Lung Cancer Detection using Matlab Image Processing Techniques

[1]Abubaker Khan, [2]Tejas Nivatkar, [3]Vivek Mishra, [4]Prof. Omprakash Yadav

[1,2,3]Students, Department of Computer Engineering, Xavier institute of Engineering, Maharashtra, India
[4]Prof. Omprakash, Department of Computer Engineering, Xavier Institute Of Engineering, Maharashtra, India

---

Abstract - *Cancer is a disease during which cells within the body grow out of control. Lung cancer begins within the lungs and will spread to lymph nodes or other organs within the body, like the brain. Cancer from other organs also may spread to the lungs. When cancer cells spread from one organ to a different, they're called metastases. In 2019 America alone recorded 228820 cases of which 135720 deaths were recorded. The existing system used for lung cancer detection uses X-ray or CT scan images identifying the excess mass of flesh in the lung. The traditional system requires you to be diagnosed by the doctor manually, which sometimes can be a lengthy process , we in this project wants to make this manual process machine based, which will facilitate the diagnosing and will result in fast treatment of the diseases. Different CT-Scan images were taken from the different hospital and And varies techniques were used like image enhancement image segmentation and gabor filter in this project.*

Key Words: Segmentation, Thresholding, Watershed, Gabor filter, CT Scan Image, inarization.

1. INTRODUCTION

Lung abnormality is a disease of abnormal cells multiplying and growing into a tumour. Abnormality cells can be carried away from the lungs in blood, or lymph fluid that surrounds lung tissue. Lymph flows through lymphatic vessels, which drain into lymph nodes located in the lungs and in the centre of the chest. Lung abnormality often spreads toward the centre of the chest because the natural flow of lymph out of the lungs is toward the centre of the chest. Metastasis occurs when abnormality cell leaves the site where it began and moves into a lymph node or to another part of the body through the blood stream. Abnormality that starts in the lung is called primary lung abnormality. About 85% male and 75% females are suffering from lung cancer due to cigarette smoking. The general survival rate of people suffering from lung cancer is 63%. There are several different types of lung abnormality, and these are divided into two main groups: Small cell lung abnormality and non-small cell lung abnormality which has three subtypes: Carcinoma, Adenocarcinoma and Squamous cell carcinomas The rank order of abnormalities for both males and females among Jordanians in 2008 indicated that there were 356 cases of lung abnormality accounting for (7.7 %) of the newly diagnosed abnormality cases in 2008. Lung abnormality affected 297 (13.1 %) males and 59 (2.5%) females with a male to female ratio of 5:1 which Lung abnormality ranked second among males and 10th among females. The disease general description of lung abnormality detection system that contains five basic stages. The first stage starts with taking a collection of CT images (normal and abnormal) from the available Database from IMBA Home (VIA-ELCAP Public Access).

1.1 Image Enhancement

Gabor filter named by Gabor, may be a linear filter is employed for edge detection. Representation of Garbor filter almost like the human sensory system. In the spatial domain,

D Gabor filter is a Gaussian filter function modulated by a sinusoidal function. In the process of this cancer detection imagery used may be a 2D image, so using 2D Gabor filter.

Gabor formula:

$$G(\sigma, \theta, \lambda, \psi, \gamma; x, y)=\exp -(x\ 02+\gamma\ 2y\ 02)\ 2\sigma 2 \bullet \cos(2\ x\ 0\ \lambda + \psi)$$

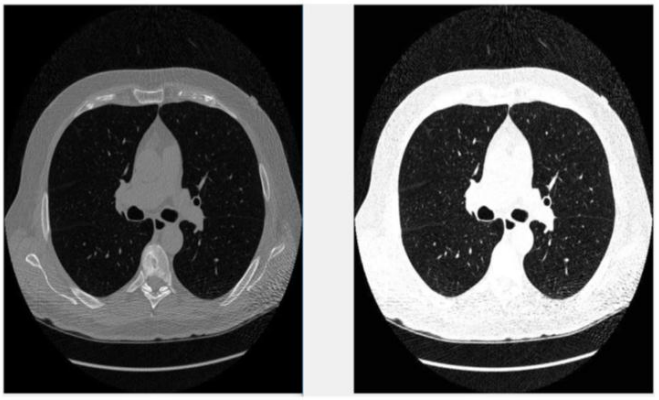

Figure 1.1 Enhanced Gabor Filter output Of Lung Cancer

1.2 Image Segmentation Based Region Growing

Region growing is the simple image segmentation method. It is also classified as a pixel-based image segmentation method since it involves the choice of initial seed points. This approach to segmentation examines neighbouring pixels of initial seed points and determines whether the pixel neighbours should be added to the region. The process to urge the segmentation by region growing method is because the following; firstly, select the world which will be the target object, which are the proper lung and the left lung, then, put the seed in this area.Region growing process in this project give us the clear idea and picture of the image this help us in the further process which will be carried out

during the process of the marked controlled watershed and the binarization process. One simple thanks to segment different objects might be to use their pixel values. An important point to note – the pixel values will be different for the objects and the image's background if there's a sharp contrast between them.We can set a threshold value. The pixel values falling below or above that threshold can be classified accordingly this technique is known as Threshold Segmentation.

Figure 1.2 Region Growing of the Gabor Filter Image

1.3 Marker Controlled Watershed Segmentation

It can be defined as a watershed transformation on the a grayscale image. The watershed transformation treats the image it operates upon sort of a topographic map, with the brightness of every point representing its height, and finds the lines that line the tops of ridges. To get the segmentation results of watershed method, the steps are as follows: firstly, calculate the space gradient for edge detection, then mark the target object by using a morphological technique called opening by reconstruction and shutting by reconstruction. After marking object is revealed, other areas are often discarded. Modules and example of marker controlled watershed segmentation method. Figure shows the result of segmentation by using Marker- Controlled Watershed.

1.4 Images for Database

This databse consist of the 900 small-cell types of lung cancer images, 900 non-small-cell types of lung cancer images, a total of 1800 images (samples) each of 200 X 200 pixels in size.This database we have got from the IMBA Home (VIA-ELCAP Public Access). And some of the image we have got from the hospital doctors and the the Ct-scan centres.

1.5 Feature Extraction

Feature extraction a kind of dimensionality reduction that efficiently represents interesting parts of a picture as a compact feature vector. This approach is beneficial when image sizes are large and a reduced feature representation is required to quickly complete tasks like image matching and retrieval. Feature detection, feature extraction, and matching are often combined to unravel common computer vision problems like object detection and recognition, content-based image retrieval is that the main aspect of any project and the face detection and recognition, and texture classification. Feature extraction a kind of dimensionality reduction that efficiently represents interesting parts of a picture as a compact feature vector. This approach is beneficial when image sizes are large and a reduced feature representation is required to quickly complete tasks like image matching and retrieval.

2. Methodology

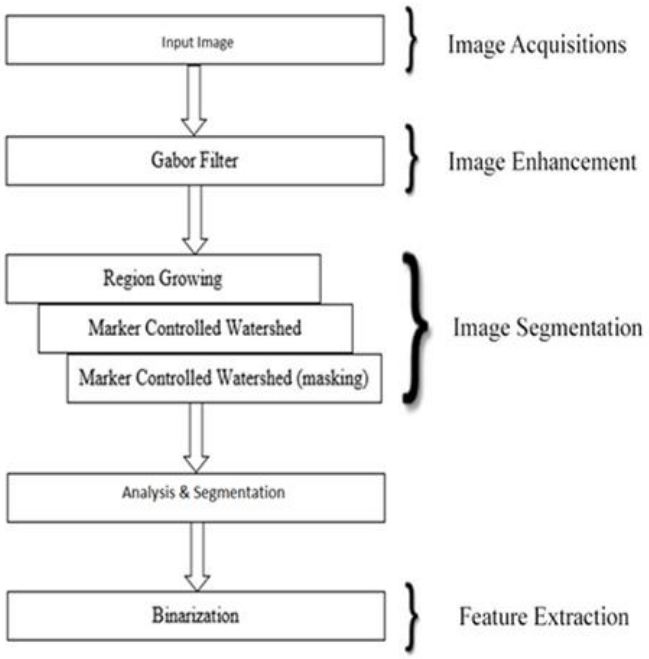

Figure 2.1 Methodology and the flow of the system

3. Activity Diagram

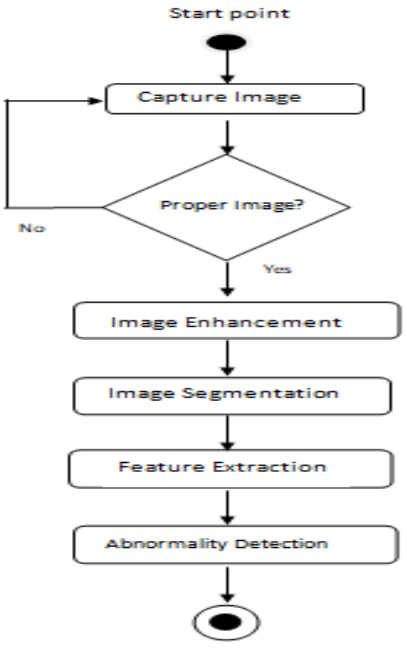

Figure 3.1 Activity Diagram and the flow of the system

4. Binarization Process

Binarization is the process of changing the color of the pixel values into two classes such as black and white. After getting the quantity of black and white pixels on segmentation results, then we compared it with a threshold value to determine the condition of lung (normal or cancer). The threshold value is obtained from observations on normal lung. If the amount of black pixels quite the edge , then we conclude the lung is normal. Image binarization is that the process of taking a grayscale image and converting it to black-and-white, essentially reducing the knowledge contained within the image from 256 reminder gray to 2: black and white, a binary image. This is sometimes referred to as image thresholding, although thresholding may produce images with quite 2 levels of gray. It is a form or segmentation, whereby an image is divided into constituent objects. This is a task commonly performed when trying to extract an object from a picture . However like many image processing operations, it is not trivial, and is solely dependent on the content within the image.

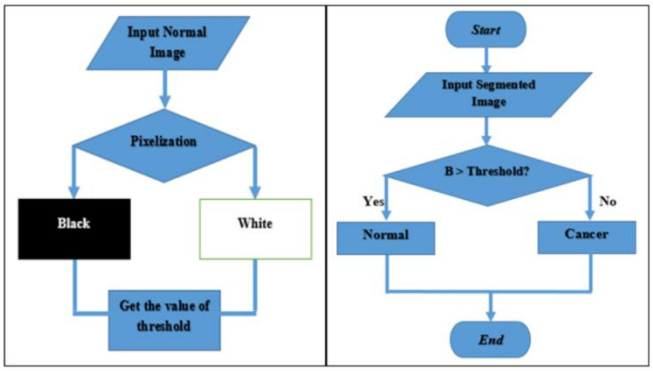

Figure 4.1 Binarization process flow diagram

4.1 Thersholding

We have used the value 17179 for the classification of our input images. After all the steps necessary for the process run successfully which includes image enhancement ,image segmentation ,feature extraction we do the classification of the imgae when binarization gives us the matrix of the final image, if the value is higher then the threshold value set then the image is having the lung cancer and if the thershld value is less as compared to the value set then the image is of the normal lung. Colour images can also be thresholded. One approach is to designate a separate threshold for every of the RGB components of the image then combine them with an AND operation. This reflects the way the camera works and the way the info is stored within the computer, but it doesn't correspond to the way that folks recognize colour. Therefore, the HSL and HSV colour models are more often used; note that since hue may be a circular quantity it requires circular thresholding.

5. Implementation And Result

We implement and analyse the image processing method for detection of lung cancer. Image processing techniques are widely utilized in several medical problems for picture enhancement within the detection phase to support the first medical treatment. In this research we proposed a detection method of carcinoma supported image segmentation. Image segmentation is one among intermediate level in image processing. Marker control watershed and region growing approach are wont to segment of CT scan image. Detection phases are followed by image enhancement using Gabor filter, image segmentation, and features extraction. The best approach for main features detection is watershed with masking method which has high accuracy and robust.

Figure 5.1 Input image to the system

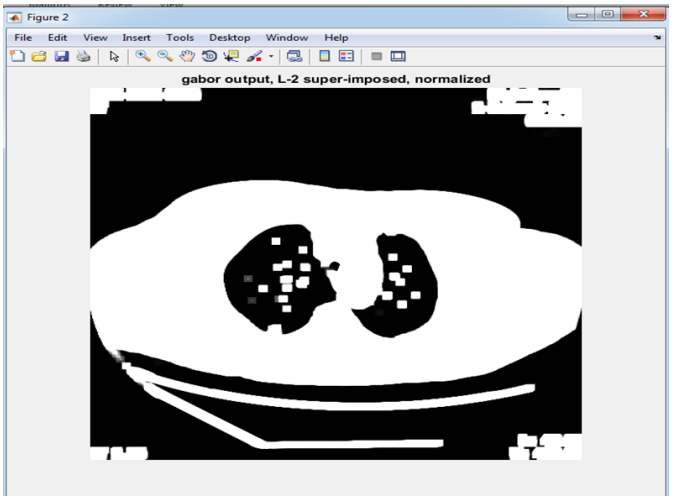

Figure 5.2 Gabor filter output

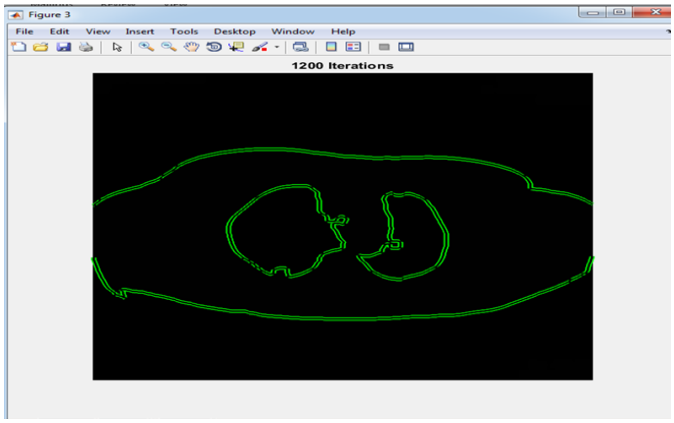

Figure 5.3 Region Growing Output

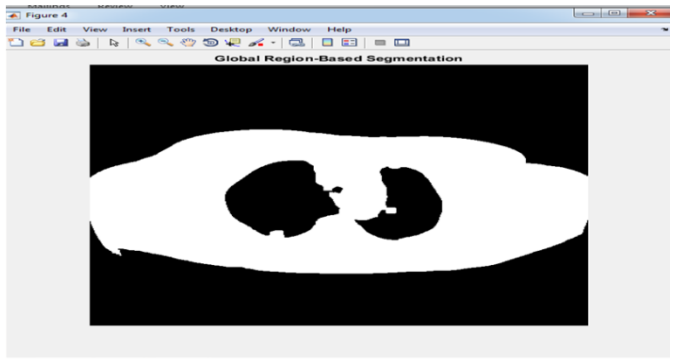

Figure 5.4 Segmentation Output

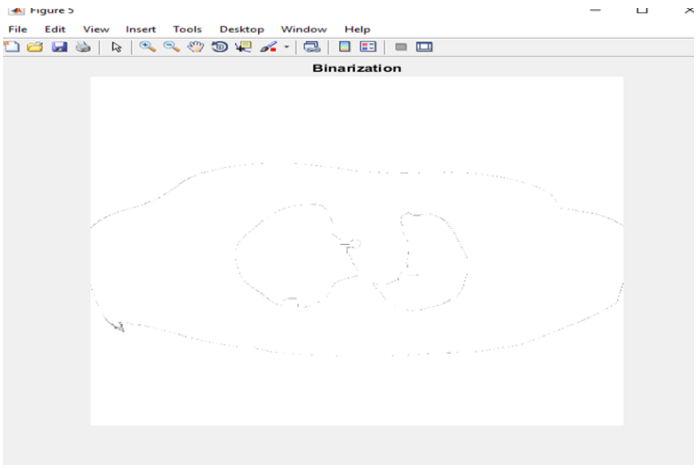

Figure 5.5 Binarization Output

The output of the image we found that the image was infected with the lung cancer. We have used different predetermined image to see and calculate the accuracy of the software by applying different approach of finding the accuracy and its precision. Precision and accuracy are both important in software engineering. Perhaps the simplest thanks to characterize the 2 concepts is that precision is that the mapping of the programmer's model of the matter to the computer's model, whereas accuracy is the mapping of the business' needs to the programmer's model

6. Accuracy evaluation

There are 3 Datasets collected by us from the Internet each containing cancerous as well as non-cancerous images. The Dataset 1 consists of 332 images, dataset 2 and 3 consist of 314 and 257 images respectively.

Table -1: Accuracy matrix

DATASET	NUMBER OF IMAGES	CORRECT OUTPUT	ACCURACY %
Dataset 1	332	268	81%
Dataset 2	314	251	80.1%
Dataset 3	257	185	79.6%

On comparing the results, the approximate correct outputs received were 268, 251 and 185 from datasets 1, 2 and 3 respectively. Hence the accuracy percent is calculated in the table given above and the Average Accuracy obtained was 81%.The highest accuracy which we got from the dataset was 81%,these dataset included both the images of the infected as well as of the healthy person.The model of classification which we build is fairly capable of detecting the lung canvcer if the value of the threshold exceeds the predefined value.

7. Application

[1] This method can be used for the fast detection of the lung cancer so that the proper care can be given to the patient.
[2]This similar method can also be used for the detection of the breast cancer with little modification in the threshold value and the change in the pattern reorganization.
[3] With particular pattern involved with the threshold value this system can be easily be used for other early cancer and tumer detection
[5]It can be used in the detection of the other diseases which are not cancer efficiently like the Lung virus and the other bacterial infection in lung if proper threshold value is given.

8. Advantages

[1] Early detection of the lung cancer which help the patient to start the early treatment of the cancer and will drastically improve the chance of the survival
[2]This method of detection is far more easier and convenient then the other traditional way of detection.
[3]This system is more effective with above 80% accuracy and gives the best result most of the time.

[4]Easy to use system and can be used by any laymen who does not have any information about the lung cancer
[5]result of the testing can come on the same day

9. Conclusion

Lung Abnormality is the most dangerous and widespread in the world according to stage discovery of the Abnormality cells in the lungs, this gives us the indication that the process of detection of this disease plays a really important and essential role to avoid serious stages and to scale back its percentage distribution within the world. Lung cancer is the main disease cause of death of among throughout the world. Lung cancer is causing very high mortality rate. There are various cancer tumours such as lung cancer, breast Cancer, etc. Early stage detection of carcinoma is vital for successful treatment. Diagnosis is based on Computed Tomography (CT) images. In this Histogram Equalization wont to pre-processing of the pictures and have extraction process and classifier to see the condition of a patient in its early stage whether it is normal or abnormal. Recently, image processing techniques are widely utilized in several medical areas for image improvement in earlier detection and treatment stages, where the time factor is extremely important to get the abnormality issues in target images, especially in various abnormality tumours such as lung abnormality, breast abnormality, etc. Image quality and accuracy is that the core factors of this research, image quality assessment also as improvement are counting on the enhancement stage where low pre-processing techniques is used based on Gabor filter within Gaussian rules

ACKNOWLEDGEMENT

I would like to express my deepest appreciation to all those who provided me the possibility to complete this project.

A special gratitude I give to our final year project Guide, Prof. Omprakash yadav whose contribution in stimulating suggestions and encouragement, helped me to coordinate my project I would again like to thank Prof. Dr. Vainkatesh, Principal of the College, for providing us the necessary help and facilities we needed. I express my thanks to all the staff members of computer Engineering department who have directly or indirectly extended their kind co-operation in the completion of my research paper.

REFERENCES

[1]Febr Mokhled S. AL-TARAWNEH, "Lung Cancer DetectionUsing Image Processing Techniques", Leonardo Electronic Journal of Practices and Technologies, June 2012.

[2]Muhammad Usman, Muhammad Shoaib and Mohamad Rahal, "Lung Cancer Detection Using Digital Image Processing", PIERS Proceedings, Stockholm, Sweden, Aug. 12-15, 2013.

[3]Anita Chaudhary and Sonit Sukhraj Singh, "Multi-resolution Analysis Technique for Lung Cancer Detection in Computed Tomographic Images", International Journal of Research in Engineering & Applied Sciences, IJREAS Volume 2, Issue 2 January 2012

[4]Disha Sharma, Gagandeep Jindal, "Computer Aided Diagnosis System for Detection of Lung Cancer in CT Scan Images", International Journal of Computer and Electrical Engineering, Vol. 3, No. 5, October 2011

[5] P. Bountris, E. Farantatos, and N. Apostolou, "Advanced Image Analysis Tools Development for the Early Stage Bronchial Cancer Detection", World Academy of Science, Engineering and Technology 9 2005

[6] Zhi-Hua Zhou, Yuan Jiang, Yu-Bin Yang, Shi-Fu Chen, "Lung Cancer Cell Identification Based on Artificial Neural Network Ensembles", Artificial Ingelligence in Medicine, 2002, vol.24, no.1, pp.25-36.

[7] Samir Kumar Bandyopadhyay, "Edge Detection from CT Images of Lung", International Journal Of Engineering Science & Advanced Technology Volume - 2, Issue - 1, pg: 34 – 37, Jan-Feb 2012.