algorithms

MDPI

Article

Pediatric Pneumonia Recognition Using an Improved DenseNet201 Model with Multi-Scale Convolutions and Mish Activation Function

Petra Radočaj [1], Dorijan Radočaj [2,*] and Goran Martinović [3]

[1] Layer d.o.o., Vukovarska cesta 31, 31000 Osijek, Croatia; radocajpetra@gmail.com
[2] Faculty of Agrobiotechnical Sciences Osijek, Josip Juraj Strossmayer University of Osijek, Vladimira Preloga 1, 31000 Osijek, Croatia
[3] Faculty of Electrical Engineering, Computer Science and Information Technology, Josip Juraj Strossmayer, University of Osijek, Kneza Trpimira 2B, 31000 Osijek, Croatia; goran.martinovic@ferit.hr
* Correspondence: dradocaj@fazos.hr; Tel.: +385-31-554-965

Abstract: Pediatric pneumonia remains a significant global health issue, particularly in low- and middle-income countries, where it contributes substantially to mortality in children under five. This study introduces a deep learning model for pediatric pneumonia diagnosis from chest X-rays that surpasses the performance of state-of-the-art methods reported in the recent literature. Using a DenseNet201 architecture with a Mish activation function and multi-scale convolutions, the model was trained on a dataset of 5856 chest X-ray images, achieving high performance: 0.9642 accuracy, 0.9580 precision, 0.9506 sensitivity, 0.9542 F1 score, and 0.9507 specificity. These results demonstrate a significant advancement in diagnostic precision and efficiency within this domain. By achieving the highest accuracy and F1 score compared to other recent work using the same dataset, our approach offers a tangible improvement for resource-constrained environments where access to specialists and sophisticated equipment is limited. While the need for high-quality datasets and adequate computational resources remains a general consideration for deep learning applications, our model's demonstrably superior performance establishes a new benchmark and offers the delivery of more timely and precise diagnoses, with the potential to significantly enhance patient outcomes.

Keywords: convolutional neural network; deep learning; pediatric pneumonia; Mish activation function; multi-scale convolution

check for updates

Academic Editor: Maryam Ravan

Received: 2 January 2025
Revised: 2 February 2025
Accepted: 7 February 2025
Published: 10 February 2025

Citation: Radočaj, P.; Radočaj, D.; Martinović, G. Pediatric Pneumonia Recognition Using an Improved DenseNet201 Model with Multi-Scale Convolutions and Mish Activation Function. *Algorithms* **2025**, *18*, 98. https://doi.org/10.3390/a18020098

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Pediatric pneumonia constitutes a significant global health challenge, with a disproportionate impact on low- and middle-income countries (LMICs) [1], where it persists as a primary cause of mortality in children under five years of age [2]. According to the World Health Organization (WHO), pneumonia is responsible for approximately 14% of all fatalities in children under five years old; this translates to an estimated 740,000 deaths each year [3]. Globally, the disease accounts for 10–20 million hospitalizations and an annual incidence of 150–156 million cases, with over 80% of pneumonia-related deaths occurring in LMICs. Regions such as South Asia and West and Central Africa bear the highest burden, with pneumonia affecting approximately 1 in 71 children annually [4,5]. This infectious disease, characterized by inflammation of the pulmonary alveoli in one or both lungs, presents clinically in children with a spectrum of signs and symptoms, including cough, fever, dyspnea, and chest pain [6]. In severe presentations, the disease can progress to

hypoxemia (decreased arterial oxygen saturation) and potentially life-threatening complications such as pleural effusion and respiratory failure [7]. The etiology of pediatric pneumonia is diverse, encompassing bacterial, viral, and fungal pathogens. While the clinical presentation may exhibit variability based on the specific causative agent, the disease typically manifests with respiratory distress and systemic symptoms, necessitating timely diagnosis and appropriate therapeutic intervention [8,9]. In high-income countries, the mortality rates are significantly lower due to better access to vaccines, diagnostic tools, and antibiotics. However, pneumonia continues to impose a significant burden on healthcare systems worldwide, accounting for millions of outpatient visits and hospital admissions annually, as well as substantial healthcare costs. The financial burden of prolonged care is particularly severe in resource-constrained areas, where limited access to affordable healthcare exacerbates the issue [10,11]. Addressing this challenge requires not only improvements in prevention and treatment but also innovative diagnostic methods to ensure early and accurate detection [12].

Deep learning has become a pivotal tool in the diagnosis of pediatric pneumonia, offering advanced capabilities for analyzing large datasets of medical images. These algorithms can identify subtle patterns and anomalies that may elude human interpretation, providing consistent and objective diagnostic outcomes [13]. By reducing diagnostic errors and operating with greater efficiency than traditional methods, deep learning addresses key limitations of conventional approaches. Specifically, conventional methods often rely on manual interpretation, which can be prone to human error and variability, particularly when identifying subtle or atypical patterns in medical images. Deep learning overcomes these challenges by leveraging automated feature extraction and pattern recognition, enabling consistent and reproducible results. Additionally, traditional diagnostic methods can struggle with large volumes of data or complex datasets, whereas deep learning excels at processing and analyzing such data efficiently [14]. Its utility is particularly pronounced in resource-constrained settings, where access to skilled radiologists and advanced diagnostic facilities is limited, thereby promoting equitable healthcare delivery and enabling timely, accurate diagnoses in underserved populations [15,16]. CNNs, a specialized architecture within deep learning, have demonstrated remarkable accuracy in medical imaging. They excel at extracting and analyzing both local and global features from radiographic images, which are essential for identifying pathologies such as pneumonia [17]. In pediatric applications, deep learning models trained on extensive radiographic datasets can reliably detect features indicative of pneumonia, including those associated with early-stage or atypical presentations. This capability not only enhances diagnostic efficiency; however, it also serves as a valuable supplementary tool for clinicians—minimizing the risk of missed diagnoses and contributing to improved patient outcomes [18]. Although these advancements are significant, they still require ongoing refinement and validation in clinical settings because the stakes in medical diagnostics are exceptionally high.

Despite the considerable advancements in deep learning applications within the realm of medical imaging, existing models exhibit numerous limitations. Many of the leading models—such as Inception, ResNet, and VGG—are computationally demanding, necessitating extensive memory and processing capabilities [19,20]. These prerequisites render them difficult to implement on edge devices or in settings with limited resources [19]. Furthermore, these models frequently possess a vast number of parameters, which heightens the risk of overfitting when trained on small or imbalanced datasets; this scenario is particularly prevalent in medical imaging [21]. Another significant challenge pertains to the interpretability of deep learning models. Although these algorithms can achieve high accuracy, their opaqueness complicates clinicians' ability to trust and effectively integrate them into their practice [22,23]. Addressing these limitations is crucial for facilitating the

widespread adoption of deep learning techniques in pediatric pneumonia diagnostics. This study intends to tackle the deficiencies of current deep learning models by developing an enhanced DenseNet201-based model for pediatric pneumonia recognition.

The primary aim of this study is to optimize the architecture of a deep-learning model for pediatric pneumonia diagnosis. This is accomplished by leveraging a pretrained DenseNet201 model and introducing multi-scale convolutional layers. These layers facilitate the model in capturing features at varying resolutions, thus enhancing its ability to identify pneumonia across diverse radiographic presentations. The model was evaluated using a publicly available dataset, which ensures transparency and reproducibility of the results. Furthermore, the Mish activation function was incorporated to provide smoother gradient propagation and improved convergence compared to traditional activation functions, such as ReLU. An essential aspect of the model's design is its parameter efficiency, achieved by minimizing the number of parameters in comparison to other pre-trained models. This design choice makes the model particularly suitable for deployment in resource-constrained environments where computational resources may be limited. However, by utilizing transfer learning and a publicly accessible dataset, we demonstrated the model's robustness and diagnostic accuracy across a variety of imaging scenarios. Although challenges remain, this approach significantly contributes to the field of pediatric pneumonia diagnosis.

The primary contributions of this study are as follows:

- This study leveraged a pretrained DenseNet201 model to develop a novel architecture optimized for pediatric pneumonia diagnosis. By introducing multi-scale convolutional layers, the model effectively captures features at varying resolutions, enhancing its diagnostic accuracy across diverse radiographic presentations.
- The incorporation of the Mish activation function provides smoother gradient propagation and improved convergence compared to traditional activation functions, such as ReLU, further enhancing the model's stability and generalization capabilities.
- An essential focus was placed on parameter efficiency. The model achieves a significantly lower parameter count compared to other pretrained architectures, ensuring computational efficiency and making it ideal for deployment in resource-constrained settings where advanced hardware is limited.
- By employing transfer learning and a publicly accessible dataset, the study demonstrates the adaptability of the model for diverse imaging scenarios, addressing the diagnostic challenges faced in both high-resource and low-resource clinical environments.

This paper is structured as follows: Section 2 reviews the relevant literature on pediatric pneumonia diagnosis using deep learning techniques. Section 3 provides a detailed description of the proposed model and methodologies. Section 4 presents the experimental results, including a comparative analysis of model performance. Finally, Section 5 outlines the conclusions and directions for future research.

2. Related Works

Although deep learning has made significant progress in advancing automated pneumonia diagnosis, achieving higher performance and ensuring the scalability and robustness of these models in real-world clinical settings remain ongoing challenges. Previous studies have demonstrated moderate success in refining deep learning techniques for detecting pneumonia from medical imaging data.

Ha Pham and Tran [24] evaluated the performance of an ensemble model combining three CNN architectures (InceptionResNetV2, DenseNet201, and VGG16) for binary pneumonia classification using 5856 chest X-ray images. Compared to single CNN models, the ensemble demonstrated superior performance, achieving an accuracy of nearly 0.95 and improving the average F1 score by 3%. Additionally, the study underscores the potential

of ensemble methods to outperform individual architectures in medical image analysis tasks. Kahwachi and Saed [25] emphasized the importance of rapid COVID-19 diagnosis due to its widespread social and economic impact and the limitations of molecular tests. Their study utilized six CNN architectures (VGG19, InceptionV3, ResNet50, InceptionResNetV2, and DenseNet121) to classify chest CT images as COVID-19 or non-COVID-19, comparing the effects of various activation functions. They found that InceptionV3 and DenseNet121 performed best when paired with non-ReLU activation functions, highlighting their potential to improve diagnostic accuracy in resource-limited areas. Lujan-Garcia et al. [26] highlighted pneumonia as a leading cause of mortality in children under five and emphasized the importance of chest X-ray imaging for its diagnosis. In research, they used transfer learning with the ImageNet pre-trained Xception Network to classify chest X-rays, distinguishing between 3883 pneumonia cases and 1349 normal images. The model demonstrated competitive performance compared to state-of-the-art methods, achieving a precision of 0.84, a recall of 0.99, an F1 score of 0.91, and an AUC of 0.97. Jain et al. [27] highlighted pneumonia as a leading cause of death among children, with 880,000 deaths reported in 2016, particularly in South Asia and Sub-Saharan Africa. Using CNNs, the researchers classified chest X-ray images to classify pneumonia. They experimented with different parameters, hyperparameters, and the number of convolutional layers within the CNN architecture. Two custom CNN models achieved validation accuracies of 0.8526 and 0.9231, while pre-trained models like VGG16, VGG19, ResNet50, and InceptionV3 attained accuracies of 0.8728, 0.8846, 0.7756, and 0.7099, respectively. Kaya [28] emphasized the importance of rapid pediatric pneumonia detection due to its seasonal association and potentially fatal outcomes. The study utilized deep tuning transfer learning of state-of-the-art CNN models, specifically DenseNet121, achieving an accuracy of 0.9503 and an F1 score of 0.9603 on a dataset of 5856 chest X-ray images. The results demonstrated the effectiveness of the approach for accurate and timely pneumonia detection in pediatric cases. Wang et al. [29] proposed a DenseNet-based method with a Squeeze and Excitation (SE) block and max-pooling to enhance pneumonia classification by better focusing on lesion regions. They selected PReLU as the activation function to prevent neuron death during training. The model outperformed DenseNet, achieving 0.928 accuracy, 0.926 precision, 0.962 recall, and 0.943 F1 score, with improvements in all metrics.

Previous studies on pediatric pneumonia recognition using CNNs face limitations, including inadequate feature extraction for detecting subtle patterns in chest X-rays. For example, while ensemble models combining architectures like DenseNet201, Inception-ResNetV2, and VGG16 demonstrated superior performance with improved F1 scores [24], their computational complexity makes them less practical for resource-constrained environments. Similarly, pre-trained CNNs such as VGG16, ResNet50, and InceptionV3 achieved validation accuracies between 0.7099 and 0.9231 in pediatric pneumonia recognition, but their reliance on extensive fine-tuning and higher resource requirements may limit their scalability in real-world applications [27]. Furthermore, despite promising results from DenseNet-based approaches incorporating SE blocks and custom activation functions [29], these methods often involve added architectural complexity, making deployment in clinical workflows challenging. Notably, many of these studies, including those by Kaya [28], Lujan-Garcia et al. [26], and Ha Pham and Tran [24], used the same publicly available Chest X-ray Images dataset as our study, highlighting dataset-specific limitations such as potential biases, variability in image quality, and limited demographic diversity. To tackle these deficiencies, this study introduces an innovative CNN model grounded in the DenseNet architecture. It integrates multi-scale convolution and the Mish activation function to improve feature detection, as well as gradient flow. The model achieves significant accuracy, sensitivity, and specificity; thus, it demonstrates a well-rounded performance in pneumonia

recognition, thereby avoiding misdiagnoses. This advancement effectively reduces false negatives, which can lead to overlooked diagnoses, and minimizes false positives, thus preventing unnecessary treatments. Although the model enhances scalability, robustness, and clinical applicability, it ultimately serves as a reliable tool to aid healthcare professionals in diagnosing pediatric pneumonia, particularly in low-resource settings.

3. Materials and Methods

We introduce a novel approach that leverages the DenseNet architecture, augmented with the Mish activation function and multi-scale convolutions (see Supplementary Materials). Through transfer learning, this method aims to enhance the recognition of pediatric pneumonia. The proposed methodology comprises three primary stages, as depicted in Figure 1: (1) preprocessing and preparation of input data, including the categorization of chest X-ray images into healthy and pneumonia classes; (2) implementation of the designed deep learning model, incorporating the DenseNet architecture with the Mish activation function and multi-scale convolutions; and (3) quantitative evaluation of the model's diagnostic performance in pneumonia detection, emphasizing metrics such as accuracy, precision, F1 score, sensitivity, and specificity.

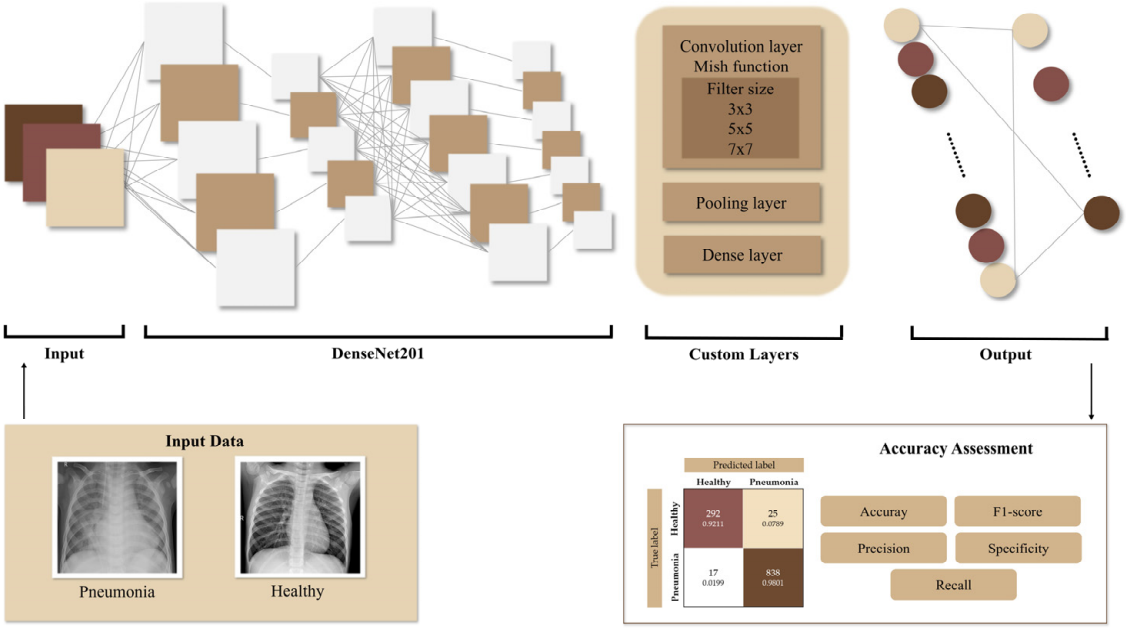


Figure 1. The study workflow for pneumonia recognition, utilizing a DenseNet architecture, Mish activation function, and multi-scale convolutions, proceeded through the following phases: (1) input data preparation, encompassing both healthy and pneumonia image classes; (2) implementation of the proposed deep learning model; and (3) evaluation of the implemented model's performance in pneumonia recognition, including accuracy assessment.

3.1. Data Preparation and Experimental Setup

We used the publicly available Chest X-ray Images (Pneumonia) dataset [30] for our experiments, which includes X-ray images of pediatric patients aged one to five years from Guangzhou Women and Children's Medical Center. The dataset contains 5856 images, categorized into two groups: pneumonia (4273 images) and healthy (1583 images) [30]. All chest radiographs were initially screened for quality control, and low-quality or unreadable scans were removed prior to analysis. Diagnoses for the images were graded by two expert physicians, with a third expert verifying the evaluation set to minimize grading errors and ensure the reliability of the dataset [30]. Samples of used images are shown in Figure 2. To ensure a balanced evaluation, we performed a stratified random split of the dataset into

training and validation subsets in an 80:20 ratio, preserving the class distribution. Due to computational resource constraints and the large dataset size, we opted for a single 80:20 split instead of k-fold cross-validation, which would have significantly increased training time without providing proportional benefits in this context. Although we did not apply explicit measures such as class weighting or oversampling to address the class imbalance, the CNN inherently adapts to the data distribution through its training process. Additionally, data augmentation techniques, including rotation, shifting, zooming, and horizontal flipping, were applied during preprocessing to increase dataset variability and improve model generalization. All images were resized to 224 × 224 pixels.

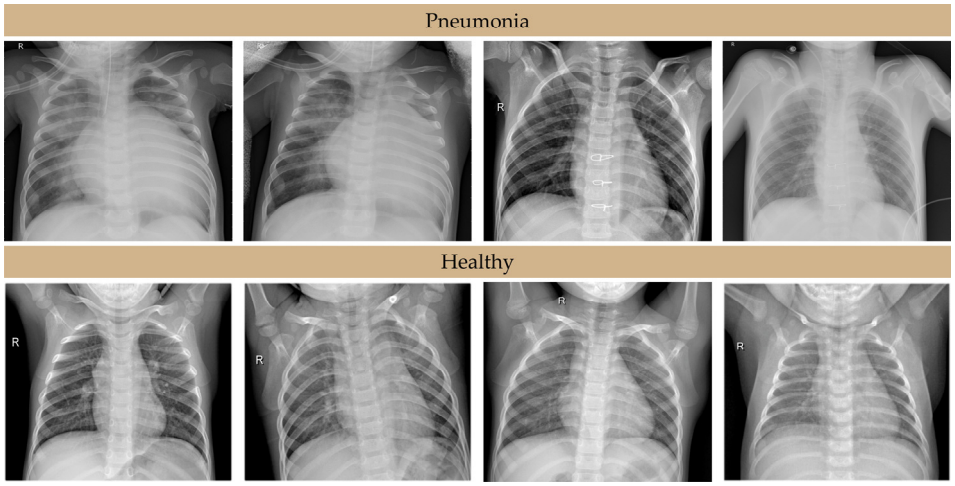


Figure 2. Samples of the pneumonia and healthy chest X-ray images used in this study.

We developed and trained the deep learning models using Python with the Keras-GPU [31] and TensorFlow-GPU [32] frameworks on the Google Colab platform. The platform's NVIDIA Tesla K80 GPUs with 12 GB of memory enabled efficient training. The models were trained over 20 epochs with a batch size of 32, utilizing the Adam optimization algorithm to update model parameters dynamically. The learning rate was automatically adjusted during training using a reduction mechanism that lowered it when validation performance plateaued, with a minimum of 0.5×10^{-6} to optimize convergence and prevent overfitting.

3.2. Model Architecture for Proposed Approach

Multi-scale convolution is a powerful technique in CNNs that enhances feature extraction by capturing information across varying spatial resolutions. Unlike standard convolutional layers that operate with fixed filter sizes, multi-scale convolution utilizes filters of different dimensions to simultaneously extract fine-grained and coarse-grained features [33]. This capability is particularly beneficial for medical image analysis, such as pediatric pneumonia recognition, where diverse patterns and textures must be identified across varying spatial scales. By combining outputs from filters of varying sizes, our model effectively captures both local and global features, making it more robust to variations in image resolution and scale [34,35]. This ability to capture multi-resolution features improves the model's generalizability, particularly in clinical datasets with heterogeneous imaging conditions [36,37]. Additionally, multi-scale convolution reduces the risk of overfitting by diversifying the spatial patterns used during feature extraction.

In this research, we utilized the pretrained DenseNet201 network as the foundational framework for feature extraction. DenseNet201, renowned for its intricately connected structure and effective feature reuse, was pretrained on ImageNet, thus providing a substantial basis for transfer learning. We chose to freeze the pretrained layers of DenseNet201

to preserve the learned features; however, we incorporated multi-scale convolutional layers on top. Specifically, we implemented filters of sizes 3×3, 5×5, and 7×7, which enabled the model to capture features across various spatial resolutions. These multi-scale outputs were then concatenated, forming a more comprehensive feature representation, which improved the model's capacity to detect pneumonia-related patterns in chest X-ray images. Additionally, we enhanced the architecture with pooling operations and a fully connected dense layer for classification because this allows for more accurate predictions. The convolutional operation in this architecture is expressed in Equation (1):

$$y(i,j) = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} x(i+m, j+n) * w(m,n), \quad (1)$$

where $x(i+m, j+n)$ represents the input feature map, $w(m,n)$ denotes the convolution filter of size $M \times N$, and $y(i,j)$ is the resulting output feature [38]. The multi-scale convolutional framework incorporates filters of different dimensions to capture diverse spatial patterns, thereby enhancing the model's ability to recognize complex and heterogeneous radiographic presentations, as demonstrated in Algorithm 1.

Algorithm 1: Multi-Scale Convolution

```
function MultiScaleConvolution (Input_Image, Kernels, Biases)
    Input:
        Input_Image: 2D matrix of pixel intensities
        Kernels: List of 2D filter matrices [Kernel_1, Kernel_2,..., Kernel_N]
        Biases: List of corresponding bias terms [Bias_1, Bias_2,..., Bias_N]
    Output: Combined_Feature_Map
    InitializeFeature_Maps as an empty list
    for each scale k in range 1, N do
        Kernel ← Kernels[k]
        Bias ← Biases[k]
        Feature_Map_k ← Convolution (Input_Image, Kernel, Bias)
        Append Feature_Map_k to Feature_Maps
    end for
    Combine all Feature_Map_k into a single representation:
    Combined_Feature_Map ← Concatenate (Feature_Maps)
    return Combined_Feature_Map
end function
```

To further improve the model's efficacy, the Mish activation function was utilized in the multi-scale convolutional layers. Mish, as mathematically defined in Equation (2):

$$Mish(x) = x * tanh(softplus(x)) = x * \tanh(\ln(1 + e^x)) \quad (2)$$

provides smoother gradient propagation and better convergence compared to traditional activation functions such as ReLU [39]. Unlike ReLU, which truncates negative values, Mish retains negative information through a smooth, non-monotonic curve, improving the model's capacity to learn subtle patterns and enhancing its generalization capabilities [39,40]. Supporting its adoption in this work, prior studies show Mish's superior performance compared to ReLU and Swish in CNN-based tasks, particularly within the field of medical imaging, including applications for pediatric pneumonia recognition [41]. Moreover, we conducted experiments comparing our model using the Mish activation function against the same model using ReLU. This comparative analysis allowed us to em-

pirically evaluate the impact of Mish on performance metrics and validate its effectiveness in our specific application.

Overall, the proposed approach is demonstrated in Algorithm 2.

Algorithm 2: Multi-Scale Convolution and Mish activation function with Pretrained DenseNet201

Step 1: Define Mish Activation Function
Define the Mish activation function:
    $Mish(x) = x * tanh(softplus(x))$, where:
    $softplus(x) = log(1 + exp(x))$,
    $tanh(x) = (exp(x) - exp(-x))/(exp(x) + exp(-x))$.
Step 2: Load Pretrained Model
Load DenseNet201 pretrained on ImageNet without the top layers.
Freeze the pretrained layers for feature extraction.
Step 3: Add Multi-Scale Convolutions
Extract features from DenseNet201 (*base_model.output*).
Apply multi-scale convolutional layers:
    conv_small with filter size 2×2, activation: *mish*.
    conv_medium with filter size 4×4, activation: *mish*.
    conv_large with filter size 8×8, activation: *mish*.
Concatenate multi-scale features.
Step 4: Global Pooling and Classification
Apply *GlobalMaxPooling2D* on concatenated features.
Add a Dense layer with sigmoid activation for binary classification.
Step 5: Compile and Train the Model
Compile the model using:
    Optimizer: *Adam*.
    Loss function: *binary_crossentropy*.
    Metrics: *Accuracy and AUC*.
Train the model with the training generator.

3.3. Pediatric Pneumonia Performance Evaluation

We assessed the performance of deep learning models for classifying pneumonia in chest X-ray images using precision, recall, specificity, F1 score, and accuracy, as defined in Equations (3)–(7). These metrics, calculated from the confusion matrix, provided a comprehensive evaluation of the models' diagnostic capabilities:

$$Precision = \frac{TP}{TP + FP}, \tag{3}$$

$$Recall = \frac{TP}{TP + FN}, \tag{4}$$

$$Specificity = \frac{TN}{TN + FP}, \tag{5}$$

$$F1\ score = 2 \times \frac{Precision \times Recall}{Precision + Recall}, \tag{6}$$

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \tag{7}$$

We defined true positives (TP) as cases where the model correctly identified pneumonia and true negatives (TN) as cases without pneumonia that were accurately classified. False

positives (FP) occurred when the model incorrectly classified non-pneumonia cases as pneumonia, and false negatives (FN) represented missed pneumonia cases. Specificity measured the model's ability to correctly identify non-pneumonia cases, complementing recall, which focused on detecting pneumonia cases. Chest X-ray images with pneumonia typically showed patterns such as increased lung density, consolidation, or pleural effusion, while diagnostic errors often resulted from overlapping features with other conditions or artifacts in the images.

4. Results and Discussion

The proposed model with the Mish activation function achieved notable performance in the diagnosis of pediatric pneumonia from radiological images, demonstrating an overall accuracy of 0.9642. A more comprehensive evaluation of accuracy assessment is presented in Table 1. This high accuracy indicates the model's effectiveness in correctly classifying both pneumonia-positive and healthy samples within the dataset, as reflected in the associated performance metrics. Furthermore, the table presented the results of the ablation study, comparing the Mish activation function against ReLU. This comparison provided the numerical evidence mentioned earlier, quantifying the impact of Mish on performance metrics and validating its effectiveness in this specific application. The F1 score was employed as a key evaluation metric due to its ability to provide a harmonic mean of precision and recall, particularly valuable in datasets characterized by class imbalance. During the research evaluation, it was concluded that accuracy is the most commonly used metric. We utilized it to enable an objective comparison with other studies. In addition to accuracy, the F1 score was also employed, although it was observed that many studies did not include this metric despite dealing with imbalanced datasets. To encourage future studies to adopt a fully objective approach to evaluating solutions, we emphasized the following metrics in our analysis: Accuracy, Precision, Recall (sensitivity), F1 score, and specificity, thereby ensuring a comprehensive and detailed evaluation. In this study, we employed a dataset consisting of 4273 chest X-ray images of pneumonia cases and 1583 images of healthy individuals. Despite the observed class imbalance, the model achieved a high F1 score of 0.9542, indicating a negligible impact of imbalance on the model's capacity for accurate classification. We extensively evaluated sensitivity and specificity to ensure accurate diagnoses, minimizing both missed cases and unnecessary treatments and ultimately improving patient outcomes. The model demonstrated a sensitivity of 0.9506, reflecting its ability to correctly identify pediatric patients with pneumonia while reducing the risk of missed or delayed diagnoses, which can have a serious impact on patient outcomes. Onwards, the model also demonstrated a high specificity of 0.9507, reflecting its ability to accurately classify healthy patients and to minimize false positives, thereby reducing unnecessary treatments such as inappropriate antibiotic use, which can lead to unwanted side effects and contribute to antibiotic resistance.

Table 1. Accuracy assessment of the proposed approach.

Metric	Model with Mish	Model with ReLU
Accuracy	**0.9642**	0.9616
Precision	**0.9580**	0.9534
Recall (Sensitivity)	**0.9506**	0.9489
F1 score	**0.9542**	0.9511
Specificity	**0.9507**	0.9489

These findings are corroborated by the confusion matrix, which is shown in Figure 3. The model correctly identified 838 pneumonia-positive cases (true positives) and 292 healthy cases (true negatives) while misclassifying 25 healthy cases (false positives) and 17 pneumonia cases (false negatives). These outcomes align with the reported performance metrics, underscoring the model's reliability and effectiveness in pediatric pneumonia diagnosis. Further analysis of these misclassified instances suggested that some errors arose from cases exhibiting borderline radiological features or overlapping radiological characteristics with other pulmonary conditions, such as bronchitis or atelectasis. The findings suggest that augmenting the training dataset and implementing region-specific feature enhancement techniques could lead to a reduction in such misclassifications.

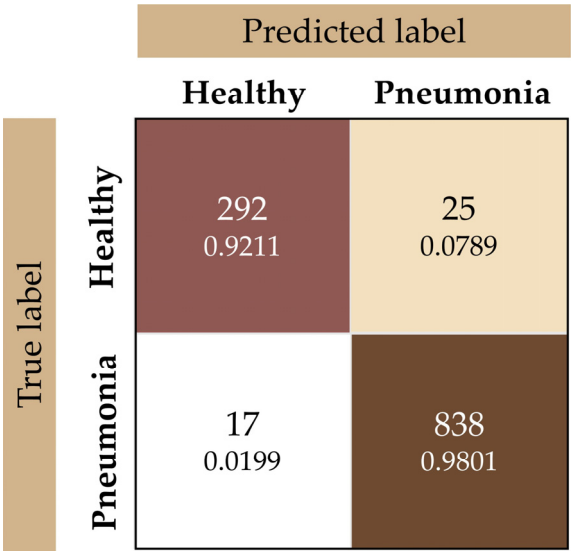


Figure 3. The confusion matrix for the proposed methodology. The classified values are represented by the upper numerical entry in each cell, with the corresponding class-specific percentage shown below.

To further validate the model, enhance interpretability, and facilitate clinical integration, Gradient-weighted Class Activation Mapping (Grad-CAM) was employed. Grad-CAM generates visual explanations by producing heatmaps that highlight the image regions most salient to the model's classification. In our architecture, this technique was applied to the last convolutional layer to ensure that the extracted features were both spatially and semantically rich. This technique is particularly valuable in the context of pediatric pneumonia diagnosis, as it offers transparency into the model's decision-making process, enabling verification of its diagnostic rationale and fostering clinical trust. In healthy cases, Grad-CAM typically generates diffuse activation patterns across the lung fields, consistent with the absence of focal pathology. Conversely, in cases of viral pneumonia, Grad-CAM heatmaps demonstrate more distributed and heterogeneous activation, often encompassing perihilar regions and reflecting the characteristic interstitial infiltrates associated with viral infection. In contrast, bacterial pneumonia exhibits more localized and intense activation patterns, concentrated over regions of consolidation, aligning with the typical lobar presentation of bacterial infection. The results are presented in Figure 4.

The proposed approach yielded an accuracy of 0.9642, placing it among the highest reported accuracies in this domain. A comparative evaluation is provided in Table 2. This result is slightly higher than the accuracy achieved by Vrbancic and Podgorelec [42], who reported an accuracy of 0.9626 using their SGDRE method on a similarly sized dataset. Other studies, such as Jaganathan et al. [43] using the LeNet-5 model on a significantly larger dataset of 84,484 images, reported an accuracy of 0.9600, which is comparable to ours despite the larger dataset. Notably, many methods using ensemble architectures, such as

Ha Pham and Tran [24], who reported an accuracy of 0.9503, and Mabrouk et al. [44], who reported an accuracy of 0.9391, achieved slightly lower accuracy values despite leveraging advanced combinations of models. This suggests that our single-model approach provides state-of-the-art performance without requiring the complexity of ensemble methods. In terms of the F1 score, our approach achieved 0.9640, outperforming most other methods that explicitly reported this metric. For example, Wang et al. [29], employing a custom CNN with a dataset of 5857 images, achieved an F1 score of 0.9430, while Mabrouk et al. [44] reported an F1 score of 0.9343 with their ensemble method on a dataset of the same size. Notably, Jaganathan et al. [43] and AlGhamdi [45] also achieved F1 scores of 0.9600 using LeNet-5 and MobileNetV3, respectively, with larger datasets. These comparisons highlight the efficiency of our approach, which performs competitively with methods that use significantly more training data. Some studies, such as Mardianto et al. [46], did not report F1 scores, focusing only on accuracy. This omission is concerning, as accuracy alone does not provide insight into the balance between false positives and false negatives, which is particularly important in medical diagnostics where class imbalances are common. By not reporting F1 scores, these studies risk underestimating the importance of false negative rates, which can have critical clinical implications. Reporting the F1 score, on the other hand, ensures a more comprehensive evaluation of a model's performance in such sensitive applications.

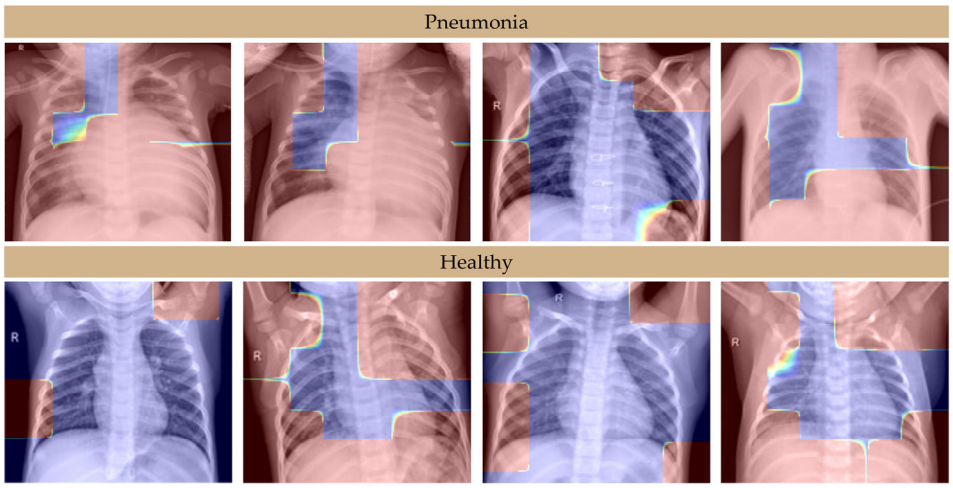


Figure 4. Grad-CAM heatmaps demonstrating model focus in pediatric pneumonia diagnosis.

Our approach, based on the improved DenseNet201 model with multi-scale convolutions and the Mish activation function, achieved superior metrics compared to both custom CNNs and hybrid combinations of CNNs with other methods. For instance, our accuracy of 0.9642 surpasses the performance of Wang et al. [29], who used a custom CNN which performed an accuracy of 0.9280 and F1 score of 0.9430, and Stephen et al. [22], who also employed a custom CNN but did not report an F1 score. Similarly, our results outperform Mardianto et al. [46], who combined a CNN with SVM and achieved an accuracy of 0.9200 without reporting an F1 score. Compared to ensemble methods, such as Ha Pham and Tran [24], our approach demonstrated a higher accuracy of 0.9642 compared to 0.9503 and similar F1 scores of 0.9640 compared to 0.9604. This suggests that our model is more streamlined while maintaining comparable or better performance than more complex architectures. Instead of developing custom CNNs, which demand extensive hyperparameter tuning and large training datasets for effective generalization, our approach leverages transfer learning with a pre-trained DenseNet201 model. The utilization of features pre-trained on the extensive ImageNet dataset facilitates efficient training and robust performance, even

with limited medical data, representing a substantial advantage in medical imaging where the acquisition of large, annotated datasets is often both challenging and costly. While custom CNNs trained on limited data are susceptible to overfitting and poor generalization, transfer learning with DenseNet201 provides a strong inductive bias, promoting better generalization and faster convergence. The incorporation of multi-scale convolutions and the Mish activation function further refine feature extraction from chest X-rays, surpassing standard transfer learning performance. This combined strategy of transfer learning with targeted architectural modifications offers a more efficient and effective alternative to building custom CNNs from scratch.

A clear progression in performance metrics can be observed over recent years as models have evolved to leverage larger datasets and more advanced architectures. For instance, early approaches, such as Manickam et al. [47] using ResNet50 in 2021, achieved an accuracy of 0.9306 and an F1 score of 0.9271 on 5232 images. Similarly, Mabrouk et al. [44] in 2022 achieved an accuracy of 0.9391 and an F1 score of 0.9343 using ensemble methods. However, more recent studies, such as Vrbancic and Podgorelec [42] in 2022, who reported an accuracy of 0.9626 accuracy and an F1 score of 0.974, and Kaya [28] in 2024, who reported an accuracy of 0.9503 accuracy and an F1 score of 0.9603, reflect a steady improvement in results, primarily due to advancements in model architectures and training strategies.

Table 2. Comparison and evaluation of the proposed approach and scientific studies in the last six years according to the data from the Web of Science Core Collection.

Reference	Year	CNN Model	No. of Images	Accuracy	F1 Score
Mardianto et al. [46]	2024	CNN+SVM	6140	0.9200	-
AlGhamdi [45]	2024	MobileNetV3	14,000>	-	0.9600
Jaganathan et al. [43]	2024	LeNet-5	84,484	0.9600	0.9600
Dzhaynakbaev et al. [48]	2024	VGG16	5228	-	-
Ha Pham and Tran [24]	2024	Ensemble (InceptionResNetV2, DenseNet201, VGG16)	5856	0.9503	0.9604
Mabrouk et al. [44]	2022	Ensemble (MobileNetV2, DenseNet169, Vision Transformer)	5856	0.9391	0.9343
Stephen et al. [22]	2019	Custom CNN	5856	0.9531	-
Lan et al. [49]	2024	DenseNet121	578	0.8100	-
Kaya [28]	2024	DenseNet121	5856	0.9503	0.9603
Manickam et al. [47]	2021	ResNet50	5232	0.9306	0.9271
Wang et al. [29]	2022	Custom CNN	5857	0.9280	0.9430
Vrbancic and Podgorelec [42]	2022	SGDRE method	5858	0.9626	0.9744
Yi et al. [50]	2023	Custom CNN	5856	0.9609	-
Our approach			**5856**	**0.9642**	**0.9640**

Hybrid models, such as the one proposed by Mardianto et al. [46], which integrate CNN with SVM, leverage the improved classification boundaries provided by SVM. However, their relatively lower accuracy of 0.9200 and omission of F1 scores raise concerns about their ability to manage class imbalances effectively. Similarly, ensemble methods such as those employed by Ha Pham and Tran [24] and Mabrouk et al. [44] leverage the diversity of models like InceptionResNetV2, DenseNet201, and Vision Transformers to

achieve competitive F1 scores of 0.9604 and 0.9343, respectively. Despite these advantages, the complexity and resource demands of ensemble architectures pose significant challenges in terms of interpretability, computational efficiency, and scalability. Custom CNNs, including those developed by Stephen et al. [22], Wang et al. [29], and Yi et al. [50], exhibit flexibility and adaptability to specific datasets, yielding accuracy values ranging from 0.9280 to 0.9609. Their performance highlights the potential of tailored architectures; however, the omission of F1 scores in some cases and slightly lower performance compared to pre-trained models suggest limited generalizability and challenges in capturing complex data patterns. Pre-trained architectures like DenseNet and ResNet demonstrate consistent performance due to their robust design. For example, Kaya [28] and Manickam et al. [47] reported accuracy values of 0.9503 and 0.9306, respectively, showcasing the efficacy of DenseNet121 and ResNet50 in feature propagation and handling gradient-related issues. However, DenseNet's performance can vary depending on dataset size and complexity, as evidenced by the low accuracy of 0.8100 achieved by Lan et al. [49] on a small dataset of 578 images. Among lightweight models, AlGhamdi [45] utilized MobileNetV3 to achieve an F1 score of 0.9600 on a larger dataset (>14,000 images), demonstrating its efficiency for resource-constrained environments. However, the absence of accuracy metrics limits a holistic evaluation of its performance. Similarly, foundational architectures like LeNet-5, used by Jaganathan et al. [43], achieved competitive accuracy and F1 scores of 0.9600 on a significantly larger dataset with 84484 images, though their simplicity may constrain their ability to capture complex features compared to modern architectures. Optimization-based approaches such as the SGDRE method by Vrbancic and Podgorelec [42] delivered an outstanding performance, achieving the highest F1 score of 0.9744 and an accuracy of 0.9626. While this underscores the potential of targeted optimization strategies, such methods might face limitations in adaptability across diverse datasets. The proposed approach, leveraging an improved DenseNet201 model enhanced with multi-scale convolutions and the Mish activation function, outperformed most compared methods, achieving the highest accuracy of 0.9642 and an F1 score of 0.9640. Its ability to deliver superior performance without requiring ensemble complexities highlights its efficiency. However, like other pre-trained architectures, its reliance on transfer learning may limit domain-specific interpretability compared to custom CNNs tailored exclusively for the task.

By utilizing a deep convolutional architecture, the model significantly reduces computational demands compared to ensemble-based methodologies, making it suitable for deployment in resource-constrained environments. This efficiency enables integration into point-of-care diagnostic tools, such as portable X-ray devices and mobile applications, thereby broadening access to accurate and timely diagnoses in low-resource and remote clinical settings. Additionally, the model has the potential to integrate seamlessly into telemedicine workflows, where portable X-ray systems combined with cloud-based analysis could provide rapid diagnostic support for patients in underserved areas. The model's resilience to class imbalance, in conjunction with training on publicly available datasets, significantly enhances its generalizability across varied patient populations. This model is also engineered to accommodate cloud-based deployment strategies, which yield several notable advantages. For instance, it offers scalability to manage substantial volumes of radiological data originating from multiple healthcare facilities. However, it also enhances accessibility, allowing clinicians in under-resourced areas to upload and analyze images in real time. Despite these strengths, limitations include the reliance on publicly available datasets, which may not fully capture the variability in image quality or geographic and demographic diversity. Plans for external validation using diverse real-world datasets are essential to further establish the model's reliability and applicability. Although it establishes a framework for continuous model refinement through iterative updates and retraining,

this ensures alignment with evolving diagnostic standards and expanding datasets. A possible solution that incorporates cloud-based infrastructure could facilitate centralized data processing, all while preserving the model's clinical relevance over time, making it an adaptable and sustainable solution for medical image analysis.

5. Conclusions and Future Work

In summary, the advanced deep learning model, which employs the DenseNet201 architecture, along with the Mish activation function and multi-scale convolutions, demonstrates impressive performance in diagnosing pediatric pneumonia from chest X-rays. Trained on a dataset comprising 5856 images, it achieved an overall accuracy of 96.42%, thereby securing its position among the top methodologies within this field. However, its performance not only exceeds that of numerous ensemble-based techniques, but it also retains a streamlined and computationally efficient design. This architectural efficiency is particularly advantageous for deployment in resource-constrained environments because ensemble methods can often be impractical due to their inherent complexity and computational requirements. The model's ability to provide timely and accurate diagnoses offers a scalable solution for addressing diagnostic gaps in low- and middle-income countries, where pneumonia remains a leading cause of childhood mortality, contributing to over 740,000 deaths annually. By reducing reliance on radiologists and advanced diagnostic infrastructure, the model has the potential to enhance healthcare delivery in under-resourced regions.

Despite its strengths, limitations include dependency on high-quality datasets, potential biases in training data, and challenges in generalizing across diverse populations. Furthermore, epidemiological and demographic distributions in LMICs may differ significantly from the public datasets used in this study. This limitation highlights the need for further validation using real-world and diverse demographic data to ensure broader applicability and reliability of the model. Future work could also explore multi-class tasks, such as distinguishing between bacterial and viral pneumonia, to enhance diagnostic precision and clinical relevance. To address these limitations, we aim to incorporate additional imaging modalities and integrate the model into clinical decision support systems. Scalability for other modalities, such as ultrasound, represents an important opportunity for expanding the model's utility across different healthcare contexts. Additionally, cloud-based deployment strategies could enhance scalability, enabling real-time analysis and continuous model refinement to adapt to evolving clinical needs. This approach represents a significant step forward in leveraging artificial intelligence for equitable and effective healthcare delivery.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/a18020098/s1, Table S1. Checklist for Artificial Intelligence in Medical Imaging (CLAIM): 2024 Update [51].

Author Contributions: Conceptualization, P.R.; methodology, P.R.; software, P.R.; validation, D.R. and G.M.; formal analysis, P.R.; investigation, P.R.; resources, P.R.; data curation, P.R.; writing—original draft preparation, P.R.; writing—review and editing, P.R., D.R. and G.M.; visualization, P.R.; supervision, G.M.; project administration, G.M.; funding acquisition, D.R. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The code developed in this study is available upon request from the corresponding author. The open access Chest X-ray Images (Pneumonia) repository containing images collected from pediatric patients, is divided into two categories (Pneumonia/Healthy) and is available

at https://www.kaggle.com/datasets/paultimothymooney/chest-xray-pneumonia (accessed on 12 December 2024).

Conflicts of Interest: Author Petra Radočaj was employed by the company "Layer d.o.o.". The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. Marangu, D.; Zar, H.J. Childhood Pneumonia in Low-and-Middle-Income Countries: An Update. *Paediatr. Respir. Rev.* **2019**, *32*, 3–9. [CrossRef] [PubMed]
2. Selvi, M.; Vaithilingan, S. Childhood Pneumonia in Low- and Middle-Income Countries: A Systematic Review of Prevalence, Risk Factors, and Healthcare-Seeking Behaviors. *Cureus* **2024**, *16*, e57636. [CrossRef]
3. Pneumonia in Children. Available online: https://www.who.int/news-room/fact-sheets/detail/pneumonia (accessed on 22 December 2024).
4. Rudan, I. Epidemiology and Etiology of Childhood Pneumonia. *Bull. World Health Organ.* **2008**, *86*, 408–416. [CrossRef] [PubMed]
5. Genie, Y.D.; Sayih, A.; Dessalegn, N.; Adugnaw, E.; Hiwot, A.Y.; Tesfa, T.B.; Kindie, K.; Gutema, L.; Ayalew, E.; Kebede, B.F. Time to Recovery from Severe Pneumonia and Its Predictors among Pediatric Patients Admitted in South West Region Governmental Hospitals, South West Ethiopia: Prospective Follow-up Study. *Glob. Pediatr.* **2024**, *9*, 100227. [CrossRef]
6. Kevat, P.M.; Morpeth, M.; Graham, H.; Gray, A.Z. A Systematic Review of the Clinical Features of Pneumonia in Children Aged 5-9 Years: Implications for Guidelines and Research. *J. Glob. Health* **2022**, *12*, 10002. [CrossRef] [PubMed]
7. Salah, E.T.; Algasim, S.H.; Mhamoud, A.S.; Husian, N.E.O.S.A. Prevalence of Hypoxemia in Under-Five Children with Pneumonia in an Emergency Pediatrics Hospital in Sudan. *Indian J. Crit. Care Med. Peer-Rev. Off. Publ. Indian Soc. Crit. Care Med.* **2015**, *19*, 203–207. [CrossRef] [PubMed]
8. Grief, S.N.; Loza, J.K. Guidelines for the Evaluation and Treatment of Pneumonia. *Prim. Care Clin. Off. Pract.* **2018**, *45*, 485–503. [CrossRef] [PubMed]
9. Zhang, J.; Zhu, Y.; Zhou, Y.; Gao, F.; Qiu, X.; Li, J.; Yuan, H.; Jin, W.; Lin, W. Pediatric Adenovirus Pneumonia: Clinical Practice and Current Treatment. *Front. Med.* **2023**, *10*, 1207568. [CrossRef] [PubMed]
10. Haggie, S.; Selvadurai, H.; Gunasekera, H.; Fitzgerald, D.A. Paediatric Pneumonia in High-Income Countries: Defining and Recognising Cases at Increased Risk of Severe Disease. *Paediatr. Respir. Rev.* **2021**, *39*, 71–81. [CrossRef]
11. Madhi, S.A.; De Wals, P.; Grijalva, C.G.; Grimwood, K.; Grossman, R.; Ishiwada, N.; Lee, P.-I.; Nascimento-Carvalho, C.; Nohynek, H.; O'Brien, K.L.; et al. The Burden of Childhood Pneumonia in the Developed World: A Review of the Literature. *Pediatr. Infect. Dis. J.* **2013**, *32*, e119–e127. [CrossRef]
12. Kanwal, K.; Asif, M.; Khalid, S.G.; Liu, H.; Qurashi, A.G.; Abdullah, S. Current Diagnostic Techniques for Pneumonia: A Scoping Review. *Sensors* **2024**, *24*, 4291. [CrossRef] [PubMed]
13. Li, Y.; Zhang, Z.; Dai, C.; Dong, Q.; Badrigilan, S. Accuracy of Deep Learning for Automated Detection of Pneumonia Using Chest X-Ray Images: A Systematic Review and Meta-Analysis. *Comput. Biol. Med.* **2020**, *123*, 103898. [CrossRef] [PubMed]
14. Hameed, B.M.Z.; Prerepa, G.; Patil, V.; Shekhar, P.; Zahid Raza, S.; Karimi, H.; Paul, R.; Naik, N.; Modi, S.; Vigneswaran, G.; et al. Engineering and Clinical Use of Artificial Intelligence (AI) with Machine Learning and Data Science Advancements: Radiology Leading the Way for Future. *Ther. Adv. Urol.* **2021**, *13*, 17562872211044880. [CrossRef] [PubMed]
15. Liu, X.; Faes, L.; Kale, A.U.; Wagner, S.K.; Fu, D.J.; Bruynseels, A.; Mahendiran, T.; Moraes, G.; Shamdas, M.; Kern, C.; et al. A Comparison of Deep Learning Performance against Health-Care Professionals in Detecting Diseases from Medical Imaging: A Systematic Review and Meta-Analysis. *Lancet Digit. Health* **2019**, *1*, e271–e297. [CrossRef]
16. Guo, Y.; Huang, C.; Sheng, Y.; Zhang, W.; Ye, X.; Lian, H.; Xu, J.; Chen, Y. Improve the Efficiency and Accuracy of Ophthalmologists' Clinical Decision-Making Based on AI Technology. *BMC Med. Inform. Decis. Mak.* **2024**, *24*, 192. [CrossRef] [PubMed]
17. Varshni, D.; Thakral, K.; Agarwal, L.; Nijhawan, R.; Mittal, A. Pneumonia Detection Using CNN Based Feature Extraction. In Proceedings of the 2019 IEEE International Conference on Electrical, Computer and Communication Technologies (ICECCT), Coimbatore, India, 20–22 February 2019; pp. 1–7.
18. Rajaraman, S.; Thoma, G.; Antani, S.; Candemir, S. Visualizing and Explaining Deep Learning Predictions for Pneumonia Detection in Pediatric Chest Radiographs. In *Medical Imaging 2019: Computer-Aided Diagnosis*; Hahn, H.K., Mori, K., Eds.; SPIE: San Diego, CA, USA, 2019; p. 27.
19. Li, J.; Wang, W. Deployment and Application of Deep Learning Models under Computational Constraints. In Proceedings of the 2023 IEEE International Conference on Big Data (BigData), Sorrento, Italy, 15–18 December 2023; pp. 2529–2533.
20. Hung, C.-L.; Hsin, C.; Wang, H.-H.; Tang, C.Y. Optimization of GPU Memory Usage for Training Deep Neural Networks. In *Pervasive Systems, Algorithms and Networks*; Esposito, C., Hong, J., Choo, K.-K.R., Eds.; Springer International Publishing: Cham, Switzerland, 2019; pp. 289–293.

21. Iqbal, S.; Qureshi, A.N.; Li, J.; Choudhry, I.A.; Mahmood, T. Dynamic Learning for Imbalanced Data in Learning Chest X-Ray and CT Images. *Heliyon* **2023**, *9*, e16807. [CrossRef] [PubMed]
22. Stephen, O.; Sain, M.; Maduh, U.J.; Jeong, D.-U. An Efficient Deep Learning Approach to Pneumonia Classification in Healthcare. *J. Healthc. Eng.* **2019**, *2019*, 4180949. [CrossRef]
23. Liz, H.; Sánchez-Montañés, M.; Tagarro, A.; Domínguez-Rodríguez, S.; Dagan, R.; Camacho, D. Ensembles of Convolutional Neural Network Models for Pediatric Pneumonia Diagnosis. *Future Gener. Comput. Syst.* **2021**, *122*, 220–233. [CrossRef]
24. Ha Pham, N.; Tran, G.S. Apply a CNN-Based Ensemble Model to Chest-X Ray Image-Based Pneumonia Classification. *J. Adv. Inf. Technol.* **2024**, *15*, 1205–1214. [CrossRef]
25. Kahwachi, W.T.; Saed, K.A. A Novel Architecture of Convolutional Neural Network to Diagnose COVID-19 Disease. *Math. Stat. Eng. Appl.* **2022**, *71*, 4831–4856. Available online: https://www.philstat.org/index.php/MSEA/article/view/1080 (accessed on 6 February 2025).
26. Luján-García, J.E.; Yáñez-Márquez, C.; Villuendas-Rey, Y.; Camacho-Nieto, O. A Transfer Learning Method for Pneumonia Classification and Visualization. *Appl. Sci.* **2020**, *10*, 2908. [CrossRef]
27. Jain, R.; Nagrath, P.; Kataria, G.; Sirish Kaushik, V.; Jude Hemanth, D. Pneumonia Detection in Chest X-Ray Images Using Convolutional Neural Networks and Transfer Learning. *Measurement* **2020**, *165*, 108046. [CrossRef]
28. Kaya, M. Feature Fusion-Based Ensemble CNN Learning Optimization for Automated Detection of Pediatric Pneumonia. *Biomed. Signal Process. Control* **2024**, *87*, 105472. [CrossRef]
29. Wang, K.; Jiang, P.; Meng, J.; Jiang, X. Attention-Based DenseNet for Pneumonia Classification. *IRBM* **2022**, *43*, 479–485. [CrossRef]
30. Kermany, D.S.; Goldbaum, M.; Cai, W.; Valentim, C.C.S.; Liang, H.; Baxter, S.L.; McKeown, A.; Yang, G.; Wu, X.; Yan, F.; et al. Identifying Medical Diagnoses and Treatable Diseases by Image-Based Deep Learning. *Cell* **2018**, *172*, 1122–1131.e9. [CrossRef] [PubMed]
31. Keras 3 API Documentation. Available online: https://keras.io/api/ (accessed on 24 September 2024).
32. Module: Tf | TensorFlow v2.16.1. Available online: https://www.tensorflow.org/api_docs/python/tf (accessed on 4 May 2024).
33. Sun, Y.; Zhang, Y.; Liu, S.; Lu, W.; Li, X. Image Super-Resolution Using Supervised Multi-Scale Feature Extraction Network. *Multimed. Tools Appl.* **2021**, *80*, 1995–2008. [CrossRef]
34. Shibu, D.S.; Priyadharsini, S.S. Multi Scale Decomposition Based Medical Image Fusion Using Convolutional Neural Network and Sparse Representation. *Biomed. Signal Process. Control* **2021**, *69*, 102789. [CrossRef]
35. Liu, J.; Zhang, L.; Guo, A.; Gao, Y.; Zheng, Y. Multi-Scale Feature Fusion Convolutional Neural Network for Multi-Modal Medical Image Fusion. In Proceedings of the 2023 4th International Conference on Computing, Networks and Internet of Things, Xiamen, China, 26–28 May 2023; Association for Computing Machinery: New York, NY, USA, 2023; pp. 913–917.
36. Haque, M.I.U.; Dubey, A.K.; Danciu, I.; Justice, A.C.; Ovchinnikova, O.S.; Hinkle, J.D. Effect of Image Resolution on Automated Classification of Chest X-Rays. *J. Med. Imaging Bellingham Wash* **2023**, *10*, 044503. [CrossRef]
37. Baltruschat, I.M.; Nickisch, H.; Grass, M.; Knopp, T.; Saalbach, A. Comparison of Deep Learning Approaches for Multi-Label Chest X-Ray Classification. *Sci. Rep.* **2019**, *9*, 6381. [CrossRef]
38. Traore, B.B.; Kamsu-Foguem, B.; Tangara, F. Deep Convolution Neural Network for Image Recognition. *Ecol. Inform.* **2018**, *48*, 257–268. [CrossRef]
39. Misra, D. Mish: A Self Regularized Non-Monotonic Activation Function. *arXiv* **2020**, arXiv:1908.08681.
40. Radočaj, P.; Radočaj, D.; Martinović, G. Image-Based Leaf Disease Recognition Using Transfer Deep Learning with a Novel Versatile Optimization Module. *Big Data Cogn. Comput.* **2024**, *8*, 52. [CrossRef]
41. Radočaj, P.; Radočaj, D.; Martinović, G. Optimizing Convolutional Neural Network Architectures with Optimal Activation Functions for Pediatric Pneumonia Diagnosis Using Chest X-Rays. *Big Data Cogn. Comput.* **2025**, *9*, 25. [CrossRef]
42. Vrbančič, G.; Podgorelec, V. Efficient Ensemble for Image-Based Identification of Pneumonia Utilizing Deep CNN and SGD with Warm Restarts. *Expert Syst. Appl.* **2022**, *187*, 115834. [CrossRef]
43. Jaganathan, D.; Balsubramaniam, S.; Sureshkumar, V.; Dhanasekaran, S. Concatenated Modified LeNet Approach for Classifying Pneumonia Images. *J. Pers. Med.* **2024**, *14*, 328. [CrossRef] [PubMed]
44. Mabrouk, A.; Díaz Redondo, R.P.; Dahou, A.; Abd Elaziz, M.; Kayed, M. Pneumonia Detection on Chest X-Ray Images Using Ensemble of Deep Convolutional Neural Networks. *Appl. Sci.* **2022**, *12*, 6448. [CrossRef]
45. AlGhamdi, A.S. Efficient Deep Learning Approach for the Classification of Pneumonia in Infants from Chest X-Ray Images. *Trait. Signal* **2024**, *41*, 1245–1262. [CrossRef]
46. Mardianto, M.; Yoani, A.; Soewignjo, S.; Putra, I.; Dewi, D.A. Classification of Pneumonia from Chest X-Ray Images Using Support Vector Machine and Convolutional Neural Network. *Int. J. Adv. Comput. Sci. Appl.* **2024**, *15*, 1015. [CrossRef]
47. Manickam, A.; Jiang, J.; Zhou, Y.; Sagar, A.; Soundrapandiyan, R.; Samuel, R. Automated Pneumonia Detection on Chest X-Ray Images: A Deep Learning Approach with Different Optimizers and Transfer Learning Architectures. *Measurement* **2021**, *184*, 109953. [CrossRef]

48. Dzhaynakbaev, N.; Kurmanbekkyzy, N.; Baimakhanova, A.; Mussatayeva, I. 2D-CNN Architecture for Accurate Classification of COVID-19 Related Pneumonia on X-Ray Images. *Int. J. Adv. Comput. Sci. Appl. IJACSA* **2024**, *15*, 905. [CrossRef]
49. Lan, X.; Zhang, Y.; Yuan, W.; Shi, F.; Guo, W. Image-Based Deep Learning in Diagnosing Mycoplasma Pneumonia on Pediatric Chest X-Rays. *BMC Pediatr.* **2024**, *24*, 720. [CrossRef] [PubMed]
50. Yi, R.; Tang, L.; Tian, Y.; Liu, J.; Wu, Z. Identification and Classification of Pneumonia Disease Using a Deep Learning-Based Intelligent Computational Framework. *Neural Comput. Appl.* **2023**, *35*, 14473–14486. [CrossRef]
51. Tejani, A.S.; Klontzas, M.E.; Gatti, A.A.; Mongan, J.T.; Moy, L.; Park, S.H.; Kahn, C.E., Jr. Checklist for Artificial Intelligence in Medical Imaging (CLAIM): 2024 Update. *Radiol. Artif. Intell.* **2024**, *6*, e240300. [CrossRef] [PubMed]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.