Deep Learning in Radiology:A Transfer-Learning Based Approach for the Identification and Classification of COVID-19 and Pneumonia in Chest X-ray Images

Conference Paper · December 2023

DOI: 10.1109/ICSTCEE60504.2023.10585226

Md Rabiul Hasan
Khulna University of Engineering and Technology

Shah Muhammad Azmat Ullah
Khulna University of Engineering and Technology

2023 Fourth International Conference on Smart Technologies in Computing, Electrical and Electronics (ICSTCEE) | 979-8-3503-8462-8/23/$31.00 ©2023 IEEE | DOI: 10.1109/ICSTCEE60504.2023.10585226

Deep Learning in Radiology:A Transfer-Learning Based Approach for the Identification and Classification of COVID-19 and Pneumonia in Chest X-ray Images

Md Rabiul Hasan
Dept. of Electronics and Communication Engineering
Khulna University of Engineering & Technology
Khulna, Bangladesh
mdrabiulhasan7890@gmail.com

Shah Muhammad Azmat Ullah
Dept. of Electronics and Communication Engineering
Khulna University of Engineering & Technology
Khulna, Bangladesh
azmat@ece.Kuet.ac.bd

Mehedi Hasan
Dept. of Electronics and Communication Engineering
Khulna University of Engineering & Technology
Khulna, Bangladesh
mehediece16@gmail.com

Abstract— **Pneumonia, a prevalent infection of the lungs, often arises from bacterial or viral sources. Its impact is notably significant in underdeveloped and developing nations characterized by high pollution levels, unhygienic living conditions, overpopulation, and limited healthcare infrastructure. Detecting pneumonia early is paramount for ensuring effective medical intervention and improving mortality rates. Radiograph of the chest (CXR) examinations represent the most widely used diagnostic way of pneumonia detection. However, interpreting CXRs is challenging and susceptible to subjectivity. Throughout our study, we have devised an automated system for pneumonia detection from CXR images. To address the data scarcity issue, we employed deep transfer learning techniques on VGG16 and ResNet50 architectures. Our proposed approach achieved impressive accuracy rates of 89.23% and 88.80% for VGG16 and ResNet50, respectively. This performance surpasses that of existing methods, underscoring the effectiveness of our approach.**

Keywords— Pneumonia, VGG16, Deep Learning, ResNet50, CNN, COVID-19, Transfer Learning, CNN.

I. INTRODUCTION

Pneumonia has emerged as one of the most perilous diseases globally, responsible for above 4 million untimely fatalities due to atmosphere pollution. Particularly alarming is its status as a main reason of children mortality worldwide, with an alarming death rate of approximately 95% in impoverished nations. In the year 2015, among the 119,000 children who tragically lost their lives across the nation, a staggering 17,850 succumbed to pneumonia in Bangladesh, accounting for 15% of the total child mortality [1]. This translates to a heart-wrenching statistic of approximately 72 children per day (24 hours x 3 children per hour), and an annual total of 24,300 lives claimed by pneumonia in Bangladesh. Notably, in 2016, pneumonia ranked as the fourth leading cause of mortality.

The rapid progress of industrialization and the resulting environmental degradation have heightened the pneumonia risk, particularly affecting developing nations. This situation is compounded by the scarcity of healthcare professionals and limited resources. Detecting pneumonia in its early stages can be challenging due to the inherent oobscured details in X-ray images. While chest X-rays are considered valuable, distinguishing pneumonia can be intricate, as it may be mistaken for conditions like heart failure or various other respiratory ailments. Therefore, we propose an enhanced deep convolutional neural network (CNN) architecture designed to facilitate precise pneumonia diagnosis, with a specific focus on addressing these challenges.

We give the highest importance to the effectiveness and reliability of the algorithm because the accuracy of the medical treatment has a substantial effect on patients' survival rates. To identify COVID-19, non-COVID-19 pneumonia caused by a virus, bacteria, and normal patients using chest X-ray (CXR) images, we suggest employing pre-trained (transfer learning) ResNet50 and VGG16 models in our proposed framework. To obtain the best result, we tuned these models with different hyperparameters. The primary objective of this study revolves around identifying distinctive patterns within CXR images and accurately categorizing them into four distinct classes.

We have outlined our paper as follows: Section I, gives a concise overview of COVID-19 and pneumonia. Section II entails a comprehensive review of models and recommended by various authors. In Section III, we elaborate on our proposed methodology, while Section IV provides a detailed exposition of our dataset. Sections V dives into data preparation while section VI focuses on CNN models. Transfer learning is examined in Section VII. To facilitate the reader's comprehension of our proposed model's performance, Section VIII offers visual representations. Finally, our study concludes after thorough exploration.

II. LITERATURE REVIEW

Recently, a wide range of work focused on the classification of pneumonia has been done. The following provides a brief overview of these studies:

Hasan et al.[2] proposed the utilization of machine learning tools, including LabelBinarizer, for one-hot encoding on labelled chest X-ray images. The transformation into categorical form is achieved using Python's categorical tool. Subsequently, a detection model is constructed employing various deep learning components such as CNN, VGG16, AveragePooling2D, dropout and input. The Adam optimizer is applied to predict pneumonia in COVID-19 patients, yielding a model with an average accuracy of 91.69%, sensitivity of 95.92%, and specificity of 100%. Xin Zhang et al.[3] created DC-Net-S, DC-Net-E, and DC-Net-R for diagnosing COVID-19 pneumonia using a private dataset, achieving an accuracy of 90.91% and a specificity of 96.13%. Similarly, Thakur et al.[4] employed the VGG16 CNN architecture for pneumonia detection , and achieved 90.54% accuracy. Bashar et al.[5] implemented a deep learning approach involving a three-stage process for chest X-ray images. The augmentation of the data comes after the initial stage of image enhancement. In the final stage, the results are input to different CNN models, like VGG16, AlexNet, and GoogleNet, for image classification and diagnosis. Extensive experiments were conducted under diverse scenarios, resulting in the highest classification accuracy of 95.63%. This achievement was realized through the application of the VGG16 transfer learning algorithm on the enhanced augmented dataset with freeze weights. Rajpurkar et al.[6] suggested the 121 layer DenseNet architecture for pneumonia identification, achieving an F1-score of 76.8%. They acknowledged that the lack of patient history significantly impacted their deep learning model's performance. Omar et al.[7] proposed a simplified CNN model featuring five convolutional layers for pneumonia detection. They divided the dataset into training (90%) and testing (10%) subsets, obtaining an accuracy of 87.65%. Hasan et al.[8] also pursued pneumonia classification, employing a customized CNN architecture that demonstrated enhanced performance[9]. Jakub Garstka et al.[10] introduced a tailored CNN model that incorporated data augmentation techniques, resulting in an achievement of approximately 85% accuracy in categorical classification. Additionally, the model demonstrated a sensitivity level of 0.95. Han et al.[11] utilized pre-trained VGG16, VGG19, and DenseNet169 networks for feature extraction and model training. Their results demonstrated an accuracy of 85%, with precision and recall values of approximately 83.14% and 83.05%, respectively.

III. Proposed Method

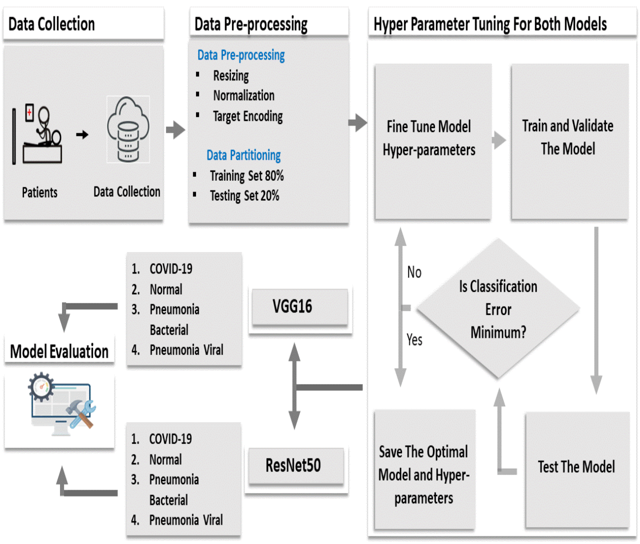

Fig. 1. The conceptual diagram of the proposed model.

The primary goal of this research is to create a system capable of identifying pneumonia in chest X-ray images by leveraging the VGG16 and ResNet50 models. To achieve this, we began by gathering a dataset and then applied preprocessing techniques to prepare the data. Subsequently, we fine-tuned the hyperparameters of these models through a series of iterative experiments, ultimately achieving minimal errors and saving the optimized models. These finely-tuned VGG16 and ResNet50 models were then employed for pneumonia detection. A visual representation of the proposed methodology is illustrated in Figure 1.

A. *Dataset*

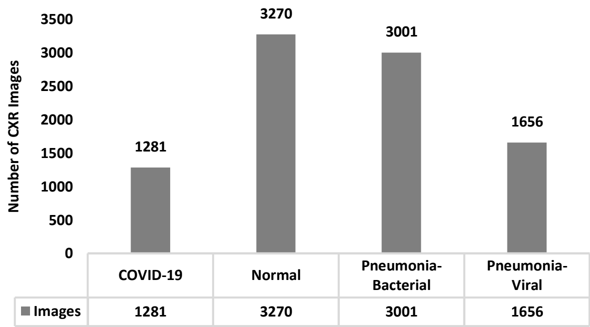

Fig. 2. Dataset summary.

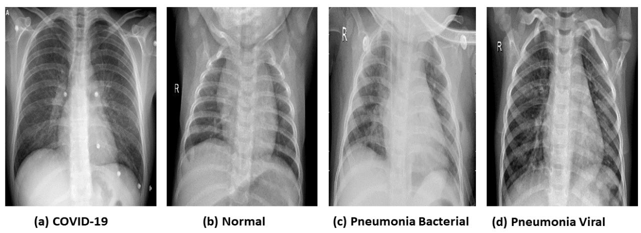

Fig. 3. Chest X-ray images of the dataset.

This compilation of chest X-ray images related to COVID-19 was meticulously curated by amalgamating data from 15 openly available datasets[12]. The present dataset encompasses 3001 X-rays depicting bacterial pneumonia cases, 1656 X-rays illustrating viral pneumonia cases, 1281 X-rays specifically depicting COVID-19 cases, along with 3270 X-rays portraying normal conditions. Additionally, there are 1281 X-rays focused on COVID-19 available within this dataset. Theis valuable resource is now conveniently accessible on Kaggle. The resolutions of the images differ greatly as a result of the various scanning devices and imaging sites.

B. *Image Pre-processing*

Dataset pre-processing holds a pivotal role in this context, as the effectiveness of image pre-processing significantly impacts the outcomes of the classification process.

1) *Resizing*

The images underwent a conversion process from their initial RGB format to grayscale as part of our experimental setup. The database contains images of different dimensions, ranging from 1040 x 664 to 1224 x 1000 and 1848 x 1632 pixels. To facilitate seamless integration into our proposed Convolutional Neural

Network (CNN) models, we resized all images uniformly to a more manageable size of 224 by 224 pixels.

2) Normalization

A fundamental technique in machine learning, known as normalization, involves the division of each pixel value by 255. This process is employed to constrain pixel values to a standardized range between 0 and 1, thereby harmonizing the scaling of features within a dataset. Addressing the challenge of substantial disparities in feature scales, data normalization emerges as a valuable strategy that expedites the convergence of machine learning algorithms. By equalizing the influence of each feature within the dataset, normalization contributes to mitigating biases, promoting fairness in the study. Furthermore, normalization enhances the interpretability of machine learning model outcomes by facilitating the comparison of the relative significance of diverse parameters.

3) Target Encoding

In our approach, we transformed the target labels from their original string representations into numerical values. Each category was assigned a corresponding integer label: 0 for COVID-19, 1 for Normal, 2 for Bacterial Pneumonia, and 3 for Viral Pneumonia.

IV. Model Description

A. VGG16

A widely recognized convolutional neural network (CNN) design utilized for image recognition is the Visual Geometry Group (VGG). Two prominent variations of this architecture are VGG16 and VGG19. While VGG19 comprises 19 layers, VGG16 is characterized by 16 layers. The architecture underwent training using the ImageNet dataset, which encompasses over 14 million images classified into 1000 categories. The VGG-16 pre-trained model is frequently employed by scholars for image analysis due to its noteworthy performance metrics, as illustrated in Figure 4.

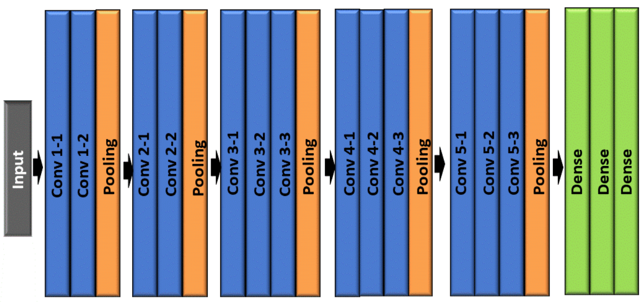

Fig. 4. VGG16 model architecture.

B. ResNet50

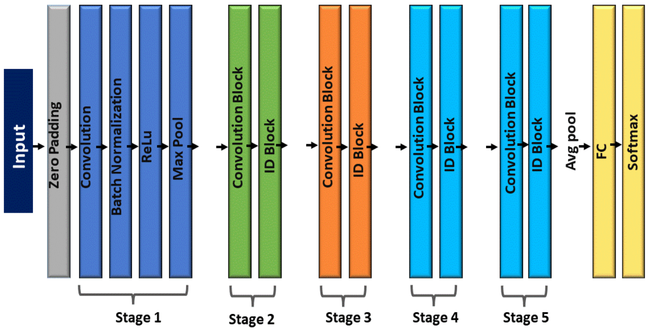

Fig. 5. ResNet50 model architecture.

ResNet-50, a convolutional neural network consisting of 50 layers, plays a crucial role in numerous computer vision tasks. Referred to as Residual Networks (ResNet), it stands as a cornerstone in neural network architecture. Initially introduced in 2015, ResNet has evolved to become a staple not only in research but also in commercial applications. It finds extensive use in image recognition and medical research, where its deep learning capabilities offer potential for the identification and treatment of various diseases. These instances serve as noteworthy examples of the deep learning integration trend.[13]

V. Transfer Learning-Fine Tuning

Transfer learning methods facilitate the acceleration of learning for distinct yet related tasks using models originally trained for a different primary task[14].This approach allows models to leverage knowledge gained from related tasks to improve performance in scenarios with limited training data. Established models like VGG16 and ResNet50, which were initially trained on a vast dataset containing 1.2 million images and 1,000 categories from ImageNet, benefit significantly from transfer learning. In this process, pre-trained ImageNet weights are utilized to initialize convolutional layer weights before complete retraining of the network's classifier component.

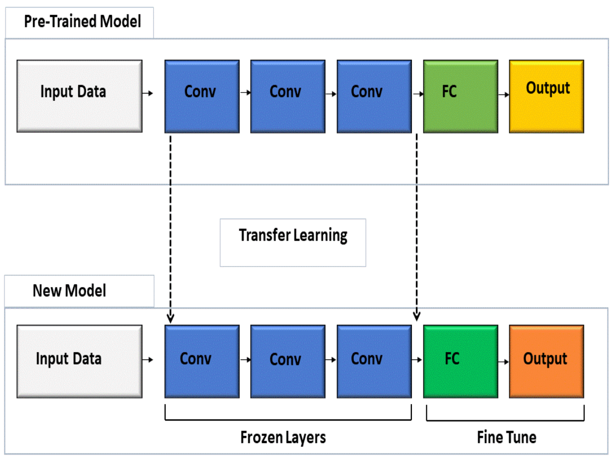

Fig. 6. Concept of transfer learning.

VI. Results and Discussion

We firstly showed the training and validation loss curves produced by ResNet50 and VGG16 models. After that confusion metrics and model image prediction has been shown.

A. Output of ResNet50 Model

1. Training and Validation Results

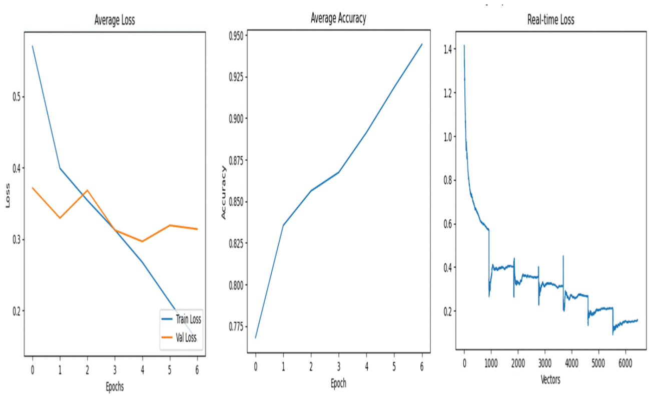

Fig. 7. Average loss, average accuracy and real time loss curve of ResNet50 model.

This study involved training the model over multiple epochs, with each epoch utilizing a batch size of 8, resulting in 991 training steps per epoch. We started with a learning rate of 0.0001, and we implemented the ReduceLROnPlateau callback to dynamically adjust the learning rate based on the model's performance. This callback monitored a specific metric, and if no upgrading was observed after a predefined epochs, it automatically reduced the learning rate by a specified factor. As a result, after 8 epochs of training, we were able to achieve the highest accuracy of 88.8%.

2. Confusion Metrics and ROC Curve

A confusion matrix offers a perspective on the ratio of accurate and inaccurate predictions generated by a classification model. It facilitates the assessment of diverse performance metrics such as accuracy, precision, recall, and others. Figure 8 depicts the confusion matrix associated with the ResNet50 model.

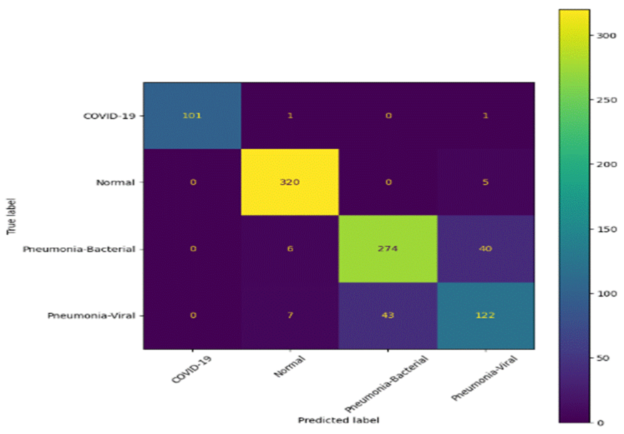

Fig. 8. Confusion matrix of ResNet50 model.

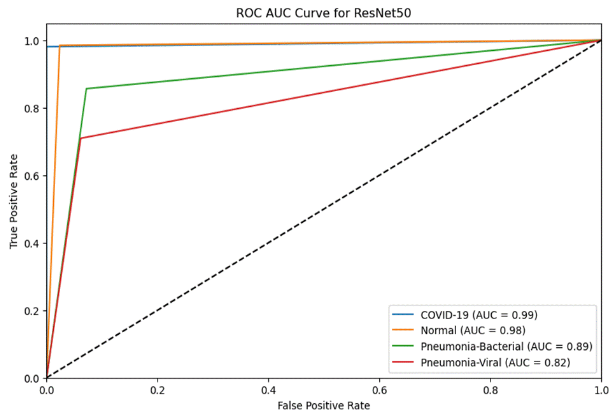

Fig. 9. ROC curve for ResNet50 model.

The ResNet50 model's ROC AUC curve in this study offers a thorough analysis of its performance in classifying medical images into several categories. Notably, the model achieves a stunning ROC AUC score of 99% when it comes to recognising COVID-19 cases, demonstrating extraordinary discrimination capabilities. With a score of 98%, it also exhibits a high degree of precision in differentiating normal cases. With ROC AUC ratings of 0.89 and 0.82, respectively, the model finds it slightly more difficult to differentiate between instances of viral pneumonia and bacterial pneumonia. The model appears to perform rather well in these categories, as indicated by these scores, but it appears that its performance is particularly high when detecting COVID-19 and normal instances. Overall, the ROC AUC curve highlights the ResNet50 model's effectiveness in distinguishing different medical conditions, with its highest accuracy observed in COVID-19 and normal classifications.

3. ResNet50 Model Image Prediction

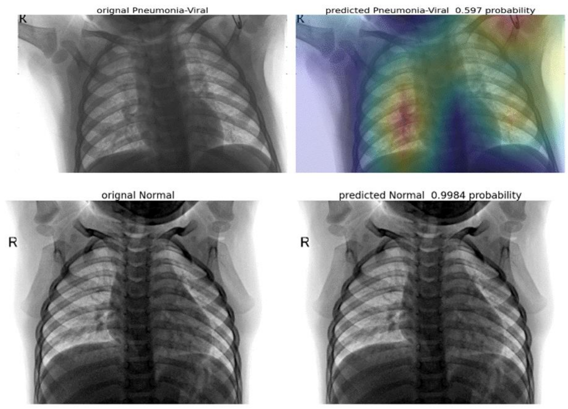

Fig. 10. ResNet50 model image prediction.

The ResNet50 model's predictive capabilities are briefly described in this section. In identifying images from the test dataset, the ResNet50 model performs very well. The model shows a probability of about 0.597 for properly detecting an original image labelled "Pneumonia-Viral" when it is displayed. Similarly, for images categorised as 'normal,' the model confidently assigns a probability of approximately 0.9984, correctly classifying them as normal

B. *Output of VGG16 Model*

1. *Training and Validation Results*

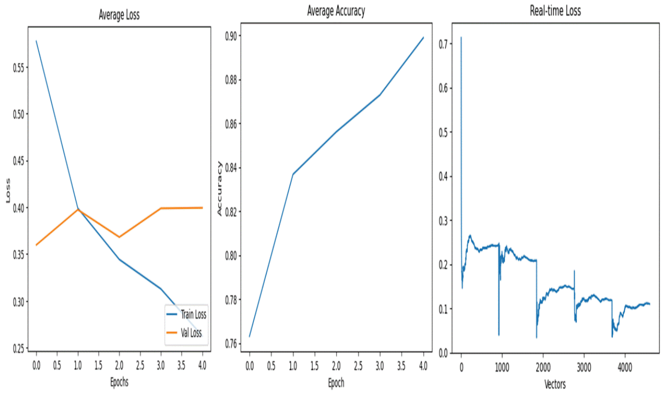

Fig. 11. Average loss, average accuracy and real time loss curve of VGG16 model.

Similarly, the Vgg16 model underwent training across multiple epochs, with each epoch employing a batch size of 8. The initial learning rate was established at 0.0001, and we incorporated the ReduceLROnPlateau callback to dynamically fine-tune the learning rate in response to the model's performance. Following 6 epochs of training, we achieved the peak accuracy of 89.23%.

2. *Confusion Metrics and ROC Curve*

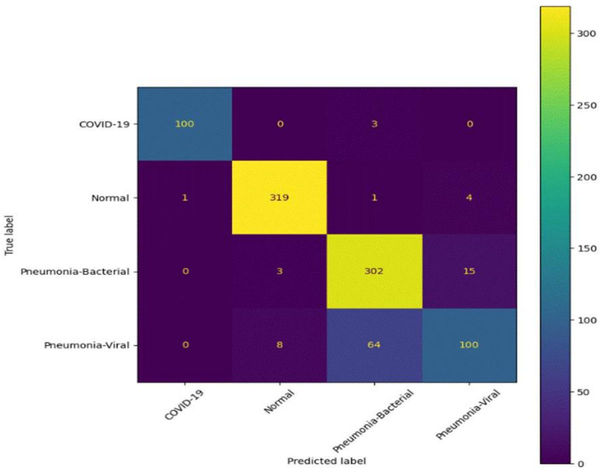

Fig. 12. Confusion matrix of VGG16 model.

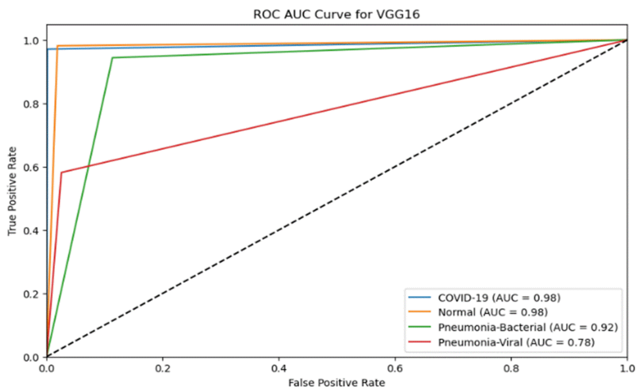

Fig. 13. ROC curve for VGG16 model.

This study's ROC AUC curve for the VGG16 model illustrates how well it can categorise medical images into different groups. With comparable ROC AUC scores of 0.98, the model impressively exhibits a high degree of accuracy in differentiating both COVID-19 and normal instances. This suggests that the VGG16 model performs exceptionally well in distinguishing COVID-19 cases from normal instances as well as between non-COVID-19 and COVID-19 cases, respectively. Additionally, the model performs well in detecting instances of bacterial pneumonia, earning a ROC AUC score of 0.92. The ROC AUC score of 0.78, which indicates that it has substantially more difficulty differentiating viral pneumonia patients, reflects this. This score denotes a decent level of accuracy, even if it shows that the model's performance in this category is less robust. In conclusion, the VGG16 model correctly classifies COVID-19 , normal cases and bacterial pneumonia cases, but it struggles to separate viral pneumonia cases.

3. *VGG16 Model Image Prediction*

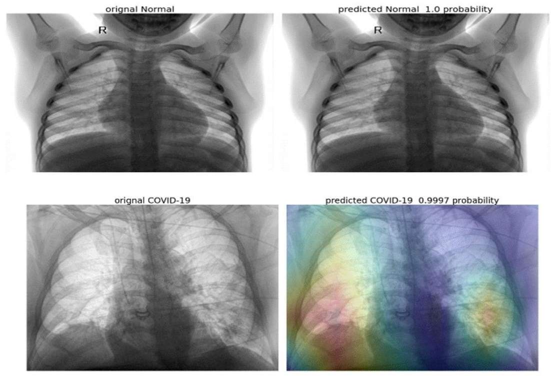

Fig. 14. VGG16 model image prediction.

The VGG16 model performs with remarkable accuracy when categorizing images from the test dataset. Notably, the model gives probability of 1, suggesting that it correctly recognizes images labelled as "Normal" when they are presented. The model also exhibits a high degree of confidence when classifying image as "COVID-19," with a probability of roughly 0.9997, guaranteeing that these images were correctly labelled as COVID-19 cases.

TABLE I. PERFORMANCE METRICS OF RESNET50 AND VGG16 MODELS.

Model	**Class**	**Accuracy (%)**	**Precision (%)**	**Recall (%)**	**F1-score (%)**
ResNet50	COVID-19	88.80	99.02	100	99.51
	Normal		97.85	96.68	97.26
	Pneumonia-Bacterial		86.40	83.49	84.92
	Pneumonia-Viral		96.06	73.85	83.62
VGG16	COVID-19	89.23	99.01	90.91	94.83
	Normal		98.61	98.39	98.50
	Pneumonia-Bacterial		82.42	94.20	87.86
	Pneumonia-Viral		90.91	43.80	59.32

Table I, delineates the performance of the ResNet50 and VGG16 models in the classification of distinct lung pathologies, including COVID-19, normal lung conditions, bacterial pneumonia, and viral pneumonia. The ResNet50 and VGG16 achieved commendable accuracy of 88.80% and 89.23%, respectively. Notably, VGG16 exhibited a notably lower recall for COVID-19 cases, while ResNet50 demonstrated exceptional precision and recall. Both models excelled in identifying patients with normal lung conditions, displaying commendable precision and recall.

TABLE II. COMPARISON OF PERFORMANCE WITH EXISTING WORKS

Authors	Models	Performance
Varshani et al.[15]	CNN models, SVM and DenseNet169	Accuracy=80.02%
Omar et al.[7]	Five convolutional layers-based CNN model	Accuracy=87.65%
Ayan et al.[16]	Xception and VGG16	Xception model accuracy=82% and 87% for Xception model accuracy=87%
Jakub Garstka et al.[10]	Customized CNN model	Accuracy=85%
Han et al.[11]	VGG16, VGG19 and DenseNet169	Accuracy=85%
Proposed Model	VGG16 and ResNet50	Accuracy 89.23% and 88.80% respectively

Table II provides a comparative analysis of multiple models for the precise classification of pneumonia and COVID-19. Various authors have employed unique methodologies in the development of their respective models. Moreover, it is evident that our proposed model has exhibited superior performance metrics when contrasted with the existing models.

VII. CONCLUSION

We have illustrated the process of categorizing X-ray images into pneumonia, non-pneumonia, and COVID-19 classes. Our approach involved leveraging transfer learning and meticulously optimizing hyperparameters for the VGG16 and ResNet50 models. We conducted a comparative analysis with existing methodologies to evaluate the performance of our proposed model. The results demonstrated the robustness of our model in accurately classifying chest X-ray (CXR) images. Notably, as the model underwent training, we observed a consistent improvement in accuracy and a reduction in loss during each epoch. Furthermore, our proposed model exhibited a notable reduction in computational complexity while maintaining high prediction accuracy. In the medical domain, this advancement is particularly advantageous for enabling prompt and precise pneumonia diagnoses, potentially contributing to timely and life-saving treatments.

REFERENCES

[1] Imran Hossain, “Every hour pneumonia kills 3 children in Bangladesh,” Nov. 2020.

[2] Hasan MK, Ahmed S, Abdullah ZE, Monirujjaman Khan M, Anand D, Singh A, AlZain M, Masud M, “Deep Learning Approaches for Detecting Pneumonia in COVID-19 Patients by Analyzing Chest X-Ray Images,” Mathematical Problems in Engineering 2021: pp.1-8, doi: 10.1155/2021/9929274.

[3] Zhang, Xin, Siyuan Lu, Shui-Hua Wang, Xiang Yu, Su-Jing Wang, Lun Yao, Yi Pan, and Yu-Dong Zhang., “Diagnosis of COVID-19 Pneumonia via a Novel Deep Learning Architecture,” Journal of computer science and technology, vol. 37, no. 2, pp. 330–343, Apr. 2022, doi: 10.1007/s11390-020-0679-8..

[4] Thakur, Saurabh, Yajash Goplani, Sahil Arora, Rohit Upadhyay, and Geetanjali Sharma, “Chest X-Ray Images Based Automated Detection of Pneumonia Using Transfer Learning and CNN,” International Conference on Artificial Intelligence and Applications, pp. 329–335. Springer Singapore, 2021.

[5] Bashar, A., Latif, G., Ben Brahim, G., Mohammad, N., & Alghazo, Jaffar., “COVID-19 pneumonia detection using optimized deep learning techniques,” Diagnostics, vol. 11, no. 11, Nov. 2021, doi: 10.3390/diagnostics11111972.

[6] Rajpurkar, Pranav, Jeremy Irvin, Kaylie Zhu, Brandon Yang, Hershel Mehta, Tony Duan, Daisy Ding, “CheXNet: Radiologist-Level Pneumonia Detection on Chest X-Rays with Deep Learning,” Nov. 2017, Available: http://arxiv.org/abs/1711.05225

[7] H. Omar and A. Babalık, “Detection of pneumonia from X-ray images using convolutional neural network,” Proceedings Book, vol. 183, 2019.

[8] Md. R. Hasan, S. M. A. Ullah, and Md. E. Karim, “An Effective Framework for Identifying Pneumonia in Healthcare Using a Convolutional Neural Network,” *2023 International Conference on Electrical, Computer and Communication Engineering (ECCE)*, pp. 1–6, Feb. 2023, doi: 10.1109/ECCE57851.2023.10101548.

[9] R. Hasan, S. Ullah, A Nandi and A “Improving Pneumonia Diagnosis: A Deep Transfer Learning CNN Ensemble Approach for Accurate Chest X-Ray Image Analysis,” *2023 International Conference on Information and Communication, 2023• doi: 10.1109/ICICT4SD59951.2023.10303322*

[10] Garstka, Jakub, and Michał Strzelecki “Pneumonia detection in X-ray chest images based on convolutional neural networks and data augmentation methods,” 2020 Signal Processing: Algorithms, Architectures, Arrangements, and Applications (SPA), pp. 18-23. IEEE, 2020.

[11] hanh, Han Trong, Pham Huong Yen, and Trinh Bich Ngoc, “Pneumonia classification in X-ray images using artificial intelligence technology,” 2020 Applying New Technology in Green Buildings (ATiGB), pp. 25-30. IEEE, 2021.

[12] Sait, Unais, K. G. Lal, S. Prajapati, Rahul Bhaumik, Tarun Kumar, S. Sanjana, and Kriti Bhalla, “Curated dataset for COVID-19 posterior-anterior chest radiography images (X-Rays),” Mendeley Data 1,2020.

[13] Alzamily, Jawad Yousef Ibrahim, Syaiba Balqish Ariffin, and Samy S. Abu-Naser, “Classification of Encrypted Images Using Deep Learning–Resnet50,” Journal of Theoretical and Applied Information Technology, 100(21), pp.6610-6620.

[14] K. Weiss, T. M. Khoshgoftaar, and D. D. Wang “A survey of transfer learning,” J Big Data, vol. 3, no. 1, Dec. 2016, doi: 10.1186/s40537-016-0043-6.

[15] D. Varshni, R. Nijhawan, K. Thakral, A. Mittal, and L. Agarwal, “Pneumonia Detection Using CNN based Feature Extraction,” IEEE international conference on electrical, computer and communication technologies (ICECCT),2019. pp. 1-7.

[16] E. Ayan and H. M. Ünver, “Diagnosis of Pneumonia from Chest X-Ray Images Using Deep Learning,” 2019 Scientific Meeting on Electrical-Electronics & Biomedical Engineering and Computer Science (EBBT), pp. 1–5, 2019.