Training

#-line no. ;)
Libraries

Imutile: for easy translation, rotation etc of the image

Mobile net_v2: A rather light deep learning module hence can be used for slower infrastructures like Raspberry pi or phones.
Alpam complex aanu, enikku thanne velya karyam ayittu ariyilla ithu

Command line arguments

- --dataset: The path to the input dataset of faces and and faces with masks
- --plot: The path to your output training history plot, which will be generated using matplotlib
- --model: The path to the resulting serialized face mask classification model

Hyper Parameters

INIT_LR: gradient descent learning rate(initial)

Epochs:One Epoch is when an ENTIRE dataset is passed forward and backward through the neural network only ONCE.
We are using a limited dataset and to optimise the learning and the graph we are using **Gradient Descent** which is an ***iterative*** process. So, *updating the weights with single pass or one epoch is not enough*

Batch Size:Total number of training examples present in a single batch.
Eg:We can divide the dataset of 2000 examples into batches of 500 then it will take 4 iterations to complete 1 epoch.

#56 Pre-processing steps include resizing to *224×224* pixels, conversion to array format, and scaling the pixel intensities in the input image to the range *[-1, 1]*

ImageDataGenerator: generates batches of tensor image data with real-time data augmentation, its used since we don't have to store the augmented data in disk/memory.
Enables rotation, width shifting, height shifting, channel shift etc.

Fine-tuning(baseModel= MobileNetV2 thottu):

- Load MobileNet with pre-trained **ImageNet**weights, leaving off head of network

- Construct a new FC head, and append it to the base in place of the old head
- Freeze the base layers of the network The weights of these base layers will not be updated during the process of backpropagation, whereas the head layer weights *will* be tuned

Why we fine tune it:

1. Greatly reduced training time
2. Improved performance
3. Counters over fitting on small datasets

#108-109
Adam as optimiser, its a learning rate decay schedule and binary cross entropy
#115
Aug will provide batches of mutated image data