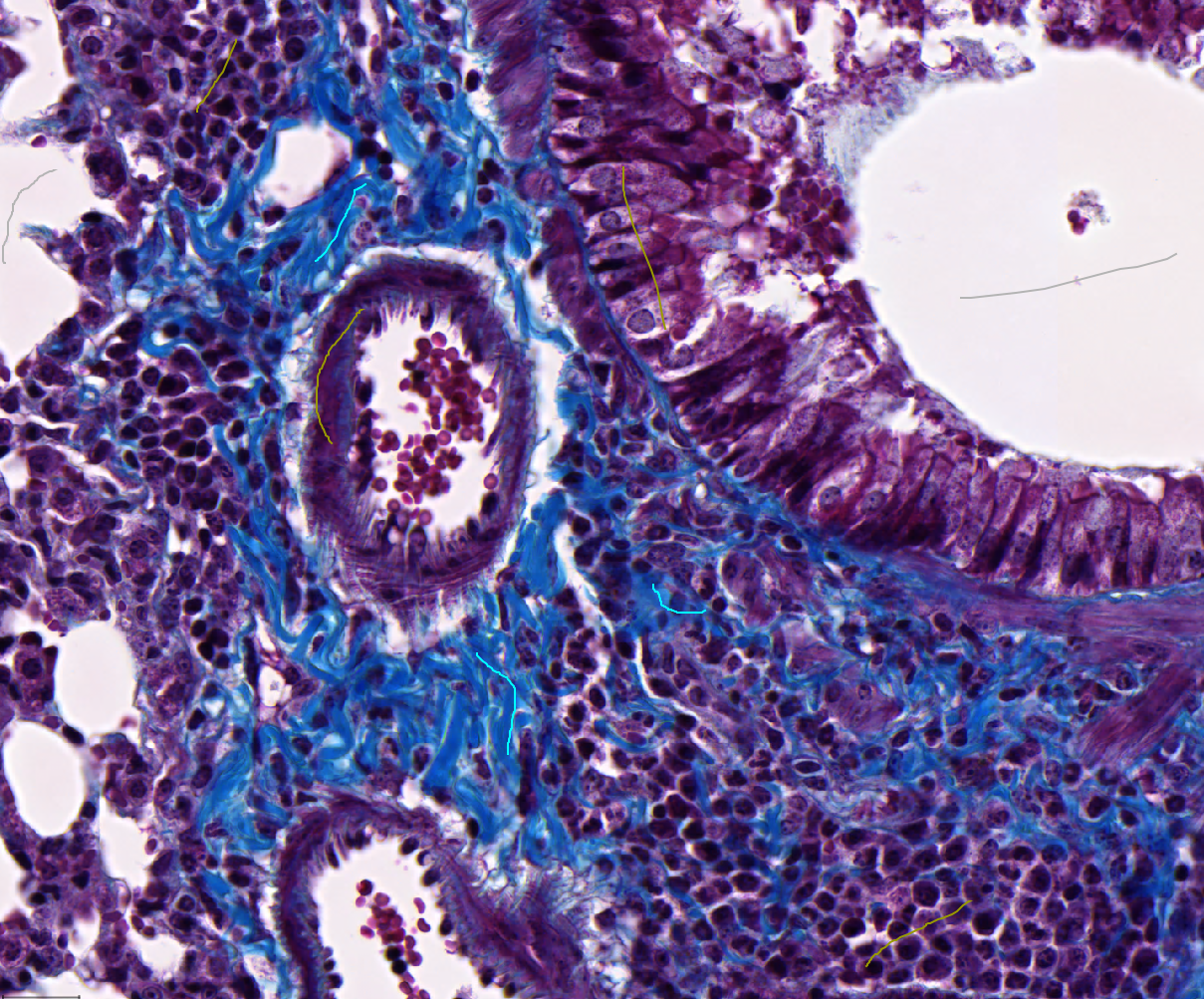
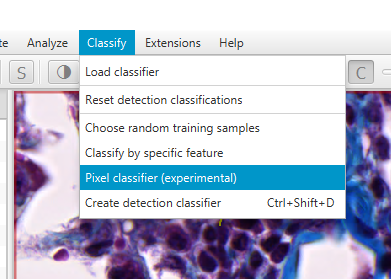
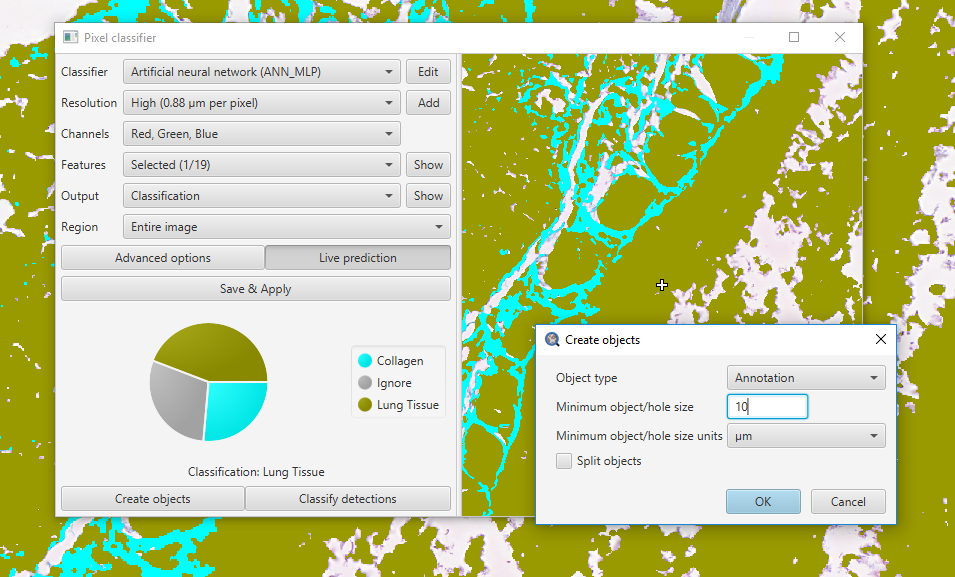
**Measuring Collagen in the Lung**

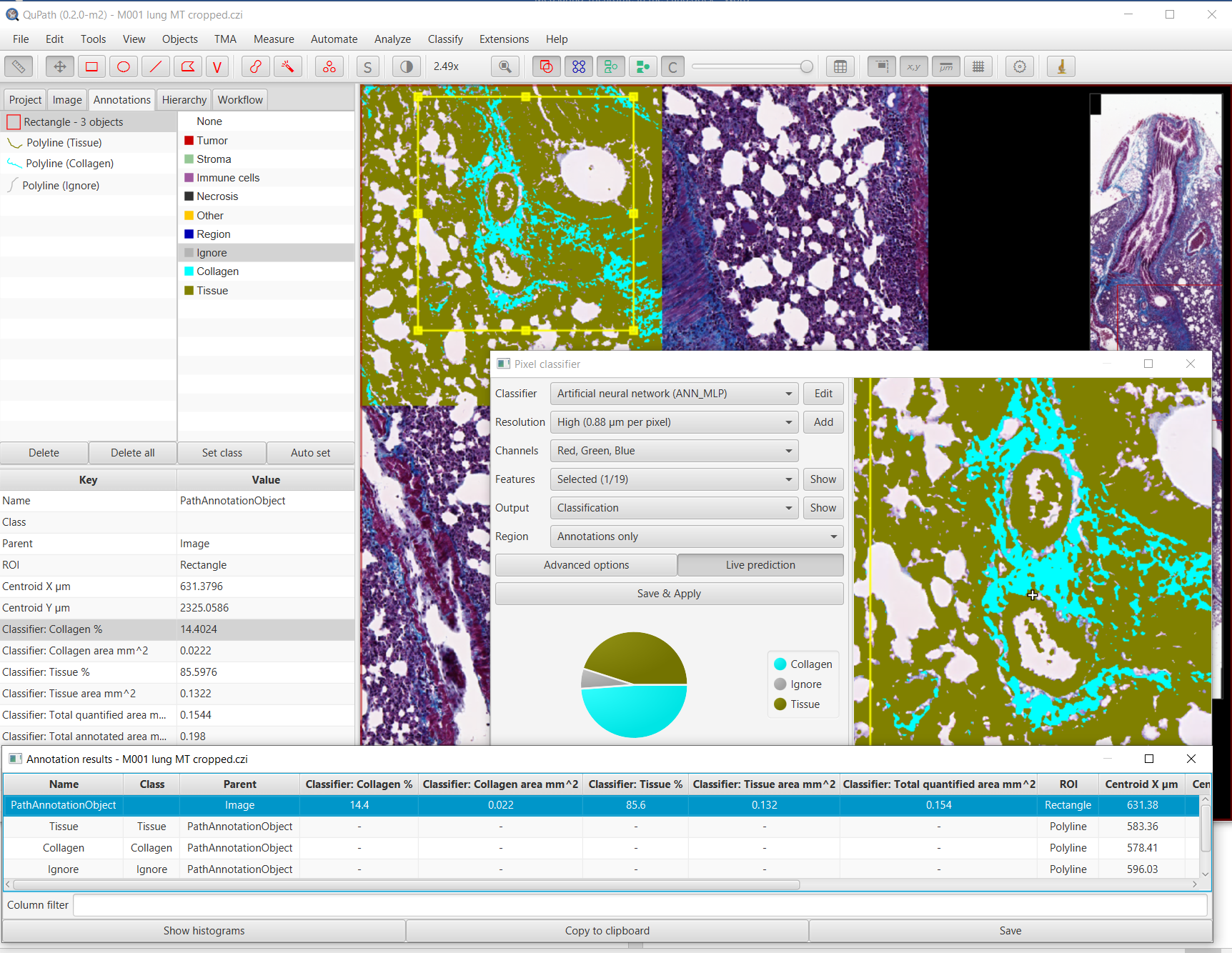
Step-by-step Protocol

**Goals:** Learn how to use the pixel classifier

Assumes you are already familiar with projects, annotations, etc

1. Use QuPath 0.2.0 (any milestone)
2. Download “M001 Lung MT cropped.czi” and “M007 Lung MT cropped.czi”.  
   These are images of inflamed mouse lungs stained with Masson’s Trichrome, which highlights collagen in blue, nuclei in brown/black and other tissue components in dark red/purple. In healthy lungs, large airways and blood vessels will have a collagen ring surrounding them. During inflammation, the amount of collagen increases and it spreads further from the large airways.
3. Create a new project
   1. Use “Add Images” to add the files to the new project
4. Create 2 new classes: “Collagen” and “Lung Tissue”. The specific names are unimportant, just choose ones that are meaningful to you. Don’t use “Region” or “Ignore”- those are special. You can change the colors for the classes.
5. Use the polyline tool to begin to annotate regions of collagen, other tissue, or background. Use the “Ignore” class for background. Don’t forget to set each annotation to a particular class. See example below, where bright blue = Collagen; green = Lung Tissue; and grey = Ignore. Start with ~5 lines of each class in different regions of the slide.   
   
6. Create a pixel classifier. Go to Classify > Pixel classifier (experimental)  
     
   A window will pop up asking you to set some of the parameters of the classifier. Here’s a quick (and incomplete) overview of what the different options mean:
   1. *Classifier*: Random Trees, K-nearest neighbor, or Artificial Neural Network. These 3 different algorithms each have their pros and cons, which can be explored at length elsewhere on the internet (for instance, with [Weka](https://www.cs.waikato.ac.nz/ml/weka/)). For this example, we found ANN to work best, but please play with this to get a feel for their different outputs.
   2. *Resolution*: The pixel size of the eventual output of the classification. Lower resolution (larger pixels) calculates faster. The image on the right side of the window demonstrates the chosen pixel size. \*Using full resolution on large images may cause the program to become unstable on weaker computers!\* Draw a small annotation (see point *e* below) and zoom in for a more smooth experience.
   3. *Channels:* Which of the input data to use. For brightfield, you typically want all 3 channels. For fluorescence images, you may consider removing channels that are irrelevant to your classes.
   4. *Features:* Which operations to calculate on the raw data to use as additional information in the classification. These “features” are different filters that are applied to the image (gaussian smoothing, median, laplacian of the gaussian, etc) that can be thought of as creating additional channels, so that there is more information for each pixel. All features will be calculated on all RGB channels. For this image, we found the best results using the original pixels and median 5x5.
   5. *Region:* Choose whether the live updates of the classifier should be applied to the entire image or just the annotated regions (only applies to annotations with an area, like polygon or rectangle). Use the second option for computers with less resources.
7. Press “live prediction” to see the current classification results
8. Inevitably, there will be some errors. Continue adding annotations over the incorrect classifications to fix them. The results will update continuously as long as “live prediction” is pressed. Also, try changing any of the parameters described above to see if those changes will improve the results. Make sure you look at different regions of the image to be sure there are no local issues.
   1. The pie chart in the window shows how many pixels were annotated as each class. Try to keep these approximately balanced (at least, for the Collagen and Lung Tissue classes. The Ignore class can be much smaller or larger).
9. If you want the entire image classified, unselect all objects by selecting the movement tool  and double clicking anywhere outside of an annotation. If you instead only want to classify a small portion of the image (i.e, an annotation drawn around an interesting piece of the tissue), select that annotation. Multiple annotations can be classified one-at-a-time. If you leave a polyline annotation selected, you will get an error in the next step. For this example, we want to classify the entire image, so deselect all annotations.
10. Once you are satisfied with the results, press “Create Objects” (bottom left corner).  
      
      
      
    A window will pop up asking a few questions:
    1. Will the final objects be annotations or detections? For this example, choose annotations.
    2. Very small regions can be ignored to make processing significantly faster. You can tell the program what is the minimum size to care about, and whether that is in pixels or microns. In this example, choose to suppress anything smaller than 10 um.
    3. With “Split objects” unchecked, you will get 1 object per class. If it is checked, each region of connected pixels will become an individual object. Leave it unchecked for this application.
    4. When you hit OK, it will confirm that it is OK to delete all annotations inside the region that it is classifying. If you are processing the entire image, that will mean all annotations. If you are processing only a smaller region, it will delete the annotations drawn inside that region, but leave the region outline. It is a good idea to save your hand-drawn annotations to a text file before this, using the scripts written by Svidro [here](https://gist.github.com/Svidro/5829ba53f927e79bb6e370a6a6747cfd#file-copying-annotations-between-images-groovy).
11. This will create 2 annotations- one for Lung Tissue and one for Collagen. The “Ignore” class was ignored. Click on each to see and record the total area classified as that type.
12. Type your results in this spreadsheet to see how your results compare with others.
13. Repeat for the next image. Or, try your own!

Notes:

You will often want to know what is the percentage of one of your classes in relation to the total area. To see that, create or select an area annotation after you clicked on *Live prediction* button. Classifier: Collagen % will be shown in Annotations tab and displayed in Annotation results table.

Use “create whole image annotation” if you want to score the whole image. Then draw some annotations, and start a pixel classifier. It might take a while until the whole area is calculated, but the results will be eventually visible as show above. You might need to ask QuPath to ignore the areas at the borders of the image – sometimes black pixels will have weird classification. You need to keep Pixel classifier window open, the results will disappear when you close it!

***Random trees*** classifier uses a random number (seed) to grow the trees. QuPath creates a new seed every time you open an image. This will often lead to a different classification result, even if you use the same annotations to train. ***Artificial Neural Networks*** will always give you consistent results.