



AI-Driven Pneumonia Diagnosis Using Deep Learning: A Comparative Analysis of CNN Models on Chest X-Ray Images

Abdullah Al Foysal^{1*}, Sumaiya Sultana²

¹Department of Computer Engineering (AI), University of Genova, Genova, Italy

²Department of Engineering Technology for Strategy and Security, University of Genova, Genova, Italy

Email: *niloyhasanfoysal440@gmail.com, sultanasumaiya38@gmail.com

How to cite this paper: Al Foysal, A. and Sultana, S. (2025) AI-Driven Pneumonia Diagnosis Using Deep Learning: A Comparative Analysis of CNN Models on Chest X-Ray Images. *Open Access Library Journal*, 12: e12899.

<https://doi.org/10.4236/oalib.1112899>

Received: January 1, 2025

Accepted: February 14, 2025

Published: February 17, 2025

Copyright © 2025 by author(s) and Open Access Library Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

Pneumonia remains a significant cause of morbidity and mortality worldwide, particularly in vulnerable populations such as children and the elderly. Early detection through chest X-ray analysis plays a crucial role in timely treatment; however, reliance on radiologists can lead to variability, delays, and diagnostic errors. This paper presents a convolutional neural network (CNN) designed to automate pneumonia diagnosis from chest X-ray images, addressing the need for faster and more consistent diagnostic solutions. The model was trained and validated on a publicly available dataset containing chest X-ray images labeled as either pneumonia or normal. Data augmentation techniques, such as rotation, scaling, and flipping, were applied to enhance generalization and mitigate class imbalance. The architecture consists of multiple convolutional layers, batch normalization, max pooling, and dropout layers to extract and classify image features effectively. The CNN achieved an accuracy of 90.22% and an AUC score of 0.96 on the test set. Precision, recall, and F1-score metrics demonstrate the model's robust performance, with a recall of 96% for pneumonia cases, indicating a low rate of false negatives. The receiver operating characteristic (ROC) curve and confusion matrix further validate the model's efficacy in distinguishing pneumonia from normal cases. This study highlights the potential of AI-driven diagnostic tools to improve pneumonia detection, particularly in resource-limited settings. The proposed model can assist radiologists by providing rapid and reliable interpretations, ultimately enhancing patient outcomes and reducing the burden on healthcare systems. Future work will focus on expanding datasets and refining the architecture for even greater accuracy and stability.

Subject Areas

Deep Learning

Keywords

Pneumonia Diagnosis, Deep Learning, Convolutional Neural Networks (CNNs), Medical Imaging, Automated Diagnostics, Artificial Intelligence (AI)

1. Introduction

Pneumonia remains one of the most common and serious respiratory infections, responsible for over 2.5 million deaths annually, including a significant proportion of children under five and elderly individuals [1]-[7]. It is a critical public health concern, particularly in low-resource environments where healthcare infrastructure is limited and access to trained radiologists is scarce. Early detection and treatment are essential to prevent severe complications, but pneumonia diagnosis relies heavily on chest X-rays, which require careful interpretation by experienced radiologists [8]-[10]. This dependency can lead to diagnostic delays, misinterpretations, and increased mortality rates, underscoring the urgent need for automated diagnostic tools. Artificial intelligence (AI) and deep learning have emerged as transformative technologies in the medical imaging domain, offering the potential to address this diagnostic gap [11] [12]. Convolutional neural networks (CNNs), a subset of deep learning models, have demonstrated remarkable success in automating image classification, segmentation, and feature extraction across various medical fields [13] [14]. Their ability to analyze complex patterns in imaging data has positioned CNNs as valuable tools for diagnosing diseases such as pneumonia, lung cancer, and diabetic retinopathy. By leveraging large datasets of chest X-rays, CNNs can learn to detect pneumonia with accuracy comparable to or surpassing that of human radiologists. This automation not only accelerates diagnostic workflows, but also enhances consistency and reduces human error. This paper focuses on developing and evaluating a CNN model tailored for pneumonia detection, contributing to the advancement of AI-driven diagnostic solutions that can operate effectively in diverse clinical settings. Our approach addresses the need for scalable, reliable, and cost-effective diagnostic tools to improve healthcare outcomes globally.

2. Related Work

The application of convolutional neural networks (CNNs) to pneumonia diagnosis has been extensively studied, demonstrating the potential of deep learning in medical imaging [15]-[17]. A pivotal contribution by Wang *et al.* (2017) [18] introduced ChestX-ray8, a large-scale dataset comprising over 100,000 chest X-ray images annotated with various thoracic diseases, including pneumonia. Their work underscored the ability of CNNs to perform multilabel classification, laying the groundwork for future research in automated disease detection. Building on this foundation, Rajpurkar *et al.* (2018) [19] introduced CheXNet, a deep CNN model

trained on the ChestX-ray14 dataset, an extension of Wang *et al.*'s dataset. CheXNet achieved state-of-the-art performance, surpassing practicing radiologists in pneumonia detection accuracy. This achievement highlighted the viability of deep learning in critical diagnostic tasks, reinforcing the notion that AI models can complement and enhance the capabilities of medical professionals. While these studies demonstrate the efficacy of deep CNNs, they often rely on highly complex architectures, which can increase computational demands and reduce interpretability. Our approach seeks to optimize this process by developing a streamlined CNN architecture designed specifically for binary classification of pneumonia and normal cases. By incorporating techniques such as data augmentation, dropout regularization, and batch normalization, we enhance model generalization while minimizing overfitting. This approach allows for efficient training on limited datasets, providing a practical and scalable solution that can be readily deployed in diverse clinical environments.

3. Methodology

3.1. Dataset

This study utilizes a publicly available dataset of chest X-ray images, categorized into two classes: pneumonia and normal. The dataset serves as a benchmark for evaluating the effectiveness of convolutional neural networks (CNNs) in automated pneumonia diagnosis. **Figure 1** illustrates the distribution of cases within the dataset, revealing a noticeable class imbalance where pneumonia cases significantly outnumber normal cases. This imbalance reflects real-world scenarios but poses challenges during model training, potentially leading to biased predictions.

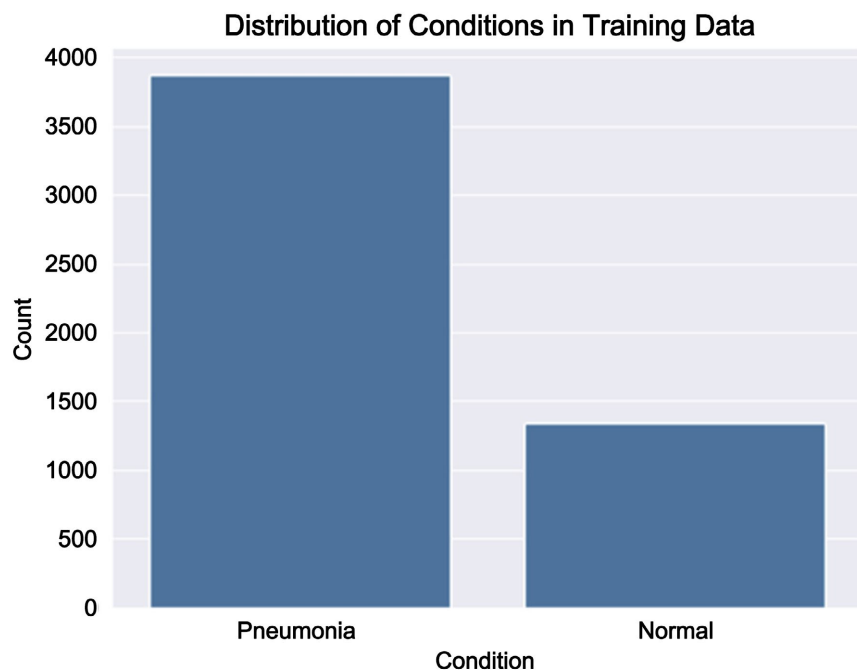


Figure 1. Distribution of pneumonia and normal conditions in the training data.

Addressing this imbalance is critical to ensuring the model generalizes well across both classes and does not disproportionately favor the majority class. While the dataset used in this study provided an essential benchmark for evaluating our model, it is derived from a single publicly available source. This limitation reduces the diversity of imaging conditions, demographics, and equipment used. Future work will incorporate additional datasets from diverse sources, such as hospitals in different geographical regions, to address potential biases and enhance generalizability. This step is critical to ensure the model's applicability across a broad range of clinical environments.

3.2. Data Preprocessing

We recommend implementing advanced techniques beyond basic data augmentation to mitigate class imbalance. Focal loss, which dynamically scales the loss for hard-to-classify examples, can increase the model's focus on minority classes. Additionally, cost-sensitive learning can assign higher weights to normal cases, balancing the loss function. Oversampling techniques like SMOTE (Synthetic Minority Oversampling Technique) could be applied to generate synthetic data for underrepresented classes. These approaches aim to enhance the model's classification accuracy and reduce misclassification rates. To standardize input data and improve model performance, several preprocessing steps were applied. Each chest X-ray image was resized to 150×150 pixels to ensure uniformity and reduce computational overhead. This dimension preserves essential features necessary for classification while balancing model efficiency. To further enhance generalization and reduce overfitting, data augmentation techniques were employed [20]-[22]. Augmentation introduces variations into the training data by applying random transformations, forcing the model to learn invariant features [23]-[25]. The following augmentation strategies were utilized:

Rotation—Random rotations up to 20 degrees to simulate variability in image orientation.

Zooming—Image zooming within a factor of 0.2 to simulate different scales of X-rays.

Horizontal Flipping—Random horizontal flips to mirror images, creating additional training samples.

Normalization was applied by scaling pixel values to the range (0, 1). This transformation ensures the model converges faster during training by stabilizing gradients and avoiding vanishing or exploding weights [26] [27]. Preprocessing is essential for minimizing discrepancies between training and validation data, thereby improving the model's generalization capabilities [28]-[30].

3.3. Model Architecture

The CNN model used for pneumonia classification was carefully designed to balance complexity, computational efficiency, and performance [31]-[33]. To enhance the robustness of the proposed model, we propose exploring advanced architectures

such as ResNet and DenseNet. These architectures offer deeper layers and skip connections, addressing issues like vanishing gradients and enabling the extraction of richer hierarchical features. Additionally, attention mechanisms like SE (Squeeze-and-Excitation) blocks or CBAM (Convolutional Block Attention Module) could further improve performance by focusing on critical regions of the chest X-rays. These techniques, combined with the current architecture, have the potential to significantly improve sensitivity and reduce false positives. Comparative experiments with this architecture will be conducted in future studies to assess their efficacy. The architecture consists of three convolutional layers, progressively increasing in depth to extract hierarchical features from chest X-ray images. A summary of the architecture is provided in **Figure 2**, with key components outlined as follows:

Model: "sequential_1"

Layer (type)	Output Shape	Param #
conv2d_2 (Conv2D)	(None, 150, 150, 32)	320
batch_normalization_2 (BatchNormalization)	(None, 150, 150, 32)	128
max_pooling2d_2 (MaxPooling2D)	(None, 75, 75, 32)	0
conv2d_3 (Conv2D)	(None, 75, 75, 16)	4,624
dropout_2 (Dropout)	(None, 75, 75, 16)	0
batch_normalization_3 (BatchNormalization)	(None, 75, 75, 16)	64
max_pooling2d_3 (MaxPooling2D)	(None, 37, 37, 16)	0
flatten_1 (Flatten)	(None, 21904)	0
dense_2 (Dense)	(None, 64)	1,401,920
dropout_3 (Dropout)	(None, 64)	0
dense_3 (Dense)	(None, 2)	130

Total params: 4,221,368 (16.10 MB)
 Trainable params: 1,407,090 (5.37 MB)
 Non-trainable params: 96 (384.00 B)
 Optimizer params: 2,814,182 (10.74 MB)

Figure 2. CNN model architecture summary.

Convolutional Layers: Three convolutional layers with filter sizes of 32, 64, and 128, each using 3×3 kernels. These layers detect spatial hierarchies, identifying features such as edges, patterns, and textures that characterize pneumonia. ReLU (Rectified Linear Unit) activation is applied to introduce non-linearity, allowing the network to model complex patterns [34]-[36].

Batch Normalization: Each convolutional layer is followed by batch normalization, which standardizes activations within the layer. This step accelerates convergence, stabilizes training, and reduces sensitivity to weight initialization [37] [38].

Max Pooling Layers: Max pooling (2×2) is applied after each convolution to

down sample feature maps, reducing spatial dimensions while retaining dominant features. This operation improves computational efficiency and controls overfitting by limiting the model's capacity [39] [40].

Flattening Layer: The output of the final convolutional block is flattened into a one-dimensional vector, preparing it for the fully connected dense layers. This vector encapsulates the learned features, acting as the input for the classification stage [41] [42].

Fully Connected Dense Layer: A dense layer with 128 units and ReLU activation integrates the extracted features to form high-level representations [43]-[45].

Dropout Layer: A dropout rate of 0.5 is applied to prevent overfitting by randomly deactivating neurons during training. This regularization technique forces the network to learn redundant representations, enhancing generalization [46]-[48].

Output Layer: The final dense layer contains two units with softmax activation, producing probability distributions for binary classification (pneumonia vs. normal) [49] [50].

The model was compiled using the Adam optimizer, chosen for its adaptability and efficient handling of sparse gradients. The initial learning rate was set to 0.001, ensuring stable convergence. Sparse categorical cross-entropy was selected as the loss function, aligning with the binary classification task. Model performance was monitored during training using accuracy and loss metrics, with early stopping applied to prevent overfitting and improve generalization. This architecture provides a balance between simplicity and performance, making it suitable for deployment in clinical environments for rapid pneumonia detection.

4. Training and Evaluation

4.1. Training Setup

The CNN model was trained over 12 epochs using a batch size of 32, which strikes a balance between computational efficiency and model convergence. A validation split of 20% was utilized, meaning 20% of the training data was set aside for validation at each epoch. This split allowed the model to be evaluated on unseen data during training, providing insight into its ability to generalize and prevent overfitting [51]-[53]. Training was conducted using the Adam optimizer, known for its adaptive learning rate and efficient handling of sparse gradients [54] [55]. The initial learning rate was set to 0.001, ensuring stable and gradual convergence [56] [57]. Sparse categorical cross-entropy was employed as the loss function, as it is well-suited for binary classification tasks where class labels are represented as integers [58]. Accuracy was used as the primary metric to track model performance across epochs. During each epoch, the model's performance on the training and validation sets was recorded. Accuracy and loss curves were plotted to visualize the progression of training (**Figure 3** and **Figure 4**). These plots serve as essential diagnostic tools, revealing whether the model experienced underfitting, overfitting, or achieved a balance between bias and variance.

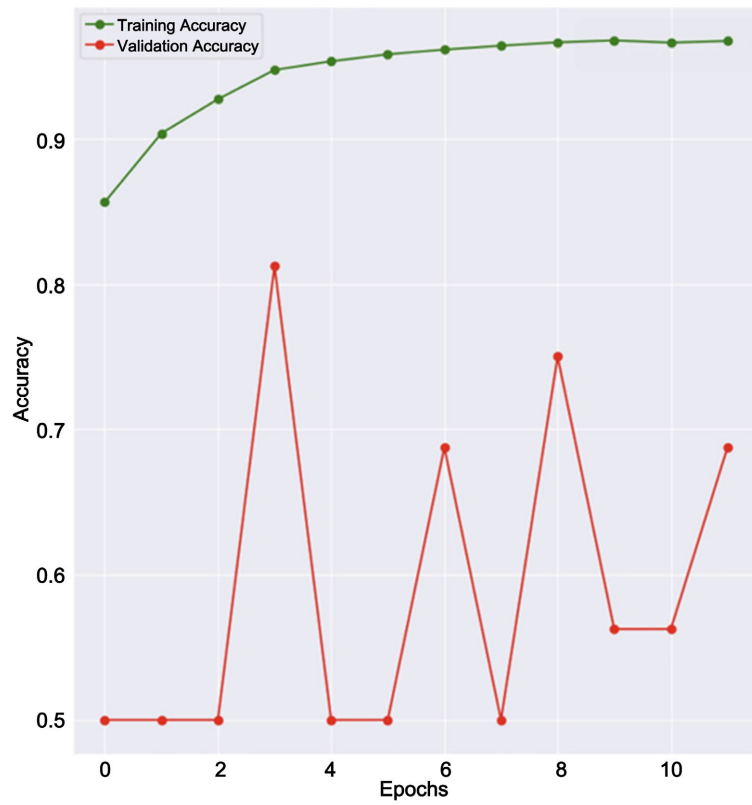


Figure 3. Training and validation accuracy over epochs.

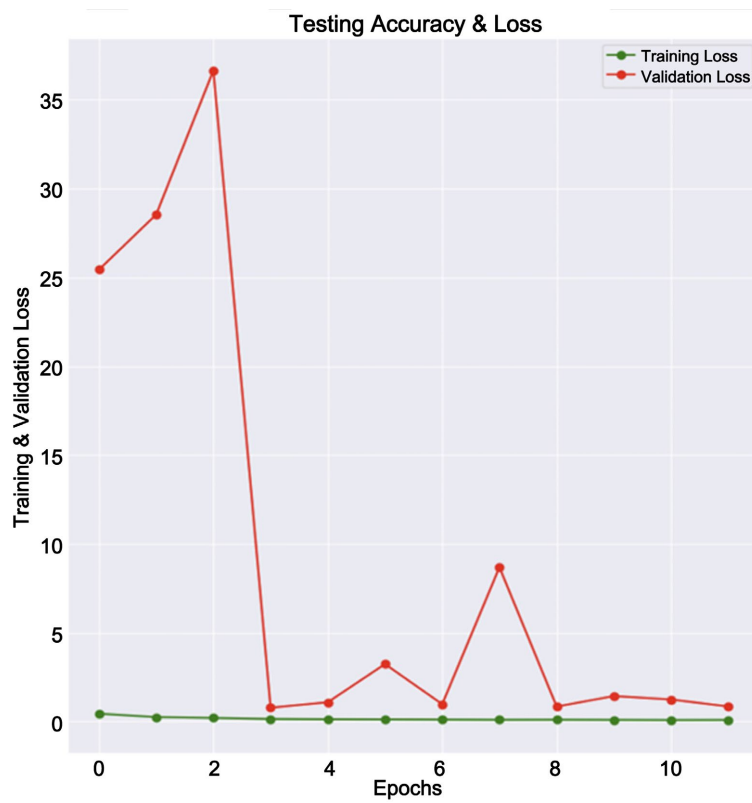


Figure 4. Training and validation loss over epochs.

The training accuracy curve showed a steady increase, while the validation accuracy exhibited fluctuations, suggesting potential overfitting. The loss curve displayed an initial spike in validation loss, stabilizing as training progressed, indicating that the model gradually learned meaningful patterns despite initial instability.

4.2. Evaluation Metrics

The validation accuracy curve exhibited fluctuations during training, indicating potential overfitting or instability. To address this, ensemble methods—combining predictions from multiple models—can reduce variance and enhance stability. Additionally, techniques like dropout regularization and weight decay can further prevent overfitting. Upon completing training, the model's performance was evaluated on the test set using a comprehensive set of metrics:

- Accuracy: Measures the overall proportion of correctly classified images.
- Precision: Reflects the proportion of positive predictions that were correct (pneumonia cases). This metric is essential to minimize false positives, which could lead to unnecessary medical intervention.
- Recall (Sensitivity): Measures the proportion of actual pneumonia cases that were correctly identified. A high recall reduces the risk of false negatives, which is critical in clinical scenarios to ensure no pneumonia cases are overlooked.
- F1-score: Represents the harmonic mean of precision and recall, providing a balanced measure that accounts for both false positives and false negatives.

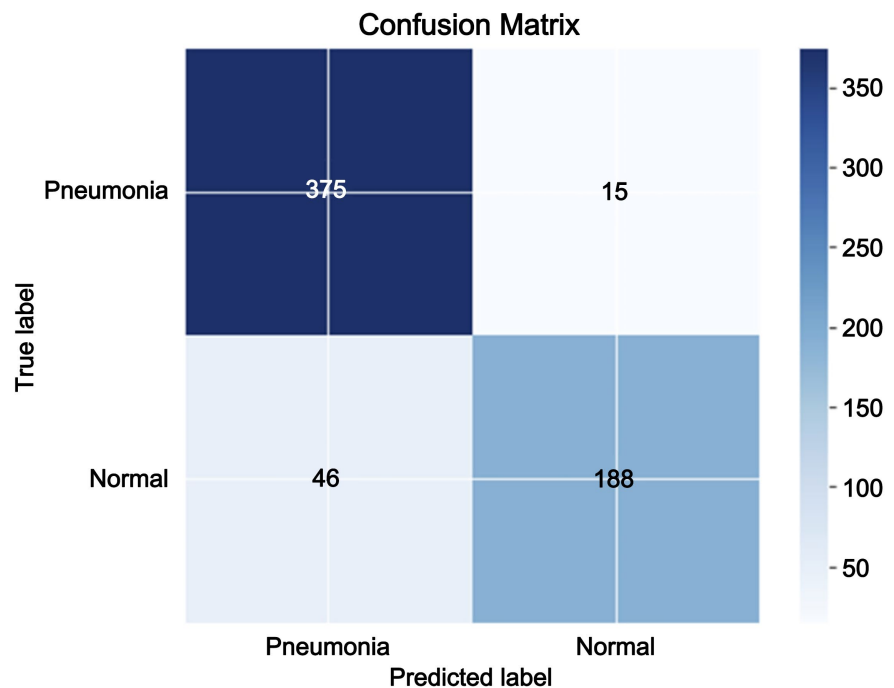


Figure 5. Confusion matrix for model predictions on test data.

The confusion matrix (**Figure 5**) visually represents the distribution of true

positives, true negatives, false positives, and false negatives. This matrix is crucial for identifying misclassification patterns. The model achieved a high recall for pneumonia cases, indicating strong sensitivity in detecting positive cases, while a slight tendency to misclassify normal cases as pneumonia was observed. In addition, the receiver operating characteristic (ROC) curve (**Figure 6**) was plotted to assess the model's discriminatory power. The ROC curve illustrates the trade-off between the true positive rate (sensitivity) and the false positive rate at various classification thresholds. The area under the curve (AUC) was calculated to be 0.96, indicating excellent performance and confirming the model's robustness in distinguishing between pneumonia and normal cases.

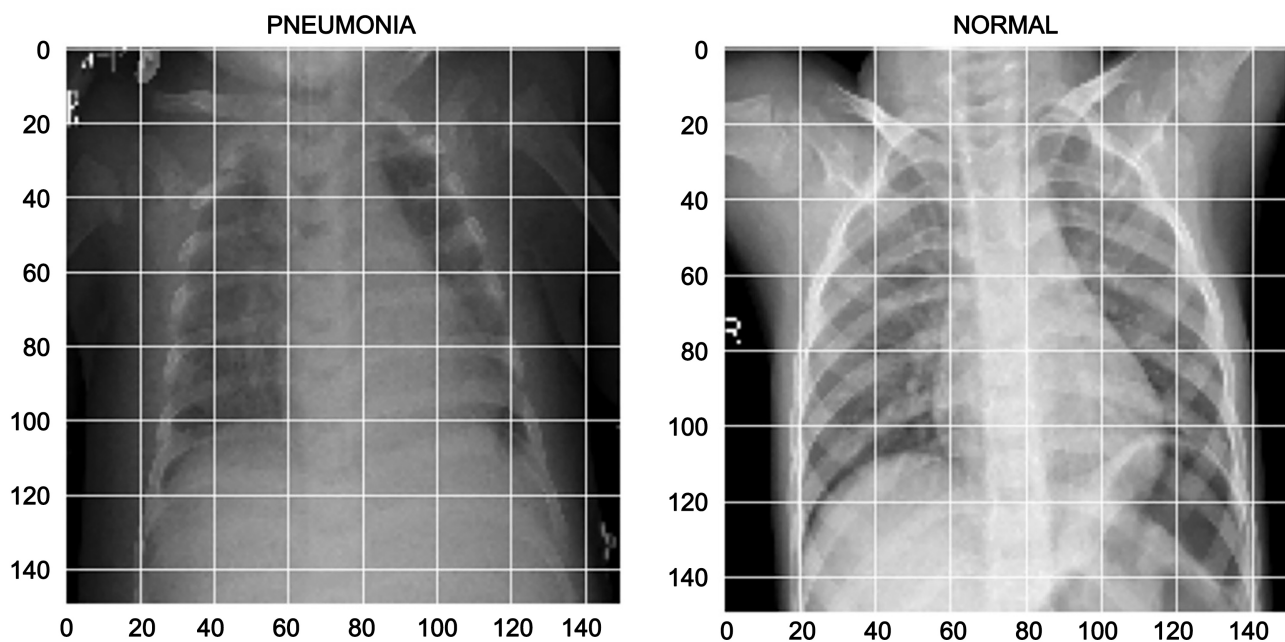


Figure 6. Comparative visualization of chest X-ray images.

The left image in **Figure 6** displays a pneumonia case characterized by irregular opacity in the lung regions, while the right image represents a normal chest X-ray with clear and symmetrical lung fields.

Figure 6 illustrates the visual differences between pneumonia and normal chest X-ray images. The top image highlights the increased opacity and irregular patterns often observed in pneumonia cases, whereas the bottom image depicts a normal X-ray with clear lung fields. This visual comparison emphasizes the challenge of distinguishing between subtle variations, reinforcing the necessity of automated CNN.

The combination of high accuracy (90.22%) in **Figure 7**, strong recall, and an impressive AUC score underscores the model's potential as a reliable tool for pneumonia diagnosis. The evaluation metrics collectively reflect the model's capacity to assist radiologists by reducing diagnostic errors and ensuring rapid detection of pneumonia in chest X-ray images.

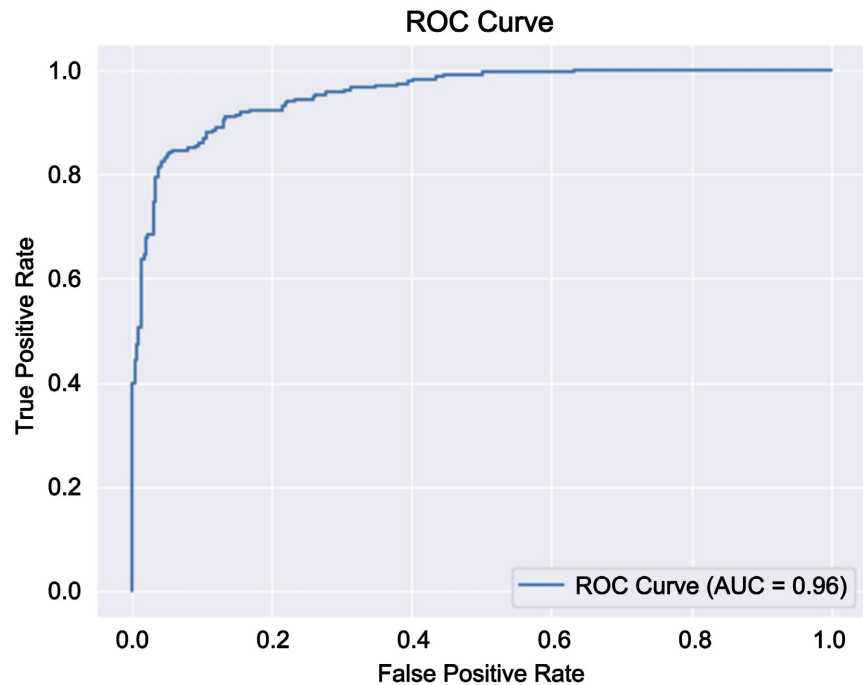


Figure 7. Receiver operating characteristic (ROC) curve with area under the curve (AUC) score.

5. Results

The final CNN model demonstrated strong performance during testing, achieving an overall accuracy of 90.22% on the test set, underscoring its effectiveness in distinguishing between pneumonia and normal chest X-ray images. This high accuracy reflects the model's ability to generalize across unseen data, validating the robustness of the training process and the architecture's capacity to extract meaningful features from the input images. A detailed breakdown of the model's performance is presented in **Figure 8**, which highlights key evaluation metrics such as precision, recall, and F1-score. For pneumonia detection, the model achieved a precision of 0.89 and a recall of 0.96. The high recall indicates that the model is highly sensitive, successfully identifying 96% of actual pneumonia cases minimizing the risk of false negatives. This is particularly crucial in clinical applications, where missing a pneumonia diagnosis can lead to severe health complications or delays in treatment. However, while the model excelled in detecting pneumonia, its performance in identifying normal cases was slightly less favorable. The recall for normal cases was 0.80, signifying that 20% of normal chest X-rays were misclassified as pneumonia. This conservative bias, while increasing false positives, is generally preferred in medical diagnostics, as it prioritizes capturing potential pneumonia cases over inadvertently dismissing them. False negatives (missed pneumonia) pose a greater clinical risk compared to false positives, which can be further evaluated by a radiologist to confirm the diagnosis. The overall F1-score, which balances precision and recall, highlights the model's stable and reliable classification performance across both classes. The ROC curve and confusion matrix

corroborate these findings, demonstrating that the model consistently performs well across varying classification thresholds.

Accuracy: 90.22%				
	precision	recall	f1-score	support
Pneumonia	0.89	0.96	0.92	390
Normal	0.93	0.80	0.86	234
accuracy			0.90	624
macro avg	0.91	0.88	0.89	624
weighted avg	0.90	0.90	0.90	624

Figure 8. Classification report—precision, recall, and F1-score for pneumonia diagnosis.

These results reflect the potential of the CNN model to serve as an auxiliary diagnostic tool, assisting radiologists in detecting pneumonia cases with high accuracy and contributing to faster, more efficient patient care. Future improvements may focus on enhancing the model's specificity to reduce misclassification of normal cases, ultimately refining its application for real-world deployment in medical environments.

6. Discussion

The results of this study underscore the significant potential of convolutional neural networks (CNNs) as reliable tools for automated pneumonia diagnosis using chest X-ray images. The model's high recall of 96% for pneumonia cases indicates its strong capability to minimize false negatives, which is crucial in clinical settings where missing a diagnosis could lead to severe patient outcomes, including disease progression and increased mortality. This sensitivity ensures that the model prioritizes detecting pneumonia cases, aligning with the primary objective of reducing undiagnosed cases in vulnerable populations. However, while the model excels at identifying pneumonia, its tendency to misclassify normal cases as pneumonia—as reflected by the lower recall of 0.80% for normal images—highlights areas for improvement. This bias toward detecting pneumonia, though beneficial in reducing missed cases, may lead to false positives and unnecessary medical interventions [59] [60]. Although false positives are less dangerous than false negatives, they can increase patient anxiety, healthcare costs, and unnecessary treatments [61] [62]. Addressing this challenge by enhancing the model's specificity without compromising sensitivity is a priority for future work. The fluctuations observed in validation accuracy throughout training indicate possible instability or overfitting to the training data. This suggests that while the model generalizes well to some extent, its performance might degrade when exposed to diverse or more complex datasets. This variability could stem from factors such as class imbalance, limited data volume, or insufficient regularization. Implementing techniques like cross-validation, increasing the size of the dataset, and further hyperparameter tuning could enhance stability and mitigate overfitting. The confusion matrix

provides deeper insight into the classification performance, visually representing the distribution of misclassified and correctly classified cases. The model demonstrates robust performance in correctly identifying pneumonia, but the misclassification of normal cases warrants further investigation. One potential solution is the application of ensemble methods that aggregate the predictions of multiple models, reducing variance and improving overall accuracy [63] [64]. Alternatively, leveraging transfer learning from pre-trained models such as ResNet or VGG could improve feature extraction and enhance performance, particularly for normal cases [65] [66]. Future work will prioritize the expansion of the dataset by incorporating more diverse chest X-ray images from different sources, ensuring the model is exposed to a broader range of cases and patient demographics. Exploring deeper and more complex architectures, or hybrid approaches that integrate attention mechanisms, could also help the model focus on the most relevant regions of the X-ray, thereby reducing misclassification [67] [68]. Additionally, techniques like focal loss may be implemented to specifically address the class imbalance, ensuring the model pays greater attention to minority classes [69] [70]. While the CNN model demonstrates strong performance and promising results for pneumonia detection, continued refinement and expansion are essential to ensure its reliability across varying clinical environments. By addressing the limitations identified, this AI-driven diagnostic tool could be deployed as a scalable solution, aiding radiologists and enhancing the early detection of pneumonia, ultimately improving patient care and outcomes.

7. Conclusion

This study highlights the effectiveness of convolutional neural networks (CNNs) in automating pneumonia diagnosis from chest X-ray images, demonstrating their potential to transform medical imaging and diagnostic workflows. The proposed model achieved a test accuracy of 90.22% and an AUC score of 0.96, indicating its robust capability to distinguish between pneumonia and normal cases. The model's high recall rate of 96% for pneumonia detection underscores its strength in minimizing false negatives, which is crucial in clinical environments where early diagnosis directly correlates with improved patient outcomes. By automating the detection process, this AI-driven approach alleviates the diagnostic burden on radiologists, enabling faster analysis and more consistent results. This is particularly valuable in resource-limited settings, where radiology expertise may be scarce. The model's performance suggests its suitability as a complementary tool to support radiologists, reducing variability and enhancing the accuracy of pneumonia diagnosis. Despite the promising results, the study also identifies areas for improvement. The model exhibited a slight tendency to misclassify normal cases as pneumonia, emphasizing the need for further refinement to enhance specificity without compromising sensitivity. Expanding the dataset, integrating more complex architectures, and leveraging techniques such as transfer learning or ensemble methods will be key to addressing these limitations. In conclusion, this research

underscores the viability of CNNs in medical diagnostics and sets the stage for future advancements that could lead to the widespread adoption of AI-assisted diagnostic tools, ultimately improving healthcare delivery and patient outcomes on a global scale.

Conflicts of Interest

The authors declare no conflicts of interest.

References

- [1] Mani, C.S. (2018) Acute Pneumonia and Its Complications. In: Long, S.S., Prober, C.G. and Fischer, M., Eds., *Principles and Practice of Pediatric Infectious Diseases*, Elsevier, 238-249.e4. <https://doi.org/10.1016/b978-0-323-40181-4.00034-7>
- [2] Marrie, T.J. (1994) Community-Acquired Pneumonia. *Clinical Infectious Diseases*, **18**, 501-515. <https://doi.org/10.1093/clinids/18.4.501>
- [3] Nascimento-Carvalho, C.M. (2020) Community-Acquired Pneumonia among Children: The Latest Evidence for an Updated Management. *Jornal de Pediatria*, **96**, 29-38. <https://doi.org/10.1016/j.jped.2019.08.003>
- [4] Eshwara, V.K., Mukhopadhyay, C. and Rello, J. (2020) Community-Acquired Bacterial Pneumonia in Adults. *Indian Journal of Medical Research*, **151**, 287-302. https://doi.org/10.4103/ijmr.ijmr_1678_19
- [5] Jean, S., Chang, Y., Lin, W., Lee, W., Hsueh, P. and Hsu, C. (2020) Epidemiology, Treatment, and Prevention of Nosocomial Bacterial Pneumonia. *Journal of Clinical Medicine*, **9**, Article 275. <https://doi.org/10.3390/jcm9010275>
- [6] Baldo, V., Cocchio, S., Baldovin, T., Buja, A., Furlan, P., Bertonecello, C., et al. (2014) A Population-Based Study on the Impact of Hospitalization for Pneumonia in Different Age Groups. *BMC Infectious Diseases*, **14**, Article No. 485. <https://doi.org/10.1186/1471-2334-14-485>
- [7] Falsey, A.R. and Walsh, E.E. (2006) Viral Pneumonia in Older Adults. *Clinical Infectious Diseases*, **42**, 518-524. <https://doi.org/10.1086/499955>
- [8] Ali, M., Shahroz, M., Akram, U., Mushtaq, M.F., Altamiranda, S.C., Obregon, S.A., et al. (2024) Pneumonia Detection Using Chest Radiographs with Novel EfficientNetV2L Model. *IEEE Access*, **12**, 34691-34707. <https://doi.org/10.1109/access.2024.3372588>
- [9] Abdelwanis, M., Alarafati, H.K., Tammam, M.M.S. and Simsekler, M.C.E. (2024) Exploring the Risks of Automation Bias in Healthcare Artificial Intelligence Applications: A Bowtie Analysis. *Journal of Safety Science and Resilience*, **5**, 460-469. <https://doi.org/10.1016/j.jnlssr.2024.06.001>
- [10] Maambo, M. (2023) Assisted Artificial Intelligence Medical Diagnosis System for Heart Disease. Ph.D. Thesis, The University of Zambia.
- [11] Najjar, R. (2023) Redefining Radiology: A Review of Artificial Intelligence Integration in Medical Imaging. *Diagnostics*, **13**, Article 2760. <https://doi.org/10.3390/diagnostics13172760>
- [12] Pillai, A.S. (2021) Utilizing Deep Learning in Medical Image Analysis for Enhanced Diagnostic Accuracy and Patient Care: Challenges, Opportunities, and Ethical Implications. *Journal of Deep Learning in Genomic Data Analysis*, **1**, 1-17.
- [13] Anwar, S.M., Majid, M., Qayyum, A., Awais, M., Alnowami, M. and Khan, M.K. (2018) Medical Image Analysis Using Convolutional Neural Networks: A Review. *Journal of Medical Systems*, **42**, Article No. 226. <https://doi.org/10.1007/s10916-018-1088-1>

- [14] Sarvamangala, D.R. and Kulkarni, R.V. (2021) Convolutional Neural Networks in Medical Image Understanding: A Survey. *Evolutionary Intelligence*, **15**, 1-22. <https://doi.org/10.1007/s12065-020-00540-3>
- [15] Szepesi, P. and Szilágyi, L. (2022) Detection of Pneumonia Using Convolutional Neural Networks and Deep Learning. *Biocybernetics and Biomedical Engineering*, **42**, 1012-1022. <https://doi.org/10.1016/j.bbe.2022.08.001>
- [16] Nessipkhanov, D., Davletova, V., Kurmanbekkyzy, N. and Omarov, B. (2023) Deep CNN for the Identification of Pneumonia Respiratory Disease in Chest X-Ray Imagery. *International Journal of Advanced Computer Science and Applications*, **14**, Article 2023. <https://doi.org/10.14569/ijacsa.2023.0141069>
- [17] Chen, J., Li, Y., Guo, L., Zhou, X., Zhu, Y., He, Q., et al. (2022) Machine Learning Techniques for CT Imaging Diagnosis of Novel Coronavirus Pneumonia: A Review. *Neural Computing and Applications*, **36**, 181-199. <https://doi.org/10.1007/s00521-022-07709-0>
- [18] Wang, X., Peng, Y., Lu, L., Lu, Z., Bagheri, M. and Summers, R.M. (2017) ChestX-ray8: Hospital-Scale Chest X-Ray Database and Benchmarks on Weakly-Supervised Classification and Localization of Common Thorax Diseases. 2017 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, 21-26 July 2017, 3462-3471. <https://doi.org/10.1109/cvpr.2017.369>
- [19] Rajpurkar, P., Irvin, J., Zhu, K., Yang, B., Mehta, H., Duan, T. and Ng, A. Y. (2018) CheXNet: Radiologist-Level Pneumonia Detection on Chest X-Rays with Deep Learning. arXiv: 1711.05225.
- [20] Shorten, C. and Khoshgoftaar, T.M. (2019) A Survey on Image Data Augmentation for Deep Learning. *Journal of Big Data*, **6**, Article No. 60. <https://doi.org/10.1186/s40537-019-0197-0>
- [21] Khosla, C. and Saini, B.S. (2020) Enhancing Performance of Deep Learning Models with Different Data Augmentation Techniques: A Survey. 2020 *International Conference on Intelligent Engineering and Management (ICIEM)*, London, 17-19 June 2020, 79-85. <https://doi.org/10.1109/iciem48762.2020.9160048>
- [22] Taylor, L. and Nitschke, G. (2018). Improving Deep Learning with Generic Data Augmentation. 2018 *IEEE Symposium Series on Computational Intelligence (SSCI)*, Bangalore, 18-21 November 2018, 1542-1547. <https://doi.org/10.1109/ssci.2018.8628742>
- [23] Liu, B., Wang, X., Dixit, M., Kwitt, R. and Vasconcelos, N. (2018) Feature Space Transfer for Data Augmentation. 2018 *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Salt Lake City, 18-23 June 2018, 9090-9098. <https://doi.org/10.1109/cvpr.2018.00947>
- [24] Li, P., Li, D., Li, W., Gong, S., Fu, Y. and Hospedales, T.M. (2021) A Simple Feature Augmentation for Domain Generalization. 2021 *IEEE/CVF International Conference on Computer Vision (ICCV)*, Montreal, 10-17 October 2021, 8866-8875. <https://doi.org/10.1109/iccv48922.2021.00876>
- [25] Hauberg, S., Freifeld, O., Lindbo Larsen, A.B., Fisher, J. and Hansen, L. (2016) Dreaming More Data: Class-Dependent Distributions over Diffeomorphisms for Learned Data Augmentation. arXiv: 1510.02795.
- [26] Huang, L., Qin, J., Zhou, Y., Zhu, F., Liu, L. and Shao, L. (2023) Normalization Techniques in Training DNNs: Methodology, Analysis and Application. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **45**, 10173-10196. <https://doi.org/10.1109/tpami.2023.3250241>
- [27] Zhang, J., Lei, Q. and Dhillon, I. (2018) Stabilizing Gradients for Deep Neural Networks via Efficient SVD Parameterization. arXiv: 1803.09327.

- [28] Fan, C., Chen, M., Wang, X., Wang, J. and Huang, B. (2021) A Review on Data Preprocessing Techniques toward Efficient and Reliable Knowledge Discovery from Building Operational Data. *Frontiers in Energy Research*, **9**, Article 652801. <https://doi.org/10.3389/fenrg.2021.652801>
- [29] Alexandropoulos, S.N., Kotsiantis, S.B. and Vrahatis, M.N. (2019) Data Preprocessing in Predictive Data Mining. *The Knowledge Engineering Review*, **34**, e1. <https://doi.org/10.1017/s026988891800036x>
- [30] Habib, M. and Okayli, M. (2024) Evaluating the Sensitivity of Machine Learning Models to Data Preprocessing Technique in Concrete Compressive Strength Estimation. *Arabian Journal for Science and Engineering*, **49**, 13709-13727. <https://doi.org/10.1007/s13369-024-08776-2>
- [31] An, Q., Chen, W. and Shao, W. (2024) A Deep Convolutional Neural Network for Pneumonia Detection in X-Ray Images with Attention Ensemble. *Diagnostics*, **14**, Article 390. <https://doi.org/10.3390/diagnostics14040390>
- [32] Akbar, W., Soomro, A., Hussain, A., Hussain, T., Ali, F., Haq, M.I.U., *et al.* (2024) Pneumonia Detection: A Comprehensive Study of Diverse Neural Network Architectures Using Chest X-Rays. *International Journal of Applied Mathematics and Computer Science*, **34**, 679-699. <https://doi.org/10.61822/amcs-2024-0045>
- [33] Jia, G., Lam, H. and Xu, Y. (2021) Classification of COVID-19 Chest X-Ray and CT Images Using a Type of Dynamic CNN Modification Method. *Computers in Biology and Medicine*, **134**, Article ID: 104425. <https://doi.org/10.1016/j.combiomed.2021.104425>
- [34] Yu, Y., Adu, K., Tashi, N., Anokye, P., Wang, X. and Ayidzoe, M.A. (2020) RMAF: Relu-Memristor-Like Activation Function for Deep Learning. *IEEE Access*, **8**, 72727-72741. <https://doi.org/10.1109/access.2020.2987829>
- [35] Karnewar, A., Ritschel, T., Wang, O. and Mitra, N. (2022) ReLU Fields: The Little Non-Linearity That Could. *Special Interest Group on Computer Graphics and Interactive Techniques Conference Proceedings*, Vancouver, 7-11 August 2022, 1-9. <https://doi.org/10.1145/3528233.3530707>
- [36] Dubey, S.R., Singh, S.K. and Chaudhuri, B.B. (2022) Activation Functions in Deep Learning: A Comprehensive Survey and Benchmark. *Neurocomputing*, **503**, 92-108. <https://doi.org/10.1016/j.neucom.2022.06.111>
- [37] Qiao, S.Y., Wang, H.Y., Liu, C.X., Shen, W. and Yuille, A. (2019) Micro-Batch Training with Batch-Channel Normalization and Weight Standardization. arXiv: 1903.10520.
- [38] Shao, J., Hu, K., Wang, C.H., Xue, X.Y. and Raj, B. (2020) Is Normalization Indispensable for Training Deep Neural Network? *Advances in Neural Information Processing Systems*, **33**, 13434-13444.
- [39] Zafar, A., Saba, N., Arshad, A., Alabrah, A., Riaz, S., Suleman, M., *et al.* (2024) Convolutional Neural Networks: A Comprehensive Evaluation and Benchmarking of Pooling Layer Variants. *Symmetry*, **16**, Article 1516. <https://doi.org/10.3390/sym16111516>
- [40] Noel, D. (2023) An Investigation of Methods for Improving Spatial Invariance of Convolutional Neural Networks for Image Classification. Ph.D. Thesis, Nova Southeastern University.
- [41] Musthafa, M.M., T R, M., V, V.K. and Guluwadi, S. (2024) Enhanced Skin Cancer Diagnosis Using Optimized CNN Architecture and Checkpoints for Automated Dermatological Lesion Classification. *BMC Medical Imaging*, **24**, Article No. 201. <https://doi.org/10.1186/s12880-024-01356-8>
- [42] Mahamud, E., Fahad, N., Assaduzzaman, M., Zain, S.M., Goh, K.O.M. and Morol,

- M.K. (2024) An Explainable Artificial Intelligence Model for Multiple Lung Diseases Classification from Chest X-Ray Images Using Fine-Tuned Transfer Learning. *Decision Analytics Journal*, **12**, Article ID: 100499. <https://doi.org/10.1016/j.dajour.2024.100499>
- [43] Huang, G., Liu, Z., Van Der Maaten, L. and Weinberger, K.Q. (2017) Densely Connected Convolutional Networks. 2017 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, 21-26 July 2017, 2261-2269. <https://doi.org/10.1109/cvpr.2017.243>
- [44] Tayara, H. and Chong, K.T. (2018) Object Detection in Very High-Resolution Aerial Images Using One-Stage Densely Connected Feature Pyramid Network. *Sensors*, **18**, Article 3341. <https://doi.org/10.3390/s18103341>
- [45] Francis, M. and Deisy, C. (2019) Disease Detection and Classification in Agricultural Plants Using Convolutional Neural Networks—A Visual Understanding. 2019 *6th International Conference on Signal Processing and Integrated Networks (SPIN)*, Noida, 7-8 March 2019, 1063-1068. <https://doi.org/10.1109/spin.2019.8711701>
- [46] Khan, S.H., Hayat, M. and Porikli, F. (2019) Regularization of Deep Neural Networks with Spectral Dropout. *Neural Networks*, **110**, 82-90. <https://doi.org/10.1016/j.neunet.2018.09.009>
- [47] Santos, C.F.G.D. and Papa, J.P. (2022) Avoiding Overfitting: A Survey on Regularization Methods for Convolutional Neural Networks. *ACM Computing Surveys*, **54**, 1-25. <https://doi.org/10.1145/3510413>
- [48] Bejani, M.M. and Ghatee, M. (2021) A Systematic Review on Overfitting Control in Shallow and Deep Neural Networks. *Artificial Intelligence Review*, **54**, 6391-6438. <https://doi.org/10.1007/s10462-021-09975-1>
- [49] Kong, L. and Cheng, J. (2021) Based on Improved Deep Convolutional Neural Network Model Pneumonia Image Classification. *PLOS ONE*, **16**, e0258804. <https://doi.org/10.1371/journal.pone.0258804>
- [50] El Asnaoui, K., Chawki, Y. and Idri, A. (2021) Automated Methods for Detection and Classification Pneumonia Based on X-Ray Images Using Deep Learning. In: Maleh, Y., Baddi, Y., Alazab, M., Tawalbeh, L. and Romdhani, I., Eds., *Artificial Intelligence and Blockchain for Future Cybersecurity Applications*, Springer, 257-284. https://doi.org/10.1007/978-3-030-74575-2_14
- [51] Shao, Y. and Lunetta, R.S. (2012) Comparison of Support Vector Machine, Neural Network, and CART Algorithms for the Land-Cover Classification Using Limited Training Data Points. *ISPRS Journal of Photogrammetry and Remote Sensing*, **70**, 78-87. <https://doi.org/10.1016/j.isprsjprs.2012.04.001>
- [52] Raschka, S. (2018) Model Evaluation, Model Selection, and Algorithm Selection in Machine Learning. arXiv: 1811.12808.
- [53] LeBaron, B. and Weigend, A.S. (1998) A Bootstrap Evaluation of the Effect of Data Splitting on Financial Time Series. *IEEE Transactions on Neural Networks*, **9**, 213-220. <https://doi.org/10.1109/72.655043>
- [54] Dogo, E.M., Afolabi, O.J., Nwulu, N.I., Twala, B. and Aigbavboa, C.O. (2018) A Comparative Analysis of Gradient Descent-Based Optimization Algorithms on Convolutional Neural Networks. 2018 *International Conference on Computational Techniques, Electronics and Mechanical Systems (CTEMS)*, Belgaum, 21-22 December 2018, 92-99. <https://doi.org/10.1109/ctems.2018.8769211>
- [55] Zaheer, M., Reddi, S., Sachan, D., Kale, S. and Kumar, S. (2018) Adaptive Methods for Nonconvex Optimization. *Proceedings of the 32nd International Conference on Neural Information Processing Systems*, Montréal, 3 December 2018, 9815-9825.
- [56] Liang Jin, and Gupta, M.M. (1999) Stable Dynamic Backpropagation Learning in

- Recurrent Neural Networks. *IEEE Transactions on Neural Networks*, **10**, 1321-1334. <https://doi.org/10.1109/72.809078>
- [57] Wu, Y., Liu, L., Bae, J., Chow, K., Iyengar, A., Pu, C., *et al.* (2019) Demystifying Learning Rate Policies for High Accuracy Training of Deep Neural Networks. 2019 *IEEE International Conference on Big Data (Big Data)*, Los Angeles, 9-12 December 2019, 1971-1980. <https://doi.org/10.1109/bigdata47090.2019.9006104>
- [58] Porter, E., Solis, D., Bruckmeier, P., Siddiqui, Z.A., Zamdborg, L. and Guerrero, T. (2021) Effect of Loss Functions in Deep Learning-Based Segmentation. In: Yang, J.Z., Sharp, G.C. and Gooding, M.J., Eds., *Auto-Segmentation for Radiation Oncology*, CRC Press, 133-150. <https://doi.org/10.1201/9780429323782-12>
- [59] Niederman, M.S., Torres, A. and Summer, W. (1994) Invasive Diagnostic Testing Is Not Needed Routinely to Manage Suspected Ventilator-Associated Pneumonia. *American Journal of Respiratory and Critical Care Medicine*, **150**, 565-569. <https://doi.org/10.1164/ajrccm.150.2.8049849>
- [60] Brusselaers, N., Labeau, S., Vogelaers, D. and Blot, S. (2012) Value of Lower Respiratory Tract Surveillance Cultures to Predict Bacterial Pathogens in Ventilator-Associated Pneumonia: Systematic Review and Diagnostic Test Accuracy Meta-Analysis. *Intensive Care Medicine*, **39**, 365-375. <https://doi.org/10.1007/s00134-012-2759-x>
- [61] Wardle, J. and Pope, R. (1992) The Psychological Costs of Screening for Cancer. *Journal of Psychosomatic Research*, **36**, 609-624. [https://doi.org/10.1016/0022-3999\(92\)90051-3](https://doi.org/10.1016/0022-3999(92)90051-3)
- [62] Wakefield, J.C. (2010) Misdiagnosing Normality: Psychiatry's Failure to Address the Problem of False Positive Diagnoses of Mental Disorder in a Changing Professional Environment. *Journal of Mental Health*, **19**, 337-351. <https://doi.org/10.3109/09638237.2010.492418>
- [63] Seni, G. and Elder, J. (2010) Ensemble Methods in Data Mining: Improving Accuracy through Combining Predictions. Morgan & Claypool Publishers.
- [64] Webb, G.I. and Zheng, Z. (2004) Multistrategy Ensemble Learning: Reducing Error by Combining Ensemble Learning Techniques. *IEEE Transactions on Knowledge and Data Engineering*, **16**, 980-991. <https://doi.org/10.1109/tkde.2004.29>
- [65] Alammari, Z., Alzubaidi, L., Zhang, J., Li, Y., Lafta, W. and Gu, Y. (2023) Deep Transfer Learning with Enhanced Feature Fusion for Detection of Abnormalities in X-Ray Images. *Cancers*, **15**, Article 4007. <https://doi.org/10.3390/cancers15154007>
- [66] Ashraf, A., Naz, S., Shirazi, S.H., Razzak, I. and Parsad, M. (2021) Deep Transfer Learning for Alzheimer Neurological Disorder Detection. *Multimedia Tools and Applications*, **80**, 30117-30142. <https://doi.org/10.1007/s11042-020-10331-8>
- [67] Siddiqi, R. and Javaid, S. (2024) Deep Learning for Pneumonia Detection in Chest X-Ray Images: A Comprehensive Survey. *Journal of Imaging*, **10**, Article 176. <https://doi.org/10.3390/jimaging10080176>
- [68] Feng, Y., Yang, X., Qiu, D., Zhang, H., Wei, D. and Liu, J. (2022) Pcxrnet: Pneumonia Diagnosis from Chest X-Ray Images Using Condense Attention Block and Multiconvolution Attention Block. *IEEE Journal of Biomedical and Health Informatics*, **26**, 1484-1495. <https://doi.org/10.1109/jbhi.2022.3148317>
- [69] Cui, L., Li, D., Yang, X., Liu, C. and Yan, X. (2024) A Unified Approach Addressing Class Imbalance in Magnetic Resonance Image for Deep Learning Models. *IEEE Access*, **12**, 27368-27384. <https://doi.org/10.1109/access.2024.3365544>
- [70] Zhang, M., Fan, B., Zhang, N., Wang, W. and Fan, W. (2021) Mining Product Innovation Ideas from Online Reviews. *Information Processing & Management*, **58**, Article ID: 102389. <https://doi.org/10.1016/j.ipm.2020.102389>