Automated pulmonary nodule detection system in computed tomography images based on Active-contour and SVM classification algorithm

F Shariaty1,2, V V Davydov1,2, V V Yushkova4, A P Glinushkin3 and V Yu Rud1,3
1Department of Electrical Engineering, University of Qom, Qom, Iran
2Peter the Great St. Petersburg Polytechnic University, St. Petersburg 195251, Russia
3All-Russian Research Institute of Phytopathology, Moscow Region 143050, Russia
4Saint Petersburg University of Management Technologies and Economics, 190109, Russia

Abstract. Lung cancer is a common type of cancer that requires early diagnosis. Computer systems by particular different image processing techniques can use for increase the speed and accuracy of lung nodule detection. CT images used in this work in order to process medical images. In this paper proposed an automatic lung nodule detection algorithm using active contour method and SVM classification method. In proposed method, at first in order to achieve better results, lung CT image pre-processing is performed. Then the lung area is segmented by thresholding method followed by some reconstruction techniques to transfer non-isolated nodules into isolated ones. In the next step the nodule candidates are determined using active contour method. Then, nodules are detected by the support vector machine (SVM) classifier using efficient 2D stochastic and 3D anatomical features. In the result, nodules are detected with an overall detection rate of 87%; the number of false positive is 7.5/scan and the location of all detected nodules are recognized correctly.

Shariaty3@gmail.com

1. Introduction
Lung cancer is the most common reason for cancer death, which leads to 26% of cancer death in the United States [1]. Small cell lung cancer (SCLC) and non-SCLC (NSCLC) are two types of lung cancer but 85% of lung cancer is related to NSCLC. Unfortunately, just 17% of the patients will survive 5 years after diagnosing [2].

Early detection of lung nodule considerably increases the survival rate. If the tumor detects while it is smaller than 3cm, the 5-year survival rate may be increased 60%-80% after surgical resection. A Computed-Tomography (CT) image of the lung consists of several slices (2-D images) and the manual examination of these slices by physicians is a time-consuming process. On the other hand, computing accurate values of parameters has the inherent error and is prone to failing to discover all data and imaging information [3-6].

With the advancement of computer systems, they can acquire the experience of radiologists to extract data from medical images. Human analysis is typically subjective and qualitative that involve carelessness. In the meantime, a comparative analysis is needed between an image with nodule and another nodule pattern, and the human observer typically provides a qualitative response. The usage and extraction of quantitative or numerical features of images definitely need exploitation computers.
condition that most analyses that conducted by humans are based on qualitative judgment; they depend on time for a specific observer or from an observer to a different observer [7]. Firstly, because of the lack of diligence or due to inadequate information and, second, due to the range of education and also the level of understanding or ability. Computers can apply a given procedure frequently in brief time and with high exactness. Additionally, data of the many specialists might be completed in a machine method, therefore computers are trained within the specific field by a team of many human specialists [8-11].

Computer Aided Detection (CADe) and Computer Aided Diagnosis (CADx) systems are two types of CAD systems. CADe is a system for locating lesions in medical pictures. While the CADx system performs a diagnosis of lesions, for example, the distinction between benign and malignant tumors. The CADe and CADx systems for the diagnosis of lung cancer in recent decades have been important areas of research. The CADe system does not show the radiological profile of the tumors, and the CADx system does not detect nodules and does not have a good level of automation, as a result, these systems are not widely used in clinical settings [12]. Some CAD systems can be both CADx and CADe; in total, this paper examines CADe systems. The diagnosis of a nodule by the CAD system consists of five important steps [7]: database access, preprocessing, segmentation, analysis, and classification. The block diagram of these steps shown in Fig. 1.

Fig. 1. Five important steps of CAD systems

2. Materials and Methods

In this paper proposed an automatic lung nodule detection algorithm using classification methods. The steps of the proposed algorithm: 1. Preprocessing of CT images, 2. Separate the lung from the surrounding area, 3. Nodule candidates detection, 4. Feature extraction from nodule candidates. 5. Use an SVM classification algorithm in order to detect nodules.

2.1. Preprocessing of CT images

In order to achieve better results in the diagnosis of the lung nodule, lung CT image pre-processing is performed. In pre-processing algorithm, the median filter used with a smoothing algorithm. These two algorithms used to remove the salt and pepper noise and pulse interference and also effectively avoid the fuzzy edges in the image noise removing process.

Figure 2 (a, b). Preprocessing CT Lung Image Processing. a) Original image b) Image with changes in opacity, color and gradient and applying a gradient filter to enhance image quality [13]
2.2. Separate the lung from the surrounding area

Lung extraction is one of the important steps in CAD systems because the lung is a complex organ that includes various structures such as vessels and gaps. Lung extraction also leads to speeds up operations in the system, as well as increasing the accuracy of nodule diagnosis and segmentation. A completely automatic algorithm for recognition lungs in 3D pulmonary X-ray CT images used in order to lung segmentation [14]. This method includes the following major processing steps: First the lung lesion segmented from surrounding area using optimal thresholding method and make a 3D mask (Fig. 3). Then each mask slice compared with the previous slice in order to find nodules attached to the lung wall. In presence of juxta-pleural nodules a reconstruction algorithm consists of several morphological operations has been applied to rectify the mask image (Fig. 4).

![Figure 3](image1)

Figure 3 (a, b, c, d, e). The steps of thresholding algorithm to reach the lung mask: a) Input Image; b) Segmentation with optimal threshold; c) Fill the holes and select largest area; d) Logical “AND” between C and inverse B; e) Fill the holes

The advantage of this method is separation of attached nodules to the lung wall which are removed in ordinary lung segmentation methods.

![Figure 4](image2)

Figure 4 (a, b, c). Segmentation of the lung with juxta-pleural nodule: a) input image; b) segmented image with error; c) segmented image after reconstruction operation.

2.3. Nodule candidates detection

The nodule candidates segmentation is performed by active contour algorithm. In the active contour algorithm, active lines or snakes generate curves that move inside the image to find out the boundaries of the object under the influence of internal and external energies [15]. Active contour algorithm needs an input mask to start segmentation operation. In this work, we want to segment automatically the similar structures to lung nodule. For this propose, the active contour algorithm starts with circular masks which have very small radiuses. After applying these steps, the similar shapes to the lung nodule (nodule candidates) specified (Fig. 5).
2.4. Feature extraction from nodule candidates.

The features of nodule candidates should extract in the next step, in order to classify them into two categories of nodules and non-nodules. Features which are extracted in this work divided to three sets: two-dimensional, three-dimensional and grayscale. Two-dimensional features: Area, Diameter, Perimeter, circularity. Three-dimensional features: volume, compactness, Mean breadth, x-y plane projection compactness. Grayscale features: minimum intensity value inside, mean contrast, variance intensity inside.

2.5. Use an SVM classification algorithm in order to detect nodules.

Many articles have been presented in the field of automatic nodule extraction by computer systems, which mainly use classification methods to divide nodule candidates into two different types of nodule and non-nodule. Classification algorithms typically use features, which described in the previous step, to detect nodules. The SVM classification algorithm used to classify nodule candidates [16]. SVM finds the maximum-margin hyper-plane that divides the data points. The SVM learned with train data, in which the nodules are determined by doctor. The proposed method was verified and evaluated with 20 datasets provided by LIDC (Lung Image Database Consortium).

3. Results and Conclusions

The experimental results indicate high performance of our method with high accuracy of nodule segmentation and an overall detection rate of 87% and FP value of 7.5/scan.

Using computer systems and image processing for medical CT images analysing has progressed a lot, and in general, many published literatures have the potential to use in medical practice. This method also can be used in the clinical practice to help radiologists more accurately and consistently detect and quantitatively assess lung cancers at an early state. It is desirable to use this method with the nuclear magnetic resonance method and optical methods, which are more safety for health compared to CT scans [17-20]. Since there are restrictions for a person on the number of CT scans per month.

Acknowledgment:
This work partly supported by Grant of RFBR No 18-29-25071

References

[1] Tsao A S, Scagliotti G V, Bunn Jr P A, Carbone D P, Warren G W, Bai C, ... & Adusumilli P S 2016 Journal of Thoracic Oncology 11(5) 613-638
[2] Molina J R, Yang P, Cassivi S D, Schild S E and Adjei A A 2008 In Mayo Clinic Proceedings 83(5) 584-594 Elsevier
[3] Nesbitt J C, Putnam Jr J B, Walsh G L, Roth J A and Mountain C F 1995 Annals of thoracic surgery 60(2) 466-472
[4] Humphrey L L, Teutsch S and Johnson M 2004 Annals of Internal Medicine 140(9) 740-753
[5] Flehinger B J, Kimmel M and Melamed M R 1992 Chest 101(4) 1013-1018
[6] Zhang Y, Oikonomou A, Wong A, Haider M A and Khalvati F 2017 Scientific reports 7 46349
[7] Shariaty F and Mousavi M 2019 International journal of Informatics in Medicine Unlocked 15(C) 100173
[8] Firmino M, Angelo G, Morais H, Dantas M R and Valentim R 2016 Biomedical engineering online 15(1) 2

[9] Sprindzuk M V, Kovalev V A, Snezhko E V and Kharuzhyk S A 2010 Polish journal of radiology 75(1) 67

[10] Doi K 2005 The British journal of radiology 78(1) 3-19

[11] Van Ginneken B, Schaefer-Prokop C M and Prokop M 2011 Radiology 261(3) 719-732

[12] Firmino M, Angelo G., Morais H, Dantas M R and Valentim R 2016 Biomedical engineering online 15(1) 2

[13] Kim H, Nakashima T, Itai Y, Maeda S, kooi Tan J and Ishikawa S 2007 Proceedings of the 2007 International Conference on Control, Automation and Systems p. 1274-1277

[14] F. Shariaty, S. Hosseinlou1, V. V. Davydov, V. Yu. Rud (2019). Automatic Lung Segmentation Method in Computed Tomography Scans. International Conference "Emerging Trends in Applied and Computational Physics 2019".

[15] Vese L A and Chan T F 2002 International journal of computer vision 50(3) 271-293

[16] Vapnik V 1998 Statistical Learning Theory (John Wiley&Sons. Inc., New York)

[17] Davydov V V, Dudkin V I, Velichko E N and Karseev A Yu 2015 Journal of Optical Technology (A Translation of Optichekii Zhurnal) 82(3) 132-135

[18] Davydov R V, Atonov V I and Moroz A V 2018 Proceedings of the 2018 IEEE International Conference on Electrical Engineering and Photonics, EExPolytech 2018 (Saint-Petersburg) 8564378 p. 236-239

[19] Myazin N S, Davydov V V, Yushkova V V and Rud V Yu 2018 Journal of Physics: Conference Series 1038 (1) 012088

[20] Davydov V V, Velichko E N, Myazin N S and Rud’ V Yu 2018 Instruments and Experimental Techniques 61 116–122