**Object detection and tracking in the**

**presence of occluded entities using DEEPSORT**

CSE541: Computer Vision

Maulik Ranadive Kathan Bhavsar Nand Patel Arsh Mansuri

AU2040117 AU2040170 AU2040183 AU2040185

*Abstract*

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 things based on their appearance. DEEPSORT has achieved high accuracy and robustness in tracking multiple objects in challenging scenarios such as occlusions and camera motion.

Keywords

Object Tracking, Object Detection, Obscureness, Occlusion, Trajectories, Kalman Filter, Yolov4, Hungarian Algorithm, Appearance Vectors, Mean Shift, Optical Flow. Introduction

# Introduction

One of the most important tasks in computer vision is object recognition and tracking, which has a variety of uses in robotics, autonomous vehicles, and surveillance. Unfortunately, it is a difficult challenge that has not yet been fully resolved to effectively and robustly detect and track objects in tough settings like snow, fog, rain, and wall occlusions. Certain kinds of occlusions can conceal objects, making it hard to detect and track them, which can result in mistakes and poor system performance.

In order to increase the precision and dependability of computer vision systems in difficult conditions, this project aims to develop an effective object recognition and tracking system that can manage occlusions. This has the potential to improve a variety of applications, from robots and autonomous cars to surveillance and security, where precise and consistent object detection and tracking are essential.

# Literature Survey

*SORT - Simple Online Real-Time Tracking*

The paper starts by describing the limitations of traditional object tracking algorithms, such as the inability to handle occlusions and re-appearances of objects. It then introduces DeepSORT, which addresses these limitations by using a deep association metric to associate object detections over time.

*YOLOv4: Optimal Speed and Accuracy of Object Detection*

## YOLOv4 is simply the extension of the previous versions of YOLO which performs object detection .One of the key contributions of YOLOv4 is the use of a hybrid backbone network that combines features from different networks to achieve high accuracy and speed. The authors also introduce several new data augmentation techniques and a novel loss function that improve training efficiency and generalization.

*Object Tracking: A Survey" by Hamed Kiani Galoogahi*

This survey provides a comprehensive review of object tracking methods, including both traditional and deep learning-based approaches, as well as evaluation metrics and datasets. It also discusses challenges such as occlusions and motion blur.

*A Survey on Object Detection in Video" by Luciano Spinello*

This survey focuses specifically on object detection methods designed for video sequences, covering both traditional and deep learning-based approaches. It also discusses evaluation metrics, datasets, and challenges such as motion blur and illumination changes.

# Implementation

* 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.
* ResNet is being used as our CNN for this implementation
* 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 called “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.

# Results

## We use an open image dataset for the labeled image dataset which is contained in the csv format, so we convert the label csv format to yolov4 format.

## Download Darknet from open source and combine with our yolov4 model and also use pre weight yolov4 model to help our model to train very well

## Train the yolov4 model using the custom label dataset and get

## For the best result, we stop the training when the average loss is less than 0.05 if possible or at least constantly below 0.3, else train the model until the average loss does not show any significant change for a while and, The map parameter gives us the Mean Average Precision. The higher the MAP the better it is for object detection.

## We can check MAP for all the weights saved every 1000 iterations for eg:- yolov4-custom\_3000.weights, yolov4-custom\_5000.weights, yolov4-custom\_6000.weights, and so on. This way we can find out which weights file gives you the best result. The higher the MAP the better it is.

## After all that we test our yolov4 detection model with few images and get the sufficient amount of accuracy.

# Conclusion

Overall, the progress of our project is evident in the results

That we have implemented the YOLOv4 model successfully with the mentioned yolo rate in the results.

We have also started implementing the other important aspects of the project such as the appearance vector, Kalman vector, and the Hungarian Algorithm in code but since we did not have the labels dataset for training the entire model we did not have time to complete that implementation.

##### References

1. Techzizou (2023) *Train a custom Yolov4 object detector (using Google Colab)*, *Medium*. Analytics Vidhya. Available at: https://medium.com/analytics-vidhya/train-a-custom-yolov4-object-detector-using-google-colab-61a659d4868 (Accessed: March 11, 2023).
2. Maiya, S.R. (2020) *DeepSORT: Deep learning to track custom objects in a video*, *Nanonets AI & Machine Learning Blog*. Nanonets AI & Machine Learning Blog. Available at: https://nanonets.com/blog/object-tracking-deepsort/ (Accessed: March 11, 2023).
3. Maiya, S.R. (2020) *DeepSORT: Deep learning to track custom objects in a video*, *Nanonets AI & Machine Learning Blog*. Nanonets AI & Machine Learning Blog. Available at: https://nanonets.com/blog/object-tracking-deepsort/ (Accessed: March 11, 2023).
4. *Object tracking using improved Deep\_Sort\_YOLOv3 architecture - researchgate* (no date). Available at: https://www.researchgate.net/publication/344099630\_Object\_Tracking\_Using\_Improved\_Deep\_Sort\_YOLOv3\_Architecture (Accessed: March 11, 2023).
5. Techzizou (2023) *Train a custom Yolov4 object detector (using Google Colab)*, *Medium*. Analytics Vidhya. Available at: https://medium.com/analytics-vidhya/train-a-custom-yolov4-object-detector-using-google-colab-61a659d4868 (Accessed: March 11, 2023).
6. theAIGuysCode (no date) TheAIGuysCode/OIDv4\_ToolKit: Download and visualize single or multiple classes from the huge open images V4 dataset, GitHub. Available at: https://github.com/theAIGuysCode/OIDv4\_ToolKit (Accessed: March 11, 2023).
7. <https://towardsdatascience.com/yolov4-in-google-colab-train-your-custom-dataset-traffic-signs-with-ease-3243ca91c81d>