APPLYING MACHINE LEARNING TO TACKLE COVID -19

The severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) pandemic that started in Wuhan, China in December 2019, has ravaged the world. This disease has already resulted in the death of approximately 3 million people worldwide[1] and it continues to have a devastating effect on the health and well-being of the global population. Our consultancy, Deep ImageJ is taking a stand against COVID-19. We are harnessing the incredible power of artificial intelligence, specifically deep learning, and taking critical steps to fight COVID-19. We are doing this by using chest X-ray (CXR) images to build and deploy highly reliable deep learning models including end-to-end convolutional neural network (CNN) model as well as five other transfer learning models, MobilNet, MobilNetV2, VGG-16, VGG-19, Inception V3 so that the radiologists can predict promptly as well as accurately whether their patient has COVID-19 or not.

The role of chest radiography in confirming COVID-19

Introduction

The COVID-19 pandemic continues to have a devastating effect on the health and well-being of the global population. A critical step in the fight against COVID-19 is effective screening of infected patients, with one of the key screening approaches being radiology examination using chest X- ray images. It was found in early studies that patients that show abnormalities in chest radiography images are characteristic of those infected with COVID-19[21]. Motivated by this and inspired by the work done by researchers so far, here we have deployed six models, namely convolutional neural network (CNN) model, MobileNet, MobileNetV2, VGG-16, VGG-19, Inception V3 for detecting COVID-19 from chest X-ray (CXR) images. The objective behind developing these models is the development of highly accurate yet practical deep learning solutions that can be helpful to radiologists in the early diagnosis of this disease while making sure that people are able to get the right treatment quickly which is indispensable in preventing the spread of COVID-19. This will also ensure that people are not tested unnecessarily using invasive tests, or loaded with medication, or unnecessarily isolated.

Methodology

Dataset: As discussed in the midterm report and keynote, the CNN model was unable to generalize to new images. Consequently, the size of the training dataset was increased from 1200 images to 1900 images. The public dataset shared on the GitHub website by Joseph Paul Cohen [19] was utilized for increasing the size of the training dataset. Also, the split ratio of the training, validation, and test datasets was updated to reflect a better splitting ratio of 75%,15%, and 10% respectively. This dataset consists of CXR images belonging of two classes, normal (1650 images) and COVID-19 (920 images). The CXR images belonging to both these classes were equally split into test, training, and validation datasets.

Data Preprocessing: The data preprocessing steps employed during the midterm phase were kept the same for all six deep learning models. Keras Image Data Generator's inbuilt image augmentation functionality was leveraged with the following augmentation parameters: rotation range (+/-0.2°), vertical flip, the zoom range of 0.2, shear range of 0.2. Rescaling was performed for all the images in test, training, and validation set by dividing the image sizes by 255 [39]. Keras Image Data Generator's inbuilt image method, flow_from_directory() was utilized for setting the input image size of 255 x 255, setting the batch size of 64, shuffling the images, as well as setting the class mode of 'binary'.

Models deployed: CNN and five other transfer learning models, MobilNet, MobilNetV2, VGG-16, VGG-19, Inception V3 were utilized to address this supervised image classification problem. The choice of these six models was based on the research already conducted by data scientists for this particular image classification problem[25][39][32][35].

- The CNN model deployed was a sequential model consisting of 3 convolutional layers, followed by 3 dropout layers, 3 max-pooling layers, and 2 dense layers at the end.
- All five transfer learning models deployed were initially built on the pre-trained ImageNet dataset. For the exact layout of all the models refer to the code posted on my Github profile [40].

<u>Building the model:</u> For all the models the following hyperparameters and metrics were employed:

- 'Adam' was used as an optimizer with a learning rate of 0.0001.
- 'Sigmoid' was used as the activation function for the output layer. ReLU was employed as activation functions for the middle layers to avoid forwarding any negative values through the network[35].
- Here the system performance was evaluated mainly by using accuracy and recall[33][34]. Furthermore, precision and F1-score were also computed for each class (COVID-19, and Normal).
- Epoch size of 500, batch size of 64, and patience of 5 were used to train the model.
- Three callbacks, namely Early Stopping, Monitor, and Learning Rate Scheduler were defined for all the models to further prevent any issues with regard to overfitting.

Results

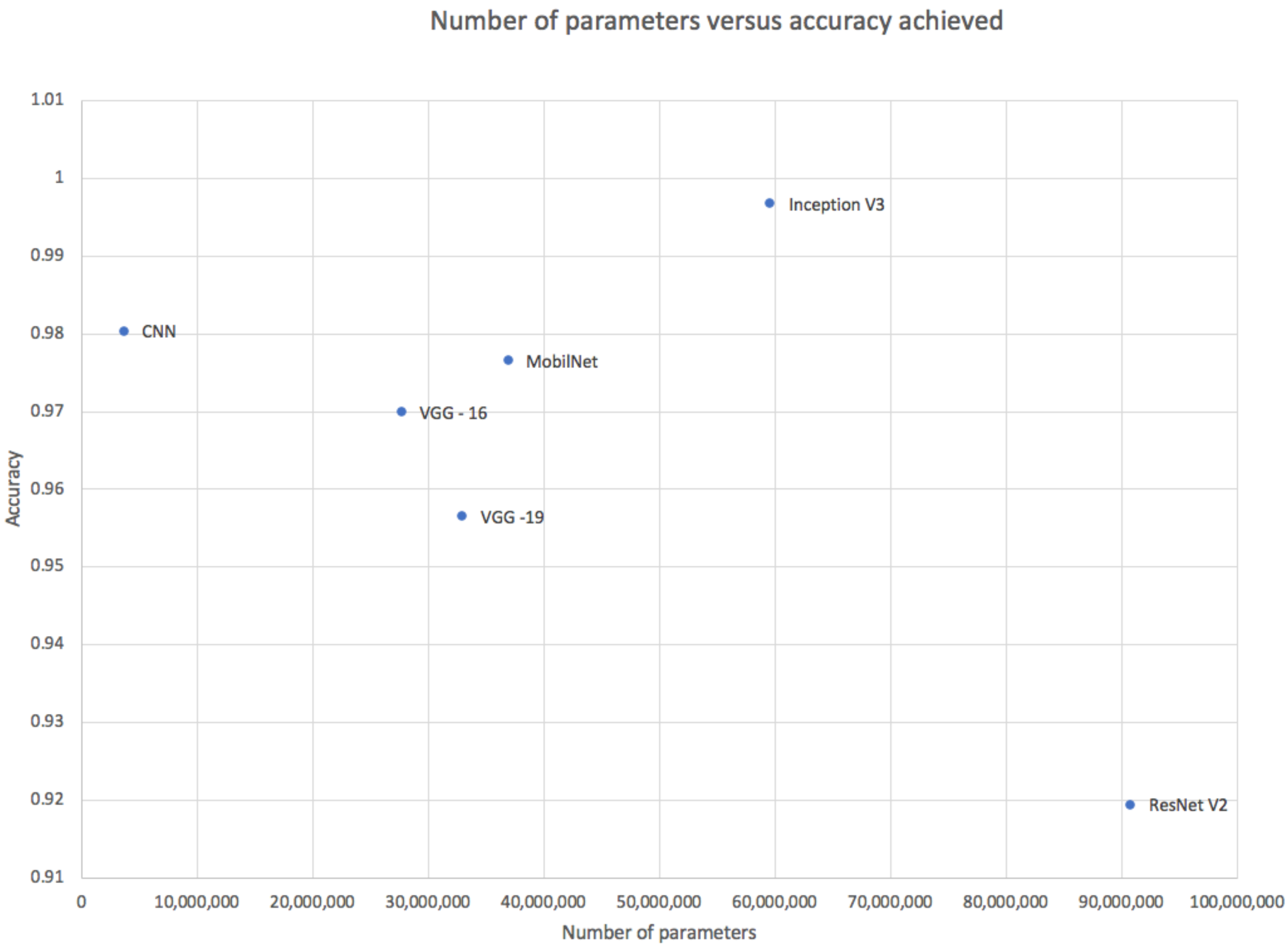

Figure 4: The scatter plot shows the number of parameters versus the accuracy achieved by the various models

The higher the number of parameters the longer the time required to train and test the model. The time required to test the images on the trained model i.e. run time is an important factor for COVID-19 detection especially considering the density in healthcare facilities during the pandemic.

The scatter plot above in Figure 3 shows a comparison of the number of parameters and the accuracy of different models. As expected the CNN model has the quickest runtime as it has the lowest count of the parameters needed to train the model. While comparing the VGG -16 and the MobileNet models, the accuracy of the MobileNet model increases by about 1% with the addition of 8 million parameters. Similarly, comparing the Inception V3 and the MobileNet models the accuracy of the Inception V3 model increases by 2% with the addition of approximately 25 million parameters.

In conclusion, as we can see in Figure 4 the best performing model is the Inception V3 model with approximately 60 million parameters and an accuracy of 99.7% on the test dataset.

Table 1: Comparison of performance metrics and loss for all six models

Model	Accuracy	Loss	Weighted Avg Recall
MobilNet	97.7%	0.246	0.73
CNN	98.0%	0.069	0.48
VGG-16	96.9%	0.088	0.66
VGG-19	95.7%	0.146	0.72
Inception V3	99.7%	0.019	0.71
ResNet V2	91.9%	0.214	0.67

The above table depicts the loss as well as the performance metrics of all the six models in terms of recall, precision, F1-score, and accuracy. The highest accuracy obtained on the test data is with the Inception V3 model. On the other hand, ResNet V2 performs worst as compared to all the other models. Classifying a patient with COVID-19 as normal is a more serious mistake than classifying a normal patient as a COVID-19 (False Positives) patient. This may lead to the rapid spread of COVID-19, especially when its contagiousness is considered. Therefore, a COVID-19 classifier needs to have a high recall. As seen above in Table 1, better recall is achieved with MobileNet, Inception V3, and VGG-19 models as compared to the other models.

Overall, keeping in mind that the best performing model needs to have the highest accuracy and recall, the best performing models here are the Inception V3 model and MobileNet model.

Model performance curves

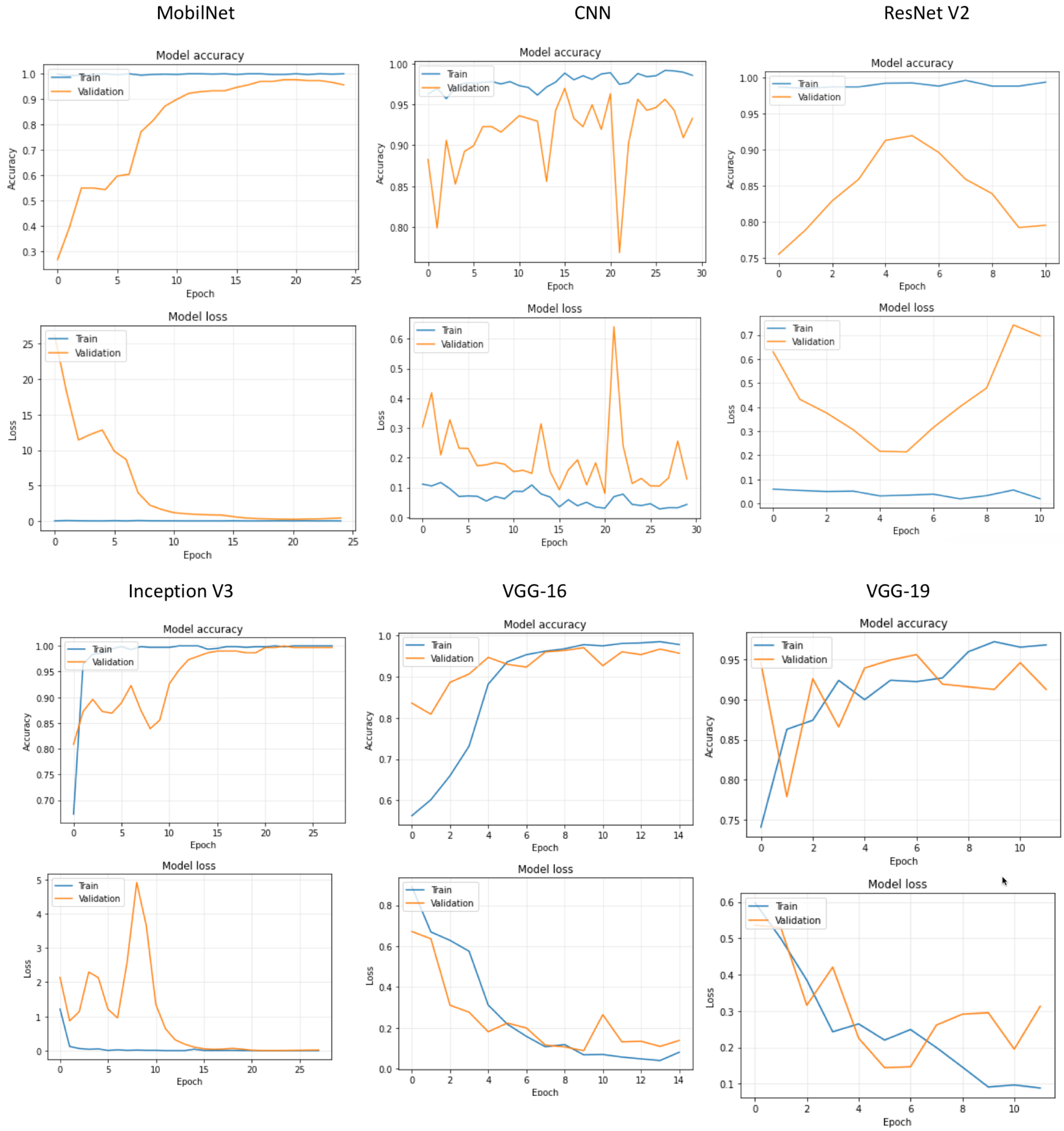

Figure 5: Performance of all six models using the CXR dataset

Here is the summary of our model performance curves.

- The Resnet V2 model performs the worst among all the models and it is able to achieve the highest validation accuracy of 92% and the lowest loss of 0.21 on the validation dataset. A difference of 5% in terms of accuracy for the validation and training is considered a sign of overfitting and this model has a problem with overfitting i.e. it is performing very well on the training dataset but it is unable to generalize well on the validation and test dataset.

- CNN, VGG-16, and VGG-19 models perform very well by demonstrating an average accuracy of 92% and an average loss of 0.15 on the validation dataset. Also, none of these models show a sign of overfitting.
- MobilNet, and Inception V2 model outperform the other four models by demonstrating an accuracy of 99% and a loss of 0.0001 on the validation dataset. Also, unlike the ResNet V2 model, there is no issue of overfitting for both these models and they are able to generalize well on the validation dataset.

Discussion

Consequently, both Inception V2 and MobileNet are good models for predicting whether a person has COVID-19 or not on this dataset mainly because of the following reasons:

- They yield a high recall and accuracy.
- They have low loss and are able to generalize well to unseen datasets.
- They both are lightweight models with few parameters that offer a fast screening tool for diagnosing COVID-19.

Out of the two models, we propose MobileNet model should be used for predicting COVID-19 in the future based on the CXR images of this dataset mainly because along with high accuracy and recall, this model demonstrates the lowest runtime i.e. it will be able to detect whether a person has COVID-19 or not at a much faster rate than Inception V2.

Conclusion

MobileNet model using the pre-trained ImageNet dataset is the best performing model for detecting COVID-19 from medical images in this case.

References:

1. https://coronavirus.jhu.edu/map.html
2.Yang, Z., & Zeng, Z. (2020). Modified SEIR and AI prediction of the epidemics trend of COVID-19 in China under public health interventions. *Journal of Thoracic Disease, 12*. View at: Publisher Site | Google Scholar
3. Ratnabali, P., Arif Ahmed, S., & Samarjit, K. (2020).Neural Network Based Country Wise Risk Prediction of COVID-19. *Applied Sciences, 10*. doi:10.3390/app10186448. View at: Publisher Site | Google Scholar
4.Arulprakash E, Aruldoss M. A(2020). Study on Fight Against COVID-19 from Latest Technological Intervention. SN Computer Science, 277, doi:10.1007/s42979-020-00301-0. View at: Publisher Site | Google Scholar
5. https://www.arcgis.com/apps/opsdashboard/index.html#/bda7594740fd40299423467b48e9ecf6
6. https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/L20LOT
7. Dong E, Du H, Gardner L. An interactive web-based dashboard to track COVID-19 in real time. Lancet Inf Dis. 20(5):533-534. doi: 10.1016/S1473-3099(20)30120-1. View at: Publisher Site | Google Scholar
8. Z. (n.d.). Cooperation or Confrontation? —— Prediction of COVID-19 Pandemic Situation via Deep Learning. doi:10.21203/rs.3.rs-128767/v1. View at: Publisher Site | Google Scholar
9. Novanto, Y. (2020). COVID-19 growth prediction using multivariate long short term memory. *IAENG International Journal of Computer Science, 47*, 829-837. View at: Publisher Site | Google Scholar

10. Shahid, {., Zameer, A., & Muneeb},, M. (2020). Predictions for COVID-19 with deep learning models of LSTM, GRU and Bi-LSTM. Chaos, Solitons & Fractals, 140, 110212. Neural Network Based Country Wise Risk Prediction of COVID-19 doi:https://doi.org/10.1016/j.chaos.2020.110212. View at: Publisher Site | Google Scholar
11. Sepp, H., & Jürgen, S. (1997). Long Short-term Memory. *Neural Computation, 9*, 1735-80. doi:10.1162/neco.1997.9.8.1735. View at: Publisher Site | Google Scholar
12. Sourabh, S., Kuljeet, S., Sachin, K., & Paramjit, K. (n.d.). Deep-LSTM ensemble framework to forecast Covid-19: An insight to the global pandemic. International Journal of Information Technology. doi:10.1007/s41870-020-00571-0. View at: Publisher Site | Google Scholar
13.Arora, A., Kumar, H., & Panigrahi, B. K. (n.d.). Prediction and analysis of COVID-19 positive cases using deep learning models: A descriptive case study of India. Chaos Solitons Fractals, 139, 110017. doi:10.1016/j.chaos.2020.110017. View at: Publisher Site | Google Scholar
14. Car, Z., Šegota, S. B., & Anđelić, N. (2020). Modeling the Spread of COVID-19 Infection Using a Multilayer Perceptron. Computational and Mathematical Methods in Medicine, 2020. doi:https://doi.org/10.1155/2020/5714714. View at: Publisher Site | Google Scholar
15. Hu, T., Guan, W., Zhu, S. (2020). Building an Open Resources Repository for COVID-19 Research, Data and Information Management. doi: https://doi.org/10.2478/dim-2020-0012.View at: Publisher Site | Google Scholar
16. Satyaki, R., & Preetam, G. (2020). Factors affecting COVID-19 infected and death rates inform lockdown-related policymaking. *PLOS ONE, 15*, 1-18. doi:10.1371/journal.pone.0241165. View at: Publisher Site | Google Scholar
17. Dandekar, R., & Barbastathis, G. (2020). Quantifying the effect of quarantine control in Covid-19 infectious spread using machine learning. *MedRxiv*. doi:https://doi.org/10.1101/2020.04.03.20052084 View at: Publisher Site | Google Scholar
18. Mathew, J., Behera, R. K., Panthakkalakath, Z. E., & N. (2020). A Deep Learning Framework for COVID Outbreak Prediction. ArXiv, Abs/2010.00382. View at: Publisher Site | Google Scholar
19. https://github.com/ieee8023/covid-chestxray-dataset
20. F. Shi, J. Wang, J. Shi et al., "Review of artificial intelligence techniques in imaging data acquisition, segmentation and diagnosis for COVID-19," *IEEE Reviews in Biomedical Engineering*, 2020.View at: Publisher Site | Google Scholar
21. H. Bai, B. Hsieh, Z. Xiong et al., "Performance of radiologists in differentiating COVID-19 from viral pneumonia on chest CT," *Radiology*, vol. 296, no. 2, pp. E46–E54, 2020.View at: Publisher Site | Google Scholar
22. Y. Fang, H. Zhang, J. Xie et al., "Sensitivity of chest CT for COVID-19: comparison to RT-PCR," *Radiology*, vol. 296, no. 2, pp. E115–E117, 2020.View at: Publisher Site | Google Scholar
23. S. S. Hare, A. N. Tavare, and V. Dattani, "Validation of the British Society of Thoracic Imaging guidelines for COVID-19 chest radiograph reporting," *Clinical Radiology*, vol. 75, no. 9, 2020.View at: Publisher Site | Google Scholar
24. F. Pesapane, M. Codari, and F. Sardanelli, "Artificial intelligence in medical imaging: threat or opportunity? Radiologists again at the forefront of innovation in medicine," *European Radiology Experimental*, vol. 2, no. 1, 2018.View at: Publisher Site | Google Scholar
25. L. Li, L. Qin, Z. Xu et al., "Using Artificial intelligence to Detect COVID-19 and Community-acquired pneumonia Based on Pulmonary CT: Evaluation of the Diagnostic Accuracy," *Radiology*, vol. 296, no. 2, pp. E65–E71, 2020.View at: Publisher Site | Google Scholar
26. M. Z. Alom, M. M. S. Rahman, M. S. Nasrin, T. M. Taha, and V. K. Asari, "COVID MTNet: COVID-19 detection with multi-task deep learning approaches," *arXiv preprint arXiv*, 2004, https://arxiv.org/abs/2004.03747.
27. R. Hu, G. Ruan, S. Xiang, M. Huang, Q. Liang, and J. Li, "Automated Diagnosis of COVID-19 Using Deep Learning and Data Augmentation on Chest CT," *medRxiv*, 2020.

28. Ai T, Yang Z, Hou H et al. Correlation of Chest CT and RT-PCR Testing in Coronavirus Disease 2019 (COVID-19) in China: A Report of 1014 Cases. **Radiology** 2020 Feb 26:200642 [Epub ahead of print]. Link, Google Scholar
30. Fang Y, Zhang H, Xie J et al. Sensitivity of Chest CT for COVID-19: Comparison to RT-PCR. Radiology 2020 Feb 19:200432 [Epub ahead of print]. Link, Google Scholar
31. Chung M, Bernheim A, Mei X et al. CT Imaging Features of 2019 Novel Coronavirus (2019-nCoV). Radiology 2020;295(1):202–207. Link, Google Scholar
32. Huang C, Wang Y, Li X et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. Lancet 2020;395(10223):497–506 [Published correction appears in Lancet 2020;395(10223):496.]. Crossref, Medline, Google Scholar
33. Kermany DS, Goldbaum M, Cai W et al. Identifying Medical Diagnoses and Treatable Diseases by Image-Based Deep Learning. Cell 2018;172(5):1122–1131.e9. Crossref, Medline, Google Scholar
34. Rajaraman S, Candemir S, Kim I et al. Visualization and interpretation of convolutional neural network predictions in detecting pneumonia in pediatric chest radiographs. Appl Sci (Basel) 2018;8(10):1715. Crossref, Medline, Google Scholar
35. Depeursinge A, Chin AS, Leung AN et al. Automated classification of usual interstitial pneumonia using regional volumetric texture analysis in high-resolution computed tomography. Invest Radiol 2015;50(4):261–267. Crossref, Medline, Google Scholar
36. Anthimopoulos M, Christodoulidis S, Ebner L, Christe A, Mougiakakou S. Lung pattern classification for interstitial lung diseases using a deep convolutional neural network. IEEE Trans Med Imaging 2016;35(5):1207–1216. Crossref, Medline, Google Scholar
37. He K, Zhang X, Ren S et al. Deep residual learning for image recognition. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016. Google Scholar
38. Ronneberger O, Philipp F, Thomas B. U-net: Convolutional networks for biomedical image segmentation. In: Navab N, Hornegger J, Wells W, Frangi A, eds. Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015. MICCAI 2015. Lecture Notes in Computer Science, vol 9351. Cham, Switzerland: Springer, 2015; 234–241. https://doi.org/10.1007/978-3-319-24574-4_28. Crossref, Google Scholar
39. Maurya, A.. "Predicting intubation support requirement of patients using Chest X-ray with Deep Representation Learning." *ArXiv* abs/2011.01787 (2020): n. pag.
40. https://github.com/aditimukerjee/DL_COVID-19