IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 10, May 2025
Impact Factor: 7.67

Automated Lung Cancer Detection Via CNN

W. P. Rahane, Maruti Ogale, Pranali Pophale, Aditi Hulawale
Department of Information Technology Engineering
NBN Sinhgad Technical Institutes Campus, Pune, India

Abstract: *Lung cancer is one of the leading causes of cancer-related deaths worldwide, primarily due to its detection at advanced stages when treatment options are limited. Early diagnosis significantly increases the chances of successful treatment and patient survival. This project proposes an advanced system for the early detection of lung cancer using a combination of image processing and machine learning techniques. The system processes lung CT scan images by applying various preprocessing steps to eliminate noise and improve image quality. It then performs segmentation to accurately identify and isolate the lung region and any suspicious masses or nodules. Once the region of interest is extracted, feature extraction techniques are employed to gather critical data such as texture, shape, and intensity. These features are then fed into a machine learning classifier that is trained to distinguish between normal and abnormal tissues, as well as to identify malignant and benign tumors. The use of artificial intelligence in this context not only improves the precision and reliability of lung cancer detection but also assists radiologists in making more informed decisions. This system aims to reduce diagnostic errors, lower analysis time, and ultimately contribute to better clinical outcomes by enabling earlier and more accurate detection of lung cancer. ..*

Keywords: Lung cancer detection, CT scan, image processing, machine learning, feature extraction, image segmentation, tumor classification, early diagnosis, medical imaging, computer-aided diagnosis, artificial intelligence

I. INTRODUCTION

Lung cancer is one of the most aggressive and life-threatening forms of cancer, responsible for a significant percentage of cancer-related deaths worldwide. It typically originates in the tissues of the lungs, most commonly in the cells lining the air passages. The primary challenge in combating lung cancer lies in its detection — in many cases, symptoms do not manifest until the disease has progressed to an advanced stage, by which point treatment becomes less effective. Therefore, early and accurate detection plays a crucial role in improving patient survival rates and treatment outcomes.

II. PROPOSED SYSTEM

The proposed system will automatically detect and classify lung cancer from medical images based on the accuracy, scalability, and real-world applicability of the CNNs. It is designed to help radiologists and medical practitioners with an efficient and reliable tool for early diagnosis, which is a critical aspect in improving patient outcomes. The system is built using Python and uses advanced machine learning libraries such as TensorFlow, NumPy, and Matplotlib for developing and visualizing the model.

Key features include:

The system processes CT scans and X-ray images using CNN models, identifying lung nodules and categorizing them based on malignancy probability. Analysis of historical and current environmental data to provide accurate crop suggestions.

Advanced preprocessing techniques, including rescaling, cropping, and noise reduction, are employed to ensure high-quality input data. Augmentation methods like rotation, flipping, and scaling enhance model generalizability.

Provides detailed visualizations of training accuracy, validation accuracy, and loss curves to monitor model performance.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429 Volume 5, Issue 10, May 2025 Impact Factor: 7.67

ALGORITHM DESCRIPTION

Convolutional Neural Networks: CNNs are a class of deep learning algorithms specifically designed for image recognition and classification tasks. They automatically extract features from images, making them ideal for medical imaging analysis like lung cancer detection.

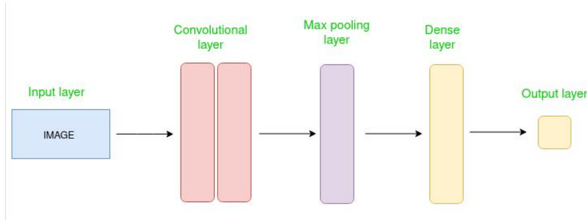

Figure 1 : Execution Process

Stochastic Gradient Decent(SGD): SGD is an optimization algorithm used to minimize the loss function during model training. It updates the model's weights iteratively based on the gradient of the loss function.

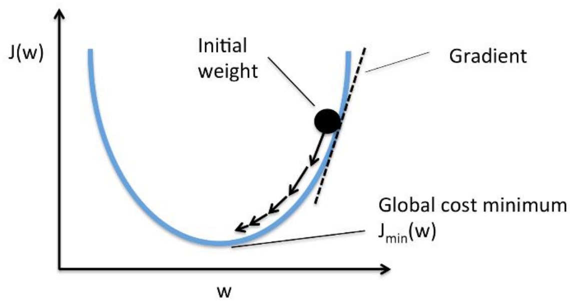

Figure 2 : SGD Algorithm

III. IMPLEMENTATION

The system was implemented using Python and powerful libraries such as ScikitLearn and TensorFlow. These libraries provide powerful tools for machine learning and data analysis, making them ideal for generating recommendations. Thi s application has four main components: data integration layer, machine learning engine, recommendation module.

Data Collection: Data collection involves gathering a large dataset of medical images, typically CT scans or X-ray images, containing both cancerous and non-cancerous lung samples. For your project, you might use publicly available datasets like the Lung Cancer Dataset or LIDC-IDRI (Lung Image Database Consortium and Image Database Resource Initiative), which contain annotated medical images for lung cancer detection. These datasets will serve as your primary data source for training, validation, and testing the model.

Feature Selection and Engineering: Feature selection is crucial for improving the model's performance by focusing on the most important features of the data. In the context of image classification, feature selection and engineering typically involve: Preprocessing, Augmentation.

Model Selection and Training: CNN model uses various layers which are available inbuilt. The model is trained using labeled images, employing backpropagation and gradient descent to minimize the loss function (e.g., categorical cross-entropy).

Deployment :Once the model is trained and optimized, it is ready for deployment. This involves integrating the trained model into a real-world application where it can make predictions on new, unseen lung scan images.

IJARSCT
ISSN: 2581-9429
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Volume 5, Issue 10, May 2025
Impact Factor: 7.67

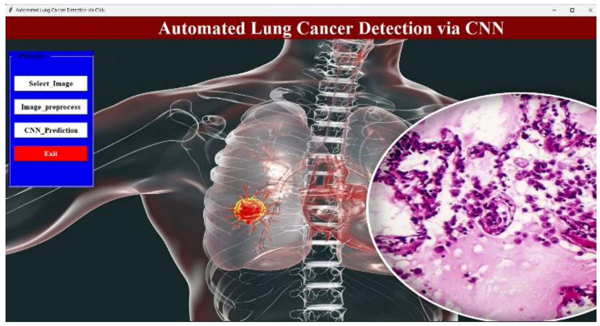

Figure 3 : Landing Page

V. ARCHITECTURE DIAGRAM

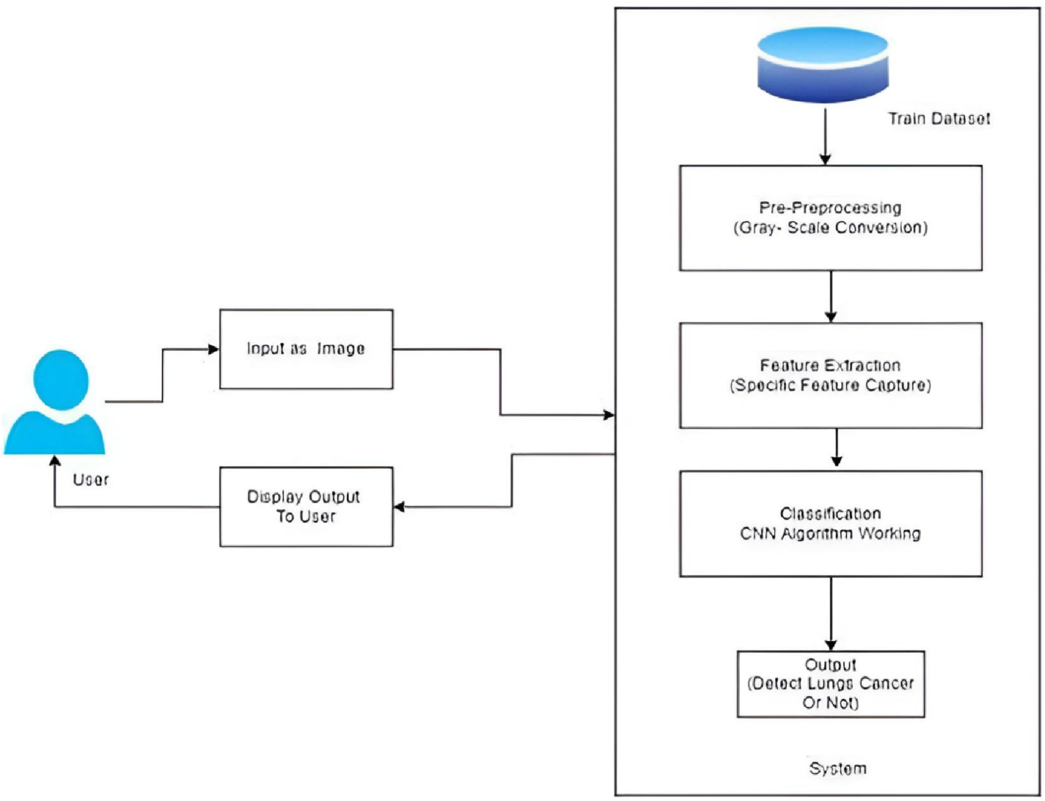

Figure 4 : System Architecture

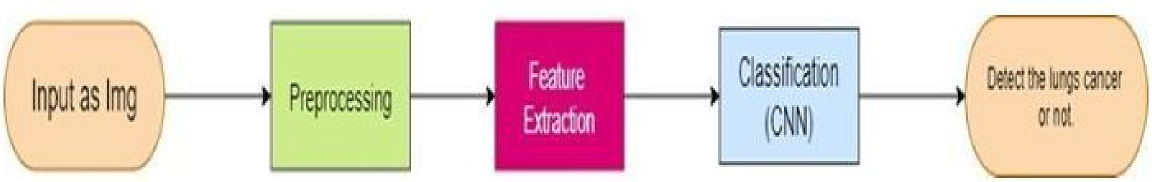

Figure 5 : Flow Diagram

VI. FUTURE SCOPE

Integration with Other Medical Imaging Modalities: The system can be expanded to incorporate additional imaging techniques such as MRI scans, PET scans, and X-rays, alongside CT scans. This would increase the robustness of the model, allowing it to detect lung cancer in various imaging contexts.

Multimodal Data Integration: Integrating the system with other types of patient data, such as genetic information, patient history, and lifestyle factors, could provide a more comprehensive and personalized approach to lung cancer detection and prognosis.

Enhanced Deep Learning Models: Future iterations of the system can explore more advanced deep learning architectures such as Transformer-based models, attention mechanisms, and reinforcement learning to improve the accuracy and efficiency of the system.

Collaborative Research and Databases: The project can be expanded to participate in global databases and collaborative research initiatives, where datasets from multiple hospitals or research institutions are pooled together, helping to train even more generalized and accurate models.

VII. CONCLUSION

In this project, we developed an approach for the early detection of lung cancer using image processing and machine learning techniques. The main objective was to identify cancerous lung tissues from CT scan images accurately and efficiently. Our system implements various preprocessing steps like noise removal and segmentation, followed by the extraction of important features that aid in the classification of malignant and benign nodules. The use of advanced algorithms helps in minimizing human error and enhances the reliability of the diagnosis, providing a vital support system for medical practitioners.

REFERENCES

[1] Shen, Y., Cheng, J., Ni, D., Tang, M., et al. (2021). Deep learning with convolutional neural network for lung cancer diagnosis based on CT images. IEEE Access.

[2] Kumar, A., Tripathi, A., & Sinha, M. (2023). Lung Cancer Detection using Deep Learning Models: A Comprehensive Review. Journal of King Saud University - Computer and Information Sciences.

[3] Wang, D., Khosravi, R., & Nahavandi, H. (2023). Automated end-to-end lung cancer detection using 3D deep learning. Computers in Biology and Medicine, 106547.

[4] Rajalakshmi, P., & Santhosh Kumar, K. (2024). An Efficient Lung Cancer Detection Model Using Deep Learning Techniques. Biomedical Signal Processing and Control, 105259.

[5] Saxena, S., Prasad, S.N., Polnaya, A.M., & Agarwala, S. (2025). Hybrid Deep Convolution Model for Lung Cancer Detection with Transfer Learning. arXiv preprint arXiv:2501.02785.

[6] Mamun, M., Mahmud, M.I., Meherin, M., & Abdelgawad, A. (2023). LCDctCNN: Lung Cancer Diagnosis of CT Scan Images Using CNN Based Model. arXiv preprint arXiv:2304.04814.

[7] Shuvo, S.B. (2023). An Automated End-to-End Deep Learning-Based Framework for Lung Cancer Diagnosis by Detecting and Classifying the Lung Nodules. arXiv preprint arXiv:2305.00046.