Hybrid Deep Learning Framework for Pneumonia detection: Integrating VGG19 and Resnet50

Muttha Navya Sri
Dept. of Information Technology
Velagapudi Ramakrishna Siddhartha Engineering College
Vijayawada, India
228w1a12h3@vrsec.ac.in

Abdul Rahaman
Dept. of Information Technology
Velagapudi Ramakrishna Siddhartha Engineering College
Vijayawada, India
228w1a12d1@vrsec.ac.in

Dr. M Ashok Kumar,
Dept. of Information Technology
Velagapudi Ramakrishna Siddhartha Engineering College
Vijayawada, India
ashockkumar.munnagi@gmail.com

Shaik Mastan
Dept. of Information Technology
Velagapudi Ramakrishna Siddhartha Engineering College
Vijayawada, India
228w1a12i8@vrsec.ac.in

Abstract—Pneumonia remains one of the serious global health problems requiring prompt and appropriate diagnosis to ensure effective cure. This project intends to introduce an advanced deep learning solution that detects pneumonia with a hybrid model that integrates the power of both VGG19 and ResNet50. Using VGG19 and combining it with the power of ResNet50 residual learning technique, this diagnostic model has given very strong accuracy and generalization. Using augmentation techniques, issues related to poor and imbalanced datasets have been dealt with, bringing about impressive improvement in terms of performance metrics, such as accuracy, recall, and F1-score. The implementation includes an elaborate evaluation process that shows that the model can classify pneumonia case instances, with special emphasis on the improvement of the detection of minority class instances. After integrating advanced augmentation methods, the accuracy of the models improved from 77% to 90%, thus proving the framework's potential and clinical applicability.

Keywords—Pneumonia Detection, Imbalanced Dataset, Deep Learning Framework, Hybrid Architecture, VGG19, ResNet50, Augmentation Strategies.

I. Introduction

Pneumonia has remained one of the major causes of loss of life globally, most predominately in children and the elderly. To this end, correct and timely diagnosis enhances patient care outcome, yet traditional techniques that rely on well- established techniques using chest X-rays are dependent on radiologists and slow and prone to errors.

Thus, to avoid such challenges, this system proposed a deep learning- based hybrid model that relied on the power of CNNs for the automated detection of pneumonia. This project is an integration of two state-of-the-art models, VGG19 and ResNet50, aimed at producing an overall robust system for extracting fine-grained and hierarchical features from chest X-ray images. The integrating of VGG19 and ResNet50 ensures that this excellent feature representation improves the distinction among the normal cases and those affected with pneumonia. Then this system approached the problem of dataset imbalance by using data augmentation techniques, thereby allowed the model to generalize well on the diverse real-world scenarios. This ensured an improvement in accuracy and not tend to overfit. Thus, combining innovative architecture with effective data preprocessing, this work presents some potential AI-driven solutions in healthcare diagnostics.

A. Objectives of Proposed system

This project aims at integrating the strengths of two pre-trained deep learning models that are VGG19 and ResNet50 toward constructing a hybrid framework which excels at pneumonia detection. One of the challenge of having fewer and unbalanced datasets was addressed by advanced data augmentation techniques within this project. This integration ensures that the model is able to capture both low-level features from VGG19 and high-level features from ResNet50.This combination will better performance since their unique capabilities can be utilize to enhance more accurate and reliable diagnostics. The hybrid model ensures precise classifications of different features of images computed, so the model is applicable in medical imaging.

A most important objective of this work is to improve performance in diagnostics through several metrics in terms of accuracy, recall, precision, and F1-score. Another important aspect that project deals with is optimization of the learning process as well as scalability. Use of transfer learning with pre-trained models significantly lowers the computational costs as well as the training time of this model. The design of the model also ensures that it is easily adaptable to scale up when needed.

II. LITERATURE SURVEY

[1] Feras Barneih, Nida Nasir, Afreen Kansal, Omar Alshaltone, Talal Bonny, and Mohammad Al-Shabi proposed a pneumonia detection model using the ResNet50 architecture trained on chest X-ray images. Their work highlights the use of residual learning as a means of feature extraction that enhances diagnosing accuracy and helps a radiologist while making decisions in the clinic. Statistical validation was performed through calculations of various metrics: accuracy, F1-score, recall, and AUC. The proposed model achieved a strong F1-score of 86% as proof of reliability. [2] Anas, Bhavye Gupta, Chirag Mishra, and Puneet Garg employed CNNs in their article on chest X-ray images for a sake of classification tasks. They used data taken from Kaggle: over 17,000 chest X-ray images categorized into two: normal and pneumonia. The researchers used this dataset to train the network, so the model will automatically feature extract and classify the data. The overall accuracy that was obtained for the model for this study was 88.62%. [3] In Kavitha et al. (2023), the authors explored the use of deep learning techniques for the detection of pneumonia from chest X-ray images. The CNN model achieved an accuracy of 89%, while VGG-16 obtained an accuracy of 85%.

[6] The method proposed in Zhang, Guotao, Xin Li, Dingfu Zhang, and Teoh Teik Toe suggests the classification of chest X-ray images using convolutional neural networks. In particular, the final CNN model has achieved training accuracy greater than 85%. [4] According to Vinay Singh et al. (2022) emphasized the use of Convolutional Neural Networks (CNNs) to diagnose pneumonia using chest X-ray pictures. The authors used huge datasets with pre-trained CNN models such as VGG-16, InceptionV3, and CheXNet as feature extractors. [5] Dhyanendra Jain, the researchers designed a tailored Conventional Neutral Networks. The objective of the study was to create and develop an automated system that makes use of deep learning technologies in assisting to diagnose, especially where access to medical facilities would be limited.

Babic et al. [9] proposed a TensorFlow system based on CNN architectures, which could accurately identify chest X-rays with more than 95% accuracy. It reduced pressure on health systems, especially in times of COVID-19 because of the prioritization of cases with higher priority. Mishra et al. [8] enhanced transfer learning by fine- tuning a ResNet model to an accuracy of 91% and demonstrated diagnostic skills. Similarly, Hossain et al. [7] introduced the "ESPD" model; it is a deep learning architecture that was trained based on 5,856 X-ray images, with 98.24% accuracy and with a low computational cost.

III. METHODOLOGY

A. Overview of Dataset

Dataset ‘Chest_X-ray’ from Kaggle with 5,856 chest X-ray images that divided into two categories: Normal (healthy) and Pneumonia (lung-affected). The dataset consists of three subsets: training, testing, and validation.

	Training Set	Testing set	Validation Set
Normal	1341	234	8
Pneumonia	3875	624	8

Table I. Dataset Description before augmentation

From Table I, The Training Set consisted of 5,216 images, including normal and pneumonia images. For this set, there were 1,341 normal images and 3,875 images of pneumonia. For Testing Set includes 624 images, 234 normal, and 390 pneumonia images. For Validation Set 8 images, divided evenly between the two classes.

The Pneumonia class has both bacterial and viral cases. Due to the significant imbalance inherent in the dataset by having much more images of pneumonia in the training set, this could lead to biased predictions. Data balancing techniques were thus applied to address this issue.

B. Data Balancing and Data Preprocessing

As an initial step, this project required balancing the data set because the data in the training set were imbalanced-between normal and pneumonia classes. For this purpose, various data augmentation techniques applied to augment normal images in the training set from 1,341 to 3,200. The applied augmentation techniques are:

Random Rotations: Added random rotations up to 20 degrees to mimic varied angles of chest X-rays for robustness of the model under positional variations.

Horizontal Flips: The horizontal flips generated the flipped version of images, mirroring the real-world variations in the imaging of X-rays.

Zooming: Zoomed in a little from 0.8x to 1.2x on lung features at different scales so the model can help pick those minor patterns.

Brightness adjustment: This made the brightness of the images fall into some controlled levels so that there could be slight variations in imaging conditions, such as X-ray intensity.

Random Cropping and Padding: This ensured the preservation of important lung features while introducing subtle changes to the image structure.

After augmentation, the dataset became heavily balanced with 3,200 normal images and 3,875 pneumonia images in the training set that was what led to a reduced level of biased prediction and even better generalized overall performance in this model. For preprocessing purposes, all images were normalized by scaling the pixel values to the range [0, 1]. All images are resized to 224x224 pixels in order to meet the requirements in our input deep learning model, which improves computational efficiency and compatibility.

VGG19:
Targeted to extract high-level features based on broad spatial and contextual information in images. Its sequential architecture provided a layer-by-layer step-by-step representation of the input images and, therefore covered all the strengths of the ResNet-50 model.

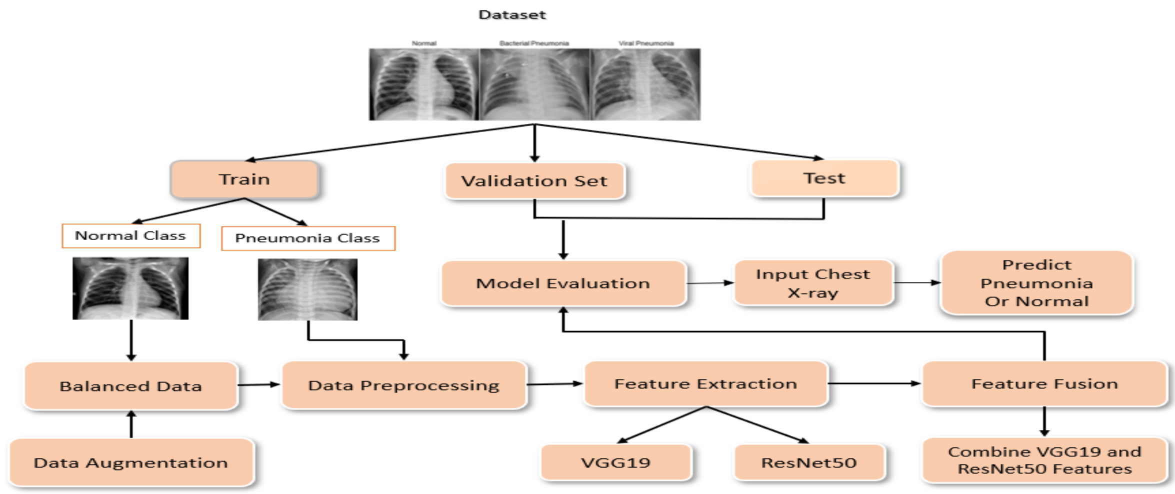

Fig. 1. Architecture Diagram

Overview of dataset after augmentation:
Following augmentation and preprocessing, generated a balanced and augmented dataset that called as the "Chest X-ray Enhanced" to reflect both the augmentation and the improvements made. The new dataset is described as follows:

	Training Set	Testing Set	Validation Set
Normal	3200	234	8
Pneumonia	3875	624	8

Table II. Dataset Description after augmentation

From Table II, In Training Set total 7,075 images, of these, we have 3,200 normal healthy images and 3,875 pneumonia images. The augmentation dataset increases the kind of chest X-ray variability that one would get in real-world deployment, thereby making the model that much more robust and accurate.

C. Model Selection

We used an integrated approach in our project: combining the feature extraction capabilities of ResNet-50 with VGG19. These two powerful pre-trained deep learning models were selected based on their capability to complement each other and improve the detection of pneumonia using chest X-ray images.

1. Feature Extraction:

ResNet-50:
Used to extract finer features that capture the subtle details and intricacies involved in the patterns and chest X-rays. Its connection reserves helped it achieve network training up to deeper layers without degradation and to learn details and complexities of features needed for accurate diagnosis.

2. Feature Fusion:
The feature vectors obtained using ResNet-50 and VGG19 were fused. This also ensured that the strengths of both models are used to create a better representation of the given input images. The ResNet-50 provided low-level, detailed features. VGG19 supplied robust high-level abstractions, thereby enabling coverage of local as well as global information. This combined the two models into a single one, using either individually, making it extremely effective for tasks such as binary classification- identifying "Normal" versus "Pneumonia".

3. Model Strengths
ResNet-50: Deep architectures and residual learning, which is perfect for complex pattern learning.
VGG19: By depth and structure uniformness is where the resulting stability and consistent performance.
This integrated model uses complementary strengths of these two architectures in bringing up an extremely reliable pneumonia detection solution through good feature representations with accurate predictions.

D. Model Architecture

From Figure 1, In this pneumonia detection project, it presents a robust model architecture obtained by fusing the balance of two fully pre-trained deep learning models, namely ResNet50 and VGG19, to classically address the challenge of classifying chest X-ray images into one of two categories: "Normal" and "Pneumonia."
Model Architecture: The data preprocessing phase is the very first phase in the construction of the model architecture. Here, this work balance the dataset using techniques like data augmentation. The class distribution

should be well balanced, and also the model should not favor one class more than the rest. Resizing all the input images to 224 pixels by 224 pixels is the first step. This is actually what was standard input for both VGG19 and ResNet50 models, and it was a necessary step in order to get meaningful features because both models were expecting a fixed size of an image. The images then had uniform dimensions, which in turn will help the models process them much more efficiently. Then, we feed these images to the preprocessed stage of the model to the feature extraction step, where we make use of VGG19 and ResNet50 models to extract features from X-ray images. VGG19 is particularly good at getting high- level features in image understanding. It catches wide understanding. Conversely, ResNet50 focuses more on the fine-grained features that make it possible to discover minor details.

The capturing of features from both VGG19 and ResNet50, feature fusion is carried out that integrates these into one unified vector. This ensures the captured features are of high-level as well as fine-grained so that maximum relevant information is extracted for the purpose of classification. Those fused features are further learned using fully connected layers with ReLU activation functions to learn the complex patterns in data and underlying relationships and inter dependencies within the data. Ultimately, the output of those layers passes through a sigmoid activation function within the output layer that predicts the probability that the chest X-ray image actually belongs to the "Pneumonia" class. It returns a probability score from 0 to 1. The classification of the input image as "Normal" or "Pneumonia" is made based on the received score. This model was optimized for performance over the task of binary classification using Binary Cross-Entropy Loss. For optimization, the Adam optimizer was used, which adjusts learning rates in the course of training in an effort always to improve convergence.

Another feature of improvement of training the model is early stopping. Early stopping prevents overfitting and optimizes the fact that the model should not waste time on extra training when already at its top performance. This saves computational resources and reduces the overfitting possibility from the provided training data. Then the performance of the model will be assessed against range criteria such as accuracy, precision, recall and F1-score, so that this model would not only perform well when being correct for classifying pneumonia cases but also minimize the number of false positives and false negatives.

E. Model Evaluation

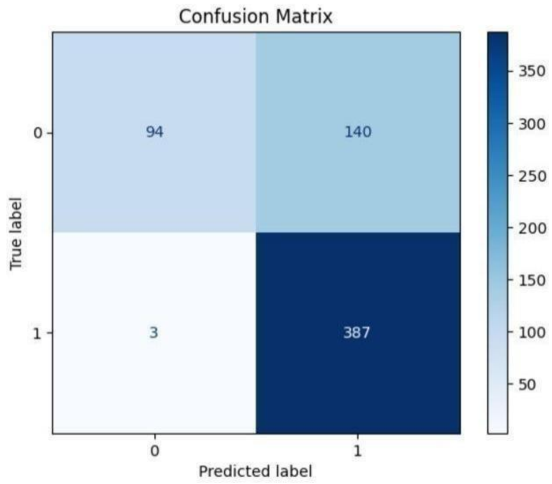

Fig. 2. Confusion Matrix (Before Augmentation)

Metric	Class 0	Class 1	Overall
Precision	0.97	0.73	0.82
Recall	0.40	0.99	0.77
F1-Score	0.57	0.84	0.74
Accuracy	0.77		
Support	234	390	624

Table III. Classification Report (before Augmentation)

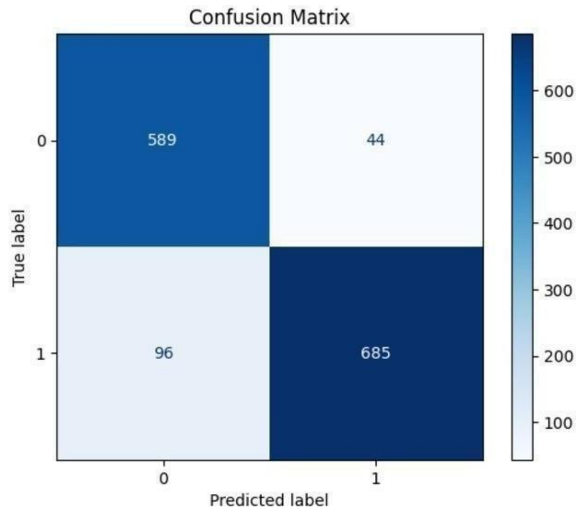

Fig. 3. Confusion Matrix (After augmentation)

Metric	Class 0	Class 1	Overall
Precision	0.86	0.94	0.90
Recall	0.93	0.88	0.90
F1-Score	0.89	0.91	0.90
	0.90		
Support	633	781	1414

Table IV. Classification Report (After Augmentation)

From Figures 3,6 and table III, IV first present an analysis of the performance metrics—Precision, Recall, F1-Score, and Support—for the pneumonia detection model before and after it was augmented. Before augmentation, the model had very high precision for Class 0 (Normal), with a value of 0.97, but its recall was lower at 0.40, with an associated F1-score of 0.57. For Class 1 (Pneumonia), the precision achieved was 0.73, but with excellent recall at 0.99, giving an F1-score of 0.84. After applying data augmentation, the performance has increased heavily on the Normal class. Precision for Class 0 is now 0.86, but its recall is now 0.93, yielding an F1-score of 0.89. For Class 1 (Pneumonia), precision has improved heavily to 0.94, and its recall has slightly dropped to 0.88, giving a score of 0.91 F1. These changes reflect that the augmentation of the dataset helped the model to distinguish better normal images and enhanced its capacity for the detection of pneumonia, ultimately enabling the overall metrics-precision, recall, and F1-score - to improve towards an F1-score of 0.90 after performing the augmentation, hereby reflecting the ability of the model to make correct decisions in the classification of both normal and pneumonia chest X-rays. Also, the support, which is the count of instances in every class, had increased after augmentation, with the total support increased to 624 to 1414 images. Conclusion of Data Augmentation resulted in a balanced dataset, thus model performance increases.

IV. RESULT ANALYSIS

Training and Validation graphs before augmentation:

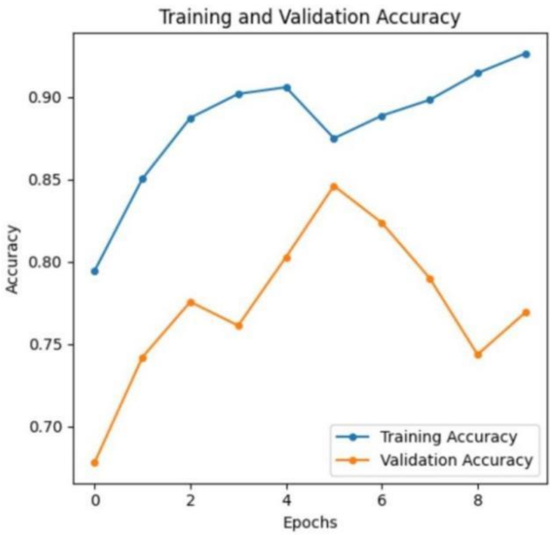

Fig. 4. Training and Validation Accuracy

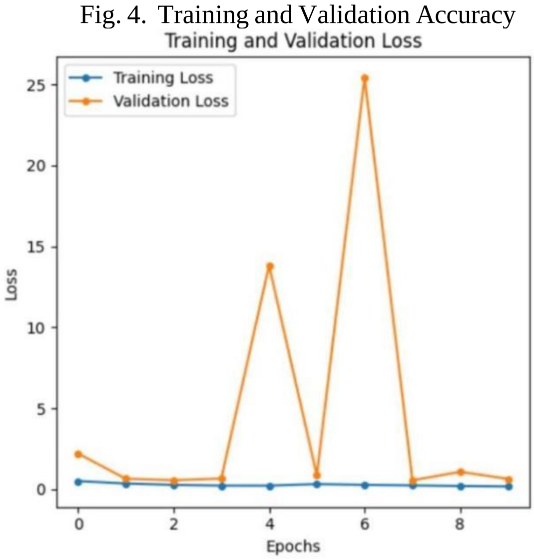

Fig. 5. Training and Validation Loss

From figure 4 and 5 The model's improvement before data augmentation in terms of training accuracy is quite good, about 90%, while its validation accuracy lags at about 80%. This presentation suggests some overfitting-the model does very well on the training data but is challenged to generalize to the validation set. Training loss is almost steadily dropping, pointing to efficient learning; validation loss appears unstable with spikes, further pointing to challenges in generalization. Data augmentation may be used to enhance the generalizability of the model by preventing overfitting in unseen data, model's total accuracy before augmentation is 77%.

Training and Validation graphs after augmentation:

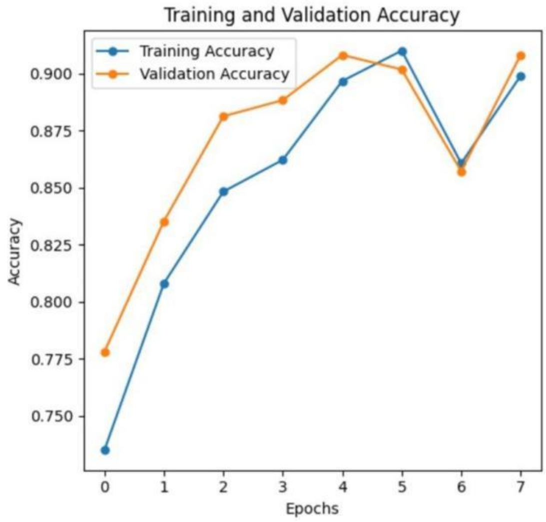

Fig. 6. Training and Validation Accuracy

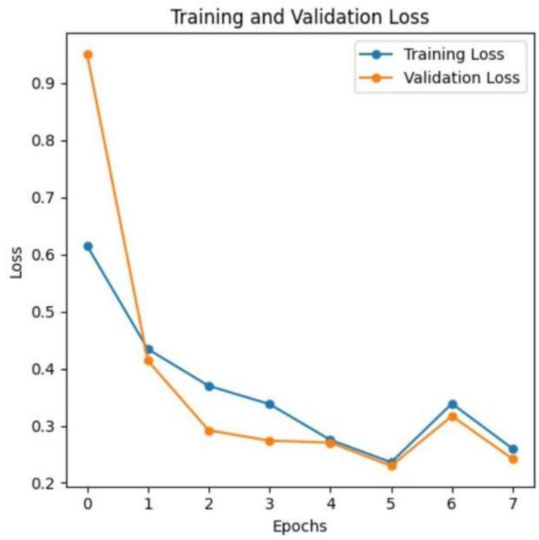

Fig. 7. Training and Validation Loss

As can be seen in Figure 6 and 7, applying data augmentation leads to a significant improvement both in training as well as in validation accuracy, reducing the validation loss. Training accuracy nears 90%, while the validation accuracy is elevated into a similar range, indicating better generalization. It indicates the model does not overfit and improves a lot on unseen data. Both the training loss and validation loss decrease steadily. So, data augmentation improves the process of the model and reduces the risk of overfitting ends, model's total accuracy after augmentation is 90%.

V. Conclusion

The combination of augmentation techniques and advanced deep learning models, VGG19 and ResNet50, has proven to be a game-changer in improving pneumonia detection accuracy and Augmentation helped to reduce overfitting and improves both training as well as validation performance ensuring the model learn the unseen data set. It further enhanced robustness using augmentation so the models accurately identified normal, as well as pneumonia, cases. Additional of VGG19 and ResNet50 added power to this project since those models combined distinct architectural advantages of the designs of themselves. VGG19 gave hierarchical feature extraction, and skip connections in ResNet50 allows the model to learn deep features. Together, augmentation and advanced model integration have allowed to create a comprehensive and increase in performance for pneumonia detection system.

References

[1] F. Barneih, N. Nasir, A. Kansal, O. Alshaltone, T. Bonny, and M. Al-Shabi, "Pneumonia Detection in Chest X-ray Images using ResNet50 Model," in Proc. 2023 Adv. Sci. Eng. Technol. Int. Conf. (ASET), Dubai, UAE, Feb. 20-23, 2023. https://ieeexplore.ieee.org/document/10180737

[2] A. Anas, B. Gupta, C. Mishra, and P. Garg, "Pneumonia detection using convolutional neural network," 2023 International Conference on Computational Intelligence, Communication Technology and Networking (CICTN), 2023. https://ieeexplore.ieee.org/document/10141492

[3] M. Kavitha, R. Srinivsan, K. Triveni, and C. Pavan Chowdary, "Pneumonia Detection Using Deep Learning Techniques," *Proceedings of the IEEE*, 2023. https://ieeexplore.ieee.org/document/10369493

[4] Vinay Singh, Aditya Manay, Raj Prakash Singh, Shaurya Tomer, Chandan A, Narayana Darapaneni, "Medical Radiology Image Processing for Pneumonia Detection Using Convolutional Neural Network," *Proceedings of the 2022 International Conference on Medical Imaging and Healthcare Technologies*, 2022. https://ieeexplore.ieee.org/document/10099292

[5] Dhyanendra Jain, Prashant Singh, Vaishnavi Rohatgi, Rohit Raj, and Akshita Sharma, "Pneumonia Detection using Customized CNN," *Proceedings of the 2022 4th International Conference on Advances in Computing, Communication Control and Networking (ICAC3N)*, 2022. https://ieeexplore.ieee.org/document/10074473

[6] G. Zhang, X. Li, D. Zhang, and T. T. Toe, "Pneumonia detection method based on convolutional neural network,"2023 6th International Conference on Artificial Intelligence and Big Data (ICAIBD), 2023.and Big Data (ICAIBD), 2023 Artificial Intelligence and Machine Learning Applications (AIMLA), IEEE. https://doi.org/10.1109/AIMLA59606.2024.10531467

[7] Hossain, M. Z., Mostafa, K. I., Rahman, M. M., & Mia,M. S. Pneumonia Detection from Chest X-Ray Images Using Convolutional Neural Network. Published in: 2024 International Conference on Advances in Computing, Communication, Electrical, and Smart Systems (iCACCESS), IEEE. https://doi.org/10.1109/iCACCESS61735.2024.10499549

[8] Mishra, S., Hazra, A., & Prakash, U. M. Pneumonia Detection Using Deep Learning. Published in: 2022 2nd International Conference on Advanced Computing and Innovative Technologies in Engineering (ICACITE), IEEE. https://doi.org/10.1109/ICACITE53722.2022.9823625

[9] Babic, D., Jovovic, I., Popovic, T., et al. Detecting Pneumonia with TensorFlow and Convolutional Neural Networks. Published in: 2022 IEEE International Conference on Omni-layer Intelligent Systems (COINS), IEEE. https://doi.org/10.1109/COINS54846.2022.9854948