AN EFFICIENT ALGORITHMIC SOLUTION FOR AUTOMATIC SEGMENTATION OF LUNGS FROM CT IMAGES

F. Shaukat and G. Raja

Faculty of Electrical and Electronics Engineering
University of Engineering and Technology, Taxila, Pakistan.
Corresponding author's email: furqan.shoukat@uettaxila.edu.pk

ABSTRACT: A novel technique for lung segmentation from input Computed Tomography (CT) images using optimal thresholding was developed. Initially, the CT image was segmented by optimal thresholding. The lung volume was obtained using connected component labeling method by removing irrelevant information. The resultant image contained holes which were filled by morphological operations. A novel technique to separate the lungs was introduced which effectively separated the right and left lungs. Finally, the lung contour was smoothed by rolling ball algorithm to include any juxta pleural nodules. The proposed system was evaluated using 84 scans of publicly available dataset Lung Image Database Consortium (LIDC). The proposed system achieved an overlap measure of 0.985 and the root mean square difference between the proposed method and ground truth was 0.47 mm.

Keywords: Lung segmentation, Optimal thresholding, Computer aided detection, Lung image database consortium

(Received 10-01-2018 Accepted 30-03-2018)

INTRODUCTION

Lung cancer is one of the leading causes of the deaths around the world with the smallest rate of survival after diagnosis. The survival rate can be increased by early nodule detection (Siegel *et al*., 2016). It is found in both developed and under developed countries (Diaz *et al*., 2014). According to an estimate, 225,000 people are diagnosed with lung cancer every year in United States costing $12 billion in health care (Mariotto *et al*., 2011). Another study shows that 433 americans died of lung cancer every day (Howlader *et al*., 2016). The situation in under developed countries is more worse. Lung cancer is the most common type of cancer in Asia with the highest risk in South East Asia (Moore *et al*., 2010). Pakistan has also been the victim of lung cancer with the danger increasing every passing day. According to a study conducted in January 2014, the lung cancer occurrence and mortalities both are increasing in the country. Lack of awareness, poor hygienic conditions and meat consumption are the main reasons apart from tobacco which is the primary source of this deadly disease (Luqman *et al*., 2014).

Lung segmentation can be defined as the process of extracting the lung volume form input CT image and removing the background and other irrelevant components (Choi *et al*., 2014). Lung segmentation serves as a prerequisite to the nodule detection. Accurate lung segmentation plays an important role to enhance the efficiency of lung nodule detection system (Valente *et al*., 2015). Numerous methods have been proposed in literature for the extraction of lung volume from CT image such as optimal thresholding, rule-based region growing, global thresholding, 3-D-adaptive fuzzy thresholding, hybrid segmentation, and connected component labeling (Dehmeshki *et al*., 2007; Suárez-Cuenca *et al*., 2009; Choi *et al*., 2012). After the initial segmentation, juxta-pleural nodules are included by refining the extracted lung volume. To do this, a chain-code method, a rolling ball algorithm, and morphological approaches have been generally used (Ali *et al*., 2008; Van Rikxoort *et al*., 2009; De Nunzio *et al*., 2011;).

Lung segmentation techniques can be broadly classified into four categories (i) Shape Models (ii) Edge-based techniques (iii) Thresholding (iv) Deformable Boundaries (El-Baz *et al*., 2013). Recently some studies (Filho *et al*., 2017; Soliman *et al*., 2017) using shape models have surpassed the conventional approaches (Shi *et al*., 2016) of lung segmentation in terms of accuracy.

In summary, each technique has its own pros and cons. Threshold based techniques are very good when it comes to high contrast CT images but the performance can vary with the low contrast pathologies. Thresholding can also be affected with different imaging protocols and image acquisition scanners. Moreover, different lung structures like blood vessels, bronchioles and bronchi have so close densities with chest tissues that it is very difficult to accurately threshold the region of interest and it requires special post-segmentation processes for accurate segmentation. Deformable boundary based techniques have the disadvantage of extra sensitivity of initialization. Further they are unable to overcome the inhomogeneity of lung volume with the use of traditional external forces like edges and gray

levels. Hence it becomes really difficult to achieve the accurate lung segmentation by guiding the deformable model. The accuracy of shape-based segmentation techniques depends on the accurate registration of prior shape-model with respect to the CT image. Poor registration in this regard can affect the overall performance and it is the main limitation of shape based techniques. Further the diversity of lung pathologies makes it difficult to accurately segment the lung fields. The aim of this study was to propose a novel technique for lung segmentation using optimal thresholding and connected component labelling with contour smoothing which could effectively separate the right and left lungs.

MATERIALS AND METHODS

Our proposed method consisted of series of steps. Initially, the Computed tomography (CT) image was segmented using optimal thresholding and the lung volume was obtained using connected component labeling method and other irrelevant information was removed (Hu *et al*., 2001). The resultant image contained holes which were filled with the hole filling algorithm e.g. morphological operations. A novel technique to separate the lungs was introduced which effectively separated the right and left lungs. Finally, the lung contour was smoothed by rolling ball algorithm to include any juxta pleural nodules (Armato *et al*., 1999). Optimal thresholding was used because in original CT images, density of the lung volume was different from the background which provided the basis of intensity based segmentation techniques to be used effectively. The proposed work flow for lung volume segmentation is shown in (Fig. 1). In the following section, each of this step was described in detail.

Thresholding: For optimal thresholding, let T^i be the threshold after the i^{th} step. The lung CT scan was divided in two density groups. The HU values in lung CT scan varied from -2000 HU to +2000 HU. The lung area also called non-body area was a low-density area which ranged from -1000 HU to -500 HU (Hu *et al*., 2001). The CT scanner area was also part of the non-body area. The body area contained the surroundings of lung lobes. Because the lungs were in non-body area, we initially selected a threshold value of -500 HU for T^o. For selection of new threshold, we applied T^i to the lung image. The μ_o and μ_b were the mean intensities of the object and background in the lung region, respectively, the new threshold was given by Hu *et al*., (2001). Results of optimal thresholding on a few sample images can be seen in column (b) of Fig-2.

Connected component labelling: After applying optimal thresholding, lung CT image containing body and non-body area was obtained. White area belonged to non-body area and black belonged to body area. We were interested in extracting the lung region from non-body voxels. To achieve this, we applied 3D connected component labeling to initially thresholded image $f(x, y, z)$ to acquire the lung region from non-body voxels. Using this technique, the first and second largest volumes were selected. Most of the unwanted non-body components were ignored in the volume selection. After the background removal, the resultant image at this stage contained holes in lung lobes, which could be potential nodules or vessels. These must be included to the lung region for accurate detection and thus filled by morphological operations. The resultant image at this stage can be seen in column (c) of Fig. (2). The hole-filled image could also contain the potential nodules at the border known as juxta-pleural nodules. These nodules must be included for accurate detection. To include these, we used a rolling ball algorithm (Armato *et al*., 1999).

Boundary repair of lung parenchyma: Various methods have been proposed for repairing gaps in the lung boundaries. The rolling ball method is one of the popular repair method. In this method, a two-dimensional ball filter was placed tangentially on the boundary of the lungs and rolled along the direction of the lung boundary. If there was a gap in the lung boundary, it was identified by the contact of rolling ball filter at more than two points. This gap was filled by the new contour segment that linearly connected the two end points of gap. The basic principle of this method was to use the change in boundary curvature and circular filter template for lung contour morphological operation to achieve the goal of smoothing.

One important aspect which needed consideration was the selection of size of the rolling ball filter. If the selected radius would be too small, smoothing of the lung contour would have no effect and will affect the desired segmentation of the lung contour which could lead to lower detection accuracy of the system.

On the contrary, if the selected radius would be too large, then there was a possibility of pleural part inclusion into the real lung contour, adding the interference to the area of lung nodules and affecting the results. So, it was very important to set the appropriate circular radius after large number of validations. Final process images after lung contour corrections are shown in column (e) of Fig. (2) from our user interface, LungCAD.

RESULTS AND DISCUSSION

We evaluated our proposed system on Lung Image Database Consortium (LIDC) (Armato *et al*., 2011). The LIDC is a publicly available database accessible from The Cancer Imaging Archive (TCIA).

We considered the 84 scans (LIDC case 13614193285030022 through LIDC case 13614193285030147, the case numbers were not consecutive but in ascending order) of this dataset. Each CT scan consisted of 150-300 slices where each slice was of size 512*512 and 4096 gray level values in HU. The pixel spacing was 0.78 mm – 1 mm and reconstruction interval varied from 1-3 mm. The results of our proposed system were compared to the manual ground truth which was available by manual contouring of these scans by an expert radiologist. We evaluated our proposed system using the two standard evaluation metrics namely "Overlap measure and Root Mean Square Difference".

The overlap measure of 0.985 was achieved using the proposed methodology. The root mean square difference between the proposed method and ground truth was 0.47 mm. These results showed the superiority of our scheme. Figure (3) presented the results of some example images from our user interface, LungCAD.

In the following section, we have presented sample results and the corresponding discussion. These samples were taken from different scans of LIDC-IDRI dataset used in the study as shown in Fig. (4). The (a), (c), (e) and (g) of Fig. (4) represented the middle slices of input images of different scans while (b), (d), (f) and (h) of Fig. (4) represented the lung segmentation output respectively. We presented two scenarios (i) First scenario in which our algorithm segmented the lung without any artifacts (ii) Second scenario in which our algorithm segmented the lung with some artifacts. The (b) of Fig. (4) represented a segmented lung as an output of our proposed lung segmentation algorithm. Close examination of this output in (b) with the input image in (a) of Fig. (4) revealed that there were almost no artifacts involved in the segmented lung. Similar situation was presented in (f) where no artifacts were found in the segmented lung when compared to the input image in (e) of Fig. (4). The second scenario was presented in (d) of Fig. (4) which represented a segmented lung as an output. Close examination of this output in (d) with the input image in (c) of Fig. (4) revealed that there were some artifacts involved in the segmented lung. On the middle of the right side of left segmented lung, a minute closing could be observed between the two disjoint regions. This artifact was highlighted with the help of a red circle in the figure. Similar situation was presented in (h) where close examination of this output in (h) with the input image in (g) of Fig. (4) revealed the same artifact involved in the segmented lung which was highlighted with the help of a red circle in the figure.

Lung separation: In this study, we have proposed a novel lung separation algorithm of right and left lung lobes which increased the system 'accuracy and removed any artifacts present in lung segmentation. After the extraction of the lung volume, in many cases the gray level thresholding failed to separate the right and left lungs completely and there could exist a junction between them which had to be removed for further processing including smoothing of lung contours.

Specific steps for the lung separation are as follows: (1) Binarization of the connected lung region and the filling algorithm was applied to obtain the lung region binary image. (2) Since the image size is 512 × 512, set the horizontal axis to 256 and set the horizontal scan area for the left and right floating 20, that is between 236 to 276. (3) Fixed horizontal coordinate value from the ordinate value of 0 to start the column scan. Record the first value of the maximum gray value coordinates and then continue to scan and then record the first value of 0 coordinates. Remember that the difference is ΔL and store it in an array. (4) To move within the set abscissa area, the same step (3) was performed and the value was stored into the array (5) Compare with the abscissa value 256 left 20 data changes and then compare the abscissa value 256 on the right side of the 20 data changes, one side of the trend for the first change became smaller. Record its value compared with the right, if still it was the smallest value, it was considered as the separation position otherwise the original axis coordinate 256 wants the array value to be smaller on the side of the transformation. The steps (3), (4) and (5) were performed again until the separation position is found. (6) After separation of left and right lungs, gray level values were converted to 0 according to the location. This process is shown in Fig. (5).

Finally, we compared the results of our proposed sytem with some present studies like Soliman *et al*. (2017) which achieved an overlap measure of 0.98 using 75 manual traced scans but the number of scans were relatively less as compared to our study in which we used 84 scans. Shi *et al*. (2016) also achieved an overlap measure of 0.984 but the number of scans used in the study were relatively low with only 23 scans used. Mansoor *et al*. (2014) used a large set of data consisting of 400 scans to evaluate his system but the overlap measure was low as compared to our system with a value of 0.955. Similarly, Besbes *et al*. (2011) proposed an automatic lung segmentation method and evaluated his system with a large dataset of 123 scans but the achieved overlap measure of the system was 0.94 which was low as compared to our system. These results showed the superiority of our scheme. We developed a fully automated system for lung segmentation. The user interface of the developed system was made in Visual Studio as shown in Figure 6.

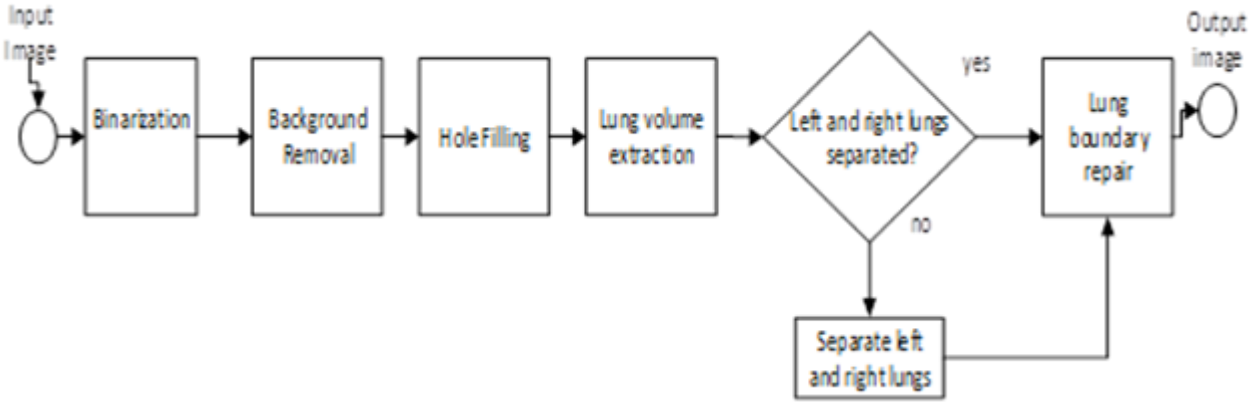
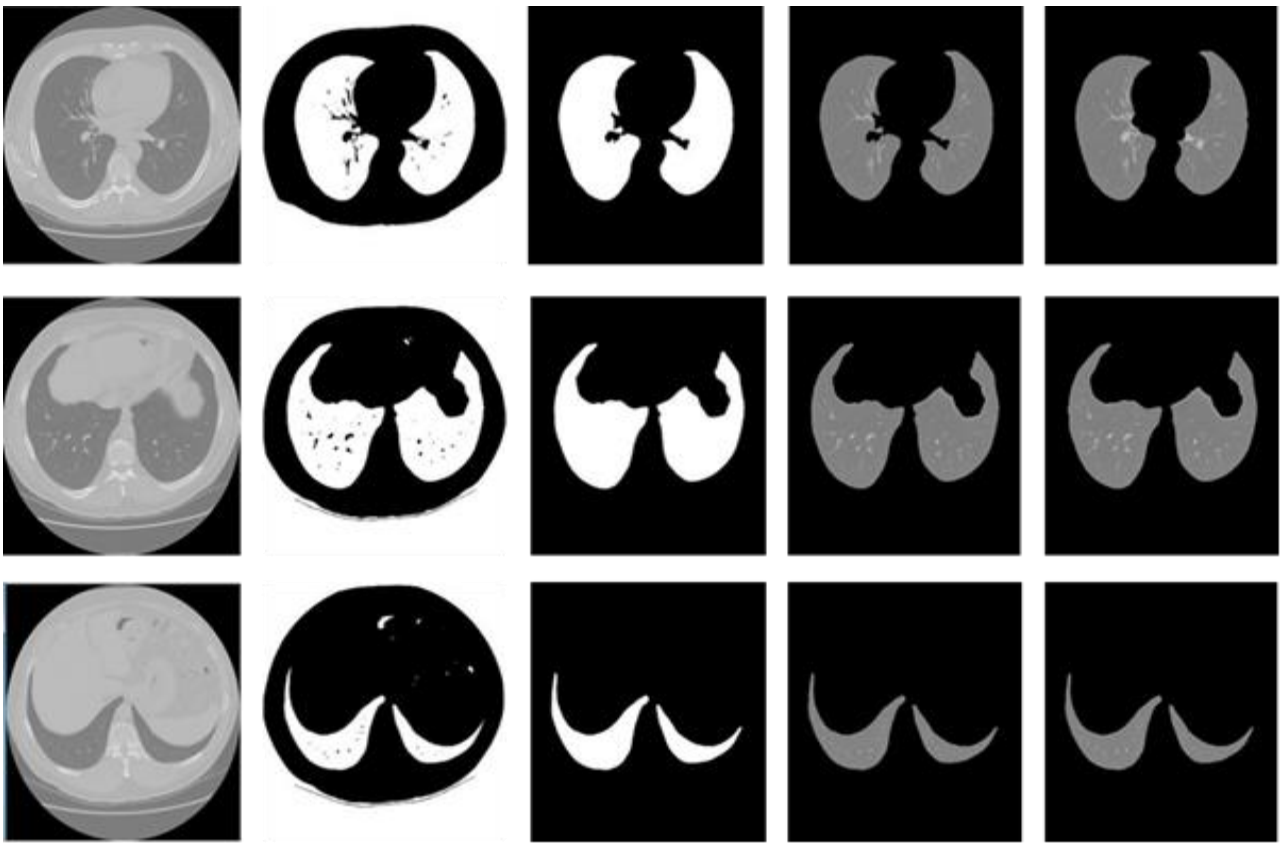
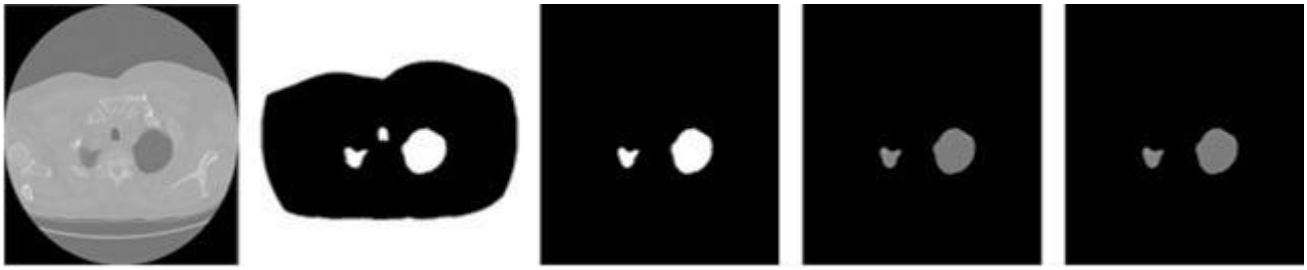


Figure-1: Lung parenchymal segmentation flow chart

(a) (b) (c) (d) (e)

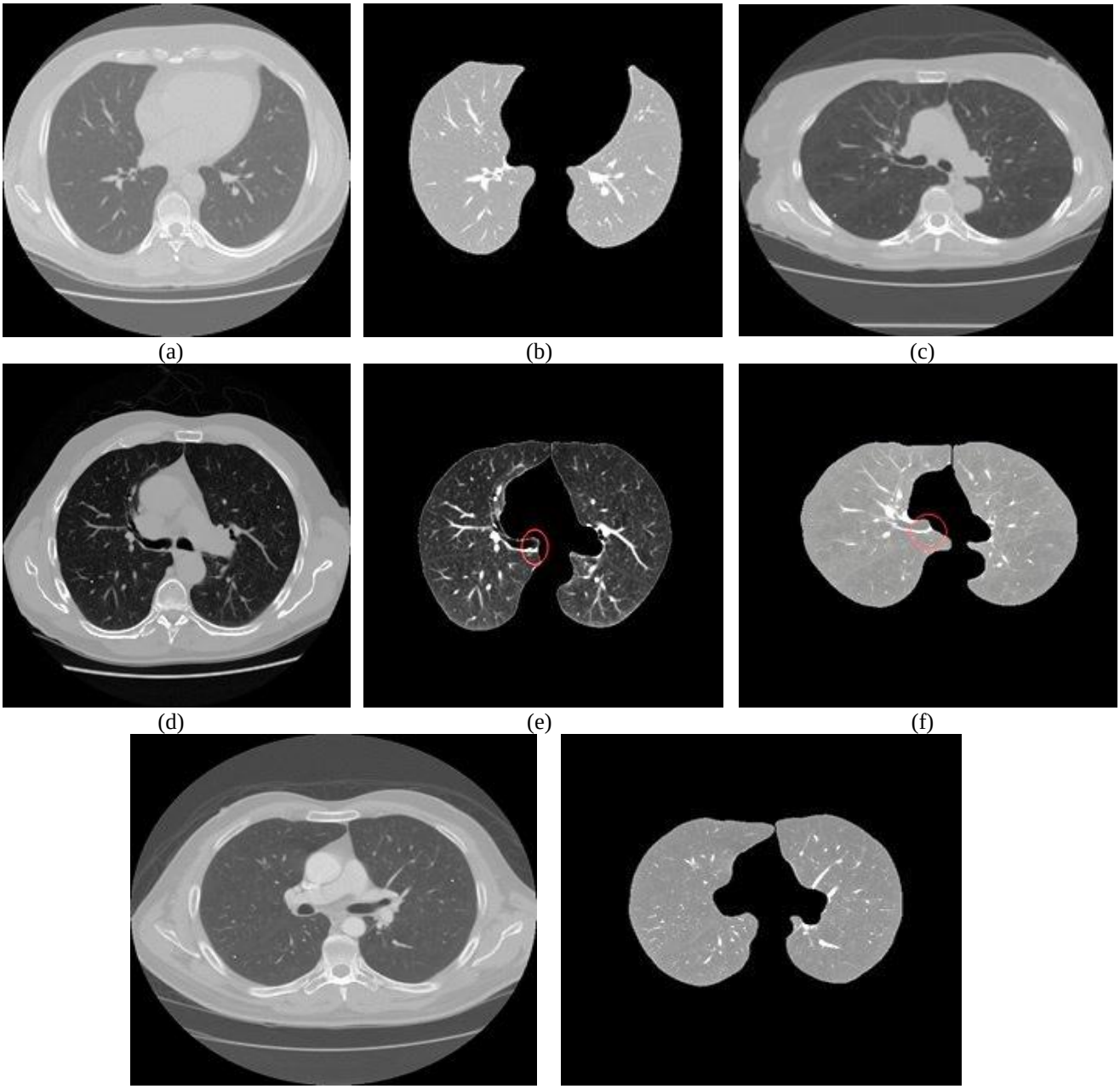
Figure-2: (a) to (e) from left to right presenting input, thresholded, hole filled, lung segmented and contour corrected images, respectively.





(a) (b) (c) (d) (e)

Figure-3: Example images of lung volume segmentation, (a) to (e) from left to right presenting input, thresholded, hole filled, lung segmented and contour corrected images, respectively.

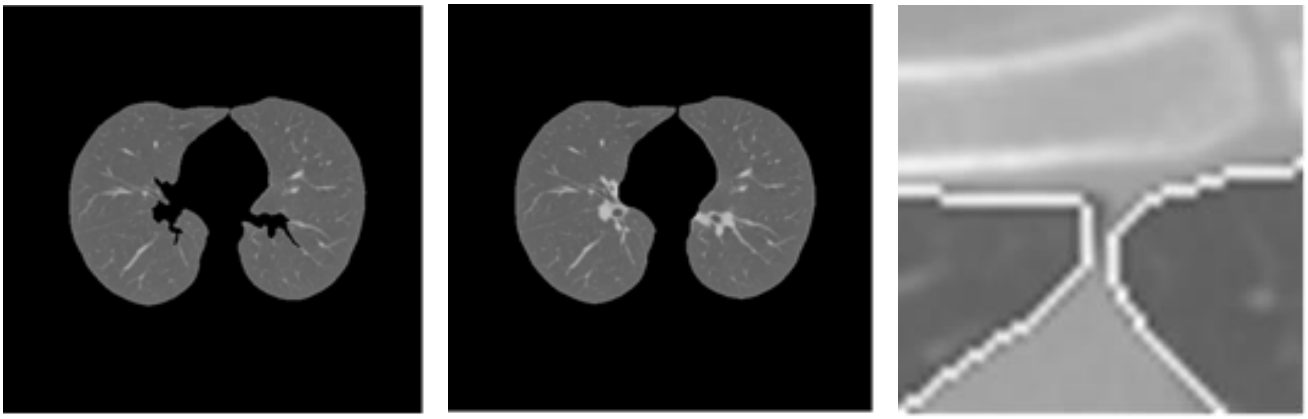


(a) (b) (c)

(d) (e) (f)

(g) (h)

Figure-4: Sample results of our proposed algorithm on middle slices of different scans.



(a) (b) (c)

Figure-5: (a) Represents a parenchymal image (b) represents an image after repair of the lung parenchyma (c) zoomed view of left and right lung contour separation.

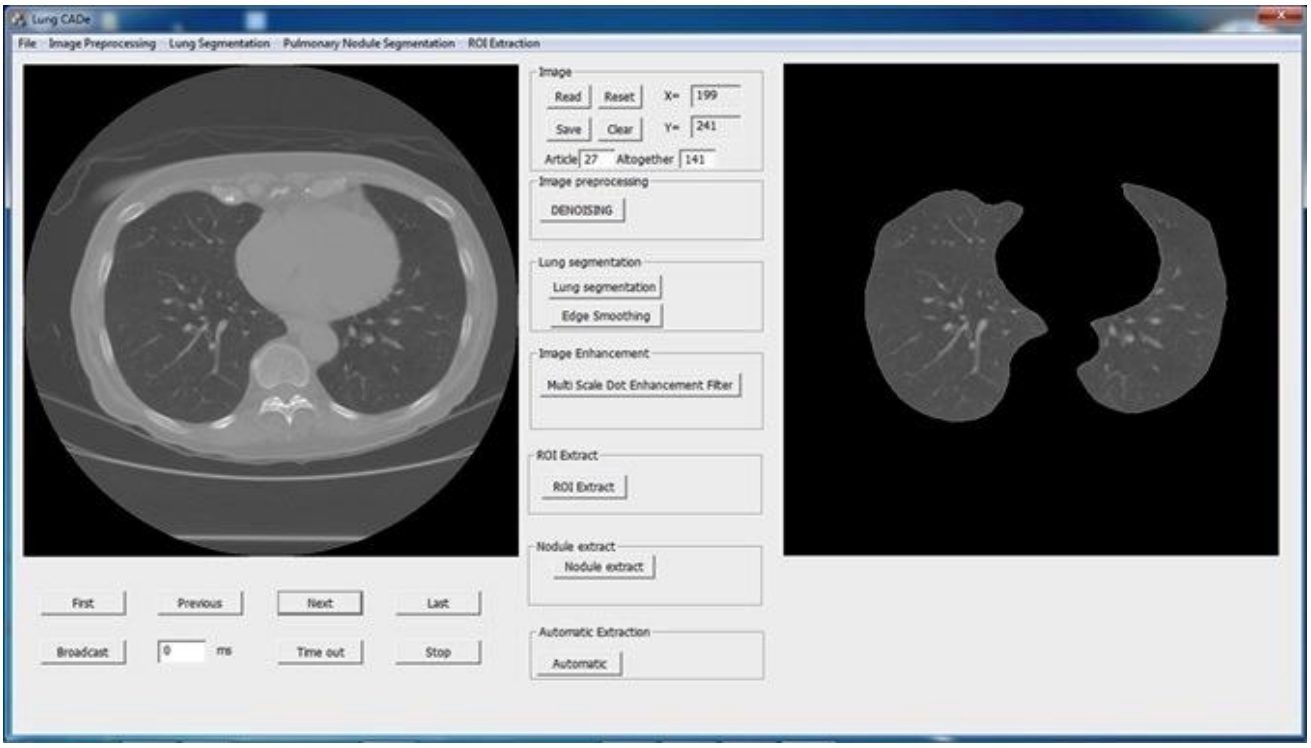


Figure-6: User interface of lung computer aided detection

Conclusion: A novel technique to separate the lungs was introduced which effectively separated the right and left lungs.

REFERENCES

Ali A.M. and A.A. Farag (2008). Automatic lung segmentation of volumetric low-dose CT scans using graph cuts, in advances in visual computing: 4th International Symposium, ISVC 2008, Las Vegas, NV, USA, December 1-3, 2008. Proceedings, Part I, G. Bebis, R. Boyle, B. Parvin, D. Koracin, P. Remagnino, F. Porikli, J. Peters, J. Klosowski, L. Arns, Y.K. Chun, T.M. Rhyne, and L. Monroe, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg. Pp: 258–267.

Annangi P., S. Thiruvenkadam, A. Raja, H. Xu, X. Sun, and L. Mao (2010). Region based active contour method for x-ray lung segmentation using prior shape and low level features. Proceedings of the 7th IEEE Int. Symp. Biomed. Imaging from Nano to Macro (ISBI '10). Pp: 892–895.

Armato S.G., M.L. Giger, C.J. Moran, J.T. Blackburn, K. Doi and H. MacMahon (1999). Computerized

detection of pulmonary nodules on CT scans. Radiographics, 19(5):1303– 11.

Armato S.G., G. McLennan, D. Hawkins (2011). The Lung Image Database Consortium (LIDC) and Image Database Resource Initiative (IDRI): A completed reference database of lung nodules on CT scans. Med. Phys. 38(2): 915–31.

Besbes A. and N. Paragios (2011). Landmark-based segmentation of lungs while handling partial correspondences using sparse graph-based priors. Proc. Int. Symp. Biomed. Imag. Pp: 989–95.

Campadelli P., E. Casiraghi, and D. Artioli (2006). A fully automated method for lung nodule detection from postero-anterior chest radiographs. IEEE Trans. Med. Imaging, 25(12): 1588–1603.

Choi W.J. and T.S. Choi (2014). Automated pulmonary nodule detection based on three-dimensional shape-based feature descriptor. Comput. Meth. Prog. Biomed. 113(1): 37–54.

Choi W.J. and T.S. Choi (2012). Genetic programming-based feature transform and classification for the automatic detection of pulmonary nodules on computed tomography images, Inf. Sci. (Ny). 212: 57–78.

De Nunzio G., E. Tommasi, A. Agrusti (2011). Automatic lung segmentation in CT images with accurate handling of the hilar region. J. Digit. Imaging 24(1): 11–27.

Dehmeshki J., X. Ye, X. Lin, M. Valdivieso and H. Amin (2007). Automated detection of lung nodules in CT images using shape-based genetic algorithm. Comput. Med. Imaging Graph. 31(6): 408–417.

Dai S., K. Lu, J. Dong, Y. Zhang and Y. Chen (2015). A novel approach of lung segmentation on chest CT images using graph cuts. Neurocomputing. 168: 799–807.

Diaz J.M., R.C. Pinon and G. Solano. (2014). Lung cancer classification using genetic algorithm to optimize prediction models. IISA 2014. 5th Int. Conf. Inf. Intel. Sys. App. Pp: 1–6.

El-Baz A., G.M. Beache, G. Gimel'Farb, K. Suzuki, K. Okada, A. Elnakib, A. Soliman and B. Abdollahi (2013). Computer-aided diagnosis systems for lung cancer: Challenges and methodologies. Int. J. Biomed. Imaging. 1-13.

El-Baz A., G. Gimel'farb, R. Falk, M. Abou El-Ghar, T. Holland and T. Shafer (2008). A new stochastic framework for accurate lung segmentation. Proceed. Int. Conf. Med. Img. Comp. Comp. Assist. Interv. (MICCAI '08), New York, NY, USA. Pp: 322–330.

Gao Q., S. Wang, D. Zhao, and J. Liu (2007). Accurate lung segmentation for X-ray CT images. Proceed. 3rd Int. Conf. Nat. Comp. 2: 275– 279.

Howlader N., A.M. Noone, M. Krapcho, D. Miller, K. Bishop, S.F. Altekruse, C.L. Kosary (2016). SEER Cancer Statistics Review, 1975-2013, National Cancer Institute. Bethesda, MD." 2016-02-16]. http://seer. cancer. gov/csr 2016.

Hu S., E.A. Hoffman and J.M. Reinhardt (2001). Automatic lung segmentation for accurate quantitation of volumetric X-ray CT images. IEEE Trans. Med. Img. 20 (6): 490–498.

Korfiatis P., S. Skiadopoulos, P. Sakellaropoulos, C. Kalogeropoulou, and L. Costaridou (2007). Combining 2D wavelet edge highlighting and 3D thresholding for lung segmentation in thin-slice CT. Brit. J. Rad. 80 (960): 996–1005.

Luqman M., M.M. Javed, S. Daud, N. Raheem, J. Ahmad and A.U.H. Khan (2014). Risk factors for lung cancer in the Pakistani population. A. Pac. J. Cancer Prev. 15(7): 3035–9.

Mansoor A., U. Bagci, Z. Xu, B. Foster, K.N. Olivier, J.M. Elinoff, A.F. Suffredini, J.K. Udupa and D.J. Mollura (2014). A Generic Approach to Pathological Lung Segmentation. IEEE Trans. Med. Imaging. 33(12): 2293–310.

Mariotto A.B., K.R. Yabroff, Y. Shao, E.J. Feuer and M.L. Brown, (2011). Projections of the Cost of Cancer Care in the United States: 2010-2020. J. Natl. Cancer Inst. 103(2): 117–28.

Mendonca A.M., J.A. da Silva and A. Campilho (2004). Automatic delimitation of lung fields on chest radiographs. Proceed. Int. Symp. Biomed. Imag. 2: 1287–90.

Moore M.A., P. Attasara, T. Khuhaprema, T.N. Le, T. H. N. Nguyen, P. P. Raingsey, S. Sriamporn, H. Sriplung, P. Srivanatanakul, D.T. Bui, S. Wiangnon and T. Sobue (2010). Cancer epidemiology in mainland South-East Asia - past, present and future. Pac. J. Cancer Prev. 11(2): 67–80.

Rebouças Filho P.P., P.C. Cortez, A.C. da Silva Barros, V.H.C. Albuquerque and J.M.R.S. Tavares (2017). Novel and powerful 3D adaptive crisp active contour method applied in the segmentation of CT lung images, Med. Image Anal. 35: 503–16.

Siegel R.L., K.D. Miller and A. Jemal (2016). Cancer statistics, 2016, CA. Cancer J. Clin. 66(1): 7–30.

Soliman A., F. Khalifa, A. Elnakib, M. Abou El-Ghar, N. Dunlap, B. Wang, G. Gimel'farb, R. Keynton, and A. El-Baz (2017). Accurate Lungs Segmentation on CT Chest Images by Adaptive Appearance-Guided Shape Modeling. IEEE Trans. Med. Imaging. 36(1): 263–76.

Sluimer I., M. Prokop and B. van Ginneken (2005). Toward automated segmentation of the pathological lung in CT. IEEE Trans. Med. Imaging. 24(8): 1025–38.

Suárez-Cuenca J.J., P.G. Tahoces and M. Souto (2009). Application of the iris filter for automatic detection of pulmonary nodules on computed tomography images. Comput. Biol. Med. 39 (10): 921–33.

Sofka M., J. Wetzl and N. Birkbeck (2011). Multi-stage learning for robust lung segmentation in challenging CT volumes. Proceed. Int. Conf. Med. 27th Img. Comp. & Comp. Assist. Interv. Pp: 667–674.

Sun S., C. Bauer and R. Beichel (2012). Automated 3-D segmentation of lungs with lung cancer in CT data using a novel robust active shape model approach. IEEE Trans. Med. Imaging. 31(2): 449–60.

Shi Y., F. Qi and Z. Xue (2008). Segmenting lung fields in serial chest radiographs using both population-based and patient-specific shape statistics. IEEE Trans. Med. Imaging. 27(4): 481–94.

Shi Z., J. Ma, M. Zhao, Y. Liu, Y. Feng, M. Zhang, L. He and K. Suzuki (2016). Many Is Better Than One: An Integration of Multiple Simple Strategies for Accurate Lung Segmentation in CT Images. Biomed Res. Int. 2016: 1–13.

Valente I.R.S., P.C. Cortez, E.C. Neto, J.M. Soares, V.H.C. de Albuquerque and J. M.R.S. Tavares (2015). Automatic 3D pulmonary nodule detection in CT images: A survey, Comput. Methods Programs Biomed.124:91–107.

Van Rikxoort E., B. de Hoop and M. Viergever (2009). Automatic lung segmentation from thoracic computed tomography scans using a hybrid approach with error detection. Med. Phys. 36(7): 2934.

Ye X., X. Lin and J. Dehmeshki (2009). Shape based computer-aided detection of lung nodules in thoracic CT images. IEEE Trans. Biomed. Eng. 56(7): 1810–20.

Yim Y., H. Hong and Y.G. Shin (2005). Hybrid lung segmentation in chest CT images for computer-aided diagnosis. 7th Int. W. Ent. Net. Com. H. Indus. Pp: 378–83.