From Pixels to Diagnosis: AI-Powered CNN for Pneumonia Detection

1st Bilal Hameed Zaid
Continuing Education Center, University of Baghdad
Baghdad, Iraq
bilalalaameri@dcec.uobaghdad.edu.iq

2nd Ahmed Hadi Al-Dubaisi
Computer Science and Information Technology Department, College of Engineering, Abu Dhabi University
UAE, Abu Dhabi
1074630@alumni.adu.ac.ae

3rd Anfal Sabeeh Hamood
Department of Journalism, College of Mass Communication, University of Baghdad
Baghdad, Iraq
anfaltamimi@comc.uobaghdad.edu.iq

***Abstract*—Pneumonia is a major respiratory infection and is one of the top causes of death among young children around the world. Traditional diagnostic approaches, including X-ray imaging, require expert analysis and are often limited by delayed intervention and misdiagnosis, especially in resource-limited settings. In this study, we propose a convolutional neural network (CNN)-based model to facilitate the automatic identification of pneumonia from chest X-rays. The model was trained on a dataset of 5,863 pediatric X-ray images and achieved an accuracy of 93%, precision of 91%, recall of 96%, and F1-score of 94%. To optimize performance, multiple preprocessing steps were performed, including grayscale conversion, image resizing, normalizing, and data augmentation. Future work should focus on scaling the model by using larger datasets and incorporating transfer learning to amplify performance and generalizability even further.**

Keywords—Pneumonia Detection, Convolutional Neural Network (CNN), Data Augmentation, Chest X-ray, Medical Imaging, Pediatric Radiology

I. INTRODUCTION

Pneumonia is a serious respiratory condition that infects and inflames the lungs' air sacs, which can fill with fluid or pus, resulting in coughing, fever, chills and trouble breathing. It is caused from a variety of pathogens, all from bacteria and viruses to fungi, most commonly transmitted when someone coughs or sneezes, as well as through direct contact with contaminated objects or people [1]. Pneumonia can affect individuals of all ages and body types, but is particularly dangerous for individuals below five and above sixty-five, as well as those with an unbalanced immune system.

According to the World Health Organization, pneumonia has caused an estimated 2.5 million deaths worldwide in 2019 alone, making it among the leading causes of infectious disease deaths globally [2]. This emphasizes how important timely and accurate diagnosis is for effective treatment and intervention, with the high mortality rate. Nonetheless, existing diagnostic methods, including visual assessment of chest X-ray images by radiologists, suffer from various drawbacks [3]. These traditional methods are not only time consuming but also prone to human error, variability in expertise and fatigue.

Diagnostic delays are particularly common in many under resourced medical settings where such specialists are in short supply. Additionally, X-rays may be unable to detect faint signs of pneumonia, leading to misdiagnosis or delayed treatment that can be a life-threatening event [4]. The interpretive process is also very subjective; for example, same image can produce different conclusions from different radiologists, and more so in borderline or early stage cases. Moreover, traditional imaging equipment may have poor image quality caused by old machines or bad imaging conditions, which makes diagnosis inaccurate. Such compounded limitations point out the urgent need of a more objective, efficient and reliable diagnostic aid which can operate in consistent manner in different clinical environments.

To help mitigate this problem, this study presents a deep learning solution that utilized a Convolutional Neural Network (CNN) approach for automated pneumonia screening from X-ray images [5]. This technique harnesses CNN's strength in feature extraction and learning of complex patterns from medical images, ultimately seeking to enhance the speed, accuracy, and consistency of diagnostics, while providing a scalable and highly effective approach for the future of healthcare systems.

The rest of the paper is structured as follows: Section II reviews related works; Section III describes the proposed methodology; Section IV presents results and discussion; and Section V summarizes the main findings and further directions.

II. LITERATURE REVIEW

As the deep learning technologies continue to grow rapidly, many studies have been performed to explore the use of convolutional neural networks (CNNs) for automated detection of pneumonia from chest X-rays. The aim of these efforts is to address the diagnostic limitations of the healthcare professionals, especially in the low resource setting. The contributions to this field which are highlighted in this section are in terms of model performance, dataset, and architectural decisions.

In a study by Yuvraj Sinha Chowdhury et al. [6], various CNN models were run on a Kaggle chest X-ray dataset with varying numbers of iterations, CNN layers and optimizers. The best performing model was achieved by the one with four hidden layers and the SGD optimizer that was at 91% accuracy. It shows that network depth and optimizer choice can have important impact on pneumonia detection performance in CNN based systems.

Another study by V. Sirish Kaushik et al. [7] used four CNN architectures for pediatric pneumonia detection and tested the models on a Kaggle dataset. The model with two convolutional layers had its lowest accuracy with 85.26%, and its highest reaching 92.31%. To prevent overfitting, dropout was used, and recall and F1 score were evaluated for each model.

According to a study by Swapnil Singh [8], which used Kaggle's chest X-ray dataset, the pneumonia detection was carried out with CNN and multilayer perceptron models. A GUI was developed to predict the pneumonia and congestion percentage. The CNN model outperformed the multilayer perceptron with 92.63% of accuracy, which was the highest and multilayer perceptron with 77.56% of accuracy was the lowest, proving that CNN is more reliable and an automated diagnosis.

Moreover, Orlando et al. [9] conducted experiments on four pre-trained CNN models, namely VGG16, VGG19, ResNet50 and InceptionV3 for detecting pneumonia from chest X-ray images. The study had sought to detect the disease early in order to reduce child mortality and according to which InceptionV3 got the highest accuracy of 72.9%.

Mudasir Ali et al. [10] explored a dataset of 5,856 chest X-rays that were used to evaluate six deep learning models including: CNN, InceptionResNetV2, Xception, VGG16, ResNet50, and EfficientNetV2L. Training was done with the Adam optimizer models, and results showed that deep learning has the potential to support accurate pneumonia diagnosis with an accuracy of 87.78% achieved by the proposed model.

Testing on an expanded dataset of 5,863 chest X-rays, Shadi A. Aljawarneh et al. [11] trained various deep learning models including Enhanced CNN, VGG-19, ResNet-50. The Enhanced CNN applied with transfer learning and fine tuning techniques on it has highest accuracy of 92.4% and ResNet-50 model achieved the least accuracy of 82.8%.

Overall, all the studies have shown promising results of detecting pneumonia from X-ray images using different variations of a CNN model, however, majority of these papers have achieved accuracies of 89% and below, which is due to the imbalance in the dataset used by all the papers. Although the accuracy score is good, it can be improved by addressing the imbalance in the dataset.

III. Methodology

This section outlines the methodological framework used for the development and evaluation of the proposed pneumonia detection system. The section is structured into four important parts discussing the dataset utilized, the preprocessing utilized on the dataset, the convolutional neural network (CNN) architecture used, and the metrics used to test the model performance.

A. Dataset

The dataset used for training and testing the CNN model is based on pediatric chest X-ray images. This dataset was sourced from the publicly available and widely known Guangzhou Women and Children's Medical Center dataset found on Kaggle [12]. There are a total of 5,863 anterior–posterior chest X-ray images from pediatric patients within the age range of one to five years. The two diagnostic categories are labeled NORMAL and PNEUMONIA; hence, supervised learning approaches for binary classification can be implemented on the dataset. A summary of the dataset is provided in table I.

TABLE I. Description Of The Chest X-ray Dataset Used For Binary Classification Of Pneumonia Presence

Feature	Description
Source	Guangzhou Women and Children's Medical Center
Total Images	5,863 X-ray images (JPEG format)
Age Group	Pediatric patients (1-5 years old)
Image Type	Anterior–Posterior chest X-ray images
Categories	Normal, Pneumonia

This is the same dataset used by all the papers we explored in the literature review section, this is done to properly analyse and compare our CNN model with their attempts, it is also done to solve the dataset imbalance hurdle previous studies faced. The dataset images were divided into training, testing, and validation sets, Fig. 1 visualizes the distribution of images in the training set, we can observe an imbalance where the number of pneumonia images are triple the number of normal images with 3875 pneumonia images to 1341 normal images.

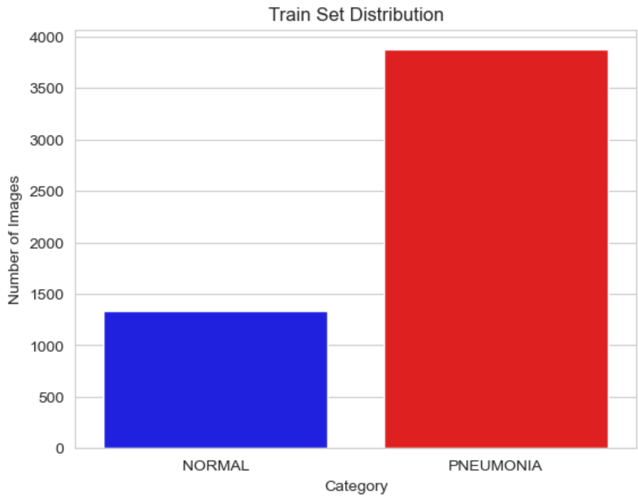

Fig. 1. Training set chest X-ray image distribution.

B. Preprocessing Phase

To prepare the is dataset for training, multiple preprocessing steps were applied to sure data uniformity, as well as help model learning and reduce the computational complexity. Grayscale conversion, image resizing, normalization and data augmentation the main milestones of this section. Each phase is crucial in determining how the input data will be shaped for the convolutional neural network and increasing the model's generalization ability.

Grayscale conversion is the first step of the preprocessing pipeline. As chest X-ray images do not have color channels for conveying important medical features, converting from RGB to grayscale helps reduce the number of channels from three to one. Moreover, it not only lowers computational cost during training but also removes the noise that is not important for discriminating pathological patterns like opacities or fluid accumulation related to pneumonia, allowing the model to concentrate on the variations based on intensity.

All images were then converted to grayscale, and all images were resized to a fixed dimension of 150x150 pixels. The sizes of the original chest X-rays in the dataset were varied, which could possibly result in inconsistencies in model input [13]. We standardized the image size so that the convolutional layers can process inputs of same dimension and the training process goes smoothly, preventing errors caused by inconsistent tensor shapes. The dimension 150x150 strikes a good compromise between keeping the key visual features and keeping the computational load as reasonable as possible.

The pixel intensity values of the images were then normalized to a consistent range using normalization. It enhances the convergence rate of the model during training and avoids the gradient related issues. In particular, pixel values were scaled down to a [0, 1] range. By normalizing the data, this transformation ensures that all input data is on a similar scale, which can stabilize training and improve its efficiency.

To enhance model robustness and prevent overfitting, data augmentation procedures were employed to augment the training images [14]. This method generates altered copies of images found in the dataset using pre-defined transformation functions, thus artificially enlarging the dataset size. Several augmentation techniques were used in this study, as shown in Fig. 2 some training images were randomly rotated up to 30 degrees, others were zoomed in by 20% and horizontal and vertical shifts by 10% were used as these levels of positional variance.

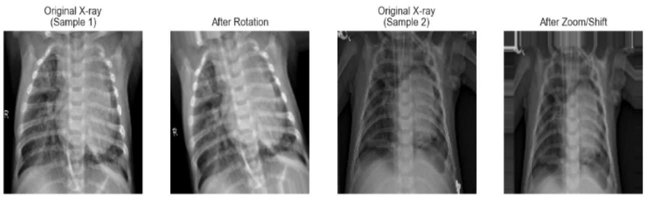

Fig. 2. Dataset changes after applying data augmentation.

Moreover, images were randomly flipped horizontally to add more variance. After defining the augmentation configurations, the data set was fitted to an augmentation generator so that these augments would be applied on the fly as the model learnt. This led to a larger and more diverse training dataset, which enabled the CNN model to better generalize to images it had not seen before.

C. Proposed CNN Model Architecture

A structured and layered architecture of the proposed Convolutional Neural Network (CNN) model for pneumonia detection is followed, to extract and learn complex features from chest X-ray images in an efficient manner [15]. As in the model shown in Fig. 3, we begin with the input layer which accepts grayscale chest X-ray images of size 150x150 pixels. The network then passes through multiple convolutional layers, where it applies learnable filters to detect low to high level features like edges, textures, and shapes that are relevant for pneumonia diagnosis.

After each convolutional layer, there is a max pooling operation performed on the feature maps, down sampling them, reducing spatial dimensions and computational load while preserving the main features [16]. Batch Normalization is added after the convolutional and max pooling stages to stabilize and speed up training and make sure that the network has the constant activation distributions in all layers. Dropout layers are introduced into the network at strategic points to prevent overfitting.

The spatially structured data is flattened into a one dimensional vector after the convolutional and max pooling layers are applied to the output. The learned features are further interpreted by these dense layers and used to make the final classification. The sigmoid activation function is used on the final dense layer to output a binary prediction of whether the chest X-ray image is normal or is infected with pneumonia. The CNN can learn robust patterns from the radiographic images and robustly make reliable diagnostic predictions in this sequential and hierarchical model design.

To ensure optimal performance of our CNN model, a set of hyperparameters were used throughout the training process. These hyperparameters influenced the model's generalization and classification capability. Table II provides a summary of these hyperparameters.

TABLE II. CORE HYPERPARAMETERS USED TO TRAIN THE PROPOSED CNN MODEL

Hyperparameter	Value	Rationale
Optimizer	Adam	Efficient for CNNs in medical imaging
Learning Rate	0.0001	Chosen through experimentation for stable convergence
Batch Size	32	Balanced memory usage
Epochs	20	To ensure convergence without overfitting
Activation Function	ReLu and Sigmoid	To prevent vanishing and for binary output
Kernal Size	3x3	Standard choice to capture local features

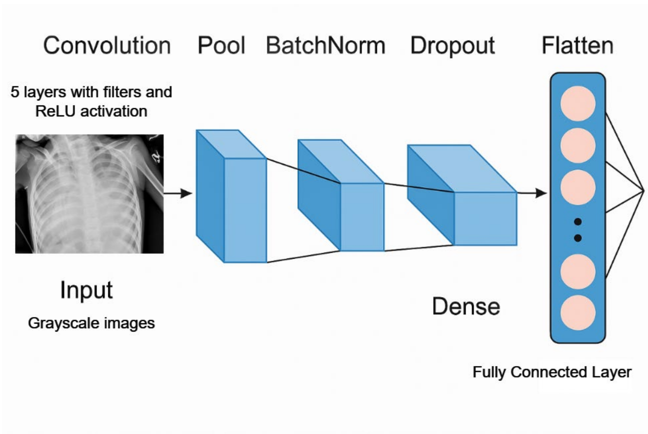

Fig. 3. Proposed CNN model architecture.

D. Evaluation Metrics

To thoroughly assess the proposed CNN model for pneumonia detection, accuracy, precision, recall and F1-score were used as evaluation metrics [17]. The accuracy, as in equation (1), is the overall correctness of the model which is measured by the ratio of correctly predicted instances to the total number of predictions made by the model. Nevertheless, in cases of class imbalance, just accuracy may not give a complete picture. Therefore, precision and recall were also used.

Precision, as in equation (2), is the proportion of true positive predictions over all positive predictions the model made, that is, how many of the model's predicted pneumonia cases were actually correct. On the other hand, recall, as in equation (3), measures the model's ability to identify all actual positive cases, meaning how well pneumonia cases were detected by the system. Finally, F1-score, as in equation (4) will help measure the balance between precision and recall. True Positive (TP), False Positive (FP), True Negative (TN) and False Negative (FN) are represented in equations (1) to (3) respectively.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \tag{1}$$

$$Precision = \frac{TP}{TP+FP} \tag{2}$$

$$Recall = \frac{TP}{TP+FN} \tag{3}$$

$$F1 = 2 \times \frac{Precision \times \text{Recall}}{Precision + \text{Recall}} \tag{4}$$

Apart from these metrics, confusion matrix was used to visualize and quantify the model performance based on true positives, false positives, true negatives and false negatives. This matrix also helps in understanding the nature of errors made by the classifier and class specific performance [18]. In addition, it was plotted using Training vs Validation Curves and analyzed to understand how a model learned over epochs. In the training process, these curves reflect the changes in accuracy and loss for both training and validation datasets, so as to detect such problems as overfitting or underfitting. The results from this curve will identify whether our data augmentation has helped in addressing the dataset imbalance.

IV. Results and Discussion

After testing the model for 20 epochs, the classification report gave very promising results as the overall accuracy of the model was 93%. The precision was also 91%, indicating a strong ability to correctly identify positive cases of pneumonia with minimal false positives. Additionally, the model had a recall value of 96%, showing the model's effectiveness in detecting real pneumonia cases. The F1-score was 94% which indicated a robust and well generalized performance in handling both classes. Fig. 4 showcases correctly predicted chest X-ray images from the test-set, the first row shows normal cases, while the second row shows pneumonia cases.

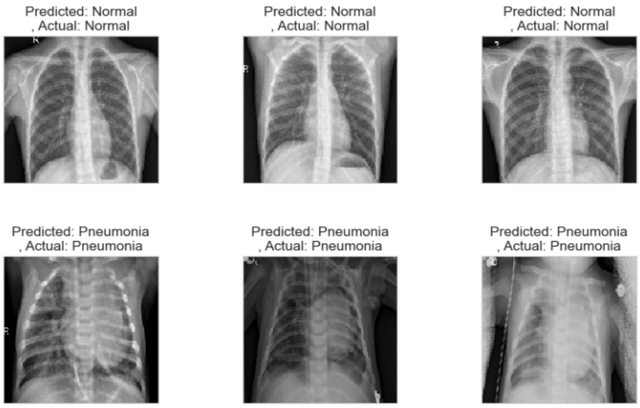

Fig. 4. Correct predictions of normal and pneumonia test cases.

Confusion matrix provides further information on the model's predictive behavior. As in Fig. 5 among all predictions, 33 predictions were misclassified as pneumonia and 201 normal cases were identified correctly. In terms of the pneumonia side, only 13 images were wrongly labeled as normal and 377 were correctly diagnosed. It is a favorable distribution in medical diagnosis, where missing a pneumonia case is more dangerous than over diagnosing. Data augmentation thus substantially aided in closing the dataset imbalance and improving the classification confidence.

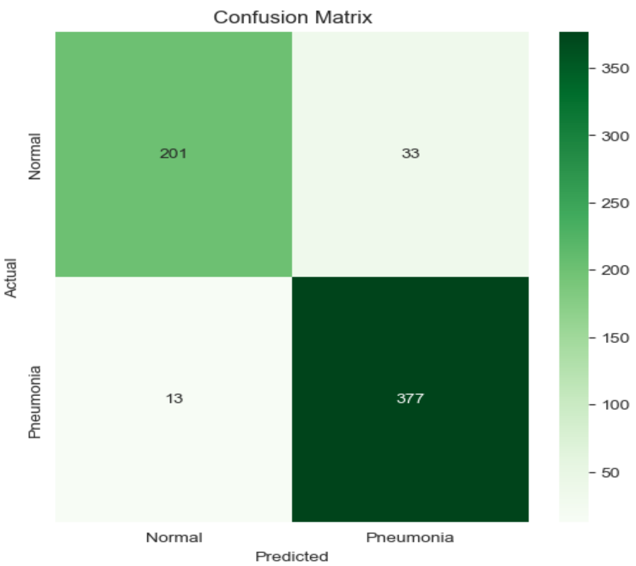

Fig. 5. CNN model confusion matrix.

The training and validation accuracy curves also showcase the benefits of data augmentation, as Fig. 6 revealed a consistent gradual increase throughout the epochs, signifying that the model was learning effectively without overfitting. Moreover; during training and validation loss values, it also showed steady decline of the model and it was adapting well to the dataset and generalized well for unseen data. These patterns are indicative of healthy convergence and stability in the training process which verify the reliability of the proposed CNN architecture.

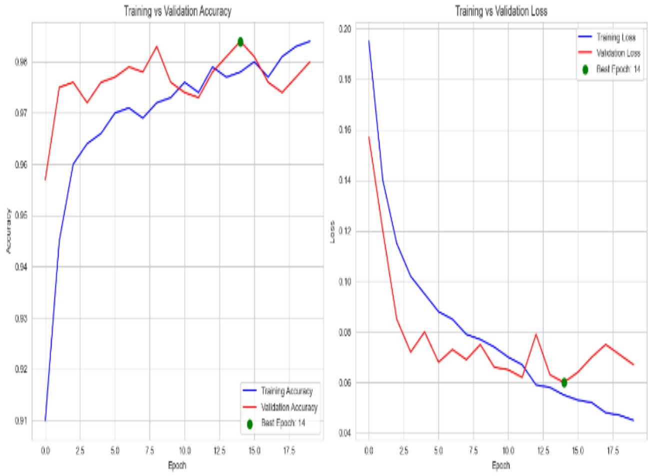

Fig. 6. Training and validation accuracy/loss curves.

A. Discussion

Overall results show that the proposed CNN model is reliable and robust in the automated detection of pneumonia from chest X-ray images, as evaluated with all the metrics. The model performed with an accuracy of 93%, precision of 91%, recall of 96% and F1-score of 94%, which is a strong and consistent classification performance. These results are further supported by the confusion matrix which shows 201 correctly classified normal images, 377 correctly classified pneumonia cases, 33 false positives and only 13 false negatives.

The significance of this low number of false negatives is especially important in medical diagnosis, as overlooking a pneumonia case can be very dangerous. The performance of this study is superior to that of all the reviewed models, having the highest classification accuracy of 93%, which is higher than best accuracies reported in previous literature ranging from 72.9% to 92.63%. It therefore places this proposed model as a superior solution in the domain of automated pneumonia detection.

The simplicity, efficient convergence, and generalization through a carefully tuned architecture and comprehensive data augmentation strategies make the model's advantage. Nevertheless, there are still limits, for instance, dependency on labeled datasets and the difficulty of generalizing to a variety of imaging conditions or patient demographics which are not present in the dataset. This can be addressed by extending to multi-class classification of other thoracic diseases to further extend its clinical utility.

The results generally validate the model's ability to be deployed in real time in medical settings to facilitate faster and more accurate pneumonia diagnosis. Table III presents a comparative analysis between the literature work and our CNN model.

TABLE III. COMPARATIVE ANALYSIS OF PNEUMONIA DETECTION ACROSS PAST STUDIES

#	Author	Model Used	Results (%)
[6]	Yuvraj Sinha	CNN with 1-5 hidden layers	91% (Best Layer)
[7]	Sirish Kaushik	Two CNN Layers	85.26% and 92.31%
[8]	Swapnil Singh	CNN	92.63%
[9]	Orlando	VGG16, VGG19, ResNest50, InceptionV3	72.9% (Best Model)
[10]	Mudasir Ali	CNN, Xception, VGG16, ResNet50	87.78% (Best Model)
[11]	Shadi A. Aljawarneh	CNN, VGG19, ResNet50	92.5% (Enhanced CNN)
Proposed CNN Model		**Custom Built CNN Model**	**93%**

V. CONCLUSION AND FUTURE DIRECTIONS

This study successfully demonstrates the effectiveness of Convolutional Neural Network (CNN)-based models in detecting pneumonia from chest X-ray images with high accuracy and reliability. The proposed model was applied to a pediatric dataset of 5,863 images and achieved 93% accuracy, 91% precision, 96% recall and 94% F1-score. The model is shown to be robust in identifying positive and negative cases correctly using these metrics.

Combinations of preprocessing steps including grayscale conversion, normalization and resizing as well as data augmentation techniques were critical in solving the problem of class imbalance as well as generalization. Additionally, the evaluation results showed that the model was compatible with unseen data and was not prone to the use of diagnostic errors for clinical applicability.

Several further improvements can be pursued to increase the impact of this work looking ahead. A promising approach is to broaden the current binary classification to a multi-class model capable of detecting other pulmonary conditions such as tuberculosis (TB) or COVID-19, thereby enabling it to serve as a useful tool in the diagnostic process. Moreover, exploring transfer learning may lead to higher accuracy scores.

Such improvements would allow the model to be more versatile and scalable across many different healthcare settings. In conclusion, the expansion of a larger and more diverse dataset and more sophisticated learning architectures can make this work grow into a complete diagnostic support system for the particular medical decisions made around the world faster and more accurate.

REFERENCES

[1] D. Varshni, K. Thakral, L. Agarwal, R. Nijhawan and A. Mittal, "Pneumonia Detection Using CNN based Feature Extraction," 2019 IEEE International Conference on Electrical, Computer and Communication Technologies (ICECCT), Coimbatore, India, 2019, pp. 1-7, doi: 10.1109/ICECCT.2019.8869364.

[2] C. Cilloniz, "Pneumonia causes 2.5 million deaths around the world each year," Hospital Clínic de Barcelona, Nov. 3, 2021. [Online]. Available: https://www.clinicbarcelona.org/en/news/pneumonia-causes-2-5-million-deaths-around-the-world-each-year

[3] T. Rahman et al., "Transfer Learning with Deep Convolutional Neural Network (CNN) for Pneumonia Detection Using Chest X-ray," Applied Sciences, vol. 10, no. 9, p. 3233, 2020, doi: 10.3390/app10093233.

[4] L. Račić, T. Popović, S. čakić and S. Šandi, "Pneumonia Detection Using Deep Learning Based on Convolutional Neural Network," 2021 25th International Conference on Information Technology (IT), Zabljak, Montenegro, 2021, pp. 1-4, doi: 10.1109/IT51528.2021.9390137.

[5] A. Tilve, S. Nayak, S. Vernekar, D. Turi, P. R. Shetgaonkar and S. Aswale, "Pneumonia Detection Using Deep Learning Approaches," 2020 International Conference on Emerging Trends in Information Technology and Engineering (ic-ETITE), Vellore, India, 2020, pp. 1-8, doi: 10.1109/ic-ETITE47903.2020.152.

[6] Y. S. Chowdhury, R. Dasgupta and S. Nanda, "Analysis of Various Optimizer on CNN model in the Application of Pneumonia Detection," 2021 3rd International Conference on Signal Processing and Communication (ICPSC), Coimbatore, India, 2021, pp. 417-421, doi: 10.1109/ICSPC51351.2021.9451768.

[7] V. Sirish Kaushik, A. Nayyar, G. Kataria, and R. Jain, "Pneumonia Detection Using Convolutional Neural Networks (CNNs)," in Proceedings of First International Conference on Computing, Communications, and Cyber-Security (IC4S 2019), Lecture Notes in Networks and Systems, vol. 121, Singapore: Springer, 2020, doi: 10.1007/978-981-15-3369-3_36.

[8] S. Singh, "Pneumonia Detection using Deep Learning," 2021 4th Biennial International Conference on Nascent Technologies in Engineering (ICNTE), NaviMumbai, India, 2021, pp. 1-6, doi: 10.1109/ICNTE51185.2021.9487731.

[9] O. Iparraguirre-Villanueva et al., "Convolutional Neural Networks with Transfer Learning for Pneumonia Detection," Science and Information Organization, 2022. [Online]. Available: https://hdl.handle.net/20.500.13053/7688

[10] M. Ali et al., "Pneumonia Detection Using Chest Radiographs With Novel EfficientNetV2L Model," IEEE Access, vol. 12, pp. 34691-34707, 2024, doi: 10.1109/ACCESS.2024.3372588.

[11] S. A. Aljawarneh and R. Al-Quraan, "Pneumonia Detection Using Enhanced Convolutional Neural Network Model on Chest X-Ray Images," Big Data, vol. 13, no. 1, pp. 16-29, Feb. 2025, doi: 10.1089/big.2022.0261.

[12] P. Mooney, "Chest X-Ray Images (Pneumonia)," Kaggle, 2018. [Online]. Available: https://www.kaggle.com/datasets/paultimothymooney/chest-xray-pneumonia

[13] M. Habib and M. Okayli, "Evaluating the Sensitivity of Machine Learning Models to Data Preprocessing Technique in Concrete Compressive Strength Estimation," Arab J Sci Eng, vol. 49, pp. 13709–13727, 2024, doi: 10.1007/s13369-024-08776-2.

[14] T. Islam, M. S. Hafiz, J. R. Jim, M. M. Kabir, and M.F. Mridha, "A systematic review of deep learning data augmentation in medical imaging: Recent advances and future research directions," Healthcare Analytics, vol. 5, p. 100340, 2024, doi: 10.1016/j.health.2024.100340.

[15] S. Sharma and K. Guleria, "A systematic literature review on deep learning approaches for pneumonia detection using chest X-ray images," Multimed Tools Appl, vol. 83, pp. 24101–24151, 2024, doi: 10.1007/s11042-023-16419-1.

[16] G. Labhane, R. Pansare, S. Maheshwari, R. Tiwari and A. Shukla, "Detection of Pediatric Pneumonia from Chest X-Ray Images using CNN and Transfer Learning," 2020 3rd International Conference on Emerging Technologies in Computer Engineering: Machine Learning and Internet of Things (ICETCE), Jaipur, India, 2020, pp. 85-92, doi: 10.1109/ICETCE48199.2020.9091755.

[17] Chatterjee, A., Ghosh, S., Ghosh, A. (2024). Modeling Uncertainty by Integrating Label Smoothing and Data Perturbation: An Application to Pneumonia Detection. In: Dehuri, S., Cho, SB., Padhy, V.P., Shanmugam, P., Ghosh, A. (eds) Machine Intelligence, Tools, and Applications. ICMITA 2024. Learning and Analytics in Intelligent Systems, vol 40. Springer, Cham. https://doi.org/10.1007/978-3-031-65392-6_26

[18] A. Ali, Y. Wang, and X. Shi, “Detection of multi-class lung diseases based on customized neural network,” Computational Intelligence, vol. 40, no. 2, p. e12649, 2024, doi: 10.1111/coin.12649.