Segmentation of CT Lung Scans Using Image Processing Techniques

Terry Griffin
COMP.5230 Computer Vision I
Spring 2018
terence_griffin@student.uml.

***Abstract*— A system for lung segmentation of Computed Tomography (CT) chest scans is presented. The approach uses typical image processing techniques, following the most successful methods in the literature. The system was tested on CT scans from three datasets. The results show that the system operates with a high level of success and compares well with other methods.**

Keywords—lung segmentation, image processing, computed tomography

I. Introduction

This project investigated the current state of the art methods for automatic segmentation of computed tomography (CT) lung scans. The focus was on approaches which make use of image processing techniques, rather than deep learning methods. A system was implemented which successfully takes a CT lung scan as input and produces labels identifying the left and right lungs. Enhancement of the system to include segmentation of the lung lobes was not successful.

Lung cancer continues to be one of the leading causes of death in the U.S. and world wide [1]. Early detection through the use of CT scans is an important step in increasing the chance of patient survival. Computer-aided diagnosis (CAD) systems can greatly improve the efficiency and effectiveness of physicians in evaluating these scans, a process which tends to be time intensive and laborious.

The first step in automatic analysis of a CT scan is separating the lung tissue from the blood and air vessels and the surrounding bone and musculature. A further, and more difficult task, is to identify the five lobes which make up the two lungs. The result of these two steps is a labeled map of the CT scan which can be used by additional systems to identify and categorize potentially cancerous tumors.

The system developed for this project makes use of image processing techniques such as thresholding, filtering, morphological operations, and watershed algorithms. One goal of this investigation was to identify the best techniques for solving this problem without using deep learning approaches. This work can then be used as a baseline for comparing deep learning approaches to the segmentation problem.

To ensure that the system provides a general solution, it was tested on multiple datasets, include those from the the LUng Nodule Analysis 2016 (LUNA16) challenge [2], the LObe and Lung Analysis 2011 (LOLA11) challenge [3], and additional scans from the Tianchi dataset on the UML server. The use of multiple datasets provided an additional challenge, as some approaches which initially showed promise with one dataset failed to generalize to the others.

The resulting system is able to identify the two lungs with a high degree of success. All of the tested scans were successfully segmented. A qualitative analysis using the reference segmentation data from the LUNA16 dataset shows that the generated segmentation is highly reliable.

Segmentation of the lungs into lobes was not successful. Details of the attempted approach and the problems encountered are given in section III.

II. Background

Lung segmentation has a relatively long history. Brown et al. [4] approached this task by combining image processing techniques with anatomical knowledge and an inference engine in 1997. After the LOLA11 dataset was made available, Po et al. demonstrated the applicability of Browns method [5] on this larger dataset.

Most approaches for lung segmentation follow methods similar to those presented by Hu [6]. A thresholding step is first used to perform a coarse segmentation, with the thresholding value often dynamically chosen using the image data. Then connected component analysis is used to identify the two lung volumes. Morphological operations are often used for separating the two lungs and smoothing the final lung mask. Application of these techniques have been shown in several studies to be effective for segmenting the lungs and current systems typically perform at near human levels, achieving mean overlap measures in the high 90s.

Lobe segmentation is a more difficult problem. Although the LOLA11 challenge was designed to help develop lobe segmentation system, only five groups submitted results for that part of the challenge, of which two are semi-automated. The two highest scoring systems, [7] and [8], employ a similar approach, using a watershed algorithm for the lobe segmentation.

It seems that using the image data directly alone provides insufficient information for a successful lobe segmentation. Both [7] and [8] develop watershed cost functions based on an analysis of the Eigen values of the image's Hessian matrix, with the goal of identifying the fissures by looking for sheet-like structures. This information is combined with the locations of blood vessels and airways, which generally do not cross lobe boundaries. [8] also includes information about fissure locations from previously analyzed data as a probabilistic prior to compensate for missing or incomplete fissures.

The current best lobe segmentation results are far below those for lung segmentation, with the best score on the LOLA11 challenge for a fully automated system being 88%. The mean score for the small and difficult to find right middle lobe tends to be much lower (80% for [8]). The mean scores on the LOLA11 challenge for each lobe for these two approaches are very close, and are shown in Table 1.

TABLE I. LOLA11 RESULTS FOR [7] AND [8]

Lobe	[7] Mean	[8] Mean
LUL	0.92	0.91
LLL	0.89	0.88
RUL	0.92	0.93
RML	0.77	0.80
RLL	0.91	0.91
Overall score	0.881	0.884

III. APPROACH

The developed system is implemented in C++ using the Insight Segmentation and Registration Tool Kit (ITK)[1]. The original plan was to use OpenCV[2]. However, OpenCV lacks support for 3 dimensional datasets in many of its operations. ITK does not have the breath of functionality that OpenCV provides, but this is outweighed by the better support for 3 dimensional images such as CT scans.

A. Input

The three datasets used all supply images in MetaImage format. The scans are all 512 x 512 slices with varying Z depths, with a typical scan having between 200 and 600 slices. The voxel data is anisotropic, with the X and Y voxel spacing being the same and the Z spacing usually larger. Typical X/Y voxel spacing is between 0.5 and 1.0 mm, and typical Z spacing is between 0.75 and 2.0 mm. Even among the three datasets used there is not a great deal of variation in the voxel spacing, so no special handling was needed to deal with differences, such as scaling the images to a fixed consistent spacing. The voxel data is signed 16-bit values in Hounsfield units.

B. Lung Segmentation

The lung segmentation process is similar to that described by Hu [6], which is also used in [7] and [8].

The first step is a thresholding to remove any background areas which are outside the scanned volume. Fig. 1. shows an axial slice of a typical scan from the LUNA16 dataset. The dark area surrounding the image has a value of -3024. These pixels need to be reassigned -1000 (the Hounsfield value for air) in order for the adaptive thresholding in the following step to work. This is accomplished by applying a high pass filter with the threshold value and replacement value set to -1000.

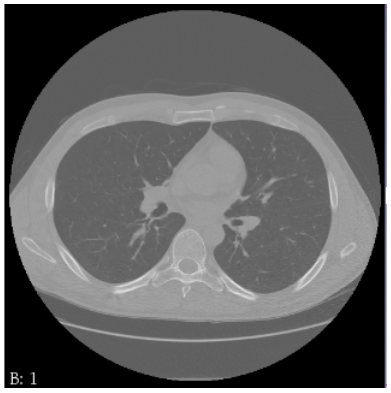

Fig. 1: Axial slice with background

After the background has been removed, a Gaussian smoothing filter with a sigma of 1 mm is used to remove noise from the image.

The next step is uses thresholding to separate the lung volume from the surrounding tissue. The lung tissue has a lower density than the body tissue and so will have lower voxel values. The threshold value is determined using an iterative process. The process attempts to find a threshold value which evenly divides the voxels into two groups, using the mean voxel values above and below the threshold. The rule for updating the threshold is

$$T^{i+1} = (\mu_b + \mu_l)/2$$

where μ_b and μ_l are the mean voxel values of the body tissue and lung voxels (those above and below the current threshold). T^0 is initialized to -500. The process stops when the update is less than 5 units, and usually converges in less than 5 iterations. The thresholding step creates a binary image, as shown in Fig. 2.

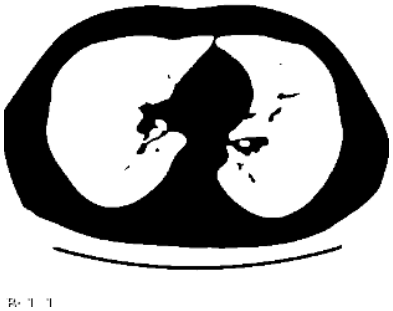
Fig. 2: After thresholding

Connected component analysis is used next to identify the two lung volumes. First, components which contain the corners of the image are removed, as these comprise the volume outside the body. Next, the two largest components (in terms of number of voxels) are analyzed. If both components are larger than 1% of the total volume then both of the components are retained, as these will be the two lung volumes. Often there is

1 https://itk.org
2 https://opencv.org/

only one large component, with the two lungs connected in the 3D space.

In places where the left and right lungs come close together, as is shown in Fig. 3, the lung mask will typically be joined at that location. In most cases the area where the two lung volumes are joined together will span between 5 and 10 slices.

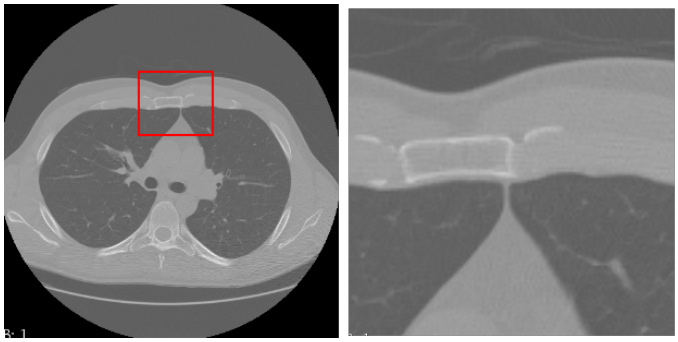

Fig. 3: Area where lungs join together

A dynamic programming approach is used to separate the two lungs. A max cost path algorithm is used to find a path from the center of the bottom to the center of the top of each axial slice. Each slice is processed in turn, with the path from the previous slice used as a starting point. If the path does not cross any pixels included in the lung mask, which is most often the case, then the path is kept and not further processing is needed for that slice.

The first time the previous path does cross pixels included in the lung mask, which will happen in processing the slice for Fig. 3, the max cost path algorithm is run. The gray scale value from the original image is used as the cost, causing the algorithm to find the small, bright gap between the two lungs. After finding the best path, any lung mask pixels on the path are removed from the lung mask.

In processing the next slice, we need to ensure that the path stays contiguous from slice to slice in places where the lung mask as changed. Otherwise the two lungs may still be connected in 3D, even though each slice has been separated. This is accomplished by adjusting the starting values in the memoization array used by the dynamic programming algorithm. For any horizontal line in which a lung mask pixel was changed, the cost values are initialized so that only a pixel value adjacent to the changed pixel will be selected. This ensures that in places where the lung mask needs to be changed to separate the two lungs the cut between the lungs is contiguous in all three dimensions.

The next step is another connected component analysis pass. At this point we may still only have one large component. Even though the lungs have been separated in each slice, they are sometimes still connected in 3D by small airways. In this case a number of erosion steps, using a 3x3x3 ball structuring element, are applied until the lungs are separated. A matching number of dilation steps are then applied to return each lung to approximately its original volume.

At this point the two lungs have been separated. The lung boundaries are rough and some internal holes often exist. Fig. 4 shows a typical axial and coronal view of the lung mask.

Fig. 4: Axial and coronal view before smoothing

A morphological opening operation is used to smooth the lung boundaries and fill in the internal holes. An example of a final lung mask is shown in Fig. 5.

Fig. 5: Final lung mask

C. Lobe Segmentation

An attempt was made to further segment the scans into the five lung lobes. This was ultimately unsuccessful. The steps taken follow those described in [7]:

1. Compute the Eigen values of the Hessian matrix of the image.

2. Compute a fissure-ness score based on the relative magnitudes to the three Eigen values.

3. Segment the blood vessels and airways and compute distance maps for each.

4. Combine the image data, fissure-ness values, and the two distance maps to create a cost image.

5. Run a watershed algorithm to segment each lung into the component lobes using the cost image.

I was unable to construct a cost image that would allow the watershed algorithm to succeed. Using simulated data I was able to confirm that the ITK watershed implementation does work in general. However, the cost images generated from the test scans I used would either find many small regions or a single large region, depending on the watershed parameters.

I believe the approach is sound in general, but experimentally determining the correct mix of parameters seems to be a significant task.

IV. Datasets

For qualitative evaluation, the LUNA16 dataset was used, as this dataset provides ground truth for the segmentation of the lungs (but not lung lobes).

The dataset from the LOLA11 challenge was also used. Although this dataset does not include the ground truth data, the organization does provide an online evaluation of the lobe segmentation results. This evaluation would have proved useful had the lobe segmentation part of the project been more successful. Even without the ground truth segmentation data, the lung segmentation of the scans from this dataset were able to be coarsely evaluated to ensure that both lungs were extracted from the full image.

Additional scans from the Tianchi dataset on the UML server were used as a third source of images. As with the LOLA11 dataset, the ground truth data is not available. A manual test to ensure that both lungs can be identified was used, rather than a more rigorous qualitative analysis.

V. Evaluation

The lung segmentation system was run on the 55 scans from the LOLA11 dataset, 100 scans from the LUNA16 dataset, and 40 scans from the Tianchi dataset.

When the lung segmentation fails, the symptom is that instead of finding two lungs either only one lung is found, one lung and part of the body or the bed is found, or one lung is split into two pieces. These problems can be detected by observing the center axial and coronal slices of the lung mask. This is the test that was done to evaluate the performance for the scans from the LOLA11 and Tianchi datasets. The system was able to segment of these scans and generate lung masks which visually appear correct.

To evaluate the scans from the LUNA16 data, I measured the intersection over union (overlap) between the generated lung mask and the reference segmentation data provided for each LUNA16 scan. In comparing the two masks, a false positive is an voxel included in the generated mask which is not in the LUNA16 segmentation, and a false negative is a voxel not included in the generated mask which is in the LUNA16 segmentation. Table 2 shows the mean, minimum, maximum, and standard deviation for the overlap, specificity, and sensitivity measures for the 100 LUNA16 scan tested.

TABLE II. Overlap, Specificity, and Sensitivity Scores on Scans from the LUNA16 Dataset

	Overlap	Sensitivity	Specificity
Mean	0.95	0.99	0.96
Min	0.67	0.91	0.80
Max	0.98	1.00	0.98
Std Dev	0.04	0.01	0.02

The high overlap score shows that the generated scans match well with the ground truth. The mean overlap is in the same range as other results reported in the literature.

There were a few scan with somewhat low scores, with the minimum being 0.67. For these scans the borders of the lungs matched well but there were holes inside the lung volumes.

Both the specificity and sensitivity numbers are reasonably high, indicating that in most cases the voxels were labeled correctly. The specificity is lower than the sensitivity, which indicates the when an error does occur, it is more likely to be a false positive than a false negative. These false positives tend to be a small number of voxels around the edge of the lung boundary. Matching the reference segmentation exactly would be unreasonable, and having a few extra pixels around the edges is preferred over missing a potentially important part of the lung.

VI. Conclusions

A system was developed using the ITK toolset for lung segmentation using typical image processing techniques. The system was tested using scans from three datasets; LOLA11, LUNA16, and Tianchi. Use of multiple datasets demonstrates the generality of the solution.

The approach is able to find the two lung volumes in all the scans tested. A quantitative evaluation of the system comparing the generated lung masks with the reference segmentations provided for the LUNA16 scans indicates that the system is able to perform at a high level.

Development of the particular steps and parameters required to achieve the eventual level of success required a large amount of trial and error. The general steps for lung segmentation are well defined in the literature. However, the details involved in selecting the exact steps and parameters needed in constructing a working system are non-trivial.

Although the current system works well for the datasets tested, it would not be surprising if it had problems with a new dataset with different voxel spacings, contrast artifacts, or other differences. Any issues could likely be accounted for with additional tweaking of the method or parameters. However, the next dataset used may uncover additional difficulties.

The attempt to extend the system to perform lobe segmentation failed. Although the approach seems reasonable, I was unable to arrive at the correct mix of parameters and/or preprocessing steps in the time given. Further investigation may lead to a solution. However, the difficulty of finding the right mix of empirically derived parameters suggests that a deep learning approach may be a better option.

References

[1] ACS, 2018. American Cancer Society, Cancer Facts and Figures. <http://www.cancer.org/research/cancerfactsstatistics/index>

[2] Setio, A. et al., Validation, comparison, and combination of algorithms for automatic detection of pulmonary nodules in computed tomography images: the LUNA16 challenge. <http://arxiv.org/abs/1612.08012>

[3] LObe and Lung Analysis 2011, <https://lola11.grand-challenge.org/home>

[4] M. Brown, M. McNitt-Gray, N. Mankovich, J. Goldin, J. Hiller, L. Wilson, and D. Aberle, “Method for segmenting chest CT image data

using an anatomical model: preliminary results," IEEE Transactions on Medical Imaging, Vol. 16, No. 6, December 1997

[5] P. Lo, J. Goldin, D. Oria, A. Banola, and M. Brown, "Historic automated lung segmentation method: performance on LOLA11 sata set," September 2011, <http://www.lungworkshop.org/2011/resources/Historic-Automated-Lung-Segmentation-Method_-Performance-on-LOLA11-Data-Set.pdf>

[6] S. Hu, E. Hoffman, and J. Reinhardt, "Automatic lung segmentation for accurate quantitation of volumetric X-Ray CT images," IEEE Transactions On Medical Imaging, Vol. 20, No. 6, June 2001

[7] B. Lassen, E. Rikxoort, M. Schmidt, S. Kerkstra, B. Ginneken, and J. Kuhnigk, "Automatic segmentation of the pulmonary lobes from chest CT scans based on fissures, vessels, and bronchi," IEEE Transactions On Medical Imaging, Vol.32, No. 2, February2013

[8] F. Bragman, J. McCelland, J. Jacob, J. Hurst, and D. Hawkes, "Pulmonary lobe segmentation with probabilistic segmentation of the fissures and a groupwise fissure prior," IEEE Transactions On Medical Imaging, Vol. 36, No. 8, August 2017