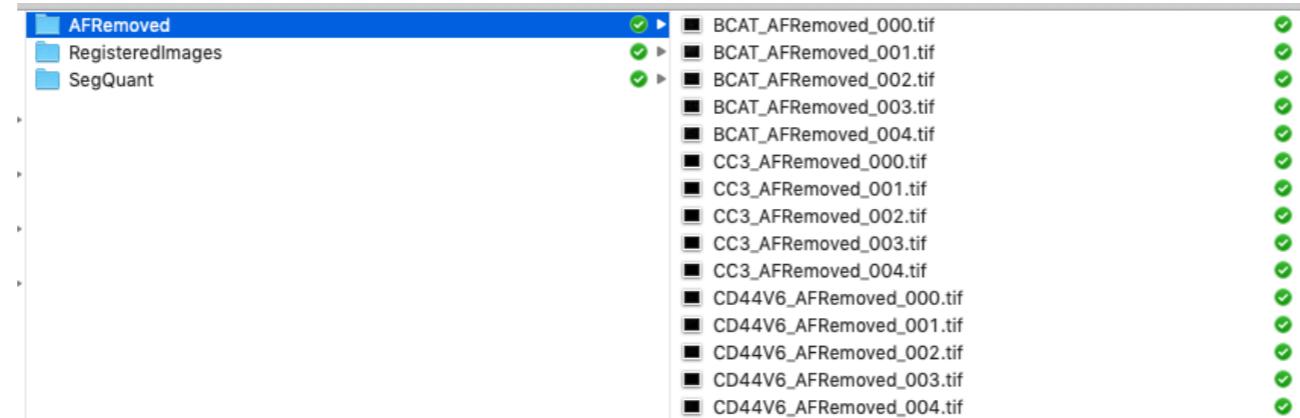


Initialize pipeline by calling CellSegQuant

The slide directory must contain a folder “AFRemoved” containing the registered and background removed images for each marker in the following format:

MarkerName_AFRemoved_XXX.tif

Where MarkerName is the marker and XXX is the position



DAPI images are located in “RegisteredImages/001” and must contain “dapi” in the file name

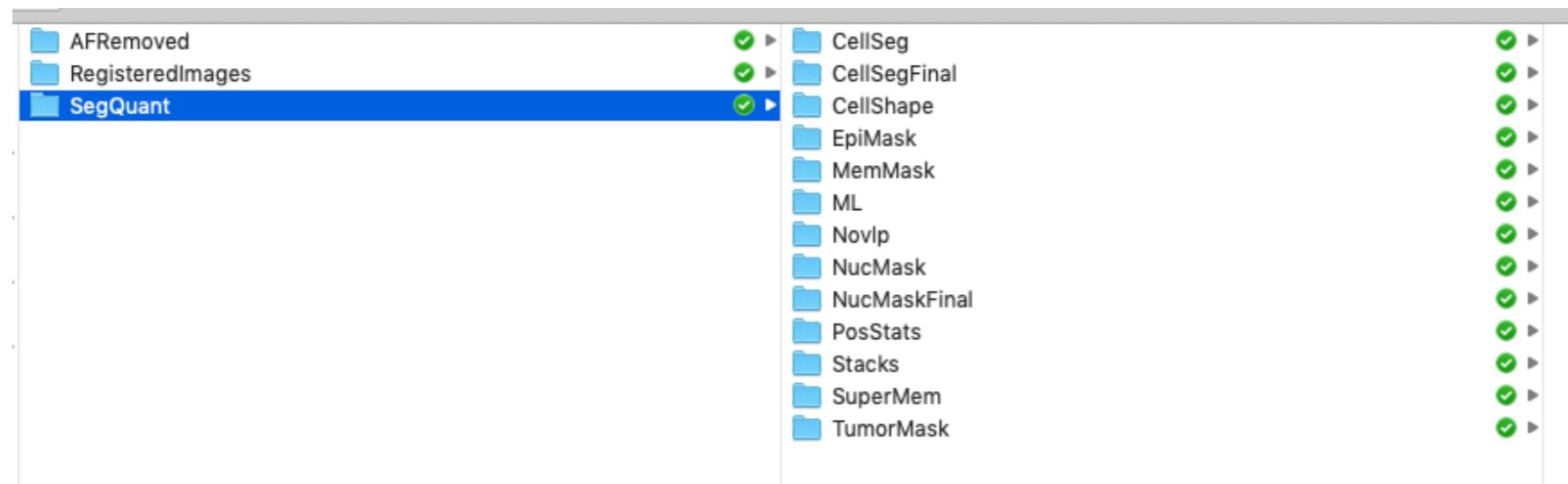


```
warning off
clear all
close all

dir='/Users/etmckinley/Dropbox (VUMC)/Research/Manuscripts/Cell Segmentation/Example MxIF Data';
CellSegQuant(dir ,1,1,1,0, 1)
```

Intermediate steps

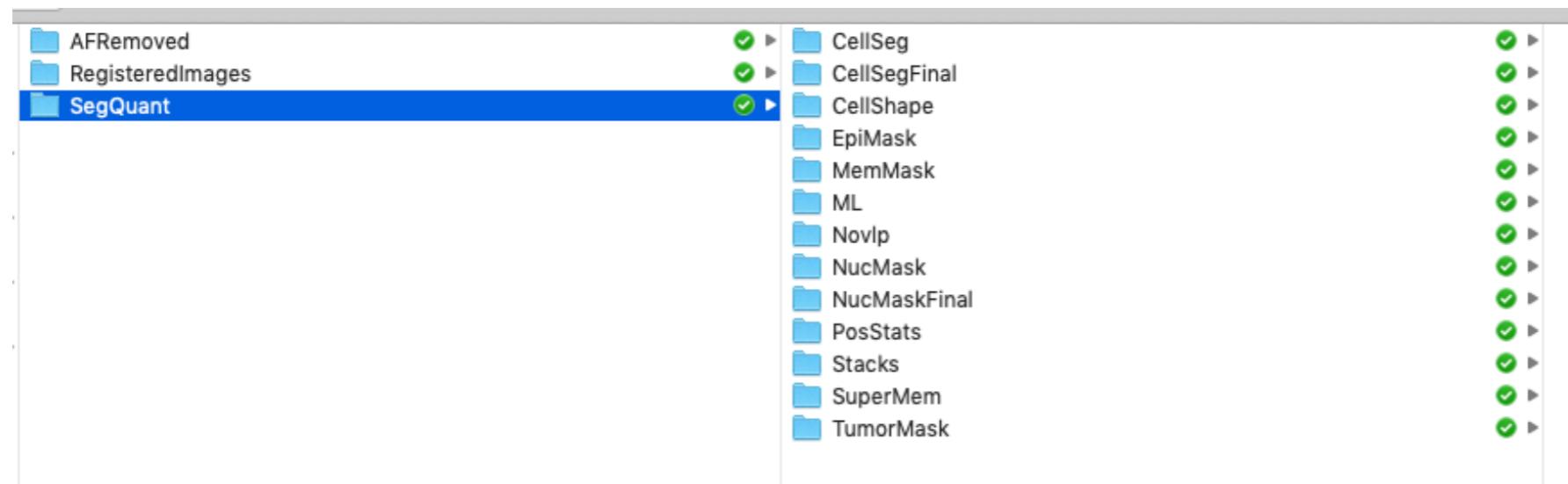
The script will create and populate a new folder, “SegQuant” to store the results into various folders. No interaction is necessary until prompted in the command window.



The pipeline first generates stacked tiffs of all the markers and DAPI for all positions, but will not continue until probability makes for epithelium/stroma, and membrane/nucleus/cytoplasm are created in Ilastik.

Generating probability masks in Ilastik

The script will create and populate a new folder, “SegQuant” to store the results into various folders. No interaction is necessary until prompted in the command window.

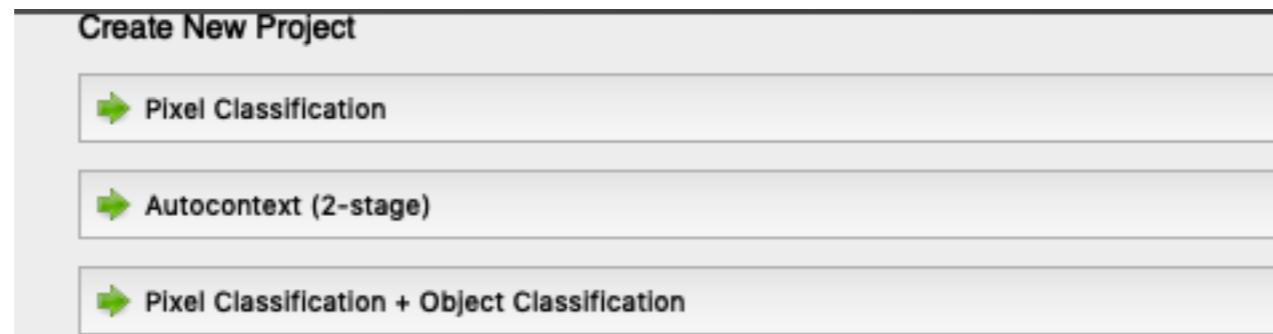


The pipeline first generates stacked tiffs of all the markers and DAPI for all positions, but will not continue until probability makes for epithelium/stroma, and membrane/nucleus/cytoplasm are created in Ilastik.

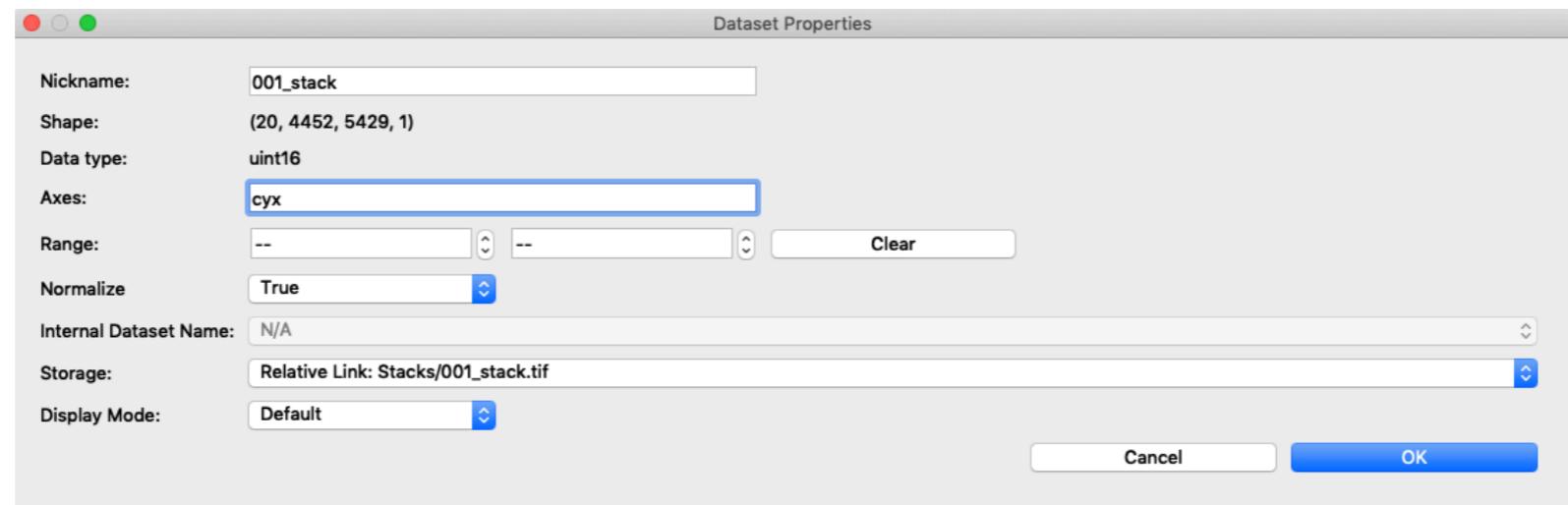
```
000: Stack: 000
No Epithelial Probability File
001: Stack: 001
No Epithelial Probability File
002: Stack: 002
No Epithelial Probability File
003: Stack: 003
No Epithelial Probability File
004: Stack: 004
No Epithelial Probability File
```

Generating probability masks in Ilastik

Once the tiff stacks are generated, open Ilastik and create a new “Pixel Classification” project for epithelial mask and save into the SegQuant folder.

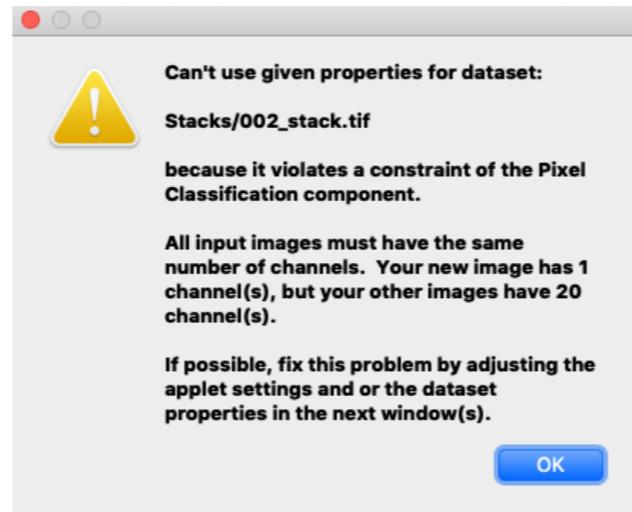


In the Ilastik project, click the “Add New” button to load in your first tiff stack. In the dataset properties pop-up window, change the axes to “cyx”.

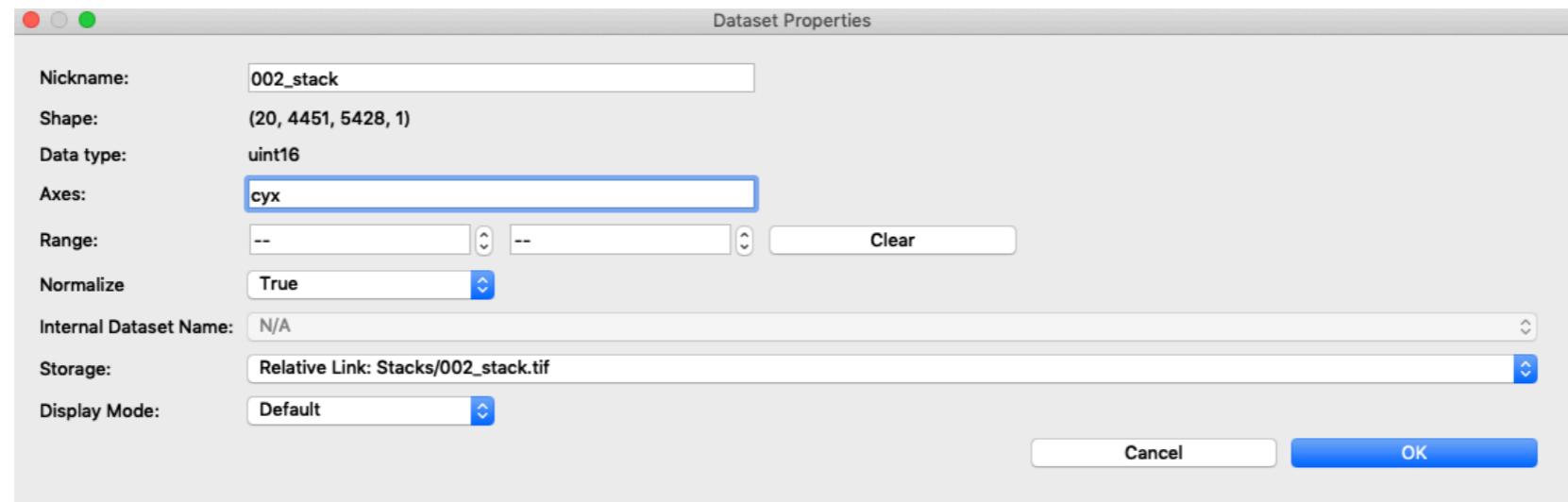


Generating probability masks in Ilastik

After loading the first stack, you can select multiple stacks at a time. Ilastik will give you a warning message stating that the dimensions do not match, press “OK”

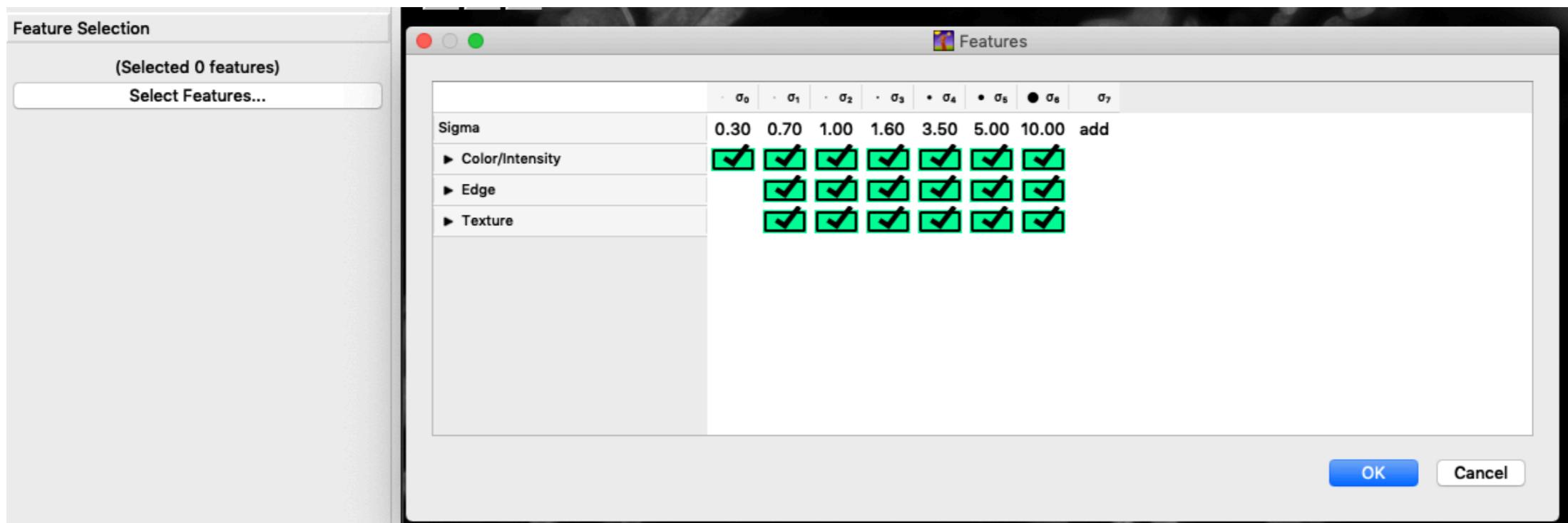


In the dataset properties pop-up window, change the axes to “cyx” as you did for the first stack. Do this for each stack you want to train on.



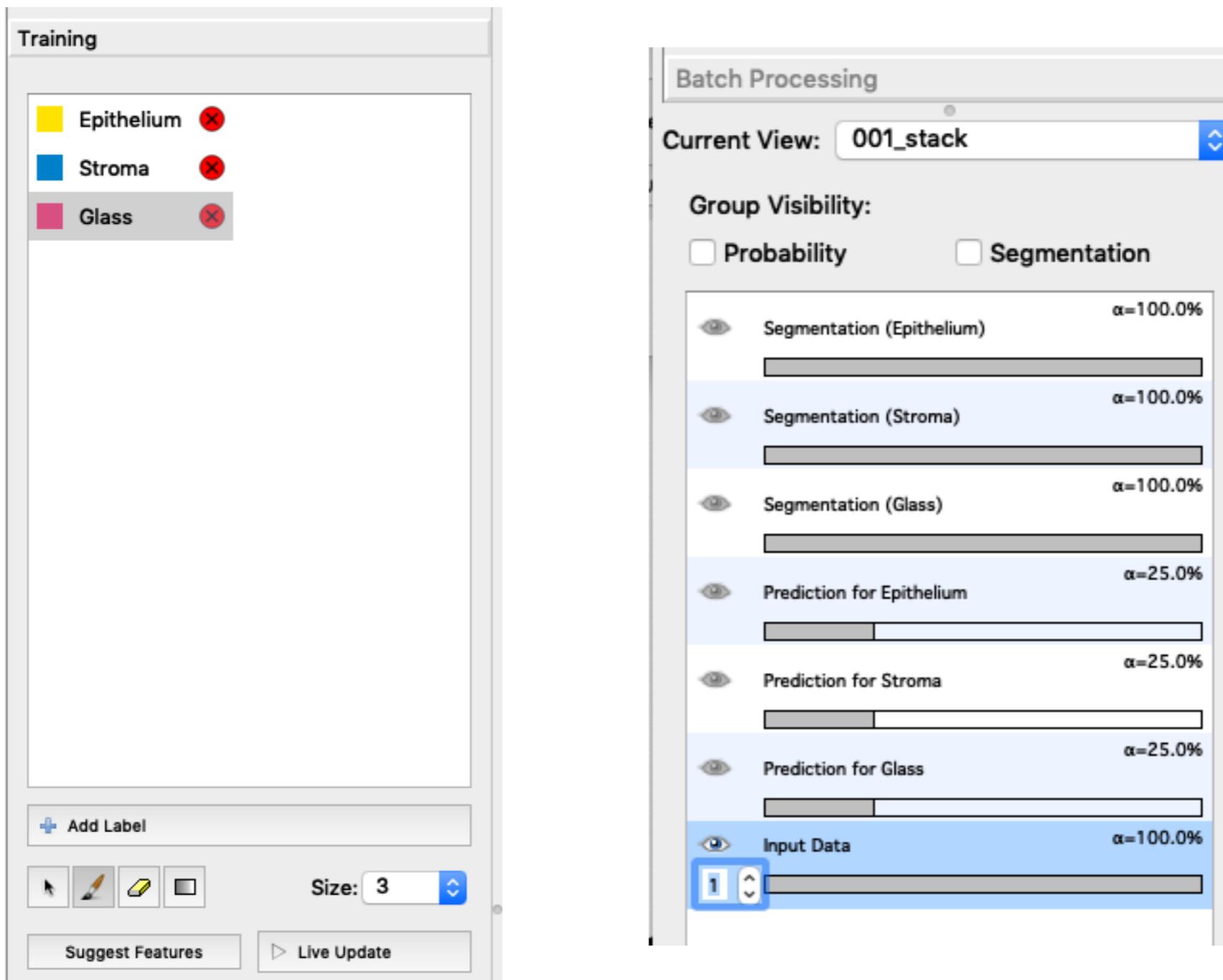
Generating probability masks in Hlastik

Once all stacks are loaded, click on feature selection and select the features you want to train on. Typically this is all of them, but can be modified based on the data set and hardware available.



Generating probability masks in Ilastik

Next, select the “Training” tab. For the epithelium mask, set the Labels as Epithelium and Stroma. Next click the “Add Label” button to add a third class, and name it glass. Now you are ready to draw on top of your stack. You can change the stack and which marker you are viewing on the lower left panel.



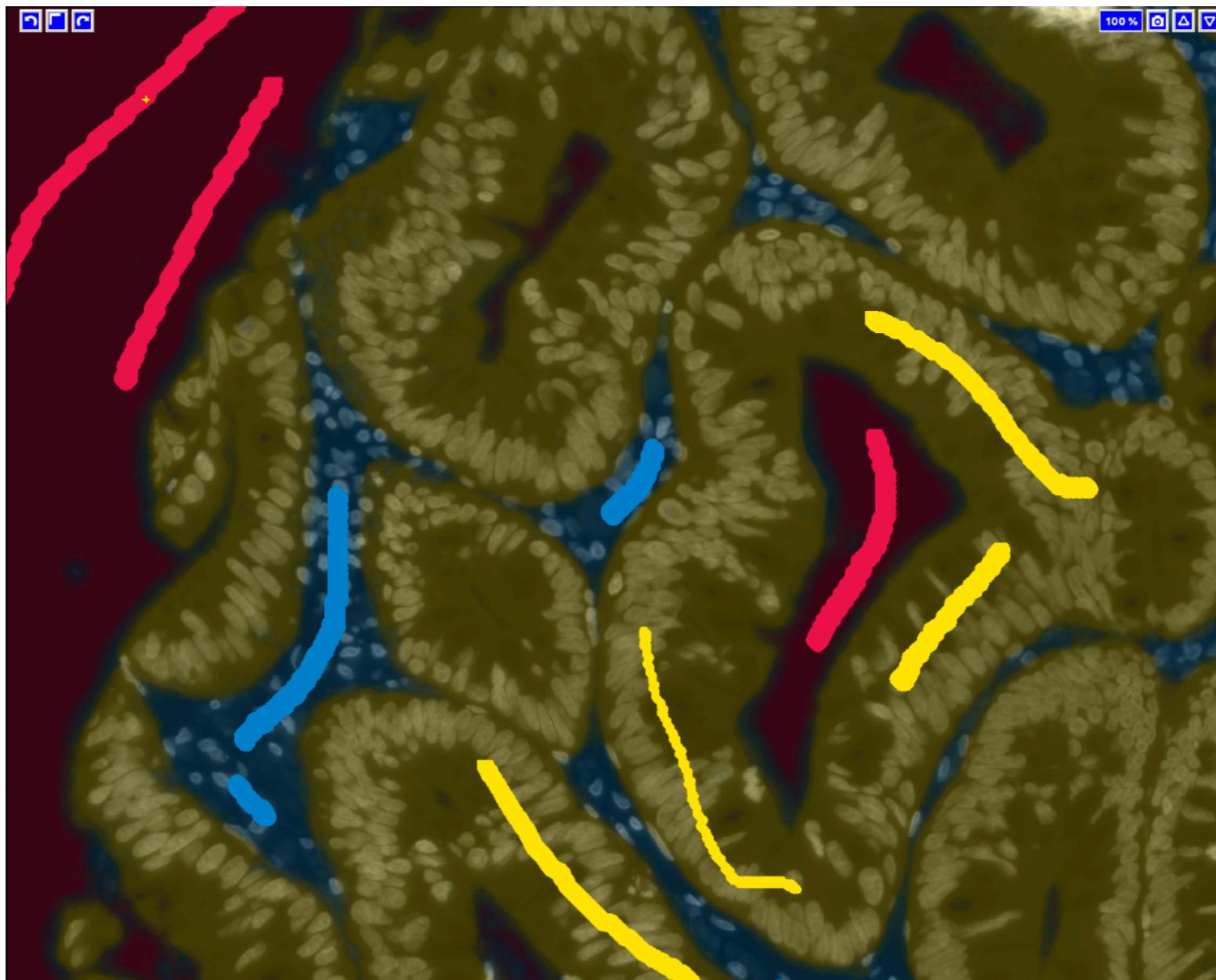
Generating probability masks in Hlastik

Annotate each label for each image stack and when finished select the “Live Update” button. This will run the random forest pixel classification algorithm. N.B.: depending on the data set, this can take a long time.



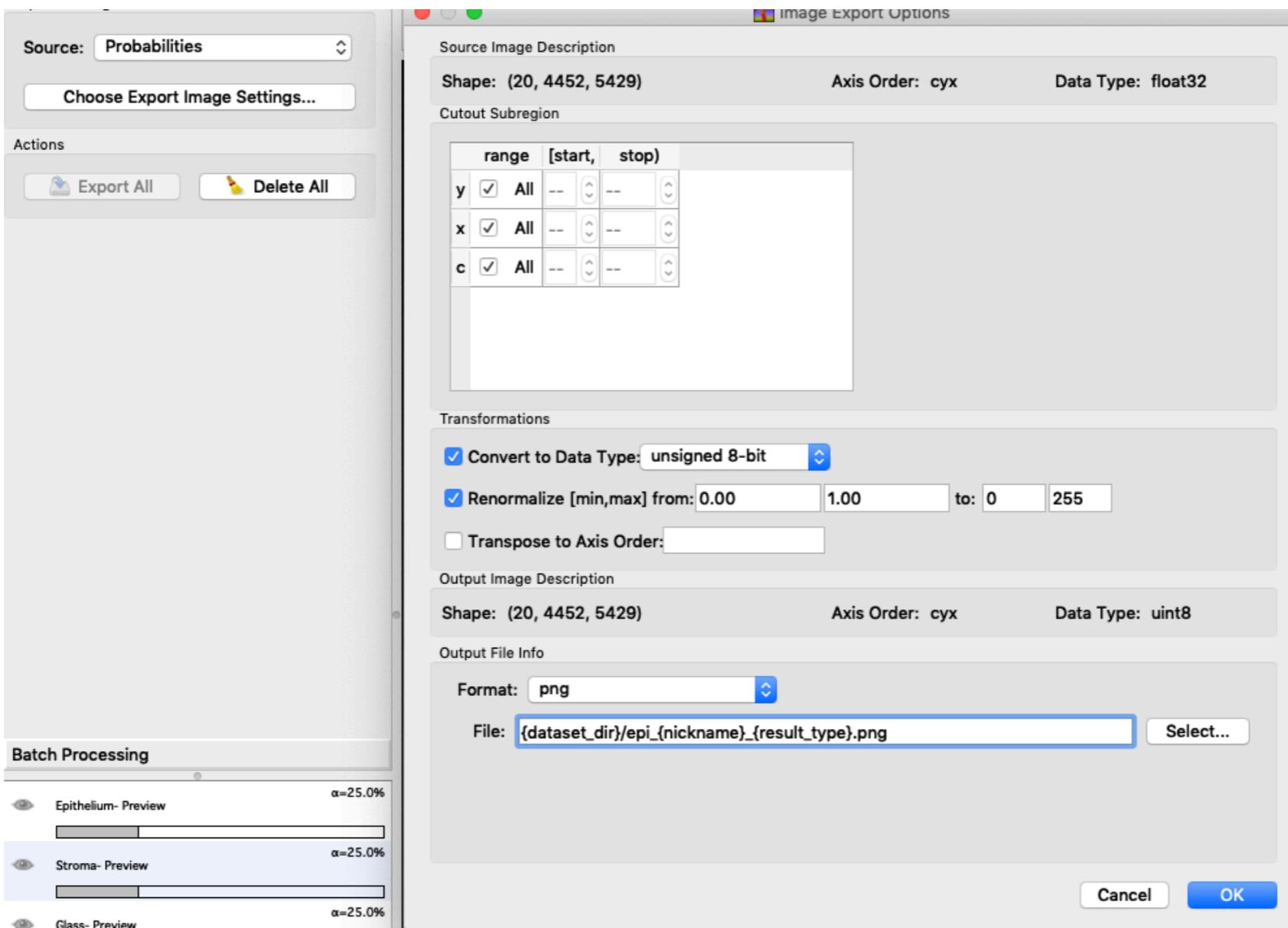
Generating probability masks in Illastik

When training is complete, click the “Live update” button again to turn off additional training and save the project. If you are satisfied with the training move to the next step. Otherwise continue to annotate the image until you are satisfied with the results.



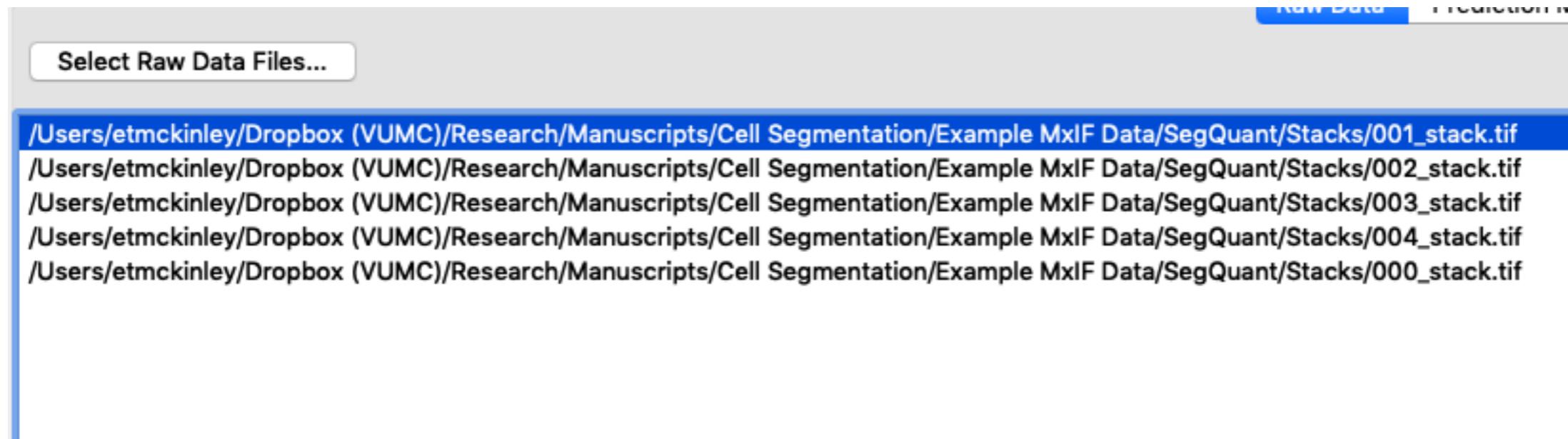
Generating probability masks in Illastik

Next we have to set the export settings. Click the “Prediction Export” tab, and then click the “Choose Export Image Settings” button. Check “Convert to Data Type” and select “unsigned 8-bit” from the dropdown menu and check the “Renormalize [min,max] from” and leave the default parameters. Change the file format to “png” and add an “epi_” prior to {nickname} in the “File” input. You can press the “Export” button on the top of the window to test the export (file will be in the “Stacks” folder in SegQuant).



Generating probability masks in Ilastik

Next, click on the “Batch Processing” tab. Press the “Select Raw Data Files” in the upper portion of the window and select the image stacks that you wish to process. Then press the “Process All Files” button. This can also take some time to complete.

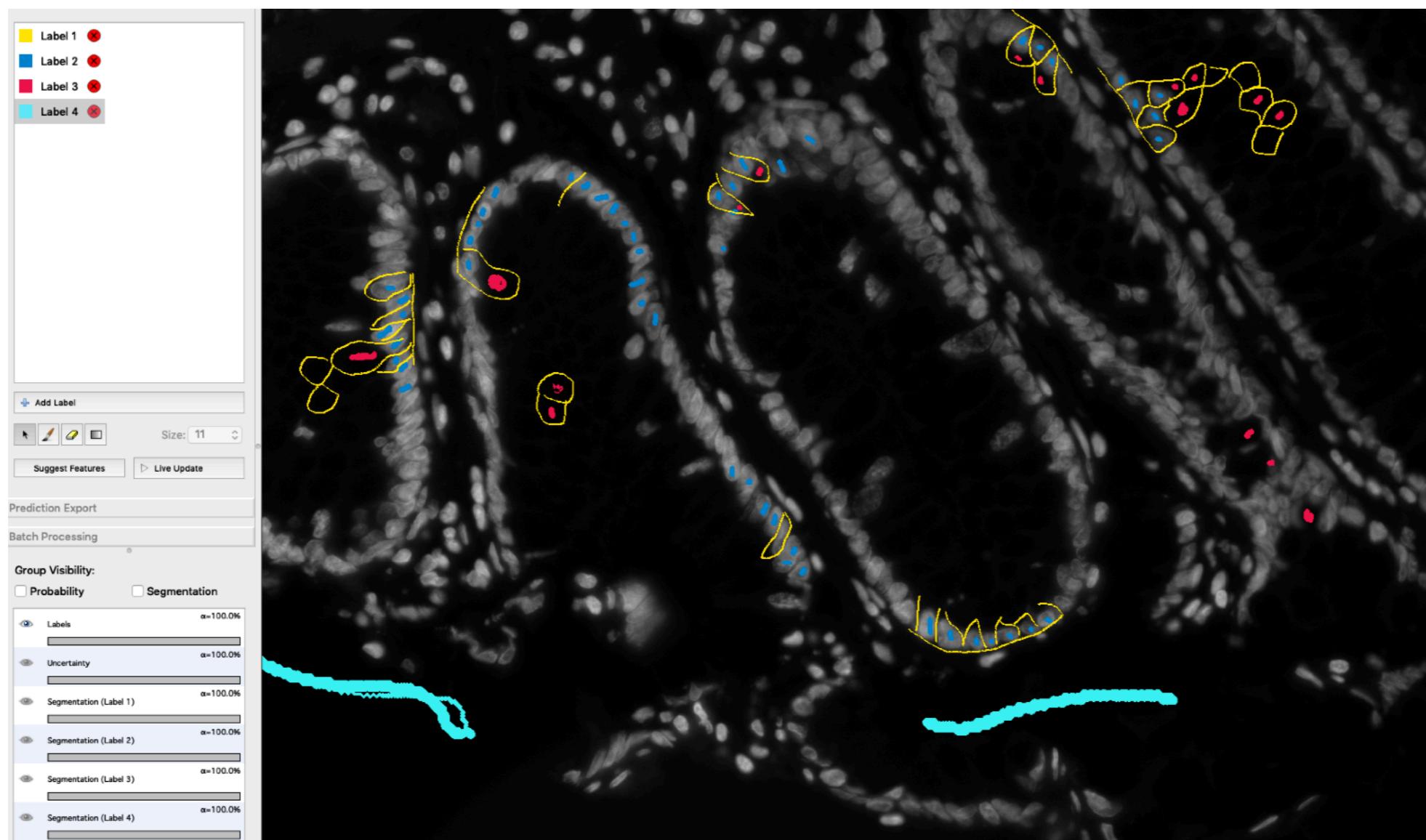


When complete, move the png probability images to the “ML” folder in SegQuant

Generating probability masks in Illastik

Repeat the training process for membrane/nucleus/cytoplasm/glass. There are small details that need to be changed from the epithelium/stroma/glass training.

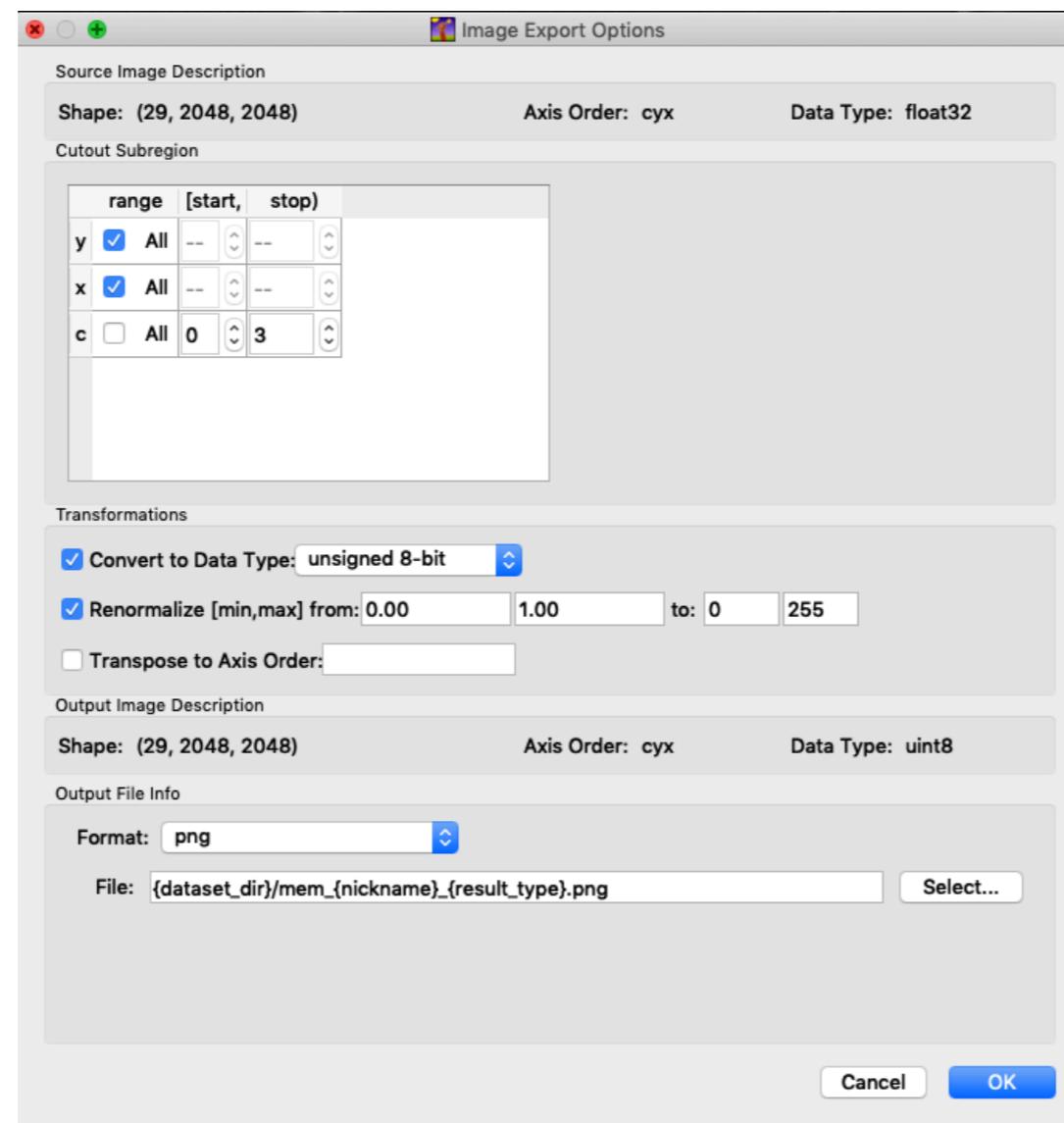
Use 4 label classes for the training. Label 1: membrane; Label 2: nucleus; Label 3: cytoplasm; Label 4: glass



Generating probability masks in Ilastik

Repeat the training process for membrane/nucleus/cytoplasm/glass. There are small details that need to be changed from the epithelium/stroma/glass training.

In export options, uncheck range “c” and set start to 0 and stop to 3. Additionally, add “mem_” in front of {nickname}. Move png files to SegQuant/ML when finished.

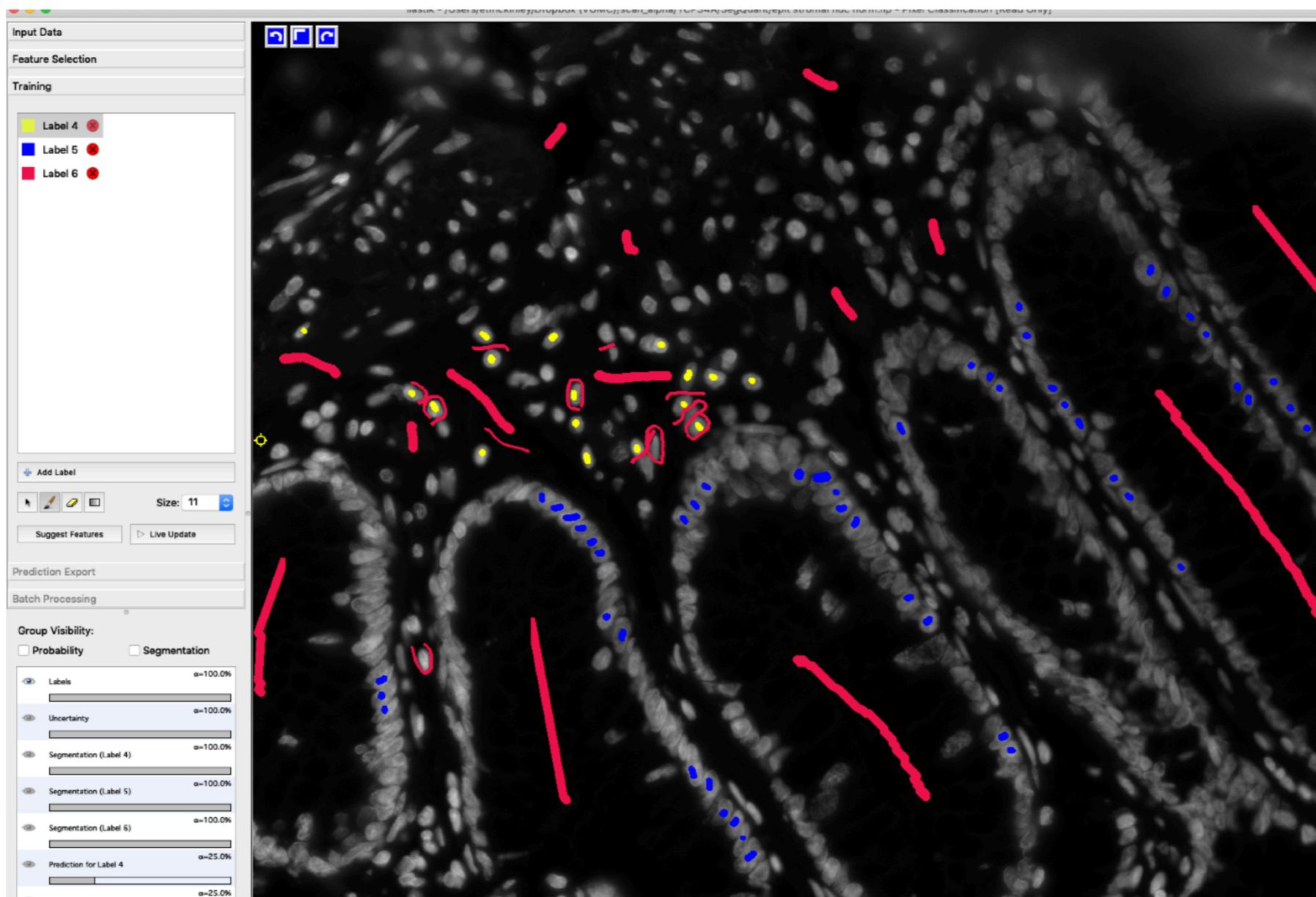


Generating probability masks in Ilastik

Optional if stromal quantification is desired:

Repeat the training process for epithelial nuclei/stromal nuclei/other. There are small details that need to be changed from the epithelium/stroma/glass training.

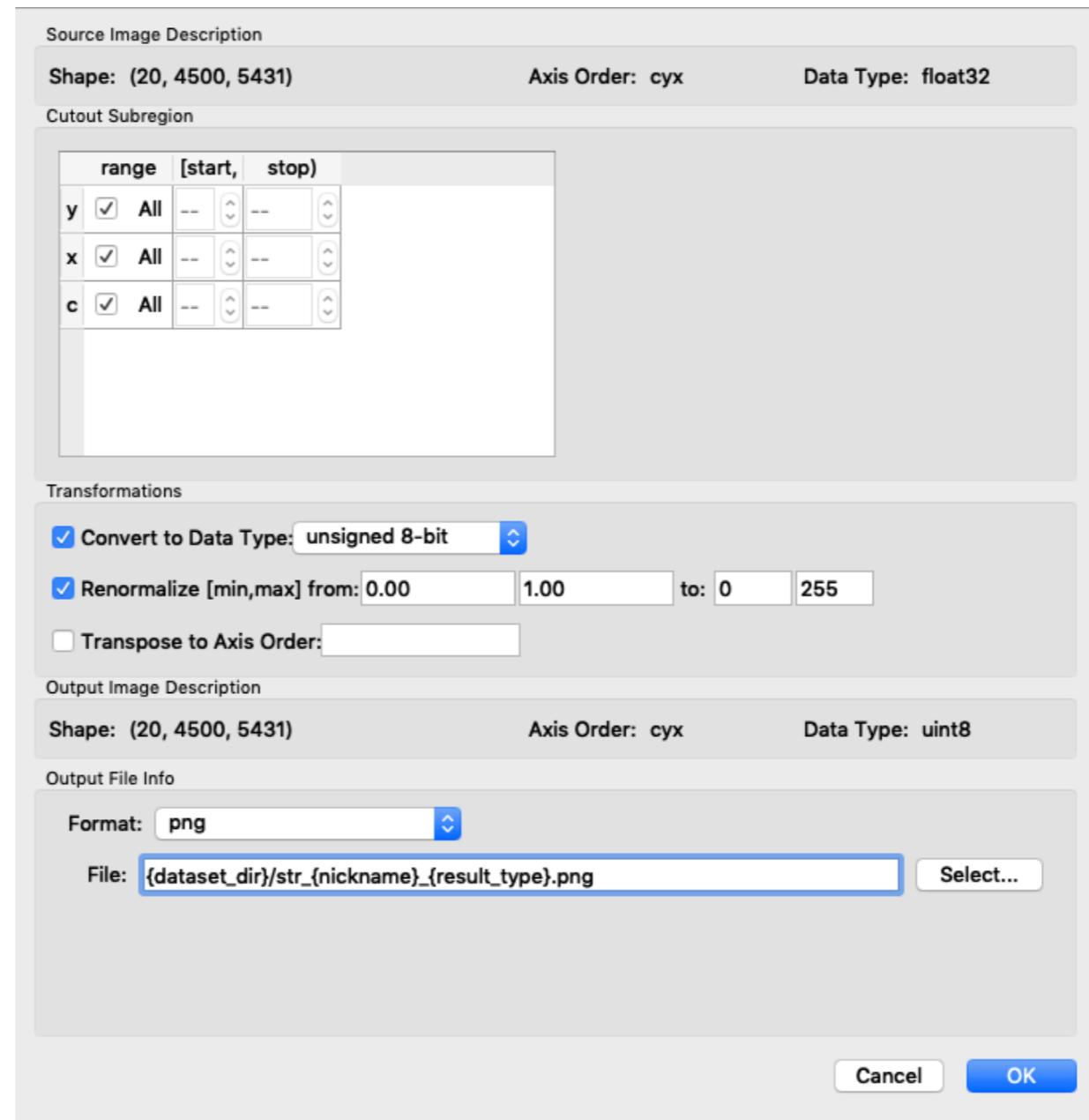
Use 3 label classes for the training. Label 1: stromal nuclei; Label 2: epithelial nuclei; Label 3: other (e.g. glass, non-nuclear cellular components)



Generating probability masks in Illastik

Repeat the training process for membrane/nucleus/cytoplasm/glass. There are a few small details that need to be changed from the epithelium/stroma/glass training.

Use the same export options as epithelium/stroma/glass, but place “str_” in front of {nickname}. Move png files to SegQuant/ML when finished.



Re-initialized pipeline in Matlab

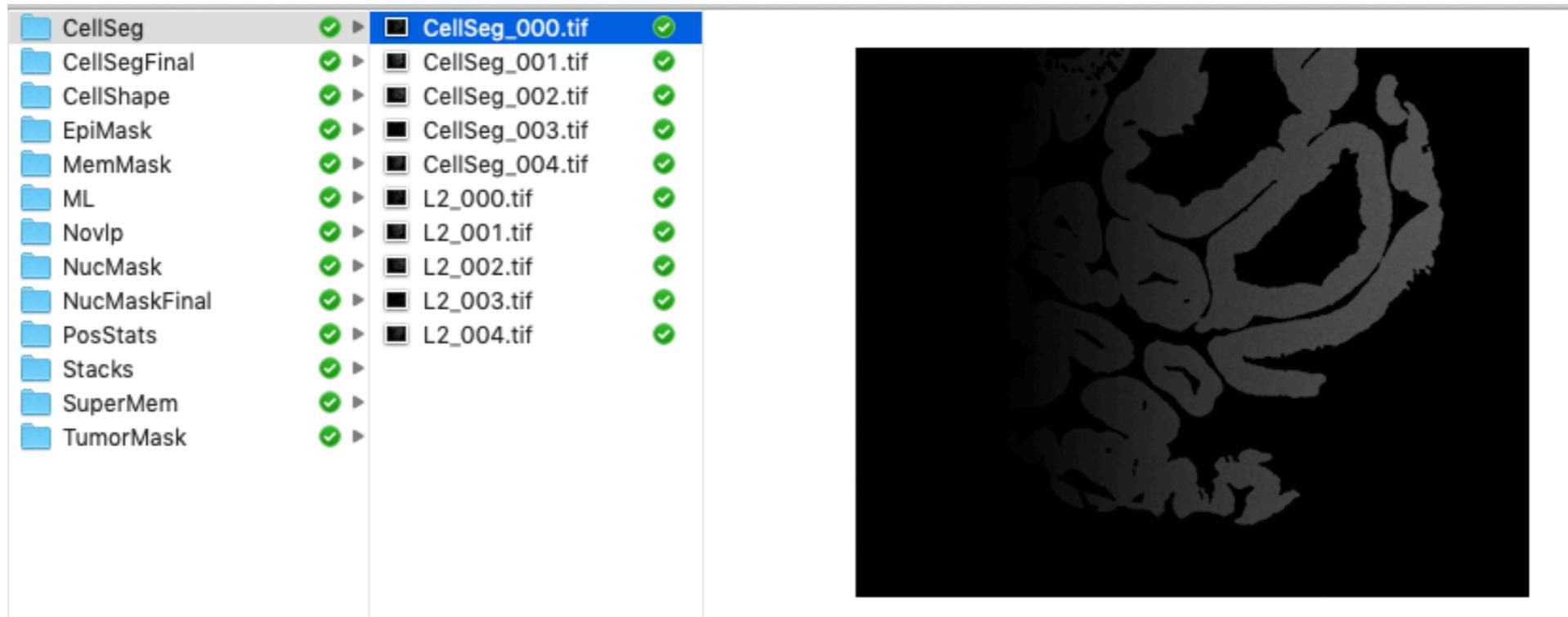
Run the script to call the segmentation pipeline again. No further interaction is needed. The script will segment and quantify each marker and measure cell shape automatically.

```
warning off
clear all
close all

dir='/Users/etmckinley/Dropbox (VUMC)/Research/Manuscripts/Cell Segmentation/Example MxIF Data';
CellSegQuant(dir ,1,1,1,0, 1)
```

Pipeline Outputs

In the end, multiple folders with results will be generated. A general overview of each follows.



CellSeg

This folder contains intermediate segmentation steps (e.g. prior to cell re-segmentation for internal membranes) and generally does not need to be looked at.

Pipeline Outputs

In the end, multiple folders with results will be generated. A general overview of each follows.

CellSeg	✓	CellSegFinal_000.tif	✓
CellSegFinal	✓	CellSegFinal_001.tif	✓
CellShape	✓	CellSegFinal_002.tif	✓
EpiMask	✓	CellSegFinal_003.tif	✓
MemMask	✓	CellSegFinal_004.tif	✓
ML	✓	StrCellSegFinal_000.tif	✓
Novlp	✓	StrCellSegFinal_001.tif	✓
NucMask	✓	StrCellSegFinal_002.tif	✓
NucMaskFinal	✓	StrCellSegFinal_003.tif	✓
PosStats	✓	StrCellSegFinal_004.tif	✓
Stacks	✓		
SuperMem	✓		
TumorMask	✓		

CellSegFinal

This folder contains the final cell segmentation for epithelial and, if defined, stromal cells. Each cell has a unique pixel value to denote its ID within each position.

Pipeline Outputs

In the end, multiple folders with results will be generated. A general overview of each follows.

CellSeg	✓	autoencoder.mat	✓
CellSegFinal	✓	CellShape_000.mat	✓
CellShape	✓	CellShape_001.mat	✓
EpiMask	✓	CellShape_002.mat	✓
MemMask	✓	CellShape_003.mat	✓
ML	✓	CellShape_004.mat	✓
Novlp	✓	encoded_cells.csv	✓
NucMask	✓		
NucMaskFinal	✓		
PosStats	✓		
Stacks	✓		
SuperMem	✓		
TumorMask	✓		

CellShape

This folder contains images of individual cell shapes for each position as well as the autoencoder in .mat files. A “encoded_cells.csv” are the latent vectors for each cell that can be used for further processing. The cell IDs match the marker quantification files.

Pipeline Outputs

In the end, multiple folders with results will be generated. A general overview of each follows.

CellSeg	✓	EpiMask_000.png	✓
CellSegFinal	✓	EpiMask_001.png	✓
CellShape	✓	EpiMask_002.png	✓
EpiMask	✓	EpiMask_003.png	✓
MemMask	✓	EpiMask_004.png	✓
ML	✓		
Novlp	✓		
NucMask	✓		
NucMaskFinal	✓		
PostStats	✓		
Stacks	✓		
SuperMem	✓		
TumorMask	✓		



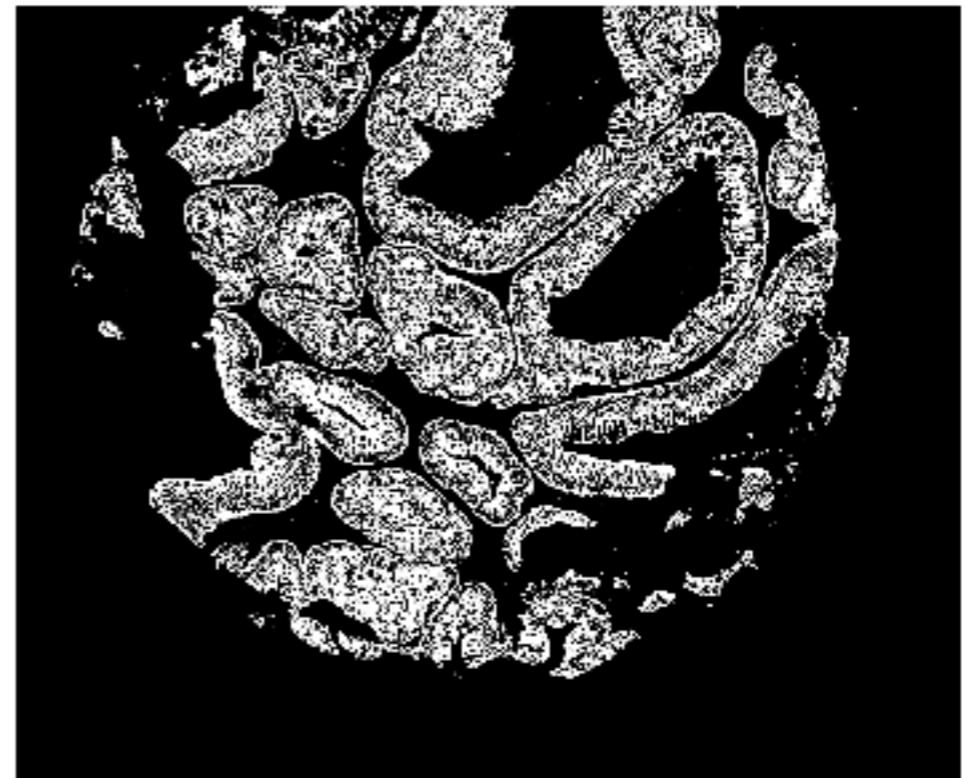
EpiMask

This folder contains the final epithelial masks that were derived from the machine learning pixel classification with some morphological filtering to remove small holes and smooth edges.

Pipeline Outputs

In the end, multiple folders with results will be generated. A general overview of each follows.

CellSeg	✓	MemMask_000.png	✓
CellSegFinal	✓	MemMask_001.png	✓
CellShape	✓	MemMask_002.png	✓
EpiMask	✓	MemMask_003.png	✓
MemMask	✓	MemMask_004.png	✓
ML	✓		
Novlp	✓		
NucMask	✓		
NucMaskFinal	✓		
PosStats	✓		
Stacks	✓		
SuperMem	✓		
TumorMask	✓		



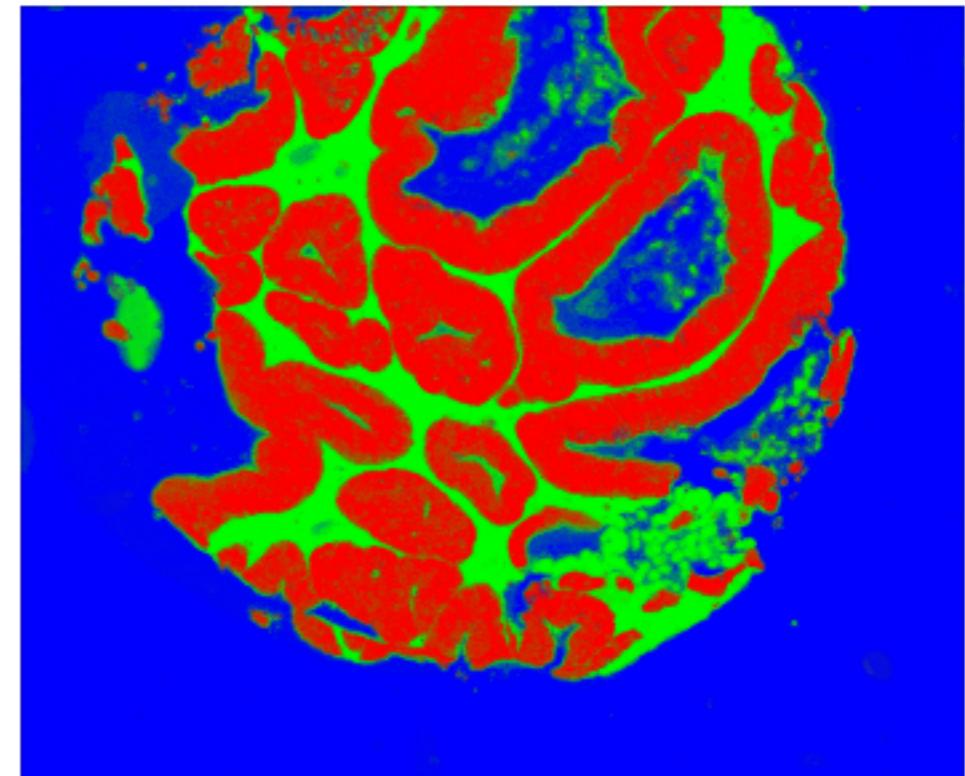
MemMask

This folder contains the final membrane masks that were derived from the machine learning pixel classification.

Pipeline Outputs

In the end, multiple folders with results will be generated. A general overview of each follows.

CellSeg	✓	► epi_000_s...ilities.png	✓
CellSegFinal	✓	► epi_001_s...ilities.png	✓
CellShape	✓	► epi_002_s...ilities.png	✓
EpiMask	✓	► epi_003_s...ilities.png	✓
MemMask	✓	► epi_004_s...ilities.png	✓
ML	✓	► mem_000...ilities.png	✓
Novlp	✓	► mem_001...ilities.png	✓
NucMask	✓	► mem_002...ilities.png	✓
NucMaskFinal	✓	► mem_003...ilities.png	✓
PosStats	✓	► mem_004...ilities.png	✓
Stacks	✓	► str_000_s...ilities.png	✓
SuperMem	✓	► str_001_s...ilities.png	✓
TumorMask	✓	► str_002_s...ilities.png	✓
		► str_003_s...ilities.png	✓
		► str_004_s...ilities.png	✓



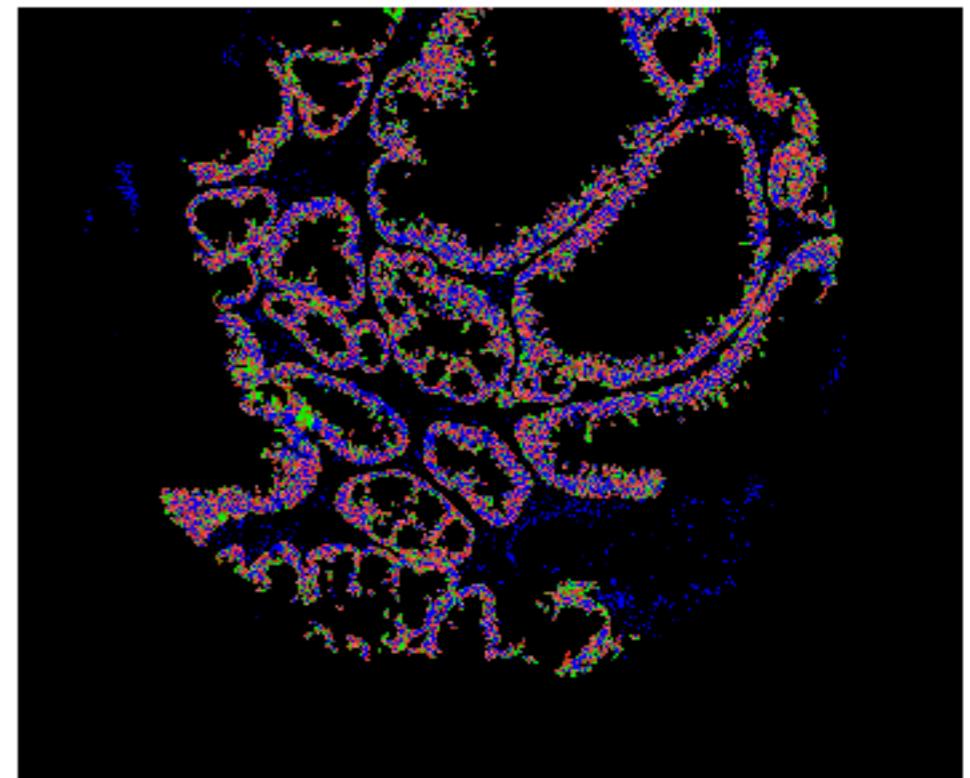
ML

This folder contains the pixel classification images that were created in Ilastik.

Pipeline Outputs

In the end, multiple folders with results will be generated. A general overview of each follows.

CellSeg	Novlp_000.png
CellSegFinal	Novlp_001.png
CellShape	Novlp_002.png
EpiMask	Novlp_003.png
MemMask	Novlp_004.png
ML	StrNovlp_000.png
Novlp	StrNovlp_001.png
NucMask	StrNovlp_002.png
NucMaskFinal	StrNovlp_003.png
PosStats	StrNovlp_004.png
Stacks	
SuperMem	
TumorMask	



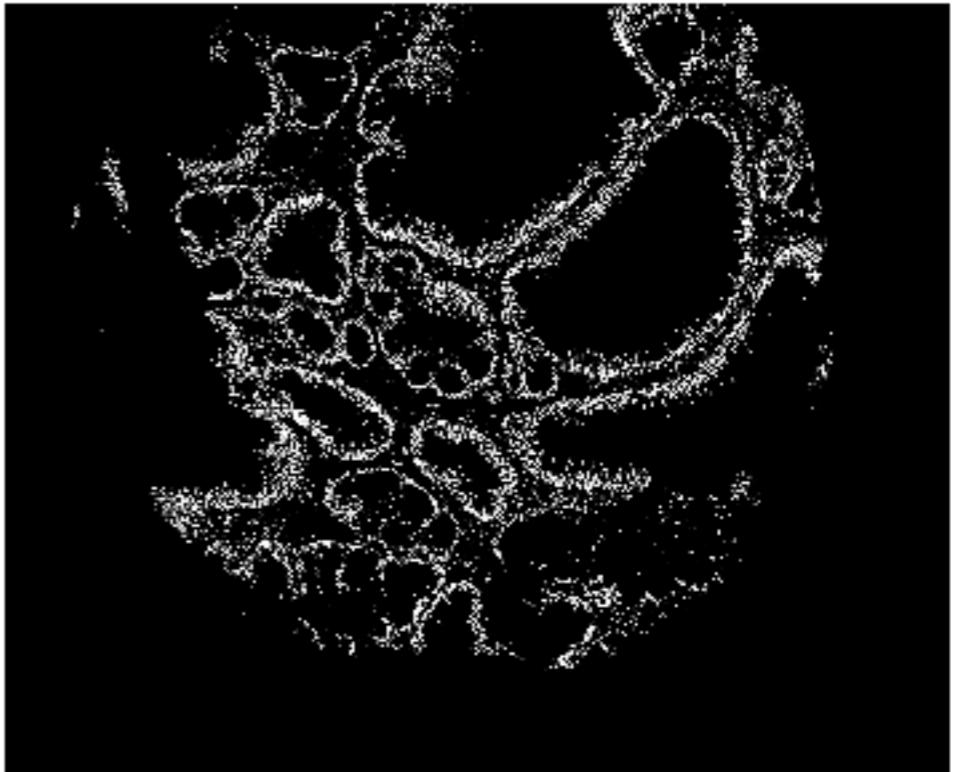
Novlp

This contains the non-overlapping (novlp) sub-cellular segmentation images. Cell borders are white lines, nucleus is blue, cytoplasm is green, and membranes are red. These images can help qualitatively assess the quality of segmentation and are not used for any calculations in the pipeline.

Pipeline Outputs

In the end, multiple folders with results will be generated. A general overview of each follows.

CellSeg	✓	►	■ NucMask_000.png	✓
CellSegFinal	✓	►	■ NucMask_001.png	✓
CellShape	✓	►	■ NucMask_002.png	✓
EpiMask	✓	►	■ NucMask_003.png	✓
MemMask	✓	►	■ NucMask_004.png	✓
ML	✓	►		
Novlp	✓	►		
NucMask	✓	►		
NucMaskFinal	✓	►		
PosStats	✓	►		
Stacks	✓	►		
SuperMem	✓	►		
TumorMask	✓	►		



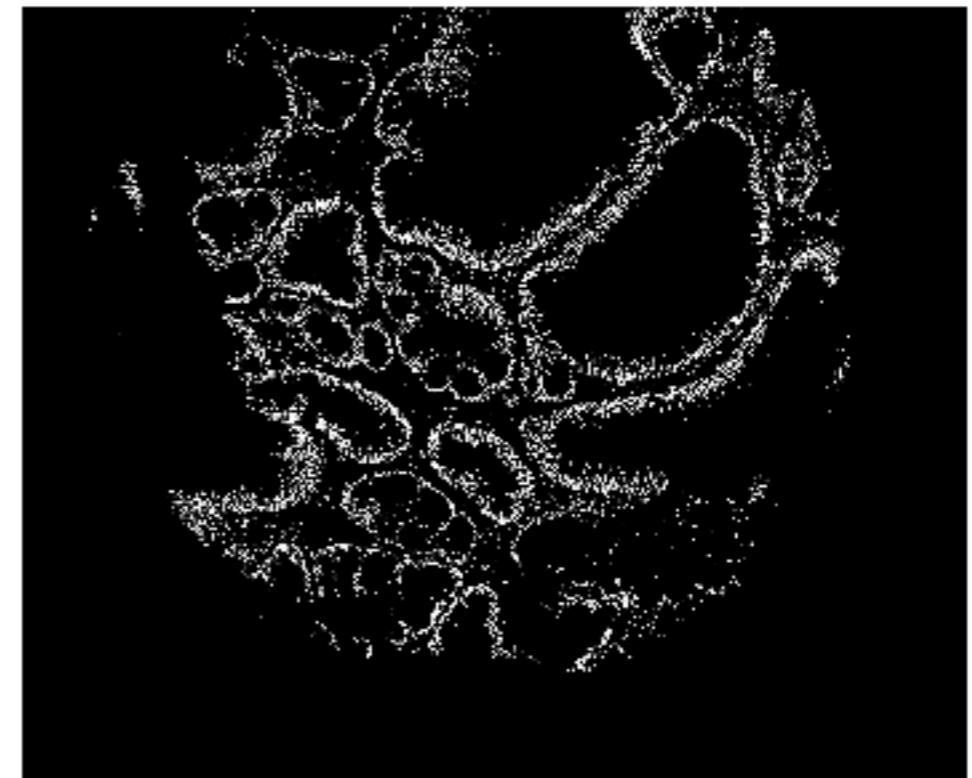
NucMask

Intermediate nuclear mask derived from the machine learning pixel classification.

Pipeline Outputs

In the end, multiple folders with results will be generated. A general overview of each follows.

CellSeg	✓ ►	NucMask...al_000.png	✓
CellSegFinal	✓ ►	NucMask...al_001.png	✓
CellShape	✓ ►	NucMask...al_002.png	✓
EpiMask	✓ ►	NucMask...al_003.png	✓
MemMask	✓ ►	NucMask...al_004.png	✓
ML	✓ ►	StrNucMa...l_000.png	✓
Novlp	✓ ►	StrNucMa...l_001.png	✓
NucMask	✓ ►	StrNucMa...l_002.png	✓
NucMaskFinal	✓ ►	StrNucMa...l_003.png	✓
PosStats	✓ ►	StrNucMa...l_004.png	✓
Stacks	✓ ►		
SuperMem	✓ ►		
TumorMask	✓ ►		

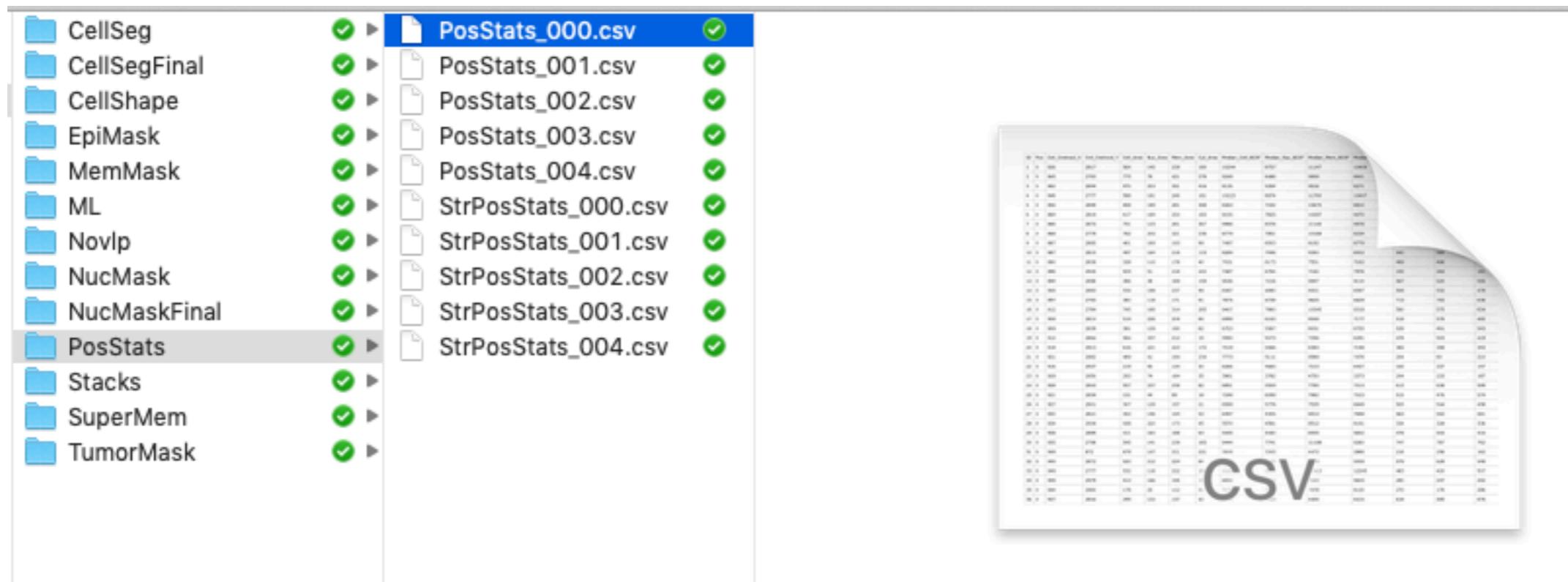


NucMaskFinal

Final nuclear mask for epithelial and stromal cells. Some filtering is done from NucMask for the epithelial cells.

Pipeline Outputs

In the end, multiple folders with results will be generated. A general overview of each follows.



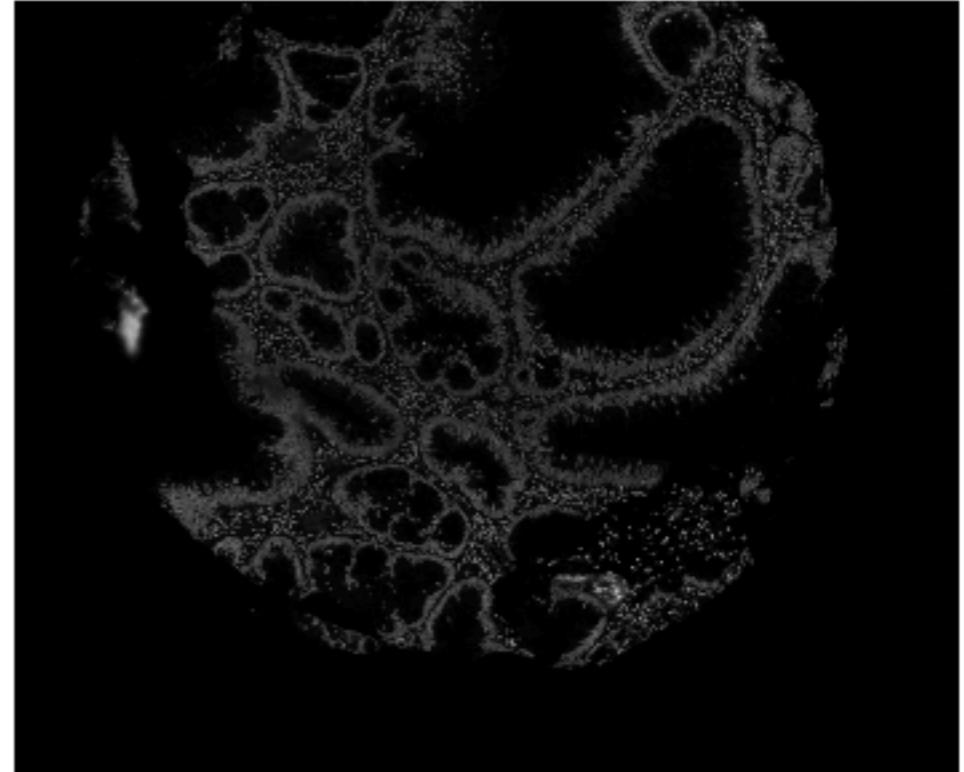
PosStats

This folder contains csv quantification files for epithelial and stromal cells for each image position. Each row is a cell, each column is a measurement. The cell IDs match the cell shape quantification files.

Pipeline Outputs

In the end, multiple folders with results will be generated. A general overview of each follows.

CellSeg	✓	000_stack.tif	✓
CellSegFinal	✓	001_stack.tif	✓
CellShape	✓	002_stack.tif	✓
EpiMask	✓	003_stack.tif	✓
MemMask	✓	004_stack.tif	✓
ML	✓		
Novlp	✓		
NucMask	✓		
NucMaskFinal	✓		
PosStats	✓		
Stacks	✓		
SuperMem	✓		
TumorMask	✓		



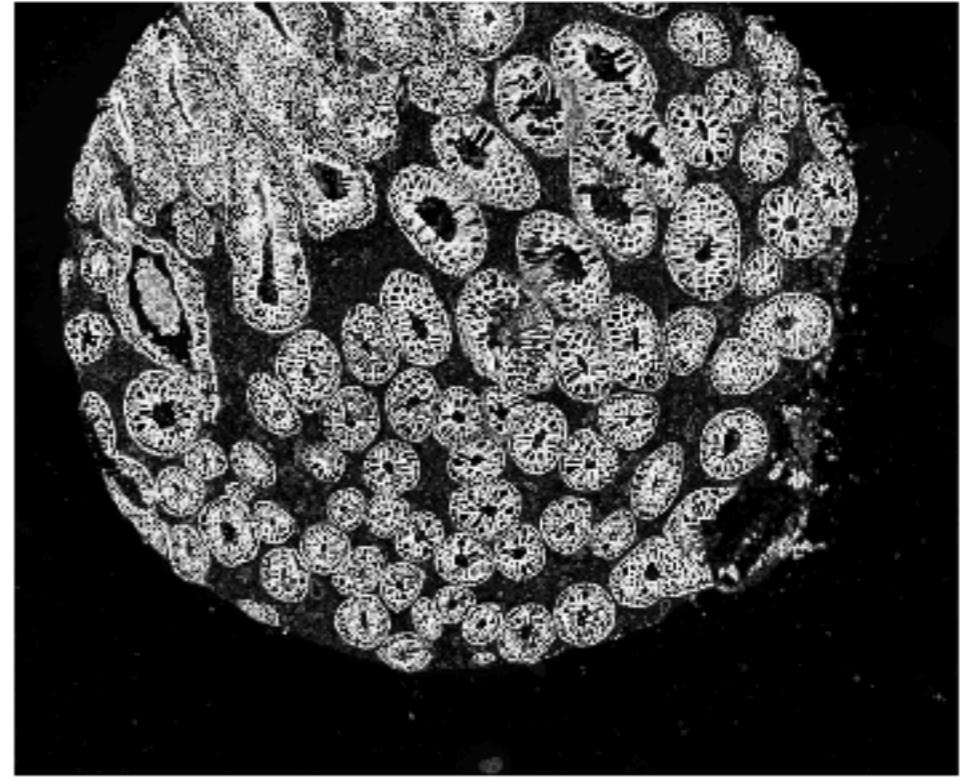
Stacks

This folder contains the tiff stacks generated from the AFRemoved images and DAPI. These are used for the machine learning only.

Pipeline Outputs

In the end, multiple folders with results will be generated. A general overview of each follows.

CellSeg	✓	SuperMem_000.tif	✓
CellSegFinal	✓	SuperMem_001.tif	✓
CellShape	✓	SuperMem_002.tif	✓
EpiMask	✓	SuperMem_003.tif	✓
MemMask	✓	SuperMem_004.tif	✓
ML	✓		
Novlp	✓		
NucMask	✓		
NucMaskFinal	✓		
PosStats	✓		
Stacks	✓		
SuperMem	✓		
TumorMask	✓		



SuperMem

This folder contains the tiff of membrane mask probabilities in gray scale.

Pipeline Outputs

In the end, multiple folders with results will be generated. A general overview of each follows.

CellSeg	✓ ▶
CellSegFinal	✓ ▶
CellShape	✓ ▶
EpiMask	✓ ▶
MemMask	✓ ▶
ML	✓ ▶
Novlp	✓ ▶
NucMask	✓ ▶
NucMaskFinal	✓ ▶
PosStats	✓ ▶
Stacks	✓ ▶
SuperMem	✓ ▶
TumorMask	✓ ▶

TumorMask

This folder contains binary images of the tumor mask if applicable. For this example they were not used, but could be generated in Ilastik if there are regions of normal and tumor tissue that need to be defined.