Detection of Covid-19

Based on CNN

Report-1

CS3006D Computer Networks (MCA)

Dr. M.Prabu

Submitted by - Group 3

- Amanpreet Kaur
- Ankit Hammad
- Chetan Patidar
- Janvi Agrawal
- Jayant Parganiha
- Sheenam Waris

What is the Application?

The increasing frequency and magnitude of viral outbreaks in recent decades, epitomized by the current COVID-19 pandemic, has resulted in an urgent need for rapid and sensitive viral diagnostic methods. The symptoms of the disease can vary and include dyspnoea, high fever, runny nose, and cough according to WHO [1]. These cases can most commonly be diagnosed using chest X-ray imaging analysis for the abnormalities. Here, we present a methodology for virus detection and identification that uses a convolutional neural network (CNN). Our approach with deep learning can help rapidly detect viruses using feature extraction and trained CNN models which observe between distinct X-ray images of different virus-positive and virus-negative samples. First by using convolution we extract the specific feature from the input image and tend to experiment with fewer layers using convolution matrices as we go inside the sub layers, the more features can be extracted. At last, we are having binary classified results which will contain only one output result. Based on our trained model we'll be able to know if a person is infected with covid-19 or not.

Why is the Algorithm best fit?

Deep learning has shown a dramatic increase in the medical applications in general and specifically in medical image-based diagnosis. Due to the very promising results provided by CNNs in medical image analysis and classification, they are inspired by the visual system of the human brain. The idea behind the CNNs is to make computers capable of viewing the world as humans view it. This way CNNs can be used in the fields of image recognition and analysis, image classification, and natural language processing. CNN has been used for a variety of classification tasks related to medical diagnosis such as lung disease, breast cancer detection, interstitial lung disease, diagnosis of skin cancer by classification, and automatic diagnosis of various chest diseases using chest X-ray image classification. The state-of-the-art CNN architectures for the automatic detection of patients with COVID-19 and some of the CNN methods used for resolving Covid-19 problems have been mentioned [2]. They appear to be suitable for application to time series, images, or video. CNNs take raw data, without the need for an initial separate pre-processing or feature extraction stage: in a CNN the feature extraction and classification stages occur naturally within a single framework.

Other existing algorithms used for the application

According to the project we need algorithms that can be used for image classification and image recognition. We find CNN algorithm as the best fit for our project as the algorithm solves the problem with 97% accuracy which is better than other mentioned algorithms.

Single-shot detector (SSD) algorithm [3]

regions with convolutional neural networks (R-CNN) algorithm[4]

You Look Only Once(YOLO) algorithm [5]

How the algorithm chosen is advantageous over existing algorithms with respect to the application

as mentioned in ii) point we can use R-CNN, SSD and YOLO algorithms also but still we are going to use CNN only in this project as CNN is easy to Understand and very fast to implement .by using CNN we can achieve the highest accuracy as compared to all other Algorithms.

Why not R-CNN?

R-CNN takes more time than CNN as the R-CNN image first divides into so many regions and also the size of all regions is also not fixed. We need uniform- size regions . after division into regions, we have to apply the CNN algorithm[6] .correct region will be inserted into the Neural Network which takes a lot of time as compared to a simple CNN algorithm. in the International Research Journal of Engineering and Technology(IRJET) report also it is that R-CNN is inefficient for convolutions on every regional proposals.

Why not SSD?

If we use the CNN algorithm we can achieve a 97% accurate result but by using SSD it will be less accurate. It required a large number of input images to get the most optimum result[7],.So that's why we are not using the SSD algorithm in our project .

Why not YOLO ?

YOLO is the best in terms of speed and accuracy together as it is the latest algorithm in this domain but due to the spatial constraints of this algorithm, it won't give good results in the case of small objects.it also struggles to generalize Objects in New aspect ratios[8]. That's why it will not be beneficial to use this algorithm in our project.

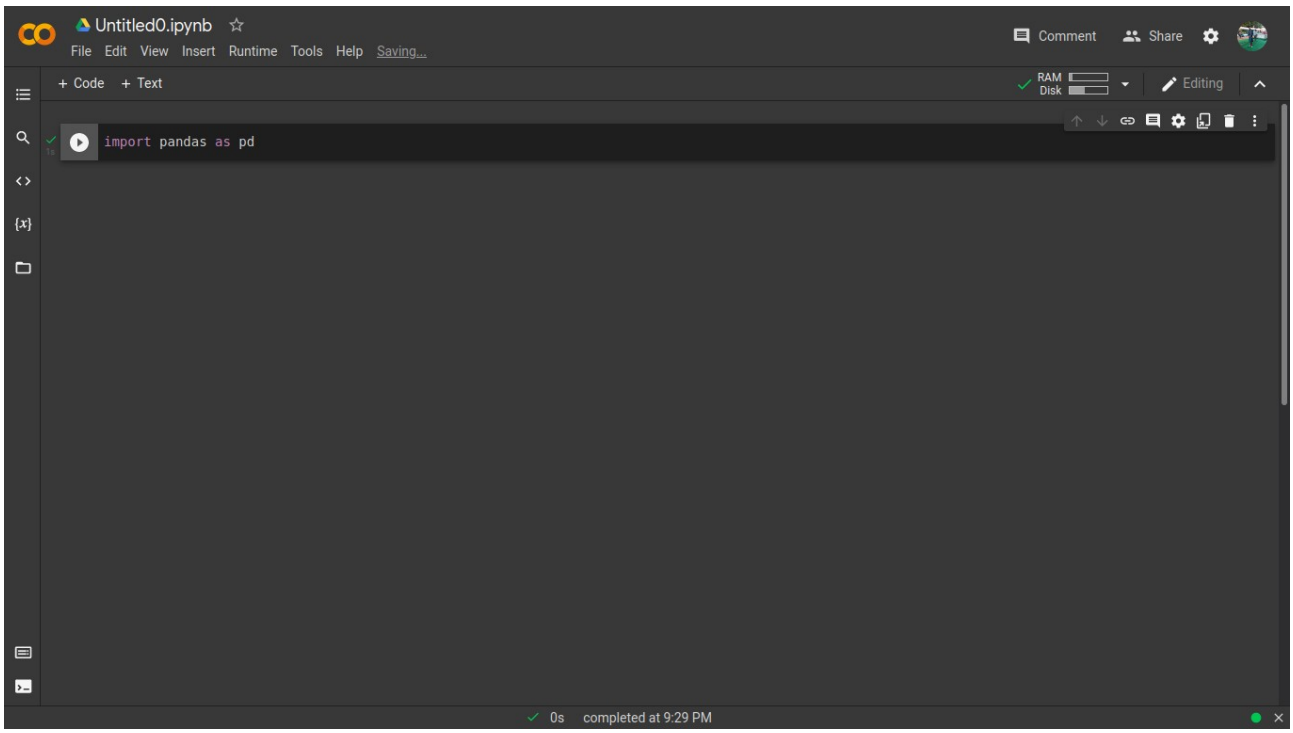
python modules that will be used for implementation :

Pandas : Pandas is a Python library used for working with data sets.It has funtions for analyze, cleaning, exploring, and manipulating data.Pandas help to clean messy dataset and make them relevant.

```
import pandas as pd
```

What Can Pandas Do?

- Read dataset.
- Quick overview of the DataFrame.
- What is average value?
- Replace empty cells .
- Remove row or coloum.



Python's standard utility modules

OS: Os module is used to intreact with operating system. Os module help to create new directorys, check current directorys,delete directory,edit directorys.

- **os.getcwd() :**To get the location of the current working directory.
- **os.chdir()** :To change the current working directory.
- **os.mkdir()** :To create new directory.
- **os.listdir()** :To get the list of all files and directories.
- **os.rmdir()** :To remove or delete an empty directory.

Random: Random modules is required to shuffle dataset and choice some random images for tranning and test dataset.

- **choice() :**Returns a random element from the given sequence
- **random() :**Returns a random float number between 0 and 1
- **uniform() :**Returns a random float number between two given parameters
- **shuffle() :**Takes a sequence and returns the sequence in a random order

Shutil: This module helps in automating process of copying and removal of files and directories.

- **shutil.copy() *:***To copy the content of source directory to destination directory
- **shutil.move() *:***To move the source directory to another location and returns the destination
- **shutil.rmtree()** :To delete an entire directory tree

API used :

Keras : Keras is an open source API(Application Programming Interface). Using class from the keras library makes it pretty simple to build a CNN. It makes the convolutional neural network more advanced and capable enough of classifying images

- Keras is a high-level neural network library that runs on top of TensorFlow.
- Keras is a simple-to-use but powerful deep learning library for Python.
- Developed with a focus on fast experimentation.

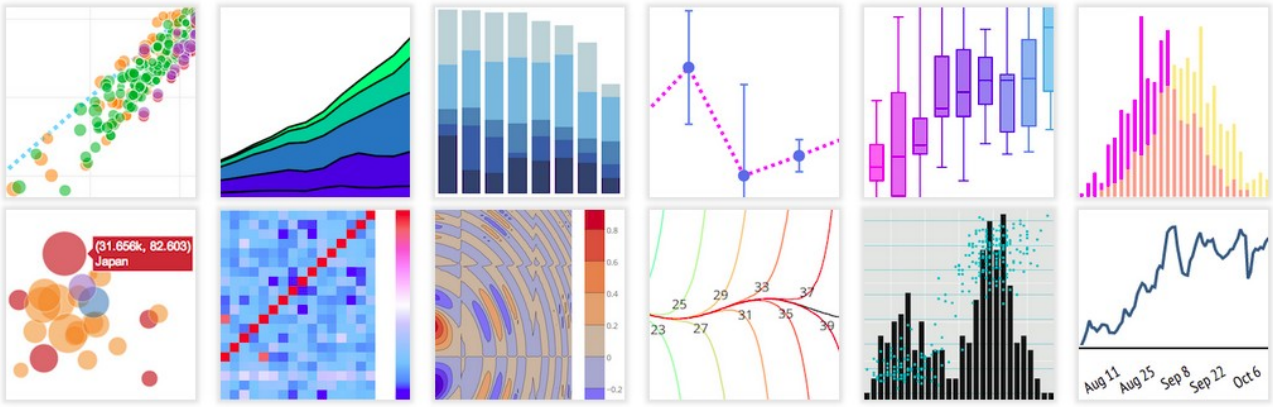
Data Visualization Libraries

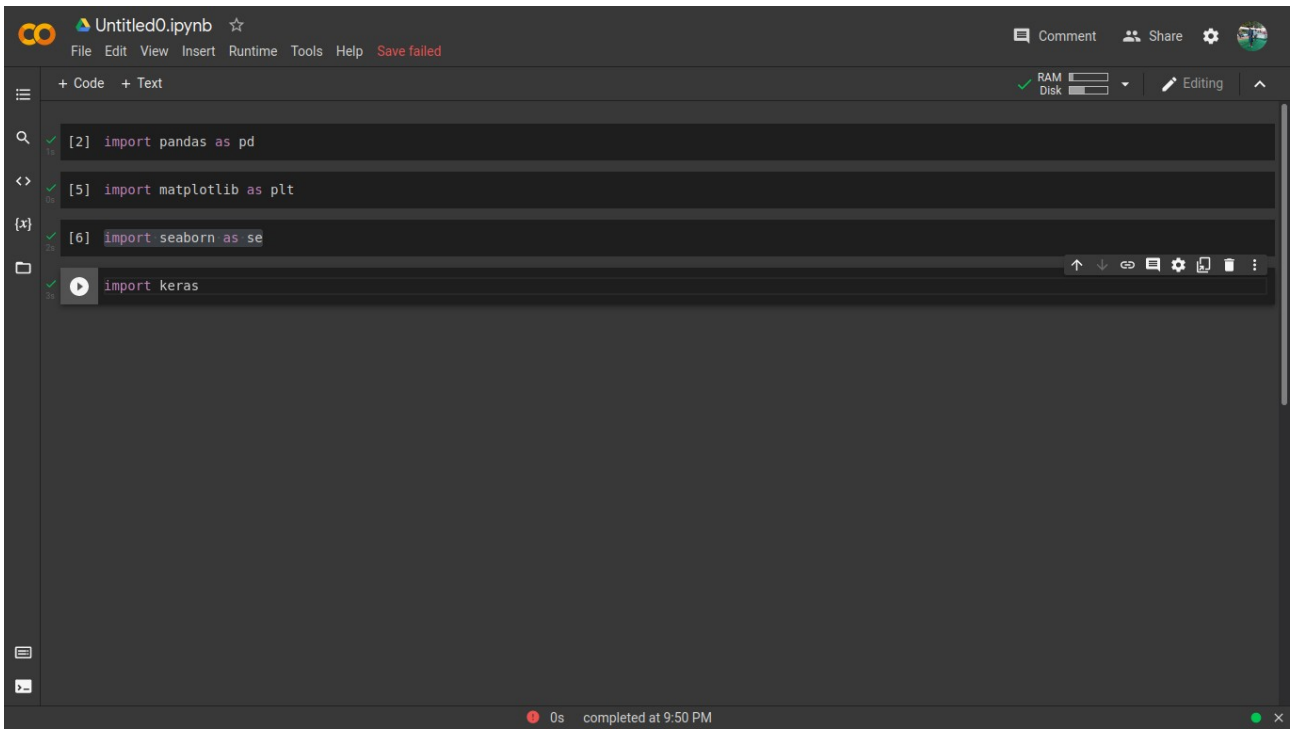
Matplotlib: Matplotlib is used for plotting graphs. It is a powerful tool for visualizing data.

Seaborn: Seaborn is superset of the Matplotlib library. It provides a variety of advanced visualization plots with simple syntax like box plots, violin plots, dist plots, Joint plots, pair plots, heatmap, and many more. Seaborn is also commonly used to plot confusion matrix.

Usage of Seaborn

- Data visualization
- Use to show data as a line plot.
- Use to create high level informative and attractive plots to show the data.





Dataset:

According to the project we need chest X-Rays images, such that there will be two sections one for training and one for validation. We will train our model with normal chest X-Rays which we got from kaggle[9].

Next, we also needed covid infected people's chest X-Ray images which we got from GitHub[10]. There is a group of doctors which keeps uploading the covid infected chest X-Ray for research purposes.

We took about 200 samples from both datasets. Image sizes were too big so we used google collab.

References

[1] World Health Organization, "who," 2020. [Online]. Available: https://www.who.int/health-topics/coronavirus#tab=tab_3.

[2] J. N. B. D. Nayak, "Springer," 15 June 2020. [Online]. Available: https://link.springer.com/article/10.1007/s42600-021-00135-6/tables/3.

[3] COVID Detection in Chest X-ray Images using a Deep Learning Approach. International Journal of Interactive Multimedia and Artificial Intelligence. InPress. 1.10.9781/ijimai.2020.04.003.

[4] A. Krizhevsky, I. Sutskever, and G. E. Hinton, “ImageNet classification with deep convolutional neural networks,” Communications of the ACM, vol. 60, no. 6, pp. 84–90, 2017.

[5] Khan AI, Shah JL, Bhat MM (2020) CoroNet: a deep neural network for detection and diagnosis of COVID-19 from chest X-ray images. Compute Methods Programs Biomed 196:105581.

[6] https://www.quora.com/What-is-the-difference-between-CNN-and-R-CNN

[7] International Research Journal of Engineering and Technology (IRJET) http://www.irjet.net/

[8] International Research Journal of Engineering and Technology (IRJET) http://www.irjet.net/

[9] https://www.kaggle.com/paultimothymooney/chest-xray-pneumonia

[10] https://github.com/ieee8023/covid-chestxray-dataset/tree/master/images