Image Segmentation of Lung Data

Name: Karthick Sivakumar
Professor: Lifang He

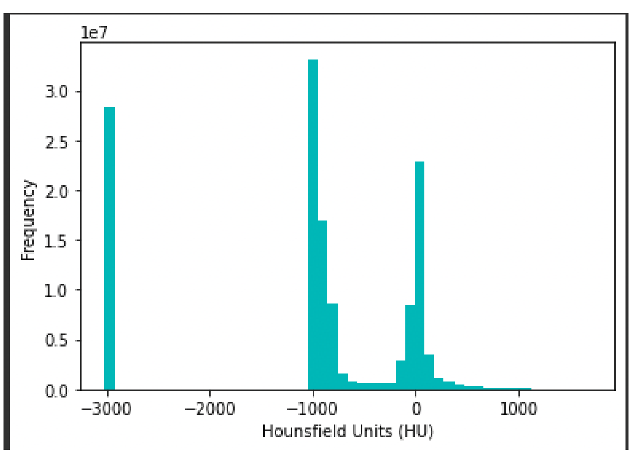
Introduction

Lung CT scans are widely used in the diagnosis and treatment of various lung diseases. These scans provide detailed images of the lungs, which are then used by medical professionals to identify and locate any abnormalities in the lung tissue. Image segmentation, a technique used to separate and extract objects or regions of interest from an image, is an important step in analyzing lung CT scans. In this paper, I will focus on image segmentation of lung CT scans using various tools and techniques. The paper is divided into two parts. In the first part, I will preprocess the images, segment the data, and plot a 3D visualization using plotly. In the second part, I will segment and visualize the lung CT scans using the 3D Slicer software. Through this paper, I aim to provide a comprehensive overview of the image segmentation techniques used in analyzing lung CT scans.

Part 1: Image Segmentation in Python

To produce a 3D CT image, individual CT images in the form of DICOM slices are compiled to form a unified whole. The DICOM header of the high-resolution CT image holds essential data on slice thickness and pixel spacing for each slice. It is worth noting that the pixel spacing is not always standardized and may differ from the desired 1 x 1 x 1 mm measurement. As a solution, we can resample the image to ensure that each voxel measures precisely 1 x 1 x 1 mm.

Another critical aspect we can obtain from the image is the distribution of Hounsfield Units for each pixel, which can reveal information about the most frequent HU and the material, such as water, air, or tissue, represented by a pixel's density. We can generate a histogram for this purpose:



Upon analyzing the histogram, we observe that the majority of the pixels in the image represent either air at approximately -1000 HU or tissue at roughly 0 HU. The number of pixels between -700 to -300 is relatively low, indicating that the image mainly contains air or tissue/bone with minimal variation in between. It's worth noting that there is a peak at -3000, which suggests the presence of noise in the image.

Further analysis of the histogram reveals that the image consists of a large quantity of air, and predominantly contains soft tissue such as muscle, liver, etc., with some fat. Bone is represented by only a small number of voxels (as seen in a tiny sliver of height between 700-3000). This observation implies that significant preprocessing is necessary if we want to analyze lung because only a small fraction of the voxels represent lung tissue.

Image Slices

By examining several slices of the CT image between z=60 to z=130, spaced 10 intervals apart, it's evident that there are distinct regions in the image containing either air or tissue/bone. The light gray regions in each image represent air, while the dark gray areas within the light gray regions indicate the presence of either lung tissue or the airway.

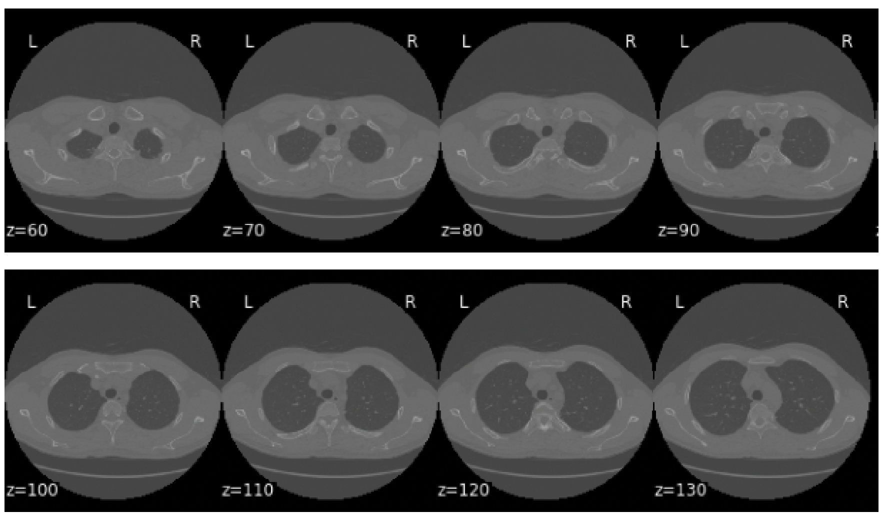


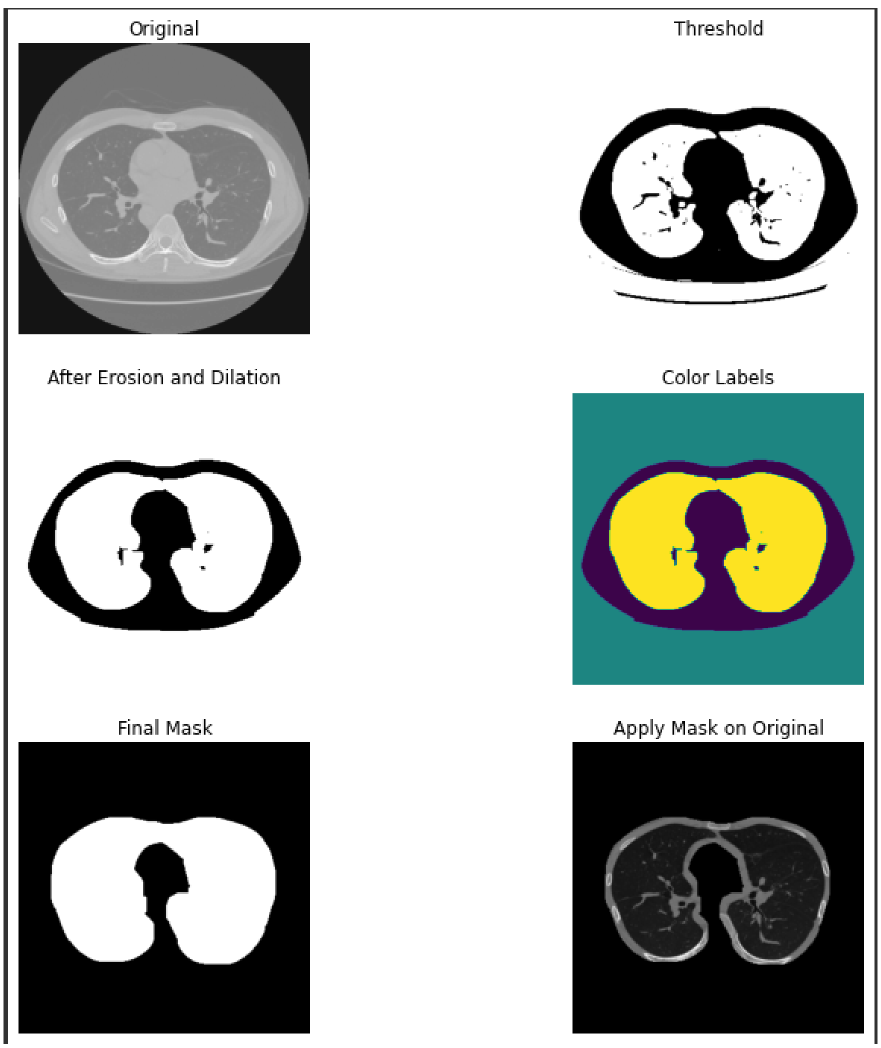
Image Segmentation and Masking

We use a hybridization of k-means clustering to cluster the background and foreground data, and thresholding to separate the foreground from background images. The resulting image is obtained after applying the threshold. The process yields several images by using a calculated threshold to segment the data.

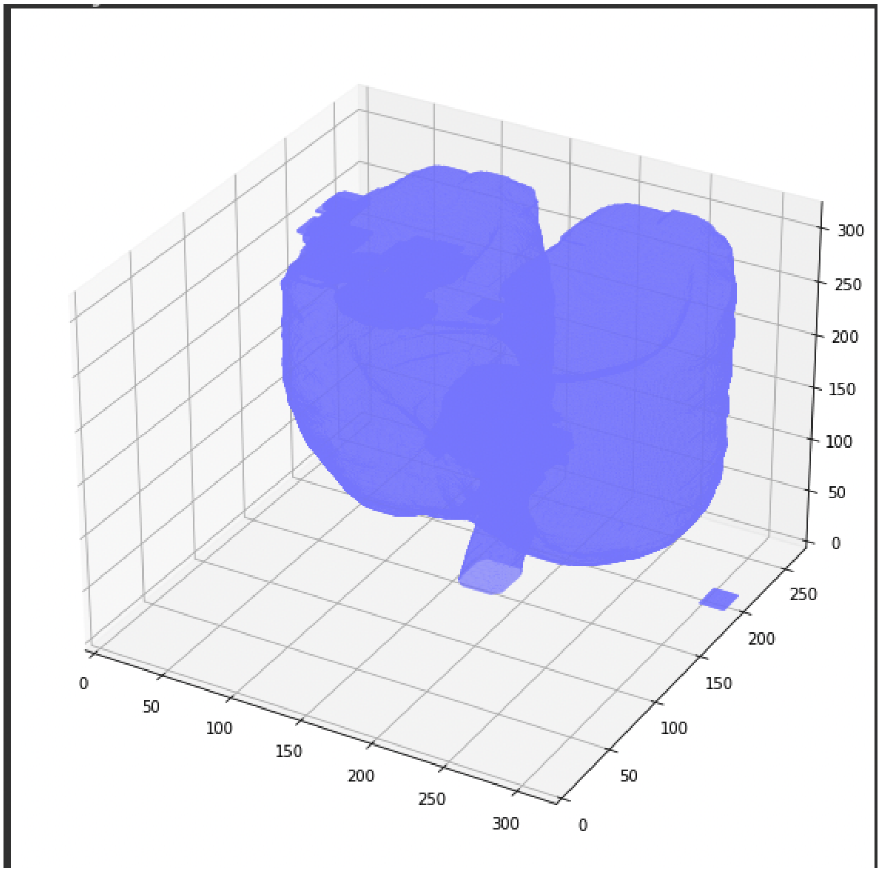
After applying the threshold, the next step is to remove minor areas within the mask that may be mistaken for significant regions. If left unremoved, these minor areas may result in the removal of essential regions. To address this issue, erosion is first applied to the threshold image to remove the minor areas. This operation involves removing pixels from the boundaries of regions until they shrink to the desired size. Then, dilation is applied to the resulting image, which is the opposite of erosion and involves adding pixels to the boundaries of regions until they reach the desired size. By combining erosion and dilation, the resulting image shows that many small black dots within the lungs are removed, leaving only the significant regions intact.

After obtaining the binary mask using thresholding and morphological operations, the next step in our image segmentation process is to apply the mask onto the original image. By doing so, we are only retaining the pixel values that correspond to the segmented regions of interest, which in this case are the lung and airway regions. The pixel values outside of these regions are removed, effectively eliminating the air and background information from the image.

The resulting output image provides us with a clearer and more defined view of the areas of interest, allowing us to focus on the lung and airway regions without any distraction from the background.



3D Image Plot

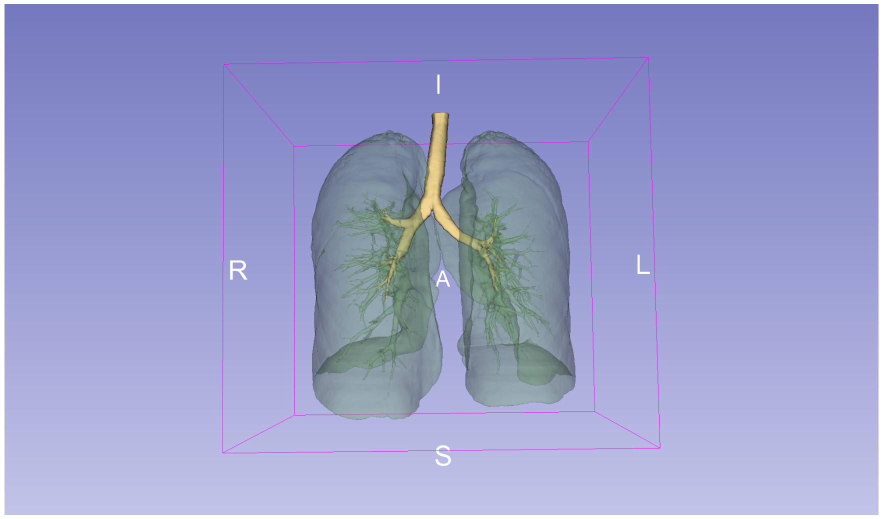


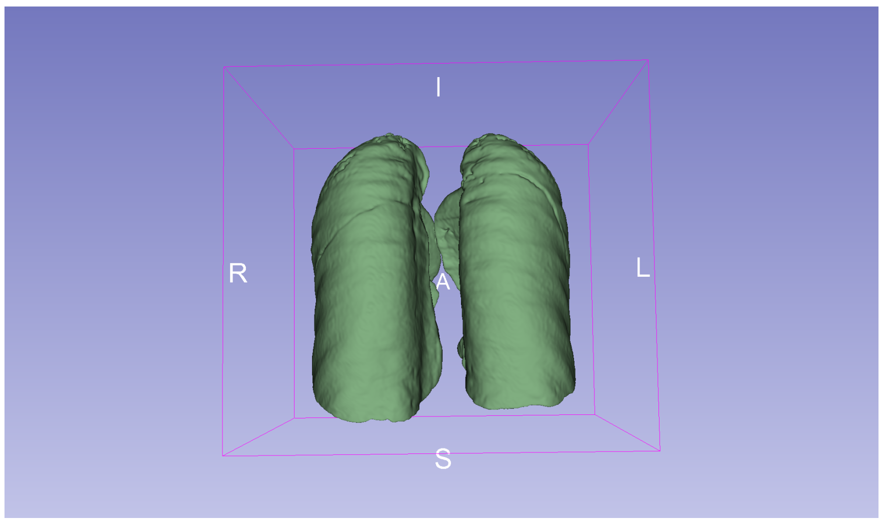
Part 2: Segmentation and Visualization in 3D Slicer

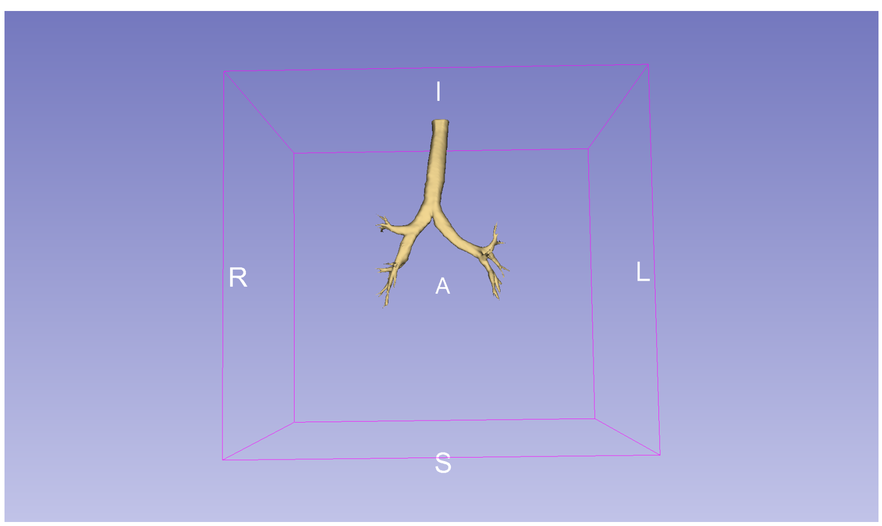
To accurately visualize and segment the lung and airway, I utilized 3D Slicer. At first, I attempted to manually segment the image using the threshold function in the segmentation editor. However, I realized that this approach was both tedious and prone to inaccuracy.

I pivoted to performing the "grow from seed" operation instead. With this method, I painted slices with lung and airway labels and used the "grow from seeds" option to automatically segment the data. This approach proved to be much more effective, requiring minimal cleaning and producing a more accurate result.

After completing the segmentation, I performed a visual inspection to ensure that the results were accurate. I found that the segmentation includes the entire regions of both the lungs and airways, even differentiating regions where the airway comes close to the lung. The segmentation was able to accurately differentiate between lung tissue and airway space, effectively removing air and background regions while conserving the lung and airway regions.


I
R
A
L
S


I
R
A
L
S


I
R
A
L
S

Conclusion

In conclusion, segmenting the lung and airway regions in CT images is a crucial step in diagnosing and treating various pulmonary diseases. The first step of the first part was to normalize the Hounsfield Units (HU) values for each pixel in the image. This is done by subtracting the mean of all pixel values from the image data and dividing the result by the standard deviation, scaling all pixel values to between 0 and 1. The second step involves applying a threshold to the image data to separate the foreground from the background. The third step involves using erosion and dilation to remove small artifacts and noise from the image. From here we can see the segmented image and the 3D visualization.

One effective tool for segmenting the lung and airway regions is 3D Slicer. A grow from seed operation can be used to automatically segment the data, which proved to be effective and required minimal cleaning. Upon visual inspection, the segmentation appears to be accurate, including the entire regions of both the lungs and airways, even differentiating regions where the airway comes close to the lung.