SPARK: Spatial-aware Online Incremental Attack Against Visual Tracking
Supplementary Material

Qing Guo[1,2⋆], Xiaofei Xie[2⋆], Felix Juefei-Xu[3], Lei Ma[4], Zhongguo Li[1], Wanli Xue[5], Wei Feng[1⋆⋆], and Yang Liu[2]

[1] College of Intelligence and Computing, Tianjin University, China
[2] Nanyang Technological University, Singapore
[3] Alibaba Group, USA
[4] Kyushu University, Japan
[5] Tianjin University of Technology, China
tsingqguo@gmail.com

1 Attacking Correlation Filter-based Trackers

Correlation filter (CF) is a dominant tracking framework that can achieves well balance between tracking speed and accuracy. However, most of the CF-based trackers are not end-to-end architectures and use hand-craft features. Hence, it is difficult to attack them via the white-box setup and is meaningful to explore if SPARK could attack CF-based trackers by using deep tracking frameworks, *e.g.*, SiamRPN-based trackers. As shown in Table I, the adversarial examples from SiamRPN-Alex can reduce all tested CF-based trackers having different features, which demonstrates that the transiferability of our attack across different trackers and features exists. In terms of different features, the HOG feature is easier attacked when compared with the gray feature, hybird feature (*i.e.*, HOG+CN), and deep feature (*e.g.*, VGG).

Table I. Untargeted attack (UA) for correlation filter-based trackers, *e.g.*, MOSSE [1], KCF [4], BACF [3], STRCF [5], and ECO [2] with the perturbations generated from SiamRPN-AlexNet.

	MOSSE	KCF	BACF	STRCF	ECO
Features	Gray	HOG	HOG+CN	HOG+CN	VGG
Org. Prec. (%)	41.7	69.2	70.5	72.3	89.6
Proc. Drop (%)	0.2	3.3	2.1	1.5	0.9

2 Speed Analysis

We have reported the time cost of our SPARK in Table. 1 in the submission and shown that SPARK is more suitable for attacking online trackers than

⋆ Qing Guo and Xiaofei Xie contributed equally to this work.
⋆⋆ Wei Feng is the corresponding author (wfeng@tju.edu.cn).

three basic attack methods due to the balance between time cost and attack Succ. Rate. Please find details in Section 3.3. Compared with trackers' cost shown in Table. III, the time cost of our attack method increases as the tracking model becomes larger under the white-box attack. In particular, when attacking SiamRPN-Alex, SPARK achieves near real-time attacking. Although the attack speed decreases with more complex models, the corresponding tracking speed is also slower and lets the influence of decreased attacking be smaller. We can

Table II. Time cost of attacks w.r.t. different trackers on OTB100 dataset.

SiamRPN	AlexNet	MobileNetV2	Res50
Track cost per frame (ms)	9.3	37.6	42.1
Attack cost per frame (ms)	41.4	126.9	156.3
Track speed (fps)	108.4	15.3	16.8
Attack speed (fps)	24.3	8.0	6.4

reduce the high time cost of attacking larger models (*e.g.*, MobileNetv2 and Res50) by using the light one (*e.g.*, AlexNet) due to the existence of the transferability between models as discussed in Section 4.3 and Table 4. Specifically, we attack three trackers, *i.e.*, SiamRPN-Alex/Mob./Res50, via SPARK with the adversarial perturbations generated from SiamRPN-Alex. Then, we calculate the attack's online speed as well as the three trackers' speed. As shown in the following Table. III, the speed of SPARK base on SiamRPN-Alex can reach near real-time speed (around 25 fps) for different trackers, which means our method is suitable for attacking real-time online trackers.

Table III. Time cost of attacking trackers on OTB100. The adversarial perturbations are generated from SiamRPN-Alex.

SiamRPN	AlexNet	MobileNetV2	Res50
Track speed (fps)	108.4	15.3	16.8
Attack speed (fps)	24.3	23.1	22.7

References

1. Bolme, D.S., Beveridge, J.R., Draper, B.A., Lui, Y.M.: Visual object tracking using adaptive correlation filters. In: CVPR (2010)
2. Danelljan, M., Bhat, G., Khan, F.S., Felsberg, M.: Eco: Efficient convolution operators for tracking. In: CVPR (2017)
3. Galoogahi, H.K., Fagg, A., Lucey, S.: Learning background-aware correlation filters for visual tracking. In: ICCV. pp. 1144–1152 (2017)
4. Henriques, J.F., Caseiro, R., Martins, P., Batista, J.: High-speed tracking with kernelized correlation filters. IEEE TPAMI **37**(3), 583–596 (2015)
5. Li, B., Wu, W., Zhu, Z., Yan, J., Hu, X.: High performance visual tracking with siamese region proposal network. In: CVPR (2018)