Adversarial Perturbations on Sensor Data for Compromising Robotic Path Planning in Simulated Environments

Remaining Challenges and Future Research Directions

Alejandro Almeida[1,2], Daniel Aviles Rueda[2]

[1]Analytics for Cyber Defense Lab
[2]Department of Electrical and Computer Engineering
Florida International University, Miami, FL, USA

{aalme034, davil050}@fiu.edu

Code: https://github.com/aalme034/adversarial-lidar-path-planning

December 2025

Despite the strong results presented in our work — where a simple MLP path planner in a 20×20 grid achieves over 96% clean success, drops to 65% under realistic PGD-$\epsilon = 0.30$ LiDAR attacks, and recovers to ∼94% with lightweight defenses — several critical limitations remain before these findings can be confidently transfer to real-world autonomous robots.

Key Remaining Limitations

- **2D grid-world abstraction** — No elevation, continuous dynamics, non-holonomic constraints, or actuator latency.
- **Extremely sparse and idealized sensor** — Only 8 perfect 2D rays vs. real 64–128-beam 3D LiDARs with complex point-cloud structure.
- **No sensor noise or environmental effects** — Missing multipath reflections, range noise, reflectivity variation, rain, fog, or dust.
- **Only white-box ℓ_∞ PGD attacks tested** — No black-box, transfer-based, query-limited, adaptive, or physical spoofing attacks evaluated.
- **Limited training data diversity and scale** — Only 25k–50k transitions from similar 20×20/30×30 maps with modest obstacle density.
- **Computational overhead of defenses** — LSTM adds ∼25% training time and ∼8 ms CPU inference latency (acceptable at 10 Hz but tight for higher control rates).

Prioritized Future Research Directions

1. **Port to 3D high-fidelity simulators** (CARLA, NVIDIA Isaac Sim, Gazebo) with realistic LiDAR physics, weather, and full vehicle dynamics.
2. **Evaluate real physical adversarial attacks** using LiDAR spoofers (lasers, retro-reflectors) on physical robots such as TurtleBot4, Clearpath Jackal/Husky, or F1TENTH with Velodyne/Ouster sensors.
3. **Test black-box and adaptive attacks** including transfer-based, score-based, and decision-based methods under realistic query budgets.
4. **Combine adversarial training with certified defenses** (randomized smoothing, IBP, CROWN) to obtain provable robustness guarantees.
5. **Deploy on real hardware** in lab and outdoor settings, measuring success rate, collision rate, energy use, and recovery behavior under live attacks.
6. **Explore recurrent RL/imitation hybrids** (SAC-LSTM, DreamerV3, TD-MPC2, IRIS) for better long-horizon robustness and sample efficiency in continuous spaces.
7. **Investigate multi-modal defenses** using sensor fusion (LiDAR + camera + IMU) and runtime anomaly detection to identify spoofed streams.

Closing Note

Our lightweight, fully open-source NumPy simulator has successfully served as an accessible educational benchmark and rapid prototyping tool that clearly exposes how fragile learned robotic planners are to small sensor perturbations — and how surprisingly effective simple defenses can be.

However, the central challenge for the robotics security community now is to **close the simulation-to-reality gap** through physical attacks, 3D environments, diverse scenarios, and provable defenses. Only then can we achieve truly trustworthy autonomy in the presence of adversarial sensor threats.

Alejandro Almeida and Daniel Aviles Rueda
Analytics for Cyber Defense Lab
Florida International University
December 2025