International Journal of Bio-Science and Bio-Technology
Vol.8, No.5 (2016), pp. 105-114
http://dx.doi.org/10.14257/ijbsbt.2016.8.5.10

Adaptive Thresholding Technique for Segmentation and Juxtapleural Nodules Inclusion in Lung Segments

Muhammad Zia ur Rehman, Syed Omer Gilani, Syed Irtiza Ali Shah, Mohsin Jamil, Irfanullah and Shahid Ikramullah Butt

Department of Robotics and intelligent Machine Engineering, School of Mechanical and Manufacturing Engineering, National University of Sciences and Technology (NUST), Islamabad 44000, Pakistan.
ziaurrehman@smme.edu.pk

Abstract

Early diagnosis of lung cancer plays crucial role in the improvement of patients' chances of survival. Computer aided detection (CAD) system has been a groundbreaking step in the timely diagnosis and identification of potential nodules (lesions). CAD system starts detection process by extracting lung regions from CT scan images, this step narrows down the region for detection. Hence saving the time consumption and reducing false positives outside the lung regions that results in the improvement of specificity of system. Proper lung segmentation significantly increases the performance of CAD systems. Different algorithms are presented by various researchers to improve segmentation results. An intensity based approach is presented in this paper for the segmentation of parenchyma and the goal is to achieve reasonable segmentation results in less time. Algorithm used in this paper is based on the Intensity based thresholding which is the fastest method for image segmentation. Images used in this research to analyze algorithm's result are taken from Lung Image Database Consortium (LIDC). Twenty random cases were picked, each having different number of slices (128 to 300). Algorithm is implemented using MatlabR2014 and a system with processor of 2.6 GHz and RAM of 4 GB. Total time taken for a single case of 128 images was 6.3 seconds and hence with an average of 49 milli sec/slice.

***Keywords**: Lung cancer, Juxta-pleural nodules, Computer aided detection (CAD), segmentation, Intensity based thresholding*

1. Introduction

Lung disease is a standout amongst the most serious growths worldwide and the aggregate deaths rate due to this specific disease is more prominent than that of other diseases like colorectal and prostate diseases. [1-2]. Early discovery and treatment of the lung potential nodules can enhance the survival rate of the patients [3]. CT (computed tomography) scan is considered as an essential instrument in the early diagnosis or detection of lesions (potential nodules) but yet the translation of the vast or broad measurement of thoracic tomographic pictures is extremely difficult chore for the radiologists [2]. According to the study, the identification rate of potential nodules utilizing computed tomography is much higher than simple radiography (the difference is 2-10 times) [4]. But as mentioned earlier, it is very difficult task for the radiologist to analyze the large number of patient’s cases and it specifically affects the workload of the doctor which increases the chances of false detection of the disease and hence increase the chances of error [5]. This leads to the requirements of a computational system that can automatically detects the lesions and can explain the different consequences and their attributes Therefore, the development of a CAD (computer aided detection) system is

ISSN: 2233-7849 IJBSBT

considered as better system that helps the radiologists in the analysis of CT of the lung nodules [2].

In a lung CAD framework, better lung segmentation [6] can further enhance the ability of the whole CAD framework and reduce misdiagnosis. Along these lines, the lung parenchyma division is a key system of a CAD framework for lung ailments and which will influence the precision of the entire lung CAD framework [7]. The two main developed such frameworks are the computer aided diagnosis system (CADx) and the purpose of it is to determine the nodule's qualities and behaviors *etc*, while the second such developed system is CADe (detection system) which is aimed to determine and locate the tumors or potential nodules utilizing image processing algorithms [5]. The presented work in this paper deals with a portion of CADe system (segmentation of lungs parenchyma) and the aims of such systems are to help in timely identification of lesions (if any) with high level of accuracy in detection and to overcome or decrease the workload of doctors in evaluation of patient's data [8]. As mentioned earlier, CADe systems are very important in a way it assist doctors and radiologists and some of the specification considered for such systems [5] are high sensitivity, computationally efficient, less false positive occurrence, cost effective, highly automated and software security assurance [9, 10] *etc*.

2. Implementation

2.1. Generalized Segmentation Algorithms

In last few years, researchers from different countries implemented different segmentation algorithms and put forwarded a progression of lung segmentation techniques [11, 12]. These algorithms can be broadly classified in to the threshold, region growing and pattern classification methods [2]. The first one which is threshold method [13] is computationally efficient and simple but some of the limitations of utilizing this basic technique are, it is very difficult to discount the tracheal/bronchial portion successfully and to include the juxtapleural nodules as a lung segment in the segmented images [2]. This research work proposes a technique that is utilized to specifically deal this problem of threshold methods. The second segmentation technique which is a semi-automatic is that of region growing technique [14, 15]. Despite the fact that this technique can overcome some of the problems of threshold methods but yet it has some limitations like failure to separation of lung lobs and also the manual selection *i.e.* the selection of the seed points. This technique is delicate to the choice of parameters concerning rules for it, which may influence the strength. The last technique is that of pattern classification method [16]. This technique obliges an extensive number of preparing sample information, furthermore, needs to concentrate the features, so this technique is not as much computationally efficient and it takes time. Hence in clinical applications, it fails to fulfill the CAD system requirements in real-time processing [2].

2.2. Proposed Algorithm

Raw images data was collected from LIDC (Lung Image Database Consortium) database. This is publically available database having large collection of up-to-date images data. Its purpose is to provide data for the development and evaluation of CAD system. The first step in this algorithm is to sequence the images of a complete case according to their metadata instance number. This step helps to correctly organizing images for further processing. After sequencing, these images of a single case are read by applying specific matlab commands which reads a complete folder sequence wise and then preprocessing is performed which purpose is to enhance the quality of raw CT scan images and to remove noise or defects (if any). Generally, this step helps in improving the sensitivity and accuracy of the next coming steps in the algorithm. As the purpose of this research paper is to achieve high speed of processing, so the contrast stre*tc*hing is specifically performed because it is a computationally efficient technique. The next and main step of this algorithm is segmentation of lung region which is performed by using threshold intensity method. Along with it, morphological opening and closing operations are also performed in order to include the juxtapleural nodules as part of the lung parenchyma.

After the segmentation, as mentioned before, these all steps are performed on sequenced images, so the finally segmented images are also written in a sequence to a new folder, which can be used for further processing or viewing and analyzing purpose. Figure 1 shows the block diagram of the algorithm and it mentions the step by step process**.**

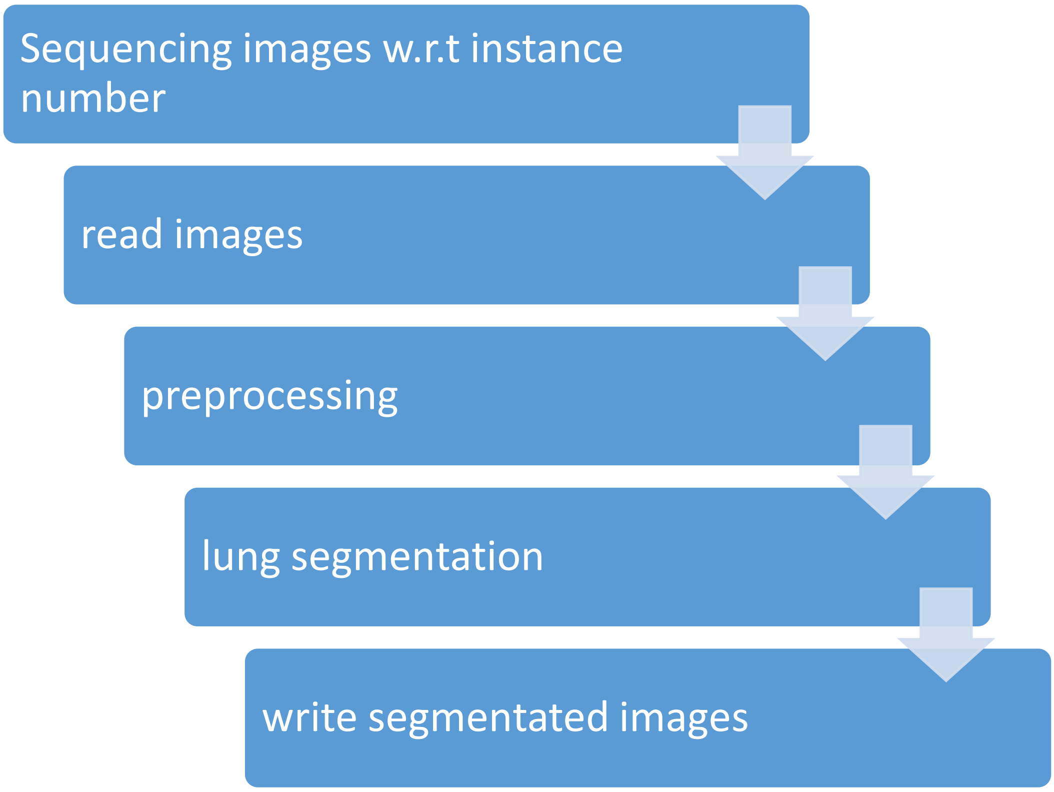

Figure 1. Show the Block Diagram of Proposed Algorithm which first reads the sequence of images of a single case according to their metadata instance number. After reading these images in matlab, the preprocessing and segmentation techniques are applied and the final segmented images are stored to a new folder which can be used for further processing or viewing and analyzing purpose

2.3. Preprocessing and Segmentation

Preprocessing of the images is an essential step that is generally utilized to remove defects and noise *etc* which are captured at the time of acquisition of images. It is used in order to enhance the contrast and quality of the images, which results in the accuracy and exactness of processing algorithms that are used later in the system [8]. Scholar around the world uses different approaches and techniques to tackle this step. Some of these methods used in the pre-processing step are [5]: Median Filtering, Auto-improvement, Wiener Filter, Smoothing filters and Noise Correction, Improvement Filter, Fast Fourier Transform, Antigeometric Diffusion, Wavelet transform and Erosion Filter *etc*.

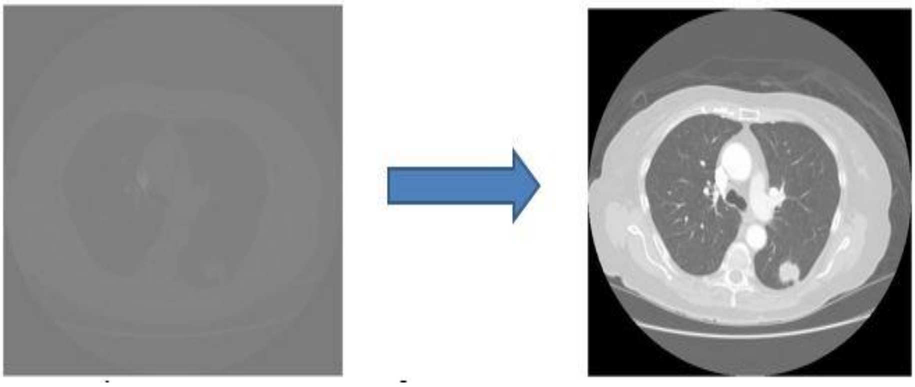

Figure 2. It shows the result of preprocessing step. The figure on the left side is the raw ct scan images in which nothing is visible. After apply the contrast str*etc*hing technique, the raw image is enhanced (as shown in right side image) and now different portions like, black corners, empty space inside ct scan, lungs parenchyma and bedsheet *etc* are visible. This step helps in accurately performance of further steps

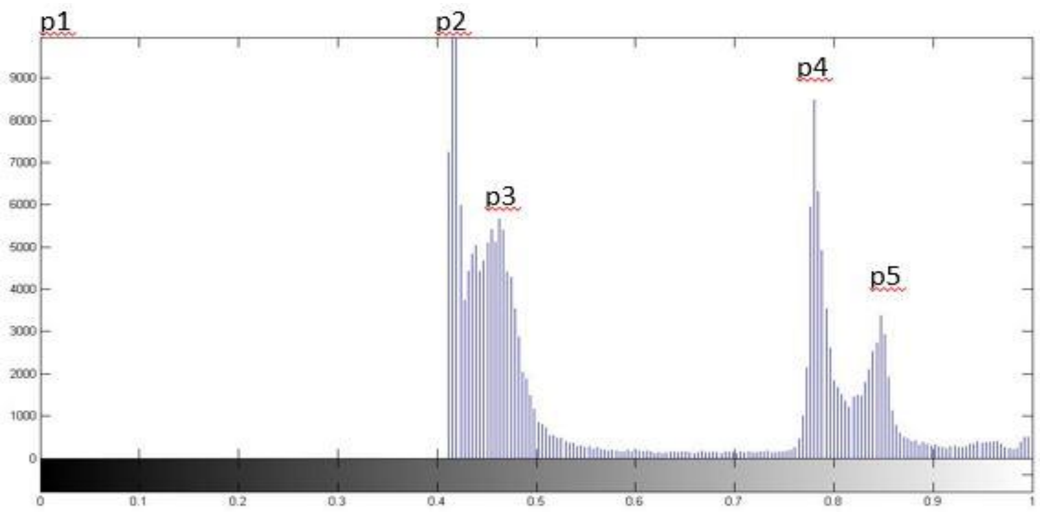

Figure 3. Histogram of a preprocessed image. This shows some intensity peaks in images which represent the black corner pixels which are outside of cylinder, inside airspace and air inside lung pipes, lungs parenchyma, soft tissue, muscles and bones *etc*

As mentioned earlier, for the timing efficiency purpose, contrast str*etc*hing technique is applied in preprocessing step and the resulting histogram of the image is shown in figure 2. The histogram of preprocessed images shows some peaks and valleys which represents the intensities of different portions in ct chest region. The peak p1 represents the black corner pixels of ct image which are outside of cylindrical region. The peak p2 represents the pixels of both space inside the cylindrical region and air inside the lungs and bronchioles *etc*. the peak p3 represents the pixel that contains the lung parenchyma. The first valley inside peak p3 and p4 shows the pixels of bronchioles and blood vessels *etc* while the second valley after the peak p5 shows the pixels of ribs, bones and hard tissues *etc*. the last peak p6 (at the end) may be the pixels for metallic or pacemaker leads. The area of focus which is lung parenchyma lies in the first valley which is between peak p3 and p4. So the thresholding value is chosen as the average of p2 and p4 which represents the lung parenchyma pixels.

The adaptive threshold was computed using fallowing formulas

$$Threshould\ value\ (T) = \frac{p2 + p4}{2}$$

Assuming f(x,y) is a pixel intensity value, every pixel is categorized into background and foreground according to fallowing decision rule

$$f(x,y) = \begin{cases} \geq T & f(x,y) = 1 \\ < T & f(x,y) = 0 \end{cases}$$

Here f(x,y)=1 represent the background and f(x,y)=0 represent the foreground pixels

Table 1. Measured Peak Values for all Scans

Scan Images	Peak p2	Peak p3	Peak p4	Peak p5
Case 1	0.422	0.469	0.785	0.858
Case 2	0.427	0.470	0.787	0.863
Case 3	0.426	0.469	0.787	0.861
Case 4	0.423	0.469	0.786	0.859
Case 5	0.422	0.468	0.784	0.858
Case 6	0.422	0.469	0.783	0.860
Case 7	0.424	0.470	0.785	0.859
Case 8	0.423	0.471	0.783	0.859
Case 9	0.425	0.469	0.784	0.862
Case 10	0.424	0.470	0.787	0.861
Case 11	0.424	0.470	0.784	0.863
Case 12	0.427	0.468	0.786	0.864
Case 13	0.422	0.467	0.783	0.859
Case 14	0.426	0.470	0.787	0.862
Case 15	0.423	0.469	0.785	0.859
Case 16	0.423	0.469	0.783	0.861
Case 17	0.425	0.471	0.784	0.861
Case 18	0.422	0.468	0.786	0.858
Case 19	0.424	0.467	0.785	0.862
Case 20	0.422	0.469	0.783	0.862
Mean $\pm$ Deviation	0.4238 $\pm$ 0.0017	0.469 $\pm$ 0.0011	0.7848 $\pm$ 0.0014	0.8605 $\pm$ 0.0018

The results of this thresholding along with some morphological opening and closing operations are shown in figure 6 to 8. Figure 2, 3, 4 and 5 shows the results of different steps which includes the preprocessing, thresholding, application of opening and closing and filling operations on binaries images after thresholding and resulting images.

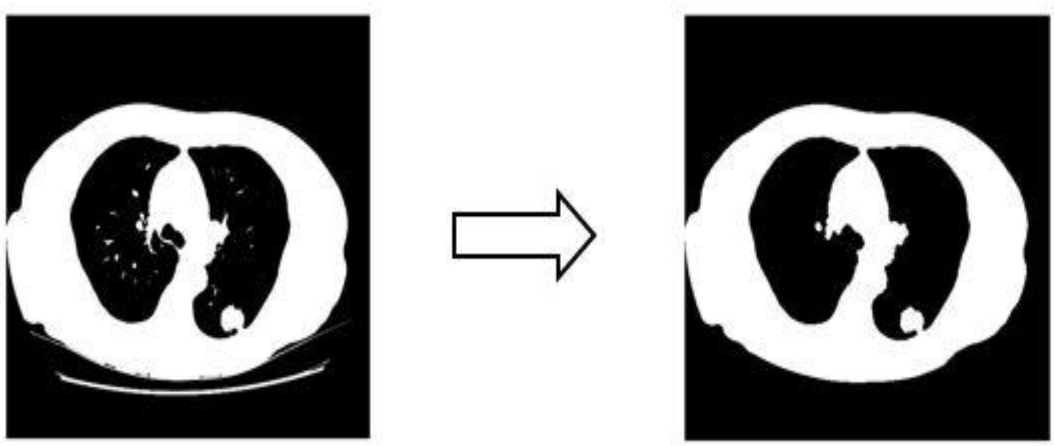

Figure 4. The right side figure shows the result of intensity thresholding while the left sided figure is the result of applying morphological operation on resulting binary images

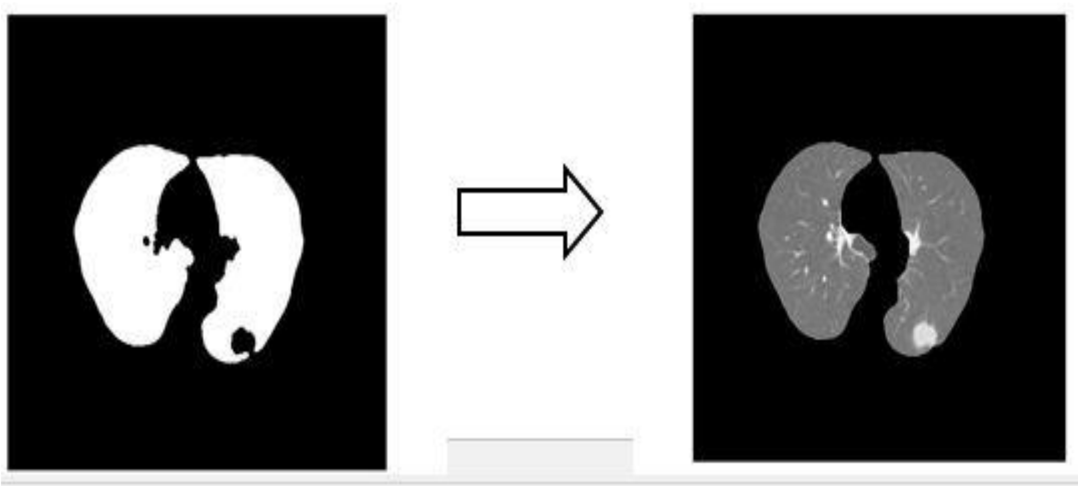

Figure 5. Shows the contour refinement process on binary images and application of morphological closing in order to include the juxtapleural nodules in lung parenchyma

3. Results and Discussions

In this work, the images of twenty cases were downloaded from Lung Image Database Consortium (LIDC). In different cases, the number of images varies from 128 to 300. All of the implementations were performed on MATLAB 2014 (The MathWorks®, Inc., Natick, MA, USA) on the system having 2.6 GHz processor with a RAM of 4 GB. The algorithm generated the 3D segment of a single case in 6.3 seconds with a total of 128 images hence with an average of 45 mili sec/image.

The main purpose of this work was to reduce the time for segmentation. Because it's a portion of a CAD system and next steps are to detect potential nodules and removing false positives and statistical analysis which also takes some time. So the literature review was done in order to choose a simple and time efficient technique which can give better approximate between the quality and time efficiency. In order to fallow this goal a better technique is presented in this paper and the resulting images of lung Parenchyma of both right and left lung are shown in figures 6 and 7. The figures are clearly indicative of the analysis on the basis of new concept provided in the proposed work and will certainly improve the performance and thus detection mechanism to curve the menace of the deceases.

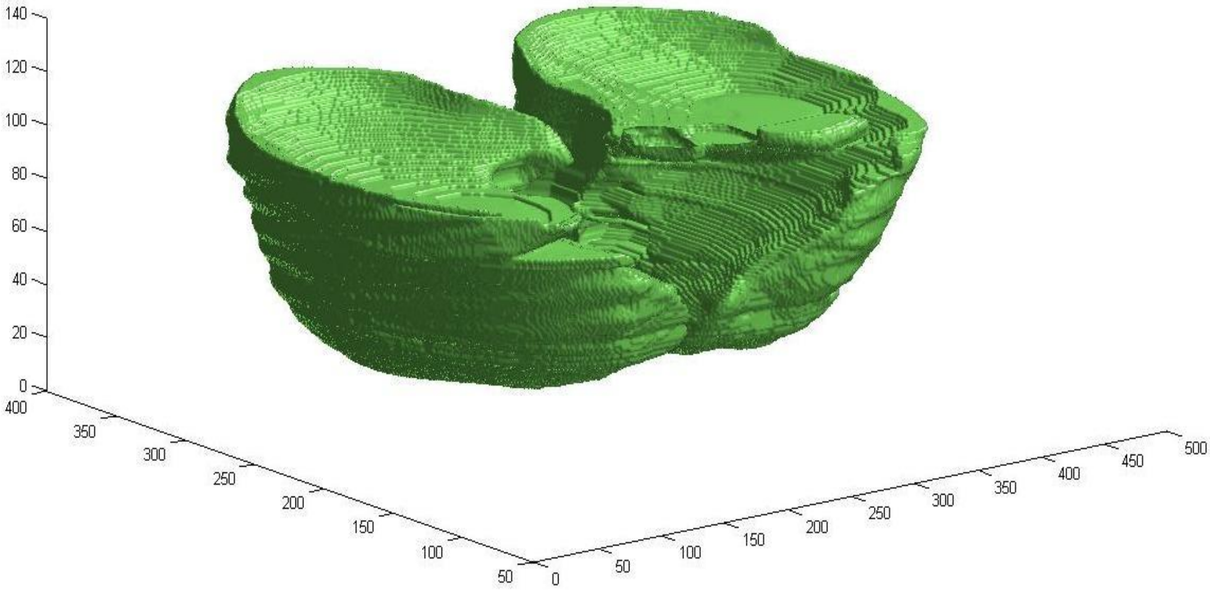

Figure 6. The Side View of 3D Lung Parenchyma. It shows the empty space between the left and right lungs where the other tissues like heart *etc* lies.

Figure 7. The Frontal View of 3D Lung Segmentation and it shows the expression of ribs on lungs

The Figure 7 is indicative of 3D view of lung segmentation along with ribs effect on the lungs to give better image for analysis purpose.

4. Conclusion

In this research, an algorithm for lung segmentation is implemented and it showed the desired results. Different image processing methods have been investigated and analysed for the purpose of implementation and validation of proposed algorithm in the work. Prominently among them is intensity based thresholding, which was selected due to its adaptability and performance. Selection criteria was based on the time consumption and segmentation results. As CT scan case consists of large number of slices, segmentation on each slice and then detection over burden the system. Detection process significantly slows down if segmentation step consumes much time. Therefore, a tradeoffs have to be done between segmentation results and time consumption. Lung segmentation based on intensity thresholding gave good results in less time. Problems in lungs segmentation like exclusion of pleura connected nodules from extracted lungs were tackled using morphological opening and closing operations. This segmentation process when implements on single case with 128 slices took 6.3 seconds, consumed 45 milliseconds/slice. Algorithm is implemented on Matlab 2014, System's processor 2.6 GHz and RAM 4GB. Test is performed on twenty cases taken from LIDC database. Algorithm presented in this research shows promising results in the segmentation of lung regions in lesser time. This segmentation step when used in CAD systems can significantly improves its performance in the detection and identification of potential nodules.

References

[1] G. D. Rubin, “Pulmonary Nodules on Multi–Detector Row CT Scans: Performance Comparison of Radiologists and Computer-aided Detection 1”, Radiology, vol. 234, no. 1, **(2005)**, pp. 274-283.

[2] Y. G. S. Wei and J. J. Li, “A fully automatic method for lung parenchyma segmentation and repairing”, Journal of digital imaging, vol. 26, no. 3, **(2013)**, pp. 483-495.

[3] D. T. Lin, C. R. Yan and W. T. Chen, “Autonomous detection of pulmonary nodules on CT images with a neural network-based fuzzy system”, Computerized Medical Imaging and Graphics, vol. 29, no. 6, **(2005)**, pp. 447-458.

[4] K. Awai, “Pulmonary Nodules at Chest CT: Effect of Computer-aided Diagnosis on Radiologists’ Detection Performance 1”, Radiology, no. 230, no. 2, **(2004)**, pp. 347-352.

[5] M. Firmino, “Computer-aided detection system for lung cancer in computed tomography scans: Review and future prospects”, Biomed Eng Online, vol. 13, **(2014)**, pp. 1-16.

[6] Q. Li, “Recent progress in computer-aided diagnosis of lung nodules on thin-section CT”, Computerized Medical Imaging and Graphics, vol. 31, no. 4, **(2007)**, pp. 248-257.

[7] S. Hu, E. Hoffman and J.M. Reinhardt, “Automatic lung segmentation for accurate quantitation of volumetric X-ray CT images”, Medical Imaging, IEEE Transactions, vol. 20, no. 6, **(2001)**, pp. 490-498.

[8] M. Gomathi and P. Thangaraj, “Computer Aided Medical Diagnosis System for Detection of Lung Cancer Nodules: A Survey”, International Journal of Computational Intelligence Research, vol. 5, no. 4, **(2009)**.

[9] B. van Ginneken, C. M. Schaefer-Prokop and M. Prokop, “Computer-aided diagnosis: how to move from the laboratory to the clinic”, Radiology, vol. 261, no. 3, **(2011)**, pp. 719-732.

[10] A. E. Baz, “Computer-aided diagnosis systems for lung cancer: challenges and methodologies”, International journal of biomedical imaging, 2013, **(2013)**.

[11] G. D. Nunzio, “Automatic lung segmentation in CT images with accurate handling of the hilar region”, Journal of digital imaging, vol. 24, no. 1, **(2011)**, pp. 11-27.

[12] X. Bian and Y. Wu, “A method of careful lung segmentation based on CT images”, Journal of Chongqing University of Posts and Telecommunications (Natural Science Edition), vol. 22, no. 5, **(2010)**, pp. 665-668.

[13] J. K. Leader, “Automated lung segmentation in X-ray computed tomography: development and evaluation of a heuristic threshold-based scheme1”, Academic radiology, vol. 10, no. 11, **(2003)**, pp. 1224-1236.

[14] L. Hedlund, "Two methods for isolating the lung area of a CT scan for density information", Radiology, vol. 144, no. 2, **(1982)**, pp. 353-357.
[15] S. G. Armato and W. F. Sensakovic, "Automated lung segmentation for thoracic CT: Impact on computer-aided diagnosis1", Academic Radiology, vol. 11, no. 9, **(2004)**, pp. 1011-1021.
[16] M. F. McNitt-Gray, "Pattern classification approach to segmentation of digital chest radiographs and chest CT image slices. in Medical Imaging 1994", International Society for Optics and Photonics, **(1994)**.