NEURAL INFORMATION PROCESSING SYSTEMS

Reproducibility Study on Adversarial Attacks Against Robust Transformer Trackers



Fatemeh Nourilenjan Nokabadi[1,2,3], Jean-François Lalonde[1,2], Christian Gagné[1,2,3,4]

[1]IID, [2]Université Laval, [3]Mila, [4]Canada CIFAR AI Chair

Adversarial Robustness

Adversarial perturbations deceive neural networks, leading to inaccurate outputs. Such adversarial attacks have been studied for various vision tasks ranging from:

Object Classification

Object Segmentation

Object Tracking

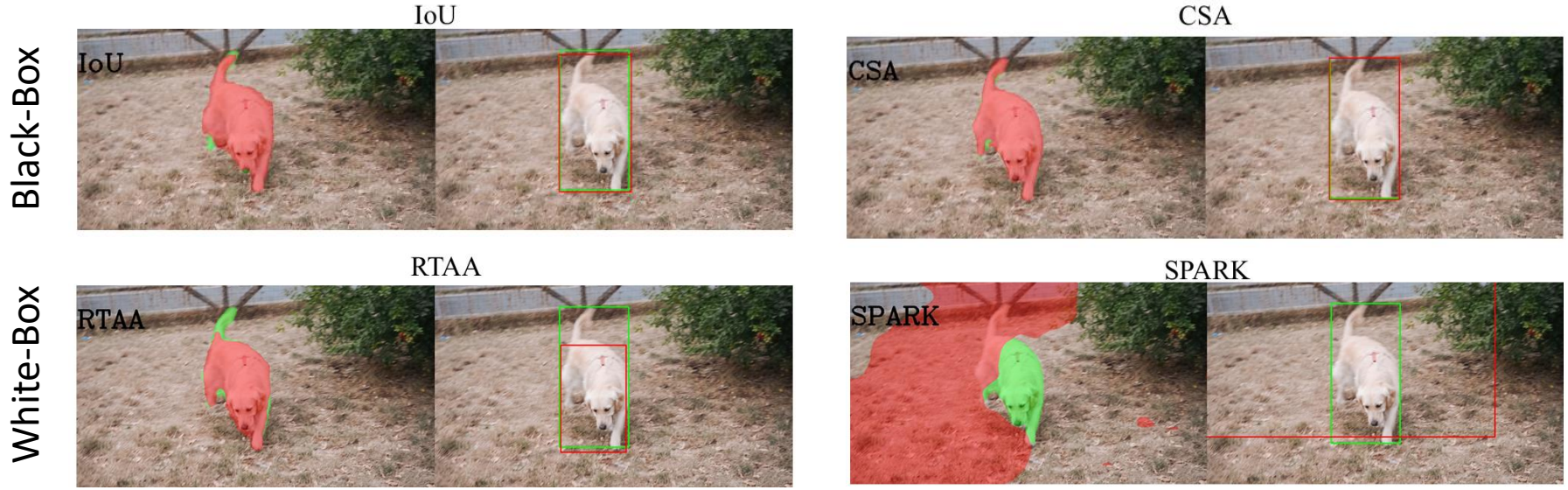
Transformers have recently boosted object tracking performance, but their robustness under adversarial attacks remains unclear.

Research Goals & Contributions

- **Goal 1.** Assess effectiveness of attacks on different trackers.
- **Goal 2.** Compare transformer and non-transformer trackers.
- **Goal 3.** Evaluate robustness across different perturbation levels and output types (bounding box vs. binary mask).

A. Adversarial Attacks per Tracker Output

- TransT-SEG Tracker on VOT2022

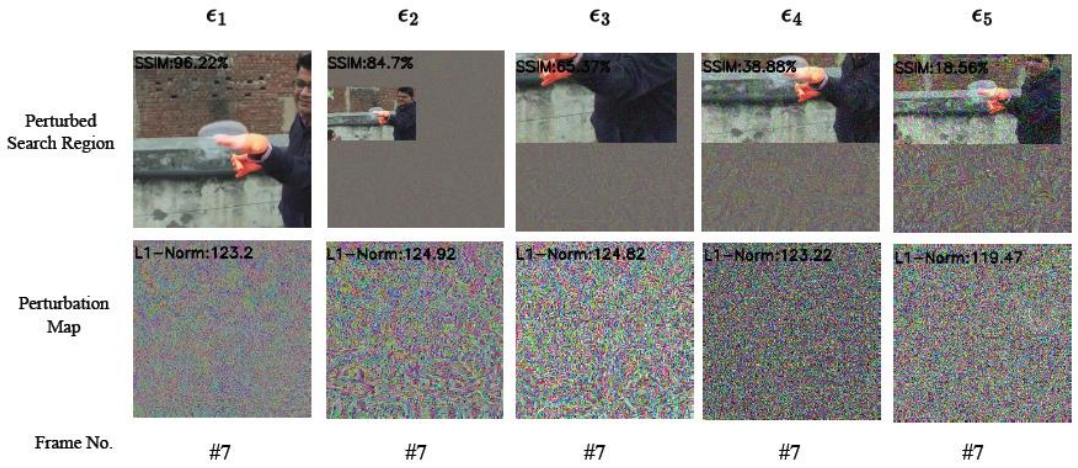


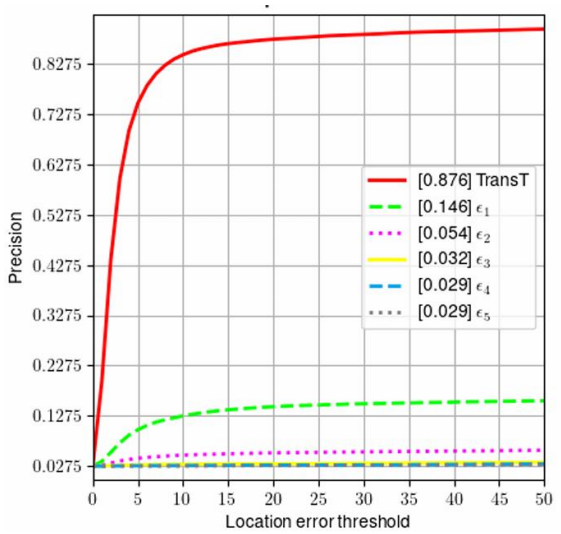
Main Takeaway: The attacks applicable to transformer trackers have more impact on the accuracy of the object mask than the bounding boxes on VOT2022ST dataset.

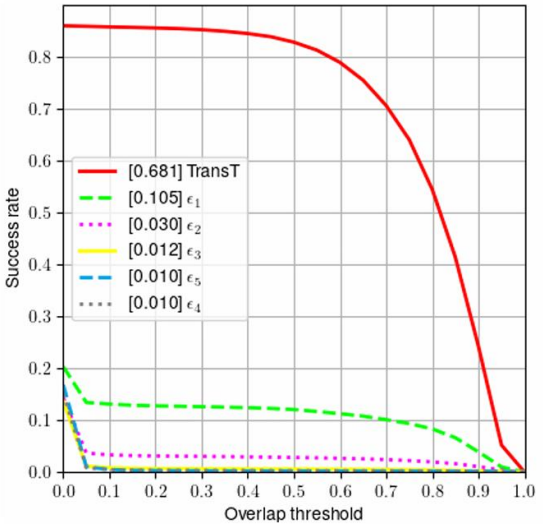
B. Adversarial Attacks per Perturbation Level

- White-box attacks under perturbation level shift

RTAA Performance against TransT Tracker on UAV123





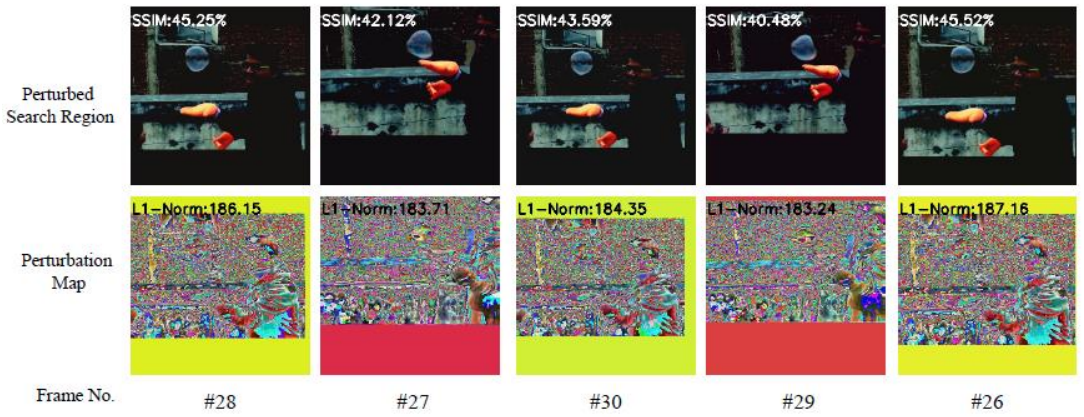


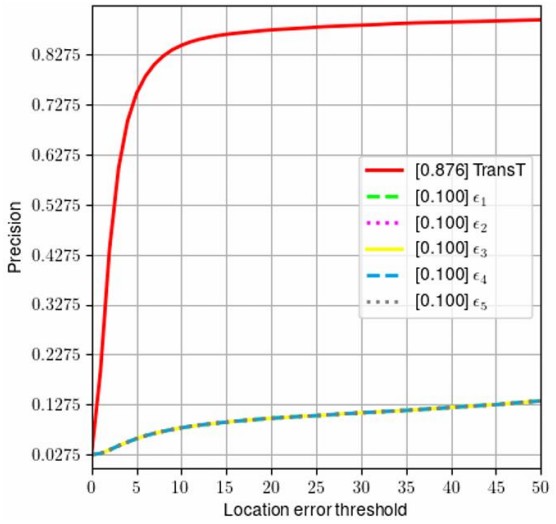
Main Takeaway: For RTAA attack, adding a higher perturbation level generates more perceptible noise for all frames, which damage more the overall tracking performance.

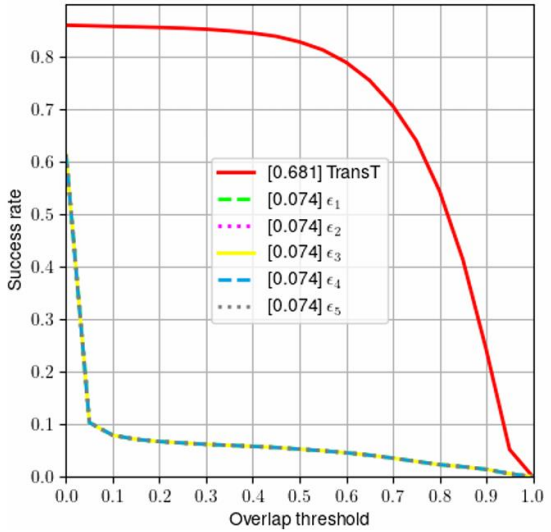
B. Adversarial Attacks per Perturbation Level

- White-box attacks under perturbation level shift

SPARK Performance against TransT Tracker on UAV123





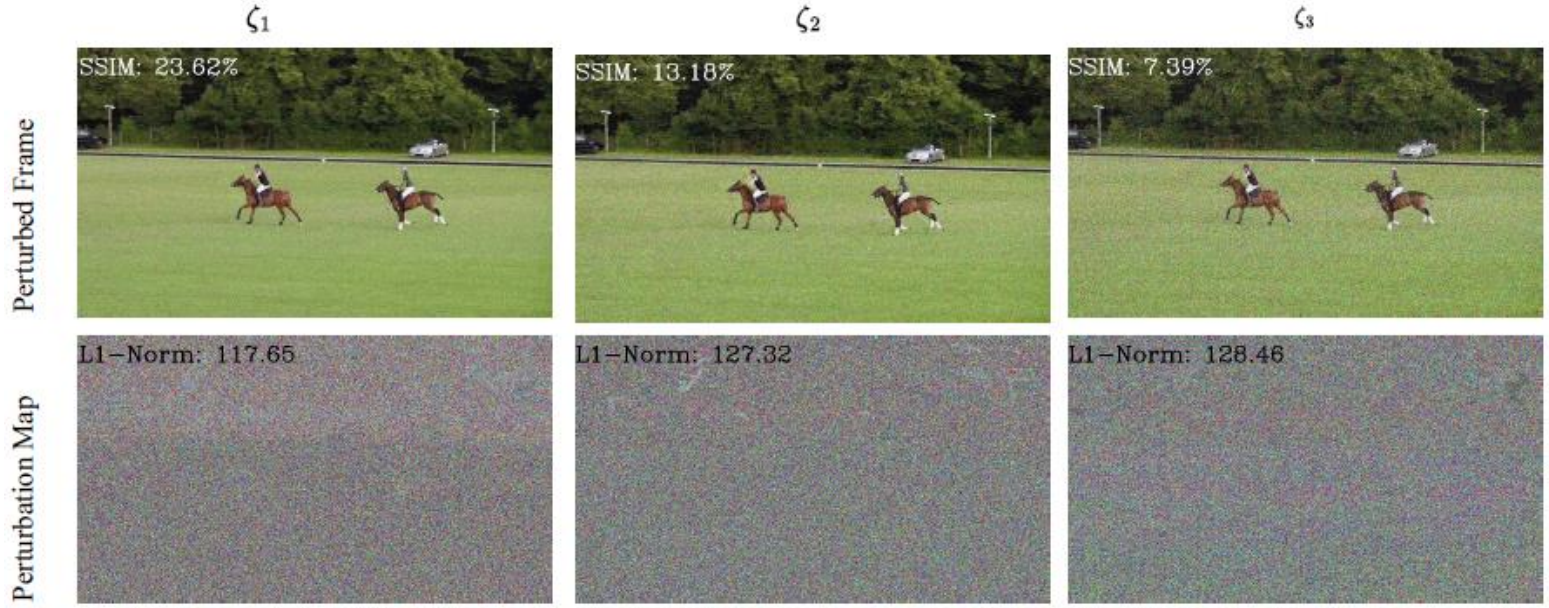


Main Takeaway: Increasing the perturbation level on SPARK attack setting results in more super-perturbed regions, i.e. regions with perceptible noise.

C. Adversarial Attack per Upper-Bound

- Black-box attack under various upper bounds

IoU Performance against ROMTrack Tracker on UAV123



Main Takeaway: The outcome of the IoU attack is sensitive to its initialization. The evaluation process may take a long time due to unsuitable initialization point.

D. Transformer versus Non-transformer Trackers

- ROMTrack and MixFormer show higher robustness.
- SPARK and RTAA substantially degrade tracking performance; transformer models are generally more resistant than non-transformers.
- Despite transformer trackers (ROMTrack, TransT, and MixFormer) showcasing the **top-3 performance**, their evaluation scores more notably decreased after applying the IoU method.

Conclusion and Future Directions

- Conclusion: Transformer trackers exhibit superior robustness but require further exploration of targeted adversarial techniques.

- Future Work: Development of more sophisticated attacks to effectively challenge these models and enhance robustness testing.