A New Hybrid Approach Using Fuzzy Clustering and Morphological Operations for Lung Segmentation in Thoracic CT Images

**SATYA PRAKASH SAHU*[1], PRIYANKA AGRAWAL[1],
NARENDRA D LONDHE[2] and SHRISH VERMA[3]**

[1]Department of Information Technology, National Institute of Technology, Raipur, India.
[2]Department of Electrical Engineering, National Institute of Technology, Raipur, India.
[3]Department of Electronics & Telecommunication Engineering,
National Institute of Technology, Raipur, India.

ABSTRACT

For computer-aided-diagnosis (CAD) System, the lung segmentation phase is having most significant role in the detection of lung cancer at initial stages. It is needed as preprocessing step for obtaining the accurate Region of Interest (ROI) area. Efficiency of CAD system is mainly depending on how the lungs are precisely segmented. The effective lung segmentation overcomes the various challenges offered in CAD system to deal with the cases of juxtapleural nodules. This paper emphasizes the proposed method for lung segmentation in CT images using clustering approach of fuzzy-c-means with automatic thresholding and morphological operations. The experimental database contains 20 patients' series (approximately 3600 images) from publically available LIDC-IDRI dataset which includes 10 general cases and 10 cases having juxtapleural nodules in pulmonary region. Reference standard contours are prepared by the expert through manually tracing the lungs boundary. The segmented lungs obtained through proposed method are compared with the reference standards with the use of various parameters. The proposed method achieved overall overlap ratio of 99.94% accuracy with 0.94 Jaccard's Index and 0.97 Dice similarity coefficient values.

Keywords: CAD system, Computed Tomography, Thresholding, Fuzzy c-means, Pulmonary nodules, Lung Segmentation, Morphological Operations.

INTRODUCTION

As per various statistics surveyed by different agencies, the lung cancer has been observed as most common cancer worldwide. In 2012, 1.8 million new lung cancer cases were estimated[1]. Deaths due to lung cancer are more as compared to other cancer related death worldwide. It contributes 13% for the cases registered, and total deaths due to lung cancer, its contribution was 19% in 2012, reported by GLOBOCAN 2012 (IARC) Section of Cancer Surveillance (Fig 1)[1]. In US, 224,400 cases are new cases of cancer which is about 14% and 158,080 cases related to cancer death are 27% as reported by American Cancer Society, Cancer facts and figure 2016[2]. In India, its

percentage was 6.9% for all new cancer cases and 9.3% for cancer related deaths. This data includes both male and female patients. It has been observed that most of the cancer cases are registered from Mizoram for both male and female (Age adjusted rate 28.3 and 28.7 per 100,000 population in males and females, respectively)[3].

Lung cancer leading to mortality rate higher in subsequent years and this encourages many researchers for the diagnosis of lung cancer in early stages. The survival rate may be increased by 70-80%, if this cancer is detected and diagnosed in prior stages[4]. The first screening test using thorax X-ray modality, the large nodule can only be detected, for small nodule of size even upto 2 cm could not be detected since of their unfavorable positions and locations[5]. With the advent of multi-slice, spiral CT, capability to perform a 3D reconstruction of the anatomical structure of human body, has explored detection of lung cancer in the early stages[6] and small nodules with a diameter lower than 5 mm (micro-nodules) can still be caught by LD-CT.

In a typical computer-aided diagnosis (CAD) system for lung cancer[7, 8], the lung segmentation is very important step for the detection of lung cancer or abnormalities and pulmonary diseases in lung area. For detecting proper lung region, CAD system have to overcome from various challenges like inhomogeneities in lung region and similar densities in other pulmonary structures like veins, bronchi, arteries, and bronchioles. Among all these challenges, the most important and difficult task is to determine the nodules attached to lung wall (i.e. juxtaplueral). Chances of missing the juxtaplueral nodules get increased if the lungs are not segmented properly. Approximately 5% to 17% true lung nodules are not detected because of poor and inaccurate lung segmentation, as stated by Armato *et al.*[9]. For the above stated reason it is necessary to effectively and accurately segment the lungs so that juxtaplueral nodules get properly detected.

In this paper, FCM algorithm is deployed for determining lung contour and to extract the pulmonary field or ROI for further detection of various types of nodules in these regions. The segmentation follows the correction of broken contour based on various morphological operations with the inclusion of juxtapluealralnodules, this phase also reduces the scope of over-segmentation. For testing the proposed method, approximately 3600 CT slices from 20 cases of LIDC-IDRI (Lung Imaging Database Consortium & Image Database Resource Initiative) dataset have been taken[10, 11]. The experts have been asked for trace the contour of lungs to create the reference images of CT slices. The obtained results have been compared and evaluated with the various performance metrics such as Jaccard's similarity index, Dice similarity coefficient, over-segmentation rate, under-segmentation rate, and overlap-ratio difference. The remaining section of this paper follows – Section II includes the review of previous work and also these research works are evaluated on the basis of challenges, issues, and effectiveness. In Section III, material and methods are presented in the following steps: data acquisition, preprocessing of data, and lung segmentation by using FCM with thresholding and morphological operations. Section IV describes experiments, evaluations of proposed method and results. In the last Section V, the results of proposed work is compared with some other works as conclusions and which is followed by limitations of the method and discussion with scope of future work in this domain.

Related work

For lung, segmentation, various automatic and semi-automatic approaches have been addressed by the authors. Automation level, processing time and the accuracy terms such as sensitivity & specificity are the measuring parameter for effectiveness and efficiency of the existing methodologies. The studies of these methods of segmentation mainly rely on the criteria like thresholds, shape based, active-contour, region based, clustering and morphological operations. For image segmentation of object, the most general method is gray level or intensity value based technique and the common known technique is Threshold based segmentation. The region of Lung lobes are having lower intensity level i.e. approximately -500 HU as compared to surrounding anatomical structure of thorax, due to this reason authors have adopted optimum threshold as their basis for segmenting the lungs in their research studies. Hu *et al.*[12] used iterative automatic threshold with morphological operation for smoothing the images in their irregular boundaries. Wei *et al.*[13]

analyses histogram for threshold to segment the pulmonary region. Ye *et al.*[14] had given the Adaptive fuzzy thresholding for lung segmentation in CT images. Helen *et al.*[15] used the particle swarm optimization (PSO) for the optimal 2D-Otsu thresholding algorithm for segmentation of lung parenchyma.

Shape based segmentation is another useful technique. From the collected dataset, prior shape information of lungs are registered and used further for extraction of ROI from the thorax region. This method uses an energy framework to help and guide deformable models in lung lobe segmentation and for such framework the parameters like edges or contour, points, etc. are characterized from previous data. For the local minima issues in lung segmentation Annangi *et al.*[16] proposed region based active contour technique for x-ray images by using initial shape and low level features. Kockelkorn *et al.*[17] took the severe abnormal data with their prior shape parameters and trained using k-NN classifier and devised interactive techniques for lung. Pattern recognition technique with the combination of statistical shape model and anatomical information for the robust lung segmentation has been adopted by Sofka *et al.*[18]. Pu *et al.*[19] proposed "break-and-repair" technique for the segmentation of medical images based on geometric modeling and shape. An approach given by Sun *et al.*[20] based on active shape model for the segmentation of lungs boundary and further to adapt a method based on constrained optimal surface using initial result of segmentation. Gill *et al.*[21] gave the feature based atlas approach for the initialization of active shape models in CT images.

The other significant method used in various articles[22-25] of lung segmentation is based on *Active contour method* (ACM). The main fundamental concept behind this method is minimizing of energy function iteratively through dynamic contour at each step. The most challenging part of this approach is the problem of initialization and convergence. Chenyang Xu and Jerry L Prince[22] proposed Gradient Vector Flow (GVF) as an external force for guiding the snake so that better convergence can be obtained. Wang *et al.*[23] further work for improving the initialization and convergence by Normally Biased Gradient Vector Flow (NBGVF).

Cui *et al.*[24] manipulate the ACM based on statistics of local region for medical images segmentation. An automatic method using ACM is exploited fully by Athertya *et al.*[25] that helps in reducing the user interaction for the segmentation of lungs.

Recently, one state-of-art method proposed by Soliman *et al.*[38] for accurate lung segmentation based on guided shape modeling. This approach provides the stack of CT images as 3D joint Markov-Gibbs random field (MGRF) model. The method segment the pathological lungs due to integrating the first and second order probabilistic descriptors of original and Gaussian scale space filtered image with prior guided adaptive shapes. However, this method requires larger training set of healthy and pathological lungs as guided landmarks for better performance.

The existing above approaches are based on general lung segmentation techniques. These methods are either not considering the cases of CT scan having the juxtapleural nodules or not so specific to deal with the complexity arises due to the presence of such nodules attached with the parenchyma of lungs. In this paper we are concerned about this complexity and prepared our database by including the 50% data with such cases and developed the framework to segment the accurate lung region even in the cases with CT images having juxtapleural nodules.

MATERIAL AND METHODS

Data

The database for the experiment contains 20 subjects and has been prepared from publically available LIDC-IDRI data set [10]. The database of LIDC-IDRI has 1018 cases of Lung CT scans under The Cancer Imaging Archive (TCIA) Public Access[26].

The collection of lung cancer thoracic CTs with marked-up annotation introduced by National Cancer Institute (NCI), later on upgraded by Foundation for the National Institutes of Health (FNIH) and Food and Drug Administration (FDA) under public-private partnership for the evaluation, training and development of CAD systems and specifies process for the detection and diagnosis of lung cancer. The annotation of this database has

been done by four experts (radiologists) and stored in XML file[11]. The database stored all Lung CT images with specifications as follows: DICOM image of size 512 X 512 pixels, tube current in the range of 265 to 570 mAs, tube voltage 120kVp and intensity level from -600 to 1600 HU[11].

Data Preprocessing

Data preprocessing is required for making the selected data as suitable for computable format. It starts with selecting the suitable cases from LIDC-IDRI database. The steps for the data preprocessing are shown in Fig. 2. In this experiment, 20 patient CT scan images have been taken with 10 cases of juxtapleural lung nodules. The contour or boundary tracing of lung region for each patient of all the slices are done by expert as the representation of reference standard or ground truth. The accuracy can be further enhanced if the image pixel value is converted to double format. The optimum level of contrast for the selected data is set by the process of image enhancement. The acquired images of LIDC possess DICOM format data so it need to be converted into Hounsfield Unit (HU) and to set

window size (level and width) appropriately for the segmentation of lungs.

Linear transformation performs the rescaling of slope and intercept and lead to defined range of HU values of the images. The final step for getting preprocessing is the conversion from 16-bits images into 8-bits images so as to get the suitable data for further process.
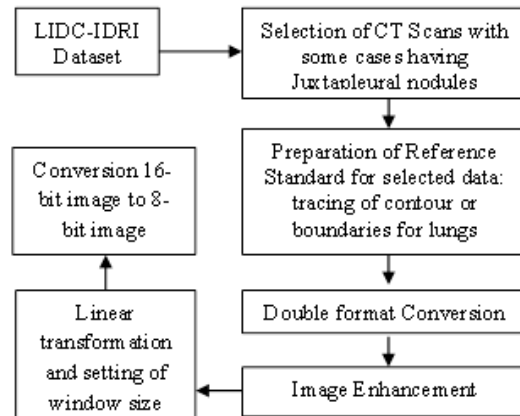
Lung Segmentation

The proposed method for lung segmentation is shown in Fig. 3 as flowchart and is implemented by the following steps:

A. Applying Fuzzy-c-means clustering based segmentation to generate the gray-scale masked image of the CT slices. The derivation & steps of the algorithm are given in later section.

B. By using automatic thresholding or otsu's algorithm, the output of fuzzy-c-means mask is converted into binary image (method
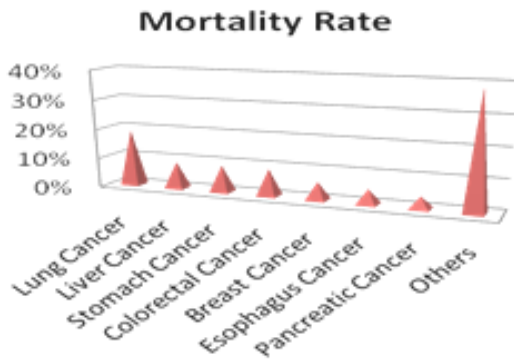
Fig. 1: Worldwide Incidence and Mortality rate [Courtesy of GLOBOCAN 2012 (IARC) Section of Cancer Surveillance]


Fig. 2: Basic preprocessing steps
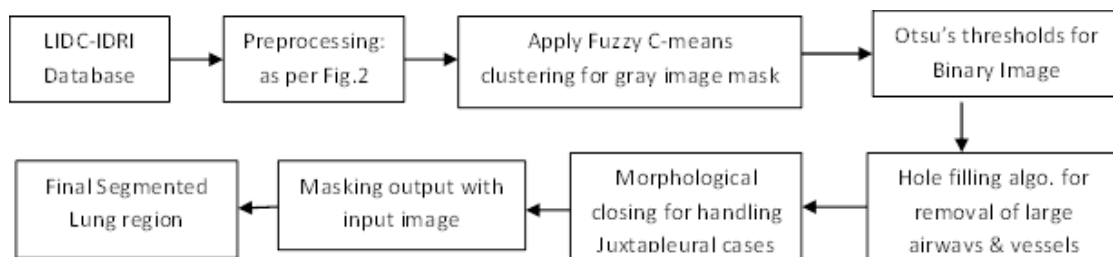
Fig. 3: Flowchart for the proposed method

Table 1: Statistical analysis of Over-segmentation, Under-segmentation, Overlap ratio difference, Dice similarity coefficient and Jaccard's similarity index of proposed method

	Over-segmentation	Under-segmentation	Overlap-ratio diff.	Dice similarity Coefficient	Jaccard similarity Index
Mean	0.039428	0.019319	0.058747	0.971036	0.944359
SD	0.026018	0.017359	0.036979	0.033143	0.017973
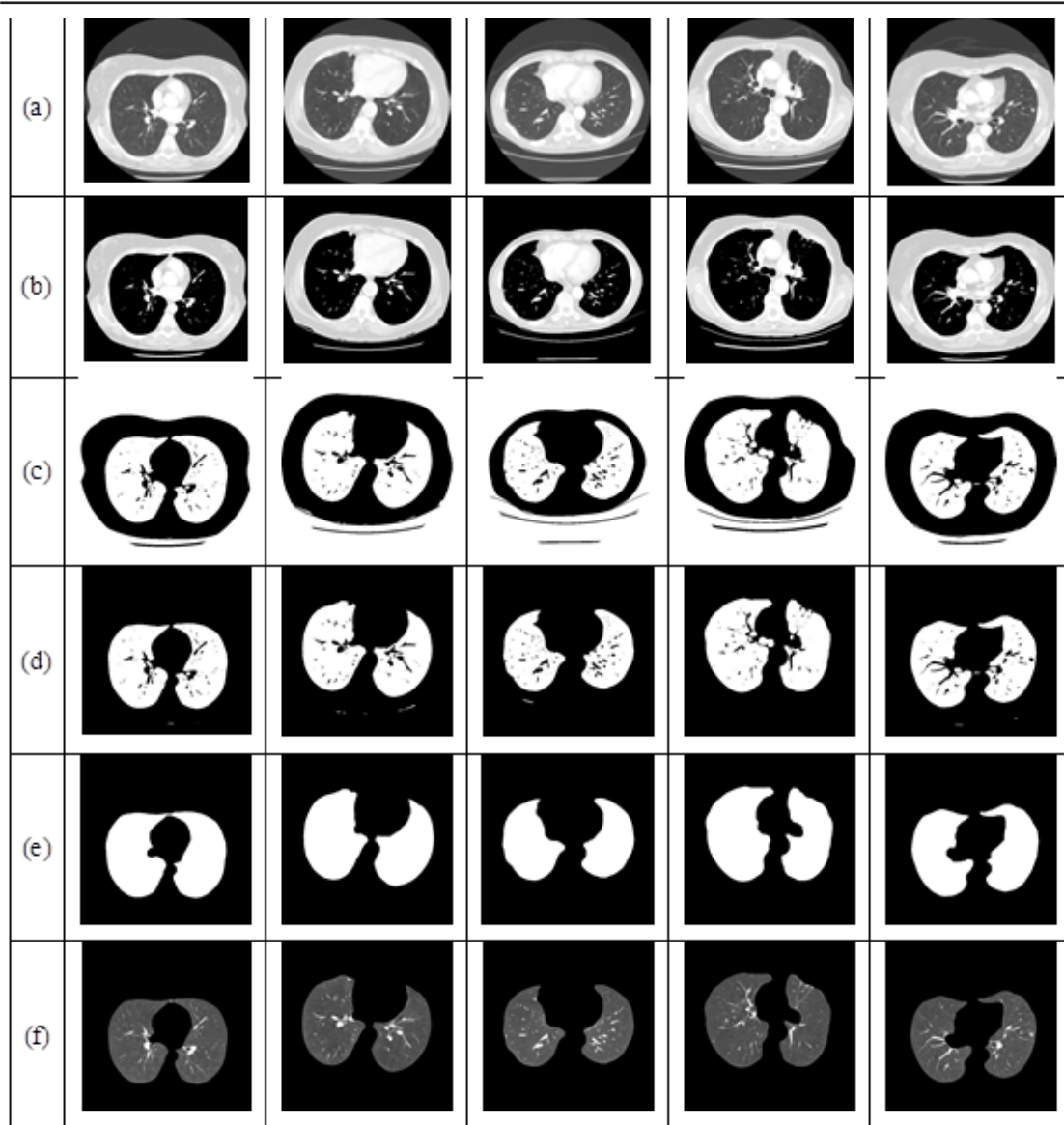
Fig. 4: The lung segmentation results of our proposed approach for the images of *general* cases. Row (a) are five original lung CT images (b) gray scale masked images after applying FCM (c) binary images of previous step using Otsu's threshold (d) background subtracted images; (e) final lung mask after removing the large airways, trachea, and blood vessels and (f) final result– segmented lungs

discussed in later section) followed by the background subtraction for ROI in image.

C.   Large airways, blood vessels and trachea are removed with the application of hole-filling algorithm. This algorithm is based on the process of dilation, intersection and complementation.

Let I be the set whose boundary pixel are labeled by 1s (considering 8-connected boundaries) & each boundary surrounding a hole (background region). With the given point x in each hole, the goal is to fill the holes by 1s. This step is followed by morphological reconstruction for removing the lighter border by reducing the overall intensity of the border structures.
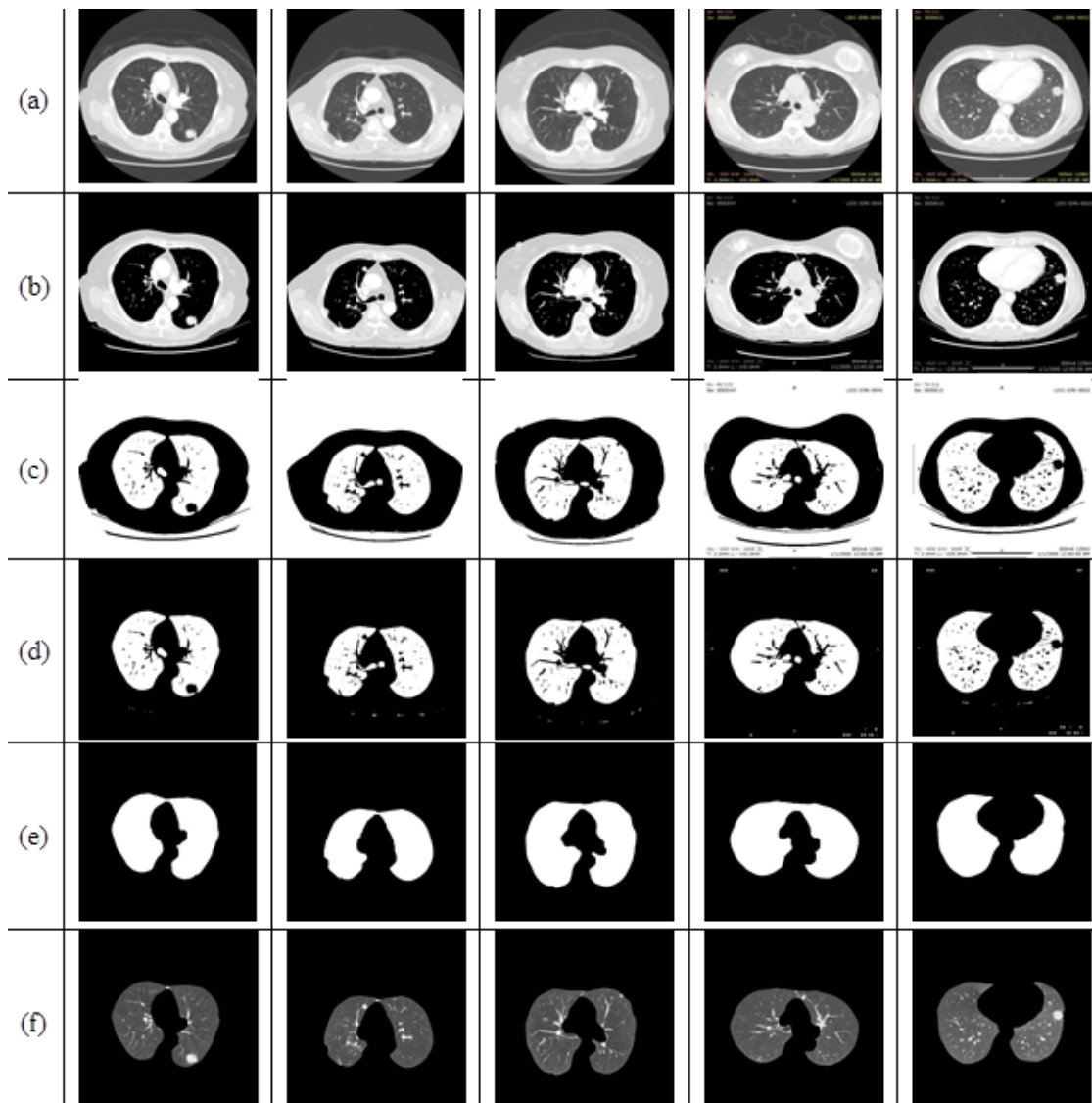
Fig. 5: The lung segmentation results of our proposed approach for the image having *juxtapleural* nodules. Row (a) are five original lung CT images (b) gray scale masked images after applying FCM (c) binary images of previous step using Otsu's threshold (d) background subtracted images; (e) final lung mask after removing the large airways, trachea, and blood vessels and (f) final result– segmented lungs

D.   The morphological closing operations are applied to deal with the cases having juxtapleural nodules.

E.   At last masking the output image with input gray-scale image for obtaining the segmented lung region.

Fuzzy-c-means Algorithm

Fuzzy c-means (FCM) is a method of clustering and is used here for lung segmentation. The flexible sense of this method is that a data point can belong to two or more cluster at a time. FCM is initially given by Dunn [27] and further enhanced by Bezdek [28]. The algorithm is based on minimization of the objective function (J_p)

$$J_p = \sum_{i=1}^{N} \sum_{j=1}^{c} v_{ij}^{p} \left\| x_i - c_j \right\|^2 \qquad ...(1)$$

where, p is any real number greater than 1 i.e. $1 \leq p < \infty$, v_{ij} is the membership matrix of x_i in the cluster *j*, is of d-dimensional data measured

(say pixel intensity), is the centroid or cluster center and ||*|| is a norm representing the similarity between any considered pixeland thecluster center.

Based on the optimization of above objective function fuzzy partitioning is performed. The membership and the centroid are updated as follows:

$$v_{ij} = \frac{1}{\sum_{k=1}^{c} \left(\frac{\|x_i - c_j\|}{\|x_i - c_k\|} \right)^{\frac{2}{p-1}}} \qquad ...(2)$$

$$c_j = \frac{\sum_{i=1}^{N} v_{ij}^{p} \cdot x_i}{\sum_{i=1}^{N} v_{ij}^{p}} \qquad ...(3)$$

The iteration will terminate when
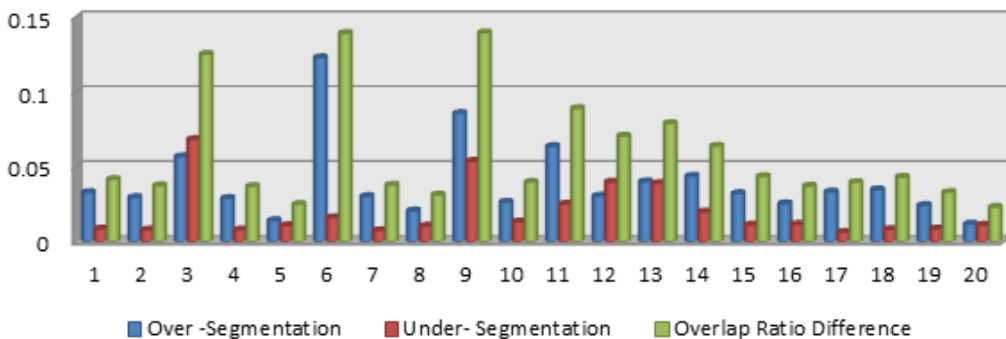
(a)
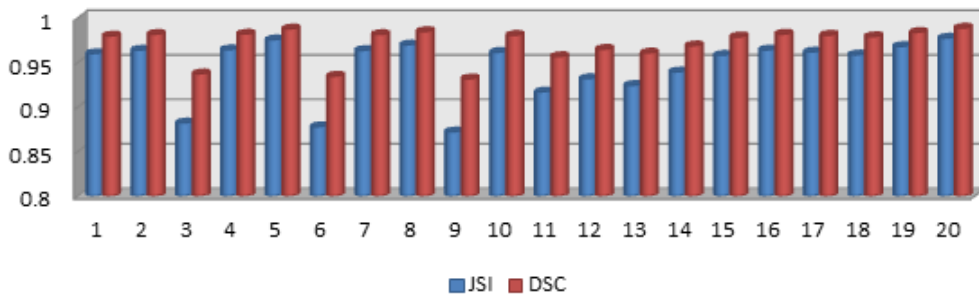
(b)

Fig. 6: Proposed method bar plots of 20CT scans (patients):
(a) Over-segmentation, Under-segmentation, Overlap ratio difference, and
(b) Jaccard'ssimilarity index (JSI)and Dice similarity coefficient (DSC)of proposed method

Table 2: Performance comparison of proposed method for segmentation of lungs.

Authors & Works	Datasets	Overlap-ratio	Over Segmen-tation	Under Segmen-tation	Dice Similarity Coefficient	Jaccard Similarity Index
Wei et al. [35]	97 nodules CT scans25 juxtapleural nodule scans	95.24%	-	-	-	-
S.Zhou et al.[36]	20 CT scans with 65juxtapleural nodules	95.81%	-	-	-	-
Liao et al. [37]	80 cases divided in 4 categories of non nodules, benign nodules, malignant SPN and pleural nodules with 20 cases in each category	92.22%	-	-	-	Non nodules 0.941 Benign nodules 0.9342Malignant SPN 0.9257 Pleural nodules 0.8821
Proposed method	20 cases of LIDC-IDRI datasets with 10 cases of juxtapleural nodules (3600 CT images approx.)	99.94%	0.0394	0.0193	0.9710	Overall 0.9444 Pleural nodules 0.9389

$$max_{ij} = \left\{ \left| v_{ij}^{(k+1)} - v_{ij}^{k} \right| \right\} < \Phi, \qquad ...(4)$$

where Φ is a stopping criterion ranges 0<<1 and k as the iteration steps. This procedure follows the converging toward a local minimum of   .

The step wise algorithm follows as:

a. *Image intensity distribution computed for different gray-scale value.*

b. *Initialize V=[] matrix, V$^{(0)}$*

c. *At kth step: calculate the centroid vectors C$^{(k)}$=[] with V$^{(k)}$as per eq. (3).*

d. *Update membership matrices V$^{(k)}$, V$^{(k+1)}$as per eq. (2).*

e. *For termination, if ||V$^{(k+1)}$ - V$^{(k)}$||< then STOP; else return to step (c).*

Automatic threshold Algorithm

Otsu's [29] had given the automatic method of thresholding. The estimated threshold of gray level value (g*) is obtained by simple and effective method that minimizes the weight within-class variance and maximizes the between class variance assuming that the histogram are bimodal. For the given image, the representation of pixels in M gray levels *[1, 2…, M]*, the normalized histogram as probability distribution (p_i) can be given as:

$$p_i = \frac{n_i}{N}, p_i \geq 0, \sum_{i=1}^{M} p_i = 1 \qquad ...(5)$$

where, n_i represents the number of pixels at the gray level *i*. and

$N = n_1 + n_2 +…+ n_M$ is the total number of pixel.

$$\sigma_c^2(g) = \frac{[\mu_G s(g) - \mu(g)]^2}{s(g)[1 - s(g)]} \qquad ...(6)$$

where, *s(g)* and $\mu(t)/s(t)$ represent class occurrence probability and gray level mean value respectively of background; *[1- s(g)] and* $(\mu_G - \mu(g)$ $(1 - s(g)$ represent class occurrence probability and gray level mean value respectively of object.

Thus with the analysis of histogram and checking for each gray level, the general form of optimal threshold *g*** that maximizes $\sigma_c^2(g^*)$ can be defined as:

$$\sigma_c^2(g^*) = \max_{1 \leq g < M} \sigma_c^2 \qquad ...(7)$$

Result Analysis

Experimental database contains 20 patients or CT series (approximately 3600 slices or images examined) out of which 10 patients having juxtaplueral nodules in their CT images, taken from the LIDC-IDRI dataset. The lung boundaries or contours were manually traced by experts under the observation or supervision of thoracic radiologists for each case. These tasks have been performed by an open-source system "The Medical Imaging Interaction Toolkit", a workbench and toolkit of MITK 2016.11[30]. This workbench has deployed here as annotation tool that includes Visualization Toolkit (VTK) and Insight Toolkit (ITK). The flexibility with this toolkit is that it support border corrections and contour adjustments if some inaccuracies are observed. These manually traced boundaries or contours were considered as ground truth for respective images. Finally, for the evaluation and comparison of experimental results for proposed method with reference standard based on evaluation measures have been presented.

Evaluation measures

The following metrics have been utilized for the performance of results of lung segmentation:*over-segmentation & under-segmentation rate, overlap ratio, Dice similarity coefficient andJaccard's similarity index.*

The over-segmentation is given as the area (number of pixel) or volume (number of voxels) that is included as a portion of the segmented ROI (result of proposed method) but are absent in reference standard or ground truth image whereas under-segmentation is given as the area or volume that are included in ground truth, but are absent in segmented ROI[31, 32]. The volume overlap ratio (*VO)* is given as the relative overlap between two volumes i.e. segmentation mask produced by our proposed method (*V$_m$)* and ground truth (*V$_g$*) and formulated by taking the intersection *V$_m$*and *V$_g$* divided by their union[33]. The definitions of these measures are as follows:

The over-segmentation rate $O_s(V_m, V_g)$ is given by

$$O_s(V_m, V_g) = \left| \frac{V_m \setminus V_g}{V_g} \right| \qquad ...(8)$$

where $V_m \setminus V_g$ is the relative compliment of V_m in V_g. In the similar manner the under-segmentation rate $U_s(V_m, V_g)$ is given by

$$O_s(V_m, V_g) = \left| \frac{V_m \setminus V_g}{V_g} \right| \qquad ...(9)$$

$$VOR(V_m, V_g) = \frac{|V_m \cap V_g|}{|V_m \cup V_g|} \qquad ...(10)$$

The other measures which have been used for evaluation of the segmentation accuracies in the terms of similarity between the segmentation results and ground truths are Dice-similarity coefficient (*DSC*) and Jaccard's similarity index (*JSI*)[34]. These methods compare the similarity between the segmented ROIs (number of pixels or voxels) of adopted method (R_m) to the ground truth region (R_g).

The Jaccard's similarity index is given as the overlap degree between R_m and R_g and follows as

$$JSI(R_m, R_g) = \frac{|R_m \cap R_g|}{|R_m| + |R_g| - |R_m \cap R_g|} \qquad ...(11)$$

Similarly Dice-similarity coefficient is defined as the overlap of two regions i.e. R_m and R_g and can be given as

$$DSC(R_m, R_g) = \frac{2|R_m \cap R_g|}{|R_m| + |R_g|} \qquad ...(12)$$

The values of *JSI* and *DSC* will be in range of [0,1] and higher values indicates better results in terms of segmentation accuracy.

RESULTS AND DISCUSSION

The performance evaluation of our proposed method has been tested on 20 CT scans (approximately 3600 CT slices) including 10 CT scans of juxtapleural nodules. Fig.4 & Fig.5 show the sample results (randomly selected slice of 5 CT scans each of the two cases) stepwise for our adopted method in lung segmentation for general cases and cases having juxtapleural nodules respectively.

With the use of evaluation measures, in order to show the average distance of segmentation error of 20 patients (each patient having 150-200 slice images) undertaken are represented in horizontal-axis (from 1 to 20) in Fig.6 (a), we depicts bar plots of cumulative probability distribution based on over-segmentation, under-segmentation and overlap ratio difference error in vertical-axis. Similarly in Fig.6 (b) the accuracies are represented in terms of similarity measures i.e. bar plots of the cumulative probability distribution based on the Dice-similarity coefficient and Jaccard's similarity index (both represented in vertical-axis) for 20 patients (represented in horizontal-axis) are shown.

In the Table 1 we have shown the overall results of our adopted method for the all data that achieved value (mean values) : over-segmentation of 0.0394, under-segmentation of 0.0193, and overlap ratio difference of 0.0587, overlap ratio of 99.9413, Dice-similarity coefficient of 0.9710 and Jaccard's index of 0.9444. For the cases having juxtapleural nodules our approach achieved an outstanding performance (in mean values) i.e overlap ratio of 99.9347% and Dice-similarity coefficient value of 0.9679 and Jaccard's index value of 0.9389.

The number of authors had given various methods and techniques for lung segmentation, but only some are effective for the cases of juxtapleural nodules. Table 2 compares the performance for our approach with the other methods for lung segmentation. The proposed method achieved the best average overlap ratio among the others listed in the table. In terms of Jaccard's similarity index the proposed method performs better than

Liao[37]. Although, for fair comparison of methods is possible if the same datasets and image standards have accessed. But due to non availability of other database we have taken the public datasets and tested our hybrid approach and hence it is not possible for complete comparison. However comparing to other approaches, our proposed method is achieving the good results and generating the better accuracies for lung segmentation.

CONCLUSION

Lung segmentation is very crucial and important step in CAD system for the detection and diagnosis of lung cancer in prior stages. For the past few decades' number of research works have been given by various authors for lung segmentation and shown their effectiveness in different cases. In this paper, we have implemented a new hybrid approach using fuzzy clustering (FCM) with morphological operations for the segmentation of lungs in CT images that have shown very effective results in the cases of lungs having juxtapleural nodules.

The efficiency of our approach is tested on 20 cases (i.e. 10 general and 10 cases of juxtapleural nodules) of lung thoracic CT scans from LIDC-IDRI a public dataset. The experimental results show that the proposed method can correctly including all juxtapleural nodules into the ROI of lungs with the minimum rate of under and over-segmentation and thereby achieved overall 99.94% of overlap ratio. In the result section, we have also given the performance in terms of of Dice similarity coefficient and Jaccard's index and shown the accuracies of 0.9710 and 0.9444 respectively. The average processing time for the adopted method is 0.80 s per CT slice using MATLAB software experimented on PC with CPU Intel Core i5 3.1 GHz and 8 GB RAM. The proposed approach can fulfill the requirements of CAD system for the lung cancer and provide the accurate search space for further processing toward the detection and classification of pulmonary nodules.

ACKNOWLEDGEMENT

REFERENCES

1. Ferlay J, Soerjomataram I, Ervik M, Dikshit R, Eser S, Mathers C, *et al.* Lyon, France: International Agency for Research on Cancer; 2013. GLOBOCAN 2012 v1.0, Cancer Incidence and Mortality Worldwide: IARC CancerBase No. 11. Available from: http://globocan.iarc.fr , accessed on 04/10/2016.
2. American Cancer Society, Cancer facts and ûgures, 2016.
3. Indian Council of Medical Research; 2013. National Cancer Registry Programme. Three Year Report of Population Based Cancer Registries: 2009-2011. Available from: http://www.ncrpindia.org.
4. Swensen SJ, Jett JR, Sloan JA, Midthun DE, Hartman TE, Sykes AM, Aughenbaugh GL, Zink FE, Hillman SL, Noetzel GR, Marks RS. Screening for lung cancer with low-dose spiral computed tomography. *American journal of respiratory and critical care medicine.;* **165**(4):508-13 (2002).
5. Flehinger BJ, Melamed MR, Zaman MB, Heelan RT, Perchick WB, Martini N. Early Lung Cancer Detection: Results of the Initial (Prevalence) Radiologic and Cytologic Screening in the Memorial Sloan-Kettering Study 1–3. *American Review of Respiratory Disease.;* **130**(4):555-60 (1984).
6. Awai K, Murao K, Ozawa A, Komi M, Hayakawa H, Hori S, Nishimura Y. Pulmonary Nodules at Chest CT: Effect of Computer-aided Diagnosis on Radiologists' *Detection Performance 1. Radiology.;* **230**(2):347-52 (2004).
7. El-Baz A, Beache GM, Gimel'farb G, Suzuki K, Okada K, Elnakib A, Soliman A, Abdollahi B. Computer-aided diagnosis systems for lung cancer: challenges and methodologies. *International journal of biomedical imaging.* **29**; (2013).

8. Antonelli M, Cococcioni M, Lazzerini B, Marcelloni F. Computer-aided detection of lung nodules based on decision fusion techniques. *Pattern Analysis and Applications.*; **14**(3):295-310 (2011).

9. Armato SG, Sensakovic WF. Automated lung segmentation for thoracic CT: Impact on computer-aided diagnosis1. *Academic Radiology;* **11**(9):1011-21 (2004).

10. Armato III, Samuel G., McLennan, Geoffrey, Bidaut, Luc, McNitt-Gray, Michael F., Meyer, Charles R., Reeves, Anthony P., … Clarke, Laurence P. (2015). Data From LIDC-IDRI. The Cancer Imaging Archive. http://doi.org/101.7937/K9/TCIA.2015.LO9QL9SX

11. Armato SG, McLennan G, Bidaut L, McNitt Gray MF, Meyer CR, Reeves AP, Zhao B, Aberle DR, Henschke CI, Hoffman EA, Kazerooni EA. The lung image database consortium (LIDC) and image database resource initiative (IDRI): a completed reference database of lung nodules on CT scans. *Medical physics.;* **38**(2):915-31 (2011).

12. Hu S, Hoffman EA, Reinhardt JM. Automatic lung segmentation for accurate quantitation of volumetric X-ray CT images. *IEEE transactions on medical imaging.;* **20**(6):490-8 (2001).

13. Wei Q, Hu Y, Gelfand G, MacGregor JH. Segmentation of lung lobes in high-resolution isotropic CT images. *IEEE Transactions on biomedical engineering.;* **56**(5):1383-93 (2009).

14. Ye X, Lin X, Dehmeshki J, Slabaugh G, Beddoe G. Shape-based computer-aided detection of lung nodules in thoracic CT images. *IEEE Transactions on Biomedical Engineering.;* **56**(7):1810-20 (2009).

15. Helen R, Kamaraj N, Selvi K, Raman VR. Segmentation of pulmonary parenchyma in CT lung images based on 2D Otsu optimized by PSO. InEmerging Trends in Electrical and Computer Technology (ICETECT), 2011 International Conference on 2011 Mar 23 (pp. 536-541). IEEE.

16. Annangi P, Thiruvenkadam S, Raja A, Xu H, Sun X, Mao L. A region based active contour method for x-ray lung segmentation using prior shape and low level features. InBiomedical Imaging: From Nano to Macro,

17. Kockelkorn TT, van Rikxoort EM, Grutters JC, van Ginneken B. Interactive lung segmentation in CT scans with severe abnormalities. InBiomedical Imaging: From Nano to Macro, 2010 IEEE International Symposium on 2010 Apr 14 (pp. 564-567). IEEE.

18. Sofka M, Wetzl J, Birkbeck N, Zhang J, Kohlberger T, Kaftan J, Declerck J, Zhou S. Multi-stage learning for robust lung segmentation in challenging CT volumes. Medical image computing and computer-assisted intervention–MICCAI 2011. 667-74 (2011).

19. Pu J, Paik DS, Meng X, Roos J, Rubin GD. Shape "break-and-repair" strategy and its application to automated medical image segmentation. *IEEE transactions on visualization and computer graphics.;* **17**(1):115-24 (2011).

20. 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 transactions on medical imaging.*; **31**(2):449-60 (2012).

21. Gill G, Toews M, Beichel RR. Robust initialization of active shape models for lung segmentation in CT scans: A feature-based atlas approach. *Journal of Biomedical Imaging.;* 2014:13.

22. Xu C, Prince JL. Snakes, shapes, and gradient vector flow. *IEEE Transactions on image processing.;* **7**(3):359-69 (1998).

23. Wang Y, Liu L, Zhang H, Cao Z, Lu S. Image segmentation using active contours with normally biased GVF external force. *IEEE Signal Processing Letters.;* **17**(10):875-8 (2010).

24. Chan TF, Vese LA. Active contours without edges. *IEEE Transactions on image processing.;* **10**(2):266-77 (2001).

25. Athertya JS, Kumar GS. Automatic initialization for segmentation of medical images based on active contour. InBiomedical Engineering and Sciences (IECBES), 2014 IEEE Conference on 2014 Dec 8 (pp. 446-451). IEEE.

26. Clark K, Vendt B, Smith K, Freymann J, Kirby J, Koppel P, Moore S, Phillips S, Maffitt D, Pringle M, Tarbox L. The Cancer Imaging

Archive (TCIA): maintaining and operating a public information repository. *Journal of digital imaging.;* **26**(6):1045-57 (2013).

27. Dunn J C. A Fuzzy Relative of the ISODATA Process and Its Use in Detecting Compact Well-Separated Clusters, *Journal of Cybernetics.* **3**(3): 32-57 (1973).

28. Bezdek, James C. *Pattern recognition with fuzzy objective function algorithms*. Springer Science & Business Media,(2013).

29. Otsu N. A threshold selection method from gray-level histograms. *Automatica.;* **11**(285-296):23-7 (1975).

30. MITK: The Medical Imaging Interaction Toolkit, German Cancer Research Center, Division of Medical and Biological Informatics, 2016, http://mitk.org/wiki/MITK.

31. Pu J, Roos J, Chin AY, Napel S, Rubin GD, Paik DS. Adaptive border marching algorithm: automatic lung segmentation on chest CT images. *Computerized Medical Imaging and Graphics.;* **32**(6):452-62 (2008).

32. Hoover A, Jean-Baptiste G, Jiang X, Flynn PJ, Bunke H, Goldgof DB, Bowyer K, Eggert DW, Fitzgibbon A, Fisher RB. An experimental comparison of range image segmentation algorithms. *IEEE transactions on pattern analysis and machine intelligence.;* **18**(7):673-89 (1996).

33. VanRikxoort EM, de Hoop B, Viergever MA, Prokop M, van Ginneken B. Automatic lung segmentation from thoracic computed tomography scans using a hybrid approach with error detection. *Medical physics.;* **36**(7):2934-47 (2009).

34. Sampat MP, Wang Z, Gupta S, Bovik AC, Markey MK. Complex wavelet structural similarity: A new image similarity index. *IEEE transactions on image processing;* **18**(11):2385-401 (2009).

35. Wei Y, Shen G, Li JJ. A fully automatic method for lung parenchyma segmentation and repairing. *Journal of digital imaging.*; **26**(3):483-95 (2013).

36. Zhou S, Cheng Y, Tamura S. Automated lung segmentation and smoothing techniques for inclusion of juxtapleural nodules and pulmonary vessels on chest CT images. *Biomedical Signal Processing and Control.;* **13**:62-70 (2014).

37. Liao X, Zhao J, Jiao C, Lei L, Qiang Y, Cui Q. A Segmentation Method for Lung Parenchyma Image Sequences Based on Superpixels and a Self-Generating Neural Forest. *PloS one.;* **11**(8):e0160556 (2016).

38. Soliman A, Khalifa F, Elnakib A, El-Ghar MA, Dunlap N, Wang B, Gimel'farb G, Keynton R, El-Baz A. Accurate lungs segmentation on CT chest images by adaptive appearance-guided shape modeling. *IEEE transactions on medical imaging;* **36**(1):263-76 (2017).