LUNG CANCER DETECTION USING DIGITAL IMAGE PROCESSING

Anita Chaudhary*

Sonit Sukhraj Singh*

ABSTRACT

In recent years the image processing mechanisms are used widely in several medical areas for improving earlier detection and treatment stages, in which the time factor is very important to discover the disease in the patient as possible as fast, especially in various cancer tumors such as the lung cancer, breast cancer. Lung cancer has been attracting the attention of medical and sciatic communities in the latest years because of its high prevalence allied with the difficult treatment. Statistics from 2008 indicate that lung cancer, throughout world, is the one that attacks the greatest number of people. Early detection of lung cancer is very important for successful treatment. Diagnosis is mostly based on CT images. Our current work focuses on finding nodules, early symptoms of the diseases, appearing in patients' lungs. We use a modified Watershed segmentation approach to isolate a lung of an CT image, and then apply a small scanning window to check whether any pixel is part of a disease nodule. Most of the nodules can be detected if process parameters are carefully selected. We are aiming at computerizing these selections. We passed the available lung cancer images and its database in basic three stages to achieve more quality and accuracy in our experimental results: pre-processing stage, feature Extraction stage and Lung cancer cell identification.

* Department of Electronics and Communication Engineering, Lovely Professional University, Phagwara, Punjab.

1.INTRODUCTION

1.1 Lung Anatomy

The lungs are a pair of sponge-like, cone-shaped organs. The right lung has three lobes, and is larger than the left lung, which has two lobes. Oxygen is brought into the lungs when air is inhaled. Lung tissue transports oxygen to the bloodstream to go to the rest of the body. Cells release carbon dioxide as they use oxygen. The bloodstream carries carbon dioxide back to the lungs and the carbon dioxide leaves the body when air is exhaled.

1.2 Lung Cancer

Lung cancer is a disease of abnormal cells multiplying and growing into a tumor. Cancer cells can be carried away from the lungs in blood, or lymph fluid that surrounds lung tissue. Metastasis occurs when a cancer cell leaves the site where it began and moves into a lymph node or to another part of the body through the bloodstream.

1.3 Lung Cancer Types

Cancer that starts in the lung is called primary lung cancer. There are several different types and these are divided into two main groups

- Small cell lung cancer
- Non small cell lung cancer

1. Small cell lung cancer

About 20 out of every 100 lung cancers diagnosed are this type. Small cell lung cancer is called this because the cancer cells are small cells that are mostly filled with the nucleus. This type of cancer is usually always caused by smoking. It is very rare for someone who has never smoked to develop it. Small cell lung cancer often spreads quite early on and so doctors often suggest chemotherapy treatment rather than surgery.

2. Non small cell lung cancer

There are three types of non small cell lung cancer. These are grouped together because they behave in a similar way and respond to treatment in a different way to small cell lung cancer.

2. LUNG CANCER DETECTION SYSTEM

LCDS system uses convolution filters with Gaussian pulse to smooth the cell images. The contrast and color of 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. The extracted morphologic features include the perimeter, area, roundness, and rectangleness of the nucleus. The extracted colorimetric features include the red component, green component, blue component, illumination, saturation, difference between red and blue components, and proportion of blue component of the nucleus.

Also the red component, green component, and blue component of the entire image are included as colorimetric features. On this basis, a lung cancer cell identification module 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 Fig 1.

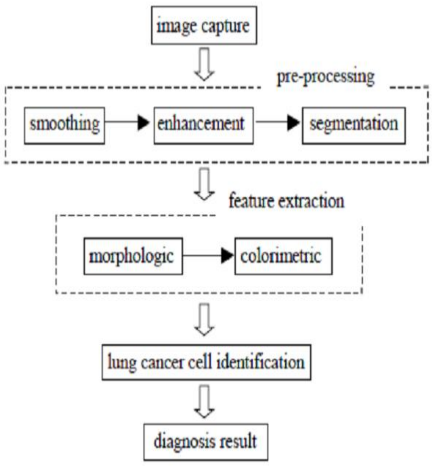

Fig.1 Lung cancer detection system

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[3].

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: Gobor filter, auto-enhancement and Fast Fourier transform techniques.

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 marker controlled 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.

Marker-driven watershed segmentation extracts seeds indicating the presence of objects or background at specific image locations.

3. 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.

Image enhancement techniques can be divided into two broad categories:

1. Spatial domain techniques, which operate directly on pixels.
2. Frequency domain techniques, which operate on the Fourier transform of an image.

As we work on medical images, we tried out three types of enhancement techniques: Gabor filter ,Fast Fourier transform ,and Auto enhancement.

To talk about the better method we should use a theory to determine this unfortunately, there is no general theory for determining what `good' image enhancement is when it comes to human perception.

Spatial domain methods

The value of a pixel with coordinates (x; y) in the enhanced image ^ F is the result of performing some operation on the pixels in the neighborhood of (x; y) in the input image, F. Neighborhoods can be any shape, but usually they are rectangular.

Frequency domain methods

Image enhancement in the frequency domain is straight forward. We simply compute the Fourier transform of the image to be enhanced, multiply the result by a filter (rather than convolve in the spatial domain), and take the inverse transform to produce the enhanced image.

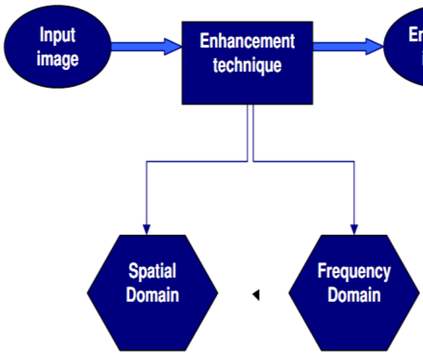

Fig.2 Flow Chart Of Image Enhancement Techniques[2].

3.1 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.

The image presentation based on Gabor function constitutes an excellent local and multi-scale decomposition in terms of logons that are simultaneously (and optimally) localization in space and frequency domains.

A Gabor filter is a linear filter whose impulse response is defined by a harmonic function multiplied by a Gaussian function. Because of the multiplication-convolution property (Convolution theorem), the Fourier transform of a Gabor filter's impulse response is the convolution of the Fourier transform of the harmonic function and the Fourier transform of the Gaussian function.

3.2 Auto Enhancement technique

Auto enhancement, automatically adjusts and enhances the image (brightness, color and contrast) to optimum levels. This method strongly depends on statistical operations such as mean, variance calculation.

3.3 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). The FFT utilizes some clever algorithms to do the same thing as the DTF, but in much less time.

The Fourier transform is used in a wide range of applications, such as image analysis, image filtering, image reconstruction and image compression.

4. IMAGE SEGMENTATION

4.1 Image segmentation technique

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.

The goal of segmentation is to simplify and/or change the representation of an image into something that is more meaningful and easier to analyze. Image segmentation is typically used to locate objects and boundaries (lines, curves, etc.) in images. More precisely, image segmentation is the process of assigning a label to every pixel in an image such that pixels with the same label share certain visual characteristics.

The result of image segmentation is a set of segments that collectively cover the entire image, or a set of contours extracted from the image (edge detection). Each of the pixels in a region is similar with respect to some characteristic or computed property, such as color, intensity, or texture. Adjacent regions are significantly different with respect to the same characteristic(s)[1].

4.2 Segmentation algorithms

The Segmentation algorithms are based on one of two basic properties of intensity values discontinuity and similarity. First category is to partition an image based on abrupt changes in intensity, such as edges in an image , Second category are based on partitioning an image into regions that are similar according to a predefined criteria. Histogram Thresholding approach falls under this category.

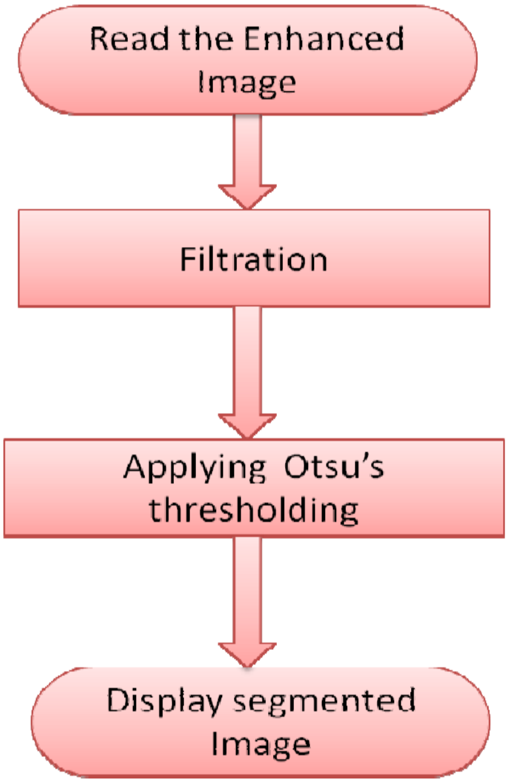

Fig.3 Otsu's Thresholding Segmentation Flow Chart

4.3 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 are below or above the specified threshold value.

We used Otsu's method using (gray thresh) function which Compute global image threshold using Otsu's method. Otsu's method is based on threshold selection by statistical criteria. Otsu suggested minimizing the weighted sum of within-class variances of the object and background pixels to establish an optimum threshold. Recall that minimization of within-class variances is equivalent to maximization of between-class variance. This method gives satisfactory results for bimodal histogram images. Threshold value based on this method will be between 0 and 1, after achieve this value we can segment an image based on it.

4.4 Marker-Controlled Watershed Segmentation approach

Separating touching objects in an image is one of the more difficult image processing operations. The water shed transform is often applied to this problem. The marker based watershed segmentation can segment unique boundaries from an image.

The strength of watershed segmentation is that it produces a unique solution for a particular image. The placement of internal and external markers into regions of interest can easily cope with the over-segmentation problem. Another disadvantage of watershed segmentation, again related to the image noise and the image's discrete nature, is that the final boundaries of the segmented region lack smoothness.

5. IMAGE FEATURES 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.

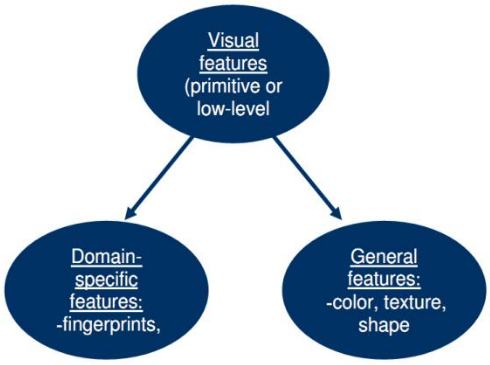

Fig.4 Visual Feature Extraction types

The issue of choosing the features to be extracted:

The features should carry enough information about the image and should not require any domain-specific knowledge for their extraction.

Sequence of stages started from image enhancement, image segmentation and cropping, finally feature extraction get introduced. Feature extraction is an essential stage that represents the final results to determine the normality or abnormality of an image .

Two approaches to predict the probability of lung cancer presence; first approach is Binarization and the second is masking, Both of these methods are based on facts strongly related to lung anatomy and related information of lung CT imaging[5].

5.1 Binarization Approach

This approach depends on the fact that the number of black pixels is much more than white ones in a normal lung image, so that we started to count the black pixels for normal and abnormal images to get the average which will be denoted later as threshold then each image black pixels will be compared to this threshold, whether it is greater, then it is normal, else the opposite.

5.2 Masking Approach

This method depends on the fact that the masses are appeared as white connected areas inside ROI (lungs) as they increase the percent of cancer presence increase.

Combining the two approaches together will lead us to take a decision weather the case is normal or not .

6. 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 three stages: Image Enhancement stage, Image Segmentation stage and Features Extraction stage.

REFERENCES

1. 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.
2. B. Zhao, G. Gamsu, M. S. Ginsberg, L. Jiang and L. H. Schwartz, "Automatic detection of small lung nodules on CT utilizing a local density maximum algorithm", journal of applied clinical medical physics, vol. 4, (2003).
3. Kakeda, S. et al., "Improved Detection of Lung Nodules on Chest Radiographs Using a Commercial Computer-Aided Diagnosis System", American Journal of Roentgenology, 182, February 2004, pp. 505-510.
4. R. Wiemker, P. Rogalla, T. Blaffert, D. Sifri, O. Hay, Y. Srinivas and R. Truyen "Computer-aided detection (CAD) and volumetry of pulmonary nodules on high-resolution CT data", (2003).
5. Raicu, D.S. Image Feature Extraction. 2004 http://facweb.cti.depaul.edu/research/vc/VC_Workshop/presentations/pdf/daniela_tutorial2.pdf.