Detection of Lung Cancer Using Marker-Controlled Watershed Transform

Sayali Satish Kanitkar
M.E. (Signal Processing),
Sinhgad College of engineering, Pune,
Maharashtra, INDIA
kanitkar.sayali@gmail.com

N. D. Thombare
Department of Electronics and Telecommunication,
Sinhgad College of engineering, Pune,
Maharashtra, INDIA
nilima_thombare@yahoo.co.in

S. S. Lokhande
Department of Electronics and Telecommunication,
Sinhgad College of engineering, Pune,
Maharashtra, INDIA
sunitalokhande@yahoo.co.in

Abstract—**The major cause of cancer death is lung cancer. Detection of cancer in the early phase can provide more treatment options, less invasive surgery and increases the survival rate. For lung cancer, if the disease is detected in time, the survival rate of patient increases from 14 to 49% in recent 5 years. It is the most dangerous and widespread disease in the world. The cancer cells present in lung causes lung cancer disease. This cells detection is very important issue for medical researchers. The chances of an effective treatment will significantly increases with early detection. The Computed Tomography (CT) images are used which are more efficient than X-ray. MATLAB is widely used software for the study of lung cancer detection from CT scan images. The process includes image pre-processing, image segmentation, feature extraction and classification technique. This present work proposes a method to detect the cancerous cells effectively from the lung CT scan images. It will minimize the detection error made by the physicians' naked eye.**

Keywords— ***Cancer Detection, Feature extraction, Image processing, Image Enhancement, Marker Controlled Watershed Transform, Thresholding.***

I. INTRODUCTION

Recently, the image processing mechanisms are used widely in different medical areas for increasing earlier detection and treatment stages. The time is very significant factor to discover the disease in the patient as possible as fast. Its early detection increases the chances of an effective treatment.

In 2005, approximately 1,372,910 new cancer cases are predictable and about 570,280 cancer deaths are expected to occur. It is expected that there will be 163,510 deaths from lung cancer, which forms 29% of all cancer deaths. [1] When cells start to grow out of control, cancer begins in a part of the body. The cancer cell starts because of out of control expansion of abnormal cells.

Lung cancer is a disease of abnormal cells multiplying and increasing into a tumour. Cancer cells continue to increase and form new, abnormal cells. Many of them notice their disease when it's too late and the surgery is not simply possible. So finding of lung cancer earlier is most important for successful treatment. Diagnosis is mainly based on CT scan images. Cancerous tumour starts in the part of lung is called primary lung cancer. Following are the types of this lung cancer and these are divided into two main types:

1. Small cell cancer
2. Non small cell cancer

In this, current work focuses on finding tumour and its stages. In this Marker-controlled Watershed segmentation is used to isolate a lung of a CT image.

II. LITERATURE SURVEY

In [2] Region growing algorithm proposed for segmentation of CT scan images of the Lung. This algorithm starts with a seed pixel and also checks other pixels that surround it. It determines the most similar one and, if it meets certain criteria, it will include in the region. The region is developed by examining all unallocated neighboring pixels to the region.

In [3] proposed an approach for detection of cancer cells from Lung CT scan images. This work presents a method to detect the cancer cells from the CT scan image. It reduces the error in the detection part made by the doctors for medical study. It is based on Sobel edge detection and label matrix. Sobel operator helps to locate the edges in an image. It does so by finding the image gradient. Image gradient gives the change in the intensity of the image. Also in [4] a system using Computer Aided Diagnosis (CAD) for finding the edges from CT scan images of lung for detection of diseases is used. Thresholding algorithm [5] gives filtering to detect the sputum cell from the raw image for early detection. A novel method, watershed transformation is presented for image segmentation in [6]. Morphological operations which are opening and closing operations are used to process the gradient image. It is used to eliminate the over segmented area and to reconstruct the morphological gradient which can maintain the shape of gradient image.

The main idea of this paper is to detect the tumour and decide whether it is cancerous or not. It also finds the lung cancer stage and gives more accurate result by using different enhancement and segmentation techniques.

III. PROPOSED SYSTEM

In this, available lung CT scan images are passed through the system which is having following stages: pre-processing stage, segmentation stage, feature Extraction stage and classification.

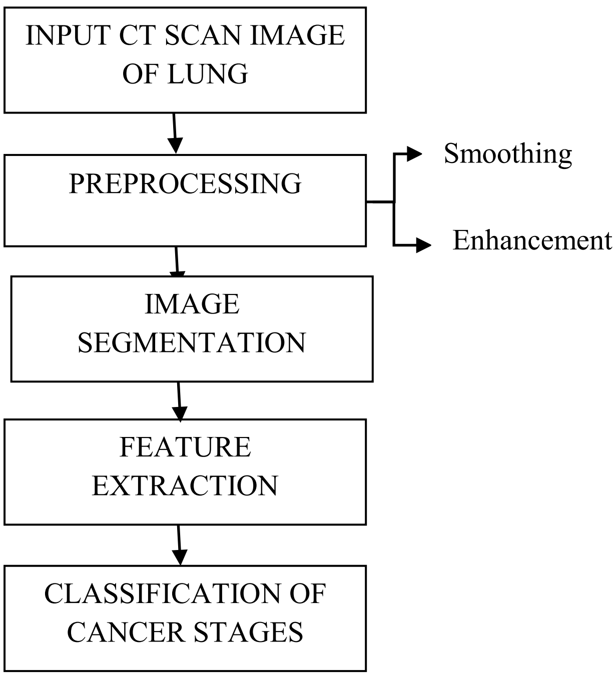
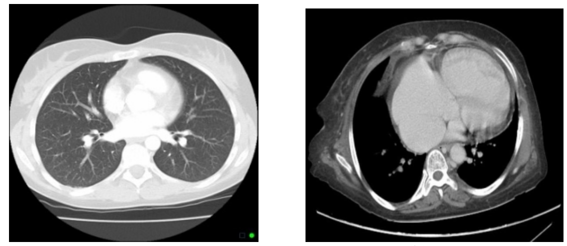


Figure 1. Block Diagram of Proposed System

The Gaussian filter is used to smooth the input image in the preprocessing stage. As well as, in the pre-processing stage, Gabor filter is used for enhancement and thresholding and Marker-Controlled watershed transform is used for the segmentation purpose. After image segmentation, the features such as average intensity, perimeter, area and eccentricity are extracted from the detected tumour. Binarization process is done to decide whether it is cancerous tumour or not. Also, if there is cancerous tumour, the cancer stage is identified.

A. Input CT scan images of lung

The CT scan images which are used for processing are collected from the hospitals. This image dataset contains lung CT scan images with tumour and without tumour. The figure 2 shows some of the lung CT scan images with tumour and without tumour.



(a) (b)

Figure 2. CT scan image (a) Lung without tumour
(b) Lung with tumour

B. Pre-processing

In this image pre-processing stage, image smoothing is the first step. For smoothing, Gaussian filter is applied on the input image. Gaussian smoothing is very effective for removing noise. Gaussian removes high frequency components from the image. So it is a low pass filter. Smoothing reduces the noise and giving us a more accurate intensity surface. The mathematical equation for the Gaussian filter is as given in equation (1).

$$G(x, y) = \frac{1}{2\pi\sigma^2} e^{\frac{-(x^2+y^2)}{2\sigma^2}} \quad (1)$$

where x is the distance from the center on the horizontal axis, y is the distance from the center on the vertical axis, and σ is denoted as the standard deviation of the gaussian distribution. It determines the amount of smoothing. The output of Gaussian filter for figure 2(b) is shown in figure 3.

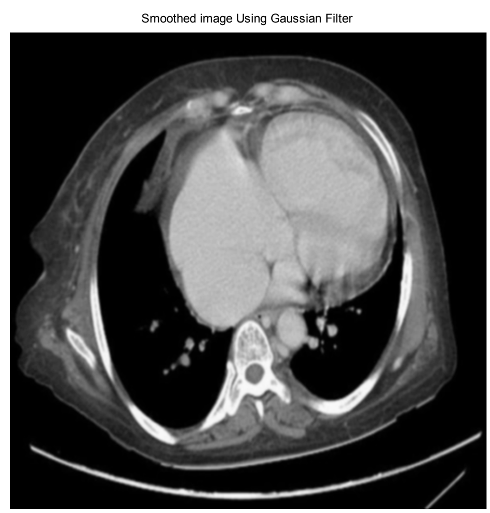


Figure 3. Smoothed Image

Next part in pre-processing is image enhancement stage. The function of image enhancement stage is to highlight the important information of image. In this step, better visual effects are performed on the image which enhance the human eyes' distinguish ability of information. It is a way to improve the class of image, so that the final output image is better than the original one. For image enhancement, Gabor filter is used.

The Gabor function is a very helpful tool in image processing, texture analysis. It is a linear filter and its impulse response is derived from the multiplication of harmonic function and gaussian function. It is a band pass filter. It is used to increase the contrast between the nodule areas and other structure around it. The mathematical expression for the gabor filter is given in equation (2).

$$g(x, y) = \exp\left(-\frac{x'^2+\gamma^2 y'^2}{2\sigma^2}\right) \cos\left(2\pi \frac{x'}{\lambda} + \varphi\right) \qquad (2)$$

where,

$$x' = x\cos\theta + y\sin\theta \qquad (3)$$

$$y' = -x\cos\theta + y\sin\theta \qquad (4)$$

λ represents the wavelength of the sinusoidal wave.

θ denotes for the orientation of normal to parallel stripes of gabor function.

φ denotes phase offset.

σ denotes standard deviation

γ is spatial aspect ratio

The output image of gabor filter is shown in figure 4.

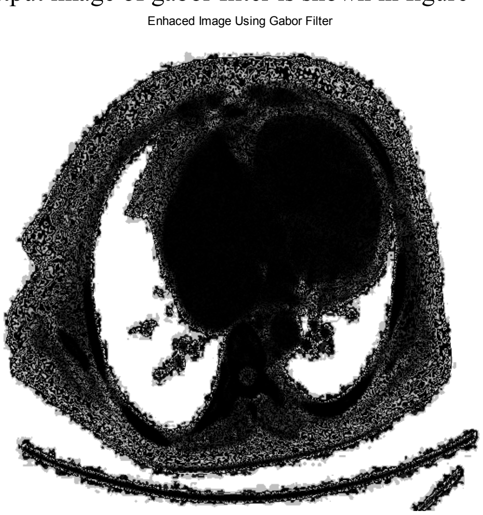


Figure 4. Enhanced Image using Gabor Filter

C. Image segmentation

Segmentation is used to divide an image into different small regions or objects. It has many applications in the medical field for the segmentation of the 2D medical images.

It is an important process for most image analysis following techniques. There are various methods available for image segmentation. In this paper, thresholding and marker controlled watershed segmentation methods are used. Thresholding is the most effective tool for the image segmentation purpose. It is used to converts a gray scale image into a binary image. These two levels are assigned to pixels, below or above the particular threshold value. The image obtained from thresholding segmentation has smaller storage space, fast processing speed and ease in manipulation compared with gray level image which usually has 256 levels. The output of threshold image is shown in figure 5.

Figure 5. Threshold based segmented Image

The concept of watershed is well known in topography. Watershed segmentation is used to extract the region minimum value from an image. It determines the corresponding to the dividing line with the least value. Dividing line in the image gives the rapid change of boundary. This transform finds catchment basins and watershed edge lines in the image. It treats the image as a plane, where light pixels are high and dark pixels are low. The important drawback associated to the watershed transform is the over segmentation that usually results. The output of watershed segmented image is shown in figure 6.

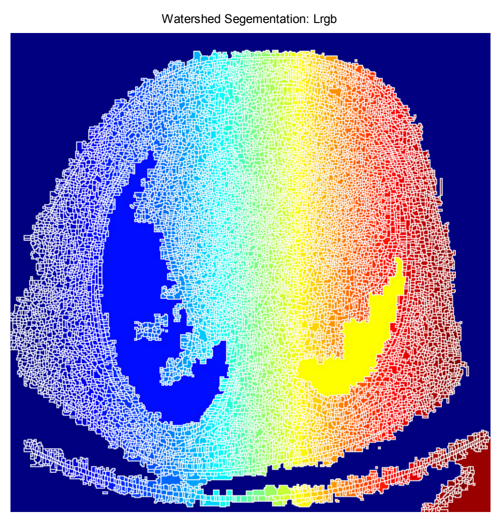


Figure 6. Watershed segmented Image

To overcome the drawbacks of this watershed segmentation i.e. over segmentation, the marker based watershed segmentation technique is used. It can segment boundaries from an image. Morphological operations are performed on the watershed segmented image to get final segmented image. Here the method is to use morphological operations called opening by reconstruction and closing by reconstruction to clean up the image. These operations will generate flat maxima inside each object which is located using imregionalmax. The output of marker based watershed segmented image is shown in figure 7.

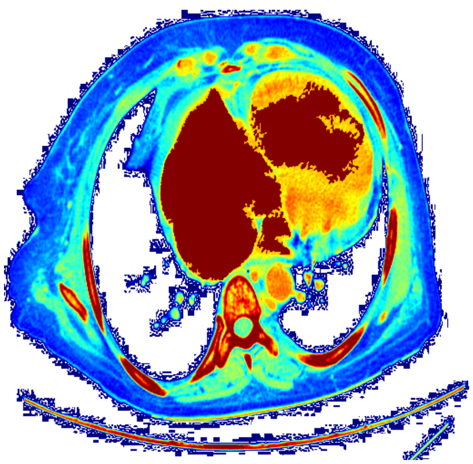


Figure 7. Marker-Controlled Watershed based segmented Image

After segmentation process, binarization process is done. In this approach, the total number of black pixels & white pixels are counted. If the total number of black pixels of input image is more than threshold, then the tumour is normal tumour. Otherwise, if the total number of the black pixels is less than the threshold then the tumour is cancerous tumour.

By using this result, sensitivity and accuracy are calculated. To find these parameters, first calculate some of the conditions like true positive, false negative, true negative and false positive.

$$\text{Sensitivity} = TP/ (TP + FN) \quad (5)$$

$$\text{Accuracy} = (TN + TP)/ (TN + TP + FN + FP) \quad (6)$$

here, TP is True Positive, TN is True Negative, FN is False Negative and FP is False Positive. Sensitivity is defined as amount of true positives that are correctly recognized by a diagnostic test. Accuracy is defined as the amount of true results, which is either true positive or true negative. It measures the degree of reliability of a diagnostic test on a condition.

In the CT scan lung image dataset, total 14 images are available. With the help of this dataset, the values of TP, FP, FN, TN, Sensitivity and accuracy are determined and given in Table 1.

Table 1. Detection accuracy of the proposed approach

Evaluation metrics	Thresholding	Marker controlled watershed transform
True Positive	13	13
False Positive	1	0
True Negative	0	1
False Negative	0	0
Sensitivity	1	1
Accuracy	0.92	1

By analysis, it is found that the Marker Controlled Watershed Transform performs superior than the existing algorithms. The results of the analysis have proved 100% of accuracy in Marker Controlled Watershed Transform algorithm by detecting the tumours from the CT scan lung images.

The features for example average intensity, area, perimeter and eccentricity are extracted and according to these features, tumour will be classified into the stages of cancer.

1. In stage I, the cancer tumour is small.
2. In stages II, the cancer tumour is limited to the chest.
3. In stage III, cancer tumour reaches to other parts of the body.

CONCLUSION

The system consists of pre-processing, segmentation, feature extraction and final classification. The proposed marker controlled watershed segmentation technique separates the touching objects in the image. It provides best identification of the main edge of the image and also avoids over segmentation. It gives 100% accuracy compared to the thresholding algorithm. So it is efficient for segmentation. The proposed technique gives very promising results comparing with other used techniques.

Acknowledgment

The author would like to thank respected mentor and guide Miss.N.D. Thombare, Assistant Professor, Department of Electronics and Telecommunication Engineering, Dr.S.S.Lokhande, Associate Professor in Department of Electronics and Telecommunication Engineering, Dr.M. B. Mali (Head of Department), Dr.S. D. Lokhande (Principal)of Sinhgad college of Engineering, Vadgaon(BK),Pune 41 for the technical advice, encouragement and constructive criticism, which motivated to strive harder for excellence.

.

REFERENCES

[1] Anita chaudhary, SonitSukhraj Singh “Lung Cancer Detection on CT Images by Using Image Processing”2012 International Conference on Computing Sciences

[2] Nihad Mesanovic, Haris Huseinagic, Matija Males, , Mislav Grgic, Emir Skejic, Muamer Smajlovic "Automatic CT Image Segmentation of the Lungs with Region Growing Algorithm"

[3] Sayani Nandy, Nikita Pandey "A Novel Approach of Cancerous Cells Detection from Lungs CT Scan Images'' International Journal of Advanced Research in Computer Science and Software Engineering Volume 2, Issue 8, August 2012

[4] Prof. Samir Kumar Bandyopadhyay "Edge Detection From Ct Images Of Lung'' International Journal Of Engineering Science & Advanced Technology Volume - 2, Issue - 1, 34 – 37

[5] FatmTaher, NaoufelWerghi and Hussain Al-Ahmad "Extraction of Sputum Cells using Thresholding Techniques for Lung Cancer Detection" 2012 International Conference on Innovations in Information Technology

[6] QinghuaJi,Ronggang Shi "A Noval Method of Image Segmentation Using Watershed Transformation"2011 International Conference on Computer Science and Network Technology

[7] Pinaki Pratim Acharjya, Santanu Santra, Dibyendu Ghoshal "An Improved Scheme on Morphological Image Segmentation Using the Gradients"(IJACSA) International Journal of Advanced Computer Science and Applications Vol.4, No.2,2013

[8] FAN Li-nan , LI Dao-jing, SUN Shen-shen, CHANG Chao-hai "Applied research on the automatic detection of lung nodules ROI based on Top-hat and Gabor filter" Third International Symposium on Information Processing IEEE 2010

[9] Rahil Garnavi, Ahmad Baraani-Dastjerdi, Hamid Abrishami Moghaddam2, Masoomeh Giti, Ali Adjdari Rad "A New Segmentation Method for Lung HRCT Images" Proceedings of the Digital Imaging Computing: Techniques and Applications (DICTA 2005) IEEE 2005

[10] Disha Sharma, Gagandeep Jindal"Identifying Lung Cancer Using Image Processing Techniques" International Conference on Computational Techniques and Artificial Intelligence (ICCTAI'2011)

[11] K. Parvati, B. S. Prakasa Rao, and M. Mariya Das Image Segmentation Using Gray-Scale Morphology and Marker-Controlled Watershed Transformation Hindawi Publishing Corporation Discrete Dynamics in Nature and Society Volume 2008