

DETECTION OF LUNG NODULE USING IMAGE PROCESSING TECHNIQUES

Ms. Sushma Nagdeote

Assistant Professor, Department of Electronics Engineering
Fr. Conceicao Rodrigues College of Engineering
Bandra, Mumbai, Maharashtra,

Abstract : Lung cancer also known as carcinoma is one of the common cause of death among people throughout the world. Early detection of lung cancer can increase the chance of survival among people. Medical Imaging plays an important role in the early detection and treatment of cancer. Computer Aided Diagnosis (CAD) system allows detection of lung cancer through analysis of chest CT scan images. This work is basically focusing on Segmentation and nodule extraction, in which nodule attributes like image intensity, shape and size have to be taken into consideration. The objective of this study is identifying all nodules from the chest CT lung images and classifying these nodules into Malignant and Benign nodules, using Image processing techniques.

IndexTerms - SCLC, NSCLC, Watershed, Malignant , Benign

I. INTRODUCTION

Of all types of cancer, Carcinoma is the most common cause of death. Lung cancer claims more lives each year than colon, prostate, ovarian and breast cancers combined. This work is based on Image Processing Techniques. The techniques used in this approach are thresholding, region growing, edge detection, ridge detection, morphological operations, Watershed algorithm and Nodule extraction. There is also an another approach used in lung regions extraction process based on pixel classifications.

Computer Aided Diagnosis (CAD) is a technology combining elements of computer vision: an application area of artificial intelligence, and digital image processing. The applications of CAD include identification of cancer in lungs, breasts, and colon, coronary artery disease, and heart. In CAD, the process of detecting lung cancer involves two tasks: identification and classification. The identification task pertains to finding lung cancer nodules in medical images. Lung cancer nodule, also referred as solitary pulmonary nodule or coin lesion, is visible as a mass in lung spanning a diameter smaller than 3 centimeter. The classification task comprises of two stages. First of all, candidate nodules are classified as nodules or non-nodules. Second, the determined nodules are further classified as benign or malignant. In 1993 Chiou YSP, Lure YMF, Ligomenides PA[1] designed an artificial neural network based hybrid lung cancer detection system, which was used to improve the accuracy of diagnosis.

II. OBJECTIVE

The Objective of this model is Segmentation and nodule extraction, and classification of Lung Nodule. Lung cancer one of the most serious health issues that cause death worldwide, due to this reason Lung image segmentation has earned a lot of importance in medical imaging. Manual segmentation of Lung is huge time consuming and intensive task. To automate the process so that large number of cases can be handled with accuracy a new fast approach for the segmentation of the lung images in order to detect and extract the tumor region is proposed.

III. TYPES OF LUNG CANCER

There are two main types are carcinoma 1) small-cell lung cancer(SCLC) and 2) non-small-cell lung cancer(NSCLC). Faster-growing types of NSCLC include large-cell carcinoma and large-cell neuroendocrine tumors. SCLC represents about 20% of lung cancers. SCLC grows and spreads faster than NSCLC. This also makes it more likely to respond to chemotherapy, but it's also less likely to be cured with treatment. In some cases, lung cancer tumors contain both NSCLC and SCLC cells. In 2011 Disha Sharma and Jagdeep Jindal [4] applied computer Aided Diagnosing (CAD) system for detection of lung cancer. This system generally first segments the lung and then analyzes the separately obtained area for nodule detection in order to diagnosis the disease, as a result 90% sensitivity with 0.05 false positives per image.

3.1. STAGES OF NSCLC

Stage 1: Cancer is found in the lung, but it has not spread outside the lung. **Stage 2:** Cancer is found in the lung and nearby lymph nodes. **Stage 3:** Cancer is in the lung and lymph nodes in the middle of the chest. **Stage 3A:** Cancer is found in lymph nodes, but only on the same side of the chest where cancer first started growing. **Stage 3B:** Cancer has spread to lymph nodes on the opposite side of the chest or to lymph nodes above the collarbone. **Stage 4:** Cancer has spread to both lungs, into the area around the lungs, or to distant organs.

3.2. STAGES OF SCLC

Small cell cancer is divided into "limited stage" generally cancer confined to the chest and "extensive stage" cancer that has spread outside the chest.

Image processing is used to analyse images at the lowest level provided any quality. The main requirement of processing is to improve pixel intensity by converting from discrete to digital image, segmenting to pixels, carrying out mathematical operations on pixels, and reconstructing of image with better quality. Preprocessing of CT images is the initial step in image analysis followed by segmentation process and ended with some morphological operations are applied to detect the cancer spots or cells in the image. Also it can be used to determine the amount of spreading of cancer i.e. what percentage of lung is affected with cancer. The morphological operations are basically applied by comparing the size and shape of the cancer cell with normal cell, and then the infected cells images displayed onto gray scale image with maximum intensity (255). The sequencing of the paper is as follows. Section III describes the prescribed methodology, in detail performed on the image using various filtering methods followed by segmentation and shows the results of the developed methodology using a MATLAB. Finally, section IV defines the conclusion.

IV. METHODOLOGY

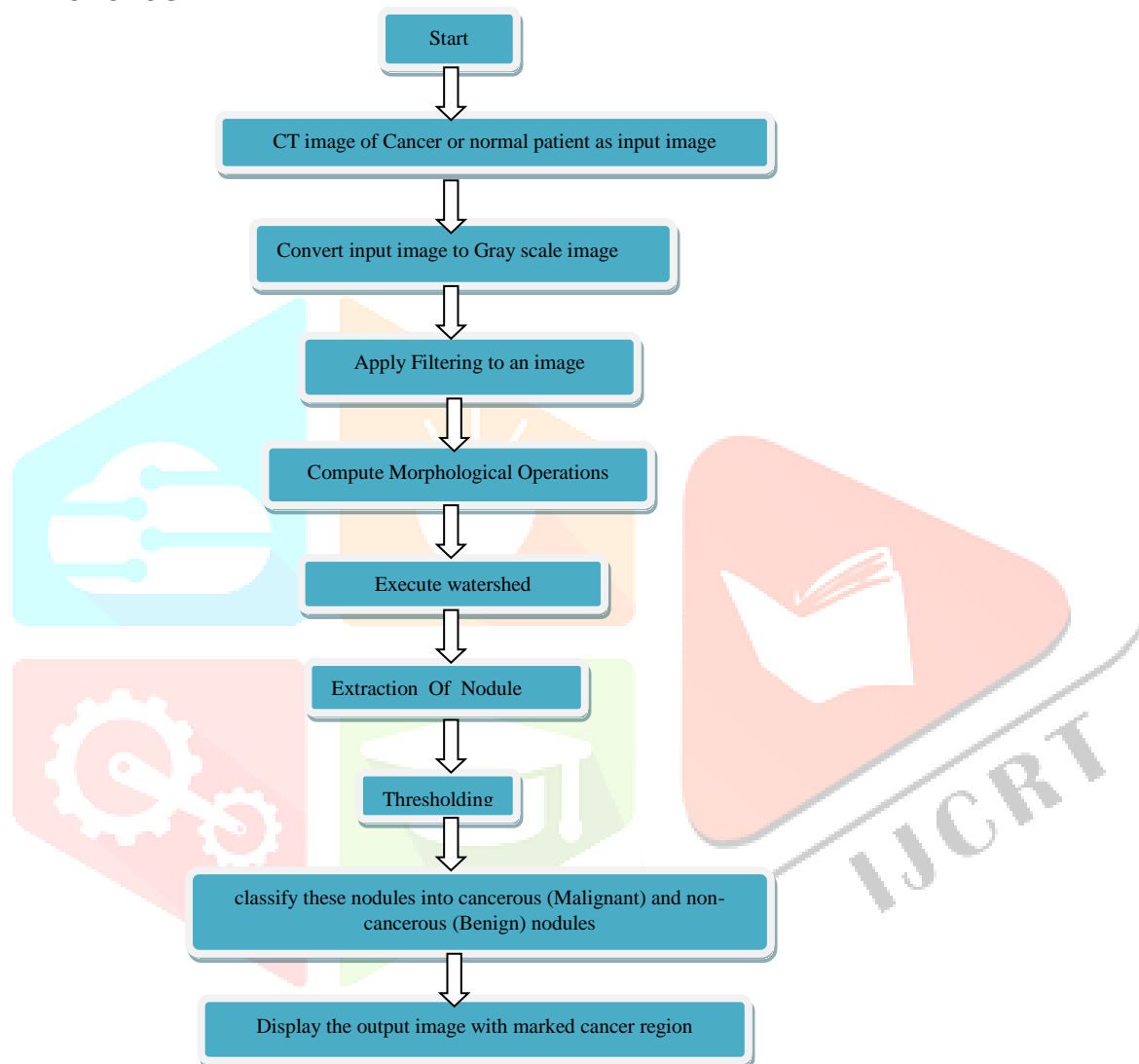


Figure. 1. Algorithm for lung cancer detection

The algorithm is explained below as follows:

4.1 Input Image

CT scan also known as computed tomography makes use of X-rays for capturing the images from various angles and merge these images to generate cross sectional tomographic image of particular areas of scanned tissues. The image of affected lungs and normal lungs is quite different and easily differentiable. These CT images are converted to grayscale images.

4.2. Converting input image to Grayscale image

A grayscale image is a digital image, it carries only intensity or amplitude. This image consist of white pixel(highest intensity), black pixel(lowest intensity) and inbetween highest and lowest intensity consists of exclusive shades of gray. A gray colour is one in which the R, G, B planes have equal intensity, the brightness levels represented as a number from decimal 0 to 255. For every pixel in an RGB grayscale image, $R = G = B$. The intensity varies in proportion with the number representing the brightness levels of the RGB colours. Black is represented by $R = G = B = 0$ and white is represented by $R = G = B = 255$.



Figure 2. Input Image

4.3. Filtering

The median filter is applied before edge detection, which helps in noise reduction. The noise should be removed initially up to some threshold value. The main feature of median filter is to remove unwanted noise without affecting the edges. Median filter is Order statistic filters are non-linear spatial filter which is sensitive to mean value of pixel intensity and less sensitive to extreme intensity of pixels. Sharpening filter is applied to enhance edged in an image. This project uses two different filters which are Gabor and Laplacian and we compare the results of two filter. The Gabor filter is band pass filter used to detect edges. Laplacian filters are derivative filters which are used to find areas of rapid change (edges) in images. The result of Gabor and laplacian filter is shown in figure 2. Mokhled S. Al-Tarawneh,[5] introduces a lung detection algorithm depends on Gabor filter and watershed algorithm. The system designed by Mokhled S. Al-Tarawneh, achieved 85.165 rate in lung cancer detection.

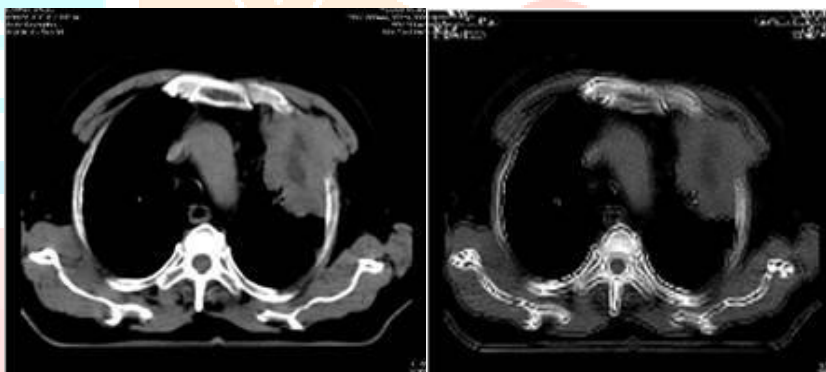


Figure 2. Filtered Image (a) Gabor (b) Laplacian

The first operation in detecting tumor is defining the segmentation function, which in our algorithm is the gradient magnitude and it has been done by Sobel filter. The Sobel operator performs a 2-D spatial gradient measurement on an image and emphasizes regions of high spatial frequency that correspond to edges. Typically it is used to find the approximate absolute gradient magnitude at each point in an input grayscale image.

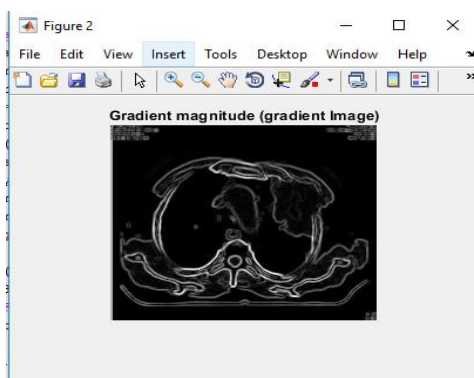


Figure 3. Gradient Image

4.4. Segmentation Function

In this algorithm, the first operation in detecting tumor is defining the segmentation function, which is the gradient magnitude and it has been done by Sobel filter.

Sobel filter is applied on vectors and rows, then the square root of results is calculated to get edges. Filtered images (G_x, G_y) of sobel and the resulted edges (G) are given as follows [6]:

$$G_y = f(x-1, y-1) + 2 * f(x-1, y) + f(x-1, y+1) - f(x+1, y-1) - 2 * f(x+1, y) - f(x+1, y+1)$$

$$G = \sqrt{G_x^2 + G_y^2}$$



Figure 4. Sobel filter output Image

4.5. Morphological Operations

Morphological image processing is a collection of non-linear operations related to the shape or morphology of features in an image. R. A. Blechschmidt, R. Werthschutzky, and U. Lorcher, [2] designed morphology based automated CT image evaluation of the lung cancer. Apply opening operation to the above image using the 20 pixel size mask of type disk on the enhanced image. The reconstructed image can be obtained by closing operation on previously opened image. So as to remove the dark regions from an image. Now apply erosion then reconstruction operations on enhanced image. The resultant image is shown in figure 5. The reconstruction step needs two images which are the marker image (eroded image) and mask image (enhanced image). Now apply dilation operation on the reconstructed image. The complement of dilated image is then fed as marker of second reconstruction operation, while the mask image is the complement of the first reconstruction image. The comparison between the previous two images shows that reconstruction method is better than the other, so we go on with the reconstruction method. The resultant image is shown in figure 6. The next step is to get the foreground pixels. For this calculate the maximum values in the reconstructed image. The foreground pixels are then plotted on the enhanced image. The resultant image is shown in figure 7. The reconstructed open-close image is transformed into binary form. The black pixels is the background but these points shouldn't touch the region of interest. Hence, we skeletonize the background area.

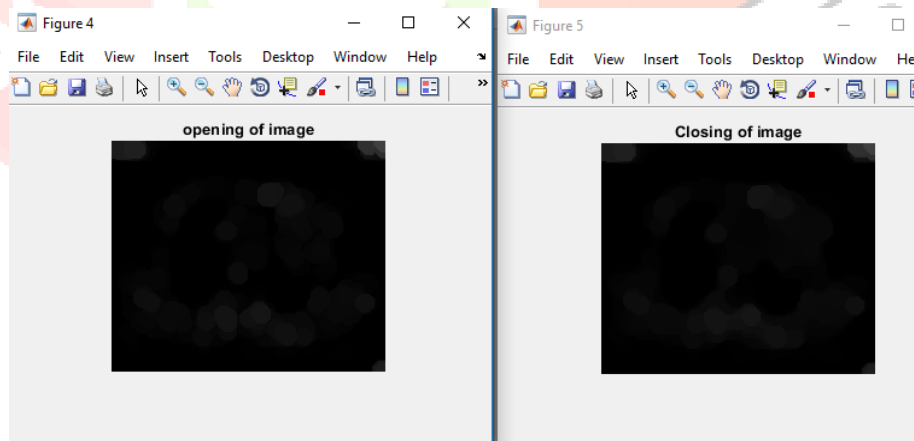


Figure 5. a) Opening of Image b) Closing of Image

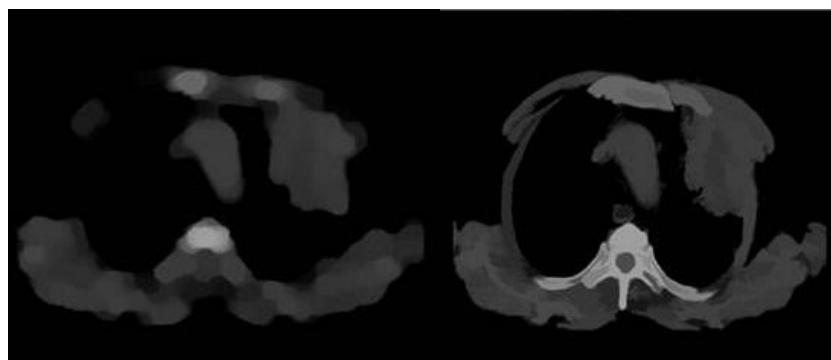


Figure 6. a) Opening and closing of enhanced image b) Reconstructed open-close image

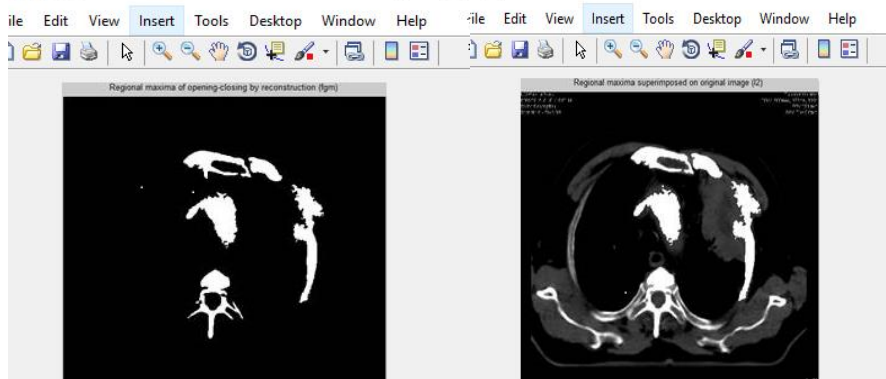


Figure 7. a) Maximum regions in reconstructed image b) Enhanced image with Foreground pixels

4.6 Watershed Segmentation

Any grayscale image can be viewed as a topographic surface where high intensity denotes peaks and hills while low intensity denotes valleys. You start filling every isolated valleys (local minima) with different colored water (labels). As the water rises, depending on the peaks (gradients) nearby, water from different valleys, obviously with different colors will start to merge. To avoid that, you build barriers in the locations where water merges. You continue the work of filling water and building barriers until all the peaks are under water. Then the barriers you created gives you the segmentation result. Ilya Levner, Hong Zhangm, [3] Classification driven Watershed segmentation presents a novel approach for creation of topographical function and object markers used within watershed segmentation. Watershed segmentation technique extract regions that indicate the presence of objects or background at specific image locations. Marker locations are then set to be regional minima within the topological surface, and the watershed algorithm is applied. Separating touching objects in an image is one of the most difficult image processing operations, where the watershed transform is often applied to such problem. Marker-controlled watershed approach has two types: External associated with the background and Internal associated with the objects of interest. Image Segmentation using the watershed transforms works well if we can identify or “mark” foreground objects and background locations, to find “catchment basins” and “watershed ridge lines” in an image by treating it as a surface where light pixels are high and dark pixels are low. Figure below shows a segmented image by watershed.

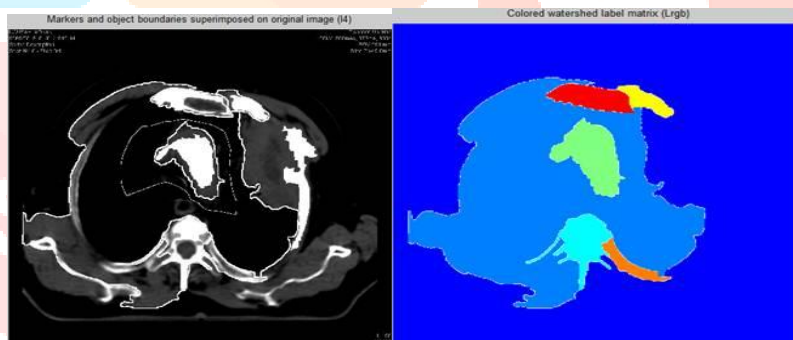


Figure 8. a) Enhanced image b) Colored Watershed Image

The multi-color Watershed image signifies that the lung image is infected. While if the lung is not infected, then the watershed result in mono-color. So, a threshold of 1 of image’s regions is taken to determine if the lung is infected or not. The non-infected lung image which includes only one region is shown. figure 9.

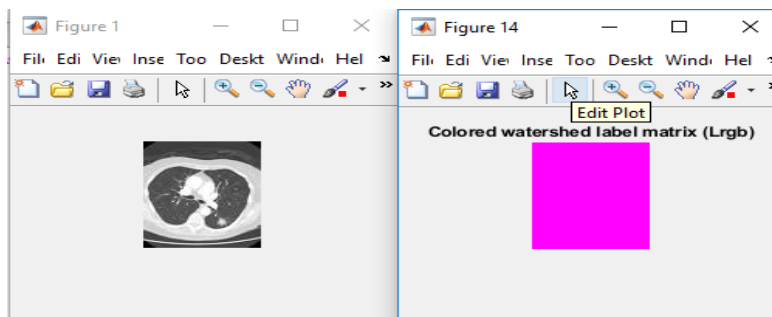


Figure. 9. a) Healthy Lung Image b) watershed image of healthy lung

4.7. Nodule Extraction

To detect the tumor region, we take the labeled watershed image as input. First, we removed the very small regions which are noise, and can’t be the tumor.

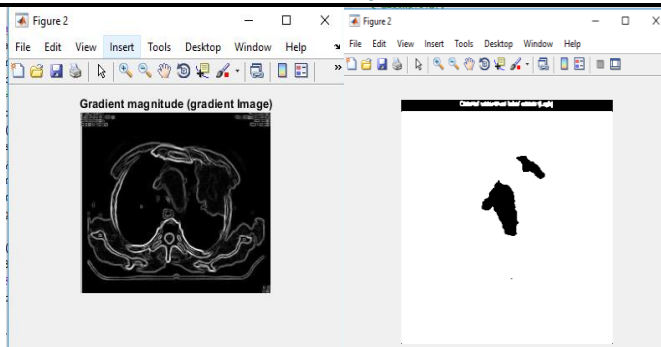


Figure. 10. Extracted Nodule

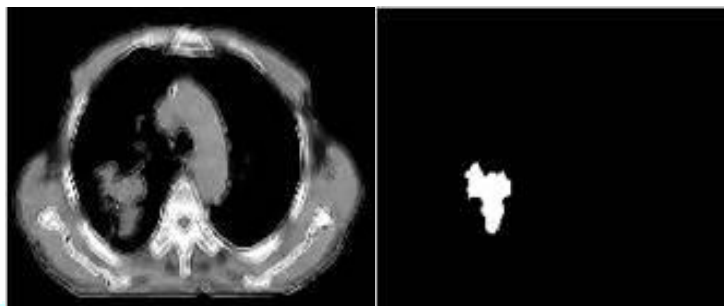


Figure. 11. Extracted Nodule

4.8. Nodule Classification

Nodule Classification Based on the features identified in the previous step, the classification step tries to minimize the number of false positive, and to differentiate between the nodules and the non-nodules. To detect the tumor region, we take the labeled watershed image as input. The threshold value for this stage is 190. Next, the area of each extracted nodule is calculated and then classifying these nodules into cancerous (Malignant) and non-cancerous (Benign) nodules.

$$T(x,y) = \begin{cases} 1 & \text{if } f(x,y) \geq 190 \\ 0 & \text{otherwise} \end{cases}$$

V. CONCLUSION

The Results were applied on 25 images. The Laplacian filter approach were better than gabor in detection and segmentation rates. The Gabor filter gives 88% detection rate while the laplacian filter gives 100% detection rate. This proposed algorithm works well with the original CT images and is able to successfully extract and classify the nodule as malignant or benign from the CT image. The result of the approach is shown in the table below.

VI. RESULTS

TABLE 1 DETECTION RATE BY GABOR FILTER AND LAPLACIAN FILTER

Patient ID	No. of Images	Detected Correctly by Gabor	Detected Correctly by Laplacian
1	5	4	5
2	5	4	5
3	5	5	5
4	5	4	5
5	5	5	5

VII. REFERENCES

[1] Chiou YSP, Lure YMF, Ligomenides PA. Neural network image analysis and classification in hybrid lung nodule detection (HLND) system. In: Proceedings of the IEEE-SP Workshop on Neural Networks for Signal Processing, 1993. p.517-526.

[2] R. A. Blechschmidt, R. Werthschutzky, and U. Lorcher, "Automated CT image evaluation of the lung: A morphological based concept," IEEE trans. Med. Imag., vol.20, no.5, pp.434-422, May 2001.

[3] Ilya Levner, Hong Zhangm, "Classification driven Watershed segmentation", IEEE transactions on image processing vol. 16, no. 5, may 2007.

[4] Disha Sharma, Gagandeep Jindal, "Identifying Lung Cancer Using Image Processing Techniques", International Conference on Computational Techniques and Artificial Intelligence, 2011, pp:115-120.

[5] Mokhled S. AL-TARAWNEH, "Lung Cancer Detection Using Image Processing Techniques", Leonardo Electronic Journal of Practices and Technologies, Issue 20, January-June 2012, p. 147-158.

[6] R.C. GONZALEZ and R. E. WOODS, "Digital Image Processing", 2nd Edition, Prentice Hall, January 2002.