INTERNATIONAL JOURNAL OF NOVEL RESEARCH AND DEVELOPMENT (IJNRD)
© 2026 IJNRD | Volume 11, Issue 2, February 2026 | ISSN: 2456-4184 | IJNRD.ORG

Deep Learning-Based Pneumonia Detection from Chest X-Ray Images

R. Bhuvaneswari[1], S. Snega[2] R. Bhuvaneswari[1], M.Sc., M.Phil., M.Tech., SET.,
Assistant Professor, Department of Computer Science and Applications D.K.M. College for Women (Autonomous)S. Snega [2], PG Student, Department of computer Science and Applications D.K.M. College for Women (Autonomous), Vellore, Tamil Nadu, India.

Abstract—This paper presents a deep learning-based automated system for pneumonia detection from chest X-ray images. Pneumonia is a serious respiratory disease requiring timely diagnosis to reduce morbidity and mortality. The proposed system leverages Convolutional Neural Networks (CNNs) with transfer learning using the pre-trained VGG16 architecture to enhance classification accuracy. Image preprocessing including resizing, normalization, and data augmentation improves model generalization. The system classifies chest X-ray images into Normal and Pneumonia categories, achieving 95.7% overall accuracy. AES-256 and SHA-512 algorithms ensure data security in the Flask-based web application. The system provides real-time clinical decision support and outperforms traditional machine learning baselines including SVM (86.2%) and CNN from scratch (91.3%).

Keywords—*Pneumonia Detection, Convolutional Neural Network, Transfer Learning, VGG16, Chest X-Ray, Deep Learning, Medical Image Classification, Flask Deployment, Data Augmentation, ResNet.*

I. INTRODUCTION

Pneumonia is a serious respiratory infection that inflames the air sacs in one or both lungs and remains one of the leading causes of death globally, particularly among children under five and elderly adults. Accurate and early diagnosis is critical for effective treatment and patient survival. Chest X-ray imaging is the primary diagnostic tool used by clinicians worldwide due to its cost-effectiveness and widespread availability.

Traditional diagnosis relies on radiologists manually examining X-ray images, a process that is time-consuming and susceptible to inter-observer variability and human error. The shortage of experienced radiologists in low- and middle-income countries further exacerbates delayed diagnosis. These limitations motivate the development of automated, AI-driven diagnostic systems that can assist healthcare professionals in making accurate, timely decisions.

Deep learning, particularly Convolutional Neural Networks (CNNs), has demonstrated remarkable success in medical image classification tasks. CNNs automatically learn hierarchical feature representations directly from raw images, eliminating the need for handcrafted feature engineering. Pre-trained architectures such as VGG16, ResNet50, and InceptionV3, trained on large-scale datasets like ImageNet, can be fine-tuned on medical imaging datasets through transfer learning to achieve high classification accuracy even with limited training data.

This paper proposes a CNN-based automated pneumonia detection system trained on a pediatric chest X-ray dataset. The system incorporates robust preprocessing, data augmentation, and VGG16 transfer learning to deliver accurate real-time clinical predictions through a secure Flask-based web application with JWT authentication.

The rest of this paper is organized as follows: Section II surveys related work; Section III analyzes the existing system and proposes improvements; Section IV describes system

VI. IMPLEMENTATION

A. Technology Stack

The system is implemented in Python 3.8 using TensorFlow 2.10 and Keras for deep learning model training and deployment. The VGG16 architecture with ImageNet weights serves as the feature extraction backbone. The custom classification head consists of Global Average Pooling, Dense (512 units, ReLU), Dropout (0.5), and Softmax output layers. Flask serves the web backend with REST API endpoints. Bootstrap 5, HTML5, and CSS3 form the responsive clinical frontend. Werkzeug handles secure file uploads, and Flask-JWT-Extended manages JWT authentication tokens stored in HTTP-only cookies.

B. Training Pipeline

Transfer learning is applied in two phases. In Phase 1, VGG16 base layers are frozen and only the custom classification head is trained for 10 epochs using the Adam optimizer with learning rate 1×10^{-3} and binary cross-entropy loss. In Phase 2, the last four convolutional blocks of VGG16 are unfrozen and the full network is fine-tuned for 20 additional epochs with a reduced learning rate of 1×10^{-5}. Early stopping with patience 5 prevents overfitting. Class weights computed inversely proportional to class frequency address the dataset imbalance.

C. Preprocessing and Augmentation

All images are resized to 224×224 pixels to match VGG16 input dimensions. Pixel values are normalized to [0, 1]. Training-set augmentation includes random horizontal flipping, rotation (±15°), width and height shifts (±10%), zoom range (±10%), and shear transformations. These strategies expand the effective training set size and improve generalization to unseen clinical images. Validation and test sets undergo only resizing and normalization without augmentation.

requirements; Section V presents the system design; Section VI details the implementation; Section VII covers testing and results; Section VIII discusses applications and future scope; and Section IX concludes the paper.

II. LITERATURE SURVEY

Rajpurkar, Irvin, and Ball (2017) introduced CheXNet, a 121-layer DenseNet trained on the ChestX-ray14 dataset, achieving performance comparable to experienced radiologists on the pneumonia detection task in terms of F1-score. The study demonstrated that deep convolutional networks can automatically learn discriminative features from raw X-ray images without manual feature engineering, validating the potential of AI for clinical radiology [1].

Wang, Peng, and Lu (2017) contributed the NIH ChestX-ray14 dataset containing over 100,000 labeled frontal-view chest X-ray images across 14 disease categories including pneumonia. They benchmarked multiple CNN architectures and demonstrated that transfer learning from ImageNet-pre-trained models significantly outperformed training from scratch, establishing a critical baseline for subsequent medical imaging research [2].

Kermany, Goldbaum, and Cai (2018) developed a CNN model specifically trained on pediatric chest X-ray images using transfer learning with ImageNet-pretrained weights. Data augmentation including rotation, flipping, and zooming was applied to reduce overfitting on the limited medical dataset. The model achieved high classification accuracy, demonstrating the critical importance of preprocessing and regularization for medical image classification [3].

Apostolopoulos and Mpesiana (2021) applied transfer learning with VGG19 and InceptionV3 architectures to multi-class chest X-ray classification, differentiating normal, bacterial pneumonia, viral pneumonia, and COVID-19 cases. Fine-tuning pre-trained models on limited medical data achieved high accuracy, highlighting the benefit of transfer learning when annotated medical datasets are scarce [4].

Roy, Chatterjee, and Das (2022) introduced a lightweight CNN architecture with optimized pooling layers specifically designed for deployment in resource-limited settings. The model addressed class imbalance using data augmentation and weighted loss functions, achieving competitive accuracy with significantly reduced computational cost, making it suitable for mobile and edge-based medical diagnostics [5].

Ahmad, Hussain, and Khan (2023) compared Vision Transformer (ViT) models with conventional CNNs for chest X-ray classification. ViTs showed promising performance especially when combined with CNN-based feature extractors, and patch-based image encoding improved focus on diagnostically relevant regions. The study informed the attention-based feature extraction strategy in the proposed system [6].

Lee, Gupta, and Han (2024) integrated attention mechanisms within a CNN framework to highlight critical areas in chest X-rays indicative of pneumonia. The attention layers improved both feature learning from diseased regions and clinical interpretability compared to baseline CNN models, contributing explainability to the diagnostic process [7].

Khan, Ali, and Rehman (2024) proposed a hybrid model combining CNN for spatial feature extraction and RNN for capturing contextual patterns across sequential image slices. This hybrid approach outperformed standalone CNN models,

D. Web Application

The Flask application defines secure routes for user login (/login), logout (/logout), image upload and prediction (/predict), and result display (index). The predict_class() function loads the saved pneumonia.h5 model, preprocesses the uploaded image using MobileNetV2-compatible preprocessing, performs inference, and returns the class label with confidence percentages for both Normal and Pneumonia classes. JWT tokens stored in HTTP-only cookies secure all clinical sessions with 1-hour expiration.

VII. TESTING AND RESULTS

A. Testing Methodology

Six testing strategies were employed. Black Box Testing verified system behavior from the clinician perspective—uploading X-ray images, viewing predictions, and checking confidence scores—without inspecting internal code, ensuring correct input-output behavior. White Box Testing examined internal logic including model inference correctness, preprocessing pipeline, and JWT validation to detect logical errors. Unit Testing validated each module independently: Preprocessing Module, Model Inference Module, Authentication Module, and Web API Endpoints were each tested with defined inputs and expected outputs. Integration Testing verified the complete pipeline from image upload through preprocessing, CNN inference, and frontend result rendering.

B. System Performance

The model achieves 95.7% overall accuracy on the 624-image test set. The pneumonia class achieves precision of 0.95, recall of 0.98, and F1-score of 0.96. The normal class achieves precision of 0.97, recall of 0.93, and F1-score of 0.95. The macro-averaged F1-score of 0.955 confirms balanced performance across both classes. The high recall of 0.98 for pneumonia is clinically critical, ensuring 98% of actual pneumonia cases are correctly identified and minimizing life-threatening false negatives.

Table I. Performance Comparison: Baseline vs. Proposed System

The proposed VGG16 transfer learning model significantly outperforms all baseline methods. SVM with handcrafted features achieved 86.2% accuracy; k-NN achieved 79.4%; CNN trained from scratch achieved 91.3%. The proposed system achieves 95.7% accuracy, 0.98 recall, and 4.5/5 user satisfaction. Average inference time is 180 ms per image on CPU. Full natural language query support and real-time prediction are qualitative advantages unavailable in baseline systems.

C. Functional Test Results

All major system functions passed testing: (1) User authentication—signup, login, and logout with JWT token validation functioned correctly; (2) Image upload with format validation (.jpg, .jpeg, .png) operated as expected; (3) Prediction endpoint returned correctly formatted JSON with label and confidence scores; (4) Both Normal and Pneumonia test images were classified accurately with high confidence; (5) Session management with JWT cookie expiration worked correctly; and (6) Performance testing confirmed average response time below 500 ms under concurrent load.

especially in complex cases where spatial context across image regions mattered for accurate diagnosis [8].

III. SYSTEM ANALYSIS

A. Existing System Limitations

Traditional computer-aided diagnosis (CAD) systems for pneumonia detection relied primarily on manual feature extraction methods such as texture analysis, edge detection, histogram features, and shape-based descriptors. These features were then fed into classical classifiers including Support Vector Machines (SVM), k-Nearest Neighbors (k-NN), Decision Trees, and Random Forest algorithms. These approaches required significant domain expertise for feature selection and were highly sensitive to the quality of handcrafted representations.

Existing systems depend heavily on structured metadata and predefined categories. Classification accuracy decreased substantially when tested on unseen hospital data due to poor generalization. The risk of overfitting was high due to the limited availability of annotated medical imaging datasets. Additionally, traditional systems lacked real-time implementation capability and failed to provide interpretable outputs that clinicians could trust for critical diagnostic decisions.

B. Proposed System

The proposed AI-powered pneumonia detection system overcomes these limitations by incorporating deep learning with transfer learning, eliminating the need for manual feature engineering. Instead of relying on handcrafted descriptors, the system leverages VGG16's convolutional layers pre-trained on ImageNet to automatically extract hierarchical features from chest X-ray images. A custom classification head with dropout regularization prevents overfitting on the limited medical dataset.

The system supports real-time diagnostic prediction through a Flask-based web application with JWT-secured authentication. Clinicians can upload chest X-ray images and receive immediate predictions with class probabilities for Normal and Pneumonia categories. Grad-CAM visualization is supported for future integration to provide attention heatmaps, improving clinical trust and interpretability of AI predictions.

IV. SYSTEM REQUIREMENTS

A. Hardware Requirements

The system requires a processor of Intel Core i5-2450M or higher running at 2.50 GHz or above. A minimum of 8 GB RAM is needed for handling transformer-based preprocessing, model inference, and Flask web server operations concurrently. A hard disk capacity of 250 GB is sufficient for storing the chest X-ray image dataset, trained model weights (.h5 file), uploaded images, and the application codebase.

B. Software Requirements

The development environment uses Anaconda with Jupyter Notebook and VS Code as the IDE. Python 3.8 is the primary programming language. Key libraries include TensorFlow 2.10 and Keras for deep learning model training and inference, NumPy and Pandas for data manipulation, Scikit-learn for evaluation metrics, Matplotlib for visualization, Flask and Flask-JWT-Extended for web deployment, and Werkzeug for secure file handling. The frontend is built with HTML5, CSS3, and Bootstrap 5.

VIII. APPLICATIONS AND FUTURE SCOPE

A. Applications

- Hospital clinical decision support for radiologists to reduce manual workload and accelerate chest X-ray diagnosis turnaround time
- Rural and remote healthcare centers where access to experienced radiologists is limited, enabling automated preliminary diagnosis
- Telemedicine platforms providing remote pneumonia screening for patients in underserved regions with standard mobile or web interfaces
- Public health screening programs for early-stage pneumonia detection in vulnerable populations including infants and elderly patients
- Medical education and training tools for radiology students to practice identifying pneumonia features in chest X-rays

B. Future Enhancements

Several enhancements are planned for future versions. Multi-class classification will extend the system to differentiate between bacterial pneumonia, viral pneumonia, COVID-19, tuberculosis, and other pulmonary conditions from a single X-ray. Integration of Grad-CAM and LIME explainable AI techniques will generate visual heatmaps highlighting diagnostically critical regions, improving clinical trust. A hybrid recommendation model combining CNN spatial features with attention mechanisms will further improve diagnostic sensitivity. Integration of real-time patient data via hospital PACS systems will automate X-ray ingestion. Additionally, a mobile-responsive Progressive Web App (PWA) version is planned for point-of-care deployment.

IX. CONCLUSION

This paper presented a deep learning-based automated pneumonia detection system using Convolutional Neural Networks with VGG16 transfer learning. The system classifies chest X-ray images as Normal or Pneumonia with 95.7% overall accuracy, precision of 0.95, and recall of 0.98 for the pneumonia class. The high recall rate is clinically significant as it minimizes missed diagnoses that could endanger patient lives.

The integration of preprocessing, augmentation, transfer learning, and a Flask-based web application produces a clinically practical, deployable tool that significantly reduces radiologist workload. Results demonstrate that deep learning with transfer learning substantially outperforms traditional machine learning methods for medical image classification. The system demonstrates practical effectiveness in transforming conventional diagnosis workflows into AI-assisted, context-aware clinical decision support that enhances patient outcomes and healthcare efficiency.

REFERENCES

[1] P. Rajpurkar, J. Irvin, and K. Ball, CheXNet: Radiologist-Level Pneumonia Detection on Chest X-Rays with Deep Learning, arXiv:1711.05225, 2017.

[2] X. Wang, Y. Peng, and L. Lu, ChestX-Ray8: Hospital-Scale Chest X-Ray Database and Benchmarks, IEEE CVPR, pp. 2097–2106, 2017.

[3] D. S. Kermany, M. Goldbaum, and W. Cai, Identifying Medical Diagnoses and Treatable Diseases by Image-Based Deep Learning, Cell, vol. 172, no. 5, pp. 1122–1131, 2018.

[4] S. Apostolopoulos and T. Mpesiana, COVID-19: Automatic Detection from X-Ray Images Utilizing Transfer Learning

V. SYSTEM DESIGN

A. System Architecture

The system architecture consists of five integrated components: (1) CineMatch-style Web Interface—a clinician-facing front-end providing authentication, X-ray image upload, and real-time diagnostic results; (2) Flask Backend—REST endpoints for login, logout, image upload, prediction, and confidence display with JWT security; (3) CNN Inference Engine—the VGG16 transfer learning pipeline combining feature extraction with classification; (4) Image Dataset—preprocessed and augmented chest X-ray images used for training and testing; and (5) Authentication Module—JWT-based session management for secure clinical access.

B. Data Flow

The data flow begins when a clinician submits a chest X-ray image through the web interface. The image is sent via POST to the /predict endpoint with a JWT token. The Flask backend validates the token, saves the image to a secure upload directory, and passes it to the preprocessing pipeline. The image is resized to 224×224 pixels and normalized using VGG16-compatible preprocessing. The CNN model generates class probabilities. The predicted label (Normal or Pneumonia), confidence score, and individual class probabilities are returned as JSON and displayed on the interface.

C. Functional Modules

The system is organized into five functional modules. The Preprocessing Module cleans and resizes raw X-ray images, applies normalization and augmentation during training. The Feature Extraction Module leverages VGG16's convolutional blocks to automatically learn spatial patterns indicative of lung infection. The Classification Module applies a custom fully connected head with dropout to produce Normal/Pneumonia probabilities. The Evaluation Module computes accuracy, precision, recall, F1-score, and confusion matrix. The Web Application Module provides the clinical interface with secure upload, real-time prediction, and result visualization.

with Convolutional Neural Networks, Medical and Biological Engineering and Computing, 2021.

[5] S. Roy, P. Chatterjee, and A. Das, Novel Pooling-Based VGG-Lite for Pneumonia and COVID-19 Detection from Imbalanced Chest X-Ray Datasets, arXiv, 2022.

[6] M. Ahmad, F. Hussain, and M. Khan, Automated Diagnosis of Lung Diseases Using Vision Transformer: A Comparative Study on Chest X-Ray Classification, arXiv, 2023.

[7] S. Lee, R. Gupta, and J. Han, Deep CNN with Attention Mechanisms for Chest X-Ray Analysis, IEEE Transactions on Medical Imaging, 2024.

[8] M. Khan, Z. Ali, and A. Rehman, Hybrid CNN-RNN Model for Pneumonia Detection, IEEE Access, 2024.

[9] R. Vyas and D. R. Khadatkar, Ensemble of Deep Learning Architectures for Pneumonia Classification Using Chest X-Rays, Journal of Imaging Informatics in Medicine, vol. 38, no. 2, 2025.

[10] H. Iqbal, M. Tariq, and S. Rafiq, Deep Learning Approaches for Chest Radiograph Interpretation: A Systematic Review, Electronics, vol. 13, no. 23, p. 4688, 2024.

Copyright & License:

© Authors retain the copyright of this article. This work is published under the Creative Commons Attribution 4.0 International License (CC BY 4.0), permitting unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.