Demo: Attacking LiDAR Semantic Segmentation in Autonomous Driving

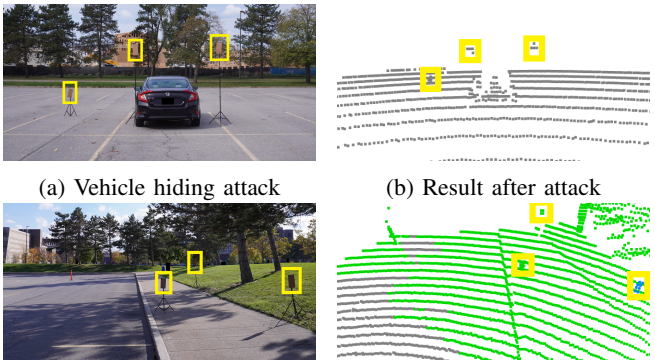
Yi Zhu[1], Chenglin Miao[2], Foad Hajiaghajani[1], Mengdi Huai[3], Lu Su[4], Chunming Qiao[1]
[1] State University of New York at Buffalo, [2] University of Georgia, [3] University of Virginia, [4] Purdue University
[1]{yzhu39, foadhaji, qiao}@buffalo.edu, [2] cmiao@uga.edu, [3] mh6ck@virginia.edu, [4] lusu@purdue.edu

***Abstract*—As a fundamental task in autonomous driving, LiDAR semantic segmentation aims to provide semantic understanding of the driving environment. We demonstrate that existing LiDAR semantic segmentation models in autonomous driving systems can be easily fooled by placing some simple objects on the road, such as cardboard and traffic signs. We show that this type of attack can hide a vehicle and change the road surface to road-side vegetation.**

The development of autonomous vehicles (AVs) has gained an increasing amount of momentum in recent years. To understand the driving environment, autonomous vehicles rely on various sensors, such as camera, LiDAR and radar. Among them, LiDAR is becoming increasingly important because it can provide accurate 3D point cloud representation of the surrounding environment. In LiDAR perception systems, LiDAR semantic segmentation is a fundamental task, which aims to assign each point with a class label such as vehicle, road, or vegetation. LiDAR semantic segmentation helps achieve semantic scene understanding and has enabled many applications in autonomous driving, such as vehicle detection, road boundary detection, and 3D environment reconstruction.

Despite of the importance of LiDAR semantic segmentation, its vulnerability has not been well studied. The state-of-the-art LiDAR segmentation models mainly rely on deep neural networks (DNNs) to achieve good performance. However, recent studies have shown that, when the inputs are images, adding some small perturbations to the input images can change the DNNs' outputs drastically, which is referred as adversarial attack. It is entirely possible for an attacker to perform a similar attack by slightly changing the driving environments and fool the LiDAR segmentation models.

This demo is built upon our previous work [1], which is the first study on the vulnerability of LiDAR semantic segmentation. We propose a practical attack that can fool the LiDAR segmentation models of the victim AV. We show that, by placing some simple objects around some specific locations on the road, the attacker can manipulate the outputs of LiDAR semantic segmentation models. These objects are referred as adversarial objects, and they can be in any shape as long as they can reflect laser. An Adversarial Location Generation framework is proposed to derive the locations of these adversarial objects in an offline manner, based on the surrogate point cloud data collected before the attack using a similar LiDAR sensor as that in victim AV.



(a) Vehicle hiding attack (b) Result after attack

(c) Road surface changing attack (d) Result after attack

Fig. 1: Visualizations of two attack scenarios [1]. The blue points belong to "Vehicle", the grey points belong to "Ground", and the green points belong to "Vegetation".

In this demo, we use an Ouster OS1-64 LiDAR mounted on top of a vehicle as the testbed. We will demonstrate two attack scenarios: vehicle hiding attack and road surface changing attack. In vehicle hiding attack, we consider a victim AV driving on the road and a vehicle parked in front of it. The parked vehicle can be parked by the attacker intentionally. The attack goal is to hide the parked vehicle from the LiDAR segmentation model used by the victim AV. To perform the attack, we use the cardboard held by poster stands as the adversarial objects and place them around some specific locations derived by our attack framework, as shown by the yellow rectangles in Figure 1a. We will show that this attack can continuously hide the parked vehicle from the victim AV. In road surface changing attack, the attack goal is to change the road surface to roadside vegetation in LiDAR semantic segmentation. We also use cardboard as the adversarial objects and constrain their locations on the roadside. We will show that the road surface in front of the victim vehicle are changed to vegetation, as shown in Figure 1d.

References

[1] Y. Zhu, C. Miao, F. Hajiaghajani, M. Huai, L. Su, and C. Qiao, "Adversarial attacks against lidar semantic segmentation in autonomous driving," in *Proceedings of the 19th ACM Conference on Embedded Networked Sensor Systems*, pp. 329–342, 2021.

Workshop on Automotive and Autonomous Vehicle Security (AutoSec) 2022
24 April 2022, San Diego, CA, USA
ISBN 1-891562-75-4
https://dx.doi.org/10.14722/autosec.2022.23022
www.ndss-symposium.org