LUNG CANCER DETECTION ON CT IMAGES BY USING IMAGE PROCESSING

Anita chaudhary
Dept. of Electronics and Communication Engineering
Lovely Professional University, Phagwara
Punjab, India
Anitachaudhary111@gmail.com

Sonit Sukhraj Singh
Dept. of Electronics and Communication Engineering
Lovely Professional University, Phagwara
Punjab, India
Enggsonit7@gmail.com

***Abstract*—Lung cancer seems to be the common cause of death among people throughout the world. Early detection of lung cancer can increase the chance of survival among people. The overall 5-year survival rate for lung cancer patients increases from 14 to 49% if the disease is detected in time. Although Computed Tomography (CT) can be more efficient than X-ray. However, problem seemed to merge due to time constraint in detecting the present of lung cancer regarding on the several diagnosing method used. Hence, a lung cancer detection system using image processing is used to classify the present of lung cancer in an CT- images. In this study, MATLAB have been used through every procedures made. In image processing procedures, process such as image pre-processing, segmentation and feature extraction have been discussed in detail. We are aiming to get the more accurate results by using various enhancement and segmentation techniques.**

Keywords-CT, LCDS, Watershed Segmentation, ROI, Thresholding, morphologic, Metastasis.

I. INTRODUCTION

The mortality rate of lung cancer is the highest among all other types of cancer. Lung cancer is one of the most serious cancers in the world, with the smallest survival rate after the diagnosis, with a gradual increase in the number of deaths every year. Survival from lung cancer is directly related to its growth at its detection time. But people do have a higher chance of survival if the cancer can be detected in the early stages[1]. Lung cancer can be divided into two main groups, non-small cell lung cancer and small cell lung cancer. These assigned of the lung cancer types are depends on their cellular characteristics. As for the stages, in general there are four stages of lung cancer; I through IV. Staging is based on tumor size and tumor and lymph node location. Presently, CT are said to be more effective than plain chest x-ray in detecting and diagnosing the lung cancer. The earlier the detection is, the higher the chances of successful treatment. An estimated 85% of lung cancer cases in males and 75% in females are caused by cigarette smoking[2].

In 2005, approximately 1,372,910 new cancer cases are expected and about 570,280 cancer deaths are expected to occur in the United States. It is estimated that there will be 163,510 deaths from lung cancer, which forms 29% of all cancer deaths. The overall survival rate for all types of cancer is 63%. Although surgery, radiation therapy, and chemotherapy have been used in the treatment of lung cancer, the five-year survival rate for all stages combined is only 14%. This has not changed in the past three decades [2]. The purpose of this paper is to find the early stage of lung cancer and more accurate result by using various enhancement and segmentation techniques.

II. METHODOLOGY

Overall, there are three main processes used throughout the report; Pre-processing, feature extraction and finally the classification process. MATLAB is used in every process made throughout the project. Process involved in the lung cancer detection system for the project can be view in Figure 1.

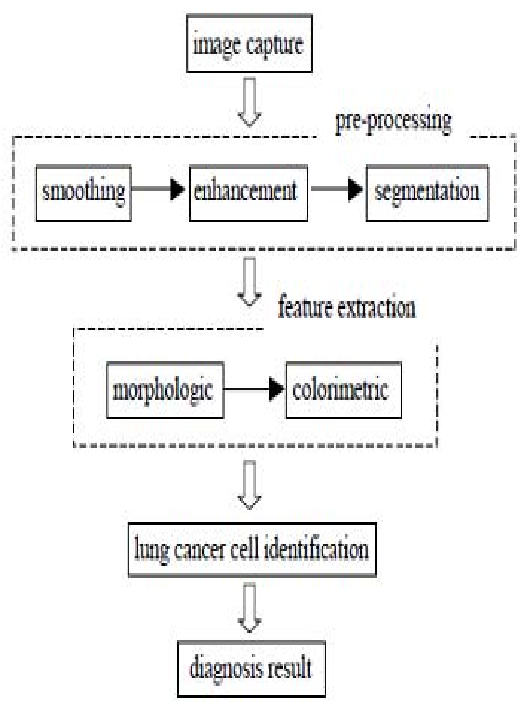

Figure.1 Entire diagnosis process of LCDS[1].

LCDS system uses convolution filters with Gaussian pulse to smooth the cell images. The contrast and color of

978-0-7695-4817-3/12 $26.00 © 2012 IEEE
DOI 10.1109/ICCS.2012.43

the images are enhanced. Then the nucleuses in the images are segmented by thresholding. All of those are simple digital image processing techniques. After that, LCDS utilizes morphologic and colorimetric techniques to extract features from the images of the nucleuses [3]. The extracted morphologic features include the average intensity, area, perimeter and eccentricity of the nucleuses. On this basis, a lung cancer cell identification nodule is employed to analyze those features to judge whether cancer cells exist in the specimens or not. Moreover, if there are cancer cells, the cancer cell type is identified. The entire diagnosis process of LCDS is shown in Figure 1.

A. *Image pre-processing*

In the image Pre-processing stage we started with image enhancement; the aim of image enhancement is to improve the interpretability or perception of information in images for human viewers, or to provide 'better' input for other automated image processing techniques Therefore, all the image have been undergoing several pre-processing process[4]. Image pre-processing process involved are smoothing, enhancement, and segmentation are done.

i). Image Enhancement

Image enhancement techniques can be divided into two broad categories: Spatial domain methods and frequency domain methods. Unfortunately, there is no general theory for determining what "good" image enhancement is when it comes to human perception. If it looks good, it is good! However, when image enhancement techniques are used as pre-processing tools for other image processing techniques, then quantitative measures can determine which techniques are most appropriate. In our image enhancement stage we used three techniques: Gabor filter, auto-enhancement and Fast Fourier transform techniques.

ii). Image Segmentation

Image segmentation is an essential process for most image analysis subsequent tasks. In particular, many of the existing techniques for image description and recognition depend highly on the segmentation results. We used Thresholding and watershed segmentation techniques.

Thresholding is one of the most powerful tools for image segmentation. The segmented image obtained from Thresholding has the advantages of smaller storage space, fast processing speed and ease in manipulation, compared with gray level image which usually contains 256 levels. Therefore, thresholding techniques have drawn a lot of attention during the past 20 years.

Watershed segmentation extracts seeds indicating the presence of objects or background at specific image locations. The marker locations are then set to be regional minima within the topological surface (typically, the gradient of the original input image) and the watershed algorithm is applied.

B. *Feature Extraction*

The Image features Extraction stage is very important in our working in image processing techniques which using algorithms and techniques to detect and isolate various desired portions or shapes (features) of an image. Feature extraction is an essential stage that represents the final results to determine the normality or abnormality of an image[7].

These features act as the basis for classification process. Only these features were considered to be extracted; average intensity, area, perimeter and eccentricity. The features are defined as follows:

1) Area: it is a scalar value that gives the actual number of overall nodule pixel. It is obtained by the summation of areas of pixel in the image that is registered as 1 in the binary image obtained.

2) Perimeter: It is a scalar value that gives the actual number of the outline of the nodule pixel. It is obtained by the summation of the interconnected outline of the registered pixel in the binary image.

3). Roundness: Eccentricity: This matric value or roundness or circularity or irregularity index (I) is to 1 only for circular and it is <1 for any other shape. Here it is assumed that, more circularity of the object[5].

III. RESULT AND DISCUSSION

CT image of Lung Cancer have successfully undergo the image pre-processing procedure with four features: average intensity, area, perimeter and eccentricity.

A. *Results for image enhancement*

We can define image enhancement as away to improve the quality of image, so that the resultant image is better than the original one , the process of improving the quality of a digitally stored image by manipulating the image with MATLAB software. It is quite easy, for example, to make an image lighter or darker, or to increase or decrease contrast. MATLAB also supports many filters for altering images in various ways.

i). Gabor filter enhancement technique

The Gabor filter was originally introduced by Dennis Gabor, we used it for 2D images (CT images). The Gabor function has been recognized as a very useful tool in computer vision and image processing, especially for texture analysis, due to its optimal localization properties in both spatial and frequency domain.

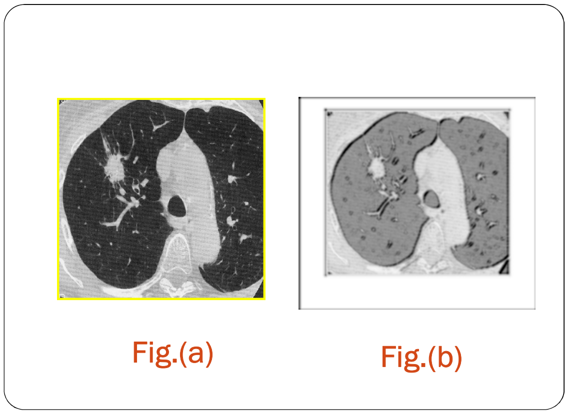

Figure.3.1 The result of applying Gabor filter enhancement technique:
Fig.(a) Original Image Fig.(b) Enhanced Image

ii). Auto Enhancement technique

Auto enhancement, automatically adjusts and enhances the image (brightness, color and contrast) to optimum levels, and this is clearly observed by following (figure 3.2) as you see (a) is the original image and (b) is the image after applying auto enhancement MATLAB code. This method strongly depends on statistical operations such as mean, variance calculation.

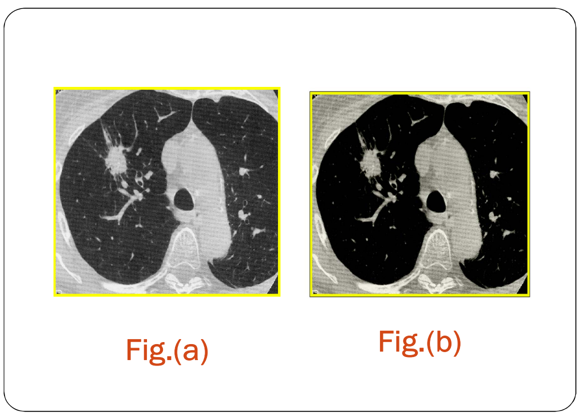

Figure.3.2 The result of applying Auto enhancement technique:
Fig.(a) Original Image Fig.(b) Enhanced Image

iii). Fast Fourier Transform technique

Fast Fourier Transform technique operates on Fourier transform of image. The frequency domain is a space in which each image value at image position F represents the amount that the intensity values in image I vary over a specific distance related to F. Fast Fourier Transform "FFT" is a faster version of the Discrete Fourier Transform (DFT).

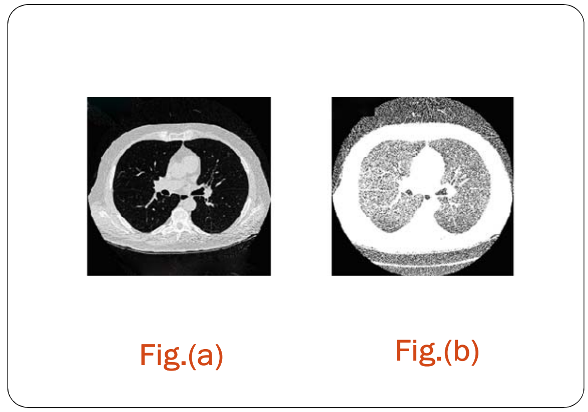

Figure.3.3 The result of applying FFT enhancement technique:
Fig.(a) Original Image Fig.(b) Enhanced Image

Gabor Enhancement is the most suitable technique that can be used to have a good quality image.

B. Results for segmentation

Segmentation divides an image into its constituent regions or objects. The segmentation of medical images in 2D, slice by slice has many useful applications for the medical professional: visualization and volume estimation of objects of interest, detection of Abnormalities (e.g. tumors, etc.), tissue quantification and classification, and more.

i). Thresholding approach

Thresholding is a non-linear operation that converts a gray-scale image into a binary image where the two levels are assigned to pixels that below or above the specified threshold value.

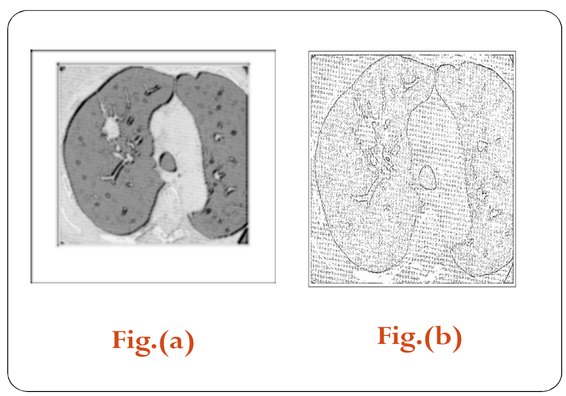

Figure.3.4 Normal Enhanced Image by Gabor filter and its Segmentation using thresholding code:
Fig.(a) Enhanced Image Fig.(b) Segmented Image

One obvious way to extract the objects from the background is to select a threshold 'T' that separates these modes. Then any point (x,y) for which $f(x,y) > T$ is called an object point, otherwise, the point is called a background point. Threshold value based on this method will be between 0 and 1, after achieve this value we can segment an image based on it.

ii). Watershed Segmentation approach

Separating touching objects in an image is one of the more difficult image processing operations, however it has no smoothing/generalization properties. The marker based watershed segmentation can segment unique boundaries from an image.

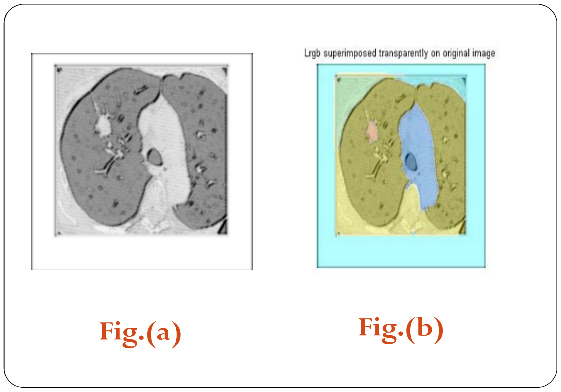

Figure3.5 Normal Enhanced Image by Gabor filter and its Segmentation using Marker-Watershaed:
Fig.(a) Enhanced Image Fig.(b)Segmented Image

According to our experimental subjective assessment in the segmentation stage the Watershed Segmentation approach has more accuracy and quality than Thresholding approach [6].

C. *Result for feature extraction*

Feature extraction is an essential stage that represents the final results to determine the normality or abnormality of an image. Features estimated for separated nodule has been found as follow:

Average intensity: 226.0;

Area: 1024;

Perimeter: 196.3;

Eccentricity: 0.78;

These features act as the basis for classification process. Regarding on the type of the image processed, which is binary image, the only color presented is black and white. Thus, only three features were considered to be extracted; area, shape and perimeter. The features are defined as follows:

1) Area: It is a scalar value that gives the actual number of overall nodule pixel. It is obtained by the summation of areas of pixel in the image that is registered as 1 in the binary image obtained.

2) Perimeter: It is a scalar value that gives the actual number of the outline of the nodule pixel. It is obtained by the summation of the interconnected outline of the registered pixel in the binary image.

3) Eccentricity: This matric value or roundness or circularity or irregularity index (I) is to 1 only for circular and it is <1 for any other shape. Here it is assumed that, more circularity of the object.

D. *Classification*

Lung nodule are smallest growths in the lung that measure between 5mm to 25mm in size. Malignant nodules tend to be bigger in size >25mm, and have a faster growth rate. In the normal images nodule size is less than 25mm. And in the abnormal images its size is greator than 25mm. In the segmentation that nodule is detected and than we use feature extraction to extract the featuers from that segmented image by which we can identify the stages of lung cancer.

Lung nodule show up as round, white opacities on chest x-rays and computed tomography scans. Previous scan x-ray or scan and the current x-ray and ct-scan is used to determine if there is any change in shape, size, or appearance of the nodules. If the nodule do not grow larger after monitoring for a 2 year period, no further treatment is necessary.

IV. STAGING OF LUNG CANCER

The stage of a cancer is a measure of the extent to which a cancer has spread in the body. Staging involves evaluation of a cancer's size and its penetration into surrounding tissue as well as the presence or absence of metastases in the lymph nodes or other organs[7]. Staging is important for determining how a particular cancer should be treated, since lung-cancer therapies are geared toward specific stages. Staging of a cancer also is critical in estimating the prognosis of a given patient, with higher-stage cancers generally having a worse prognosis than lower-stage cancers.

A. *Results of image segmentation for normal, stage-I, stage-II*

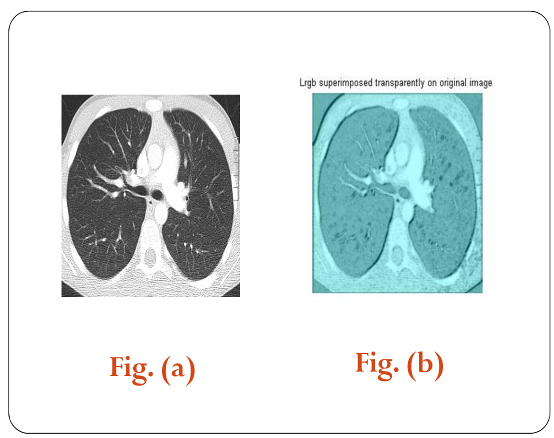

Figure.4.1 Results of image segmentation for normal image

Fig.(a) Normal CT Image Fig.(b) Segmented Image

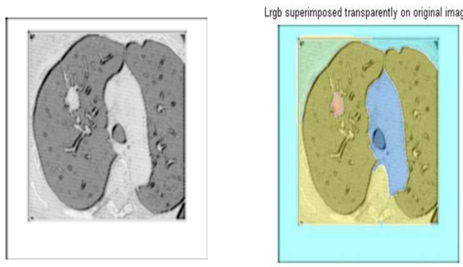

Fig.(a) Fig.(b)

Figure.4.2 Results of image segmentation for stage-I

Fig.(a) Normal CT Image Fig.(b) Segmented Image

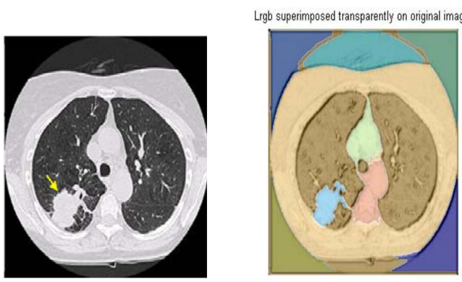

Fig.(a) Fig.(b)

Figure.4.3 Results of image segmentation for stage-II

Fig.(a) Normal CT Image Fig.(b) Segmented Image

B. Results of Feature Extraction for stage-I and stage-II

Table.4.1 Results of feature extraction for stage-I and stage-II

Stage-I	Stage-II
• Average intensity: 226.0;	• Average intensity: 247.56;
• Area: 1024;	• Area: 2438;
• Perimeter: 196.3;	• Perimeter: 294.50;
• Eccentricity: 0.78;	• Eccentricity: 0.86;

Results of Feature Extraction for stage-I and stage-II

Stages from I to IV in order of severity:

- In stage I, the cancer is confined to the lung.
- In stages II and III, the cancer is confined to the chest (with larger and more invasive tumors classified as stage III).
- Stage IV cancer has spread from the chest to other parts of the body.

V. CONCLUSION

Lung cancer is the most dangerous and widespread in the world according to stage the discovery of the cancer cells in the lungs, this gives us the indication that the process of detection this disease plays a very important and essential role to avoid serious stages and to reduce its percentage distribution in the world. To obtain more accurate results we divided our work into three stages: Image Enhancement stage, Image Segmentation stage and Features Extraction stage. Lung Nodule Detection in CT Scans is an active area of research which is continuously emerging and there are many enhancements that can be included to make more efficient.

REFERENCES

[1] American Cancer Society, "Cancer Statistics, 2005", CA: A Cancer Journal for Clinicians, 55: 10-30, 2005, "http://caonline.amcancersoc.org/cgi/content/full/55/1/10".

[2] D. Lin and C. Yan, "Lung nodules identification rules extraction with neural fuzzy network", IEEE, Neural Information Processing, vol. 4, (2002).

[3] A. El-Baz, A. A. Farag, PH.D., R. Falk, M.D. and R. L. Rocco, M.D., "detection, visualization, and identification of lung abnormalities in chest spiral CT scans: phase I", Information Conference on Biomedical Engineering, Egypt (2002).

[4] B.V. Ginneken, B. M. Romeny and M. A. Viergever, "Computer-aided diagnosis in chest radiography: a survey", IEEE, transactions on medical imaging, vol. 20, no. 12 (2001).

[5] Beucher, S. and Meyer, F., "The Morphological Approach of Segmentation: The Watershed Transformation," Mathematical Morphology in Image Processing, E. Dougherty, ed., pp. 43-481, New York: Marcel Dekker, 1992.

[6] Nguyen, H. T., et al "Watersnakes: Energy-Driven Watershed Segmentation", IEEE Transactions on Pattern Analysis and Machine Intelligence, Volume 25, Number 3, pp.330-342, March 2003.

[7] Suzuki K., et al., "False-positive Reduction in Computer-aided Diagnostic Scheme for Detecting Nodules in Chest Radiographs by Means of Massive Training Artificial Neural Network", Academic Radiology, 12, No 2, February 2005, pp. 191-201.