ENSC 474

Final Project Report

Han, Qiao Chu

301187810

Email: qiaochuh@sfu.ca

Section 1: Exploration

1. a. Date： April 1st.
   b. Method： Graph Cut
   c. Results:

   Most of the lung tissue have been cut out as the mask. Though there are lines for Foreground and Background Mask. Some edge of the lung was not able to be included to the front mask. Partial success.

   d. Explanation:
   This tool requires input such as Foreground Mask and Background Mask definition aka manual input. To be precise, this tool requires multiple inputs to fully include the lung and exclude the background.

2. a. Date： April 2nd.
   b. Method： Global image threshold using Otsu's method (Graythresh)
   c. Results:
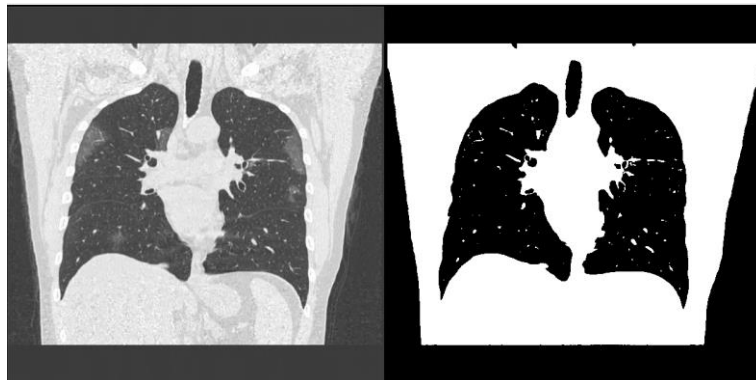
Partial success. Background of body tissue has been removed, though most of black borders are still there.
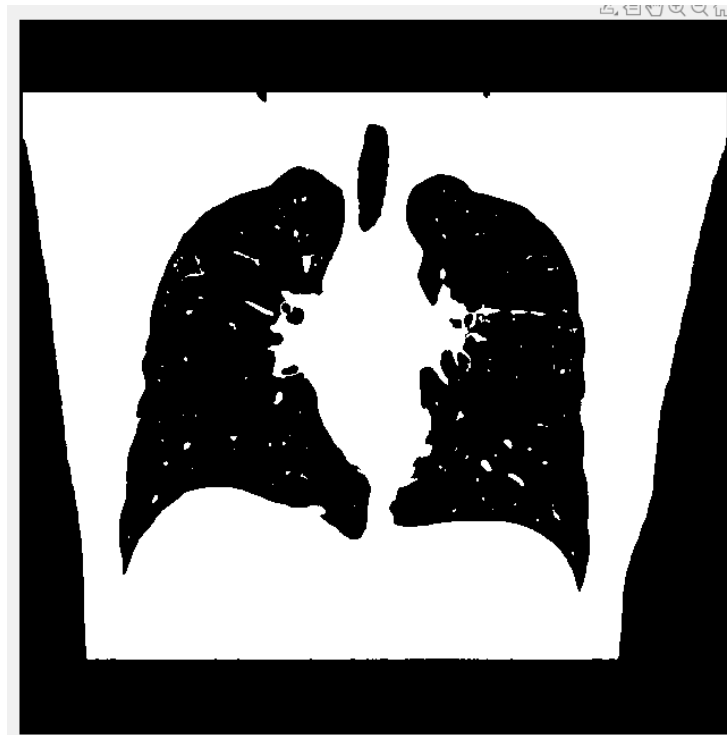
d. Explanation:

Border of the image has similar intensity as the lung tissue. They are recognized as the part that being kept.

3. a. Date：April 5th.
   b. Method：K-means clustering based image segmentation
   c. Results:

Partial success. Similar to Graythresh. Background of body tissue has been removed, though most of black borders are still there.
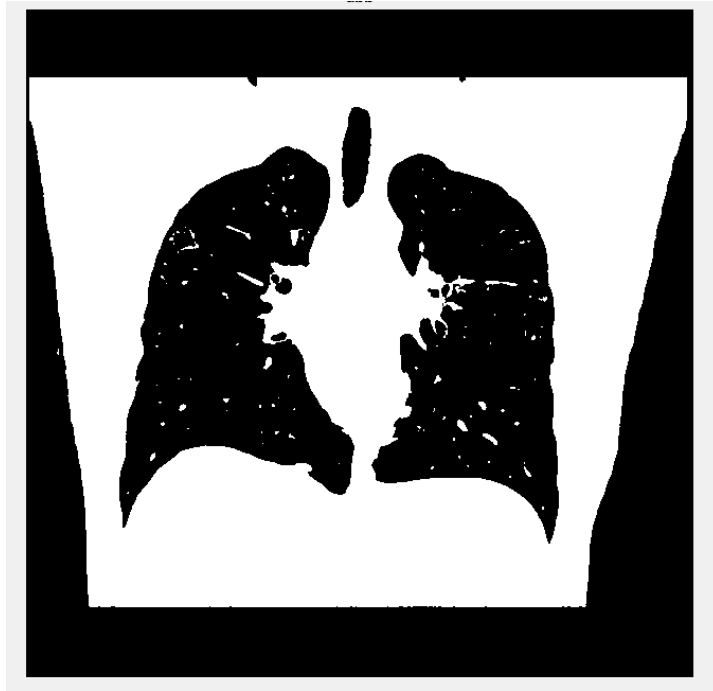
d. Explanation:

Same as Graythresh.

4. a. Date：April 6th.
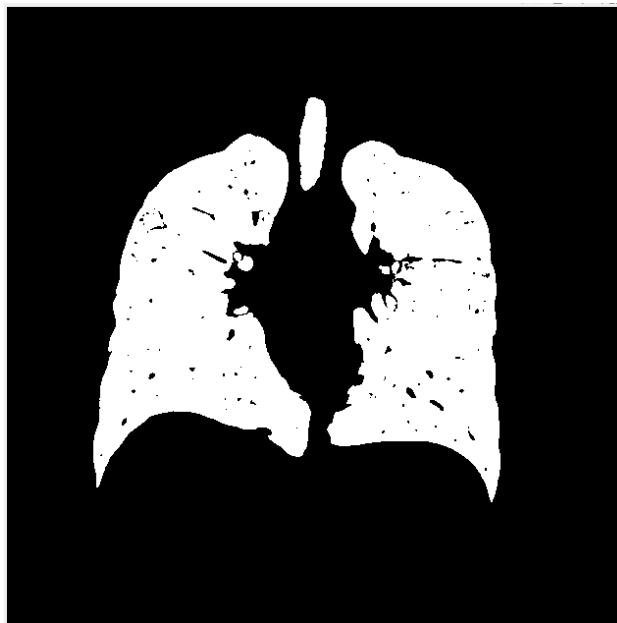   b. Method：Remove border： imclearborder
   c. Results:

Fail, borders were not removed.

d. Explanation:
'imclearborder' removes the high values on the image border. Borders on this image appears black. Image have to be flipped (complement) before the border clearing process.
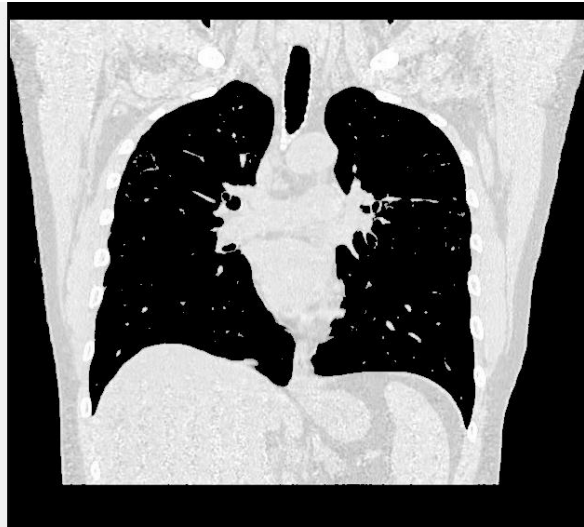
5. a. Date：April 7th.
   b. Method：Remove border: imclearborder
   c. Results:

Partial Success, borders were removed. However, some of the lung tissue has been removed as well.

d. Explanation:

Some of the lung segment such as the blood vessels are removed in the previous process. It's because the blood vessel tend to be have similar image intensity values as the body tissue.
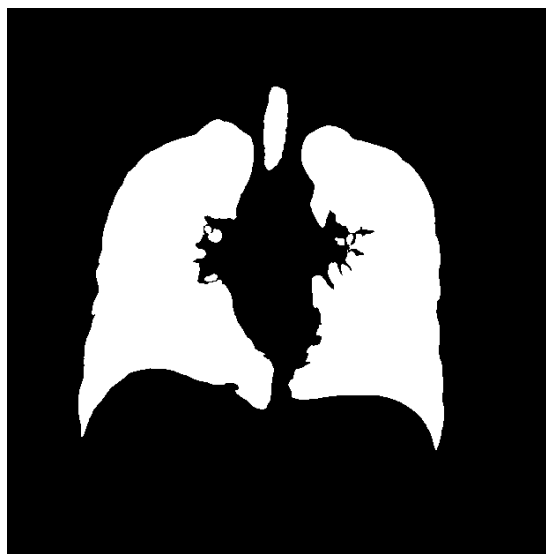
6. a. Date：April 10th.
   b. Method：Apply mask on to original image Trial
   c. Results:

Fail, Body tissues were kept but lung tissues were removed. Also, borders haven't been removed.

   d. Explanation:

I used the wrong mask, I used the complement value of the mask values; the lung area mask should be 1 and body should be 0.

7. a. Date：April 11th.
   b. Method：Use 'bwareaopen' to fill the blood vessel in the lung mask
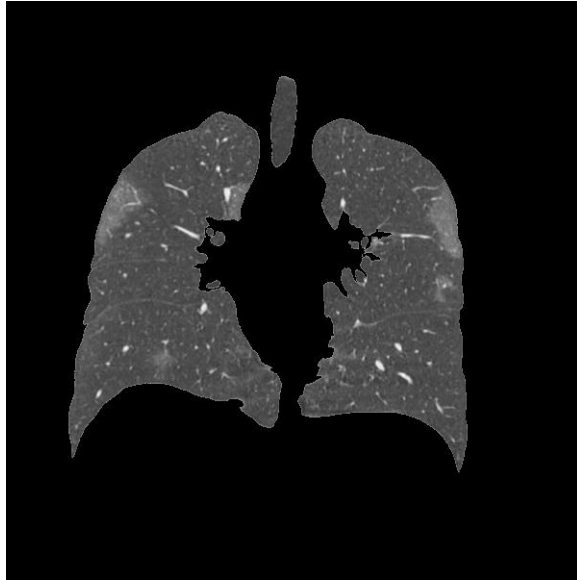   c. Results:

Success.

d. Explanation:

'bwareaopen' is very good at removing small objects from binary image.

8. a. Date：April 12th.
   b. Method：Apply mask on to original Image again
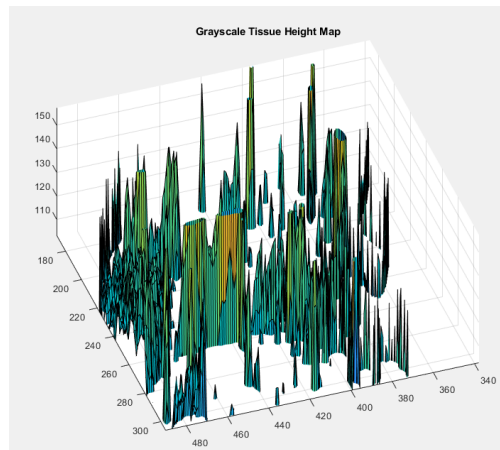   c. Results:

Success.

d. Explanation:

Dot product of the original image and the mask can create the segmented lung image I want. In the mask, the lung area has the value of '1', every where else have the value of '0'.

9. a. Date：April 16th.
   b. Method：Histogram of the Image
   c. Results:

Grayscale Tissue Height Map
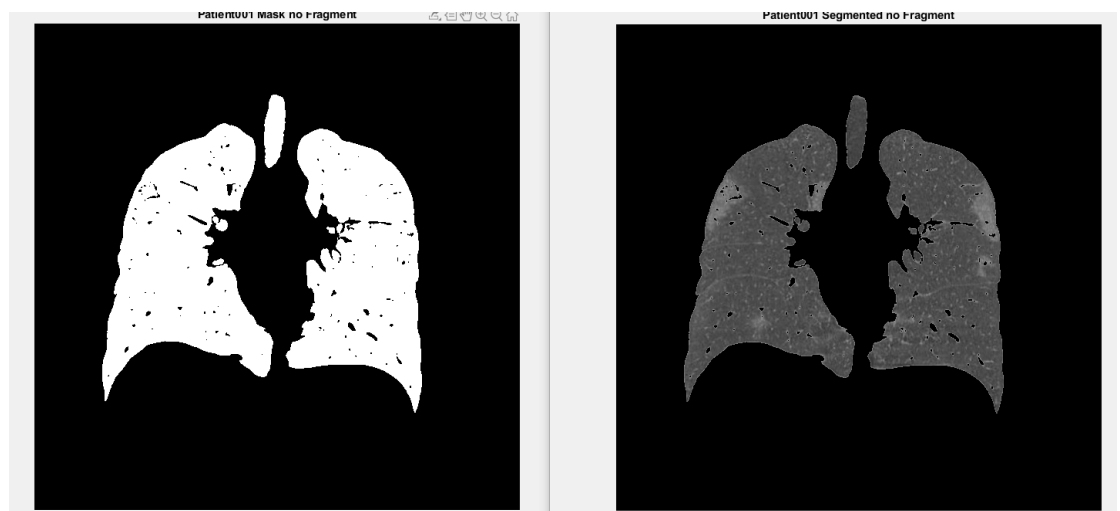
Success.

d. Explanation:

I used 'surf' to inspect the histogram of the segmented lung.

It is clear that most of the infected areas have intensities above 100 to 120.

10. a. Date：April 17th.
    b. Method：Apply new mask without blood vessel on to original Image
    c. Results:

Success.

d. Explanation:

Even though using the 'bwareaopen' can help to segment the whole area of the lung. But blood vessels were also kept on the image. Those blood vessels might interfere the calculation of infected area in the next step. So I used the mask before the 'bwareaopen' process for the next step.

11. a. Date：April 18th.
    b. Method：Testing thresh hold for healthy and infected area.
    c. Results:

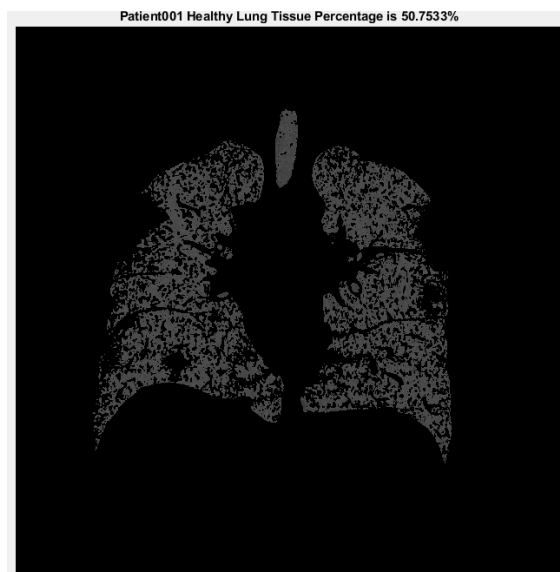Fail, some of the infected area was identify as the healthy tissue

    d.   Explanation:

       I set the threshold to 120, any area with intensity greater than 120 will be identify as infected and removed. Apparently, the threshold is too high.

12. a.   Date：April 18th.

    b.   Method：Testing thresh hold for healthy and infected area.

    c.   Results:

Patient001 Healthy Lung Tissue Percentage is 50.7533%

Fail, some of the heathy areas were identify as the infected tissues.
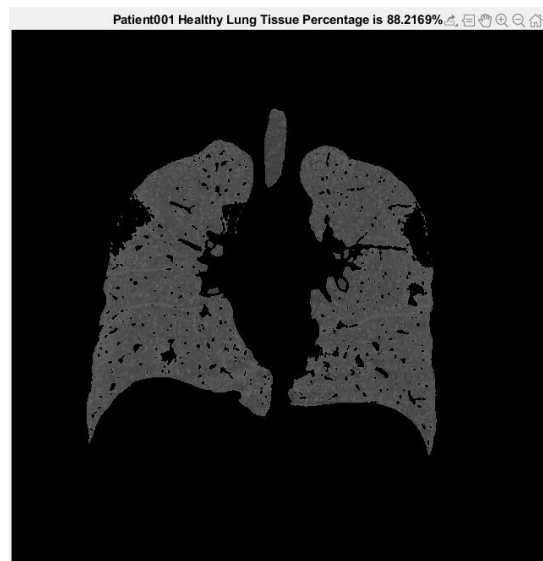
    d.   Explanation:

   Same as before, but the threshold is too low.

13. a.   Date：April 19th.

    b.   Method：Testing thresh hold for healthy and infected area

    c.   Results:

Patient001 Healthy Lung Tissue Percentage is 88.2169%

Success.　Heathy area was successful detected and segmented as well as the percentage.

14. a.　Date：April 20th.
    b.　Method：Testing recognition of infected area.
    c.　Results:

Patient001 Infected Lung Tissue Percentage is 12.3659%
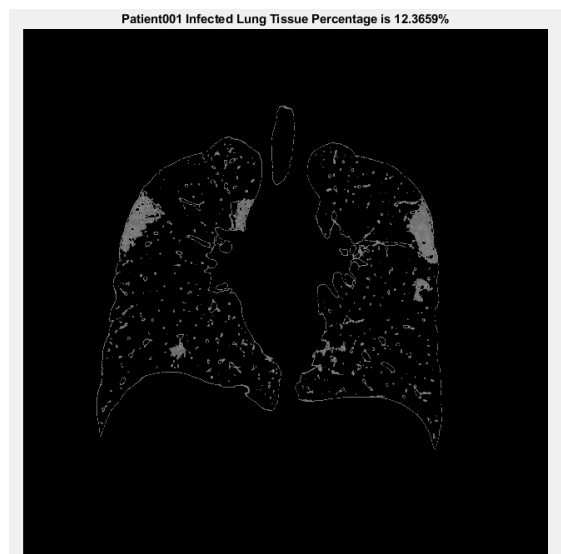
Partial Success. Even though most of the infected areas were successfully recognized. The edges of the lung and the remaining small blood vessels were also identified as the infected area.

    d.　Explanation:
    Those false positive areas have intensity above 100, thus identify as the infected area.

15. a.　Date：April 21th.
    b.　Method：Removing edges and blood vessels with 'imerode'
    c.　Results:

Patient001 Infected Lung Tissue Percentage is 2.0334%
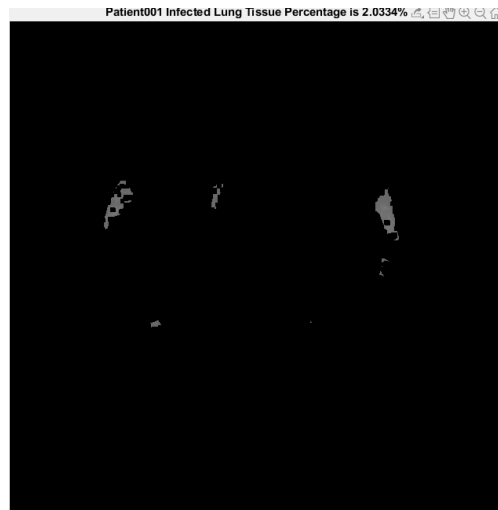
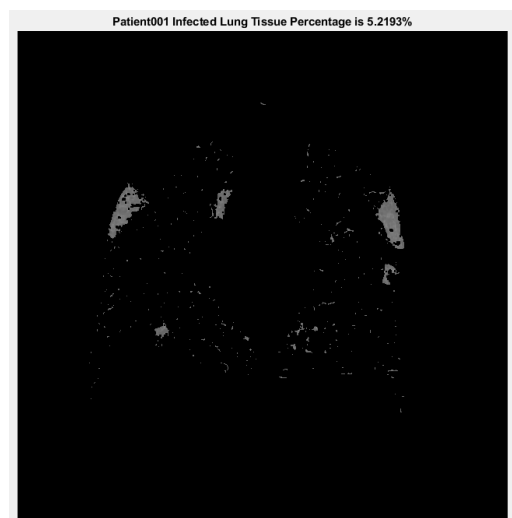Fail, too much contents were removed

    d.   Explanation:

That's because the 'imerode' mask was set to high.

16. a.   Date：April 21th.

    b.   Method：Removing edges and blood vessels with 'imerode'

    c.   Results:

Patient001 Infected Lung Tissue Percentage is 5.2193%

Success, all edges has been removed, only small portion of the blood vessels remain.

    d.   Explanation:

The 'imerode' mask was set to right size.

Section 2: Proposed Solution

By observing the patients' images such as figure 1, we can see that those images are mostly divided into two type of areas. The lighter areas contain human body except the lung area and darker areas contain human lung and background. In order to segment the lung area, we should consider clustering method to segment the image by its pixel values or intensity of each elements in the image matrix.

In the image segmentation process is partitioning the digital image into distinct regions containing each pixel with similar attributes, in this case it's the lung area with darker color (low intensity value).
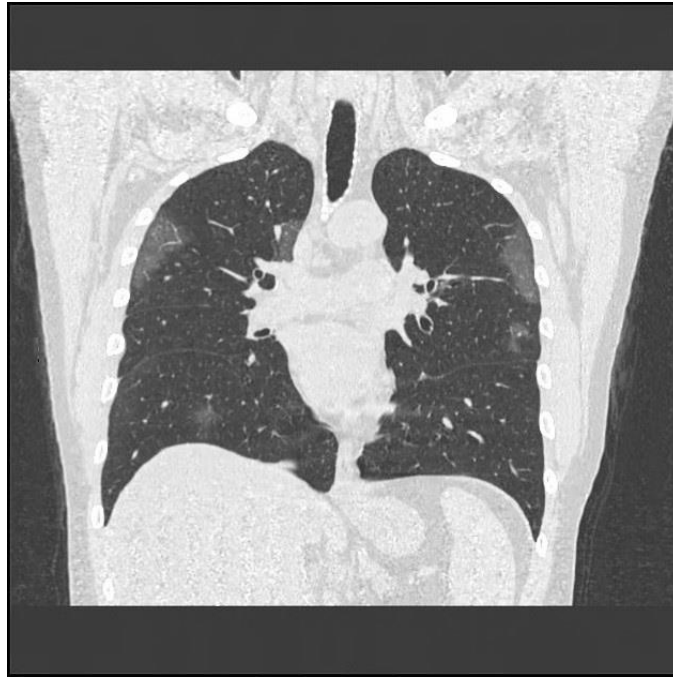

Figure 1-1: Patient 001 Original Lung Image

The method I used for lung tissue segmentation is called K-means clustering based image segmentation.

K-Means clustering algorithm is an unsupervised algorithm and it is used to segment the interest area from the background. It clusters or partitions the given data into K-clusters or parts based on the K-centroids.

The algorithm is used when you have unlabeled data. The goal is to find certain groups based on similarity in the data with the number of groups represented by K.

Steps in K-Means algorithm:

1. Choose the number of clusters K.

2. Select at random K points as the centroids.

3. Assign each data point to the closest centroid which forms K clusters.

4. Compute and place the new centroid of each cluster.

5. Reassign each data point to the new closest centroid.

6. If any reassignment took place, go to step 4, otherwise, the model is ready.

After the K-means clustering, I obtained the binary image mask as figure 2. In this mask

the pixels in lung areas and background have the value of 1 and pixels in blood vessel areas and other body areas have the value of 0. We can apply this mask to original image to obtain the segmented lung. But before that, we have to remove the background and the blood vessels.
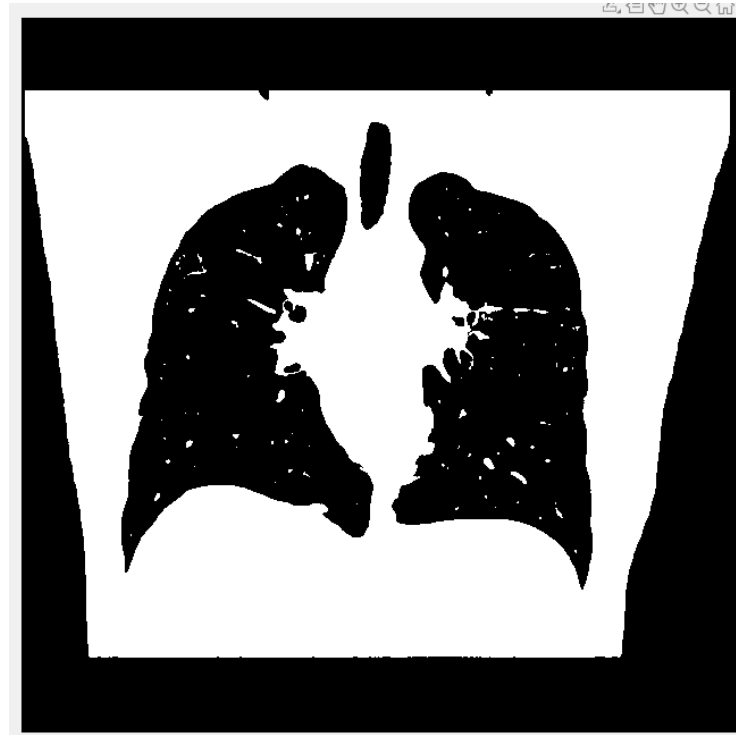

Figure 2: binary image mask.

In order to remove the image black background, I used the MatLab build in function called 'imclearborder' to suppress light structures connected to image border.

J = imclearborder(I) suppresses structures in image I that are lighter than their surroundings and that are connected to the image border. Use this function to clear the image border. For grayscale images, imclearborder tends to reduce the overall intensity level in addition to suppressing border structures. The output image, J, is grayscale or binary, depending on the input. In this case both the input and output are binary image. The output is as shown below in figure 3.
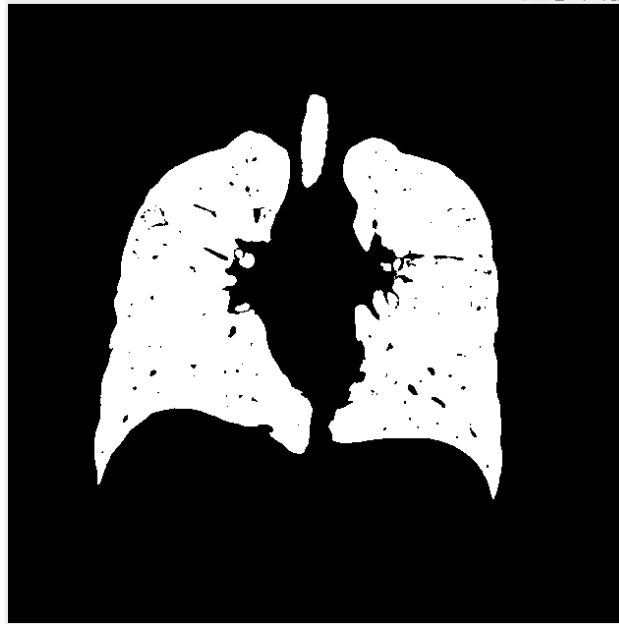
Figure 3: complement binary image mask without the border.

In order to remove the blood vessels, I used the MatLab build in function called 'imclearborder' to remove small objects from binary image.

BW2 = bwareaopen(BW,P) removes all connected components (objects) that have fewer than P pixels from the binary image BW, producing another binary image, BW2. This operation is known as an area opening. The output is as shown below in figure 4.
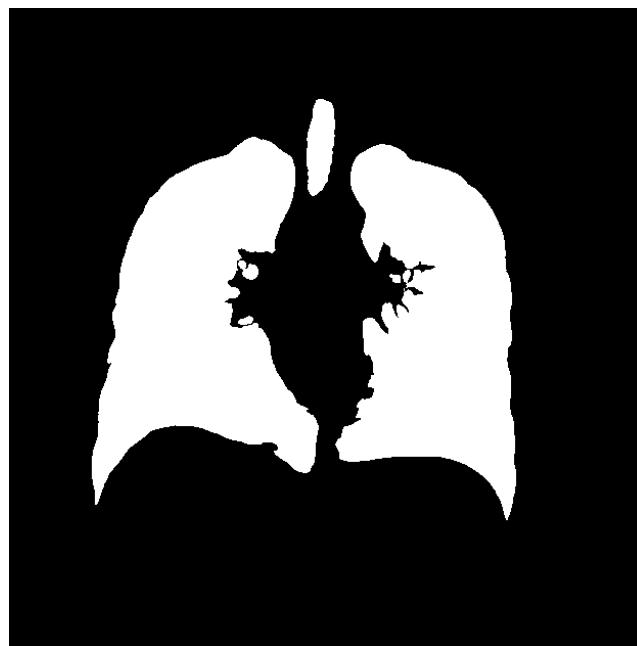

Figure 4: complement binary image mask without the border and vessels.

After flipped the mask image in figure 4 (find the complement) and dot product with the original image in figure 1, we can obtain the segmented lung tissue in figure 5.
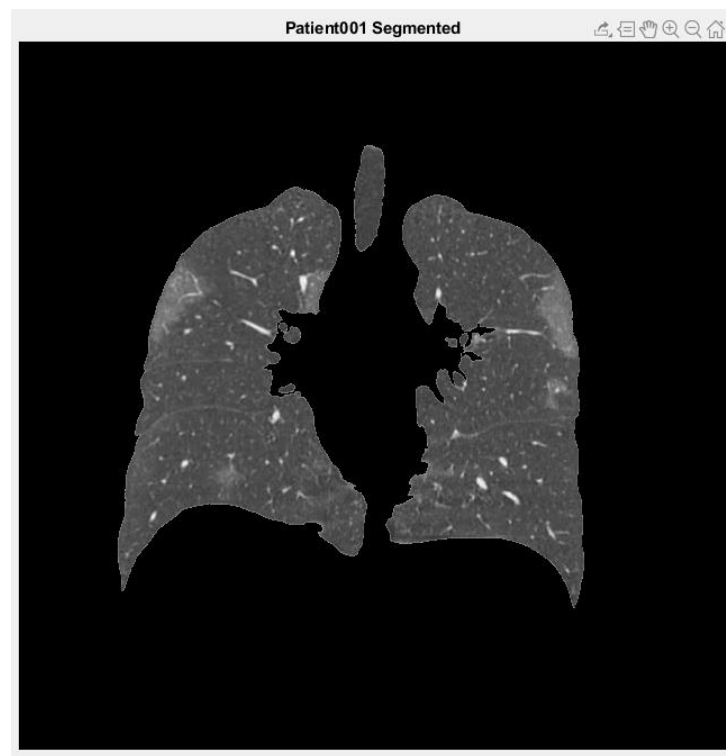
Figure 4: Segmented Lung of Patient 001.

Before We proceed into the next step, I have to remove the large vessels from the segmented lung to prevent them interfere with the infected area detection. I applied the mask before 'bwareaopen' process onto the original image figure 1 to obtain a segmented lung image without the large blood vessels.
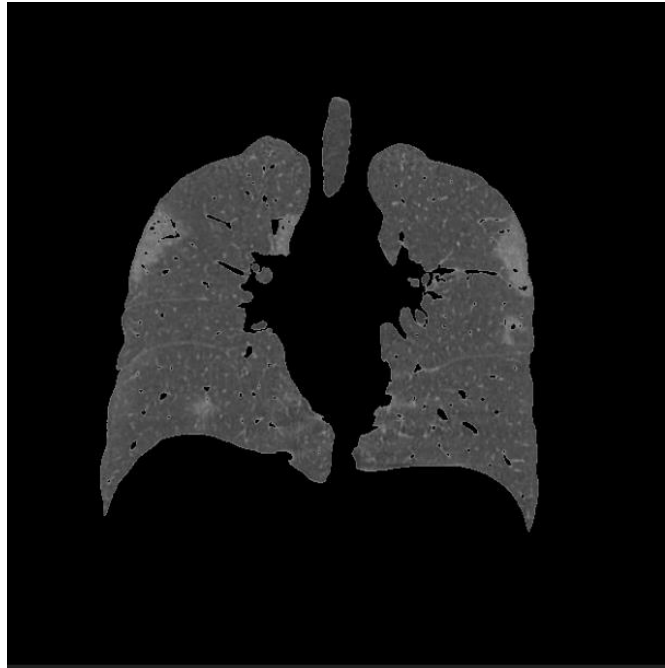
Figure 5: Segmented Lung of Patient 001 without Large Blood Vessels.

For next step we need to investigate the segmented lung image in figure 5. By using the 'surf' function we can obtain the histogram of the lung image in figure 6.
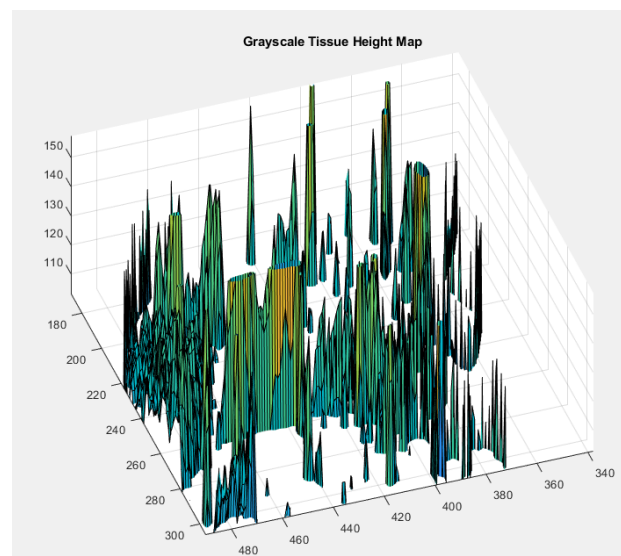
Figure 6: Partial Height Map of Lung Tissue Image

. As we can see in figure 6. Most of pixels in the infected areas have values greater than 100.

Therefore, for segmented healthy tissues, I find all the pixels in with value higher than 100 and replace them with 0, then generate new binary mask and apply it onto original image in figure 1. The results are shown in figure 7.
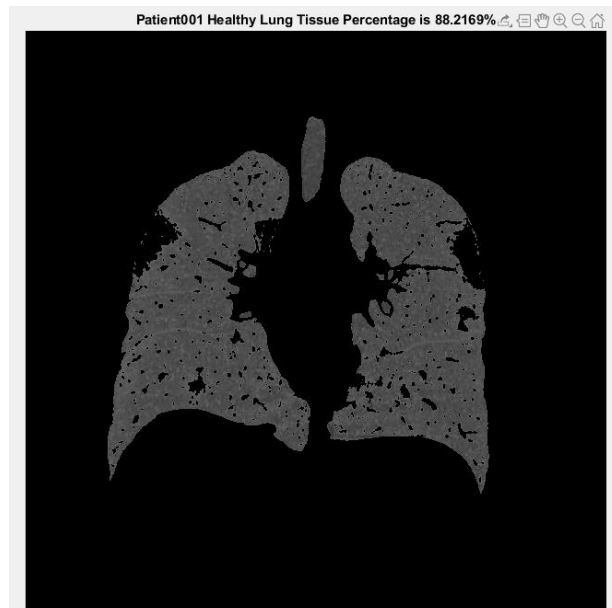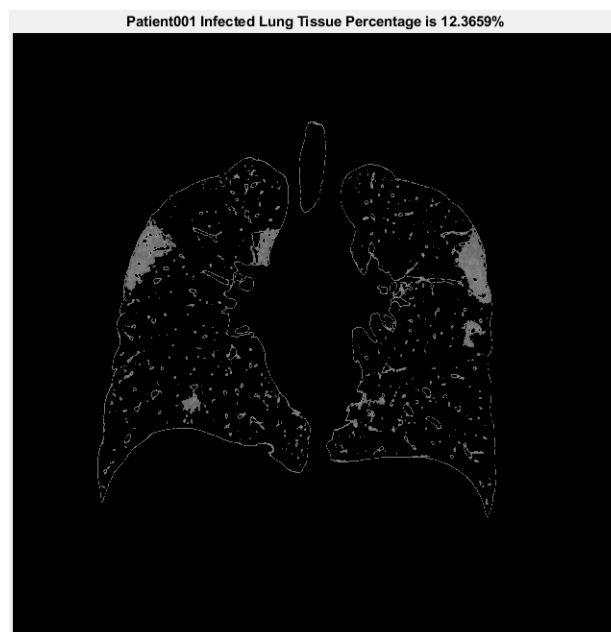
Figure 7: Segmented Healthy Lung of Patient 001 With its percentage.

I am also able to calculate the percentage of the healthy lung tissues by compare the number of the non-zero pixels in figure 7 and figure 4. As you can see in figure 7, this lung is only 88% healthy, rest of them are infected.

For segmented infected tissues, I encountered some issues. I find all the pixels in with value less than 100 and replace them with 0, then generate new binary mask and apply it onto original image in figure 1. The results are shown in figure 8.

Figure 8: Testing Segmented Infected Lung of Patient 001 With its percentage.

As you can see in figure 8, the edges of the lung and some small blood vessels are also detected as infected areas. In order to remove those false positive area from this image. I used the MatLab build in function called 'imerode' to erode this image. J = imerode(I,SE) erodes the grayscale, binary, or packed binary image I, returning the eroded image, J. Correct result image can be seen in figure 9. Though this method is not 100% perfect, there are some tiny vessels remain in the output image, but most of the blood vessels and all the edges have been removed.
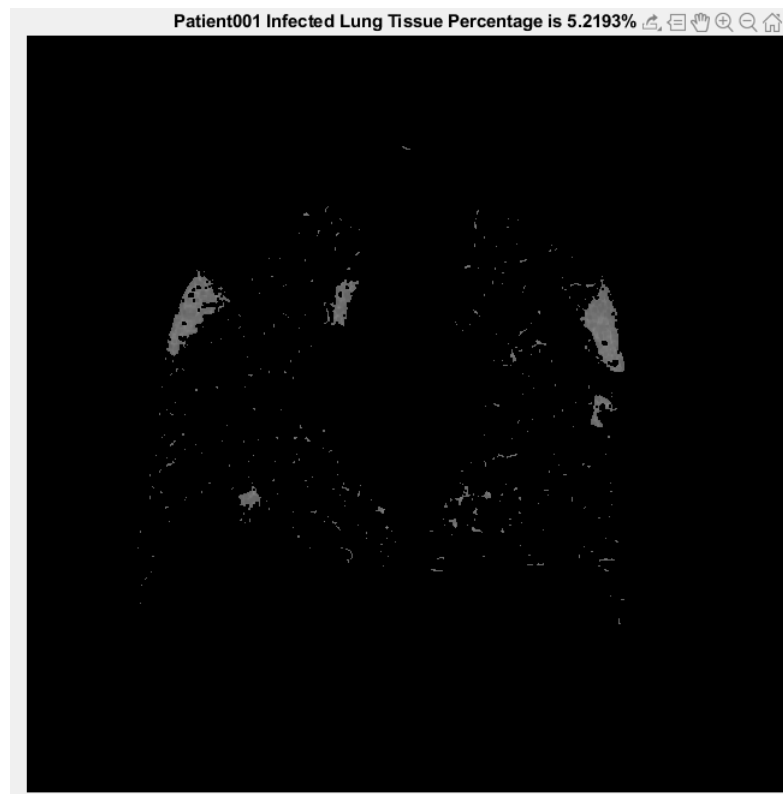
Figure 9: Segmented Infected Lung of Patient 001 With its percentage.

From the results above, I am certain that patient 001 is infected with COVID-19, though his situation is not severe, the infected percentage is only around 5%.

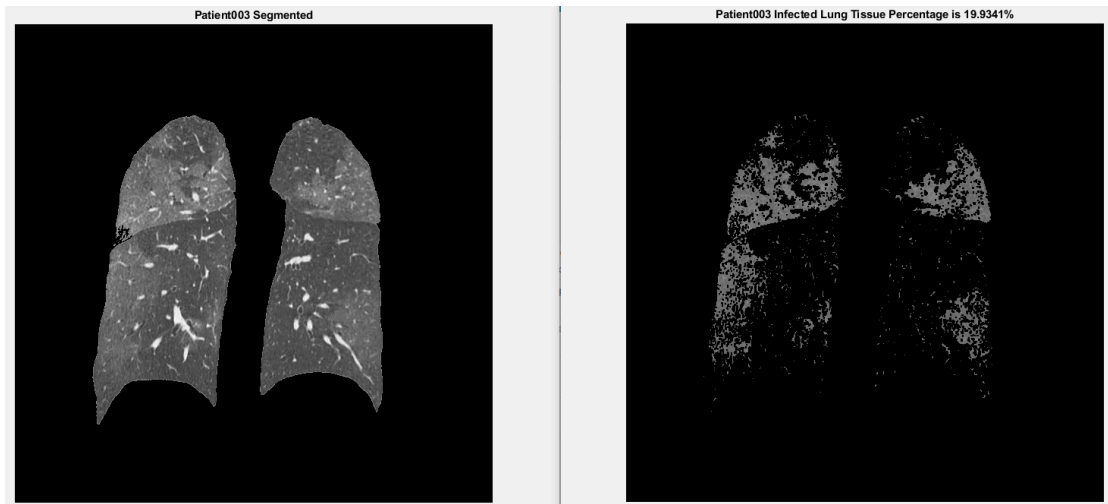Medical Reports

Patient 001:
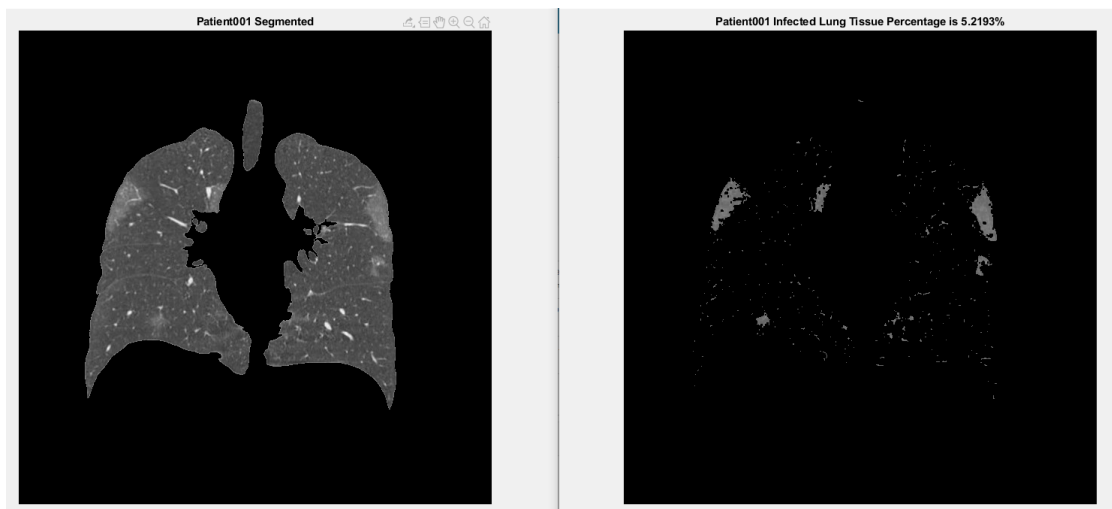
Patient 001 is infected with COVID-19, with the infected rate of 5.2193%.
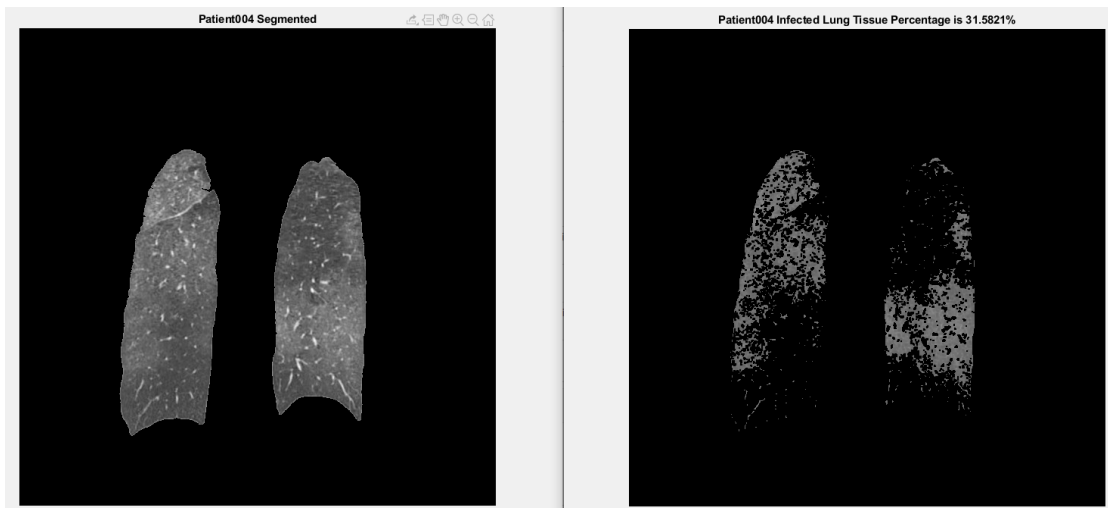
Patient 002:
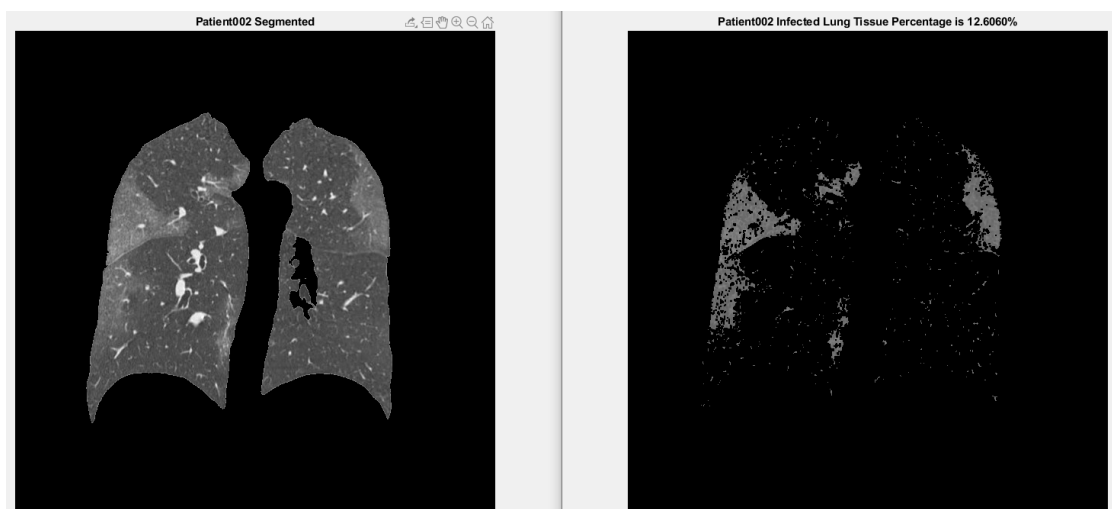
Patient 002 is infected with COVID-19, with the infected rate of 12.6060%.
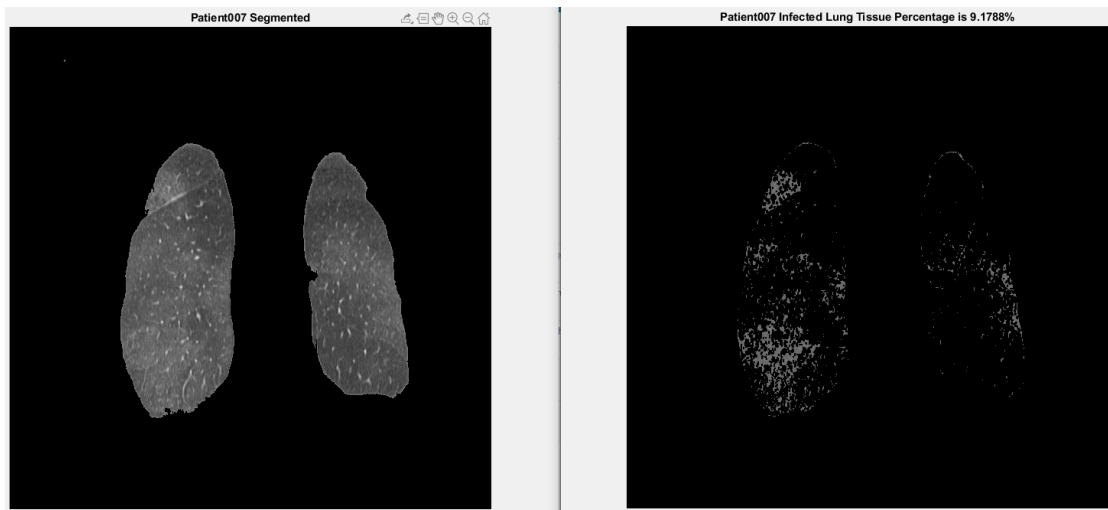
Patient 003:

Patient 003 is infected with COVID-19, with the infected rate of 19.9341%.
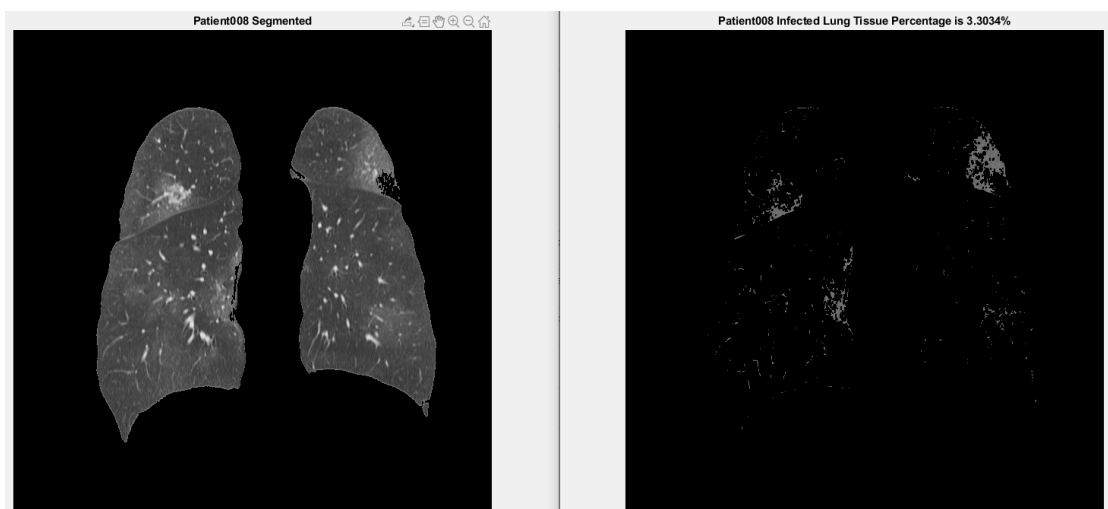
Patient 004:

Patient 004 is infected with COVID-19, with the infected rate of 31.5821%.
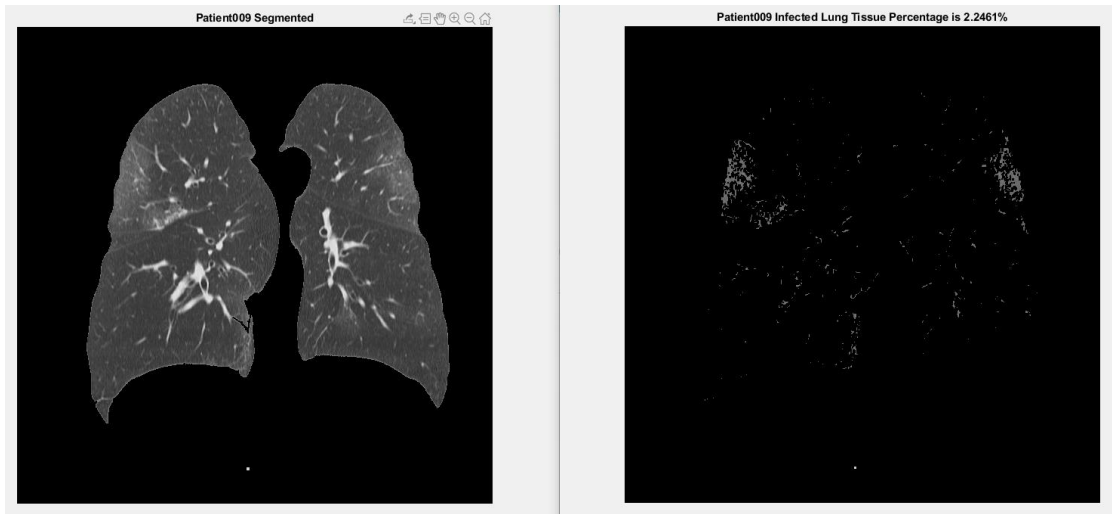
Patient 007:

Patient 007 is infected with COVID-19, with the infected rate of 9.1788%.
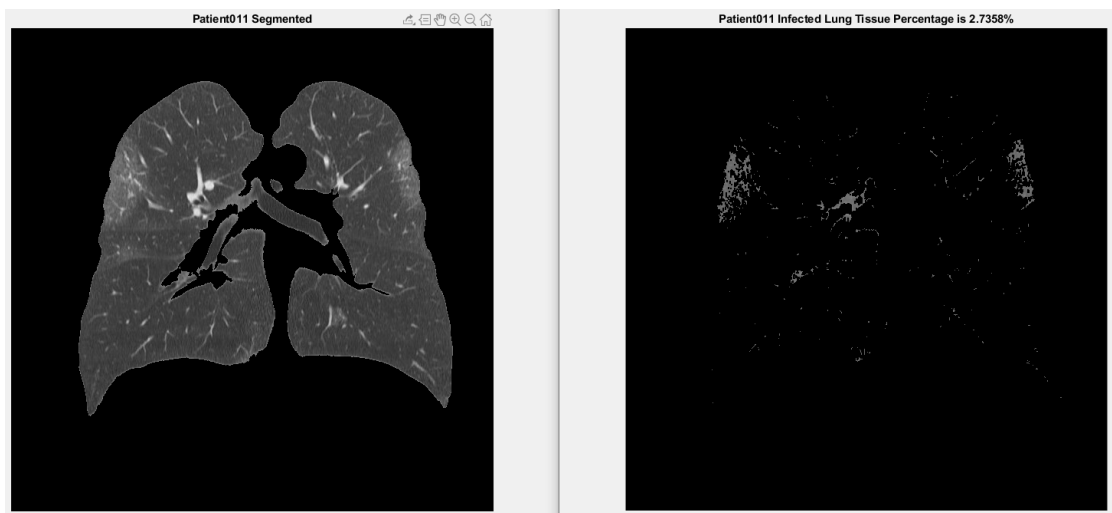
Patient 008:

Patient 008 is infected with COVID-19, with the infected rate of 3.3034%.
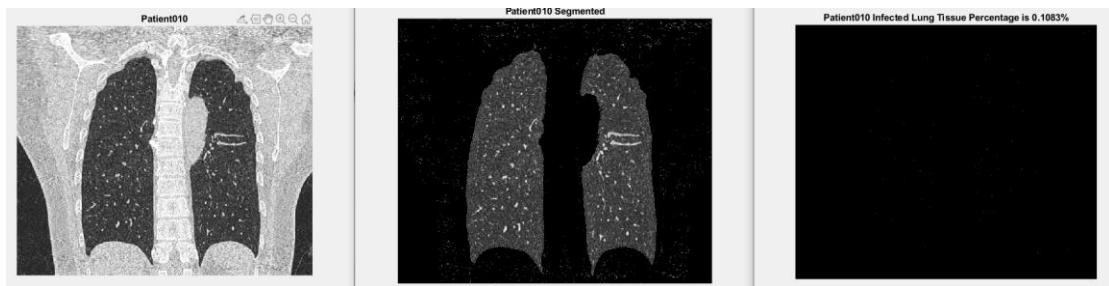
Patient 009:

Patient 009 is infected with COVID-19, with the infected rate of 2.2461%.

Patient 011:

Patient 011 is infected with COVID-19, with the infected rate of 2.7358%.

Patient 010:

Patient010 seems not infected, but its hard to tell since the original image has too much noise at the first place.

Failed Cases:

The lung segmentation of patient 005, 006, and 012 has failed, the segmentation for patient 010 partially successes but a lot of noise remains.

That's because in those original images, the gray scaled values from the lung areas and other areas have very small difference, it is very hard for the function to distinguish them apart during clustering process. K-means process requires setting initial centroids and converge after each iteration. For those cases, the processes failed to find correct centroids and stopped at local optimal solution.
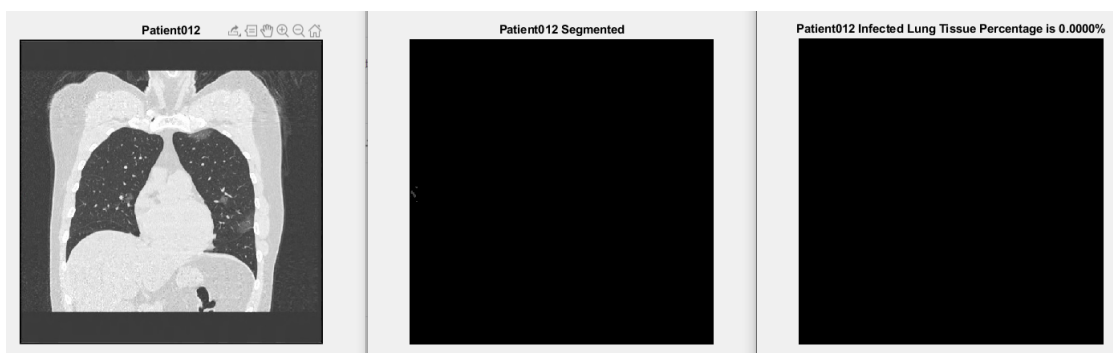
Patient 005:
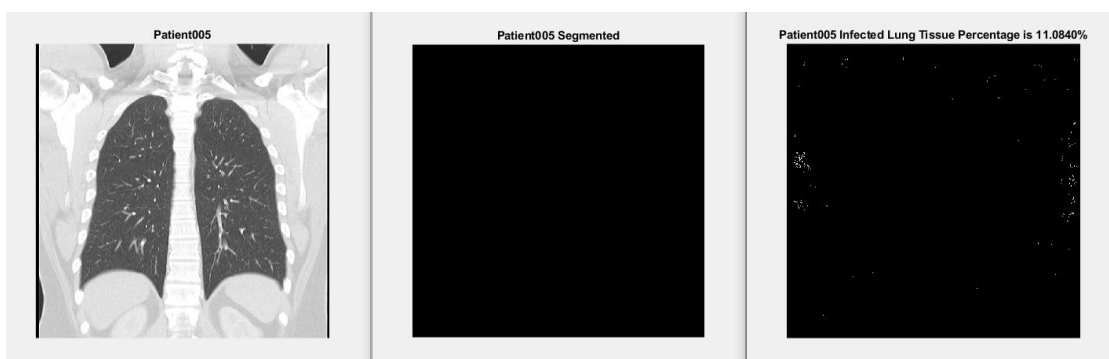
Patient 006:
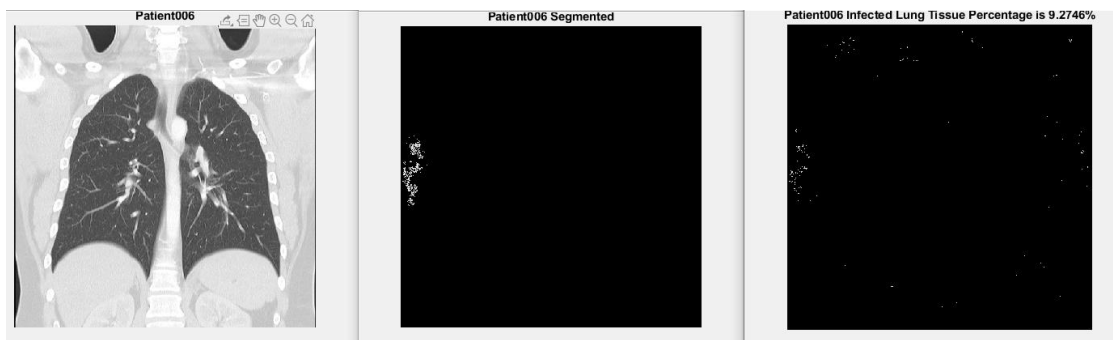
Patient 012:

Reference:

segmentImage.m is Auto-generated by MatLab imageSegmenter app on 05-Apr-2020