Mini Project Report on

LUNG CANCER DETECTION USING IMAGE PROCESSING

Submitted in partial fulfillment of the requirement for the award of the degree of

BACHELOR OF TECHNOLOGY

IN

COMPUTER SCIENCE & ENGINEERING

Submitted by:

Student Name	University Roll No.
Utkarsh Shukla	2021765

Under the Mentorship of

Dr. Manoj Diwakar
Professor
Computer Science and Engineering
Graphic Era Deemed to be University Dehradun

Department of Computer Science and Engineering
Graphic Era (Deemed to be University)
Dehradun, Uttarakhand
January-2024

CANDIDATE'S DECLARATION

I hereby certify that the work which is being presented in the project report entitled **"Lung Cancer Detection using Image Processing "** in partial fulfillment of the requirements for the award of the Degree of Bachelor of Technology in Computer Science and Engineering of the Graphic Era (Deemed to be University), Dehradun shall be carried out by the under the mentorship of **Dr. Manoj Diwakar, Professor**, Department of Computer Science and Engineering, Graphic Era (Deemed to be University), Dehradun.

Name-Utkarsh Shukla University Roll no-2021765

Table of Contents

Chapter No.	Description	Page No.
Chapter 1	Introduction	1-2
Chapter 2	Literature Survey	**3-4**
Chapter 3	Methodology	**4-5**
Chapter 4	Result and Discussion	**5**
Chapter 5	Conclusion and Future Work	**6-7**
	References	**7-8**

Chapter 1

Introduction

In the following sections, a brief introduction and the problem statement for the work has been included.

1.1 Introduction

Lung cancer remains one of the leading causes of cancer-related deaths worldwide, accounting for approximately 18% of all cancer deaths. Despite advancements in medical treatments and interventions, the survival rate for lung cancer patients remains alarmingly low, primarily due to late-stage diagnosis. Early detection of lung cancer significantly enhances the prospects for successful treatment and improves patient survival rates. Consequently, there is a pressing need for reliable, accurate, and non-invasive diagnostic methods to detect lung cancer at its nascent stages.

In recent years, image processing has emerged as a pivotal technology in the field of medical diagnostics, offering promising solutions for the early detection of various diseases, including lung cancer. Image processing techniques involve the manipulation and analysis of medical images to extract meaningful information that can aid in the diagnosis and treatment of diseases. These techniques leverage advanced algorithms and computational methods to enhance image quality, identify patterns, and detect abnormalities that may be indicative of lung cancer.

The application of image processing in lung cancer detection involves several stages, including image acquisition, pre-processing, segmentation, feature extraction, and classification. High-resolution imaging modalities such as computed tomography (CT) scans and chest X-rays provide detailed visual representations of the lungs, which are then processed to identify potential cancerous lesions. Pre-processing techniques are employed to enhance image quality by reducing noise and improving contrast, facilitating more accurate analysis. Segmentation methods are used to isolate regions of interest, such as nodules or tumors, from the surrounding tissues. Subsequently, feature extraction algorithms analyze these regions to identify specific characteristics, such as shape, size, and texture, which may suggest malignancy. Finally, classification algorithms, often powered by machine learning and artificial intelligence, are applied to differentiate between benign and malignant lesions, providing a diagnostic outcome.

This paper explores the advancements in image processing techniques for lung cancer detection, highlighting the efficacy and accuracy of these methods. We review recent studies and innovations in the field, discuss the challenges associated with image-based lung cancer diagnostics, and propose potential solutions to overcome these hurdles. By leveraging state-of-the-art image processing technologies, we aim to contribute to the ongoing efforts to improve early lung cancer detection and ultimately enhance patient outcomes.

1.2 Definition of Image Processing

Image processing is a field of computer science and engineering that focuses on the manipulation, analysis, and interpretation of digital images. It involves several stages:

1.) Image Acquisition: Capturing images using various modalities like X-rays, CT scans, MRI, and ultrasound.
2.) Pre-Processing: Enhancing image quality through noise reduction, contrast adjustment, and normalization.
3.) Segmentation: Partitioning the image into distinct regions or objects of interest.
4.) Feature Extraction: Quantifying attributes such as shape, size, texture, and intensity from segmented regions.
5.) Classification: Using machine learning and AI to label or diagnose based on extracted features.

In medical imaging, image processing enhances diagnostic accuracy and efficiency, playing a crucial role in early disease detection and treatment planning.

1.3 Stages of Image Processing

Image Processing Techniques have various Layers such as Acquisition, Preprocessing, Segmentation, Recognition, Feature Extraction and Interpretation.

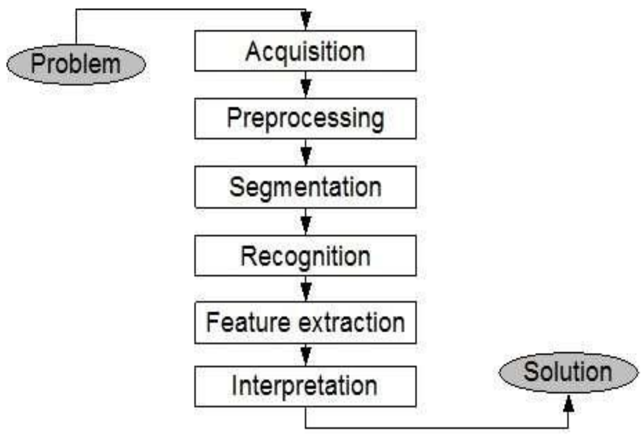

Fig 1.1- Stages of Image Processing System

Chapter 2

Literature Survey

Several experimenters has proposed and enforced discovery of lung cancer using different approaches of image processing and machine literacy. Aggarwal, Furquan and Kalra(4) proposed a model that provides bracket between nodes and normal lung deconstruction structure. Although the system detects the cancer bump, its delicacy is still inferior. No any machine literacy ways has been used to classify and simple segmentation ways is used. thus, a combination of any of its ways in our new model doesn't give the probability of enhancement. Roy, Sirohi, and Patle(6) developed a system to descry lung cancer bump using a fuzzy interference system and active figure model. This system uses gray transformation for image distinction enhancement. Image binarization is performed before segmentation and the reacted image is segmented using an active figure model. Cancer type is performed using a fuzzy conclusion system. Features like area, mean, entropy, correlation, major axis length, and minor axis length are pulled to train the classifier. Overall, the delicacy of the system is 94.12. Counting its limitations it does not classify the cancer as benign or nasty which is the future compass of this proposed model.

Gonzalez and Ponomaryvo(7) proposed a system that classifies lung cancer as benign or nasty. The system uses the priori information and HousefieldUnit(HU) to calculate the Region of Interest(ROI). Shape features like area, curiosity, circularity, fractal dimension, and textural features like mean, disunion, energy, entropy, skewness, distinction, and smoothness are pulled to train and classify the support vector machine to identify whether the bump is benign or nasty. The advantage of this model is that classifies cancer as benign or nasty still the limitation of it is that former information is demanded about the region of interest. The model's type of benign or nasty using a support vector machine can be useful in our new model.

Jin, Zhang and Jin(8) used a complication neural network as classifier in his CAD system to descry lung cancer. The system has 84.6 of delicacy,82.5 of perceptivity, and 86.7 of particularity. The advantage of this model is that it uses indirect sludge in the Region of interest(ROI) birth phase which reduces the cost of training and recognition way. Although perpetration cost is reduced, it has still the wrong delicacy.

It shows the comparison with other preliminarily proposed models on colorful factors.

Chapter 3

Methodology

3.1 Overview

The following figure represents the block diagram for the proposed model of Lung Cancer detection using Image Processing.

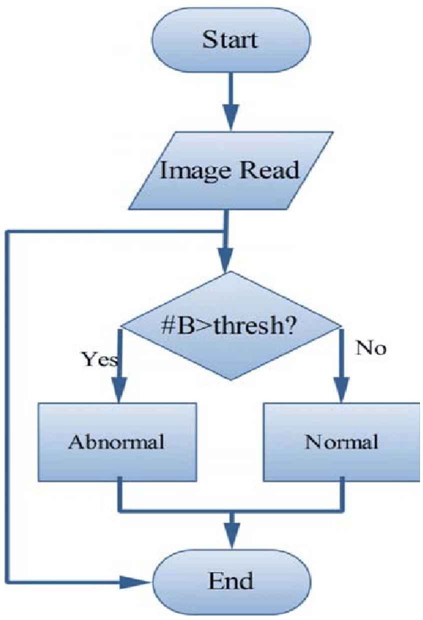

Fig-3.1 Block diagram for the program

3.2 Steps Involved in Data Preprocessing

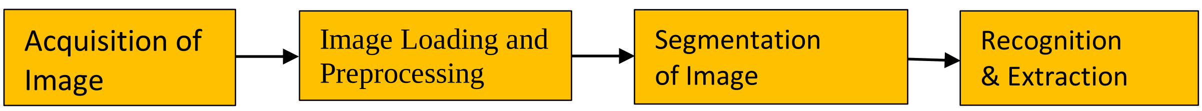

Fig-3.2 Diagram representing Input and Output Cycle.

1)**Acquisition of Image-** The dataset is loaded from a zip file to the working directory and divided into normal, abnormal, and test sets, maintaining the directory structure.

2)Image loading and Preprocessing-

- Images loaded using OpenCV 'cv2.imread(path_of_the_Directory)' function.
- Loaded with parameters ensuring correct labeling, RGB color mode, and resizing to high-revolution pixels.

3)Segmentation of Image- The image is Divided into various Segments, to simplify its representation and make it more meaningful. It is Commonly used in OpenCV to descry objects and boundaries in images.

4)Recognition and Extraction- The Abnormalities are Recognized by the model and extracted. Now based on the Comparison rate, we check if the CT image is Normal or Abnormal.

Chapter 4

Result and Discussion

4.1 Results

The image-processing stage starts with image improvement; the end of image improvement is to ameliorate the interpretability or perception of information included in the image for mortal observers or to give better input for other automated image-processing ways.

Image Improvement ways can be divided into two broad orders Spatial sphere styles and frequency sphere styles. Unfortunately, there's no general proposition for determining what "good" image improvement is when it comes to mortal perception. However, it's good, If it looks good. still, when image improvement ways are used as pre-processing tools for other image processing ways, the quantitative measures can determine which ways are most applicable.

Chapter 5

Conclusion and Future Work

5.1 Conclusion

An image enhancement fashion is developing for earlier complaint discovery and treatment stages; the time factor was taken into account to discover the abnormality issues in target images. Image quality and delicacy are the core factors of this exploration, image quality assessment as well as improvement stage where were espoused on lower-processing ways. The proposed fashion is effective for segmentation principles to be a region of interest foundation for point birth carrying. The proposed fashion gives veritably promising results compared with other habituated ways. counting on general features, a normal comparison is made. The main detected features for accurate image comparison are pixel chance and mask-labeling with high delicacy and robust operation.

5.2 Future Work

The landscape of lung cancer detection through image processing continues to evolve, driven by advancements in computational methods, imaging technologies, and interdisciplinary collaborations. Future research in this domain promises to push the boundaries of what is currently achievable, paving the way for more accurate, efficient, and early detection methodologies. Several key areas hold particular promise for future exploration: **Multi-scale and Multi-modal Approaches:** Integrating multi-scale and multi-modal information, such as infrared or hyperspectral imaging, could improve disease detection accuracy and robustness across different environmental conditions.

1.) Real-Time Image Processing: The development of real-time image processing capabilities can significantly enhance the diagnostic workflow. Implementing advanced algorithms that can process and analyze images instantaneously during imaging procedures can assist radiologists in making immediate, informed decisions, potentially improving patient outcomes through timely interventions.

2.) Cloud-Based Collaborative Platforms: Developing cloud-based platforms for collaborative research and clinical practice can facilitate the sharing of large datasets, computational resources, and expertise across institutions. Such platforms can accelerate the development and validation of novel image processing algorithms and promote the adoption of best practices in lung cancer detection.

3.) Deep Learning and Explainable AI: While deep learning has already shown substantial promise in image processing, future work should focus on making these models more interpretable and transparent. Explainable AI techniques can help clinicians understand the rationale behind AI-driven diagnostic decisions, increasing trust and facilitating the integration of AI into clinical practice.

References

[1] Makaju, S., Prasad, P. W. C., Alsadoon, A., Singh, A. K., & Elchouemi, A. (2018). Lung cancer detection using CT scan images. *Procedia Computer Science*, *125*, 107-114.

[2] Sharma, D., & Jindal, G. (2011). Identifying lung cancer using image processing techniques. In *International Conference on Computational Techniques and Artificial Intelligence (ICCTAI)* (Vol. 17, pp. 872-880). Citeseer.

[3] Al-Tarawneh, M. S. (2012). Lung cancer detection using image processing techniques. *Leonardo electronic journal of practices and technologies*, *11*(21), 147-58.

[4] Aggarwal, T., Furqan, A., & Kalra, K. (2015) "Feature extraction and LDA based classification of lung nodules in chest CT scan images." 2015 International Conference On Advances In Computing, Communications And Informatics (ICACCI), DOI: 10.1109/ICACCI.2015.7275773.

[5] Makaju, S., Prasad, P. W. C., Alsadoon, A., Singh, A. K., & Elchouemi, A. (2018). Lung cancer detection using CT scan images. *Procedia Computer Science*, *125*, 107-114.

[6] Roy, T., Sirohi, N., & Patle, A. (2015) "Classification of lung image and nodule detection using fuzzy inference system." International Conference On Computing, Communication & Automation. DOI: 10.1109/CCAA.2015.7148560.

[7] Rendon-Gonzalez, E., & Ponomaryov, V. (2016) "Automatic Lung nodule segmentation and classification in CT images based on SVM." 2016 9Th International Kharkiv Symposium On Physics And Engineering Of Microwaves, Millimeter And Submillimeter Waves (MSMW). DOI: 10.1109/MSMW.2016.7537995.

[8] Jin, X., Zhang, Y., & Jin, Q. (2016) "Pulmonary Nodule Detection Based on CT Images Using Convolution Neural Network." 2016 9Th International Symposium On Computational Intelligence And Design (ISCID). DOI: 10.1109/ISCID.2016.1053.

[9] Open Source Computer Vision (OpenCV) [Online]. Accessed on 01st June 2024: http://opencv.willowgarage.com/wiki/ **(Example : Website)**

[10] Kaggle [Online]. Used for Dataset of CT images. Accessed on 09th June 2024:
Find Open Datasets and Machine Learning Projects | Kaggle (**Example : Website**)