



Volume Estimation of Pulmonary Lesion Using Chest CT Sequence

Ramyasri Nayak¹, Nandish S^{*2}, Prakashini Koteshwara³

¹Department of Biomedical Engineering, Manipal Institute of Technology, MAHE, Manipal, India

² School of Information Sciences, MAHE, Manipal, India

³ Department of Radiodiagnosis, Kasturba Medical College, MAHE, Manipal, India

*Corresponding author E-mail: nandish.s@manipal.edu

Abstract

Many of the imaging modalities like X-ray, Computed Tomography (CT), Magnetic Resonance Imaging (MRI), fMRI have emerged to capture high quality images of anatomical structures of the human body. Radiologist can also have a better visualization if the regions of interest in the images are extracted and visualized 3D. To extract region of interest, sometimes preprocessing steps are performed on the input data. Pulmonary lesion is a small round or oval-shaped growth in the lung. It consists of solid and non-solid portion. The estimation of solid and non-solid portion of the pulmonary nodules will help the clinicians in the diagnosis and to suggest the appropriate treatment methodology. Lesion volume estimation gives a brief idea about the area occupied by the lesion tissues, which in turn can help the radiologist to plan treatment accordingly. In proposed work, lesion is segmented using K-means algorithm and then volume of the lesion is estimated. The slices which have segmented lesion with solid and non-solid regions is used for 3D visualization. The results obtained using the proposed methodology is validated with the Slicer 3D software. Error in the estimated volume of the solid and non-solid portion of the lesion was found to be in the range of 1.11% - 3.30% and 0.1% to 4.55% respectively. Results from the proposed methodology, lesion extraction with solid and non-solid, 3D visualization of the same and volume estimation respectively are validated by taking feedback from the radiologists and segmented lesion slices are used to estimate the volume and 3D visualization in Slicer 3D software for validation.

Keywords: Threshold, K-means, Lesions, Volume, Slicer 3D.

1. Introduction

Currently, in the area of medical domain, image segmentation has become a useful tool to extract the region of interest. There are many methods used for segmentation of medical images and further the segmented images can be analyzed for precise diagnosis. Usually, an array of numbers represents the medical image in its raw form, which are the intensity values of the respective physical quantity which in turn exhibits the contrast difference among the different body parts in the image. To identify the boundaries of the abnormal structure present in an image, the fundamental requirement is image segmentation. Features like shape, volume etc. can be effectively obtained from the segmented image. These results can also be further used for classification and diagnosis of different lesions [1]. Lesion volume estimation gives a brief idea about the area occupied by the lesion tissues, which in turn can help the radiologist to plan treatment accordingly. In this work, volume estimation of the solid and non-solid portions, as well as that of the entire lesion is performed. Volume estimation of the pulmonary sub-nodules provides information as to whether the tissues are malignant. This is one of the important feature, which allows doctors to decide whether the surgical resection is required or not. Volume of the entire lesion is calculated by the area covered by the lesion in all the three planes: axial, coronal and sagittal. Pulmonary nodule (lesion) is a small round or oval-shaped growth in the lung. It consists of solid and non-solid portion. The estimation of solid and non-solid portion of the pulmonary nodules will help

the clinicians in the diagnosis and to suggest the appropriate treatment methodology. The classification also provides the diagnosis of benign versus malignant lesion and aids in targeting the solid component of lesion for diagnostic biopsy/FNAC. The primary purpose of this study is to identify an algorithm to extract the pulmonary nodules into sub-nodules and to deliberate technology advancement in medical imaging [1, 2].

2. Literature Review

Mahmood et al., proposed a method which started with preprocessing of the lung (CT) image to enhance its contrast and to make them ready for image segmentation by implementing K-means classification. The area of the lesion in the CT slices were estimated and used to determine the volume by overlapping the slices one over the other [3]. T.S. Sakthi et al., 2004 used Otsu's thresholding method to segment the lung cancer nodules from CT images. Preprocessing was performed using median filters for denoising the images. Otsu method was applied on gray scale images for thresholding and for feature extraction, gray-level co-occurrence matrix (GLCM) method was used [4]. Kaur et al., implemented computer aided diagnosis tool for detection of lung cancers. Primarily, they extracted the lung image from the background based on thresholding method. Segmentation of lesion was done using region growing algorithm which eliminates the area of pixels which are not bearing the similar intensity as that of the initial seed pixel [5]. Kothawari et al., implemented various im-



age segmentation techniques to extract lung lesions. They used threshold, watershed and clustering techniques to extract lesions present in lung for gray scale images. The algorithm that was used in this work also had few drawbacks such as; the algorithm needs initial markers and seed points to be drawn carefully around the region of interest making the method semi-automatic [6].

Praseeja et al., used a linear Vector Quantization neural network (LVQNN) with a preprocessing procedure with cellular automata based seeded segmentation method for segmentation and volume estimation of thyroid gland using CT images. The proposed method extracted the thyroid region effectively and the estimated volume was compared with the volume estimated from the existing invasive methods like biopsy/FNAC [7]. Jose et al., performed K-means and Fuzzy C means clustering techniques to segment out the lesions in brain region. In the paper, the techniques were performed on gray scale noisy images which were preprocessed and directed for segmentation process to extract region of interest. They also extracted the features and estimated the size of brain lesion in MR images [8]. Morey et al., studied the manual segmentation process for extracting the region of interest. The manual process has been utilized in many medical segmentation software systems, which helped medical practitioner with the knowledge of anatomical structures to extract the region of interest by tracing the boundaries of the lesion. The main drawback of this process is that the user has to view axial, coronal and sagittal planes for drawing/filling the region of interest, which was although efficient but time consuming [9]. Kostis et al. developed 3D segmentation techniques like morphological processing, connected component analysis to segment the pulmonary nodule's volume using CT. The results were displayed in three dimensions. The segmented 3D nodules were then used for classification of benign and malignant lesions based on size of the nodules extracted [10]. P. Daniel et al., proposed histogram based threshold for the segmentation process. In this paper, the histogram of the image was plotted, based on which the threshold was selected for the segmentation process. Results obtained showed better segmentation and the attractive side of the proposed approach is the ease of implementation. The main advantage of this method was that it was robust and automatic [11].

3. Methodology

Proposed research follows methodology shown in figure 1.

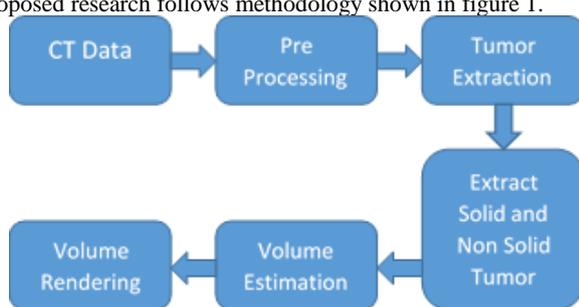


Figure 1: Proposed Methodology

In proposed methodology, plain CT sequence is used. Any noise present in the image was removed by pre-processing the image (the images were subjected to Median filtering in order to remove the noise as well as to preserve the boundary details). The pre-processed images were then subjected to extract the lesion. Proposed methodology used three segmentation techniques; initially 'thresholding' method, manual segmentation and then K-means segmentation technique to extract lesion.

Thresholding: It was observed that normally the lesion had higher intensity compared to its background; hence thresholding was initially tried for segmentation.

- DICOM sequence is read and the color bar obtained based on the pixel intensity distribution over the image was examined, and the threshold value was chosen accordingly.

- With respect to the threshold value, the pixels in the lesion region were assigned a value of '1', and the surrounding pixels were assigned a value of '0'.
- The pixel intensities masked with a value of '1' were extracted to represent the lesion.

However, in some cases where the lesion was attached to the organ, segmentation by this method did not provide good results. Thus manual segmentation method was implemented.

Manual segmentation: Image segmentation can also be performed manually to extract the lesion. The algorithm is explained as follows:

- DICOM CT sequence is read and the region of interest in the image was manually drawn.
- Then, the manually selected region was extracted and the process was repeated for each slice containing lesion.

Since the process is tedious and time consuming, the K-means clustering method is chosen for extracting lesion.

One of the popular methods used for segmentation is the K-means clustering. During clustering, a set of data is divided into specific number of groups. K numbers of disjoint clusters are formed from the given dataset. There are two separate phases for K-means algorithm. It calculates the k centroid in the first phase, and each point which has smallest distance to the centroid is clustered together in the second phase [24]. There are different methods to define the nearest centroid distance, out of which Euclidian distance is the one commonly used.

Once grouping is done, a new centroid is recalculated. Euclidean distance between the new centroid and each data point is calculated, and the data points having minimum Euclidean distance are clustered together. Each cluster is defined by its member objects and by its centroid [24].

The algorithm of K-means clustering is as follows:

- The two dimensional dicom image $I(x,y)$ was used as input in this algorithm.
- The number of clusters 'k' was initialized. Let $p(x, y)$ be an input pixels to be clustered and C_k be the cluster center.
- For each pixel of an image, the Euclidean distance d is calculated using the equation below,

$$d = \|p(x, y) - C_k\| \quad (1)$$

- Based on the distance d , all the pixels having minimum Euclidean distance are clustered together.
- After assigning a pixel to a particular cluster, the new centroid was calculated using the equation below, and the process is repeated until the cluster labels of the image do not change anymore

$$C_k = \frac{\sum_{y \in C_k} \sum_{x \in C_k} p(x, y)}{k} \quad (2)$$

- After K-means clustering, the cluster containing the lesion was labeled using `bwlabel` function and extracted out by selecting the index value of the lesion region.

If the file imported to MATLAB is in .jpg format, there are chances of losing the required medical information. However to retain the original information, .jpg image can be reimported and can be superimposed on the dicom file to preserve the original information. However, it is not possible to assign proper spacing between the slices while saving the dicom files as stack of images. In order to overcome this, the entire subject database was imported in analyze format (.img), which helped to save the slices containing lesions with proper pixel spacing in three planes (axial, coronal and sagittal).

Two groups of color intensities were observed in the segmented images. Brighter intensities were observed for the pixels located in the center of the lesion (solid contents of the lesion) and the peripheral pixels (non-solid contents) were of lesser intensities.

Solid and Non solid lesion is extracted by creating a histogram of the image and use the mean value of intensities as the threshold. The histogram clearly shows that there are some differences in the intensity of the pixel belonging to the background and the object. The choice of the threshold value is the key parameter in the thresholding process. Users can choose a threshold value manually. Although there are different methods to choose the threshold automatically, selecting the threshold based on the mean and median values are the simplest methods used. The rationale being that, if the objects are brighter than the background, they should also be brighter than the average.

Histogram based thresholding method is explained as follows:

- Initially, the segmented lesion images are imported to MATLAB.
- Histogram of each image can be constructed using `imhist` command.
- Then, the threshold(th) is calculated based on the histogram plot of each image using the formula given below

$$th = \frac{(x + y)}{2} \quad (3)$$

Where, th-threshold of each image

x- Minimum intensity value of the lesion image

y- Maximum intensity value of the lesion image

- Thereby, the final threshold (th) is calculated by taking the mean value of the threshold of each image obtained earlier.
- The intensities that are falling less that the threshold value (th) are grouped as non-solid portion and the remaining intensities are grouped as solid portion of the lesion.

Volume estimation in proposed research is found by the segmented lesion and segmented pulmonary sub-nodules are subjected to volume measurement. In MATLAB, the area of each slice containing the lesion and sub-nodules are calculated using `bwarea` function.

The pixel spacing between the axial, sagittal, coronal slice planes was 0.69mm, 0.69mm, 5mm respectively. The volume in the voxel spaces is then calculated by taking the product of the area into the spacing between the different planes as given in equation 4.

$$volume = \sum_{i=1}^N area(i) * A * S * C \quad (4)$$

Where, area (i)-count of pixels in the region of lesion (in all the slices), A, S, C- Space between axial, sagittal and coronal slices (in millimeters) respectively. The area of the extracted lesion can be computed by multiplying the number of pixels present in the lesion region by the size of each pixel. Thereby using the above formula volume, of the lesion and pulmonary sub-nodules are estimated the segmented slices were saved as stack of images with pixel spacing of [0.69 0.69 5] mm in axial, sagittal and coronal planes respectively in analyze format. These stacks of images were imported in slicer 3D software and the volumes of the lesion are rendered with the help of model maker module. The volume is estimated using label statistics module in quantification tool of slicer 3D software. These volumes obtained using the developed method was then compared with the results obtained using the Slicer3D software and the errors were tabulated.

4. Results

The CT images were obtained from Radiology department is shown in the figure 2. The pre-processed CT images were subjected to segmentation process using thresholding method. Although thresholding method was able to extract the lesion in cases where the lesion was not attached to the organ (as shown in the fig 2), it was not able to provide significant results especially when the lesion was attached to the organ.

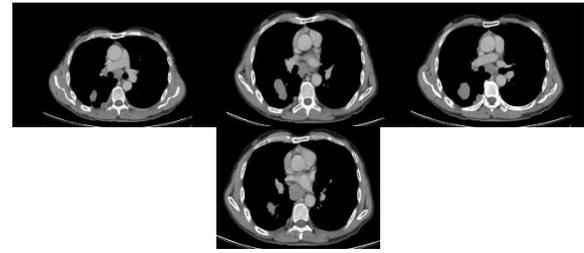


Figure 2: CT images of lungs showing the presence of lesion

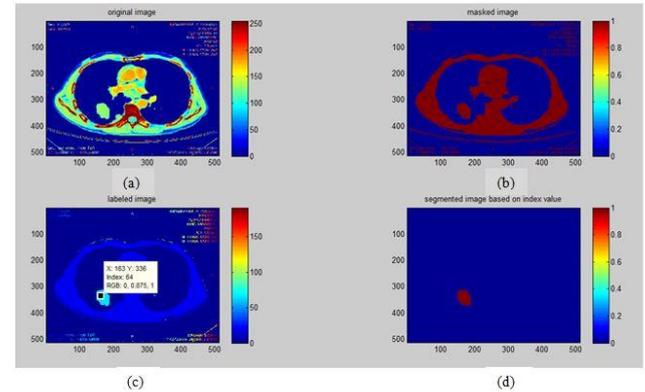


Figure 3: Threshold based image segmentation (a) Original image, (b) Masked image, (c) Labeled image, (d) segmented image based on index value.

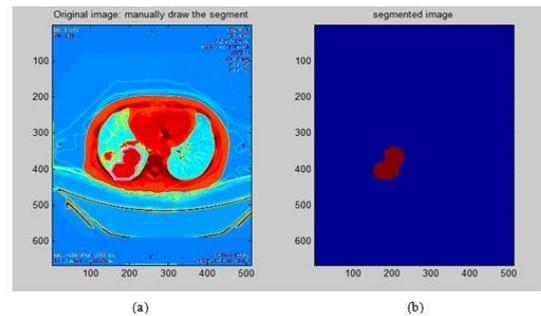


Figure 4: Manual segmentation (a) original image: manually draw to segment, (b) Segmented image

To overcome this problem, manual segmentation was opted which would segment out the lesions effectively. But, this manual process is a very tedious work and time consuming, which includes drawing the image patterns of lesions for each and every CT slices (as shown in Figure 4). To overcome the drawbacks of manual segmentation process, a semi-automatic segmentation method using K-means clustering was implemented. The clustered images were assigned to the labels which were then extracted accordingly using index values to extract only the clusters containing lesions as shown in Figure 5.

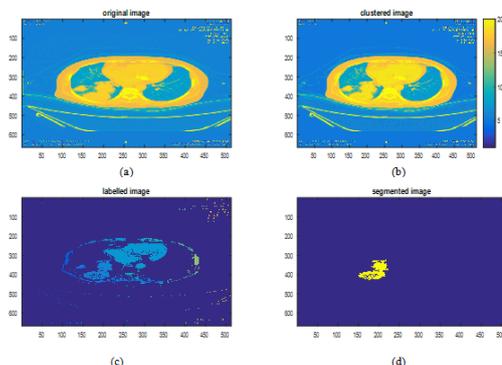


Figure 5: K-means image segmentation (K=20, idx1=19, idx2==208) (a) original image, (b) clustered image, (c) labeled image, (d) segmented image

Once, the lesions are segmented from its background, the segmented images are subjected to histogram thresholding to obtain the solid and non-solid portion of the lesion. The segmented portion of the lesion and the corresponding histogram is shown in the Figure 6. For a particular data set, the final threshold T was obtained as 60. Pixels having intensities above 60 were marked as the solid portion of the lesion and the remaining pixels in the lesion were considered to be non-solid portion. The histogram of segmented solid and non-solid portion of the lesion is as shown in the Figure 7.

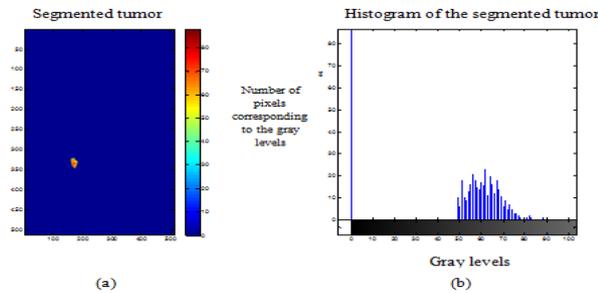


Figure 6: (a) segmented lesion, (b) histogram of the corresponding lesion

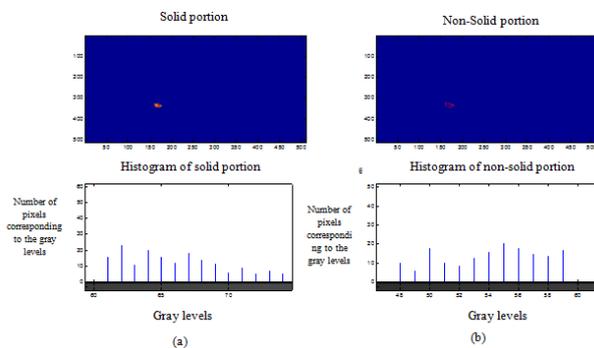


Figure 7: The histogram of (a) solid portion, (b) non-solid portion of the lesion

The entire lesion, as well as solid and non-solid portion of the lesion is subjected to volume estimation. The obtained results are tabulated (as shown in Table 1). The volume data provides useful information to the clinicians for diagnosis (higher percentage of solid portion in the lesion is indicative of a malignant lesion).

Table 1: volume of entire lesion and solid and non-solid portion of the lesion

Data sets	Volume of extracted lesion(mm ³)	Volume of the solid portion (mm ³)	Volume of the non-solid portion (mm ³)
1	13180.82	6160.73	7020.09
2	32149.69	13428.81	18720.87
3	1253.63	938.51	315.118
4	3250.27	3083.04	167.23
5	22688.55	10605.56	12082.99
6	32397.92	21601.36	10796.55

Figure 8 shows the GUI developed using the MATLAB. The first coordinate in the GUI has the option of viewing all the slices in a particular data set by moving the 'slider'. The second coordinate has the option of selecting the slices containing lesion from the dataset by specifying the slice numbers. The third coordinate shows the output of the segmentation of the lesion from a particular slice selected, by clicking on the 'segment' button. The entire lesion volume will also be displayed on clicking the 'lesion volume' button. The fourth coordinate on the screen displays the segmented solid portion of the lesion. And the volume of the solid and non-solid portion of the lesion will also be displayed to its right.

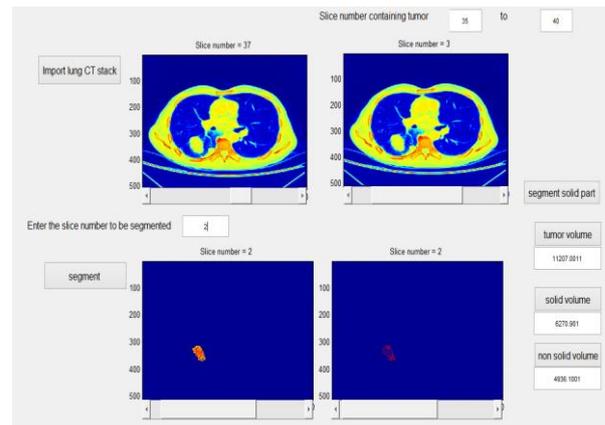


Figure 8: Screenshot of GUI developed

For 3D visualization of segmented lesion, the images obtained after segmentation were imported to the *licer3D* software, where in by using the 'editor' and 'make model tools', the segmented lesions could be visualized in three dimensions. Figure 9 (a) to (f) shows the 3D visualization of the segmented lesions of different subjects

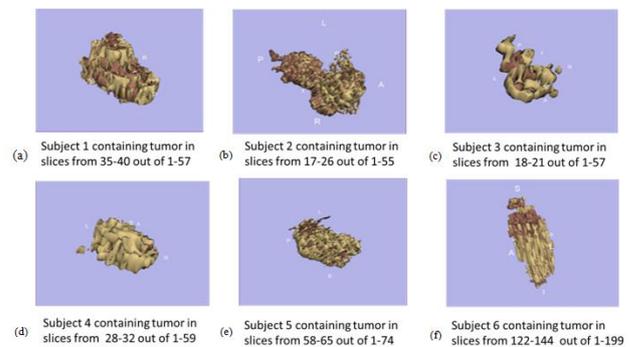


Figure 9: (a-f): 3D visualization of segmented lesion of 6 different subjects

5. Discussion

Feedback was taken from 5 radiologist on the work implemented. The results were presented to the radiologist along with the GUI that was developed and they were requested to provide feedback by answering the following questions:

1. Are you satisfied with the user interface built?
2. What is the scope of the semi-automatic method used for lesion extraction?
3. Is the solid and non-solid portion of the lesion extracted effectively?
4. Has the final output met the requirements?
5. Will the final output (i.e. volume estimation of the pulmonary sub-nodules) help in clinical diagnosis?
6. Is the GUI developed user friendly?

The feedback obtained was tabulated in the form of a bar graph as shown in Figure 10.

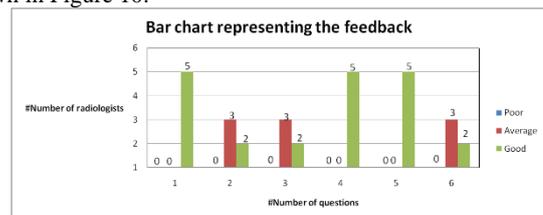


Figure 10: Bar charts of the remarks obtained from Radiologists.

The x-axis of the graph represents the questions that were asked and the y-axis represents the response of the radiologist. According to the feedback, the GUI implemented received good remarks. The final output obtained had also met the requirements and

would also help the radiologist in clinical diagnosis. However, physicians felt that the segmentation method used for lesion detection could be made fully automatic, thus eliminating the need to select the region of interest. For segmenting the solid and non-solid portion of the lesion in the 3DSlicer, the threshold value (obtained from the histogram) had to be manually entered to the software, so the physicians felt that the validation could have been done with some other software, which would automatically segment the solid and non-solid portion without one feeding the threshold values.

In order to compare the results, the volume obtained using the developed method was compared with the volume obtained using the 3DSlicer. The volume of the extracted lesion could be estimated in the Slicer using the 'label statistics tool' and are tabulated (as shown in Table 2).

From the Table 2, it is observed that the error between the volume estimated from the 3DSlicer and that obtained from the developed methodology was within the permissible limits (5-10%). For the subject 2 and 6, the error was found to be slightly higher as in the case of subject 2 the lesion was attached to the organ and in the case of subject 6, there were air gaps in the lesion region. The volume of solid and non-solid content of the lesion estimated with the developed method was compared with that obtained using the 3DSlicer and the percentage error is tabulated in the Table 3. From the Table 3, it can be clearly observed that the percentage error of solid content was found to be in the range of 1-3% and non-solid content was from 0-4.5%. This error range represents efficiency of the developed method using K-means segmentation.

Table 2: Comparison of the estimated volumes obtained using the developed method and the 3DSlicer

Patient no.	Volume of segmented lesion obtained using current method (mm ³)	Volume of segmented lesion obtained using SLICER (mm ³)	Error (%)
1	13180.82	13362.087	1.37
2	32149.69	30568.49	4.91
3	1253.63	1258.588	0.39
4	3250.27	3190.544	1.83
5	22688.55	22279.093	1.80
6	32397.92	30938.23	4.50

Table 3: Percentage error of solid and non-solid volume validated with 3DSlicer

Subject no.	Volume of solid portion of the lesion estimated using developed method (mm ³)	Volume of solid portion of the lesion estimated using Slicer (mm ³)	Volume of non-solid portion of the lesion estimated using developed method (mm ³)	Volume of non-solid portion of the lesion estimated using Slicer (mm ³)	%Error (solid portion)	%Error (non-solid portion)
1	6160.73	6405.57	7020.09	6956.51	1.85	0.482368
2	13428.81	14297.45	18720.87	17284.76	2.70	4.466948
3	938.51	952.511	315.118	372.189	1.11	4.55246
4	3083.04	3190.54	167.23	181.19	3.30	0.4295
5	10605.56	9725.89	12082.99	12554.04	3.87	2.07616
6	21601.36	22071.80	10796.55	10830.22	1.45	0.10393

6. Conclusion

Since, automatic segmentation of lesion from lung CT images is a very challenging task; an attempt is made to segment the lesions semi-automatically. And also solid and non-solid contents were extracted. Even though most of the medical imaging software like the 3DSlicer performs segmentation, the MATLAB codes and GUI developed in this work also shows their competency in segmenting the images in an efficient way. The idea behind this work is to develop a semi-automatic segmentation process which clusters out the intensity values based on K-means algorithm. The clusters can then be labeled and segmented. The slices containing the segmented portions of the lesion are saved and then represented as the 3D image which helps in better visualization to the radiologists. The only drawback of K-means segmentation method was that it required initializing the number of clusters to be used in the process, thus making it semi-automatic. However, there is a need to develop fully automatic method for segmentation, eliminating the need to select the region of interest, making the work much easier for the radiologist. The work can be further continued to separate the various components like lipids, calcium, hemorrhage and air space within the segmented lesion.

References

- [1] P.S. Gurudeep, N. Khandelwal, "Computed Tomography: An Increase Source of Radiation Exposure", International Conference on Innovations & Advances in Science, Engineering And Technology, volume 2(1), 2010, 119-131.
- [2] A. Fedorov, R. Bichel, J. Kalpathy, "3D Slicer as an Image Computing Platform for the Quantitative Imaging Network", International Journal of Radiology and Radiation Therapy, volume 1(2), 2003, 89-103.
- [3] F.H.Mahmood, W.A.Abbas, S.M.Ali, "Estimating the Lung Lesion Size in CT Images Using Image Segmentation Techniques", International Journal of Emerging Technology and Advanced Engineering, volume 4(7), 2003, 509-517
- [4] T. S. Sakti, K. Parasuraman, T. A. M. Devi, "Implementation of Lung Cancer Nodule Feature Extraction Using Threshold Technique", International Advanced Research Journal in Science Engineering and Technology, volume 3(8), 2004, 29-33.
- [5] J.Kaur, N.Garg, D.Kaur, "Segmentation and Feature Extraction of Lung Region for the Early Detection of Lung Lesion", International Journal of Science and Research, volume 3(6), 2007, 2327-2330.
- [6] Kothawari and S. N. Deepa, "A Study of Image Segmentation Techniques on CT lung Images", International Journal of Radiology Department, 129(2), 2008, 109-20.
- [7] P.S.Praseeja, K.Gangadharan, "Thyroid Segmentation and Volume Estimation Using CT Images", International Conference on Innovations & Advances in Science, Engineering And Technology, volume 3(5), 2010, 334-339
- [8] A. Jose, S. Ravi, M. Sambath, "Brain Lesion Segmentation using K-means Clustering and Fuzzy C-means Algorithm and its Area Calculation", In International Journal of Innovative Research in Computer and Communication Engineering, volume 2(2), 2011, 177-198.
- [9] R.A.Morey, C.M.Pettya, Y.Xua, J.P.Hayesa, H.R.Wagner, D.V.Lewisa, et al., "A Comparison of Automated Segmentation and Manual Tracing for Quantifying Hippocampal and Amygdala Volumes", Neuroimage, volume 45(3), 2011, 855-866.
- [10] W. J. Kostis, A. P. Reeves, David, F. Yankelevitz, I. Henschke, "Three-Dimensional Segmentation and Growth-Rate Estimation of Small Pulmonary Nodules in Helical CT Images", IEEE Transactions on Medical Imaging, volume 22(10), 2011, 1259-1273.
- [11] P. Daniel, S. Pieper, "3D Slicer as a Tool for Interactive Brain Lesion Segmentation", International Journal of Medical Imaging and Technology, volume 2(5), 2012, 6982-6984.