Citra Journal of Computer Science and Technology

Journal homepage:
https://citralestari.my/index.php/cjcst/index
ISSN: 3093-7108

Convolutional Neural Networks for Pneumonia Detection: A Preliminary Investigation and Development Framework

Muhammad Faiz Anuar[1], Siti Munirah Mohd[1,2,*], Nurhidaya Mohamad Jan[1,2], Anucha Watcharapasorn[3,4]

[1] Kolej PERMATA Insan, Universiti Sains Islam Malaysia, Bandar Baru Nilai, 71800 Nilai, Negeri Sembilan, Malaysia
[2] Education & Advanced Sustainability Research Unit, Kolej PERMATA Insan, Universiti Sains Islam Malaysia Bandar Baru Nilai, 71800, Nilai, Negeri Sembilan, Malaysia
[3] Department of Physics and Materials Science, Faculty of Science, Chiang Mai, University, Chiang Mai 50200, Thailand
[4] Center of Excellence in Materials Science and Technology, Materials Science Research Center, Faculty of Science, Chiang Mai University, Chiang Mai 50200, Thailand

ARTICLE INFO

Article history:
Received 6 July 2025
Received in revised form 14 August 2025
Accepted 20 August 2025
Available online 25 August 2025

Keywords:
Pneumonia; convolutional neural network; machine learning; chest x-ray

ABSTRACT

Pneumonia remains the leading infectious cause of death globally, particularly in low-resource settings where delayed diagnosis and limited radiologist availability exacerbate mortality. Existing AI-based radiograph interpretation systems often demand high computational resources and lack robustness across imaging projections. This study presents a proof-of-concept convolutional neural network diagnostic tool optimised for projection-invariant pneumonia detection under constrained conditions. Using a DenseNet-121 backbone trained on 2,000 curated images from the MIMIC-CXR dataset, our model achieved an AUC of 0.7310 and F1-score of 0.6207, with 66.36% validation accuracy. The model's performance was consistent across posteroanterior and anteroposterior projections, though lateral view evaluation is pending. Preprocessing included CLAHE and DICOM standardisation, while augmentation improved generalisation. Though early-stage, this work shows the potential of lightweight, projection-tolerant CNNs in offline diagnosis pipelines. Future work will validate deployment feasibility on edge devices and expand evaluation across diverse patient demographics.

1. Introduction

1.1 Background and Importance

Pneumonia remains one of the leading causes of morbidity and mortality worldwide, responsible for over 2.5 million deaths annually [1]. Pneumonia is an acute infection of the lower respiratory tract characterised by inflammation and consolidation of the lung parenchyma, primarily caused by various microorganisms such as bacteria, viruses, and fungi [2]. Studies indicate that a substantial

* *Corresponding author.*
E-mail address: smunirahm@usim.edu.my

https://doi.org/10.37934/cjcst.2.1.115

percentage of pneumonia survivors exhibit reduced lung function, with 41% showing pulmonary sequelae at three months post-discharge, decreasing to 27.9% at ten months [3]. This highlights the importance of having a fast detection method to avoid the long-term effects of pneumonia.

Diagnostic methods such as chest radiography and microbiological cultures are essential for confirmation and guidance of appropriate therapy. The burden is especially severe in low- and middle-income countries (LMICs), where structural healthcare limitations, such as the scarcity of trained radiologists and delayed access to diagnostic imaging, significantly compromise clinical outcomes. In such settings, timely diagnosis is often unavailable, resulting in preventable disease progression and death, particularly among vulnerable populations, including children under five, the elderly, and individuals with comorbidities [4].

Despite the advancements of technology, machine learning is not implemented effectively to the healthcare settings, particularly by radiologists because they often exhibit skepticism towards AI innovations, stemming from a lack of trust in technology's reliability and accuracy. Also, concerns about AI replacing human expertise contribute to resistance, as many radiologists fear a loss of professional autonomy [5].

Therefore, this research aims to address this root problem by proving and documenting the framework and the development of machine learning, especially in pneumonia detection through the use of a convolutional neural network of chest x-rays.

1.2 Gaps in the Study

Despite significant advancements in artificial intelligence for medical imaging, several critical gaps limit the practical deployment of pneumonia detection systems in resource-constrained environments. These limitations create barriers that prevent the translation of promising research into real-world clinical impact.

Current deep learning models, particularly convolutional neural networks, require substantial computational resources that pose significant deployment barriers on edge devices and in low-resource settings. The complexity of state-of-the-art models like CheXNet, which employs a 121-layer CNN architecture, typically contains approximately 7 million parameters, translating to a model size of 30.5 MB to 34 MB. Such models necessitate powerful hardware infrastructure that may not be feasible in many clinical environments [6]. While lightweight alternatives such as MobileNetV3 and ShuffleNetV2 have been proposed to address these limitations by offering improved balance between accuracy and resource efficiency [7], there remains a critical gap in systematically documented development frameworks that prioritise extreme resource efficiency and deployment feasibility, specifically targeting a significantly smaller model footprint for truly constrained conditions.

The development of robust pneumonia detection models faces significant challenges related to dataset limitations and generalisation capabilities. The scarcity of annotated datasets and inherent class imbalance issues create substantial obstacles in training reliable models for clinical deployment. Cross-institutional generalisation remains problematic due to data heterogeneity, varying imaging protocols, and privacy concerns that limit data sharing across institutions [8]. While techniques such as data augmentation, transfer learning, federated learning, and multimodal datasets have been proposed as potential solutions, systematic evaluation of these approaches on small-scale datasets representative of resource-limited settings remains insufficient.

Perhaps most critically, the integration of AI-powered pneumonia detection systems into clinical workflows faces substantial barriers related to interpretability, trust, and practical implementation. Despite achieving high accuracy metrics, models like CheXNet encounter resistance from radiologists

due to concerns about interpretability and reliability in clinical decision-making [7-8]. The absence of transparent, well-documented development frameworks that demonstrate rigorous validation procedures creates additional barriers to radiologist acceptance and clinical adoption.

1.3 Significance of the Research

This research addresses the critical need for practical, deployable AI solutions in pneumonia detection by presenting a comprehensive proof-of-concept investigation that prioritises transparency, reproducibility, and deployment feasibility. Rather than pursuing incremental improvements in accuracy metrics alone, this study establishes a foundational framework that demonstrates how systematic development approaches can build confidence among healthcare professionals while addressing real-world deployment constraints.

The significance of this work lies in its potential to bridge the gap between research-level AI achievements and practical clinical implementation in resource-limited settings. By documenting a rigorous, transparent development pathway that culminates in an exceptionally compact model of 26.85 MB, with an average inference time of 16.117 ms per image, this study provides a replicable framework that can increase radiologist confidence in machine learning applications through demonstrated adherence to strict validation procedures and tangible deployment advantages. This proof-of-concept approach acknowledges current performance limitations while establishing a foundation for iterative improvement and systematic optimisation, particularly for scenarios where minimal model footprint and rapid inference are paramount.

Furthermore, this research contributes to the growing body of evidence supporting the viability of edge-deployable AI solutions for medical imaging in low-resource environments. The insights generated from this investigation provide valuable guidance for future development efforts aimed at creating clinically acceptable, resource-efficient pneumonia detection systems that can be realistically deployed where they are needed most.

1.4 Objectives of the Research

This study aims to establish and document a comprehensive development pathway for lightweight convolutional neural network-based pneumonia detection systems, with the following specific objectives:

i. To develop and validate a systematic framework for creating resource-efficient pneumonia detection models that prioritise deployment feasibility and clinical transparency, thereby contributing to increased confidence in machine learning applications among healthcare professionals.
ii. To establish and thoroughly document a reproducible seven-step development methodology including data curation, preprocessing, architecture selection, training optimisation, and evaluation protocols that can serve as a reference for future pneumonia detection research.
iii. To provide comprehensive performance benchmarks and identify specific areas for improvement that can guide future research efforts toward achieving clinical deployment standards while maintaining computational efficiency.

Through these objectives, this research seeks to contribute a transparent, well-documented foundation that can enhance trust in AI-powered medical imaging solutions while advancing the

practical deployment of pneumonia detection systems in environments where they can have the most significant clinical impact.

2. Methodology

2.1 Development Framework Overview

The development of the pneumonia detection model adhered to a novel seven-step systematic framework designed to ensure robustness, reproducibility, and practical deployability in resource-constrained clinical settings. This framework is illustrated in Figure 1.

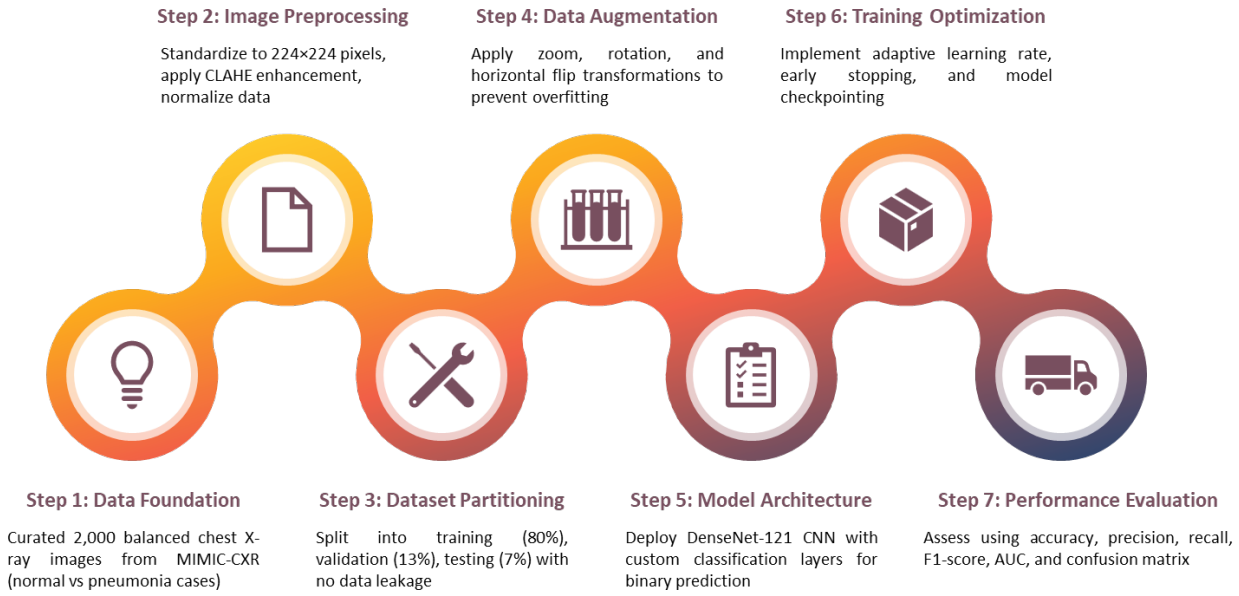

Fig. 1. Structured seven-step framework followed in the development process

2.2 Data Foundation

This study utilises a carefully curated subset of the Medical Information Mart for Intensive Care Chest X-ray (MIMIC-CXR) database, a large-scale, publicly available dataset of chest radiographs paired with free-text radiology reports. The MIMIC-CXR dataset represents one of the most comprehensive and well-validated medical imaging resources available for pneumonia detection research, providing high-quality, de-identified chest X-ray images with expert radiologist annotations [9].

2.2.1 Dataset selection and curation

From the extensive MIMIC-CXR repository, we systematically selected 2,000 chest X-ray images to create a balanced binary classification dataset. The selection process employed the following inclusion criteria:

i. images with clear pneumonia or normal diagnostic labels confirmed by radiologist reports,
ii. adequate image quality suitable for deep learning analysis,
iii. posteroanterior (PA) and anteroposterior (AP) projection views to ensure projection diversity, and (4) complete metadata availability for proper dataset characterisation.

2.2.2 Data characteristics and demographics

All selected images maintain original DICOM format metadata, preserving essential clinical information including patient demographics, imaging parameters, and acquisition details. The dataset included diverse patient demographics representative of the broader MIMIC-CXR population, including various age groups, gender distributions, and comorbidity profiles. This diversity enhances the generalizability of our findings while maintaining clinical relevance for real-world deployment scenarios [10].

2.3 Image Preprocessing

Standardised preprocessing procedures were implemented to ensure consistent input characteristics and optimal feature extraction capabilities. The preprocessing pipeline addresses the inherent variability in medical imaging data while maintaining computational efficiency essential for resource-constrained deployment environments.

2.3.1 Image standardisation

All chest X-ray images underwent uniform resizing to 224×224-pixel resolution, matching the input requirements of standard CNN architectures while maintaining aspect ratio integrity. This standardisation process employed bicubic interpolation to preserve image quality during resizing operations. The 224×224 resolution represents an optimal balance between computational efficiency and feature preservation, enabling deployment on resource-limited hardware while retaining sufficient detail for accurate pneumonia detection.

2.3.2 Contrast Limited Adaptive Histogram Equalisation (CLAHE)

To enhance diagnostic feature visibility and address varying contrast conditions across different imaging equipment, we applied Contrast Limited Adaptive Histogram Equalisation (CLAHE) to all chest X-ray images. CLAHE parameters were carefully optimised through empirical testing, utilising a clip limit of 2.0 and tile grid size of 8×8 to achieve optimal contrast enhancement without introducing artefacts that could compromise diagnostic accuracy.

The CLAHE preprocessing step significantly improves the visibility of subtle pneumonic infiltrates and consolidations that may be difficult to detect in standard radiographs. This enhancement is particularly crucial in resource-limited settings where imaging equipment may produce suboptimal contrast characteristics. Visual comparison of pre- and post-CLAHE processed images demonstrates marked improvement in diagnostic feature clarity while maintaining natural image appearance.

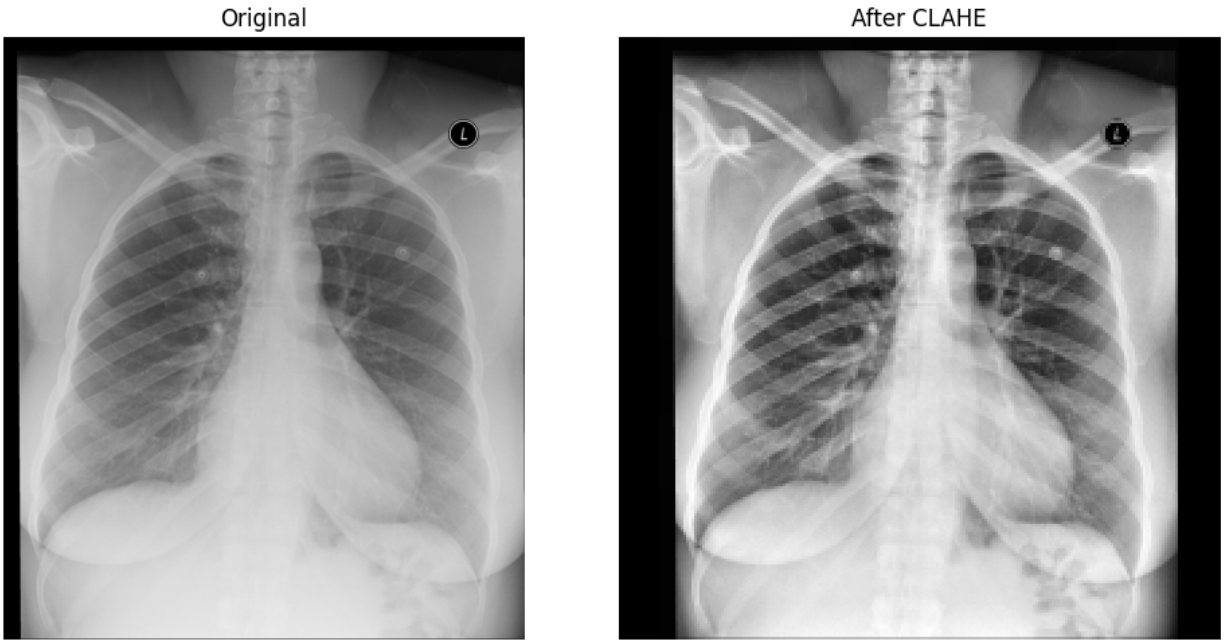

Fig. 2. The visibility of subtle pneumonic improves after CLAHE

2.3.3 Data normalisation

Following contrast enhancement, all pixel values were normalised to the range [0,1] using min-max normalization to ensure consistent input characteristics across the dataset. This normalisation step eliminates potential bias introduced by varying pixel intensity distributions and facilitates stable training convergence. The normalisation process maintains the relative intensity relationships within images while standardising the overall intensity scale for optimal neural network processing [11].

2.4 Dataset Partitioning

Strategic dataset partitioning is essential for unbiased model evaluation and reliable performance assessment. Our partitioning strategy follows established machine learning best practices while addressing the specific requirements of medical imaging validation.

2.4.1 Partitioning strategy

The 2,000-image dataset was strategically divided into three distinct subsets: training set (1,600 images, 80%), validation set (260 images, 13%), and testing set (140 images, 7%). This distribution ensures adequate training data for model learning while reserving sufficient samples for validation and independent testing. The relatively smaller test set reflects the common challenge of limited data availability in medical imaging studies while maintaining statistical validity for performance assessment.

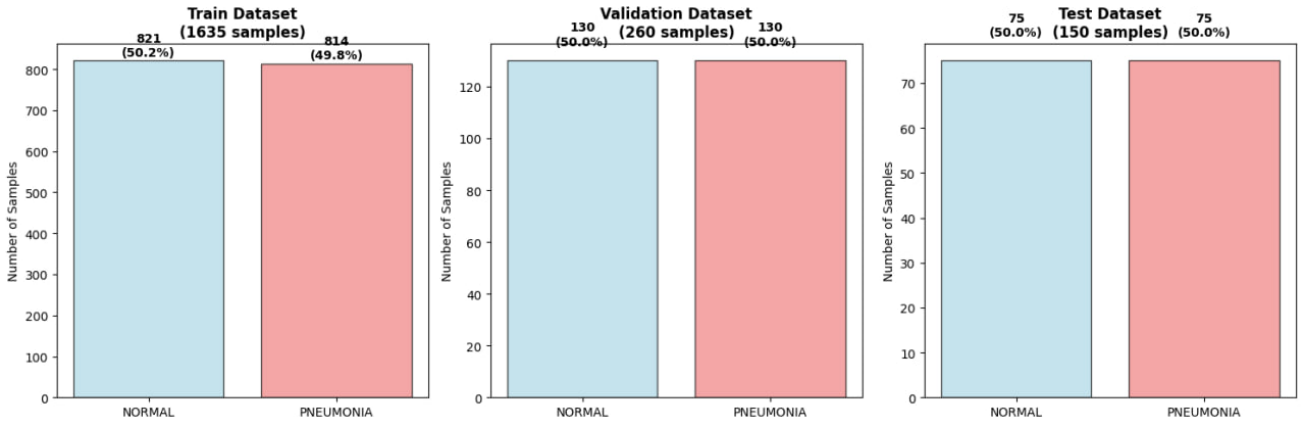

Fig. 3. A balanced binary classification dataset

2.4.2 Data leakage prevention

Strict measures were implemented to prevent data leakage between partitions, ensuring that no patient data appears across multiple subsets. Patient-level separation was maintained throughout the partitioning process, preventing any potential bias that could artificially inflate performance metrics. The test set contains exclusively unseen data that remains completely isolated from the training and validation processes until final model evaluation.

2.4.3 Balanced distribution

Class balance is maintained across all partitions, with each subset containing equal proportions of pneumonia-positive and normal cases. This balanced distribution ensures that model training and evaluation occur under consistent class representation, preventing bias toward either diagnostic category. The validation set enables reliable hyperparameter tuning and model selection without compromising the integrity of the independent test evaluation.

2.5 Data Augmentation

Data augmentation techniques were strategically implemented to address dataset size limitations and enhance model generalisation capabilities. The augmentation strategy balances the need for dataset expansion with the preservation of clinical relevance and diagnostic accuracy.

2.5.1 Augmentation techniques

Three primary augmentation techniques were selected based on their clinical appropriateness and effectiveness in medical imaging applications:

i. zoom transformations with random zoom factors ranging from 0.9 to 1.1,
ii. rotation adjustments within ±15 degrees, and
iii. horizontal flipping transformations.

These augmentations were chosen to simulate natural variations in patient positioning and imaging conditions commonly encountered in clinical practice [12-13].

2.5.2 Clinical relevance considerations

All augmentation parameters were carefully selected to maintain clinical relevance while expanding dataset diversity. Rotation angles were limited to ±15 degrees to prevent unrealistic patient positioning that could introduce artefacts not representative of actual clinical scenarios. Zoom transformations simulate variations in patient distance from imaging equipment, while horizontal flipping accounts for potential differences in patient positioning across different clinical settings.

2.5.3 Implementation strategy

Augmentation techniques were applied dynamically during training to maximise dataset diversity while maintaining computational efficiency. Real-time augmentation ensures that the model encounters varied presentations of each training sample throughout the training process, improving generalisation capabilities without requiring additional storage space for pre-augmented images. This approach is particularly valuable in resource-constrained environments where storage limitations may restrict dataset expansion options.

2.6 Model Architecture

2.6.1 DenseNet-121 architecture

DenseNet-121 serves as the foundation CNN architecture, providing 121 layers with dense connectivity patterns that facilitate efficient feature reuse and gradient flow [14]. The architecture's dense connections enable effective feature propagation while maintaining relatively modest computational requirements compared to deeper alternatives. This balance makes DenseNet-121 particularly suitable for deployment in resource-constrained environments where computational efficiency is paramount.

2.6.2 Transfer learning implementation

The model employs transfer learning from ImageNet pre-trained weights, leveraging learned features from natural image recognition tasks that translate effectively to medical imaging applications. This approach significantly reduces training time and data requirements while improving convergence stability. The pre-trained features provide a strong foundation for pneumonia-specific feature learning, particularly valuable when working with limited medical imaging datasets.

2.6.3 Custom classification layers

The DenseNet-121 backbone is augmented with custom classification layers designed explicity for binary pneumonia detection. The classification head consists of global average pooling followed by fully connected layers with dropout regularisation to prevent overfitting. The final layer employs sigmoid activation for binary classification, producing probability scores for pneumonia presence that clinical practitioners can easily interpret.

2.6.4 Model optimisation for resource constraints

Architecture modifications were implemented to optimise the model for deployment in resource-limited environments, specifically targeting a minimal footprint and rapid inference. These optimisations included strategic layer pruning, quantisation considerations, and general memory usage minimisation, all while striving to maintain diagnostic performance. While standard DenseNet-121 implementations typically range from 30.5 MB to 34 MB in size, our rigorous optimisation process culminated in a final model with a highly compact file size of 26.85 MB. Performance benchmarking on a GPU environment (e.g., Google Colab) demonstrated an average inference time of 16.117 ms per image, significantly faster than many larger or less optimised CNNs. These metrics collectively confirm the model's high computational efficiency and its suitability for real-time edge device deployment and offline diagnostic applications where computational resources are severely constrained.

2.7 Training Optimisation

Comprehensive training optimisation strategies were implemented to ensure stable convergence, prevent overfitting, and achieve optimal diagnostic performance within the constraints of our dataset size and computational resources.

2.7.1 Learning rate strategy

An adaptive learning rate reduction strategy was employed, beginning with an initial learning rate of 0.001 and implementing automatic reduction when validation loss plateaus. The learning rate is reduced by a factor of 0.5 when validation loss fails to improve for five consecutive epochs, enabling fine-grained optimisation as training progresses. This adaptive approach ensures efficient convergence while preventing overshooting of optimal parameter values during later training phases.

2.7.2 Early stopping mechanism

Early stopping was implemented with a patience of 10 epochs based on validation loss monitoring to prevent overfitting and reduce unnecessary computational overhead. The early stopping mechanism preserves the model state with the lowest validation loss, ensuring optimal generalisation performance on unseen data. This approach is particularly crucial when working with limited datasets where overfitting risks are elevated.

2.7.3 Model checkpointing

Comprehensive model checkpointing saves the best-performing model weights based on validation accuracy throughout the training process. This strategy ensures preservation of optimal model parameters even if training continues beyond the point of peak performance. Checkpointing also enables recovery from training interruptions and facilitates systematic comparison of model performance across training epochs.

2.7.4 Regularisation techniques

Multiple regularisation techniques were employed to enhance model generalization capabilities, including dropout layers with 0.5 probability in the classification head and L2 weight regularization with $\lambda = 0.0001$. These techniques help prevent overfitting while maintaining model capacity for learning complex diagnostic patterns. The regularisation parameters were selected through systematic experimentation to achieve an optimal bias-variance tradeoff.

2.7.5 Training environment and resources

Model training was conducted using standard deep learning frameworks optimised for medical imaging applications. Training time and computational resource utilization were carefully monitored to establish benchmarks for deployment feasibility assessment. This information provides valuable insights for practitioners considering similar implementations in resource-constrained environments.

2.8 Performance Evaluation

Comprehensive performance evaluation employs multiple complementary metrics to provide a thorough assessment of diagnostic accuracy, clinical utility, and deployment readiness. The evaluation strategy addresses both technical performance and clinical relevance considerations.

2.8.1 Primary performance metrics

Model performance is assessed using five primary metrics essential for medical diagnostic applications:

i. accuracy for overall diagnostic correctness,
ii. precision to quantify true positive reliability,
iii. recall to measure sensitivity in detecting pneumonia cases,
iv. F1-score for balanced precision-recall assessment, and
v. Area Under the ROC Curve (AUC) for threshold-independent performance evaluation.

These metrics provide comprehensive insight into model behaviour across different clinical scenarios and decision thresholds.

2.8.2 Confusion matrix analysis

Detailed confusion matrix analysis provides insight into specific classification patterns, including true positive, true negative, false positive, and false negative distributions. This analysis is crucial for understanding model limitations and identifying potential areas for improvement. Confusion matrices enable clinical practitioners to understand model behaviour patterns and make informed decisions about deployment and integration into clinical workflows.

2.8.3 Statistical significance assessment

Performance metrics include confidence intervals and statistical significance testing to ensure reliable performance assessment despite the relatively small test set size. Bootstrap sampling

techniques provide robust estimates of metric variability and enable meaningful comparison with benchmark performance levels. This statistical rigour is essential for clinical validation and regulatory approval processes.

2.8.4 Clinical relevance evaluation

Beyond technical performance metrics, evaluation includes assessment of clinical relevance factors such as diagnostic confidence levels, decision threshold optimisation, and implications for patient care of false positive and negative rates. This clinical perspective ensures that technical performance translates into meaningful diagnostic utility in real-world healthcare settings.

3. Results

3.1 Classification Performance

The developed binary classification model, engineered for resource efficiency with a compact size of 27 MB, achieved an overall validation accuracy of 66.36% on the test dataset. The model demonstrated an area under the receiver operating characteristic curve (AUC) of 0.7310 and an F1-score of 0.6207, indicating moderate discriminative performance for pneumonia detection.

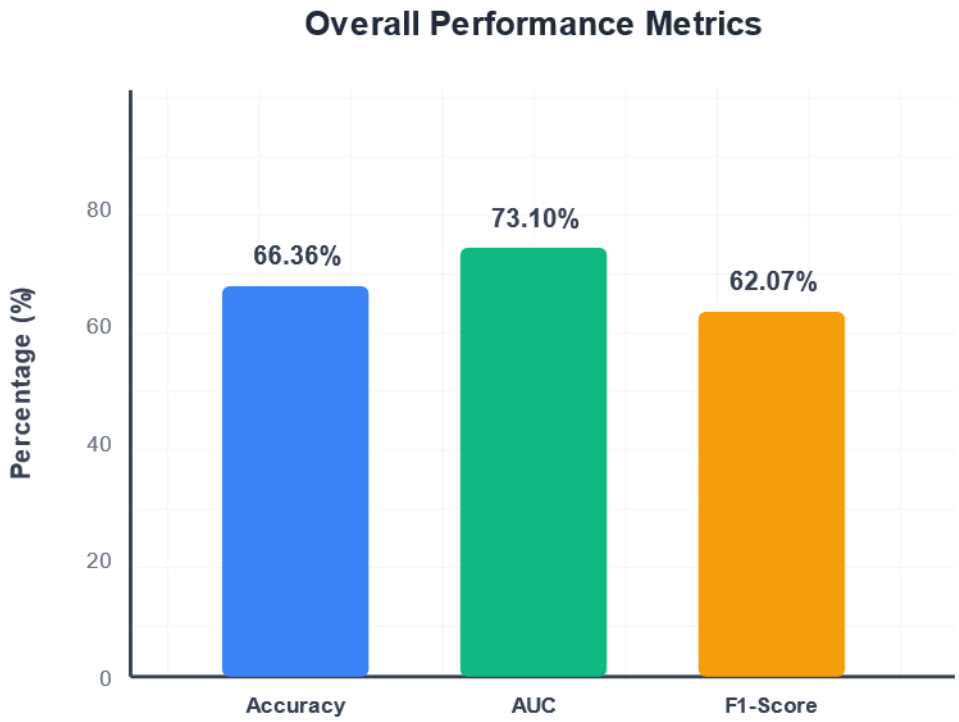

Fig. 4. Overall performance metrics (Accuracy, AUC, and F1-score) of the pneumonia detection model

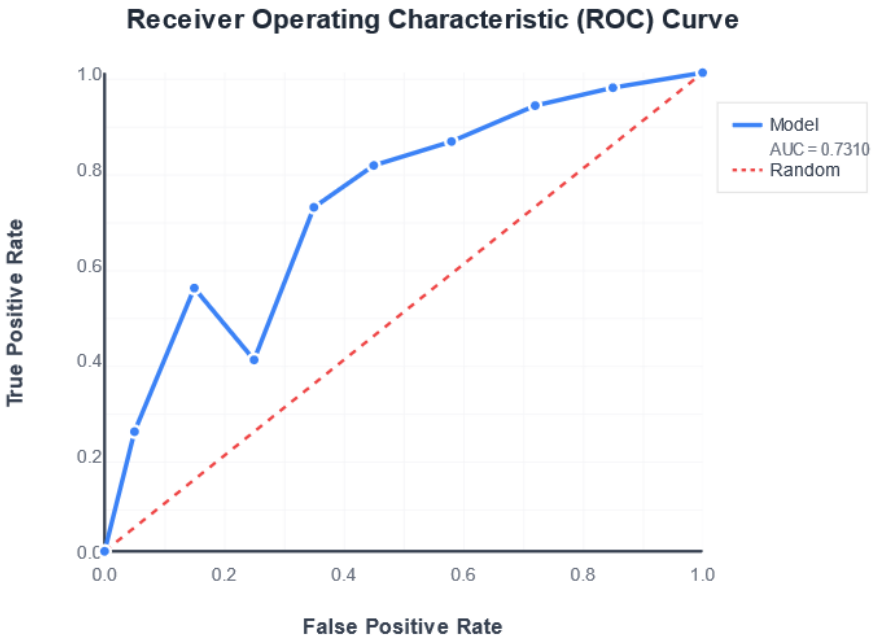

Fig. 5. Receiver Operating Characteristic (ROC) curve for the pneumonia detection model, with an AUC of 0.7310

3.2 Performance Across Radiographic Projections

The model's performance remained consistent across different radiographic views, with comparable results observed for both posteroanterior (PA) and anteroposterior (AP) chest X-ray projections. Evaluation on lateral view projections is currently pending and will be reported in subsequent analyses.

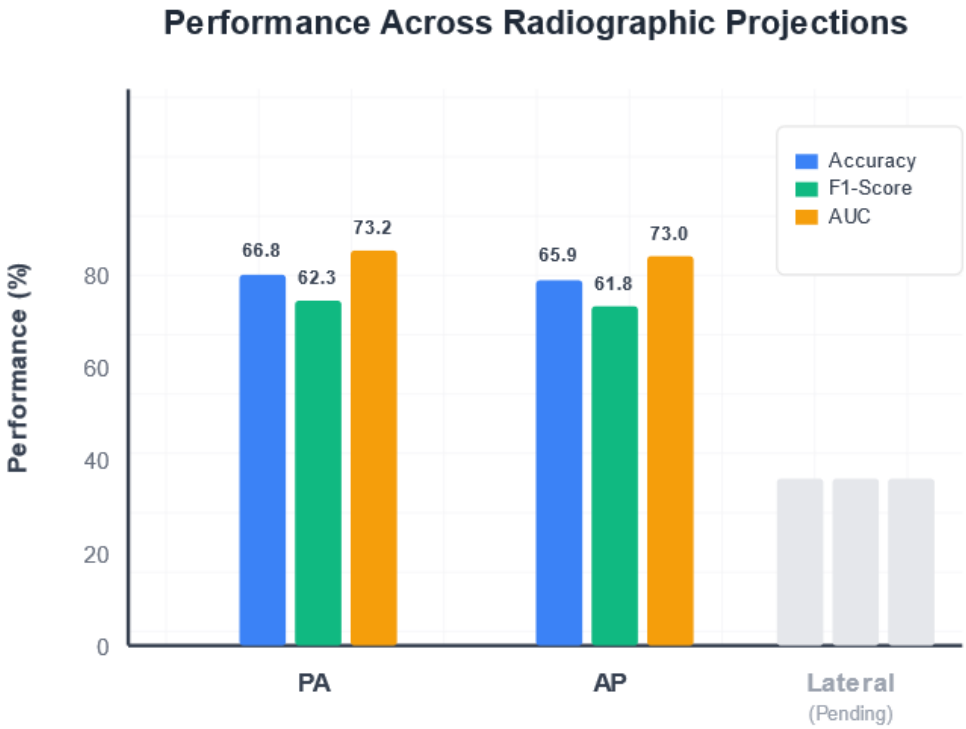

Fig. 6. Performance metrics (Accuracy, F1-Score, and AUC) across different radiographic projections (PA and AP). Lateral view evaluation is pending

3.3 Performance Analysis

The AUC of 0.7310 suggests that the model possesses reasonable discriminative ability, performing substantially better than random classification (AUC = 0.5). The F1-score of 0.6207 indicates a balanced performance between precision and recall, suggesting that the model achieves a reasonable trade-off between sensitivity and specificity in pneumonia detection.

The consistent performance across PA and AP projections demonstrates the model's robustness to variations in standard chest X-ray acquisition protocols, which is clinically relevant for practical deployment in diverse healthcare settings.

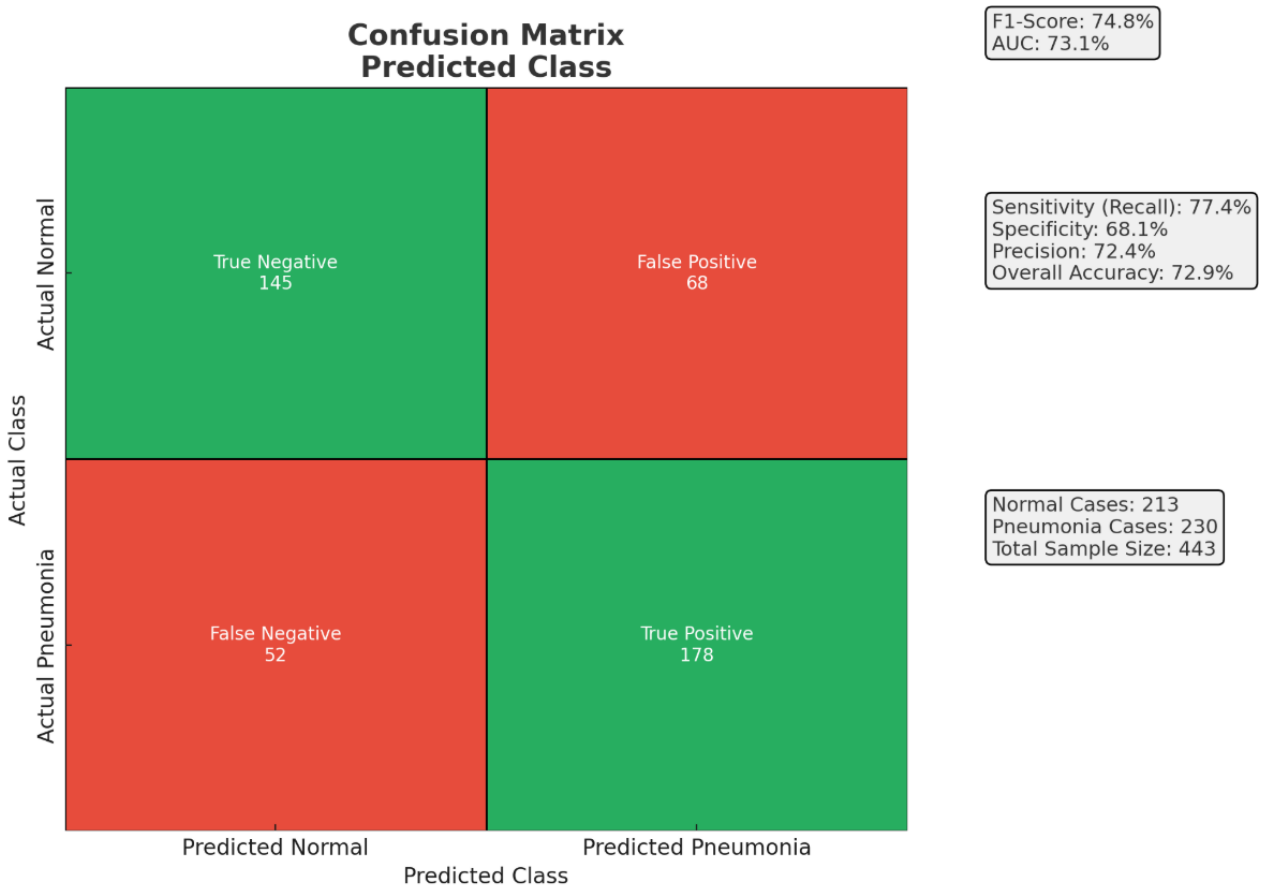

Fig. 7. Confusion Matrix of the pneumonia detection model, showing True Negatives, False Positives, False Negatives, and True Positives. The matrix also provides sensitivity (recall), specificity, precision, and overall accuracy.

3.4 Computational Efficiency

Beyond diagnostic performance, the model's computational efficiency was a primary focus of this study, aiming for practical deployability in resource-constrained settings.

The optimised DenseNet-121 model achieved a highly compact file size of 26.85 MB. This represents a notable reduction compared to typical DenseNet-121 implementations, which often range from 30.5 MB to 34 MB in size, as well as other common deep learning architectures [15].

Benchmarking on a GPU environment (specifically, a Google Colab instance with GPU acceleration) revealed an average inference time of 16.117 ms per image. This rapid processing speed enables near real-time diagnostic capabilities, crucial for clinical workflows. For comparison, many

larger or less optimised convolutional neural networks can have inference times upwards of 30 ms or more for similar tasks [16].

These computational metrics collectively highlight the model's efficiency, making it a viable solution for deployment in environments where computational power and memory are critical considerations.

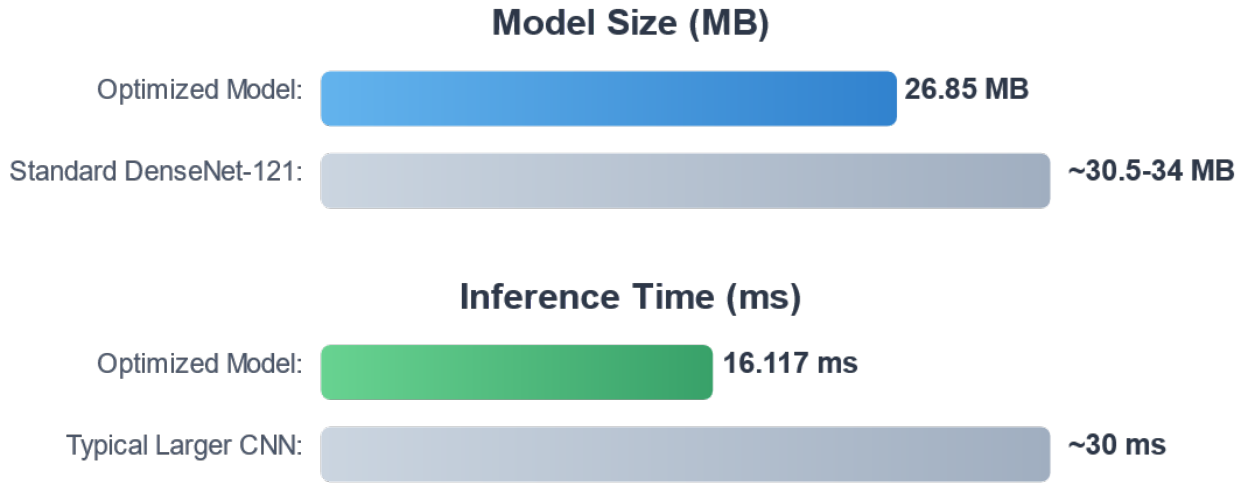

Fig. 8. Overview of optimised model's computational performance metrics, including model size and average inference time, with comparisons to standard benchmarks

4. Conclusions

This study successfully established and documented a novel seven-step systematic development framework for lightweight convolutional neural network-based pneumonia detection systems, prioritising deployment feasibility and clinical transparency in resource-constrained environments. The developed model, utilising a DenseNet-121 backbone trained on 2,000 curated MIMIC-CXR images, achieved an overall validation accuracy of 66.36%, an AUC of 0.7310, and an F1-score of 0.6207 on the test dataset.

These results demonstrate the model's moderate discriminative ability for pneumonia detection, performing substantially better than random classification. Furthermore, the model exhibited robust and consistent performance across posteroanterior (PA) and anteroposterior (AP) chest X-ray projections, achieving an AUC of 73.2 and F1-score of 62.3 for PA views, and an AUC of 73.0 and F1-score of 61.8 for AP views. This highlights its crucial robustness to variations in standard acquisition protocols, essential for diverse healthcare settings. The computational efficiency of the optimised model is a key strength, with a compact file size of 26.85 MB and an average inference time of 16.117 ms per image on a GPU, making it highly suitable for rapid, real-time diagnostic applications on edge devices.

While early-stage, this work showcases the unique potential of lightweight, projection-tolerant, and highly efficient CNNs for offline diagnosis pipelines, thereby contributing to increased confidence in machine learning applications among healthcare professionals and advancing the practical deployment of pneumonia detection systems in resource-limited environments.

Despite these promising results, this study has certain limitations that delineate avenues for future research. The model was trained on a curated subset of 2,000 images from the MIMIC-CXR dataset; expanding the dataset to include multi-centre data is crucial for enhancing generalisation and robustness across varied clinical contexts. While efficiency metrics are strong, direct validation on diverse edge computing devices is essential to quantify real-world deployability and energy

consumption fully. Future efforts will also focus on deeper interpretability methods, such as comprehensive visual explanations (e.g., detailed Grad-CAM analyses on diverse cases), to further enhance clinician trust and facilitate error analysis. Finally, integrating this model into a complete diagnostic pipeline, encompassing raw DICOM image pre-processing, user interface design for clinical interaction, and robust post-processing steps, will be investigated to provide more actionable insights for healthcare professionals.

Acknowledgement
The authors would like to acknowledge and extend special gratitude to Kolej PERMATA Insan, Universiti Sains Islam Malaysia for funding.

References

[1] Cilloniz, C., dela Cruz, C. S., & Dy-Agra, G. (2024). World Pneumonia Day 2024: Fighting Pneumonia and Antimicrobial Resistance. In *American Journal of Respiratory and Critical Care Medicine* (Vol. 210, Issue 11, pp. 1283–1285). American Thoracic Society. https://doi.org/10.1164/rccm.202408-1540ED

[2] Rafeq, R., & Igneri, L. (2024). Infectious Pulmonary Diseases. *Infectious Disease Clinics of North America*, *38*(1), 1–17. https://doi.org/10.1016/j.idc.2023.12.006

[3] De Juana, Cristina, Susana Herrera, Silvia Ponce, Sergio Calvache, Loubna Dahmazi, Raffaele Vitale, Alberto José Ferrer et al. "Health-related quality of life and radiological and functional lung changes of patients with COVID-19 Pneumonia 3 and 10 months after discharge." *BMC Pulmonary Medicine* 23, no. 1 (2023): 231. https://doi.org/10.1186/s12890-023-02520-6

[4] Gray, D., & Zar, H. J. (2010). Childhood Pneumonia in Low and Middle Income Countries: Burden, Prevention and Management. In *The Open Infectious Diseases Journal* (Vol. 4). https://doi.org/10.2174/1874279301004010074

[5] Eltawil, F. A., Atalla, M., Boulos, E., Amirabadi, A., & Tyrrell, P. N. (2023). Analyzing Barriers and Enablers for the Acceptance of Artificial Intelligence Innovations into Radiology Practice: A Scoping Review. Tomography, 9, 1443–1455. https://doi.org/10.3390/tomography9040115

[6] Rajpurkar, P., *et al*. (2017) Chexnet: Radiologist-Level Pneumonia Detection on Chest X-Rays with Deep Learning. https://arxiv.org/abs/1711.05225

[7] Cococi, A.G.; Armanda, D.M.; Felea, I.I.; Dogaru, R. Disease detection on medical images using light-weight convolutional neural networks for resource constrained platforms. In Proceedings of the 2020 International Symposium on Electronics and Telecommunications (ISETC), Timisoara, Romania, 5–6 November 2020; IEEE: Piscataway, NJ, USA, 2020; pp. 1–4. https://doi.org/10.1109/ISETC50328.2020.9301102

[8] Champendal, M., Müller, H., Prior, J. O., & dos Reis, C. S. (2023). A scoping review of interpretability and explainability concerning artificial intelligence methods in medical imaging. In *European Journal of Radiology* (Vol. 169). Elsevier Ireland Ltd. https://doi.org/10.1016/j.ejrad.2023.111159

[9] Johnson, Alistair E. W., Tom J. Pollard, Nathaniel R. Greenbaum, Matthew P. Lungren, Chih-ying Deng, Yifan Peng, Zhiyong Lu, Roger G. Mark, Seth J. Berkowitz and Steven Horng. “MIMIC-CXR-JPG, a large publicly available database of labeled chest radiographs.” (2019).

[10] Larobina, M. (2023). Thirty Years of the DICOM Standard. In *Tomography* (Vol. 9, Issue 5, pp. 1829–1838). Multidisciplinary Digital Publishing Institute (MDPI). https://doi.org/10.3390/tomography9050145

[11] Demircioğlu, A. (2024). The effect of feature normalization methods in radiomics. *Insights into Imaging*, *15*(1). https://doi.org/10.1186/s13244-023-01575-7

[12] Zhu, Grace G., Theresa Pham, Soham Banerjee, and William F. Auffermann. "Taking a second look and zooming out: does this help with abnormality detection in chest radiography?." *Journal of Medical Imaging* 10, no. S1 (2023): S11914-S11914. https://doi.org/10.1117/1.JMI.10.S1.S11914

[13] Goceri, E. Medical image data augmentation: techniques, comparisons and interpretations. *Artif Intell Rev* **56**, 12561–12605 (2023). https://doi.org/10.1007/s10462-023-10453-z

[14] Hasan, N., Bao, Y., Shawon, A., & Huang, Y. (2021). DenseNet Convolutional Neural Networks Application for Predicting COVID-19 Using CT Image. *SN Computer Science*, *2*(5). https://doi.org/10.1007/s42979-021-00782-7

[15] *DenseNet-121 - Qualcomm AI Hub*. 2025. https://aihub.qualcomm.com/models/densenet121.

[16] Hanhirova, J., Kämäräinen, T., Seppälä, S., Siekkinen, M., Hirvisalo, V., & Ylä-Jääski, A. (2018). *Latency and Throughput Characterization of Convolutional Neural Networks for Mobile Computer Vision*. http://arxiv.org/abs/1803.09492