Lung Nodule Detection in CT Images using Thresholding and Morphological Operations

Sudha. V, Jayashree. P

Abstract— Lung cancer which is among the five main types of cancer is a leading one to overall cancer mortality contributing about 1.3 million deaths/year globally. Lung cancer is a disease and it is characterized by uncontrolled cell growth in tissues of the lung. Lung nodule is an abnormality that leads to lung cancer, characterized by a small round or oval shaped growth on the lung which appears as a white shadow in the CT scan. An effective computer aided lung nodule detection system can assist radiologists in detecting lung abnormalities at an early stage. If defective nodules are detected at an early stage, the survival rate can be increased up to 50%. This paper aims to develop an efficient lung nodule detection system by performing nodule segmentation through thresholding and morphological operations. The proposed method has two stages: lung region segmentation through thresholding and then segmenting the lung nodules through thresholding and morphological operations.

Index Terms—Computed Tomography, Morphological Operations, Segmentation, Thresholding.

I. INTRODUCTION

According to statistics, lung cancer is the leading cause of cancer related deaths compared to any other type of cancer in the world. Lung cancer is contributing about 1.3 million deaths/year globally [20]. Further, these reports indicate that the survival rate of lung cancer is only 14%; but still, if defective nodules are detected at an early stage, the survival rate can be increased up to 50%. Thus the early detection of lung nodules is important in the treatment of lung cancer [18].

In lung cancer research, one of the most sensitive methods for detecting pulmonary nodules is Computed Tomography (CT), in which a nodule is defined as a rounded and irregular opaque figure on a CT scan, with a diameter up to 30mm. Each scan contains hundreds of images that must be evaluated by a radiologist, which is a difficult process. So for this reason, the use of a Computer-Aided Detection (CAD) system can provide an effective solution by assisting radiologists in increasing the scanning efficiency and potentially improving nodule detection [17].

In general, a nodule detection system consists of two steps: lung segmentation and nodule candidate detection. In literature, the use of optimal thresholding is a common method for segmenting a lung volume [1, 6]. For example, Xujiong Ye et al. [18] used a 3D-adaptive fuzzy thresholding method to segment the lung region, since a fixed threshold value has also been used [16]. After thresholding, the lung volume was then extracted from the segmented images using 3D approaches. A 3D connectivity with a seed point in the initial lung region, and 3D-connected component labelling techniques have also been used to segment the lung volume without artefacts [17]. Note that in these cases, the extracted lung volume needs to be refined to include juxta-pleural nodules. Subsequently, due to the complexity of these approaches, several methods have been presented for refining a lung mask. Notably, a rolling ball algorithm has been used for effective lung mask correction [1, 17], in which the rolling ball algorithm is equivalent to the combination of two fundamental morphological operations: erosion and dilation. More recently, the application of a chain code representation over a lung mask was also proposed an attempt to correct the contours [19].

From the segmented lung volume, nodule candidates have been detected using a number of methods. As examples, Armato et al. [1] and Messay et al. [17] applied multiple gray-level thresholds to the volumetric lung regions to identify nodule candidates. In addition, a number of template-matching based methods have been studied. Brown et al. [8] developed a patient- specific model that could be used in combination with image primitive-matching to find nodules. And Ye et al. [7] proposed a novel template-matching technique based on a genetic algorithm template-matching (GATM) technique to detect nodules within the lung area; the GA was used to determine the required target position in the observed image. Based on further research, Dehmeshki et al. [7] then improved upon this method by adding a shape-based approach to detect nodules having spherical elements. Shape-based methods have also been popular in nodule detection. And Ye et al. [19] proposed a shape-based detection method by combining the shape index (local shape information) and ‘‘dot’’ features (local intensity dispersion information); this detection method has provided a good structural descriptor for nodule detection, and other filtering-based method have been used to detect spherical objects as nodule candidates. Retico et al. [21] subsequently reported an automated procedure for selecting the nodule candidates, based on a filter that enhances the shape of spherical objects.

II. PROPOSED SYSTEM FOR LUNG NODULE DETECTION

An automated pulmonary nodule detection system is generally considered advantageous to detect lung abnormalities at an early stage. The basic layout of the proposed system is shown in Fig.1.

Manuscript Received on December 2012.

Sudha.V, Electronics and Instrumentation Engineering, Kongu Engineering College, Erode, India.

Jayashree.P, Electronics and Instrumentation Engineering, Kongu Engineering College, Erode, India.

Initially the lung region is extracted from the CT image by performing thresholding and morphological reconstruction. Then from the extracted lung region we are segmenting the nodules by using global thresholding and morphological operation [17]. The segmented nodules are used for feature extraction. The features like geometric and intensity-based statistical features are extracted. The features extracted are given for classification. Genetic programming-based classifier is used to classify the cancerous and non-cancerous nodules.

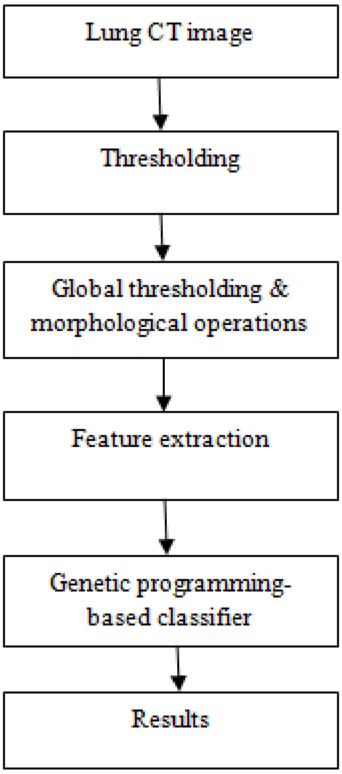


Fig.1. Block diagram of proposed lung nodule detection system

III. METHODOLOGY

Fig.1 gives an overview of our proposed lung nodule detection scheme. In the following sections, each stage is described in detail.

A. Lung Region Segmentation

Lung region segmentation is an essential preprocessing step in lung nodule detection systems. The main purpose of lung segmentation is to separate the voxels corresponding to the lung cavity in axial CT scan slices from the surrounding lung anatomy. As such, the accuracy of lung segmentation largely influences the nodule detection results. Our proposed CAD system includes a fully automated method for segmenting lung regions in CT scans: global thresholding is used to obtain the lung region and the thin structures attached to the lung region are cleared using morphological reconstruction.

1) Thresholding

The CT scan can be separated into two types of voxels, characterized by the density differences between the two anatomical structures. The high-density regions primarily consist of the body surrounding the lung cavity, whereas the low-density regions contain the lung cavity, the air surrounding the body, and other low-intensity regions. To extract the lung volume, we need to segment the low-density regions in the initial stage. For lung image segmentation, when using a histogram method, a fixed threshold value is needed in order to separate the low-density lung parenchyma from the surrounding lung anatomy, though the availability of different scanning protocols makes the selection of an appropriate threshold a challenging task [5]. The appropriate threshold is selected using global thresholding algorithm. Therefore, after applying a fixed threshold, subsequent procedure is used to refine the initial segmentation results. A thresholded image g(x, y) is defined as

$$g(x,y) = \begin{cases} 1 \text{ if } f(x,y) < T \\ 0 \text{ if } f(x,y) \geq T \end{cases} \tag{1}$$

The following iterative procedure is one such approach:

1. Select an initial estimate for T.
2. Segment the image using T. This will produce two groups of pixels: G_1 consisting of all pixels with gray level values <T and G_2 consisting of pixels with values $\geq$ T.
3. Compute the average gray level values μ_1 and μ_2 for the pixels in the regions G_1 and G_2.
4. Compute a new threshold value:
$$T = \frac{1}{2}(\mu_1 + \mu_2) \tag{2}$$
5. Repeat steps 2 through 4 until the difference in T in successive iterations is smaller than a predefined parameter T_0.

When there is a reason to believe that the background and object occupy comparable areas in the image, a good initial value for T is the average gray level of the image. When objects are small compared to the area occupied by the background, then one group of pixels will dominate the histogram and the average gray level is not as good an initial choice. A more appropriate initial value for T in cases such as this is a value midway between the maximum and minimum gray levels. The parameter T_0 is used to stop the algorithm after changes become small in terms of this parameter. This is used when speed of iteration is an important issue [11]. The threshold values obtained using the above algorithm is shown in the Table (i).

Table (i): Threshold Values for Lung Region Extraction

PATIENTS	IMAGES	THRESHOLD VALUE
PATIENT 1	Image 1	133.9786
	Image 2	133.2739
	Image 3	131.3092
	Image 4	131.0963

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

PATIENT 2	Image 1	135.6461
	Image 2	135.5892
	Image 3	136.6548
	Image 4	134.6029
PATIENT 3	Image 1	136.4834
	Image 2	137.3816
	Image 3	135.7863
	Image 4	131.0689
PATIENT 4	Image 1	138.0430
	Image 2	140.4114
	Image 3	139.3237
	Image 4	136.2806

Based on the threshold T given in the table, the lung region is identified from the original CT image as shown in Fig. 2(b). A morphological reconstruction operation is applied after thresholding to extract the lung region alone.

2) Morphological Reconstruction

The morphological reconstruction operation is used here to remove the objects that are connected to the image border [11].

i) Clearing Border Objects

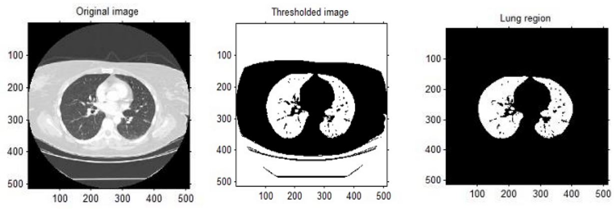
One of the useful applications of reconstruction is removing objects that touch the border of an image. Here, the key task is to select the appropriate marker to achieve the desired effect. The marker image F, is defined as

$$F(x, y) = \begin{cases} I(x,y) & \text{if } (x,y) \text{ is on the border of } I \\ 0 & \text{otherwise} \end{cases} \tag{3}$$

where I is the original image. Then by using I as the mask image, the reconstruction

$$H = R_I\,(F) \tag{4}$$

The Equation 4 yields an image, H that contains only the objects touching the border. The difference, 1-H, contains only the objects from the original image that do not touch the border as shown in the Fig 2(c).



(a) (b) (c)

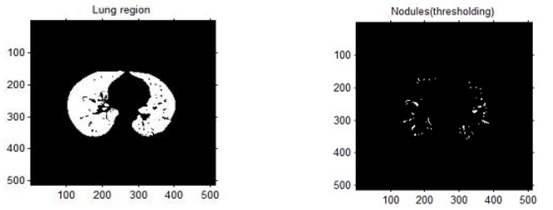
Fig.2. Lung segmentation based on thresholding and morphological reconstruction. (a) Original CT lung image; (b) lung region extracted using thresholding; (c) final segmented lung region by morphological reconstruction.

B. Lung Nodule Segmentation

After extracting the lung region from the original image, the lung nodules are segmented using thresholding operation [11]. The threshold values obtained using global thresholding algorithm for lung nodule segmentation is tabulated in the Table (ii). The results obtained by using these threshold values are shown in the Fig.3.

Table (ii) Threshold Values for Lung Nodule Segmentation

PATIENT	IMAGES	THRESHOLD VALUE
PATIENT 1	Image 1	0.0795
	Image 2	0.1248
	Image 3	0.0467
	Image 4	0.0075
PATIENT 2	Image 1	0.0559
	Image 2	0.1491
	Image 3	0.1610
	Image 4	0.0471
PATIENT 3	Image 1	0.1753
	Image 2	0.2242
	Image 3	0.1330
	Image 4	0.0139
PATIENT 4	Image 1	0.1439
	Image 2	0.1844
	Image 3	0.1759
	Image 4	0.0106



(a) (b)

Fig.3 lung nodules segmented using thresholding operation. (a) Lung region; (b) nodules segmented by thresholding

1) Mathematical morphology

Mathematical morphology is a tool for extracting image components. The operations of MM are originally defined as set operations and are used to extract the edges of an image, to filter an image and to skeletonise an image. Here we have used this for filtering the false positives from the detected nodules [12]. Morphological Opening Morphological opening removes completely removes completely regions of an object that cannot contain the structuring element, smooth's object contours, breaks thin connections and removes thin protrusions.

The morphological opening of A by B, denoted $A \circ B$, is defined as the erosion of A by B, followed by a dilation of the result by B:

$$AoB = \cup\left\{(B)_Z \,\middle|\, (B)_Z \subseteq A\right\} \tag{5}$$

Structuring Elements

The structuring element is a small binary image, i.e. a small matrix of pixels, each with a value of zero or one. The matrix dimensions specify the size of the structuring element. The pattern of ones and zeros specifies the shape of the structuring element. An origin of the structuring element is usually one of its pixels, although generally the origin can be outside the structuring element. The structuring element used in the morphological operation plays a vital role with its different shapes and sizes. The resultant value is applied to the centre pixel and it can be anywhere in the structuring element according to the applications [12].
There are many kinds of structuring element: Disk-shaped, Diamond-shaped, Ball-shaped, Square-shaped, Flat linear with length LEN, arbitrary with the specified neighborhood.
Here disk-shaped structuring element of size 2 is used. The results obtained using these operations are shown in the Fig 4.4.

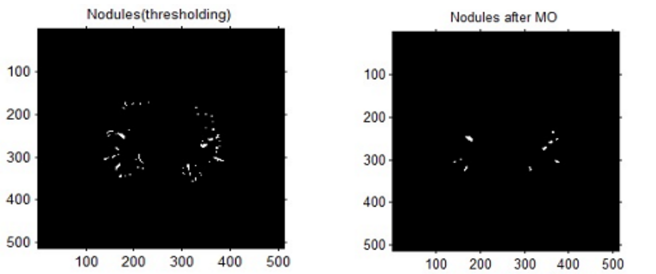


Fig.4 Nodules Detected after Morphological Opening

IV. EXPERIMENTAL RESULTS

The proposed CAD system is evaluated using the LIDC database [2, 17], a publicly available database from the National Biomedical Imaging Archive (NBIA), and its nodules have been fully annotated by multiple radiologists. In this database, four expert chest radiologists drew outlines for nodules having effective sizes of 3 mm or greater. The LIDC database consists of 84 CT scans, but only 58 CT scans contain nodules. In the nodule containing CT scans, we randomly collected 20 CT scans in order to evaluate the proposed system. All annotated nodule segmentations were used in the evaluation of the proposed method [17].

In the first step, the lung region is segmented based on thresholding and morphological reconstruction. In the second step the nodules are segmented using thresholding and morphological operation is used to reduce the number of false positives. The results are shown in the Fig.5.

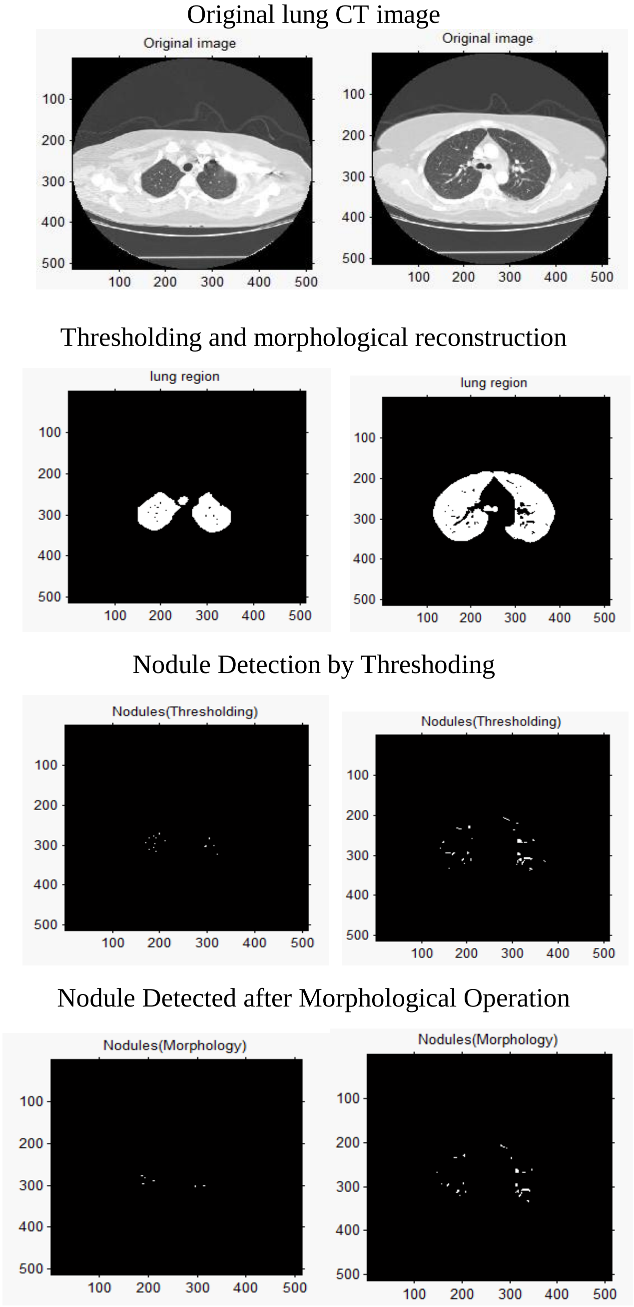


Fig.5. Results of lung nodule detection

Extracted feature values
The extracted features for the nodule images are tabulated and given in the Table (iii). A set of images are analyzed and the features such as entropy, correlation, energy, contrast, homogeneity, standard deviation, mean, skewness and kurtosis are extracted.

Table (iii) Extracted Feature Values from the Lung Nodule Images

FEATURES	IMAGE 1	IMAGE 2	IMAGE 3
Entropy	-0.6928	-0.6906	-0.6899
Correlation	0.9999	0.9996	0.9995
Energy	0.9994	0.9958	0.9945
Contrast	1.4971e-004	7.147e-004	9.6102e-004
Homogeneity	0.9999	0.9996	0.9995
Mean	2.4414e-004	0.0018	0.0023

V. CONCLUSION AND FUTURE WORK

Lung cancer is a major cause of cancer-related deaths; it can be detected early by detecting the lung nodules. Early detection can improve the survival rate of lung cancer patients. The main idea of this project is to detect lung nodule and to classify nodules as cancerous and non-cancerous using Genetic Programming-based Classifier (GPC) technique. Thus the lung CT image is subjected to various processing steps and features are extracted for a set of images. The processing steps include thresholding, morphological operations and feature extraction. By using these steps the nodules are detected and segmented and some features are extracted. The extracted features are tabulated for future classification.

The future work of this project is to identify the effective features for further classification. Genetic Programming-based Classifier will be used for classification of lung CT images as cancerous and non-cancerous by using the identified effective features.

REFERENCES

1. S.G.Armato, M.L.Giger, C.J.Moran, J.T.Blackburn, K.Doi, H.MacMahon (1999) ‘Computerized detection of pulmonary nodules on CT scans’, Radiographics 19 1303-1311.
2. S.G.Armato, G.McLennan, M.F. McNitt-Gray, C.R.Meyer, D.Yankelevitz, D.R.Aberle, C.I.Henschke, E.A.Hoffman, E.A.Kazerooni, H.MacMahon, A.P.Reeves, B.Y.Croft, L.P.Clarke (2004) L.I.D.C.R. Group, Lung image database consortium: developing a resource for the medical imaging research community, Radiology 232, 739-748.
3. Eva M. van Rikxoort, Mathias Prokop, Bartjan de Hoop, Max A. Viergever, Josien P. W. Pluim, and Bram van Ginneken (2010) ‘Automatic Segmentation of Pulmonary Lobes Robust Against Incomplete Fissures’, IEEE transactions on medical imaging, vol. 29, no. 6.
4. Jan-Martin Kuhnigk, Volker Dicken, Lars Bornemann, Annemarie Bakai, Dag Wormanns, Stefan Krass, and Heinz-Otto Peitgen (2006) ‘Morphological Segmentation and Partial Volume Analysis for Volumetry of Solid Pulmonary Lesions in Thoracic CT Scans’, IEEE transactions on medical imaging, vol. 25, no. 4.
5. Jamshid Dehmeshki, Hamdan Amin, Manlio Valdivieso, and Xujiong Ye (2008) ‘Segmentation of Pulmonary Nodules in Thoracic CT Scans: A Region Growing Approach’, IEEE transactions on medical imaging, vol. 27, no. 4, 467
6. Jiantao Pu, David S. Paik, Xin Meng, Justus E. Roos, and Geoffrey D. Rubin (2011) ‘Shape “Break-and-Repair” Strategy and Its Application to Automated Medical Image Segmentation’, IEEE transactions on visualization and computer graphics, vol. 17, no. 1.
7. Jamshid Dehmeshki, X. Ye, X. Lin, M. Valdivieso, H. Amin (2007) ‘Automated detection of lung nodules in CT images using shape-based genetic algorithm’, Computerized Medical Imaging and Graphics 31, 408–417.
8. Matthew S. Brown, Michael F. McNitt-Gray, Jonathan G. Goldin, Robert D. Suh, James W. Sayre, and Denise R. Aberle (2001) ‘Patient-Specific Models for Lung Nodule Detection and Surveillance in CT Images’, IEEE transactions on medical imaging, vol. 20, no. 12.
9. Panayiotis D. Korfiatis, Anna N. Karahaliou, Alexandra D. Kazantzi, Cristina Kalogeropoulou, and Lena I. Costaridou (2010) ‘Texture-Based Identification and Characterization of Interstitial Pneumonia Patterns in Lung Multidetector CT’ , IEEE transactions on information technology in biomedicine, vol.14, no. 3.
10. Pedro G. Espejo, Sebasti´an Ventura, and Francisco Herrera (2010) ‘A Survey on the Application of Genetic Programming to Classification’, IEEE transactions on systems, man, and cybernetics—part c: applications and reviews, vol. 40, no. 2.
11. Rafael C. Gonzalez, Richard E. Woods and Steven L.Eddins (2010) ‘Digital Image Processing Using MATLAB’, second edition.
12. Rafael C. Gonzalez and Richard E. Woods (2002) ‘Digital Image Processing’, Prentice Hall, second edition.
13. Sang Cheol Park, Brian E. Chapman, Bin Zheng (2011) ’ A Multistage Approach to Improve Performance of Computer-Aided Detection of Pulmonary Embolisms Depicted on CT Images: Preliminary Investigation’, IEEE transactions on biomedical engineering, vol. 58, no. 6.
14. Shanhui Sun, Christian Bauer, and Reinhard Beichel (2012) ‘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, vol. 31, no. 2.
15. Stefano Diciotti, Giulia Picozzi, Massimo Falchini, Mario Mascalchi, Natale Villari, and Guido Valli (2008) ’ 3-D Segmentation Algorithm of Small Lung Nodules in Spiral CT Images’, IEEE transactions on information technology in biomedicine, vol. 12, no. 1.
16. Tao Xu, Mrindal Mandal, Richard Long, Irene Cheng and Anup Basu, (2012) ‘An edge-region force guided active shape approach for automatic lung field detection in chest radiographs’, Computerized Medical Imaging and Graphics .
17. Temesguen Messay, Russell C. Hardie, Steven K. Rogers (2010) ‘A new computationally efficient CAD system for pulmonary nodule detection in CT imagery’, Medical Image Analysis 14 390–406.
18. Wook-Jin Choi, Tae-Sun Choi (2012) ‘Genetic programming-based feature transform and classification for the automatic detection of pulmonary nodules on computed tomography images’, Information Sciences 212 57–78
19. [19] Xujiong Ye, Xinyu Lin, Jamshid Dehmeshki, Greg Slabaugh, Gareth Beddoe (2009) ‘Shape-Based Computer-Aided Detection of Lung Nodules in Thoracic CT Images’ , IEEE transactions on biomedical engineering, vol. 56, no. 7.
20. World health organization Cancer, accessed on February 02(2010) http://www.who.int/mediacentre/factsheets/fs297/en/index.html .
21. A.Retico, P.Delogu, M.Fantacci, I.Gori, A. Preite Martinez, ‘Lung nodule detection in low- dose and thin-slice computed tomography’, Computers in Biology and Medicine 38 (2008) 525-534.