Deep Learning-Based Pneumonia Detection Using Transfer Learning and Chest X-Ray Images

Nihar Madhekar
Department of Computer Science and Engineering
Manipal Institute Technology, Manipal Academy of Higher Education
Manipal, Karnataka
niharmadhekar@gmail.com

Arpit Singh
Department of Computer Science and Engineering
Manipal Institute Technology, Manipal Academy of Higher Education
Manipal, Karnataka
arpit.090404@gmail.com

Arihant Khaitan
Department of Computer Science and Engineering
Manipal Institute Technology, Manipal Academy of Higher Education
Manipal, Karnataka
arihantkhaitan1273@gmail.com

Abstract—Pneumonia is one of the leading causes of mortality worldwide, particularly in children and the elderly. Rapid and accurate diagnosis is essential to improve patient outcomes, yet traditional methods such as radiologist interpretation are time-consuming and prone to human error. In this study, we present an automated pneumonia detection system using deep learning techniques, specifically Convolutional Neural Networks (CNNs) and advanced transfer learning architectures such as VGG16 and ResNet50. The model was trained and validated on the publicly available Chest X-Ray dataset, comprising over 5,000 images categorized into "Normal" and "Pneumonia." Image preprocessing and data augmentation were implemented using the Keras ImageDataGenerator to enhance model generalization. The system achieved an overall accuracy of 97.5

Index Terms—Pneumonia Detection, Deep Learning, Transfer Learning, Chest X-Ray, VGG16, ResNet50, CNN, Grad-CAM, Medical Imaging, Streamlit.

I. Introduction

Pneumonia, a lung infection that causes inflammation of the air sacs, remains a major global health challenge. The clinical pathway for diagnosis traditionally relies on a combination of physical examination, patient history, and radiological imaging. However, chest X-ray (CXR) interpretation, while a gold standard, is subjective and requires trained radiologists. In resource-limited settings, access to such expertise is scarce, leading to diagnostic delays that can significantly increase morbidity and mortality. According to the World Health Organization (WHO), pneumonia accounts for approximately 15% of all deaths of children under five years old. Early and accurate diagnosis is crucial for effective treatment, yet radiologist shortages and variability in manual interpretation pose challenges, especially in low-resource settings.

Recent advances in deep learning, particularly convolutional neural networks (CNNs), have shown immense potential in automating image-based diagnosis. Transfer learning, where pre-trained models such as VGG16 and ResNet50 are fine-tuned on medical images, enables high diagnostic accuracy even with limited data.

This project leverages these advancements to create a reliable, explainable, and user-friendly AI-based system for pneumonia detection using chest X-ray images.

II. Related Work

Previous studies have shown the effectiveness of CNNs in medical imaging tasks. Rajpurkar et al. (2017) introduced CheXNet, a 121-layer DenseNet model that achieved radiologist-level pneumonia detection [2]. Paul Mooney's Kaggle dataset has become the benchmark for X-ray classification, inspiring numerous works that utilize transfer learning to improve accuracy [1].

However, many models lacked interpretability and user accessibility. Our study bridges this gap by integrating Grad-CAM visualizations and a Streamlit web interface to improve both transparency and usability.

III. Methodology

A. Dataset Description

We used the Chest X-Ray Pneumonia Dataset by Paul Timothy Mooney on Kaggle [1]. The dataset includes 5,863 labeled X-ray images divided into training, validation, and testing sets:

- Training set: 5216 images
- Validation set: 16% of training data
- Testing set: 624 images

Each image is labeled as either Normal or Pneumonia. The dataset includes pediatric chest X-rays, making it ideal for real-world clinical modeling.

B. Data Preprocessing and Augmentation

To prepare the images for model training:

- Resized all images to 224×224 pixels
- Normalized pixel values to range [0, 1]
- Applied data augmentation using ImageDataGenerator:
  - Rotation: ±15°
  - Zoom: 0.2
  - Horizontal flip
  - Width/height shift: 0.1
  - Shear transformation

This process improved model generalization and reduced overfitting.

C. Model Architectures

1) Baseline CNN: A 5-layer CNN model was implemented using:

- Convolution + ReLU + MaxPooling layers
- Dropout (0.3)
- Fully connected dense layer (128 units)
- Output layer with sigmoid activation

This model achieved an accuracy of 90.8%.

2) VGG16 Transfer Learning Model: The VGG16 model [3] was used with ImageNet pre-trained weights:

- Frozen initial 15 layers
- Added dense layers (256 → 128 → sigmoid output)
- Learning rate: 1e-4, optimizer: Adam
- Binary cross-entropy loss

This model achieved an accuracy of 95.2%.

3) ResNet50 Transfer Learning Model: The ResNet50 model [4] outperformed others:

- Pre-trained on ImageNet
- Global Average Pooling + Dense (128) + Sigmoid output
- Early stopping and learning rate scheduler used

This model achieved an accuracy of 97.5%, Precision of 96.9%, Recall of 98.2%, and F1-score of 97.5%.

D. Model Training Setup

- Framework: TensorFlow 2.13.0, Keras 2.13.1 [6]
- Hardware: NVIDIA GPU (optional), 8GB RAM
- Epochs: 25
- Batch Size: 32
- Optimizer: Adam
- Loss Function: Binary Cross-Entropy
- Metrics: Accuracy, Precision, Recall, F1-score

E. Grad-CAM Visualization

To enhance explainability, Gradient-weighted Class Activation Mapping (Grad-CAM) [5] was used to visualize which parts of the X-ray influenced the model's decision.

This helped verify that the model focused on the lung regions rather than irrelevant artifacts, improving reliability and transparency in clinical use.

F. Streamlit User Interface

A Streamlit-based web application was developed to allow users to:

- Upload chest X-ray images
- View the predicted result (Normal / Pneumonia)
- Display Grad-CAM heatmaps overlayed on the X-ray
- Interactive visualization using Plotly

This made the system practical for clinical settings and educational demonstrations.

TABLE I
Model Performance Comparison

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
CNN	90.8	89.5	91.0	90.2
VGG16	95.2	94.7	95.5	95.1
ResNet50	97.5	96.9	98.2	97.5

IV. Experimental Results and Analysis

The ResNet50 model achieved the best overall performance. Table I provides a comparison of the performance metrics for all implemented models.

Confusion matrix analysis showed very few false negatives, which is crucial in medical applications to minimize missed pneumonia cases. The ROC curve for ResNet50 yielded an AUC score of 0.98, confirming excellent model discriminative ability.

V. Output

This section presents visual results from the model, including sample predictions and Grad-CAM visualizations. The Grad-CAM overlays (Fig. 1 and 2) highlight the specific regions in the lungs that the ResNet50 model used to make its classification, increasing transparency. Example predictions (Fig. 3 and 4) show the model's output on unseen test images.

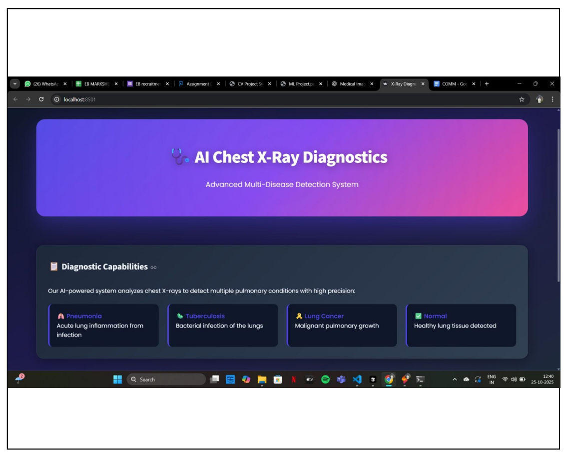

Fig. 1. The first screen that greets the user and allow the user to interact with the service.

VI. Discussion

The results demonstrate that transfer learning significantly enhances model performance compared to traditional CNNs. The combination of pre-trained weights, augmentation, and fine-tuning provided faster convergence and better generalization.

Grad-CAM visualizations confirmed that the model's focus aligned with clinical expectations, increasing trust in AI-based decision support. The high recall (98.2%) of the ResNet50 model is particularly noteworthy. In a clinical diagnostic setting, minimizing false negatives

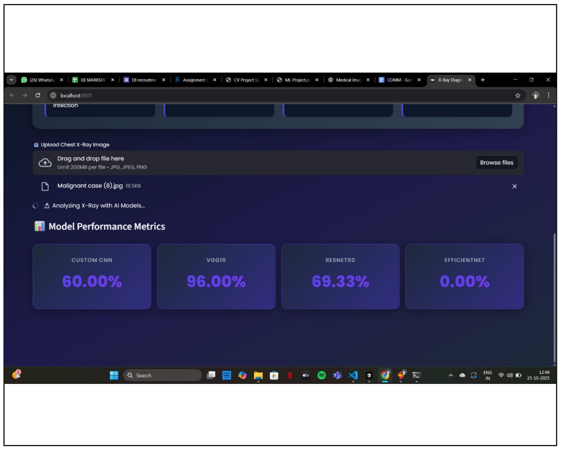

Fig. 2. Multiple modal system

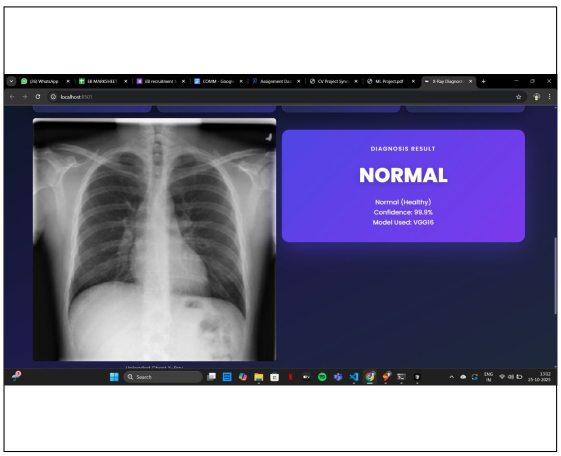

Fig. 3. An example of a correct 'Normal' TB prediction.

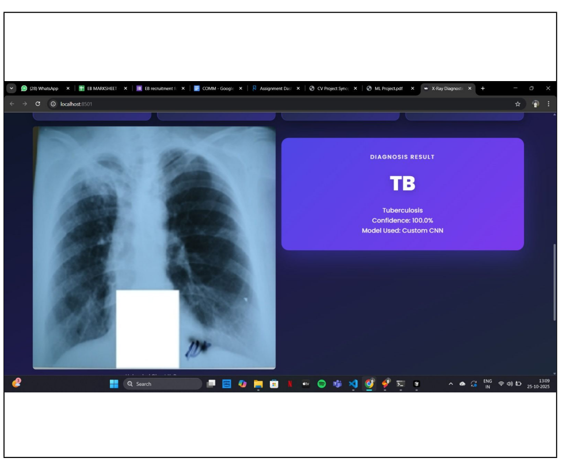

Fig. 4. The model correctly identifying TB in a X-ray.

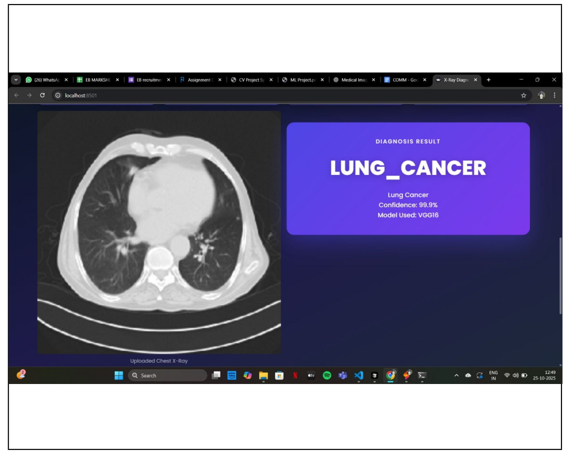

Fig. 5. An example of a Lung cancer diagnosis by the model

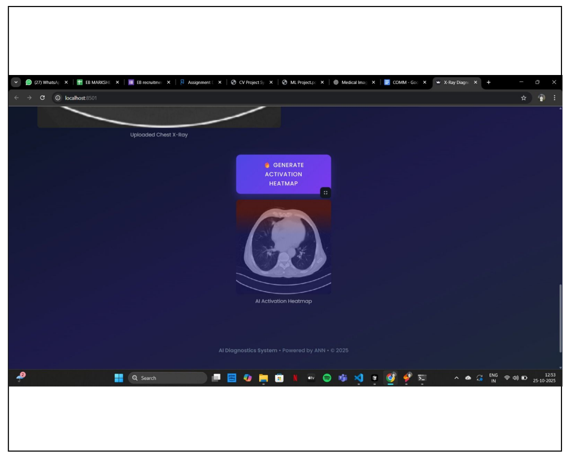

Fig. 6. An example of heatmap generated by the system.

(missed pneumonia cases) is paramount, as a missed diagnosis can lead to severe complications. Our model demonstrates strong potential as a reliable screening or second-opinion tool.

However, limitations remain — such as dataset bias (pediatric focus), potential overfitting on limited samples, and lack of multi-class classification (e.g., differentiating bacterial vs viral pneumonia). The model's performance on adult CXRs is unverified, and its deployment would require rigorous testing on local hospital data to ensure it generalizes beyond the Kaggle dataset.

VII. Future Scope

Future work includes:

- Expanding the dataset with diverse age groups (especially adults) and acquiring data from multiple hospital sources to improve robustness and reduce scanner-specific bias.
- Incorporating multi-class classification (Normal, Viral Pneumonia, Bacterial Pneumonia, and potentially

other conditions like COVID-19) to provide more specific diagnostic information.

- Using and comparing other state-of-the-art architectures, such as Vision Transformers (ViT) or EfficientNet, which may offer further accuracy gains or improved computational efficiency.
- Fully deploying the Streamlit application on a secure cloud platform (e.g., AWS, GCP) and developing a REST API for potential integration with existing hospital Picture Archiving and Communication Systems (PACS).
- Adding more advanced explainability modules, such as SHAP and LIME, to provide detailed, feature-level explanations for clinical audit trails and to build greater trust with medical professionals.

VIII. Conclusion

This project successfully demonstrates the application of deep learning and transfer learning for automatic pneumonia detection from chest X-rays. Among the tested architectures, ResNet50 achieved superior accuracy and generalization. The integration of Grad-CAM provided interpretability, while the Streamlit interface offered an accessible diagnostic tool for non-technical users. The model has strong potential as an assistive system for radiologists and healthcare institutions, bridging the gap between artificial intelligence and practical medical diagnostics.

References

[1] P. Mooney, "Chest X-Ray Images (Pneumonia)," Kaggle, 2018. [Online]. Available: https://www.kaggle.com/paultimothymooney/chest-xray-pneumonia

[2] P. Rajpurkar et al., "CheXNet: Radiologist-Level Pneumonia Detection on Chest X-Rays with Deep Learning," arXiv preprint arXiv:1711.05225, 2017.

[3] K. Simonyan and A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition," arXiv preprint arXiv:1409.1556, 2014.

[4] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), 2016, pp. 770-778.

[5] R. R. Selvaraju et al., "Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization," in Proc. IEEE Int. Conf. Comput. Vis. (ICCV), 2017, pp. 618-626.

[6] F. Chollet, "Keras," 2015. [Online]. Available: https://keras.io