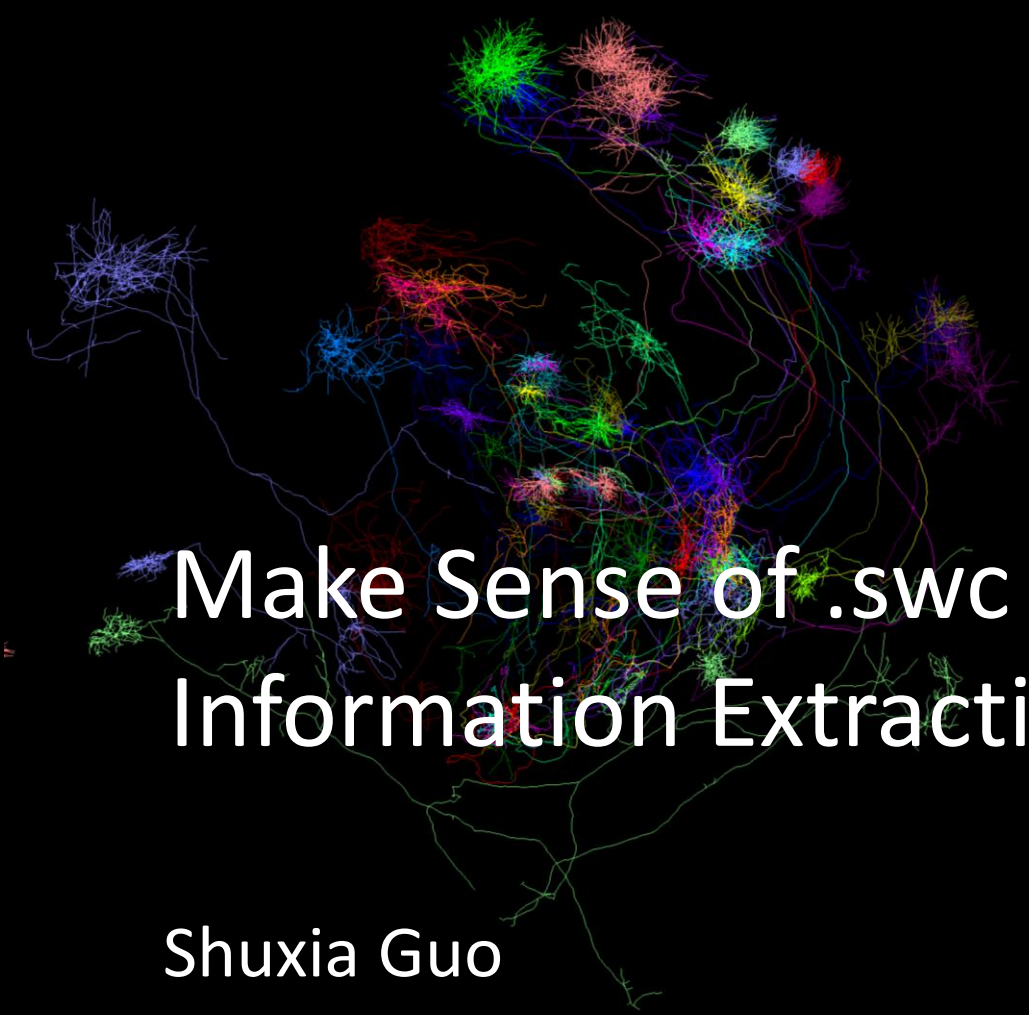


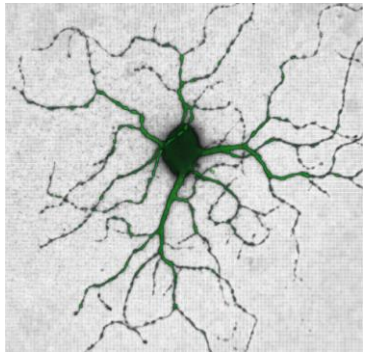
Make Sense of .swc Files: Information Extraction via Vaa3D Platform

Shuxia Guo

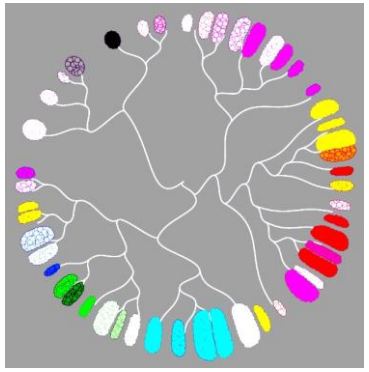
Aug 16th, 2021

ALLEN INSTITUTE *for* BRAIN SCIENCE

Key Questions in Neuroscience



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

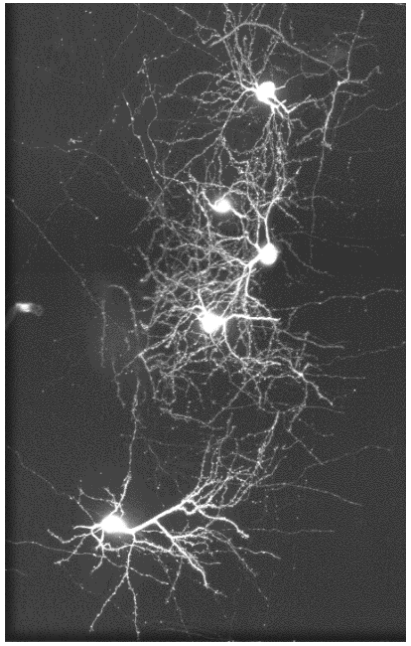


- **Where neurons come from and how they change over time?**
  (lineage, neuron growth/pruning/aging/plasticity,…)

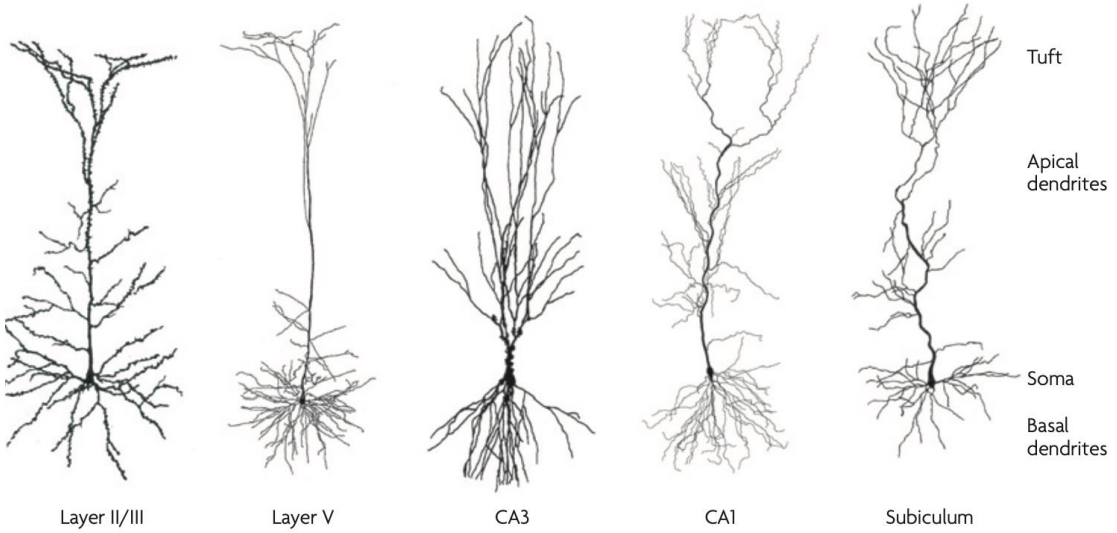
How neurons connect and work together?
(connectivity/connectome/networks, information processing, dynamics, behaviors, …)

pic courtesy of cosmosmagazine.com

Key Questions in Neuroscience



Dendritic morphology



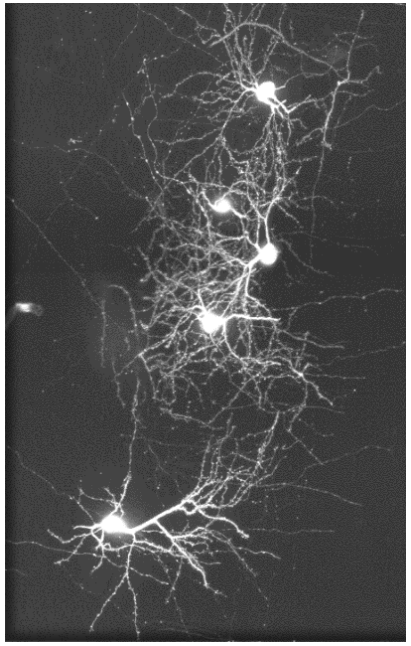
(Spruston, 2008)

.swc files

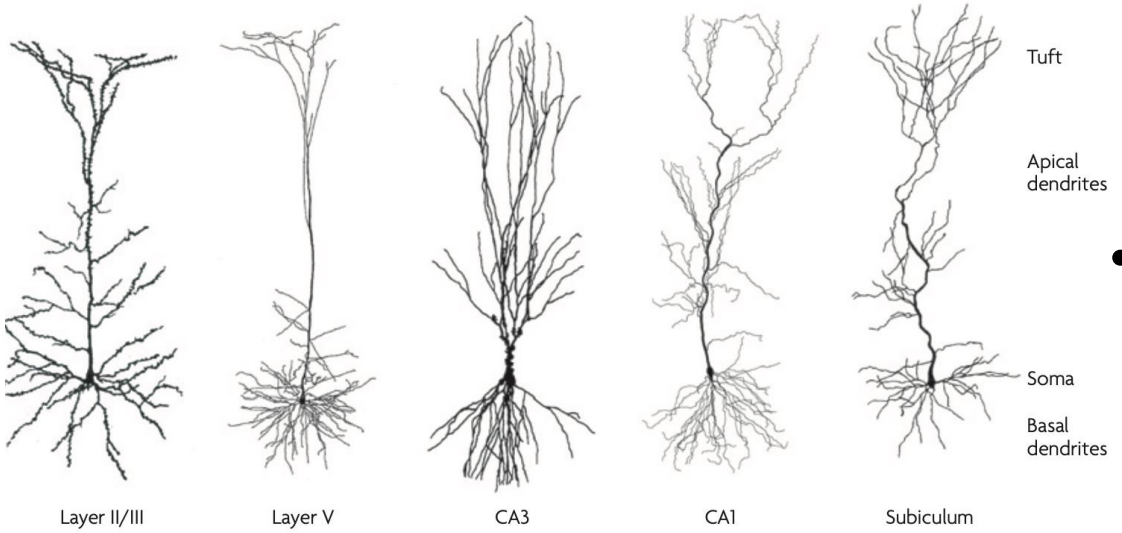
```
##n type x y z r parent
1 1 513.69 514.30 258.00 1.0 -1
2 3 513.50 515.97 258.40 1.0 1
3 3 512.86 520.58 259.52 1.0 2
4 3 512.28 525.18 260.62 1.0 3
5 3 511.70 529.73 261.73 1.0 4
6 3 511.06 534.34 262.83 1.0 5
7 3 511.00 539.52 263.58 1.0 6
8 3 511.00 545.15 264.10 1.0 7
9 3 510.87 549.89 265.01 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.

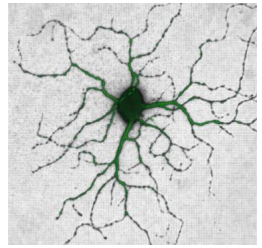
Key Questions in Neuroscience

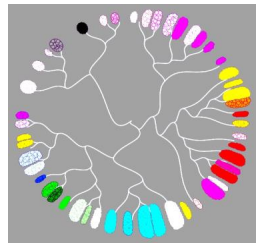


Dendritic morphology

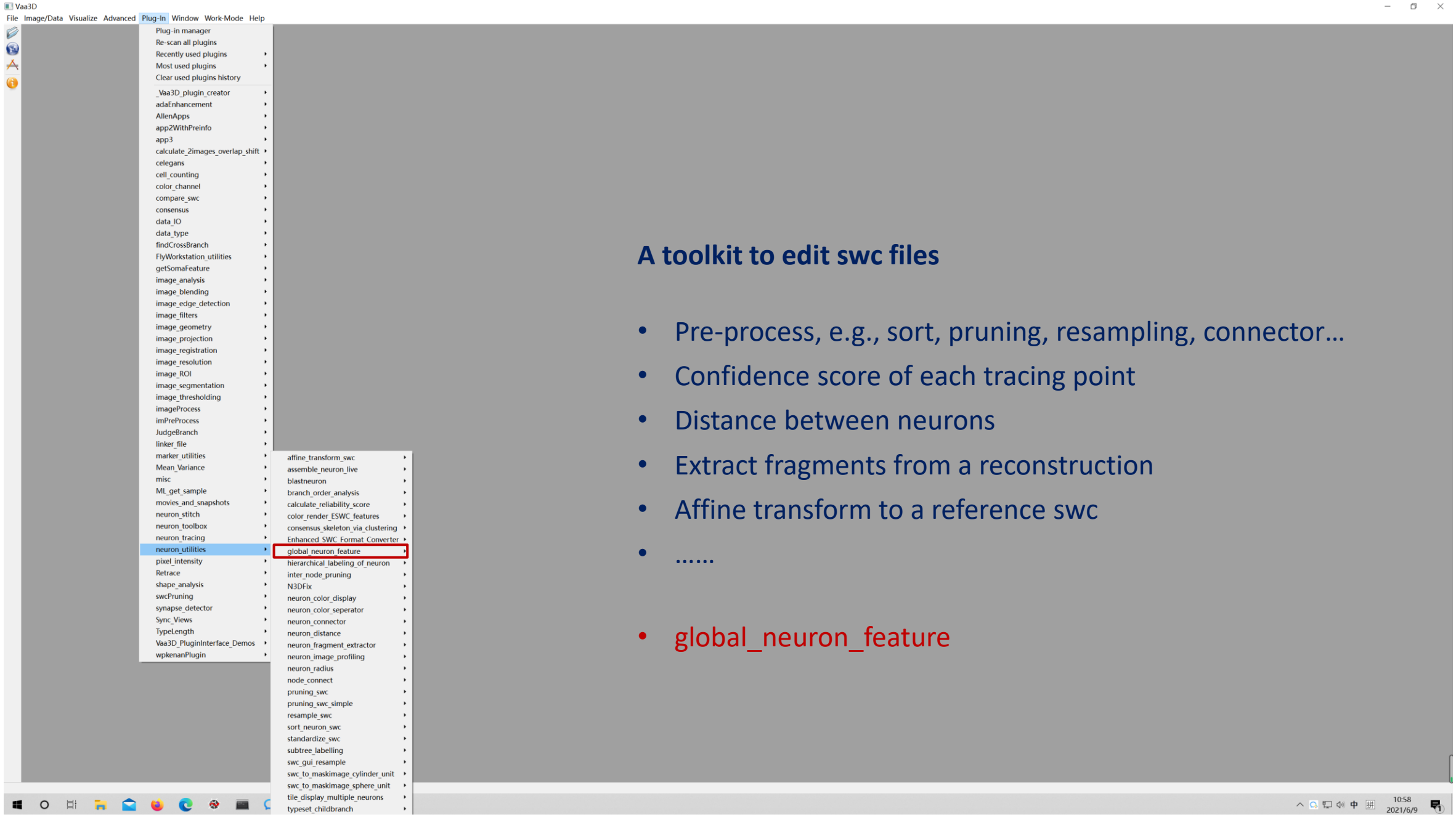


(Spruston, 2008)





- 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
- ……

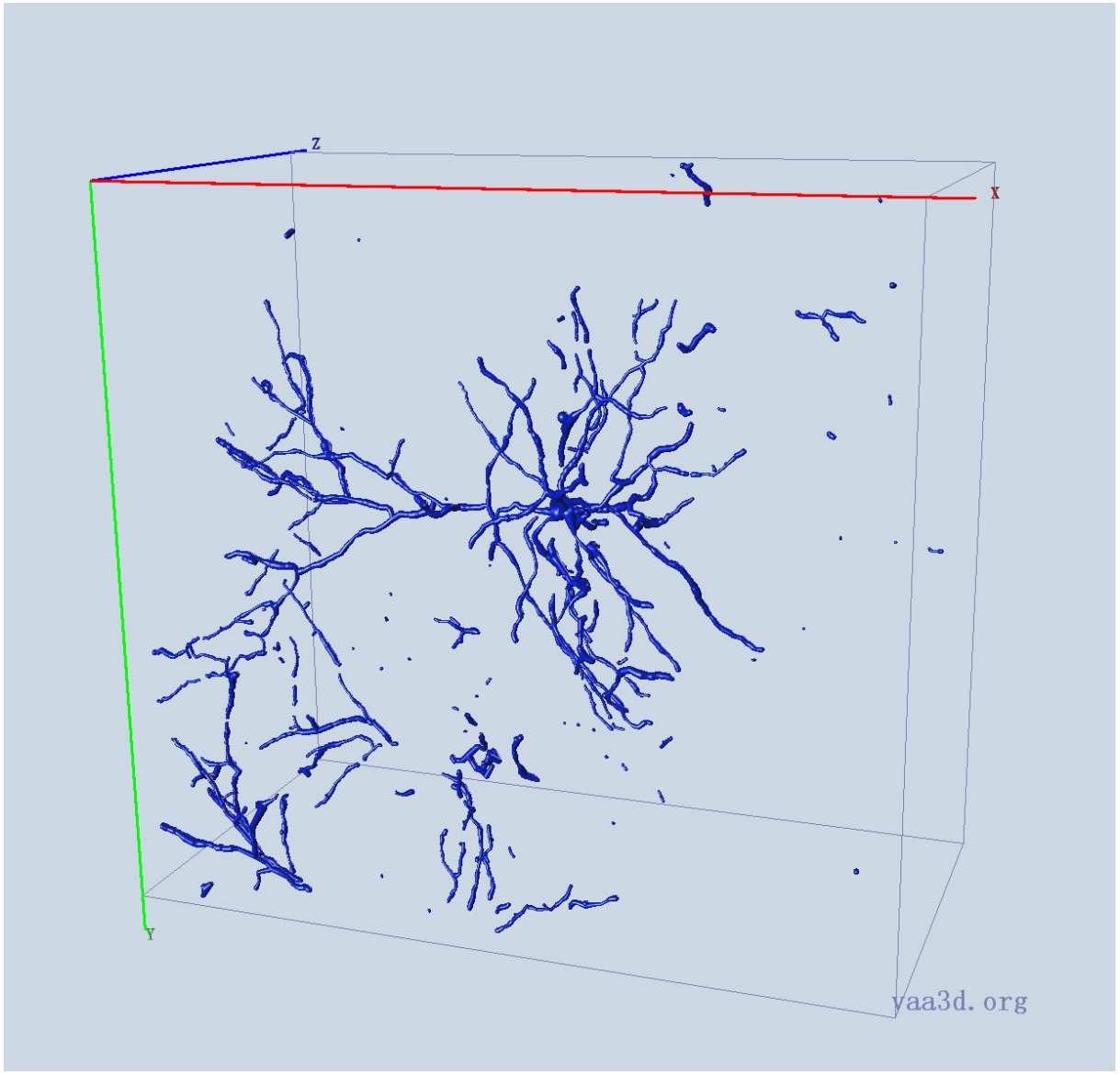
- global_neuron_feature

L-Measure Features

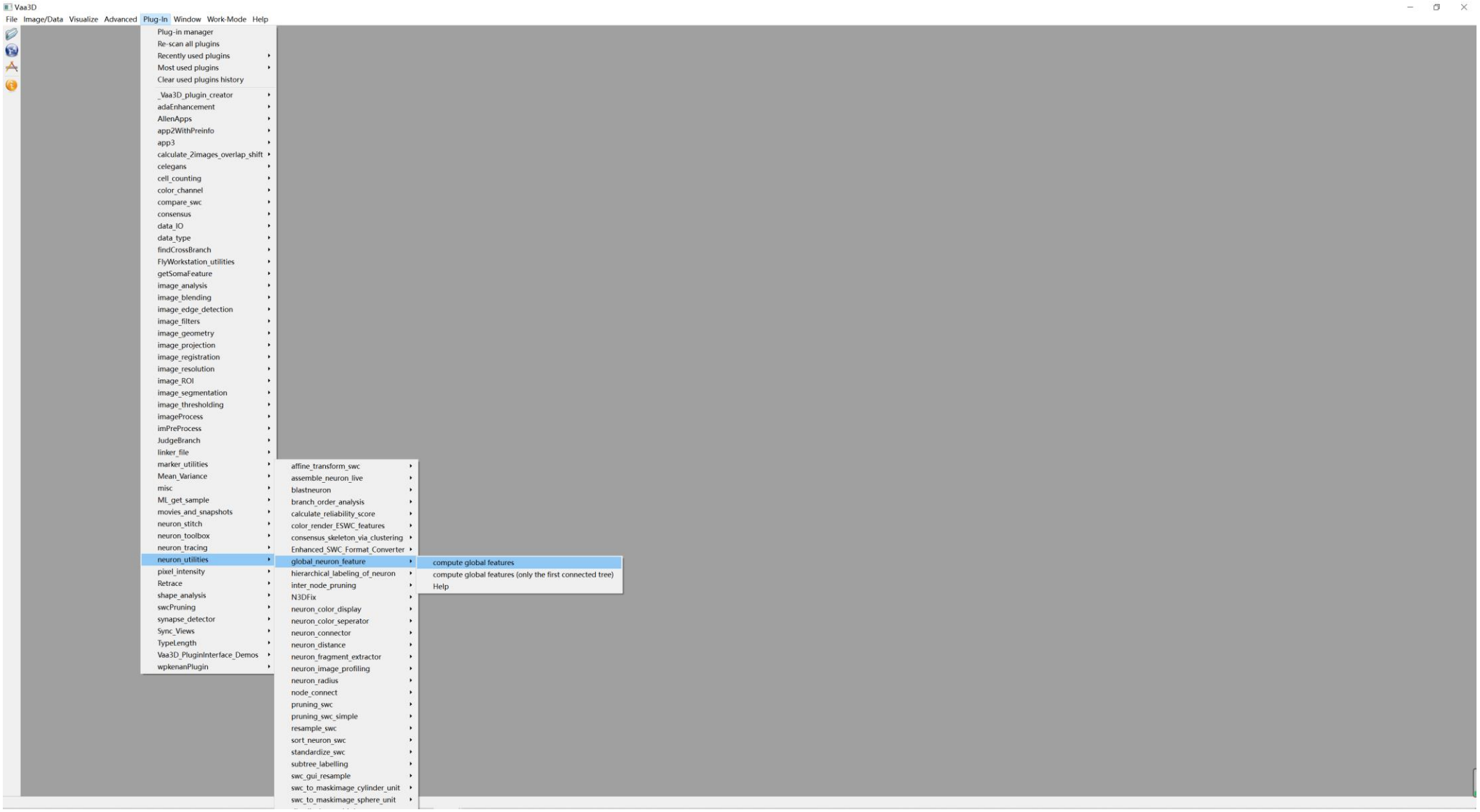
- Number of Nodes
- Soma Surface
- Number of Stems
- Number of Bifurcations
- Number of Branches
- Number of Tips
- Overall Width
- Overall Height
- Overall Depth
- Average Diameter
- Total Length
- Total Surface
- Total Volume
- Max Euclidean Distance
- Max Path Distance
- Max Branch Order
- Average 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

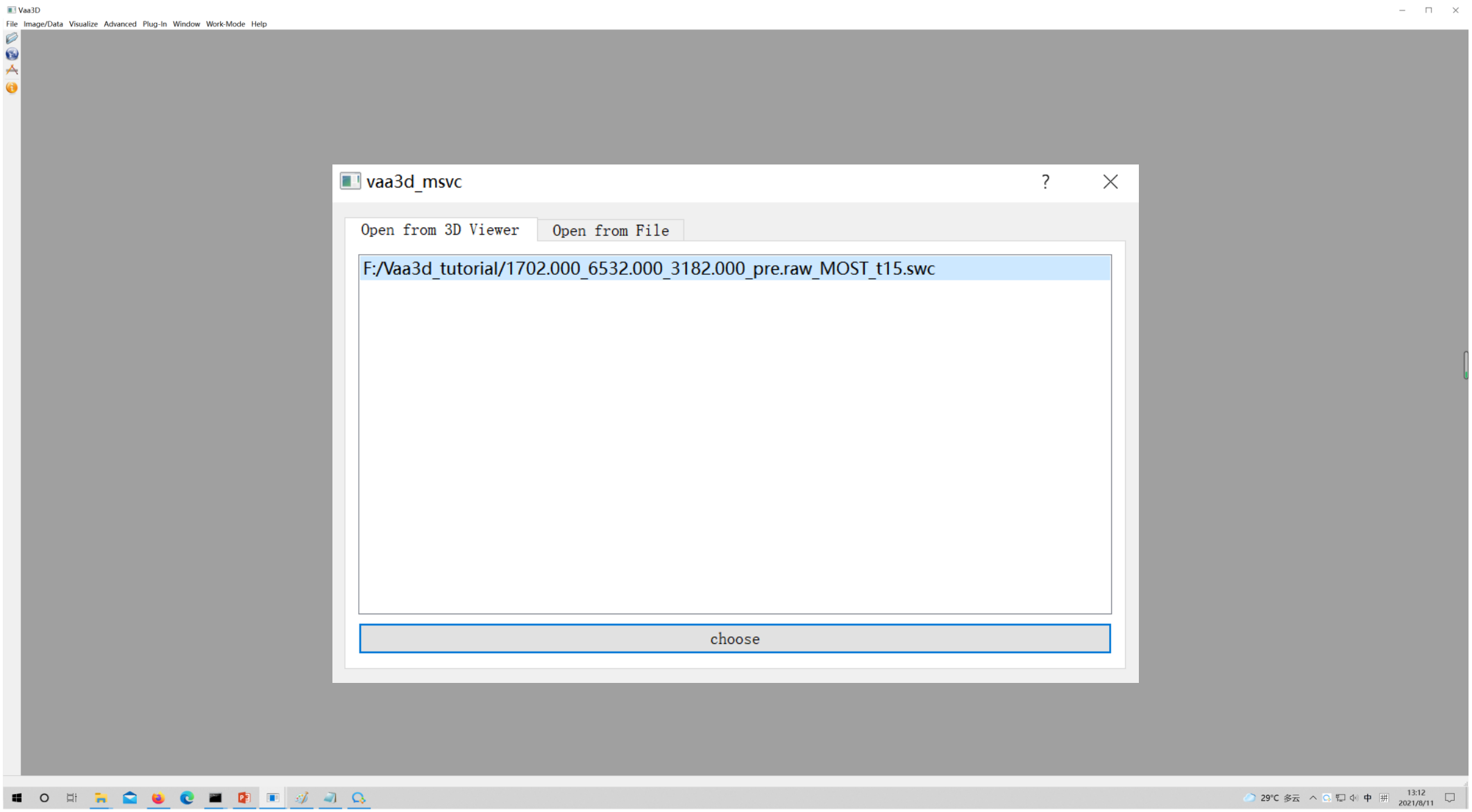
Global Features via GUI



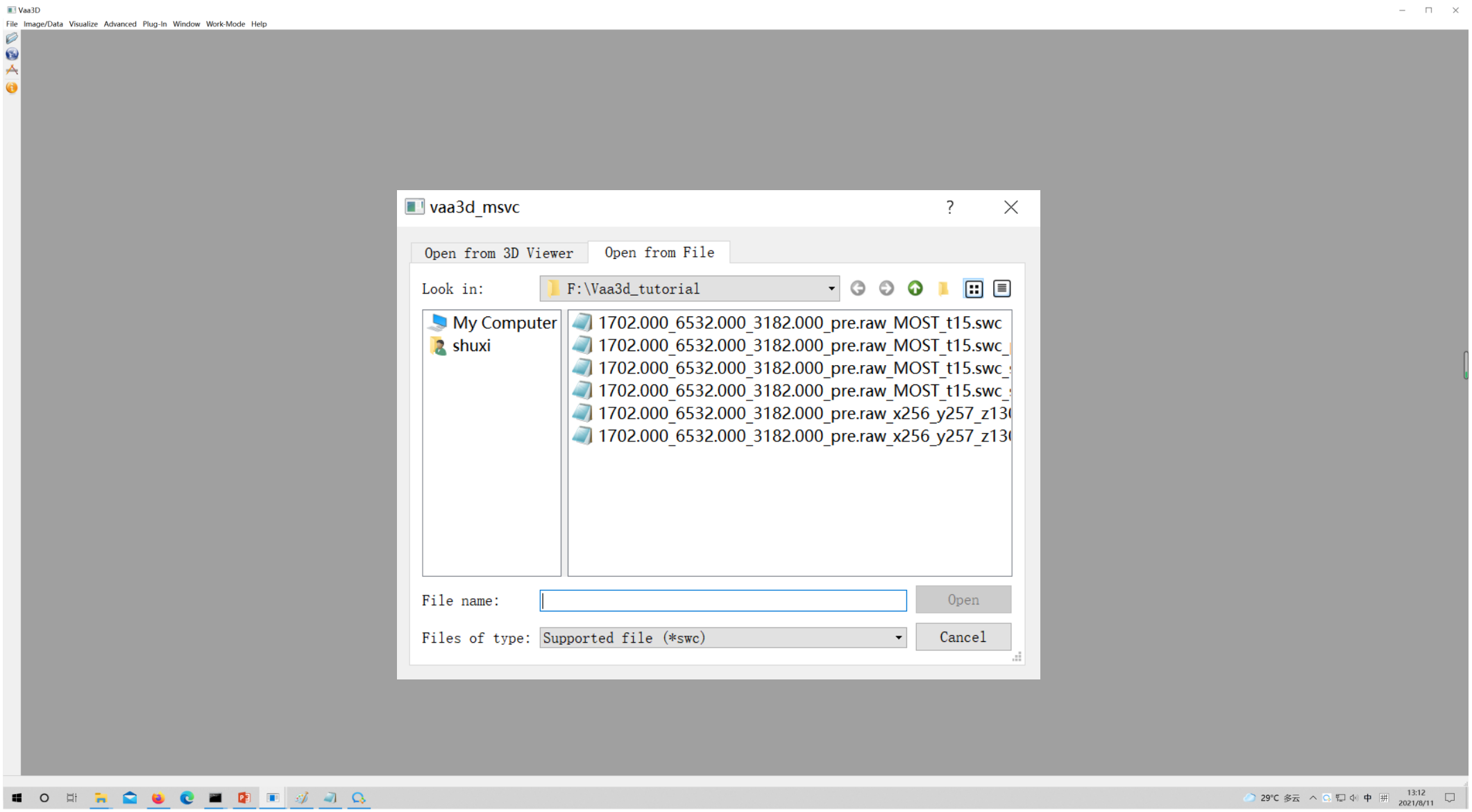
Global Features via GUI



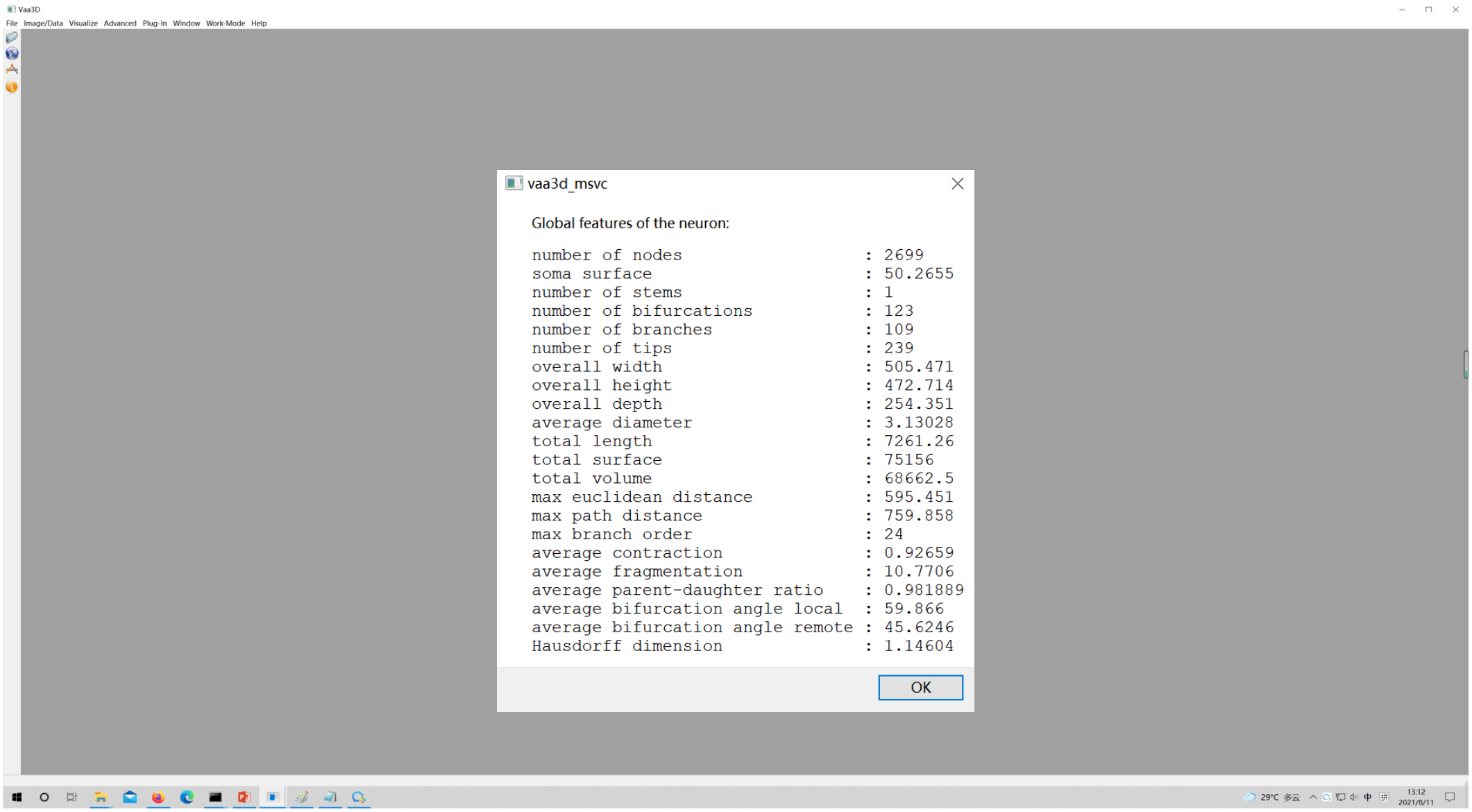
Global Features via GUI



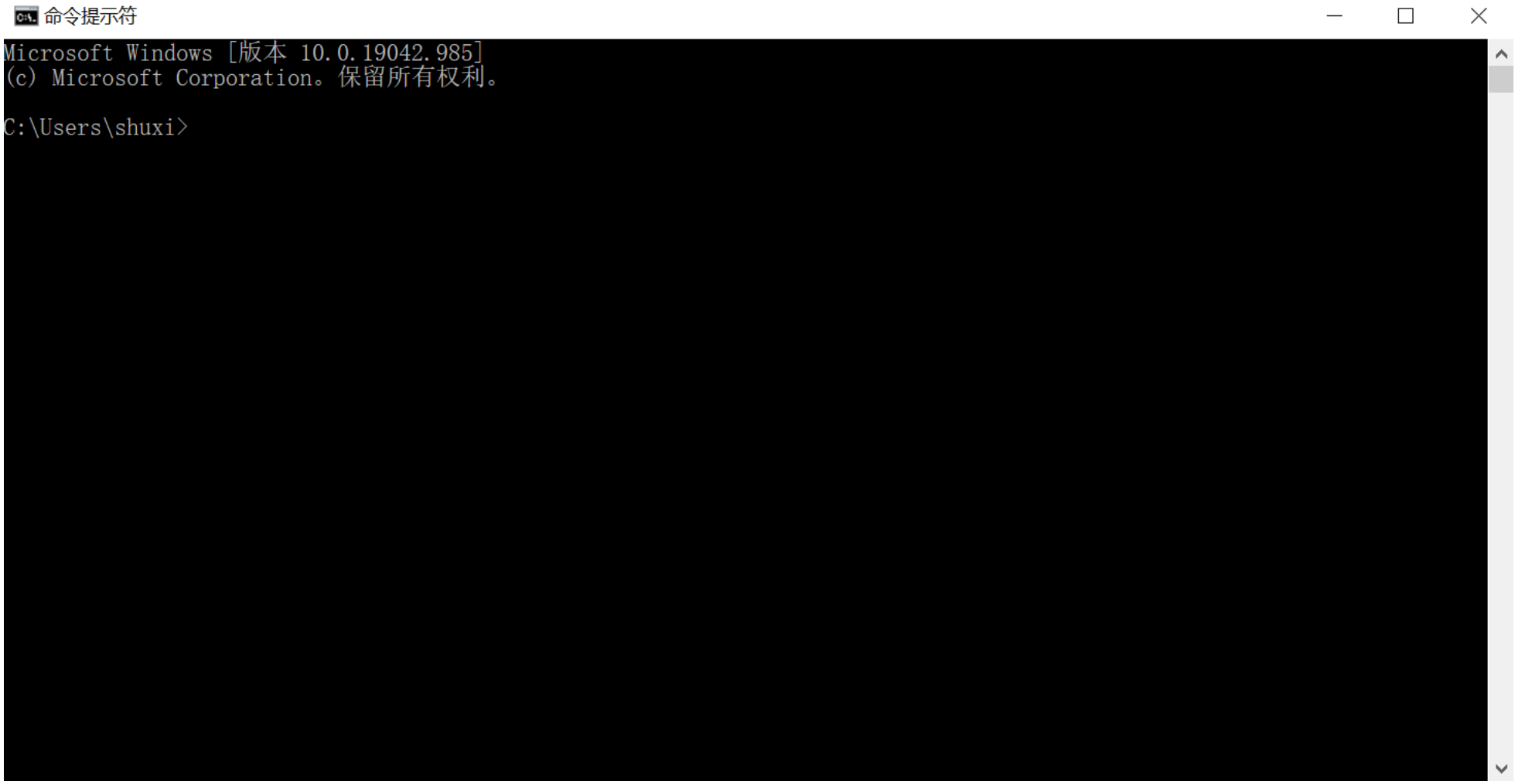
Global Features via GUI



Global Features via GUI



Run Functions Via Command Line

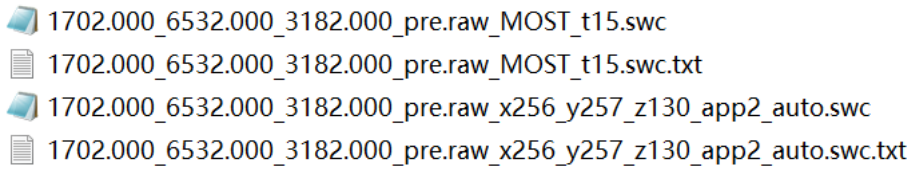


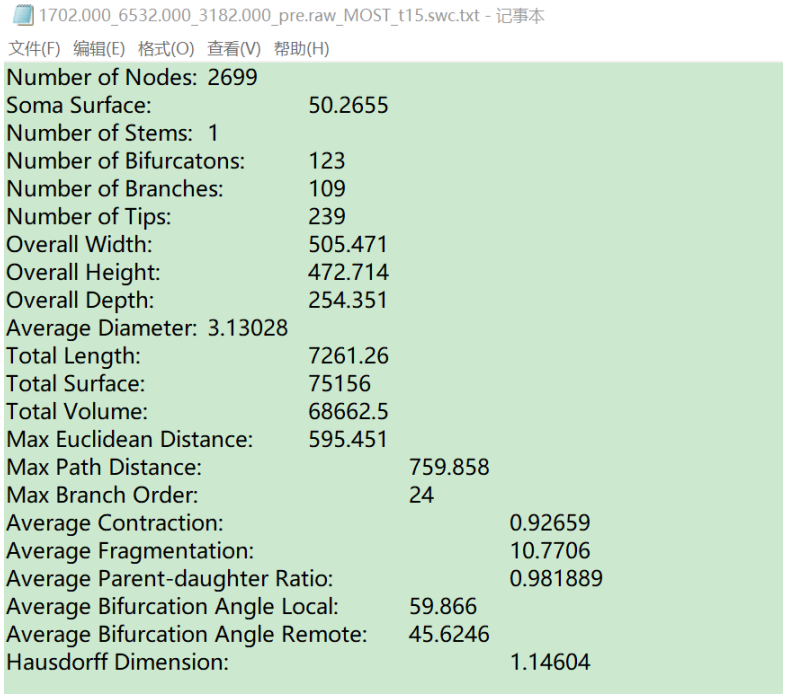
Global Features Via Command Line

v3d software → v3d | plugin name → global_neuron_feature | function name → compute_feature_batch | path of swc files → path_of_swc_files

W v3d /x global_neuron_feature /f compute_feature_batch /i path_of_swc_files

L/M v3d –x global_neuron_feature –f compute_feature_batch –i path_of_swc_files



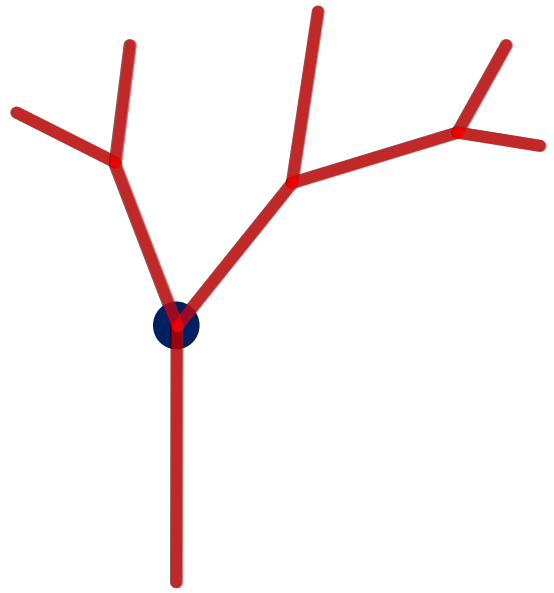


Find More from Source Code

- vaa3d_tools/released_plugins/v3d_plugins/global_neuron_feature/

.swc files

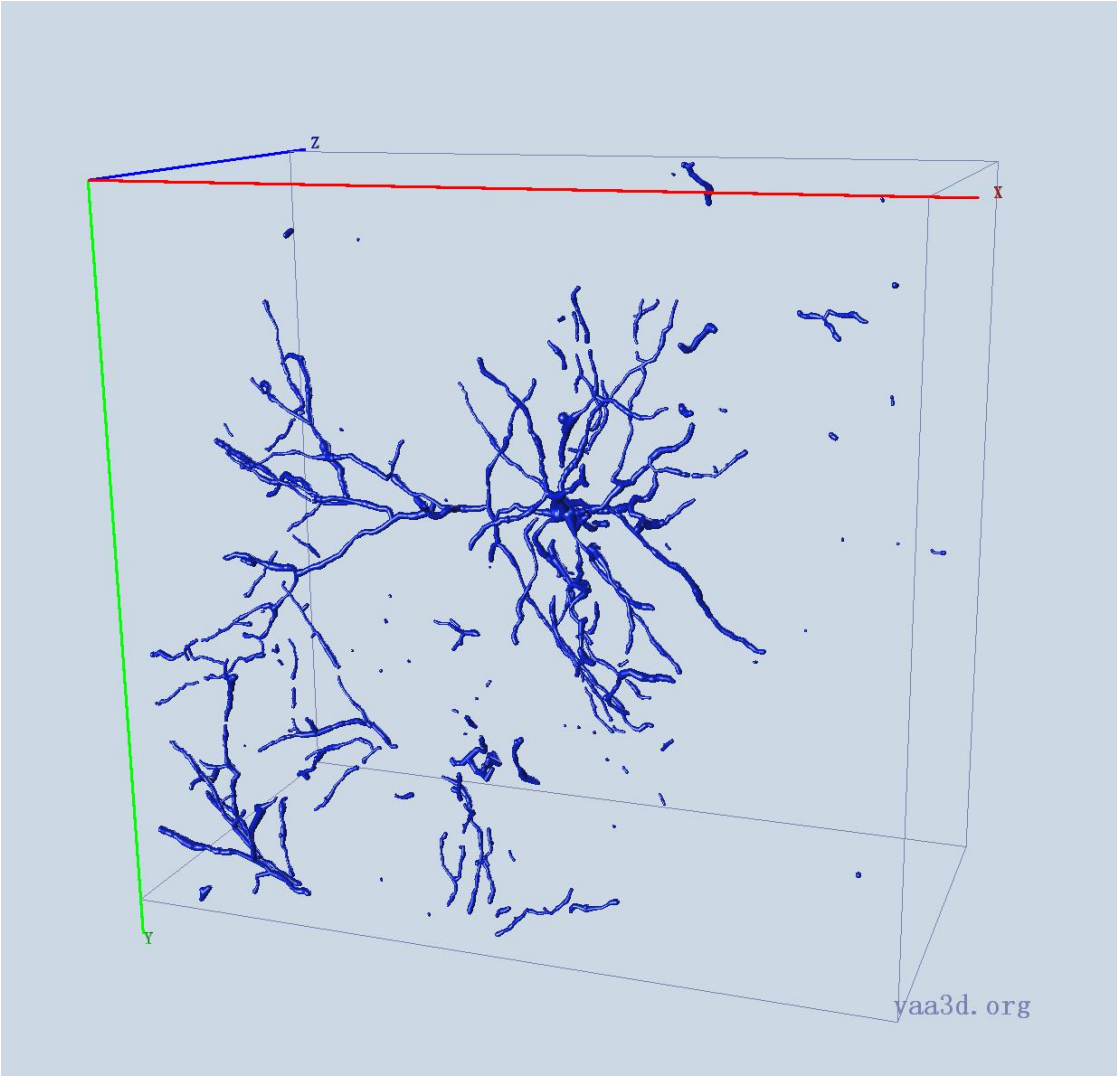
```
##n type x y z r parent
1 1 513.69 514.30 258.00 1.0 -1
2 3 513.50 515.97 258.40 1.0 1
3 3 512.86 520.58 259.52 1.0 2
4 3 512.28 525.18 260.62 1.0 3
5 3 511.70 529.73 261.73 1.0 4
6 3 511.06 534.34 262.83 1.0 5
7 3 511.00 539.52 263.58 1.0 6
8 3 511.00 545.15 264.10 1.0 7
9 3 510.87 549.89 265.01 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
```



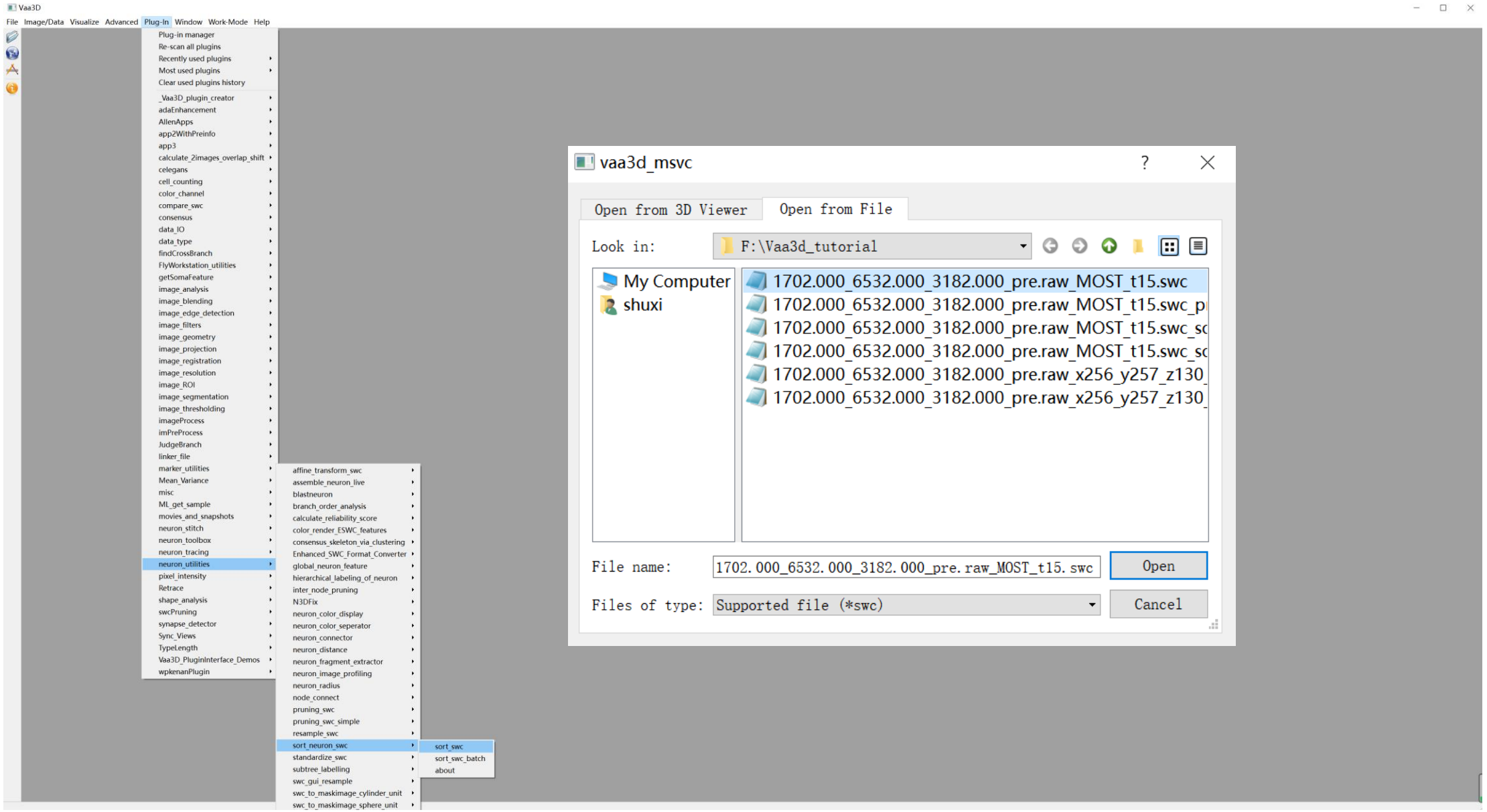
Pre-process .swc Files

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

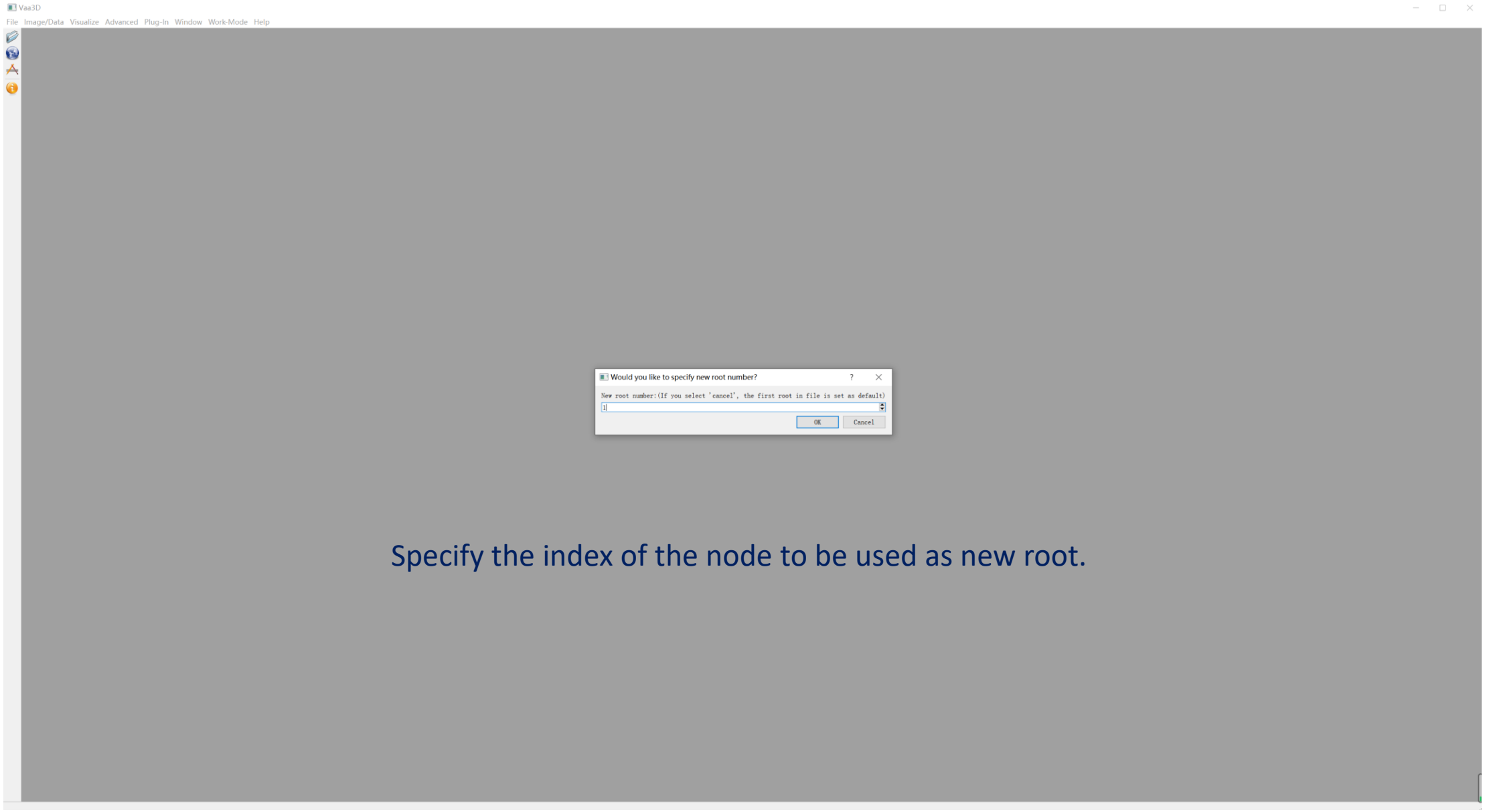
- sort → prune



Sort swc via GUI

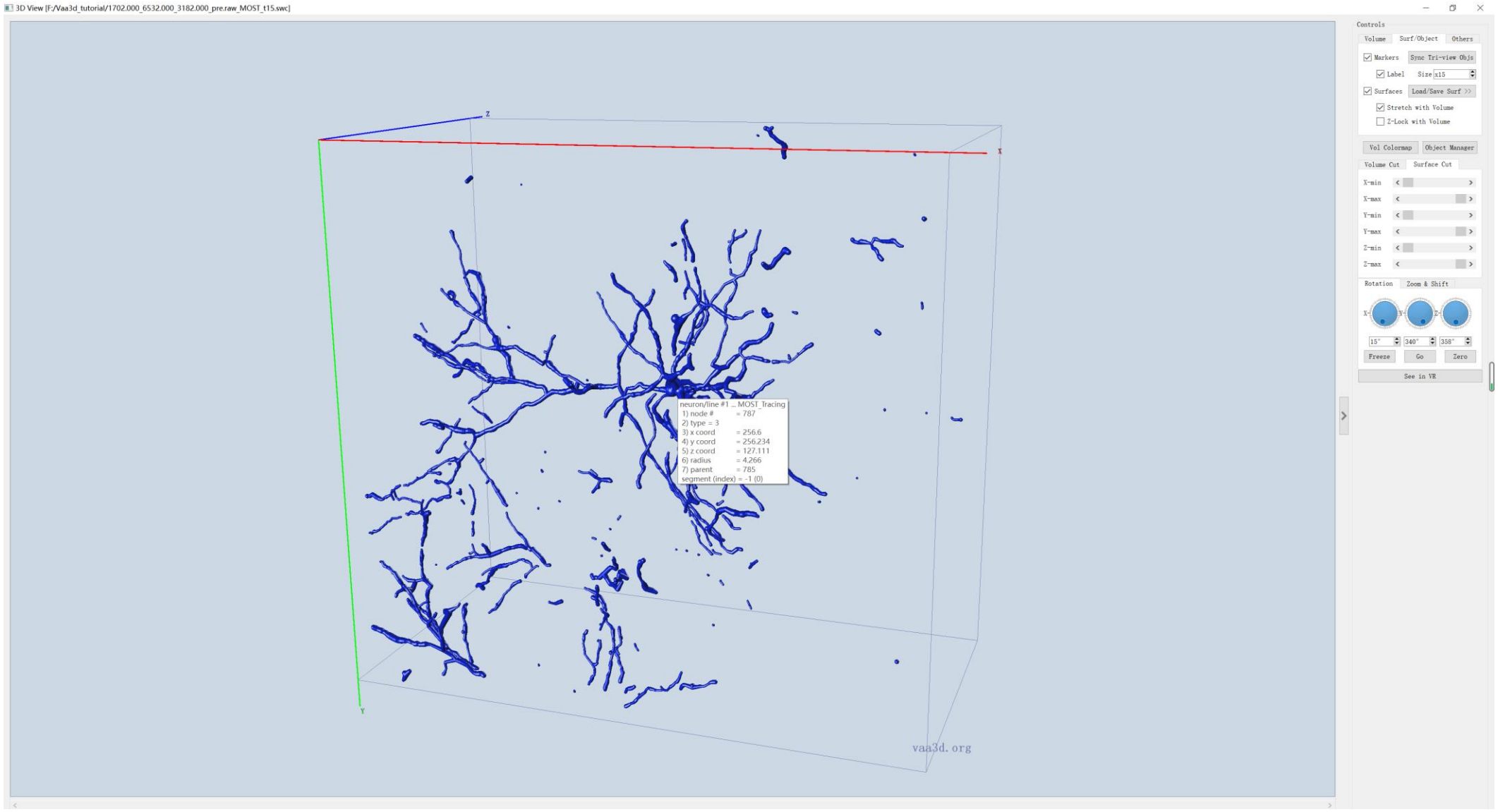


Sort swc via GUI

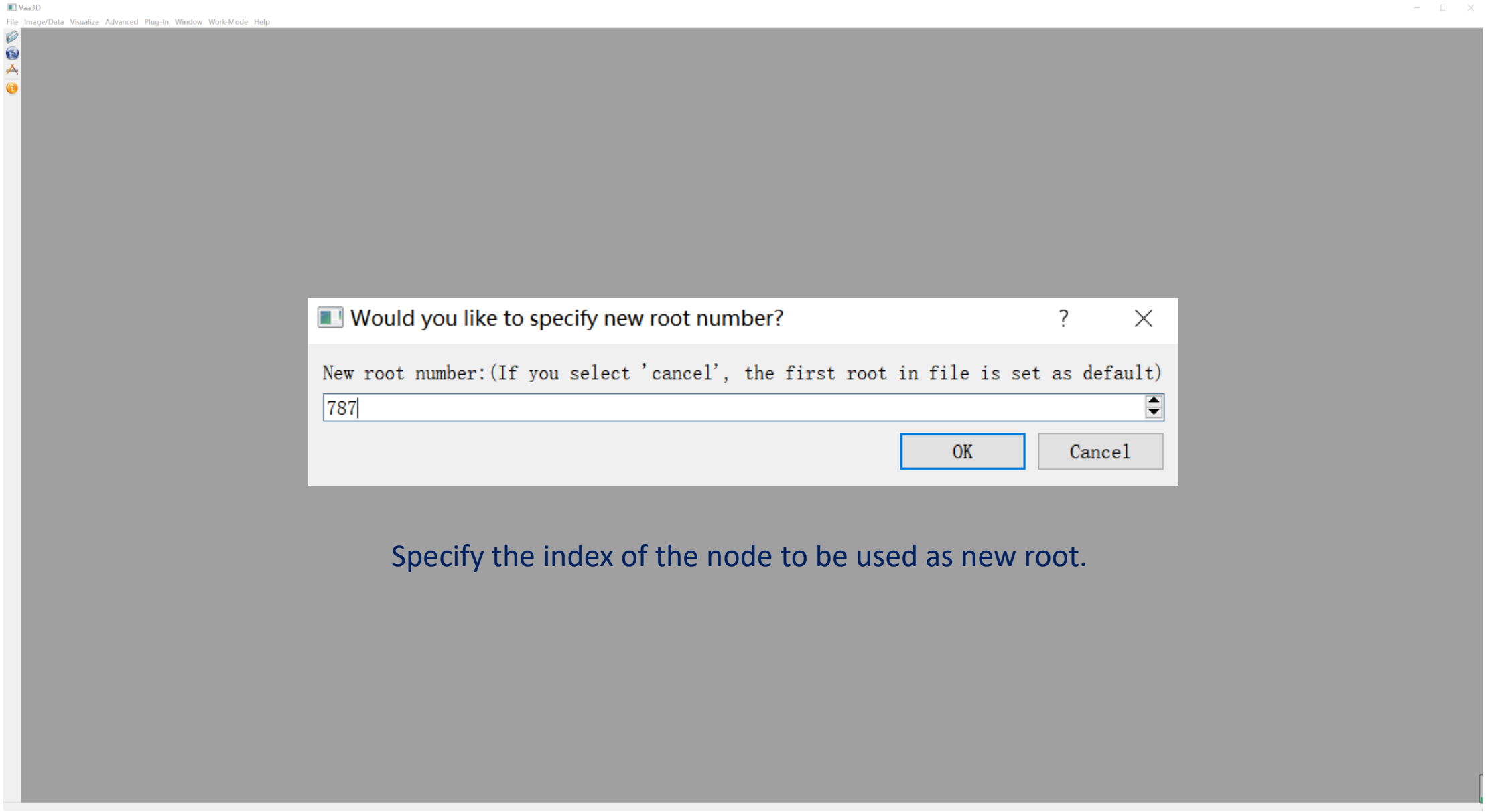


Specify the index of the node to be used as new root.

Sort swc via GUI

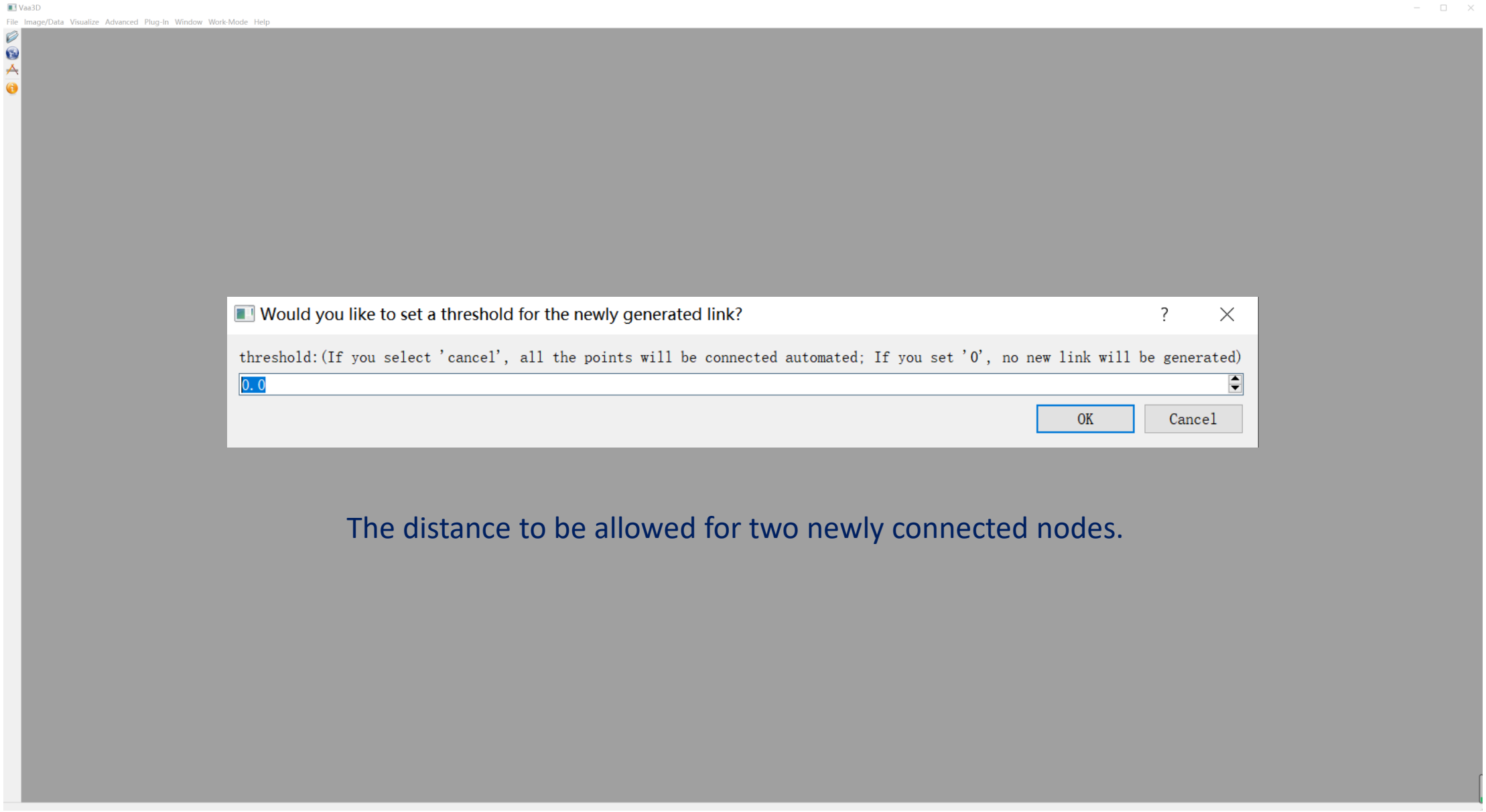


Sort swc via GUI



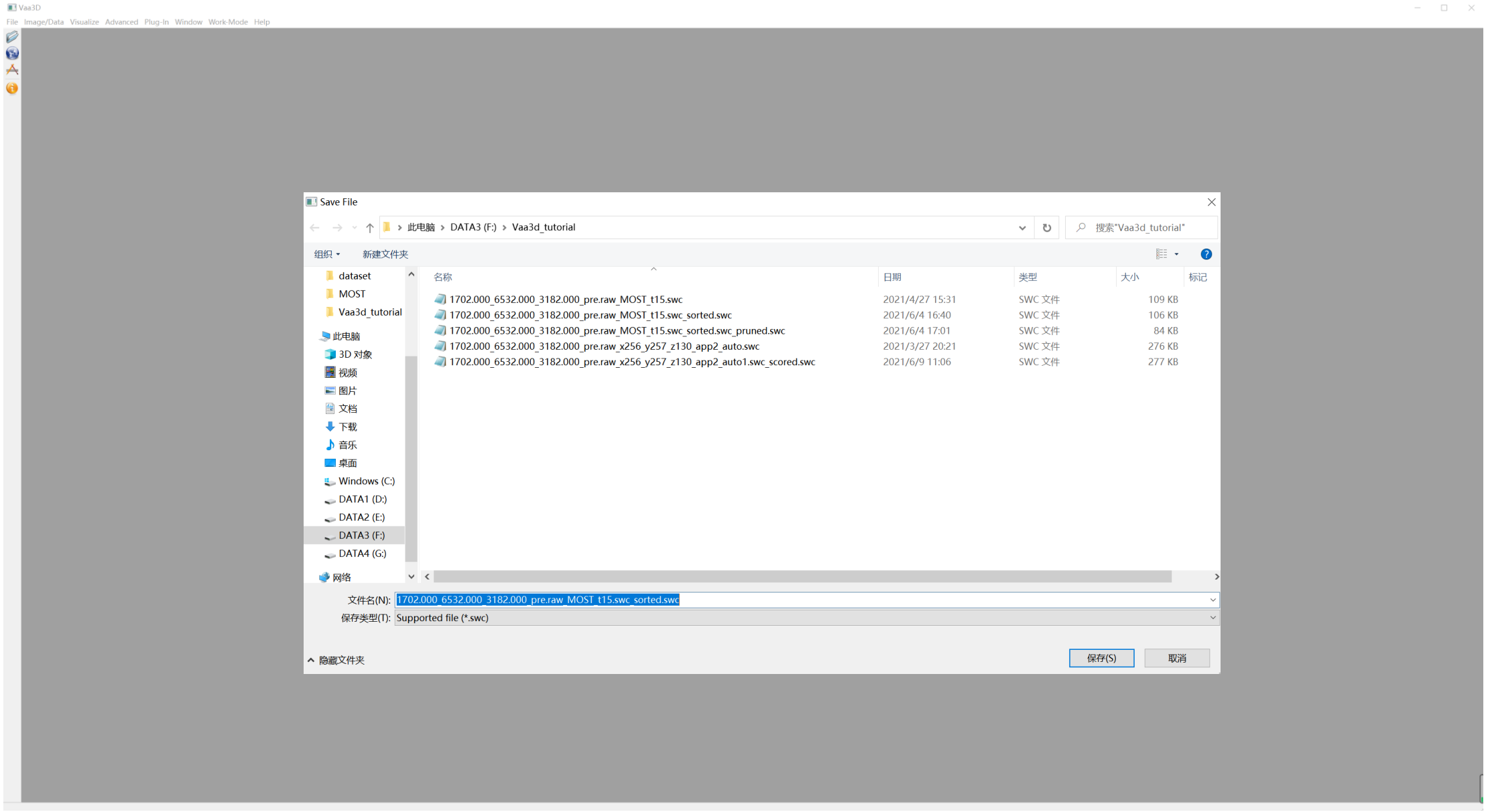
Specify the index of the node to be used as new root.

Sort swc via GUI



The distance to be allowed for two newly connected nodes.

Sort swc via GUI



Example Result

1702.000_6532.000_3182.000_pre.raw_MOST_t15.swc - 记事本

文件(F) 编辑(E) 格式(O) 查看(V) 帮助(H)

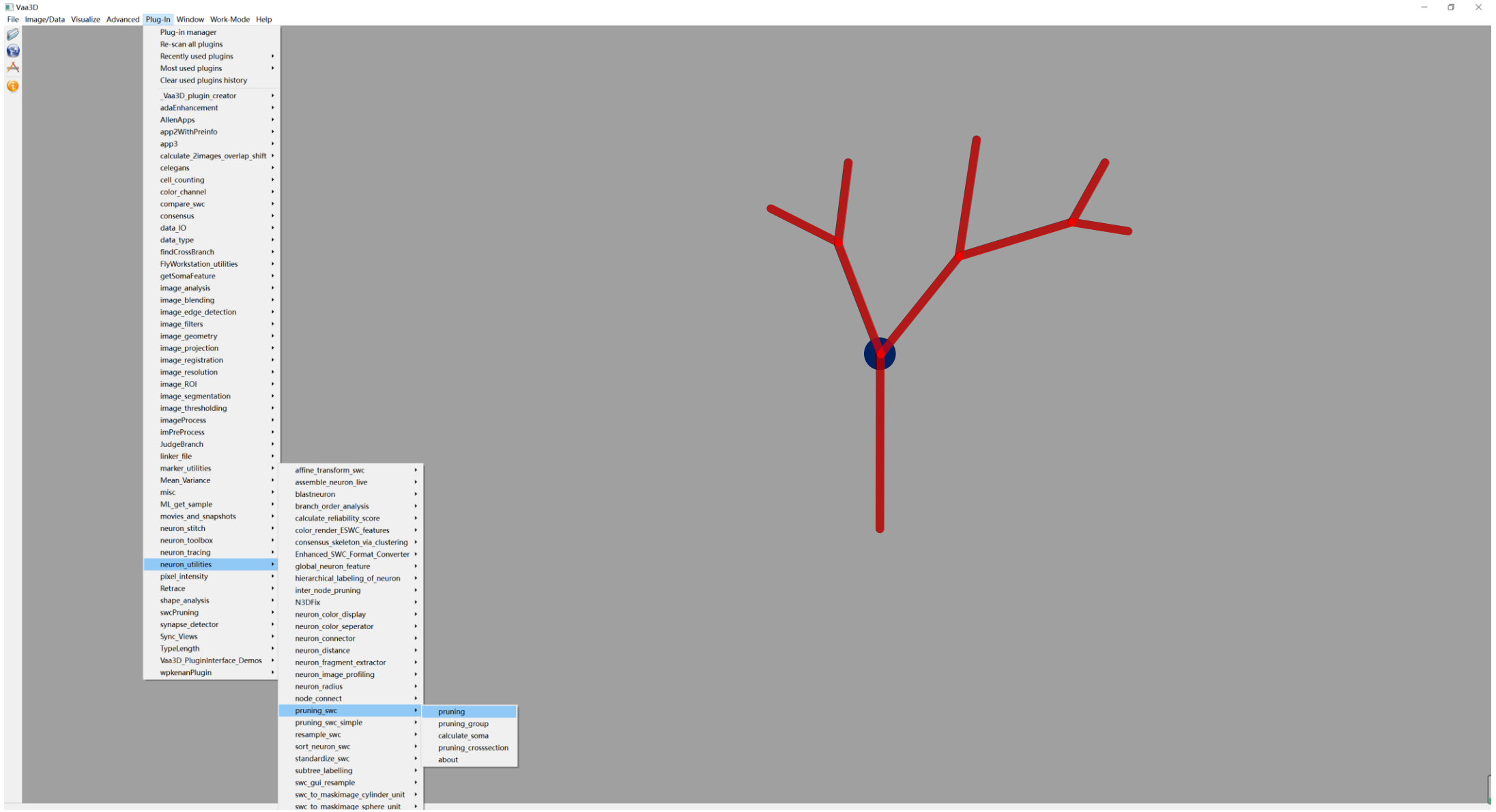
```
#name MOST_Tracing
#comment
##n,type,x,y,z,radius,parent
1 3 1.000 138.000 201.000 2.000 -1
2 3 1.800 138.646 201.031 2.000 1
3 3 3.024 142.238 199.952 1.500 2
4 3 3.977 144.409 199.568 1.250 3
5 3 5.310 146.429 199.405 1.125 4
6 3 6.417 148.500 199.306 1.063 5
7 3 8.348 150.478 198.565 1.031 6
8 3 9.000 152.500 198.500 1.016 7
9 3 10.407 154.556 198.111 1.008 8
10 3 11.276 156.552 197.586 1.004 9
11 3 11.758 158.515 196.970 1.002 10
12 3 11.871 160.452 196.581 1.001 11
13 3 12.091 162.545 195.758 1.000 12
14 3 13.182 164.500 194.682 1.000 13
15 3 14.000 166.533 193.822 1.000 14
16 3 15.450 168.550 192.633 1.000 15
17 3 16.625 170.500 191.719 1.000 16
18 3 17.650 172.483 190.950 1.000 17
19 3 18.475 174.475 190.148 1.500 18
20 3 19.745 176.451 188.765 1.250 19
21 3 20.818 178.545 187.636 1.125 20
22 3 22.038 180.500 186.635 1.563 21
23 3 22.618 182.491 186.327 1.781 22
24 3 23.548 184.429 186.000 1.891 23
25 3 25.233 186.467 185.133 1.445 24
26 3 25.886 188.543 184.457 1.223 25
27 3 27.000 190.542 183.313 1.611 26
28 3 28.750 192.456 182.265 1.306 27
29 3 31.021 194.277 181.000 2.153 28
30 3 33.591 196.018 179.709 2.076 29
31 3 36.121 197.822 177.879 2.038 30
```

1702.000_6532.000_3182.000_pre.raw_MOST_t15.swc_sort

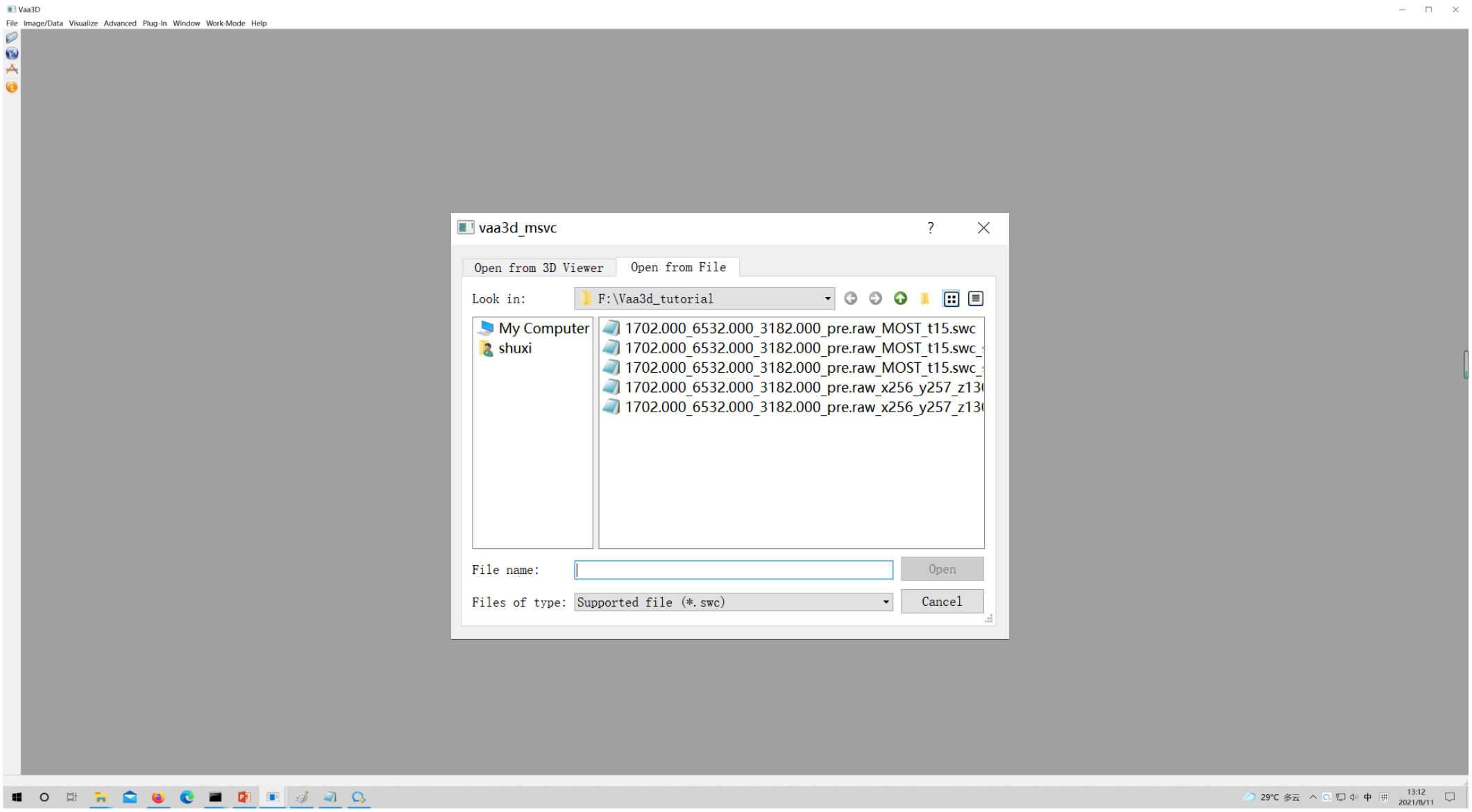
文件(F) 编辑(E) 格式(O) 查看(V) 帮助(H)

```
# generated by Vaa3D Plugin sort_neuron_swc
# source file(s): F:/Vaa3d_tutorial/1702.000_653
# id,type,x,y,z,r,pid
1 3 256.6 256.234 127.111 4.266 -1
2 3 258.943 257.693 127.028 4.531 1
3 3 260.793 256.027 127.5 4.063 2
4 3 262.592 256.224 127.924 3.125 3
5 3 264.507 256.364 128.396 3.25 4
6 3 266.488 255.837 128.576 2.5 5
7 3 268.726 255.164 128.712 2 6
8 3 270.492 256.033 133.27 1 7
9 3 272.516 255.59 134.336 3.052 8
10 3 274.473 254.427 133.673 3.105 9
11 3 276.473 253.352 133.286 2.21 10
12 3 279.418 253.022 132.571 1.42 11
13 3 282.007 250.248 133.248 1.84 12
14 3 282.81 251.43 133.886 1.68 13
15 3 284.417 253.854 134.729 1.359 14
16 3 286.545 255.697 135.364 1.718 15
17 3 288.69 257.048 135.833 1.436 16
18 3 291.055 258.255 136.182 1.872 17
19 3 292.981 259.83 136.679 1.744 18
20 3 295.755 264.638 137.489 1.488 19
21 3 300.711 269.132 138.86 1.977 20
22 3 302.021 267.715 140.535 2.953 21
23 3 302.716 266.938 142.481 2.906 22
24 3 301.794 267.221 144.529 2.813 23
25 3 302.829 265.095 146.514 2.625 24
26 3 303.081 264.244 148.488 3.25 25
27 3 306 266.767 152.401 3.5 26
28 3 305.239 267.304 153.79 4 27
29 3 304.464 266.75 155.768 1.252 28
30 3 302.56 264.84 156.64 1.504 29
31 3 300.444 263.63 156.222 2.009 30
```

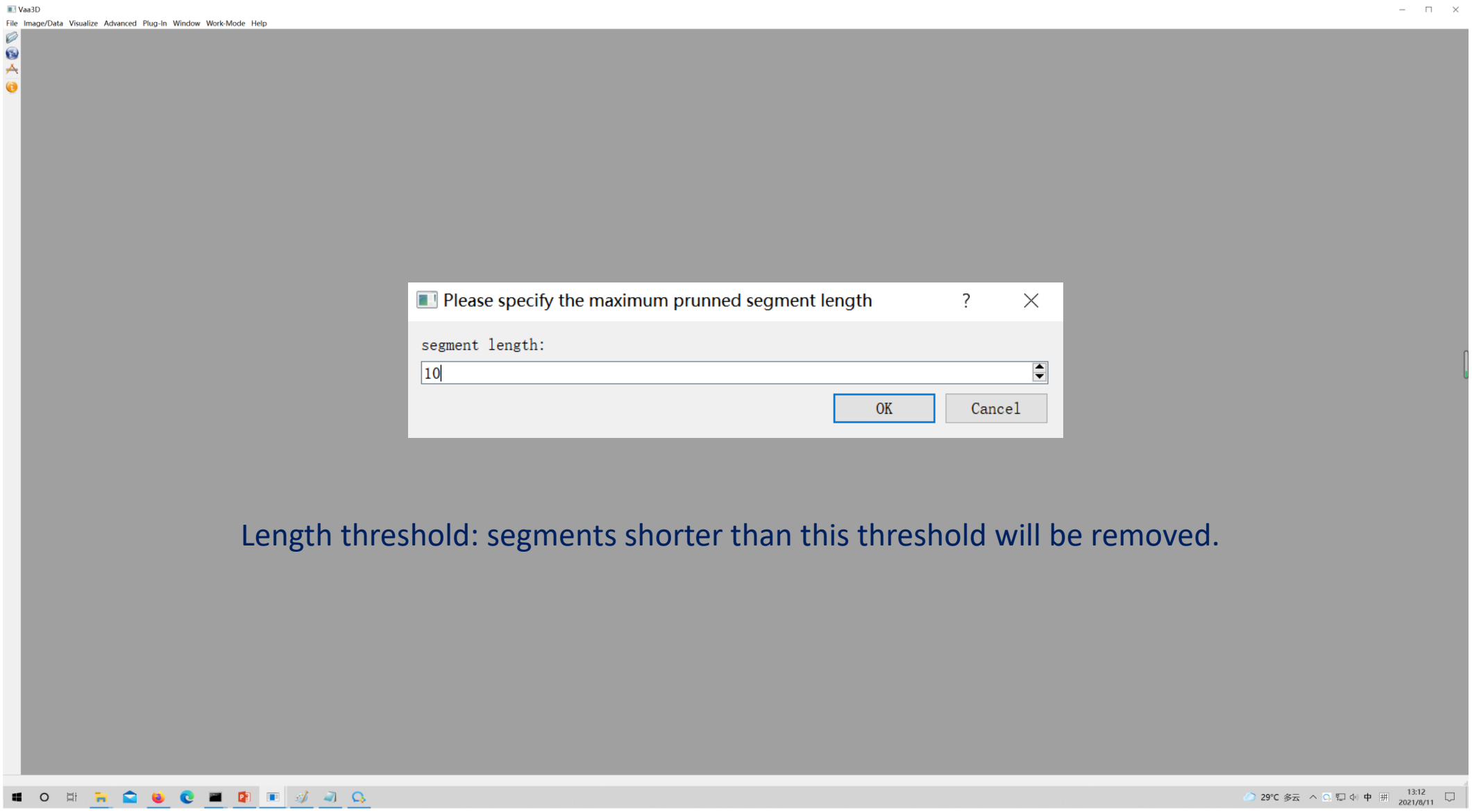
prun .swc files



prun .swc files

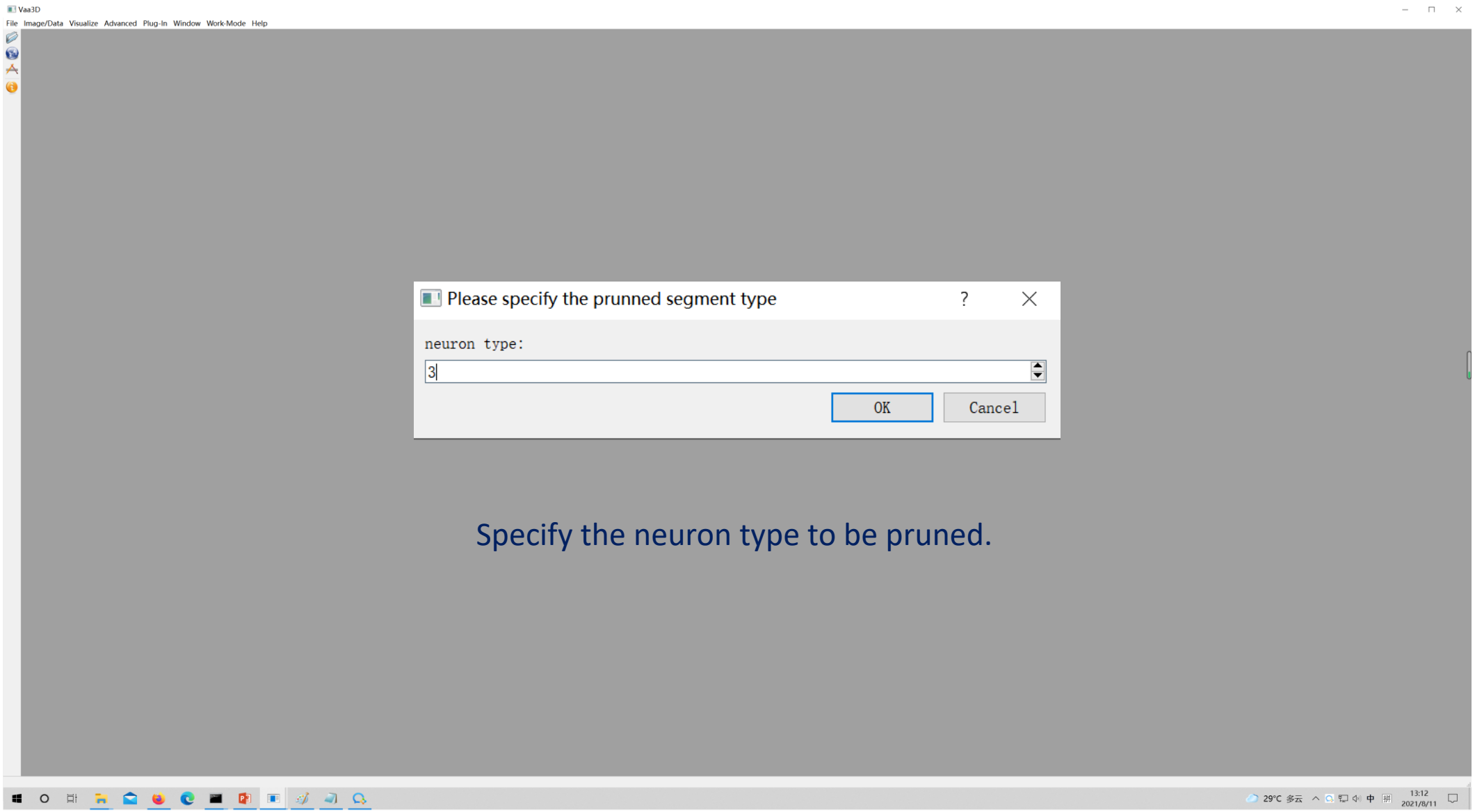


prun .swc files



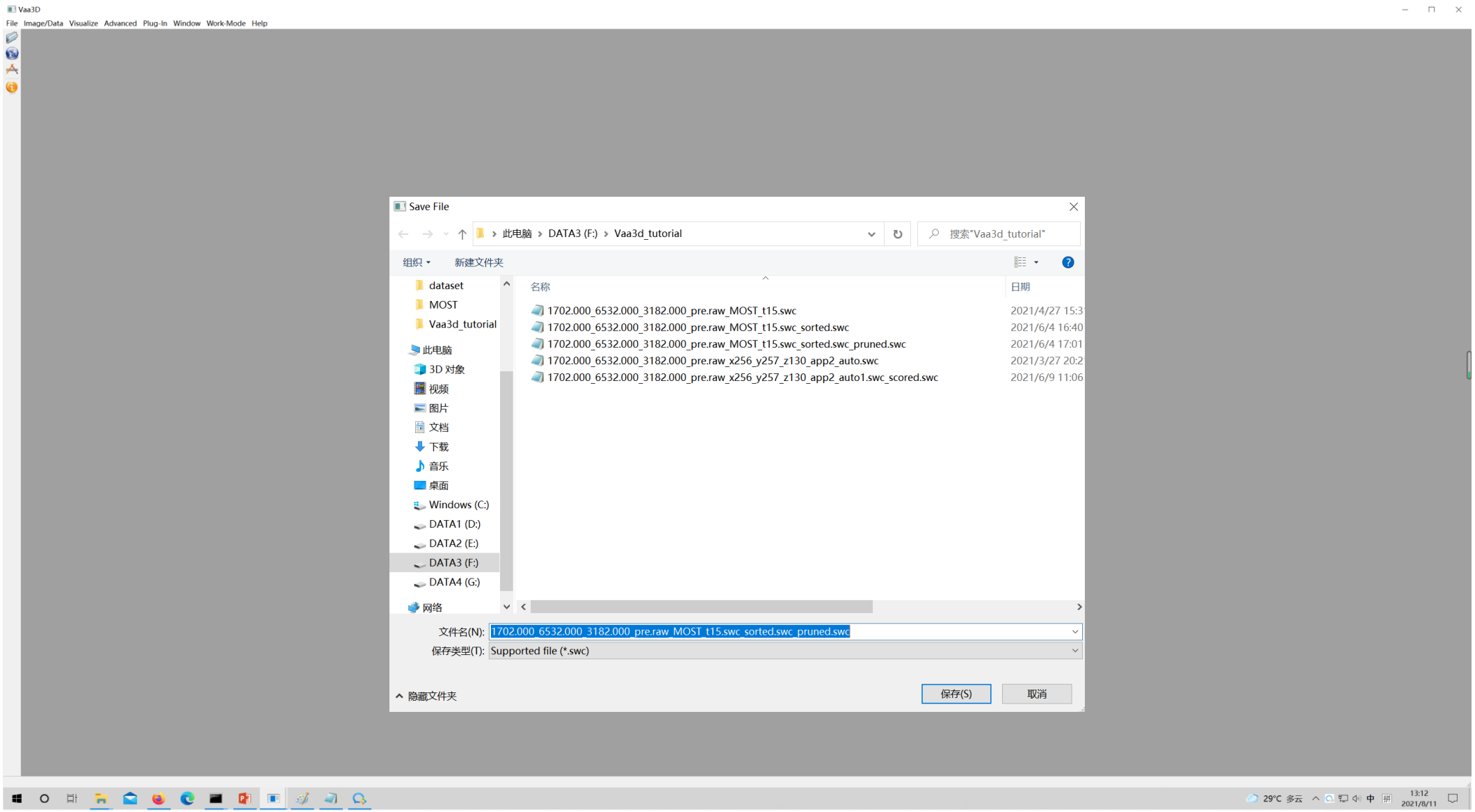
Length threshold: segments shorter than this threshold will be removed.

prun .swc files

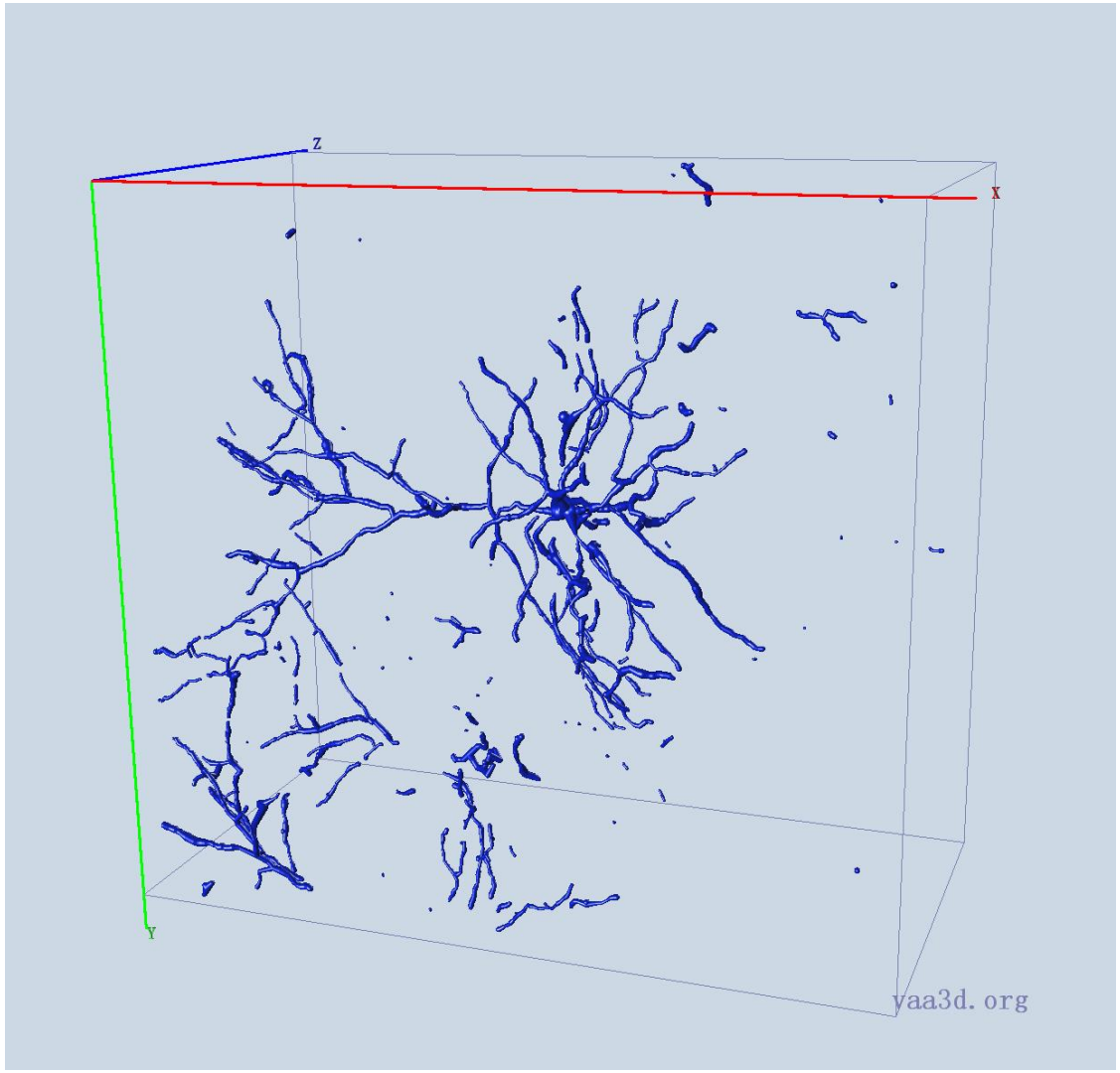


Specify the neuron type to be pruned.

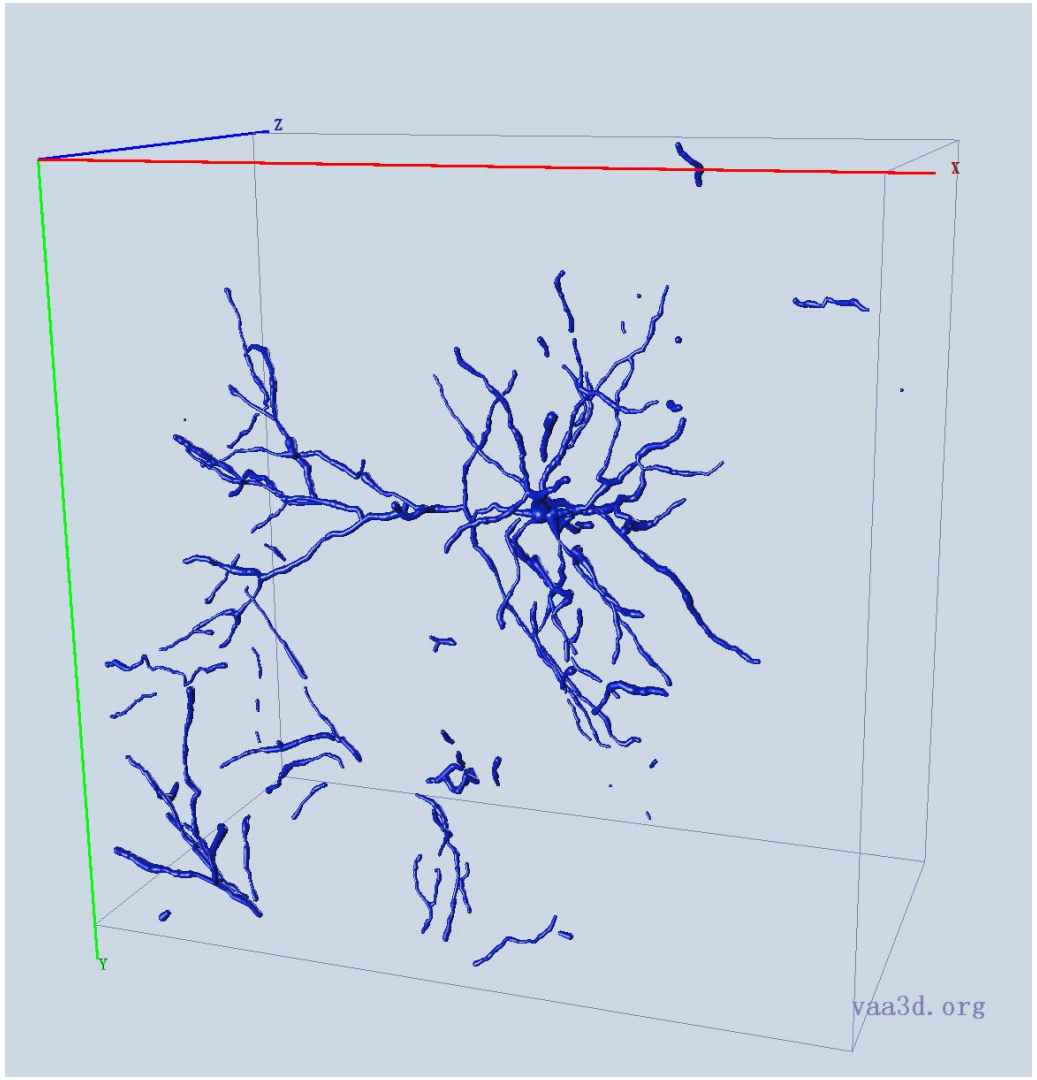
prun .swc files



Results

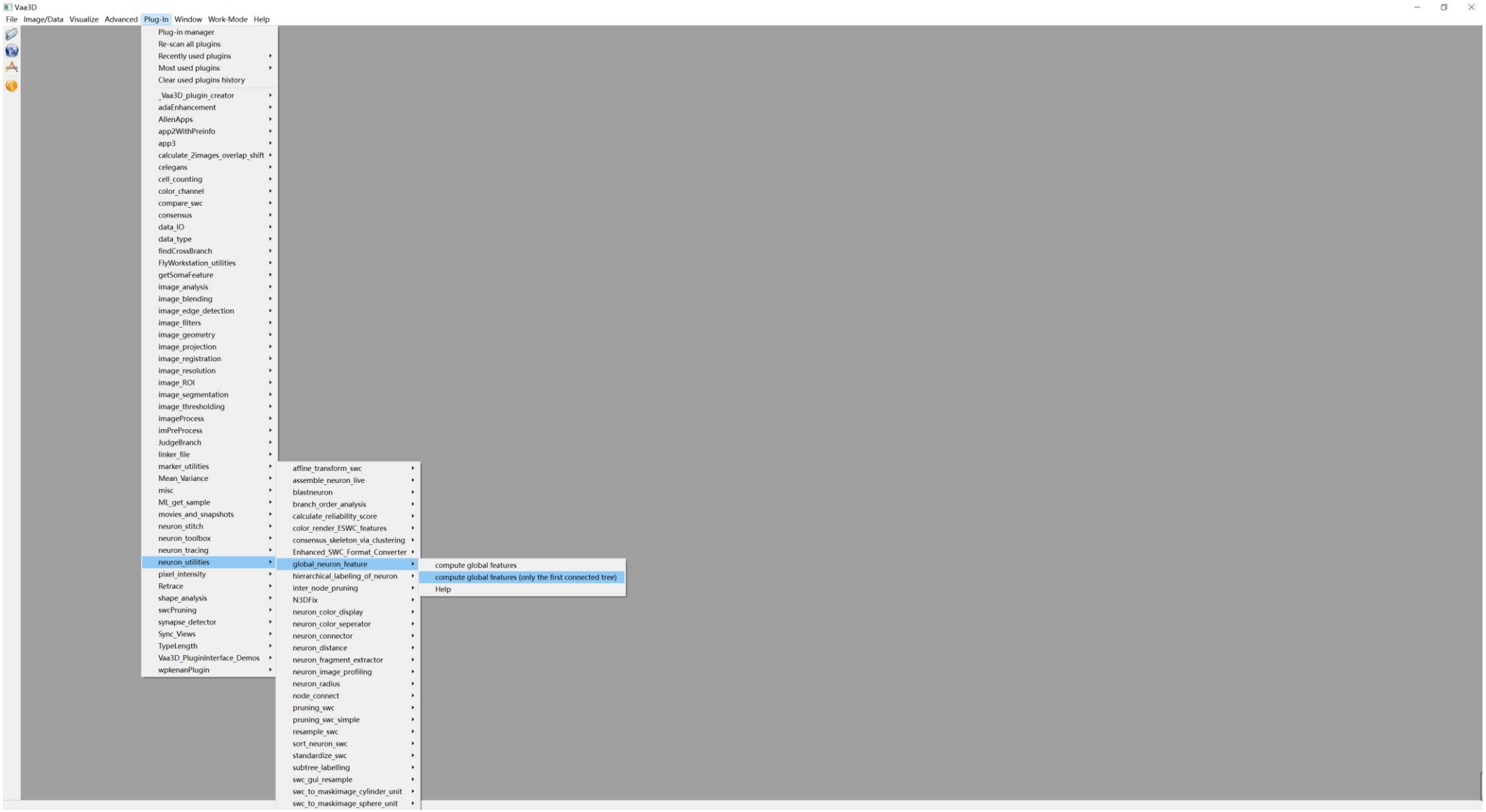


raw

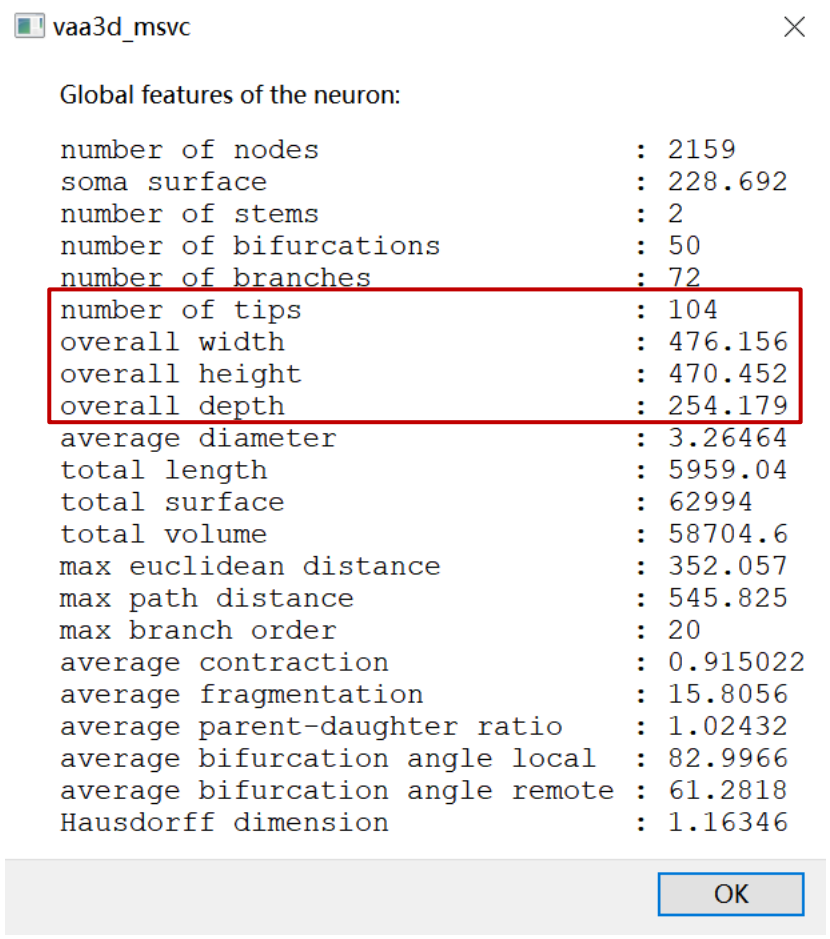


pruned

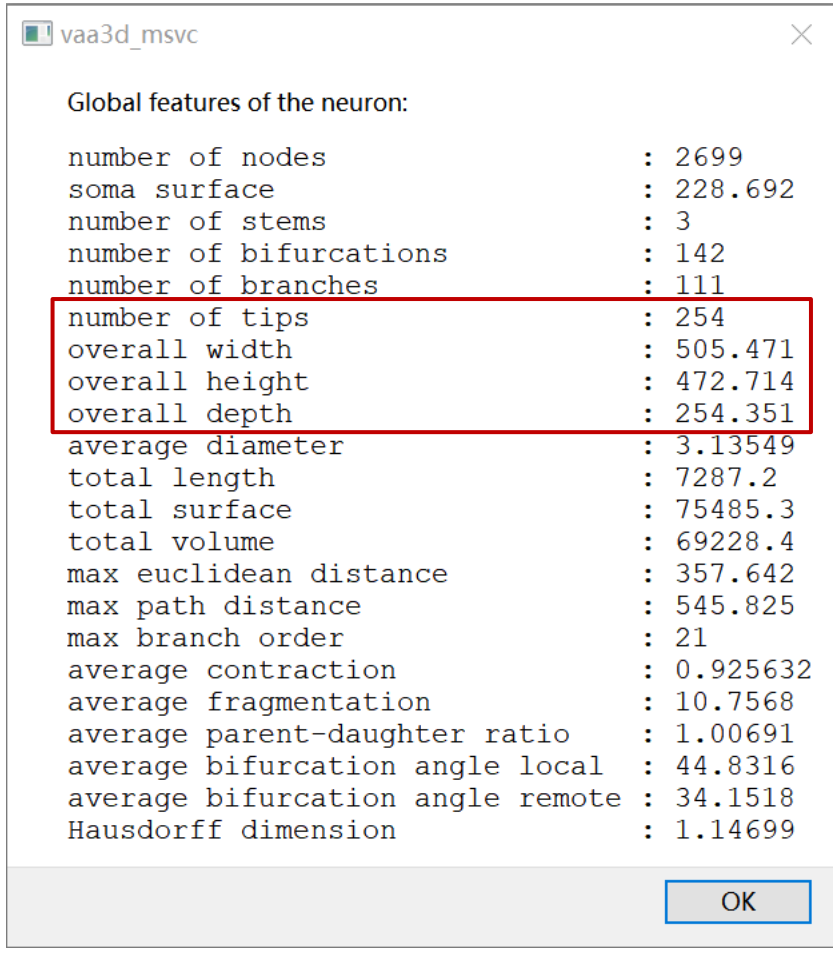
Global Features



Global Features



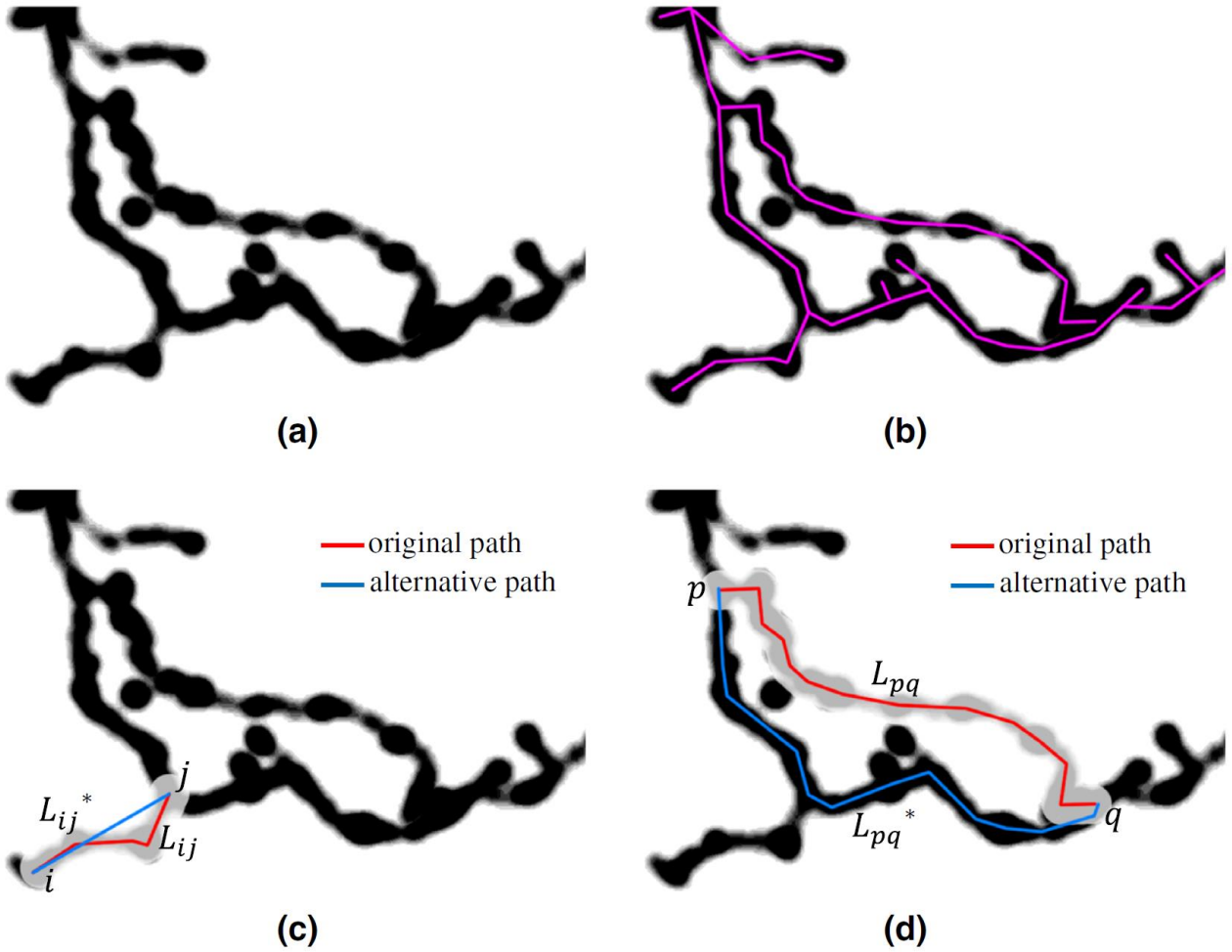
Pre-processed .swc



raw .swc

Confidence (Reliability) Score

Fig. 2 Illustration of alternative path. For each segment in the reconstructions, after masking the image along the segment, the alternative path will be searched by fast marching from one end to the other end of the segment based on intensity. **a** neuron to reconstruct, **b** initial reconstructions, **c** alternative path of L_{ij}, **d** alternative path of L_{pq}



- To what extend can I trust the reconstruction
- Whether or not a segment is trustworthy can be tested by checking if there is an alternative path connecting the two ends of the segment.
- A segment with no better alternative pathway is more reliable in comparison with a segment with alternative pathway.

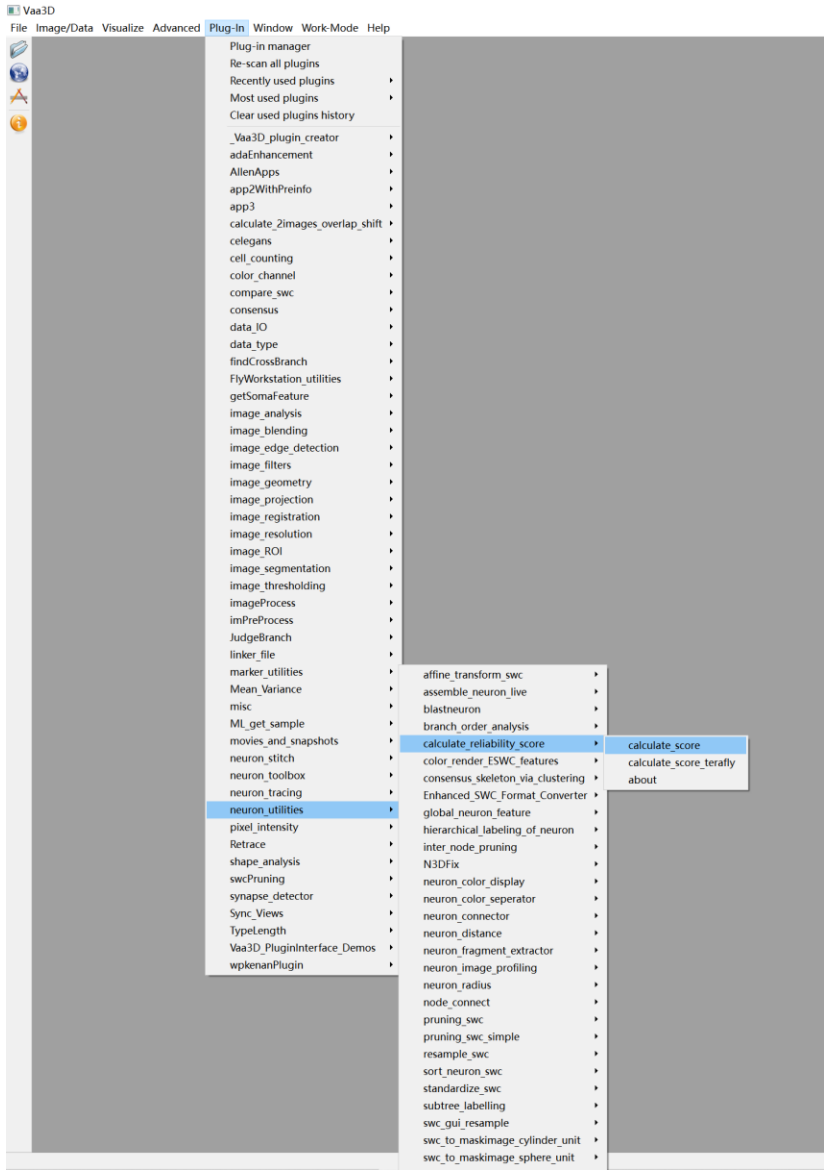
Brain Informatics (2015) 2:135–144

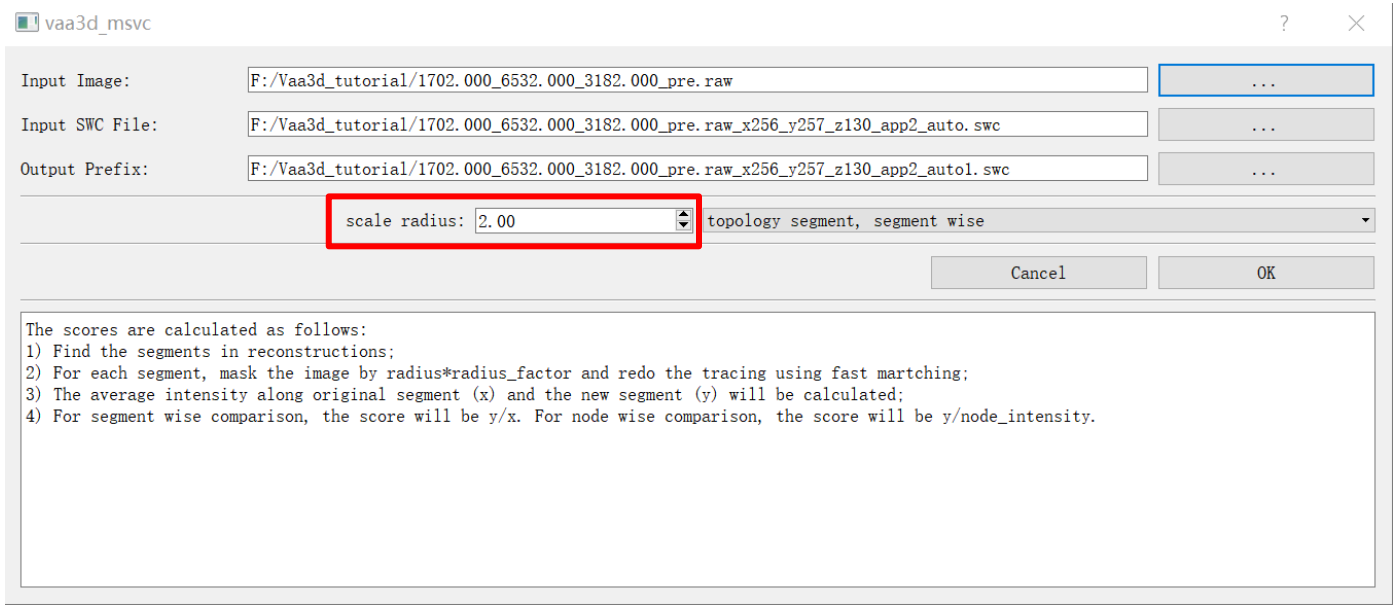
Confidence Score

Steps to calculate the score:

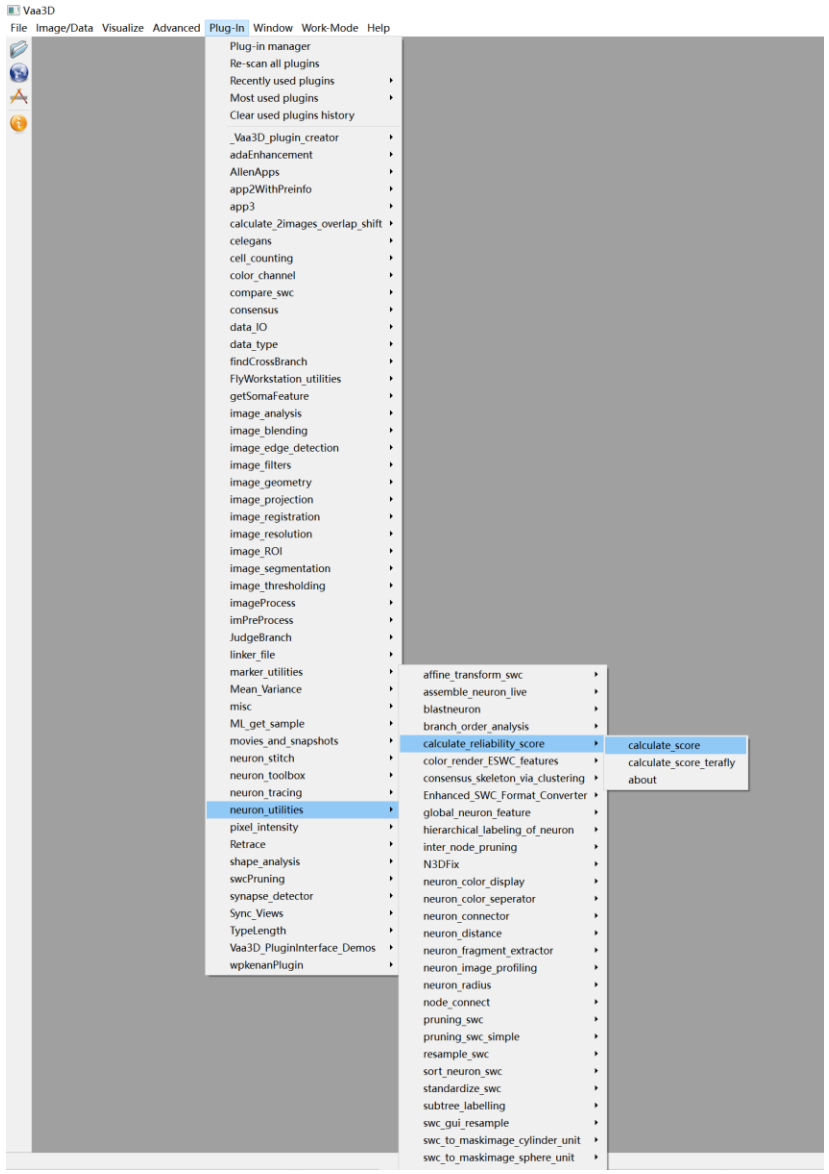
1. Find the segments in reconstructions;
2. For each segment, mask the image by radius*radius_factor and redo the tracing using fast marching;
3. The average intensity along original segment (x) and the new segment (y) will be calculated;
4. For segment wise comparison, the score will be y/x; For node wise comparison, the score will be y/node_intensity.

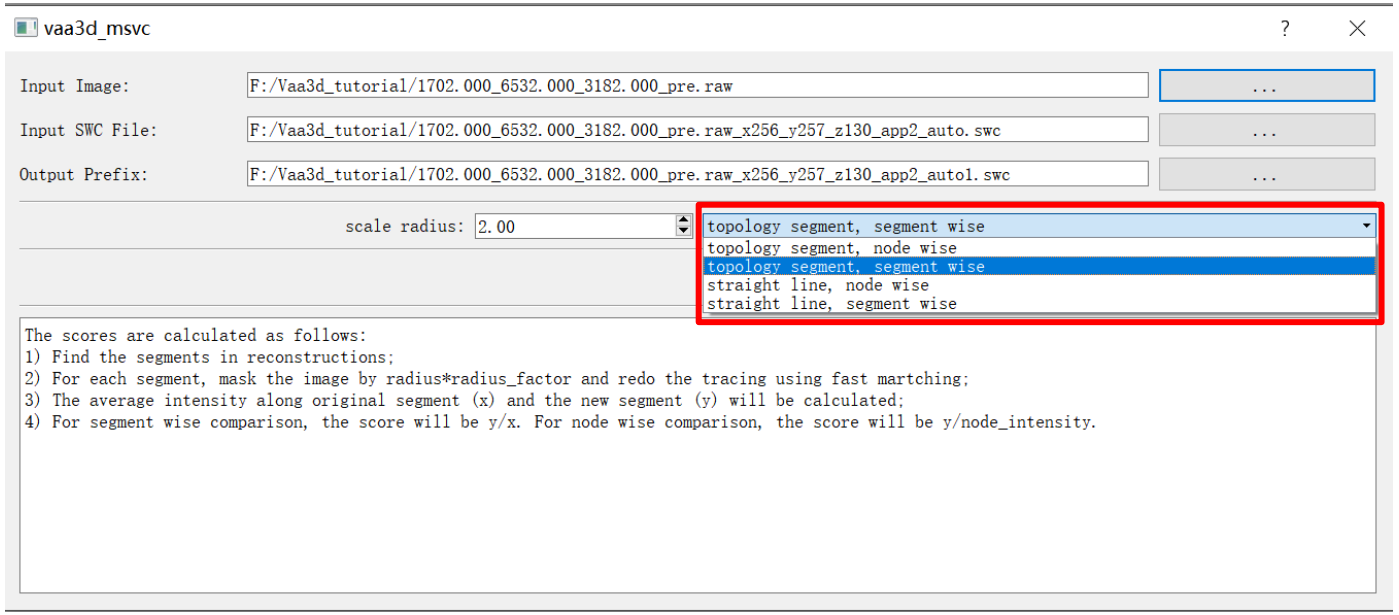
Calculate Confidence Score in Vaa3D



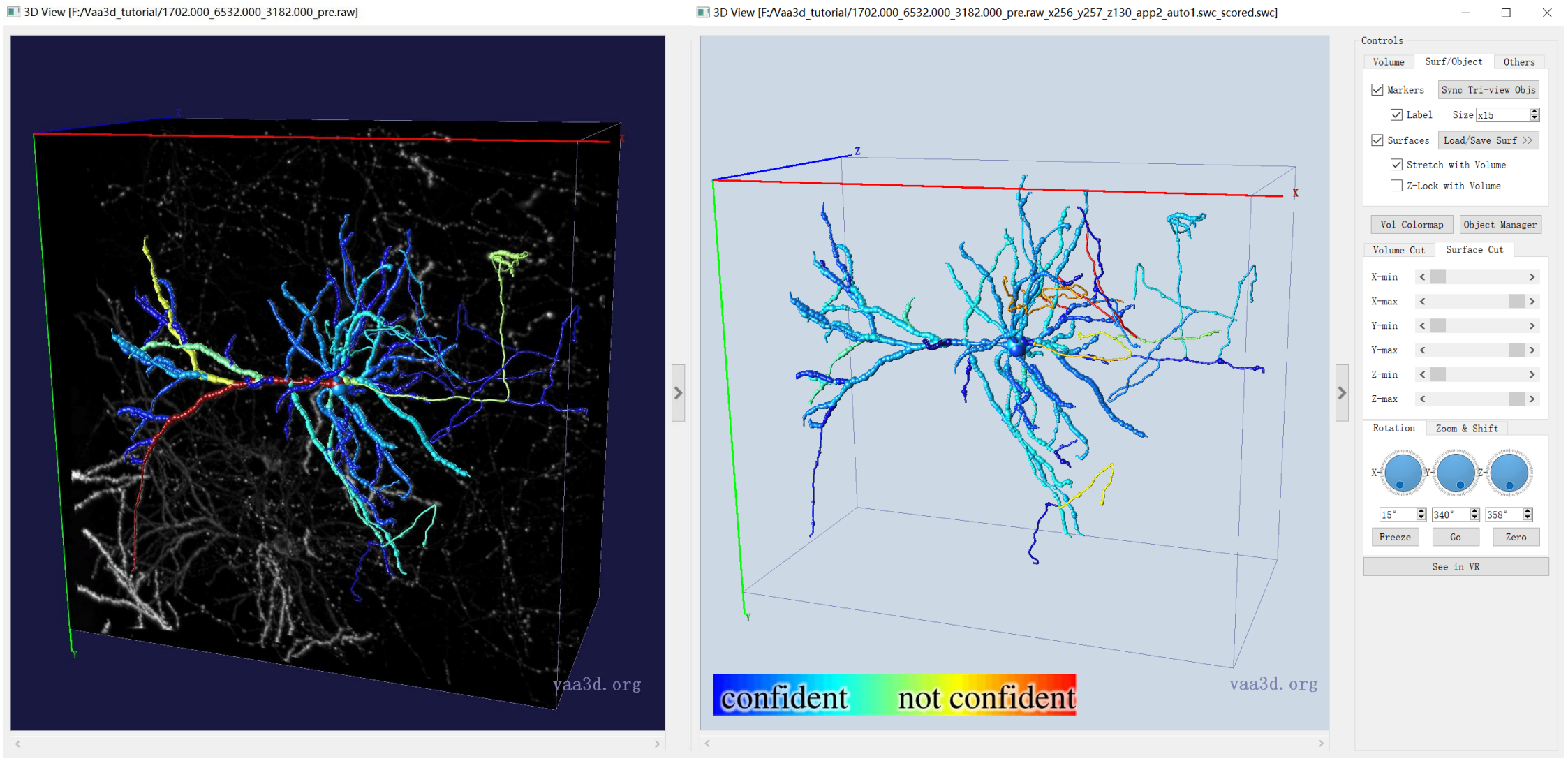


Calculate Confidence Score in Vaa3D

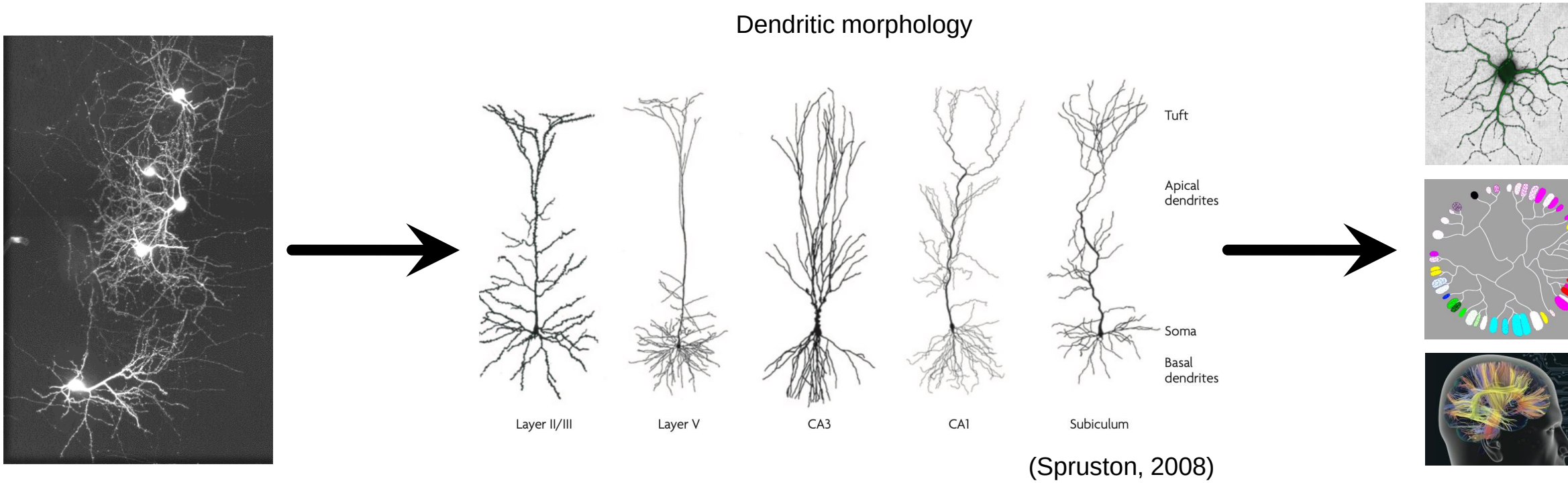




Calculate Confidence Score in Vaa3D



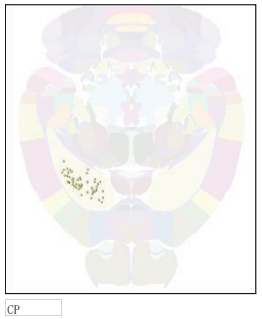
Key Questions in Neuroscience



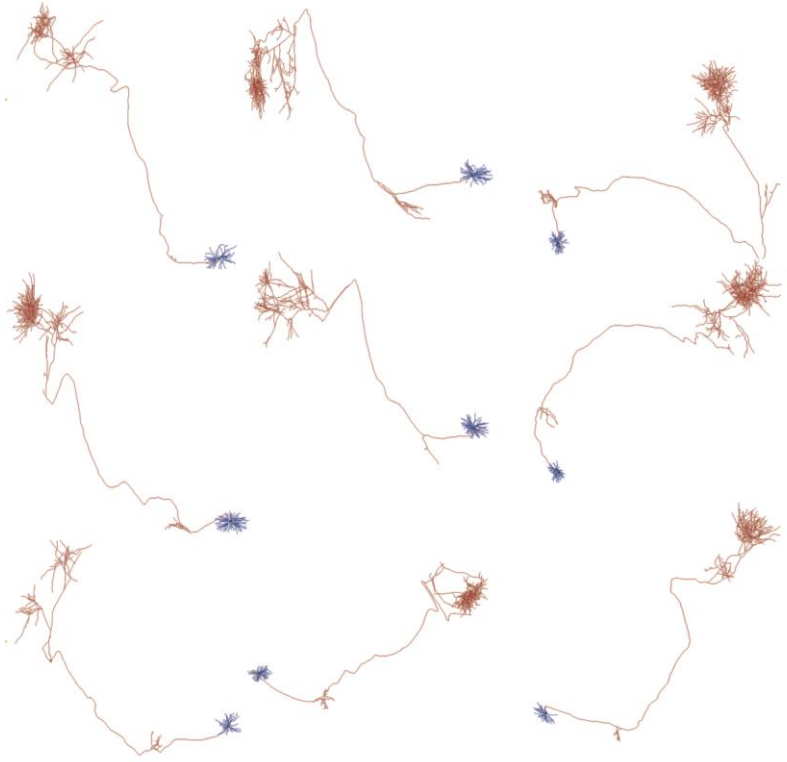
- Quantify/characterize neuron morphologies

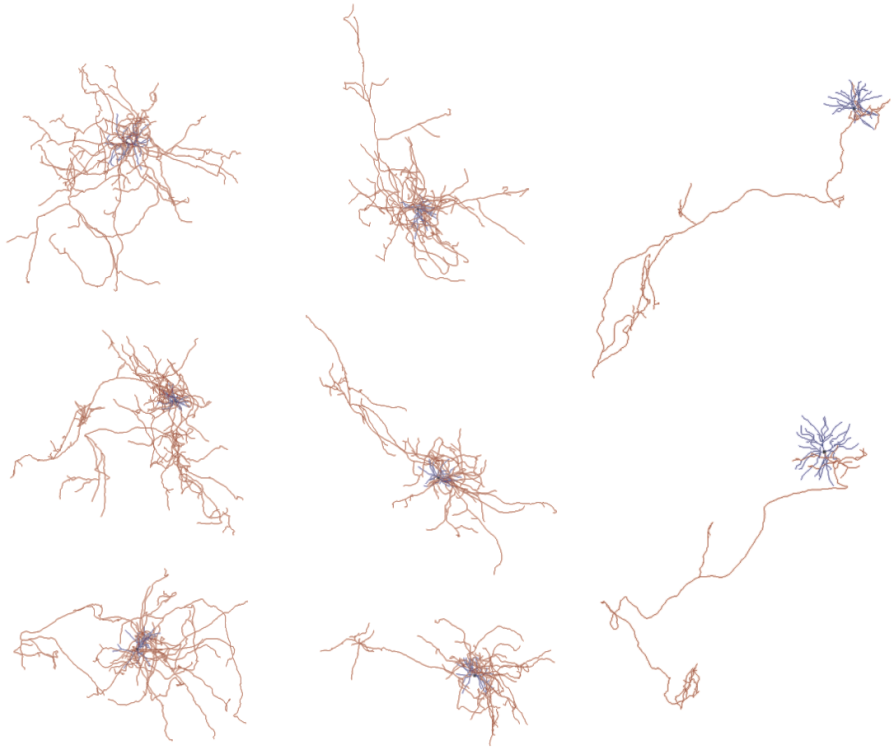
Single Neurons

VPM



CP

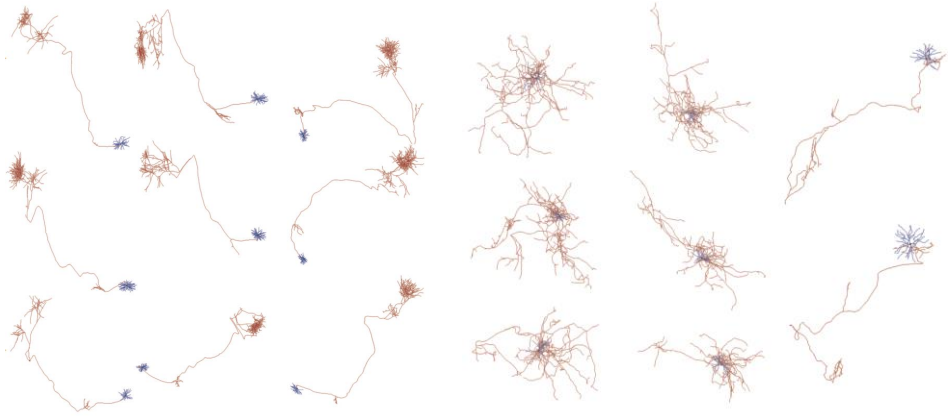


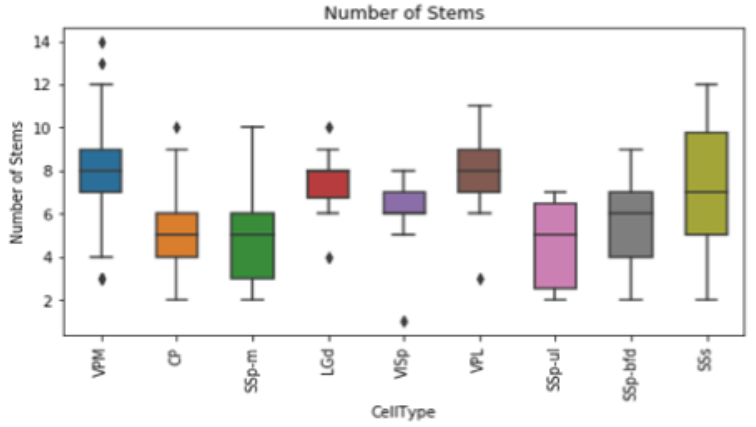


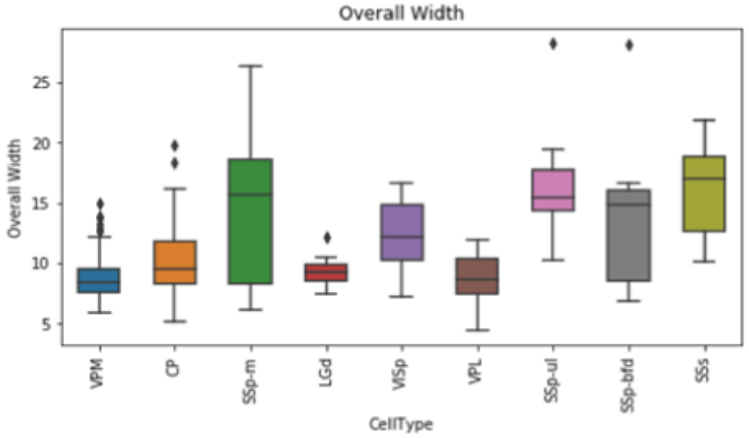
Morphological Features

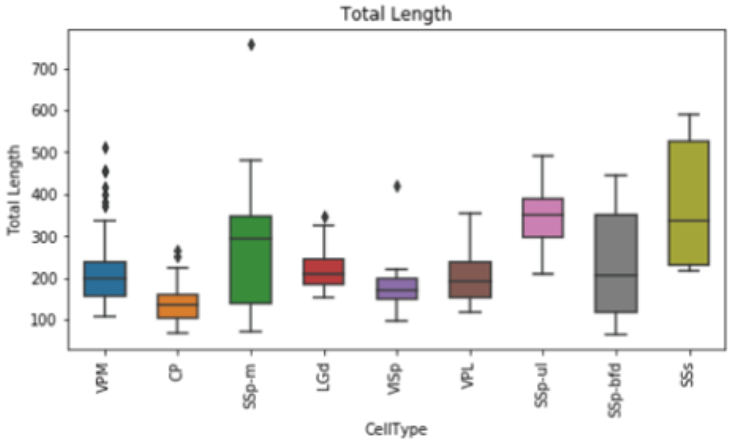
VPM

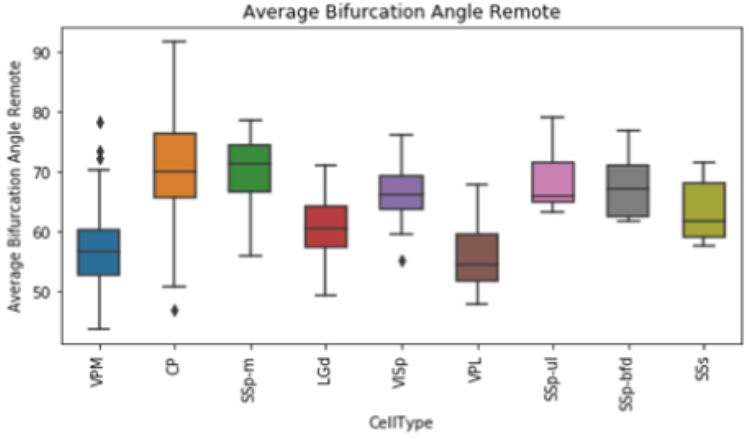
CP



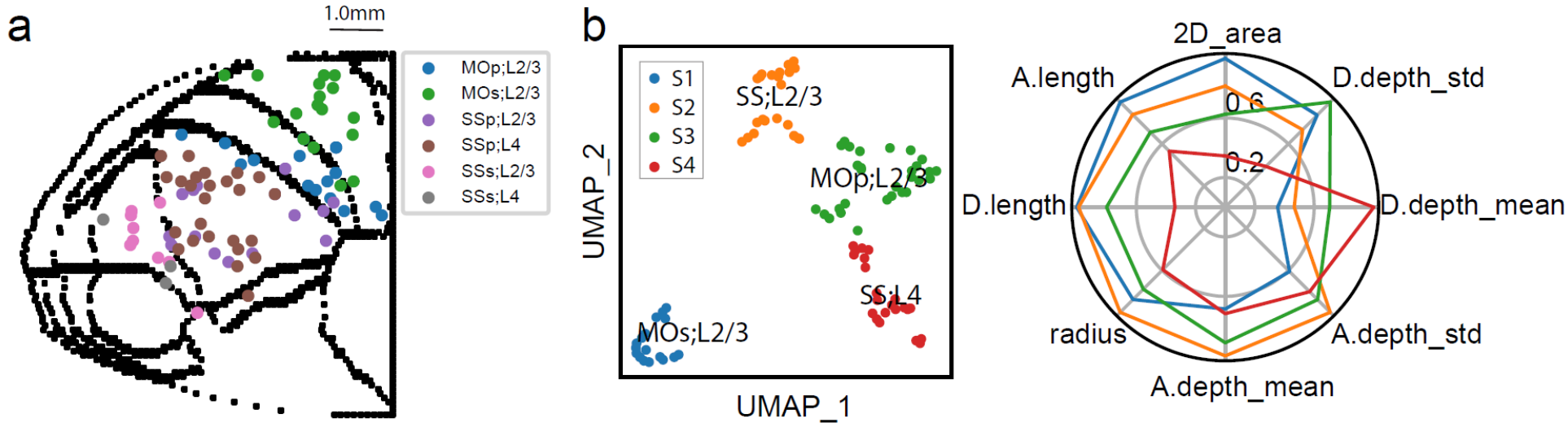


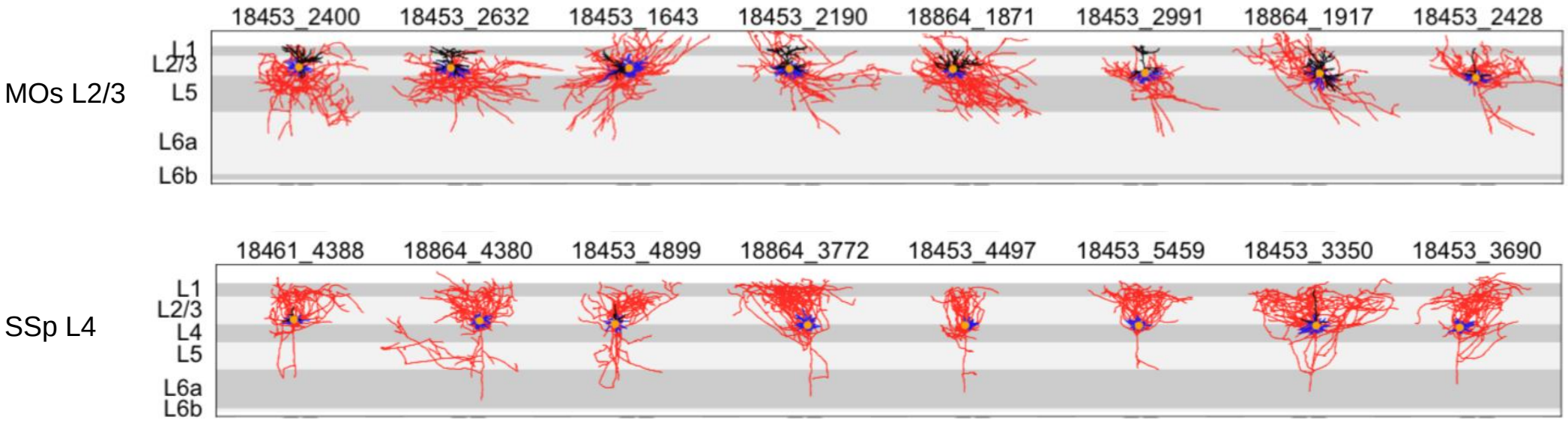






Local Arborization Patterns of Cortical L2/3/4 IT Neurons

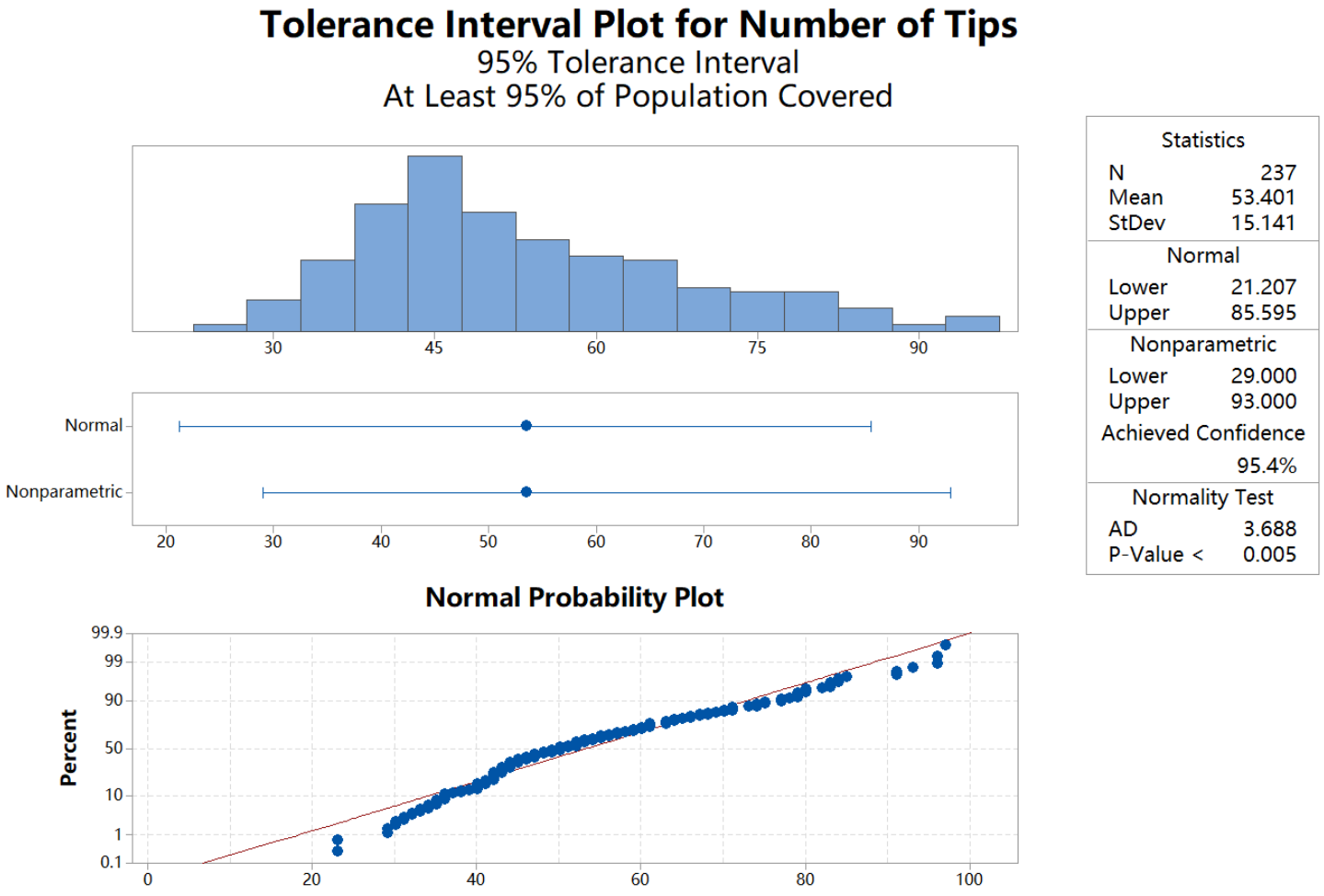




(Peng *et al*, 2020)

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
- To check the reliability of reconstructed segments/nodes

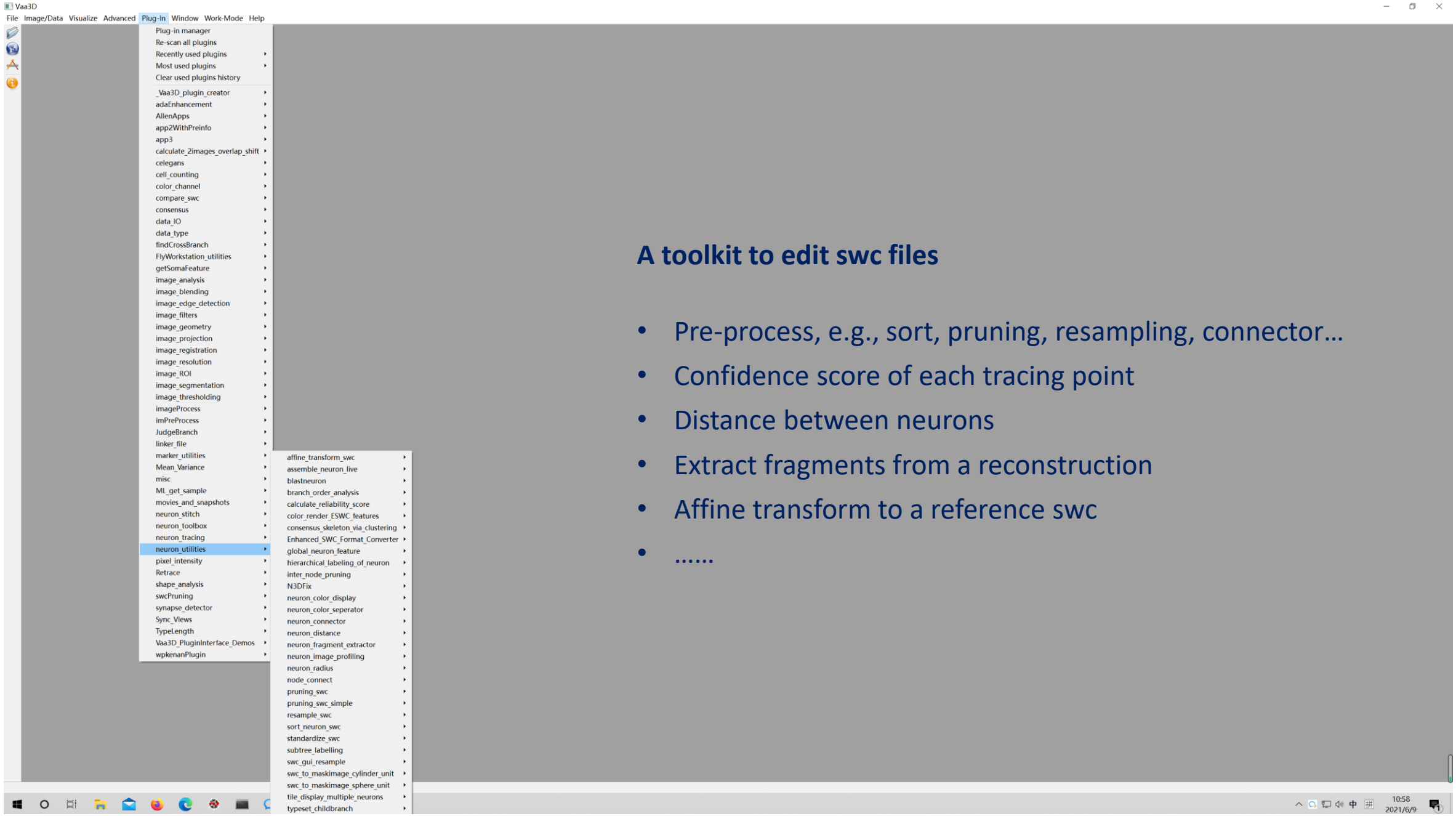
- Brief showcase where we can make use of the morphological features

Reference list

- L-Measure: http://cng.gmu.edu:8080/Lm/help/index.htm
- H. Chen *et al*, SmartTracing: self-learning-based Neuron reconstruction, *Brain Informatics*, 2015, 2:135–144, DOI: 10.1007/s40708-015-0018-y
- H. Peng *et al*, Brain-wide single neuron reconstruction reveals morphological diversity in molecularly defined striatal, thalamic, cortical and claustral neuron types, 2020, DOI: 10.1101/675280

Find More from Source Code

- vaa3d_tools/released_plugins/v3d_plugins/global_neuron_feature/
- vaa3d_tools/released_plugins/v3d_plugins/sort_neuron_swc/
- vaa3d_tools/released_plugins/v3d_plugins/pruning_swc/
- vaa3d_tools/released_plugins/v3d_plugins/neuron_reliability_score/



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
- ……