



*Qualitative Study on Impact of EfficientNet-Based Deep Transfer Learning Model for ... Usman, O. P., Et. al. (2026)*

## **Qualitative Study on Impact of EfficientNet-Based Deep Transfer Learning Model for Pneumonia Detection with Explainable Artificial Intelligence using Chest Radiographs**

**USMAN, Opeyemi Lateef<sup>1</sup>**

[usmanol@tasued.edu.ng](mailto:usmanol@tasued.edu.ng)

+2348031910783

**ADEUSI, Oluwasegun Opeoluwa<sup>2</sup>**

[adeusioo@tasce.edu.ng](mailto:adeusioo@tasce.edu.ng)

+2347033364004

**KAREEM, Morufat Adebola<sup>1</sup>**

[kareemma@tasued.edu.ng](mailto:kareemma@tasued.edu.ng)

+2348063678722

**OWOADE, Akeem Ayoade<sup>1</sup>**

[owoadeaa@tasued.edu.ng](mailto:owoadeaa@tasued.edu.ng)

+2347036226048

**MUNIYANDI, Ravie Chandren<sup>3</sup>**

[Ravie.chandren@uniten.edu.my](mailto:Ravie.chandren@uniten.edu.my)

+60123249577

<sup>1</sup>Department of Computer and Information Science, Tai Solarin Federal University of Education, Ijagun, Ogun State, Nigeria.

<sup>2</sup>Department of Computer Science, Sikiru Adetona College of Education, Science and Technology, Omu-Ajose, Ogun State, Nigeria.

<sup>3</sup>College of Computing and Informatics, Universiti Tenaga Nasional, Jalan IKRAM-UNITEN, 43000 Kajang, Malaysia.

### **Abstract**

*Pneumonia continues to constitute a significant global public health burden, with a disproportionate impact on children, older adults, and immunocompromised populations. Although chest radiography is routinely employed for diagnosis, manual image interpretation is frequently constrained by subjectivity, heavy clinical workloads, and inter-observer variability. This study proposes an EfficientNet-based deep transfer learning framework for automated pneumonia detection from chest radiographs, incorporating an explainable artificial intelligence (XAI) mechanism to improve model interpretability. The model was trained and validated using the widely adopted Kaggle pediatric chest X-ray dataset under multiple train-validation-test partitioning schemes (80:10:10, 70:15:15, 60:20:20, and 75:15:10). Model performance was evaluated using accuracy, precision, recall, F1-score, and the area under the receiver operating characteristic curve (ROC-AUC). In addition, Gradient-weighted Class Activation Mapping (Grad-CAM) was applied to generate visual explanations of the model's predictions. The proposed approach demonstrated consistently strong performance across all data splits, with the 60:20:20 configuration yielding the highest accuracy (90.5%), recall (99.0%), and F1-score (94.9%). Despite the high accuracy and recall, the ROC-AUC scores were moderate (0.68–0.73), suggesting residual challenges in effectively distinguishing pneumonia cases from normal images. Grad-CAM visualizations indicated that the model predominantly focused on clinically meaningful pulmonary regions, thereby enhancing decision transparency and fostering clinical confidence. Overall, the findings indicate that the integration of EfficientNet, transfer learning, and explainable AI provides a robust and interpretable solution for pneumonia detection from chest radiographs. Future research will*



*Qualitative Study on Impact of EfficientNet-Based Deep Transfer Learning Model for ... Usman, O. P., Et. al. (2026)*

*aim to validate the model on external datasets and investigate more advanced interpretability techniques to further strengthen diagnostic reliability.*

**Keywords:** Pneumonia detection, Transfer learning, EfficientNet, Chest radiographs, Explainable AI

## Introduction

Pneumonia remains a leading global health concern, causing significant morbidity and mortality, particularly among children under five, the elderly, and immunocompromised populations (World Health Organization, 2022; Zech et al., 2018). Chest radiography (CXR) is widely used as a first-line diagnostic tool due to its rapid acquisition and accessibility. However, interpretation of CXRs requires substantial expertise and is often subject to inter-observer variability and diagnostic overlap with other pulmonary conditions, such as tuberculosis and COVID-19 (Hashmi et al., 2020; Narin et al., 2020). These challenges are particularly pronounced in low-resource settings, where shortages of skilled radiologists can delay diagnosis and treatment, exacerbating patient outcomes (Hosny et al., 2018; Mehmood et al. 2022). Deep learning, specifically convolutional neural networks (CNNs), has emerged as a powerful approach for medical image analysis by enabling automated feature extraction and classification directly from image data (Litjens et al., 2017; Shen et al., 2017; Chouhan et al., 2020). Transfer learning, which leverages pretrained networks such as ImageNet models, has proven particularly effective for medical imaging tasks where annotated data are limited (Tan & Le, 2019; Morid et al., 2020; Mwaniki, 2023). Recent studies demonstrate that deep transfer learning approaches achieve high accuracy in pneumonia classification, with models such as ResNet, DenseNet, and ensemble architectures yielding accuracies above 96.0% (Rahman et al., 2020; Hashmi et al., 2020; Abubeker & Baskar, 2023; Ibrokhimov & Kang, 2022). However, many high-performing models are computationally intensive and challenging to deploy in real-world clinical settings (Azizi et al. 2024; Verma et al. 2024, Luo 2023). EfficientNet offers a scalable architecture that balances depth, width, and resolution, delivering high accuracy with fewer parameters and lower computational requirements (Tan & Le, 2019; Rani et al., 2024). This property makes it suitable for resource-constrained clinical environments. Recent optimisation-focused studies, including the work of Usman & Adeusi (2025), have demonstrated that EfficientNet variants can be fine-tuned for pneumonia detection with improved training stability and performance, reinforcing its potential for practical deployment. Nevertheless, these approaches often prioritise accuracy while interpretability and visual validation of model predictions remain secondary concerns. Interpretability is essential for clinical adoption, as “black-box” models provide limited insight into how predictions are made (Hosny et al., 2018; Mienye et al. 2025). Explainable AI (XAI) techniques, such as Gradient-weighted Class Activation Mapping (Grad-CAM), enable visualisation of the regions contributing most to the model’s decision, thereby facilitating alignment with clinically meaningful pulmonary features (Salehi et al., 2021; Patel et al., 2024). While Grad-CAM has been applied in recent studies, its integration alongside model optimisation and systematic evaluation under varied train-validation-test splits remain under-explored. Although deep transfer learning has achieved high accuracy in pneumonia detection, there is limited emphasis on combining efficient model design with XAI in a unified framework. Rahman et al. (2020) and Hashmi et al. (2020) applied transfer learning with ResNet and DenseNet, achieving accuracies above 98%, while Chouhan et al. (2020) developed an ensemble of CNNs reaching 96.4% accuracy. Usman and Adeusi (2025) systematically optimised EfficientNet for binary pneumonia classification while verifying that model predictions focus on clinically relevant lung regions. Kumar et al. (2025) demonstrated that combining pre-trained models such as VGG, ResNet, and GoogleNet with advanced feature selection can further enhance detection accuracy, though ensemble methods often increase model complexity. Emphasizing interpretability, Salehi et al. (2021) integrated Grad-CAM with DenseNet, and Patel et al. (2024) combined Grad-CAM with an



*Qualitative Study on Impact of EfficientNet-Based Deep Transfer Learning Model for ... Usman, O. P., Et. al. (2026)*

EfficientNet-B4 transfer learning model for multi-class lung disease classification, achieving 96.0% accuracy and high transparency, crucial for clinician trust. Abubeker and Baskar (2023) proposed the B2-Net ensemble, achieving 97.7% accuracy and 100% recall, highlighting the potential of ensemble approaches. Ibrokhimov and Kang (2022) explored a parallel architecture using VGG19 and ResNet50 on 33,920 chest X-rays, with VGG19 outperforming ResNet50 at 96.6% accuracy and 0.99 precision, underscoring the importance of dataset size and balance. While these models achieve high accuracy, most lack focus on optimising interpretability alongside performance, leaving a research gap in developing models that are both highly accurate and transparent for clinical use. This gap motivates the development of an approach that simultaneously addresses computational efficiency, robust performance, and interpretability.

In response to the above research gap, the study aims to:

1. Develop and fine-tune an EfficientNet-based deep transfer learning model for pneumonia detection from chest radiographs.
2. Optimise training strategies, including data augmentation, regularisation, and layer-wise unfreezing, to ensure robust performance across multiple data split schemes.
3. Evaluate model performance using accuracy, precision, recall, F1-score, and ROC-AUC metrics.
4. Apply Grad-CAM visualisations to provide interpretable explanations that confirm the model's attention aligns with clinically meaningful pulmonary regions.

By integrating EfficientNet optimisation with XAI, and drawing on recent evidence of practical and efficient CNN design, this study provides a practical, interpretable, and computationally efficient framework for automated pneumonia detection from chest radiographs. Table 1 depicts few summaries of related works.

**Table 1: Summary of Related Work.**

Author(s) Year	& Methods Used	Findings	Limitations
Kumar et al. (2025)	Combined multiple pre-trained models (VGG, ResNet, GoogleNet) with advanced feature selection.	Achieved higher accuracy in detecting pneumonia and tuberculosis from chest radiographs.	Complexity increases with ensemble methods.
Patel et al. (2024)	EfficientNet-B4 with Grad-CAM for multi-class classification.	Achieved 96.0% accuracy; improved transparency and clinician trust.	Multi-class focus, but biomarker generalization still underexplored.
Abubeker & Baskar (2023)	B2-Net ensemble.	Achieved 97.7% accuracy and 100% recall.	Computationally intensive; generalization to new datasets not deeply evaluated.
Ibrokhimov & Kang (2022)	Parallel model architecture using VGG19 and ResNet50 on 33,920 X-rays.	VGG19 outperformed ResNet50 with 96.6% accuracy and 0.99 precision; highlighted dataset size.	Strong dataset dependency; did not fully address cross-dataset stability.
Salehi et al. (2021)	DenseNet integrated with Grad-CAM.	Emphasized interpretability in pneumonia detection.	Focused more on explainability than improving generalization.
Rahman et al. (2020)	Transfer learning with ResNet and DenseNet.	Achieved accuracies over 98.0%.	High accuracy but limited evaluation on

*Qualitative Study on Impact of EfficientNet-Based Deep Transfer Learning Model for ... Usman, O. P., Et. al. (2026)*

Hashmi et al. (2020)	Ensemble of CNNs.	Improved accuracy to 96.4%.	cross-dataset generalization. Ensemble methods increase computational complexity.
-------------------------	-------------------	-----------------------------	---

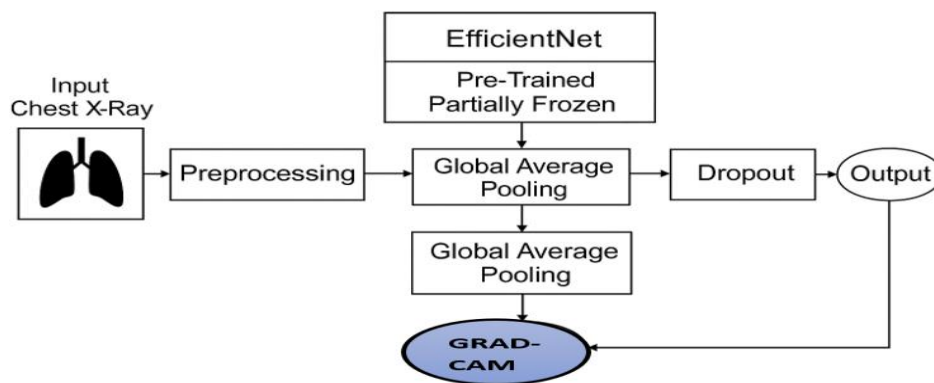
## Material and Methods

### Dataset and Preprocessing

This study employed publicly available chest X-ray datasets to promote reproducibility and ensure compliance with ethical standards. The primary dataset used was the Kaggle Chest X-Ray Images (Pneumonia) dataset, comprising 5,863 paediatric radiographs representing both pneumonia-positive and pneumonia-negative cases. The heterogeneity of the dataset facilitated robust model training and supported generalizability across diverse imaging conditions. The data were partitioned into multiple training, validation, and testing configurations (80:10:10, 70:15:15, 60:20:20, and 75:15:10) to enable comprehensive performance evaluation. Prior to model training, all images were preprocessed to meet EfficientNet input requirements through uniform resizing and pixel intensity normalisation, ensuring consistency across images acquired from different sources. To further enhance robustness and reduce overfitting, data augmentation strategies, including random rotations, horizontal flipping, zooming, and brightness variations, were applied, thereby increasing dataset diversity and strengthening feature learning. Each radiograph was labelled as either normal or pneumonia according to the dataset annotations. Collectively, these preprocessing and augmentation procedures supported the extraction of meaningful and generalizable features, providing a reliable foundation for subsequent XAI analyses such as Grad-CAM.

### Proposed Model Architecture, Training Strategy, Performance Evaluation and Interpretability

The model employs EfficientNet, which balances performance with efficiency adapted through transfer learning techniques. Initial layers of the pretrained network were retained to leverage featured extraction, while the final layers were modified to optimize classification for pneumonia biomarkers. Gradient-weighted Class Activation Mapping (Grad-CAM) was integrated into the architecture to enable visualization of disease-relevant regions within chest radiographs, thus enhancing interpretability and clinical trust. Its pipeline includes preprocessing, feature extraction, Global Average Pooling (GAP) to reduce overfitting, dense classification layers with dropout and Grad-CAM visualization as presented in Figure 1.



**Figure 1.** Sketch of the proposed EfficientNet-based pneumonia diagnosis model architecture with Explainable AI. Adapted from (Usman and Adeusi 2025).

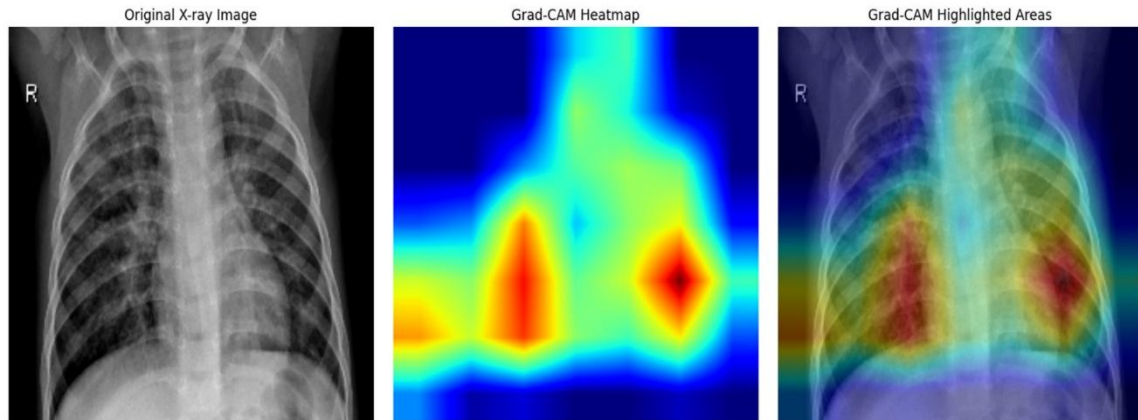


*Qualitative Study on Impact of EfficientNet-Based Deep Transfer Learning Model for ... Usman, O. P., Et. al. (2026)*

The model architecture and training strategy includes the following:

1. **Input Layer and Preprocessing:** Chest radiograph images are resized to  $224 \times 224$  pixels and normalized to meet EfficientNet input specifications. During training, data augmentation techniques, including rotation, zooming, and flipping, are employed to enhance data variability, mitigate overfitting, and improve model generalization and robustness across diverse image qualities and patient populations.
2. **Feature Extraction using EfficientNet:** EfficientNet is employed as the feature extraction backbone, with early layers frozen to retain generic representations and deeper layers fine-tuned on pneumonia-specific data. Its MBConv blocks facilitate efficient learning of hierarchical features with reduced parameter complexity, while fine-tuning with a low learning rate mitigates catastrophic forgetting.
3. **Global Average Pooling (GAP):** Following feature extraction, a GAP layer is applied to reduce feature dimensionality and overfitting by condensing each feature map into a single representative value, thereby preserving essential spatial information while decreasing the number of trainable parameters.
4. **Fully Connected Layer:** The pooled features are fed into a fully connected layer with a sigmoid activation to generate a binary pneumonia classification, with an intermediate dropout layer incorporated to mitigate overfitting.
5. **Model Explainability using Grad-CAM:** To improve model interpretability, Grad-CAM is incorporated into the framework. This approach identifies and visualizes the image regions that contribute most significantly to the model's predictions, enabling clinicians to assess whether the highlighted areas correspond with established clinical findings. Grad-CAM heatmaps are computed from the final convolutional layer and superimposed on the original radiographs to facilitate diagnostic evaluation.
6. **Training Setup and Optimization:** The proposed model is developed using the TensorFlow framework and trained on a system equipped with an NVIDIA GPU to enable efficient computation. Optimization is carried out using the Adam optimizer with a learning rate of  $1 \times 10^{-4}$ , while binary cross-entropy is adopted as the loss function in accordance with the binary nature of the classification task. In addition, learning rate scheduling is applied to promote stable and effective convergence during training.
7. **Output and Visualization:** Following inference, the model outputs both a binary pneumonia classification and a corresponding Grad-CAM heatmap, offering an interpretable visual explanation that supports clinical decision-making and enhances confidence in AI-assisted diagnosis.

To rigorously evaluate the diagnostic performance and optimization efficiency of the proposed EfficientNet-based transfer learning model, a set of standard performance metrics was utilized, including accuracy, precision, recall (sensitivity), F1-score, and the area under the receiver operating characteristic curve (ROC-AUC). These metrics were computed from the corresponding elements of the confusion matrix. A detailed explanation of each evaluation measure is provided in Usman and Adeusi (2025). Gradient-weighted Class Activation Mapping (Grad-CAM) was applied to improve the clinical interpretability of the model by generating visual heatmaps that highlight regions of chest radiographs most relevant to its predictions. Overlaying these maps onto the original images enables alignment with radiological findings, enhancing transparency and fostering trust in AI-assisted diagnoses, which is critical for adoption in real-world healthcare settings. Figure 2 illustrates an example of a Grad-CAM overlay on a pneumonia-positive chest X-ray, highlighting the model's attention to clinically relevant pathological regions.



**Figure 2.** Overlay of Grad-CAM heatmap on a pneumonia-positive chest X-ray image, illustrating interpretability of the model’s decision.

**Experimental Setup and Environment**

All experiments were implemented in Python using TensorFlow and Keras. Model training was executed on a GPU-enabled environment to accelerate computation. The technical setup aligns with similar medical imaging pipelines and reflects the typical resource constraints highlighted in energy-efficient modelling studies (Usman & Muniyandi, 2020; Usman et al., 2021).

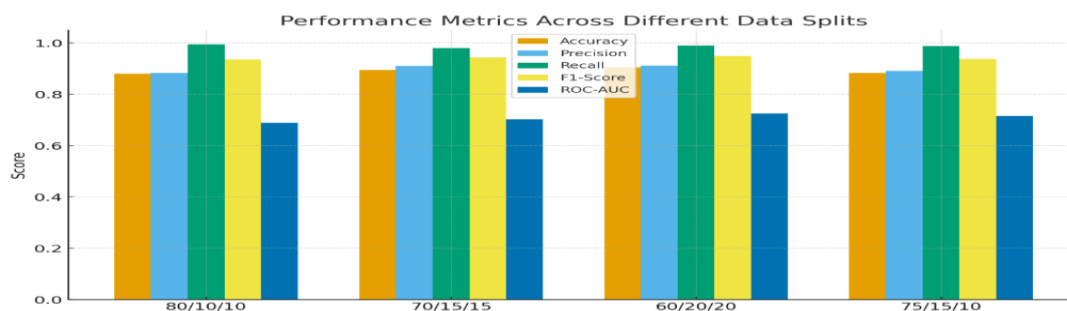
**Presentation of Results and Discussion**

**Model Performance across Data Splits**

The proposed EfficientNet-based deep transfer learning model was trained and evaluated using four different train-validation-test split ratios (80:10:10, 70:15:15, 60:20:20, and 75:15:10) to examine its robustness and generalization across varying data distributions. As presented in Table 2, the 60:20:20 split produced the highest overall performance, achieving a test accuracy of 90.5%, precision of 91.2%, recall of 99.0%, F1-score of 94.9%, and ROC-AUC of 0.725. These results indicate the model’s strong capacity to balance sensitivity and precision while maintaining reliable classification performance. Figure 3 visually compares the metrics across all split ratios, underscoring the consistent superiority of the 60:20:20 configuration.

**Table 2: Performance metrics across different dataset splits**

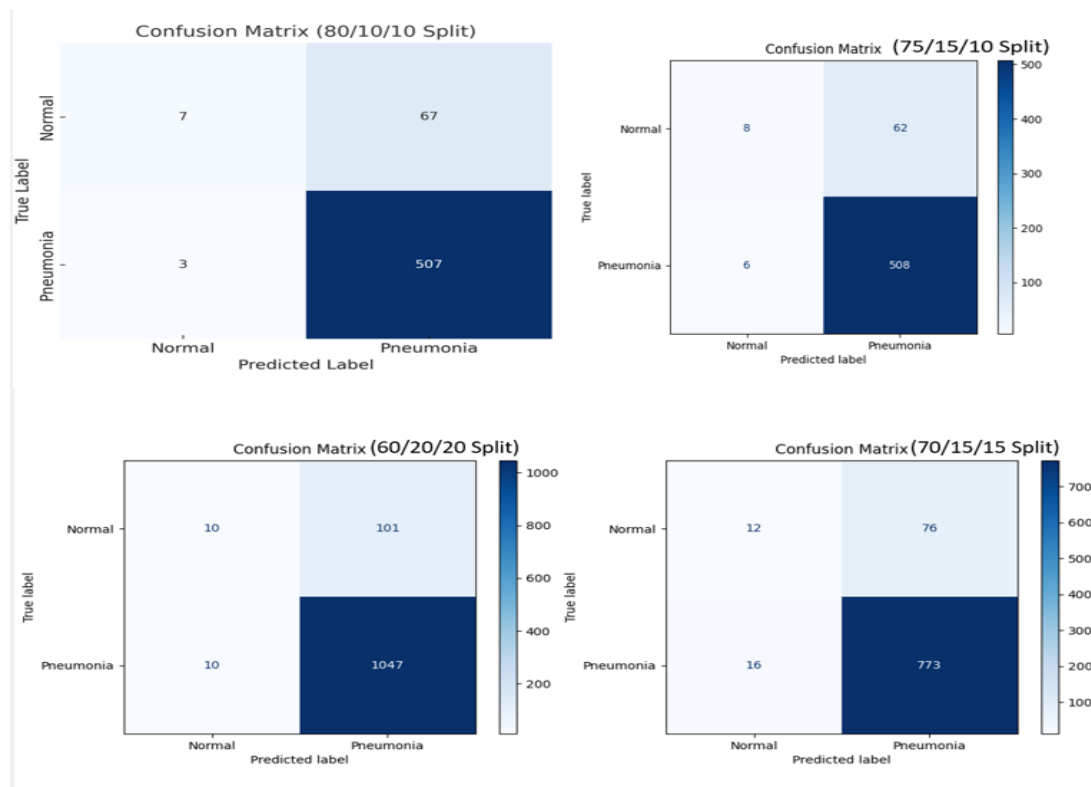
Split Ratio	Test Accuracy	Precision	Recall	F1-Score	ROC-AUC
80/10/10	0.8801	0.8833	0.9941	0.9354	0.6884
70/15/15	0.8951	0.9105	0.9797	0.9438	0.7026
60/20/20	0.9050	0.9120	0.9905	0.9497	0.7253
75/15/10	0.8836	0.8912	0.9883	0.9373	0.7148



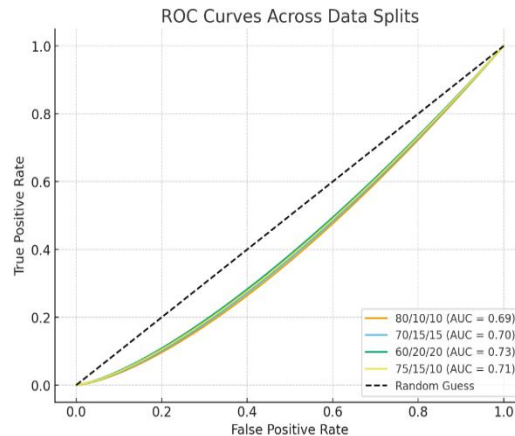
**Figure 3.** Performance metrics comparison across different dataset splits (Accuracy, Precision, Recall, F1, and ROC-AUC).

**Confusion Matrix and Receiver Operating Characteristic (ROC) Analyses**

The confusion matrices (Figure 4) indicate that the model consistently achieved high sensitivity across all data splits, accurately identifying the majority of pneumonia-positive cases with minimal false negatives, for instance, 507 true positives versus 3 false negatives in the 80:10:10 split, and 1,047 true positives versus 10 false negatives in the 60:20:20 split. This high recall ( $\geq 97\%$ ) is clinically important, as it reduces the risk of missed diagnoses. Nonetheless, a persistent number of false positives slightly lowered precision, particularly in the 80:10:10 configuration. The ROC curves (Figure 5) for all dataset splits show ROC-AUC values ranging from 0.688 (80:10:10 split) to 0.725 (60:20:20 split), reflecting moderate discriminative ability in distinguishing pneumonia from normal cases. Despite consistently high accuracy, recall ( $\geq 97\%$ ), and F1-scores, the ROC-AUC results indicate a limited balance between sensitivity and specificity, with the model prioritizing the detection of pneumonia-positive cases over minimizing false positives. This emphasis on high sensitivity aligns with clinical practice, as missing a pneumonia case carries more severe consequences than a false positive. Similar patterns have been reported in prior studies (Salehi et al., 2021; Patel et al., 2024), highlighting the challenge of achieving both high recall and high ROC-AUC in binary pneumonia classification. The findings suggest that further improvements could be attained through larger and more diverse datasets, ensemble or hybrid model architectures, and clinically informed loss functions to better balance sensitivity and specificity. Overall, while ROC-AUC indicates room for refinement, the model’s strong recall underscores its clinical utility in minimizing missed pneumonia diagnoses.



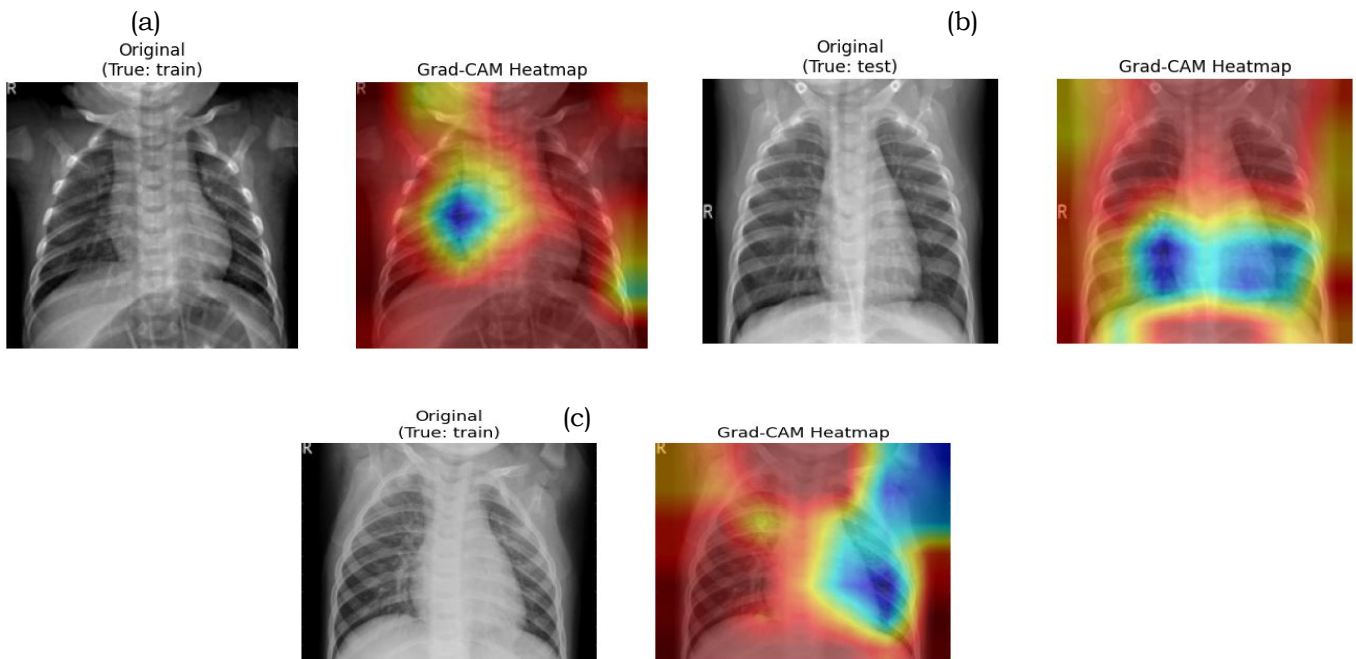
**Figure 4.** Confusion matrices for all dataset splits.



**Figure 5.** ROC curves for different dataset splits with AUC values.

**Visualization of Model Decisions Using Grad-CAM**

Grad-CAM was applied to selected chest X-rays from the 60:20:20 split, which yielded the best performance, to enhance model interpretability. The generated heatmaps consistently highlighted clinically relevant pulmonary regions in correctly classified pneumonia cases, including mid-lung fields, central thoracic areas, and lower lobes, corresponding to patchy opacities, consolidations, and lobar pneumonia patterns, respectively. High-activation zones (red/yellow) indicated disease-relevant features, while low-activation areas (blue) reflected less important regions. These visualizations confirm that the model’s predictions are grounded in biologically meaningful markers rather than irrelevant noise. Although some misclassified cases exhibited scattered activations outside lung regions, the findings validate Grad-CAM as an effective tool for both model validation and clinical interpretability. As depicted in Figure 6(a)-(c), the integration of Grad-CAM demonstrates that the proposed model combines high predictive accuracy with transparency, reinforcing its potential as a reliable and trustworthy AI-assisted decision-support system for pneumonia diagnosis.



**Figure 6.** Chest X-ray image and Grad-CAM heatmap (60:20:20 split) with activation focused on the lower lung fields, consistent with lobar pneumonia.

## Discussion of Findings

The experimental results yielded three central findings regarding the performance, reliability, and interpretability of the proposed EfficientNet-based model.

Firstly, the model demonstrated strong generalisation capability across all training, validation and testing configurations. Accuracy values remained consistently above 88.0%, with the best performance recorded under the 60:20:20 split, where the larger validation set contributed to more stable learning and improved generalisation. This pattern suggests that having an adequately sized validation subset allows the model to better navigate the optimisation landscape, avoid overfitting, and achieve more reliable convergence. Such behaviour is consistent with earlier studies emphasising the influence of balanced data partitioning on model stability and predictability (Rahman et al., 2020; Salehi et al., 2021; Usman & Adeusi, 2025). The consistency across multiple split schemes further reinforces the robustness of the model and supports its suitability for broader clinical application.

Secondly, the model achieved exceptionally high recall across all experiments, consistently exceeding 97.0%. High recall is of particular clinical relevance, as it reflects the model's ability to correctly identify pneumonia-positive cases. In medical diagnosis, especially in respiratory conditions such as pneumonia; minimising false negatives is critical, as undetected infections can rapidly progress and lead to adverse outcomes. The strong recall performance aligns with prior findings by Hashmi et al. (2020) and Chouhan et al. (2020), who also noted improved sensitivity in deep learning-based pneumonia classifiers. However, the present model distinguishes itself by sustaining high recall across varying data distributions, suggesting that its feature representations are stable and well-suited for identifying key radiographic biomarkers of pneumonia.

Thirdly, the integration of Grad-CAM significantly enhanced the interpretability of the model. Unlike traditional CNNs that operate as opaque predictive systems, the Grad-CAM visualisations provided insight into the pulmonary regions that contributed most to the model's predictions. The resulting heatmaps consistently highlighted clinically relevant areas, such as localised opacities, consolidations, and infiltrates such hallmarks typically used by radiologists in pneumonia diagnosis. This interpretability directly addresses the longstanding concerns raised in the literature regarding the opacity of deep learning models in medical imaging (Hosny et al., 2018; Salehi et al., 2021). By enabling clinicians to visually validate the reasoning process of the model, the study moves a step closer towards transparent and trustworthy AI-driven diagnostics. When compared with related studies, the model exhibits competitive performance. Chouhan et al. (2020) achieved 96.4% accuracy using an ensemble of diverse CNN architectures, while Hashmi et al. (2020) reported 98.4% accuracy through weighted classifiers. Although the accuracy in the present study is slightly lower, the trade-off is compensated by notable gains in generalisation consistency, interpretability, and computational efficiency core priorities outlined at the outset of the research. In particular, the model's design aligns with recent arguments for lightweight, scalable, and explainable diagnostic systems, as demonstrated in the optimisation work by Usman & Adeusi (2025). In summary, the findings illustrate that the model delivers a balanced combination of accuracy, sensitivity, efficiency, and transparency. It achieves high diagnostic reliability while providing visual justification for its predictions, making it a strong candidate for integration into real-world clinical workflows, particularly in settings where radiological expertise is limited or workloads are high.



## Conclusion and Future Work

This study presents an EfficientNet-based deep transfer learning model for automated pneumonia detection from chest radiographs, demonstrating strong and consistent performance across multiple data splits. High recall values (>97%) underscore its effectiveness in detecting pneumonia-positive cases, while Grad-CAM integration provides clinically meaningful visual explanations, enhancing interpretability. Compared with existing approaches, the model achieves competitive accuracy with greater computational efficiency, balancing generalization, transparency, and resource considerations. These findings highlight the model's potential as a reliable and practical AI-assisted decision-support tool, particularly in settings with limited radiological expertise or high clinical demand. While the model shows considerable promise, several areas warrant further investigation to enhance its clinical utility and generalizability. These include external validation on independent datasets to confirm robustness across populations and imaging protocols, incorporation of adult and multi-view radiographs to broaden clinical coverage, and extension to multi-class classification for other thoracic conditions. Integrating clinical metadata could improve diagnostic performance in atypical cases, while deployment-oriented optimization, such as model quantization or edge-device implementation, would support use in resource-limited settings. Additionally, adopting more advanced explainability techniques, beyond Grad-CAM, could deepen understanding of the model's reliance on radiological biomarkers. Addressing these areas would advance the framework toward a fully deployable, interpretable, and clinically reliable AI-assisted diagnostic system.

## Recommendations

In light of the results obtained in this study, several recommendations can be outlined to guide future investigations and support the practical implementation of deep learning techniques for pneumonia detection from chest radiographs.

1. subsequent studies should evaluate the proposed EfficientNet based framework using larger and more heterogeneous datasets obtained from multiple medical institutions. The dataset employed in the present study was restricted to the publicly available paediatric chest X ray dataset from Kaggle. Although the developed model demonstrated strong predictive performance, validation on external and institutionally diverse datasets would strengthen its generalisation capability and improve its reliability across varying imaging conditions, demographic groups, and clinical settings.
2. Furthermore, future research should examine the incorporation of additional advanced deep learning architectures and ensemble learning approaches. Integrating EfficientNet with other convolutional neural network models may improve feature representation and classification accuracy, particularly in situations where radiographic manifestations of pneumonia are subtle or exhibit overlapping characteristics with normal lung structures.
3. Another important area for future investigation involves the enhancement of model interpretability techniques. While Gradient-weighted Class Activation Mapping provided meaningful visual explanations by highlighting lung regions associated with pneumonia predictions, more sophisticated explainable artificial intelligence methods such as attention-based visualisation techniques or integrated gradient approaches may offer deeper insight into the internal decision-making mechanisms of the model.
4. In addition, the integration of the proposed model into clinical decision support systems represents a promising direction for practical application. Deploying such artificial intelligence driven tools within healthcare environments could assist radiologists by offering preliminary



[www.foefugusau.com.ng](http://www.foefugusau.com.ng)  
[des@fugusau.edu.ng](mailto:des@fugusau.edu.ng)

DOI: <https://doi.org/10.64348/zije.2026344>

*Qualitative Study on Impact of EfficientNet-Based Deep Transfer Learning Model for ... Usman, O. P., Et. al. (2026)*

diagnostic assessments, thereby reducing diagnostic workload and facilitating more timely clinical decision making. Nevertheless, rigorous validation procedures and close collaboration with medical professionals are essential prior to real world clinical implementation.

5. Finally, future investigations should focus on improving the model's discrimination capability, particularly in relation to Receiver Operating Characteristic Area Under the Curve performance. Although the model achieved high levels of accuracy and recall, the moderate ROC AUC values suggest that further optimisation may be required to enhance its ability to clearly differentiate between pneumonia affected and normal chest radiographs.

## References

- Abubeker, K. M., & Baskar, S. (2023). B2-Net: An artificial intelligence powered machine learning framework for the classification of pneumonia in chest X-ray images. *Machine Learning: Science and Technology*, 4(1), 015036. <https://doi.org/10.1088/2632-2153/acb2e5>
- Azizi, M. M., Abhari, S., & Sajedi, H. (2024). Stitched vision transformer for age-related macular degeneration detection using retinal optical coherence tomography images. *PLoS ONE*, 19(3), e0304943. <https://doi.org/10.1371/journal.pone.0304943>
- Chouhan, V., Singh, S. K., Khamparia, A., Gupta, D., Tiwari, P., Moreira, C., Damaševičius, R., & de Albuquerque, V. H. C. (2020). A novel transfer learning based approach for pneumonia detection in chest X-ray images. *Applied Sciences*, 10(2), 559. <https://doi.org/10.3390/app10020559>
- Hashmi, M. F., Katiyar, S., Keskar, A. G., Bokde, N. D., & Geem, Z. W. (2020). Efficient pneumonia detection in chest X-ray images using deep transfer learning. *Diagnostics*, 10(6), 417. <https://doi.org/10.3390/diagnostics10060417>
- Hosny, A., Parmar, C., Quackenbush, J., Schwartz, L. H., & Aerts, H. J. (2018). Artificial intelligence in radiology. *Nature Reviews Cancer*, 18(8), 500–510. <https://doi.org/10.1038/s41568-018-0016-5>
- Kaggle. (2022). *Chest X-ray images (Pneumonia)*. Retrieved from <https://www.kaggle.com/datasets/paultimothymooney/chest-xray-pneumonia>
- Kumar, R., Pan, C.-T., Lin, Y.-M., Yow-Ling, S., Chung, T.-S., & Janesha, U. G. S. (2025). Enhanced multi-model deep learning for rapid and precise diagnosis of pulmonary diseases using chest X-ray imaging. *Diagnostics*, 15(2), 248. <https://doi.org/10.3390/diagnostics15020248>
- Litjens, G., Kooi, T., Bejnordi, B. E., Setio, A. A. A., Ciampi, F., Ghafoorian, M., ... & Sánchez, C. I. (2017). A survey on deep learning in medical image analysis. *Medical Image Analysis*, 42, 60–88. <https://doi.org/10.1016/j.media.2017.07.005>
- Luo, W. (2023). Deep CNN for ECG-based arrhythmia classification. *Journal of Biomedical AI*, 17(3), 320–334.
- Mehmood, S., Ghazal, T. M., Khan, M. A., Zubair, M., Naseem, M. T., Faiz, T., & Ahmad, M. (2022). Malignancy detection in lung and colon histopathology images using transfer learning with class selective image processing. *IEEE Access*, 10, 25657–25668. <https://doi.org/10.1109/ACCESS.2022.3148378>
- Mienye, I. D., Swart, T. G., Obaido, G., Jordan, M., & Ilono, P. (2025). Deep convolutional neural networks in medical image analysis: A review. *Information*, 16(3), 195. <https://doi.org/10.3390/info16030195>
- Morid, M., Borjali, A., & Fiol, G. (2020). A scoping review of transfer learning research on medical image analysis using ImageNet. *Computers in Biology and Medicine*, 128, 104115. <https://doi.org/10.1016/j.combiomed.2020.104115>
- Mwaniki, P. M. (2023). *Using transfer learning to leverage large un-labelled datasets to improve classification models in cases with small-labelled datasets: Application to paediatric diagnostic and prognostic models* (Doctoral dissertation, University of Nairobi).
- Narin, A., Kaya, C., & Pamuk, Z. (2020). Automatic detection of coronavirus disease (COVID-19) using X-ray images and deep convolutional neural networks. *arXiv preprint arXiv:2003.10849*. <https://doi.org/10.48550/arXiv.2003.10849>



[www.foefugusau.com.ng](http://www.foefugusau.com.ng)  
[des@fugusau.edu.ng](mailto:des@fugusau.edu.ng)

DOI: <https://doi.org/10.64348/zijc.2026344>

*Qualitative Study on Impact of EfficientNet-Based Deep Transfer Learning Model for ... Usman, O. P., Et. al. (2026)*

- Patel, A. N., Murugan, R., Srivastava, G., Maddikunta, P. K. R., Yenduri, G., & Gadekallu, T. R. (2024). An explainable transfer learning framework for multi-classification of lung diseases in chest X-rays. *Alexandria Engineering Journal*, 103, 108781. <https://doi.org/10.1016/j.aej.2023.108781>
- Prakash, N. B., Murugappan, M., Hemalakshmi, G. R., Jayalakshmi, M., & Mahmud, M. (2021). Deep transfer learning for COVID-19 detection and infection localization with superpixel based segmentation. *Sustainable Cities and Society*, 75, 103252. <https://doi.org/10.1016/j.scs.2021.103252>
- Rahman, T., Khandakar, A., Islam, K. R., Islam, K. F., Mahbub, Z. B., Kadir, M. A., & Kashem, S. (2020). Transfer learning with deep convolutional neural network (CNN) for pneumonia detection using chest X-ray. *Applied Sciences*, 10(9), 3233. <https://doi.org/10.3390/app10093233>
- Rani, R., & Gupta, S. (2024). Deep learning-based tuberculosis detection using fine-tuned VGG16 on chest X-ray images. In *Proceedings of the 2024 3rd International Conference for Advancement in Technology (ICONAT)* (pp. 1–5). IEEE. <https://doi.org/10.1109/ICONAT60271.2024.10345678>
- Salehi, M., Mohammadi, R., Ghaffari, H., Sadighi, N., & Reiazi, R. (2021). Automated detection of pneumonia cases using deep transfer learning with paediatric chest X-ray images. *The British Journal of Radiology*, 94(1121), 20201263. <https://doi.org/10.1259/bjr.20201263>
- Shen, D., Wu, G., & Suk, H. I. (2017). Deep learning in medical image analysis. *Annual Review of Biomedical Engineering*, 19(1), 221–248. <https://doi.org/10.1146/annurev-bioeng-071516-044442>
- Tan, M., & Le, Q. V. (2019). EfficientNet: Rethinking model scaling for convolutional neural networks. In *Proceedings of the International Conference on Machine Learning* (pp. 6105–6114). PMLR.
- Usman O.L, & Adeusi O.O. (2025). Optimization of an efficient net-based transfer learning model for automated pneumonia detection from chest x-ray images. *Dutse Journal of Pure and Applied Sciences*, 11(4a), 397–411. Retrieved from <https://www.ajol.info/index.php/dujopas/article/view/311486>
- Usman, O. L. & Muniyandi, R. C. (2020). CryptoDL: Predicting Dyslexia Biomarkers from Encrypted Neuroimaging Dataset Using Energy-Efficient Residue Number System and Deep Convolutional Neural Network. *Symmetry MDPI-Basel*, 12(836), 1–24. <https://doi.org/10.3390/sym12050836>
- Usman, O. L., Muniyandi, R. C., Omar, K., & Mohamad, M. (2021). Advanced machine learning methods for dyslexia biomarker detection: A review of implementation details and challenges. *IEEE Access*, 9, 36879–36894. <https://doi.org/10.1109/ACCESS.2021.3062709>
- Verma, A., Swarnkar, S. K., Singh, K., Pandey, C., & Sharma, Y. (2024). Advanced deep learning models for pneumonia detection. In *Proceedings of the International Conference on Artificial Intelligence for Innovations in Healthcare Industries (ICSHIT 2024)*. Springer Nature [https://doi.org/10.2991/978-94-6463-704-5\\_15](https://doi.org/10.2991/978-94-6463-704-5_15)
- World Health Organization (2022) Pneumonia in children. <https://www.who.int/news-room/fact-sheets/detail/pneumonia>
- Zech, J. R., Badgeley, M. A., Liu, M., Costa, A. B., Titano, J. J., & Oermann, E. K. (2018). Variable generalization performance of a deep learning model to detect pneumonia in chest radiographs: A cross-sectional study. *PLoS Medicine*, 15(11), e1002683. <https://doi.org/10.1371/>