An Adversarial Attack on DNN-based Adaptive Cruise Control Systems

Yanan Guo (University of Pittsburgh), Christopher DiPalma, Takami Sato (UC Irvine),
Yulong Cao (Univerisity of Michigan), Qi Alfred Chen (UC Irvine), Yueqiang Cheng (NIO)

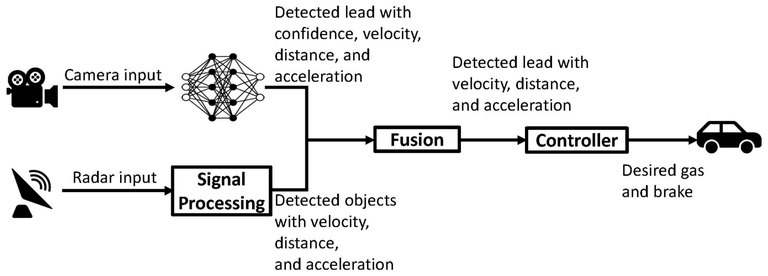
2021 ICCV OCTOBER 11-17 VIRTUAL

1. Introduction

- ACC is extremely safety critical since if it makes the wrong driving decision and accelerates/brakes unsafely, it may rear-end the vehicle in front or get rear-ended by the vehicle behind.
- Most ACC systems use camera as the only or major sensor to perceive the driving experiment, and use DNNs to extract the statues of the lead from the image captured by camera.
- We aim at building the first end-to-end ACC attack.
- We demo our attack on the OpenPilot+CARLA platform.
- A demo of the attack can be found at:
  https://sites.google.com/view/acc-adv/

2. OpenPilot

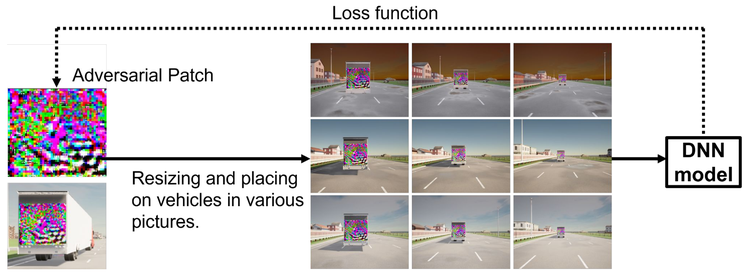
- OpenPilot is the state-of-the-art assisted driving system.
- Many studies have shown that OpenPilot has better performance that other systems.
- ACC in OpenPilot:
  - Keep a safe distance from the leading vehicle (if exists)
  - Maintain the same velocity with the leading vehicle (if exists)



3. Attack Overview

- Threat Model.
  - We assume a threat model that the attacker can put a patch on the leading vehicle.
  - This can be achieved in the scenario with truck back advertisement.

- Attack Method.
  - We collect real driving scenarios from CARLA, and based on those images, we use EOT to generate a robust adversarial patch.



- The lossfunction consists of three parts.
  - The loss function of the confidence score of the lead.
  - The loss function of the detected distance between the ego vehicle and the lead.
  - The loss function of total variance.

1. Attack Results

- Attack Scenario (4 phases)
  - The ACC vehicle is following the lead as normal at a constant velocity which is 30-80 mph, and the following distance is thus 28-70 m.
  - The lead starts to slow down due to some special situation in front of it. Later, the ACC vehicle starts to slow down accordingly, and it will get closer to the lead.
  - When the distance between the two vehicles is smaller than a threshold (25 m), the adversarial patch starts to take effect and the ACC vehicle starts to accelerate until hitting the lead.
- The following figure is showing the benign/attack process where the ACC vehicle is initially following the lead at the velocity of 40 mph. dis denotes the real distance between the two vehicles and d^c_t denotes the detected distance.