Detection of the Covid-19 using the X-ray images

Saleh Sereshki, Ceren Sevinc, Mohammad Reza Zare Shahneh, Sharzad Haji Amin Shirazi

I. Introduction

The coronavirus (COVID-19) pandemic has infected more than 7 million people and caused a death toll so far above $400,000$ and it appears to be staying with us for the foreseeable future. Real time polymerase chain reaction (RT-PCR) diagnostic test is the primary testing method to detect COVID-19-positive patients. Even though the analytical sensitivity of the RT-PCR test is high, the clinical sensitivity can be as low as 59% [1] due to difficulties in collecting good samples. Such difficulties stem from the intricacies of sample collection or from insufficient number of viruses in the areas where samples are collected, which usually happens in the later stages of the disease where most of the viruses are in the lungs. Moreover, many developing and underdeveloped countries are not able to conduct PCR tests widely since it requires relatively expensive laboratory equipment and supplies. However, almost every country has access to X-ray devices. In many works like [1],[2], it is shown that CT scans or X-ray images can provide a good alternative to PCR tests with high sensitivity to detect COVID-19-positive patients. In this project, our main objective is to detect positive COVID-19 cases and differentiate them from viral pneumonia cases while maximizing accuracy and sensitivity (a.k.a. recall) by only using chest X-ray images and deep learning techniques.

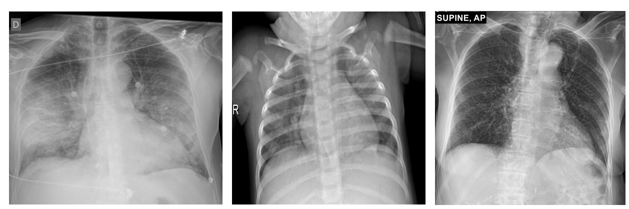
Detection of COVID-19 with deep learning has been the focus of many recent works [3], [4], [5], [6], [7] to list a few. Also, a review of artificial intelligence techniques in imaging data acquisition, segmentation and diagnosis for COVID-19 can be found in [8]. We refer the interested readers to this work and the references therein for further literature review. The most commonly used dataset in these works is provided in [9], which consists of 109 COVID-19 positive chest x-ray images. A recent work [7] provides more samples of COVID-19 positive chest x-ray images, thus we used the latter dataset in our work.

In this project, we use models of VGG16, ResNet50, and DenseNet121 with hyper-parameter optimization and AdaBoost algorithm, and InceptionV3 with transfer learning. Works in [10] and [11] have also used transfer learning for the detection of COVID-19.

This report is organized as follows. Section II introduces dataset and pre-processing techniques that were used. Section III describes the transfer learning approach, while Section IV discusses hyper-parameter optimization and the AdaBoost algorithm. Section V summarizes the results and concluding remarks are given in Section VI.

II. Data and Pre-processing

The dataset in [7] consists of 219 COVID-19, 1345 viral pneumonia, and 1341 normal chest x-ray images. Fig. 1 illustrates an example from each class.



COVID-19 Normal Pneumonia

Fig. 1: Classes in the dataset

We also test our models with augmented data by using horizontal flip, shift, rotation, shear, and zoom. The dataset is split for training, validation and testing as shown in TABLE I.

	Raw Dataset			Augmented Dataset		
Class Label	Train	Valid.	Test	Train	Valid.	Test
COVID-19	141	35	43	2820	700	43
Normal	858	215	268	2574	645	268
Pneumania	861	215	269	2583	645	269

TABLE I: COVID-19 dataset with class distributions.

III. Transfer Learning

A. XGBoost

We chose to work on XGBoost with transfer learning. Multiple models were used to see which was best with XGBoost. Boosting is an ensemble technique where new models are added to correct the errors made by existing models. Models are added sequentially until no further improvements can be made. XGBoost performs as well as a classifier for data with relatively high dimensions compared to the size of the dataset. We also applied principal component analysis (PCA) on the output of the feature extractor of the InceptionV3, which was the best convolutional neural network (CNN) model that worked with XGBoost. A block diagram of a CNN is shown in Fig. 2.

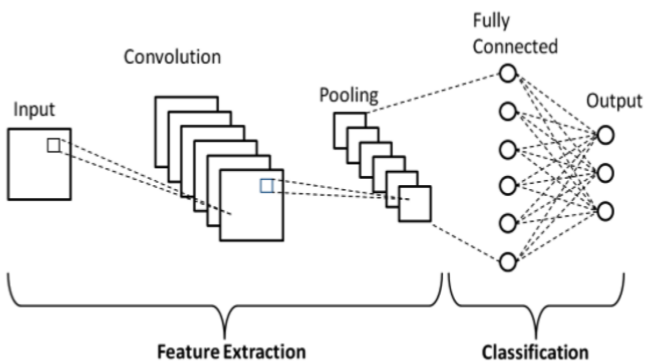


Fig. 2: Block diagram of a CNN

In [12], a novel method combining CNN and the XGBoost classifier is presented. CNNs are known to be successful at feature extraction by using filters for softening, edge detection, etc. XGBoost is known as a highly efficient and accurate gradient boosting classifier, especially for sparse datasets. Gradient boosting is a model of prediction in the form of ensemble models, which can transform several weak classifiers into a strong classifier for better classification performance. Therefore, combining CNN and XGBoost yields favorable results for sparse datasets, as is the case herein after feature selection. For the XGBoost we used a max depth 3 for the trees and number of estimators was 200.

B. PCA

We also applied PCA on the output of the feature extractor part of InceptionV3. The main idea of PCA is to reduce the dimensionality of a dataset consisting of many variables correlated with each other, either heavily or lightly, while retaining the variation present in the dataset, up to the maximum extent. The setup for the model is the same as XGBoost: we use the same 2000 samples and reshape them to 256×256 with data augmentation. After flattening the feature extractor data as part of XGBoost, we obtained $73,728$-dimensional data. Since the size of the dataset is small compared to the dimension of the data, we reduced the dimension of the data to 59 using PCA, in the hope of obtaining a more adequate dataset to train a classifier. A k-nearest neighbor (KNN) classifier with searching among 3 neighbors of a data point was used for comparison against XGBoost. Both the KNN and XGBoost classifiers were given the output of the PCA, with 59 PCA components.

C. Transfer Learning Results

The results of transfer learning (TL) by using PCA and KNN, PCA and XGB, and XGB are shown in TABLE II.

Method	Accuracy
PCA(59)+KNN	0.86
PCA(59)+XGB	0.84
XGB	0.88

TABLE II: Results of transfer learning

For comparison purposes, a CNN was trained to classify COVID-19 cases. We used InceptionV3 model with Adam optimizer and learning rate 0.000005 and categorical cross-entropy for the loss function. We trained our model in 40 epochs. The training and test loss and accuracy curves are shown in Fig. 3 and Fig. 4, respectively. It can be seen from Fig. 4 that InceptionV3 alone achieves 95% classification accuracy, a 7% increase over the highest transfer learning accuracy in TABLE II.

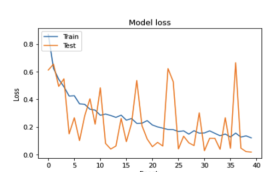


Fig. 3: Loss function

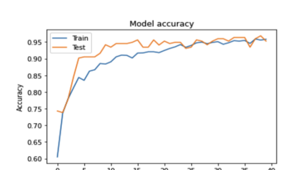


Fig. 4: Accuracy

To make sense of the performance of the feature extractor of the CNN, the first two principal components at epochs 5 and 40 are illustrated in Fig. 5 and

Fig. 6, where class 0 (red), class 1 (green), and class 2 (blue) represent normal, pneumonia, and COVID-19 cases, respectively. It can be seen from these figures that it is hard to separate between the COVID-19 and Pneumonia cases, while separation between a healthy and sick point is easier.

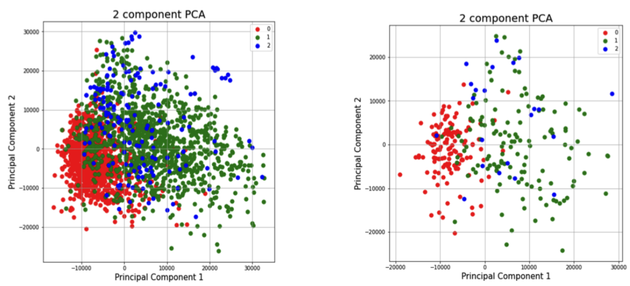


Fig. 5: Epoch 5

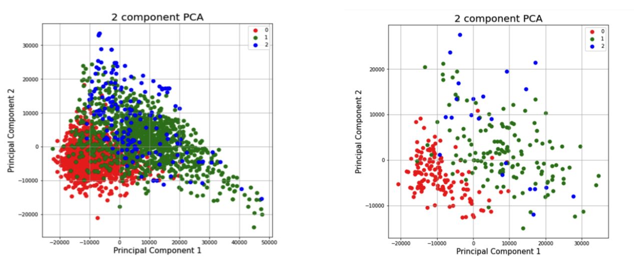


Fig. 6: Epoch 40

IV. Hyper-parameter optimization

In order to find the correct configurations for the hyper parameters of our algorithm, we used the Hyperband algorithm[13]. The hyper-parameters we were trying to optimize and the values we used for them are listed in TABLE III. For each configuration we would assign 3 epochs and then at the end of each step we would throw two-third of the configurations out.

Parameters	Range/Methods
Learning Rate	$[10^{-6}, 10^{-1}]$
Optimizer Algorithm	Adam , SGD
Momentum for SGD	$[0, 0.99]$
Augmentation Algorithms	Horizontal Flip, Shift, Shear, Zoom

TABLE III: Hyper-parameter optimization

The Adam optimizer was chosen for all of the VGG16, DenseNet121 and ResNet50. We suspect that the reason the Adam optimizer was chosen is that it has a faster movement towards the optimum at the beginning and therefore the SGD would be ruled out by the Hyperband. TABLE IV shows the hyper parameters chosen for the VGG16 and DenseNet121 algorithms.

Model	Optimizer	Learn. rate	Augmentation
VGG	Adam	0.0011	rotation
DenseNet	Adam	0.0030	rotation, sheer, zoom

TABLE IV: Hyper-parameters

A. *AdaBoost Algorithm*

An important indicator of a good COVID-19 detector is a high recall. This is due to the fact that a false negative can have catastrophic result of returning a sick person back to society. Therefore, we wanted to ensure good recall while maintaining a good accuracy. For that we used a modification of the AdaBoost algorithm [14]. As you can see in the Fig. 7, the naive implementation of the AdaBoost algorithm gives more weight to the miss-classified observation at each iteration.

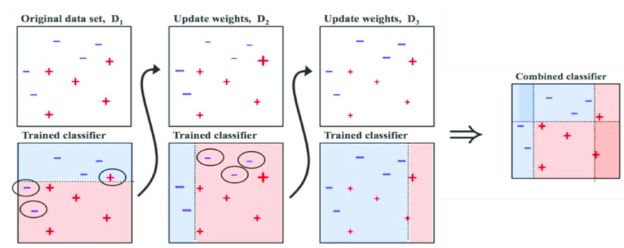


Fig. 7: AdaBoost algorithm [14]

We used our models to extract features from images (removing the detection layer of the model and using the result of the last remaining layer as features). Then, we would find at each iteration a linear classifier and move it to the point where recall becomes 100%. This way we maintain 100% recall on the training data. As shown in TABLE IX, our algorithm achieves 95% recall on COVID-19.

V. Results and Discussion

Performance results of VGG16, ResNet50, and DenseNet121 models along with hyper-parameter optimization are shown in TABLE VI, TABLE VII, and TABLE VIII, respectively. As it is clear from TABLE VII, ResNet50 model implementation performed

Study	Method	Accuracy (%)	COVID-19 Recall (%)	COVID-19 Specificity (%)
Apostolopoulos et al. [10]	MobileNet+ TL	94.72	98.66	96.46
Abbas et al. [5]	ResNet18-DeTrac	95.12	97.91	91.87
Wang et al. [6]	COVID-Net -Tailored CNN	93.3	91	-
Ucar et al. [3]	Bayes + SqueezeNet	98.3	98.3	99.1
Ozturk et al. [4]	DarkCovidNet	87.02	85.35	92.18
Chowdhury et al. [7]	AlexNet + Data Aug.	95.4	93	95.8
Chowdhury et al. [7]	ResNet18 + Data Aug.	95	95	96
Chowdhury et al. [7]	DenseNet201 + Data Aug.	96.7	96	96
Chowdhury et al. [7]	SqueezeNet + Data Aug.	98.3	96.7	99
Proposed Method	VGG16	93.45	95.35	-
Proposed Method	DenseNet201	94	95	-
Proposed Method	AdaBoost	97	95	-

TABLE V: Performance Comparison

poorly implying that it is open for further improvements.

VGG16	Precision	Recall.	F-1	Accuracy
COVID-19	0.91	0.95	0.93	0.9345
Normal	0.91	0.97	0.94	
Pneumania	0.96	0.90	0.93	

TABLE VI: Results of VGG16 model.

ResNet50	Precision	Recall.	F-1	Accuracy
COVID-19	0.11	1	0.2	0.38
Normal	0.99	0.49	0.66	
Pneumania	0.68	0.17	0.27	

TABLE VII: Results of ResNet50 model.

DenseNet121	Precision	Recall.	F-1	Accuracy
COVID-19	0.89	0.95	0.92	0.94
Normal	0.92	0.98	0.95	
Pneumania	0.98	0.90	0.94	

TABLE VIII: Results of DenseNet121 model.

Performance result of the AdaBoost algorithm is encapsulated in TABLE IX. Using the AdaBoost algorithm evidently improved the results that were obtained with VGG16, ResNet50, and DenseNet121 models. The highest accuracy, i.e. 97% is achieved by AdaBoost algorithm.

AdaBoost	Precision	Recall.	F-1	Accuracy
COVID-19	1	0.95	0.98	0.97
Normal	0.97	0.97	0.97	
Pneumania	0.97	0.97	0.97	

TABLE IX: Results of AdaBoost algorithm.

In order to compare our results with some of the successful previous works, we summarized them along with our results in TABLE V. There are only 2 results that achieve higher accuracy than our best performing proposed method. Both of these methods [3], [7] use SqueezeNet as a model and achieve an accuracy of 98.3%, while the proposed AdaBoost algorithm achieves an accuracy of 97%. Since the AdaBoost algorithm increased the accuracy of our base models, we expect the SqueezeNet model from [3] and [7] to improve when combined with the AdaBoost algorithm.

VI. Conclusion

In this project, we explored several approaches to detect COVID-19 cases from chest X-ray images. Among the approaches and models we implemented, the AdaBoost algorithm achieved the highest accuracy. A major limiting factor on the achievable accuracy and the choice of models is the size of the dataset. With more data, our current models can be trained to detect COVID-19 cases with higher accuracy. Furthermore, more sophisticated approaches may be employed with a larger dataset, such as self-supervised learning. A preliminary implementation of a self-supervised algorithm yielded poor accuracy with the existing dataset. In the future, as the dataset grows, we expect self-supervised learning approaches to perform adequately, as it already does in other medical applications.

References

[1] T. Ai, Z. Yang, H. Hou, C. Zhan, C. Chen, W. Lv, Q. Tao, Z. Sun, and L. Xia, “Correlation of chest ct and rt-pcr testing in coronavirus disease 2019 (covid-19) in china: A report of 1014 cases,” *Radiology*, p. 200642, 02 2020.

[2] J.-L. He, L. Luo, Z.-D. Luo, J.-X. Lyu, M.-Y. Ng, X.-P. Shen, and Z. Wen, “Diagnostic performance between ct and initial real-time rt-pcr for clinically suspected 2019 coronavirus disease (covid-19) patients outside wuhan, china,” *Respiratory Medicine*, vol. 168, p. 105980, 04 2020.

[3] F. Ucar and D. Korkmaz, “Covidiagnosis-net: Deep bayes-squeezenet based diagnosis of the coronavirus disease 2019 (covid-19) from x-ray images,” *Medical Hypotheses*, vol. 140, p. 109761, 2020. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0306987720307702

[4] T. Ozturk, M. Talo, E. A. Yildirim, U. B. Baloglu, O. Yildirim, and U. R. Acharya], “Automated detection of covid-19 cases using deep neural networks with x-ray images,” *Computers in Biology and Medicine*, vol. 121, p. 103792, 2020. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0010482520301621

[5] A. Abbas, M. M. Abdelsamea, and M. M. Gaber, “Classification of covid-19 in chest x-ray images using detrac deep convolutional neural network,” 2020.

[6] L. Wang and A. Wong, “Covid-net: A tailored deep convolutional neural network design for detection of covid-19 cases from chest x-ray images,” 2020.

[7] 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, and M. B. I. Reaz, “Can ai help in screening viral and covid-19 pneumonia?” 2020.

[8] F. Shi, J. Wang, J. Shi, Z. Wu, Q. Wang, Z. Tang, K. He, Y. Shi, and D. Shen, “Review of artificial intelligence techniques in imaging data acquisition, segmentation and diagnosis for covid-19,” *IEEE Reviews in Biomedical Engineering*, pp. 1–1, 2020.

[9] J. P. Cohen, P. Morrison, and L. Dao, “Covid-19 image data collection,” *arXiv 2003.11597*, 2020. [Online]. Available: https://github.com/ieee8023/covid-chestxray-dataset

[10] I. Apostolopoulos and M. Tzani, “Covid-19: Automatic detection from x-ray images utilizing transfer learning with convolutional neural networks,” *Australasian physical and engineering sciences in medicine / supported by the Australasian College of Physical Scientists in Medicine and the Australasian Association of Physical Sciences in Medicine*, 03 2020.

[11] M. S. S. Y. Shervin Minaee, Rahele Kafieh and G. J. Soufi, “Deep-covid: Predicting covid-19 from chest x-ray images using deep transfer learning,” *arXiv 2004.09363*, 2020.

[12] X. Ren, H. Guo, S. Li, S. Wang, and J. Li, “A novel image classification method with cnn-xgboost model,” in *International Workshop on Digital Watermarking*. Springer, 2017, pp. 378–390.

[13] L. Li, K. G. Jamieson, G. DeSalvo, A. Rostamizadeh, and A. Talwalkar, “Efficient hyperparameter optimization and infinitely many armed bandits,” *CoRR*, vol. abs/1603.06560, 2016. [Online]. Available: http://arxiv.org/abs/1603.06560

[14] R. E. Schapire, *Explaining AdaBoost*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2013, pp. 37–52. [Online]. Available: https://doi.org/10.1007/978-3-642-41136-6_5