(2020 USENIX)Towards Robust LiDAR-based Perception in Autonomous Driving: General Black-box Adversarial Sensor Attack and Countermeasures

1 Summary

In autonomous driving cars, the LiDAR sensor plays pivotal role in the perception module. In this paper, the author proposed the first black-box LiDAR spoofing attack on state-of-art LiDAR sensor and the neural network model behind it. To achieve this goal, the author first analyzed the occlusion pattern based on previous study. After identifying this phenomenon, the author conduct the attack with a high success rates. The author also devised two effective countermeasures which don't need to modify the existing hardware.

2 Challenge

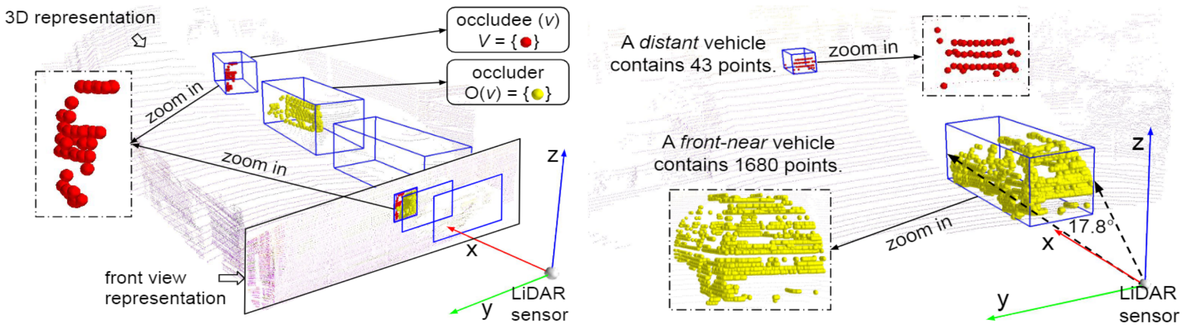
Successful LiDAR spoofing attack has been realized in previous study. However, previous work has an assumption that attackers must know exactly how the neural networks model behind the LiDAR are constructed, which lose generality and need to construct adversarial examples for every specific models. Currently, some methods to conduct black-box attack on camera have emerged, but these method need the attack to construct an identical model by themselves which is hardly to achieve.

3 Main Idea

3.1 Design-level Vulnerability

Unlike previous works that conduct spoofing attack blindly, the author first analyze the reason why previous work can be successful with such a little number of spoofed points.
The author first identified two different scenarios where a physical car has much fewer points in the output of LiDAR.

- an occluded vehicle
- a distant vehicle
  the picture below show the two different scenarios.



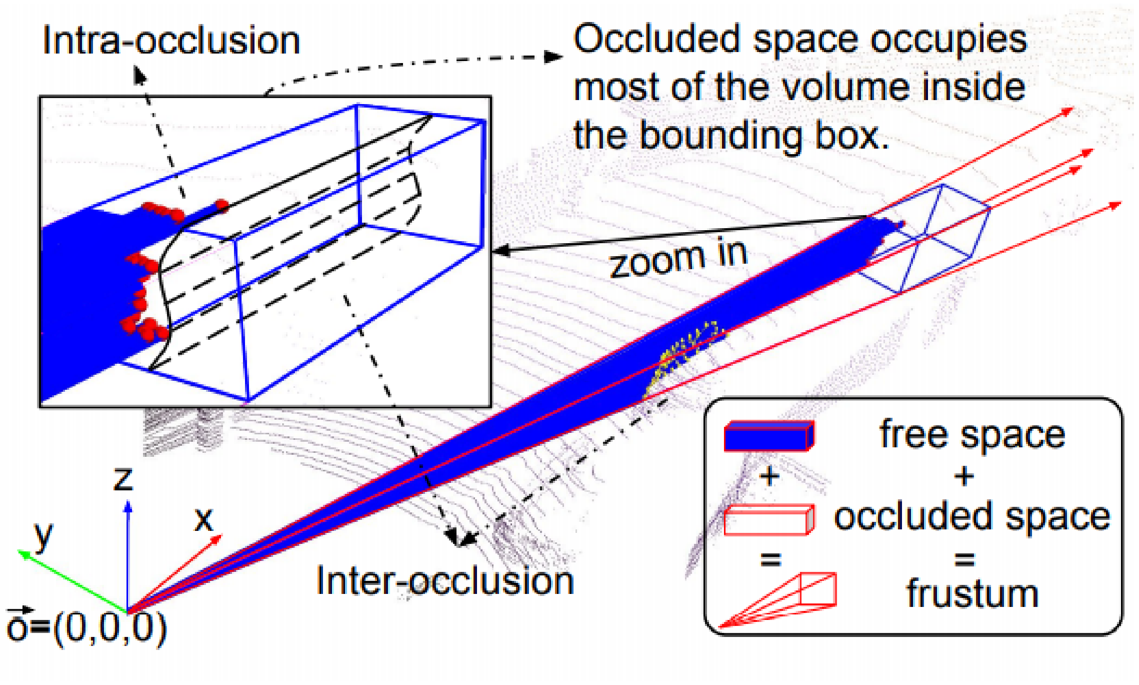
The author made an assumption which is the central idea in this paper: the two categories of points cloud can achieve the same results without factors that cause the decrement of points. So, attackers can also achieve their goals by imitating these pattern.

3.2 Black-box Spoofing Attack

Based on the idea above, the author proposed a black-box spoofing attack algorithm that works well for all state-of-art perception model.
First, the author used real-world physical traces of occluded or distant vehicles to perform the attack. Specifically, the author pick points clouds that conform the pattern above and do spatial transformation to make target points closer to the vehicle. With the assumption proposed above, the removed points clouds can also fool the perception model.
To make the attack more flexible, the author also proposed a method which generates spoofed points from scratch. That is, get a 3D mesh firstly and then use a renderer to simulate the function of a LiDAR.

3.3 Physics-Informed Anomaly Detection

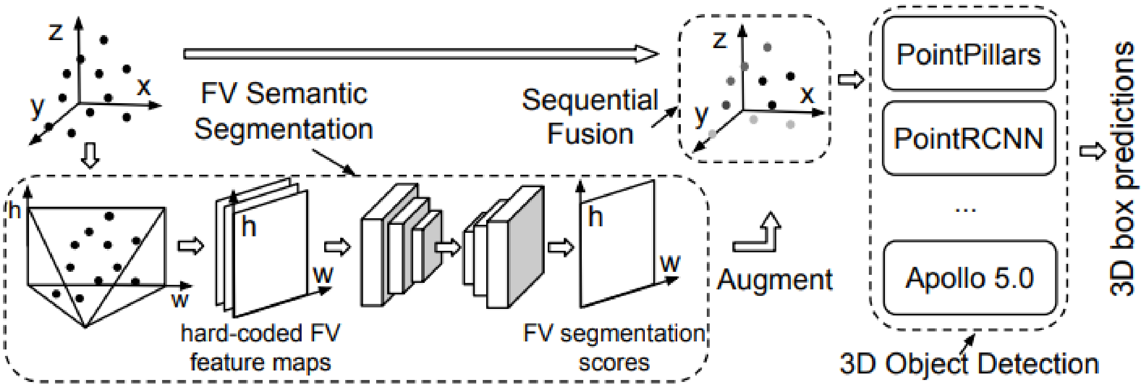
To detect the occlusion patterns in real-time the author design CARLO(oCclusion Aware hieRarchy anomaLy detectiOn) that detects the violation of physics.



Specifically, the author design two metrics, FSD and LPD which denote the ration of the occluded and unoccluded points, If these metrics is under some threshold, we can know clearly that it's a fake object. However the computational cost of the FSD is too high and the LPD can't get a high accuracy standalone, so the author design a hierarchical structure that uses LPD as a rough filter. The result of LPD is then sent to the FSD to get the final result.

3.4 Physical-embedded Perception Architecture

CARLO can effectively detect the spoofed objects. However, it would be better if we can eliminate these vulnerabilities from the model aspect. To achieve this goal, we must change the design of the perception algorithm.
The popular bird-eye view based algorithms have its inherent limitation to distinguish the spoofed points with occluded objects. It is obvious that the front view can mitigate the effect of occlusion. However, compared with BEV based methods, current FV based models have a lower accuracy. So, to use FV in perception module, it's inevitable to devise a new model.



To address this problem, the author proposed sequential view fusion illustrated in the picture above, the sequential view fusion consists of three main components.

- Semantic segmentation: utilize FV representation to compute the point-wise confidence score
- View fusion: 3D representation is augmented with the semantic segmentation score
- 3D object detection: the augmented point cloud is fed into the state-of-art attacks to get the perception result.

Adversarial Objects Against LiDAR-Based Autonomous Driving Systems

1 Solved Problem

Deep neural network has been proved to be vulnerable to adversarial sample attacks. However, tradition gradient based algorithm in the image domain can't get a great result because of some natures of the 3D points cloud.
In this paper, the author proposed the first attack method that generating spoofed LiDAR points using a render that mitigate the behavior of the laser scan. To evaluate the invented method, the author bring the adversarial object in reality using 3D print. The result shows that the attack is effective to state-of-art industrial autonomous driving platforms such as Baidu Apollo.

2 Main Idea

At first, a mesh of 3D object is created to be the input of the total pipeline. The mesh is processed by a renderer which simulate the lase beam on the surface of the object. After that, the object is transformed to a set of points in the final points cloud.
The renderer proposed above is differentiable so that it can use optimization-based algorithm to get adversarial examples.
After generating the spoofing object, the author used 3D printing technology to get the real object. The author then evaluated its performance in real-world cars and the result shows that it can foo the state-of-art LiDAR perception algorithm effectively.

3 HighLights Worth Learning

1. perform the real-world experiments
2. the idea to generating spoofed points using 3D modeling
3. the evaluation black box algorithm.

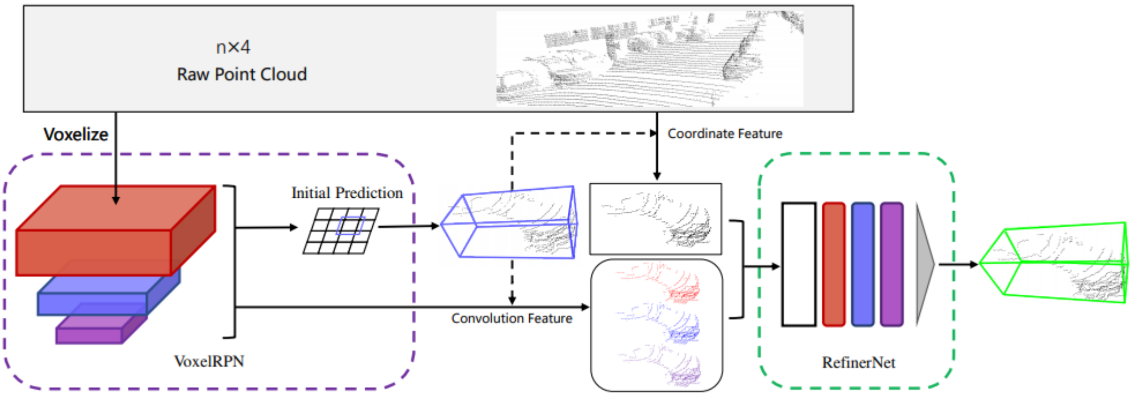
(2019 ICCV) Fast Point R-cnn

1 Solved Problem

State-of-art 3D object detection algorithms based on bird-eye view and voxel view are likely to loss information contained in the raw points cloud. However, applying CNN directly on the 3D points cloud has been proved computational expensive.
In this paper, the author proposed a novel two-stage neural network to try to get both high quality and efficiency.

2 Main Idea

In this paper, the author presented a two-stage 3D object detection framework which hybrids the voxel and raw point cloud.



The method first take the voxel represent of raw point cloud to a VoxelRPN network to get a set of initial prediction in high speed. In the second stage, the initial prediction is fused with the feature extracted and taken as input to RefinerNet to get the final results with better accuracy.

3 Highlights Worth Learning

1. the idea to fuse voxel representation with the raw point cloud to make use of uncompressed information.