**Problem Statement:** Real time object detection using YOLO v4 and DeepSort Mobile GPU.

**Description:**

Object detection being a part of computer vision has been a major domain in development of AI models. From the domain of Computer vision this model takes the concepts of Object detection, Instance Segmentation and object tracking in the given video stream for identifying the instances of objects listed under COCO dataset. This model processes the video to provide areas of interest with objects which can be used for further decision-making process in AI models.

**Proposed Work:**

The model is designed to detect the instances of objects and provide a bounding box along with the class label as they appear in the frames. The model has the following phases:

1. Object detection with YOLO v4:

For purpose of Detecting the instances of objects, YOLOv4 (You Look Only Once) based on Darknet DNN. With the help of Transfer learning , a pre-trained model weights with around 64,429,405 parameters were taken in for building the model. This pre-trained YOLOv4 model can detect around 80 classes like person, car, apple, traffic signs etc. This YOLOv4 model can take 480\*480 input size and returns bounding box for detected instances of the class of the object.

1. Object Tracking with DeepSORT:

After the bounding box is returned from yolov4 model, the detection results are then passed to DeepSORT tracker. DeepSORT uses COSINE distance to identify the instance throughout the frames. DeepSORT assigns the bounding box along with a tracking ID for objects which can be seen in the output video stream.

**DEMO:**

* **Parameters:**

1. Object Detection is done through YOLOv4 for the input size 480\*480 is based on DarkNet weights for around 64,429,405 parameters.
2. Object tracking by DeepSORT is based on maximum Cosine distance=0.4, IOU threshold= 0.4.
3. At first the Darknet weights for YOLOv4 model is saved to TensorFlow using save\_model.py with the following syntax:

python save\_model.py --model yolov4

Graphical user interface, text

Description automatically generated

A screenshot of a computer

Description automatically generated with medium confidence

* **Running the Algorithm:**

DeepSort dependencies and yoloV4 dependencies are loaded and object detection is done by running the object\_tracker.py with the following syntax:

python object\_tracker.py --video ./data/video/test.mp4 --output ./outputs/demo.avi --model yolov4

arguments:

--video: for input video

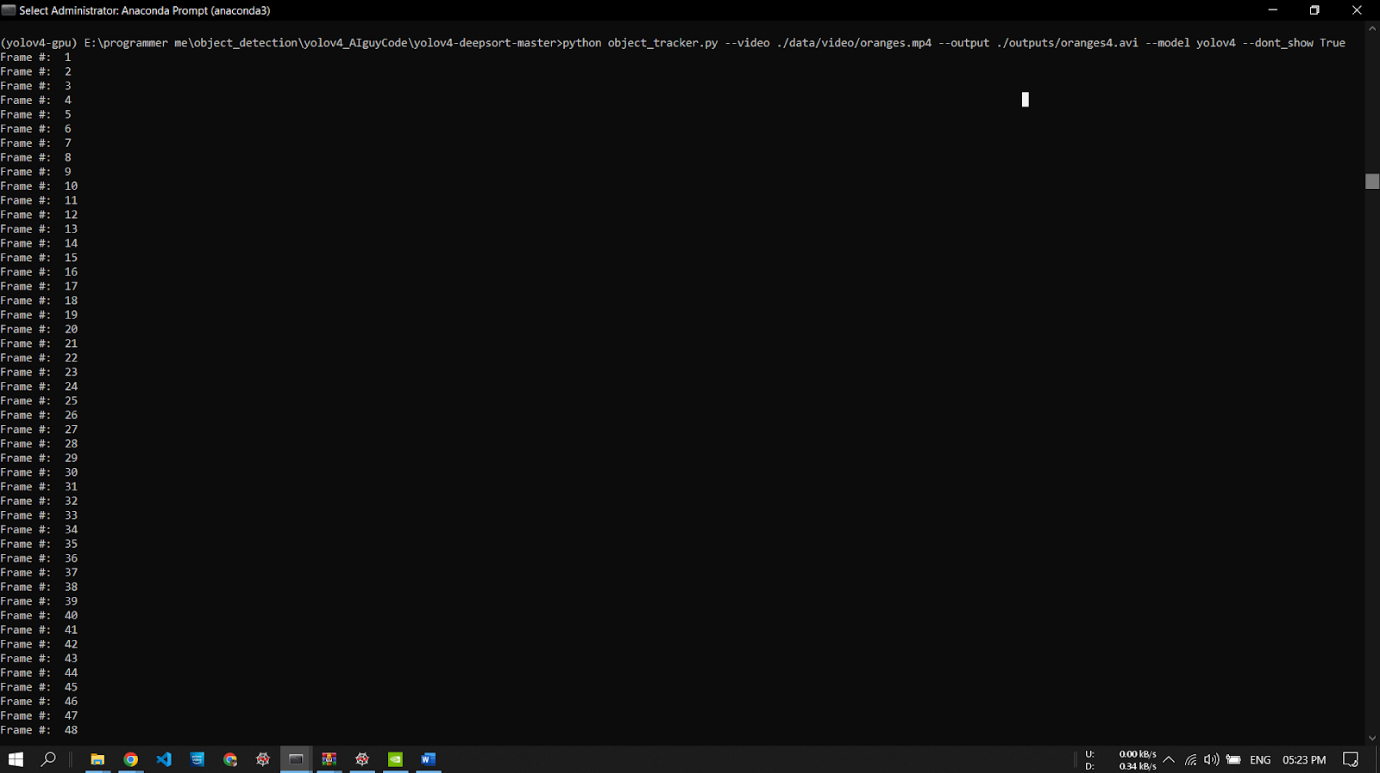
--output: for output video containing the detected objects and final count

* **Performance Results:**

All the tasks have been carried with GPU processing on NVIDIA RTX 1650 Ti mobile GPU with an average 18 FPS as processing speed. Requirements and libraries for GPU processing are installed accordingly.

A screenshot of a computer

Description automatically generated with medium confidence

Text

Description automatically generated