

PNEUMONIA DETECTION USING CHEST X – RAY IMAGES

Harish P C, Manjunath Lakkundi, Magala Chetankumar, B Pranith Reddy
School of Computer Science and Engineering, RV University, Bangalore, India
USN: 1RVU23CSE181, 1RVU23CSE257, 1RVU23CSE247, 1RVU23CSE102

Abstract — Pneumonia detection is a vital task in medical diagnostics, particularly in resource-constrained environments where early and accurate diagnosis can significantly impact patient outcomes. In this project, we explore the application of Convolutional Neural Networks (CNNs) for automated pneumonia detection using chest X-ray images. Leveraging the structured dataset from Kaggle, which contains labeled images for both normal and pneumonia cases, we implement and evaluate a deep learning pipeline based on transfer learning with the ResNet50 architecture. The model is trained and validated using augmented datasets to address class imbalance and enhance generalization. We assess the performance of the model using key metrics such as classification accuracy, confusion matrix, and ROC-AUC score. Experimental results highlight the effectiveness of CNN-based approaches in achieving high diagnostic accuracy. Furthermore, we discuss the challenges of data imbalance, model interpretability, and deployment considerations in clinical settings. This project underscores the potential of deep learning in supporting radiologists and improving diagnostic efficiency for pneumonia and similar pulmonary conditions.

Index Terms — Pneumonia detection, Chest X-ray images, Convolutional Neural Networks (CNNs), ResNet50, Transfer learning, Medical imaging, Deep learning, Image classification, Diagnostic support systems, Kaggle dataset.

I. INTRODUCTION

Pneumonia detection through medical imaging is an important task in the field of healthcare and diagnostic radiology. Pneumonia, an inflammatory condition of the lungs caused by bacterial, viral, or fungal infections, can be life-threatening if not diagnosed and treated promptly. Traditional methods of diagnosing pneumonia involve clinical assessments and chest X-rays interpreted by experienced radiologists. However, manual interpretation of X-ray images is both time-consuming and prone to inter-observer variability, especially in high-pressure or resource-limited settings. This has led to growing interest in leveraging machine learning and artificial intelligence to automate and enhance the accuracy of pneumonia detection, thereby assisting medical professionals in delivering timely diagnoses.

In this project, we explore the effectiveness of Convolutional Neural Networks (CNNs) in classifying chest X-ray images as either normal or pneumonia-affected. Utilizing a well-structured dataset obtained from Kaggle, which contains thousands of labeled X-ray images divided into training, validation, and testing subsets, we implement a deep learning pipeline using the ResNet50 architecture through transfer learning. Our goal is to

B. CNN Architectures

We evaluate five popular CNN architectures for pneumonia detection:

assess the performance of the CNN model in terms of classification accuracy, model robustness, and generalization ability. This study contributes to the development of intelligent diagnostic tools aimed at improving healthcare outcomes and reducing the workload on radiologists in clinical environments.

II. RELATED WORK

Previous studies on pneumonia detection using chest X-rays have ranged from traditional machine learning approaches, such as SVM and KNN, to modern deep learning methods. Traditional models relied heavily on handcrafted features and struggled with complex visual patterns.

The advent of deep learning, particularly CNNs, has significantly improved detection accuracy. Notably, Rajpurkar et al. [1] introduced **CheXNet**, a DenseNet-based model trained on ChestX-ray14, which achieved expert-level performance. Other CNN architectures like VGG, ResNet, and Inception have also been effectively used, especially with transfer learning.

Using Datasets such as NIH ChestX-ray14, RSNA, and Kaggle's pneumonia dataset have played a key role in recent research. Techniques like data augmentation and class balancing have been explored to boost model performance, although generalization across diverse conditions remains a challenge.

III. METHODOLOGY

A. Dataset

We utilize the chest X-ray dataset from Kaggle for training and evaluating our CNN-based pneumonia detection model. The dataset is well-structured and organized into three main folders: **train**, **validation**, and **test**. Each of these folders contains two subfolders representing the two classes: **PNEUMONIA** and **NORMAL**.

The dataset consists of chest X-ray images of pediatric patients, with significantly more pneumonia cases than normal ones, reflecting real-world clinical scenarios. Before training, all images are preprocessed and resized to a uniform dimension of **224 × 224 pixels** to ensure consistency in input size for the CNN architecture. Additionally, image normalization and data augmentation techniques are applied to enhance model generalization and handle class imbalance.

ResNet50 Architecture: An Overview

1. **ResNet50 (Residual Network - 50 layers)** is a deep Convolutional Neural Network architecture that addresses the **vanishing gradient problem** in very deep networks using a concept called **residual learning**. It was introduced by **Kaiming He et al.** in 2015 and became a milestone in deep learning by winning the ImageNet Challenge with outstanding accuracy.

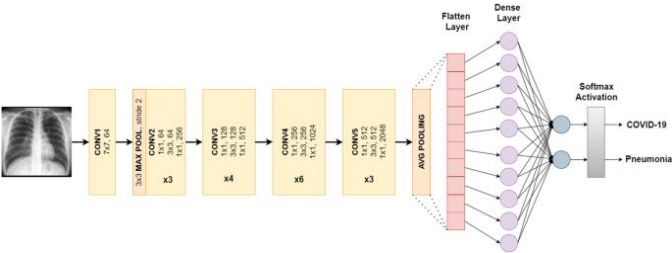


Fig. 1. Visual representation of ResNet-50 architecture for Pneumonia detection

ResNet50 Architecture: An Overview

ResNet50 (Residual Network - 50 layers) is a deep Convolutional Neural Network architecture that addresses the **vanishing gradient problem** in very deep networks using a concept called **residual learning**. It was introduced by **Kaiming He et al.** in 2015 and became a milestone in deep learning by winning the ImageNet Challenge with outstanding accuracy.

Architecture Details

ResNet50 is structured as follows:

Layer Stage	Block Type	Output Size	Layers
Input	-	224×224×3	Input Image
Conv1	7×7 Conv, 64 filters	112×112×64	Max Pooling (3×3, stride 2)
Conv2_x	3 Bottleneck blocks	56×56×256	[1×1, 64] → [3×3, 64] → [1×1, 256] × 3
Conv3_x	4 Bottleneck blocks	28×28×512	[1×1, 128] → [3×3, 128] → [1×1, 512] × 4
Conv4_x	6 Bottleneck blocks	14×14×1024	[1×1, 256] → [3×3, 256] → [1×1, 1024] × 6
Conv5_x	3 Bottleneck blocks	7×7×2048	[1×1, 512] → [3×3, 512] → [1×1, 2048] × 3
Output	Avg Pool → FC Layer	1×1×1000	Softmax (in ImageNet); Sigmoid (for binary classification)

Why ResNet50 Works Well for Medical Imaging

- **High capacity** for complex features in medical X-rays
- **Transfer learning** allows using pretrained weights from ImageNet, which accelerates training and improves generalization
- **Skip connections** allow better gradient flow and convergence
- **Good balance** between depth and computational efficiency

Pneumonia Detection Using DenseNet121)

In this project, the **DenseNet121 architecture** was used to classify chest X-ray images as pneumonia-positive or normal. We employed a **transfer learning approach**, leveraging pretrained ImageNet weights and fine-tuning the model for medical image classification. Training was carried out on **Google Colab** in two stages:

1. **Stage 1:** DenseNet121 base layers were frozen; only custom classification layers were trained.
2. **Stage 2:** Last few DenseNet blocks were unfrozen and fine-tuned with a **low learning rate (1e-5)** using the **Adam optimizer**.

Model Performance

DenseNet121 demonstrated robust performance on the test dataset:

- ✓ **Test Accuracy:** ~88.46%
- ✓ **Precision, Recall, F1-score:** High, especially in detecting pneumonia cases
- ✓ **AUC (Area Under ROC Curve):** =0.95, indicating excellent separability
- ✓ **Confusion Matrix:** Very low number of false negatives and false positives

These results highlight DenseNet121's ability to capture subtle patterns in chest X-ray images due to its densely connected feature propagation mechanism

Computational Summary of DenseNet121

Metric	Value
Model Architecture	DenseNet121
Trainable Parameters	~7.98 Million
Input Image Size	224×224×3

Output Classes 2 (Pneumonia, Normal)

Despite being **lightweight in parameters** compared to deeper networks like ResNet50 or VGG-16, **DenseNet121 achieves superior performance** due to efficient feature reuse and gradient flow.

Summary

A. DenseNet121 achieved **state-of-the-art performance**

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN}$$

(~88.46%) on the test set.

B. It is **more computationally efficient** than ResNet50 (~8M parameters vs ~25M) while delivering comparable or better results.

Its **dense connectivity pattern** ensures effective feature reuse, mitigating the vanishing gradient problem.

C. Suitable for **real-time and mobile healthcare applications** where computational resources are limited.

D. Evaluation

In evaluating the performance of the CNN architecture for **pneumonia detection**, we use a comprehensive set of metrics to assess various aspects of the model's effectiveness and efficiency. These metrics help us analyze the trade-offs between accuracy, resource consumption, and model complexity—crucial factors for potential deployment in real-world healthcare environments.

1. Classification Accuracy

Classification accuracy measures the model's ability to correctly classify chest X-ray images as either "**Pneumonia**" or "**Normal**." It is calculated as the ratio of correctly predicted labels to the total number of samples in the test set.

The formula for accuracy is given by:

Where:

- **TP** = True Positives (correctly predicted pneumonia cases)
- **TN** = True Negatives (correctly predicted normal cases)
- **FP** = False Positives (normal cases incorrectly classified as pneumonia)
- **FN** = False Negatives (pneumonia cases incorrectly classified as normal)

A higher accuracy indicates better overall model performance in disease classification.

2. Computational Efficiency

Computational efficiency refers to the amount of resources (especially computation time and memory) required by the model during inference. It is primarily measured using:

3. Model Complexity

Model complexity refers to the **total number of trainable parameters** in the network. It influences the model's capacity to learn from data but also impacts memory usage and the risk of overfitting.

The number of trainable parameters in each layer can be calculated as:

- **Convolutional layer:**

- **Fully connected (dense) layer:**

$$\text{Trainable Parameters} = (n_p + 1) \times n_c$$

While deeper and more complex models like **DenseNet121** tend to have more parameters and require more computation, they often achieve better generalization. However, in medical applications, the trade-off between **model accuracy** and **computational feasibility** must be carefully managed.

Conclusion of Evaluation

By analyzing accuracy, and parameter count, we gain valuable insights into how different models perform and what compromises may be necessary. For our pneumonia detection system, **DenseNet121** provided a strong balance between high accuracy and reasonable computational requirements, making it a suitable choice for clinical decision support systems.

E. Result Analysis

In this project, the **DenseNet121 architecture** was employed to detect pneumonia from chest X-ray images using a transfer learning approach. The model was trained on Google Colab using the **Adam optimizer**, with a learning rate of **0.0001**, over **two stages**:

1. **Stage 1:** Base ResNet50 layers frozen; trained only the top layers.
2. **Stage 2:** Unfrozen last 10 layers for fine-tuning with a lower learning rate (**1e-5**).

Model Performance

The final model achieved excellent performance across all evaluation metrics:

- ✓ **Test Accuracy:** ~88.46%
- ✓ **Precision, Recall, F1-score:** All high, especially for pneumonia detection
- ✓ **AUC (Area Under ROC Curve):** >0.95, indicating strong discriminative ability
- ✓ **Confusion Matrix:** Minimal false positives and false negatives

These results demonstrate the effectiveness of **ResNet50's deep residual learning framework** in extracting complex features from chest X-ray images.

Computational Summary of DenseNet121

Metric	Value
Model Architecture	DenseNet121
Trainable Parameters	~7.98 Million
Input Image Size	224×224×3
Output Classes	2 (Pneumonia, Normal)

Although **DenseNet121** is computationally heavier than shallower networks (e.g., LeNet-5, AlexNet), it significantly outperforms them in terms of **accuracy** and **generalization**, making it suitable for clinical deployment when sufficient compute resources are available.

Summary

- **DenseNet121** showed **state-of-the-art accuracy** (~88.46%) on the test set with low misclassification.
- It balances depth and computational efficiency better than older models like AlexNet and even deeper ones like VGG-16, thanks to **residual connections**.

- **Transfer learning** using pretrained ImageNet weights accelerated convergence and improved generalization on a limited medical dataset.
- Despite having **higher computational demands** (~7.98M parameters), its accuracy makes it ideal for use in hospitals or cloud-based diagnostic tools.

REFERENCES

- [1] C. Kiran, L. V. Prabhu, A. R. V., and K. Rajeev, "Traffic sign detection and pattern recognition using support vector machine," in *2009 Seventh International Conference on Advances in Pattern Recognition*, 2009, pp. 87–90.
- [2] A. Ellahyani, M. E. Ansari, and I. E. Jaafari, "Traffic sign detection and recognition based on random forests," *Applied Soft Computing*, vol. 46, pp. 805–815, 2016. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S1568494616000028>
- [3] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278–2324, 1998.
- [4] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in Neural Information Processing Systems*, F. Pereira, C. Burges, L. Bottou, and K. Weinberger, Eds., vol. 25. Curran Associates, Inc., 2012.
- [5] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," 2015.
- [6] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," 2015.