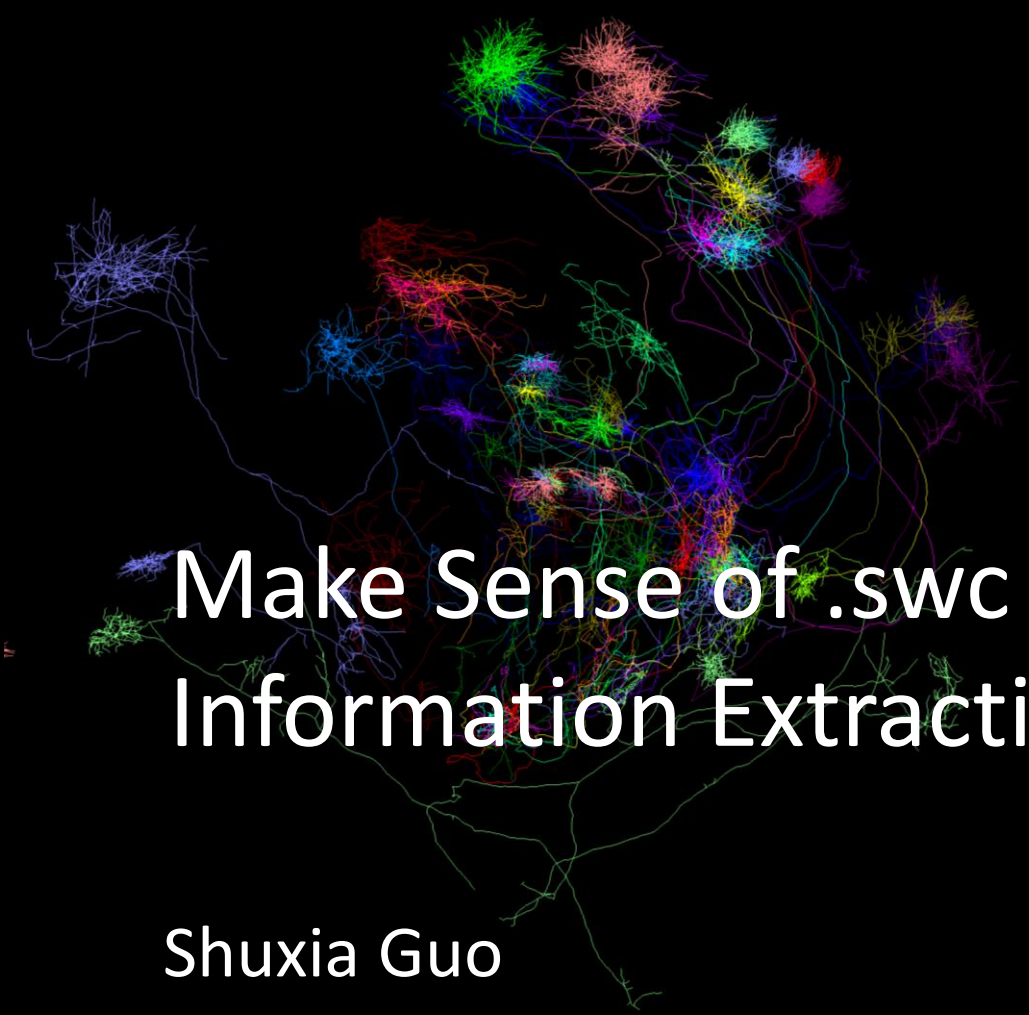


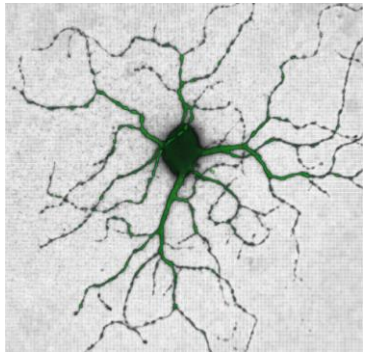
Make Sense of .swc Files: Information Extraction via Vaa3D Platform

Shuxia Guo

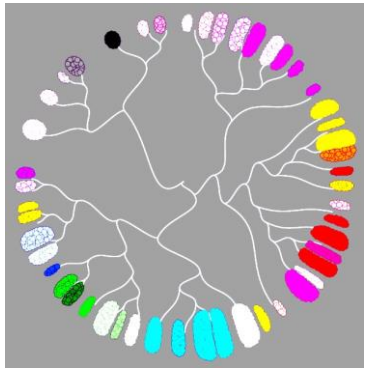
June 28th, 2021

ALLEN INSTITUTE *for* BRAIN SCIENCE

Key Questions in Neuroscience



- **What properties neurons have?**
  (morphology, physiology, gene expression, locations, …)

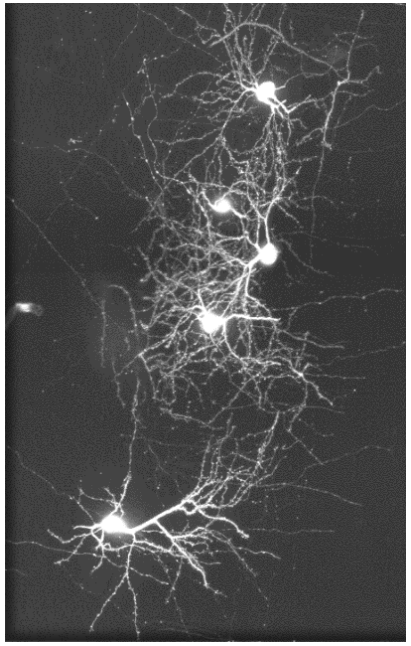


- **Where neurons come from and how they change over time?**
  (lineage, neuron growth/pruning/aging/plasticity,…)

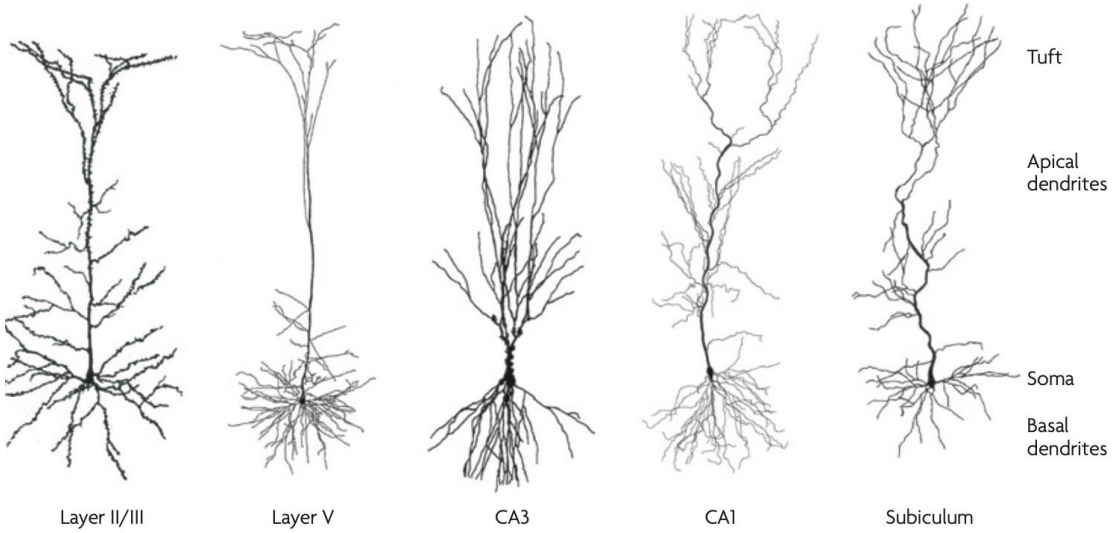
How neurons connect and work together?
(connectivity/connectome/networks, information processing, dynamics, behaviors, …)

pic courtesy of cosmosmagazine.com

Key Questions in Neuroscience



Dendritic morphology



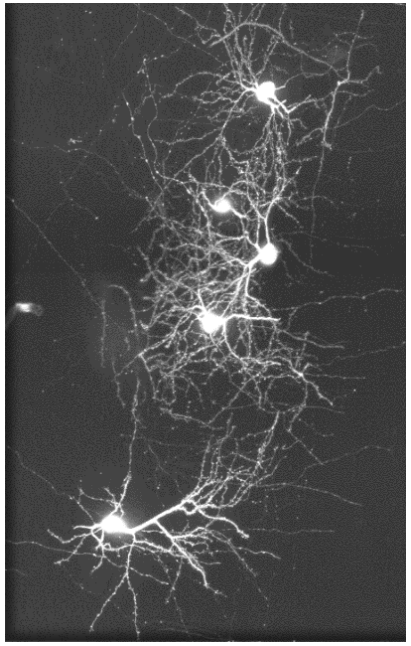
(Spruston, 2008)

.swc files

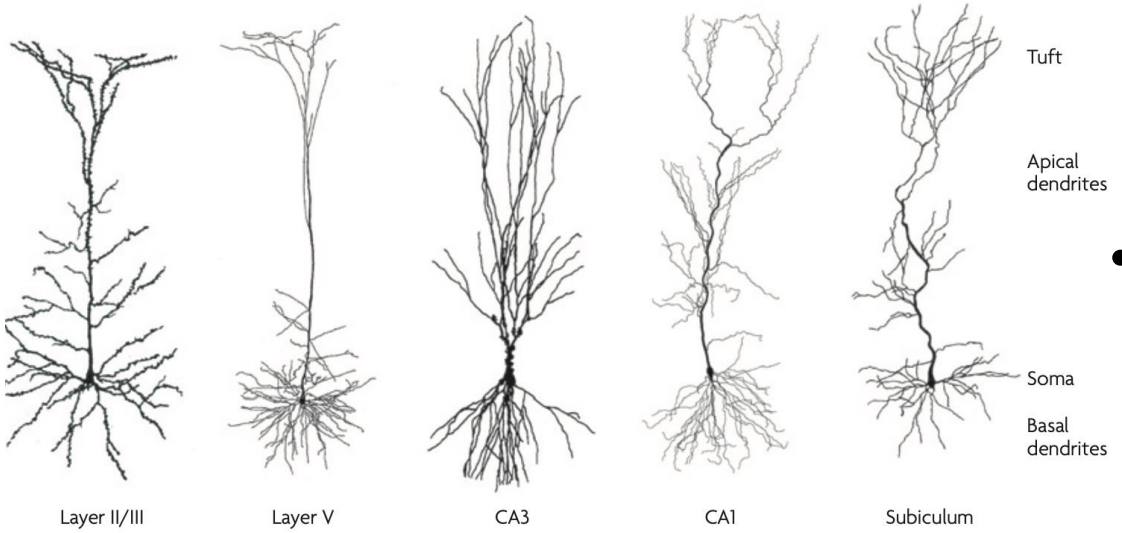
```
##n type x y z r parent
1 1 513.69 514.30 258.00 1.0 -1
2 3 513.50 515.97 258.40 1.0 1
3 3 512.86 520.58 259.52 1.0 2
4 3 512.28 525.18 260.62 1.0 3
5 3 511.70 529.73 261.73 1.0 4
6 3 511.06 534.34 262.83 1.0 5
7 3 511.00 539.52 263.58 1.0 6
8 3 511.00 545.15 264.10 1.0 7
9 3 510.87 549.89 265.01 1.0 8
10 2 509.98 552.90 265.10 1.0 9
11 2 510.04 554.37 264.64 1.0 10
12 2 511.00 559.23 263.78 1.0 11
13 2 510.10 564.03 263.01 1.0 12
14 2 510.10 570.43 262.99 1.0 13
15 2 510.49 576.83 262.99 1.0 14
16 2 512.02 582.98 262.99 1.0 15
17 2 512.09 589.38 262.99 1.0 16
18 2 512.79 595.71 262.99 1.0 17
19 2 514.01 601.92 262.96 1.0 18
20 2 514.14 608.13 262.62 1.0 19
21 2 514.97 612.10 261.44 1.0 20
22 2 515.99 617.86 261.01 1.0 21
23 2 516.06 624.26 260.99 1.0 22
24 2 516.57 630.59 260.96 1.0 23
25 2 518.94 636.29 260.59 1.0 24
26 2 519.96 640.26 259.41 1.0 25
27 2 520.98 646.02 258.99 1.0 26
28 2 522.97 652.10 258.99 1.0 27
29 2 524.25 656.96 258.34 1.0 28
30 2 525.98 662.59 258.00 1.0 29
31 2 527.00 668.80 257.89 1.0 30
32 2 527.00 673.60 257.06 1.0 31
33 2 527.00 678.59 256.26 1.0 32
34 2 526.36 684.10 255.79 1.0 33
35 2 525.08 688.45 254.99 1.0 34
36 2 525.02 693.50 254.27 1.0 35
37 2 524.12 698.11 254.00 1.0 36
38 2 522.90 703.17 253.44 1.0 37
39 2 521.75 708.22 252.88 1.0 38
```

	1	2	4	5	6	7	8
data type	Sample number	Structure Identifier	x position	y position	z position	radius	parent sample
data value	integer value, generally continuous, starting from '1', though this is not required.	Standardized swc files (www.neuromorpho.org) – 0 – undefined 1 – soma 2 – axon 3 – (basal) dendrite 4 – apical dendrite 5+ – custom A lot of data does not conform exactly to this standard however e.g. CNIC data – 0 – undefined 1 – soma 2 – axon 3 – (basal) dendrite 4 – apical dendrite 5 – fork point 6 – end point 7 – custom VNED data – seems to be standard, but uses 10 – related to soma ? Gulyas data – each number represents structure with same diameter. Other data has been observed with -1 – also possibly related to soma ?	'x', 'y', 'z' are spatial co-ordinates, given in micrometers. 'radius' is half the dendrite thickness, also given in micrometers (note this is one of the few formats which use radius instead of diameter)				The sample number. Connectivity is expressed with this value. Parent samples should appear before any child samples.

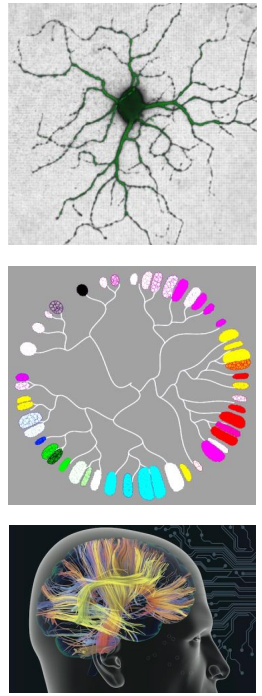
Key Questions in Neuroscience



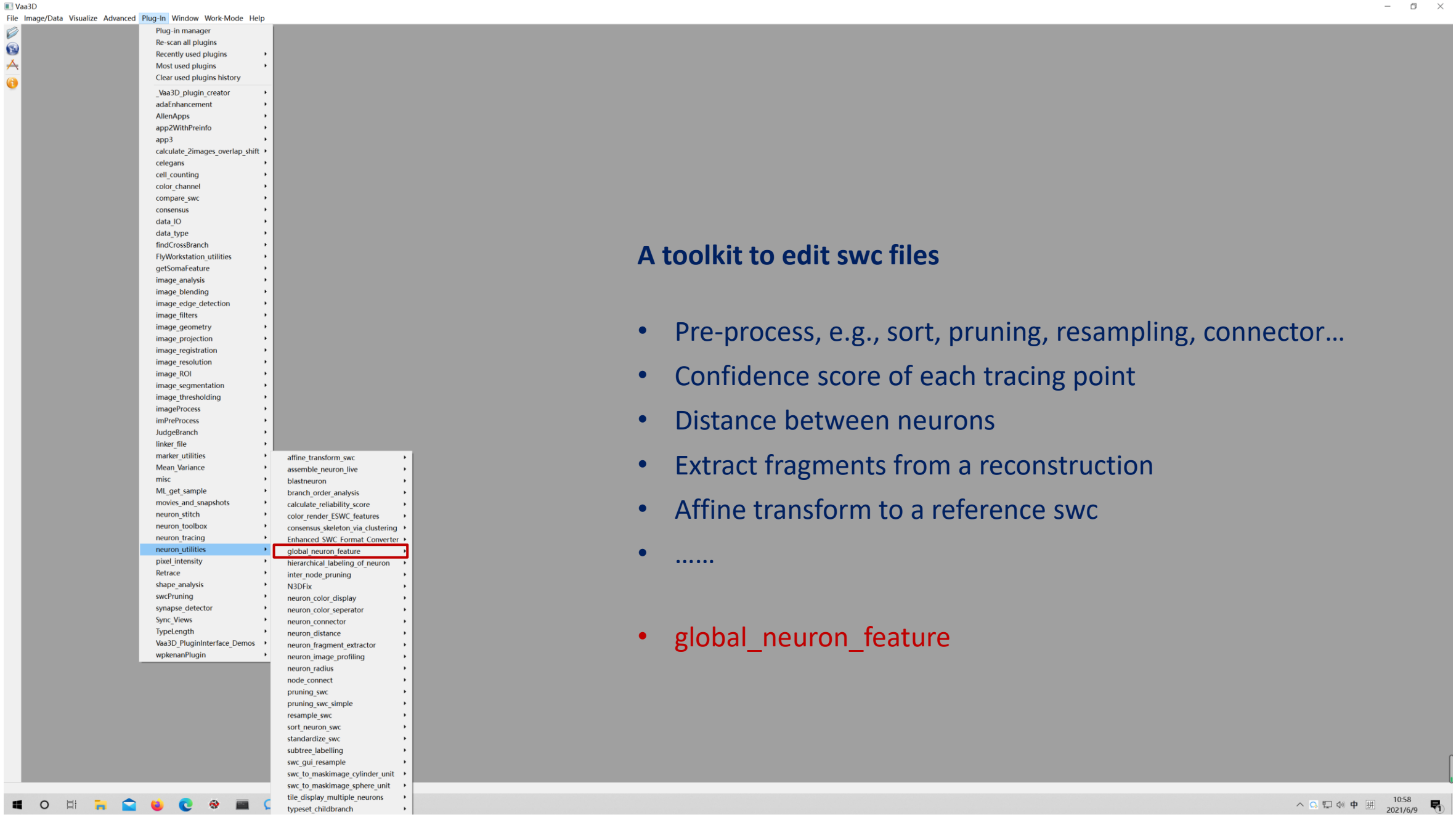
Dendritic morphology



(Spruston, 2008)



- Quantify/characterize neuron morphologies



A toolkit to edit swc files

- Pre-process, e.g., sort, pruning, resampling, connector…
- Confidence score of each tracing point
- Distance between neurons
- Extract fragments from a reconstruction
- Affine transform to a reference swc
- ……

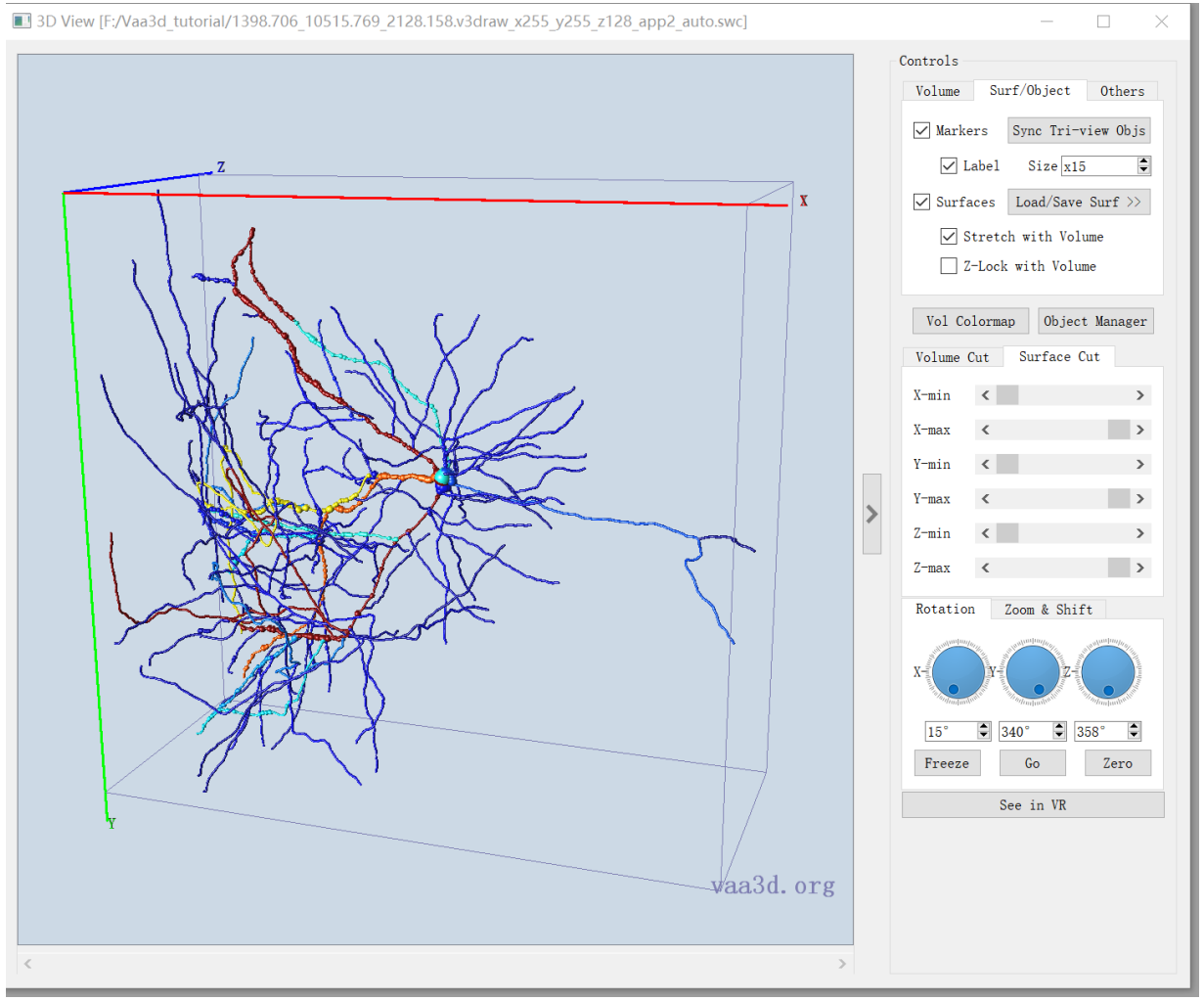
- global_neuron_feature

L-Measure Features

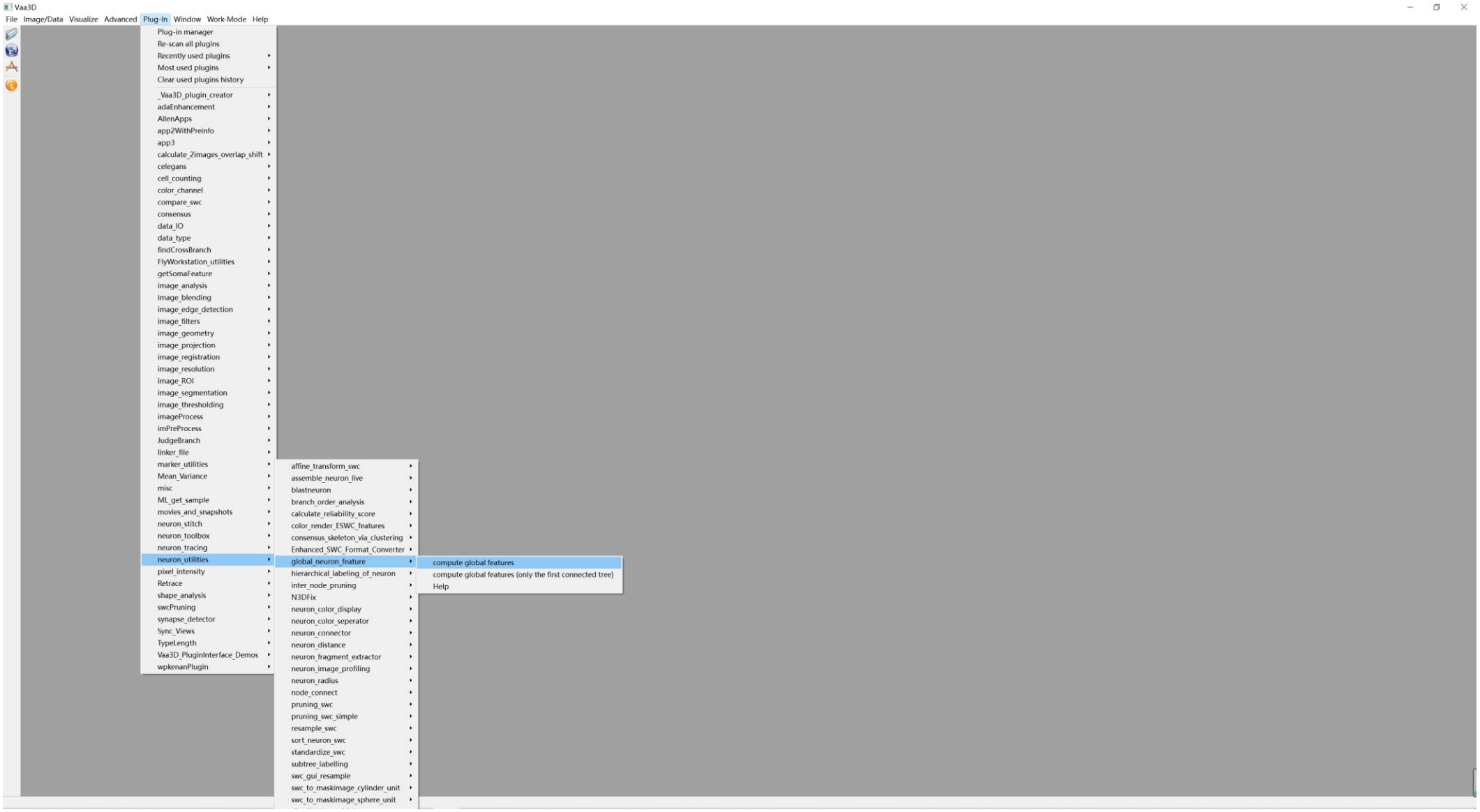
- Number of Nodes
- Soma Surface
- Number of Stems
- Number of Bifurcations
- Number of Branches
- Number of Tips
- Overall Width
- Overall Height
- Overall Depth
- Average Diameter
- Total Length
- Total Surface
- Total Volume
- Max Euclidean Distance
- Max Path Distance
- Max Branch Order
- Average Contraction
- Average Fragmentation
- Average Parent-daughter Ratio
- Average Bifurcation Angle Local
- Average Bifurcation Angle Remote
- Hausdorff Dimension

http://cng.gmu.edu:8080/Lm/help/index.htm

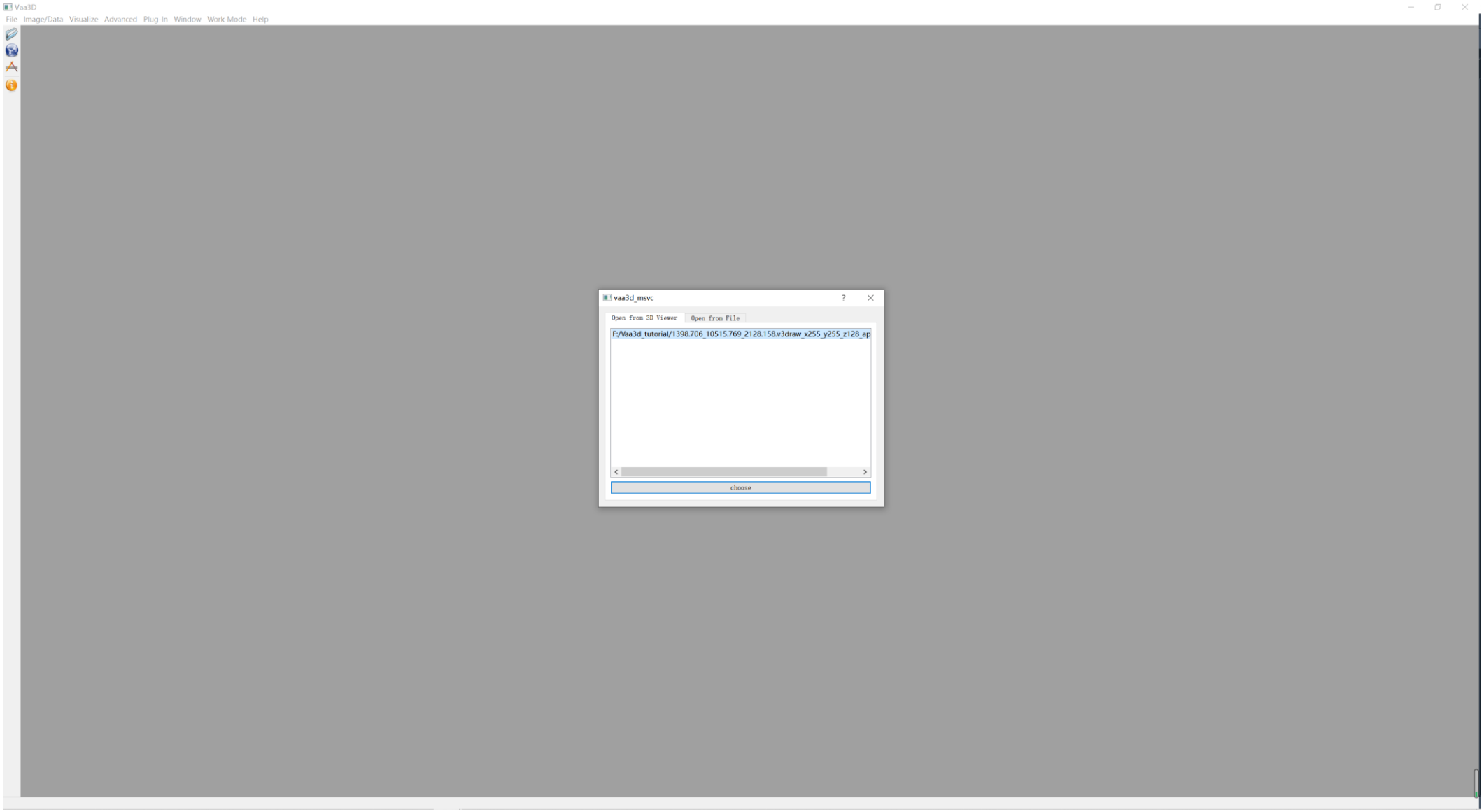
Global Features via GUI



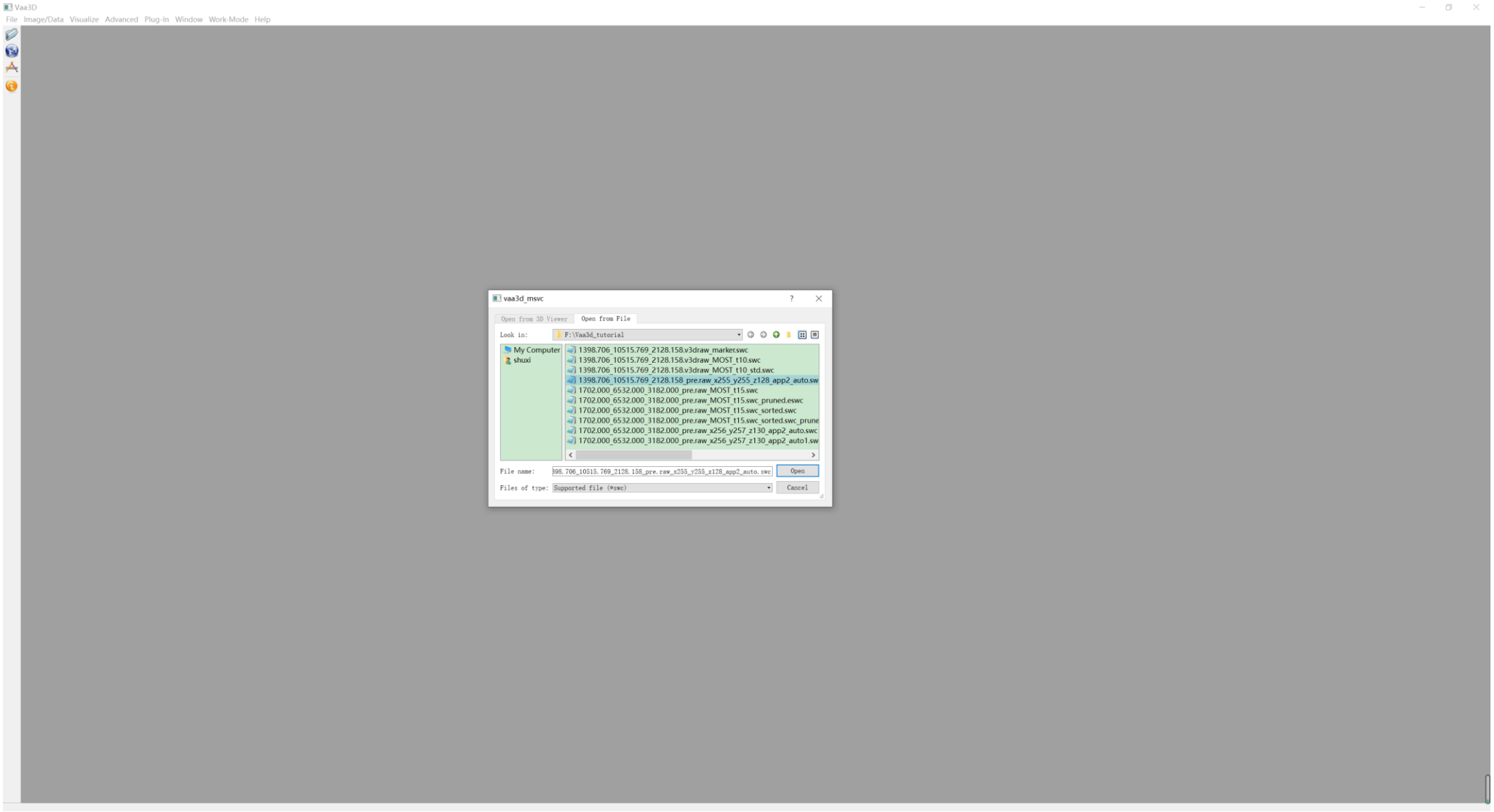
Global Features via GUI



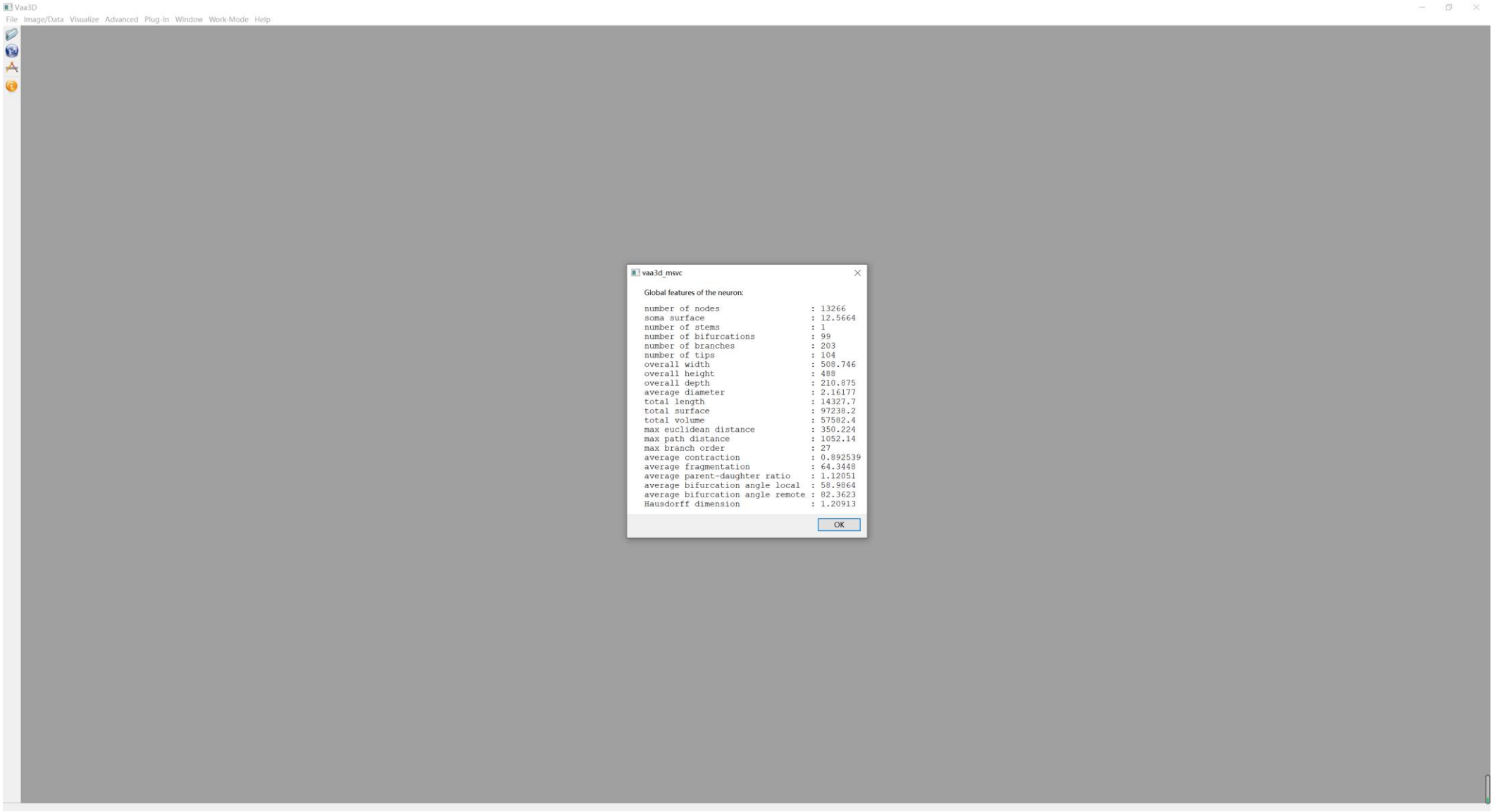
Global Features via GUI



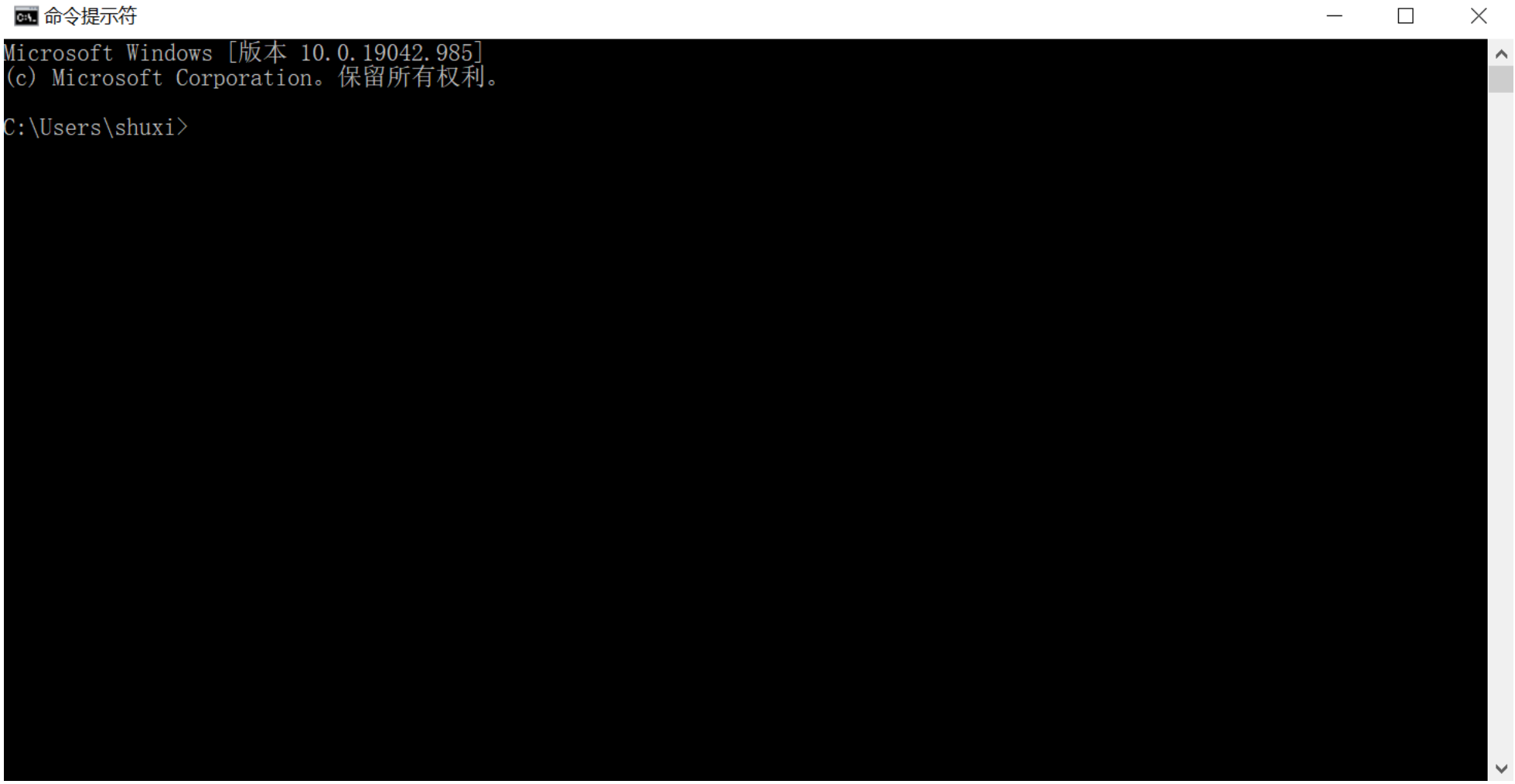
Global Features via GUI



Global Features via GUI



Run Functions Via Command Line

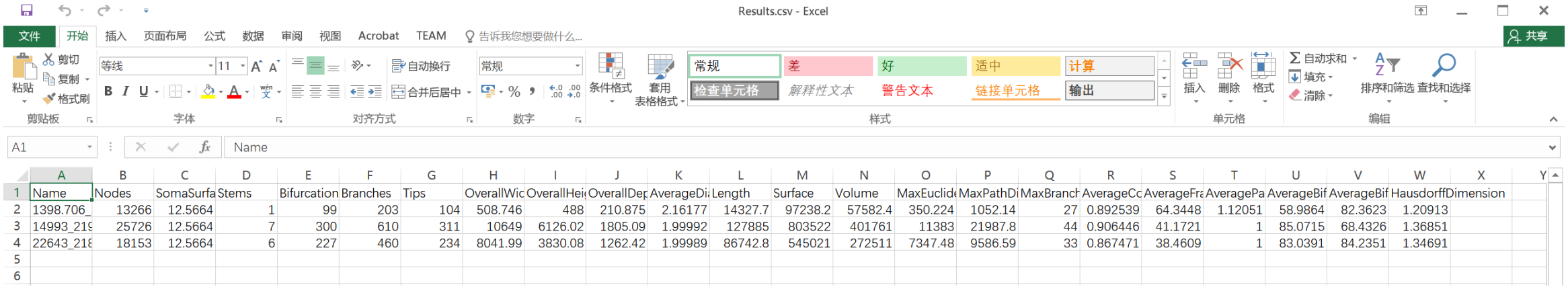


Global Features Via Command Line

v3d software → plugin name → function name → path of swc files → where to save

W v3d /x global_neuron_feature /f compute_feature_in_folder /i path_of_swc_folder /o csv_to_save

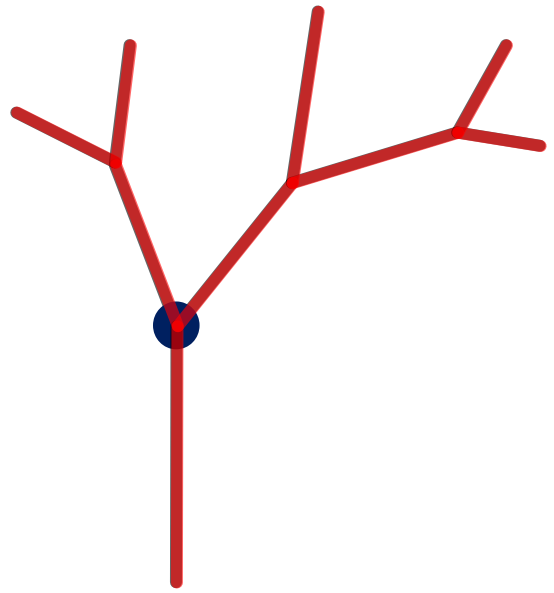
L/M v3d –x global_neuron_feature –f compute_feature_in_folder –i path_of_swc_folder –o csv_to_save



	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y
1	Name	Nodes	SomaSurfa	Stems	Bifurcation	Branches	Tips	OverallWic	OverallHei	OverallDep	AverageDi	Length	Surface	Volume	MaxEuclide	MaxPathDi	MaxBranch	AverageCc	AverageFra	AveragePa	AverageBif	AverageBif	HausdorffDimension		
2	1398.706_	13266	12.5664	1	99	203	104	508.746	488	210.875	2.16177	14327.7	97238.2	57582.4	350.224	1052.14	27	0.892539	64.3448	1.12051	58.9864	82.3623	1.20913		
3	14993_219	25726	12.5664	7	300	610	311	10649	6126.02	1805.09	1.99992	127885	803522	401761	11383	21987.8	44	0.906446	41.1721	1	85.0715	68.4326	1.36851		
4	22643_218	18153	12.5664	6	227	460	234	8041.99	3830.08	1262.42	1.99989	86742.8	545021	272511	7347.48	9586.59	33	0.867471	38.4609	1	83.0391	84.2351	1.34691		
5																									
6																									

.swc files

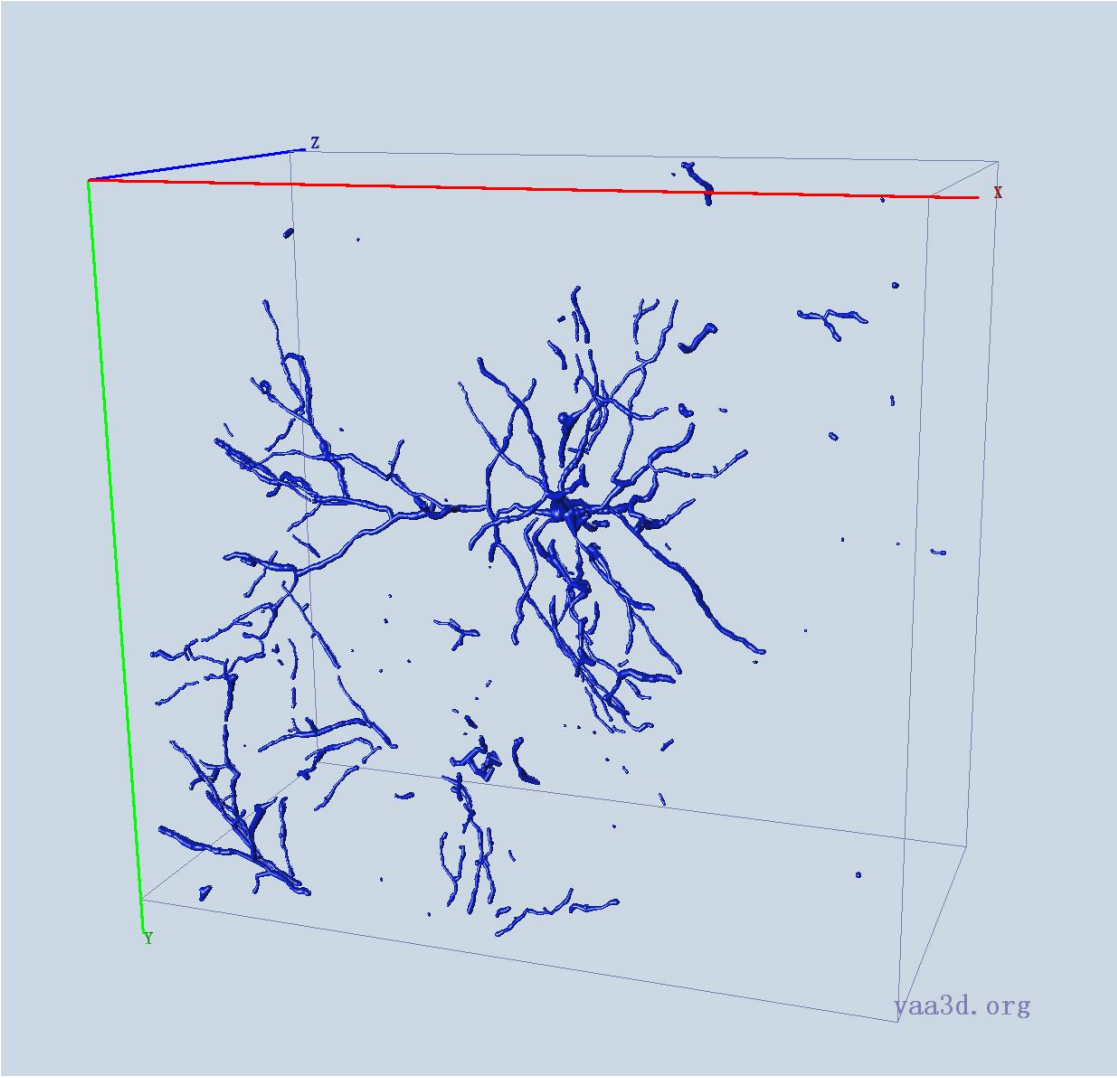
```
##n type x y z r parent
1 1 513.69 514.30 258.00 1.0 -1
2 3 513.50 515.97 258.40 1.0 1
3 3 512.86 520.58 259.52 1.0 2
4 3 512.28 525.18 260.62 1.0 3
5 3 511.70 529.73 261.73 1.0 4
6 3 511.06 534.34 262.83 1.0 5
7 3 511.00 539.52 263.58 1.0 6
8 3 511.00 545.15 264.10 1.0 7
9 3 510.87 549.89 265.01 1.0 8
10 2 509.98 552.90 265.10 1.0 9
11 2 510.04 554.37 264.64 1.0 10
12 2 511.00 559.23 263.78 1.0 11
13 2 510.10 564.03 263.01 1.0 12
14 2 510.10 570.43 262.99 1.0 13
15 2 510.49 576.83 262.99 1.0 14
16 2 512.02 582.98 262.99 1.0 15
17 2 512.09 589.38 262.99 1.0 16
18 2 512.79 595.71 262.99 1.0 17
19 2 514.01 601.92 262.96 1.0 18
20 2 514.14 608.13 262.62 1.0 19
21 2 514.97 612.10 261.44 1.0 20
22 2 515.99 617.86 261.01 1.0 21
23 2 516.06 624.26 260.99 1.0 22
24 2 516.57 630.59 260.96 1.0 23
25 2 518.94 636.29 260.59 1.0 24
26 2 519.96 640.26 259.41 1.0 25
27 2 520.98 646.02 258.99 1.0 26
28 2 522.97 652.10 258.99 1.0 27
29 2 524.25 656.96 258.34 1.0 28
30 2 525.98 662.59 258.00 1.0 29
31 2 527.00 668.80 257.89 1.0 30
32 2 527.00 673.60 257.06 1.0 31
33 2 527.00 678.59 256.26 1.0 32
34 2 526.36 684.10 255.79 1.0 33
35 2 525.08 688.45 254.99 1.0 34
36 2 525.02 693.50 254.27 1.0 35
37 2 524.12 698.11 254.00 1.0 36
38 2 522.90 703.17 253.44 1.0 37
39 2 521.75 708.22 252.88 1.0 38
```



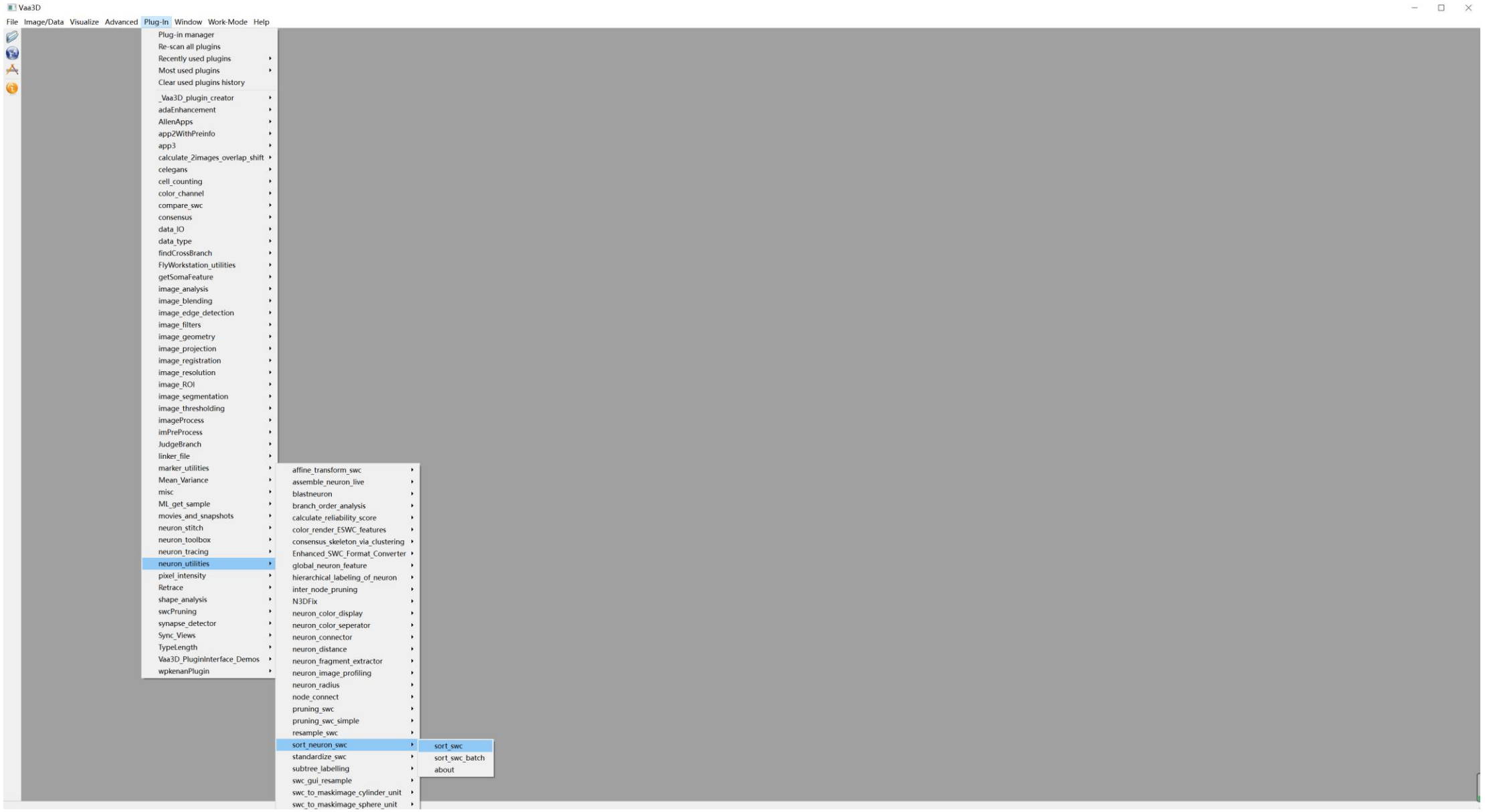
Pre-process .swc Files

- not saved in the order of root-branches
- more than one roots
- not well-connected tree (i.e., there are gaps)
- small branches

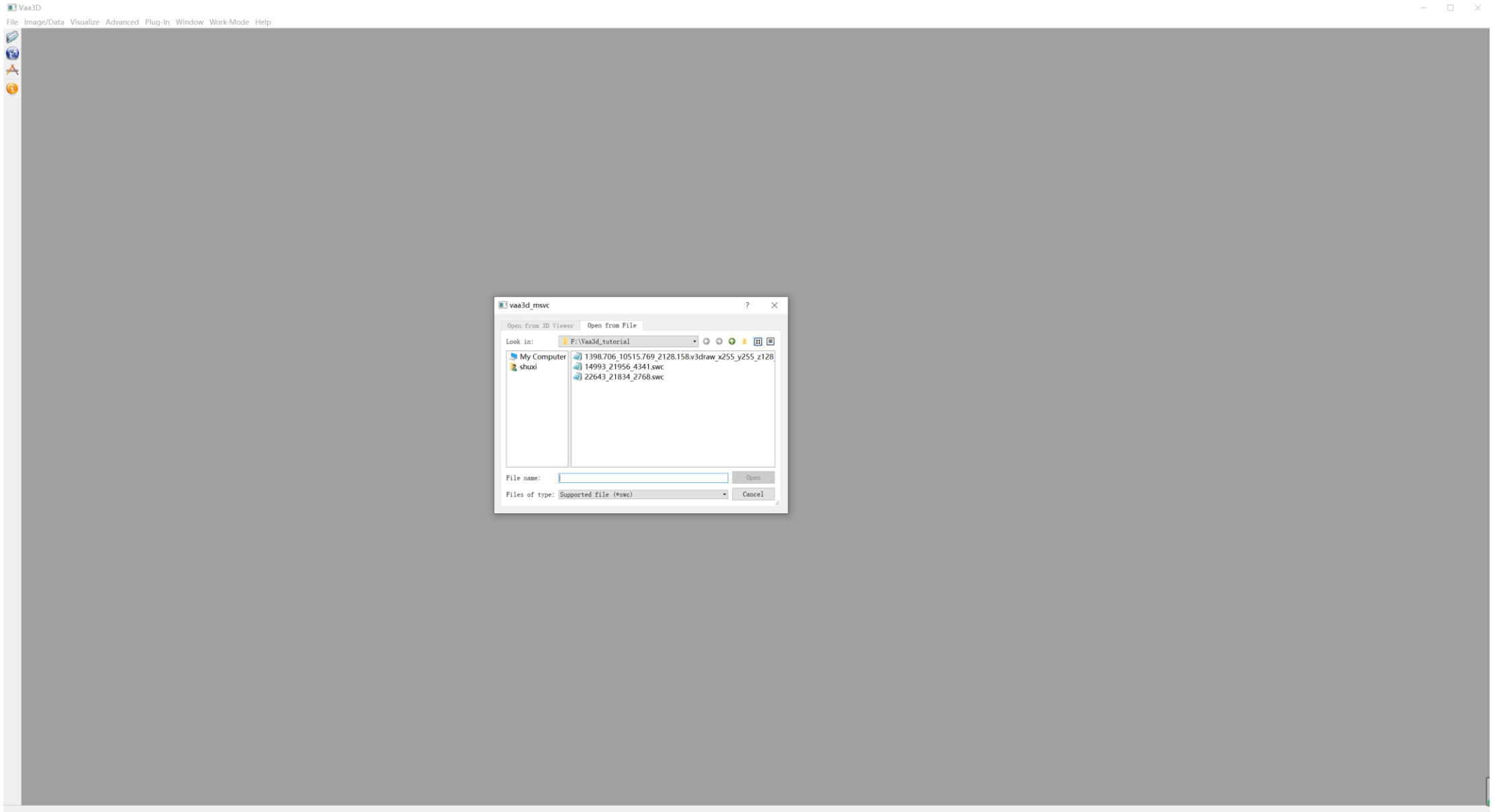
- sort → prune



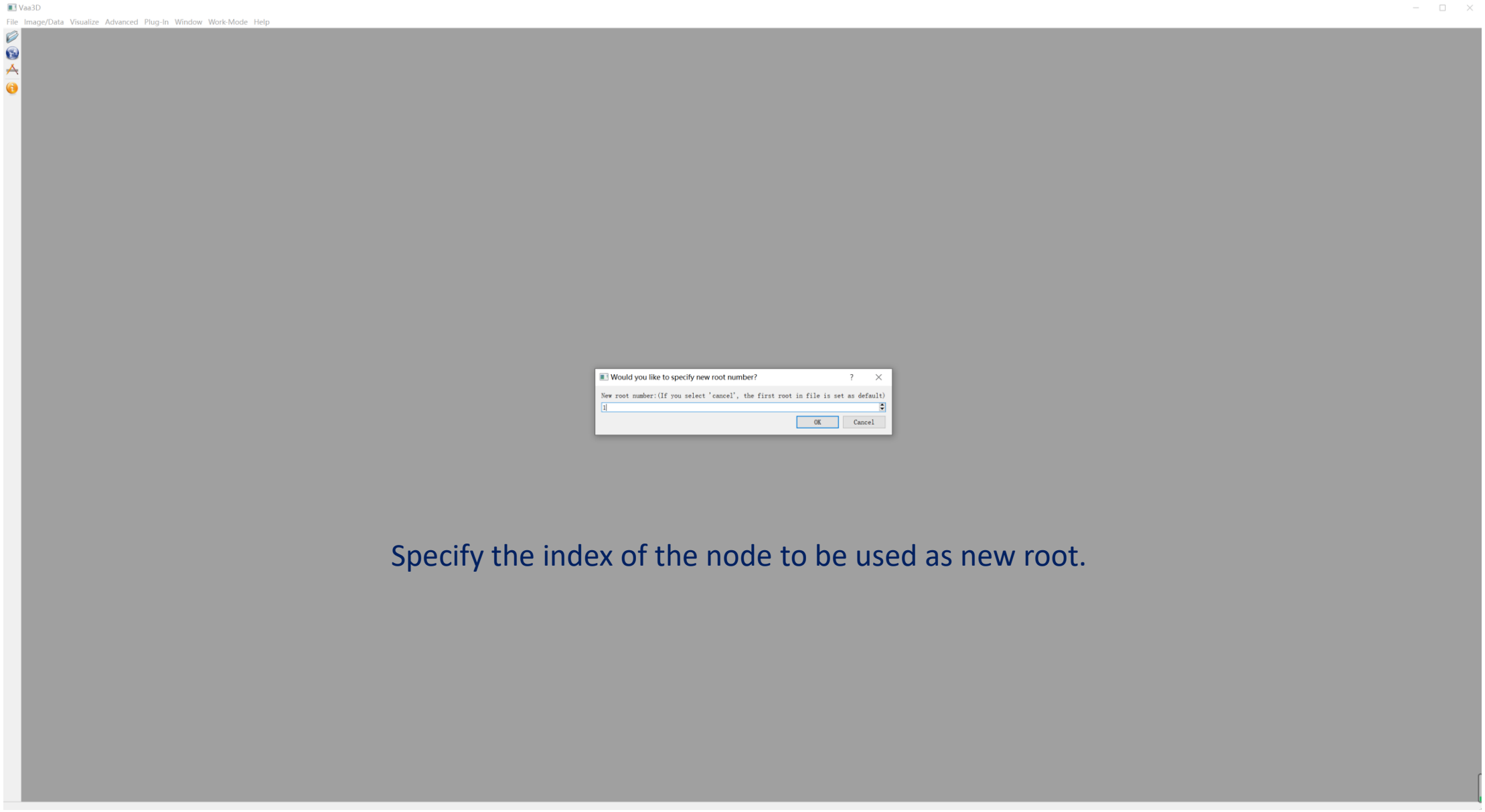
Sort swc via GUI



Sort swc via GUI

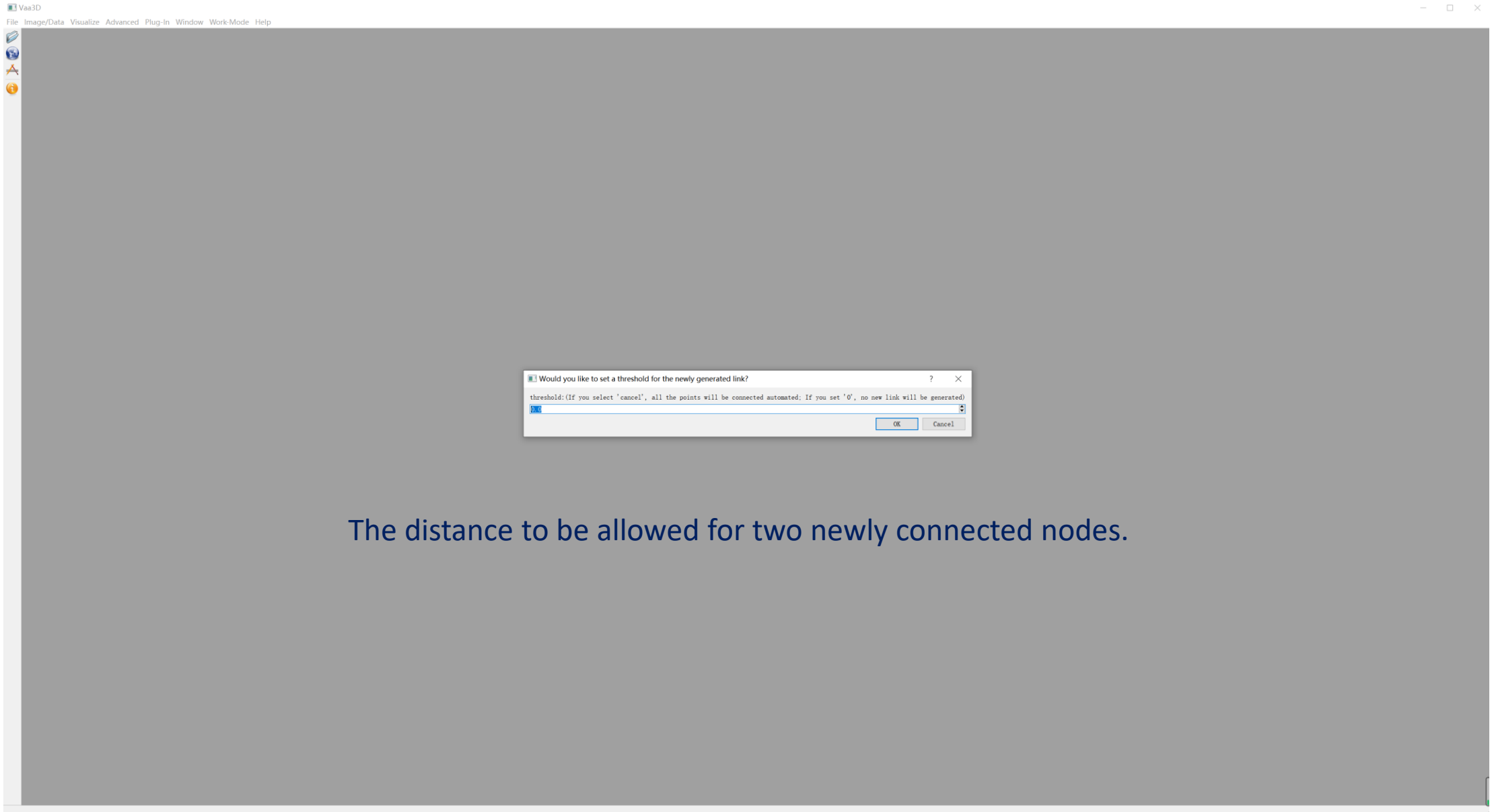


Sort swc via GUI



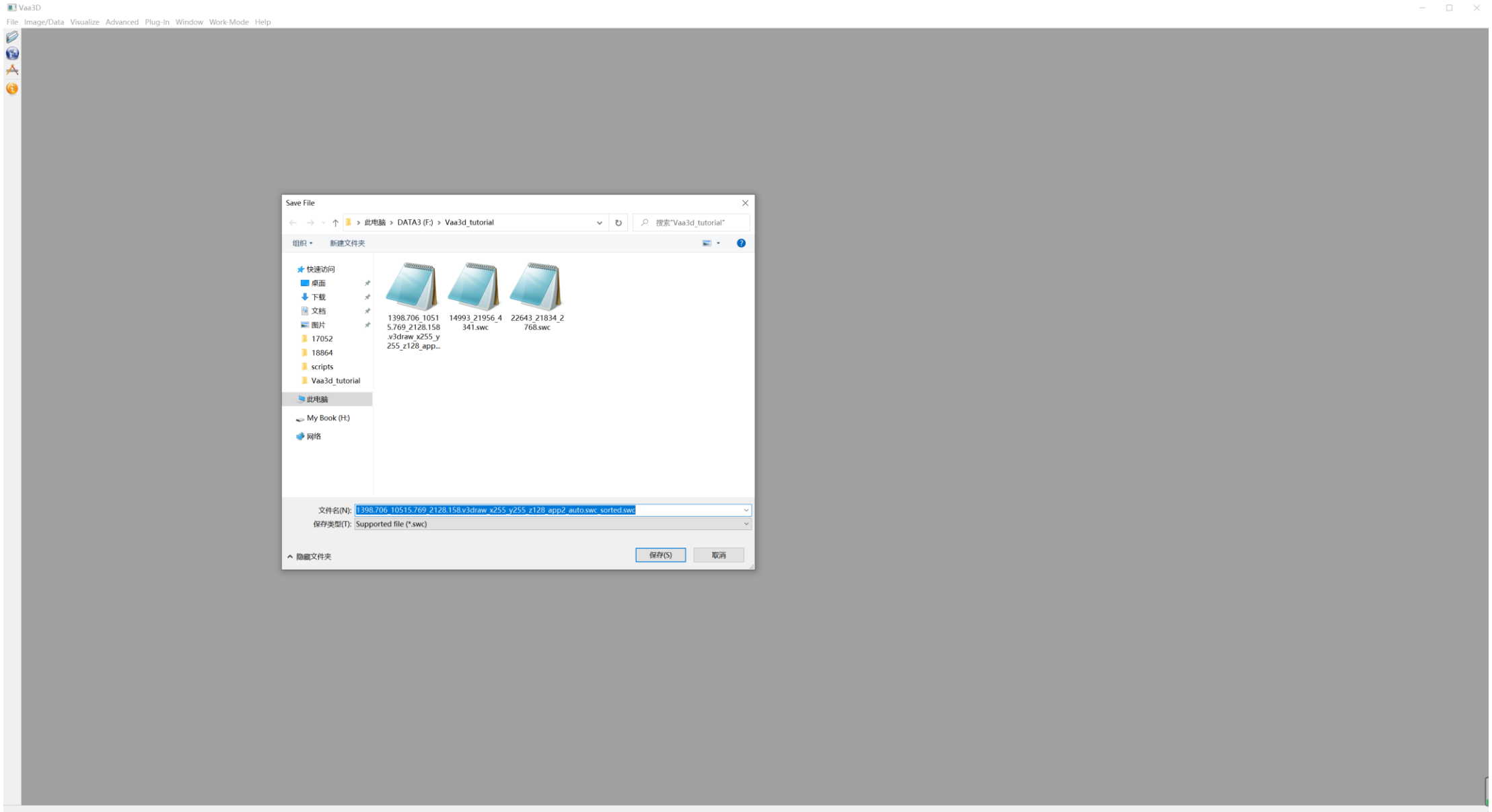
Specify the index of the node to be used as new root.

Sort swc via GUI



The distance to be allowed for two newly connected nodes.

Sort swc via GUI



Example Result

1398.706_10515.769_2128.158.v3draw_x255_y255_z128_app2_auto.swc - 记事本

文件(F) 编辑(E) 格式(O) 查看(V) 帮助(H)

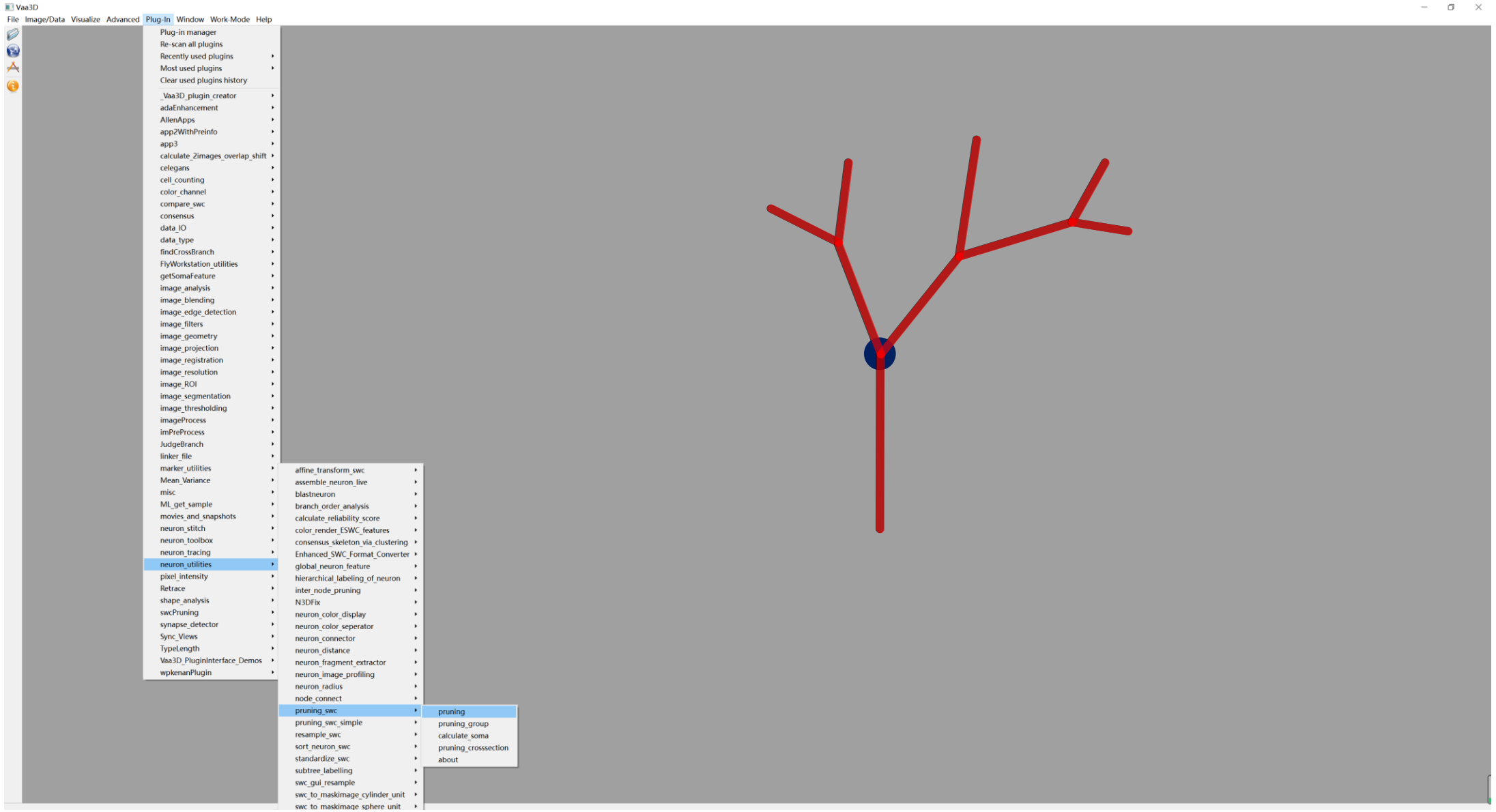
```
##n,type,x,y,z,radius,parent
1 274 20 287 0 1 2
2 274 20 288.111 0 1 3
3 274 20 289.025 0 1 4
4 274 20 290.018 0 1 5
5 274 20 291.007 0 1 6
6 274 20 292.003 0 1 7
7 274 20 293.002 0.111111 1 8
8 274 20 294.001 0.358025 1 9
9 274 19.8889 295 0.758573 1 10
10 274 19.642 296 0.875019 1 11
11 274 19.2414 297 0.945401 1 12
12 274 19.0139 297.889 1.08509 1 13
13 274 18.6966 298.642 1.34618 1 14
14 274 18.2674 299.241 1.75305 1 15
15 274 18.1368 300.125 1.87247 1 16
16 274 18.0601 301.055 1.94422 1 17
17 274 17.9175 302.026 1.97344 1 18
18 274 17.655 303.012 1.9879 1 19
19 274 17.2475 304.006 1.99436 1 20
20 274 17.1278 305.003 1.9974 1 21
21 274 17.0559 306.001 1.9988 1 22
22 274 17.0266 307.001 1.99944 1 23
23 274 17.0121 308 2.11085 1 24
24 274 17.0057 309 2.35791 1 25
25 274 17.0026 310 2.75852 1 26
26 274 17.0012 311 2.87499 1 27
27 274 17.0006 312 2.94539 1 28
28 274 17.0003 313 2.97397 1 29
29 274 17.0001 314 2.98815 1 30
30 274 17.0001 315 2.99447 1 31
31 274 17.1111 316 2.99746 1 32
32 274 17.358 317 2.99882 1 33
```

1398.706_10515.769_2128.158.v3draw_x255_y255_z128_app2_auto.swc_sorted.swc - 记事本

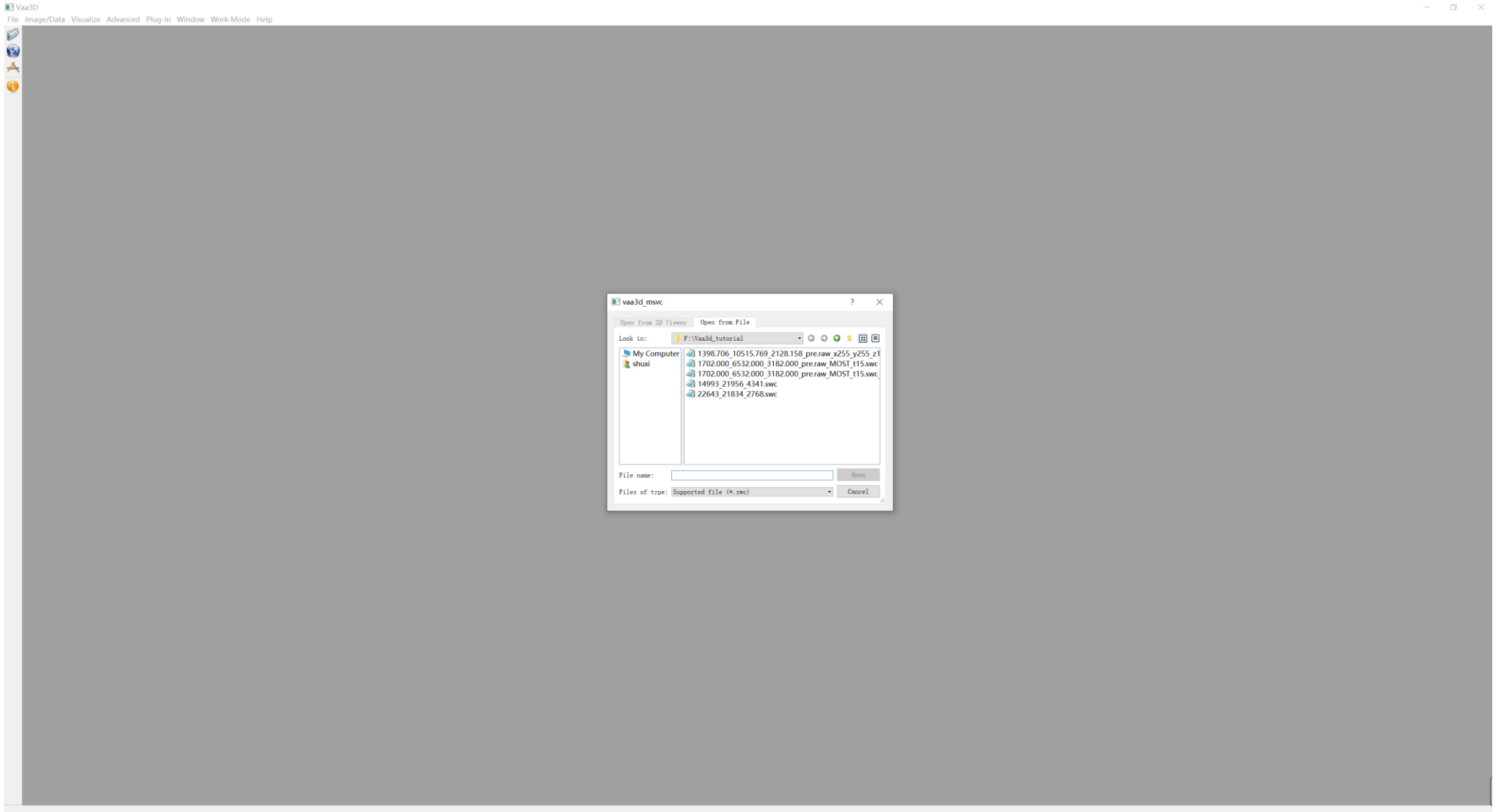
文件(F) 编辑(E) 格式(O) 查看(V) 帮助(H)

```
# id,type,x,y,z,r,pid
1 274 255 255 128 1 -1
2 274 254.553 255.103 127.358 7 1
3 274 253.937 255.228 127.111 6 2
4 274 253.103 255.469 127 1 3
5 274 252.229 256.111 127 1 4
6 274 251.472 257 127 1 5
7 274 251.116 258 127 1 6
8 274 251.011 259 127 1 7
9 274 251.025 260 127 1 8
10 274 251.052 261 126.999 1 9
11 274 251.119 262 126.999 1 10
12 274 251.228 263 126.997 1 11
13 274 251.613 264 126.994 1 12
14 274 251.828 265 126.988 1 13
15 274 251.861 266 126.974 1 14
16 274 251.728 267 126.945 1 15
17 274 251.295 268 126.875 2 16
18 274 250.967 269 126.758 1 17
19 274 250.717 270 126.357 1 18
20 274 250.269 271 126.11 2 19
21 274 249.913 272 125.997 1 20
22 274 249.594 273 125.994 1 21
23 274 249.032 274 125.988 1 22
24 274 248.28 275 125.974 1 23
25 274 247.725 275.999 125.945 1 24
26 274 247.074 276.999 125.875 1 25
27 274 246.379 277.997 125.759 1 26
28 274 245.91 278.994 125.358 1 27
29 274 245.587 279.988 125.111 1 28
30 274 245.016 280.974 125 1 29
31 274 244.247 281.945 125 1 30
32 274 243.654 282.875 125 1 31
```

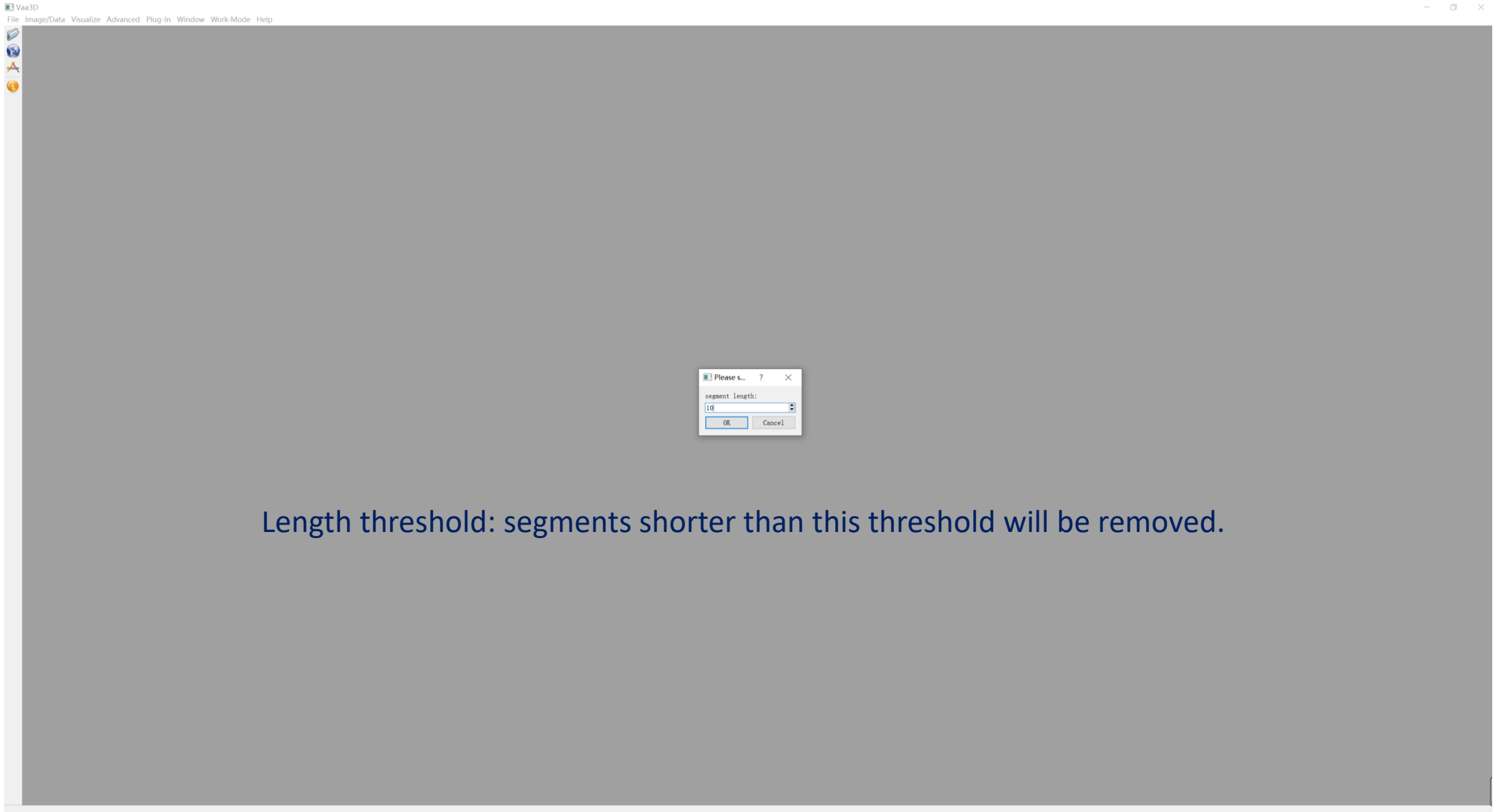
prun .swc files



prun .swc files

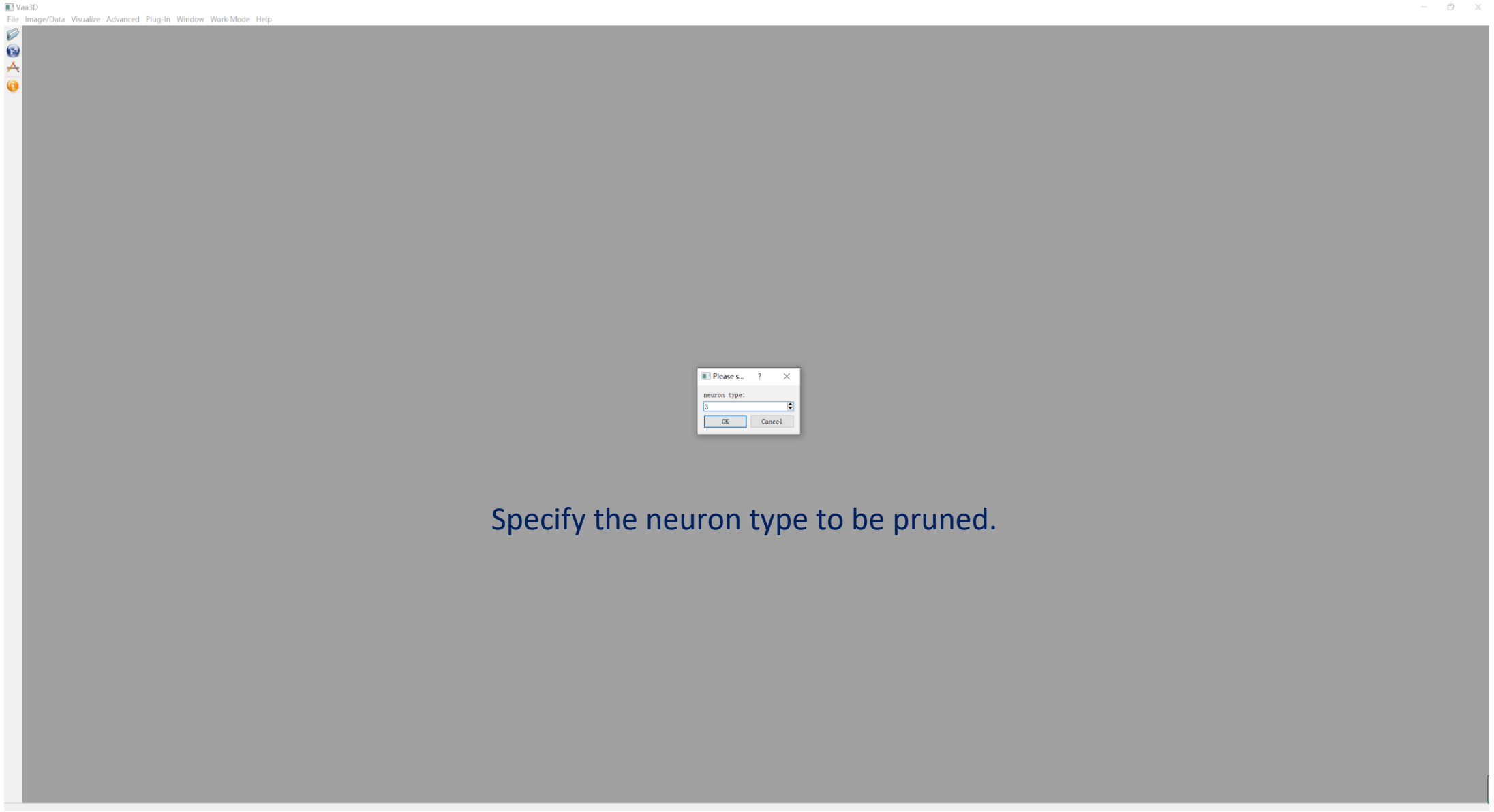


prun .swc files

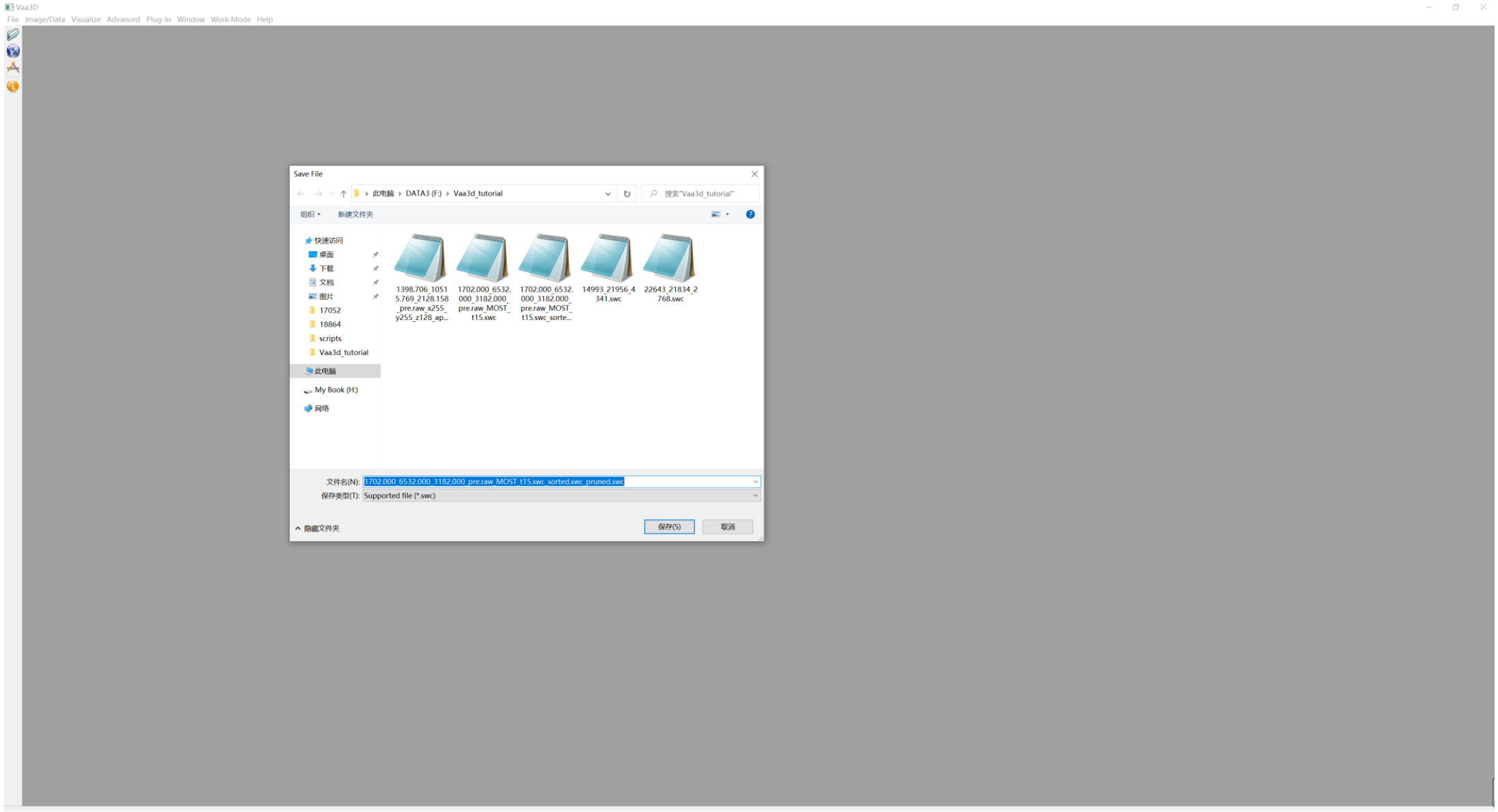


Length threshold: segments shorter than this threshold will be removed.

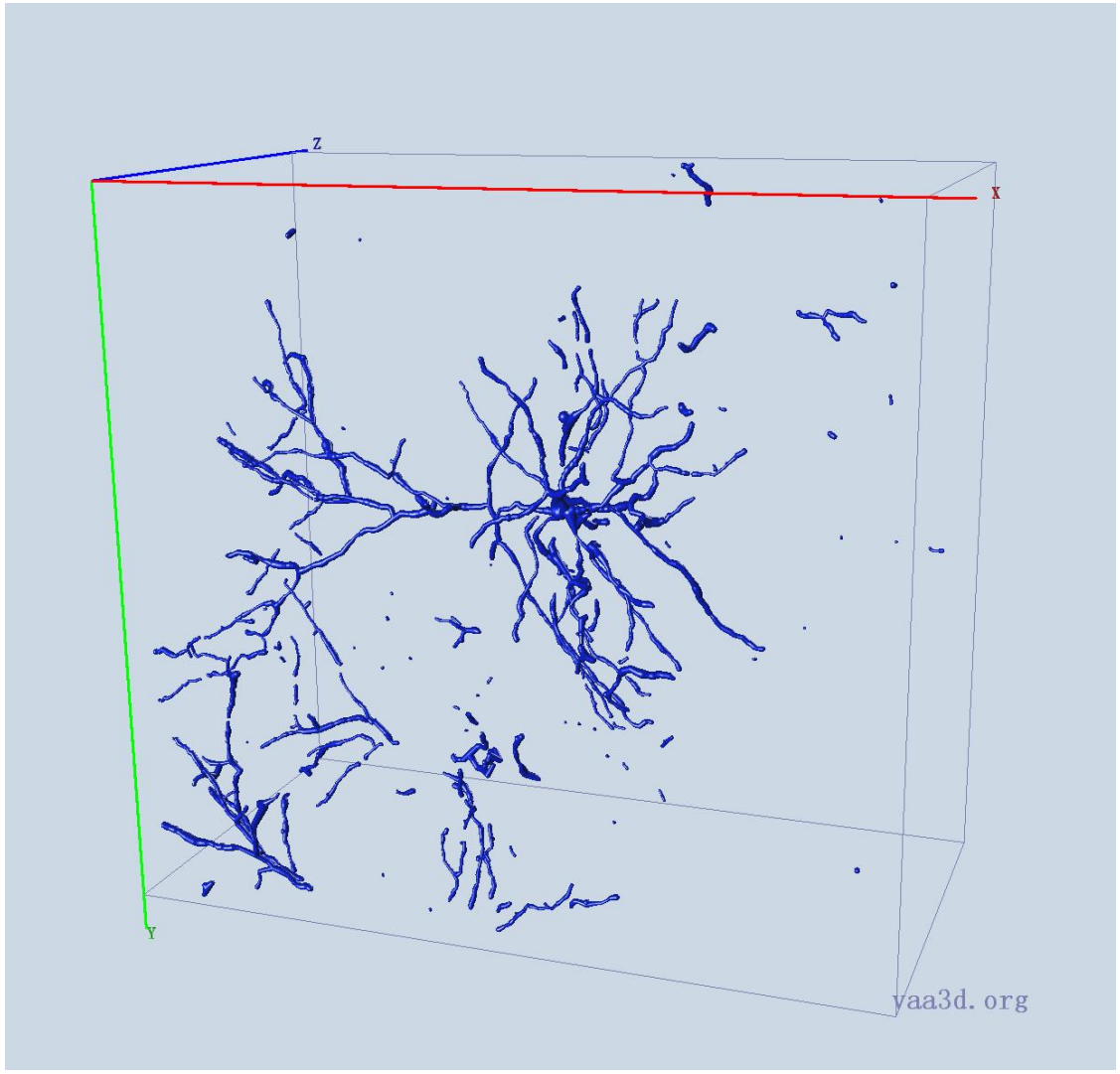
prun .swc files



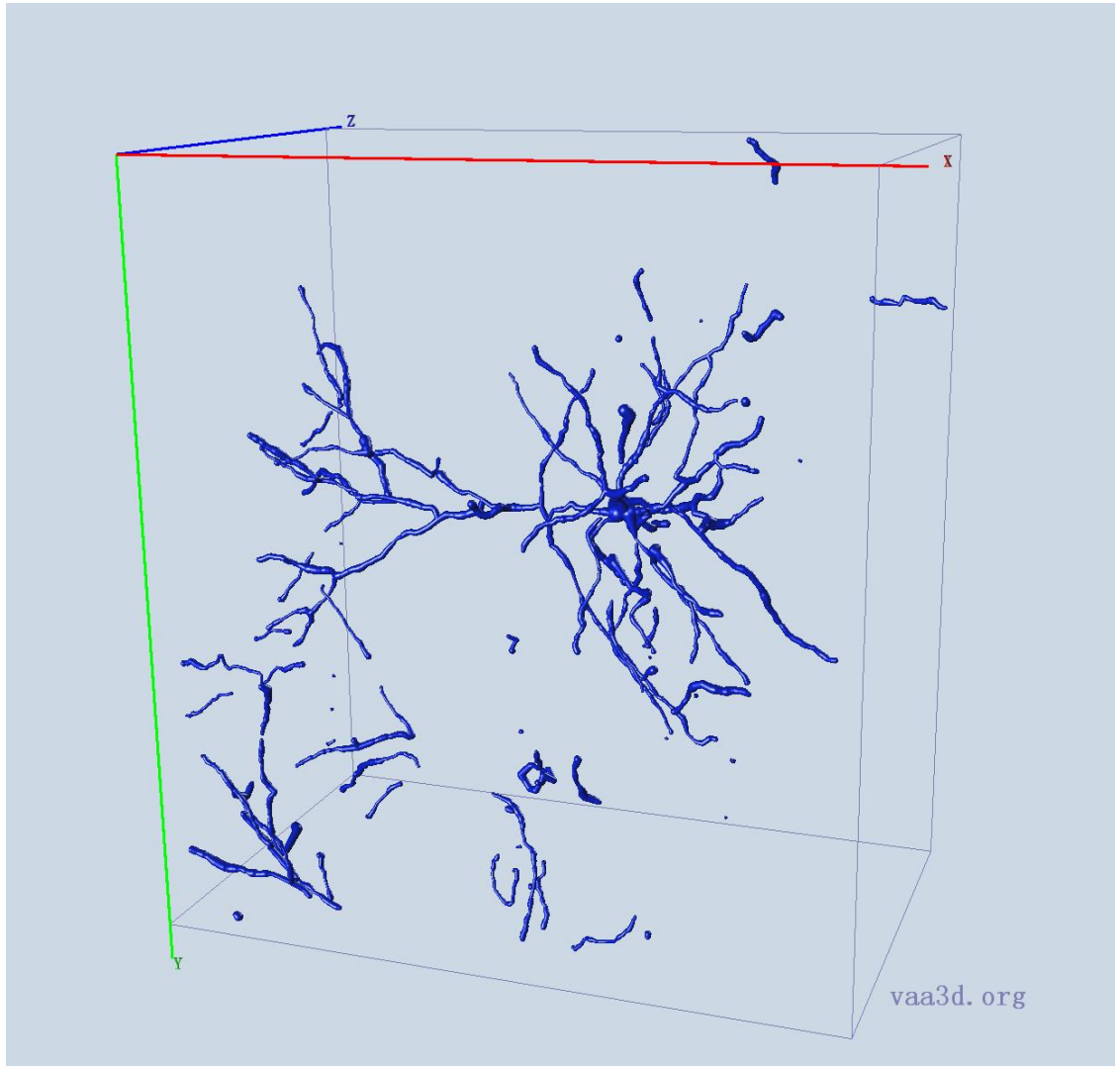
prun .swc files



Results

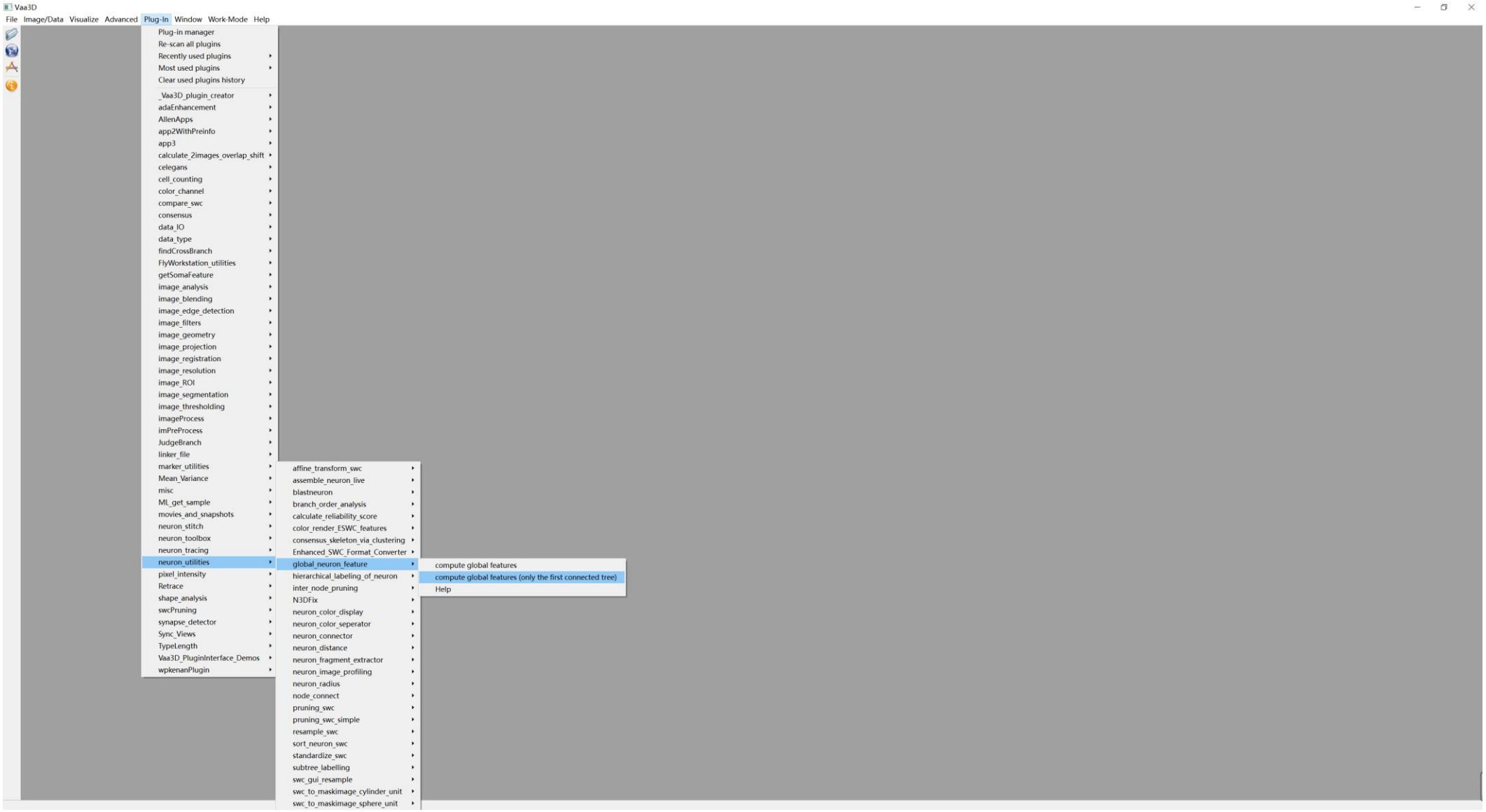


raw

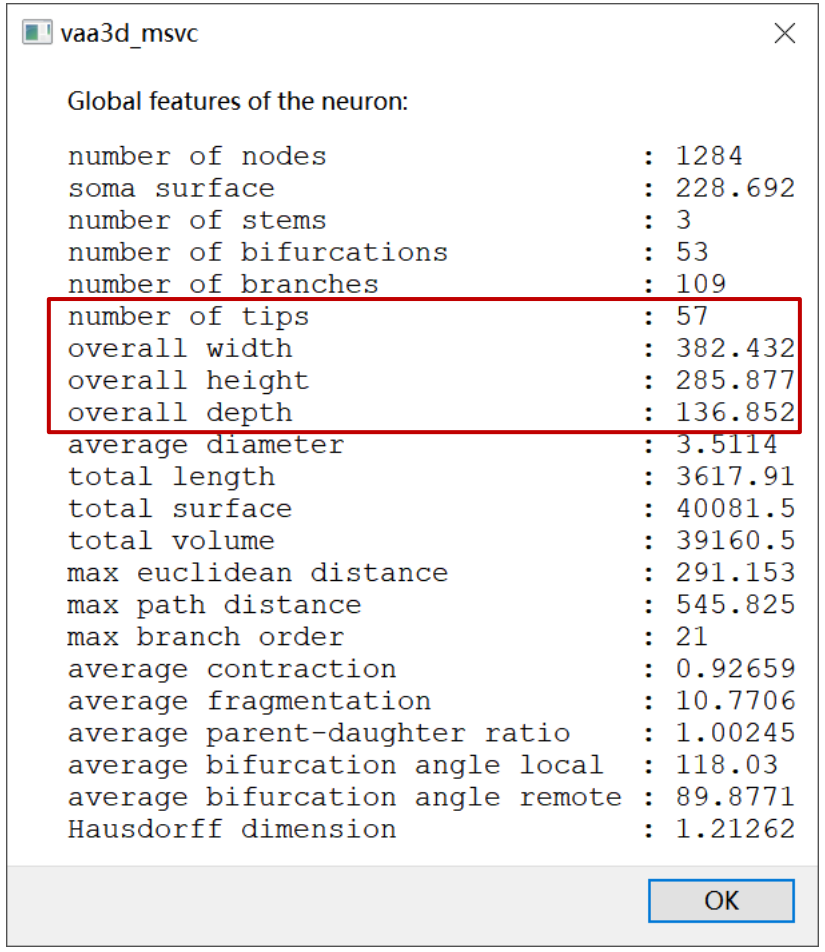


pruned

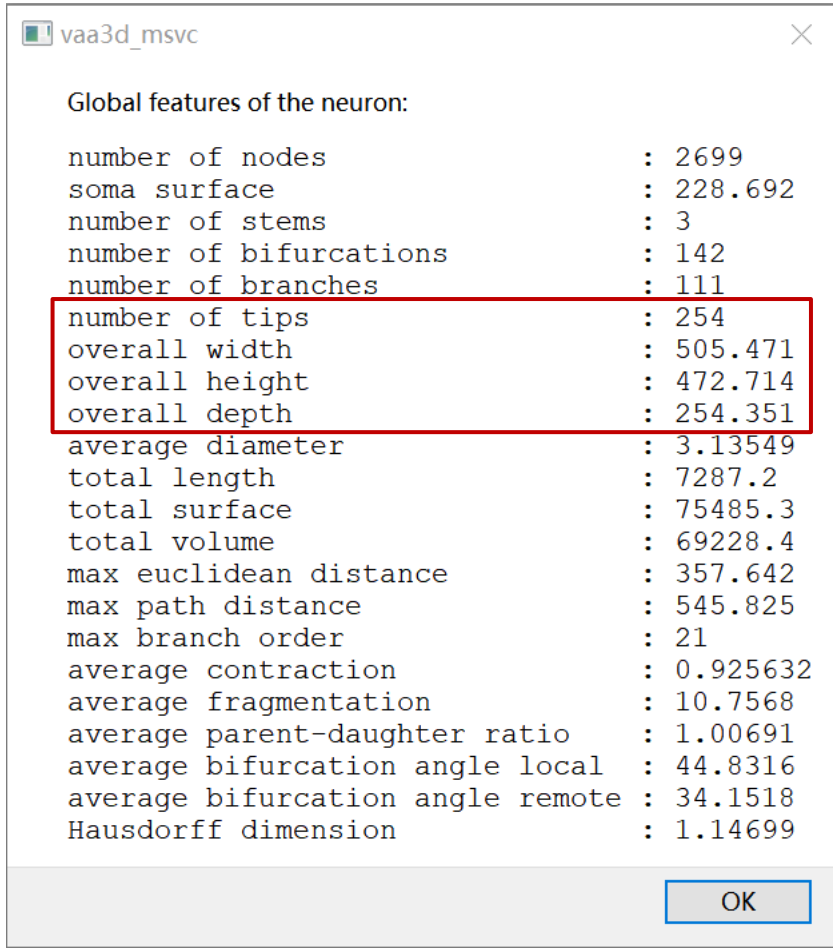
Global Features



Global Features

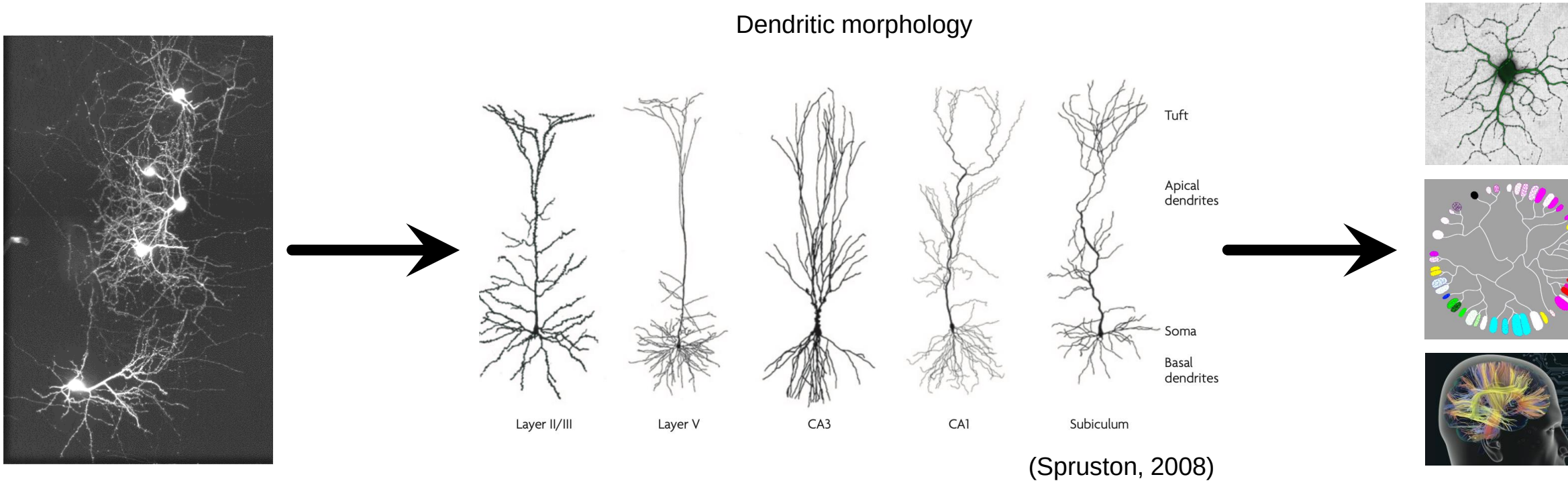


Pre-processed .swc



raw .swc

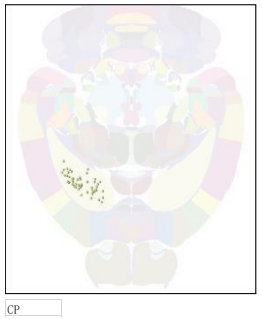
Key Questions in Neuroscience



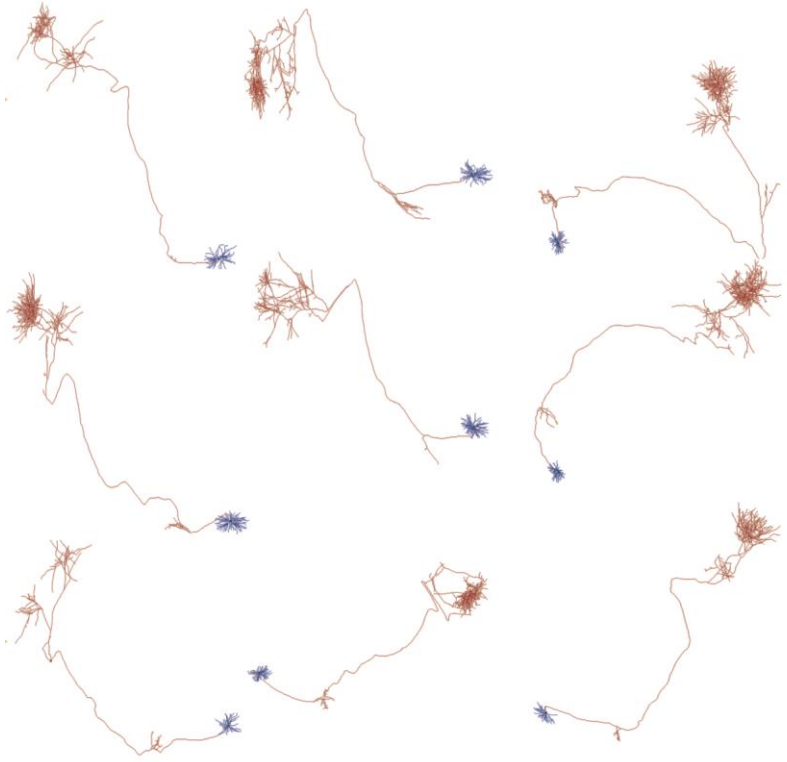
- Quantify/characterize neuron morphologies

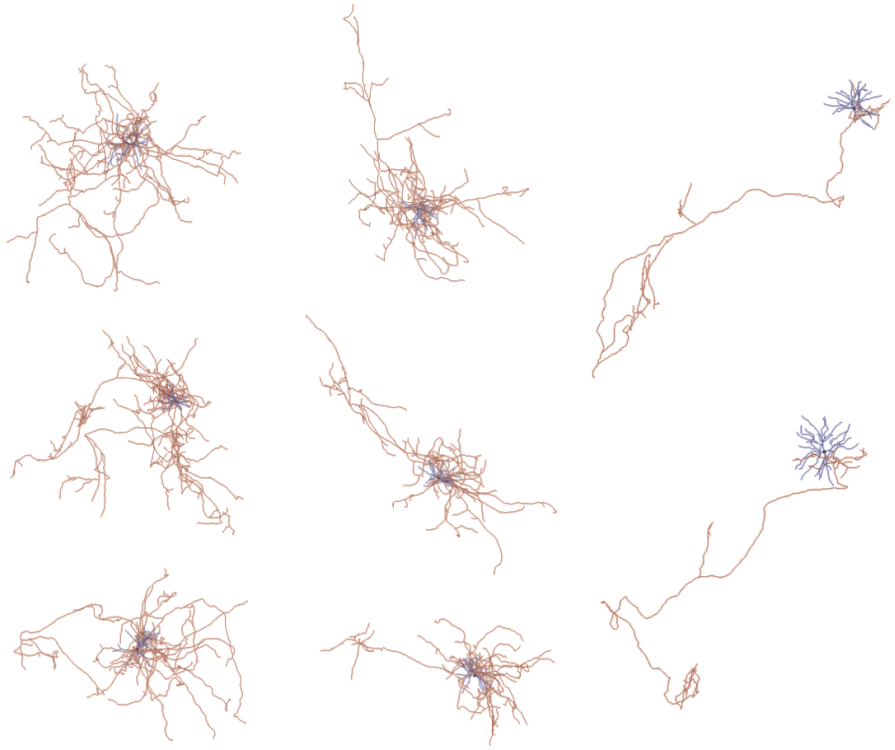
Single Neurons

VPM



CP

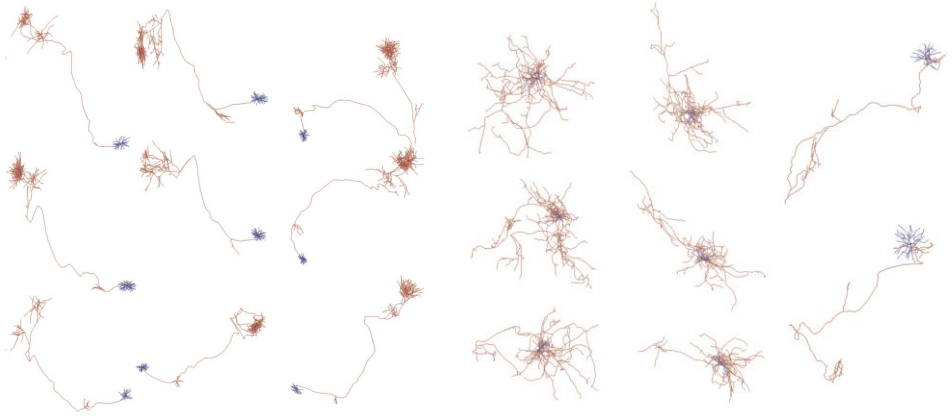


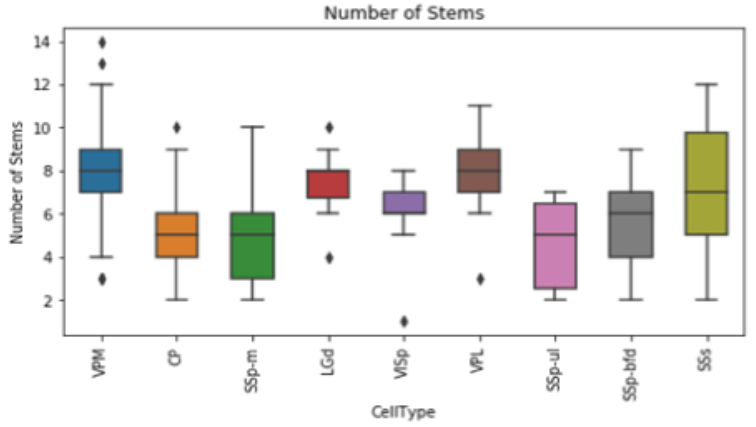


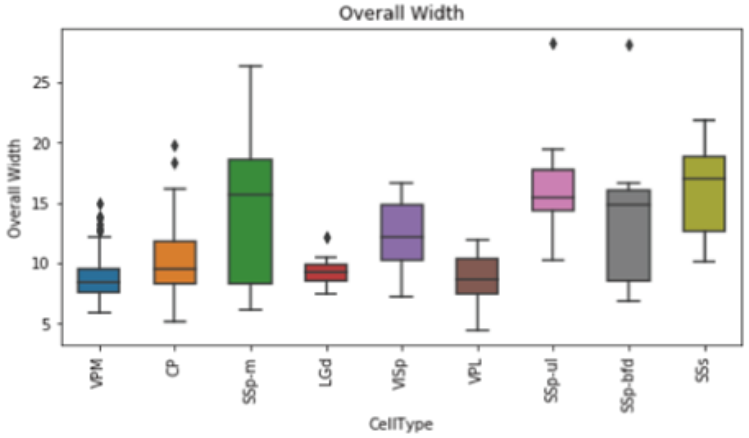
Morphological Features

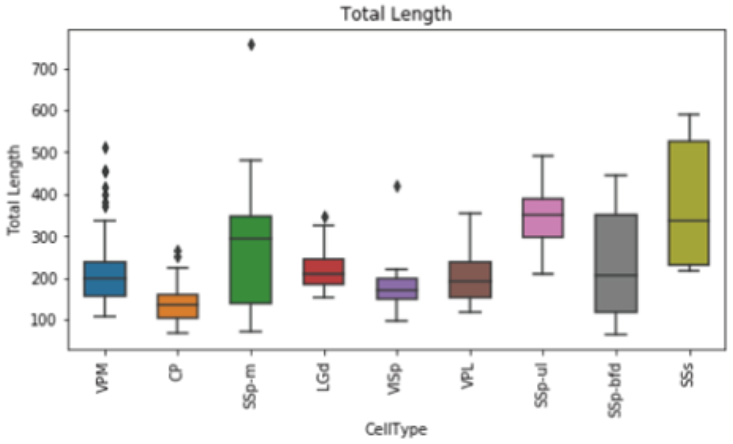
VPM

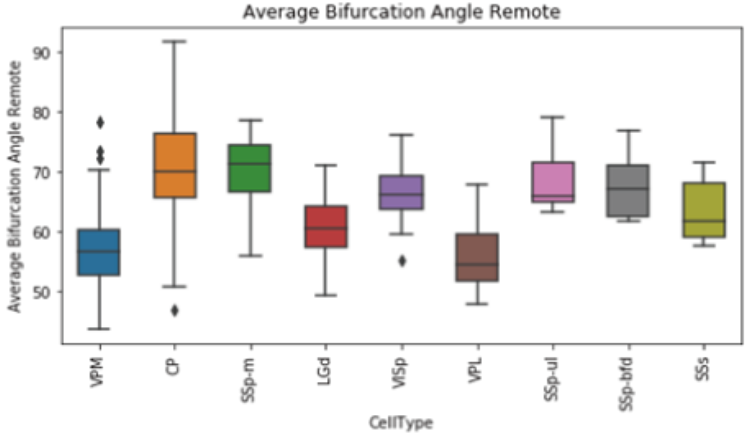
CP



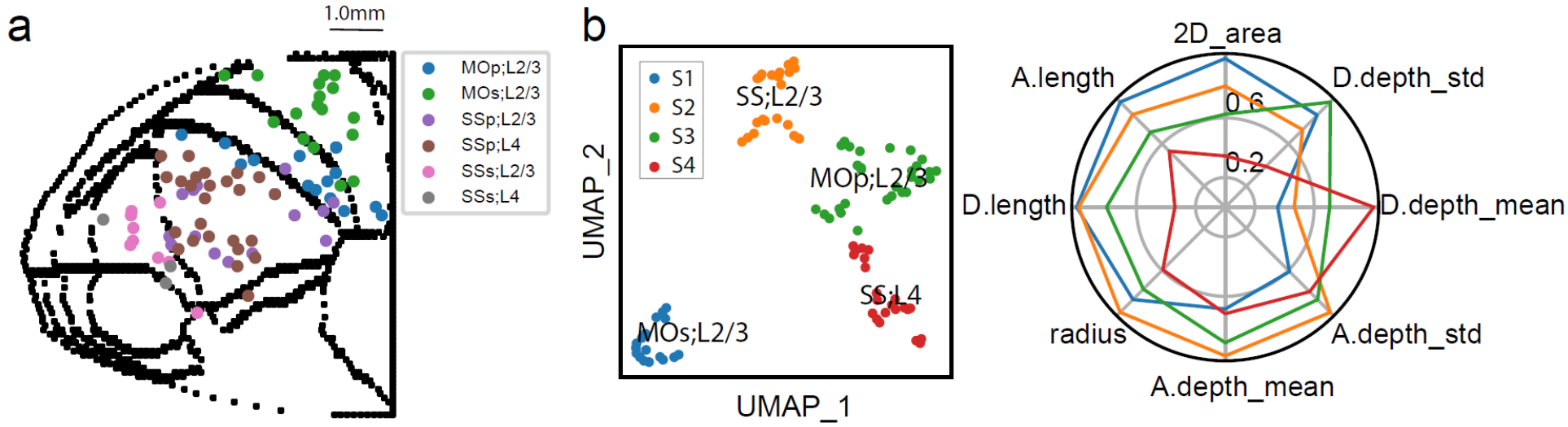


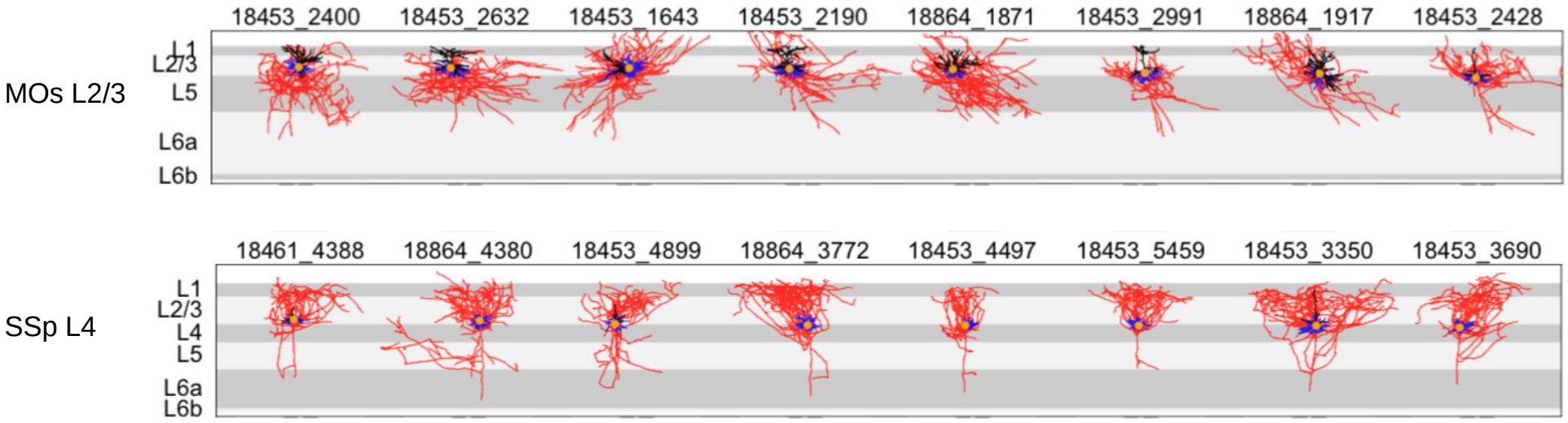






Local Arborization Patterns of Cortical L2/3/4 IT Neurons

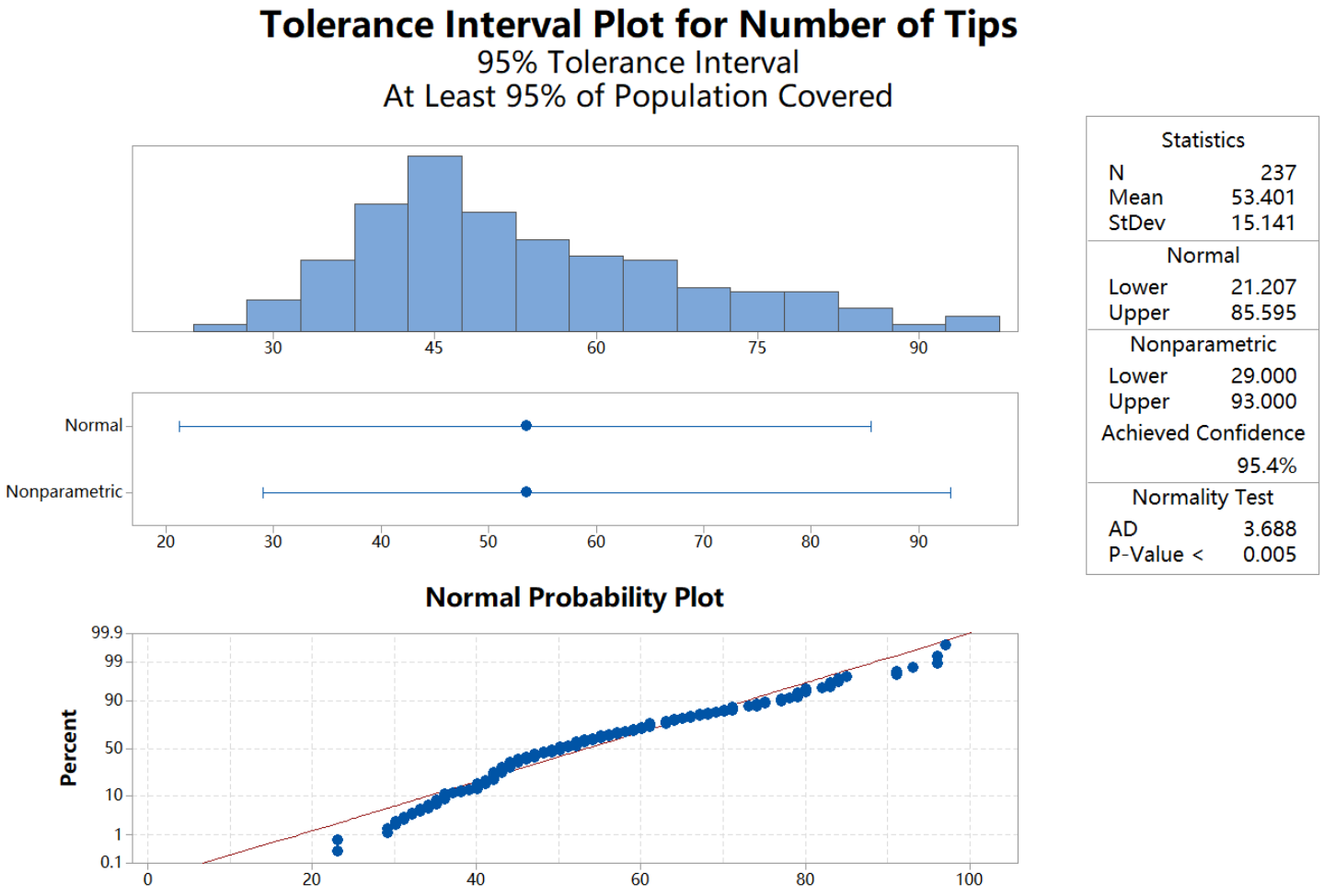




(Peng *et al*, 2020)

Automatic Quality Control

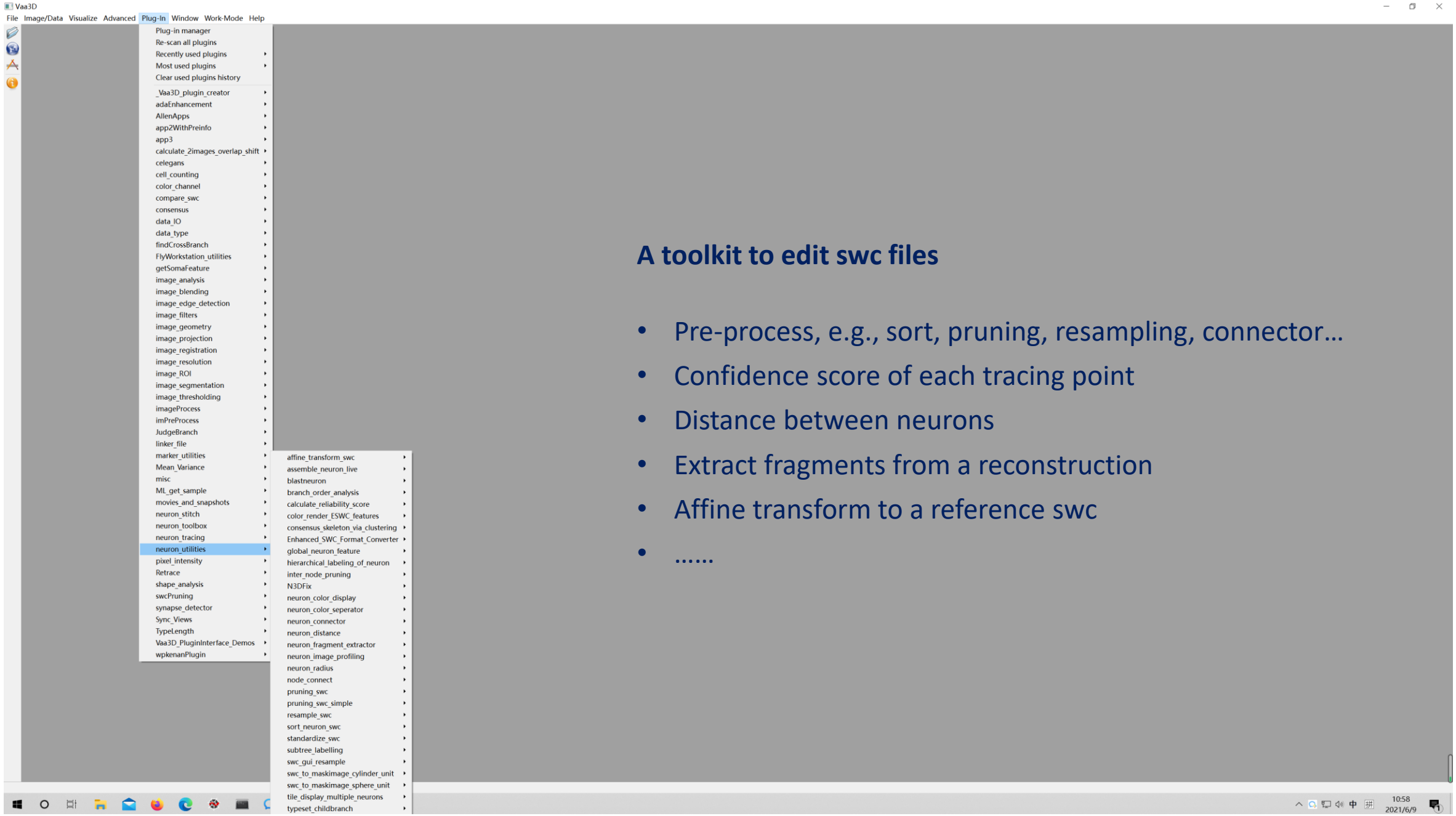
features after initial screening
Number of Tips
Overall Width
Overall Height
Overall Depth
Total Length
Max Branch Order
Number of Stems
Number of Bifurcatons
Number of Branches
Max Euclidean Distance
Max Path Distance
Average Contraction



Summary

- To calculate morphological features of a reconstruction
- To sort and prune .swc files as preprocessing steps

- Brief showcase where we can make use of the morphological features



A toolkit to edit swc files

- Pre-process, e.g., sort, pruning, resampling, connector…
- Confidence score of each tracing point
- Distance between neurons
- Extract fragments from a reconstruction
- Affine transform to a reference swc
- ……