A Deep Learning Approach for Multiclass Pneumonia Detection in Chest X-Ray Images

Timothy Karani[1]* & Stephen Waithaka[2]

[1]*Department of Computing and Information Science, Kenyatta University, Kenya*

[2]*Department of Computing and Information Science, Kenyatta University, Kenya*

**Corresponding author: kartimothy@gmail.com*

https://doi.org/10.62049/jkncu.v6i2.601

Abstract

Pneumonia is a significant cause of mortality, particularly in children under five. Accurate detection of pneumonia from Chest X-ray (CXR) images is crucial in mitigating diagnostic errors common in manual radiographic analysis. This study leverages deep learning models to enhance the detection of multiclass pneumonia (normal, bacterial, and viral) using CXR images. We utilized a dataset comprising 5,863 multiclass pneumonia CXR samples. Data augmentation and regularization techniques were applied to address class imbalance and overfitting. Pre-trained models, including EfficientNet, MobileNet, RegNet, and ViT, were fine-tuned using the PyTorch framework, with transfer learning employed to optimize training. Model performance was assessed using accuracy, precision, recall, and specificity. The fine-tuned models achieved high classification accuracy, with EfficientNet and ConvNext models achieving accuracy scores of 83% and 82%, respectively. Data augmentation and regularization significantly improved the models' generalization, reducing overfitting and improving predictive accuracy. The proposed deep learning models provide an efficient and accurate tool for multiclass pneumonia detection from CXRs. These models have the potential to support healthcare professionals in making more accurate diagnoses.

Keywords: Multiclass Detection, Class Imbalance, Transfer Learning, Data Augmentation

Introduction

Pneumonia is a respiratory infection that predominantly affects the lungs, causing inflammation in the alveoli and bronchial tree. It is commonly categorized into two types: community-acquired pneumonia (CAP) and hospital-acquired pneumonia (HAP) (Torres et al., 2021). Globally, pneumonia remains one of the leading causes of death in children under the age of five, particularly in Sub-Saharan Africa. In 2015, of the 5.9 million deaths of children under five, pneumonia accounted for approximately 920,000 deaths, with countries such as Nigeria, the Democratic Republic of Congo, and Angola being the most affected. (Liu et al., 2016). The number of lower respiratory infections (formerly referred to as pneumonia) decreased in 2019 to 740,000 (Perin et al., 2022). Although progress has been made in reducing the overall number of pneumonia-related deaths, the disease continues to be a public health burden, especially in low-resource settings.

Traditionally, the diagnosis of pneumonia has relied on clinical evaluations and radiographic techniques such as Chest X-rays (CXRs). However, the manual interpretation of CXRs is susceptible to perceptual and interpretive errors, often resulting in inaccurate diagnoses. These errors are exacerbated by the high workload of radiologists and the limitations of human perception in detecting subtle variations in medical images. (Maskell, 2019). Recent advancements in deep learning (DL) and computer vision have opened new opportunities for automating the analysis of medical images, providing more accurate and timely diagnoses.

Despite the promising potential of DL models in medical imaging, several challenges persist. Class imbalance in medical datasets, overfitting due to large parameter counts in deep neural networks, and the high computational costs associated with training complex models are significant obstacles (Voulodimos et al., 2018). Addressing these challenges is essential to developing reliable and efficient models that can be deployed in clinical settings, especially in low-resource environments where computational power and quality datasets may be limited.

This study aims to develop a deep learning model for multiclass classification of pneumonia in CXRs. By employing techniques such as transfer learning, data augmentation, and regularization, the model seeks to improve the accuracy and generalizability of pneumonia detection, particularly in differentiating between normal, bacterial, and viral pneumonia cases. The research also investigates the performance of lightweight deep learning models that can be efficiently trained and deployed without sacrificing accuracy.

The study addresses the following research questions:

- What are the current deep learning models used for pneumonia detection in CXR images?
- How can a lightweight deep learning model be developed for multiclass pneumonia classification?
- What techniques can effectively reduce overfitting in deep learning models for pneumonia detection?

This research aims to take advantage of deep learning while helping to overcome some of the limitations of existing DLAs in the context of multilabel classification, making it a more robust and reliable tool for automated pneumonia detection.

Literature Review

Overview of Deep Learning in Medical Imaging

Deep learning (DL) has revolutionized medical imaging, providing advanced methods for automated analysis and classification of complex medical data. Among the most effective models in this field are Convolutional Neural Networks (CNNs), which have achieved significant success in tasks like detecting pneumonia, tumours, and other medical conditions. CNNs, by learning image features directly from raw pixel data, offer superior accuracy compared to traditional methods, particularly when analysing chest X-rays (CXRs) (Hatcher & Yu, 2018).

For pneumonia detection, CNNs have consistently demonstrated high accuracy in distinguishing between normal and infected lungs, addressing some of the limitations faced by radiologists. This capability is crucial in resource-constrained settings, where access to specialized healthcare professionals is limited. Recent studies, such as those involving CNN-based models like DenseNet and ResNet, highlight how these architectures enhance diagnostic efficiency, thereby minimizing human error in interpreting CXRs (Kermany et al., 2018).

Despite their success, DL models in medical imaging still face challenges, particularly in handling small datasets and imbalanced classes. Overfitting is a common issue, especially when training deep networks on limited data. To combat this, researchers often employ techniques like data augmentation and transfer learning to enhance model generalization. (Voulodimos et al., 2018). These approaches, when integrated with CNNs, reduce model complexity and improve performance across a range of medical imaging tasks.

Deep Learning Models for Pneumonia Detection

Several state-of-the-art models, including ResNet, DenseNet, and MobileNet, have been widely adopted in pneumonia detection tasks. These models are often pre-trained on large-scale datasets such as ImageNet, which allows them to capture a variety of image features before being fine-tuned on medical datasets. For example, Rahman et al. (2020) successfully fine-tuned DenseNet201 for the multiclass classification of normal, bacterial, and viral pneumonia, achieving an accuracy of 95% in the process. Similarly, other studies have used models like MobileNet due to its efficiency and relatively low computational cost, making it a suitable option for deployment in clinical settings (Reshan et al., 2023). Despite these successes, most research focuses on binary classification, which ignores the challenge of distinguishing between bacterial and viral pneumonia. This gap has motivated the current study to focus on multiclass classification for improved diagnostic accuracy.

Transfer Learning and Regularization in Medical Imaging

Transfer learning is a widely used technique in DL, especially in medical imaging, due to the scarcity of large, labelled datasets. This method allows models trained on general datasets, such as ImageNet, to be fine-tuned for specific medical tasks, saving time and computational resources. In pneumonia detection, transfer learning has shown significant promise. For example, Hashmi et al., (2020) Demonstrated the effectiveness of fine-tuning pre-trained models for pneumonia classification, achieving test accuracy as high as 98.43%. In this study, we similarly apply transfer learning to models like EfficientNet and ConvNeXt, which allows us to leverage their pre-learned features and adapt them for pneumonia classification.

In addition to transfer learning, regularization techniques such as weight decay have been explored to prevent overfitting, especially in deep models with numerous parameters. Weight decay penalizes large weights, promoting simpler models that generalize better to unseen data. This approach proved effective in our study, improving the performance of models like EfficientNet B0 from an accuracy of 80.83% to 82% by reducing overfitting during training.

Data Augmentation for Addressing Class Imbalance

Class imbalance, where certain classes (such as bacterial or viral pneumonia) are underrepresented, is a common issue in medical imaging datasets. If not addressed, class imbalance can bias models towards overrepresented classes (Li et al., 2019). Data augmentation is one of the most effective techniques for addressing this issue. By applying transformations such as rotation, scaling, and flipping, the training dataset can be artificially expanded, improving model generalization.

In this study, we applied several augmentation techniques to balance the dataset, which originally had a disproportionate number of normal and bacterial pneumonia cases compared to viral pneumonia. Augmentation helped mitigate overfitting and enhanced model accuracy. ConvNeXt Base, for example, achieved an accuracy of 83% after augmentation, compared to 70% before augmentation, demonstrating the importance of these techniques in handling class imbalance.

Multiclass Pneumonia Classification

While most existing research focuses on binary classification, this study aims to address the more complex task of multiclass pneumonia classification. Differentiating between bacterial and viral pneumonia is crucial because treatment approaches for these types of pneumonia differ significantly. Rahman et al., (2020) Explored multiclass classification using DenseNet201 and achieved a classification accuracy of 95%. However, DenseNet201's high computational requirements make it impractical for real-time use in resource-constrained environments. In this research, we experimented with more lightweight models such as EfficientNet and ConvNeXt, which require less computational power while still delivering high accuracy.

The focus on lightweight models like EfficientNet is motivated by the need for practical, real-time diagnostic tools in low-resource settings. EfficientNet B0, while smaller in size compared to DenseNet201, performed similarly in our multiclass classification task, achieving an accuracy of 82%. These results suggest that smaller, less resource-intensive models can be just as effective in pneumonia detection, offering a more accessible solution for real-time clinical use.

Related Works

Various studies have explored the use of DL models for pneumonia detection, with some focusing on ensembles of multiple models to improve accuracy. For example, Hashmi et al., (2020) Combined five state-of-the-art models in a weighted ensemble and achieved an accuracy of 98.43%. However, the use of multiple models significantly increases computational requirements, making these solutions less feasible for real-time deployment. In contrast, this study focuses on optimizing single models like ConvNeXt and EfficientNet for multiclass classification to ensure both high accuracy and computational efficiency.

In addition, Varshni et al., (2019) evaluated a DenseNet model for binary pneumonia classification and achieved an AUC score of 80.02%. However, their focus was limited to binary classification, and the model's complexity and computational cost hindered its scalability. This study addresses those limitations by focusing on lightweight, efficient models that can handle multiclass classification, which is more representative of clinical diagnostic challenges.

Gaps in Existing Research

Despite significant progress in using DL for pneumonia detection, several gaps remain. First, most studies have concentrated on binary classification, neglecting the need for models that can distinguish between bacterial and viral pneumonia. Additionally, the issue of class imbalance has been inadequately addressed in many studies, leading to models that perform well on the majority class but poorly on minority classes.

This study seeks to fill these gaps by focusing on multiclass classification and addressing class imbalance through data augmentation and regularization techniques. The use of lightweight models like EfficientNet B0 and ConvNeXt also addresses the need for efficient models that can be deployed in low-resource settings, providing a practical solution for real-time pneumonia detection in clinical environments.

Methodology

Research Design

This study employs a quantitative research design utilizing deep learning models for multiclass classification of pneumonia using chest X-ray (CXR) images. The quantitative approach is appropriate for this study as it allows for the development and evaluation of a machine learning model based on numerical data (pixel values of the images) and the measurement of model performance using established metrics such as accuracy, precision, recall, and specificity. The research focuses on improving model accuracy and efficiency through data augmentation, transfer learning, and regularization techniques.

Research Setting

The research is conducted in a controlled computational environment, with the dataset sourced from publicly available chest X-ray repositories. The dataset consists of images labelled into three categories: normal, bacterial pneumonia, and viral pneumonia. No personal or identifying information is included in the dataset, ensuring confidentiality and compliance with ethical standards. The dataset was pre-processed and augmented within a high-performance computing environment equipped with a GPU (Nvidia RTX A6000), enabling faster training of deep learning models.

Sampling

The study uses the Labelled Optical Coherence Tomography and Chest X-Ray dataset, which contains 5,863 labelled CXR images across three classes: normal, bacterial pneumonia, and viral pneumonia (Kermany et al., 2018). The dataset is randomly split into training, validation, and test sets, with 70% of the images used for training, 15% for validation, and 15% for testing. To address the issue of class imbalance—where the number of normal images exceeds the number of pneumonia cases—data augmentation techniques are applied to increase the number of pneumonia samples in the training set.

Instruments and Data Collection

The primary instruments used in this research are deep learning algorithms implemented using the PyTorch machine learning framework. Training models from scratch however, is expensive, so the use of pretrained models can help in this regard. (Burkov, 2020). The pre-trained models—EfficientNet, MobileNet, RegNet, and ViT—are fine-tuned for pneumonia detection using transfer learning. (Russakovsky et al., 2015). The dataset is augmented through techniques such as rotation, horizontal and vertical flips, and scaling, which increase the variability of the training data and help mitigate the problem of class imbalance. The models are trained on the augmented dataset and evaluated based on key performance metrics.

Data Preprocessing

Preprocessing steps involve resizing all images to a uniform size (224 x 224 pixels) to match the input requirements of the pre-trained models. Images are normalized to improve model convergence and enhance training stability. Data augmentation techniques are then applied to the training set to alleviate class imbalance. Horizontal and vertical flips, rotation (up to 45 degrees), and scaling are used to artificially increase the size of the minority class (bacterial and viral pneumonia).

Data Analysis

The performance of the deep learning models is evaluated using cross-entropy loss as the loss function and the Adam optimizer for optimization. The models are fine-tuned for 32 epochs, with early stopping employed to prevent overfitting. Key performance metrics are calculated to assess the models' ability to classify normal, bacterial, and viral pneumonia cases. The results are recorded for each model, and the performance will be evaluated using the metrics listed below:

$$Accuracy = \frac{TP+TN}{(TP+FN)+(FP+TN)} \quad (1)$$

$$Sensitivity = \frac{TP}{(TP+FP)} \quad (2)$$

$$Precision = \frac{TP}{(TP+FP)} \quad (3)$$

$$Specificity = \frac{TN}{(FP+TN)} \quad (4)$$

$$F1\,Score = \frac{(2*TP)}{(2*TP+FN+FP)} \quad (5)$$

Where TP, FP, TN, and FN are True and False Positives, and True and False Negatives, respectively.

Ethical Considerations

This study adheres to ethical standards by using a publicly available dataset that contains no personally identifiable information. All data used in this research is anonymized, ensuring compliance with privacy and data protection regulations. Additionally, no direct human or animal subjects were involved in this research, and ethical approval was not required. The study aims to contribute to the public good by advancing research on automated pneumonia detection in medical imaging, potentially improving healthcare outcomes.

Results

This section presents the key findings from the experiments conducted with different deep learning models on the multiclass pneumonia classification task. The models were evaluated based on their performance in classifying normal, bacterial, and viral pneumonia using accuracy, precision, recall, specificity, and F1 score as the primary evaluation metrics.

Dataset and Augmentation

The dataset used for this study consisted of 5,863 chest X-ray (CXR) images, which were divided into three classes: normal, bacterial pneumonia, and viral pneumonia. The data was split into training (70%), validation (15%), and testing (15%) sets. Data augmentation techniques such as horizontal and vertical flips, rotation (up to 45 degrees), and scaling were applied to mitigate class imbalance. The augmented dataset improved the generalization of the models and helped reduce overfitting.

Table 1: Distribution of Images in the Original and Augmented Datasets

Class	Original Dataset	Augmented Dataset
Normal	1,183	5,000
Bacterial Pneumonia	2,380	5,000
Viral Pneumonia	1,093	5,000

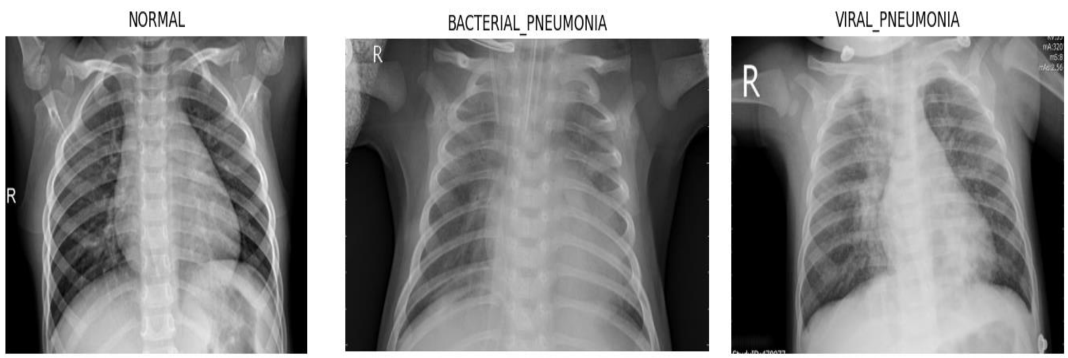

Figure 1: Samples from the base dataset

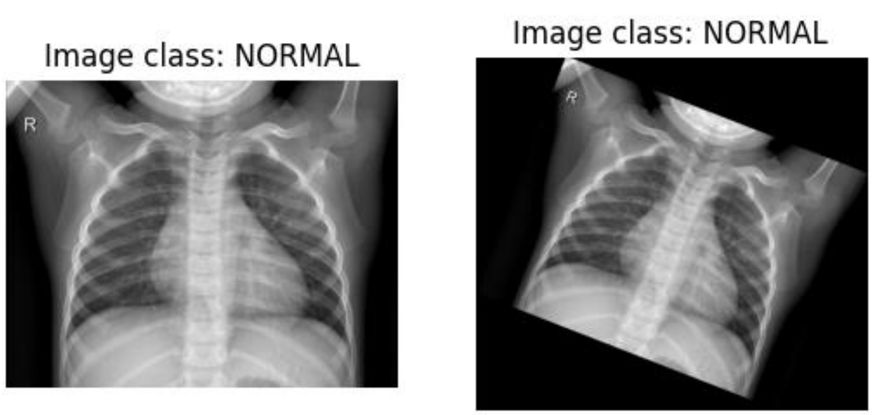

Figure 2: Sample image after transforming.

Model Performance

Eight pre-trained deep learning models—EfficientNet (B0, B2), MobileNet (v3 small, v3 large), RegNet (x1.6gf, y1.6gf), ConvNeXt Base, and ViT (B_16)—were fine-tuned using transfer learning on the augmented dataset. The results for each model are presented in Table 2, showing the accuracy, specificity, precision, recall, and F1 score for each model.

Table 2: Model Performance on the Test Set

Model	Accuracy (%)	Specificity (%)	Precision (%)	Recall (%)	F1 Score (%)
EfficientNet B0	82.0	91.0	80.5	83.0	81.7
EfficientNet B2	81.8	90.9	80.3	82.8	81.5
MobileNet v3 small	81.0	90.5	79.8	81.0	80.4
MobileNet v3 large	80.0	90.0	79.5	80.0	79.7
RegNet x1.6gf	79.4	89.7	78.6	79.4	79.0
RegNet y1.6gf	77.6	88.8	77.0	77.6	77.3
ConvNeXt Base	83.0	91.5	82.0	83.5	82.7
ViT B_16	80.1	90.1	79.2	80.1	79.6

Loss and Accuracy Curves

The training and validation loss and accuracy curves for the ConvNeXt Base model are depicted in Figure 3. The curves show that the model converged well, with no significant overfitting observed. Early stopping was employed during training to prevent overfitting, and the final model was selected based on its best performance on the validation set.

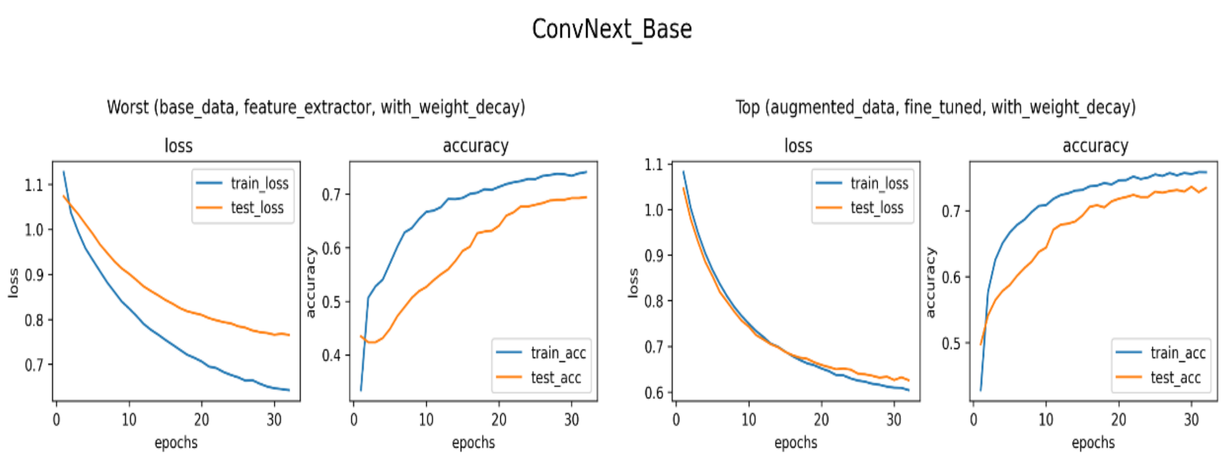

Figure 3: Training and Validation Curves for ConvNeXt Base Model

Comparison with State-of-the-Art Models

The performance of the ConvNeXt Base model was compared to state-of-the-art models from related studies. Table 3 presents a comparison of the best-performing model from this study with previous work.\

Table 3: Comparison with SOTA Models

Study	Model	Accuracy (%)	Specificity (%)
Kermany et al. (2018)	DenseNet169	95.7	93.0
Hashmi et al. (2020)	Weighted Ensemble	98.4	99.7
Ukwuoma et al. (2022)	Transformer	98.2	97.2
This Study	ConvNeXt Base	83.0	91.5

Although the ConvNeXt Base model did not achieve the same level of accuracy as some of the more complex models in previous studies, it provides a balance between accuracy and computational efficiency. This model is more suitable for deployment in low-resource settings where computational resources are limited.

Summary of Results

The ConvNeXt Base model demonstrated the highest accuracy (83%) and specificity (91.5%) among the tested models, proving to be a promising tool for multiclass pneumonia detection. The use of transfer learning, data augmentation, and regularization techniques contributed to the improved generalizability and performance of the model, particularly in addressing class imbalance and reducing overfitting. The results also highlight the challenge of distinguishing between bacterial and viral pneumonia, suggesting that further refinement of the model is needed for more precise classification.

Discussion

Transfer Learning

The results of this study demonstrate the effectiveness of transfer learning in multiclass pneumonia detection. By utilizing pre-trained models, the research was able to achieve high accuracy scores without needing to train models from scratch, which is beneficial in medical domains where large labelled datasets are scarce. Fine-tuning the pre-trained models consistently yielded better results than feature extraction alone. For instance, the ConvNeXt Base model achieved the highest accuracy (83%) and specificity (91.5%) after fine-tuning, outperforming the feature extractor configurations. These findings are in line with previous research that highlights the advantages of fine-tuning in improving model generalization, particularly when applied to small datasets. (Farahani et al., 2021).

Data Augmentation

Data augmentation proved crucial in addressing class imbalance and reducing overfitting in this study. As seen with models like ConvNeXt Base, applying augmentation techniques such as flipping and rotation improved performance metrics significantly. Without data augmentation, models such as MobileNet v3 Large experienced overfitting, as evidenced by lower test set accuracies and specificity scores. For ConvNeXt Base, data augmentation led to an increase in accuracy from 70.03% to 83% and specificity from 85.02% to 91.5%. This improvement aligns with the literature, where data augmentation has been shown to enhance the robustness of deep learning models in imbalanced datasets. (Gautam et al., 2022).

Regularization and Weight Decay

Regularization techniques, specifically weight decay, were also explored in this study. The application of weight decay reduced overfitting, particularly in models with higher parameter counts. For example, EfficientNet B0 demonstrated an accuracy improvement from 80.83% to 82% and specificity from 90.42% to 91% when weight decay was applied. Similarly, the ConvNeXt Base model exhibited a marked increase in performance with weight decay, showcasing the importance of regularization in optimizing deep learning models for medical image classification. These results support the findings of previous studies that emphasize the role of weight decay in controlling overfitting in neural networks (Ying, 2019).

Comparison with State-of-the-Art Models

The results achieved in this study, while promising, did not surpass those of some state-of-the-art (SOTA) models, such as DenseNet169, which achieved an accuracy of 95.72% (Hammoudi et al., 2021). Nevertheless, the ConvNeXt Base model outperformed several other models, such as VGG19 and ResNet variants, in terms of accuracy and specificity. Although ConvNeXt Base achieved the highest accuracy (83%), EfficientNet B0 was only slightly behind (82%) while being more computationally efficient. This finding supports the growing body of research suggesting that smaller models can be equally effective in medical image classification tasks, particularly in resource-constrained environments. The results of this study demonstrate that relatively smaller models like EfficientNet B0 can achieve comparable performance to larger models such as ConvNeXt Base, but with significantly reduced computational requirements.

Limitations

While this research achieved promising results, there are several limitations to consider. First, the dataset used in this study is relatively small compared to other large-scale datasets used in medical image analysis. Although data augmentation helped mitigate this issue, a larger and more diverse dataset would provide more robust results. Additionally, the study focused on pre-trained models and did not explore models trained from scratch, which could potentially lead to different outcomes. Finally, while weight decay improved performance in most models, further experimentation with other regularization techniques, such as dropout, could lead to further improvements.

Conclusion

This study explored the application of deep learning models for multiclass pneumonia classification in chest X-ray images. Using transfer learning, data augmentation, and regularization techniques such as weight decay, we aimed to enhance model performance while addressing issues like class imbalance and overfitting. The best-performing model, ConvNeXt Base, achieved an accuracy of 83% and a specificity of 91.5%, demonstrating that smaller, fine-tuned models can perform just as well as, or even better than, larger and more complex models.

Data augmentation proved to be a crucial technique for improving model generalization, particularly in mitigating the effects of overfitting. Fine-tuning the pre-trained models on the augmented dataset led to significant improvements in accuracy and specificity. Regularization via weight decay further helped in reducing overfitting, especially in fine-tuned models with larger sample sizes.

In conclusion, this research confirmed the effectiveness of using pre-trained models for multiclass pneumonia classification. Models like EfficientNet B0, despite being smaller, exhibited performance close to that of larger models like ConvNeXt Base, which suggests that model size does not necessarily correlate with performance. Pre-trained models provide a cost-effective way to achieve high accuracy in medical imaging tasks, making them highly suitable for resource-constrained environments.

Recommendations

By addressing the areas mentioned below, future studies can contribute to the ongoing efforts to improve deep learning models for medical image analysis and, more specifically, the detection and classification of pneumonia in chest X-rays.

- Future exploration of experiment tracking - particularly focusing on learning rate schedules and weight decay configurations. These parameters can significantly impact model performance, and tracking their effects over time would provide deeper insights into the optimal configuration for multiclass pneumonia detection.
- Data augmentation techniques - exploring a wider range of augmentation strategies, such as contrast adjustments, zooming, and adding synthetic noise, may help to create more robust models capable of generalizing better across different datasets.
- Lightweight models for deployment in low-resource settings - while EfficientNet and ConvNeXt Base showed strong performance in this study, further work could focus on developing lightweight models that maintain high accuracy while reducing computational overhead.
- Investigation of overfitting mitigation techniques - future studies could experiment with additional overfitting prevention techniques such as dropout, early stopping, and batch normalization. These methods could further enhance model generalizability, especially in datasets with class imbalance.

References

Burkov, A. (2020). *Machine learning engineering*. True Positive Inc.

Farahani, A., Pourshojae, B., Rasheed, K., & Arabnia, H. R. (2021). *A Concise Review of Transfer Learning* (arXiv:2104.02144). arXiv. http://arxiv.org/abs/2104.02144

Gautam, S., Höhne, M. M.-C., Hansen, S., Jenssen, R., & Kampffmeyer, M. (2022). *Demonstrating The Risk of Imbalanced Datasets in Chest X-ray Image-based Diagnostics by Prototypical Relevance Propagation* (arXiv:2201.03559). arXiv. http://arxiv.org/abs/2201.03559

Hammoudi, K., Benhabiles, H., Melkemi, M., Dornaika, F., Arganda-Carreras, I., Collard, D., & Scherpereel, A. (2021). Deep Learning on Chest X-ray Images to Detect and Evaluate Pneumonia Cases in the Era of COVID-19. *Journal of Medical Systems*, *45*(7), 75. https://doi.org/10.1007/s10916-021-01745-4

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

Hatcher, W. G., & Yu, W. (2018). A Survey of Deep Learning: Platforms, Applications and Emerging Research Trends. *IEEE Access*, *6*, 24411–24432. IEEE Access. https://doi.org/10.1109/ACCESS.2018.2830661

Kermany, D., Zhang, K., & Goldbaum, M. (2018). *Labeled Optical Coherence Tomography (OCT) and Chest X-Ray Images for Classification*. *2*. https://doi.org/10.17632/rscbjbr9sj.2

Liu, L., Oza, S., Hogan, D., Chu, Y., Perin, J., Zhu, J., Lawn, J. E., Cousens, S., Mathers, C., & Black, R. E. (2016). Global, regional, and national causes of under-5 mortality in 2000–15: An updated systematic analysis with implications for the Sustainable Development Goals. *The Lancet*, *388*(10063), 3027–3035. https://doi.org/10.1016/S0140-6736(16)31593-8

Maskell, G. (2019). Error in radiology—Where are we now? *The British Journal of Radiology*, *92*(1096), 20180845. https://doi.org/10.1259/bjr.20180845

Perin, J., Mulick, A., Yeung, D., Villavicencio, F., Lopez, G., Strong, K. L., Prieto-Merino, D., Cousens, S., Black, R. E., & Liu, L. (2022). Global, regional, and national causes of under-5 mortality in 2000–19: An updated systematic analysis with implications for the Sustainable Development Goals. *The Lancet Child & Adolescent Health*, *6*(2), 106–115. https://doi.org/10.1016/S2352-4642(21)00311-4

Rahman, T., Chowdhury, M. E. H., 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

Reshan, M. S. A., Gill, K. S., Anand, V., Gupta, S., Alshahrani, H., Sulaiman, A., & Shaikh, A. (2023). Detection of Pneumonia from Chest X-ray Images Utilizing MobileNet Model. *Healthcare*, *11*(11), Article 11. https://doi.org/10.3390/healthcare11111561

Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., Huang, Z., Karpathy, A., Khosla, A., Bernstein, M., Berg, A. C., & Fei-Fei, L. (2015). *ImageNet Large Scale Visual Recognition Challenge* (arXiv:1409.0575). arXiv. https://doi.org/10.48550/arXiv.1409.0575

Torres, A., Cilloniz, C., Niederman, M. S., Menéndez, R., Chalmers, J. D., Wunderink, R. G., & van der Poll, T. (2021). Pneumonia. *Nature Reviews Disease Primers*, *7*(1), 25. https://doi.org/10.1038/s41572-021-00259-0

Varshni, D., Thakral, K., Agarwal, L., Nijhawan, R., & Mittal, A. (2019). Pneumonia Detection Using CNN-based Feature Extraction. *2019 IEEE International Conference on Electrical, Computer and Communication Technologies (ICECCT)*, 1–7. 2019 IEEE International Conference on Electrical, Computer and Communication Technologies (ICECCT). https://doi.org/10.1109/ICECCT.2019.8869364

Voulodimos, A., Doulamis, N., Doulamis, A., & Protopapadakis, E. (2018). Deep Learning for Computer Vision: A Brief Review. *Computational Intelligence and Neuroscience*, *2018*, 1–13. https://doi.org/10.1155/2018/7068349

Ying, X. (2019). An Overview of Overfitting and Its Solutions. *Journal of Physics: Conference Series*, *1168*(2), 022022. https://doi.org/10.1088/1742-6596/1168/2/022022