Detection of Covid-19 on chest x-ray images using Convolutional Neural Networks

O. A. Esteban, S. H. Serrano

Facultad de Ingeniería y Tecnología, Universidad de Montemorelos

Abstract — *COVID-19 (Coronavirus disease 2019), is an infectious disease caused by the SARS-CoV-2 virus. It was first detected in late 2019, in the city of Wuhan, China. Three months later, in March 2020, it was declared a pandemic by the WHO because it was already affecting several countries. The virus spreads easily through droplets released from the nose or mouth of an infected person, making it difficult to contain worldwide. The goal of this research is to predict whether a person has COVID-19, pneumonia, or not, by evaluating chest x-ray images using CNNs. In this paper, we proposed a model based on CNNs and used a dataset with chest x-ray images of COVID-19, pneumonia and healthy. The final model's results shows a 97% accuracy, precision, recall and f1 score.*

I. INTRODUCTION

a. Describing the problem

COVID-19 took the world by surprise in 2020. Almost 5 months after the first infection was believed to have occurred, more than 3 million people have been infected and 200,000 have died. Just in the US, there are more than a million infections. This disease, which originated in Wuhan, China, causes people to be feverish, tired, have a dry cough or even a sore throat and nasal congestion. You can get COVID-19 from contact with another person who is infected with the virus [1]. In some cases, COVID-19 can cause pneumonia as a complication of itself. This is more likely in people who are in third age, and those who are 85 or older are at the highest risk, but also affects someone who has a weakened immune system, like smokers, people being treated for cancer, diabetes HIV or AIDS, people who have had a bone marrow transplant and people who takes specific medications. Diagnosing COVID-19 then becomes critical to tracking and controlling its transmission.

PCR (Polymerase chain reaction) are the standard COVID-19 tests. There are a few issues with the test because antibodies are produced over days to weeks after infection with the virus[2]. Laboratory testing is a better diagnostic but it takes about 2 to 5 days to get a diagnosis with significant false-negative results[3].

COVID-19 standard tests [2] and [3] are extremely expensive and a large scale implementation cannot be afforded by many of the developing and underdeveloped countries. Fast and accurate diagnostic methods are urgently needed to combat the disease.

In recent years, machine learning has been used more and more in the medical area for disease diagnosis, personalized treatment, clinical research or even predictions. These and many other applications cause an improvement in results and a reduction of time of other processes.

b. Related works

Lately, several researches have been published proposing one or many classification models to detect COVID-19. For instance [4], shows the performance of thirteen models to detect COVID-19, Pneumonia and Normal patients using chest x-ray images. It measures using deep features, like: Sensitivity, FPR and F1 Score. However, they used 127 images per class. Also, the model with higher accuracy and f1 score gave 95% in both metrics.

In [5], a CNN previously developed by the authors, called Decompose, Transfer, and Compose (DeTraC), is validated and adapted for the classification of chest X-ray images dataset collected

from several hospitals around the world. However, it presents a high accuracy of 95.12% in the detection of COVID-19 X-ray images from normal, and severe acute respiratory syndrome cases. It also does not specify how many images were used and is not open data. This work is structured as follows: The second section presents Methodology, the third section, Experimental results, and the fourth section, Conclusions and Future work.

II. METHODOLOGY

a. Data collection

A total of 534 x-ray images were used. COVID-19 images were extracted from [6]. This dataset contains images mostly from male than female patients. It also has more labeled patients over their 50's. Italy is the most labeled country. Still, there are from another countries including: China, Korea, USA or Taiwan.

Pneumonia and healthy images were extracted from [7]. Images we used are also available in a Jupyter folder for COVID[1], Pneumonia[2] and Healthy[3] images.

b. Data cleaning

The pandas library was used to read the metadata[4] in csv format. To get the images from the repositories the glob library was used. Images of datasets of Pneumonia and Healthy patients are taken from a frontal perspective. COVID-19 dataset is composed of more images taken from a frontal perspective (See Figure 1). Although, it also contains images from other perspectives (See Figure 2).

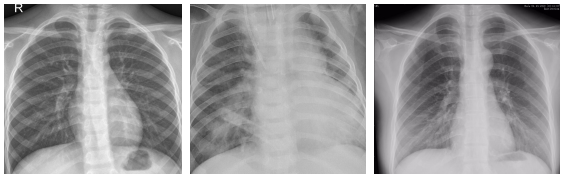

Figure 1. Frontal perspective. From left to right: Healthy, Pneumonia and COVID-19.

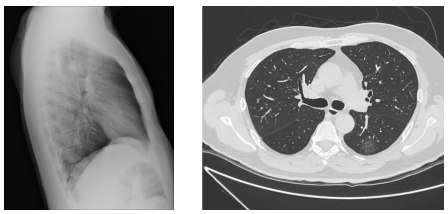

Figure 2. Non-frontal perspective. Only from COVID-19.

The images of the 3 classes were converted to grayscale and resized to 128x128 pixels. To do this, the cv2 library was used. Each class had 178 images. A total of 427 images (80%) were used for training and 107 (20%), for testing, which already included the 3 classes.

c. Topology creation

To create the topology, the Tensorflow library was used, version 2.0.0. Also, the Keras library was used, version 2.3.1. All created and evaluated on an alienware server through Jupyter.

The topology is sequential, consisting of 4 stages. Each stage consists of a padding of 2, a convolutional layer and a max pooling of 2x2. Then 2 layers completely connected, and an output with 3 neurons, one for each class. The layers of the topology are shown in Table 1.

	Layer (Type)	Output shape	Parameters
1	Zero Padding	(None, 132, 132, 1)	0
2	Conv2D	(None, 130, 130, 8)	80
3	MaxPool2D	(None, 65, 65, 8)	0
4	Zero Padding	(None, 69, 69, 8)	0
5	Conv2D	(None, 67, 67, 8)	584
6	MaxPool2D	(None, 33, 33, 8)	0
7	Zero Padding	(None, 37, 37, 8)	0
8	Conv2D	(None, 35, 35, 16)	1168
9	MaxPool2D	(None, 17, 17, 16)	0
10	Zero Padding	(None, 21, 21, 16)	0

[1] http://201.134.41.15:8889/tree/ANN/Oscar/Ex3/Covid

[2] http://201.134.41.15:8889/tree/ANN/Oscar/Ex3/Pneumonia

[3] http://201.134.41.15:8889/tree/ANN/Oscar/Ex3/Normal

[4] http://201.134.41.15:8889/edit/ANN/Oscar/Ex3/metadata.csv

11	Conv2D	(None, 19, 19, 16)	2320
12	MaxPool2D	(None, 9, 9, 16)	0
13	Dropout	(None, 9, 9, 16)	0
14	Flatten	(None, 1296)	0
15	Dense	(None, 100)	129700
16	Dense	(None, 3)	303

Table 1. Layers of the topology.

The first stage is described below:

On line 1, a 2-pixel padding is added for each side of the input image.

On line 2 the input passes through a convolutional layer with 8 3x3 pixel filters, a hyperbolic tangent trigger function and the he_uniform initializer from the Keras library for the weight matrix kernel.

In line 3 a 2x2 max pooling is applied for the output of the convolutional layer.

The second stage does the same as the first stage in lines 4, 5, and 6 respectively.

The third stage is described below:

In line 7 a 2 pixel padding is added for each side of the input image which is the output of the second stage max pooling.

On line 8 the input passes through a convolutional layer with 16 3x3 pixel filters, a hyperbolic tangent trigger function and the he_uniform initializer for the weight matrix kernel.

In line 9 a 2x2 max pooling is applied for the output of the convolutional layer.

The fourth stage does the same as the third stage in lines 10, 11, and 12 respectively.

In line 13 a dropout of 0.2 is applied to the output of line 12, which reduces the number of neurons to 20% at random.

In line 14 the flatten allows the image to be reduced to a single dimension.

On line 15 a fully connected layer with 100 neurons uses the hyperbolic tangent activation function and the he_uniform initializer for the weight matrix kernel.

Line 16 shows the output layer which is a fully connected layer with three neurons and the softmax activation function.

d. Topology training

There were used the following parameters to compile and train the topology:

- Loss function: sparse categorical cross entropy
- Optimizer: Adam

It was trained using a Windows platform, version 10.0.18362. An Intel64 Family 6 Model 85 Stepping 4, GenuineIntel processor was used. Architecture AMD64.

III. EXPERIMENTAL RESULTS

a. Metrics definition

The efficiency of the model is measured using the metrics: Accuracy, Precision, Recall and F1 Score. To measure these metrics, we make use of True positive (TP), False positive (FP) and False negative (FN) cases. TP is when the model correctly predicts the positive class. FP is when the model incorrectly predicts the positive class. FN is when the model incorrectly predicts the negative class. You can understand the metrics applying the following questions:

Accuracy: How many predictions the model classified right? It is defined by the formula:

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN}$$

Precision: Of the instances classified as #, How many are actually #? It is defined by the formula:

$$Precision = \frac{TP}{TP+FP}$$

Recall: Of the instances that are actually #, How many are classified as #? It is defined by the formula:

$$Recall = \frac{TP}{TP+FN}$$

F1 Score: Balance between Precision and Recall. It is defined as the harmonic mean of Precision and Recall, and its formula is:

$$F1\ Score = 2 \times \frac{Precision \times Recall}{Precision+Recall}$$

To measure these metrics, the scikit-learn library version 0.22.1 was used.

b. Obtained results

The model was trained with a number of 12 periods and a batch size of 60.

Figure 3 shows how accuracy of training and validation sets increases through 12 epochs. The accuracy was 97.2%.

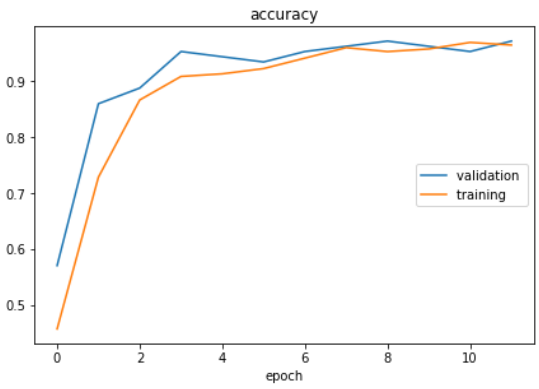

Figure 3. Accuracy per epoch.

Figure 4 shows how loss of training and validation sets decreases through 12 epochs. The loss was 14.5%.

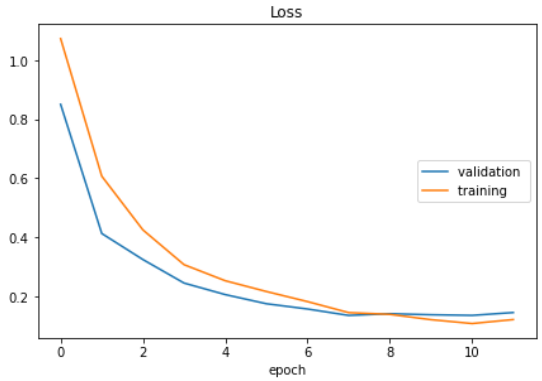

Figure 4. Loss per epoch.

Table 2 shows how many classes were actually predicted as they should be, and how many of them were not. Of the 107 images used in training, 104 images were predicted correctly and 3 images were not. It predicted one case as COVID that is actually healthy, and two cases predicted as healthy that are actually pneumonia. In this case, the model did not predict, for example, a case with healthy or pneumonia, when it is actually COVID. However, it predicted as COVID one case that is actually healthy.

		Predicted		
		Covid	Pneumonia	Normal
True	Covid	30	0	0
	Pneumonia	0	36	2
	Normal	1	0	38

Table 2. Confusion matrix.

Table 3 shows the percentage of Precision, Recall and F1 Score of the three classes. Precision in pneumonia is 100%, which means that all predictions as pneumonia were actually pneumonia. Recall is also 100% in COVID, which means that, of all the images that were actually COVID, all were predicted as COVID.

[1] http://201.134.41.15:8889/tree/ANN/Oscar/Ex3/Covid

[2] http://201.134.41.15:8889/tree/ANN/Oscar/Ex3/Pneumonia

[3] http://201.134.41.15:8889/tree/ANN/Oscar/Ex3/Normal

[4] http://201.134.41.15:8889/edit/ANN/Oscar/Ex3/metadata.csv

	Precision	Recall	F1 Score
Covid	97%	100%	98%
Pneumonia	100%	95%	97%
Normal	95%	97%	96%
Avg	**97%**	**97%**	**97%**

Table 3. Precision, Recall and F1 Score.

IV. CONCLUSIONS AND FUTURE WORK

In this paper, a classification model of a CNN was proposed. The topology was created using Tensorflow and it is composed by 16 layers. A total of 534 images were used, which include the 3 classes: COVID-19, Pneumonia and Healthy patients. 178 images per class. The results shows an accuracy, precision, recall and f1 score of 97% through 12 epochs.

As future work, we expect to recollect more x-ray images of COVID-19, specifically, when it is in early stages. Also, recollect more varied x-ray images of Pneumonia and Healthy patients.

REFERENCES

[1]"Preguntas y respuestas sobre la enfermedad por coronavirus (COVID-19)", *World Health Organization*, 2020. [Online]. Available: https://www.who.int/es/emergencies/diseases/novel-coronavirus-2019/advice-for-public/q-a-coronaviruses#:~:text=sintomas. [Accessed: 28- Apr- 2020].

[2]"Advice on the use of point-of-care immunodiagnostic tests for COVID-19", *World Health Organization*, 2020. [Online]. Available: https://www.who.int/docs/default-source/coronaviruse/sb-2020-1-poc-immunodiagnostics-2020-04-08.pdf?sfvrsn=5b6820_2. [Accessed: 29- Apr- 2020].

[3]World Health Organization, 2020. [Online]. Available: https://apps.who.int/iris/bitstream/handle/10665/331501/WHO-COVID-19-laboratory-2020.5-eng.pdf?sequence=1&isAllowed=y. [Accessed: 29- Apr- 2020].

[4]P. Sethy, S. Behera, P. Ratha and P. Biswas, "Detection of Coronavirus Disease (COVID-19) Based on Deep Features and Support Vector Machine", Preprints.org, 2020. [Online]. Available: https://www.preprints.org/manuscript/202003.0300/v2. [Accessed: 13- May- 2020].

[5]A. Abbas, M. Abdelsamea and M. Gaber, "Classification of COVID-19 in chest X-ray images using DeTraC deep convolutional neural network", *Cornell University*, 2020. [Online]. Available: https://arxiv.org/abs/2003.13815. [Accessed: 13- May- 2020].

[6]J. Cohen, P. Morrison and L. Dao, "COVID-19 image data collection", *GitHub*, 2020. [Online]. Available: https://github.com/ieee8023/covid-chestxray-dataset. [Accessed: 02- Apr- 2020].

[7]P. Mooney, "Chest X-Ray Images (Pneumonia)", *Kaggle*, 2018. [Online]. Available: https://www.kaggle.com/paultimothymooney/chest-xray-pneumonia. [Accessed: 06- Apr- 2020].

[8]S. Chakraborty, "Detection of COVID-19 presence from Chest X-ray scans using CNN & Class Activation Maps", *Medium - towards data science*, 2020. [Online]. Available: https://towardsdatascience.com/detection-of-covid-19-presence-from-chest-x-ray-scans-using-cnn-class-activation-maps-c1ab0d7c294b. [Accessed: 14- Apr- 2020].