
Benchmark Analysis of Deep Learning Models and MedImageInsight Embeddings for Pneumonia Detection from Chest X-Rays

Aldrin Joan Pandian W¹, Saravanan S², Raghavendra Mishra³, Jasmine Selvakumari
Jeya I⁴, Indirani M⁵, J Manikandan⁶

^{1,3}*School of Computing Science Engineering and Artificial Intelligence, VIT Bhopal University, Sehore, Madhya Pradesh, India, aldrinjoan6@gmail.com, mraghavendra86@gmail.com*

²*School of Computing, SRM Institute of Science and Technology, Tiruchirappalli, India, jeyasaraa@gmail.com*

^{4,6}*School of Computing Science and Engineering, VIT Bhopal University, Sehore, Madhya Pradesh, India, wjasminejeya@gmail.com, jdkmani@gmail.com*

⁵*Department of Computer Science and Business Systems, Vel Tech Multi Tech Dr. Rangarajan Dr. Sakunthala Engineering College, Avadi, Chennai, India, indugovind10@gmail.com*

Abstract.

Premature and reliable identification of pneumonia from chest X-ray images is critical for clinical diagnosis and intervention; however, accurate automated classification persists a challenge due to variability in image features and class imbalance. To overcome this challenge, this study measures three approaches used for bi-nary predictions of chest X-ray images (Normal and Pneumonia) using enhanced Kaggle Chest X-Ray Pneumonia dataset. Two transfer learning models, ResNet-50 and DenseNet-121, and an embedding-based prototype classifier using MedImageInsight embedding were optimized using label weighting, early termination and learning rate control. An autonomous dataset comprising 234 normal and 390 pneumonia cases. DenseNet-121 acquired the leading score using 91% accuracy and ROC of 0.96. ResNet-50 obtained a significantly high pneumonia recall of 0.995 but obtained reduced overall accuracy (63.6%) caused by misclassification of normal cases. The embedding-based prototype performed close to random (~50% accuracy), and PCA and t-SNE visualization indicated limited class (Silhouette score = 0.075). These Inferences suggest that deep convolutional feature descriptions are more optimal for pneumonia detection. The proposed evaluation framework can validate the implementation of computer-aided diagnostic frameworks for clinical screening and medical imaging analysis.

Keywords. Pneumonia Detection, Chest X-ray Image Analysis, Deep Learning, Transfer Learning, Computer-Aided Diagnosis (CAD).

1. INTRODUCTION

Pneumonia is a major global health concern, and early detection is essential for effective treatment. Deep learning-based chest X-ray analysis can support diagnosis, especially where radiologists are limited. However, class imbalance and dataset variability can reduce model

reliability. This study compares ResNet-50, DenseNet-121, and MedImageInsight using the Kaggle Chest X-Ray Pneumonia dataset, evaluating both transfer learning and embedding-based classification approaches. PCA and t-SNE are also used to analyze feature representations and class separation.

2. METHODOLOGY

The Kaggle Chest X-ray Pneumonia dataset was used, containing 1,342 normal and 3,876 pneumonia images. Data augmentation increased the normal class to 3,688 images, producing 7,564 balanced samples. Images were resized to 224×224 pixels for CNN training, with 20% validation. An independent test set of 624 images (234 normal, 390 pneumonia) was excluded from training. PCA and t-SNE were used for feature visualization.

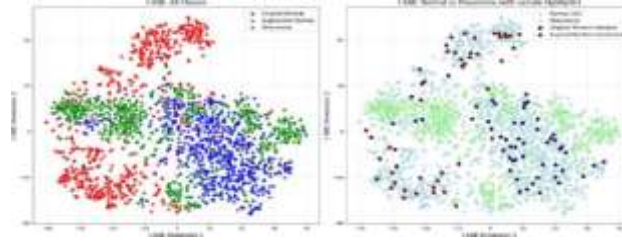


Figure 2.1. t-SNE Feature Space Visualization

Training images were augmented using rotation, flipping, shifting, zoom, shear, and brightness adjustments, with conservative augmentation for MedImageInsight. Class weights were applied to reduce imbalance. ResNet-50 with ImageNet weights used global average pooling, dense layers, normalization, dropout, and sigmoid classification. The backbone was initially frozen and later fine-tuned by unfreezing the last 50 layers with a reduced learning rate.



(a) DenseNet-121

(b) ResNet-50

Figure 2.2. Transformation sequence process of DenseNet-121 and ResNet-50 framework

DenseNet-121 using ImageNet pre-trained network was used as the base model. The framework hybrid the custom prediction head and applied a uniform two-stage training framework, comprising of initial training of the classification head succeeded by particular fine-tuning of the upper layers of the back-bone framework. Figure 2(b) indicates the configuration structure of DenseNet-121 using dense net-work that supports parameter reutilize over layers and enhanced gradient flow. The Vision Transformer (ViT) design module to construct a transformer based image encoder that derived equal length image encoded as 1,024 spatial parameter sequences. We then trained a shallow multilayer

perceptron (MLP) classifier on these embeddings. The MLP had fully connected layers using 512 and 256 units and the last dense layer had a sigmoid activation for binary classification.

Table 1. Outline of prediction result on the unseen sample (Normal: 234, Pneumonia: 390)

Models	Recall	AUC	Accuracy	Precision	F1	Test Loss
	Primary Metrics			Secondary Metrics		
ResNet-50 [2]	0.9949	0.8005	0.6362	0.6330	0.510	1.1090
DenseNet-121 [3]	0.9615	0.9608	0.9103	0.9014	0.9305	0.2276
MedImageInsight (Placeholder) [7]	0.5154	0.5070	0.5064	0.6281	0.510	0.7327

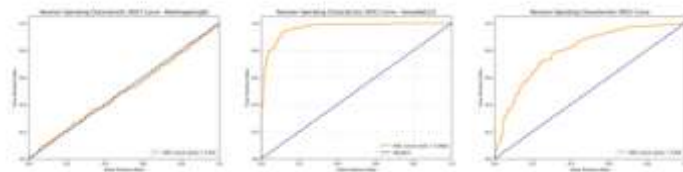


Figure 2.3. Evaluation of ROC performance curves of three framework

Normalization and dropout were used for regularization. Due to the unavailable MedImageInsight checkpoint, placeholder embeddings were used, so its results require validation with authentic embeddings. All models were trained for 50 epochs using Adam, cross-entropy loss, batch size 16, early stopping, and learning-rate reduction. DenseNet-121 achieved the best results, with 91% accuracy, 90% precision, 93% F1-score, and 0.96 ROC-AUC.

3. CONCLUSIONS

This study evaluated pneumonia detection using ResNet-50, DenseNet-121, and MedImageInsight. Fine-tuned CNNs achieved better discrimination than the embedding-based approach, highlighting the importance of deep feature representations and suitable augmentation. Future work should focus on multicenter validation, ensembling, interpretability, and attention-based learning.

4. REFERENCES

- [1] A.Y. Nageye, A.D. Jimale, M.O. Abdullahi, Y.A. Ahmed, M.A. Addow, 'Enhancing deep learning for pneumonia detection: Developing web based solution for Dr. Sumait Hospital in Mogadishu Somalia', Discover Applied Sciences, vol. 7, 2025.
- [2] R. Rajagopal, R. Karthick, P. Meenalochini, T. Kalaichelvi, 'Deep convolutional spiking neural network optimized with arithmetic optimization algorithm for lung disease detection using chest X-ray images', Biomedical Signal Processing and Control, vol. 79, 2023.
- [3] P. Verma and R. Verma, "Predictive analytics for sustainable investment portfolios in banking," Multidisciplinary Science Journal, vol. 7, 2025, Art. no. 2025ss0231.
- [4] M. Mittal, J.P. Gujjar, M.S.G. Prasad, R.M. Devadas, L. Ambreen, V. Kumar, 'Dimensionality reduction using UMAP and TSNE technique', Proc. In 2024

- Second International Conference on Advances in Information Technology (ICAIT), pp. 1-5, 2024.
- [5] S.L. Krishna, I.J.S. Jeya, S.N. Deepa, 'Fuzzy-twin proximal SVM kernel-based deep learning neural network model for hyperspectral image classification', Neural Computing and Applications, vol. 34, pp. 19343-19376, 2022.
 - [6] J.S. Jeya, W.A.J. Pandian, P.K. Mishra, 'A next-generation computing approach for early and accurate lung cancer detection using an integrated hybrid deep learning framework', Proc. In 2025 International Conference on Next Generation Computing Systems (ICNGCS), pp. 1-8, 2005.
 - [7] E. P. Pathrose and P. Verma, "The role of higher education in UAE: Narrowing the gap between employers' expectations and students' perceptions," Electrochemical Society Transactions, vol. 107, no. 1, pp. 4213–4231, 2022.

Biographies



Aldrin Joan Pandian W is an undergraduate student at the School of Computing Science, Engineering and Artificial Intelligence, VIT Bhopal University, Madhya Pradesh. His research interests include Machine Learning, Generative AI and Deep Learning.



Saravanan S is currently an Associate Professor, School of Computing, SRM Institute of Science and Technology, Tiruchirappalli, India. He brings over 20 years of experience in teaching and research. He has an impressive portfolio with more than 30 research articles published. His research interests include Machine Learning, Image Processing and Deep Learning.



Raghavendra Mishra is currently an Assistant Professor with the School of Computing Science Engineering and Artificial Intelligence, VIT Bhopal University, Madhya Pradesh. He brings over 6 years of experience in teaching and research. He has an impressive portfolio

with more than 16 research articles published. His research area are Computer Vision and Explainable Artificial Intelligence.



Jasmine Selvakumari Jeya I is currently a Senior Associate Professor with the School of Computing Science and Engineering, VIT Bhopal University, Madhya Pradesh. She brings over 24 years of experience in teaching and research. She has an impressive portfolio with more than 60 research articles published, 30 patents filed, and two patents granted.



Indirani M is currently an Associate Professor, Department of Computer Science and Business Systems, Vel Tech Multi Tech Dr. Rangarajan Dr. Sakunthala Engineering College, Avadi, Chennai, India. She brings over 21 years of experience in teaching and research. She has an impressive portfolio with more than 20 research articles published. Her research interests include Machine Learning, Image Processing and Deep Learning.



J Manikandan is currently an Associate Professor with the School of Computing Science and Engineering, VIT Bhopal University, Madhya Pradesh. He brings over 10 years of experience in teaching and research. He has an impressive portfolio with more than 17 research articles published. His research area are Machine Learning, Soft Computing and Artificial Intelligence.