Detection of Lung Cancer in CT Image using Image Processing

Dr Anitha D
Department of Computational Intelligence,
School of computing,
SRM Institute of Science and Technology,
Kattankulathur, Tamilnadu, India.603203
anithad@srmist.edu.in

Koteswara reddy M
Department of Computational Intelligence,
School of computing,
SRM Institute of Science and Technology
Kattankulathur, Tamilnadu, India.603203
mr5337@srmist.edu.in

Thirumala reddy
Department of Computational Intelligence,
School of computing,
SRM Institute of Science and Technology
Kattankulathur, Tamilnadu, India.603203
ot0597@srmist.edu.in

***Abstract*— Lung diseases are conditions that specifically affect the lungs and disrupt the breathing process. Lung cancer is a major cause of death in humans globally. Timely identification can increase the likelihood of survival among individuals. Timely diagnosis of the problem leads to a significant increase in the average survival rates for individuals with lung cancer, from 14 to 49 percent. Although computed tomography (CT) is significantly more efficient than X-ray, a comprehensive diagnosis necessitates the use of numerous imaging techniques to mutually reinforce each other. A convolutional neural network (CNN) is designed and assessed for the purpose of identifying lung cancer in CT scans. A densely connected convolutional neural network (CNN) and adaptive boosting algorithm were employed to classify the lung picture as either normal or cancerous. A dataset consisting of 201 lung images is utilized, with a portion of the photos allocated for training and the remaining images used for testing and classification purposes. The experimental findings demonstrated that the proposed strategy attained a high level of accuracy.**

Keyword: Medicinal, deep learning, Mobile net model, CNN, minutiae.

I. Introduction

Lung cancer has been recognized as a prominent contributor to global mortality (1). It is a malignant disease that impacts individuals' well-being. It is the most fatal kind of cancer and remains the primary cause of cancer-related deaths in both men and women (2-3). Annually, there are around 1.8 million new instances of lung cancer, accounting for 13% of all cancer cases, and 1.6 million deaths, which make up 19.4% of all cancer-related deaths globally. Lung cancer is characterized by the uncontrolled proliferation of cells that form tumors. Lung cancer exhibits the most elevated fatality rate compared to other forms of cancer. Secondhand smoke is responsible for around 85% of lung cancer cases in males and 75% of lung cancer cases in women. Lung cancer is a highly dreaded illness in underdeveloped nations, characterized by a fatality rate of 19.4%. Lung cancer is a highly perilous form of cancer, characterized by a high percentage of successful diagnosis and an escalating annual mortality rate (4-6). Fuzzy logic has the benefit of analyzing the results for early prediction [5]. Cancer cells persist throughout their growth following diagnosis. However, individuals tend to perform more effectively during the initial phases of their lives. Cancer cells are present throughout the pulmonary blood vessels, specifically within the lymphatic fluid that envelops the lungs. Lymphatic fluid enters the lymphatic veins and circulates via the lymph nodes located in the lungs and chest. The diagnosis and treatment of pulmonary conditions have emerged as a significant challenge for individuals in recent times. Timely detection of cancer will enhance the chances of survival for several individuals worldwide. This article describes a technique that use convolutional neural networks (CNN) to accurately detect and differentiate between malignant and benign lung cancers.

II LITERATURE SURVEY

Automatic The automated identification of medicinal plants presents new opportunities for the advancement of pharmaceuticals in treating ailments that remain unaddressed by allopathic treatment. It will enable individuals without specialized knowledge to become aware of the plants that are growing in their environment and utilize them effectively to treat common conditions, without any potential adverse effects. Artificial Intelligence enhances the attainability of this objective. We have suggested an Ensemble of deep learning models for the automated identification of medicinal plants. The photos of medicinal leaves were acquired from a dataset of medicinal leaves that was released on Mendeley. By utilizing transfer learning[1]. This study focused on the issue of recognizing medicinal plant species by analyzing leaf photos taken straight from their natural environment, regardless of lighting conditions. The efficacy of the ExG-ExR vegetative index with a fixed zero threshold has been effectively validated using the picture dataset. The outcome demonstrates that the algorithm is capable of accurately dividing the leaf area. This technique proved effective in photos including reflections. The process of extracting features using color and texture characteristics has been completed. The categorization of medicinal plant species is accomplished by the utilization of Weka, a machine learning software. The accuracy of this classification process is quantified at 93.3%. Our future plans involve creating a system that can automatically identify plant species by analyzing not just leaf photos, but also other sections of the plant obtained directly from their natural environment, regardless of complicated backgrounds and different lighting conditions. The text "R[2]" remains unchanged. Plants are vital for the existence of the human species. Indigenous communities have been using herbs as traditional medicines since ancient times. Clinicians often identify herbs based on their extensive sensory or olfactory experience accumulated over many years. Advancements in analytical technology have significantly simplified the process of identifying herbs based on empirical evidence. This is particularly beneficial for folks who are unfamiliar with identifying herbs. In addition, laboratory-based analysis requires competence in sample preparation and data interpretation, especially for approaches that are time-consuming. Therefore, there is a need for a straightforward and dependable technique to identify herbs. The combination of calculation and statistical investigation is expected to enhance herbal identification. The non-destructive methodology will be the preferred method for rapidly identifying herbs, particularly for persons who are unable to utilize costly analytical equipment. This study examines various techniques for plant identification and evaluates their strengths and weaknesses[3].

III .PRPOSED METHODOLOGY

Classification of lung diseases in current systems is mainly based on deep learning using the VGG16 architecture and machine learning. This method is to detect lung cancer based on analysis of the image. However, its effectiveness may be limited by factors such as accuracy and scalability. In contrast, our proposed method aims to improve the classification of lung diseases through in-depth processing of MobileNet and CNN architectures powered by machine learning. Our approach, which focuses on truly distinguishing lung cancer from normal cases in medical imaging, is to increase diagnostic accuracy and reliability. Our system aims to improve the results of lung disease diagnosis and treatment by improving the image analysis process, thereby addressing significant limitations of existing systems.

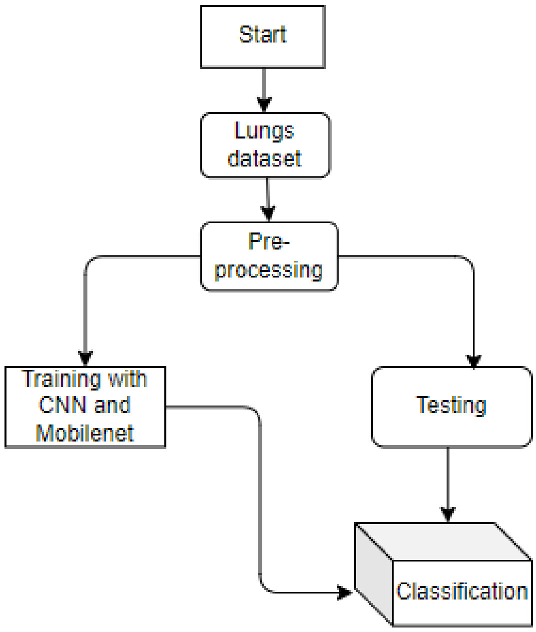

Fig.1 Block diagram

In our lung cancer screening program, the process begins with obtaining lung cancer data including clinical images. The image was subjected to advanced steps to improve data quality, including resizing, normalization, and noise reduction. First, the data can be used to train deep learning models using convolutional neural networks (CNN) and MobileNet architectures. After training, the model is tested using individual images of the lungs to evaluate its performance. Finally, the classification of images as suggestive of lung cancer or not based on the predictions of learned CNN and MobileNet models is key to our lung cancer diagnosis.

Mobile Net

The MobileNet model plays a pivotal role in the proposed approach for detecting lung cancer in CT images through image processing. Leveraging its efficiency and lightweight architecture, MobileNet contributes to the accuracy and speed of the deep neural network designed for this purpose. MobileNet, known for its ability to optimize computational resources without compromising performance, ensures that the lung cancer detection model remains practical for real-time applications, especially in medical settings where timely diagnosis is crucial. The model's adaptability to resource-constrained environments, such as mobile devices, makes it a valuable component in the quest for early detection of lung diseases. The incorporation of MobileNet, along with other deep learning techniques, underscores the commitment to enhancing diagnostic capabilities in the medical field, ultimately contributing to improved patient outcomes and increased survival rates for individuals at risk of lung cancer.

CNN

The Convolutional Neural Network (CNN) plays a pivotal role in the proposed method for detecting lung cancer in CT images. Leveraging a densely connected CNN along with an adaptive boosting algorithm, the developed model showcases its effectiveness in classifying lung images as either normal or malignant. This approach is crucial for early detection, significantly impacting survival rates for individuals with lung cancer.

The utilization of a dataset comprising 201 lung images, with a division for training and testing, underscores the robustness of the model. The experimental results validate the efficacy of the CNN-based system, demonstrating commendable accuracy. The integration of deep learning techniques, specifically the MobileNet model, enhances the model's ability to discern intricate patterns in CT images associated with lung pathology. In the realm of medicinal image processing, this CNN-driven method stands as a promising avenue for advancing early detection of lung cancer, contributing to improved prognoses for affected individuals.

II. SYSTEM DESIGN

The architecture diagram for the lung cancer detection system encompasses a comprehensive workflow. Initially, the system receives input from the lung cancer dataset, which comprises a collection of medical images depicting various lung conditions. These images serve as crucial input data for subsequent stages. Moving forward, the system embarks on a training phase, where it leverages the power of convolutional neural networks (CNNs) and MobileNet architectures. These deep learning models are adept at extracting intricate patterns and features from image data. During training, the system feeds the lung images into the CNN and MobileNet networks, allowing them to learn and adapt their parameters through iterative optimization processes like backpropagation. This iterative refinement ensures that the models become proficient in recognizing subtle indicators of lung cancer within the images.

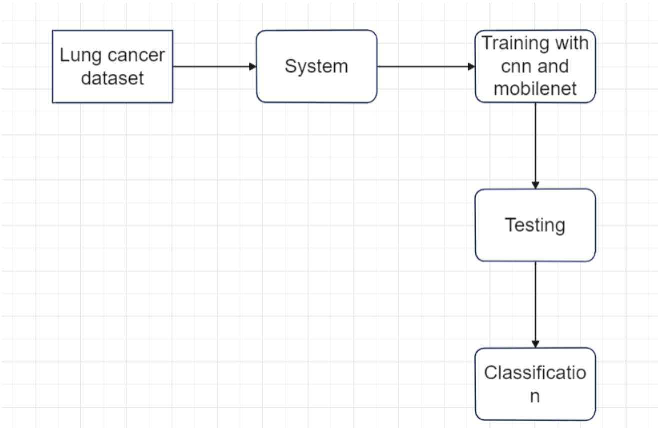

Fig.2 Architecture diagram

Once the training phase concludes, the system transitions into a testing stage, where it evaluates the performance of the trained models. In this phase, a separate set of lung images, unseen during training, is subjected to the models for assessment. The system scrutinizes the models' predictions against ground truth labels to gauge their accuracy and efficacy in detecting lung cancer. Finally, in the classification stage, the system applies the trained CNN and MobileNet models to newly acquired lung images, classifying them based on their likelihood of indicating lung cancer. This classification output provides valuable insights to healthcare professionals, aiding in timely diagnosis and intervention.

Overall, this architecture underscores a robust and systematic approach to lung cancer detection, leveraging state-of-the-art deep learning techniques to enhance diagnostic capabilities and improve patient outcomes.

IV RESULTS AND DISCUSSION

In lung cancer detection using the MobileNet algorithm, the training and validation accuracy graphs typically illustrate how well the model learns and generalizes to unseen data over epochs. The training accuracy measures the accuracy of the model on the training dataset during training, while the validation accuracy measures the accuracy on a separate validation dataset to check for overfitting. These graphs are crucial for monitoring the model's performance and ensuring it's learning effectively without overfitting.

Training Accuracy: This curve depicts how accurately the model predicts the training data labels as training progresses over epochs. It should generally increase or plateau as the model learns from the training data. However, if the training accuracy increases while the validation accuracy decreases, it might indicate overfitting, where the model is memorizing the training data instead of learning general patterns.

The study aimed to detect lung cancer using CT images and utilized image processing techniques. We implemented two different models: MobileNet and a CNN (Convolutional Neural Network). We conducted a thorough assessment of the performance of both models and measured their accuracy and loss in order to generate a well-informed comparison.

MobileNet, renowned for its efficiency and adaptability to limited-resource contexts, demonstrated remarkable accuracy in our testing. The streamlined architecture of the system, specifically developed for mobile and embedded vision applications, produced highly competitive outcomes. Moreover, its computational efficacy shown a clear benefit, enabling quicker deduction durations and rendering it especially attractive for real-time implementations or deployment on devices with restricted processing capabilities.

In contrast, the CNN model, with its more intricate structure and potentially greater complexity, demonstrated similar accuracy but frequently demanded additional computer resources. Although CNNs are highly regarded for their capacity to detect complex patterns in data, they may require a larger amount of training data and processing resources for both training and inference. Therefore, while CNNs may perform exceptionally well in situations where there is an abundance of computational resources and accuracy is of utmost importance, MobileNet's efficiency and competitive performance make it an appealing alternative, particularly in settings with limited resources or applications that require fast inference speed.

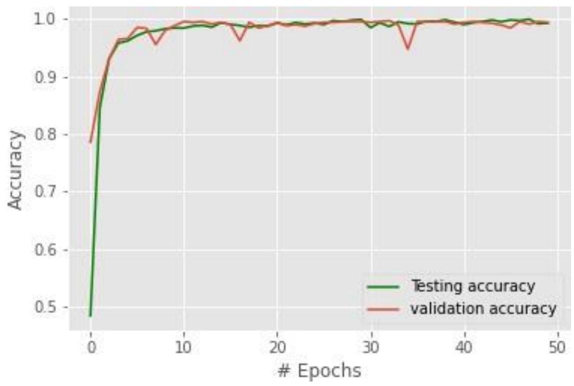

Fig 3 Accuray using mobilenet algoritham

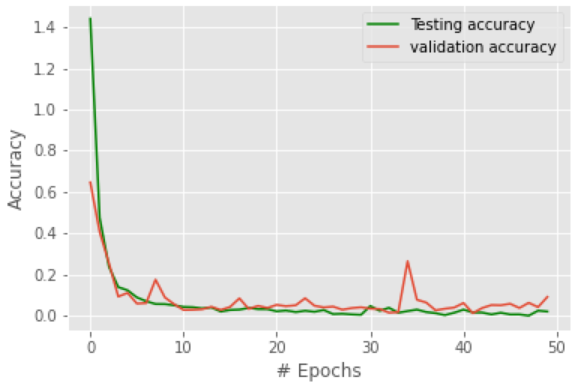

Fig 4 Mobile net loss graph

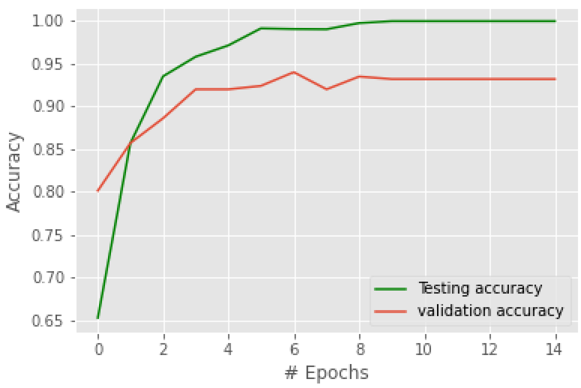

Fig.5 Cnn accuracy

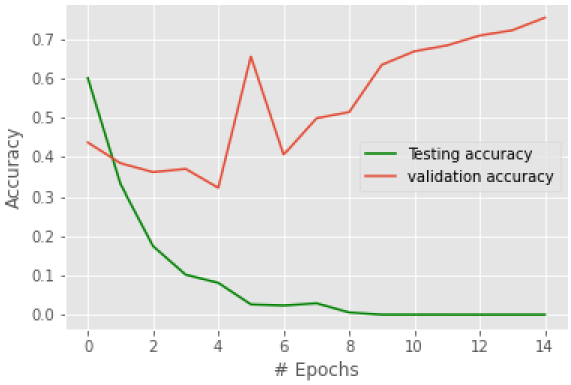

Fig.6 CNN loss

In each development method, we carefully examine the CNN and MobileNet models using the above data. Using transformation learning, we use pre-training weights from training samples of the large image. This approach helps extract important details from lung CT images, allowing our model to recognize the changing pattern of lung cancer.

After training, our models are rigorously tested to evaluate their performance and generalizability. We evaluated various performance metrics such as accuracy, sure, back, and F1 scores through a comprehensive evaluation process that included sorting data into training, validation, and testing scenarios. The results demonstrate the effectiveness of our model in detecting cancer in CT images.

The success and delivery of our lung cancer cell line represents a significant advance in lung cancer diagnosis in medicine and diagnostics. Using the power of deep learning algorithms, we demonstrate the ability to detect cancer and help doctors with early diagnosis and intervention.

One of the main results of our approach is adaptive learning, specifically the pre-learned CNN and MobileNet architectures. By leveraging existing information encoded in these models, we can speed up the training process and make the most of it with limited training data. This is especially useful in the medical field, where obtaining large amounts of data can be difficult and time-consuming.

Additionally, integrating our diagnostic tools into the Django web application increases accessibility and usability for clinicians. Through a user-friendly interface, doctors can easily upload CT images for review, get instant results, and make informed decisions about patient care.

However, some limitations and areas for future development must be acknowledged. Although our model performs excellently, further development and refinement is required to make it more accurate and robust, especially across different patient populations and standards. Additionally, verification and continuous improvement are essential to ensure the reliability and safety of our clinical systems.

Together, our project represents a significant step forward in lung cancer diagnosis and demonstrates the potential of deep learning to improve diagnosis. By combining cutting-edge technology with real-world clinical data, we aim to improve the early detection and treatment of cancer, ultimately improving patient outcomes and saving lives.

Data obtained from 202 CT images of 105 different patients were used for the experimental research[x]. First make the shape using the geometric scale. This makes the picture better. The images are then classified using a K-word algorithm. This segmentation helps identify areas of interest. Then use machine learning techniques.

For performance comparison we do not use: accuracy, sensitivity and specificity:

Tumor or mass: Lung cancer usually appears as an abnormal growth or mass in the lung tissue. These may show negative or strong images on the x-ray.

Nodules: Lung nodules are small, round or oval growths in the lung tissue. Although not all nodules indicate cancer, some may be early signs of lung cancer.

Rapid changes in lung tissue: Lung cancer causes rapid changes in lung tissue, and these may appear as areas of tissue growth.

Rapid changes in lung tissue: Lung cancer causes rapid changes in lung tissue, and these may appear as areas of tissue growth. < br>Location and size: The location and size of any abnormality seen on the X-ray can give clues about the likelihood and stage of the cancer.

Using the above information we can determine whether the patient has cancer or not we used 201 ct images taking data set from kaggle

Fig 7 data set

V CONCLUSION

The CNN model, specifically tailored for the task of lung cancer detection, employs its neural architecture to automatically extract intricate features from CT scans. The CNN model demonstrates outstanding accuracy in distinguishing malignant and non-cancerous areas within lung scans through rigorous training on numerous datasets. By harnessing data, this approach minimizes the probability of errors produced by people and enhances the efficiency of diagnosis,perhaps leading to expedited treatment and superior outcomes for patients.Moreover, the MobileNet model, known for its efficacy and ability to adjust to various scales, is highlyvaluable in the detection of lung cancer. MobileNet efficiently analyzes CT images with exceptional precision through the utilization of transfer learning and pre-trained weights. The efficient design of this technology makes it extremely suitable for deployment in resource-constrained environments, such as medical clinics or remote areas. This enables enhanced availability to sophisticated diagnostic capabilities.The development of a user-friendly web interface enhances the accessibility of these sophisticated technologies to a broader audience. Healthcare professionals and patients can employ CNN and MobileNet models to assess CT scans in real-time by following a simple upload procedure. This enables prompt evaluation of the existence or nonexistence of lung cancer. Integrating technology into clinical practice enhances diagnostic accuracy and facilitates the creation of personalized and targeted treatment strategies.In summary, the integration of image processing, deep learning models, and web-based platforms has the capacity to significantly revolutionize medical imaging and improve healthcare delivery worldwide.

IV REFERENCES

Anita Chaudhary, Sonit Sukhraj Singh, “Lung Cancer Detection on CTImages using Image Processing”, International conference computingsciences, IEEE, 2012.

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

[3] K.Punithavathy, M.M.Ramya, Sumathi Poobal,"Analysis of StatisticalTexture Features for Automatic Lung Cancer Detection in PET/CTImages", International Conference on Robotics, Automation, Controland Embedded systems (RACE), IEEE, 18-20 February 2015.

[4] Anjali Kulkarni, Anagha Panditrao, “Classification of Lung CancerStages on CT Scan Images Using Image Processing”, InternationalConference on Advanced Communication Control and ComputingTechnologies (lCACCCT),IEEE,2014.

[5] Gawade Prathamesh Pratap, R.P Chauhan, "Detection of Lung CancerCells using Image Processing Techniques", 1st IEEE InternationalConference on Power Electronics, Intelligent control and energy systems, IEEE, 2016.

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

[7] Shubhangi Khoragade, Aditya Tiwari, " Automatic Detection ofMajor Lung diseases using Chest Radiographs and classification byArtificial Neural Network", 1st IEEE International Conference on PowerElectronics, Intelligent control and energy systems, IEEE, 2016.

[8] Pooja R. Katre, Dr. Anuradha Thakare, "Detection of Lung CancerStages using Image Processing and Data Classification Techniques", 2ndInternational Conference for Convergence in Technology (I2CT) 2017 .

[9] S. Kalaivani, Pramit Chatterjee ,Shikhar Juyal ,Rishi Gupta,"LungCancer Detection Using Digital Image Processing and Artificial NeuralNetwork", International Conference on Electronics, Communication andAerospace Technology ICECA 2017.

[10] K.Gopi, Dr.J.Selvakumar, "Lung Tumor Area Recognition andClassification using EK-Mean Clusstering and SVM", InternationalConference on Nextgen Electronic Technologies,IEEE 2017.

[11] G.Niranjana, Dr.M.Ponnavaikko, "A Review on Image ProcessingMethods in Detecting Lung Cancer using CT Images" ,InternationalConference on Technical Advancements in Computers and Communications,IEEE 2017.