Convolutional Neural Networks and its Applications: Covid 19 Detection

Mentors: Aadarsh Raj, Gudipaty Aniket, Sibasis Nayak, Sahasra Ranjan

Keywords: Convolutional Neural Networks, Image Processing, Machine Learning, Covid19, Deep Learning, Supervised Learning, Tensorflow

Team Member Name	Roll Number	Email-Id
Jash Kabra	200050054	jashkabra@gmail.com
Subarno Nath Roy	200050139	subarno.nathroy@gmail.com
Akshat Kumar	200050004	akshatkumar000@gmail.com
Bhavya Kohli	20D070021	bhavyakohli2024@gmail.com

Brief Description

This project's aim was to introduce use of Convolutional Neural Networks (CNNs) which is the state-of-the-art technique for various image-based Machine Learning problems. The project first covered learning the basics of Machine Learning where we were introduced to various libraries in our assignments used for data processing. We then moved on to specialise in CNNs, where we covered its structure and implementation from basics.
Particularly our sub project was to train and create a model to analyse the features of Chest Radiographs for Covid-19 detection with high accuracy using CNNs. Further, we expanded the model to predict if a person is diagnosed with Pneumonia or with Covid-19 from those Lung X-Rays.

Progress

The first few weeks of the project were aimed at introducing us firstly to Python and GitHub and then to the basics of Machine Learning like Linear Regression, Logistic Regression, Neural Networks etc. For this, we completed the Machine Learning course by Andrew Ng. Our learning was supplemented by assignments created by the mentors. This was completed before the first checkpoint.

In the coming weeks, we started learning about CNNs from the Deep Learning Specialization- Convolutional Neural Networks course by Andrew Ng. This involved learning about convolution and pooling operations, building a basic network, learning about and the case studies of popular deep CNNs used widely in research papers and implementing them. This completed our individual learning phase and marked the conclusion of the second checkpoint.

After this, we were split into 2 teams each of which implemented these models-binary and multiclass classification.The binary classification task involved classifying a radiograph as Covid +ve or Covid -ve whereas the multiclass classifier involved classifying a radiograph as Normal,

Covid +ve and Pneumonia. The work flow involved finding a suitable dataset, importing it, preprocessing the images, stratified splitting of the data into training and test data, setting up the model, training it on the train_data and finally evaluating it on the test data.

Challenges Faced:

1. Finding suitable dataset- We couldn't easily get a dataset having radiographs of all the three classes i.e. Covid-19 +ve, Normal, Pneumonia.
2. Finetuning Hyperparameters- We realised that a small change in the hyperparameters can create a large change in the result. Thus, a lot of time was spent in finding the ideal hyperparameters for both classifications.
3. Google Colab- Crashing of Runtime due to limited GPU and RAM. We couldn't do much about it and just used separate notebooks for each task.
4. Using grouped convolutions- Grouped convolutions requires a specific input shape and channels which we had to forcefully convert our images into.

Learning Value

Generalizing steps which require repetition or changing parameters depending on situations using functions.
Identifying separable steps and splitting the whole work into different steps to avoid crossing available RAM limits (on Colab).
Using image pre-processing techniques like blurring and resizing
Experimenting with model hyperparameters like layer arrangements, convolution filter size, number of filters/nodes, etc., to obtain the best results.
Using sufficient Dropout layers to minimize overfitting.
Training the model with callbacks to improve results (EarlyStopping, LearningRateScheduler), user experience (TqdmCallback) and ease in visualizing results (TensorBoard).

Results

Team 1: Subarno Nath Roy and Jash Kabra

Accuracy achieved:

i. Binary Classification= 97.42%

ii. Multiclass Classification= 97.09%

Team 2 : Akshat Kumar and Bhavya Kohli

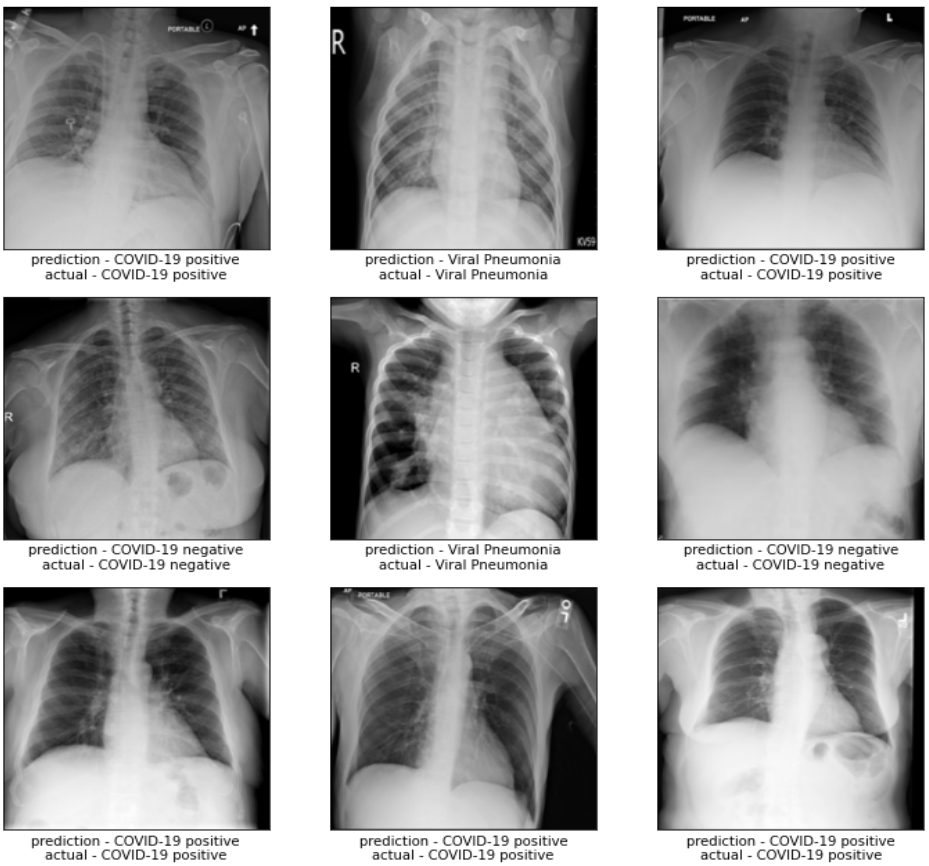
Accuracy achieved:

i. Binary Classification = 98.11%

ii. Multiclass Classification = 95.92%

Github Repository containing models of both teams-
https://github.com/adarsh0raj/SoC-CNN/tree/master/COVID19%20Disease%20Detection
Project Video-
https://drive.google.com/file/d/1AIuRVGR954HVUxuCy1FqnC4Xazl6idGi/view?usp=sharing

Working Prototype Image:



Software used

Google Colab
Jupyter Notebook
DeepNote

Suggestions for others

We think that just watching the ML course by Andrew Ng is not enough. Our mentors made some really great assignments for us as a supplement to the course and we highly recommend everyone who's doing the course to try out the assignments. They also cover basics of Github and Python libraries as a bonus point.
We would also recommend people learning about CNNs to not be afraid to try out many different architectures not necessarily given in a research paper as you never know which may work best for your model.

References and Citations

https://journals.physiology.org/doi/pdf/10.1152/physiolgenomics.00084.2020

(https://www.kaggle.com/tawsifurrahman/covid19-radiography-database)
DATA ACCESS AND USE: Academic/Non-Commercial Use
COVID-19 RADIOGRAPHY DATABASE from kaggle
M.E.H. Chowdhury, T. Rahman, A. Khandakar, R. Mazhar, M.A. Kadir, Z.B. Mahbub, K.R. Islam, M.S. Khan, A. Iqbal, N. Al-Emadi, M.B.I. Reaz, M. T. Islam, "Can AI help in screening Viral and COVID-19 pneumonia?" IEEE Access, Vol. 8, 2020, pp. 132665 - 132676. Paper link
Rahman, T., Khandakar, A., Qiblawey, Y., Tahir, A., Kiranyaz, S., Kashem, S.B.A., Islam, M.T., Maadeed, S.A., Zughaier, S.M., Khan, M.S. and Chowdhury, M.E., 2020. Exploring the Effect of Image Enhancement Techniques on COVID-19 Detection using Chest X-ray Images. Paper Link