Pneumonia Detection using Deep Learning Techniques

Himashaili Donavalli
dept. of Computer Science
Middle Tennessee State University
Murfreesboro, USA
hd3q@mtmail.mtsu.edu

***Abstract*—An acute respiratory infection that affects the lungs, pneumonia is a major global health concern. It is still a major cause of death, which highlights how important early detection is to lowering the death rate. Three convolutional neural network (CNN) models were used in this study to identify pneumonia from chest X-ray images: VGG16, VGG19, and ResNet50. Using numerous convolutional layers, the models were trained to classify images into three categories: normal, viral pneumonia, and bacterial pneumonia. Evaluation metrics including F1-Score, recall, precision, and accuracy were used. The results showed that the ResNet50 model performed better than the other models, with an accuracy of 91.65These results demonstrate CNN models' effectiveness in correctly identifying pneumonia and point to their potential influence on enhancing diagnostic results.**

***Index Terms*—Pneumonia, Convolutional Neural Networks (CNN), ResNet50 Accuracy**

I. INTRODUCTION

Rapid advances in computational techniques are required to improve diagnostic precision since pneumonia detection confronts significant obstacles in receiving timely and accurate diagnoses. Modern Convolutional Neural Networks (CNNs), namely VGG16, VGG19, and ResNet50, are used in this study to classify pneumonia in chest X-ray pictures, thereby addressing contemporary problems. Viral pneumonia, bacterial pneumonia, and normal patients are all represented in the dataset that was obtained via Kaggle, which guarantees strong model adaptation to a range of clinical situations.

The approach begins with an image processing pipeline that is quite detailed and uses methods like Gaussian blur, median blur, Sobel filtering, and HSV color space transformation to highlight important elements in medical images. To enable efficient model learning and evaluation, the dataset is carefully pre-processed using a 75%–25% training-testing split. Image data generators are developed using the Keras library to optimize training and testing procedures, hence increasing the effectiveness of feeding data into the neural network.

The core of the methodology is the application of VGG16, VGG19, and ResNet50 architectures, which are specifically designed to extract hierarchical features that are essential for the classification of pneumonia. To overcome the difficulties of training deep networks, ResNet50 presents residual learning. To shape the learning process, the neural network models are created using the 'accuracy' measure, 'categorical crossentropy' loss function, and 'adam' optimizer. The models that have been trained undergo extensive testing using fresh chest X-ray pictures to assess their capacity for generalization and usefulness in the identification of pneumonia. By developing more precise and effective diagnostic instruments, this research hopes to enhance respiratory health outcomes in the long run.

II. RELATED WORK

The issue of using neural networks for pneumonia diagnosis has been the subject of research publications published by medical professionals and academic scholars.

An automatic feature extraction function is a feature of CNNs. At the moment, radiographic image analysis—which aims to help detect signs and risk factors of the Covid-19 virus early on—is the preferred field for sick diagnosis. Also, [1] investigated the efficacy of AI in identifying COVID-19 from a collection of photos fast and precisely using several deep learning (DL) models and modifying them to improve detection accuracy of the optimal methods. They also assessed the effectiveness of CNN architectures for classifying X-ray images in [2], coming to the conclusion that the technique known as TL yields the best results for the identification of different abnormalities. Parallel to this, they suggested in [3] a DL method combining the pre-trained VGG19 model with CNN, TL, and object detection. They used 5273 samples that had pneumonia and 1583 samples that were healthy for this. With over 99% accuracy, the suggested model performed well.

In order to accurately identify pneumonia from chest pictures, the authors of [4] introduced a CNN model. At an accuracy of 89.67%, the experiment was conducted using chest X-ray pictures. In the same way, in [5][6], using 5247 chest X-ray pictures, they created work to automatically identify bacterial pneumonia using the TL. Achieving classification accuracy of 98%, 95%, and 93%, respectively, by grouping pneumonia cases into three categories: viral, bacterial, and normal. In the same way, they utilized TL methods with CNN to create four models in [7]: CNN, VGG16, VGG19, and InceptionV3. There were 2972 pneumonia cases and 9992 normal chest radiographs used. Over 97% accuracy was achieved in all four models when testing the results using 854 pneumonia and 849 normal chest pictures.Moreover, in [8] four pre-trained models (VGG19, DenseNet121, Xception, and ResNet50) were used in an automated technique to CNN for the identification of pneumonia. 83.0% or more accuracy was produced by the four models' combined performance. In addition, many CNNs were trained in [9] to distinguish between two categories of X-ray images: pneumonia and normal. Furthermore, a learning framework incorporating residual thinking and convolution was developed in [10] for the diagnosis of pediatric pneumonia.

III. Methodology

The cornerstone of constructing a CNN lies in convolution, which comprises multiple layers of convolutional filters. These filters are complemented by activation functions and optimizers to enhance overall performance. The workflow unfolds through various stages, commencing with the importation of the dataset from Kaggle, followed by dataset processing. Subsequently, the dataset undergoes training using classification models such as VGG16, VGG19, and ResNet50. The detailed steps of this process are elaborated in the subsequent sections, as depicted in Figure 1.

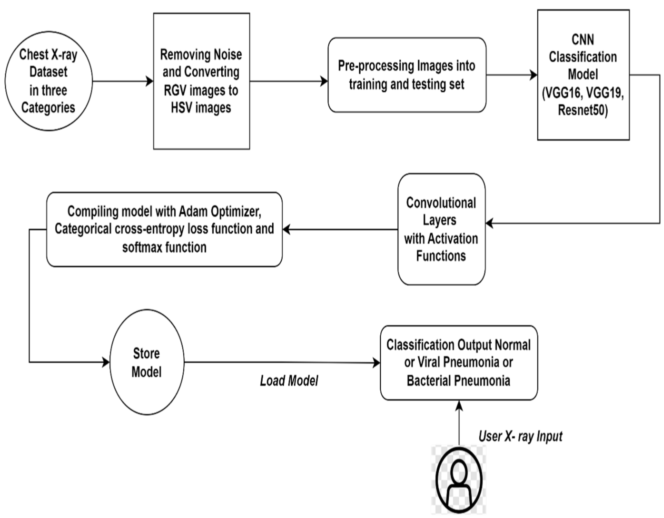

Fig. 1. Proposed workflow.

A. Dataset

The dataset, sourced from Kaggle, encompasses images categorized into three groups: normal, viral pneumonia, and bacterial pneumonia. Specifically, there are 1990 normal images, 1980 bacterial pneumonia images, and 1343 viral pneumonia images within this dataset. Notably, all these images are in grayscale, representing a monochromatic spectrum without the use of color channels. Shown in Figure 2

B. Image Processing

In this image processing pipeline, various filters are applied to enhance the features of a medical image depicting lung opacity. The process begins with the application of Gaussian blur to reduce noise, followed by median blur to further smooth the image. The subsequent step involves the use of a Sobel filter to highlight edges, contributing to the extraction of meaningful structures. Additionally, the image is transformed

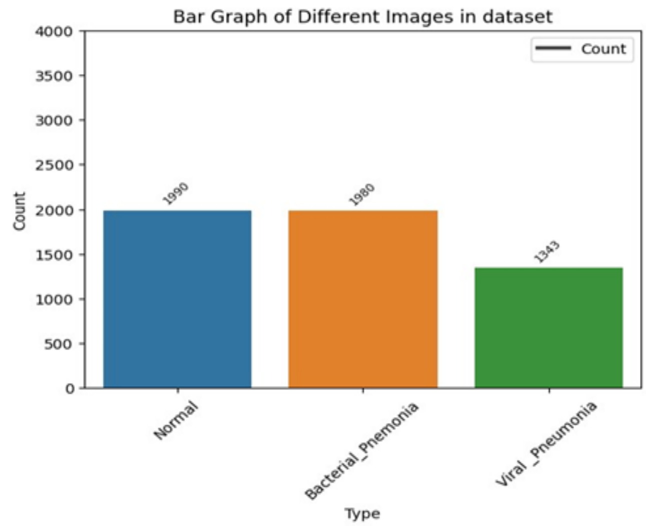

Fig. 2. Count of Images in Dataset.

into the HSV color space after median blur, providing a different perspective for analysis. This comprehensive approach aims to improve the visibility of relevant details in medical imaging for accurate diagnosis and interpretation. Shown in Figure 3

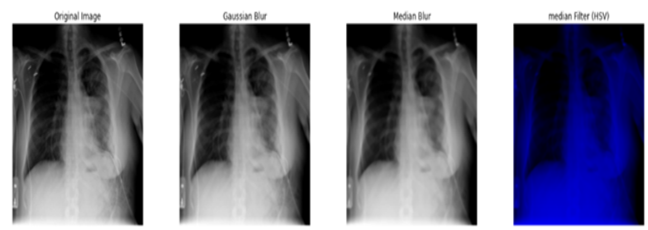

Fig. 3. Processing Image through Filters.

C. Pre-processing Image

In my research, I divided the data into a 75% training set and a 25% testing set using scikit-learn. This ensured that the computer model learned from most of the data and was then evaluated on unseen information to gauge its real-world performance. Shown in Figure 4.

Fig. 4. Different Categories of Image.

Next utilized Keras library for creating image data generators to facilitate efficient training and testing of a neural network. The ImageDataGenerator is configured with a preprocessing function and a validation split of 20% . It is then applied to create generators for training, validation, and testing datasets, specifying parameters such as file paths, labels, target size, class mode, batch size, and shuffling preferences. These generators streamline the process of feeding image data into the neural network during model training and evaluation.

D. Applying CNN model

1) VGG16: VGG16 is a renowned convolutional neural network (CNN) design recognized for its straightforward yet powerful approach to image classification. Originating from the Visual Graphics Group at the University of Oxford, the VGG16 architecture comprises 16 layers, featuring 13 convolutional layers and three fully connected layers. Its simplicity, characterized by the consistent use of 3x3 convolutional filters and max-pooling layers, has made VGG16 a widely adopted model in the field of computer vision for diverse applications. Shown in Figure 5.

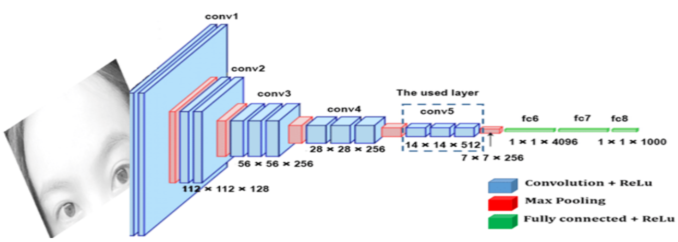

Fig. 5. VGG16 Model Architecture.

2) VGG19: VGG19, an extension of VGG16, comprises 19 layers, including 16 convolutional and 3 fully connected layers. Like VGG16, it employs 3x3 convolutional filters with a stride of 1 and max-pooling layers with 2x2 filters and a stride of 2. The convolutional layers use 'same' padding, maintaining spatial resolution. Rectified linear units (ReLU) serve as activation functions throughout the network. The final layers are fully connected, typically with softmax activation for classification tasks. VGG19's increased depth allows it to capture more complex features, making it suitable for tasks requiring a deeper understanding of visual data. Shown in Figure 6.

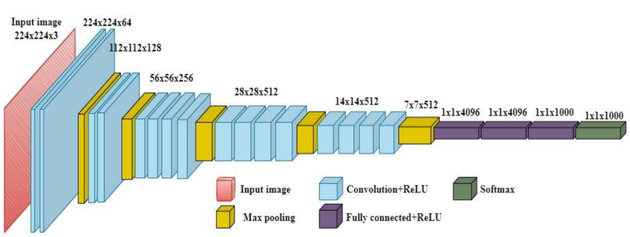

Fig. 6. VGG19 Model Architecture.

3) ResNet50: ResNet-50 is a deep convolutional neural network with 50 layers. It introduces residual learning through skip connections, aiding the training of very deep networks. The architecture includes 3x3 and 1x1 convolutions, bottleneck structures for efficiency, and ends with global average pooling and a softmax-activated fully connected layer. ResNet-50 is known for its success in image classification and computer vision tasks. Shown in Figure 7.

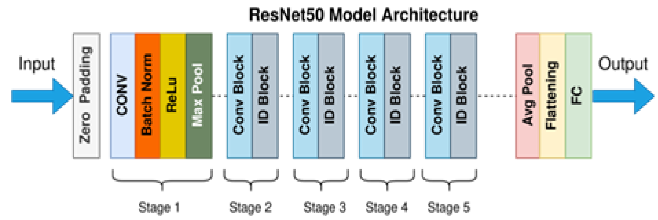

Fig. 7. ResNet50 Model Architecture.

E. Convolution layers with activation functions

All three CNN models utilizes a series of convolutional layers to extract hierarchical features from input images. Residual blocks within the architecture, featuring skip connections, mitigate the vanishing gradient problem and facilitate more effective gradient flow during training. Subsequent to these convolutional layers, custom dense layers are added, commencing with a global average pooling layer to summarize each feature map globally. Rectified Linear Unit (ReLU) activation functions are applied to these dense layers, introducing non-linearity crucial for capturing intricate patterns in the data. The output layer employs the softmax activation function, tailored for multi-class classification tasks, to convert model outputs into probability scores for each class. This comprehensive architecture, combining convolutional layers, ReLU activation, and dense layers, enables the model to discern and classify intricate patterns, rendering it adept for image classification endeavors.

F. Compiling Model

The research configures the compilation of a neural network model using the Keras library. The 'adam' optimizer is chosen for efficient weight updates during training, 'categorical crossentropy' serves as the loss function to measure dissimilarity between predicted and true class distributions in multi-class classification, and 'accuracy' is specified as the metric for evaluating the model's performance. This compilation step is crucial in preparing the model for training, influencing its learning process and enabling effective classification on unseen data by minimizing the defined loss.

G. Classification of image

Once trained, the model's architecture and learned weights can be saved to disk for future use. In a testing phase, the saved model is loaded, and X-ray images are input for classification. The model utilizes its learned features to predict whether the input X-ray depicts a normal condition, viral pneumonia, or bacterial pneumonia. This classification step is pivotal in assessing the model's generalization to new, unseen data and evaluating its practical utility in medical image analysis.

IV. Results and Discussions

In this segment, we showcase the outcomes of training and validating three models using a dataset comprising 5,313 chest X-ray images, both with and without pneumonia. The organization is as follows: 75% of the dataset, equivalent to 3,984 chest X-ray images, was allocated for training, while the remaining 25%, totaling 1,329 images, was reserved for model validation. Consistency was maintained across all models—including VGG16, VGG19, and ResNet50—in terms of processing techniques, data quantity, and partition numbers.

This section delves into the evaluation of model performance throughout the epochs, with graphical representations in Figures 8, 9, and 10. The analysis of the models in this study is as follows.

These figures depict a comparison between accuracy and loss, reflecting the error concerning both training and validation data. The error in training data diminishes notably with an increasing number of epochs. Conversely, the validation data's loss exhibits significant fluctuations, deviating considerably from its true value as epochs progress. Regarding accuracy, the blue line suggests that the model's accuracy approaches one with more epochs, in contrast to the accuracy in validation data, which shows variability and occasionally reaches high values.

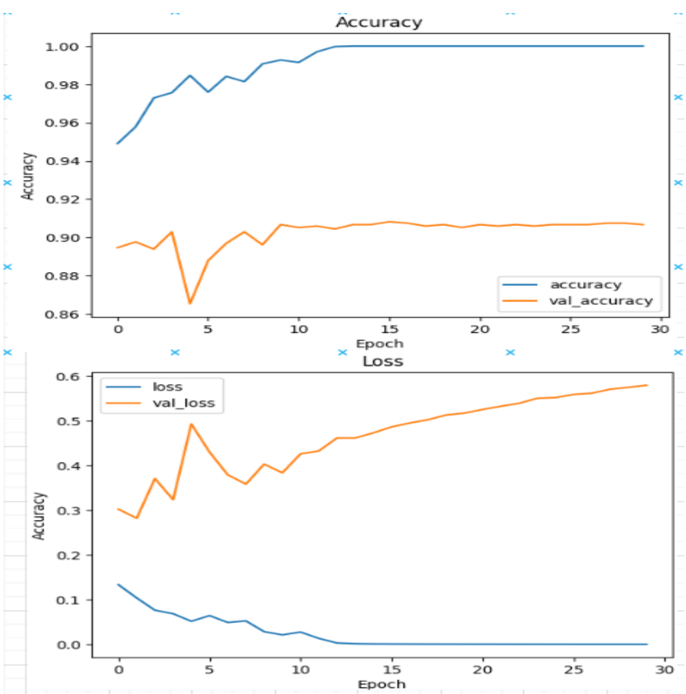

Fig. 8. Accuracy and Loss of the VGG16 Model

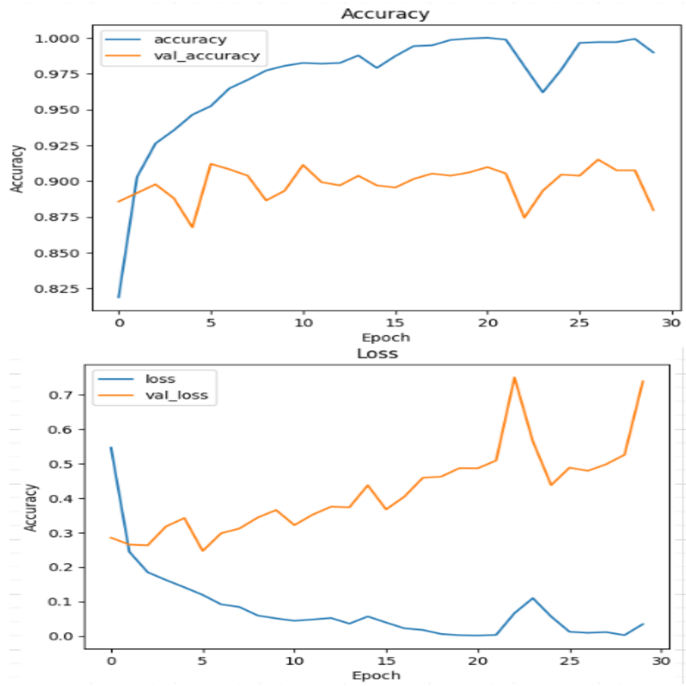

Fig. 9. Accuracy and Loss of the VGG19 Model

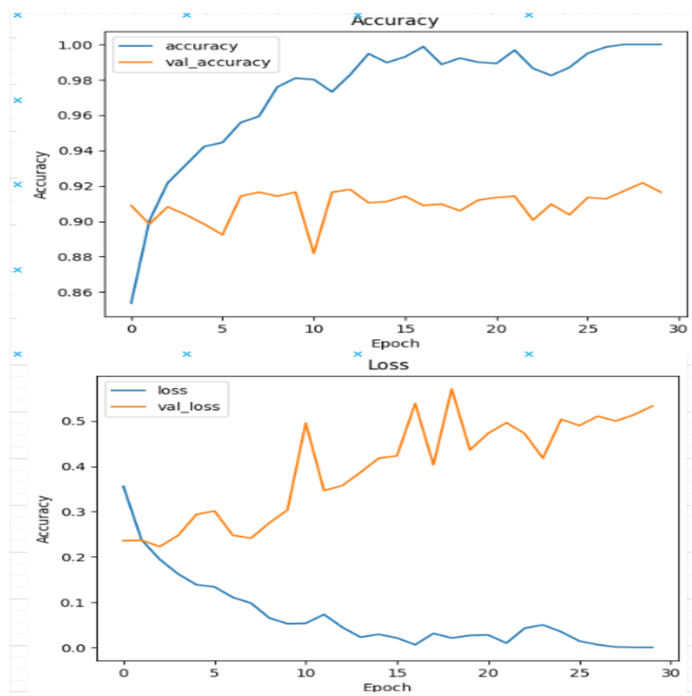

Fig. 10. Accuracy and Loss of the ResNet50 Model

In this research, employing machine learning models like VGG16, VGG19, and ResNet50, predictions are made on a test dataset. The numerical predictions are then mapped to class names, and a comprehensive classification report is generated. This report includes metrics such as precision, recall, F1-score, and support for each class. This analysis provides a detailed and balanced assessment of the performance of these specific deep learning models on the test data. Shown in Figures 11, 12, and 13.

The table includes information on the validation, training, and testing accuracies, as well as the losses, for three different models such as VGG16, VGG19 and ResNEt50. Shown in Figure 14.

Utilizes a pre-trained neural network to categorize a resized

	precision	recall	f1-score	support
Bacterial_Pnemonia	0.36	0.37	0.36	482
Normal	0.36	0.34	0.35	504
Viral _Pneumonia	0.25	0.25	0.25	343
accuracy			0.33	1329
macro avg	0.32	0.32	0.32	1329
weighted avg	0.33	0.33	0.33	1329

Fig. 11. Classification Report of VGG16 Model

	precision	recall	f1-score	support
Bacterial_Pnemonia	0.92	0.91	0.92	501
Normal	0.90	0.78	0.84	488
Viral _Pneumonia	0.81	0.98	0.89	340
accuracy			0.88	1329
macro avg	0.88	0.89	0.88	1329
weighted avg	0.88	0.88	0.88	1329

Fig. 12. Classification Report of VGG19 Model

	precision	recall	f1-score	support
Bacterial_Pnemonia	0.90	0.93	0.92	494
Normal	0.91	0.87	0.89	493
Viral _Pneumonia	0.95	0.96	0.95	342
accuracy			0.92	1329
macro avg	0.92	0.92	0.92	1329
weighted avg	0.92	0.92	0.92	1329

Fig. 13. Classification Report of ResNet50 Model

Model Name	Training Accuracy(%)	Training Loss(%)	Validation Accuracy(%)	Validation Loss(%)	Testing Accuracy(%)	Testing Loss(%)
VGG16	100	12.82	**90.67**	57.93	90.67	57.932
VGG19	98.97	35.0	**87.96**	52.50	87.96	73.796
ResNet50	100	27.51	**91.65**	53.29	91.65	53.286

Fig. 14. Accuracy and Loss percentage comparisons among models

lung image into classes like bacterial pneumonia, normal, or viral pneumonia, and subsequently displays the label with the highest predicted probability. Shown in Figure 15.

```
labels = ["bacterial_pneumonia", "normal", "viral_pneumonia"]

[[99  0  0]]
Predicted Label: bacterial_pneumonia

[[ 0 99  0]]
Predicted Label: normal

[[ 0 41 58]]
Predicted Label: viral_pneumonia
```

Fig. 15. Classifying the Type of Cancer(Output)

V. Conclusion

In the evaluation of various convolutional neural network (CNN) models, including VGG16, VGG19, and ResNet50, distinct patterns in training accuracy and loss metrics emerge. ResNet50 demonstrates the highest training accuracy (100%) and impressive validation accuracy (91.65%), making it the most robust model among VGG16, VGG19, and ResNet50. VGG16 follows closely with a training accuracy of 100% and a validation accuracy of 90.67%. Meanwhile, VGG19 lags behind, with a slightly lower training accuracy (98.97%) and the lowest validation accuracy (87.96%) among the three models. ResNet50 emerges as the best model among the three, showcasing both high training accuracy (100%) and superior validation accuracy (91.65%).

References

[1] M. M. Taresh, N. Zhu, T. A. A. Ali, A. S. Hameed, and M. L. Mutar, "Transfer Learning to Detect COVID-19 Automatically from X-Ray Images Using Convolutional Neural Networks," Int J Biomed Imaging, vol. 2021, 2021, doi: 10.1155/2021/8828404.

[2] I. D. Apostolopoulos and T. A. Mpesiana, "Covid-19: automatic detection from X-ray images utilizing transfer learning with convolutional neural networks," Phys Eng Sci Med, vol. 43, no. 2, pp. 635–640, Jun. 2020, doi: 10.1007/S13246-020-00865-4/TABLES/6.

[3] O. Dahmane, M. Khelifi, M. Beladgham, and I. Kadri, "Pneumonia detection based on transfer learning and a combination of VGG19 and a CNN built from scratch," Indonesian Journal of Electrical Engineering and Computer Science, vol. 24, no. 3, pp. 1469–1480, Dec. 2021, doi: 10.11591/IJEECS.V24.I3.PP1469-1480.

[4] V. Sirish Kaushik, A. Nayyar, G. Kataria, and R. Jain, "Pneumonia Detection Using Convolutional Neural Networks (CNNs)," in Lecture Notes in Networks and Systems, vol. 121, Springer, 2020, pp. 471–483. doi: 10.1007/978-981-15-3369-3-36

[5] T. Rahman et al., "Transfer Learning with Deep Convolutional Neural Network (CNN) for Pneumonia Detection Using Chest X-ray," Applied Sciences 2020, Vol. 10, Page 3233, vol. 10, no. 9, p. 3233, May 2020, doi: 10.3390/APP10093233.

[6] P. Chhikara, P. Singh, P. Gupta, and T. Bhatia, "Deep convolutional neural network with transfer learning for detecting pneumonia on chest x-rays," Advances in Intelligent Systems and Computing, vol. 1064, pp. 155–168, 2020, doi: 10.1007/978-981-15-0339-9-13/COVER

[7] 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," Proceedings of 3rd International Conference on Emerging Technologies in Computer Engineering: Machine

Learning and Internet of Things, ICETCE 2020, pp. 85–92, Feb. 2020, doi: 10.1109/ICETCE48199.2020.9091755.

[8] M. Salehi, R. Mohammadi, H. Ghaffari, N. Sadighi, and R. Reiazi, "Automated detection of pneumonia cases using deep transfer learning with pediatric chest X-ray images," British Journal of Radiology, vol. 94, no. 1121, May 2021, doi: 10.1259/BJR.20201263.

[9] R. Jain, P. Nagrath, G. Kataria, V. Sirish Kaushik, and D. Jude Hemanth, "Pneumonia detection in chest X-ray images using convolutional neural networks and transfer learning," Measurement (Lond), vol. 165, Dec. 2020, doi: 10.1016/j.measurement.2020.108046.

[10] G. Liang and L. Zheng, "A transfer learning method with deep residual network for pediatric pneumonia diagnosis," Comput Methods Programs Biomed, vol. 187, p. 104964, Apr. 2020, doi: 10.1016/J.CMPB.2019.06.023.