Diagnosing COVID-19 with Deep Learning

Yashika Khurana, Raj Oza, Amey Kolhe
New York University
yk2773@nyu.edu, ro2151@nyu.edu, apk9563@nyu.edu
Codebase: https://github.com/khuranayashika31/DL-Final-Project/tree/main

Abstract

The COVID-19 pandemic has presented a significant global healthcare challenge, necessitating swift and accurate diagnosis to mitigate its spread. Chest X-ray and CT scan images have emerged as valuable diagnostic tools for COVID-19 due to their accessibility and ability to detect characteristic lung abnormalities. This study conducts a comparative analysis of multiple neural network architectures for diagnosing COVID-19 using chest X-ray and CT scan images. The research focuses on evaluating the performance of different deep learning models, including ResNet-50, MobileNetV2, VGG-16, and EfficientNetB0. Publicly available datasets are utilized for training and evaluating these neural network models. Through extensive experimentation, we examine the classification accuracy, precision, recall, F1-score, Cohen Kappa score, and overall performance metrics of each architecture, and present visualizations of our findings. Our results demonstrate the effectiveness of deep learning models in COVID-19 diagnosis from radiographic images.

Introduction

COVID-19 has had a profound impact on global healthcare systems, posing unprecedented challenges in terms of diagnosis and containment. Although laboratory testing methods are available, they were recently known to be reporting a huge number of false negatives. In this context, chest X-ray and CT scan images have emerged as valuable diagnostic tools due to their ability to detect characteristic lung abnormalities, particularly, ground glass opacities & abnormalities associated with the disease. Thus, this research work is an attempt to perform a comparative analysis of four neural network architectures with varying depths and choice of hyperparameters to diagnose COVID-19 in chest X-ray and CT scan images. Our objective is to propose an effective model and that has low false negative rate to consider it suitable for clinical settings. We aim to assist physicians in confirming or consulting AI along with their manual diagnosis.

The dataset used for this work is publicly available. We have acquired data from two sources. The X-ray images have been taken from [1,2] and the CT scan images have been procured from [3]. We have not used the entire dataset for this study, due to the availability limited computing resources. The figures below represent sample images from the dataset.

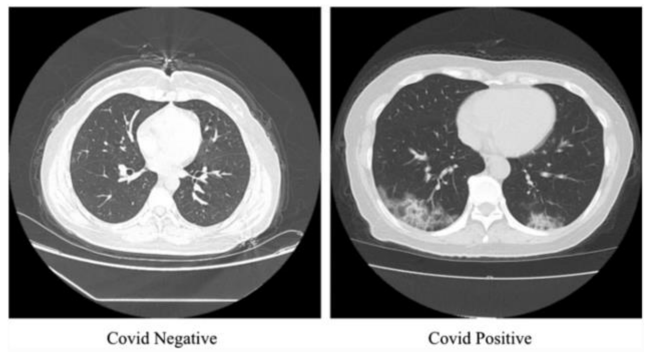

Figure 1: Sample from CT scan data

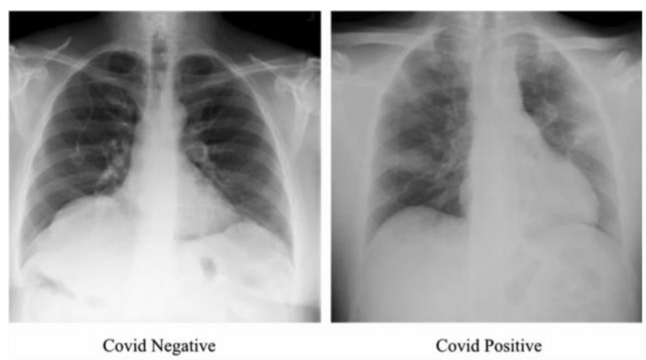

Figure 2: Sample from X-ray data

Thus, the study has been performed over 4000 CT scan images and 4000 X-ray images. The datasets were sliced to contain 3000 images in the training set and 1000 images in the test set. Further, while training, 10% of the training data is used for validation. The truncated dataset includes balanced (equal) representation from both the classes i.e. COVID positive and COVID negative, which are addressed as 1 and 0, respectively. Also, we resized the images in the dataset to have width and height of 224 by inter-area interpolation and shuffled it prior to training and testing. After the aforementioned steps, we used the following neural network architectures on both the CT scan & X-ray datasets to evaluate and compare their performance.

1. MobileNet-V2
2. VGG-16
3. EfficientNet-B0
4. ResNet-50

All these networks use the pre-trained weights from the ImageNet dataset [3] which is a very large database containing over 14 million images across 20k categories. The specifications for each model has been discussed in the following sections of the paper. Further, we measured the performance of the above models by calculating the test accuracy, precision, recall, F1-score and Cohen Kappa score. Finally, we discuss the performance of the employed models on our dataset and visualize our findings. The results and conclusion section will discuss the objectives achieved by our study.

Literature Review

Numerous studies have explored the use of deep learning models for COVID-19 diagnosis. E.H Chowdhury et al. used Deep CNN models like SqueezeNet, Inception-V3 etc. to classify X-ray images into three categories- COVID-19 pneumonia, viral pneumonia, and normal to obtain above 99% accuracy and precision. Mohammad Jamshidi et al. [4] illustrated the use of Generative Adversarial Networks (GANs), Extreme Learning Machine (ELM), and Long/Short Term Memory (LSTM) models to identify COVID-19 in radiographic images. Javiar Masot et al. [5] use torso radiographs to determine COVID-19 using VGG-16 network& achieve high specificity. Xin Zhang et al. [6] modify the AlexNet deep learning network by using batch normalization and replacing its fully connected layers to obtain over 91% accuracy on a private dataset. Hussain et al. [7] used a 22 layers deep architecture for 3-class classification on a self-constructed database.

Overall, we did not find a study that employs the same architectures on both CT scan and X-ray images and analyze their performance. Thus, we wanted to bridge that gap. Also, we chose ResNet-50 because it is known to demonstrate excellent performance in various classification tasks. It uses the residual connections to overcome the vanishing gradient issue, which makes it an effective tool for diagnosis. MobileNet-V2 has low computational requirements which makes it a compact model with separable convolutions. So, we were particularly interested in understanding its performance and deploy it in constrained situations in the future work. VGG-16 has a deep architecture with stacked convolutional layers that help capture intricate features. Several studies have investigated the performance of VGG-16 in detecting COVID-19 from radiographic images, highlighting its effectiveness in achieving high classification accuracy. EfficientNetB0, a recent addition to the family of EfficientNet models, has gained attention for its impressive performance in image classification tasks. It balances model size and computational efficiency through compound scaling techniques. Although its application in COVID-19 diagnosis is relatively nascent, initial studies in this aspect have shown promising results. Further, we have used pre-trained ImageNet weights to extend the feature rich knowledge to our models. Further, we used random search hyperparameter tuning approach in two of the models. It is known to optimize the performance of machine learning models by searching for the best combination of hyperparameters within a predefined search space. The corresponding observations are discussed in the next sections. Collectively, the literature demonstrates the effectiveness of deep learning models, such as ResNet-50, MobileNetV2, VGG-16, and EfficientNetB0, in COVID-19 diagnosis using radiographic images. These models have shown promising results and hold great potential to contribute to the development of accurate and efficient diagnostic tools, aiding healthcare practitioners and researchers in the fight against the COVID-19 pandemic.

Models Deployed

This section discusses the various neural network architectures used in our research work. We deployed 8 models in total, 4 for X-ray and 4 for CT scan images. However, the configurations are kept same across them to facilitate comparison between results derived from the two different datasets.

ResNet-50
Optimizer: Adam
Hyperparameters tuned (using Random Search):
Learning rate:0.01
Batch size: 16
Inner layer activation: ReLU
Output layer activation: Sigmoid
Loss: Binary cross entropy
Epochs: 40

MobileNet-V2
Optimizer: Adam
Hyperparameters tuned (using Random Search):
Learning rate: 0.001
Number of filters:128
Inner layer activation: ReLU
Output layer activation: Sigmoid
Loss: Binary cross entropy
Epochs: 40

EfficientNet-B0
Optimizer: Adam
Learning rate: 0.0001
Dropout:0.5
Inner layer activation: ReLU
Output layer activation: Sigmoid
Loss: Binary cross entropy
Epochs: 40

VGG-16
Optimizer: Adam
Learning rate: 0.0001
Dropout:0.5
Inner layer activation: ReLU
Output layer activation: Sigmoid
Loss: Binary cross entropy
Epochs: 40

Results

We evaluated our models on the metrics of precision, recall, F1-score and Cohen Kappa score. The confusion matrix corresponding to each model is also presented below.

Table 1: Results obtained

		ResNet-50	MobileNet-V2	VGG-16	EfficientNet-B0
X-ray images	Test Accuracy	0.970	0.825	0.958	0.967
	Recall	0.972	0.652	0.934	0.982
	Precision	0.968	1.0	0.981	0.953
	F1-score	0.970	0.789	0.956	0.967
	Cohen Kappa score	0.940	0.652	0.916	0.934
CT scan images	Test accuracy	0.998	0.999	0.999	0.999
	Recall	0.998	0.998	1.0	0.998
	Precision	0.998	1.0	0.998	1.0
	F1 score	0.998	0.998	0.999	0.998
	Cohen Kappa score	0.996	0.998	0.998	0.998

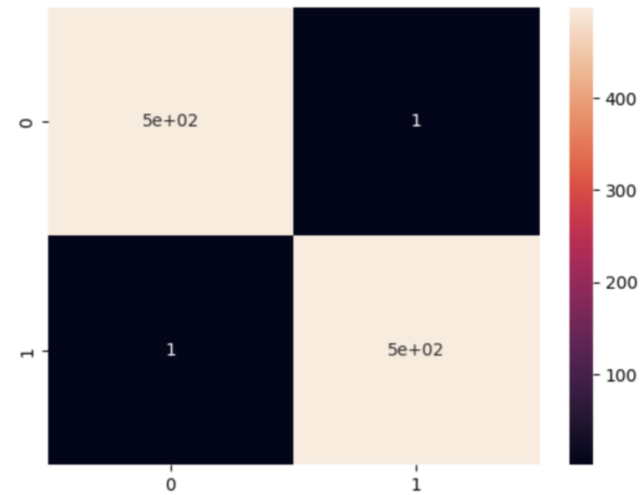

Figure 3: ResNet CT

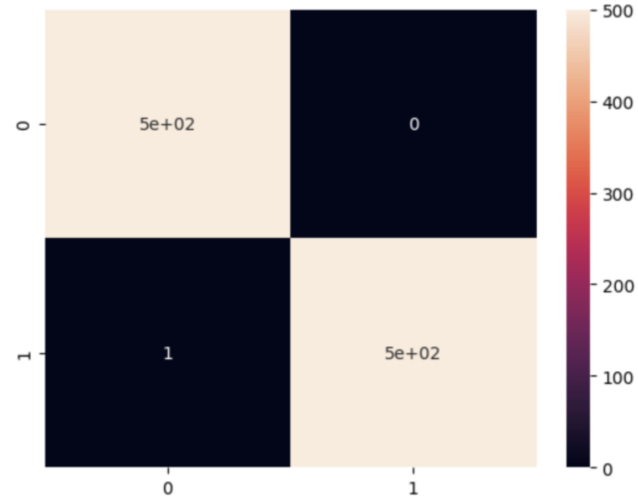

Figure 4: MobileNet CT

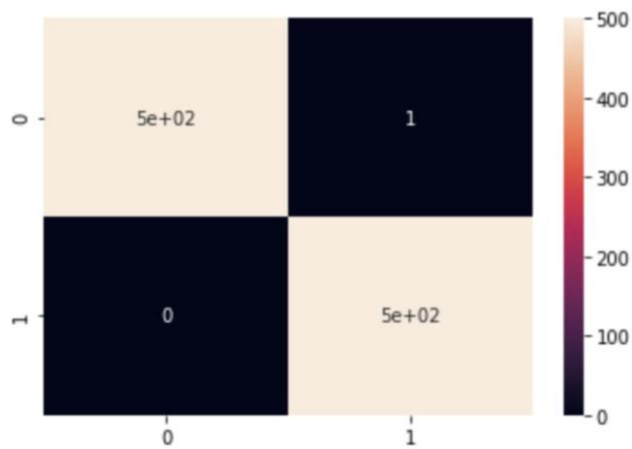

Figure 5: VGG CT

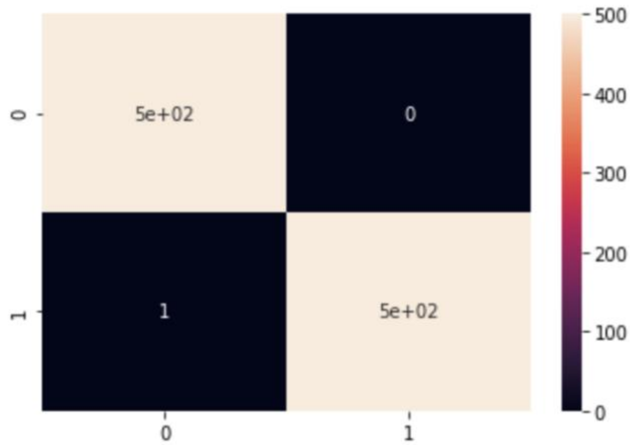

Figure 6: EfficientNet CT

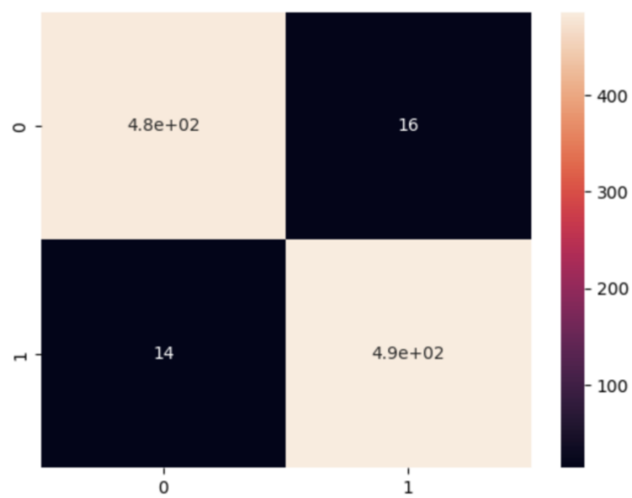

Figure 7: ResNet XR

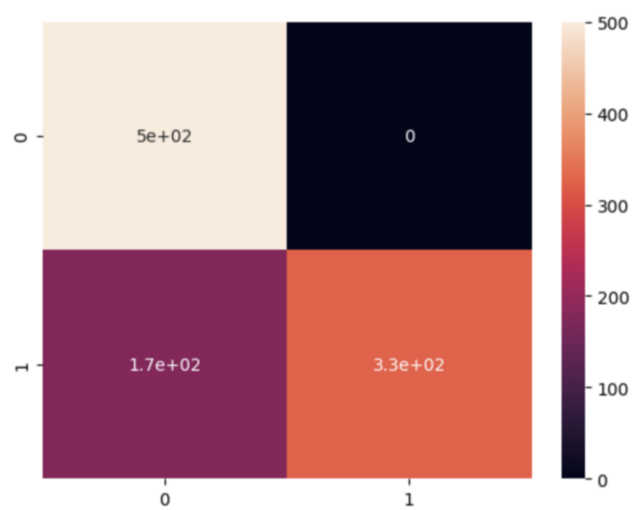

Figure 8: MobileNet XR

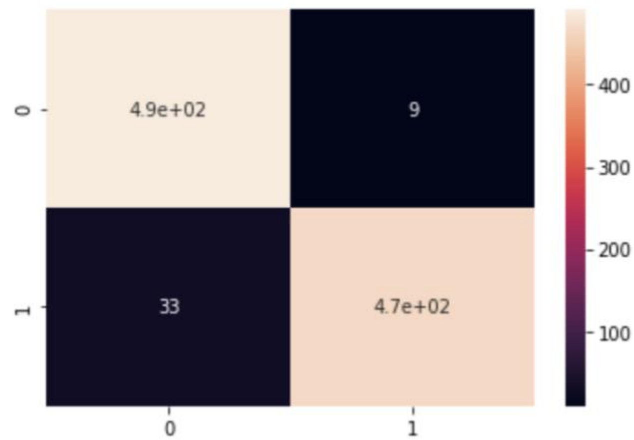

Figure 9: VGG XR

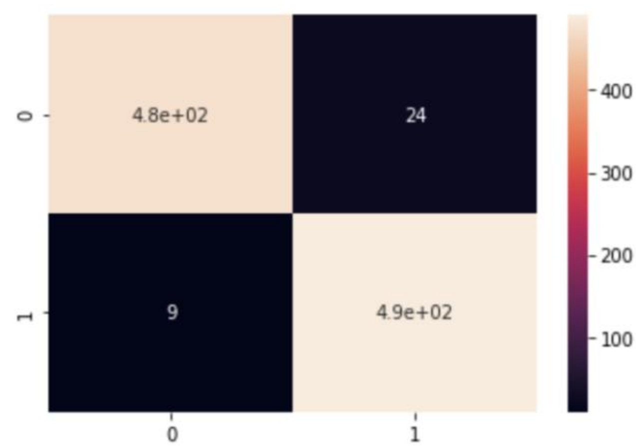

Figure 10: EfficientNet XR

The following curves represent the Receiver Operating characteristic curves demonstrate the performance and

AUC of the ResNet-50 model on both X-ray and CT scan images.

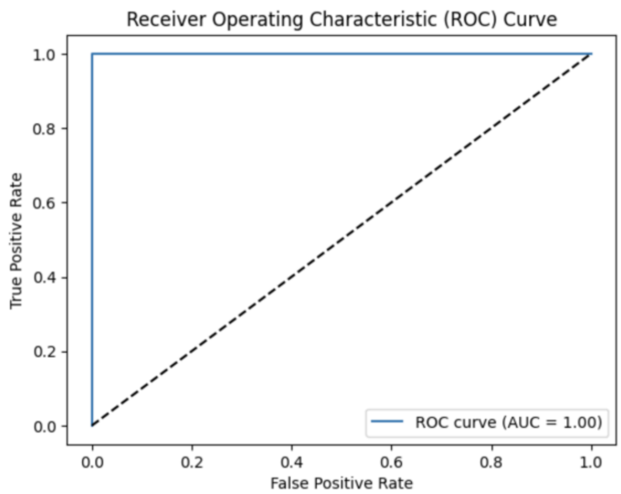

Figure 11: ROC curve for ResNet-50 on CT scan images

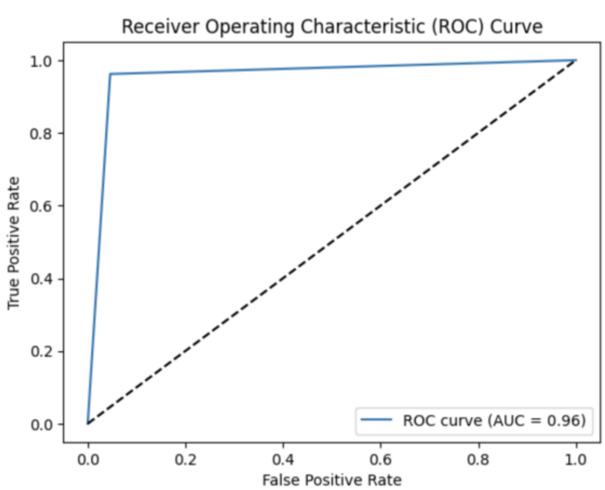

Figure 12: ROC curve for ResNet-50 on X-ray images

Discussion

For X-ray images, the ResNet-50 model, pretrained on ImageNet weights and after hyperparameter tuning, demonstrates high performance in terms of test accuracy, precision, recall, F1-score and Cohen Kappa score. Highest recall value has been reported by the EfficientNet-B0 model. Please refer to Table 1 above.

For CT scan images however, all the employed models report 0.99 above on all calculated metrics. Most papers we referred for our project demonstrate a very high accuracy in COVID-19 classification problem on CT scan images. Our base paper [3] also reports above 0.99 accuracy and other calculated metrics. We observe comparable results in our work.

Conclusion

The research work propounds the use of the four neural network architectures: ResNet-50, MobileNet-V2, EfficientNetB0 and VGG-16 for COVID-19 diagnosis using radiographic images, chest CT scans and X-ray, especially the ResNet-50 model, with hyperparameter tuning, pre-trained on ImageNet. Our base paper utilized the chest X-ray images for their study, however we also used the CT scan dataset to understand its applicability in clinical settings. We proposed to achieve performance comparable to that obtained by our base paper, which has been accomplished. Our base paper reports the best test accuracy, precision, recall and F1 score of 99.42, 99.42, 99.41 respectively. However, we achieved 99.8% for all the metrics in case of CT scan images. Comparing the performance of the models, it was observed that CT scans are better suitable for the purpose of diagnosis because the same models achieve better results in terms of precision, recall, test accuracy, F1-score & Cohen Kappa score on CT scan images than in X-ray images. Also, we achieved a high AUC of 0.96 on X-ray and 1.0 on CT scan, which indicate a good performance. We proposed to reduce the false positive rate and the highest precision for X-ray images 0.968 on X-ray and 0.998 on CT scan, which is promising. The localization of the infected region is in the scope of our future work. Thus, we promote the usage of CT scan images as a supplement to manual diagnosis by a healthcare practitioner.

References

[1] M.E.H. Chowdhury, T. Rahman, A. Khandakar, R. Mazhar, M.A. Kadir, Z.B. Mahbub, K.R. Islam, M.S. Khan, A. Iqbal, N. Al-Emadi, M.B.I. Reaz, M. T. Islam, “Can AI help in screening Viral and COVID-19 pneumonia?” IEEE Access, Vol. 8, 2020, pp. 132665 - 132676.

[2] Rahman, T., Khandakar, A., Qiblawey, Y., Tahir, A., Kiranyaz, S., Kashem, S.B.A., Islam, M.T., Maadeed, S.A., Zughaier, S.M., Khan, M.S. and Chowdhury, M.E., 2020. Exploring the Effect of Image Enhancement Techniques on COVID-19 Detection using Chest X-ray Images

[3] Wanshan Ning, Shijun Lei, Jingjing Yang, Yukun Cao, Peiran Jiang, Qianqian Yang, Jiao Zhang, Xiaobei Wang, Fenghua Chen, Zhi Geng, Liang Xiong, Hongmei Zhou,

Yaping Guo, Yulan Zeng, Heshui Shi, Lin Wang, Yu Xue, Zheng Wan. Open resource of clinical data from patients with pneumonia for the prediction of COVID-19 outcomes via deep learning

[4] M. Jamshidi et al., "Artificial Intelligence and COVID-19: Deep Learning Approaches for Diagnosis and Treatment," in IEEE Access, vol. 8, pp. 109581-109595, 2020/

[5] Civit-Masot, J.; Luna-Perejón, F.; Domínguez Morales, M.; Civit, A. Deep Learning System for COVID-19 Diagnosis Aid Using X-ray Pulmonary Images. Appl. Sci. **2020**, 10, 4640.

[6] Zhang, X., Lu, S., Wang, SH. et al. Diagnosis of COVID-19 Pneumonia via a Novel Deep Learning Architecture. J. Comput. Sci. Technol. **37**, 330–343 (2022).

[7] Emtiaz Hussain, Mahmudul Hasan, Md Anisur Rahman, Ickjai Lee, Tasmi Tamanna, Mohammad Zavid Parvez. CoroDet: A deep learning based classification for COVID-19 detection using chest X-ray images. Elsevier Chaos, Solitons & Fractals Vol. 142