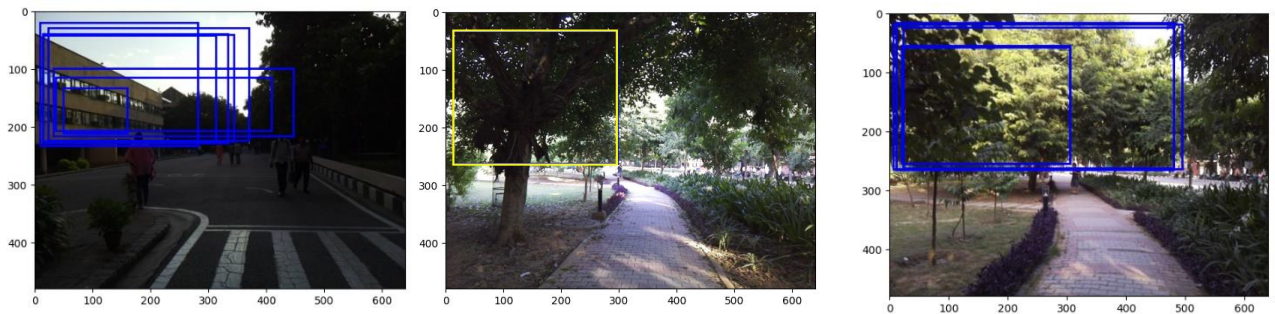# CV Assignment

DINO on Pedestrian Data

By Geetika Vadali, 24th September 2024

The dataset is loaded and then converted to a pandas data frame for easier analysis. By merging the information about the image and its constituent bounding boxes, I visualised the initial bounding boxes on respective images. Like so –


Then this data frame is sampled randomly, to facilitate the splitting. And then 160 images make the training set. And 40 images make the validation set. New json files are created for both sets' annotations. Format of theses instances are same as the dataset link provided, with respective data points only.

The evaluation is done using the "checkpoint0011_4scale.pth" on our pedestrian data. The final values for Average Precision are so –

```
Average Precision  (AP) @[ IoU=0.50:0.95 | area=   all | maxDets=100 ] = 0.486
Average Precision  (AP) @[ IoU=0.50      | area=   all | maxDets=100 ] = 0.823
Average Precision  (AP) @[ IoU=0.75      | area=   all | maxDets=100 ] = 0.509
Average Precision  (AP) @[ IoU=0.50:0.95 | area= small | maxDets=100 ] = 0.372
Average Precision  (AP) @[ IoU=0.50:0.95 | area=medium | maxDets=100 ] = 0.599
Average Precision  (AP) @[ IoU=0.50:0.95 | area= large | maxDets=100 ] = 0.697
Average Recall     (AR) @[ IoU=0.50:0.95 | area=   all | maxDets=  1 ] = 0.108
Average Recall     (AR) @[ IoU=0.50:0.95 | area=   all | maxDets= 10 ] = 0.506
Average Recall     (AR) @[ IoU=0.50:0.95 | area=   all | maxDets=100 ] = 0.602
Average Recall     (AR) @[ IoU=0.50:0.95 | area= small | maxDets=100 ] = 0.545
Average Recall     (AR) @[ IoU=0.50:0.95 | area=medium | maxDets=100 ] = 0.663
Average Recall     (AR) @[ IoU=0.50:0.95 | area= large | maxDets=100 ] = 0.750
```

Given the IoU metric as bbox.

To visualise the predictions then –
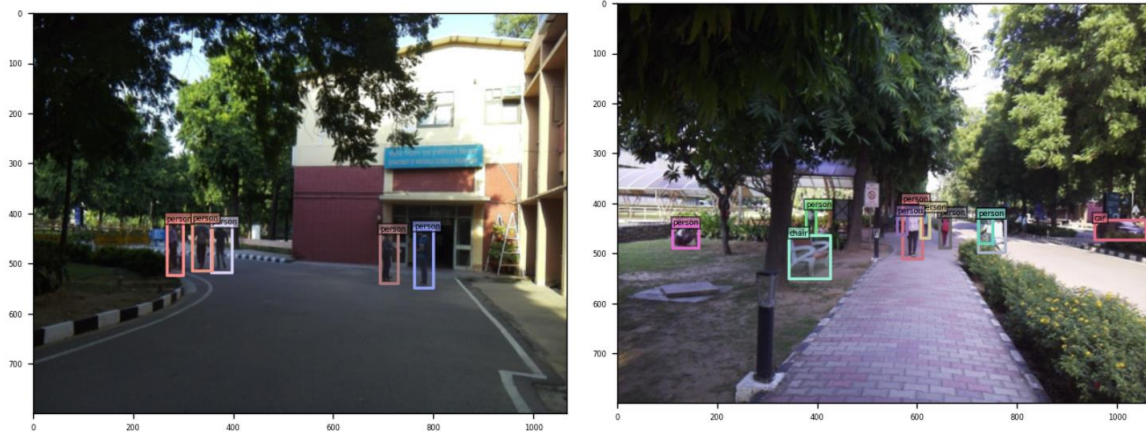
Finetuning the model has currently posed a problem. In accordance with the repository authors, I ran this script with respective paths –

!python main.py

      --config_file /kaggle/input/config-dino-ped/config-dino/DINO_4scale_pedestrian.py

      --options num_classes=1 dn_labebook_size=2

      --coco_path /kaggle/input/cocodir-pedestrian/COCODIR

      --pretrain_model_path /kaggle/input/dino0011/pytorch/default/1/checkpoint0011_4scale.pth

      --finetune_ignore label_enc.weight class_embed --output_dir /kaggle/working/finetunedmodel.pth

Which ends up on an eventual cuda error that is possibly because of index mismatch.

Weights after running the finetuning script – https://drive.google.com/file/d/1xxcAogJoZjrOogfCoYOYd7t4CRoHPSBu/view?usp=sharing