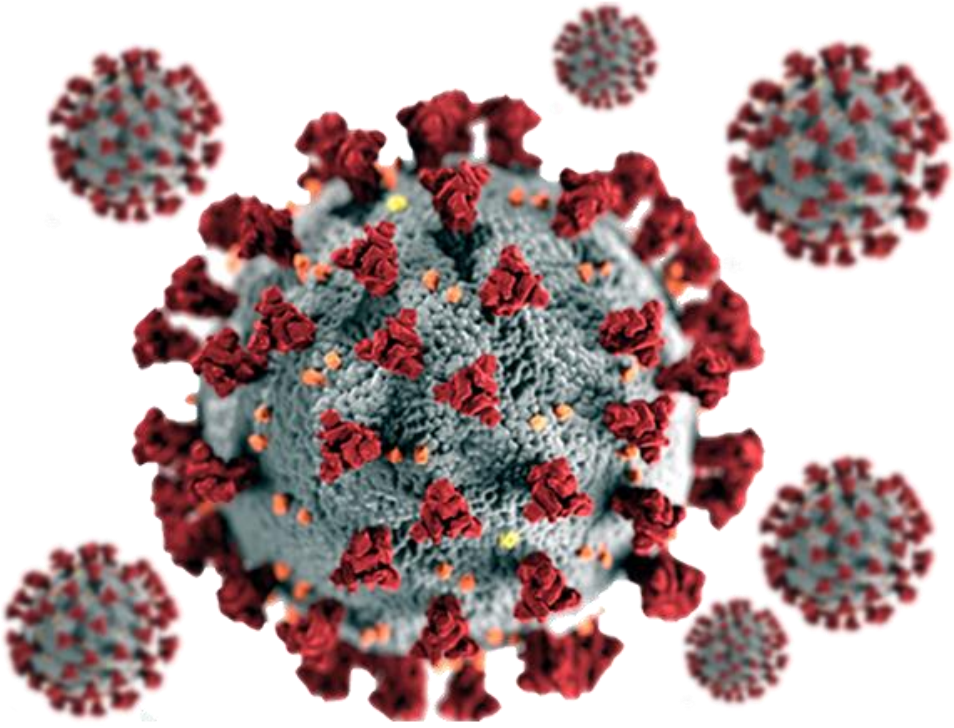
Detect Covid-19 via Chest X-Ray

Using convolutional neural network (CNN)

Ali Altamimi, Faisal Alasgah, Ghanim Alghanim, Omran Fallatah



Abstract

COVID-19 has caused many hospitals to be overcrowded and taken many lives worldwide. Our team used deep learning to help reduce the effects of COVID-19. We believe that deep learning would provide value in smart cities and hospitals where it can reduce the exposure of health workers to the virus and give faster and better detection results. Our idea is to use AI and deep learning to have zero exposure to patients testing for COVID-19, which is achieved by introducing an X-Ray room that has a robot that will scan patients and provide immediate results and further instructions to the patient. Also, in the future we can train the model to predict other diseases.

Design

The dataset was sourced from Kaggle. We started with a purpose to help our nation and the world against COVID-19 and arrived with a model that'll detect COVID-19 cases by using X-Ray in a scanning station by a robot. During our search we found that X-Ray gave better results for COVID-19 detection but exposed the health workers to possible positive cases and that was our inspiration.

Our design can be further developed in the future by integrating multiple models into one system that can detect multiple diseases.

Data

The original dataset had 30,000 labeled X-Ray images (data points) and four columns (ID, Data Source, File Path and Labels).

Algorithms

Data manipulation and cleaning.

- Applied feature Selection by dropping unnecessary columns.
- Applied feature reduction due to computational limitations and selected 10,000 stratified images out of 30,000 images.
- Applied data cleaning & Preprocessing by reshaping the images and converting them into grayscale.

After applying the previous steps to prepare the dataset, we end up with:

- 9,000 Images for Training.
- 1,000 Images for Validation.
- 400 Images for Testing.

Model Evaluation and Selection

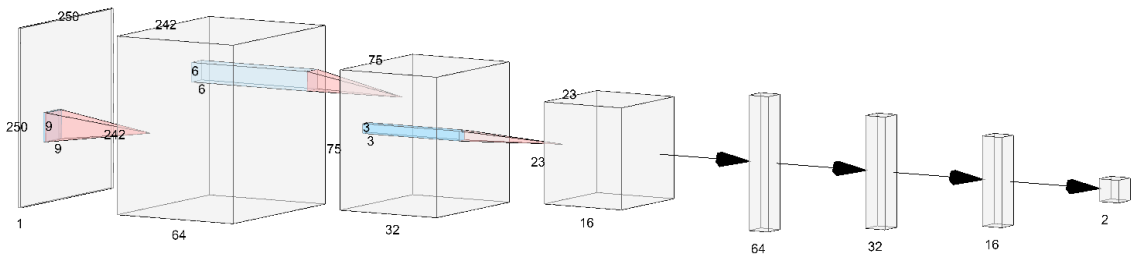
Since we are dealing with image classification the best course of action is to use convolutional neural networks since it's very effective in reducing the number of parameters without losing on the quality of models. Images have high dimensionality (as each pixel is considered as a feature).

Optimization Algorithm: Adaptive Moment Estimation (Adam).

Adam is an optimization algorithm that can be used instead of the classical stochastic gradient descent procedure to update network weights iteratively based on training data.

loss function: sparse categorical cross entropy.

Sparse categorical cross entropy is a loss function that is used in multi-class classification tasks. These are tasks where an example can only belong to one out of many possible categories, and the model must decide which one.



Model: "sequential_1"		
Layer (type)	Output Shape	Param #
conv2d_2 (Conv2D)	(None, 242, 242, 64)	5248
activation_3 (Activation)	(None, 242, 242, 64)	0
max_pooling2d_2 (MaxPooling2	(None, 80, 80, 64)	0
conv2d_3 (Conv2D)	(None, 75, 75, 32)	73760
activation_4 (Activation)	(None, 75, 75, 32)	0
max_pooling2d_3 (MaxPooling2	(None, 25, 25, 32)	0
conv2d_4 (Conv2D)	(None, 23, 23, 16)	4624
activation_5 (Activation)	(None, 23, 23, 16)	0
max_pooling2d_4 (MaxPooling2	(None, 7, 7, 16)	0
flatten_1 (Flatten)	(None, 784)	0
dense_3 (Dense)	(None, 64)	50240
dense_4 (Dense)	(None, 32)	2080
dense_5 (Dense)	(None, 16)	528
dense_6 (Dense)	(None, 2)	34
activation_6 (Activation)	(None, 2)	0

Total params: 136,514
Trainable params: 136,514
Non-trainable params: 0

Tools

- Pandas: Data Cleaning and manipulation
- Numpy: Mathematical operations
- Tensorflow: Deep Learning
- Seaborn: Data Visualization
- Sklearn: Testing the results

Communication

In addition to the slides and the visuals included in the presentation, we will submit our code and proposal. Below are our model scores.