Assignment 1

Name: Yang Hao-Wei, Student ID: 109062526 03/22/2021

1 Part 1

1.1 Read in and print out all the data fields in a DICOM file

Inthis section, I use the patient of folder "0a0c32c9e08cc2ea76a71649de56be6d" and the first slice "0a67f9edb4915467ac16a565955898d3.dcm" to print its dataFields. The result is shown below.

```
---------------------------- Part 1: Read and Print Dicom DataFields ------------------------------
Dataset.file_meta -------------------------------
(0002, 0000) File Meta Information Group Length  UL: 192
(0002, 0001) File Meta Information Version       OB: b'\x00\x01'
(0002, 0002) Media Storage SOP Class UID         UI: CT Image Storage
(0002, 0003) Media Storage SOP Instance UID      UI: 1.2.840.113654.2.55.240087524148038410985780799448670801102
(0002, 0010) Transfer Syntax UID                 UI: Explicit VR Little Endian
(0002, 0012) Implementation Class UID            UI: 1.2.40.0.13.1.1.1
(0002, 0013) Implementation Version Name         SH: 'dcm4che-1.4.31'
-------------------------------------------------
(0008, 0005) Specific Character Set              CS: 'ISO_IR 100'
(0008, 0016) SOP Class UID                       UI: CT Image Storage
(0008, 0018) SOP Instance UID                    UI: 1.2.840.113654.2.55.240087524148038410985780799448670801102
(0008, 0060) Modality                            CS: 'CT'
(0008, 103e) Series Description                  LO: 'Axial'
(0010, 0010) Patient's Name                      PN: '0a0c32c9e08cc2ea76a71649de56be6d'
(0010, 0020) Patient ID                          LO: '0a0c32c9e08cc2ea76a71649de56be6d'
(0010, 0030) Patient's Birth Date                DA: '19000101'
(0018, 0060) KVP                                 DS: None
(0020, 000d) Study Instance UID                  UI: 2.25.60037070027156423276159501017920151735078954137544798194660
(0020, 000e) Series Instance UID                 UI: 2.25.58703274222857573910779974742342423982066946347485459782406
(0020, 0011) Series Number                       IS: "1"
(0020, 0012) Acquisition Number                  IS: "1"
(0020, 0013) Instance Number                     IS: "45"
(0020, 0020) Patient Orientation                 CS: ''
(0020, 0032) Image Position (Patient)            DS: [-160.100006, -142.500000, -93.410004]
(0020, 0037) Image Orientation (Patient)         DS: [1, 0, 0, 0, 1, 0]
(0020, 0052) Frame of Reference UID              UI: 2.25.10816424791949633522058525471061381137400269917518537485573
(0020, 1040) Position Reference Indicator        LO: 'SN'
(0020, 1041) Slice Location                      DS: "-93.410004"
(0028, 0002) Samples per Pixel                   US: 1
(0028, 0004) Photometric Interpretation          CS: 'MONOCHROME2'
(0028, 0006) Planar Configuration                US: 0
(0028, 0010) Rows                                US: 512
(0028, 0011) Columns                             US: 512
(0028, 0030) Pixel Spacing                       DS: [0.664062, 0.664062]
(0028, 0100) Bits Allocated                      US: 16
(0028, 0101) Bits Stored                         US: 12
(0028, 0102) High Bit                            US: 11
(0028, 0103) Pixel Representation                US: 0
(0028, 0120) Pixel Padding Value                 US: 0
(0028, 1050) Window Center                       DS: "40.0"
(0028, 1051) Window Width                        DS: "400.0"
(0028, 1052) Rescale Intercept                   DS: "-1024.0"
(0028, 1053) Rescale Slope                       DS: "1.0"
(0028, 1054) Rescale Type                        LO: 'HU'
(7fe0, 0010) Pixel Data                          OW: Array of 524288 elements
```

Figure 1: all datafields of a single DICOM file

1.2 Statistics for raw data and Hounsfield Unit data

I use the formula, HounsfieldUnit = RescaleScope * Raw + RescaleIntercept, to convert raw data to Hounsfield units. Since RescaleSlope is 1 and RescaleIntercept is -1024, basically the entire distribution only shifts to the left by 1024 units. To be noted, based on the hint from TAs, I set the value of raw data pixel array to 0 if the value of the raw data < PixelPaddingValue while transforming to Hounsfield Unit.

```
------------------------------------- Part 1: Print Statistics -------------------------------------
Data Type       Min Data        Max Data        Mean Data       Std Data
Raw Data        0               2362            471.449         495.317
Housefield Data -1024.0         1338.0          -552.551        495.317
```

Figure 2: the statistic of raw data and Hounsfiled

2 Part 2

2.1 Read in and sort DICOM files

I first read in 25 slices from the first folder. Then, I found that the attribute "ImagePositionPatient" has three dimensions, the first two dimensions are the same in all 25 slices, but the third dimension(z-axis) is different. Therefore, I sort the slices by ImagePositionPatient[2] from small to large.

2.2 Print out 25 slices in correct order

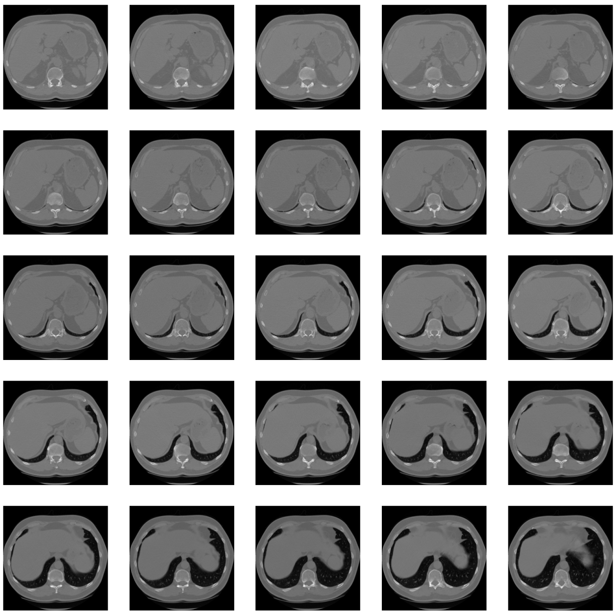

Figure 3: 25 images of sorted CT slices

3 Part 3

I tried "local mean" and "local median" to segment the lung by applying "thresholdlocal" from skimage to calculate the thresholds of each slice, and got a two-dimensional array. Then, I calculate the average of the array to get a float number, which is the red line drawn in the histogram of the figures below.

3.1 Segmentation - local mean

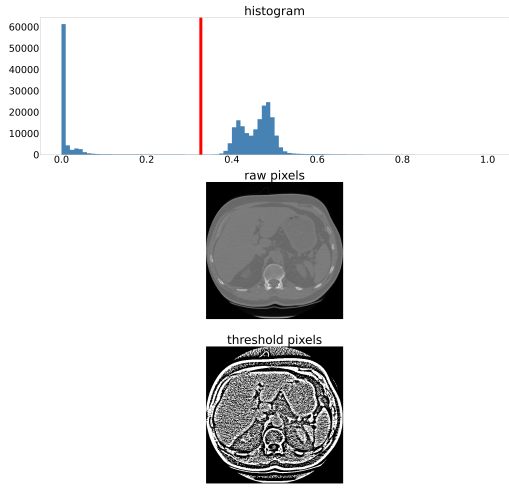

Figure 4: Segmentation by mean

3.2 Segmentation - local median

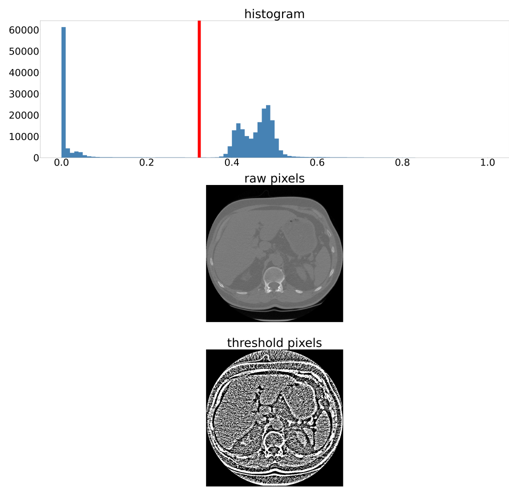

Figure 5: Segmentation by median

4 Bonus

By using the toolkit "Poly3DCollection" from "mplot3d.art3d.", I plot a 3D image of the CT slices from the first folder. I use mean as the threshloding strategy so that each slice has a mean threshold return, and then I calculate the average these thresholds as the overall threshold. The results are shown below.

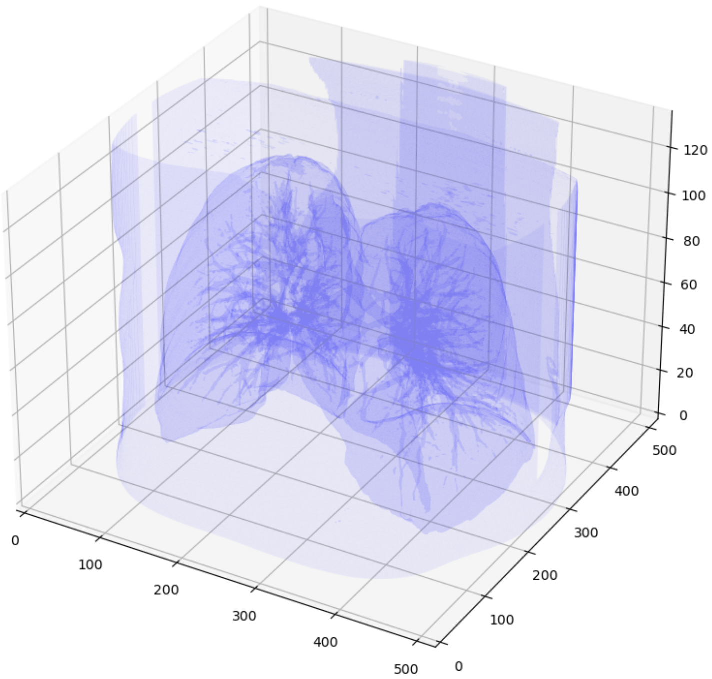

Figure 6: 3D visualization of the CT slices in the first folder

5 Question and Summary

In this assignment, I have learned the technique to process DICOM files. Also, I've tried several thresholding technique to segment CT slices. Moreover, I've learned how to visualize CT slices into 3D view point with the toolkit "Poly3DCollection". Last but not least, this assignment help me get more familiar with writing reports with Latex, which is what I had been planning to learn but not really trying.