Semantic Segmentation of Anatomical Regions of Interest in Chest CT Scan Images

Anushka Hebbar*, Naman Jain*, Dheeraj Reddy*, Mehul Bhandari*
*Department of Computer Science and Engineering, PES University,
Bengaluru, India
{ anushkahebbar@pesu.pes.edu, pes1ug19cs288@pesu.pes.edu,
pes1ug19cs142@pesu.pes.edu, pes2ug19cs909@pesu.pes.edu }

***Abstract*—Chest CT scans are vital for the diagnosis, treatment and research of conditions like lung cancer, heart disease, blood clots and pneumonia. The segmentation of the heart and lungs is a necessary prerequisite step for further diagnostic analysis. Manual, supervised methods for anatomical segmentation require extensive time, labor and effort. Unsupervised, image-driven approaches are required for efficient and accurate results. This paper presents a comprehensive workflow for semantic segmentation of anatomically meaningful structures in chest CT images, more specifically, generation of binary masks for the trachea, lungs, spine and the heart. KMeans clustering, region filling and connected component analysis are the major techniques used for mask generation; mathematical morphological operations are used for cleaning and filtering processes.**

***Index Terms*—chest CT scans, image processing, edge-based segmentation, KMeans clustering, morphological operations, region properties, connected components, region filling**

I. Introduction

Computed tomography, or CT, is a medical imaging modality that creates a three-dimensional image of the internal anatomy of a patient [1]. It uses non-invasive techniques, and generally provides more information compared to traditional X-rays. Over 75 million CT scans are performed each year in the United States [2]. CT scans of the chest are obtained using a rotating X-ray tube emitting X-ray beams through the chest and detecting them through the other side. Multiple cross-sectional views of images, also known as slices, are obtained. These sets of slices are processed by computers to build three-dimensional images of the body. These images are composed of voxels. The intensity of a voxel depends on the density of the tissues between the emitter and the detector. Air pockets have low density and as a result, most of the X-ray beams pass through, and the areas are displayed on the final image as darker shades; bone structures obstruct X-ray beams and the resultant areas of the image are relatively lighter.

CT scan tests are crucial in diagnosing, treating and researching pulmonary and cardiovascular diseases and cancer. Chest CT scans are performed to check for blockages, injuries and tumors in chest anatomy [3]. Identifying anatomical structures in CT scan images is necessary in all further processing and diagnostic model building. Furthermore, the segmentation and detection of the entire heart in CT scan images of the chest is an important procedure that acts as prerequisite to the training of models that detect the presence of heart disease or other cardiovascular functional abnormalities. Regions of interest or ROIs, are anatomically meaningful areas of partitioning in a chest CT image. These are to be extracted from CT scans for further analysis. Manual approaches for anatomical ROI labeling require extremely long periods of time. Computed tomography is performed by radiologists or radiographers. Supervised methods of segmentation require the skills of these experts for accurate segmentation. Such approaches require extensive time, effort, labor and expert supervision, and therefore, automated methodologies that are image-driven, unsupervised, and yet, accurate and efficient, are required.

In this paper, we apply an unsupervised, automated workflow of segmenting anatomical regions of interest, more specifically, the lungs, spine, trachea or bronchi and the heart, from chest CT scans. We also elucidate a more robust segmentation algorithm for our data, as compared to current methods in this domain.

We assume that significant lung or heart cancer masses are not present in our data. High-density pathologies would obstruct our methodology of segmentation and these are handled by more suitable approaches as in [4].

II. Related Works

In recent works, numerous papers address this problem domain [5]. They employ a wide range of techniques including - image-driven algorithms, mean and multiple thresholding, and deep learning methods. We examine two that are most relavent to our objective.

Rim et al. [6] propose a method of cardiac segmentation whose goal is to partition the whole chest CT image into cardiac anatomical regions of interest, using the dissimilarity of each pixel's value, and meaningful structure or geometric position of cardiac anatomical ROIs. The threshold criteria used is based on K-means clustering, which automatically clusters pixels. Since the anatomical structure of cardiac tissues and the quantitative scale of CT images (hounsfield units or HU) are complicated, directly applying K-Means is hard, and as a result, mathematical morphology methods are used as threshold shifting enhancers. This methodology is extensive as masks for many anatomical parts are created.

Larrey-Ruiz et. al. [7] propose an image-driven method for the accurate segmentation of heart from CT scans. The

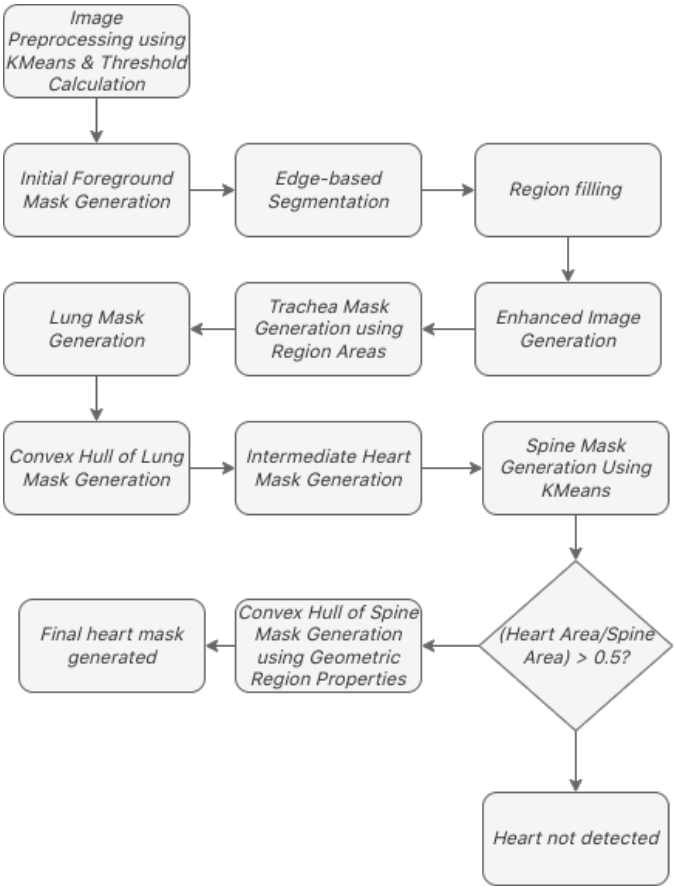


Fig. 1. Flowchart of Method Workflow

methodology proposed is automatic and significantly faster relative to its predecessors. The idea is to employ statistical local and global parameters of an image to identify ROIs and use the prior knowledge about cardiac structures for morphological operations and filtering techniques to segment the cardiac regions.

Although in recent years numerous deep learning based strategies have gained popularity, not all are suitable for routine clinical practice - they are either computationally intensive or require an extensive amount of data. The algorithmic approach is much more computationally efficient, and performs on par with neural network architectures.

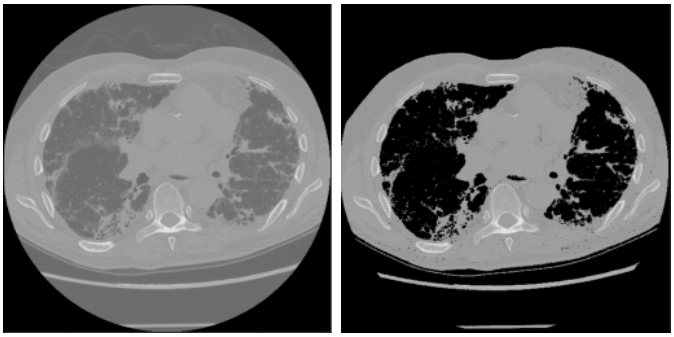
In this paper, the proposed method builds on top of the workflow presented by Rim et al. to provide more effective and accurate results on the data we work with.

III. Proposed Methodology

The workflow that is presented in this paper is depicted in Fig. 1.

A. About the Data

Data is collected from a Kaggle repository [8] that consists of 16,708 images, each of size 512x512 pixels. Each patient is associated with an ID and a volume, or a set of CT scan slices. All images are in the RGB format.



(a) Original CT Image Slice (b) Filtered Image

Fig. 2. Before and After Preprocessing for Filtering ROIs

B. KMeans Clustering for CT Scan Image Preprocessing

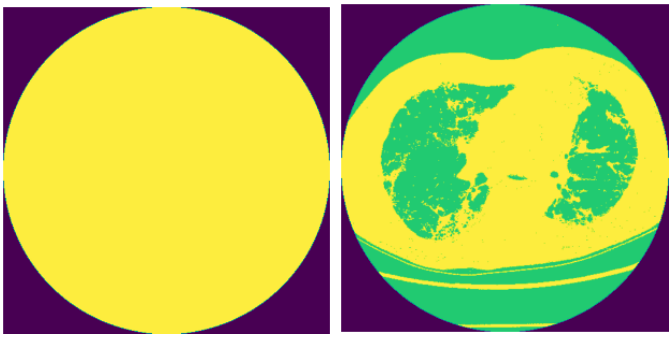
Every slice image from the set of CT scans corresponding to any patient is structured as a disk, framed by a black background as shown in Fig. 2a. For any further processing, the regions of interest, or the anatomical structures in the chest that constitute the foreground, have to be distinguished from the rest of the image. In order to acquire only pixels that make up body tissue and bone, both air pockets (represented by pixels of relatively darker intensity) and the black frame are treated as the background for all further processing; pixels constituting these are assigned a uniform intensity of 0. All other anatomical structures are filtered to arrive at a segmented image consisting of body tissues, bone and other foreground structures only, and the model workflow utilizes this filtered image as input. This process is depicted in Fig. 2.

Filtering the ROIs is achieved by transforming intensity values lower than a certain threshold to zero, and retaining values greater than the threshold to remain the same. The thresholding process [9] follows (1), where $dst(x, y)$ is a pixel intensity on the resultant image after thresholding, and $src(x, y)$ is the intensity of a pixel at (x, y) in the input image.

$$dst(x,y) = \begin{cases} src(x,y) & \text{if } src(x,y) > \text{threshold} \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

This truncated thresholding to zero technique uses a threshold value calculated as the average of two cluster center values derived by applying KMeans clustering with K = 3, on the original image. The two cluster centers mark the clusters of pixels that depict the air pockets and foreground material. By calculating the average of the centers, the resulting value is able to distinguish between intensity values that belong to the background and air, and those of foreground structures. In Fig 4, the three clusters and their center values are depicted in different colors, and the calculated threshold is used to demarcate the boundary between the foreground and background.

To calculate the threshold, the average of the two greater cluster centers are used instead of KMeans with K = 2, since the latter approach reduces the demarcation of the image to the background frame and the spotlight disc as seen in Fig.3, which is not our objective.



(a) KMeans with K = 2 (b) KMeans with K = 3

Fig. 3. KMeans Clustering on the Original Image

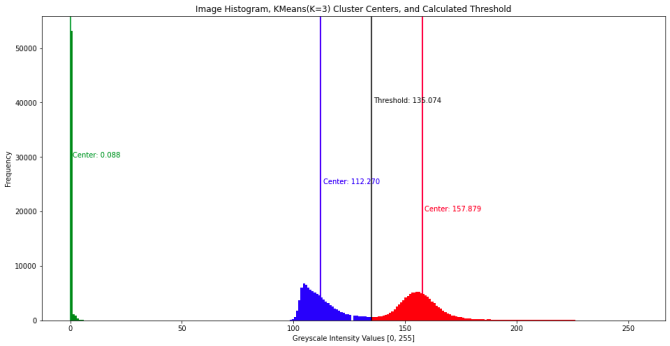


Fig. 4. Image Histogram with Center Clusters and Calculated Threshold

C. Edge-based Segmentation for Foreground Binary Mask Generation

To drill down to individual anatomical structures in the preprocessed, filtered image, it is vital to remove all extraneous curves that lie outside the boundary of the patient's body. Therefore, a foreground mask is obtained; in this section, the foreground constitutes only the structures bounded by the chest.

The objective in this phase of the workflow is to remove extraneous objects, and so a three-part procedure is followed:

1) Initial Foreground Binary Mask: The threshold calculated to discriminate between foreground structures and the background in the previous section is used to create an initial binary mask of foreground structures using the following equation:

$$dst(x,y) = \begin{cases} 1 & \text{if } src(x,y) > \text{threshold} \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

Then, two morphological operations, one with square and the other with disk structuring elements are carried out to remove extraneous curves and elements that lie outside the chest.

2) Edge-based Segmentation: Edge-based segmentation is used to find contours that delineate the outer boundary of the chest structure. The Canny edge detector is used to distinguish all edges. Since the background is uniformly black, all edges are either at the boundary of the chest, are any extraneous curves or are found inside the chest.

3) Region Filling: The region filling operation is used to segment the inside of the chest from the outside. Since the chest is a closed structure in almost every slice of a patient, this operation is bound to segment only the inner part of the chest from the outer.

The output of this procedure resembles that of Fig. **??**b. In the next section of the workflow, this binary mask operates on the original image to retain only structures inside the chest.

D. Lung and Trachea / Primary Bronchi Masks Generation using Connected Components and Region Properties

The foreground mask generated in the previous section is applied onto the original image to create a segmented image. Consequently, chest structures are extracted and an enhanced image is created as in Fig. 7b. This enhanced image distinguishes the black background pixels from the black pixels of the lungs. This is accomplished by setting pixels in the enhanced image to white if the corresponding pixel in the foreground mask is black (since these represent the background).

Inverse thresholding following (3) uses the previously calculated threshold from KMeans to segment the lungs from the enhanced image. The result of this is shown in Fig. 7c.

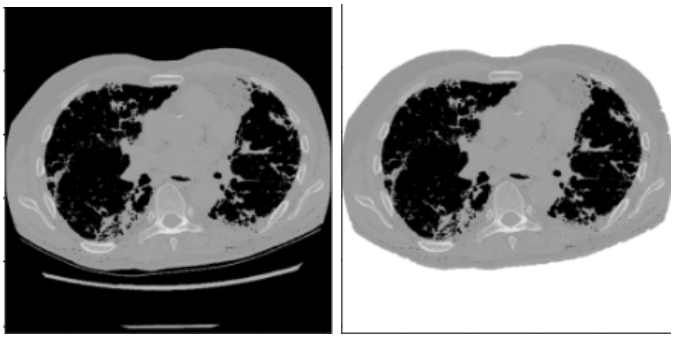
$$dst(x,y) = \begin{cases} 0 & \text{if } src(x,y) > \text{threshold} \\ 1 & \text{otherwise} \end{cases} \tag{3}$$

A sequence of mathematical morphology operations: (1) opening using a square structuring element, (2) closing with a disk structuring element, and (3) the region filling operation are used to clean the segmented image, and fill in holes caused by capillaries. This results in images similar to Fig. 7d.

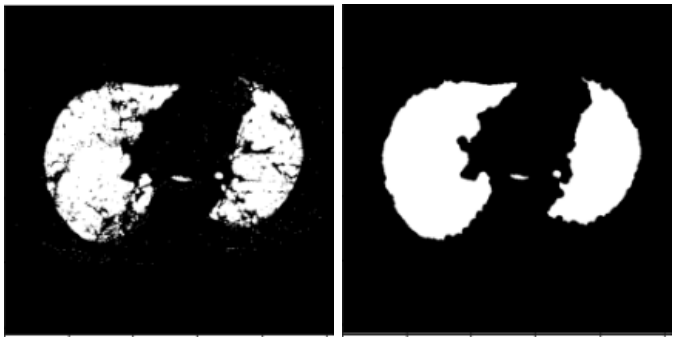
To segment out wind pipe structures, all connected components are extracted and the region properties are computed. Upon empirical results, regions with areas lesser than 1500 pixels are set to be wind pipe structures, either the trachea or bronchi. The final trachea mask in Fig. 7f is created with the summation of all these regions, and the rest of the regions make up the final lung mask of Fig. 7e.

E. Convex Hull of Lung Mask and Intermediate Heart Mask Generation

For purposes of heart segmentation, a convex hull of the lung mask is computed. The inverse image of the initial lung mask is applied on the convex hull to segment regions of the heart. This makes sure regions of the trachea also are segmented out of the heart regions. Consequently, morphological opening operations: (1) opening with a square structuring element and (2) opening with a disk structuring element, are applied in order to remove unnecessary pixels. This results in an intermediate heart mask. The procedure is depicted in Fig. 6.



(a) Original Preprocessed Image (b) Enhanced Image



(c) Initial Lung Mask (d) After Morphological Operations

(e) Final Lung Mask (f) Final Trachea Mask

Fig. 5. Lung and Trachea Mask Generation

(a) Convex Hull of Lung Mask (b) Convex Hull on Inverted of Lung Mask (c) Intermediate Heart Mask

Fig. 6. Intermediate Heart Mask Generation

F. Spine and Heart Masks Generation using Convex Hull and Geometric Region Properties

The intermediate heart mask is applied on the original image to segment out portions of the heart and the spine. Applying KMeans with K = 3, clusters pixels belonging to the background, to the spine and to body muscles or tissue. Since heart tissue lies in front of the spine, the following procedure is used to segment heart tissue from the cluster segmented by KMeans:

- Segment the spine pixels using the KMeans cluster.
- Create a convex hull of the spine and perform morphological operations to smoothen the image.
- Find the region properties for this convex hull - specifically, the centroid, orientation and major axis length.
- Using these geometric properties, obtain the uppermost point and the centroid point of the convex hull, using a bit more than half of the major axis length, as in (4). This is done since the centroid does not lie exactly on the midpoint of the major axis.

$$\begin{aligned} x_1 &= x_0 - 0.6 * sin(orientation) * majorAxisLen \\ y_1 &= y_0 - 0.6 * cos(orientation) * majorAxisLen \\ &\text{where, } (x_0, y_0) \text{ are centroid coordinates} \\ &(x_1, y_1) \text{ are uppermost point coordinates.} \end{aligned} \tag{4}$$

- Set all pixels to the left of the uppermost point to white and all pixels to the right of the centroid to black. This demarcates the heart region and the spine region.
- If the fraction of the area of the intermediate heart mask compared to the spine mask is less than one-half, we set the final heart mask to a blank image, depicting that no heart is detecting in that slice image. Otherwise, we perform the next step.
- Apply this mask on the intermediate heart image to arrive at the final heart mask, only if the ratio of the area of heart to spine masks is greater than one-half.

The procedure is depicted in Fig. 7.

IV. Experimental Results

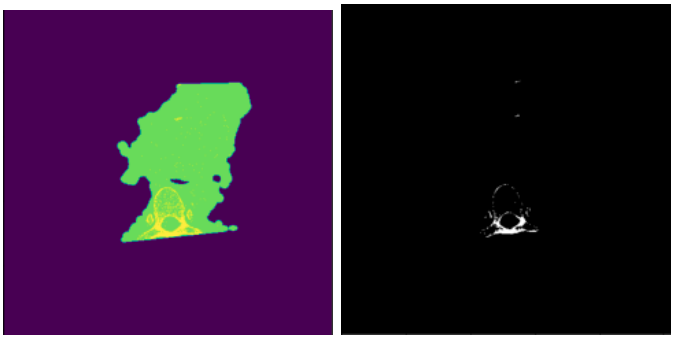
We implement both the workflow implemented by Rim et al. and our workflow to compare and contrast the effectiveness among different kinds of slices.

A. Initial Slices: Prevalent Trachea and Small Lungs

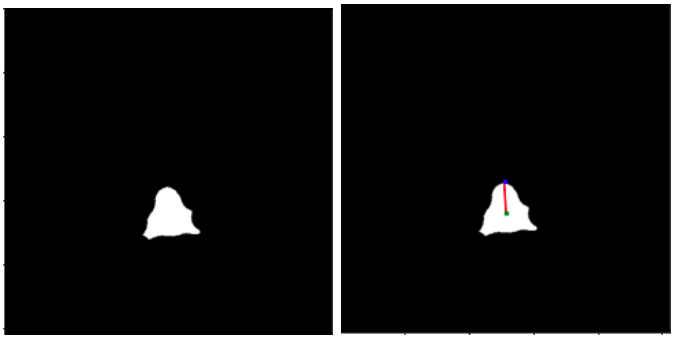
In the initial slices of CT scans, the trachea is more prevalent and small portions of the lungs are seen. Using the method presented by Rim et al., the trachea is labeled as part of the lungs during the convex hull of lung mask creation. This is because the trachea is constituted by dark pixels in the original image. As a result, the convex hull becomes triangular and the heart is inaccurately segmented as seen in Fig. 8.

Our method segments the trachea as a separate entity using the area of the connected components. This ensures the lung mask and the convex hull of the lung mask are accurate, which leads to accurate heart segmentation.

(a) Intermediate Heart Mask (b) Intermediate Heart Segment



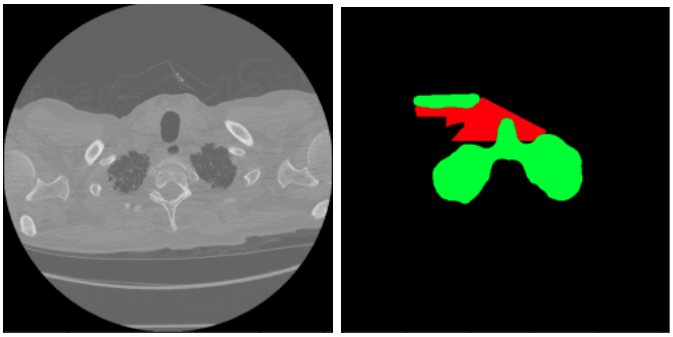
(c) On Applying KMeans (K=3) (d) Spine Mask Binary Image



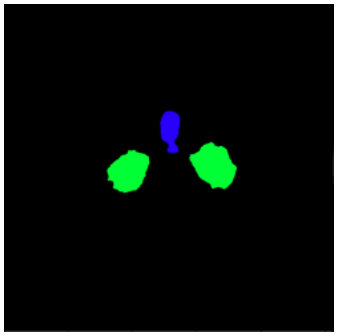
(e) Convex Hull of Spine Mask (f) Geometric Measures of Centroid and half Major Axis

(g) Mask for Heart Tissue (h) Final Heart Mask

Fig. 7. Spine and Heart Mask Generation



(a) Initial CT Scan Image (b) Original Workflow Segmentation



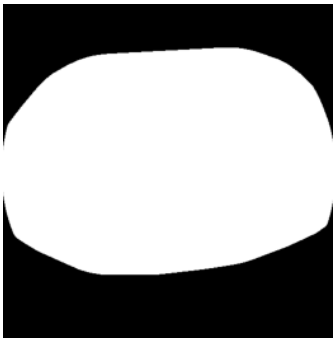
(c) Improved Workflow Segmentation

Fig. 8. Difference in Segmentation Workflows. Red: Heart, Blue: Trachea, Green: Lungs

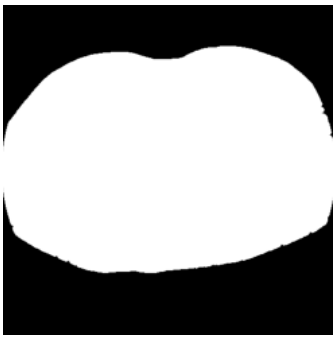
B. Middle Slices: Prevalent Heart and Lungs

In these slices, both the heart and lungs are prevalent. In the case of lung segmentation, we arrived at much more effective results that retained more information. Rim et al. used the method of KMeans clustering to arrive at a cluster center averaged threshold to distinguish between the foreground and the background. Consequently, two iterations of morphological opening operations with the standard structuring element of [10] were performed on the thresholded binary image to arrive at a cleaned foreground mask for body muscle and bone. The convex hull of the foreground mask mask was then calculated in order to distinguish the lung (which is represented by black pixels bounded by the chest structure) from background black pixels.

On applying this methodology on our data, we notice that significant area of the background remain along with the lungs as a result of more pronounced upper curvatures in the body structures of the data we work with as seen in Fig. 9. This calls for much more aggressive morphological operations, which loses the sanctity of information contained in the actual regions of interest. Since our data show imperfect results with this method, we finalize a better process of lung segmentation that is much more effective for our data, upon experimentation with the choice of structuring element and number of opening operations.

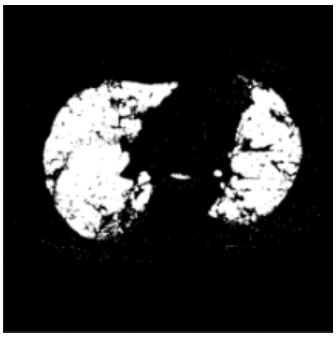


(a) Convex Hull of the Foreground



(b) Foreground Mask using Region Filling

(c) Applying convex hull to segment lungs



(d) Applying region filled mask to segment lungs

Fig. 9. Effectiveness of Lung Segmentation using a Foreground Mask created with Region Filling compared to a Convex Hull of Foreground Mask

C. Final Slices: No prevalent structures

In the final set of slices of a patient's CT scans, the proposed method performs better at appropriately distinguishing between heart and non-heart tissue. However, both models take a hit when compared to performances seen in the earlier slices. In these slices, other anatomical structures like the liver take center stage, and the models fail to detect this. However, the proposed method is more robust when compared to the other workflow, and stops falsely classifying the heart after just one slice. An example of this is presented in Fig. 10.

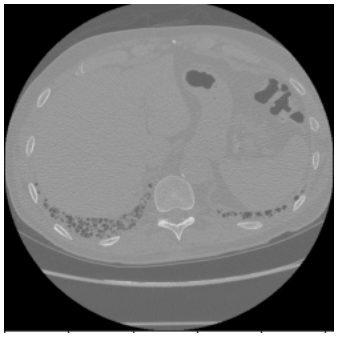
D. Future Steps and Improvements

In the initial slices of the image, the presented workflow labels the trachea as lungs for a couple of slices, when the trachea is extremely close to the lungs. This occurs due to the morphological closing operation connecting the trachea and the lungs during lung mask generation. This can be solved using a template of the trachea and performing template matching. An example of this is presented in Fig. 11.

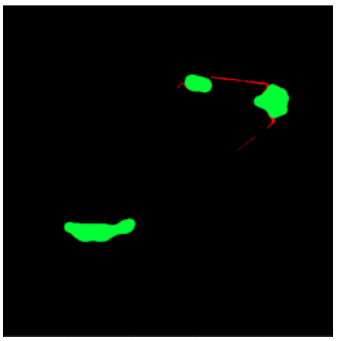
V. Conclusion

Contribution of Team Members

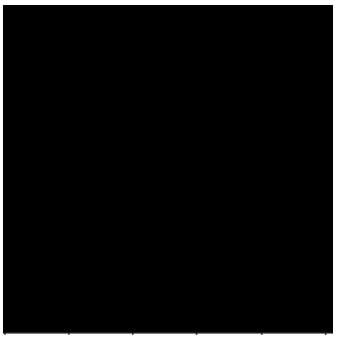
- Anushka Hebbar : Workflow building+implementation, Rim et al. implementation, Paper Writing, Report Writing, Video making



(a) Final Slice Sample

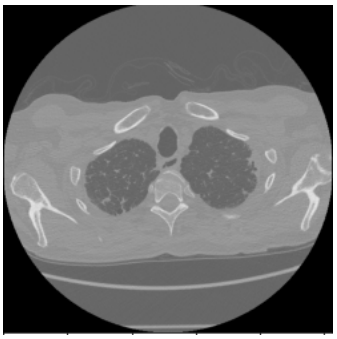


(b) Original Workflow's Performance



(c) Presented Workflow's Performance

Fig. 10. Performance of Workflows on Final Slices. Red: Heart, Blue: Trachea, Green: Lungs



(a) Initial Slice Sample

(b) Presented Workflow's Performance

Fig. 11. Classifying the Trachea as Lungs. Red: Heart, Blue: Trachea, Green: Lungs

- Naman Jain : Larrey-Ruiz implementation, Video making, Report Writing
- Dheeraj Reddy : Larrey-Ruiz implementation, Video making, Report Writing
- Mehul Bhandari : Larrey-Ruiz implementation, Video making, Report Writing

References

[1] Wikipedia.org, "CT Scan", *Wikipedia*. Accessed: Apr. 29 2022. [Online]. Available: https://en.wikipedia.org/wiki/CT_scan

[2] Idataresearch.com, "Over 75 Million CT Scans Are Performed Each Year and Growing Despite Radiation Concerns", *iData Research*. Accessed: Apr. 29,2022. [Online]. Available:

https://idataresearch.com/over-75-million-ct-scans-are-performed-each-year-and-growing-despite-radiation-concerns

[3] Hopkinsmedicine.org, "Computed Tomography (CT) Scan of the Chest", *John Hopkins Medicine*. Accessed: Apr. 29 2022. [Online]. Available: https://www.hopkinsmedicine.org/health/treatment-tests-and-therapies/ct-scan-of-the-chest

[4] Sun S, Bauer C, Beichel R. Automated 3-D segmentation of lungs with lung cancer in CT data using a novel robust active shape model approach. IEEE Trans Med Imaging. 2012;31(2):449-460. doi:10.1109/TMI.2011.2171357.

[5] Kang et. al. , "Heart chambers and whole heart segmentation techniques: review", 2012.

[6] Rim, B.; Lee, S.; Lee, A.; Gil, H.-W.; Hong, M. Semantic Cardiac Segmentation in Chest CT Images Using K-Means Clustering and the Mathematical Morphology Method. Sensors 2021, 21, 2675. https://doi.org/10.3390/s21082675.

[7] Larrey-Ruiz et al.: Automatic image-based segmentation of the heart from CT scans. EURASIP Journal on Image and Video Processing 2014 2014:52.

[8] Kaggle.com, "Dataset - Chest CT Segmentation", *Marco Polo*. Accessed: April 29, 2022. [Online]. Available: https://www.kaggle.com/datasets/polomarco/chest-ct-segmentation

[9] Docs.opencv.org, "Miscellaneous Image Transformations", *OpenCV*. Accessed: Apr. 29,2022. [Online]. Available: https://docs.opencv.org/3.4/d7/d1b/group__imgproc__misc.html

[10] Boulogne, F.; Warner, J.D.; Neil Yager, E. Scikit-image: Image processing in Python. J. PeerJ 2014, 2, 453.