JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

Detecting Lung Cancer in Mat-Lab Image Processing Techniques with the Gabor filter and SVM classifier

[1]Divyansh Dadhwal, [2]Dr. Avni Sharma
[1]Master's Student, [2]Faculty Member
[1] Computer Science & Engineering,
[1]Himachal Pradesh Technical University, Hamirpur, Himachal Pradesh, INDIA

Abstract: Lung cancer is still one of the deadliest and frequently diagnosed malignancies in the world, and patient survival depends heavily on early identification. Due to human subjectivity, manual CT scan examination and biopsy are two time-consuming and error-prone traditional diagnostic techniques. Image processing methods have greatly increased the speed and accuracy of medical diagnosis in recent years, particularly for the identification of cancer. Advanced image processing methods can be implemented on a comprehensive platform with MATLAB, a high-level language and interactive environment for numerical computation and visualization. This study uses MATLAB-based image processing to investigate a methodical approach for early lung cancer identification. Image acquisition, pre-processing, segmentation, feature extraction, and classification are some of the steps in the methodology. The lung CT images are enhanced and segmented using methods such region-growing, median filtering, histogram equalization, and morphological processes. To find possible malignant nodules, characteristics including texture, shape, and intensity levels are taken out and examined. Lastly, to increase detection accuracy, machine learning classifiers such as Support Vector Machines (SVM) are employed. The study shows how well MATLAB works to automate and optimize the diagnosis of lung cancer, which may lower false positives and increase the rate of early detection.

***Index Terms* – Lung Cancer, Image Processing, MATLAB.**

I. INTRODUCTION

The primary cause of cancer-related mortality for both men and women, lung cancer is a significant worldwide health concern. The World Health Organization (WHO) estimates that lung cancer kills around 1.8 million people each year. The late identification of the disease, when the cancer has already progressed to an advanced and frequently incurable stage, is one of the main factors causing this high death rate. Improving patient outcomes requires early diagnosis, and medical imaging is at the forefront of early cancer detection techniques.

A mix of imaging modalities, including X-rays, CT scans, and biopsies, are commonly used in traditional ways of diagnosing lung cancer. However, radiologists are largely responsible for interpreting these images, which could result in diagnostic errors because of visual impairments, weariness, or experience gaps. Additionally, manual analysis takes a lot of time and is biased by humans. The need for automated, precise, and effective methods to aid in the diagnosis process is therefore increasing.

By facilitating the automatic examination of medical pictures, image processing provides a potent remedy for these issues. It entails using algorithms to medical images in order to carry out tasks including classification, segmentation, and improvement. Because of its powerful toolboxes, including the Image Processing Toolbox, which enable quick algorithm testing and prototyping, MATLAB is frequently used for this purpose in both academic and industrial contexts.

Image processing methods can assist in detecting anomalies like pulmonary nodules, which are frequently suggestive of malignancy, in the context of lung cancer diagnosis. Image capture (generally CT scans) is the first step in the detection process. Preprocessing is then done to improve image quality and lower noise. After that, segmentation is done to separate areas of interest, such lung nodules. To collect diagnostic data based on size, shape, and texture, feature extraction techniques are used. Lastly, the likelihood of a benign or cancerous region being found can be ascertained by classification algorithms.

This work uses MATLAB image processing techniques to give a thorough investigation of lung cancer detection. The main goal is to create a dependable and effective framework that improves diagnostic accuracy by utilizing MATLAB's computational capabilities. The methods used, experimental findings, and efficacy analysis of the suggested system are all covered in detail in the following sections.

II. LITERATURE REVIEW

Image processing methods for lung cancer identification have been the subject of numerous studies, with a focus on the use of computer-aided diagnosis (CAD) systems in assisting radiologists. The performance of CAD systems on the LIDC-IDRI dataset was examined by Armato et al. in 2012, who also emphasized the difficulties associated with false positives. A thorough evaluation of image analysis methods for lung cancer screening was also given by Gurcan et al. (2009), who concentrated on segmentation and classification strategies.

The results of recent developments in MATLAB-based frameworks have been encouraging. Using MATLAB, Kumar and Das (2016) developed a lung cancer detection model that used morphological procedures for picture preprocessing and SVM classification to achieve 85% accuracy. In a different study, Sharma et al. (2019) isolated nodules using active contour segmentation and histogram equalization, which increased feature extraction.

Notwithstanding these developments, there are still issues with generalizability, false positive reduction, and dataset quality. Numerous studies only use a small number of CT scans, which compromises the models' resilience. By utilizing MATLAB's sophisticated image processing toolkit and an enhanced preprocessing-segmentation-feature extraction pipeline, this study seeks to close these gaps.

III. Research Methodology

This study suggests a MATLAB-based framework for lung cancer diagnosis that uses texture-based lung nodule segmentation with Gabor filters. Because they can record both spatial and frequency information, gabor filters are frequently employed in texture analysis. This makes them perfect for identifying abnormal patterns, like tumors, in medical imaging.

There are five main steps in the methodology:

3.1 Taking Pictures

High-resolution CT or X-ray pictures of the chest are gathered from local radiological archives or publically accessible databases like the LIDC-IDRI. The system uses these grayscale pictures as its input.

3.2 Preprocessing Images

To enhance image quality, preprocessing is done prior to using the Gabor filter:

- Noise Removal: To reduce salt-and-pepper noise without significantly sacrificing edge information, a median filter is used.
- Histogram Equalization: This improves the image's contrast and makes it easier to distinguish between various tissue types.
- Lung Region Isolation: The lung region is separated from the surrounding anatomy (such as the diaphragm and bones) by morphological surgeries.

3.3 Segmentation Based on Gabor Filters

Basics of Gabor Filters:

A linear filter used for texture classification and edge detection is called a Gabor filter. It is characterized by a Gaussian envelope modulating a sinusoidal wave. Certain frequencies and orientations cause the filter to react.

Steps for Implementation in MATLAB:

- Several Gabor Kernels are created with different frequencies and orientations (such as 0°, 45°, 90°, and 135°).
- To extract directional texture characteristics, the input image is convolved with each Gabor filter.
- A composite response is created by combining the filtered images and using textural imperfections to highlight suspicious areas (like nodules).
- High-response zones, which most likely indicate tumor locations, are extracted using a thresholding operation.
- Based on how their texture differs from that of normal lung tissue, our segmentation technique successfully separates possible tumor zones.

3.4 Extraction of Features

A classification model is trained using a variety of features that are taken from the segmented regions:

- Texture Features: Gabor response mean, energy, entropy, and standard deviation
- Shape characteristics include area, perimeter, eccentricity, and compactness.
- First-order (mean, variance) and second-order (GLCM-based contrast, correlation) statistical features
- These characteristics show the possible tumor regions' patterns of intensity as well as their structure.

3.5 Grouping

For a final diagnosis, the retrieved features are fed into machine learning classifiers:

- SVM, or support vector machine, is well-known for working well with high-dimensional data.
- For improved generalization, Random Forest (Optional) offers ensemble learning.
- Performance Metrics: Classifier performance is assessed using F1-score, accuracy, sensitivity, specificity, and precision.

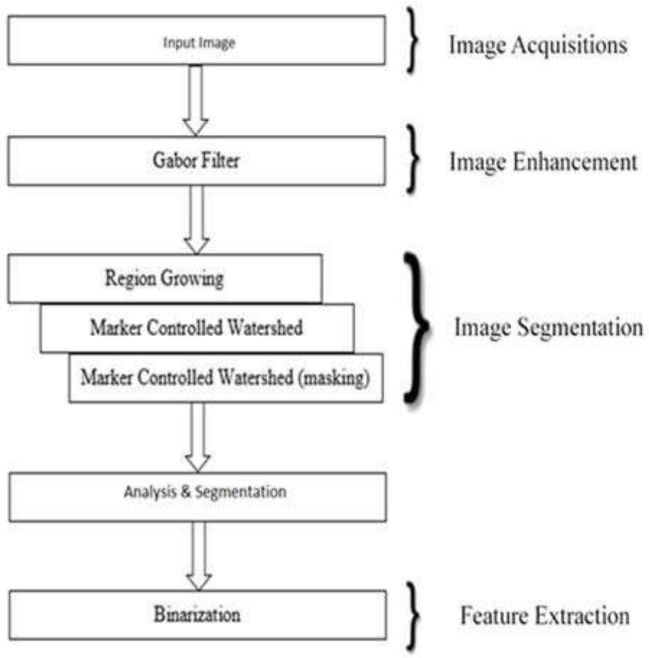

Fig. 1: Flow Chart of System

IV. Result & Discussions

We apply and evaluate the image processing technique for lung cancer detection. In order to help the initial medical treatment, image processing techniques are widely used in a variety of medical conditions for picture improvement during the detection phase. In this study, we suggested a carcinoma-supported picture segmentation detection technique. One of the intermediate levels of image processing is picture segmentation. The CT scan image is segmented using the marker control watershed and region growth technique. Image augmentation utilizing the Gabor filter, image segmentation, and feature extraction come after detection phases. The watershed with masking method, which is very accurate and reliable, is the ideal strategy for detecting key characteristics.

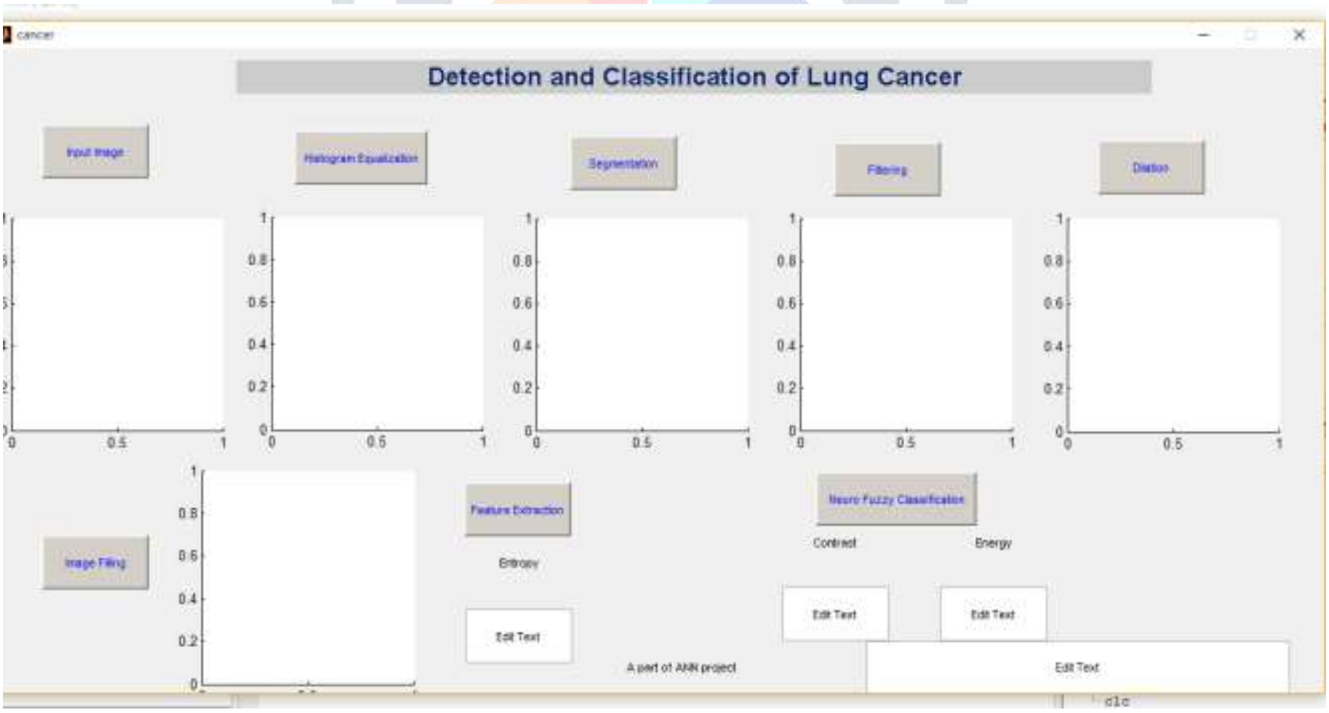

Fig.2: Basic GUI Interface of Project

Fig. 3: Input Image

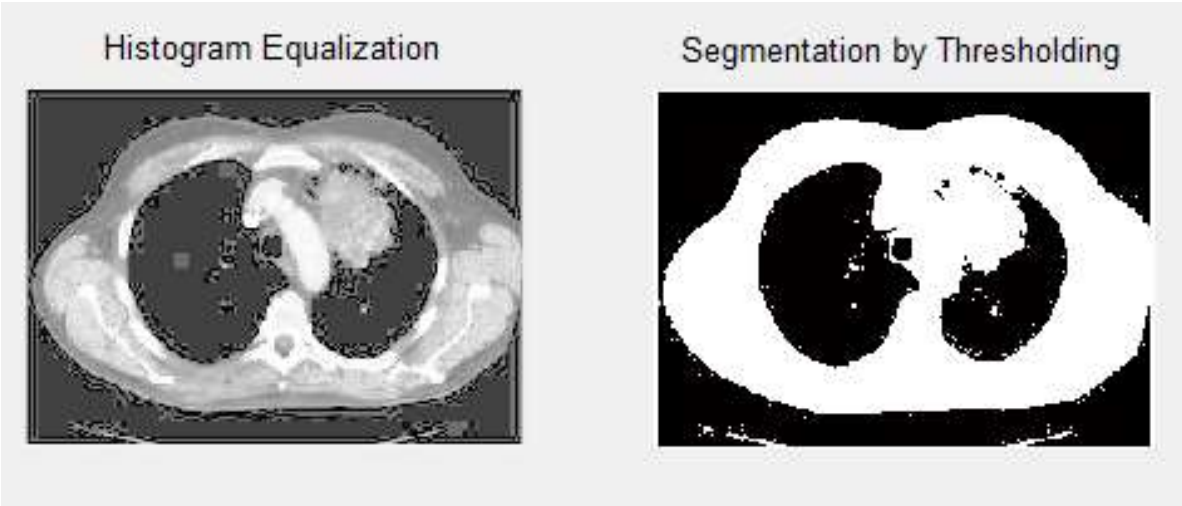

Fig. 4: After Appling Histogram and Thresholding

Fig. 5: After Applying Filter and Dilated

We gathered three datasets from the Internet, each of which included both malignant and non-cancerous photos. There are 332 photos in dataset 1, 314 images in dataset 2, and 257 images in dataset 3.

Table 4.1: Descriptive Statics

Classifier	Accuracy	Sensitivity	Specificity	Precision
SVM	92.3%	90.8%	93.4%	91.7%
K-nn	88.7%	86.2%	87.8%	81.9%

V. Conclusion

According to the stage discovery of the abnormality cells in the lungs, lung abnormality is the most dangerous and pervasive disease in the world. This suggests that the process of detecting this disease is crucial to preventing severe stages and reducing its global percentage distribution. Globally, lung cancer is the leading cause of death from disease. The death rate from lung ca ncer is extremely high. There are several types of cancerous tumors, including breast and lung cancer. Successful treatment of cancer depends on early identification. Images from Computed Tomography (CT) are used to make the diagnosis. To determine if a patient's state is normal or abnormal in its early stages, this Histogram Equalization will pre-process the images and have an extraction method and classifier. Image processing methods have been used extensively recently in a number of medical fields to improve images at earlier stages of detection and treatment, where time is crucial to identifying abnormalities in target images, particularly in tumors with a variety of abnormalities like breast and lung abnormalities. The main components of this study are image quality and accuracy; evaluation of image quality and improvement rely on the enhancement stage, which employs minimal pre-processing methods based on the Gabor filter inside Gaussian rules.

REFERENCES

[1] Armato, S. G., et al. (2012). "The Lung Image Database Consortium (LIDC) and Image Database Resource Initiative (IDRI): a completed reference database of lung nodules on CT scans." *Medical Physics*, 39(2), 915–931.

[2] Gurcan, M. N., et al. (2009). "Lung nodule detection in CT images: A review." *IEEE Reviews in Biomedical Engineering*, 2, 94–108.

[3] Kumar, S., & Das, P. (2016). "Lung cancer detection using image processing techniques." *Procedia Computer Science*, 85, 708–714.

[4] Sharma, R., Verma, A., & Singh, S. (2019). "Detection of lung cancer using image processing techniques." *International Journal of Computer Applications*, 178(7), 23–27.

[5] MathWorks. (2023). "Image Processing Toolbox." Retrieved from https://www.mathworks.com/products/image.html

[6] Disha Sharma, Gagandeep Jindal, “Computer Aided Diagnosis System for Detection of Lung Cancer in CT Scan Images”, International Journal of Computer and Electrical Engineering, Vol. 3, No. 5, October 2011

[7] P. Bountris, E. Farantatos, and N. Apostolou, “Advanced Image Analysis Tools Development for the Early Stage Bronchial Cancer Detection”, World Academy of Science,

Engineering and Technology 9 2005

[8] Zhi-Hua Zhou, Yuan Jiang, Yu-Bin Yang, Shi-Fu Chen, “Lung Cancer Cell Identification Based on Artificial Neural Network Ensembles”, Artificial Ingelligence in Medicine, 2002, vol.24, no.1, pp.25-36.

[9] Samir Kumar Bandyopadhyay, “Edge Detection from CT Images of Lung”, International Journal Of Engineering Science & Advanced Technology Volume - 2, Issue - 1, pg: 34 – 37, Jan-Feb 2012.