Depth Priors-informed Purification Defense for Car-Borne LiDAR Vehicle Detection

Yihan Xu
Nanyang Technological University
Singapore

Zimo Ma
Nanyang Technological University
Singapore

Qun Song
City University of Hong Kong
Hong Kong SAR, China

Jianping Wang
City University of Hong Kong
Hong Kong SAR, China

Rui Tan
Nanyang Technological University
Singapore

Abstract

LiDAR-based 3D vehicle detection is a fundamental perception task of autonomous driving systems. However, recent studies show that physical or physically plausible adversarial examples can severely degrade the vehicle detection performance and then affect downstream tasks such as motion planning. However, the existing defense methods do not achieve satisfactory trade-offs between computational efficiency and defense effectiveness. In this paper, we identify two attack-indicative priors in the depth of the perturbed point cloud area and propose a two-stage informed purification algorithm to remove adversarial points while keeping essential benign points for vehicle detection. With low computational overhead, this new input purification achieves defense performance comparable to the state-of-the-art neural network-based methods while remaining highly efficient.

CCS Concepts

• **Computer systems organization** → **Embedded and cyber-physical systems**; • **Security and privacy** → *Systems security*.

ACM Reference Format:
Yihan Xu, Zimo Ma, Qun Song, Jianping Wang, and Rui Tan. 2026. Depth Priors-informed Purification Defense for Car-Borne LiDAR Vehicle Detection. In *The 24th Annual International Conference on Mobile Systems, Applications and Services (MobiSys Workshop '26), June 21–25, 2026, Cambridge, United Kingdom.* ACM, New York, NY, USA, 5 pages. https://doi.org/10.1145/3812836.3815164

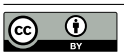
1 Introduction

LiDAR-based 3D object detection is a fundamental perception task for autonomous driving [1, 2]. It recognizes surrounding objects and traffic participants and provides inputs for planning and control. Among various on-road objects, the nearby vehicles form a critical class of objects because the interactions with them are tightly coupled with the ego vehicle's routine driving behaviors, such as car-following, braking, lane changing, and merging.

Despite its good performance in benign settings, LiDAR-based vehicle detection is vulnerable in adversarial environments. Recent work has shown that LiDAR detectors can be misled by adversarial attacks that modify the observed point cloud [6, 12, 13, 18, 21, 24]. This poses a substantial security risk for autonomous driving systems. Among various attack goals, vehicle hiding [3, 5, 9, 10, 14–17, 20, 26–28] is concerning because it is the most widely studied and practical goal in real-world settings. The vehicle hiding attacks place adversarial objects [3, 17, 26, 28] or inject points with a laser transmitter [10] around a *target vehicle*. As a result, a *victim vehicle* may fail to recognize the target vehicle in front, leading to missed reactions and therefore collisions.

To counteract adversarial examples, various defense mechanisms have been proposed and some of them are specifically devised for LiDAR-based object detection. Model enhancement methods, such as adversarial training [8], improve robustness by augmenting training data, but often degrade clean detection performance. Recent work employs ensemble-based defenses [19], where ensemble and dynamic parameter generation techniques are combined to enhance robustness. Although such methods can achieve strong robustness, their compute cost is proportional to the number of ensemble members. Purification methods aim to remove adversarial points before object detection and have shown effectiveness against object hiding attacks. However, recent purification methods typically rely on complex neural networks, such as reinforcement learning policies, reconstruction networks, or diffusion models [4, 22, 23], which result in substantial compute overhead. A notable exception among these purification methods is Simple Random Sampling (SRS) [25], a purification defense that can operate entirely on the central processing unit (CPU). However, as SRS removes adversarial and benign points indiscriminately, it is subjected to a tussle between retaining clean accuracy and pursuing defense performance. This limitation highlights the need for a lightweight, targeted purification mechanism that can distinguish adversarial points and benign points.

In this paper, we propose a fast purification method called Depth Priors-Informed Purification (DPP) against vehicle hiding attacks. DPP is built on two geometric metrics of LiDAR returns, namely, *depth gradient* and *scanline consistency*, that can be computed readily and exhibit different value ranges for adversarial and benign points. After projecting suspicious point clouds onto a depth map, DPP performs targeted removal of the adversarial points through a two-stage filtering process based on the two metrics. Evaluation shows that DPP achieves effectiveness comparable to the state-of-the-art ensemble-based defense [19], while requiring 1/10 compute time, making it well-suited for deployment on resource-constrained autonomous platforms. Besides, DPP improves the vehicle detection



MobiSys Workshop '26, Cambridge, United Kingdom

ACM ISBN 979-8-4007-2712-2/2026/06
https://doi.org/10.1145/3812836.3815164

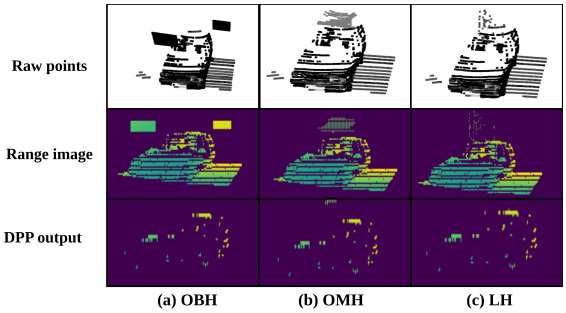


Figure 1: Three attacks and DPP's output range images

accuracy in the presence of attack by up to 80%, compared with the lightweight defense baseline SRS.

Paper organization: §2 presents key observations and opportunities of DPP. §3 and §4 present the design and evaluation results, respectively. §5 concludes this paper and outlines future directions.

2 Observations and Opportunities

To design an effective and efficient defense against vehicle hiding attacks, we analyze the geometric characteristics of adversarial points in LiDAR range images and identify two key observations. These observations reveal exploitable differences between adversarial points and genuine vehicle structures, which motivate our purification design.

Observation 1: Adversarial points tend to exhibit significantly weaker local depth gradient than genuine vehicle surfaces in the range image. A LiDAR point is considered valid if its corresponding laser ray returns a non-zero depth value. Due to the complex 3D geometry and large size of real vehicles, adjacent valid pixels on visible surfaces or object boundaries often exhibit noticeable depth gradient after projection onto the range image. These pixels usually preserve the structural outline of the vehicle and retain sufficient geometric information for 3D detection. Adversarial points, in contrast, have small or near-zero depth gradient. In this paper, we consider two representative object hiding attacks, as illustrated in the first two columns of Figure 1. The Object-Board Hiding (OBH) attack [28] introduces planar cardboard surfaces; the Object-Mesh Hiding (OMH) attack [17] produces mesh-like structures composed of locally smooth facets. These two attacks lead to small depth variation within local neighborhoods.

Opportunity 1: A filter may enable the separation of adversarial points from genuine vehicle structures in mixed point clouds with depth gradient information. Based on Observation 1, we first apply a filtering step (referred to as *first-stage filtering*) that retains only valid points with sufficiently large depth differences from adjacent valid pixels. This strategy removes adversarial points that form locally smooth structures while preserving structurally informative vehicle points. To quantitatively evaluate this effect, we use the *retained ratio*, defined as the fraction of points preserved after filtering. In Table 1, we report both the retained ratio of clean points, which include vehicle and surrounding non-adversarial points within the region of interest (ROI), and adversarial points after the first-stage filtering (F1). The filtering effect

Table 1: Retained points ratio after filtering stages (F1: Based on Depth Variation, F2: Based on Scanline Consistency).

Attack	Filtering	Clean ↑	Attack ↓
OBH	F1	0.373	0.044
OBH	F1+F2	0.288	0.044
OMH	F1	0.373	0.213
OMH	F1+F2	0.288	0.070
LH	F1	0.373	0.224
LH	F1+F2	0.288	0.013

is evident for OBH, since points inside each cardboard plane usually have small local depth variation. OMH shows a similar trend, although residual points on mesh regions that produce relatively large local depth variation. Quantitatively, as shown in Table 1, the first-stage filtering (F1) reduces the retained adversarial points to only 4.4% for OBH and around 21% for OMH.

However, the depth gradient-based criterion is less effective for the third attack called Laser Hiding (LH) [10], where injected points can be independently controlled in position and depth. As a result, a non-negligible portion of adversarial points (22.4%) survives after F1, motivating an additional filtering mechanism.

Observation 2: Adversarial points often violate the local scanline consistency observed on real object surfaces in the range image. Along each LiDAR scanline, valid points from common objects typically exhibit locally consistent depth patterns, where depth values change smoothly in a fixed direction or remain nearly constant within a short segment. We refer to this property as *scanline consistency*. In contrast, adversarial points, especially those generated independently, often introduce irregular depth fluctuations and break this consistency. This effect is particularly pronounced in LH, where injected laser points can be independently controlled in both position and depth, and also appears in OMH, where sharp transitions between mesh facets create inconsistent depth patterns.

Opportunity 2: Scanline consistency can serve for removing independent injected adversarial points. Based on Observation 2, we introduce a second filtering stage that further refines the retained points by checking whether each scanline contains a sufficiently continuous monotonic or near-constant depth trend. Again, we use the retained ratio to evaluate this effect, as defined previously. As shown in Table 1, applying both filtering stages (F1+F2) significantly reduces the retained adversarial points compared to F1 alone. In particular, the retained ratio for OMH drops from around 21% to 7%, and for LH from 22.4% to only 1.3%, demonstrating that scanline consistency effectively removes adversarial points that evade depth gradient-based filtering.

Although F2 further reduces the number of retained clean points, the preserved points remain aligned with detection-critical structures. To verify this, we conduct an evaluation experiment on 400 clean frames and find that more than 95% of vehicles remain detectable after applying both filtering stages.

We further quantify point importance using a gradient-based saliency score. For each input point, we compute the logarithm of the gradient magnitude of the detection loss. This score measures how strongly the predicted bounding box depends on that point.

Table 2: Average saliency score.

Original points	After SRS	After DPP
0.34	0.34	0.43

Table 2 reports the average per-point saliency score for three point sets: original points, points retained by DPP, and points retained by SRS. For a fair comparison, SRS is adjusted to use the same drop ratio as DPP for each sample. DPP retains points with an average contribution score that is 24% higher than that of randomly selected SRS points. This allows us to directly assess whether DPP preserves more influential points rather than simply retaining random points.

The two physically grounded filtering criteria form the complete DPP pipeline, whose detailed algorithm is presented in §3. As shown in Figure 1, this two-stage filtering process largely preserves structurally informative vehicle points while removing most adversarially inserted points.

3 Design of DPP

DPP is designed based on the observations in §2, i.e., the adversarial points differ from genuine vehicle surfaces in both local depth variation and scanline structure. Guided by these insights, the core of DPP is a *Depth Priors-informed Filtering* module that selectively preserves structurally informative vehicle points while removing adversarial points. To integrate this filtering mechanism into a practical 3D detection pipeline, we further introduce two auxiliary components: *Attack ROI Localization* to restrict filtering to suspicious regions and reduce unnecessary modification; and *Back-projection Replacement* to reconstruct the filtered point cloud.

3.1 Depth Priors-informed Filtering

This module is the core of DPP and is directly motivated by the observations and opportunities described in §2. It consists of two complementary steps.

Algorithm 1 Depth gradient-based filtering

Require: Range image R, threshold τ_{abs}
Ensure: Candidate mask E
$E \leftarrow 0$
for each valid pixel (i, j) with $R(i, j) > 0$ **do**
  $g \leftarrow 0$
  for each valid 4-neighbor (i', j') of (i, j) **do**
    $d \leftarrow |R(i, j) - R(i', j')|$
    $g \leftarrow \max(g, d)$
  end for
  if $g > \tau_{\text{abs}}$ **then**
    $E(i, j) \leftarrow 1$
  end if
end for
return E

3.1.1 Depth gradient-based filtering. Motivated by Observation 1 and Opportunity 1, we use depth gradient to distinguish structurally informative vehicle points from adversarial ones. Given a suspicious region that is identified by the approach to be described in §3.2, we project its point cloud onto a range image $R \in \mathbb{R}^{H \times W}$. Each valid pixel stores the depth of its nearest LiDAR return, and empty pixels are set to zero. Then, we compute the depth gradient for each valid pixel by comparing it with its four neighboring pixels. This stage of filtering retains pixels whose maximum depth difference with neighbors exceeds a threshold τ_{abs}. The procedure is summarized in Algorithm 1. In our experiments, τ_{abs} is set to 0.2 m.

3.1.2 Scanline consistency-based filtering. While the previous filtering step removes most object-based adversarial points, it may still retain sparse or isolated points injected by laser-based attacks. Motivated by Observation 2, we further enforce scanline-level structural consistency. For each row in the range image, we examine the sequence of valid depth values. A row is preserved only if it contains a sufficiently long segment (length $\geq L$) that is monotonic or approximately constant (within tolerance ϵ). Otherwise, the entire row in the candidate mask is discarded. This step effectively removes isolated injected points while preserving coherent object structures. The procedure is described in Algorithm 2. In our implementation, $L = 10$ and $\epsilon = 0.05$ m.

Algorithm 2 Scanline consistency-based filtering

Require: Range image R, candidate mask E
Require: minimum length L, tolerance ϵ
Ensure: Final mask $\tilde{E}$
$\tilde{E} \leftarrow 0$
for each row i **do**
  Extract valid depths $\{r_1, \ldots, r_n\}$ from row i of R
  if there exists a run of at least L pixels that is monotonic or satisfies $|r_k - r_{k-1}| < \epsilon$ **then**
    $\tilde{E}(i, :) \leftarrow E(i, :)$
  end if
end for
return $\tilde{E}$

3.2 Attack ROI Localization

To avoid unnecessary modification of the entire scene and reduce potential performance degradation on benign regions, we restrict the filtering process to suspicious areas, referred to as region of interest (ROI). Following [23], we apply the DBSCAN clustering algorithm on the input point cloud and obtain a set of clusters before the filter. Then, we compare these clusters with the predicted bounding boxes from the current detector. Clusters that do not match any detection are regarded as suspicious and defined as attack ROIs. DPP is applied only within these ROIs.

3.3 Back-projection Replacement

After filtering, we obtain purified pixels in the range image. These pixels are back-projected to 3D LiDAR coordinates to recover the corresponding points. Then, we replace the original points within each suspicious ROI with the filtered points to construct a new point cloud frame. This updated frame is finally fed into the 3D object detector to produce the detection results.

4 Evaluation

4.1 Evaluation Setup

We evaluate DPP in both trace-driven simulations and real-world experiments. This section describes the object detector, attacks, baseline defense approaches, dataset, and evaluation metrics.

Object detection model: We adopt PointPillars [11] as the base 3D object detector. The confidence threshold for the vehicle class is set to 0.3, which is a common setting.

Attacks: We evaluate the three attacks introduced in §2, i.e., OBH [28], OMH [17], and LH [10] attacks, which capture the main attack strategies considered in literature.

Baselines: We compare DPP with two representative defense methods. SRS [25] performs random point dropping, where we set the drop rate to 50%. Hyper3Def [19] adopts an ensemble strategy, and we use an ensemble size of 4 as suggested in the original work.

Dataset. We conduct evaluation on the KITTI 3D object detection dataset [7]. We randomly select 400 frames containing vehicles approaching from the opposite direction, which serve as target vehicles for mounting attacks. We add adversarial points planned by the three attacks to the clean point cloud frames.

Evaluation metrics: We evaluate attack and defense performance based on the bird's-eye-view IoU (IoU_{BEV}). A hiding attack is considered successful if no detected bounding box attains an IoU_{BEV} value greater than a predefined threshold η, with respect to the ground-truth bounding box of the target vehicle. We set η to 0.1, by following the common setting in 3D object detection performance evaluation. We further define *defense success rate* (DSR) as the ratio of the number of frames in which a previously successful attack becomes unsuccessful after defense to the total number of frames in which the attack succeeds before defense.

4.2 Trace-driven Simulations

As shown in Table 3, DPP consistently achieves strong defense performance across the three attacks, with DSR over 88%. Compared with lightweight defense SRS, DPP achieves more than 51.2% DSR across all the attacks. Compared with Hyper3Def, DPP is more effective in counteracting laser-based attack LH, which can be attributed to the scanline consistency filtering that explicitly suppresses irregular injected points. In counteracting object-based attacks (i.e., OBH and OMH), DPP achieves competitive but slightly lower performance than Hyper3Def. This is because DPP focuses on preserving geometrically informative points, and some adversarial points that partially satisfy geometric constraints may still be retained. While not always the best in terms of performance, DPP has significantly shorter latency, as presented in §4.4 shortly.

4.3 Real-World Experiments

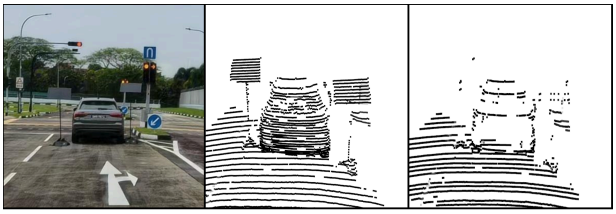
In the real-world experiments, a SUV shown in Figure 2 serves as the victim vehicle and another crossover SUV shown in Figure 3(a) as the target vehicle. We compute the placement positions of the cardboards generated by the OBH attack and place them around the target vehicle, as shown in Figure 3. An OS1-128 LiDAR is mounted on the victim vehicle, which moves toward the target vehicle.

We select 10 frames in which the attack successfully hides the target vehicle to evaluate the defense performance of DPP. The results show that DPP achieves a DSR of 80% in this scene. Figure 3(c) shows an output of DPP. The defense failures are caused by the long distances between the victim vehicle and the target vehicle so that the remained vehicle points are not enough to form a complete structure. For comparison, SRS and Hyper3Def achieve DSRs of 0% and 90%, respectively. We also conduct a simulation of OMH on the collected clean frames by our LiDAR. Results show that DPP achieves a DSR of 100% against OMH.

Table 3: DSR (%) comparison on KITTI.

Attack	DPP	SRS	Hyper3Def
OBH	89.65	2.15	99.65
OMH	88.28	29.72	89.70
LH	94.55	43.35	72.06

Figure 2: Victim vehicle used in physical experiments.



(a) Attack setup **(b) Attack points** **(c) DPP output**

Figure 3: DPP defense against Object-Board Hiding attack.

4.4 Compute Time

We measure the per-frame compute time on a workstation equipped with an AMD Ryzen 9950X CPU and an NVIDIA RTX 4090 GPU, as well as on an NVIDIA Jetson Orin AGX 32GB embedded platform with an 8-core ARM Cortex-A78AE CPU. SRS and DPP are implemented using only the CPU, whereas RL-remove [22] and Hyper3Def [19] rely on both CPU and GPU. The reported runtime includes only the stages after ROI localization, since ROI localization is shared by all methods.

As shown in Table 4, DPP requires only about 10 ms per frame on the Orin CPU, compared with 167 ms for Hyper3Def and 7623 ms for RL-remove on the same platform. This large gap highlights the practicality of our method for embedded deployment. For Surface-remove [23], whose code is not yet publicly available, we estimate its runtime by scaling with the relative GPU computational capability. The estimated runtime of Surface-remove on the workstation is approximately 50 ms, while our method requires only 2 ms. DPP achieves a much more favorable balance between defense effectiveness and computational cost than existing methods.

Table 4: Per-frame defense runtime (ms).

Method	Workstation	Orin
SRS (CPU only)	0.56	1.18
RL-remove (CPU+GPU)	2523.12	7623.13
Hyper3Def (CPU+GPU)	20.43	167.95
DPP (CPU only)	2.23	10.23

In general, DPP can achieve strong defense performance across multiple vehicle hiding attacks with short compute times, which makes it suitable for real-world mobile deployments.

5 Conclusion and Future Work

In this paper, we present DPP, a lightweight and model-agnostic data purification module for improving the robustness of LiDAR-based object detection systems against adversarial example attacks targeting vehicle objects. DPP uses two complementary geometric cues, depth gradient and scanline consistency, to effectively distinguish adversarial points from genuine vehicle surfaces in the range image. Its low computational overhead and plug-and-play design make it suitable for resource-constrained deployment scenarios.

Several directions remain open for further exploration. (1) Future work can further improve robustness against adaptive attacks, where the adversary has access to the design and thresholds of DPP. In such scenarios, attackers may construct structured objects and carefully control their spatial layout allowing some adversarial points to bypass filtering. (2) It is important to handle long-range sparse observations. Although DPP preserves the most informative vehicle points, LiDAR measurements naturally become sparse at longer distances, which may affect detection reliability after filtering. Incorporating point completion techniques for distant objects could help alleviate this issue. However, such methods need to be carefully designed to avoid reintroducing adversarial points.

Acknowledgments

This research is supported in part by the National Research Foundation, Singapore under its AI Singapore Programme (AISG Award No: AISG4-GC-2023-006-1B), in part by a grant from City University of Hong Kong (Project No. 9610753), and in part by JC STEM Lab of Future Energy Systems (20250039).

References

[1] 2025. VueOne. https://www.vueron.com/vueone/.
[2] 2025. Waymo Driver. https://waymo.com/waymo-driver/.
[3] Mazen Abdelfattah, Kaiwen Yuan, Z Jane Wang, and Rabab Ward. 2021. Towards universal physical attacks on cascaded camera-lidar 3d object detection models. In *IEEE International Conference on Image Processing (ICIP)*.
[4] Mumuxin Cai, Xupeng Wang, Ferdous Sohel, and Hang Lei. 2025. LiDAR-SPD: Improving adversarial robustness of 3D object detection via spherical projection and diffusion. In *ICASSP 2025-2025 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 1–5.
[5] Yulong Cao, Ningfei Wang, Chaowei Xiao, Dawei Yang, Jin Fang, Ruigang Yang, Qi Alfred Chen, Mingyan Liu, and Bo Li. 2021. Invisible for both camera and lidar: Security of multi-sensor fusion based perception in autonomous driving under physical-world attacks. In *IEEE Symposium on Security and Privacy (SP)*.
[6] Yulong Cao, Chaowei Xiao, Benjamin Cyr, Yimeng Zhou, Won Park, Sara Rampazzi, Qi Alfred Chen, Kevin Fu, and Z Morley Mao. 2019. Adversarial sensor attack on lidar-based perception in autonomous driving. In *ACM SIGSAC Conference on Computer and Communications Security (CCS)*.
[7] Andreas Geiger, Philip Lenz, Christoph Stiller, and Raquel Urtasun. 2015. The kitti vision benchmark suite. *URL http://www. cvlibs. net/datasets/kitti* 2, 5 (2015), 1–13.
[8] Ian J Goodfellow, Jonathon Shlens, and Christian Szegedy. 2015. Explaining and harnessing adversarial examples. In *International Conference on Learning Representations (ICLR)*.
[9] Zhongyuan Hau, Kenneth T Co, Soteris Demetriou, and Emil C Lupu. 2021. Object removal attacks on lidar-based 3d object detectors. In *Automotive and Autonomous Vehicle Security Workshop (AutoSec)*.
[10] Zizhi Jin, Xiaoyu Ji, Yushi Cheng, Bo Yang, Chen Yan, and Wenyuan Xu. 2023. Pla-lidar: Physical laser attacks against lidar-based 3d object detection in autonomous vehicle. In *IEEE Symposium on Security and Privacy (SP)*.
[11] Alex H Lang, Sourabh Vora, Holger Caesar, Lubing Zhou, Jiong Yang, and Oscar Beijbom. 2019. Pointpillars: Fast encoders for object detection from point clouds. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*.
[12] Daizong Liu and Wei Hu. 2022. Imperceptible transfer attack and defense on 3d point cloud classification. *IEEE transactions on pattern analysis and machine intelligence* 45, 4 (2022), 4727–4746.
[13] Daniel Liu, Ronald Yu, and Hao Su. 2019. Extending adversarial attacks and defenses to deep 3d point cloud classifiers. In *IEEE International Conference on Image Processing (ICIP)*.
[14] Yang Lou, Yi Zhu, Qun Song, Rui Tan, Chunming Qiao, Wei-Bin Lee, and Jianping Wang. 2024. A First {Physical-World} Trajectory Prediction Attack via {LiDAR-induced} Deceptions in Autonomous Driving. In *33rd USENIX Security Symposium (USENIX Security 24)*. 6291–6308.
[15] Jiachen Sun, Yulong Cao, Qi Alfred Chen, and Z Morley Mao. 2020. Towards robust LiDAR-based perception in autonomous driving: General black-box adversarial sensor attack and countermeasures. In *USENIX Security Symposium*.
[16] James Tu, Huichen Li, Xinchen Yan, Mengye Ren, Yun Chen, Ming Liang, Eilyan Bitar, Ersin Yumer, and Raquel Urtasun. 2022. Exploring adversarial robustness of multi-sensor perception systems in self driving. In *Conference on Robot Learning (CoRL)*.
[17] James Tu, Mengye Ren, Sivabalan Manivasagam, Ming Liang, Bin Yang, Richard Du, Frank Cheng, and Raquel Urtasun. 2020. Physically realizable adversarial examples for lidar object detection. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*.
[18] Chong Xiang, Charles R Qi, and Bo Li. 2019. Generating 3d adversarial point clouds. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*.
[19] Yihan Xu, Dongfang Guo, Qun Song, Yang Lou, Yi Zhu, Jianping Wang, Chunming Qiao, and Rui Tan. 2025. Dynamic Defense for Car-Borne LiDAR Vehicle Detection. In *Proceedings of the 23rd Annual International Conference on Mobile Systems, Applications and Services (Mobisys)*. 431–444.
[20] Kaichen Yang, Tzungyu Tsai, Honggang Yu, Max Panoff, Tsung-Yi Ho, and Yier Jin. 2021. Robust roadside physical adversarial attack against deep learning in lidar perception modules. In *ACM Asia Conference on Computer and Communications Security (AsiaCCS)*.
[21] Jinlai Zhang, Lyujie Chen, Binbin Liu, Bo Ouyang, Qizhi Xie, Jihong Zhu, Weiming Li, and Yanmei Meng. 2023. 3d adversarial attacks beyond point cloud. *Information Sciences* 633 (2023), 491–503.
[22] Yan Zhang, Zihao Liu, Chongliu Jia, Yi Zhu, and Chenglin Miao. 2024. An online defense against object-based lidar attacks in autonomous driving. In *Proceedings of the 22nd ACM Conference on Embedded Networked Sensor Systems (Sensys)*. 380–393.
[23] Yan Zhang, Zihao Liu, Yi Zhu, and Chenglin Miao. 2025. Towards Real-Time Defense against Object-Based LiDAR Attacks in Autonomous Driving. In *Proceedings of the 2025 ACM SIGSAC Conference on Computer and Communications Security (CCS)*. 3825–3839.
[24] Tianhang Zheng, Changyou Chen, Junsong Yuan, Bo Li, and Kui Ren. 2019. Pointcloud saliency maps. In *IEEE/CVF International Conference on Computer Vision (ICCV)*.
[25] Hang Zhou, Kejiang Chen, Weiming Zhang, Han Fang, Wenbo Zhou, and Nenghai Yu. 2019. DUP-Net: Denoiser and Upsampler Network for 3D Adversarial Point Clouds Defense. In *2019 IEEE/CVF International Conference on Computer Vision (ICCV)*.
[26] Shenchen Zhu, Yue Zhao, Kai Chen, Bo Wang, Hualong Ma, et al. 2024. AE-Morpher: Improve Physical Robustness of Adversarial Objects against {LiDAR-based} Detectors via Object Reconstruction. In *33rd USENIX Security Symposium (USENIX Security 24)*. 7339–7356.
[27] Yi Zhu, Chenglin Miao, Foad Hajiaghajani, Mengdi Huai, Lu Su, and Chunming Qiao. 2021. Adversarial attacks against lidar semantic segmentation in autonomous driving. In *ACM Conference on Embedded Networked Sensor Systems (SenSys)*.
[28] Yi Zhu, Chenglin Miao, Tianhang Zheng, Foad Hajiaghajani, Lu Su, and Chunming Qiao. 2021. Can we use arbitrary objects to attack lidar perception in autonomous driving?. In *ACM SIGSAC Conference on Computer and Communications Security (CCS)*.