DETECTION OF LUNG CANCER IN CT IMAGE USING IMAGE SEGMENTATION TECHNIQUES

Dissertation Submitted to the Panchayat College in Partial Fulfilment of the Requirements for the Degree of

BACHELOR OF SCIENCE
IN
COMPUTER SCIENCE

ESTD: 1980

Submitted By

SAMIKSHYA PATTJOSHI

Roll No- 2215CSC011

Under the guidance
Of

Mr. DHIRAJ NAIK

Lecturer in Computer Science

DEPARTMENT OF COMPUTER SCIENCE
PANCHAYAT COLLEGE, DHARAMGARH
KALAHANDI- 766015.

(2022-2025)

CERTIFICATE

This is to clarify that the dissertation entitled **“DETECTION OF LUNG CANCER IN CT IMAGE USING IMAGE SEGMENTATION TECHNIQUES”** being submitted by **Miss Samikshya Pattjoshi** bearing roll no.2215CSC011 to the Panchayat College, Dharamgarh, Kalahandi for the degree of Bachelor of Science, is a record of bona-fide research and authentic work carried out by her under my supervision and guidance.

The matter presented in this dissertation has not been submitted for the award of any other degree of this or any other institute.

MR DHIRAJ NAIK
(Supervisor)
Dept. of Computer Science

DECLARATION

I declared that, present study entitled **"DETECTION OF LUNG CANCER IN CT IMAGE USING IMAGE SEGMENTATION TECHNIQUES"** submitted by me for the partial fulfilment of the requirements for the degree of Bachelor of Science in Computer Science is the original work carried out by me under the guidance and supervision of **Mr Dhiraj Naik**, Lecturer in Dept. of Computer Science, Panchayat College, Dharamgarh.

This work is based on my original research work and no part of the thesis has so far been submitted for the award of any other degree or diploma to Panchayat College, Dharamgarh or elsewhere.

Dharamgarh **Samikshya Pattjoshi**

Date: **Roll No-**2215CSC011

APPROVAL CERTIFICATE PAGE

This thesis entitled "Detection of lungs cancer in CT image using image segmentation techniques" prepared and submitted by Samikshya Pattjoshi in partial fulfillment of the requirements in research for the Degree of Bachelor of Secondary Education major in Computer Science has been examined and is recommended for approval and acceptance.

Signature of guide Mr Dhiraj Naik Lecturer in CSC	**Signature of HOD** Mr Binayak Panda Lecturer in CSC	**Signature of External**

ACKNOWLEDGEMENT

The successful completion of this dissertation would not have been possible without the support and cooperation of other. I know that this thesis was not my individual achievement, but this result of many people to whom I will for be forever grateful.

First and foremost, I would like to thank respectable and learned guide, **Mr Dhiraj Naik**, Lecturer in Computer Science for who's untiring, preserving and unflinching help and for being a constant source of inspiration for and sharing their ideas with me.

I am also grateful to **Mr Binayak panda**, HOD, Lecturer in Computer Science, who have taken and given the work final shape in a short manner. The project is thankful to them. The dissertation is thankful to them.

I would like to express my heartily devotion and foremost indebtedness to my parents, my teachers, my friends for their constant supports and encouragement throughout the research work.

Dharamgarh **Samikshya Pattjoshi**

Date: Roll No- 2215CSC011

ABSTRACT

Cancer refers to any one of a large number of diseases characterized by the development of abnormal cells that divide uncontrollably and have the ability to infiltrate and destroy normal body tissue. Cancer often has the ability to spread throughout your body. Among all the cancers lung cancer was the most common and fearful one. Most of the lung cancers are detected at the later stage this is causing increase in death rate. Cancer is detected using CT scan in most of the cases.

This project is based on detection of lung cancer in CT image using image segmentation by using Matlab algorithms . For this process a medical image will be considered, and then the image will be pre-processed for noise removal. Further segmentation is done to identify and separate desired tumour nodule using the watershed method and Matlab algorithms to detect the tumour cells.

TABLE OF CONTENTS PAGE No

CHAPTER 1

INTRODUCTION

Cancer is a noteworthy general health issue worldwide with mortality rates increasing day by day. Lung cancer among all other cancer types is the most common and deadly that occurs both in men and women. However, the early detection of lung cancer significantly improves survival rate. Detection of early stage is necessary for the crucial prognosis as early stage cancerous are very much similar to non cancerous lung nodules and can be treated lung cancer. Lung cancer, additionally known carcinoma is formation of malignant lung tumors (cancerous nodules) due to uncontrolled growth of cells in lung tissues.

Eating tobacco and smoking are the leading risk factors for causing cancerous lung nodules. The survival rate of lung cancer patients combining all stages is very less roughly 14% with a time span of about 5-6 years. The main problem with lung cancer is that most of these cancer cases are diagnosed in later stages of cancer making treatments more problematic and significantly reducing the survival chances. Hence detection of lung cancer in its earlier stages can increase the survival chances up to 60-70% by providing the patients necessary fast treatment and thus it curbs the mortality rate. Small cell lung cancer and non-small cell lung cancer are two main types of lung cancer classifications based on cell characteristics. The most commonly occurring is non-small cell lung cancer that makes up about 80-85% of all cases, whereas 15-20% of cancer cases are represented by small cell lung cancer. Lung cancer staging depends on spread of cancer in the lungs and tumor size. Lung cancer is mainly classified into 4 stages in order of seriousness: Stage I-Cancer is confined to the lung, Stage II and III-Cancer is confined within the chest and Stage IV-Lung cancer has spread from the chest to other parts of the body.

The most suitable method for the Investigation of lungs cancer disease is computer tomography (CT) imaging. However, CT scan investigation has a high rate of false positive findings, with carcinogenic effects of radiation. Low dose CT uses considerably lower power radiation contact than standard dose CT. The result shows that there is no significant difference in detection sensitivities between low dose and standard dose CT images. However, cancer related death were significantly reduced in the selected population that were exposed to low dose CT scans as compared to chest radiographs, which is depicted in the national Lung screening trail (NLST). The detection sensitivity of lungs nodules improves with sophisticated anatomical details and better image registration techniques.

Lung cancer diagnosis can be done by using various imaging modalities such Positron Emission Tomography (PET), Magnetic Resonance Imaging (MRI), Computed Tomography (CT) and Chest X-rays. CT scan images are mostly preferred over other modalities because they are more reliable, have better clarity and less distortion. Visual interpretation of database is a tedious procedure that is time consuming and highly dependent on given individual. This introduces high possibility of human errors and can lead to misclassification of cancer. Hence an automated system is of utmost importance to guide the radiologist in proper diagnosis of lung cancer. The methodology developed for this system includes dataset collection, pre-processing, lung segmentation, feature extraction and classification.

1.What is image?

An image is a visual representation of something, which can be in various forms such as a painting, photograph, or a graphical representation. In a broader sense, images can be analog or digital. Analog images are continuous representations, like photographs on film or hand-drawn pictures. It defined as a two-dimensional function f(x,y), Where x and y are spatial (plane) coordinates, and the amplitude of f at any pair of coordinates (x,y) is called the intensity or graylevel of the image.

Digital image -A digital image is an image composed of picture elements, also known as pixels, each with finite, discrete quantities of numeric representation for its intensity or gray level that is an output from its two-dimensional functions fed as input by its spatial coordinates denoted with x, y on the x-axis and y-axis. It is a type of image which is formed from pixels. Each pixel has some finite size and is represented by some finite intensity to show the image. The pixels are set of a properly arranged rectangular array. The image size can be determined by the pixel array dimensions. It has finite coordinates i.e. x-coordinate and y-coordinate. It is of two types i.e. raster type and vector type. The raster image is basically used for reference to a digital image.

There are many ways of encoding the information in an image.

1. Binary image -Each pixel is just black or white. Since there are only two possible values for every pixel (0, 1), we only need one bit per pixel.

2. Grayscale image- Each pixel may be a shade of gray, normally from 0 (black) to 255 (white). This range means that each pixel can be represented by exactly one byte or eight bits. Other grayscale ranges that are used, but generally they are a power of Grayscale Image.

3. Indexed image- An indexed image consists of an colour map matrix and array. In the array the pixel values are direct indices into a colour map. By convention, this documentation uses the variable name map to refer to the colour map and X to refer to the array.

4. True color or RGB image Binary image- True color or RGB image Each pixel features a particular color; that color is described by the quantity of red, green and blue in it. If each of those components features a range 0–255, this provides a complete of 2563 different possible colors. Such apicture may be a "stack" of three matrices; representing the red values, green values and blue values for every pixel. This means that for each pixel there correspond 3 values.

1.1 Image processing:

Image processing is a method to perform some operations on an image, in order to get an enhanced image or to extract some useful information from it. It is a type of signal processing in which input is an image and output may be image or characteristics/features associated with that image. Nowadays, image processing is among rapidly growing technologies. It forms core research area within engineering and computer science disciplines too. Image processing basically includes the following three steps: Importing the image via image acquisition tools; Analysing and manipulating the image; Output in which result can be altered image or report that is based on image analysis. There are two types of methods used for image processing namely, analogue and digital image processing. Analogue image processing can be used for the hard copies like printouts and photographs. Image analysts use various fundamentals of interpretation while using these visual techniques. Digital image processing techniques help in manipulation of the digital images by using computers. The three general phases that all types of data have to undergo while using digital technique are pre-processing, enhancement, and display, information extraction. Image processing is a method to convert an image into digital form and perform some operations on it, in order to get an enhanced image or to extract some useful information from it. It is a type of signal dispensation in which input is image, like video frame or photograph and output may be image or characteristics associated with that image. Usually Image Processing system includes treating images as two dimensional signals while applying already set signal processing methods to them. It is among rapidly growing technologies today, with its applications in various aspects of a business. Image Processing forms core research area within engineering and computer science disciplines too. Image processing basically includes the following three steps. · Importing the image with optical scanner or by digital photography. · Analyzing and manipulating the image which includes data compression and image enhancement and spotting patterns that are not to human eyes like satellite photographs. · Output is the last stage in which result can be altered image or report that is based on image analysis.

There are five main types of image processing:

Visualization – Find objects that are not visible in the image

Recognition – Distinguish or detect objects in the image

Sharpening and restoration – Create an enhanced image from the original image

Pattern recognition – Measure the various patterns around the objects in the image

Retrieval – Browse and search images from a large database of digital images that are similar to the original image

1.1.1 Application of image processing:

Medical Image Retrieval- Image processing has been extensively used in medical research and has enabled more efficient and accurate treatment plans. For example, it can be used for the early detection of breast cancer using a sophisticated nodule detection algorithm in breast scans. Since medical usage calls for highly trained image processors, these applications require significant implementation and evaluation before they can be accepted for use.

Traffic Sensing Technologies- In the case of traffic sensors, we use a video image processing system or VIPS. When capturing video, a VIPS has several detection zones which output an "on" signal whenever a vehicle enters the zone, and then output an "off" signal whenever the vehicle exits the detection zone. These detection zones can be set up for multiple lanes and can be used to sense the traffic in a particular station.

Image Reconstruction- Image processing can be used to recover and fill in the missing or corrupt parts of an image. This involves using image processing systems that have been trained extensively with existing photo datasets to create newer versions of old and damaged photos.

Face Detection- One of the most common applications of image processing that we use today is face detection. It follows deep learning algorithms where the machine is first trained with the specific features of human faces, such as the shape of the face, the distance between the eyes, etc. After teaching the machine these human face features, it will start to accept all objects in an image that resemble a human face. Face detection is a vital tool used in security, biometrics and even filters available on most social media apps these days.

1.2.Image processing techniques:

Some image processing techniques are as follows:-

1.2.1: Image Enhancement

Operations improve the qualities of an image like improving the brightness characteristics of the image and also improving image's contrast. Reducing it sharpen the details or reducing its noise content. Therefore this enhances the image and it reveals the same information in more understandable image. Where it does not add any information to it.

1.2.2: lmage Restoration

Same like enhancement, image restoration improves the qualities of image but all the operations that are mainly based on known, degradations or measured of the original image. Image restorations that restores the images. Problems such as geometric distortion, repetitive noise, camera motion and improper focus. Image restoration helps to correct images for known degradations.

1.2.3: Image Analysis

Numerical or graphical information is produced by the image analysis based on characteristics of the original image. They break into objects and then classify them. As they depend on the image statistics. Extraction and description of scene and image features, object classification and automated measurements are the common operations. In machine vision applications mainly image analyze is used.

1.2.4: Image Compression

lmage compression and decompression that reduce the data content which is necessary to describe the image. Many of the images contain lot of redundant information, so the compression it removes all the redundancies, the size is reduced because of the compression,so stored or transported efficiently. The compressed image that is decompressed when displayed. Exact data in the original image is preserved by the lossless compression, Lossy compression provide excellent compression but it does not represent the original image.

1.2.5: Image Synthesis:

Image synthesis operations create images from non-image data or from other images. Image synthesis operations generally create images that are either impractical to accumulate or physically impossible.

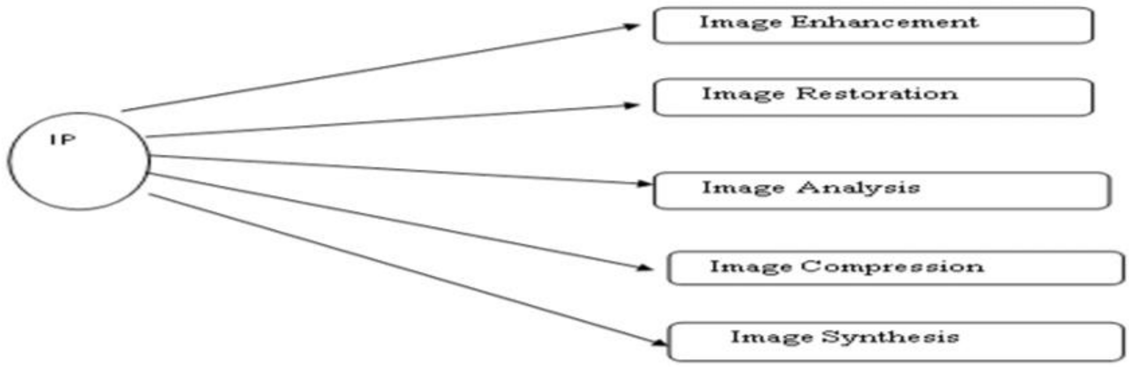


Figure 1.1 Image processing techniques

1.3 Image Segmentation:

Image segmentation is the process of breakdown of digital image into image segments. Image segmentation is the method helps to better analysing of digital image. Image segmentation method is used to locate the boundaries of image. Every pixel in the image is allocated to one of a number of these categories. Image segmentation is the process of assigning a label to every pixel in an image in such way that pixels with the label share certain characteristics. It is mainly used to locate objects and boundaries like lines and curves in the images. In semantic segmentation is basically used for more accurate view of an image. Segmentation is an important stage of the image recognition system, because it extracts the objects of our interest, for further processing such as description or recognition.

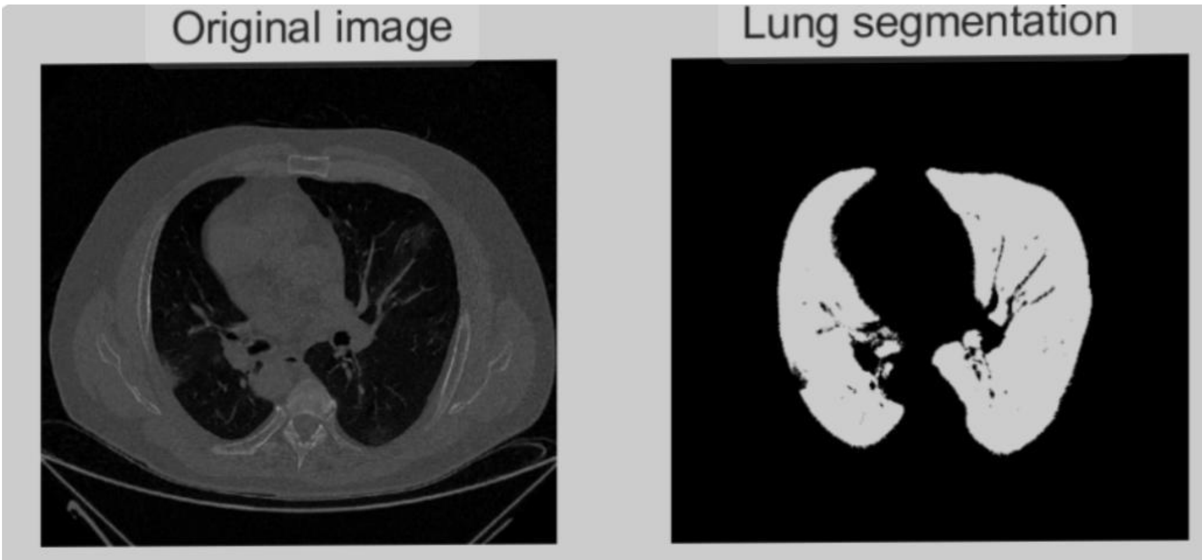


Figure 1.2 Segmented image of lungs

Segmentation of an image is in practice for the classification of image pixel. Cancer has long been a deadly illness. Even in today's age of technological advancements, cancer can be fatal if we don't identify it at an early stage. Detecting cancerous cell(s) as quickly as possible can potentially save millions of lives. The shape of the cancerous cells plays a vital role in determining the severity of the cancer. We might have put the pieces together – object detection will not be very useful here. We will only generate bounding boxes which will not us in identifying the shape of the cells. Image Segmentation techniques make a massive impact here. They help us approach this problem in a more granular manner and get more meaningful results.

1.3.1: Approaches of segmentation

Threshold Approach-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. In this we have a gray scale image given for a thresholding procedure it converts the rgb image into a binary image i.e black and white image which has only two shades i.e black and white which represent the level 0 and 1 only.the threshold value for this will be lies between 0 and 1 because it has only two levels., after achieving the threshold value; image will be segmented based on it.

Market watershed Approach- Marker watershed segmentation technique indicate the presence of objects or background at specific image locations Marker-Controlled Watershed

Segmentation Approach has two types: External associated with the background and Internal associated with the object of interest. Image segmentation using the Watershed transform works well if we can find 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 the light pixels are high and dark pixels are low.

1.4 What is Lungs Cancer ?

Lung cancer is a kind of cancer that starts as a growth of cells in the lungs. The lungs are two spongy organs in the chest that control breathing. Lung cancer is the leading cause of cancer deaths worldwide. People who smoke have the greatest risk of lung cancer. The risk of lung cancer increases with the length of time and number of cigarettes smoked. Quitting smoking, even after smoking for many years, significantly lowers the chances of developing lung cancer. Lung cancer also can happen in people who have never smoked.

1.5 Application of lungs cancer:

Some applications are as follows:

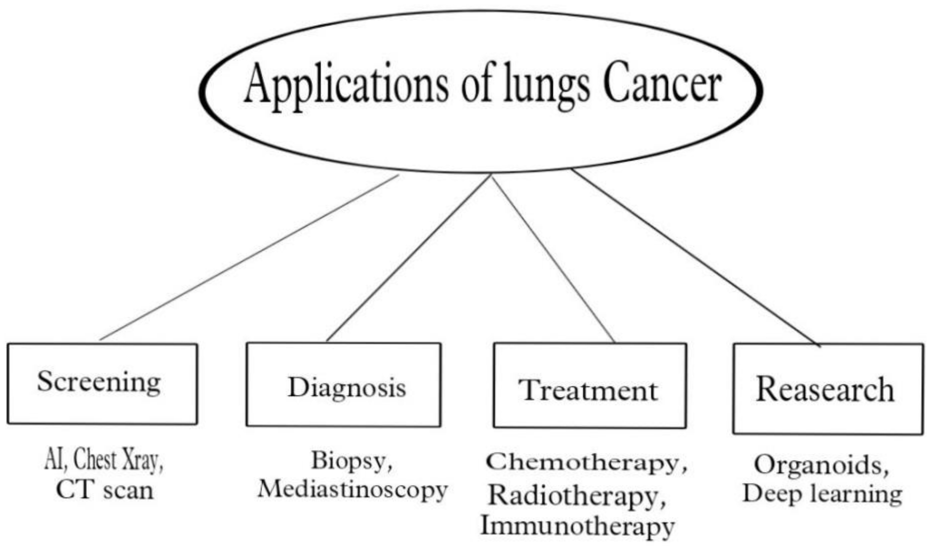


Figure 1.3 Block diagram of application of lungs cancer

1.5.1 Screening

Artificial intelligence – Artificial intelligence (AI) is good at handling a large volume of computational and repeated labor work and is suitable for assisting doctors in analyzing image-dominant diseases like lung cancer. Scientists have shown long-standing efforts to apply AI in lung cancer screening via CXR and chest CT since the 1960s. Several grand challenges were held to find the best AI model. Currently, the FDA have approved several AI programs in CXR and chest CT reading, which enables AI systems to take part in lung cancer detection.

Chest Xray – A chest X-ray is one the first imaging tests performed by a doctor to diagnose lung cancer. It exposes the chest to controlled beams of radiation. Tissue blocks radiation to different degrees based on how dense it is, which shows up on a black and white image. A normal, healthy lung will appear mostly dark gray. However, tumor tissue is denser and blocks more radiation, resulting in a lighter gray area relative to the rest of the lung.

CT scan -Computed tomography (CT) is another imaging technique. It uses X-rays to take images at multiple angles. This results in a more detailed, cross-sectional image of the body. Similar to a standard X-ray, areas with dense tissue (such as a tumor) appear on CT scans as a light gray to white mass on the image. A lung without disease will appear only as dark gray to black, with no evidence of a mass.

1.5.2 Diagnosis

Biopsy – These are the most common tool to obtain tissue for diagnosing lung cancer. Depending on where the nodule is located and the patient's physical condition, the doctor will do either a needle biopsy or a bronchoscopy. During a needle biopsy, the surgeon uses a syringe to remove tissue from the nodule. A CT scan guides the surgeon to the nodule. This type of test is usually done under sedation rather than under general anesthesia so that it can be done as an outpatient procedure without a hospital stay.

Mediastinoscopy – Mediastinoscopy is a surgical procedure that requires general anesthesia. An incision is made in the neck so that a lighted instrument called a mediastinoscope can be inserted to examine the area between the lungs known as the mediastinum. Biopsies of the mediastinal lymph nodes are taken for cancer staging.

1.5.3 Treatment

Chemotherapy -Lung cancer chemotherapy can be the main type of treatment. Sometimes lung cancer chemotherapy is used to shrink the tumor before surgery. This is called preoperative or neoadjuvant chemotherapy. Sometimes it is used to kill any cancers cells that remain after surgery. In some patients who are not eligible for surgery, but whose cancer has not spread outside the chest.

Radiotherapy – Lung cancer radiation therapy uses powerful, high-energy X-rays to kill cancer cells or keep them from growing. Radiation may come from outside the body (external) or from radioactive materials placed directly inside the lung cancer tumor (internal/implant). External radiation is used most often. The radiation is aimed at the lung cancer tumor and kills the cancer cells only in that area of the lungs.

Immunotherapy – Immunotherapy uses medications to stimulate your immune system to kill lung cancer cells. The drugs boost, direct or restore the body's natural defenses against cancer. Immunotherapy is also called biotherapy or biologic immunotherapy.

1.5.4. Research

Organoids- This technique offers a novel platform for disease modeling , biobanking and drug development. The expected benefit of organoids is for cancer patient as the subsequent precision medicine tecnology. Over the past few years, numerous basic and clinical studies have been conducted on lung cancer organoids, highlighting the significant contribution of this technique.

Deep learning- The task of deep learning application in lungs cancer include nodules detection on chest radiographs or lung cancer CT screening, potential candidate selection in lung cancer screening, malignancy prediction in indeterminate pulmonary nodules, lung cancer staging, treatment response prediction, prognostication, and prediction of genetic mutation in lung cancer.

1.6 Challenges of lungs cancer:

Some challenges are as follows:-

1.6 .1. Physical challenges

Shortness of breath: Tumors can block airways, put pressure on lungs, or cause inflammation in the respiratory system

Coughing up blood: Also known as hemoptysis, this can occur when airways are bleeding

Side effects from treatment: These can include fatigue and pain

1.6.2 Emotional challenges

Psychological distress: Lung cancer is associated with a high symptom burden and poorer prognosis, which can negatively impact quality of life

Stigma and social isolation: Patients may experience stigma and social isolation

1.6.3. Social challenges

Financial challenges: Cancer survivors may face financial challenges due to lack of money for treatment.

1.7 Objectives:

The objective of Our work is to take any input CT scan images and identify the stage of cancer based on the size of tumour. Lung cancer is a death taking disease in the present day. The identification of lung cancer at initial stages is of extreme importance because earlier detection is the only method to improve the survival rate. The presence of lung cancer can be diagnosed with the help of a CT image of the lungs. The manual detection by doctor may result in false reports. So, the objective in this project is need of computerized method for cancer detection. This project analyzes CT images using different image processing operations on them to get accurate detections.

In this project MATLAB has been used through every procedure made.

- To study the medical image processing under the concept of MATLAB.

-The main objective of this project is to find out early stage of lung cancer using image processing technique.-Enhancement of the CT image is done.-Image segmentation using thresholding approach.

1.8. Organisations of thesis:

Chapter 1- It is the introduction part that provides background on lung cancer, the importance of early detection, and an overview of medical image processing. It also outlines the objectives and scope of the project.

Chapter 2 – Literature review sections that reviews existing techniques for lung cancer detection, with a focus on image segmentation methods, especially Watershed segmentation. It also discusses challenges and gaps in current approaches.

Chapter 3 - Describes the proposed system in detail, including data acquisition, preprocessing steps, flow chart, Watershed segmentation, and post-processing techniques and explains how the algorithm was implemented.

Chapter 4- Explains how the algorithm was implemented, presents experimental results, and evaluates the performance using metrics such as accuracy, precision, and recall.

Chapter 5- Summarizes the key findings, highlights the strengths and limitations of the approach, and suggests directions for future research and improvements.

Chapter 2

LITERATURE REVIEW

1-''Detection of lung cancer in CT images using image processing" Nidhi S. Nadkarni & Prof. Sangam Borkar .

-Cancer is the one of the most serious and widespread disease that is responsible for large no of deaths every year. This paper presents a system for automatic detection of lung cancer in CT images was successfully developed using image processing technique as CT scan images may be an error prone task and can cause delay in lung cancer detection. Therefore, image processing technique are used widely in medical field for early stages detection of lung cancer. The methodology is carried out well in enhancing, segmenting and extracting features from CT images and implementation was done using MATLAB software. Median filtering technique was effective in eliminating impulse noise from the images without blurring the image. Mathematical morphological operations enable accurate segmentation of lung and tumor region. Three geometrical features i.e area, perimeter and eccentricity were extracted from segmented tumor region and classified it into normal and abnormal. Hence this proposed methodology helps in accurate and early stage detection of lung cancer.

2-''Image processing based detection of lung cancer on CT scan images" Bariqi Abdillah , Alhadi Bustamam and Devvi sawinda .

-Cancer is one of most dangerous disease that causes deaths which occupies the first position in the whole world. In this paper, we implement and analyze the image processing method for detection of lung cancer. Image processing technique are widely used in several medical problems for picture enhancement in the detection phase to support the early medical treatment. We implement and evaluate three image segmentation methods for analyzing lung cancer such as Region Growing, Marker Controlled Watershed, Marker Controlled Watershed with Masking. The results show that Marker Controlled Watershed with Masking gives them the best performance in term of segmentation result and running time. Therefore, We select Marker Controlled Watershed with Masking method in image segmentation stage. Furthermore, in the feature extraction stage, we use color attribute for the analysis of lung cancer using binarization. Finally, the binarization method was successfully determined condition of lung cancer from the CT scan image.

3-''Lung Cancer detection using image processing and machine learning healthcare" Wasdeo Rahane , Yamini Magar, Himali Dalvi, Anjali Kalane and Satyajeet Jondhale.

-Lungs cancer is a serious disease that described by unlimited growth of cells in tissues of the lungs. Since the cause of lung cancer stay obscure, prevention become impossible. In our proposed system we are describing the lungs cancer and its stages using different image processing and machine learning algorithms such as, grayscale conversion, noise reduction and binarization. All algorithms are used for the pre-processing stages, median filter and segmentation gives accurate results. From the extracted ROI some features are extracted i.e , Area, Perimeter, and Eccentricity. This characters are helpful for defining the lungs cancer at earlier stages. For grouping purpose Support Vector Machine (SVM) classifier classifies the positive and negative samples of lungs cancer images in this system.

4-''Lung cancer detection and classification based on image processing and statictical learning"Md Rashidul Hasan, Muntasir Al Kabir.

-Lung cancer is one of the death threatening diseases among human beings. Early and accurated detection of lung cancer can increased the survival rate. The proposed lung cancer detection the tumor within the lung. The CT image is pre-processed and the pre-processed image is then subjected to segmentation by using Marker Controlled Watershed segmentation. Segmented image is used for feature extraction. With the extracted feature the tumor is detected within the lung. Both supervised and unsupervised classifier is used for the identifying of the cancer. The accuracy rate of the proposed system is 72.2 percent by using support vector machine. Thus this system helps the radiologist to identify the stage of the tumor and increase the accuracy.

5-''Enhanching lung cancer detection from lung CT scan using image processing and deep neural networks" Pavan Kumar Pagadala, Sree Lakshmi Pinapatruni, Chanda Raj Kumar, Srinivas Katakam, Lalitha Surya Kumari Peri, Dasari Anantha Reddy.

-The proposed methodology employs a variety of image processing and analysis techniques to achieve accurate detection result. To begin the acquired lung cancer images and pre-processed with a multidimentional filter and histogram equalization in order to improve their quality and subsequent analysis. Histogram equalization optimizes an images dynamic range, enhancing visibility of structures and abnormalities. This technique proves invaluable in medical imaging, revealing subtle features for accurate anomaly detection. For accurate lung cancer detection, a deep neural network is trained all tested and achieved 99.1251% specificity, 99.1121% sensitivity, and 99.269% accuracy. The use of MATLAB as a computing tool ensured efficient implementation. Finally this study demonstrate the utility of image processing and deep learning technique in the detection of lung cancer.

6-''An effective method for lung cancer diagnosis from CT scan using deep learning-based support vector network"Imran Shafi, Sadia Din, Asim Khan, Isabel De La Toree Diez, Ramon Del Jesus Pali Casanova, Kilian Tutusaus Pifarre and Imran Ashraf.

-The study involves applying various supervised learning technique for cancer detection on the LUNA16 dataset. It envisions that the proposed model SVM + CNN, deep learning enabled SVM outperforms all other methodologies. Understanding the dataset and extracting features from it prior to inputting data into the machine learning algorithm is necessary. Ultimately, this feature selection impacts the final accuracy of the machine learning models. Contrarily, a deep learning algorithm also has a final classification layer. The hybrid deep learning enabled SVM uses the advantages of both the deep and machine learning techniques. In this approach CNN is used for features selection and the machine learning model SVM is used for classification. The advantages of both type of learning are combined in the proposed scheme and hence a definitive outcome is achieved that is more accurate and fast compared to other existing approaches.

7-''Lung cancer detection system using image processing and machine learning techniques" Vikul J. Pawar, Kailash D. Kharat, Suraj R, Pardeshi, Prashant D. Pathak.

-In recent time the lungs cancer diseases are increases widely, the human body made up of with diverse fundamental organs, one of those is lungs. The analysis of the CT scan we observe that the effective use of different techniques such as thresholding, classification techniques such as ANN, SVM, and etc. in the several research it is found that lung cancer is detected in the last stage of it which may cause the risk of death. There is a need of computer aided diagnosis system (CAD) to work on robustness of preprocessing techniques, segmentation techniques, noise removal method, and accurate diagnosing classifier to achieve the objective of medical imaging field by applying image processing techniques. The proposed system overcome the challenges of preceding methods which are used in detection system and exploit the robust noise filtering method using autoencoder system, segmentation and classification techniques such as OTSU algorithm for segmentation and classification techniques are decision tree and CNN. This paper is bringing the emerging image processing and machine learning for implementing purpose.

8-''Lung cancer detection using image processing techniques" Shradha fule.

-An image processing technique is built to detect diseases at early stage of cancer and the time factor is very important to discover the abnormal tissue in target x-ray images. In this project we access cancer image into MATLAB collected from different hospitals where present work is going on and this available image are colour image so we have access that image into MATLAB and followed conversion. Image quality and accuracy is the core factor of this research, image quality assessment as well as improvement are depending on the enhancement stage where low pre processing techniques is used based on Gabor filter within Gaussian rules. The segmentation and enhancement procedure is used to obtain the feature extraction of normal and abnormal image. In this research, the main detected feature for accurate images comparison are pixel percentage and mask labelling.

9-''IOT based lung cancer detector using image processing" Suhasini Jotirning Utture, Apurva Malgonda Bhokare, Aarti Sanjay Patil, Rutuja Rahul Chothe and Mr. A.J.Chinchawade.

-The proposed cancer detector detect the lung cancer from CT scan images. The overall 5 years of survival rate for lungs cancer patients increases from 14 to 49 % if the disease is detected in time. Although computer tomography (CT) can be more efficient than x ray. However, problem seems to merge due to time constraint in detecting the present of lung cancer regarding the several diagnosis method used. Hence, a lung cancer detection system using image processing is used to classify the present of lung cancer in an CT images. In this study, MATLAB has been used through every procedure made. In order to improve the classification performance of the large scale of data, Generic algorithm is proposed. The system detect the different stages of lung cancer which will assist the doctor to detect lung cancer accurately and fast from large amount of data.

10-''Lungs cancer detection using image processing and machine learning healthcare" Wasudeo Rahane, Himali Dalvi, Yamini Magar, Anjali Kalane, and Satyajeet Jondhale.

-Lung cancer is one of the significant reasons for death among India. Many diagnosis and detection of lung cancer has been done using various data analysis and classification techniques. Since the cause of lung cancer stay obscure, prevention became impossible, thus early detection of tumor in lung is the only way to cure lung cancer. Hence, lung cancer detection system using image processing and machine learning is used to classify the presence of lung cancer in a CT images and blood samples. In spite of CT scan report are more effective than mammography, therefore patient CT scan images are categorized in normal and abnormal. The abnormal images are subjected to segmentation to focus on tumor portion. Classification done on features extracted from the images. The efficient method to detect the lung cancer and its stages successfully and also aim to have more accurate results by using SVM and image processing techniques.

Table 2.1 Comparison of the Related work.

Sl.no.	Reference	Method	Advantages	Disadvantages
1	Nidhi S. Nadkarni and prof. Sangam Borkar.	Morphological Method	Early stage detection of lung cancer in CT image.	May have high false positive rates.
2	Bariqi Abdillah , Alhadi Bustamam and Devvi sawinda .	Marker controlled watershed with masking.	Determined the condition of lungs cancer from CT scan.	Need for accurate marker replacement to achieve good result.
3	Wasdeo Rahane , Yamini Magar, Himali Dalvi, Anjali Kalane and Satyajeet Jondhale.	Median Filter	Classifies the positive and negative sample of lung cancer images in this system.	Blurring of edges and fine details.
4	Md Rashidul Hasan, Muntasir Al Kabir.	Marker controlled watershed	Early and accurate detection of tumour within the lungs.	A robust method for delineating objects in complex image datasets.
5	Pavan Kumar Pagadala, Sree Lakshmi Pinapatruni, Chanda Raj Kumar, Srinivas Katakam, Lalitha Surya Kumari Peri, Dasari Anantha Reddy.	Histogram equalization	It employs a variety of image processing and analysis technique to achieve accurate result.	Creating unnatural looking result in certain areas.
6	Imran Shafi, Sadia Din, Asim Khan, Isabel De La Toree Diez, Ramon Del	CNN and SVM	Outcome is more accurate and fast compared to	Required large amount of labelled training data set.

	Jesus Pali Casanova, Kilian Tutusaus Pifarre and Imran Ashraf.		other approaches.	
7	Vikul J. Pawar, Kailash D. Kharat, Suraj R, Pardeshi, Prashant D. Pathak.	Thresholding	The effective use of different techniques and early detection.	It takes time but gives accurate result.
8	Shradha fule.	Marker controlled watershed	The main detection feature for accurate images comparison and pixel and mask labelling.	High sensitivity to noise.
9	Suhasini Jotirning Utture, Apurva Malgonda Bhokare, Aarti Sanjay Patil, Rutuja Rahul Chothe and Mr. A.J.Chinchawade.	Generic algorithm	Detect the different stages of lung cancer which will assist the doctor within less time.	Slow convergence and high computational costs.
10	Wasudeo Rahane, Himali Dalvi, Yamini Magar, Anjali Kalane, and Satyajeet Jondhale.	Grey scale and binarization	Helps in reducing the growth of abnormal cell or spreading to other part of the body.	Loss of detail due to reducing colour information to said off grey or black and white.

I read thoroughly total 10 Research paper of detection of lungs cancer in CT images using various image segmentation techniques related papers and from all this I have concluded that watershed method is the best one for detecting lungs cancer because it is excellent at separating and identifying different areas in medical images, like CT scan which is crucial for detecting tumours. This method helps in clearly distinguish the boundaries of tumors from surrounding tissues, making it easier to spot cancel and ensures accurate results compared to other methods, the watershed method provides a more precise segmentation, which is important for better diagnosis and treatment planning.

CHAPTER 3

Detection of lung cancer

Lung cancer varies differently from person to person, depending on the size of the tumour and the stage it is in. Stage I is considered as when the cancer is restricted to the lung. Stage II is when the cancer is limited to the chest. Stage III is when the tumour grows larger and appears in the CT scan. Stage IV is confined to spreading cancer cells to other parts of the body and growth of tumours in other parts as well. Analysing CT scan image of lungs and predicting stage of cancer based on tumour requires a high level of skill and concentration, and is possible only by the expert doctors or radiologists. CT stands for computerized tomography where passing of X-rays inside the human body takes place. There are many other image processing methods and techniques such as MRI, Ultrasound, DEXA, X-ray and PET, but CT scan is best recommended because CT scan is best accurate and noise is very less as compared to others. As medical images may have noise, the CT image is passed to the second step, i.e., pre-processing where median filter is used in this paper for elimination of noise.

Then conversion of image into binary followed by segmentation is done where CT image is partitioned into some sets of pixels and obtaining of tumour (white) region pixels i.e., extraction of required and interested region (which are the white regional pixels in different places present in the lungs, eliminating the other part of the lung which is not affected takes place). Various simple segmentation techniques such as watershed algorithm and different morphological operations like dilation, erosion, opening to apply big mask were used in this paper. Last step is feature extraction to decide whether those white regional pixels were still initial cancer cells or tumour is decided here and area and diameter of those white regional pixels are extracted and based on the size value of the white grouped pixels, exact tumours are separated out and based on radius value of tumour, classification of lung cancer into various stages is decided in this project. In feature extraction fore ground -background and SIFT (scale-invariant feature detection) techniques are used to extract required features of tumour module. Based on size of tumour module stage of cancer is detected. By knowing size of tumour nodule it becomes easier for the radiologists/doctors to easily find the stage of cancer instantly. Proper medication like radiotherapy or surgery can be done if the tumour's size is large. Early detection of stage and diagnosis are the only prevention steps of lung cancer.

3. Introduction to image pre-processing and segmentation :

Image pre-processing is a set of techniques employed to enhance the quality and extract relevant information from digital images before they are further analysed and processed by computer vision or machine learning algorithms. It involves a series of operations that aim to correct, filter, normalize, or enhance the images in order to improve their suitability for subsequent analysis. These pre-processing steps include various operations such as noise reduction, which reduces unwanted artifacts or random variations in the image; image resizing, which adjusts the image dimensions to meet specific requirements; contrast enhancement, which adjusts the image to improve visibility of objects or features; and image normalization, which ensures that the pixel values are scaled to a consistent range. Other operations may involve color space conversion, edge detection, morphological operations, or image segmentation, depending on the specific requirements of the application. By applying image pre-processing techniques, the quality and reliability of subsequent computer vision or machine learning tasks, such as object detection, classification, or recognition, can be significantly improved. Image pre-processing may have beneficial effects on the excellence of feature extraction and the outcomes of image analysis. Image pre-processing is similar to the scientific standardization of a data set, which is a general step in many feature descriptor techniques. Image pre-processing is used to correct the degradation of an image. In that case, some prior data or information is important such as information about the nature of the degradation, information about the Feature of images.

Image segmentation is a technique that partitions a digital image into discrete groups of pixels—image segments—to inform object detection and related tasks. By parsing an image's complex visual data into specifically shaped segments, image segmentation enables faster, more advanced. Image segmentation is a commonly used technique in digital image processing and analysis to partition an image into multiple parts or regions, often based on the characteristics of the pixels in the image. Image segmentation could involve separating foreground from background or clustering regions of pixels based on similarities in color or shape. For example, a common application of image segmentation in medical imaging is to detect and label pixels in an image or voxels of a 3D volume that represent a tumor in a patient's brain or other organs.

3.1 Pre-processing:

Pre-processing is a typical name for procedures applied to both input and output intensity images. These images are indistinguishable from the original data taken by the sensors.

Basically, image pre processing is a method to transform raw image data into a clean image data, as most of the raw image data contain noise and contain some missing values or incomplete values, inconsistent values, and false values. Missing information means lacking of certain attributes of interest or lacking of attribute values. Inconsistent information means there are some discrepancies in the image. False value means error in the image value. The purpose of pre-processing is an enhancement of the image data to reduce reluctant falsifications or to improve some image features vital for additional processing. Some will contend that image pre-processing is not a smart idea, as it alters or modifies the true nature of the raw data. Nevertheless, smart application of image pre-processing can offer benefits and take care of issues that finally produce improved global and local feature detection.

3.1.1 Techniques used for pre-processing:

The choice of techniques depends on the nature of the image and the application. Here are some techniques used in image pre-processing:-

Noise Reduction: Noise in an image can be caused by various factors such as low light, sensor noise, and compression artifacts. Noise reduction techniques aim to remove noise from the image while preserving its essential features. Some common noise reduction techniques include Gaussian smoothing, median filtering, and wavelet denoising.

Contrast Enhancement: Contrast enhancement techniques aim to increase the contrast of an image, making it easier to distinguish between different image features. These techniques can be helpful in applications such as medical imaging and surveillance. Some standard contrast enhancement techniques include histogram equalization, adaptive histogram equalization, and contrast stretching.

Image Resizing: Image resizing techniques are used to adjust the size of an image. Resizing can be done to make an image smaller or larger or to change its aspect ratio. Some typical image resizing techniques include nearest neighbour interpolation, bilinear interpolation, and bicubic interpolation.

Color Correction: Color correction techniques are used to adjust the color balance of an image. Color correction is important in applications such as photography, where the color accuracy of an image is critical. Some common color correction techniques include grey world assumption, white balance, and color transfer.

Segmentation: Segmentation techniques are used to divide an image into regions based on its content. Segmentation can be helpful in applications such as medical imaging, where specific structures or organs must be isolated from the image. Some standard segmentation techniques include thresholding, edge detection, and region growing.

Feature Extraction: Feature extraction techniques are used to identify and extract relevant features from an image. These features can be used in object recognition and image classification applications. Some standard feature extraction techniques include edge detection, corner detection, and texture analysis.

3.2 Segmentation:

Segmentation is the process of breakdown of digital image into segments. Image segmentation is the method helps to better analysing of digital image. Image segmentation method is used to locate the boundaries of image. Every pixel in the image is allocated to one of a number of these categories. Image segmentation is the process of assigning a label to every pixel in an image in such way that pixels with the label share certain characteristics. It is mainly used to locate objects and boundaries like lines and curves in the images. In semantic segmentation is basically used for more accurate view of an image. Segmentation is an important stage of the image recognition system, because it extracts the objects of our interest, for further processing such as description or recognition. Segmentation of an image is in practice for the classification of image pixel.

3.2.1 How image segmentation works?

The first step in image segmentation involves converting an image into a collection of regions of pixels that are represented by a mask or a labeled image. By dividing an image into segments, you can select and then process only the important segments of the image instead of processing the entire image.

A common technique is to look for abrupt discontinuities in pixel values, which typically indicate edges that define a region Another common approach is to detect similarities in the regions of an image. Some techniques that follow this approach are region growing, clustering, and thresholding.

3.2.2 Image segmentation techniques:

1-Thresholding: Thresholding is simplest method of segmenting images .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. In this we have a gray scale image given for a thresholding procedure it converts the rgb image into a binary image i.e black and white image which has only two shades i.e black and white which represent the level 0 and 1 only threshold value for this will be lies between 0 and 1 because it has only two levels., after achieving the threshold value; image will be segmented based on it.

2-Region based segmentation: Region-based segmentation is a technique in image processing and computer vision that identifies regions of an image that are similar according to a set of features, and groups them together.This grouping is based on predefined criteria like intensity, color, texture, or pixel boundaries. The primary goal is to transform an image into meaningful data that helps identify and distinguish objects or areas of interest, and simplifies further analysis and processing. It is particularly useful in medical imaging, robotics, video surveillance, and any use case that requires an understanding/recognition of objects or scenes.

3-Edge based segmentation: Edge-Based Segmentation is a technique in image processing used to identify and delineate the boundaries within an image. It focuses on detecting edges, which are areas in an image where there is a sharp contrast or change in intensity, such as where two different objects meet. Edge-based segmentation techniques work by identifying areas in an image where there is a rapid change in intensity or color. These changes often mark the edges of objects or regions within the image.

4-Watershed: Watershed is a ridge approach, also a region-based method, which follows the concept of topological interpretation. We consider the analogy of geographic landscape with ridges and valleys for various components of an image. The slope and elevation of the said topography are distinctly quantified by the gray values of the respective pixels – called the gradient magnitude. Based on this 3D representation which is usually followed for Earth landscapes, the watershed transform decomposes an image into regions that are called “catchment basins”. For each local minimum, a catchment basin comprises all pixels whose path of steepest descent of gray values terminates at this minimum.

5-Clustering: The clustering based techniques segment the Image into clusters or disjoint groups of pixels with similar characteristics. By the virtue of basic Data Clustering properties, the data elements get split into clusters such that elements in same cluster are more similar to each other as compared to other clusters. Some of the more efficient clustering algorithms such as k-means, improved k means, fuzzy c-mean (FCM) and improved fuzzy c mean algorithm (IFCM) are being widely used in the clustering based approaches proposed. K means clustering is a chosen and popular method because of its simplicity and computational efficiency. The Improved K-means algorithm can minimize the number of iterations usually involved in a k-means algorithm. FCM algorithm allows data points, (pixels in our case) to belong to multiple classes with varying degrees of membership. The slower processing time of an FCM is overcome by improved FCM.

3.2.3 Segmentation using Watershed:

A watershed is a transformation defined on a greyscale image. The name refers metaphorically to a geological watershed, or drainage divide, which separates adjacent drainage basins. The watershed transformation treats the image it operates upon like a topographic map with the brightness of each point representing its height, and finds the lines that run along the tops of ridges. There are different technical definitions of a watershed. In graphs, watershed lines may be defined on the nodes, on the edges, or hybrid lines on both nodes and edges. Watersheds may also be defined in the continuous domain. There are also many different algorithms to compute watersheds. Watershed algorithms are used in image processing primarily for object segmentation purposes, that is, for separating different objects in an image. This allows for counting the objects or for further analysis of the separated objects. The algorithm based on the concept of “immersion”. Each local minima of a gray-scale image I which can be regarded as a surface has a hole and the surface is immersed out into water. Then, starting from the minima of lowest intensity value, the water will progressively fill up different catchment basins of image (surface) I. Conceptually the algorithm then builds a dam to avoid a situation that the water coming from two or more different local minima would be merged. At the end of this immersion process, each local minimum is totally enclosed by dams corresponding to watersheds of image (surface).

The watershed transform has been widely used in many fields of image processing, including medical image segmentation, due to the number of advantages that it possesses: it is a simple intuitive method, it is fast and can be parallelized and an almost linear speedup was reported for a number of processors up to 64 and it produces a complete division of the image in separated regions even if the contrast is poor, thus avoiding the need for any kind of contour joining. It is appropriate to use this method to segment the high-resolution remote sensing image.

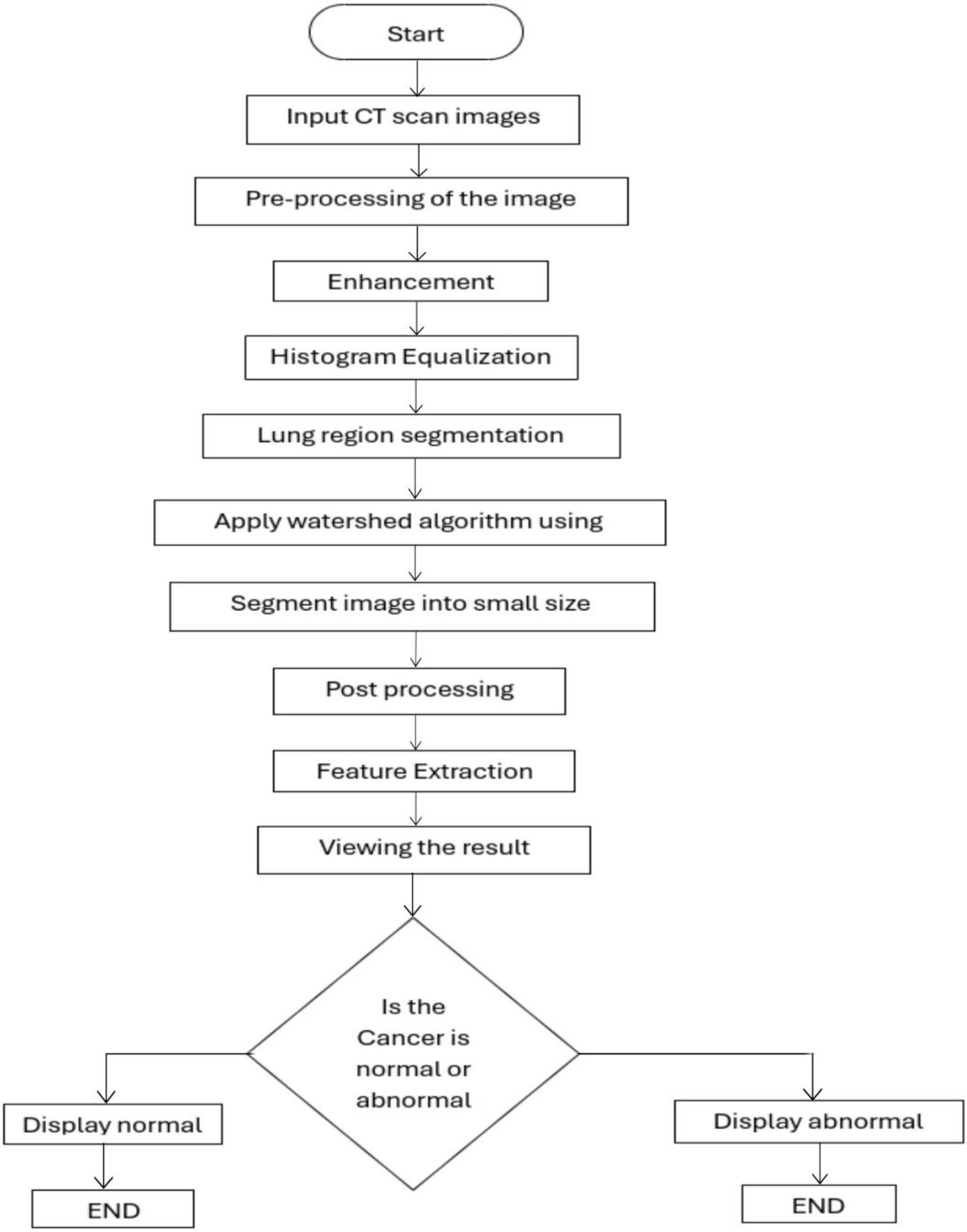
Watershed implementation methods are as follows :-

1-Distance Transform Approach: A classic way of separating touching objects in binary images makes use of the distance transform and the watershed method. The idea is to create a border as far as possible from the center of the overlapping objects. This strategy works very well on rounded objects and it is called Distance Transform Watershed. It consists on calculating the distance transform of the binary image, inverting it (so the darkest parts of the image are the centers of the objects) and then applying watershed on it using the original image as mask (see figure below). In our implementation, we include an option to use watershed with extended minima so the user can control the number of object splits and merges.

2- Gradient method: The gradient magnitude is used to pre process a gray-scale image prior to using the watershed transform for segmentation. The gradient magnitude image has high pixel values along object edges and low pixel values everywhere else. Watershed transform would result in watershed ridge lines along object edges. There is a problem of over segmentation in this method. The topological gradient provides a global analysis of the image then the almost unwanted contours due to the noise added to a given image can be significantly reduced by our approach. The experimental results show that the over segmentation problem, which usually appears with the watershed technique, can be attenuated, and the segmentation results can be performed using the topological gradient approach. Another advantage of this method is that it splits the segmentation process into two separate steps: first we detect the main edges of the image processed, and then we compute the watershed of the gradient detected.

3-Marker watershed: Marker watershed segmentation technique indicate the presence of objects or background at specific image locations Marker-Controlled Watershed Segmentation Approach has two types: External associated with the background and Internal associated with the object of interest. Image segmentation using the Watershed transform works well if we can find 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 the light pixels are high and dark pixels are low. According to the experimental subjective segmentation stage assessment Marker-Controlled during the Watershed Segmentation approach has more accuracy and quality than Thresholding approach.

3.3 Flow chart for proposed method:



3.4 Algorithm for the present work:

Initialization

Step 1: Input a CT scan images.

Step 2: pre-processing of the image.

Pre-processing is a crucial step in image analysis, particularly in tasks such as object detection, classification, segmentation, and recognition. The goal of image pre-processing is to improve the quality and relevance of the image data for further analysis, making it easier for algorithms to extract meaningful information. It encompasses a range of techniques, each tailored to address specific challenges in image quality and structures . Some common techniques of pre-processing in image are below:-

1. Grayscale Conversion

Purpose: Color information in an image may not always be necessary for certain image analysis tasks, especially when the focus is on intensity-based features (such as edges and shapes). Converting an image to grayscale reduces the complexity of the data by eliminating the color channels (Red, Green, and Blue), retaining only the intensity or luminance.

How it works: Each pixel's RGB values are converted to a single intensity value, typically using a weighted sum (e.g., 0.2989R + 0.5870G + 0.1140*B for standard luminance).

2. Noise Removal (Denoising)

Purpose: Images may contain noise, which is unwanted random variations in pixel values that obscure important features. Denoising techniques aim to reduce this noise without losing key image information.

Techniques:

Median Filtering: In this technique, each pixel is replaced with the median value of its neighboring pixels. It is particularly effective at removing "salt-and-pepper" noise.

Gaussian Blurring: Involves convolving the image with a Gaussian filter, which helps smooth the image and reduce noise by averaging neighboring pixels.

Bilateral Filtering: A more sophisticated technique that preserves edges while reducing noise by considering both spatial proximity and intensity differences.

3. Image Resizing

Purpose: Standardizing the size of images is necessary, especially for machine learning tasks where a fixed input size is required. It helps improve the efficiency of algorithms and ensures consistency across datasets.

How it works: The image is scaled by either shrinking or enlarging it. Resizing techniques include nearest-neighbor interpolation, bilinear interpolation, and bicubic interpolation, each balancing speed and quality.

4. Histogram Equalization

Purpose: This technique improves the contrast of an image by redistributing the intensity levels across the entire range. It helps make features in the image more distinguishable, particularly in images with low contrast.

How it works: Histogram equalization uses the cumulative distribution function (CDF) of pixel intensities to adjust pixel values so that the output image has a uniform distribution of intensities.

5. Edge Detection

Purpose: Edge detection identifies boundaries within an image, where there is a significant change in pixel intensity. This is useful for feature extraction, object recognition, and shape analysis.

Techniques:

Sobel Operator: Uses convolution with Sobel kernels to detect edges by approximating the gradient of intensity at each pixel.

Canny Edge Detector: A multi-step algorithm that detects edges by first smoothing the image, finding intensity gradients, applying non-maximum suppression, and using edge tracing by hysteresis.

Prewitt Operator: Similar to Sobel, it highlights edges by computing approximations of the gradient of pixel intensities.

6. Thresholding

Purpose: Thresholding converts a grayscale image into a binary image, simplifying the analysis of objects or regions of interest. It is useful for separating the foreground from the background.

Techniques:

Global Thresholding: A fixed threshold is applied to the entire image, where pixels above the threshold are set to one value (typically white), and pixels below are set to another (black).

Adaptive Thresholding: The threshold is computed based on local regions of the image, which is useful when lighting conditions vary across the image.

7. Morphological Operations

Purpose: Morphological operations are typically applied to binary images to manipulate their structure. These operations help remove small noise, fill holes, or enhance specific structures.

Common Operations:

Erosion: Reduces the size of foreground objects by removing boundary pixels. It is often used to remove small noise.

Dilation: Expands the size of foreground objects by adding pixels around them. It is useful for closing gaps or filling holes in objects.

Opening: Erosion followed by dilation. It removes small objects or noise.

Closing: Dilation followed by erosion. It fills small holes or gaps within objects.

8. Image Rotation and Flipping

Purpose: Geometric transformations such as rotation and flipping are useful for data augmentation, normalizing image orientation, or correcting image alignment.

How it works:

Rotation: The image is rotated by a specified angle, either clockwise or counterclockwise.

Flipping: The image is reflected horizontally or vertically.

9. Contrast Adjustment

Purpose: Adjusting the contrast of an image improves the visibility of details in areas that may be too dark or too light.

How it works: Contrast stretching or linear contrast adjustment is typically used to remap the intensity values of an image to utilize the full dynamic range.

10. Binarization

Purpose: Binarization simplifies the image into two distinct categories: foreground and background, usually by turning every pixel into either black or white.

How it works: A threshold value is applied to each pixel. If the pixel's intensity is above the threshold, it is set to white; otherwise, it is set to black.

Step 3: Enhancement of the image.

Image enhacement is the process of sharpening or smoothen the image. It improves the image quality and remove the noise from the image.It provides the better input for the digital image processing. Image enhancement belongs to image preprocessing methods. Objective of image enhancement – process the image (e.g. contrast improvement, image sharpening ,…) so that it is better suited for further processing or analysis Image enhancement techniniques classified into two main parts:-

Spatial domain methods- which directly operates on a pixel of a digital image.

Frequency domain method-which operates on a fourior transform of a image. It is the low level processing technique. No mathematical criteria are used for optimizing processing results. This can be achieved through various techniques, depending on the goal—whether to highlight specific details, reduce noise, improve contrast, or sharpen edges.

Some common image enhancement techniques include:

1. Contrast Adjustment

This involves changing the difference between light and dark regions in an image, making details more visible. This can be done using histograms, where the contrast can be adjusted by stretching or compressing the intensity levels.

2. Histogram Equalization

This method adjusts the image's contrast by redistributing pixel intensity values. The goal is to make the intensity distribution as uniform as possible, enhancing the visibility of features across different brightness levels.

3. Noise Reduction (Smoothing)

Noise in images can obscure important details. Techniques like Gaussian smoothing, median filtering, or bilateral filtering are used to reduce noise while preserving edges and important features.

4. Sharpening

This technique enhances edges in an image, making it appear clearer. It is typically done by using convolution kernels like the Laplacian or Sobel operators, which detects the edges and make them more prominent.

5. Edge Enhancement

Edges represent important boundaries in an image. Techniques like Sobel or Canny edge detection are used to highlight the edges, which can be useful for applications like object recognition.

6. Gamma Correction

This process adjusts the brightness of an image. It's especially useful for correcting images that are either too dark or too bright, ensuring that all features are visible.

7. Color Enhancement

This involves adjusting the colors in an image for more vibrant or accurate color representation. Techniques like saturation enhancement, color balance adjustment, or white balance correction are used to improve the overall color quality.

8. Morphological Operations

These operations (like dilation, erosion, opening, and closing) are used for shape processing, especially in binary or grayscale images. They can enhance image structures, remove small artifacts, or smooth boundaries.

9. High-Pass Filtering

High-pass filters allow high-frequency components (like edges) to pass while filtering out lower frequencies (like smooth areas), resulting in sharper images with enhanced details.

10. Dynamic Range Compression

This technique compresses the dynamic range of an image to make the bright and dark areas more visible, which is useful in images with extreme lighting conditions.

Step 4: Lung segmentation is done as the initial segmentation.

Lungs cancer be Segmented in various Segmentation Techniques such as

1.Thresholding Based Segmentation: - Use thresholding to segment different tissue types (lung tissue vs. tumor tissue) based on intensity or contrast differences. For instance, use higher threshold values for tumors or abnormal regions.

2. Edge Detection: - Use edge detection algorithms like Canny Edge Detection or Sobel Operators to detect boundaries between regions (tumor vs. normal tissue).

3. Region Growing Segmentation: - Start from seed points located near potential tumors and grow regions based on intensity or color homogeneity. This technique can help detect irregular shapes or complex.

4. Watershed Segmentation: - Watershed segmentation can help in distinguishing connected but different objects (e.g., tumors, vessels). It can be useful in identifying cancerous regions by detecting changes in intensity gradients.

Step 5: Segmented the image using Watershed segmentation techniques.

To segment an image using the Watershed segmentation technique, you can follow these steps.

1.Read the image:

Load the image into your environment using image-processing libraries.

2.Convert the image to grayscale:

Watershed works better on grayscale images, so convert the image to a grayscale format.

3.Apply Gaussian Blur:

To reduce noise and make the segmentation process more efficient, apply a Gaussian blur to the images.

4.Finding Sure Background and Sure Foreground:

Watershed segmentation requires identification of sure background and sure foreground. The sure background is usually obtained by dilating the image, and the sure foreground is obtained by applying distance transform and thresholding.

5.Apply Watershed:

Use the function that to perform the segmentation in watershed approach.

6.Visualization:

Once the segmentation is complete, visualize the segmented image and boundaries.

Step 6: Features Extraction.

After segmentation, the next step is extracting meaningful features from the segmented regions to distinguish normal tissues from abnormal areas.

1.Shape Features:

Area, perimeter, eccentricity, and circularity of detected regions can help in classifying potential tumor regions.

2.Texture Features:

Use texture featureslike, Gabor filters, or Local Binary Patterns (LBP) to extract textural information from the segmented regions, which can indicate malignancy.

3.Edge-based Features:

Analyze the edges of segmented tumor regions, using algorithms like Laplacian of Gaussian (LoG) or Canny edge detection, to capture details about the tumor's boundary.

4.Statistical Features:

Calculate mean, standard deviation, skewness, and kurtosis of pixel intensities within the segmented.

Step 7: Results Visualization

1.Overlay Segmented Tumor on Original Image:

Visualize the detected lung cancer tumor(s) by overlaying the segmented tumor regions over the original CT image.

2.Output a Segmentation Mask:

Produce a binary mask that highlights the tumor regions for easy analysis or further clinical review.

Termination

CHAPTER 4

RESULT AND DISCUSSION

Implementing the following steps the cancer cells are identified. This figure shows cancerous lungs CT scan image which was collected from Cancer Imaging Archive (CIA) database. The various experiments/processes proposed in the above sections for the lung cancer detection were implemented using MATLAB which is necessary and suitable for better classification of the stage of cancer and accuracy in the process of prediction using segmentation

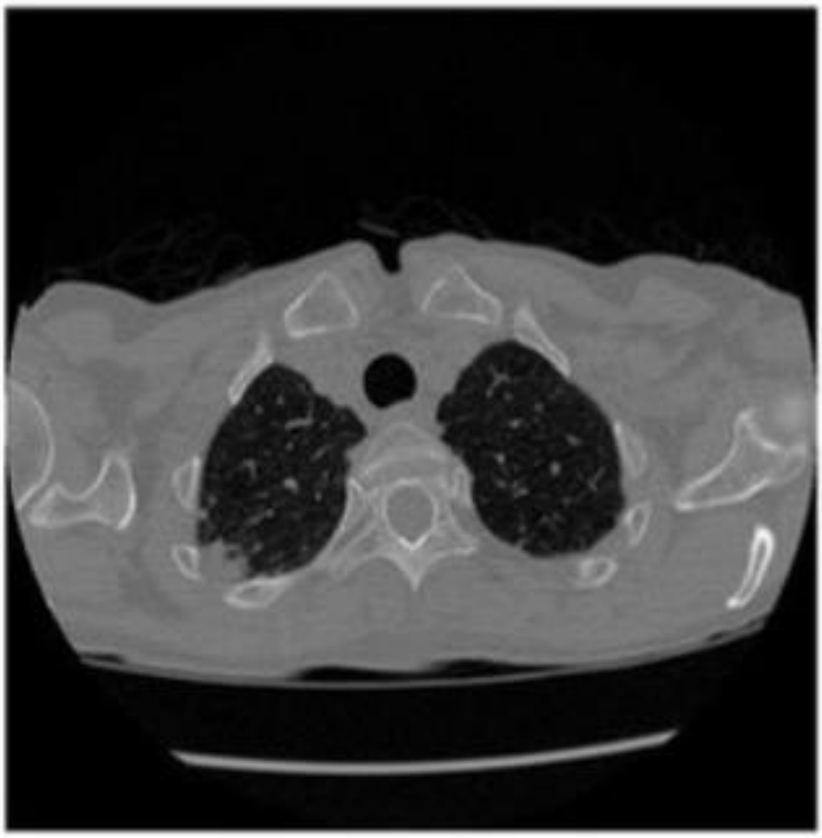


Fig: 4.1 (a) CT scan image

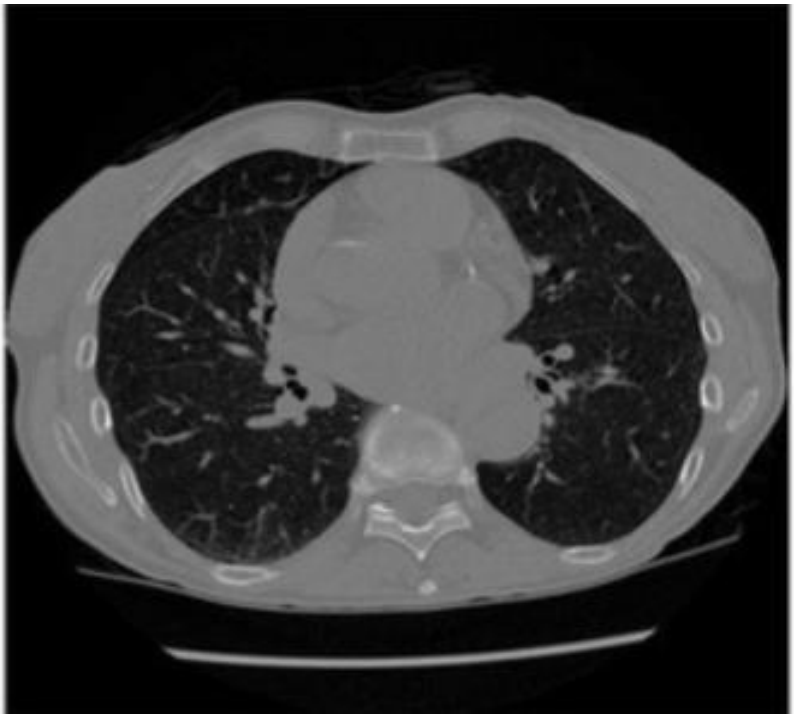


Fig:4.1 (b) Segmented image

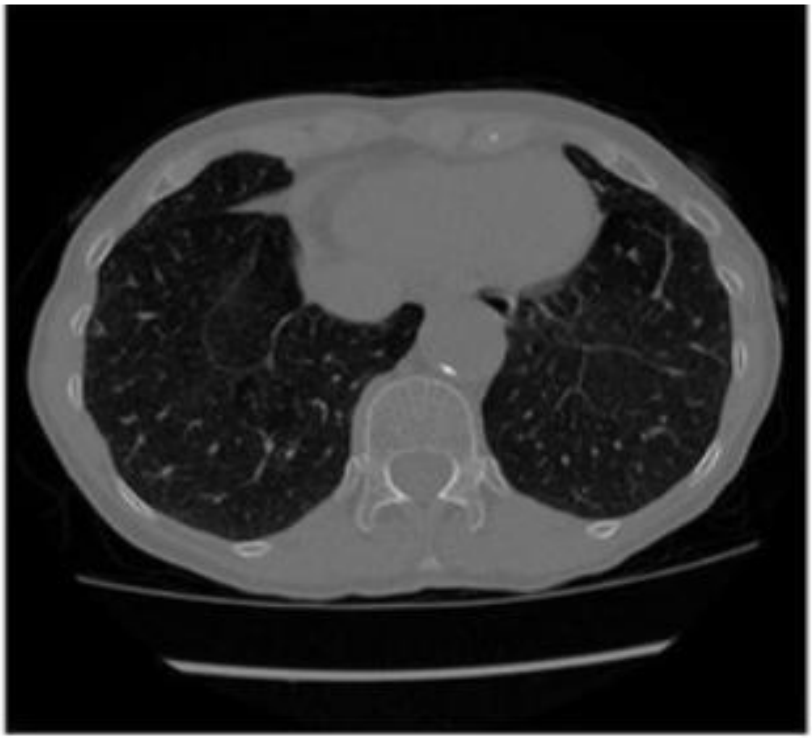


Fig: 4.1 (c) Cancer tissue

Fig: 4.2 (a) Original image

Fig: 4.2 (b) Segmented image

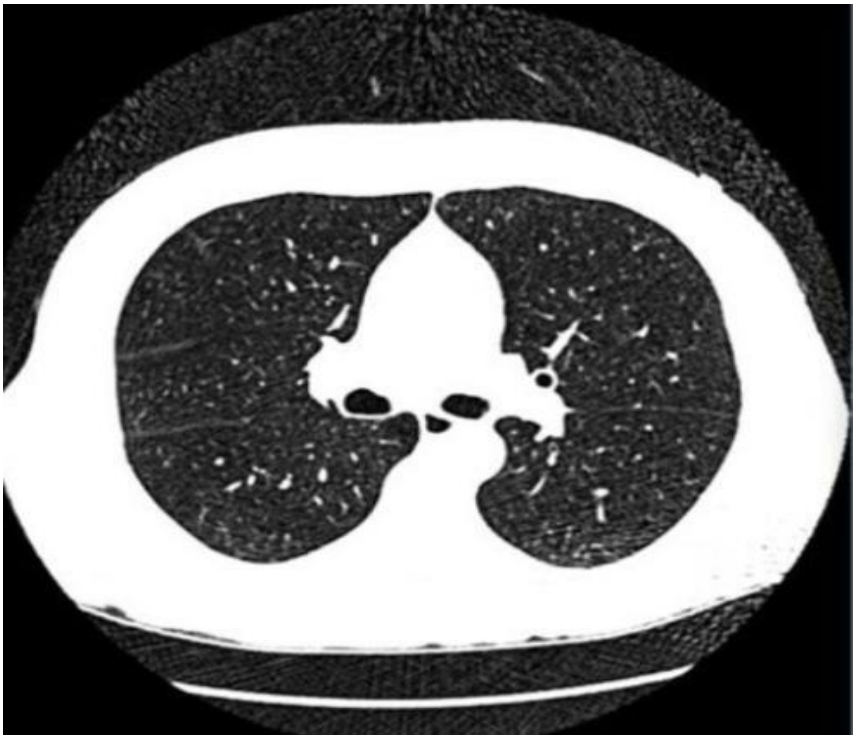


Fig: 4.3 (a) original image

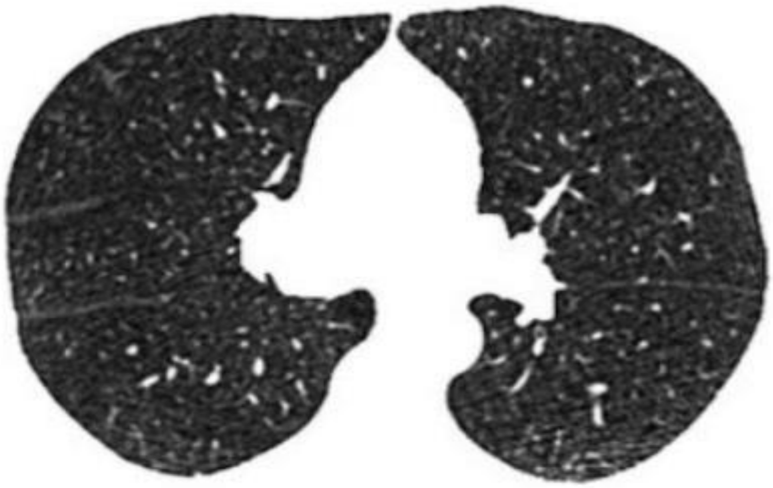


Fig: 4.3 (b) Segmented image

Fig: 4.4 (a) original image

Fig: 4.4 (b) Segmented image

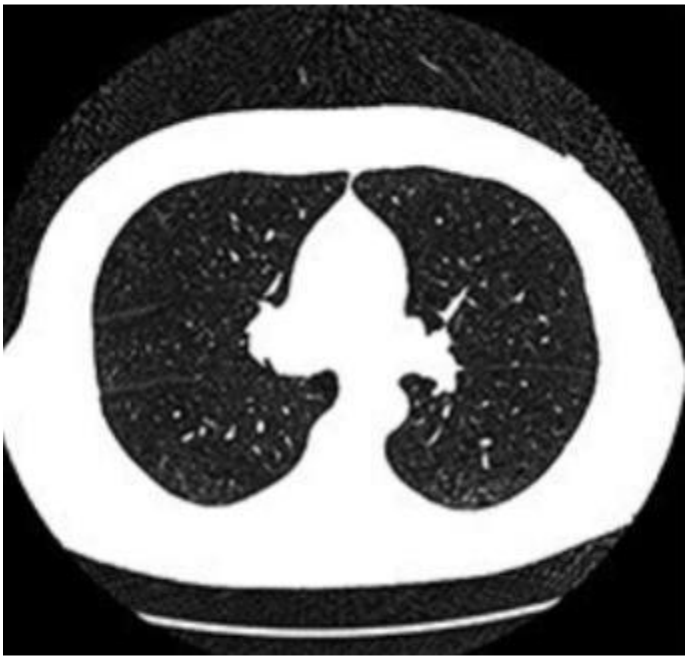


Fig: 4.5 (a) original image

Fig: 4.5 (b) Segmented image

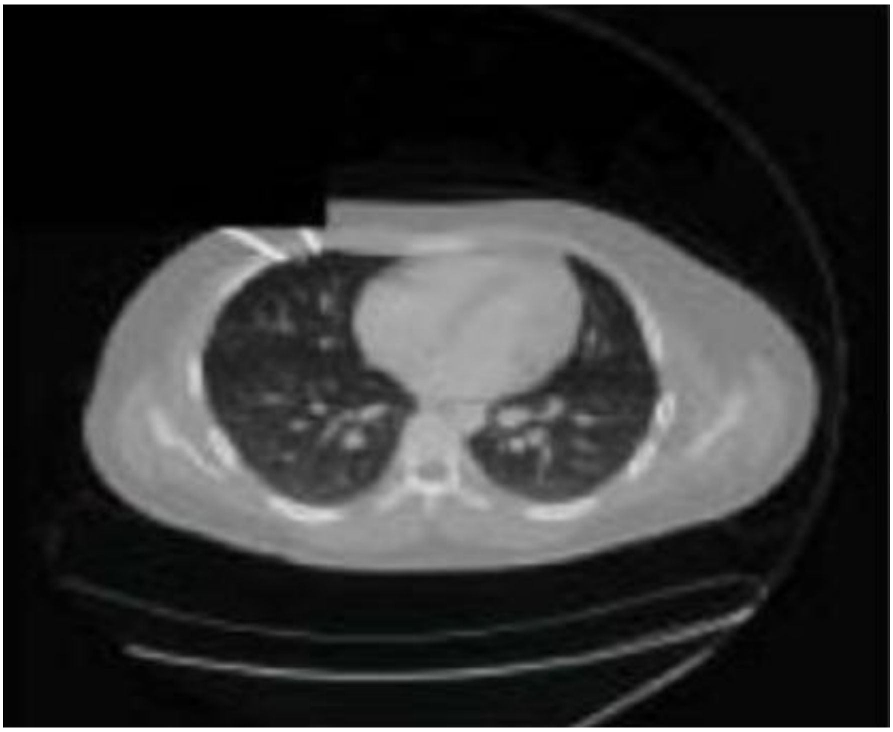


Fig: 4.6 (a) original image

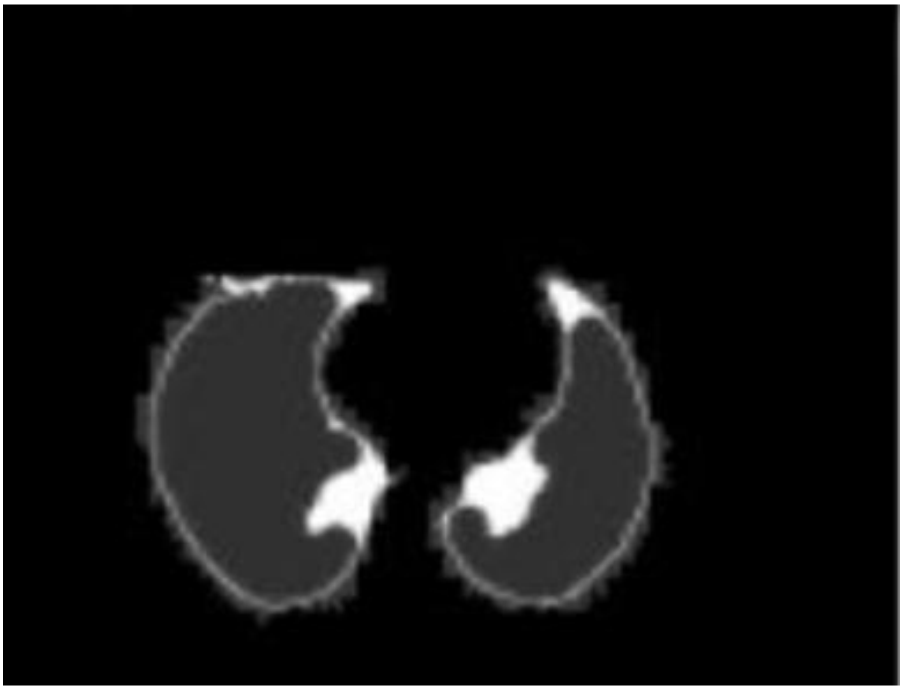


Fig: 4.6 (b) Segmented image

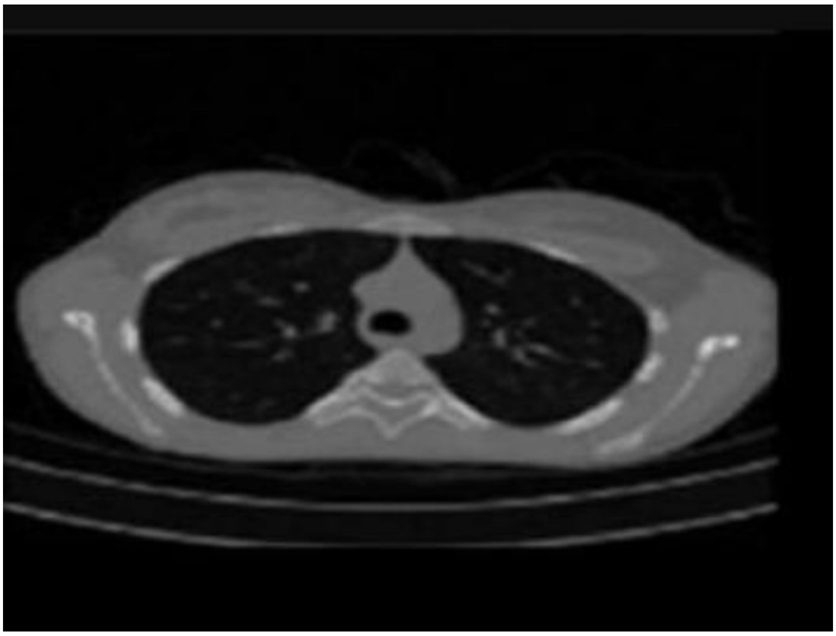


Fig: 4.7 (a) original image

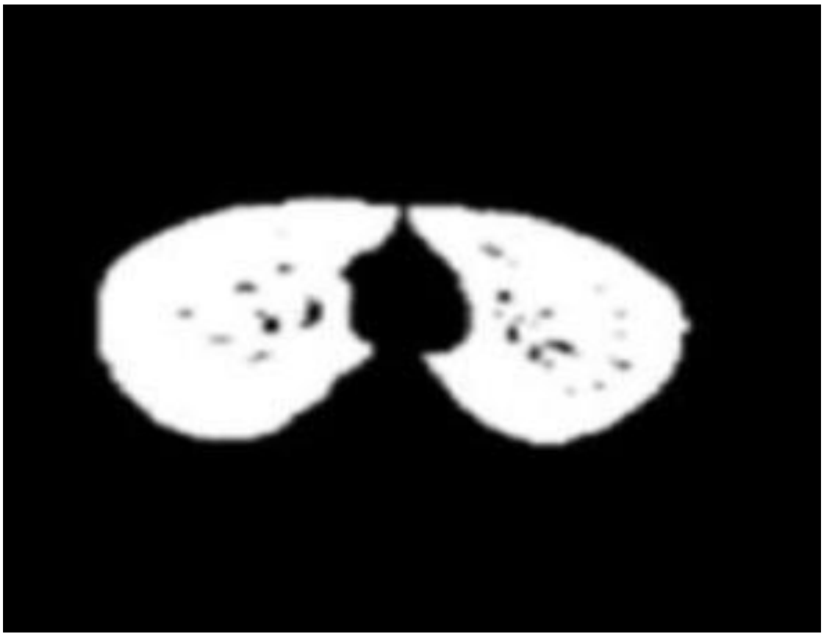


Fig: 4.7 (b) Segmented image

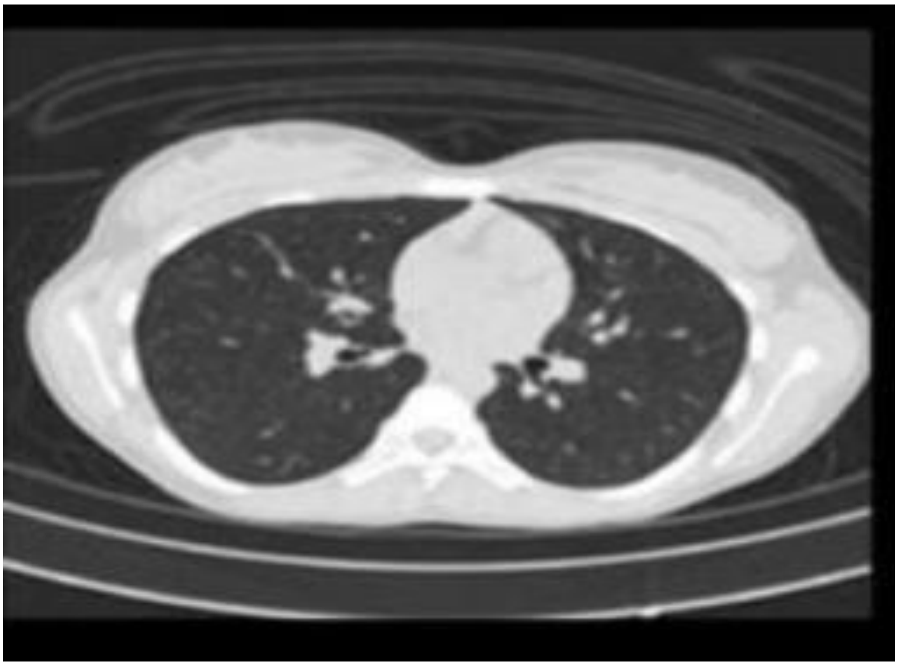


Fig: 4.8 (a) original image

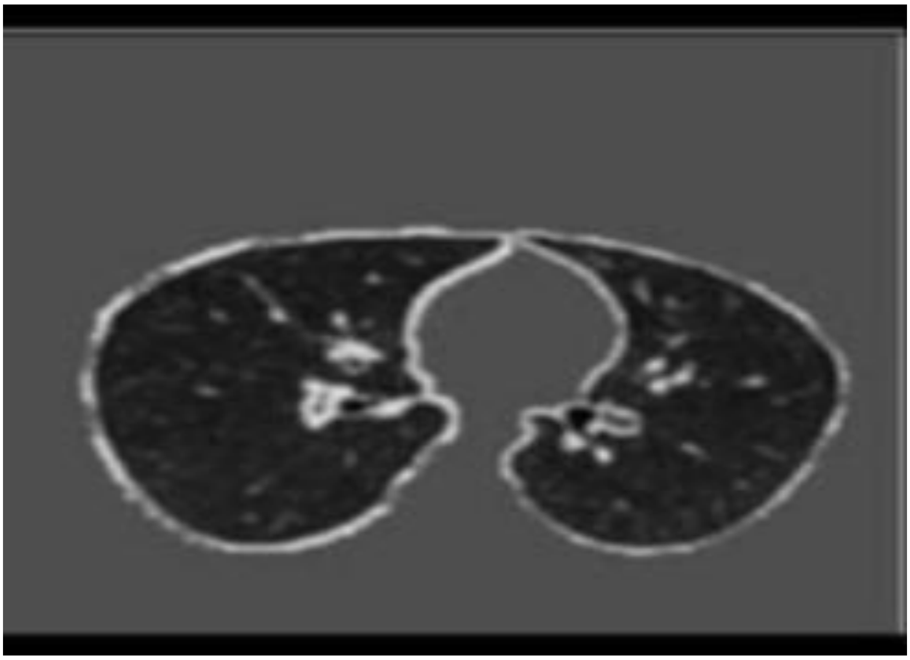


Fig: 4.8 (b) Segmented image

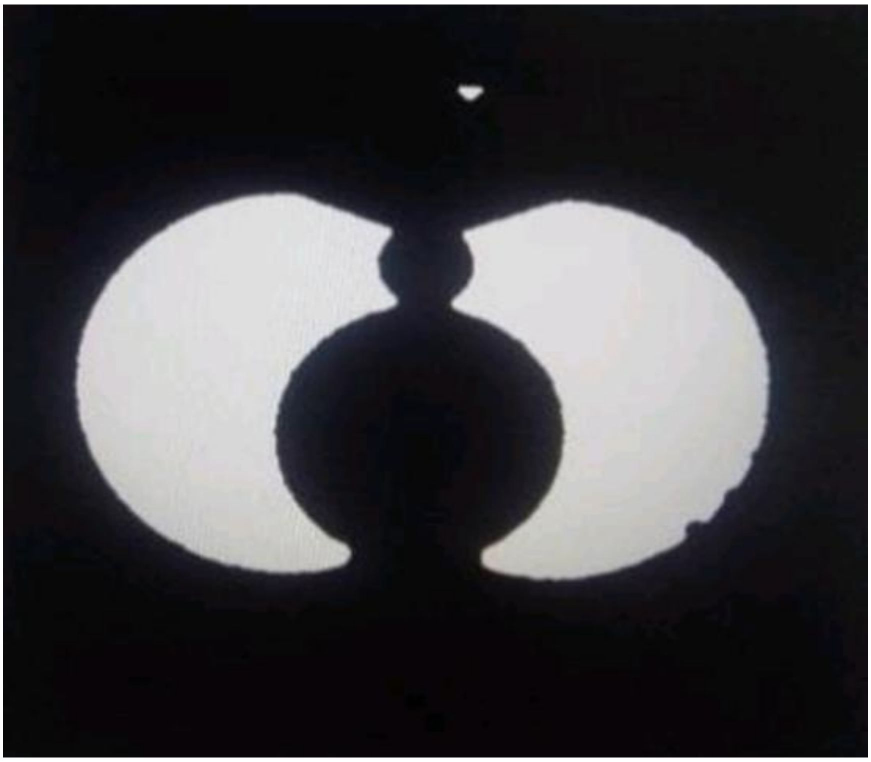


Fig: 4.9 (a) original image

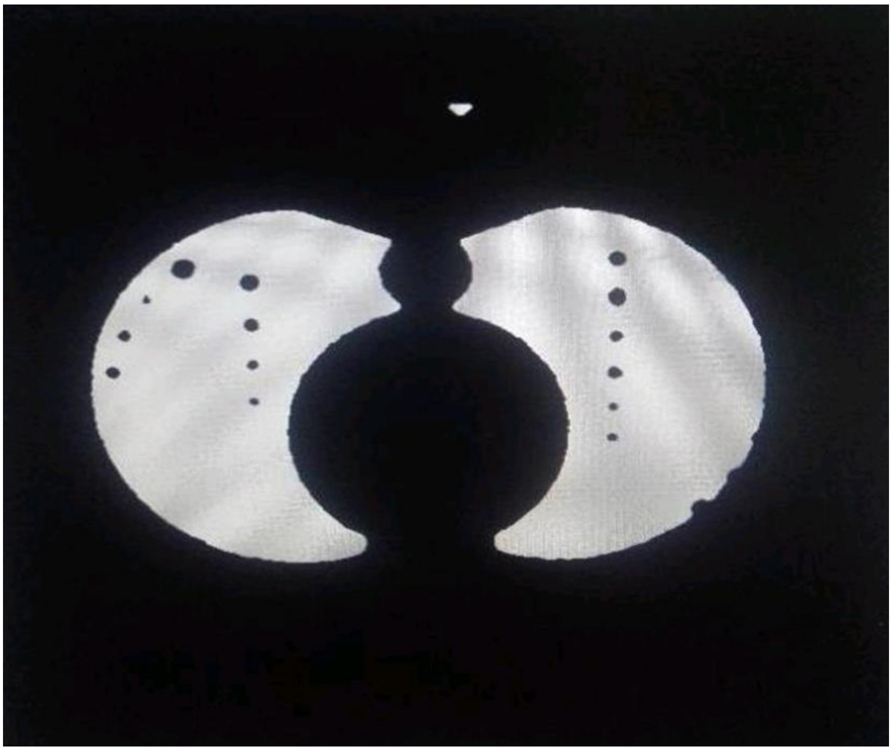


Fig: 4.9 (b) Segmented image

. **Fig: 4.10 (a) original image**

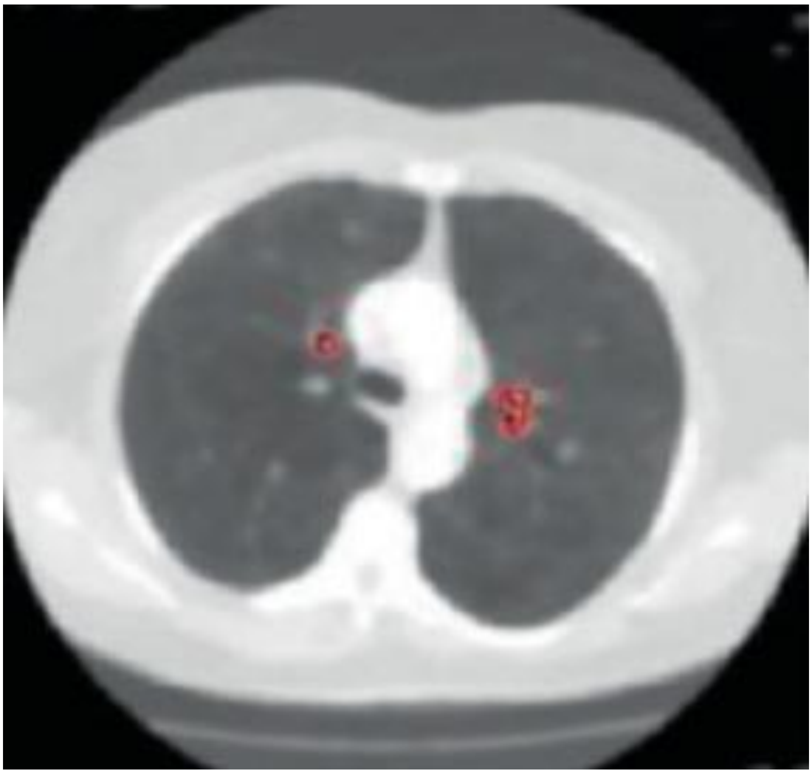


Fig: 4.10 (b) cancer found image

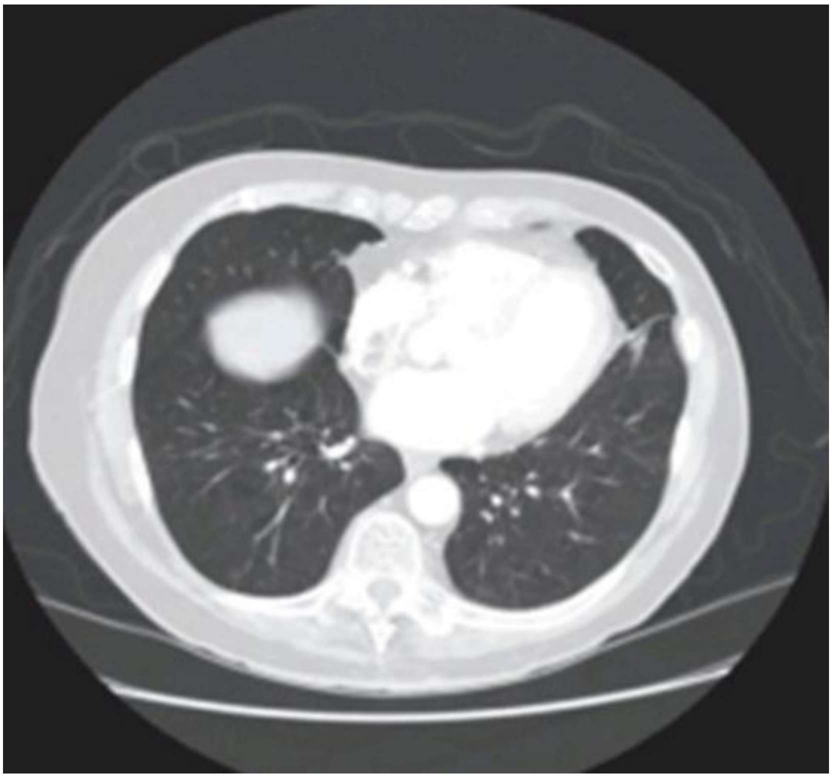


Fig: 4.11 (a) original image

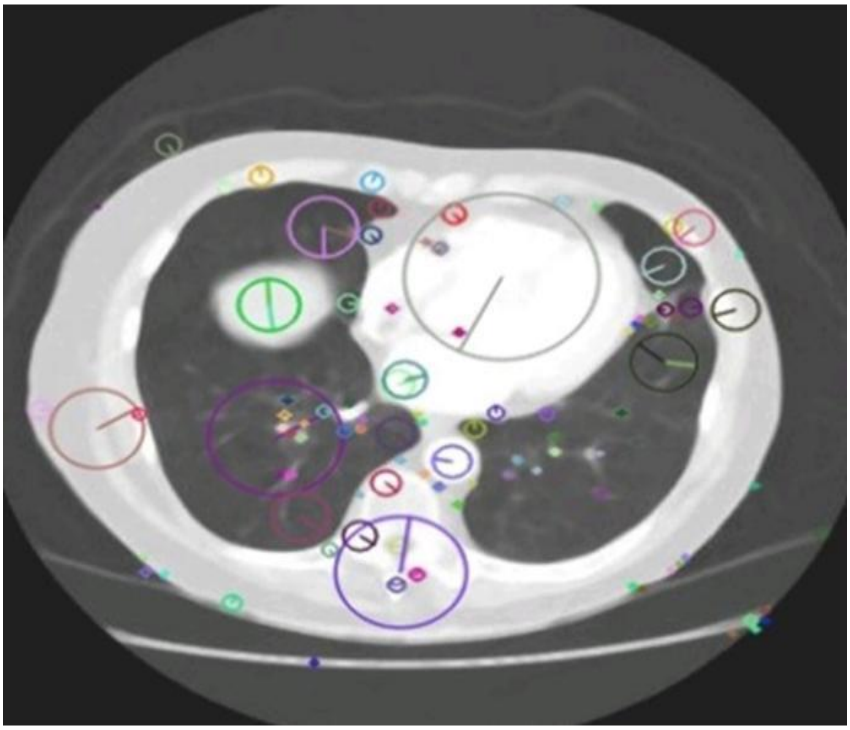


Fig: 4.11 (b) Tumour found marked

CHAPTER 5

CONCLUSION AND FUTURE WORK

The project is based on detection of lung cancer by Machine Learning algorithms. In this project a medical CT image has been considered and this image has been preprocessed to increase the quality. The image is segmented into various sub groups called image segmentation then watershed algorithm has been applied to the image. The project has successfully identified the tumour cell. This project demonstrated that Watershed segmentation is effective for detecting lung cancer in CT images by accurately identifying and segmenting suspicious regions.

Future improvements could include combining Watershed with deep learning, automating marker selection, extending to 3D segmentation, and validating the approach on larger datasets for clinical use.

REFERENCE

[1] Shradhe fule(2017), “Lung cancer detection using image processing techniques”, International research journal of engineering and technology (IRJET).

[2] Bariqi Abdillah(2017), “Image processing based detection of lungs cancer on CT scan images”, Journal of physics.

[3] Nidhi s. Nadkarni and prof. Sangam Borkar (2019), “Detection of lungs cancer in CT images using image processing” The third international conference on trends in electronics and informatics IEEE.

[4] Wasudeo Rahane,Himali dalvi, Yamini Magar, Anjali Kalane, Satyajeet Jondhale(2018), “Lung cancer detection using image processing and machine learning healthcare” International conference on current trends toward converging technologies IEEE.

[5] Pavan Kumar pagadala,Sree lakshmi pinapatruni,Chanda raj kumar, Srinivas katakam, Lalitha Surya kumari peri, Dasari anantha Reddy(2023), “Enhancing lung cancer detection from lung CT scan using image processing and Deep neural network”, International information and engineering technology association.

[6] Anita Chaudhary, Sonit Sukhraj Singh, “Lung Cancer Detection on CT Images using Image Processing”, International conference computing sciences, IEEE, 2012.

[7] Nooshin Hadavi, Md Jan Nordin, Ali Shojaeipour, “Lung Cancer Diagnosis using CT-scan Images based on Cellular Learning Automata”, International conference on Sciences(ICCOINS), IEEE , 2014. Computer and Information

[8] K.Punithavathy, M.M.Ramya, Sumathi Poobal,”Analysis of Statistical Texture Features for Automatic Lung Cancer Detection in PET/CT Images”, International Conference on Robotics, Automation, Control and Embedded systems (RACE), IEEE, 18-20 February 2015.

[9] Anjali Kulkarni, Anagha Panditrao, “Classification of Lung Cancer Stages on CT Scan Images Using Image Processing”, International Conference on Advanced Communication Control and Computing Technologies (lCACCCT),IEEE,2014.

[10] Gawade Prathamesh Pratap, R.P Chauhan, “Detection of Lung Cancer Cells using Image Processing Techniques”, 1st IEEE International Conference on Power Electronics, Intelligent control and energy systems, IEEE, 2016.

[11] Moffy Crispin Vas, Amita Dessai, “Classification of Cancerous and Non-cancerous Lung Cancer Nodule using Image Processing Technique” International Conference on Academic Research in Engineering and Management(ICAREM-17).

[12] Shubhangi Khoragade, Aditya Tiwari, “ Automatic Detection of Major Lung diseases using Chest Radiographs and classification by Artificial Neural Network”, 1st IEEE International Conference on Power Electronics, Intelligent control and energy systems, IEEE, 2016.

[13] Pooja R. Katre, Dr. Anuradha Thakare, “Detection of Lung Cancer Stages using Image Processing and Data Classification Techniques”, 2nd International Conference for Convergence in Technology (I2CT) 2017 .

[14] S. Kalaivani, Pramit Chatterjee ,Shikhar Juyal ,Rishi Gupta,“Lung Cancer Detection Using Digital Image Processing and Artificial Neural Network”, International Conference on Electronics, Communication and Aerospace Technology ICECA 2017.

[15] K.Gopi, Dr.J.Selvakumar, “Lung Tumor Area Recognition and Classification using EK-Mean Clusstering and SVM”, International Conference on Nextgen Electronic Technologies,IEEE 2017.

[16] G.Niranjana, Dr.M.Ponnavaikko, “A Review on Image Processing Methods in Detecting Lung Cancer using CT Images” ,International Conference on Technical Advancements in Computers and Communications,IEEE 2017.