DETECTION METHODS OF COVID-19 UTILIZING CONVOLUTIONAL NEURAL NETWORKS WITH DIFFERENT LAYERS MAXPOOLING 2D AND ITS IMPLIMENTATION: A STUDY

MANAV RACHNA INTERNATIONAL INSTITUTE OFRESEARCH AND STUDIES

AkashVerma and MugdhaJaiswal

Abstract:

This study will explore various convolutionary neural network architectures tested on posteroanterior chest X-rays of several stable, COVID-19 diagnosed patients and other forms of pneumonia, including Extreme Acute Respiratory Syndrome, Streptococcus, Klebsiella, Legionellosis, Pneumocystis Jiroveci Pneumonia, Acute Respiratory Distress Syndrome not triggered by COVID-19 In particular , this study explores the deep convolutionary neural networks with 2D maxpooling, various layers using VGG16 and VGG19 as provided by earlier researchers, all of whom followed by a multi-layer perceptron and a final 30% drop-out. The study found that VGG16 is the best performing network, rather than any other.

Introduction

Around in between 1.2 percent and 5.4 percent of the population in March 2020 had been diagnosed with COVID-19 in the United Kingdom (i.e. between around 800,000 and 3,600,000 inhabitants of the United Kingdom) as well as in other countries such as India. Studies of antibody seroprevalence found that, between 2.24 percent and 3.37 percent of the Santa Clara population, CA contracted COVID-19 until April 4 , 2020 (Bendavid et al., 2020), and about 21 percent of the New York City population contracted the disease according to the New York governor. Estimating the total number of people infected with COVID-19 by statistical models (Ferguson et al., 2020; Flaxman et al., 2020) or by studies of seroprevalence (Bendavid et al., 2020).

The most advanced clinical trials for a potential vaccine are currently based in the United States , the United Kingdom and Germany, and such a vaccine will only be available to healthcare practitioners at the earliest in the fall of 2020. Wang et al. (2020) applied deep convolutionary neural networks on 1,065 CT images of pathogen-confirmed COVID-19 cases (325 patients) in the lungs along with those previously diagnosed with standard viral pneumonia (740 patients) not caused by coronavirus. The internal validation of that study obtained a total accuracy of 89.5% with a specificity of 88% and a sensitivity of 87%, while the external test data set showed a total accuracy of 79.3% with 83% specificity and 67% sensitivity. The first two nucleic acid test results for COVID-19 were negative in 54 patient’s part of the sample, 46 of whom were expected to be COVID-19 positive by an algorithm with a likelihood of 85.2 percent, which may further illustrate the very low accuracy of these studies. However, Ai et al. (2020) found that 75 percent of patients with negative RT-PCR results had positive chest CT findings, 48 percent of

which were regarded as highly likely COVID-19 cases. This work has two main objectives to diagnose possible pneumonia using machine learning software, and, if detected, if it is related to a coronavirus infection or another, allowing for the identification of new potential coronavirus foyers after possible 'stop-and-go' lockdowns as predicted by Ferguson et al. (2020) before a vaccine is identified and in addition, in the short term, practitioners will be able to compare the diagnosis of deep convolutional neural networks in this paper with potential RT-PCR test results {if clashing, the chest CT may be done as it is more reliable to display COVID-19 pneumonia as shown by Ai et al. (2020); Wang et al. (2020); Xu et al. (2020).

Research Methodologies

This is the work descriptive of it. We start with the Report description taken after with the secondary collection of data and their management. We will be investigating this background after data is obtained. After the following data collection approach, we will take a look at how to develop new technologies used in the engineering and medical sectors.

Data Collection Method

The data will be obtained from Cohen et al. (2020) and from Kermany et al. (2018). The FIrst source (Cohen et al ., 2020) had primarily data from ill patients and thus the paper partially uses the second data collection (Kermany et al., 2018) which consists of healthy patients only (there would be an over-representation only partially otherwise) .

Secondary Data: Secondary data collection will be done through internetpublished by the organizations; like journals, books, internet and articles.

Tool Used in Data Analysis: Methods will be implemented as profoundly convolutionary neural network follows the VGG16 architecture proposed by Simonyan and Zisserman (2015) with 2D maxpooling, various layers. Narin (2020) et al. Automatic Coronavirus Virus Detection (COVID-19) Using X-ray images and a Deep Convolutionary Neural Networks

VGG16

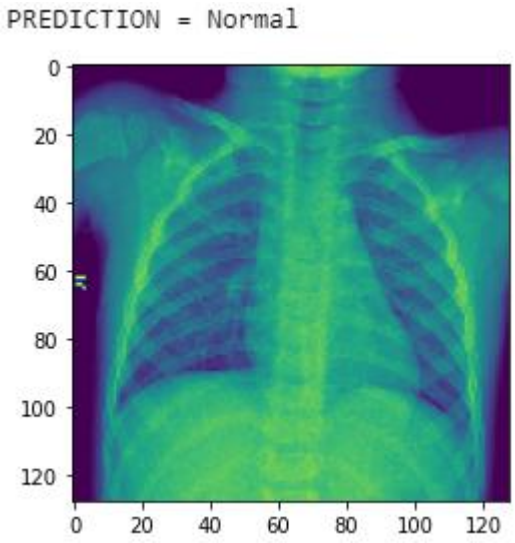
The architecture of the first called deep convolutionary neural network follows Simonyan and Zisserman 's proposed VGG16 architecture (2015), and is outlined in Table 1. Having trained more than 200 epochs on a data set of over 14 million 224x224 images belonging to 1,000 classes based on the 2014 Large Scale Visual Recognition Challenge (ILSVRC2014), this deep convolutionary neural network achieved top-5 test accuracy of 92.7 per cent. Following the ILSVRC2014 we kept the weights for pre-training in this study. The optimizer is Adam (Kingma and Lei-Ba, 2015), with a categorical loss of cross-entropy, learning rate 0.0001, and over 200 epochs worked.

Table 1: VGG16 configuration (Simonyan and Zisserman, 2015) with portion added with some rate of drop-out p. The activation function is all through the Rectified Linear Unit (ReLU). In this architecture the total number of parameters is 515,505in 182x182 photo.

Layer (type)	Output Shape	Param #
conv2d_1 (Conv2D)	(None, 128, 128, 16)	448
max_pooling2d_1 (MaxPooling2	(None, 64, 64, 16)	0
conv2d_2 (Conv2D)	(None, 64, 64, 32)	4640
max_pooling2d_2 (MaxPooling2	(None, 32, 32, 32)	0
conv2d_3 (Conv2D)	(None, 32, 32, 64)	18496
max_pooling2d_3 (MaxPooling2	(None, 16, 16, 64)	0
flatten_1 (Flatten)	(None, 16384)	0
dense_1 (Dense)	(None, 30)	491550
dropout_1 (Dropout)	(None, 30)	0
dense_2 (Dense)	(None, 10)	310
dropout_2 (Dropout)	(None, 10)	0
dense_3 (Dense)	(None, 5)	55
dense_4 (Dense)	(None, 1)	6
activation_1 (Activation)	(None, 1)	0

Total params: 515,505
Trainable params: 515,505
Non-trainable params: 0

Tables and figure show the Neural networks, convolutional neural networks, maxpooling, dense,conv2d, dropout with Total params: 515,505



We will conduct the VGG16 stratified 5-Fold cross-validation with a final drop-out of 30 per cent, and the results are as shown in Table 2.

Table 2 The VGG16 neural network's cross-validated output metrics (rounded to the thousandth) illustrated in Table 2 by Stratified 5-Fold and 200 epochs, learning levels of 0.0001. The 95 per cent confidence interval was determined by 4 degrees of freedom t-distribution.

Internal / Training Set			
Measure	**LHS 95% CI**	**Value**	**RHS 95% CI**
Loss	-0.135	0.105	0.339
Accuracy	0.905	0.939	0.973
Flat AUC	0.971	0.974	0.976
COVID-19 Recall	0.858	0.877	0.897
No Finding Recall	0.960	0.968	0.976
Other Pneumonia Recall	0.500	0.534	0.568
External / Testing Set			
Measure	**LHS 95% CI**	**Value**	**RHS 95% CI**
Loss	-0.324	0.337	0.998
Accuracy	0.706	0.841	0.976
Flat AUC	0.970	0.974	0.977
COVID-19 Recall	0.858	0.877	0.897
No Finding Recall	0.960	0.968	0.976
Other Pneumonia Recall	0.500	0.534	0.568

On the other hand, the sensitivity to pneumonia diseases other than COVID-19 is low (at 53 percent for both the training and test sets), suggesting that the neural network has separation separation difficulties between COVID-19 and other pneumoniae. Wang et al. (2020) found 67 percent sensitivity for such distinction, which is greater, but they used only two grades, COVID-

19 and Other Pneumonia. This increased sensitivity may be due to the fact that Wang et al. (2020) trained their neural networks on CT images, rather than on PA X-rays, which appear to discern pathologies better. In addition, such a mark is under-represented in our data collection, although we will find it to be indicative of data sets observed in hospital during a COVID-19 epidemic.

The Flat AUC (Area Under the ROC Curve) is not identical to the regular binary AUC, since we are dealing with a multi-class problem. Before AUC computation, we attended the data in a single mark here. In that case, every pair of label-predictions is viewed as an individual data point. This calculation will therefore not be very accurate, although it offers some details on the neural network's "overall" ROC. For both the training and research data-sets it is found to be 97:4 percent. Because of the construction of this output metric, it is easier to equate it with other networks and does not bear network significance alone {which is what we are going to do in the sections below.

InceptionResNetV2 & InceptionV3

The second architecture considered by Szegedy et al. (2016) is a very deep convolutionary neural network called In-ceptionResNetV2 and is represented in the diagram shown in Figure 1. When trained on the ImageNet data set for over 200 epochs, this very deep convolutionary neural network achieved top-5 test accuracy of 95.3 per cent. In this study, after the above preparation, we kept the weights for pre-training, using an Adam optimizer (Kingma and Lei-Ba, 2015) with a learning rate of 0.0001, a categorical cross-entropy loss and 200 epochs. Max-pooling was done in the architecture for extraction of the features.

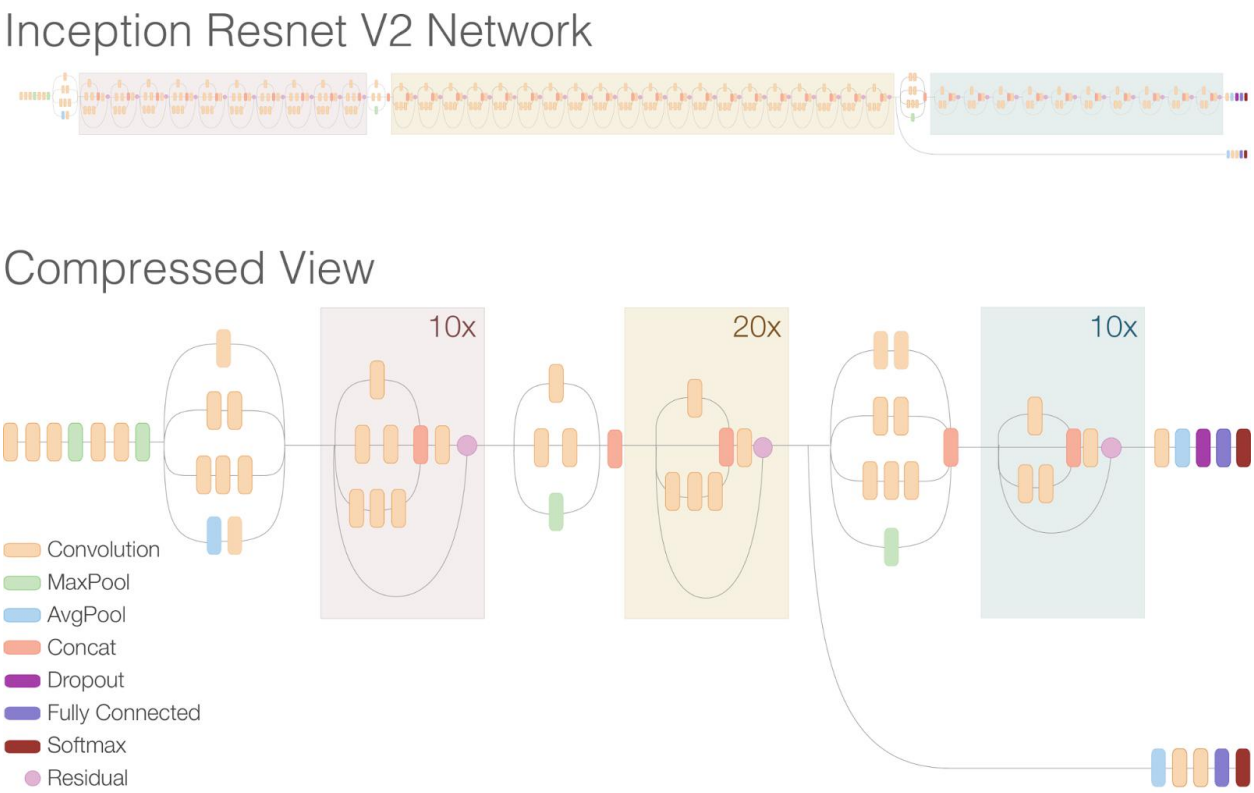


Figure 1 Definition diagram for InceptionResNetV2 architecture, as outlined in Szegedy et al. (2016). This neural network has 54,737,123 parameters on the data-set 182 x 182.

The cross-validated performance measurements obtained by Stratified 5-Fold are shown in Table 4 for the In-ceptionResNetV2 with final 30% drop-out.Here the performance results for InceptionResNetV2 appear to be somewhat disappointing, mostly because there are several batch normalizations in this architecture (Ioffe and Szegedy, 2015). The batches are normalized by mean and variance during model preparation, but in the testing process, the batches are normalized with respect to the increasing average of observed mean and variance, resulting in lower accuracies for the type of data set we are considering. In particular, for the InceptionResNetV2 network, we found COVID-19 sensitivities of 70:8(±3:5;+3:6) per cent and 71:0(±3:5) per cent respectively for internal and external validations.

It is considered fairly weak in medical terms, and still seems small even if we will take into account errors from RT-PCR testing or studies of seroprevalence as shown in Xu et al. (2020). In addition, the 95 per cent confidence interval indicated by the Stratified 5-Fold cross-validation is broad for the accuracies and sensitivities of COVID-19, suggesting greater variance. Therefore, training and testing on separate but identical data-sets is likely to result in very separate performance outcomes.

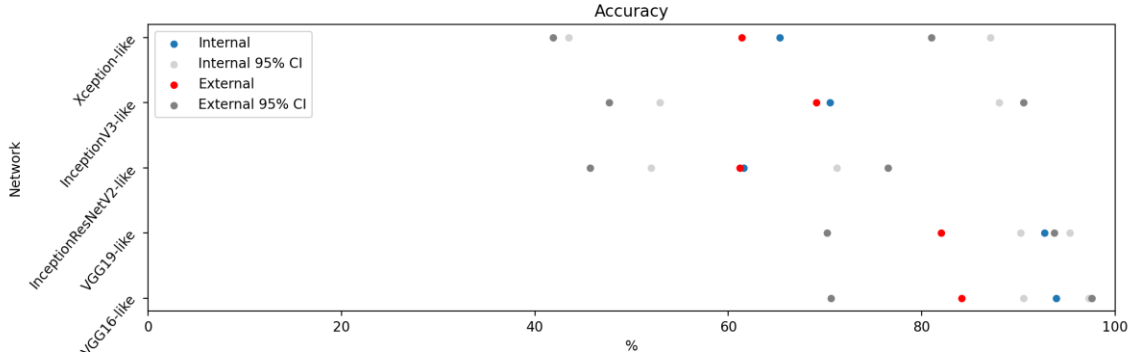
Table 3 offers a review table of some of the success indicators for the neural networks that are all trained over 200 epochs. Figure 2 further indicates some of those steps. As we can see from

Table 3 and Figure 2, VGG16 performed best on all scales, al- though VGG19 performed similarly with lower variability. In particular, for the internal data set, VGG16 had stratified 5-Fold cross-validated accuracies of 93:9(±3:4) per cent and for the external data set of 84:1(±13:5).

Table 3: Cross-validated internal and external steps (Strati ed 5-Fold) for each of the models considered, rounded to nearest one thousandth. Here, # 1 Recall refers to the neural network's sensitivity to the categorical variable COVID-19, and # 2 Recall the sensitivity to the categorical variable No Find. Networks: (1) VGG16 (Simonyan and Zisserman, 2015), (2) VGG19 (Simonyan and Zisserman, 2015), (3) InceptionResNetV2 (Szegedy et al., 2016), (4) InceptionV3 (Szegedy et al., 2016), and (5) Xception (Chollet, 2017).

-	Internal / Training Set			External / Testing Set		
Net	Accuracy	#1 Recall	#2 Recall	Accuracy	#1 Recall	#2 Recall
(1)	$93.9^{+3.4}_{-3.4}\%$	$87.7^{+2.0}_{-1.9}\%$	$96.8^{+0.8}_{-0.8}\%$	$84.1^{+13.5}_{-13.5}\%$	$87.7^{+2.0}_{-1.9}\%$	$96.8^{+0.8}_{-0.8}\%$
(2)	$92.7^{+2.6}_{-2.5}\%$	$86.1^{+1.8}_{-1.8}\%$	$96.4^{+0.5}_{-0.4}\%$	$82.0^{+11.7}_{-11.8}\%$	$86.1^{+1.8}_{-1.8}\%$	$96.4^{+0.4}_{-0.4}\%$
(3)	$61.6^{+9.6}_{-9.6}\%$	$70.8^{+3.6}_{-3.5}\%$	$91.8^{+0.7}_{-0.7}\%$	$61.2^{+15.3}_{-15.5}\%$	$71.0^{+3.5}_{-3.5}\%$	$91.8^{+0.7}_{-0.7}\%$
(4)	$70.5^{+17.5}_{-17.6}\%$	$77.9^{+9.1}_{-9.0}\%$	$94.1^{+1.2}_{-1.3}\%$	$69.1^{+21.4}_{-21.4}\%$	$78.0^{+8.9}_{-8.9}\%$	$94.1^{+1.2}_{-1.3}\%$
(5)	$65.3^{+21.8}_{-21.8}\%$	$75.0^{+8.9}_{-8.8}\%$	$89.5^{+2.9}_{-2.9}\%$	$61.4^{+19.6}_{-19.5}\%$	$75.1^{+8.8}_{-8.8}\%$	$89.5^{+2.9}_{-2.8}\%$

Remember that we must run all the training sessions with 200 epochs for comparability between the neural networks, and therefore this discrepancy between the internal and external accuracies may be clarified by overfit. In addition, the VGG16 neural network's COVID-19 sensitivity is 87:7(-1:9; 2) percent for both internal and external data sets, indicating the network 's efficiency in correctly identifying true positive COVID-19. Idem for the response to no find that is 96:8(±0:8) percent.



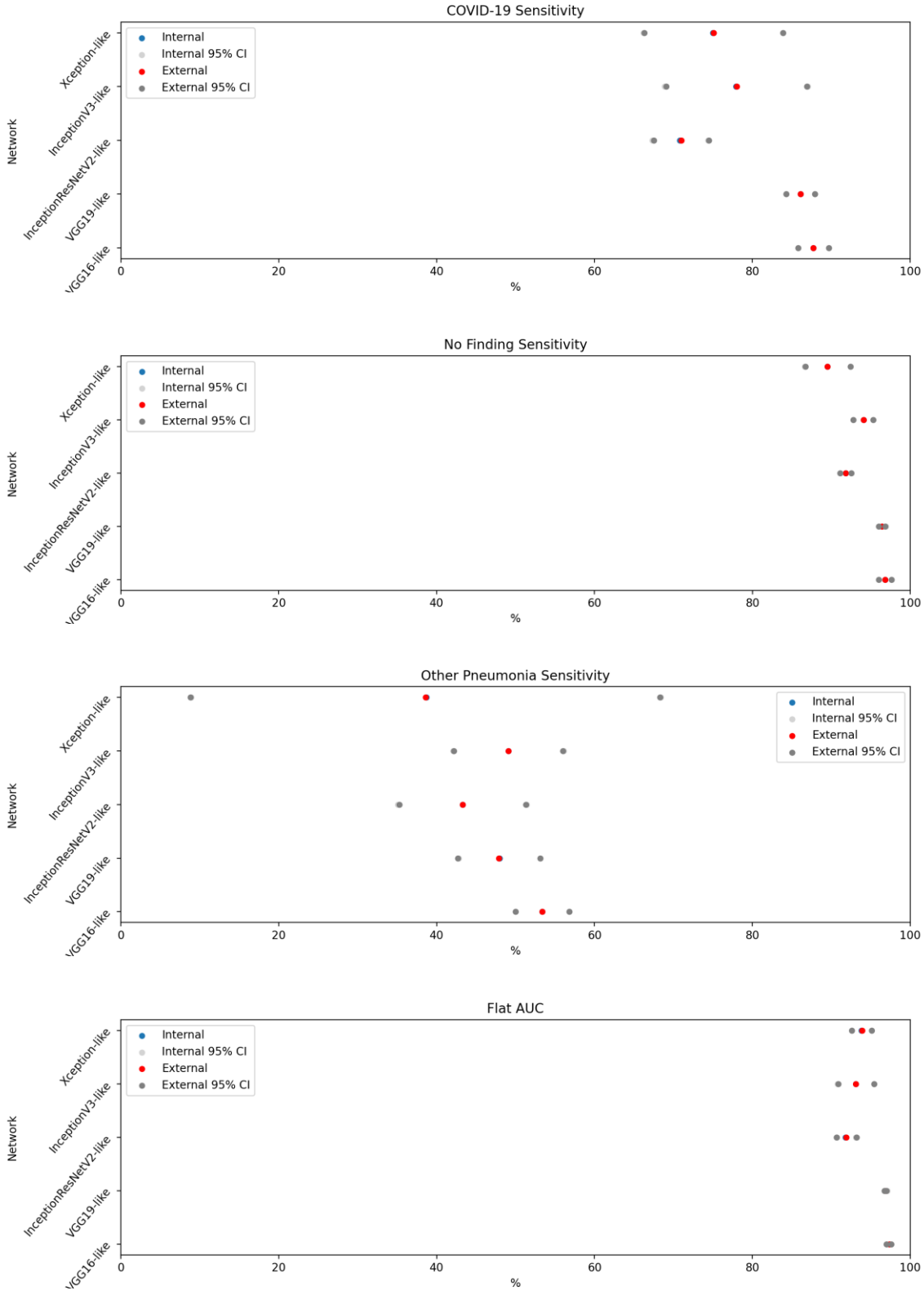


Figure 2: Output measurements of different neural networks on PA chest X-ray data collection

One research area that could be explored following this research is the development of a particular deep neural network tailored to the data set made of X-rays of the PA chest. This study will be based on the easy access to a broader database to PA chest X-rays COVID- 19 positive patients (through CT diagnosis, RT-PCR or seroprevalence testing), as well as computational capacities of very deep convolutionary networks training on a large scale. In addition, such a tailor-made network could make it more difficult to share it with hospitals, thereby making it less vulnerable to front-line applications.

Limitation of the Study

The biggest limit to this work is the data collection, by far. Despite accounting for the number of COVID-19 positive cases, the data-set obtained was low and was also somewhat unbalanced as the "Other Pneumonia" class was underrepresented. In fact, the data contains PA chest X-rays from a multitude of hospitals that may have decreased the accuracies and sensitivities. It is also clear that the predictions provided by the qualified profound neural networks on a PA chest X-ray can not be relied on alone. This is best used in conjunction with clinical testing (such as pathogen-RT-PCR or seroprevalence testing) or qualified diagnosis based on other simpler forms of medical imaging, such as CT scans.

Objective of the Study

- To proposed new detection methods of COVID-19 utilizing convolutional neural networks with different layers maxpooling 2D
- To compare between VGG16 and VGG19 for the detection methods of COVID-19 and its benifits

Scope and Significance of the Study

This work focuses on PA chest X-rays because while COVID-19 can be diagnosed less reliably than CT imaging, they are more popular in hospitals and comparatively cheaper to administer. Therefore, training of neural networks on PA chest X-rays could lead to greater generalization and use in hospitals around the world, and CT imaging will only be used in cases where patients are pathogen-confirmed COVID-19 by RT-PCR testing but do not consider AI diagnoses or where the AI diagnoses patients as highly likely COVID-19 but the RT-PCR tests are negative.

Expected Outcomes

It is hoped that if this work is put into action in hospitals, medium to long-term healthcare professionals will be able to diagnose potential pneumonia using machine learning software, and if it is identified, if it is related to a COVID-19 infection, allowing for the identification of new potential COVID-19 foyers during the possible "stop-and-go" lockdowns as predicted. Before establishing and distributing the vaccine. Furthermore, in the short term, practitioners are hoped to be able to correlate the diagnosis from deep convolutionary neural networks with potential

RT-PCR test results obtained by (Ai et al., 2020; Wang et al., 2020; Xu et al., 2020) and, if clashing, a computed tomography may be performed as it is more reliable to display COVID-19. Research has shown that VGG16 is the highest performing network, with a final 30 percent drop-out educated over 3 groups (COVID-19, No Finding, and Other Pneumonia). This has 93:9(±3:4) percent cross-validated training accuracy, 87:7(-1:9;+2) percent COVID-19 sensitivity, and 96:8(±0:8) percent No Finding sensitivity. The test set 's respective cross-validated values are 84:1(±13:5) percent, 87:7(±1:9; 2) percent, and 96:8(±0:8) percent.2 The model optimizer was Adam (Kingma and Lei-Ba, 2015) with a learning rate of 0.0001 and a categorical cross-entropy loss.

Bibliography

Ai, T., Yang, Z., Hou, H., Zhan, C., Chen, C., Lv, W., Tao, Q., Sun, Z., et al. (2020).Correlation of Chest CT and RT-PCR Testing in Coronavirus Disease 2019 (COVID-19)in China: A Report of 1,014 Cases. Radiology,

Bendavid, E., Mulaney, B., Sood, N., Shah, S., Ling, E., Bromley-Dulfano, R., Lai, C.,Weissberg, Z., et al. (2020).COVID-19 Antibody Seroprevalence in Santa Clara County,California.Stanford University, Stanford, CA.

Chollet, F. (2017).Xception: Deep Learning with Depthwise Separable Convolutions. 2017IEEE Conference on Computer Vision and Pattern Recognition (CVPR).

Cohen, J. P., Morrison, P., and Dao, L. (2020).COVID-19 image data collection.arXiv:2003.11597.Ferguson, N. M., Laydon, D., Nedjati-Gilani, G., Imai, N., Ainslie, K., Baguelin, M., Bhatia,S., Boonyasiri, A., et al. (2020).Impact of non-pharmaceutical interventions (NPIs)to reduce COVID-19 mortality and healthcare demand. Imperial College COVID-19Response Team, Imperial College London.

Flaxman, S., Mishra, S., Bhatt, S., Gandy, A., Unwin, J. T., Coupland, H., Mellan, T. A.,Zhu, H., et al. (2020). Estimating the number of infections and the impact of non-pharmaceutical interventions on COVID-19 in 11 European countries. Imperial CollegeCOVID-19 Response Team, Imperial College London.

Ioffe, S. and Szegedy, C. (2015). Batch Normalization: Accelerating Deep Network Trainingby Reducing Internal Covariate Shift. arXiv:1502.03167.

Kermany, D. S., Goldbaum, M., Cai, W., Valentim, C. C. S., Liang, H., Baxter, S. L.,McKeown, A., Yang, G., et al. (2018). Identifying Medical Diagnoses and TreatableDiseases by Image-Based Deep Learning. Cell, 172(5):1122{1131.

Kingma, D. P. and Lei-Ba, J. (2015). Adam: A Method for Stochastic Optimization.Published as a conference paper at ICLR 2015.

Narin, A., Kaya, C., and Pamuk, Z. (2020).Automatic Detection of CoronavirusDisease (COVID-19) Using X-ray Images and Deep Convolutional Neural Networks.arXiv:2003.10849.

Simonyan, K. and Zisserman, A. (2015).Very Deep Convolutional Networks for Large-ScaleImage Recognition.International Conference on Learning Representations.

Szegedy, C., Io_e, S., Vanhoucke, V., and Alemi, A. (2016). Inception-v4, Inception-ResNetand the Impact of Residual Connections on Learning. Proceedings of the Thirty-FirstAAAI Conference on Arti_cial Intelligence (AAAI-17), Google Inc.16

Szegedy, C., Vanhoucke, V., Io_e, S., Shlens, J., and Wojna, Z. (2015). Rethinking theInception Architecture for Computer Vision.arXiv:1512.00567.

Wang, S., Kang, B., Ma, J., Zeng, X., Xiao, M., Guo, J., Cai, M., Yang, J., et al. (2020). Adeep learning algorithm using CT images to screen for Corona Virus Disease (COVID-19).Pre-publication - medRXiv.

Xu, X., Jiang, X., Ma, C., Du, P., Li, X., Lv, S., Yu, L., Chen, Y., et al. (2020).DeepLearning System to Screen Coronavirus Disease 2019 Pneumonia.arXiv:2002.09334.17