Lung Cancer Detection using MATLAB

Mansi M. Dayma

(MSc Computer Science)

Abstract
Cancer is a condition in which cells in the body become uncontrolled. Lung cancer starts in the lungs and spreads to lymph nodes or other organs, such as the brain, inside the body. Cancer can potentially spread to the lungs from other organs. When cells of cancer travel from one organ to another, metastases are termed. In 2019 alone, America registered 228820 instances, 135720 of them. The current technique for the identification of lung cancer employs X-ray or CT scans to recognize the extra weight of the pulmonary flesh. The traditional method needs that you be diagnosed manually by your doctor, and that sometimes is a long process, and we want to build a manual process machine based on this project which will help you to diagnose your condition quickly. Various CT-Scan pictures were obtained from various hospitals and various techniques such as image segmentation and Gabor filter have been utilized in this research.
Keywords: *CT Scan Image, irisation, segmentation, thresholding, watershed, Gabor filter.*

Date of Submission: 28-09-2021   Date of Acceptance: 12-10-2021

I.   Introduction

Lung abnormality is a tumor-growing illness with aberrant cells. The lungs or the lymph fluid that surrounds lung issues may take abnormal cells in the blood. Lymphatic vessels that drain into a lymph node in the lung and center of the chest run through the lymphatic vessels. Lung irregularity typically stretches to the middle of the chest because the lymph's natural flow from the lungs goes to the middle of the chest. Metastasis happens when an aberrant cell leaves the place where it started and travels via the bloodstream to a lymph node or to an other area of the body. The main lung abnormality which begins in the lungs. Approximately 85% of men and 75% of women suffer from pulmonary cancer because of cigarette smoking.

The overall survival percentage for lung cancer patients is 63%. Lung irregularity exists in many kinds and is classified into two main groups: Small lung defects and non-small pulmonary defects with three subtypes: Squamous cell carcinoma and adenocarcinoma.

In Jordan, in the ranks of anomalies between men and females in 2008, 356 instances (7.7%) of newly confirmed cases of lung abnormality were reported in 2008.

297 (13.1%) males and 59 (2.5%) females with a 5:1 male-to-female ratio, with Lungs being the second-most aberrant among males and 10th among females, were afflicted by lung abnormalities. The illness description of a system containing five main phases of lung abnormalities. The first stage is starting with the use of an IMBA Home database to gather CT scans (normal and aberrant) (VIA-ELCAP Public Access).

Lung cancer can be classified into several stages, including:

Stage I: Cancer cells are exclusively seen in the lungs and have not migrated to lymph nodes. The tumor is around 3 cm in diameter.

Stage II: Cancer cells have spread across the lung and surrounding lymph nodes. The tumor is between 3 cm and 7 cm in size at this stage.

Stage III: Cancer cells appear in the lungs and lymph nodes in the center of the chest at this Stage. The tumor has grown to a size of more than 7 cm at this point.

Stage IV: This stage of cancer is extremely deadly and incurable. Cancer cells are at this time disseminated to the lungs or to a different area of the body, such the liver or another organ.

II.   Methodology

Image Pre-Processing

In the first place we resize the picture since the different sources of pictures are different. Image resizing is the most significant component in determining the stage of cancer in image processing. The image is transformed into a grey image after resizing. The value of each pixel in the grey image is a single sample, which solely provides pixel intensity information. The picture consists of unique grey colors. Such images are the result of the measurement of the light intensity at each pixel in a single light spectrum band, such as infrared, ultraviolet and visible light, etc.

Image Enlargement

Image enhancement is used to increase the quality of the image. This method may be used in both spatial and frequency domains. The goal of Image Enhancement is to make the image better by enhancing its quality. An enhancement algorithm is employed for a specific application, and it can either increase or decrease the image's noise or contrast. Image enhancement allows us to adjust the brightness of an image according to our needs. The image's visual perception is improved with this method. In general, this approach may be divided into two groups.
1.   Spatial Domain Method
2.   Transfer Domain Method
A spatial domain technique is operated directly on the picture, whereas the domain method transforms the image from Fourier to the spatial domain again. Techniques for improvement are based essentially on histograms, as it is extremely easy and quick. We can achieve some application that we desire with acceptable outcomes. This approach emphasizes and sharpens picture characteristics including edges, borders and contrasts, so that graphical data may be easily viewed inside the image. We utilized the Median filter for improvement.

Median Filter

To minimize picture noise, the median filter is a highly helpful and successful approach. This filter identifies edges; first, the noise in the picture should be eliminated, and then the edges should be removed. The median filter's primary function is to reduce noise while leaving edges intact. Average filter is also used to reduce noise, however because the average of pixels in the window replaces the mid pixel intensity, some data is lost. So, for noise removal, we utilized the Median filter. As with the linear Gaussian filter, the median filter is a smoothing method.

The median filter's primary concept is to run each input and replace it with the median of surrounding pixels' intensity entries.

The "window" refers to the arrangement of neighbors. The most evident window for a 1D signal is only the first few preceding and following entries, but more complicated window patterns are available for 2D (or higher-dimensional) pictures.
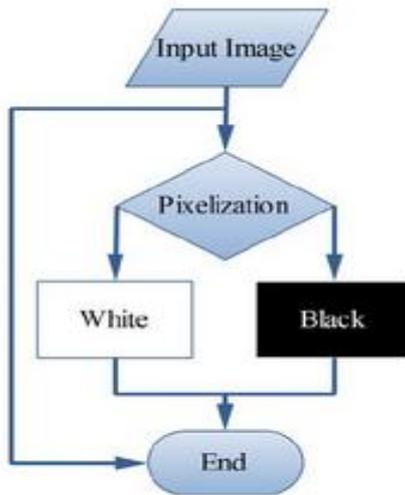
Binary Image Processing

We utilize morphological operations for this phase of our study because it includes certain binary operations. The term "morphology" refers to the study of an object's shape, size, and structure. The structural element is a notion we employ. The structuring element is a mask or window that is placed in the original picture to locate the required output and is significantly smaller than the original image. Dilation, erosion, opening & closure are the four basic morphological processes. We utilized opening and shutting in this example, and in order to do so, we must first grasp dilation and erosion morphological procedures.

Binarization Approach :

The approach of binarization depends on the fact that in a normal pixel we start counting the black pixel to normal and abnormal images to get an average number later on as a threshold if there are more Black Pixels in a new picture than the threshold, otherwise the image will be normal if the pixels of a new image are greater than the threshold if it's average, otherwise it's t normal. The threshold for the research was 17178.48, with the true rate of acceptance (TRA) being(92.86%) and the fake rate of acceptance (FAR) being (7.14 percent).

**Binary Method Procedure** **Binarization Check method Flowchart**
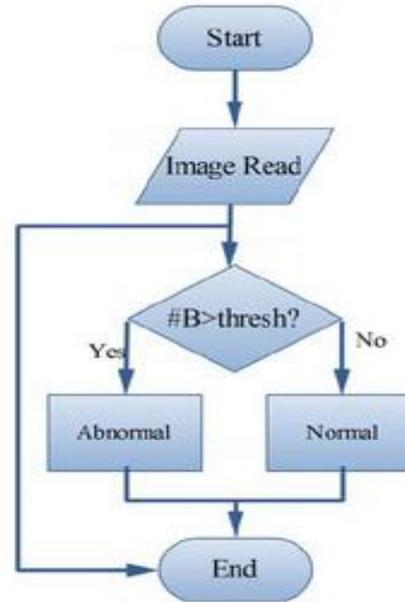
Thresholding:

For the categorization of the input pictures, we've utilized the number 17179. After all of the necessary steps for the process have been completed, such as image enhancement, image segmentation, and feature extraction, we classify the images using binarization, which gives us the matrix of the final image. If the value is greater than the threshold value set, the image has lung cancer, and if the threshold value is less, the image does not. Thresholding may also be done in colour pictures. A distinct threshold is designated, and a combination of each RGB component of the picture with an AND function is used. This represents the way the camera works and the way the information is stored on the computer, but it does not match the way people know the colour. The colour patterns of HSL and HSV are increasingly often employed, and because the hue might be a circular number, circular thresholds are required.

Masking Approach :

The masking technique hinges on the fact that masses appear in ROI (lungs), as they increase the percentage of malignancy. The appearance of solid blue colors is normal, whereas RGB masses show the presence of cancer, TAR is (85.7%) and FAR is (85.7%) (14.3 percent).

Figure below illustrates typical and aberrant pictures that come from MATLAB implementation.
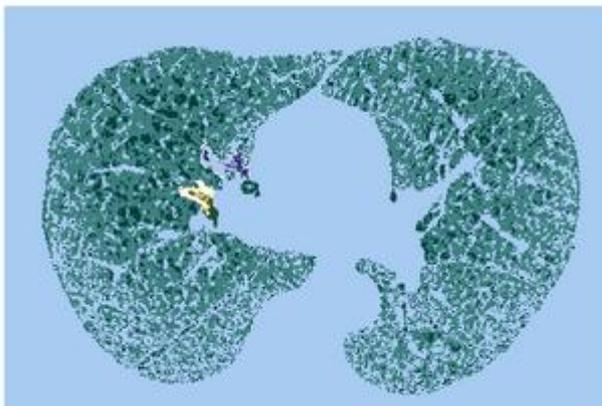
Normal Image Enhancement by Gabor,   The Resulted Image Indicates Normality
  Segmented by watershed

Abnormal Image   The results Indicate Abnormality
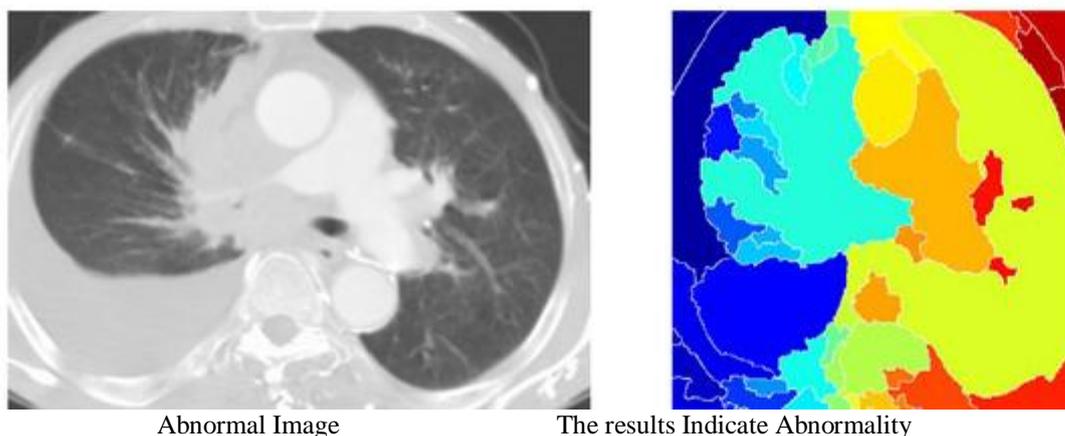
By combining the Binarization and Masking methods, we can determine if the situation is normal or aberrant. We may infer that a picture with a higher number of black pixels than white pixels suggests normalcy, while an image with a smaller number of black pixels shows abnormality, based on the assumptions stated in the preceding two techniques.

Image Segmentation Based Region Growing

Growing a region is the easy way of dividing images. The pixel segmentation technique is also categorized since the choosing of starting seed points is required.

The segmentation technique evaluates the neighboring pixels of the original seed points and assesses if the neighbors to the region should be added. It is because first choose the world that will be the goal object, which is the correct lung and left lung, then place seed on this area, that this is the procedure to urge segmentation by region growth technique. The developing process of this project in the region gives us a clear concept and picture of the picture, which helps us to further process the defined regulated watershed and the binarization process. Different objects might be used to utilize their pixel values owing to segments. An essential item to note – if there is a strong difference between the pixel values of the objects and the backdrop in the image. A threshold value can be set. This approach may be characterized as the threshold segmentation of the pixel values below or above this threshold.

(a) Enhance by Gabor   (b) Segmented Image

Fig : 2 Region Growing of the Gabor Filter Image

Marker Controlled Watershed Segmentation

The transformation on the grayscale can be characterized as a watershed. The change in the water-shed deals with the picture it works on a topographical map, in a manner that reflects the brightness of each point and the lines that line the crests of ridges. To obtain the watershed method's segmentation results the following procedures are: Firstly, compute the edge detection space gradient, then choose the target object using a morphological approach called reconstruction opening and reconstruction shutting.

Other regions are frequently removed when the marking object is exposed. Controlled watershed segmentation modules and example of the marker. The outcome of segmentation using Marker- Controlled Watershed is shown in the figure.

Image of Database

This database includes 900 photos of lung cancer kinds, 900 images of lung cancer of non-small cells, a total of 1800 images each of 200×200 pixels. We have this IMBA Home database (VIA-ELCAP Public Access). Some of the image we obtained from the clinicians and the Ct-scan centers.

Feature Extraction

Function extraction is a sort of reduction in dimensionality which is an important component of an image as a compact vector. This method is useful when the picture files are big and feature representation is minimized for activities such as image matching and recovery to be completed fast. Feature detection, extraction and correlation are typically associated with unravelling common computer vision issues such as object identification and recognition, content-based picture recovery, and face detection and recognition as well as texture classification are important features of all projects. Extraction of a feature is a sort of reduction in dimension which effectively depicts interesting pieces of an image as a compact feature vector.

That method is advantageous when picture sizes are big and feature displays are minimized to accomplish tasks such as image matching and recovery fast.

III. Conclusion

According to stage discovery of abnormality cells in the lungs, Lung Abnormality is the most dangerous and widespread disease in the world; this indicates that the process of disease detection is extremely important and essential in avoiding serious stages and reducing the disease's percentage distribution around the world. Lung cancer is the most common cause of mortality among people all over the world. Lung cancer is associated with a high death rate. Lung cancer, breast cancer, and other cancer tumors are examples of cancer tumors.

For effective therapy, early cancer diagnosis is essential. Computed tomography (CT) scans are used for diagnosis. The images were pre-processed and the method of extraction and classification were used to see a patient's condition whether it was normal or anomalous in its early stage.

Recently, image processing techniques have been widely used in several medical areas for image improvement in earlier detection and treatment stages, where the time factor is extremely important to get the abnormality issues in target images, especially in various abnormality tumors such as lung abnormality, breast abnormality, and so on, where the time factor is extremely important to get the abnormality issues in target images, especially in various abnormality tumors such as lung abnormality, breast abnormality, and so on.

The key components of this research are picture quality and accuracy; image quality evaluation and improvement are based on the enhancement stage, where low-pre-processing approaches based on Gabor filters inside Gaussian rules are utilized.

References

[1]. Introduction of lung cancer is available: https://en.wikipedia.org/wiki/Lung_cancer
[2]. Makaju, Suren, P. W. C. Prasad, Abeer Alsadoon, A. K. Singh, and A. Elchouemi. "Lung Cancer Detection using CT Scan Images." Procedia Computer Science 125 (2018): 107-114.
[3]. El-Regaily, Salsabil A., Mohammed A. Salem, Mohammed H. Abdel Aziz, and Mohammed I. Roushdy. "Survey of Computer Aided Detection Systems for Lung Cancer in Computed Tomography." Current Medical Imaging Reviews 14, no. 1 (2018): 3-18.
[4]. Kalaivani, S., Pramit Chatterjee, Shikhar Juyal, and Rishi Gupta. "Lung cancer detection using digital image processing and artificial neural networks." In Electronics, Communication and Aerospace Technology (ICECA), 2017 International conference of, vol. 2, IEEE, (2017):100-103.
[5]. Kumar, S. Logesh, M. Swathy, S. Sathish, J. Sivaraman, and M. Rajasekar. "Identification of lung cancer cell using watershed segmentation on CT images." Indian Journal of Science and Technology 9, (2016)
[6]. Dimililer, Kamil, Buse UGUR, and K. Yoney. "Tumor detection on CT lung images using image enhancement." In International Science and Technology Conference (ISTEC 2016). (2017).
[7]. Kumar, Devinder, Alexander Wong, and David A. Clausi. "Lung nodule classification using deep features in CT images." In Computer and Robot Vision (CRV), 2015 12th Conference on, IEEE, (2015): 133-138.
[8]. Hua, Kai-Lung, Che-Hao Hsu, Shintami Chusnul Hidayati, Wen-Huang Cheng, and Yu-Jen Chen. "Computer-aided classification of lung nodules on computed tomography images via deep learning technique." OncoTargets and therapy 8 (2015).
[9]. Gajdhane, M. V. A., & Deshpande, L. M. (2014). Detection of Lung Cancer Stages on CT scan Images by Using Various Image Processing Techniques. IOSR Journal of Computer Engineering (IOSR-JCE) e-ISSN, 2278-0661.
[10]. Introduction of lung cancer available: https://www.medicinenet.com/lung_cancer/article.ht m#lung_cancer_facts
[11]. Stages of lung cancer is available:https://www.healthline.com/health/lung-cancer#back-pain
[12]. Grayscale Image available: https://en.wikipedia.org /wiki/Grayscale
[13]. Chaudhary, Anita, and Sonit Sukhraj Singh. "Lung cancer detection on CT images by using image processing." In Computing Sciences (ICCS), 2012 International Conference on, IEEE, (2012): 142-146.

[14]. Gonzalez, Rafael C., and Richard E. Woods. "Digital image processing prentice hall." Upper Saddle River, NJ (2002).
[15]. Van Ginneken, Bram, BM Ter Haar Romeny, and Max A. Viergever. "Computeraided diagnosis in chest radiography: a survey." IEEE Transactions on medical imaging 20, no. 12 (2001):1228-1241.
[16]. Median filter available: https://en.wikipedia.org /wiki/Median_filter
[17]. Pratap, Gawade Prathamesh, and R. P. Chauhan. "Detection of Lung cancer cells using image processing techniques." In Power Electronics, Intelligent Control and Energy Systems (ICPEICES), IEEE International Conference on, IEEE, (2016):1-6.
[18]. Jayaraman, S., and S. Esakkirajan. "Digital image processing." (2009).
[19]. Gonzalez, Rafael C., and Richard E. Woods. "Digital image processing prentice hall." Upper Saddle River, NJ (2002).
[20]. Armato, Samuel G., Maryellen L. Giger, and Heber MacMahon. "Automated detection of lung nodules in CT scans: preliminary results." Medical physics 28, no. 8 (2001):1552-1561.
[21]. Mokhled, S. "Lung Cancer Detection Using Image Processing Techniques." Leonardo Electronic Journal of Practices and Technologies 11, no. 20 (2012): 147-158.
[22]. ROI is available: https://en.wikipedia.org/wiki/ Region_of_interest
[23]. https://angeljohnsy.blogspot.com/2012/05/find-area-perimeter-centroid.html
[24]. Figure2 is available: https:// www.emedicinehealth. com/lung_cancer/article_em.htm
[25]. Pixel to Cm conversion available :http://www.unitconversion.org/typography/pixels-x-to-centimeters-conversion.html.