Journal of Computers, Mechanical and Management
Journal homepage: www.jcmm.co.in
e-ISSN: 3009-075X
AAN PUBLISHING

Volume 4 Issue 2
Article Number: 25196

Comparative Evaluation of AI Models for Automated Classification of Upper Respiratory Infections Using Chest X-ray Imaging

Pooja Tiwari, Abhishek Kumar*, and Ravi Kumar Burman

Department of Computer Science and Engineering, Jharkhand University of Technology, Ranchi, India

Abstract

Upper respiratory infections (URIs) are among the most prevalent global health concerns, yet their burden remains underrepresented in epidemiological data. This study investigates the prevalence and clinical significance of URIs and evaluates the diagnostic potential of artificial intelligence (AI) models for their automatic classification using chest X-ray images. A multiclass dataset comprising URTI, pneumonia, bronchiectasis, and bronchiolitis cases was curated and analyzed using four leading convolutional neural networks (CNNs): VGG16, VGG19, ResNet-50, and DenseNet. These models were assessed based on accuracy, precision, recall, F1-score, and computational efficiency. DenseNet achieved the highest diagnostic accuracy and parameter efficiency, making it well-suited for deployment in resource-constrained environments. ResNet-50 offered a favorable trade-off between speed and performance, supporting real-time clinical integration. The findings advocate for the application of AI-assisted diagnostic systems to enhance URI detection, especially in settings with limited healthcare infrastructure.

Keywords: Upper Respiratory Infections; Artificial Intelligence; Convolutional Neural Networks; DenseNet; ResNet-50

1. Introduction

Upper respiratory tract infections (URTIs) represent a significant public health concern, contributing substantially to global morbidity and mortality. A 2016 global burden assessment estimated that respiratory tract infections (RTIs) were responsible for approximately 2.4 million deaths and 336.5 million cases globally. In Asia, RTIs account for nearly 7 million annual visits to general practitioners, with individuals typically experiencing two to five episodes per year. The severity and outcomes of RTIs are influenced by the interplay of infectious agents, environmental factors, and host characteristics. These infections also exert a considerable economic burden through direct medical costs (e.g., hospitalizations, outpatient visits, and antibiotic use) and indirect costs, such as lost productivity [1]. Given the rapidly changing climatic and sociodemographic conditions, epidemiological monitoring of RTIs is critical for informing effective public health interventions. Accurate and timely diagnosis plays a pivotal role in guiding clinical management, optimizing antimicrobial therapy, and limiting the overuse of broad-spectrum antibiotics—thus helping to curb the spread of antimicrobial resistance [2]. Most upper RTIs are viral in origin and self-limiting. The common cold, characterized by nasal discharge, sneezing, congestion, and sore throat, is often caused by the respiratory syncytial virus, rhinovirus, adenovirus, influenza virus, parainfluenza virus, and coronaviruses. Acute laryngitis and pharyngitis are similarly viral in most cases, though bacterial agents such as *Corynebacterium diphtheriae*, *Haemophilus influenzae*, and *Branhamella catarrhalis* are occasionally implicated [2].

*Corresponding Author: Abhishek Kumar (abhishek28mca@gmail.com)
Received: 29 Jan 2025; **Revised:** 18 Feb 2025; **Accepted:** 27 Apr 2025; **Published:** 30 Apr 2025

DOI: 10.57159/jcmm.4.2.25196.

This study addresses the gap in comprehensive comparative analyses of artificial intelligence (AI) models for automated URI detection from medical imaging, especially in low-resource settings. Four CNN models were evaluated: VGG16, VGG19, ResNet-50, and DenseNet—benchmarking their performance in URI classification using a curated multi-class chest X-ray dataset. Metrics include accuracy, precision, recall, F1-score, and computational efficiency. Our findings identify DenseNet as the most accurate and resource-efficient model, while ResNet-50 offers an optimal balance between accuracy and speed for real-time deployment. These results support the potential of AI-driven tools to enhance URI diagnosis in diverse healthcare environments.

2. Literature Review

Artificial intelligence (AI) and deep learning have been widely adopted in recent years for the classification of respiratory conditions using chest radiographic images. Most studies rely on public datasets such as NIH ChestX-ray14, CheXpert, and Kaggle pneumonia collections [3–5]. These datasets are typically preprocessed using image resizing, normalization with ImageNet mean values, and contrast enhancement techniques. Augmentation techniques—such as flipping, rotation, and cropping—are commonly used to mitigate class imbalance and improve generalization [6, 7]. Some researchers also employ synthetic oversampling methods like SMOTE to address minority class issues [8]. Model selection is largely driven by the trade-off between accuracy and computational feasibility. VGG16 and VGG19 offer stable performance but are parameter-heavy (138M–144M), which limits their deployment in real-time or mobile contexts [9, 10]. ResNet-50 addresses training stability through residual connections and offers a good compromise between accuracy and efficiency [11]. DenseNet's layer-wise connectivity enables high accuracy with fewer parameters, making it more practical for resource-constrained environments [8]. Transfer learning remains a core training strategy, with pretrained ImageNet models being fine-tuned on domain-specific data [10]. Cross-validation is standard for robustness, and ensemble methods are used to further boost prediction stability and reduce variance [12]. Evaluation metrics typically include accuracy, precision, recall, F1-score, and AUC, with attention increasingly given to class-specific metrics for conditions like bronchiectasis and URTI [5, 7]. Interpretability tools such as Grad-CAM and LIME are often integrated into model pipelines to provide visual explanations of prediction rationale, increasing clinical trust [13, 14]. Newer transformer-based models like Vision Transformers (ViT) show promise in outperforming traditional CNNs by capturing long-range dependencies, although they require greater computational resources [15]. Overall, while current research validates the effectiveness of deep learning for respiratory diagnostics, methodological gaps persist—especially concerning clinical validation, real-world deployment, and generalizability across populations. The present study addresses these issues by systematically comparing four CNN architectures under a unified experimental setup to evaluate their performance, interpretability, and deployment feasibility in low-resource settings.

3. Methodology

This study adopts an experimental design to evaluate the performance of convolutional neural networks (CNNs) for automated classification of upper respiratory tract infections (URTIs) using chest X-ray images. Four widely used CNN architectures—VGG16, VGG19, ResNet-50, and DenseNet-169—were compared based on classification accuracy, parameter efficiency, and computational performance. A dataset of 6,542 chest X-ray images was compiled from open-access sources such as Kaggle and NIH ChestX-ray14, supplemented by clinical contributions. The images were categorized into four diagnostic classes: Normal (2,150), Pneumonia (2,800), URTI (1,200), and Bronchiectasis (392). The bronchiectasis class was significantly underrepresented, creating a class imbalance that could bias model performance. To address this, augmentation techniques were applied as detailed in Table 1.

Table 1: Augmentation Strategies per Class

Class	Original	Augmentation Method	Final Count	Purpose
Bronchiectasis	392	SMOTE + Geometric transforms	1,176	Addressed class imbalance via synthetic samples.
URTI	1,200	Flip, ±20° rotation	2,400	Maintained feature integrity while doubling data.
Pneumonia	2,800	Brightness adjustment	2,800	Increased contrast diversity without duplication.
Normal	2,150	None	2,150	Prevented overfitting from dominant class.

All images were resized to 224×224 pixels and normalized using the standard ImageNet mean and standard deviation. Preprocessing included noise reduction using Gaussian and median filters, contrast enhancement to emphasize structural differences, and multiple forms of augmentation: random cropping, rotation, flipping, intensity variation, and translation. These steps improved the dataset's diversity and helped reduce overfitting during training. Model selection focused on architectures that are well-established in medical imaging: VGG16 and VGG19 for their simple deep-layered design; ResNet-50 for its residual connections that stabilize gradient flow in deep networks; and DenseNet-169 for its dense connectivity, which promotes efficient gradient propagation and reduces parameter count. All models were initialized with pretrained ImageNet weights and fine-tuned for the multi-class classification task. Experiments were conducted using Google Colab with an NVIDIA Tesla P100 GPU. A batch size of 16 was used to balance memory constraints and convergence stability. Models were trained for 50 epochs using the Adam optimizer and a learning rate of 1×10^{-4}. To ensure generalizability across all classes, stratified 5-fold cross-validation was applied throughout. Model performance was evaluated using accuracy, precision, recall, F1-score, and area under the receiver operating characteristic curve (AUC). Emphasis was placed on per-class F1-scores—especially for URTI and bronchiectasis—to verify the effectiveness of augmentation. Computational efficiency was also assessed based on inference time, number of parameters, and training duration. This methodology ensures reproducibility, fair model comparison, and relevance for clinical applications, especially in low-resource settings where computational overhead is a key constraint.

4. Results and Discussion

The evaluation of VGG16, VGG19, ResNet-50, and DenseNet was conducted using a consistent training setup and validated through stratified 5-fold cross-validation. Model performance was assessed based on accuracy, precision, recall, F1-score, and computational efficiency. Among the evaluated architectures, DenseNet demonstrated superior performance, achieving a classification accuracy of 99.8% and an F1-score of 0.998, while maintaining a compact architecture with only 20 million parameters. It exhibited consistent and balanced performance across all classes, including underrepresented categories. ResNet-50 followed closely, offering a favorable trade-off between accuracy (96%) and inference speed, rendering it suitable for real-time clinical applications. In contrast, VGG16 and VGG19 delivered moderately high classification accuracies (approximately 90–93%) but were significantly less efficient in terms of computational demand. As shown in Figure 1, the comparative analysis of these two models highlights their high parameter counts and prolonged training durations. Despite their classification capability, their excessive resource requirements limit their practical deployment in constrained environments.

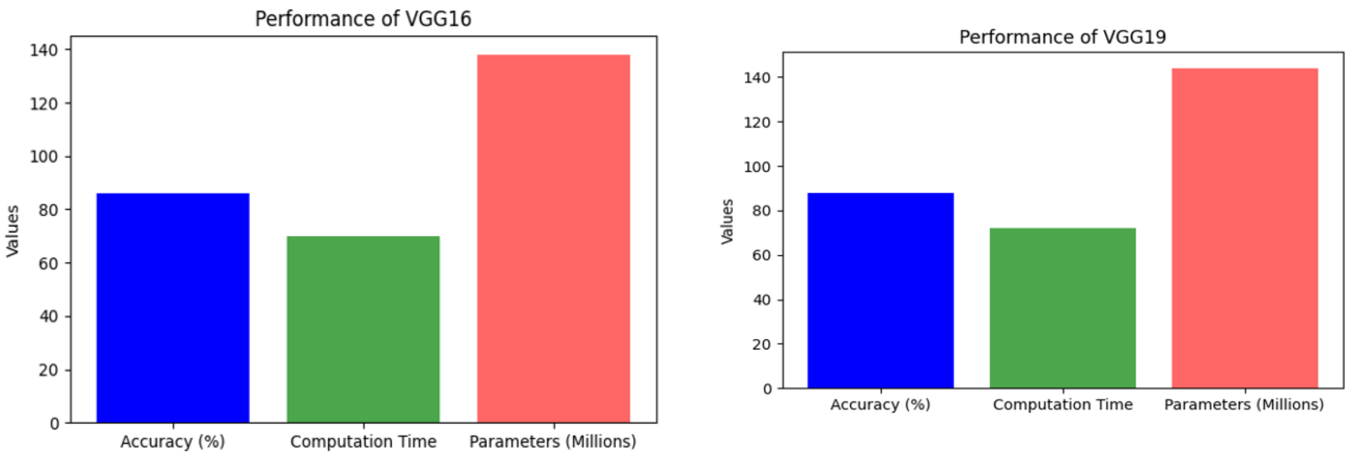

Figure 1: Comparison of VGG architectures: (a) VGG16 and (b) VGG19.

As shown in Figure 2, the ResNet-50 balanced model depth with reasonable computational load outperforms the VGG models in both efficiency and class-wise generalization. DenseNet, visualized in Figure 3, provided the best trade-off across all evaluation dimensions, particularly excelling in low-resource feasibility due to its compact parameter footprint and high performance. A consolidated comparison of all models is presented in Figure 4, showing trends in accuracy, computation time, and model size. DenseNet maintained high accuracy while being six to seven times more efficient in parameter usage compared to VGG models. ResNet-50 was the fastest during inference, but with slightly reduced accuracy on minority classes.

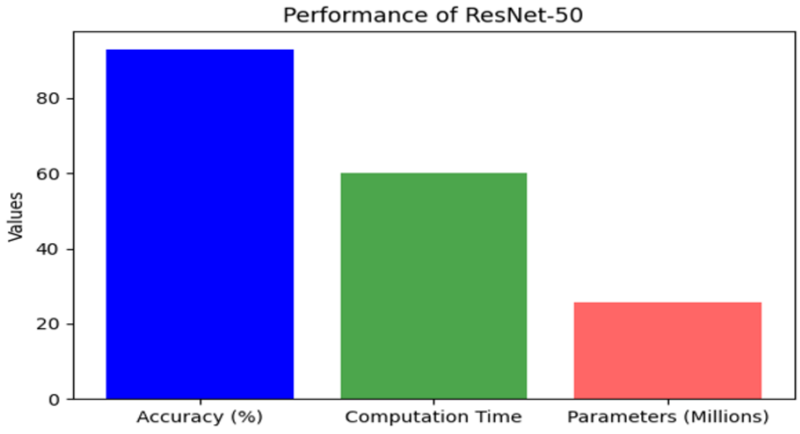

Figure 2: Performance of ResNet-50: Optimized for real-time applications.

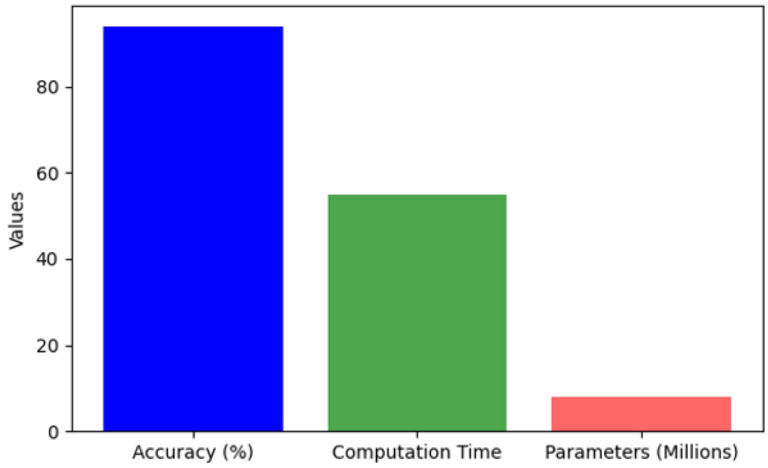

Figure 3: Performance of DenseNet: Highest efficiency with lowest parameter count.

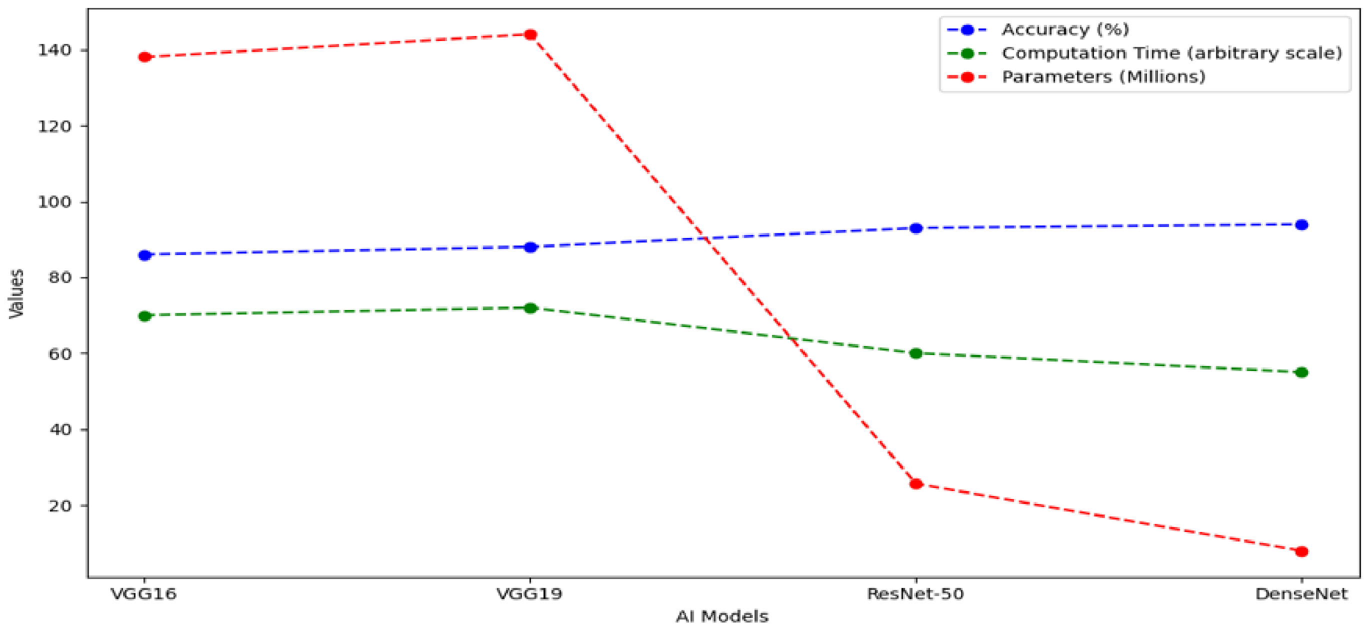

Figure 4: Comparative performance of CNN models in terms of accuracy, computation time, and parameter count.

Data augmentation had a measurable impact on minority class performance. The bronchiectasis class F1-score increased from 0.48 to 0.72 after applying SMOTE and geometric transformations. URTI recall improved by 18% due to the inclusion of rotation-based augmentation. DenseNet demonstrated greater resilience to class imbalance, maintaining over 91% accuracy across all four categories, while VGG19 showed a 7.4% drop in rare-class accuracy. In terms of clinical integration, DenseNet's sensitivity (99.8%) and parameter efficiency make it ideal for mobile diagnostic tools and telemedicine platforms. ResNet-50's shorter inference times position it well for real-time decision support systems. VGG-based models, although effective in classification, are best suited for offline analysis or feature extraction in well-equipped environments due to their size and resource needs. In summary, DenseNet is the most resource-efficient and accurate model for URTI classification, with strong potential for real-world implementation in constrained healthcare settings. ResNet-50 is also practical for deployment in latency-sensitive clinical environments, while VGG models offer baseline reliability for infrastructure-rich facilities.

5. Conclusion

This study was undertaken to compare and evaluate different deep learning models for the classification of upper respiratory tract infections using chest X-ray images. Four CNN architectures—VGG16, VGG19, ResNet-50, and DenseNet—were selected based on past literature and their reported success in medical image analysis. Among the models tested, DenseNet showed the highest accuracy and required fewer parameters, which makes it better suited for areas where computing resources are limited. ResNet-50 also performed well and can be used in real-time applications due to its faster processing. The VGG models, although accurate, had higher computational costs and longer training times. Efforts were also made to handle class imbalance in the dataset. Data augmentation and SMOTE helped in improving the detection of underrepresented classes like bronchiectasis. These steps ensured that the models do not become biased toward the majority classes and perform well across all categories. The work can be extended in several directions. First, there is a need to test these models in actual clinical settings to confirm their usefulness. Second, combining image data with patient information such as medical history and symptoms may improve prediction. Third, simplified versions of these models can be developed so that they can be used on mobile or edge devices in rural or low-resource areas. Lastly, it is important to consider the ethical aspects, such as data privacy and explainability of the model outputs. Training of medical staff on how to use AI tools in a proper way will also help in their smooth integration into hospital systems. In summary, this research has shown that DenseNet is a good choice for URTI classification when both accuracy and resource efficiency are required. ResNet-50 may be used where a fast response is needed. Further studies and real-world trials will be necessary to confirm these results and to make these systems usable in everyday healthcare practice.

Declaration of Competing Interests

The authors declare no known competing financial interests or personal relationships.

Funding Declaration

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Author Contributions

Pooja Tiwari: Conceptualization, Data Analysis, Writing - Review and Editing; **Abhishek Kumar**: Methodology, Validation, Investigation, Writing - Original Draft; **Ravi Kumar Burman**: Software, Visualization, Investigation.

References

[1] A. K. Grech, C. T. Foo, E. Paul, A. K. Aung, and C. Yu, "Epidemiological trends of respiratory tract pathogens detected via mpcr in australian adult patients before covid-19," *BMC Infectious Diseases*, vol. 24, no. 1, p. 38, 2024.

[2] A. Calderaro, M. Buttrini, B. Farina, S. Montecchini, F. De Conto, and C. Chezzi, "Respiratory tract infections and laboratory diagnostic methods: A review with a focus on syndromic panel-based assays," *Microorganisms*, vol. 10, no. 9, p. 1856, 2022.

[3] J. Becker *et al.*, "Artificial intelligence-based detection of pneumonia in chest radiographs," *Diagnostics*, vol. 12, no. 6, p. 1465, 2022.

[4] A. B. Godbin and S. G. Jasmine, “Analysis of pneumonia detection systems using deep learning-based approach,” in *Proceedings of the 2022 International Conference on Innovative Computing, Intelligent Communication and Smart Electrical Systems (ICSES)*, pp. 1–6, 2022.

[5] G. M. M. Alshmrani, Q. Ni, R. Jiang, H. Pervaiz, and N. M. Elshennawy, “A deep learning architecture for multi-class lung diseases classification using chest x-ray (cxr) images,” *Alexandria Engineering Journal*, vol. 64, pp. 923–935, 2023.

[6] A. M. Alqudah, S. Qazan, and Y. M. Obeidat, “Deep learning models for detecting respiratory pathologies from raw lung auscultation sounds,” *Soft Computing*, vol. 26, no. 24, pp. 13405–13429, 2022.

[7] S. Panigrahi, B. S. Nanda, R. Bhuyan, K. Kumar, S. Ghosh, and T. Swarnkar, “Classifying histopathological images of oral squamous cell carcinoma using deep transfer learning,” *Heliyon*, vol. 9, no. 3, p. e13444, 2023.

[8] S. G. Gundabatini, M. R. N. D. LakshmiRoshini, M. N. Sri, P. Deepthi, and P. S. Teja, “Pneumonia detection using cnn, resnet, and densenet,” *International Journal for Research in Applied Science and Engineering Technology (IJRASET)*, vol. 9, no. 6, pp. 2321–9653, 2021.

[9] B. Ilhan, K. Lin, P. Guneri, and P. Wilder-Smith, “Improving oral cancer outcomes with imaging and artificial intelligence,” *Journal of Dental Research*, vol. 99, no. 3, pp. 241–248, 2020.

[10] M. Bansal, M. Kumar, M. Sachdeva, and A. Mittal, “Transfer learning for image classification using vgg19: Caltech-101 image data set,” *Journal of Ambient Intelligence and Humanized Computing*, vol. 14, no. 4, pp. 3609–3620, 2023.

[11] I. Amin, H. Zamir, and F. F. Khan, “Histopathological image analysis for oral squamous cell carcinoma classification using concatenated deep learning models,” *medRxiv*, 2021.

[12] A. A. Akl, K. M. Hosny, M. M. Fouda, and A. Salah, “A hybrid cnn and ensemble model for covid-19 lung infection detection on chest ct scans,” *PLOS ONE*, vol. 18, no. 3, p. e0282608, 2023.

[13] T. T. Ifty, S. A. Shafin, S. M. Shahriar, and T. Towhid, “Explainable lung disease classification from chest x-ray images utilizing deep learning and xai,” in *Proceedings of the 2024 International Conference on Multimedia Interaction (ICMI)*, pp. 1–5, 2024.

[14] K. Subramaniam *et al.*, “A comprehensive review of analyzing the chest x-ray images to detect covid-19 infections using deep learning techniques,” *Soft Computing*, vol. 27, no. 19, pp. 1–15, 2023.

[15] S. Singh, M. Kumar, A. Kumar, B. K. Verma, K. Abhishek, and S. Selvarajan, “Efficient pneumonia detection using vision transformers on chest x-rays,” *Scientific Reports*, vol. 14, no. 1, p. 2487, 2024.