STEVENS
INSTITUTE OF TECHNOLOGY
1870

Ph.D. DISSERTATION DEFENSE

Candidate:	Haotian Gu
Degree:	Doctor of Philosophy
School/Department:	Charles V. Schaefer, Jr. School of Engineering and Science / Mechanical Engineering
Date:	Tuesday, February 24th, 2026
Time/Location:	10:00 am / Gateway North 103
Title:	Robust Vision-Based Object Tracking System: Perturbation Generation, Adversarial Defense and Resilient Tracking
Chairperson:	Dr. Hamid Jafarnejad Sani, Department of Mechanical Engineering, School of Engineering & Sciences
Committee Members:	Dr. Yi Guo, Department of Electrical and Computer Engineering, School of Engineering & Sciences Dr. Yong Shi, Department of Mechanical Engineering, School of Engineering & Sciences Dr. Damiano Zanotto, Department of Mechanical Engineering, School of Engineering & Sciences

ABSTRACT

Vision-based object detectors can enhance the safety of autonomous navigation systems by perceiving the surrounding environment, but they could be susceptible to adversarial attacks generated by artificial intelligence (AI). In this work, we developed novel detection and defense mechanisms against adversarial perturbations in vision-based perception for autonomous robotic systems. Particularly, we proposed approaches to defend against visible and invisible white-box image perturbations in robot navigation scenarios. For invisible image perturbations, we developed a tracking- and navigation-optimized neural network-based image restoration technique, namely TANGO-ESRGAN, to defend against such attacks in dynamic and real-time settings. In a photo-realistic simulation environment, we tested the proposed defense scheme and compared it with other approaches. Our method shows several desirable properties for real-time implementation in autonomous systems such as self-driving cars and aerial drones, including faster runtime, lower computational load, and adaptability to rectify a wide range of perturbation intensities.

For visible perturbations, we developed Ad_YOLO+, an object detection model robust to adversarial patches of varying sizes and locations in a vision-based tracking task. By treating adversarial patches as a distinct category during training, Ad_YOLO+ effectively identifies both adversarial patches and target objects simultaneously. In the AirSim virtual simulation environment, we implemented a verification procedure to assess whether Ad_YOLO+ provides provable robustness against adaptive adversarial patches in real-time under a defined threat model. Moreover, by integrating the patch-agnostic defense-based frontend with an additional broken-pixel restoration backend, we developed Segment and Recover (SAR) to detect adversarial image patches and restore object detection accuracy. We revealed adversarial patches in the high-frequency domain and proposed a recompression-based patch-localization frontend that is agnostic to patch appearance, shape, and location. Our evaluation results demonstrate the performance of our method across adversarial patches of varying sizes, tasks, and attack models, particularly in terms of object detection accuracy.