

# An Improved Pulmonary Nodule Detection Scheme based on Multi-Layered Filtering and 3d Distance Metrics

Baber Jahangir

International Islamic University  
Islamabad Pakistan

Muhammad Imran

<sup>2</sup>Shaheed Zulfikar Ali Bhutto  
Institute of Science and Technology

Qamar Abbas

Iqra university Islamabad

Shahina Rabeel

International Islamic University Islamabad Pakistan

Ayyaz Hussain

International Islamic University Islamabad Pakistan

**Abstract**—This paper proposed a computer-aided detection (CAD) system to automatically detect pulmonary nodules from thoracic computed tomography (CT) images. Automatically detect pulmonary nodules is a difficult job because of the large deviation in size, shape, location and density of nodules. The proposed CAD scheme applies multiple 3D disk-shaped laplacian filters to enhance the shape of spherical regions. Optimal multiple thresholding and 3D distance mapping is used to extract regions of interest and separate nodules. Finally, rule-based pruning removes easily dismissible false positive structures. The proposed system provides an overall nodule detection rate of 80% with an average of 12.2 false positives per scan. The experimental results reveals that the proposed CAD can attain a comparatively high performance.

**Keywords**—computer-aided detection system; nodule detection lung nodule

## I. INTRODUCTION

Among all types of cancer, lung cancer presents the leading rates of mortality around the world [1]. To detection the lung cancer at the earliest stages, computed tomography is known to be the most accurate, non-invasive imaging technique, which motivate the researchers to use computer for the early detection of lung cancer [2–3]. Lung cancer manifests itself in the form of pulmonary nodules which are visible as anatomic structures having a radio density greater than the lung parenchyma. Due to the large number of images per scan, radiologists can make errors while evaluating them. To aid the radiologists, there has been a growing interest in developing automatic computer-aided detection (CAD) methods [4,5]. Sahiner et al. [6] proved through experimental study that a medical physician with the help of CAD system can easily diagnose cancer rather than without CAD system. similarly a CAD system alone cannot perform well without the help of medical physician. The automatic detection of lung nodules, however, is a challenging task due to the variance in shape, size, location and density of the nodules Various researchers proposed different techniques for computer-aided nodule detection system. Because lung nodules have higher intensity values than those of the surrounding lung parenchyma, intensity based schemes provide the advantage of simplicity. Armato et al. [4] and Messay et al.

[5] used multiple gray-level thresholds to extract the nodule candidates from the segmented lung volume. Kostis et al. [6] applied morphological opening along with iterative dilation procedure to separate small nodules from the attached vessels. The study by Brown et al. [7] involved intensity thresholding, region growing and mathematical morphology to identify regions of interest. Similarly, the technique by Choi et al. [8] uses optimal multiple thresholding and rule-based pruning that uses local shape features to extract potential nodules from the lung region. Upon analysis, it was observed that the algorithm provides good performance in the detection of high-contrast, well shaped nodules but has low sensitivity for irregularly shaped non-solid nodules. In addition, the number of omissions for nodules that touched or infiltrated vascular structures was high. The objective of the proposed study is to improve upon this technique by adding filtering-based method and distance measures to improve the detection of such nodules.

Studies have demonstrated the advantages of using such techniques. Li et al. [9] obtained high sensitivity with low false positive detections per scan by using selective rule based segmentation based and enhancement. Suzuki et al. [10] enhanced the nodule intensity and remove non-nodules through a machine learning technique which is based on based filter. Retico et al. [11] proposed an automatic filter to enhance the shape of spherical structures. Additionally, using a signed distance field, Pu et al. [12] detected maximums found as nodule candidates. Ozekes et al. [13], presented a rule based segmentation technique with the help genetic algorithm. the proposed system achieved good results i.e. 93.4% sensitivity and 0.594 false positives on each examined. Similarly, Netto et al. [14] applied 3D distance transform and region growing to separate the lung nodules from vessels and bronchi. Ye et al. [15] extract five features such as "containing intensity information, shape index, and 3D spatial location".

Model based methods also provide promising results for nodule detection. Cascio et al. [16] use a stable 3D mass-spring model combined with a spline curve reconstruction process in order to segment and extract the nodular candidate regions. Lee et al. [17] presented a novel template-matching technique with the help of genetic algorithm template-matching technique for nodule detection. Dehmehki et al. [18] enhanced this method

by adding a shape-based methodology for nodule detection from spherical elements. Tan et al. [19] utilized the three classifiers such as artificial neural network and genetic algorithm for lung nodule detection. To validate the proposed model, results were compared with SVM and fixed-topology neural networks based models.

The CAD system proposed by Tan et al. [19] utilized three classifier; two of them were artificial neural network and genetic algorithms and compared the results of these classifier with SVM and fixed-topology neural networks. Suárez-Cuenca et al. [20] have explored discriminant analysis (LDA), three types of support vector machines (SVM), artificial neural network (ANN) and quadratic discriminant analysis (QDA). These all six classifiers were tested combined and separately as well on LIDC base having 85 samples.

The objective of this research work is to present a new and improved pulmonary nodule detection system that uses multiple sharpening filters along with intensity thresholding and 3D shape-based metrics to improve accuracy of nodule detections while reducing the number of false positives.

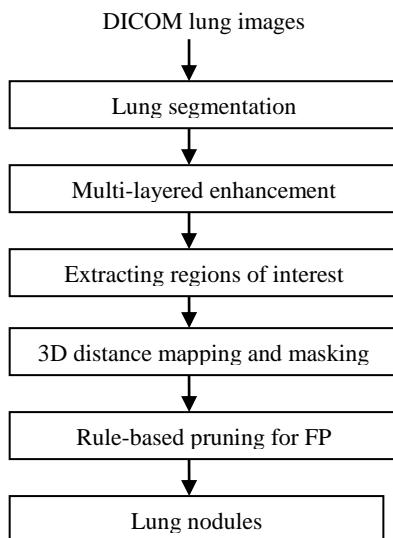


Fig. 1. Overview of proposed CAD system

## II. METHODOLOGY

The research work conducted by Choi et al. [8], is closely related to this research work in which automatic detection of lung nodules is performed by using multiple intensity thresholds and rule-based pruning based on shape-based features of the nodule candidates. However, our work has some major differences. First, structures in the lung image that have a spherical appearance are enhanced by using 3D circular laplacian filters of varying sizes. This highlights nodular regions and de-emphasizes non-nodule objects in the segmented lung volume. Second, apart from using multiple thresholding to extract initial possible nodule, the 3D distance map is computed in each of the segmented structures. The distance map gives a measure of the radius of objects, which is the distance from the boundary to the centroid. Masking the 3D image to extract the inner most layers of each structure separates the lung nodules from vessels and bronchi. Finally,

easily dismissible false positive objects are removed using rule-based pruning based on local shape-based features.

### A. Lung segmentation

The algorithm for lung volume segmentation has three stages. In order to separate low density non-body voxels (nodule candidate) from the high density body voxels (non-nodule candidate) a fixed threshold value is used first. Second, 3D connected component labeling is used to extract the lung regions from non-body voxels. To make sure that the extracted lung mask volume includes the juxtapleural nodules, a third step of contour correction is applied. Critical section is removed by using chain code representation in this step. Fig.2 shows the various stages of the lung segmentation process.

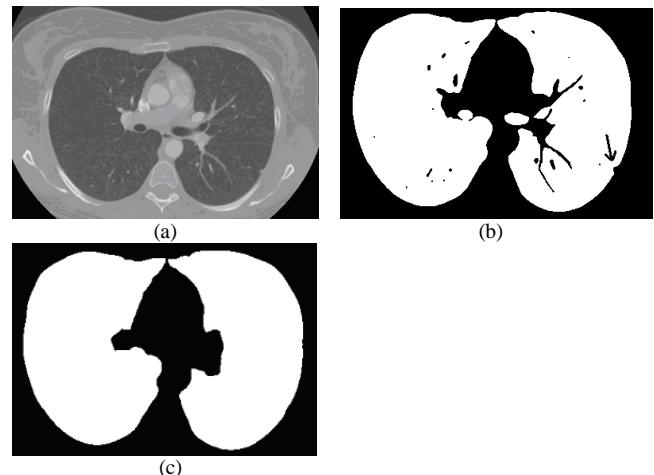


Fig. 2. Stages of lung segmentation process. (a) Original CT lung image; (b) Extracted lung volume (with concave region on the pleural boundary); (c) Final segmented lung mask with refined boundary

### B. Multi-layered enhancement of disk-shaped regions

Before beginning with nodule detection, multiple linear 3D spatial filters of varying sizes are used to convolve the original 3D image. The result of displacing the convolution filter over the image is a new 3D image in which the contrast of structures that match the size of the disk-shaped kernel is increased. The enhanced image for kernel size  $3 \times 3 \times 3$  is obtained by:

$$g(x, y, z) = f(x, y, z) + c[\nabla^2 f(x, y, z)] \quad (1)$$

where  $f(x, y, z)$  and  $g(x, y, z)$  are the input and sharpened images, respectively. The value of constant  $c = 1$ , and  $\nabla^2 f(x, y, z)$  is the discrete laplacian of three variables:

$$\begin{aligned} \nabla^2 f(x, y, z) = & -[f(x+q, y, z) + f(x-q, y, z) + \\ & f(x, y+q, z) + f(x, y-q, z) + \\ & f(x, y, z+q) + f(x, y, z-q) - \\ & f(x, y, z)] \end{aligned} \quad (2)$$

where  $q = 1, 2, 3, \dots, n$  ( $n$  = number of kernels). Similar enhanced images as in (1) are obtained by progressively increasing the kernel size in odd increments. The coefficient values are determined by varying the value of  $q$  and inserting respective terms in (2).

In the next step, all sharpened images are compared and only the highest intensity value between corresponding voxels is retained to obtain a final enhanced image. This preprocessing step provides two-fold advantage of making the system more sensitive to detection of non-solid nodules and decreasing the likelihood of oblong shapes passing the gray-value thresholds. Fig.3 shows an example of a non-solid, juxtapleural nodule that was successfully detected as a true positive structure.

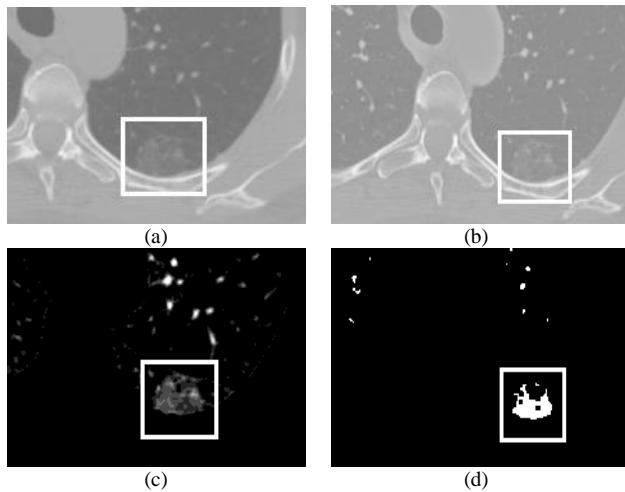


Fig. 3. An example of a non-solid nodule successfully detected by the proposed system. (a) Original CT lung image; (b) Corresponding enhanced lung image; (c) Separated regions of interest; (d) Potential nodule candidates

### C. Extracting regions of interest

multiple threshold is used to extract regions of interest from the segmented lung volume, which uses an optimal threshold as the base threshold. The optimal threshold is calculated by using an iterative approach. The average intensity value of 3D image is selected as the primary threshold  $T$ . Applying this threshold to the image segments the image into two groups of voxels: voxels with intensity values higher than the threshold, and voxels with grayscale values less than or equal to the threshold. Let  $\mu_a$  and  $\mu_b$  represent the mean intensity of the first and second group, respectively. The new threshold is calculated by:

$$T = \frac{\mu_a + \mu_b}{2} \quad (3)$$

The above process is repeated until the threshold converges.

The final obtained threshold becomes the base threshold  $T_{base}$  for multiple thresholding. Instead of using a fixed value threshold as the base threshold, the use of optimal threshold is more suitable as it adapts the input CT image, taking into consideration the wide intensity range of nodules. A total of seven thresholds are used to obtain one image containing regions of interest. The threshold values are listed as follows:  $T_{base}-200$ ,  $T_{base}-100$ ,  $T_{base}$ ,  $T_{base}+100$ ,  $T_{base}+200$ ,  $T_{base}+300$ , and  $T_{base}+400$ . The resultant image preserves nodule opacity information. The separated regions of interest are shown in Fig.4.

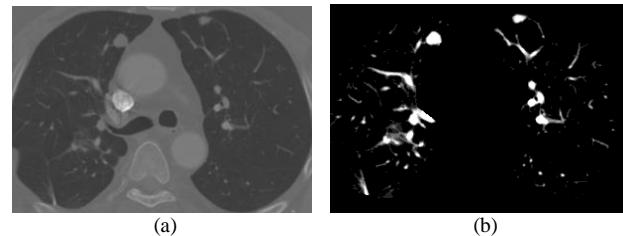


Fig. 4. An example of nodule candidates extraction. (a) Original lung image; (b) Segmented regions of interest

### D. Segmentation of nodular structures

The structures that are segmented from the lung volume through multiple thresholding include nodules, vessels, and the bronchi. For a CAD system to give efficient results, it is important that the nodules are correctly separated from tubular, elongated structures.

For each of the segmented regions, a 3D distance map is calculated. This map assigns ascending weighted values to voxels from the boundary of the region to the centroid giving a measure of the radius of both spherical and cylindrical structures. Spherical structures have a higher distance value at their innermost layer, as compared to oblong structures. The separation is performed by removing the outermost layer of each region. Fig.5 shows a 3D image of a large juxtapavascular nodule that was correctly separated from the attached vascular tree. Finally, shape based features of the nodule candidates, namely, "diameter, area, volume, elongation, and circularity" are computed to prune nodules from the non-nodules.

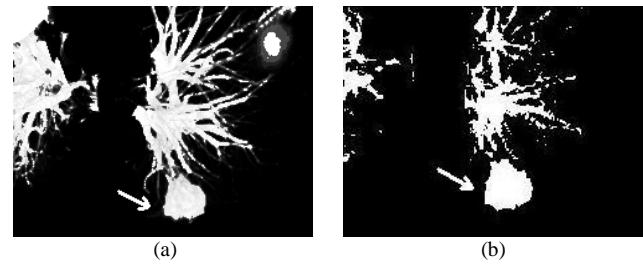


Fig. 5. Nodule segmentation from the attached anatomical structure (a) Nodule and the attached vascular tree; (b) Separated nodule

## III. EXPERIMENTAL RESULTS

The proposed CAD system is tested using the CT scans from the Lung Image Database Consortium (LIDC). LIDC is a publicly available database from the National Biomedical Imaging Archive (NBIA). The nodules in this database are annotated by four radiologists. For nodules that are smaller than 3mm in size, only the approximate centroid location that identifies the nodule is given. But for nodules that are more than or equal to 3mm, the complete boundary outline of the nodule is drawn by each radiologist. In addition, each radiologist classifies the nodule as either benign or malignant. Benign nodules appear solid in the CT images, whereas malignant nodules are manifested as semi-solid and non-solid structures.

To evaluate and compare performance, the algorithm presented by Choi et al. [8] is applied to the CT images. Only the performance of the nodule detection stage was evaluated,

and the classification step is not considered for comparison. The technique provided a nodule detection rate of 60% with a false positive rate of 12.6 per scan. The low detection rate is due to the fact that the CT images contained a large number of non-solid nodules, which were not successfully detected by the system. In addition, juxtapavascular nodules presented a greater challenge because they get wrongly classified as part of the vessel. However, the system provided very good detection results for high intensity solid nodules of varying sizes.

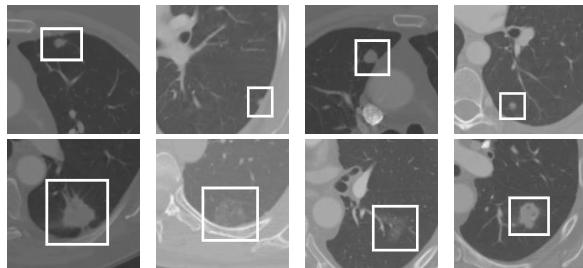


Fig. 6. Examples of detected nodules. Top Row: Nodules detected by both systems. Bottom Row: Previously missed nodules detected by the proposed method

In comparison, the nodule detection method proposed in this paper provides a higher nodule detection rate of 80% and a relatively lower false positive rate of 12.2 per scan. It was observed that the increase in nodule detection rate is a result of successful detection of a higher number of non-solid nodules. In addition, a few juxtapavascular nodules that were missed previously by Choi et al. [8] and Suzuki et al [10] techniques were also detected by the proposed system. The detection performance on high intensity and well-shaped, dense nodules was unchanged. Fig.6 shows examples of detected nodules.

TABLE I. COMPARISON OF OVERALL NODULE DETECTION PERFORMANCE

Methods	Nodule detection rate (%)	Average FPs per scan
Suzuki et al. [10]	80	16.1
Choi et al. [8]	60	12.6
Proposed method	80	12.2

#### IV. CONCLUSIONS

In this paper, a filtering-based method is proposed for the automatic detection of pulmonary nodules. A multi-layered enhancement scheme is applied before extracting potential nodule candidates. 3D disk-shaped kernels of varying sizes sharpen spherical nodules, irregularly shaped nodules with spherical elements, and non-solid nodules. Optimal multiple thresholding segments regions of interest and masking the 3D distance map separates spherical nodules from the attached vessels. Further, rule-based pruning removes false positive regions.

Non-solid nodules can be an indication of malignancy. Therefore, it is important for a nodule detection system to provide good detection results for these nodules. Sharpening non-solid regions provide better detection rate. Another advantage of enhancement is that it de-emphasizes vascular structures. Hence, a decrease in the overall rate of false positives per scan is observed. In addition, juxtapavascular nodules get omitted because they are erroneously classified as

part of the vessel. Masking a 3D distance map successfully separates spherical structures from the attached vessels. The proposed method provides a nodule detection rate of 80% and an average rate of 12.2 false positives per scan. A higher detection performance for non-solid and juxtapavascular nodules is observed.

#### REFERENCES

- [1] A. Jemal, R. Siegel, E. Ward, Y. Hao, J. Xu, M.J. Thun, "Cancer statistics, 2009," *CA: A Cancer Journal for Clinicians*, vol. 59, pp. 225-249, 2009.
- [2] S. Matsuoka, Y. Kurihara, K. Yagihashi, H. Niimi, Y. Nakajima, Peripheral solitary pulmonary nodule: CT findings in patients with pulmonary emphysema, *Radiology* 235 (1) (2005) 266-273, <http://dx.doi.org/10.1148/radiol.2351040674>. URL <http://radiology.rsna.org/content/235/1/266.abstract>.
- [3] K. Awai, K. Murao, A. Ozawa, M. Komi, H. Hayakawa, S. Hori, Y. Nishimura, Pulmonary nodules at chest CT: effect of computer-aided diagnosis on radiologists' detection performance, *Radiology* 230 (2) (2004) 347-352, <http://dx.doi.org/10.1148/radiol.2302030049>. URL <http://radiology.rsna.org/content/230/2/347.abstract>.
- [4] S.G. Armato, M.L. Giger, C.J. Moran, J.T. Blackburn, K. Doi, H. MacMahon, "Computerized detection of pulmonary nodules on CT scans," *Radiographics*, vol. 19, pp. 1303-1311, 1999.
- [5] T. Messay, R. Hardie, S. Rogers, "A new computationally efficient CAD system for pulmonary nodule detection in CT imagery," *Medical Image Analysis*, vol. 14, pp. 390-406, 2010.
- [6] W. Kostis, A. Reeves, D. Yankelevitz, C. Henschke, "Three dimensional segmentation and growth-rate estimation of small pulmonary nodules in helical CT images," *IEEE Trans. Med. Imaging*, vol. 22, no. 10, pp. 1259-1274, 2003.
- [7] M.S. Brown, J.G. Goldin, S. Rogers, H.J. Kim, R.D. Suh, M.F. McNitt-Gray, S.K. Shah, D. Truong, K. Brown, J.W. Sayre, D.W. Gjertson, P. Batra, D.R. Aberle, "Computer-aided lung nodule detection in CT: results of large-scale observer test," *Academic Radiology*, vol. 12, pp. 681-686, 2005.
- [8] W.-J. Choi, T.-S. Choi, "Genetic programming-based feature transform and classification for the automatic detection of pulmonary nodules on computed tomography images," *Information Sciences*, vol. 212, pp. 57-78, 2012.
- [9] Q. Li, F. Li, K. Doi, "Computerized detection of lung nodules in thin-section CT images by use of selective enhancement filters and an automated rule-based classifier," *Academic Radiology*, vol. 15, no. 2, pp. 165-175, 2008.
- [10] K. Suzuki, S.A. ArmatoIII, F. Li, S. Sone, K. Doi, "Massive training artificial neural network (MTANN) for reduction of false positives in computerized detection of lung nodules in low-dose computed tomography," *Medical Physics*, vol. 30, pp. 1602-1617, 2003.
- [11] 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*, vol. 38, pp. 525-534, 2008.
- [12] J. Pu, B. Zheng, J.K. Leader, X.-H. Wang, D. Gur, "An automated CT based lung nodule detection scheme using geometric analysis of signed distance field," *Medical Physics*, vol. 35, no. 8, pp. 3453-3461, 2008.
- [13] S. Ozekes, Rule based lung region segmentation and nodule detection via genetic algorithm trained template matching, *Istanbul Commer. Univ. J. Sci.* 6 (11) (2007) 17-30.
- [14] S.M.B. Netto, A.C. Silva, R.A. Nunes, M. Gattass, "Automatic segmentation of lung nodules with growing neural gas and support vector machine," *Computers in Biology and Medicine*, vol. 42, pp. 1110-1121, 2012.
- [15] X. Ye, G. Beddoe, G. Slabaugh, Graph cut-based automatic segmentation of lung nodules using shape, intensity, and spatial features, in: The Second International Workshop on Pulmonary Image Analysis, 2009, pp. 103-113.
- [16] D. Cascio, R. Magro, F. Fauci, M. Iacomi, G. Raso, "Automatic detection of lung nodules in CT datasets based on stable 3D mass-spring

- models," *Computers in Biology and Medicine*, vol. 42, pp. 1098-1109, 2012.
- [17] Y. Lee, T. Hara, H. Fujita, S. Itoh, T. Ishigaki, "Automated detection of pulmonary nodules in helical CT images based on an improved template matching technique," *IEEE Transactions on Medical Imaging*, vol. 20, pp. 595-604, 2001.
- [18] J. Dehmeshki, X. Ye, X. Lin, M. Valdivieso, H. Amin, "Automated detection of lung nodules in CT images using shape-based genetic algorithm," *Computerized Medical Imaging and Graphics*, vol. 31, pp. 408-417, 2007.
- [19] M. Tan, R. Deklerck, B. Jansen, M. Bister, J. Cornelis, A novel computer-aided lung nodule detection system for CT images, *Med. Phys.* 38 (10) (2011) 5630–5645, <http://dx.doi.org/10.1118/1.3633941>. ISSN: 0094-2405. URL <http://www.ncbi.nlm.nih.gov/pubmed/21992380>.
- [20] Sua rez-Cuenca, W. Guo, Q. Li, Automated detection of pulmonary nodules in CT: false positive reduction by combining multiple classifiers, in: Proceedings of the SPIE, vol. 7963, 2011, pp. 796338–796338-6. URL <http://dx.doi.org/10.1117/12.878793>.