DS 303 PROJECT

FaceMask Detection via CNN Frameworks in a Late-COVID-19 era

HONOR CODE

All creatives, ideation and technical work done in this this Project Assignment was carried forward by the three undersigned and no external persons were consulted.

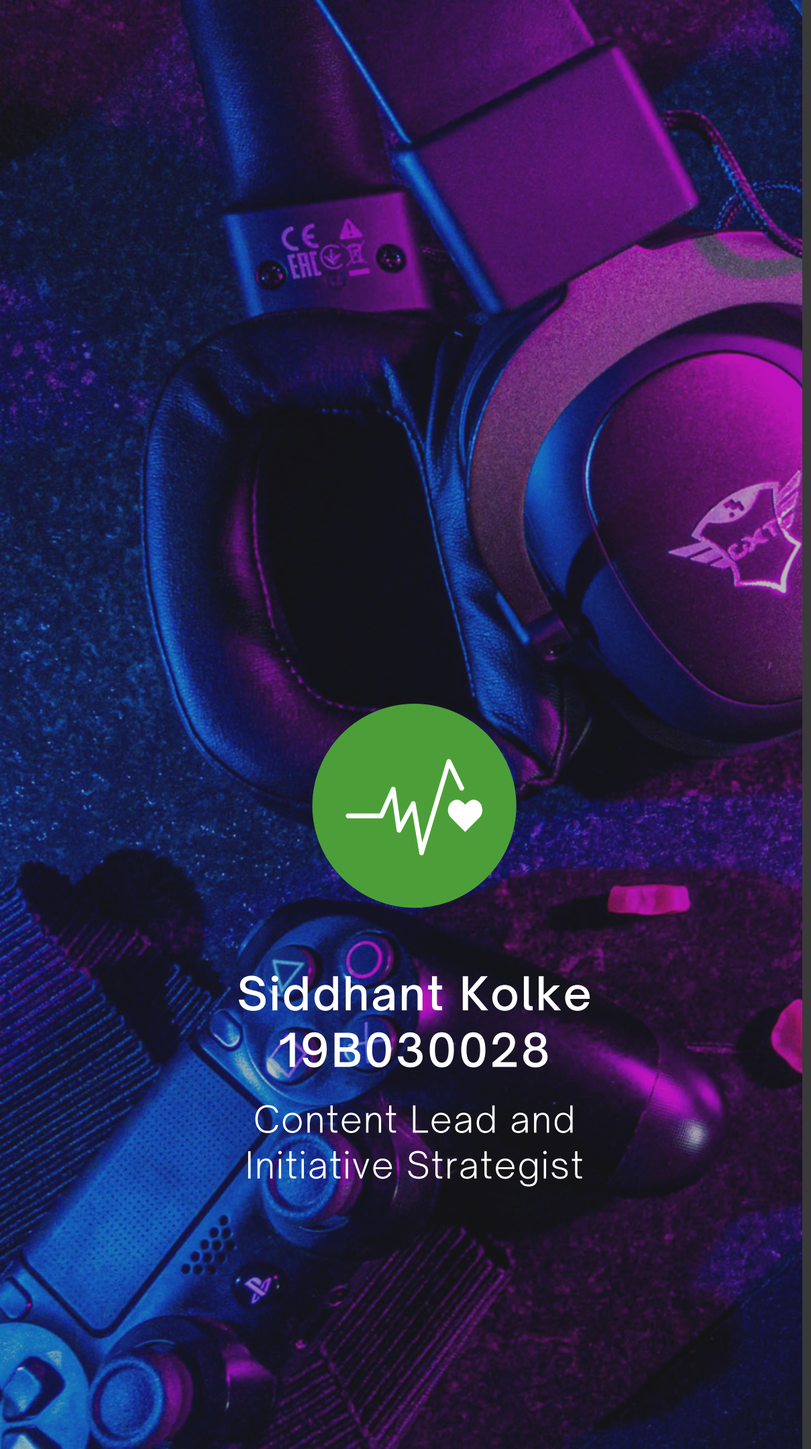
All internet resources, datasets and materials found online have been referenced in the last slide.

SIGNED, Aniket Yadav, Siddhant Kolke and Akshat Agarwal

5th May

Table of Contents

Group Members

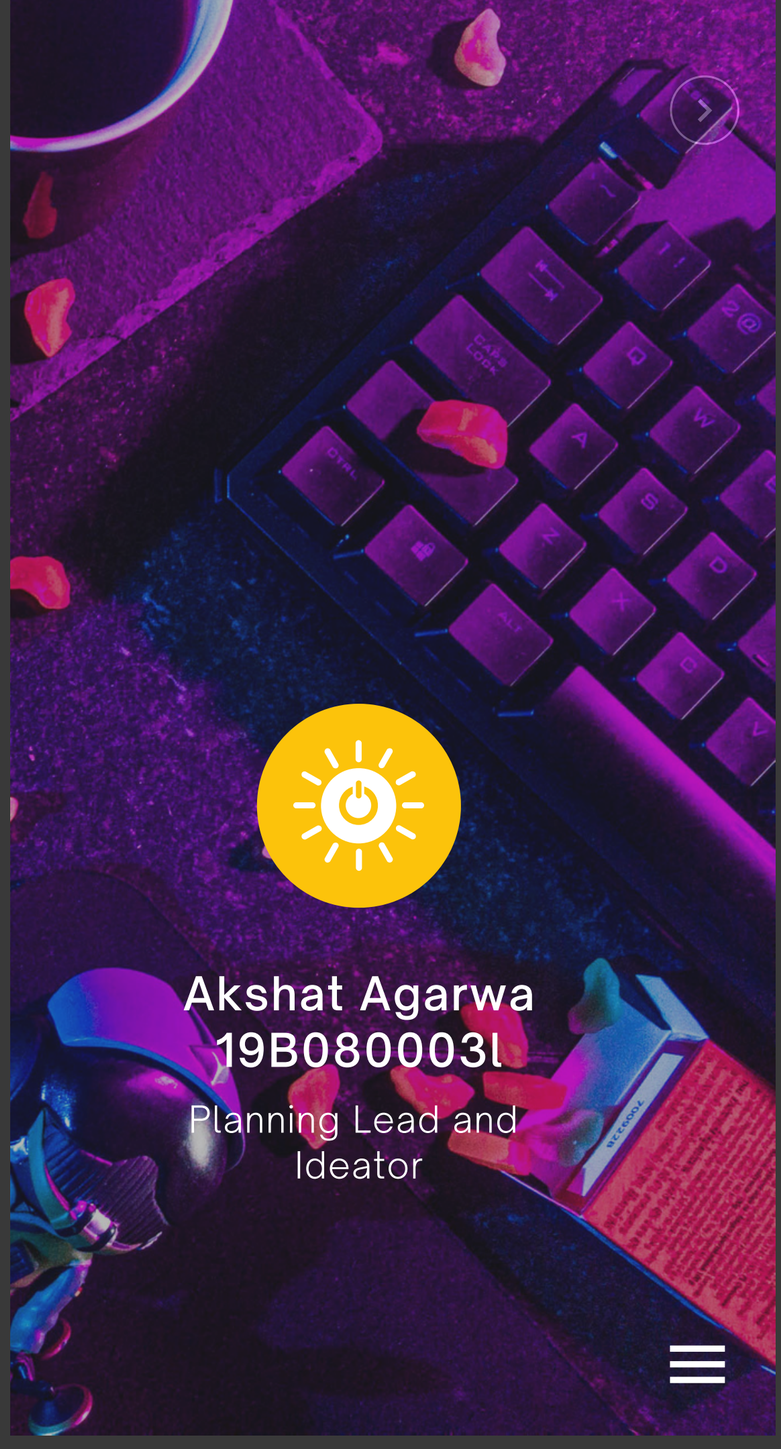


Siddhant Kolke
19B030028

Content Lead and
Initiative Strategist

Aniket Yadav
19D070009

Development Lead
and Executor



Akshat Agarwa
19B080003l

Planning Lead and
Ideator

PROBLEM STATEMENT

Can we implement neural networks, specifically CNNs to target distinguish between mask-wearers and defaulters?

PROBLEM STATEMENT WAS IDEATED AMIDST VARIOUS MEETS HELD, TAKING INTO ACCOUNT RELEVANCE AND TECHNOLOGY AT DISPOSAL.

FACE MASKS:

As 2021 heralds us with a ghastlier version of last year's outbreak in the form of a second, and probably third wave, it looks like we aren't getting to normal business anytime.

As marketplaces and eventually areas of commerce are re-populated, there's a growing need for public safety protocol such as:

-social distancing
-observing quarantine

arguably the most important, so as to avoid transmission of the virus through unavoidable human contact,

-the appropriate use of masks.

INTEGRATED TECHNOLOGY

Now let's look at the technology we have at our disposal! **CCTV cameras** have become ubiquitous and security has never been more proactive when it comes to recording registry and ensuring COVID protocol is followed at all times lest should a single bad apple spoil the bunch. Given that we have **access to video/ snapshots of entering or leaving individuals**, an **automated system** can be adopted in places where manpower can be replaced.

These cameras can be integrated with tools for computer vision in order to identify wrongly-worn masks and pick out the defaulting person. Since a registry always comes with a physical or even better, a digital timestamp, these can always be traced back to the person of interest, and hence time and money is saved in the overall picture of a minimal movement scenario where it is not feasible to place manpower or have consistency of rules be imposed at all people and at all times.

FInding a solution for: Detecting via photo or video input, NON-MASKED or MASKED test subject

Motivations

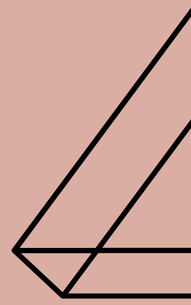
01 This has been increasingly challenging to enforce over the last year, police-work seems to be **inconsistent** and shopkeepers often have a hard time refusing business to customers for the ill-wearing of masks.

02 **Corporate areas** too, are squeezed for profits and will have to eventually transgress from the work-from-home phase.

03 Societies and residential colonies are THE areas where bubbles are bound to arise due to frequent mingling of residents and proximity. There is a strong demand for enforcement and identification for miscreants and defaulters as far as mask-wearing goes.

Source: Centers for Disease Control and Prevention

MODEL AND MATHEMATICAL CHARACTERISTICS



Neural Networks

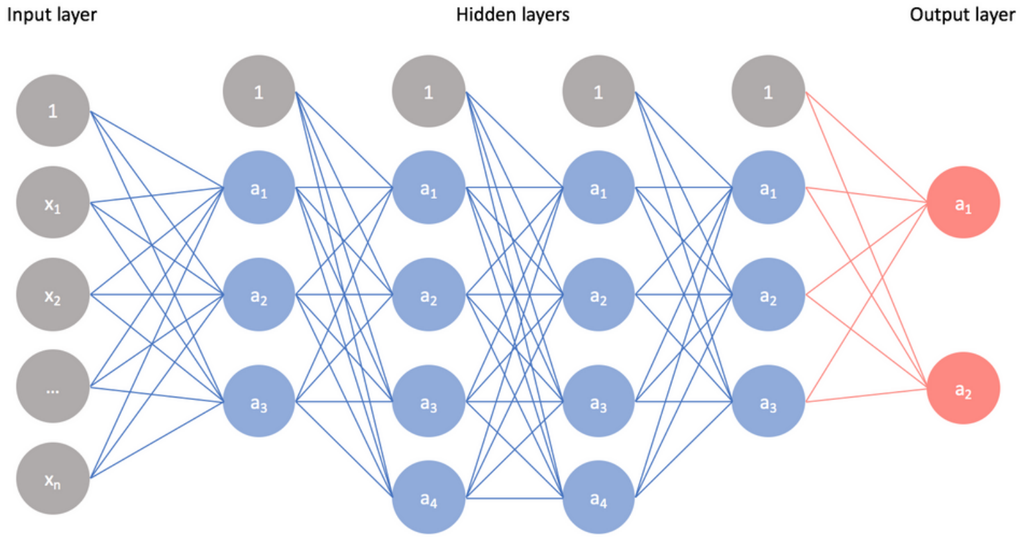
Write the first question here. It can be answerable with a multiple choice, an opinion, or a yes or a no followed by a brief explanation.

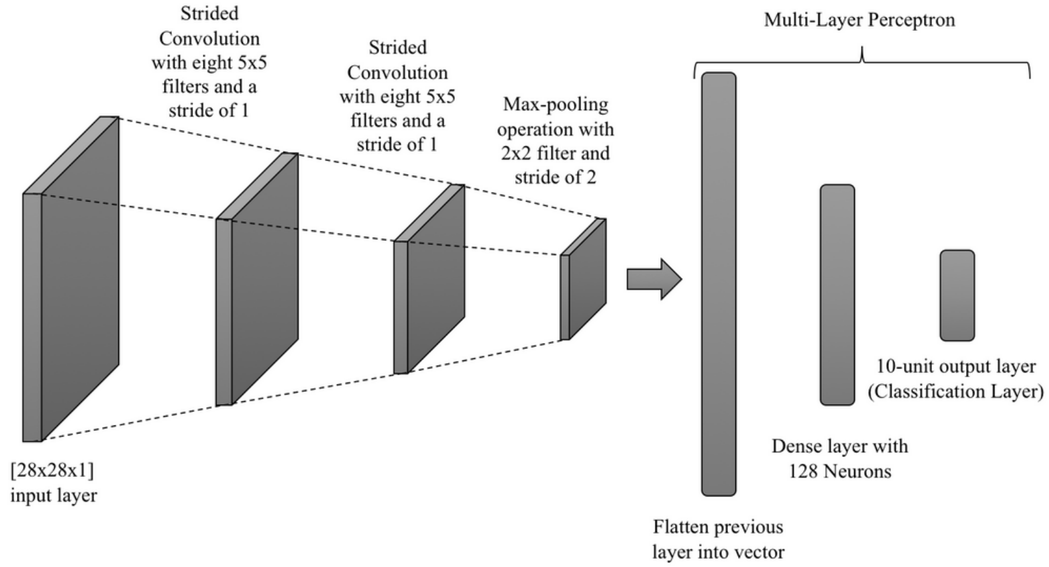
Convolutions

Write the first question here. It can be answerable with a multiple choice, an opinion, or a yes or a no followed by a brief explanation.

ReLU and SoftMax

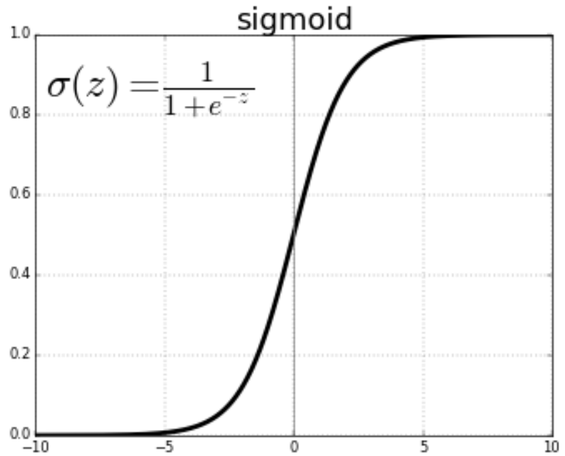
Write the first question here. It can be answerable with a multiple choice, an opinion, or a yes or a no followed by a brief explanation.

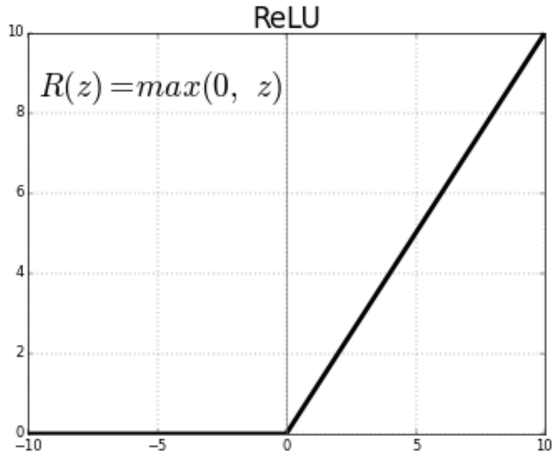




An ANN is based on a collection of connected units or nodes called artificial neurons.The weight increases or decreases the strength of the signal at a connection. Typically, neurons are aggregated into layers. Different layers may perform different transformations on their inputs. Signals travel from the first layer (the input layer), to the last layer (the output layer), possibly after traversing the layers multiple times.

Convolutions in a neural network are ideally implemented to deal with image vector data and such, where neighbouring pixels may be correlated and data is large in number (2D pixels and such)

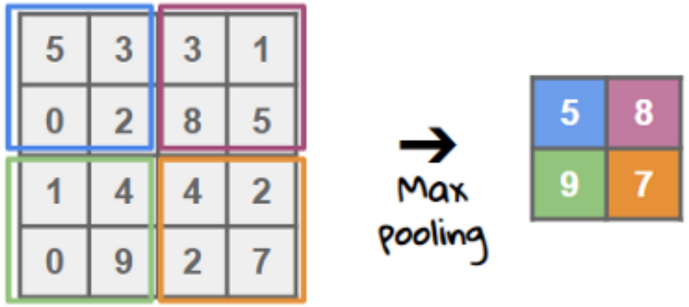




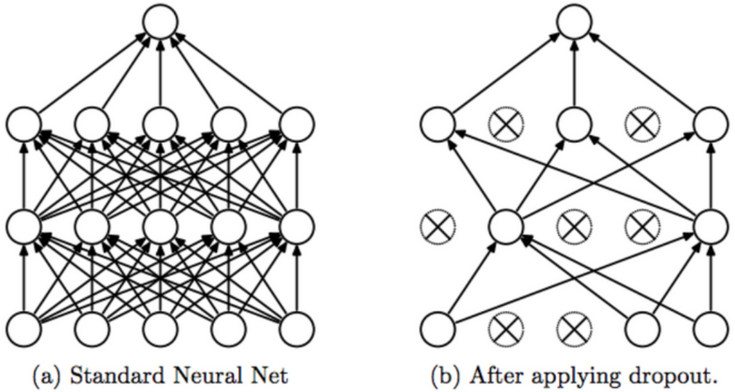
The signals from each node are weighted and sent through activation functions such as sigmoid and ReLU. As you can see, the ReLU is half rectified (from bottom). f(z) is zero when z is less than zero and f(z) is equal to z when z is above or equal to zero.

The final classification is probabilistically soft guided by soft-max sigmoid function. Cross entropy loss is then applied in our model as a learning metric for optimizing the neural net by backpropagation

Cross Entropy $J = y^* \log(y) + (1 - y^*) \log(1 - y)$

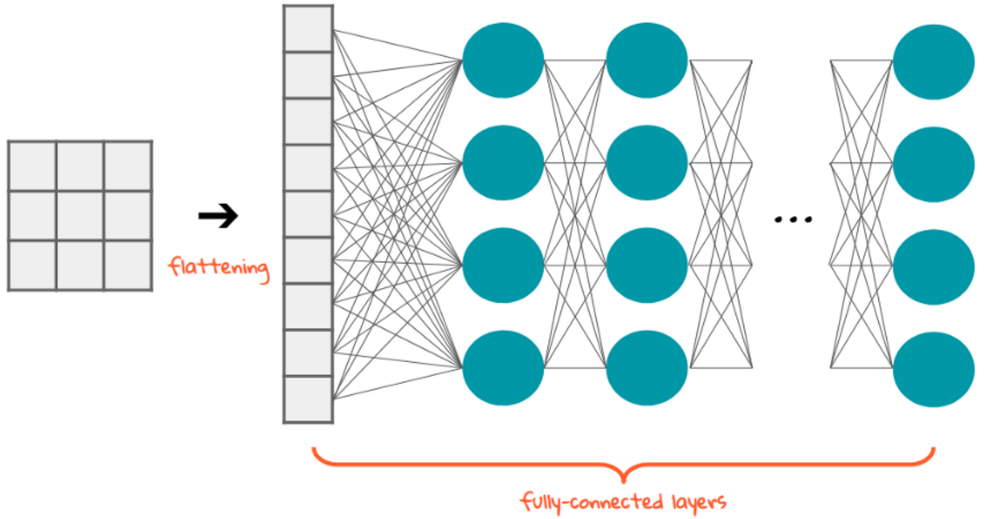


Sliding a window, we only take the maximum value inside the box on the left case. This is 'max pooling.'



(a) Standard Neural Net (b) After applying dropout.

Dropout is a technique used to prevent a model from overfitting. Dropout works by randomly setting the outgoing edges of hidden units (neurons that make up hidden layers) to 0 at each update of the training phase.



Flattening is converting the data into a 1-dimensional array for inputting it to the next layer. We flatten the output of the convolutional layers to create a single long feature vector. And it is connected to the final classification model, which is called a fully-connected layer.

THE DATASET

The dataset consisted of 1376 images, 690 face images with masks and 686 without masks.

METHOD W/ IMPLEMENTATION

Step 1

Data Pre-processing

The area of interest from the captured image and converted into a grayscale image It was resized into a 100*100 image. Now, it was normalized by dividing the image array size by 255

Step 2

Keras and CNN

Keras Libraries for 2D convolution and MaxPooling are used to implement the neural net. Output convolutions are flattened(to reduce overfitting) and fit into ReLU activation.

Step 3

Applying neural net

This is subsequently converted to a dense layer of 50 neurons. Finally Softmax activation is used to convert it into a dense layer of 2 neurons

STEP 4

Training CNN

compile the model for cross entropy loss with optimizer ”adam” and metrics Accuracy to get loss and accuracy. A train-test split of 0.9-0.1 was used

STEP 5

Selection of the Best model

20 models are prepared by setting epoch, and each were pit against each other to have minimum validation loss while ratining a high accuracy

STEP 6

Application towards detection

STEP 1 is repeated to pre-process the data from the camera (using OpenCV) and argmax is used to make a hard classification as follows:

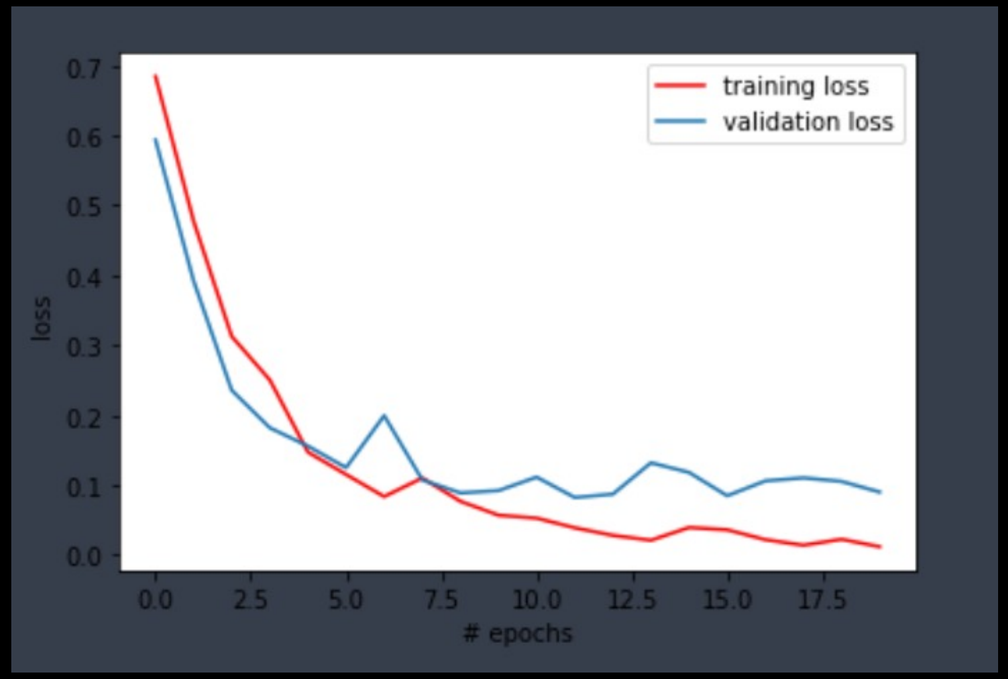
‘0’ for mask, gives “with mask”

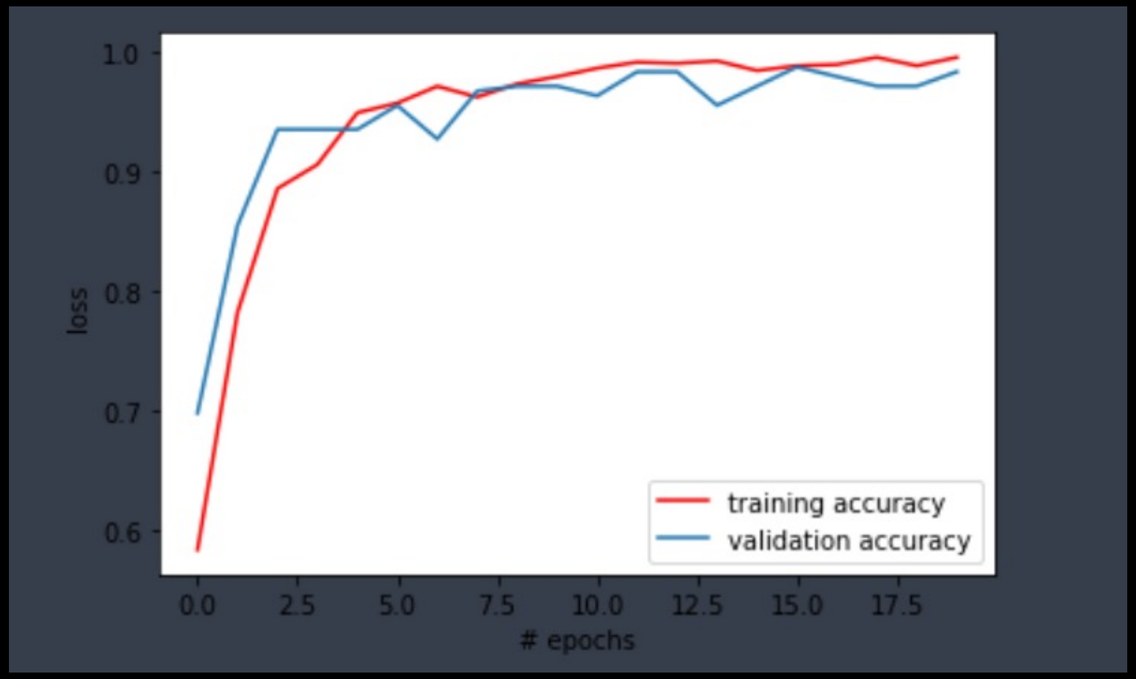
‘1’ for without mask, gives “without mask”

Results and Analysis

Out of the 20 epochs, the models with the minimum validation loss yet having a high accuracy was selected (validation set contained 20% of the data). The ultimate five models were found to have optimum values with the best one being Model 17. with validation accuracy at a high 0.9879 while maintaining a high overall accuracy of 0.9868.

```
Epoch 15/20
31/31 [==============================] - 57s 2s/step - loss: 0.0264 - accuracy: 0.9921 - val_loss: 0.1184 - val_accuracy: 0.9718
Epoch 16/20
31/31 [==============================] - 57s 2s/step - loss: 0.0338 - accuracy: 0.9894 - val_loss: 0.0850 - val_accuracy: 0.9879
Epoch 17/20
31/31 [==============================] - 57s 2s/step - loss: 0.0223 - accuracy: 0.9868 - val_loss: 0.1063 - val_accuracy: 0.9798
Epoch 18/20
31/31 [==============================] - 58s 2s/step - loss: 0.0136 - accuracy: 0.9990 - val_loss: 0.1106 - val_accuracy: 0.9718
Epoch 19/20
31/31 [==============================] - 58s 2s/step - loss: 0.0205 - accuracy: 0.9914 - val_loss: 0.1055 - val_accuracy: 0.9718
Epoch 20/20
31/31 [==============================] - 57s 2s/step - loss: 0.0134 - accuracy: 0.9937 - val_loss: 0.0903 - val_accuracy: 0.9839
```



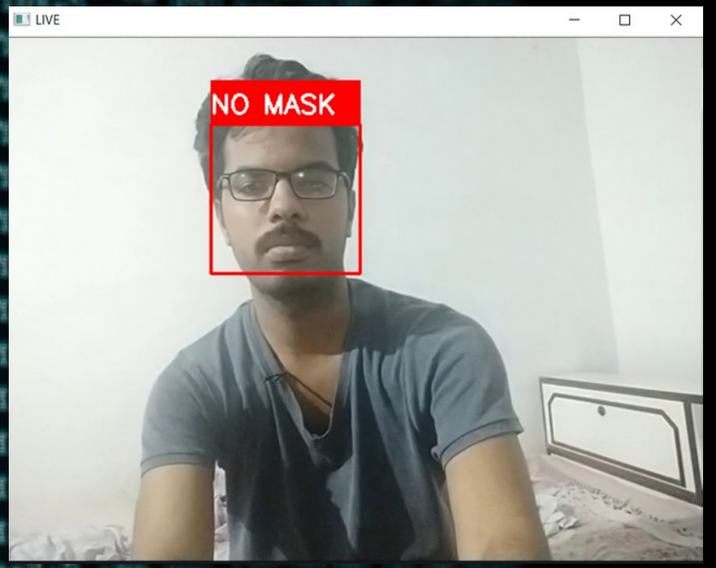


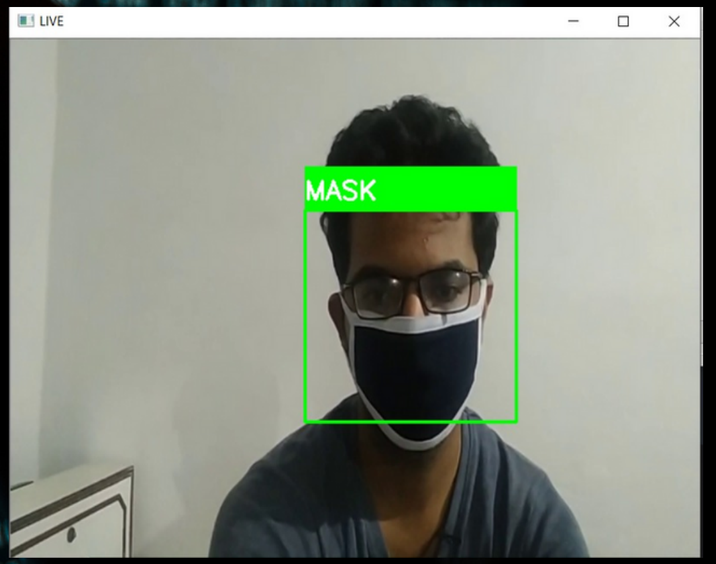
Training Loss for each model was plotted against epochs, a mental superimposition of the two graphs maiy lead us to the best model (epoch=17)

Sample Test

To test the efficacy of our model, we integrated via OpenCV, the webcam to our CNN framework. As you can see, Aniket-a valued member of our team was found to be a responsible citizen in Exhibit A (top right image) and defaulting upon the COVID-19 Rules in exhibit B (bottom left image)

Ruling the following, and with a ~98% accurate model, we regard this project a success!





References and Acknowledgements

1) Canva.com for design and vectors.

2) Study Material: https://towardsdatascience.com/the-most-intuitive-and-easiest-guide-for-convolutional-neural-network-3607be47480

3) DataSet, by Prajna Bhandari: https://www.linkedin.com/in/prajna-bhandary-0b03a416a?lipi=urn%3Ali%3Apage%3Ad_flagship3_profile_view_base_contact_details%3BIEyykOCzQGeLmarKAAY4lg%3D%3D