Lung CT Scan-based Covid-19 Detection

Najeebuddin Ahmed, Khanjan Dabhi and Prateek Asthana

***Abstract*—The coronavirus continues to disrupt our everyday lives as it spreads at an exponential rate. It needs to be detected quickly in order to quarantine positive patients so as to avoid further spread. This paper compares different methods by which COVID-19 classification is performed on lungs CT scan images. The performance between VGG16, ResNet50 and MobileNet is compared. The effects of image segmentation and adaptive histogram equalization on the performance of these models are analyzed. This paper also explores various problems which were tackled by the classification models from a data science perspective and outlines our proposed method for the solution.**

***Index Terms*—COVID-19, Literature review, Classification, CNN, Adaptive Histogram Equalization, Image Segmentation, Diagnosis.**

I. Introduction

Coronavirus, also known as COVID-19 was discovered in Wuhan, China, in December of 2019. COVID-19 has many strains and can infect animals and humans. COVID-19 is hard to detect because it has common symptoms such as cold and flu. The symptoms also range in seriousness depending on the person's immune system. Symptoms can take up to 14 days to appear after exposure. Because of this, the public disregards them as everyday common flu or cold. COVID-19 is spread through respiratory droplets when you cough, sneeze, and touch [1]. COVID-19 spread has become so severe it is shutting down our economies. There are over 127 million worldwide cases and over 2.7 million deaths as of March 29, 2021 and rising daily [2]. CT scans can be conducted quickly and efficiently for detecting COVID-19. For this project we are going to use the data-set provided by Luis Blanche, Alexandra Lorenzo, COVID-19 Lung CT Scans [3] to develop a deep convolutional neural network model which will predict if the person is affected with COVID-19. The quicker the detection, the quicker the patient will receive treatment and can be put in quarantine to avoid further spread. The data set comprises of over 300 CT scans of COVID-19 positive lungs collected from 216 patients. It also consists of approximately 400 CT scans of Non COVID-19 lungs. We are going to test the data-set with several pre-trained CNN models and try to modify them in order to achieve high accuracy for classification. Before classification is done the images will be preprocessed and prepared to be well suited for the models. We will implement the classification on adaptive histogram equalized and segmented images and compare them in order to give us an idea on how image segmentation can affect our classification model. We are going to record the performance metrics for these models such as accuracy, and AUC to see which model performs better. Furthermore, the paper is organized as follows: Section II presents the related works, Section III discusses the selected data-set, Section IV walks through the prepossessing of our selected data-set, Section V shows sanity tests, Section VI talks about the proposed model, Section VII contains the experimental analysis and finally, Section VIII is our conclusion.

Najeebuddin Ahmed, Khanjan Dabhi and Prateek Asthana are with the Department of Software Engineering, Lakehead University, Thunder Bay, ON, Canada. (e-mail: {nahmed4,kdabhi,pasthana}@lakeheadu.ca).

II. Literature Review

There has been much research done on classification of COVID-19 using machine learning and artificial intelligence. Mostly this research is focused on classification and forecasting of the disease using several ML techniques. Some of the papers published on classification problems for the COVID-19 disease are discussed below.

The paper published by Gozes et al.[4]. focuses on disease detection using CNN models. They used 453 CT scan images of lungs from 99 patients in China. They used UNet and ResNet50-2D architecture for classification, quantification and tracking for the patients. They achieved 0.996 AUC, 98.2% Sensitivity and 92.2% Specificity.

The model proposed by Wang et al.[5]. uses ResNet 18 architecture which obtained 73.1% Accuracy, 67% Specificity and 74% Sensitivity. The main aim is to do feature extraction from image data. They made a CNN based algorithm leveraging decision trees and SVM. In order to do so they modified the Inception transfer-learning model to establish the algorithm and performed internal and external validation. The results showed that they can use AI to aid decision making for radiology features in a really quick and organized manner for COVID-19.

The paper by Ali et al.[6]. compared ResNet50, InceptionV3 and Inception-ResNetV2 models for their classification performance. The three different binary classification were done with four classes as COVID, normal, pneumonia both viral and bacterial using 5-fold cross validation. They observed that the pre-trained ResNet50 model provided the highest classification performance.

Similarly, the article by Ezz et al.[7]. introduced a new framework for automatic diagnosis of COVID-19 in X-ray images they are using seven architectures of deep CNN models such as VGG19, DenseNet121, ResNetV2, InceptionV3, InceptionResNetV2, Xception, and MobileNetV2. They performed model evaluation on these models to find the best deep learning network. They found out that VGG19 and DenseNet were the better performance models for the task at hand.

Paper published by Sethy and Behera[8]. use Resnet50 with SVM to create a classification model for the diagnosis of COVID infected patients. The model achieved 95.38%

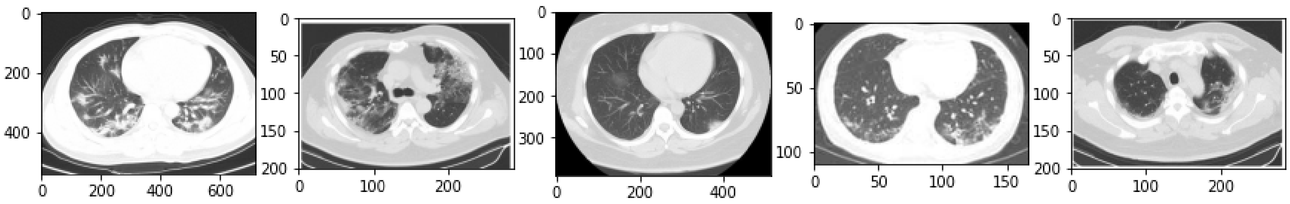


Fig. 1: Sample images of COVID-19 positive lungs CT scans from data-set.

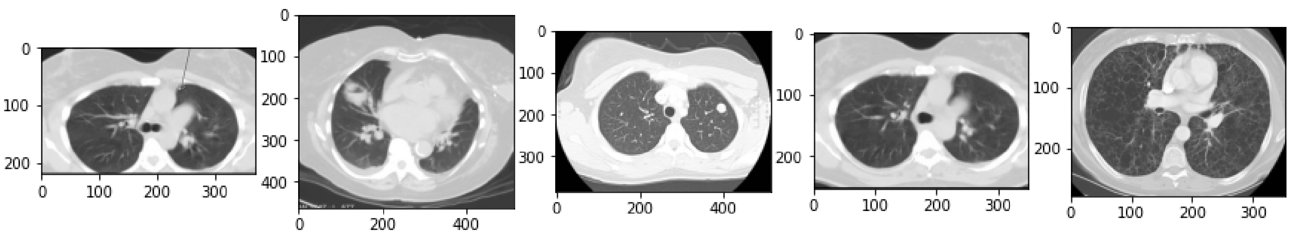


Fig. 2: Sample images of COVID-19 negative lungs CT scans from data-set.

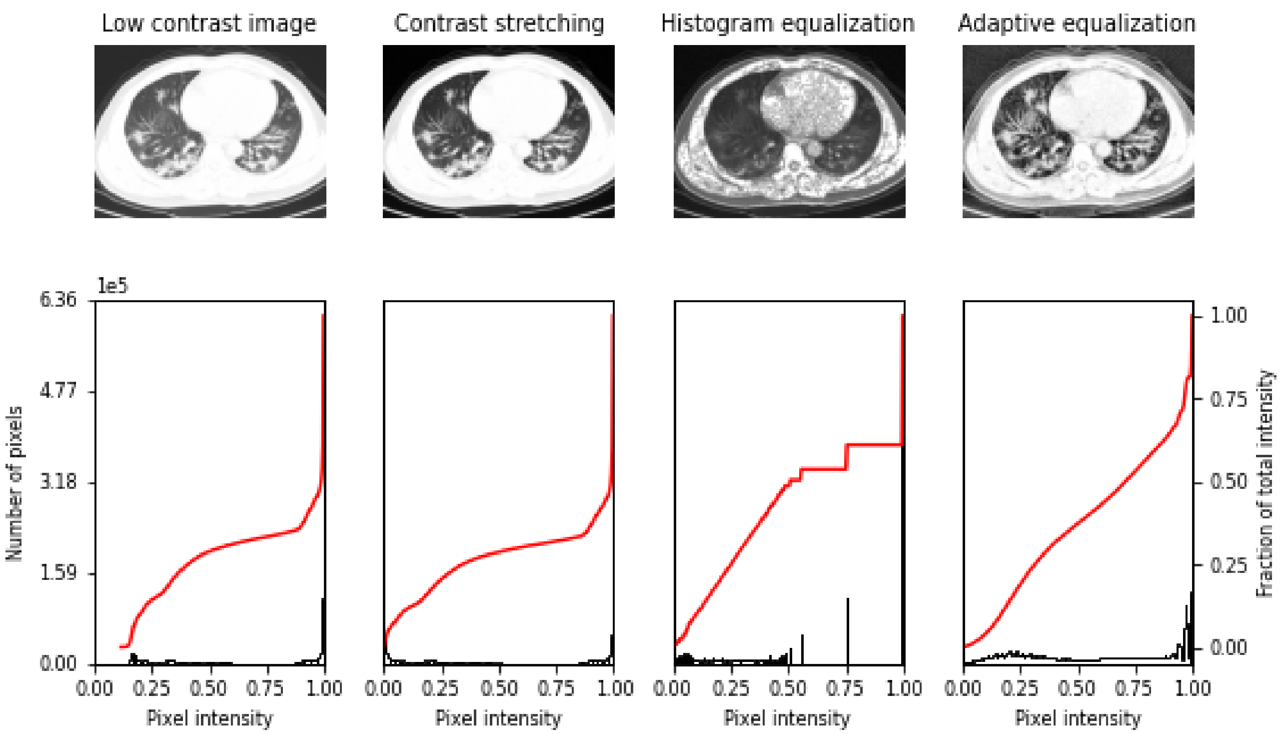


Fig. 3: Raw CT scan Transformation: Low contrast vs. Contrast stretching vs. Histogram equalization vs. Adaptive equalization.

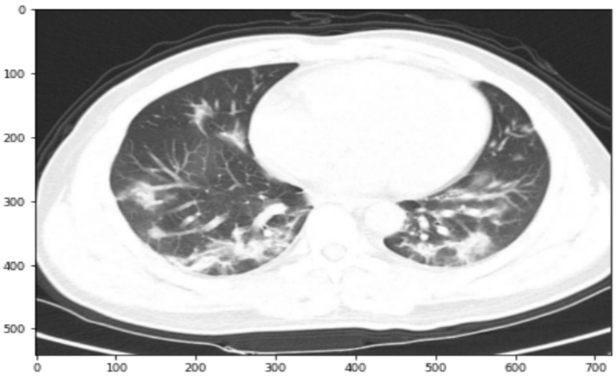
Accuracy, 95.52% FPR, 91.41% F1-score and 90.76% Kappa. This paper concluded that Resnet50 with SVM performs better than other classification models.

III. Data-set

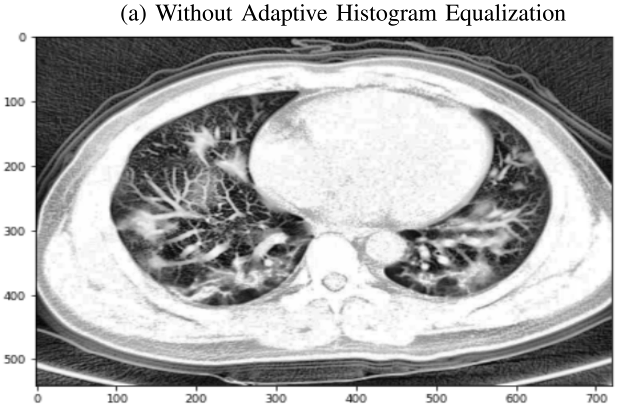
The COVID-19 Lung CT Scans data-set [3] which we used was retrieved from Kaggle and has a total of 746 images comprising of 349 COVID positive image files and 397 COVID negative image files derived from over 200 patients. Figure 1 shows sample images retrieved from the data-set that belong to the classification label of COVID (CT scan images of patients with COVID) and Figure 2 shows sample images retrieved from the data-set that belong to the classification label of NonCOVID (CT scan images of patients without COVID).

IV. Data Preprocessing

Data preprocessing is the method of preparing and converting the given data into a format that suits the machine learning models for better results.



(a) Without Adaptive Histogram Equalization

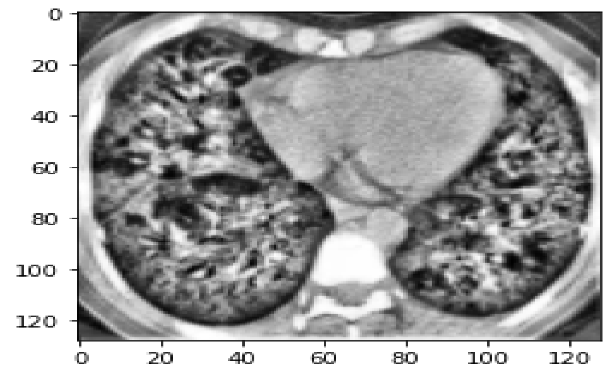


(b) With Adaptive Histogram Equalization

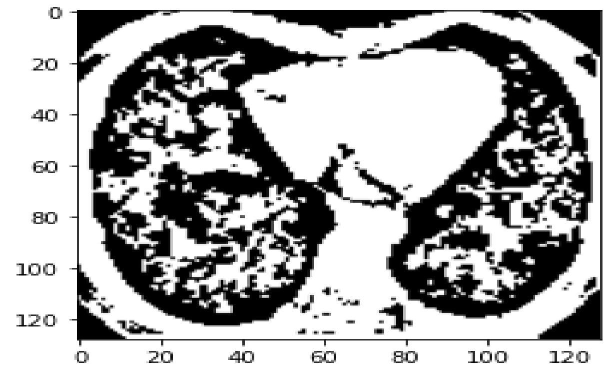
Fig. 4: Raw CT scan Transformation: Without Adaptive Histogram Equalization vs. With Adaptive Histogram Equalization.

A. Raw CT scan Transformation

Image processing techniques need to be applied in-order to get in depth and high contrast images. Some of these techniques include Contrast Stretching, Histogram Equalization and Adaptive Histogram Equalization. Contrast stretching also known as normalization, is the lengthening of intensity values to improve the image's contrast. Histogram Equalization allows the lower contrast areas of the image to get a higher contrast by spreading the intensity values that are more frequent. Adaptive Histogram Equalization sharpens the edges and betters the contrast of every region in the image by calculating a number of histograms for each region and redistributing the lightness value. In order to apply these techniques to our images in our data-set we utilized the exposure module in skimage python image processing package. Figure 3 shows the quality of the CT scan when Contrast stretching, Histogram equalization and Adaptive equalization is applied to the original raw low contrast image. The image becomes more detailed. Finally, in Figure 4 the contrast change can clearly be seen between the original image 4a and the Adaptive Histogram Equalized image 4b.



(a) Without Segmentation



(b) With Segmentation

Fig. 5: Adaptive Histogram Equalized CT scan Image: Without Segmentation vs. With Segmentation.

B. Adaptive Histogram Equalized CT scan Segmentation

K-mean clustering for image segmentation is an unsupervised algorithm which divides the image into K groups. First of all the image is converted to RGB channels instead of BRG channels as OpenCV takes that format of images. Then the image is resized to 2D array of pixels and 1 colour value. The data type of the array is the converted to float as the cv2.kmeans function only accepts that data type as input. After that the cv2.kmeas function is applied to the image array to get compactness, labels and centers from the function as output. This is then converted into 8-bit values and then back into the original image dimensions. Figure 5 above compares the adaptive histogram equalization image 5a vs. it's segmentation 5b. This complete process can visually be seen in the Segmentation section of Figure 6.

C. CT scan Augmentation

The COVID-Net: a tailored deep convolutional neural network design for detection of COVID-19 cases from chest X-ray images [9] implemented data augmentation of the images as one of their preprocessing steps and achieved a 98.9% positive predictive value for COVID-19 using COVID-Net model.

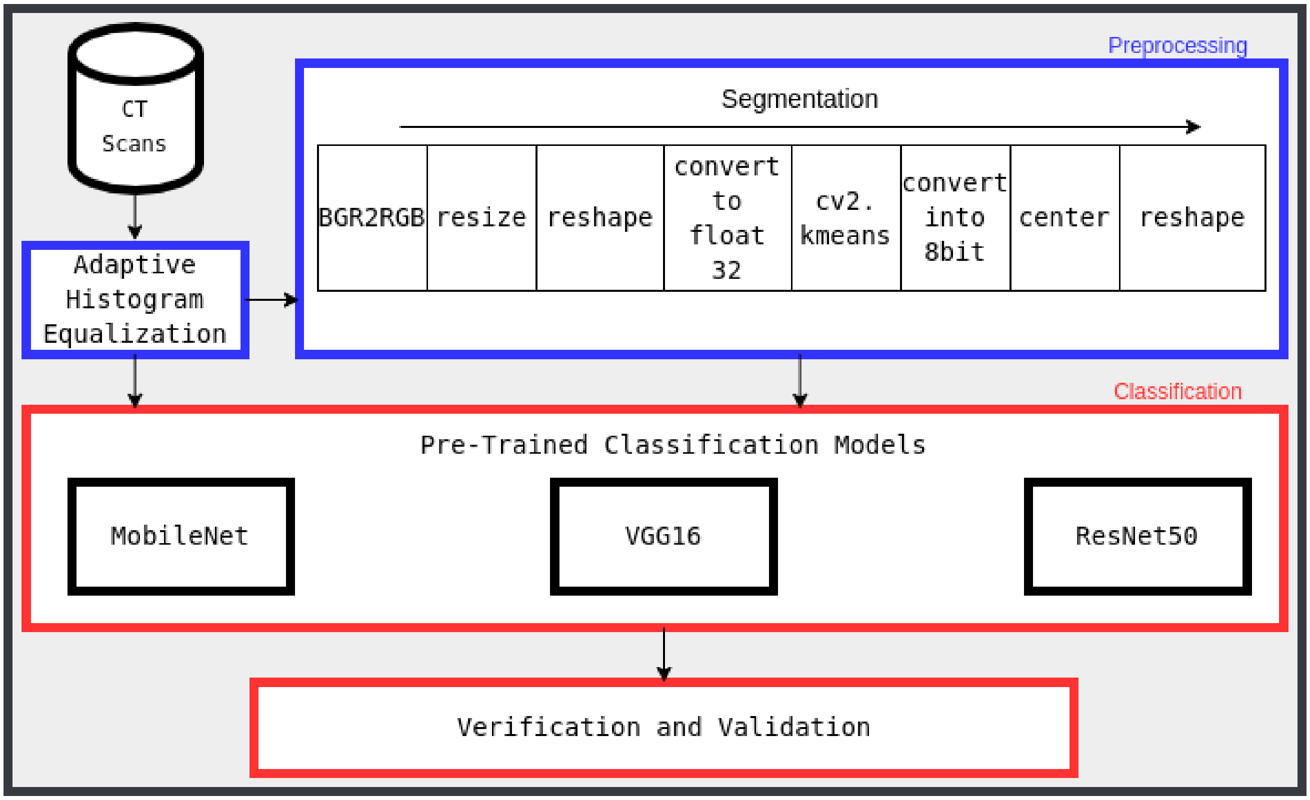


Fig. 6: The Proposed Approach.

The following data augmentation types were implemented on our images: rescale, brightness (± 10%), rotation (± 10○), translation (± 5% in x and y directions), horizontal flip, and zoom (± 15%).

The data-set's COVID-19 and Non-COVID-19 images were then each split into training(80%) and validation(20%) data-sets.

V. Sanity Tests

60% of the data-set were used as sample data, split into 40% training and 20% testing to perform our sanity tests using VGG16, ResNet, and MobileNet pre-trained models.

A. VGG16

The VGG16 [10] model was implemented with Image-Net weights in our sanity test. VGG16 is a commonly used image classification model with a lot of research done on it. It has 13 convolution layers, 5 pooling layers and 3 dense layers. It is a sequential model and has a lot of filters. We added the last fully connected layer and tested it on the sample data-set of 160 images. After training for 10 epochs we got a training accuracy of 0.50. We believe that we can improve on that accuracy by making changes in the layers for this model.

B. ResNet50

The ResNet50 model was implemented with Image-Net weights and max pooling. ResNet50 is a deep CNN that is 50 layers deep; we can load the pre-trained version of ResNet50 which is trained on ImageNet database which can classify 1000 objects [11]. When trained on our sample data-set we got around 0.50 on training accuracy. We used that to make a binary classification for this project. Similar to VGG16 after making more changes in the working layers, we can increase the training accuracy.

C. MobileNet

The MobileNet architecture uses depth-wise separable convolutions to make lightweight deep CNN. It is usually used for ML applications in embedded or mobile applications [12]. When trained on our sample data-set we got around 0.94 on training accuracy; hence, it became the most successful model of the three implemented. If we were to create an automatic diagnosis tool for CT scans then we can use such an architecture for it.

VI. The Proposed Model

The process in Figure 6 aims to reach the verification and validation state to check if a CT scan can be tested positive or negative for COVID-19. In the process, we begin with the transformation of the raw images and finally, in verification and validation, comparing the performance of different models such as VGG16, ResNet50 and MobileNet for classification.

The images are firstly segmented using K-means Clustering to enhance classification performance by the Pre-trained VGG16 model. Figure 7 below shows the comparison of segmented and adaptive histogram equalized images with

Final Results						
Model	**Parameters**	**No. of Layers**	**Parameter memory**	**FLOPS**	**Accuracy**	**AUC**
Normal VGG16	63M	16	528 MB	16 GFLOPs	53.25%	0.50
Segmentation VGG16	14M	16	528MB	16 GFLOPs	56.67%	0.31
Pre-Trained VGG16	14M	16	528 MB	16 GFLOPs	70.54%	0.69
Pre-Trained MobileNet	5.8M	28	16MB	579 MFLOPs	60.16%	0.55
Pre-Trained ResNet50	25.5M	50	98MB	4 GFLOPs	73.63%	0.81

TABLE I: Model Comparison w.r.t number of layers, number of parameters, memory footprint, performance, accuracy & AUC.

respect to training and validation accuracy. To improve the performance of the VGG16, various tuning methods, including increasing epochs, changing optimizer and reducing learning rate are performed and evaluated.

By observing the performance graphs of the model in Figure 7, the image segmentation proves to be a better power tool for learning for each models with reducing variance. However, the common validation accuracy for the segmented pictures is over the normal pictures. This could be due to the reduction in image options throughout the image because of segmentation. In addition, a straight line for the validation accuracy of the VGG16 model on the segmented pictures indicates that the model predicts identical categories for all iterations. This could ensue due to a scarcity of training or incorrect settings for weight formatting which need further improvement. Therefore, based on Figure 7 we concluded that use of normal images with adaptive histogram equalization is more appropriate than using it's segmentation due to the sole reason that the number of features get reduced in segmentation which can hinder a model's learning capabilities.

VII. Experimental Analysis

We have used three pre-trained models, i.e., VGG16, MobileNet and ResNet50 in order to perform the classification task. The pre-trained weights from the Image-Net competition are used, and only the end layers are set to be trainable for CT scan images in each of the three models. All the models have a very low learning rate (1e-5) and used sigmoid activation. The reason for using pre-trained models for our training is due to the lack of a large data-set since all the three models have fairly deep architecture, so having to train these models from scratch proved to be an ineffective strategy. Hence, the use of pre-trained weights can overcome the problem of deep architecture and small data-set. After training for 30 epochs VGG16 got an accuracy of 0.77, MobileNet had an accuracy of 0.54 and ResNet50 achieved an accuracy of 0.84. We can see that ResNet50 had the better accuracy of all the models. The reason for this is that the problem of vanishing gradient is adeptly handled by this model when compared to VGG16 or lightweight MobileNet models. However, the validation accuracy for these models range between 50% to 80% as it can be seen in Table I. This can be caused as there is a limitation of more learnable features from the small data-sets. Since we have a small data-set we also tried using K-Fold cross-validation instead of using train-test split. Upon experimenting with 5 splits our results for each of the models are outlined in Table II. It can be seen that MobileNet outperformed the other models with a mean accuracy of 90.6% due to it being over-trained. Summarized Table I shows the details of each of the model trained for this project w.r.t. number of layers, number of trainable parameters, memory footprint and their performance. Also, Figure 8 below shows the comparison of training and validation loss and accuracy for each of the three models.

K-Fold Cross-Validation Results	
Model	**Mean Accuracy**
Pre-Trained VGG16	44.43%
Pre-Trained MobileNet	90.60%
Pre-Trained ResNet50	44.42%

TABLE II: Model Comparison using K-Fold Cross-Validation w.r.t Mean Accuracy.

.

VIII. Conclusion

From Table I it can be seen that when we compare normal images with segmented images there is a noticeable loss of accuracy, this can be due to reduced image features during segmentation. Thus, only raw images with adaptive histogram equalization were used moving forward for our models. When comparing all the models we can see from Table II that ResNet50 has achieved the highest performance. We can see that having around 600 images in the data-set there is a clear limit for learnable features from this small data-set. For future application, increasing data-set size can be looked into to fully achieve the potential of the deep architectural neural networks.

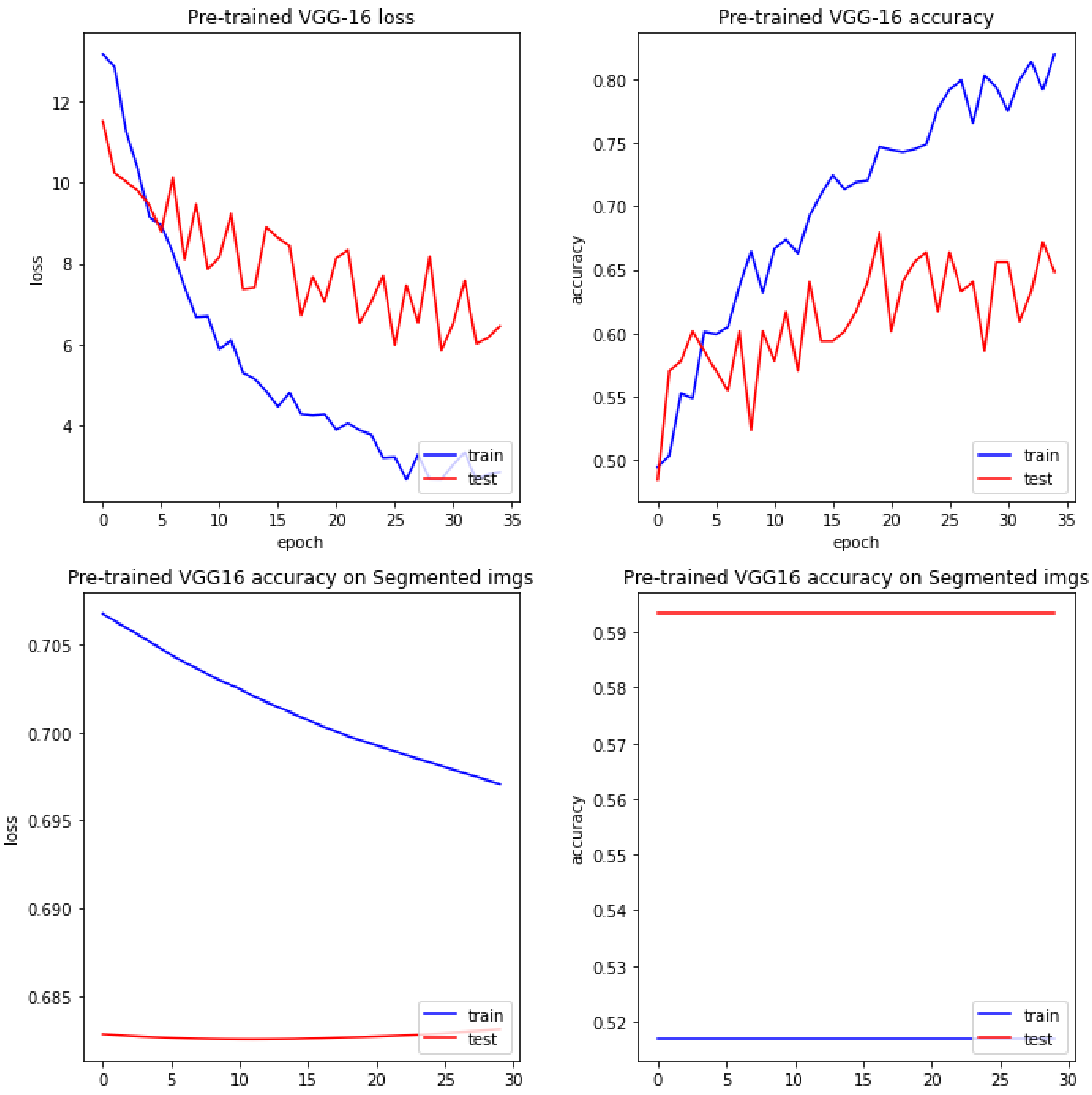


Fig. 7: Pre-trained VGG16 Performance on Transformed vs. Segmented images w.r.t. loss & accuracy graphs.

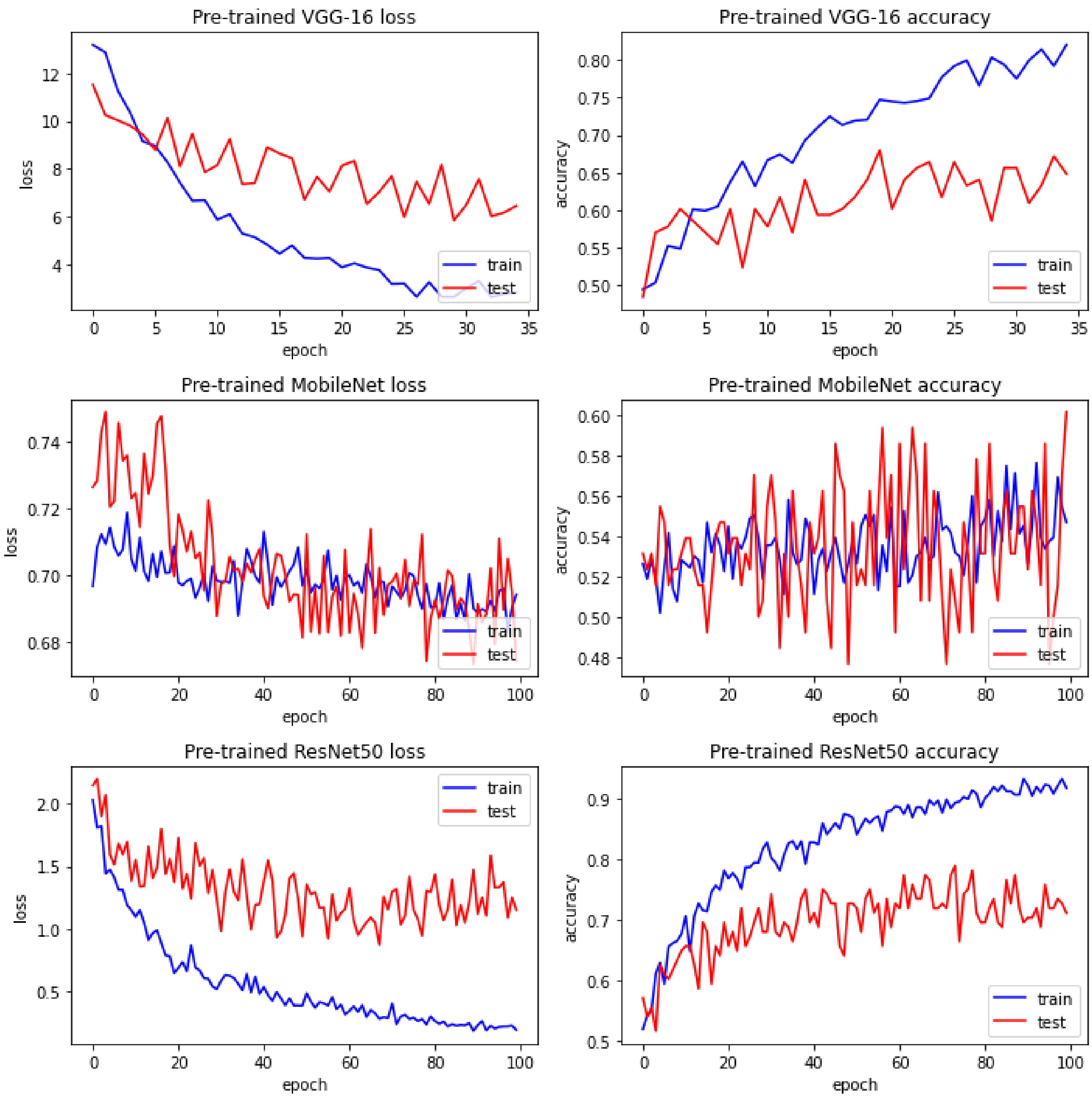


Fig. 8: Pre-trained VGG16 vs. MobileNet vs. ResNet50 performances w.r.t. loss & accuracy graphs.

ACKNOWLEDGEMENTS

This work acknowledges Luis Blanche and Alexandra Lorenzo, the organizers of COVID-19 Lung CT Scans data-set, Google Colab as the driving operation tool for this project, and Dr. Thangarajah Akilan for the technical suggestions and feedback.

REFERENCES

[1] Government of Canada, Know the facts about COVID-19; https://www.canada.ca/en/public-health/services/publications/diseases-conditions/know-facts-about-coronavirus-disease-covid-19.html
[2] Google News, Real-Time Coronavirus (COVID-19) Tracker
[3] Luis Blanche and Alexandra Lorenzo, COVID-19 Lung CT Scans, https://www.kaggle.com/luisblanche/covidct
[4] O. Gozes, M. Frid-Adar, H. Greenspan, P. D. Browning, H. Zhang, W. Ji, A. Bernheim, and E. Siegel, "Rapid ai development cycle for the coronavirus (covid-19) pandemic: Initial results for automated detection patient monitoring using deep learning ct image analysis," arXiv, 2020. [Online]. Available: https://arxiv.org/abs/2003.05037
[5] S. Wang, B. Kang, J. Ma, X. Zeng, M. Xiao, J. Guo, M. Cai, J. Yang, Y. Li, X. Meng, and B. Xu, "A deep learning algorithm using ct images to screen for corona virus disease (covid-19)," medRxiv, 2020. [Online]. Available: https://www.medrxiv.org/content/early/2020/04/24/2020.02.14.20023028
[6] A. Narin, C. Kaya, and Z. Pamuk, "Automatic detection of coronavirus disease (covid-19) using x-ray images and deep convolutional neural networks," 03 2020.
[7] E. E.-D. Hemdan, M. A. Shouman, and M. E. Karar, "Covidx-net: A framework of deep learning classifiers to diagnose covid19 in x-ray images," arXiv, 2020. [Online]. Available: https: //arxiv.org/abs/2003.11055
[8] P. K. Sethy and S. K. Behera, "Detection of coronavirus disease (covid19) based on deep features," Preprints, vol. 2020030300, p. 2020, 2020
[9] Wang, L., Lin, Z.Q. Wong, A. COVID-Net: a tailored deep convolutional neural network design for detection of COVID-19 cases from chest X-ray images. Sci Rep 10, 19549 (2020). https://doi.org/10.1038/s41598-020-76550-z
[10] Muneeb ul Hassan, VGG16 - Convolutional Network for Classification and Detection
[11] Purva Huilgol, Top 4 Pre-Trained Models for Image Classification with Python Code
[12] Ferhat Culfaz, Transfer Learning using Mobilenet and Keras

APPENDIX

Overleaf Project and Google Colaboratory Notebooks:

- Project Final report.
- Data prepossessing.
- Segmentation of data-set.
- K-Fold Prediction models.
- Final Prediction models.