OPEN ACCESS at journalijcir.com

Research Article

ISSN: 2395 -5775

PREDICTION OF LUNG CANCER USING FUZZY INFERENCE SYSTEM

ShristiTiwari*, DeeptiChoudhary and Shubi Sharda

Abstract

The most leading cancer among all in human beings is lung cancer. There are more than 17% percent of the total cancer related deaths so the fast and early detection of lung cancer can help in a sharp decrease in the lung cancer death rate. The CT image of lung helps us to find the presence of lung cancer. Doctor analyses the CT image and predicts the presence of cancer nodule. There is a chances of false recognition in the manual identification of cancer. So there is a need of automated approach of lung cancer detection Image processing technique can be used for this purpose. In this paper we propose a Lung cancer identification system that uses a fuzzy inference system to spot the most prominent cancer cells. The approach has four stage to detect the existence of cancer nodule in lung. Pre-processing stage, Segmentation stage, feature extraction stage and fuzzy inference rules to identify lung cells. Pre-processing step includes image enhancement. Enhanced CT image of lung is then passed through segmentation phase. From the segmented output features are extracted to predict the existence of abnormality of lung. On these extracted features fuzzy rules are applied to identify the possibility of cancer cells.

Key Words: Lung nodule, Fuzzy Inference System, Image Segmentation.

INTRODUCTION

Cancer is now the biggest reason of death in the world. Lung cancer is one of the most common cancers in present days. Due to the lifestyle of people there is a steady increase in cancer patient. Pain, breathlessness, cough, weight loss and fatigue are the general symptoms of cancer. Survival from the disease is not easy if it is not identified at the early stage. Only 15% of lung cancer is recognized at the early stage [1]. The main reason because of which it is difficult to find lung cancer in early stages is that there is only a dime-sized lesion growth called nodule, inside the lung, and by the time it is identified it is already too late for the patient. Also, these small lesions cannot be identified by X-rays and are only noticed by a CT scan. Even after the detection, it takes a considerable amount of effort and experience of radiologists to identify and mark the nodules as benign or as a probable case of malignancy. Normally the lung cancer can be divided into two groups. First one is non-small cell lung cancer and second one is small cell lung cancer.

The CT image of lung is used to diagnose the lung cancer. Normally by a doctor, the CT image of lung is analyses first and then they detect the presence of cancer in lung. False detection is very big problem which are faced in this approach. False detection is caused due to following reason like the presence of air in bronchi, presence of ribs and blood vessels etc [2]. So there is a need to develop approach of automated identification of cancer. Image processing tools are the best tool for developing such aautomated approach for lung cancer identification. CT image of lung is processed and finds whether the presence of cancer nodule is ther or not.There are many image processing tools [3] [4] for this purpose. This paper focuses to build an efficient and accurate automated approach for lung cancer detection. This paper proposed a Lung cancer detection system that uses a fuzzy inference system to identify the cancer cells.

This approach has four stages to detect the presence of cancer nodule in lung. The first one is Pre-processing stage, then the Segmentation is done, then the features of cancer cells are extracted in next stage i.e. in feature extraction stage and fuzzy inference rules to identify lung cells. Pre-processing step is used to enhance the image. Enhanced CT image is then passes through segmentation phase. From the segmented image some features are extracted to guess the existence of abnormality of lung. On these extracted features fuzzy rules are applied to identify the possibility of cancer cells.

Many of the different methods have been proposed and used to detect and classify lung cancer in CT images using different algorithms. For example,

Department Of Computer Science & Engineering Kalinga University, Naya Raipur, (C.G), India
Correpondence and Reprint Requests: ShristiTiwari
Received: April 19, 2016 | **Accepted:** May 04, 2016 | **Published Online:** June 28, 2016

Conflict of interest: None declared | **Source of funding:** Nil

Camarlinghi *et al.* [5] have used three different computer aided detection techniques for identifying pulmonary nodules. Abdulla and Shaharum [6] used feed forward neural networks to classify lung nodules in X-Ray images even if with only a small number of features such as area, perimeter and shape. Kuruvilla *et al.* [7] have taken six distinct parameters including skewness and fifth & sixth central moments extracted from segmented single slices containing two lungs along with the features mentioned in [6].

This paper reports a high detection rate of 88.5% with a average of 6.6 false positives (FPs) per CT scan on 15 CT scans. Hayashibe *et al* [9] projected an automatic method based on the subtraction between two serial mass chest radiographs, which is used in the detection of new lung nodules. Kanazawa *et al* [10] presented a system which extract and used to analyze the features of the lung and pulmonary blood and then utilizes defined rules to perform identification. Then it was used in the detection of tumor candidates from helical CT images. Naseer Salman [11] suggested that Marker-driven watershed segmentation extracts seeds indicating the presence of objects or background at particular image locations. The marker locations then set to be regional minima within the topological surface. Then the watershed algorithm is applied. Mori *et al* [13] projected a method to extract bronchus area from 3-D chest X-ray CT images, which is used in a virtualized bronchoscope system.

METHODOLOGY

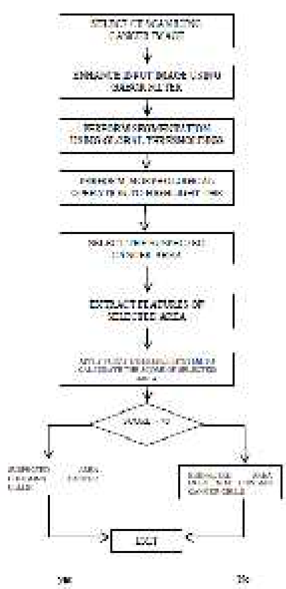

Figure 1

Step 1 Select Input Image Input image is selected from Database of CT scan lung cancer images taken from website of National Cancer Institute which contains Lung Image Database Consortium (LIDC) and Image Database Resource Initiative (IDRI)[14]. All the images are indicom format (Digital Imaging and Communications in Medicine) so conversion has been performed to convert it into JPG format.

The image database contain 3 types of images

- Images having no nodule (no cancer cell)
- Images having nodule size <= 3mm.
- Images having nodule size >3 mm.

Step 2 Enhancement image using Gabor Filter

Gabor filter is named after its developer Dennis Gabor. It is a linear filter used for edge detection. In Gabor filter Frequency and orientation representations very much are identical to those of the human visual system. And it is found that it is particularly appropriate for texture representation and discrimination. In the spatial domain, a 2D Gabor filter is a Gaussian kernel function multiplied by a sinusoidal plane wave. The Gabor filters are self-similar: all filters can be formed from one mother wavelet by dilation and rotation [13].

Based on Gabor function, the image presentation obtained consist of an excellent local and multi-scale decomposition in terms of logons that are specially localization in space and frequency domains. Gabor filter is a type of linear filter whose impulse response is defined by a function called harmonic function and then it is multiplied by a Gaussian function. Due to the property of multiplication-convolution, the Fourier transform of a Gabor filter's impulse response is the convolution between two function one is the Fourier transform of the harmonic function and other is the Fourier transform of the Gaussian function [11].

Step 3 Segmentation Using Thresholding

Three sholding is the most dominant tool for image segmentation. Thresholding operation converts the grey scale image into binary image. In this operation, we selects a threshold value T and allocate or set the two levels to the images that is one is above the threshold value and the other is below the threshold value. Now the object can be separated from the background by the help of using the threshold value T. Taking any point (x,y) for which if $f(x,y) > T$ is called an object point, otherwise the point is a background point.Select a global threshold value for the entire CT Lung image. By Appling the threshold value to the preprocessed image, it can be converted to binary and the threshold image is obtained.

$$g(x, y) = 1, \text{ if } f(x, y) > T$$
$$g(x, y) = 0, \text{ if } f(x, y) \leq T$$

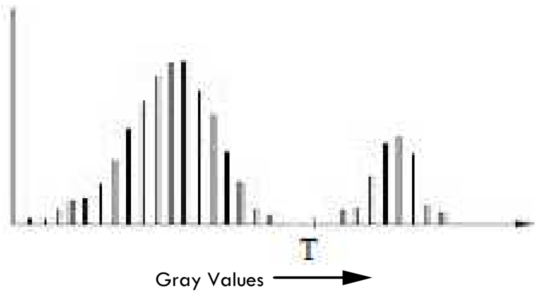

Figure 2

Step 4: Morphological Operation

To fill in holes and small gaps in the image morphological close operation is applied on the threshold image. Reserve the block whose area is the largest and set the others to zero using 8-connected neighbors. The binary lung mask is obtained using the defined step Extract the lung boundary by setting a pixel to 0. If it is 4-connected neighbors are all 1's, thus leaving only edge pixels. Original Lung CT image is multiplied with the lung masked image to get the final segmented lung area with gray level values as those of original image.

Step 5 Select Region of Interest

A part of an image that we want to filter or perform some other operation on is known as Region of interest. An ROI is a type of a binary image which is the same size as the image we want to process. The pixels that define the ROI are set to 1 and all other pixels set to 0.Here, our region of interest is suspected cancer area, Random selection is performed by user as region of interest. Then feature is extracted from selected region to identify probability of cancer in selected region.

Step 6 Feature Extraction

Texture Features The most frequently used measures of texture, in particular of random texture, are the statistical measures planned by Haralick. Some of the Haralick's measures may not be directly related to the intersecting structures, speculations, and node-like patterns of distortion, but they may provide useful information regarding the properties of the given ROI or image. Haralick's texture measures is usually rely on the moments of a joint probability density function, were this function is estimated using the joint occurrence or co-occurrence of gray levels, called as the gray-level co-occurrence matrix(GCM), o by the help of various directions and distances. GCMs are calculated with unit pixel distance for the angles of 0, 45, 90 and 135. So total 4 GCM features can be calculate as texture feature of image in different direction.

Local Binary Pattern (LBP)

The local binary pattern (LBP) feature has a significant importance in the field of texture classification and retrieval. Ojala*et al.* proposed LBPs, which are converted to a rotational invariant version for texture classification.

Area and Perimeter

Area One of the main features of cancer cell is its area and perimeter. There are two values of segmented images, 1 and 0. Nodule part will be represented with value 1. Then by finding number of pixel with value, the area can be calculated.
Area = ∑ White Pixel in image I(x,y)

Perimeter If the pixel value in nonzero then it is a part of the perimeter and it is connected to at least one zero-valued pixel. For two dimensions the default value is 4 and for three dimensions is 6.

Perimeter = ∑ Pixels in the Boundary of Image I(x,y)

Step 7 Apply Fuzzy Inference System

Using all above feature fuzzy rules are applied to identify that suspected area may contain cancer cells or not, if fuzzy output score is greater than 70 then suspected area having highest chance of cancer cells, if it is in between 50 to 70 then it has moderate chance of cancer and score less than 50 has no chance of cancer.

RESULT

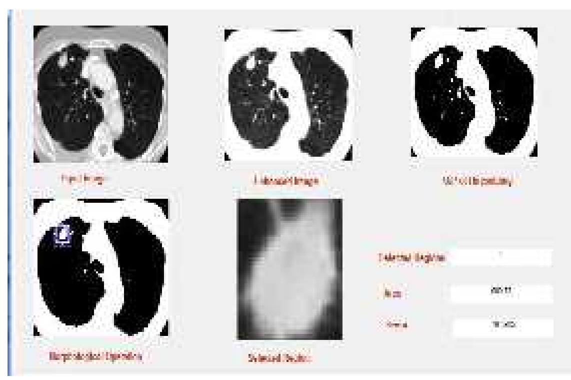

Figure 3 Output score of cancer region

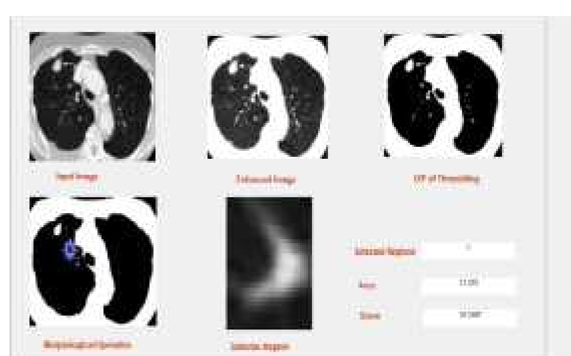

Figure 2 Output score of non-cancer region

CONCLUSION

The results shows for lungs affected by cancer the fuzzy interference system provides appropriate output. For improving the cancer detection in lungs new and improved algorithm can be developed to obtain the better detection of cancer in lungs.

References

1. American Cancer Society, "Cancer Statistics, 2005", CA: A Cancer Journal for Clinicians, 55: 10-30,2005, "http: //caonline. Amcancer soc. org/cgi/content/full/55/1/10".
2. D. S. PATIL and M. Kuchanur, "Lung cancer classification using image processing," *International Journal of Engineering and Innovative Technology* (IJEIT) Volume, vol. 2, 2012.

A. Chaudhary and S. S. Singh, "Lung cancer identification on ct images by using image processing," in Computing Sciences (ICCS), 2012 International Conference on. IEEE, 2012, pp. 142-146.

3. N. Hadavi, M. Nordin, A. Shojaeipour *et al.*, "Lung cancer diagnosis using ct-scan images based on cellular learning automata," in Computer and Information Sciences (ICCOINS), 2014 International Conference on. IEEE, 2014, pp. 1-5.
4. N. Camarlinghi, I. Gori, A. Retico, R. Bellotti, P. Bosco, P. Cerello, G. Gargano, E. L. Torres, R. Megna, M. Peccarisi *et al.*, "Combination of computer-aided detection algorithms for automatic lung nodule identification," *International journal of computer assisted radiology and surgery*, vol. 7, no. 3, pp. 455–464, 2012. [6] A. A. Abdullah and S. M. Shaharum, "Lung cancer cell classification method using artificial neural network," Information Engineering Letters, vol. 2, no. 1, pp. 49–59, 2012.
5. J. Kuruvilla and K. Gunavathi, "Lung cancer classification using neural networks for ct images," Computer methods and programs in biomedicine, vol. 113, no. 1, pp. 202–209, 2014.
6. R. Bellotti, F. De Carlo, G. Gargano, S. Tangaro, D. Cascio, E. Catanzariti, P. Cerello, S. C. Cheran, P. Delogu, I. De Mitri *et al.*, "A cad system for nodule detection in low-dose lung cts based on region growing and a new active contour model," Medical Physics, vol. 34, no. 12, pp. 4901–4910, 2007.
7. Hayashibe R, *et al.* "Automatic lung cancer detection from X-ray images obtained through yearly serial mass survey." in International Conference on Image Processing. 1996.
8. Kanazawa K, Kubo M and N. N. "Computer aided diagnosis system for lung cancer based on helical CT images." in the 13th International Conference on Pattern Recognition. 1996.
9. N Salman "Image Segmentation Based on Watershed and Edge Detection Techniques". The International Arab Journal of Information Technology Vol. 3, No. 2, April 2006.
10. A Kumar, P Kumar, "A New Framework for Color Image Segmentation Using Watershed Algorithm."Computer Engineering and Intelligent Systems, Vol 2, No.3, ISSN 22221719 (Paper) ISSN 2222-2863 (Online).
11. Mori K, KitasakaT, HagesawaJI, Toriwaki JI *et al.* "A Method for Extraction of Bronchus Regions from 3D Chest X-ray CT Images by Analyzing Structural Features of the Bronchus." in the 13th International Conference on Pattern Recognition. 1996. Page 69.
12. 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.

****