ISSN 2278-3083

International Journal of Science and Applied Information Technology (IJSAIT), Vol. 3 , No.3, Pages : 16 - 19 (2014)
Special Issue of ICCET 2014 - Held during July 07, 2014 in Hotel Sandesh The Prince, Mysore, India

Automation of Lung Nodule Segmentation Using Artificial Neural Network in CT Lung Images

S.Saravanan[1], G.Selvakumar[2], C.Amarnath[3], S.Manikandan[4]
[1]Associate Professor, PERI Institute of Technology, Chennai-600048, India,ss_saravanan@yahoo.com
[2]Professor and Vice Principal, Excel Engineering College, Komarapalayam-637303, India
[3]Professor and Head, Department of Radialogy, Stanley Medical College, Chennai-600001
[4]Student, Department of Computer Science, Narasu's Sarathy Institute of Technology, Tamil-Nadu-636305, India

Abstract *:* Lung cancer is the severe disease where 1.61 million, 12.7% of the total people are affected in India and abroad as per GLOBOCAN series Published by the International Agency for Research on cancer recent. The data load for Radiologists is increased due to current computed tomography (CT) technology which allows for near isotropic, sub-millimeter resolution acquisition of the complete chest in a single breath hold. These thin-slice chest scans have become indispensable in thoracic radiology. Automating the analysis of such data is therefore, a necessity and this has created a rapidly developing research area in medical imaging. Segmentation of Region of Interest (ROI) is often a necessary first step to computer analysis. This paper presents a novel method of automizing the segmentation process of the lung nodules in computer aided diagnosis (CAD).

Key words : CAD-Computer Aided Diagnosis, ROI-Region of Interest, GLCM –Grey Level Co-Occurrence matrix, ,Lung Cancer.

INTRODUCTION

Computer tomography (CT) for the body has been available since 1975. Originally, CT was not considered a technique particularly well suited for thorax. Low resolution resulted in large difference in attenuation values between tissue and air, which is currently a main reason for the effectiveness of CT in thorax imaging, made it difficult to correctly interpret small lesions. In neuro radiology it is well proved that CT had revolutionized and its impact in abdominal and pelvic imaging had been similarly great, but 'the ultimate role of computer tomography in the study of diseases of the chest is not as certain". Two advances in particular have had the most repercussions for CT of the chest. About twenty years ago, improved axial resolution made HRCT scans possible. Slices of 0.4 mm thick could provide anatomical detail of the lungs similar to that available from gross pathological specimens. However, scanner speed limitations at the time meant that a 1-cm gap between slices was necessary to cover the entire thorax while avoiding breathing artifacts. This limitation has been effectively removed in the last decade with the advent of multi-detector-row scanners that can acquire more number of slices per scan per second.

In CAD (Computer Aided Diagnosis), Segmentation is often a necessary first step to computer analysis. In detecting the nodule candidates, the following techniques have been reported : multiple grey level thresholding , mathematical morphology, genetic algorithm template matching of Gaussian spheres and disks, clustering, connected component analysis of threshold images , thresholding , detection of (half) circles in threshold images , grey level distance transform and filters enhancing (spherical) structures. This paper presents an algorithm which learns the patterns of a nodule while manual segmentation and feature extraction under supervisory mode. Apart from learning the algorithm reduces the false positives also thus eliminating the unnecessary Biopsy.

MATERIALS AND METHODS

Fig. 1 represents the algorithm for automation of lung nodules. Each block in the algorithm is explained extensively in the forthcoming sections.

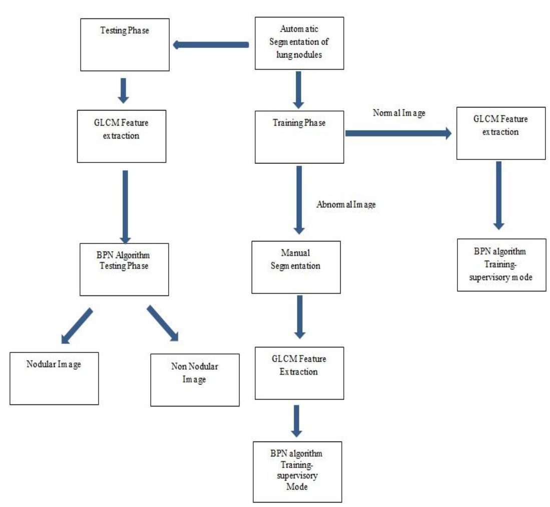

Fig 1: Algorithm for the Materials and Methods

Manual Segmentation

Any computer system that analyses the lungs and does not work on manually delineated Regions of interest must incorporate automatic lung segmentation. The importance of accurate segmentation is illustrated as a pre-processing step in a CAD scheme [1]. In a nodule detection setting, they showed that 5%-17% of the lung nodules in their test data

ISSN 2278-3083

International Journal of Science and Applied Information Technology (IJSAIT), Vol. 3 , No.3, Pages : 16 - 19 (2014)
Special Issue of ICCET 2014 - Held during July 07, 2014 in Hotel Sandesh The Prince, Mysore, India

were missed due to the pre-processing segmentation, depending on whether or not the segmentation algorithm was adapted specifically to the nodule detection task [2].

- To develop an algorithm for automatic segmentation, first step is to perform Manual Segmentation using any of the preferred method.
- To determine the textural features of the node.
- To attain the neural network algorithm for segmenting nodular and non-nodular images.

On testing phase if an unknown image is given the algorithm will classify and segment nodular and non-nodular image. For manual segmentation Region Growing Algorithm is followed.

Feature Extraction

Grey-level Co-occurrence Matrix texture measurements, have been the workhorse of image texture, since they were proposed [3]. Texture refers to ‘the spatial distribution of tonal variations’. It is discussed that there are two approaches to textural description: Statistical and Structural [4]. Statistical Models obtain image texture measurements by calculating the spatial dependency of grey-level tones, while structural models describe texture as primitives and their attributes and relationships. Statistical approaches for texture are more widely used in image analysis of medical images because statistical textural features are easier to calculate than structural primitives, with the exposition of analysis of edges as a structural approach.

There are several ways to quantify texture of an image, such as the GLCM, Grey-Level Run Length Matrix (GLRLM) and Gray-Level Difference Vector (GLDV) [4]-[6]. In order to calculate texture features for each pixel in an image, a moving window is usually used to define the neighborhood of a pixel, and the texture measurement calculated from the window is assigned to the center pixel. Among many methods, the GLCM method was found to be more effective.

GLCM is a matrix derived from the grey-level image, which shows the joint-probability of distribution of a pair of grey levels, separated at a certain distance and a certain orientation. Eight texture features, evaluated in this study are calculated from the GLCM method and shown in Table I.

Table I : Texture Features and their Equations

Se.No	Textural Features	Equations
1.	Homogeneity (HOM)	$\sum_{i=0}^{N-1}\sum_{j=0}^{N-1}\frac{P_{i,j}}{1+(i-j)^2}$
2.	Contrast (Con)	$\sum_{i=0}^{N-1}\sum_{j=0}^{N-1}P_{i,j}(i-j)^2$
3.	Dissimilarity (DIS)	$\sum_{i=0}^{N-1}\sum_{j=0}^{N-1}p_{i,j}\lvert i-j\rvert$
4.	Mean (MEAN)	$\sum_{i=0}^{N-1}\sum_{j=0}^{N-1}ip_{i,j}$
5.	Standard Deviation (SD)	$\sigma_I=\sqrt{\sum_{i=0}^{N-1}\sum_{j=0}^{N-1}p_{i,j}(i-\mu_I)^2}$
6.	Entropy(ENT)	$\sum_{i=0}^{N-1}\sum_{j=0}^{N-1}-p_{i,j}\ln p_{i,j}$
7.	Angular Second Moment(ASM)	$\sum_{i=0}^{N-1}\sum_{j=0}^{N-1}P_{i,j}^2$
8.	Correlation (COR)	$\sum_{i=0}^{N-1}\sum_{j=0}^{N-1}\frac{(i-\mu_I)(j-\mu_J)}{\sigma_I\sigma_J}$

Where N is the number of grey levels, P is the normalized symmetric GLCM of dimension NXN; and $P_{i,j}$ is the (i,j)th element of P.

$$\mu_j=\sum_{i=0}^{N-1}\sum_{j=0}^{N-1}jp_{i,j} \quad \text{and} \quad \sigma_j=\sqrt{\sum_{i=0}^{N-1}\sum_{j=0}^{N-1}p_{i,j}(j-\mu_j)^2}$$

HOM measures local homogeneity and results in a large value if the elements of the GLCMs are concentrated on the main diagonal. CON measures local spatial frequency if the GLCM has large off-diagonal elements, the local window has high contrast. DIS is similar to CON-high contrast of the local window indicates high DIS value. MEAN and SD measure the mean and the standard deviation in terms of the GLCM. ENT measures disorder of the image, while ASM indicates local uniformity. Window size, texture statistics, quantization level, the inter-pixel distance and orientation, and the spatial resolution of the image are the factors that contribute to the performance of textural statistics from GLCM method [7].

General discussion on the selection of GLCM texture features

The performance of the eight GLCM texture features is evaluated in this study. Since, some of the texture features are highly correlated, usually we do not need all of them to perform and classification. The question is, then, how do we select from the eight. Though this study does not focus on which texture features are better than others, there are some guidelines. The usefulness of ASM, CON, COR, and ENT among the set of texture features is demonstrated [3]. Three GLCM texture features including ASM, CON and COR were found to be less correlated in the study [8]. The eight GLCM texture features [9] used in this study is divided into three groups:

- The Contrast Group (CON, DIS and HOM)
- The Orderliness Group (ASM and ENT) and the
- Descriptive Statistics Group (MEAN, SD and COR)

The texture features in the Contrast group are correlated with each other, so are the features in the orderliness group. MEAN and COR are generally not correlated with other features. She suggested that researchers choose one measure from the contrast group, one from the orderliness group and two (or) three from the descriptive statistics group for classification purposes. This would result in a combination of four (or) five texture features that capture different aspects of the GLCM.

Training Phase

After extracting features from the nodules, the feature vector is given as input to the BPN algorithm and trained in supervisory mode. For nodule image, output is given as 1 and for non-nodule image the output is given as 0. For normal image, output is given as 0.

Testing Phase

During testing phase, GLCM matrix is calculated for every pixel and the features are calculated from the matrix. These features are given as input to the BPN algorithm. If the output is closer to 1 it is nodule pixel and this pixel is tagged separately and kept in jpg format as nodule. If the output of the algorithm is closer to 0, the pixel is tagged separately and kept in jpg format as non-nodule. If the input image is normal image then also the output of the algorithm is closer to zero. In that case dark image is got as output in nodule image.

RESULTS AND DISCUSSIONS

Fig. 2, Fig. 5, and Fig. 8 are the abnormal input images to the Region Growing Algorithm for Nodular segmentation. After segmentation, the nodular image Fig. 4, Fig. 7, and Fig. 10 and non-nodular image Fig. 3, Fig. 6, and Fig. 9 is fed into the GLCM algorithm and the textural features are calculated. These features are given as input to the BPN algorithm by giving 0 and 1 as output under supervisory mode. For normal lung image the output is given as 0 and the network is trained. If the unknown image is given and the network results in the value closer to zero then it is non-nodular pixel and if it is closer to one then it is nodular pixel and it is calculated for entire image. The threshold range is selected by trial and error method.

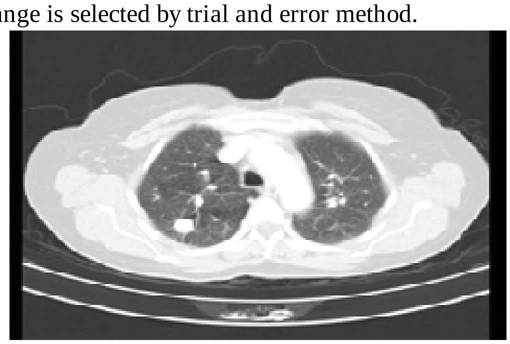

Fig 2: Lung CT image.

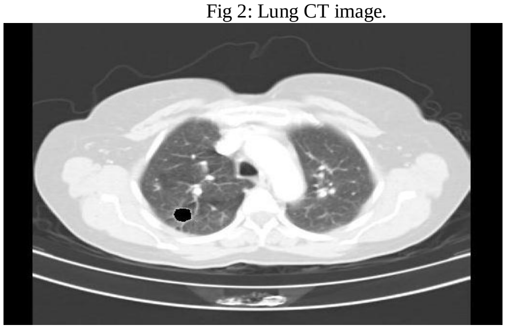

Fig 3: Lung CT image without nodule from Right Lung.

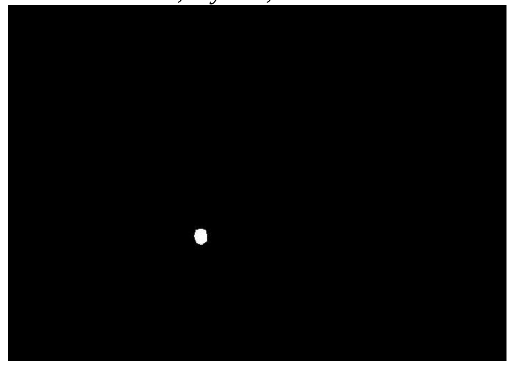

Fig 4: Segmented Lung Nodule.

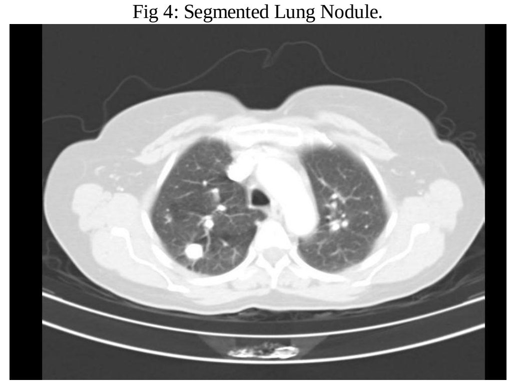

Fig 5: Lung CT image

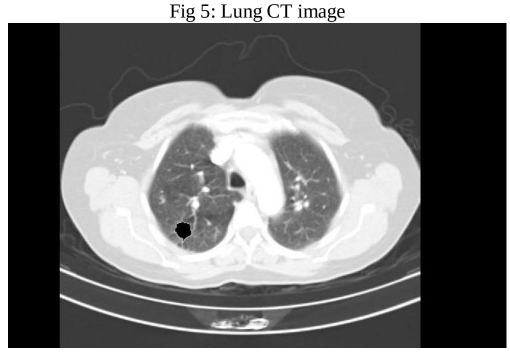

Fig 6: Lung Image without Nodule.

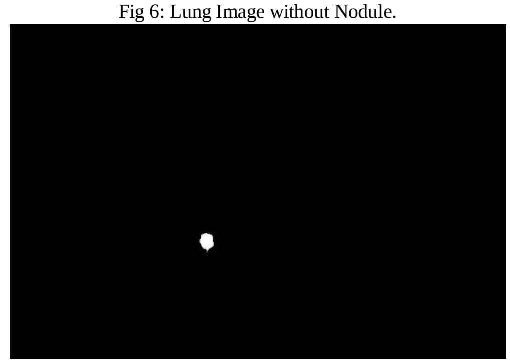

Fig 7: Segmented Nodule from CT image.

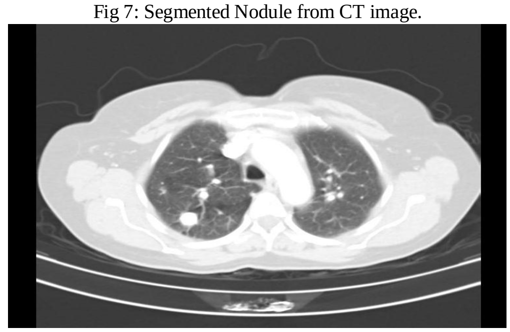

Fig 8: CT Lung image.

ISSN 2278-3083

International Journal of Science and Applied Information Technology (IJSAIT), Vol. 3 , No.3, Pages : 16 - 19 (2014)
Special Issue of ICCET 2014 - Held during July 07, 2014 in Hotel Sandesh The Prince, Mysore, India

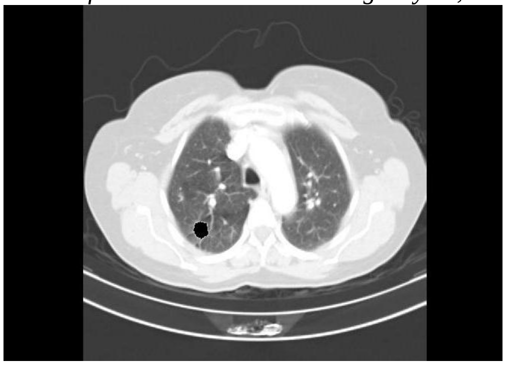

Fig 9: CT lung image without Nodule.

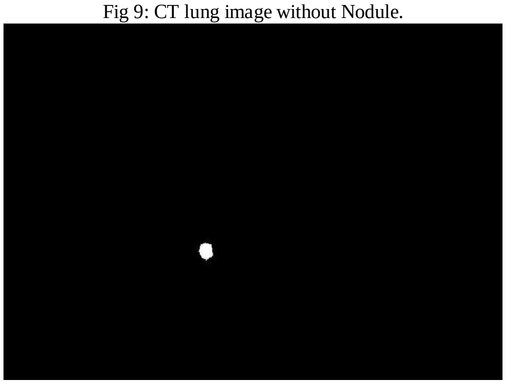

Fig 10: Segmented Nodule.

CONCLUSION

As per the general discussion on the selection of GLCM texture features various combinations were explored and the combination of CON. ASM, MEAN, SD AND COR have given good segmentation results.

FUTURE WORK

Automatic segmentation of pleural tail and sub-pleural image has to be done. Further, features have to be extracted for Malignancy and Benignity.

REFERENCES

[1] S.G. Armato, and W.F. Sensakovic, " Automated lung segmentation for thoracic CT impact on computer-aided diagnosis," Academics of Radiology, vol.11, no . 9, pp. 1011-1021, Sep 2004.

[2] I.Sluimer, A. Schilham,M. Prokop, and B. Van Ginneken, "Computer Analysis of Computed Tomography Scans of the Lung: A Survey," IEEE Trans. on Medical Imaging ,vol. 25,no.1, pp. 385-405, Apr 2006.

[3] R.M. Haralick, "Statistical and Structural approaches to texture," Proceedings of the IEEE, vol. 67, no.5, pp. 786-804, June 2005.

[4] R.M. Haralick, K. Shanmugam, and I. Dinstein, "Textural Features for Image classification," IEEE Trans. on Systems, Man and Cybernetics, vol. 3, no. 6, pp. 610-621, Nov 2007.

[5] J.S. Weszka, C.R. Dyer, and A. Rosenfeld, " A Comparitive study of texture measures for terrain classification," IEEE Trans. on Systems, Man and Cybernetics, vol. 6,no. 4, pp. 269-285, Feb 2010.

[6] R.W. Conners, and C.A. Harlow, " A Theoretical comparison of texture algorithms," IEEE Trans. on Pattern Analysis and Machine Intelligence,vol.2, no.2, pp. 204-222, Jan 1980.

[7] J. Zhang, P. Wang, P. Gong, and P. Shi, "Study of urban spatial patterns from SPOT panchromatic imagery using textural analysis," International Journal of Remote Sensing, Taylor Francis Online, vol. 24, pp. 4137-4160, June 2010.

[8] P. Gong, T.J. Marceau, and P.J. Howarth, "A Comparison of spacial feature extraction algorihms for land-use classification with SPOT HRV data," Remote Sensing of Environment,Elsevier, vol.40, no.2, pp.137-151, May 1992.

[9] M. Hall-Beyer. (Feb 2000). The GLCM Texture Tutorial. Available : http://www.fp.ucalgary.ca/mhallbey/tutorial.htm.