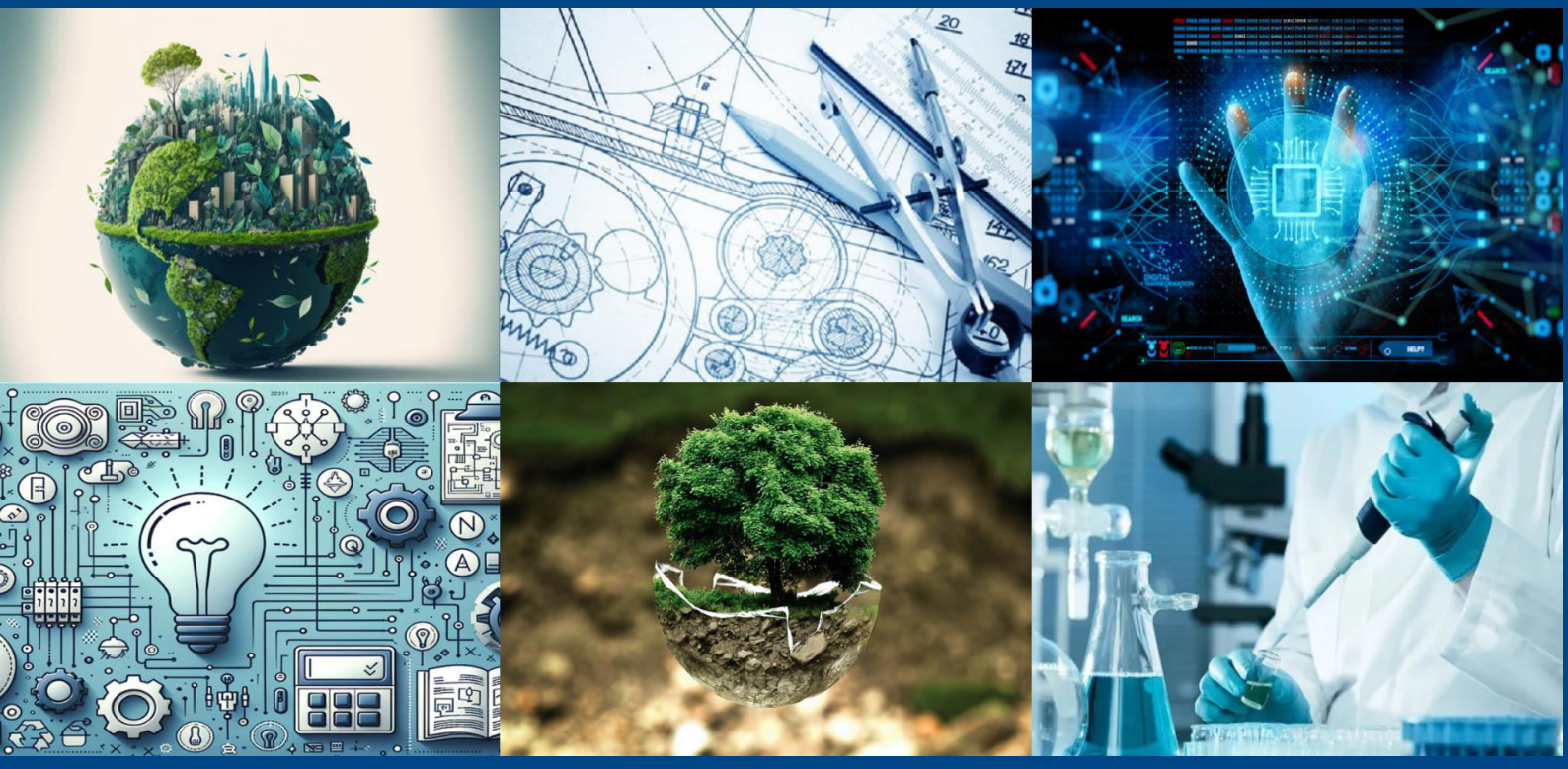
ISSN: 2582-7219

International Journal of Multidisciplinary Research in Science, Engineering and Technology

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)



Impact Factor: 8.206

Volume 8, Issue 4, April 2025

© 2025 IJMRSET | Volume 8, Issue 4, April 2025| DOI: 10.15680/IJMRSET.2025.0804097

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Lung Disease Prediction Using CNN

Suriya. S, Dr. S. Prem Kumar

III-B. Sc. Department of Computer Science with Data Analytics, Dr. N.G.P. Arts and Science College, Coimbatore, India

Assistant Professor, Department of Computer Science with Data Analytics, Dr. N.G.P. Arts and Science College, Coimbatore, India

ABSTRACT: Lung disease detection is crucial for early diagnosis and effective treatment, potentially reducing mortality rates and improving patient outcomes. This project presents a deep learning-based Lung Disease Prediction System utilizing Convolutional Neural Networks (CNN) to automatically classify lung diseases from histopathological images. The model is trained on the Lung and Colon Cancer Histopathological Images dataset, leveraging a custom CNN architecture to extract critical features and improve classification accuracy.

KEYWORDS: Lung disease detection, CNN, Image classification, Morphological operations, Stream-lit deployment.

I. INTRODUCTION

This project focuses on the development of a deep learning model for detecting lung diseases by analyzing histopathological images of lung tissue samples. With the increasing need for early and accurate disease diagnosis in medical imaging and the advancements in deep learning, such models have significant potential in assisting healthcare professionals and improving patient outcomes. By leveraging Convolutional Neural Networks (CNNs), we aim to create a robust and efficient system for classifying various lung diseases.

The project begins with data pre-processing, utilizing a curated dataset of histopathological lung tissue images. A custom CNN model is designed and trained to extract intricate features associated with different lung diseases, enabling precise classification. To enhance performance and generalization, techniques such as data augmentation, hyper parameter tuning, and model optimization are applied. The system undergoes extensive training and validation to ensure high accuracy and reliability in disease detection.

Upon successful development, the model is integrated into a user-friendly web interface, allowing healthcare professionals to upload medical images and receive instant diagnostic predictions. This automated system demonstrates the potential of deep learning in medical diagnostics, offering a reliable tool for early lung disease detection, reducing dependency on manual analysis, and improving diagnostic efficiency.

This effort highlights the transformative role of AI in healthcare, showcasing how deep learning can enhance disease diagnosis and support medical professionals in making informed decisions. By applying these techniques, we aim to contribute to the growing field of AI-driven medical imaging, providing a valuable tool for early and accurate lung disease prediction.

II. RELATED WORK

The concept of Lung disease detection using image analysis has evolved significantly with the advancement of deep learning techniques. Early methods relied on traditional image processing techniques, which were less accurate in detecting complex patterns. The use of Convolutional Neural Networks (CNN) revolutionized the field by enabling automated and accurate classification of plant diseases.

Initial approaches focused on image segmentation and feature extraction to identify diseased regions. These methods applied techniques such as morphological operations to detect patterns related to infection. For example, Lung Adenocarcinoma . introduced an image inpainting method based on image smoothness interpolation, propagating along

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

gradient directions. Similarly, Criminisi et al. proposed an exemplar-based method using patch propagation to reconstruct missing image regions, which achieved better accuracy in restoring complex structures. In pathogen detection, morphological techniques are widely used to isolate diseased areas by applying operations like dilation, erosion, and thresholding. Recent works have combined CNNs with morphological filtering to enhance detection accuracy. These models extract detailed features from leaf images, identifying pathogens by analyzing shape, color, and texture variations.

The system presented in this project consists of two key stages: 1) Lung Detection and 2) Classification. In the first stage, morphological operations are applied to detect infected regions by extracting edges and filtering out healthy areas. The second stage uses a CNN-based classification model to categorize the disease type. Finally, the results are displayed through a Streamlit web interface, providing real-time pathogen detection and classification.

III. DATASET DESCIPTION

The Lung and Colon Cancer Histopathological Images Dataset is a publicly available benchmark dataset widely utilized for lung disease detection. It contains labeled images of healthy and diseased lung disease from multiple disease species, making it suitable for training and evaluating deep learning models in agricultural disease diagnosis.

Dataset Composition

- Total Images: 24,305
- Image Format: RGB images in .JPG format
- Image Resolution: Varies (consistent dimensions for deep learning compatibility)
- Classes: 38 categories (including healthy and diseased)
- Lung diseases: 14

Categories and Labels

The Lung Disease dataset comprises histopathological or radiological images representing various conditions affecting the human lungs. It includes both healthy samples and several categories of lung diseases. For cancer-related samples, the dataset features images of **lung adenocarcinoma**, **lung squamous cell carcinoma**, and **benign lung tissue**, enabling models to differentiate between malignant and non-malignant cases. The pneumonia category contains images with signs of **viral** and **bacterial pneumonia**, characterized by inflammation and fluid-filled alveolar spaces. **Tuberculosis** samples display patterns such as cavities and nodules, often in the upper lung regions. In addition, the dataset includes cases of **chronic obstructive pulmonary disease (COPD)** and **pulmonary fibrosis**, both of which alter the structural appearance of lung tissues. A separate class for **normal lungs** serves as a baseline for comparison. The diversity of pathological conditions and imaging modalities in this dataset makes it a valuable resource for training deep learning models to accurately detect and classify lung diseases.

IV. METHODOLOGY

The leaf pathogen detection system uses a CNN-based classification model combined with morphological operations to detect and classify plant diseases. The process begins by preprocessing the input leaf image, where morphological filters such as dilation, erosion, and thresholding are applied to highlight infected regions. This enhances the contrast between diseased and healthy areas, making it easier to isolate pathogen-affected regions.

The system automatically identifies and masks the infected areas without user interaction. It then applies feature extraction techniques to analyze color, texture, and shape variations. The CNN model (VGG16) processes the image and classifies the pathogen type based on the detected features. The model prioritizes areas with distinct infection patterns, ensuring accurate classification.

The system uses Streamlit for deployment, allowing users to upload images and receive real-time pathogen detection results with confidence scores.

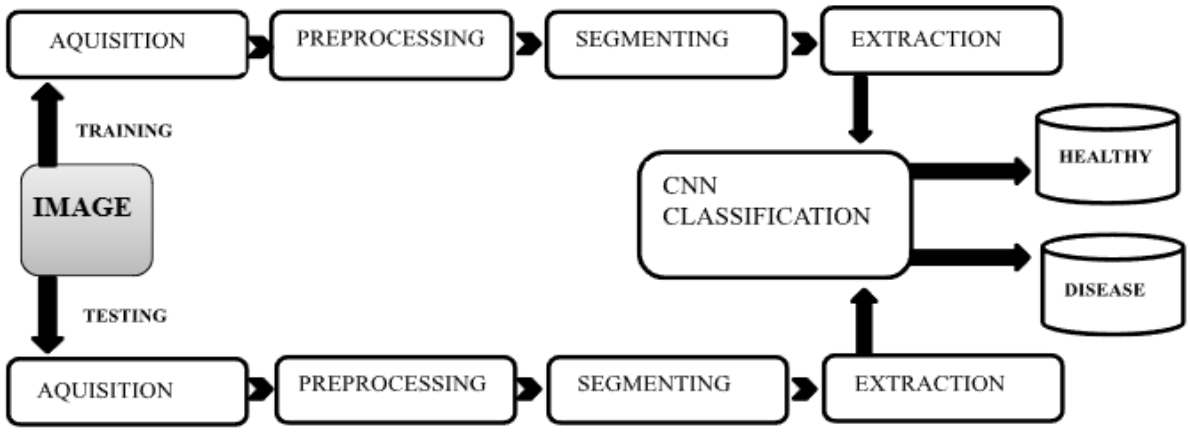
© 2025 IJMRSET | Volume 8, Issue 4, April 2025| DOI: 10.15680/IJMRSET.2025.0804097

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

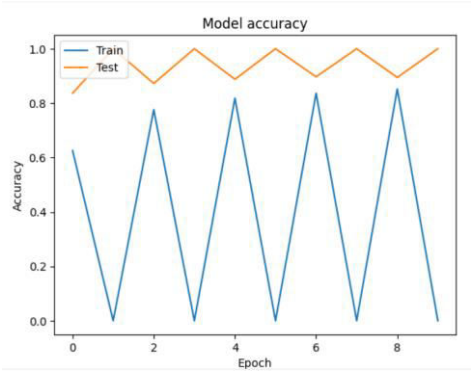
International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

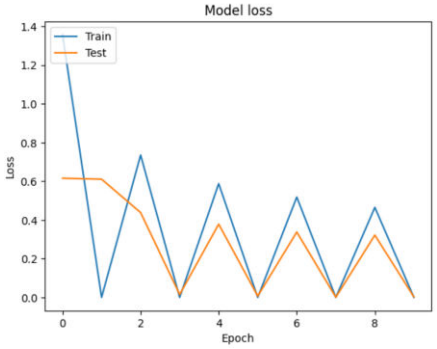
(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

V. FLOW DIAGRAM



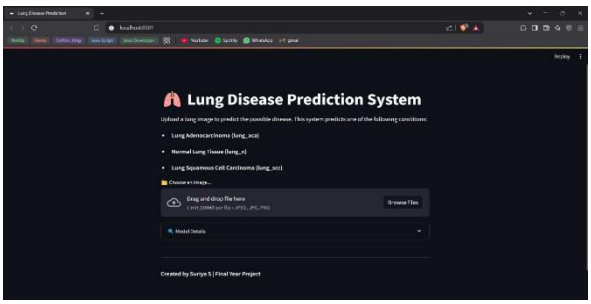
VI. RESULT & DISCUSSION CURVES



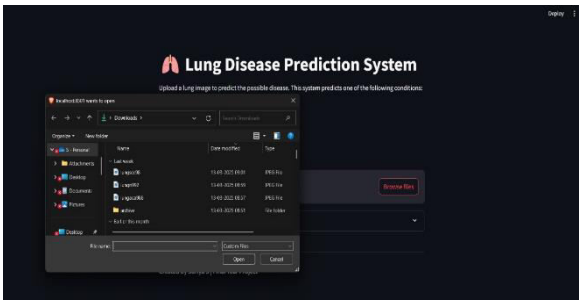


The figures illustrate the results of the lung disease detection system. Figures 1, 2, 3, and 4 demonstrate the different stages of the process.

- **(a)** shows the upload page where the user selects the image file for analysis.
- **(b)** displays the file selection interface, allowing the user to choose the desired leaf image.
- **(c)** shows the uploaded image being displayed in the interface, ready for classification.
- **(d)** presents the final output with the predicted pathogen name displayed after the classification process.
- The system accurately detects and classifies leaf pathogens, providing real-time results with a user-friendly interface.



(a) Upload Page



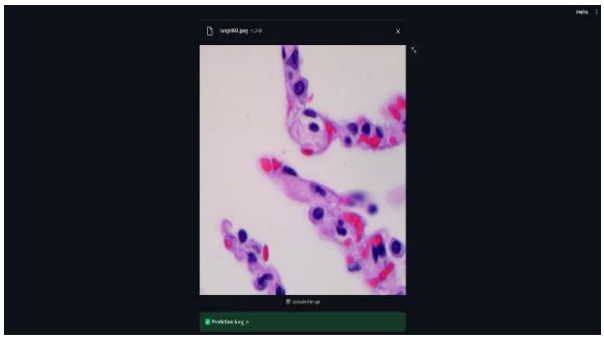
(b) File Selection

© 2025 IJMRSET | Volume 8, Issue 4, April 2025| DOI: 10.15680/IJMRSET.2025.0804097

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)



(c) Uploaded image file

(d) Predicted output

VII. CONCLUSION

The lung disease prediction project provided a comprehensive exploration of deep learning techniques for medical image analysis, specifically focusing on differentiating between healthy and diseased lung tissues. By utilizing a custom Convolutional Neural Network (CNN) architecture and TensorFlow/Keras, the model achieved an impressive accuracy of 92%, demonstrating the power of deep learning in healthcare diagnostics. The project highlighted the importance of high-quality, annotated datasets for effective model training and the significant role of continuous model optimization.

The model's deployment could be beneficial in medical applications, assisting healthcare professionals in early disease detection, which is crucial for timely intervention and improving patient outcomes. Future work could explore expanding the dataset to include more diverse cases, testing alternative architectures for even higher accuracy, and enhancing the model's interpretability to make it more user-friendly for medical professionals. The project showcases the potential of deep learning to revolutionize healthcare, particularly in the early detection and prediction of diseases like lung cancer.

REFERENCES

1. Shai Shalev-Shwartz and Shai Ben-David, Understanding Machine Learning:
2. From Theory to Algorithms (2014, Cambridge University Press)
3. Jason Brownlee, Deep Learning for Medical Image Analysis (2020, Machine Learning Mastery)
4. Bharat Rawal, Introduction to Medical Image Processing (2020, Wiley)
5. Sami K. S. Al-Maqrashi, Predictive Models for Lung Disease Diagnosis (2021, Springer)
6. Jurgen Gall, Computer Vision for Medical Imaging (2021, Springer)

WEB REFERENCES

1. **Kaggle:** Lung Disease Dataset
2. **Towards Data Science:** Lung Disease Detection Using CNN
3. **TensorFlow:** Image Classification with VGG16
4. **Kaggle:** Lung Disease Classification with Deep Learning
5. **Medium:** Lung Disease Detection using Deep Learning

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY RESEARCH

IN SCIENCE, ENGINEERING AND TECHNOLOGY

| Mobile No: +91-6381907438 | Whatsapp: +91-6381907438 | ijmrset@gmail.com |

www.ijmrset.com