USENIX

VehicleSec '26

Seeing is Deceiving

Systematic Vulnerability Analysis of LiDAR-Based Autonomous Driving to Mirror-Induced Perception Failures

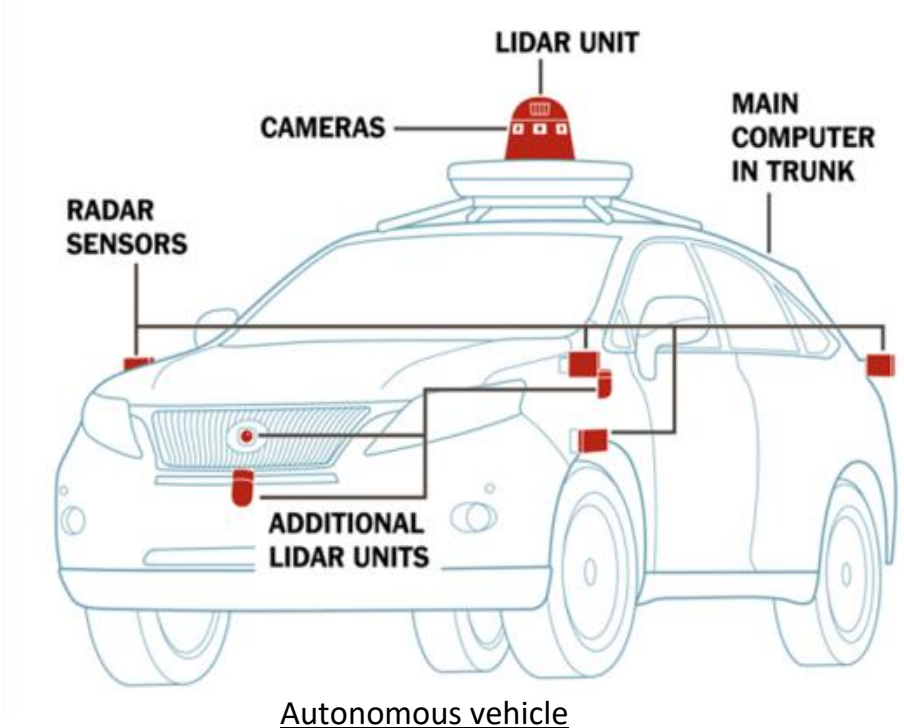
Selma Yahia[1], Ildi Alla[1], **Girija B. Mohan**[2], Daniel Rau[2], Mridula Singh[3], Valeria Loscri[4]
SnT, Univ. of Luxembourg[1] | htw saar[2] | CISPA[3] | Inria Lille - Nord Europe[4]

Contents

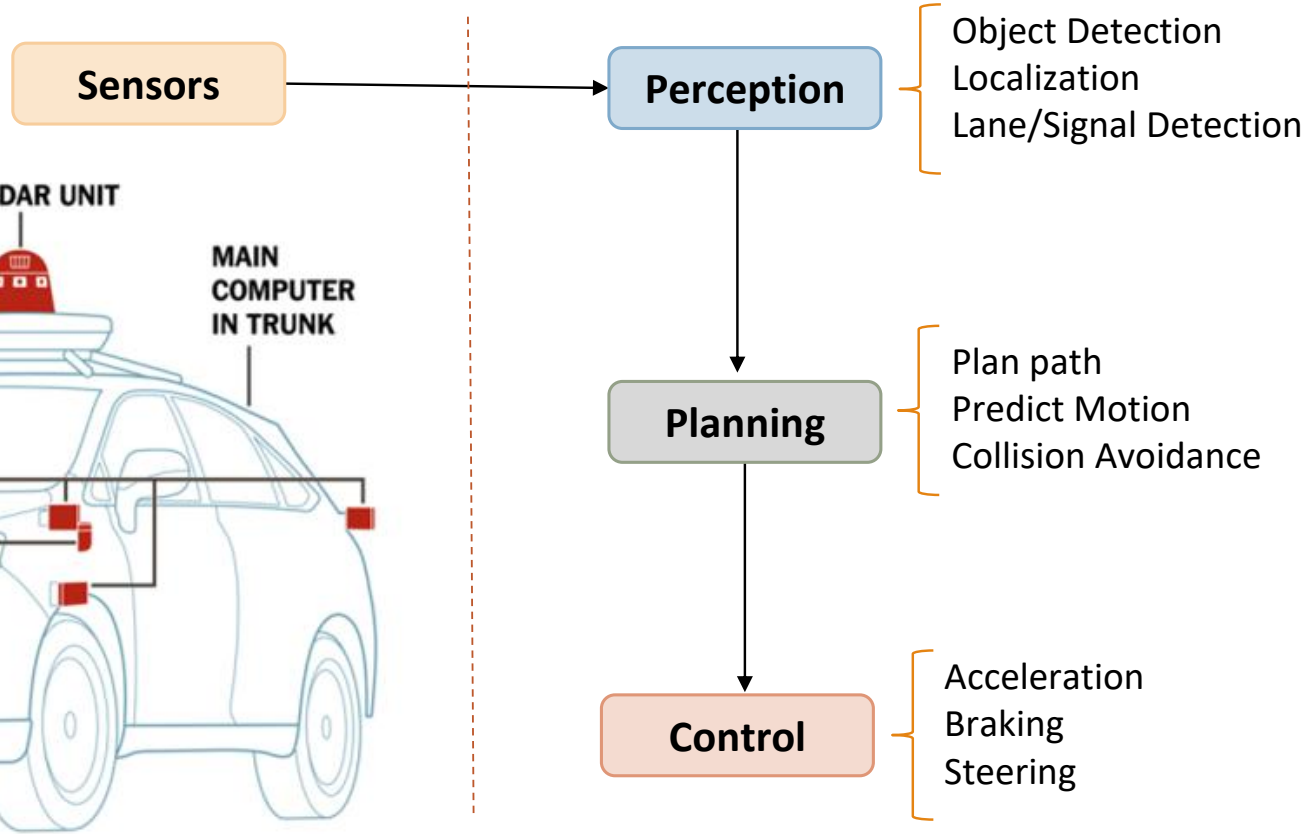
- Introduction
- Related Work
- Motivation
- Real-World (RW) Vulnerability Discovery
- Object Removal Attack (ORA)
- Object Addition Attack (OAA)
- Vulnerability Assessment on Autoware
- Empirical Modeling
- CARLA Simulation
- Defense Strategies
- Limitations and Future Work
- Discussion

Introduction (1)

Autonomous vehicle (AV) perception



Autonomous vehicle

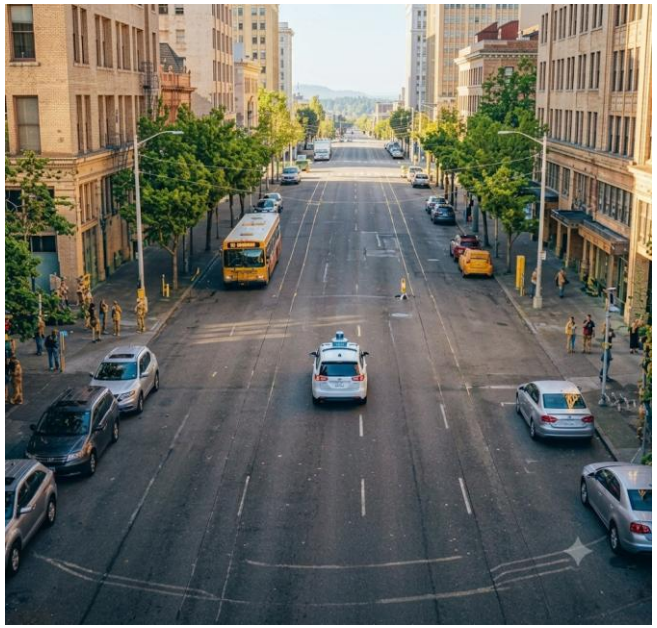


Introduction
Related work
Motivation
RW Threat
ORA
OAA
Autoware
Modelling
Simulation
Defence
Limitations
Discussion

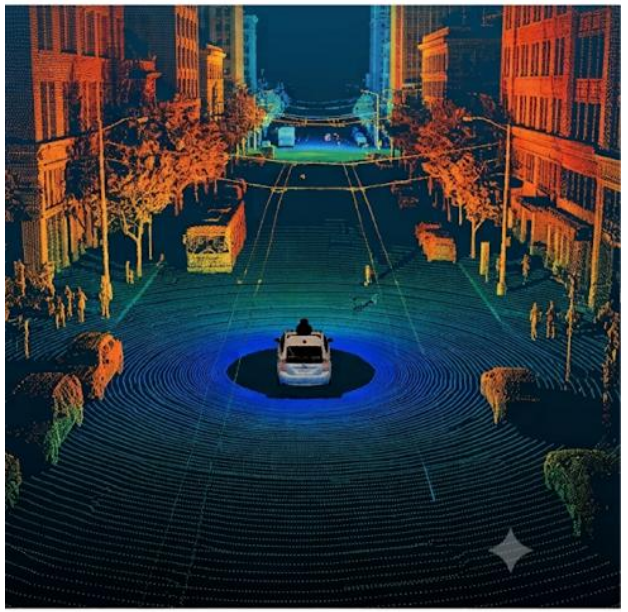
Introduction (2)

LiDAR - Point Cloud Visualization

Introduction
Related work
Motivation
RW Threat
ORA
OAA
Autoware
Modelling
Simulation
Defence
Limitations
Discussion



Camera Image



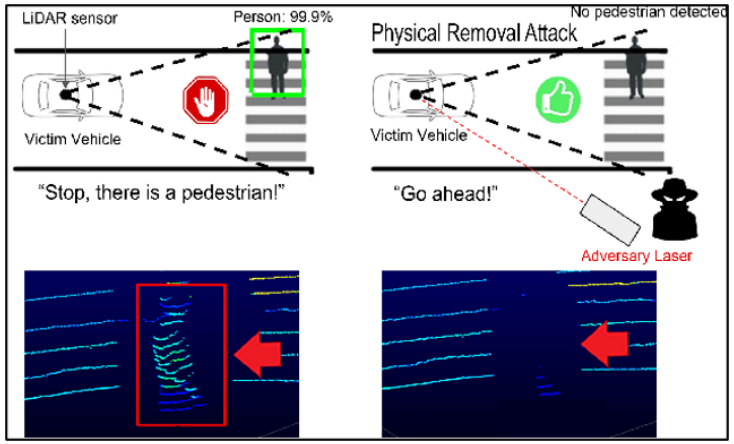
3D Point Cloud - LiDAR

AI Generated Images

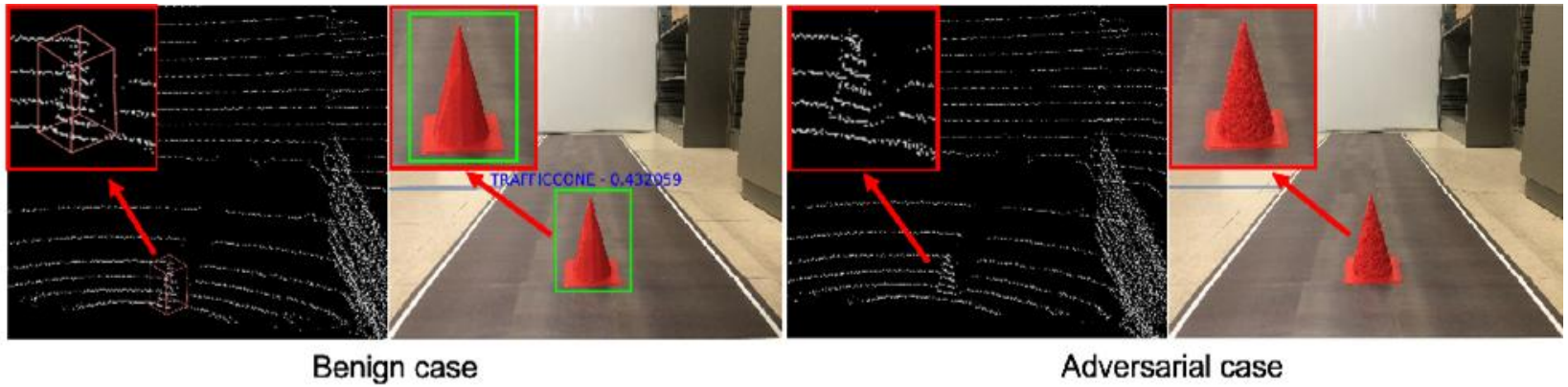
Related Work (1)

Security of LiDAR-based perception

Introduction
Related work
Motivation
RW Threat
ORA
OAA
Autoware
Modelling
Simulation
Defence
Limitations
Discussion



Active attack example [1]



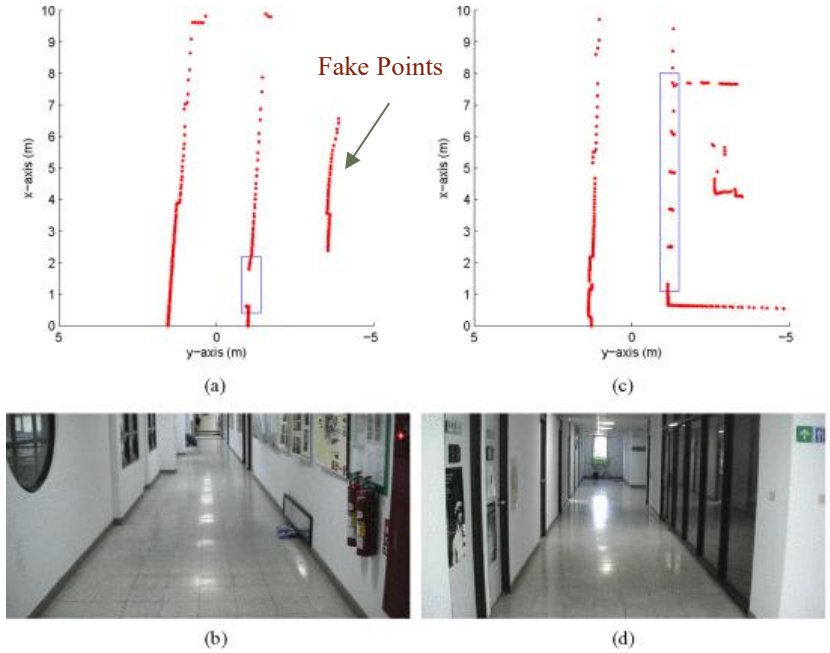
Passive attack example [2]

[1] Yulong Cao, S Hrushikesh Bhupathiraju, Pirouz Naghavi, Takeshi Sugawara, Z Morley Mao, and Sara Rampazzi. You can't see me: Physical removal attacks on {lidar-based} autonomous vehicles driving frameworks. In 32nd USENIX Security Symposium (USENIX Security 23), pages 2993–3010, 2023.
[2] Yulong Cao, Ningfei Wang, Chaowei Xiao, Dawei Yang, Jin Fang, Ruigang Yang, Qi Alfred Chen, Mingyan Liu, and Bo Li. Invisible for both camera and lidar: Security of multi-sensor fusion based perception in autonomous driving under physicalworld attacks. In 2021 IEEE symposium on security and privacy (SP), pages 176194. IEEE, 2021.

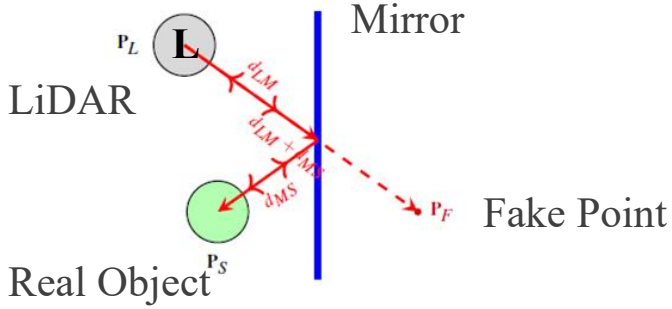
Related Work (2)

LiDAR Perception of Mirror

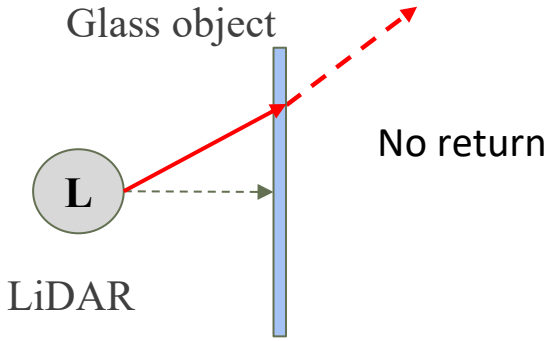
Objective: To Identify mirrors in LiDAR working environment by property of mirror symmetry.



Laser scans (a) and (c), and camera images (b) and (d) in environments with mirrors or windows



Data fabrication: two-hop path creates a virtual point at PF.



Data omission: beam reflected away, no return.

Introduction
Related work
Motivation
RW Threat
ORA
OAA
Autoware
Modelling
Simulation
Defence
Limitations
Discussion

Yang, Shao-Wen, and Chieh-Chih Wang. "On solving mirror reflection in lidar sensing." *IEEE/ASME Transactions on Mechatronics* 16.2 (2010): 255-265.

Motivation

Introduction
Related work
Motivation
RW Threat
ORA
OAA
Autoware
Modelling
Simulation
Defence
Limitations
Discussion

How do mirror-induced artifacts impact full-stack autonomous driving?

Contributions:



1. Real-world vulnerability Discovery



2. Physics-informed modeling



3. Simulation-based scalable evaluation



4. Full-stack impact validation.

Real-world vulnerability Discovery

Introduction
Related work
Motivation
RW Threat
ORA
OAA
Autoware
Modelling
Simulation
Defence
Limitations
Discussion



Object Removal Attack (ORA)



Object Addition Attack (OAA)

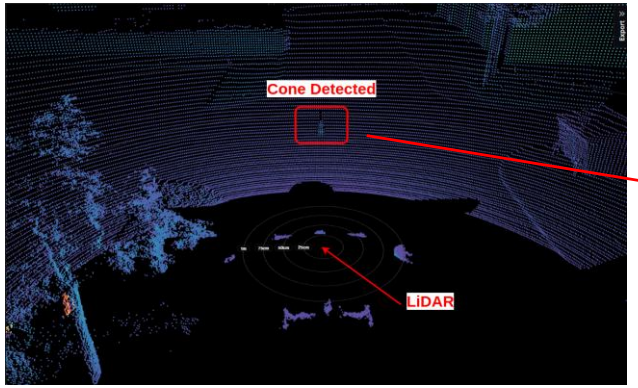
Object Removal Attack

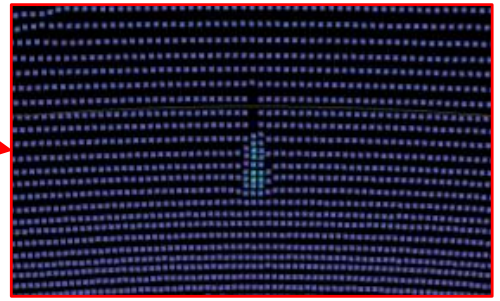
Outdoor Experimentation

Vehicle: Kia Soul EV **LiDAR**: Ouster OS1-128 (128-channel, 10Hz) **Mirror Configuration:** High-reflectivity glass panels with silver backing	**Computer**: Neousys **OS**: Kubuntu 22.04 (ROS2 Humble) **Data Collection:** ROS 2 logging with synchronized sensor data	**GPU**: Nvidia GeForce RTX 3070. **AD software**: Autoware **Test Environment:** Controlled outdoor environment, clear weather conditions



Outdoor Experimental Setup without ORA.

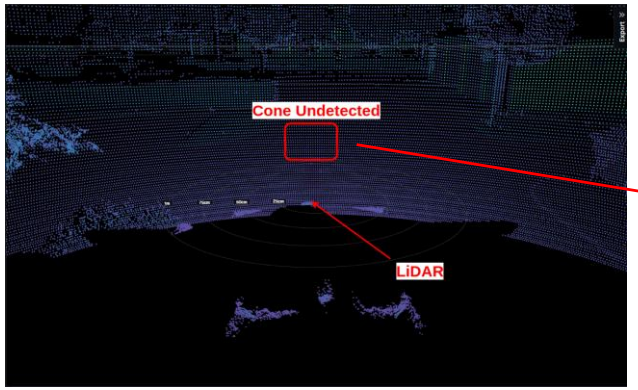


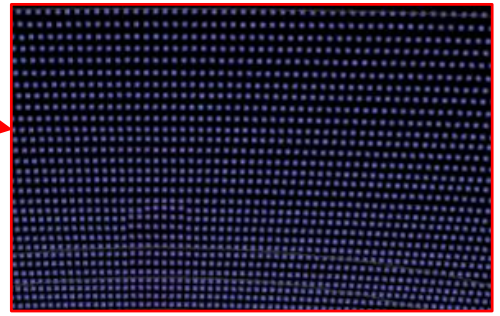


LiDAR point cloud without ORA.



Outdoor Experimental Setup of ORA.





LiDAR point cloud of ORA.

Introduction
Related work
Motivation
RW Threat
ORA
OAA
Autoware
Modelling
Simulation
Defence
Limitations
Discussion

Object Addition Attack (1)

Attack Idea

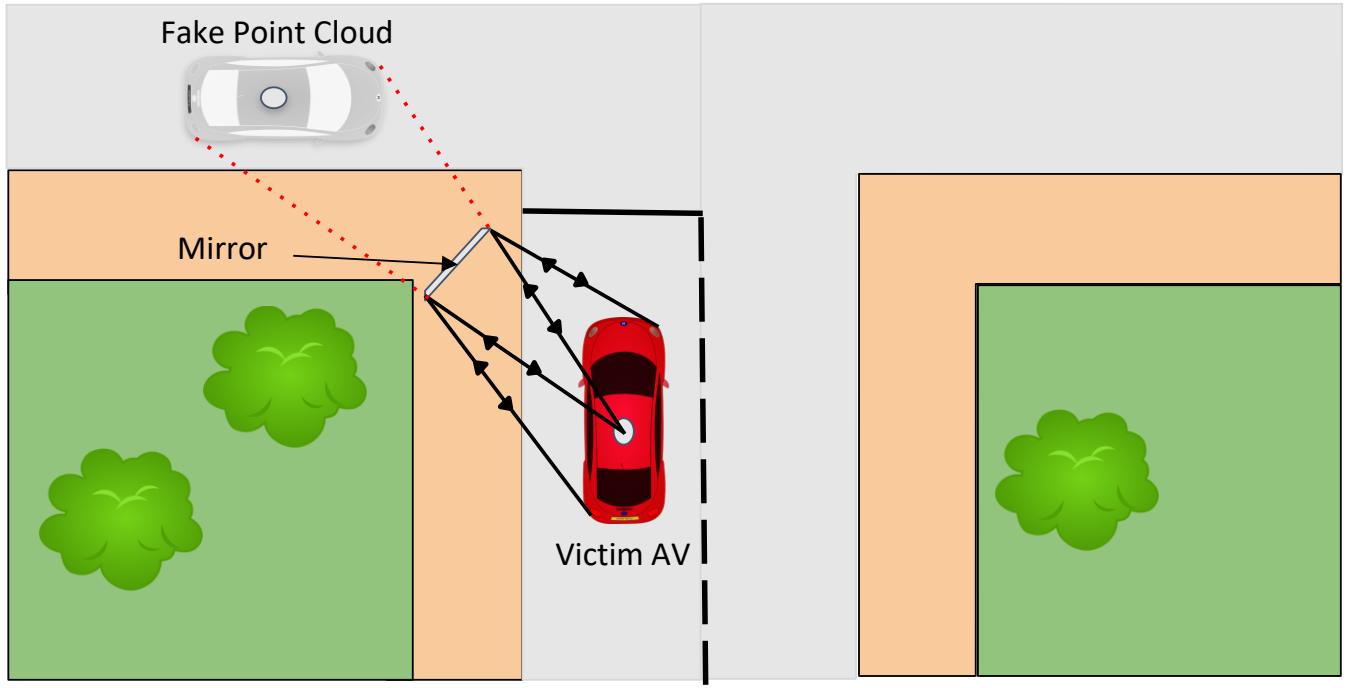


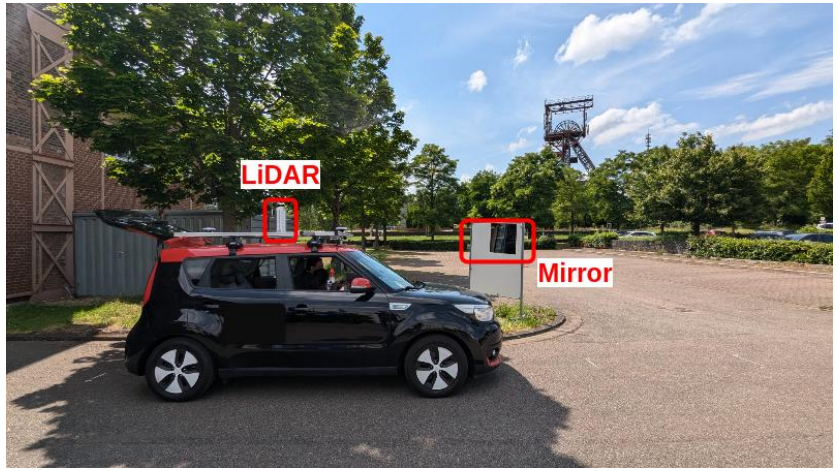
Illustration of Object Addition Attack

Introduction
Related work
Motivation
RW Threat
ORA
OAA
Autoware
Modelling
Simulation
Defence
Limitations
Discussion

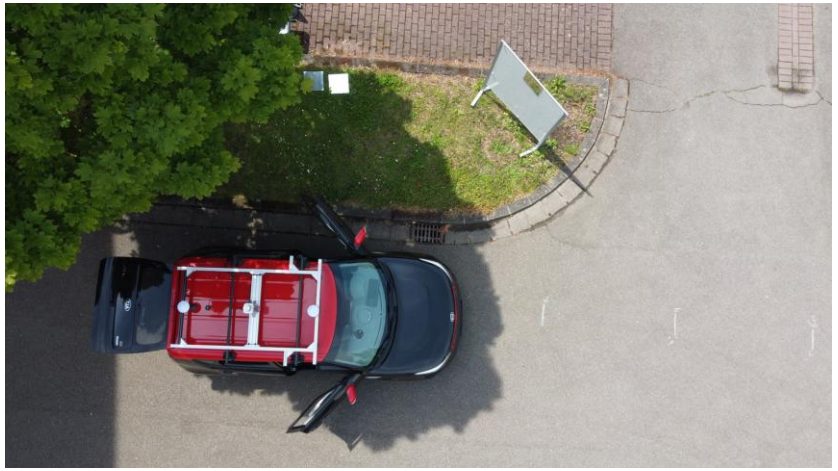
Object Addition Attack (2)

Outdoor Experimentation

Introduction
Related work
Motivation
RW Threat
ORA
OAA
Autoware
Modelling
Simulation
Defence
Limitations
Discussion



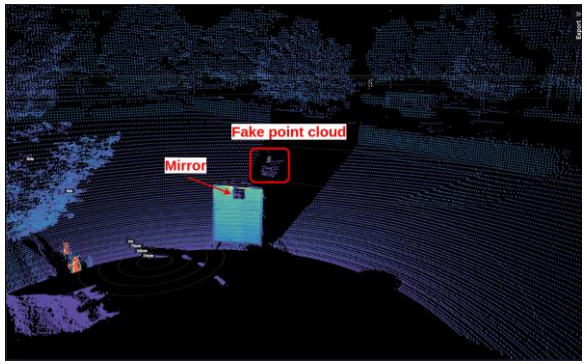
Side view of outdoor experimental setup of OAA.

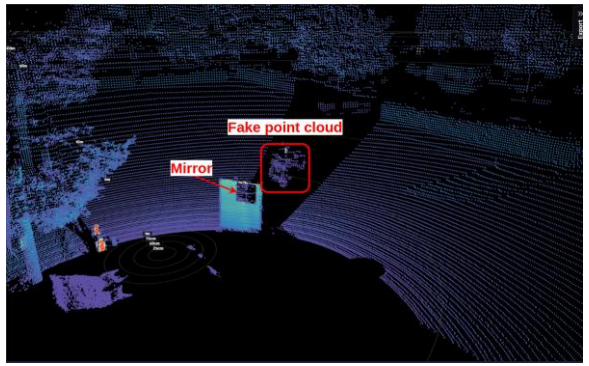


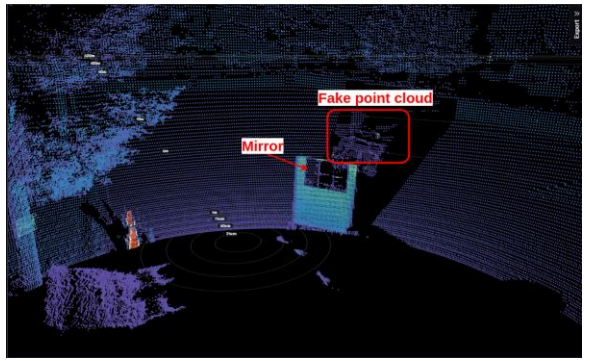
Top view of outdoor experimental setup of OAA.

Object Addition Attack (3)

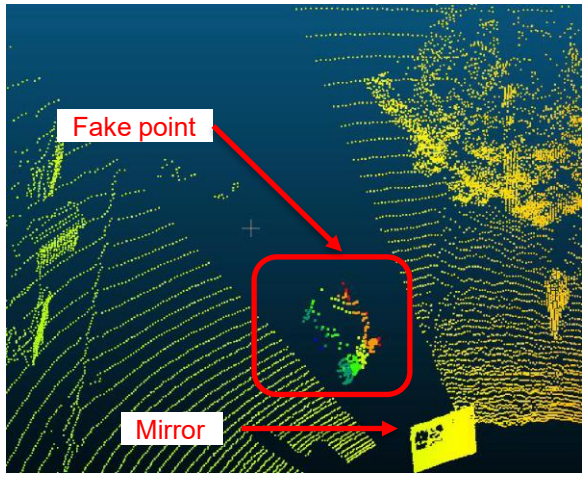
Outdoor Experimentation

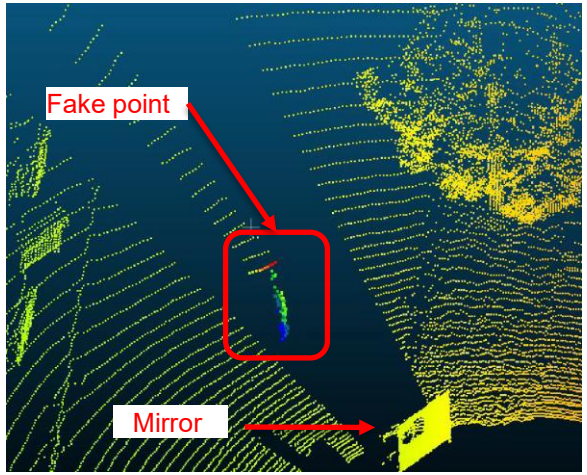


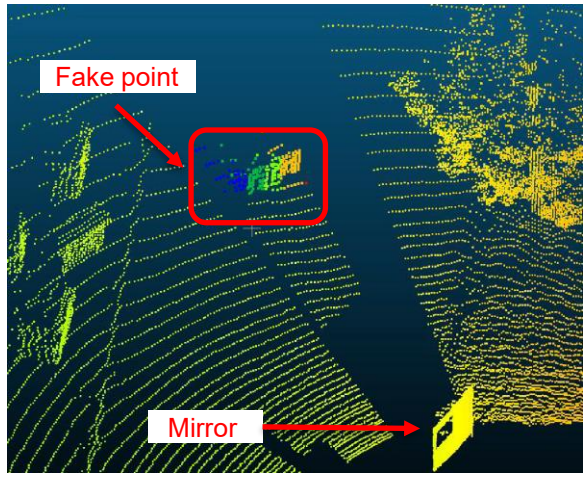




OA point clouds with varying mirror area. (a) 2 tiles: sparse cluster . (b) 4 tiles: denser vertical surface. (c) 6 tiles: high-density cluster approximating a vehicle profile







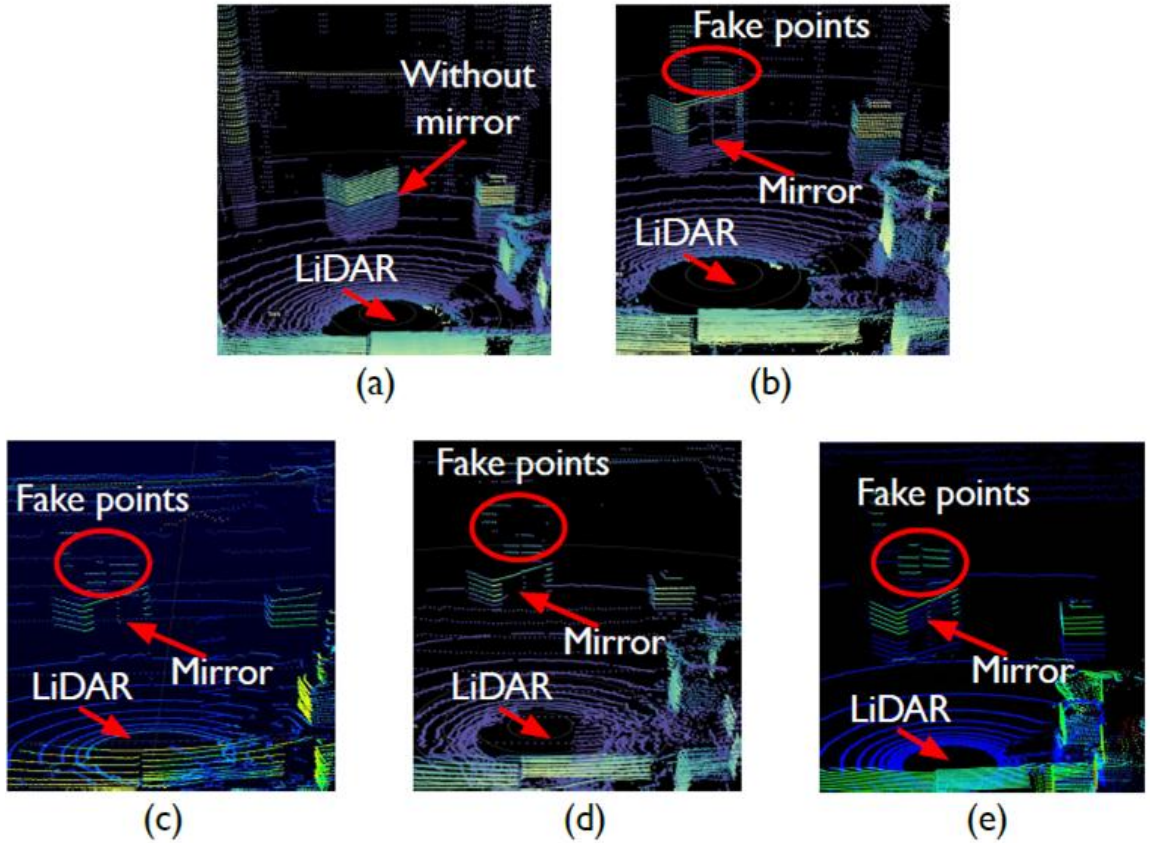
Impact of mirror tilt angle on phantom location: (a)30◦; (b) 45◦; (c) 60◦.

Introduction
Related work
Motivation
RW Threat
ORA
OAA
Autoware
Modelling
Simulation
Defence
Limitations
Discussion

Object Addition Attack (4)

Generality Across LiDAR Platforms

Introduction
Related work
Motivation
RW Threat
ORA
OAA
Autoware
Modelling
Simulation
Defence
Limitations
Discussion



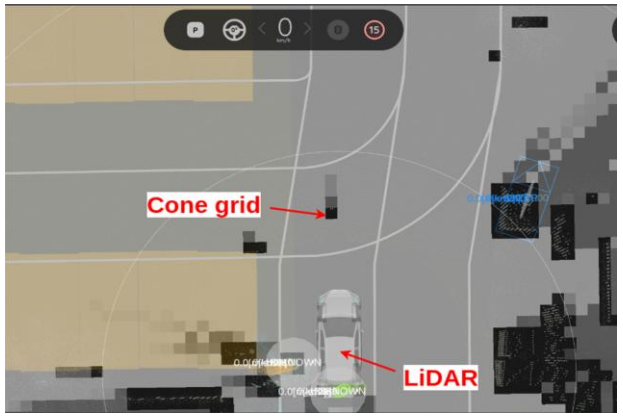
Object Addition across four LiDAR platforms. (a) Baseline (no mirror). (b) Ouster OS1-64. (c) Robosense RS-Helios 5515. (d) Ouster OS0-64. (e) Hesai XT32M2X

Vulnerability Assessment on Autoware (1)

Occupancy Grid Map - ORA



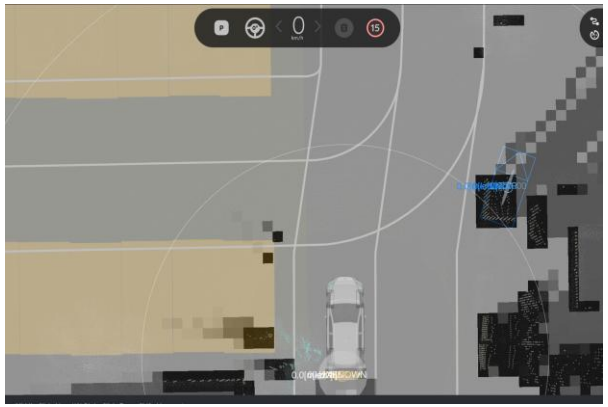
Outdoor Experimental Setup without ORA.



Occupancy Grid Map without ORA.



Outdoor Experimental Setup of ORA.

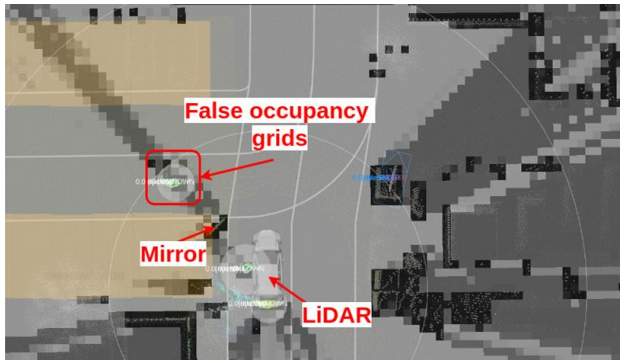


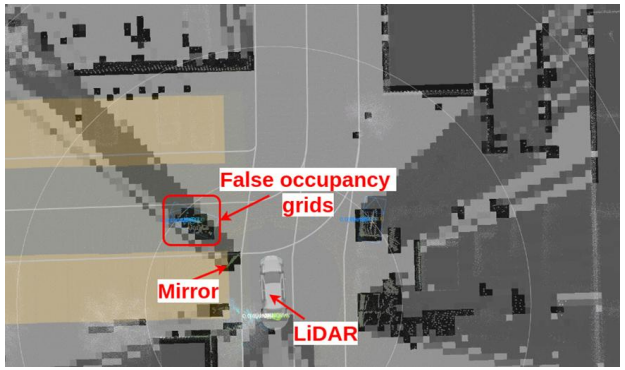
Occupancy Grid Map with ORA.

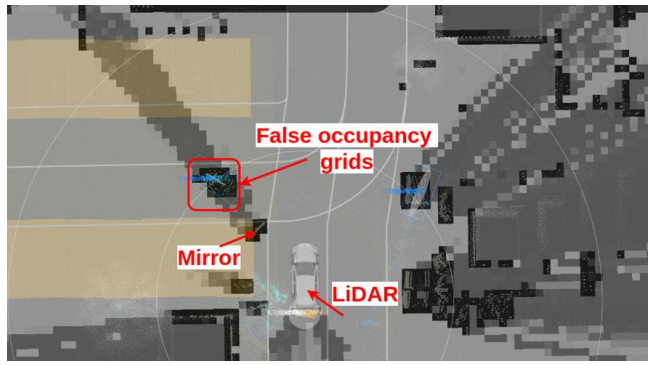
Introduction
Related work
Motivation
RW Threat
ORA
OAA
Autoware
Modelling
Simulation
Defence
Limitations
Discussion

Vulnerability Assessment on Autoware (2)

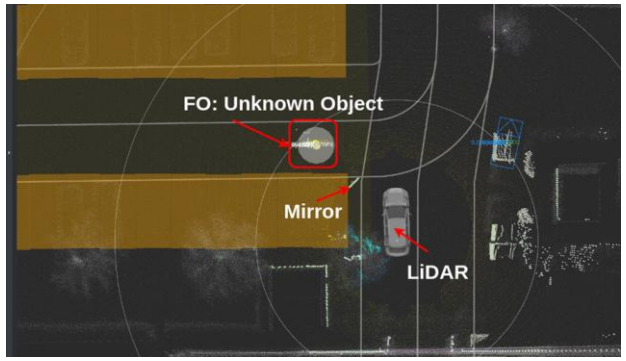
Occupancy Grid Map - OAA

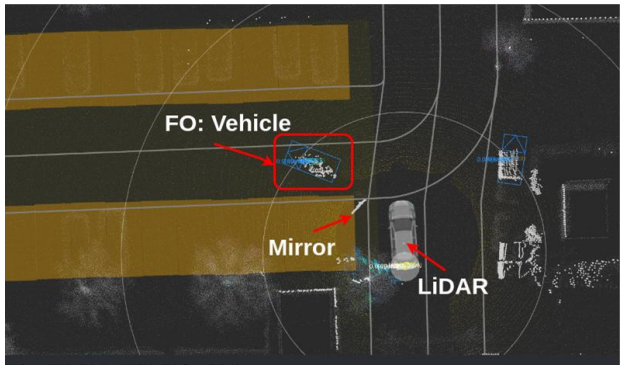


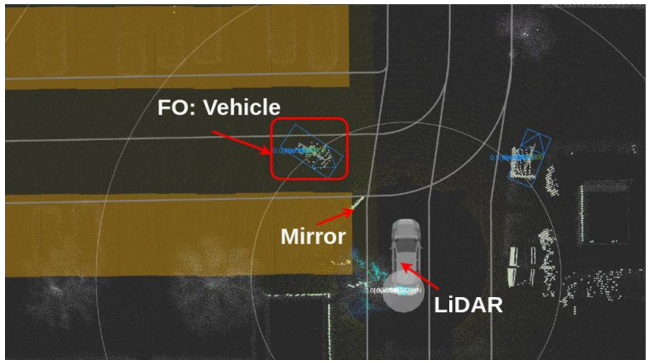




Occupancy grid corruption under Object Addition. (a) 2 tiles: thin line of occupied cells behind mirror. (b) 4 tiles: dense cluster triggers “CAR” detection. (c) 6 tiles: occupied area matches a real vehicle footprint.







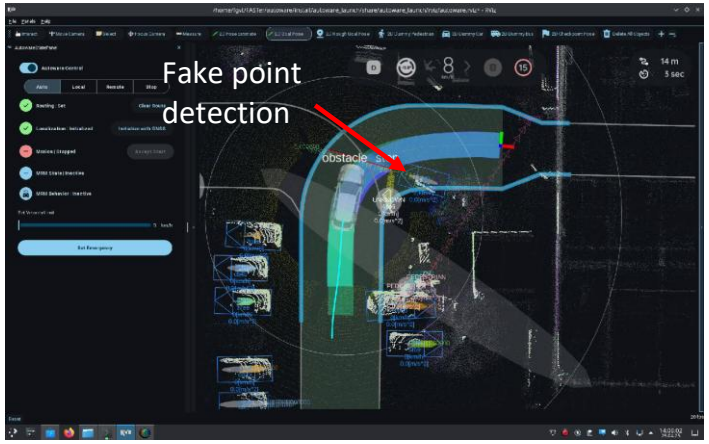
CenterPoint detection outcomes under Object Addition. (a) 2 tiles: “UNKNOWN” detection (white circle). (b) 4 tiles: “Vehicle” at 65% confidence (blue box). (c) 6 tiles: “Vehicle” at 74% confidence.

Introduction
Related work
Motivation
RW Threat
ORA
OAA
Autoware
Modelling
Simulation
Defence
Limitations
Discussion

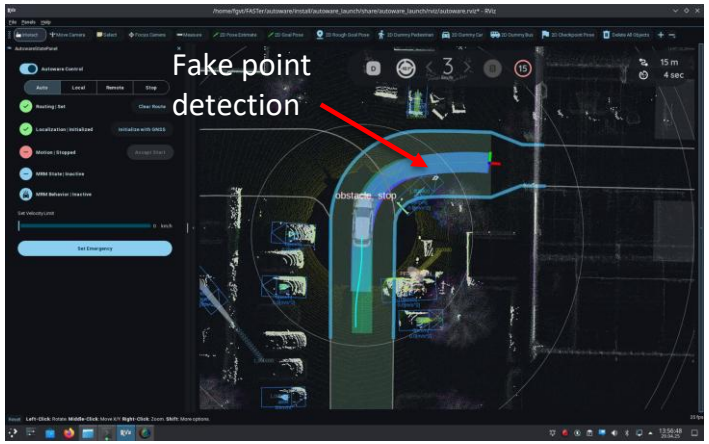
Vulnerability Assessment on Autoware (3)

Autonomous mode - OAA

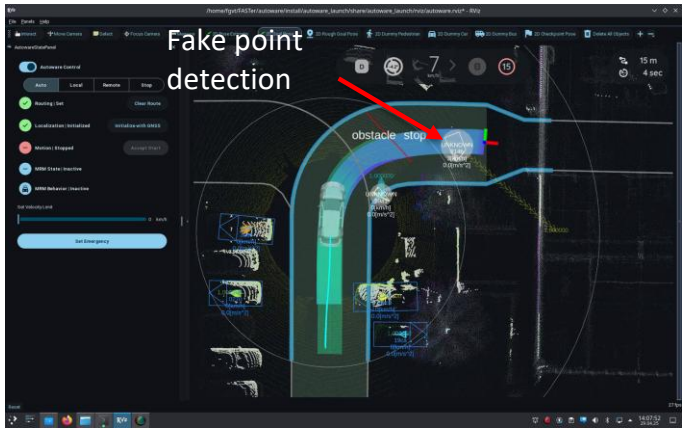
Introduction
Related work
Motivation
RW Threat
ORA
OAA
Autoware
Modelling
Simulation
Defence
Limitations
Discussion



(a)



(b)



(c)

Autoware's response to phantom detections with the 6-tile mirror at different angles. (a) 30◦: late-onset phantom triggers immediate emergency stop. (b) 45◦: earlier detection causes mid-turn braking. (c) 60◦: early detection triggers pre-turn full stop.

Empirical Modelling

Introduction
Related work
Motivation
RW Threat
ORA
OAA
Autoware
Modelling
Simulation
Defence
Limitations
Discussion

Purpose of analytical model:

- Provide a quantitative understanding of the vulnerability's underlying physics.
- Enable realistic artifact injection in simulation.

Feature extraction:

1. Radial distance (R_art) : Artifact's apparent distance from the sensor.
2. Lateral offset (X_art) : Displacement of the centroid from the LiDAR's fwd axis.
3. Point count (N_art) : the number of points per frame.
4. Appearance probability (P_app) : the fraction of frames in which an artifact cluster appears under a given configuration.

Target parameters:

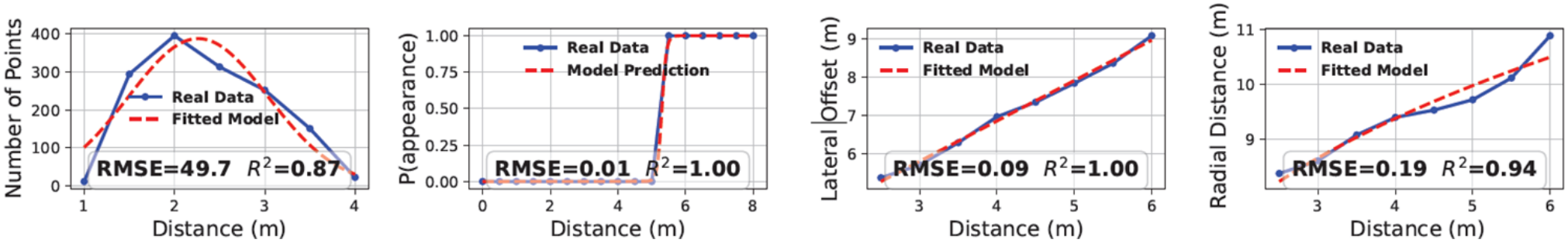
1. d: LiDAR-to-mirror distance
2. θ: Mirror tilt angle
3. A: Mirror surface area

$$R_{\text{art}}(d,\theta) = c_R\, d^{n_d}\,(1 + a_1\theta + a_2\theta^2)$$

$$X_{\text{art}}(d,\theta) = c_X\, d \tan(2\theta) + \delta_X$$

$$N_{\text{art}}(d,\theta,A) = c_0 A^{\beta} \cos^{\gamma}(\theta) \exp\left(-\frac{(d-\mu(\theta))^2}{2\sigma^2}\right)$$

$$P_{\text{app}}(d,\theta,A) = \frac{1}{1+e^{-k(d-d_{\min})}} \cdot \frac{1}{1+e^{k(d-d_{\max})}}$$

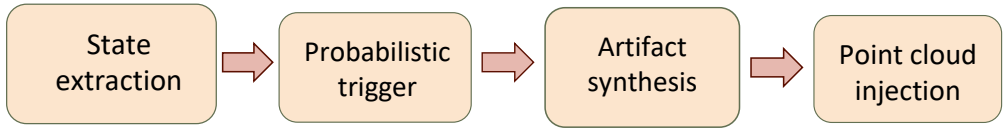


Model validation against experimental data for representative configurations.

CARLA Simulation Framework

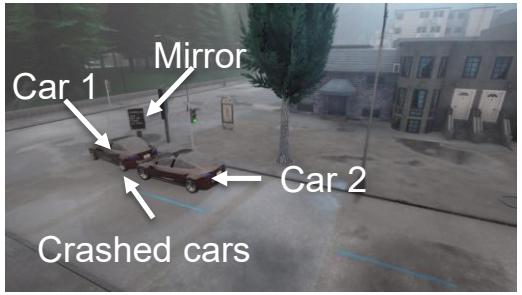
Physics-Informed Attack Injection Pipeline

Injection Pipeline:

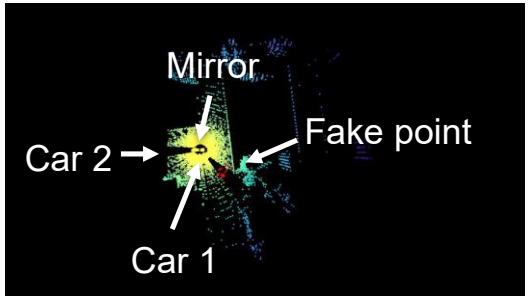


Point count and attack outcomes across mirror configurations:

Area (tiles)	Pts/Frame	Perception	Outcome
0.18 m^2 (2)	~ 25	UNKNOWN	Brake + Crash
0.36 m^2 (4)	~ 55	Car (65%)	Brake + Crash
0.60 m^2 (6)	~ 200	Car (74%)	Brake + Crash



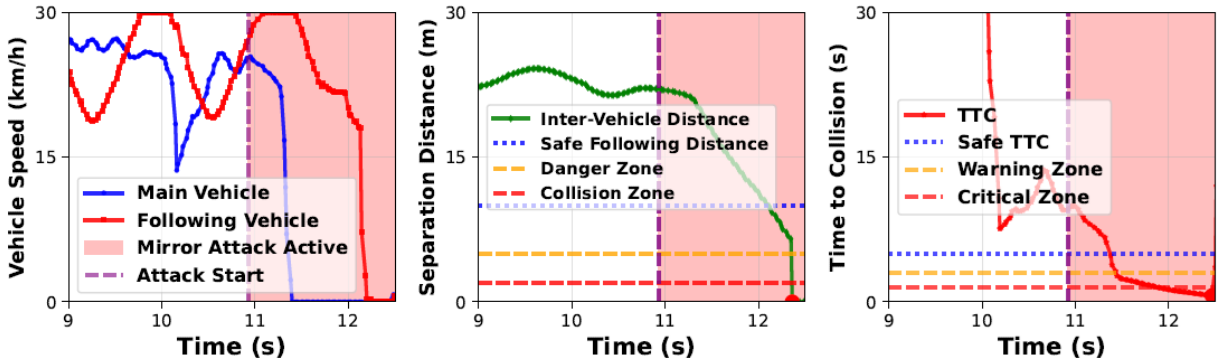
(a)



(b)

Cascade collision in CARLA. The Ego AV (Car 1) emergency-brakes for a phantom obstacle; the Follower (Car 2) rear-ends it. (a) Collision moment. (b) Ego AV's LiDAR view showing injected phantom points.





System-level dynamics following artifact injection at t ≈ 11.0 s. Ego AV emergency brakes; Follower Vehicle fails to react, leading to collision (separation and TTC reach zero).

Mirror attack activated **11.0s** → emergency braking **11.2s** → Insufficient reaction **11.8s** → rear-end collision **12.0s**

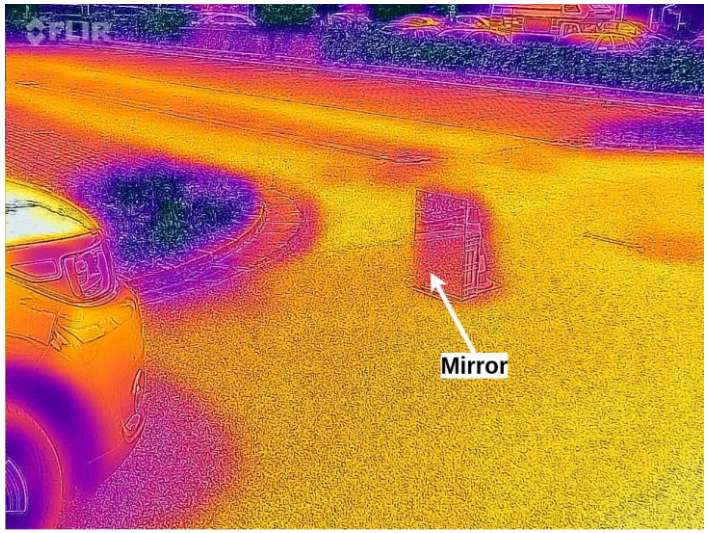
Introduction
Related work
Motivation
RW Threat
ORA
OAA
Autoware
Modelling
Simulation
Defence
Limitations
Discussion

Defence Strategies (1)

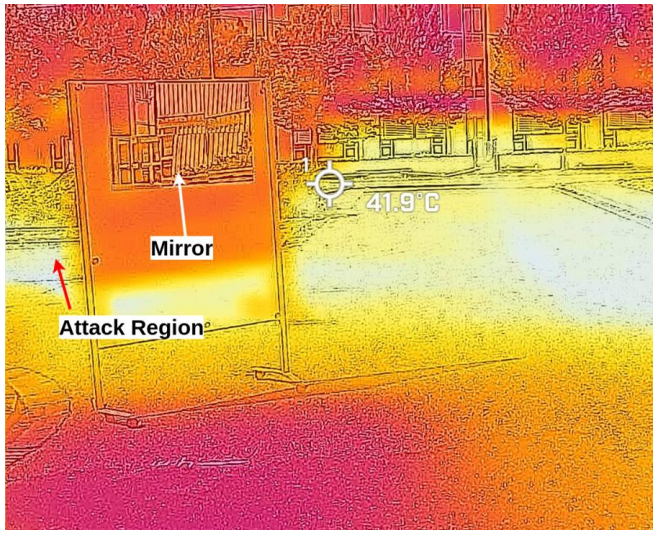
Thermal Camera Augmentation

Introduction
Related work
Motivation
RW Threat
ORA
OAA
Autoware
Modelling
Simulation
Defence
Limitations
Discussion

- **Core Principle:** Uses long-wave infrared (thermal) sensing to detect heat rather than reflected light.
- **Object Removal (OR) Defense:** Identifies mirrors directly as cool thermal anomalies contrasting against warm road surfaces.
- **Object Addition (OA) Defense:** Invalidate phantom objects by confirming the absence of a localized heat source.
- **Key Limitations:** Reduced thermal contrast in high ambient temperatures, lower resolution, high sensor cost, and lens contamination from rain or dirt.



Thermal camera images for ORA.



Thermal camera images for OAA.

Defence Strategies (2)

Intensity-Based Anomaly Detection

Introduction
Related work
Motivation
RW Threat
ORA
OAA
Autoware
Modelling
Simulation
Defence
Limitations
Discussion

- **Active Spoofing vs. Mirror Attacks:** Active lasers create abnormally bright returns, whereas mirror reflections use the vehicle's own pulses.
- **Physics Consistency:** Intensity and calculated distance remain fully consistent because light follows a natural two-hop path.
- **High-Intensity Detection Fails (AUC ≅ 0.12- 0.17):** Bouncing off the mirror slightly dims the signal, making artifacts look darker, not brighter.
- **Low-Intensity Detection Fails:** Real-world objects (dark pavement, shadows) share the same low-reflectivity tail, causing unacceptably high false-alarm rates (52 - 86 %).

Config	μ_{art}	μ_{base}	AUC	TPR @ 5% FPR
$(45°, 0.36\,\text{m}^2)$	3.1	12.9	0.12	0.7%
$(30°, 0.60\,\text{m}^2)$	4.1	12.8	0.13	1.6%
$(60°, 0.18\,\text{m}^2)$	10.4	12.7	0.17	4.4%

Intensity-based detection performance against mirror artifacts.

Limitations and Future Work

Introduction
Related work
Motivation
RW Threat
ORA
OAA
Autoware
Modelling
Simulation
Defence
Limitations
Discussion

- **Sensor Generality:** Tested on spinning ToF LiDAR; FMCW and full-waveform sensors require further validation.
- **Environmental Conditions:** Experiments ran in clear weather; rain, fog, and humidity may alter mirror reflectivity and beam dynamics.
- **Vehicle Speed Constraints:** Real-world tests were capped at ≤ 8km/h; high-speed behavior needs further empirical evaluation despite model predictions.
- **Defense Integration:** Proposed countermeasures are prototypes requiring full integration into production autonomy stacks with rigorous false-positive analysis.

Discussion

Introduction
Related work
Motivation
RW Threat
ORA
OAA
Autoware
Modelling
Simulation
Defence
Limitations
Discussion

Questions?

Contact:
gmohan2@scu.edu
selma.yahia@uni.lu

Thank You