Auto Rickshaw detection using Yolov5

Task:

“Auto” Detection using YOLO. In this task, we have to detect auto rickshaws present within an image using Faster RCNN/YOLO based methods.

Dataset:

For this task, we made use of the given dataset which contains a total of 152 images some containing and not containing autos. However, the dataset was unannotated, i.e. the bounding rectangles of autorickshaws weren’t given.

Hence, we used an annotation tool, ‘makesense.ai’ to manually annotate all the images with the label ‘Auto‘ essentially creating a bounding box around the auto(s) if present within the image.

We then exported this data to the YOLO v5 format. The labels were stored in the YOLO v5 format in which they were stored in a separate text file and each line of the text file contained the following in order:

- Class Label: In our case, since there's only 1 class to detect (Auto), the class label is always zero
- Normalised x coordinate (by image width) of the centre of the bounding rectangle
- Normalised y coordinate (by image height) of the centre of the bounding rectangle
- Width of the rectangle as a fraction of the width of the image
- Height of the rectangle as a fraction of the height of the image

Model Parameter Tuning:

Batch Size:

We observed that higher batch size gave us better results. Upon further exploration, we found that small batch sizes produce poor batchnorm statistics, as given in [1]. After 50 epochs for each batch size, the following was observed:

Batch Size	Precision	Recall	mAP@.5	Train Box Loss	Test Box Loss
4	0.8047	0.7277	0.7924	0.04117	0.03827
16	0.958	0.828	0.931	0.02113	0.02489

Number of Epochs:

As we increase the number of epochs, the metrics get better up to a certain point after which overfitting occurs.

Results:

Inlier Images (Those with the most number of correct detections):

Even though this image is very crowded, we're still successfully detecting the Autos. Hence, an inlier

Even though this image is very crowded, we're still successfully detecting the Autos. Hence, an inlier

In an image full of noise we're still successfully detecting (not all) a few Autos. Hence, it can be considered to be an inlier.

In an image full of noise we're still successfully detecting (not all) a few Autos. Hence, it can be considered to be an inlier.

Outlier Images (False positives):

There is clearly a false positive in this image.

A taxi and bus are being detected as an Auto. There is clearly a false positive in this image.

A white car is being detected as an Auto. There is clearly a false positive in this image.

A grey car is being detected as an Auto. There is clearly a false positive in this image.

References:

- Tips for Best Training Results - YOLOv5 Documentation: https://docs.ultralytics.com/tutorials/training-tips-best-results/