Long-Term Tracking

Ivan Nikolov

I. Introduction

For this assignment, a short-term deep CNN-based tracker SiamFC is evaluated on the `car9` sequence of the long-term dataset provided. The tracker is also upgraded to enable long-term tracking after the object is occluded or simply the tracking fails. We ran multiple experiments in order to find the optimal values for the threshold and number of samples. The final results are reported as precision, recall and F-score on the `car9` sequence only.

II. Experiments

A. SiamFC: short-term tracker

For evaluation of the SiamFC short-term tracker, we used the default parameters provided with the scripts. The metrics were identical to the ones provided in the instructions.

Precison	Recall	F-score
0.65	0.27	0.38

Table I: SiamFC metrics on the sequence `car9`.

B. Long term tracking

For the second part of the assignment, we modified the `update` method of the tracker and implemented a redetection algorithm when the short term tracker fails. A failure is detected if the correlation response is below a predefined threshold. When this criteria is met, multiple patches are sampled uniformly from the whole image. If the patch with the highest response is over the threshold, the algorithm goes back to short-term tracking mode.

On Figure 1 and Figure 2 we can see execution of the method on two frames, on the first one the short-term tracking fails and we activate redection. On the second, the object is found, and we go back to short-term tracking.

The long-term tracking depends on two predefined parameters: correlation threshold and number of samples used. On the `car9` sequence which we evaluated the performance on, we tried out multiple values. Based on the experiments, we found out if the threshold is set to high values, the algorithm detects failures even when the object is not occluded or lost in the frame. Too low values on the other side cause behavior similar to the short term tracker and the redetection is not activated on time. These values need to be determined empirically and depend on the use case. From the results in Table II, the best performance was achieved with a correlation threshold value of 5.

We also did an analysis on how different number of samples affect the performance of the tracker. Based on our observations, increasing the number of samples boosts the performance in terms of metrics, however it lowers the speed of execution because more processing is needed.

On Table III metrics for different experiments with different number of samples are shown. Differences between the performance in the experiemts with 10 samples and 6 samples, were only on the third decimal. This means that for our use-case it will be better to run the tracker with 6 number of samples

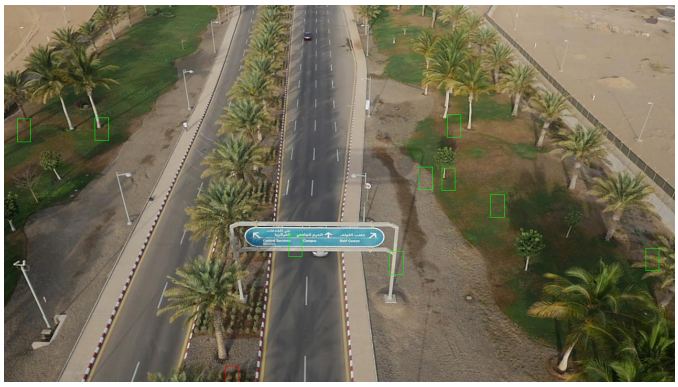

Figure 1: After occlusion, the algorithm does not detect the car. Multiple patches from the whole image are sampled.

Figure 2: The tracked object is found again. The red patch has a response over the threshold.

Threshold	Precision	Recall	F-score
1	0.59	0.48	0.53
4	0.60	0.59	0.59
5	0.60	0.59	0.60
6	0.61	0.56	0.58

Table II: Long-term tracker metrics on the sequence `car9` with different thresholds on correlation response.

becasue the execution is faster. For a different scenairios and use-cases where we would like to redetect the object with higher certanty, it would be more appropriate to run the tracker with higher number of samples set.

Threshold	Precision	Recall	F-score
1	0.60	0.56	0.58
6	0.60	0.59	0.60
10	0.60	0.59	0.60

Table III: Long-term tracker metrics on the sequence `car9` with different number of samples for redetection.

III. Conclusion

In this assignment, we evaluated SiamFC, a short term tracker on the `car9` sequence form the long-term tracking

dataset provided. We also upgraded the SiamFC tracker to a long-term tracker with adding a redetection functionality. We ran multiple experiments and found out that the tracker works best with a threshold of 5, and 6 sampled patches during redetection.

Further improvements can be made in parallelizing the sampling process and correlation response computation and adding Gaussian sampling.