**Automated tool for lung cancer lesion detection on histological images.**

-Developed by Wenfei Tang, Galban Lab, University of Michigan.

**1. Run your images on our pre-trained model**

We have provided a pre-trained model in the path “trained\_model\_20\_5\_1\_8\_45\_9\_updated.mat", which is ready to use. Follow the following steps to obtain prediction result generated from this model.

1. Create a folder to hold all the images you would like to run our model on. **Only include valid image files in this folder.** Our code does support all image formats accepted in “imread” (e.g. jpg, png, tif, etc.).
2. Run the script “generate\_overlaid\_result.m” under the root folder. It will prompt you to select a folder. Select folder you create from step (1).
3. It will automatically create two folders under the folder you provided, named “pred\_result” and “binary\_maps”. Under “pred\_result”, you can find the overlaid visualization of the prediction result; under “binary\_maps”, you can find the binary maps of the prediction result.
4. It may take a while to finish running the model. It takes about 30 minutes to run the model on one image from our example dataset.

**2. Train a new model**

If you would like to train a new model using a new dataset, follow the below steps:

1. Divide your images into two groups: images with cancerous lesions and images without cancerous lesions;
2. For cancerous group: go to the folder "training\_data/cancer\_images/real\_images", delete any existing images and put the original cancerous images here; go to the folder "training\_data/cancer\_images/binary\_maps", delete any existing images and put the binary maps of cancerous images here; (Note: You can generate the binary maps using the semi-automated tool)
3. For non-cancer group: go to the folder "training\_data/noncancer\_images", delete any existing images and put the original non-cancerous images here;
4. Provide a txt file named "file\_list.txt" under the folder "training\_data". We have provided an example “file\_list.txt”. See Appendix A for more clarification on the file format.
5. Run "train\_model.m" under the root folder;
6. You will find the model ready under the folder "model". named after date.

**Appendix A**

<#Number of non-cancerous images>

<img\_1>

<img\_2>

... (original non-cancerous lung images filenames)

<#Number of cancerous images>

<img\_1>

<img\_2>

... (original cancerous lung images filenames)

<binary\_map\_1>

<binary\_map\_2>

... (binary prediction maps for cancerous images from the semi-automated tool filenames)

Example:

5

LC-2\_Lung\_433.jpg

LC-2\_Lung\_490.jpg

LC-3\_523\_Lung\_1x.jpg

LC-3\_532\_lung\_1x.jpg

LC-3\_607\_Lung\_1x.jpg

6

LC-2\_Lung\_434.jpg

LC-2\_Lung\_444.jpg

LC-2\_Lung\_451.jpg

LC-2\_Lung\_502.jpg

LC-3\_517\_Lung\_1x.jpg

LC-3\_527\_Lung\_1x.jpg

434.png

444.png

451.png

502.png

517.png

527.png

Note:

Example provided under the same path;

Make sure the filenames match. i.e. img\_1 must match the order of binary\_map\_1;