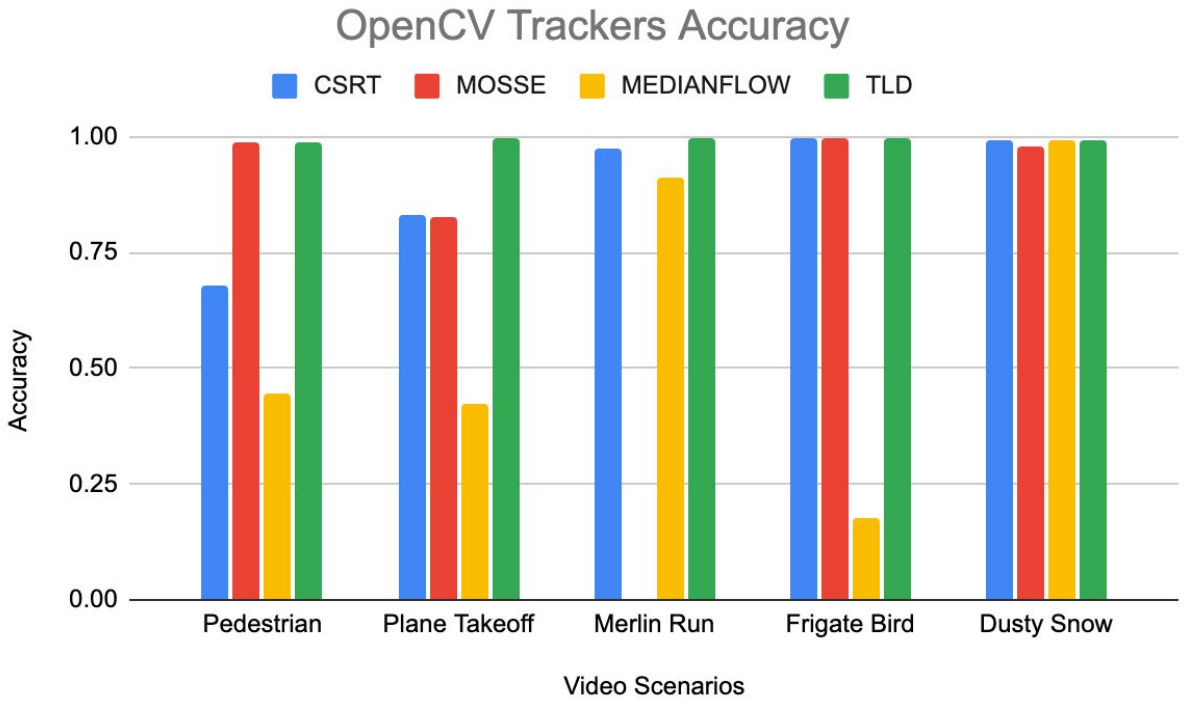
Assignment 1: Object Tracking
James Yost & Matthew Ernst
CS510

Overview:
This project takes a look at multiple different object trackers within the OpenCV library. Specifically, the trackers investigated were CSRT, MOSSE, MEDIANFLOW, and TLD. These trackers were used on multiple videos of different scenarios. Discussed below is their tracking accuracy on videos and details on why or why not they tracked the way they did.



CSRT
From the graph above we can see the results of the CSRT tracker. The CSRT did the best in the Frigate Bird, Dusty Snow, and Merlin Run videos. However, it did not do well in the Pedestrian video. This is not surprising, as the CSRT tracker is trained with a correlation filter with compressed features (HoG and Colornames). Because of this type of technique, the tracker does not do well with occlusion. This can be seen in the pedestrian video as when a car passes in front of the person tracked, the tracker loses the person and is never able to find the person again. The graph states that it was roughly 70% accurate at detection but this is not entirely the case. In order to see the accuracy we used code to count the number of frames that are being “tracked”. However, the tracker sometimes thinks it is tracking the correct object but it is not. In this case, the tracker is able to track the pedestrian and loses all tracking after the car. This should bring the accuracy of this video down to roughly 50%. Contrastly, looking at the accuracy of the Frigate Bird, it does very well. This is because of the less occlusion. This video only has the bird been tracked and a sky background. The other interesting note about this type of tracker is that it can adapt to scale, deformation and rotation. This can be seen with the

Merlin Run video as the object being tracked, the dog, changes in scale very quickly. However, the tracker is able to adapt to this and correctly track.

MOSSE

Out of all of the trackers, this was our favorite one. Simply due to the quickness of its adaptation on the Pedestrian video. As the car passes in front of the pedestrian, the tracker loses its sight for a quick second and is able to recover extremely fast. The bounding box is also stable throughout the video with very subtle changes. From the results above, it shows that the worst tracking it did was on the Plane Takeoff video. However, this video is interesting. It is able to track the plane through the whole video, but the bounding box slightly moves and doesn't encompass the whole plane towards the end of the video. This contradicts the original statements by OpenCV as it is a great tracker for changing in scale and lighting. Seeing as this video has both of those elements and it struggles to fully find encompass the plane as it takes off is surprising.

MEDIANFLOW

For the above experiments, this tracker followed the descriptions of its capabilities the best. In the OpenCV writings, it is noted that it is best suitable for smooth and predictable movements that are visible throughout the whole sequence of a video. This is shown perfectly in the Aerobatic video. In this video, a plane is moving in random fast sharp directions while clouds pass in front of it. The tracker does excellent in the beginning. However, as the plane turns and falls behind a cloud, the bounding box begins to expand very quickly and reaches a point where it cannot track the plane anymore. The tracker was also not able to recover and find the plane as it comes back from behind the clouds. Comparatively, this tracker did excellent on the Dusty Snow video. As in the previous video, it also adapts its bounding box in this video as the dog comes closer to the frame. This was the best video with 99% accuracy of tracking.

TLD

For our experiments, this is the tracker that did the best by far. The most notable for this tracker was its ability to recover from occlusion. This can be seen in the Pedestrian video. In the video, as multiple cars pass in front of the pedestrian, the tracker is still able to find the person after the occlusion. However, the tracker does change the bounding box on the person. The person is still able to be tracked, but the bounding box changes many times as the video progresses. Again, the graph does not depict the fully conclusive results. This is shown in the Merlin Run video. The tracker fails immediately in the beginning as the dog begins running down the hill. However, once the dog becomes mostly in the frame. The dog can be tracked very easily. Overall, this tracker did very well.

Conclusion

As detailed above, each tracker has its own pros and cons when it comes to object tracking. Our conclusion is that the tracker wanted to be used is highly based on the environment it will be used. If the conditions are known pre-tracking, there is a high likelihood that a person will be able to successfully track whatever they like. Overall, the TLD tracker did the best from our analysis, however, the best in terms of speed and overall usability is the MOSSE tracker.