Machine Learning COSC 6342

Fall 2020

Final Project

Project Title:

Automatic Detection of Covid-19 using Convolutional Neural Networks

Team members:

Parisa Mehdi Gholampour

Zarif Momin

Soumaya El Barrak

I. Literature review:

Covid-19 is a new deadly virus that cost millions of lives across the globe. Cases keep increasing worldwide, which challenges the medical systems. Diagnosis of the positive cases is a major task that requires special care. Providing an automatic diagnosis system will speed up the diagnosis and decrease the spread of the disease. In this direction, several studies have recently been performed in the context of covid-19 diagnosis. Ali Narin et al. in [1] have proposed a deep transfer learning technique for Covid-19 automatic diagnosis, where they have used five pre-trained Convolutional Neural Network models with 5-fold cross validation to perform binary classification. They have studied three datasets of chest X-ray images of patients suffering from COVID-19, viral pneumonia and bacterial pneumonia compared to normal cases, and therefore they end up with four classes. They reserved 80 % of the data for training and 20 % for testing. ResNet50 pre-trained model outperformed the rest of the classifiers. The accuracy achieved was 96.1 % for Covid-19/ normal classification, 99.5% for Covid-19/viral pneumonia classification and 99.7% for covid-19/bacterial pneumonia classification. Mahesh Gour et al. in [2] have come up with an ensemble learning based-approach to classify X-ray images for covid-19 diagnosis. They have combined chest X-ray datasets and implemented a stacked convolutional neural network to classify the cases into COVID-19, Normal, and Pneumonia classes. The first processing on the dataset has been done using a fine-tuned pre-trained network, the VGG19 [3], and a newly implemented 30-layered CNN model. The first step generated sub-models on which linear regression classifier was applied. The approach achieved an accuracy of 92.74% and a F1-score of 0.93 for the classification of X-ray images. Another promising approach has been proposed in [4], where a Decompose, Transfer and Compose (DeTraC) deep neural network model was adopted. The main goal behind this study is handling the problem of data irregularity and the limited number of training images that are facing medical images recognition. The authors worked on a combined dataset of X-ray images of Covid-19 positive cases, SARS positive cases and normal cases. The key contribution of this work is the decomposition of each class into sub-classes and recombining them to perform the last prediction. The transfer learning phase of the work was done using ImageNet pre-trained ResNet [5] model. The accuracy achieved was 95.12%.

In this paper [6], the authors present a framework for detecting COVID-19 from Chest X-Ray Images using convolutional neural. They built up a data-set consisting of 2000 non COVID, 84 COVID images for train data and 3000 non-COVID, 100 COVID images for test data. Augmentation techniques flipping and adding noise was done to create new images to overcome the limited number of COVID images. Furthermore, To build up their COVID detection framework, 4 popular pre-trained deep convolutional neural networks, ResNet18, ResNet50, SqueezeNet, denseNet-121, on training data set were evaluated. They used Cross entropy loss to train the models. Additionally Stochastic gradient descent algorithm was utilized to minimize the loss function. Since the number of COVID cases compared to other data were smaller in their test data set, Sensitivity and specifically measures were used to evaluate the performance of the models. Models were tested with different cutoff thresholds, Since the outcome of their models were probability scores and sensitivity and specificity measures were recorded for each of them. The SqueezeNet model showed more promising results compared to other models. It acquired a sensitivity rate of 98% with threshold number of 0.5 and specificity rate of 98% with threshold number of 0.15. ResNet18 took second place with a sensitivity rate of 98% with threshold number of 0.17 and specificity rate of 98% with threshold number of 0.35. ROC curve and precision Recall Curve were also presented to compare results of different models on a test set. In overall SqueezeNet gained better performance among other CNNs.

Bharadwaj et al. [7] created a deep stacked ensemble as a novel approach to classify chest X-ray images into 3 categories: COVID-19, viral phenomena and healthy ones. Their approach aims to achieve a model that has a faster training process and results in a more accurate model. Dataset[8] is used to examine the proposed model. They selected Inception, Vgg-Nets, Res-Nets, Inception-Res-Nets, Xception, Dense-Nets, Mobile-Nets and Nas-nets as pre trained convolutional neural network models. Furthermore, they developed a technique to have faster convergence in their classifier models. Models were trained in 3 steps. For each step learning rate, number of iteration and number of batches are incremented with predefined ratios to feed to the neural network. After training different models and evaluating them on test data base on their accuracy, precision and sensitivity, it is concluded that VGG-Nets outperform other networks with error rate of 4.7-4.9 on test data. Additionally, a stacked- ensemble containing Vgg-16 and Vgg-19 has been implemented as a novel approach in this work. The predictions of these two networks are concatenated and a new stack is created. This sack is used as an input to another deep neural network for final classification step. The authors argued that they were able to reduce the error of classification and achieve a model with high accuracy. Their accuracy score is 97.4 % for ternary classification, 99.48 for binary classification, sensitivity of 100% for and 99.44% for specificity.

The objective of this paper [9] is to examine the application of convolutional neural networks in medical image classification. The quality of their COVID-19 detection models have been examined on two datasets. First data set includes 1427 X-ray images which 224 of them are COVID-19, 700 of them are bacterial pneumonia and 504 of them are normal cases. The second one contains the same number of COVID-19 and normal cases of the previous dataset. The only difference is the number bacterial and viral pneumonia cases which is 714 images. Data is collected from available public resources. To preprocess the data, all the images were rescaled to 200×266 pixels and black background was added to prevent distortion. Furthermore, different types of CNN such as VGG19, MobileNet v2, Inception, Xception and Inception ResNet were evaluated. At the top of these networks an extra Neural Network with two hidden layers and a dropout layer has been placed.Prediction results on the first test data set revealed that VGG19 and MobileNet v2 have better accuracy compared to the rest of CNNs. VGG19 gained 98.75 accuracy for binary classification, 93.48 for ternary classification, 92.85 for sensitivity and 98.75 specificity. On the other hand, MobileNet v2 achieved 97.40 accuracy for binary classification, 92.85 for ternary classification, 99.10 for sensitivity and 97.09 for specificity. Although VGG19 provides better accuracy in this experiment, they argued that in medical diagnosis sensitivity measure which reflects false negative in its denominator is a better criterion to evaluate a better CNN model. For this matter MobileNet v2 was considered a better candidate. The result of Second dataset using MobileNet v2 as follows, 96.78 accuracy for binary classification, 94.72 for ternary classification, 98.66 for sensitivity and 96.46 for specificity.

The main goal of this paper [10] is to build an open-source dataset that can aid and be universally used by researchers/professionals aiming to build machine learning models for the covid-19 detection. The dataset collects

CT images of patients who were diagnosed with viral (incl. COVID-19), bacterial, fungal, and lipoid diseases. The wide-ranging diseases encompassed by this data-set helps to distinguish COVID-19-positive CT images from the other diseases, ensuring that CT scans depicting patterns more similar other diseases aren't misclassified as COVID-19, which is more likely to happen if the dataset only encompassed viral diseases. This dataset is very beneficial for anyone looking to build a multi-class classifiers that can be implemented on a variety of CT scans that result in more accurate COVID-19 detection than a binary (i.e. COVID-19 & normal) or ternary (i.e. pneumonia, COVID-19 & normal) classifiers that might misclassify a non-viral disease (i.e. bacterial, fungal, etc.) as COVID-19 due to the narrowness of the classes it has been trained and validated with.

In this paper [11], the authors propose the use of a convolution neural network based on the EfficientNet architecture and compare their results with other CNNs proposed in other research papers they list in related works. Their CNN, as well as the CNNs they use of comparison, use chest X-Ray image data as input and output either a binary classification ('Normal', or 'COVID-19'), or ternary classification (namely, 'Normal', 'Pneumonia', or 'COVID-19'). For training/testing, the authors used dataset with equal number of examples/images from all three classes to ensure proper training of the model. The authors used the B4 model of EffientNet (not B5-B7, which can produce higher accuracies, due to hardware/system constraints) which provided better /higher accuracies than other competing models (i.e. VGG, GoogLeNet, and Residual Network). The resulting accuracies do back up the claim that the model produces better accuracies, generating the top accuracy for binary classification (among those compared) at 99.62%, while generating 96.70% accuracy for ternary classification, being second only to a custom CNN proposed by Nour et. al[12], which outputted 97.14%.

In this paper [13], the authors have built a CNN architecture from the ground up, specifically for the covid-19 detection. The architecture is named 'COVIDNet-CT deep neural network architecture', and was built using a machine-driven design exploration strategy (not a novel idea), namely, generative synthesis, pretrained on ImageNet dataset[14], and trained on their custom dataset (COVIDx-CT dataset) via stochastic gradient descent with momentum, using custom (specified) hyperparameters [13]. The resulting (ternary classification, same as in paper 2, no binary classification was performed for this paper) accuracy of this model (as depicted in the paper) is very high at 99.1% accurate, and is *seemingly* higher than the one depicted in the previous paper[11], but due to the variation in datasets and their respective test/train/validation ratios, the two accuracies can't be compared directly. However, the importance of this architecture is that it is much less architecturally complex, as it only needs to factor in ~1.54 million parameters, whereas ResNet-50 (used as a comparison in this paper) needs to factor in ~23 million parameters, and the model proposed in the previously discussed paper [11] needed to factor in ~18 million parameters. Hence, the computational complexity (in terms of floating point operations per second, aka flops, and 1000 flops = 1 GFLOP) of COVIDNet-CT model (which is ~4.18 GFLOPS) is significantly less than the ResNet-50 model's (which is ~42.72 GFLOPS), the former's computation complexity being less than a tenth of the latter's.

II. Project implementation

1. Methodology

In our project, we have chosen to work on the area of computer vision, especially chest x-ray medical images analysis. Our study aims at enabling automatic prediction of COVID-19 using convolutional neural network (CNN) models. For this purpose, we have used AlexNet, Squeeznet, ResNet18, ResNet34, VGG16 and EfficientNetB4 pre-trained models. We performed two types of classifications, ternary and binary on two different datasets. The ternary classifies the images into three classes, as COVID-19 positive, viral pneumonia and normal. The binary classifies the images into COVID-19 positive and normal.

To perform ternary classification, we have worked with the COVID-19 RADIOGRAPHY DATABASE available at [8]. We used the models AlexNet, Squeeznet, ResNet18, ResNet34, VGG16 to classify the images in this database.

Table (1) describes the distribution of the three classes of images, as COVID-19 positive, viral pneumonia and normal.

COVID-19 positive	Viral pneumonia	Normal
219	1345	1341
		

Table 1: Chest X-ray images for COVID-19 positive, Viral pneumonia disease and Normal cases[8]

To perform binary classification, we mixed two datasets, the first one available at [15]. We only used the directory that contains normal cases from this dataset. The second one is available at [16] which contains COVID-19 positive cases. The model we used to perform the binary classification is the EfficientNetB4 network. Table (2) describes the distribution of the two classes of images, as COVID-19 positive and normal

COVID-19 positive	Normal
555	540

Table 2: Number of images in the mixed dataset for COVID-19 positive and normal cases

As for our knowledge, the available COVID-19 datasets are limited in terms f the number of samples which limits the performance of the models. For this reason, we opt to use transfer learning, which is a mechanism that serves to transfer knowledge from generic prediction task, which is object recognition in our context, to domain-specific task which is chest x-ray images classification. In this direction, the models we used were pre-trained on the ImageNet dataset [17], which achieved state-of-the art performance.

To implement our project, we have used Google Collaboratory platform, Python programming language and PyTorch and deep learning framework for the AlexNet, Squeeznet, ResNet18, ResNet34 and the VGG16, and Keras for EfficientNetB4. The pre-trained CNN models are fixed features extractors, and the final layer performs classes prediction.

2. Data pre-processing

II.1.Single dataset

The processing we performed to create our custom dataset from [8] includes images conversion from Black & White to RGB. Changing the size of images to 224x224 pixels to be the same as the size in the ImageNet dataset. Data augmentation to handle classes imbalance, which consists on performing random horizontal flip on the data. Transforming images to tensor format which is a generic n-dimensional array format interpreted by PyTorch framework and data normalization. The afore-mentioned steps were applied on the training set. For the test set, we excluded data augmentation step and applied the rest of the steps as for the training set.

II.2.Mixed datasets

The processing we performed to mix the datasets from [15], [16] includes resizing the images to 225x225 and applying many augmentation techniques, including, but not limited to, flip, transpose, rotate, and vary from loop to loop based on predefined probabilities. Training and testing data are split using Stratified 10-fold cross-validation, with shuffling.

3. Model's architectures and training

3.1. ResNet18 & ResNet34

ResNet refers to residual neural network, which is a type of convolutional neural networks. The numbers 18 and 34 refers to the depth of each model, e.g. the number of layers. There two models were trained on ImageNet dataset which has 1000 classes. In our study case we only have three classes, COVID-19 positive, viral pneumonia and normal. We needed to reshape the last fully connected layer for both models to have three output features. The parameter used for evaluation of the loss function is the cross-entropy loss and the optimization was done using Adam optimizer.

3.2. AlexNet

AlexNet model has different architecture which consists of eight layers: five convolutional layers and three fully-connected layers. It uses Rectified Linear Units (ReLU), allows for multi-GPU training and overlapping pooling.To reduce overfitting, two methods were employed, data augmentation and dropout. The output in the pre-training comes from the sixth layer of the classifier. For this reason, we reshaped the sixth layer to serve a ternary classification.

3.3. VGG16

VGG16 is a network with 16 layers proposed by the Visual Geometric Group from Oxford University. The input layer accepts RBG images with size 224x224. The images pass through a stack of convolution layers. It is performed over a max-pool window of size 2 x 2 with stride equals to 2. ReLU is the activation function used in the hidden layers. For our purpose, we disabled training for the convolutional layers setting, as we will only train the fully connected classifier to perform ternary classification.

3.4. SqueezeNet:

SqueezeNet network begins with a standalone convolution layer, followed by 8 Fire modules, ending with a final conv layer. A Fire module is comprised of a squeeze convolution layer (which has only 1×1 filters), feeding into an expand layer that has a mix of 1×1 and 3×3 convolution filters. The number of filters per fire module is gradually increased from the beginning to the end of the network. Max-pooling with a stride of 2 is performed after the first convolutional layer, fire4, fire8, and last convolutional layer. Torchvision module from PyTorch has two versions of Squeezenet, we use version 1.0. In our study, the output comes from a 1x1 convolutional layer which is the 1st layer of the classifier.

We reinitialize it to have an output feature map of depth 3.

3.5. EfficientNetB4

EfficientNetB4 is a model created by proposing a new scaling technique for CNN, that uniformly scales all dimensions of depth/width/resolution using a simple yet highly effective compound coefficient[18]. We used it for our purpose followed by a layer of global average pooling 2d to reduce overfitting, and 3 inner dense layers with RELU activation functions and dropout layers.

4. Result

The performances of the studied models on the ternary and the binary classifications were evaluated using plots of the the accuracy, the confusion matrix and the ROC curves. The results are shown for the two datasets. As for the single dataset [8], we show the performance of the ternary classification of the models Vgg16, ResNet18, ResNet34, AlexNet and SqueezNet, whereas for the mixed dataset [15], [16], we show the performance of the binary classification of the model EfficientNetB4.

Figure (1) illustrates the achieved validation accuracy at each step of the training phase on the balanced validation data of the same dataset of the models Vgg16, ResNet18, ResNet34, AlexNet and SqueezNet. The results are obtained using one epoch, a batch size equal to six, 470 training batches and 15 test batches with shuffled indices for the inputs.

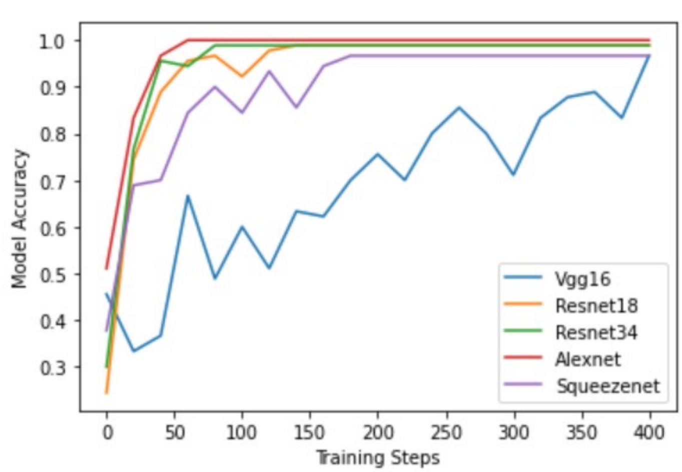

Figure 1: Validation accuracies for ternary classification by the models Vgg16, Rsenet18, Resnet34, Alexnet and Squeezenet

The AlexNet model outperforms the rest of the models since it reaches an accuracy of 100% after processing less than 100 training samples. ResNet34 and ResNet18 achieved accuracies of 98% as well after processing less than 100 training samples. Squeezenet achieved an accuracy of 98% after processing less than 200 samples. Vgg16 struggled to achieve an accuracy of 96% after processing 400 samples. Table (3) illustrate the achieved accuracy on the test set for the five models.

Model	**Accuracy**
AlexNet	**92.2%**
ResNet18	**91.1%**
ResNet34	**92.2%**
SqueezeNet	**93.3%**
Vgg16	**90.0%**

Table 3: Accuracies achieved on the test set

On the test set, the Squeezenet model achieved the highest accuracy of 93.3% as illustrated in the confusion matrix in figure (2).

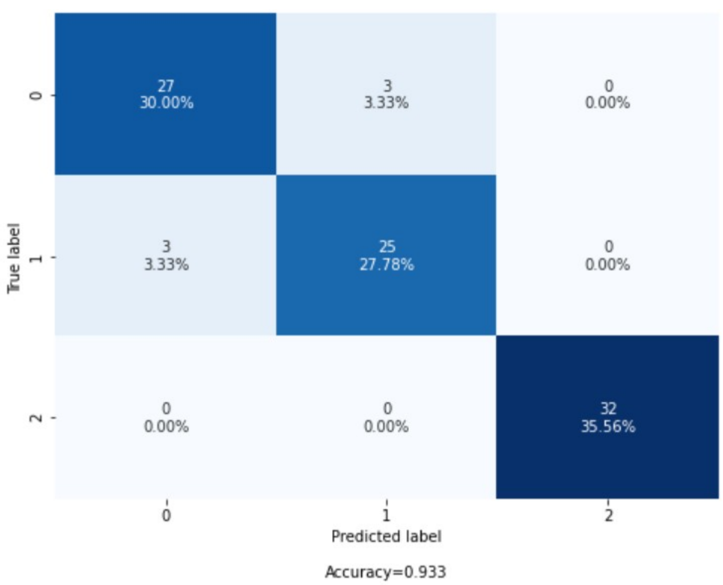

Figure 2: Confusion matrix of the model SqueezeNet

Regarding the mixed data set for the binary classification of COVID-19 positive vs Normal cases, we applied the EfficientNetB4 model. The accuracy achieved on the test set reached 99.49% after 10 fold cross-validation, as illustrated in the confusion matrix in figure (3). The same behavior was proven by the ROC curves as illustrated in figure (4) that shows the trade-off between sensitivity (or True Positive Rare) and specificity (1 – False Positive Rate). The obtained trade-off between sensitivity and specificity is minimal for the binary classification.

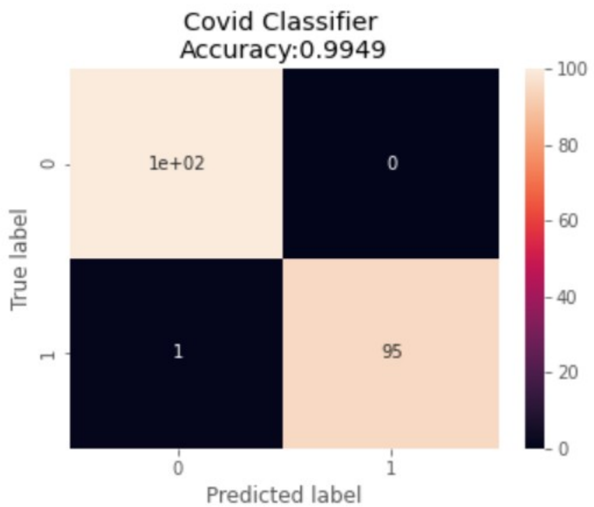

Figure 3: Confusion Matrix for EfficientNetB4 model

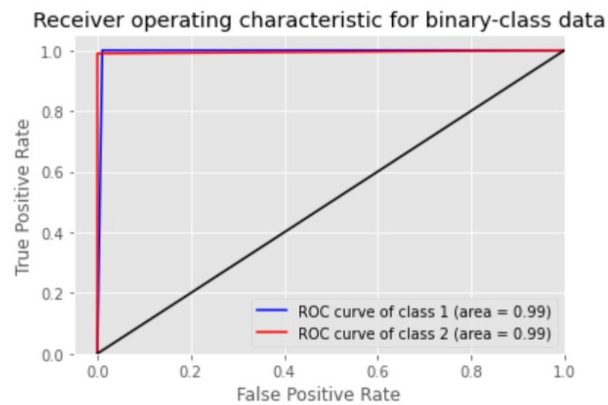

Figure 4: ROC curves for the binary classification of covid-19 for binary classification of covid-19 positive and normal cases

III. Conclusions

In our final project, we performed an inclusive comparative study of six different convolutional neural network based pre-trained models. We took advantage of transfer learning and reshaped the models to suit our chosen datasets. We worked with two kinds of dataset to enrich our study, a single dataset and a mixed one. The data pre-processing phase of the project took an interesting part of the workload as we had to deal with three different datasets. The choice of the models was rational. Two models, the ResNet18 and ResNet34 were similar in the level of the architecture and differ in terms of the number of layers to study the effect of the number of layers on the performance. The other models have different architectures. Five models were applied on a similar dataset. SqueezeNet outperforms the rest of the models. The EfficientNetB4 was applied on the mixed dataset. This latter has achieved outstanding performance.

References

[1] A. Narin, C. Kaya, and Z. Pamuk, *Automatic Detection of Coronavirus Disease (COVID-19) Using X-ray Images and Deep Convolutional Neural Networks*. 2020.

[2] M. Gour and S. Jain, "Stacked Convolutional Neural Network for Diagnosis of COVID-19 Disease from X-ray Images," *ArXiv200613817 Cs Eess*, Jun. 2020, Accessed: Nov. 24, 2020. [Online]. Available: http://arxiv.org/abs/2006.13817.

[3] A. Krizhevsky, I. Sutskever, and G. E. Hinton, “ImageNet classification with deep convolutional neural networks,” *Commun. ACM*, vol. 60, no. 6, pp. 84–90, May 2017, doi: 10.1145/3065386.

[4] A. Abbas, M. M. Abdelsamea, and M. M. Gaber, “Classification of COVID-19 in chest X-ray images using DeTraC deep convolutional neural network,” *ArXiv200313815 Cs Eess Stat*, May 2020, Accessed: Nov. 24, 2020. [Online]. Available: http://arxiv.org/abs/2003.13815.

[5] K. He, X. Zhang, S. Ren, and J. Sun, “Deep Residual Learning for Image Recognition,” *ArXiv151203385 Cs*, Dec. 2015, Accessed: Nov. 25, 2020. [Online]. Available: http://arxiv.org/abs/1512.03385.

[6] S. Minaee, R. Kafieh, M. Sonka, S. Yazdani, and G. J. Soufi, “Deep-COVID: Predicting COVID-19 From Chest X-Ray Images Using Deep Transfer Learning,” *ArXiv200409363 Cs*, Jul. 2020, Accessed: Dec. 06, 2020. [Online]. Available: http://arxiv.org/abs/2004.09363.

[7] L. B. B, R. Boddeda, S. V. K, S. M. G, S. Deepak V, and S. A. Patel, “A Novel Strategy for COVID-19 Classification from Chest X-ray Images Using Deep Stacked-Ensembles,” *ArXiv201005690 Cs Eess*, Oct. 2020, Accessed: Dec. 06, 2020. [Online]. Available: http://arxiv.org/abs/2010.05690.

[8] “COVID-19 Radiography Database.” https://kaggle.com/tawsifurrahman/covid19-radiography-database (accessed Dec. 05, 2020).

[9] I. D. Apostolopoulos and T. Bessiana, “Covid-19: Automatic detection from X-Ray images utilizing Transfer Learning with Convolutional Neural Networks,” *Phys. Eng. Sci. Med.*, vol. 43, no. 2, pp. 635–640, Jun. 2020, doi: 10.1007/s13246-020-00865-4.

[10] J. P. Cohen, P. Morrison, and L. Dao, “COVID-19 Image Data Collection,” *ArXiv200311597 Cs Eess Q-Bio*, Mar. 2020, Accessed: Dec. 06, 2020. [Online]. Available: http://arxiv.org/abs/2003.11597.

[11] “Automated medical diagnosis of COVID-19 through EfficientNet convolutional neural network - ScienceDirect.” https://www.sciencedirect.com/science/article/pii/S1568494620306293?via%3Dihub (accessed Dec. 06, 2020).

[12] “A Novel Medical Diagnosis model for COVID-19 infection detection based on Deep Features and Bayesian Optimization.” https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7385069/ (accessed Dec. 06, 2020).

[13] H. Gunraj, L. Wang, and A. Wong, “COVIDNet-CT: A Tailored Deep Convolutional Neural Network Design for Detection of COVID-19 Cases from Chest CT Images,” *ArXiv200905383 Cs Eess*, Sep. 2020, Accessed: Dec. 06, 2020. [Online]. Available: http://arxiv.org/abs/2009.05383.

[14] “ImageNet: A large-scale hierarchical image database - IEEE Conference Publication.” https://ieeexplore.ieee.org/document/5206848 (accessed Dec. 06, 2020).

[15] “Chest X-Ray Images (Pneumonia) | Kaggle.” https://www.kaggle.com/paultimothymooney/chest-xray-pneumonia (accessed Dec. 06, 2020).

[16] J. P. Cohen, *ieee8023/covid-chestxray-dataset*. 2020.

[17] “ImageNet.” http://www.image-net.org/ (accessed Dec. 05, 2020).

[18] M. Tan and Q. V. Le, “EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks,” *ArXiv190511946 Cs Stat*, Sep. 2020, Accessed: Dec. 06, 2020. [Online]. Available: http://arxiv.org/abs/1905.11946.