SCOPE

The code comprised herein allows for: (1) the development of an itemized histopathology label set to select pathological features of interest for annotation, (2) a tool to import and label 'ground truth' reference images with histopathology, and (3) a tool to visualize annotated images. The purpose of this code is to create 'ground truth' reference images to create data masks which allow for the extraction of optical data associated to histopathology status.

In the referenced work, a 720x1280 pixel video is obtained of the data collection process for a point scanning Fluorescence Lifetime Imaging (FLIm) system. Out of all video frames obtained during an acquisition, 1 frame is selected as the reference image, which shows a clear view of the surgical region of interest; this region typically comprises the main tumor for surgical excision as well as neighboring peripheral healthy tissue. This reference image is then overlaid with 'ground truth' histopathology in PowerPoint. An identical 720x1280 pixel image is then exported from PowerPoint and loaded into the *Data Annotation Tool*. This tool allows for histopathology labels, such as benign tissue, dysplasia, and cancer to be annotated using the labels created in the *Histopathology Labels* toolkit. After creating labels in the *Data Annotation Tool*, a 720x1280 reference mask is created, where all 720x1280 pixels are assigned to a value of zero, except for areas annotated with labels. For example, label 1 might be benign tissue, label 2 might be cancer, etc. With this system, all areas annotated with benign tissue will have a pixel value of 1 in the reference mask. This allows for matching of optical data at known coordinates obtained within the 720x1280 pixel space to pathology labels annotated in the *Data Annotation Tool*.

A representative label set, surgical image, ground truth image, and output (data mask) from the data annotation and visualization tool is provided as an example.

PREREQUISITES

Code was designed and developed with MATLAB Version: 9.5.0.944444 (R2018b) and Operating System: Microsoft Windows 10 Enterprise Version 10.0 (Build 18363). This code should be compatible with future MATLAB releases and operating systems compatible with MATLAB.

The MATLAB Computer Vision System Toolbox, Version 8.2 (R2018b), was used to perform data annotations with modifications made in the attached code. This Computer Vision System Toolbox is required to run the code, as the Image Labeler application is used to label images for the computer vision applications presented in this code.

LAUNCHING TOOLS & USE

(1) Run *main.m* to launch the *Histopathology Data Annotation Toolkit V0.1*

(2) Begin by selecting the *Histopathology Labels* button on the Graphical User Interface (GUI), depicted in Figure 1.
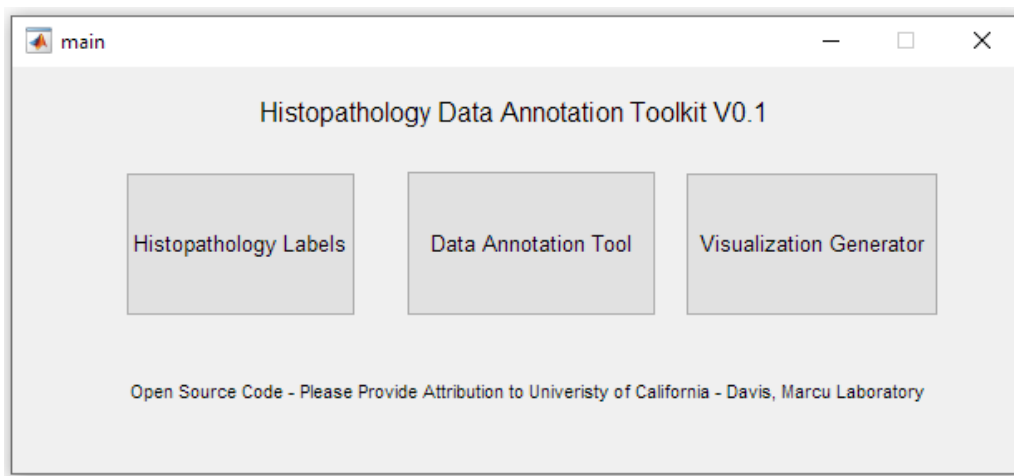

Figure 1: Graphical user interface tool selection window which is generated when running the MATLAB *main.m* code.

(3) After selecting the *Histopathology Labels* button, the following label management interface appears as demonstrated in Figure 2.
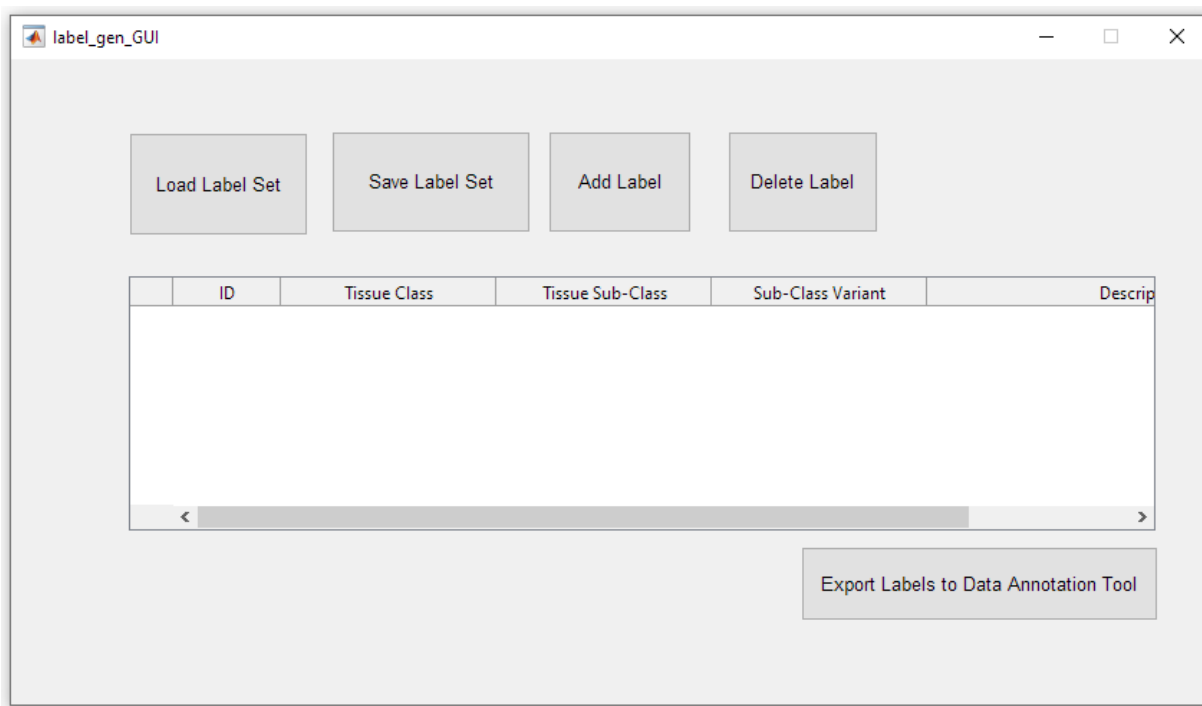

Figure 2: Label management interface to load, modify, and save histopathology label types.

(4) The label management tool allows for the labels used during image annotation to be defined and updated to over time as new tissue types are encountered during a project. Each label type is given a unique numeric ID to ensure machine readability. High level tissue class labels can be defined as well as sub-class labels and sub-class variants (if needed). This hierarchical structure can allow both coarse- and fine-grained comparisons to be made. The generated set of labels can be exported directly into the accompanying *Data Annotation Tool*.

Figure 3 depicts five representative labels in this interface as an example.

To add new entries:

(a) Click *Add Label*

(b) Choose to add a new or existing tissue class label using the generated dialogue

(c) Choose to include a sub-class label (Yes/No); if yes, select a new or existing subclass label

(d) Choose to include a sub-class variant label (Yes/No); if yes, select a new or existing sub class variant label. You cannot have a sub class variant without a sub-class.

(e) Enter a text description for the newly created label

(f) The new label will be added to the table and given a unique ID number (there is a maximum of 255 unique tissue labels in a set)

To delete entries:

(a) Click *Delete Label*

(b) Choose the desired label to delete by its ID number from the list

To save datasets:

1. Click *Save Label Set*

2. Choose a name and location to save the .mat file containing the current label set

To load existing datasets:

1. Click *Load Label Set*

2. Choose a .mat file containing a previously saved label set

3. The table will be populated with the loaded label set (all current labels will be removed)

To export labels to image labeling tool:

1. Click *Export Label Set to Data Annotation Tool*

2. Choose a name and location to save the exported .mat file

3. The saved file stores the current set of labels in a format that can be loaded directly into the image labeler tool
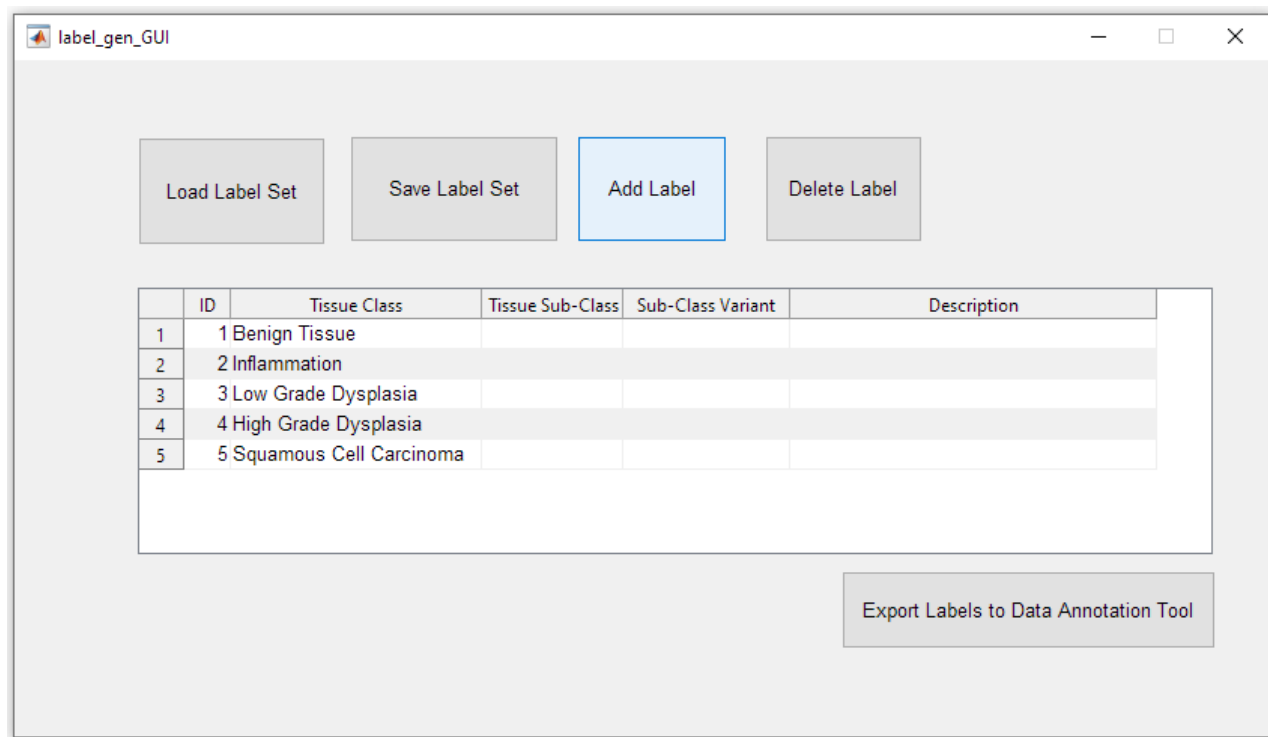

Figure 3: Representative pathology labels created in the label tool for demonstration. Tissue sub-class information, sub-variants, and descriptions left blank. Sub-class designations may be used to differentiate between nuanced tissue conditions, such as (1) Ulcerative Squamous Cell Carcinoma, (2) Invasive Squamous Cell Carcinoma, & (3) *In Situ* Squamous Cell Carcinoma as an example.

(5) After creating pathology labels, on the main *Histopathology Data Annotation Toolkit V0.1* window, select the *Data Annotation Tool* button to launch the image labeling interface demonstrated in *Figure 4*. The Image labeler provided with this code is a modified version of the built in Image Labeler application included within the MATLAB Computer Vision System Toolbox, Version 8.2 (R2018b). Please ensure that you have an updated version of MATLAB installed with this toolbox included. Modifications were made to the importing/exporting of label definitions and annotation maps to allow the tool to be better suited to the optical imaging research that our group is conducting.
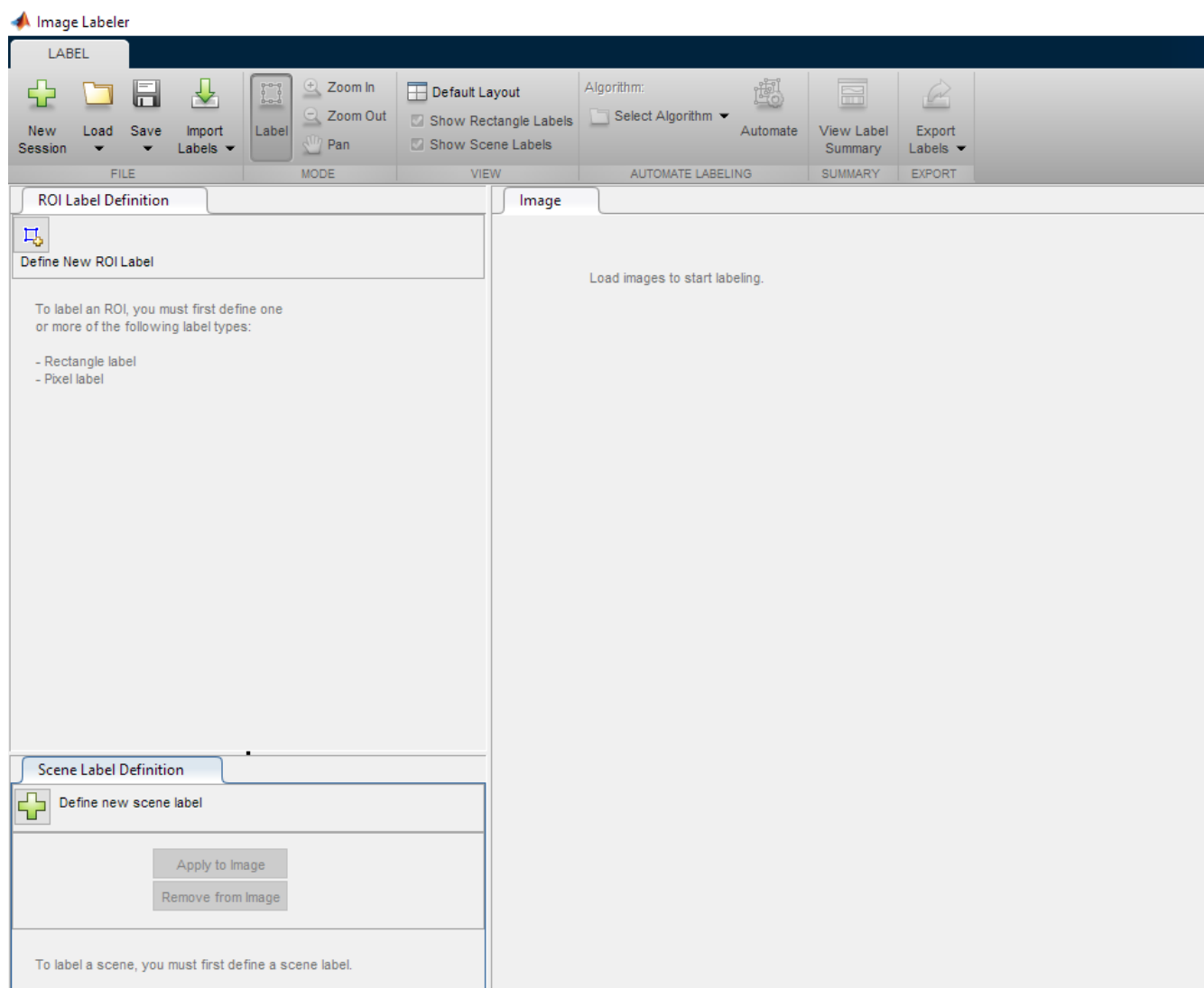
Figure 4: Main data annotation tool image labeler graphical user interface.

(6) After launching the image labeling tool, the next step is to import a label set and load an image to annotate.

To import a label set:

(a) Within the Labels Tab, Click Load >Label Definitions

(b) Select a previously exported .mat file containing a label set

(c) The set of labels will populate the ROI Label Definitions list on the left-hand side (a warning will be written to the command window but can be ignored; this will be removed in a future version)

(d) Clicking into a given class label via the drop-down button presents the class name, sub-class and variant as well as class description within the description box

To load an image to annotate:

(a) Within the Labels Tab, Click Load >Add Images from Folder

(b) Select the desired white light image to annotate

(c) The image will be presented in the main window

With these steps performed, using the test dataset provided, *Figure 5* demonstrates the imported labels and reference image that would be visible in the *Image Labeler* GUI.
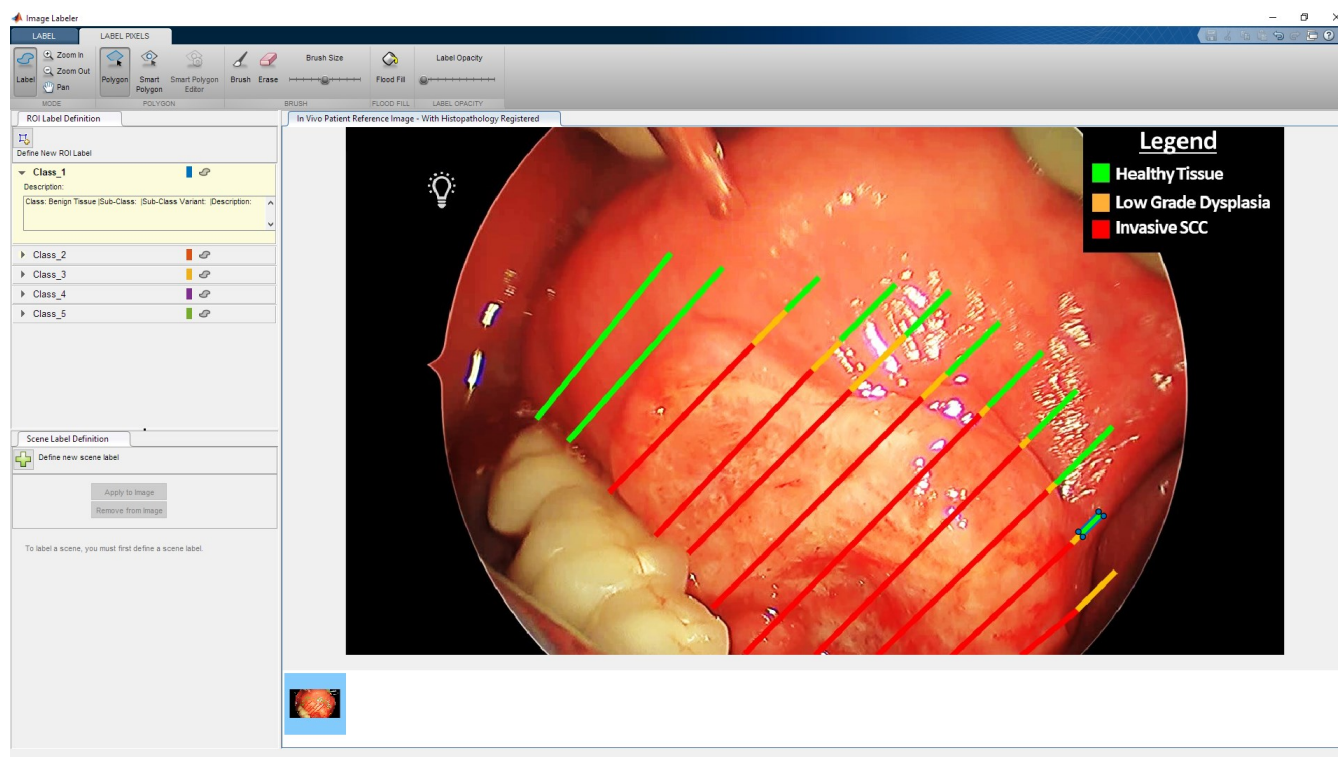

<u>Figure 5</u>*: Data Annotation Tool* with the reference image and label set loaded in, prior to annotations being performed.

(7) Annotations using the reference image and the labels can next be applied.

<u>To apply annotations</u>:

(a) Select the desired class label within the ROI Label Definitions list (it will become highlighted in yellow)

(b) Using either the polygon or brush tool, annotate one or more ROIs for the current class label (the eraser tool can be used to remove mistakes)

(c) The applied annotations will be visualized as a transparent overlay in the main window

Figure 6 illustrates a representative image of this process with the annotations performed on the *Test Dataset* reference image.
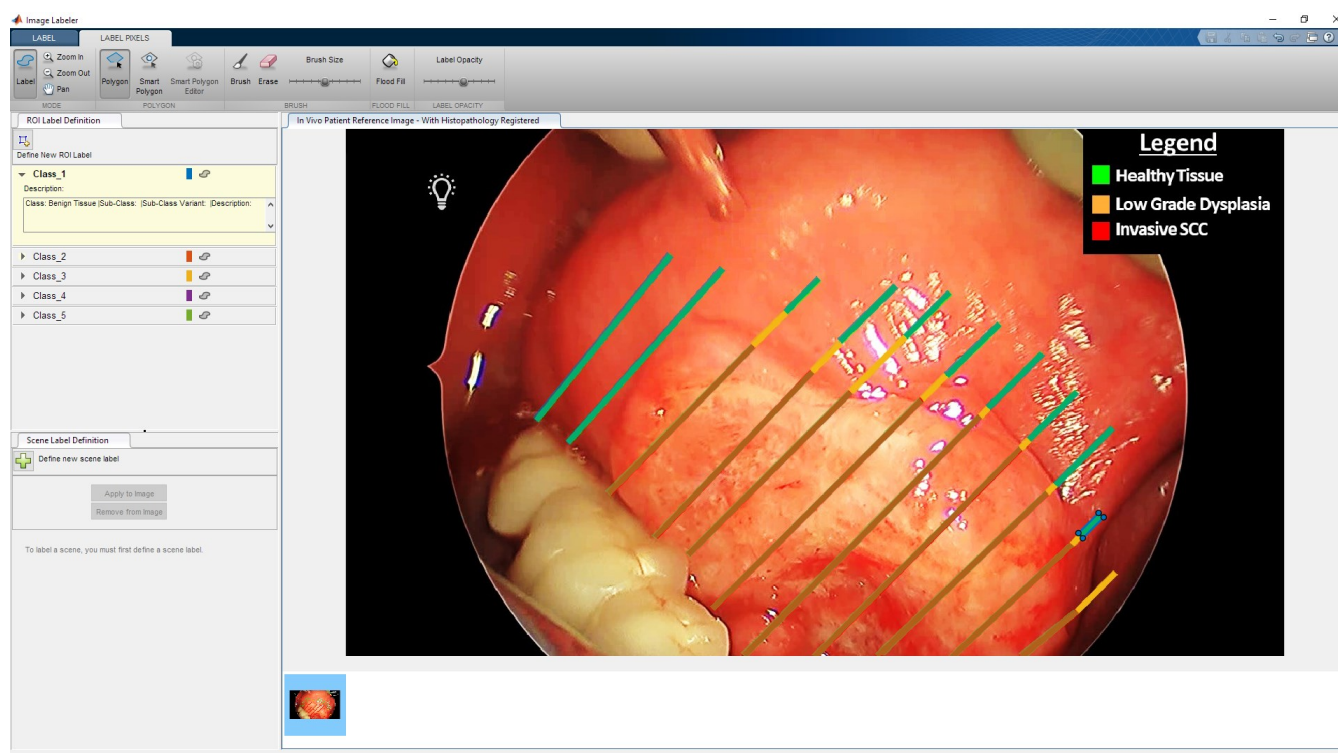
Figure 6: Representative image of the data annotation process, where annotations are performed on the *Test Dataset* reference image.

(8) After all annotations have been performed, the annotation data mask is ready to be created as illustrated in Figure 7.

To export a set of annotations:

(a) Within the Labels Tab, Click Export Labels >To Workspace

(b) Select a name and location for the exported PNG file containing the set of pixel-wise annotations

(c) The saved PNG file contains the set of annotations for this image in an efficient, language-agnostic format, where each pixel is given a grayscale value (1-256) associated with the label ID. 0 corresponds to "No Label". The PNG may appear to be just a black image when low grayscale values are present (labels such as 1-30) as the more vibrant white grayscale pixels appear as the value approaches 256. In MATLAB, the function '*imshow*' can be used, along with multiplying the values of the PNG file, for appropriate visualization.
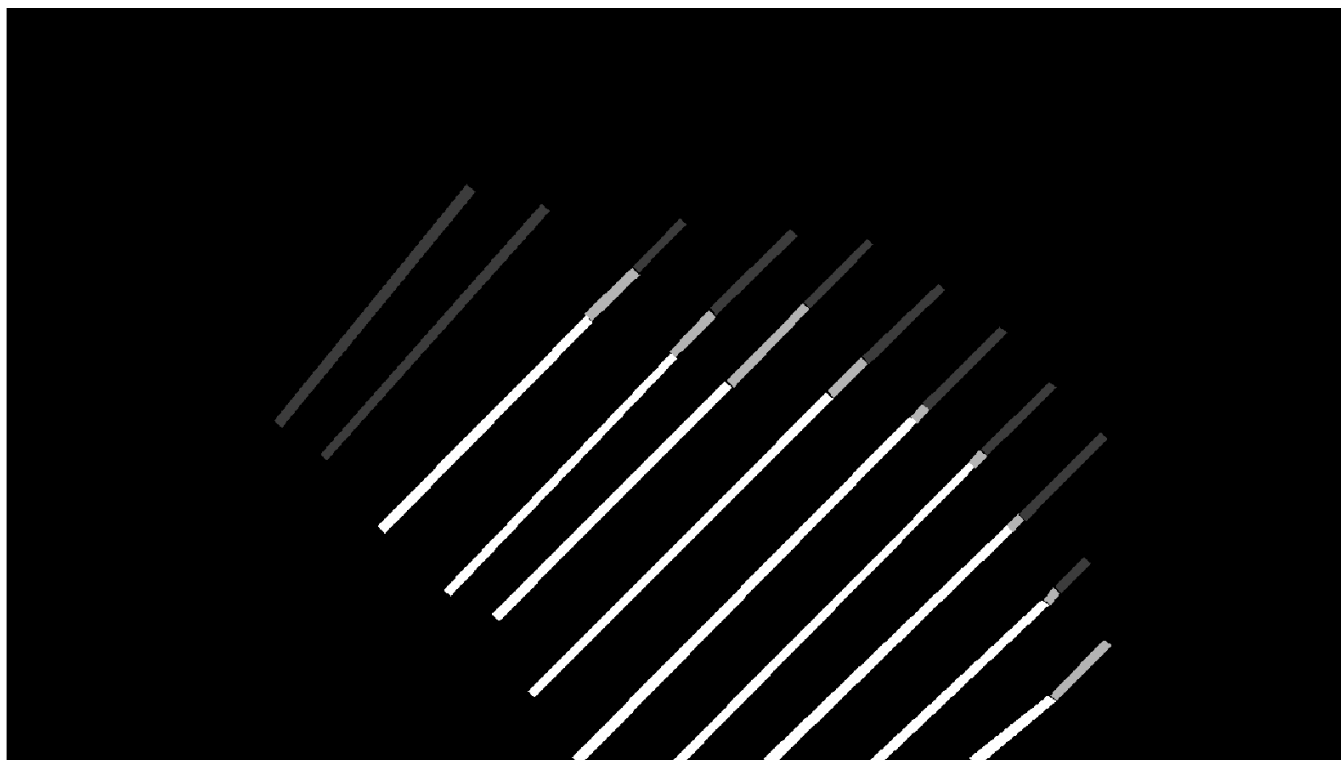
Figure 7: Data 720x1280 annotation mask exported from the *Data Annotation Tool* using the *Test Dataset*.

(9) The final tool comprised in this code is the *Visualization Generator* tool which can be accessed by launching the *Histopathology Data Annotation Toolkit V0.1* and selecting the *Visualization Generator* button on the Graphical User Interface (GUI).

The visualization generation tool allows for an exported set of annotations to be overlayed onto a white light image. An example generated from the *Test Dataset* is shown in Figure 8.

To generate an overlay for a set of labels:

(a) Click Generate Visualization from the overall tool selection window

(b) Select an exported label definition file (as used for the Image Labeler)

(c) Select a label annotations PNG image (generated using the labeling tool)

(d) Select a white light image jpg to overlay onto

(e) Select a name and location to save the visualization image

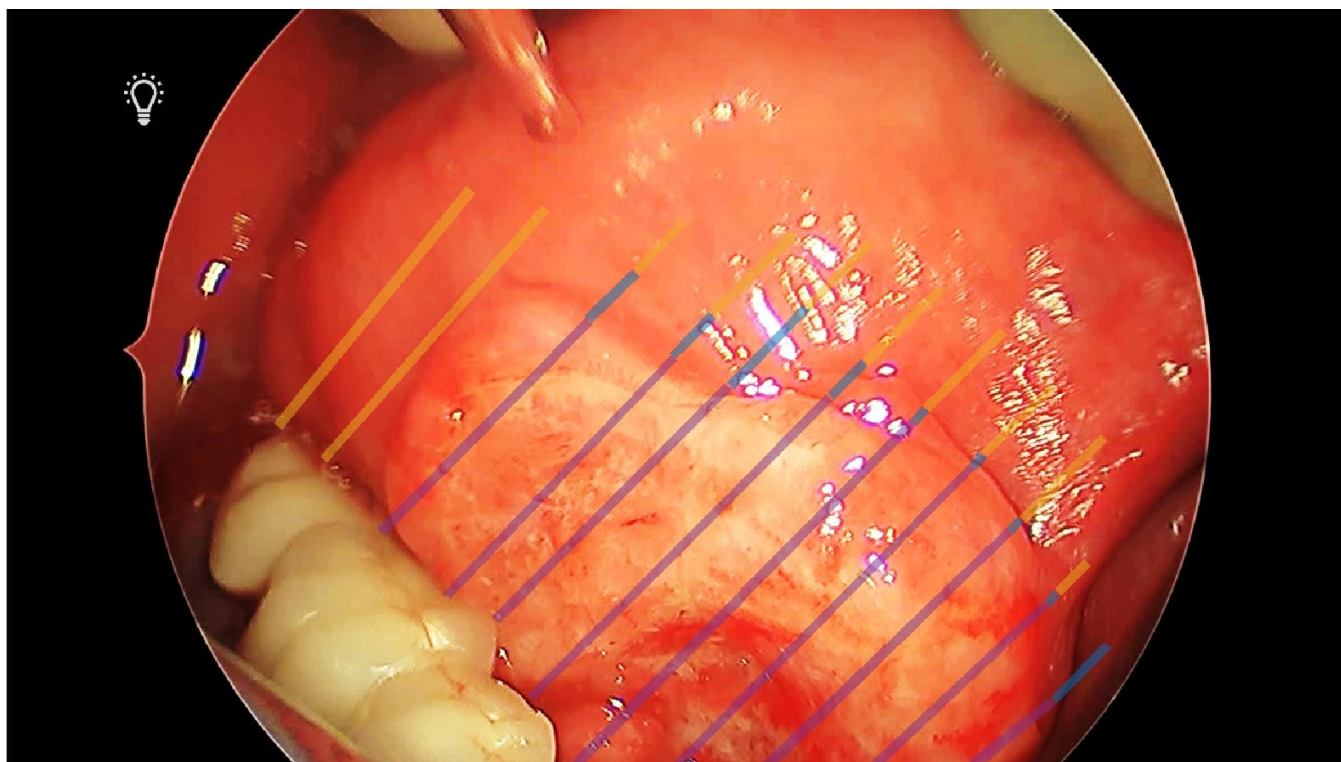(f) After being saved the visualization is saved to the code's '*Imtool*' window.

Figure 8: Representative example output from the visualization generation tool using the test dataset.

MODIFYING POWERPOINT EXPORT IMAGE DIMENSIONS & RESOLUTION

When using PowerPoint as a tool to create reference images (prior to loading such images in the *Data Annotation Tool*), precise selection of PowerPoint export image dimensions and resolutions can be performed to match pixel count and image dimensions of acquired surgical media.

Under PowerPoint >Design >Slide Size, default dimensions are a width of 13.333 in and height 7.5 inch. If alternate image dimensions are desired, these dimensions can be tailored here.

To control export resolution, type 'regedit' under the run windows on a windows platform. This pulls up the registry editor. Navigating to the path of Computer\HKEY_CURRENT_USER\SOFTWARE\ Microsoft\Office\16.0\PowerPoint\Options\ExportBitmapResolution, and entering a decimal value of 96 when PowerPoint default dimensions are used, will always yield and exported PowerPoint image of 720x1280 pixels. These dimensions and pixel count matched the video format our surgical videos generated and thus allowed for precise matching to the reference image submitted to the *Data Annotation Tool*.