

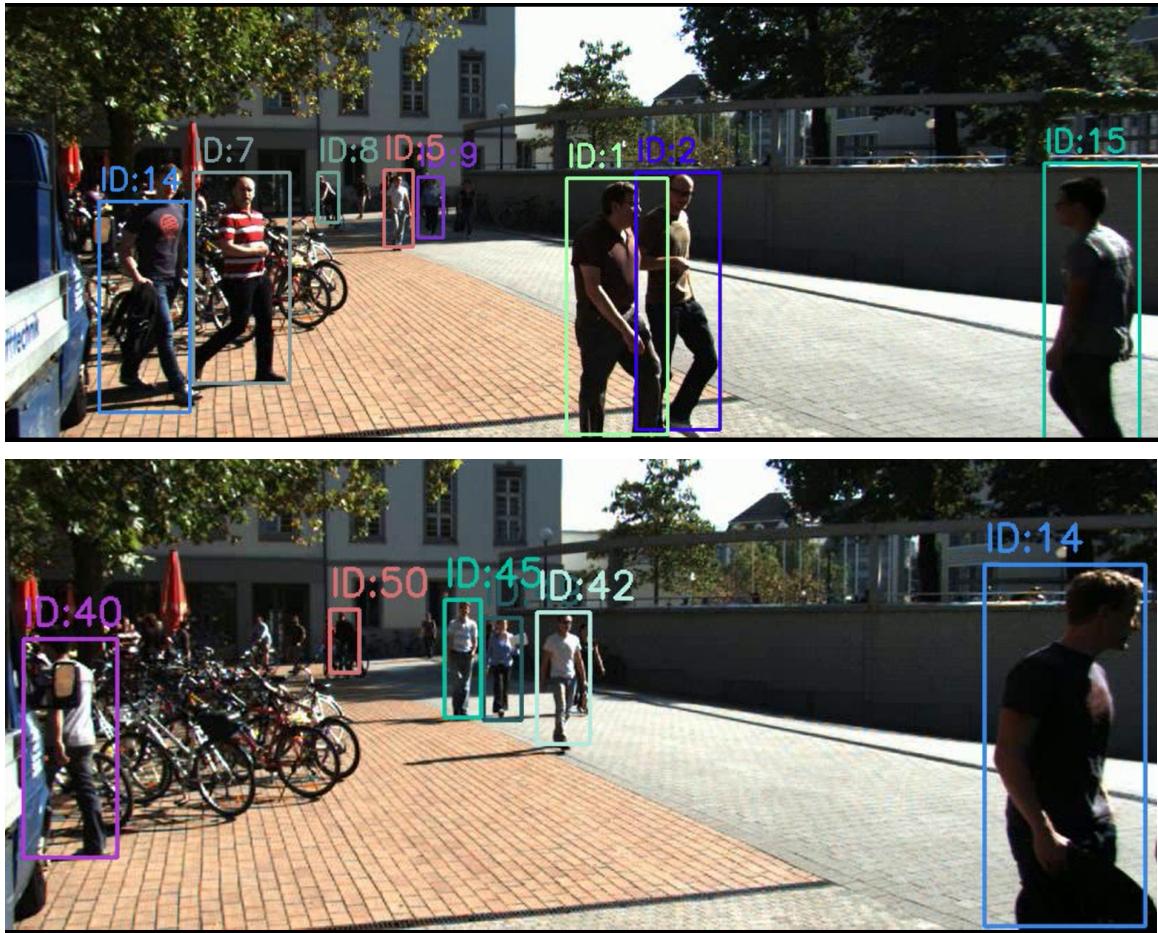
Computer Vision

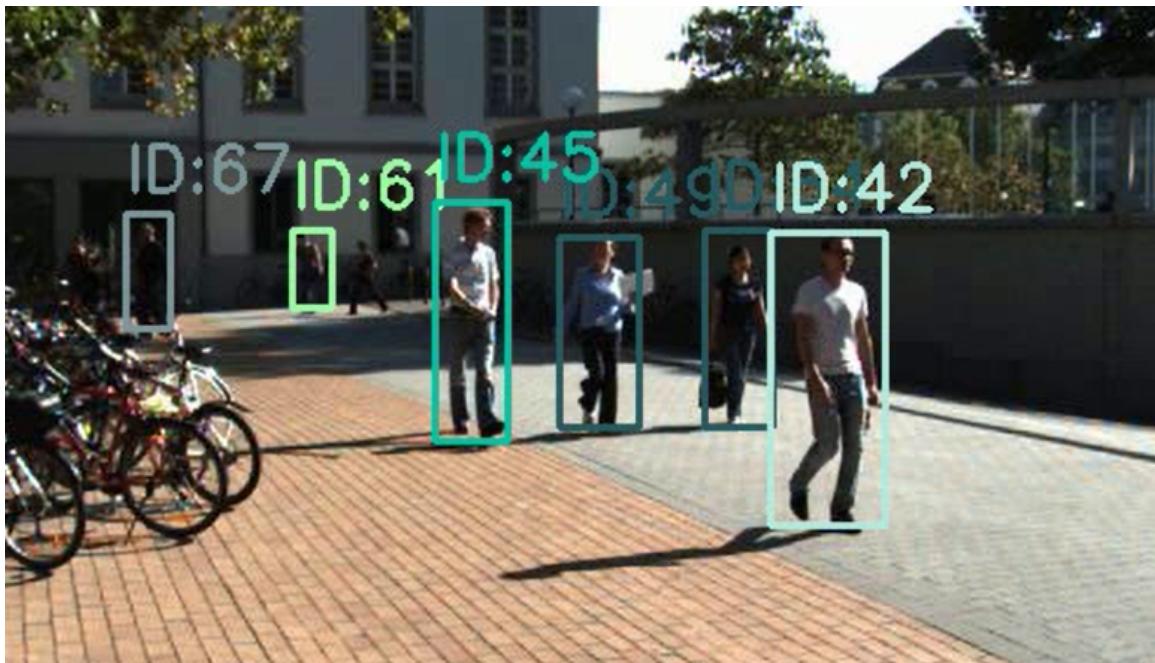
Assignment 4 : Multiple Object Tracking with Detection

Arjun Temura
2020497

Q1 .

Video file obtained after using yolov5 model and SORT on the given images. (result.avi)





output.txt file: (for first 10 frames)

```
1, 3, 561,143,17,44,1,-1,-1,-1  
1, 2, 505,140,46,194,1,-1,-1,-1  
1, 1, 392,150,104,206,0,-1,-1,-1  
2, 2, 510,139,58,198,1,-1,-1,-1  
2, 1, 413,154,87,196,1,-1,-1,-1  
3, 2, 514,135,68,209,0,-1,-1,-1  
3, 1, 442,161,67,177,1,-1,-1,-1  
4, 2, 523,132,77,218,1,-1,-1,-1  
4, 1, 456,163,64,182,0,-1,-1,-1  
5, 2, 530,135,81,220,1,-1,-1,-1  
5, 1, 464,153,74,201,1,-1,-1,-1  
6, 2, 535,139,83,218,1,-1,-1,-1  
6, 1, 471,153,76,206,1,-1,-1,-1  
7, 2, 554,141,73,209,1,-1,-1,-1  
7, 1, 483,156,77,207,1,-1,-1,-1  
8, 2, 570,141,68,206,1,-1,-1,-1  
8, 1, 503,160,70,201,1,-1,-1,-1  
9, 5, 495,145,24,52,1,-1,-1,-1  
9, 2, 583,143,71,210,1,-1,-1,-1  
9, 1, 521,157,70,205,0,-1,-1,-1  
10, 5, 496,145,24,54,1,-1,-1,-1  
10, 2, 599,143,77,213,1,-1,-1,-1  
10, 1, 537,153,79,211,0,-1,-1,-1
```

Q2. Metric Analysis of output generated:

a) command used:

```
!python /content/TrackEval-master/scripts/run_mot_challenge.py --DO_PREPROC False  
--BENCHMARK arjun --SPLIT_TO_EVAL train -TRACKERS_TO_EVAL MyTracker -METRICS  
HOTA CLEAR Identity VACE --USE_PARALLEL False --NUM_PARALLEL_CORES 1
```

output:

All sequences for MyTracker finished in 0.11 seconds									
	HOTA	DetA	AssA	DetRe	DetPr	AssRe	AssPr	LocA	
HOTA: MyTracker-pedestrian									
arjun-01	51.486	48.211	55.087	62.141	60.373	58.056	79.824	80.39	
COMBINED	51.486	48.211	55.087	62.141	60.373	58.056	79.824	80.39	
CLEAR: MyTracker-pedestrian	MOTA	MOTP	MODA	CLR_Re	CLR_Pr	MTR	PTR	MLR	
arjun-01	49.195	77.38	49.927	76.428	74.253	44.444	55.556	0	
COMBINED	49.195	77.38	49.927	76.428	74.253	44.444	55.556	0	
Identity: MyTracker-pedestrian	IDF1	IDR	IDP	IDTP	IDFN	IDFP			
arjun-01	69.553	70.571	68.563	482	201	221			
COMBINED	69.553	70.571	68.563	482	201	221			
VACE: MyTracker-pedestrian	SFDA	ATA							
arjun-01	61.579	27.467							
COMBINED	61.579	27.467							
Count: MyTracker-pedestrian	Dets	GT_Dets	IDs	GT_IDs					
arjun-01	703	683	36	9					
COMBINED	703	683	36	9					

- b) We can clearly observe that the MOTA(Multiple Object Tracking Accuracy) associated with our model is 49.195 and MOTP(Multiple Object Tracking Precision) is 77.38.

MOTA is defined as the percentage of correctly identified objects (true positives) out of the total number of objects present in the ground truth dataset. It takes into account false positives, false negatives(miss), and mismatch errors.

In this case, the MOTA value of 49.195% indicates that the algorithm correctly identified 49.195% of the total number of objects in the dataset. Higher the MOTA score, better is the tracking.

Based on the MOTA results, we can conclude that our model is a moderate predictor of pedestrian tracking because of moderate accuracy. There still is a scope of improvement in our model. There are many entries that are inconsistent with the ground truth of the model.

MOTP gives a measure of the average distance between the localization of objects in the ground truth and the detection output.

The MOTP score ranges from 0-100 (in %). If the MOTP value is 100, then the precision of the system is poor. And if it is close to zero, then the precision of the system is good.

A MOTP value of 77.38 implies that the average distance between bounding boxes in the ground truth and the output is 77.38 pixels. Our model within this range of MOTP values has a moderately low precision for detection of the bounding boxes. Our model is not able to detect and track pedestrians with good precision.

There could be many reasons for this moderately low accuracy and precision:

1. **Occlusion:** SORT is prone to errors in tracking as a result of occlusions in the frames. Due to this some pedestrians might not have been detected hence low MOTA values.
2. **Distance from camera:** Some pedestrians that were present very far away from the center of projection of the camera might not have been detected, hence low MOTA values. There could also have been distortion effects that could have affected the results
3. **Ineffective Localization:** It is possible that the bounding boxes associated with the pedestrians were not able to effectively isolate pedestrians and led to high MOTP values .