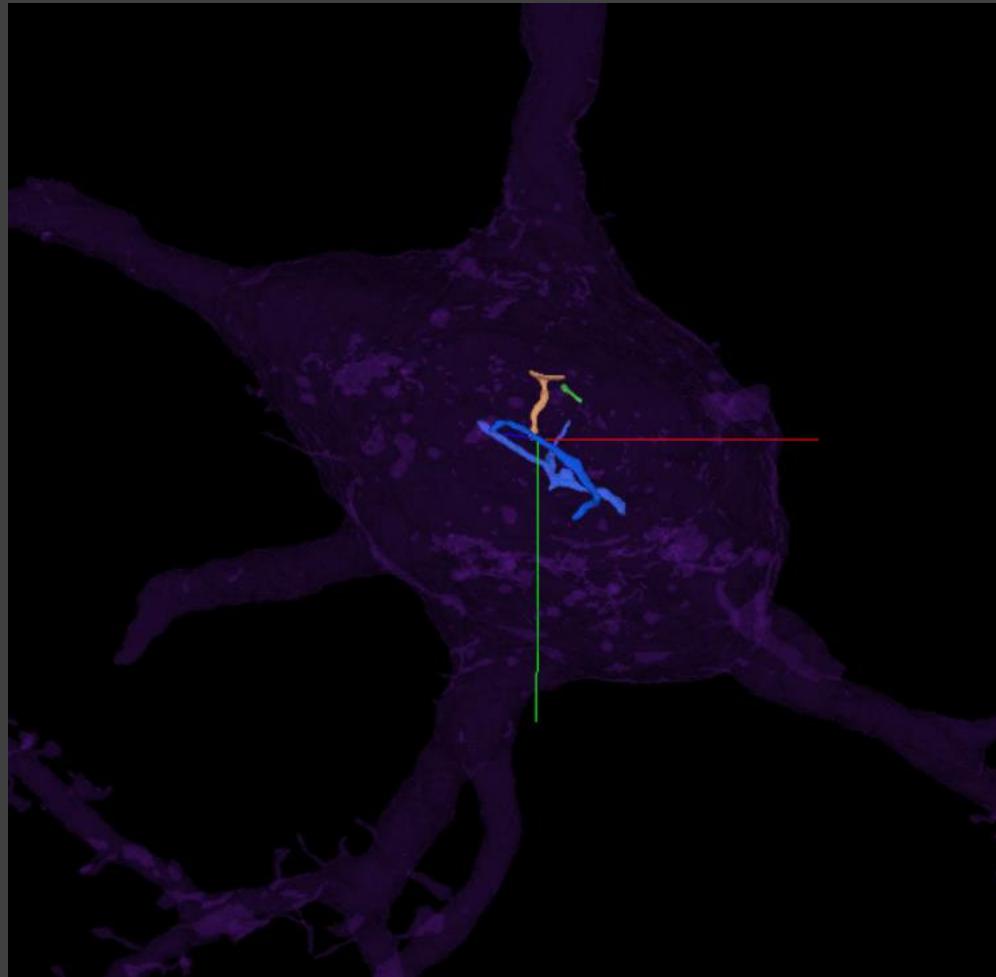
Mito distance from soma Boolean threshold



Plotly

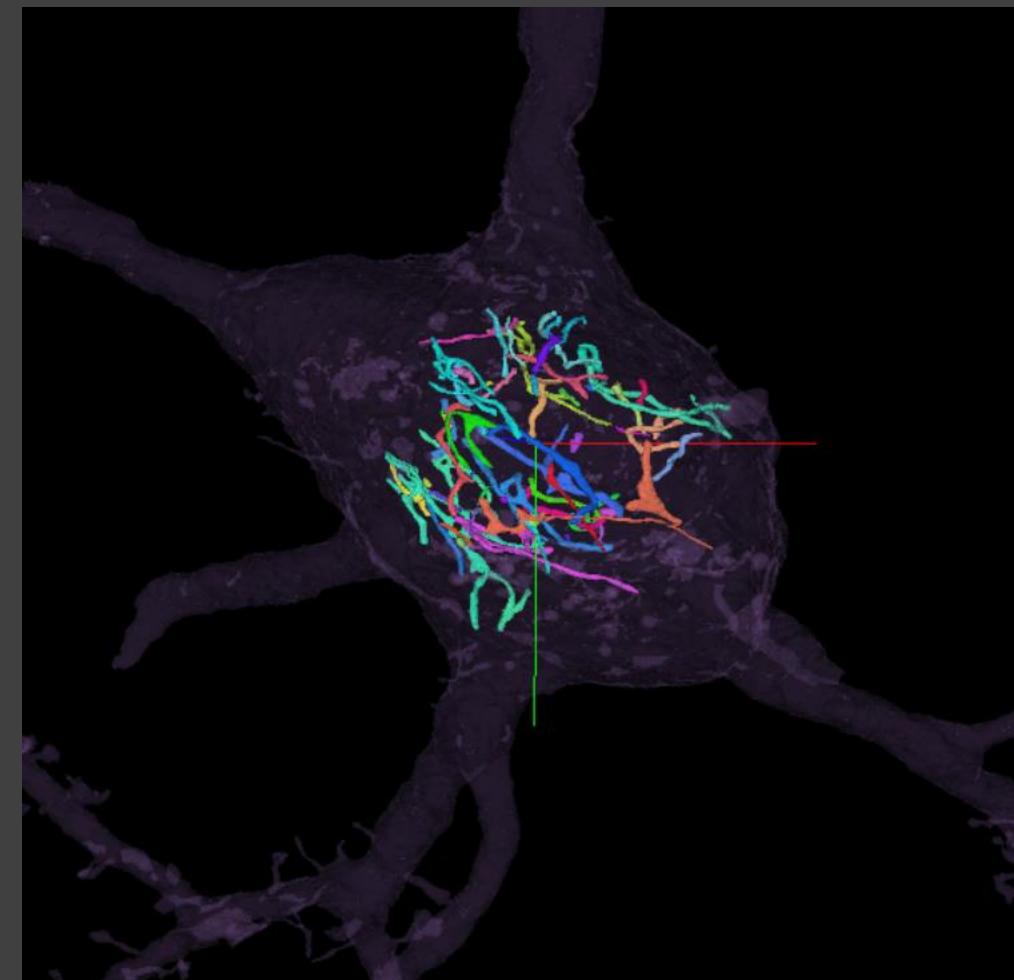


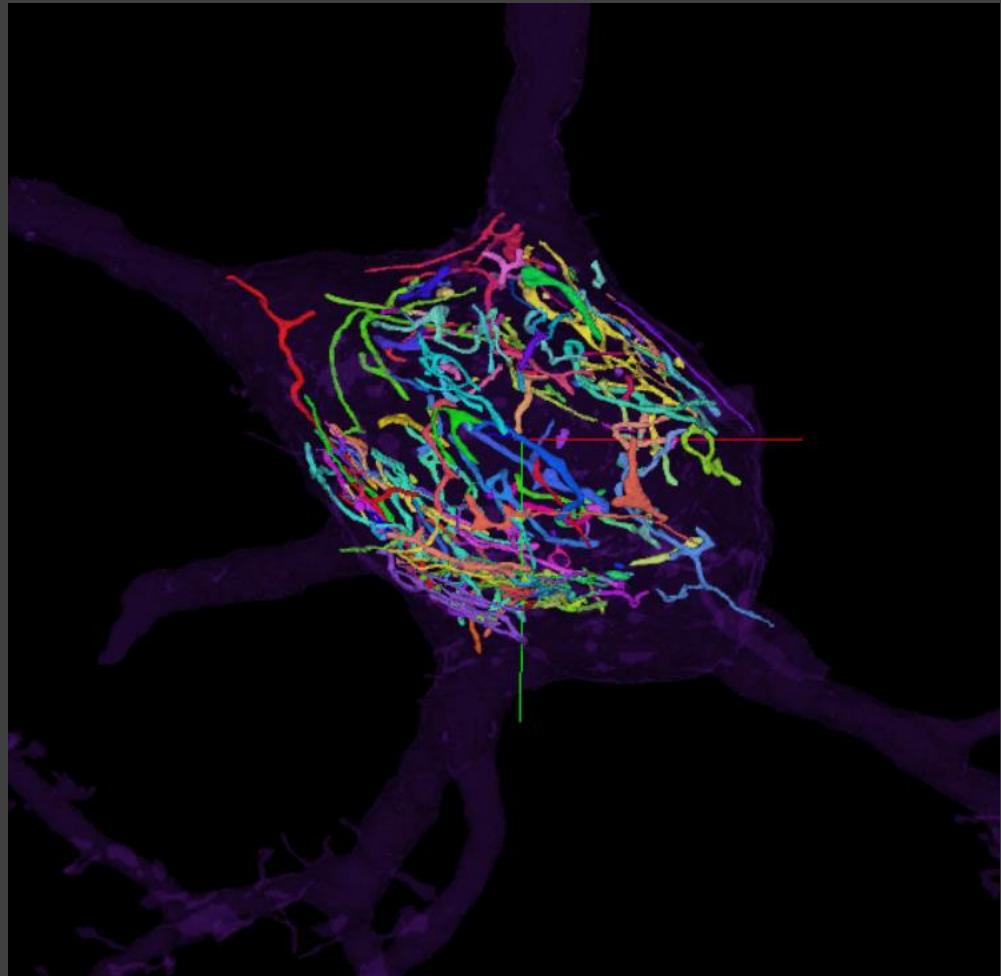
matplotlib



within 500 voxels

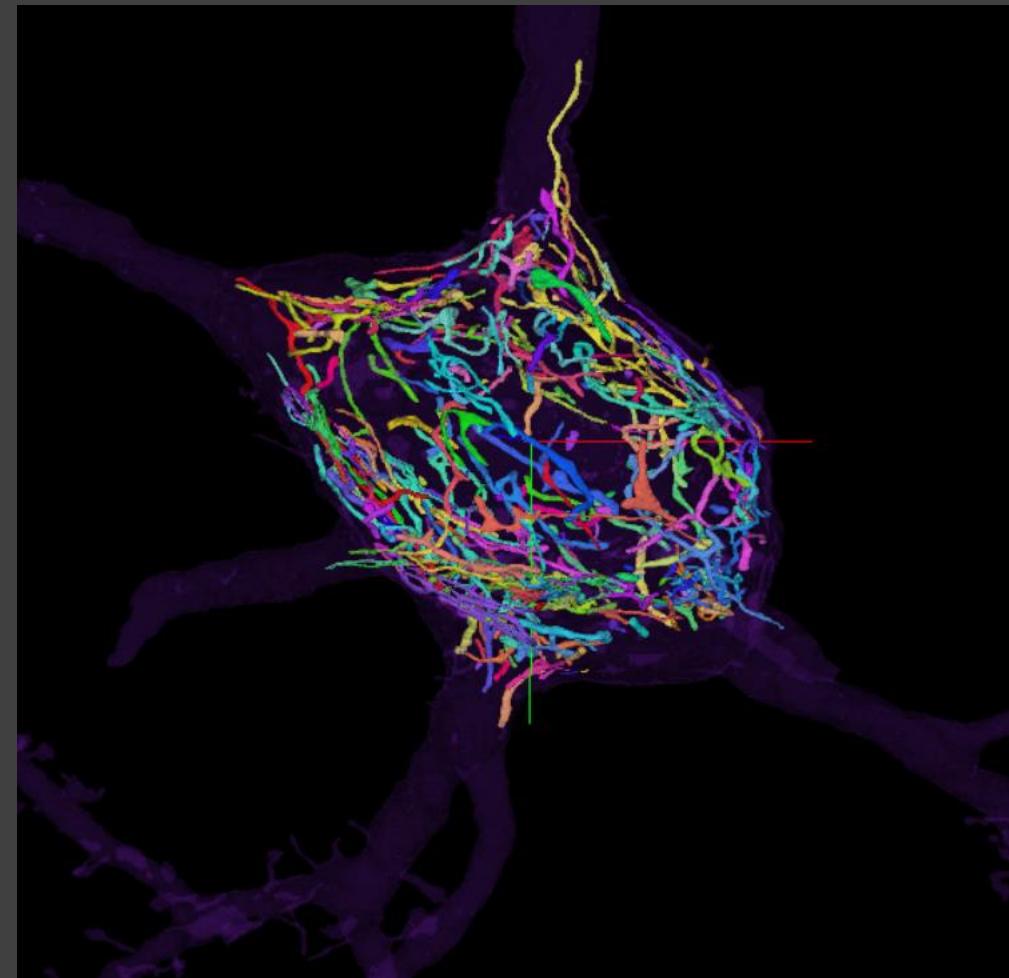
within 1000 voxels



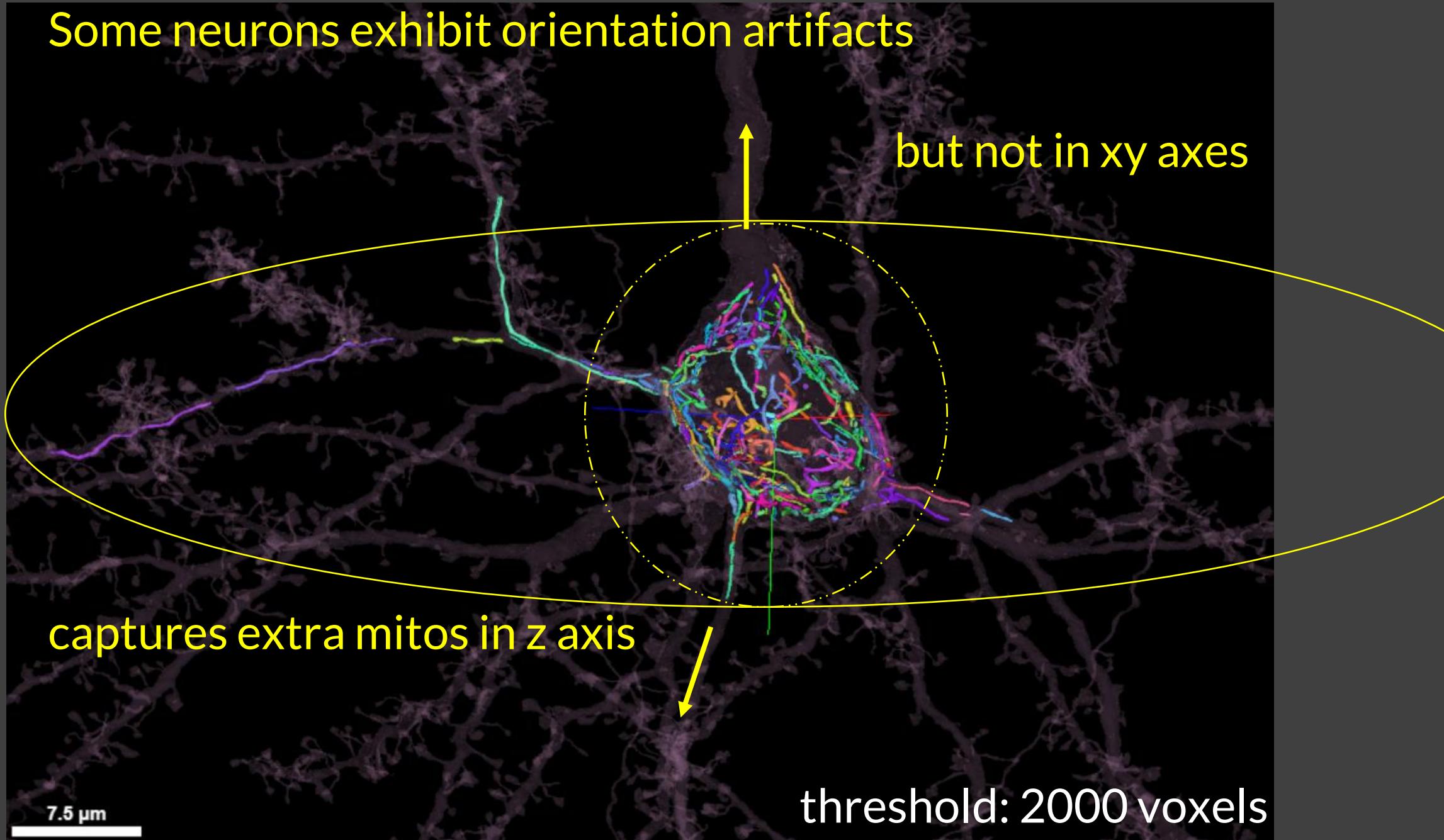


within 1500 voxels

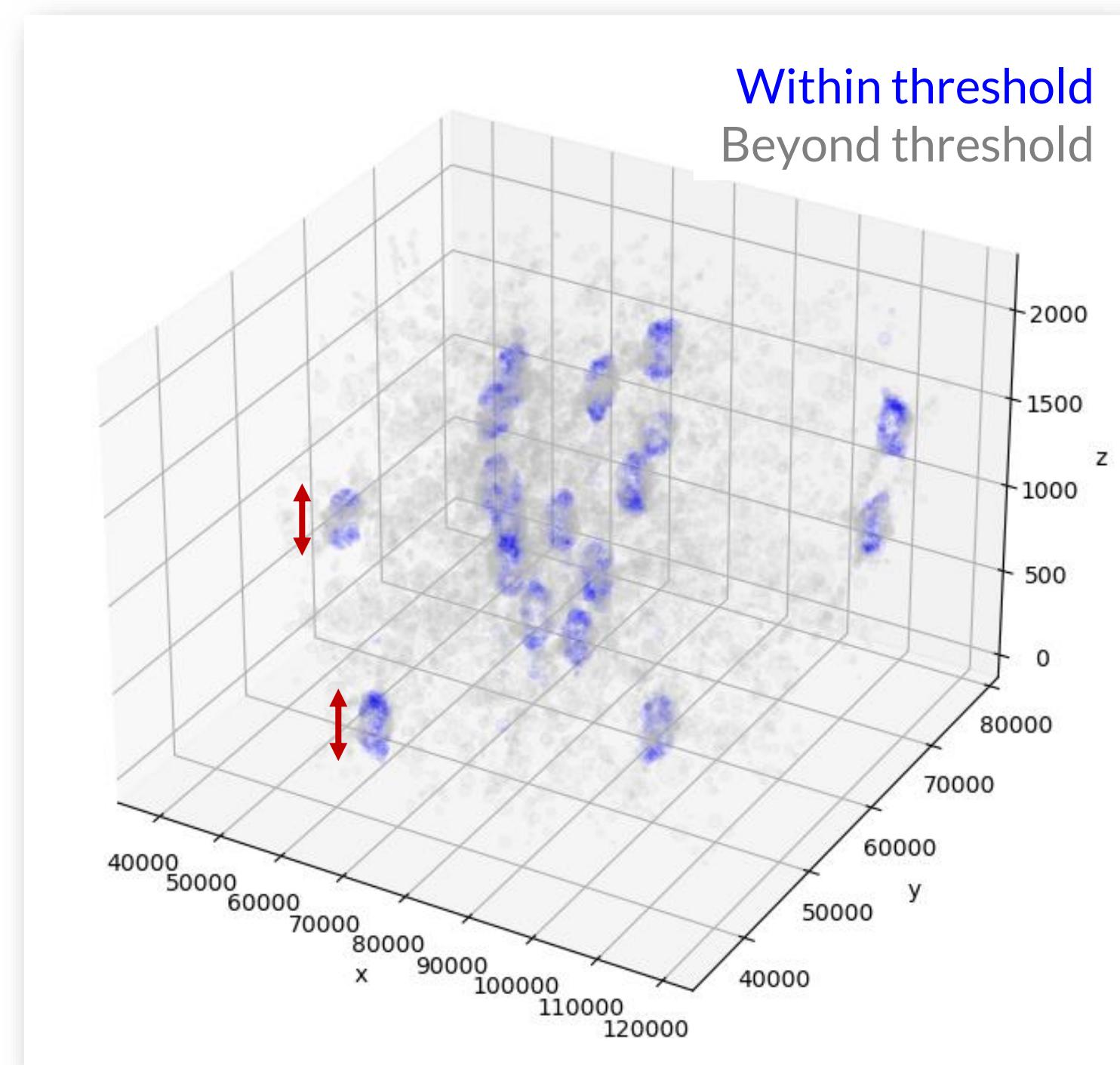
within 2000 voxels



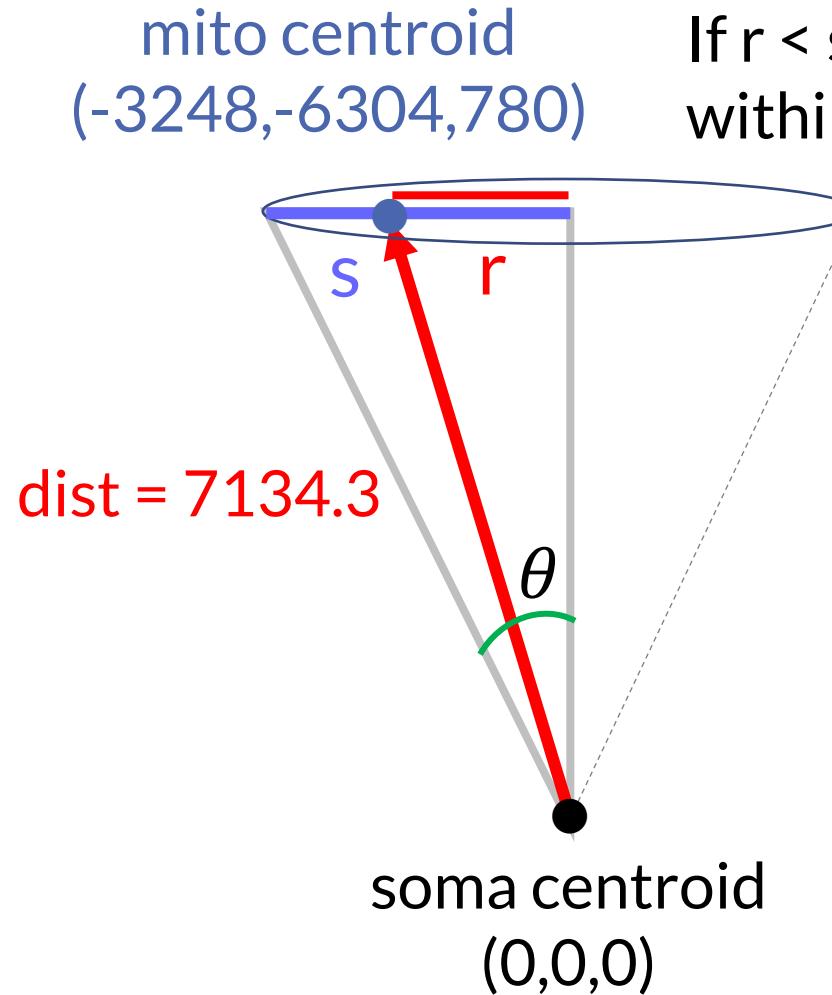
Some neurons exhibit orientation artifacts



There does
appear to be a
“stretching”
artifact in the
z-axis

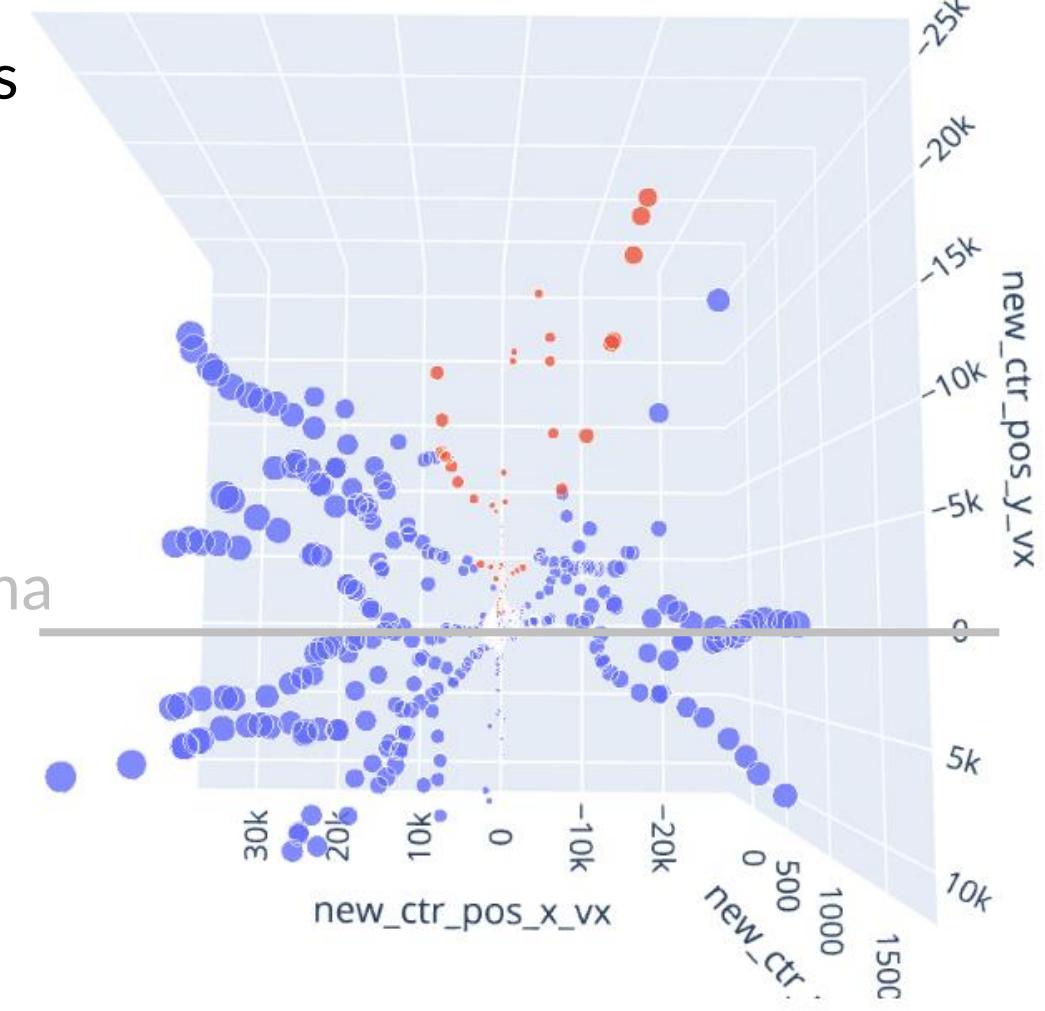


Mito centroid within a 45° cone above soma



If $r < s$, mito ctr is
within the cone

above soma
threshold



mito-to-mito distance summary stats

Data reduction approach calculating summary stats for each mito relative to others in cell

Create 363 dataframes: one per pyr neuron each with all its respective mitos

Create 363 new dfs where mito-to-mito (ctr-to-ctr) is calculated; one column per mito

Apply describe(), kurtosis & skew; transpose dfs, run a concat script to create final df

Creates summary stats for each mito (mean, std, n, min, Q1, m, Q3, max, kurt, skew)

	mito_id	mito_vx	ctr_pos_x_vx	ctr_pos_y_vx	ctr_pos_z_vx	bbox_beg_x_vx	bbox_beg_y_vx	bbox_beg_z_vx	bbox_end_x_vx	bbox_end_y_vx	bbox_e
0	1783325	527936	75186	43844	1430	74824	43668	1323	75472	44120	
1	1780086	1597632	75086	43678	1032	74168	43342	675	76458	44298	
2	2026719	141924	79378	47892	742	79192	47734	738	79532	48034	
3	1540216	2780	69738	42838	21	69708	42816	21	69770	42862	
4	1547760	211468	70706	45080	1364	70248	44874	1328	71030	45284	
...
210970	1649818	287444	72610	37220	2135	72356	37078	2095	72840	37450	
210971	2397704	336092	84174	56050	1975	83712	55254	1936	84612	56830	
210972	1667802	337984	72740	46280	1200	72594	45918	1153	72916	46642	
210973	1654227	495816	72104	39630	1908	72006	38894	1826	72178	40170	
210974	1088579	1012668	61868	43872	755	61088	43396	586	62586	44688	

183663 rows × 33 columns

Dataframe generated by Turner et al. Cell 2022
 183,663 rows of mito spatial data and mito volume, sa and complexity index

	cellid	mito_id	ctr_pos_x_vx	ctr_pos_y_vx	ctr_pos_z_vx	mito_id_2635064_ctr_dist	mito_id_895357_ctr_dist	mito_id_2151099_ctr_dist
12110	648518346349538440	2635064	90120	59784	1155	0.000000	29444.611205	11974.253380
12111	648518346349538440	895357	61000	55436	1480	29444.611205	0.000000	19947.458510
12112	648518346349538440	2151099	80708	52382	1069	11974.253380	19947.458510	0.000000
12113	648518346349538440	1813502	75278	58758	1026	14877.979735	14666.393013	8374.970149
12114	648518346349538440	2158762	81460	53710	1219	10577.956892	20534.332154	1533.488833
...	pyr 1	mito 1	mito 2	mito 3
13074	648518346349538440	3535921	105204	66310	1865	16450.526800	45523.463368	28190.012700
13075	648518346349538440	1009531	59532	58304	669	30627.640131	3322.374001	21992.113586
13076	648518346349538440	2429032	84430	71346	193	12922.127843	28350.473876	19345.644368
13077	648518346349538440	2047173	78534	58314	731	11686.576573	17784.404432	6326.858936
13078	648518346349538440	2419754	84528	67414	488	9483.261728	26420.131188	15520.663162

969 rows × 974 columns

362 rows × 367 columns

← pyr 2
← pyr ...
← pyr 363

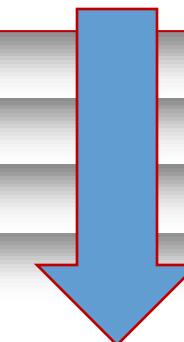
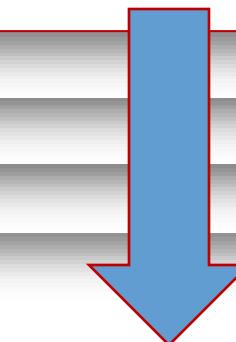
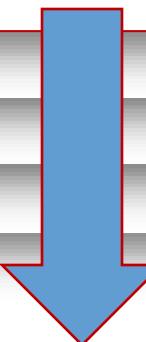
Generate 363 dfs (one for each pyr neuron) with
 All the mitochondria for that pyr neuron in the rows (range: 40-969)
 Each mito has a column for mito-to-mito distance (ctr-ctr) to other mitos

	cellid	mito_id	ctr_pos_x_vx	ctr_pos_y_vx	ctr_pos_z_vx	mito_id_2635064_ctr_dist	mito_id_895357_ctr_dist	mito_id_2151099_ctr_dist
12110	648518346349538440	2635064	90120	59784	1155	0.000000	29444.611205	11974.253380
12111	648518346349538440	895357	61000	55436	1480	29444.611205	0.000000	19947.458510
12112	648518346349538440	2151099	80708	52382	1069	11974.253380	19947.458510	0.000000
12113	648518346349538440	1813502	75278	58758	1026	14877.979735	14666.393013	8374.970149
12114	648518346349538440	2158762	81460	53710	1219	10577.956892	20534.332154	1533.488833
...	pyr 1	mito 1	mito 2	mito 3
13074	648518346349538440	3535921	105204	66310	1865	16450.526800	45523.463368	28190.012700
13075	648518346349538440	1009531	59532	58304	669	30627.640131	3322.374001	21992.113586
13076	648518346349538440	2429032	84430	71346	193	12922.127843	28350.473876	19345.644368
13077	648518346349538440	2047173	78534	58314	731	11686.576573	17784.404432	6326.858936
13078	648518346349538440	2419754	84528	67414	488	9483.261728	26420.131188	15520.663162

969 rows × 974 columns

362 rows × 367 columns

← pyr 2
← pyr ...
← pyr 363



Generate summary stats (mean, std, n, min, Q1, m, Q3, max, kurt, skew)

Transpose each column (mito 1 to mito n) into a row

Save to new df



	count	mean	std	min	25%	50%	75%	max	kurtosis	skew	mito_id	compartment	mito_vx				
0	642.0	16408.245694	4767.637466	0.0	14753.795255	16409.924902	18589.884046	31352.017415	-0.189384	0.339963	3205040	Basal	1639148				
1	642.0	10049.082873	6470.362301	0.0	5552.398179	7309.785638	14137.185185	33252.046493	3.987859	1.837180	3156851	Unknown dendritic	519460				
2	642.0	17837.727366	5948.256588	0.0	15439.063309	17044.807566	21485.449014	32015.629433	-0.576927	0.173319	2508608	Basal	1263852				
3	642.0	23399.181327	6357.203746	0.0	22385.160397	24207.483398	26411.792909	41894.669589	-0.758795	0.083111	3551417	Axonal	14828				
4	642.0	7531.763114	6570.694758	0.0	2141.664044	5598.224813	11325.461225	29030.409505	4.415615	1.967081	3491756	Somatic	22924				
...
637	642.0	9523.981739	5946.857141	0.0	5125.800306	7001.471445	13300.321068	30649.243270	3.830654	1.797771	3162450	Basal	1634616				
638	642.0	7321.702610	6403.291539	0.0	2021.759787	5600.611597	11426.506090	27796.442506	4.129071	1.902139	3495831	Somatic	666264				
639	642.0	7329.623726	6130.825303	0.0	2084.369692	5970.653169	11502.925033	26879.939509	3.874525	1.837272	3498813	Somatic	728844				
640	642.0	7351.699082	6501.843785	0.0	1858.929861	5612.884984	11399.974104	28395.032048	4.243022	1.929288	3495186	Somatic	797712				
641	642.0	7311.020300	6368.708339	0.0	2079.267560	5582.118821	11406.012494	27621.678298	4.109612	1.897274	3495579	Somatic	1310224				

642 rows × 18 columns

531 rows × 18 columns

363 new dfs (one for each pyr neuron)

Summary stats of mito-to-mito dist are merged with the other mito metrics (volume, sa, complexity, soma distance & above soma

	count	mean	std	min	25%	50%	75%	max	kurtosis	skew	mito_id	compartment	mito_vx				
0	642.0	16408.245694	4767.637466	0.0	14753.795255	16409.924902	18589.884046	31352.017415	-0.189384	0.339963	3205040	Basal	1639148	vx			
1	642.0	10049.082873	6470.362301	0.0	5552.398179	7309.785638	14137.185185	33252.046493	3.987859	1.837180	3156851	Unknown dendritic	519460	340			
2	642.0	17837.727366	5948.256588	0.0	15439.063309	17044.807566	21485.449014	32015.629433	-0.576927	0.173319	2508608	Basal	1263852	392	nito_		
3	642.0	23399.181327	6357.203746	0.0	22385.160397	24207.483398	26411.792909	41894.669589	-0.758795	0.083111	3551417	Axonal	14828	104	nito_		
4	642.0	7531.763114	6570.694758	0.0	2141.664044	5598.224813	11325.461225	29030.409505	4.415615	1.967081	3491756	Somatic	22924	628	nito_		
...
637	642.0	9523.981739	5946.857141	0.0	5125.800306	7001.471445	13300.321068	30649.243270	3.830654	1.797771	3162450	Basal	1634616	...			
638	642.0	7321.702610	6403.291539	0.0	2021.759787	5600.611597	11426.506090	27796.442506	4.129071	1.902139	3495831	Somatic	666264	508			
639	642.0	7329.623726	6130.825303	0.0	2084.369692	5970.653169	11502.925033	26879.939509	3.874525	1.837272	3498813	Somatic	728844	992			
640	642.0	7351.699082	6501.843785	0.0	1858.929861	5612.884984	11399.974104	28395.032048	4.243022	1.929288	3495186	Somatic	797712	696			
641	642.0	7311.020300	6368.708339	0.0	2079.267560	5582.118821	11406.012494	27621.678298	4.109612	1.897274	3495579	Somatic	1310224	256			

642 rows × 18 columns

531 rows × 18 columns

Concatenate these 363 dfs into one single df



```
In [2]: mito_to_mito_dist_all = pd.read_csv('mito_to_mito_distance_statistics/compiled_mito_to_mito_distance_stats_cellid.csv')
mito_to_mito_dist_all.rename( columns={'Unnamed: 0':'nested_index'}, inplace=True )
mito_to_mito_dist_all
```

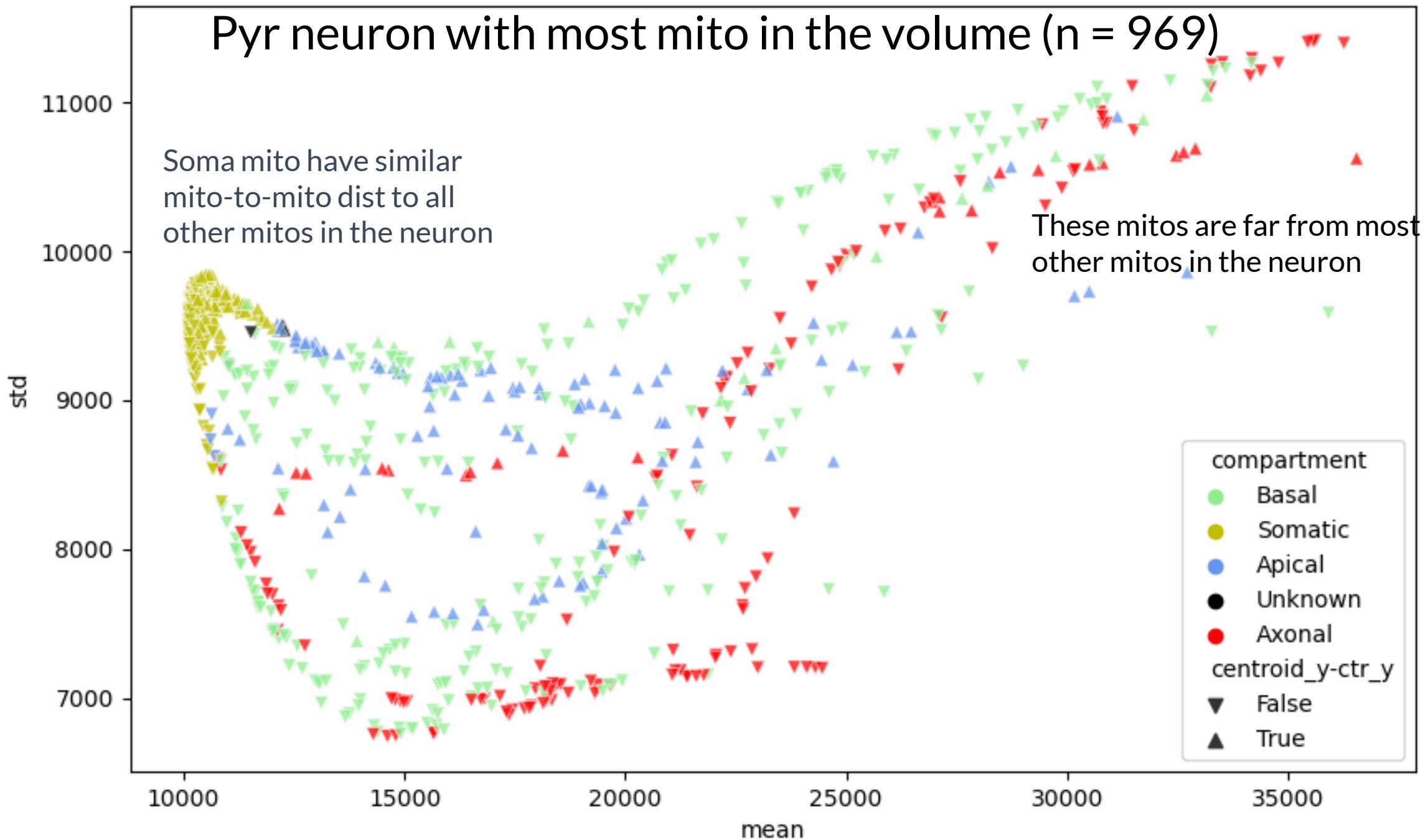
Out[2]:

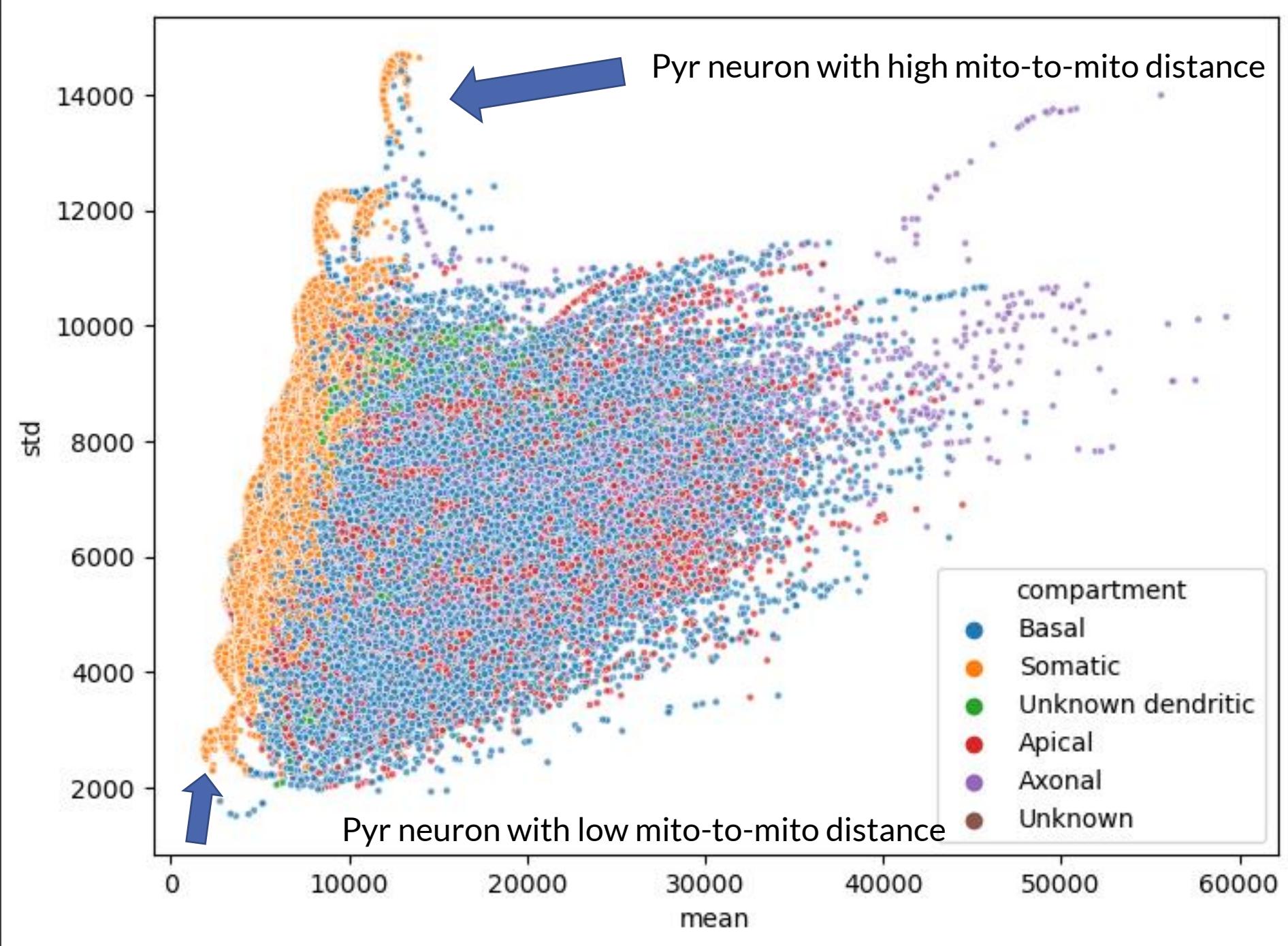
	nested_index	count	mean	std	min	25%	50%	75%	max	kurtosis	skew	mito_id	compartr
0	0	196.0	7493.879886	5358.067644	0.0	4229.199053	5547.331920	9169.323781	26461.627614	4.877342	2.025797	1783325	E
1	1	196.0	7403.495975	5418.682808	0.0	4052.491869	5378.002393	9101.405330	26497.305071	4.935129	2.044597	1780086	E
2	2	196.0	10724.406356	4860.556101	0.0	8462.439730	10156.453365	11852.577621	27373.207485	2.589345	1.296582	2026719	E
3	3	196.0	6237.536731	6097.146769	0.0	1940.415856	4356.927020	8496.087586	28787.191892	5.617236	2.248441	1540216	Sor
4	4	196.0	6307.454632	5294.853299	0.0	2791.498779	4378.841631	8158.576625	26607.713618	5.541097	2.220513	1547760	Unkn den
...
183658	183658	209.0	7130.690469	7336.003350	0.0	994.276119	4303.812844	12012.871930	31113.030132	4.364447	1.981548	1649818	Sor
183659	183659	209.0	21084.467285	6155.821683	0.0	20184.581269	21733.447403	23581.761872	36894.681405	-0.833545	0.013475	2397704	E
183660	183660	209.0	9532.196738	4332.100257	0.0	8150.885903	9234.419310	10328.859376	24579.066398	2.620475	1.290201	1667802	E
183661	183661	209.0	7038.485608	6170.602395	0.0	2679.756705	3817.004061	10301.231431	29525.049179	5.046709	2.128451	1654227	Sor
183662	183662	209.0	11538.665707	5220.474646	0.0	8690.312998	12376.476114	13501.446737	34503.112323	3.378458	1.578085	1088579	E

183663 rows × 20 columns

183,663 rows × 20 columns of mito identifiers & metrics

Pyr neuron with most mito in the volume (n = 969)





Pyr neuron with high mito-to-mito distance



Pyr neuron with low mito-to-mito distance



Training dataframe

- All known compartments
 - Axonal 12,414
 - Somatic 93,616
 - Apical dendrite 19,970
 - Basal dendrite 54,904
 - Total: 180,904
- Unknown compartments
 - Unknown dendritic: 1,137
 - Unknown: 1,622



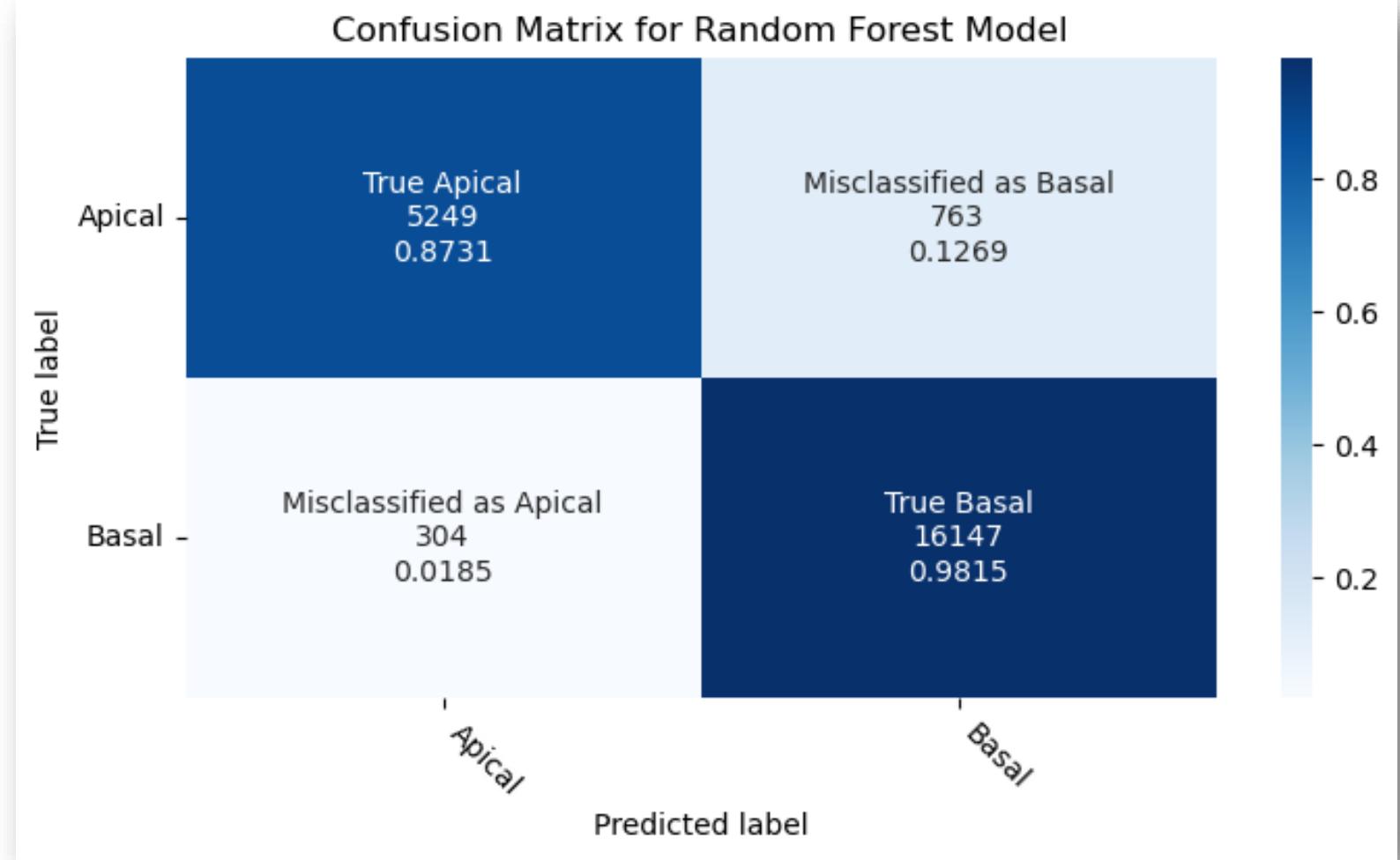
Simplify first

Instead of training on all four compartments, first train the two dendritic compartments so that the unknown dendritic mitos can be assigned to a dendritic compartment

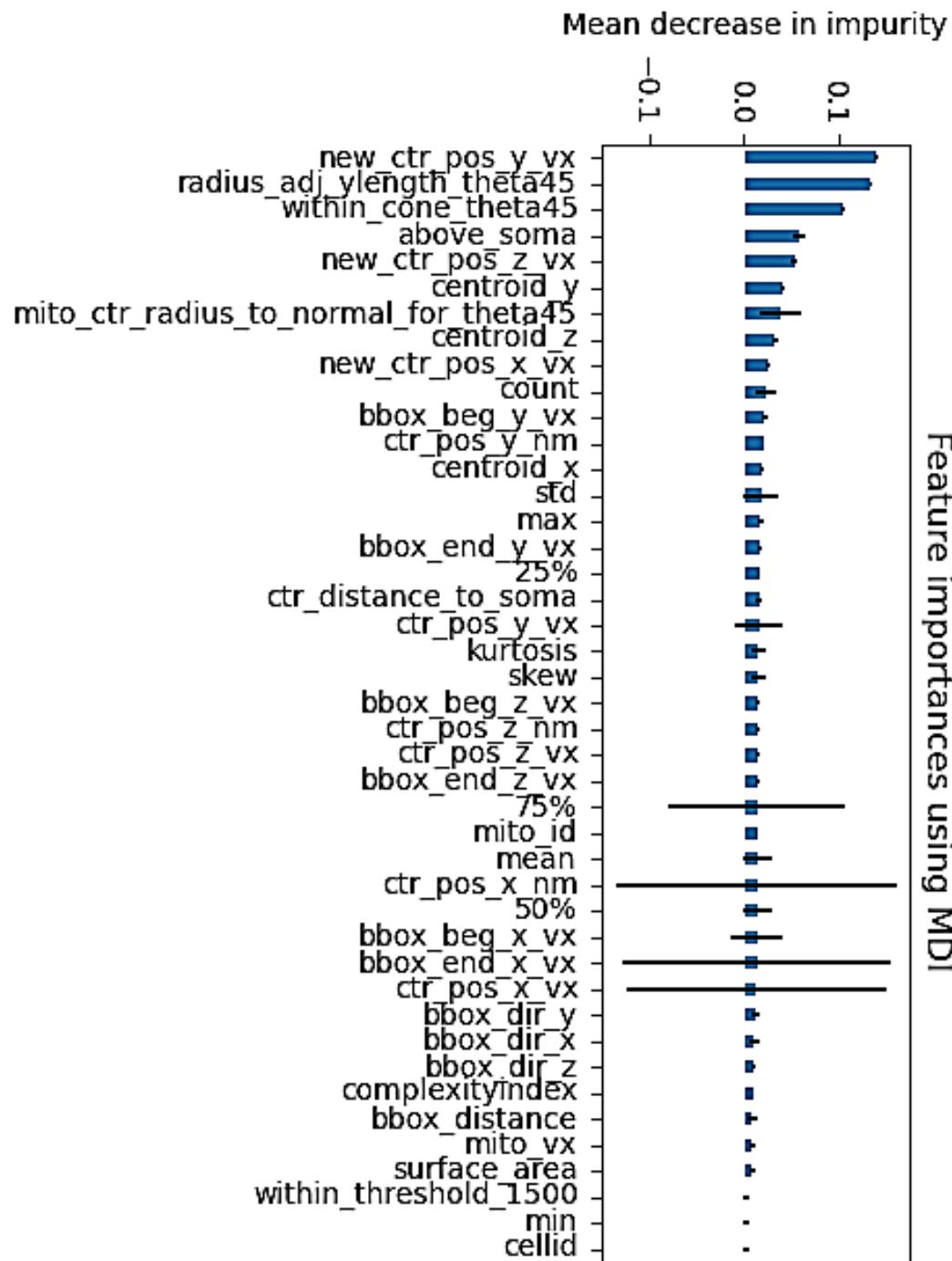
Training & Test dataframes

- Train compartments df
 - Apical dendrite 19,970
 - Basal dendrite 54,904
 - Total:** 74,874
- Test compartment df
 - Unknown dendritic 1,137

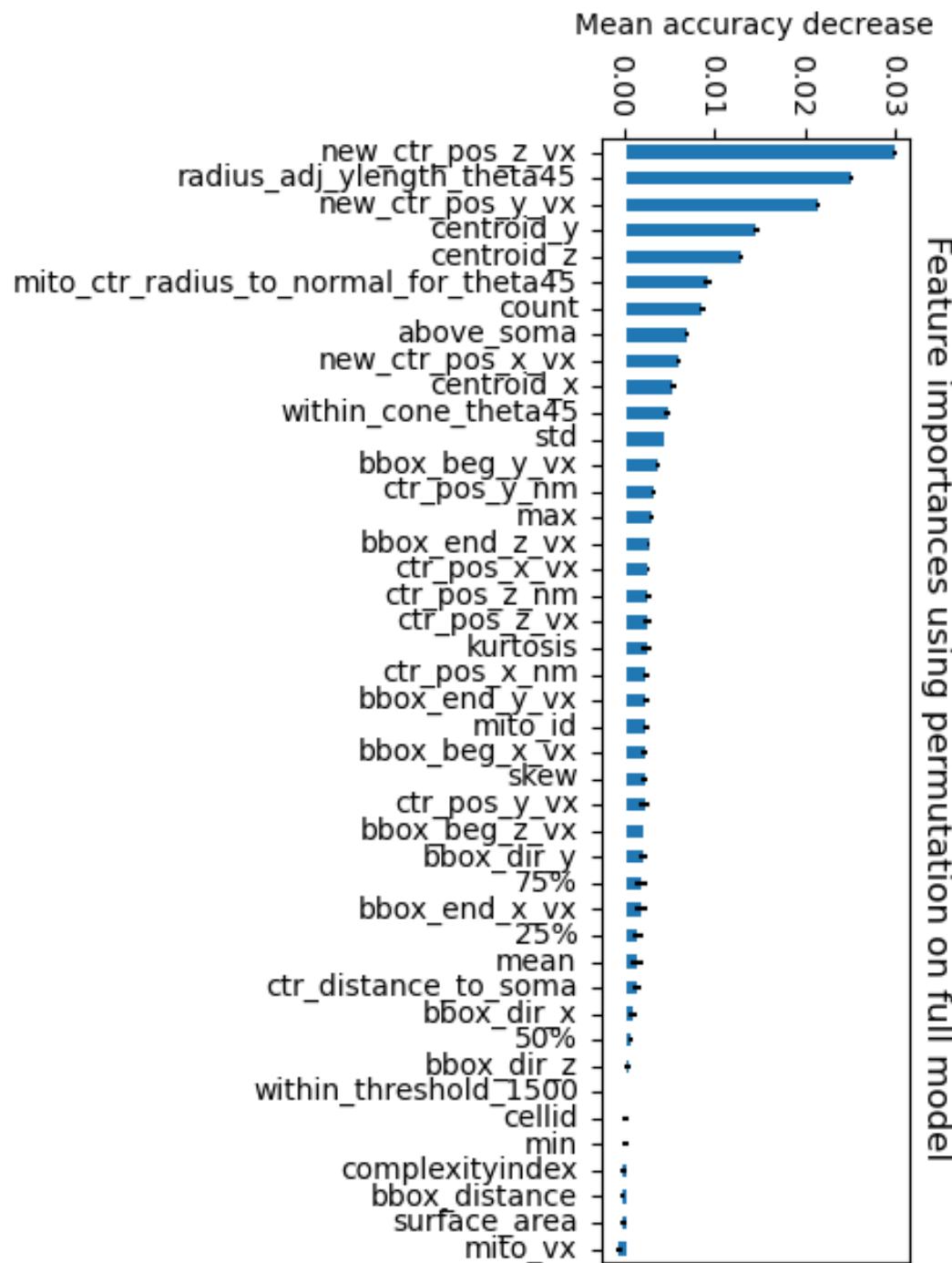
Apical vs Basal two compartment classifier (w/ withincone metric)



Apical vs Basal two- compartment classifier:by impurity



Apical vs Basal two- compartment classifier: by permutation



Validate 1,137 predictions for unknown dendritic compartment by manual inspection in Neuroglancer

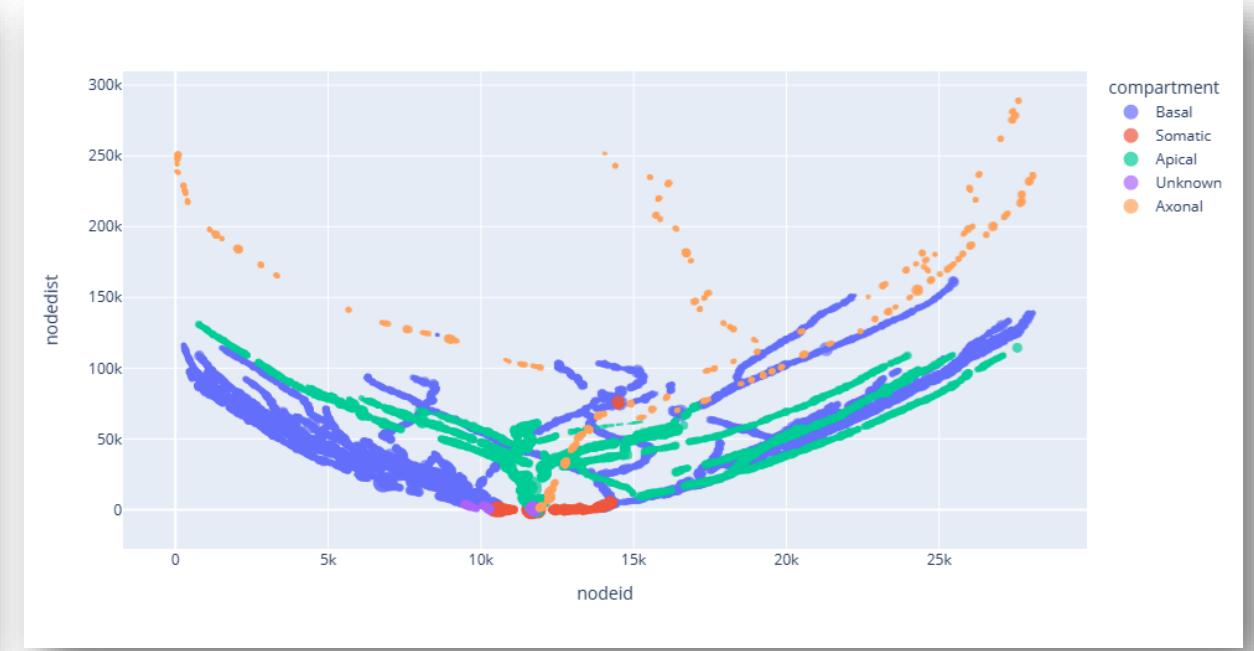
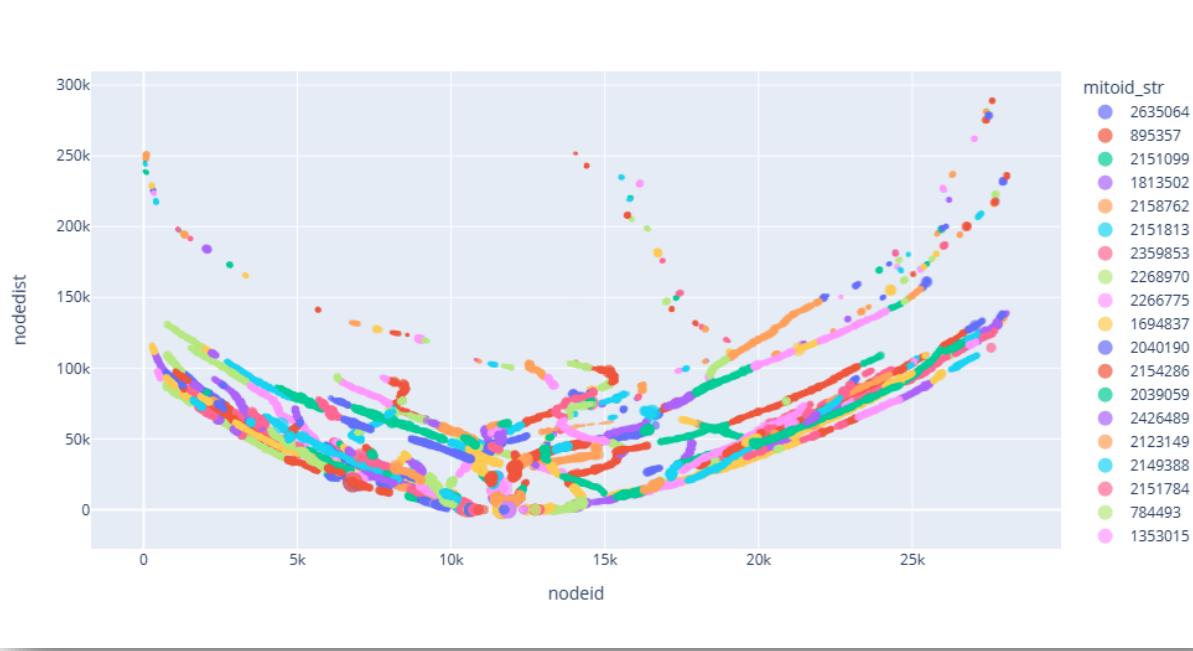
	cellid_str	mito_id	rfc_assignment	manual_inspection	correct_assignment	notes	ng_link
0	648518346349533350	3153917	Apical	Basal	False	apical dendrite is rotated off the vertical axis	https://neuromancer-seung-import.appspot.com/#...
1	648518346349533350	3153502	Apical	Basal	False	apical dendrite is rotated off the vertical axis	https://neuromancer-seung-import.appspot.com/#...
2	648518346349533350	2935817	Apical	Basal	False	apical dendrite is rotated off the vertical axis	https://neuromancer-seung-import.appspot.com/#...
3	648518346349533350	3158686	Apical	Basal	False	apical dendrite is rotated off the vertical axis	https://neuromancer-seung-import.appspot.com/#...
4	648518346349539845	3044661	Apical	Apical	True	correct assignment	https://neuromancer-seung-import.appspot.com/#...
5	648518346349539845	2825184	Apical	Basal	False	this dendrite is close to the apical dendrite	https://neuromancer-seung-import.appspot.com/#...
6	648518346349539845	2825064	Apical	Basal	False	this dendrite is close to the apical dendrite	https://neuromancer-seung-import.appspot.com/#...

Summary of Two-Compartment Classifier

- 65% success using most stringent verification (human inspection)
- Normalization – substantially worse due to loss of spatial info
- “Clean compartment” dataset does not improve accuracy
- Hyperparameter optimization slightly reduced accuracy
- “Within 45° cone” metric improved classifier to 71% accuracy
- Removing confounders (mitos that are soma or true unknown dendritic) further increased accuracy to 78%
- A much higher accuracy is needed before continuing to additional classification applications as the assignment errors propagate

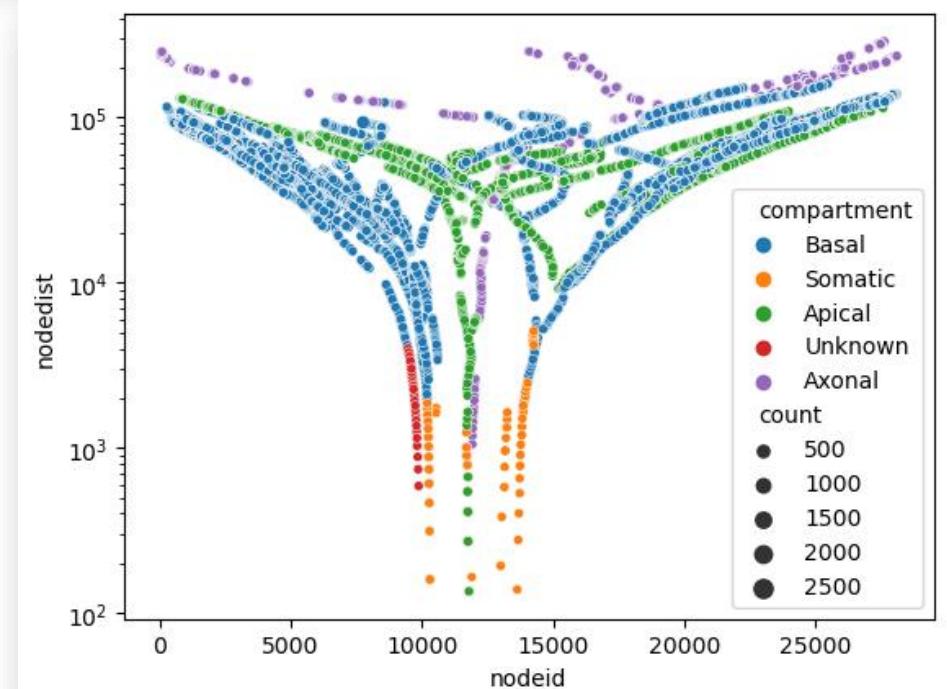
Improving the Pyr neuron mito classifier

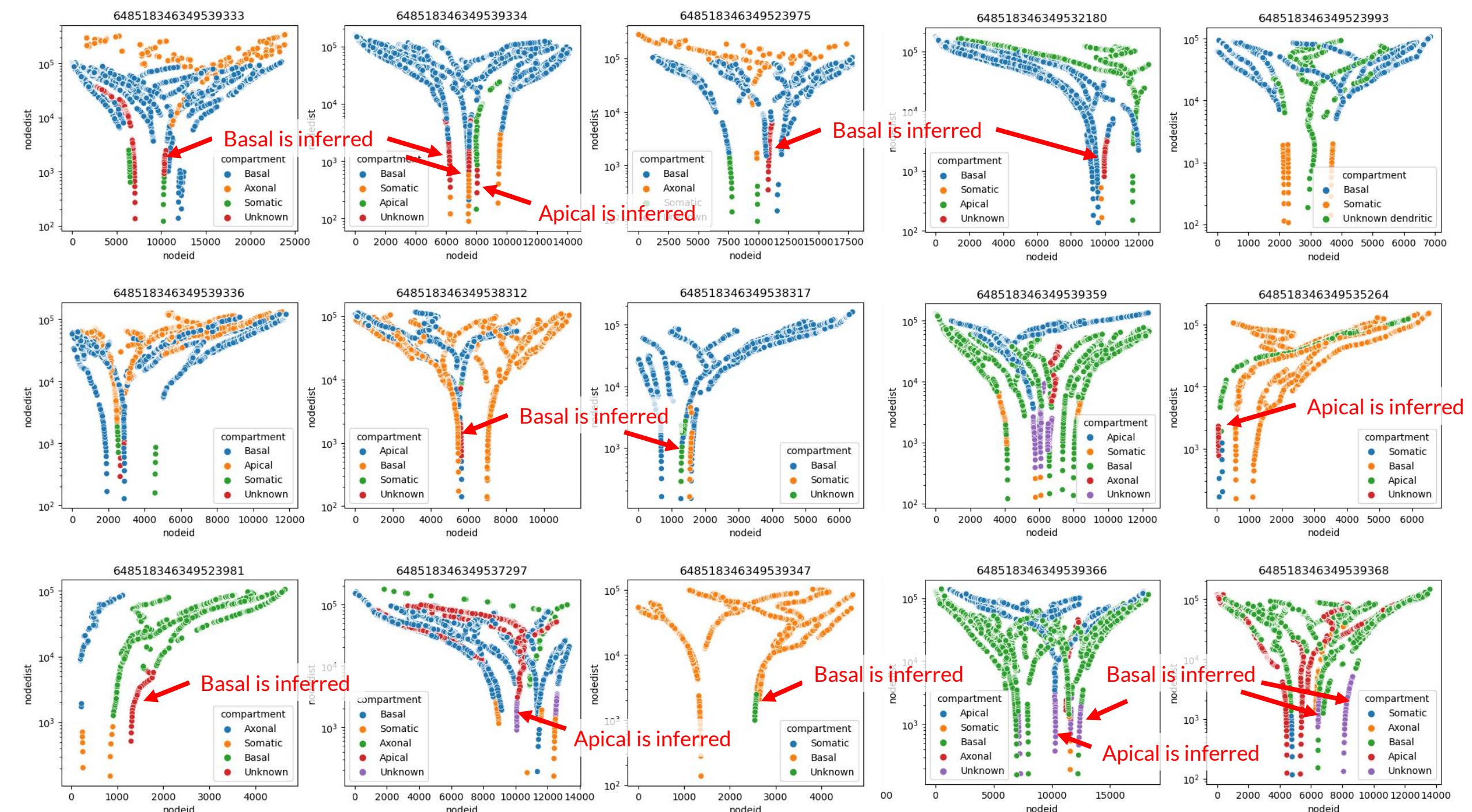
- Node metrics may improve performance as the compartment assignments for unknowns can often be inferred, however, this may need to be done through time-consuming manual inspection
- Metrics to improve axonal assignment are particularly needed; number of nearby synapses is a good candidate as axonal mito have sparse synapses nearby vs. high synaptic density in dendrites

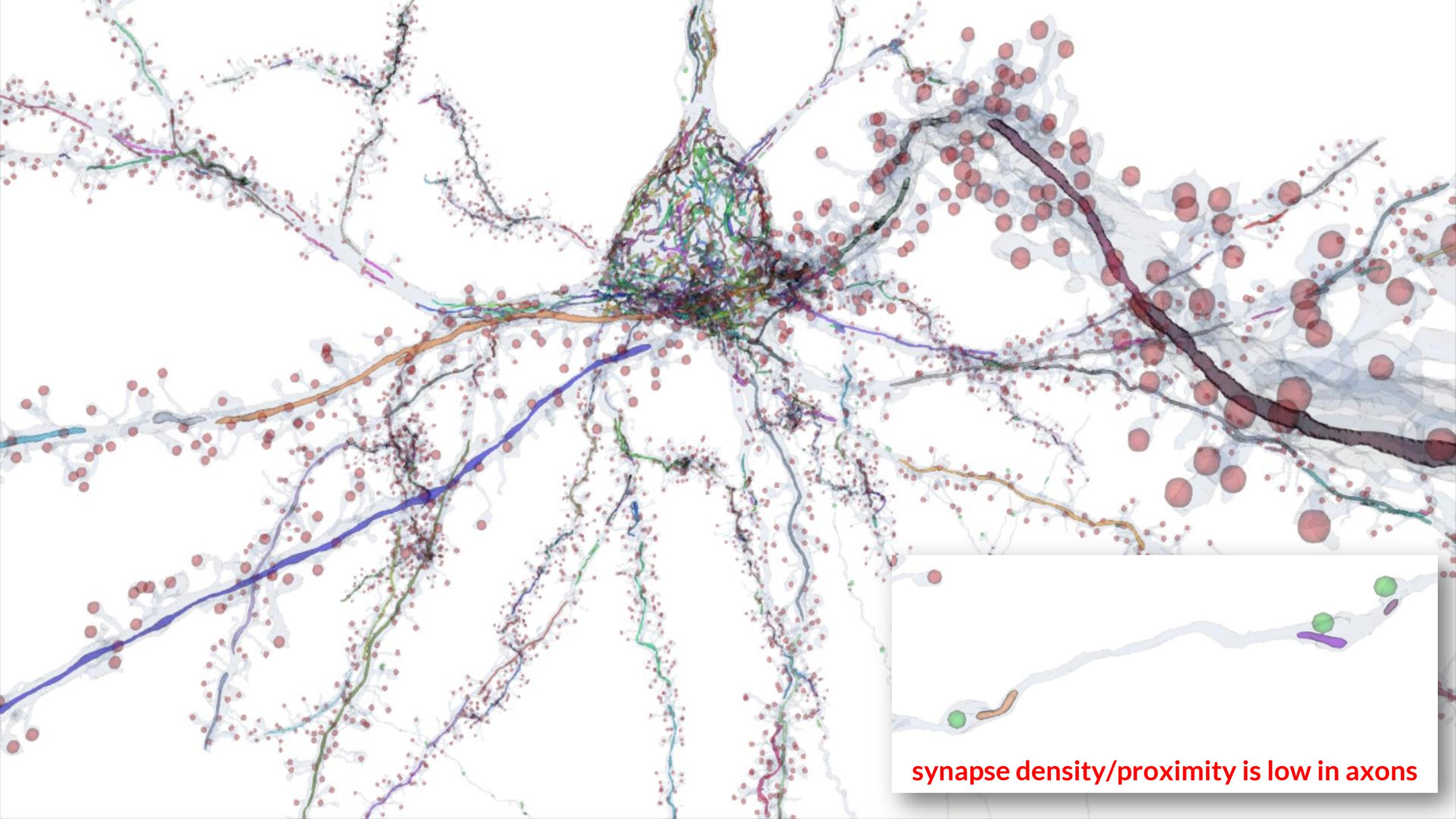


Node metrics

- Each mito has four node metrics (diameter, node count, node distance, node id)
- Each pyr neuron therefore has an array of $(x \times 4)$ node metrics
- Every pyr neuron has a different x number of mitochondria, making incorporation into ML classifier algorithms problematic



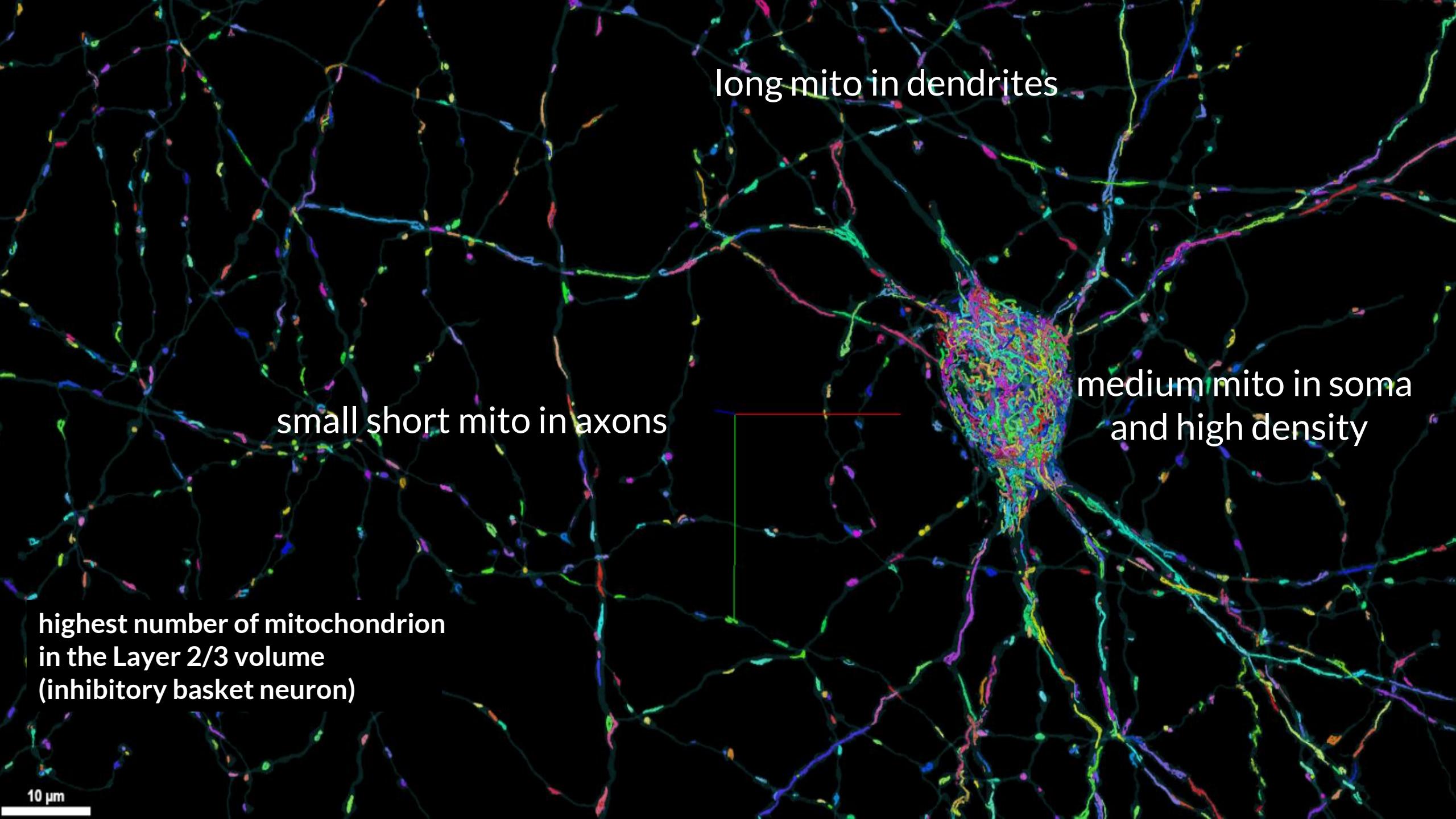




synapse density/proximity is low in axons

Classifier for non- pyramidal neurons

- Can the pyramidal mitochondria classifier be used on the mitos of inhibitory neurons to classify into four compartments?
- Can the mitochondria patterns in inhibitory neurons be used classify subtype (basket, bipolar, etc.)?



highest number of mitochondrion
in the Layer 2/3 volume
(inhibitory basket neuron)

small short mito in axons

long mito in dendrites

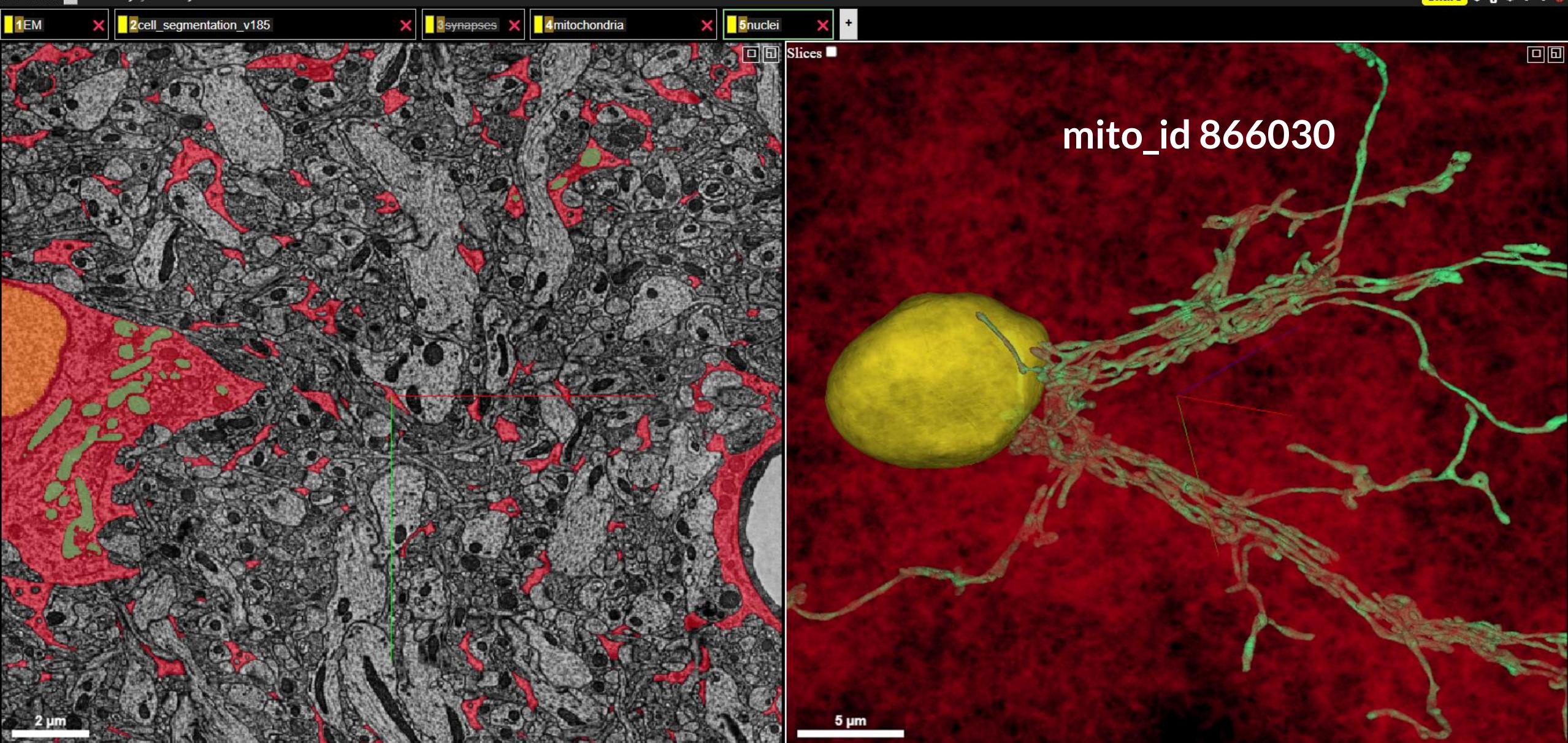
medium mito in soma
and high density

10 μm

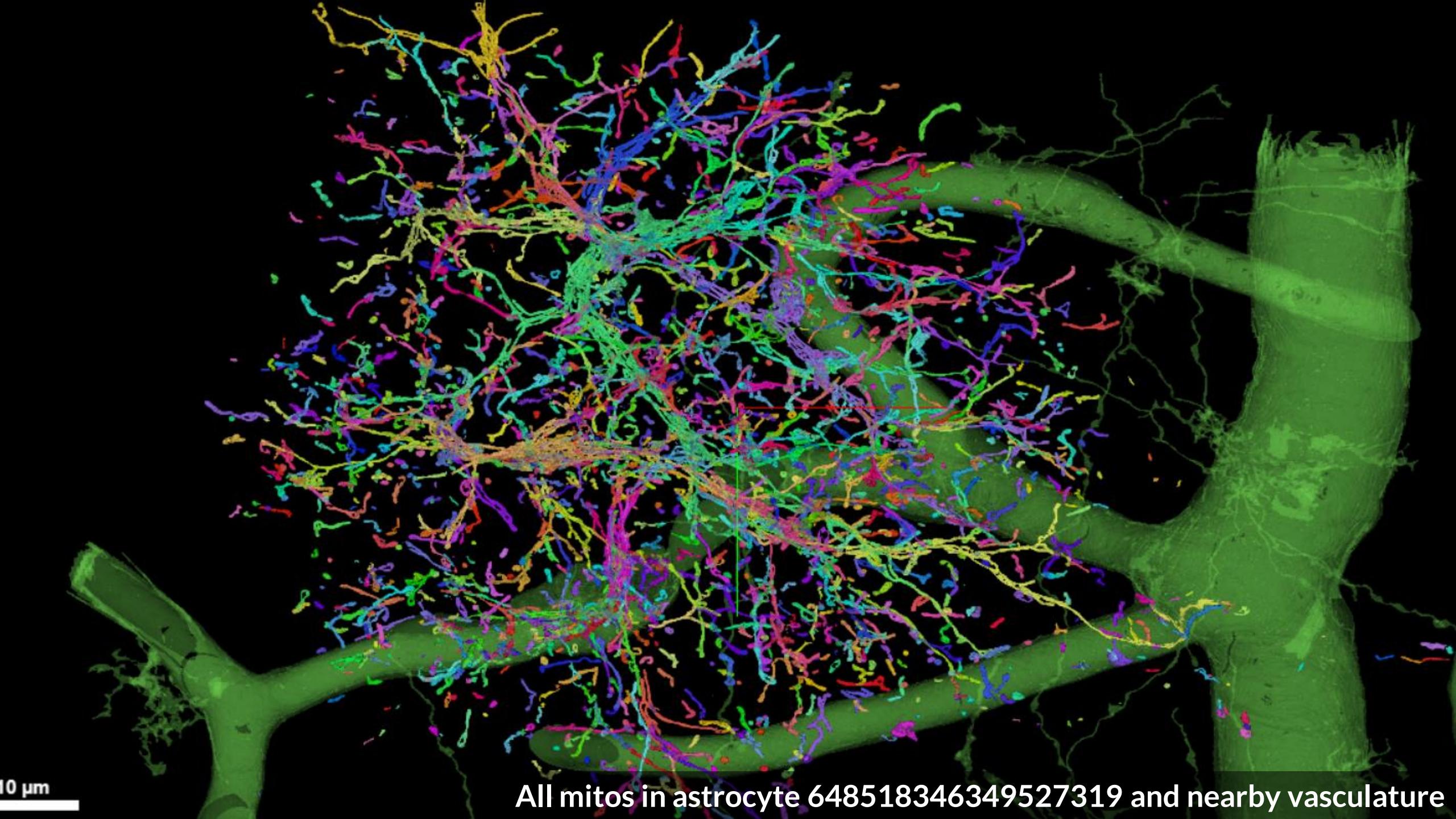
Astrocyte mitochondria

Do features in the volume and functional data (pre and/or post synapses, cell type, calcium spike patterns, proximity to vasculature, etc.) correlate with large mitochondria in astrocytes?

4x4x40 nm³ x 64613, y 40665, z 1446

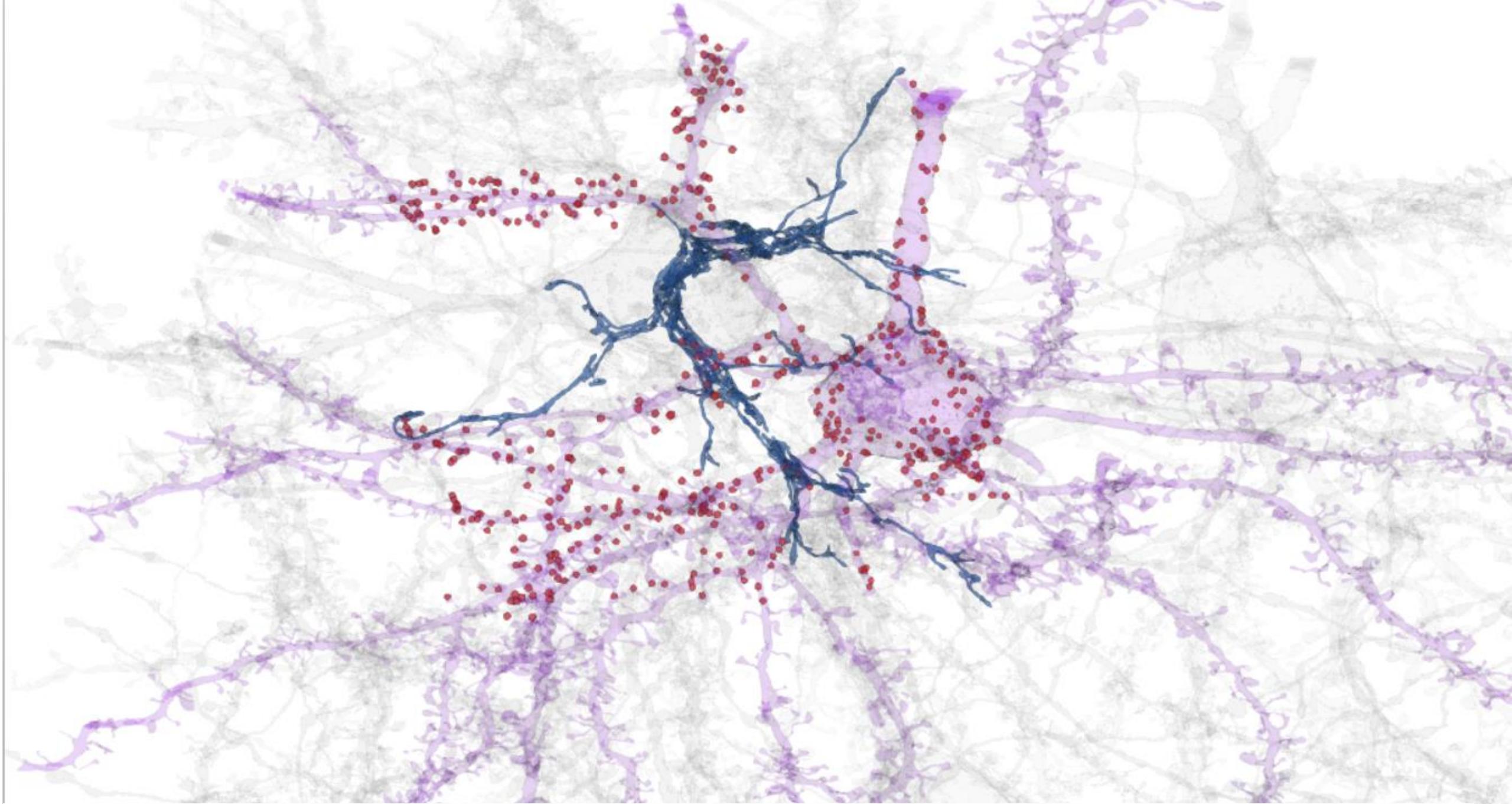


largest continuous mitochondrion in the entire volume (astrocyte 648518346349527319)



10 μm

All mitos in astrocyte 648518346349527319 and nearby vasculature



Synapses from pyramidal neuron 648518346349538100 that are in close proximity to astrocyte mito_id 866030

Thank you

- I thank Wing Wong at Stanford for the opportunity to conduct this sabbatical research in his laboratory. I thank Qiao Liu and Wanwen Zeng (post-doctoral fellows in the Wong lab) for their assistance in Python environment troubleshooting, code debugging, and implementation and interpretation of the machine learning mitochondria classifier
- I thank Nicolas Turner (first author of Cell 2022 paper reporting the mitochondria segmentation primary data from which this analysis was derived) for his invaluable assistance and advice in analyzing the mitochondria data from the Microns Consortium Layer 2/3 volume
- I thank Forrest Collman for creating the tutorial Jupyter notebooks from which the 3d and synapse visualization tools were derived
- I thank Jose Portilla for his Udemy course on Machine Learning from which the Random Forest code was derived.

Micron Layer 2/3 volume



Read the original research papers on
the [Citation](#) page at Allen Institute



Read the [Terms and Conditions](#) page



Use under Creative Commons
by Attribution 4.0 International



Code Availability

Visit my [GitHub repository](#) to view how the images and data in this presentation were generated