Ahmedabad University

CSE 541 - Computer Vision

Weekly Report 11_02_2023

Team: - Pixel Pioneers

Group Member Details

AU2040170	Kathan Bhavsar
AU2040117	Maulik Ranadive
AU2040183	Nand Patel
AU2040185	Arsh Mansuri

- Prof Mehul Raval

Tasks Performed in the week

We researched and analyzed our project this week, we explored the principles that we are going to use for our project and what datasets and technologies we will be using during the course of this project. We defined the problem more clearly, mentioned the list of tech stacks and we created the entire flowchart on how our group would be moving forwards with this project and implementation of the DEEPSORT algorithm.

Outcomes of the Tasks Performed

Here are the outcomes of the tasks that we performed,

Problem Definition & Statement

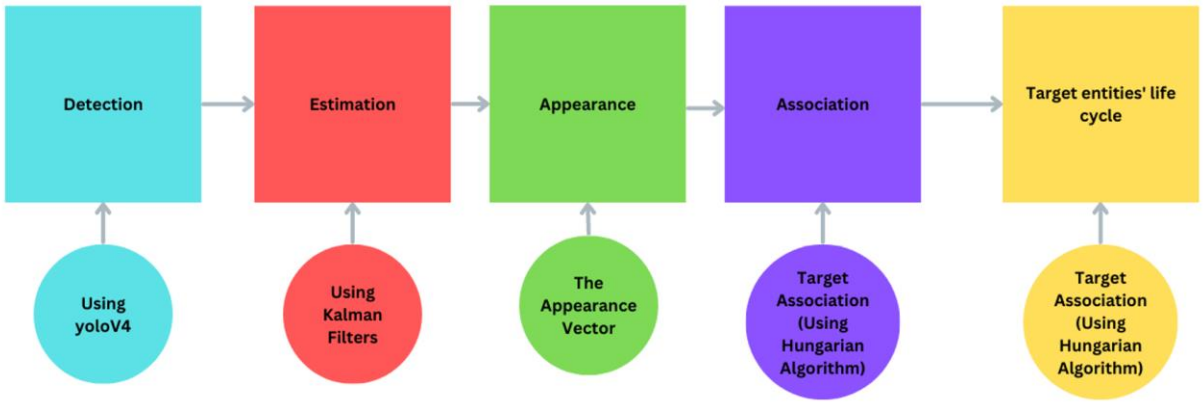
Object detection and tracking in the presence of occluded entities such as Clouds, Fog, Rain, and Walls. (Using DEEPSORT) for the BDD100k dataset.

It is hard to detect objects when there are occluded substances. DEEPSORT is a real-time multiple object tracking algorithm that combines a deep convolutional neural network (CNN) for object detection and appearance-based tracking using a Kalman filter and Hungarian algorithm. The CNN is trained to detect objects in an image and the Kalman filter and Hungarian algorithm are used to track the objects and things based on their appearance. DEEPSORT has been shown to achieve high accuracy and robustness in tracking multiple objects in challenging scenarios such as occlusions and camera motion.

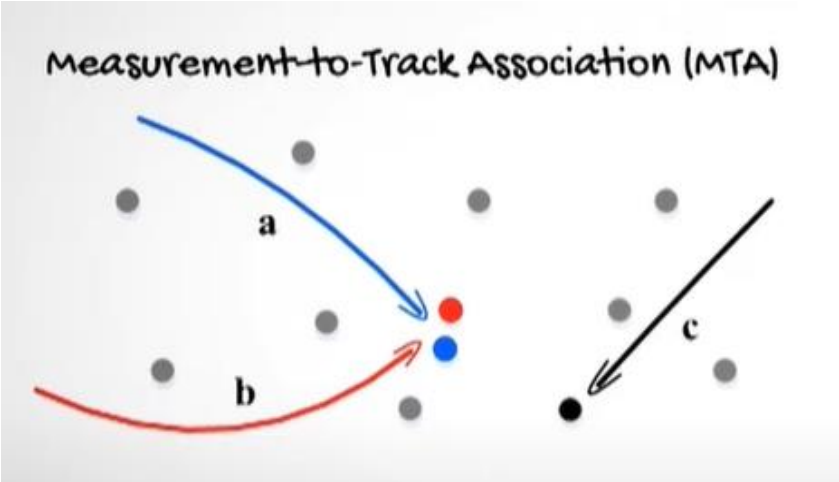
Dataset

https://www.kaggle.com/datasets/solesensei/solesensei_bdd100k

Flowchart



The flowchart provides a clear and precise understanding of how deep sort works. Firstly, the key principle is object tracking, which will be done by YoloV4. We use YoloV4 because we want to track the object in a single shot while considering that the speed of the sequential images and videos is important. Yolov4 provides very accurate results on everyday objects. The traditional tools like mean shift and optical flow are computationally complex and prone to noise, so we use the Kalman Filter, which plays a vital role while detecting objects while they are occluded. Assuming a constant velocity model and Gaussian distribution, the filter recursively helps to estimate the model by its motion and sensor confidence values. Deepsort in turn provides a very accurate distance metric which is “Appearance Vectors”.



Appearance Vectors provides us with MTA (Measurement to Track Association) which defines a relationship between a measurement and an existing track. While we now have the Target Associations, we compute the matrix of the object that is detected by every frame using Iou distance and the distance of occluded objects will be computed through the Hungarian Algorithm. The objects will have unique identifiers which will be created and destroyed accordingly. Tracks will be terminated if they are not detected by Tlost frames and the life cycle of the object will end respectively. This is how the working principle of Deep Sort works.

Tech Stack *(Technologies and Principles we will be using)*

1) **YoloV4 (You only look once version 4)** - YOLOv4 uses a single forward pass over the neural network to evaluate a whole image and generate predictions, which makes it a single shot object detection system. This allows it for real-time object detection jobs, where speed is crucial. The program can accurately identify a wide range of items, including people, vehicles, and other everyday objects.

2) **Kalman filtering** - It is a mathematical method used to estimate the state of the system based on measurements over time. It is used in object tracing to improve the accuracy of the estimated position of the objects, especially in the presence of occluded entities such as Clouds, Fog, Rain, and walls. It also helps to reduce the error in the object position estimation over time by incorporating the measurements and model of the system's dynamics.

3) **Hungarian algorithm:-** It is an optimization algorithm that is used to solve the assignment problem. In object tracking, the Hungarian algorithm is used to associate the detected objects with their corresponding tracks. The algorithm assigns the detected objects to the tracks in such a way that the overall cost, defined as the sum of the distances between the detected objects and their corresponding tracks, is minimized.

4) **CNNs: -** Convolutional Neural Networks are a type of deep learning algorithm that is widely used for image classification and object detection tasks. In object tracking, CNNs are used to detect objects in the input frames and estimate their position. The CNNs are trained on large datasets of annotated images, which enables them to learn to detect objects with high accuracy. The use of CNNs in object detection and tracking allows for robust performance in the presence of occlusions, such as clouds, fog, rain, and walls, as they are able to learn the underlying features of the objects and identify them even in challenging conditions.

Answers to the comments by TAs

1) ***When you say tracking it refers to real-time video tracking. Will the dataset which you will be using consist of images or videos? Kindly highlight that and also mention the dataset which you will be using to achieve these results.***
    - Will be using the BDD100k dataset, which consists of 100,000 high-resolution images captured from various driving scenarios in different weather conditions and times of day, along with annotations for various objects such as vehicles, pedestrians, traffic signs, and lane markings.

2) ***Are you going to deploy and implement a deep sort algorithm from scratch? and secondly, how are you going to deal with the problem of localization in such a problem?***
    - Yes, we will be implementing the deep sort algorithm from scratch. As for the problem of localization, Deep SORT makes use of bounding box predictions from a detection network. The detection network that has been trained to identify items in a picture returns bounding boxes and class labels for each object that it has identified. The tracking algorithm uses these bounding boxes to estimate where items will be in the following frame after receiving them from the user.

Tasks to be performed in the upcoming week.

- Implementation of DEEPSORT from scratch.
- Training yoloV4 for BDD100k dataset to obtain custom weights for object tracking and recognition.

References

1. Augmented Startups. (2020, August 31). *How DeepSORT Works?* [Video]. YouTube. https://www.youtube.com/watch?v=LbyqsoLJu5Q
2. T. (2023, February 2). *TRAIN A CUSTOM YOLOv4 OBJECT DETECTOR (Using Google Colab)*. Medium. https://medium.com/analytics-vidhya/train-a-custom-yolov4-object-detector-using-google-colab-61a659d4868
3. Maiya, S. R. (2020, April 24). *DeepSORT: Deep Learning to track custom objects in a video*. Nanonets AI & Machine Learning Blog. https://nanonets.com/blog/object-tracking-deepsort/
4. https://www.researchgate.net/publication/344099630_Object_Tracking_Using_Improved_Deep_Sort_YOLOv3_Architecture
5. S. (2022, November 11). *Understanding Multiple Object Tracking using DeepSORT*. LearnOpenCV – Learn OpenCV, PyTorch, Keras, Tensorflow With Examples and Tutorials. https://learnopencv.com/understanding-multiple-object-tracking-using-deepsort/