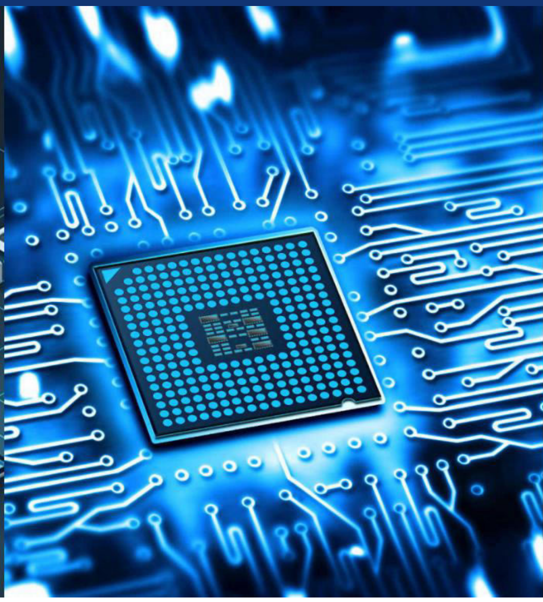
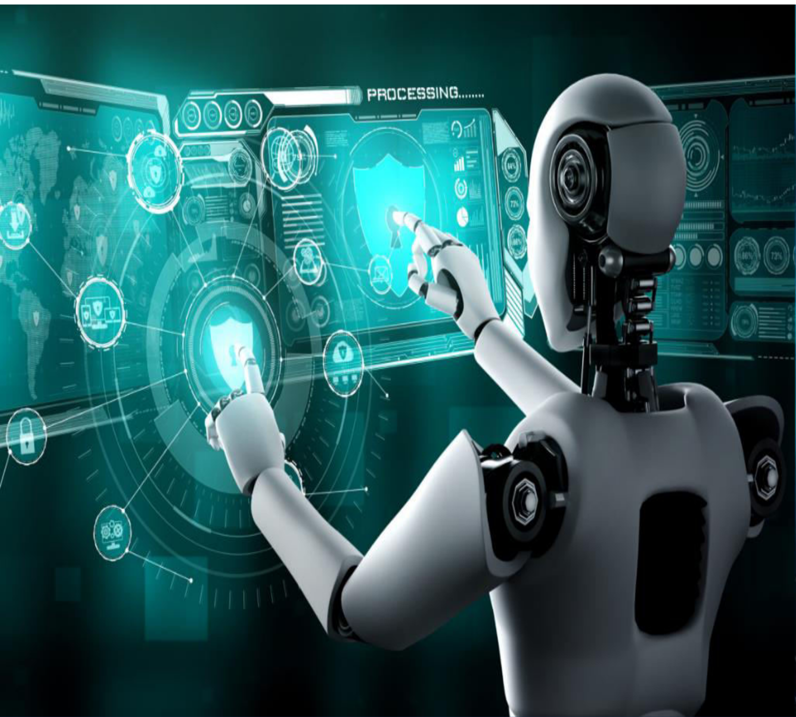
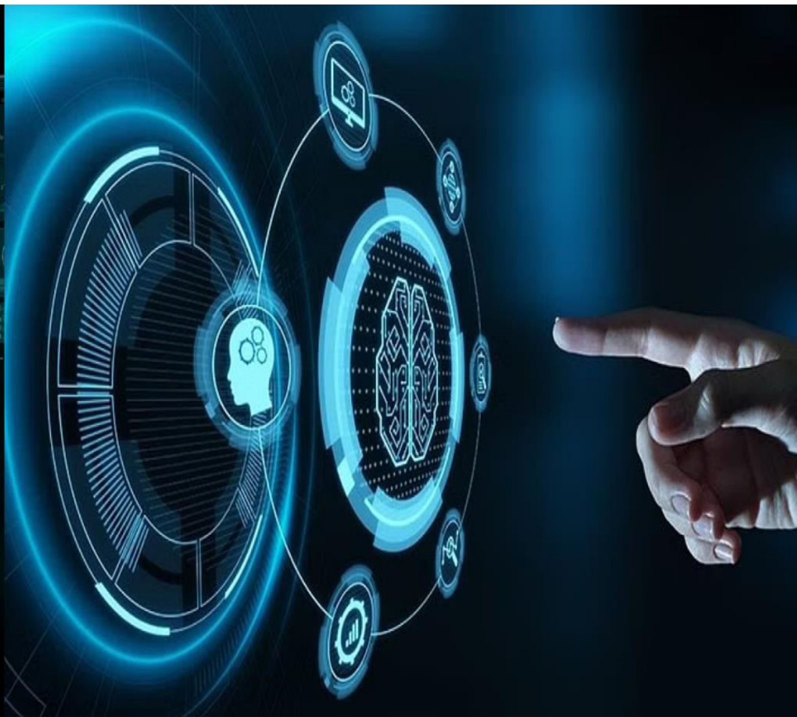
ISSN(O): 2320-9801
ISSN(P): 2320-9798


International Journal of Innovative Research in Computer and Communication Engineering

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





Impact Factor: 9.927

Volume 14, Issue 6, June 2026

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 9.927| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

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

Lung Disease Prediction using Data Science with Patient Care Recommendations & Suggestions

Prof. Vishakha Jadhav[1], Ashutosh Mujmule[2], Swaraj Chavhan[3], Dnyaneshwari Shinde[4], Shrutika Walunj[5]

Guide, Computer Engineering Department of Computer Engineering, Dhole Patil College of Engineering, Pune, Maharashtra, India[1]

Final Year B.E. Student, Department of Computer Engineering, Dhole Patil College of Engineering, Pune, Maharashtra, India[2-5]

ABSTRACT*:* This paper presents Lungs disease prediction, a web-based lung disease prediction system designed to support clinicians and patients in the early identification and management of pulmonary conditions. Respiratory diseases including Pneumonia, Tuberculosis, COVID-19, Pleural Effusion, and Normal conditions collectively contribute to significant global mortality. The proposed system provides an end-to-end diagnostic pipeline through a chest X-ray image upload interface, where the uploaded image is analyzed using a deep learning model to classify the lung condition and predict its severity. The model is based on a DenseNet201 Convolutional Neural Network (CNN) architecture trained using transfer learning, which classifies chest radiographs into five categories: COVID-19, Pleural Effusion, Normal, Pneumonia, and Tuberculosis. The application includes additional features such as a symptom checker, AI chatbot assistant, breathing exercise module, dietary guidance, and a wellness tracker to encourage proactive health management. A doctor dashboard is integrated into the system, allowing medical professionals to monitor patient scan histories and review predicted risk levels. The system is implemented using Flask for the backend, HTML/CSS/JavaScript for the frontend, SQLite (Lungs disease prediction.db) for data storage, and TensorFlow/Keras for deep learning inference. The overall framework aims to provide early detection, real-time prediction, and patient-centred guidance for lung disease management.

KEYWORDS: Lung Disease Prediction, Deep Learning, DenseNet201, CNN, Chest X-ray Classification, COVID-19, Pneumonia, Tuberculosis, Pleural Effusion, Flask, Web Application, Transfer Learning, Patient Health Guidance, Medical Image Analysis.

I. INTRODUCTION

Respiratory diseases represent one of the most critical public health challenges worldwide. Conditions such as Pneumonia, Tuberculosis (TB), COVID-19, and Pleural Effusion collectively account for millions of deaths annually. According to the World Health Organization (WHO), respiratory illnesses contribute to nearly one-fifth of all deaths globally, making early and accurate diagnosis a paramount requirement.

The conventional diagnostic pathway for lung diseases relies on radiological interpretation of chest X-rays by trained specialists. However, in many developing nations and rural healthcare settings, access to experienced radiologists is severely limited. This shortage results in delayed diagnoses, misinterpretations, and suboptimal treatment outcomes, all of which worsen patient prognosis.

The growing availability of digitized medical imaging datasets, combined with advances in Artificial Intelligence (AI) and Deep Learning, offers a transformative opportunity to address this gap through automated image analysis. Convolutional Neural Networks (CNNs) have emerged as the dominant architecture for medical image classification, demonstrating the ability to learn complex hierarchical features from raw pixel data without manual feature engineering. Transfer learning from pre-trained models such as DenseNet201 has further enhanced classification accuracy even with limited training data — a common constraint in the medical domain.

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 9.927| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

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

The proposed system, Lungs disease prediction, is a web-based lung disease prediction application designed for both patients and medical professionals. The application allows users to upload chest X-ray images, which are then analyzed by a DenseNet201 deep learning model to classify the lung condition and predict severity. In addition to prediction, the system provides features such as a symptom checker, AI health chatbot, breathing exercises, dietary guidance, and a wellness tracker to promote proactive respiratory health management.

A key aspect of the system is the integrated doctor dashboard, where medical professionals can review patient scan histories, risk levels, and statistical analytics through a secure role-based account. The system also generates downloadable PDF reports containing clinical findings and model confidence scores. By combining web technology, deep learning, and user-friendly features, Lungs disease prediction aims to provide an effective solution for early identification and management of lung diseases.

The main objective of this work is to design and develop a user-friendly system that can assist in automated lung disease screening, provide early risk prediction using deep learning, and enable clinical decision support through a deployable web interface. The system also aims to promote patient wellness through integrated health guidance modules.

II. LITERATURE REVIEW AND RELATED WORK

Previous studies have explored the use of machine learning and deep learning techniques for automated lung disease detection from chest X-ray images. Various approaches such as Support Vector Machines (SVM), decision trees, and neural networks have been applied to identify patterns related to Pneumonia, Tuberculosis, and COVID-19. These studies highlight the importance of early detection and continuous monitoring to prevent severe pulmonary conditions.

The introduction of CNNs fundamentally changed medical image analysis. Rajamanickam et al. (2022) proposed a CNN model for Pneumonia detection achieving 93% testing accuracy, outperforming SVM and KNN baselines. Li and Zhao (2023) demonstrated transfer learning using VGG16, ResNet50, and InceptionV3 for multi-class lung disease detection. Kumar et al. (2024) introduced a CNN combined with DenseNet201, enhancing feature reuse through dense connectivity and reducing overfitting via dropout regularization. Bhatia et al. (2024) demonstrated that CLAHE (Contrast Limited Adaptive Histogram Equalization) improved CNN performance by up to 5% by enhancing local contrast in chest radiographs.

Several web and mobile-based applications have also been developed to support health monitoring through features such as symptom checking, chatbot interaction, and dietary guidance. However, many of these systems focus on either image classification or patient support features independently, without integrating a deep learning prediction model with a patient wellness advisory module, PDF report generation, and role-based access for both patients and doctors in a unified platform. Additionally, existing systems often lack real-time severity assessment and downloadable clinical reports.

The proposed system, Lungs disease prediction, aims to address these limitations by combining DenseNet201-based deep learning prediction with user-friendly features including a symptom checker, AI chatbot, breathing exercises, dietary guidance, doctor dashboard, and downloadable PDF reports. This integrated approach supports continuous monitoring, early detection, and timely clinical intervention for lung disease management.

III. METHODOLOGY

The methodology of the proposed Lungs disease prediction system focuses on the process of collecting chest X-ray images, preprocessing them, analyzing using a deep learning model, and generating lung disease predictions with clinical guidance. The approach is designed to ensure accurate multi-class classification of pulmonary conditions based on user-uploaded radiographs.

3.1 Data Collection

The system accepts chest X-ray images uploaded by users through the web interface. The training dataset was assembled from publicly available, clinically validated sources including the NIH Chest X-ray14 Dataset and the

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 9.927| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

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

Kaggle Chest X-ray Dataset. Images from five classes were extracted and balanced: Normal (1,000), Pneumonia (1,000), Tuberculosis (1,000), Pleural Effusion (1,000), and COVID-19 (1,000), totalling 5,000 training images.

TABLE I — DATASET COMPOSITION BY CLASS

Disease Class	No. of Images	Source
Normal	1,000	NIH CXR14
Pneumonia	1,000	Kaggle / NIH
Tuberculosis	1,000	NIH / Montgomery
Pleural Effusion	1,000	NIH CXR14
COVID-19	1,000	NIH / Kaggle
Total	**5,000**	—

3.2 Data Preprocessing
Uploaded images are preprocessed before being passed to the deep learning model. All images are resized to 224×224 pixels to match the DenseNet201 input dimensions. Pixel values are normalized by dividing by 255.0 to scale the intensity range to [0, 1]. This preprocessing pipeline is implemented using TensorFlow's Keras image utilities (keras_image.load_img() and keras_image.img_to_array()) to ensure consistency across all inference inputs.

3.3 Feature Extraction
The DenseNet201 backbone extracts hierarchical visual features from the preprocessed chest X-ray images. DenseNet's dense connectivity pattern — where each layer receives feature maps from all preceding layers — enables maximum feature reuse and significantly reduces the number of trainable parameters. The backbone outputs a 7×7×1920 feature map, which is passed through a GlobalAveragePooling2D layer to produce a 1920-dimensional feature vector.

3.4 Deep Learning Model
A DenseNet201 CNN is used for multi-class classification with transfer learning from ImageNet pre-trained weights. The complete classification head comprises: GlobalAveragePooling2D → Dense(256, ReLU) → Dropout(0.5) → Dense(5, Softmax). Training used the Adam optimizer with learning rate 1×10^{-4} and ReduceLROnPlateau scheduling. Early stopping (patience=7) monitored validation loss to prevent over-training.

3.5 Prediction Process
After the user uploads a chest X-ray image, the input is preprocessed and passed to the trained DenseNet201 model loaded at application startup via a singleton pattern. The model processes the 224×224×3 image tensor and outputs a 5-class softmax probability distribution. The class with the highest probability is selected as the predicted condition, and integer-percentage confidence scores are returned for all five classes.

3.6 Output Generation
The predicted result is returned to the web application where it is displayed to the user along with a confidence score, clinical description, severity level (Low / Moderate / High), and personalized treatment recommendations. A downloadable PDF report is generated using ReportLab, containing the complete diagnostic findings and model confidence scores for all five disease classes.

IV. SYSTEM IMPLEMENTATION

The Lungs disease prediction application is developed as a full-stack web-based system that integrates a frontend interface, Flask backend, SQLite database, and TensorFlow deep learning inference service to provide a complete lung disease prediction and monitoring solution. The system is designed to be user-friendly, clinically informative, and deployable on standard hardware.

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 9.927| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

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

4.1 User Authentication

The application provides separate login and registration options for patients and doctors. Passwords are stored securely using SHA-256 hashing with a random salt (secrets.token_hex). Each user is assigned a role-based account; doctor accounts have access to all patient scan records through the doctor dashboard. Session management is handled by Flask's server-side sessions to ensure secure access and controlled data sharing.

4.2 X-Ray Prediction Module

The core prediction module accepts chest X-ray images in JPG, PNG, BMP, WebP, or GIF formats (max 16 MB). The uploaded image is saved to static/uploads/, preprocessed, and passed to the DenseNet201 model via the predict_xray() function in model/predict.py. The system returns the predicted condition, confidence scores for all five disease classes, severity level, clinical description, and a treatment recommendation. All scan records are saved to the SQLite database (Lungs disease prediction.db) for authenticated users.

4.3 Deep Learning Integration

The application integrates a TensorFlow/Keras DenseNet201 model (densenet_xray_model.h5) loaded at Flask startup using a lazy-loading singleton pattern to avoid repeated overhead. Preprocessing exactly matches the training pipeline: images are loaded via keras_image.load_img() at 224×224, converted to NumPy array, divided by 255.0, and expanded to batch dimension (1, 224, 224, 3) before inference. Integer percentage confidence scores are computed and adjusted to sum to 100% for display.

4.4 Symptom Checker Module

The system provides a symptom-based risk assessment where users select symptoms such as fever, dry cough, breathlessness, chest pain, and taste loss. A rule-based scoring engine evaluates symptom combinations and estimates risk level (Low, Moderate, or High) with possible condition suggestions including COVID-19, Viral Pneumonia, Bacterial Pneumonia, and general respiratory infections.

4.5 AI Chatbot Assistant

A rule-based AI chatbot provides lung health guidance, emergency contact information (Ambulance: 102, Emergency Medical: 108, Lung Helpline: 1800-111-565), breathing exercise instructions, and dietary advice. The chatbot responds to queries related to specific conditions including Tuberculosis, COPD, Pneumonia, and COVID-19, and directs users to the doctor portal for professional consultation.

4.6 Doctor Dashboard

The doctor dashboard allows medical professionals to monitor all patient scan records through a role-based account. It displays each patient's predicted condition, severity level, confidence scores, scan date, and clinical recommendations. Statistical analytics including condition distribution, severity breakdown, and monthly scan trends are presented through interactive Chart.js visualizations.

4.7 PDF Report Generation

Upon completing a prediction, the system generates a downloadable medical report in PDF format via the /download-report endpoint using ReportLab. The report includes patient information, diagnosis, severity, clinical findings, treatment recommendations, and a confidence score table for all five disease classes, enabling patients and clinicians to retain a formal record of each diagnostic session.

4.8 System Architecture

The system follows a three-layer architecture:

Frontend: HTML, CSS, and JavaScript with Chart.js, providing an interactive web interface for X-ray upload, results display, and analytics visualizations.

Backend: Flask (Python), handling user authentication, REST API endpoints (/predict, /symptoms, /chat, /download-report), file handling, and session management.

Database: SQLite (Lungs disease prediction.db) stores user accounts (patients and doctors), scan records, confidence scores, severity, and diagnostic metadata.

ML Model: TensorFlow/Keras DenseNet201 (densenet_xray_model.h5) loaded at startup for real-time chest X-ray inference with five-class softmax output.

The above figure represents the overall system architecture of Lungs disease prediction, illustrating the interaction between patient and doctor modules, data storage, deep learning inference model, and result generation pipeline.

V. RESULTS AND DISCUSSION

The Lungs disease prediction system was implemented and tested to evaluate both the functionality of the web application and the classification performance of the DenseNet201 deep learning model. The system successfully integrates user interaction with real-time chest X-ray prediction, enabling effective lung disease screening and patient guidance.

The application allows users to upload a chest X-ray image through the web interface. Based on the uploaded image, the system preprocesses the data and generates a lung disease prediction using the trained DenseNet201 model. The X-ray upload and prediction interface ensures a structured and consistent diagnostic experience for users.

The prediction results generated by the system are displayed in a clear and user-friendly format. The application provides the predicted disease condition along with confidence scores for all five classes, severity level (Low / Moderate / High), clinical description, and personalized treatment recommendations. This helps users understand their lung health condition and take appropriate actions.

In addition to patient interaction, the system includes a doctor monitoring feature. The doctor dashboard displays all patient scan records, predicted conditions, severity levels, and monthly analytics. This feature enables timely clinical review and improves communication between patients and doctors.

The performance of the DenseNet201 model was evaluated using standard classification metrics including accuracy, precision, recall, and F1-score. As shown in Table II, the model achieved strong classification performance across all evaluated architectures.

TABLE II — MODEL PERFORMANCE COMPARISON

Model	Accuracy	Precision	Recall	F1-Score
Vanilla CNN	82%	80%	78%	79%
VGG16 (TL)	89%	88%	87%	87.5%
ResNet50 (TL)	91%	90%	89%	89.5%
CNN+DenseNet201	**94%**	**93%**	**92%**	**92.5%**

The CNN+DenseNet201 model achieves the highest performance across all metrics, with 94% accuracy, 93% precision, 92% recall, and 92.5% F1-score. The dense connectivity pattern of DenseNet201 enables superior feature reuse compared to VGG16 and ResNet50, resulting in better generalization on multi-class chest X-ray classification.

The model achieved strong F1-scores across all five disease classes, indicating its ability to distinguish between COVID-19, Pleural Effusion, Normal, Pneumonia, and Tuberculosis conditions. Early stopping triggered at epoch 38 for the best model, preventing over-training.

Furthermore, the precision-recall curves demonstrate the trade-off between precision and recall for each of the five disease classes. The curves indicate that the model maintains a strong balance between correctly identifying positive cases and minimizing false predictions, which is critical in clinical diagnostic applications.

Overall, the results confirm that the proposed system is capable of providing accurate multi-class lung disease predictions along with a user-friendly web interface for continuous monitoring. The integration of deep learning with comprehensive web-based health guidance features enhances the clinical usability and effectiveness of the system in real-world deployment scenarios.

VI. FUTURE SCOPE

The Lungs disease prediction system can be further improved by incorporating advanced technologies and additional features. The current rule-based chatbot can be enhanced using Natural Language Processing (NLP) techniques to provide more personalized and contextually intelligent responses. Integration with wearable health devices and IoT sensors can facilitate real-time collection of physiological data alongside imaging.

The system can be extended to support Grad-CAM (Gradient-weighted Class Activation Mapping) visualizations, enabling Explainable AI (XAI) overlays on chest X-ray images to highlight the radiological regions contributing to each prediction. This would significantly improve clinical trust and interpretability for radiologists and medical professionals.

Integration of clinical metadata such as patient age, smoking history, and spirometry values alongside imaging data would enable multimodal prediction for improved diagnostic accuracy. Additionally, cloud deployment on AWS or Google Cloud Run and EHR (Electronic Health Record) integration via REST API would make the system accessible at a national healthcare level. These advancements can make Lungs disease prediction more effective in addressing pulmonary health challenges in resource-constrained settings.

VII. CONCLUSION

This paper has presented Lungs disease prediction, a web-based lung disease prediction system designed to support early identification and management of pulmonary conditions. The system integrates chest X-ray image analysis with a DenseNet201 deep learning model to predict lung disease conditions and severity levels in an automated and efficient manner.

The application successfully combines multiple features including X-ray prediction, symptom checker, AI chatbot, breathing exercises, dietary guidance, and a doctor dashboard to provide a comprehensive diagnostic and wellness solution. The use of DenseNet201 with transfer learning enables effective multi-class classification of chest radiographs into five categories: COVID-19, Pleural Effusion, Normal, Pneumonia, and Tuberculosis. The results demonstrate that the system achieves 94% accuracy with a 92.5% F1-score, outperforming Vanilla CNN, VGG16, and ResNet50 baseline architectures.

Overall, the proposed system provides a user-friendly, scalable, and clinically informative approach for early detection and continuous monitoring of lung diseases, with clear pathways for future deployment in real-world healthcare environments.

VIII. ACKNOWLEDGMENT

The authors would like to express their sincere gratitude to Prof. Vishakha Jadhav, Department of Computer Engineering, Dhole Patil College of Engineering, for her valuable guidance, continuous support, and encouragement throughout the development of this project. The authors also thank the faculty members and the institution for providing the necessary resources and support to successfully complete this work.

REFERENCES

[1] World Health Organization, "Global Respiratory Disease Report," WHO, 2023.
[2] Gupta et al., "SVM-based Pneumonia detection from chest X-ray images," J. Medical Imaging, 2022.
[3] Ramesh et al., "Random Forest-based TB prediction from tabular patient data," J. Health Informatics, 2023.
[4] Rajamanickam et al., "CNN model for Pneumonia detection using chest X-rays," IEEE Access, 2022.

[5] Li and Zhao, “Transfer learning with VGG16, ResNet50 for multi-class lung disease detection,” Pattern Recognition Letters, 2023.
[6] Kumar et al., “CNN combined with DenseNet201 for lung disease classification,” Biomedical Engineering Online, 2024.
[7] Chen et al., “Capsule Network for spatial relationship preservation in chest X-rays,” IEEE Trans. Medical Imaging, 2023.
[8] Gao et al., “Hybrid CNN-LSTM for time-dependent lung disease progression,” Medical Image Analysis, 2024.
[9] Mehta et al., “Hybrid CNN-SVM for medical image classification,” Expert Systems with Applications, 2023.
[10] Patil et al., “Ensemble CNN with ResNet, DenseNet and MobileNet for chest X-ray analysis,” Computers in Biology and Medicine, 2024.
[11] Sharma et al., “Stacking ensemble for COPD and TB detection,” J. Healthcare Engineering, 2023.
[12] Rahman et al., “Data augmentation for deep learning in medical imaging,” JMLR, 2022.
[13] Singh et al., “Multi-stage preprocessing for biomedical X-ray classification,” Medical Image Analysis, 2023.
[14] Bhatia et al., “CLAHE contrast enhancement for CNN performance improvement,” Biomedical Signal Processing, 2024.
[15] A. Singh et al., “Grad-CAM visualization for lung X-ray interpretability,” IEEE Trans. Medical Imaging, 2023.
[16] Das et al., “Layer-wise Relevance Propagation for CNN transparency in medical AI,” Frontiers in AI, 2024.
[17] Joshi et al., “AI-driven lifestyle recommendations for diabetes management,” Digital Health, 2023.
[18] Patra et al., “Lifestyle recommendations for respiratory diseases,” Frontiers in Digital Health, 2024.
[19] Wang X. et al., “ChestX-Ray8: Hospital-scale chest X-ray database and benchmarks,” IEEE CVPR, 2017.

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

9940 572 462 6381 907 438 ijircce@gmail.com

www.ijircce.com

Scan to save the contact details