Assignment 4

Aman Tiwari(160094),Ankush Tiwari(160130)

April 2019

1 Problem Statement

We tried to solve the object tracking problem for the given dataset.We tried to track objects frame by frame using object detection inputs.Our approach was to build an object tracking algorithm that runs on top of a object detection algorithm output(such as YOLO , FrRCNN etc.) to track objects persistently across multiple frames.

2 Methods Used

We used the principle of tracking by detection to solve the problem. For detection part we made use of YOLOv3 pretrained on COCO dataset.We used YOLO for detection since we wanted real time object tracking. For tracking we tried a number of algorithms to solve the problem:

1. Centroid Tracking Algorithm
   This algorithm works by computing centroids of the bounding boxes returned by the detection algorithm. The algorithm computes euclidean distance between the new bounding boxes and existing objects and matches the nearest neighbour bounding box for a particular object being tracked. The bounding boxes which are not matched are assigned new object ids. We deregister an object if it is not matched for some N consecutive frames. We also store a "velocity" paprameter for each tracked object, we implemented it as the distance covered by the object in x and y directions per frame. The idea is to be able to predict the next position of the tracked object if it missing in some previous consecutive frames, so it moves to its "theorical" position and will be more likely to be re-matched on the next frame.

2. SORT : Simple Online and Realtime Tracking
   This algorithm works by maintaining the motion information for each bounding box. The state of each bounding box is modelled by the values of its center,height,aspect ratio and their respective velocities. It uses Kalman filters to compute velocities. It uses an assignment cost matrix that is computed as the intersection-over-union (IOU) distance between

each detection and all predicted bounding boxes from the existing targets. Then the best match is found using the Hungarian algorithm.

3. Simple Online and Realtime Tracking with a Deep Association Metric
   This algorithm is an improvement over the SORT algorithm.It uses appearance information in addition to the motion information.So for each bounding box an appearance descriptor is calculated which comes from a convolutional neural network. Then the motion and appearance information is combined and the best match is found using the Hungarian algorithm.

4. IOU tracker
   This method is based on the assumption that the detector produces a detection per frame for every object to be tracked, i.e. there are none or few "gaps" in detections. This method essentially continues a track by associating a detection with the highest IOU to the last detection in the previous frame if a certain threshold is met.All detections not assigned to an existing track will start a new one. All tracks without an assigned detection will end.

3 Challenges Faced

1. The basic centroid tracking algorithm that we used did not perform so well. Our implementation of the velocity vector was not that good and there were a lot of identity switches.

2. IOU performed better than the basic centroid algorithm. But this algorithm is based on the assumption that there are very few gaps in detections. This was not fulfilled since in case of YOLO detections were lost at various instants. Also this algorithm is not good for slow frame rates since the overlapping areas can be very low between consecutive frames.

3. SORT algorithm performed relatively well but in this case also there were a number of identity switches. Another problem we faced was related to detection, YOLO was not good at detecting small objects, hence these were not tracked by the algorithm or their track was lost in between.

4. Deep SORT solved the problem of identity switches to some extent. However the detection problem caused by YOLO was still present.

4 References

- Simple Online and Realtime Tracking(Paper)
- Simple Online And Realtime Tracking With A Deep Association Metric(Paper)

- High-Speed Tracking-by-Detection Without Using Image Information(Paper)
- Tracking Things in Object Detection Videos

5 Model Hashes

We have only used pretrained models.The model hashes are:

mars-small128.pb	0c3fc16ea01be53ea4358f97462b3b7f
yolov3.weights	c84e5b99d0e52cd466ae710cadf6d84c