Face Mask Detection
Yolov5_Tuned

Abstract- The task of finding proportion of people wearing masks in a given image can be considered as a task fine tuned object detection on face detection with classes as wearing mask or not wearing mask.
Recent studies show that deep learning approaches can achieve impressive performance on these tasks with fine tuning.
For the purpose of object detection we use Yolov5s by Ultralytics pretrained on COCO dataset and for fine tuning purpose kaggle dataset is used.

1. Architecture

The paper for Yolov5 is not released till now, but according to [1][1] nn.SiLU() activations is replaced nn.LeakyReLU(0.1) and nn.Hardswish() activations used in previous versions simplifying the architecture as now only one single activation function used everywhere rather than the two types before.
The Yolov5s contains 7.3M parameters with a speed of 2.2.ms.

2. Training Specifications

For fine tuning purpose kaggle Face Mask Detection dataset is used containing 3 classes as correctly wearing a mask, not wearing a mask and incorrectly wearing a mask.
As a preprocessing step the kaggle dataset which is available in yaml format is converted to Densenet format as required for Yolo training.
Model is trained for 200 epochs with batch size of 16 with SDG as optimizer. Learning rate of 0.01 is used with momentum of 0.937 and nesterov as True.
Best Recall of 0.859 and best precision of 0.692 is obtained during 200 epochs. Final total loss is 0.0335 is obtained with box regression, object detection and class loss as 0.01866, 0.01393, 0.0009996 respectively.
On testing confidence threshold of 0.3 and iou_threshold of 0.45 is used.
Even though of such small dataset for fine tuning Yolov5 performed good but as much of the pretrained for object detection.

[1] https://zenodo.org/record/4418161#.X_1ZYdgzZPY

3. Results

Green box for wearing mask correctly otherwise red box