Available online at www.sciencedirect.com

ScienceDirect

Procedia Technology 24 (2016) 957 – 963

Procedia Technology

International Conference on Emerging Trends in Engineering, Science and Technology (ICETEST - 2015)

Multilevel Thresholding Based Segmentation and Feature Extraction for Pulmonary Nodule Detection

Jibi John[a]* , M G Mini[b]

[a]*Model Engineering College, Thrikkakkara, Kochi, 682021, Kerala, India*
[b]*College of Engineering, Cherthala, ,Cherthala, 688541,Kerala, India*

Abstract

The identification of pulmonary nodules in humans has always been a vexing problem in the field of medical electronics. In this paper, an approach is proposed for pulmonary nodule segmentation and feature extraction using multilevel thresholding. The suitably extracted features can go a long way in the efficient detection of pulmonary nodules, which in turn can improve the chances for successful classification of nodules. The proposed segmentation with three level thresholding along with the features extracted can be incorporated to any suitable classification architecture to detect pulmonary nodules with better accuracy.

Peer-review under responsibility of the organizing committee of ICETEST – 2015

*Keywords:*Computed Tomography(CT) scans; Lung Cancer; Medical images; Multilevel thresholding; Segmentation; Pulmonary Nodule Detection

1. Introduction

The accurate investigation and diagnosis is a challenging task in medical field in managing various diseases. Early diagnosis can greatly improve the effectiveness of the treatment and thus increase the 5 year survival rate of the lung cancer patients from 15% to 65%-80% [1]. Medical imaging provides a non invasive method for detection and characterization of the pulmonary abnormalities. The pulmonary nodule is a frequent finding on Computed Tomography (CT) scans and it is of high clinical importance as it may prove to be an early manifestation of lung

* Corresponding author. Tel.: +91-984-646-6044;
*E-mail address:*jibi@mec.ac.in

Peer-review under responsibility of the organizing committee of ICETEST – 2015
doi:10.1016/j.protcy.2016.05.209

cancer [2]. The problem of efficient detection and classification of cancerous pulmonary nodules is of great significance in improving the general health of any society. CT provides better spatial and contrast resolution and thus CT image based study of pulmonary nodules has been a common approach practiced over the years in the field of lung cancer detection.

An automated lung segmentation method based on thresholding with iteratively found optimal threshold was published by Hu et al. in [3].This method was further used as a preprocessing step for lung lobe segmentation by Van Rikxoort [4]. Armato et al. [5] used grey level thresholding with rolling ball filter to include juxtapleural nodules. Wei et al. [6] selected a threshold to extract lung region based on histogram analysis. N Homma et al. [7] proposed a lung area extraction technique for pulmonary nodule detection using Active Contour Model. Another approach for lung area extraction and nodule detection proposed by S.Sousa et al. [8] used multiple region growing and thresholds was selected based on histogram analysis of the image. A fully automated approach for lung segmentation with Active Shape Model was proposed by S. Sun. et al in [9].

In this research paper, an approach is proposed for feature extraction of prospective pulmonary nodules and initial candidate nodule identification for consequent classification. This approach makes use of CT images from publically available lung imaging database for cancer studies. After employing image enhancement techniques to the extracted lung image, an intermediate thresholding followed by a micro level thresholding is performed. In addition, various shape and intensity based features are extracted and a preliminary rule based filtering is employed to reduce the false nodules.

2. General Architecture of Pulmonary Nodule Detection System

Computer Aided Detection system for pulmonary nodules consists of various subsystems like image acquisition, preprocessing, segmentation, nodule detection and false positive elimination [10]. Image acquisition subsystem is responsible for obtaining suitable medical images and image preprocessing techniques are generally applied to improve the quality of the image for display and analysis.

2.1. Pre-processing of images

Pre-processing aims to improve the image data to enhance certain features of the image, which in turn enhances the accuracy and precision of detection algorithm. Preprocessing techniques include removal of defects such as noise, lack of contrast and removal of the unwanted parts like labels, images of other organs. In general, noise in an image has higher frequency spectrum. So low pass filtering may be effective in removing the noise. The impulse noise can be removed by simple median filtering. Histogram equalization may improve the contrast of the image. Selective enhancement filters and suppression filters also can be employed to enhance the image for further processing.

2.2. Lung Segmentation

CT scan provides detailed views of the body's soft tissues, including blood vessels, muscle tissue, and organs. So Extraction of lung area from CT image reduces search space for nodules. Lung segmentation is a challenging problem due to in-homogeneities in the structure, different available scanners and scanning protocols. Size, shape and texture of lungs image varies in different CT slices. Contrast between lungs and neighbor tissue is the basis for the most of the lung segmentation methods. Exclusion of the nodules connected to the boarder (juxta-pleural nodule) is a crucial issue at this stage.

2.3. Lung Nodule Detection

This stage aims at determining the location of pulmonary nodules. Since nodules represent the initial radiographic finding of lung cancer, early diagnosis can increase the patient's survival rate. Pulmonary nodule detection is always a difficult task due to its varying size, contrast and location within an area of complex anatomy. The nodules can be classified based on the location as well-circumscribed nodules, juxta-pleural nodules, nodule with pleural tail and juxta-vascular nodules and those based on intensity as solid and non solid nodules. Nodule detection generally consists of two stages: (i) selection of initial candidate nodules and (ii) partial elimination of the false positive nodules (FPNs).

2.4. Feature Extraction and Non-Nodule Filtering

The objective of this stage is to extract features that help to remove highly non-nodule candidates segmented. A pulmonary nodule is defined as approximately round opacity, at least moderately well margin and not greater than 3 cm in the maximum diameter [2]. Since pulmonary nodules are appeared as circular or ellipsoid in nature, features selected for initial filtering stage can be shape based. After the preliminary filtering stage a feature-based classifier also can be used for further false positive nodule elimination.

3. Proposed Method

A block diagram of the proposed work is given in Fig (1). Preprocessed CT image is given to the Thorax Extraction block. Lung area is then extracted from this thorax region. Pulmonary nodules are segmented from lung image and various shape and intensity based features are extracted. A rule based filtering is employed to reduce a large number of non-nodule candidates.

3.1. . Pre-processing of CT image and Global Thresholding based Lung Area Segmentation

In the proposed method, CT image is initially rescaled without losing the information. A smoothing operation is then performed using a 2D median filter of mask size 3x3 for removing the noise.

Removal of unwanted parts is carried out in two steps [5]. In the first step all the artifacts external to the patient's body are removed by extracting the Thorax region using a Thorax Mask. In the second step, images of other organs are removed by extracting Lungs region from the Thorax region using a Lung Mask. In order to include images with different contrast from various databases a contrast adjustment is performed.

The Thorax Mask is prepared by Thresholding and morphological operations. For thresholding, density difference of the anatomical structure is made use of. Global Thresholding technique is employed and an optimum threshold is obtained iteratively [3] for each slice since the size of the object, background and contrast vary with each slice. The mean intensity of the image is set as the initial threshold. The segmentation threshold selection is described as an iterative procedure. Let T^i be the threshold in the i^{th} step, μ_o and μ_b be the mean intensity value of the object and background in the i^{th} step and T^{i+1} be the threshold for the $i+1^{th}$ step then

$$T^{i+1} = \frac{1}{2}\left(\mu_o + \mu_b\right) \quad (1)$$

The iterative steps continue until the threshold converges to a defined small difference which is fixed by trial and error method in this case. The final threshold can be used for thresholding the CT image.

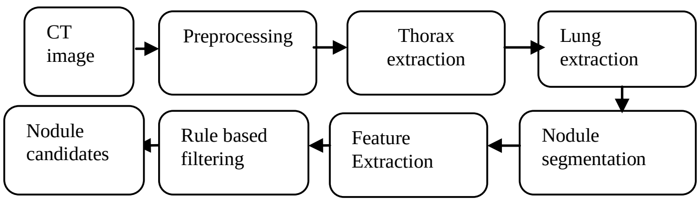


Fig. 1. Block diagram of the proposed Nodule candidate detection system.

Morphological filling operation is employed to get the Thorax Mask and small objects are removed by morphological opening. The Lung Mask is prepared by performing image subtraction of Thorax Mask and the thresholded image. After that morphological filling of the resultant image is performed. Morphological closing using circular structuring element is performed to avoid the exclusion of nodules connected to the boarder. By employing masking technique using Thorax Mask and Lung Mask with CT image Thorax region and lung region is extracted. Image is then enhanced with unsharp masking technique which helps the detection of pulmonary nodules with better accuracy.

3.2. Intermediate Thresholding

After applying a two step processing for lung segmentation, an additional two step Thresholding technique also been proposed in this paper for successful detection of pulmonary nodules. A single step intensity based thresholding may not include different density nodules. After observing intensity distribution within the lung region, a contrast enhancement has performed. In the proposed Intermediate Thresholding, the threshold is fixed with the knowledge of maximum and mean intensity within the lung region. The resultant binary image is used as the mask for obtaining Intermediate level nodule locations.

3.3. Micro- level Thresholding

By considering grey level information within the intermediate nodule locations a Micro-level threshold is selected for the next step thresholding for the enhanced image. The binary image thus obtained acts as the mask for obtaining initial nodule candidates.

3.4. Extracted Features

The following eleven morphological and intensity features are extracted using connected component analysis for the purpose of study.

- Area: Area is defined as the number of pixels in the nodule region. Since the size of the nodule is considered as 3mm to 3cm this feature is a relevant one.
- Solidity: The feature solidity is computed as ratio of the Area and Area of the smallest convex polygon that can contain the nodule region. This gives a measure of compactness.
- Eccentricity: Eccentricity is the ratio of the distance between the foci of the ellipse and its major axis length, the value 0 is for circle and 1 for line segment. Eccentricity here calculated for the ellipse that has same second moment as the nodule region, gives the circularity measure.
- Extent: It is the ratio of the Area and the Area of the smallest rectangle containing the nodule region. It is a measure of compactness.
- Equivalent diameter: It specifies the diameter of the circle with the same area of the nodule region and is calculated as :

$$d = \sqrt{4 * \text{Area}/\pi} \qquad (2)$$

- Perimeter: It is the distance around the boundary of the nodule region and computed by calculating the distance between each adjoining pair of pixels around the border of the region.
- Major axis length: Specifies the length (in pixels) of the major axis of the ellipse that has the same normalized second central moments as the nodule region.
- Minor axis length: Specifies the length (in pixels) of the minor axis of the ellipse that has the same normalized second central moments as the region.

- Minimum intensity: It is the minimum value of the intensity in the nodule region.
- Maximum intensity: It is the maximum value of the intensity in the nodule region.
- Mean intensity: It is the mean value of the intensity in the nodule region.

3.5. Filtering

After analyzing the values of extracted features, a suitable fixed value or threshold is selected for certain features and removal of highly non- nodule candidates is carried out. The features selected are Area, Solidity, Extent, Eccentricity and the ratio of minor axis length to major axis length.

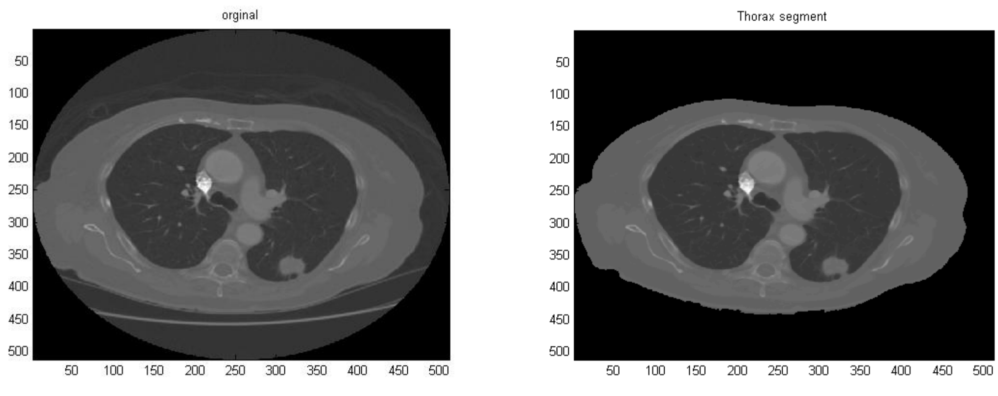
The number of false nodules is usually large in pulmonary nodule detection system. Hence the initial filtering process effectively reduces non-nodule candidates. In addition, at later stage, false nodules can be further eliminated with a suitably designed feature based classifier.

4. Results and Discussions

Lung CT image database available in open domain can be used to develop, train and validate an efficient Computer Aided Detection system for lung nodules. In this paper, two such databases have been made use of. They are the Lung Image Database Consortium (LIDC) database and Vision and Image Analysis Group - The Public Lung database to address Drug response (VIA/PLD). LIDC database contains 1018 cases, along with the images, radiological annotations are also provided to identify the location and radiological characteristics of the lesions [11]. VIA/PLD database contains 100 cases with location, size and characteristics [12].

In this study the clinical dataset used for experiments consists of 60 CT slices of 10 different CT scans from PLD and LIDC which are in Digital Imaging and Communications in Medicine (DICOM) format with resolution 512 x 512 and with a wide range of scanning parameter values. (Slice thickness 0.6mm to 2.5mm, In-plane pixel size 0.461mm to 0.977mm).

A Computed Tomography image from the LIDC database with a solid nodule connected to the border is shown in Fig 2(a). The thorax mask obtained from the preprocessed image by thresholding and morphological operations is used to remove the artifacts external to the patient’s body. Thorax extraction is the first step in the proposed lung segmentation method, and the extracted thorax region using the thorax mask is shown in Fig 2(b). This step makes the lung region extraction more accurate. The lung mask is prepared so as to include nodules connected to the boarder. The segmented lung region using lung mask is shown in Fig 3 (a). After the successful extraction of lung region without eliminating the nodule connected to the boarder, next step is to identify the possible nodule locations.



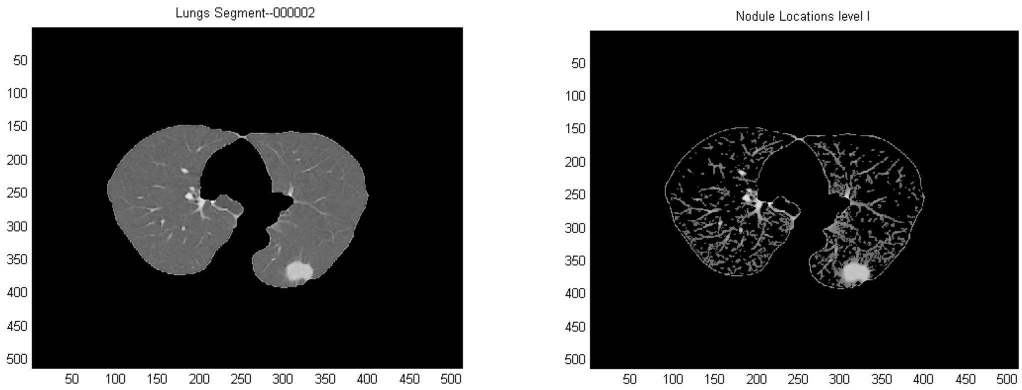
(a) (b)

Fig. 2. (a) Computed Tomography image; (b) Thorax region extracted.

An Intermediate level thresholding is proposed and employed for this and the threshold is selected by giving an offset to the maximum value of intensity which is decided from mean intensity value within the extracted lung region. The resultant mask is used to obtain the intermediate level pulmonary nodule locations. Fig 3(b) shows the possible nodule locations obtained by the proposed method.

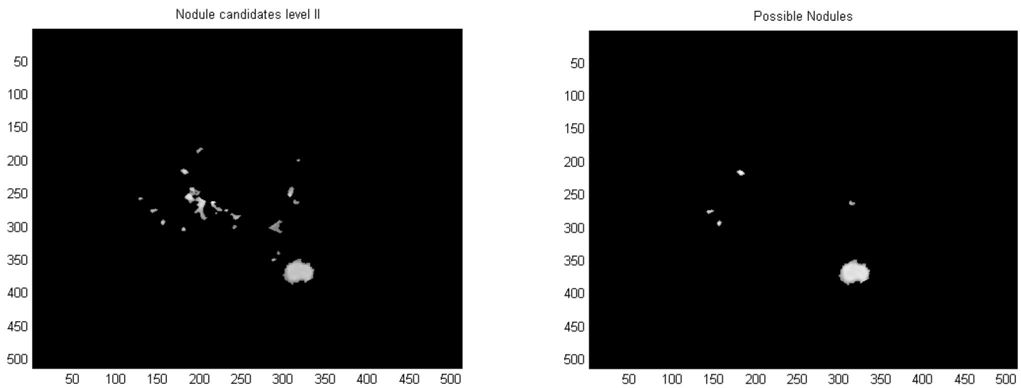
After segmenting possible nodule locations, the next step is to identify the pulmonary nodule candidates. A Micro-level threshold is selected based on the maximum, mean and minimum intensity values. The pulmonary nodule candidates identified in the micro level thresholding is shown in Fig 4(a). As explained earlier, the number of false nodules may be large and prospective non- nodule candidates are eliminated by an initial filtering using shape features. The remaining highly prospective nodule candidates are given in Fig 4(b).

Simulation studies are carried out and Table 1 summarizes the performance of the proposed method. Number of nodule candidates detected at the Micro Level thresholding and number of nodules after filtering are given. Number of false nodules can be greatly reduced by the filtering operation but the number is still large.



(a) (b)

Fig. 3. (a) Segmented Lung image; (b) Possible Pulmonary nodule locations.



(a) (b)

Fig. 4. (a) Nodule candidates identified in the Micro Level Thresholding; (b) Possible Pulmonary nodule candidates after rule based filtering.

Table 1. Pulmonary nodule Candidate detection results of the proposed study.

Database for the study	No. of Slices used	Type of the nodule	No. of Nodules Micro level	No. of Nodules after filtering
LIDC	7	Juxta-pleural	56	11
PLD-SM	37	Solid	869	221
	11	Non solid	82	11
PLD-SL	5	Solid	93	32

The proposed method is best suited for detection of isolated solid nodules. In addition, detection of non solid nodules is also possible to certain extent because of two level thresholding. Small nodules connected to the boarders are also included by morphological closing operations. But larger size juxta-pleural nodules are eliminated because of the selected size of structuring element. This can be overcome by appropriate algorithmic modifications in the proposed method.

5. Conclusion

In this paper, a novel approach making use of three level Thresholding for efficient segmentation and feature extraction of pulmonary nodule candidates has been proposed. The technique employed Global, Intermediate and Micro -level Thresholdings for segmentation and connected component analysis for extraction of features from segmented pulmonary nodules. Prospective non- nodule candidates are eliminated by an initial filtering. Feature-based classifier also can be used for further false nodule elimination.

References

[1] Yoshino.I, Yamaguchi.M, Tagawa.T, Fukuyama.S, Kameyama.T, Osoegawa.A, " Operative results of clinical stage I non small cell lung cancer", Lung cancer, vol 42, (2),pp. 221-225, May 2003.

[2] Girvin.F, Ko.J P, "Pulmonary nodule: Detection, Assesssment, and CAD", AJR, 191,pp. 1057-1069,October 2008.

[3] Hu.S, Hoffman.EA, Reinhardt.JM, "Automatic lung segmentation for accurate quantitation of volumetric X-ray CT images," IEEE Transactions on Medical Imaging, vol. 20, no. 6, pp. 490–498, 2001

[4] Van Rikxoort.EM, De Hoop. B, Van de Vorst.S, Prokop.M, Van Ginneken.B, " Automatic Segmentation of pulmonary segments from Volumetric lung CT scans", IEEE Transaction s on Medical Imaging , vol.28, no.4, pp. 621-630,2009

[5] Armato III SG, Sensakovic.WF, "Automated lung segmentation for thoracic CT: impact on computer-aided diagnosis, "Academic Radiology, vol. 11, no. 9, pp. 1011–1021, 2004

[6].Wei.W, Hu.Y, Gelfand.G, .MacGregor.JH, "Segmentation of lung lobes in high-resolution isotropic CT images," IEEE Transactions on Biomedical Engineering, vol. 56, no. 5, pp. 1383– 1393, 2009

[7] Homma.N, Shimoyama.S, Ishibashi.T, Yoshizawa.M, "Lung Area Extraction from X-ray CT Images for Computer-aided Diagnosis of Pulmonary Nodules by using Active Contour Model", WSEAS Transactions On Information Science And Applications, Issue 5, Volume 6, May 2009, 746- 755

[8] Silva da Sousa.JRF, Silva,AC, De Paiva.AC, Nunes.RA, "Methodology for automatic detection of lung nodules in computerized tomography images", Elsevier. computer methods and programs in biomedicine, 98, pp.1–14, 2010.

[9] Sun.S, Bauer.C, Beichel.R," Automated 3D Segmentation of lung with lung cancer in CTusing anovel robust active shape model approach" IEEE Transactions on Medical imaging, vol.31, no.2, pp. 449-460, 2012

[10] Ashiskumar.D, Sudiptha.M, .Niranjanan.K, "Computer –aided Detection and Analysis of Pulmonary nodule From CT Images : A Survey", IETE Technical Review, Vol 29, Issue 4, pp. 265-275, Jul-Aug 2012

[11] Armato.SG, McLennan.G, Bidaut.L, McNitt-Gray.MF, Meyer.CR, Reeves.AP, Zhao.B , Aberle.DR, Henschke.CI, Hoffman.EA, Kazerooni.EA, MacMahon.H, van Beek.EJR , Yankelevitz.D, et al.: " 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: 915--931, 2011

[12] Reeves.AP, Biancardi.AM, Yankelevitz.D, Fotin.S, Keller.BM, Jirapatnakul.A, Lee.J , "A Public Image Database to Support Research in Computer Aided Diagnosis".Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 2009. EMBC 2009. Sept. 2009