



IMPACT OF IMAGE PREPROCESSING AND MODEL FINE-TUNING ON CHEST X-RAY BASED PNEUMONIA CLASSIFICATION

Muhammad Ramzan Yasin

ramzan4167768@gmail.com

Faculty of Computer Science and Information Technology, The Superior University Lahore

H. M. Shahzad

shahzad.dar@superior.edu.pk

Faculty of Computer Science and Information Technology, The Superior University Lahore

Meer Usman Amjad

meerusman73536@gmail.com

Faculty of Computer Science and Information Technology, The Superior University Lahore

Abstract

Despite improvements in clinical management, pneumonia is still a significant respiratory infection that demands timely radiological evaluation for early clinical intervention and minimizing preventable complications. While chest X-ray imaging is widely used as it is accessible, inexpensive, and informative for initial diagnosis, knowledge of pattern recognition is not commonplace due to subtle patterns of infection, variation between images and there is a lack of radiological skills in resource constrained areas. In this study, the impact of image preprocessing and fine-tuning on automated pneumonia categorization on chest X-ray pictures is explored. For dichotomous classification of pneumonia and normal, a transfer learning framework was designed based on the models: Xception, EfficientNetB0, and DenseNet201. Our pipeline involved image resizing, model-specific normalization, data augmentation, oversampling, deep radiographic feature extraction and fine-tuned classification layers to promote generalization. The metrics used to assess performance were based on training-validation curves, precision, accuracy, recall, F1 score and a confusion matrix. Experimental results show that Xception achieves the highest accuracy of 94.23% in the test when compared to EfficientNetB0 and DenseNet201. The results show that with proper preprocessing and selective fine-tuning, it is possible to enhance the performance of computer-aided pneumonia classification for reliability and accuracy.

Keywords: *Pneumonia Classification, Chest X-Ray, Deep Learning, Transfer Learning, Image Preprocessing, Fine-Tuning, Xception, EfficientNetB0, DenseNet201*

1. Introduction

Pneumonia is a serious viral and inflammatory condition of the parenchyma of the lungs that primarily impacts the alveoli and is a significant burden of health care globally [1]. They can result from a bacterial, viral or fungal infection and are often linked to coughs, fever, a shortness of breath, respiratory problems and discomfort in the chest. Infection causes filling of the alveoli with fluid, pus and inflammatory exudates leading to the normal function of gas exchange being impeded and can cause severe complications if not diagnosed in time. Children, elderly people, immune compromised patients [2] and communities with limited resource access are most at risk as experts are often not available to provide professional radiological analysis.

When it comes to the first evaluation of pneumonia, one of the most widely-used diagnostic tools is chest X-ray imaging due to its economy, speed, non-invasiveness and wide availability during clinical practice [3]. It allows for the identification of the radiographic characteristics of pneumonia (pulmonary opacities, consolidation, infiltrative pattern and abnormal lung-field structure). The ability to interpret chest radiographs is still difficult, however, as pneumonia changes may be subtle with the features often being difficult to distinguish from other thoracic conditions such as pulmonary oedema, tuberculosis, atelectasis or malignant disease [4]. Even higher diagnostic uncertainty exists due to the possible differences in image contrast, exposure,



patient positioning, acquisition protocol, and resolution. The problems illustrate the need for automated computer aided diagnosis solutions, which can offer the doctors consistent and objective assistance.

Medical image processing has greatly benefited from deep learning, which extracts image discriminative visual features from image data automatically. A CNN is extremely well suited for the task of categorizing chest X-ray images, as it can learn both hierarchical representations (low level primitive edge features and textured surfaces, as well as intermediate level anatomical patterns, and high-level level signs of disease), and disease-specific radiographic cues. CNN-based models are considered suitable for pneumonia detection as clinically relevant features might be low-contrast, spatially distributed, but hard to describe by manual feature engineering.

When annotated clinical image datasets are scarce, transfer learning enhances applications of deep learning in medical imaging further. However, pre-trained convolutional models take advantage of a reliable visual representation learned on large datasets and avoid the process of building deep structures from scratch [6]. Such models can be adapted to pneumonia-specific radiograph patterns, discarding the original classification layers after making a few additional layers modifications, and still have useful low-level functionality transferred to the model. This helps alleviate the risk of overfitting, boosts convergence, and better generalizes.

In this research paper, the effect of image pretraining and model fine-tuning on Xception was investigated to categorize pneumonia from chest X-rays, EfficientNetB0 and DenseNet201. The proposed framework enlarges image dimensions, normalizes data for the individual network, augments the dataset, oversampling, transfers the model from other domains, and selectively fine-tunes to enhance the reliability of classification. The primary goal is to find the best transfer learning architecture for binary classification of Pneumonia and Normal chest X-rays in a reproducible and controlled setting that is clinically relevant.

2. Literature Review

Diagnosis support through image analysis of chest x-ray image is one of the important fields of research in the medical area since timely, reliable and cost effective help is vital for the effective medical decision making process. Despite advancements in antibiotics, pneumonia continues to be a major respiratory disease and an appropriately early radiographic case interpretation is crucial to efficient treatment planning and enhanced clinical result [7]. Previous computer-aided diagnostic methods used only manually-defined characteristics of form, texture, edge, intensity, and statistics. While these approaches enabled the early automated screening measures, these approaches were restricted due to requiring manual intervention for feature design, and with very low adaptability to complex variations in the radiographic images [8]. Assessing patterns of pneumonia can be tricky and include a few found patterns such as faint opacity, irregular consolidation, low-contrast infiltrate, and lung-field abnormalities that are spatially distributed across the image, which makes it hard to resolve by just 'looking' at fixed handcrafted representations.

The role of deep learning has been brought to the fore by its ability to extract some clinically meaningful and hierarchical representations directly from chest x-rays automatically, thus improving pneumonia detection [9]. Convolutional Neural Networks are able to learn discriminative features at multiple levels, with deeper layers learning while lower layers aggregate edges, gradients and disease-specific (pulmonary opacity, consolidation and inflammatory infiltrates) textures. The representation learning feature enables CNN based



systems to detect nuclear abnormalities with less ability to describe them by a traditional feature engineering. [10]. But the quality of the dataset, preprocessing consistency, augmentation strategy, class balance, depth of fine-tuning and evaluation of the model performance on unseen images have a great impact on the performance of the model.

Various convolutional models are developed for pneumonia classification, such as the VGG16, ResNet, InceptionV3 model, DenseNet, EfficientNet, MobileNet and Xception architectures. The structural designs, the costs of computing roughly with these, the propagation of selected features and the representation capabilities differ for these models. Residual connections (ResNet) are used for deeper learning, and DenseNet for feature dissemination and depiction. DenseNet improves the flow of gradients and the re-use of features by allowing dense inter-layers connections between them, while ResNet allows more sophisticated learning due to the residual connections [11]. EfficientNet reaches an efficient corner between accuracy and computation economy by having compound scaling [12] stating the width, the depth and the input precision used. Xception is suitable for enhancing the feature representation of an improved chest X-ray image due to its ability to learn the representation in spatial and channel respective manner through the depth wise separable convolutions.

Because of the extensive application of transfer learning in medical image classification, medical X-rays labeled with experts are sometimes scarce. Instead of Transfer learning uses convolution backbones already developed and that have a wide visual representation from large numbers of data, and deep networks are developed from scratch [13]. These architectures can be adapted to the pneumonia-specific pattern of radiographic findings by changing the classification layers, and fine-tuning rather deeper layers, which can help reduce overfitting and improve convergence.

In addition, image processing, augmentation and imbalanced treatment are also very important for getting stable classification of pneumonia. Pictures taken of a patient's chest using X-rays often vary in many characteristics, such as quality, brightness, contrast, noise and location of the patient. Preprocessing, e.g., resizing, normalization, and model-specific preprocess, standardizes the input images, and augmentation is used to increase the variation of the input images for training to improve the generalization ability. Due to class imbalance, models could be overfitted towards the dominant class and making predictions that may not be sufficiently reliable to be used clinically [14]. Oversampling, class weighting, balanced sampling and focal loss are therefore common methods to control the imbalance between classes to reduce imbalance-related bias and enhance the learning of minority classes.

Results were deemed satisfactory for the selected basis research (Interpretable Deep Learning for Pneumonia Detection Using Chest X-Ray Images), a ResNet50 based model with the addition of Layer-wise Relevance Propagation achieved 91.00% accuracy. It gave some interpretations, but it concentrated primarily on the single architecture. Focusing on the same framework with preprocessing, augmentation and oversampling and selective refinement and standard transfer learning, the present study compares Xception, EfficientNetB0 and DenseNet201. The present study is a controlled comparison providing a firmer base for evaluation of the effect of architects choice and training approach on classification accuracy for pneumonia.

2.1 Contributions

The following is a summary of this research work's main contributions:

1. An automatic pneumonia classification framework based on transfer learning was developed to do binary classification tasks for chest X-ray images (Healthy and Pneumonia categories).



2. Systematically preprocessing strategy was adapted, including image normalisation according to the model, image scaling, augmentation and oversampling to enhance input uniformity and mitigate class imbalance.
3. Four highly structured CNN architectures, i.e. Xception, EfficientNetB0 and DenseNet201 were implemented and trained, with the same experimental setup.
4. Convolutional feature maps, dropout regularization, thick layers, batch normalization, global average pooling were used for deep radiographic feature extraction and feature fusion.
5. Comparative analysis was performed and the performance of Xception got the highest testing accuracy rating of 94.23%, which proves that the system performs well in the categorization of pneumonia.

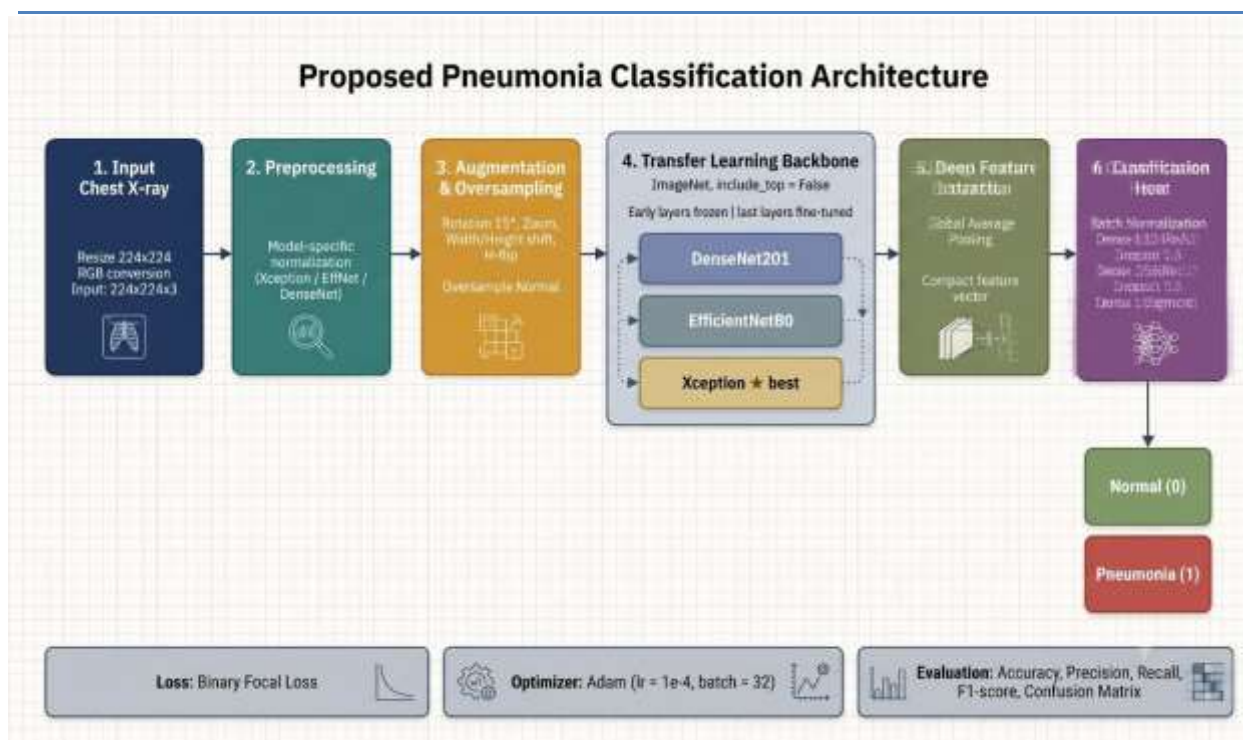
3. Materials and Methods

The proposed methodology was targeted at exploring the effect of image processing and fine-tuning of models on the classification of pneumonia using chest X-rays. The entire set-up encompassed the steps of dataset acquisition, image preprocessing, classes balancing, data augmentation, transfer learning, selective fine-tuning, deep radiologic feature extraction, binary classification, performance evaluation. Three pre-trained convolutional models (Xception, EfficientNetB0, DenseNet201) were used to classify chest radiographs as either pneumonia or normal.

The methodology was set up to ensure a fair, reproducible and controlled comparison of all the selected models. To pass to the transfer learning networks, the images were preprocessed first by resizing and preprocessed by design of the architecture. In order to be passed in to the transfer learning networks, all the images were preprocessed through resizing and architecture-specific preprocessing. To overcome imbalanced training data, oversampling was performed and performed only for the training set to enhance the image variability in the training data to prevent overfitting. Implemented the experiments in python and using libraries TensorFlow, Keras, NumPy, Pandas, Scikit-learn, Matplotlib, and Seaborn. A shared pipeline and different interchangeable transfer learning backbones are included in the overall proposed architecture depicted in Figure 1.

Figure 1. Overall Projected Pneumonia Classification Construction.

The features after extracting and classification head has been designed in a modified way, and deep Xception model proved to be the best performing model, ultimately using a single unified pipeline: input, preprocessing, augmentation and oversampling, a transfer learning backbone (Xception, EfficientNetB0, or DenseNet201), and a designed-and-modified classification head.



3.1 Dataset Overview

The Pneumonia information from chest X-ray published in the open Kaggle was used in this study. The pictures of the chest radiography in the dataset are segmented into two different diagnostic classes, namely pneumonia and normal [15]. The first training set contained 5216 images with an imbalance in the classes; with 1341 Normal pictures and 3875 Pneumonia images. To overcome this problem only the minority Average class was oversampled. The image sets were balanced and both sets after oversampling had 3,875 images each yielding a training set of 7,750 images.

The test set remained completely independent for the independent evaluation, the training set was saved separately from the validation set and all the test sets were balanced. The test set was only compiled of 624 unseen chest X-ray images including 234 healthy and 390 images for the pneumonia dataset. This dataset setup allowed a controlled evaluation of the parameters related to the preprocessing, oversampling, transfer learning and fine-tuning in order to correctly classify pneumonia as binary classification.

3.2 Data Splitting and Validation Strategy

Structured data splitting was employed to ensure fairness in assessment and prevent data leakage. The original training dataset included 5,216 images comprising 1,341 Normal images and 3,875 images of Pneumonia. Due to class imbalance, oversampling was applied only to the minority Normal class, producing a 7,750-image training dataset that is balanced, with 3,875 photos in each class.

The balanced dataset was divided into two sets – a validation set of 85% and a training set of 15%. 624 unseen photos from the first test set was kept completely separate from training, validation, augmentation, and oversampling processes.



Table 1. Data Splitting Strategy.

Dataset Stage	Split Ratio	Normal Images	Pneumonia Images	Total Images	Purpose
Original Training Dataset	—	1,341	3,875	5,216	Dataset before class balancing
After Oversampling	—	3,875	3,875	7,750	Balanced dataset before train-validation split
Training Set	85%	3,293	3,294	6,587	Model training and feature learning
Validation Set	15%	582	581	1,163	Hyperparameter tuning and overfitting monitoring
Test Set	Separate test folder	234	390	624	Final unbiased evaluation on unseen images

3.3 Methodology

The proposed methodology was developed as a controlled deep learning framework to investigate the impact of image preprocessing and model fine-tuning on chest X-ray based pneumonia classification [16]. The complete workflow included image acquisition, diagnostic label assignment, preprocessing, augmentation, class balancing, transfer learning, selective fine-tuning, deep radiographic feature extraction, feature fusion, binary classification, and model evaluation [17]. Xception, EfficientNetB0, and DenseNet201 were trained in the same experimental setup to provide an impartial and repeatable comparison.

Initially, Binary labels were applied to chest X-ray pictures that were obtained from structured dataset directories representing Normal and Pneumonia classes. The images were resized and preprocessed according to the particular architecture in order to standardize all images. The split proportion of the dataset between training and test set was used to augment the training set to improve image variance but there was an oversampling process to reduce the class imbalance and to assist balanced learning across both classes [18]. The pre-trained convnet backbones were modified by discarding their existing classification head, and then substituted with custom classification heads that utilized binary classification layers. To assess the standard performance metrics and confusion matrix analysis were used on the final an autonomous test sets.

3.3.1 Comparative Base Study Selection

The chosen comparative study was: “Interpretable Deep Learning for Pneumonia Detection Using Chest X Ray Images” in order to set a threshold. The base study used ResNet50 model with LwRP and was able to achieve 91.00% accuracy for the pneumonia detection task on chest X-Rays. It was chosen since both the research domains and Pneumonia and normal examples are classified in binary classification. In the light of the previous direction, the present research work extends this approach by comparing the three models being studied (Xception, EfficientNetB0, DenseNet201) on a controlled pipeline that considers preprocessing,



augmentation and handling of oversampled data, transfer learning, selective fine-tuning and independent testing.

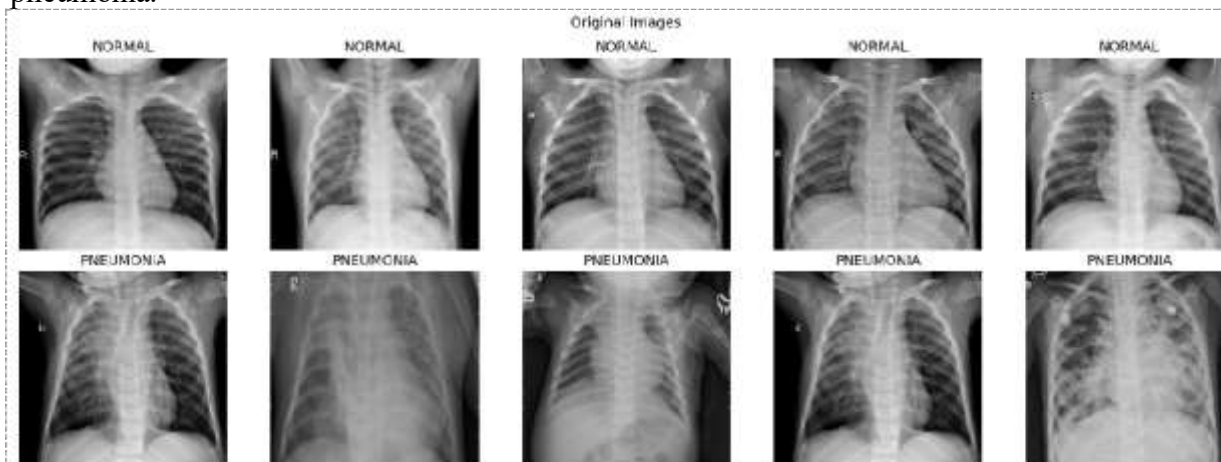
3.3.2 Image Preprocessing

To enable advanced convolutional learning with chest radiographs, image preprocessing was conducted to create a common structure for the images [19]. Transfer learning training of the pictures was performed on a square size of 224×224 pixels for Xception, EfficientNetB0 and DenseNet201 [20] to ensure that the input values remained constant. Architecture-specific preprocessing functions were used to match the distributions of input pixels with the requirements of each pre-trained model, since transfer learning architectures are sensitive to pixel scaling and normalization range.

Data augmentation was applied exclusively to the subset used for training. Rotation, zooming, width and height shifts, shearing, brightness modulation, horizontally flipping, and nearest-neighbor fills were all part of the augmentation method.

These changes resulted in realistic variations in the radiograph and involved a greater diversity in the training process to reduce overfitting. There wasn't an enhancement of test and validation images to maintain an unbiased level of performance evaluation. The results are presented in the form of Representative Samples for both types of pneumonia (NORMAL and PNEUMONIA) are depicted in Figure 2.

Figure 2. Normal and Pneumonia Chest X-ray Samples. Images of chest radiographs presented as a visual comparison of the characteristics of areas of healthy lung and areas of pneumonia.



3.3.3 Transfer Learning

The representation-learning ability of convolutional networks (CNNs) that had been pretrained on large scale data were utilized through transfer learning. Training deep models from scratch with limited medical data could expand the computational costs along with an over-fitting dangers [21]. As a result, pre-trained bases based on convolution (ConvNets) were chosen as feature extractors, and task-specific classification heads were applied for the binary pneumonia detection task.

The following three architectures were implemented: Xception, EfficientNetB0 and DenseNet201. Spatially and channel-wise information is learnt by the Xception network using depthwise separable convolutions, which are good at extracting features from the radiographic image. EfficientNetB0 provides a useful alternative between accuracy and processing performance by using compound scale to regulate depth, width, and resolution. DenseNet201 uses dense inter-layer communication to minimize information loss, enhance gradient flow,



and encourage feature reuse [22]. Selective fine-tuning was performed by freezing earlier layers and unfreezing selected deeper layers, allowing the models to adapt pre-trained ImageNet representations to pneumonia-specific radiographic patterns.

3.3.4 Deep Radiographic Feature Extraction and Fusion

In this study, separate handcrafted morphological operations such as erosion, dilation, opening, closing, thresholding, contour detection, or shape descriptor extraction were not explicitly applied. Instead, feature extraction was performed through automated deep radiographic representation learning using convolutional layers. Therefore, feature extraction refers to CNN-based hierarchical radiographic learning rather than manual morphological processing.

The convolutional layers extracted multi-level chest X-ray information. Early layers captured intermediate layers, low-level patterns like texture and outlines learned spatial and tissue-level variations, and deeper layers extracted high-level characteristics of pneumonia, such as pulmonary impenetrability, consolidation, infiltrative regions, and abnormal lung-field representations [23]. Feature fusion was performed through the customized classification head, where Global Average Pooling compressed feature maps into compact vectors, followed by Batch Normalization, Dense layers, and Dropout for discriminative learning and overfitting control.

3.3.5 Classification

Chest X-ray pictures were binary predicted into the Healthy and Pneumonia categories during the classification step. After deep feature extraction and feature fusion, the final dense layer used sigmoid activation is used to provide a probability score in the range of 0 to 1. values over or equivalent to 0.5 were categorized as pneumonia, while values below 0.5 were categorized as normal. A threshold value of 0.5 was taken.

The classification head was modified with the use of Global Average Pooling, Batch Normalization, Dense Layers as well as Dropout Regularization, which are included under Global Average Pooling. Global Average Pooling (including Batch Normalization, Dense Layers, and Dropout Regularization) was applied to the classification head. This design minimized the dimensionality of features, influenced the stability of training, found non-linear decision spaces and minimized over fitting.

3.3.6 Hyperparameter Settings

For the models trained in the totally models, fully-founded parameters were used. The number of images per batch, $N = 32$, the size of the input picture was standardized to 224×224 pixels, and a reliable convergence property was ensured by adopting the Adam optimization program that possesses adaptive learning capability. The learning rate of 0.0001 was employed when training and fine tuning. A binary focal hurt, which is used to enhance learning from challenging samples and minimize class sensitive misclassification, was used. The sigmoid activation function was used in the output layer to perform a binary classification.

In the case of Xception, all the previous layers were frozen, but the last 40 layers were optimized. In DenseNet201, while all layers prior to it were kept frozen, 50 layers following were adjusted. The classification head consisted of two layers with 0.5 and 0.3 dropping out rates, and 512 and 256 neurons, respectively. Advanced stopping methods used were EarlyStopping, ReduceLROnPlateau, and ModelCheckpoint, to adjust the learning rate and to avoid overfitting issues. when validation performance plateaued [24], and preserve the best-performing model weights.

3.4 Mathematical Formulation

One way to construct the suggested framework is as a binary based supervised classifying issue. Let's describe the chest X-ray information as:



$$D = \{(x_i, y_i)\}_{i=1}^N, y_i \in \{0, 1\} \quad (1)$$

With $y_i = 0$ and x_i is the input chest X-ray picture denotes Normal, and $y_i = 1$ denotes Pneumonia. Each image was resized and preprocessed as:

$$x_i^{\wedge} = f_{pre}(x'_i), x'_i \in \mathbb{R}^{224 \times 224 \times 3} \quad (2)$$

where x'_i is the resized image and $f_{pre}(\cdot)$ represents the model-specific preprocessing function for Xception, EfficientNetB0, or DenseNet201. The balanced training set after oversampling can be represented as:

$$D_{bal} = D_{pneumonia} \cup D_{normal}^{os} \quad (3)$$

where D_{normal}^{os} denotes the oversampled Normal class. The preprocessed image was then passed through a pre-trained convolutional backbone:

$$F_i = B(x_i^{\wedge}; \theta_b) \quad (4)$$

where $B(\cdot)$ denotes the backbone network and F_i represents the extracted feature maps. The feature maps were transformed into a condensed representation using Global Average Pooling:

$$z_k = \frac{1}{H \times W} \sum_{h=1}^H \sum_{w=1}^W F_{h,w,k} \quad (5)$$

The fused feature vector was passed through dense layers, and the final sigmoid function generated the pneumonia probability:

$$y_i^{\wedge} = \frac{1}{1 + e^{-s_i}} \quad (6)$$

The predicted class was assigned as:

$$c_i^{\wedge} = 1, \text{ if } y_i^{\wedge} \geq 0.5 \quad (Pneumonia) \quad (7a)$$

$$c_i^{\wedge} = 0, \text{ if } y_i^{\wedge} < 0.5 \quad (Normal) \quad (7b)$$

where $\hat{c}_i = 1$ indicates Pneumonia and $\hat{c}_i = 0$ indicates Normal. Binary focal loss was used during training:

$$FL(p_t) = -\alpha(1 - p_t)^{\gamma} \log(p_t) \quad (8)$$

Here γ is the concentrating parameter, α is balance factor and p_t is expected frequency of the true class. The key steps of the proposed pipeline are mathematically formulated such as preprocessing, class balancing, feature extraction using the transfer learning model, feature fusion, sigmoid class, and loss optimization.

3.5 Evaluation Strategy

The method of assessment was developed to determine the validity and suitability of clinical reliability and performance in classification. All the training models were assessed using an independent test set of 624 unseen chest X-rays, of which there were 234 normal and 390 pneumonia cases. This evaluation set was neither used for model augmentation or oversampling nor used for validation, thus yielding an unbiased estimate of the model's generalization.

The results were analysed using accuracy, precision, recall, f1 score, confusion matrix and training-validation curves to determine accuracy. Real pneumonia were successfully recognized by recall, while accuracy assessed the overall correctness, precision assessed its dependability with predictions of pneumonia and F1 score provided a balanced assessment of precision and recall. Assessment scores were described as:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (9)$$

$$Precision = \frac{TP}{TP + FP} \quad (10)$$

$$Recall = \frac{TP}{TP + FN} \quad (11)$$

$$F1 - Score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (12)$$



TP = true positive; TN = true negative; FP = false positive; FN = false negative. The prediction behavior was explored class-wise by confusion matrix and training-validation curves were used to analyze convergence, over-fitting, under-fitting and stability of generalization.

4. Results and Discussion

Tests revealed that preparation, enhancement, oversampling, transfer learning and some fine-tuning performed in the chest X-ray domain were helpful in improving the performance of the X-rays for pneumonia classification. Irrespective of the models, all had been subjected to the same experimental conditions. The comparative results are shown in Table 2.

Table 2. Comparative Performance of Transfer Learning Models.

Model	Testing Accuracy	Normal Precision	Pneumonia Precision	Normal Recall	Pneumonia Recall	Weighted F1-Score
Xception	94.23%	0.99	0.92	0.86	0.99	0.94
EfficientNetB0	92.31%	0.95	0.91	0.84	0.97	0.92
DenseNet201	90.00%	0.87	0.91	0.85	0.93	0.90

The architecture with the highest number of successes to test was Xception with success rate of 94.23% and pneumonia recall of 0.99, which demonstrates excellent sensitivity in diagnosis of infected cases. EfficientNetB0 achieved 92.31% accuracy and provided a balanced performance with computational efficiency. DenseNet201 had an accuracy of 90.00%, which demonstrated reasonable, but lower level of performance.

4.1 Comparison with the Selected Base Study

While comparing the proposed model with the selected base study (ResNet50 with LRP) which achieved the accuracy of 91.00%, it was found that CNN models with multi-level ResNet50 can outperform the selected base study model on the basis of the achieved accuracy. From the observations mentioned in Table 3, Xception model has higher accuracy than the base study by 3.23 percentage points and EfficientNetB0 model also show higher accuracy as compared to the benchmark [25]. In order to compare the proposed Xception model with the selected base study, “Interpretable Deep Learning for Pneumonia Detection Using Chest X-Ray Images”, we utilized one of the most common ResNet50-based interpretable model with 91.00% accuracy. The fine-tuned Xception model was suggested as getting the best test result (94.23%), with a 3.23% increase compared to the base study. Another, EfficientNetB0, also has an accuracy higher than the base study at 92.31%, while DenseNet201 has an accuracy of 90.00%. The results support the ability of the proposed transfer learning paradigm with respect to pneumonia classification accuracy and reliability using Xception.

Table 3. Comparison with the Selected Base Study.

References	Year	Method	Dataset	Zone Analysis	Level	Accuracy (%)	F1-score (%)
[21]	2022	Attention + Modified DenseNet	Chest X-Ray Images (Pneumonia)	Only feature analysis	global	93	94.5
[22]	2022	DenseNet169 + MobileNetV2 + ViT	Chest X-Ray Images (Pneumonia)	Only feature analysis	global	94	93



[23]	2023	VGG16 with dense CNN	Chest X-Ray Images (Pneumonia)	Only global feature analysis	92	93
[24]	2023	CNN Ensemble Model	Chest X-Ray Images (Pneumonia)	Only global feature analysis	84.4	88.56
[25]	2024	EfficientNetB0 + DenseNet121 with Attention	Chest X-Ray Images (Pneumonia)	Restricted focus on the picture as a whole rather than only certain areas	95.1	96.0
[26]	2025	ResNet50 + interpretability (LRP, CAMs, adversarial training, spatial attention)	Chest X-ray Images (Pneumonia)	Highlighting of coarse areas across the picture	91	—
Proposed Xception	2026	Transfer Learning + Fine-tuning (Xception)	Chest X-Ray Images (Pneumonia)	Global feature analysis (whole image)	94.23	94.0

To evaluate the results, training accuracy was compared with the selected base study which gave 91.00% accuracy and the accuracy achieved by the proposed model Xception-based model was 94.23% with an improvement of 3.23% in accuracy compared with the selected base study.

This enhancement in performance of Xception is due to its use of depthwise separable convolution, which efficiently extracts both the spatial aspects and channel-wise radiographic properties. Infection Patterns are frequently distributed across multiple regions in the lung, have a low contrast, and are weak, which is beneficial for pneumonia classification: Chest X-rays [26]. The confusion matrix also revealed that Precision of Xception improved significantly, especially in minimising false-negative predictions, essential in providing a reliable support for pneumonia screening. The confusion matrices for Xception, EfficientNetB0 and DenseNet201 are shown in Figures 3, 4 and 5 respectively.

Figure 3. Confusion Matrix of Xception. The prediction distribution for the top performing Xception model on previously unseen experiment set.

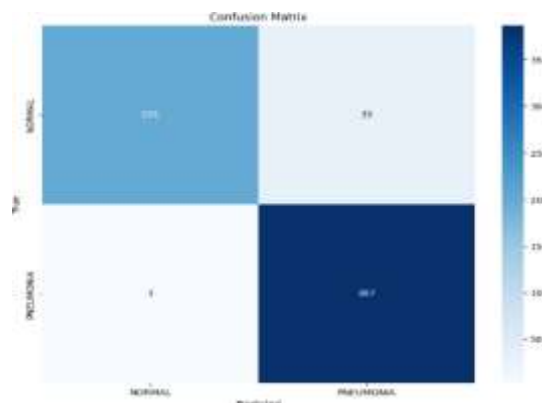




Figure 4. Confusion Matrix of EfficientNetB0. The assessment set had an efficientNetB0 classification accuracy of 98.8%.

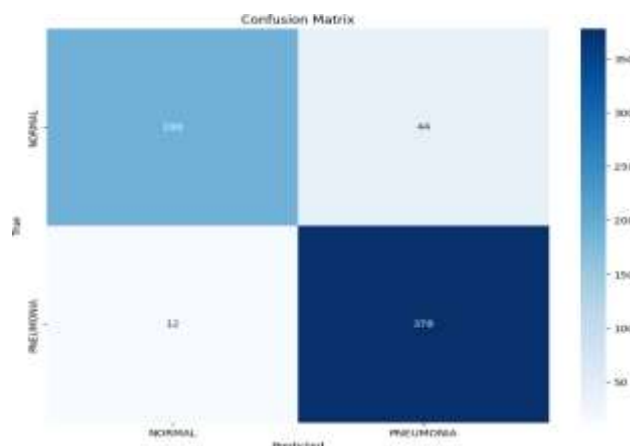
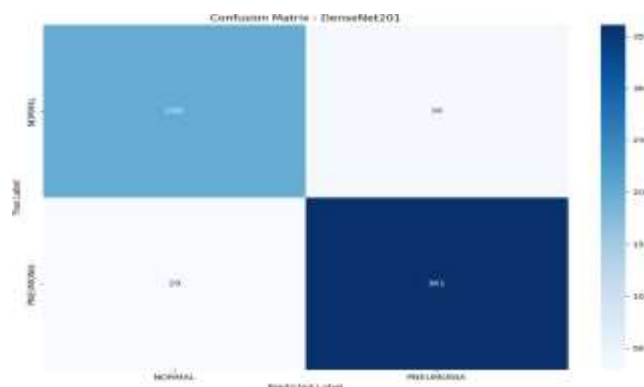


Figure 5. Confusion Matrix of DenseNet201. Classification accuracy on unseen test images of DenseNet201.



The Prediction behavior of the three models towards the unseen test images (624 images with 390 Normal + 234 Pneumonia) is summarized by the confusion matrices given in figure 3, 4 and 5. Xception gave the most clinically favorable distribution with-wise 387 Pneumonia cases and only 3 false-negative cases, which accounts for the high Pneumonia recall it achieved of 0.99. EfficientNetB0 correctly classified 378 images including Pneumonia and 190 images including Normal, and had a comparatively lower sensitivity on the Normal class with 12 FN, 44 FP. DenseNet201 achieved the highest accuracy in predicting the Pneumonia and Normal images (361 images for Pneumonia and 200 for Normal), while it gave the most number of false negatives (29) meaning it failed to predict the Pneumonia with the highest accuracy. Overall, both the matrices showed that all three models exhibited good sensitivity on pneumonia while minimizing the most significant type of inaccuracy in pneumonia screening (false negative): Xception did this best.

The training loss curve and precision curves for the three models are presented in Figures 6 and 7, respectively. Figure 8 presents the validation loss curve and the prediction precision curves for the three models. They were used to examine the convergence behavior, overfitting and generalization stability over the number of training epochs, and proved that it facilitated stable learning with the help of both augmentation and oversampling techniques.



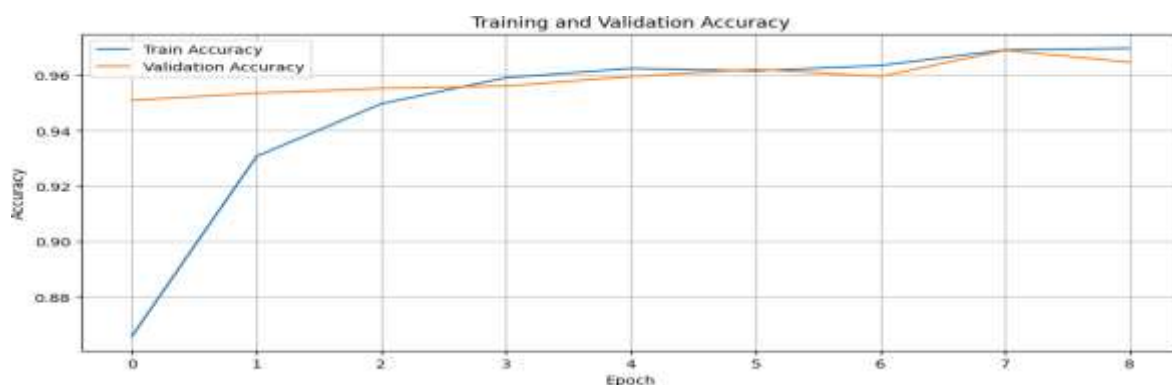
Figure 6. Training and Validation Curves of Xception. The top performing model, with regards to accuracy (left) and loss (right) across training time.



Figure 7. Training and Validation Curves of EfficientNetB0. Accuracy (left) and loss (right) over time.



Figure 8. Training and Validation Curves of DenseNet201. Accuracy (left) and losses (right) by each era of training.



In general, the results validate the effectiveness of data augmentation, overcoming overfitting, a successful Prey improvement of class balance through oversampling, and the capacity of the



models to fine-tune pre-trained ImageNet features with pneumonia-related chest X-ray patterns. The proposed framework based on Xception was thus found as the most promising framework for diagnosis, however should not be used exclusively for the radiological diagnosis but in addition as an assistive tool to the radiologist.

5. Conclusion

The impact of preprocessing of chest X-rays and model fine-tuning on pneumonia detection task using Xception, EfficientNetB0 and DenseNet201 was the focus of this study. The pipeline proposed involved augmenting the data set [27], oversampling, transfer learning, resizing, model-specific handling, model-specific preprocessing, selective fine-tuning, deep feature extraction and binary classification.

Based on findings, Xception was found to have got the highest accuracy during testing at 94.23%. EfficientNetB0 with 92.31% and DenseNet201 with 90.00%. The pneumonia recall of 0.99 indicates good potential for clinical screening support if reliably recognising infected individuals is crucial. The model proposed using the Xception network showed a significant improvement as compared to the selected base study with accuracy of 91.00%.

Overall, the study results indicate that the addition of architecture-specific pre-trained models, class balance, augmentation, transfer learning, and selective fine-tuning to computer-assisted pneumonia diagnosis using X-rays of the chest can be beneficial. These are just some of the possibilities for further study that could result in bigger multi-center databases, lung-region segmentation, and explainable artificial intelligence for better clinical reliability.

6. References

- [1] P. S. Hill, E. Pavignani, M. Michael, M. Murru, and M. E. Beesley, "The 'empty void' is a crowded space: health service provision at the margins of fragile and conflict affected states," *Conflict and Health*, vol. 8, no. 1, p. 20, Oct. 2014.
- [2] S. C. Castle, "Impact of age-related immune dysfunction on risk of infections," *Zeitschrift für Gerontologie und Geriatrie*, vol. 33, no. 5, pp. 341–349, 2000.
- [3] S. Sharma and K. Guleria, "A systematic literature review on deep learning approaches for pneumonia detection using chest X-ray images," *Multimedia Tools and Applications*, vol. 83, no. 8, pp. 24101–24151, 2024.
- [4] K. Li et al., "Comprehensive use of a high-frequency electric knife, balloon dilatation, and cryotherapy for tuberculous central tracheobronchial cicatricial constriction," *BMC Surgery*, vol. 22, no. 1, p. 412, 2022.
- [5] G. Shingan, A. Jagtap, and S. Chaudhary, "Early fusion of texture-encoded CT with CXR images for pulmonary disease classification," *SN Computer Science*, vol. 7, no. 4, p. 315, 2026.
- [6] Y. Ge, S. Jiang, Q. Xu, C. Jiang, and F. Ye, "Exploiting representations from pre-trained convolutional neural networks for high-resolution remote sensing image retrieval," *Multimedia Tools and Applications*, vol. 77, no. 13, pp. 17489–17515, Jul. 2018.
- [7] M. Ali, M. Shahroz, U. Akram, M. F. Mushtaq, S. C. Altamiranda, S. A. Obregon, I. De La Torre Díez, and I. Ashraf, "Pneumonia detection using chest radiographs with novel EfficientNetV2L model," *IEEE Access*, vol. 12, pp. 34691–34707, 2024.
- [8] Y. P. Zhang, X. Y. Zhang, Y. T. Cheng, B. Li, X. Z. Teng, J. Zhang, S. Lam, T. Zhou, Z. R. Ma, J. B. Sheng, and V. C. Tam, "Artificial intelligence-driven radiomics study in cancer: The role of feature engineering and modeling," *Military Medical Research*, vol. 10, no. 1, p. 22, May 2023.



- [9] R. Parekh, J. Yang, and V. Honavar, "Constructive neural-network learning algorithms for pattern classification," *IEEE Transactions on Neural Networks*, vol. 11, no. 2, pp. 436–451, Mar. 2000.
- [10] A. Jafar and M. Lee, "High-speed hyperparameter optimization for deep ResNet models in image recognition," *Cluster Computing*, vol. 26, no. 5, pp. 2605–2613, 2023.
- [11] S. Babu, K. Thawait, A. P. Gopi, U. Srinivasarao, and K. J. Naik, "An EfficientNet-B0 framework for nitrogen stress identification in sustainable precision farming plants," *Cognitive Computation*, vol. 17, no. 5, p. 148, 2025.
- [12] A. Hosna, E. Merry, J. Gyalmo, Z. Alom, Z. Aung, and M. A. Azim, "Transfer learning: A friendly introduction," *Journal of Big Data*, vol. 9, no. 1, p. 102, Oct. 2022.
- [13] F. Alshanketi, A. Alharbi, M. Kuruvilla, V. Mahzoon, S. T. Siddiqui, N. Rana, and A. Tahir, "Pneumonia detection from chest X-ray images using deep learning and transfer learning for imbalanced datasets," *Journal of Imaging Informatics in Medicine*, vol. 38, no. 4, pp. 2021–2040, 2025.
- [14] M. El-Ghandour and M. I. Obayya, "Pneumonia detection in chest X-ray images using an optimized ensemble with XGBoost classifier," *Multimedia Tools and Applications*, vol. 84, no. 9, pp. 5491–5521, 2025.
- [15] P. Dhal and C. Azad, "A fine-tuning deep learning with multi-objective-based feature selection approach for the classification of text," *Neural Computing and Applications*, vol. 36, no. 7, pp. 3525–3553, 2024.
- [16] F. Z. Solak, "Comprehensive evaluation of preprocessing pipeline depth in deep learning-based brain tumor classification using CNN and vision transformer architectures," *Cluster Computing*, vol. 29, no. 3, p. 155, Jun. 2026.
- [17] S. S. Yadav and S. M. Jadhav, "Deep convolutional neural network based medical image classification for disease diagnosis," *Journal of Big Data*, vol. 6, no. 1, pp. 1–18, 2019.
- [18] N. Spolaor, H. D. Lee, A. I. Mendes, C. V. Nogueira, A. R. S. Parmezan, W. S. R. Takaki, and R. Fonseca-Pinto, "Fine-tuning pre-trained neural networks for medical image classification in small clinical datasets," *Multimedia Tools and Applications*, vol. 83, no. 9, pp. 27305–27329, 2024.
- [19] M. Rotkevich, C. Viana, M. V. Neguembor, and M. P. Cosma, "Deep learning in chromatin organization: From super-resolution microscopy to clinical applications," *Cellular and Molecular Life Sciences*, vol. 82, no. 1, p. 323, Aug. 2025.
- [20] A. Thakur, M. Gupta, D. K. Sinha, K. K. Mishra, V. K. Venkatesan, and S. Guluwadi, "Transformative breast cancer diagnosis using CNNs with optimized ReduceLROnPlateau and early stopping enhancements," *International Journal of Computational Intelligence Systems*, vol. 17, no. 1, p. 14, Dec. 2024.
- [21] K. El Asnaoui, "Design ensemble deep learning model for pneumonia disease classification," *International Journal of Multimedia Information Retrieval*, vol. 10, no. 1, pp. 55–68, Mar. 2021.
- [22] L. Puybasset, P. Cluzel, P. Gusman, P. Grenier, F. Preteux, J. J. Roubey, and CT Scan ARDS Study Group, "Regional distribution of gas and tissue in acute respiratory distress syndrome. I. Consequences for lung morphology," *Intensive Care Medicine*, vol. 26, no. 7, pp. 857–869, 2000.
- [23] H. Ugail, H. M. Alawar, A. A. Zehi, A. M. Alkendi, and I. L. Jaleel, "Evaluation of latent diffusion enhanced face recognition under forensic image degradations," *Discover Computing*, vol. 29, no. 1, p. 193, 2026.



- [24] K. Wang, P. Jiang, J. Meng, and X. Jiang, "Attention-based DenseNet for pneumonia classification," *IRBM*, vol. 43, no. 5, pp. 479–485, 2022.
- [25] S. Soundararajan, C. P. Shirley, B. Mallala, and K. Padmanaban, "Cloud-powered efficiency: A mobile application for agricultural pest identification using cycle-consistent generative adversarial networks," *Environment, Development and Sustainability*, pp. 1–28, 2025.
- [26] S. Sharma and K. Guleria, "A deep learning based model for the detection of pneumonia from chest X-ray images using VGG-16 and neural networks," *Procedia Computer Science*, vol. 218, pp. 357–366, 2023.
- [27] H. Bhatt and M. Shah, "A convolutional neural network ensemble model for pneumonia detection using chest X-ray images," *Healthcare Analytics*, vol. 3, p. 100176, 2023.