Title:

A Computer Aided Medical Classification System of COVID-19 CT Lung Scans using Convolution Neural Networks (CNN)

Authors:
Miri Weiss Cohen, miri@braude.ac.il, Braude College of Engineering
Oriel Gilo, orielgilo@gmail.com, Braude College of Engineering
Liron David, lirondavid.work@gmail.com, Braude College of Engineering

Keywords:
COVID-19, CT Lung scans, Classification methods, Convolution Neural Networks (CNN)

DOI: 10.14733/cadconfP.2021.41-45

Introduction:

The coronavirus infection decease (COVID-19), has surprised the world with its rapid spread, potential virulence and overall impact on the lives of billions of people from both a safety and economic perspective. As of January 11, 2021, there have been 90,243,267 confirmed cases of COVID-19, including 1,934,216 deaths, reported to the world health organization[1].

The diagnosis of COVID-19 relies on the following criteria: clinical symptoms, epidemiological history and positive CT images and positive pathogenic testing. The clinical characteristics of COVID-19 include respiratory symptoms, fever, cough, dyspnea, and pneumonia [2]. However, these symptoms are nonspecific, as there are isolated cases where, in an asymptomatic infected family a chest CT scan revealed pneumonia and the pathogenic test was found positive. Once a person is identified as a PUI (Person Under Investigation), lower respiratory specimens, such as Broncho alveolar lavage, tracheal aspirate or sputum, will be collected for pathogenic testing.

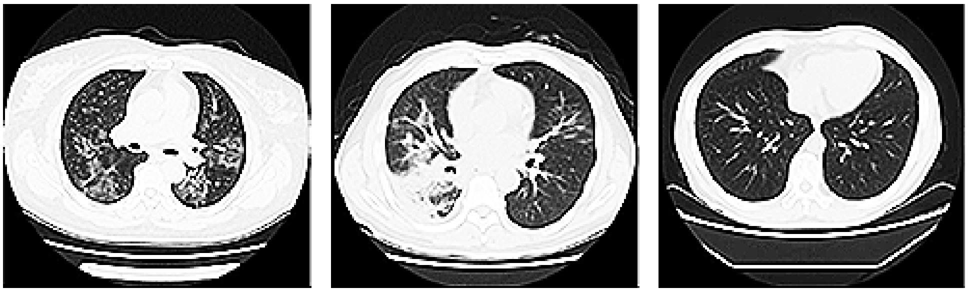

Fig. 1: CT scans COVID-19 (left), pneumonia (middle) and normal (right)

X-Ray and CT imaging is a major diagnostic tool for COVID-19. The majority of COVID-19 cases have similar features on CT images including ground-glass opacity in the early stage and pulmonary consolidation in the late stage. In some cases, a rounded morphology and a peripheral lung distribution change. Although typical CT images may help early screening of suspected cases, the images of various viral pneumonia are similar, and they overlap with other infectious and inflammatory lung diseases. In some cases it is difficult for radiologists to distinguish COVID-19 from other viral pneumonia's [3], therefore making a automatic medical classification system a challenge. Figure 1 depicts three cases of CT scans, on the left a COVID-19, in the middle a case of pneumonia and on the right healthy lungs. Using X-Ray imaging as an input for diagnosing cases of COVID-19 can be implemented and used in rural areas and less developed countries, where the source of scans is more available. The system will demonstrate the feasibility of establishing an early screening model to assist doctors in distinguishing COVID-19 from healthy cases using deep learning techniques. Three Convolutional Neural Networks (CNN) architectures were trained, validated and tested to achieve the highest accuracy in a computer aided medical classification system.

Convolutional Neural Networks (CNN):

Deep learning is a methodology used to implement supervised learning for classification, which mirrors the workings of the human brain in processing data and creating patterns for decision-making[4]. Convolution neural network (CNN) are deep graph networks developed for learning by the process of training and testing with input labeled data. In this work the input is a RGB image of CT scans of lungs. Each layer of the CNN is composed of several filter layers which extract local features. In each forward pass each input image through a sequence of multiply layers that are composed of convolution layers, activation function, pooling, fully connected layers and a softmax function that classifies an object with probabilistic values, moreover, back-propagation is used to regulate the change in weights in relation to a loss value. Figure 2 depicts the mail idea of developing the medical diagnostic system. where the input were images taken from the SARS-COV-2 Ct-Scan Dataset [7], This data-set included of 2482 CT scans, which 1252 corresponds to 60 patients identified with SARS-CoV-2, and 1230 CT scans correspond to 60 patients not identified as normal. These data have been collected from public hospitals in Brazil. In our work we trained three

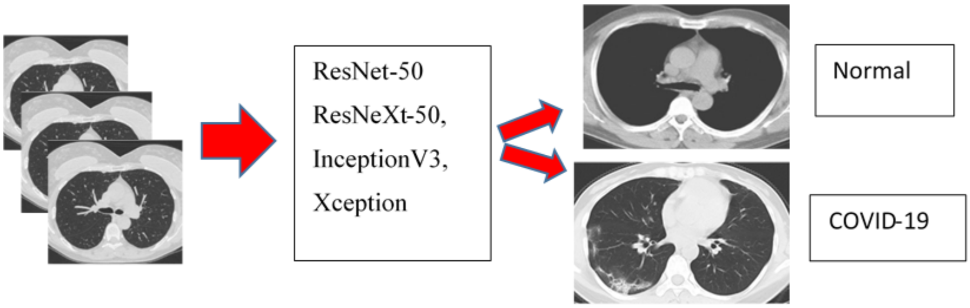

Fig. 2: Basic concept of the learning process

different CNN architectures: ResN50[5], Inception V3 and Xception[6]. ResNet introduced the residual learning in CNNs and devised an efficient methodology. ResNet with 50/101/152 layers has fewer failures on image classification and gained 28percent improvement on the famous image recognition benchmark. In Inception V3 several techniques for optimizing the network were implemented to loosen constraints for easier model adaptation. The Xception architecture is a modification of the original inception block by making it wider and replacing the different spatial dimensions, a single dimension (3x3) followed by a (1x1) convolution to regulate computational complexity.

Due to the pandemic being a new disease and the lack of quality data-sets, we use transfer learning to extract knowledge from one or more applications in this field of study and help to boost the learning performance of our system. This method is much faster and reliable than training a network from scratch with randomly initialized weights. During the training phase, we focused the research on the influence of hyper-parameters on the accuracy of each model, searching for minimal loss values. The hyper-parameters evaluated are Learning rate and Epoch size, while the batch size was fixed to 32. The Learning rates range are: 0.0005, 0.00005 and 0.000005, and the Epoch sizes tested are: 20, 40 and 60. As batch size is hardware dependent, we chose a batch size adaptive local GPU use. When training deep networks, it is often useful to reduce the learning rate as the training progresses. This can be done by using a pre-defined learning rate schedules or adaptive learning rate methods. In this system development we used a schedule learning rate. The learning rate was dropped by a factor every few epochs - when a plateau in the optimization model performance is detected. This callback is designed to reduce the learning rate after the model stops improving and suffices the fine-tuning of the weights.

Results

Using the SARS-COV-2 Ct-Scan data-set[7], for the learning process, we divided the images into three different parts: training set, validation set and testing set. We found that 71.4% of the data for training, 14.3% for validation 14.3% for testing yielded the best results.

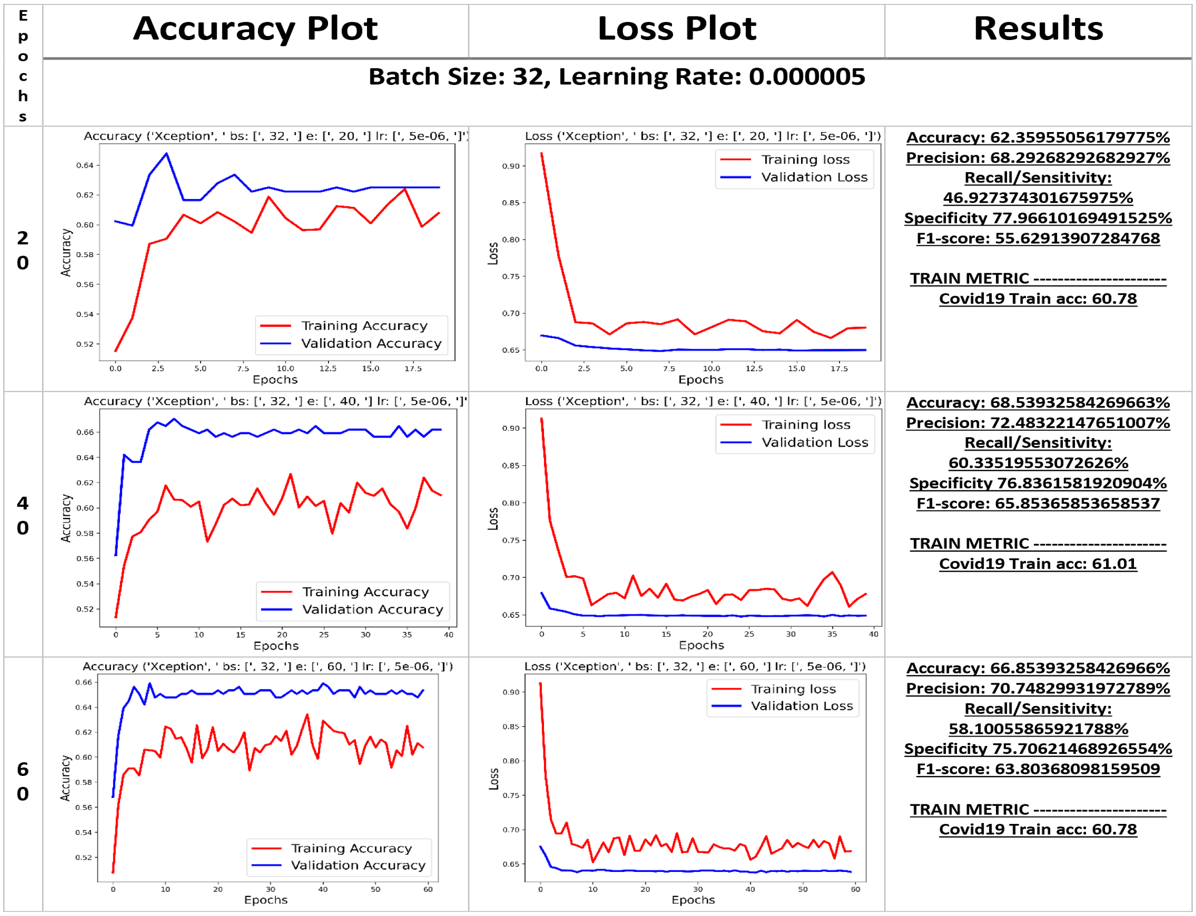

Fig. 3: Training and validation accuracy and loss graphs Xception - LR=0.00005 Epochs=20,40,60 and batch-size = 32

For simplicity in presenting the results we will define the following tested parameters:

- tp = model prediction was COVID-19, and the CT scan was classified as COVID-19
- tn = prediction was NORMAL lungs, and the CT scan was classified as NORMAL lungs
- fp = prediction was COVID-19, and the CT scan was classified as NORMAL lungs
- fn = prediction was NORMAL lungs, and the CT scan was classified as COVID-19
- Accuracy rate prediction $Accuracy = (t_p + t_n)/(t_p + t_n + f_p + f_p)$
- Rate of correctly classifying a CT scan as a COVID-19 $Prescision = (t_p)/(t_p + f_p)$
- Rate of correctly classifying a CT scan as a COVID-19 $Recall = (t_p)/(t_p + f_n)$
- Rate correctly classifying the CT scan of a Normal lungs $Sensitivity = (t_n)/(t_n + f_p)$

We associated the most important aspect of Coronavirus detection to alert those who are infected with the virus rather than informing healthy people that are negative for the virus. Hence, the term Recall which indicates all the verified patients out of all the infected peoples with the virus, is the parameter we will focus on. The following result is an example of many learning procedures performed. Figure 3 depicts a case of Xeception architecture, examining the influence of epochs on a 0.00005 learning rate (batch-size = 32).All the experiments in each of the revolved case a static learning rate and batch size were defined, while the epoch size was test in a continuously changing manner.
For the 20 epochs the results:
Accuracy: 62.4% Precision: 68.3% Recall/Sensitivity: 46.9% Specificity 78% Covid19 Train acc: 60.78
For the 40 epochs the results:
Accuracy: 66.68% Precision: 70.5% Recall/Sensitivity: 58.4% Specificity 75.7% Covid19 Train acc: 61.01
For the 40 epochs the results:

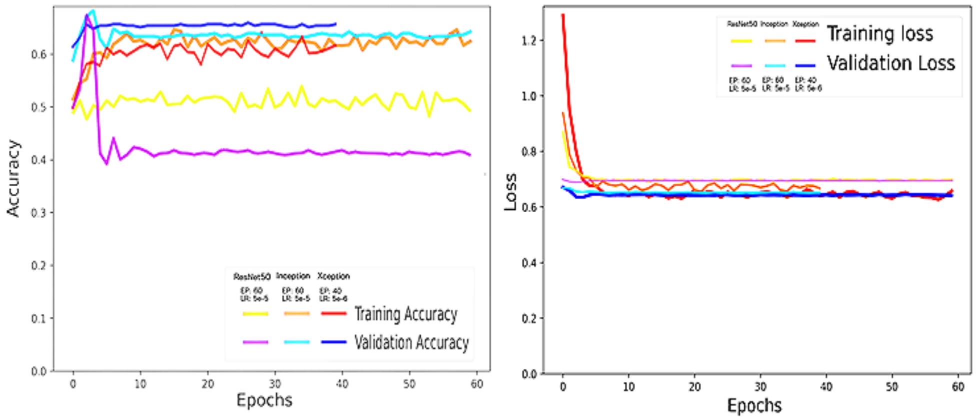

Fig. 4: Combined training and validation accuracy and loss graphs for the 3 architectures

Figure 4 depicts combined results of the three best results of each of the architectures. Xception, Inception V3 and ResNet50. The Xception is marked with red for the training and blue for the validation, the chosen hyper-parameters for this architecture were Learning rate of 5e-06, batch size of 32 and 40 epochs. The Inception V3 is marked with orange for the training and cyan for the validation, the chosen hyper-parameters for this architecture were Learning rate of 5e-05, batch size of 32 and 60 epochs. The ResNet50 is marked with yellow for the training and purple for the validation, the chosen hyper-parameters for this architecture were Learning rate of 5e-05, batch size of 32 and 60 epochs. Notice that the Xception graph is shorter than the other 2, but still yielded the highest Validation Accuracy rate and the lowest Validation Loss rate out of the 3 best results of: Accuracy: 68.54Precision: 72.48Recall/Sensitivity: 60.34Specificity

76.84F1-score: 65.85.

Conclusions

We demonstrated the feasibility of establishing a computer aided medical classification system of COVID-19 CT Lung Scans to assist doctors in distinguishing between COVID-19 and healthy cases with pulmonary CT slice images using CNN. Researching three different CNN models, with various learning rates, epochs and a static batch size has on the same data-set provided an understanding of the influence of each hyper-parameter in the learning process. We noticed that all the major changes in accuracy were dependent on variations in epochs. It can be concluded that the learning rate volatility did not affect the accuracy when defined in the same epoch size. We have concluded that the best accuracy/minimal loss was found in Xception model in the case: epoch size = 40, learning rate = 0.000005 and batch size = 32.

References:

[1] https://www.who.int/emergencies/diseases/novel-coronavirus-2019

[2] Bernheim, A.; Mei, X.; Huang M.; Yang, Y.; Fayad, Z. A.; Zhang, N.; Chung, M.: Chest CT findings in coronavirus disease-19 (COVID-19): relationship to duration of infection. Radiology, 200463. relationship to duration of infection, Radiology, 200463, 2020.

[3] Wang, S.; Kang, B.; Ma, J.; Zeng, X.; Xiao, M.; Guo, J.;Xu, B.: A deep learning algorithm using CT images to screen for Corona Virus Disease (COVID-19), MedRxiv, 2020

[4] Gozes, O.; Frid-Adar, M.; Greenspan, H.; Browning, P. D.; Zhang, H.; Ji, W.; Siegel, E. Rapid AI development cycle for the coronavirus (covid-19) pandemic: Initial results for automated detection patient monitoring using deep learning ct image analysis. arXiv:2003.05037, 2020.

[5] Khan, A.; Sohail, A.; Zahoora, U.; Qureshi, A.S.: A survey of the recent architectures of deep convolutional neural networks. Artificial Intelligence Review, pp.1-62. 2017.

[6] Rawat, W.; Wang, Z.; Deep convolutional neural networks for image classification: A comprehensive review. Neural computation, 29(9), 2352-2449, 2017.

[7] SARS-COV-2 Ct-Scan Dataset: https://www.kaggle.com/plameneduardo/sarscov2-ctscan-dataset.