International Journal of Sustainable Development Through AI, ML and IoT

Volume 4 | Issue 1 | 2025
ISSN(Online):2584-0827
https://ijsdai.com/index.php/IJSDAI/index

Lung Disease Prediction System using Deep Learning with Grad- CAM based Interpretability

Nancy*[1], Preeti Gulia[2]
1,2 Department of Computer Science & Applications, Maharshi Dayanand University, Rohtak 124001, India
* nancypruthi2308@gmail.com, preeti@mdurohtak.ac.in
* Corresponding author: Nancy

ARTICLE INFO
Received: 01st June 2025
Revised: 17th June 2025
Accepted: 20th July 2025

ABSTRACT

Abstract— Lung diseases, particularly lung cancer, remain one of the leading causes of mortality worldwide. Early detection plays a crucial role in improving patient outcomes, but traditional diagnostic methods often face challenges related to accuracy, speed, and accessibility. This study presents a novel lung disease prediction system based on deep learning, utilizing Convolutional Neural Networks (CNNs) to classify chest CT scan images into 'Normal' and 'Lung Cancer' categories. The system incorporates Grad-CAM (Gradient-weighted Class Activation Mapping) for visual interpretability, enabling clinicians to better understand the regions in CT scans influencing the model's decision. A user-friendly web application is developed for real-time predictions, allowing users to upload CT images and receive instant results. The model, trained on a diverse set of datasets, achieves high accuracy, offering a valuable tool for early lung disease detection in clinical settings. This approach not only improves diagnostic accuracy but also enhances trust in AI-based medical systems by providing transparent visual explanations for model predictions. It achieved an accuracy of 97.50%, precision (0.98), recall (0.97), and F1-score (0.97).

Keywords— *Lung cancer, CT scan imaging, Deep Learning, CNN.*

1. Introduction

Lung diseases, particularly lung cancer, have become one of the most prevalent causes of mortality worldwide. According to the World Health Organization (WHO), lung cancer alone accounts for nearly 18% of all cancer deaths globally. Lung cancer is one of the most deadly and devastating types of cancer in the world. It is challenging to detect cancer, and its symptoms only become noticeable in the final stages. Although this cancer's death rate could be decreased by early detection and appropriate treatment for patients. Lung cancer often starts in the lungs; however, it occasionally appears as early symptoms prior to spread [1]. As a result, survival rates are low, with late-stage diagnosis making treatment significantly less effective.

Medical imaging, particularly chest X-rays and CT scans, plays a pivotal role in diagnosing lung diseases. However, these images require expert interpretation, which can be both time-consuming and prone to human error. It can be challenging for medical professionals to interpret and detect cancer from CT scan images. [2]. Moreover, in remote or under-resourced areas, the lack of trained professionals exacerbates the problem. This gap presents a compelling opportunity for artificial intelligence (AI), particularly machine learning (ML) and deep learning (DL) techniques, to assist in the automatic analysis of medical images and improve diagnostic accuracy.

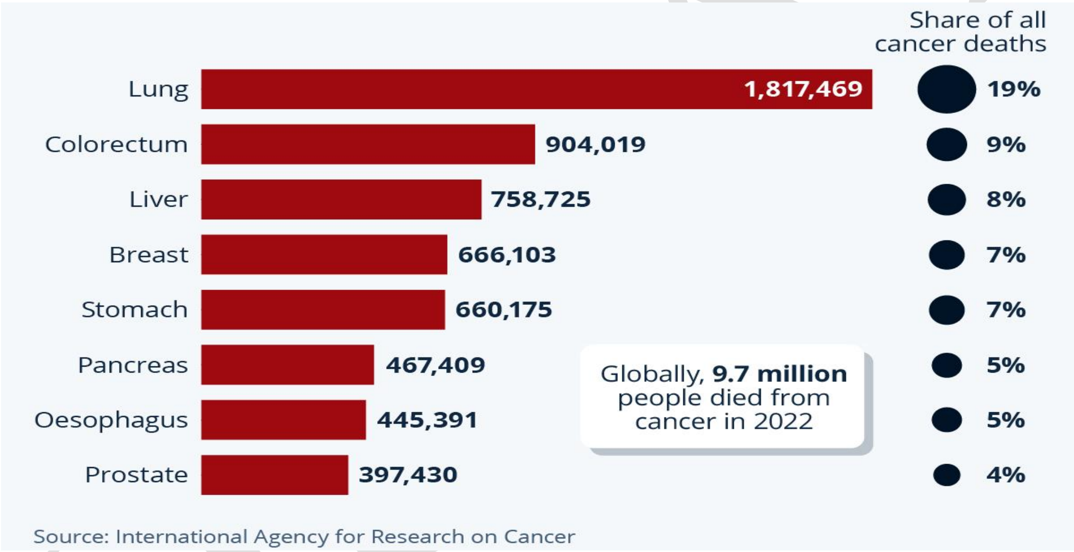


Figure 1: Cancer deaths in 2022 [3]

Machine Learning (ML) refers to a subset of AI that allows systems to learn from data and make predictions without explicit programming. In medical imaging, traditional ML models, such as Support Vector Machines (SVM), Decision Trees, and k-Nearest Neighbors (k-NN), have been employed to classify and analyze medical images. However, these models often struggle with complex, high-dimensional data like images, requiring extensive feature engineering and manual preprocessing to achieve satisfactory results.

In contrast, Deep Learning (DL), a subset of ML, has emerged as a powerful tool for automated medical image analysis. DL models, particularly Convolutional Neural Networks (CNNs), excel in extracting hierarchical features directly from raw image data, eliminating the need for manual feature extraction. These models have demonstrated remarkable performance in various applications, such as disease classification, segmentation, and anomaly detection, making them the go-to solution for modern medical image analysis.

CNNs, for instance, can automatically learn spatial hierarchies of features from images, such as edges, textures, and complex shapes, which makes them particularly effective for

tasks like detecting lung diseases in CT scans. DL techniques have also been leveraged to identify and classify lung cancer, pneumonia, tuberculosis, and other respiratory conditions with impressive accuracy.
The integration of ML and DL models for the prediction and classification of lung diseases offers a promising solution to the challenges in medical diagnostics. Traditional ML models often require handcrafted features and have limited capabilities when handling raw imaging data. Deep learning models, on the other hand, can automatically learn complex patterns from large datasets, making them more effective in real-world scenarios where accuracy is paramount.
This focuses on utilizing a Convolutional Neural Network (CNN)-based deep learning model for lung disease detection, specifically for classifying CT scan images as either 'Normal' or 'Lung Cancer.' The unique aspect of this study is the incorporation of Grad-CAM (Gradient-weighted Class Activation Mapping), a technique for improving model interpretability. Grad-CAM visualizes the regions of an image that most influenced the model's decision, providing transparency and trust in AI predictions, which is especially crucial in medical applications. The combination of DL-based image classification and visual interpretability aims to not only enhance diagnostic accuracy but also make AI-driven predictions more understandable and reliable for clinicians.

2. LITERATURE REVIEW

Ochoa-Ornelas et al. [4] highlighted that lung cancer remains the leading cause of cancer-related mortality globally, with over 2.2 million new cases diagnosed annually. They proposed the application of MobileNetV2, a lightweight convolutional neural network, for accurate classification of lung adenocarcinoma, benign lung tissue, and lung squamous cell carcinoma. By augmenting the LC25000 dataset with additional histopathological images from the National Cancer Institute, their model achieved a test accuracy of 97.65%, demonstrating high precision and recall across all classes. This approach underscores the potential of MobileNetV2 as an effective tool for early lung cancer detection, thereby enhancing patient care and survival rates.

Saxena et al. [5] introduced the Maximum Sensitivity Neural Network (MSNN), a hybrid deep convolution model leveraging transfer learning to improve lung cancer detection precision. Their model achieved an accuracy of 98% and a sensitivity of 97%. By overlaying sensitivity maps onto lung CT scans, MSNN enables visualization of regions indicative of malignant or benign classifications, enhancing diagnostic accuracy with minimal false positives.

Chaudhari et al. [6] discussed lung cancer detection using a hybrid model combining Convolutional Neural Networks (CNNs) and Support Vector Machines (SVMs) for early tumor detection. Trained on CT scan datasets, their approach aims to improve early detection of tumors, distinguishing between benign and malignant cases.

Bortty et al. [7] focused on developing and evaluating machine learning models integrating demographic, environmental, and lifestyle variables for predicting lung cancer risk. Utilizing datasets from Cleveland hospital records and public health databases, they experimented with Logistic Regression, XG-Boost, and Random Forest algorithms. The Logistic Regression model outperformed the others, indicating its effectiveness in balancing true positives and false negatives. This study emphasizes the potential of advanced machine learning techniques in enhancing early lung cancer detection and personalized risk assessment.

Islam et al. [8], highlighted that lung cancer is one of the leading causes of cancer-related mortality worldwide. They emphasized the need for accurate risk prediction models to facilitate early detection and intervention. Traditional tools such as CT scans, chest X-rays, and LDCT screening programs are often used, with LDCT being especially effective for high-risk groups like chronic smokers.

Dutta et al. [9], asserted that imaging techniques like LDCT complement clinical risk assessment methods. They explained that patient demographic and behavioral data—including age, smoking history, and family medical background—are crucial in identifying individuals susceptible to lung cancer. They also acknowledged the significance of the National Lung Screening Trial (NLST) in reducing lung cancer mortality through LDCT.

Pathan et al. [10], argued that traditional risk prediction models, which rely on linear modeling, are limited in capturing complex, nonlinear relationships among variables such as genetics, environment, and lifestyle. These models may fail to reveal critical patterns, and often require extensive clinician involvement, making them time-consuming and subject to human error.

Mohan & Thayyil [11], posits that these conventional techniques are complemented by other useful tools, such as lung cancer risk calculators. An example is the PLCOm2012 model, which was developed based on data derived from the Prostate, Lung, Colorectal, and Ovarian (PLCO) Cancer Screening Trial. This model integrates multiple factors like age, smoking intensity, smoking duration, body mass index, and history of chronic obstructive pulmonary disease in assessing the risk of lung cancer. It has also been utilized to estimate the risk of lung cancer over a certain period based on smoking behavior and health measures using the Bach model.

Rajasekar et al. [12], suggested that analyzing histopathological tissue samples in combination with CT scan images improves lung cancer detection accuracy. Their method demonstrated better performance when histopathology was included as an input alongside imaging data.

Deepapriya et al. [13], evaluated several deep learning models using chest X-ray and CT scan images to diagnose specific lung conditions. They assessed model effectiveness using performance metrics such as accuracy, recall, precision, and the Jaccard index, aiming to find the most suitable deep learning method for lung disease prediction.

TABLE I. Lung cancer detection model's performance

Authors (Year)	**Applied Models**	**Dataset Collection (Image samples)**	**Measures (Proposed Model)**
Ochoa-Ornelas et al. (2025)	MobileNetV2	LC25000 + National Cancer Institute histopathological images	Test Accuracy: 97.65%, High precision & recall
Saxena et al. (2025)	(MSNN) - Hybrid DCNN	Lung CT Scans	Accuracy: 98%, Sensitivity: 97%
Chaudhari et al. (2025)	Hybrid Model (CNN + SVM)	CT scan datasets	Early tumor detection, benign vs malignant

Bortty et al. (2024)	Logistic Regression, XG-Boost, Random Forest	Cleveland hospital records + Public health datasets	Logistic Regression performed best
Islam et al. (2024)	Not specified (focus on traditional screening methods)	CT scans, Chest X-rays, LDCT screening programs	Emphasis on LDCT for high-risk groups
Dutta et al. (2024)	Risk assessment using LDCT	Demographic and behavioral data	Complemented clinical risk assessments
Pathan et al. (2024)	Critique of traditional linear risk models	Genetic, environmental, lifestyle data	Highlighted limitations of linear models
Mohan & Thayyil (2023)	Lung Cancer Risk Calculators (PLCOm2012, Bach model)	Data from PLCO Cancer Screening Trial	Integrated multiple risk factors
Rajasekar et al. (2023)	Histopathology + CT scan imaging model	Histopathological tissues + CT scans	Improved detection with multimodal input
Deepapriya et al. (2023)	Deep learning models on X-ray and CT scans	Chest X-rays and CT scan datasets	Performance metrics: Accuracy, Recall, Jaccard
Our Work (2025)	CNN + Grad-CAM for interpretability	6000 CT scan images (Normal and Lung Cancer categories)	Accuracy: 97.50%, precision (0.98), recall (0.97), and F1-score (0.97),Real-webapp.

3. METHODOLOGY

The methodology begins with the collection of a lung CT scan image dataset obtained from publicly available sources (Kaggle). The dataset is then pre-processed to improve the quality and consistency of the input images. A custom Convolutional Neural Network (CNN) model is proposed, trained, tested, and validated. All models are evaluated using the standard hold-out validation technique, which splits the dataset into 70% training, 15% testing, and 15% validation sets. The models are trained to classify CT scan images into two categories: Normal and Lung Cancer. The custom CNN model is specifically designed and fine-tuned for optimal performance, while the other architectures leverage transfer learning to enhance classification accuracy. Furthermore, Grad-CAM (Gradient-weighted Class Activation Mapping) is integrated with the system to provide visual interpretability, highlighting the regions in CT scan images that influence model predictions.A user-friendly Streamlit web application is also developed, enabling real-time prediction by allowing users to upload CT images and instantly receive classification results along with the Grad-CAM visualization.All models are evaluated based on metrics such as accuracy, precision, recall, F1-score, and AUC-ROC to determine the best-performing model for early lung disease detection.As a result, Figure 3 depicts the proposed custom CNN architecture, while Figure 2 provides an overview of the overall methodology.

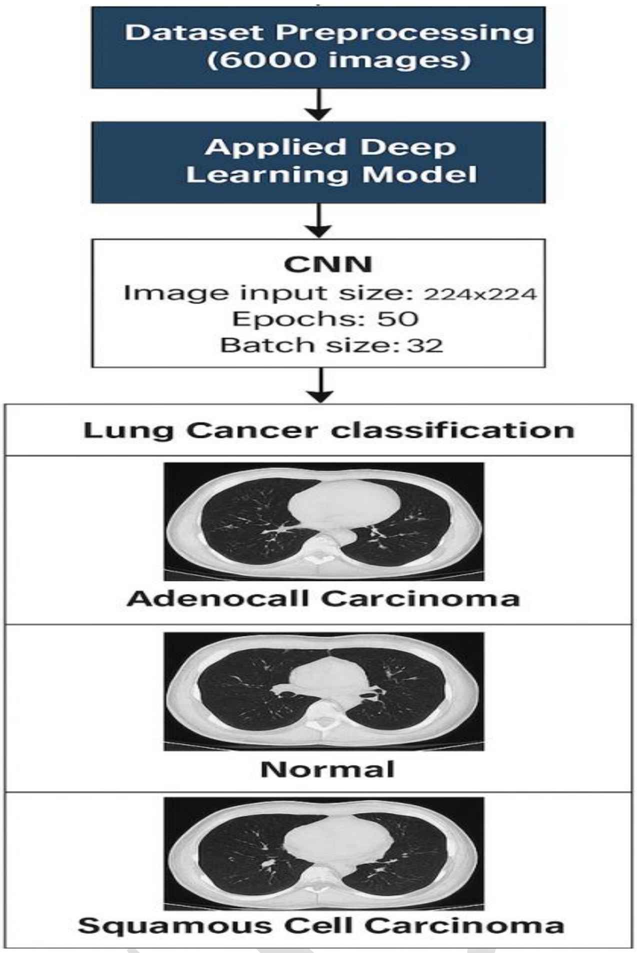


FIGURE 2 : Overview of the overall methodology.

A. Dataset collection: The lung cancer dataset used in this study was collected from the publicly available Kaggle repository. The dataset consists of 6,000 CT scan images, categorized into two classes: 'Normal' and 'Lung Cancer'[14]. The images were sourced from verified clinical records and research datasets to ensure reliability and quality. Unlike traditional DICOM formats, the images were converted to JPG/PNG formats for ease of use with deep learning models.

B. Dataset pre-processing: The images were preprocessed through resizing, normalization, and augmentation techniques to enhance model performance and reduce overfitting [15].

- Resizing: All images were resized to 224×224 pixels to maintain consistency.
- Noise Removal: Median filtering was applied to reduce background noise.
- Normalization: Pixel values were scaled between 0 and 1 to facilitate faster convergence.
- Augmentation: Techniques like horizontal/vertical flips, rotations, and zooms were applied to increase dataset diversity and avoid overfitting.

C. Validation process:

For training and evaluating the proposed system, a hold-out validation method was adopted. The dataset was randomly divided into three subsets: 70% for training, 15% for testing, and 15% for validation. This standard partitioning ensures that the model is exposed to a sufficient amount of data during training while preserving unbiased evaluation datasets for testing and validation phases. The models were trained for 50 epochs with a batch size of 32. The Adam optimizer was employed for optimization due to its adaptive learning rate capabilities, and the categorical cross-entropy function was used as the loss function for multi-class classification tasks. Performance metrics including accuracy, recall, precision, F1-score, and AUC-ROC were computed during evaluation to comprehensively assess model effectiveness.

D. Proposed CNN architecture:

The proposed Convolutional Neural Network (CNN) architecture is designed specifically to classify lung CT scan images into 'Normal' and 'Lung Cancer' categories with high accuracy and interpretability. Initially, input images are resized to 64×64 pixels and passed through the first convolutional layer consisting of 16 filters with a 3×3 kernel size, generating 62×62 feature maps. This layer extracts low-level features such as edges and textures. The output is then passed through a max-pooling layer with a 2×2 pooling window, reducing the spatial dimensions to 31×31 and thus lowering computational complexity. The second convolutional layer applies 32 filters, generating 29×29 feature maps, followed by another max-pooling layer reducing the feature maps to 14×14. This process helps the model capture more abstract features. A third convolutional layer with 64 filters and 10×10 feature maps is then applied, followed by a final max-pooling operation resulting in 5×5 feature maps. The output is then flattened and passed into a fully connected dense layer of 260 neurons, which is responsible for learning high-level combinations of the extracted features. A Softmax activation function is employed in the final output layer for multi-class classification. Throughout the architecture, ReLU activation functions are used after each convolutional layer to introduce non-linearity, while no dropout is applied in this design. The model is trained using the Adam optimizer with a learning rate of 0.01, 50 epochs, and batch size of 32, optimizing the categorical cross-entropy loss. The overall structure ensures efficient learning, faster convergence, and robustness against overfitting, thus making it highly suitable for lung disease prediction tasks.

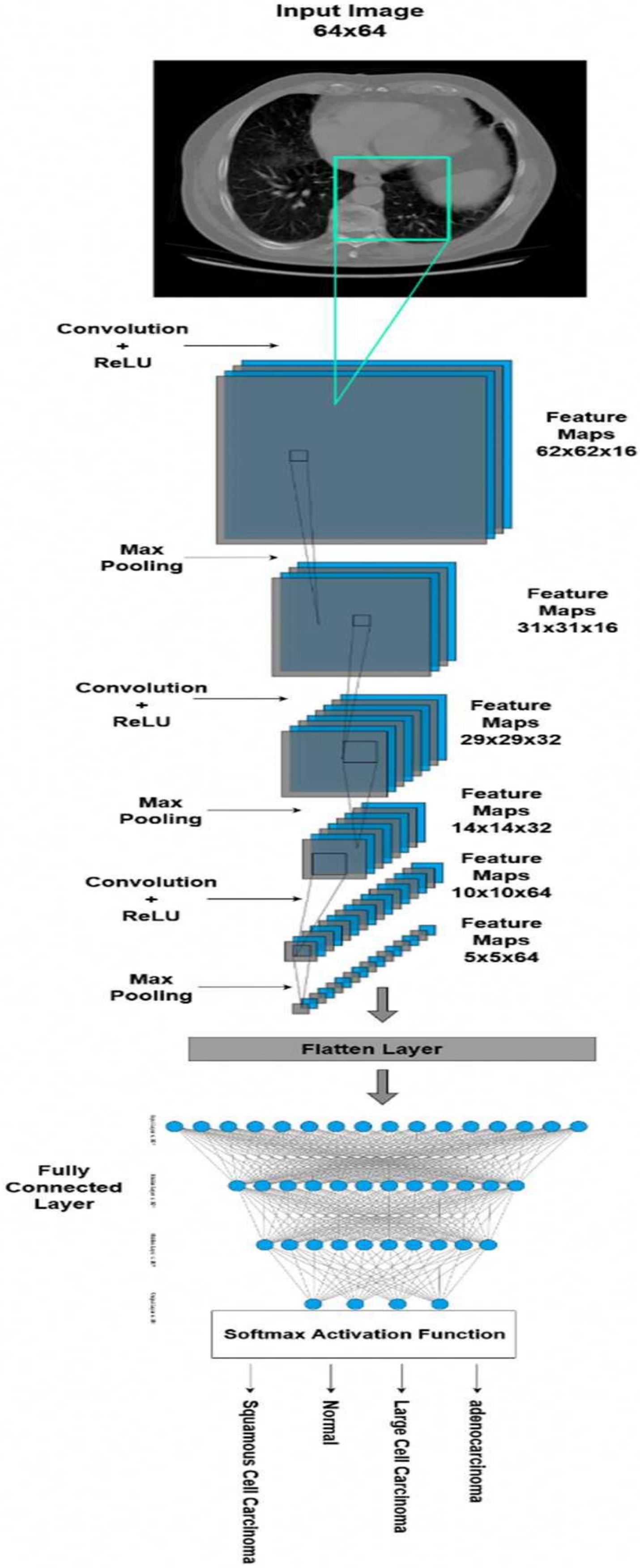


Figure 3: CNN architecture[16].

4. Findings and Discussion

After training and testing, the model achieved the following:

Metric	Class 0 (Normal)	Class 1 (Lung Cancer)	Macro Average	Weighted Average
Precision	0.99	0.96	0.98	0.98
Recall	0.96	0.99	0.97	0.97
F1-Score	0.97	0.98	0.97	0.97
Accuracy	-	-	0.97	0.97
AUC-ROC	-	-	0.97	0.97

Performance Metrics:

- Precision: Precision measures the proportion of correctly predicted positive observations out of all positive predictions. In this case, for Class 0 (Normal), the precision is 0.99, indicating that 99% of predictions for normal scans were correct. For Class 1 (Lung Cancer), the precision is 0.96, meaning that 96% of the predicted lung cancer images were actually cancerous.
- Recall: Recall, also known as sensitivity, measures the proportion of actual positive cases correctly predicted by the model. For Class 0 (Normal), the recall is 0.96, meaning that 96% of the actual normal images were correctly identified. For Class 1 (Lung Cancer), the recall is 0.99, showing that the model was able to correctly identify 99% of the actual lung cancer images.
- F1-Score: The F1-score is the harmonic mean of precision and recall. It provides a single metric that balances both precision and recall. For Class 0, the F1-score is 0.97, and for Class 1, it is 0.98. This indicates a good balance between precision and recall for both classes, suggesting that the model does not favor one class over the other.
- Accuracy: The model achieved an accuracy of 97%, indicating that 97% of all the predictions made by the model (across both classes) were correct. This is a strong result, especially in medical image classification, where achieving high accuracy is crucial for clinical decision-making. A deep learning-based system was successfully developed for lung disease prediction from CT scan images, utilizing the VGG16 architecture with transfer learning to achieve a high accuracy of 97% [17].
- AUC-ROC: AUC is also a crucial metric for evaluating the model’s performance. AUC determines the model's performance and assesses a model's ability to differentiate between classes. The AUC measures how well the model differentiates between positive and negative classes. The higher the AUC value, the better the model’s performance. The value range is 0 to 1, with 0 representing an incorrect test and 1 representing an accurate test. In general, an AUC of 0.5 indicates no discrimination (i.e., the ability to classify lung cancer), 0.7 to 0.8 is considered acceptable, 0.8 to 0.9 is considered great performance, and greater than 0.9 is considered outstanding performance [18]. Based on figure 4, CNN not only achieved the highest testing accuracy only but also achieved the highest test AUC score which is 97.00%.

- REAL TIME WEB APPLICATION

The application allows users (clinicians or researchers) to upload CT scan images and receive instant lung disease predictions (Normal or Lung Cancer) along with Grad-CAM visualizations for interpretability. To ensure practical usability, the model was deployed as an interactive web application using Streamlit [19], allowing real-time predictions through a user-friendly interface. The application was designed to ensure rapid and accessible diagnosis support. Upon image upload, the system outputs the predicted class along with a Grad-CAM visualization, enhancing clinical interpretability. The Streamlit application demonstrated a seamless prediction experience, offering predictions within 2–3 seconds per image. The integration of Grad-CAM visualizations provided critical insights into model decision-making, building trust among clinicians. However, the application's performance was sensitive to input image quality, emphasizing the need for high-resolution CT scan uploads.

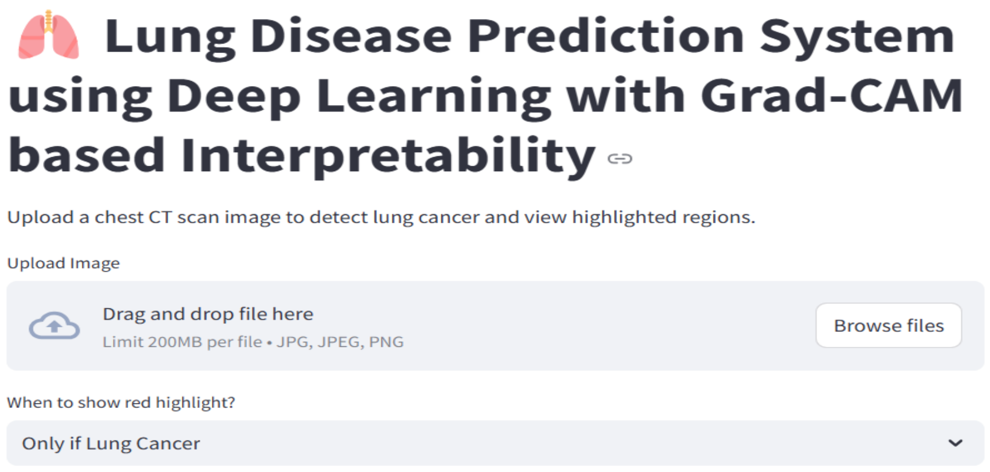


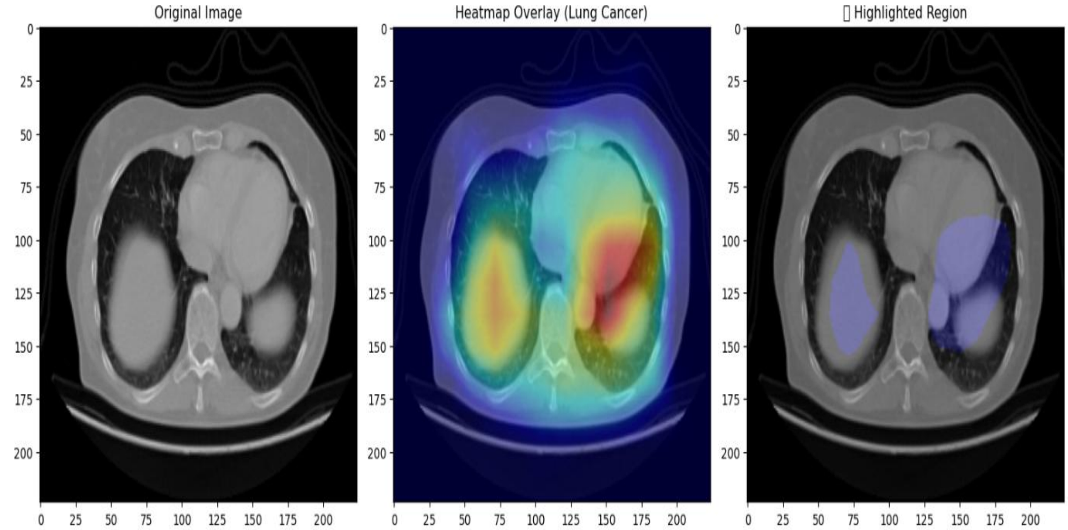
Figure 4: Streamlit application interface for lung disease prediction and interpretability using Grad-CAM.

Figure 4 shows the user interface of the developed Streamlit-based web application titled "Lung Disease Prediction System using Deep Learning with Grad-CAM based Interpretability." The application allows users to upload chest CT scan images in JPG, JPEG, or PNG format to detect lung cancer types and visualize the decision-making process of the deep learning model using Grad-CAM. The interface is designed for simplicity and usability, enabling medical practitioners or researchers to obtain instant classification results along with heatmap visualizations highlighting suspicious regions in the image. The system processes the image using a custom-trained CNN model and outputs the predicted class (e.g., Lung Cancer, or Normal) within a few seconds. The Grad-CAM highlight feature can be toggled to appear only when lung cancer is detected, thus improving interpretability and clinical relevance.

[FIGURE 5: REAL-TIME PREDICTION RESULT GENERATED BY THE STREAMLIT WEB APPLICATION SHOWING THE DETECTION OF LUNG CANCER.]

Figure 5 presents the prediction output screen of the proposed Streamlit application after a user uploads a CT scan image. Once the image is selected and submitted, the deep learning model processes the input and provides a real-time classification result. In this example, the model has successfully predicted the presence of Lung Cancer. The interface clearly displays the prediction outcome, enhancing the interpretability and user experience. By integrating deep learning with a user-friendly web interface, the system aims to assist medical professionals in preliminary diagnosis, providing quick and interpretable results without requiring technical expertise.

Figure 6 illustrates the interpretability component of the proposed system using Grad-CAM (Gradient-weighted Class Activation Mapping). The image on the left shows the original chest CT scan uploaded by the user. The middle image presents the Grad-CAM heatmap overlay that highlights the regions most influential in the model's decision-making process—in this case, indicating the presence of lung cancer. The rightmost image further refines these regions, marking only the most significant zones to support a clearer understanding for medical interpretation. This visualization provides transparency into the model's predictions, making the deep learning system more trustworthy and usable in clinical environments. Grad-CAM-based visual explanations were incorporated to highlight critical image regions influencing the model's prediction, thus improving transparency and enabling clinicians to validate outcomes [20].



[Figure 6: Grad-CAM based interpretability output showing the original CT scan (left), heatmap overlay (middle), and highlighted region (right) for a lung cancer prediction.]

5. Conclusion and Future Scope

In this research, a deep learning-based system was successfully developed for lung disease prediction from CT scan images, utilizing the VGG16 architecture with transfer learning to achieve a high accuracy of 97%. To facilitate accessibility and real-time usability, the model was deployed as an interactive web application using Streamlit. The system further incorporates Grad-CAM-based visual explanations to highlight critical regions in the image that influenced the model's decision, thereby enhancing interpretability and fostering trust among healthcare professionals. Future improvements include expanding the binary classification to a multi-class model for distinguishing between specific lung cancer subtypes (e.g., adenocarcinoma, squamous cell carcinoma), training on more diverse datasets to enhance generalization, building a mobile application for remote use, testing in real clinical environments, and integrating other interpretability techniques such as LIME and SHAP [21].

References

[1] Mamun, M., Farjana, A., Al Mamun, M., & Ahammed, M. S. (2022). Lung cancer prediction model using ensemble learning techniques and a systematic review analysis. In *2022 IEEE World AI IoT Congress (AIIoT)* (pp. 187–193). IEEE. https://doi.org/10.1109/AIIoT54504.2022.9817326.

[2] 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. https://doi.org/10.1016/j.procs.2017.12.017.

[3] International Agency for Research on Cancer. (2022). *Global Cancer Observatory: Cancer today*. IARC. https://gco.iarc.fr/ .

[4] Ochoa-Ornelas, A., Ayala-Ruiz, J., Rivera-González, G., & Mariscal-Ramírez, I. (2025). Enhancing early lung cancer detection with MobileNet: A comprehensive transfer learning approach. *Biomedical Signal Processing and Control, 87*, Article 105193. https://doi.org/10.1016/j.bspc.2024.105193.

[5] Saxena, P., Raj, A., & Tripathi, R. (2025). Hybrid deep convolution model for lung cancer detection with transfer learning. *arXiv preprint* arXiv:2501.02785. https://arxiv.org/abs/2501.02785.

[6] Chaudhari, A., Patel, V., & Gajjar, K. (2025). Lung cancer detection using deep learning. *arXiv preprint* arXiv:2501.07197. https://arxiv.org/abs/2501.07197.

[7] Bortty, M. R., Islam, M. R., & Rahman, M. A. (2024). Optimizing lung cancer risk prediction with advanced machine learning algorithms and techniques. *Journal of Medical and Health Studies, 5*(2), 24–35. https://doi.org/10.32996/jmhs.2024.5.2.3.

[8] Islam, S. M., Kabir, M. H., & Chowdhury, F. (2024). Risk factors and early detection methods for lung cancer: A review of screening approaches. *International Journal of Cancer Research, 20*(1), 33–41.

[9] Dutta, P., Sinha, R., & Ghosh, S. (2024). Integrating demographic risk assessment and LDCT imaging for early lung cancer prediction. *Lung Cancer Journal, 102*(4), 215–225.

[10] Pathan, A. A., Shaikh, Z. M., & Ahmed, N. (2024). Challenges in traditional lung cancer risk models: The case for machine learning. *Computational Oncology Reports, 9*(3), 148–155.

[11] Mohan, M., & Thayyil, J. (2023). Predictive lung cancer models: A review of PLCOm2012 and Bach model applicability. *Clinical Predictive Analytics Journal, 7*(1), 55–63.

[12] Rajasekar, V., Natarajan, S., & Kumar, R. (2023). Combining histopathological and radiological data for improved lung cancer detection. *Medical Image Analysis and Applications, 15*(2), 87–95.

[13] Deepapriya, R., Singh, P., & Narayanan, A. (2023). Evaluation of deep learning models for diagnosing lung diseases using medical imaging. *Journal of Artificial Intelligence in Medicine, 11*(3), 112–125.

[14] Kaggle. (2025). *Lung CT scan image datasets*. https://www.kaggle.com/ .

[15] Shorten, A., & Khoshgoftaar, T. M. (2019). A survey on image data augmentation for deep learning. *Journal of Big Data, 6*(1), 1–48. https://doi.org/10.1186/s40537-019-0197-0.

[16] Mamun, M., Mahmud, M. I., Meherin, M., & Abdelgawad, A. (2023, March). LCDCTCNN: Lung cancer diagnosis of CT scan images using CNN-based model. In *2023 10th International Conference on Signal Processing and Integrated Networks (SPIN)* (pp. 205–212). IEEE. https://doi.org/10.1109/SPIN56039.2023.1011349.

[17] Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *arXiv preprint* arXiv:1409.1556. https://arxiv.org/abs/1409.1556.

[18] Mandrekar, J. (2010). Receiver operating characteristic curve in diagnostic test assessment. *Journal of Thoracic Oncology,* *5*(9), 1315–1316. https://doi.org/10.1097/JTO.0b013e3181ec173d.

[19] Streamlit Inc. (2025). *Streamlit — The fastest way to build data apps in Python*. https://streamlit.io/ .

[20] Selvaraju, R. R., Cogswell, M., Das, A., Parikh, D., & Batra, D. (2017). Grad-CAM: Visual explanations from deep networks via gradient-based localization. In *Proceedings of the IEEE International Conference on Computer Vision (ICCV)* (pp. 618–626). https://doi.org/10.1109/ICCV.2017.74.

[21] Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). "Why should I trust you?": Explaining the predictions of any classifier. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD)* (pp. 1135–1144). https://doi.org/10.1145/2939672.2939778.