



## Lung Cancer Detection on CT Images by using Image Processing

Bindiya Patel<sup>1</sup>, Dr. Pankaj Kumar Mishra<sup>2</sup>, Prof. Amit Kolhe<sup>2</sup>

<sup>1</sup>Department of Digital Electronics, <sup>2</sup>Department of Electronics and Telecommunication

<sup>1,2</sup>Rungta College of Engineering and Technology, Bhilai, Chhattisgarh, India

### ABSTRACT

This project is mainly based on image processing technique. In this work MATLAB have been used through every procedure made. Image processing techniques are widely use in bio-medical sector. The objective of our work is noise removal operation, thresholding, gray scale imaging, histogram equalization, texture segmentation, and morphological operation. Detection of lung cancer from computed tomography (CT) images is done by using MATLAB software. By using these methods the work has been done on CT images and the final tumor area has been shown with pixel values.

**Keywords:** Image Acquisition, Image enhancement, Image Segmentation, Morphological operation

### I. INTRODUCTION

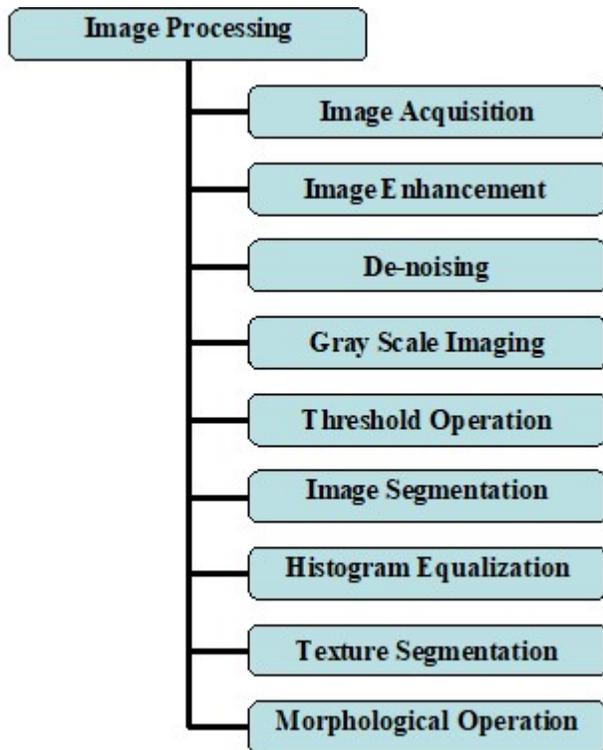
In this project we are detecting the lung cancer from the computed tomography (CT) images by using image processing technique in MATLAB. First of all we must know that what lung cancer is, so Lung cancer is a disease in which abnormal cells multiplying and growing and forms a tumor in lungs. There are different types of tumor and not all tumors are cancerous some are the basic tumor which can be cure by some basic treatments. Also Some cancer cells can be spread to other body parts and some are not for example: - Survival Cancer cells can be carried away from the lungs in blood, or lymph fluid that surrounds with lung tissue, on the other hand benign tumors are the tumors which do not spread to other part of the body. There are several types of lung cancer and these are mainly divided into two main categories this are: - small cell and non-small cell lung cancer. But people do have a higher chance of

survival from the lung cancer if the cancer can be detected in the early stages. Survival from lung cancer is directly related to its speed of growth and at its detection time as soon as it can be detected the chance of survival will be increase. This project is starts with collecting a number of computed tomography (CT) scanned images from the available data base. This images will be further being processed, enhanced, and segmented than load the images into mat lab for cancer detection and then after comparison classify into normal and abnormal tumor. This techniques helps to detects cancer and help us for diagnosis solution. This computed tomography (CT) scanned images are used as an input image, after getting the input image we removed the noise from the input image by using different filtration technique. In next step we do the gray scale imaging and then thresholding operation is done and after that we apply the histogram equalization, these all above operations are come under the image acquisition and image enhancement. In next step image segmentation will be done the segmentation is done, there are different types of image segmentation are available. We are computing the texture segmentation technique to the image. After that we do the morphological operation to the image so that we can get a clear and accurate region of the tumor. We are aiming to get more accurate result by using image enhancement technique and image segmentation operation and by the comparison of effected area so that intensity of cancer can be classified. We can also use MRI images, X-ray images of lung for the cancer detection as an input image instead of using computed tomography (CT) .

## 2. METHODOLOGY

This project consists of four major stages, the first stage is Image Acquisition, the second stage is Image Processing techniques, third stage is consist of image segmentation operation and the fourth or last stage is image extraction, and comparison. All the four stages are having some basic operations and steps which are necessary to full fill the requirements and to complete the stage step by step.

Block diagram of proposed system is shown below:-



The description of all the stages and steps are given below:-

### A.) IMAGE ACQUISITION

The first stage of any image processing system involves image acquisition, after the image has been obtained further operations are applied. The aim of image acquisition is to get the image of required area or effected region so that the detection can be done. It starts with collecting a computed tomography (CT) images of lung of different person from the record or available data base. This computed tomography (CT) images are further used as input to the system. After image acquisition we can proceed to image processing stage for further operations.

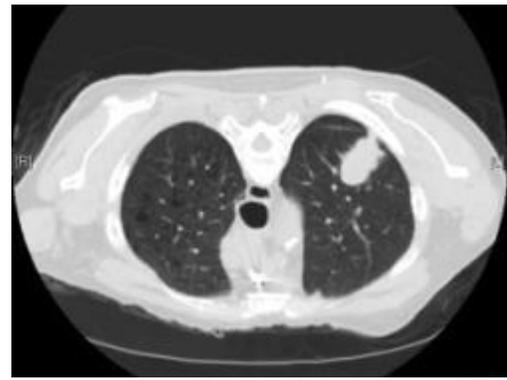


FIGURE1-LUNG CT IMAGE

### B.) IMAGE ENHANCEMENT

The second stage is an image enhancement. Image enhancement is a technique which is used to improve the quality of the image and to get the better image than the provided one, it provides a clear better and the accurate parameter of the desired region. For this purpose noise removal from the images, image filtering, techniques are use, which will helpul to detect cancer parameter during processing. Image processing involves two main steps that are; image enhancement technique and image segmentation technique both are having their own properties and important role for improving the quality of the image. Both the process having a different- different technique for the image enhancement and segmentation for the more accurate result the best one will be choose. In this stage we use the different techniques to make the image better and enhance it from noising, corruption or interference. Enhancement technique provides better input for other automated image processing technique. For image enhancement first we use different types of filtration methods for the removal of noise from the image Ex Linear filter, median filter, high pass filter and adaptive filter, in next step thresholding operation has been done and then we convert the input image to gray scale image.

#### a.) DE-NOISING

Digital images can have various types of noise. This noise can be the result of error in the image processing and segmentation and some other further operations that result in the pixel values that do not true intensity of real image. This noise may leads to interrupted o false values which may give the false information about the tumor and the person can be misguided so the removal of noise is necessary. There are several ways by which the noise can be introduced into the images, depending on the image is created for example: - (1) noise can be the result of damaged

computed tomography scanned film or introduced by scanner itself. (2) The another reason can be the mechanism for gathering the data. (3) An electronic transmission of image data can introduce the noise etc. As we know the removal of noise is very important so we use different types of filtration method to remove the noise from the picture for example removal of noise by linear filter, removal of noise from an average filter and median filter, removal of noise by using adaptive filter. We are using median filter for the removal of noise, but we can use any of the method from above three. The median filter is a nonlinear digital filtering technique, often used to remove noise. Such noise reduction is a typical pre-processing step to improve the results of later processing for example, edge detection on an image. In median filtering the value of an output pixel is determined by the median of neighbourhood pixels, rather than the mean values of the pixels. The median is much less sensitive than the mean to extreme values therefore it is better able to remove outliers without reducing the sharpness of the image. So that Median filtering is very widely used in digital image processing because, under certain conditions, it preserves edges while removing noise. The main idea of the median filter is to run through the signal entry by entry, replacing each entry with the median of neighbouring entries. The pattern of neighbours is called the "window", which slides, entry by entry, over the entire signal. For 1D signal, the most obvious window is just the first few preceding and following entries, whereas for 2D (or higher-dimensional) signals such as images, more complex window patterns are possible (such as "box" or "cross" patterns). The concept of median filter is that if the window has an odd number of entries, then the median is simple to define: it is just the middle value after all the entries in the window are sorted numerically. For an even number of entries, there is more than one possible median. This filter enhances the quality of the MRI image. (REFERENCE- International Journal of Electronics, Communication & Soft Computing Science and Engineering Rajesh C.Patil, Dr. A. S. Bhalchandra ISSN: 2277-9477, Volume2, Issue1). In our project we are using the median filter for the removal of noise.

## FIGURE2- FILTERED IMAGE

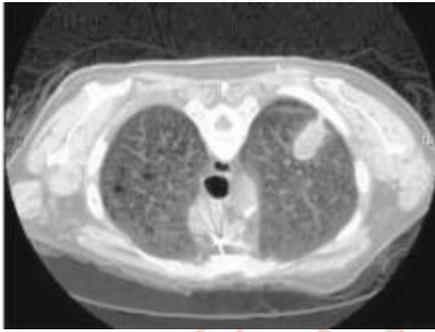
### b.) GRAYSCAL IMAGING

Computed tomography (CT) scanned images are combination of a series of x-ray images taken from the different angles and uses some computational processing to create cross sectional images of specified area or the image of required body part. The computed tomography (CT) images are black and white images in general. When we take these images as input images on computer, computer considers these images as a black and white image. So we apply gray scale imaging to the image.

Gray scaled images are not like simple black and white images it provides a combination of black and white or we can say a gray shade instead of providing only two shades black and white. On images gray scale or grey scale is one in which the amount of each pixel is single sample represents the amount of light it contains or we can say that, it carries only the intensity information. These are from black and white to exclusive shades of gray, varying from black at the lowest intensity to white at the highest intensity. Gray scale images are distinct from bit by bit on black and white images. The illusion of gray scale shading in a half tone image is obtain by rendering the image as a combination of black dots on white background or vice versa, and these gray scale images are result of measuring the intensity of light at each pixel according to a particular weighted combination of frequencies or wavelengths.

The RGB is a primary color brightness level in RGB are represented in number form as 0 to 255 in analog or we can say it has 255 level and in digital it represents in binary form as 00000000 to 11111111, where black is represented by  $R=B=G=0$  or  $R=G=B=00000000$ , and white is represented by  $R=G=B=1$  or  $R=G=B=1$ . Because there is 8bit in binary representation of the gray level so that it is also called as 8-bit gray scale. Array of class uint8, uint16, int16, single or double whose pixel values specify intensity value. For single or double arrays, values

range from [0, 1]. For uint8, values range from [0,255]. For uint16, values range from [0, 65535]. For int16, values range from [-32768, 32767]. At present, the most commonly used storage method is 8-bit storage, which have 256gray level intensity of each pixel can have from 0 to 255, with 0 being black and 255 being white. (REFERENCE- "A theory based on conversion of RGB image to Gray image" by- Tarun Kumar and Karun verma, computer science and engineering department, @ International journal of computer application sep2010).



**FIGURE3- GRAY SCALE IMAGE**

### C.) THRESHOLDING

Image thresholding is a simple way of partitioning an image into a foreground and background. Common image thresholding algorithms include histogram and multilevel thresholding. Now we know the main purpose of thresholding, so the working of thresholding is – as we know the simplest property that pixels in a region can share is intensity. So, thresholding operation segments such regions and separate light and dark regions. It creates binary images from grey-level ones by turning all pixels below some threshold to zero and all pixels about that threshold to one or apart the dark and lighter area from each other. Let's assume if  $g(x, y)$  is a threshold version of  $f(x, y)$  at some global threshold 'T' that separates these modes. Then any point  $(x, y)$  for which  $f(x, y) > T$  is called any object point; otherwise it is back ground point. High intensity areas mostly Comprises of cancer cell.(REFERENCE- International Journal of Emerging Technology and Advanced Engineering Website: [www.ijetae.com](http://www.ijetae.com), ISSN 2250-2459, ISO 9001:2008 Certified Journal, Volume 7, Issue 7, July2017). Image thresholding is most effective in images with high level of contrast. This image analysis technique is a type of image segmentation that isolates objects by converting Gray scale images into binary images which is the next step of our project.

### 1) 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 Segmentation divides an image into its constituent regions or objects as well as it can detect the edge of the images. Image segmentation is a technique which is used for separating the image from the background as well as from each other or we can say that to separate the image we are determining the outline of the image using threshold operation., this process is done by classified the pixels into objects. To divide and segment the enhanced image generally histogram equalization, threshold segmentation, region based segmentation method and either watershed approaches or texture segmentation can be used here we are using histogram technique and after histogram we will go for texture segmentation technique.

#### a.) HISTOGRAM TECHNIQUE

Histogram equalization technique is used for the segmentation of the image; it is one of the most effective techniques for segmentation. Histogram equalization of an image shows the pixels intensity values. For example generally it forms a graph in which x-axis shows the gray level intensities and the y-axis shows the frequency of these intensities. In general, a histogram is the estimation of the probability distribution of a particular type of data. An image histogram is a type of histogram which offers a graphical representation of the tonal distribution of the grey values in a digital image. To improve the contrast of the image through histogram equation, it spreads out intensity values along the total range of value in order to achieve higher contrast. The methods of histogram equation are: histogram expansion, local area histogram equalization (LAHE), cumulative histogram equalization, par sectioning, and odd sectioning. (REFERENCE- Histogram Equalization, by- Robert Krutch and David Tenorio, Microcontroller Solution group Guadalajara@ June 2011, free scale semiconductor, Inc.) The histogram can have many uses in image processing apart from image segmentation for example it can be used for image processing, can be used for brightness purpose not only for brightness purpose can also be used for adjusting the contrast level, and last but no not the least it is widely used for segmentation.

#### b.) TEXTURE SEGMENTATION

The texture is most important attribute in many image analysis or computer vision applications. It is a set of metrics calculated in image processing to quantify the

texture of an image. Texture of image gives us information the spatial arrangement of colour or selected region of an image. The procedures developed for texture problem can be subdivided into four categories: structural approach, statistical approach, model based approach and filter based approach. Different definitions of texture are described, but more importance is given to filter based methods. Such as Fourier transform, Gabor, Thresholding, Histogram and wavelet transforms. An image texture can be used in segmentation or classification of an image, or to extract boundaries between major texture regions. For more accurate result in segmentation the most useful features are spatial frequency and an average gray level. Texture is a difficult concept to represent. The identification of specific textures in an image is achieved primarily by modelling texture as a two-dimensional gray level variation. The relative brightness of pairs of pixels is computed such that degree of contrast, regularity, coarseness and directionality. There are two main types of texture segmentation that are region based and boundary based texture segmentation.

**Region Based-** it attempts to group or cluster pixels based on texture property. Segmentation algorithms operate iteratively by grouping together pixels which are neighbours and have similar values and splitting groups of pixels which are dissimilar in value.

**Boundary Based-** Edges contain some of the most useful information in an image. We may use edges to measure the size of objects in an image; to isolate particular objects from their background; to recognize or classify objects. In boundary based it attempts to group or cluster pixels based on edges between pixels that come from different texture properties. (REFERENCE- Texture segmentation: different methods, by- Vajjinath V. Bhosle, V Rushsen P. Pawar @ nov2013 International journal of soft computing and engineering.)

#### **D.) MORPHOLOGICAL OPERATION**

This is the last step for the detection of lung cancer. This stage is an important stage that uses algorithms and techniques to detect and isolate various desired portions or shapes of a given image. It is used to predict the probability of lung cancer presence when the input data to an algorithm is too large to be processed and it is suspected to be notoriously redundant, then the input data will be transformed into a reduced representation set of features. From all of the above steps like image processing, image segmentation, we get the clear image of the tumor region in lung, so differentiate the tumor in lung are called morphological operation. The basic characters for the morphological operation are area for which the numbers of iterations are performed. This are the values which we calculate or the area or region of the tumor which we are obtained from enhanced and segmented images and also from morphological or thresholding. These features are measured in scalar. After getting the tumor region we compare the tumor with the standards and try to find the type of the tumor and from the size of the tumor we try to find the stage of the cancer, because from all of this information are very important because it will use in the treatment of the cancer and from this information the required steps and cure will be taken for example Lung nodule is defined as 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 greater than 25mm. With the help of classifications and comparison in the classification stage Tumor is classified as normal Cancer Tumor or abnormal Cancer Tumor.

#### **FIGURE5- FINAL TUMOR AREA**

### **3. RELEATED WORK**

#### **FIGURE4- TEXTURE SEGMENTATION**

AUTHOR	IMAG E	TECHNIQUE	ACCUR ACY
Siva Sakthi, Kumar Parasuraman, Arumuga Maria Devi	CT	Otsu threshold, watershed segmentation	90.90%
S.shiva kumar	CT	SVM,RBF kernel	80.36%
Disha sharma, Gagandeep Jindal	CT	CAD, Weiner filter	80%
Aniket Gaikwad, Azharuddin Inamdar, Vikas Behera	CT	Histogram and watershed segmentation	84.55%
M. Premchander, Dr. .M. Venkateshwarra, dr. T.V. Rajinikanth	CT	Gabor filter, watershed segmentation	86.39%
Anuradha S. Deshpande, Dhanesh D. Lokhande, Rahul P. Mundhe, Juilee M.Ghatole	CT, MRI	Watershed segmentation, SVM algorithm	90.90%
J.R Marsilin	CT	SVM algorithm	78.00%
Yaoying Huang, Wangsen Li, Xiaojiaoye	CT	Genetic algorithm, feature selection	99.1%
Fatma Taher	Sputum	Bayesian	88.62%
Afazan Adam	CT	Genetic algorithm, back propagation neural network	83.86%
Yang Hiu	CT	SVM(GBRF kernel type)	87.82%

#### 4. RESULT AND DISCUSSION

As we know lung cancer is one of the most dangerous diseases in the world. An image improvement technique is developing for earlier disease detection and treatment stages. Correct Diagnosis and early detection of lung cancer can increase the survival rate. Image quality and accuracy is the core factors of this research, image quality assessment as well as enhancement stage. This project is based on the processing of computed tomography (CT) images. We also conclude that the lung cancer can be detected in an early stage by using any one of this method and by following all the four steps mentioned above. As we can see the median filter is chosen. After image processing we go for segmentation approach here we use histogram and texture segmentation. After segmentation of image, morphological operation is used to get individual lung and to eliminate unnecessary parts. By doing morphological operations, we get not only the individual lung but also apparent the lung nodule then we extracts the Tumor by comparison. The area calculated by the process is 1488 pixels. In this the resulting tumours are of different dimensions by measuring the area of Tumor, so the lung cancer stage can be detected accurately in early stage using the proposed methodology cancer detection and respective diagnosis measure which will helps to clear cancer Parameters permanently. The result are analysed graphically as well as numerically.

#### 5. FUTURE SCOPE

We are aiming to get the more accurate results by using various enhancement and segmentation technique, different segmentation strategies and calculations are the root idea of digital image processing the more accurate result will be more helpful and good for the diagnosis solution and the person can have more chances of survival from this dangerous disease and we can do one more thing apart from using different strategies we can use a fusion method of all this techniques or the hybrid methods to get the more accurate result. We can also develop the system as a real time system in future. It means the system will work at the time of diagnosis as well as with the time when we take the computed tomography (CT) images, the advantage of the real time system will be that it helps the person to cure the disease as soon as possible and provides a help for early treatment so the survival chance can be increase. In future by parameter and area calculation of the

tumor at the time of detection we can also find that tumor has been in which stage.

## References

- 1) International Journal of Electronics, Communication & Soft Computing Science and Engineering Rajesh C.Patil, Dr. A. S. Bhalchandra ISSN: 2277-9477, Volume2, Issue1.
- 2) International Journal of Emerging Technology and Advanced Engineering, Detection of Lung Cancer in Medical Images Using Image Processing Techniques, M. Premchander, Dr. M. Venkateshwara Rao, Dr. T. V. Rajinikanth, Certified Journal, Volume 7, Issue 7, July2017.
- 3) Histogram Equalization, by- Robert Krutch and David Tenorio, Microcontroller Solution group Guadalajara@ June 2011, free scale semiconductor, Inc.
- 4) Image Segmentation, Yu-Hsiang Wang, Graduate Institute of Communication Engineering, National Taiwan University, Taipei, Taiwan, ROC.
- 5) A theory based on conversion of RGB image to Gray image” by- Tarun Kumar and Karun verma, computer science and engineering department, @ International journal of computer application, volume-7, no-2, sep2010.
- 6) Book- MATLAB a practical approach by STORMY ATTAWAY @ 2009.
- 7) Anita Chaudhary, Sonit Sukhraj Singh “Lung Cancer Detection on CT Images by Using Image Processing Technique.” Published on -24<sup>th</sup> DECEMBER, 2012.
- 8) Gawade Prathmesh Pratab, R.P Chauchan “Detection of Lung Cancer Cells using Image Processing Techniques.” Published on- 16<sup>th</sup> FEBRUARY, 2017.
- 9) Gonzalez R.C., Woods R.E., Digital Image Processing, Upper Saddle River, NJ Prentice Hall, 2008.
- 10) International Research Journal of Engineering and Technology Volume: 03 Apr-2016 Lung cancer detection using digital Image processing On CT scans Images, Aniket Gaikwad, Azharuddin Inamdar, and Vikas Behera.
- 11) Nunes É.D.O., Pérez M.G., Medical Image Segmentation by Multilevel Thresholding Based on Histogram Difference, presented at 17th International Conference on Systems, Signals and Image Processing, 2010.
- 12) Lung Cancer Detection Using Image Processing Techniques Mokhled S. AL-TARAWNEH Computer Engineering Department, Faculty of Engineering,@ Leonardo Electronic Journal of Practices and Technologies, January-June 2012.
- 13) Detection of Lung Cancer Using Marker-Controlled Watershed Transform, sayali satish kanitkar, N.D thombare, S.S lokhande@ 2015 International Conference on Pervasive Computing (ICPC).
- 14) International journal of engineering sciences & research technology “Review on lung cancer detection using image processing technique”, by- Anam Quadri, Rashida Shujae, Nishat Khan, feb2016.
- 15) Math Work and Wikipedia.
- 16) “Lung Tumor Detection and Segmentation in CT Images”, by- Preeti Katiyar, Dr. Krishna Singh, International Journal of Innovations & Advancement in Computer Science IJIACS, Volume 6, Issue 7 July 2017.
- 17) Texture segmentation: different methods, by- Vaijinath V. Bhosle, V Rushsen P. Pawar @ nov2013, International journal of soft computing and engineering.
- 18) International Advanced Research Journal in Science, Engineering and Technology Vol.3, August2016, Implementation of Lung Cancer Nodule Feature Extraction using Threshold Technique, T.Siva Sakthi, Kumar Parasuraman, Arumuga Maria Devi.
- 19) Lung Cancer Detection with fusion of CT and MRI Images Using Image Processing Prof. Anuradha S. Deshpande, Dhanesh D. Lokhande,Rahul P. Mundhe, Juilee M.Ghatole @International Journal of Advanced Research in Computer Engineering & Technology (IJARCET) Volume 4 Issue 3, March 2015.
- 20) Identifying Lung Cancer Using Image Processing Techniques, by-Disha Sharma, Gagandeep Jindal, @ International Conference on Computational Techniques and Artificial Intelligence2011.
- 21) A survey on threshold based segmentation technique in image processing, by- K.Bhargavi and S.Joyti @ nov2014 International journal of innovative research and development.