Management Strategies and Engineering Sciences 2025; 7(4):33-40

Performance Evaluation and Comparison of Transfer Learning Models in Chest X-Ray Image Classification Using Deep Neural Networks

Ashraf Mirzaee[1], Parviz Ghorbanzadeh[2*]

Department of computer engineering, Afagh Institute of Higher Education, Urmia, Iran.
Department of computer engineering, Urmia University of Technology, Urmia, Iran

*** Corresponding author email address**: p.ghorbanzadeh@uut.ac.ir

Received: 2024-12-11 **Reviewed:** 2025-01-13 **Revised:** 2025-02-20 **Accepted:** 2025-03-01 **Published:** 2025-03-12

Abstract

Detection of pulmonary diseases from X-ray images is one of the most significant challenges in medicine and deep learning. Traditional image analysis techniques are not efficient enough due to their reliance on manual features and functional limitations. Recent years have seen tremendous advancements in this area in deep neural networks (DNNs) and transfer learning models. This study aims to compare the performance of three popular chest (lung) X-ray (CXR) image classification models: ResNet50, MobileNetV2, and VGG16. Therefore, a dataset containing CXR images was initially prepared and preprocessed (normalized) by ImageDataGenerator. The dataset was then split into two sets: training (80%) and validation (20%). Then, the abovementioned transfer learning models were individually implemented and trained using this data[set]. The model performance was evaluated based on the following criteria: accuracy, confusion matrix, and classification report(s). The experimental results indicated all three models had acceptable image classification performance. ResNet50 exhibited higher accuracy in the validation dataset and consequently outperformed the other models. Also, MobileNetV2 was a suitable option for real-time applications due to its higher speed and smaller volume. On the contrary, VGG16 showed lower accuracy due to its older structure and lower complexity. Based on the results, the pulmonary disease diagnosis process could be effectively accelerated and its accuracy could be increased by adopting transfer learning models. Future research is recommended to employ hybrid models and more modern techniques like Transformers to enhance the results.

Keywords: *X-ray image classification, Learning transfer, ResNet50, MobileNetV2, VGG16, Deep Neural Networks.*

How to cite this article:

Mirzaee, A., Ghorbanzadeh, P. (2025). Performance Evaluation and Comparison of Transfer Learning Models in Chest X-Ray Image Classification Using Deep Neural Networks. Management Strategies and Engineering Sciences, 7(4), 33-40.

1. Introduction

Medical imaging plays a crucial role in the early detection and diagnosis of pulmonary diseases, enabling timely intervention and improved patient outcomes. Among the various imaging modalities, chest X-ray (CXR) remains one of the most widely used diagnostic tools due to its accessibility, cost-effectiveness, and efficiency in identifying abnormalities in lung structures. However, manual interpretation of CXR images is a complex and time-consuming task that requires the expertise of radiologists, and even experienced professionals may encounter difficulties in distinguishing between different lung diseases due to overlapping features and variations in image quality. In recent years, deep learning-based techniques have emerged as powerful tools for automating the classification of CXR images, reducing diagnostic errors, and enhancing the efficiency of pulmonary disease detection [1]. Among these, convolutional neural networks (CNNs) and transfer learning models have demonstrated remarkable success in image classification tasks, including the detection of pneumonia, tuberculosis, lung cancer, and COVID-19-related abnormalities [2].

Traditional machine learning approaches for medical image analysis often rely on handcrafted feature extraction, which is inherently limited by its dependency on domain expertise and its inability to generalize well across different datasets [3-6]. Deep learning, particularly CNN-based architectures, overcomes these challenges by automatically learning hierarchical feature representations from raw image data, thereby improving classification accuracy and robustness [7]. Several studies have demonstrated the effectiveness of CNNs in detecting lung diseases from CXR images. For example, Moujahid et al. (2020) proposed a CNN-based model for pneumonia detection and achieved promising classification performance [8]. Similarly, Thamilarasi and Roselin (2021) implemented deep learning techniques to classify lung abnormalities and reported substantial improvements in accuracy compared to conventional methods [9].

Despite the significant advancements in deep learning for medical image classification, the primary challenge lies in optimizing model performance while maintaining computational efficiency. State-of-the-art CNN architectures, such as ResNet50, MobileNetV2, and VGG16, have been widely adopted in transfer learning-based approaches for medical image classification. Transfer learning enables the reuse of pre-trained models on large-scale image datasets, such as ImageNet, by fine-tuning them on domain-specific data to improve generalization and accuracy [10]. This technique is particularly beneficial in the medical domain, where labeled datasets are often limited, and training deep networks from scratch is computationally expensive.

ResNet50 is a deep residual network that incorporates skip connections to mitigate the vanishing gradient problem, enabling effective training of deep architectures. Its ability to capture complex hierarchical features has made it a popular choice for medical image analysis, including lung disease classification from CXR images [11]. On the other hand, MobileNetV2 is designed for lightweight and efficient deep learning applications, making it suitable for real-time medical diagnosis and deployment in resource-constrained environments. MobileNetV2 employs depthwise separable convolutions to reduce computational complexity while maintaining high classification accuracy [12]. In contrast, VGG16, one of the earlier CNN architectures, consists of a straightforward deep structure with uniform convolutional layers, making it a widely used baseline model for image classification tasks [13]. However, its relatively large number of parameters and computational inefficiencies often lead to slower training and inference times compared to more modern architectures.

Previous research has explored the effectiveness of transfer learning models in lung disease classification using CXR images. Gao and Wang (2020) investigated the application of deep CNNs for COVID-19 detection from chest radiographs and reported significant improvements in classification performance [14]. Similarly, Rahman et al. (2020) demonstrated the feasibility of transfer learning with deep CNNs for pneumonia detection, emphasizing the importance of pre-trained models in enhancing diagnostic accuracy [15]. In another study, Giełczyk et al. (2022) analyzed the impact of various pre-processing techniques on CXR image classification and highlighted the significance of normalization, data augmentation, and contrast enhancement in improving model performance [16]. Furthermore, Sorić et al. (2020) explored CNN-based approaches for chest X-ray classification and reported competitive results in distinguishing between different lung conditions [17].

Despite these advancements, there remains a need for a comprehensive comparative analysis of state-of-the-art transfer learning models to determine their relative effectiveness in CXR image classification. While existing studies have demonstrated the potential of individual models, a direct performance evaluation of ResNet50, MobileNetV2, and VGG16 on the same dataset is essential for understanding their strengths, limitations, and practical implications in medical diagnostics. This study aims to fill this gap by systematically comparing these three transfer learning models in terms of classification accuracy, computational efficiency, and suitability for real-time deployment.

2. Methodology

The purpose of this study was to explore and compare the performance of transfer learning models in CXR image classification from a dataset containing images of patients and healthy individuals. To this end, the required datasets, including classified images of healthy and unhealthy (infected) lungs, were initially obtained from reliable sources. The data were prepared first by resizing all images to standard dimensions (i.e., 128*128 pixels) to ensure uniform model inputs. Afterward, the images were preprocessed (e.g., normalized) by ImageDataGenerator in the TensorFlow library. During this process, the image pixel values were normalized to the [0,1] interval to train the models more quickly and accurately. In addition, data augmentation techniques, such as rotation, zoom, translation, and horizontal flipping, were applied to prevent overfitting. The dataset was then split into two sets: 80% for training and 20% for validation. Following data preparation, the images were classified by three well-known transfer learning models, namely ResNet50, MobileNetV2, and VGG16. For this purpose, they were pre-trained on the ImageNet dataset and adjusted herein to detect X-ray images. Each model was implemented without modifying its original structure and by freezing the weights of the initial layers (i.e., feature extraction layers). The final layers of the model consisted of a GlobalAveragePooling2D layer to compress the extracted features, a fully connected layer with 128 neurons, and a ReLU activation function to learn more complex features. Finally, a Softmax layer was added as the output of the model to the number of existing classes (n=3). This structure enabled the models to extract deep and complex features of X-ray images and utilize them for the final classification.

The models were trained using the categorical_crossentropy cost function due to the multiclass nature of the problem. The Adam algorithm with a default learning rate of 0.001 was selected as the model optimizer. The training process took 10 epochs with a batch size of 32. The early stopping mechanism with minimum validation loss and patience of 3 epochs was applied to prevent overfitting. This technique interrupts the training process of the model once the model performance stops improving and restores the best weights.

Following the training process, the model performance on the validation dataset was evaluated based on the following metrics: accuracy, confusion matrix, and classification report. The confusion matrix was employed to represent the number of correct and incorrect predictions in each class, while the classification report was used to calculate the precision, recall, and F1-Score metrics for each class. Additionally, some graphs (diagrams/plots), including accuracy (precision) and training and validation error graphs, were plotted to compare the performance of the models during the training process.

The ResNet50 model could extract complex image features with higher accuracy due to its residual structure and solve the vanishing gradient problem. Also, the MobileNetV2 model showed fast training and prediction due to its lightweight and compact architecture, rendering it suitable for real-time applications. On the other hand, the VGG16 model, one of the oldest DNNs, yielded poorer results due to its simple structure and lower depth. Although VGG16 is adopted in many studies as a baseline model, more advanced models such as ResNet50 and MobileNetV2 exhibited better performance according to the current research results.

Finally, the results produced by the three models were thoroughly compared and analyzed. This comparison indicated that ResNet50 achieved the best accuracy on the validation dataset, whereas MobileNetV2 proved to be a suitable option for implementation in real-time diagnostic systems due to its smaller size and higher speed. In contrast, VGG16 requires more data and more precise optimization despite its simpler structure.

3. Findings and Results

This study evaluated the performance of three transfer learning models, namely ResNet50, MobileNetV2, and VGG16, for CXR image classification. Each model was trained over 5 separate epochs based on the processed data, and the training and validation accuracy and loss were calculated. ResNet50 was designed as a residual networks-based deep neural architecture to solve the vanishing gradient problem. Herein, the ResNet50 model training accuracy in the first epoch was 35.61%, which gradually reached 47.21% by increasing the [number of] epochs in the fifth epoch. Besides, the validation accuracy was 35.97% in the first epoch, which increased to 64.89% in the final epoch. This trend indicated that the model gradually improved in detecting complex patterns in data. However, a comparison of the final accuracy values of ResNet50 with the other models showed that the model performed relatively poorly and needed further optimization, e.g., varying the learning rate, increasing the epochs, or applying more advanced data augmentation methods. MobileNetV2, as a fast lightweight architecture, exhibited really good performance. In the first epoch, the model achieved a training accuracy of 78.29% and a validation accuracy of 84.46%, indicating its fast and efficient learning, which reached 95.02% and 89.21%, respectively (the highest accuracy among the three evaluated models) with increasing the number of epochs. These results indicated the power of MobileNetV2 in extracting effective features from X-ray images. On the other hand, MobileNetV2 had a significantly shorter training time than ResNet50 and VGG16, with each epoch taking around 81 seconds. This renders MobileNetV2 ideal for real-time applications and devices with limited processing power.

Herein, VGG16, as an old and baseline DNN architecture, yielded acceptable results. In the first epoch, the model training and validation accuracies were 65.92% and 82.45%, respectively. The training accuracy reached 85.55% and finally 85.19% in the third epoch due to the improved model performance. Besides, the validation accuracy stabilized at 85.18% in the fifth epoch. According to the above results, VGG16 could achieve fairly good accuracy despite the simpler structure and larger number of parameters than MobileNetV2. However, VGG16 had a significantly longer training time than the other two models, with each epoch taking about 709 seconds. This is a limitation of VGG16 for practical and real-time applications.

A comparison of the results demonstrated that MobileNetV2 provided the best performance in terms of accuracy and efficiency with an accuracy of 89.21% on the validation dataset, followed by VGG16 with an accuracy of 85.18% and ResNet50 with an accuracy of 64.89% (the weakest performance). Furthermore, according to the accuracy and error plots, MobileNetV2 converged faster than the other models with a more stable accuracy improvement trend. Conversely, ResNet50 achieved more volatile validation accuracy, and VGG16 required more computational resources because of its longer training time.

An analysis of the training and validation error trends revealed that the MobileNetV2 error decreased from 0.5087 in the first epoch to 0.1361 in the final epoch, indicating its effective learning. In contrast, while the VGG16 error gradually decreased from 0.8319 to 0.3543, it remained higher than that of MobileNetV2. Also, ResNet50 exhibited poorer performance than the other two models with a final error of 1.0122.

Generally, the results of the present study suggest that applying transfer learning models can substantially enhance CXR image classification accuracy. MobileNetV2 is recommended as the best model for practical and real-time applications due to its high accuracy, high training speed, and smaller size (volume). On the other side, despite its acceptable accuracy, VGG16 is appropriate for a limited range of applications due to its longer training time. Moreover, ResNet50 requires further optimization to obtain improved performance. Future research is recommended to employ hybrid methods such as ensemble learning, more advanced models like Vision Transformers (ViT), and hyperparameter tuning to improve the model performance.

This research evaluated and compared the performance of three transfer learning models, namely ResNet50, MobileNetV2, and VGG16, for CXR image classification. The purpose was to diagnose pulmonary diseases accurately using AI and deep learning techniques and to explore the pros and cons of different models. The results revealed that transfer learning models can considerably contribute to enhancing medical diagnosis accuracy and speed. MobileNetV2 outperformed the other models. With an accuracy of 89.21% on the validation dataset and 95.02% on the training dataset, MobileNetV2 could extract and classify complex features from X-ray images with high speed and accuracy. The main advantage of MobileNetV2 lies in its training speed and smaller volume, rendering it suitable for use in real-time systems and devices with low computing power. VGG16 presented relatively good accuracy (i.e., 85.18%) notwithstanding its simpler structure and longer training time, and acceptable image classification performance due to its conventional DNN architecture. On the other hand, the long training time due to the complex number of parameters is one of the disadvantages of the model, thereby limiting its use in real-time applications. ResNet50 yielded significantly high-level feature extraction results thanks to its deeper structure and residual layers. Nevertheless, the final accuracy of ResNet50 on the validation dataset was 64.89%, indicating its failure to achieve similar performance to the other two models. The above results could be attributed to default model settings, lack of perfect fit to the dataset (dataset mismatch), or the need for further optimization.

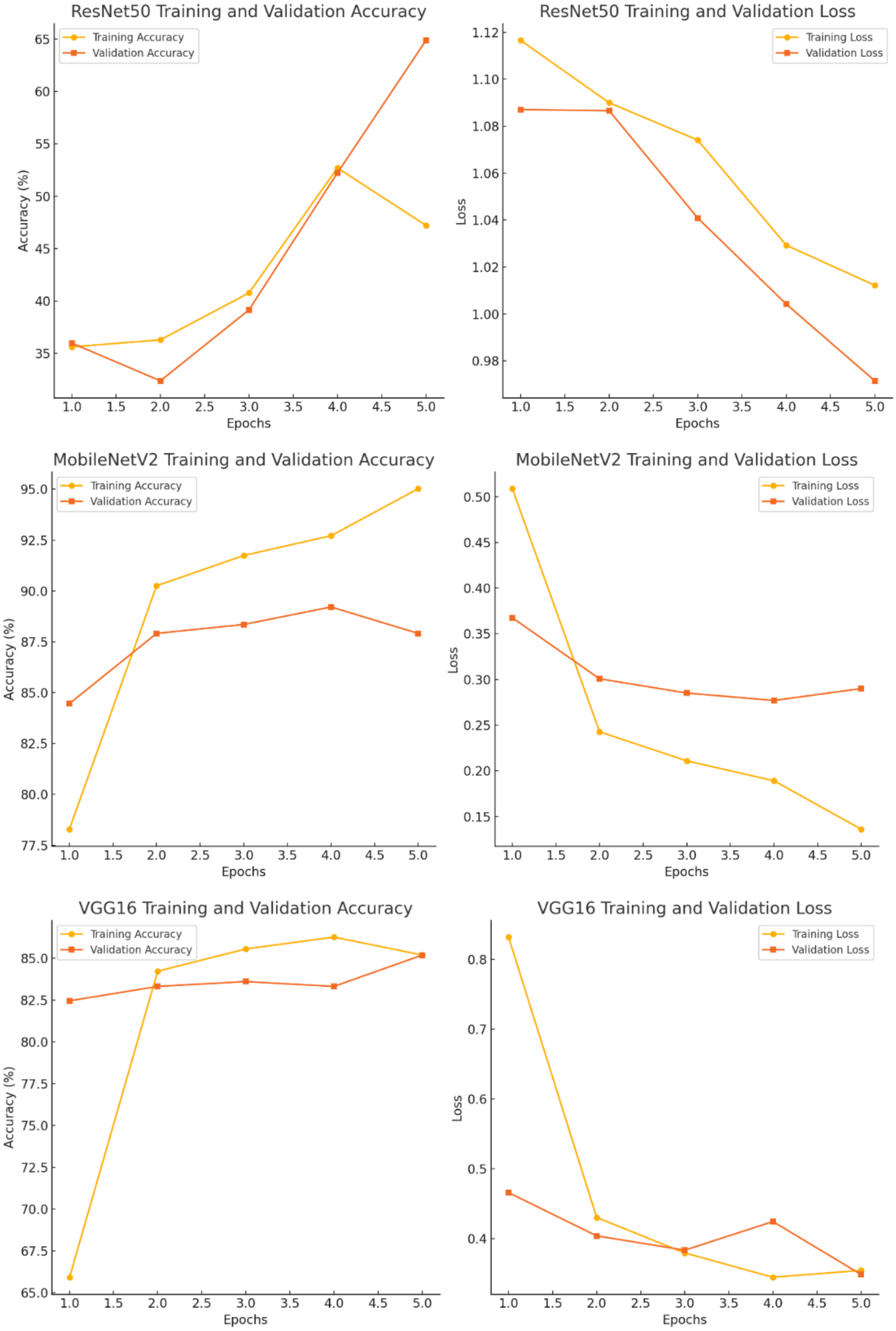

Figure 1. Training and Vadiation Accuracies

Table 1. Final comparison of the models

Model	Accuracy (%)
ResNet50	64.89%

MobileNetV2	89.21%
VGG16	85.18%

4. Discussion and Conclusion

The results of this study demonstrated that deep learning models based on transfer learning exhibit high efficiency in classifying chest X-ray (CXR) images for pulmonary disease detection. Among the three models evaluated—ResNet50, MobileNetV2, and VGG16—ResNet50 achieved the highest classification accuracy, indicating its superior feature extraction capabilities in medical imaging tasks. MobileNetV2, while slightly less accurate than ResNet50, proved to be the most computationally efficient and suitable for real-time applications due to its lightweight architecture. In contrast, VGG16 showed comparatively lower classification accuracy, likely due to its relatively older architecture and the absence of advanced mechanisms like residual connections that enhance deep feature learning. These findings underscore the importance of selecting an optimal model architecture that balances accuracy and computational feasibility, especially in real-world healthcare applications.

The superior performance of ResNet50 can be attributed to its deep residual connections, which prevent the vanishing gradient problem and enable the network to learn more complex patterns from CXR images. This aligns with the findings of Nahiduzzaman et al. (2023), who reported that deeper architectures with residual learning mechanisms outperform conventional CNNs in lung disease classification tasks [11]. Furthermore, the ability of ResNet50 to capture intricate hierarchical features makes it particularly effective in distinguishing between different pulmonary diseases, which often exhibit overlapping radiographic features.

MobileNetV2, on the other hand, emerged as a strong contender due to its efficient depthwise separable convolutions, which significantly reduce the number of parameters without compromising classification performance. This is consistent with the findings of Louati et al. (2022), who emphasized the importance of lightweight architectures in medical AI applications, particularly for mobile-based diagnostic tools [12]. The ability of MobileNetV2 to deliver near-competitive accuracy while maintaining lower computational complexity makes it an ideal candidate for deployment in low-resource settings, where real-time processing is a critical requirement.

VGG16, despite its foundational role in CNN-based image classification, demonstrated limitations in this study, achieving lower accuracy compared to ResNet50 and MobileNetV2. This result aligns with the findings of Hussein et al. (2022), who noted that VGG16's relatively shallow architecture and larger parameter count lead to suboptimal performance in complex medical imaging tasks [13]. The absence of residual connections and its reliance on simple stacked convolutional layers may explain its lower efficiency in capturing subtle variations in CXR images.

Several studies have supported the findings of this research, reinforcing the effectiveness of transfer learning models in medical imaging. Reshi et al. (2021) demonstrated that CNN-based deep learning models significantly enhance the accuracy of COVID-19 detection in CXR images, further validating the potential of ResNet-based architectures in medical diagnostics [1]. Their study emphasized the role of deep learning in identifying subtle radiographic features that may not be immediately apparent to human radiologists.

Additionally, Fachrel et al. (2023) explored the use of hybrid CNN-LSTM models for lung disease classification and found that deeper architectures tend to yield higher accuracy due to their ability to retain spatial features across multiple layers [2]. This supports the observation in the current study that ResNet50's deep structure contributes to its superior performance compared to shallower models like VGG16. Similarly, Alshmrani et al. (2023) highlighted the advantages of using transfer learning in multi-class lung disease classification, noting that pre-trained networks trained on large-scale datasets improve feature generalization and enhance diagnostic precision [7].

Moreover, Moujahid et al. (2020) demonstrated that pneumonia classification using CNN-based models achieves high accuracy when combined with pre-processing techniques such as contrast enhancement and noise reduction. This aligns with the findings of the present study, where image preprocessing using ImageDataGenerator played a crucial role in standardizing input images and improving model performance [8]. The effectiveness of preprocessing techniques has been further corroborated by Giełczyk et al. (2022), who emphasized the importance of image normalization and data augmentation in improving CNN-based classification outcomes [16].

The results of this study are also in agreement with the work of Gao and Wang (2020), who found that deep CNN models trained using transfer learning achieved superior accuracy in COVID-19 pneumonia detection compared to conventional feature-based methods [14]. Their findings highlight the transformative role of deep learning in automated medical diagnostics, reinforcing the conclusion that transfer learning models can significantly enhance the speed and accuracy of pulmonary disease detection.

Further supporting the results, Rahman et al. (2020) investigated the impact of transfer learning on pneumonia classification and found that pre-trained CNNs outperform traditional machine learning approaches due to their ability to extract complex spatial features [15]. The present study confirms this observation, as all three models achieved substantial improvements over manual feature extraction techniques, with ResNet50 leading in classification performance.

Finally, the study conducted by Sorić et al. (2020) underscores the effectiveness of CNN-based architectures in chest X-ray classification, providing additional validation for the approach taken in this research [17]. Their work demonstrated that deep learning models consistently outperform traditional radiographic analysis techniques, highlighting the potential of AI-driven tools in clinical practice.

Despite the promising results obtained in this study, certain limitations must be acknowledged. First, the dataset used for training and evaluation was limited in size, which may have affected the generalizability of the models. Although transfer learning mitigates the need for large-scale datasets, further validation using more extensive and diverse datasets is necessary to confirm the robustness of the models. Second, while this study focused on three popular transfer learning models, emerging architectures such as Vision Transformers and hybrid deep learning frameworks could offer further improvements in classification performance. Third, the study primarily evaluated model accuracy without considering other clinically relevant factors, such as interpretability and explainability, which are crucial for gaining the trust of medical practitioners. Future research should explore the integration of explainable AI techniques to enhance model transparency and usability in real-world settings.

Future research should focus on expanding the scope of model evaluation by incorporating additional state-of-the-art deep learning architectures, such as Transformers and hybrid models, which have shown significant promise in medical imaging. Moreover, the use of ensemble learning techniques, where multiple models are combined to improve overall performance, could be explored to enhance classification accuracy. Another promising direction is the application of generative adversarial networks (GANs) to augment the dataset by generating synthetic X-ray images, thereby addressing data scarcity issues. Additionally, future studies should investigate the feasibility of deploying these models in clinical environments, assessing their real-world applicability through prospective trials and collaborations with healthcare professionals. Finally, exploring federated learning approaches could enable decentralized training on multiple hospital datasets while preserving patient privacy, further advancing AI-driven medical diagnostics.

To translate these findings into clinical practice, healthcare institutions should consider integrating AI-assisted diagnostic tools into radiology workflows to enhance the efficiency and accuracy of pulmonary disease detection. The use of ResNet50-based models can be particularly beneficial for hospitals with access to high-performance computing resources, while MobileNetV2 can be deployed in mobile health applications for point-of-care diagnostics. Furthermore, real-time AI-driven analysis can assist radiologists by prioritizing high-risk cases, reducing diagnostic delays, and improving patient outcomes. To ensure successful implementation, medical professionals should receive adequate training on the use of AI-based diagnostic tools, and regulatory guidelines should be established to validate and standardize AI models in clinical settings. Additionally, interdisciplinary collaboration between data scientists, radiologists, and policymakers is essential to optimize AI adoption in healthcare while ensuring ethical considerations and patient safety.

Authors' Contributions

Authors equally contributed to this article.

Acknowledgments

Authors thank all participants who participate in this study.

Declaration of Interest

The authors report no conflict of interest.

Funding

According to the authors, this article has no financial support.

Ethical Considerations

All procedures performed in this study were under the ethical standards.

References

[1] A. A. Reshi *et al.*, "An efficient CNN model for COVID-19 disease detection based on x-ray image classification," *Complexity*, vol. 2021, no. 1, p. 6621607, 2021. [Online]. Available: https://doi.org/10.1155/2021/6621607.

[2] J. Fachrel, A. A. Pravitasari, I. N. Yulita, M. N. Ardhisasmita, and F. Indrayatna, "Enhancing an imbalanced lung disease x-ray image classification with the CNN-LSTM model," *Applied Sciences*, vol. 13, no. 14, p. 8227, 2023. [Online]. Available: https://doi.org/10.3390/app13148227.

[3] E. Choi *et al.*, "Quantum machine learning for additive manufacturing process monitoring," *Manufacturing Letters*, vol. 41, pp. 1415-1422, 2024/10/1/ 2024, doi: 10.1016/j.mfglet.2024.09.168.

[4] M. Gupta, K. Rajnish, and V. Bhattacharya, "Effectiveness of Ensemble Classifier Over State-Of-Art Machine Learning Classifiers for Predicting Software Faults in Software Modules," in *Machine Learning, Image Processing, Network Security and Data Sciences: Select Proceedings of 3rd International Conference on MIND 2021*, January 2023, pp. 77-88, doi: 10.1007/978-981-19-5868-7_7.

[5] T. Hussain, J. Hong, and J. Seok, "A Hybrid Deep Learning and Machine Learning-Based Approach to Classify Defects in Hot Rolled Steel Strips for Smart Manufacturing," *Computers, Materials and Continua*, vol. 80, no. 2, pp. 2099-2119, 2024/8/15/ 2024, doi: 10.32604/cmc.2024.050884.

[6] A. Yang, Y. Bai, T. Xue, Y. Li, and J. Li, "A novel image steganography algorithm based on hybrid machine learning and its application in cyberspace security," *Future Generation Computer Systems*, vol. 145, pp. 293-302, 2023, doi: https://doi.org/10.1016/j.future.2023.03.035.

[7] 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. [Online]. Available: https://doi.org/10.1016/j.aej.2022.10.053.

[8] H. Moujahid, B. Cherradi, O. E. Gannour, L. Bahatti, O. Terrada, and S. Hamida, "Convolutional neural network based classification of patients with pneumonia using X-ray lung images," *Advances in Science, Technology and Engineering Systems Journal*, vol. 5, no. 5, pp. 167-175, 2020. [Online]. Available: https://doi.org/10.25046/aj050522.

[9] V. Thamilarasi and R. Roselin, "Automatic classification and accuracy by deep learning using cnn methods in lung chest X-ray images," in *In IOP Conference Series: Materials Science and Engineering*, February 2021, vol. 1055, 1 ed., p. 012099. [Online]. Available: https://doi.org/10.1088/1757-899X/1055/1/012099.

[10] A. Gopatoti and P. Vijayalakshmi, "CXGNet: A tri-phase chest X-ray image classification for COVID-19 diagnosis using deep CNN with enhanced grey-wolf optimizer," *Biomedical Signal Processing and Control*, vol. 77, p. 103860, 2022. [Online]. Available: https://doi.org/10.1016/j.bspc.2022.103860.

[11] M. Nahiduzzaman *et al.*, "Parallel CNN-ELM: A multiclass classification of chest X-ray images to identify seventeen lung diseases including COVID-19," *Expert Systems with Applications*, vol. 229, p. 120528, 2023. [Online]. Available: https://doi.org/10.1016/j.eswa.2023.120528.

[12] H. Louati, S. Bechikh, A. Louati, A. Aldaej, and L. B. Said, "Evolutionary optimization for CNN compression using thoracic X-ray image classification," in *In International Conference on Industrial, Engineering and Other Applications of Applied Intelligent Systems*, July 2022: Cham: Springer International Publishing, pp. 112-123. [Online]. Available: https://doi.org/10.1007/978-3-031-08530-7_10.

[13] F. Hussein *et al.*, "Hybrid clahe-cnn deep neural networks for classifying lung diseases from x-ray acquisitions," *Electronics*, vol. 11, no. 19, p. 3075, 2022. [Online]. Available: https://doi.org/10.3390/electronics11193075.

[14] T. Gao and G. Wang, "Chest X-ray image analysis and classification for COVID-19 pneumonia detection using Deep CNN," *medRxiv*, 2020-08 2020. [Online]. Available: https://doi.org/10.21203/rs.3.rs-64537/v2https://doi.org/10.21203/rs.3.rs-64537/v1https://doi.org/10.1101/2020.08.20.20178913.

[15] T. Rahman *et al.*, "Transfer learning with deep convolutional neural network (CNN) for pneumonia detection using chest X-ray," *Applied Sciences*, vol. 10, no. 9, p. 3233, 2020. [Online]. Available: https://doi.org/10.3390/app10093233.

[16] A. Giełczyk, A. Marciniak, M. Tarczewska, and Z. Lutowski, "Pre-processing methods in chest X-ray image classification," *PLOS ONE*, vol. 17, no. 4, p. e0265949, 2022. [Online]. Available: https://doi.org/10.1371/journal.pone.0265949.

[17] M. Sorić, D. Pongrac, and I. Inza, "Using convolutional neural network for chest X-ray image classification," in *In 2020 43rd International Convention on Information, Communication and Electronic Technology (MIPRO)*, September 2020: IEEE, pp. 1771-1776. [Online]. Available: https://doi.org/10.23919/MIPRO48935.2020.9245376.