Capstone Project

On

CT Scan Image Classification

Data Overview:

This dataset contains 1252 CT scans that are positive for SARS-CoV-2 infection (COVID-19) and 1230 CT scans for patients non-infected by SARS-CoV-2, 2482 CT scans in total. These data have been collected from real patients in hospitals from Sao Paulo, Brazil. The aim of this dataset is to encourage the research and development of artificial intelligent methods which are able to identify if a person is infected by SARS-CoV-2 through the analysis of his/her CT scans.

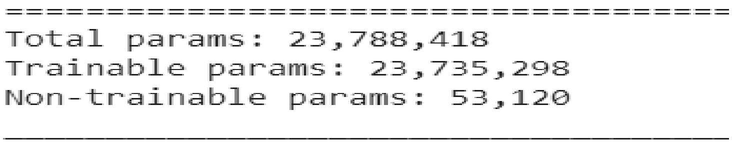
This Python 3 environment comes with many helpful analytics libraries installed. Google compute engine backend GPU.

Here's several helpful packages to load :

```
import numpy as np # linear algebra

import pandas as pd # data processing, CSV file I/O (e.g. pd.read_csv)

import matplotlib.pyplot as plt

%matplotlib inline

import seaborn as sns

import cv2

import os

from tqdm import tqdm

from sklearn.metrics import confusion_matrix

from sklearn.model_selection import train_test_split

from keras.utils.np_utils import to_categorical

from keras.models import Model,Sequential, Input, load_model

from keras.layers import Dense, Dropout, Flatten, Conv2D, MaxPool2D, BatchNormalization, AveragePooling2D, GlobalAveragePooling2D

from keras.optimizers import Adam

from keras.preprocessing.image import ImageDataGenerator

from keras.callbacks import ModelCheckpoint, ReduceLROnPlateau
```

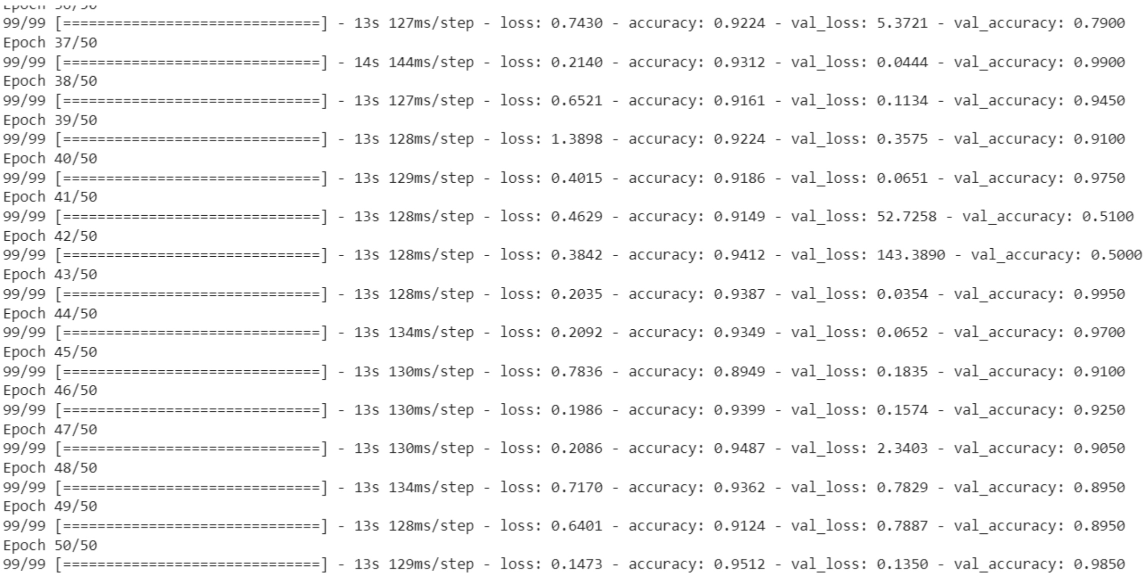
- Displaying images of Covid and Non-Covid
- Image resize and read function / Image augmentation
- Train the images
- Convert labels in to categorical data
- Splitting data in train and test
- 64*64 training data/images
- Data Augmentation and fitting model
- Loss and accuracy : Loss – 0.13 and **accuracy – 0.98**
- Confusion matrix
- Activation function : Relu and Sigmoid
- Early stopping
- Plot ROC curve
- Resnet50
- Plotting accuracy and loss plot
- Classification report
- Finally prediction from Image

Model Summary :

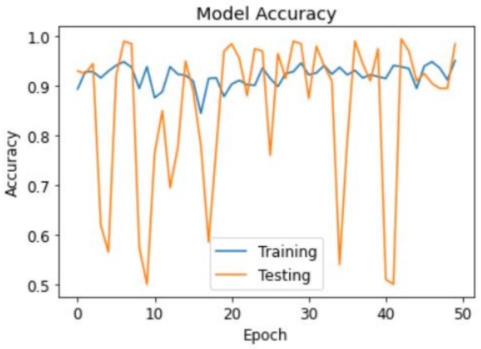


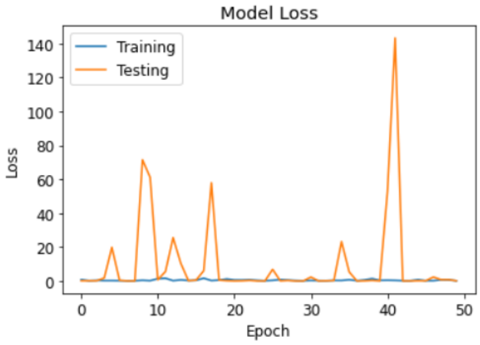
```
=================================================================
Total params: 23,788,418
Trainable params: 23,735,298
Non-trainable params: 53,120
_________________________________________________________________
```

Fitting model :



```
Epoch 36/50
99/99 [==============================] - 13s 127ms/step - loss: 0.7430 - accuracy: 0.9224 - val_loss: 5.3721 - val_accuracy: 0.7900
Epoch 37/50
99/99 [==============================] - 14s 144ms/step - loss: 0.2140 - accuracy: 0.9312 - val_loss: 0.0444 - val_accuracy: 0.9900
Epoch 38/50
99/99 [==============================] - 13s 127ms/step - loss: 0.6521 - accuracy: 0.9161 - val_loss: 0.1134 - val_accuracy: 0.9450
Epoch 39/50
99/99 [==============================] - 13s 128ms/step - loss: 1.3898 - accuracy: 0.9224 - val_loss: 0.3575 - val_accuracy: 0.9100
Epoch 40/50
99/99 [==============================] - 13s 129ms/step - loss: 0.4015 - accuracy: 0.9186 - val_loss: 0.0651 - val_accuracy: 0.9750
Epoch 41/50
99/99 [==============================] - 13s 128ms/step - loss: 0.4629 - accuracy: 0.9149 - val_loss: 52.7258 - val_accuracy: 0.5100
Epoch 42/50
99/99 [==============================] - 13s 128ms/step - loss: 0.3842 - accuracy: 0.9412 - val_loss: 143.3890 - val_accuracy: 0.5000
Epoch 43/50
99/99 [==============================] - 13s 128ms/step - loss: 0.2035 - accuracy: 0.9387 - val_loss: 0.0354 - val_accuracy: 0.9950
Epoch 44/50
99/99 [==============================] - 13s 134ms/step - loss: 0.2092 - accuracy: 0.9349 - val_loss: 0.0652 - val_accuracy: 0.9700
Epoch 45/50
99/99 [==============================] - 13s 130ms/step - loss: 0.7836 - accuracy: 0.8949 - val_loss: 0.1835 - val_accuracy: 0.9100
Epoch 46/50
99/99 [==============================] - 13s 130ms/step - loss: 0.1986 - accuracy: 0.9399 - val_loss: 0.1574 - val_accuracy: 0.9250
Epoch 47/50
99/99 [==============================] - 13s 130ms/step - loss: 0.2086 - accuracy: 0.9487 - val_loss: 2.3403 - val_accuracy: 0.9050
Epoch 48/50
99/99 [==============================] - 13s 134ms/step - loss: 0.7170 - accuracy: 0.9362 - val_loss: 0.7829 - val_accuracy: 0.8950
Epoch 49/50
99/99 [==============================] - 13s 128ms/step - loss: 0.6401 - accuracy: 0.9124 - val_loss: 0.7887 - val_accuracy: 0.8950
Epoch 50/50
99/99 [==============================] - 13s 129ms/step - loss: 0.1473 - accuracy: 0.9512 - val_loss: 0.1350 - val_accuracy: 0.9850
```





Conclusion: In this developed model, Resnet50, one of the CNN architectures, was used as the base. By using "Adam" optimizer and "sigmoid" activation function an accuracy rate of 98% was achieved.