Comprehensive Robustness Analysis of LiDAR-based 3D Object Detection in Autonomous Driving

Adwait Chandorkar[1], Kai Krink[1*], Yerdana Maulenbay[1,2*,†], Hasan Tercan[1], and Tobias Meisen[1]

[1] Institute for TMDT, University of Wuppertal, Germany
{chandorkar, kai.krink, tercan, meisen}@uni-wuppertal.de
[2] IKB Faculty of Science, University of British Columbia, Canada
yerdana@student.ubc.ca

Abstract. Recent advancements in LiDAR-only 3D object detection have demonstrated improved detection accuracy over benchmark datasets. However, the adversarial robustness of these models remains untested. Very few adversarial robustness studies exist for LiDAR-only 3D object detection and unfortunately, even they are limited to legacy models. Moreover, there is a systemic gap in the existing evaluation frameworks that rely simply on mAP ignoring other structural and predictive factors. To fill this gap, we propose a holistic framework that evaluates adversarial robustness using two structural factors (point cloud density and point cloud localization) and three predictive factors (misclassification, localization error, distance from ego). Using this framework, we perform an empirical study and critical analysis on recent and legacy state-of-the-art models using adversarial attacks specifically designed for LiDAR-based models. Our key finding is that high-capacity, voxel-based detectors are more susceptible to structured coordinate perturbations than pillar-based detectors. Additionally, non-anchor-based detectors demonstrate poor adversarial robustness, which necessitates rethinking model training techniques. Overall, our results demonstrate that recent models are as vulnerable to adversarial attacks as their predecessors. Therefore, we argue that there is a need to improve the evaluation benchmarks for 3D object detection that not only reward architectural modifications for improving detection accuracy, but also evaluate whether the design choices improve adversarial robustness.

Keywords: Adversarial attacks · 3D object detection · Robustness analysis

1 Introduction

Perception has long been one of the most critical tasks in autonomous driving systems. Over the years, we have seen significant progress in perception systems,

* Equal contribution
† The author contributed to this work as part of the DAAD RISE program.

both in standalone sensor systems (*e.g.* LiDAR-only or camera-only) and in multi-sensor systems. That said, LiDAR-based systems remain the preferred choice, thanks to the robustness of the sensor and its ability to efficiently capture depth information [3, 34, 47]. From the perspective of representation learning, LiDAR-only object detection (LiDAR-OD) models can be broadly classified into point-based methods [32,45], grid-based methods [5,8,9,20,21,27,29,43,51], and hybrid methods [30,31,33,39] based on how the raw point clouds are processed. Grid-based methods, mainly, have outperformed other methods on demanding benchmark datasets such as KITTI [16], nuScenes [4] and Waymo [35]. These methods have not only demonstrated higher detection accuracy but also achieved higher inference speeds demonstrating significant progress toward the realization of fully autonomous driving in real-world environments.

Nevertheless, as previous works [19, 34, 52] have reported, model improvements often deliver strong results on clean datasets, but their robustness remains significantly compromised when subjected to corrupt or out-of-distribution (OoD) examples. One way of evaluating this is to test the robustness of these models against adversarial attacks. Robustness analysis using adversarial attacks was initially rooted in the 2D image domain [17, 26]. However, recently there has been increased research in the 3D domain [6, 7, 36, 38, 40, 53]. Although most of these works have targeted the exploitation of vulnerabilities in existing 3D-OD and measured the resulting performance degradation, the current literature fails to investigate whether recent architectures [8, 21, 27, 29] have actually improved robustness or if they inherit similar structural vulnerabilities from their predecessors. In addition, we identify three more gaps in existing studies: *first*, an over-reliance on benchmarking legacy 3D-OD models; *second*, an evaluation scope restricted almost exclusively to Average Precision (AP), neglecting critical localization errors (*e.g.*, translation, scale) and factors like point cloud density or proximity to the ego vehicle; and *third*, a dependence on classical adversarial attacks such as FGSM [17] and PGD [25] rather than attacks designed for LiDAR-based 3D-OD.

In this work, we address these shortcomings by presenting a comprehensive benchmark suite to evaluate the adversarial robustness of 3D-OD across commonly benchmarked datasets such as nuScenes and Waymo. We hypothesize that, first, *SoTA 3D-OD architectures exhibit a fundamental lack of resilience against coordinated PC perturbations* and second, *a holistic evaluation requires decoupling structural robustness (geometric precision) from predictive robustness (detection confidence)*. To our knowledge, this represents the first exhaustive study of its kind for LiDAR-based 3D-OD. We evaluate different white-box and black-box attacks, designed specifically for LiDAR-based 3D-OD, to expose inherent model weaknesses. Moving beyond standard metrics like Attack Success Rate (ASR) or mAP degradation, we propose five distinct benchmarks to quantify the impact across point cloud (PC) density, PC localization, object classification, object localization, and the object's proximity from the ego vehicle. Our study seeks to establish whether the architectural improvements in latest SoTA

3D-OD yield improved adversarial robustness or the models remain vulnerable like previous architectures. Our main contributions are summarized as follows:

- We propose a comprehensive adversarial robustness analysis of LiDAR-based 3D-OD models by benchmarking our results on two *latest* 3D-OD and comparing their results with legacy models.
- Our study analyzes both the structural robustness, the impact of PC modifications, and the predictive robustness of the baselines and addresses the gaps that arise from architectural design choices.
- We identify the main reason for this adversarial susceptibility to *Representation Fragility* in the existing datasets: the models prioritize the memorization of specific geometric templates and canonical point-cloud densities prevalent within the training distribution, rather than the learning geometric inconsistencies. This approach yields superior performance on clean datasets but collapses under structured noise.

2 Related Work

2.1 Advancements in grid-based 3D object detection

The grid-based architecture comprises an *encoder* that transforms sparse point clouds into either voxels or pillar-based pseudo-images, followed by a *backbone* for multi-level feature extraction, a *neck* for feature fusion, and a *head* for regressing 3D bounding boxes and class probabilities. Voxel-based methods [9, 43, 51] partition the point clouds into 3D voxels using either 3D CNNs as proposed by Zhou *et al.* in VoxelNet [51] or 3D SpConv proposed by Yan *et al.* in SECOND [43]. A recent voxel-to-object method VoxelNext [9], introduced by Chen *et al.*, proposes direct end-to-end 3D-OD from voxel features, bypassing the need for post-processing. Despite advancements in reducing computational demands, voxel-based detectors still entail high computational costs and lower inference speeds. On the other hand, pillar-based methods provide a more efficient balance between inference speed and detection accuracy. Building on PointPillars [20], PillarNext [21] and PillarNest [27] replace pure max pooling with combined max–average pooling to mitigate feature loss. They further substitute conventional CNN backbones with sparse convolutions (SpConv), improving computational efficiency [21, 29]. To bypass the deployment challenges of SpConvs — specifically irregular memory access and quantization — PillarNest [27] utilizes a ConvNext [22] backbone pre-trained on images and scaled for point clouds, while FastPillars [50] implements a re-parameterized ResNet-34 for efficient scaling. Although the *encoder* varies significantly for voxel and pillar-based methods, the *head* is mostly similar. Initially, grid-based methods [20, 43, 51] relied on computationally heavy anchor-based detectors such as RPN or SSD. Subsequently, CenterPoint [46] proposed an anchor-free design, *CenterHead*, which detects objects through center point estimation. More recently, transformer-based heads (*e.g.*, FocalFormer3D [8], PillarTrack [42]) have achieved superior performance on nuScenes [4] and Waymo [35]. In this study, we utilize 3D-OD models with

enhanced *encoder* and *decoder* structures as baselines, with further justification provided in Sec. 3.1. On these baselines, we evaluate whether the enhancements have improved adversarial robustness or if it is similar to their predecessors.

2.2 Adversarial attacks on 3d object detection

While classical gradient-based attacks like FGSM and PGD provide a baseline for adversarial robustness, they fail to account for the sparsity, sensor physics, and non-differentiable pre-processing (*e.g.*, voxelization and pillarization) inherent in LiDAR-based 3D-OD pipelines. Classical methods often produce *point-wise noise* that is filtered out by the structural encoding in the *encoder*. In contrast, Abraha *et al.* [1] introduce a gradient-free attack that exploits detector knowledge via a ConvNet-based perturbation generator. Chen *et al.* [6] propose a white-box method optimizing targeted perturbations through IoU minimization. Chen *et al.* [7] propose a black-box genetic simulated annealing (GSA) attack that generates adversarial objects camouflaged within the scene. Furthermore, Long *et al.* [23] introduced a white-box non-end-to-end attack that manipulates feature saliency by restraining critical detection features while amplifying non-contributory features to diminish accuracy with reduced reliance on specific network architectures. We choose attacks that are specifically designed for LiDAR-based 3D-OD as our baseline and provide additional reasoning in Sec. 3.2.

2.3 Robustness analysis of 3d object detection

Zhang *et al.* [48] provided the first systematic analysis of LiDAR-only architectures, assessing attack efficacy, transferability, and defense performance. Further, Dong *et al.* [11] benchmarked various modalities against 27 physically realizable environmental corruptions. Recently, Eskandar [14] investigated domain adaptation factors that affect robustness, such as voxel encoding, resolution, and data augmentation. Despite these advancements, current adversarial robustness analysis often suffers from two critical limitations: the use of legacy adversarial attacks instead of recent methods and the lack of benchmarking on the latest SoTA 3D-OD models for evaluation. Our work overcomes these limitations and identifies the architectural choices that make modern SoTA 3D-OD vulnerable to adversarial attacks.

3 Methodology

We design our study based on three foundations: 1) we use *four* baseline models (*two* latest 3D-OD models and *two* classical 3D-OD models) for improvement comparison, as explained in Sec. 3.1; 2) we use *five* latest adversarial attacks that have demonstrated improved performance over classical attacks, as explained in Sec. 3.2, and 3) we propose a comprehensive benchmark involving *five* criteria to study the impact on our baselines, as explained in Sec. 3.3. Fig. 1 demonstrates

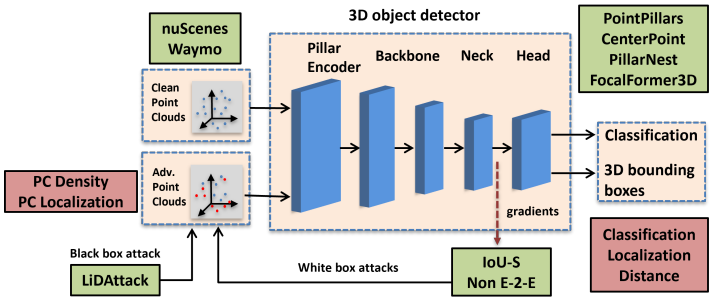


Fig. 1: Methodology

Table 1: Taxonomy of the models and datasets used in our study. For the models where the original implementation (O) was available, we use the pre-trained weights. For others, we train (T) them from scratch.

Model	Encoder	Head	nuScenes		Waymo	
			Impl.	mAP ↑	Impl.	mAP ↑
PointPillars [20]	Pillar	RPN	T	35.9	T	58.4
PillarNeSt [27]	Pillar	CenterHead	O	66.9	T	66.2
CenterPoint [46]	Voxel	CenterHead	O	58.0	O	67.8
FocalFormer3D [8]	Voxel	Transformer	O	70.5	O	71.9

our entire setup. Our investigation is conducted on two standard benchmarks: nuScenes and Waymo. For baseline models that were originally not trained on these datasets, we train the models from scratch and report the resultant mean Average Precision (mAP) in Tab. 1 to ensure performance parity across all baselines. These initial metrics establish an upper-bound performance reference under clean point cloud conditions. We subsequently expose all baselines to four white-box and one black-box adversarial attack to generate perturbed inputs. Model robustness is then evaluated along two complementary dimensions: point cloud (PC) density and PC localization capture *structural robustness*, whereas classification, bounding-box localization, and distance-based metrics quantify degradation in *predictive performance*.

3.1 Baseline 3d object detectors

Our experimental design is based on a strategic selection of baselines that serve as critical testbeds for our two hypotheses. To evaluate first hypothesis, the fundamental vulnerability of SoTA 3D-OD to coordinated PC perturbations, we incorporate PillarNeSt [27] and FocalFormer3D [8]. As an evolution of the pillar-based paradigm, PillarNeSt demonstrates a strong feature extraction efficiency, utilizing a ConvNeXt backbone and a dual-path (max and average) pooling strategy to minimize information entropy in *encoder*. It also outperforms other recent pillar-based OD, such as PillarNet [29], PillarNext [21], and FastPillars [50] on

nuScenes dataset. By selecting the leading model of this lineage, we aim to determine if enhanced *encoder* architecture increases resilience to PC perturbations. FocalFormer3D represents a shift toward transformer-based detection *head*, specifically designed to mitigate false negatives (FN) through a Hard Instance Probing (HIP) pipeline. FocalFormer3D has significantly outperformed other transformer-based OD such as OcTr [49], Li3DeTr [13] and FusionViT [41] on both nuScenes and Waymo. This selection allows us to investigate whether architectural improvements in *head* improve geometric precision or merely mask inherent structural vulnerabilities. Finally, we compare the robustness of these models with CenterPoint [46] and PointPillars [20]. These models provide the necessary contrast between modern and legacy models.

3.2 Baseline adversarial attacks

For the white-box setting, we first use the IoU-S [6] attack. The central idea is to target the localization capability. A 3D-OD predicts an object class along with 3D bounding box co-ordinates, i.e. localize the objects. Based on the predicted class probability and the IoU-score, a *confidence* score is generated using the formula confidence $= \sqrt{P_{\text{class}} \times \text{IoU}_{\text{score}}}$. The IoU-S attack learns to generate adversarial point clouds that minimize this confidence of the attacked model. We use all the three attack paradigms proposed by IoU-S - attachment (PA), detachment (PD), and perturbation (PB). The attachment attack injects new points into the existing point cloud, the detachment attack drops points, while the perturbation attack shifts the coordinates of the raw point clouds. Then, we use the Non End-to-End (NE) [23] as the second white-box attack. The method overcomes the limitations of existing white-box methods: dependence on network structure and the resulting inefficient transferability of attacks. Additionally, it proposes an adversarial loss function on the feature space that restrains the features with a high contribution towards the detection task and simultaneously amplify the features with a lower contribution. Finally, we use LiDAttack (LiD) [7] for the black-box setting. It is a novel approach that uses genetic simulated annealing (GSA) to generate a stealthy and concealable attack by strictly constraining the distance between the perturbed points.

An inherent challenge with adversarial attacks is that they are not architecture-agnostic and possess bias against certain design choices. To address this bias, we ensure that our attack selection remains fair and there exists at least one attack that is biased towards each of our baseline models. The PA attack affects the depth information in the point cloud (PC) and is hence particularly effective against pillar-based models. The PB attack alters the structural representation, making them effective against voxel-based models whereas the PD attack alters the density of PC representations, effectively posing higher challenges for heatmap-based detectors such as CenterPoint. The NE attack targets the intermediate feature maps rather than the PCs and thus disrupts the spatial continuity. This makes the attack especially stronger against heatmap-based detectors that rely on local maximum (*peak*) of a predicted heatmap. Finally, the black-box LiD attack is a more generic attack that is not strongly biased towards

any model. However, we hypothesize that the *max-pooling* used in pillar *encoder* makes them more vulnerable to LiD attack. The detailed formula, methodology and budget of these attacks is attached in the Appendix.

3.3 Evaluation Criteria

Impact on Classification. This is evaluated based on two factors: firstly, classical misclassification, and secondly, the impact on confidence. As highlighted by Wang *et al.* [37], unstable confidence scores result in flickering predictions thereby impacting detection and tracking performance of 3D-OD. Hence, we assess the average drop in confidence along with average drop in mAP after subjecting the model to an adversarial attack. We also evaluate the models based on change in false positives (FP) and false negatives (FN). Our goal is to evaluate whether models are especially vulnerable to certain classes. Hence, following the works of [12, 15, 28], we first categorize Cyclists, Pedestrians and Motorcyclists as **Safety Critical** classes, since they are most involved in accidents and carry a high chance of human fatality. Then, we report the impact on *Cars* and *Safety Critical* classes in the Sec. 4.1.

Impact on Localization. A high IoU does not always indicate accurate localization. The bounding box could be either shifted, incorrectly scaled, or have a shifted orientation. While evaluating the performance of 3D-OD, these factors are addressed to some extent in nuScenes and Waymo. However, they are not uniform across all datasets. To address this, we calculate the *translation*, *scale*, and *orientation* errors based on the predicted bounding boxes for all the models. To the best of our knowledge, our work is the first to factor in bounding box localization errors to evaluate the effectiveness of adversarial attacks.

Impact on object distance from ego. An inherent problem with LiDAR is that the PC of objects near the ego-vehicle are denser than that of objects which are further away. Therefore, the detection accuracy of near-ego objects is naturally higher than that of far-away objects. By assessing the impact of attacks relative to the object's distance from the ego vehicle, we can evaluate whether the models are robust enough to detect *near-ego* objects.

Impact of Point Cloud Density. The adversarial perturbations generated by our baseline, especially attachment (PA) and detachment (PD) attacks, inherently alter the spatial distribution of the input, directly modifying the PC density. From a conventional perspective, 3D-OD performance is expected to improve with a high PC density and degrade under conditions of high sparsity. By systematically evaluating model efficacy against varying input densities, we investigate whether this heuristic holds under adversarial conditions or if the induced structural modifications expose vulnerabilities.

Impact of Point Cloud localization. In order to assess the maximum impact of the modifications to the PC, we split the PCs within the ground truth box into two categories before generating the adversarial PC. As shown in Fig. 2, we create a secondary 3D bounding box within the ground truth, downscaled by an empirically determined factor based on each dataset. The scaling threshold for the inner bounding boxes is set to *0.8* and the empirical study to determine these thresholds is presented in Sec. 4.6. We classify the points outside this scaled box as *outer* points and the ones within the box as *inner* points. As observed in the figure, the *outer* points generally represent the object's structural contour and extremities (edges and corners), whereas the *inner* points represent the volumetric density of the object. We then evaluate the impact on model performance when the attacks affect the inner and outer points.

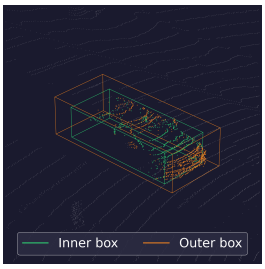


Fig. 2: Visualization of inner and outer PCs for a *Car* in nuScenes.

4 Results

Numerous abbreviations are used throughout the following sections. To provide a single reference point, we provide a brief overview of these abbreviations here. Model abbreviations - **CP** : CenterPoint, **FF**: FocalFormer3D, **PN**: PillarNest, **PP**: PointPillars. The attack abbreviations - **PA**: Attachment, **PD**: Detachment, **PB**: Perturbation, **LiD**: LiDAttack and **NE**: Non End-2-End. We report the success of attacks using the attack success ratio (ASR). However, our method of calculating ASR is different from the method used by current approaches. We observed that existing studies report attacks as successful only if they induce misclassification but not if there is a significant drop in confidence albeit with a correct classification. Secondly, ASR calculation involves all missed detections even though they were missed by the 3D-OD on clean dataset. We argue, this results in inflated ASR and does not factor in genuine impact. To overcome these challenges, the first distinction is that we exclude the objects that were undetected by the model on the clean point cloud from the calculation. Second, we consider an attack successful if it induces misclassification or if the prediction is correct but confidence < 0.15. Otherwise, the attack is unsuccessful. In the subsequent sections, we present the most relevant results and visualizations, while additional supporting visualizations can be found in the Appendix.

Experimental Setup. All models were trained and evaluated in PyTorch [2] using the mmdetection3d [10] codebase. Additional details associated with adversarial training, model training and configurations are available in the Appendix.

4.1 Impact on Classification

Tab. 2 lists the main results of our experiments. Our evaluation across the nuScenes and Waymo datasets reveals a bifurcated landscape of vulnerability

Table 2: Comprehensive Adversarial Robustness Evaluation: Attack results across nuScenes and Waymo datasets. The values represent relative AP drop / relative confidence drop across different object classes and attack types. Negative values indicate improvement.

(a) nuScenes

Model	Car					Safety Critical				
	LiD (%)	PA (%)	PB (%)	PD (%)	NE (%)	LiD (%)	PA (%)	PB (%)	PD (%)	NE (%)
CP	-0.1/-0.3	3.9/6.2	65.5/41.8	15.1/2.5	16.6/25.7	0.9/-0.8	21.9/9.9	85.8/15.5	31.7/2.9	41.9/7.7
FF	0.3/-2.0	7.8/17.7	38.9/59.2	20.6/2.5	17.5/37.1	2.6/-3.4	31.8/21.5	81.5/62.0	34.8/4.8	30.1/14.3
PN	0.0/-0.2	32.2/4.6	20.2/6.4	24.6/1.3	8.0/12.6	0.4/-0.4	54.8/6.4	38.9/10.1	41.1/1.1	19.4/3.7
PP	0.0/-0.1	85.4/42.2	10.0/18.7	25.9/5.9	8.6/19.3	1.0/-1.0	36.5/9.2	23.9/8.4	17.7/0.3	26.1/4.1

(b) Waymo

Model	Car					Safety Critical				
	LiD (%)	PA (%)	PB (%)	PD (%)	NE (%)	LiD (%)	PA (%)	PB (%)	PD (%)	NE (%)
CP	0.0/-0.0	0.0/14.8	91.0/25.6	47.8/21.3	82.0/21.6	0.0/-0.1	25.3/17.6	96.3/28.9	71.3/13.8	81.0/20.1
FF	0.2/-0.0	8.0/4.5	88.5/11.5	7.8/0.3	82.1/9.7	0.1/-0.2	19.4/19.2	61.1/12.5	24.5/7.6	56.7/8.3
PN	1.6/-0.3	24.0/6.2	18.6/7.3	2.3/-0.7	12.8/5.4	1.3/-0.3	24.5/19.7	13.8/24.7	18.8/6.1	11.1/21.3
PP	0.6/-0.2	57.5/13.2	9.4/7.4	17.7/-0.5	7.1/6.5	1.5/-0.2	45.1/13.7	15.8/18.6	12.1/3.2	12.2/16.8

between pillar-based models, PP and PN, and high-capacity detectors, CP and FF. Our key findings are:

Fragility of high-capacity detectors to geometric perturbations. The sensitivity of both legacy (CP) and contemporary (FF) architectures to subtle geometric coordinate shifts, induced by PB and NE attacks, substantiates our hypothesis that high-capacity detectors over-rely on canonical spatial distributions and fine-grained local features to achieve peak performance. This dependency creates a critical representational bottleneck, whereby models prioritize the memorization of specific PC representations over the learning of robust volumetric invariants. This vulnerability is particularly acute within safety-critical classes, where inherent semantic sparsity and structural rigidity make detectors highly sensitive to coordinate shifts.

Sensitivity of pillar-based detectors to feature injection. The pillar-based models exhibit strong robustness against perturbation attacks, but the coarser spatial quantization in pillar encoding makes them susceptible to point injections. The models demonstrate high sensitivity against PA attacks and while the failure of PP is anticipated due to its simplified encoder, the parallel degradation of the more sophisticated PN architecture suggests that the underlying pillar-based abstraction, inherited from its predecessor—constitutes a fundamental and shared structural vulnerability. The poor performance on safety-critical classes is attributed to the low geometric complexity of these objects. The pedestrian and cyclist classes occupy few pillars after encoding and have distinct vertical signatures. Post point-injection, this latent representation

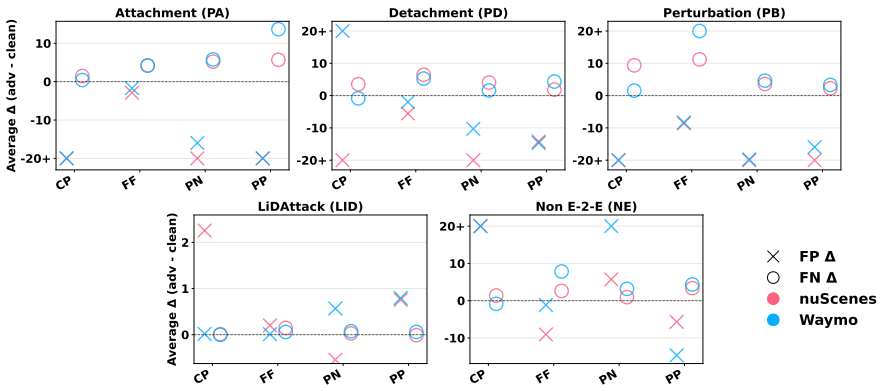


Fig. 3: Visualization of average change (△) in false positives (FP) and false negatives (FN) for each model on nuScenes and Waymo datasets.

is disturbed, resulting in higher misclassifications. On the bright side, we observe that although the PN model uses CenterHead as *detection head*, which demonstrated poor resilience against NE attack, PN exhibits strong robustness against NE attack. This underlines that pillar encoding combined with tailored architectural choices, which in case of PN is pre-trained ConvNeXt *backbone*, effectively addresses the bottlenecks of high-capacity detection frameworks.

PC density as a double-edge sword. Contrary to the intuition that a higher PC density might bolster model resilience, our findings establish that density and robustness are largely orthogonal. While the sparsity of nuScenes, with a 32-beam LiDAR, makes it fragile to PD attack due to the loss of essential geometric landmarks, the rich spatial structure of Waymo, with a 64-beam LiDAR, fails to mitigate perturbation attacks (PB, NE). In fact, we argue that a denser PC offers a richer geometric landscape for adversarial manipulation without significantly altering its global occupancy. The empirical results confirm that, while increased sensor resolution improves performance on clean datasets, the resulting dense PCs are paradoxically more susceptible to fine-grained structural transformations.

Strong robustness against black-box attack Across all models and datasets, the LiD attack remains the least effective, with performance drops typically staying below 2.6%. This shows that the models are robust to black-box attacks and cannot be easily fooled by a universal noise pattern generated in such attacks.

Adversarial Erasure and Systemic Blindness. We quantify the shift in FPs and FNs (△) in Fig. 3. To establish a robust baseline, we first compute the mean FP and FN counts across the benign (clean) data distribution for each architecture. Subsequently, we evaluate these metrics under adversarial conditions to

determine the directional bias of the induced errors. Our results indicate that these attacks are predominantly erasure-dominant, resulting in a state of systemic blindness. The transformer-based FF designed to mitigate FN in high capacity detectors, in contrast, generates high FNs in PB attack. Although FF leverages long-range attention to refine predictions and identify *hard instances*, its global attention introduces a structural vulnerability. The model's expansive receptive field allows adversarial noise from distant, non-target points to propagate into the feature representations of legitimate objects. Consequently, the same mechanism intended to improve model's recall facilitates the transmission of noise, effectively compromising the model's robustness.

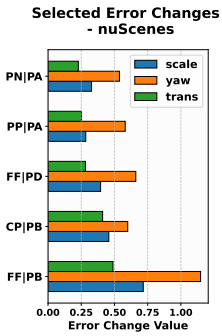


Fig. 4: Comparison of geometric errors for each model|attack combination on nuScenes.

Discussion. The heatmap-based (CP) or hard instance-based (FF) detection approach used by high capacity detectors shows strong limitations in identifying perturbed representations. Contrary to the performance on clean dataset, pillar-based models are more robust than voxel-based high-capacity detectors demonstrating that the vertical aggregation inherent in pillar encoding provides a degree of structural invariance that more granular voxel features lack. In general, advanced 3D-ODs suffer from systemic blindness when the underlying PC structure is transformed and are particularly susceptible to objects possessing distinct vertical representations, such as *safety critical* classes.

4.2 Impact on Localization

Beyond categorical accuracy, to investigate the geometric fidelity of true positives, we quantify the localization error induced by adversarial perturbations on correctly predicted objects. This error is decomposed into three fundamental spatial components: volumetric scaling (scale), angular orientation (yaw) and spatial translation (trans). As visualized in Fig. 4, our analysis on nuScenes reveals that yaw estimation is disproportionately sensitive to adversarial noise, which disrupts orientation reasoning even when the semantic identity of an object is preserved. Notably, we observe an architectural divergence in robustness: while FF under PD attack yields fewer mispredictions, the resulting bounding boxes exhibit profound misorientation. In contrast, CP and FF, despite their extreme mAP sensitivity to NE attacks, maintain higher localization precision on the true positive detections. This confirms that mAP-centric evaluation for adversarial robustness is insufficient and an efficient benchmark must include the impact on localization errors.

Discussion. The evaluation of impact on localization errors exposes how modern 3D-OD's vulnerabilities lead to spatial misorientation in predicted boxes. From a system perspective, these errors are critical; downstream motion planners rely on precise orientation to infer an object's intent and predicted trajectory. Con-

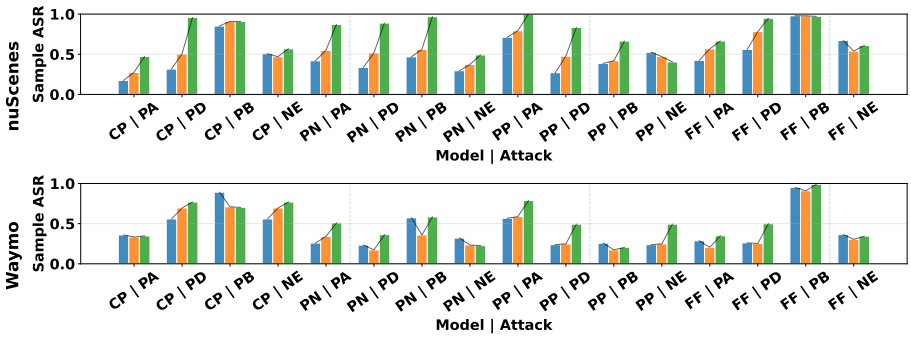


Fig. 5: Impact on ASR for near-far objects for each model-attack pair. The proximity of the objects is represented as ▮ : Near, ▮ : Mid and ▮ : Far.

sequently, a failure of 3D-OD propagates through the stack, potentially leading to risky planning decisions.

4.3 Impact on Distance

Since the PC density has inverse correlation to the proximity of object from ego, it is hypothesized that objects in the immediate vicinity, characterized by higher sampling resolution, should demonstrate superior resilience to adversarial noise. To empirically validate this hypothesis, we categorize the objects into three bins - near, mid and far and then evaluate the impact on their prediction. Referring to the official nuScenes documentation for detection range, we classify objects within 20m from the ego as *near*, 20-35m as *mid* and objects within 35-50m as *far* for both nuScenes and Waymo dataset. The results of our study, as visualized in Fig. 5. Although the initial hypothesis holds in most cases, for perturbation attacks (PB and NE), the adversarial vulnerabilities are distance-agnostic. These findings underscore the limitations of mAP-centric evaluation methods to measure adversarial robustness.

Discussion. Our empirical analysis refutes the intuitive assumption, that Near-ego objects are imminently less vulnerable. The attacks expose vulnerabilities in recent 3D-OD which are sensitive to subtle shifts in geometric representations irrespective of their distance from ego. The global attention and iterative refinement (HIP) of FF makes it equally vulnerable to attacks on near-ego objects, similar to far-away objects. Based on these findings, we strongly underscore the necessity for distance-aware learning strategies [18,24,44] that reinforce both the high-density near-field and the sparse far-away representations.

4.4 Impact of PC Density

To validate our hypothesis that a *holistic robustness evaluation requires the bifurcation of structural precision from predictive confidence*, we analyze performance degradation subject to PC density of adversarial samples in Fig. 6. In this

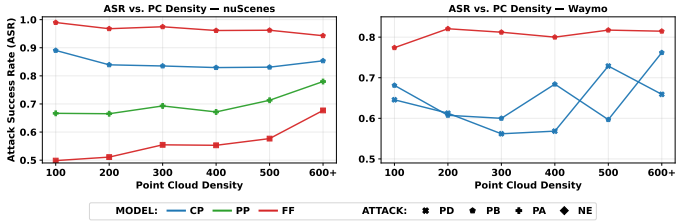


Fig. 6: Visualization of ASR vs point cloud density of models under different attacks.

context, density is defined as the cardinal point count within an object's (*e.g. Car*) volumetric bounds following adversarial manipulation. We observe that few model-attack pairs (FF-PB, CP-PB and PP-PA) report density-invariant susceptibility, analogous to observations in Tab. 2. We observe a paradoxical inverse correlation between PC density and adversarial vulnerability in specific model-attack combinations (PP-NE and CP-NE for nuScenes and CP-PA and PN-PB for Waymo), which substantiates that structural robustness is an inherent property of architectural design rather than a simple byproduct of input sparsity. For most other combinations, the robustness improves as the PC becomes denser (check Appendix). Although this sounds intuitive, we argue that *mAP* alone cannot provide these insights, therefore validating our hypothesis.

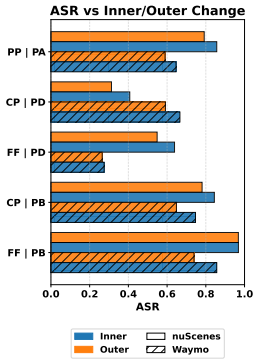


Fig. 7: Impact of PC localization on ASR.

4.5 Impact of PC localization

To further evaluate structural robustness, we analyze the impact on ASR by partitioning the PCs into concentric spatial regions *inner* and *outer*. Our hypothesis posits that perturbations along the points in the inner box strongly degrade performance as they provide volumetric density to the structural contours. The empirical results in Fig. 7, representing top-5 most affected combinations, validate our hypothesis. In most cases, ASR is similarly high for *inner* and *outer* points, indicating perturbations along edges and surface result in similar degradation. However, when the points are dropped along the surface, PD attack, induces significantly higher degradation compared to dropping peripheral points. These results once again validate our previous claim - evaluating structural robustness is equally important for a holistic evaluation of adversarial robustness.

4.6 Sensitivity Analysis

Scaling threshold for Impact of Point Cloud localization. We parameterize the inner-outer point cloud split using a concentric scaling threshold applied to the ground truth bounding box. To ensure a balanced representation across

Table 3: Sensitivity analysis of *Confidence* threshold.

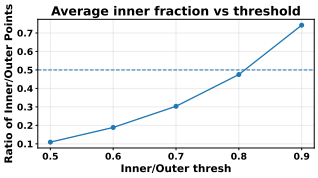
(a) Avg. ASR over all models at various *Confidence* thresholds.

Attack	Confidence Threshold				
	.10	**.15**	**.17**	**.20**	**.25**
NE	.481	.481	.493	.512	.543
PA	.498	.498	.511	.530	.562
PB	.702	.702	.711	.724	.742
PD	.488	.488	.499	.515	.541

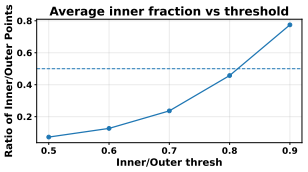
(b) Absolute difference b/w standard ASR and our ASR on Waymo.

	Threshold = 0.15				Threshold = 0.20			
Model	**NE**	**PA**	**PB**	**PD**	**NE**	**PA**	**PB**	**PD**
CP	.01	.00	.01	.01	.03	.01	.04	.07
FF	.01	.00	.02	.00	.04	.01	.05	.03
PN	.00	.02	.00	.00	.01	.07	.00	.00
PP	.00	.01	.00	.00	.01	.04	.00	.00

varying levels of LiDAR sparsity, we calibrate the threshold to equi-partition the points between the internal core and the peripheral volume. Empirical results in Fig. 8 across both datasets indicate that a threshold of 0.8 successfully yields a near-equal point distribution, mitigating density-driven bias in our localized robustness evaluation.



(a) nuScenes



(b) Waymo

Fig. 8: Mean of ratio between inner and outer point clouds across all objects for nuScenes and Waymo dataset.

Selection of *Confidence* threshold for ASR calculation. To justify the choice of the 0.15 confidence threshold within our ASR metric, we perform a comprehensive sensitivity analysis across a spectrum of thresholds from 0.10 to 0.25 as shown in Tab. 3. Setting an excessively high threshold, such as 0.17 or 0.20, introduces an overly pessimistic bias that misclassifies functional detections as total failures. In practical autonomous driving architectures, downstream planning systems typically ingest proposals down to a 0.10 or 0.15 cutoff to maintain high recall. If a higher threshold is enforced during evaluation, any adversarial perturbation causing a minor, non-critical confidence drop (e.g., to 0.18) is incorrectly penalized as a complete detection failure, artificially inflating the ASR. As demonstrated in Tab. 3a, the average ASR remains completely flat between 0.10 and 0.15, validating 0.15 as the optimal empirical boundary that tightly couples adversarial success with genuine operational risk. This is further supported by Tab. 3b, where the absolute difference between our metric and the standard ASR at the 0.15 threshold is negligible (≤ 0.02) across all models.

5 Conclusion

In this study, we conduct a comprehensive adversarial robustness nalysis of SoTA 3D-OD models across commonly benchmarked datasets using latest adversarial

attacks. Our analysis, which holistically evaluates the impact of adversarial attacks on various structural and predictive factors, exposes the limitations of existing *mAP-dominant* frameworks to measure adversarial robustness. The key finding from our empirical results is that the path toward safe autonomous perception requires a shift beyond maximizing *mAP* on clean data and prioritizing structurally robust *encoder* and *head* architectures. Our study reveals that, although recent models significantly outperform classical models on clean data, they lack the ability to detect subtle volumetric inconsistencies in object representations. The empirical results on predictive robustness validate our first hypothesis that SoTA 3D-OD architectures are fragile against coordinated point cloud perturbations. Furthermore, on the dataset-level, our investigation highlights that adversarial robustness is orthogonal to the point cloud density as the models trained on Waymo, providing richer spatial features, experience high sensitivity to point cloud transformations. Finally, our experiments on distance-specific robustness reveal that when subjected to point cloud perturbation attacks, models exhibit low robustness even for near-ego objects. Based on our extensive study, we propose the following recommendations to the research community: 1) Incorporating *distance-aware* learning strategies during model training. 2) To compensate for the geometric homogeneity in point cloud representations, future models should incorporate adversarial data augmentation techniques in their training. We hope that our analysis provides a strong foundation for holistically evaluating adversarial robustness in LiDAR-based 3D object detectors.

Limitations. We trained some models from scratch, where implementation from original authors was missing, under commonly used training configurations. Although we assume that the results will not be fundamentally different, we acknowledge that the training settings might not be optimal, and there is certainly room for improvement.

Acknowledgements

The project on which this publication is based was co-funded by the European Union under grant number EFRE-20801104. We are also grateful to the support provided by Intelligent Sensing, Diagnostics and Prognostics Research Lab ISDPRL and the Digital Research Alliance of Canada AllianceCan. We thank Mitacs for supporting Mr. Yerdana Maulenbay through the RISE-Globalink Research Internship program [IT45510]. Additionally, we thank the RISE (Research Internships in Science and Engineering) Germany program organized by DAAD (Deutscher Akademischer Austauschdienst). Finally, we would like to thank Dr. Ling Bai for their review of the paper, and for offering constructive criticism which helped to address shortcomings in this paper. The computations at the University of Wuppertal were carried out on the PLEIADES cluster, which was supported by the Deutsche Forschungsgemeinschaft (DFG, grant No. INST 218/78-1 FUGG) and the Bundesministerium für Bildung und Forschung (BMBF).

References

1. Abraha, T.M., Graham-Knight, J.B., Lasserre, P., Najjaran, H., Lucet, Y.: Convnet-generated adversarial perturbations for evaluating 3d object detection robustness. Sensors **25**(19) (2025). https://doi.org/10.3390/s25196026, https://www.mdpi.com/1424-8220/25/19/6026
2. Ansel, J., Yang, E., He, H., Gimelshein, N., Jain, A., Voznesensky, M., Bao, B., Bell, P., Berard, D., Burovski, E., Chauhan, G., Chourdia, A., et al.: PyTorch 2: Faster Machine Learning Through Dynamic Python Bytecode Transformation and Graph Compilation. In: 29th ACM International Conference on Architectural Support for Programming Languages and Operating Systems, Volume 2 (ASPLOS '24). ACM (Apr 2024). https://doi.org/10.1145/3620665.3640366, https://pytorch.org/assets/pytorch2-2.pdf
3. Aung, N.H.H., Sangwongngam, P., Jintamethasawat, R., Shah, S., Wuttisittikulkij, L.: A review of lidar-based 3d object detection via deep learning approaches towards robust connected and autonomous vehicles. IEEE Transactions on Intelligent Vehicles **10**(1), 526–547 (2025). https://doi.org/10.1109/TIV.2024.3415771
4. Caesar, H., Bankiti, V., Lang, A.H., Vora, S., Liong, V.E., Xu, Q., Krishnan, A., Pan, Y., Baldan, G., Beijbom, O.: nuscenes: A multimodal dataset for autonomous driving (2020), https://doi.org/10.1109/CVPR42600.2020.01164
5. Chandorkar, A., Tercan, H., Meisen, T.: Rethinking backbone design for lightweight 3d object detection in lidar. In: Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV) Workshops. pp. 1698–1706 (October 2025). https://doi.org/https://doi.org/10.48550/arXiv.2508.00744
6. Chen, H., Yan, H., Yang, X., Su, H., Zhao, S., Qian, F.: Efficient adversarial attack strategy against 3d object detection in autonomous driving systems. IEEE Transactions on Intelligent Transportation Systems **25**(11), 16118–16132 (2024). https://doi.org/10.1109/TITS.2024.3410038
7. Chen, J., Liao, D., Yan, Y., Xiang, S., Zheng, H.: Lidattack: Robust black-box attack on lidar-based object detection. IEEE Transactions on Intelligent Transportation Systems **26**(9), 13563–13572 (2025). https://doi.org/10.1109/TITS.2025.3573055
8. Chen, Y., Yu, Z., Chen, Y., Lan, S., Anandkumar, A., Jia, J., Alvarez, J.M.: FocalFormer3D : Focusing on Hard Instance for 3D Object Detection . In: 2023 IEEE/CVF International Conference on Computer Vision (ICCV). pp. 8360–8371. IEEE Computer Society, Los Alamitos, CA, USA (Oct 2023). https://doi.org/10.1109/ICCV51070.2023.00771, https://doi.ieeecomputersociety.org/10.1109/ICCV51070.2023.00771
9. Chen, Y., Liu, J., Zhang, X., Qi, X., Jia, J.: Voxelnext: Fully sparse voxelnet for 3d object detection and tracking (2023). https://doi.org/10.1109/CVPR52729.2023.02076
10. Contributors, M.: MMDetection3D: OpenMMLab next-generation platform for general 3D object detection. https://github.com/open-mmlab/mmdetection3d (2020)
11. Dong, Y., Kang, C., Zhang, J., Zhu, Z., Wang, Y., Yang, X., Su, H., Wei, X., Zhu, J.: Benchmarking robustness of 3d object detection to common corruptions in autonomous driving. In: 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). pp. 1022–1032 (2023). https://doi.org/10.1109/CVPR52729.2023.00105

12. Du, B., Zhang, C., Sarkar, A., Shen, J., Telikani, A., Hu, H.: Identifying factors related to pedestrian and cyclist crashes in act, australia with an extended crash dataset. Accident Analysis Prevention **207**, 107742 (2024). https://doi.org/https://doi.org/10.1016/j.aap.2024.107742, https://www.sciencedirect.com/science/article/pii/S0001457524002872
13. Erabati, G.K., Araujo, H.: Li3detr: A lidar based 3d detection transformer (2022), https://arxiv.org/abs/2210.15365
14. Eskandar, G.: An empirical study of the generalization ability of lidar 3d object detectors to unseen domains. In: 2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). pp. 23815–23825 (2024). https://doi.org/10.1109/CVPR52733.2024.02248
15. Gamerdinger, J., Teufel, S., Amann, S., Bringmann, O.: Criticality metrics for relevance classification in safety evaluation of object detection in automated driving (2025), https://arxiv.org/abs/2512.15181
16. Geiger, A., Lenz, P., Stiller, C., Urtasun, R.: Vision meets robotics: The KITTI dataset. International Journal of Robotics Research (IJRR) **32**(11), 1231–1237 (2013). https://doi.org/https://dl.acm.org/doi/10.1177/02783649134912
17. Goodfellow, I.J., Shlens, J., Szegedy, C.: Explaining and harnessing adversarial examples. arXiv preprint arXiv:1412.6572 (2014). https://doi.org/10.48550/arXiv.1412.6572
18. Hariya, K., Inoshita, H., Fukuda, Y., Yoneda, K., Suganuma, N.: Real-time 3d object detection with distance-aware hybrid point cloud representation toward long range detection. International Journal of Intelligent Transportation Systems Research (12 2025). https://doi.org/https://doi.org/10.1007/s13177-025-00598-2
19. Kong, L., Liu, Y., Li, X., Chen, R., Zhang, W., Ren, J., Pan, L., Chen, K., Liu, Z.: Robo3d: Towards robust and reliable 3d perception against corruptions. In: Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV). pp. 19994–20006 (October 2023), https://doi.ieeecomputersociety.org/10.1109/ICCV51070.2023.01830
20. Lang, A.H., Vora, S., Caesar, H., Zhou, B., Yang, J., Beijbom, O.: Pointpillars: Fast encoders for object detection from point clouds. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp. 12697–12705 (2019). https://doi.org/10.1109/CVPR.2019.01298
21. Li, J., Luo, C., Yang, X.: Pillarnext: Rethinking network designs for 3d object detection in lidar point clouds. In: 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). pp. 17567–17576 (2023). https://doi.org/10.1109/CVPR52729.2023.01685
22. Liu, Z., Mao, H., Wu, C.Y., Feichtenhofer, C., Darrell, T., Xie, S.: A convnet for the 2020s. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). pp. 11976–11986 (June 2022). https://doi.org/10.1109/CVPR52688.2022.01186
23. Long, H., Chen, H., Xu, M., Zhang, C., Qian, F.: Crafting transferable adversarial examples against 3d object detection. IET Computer Vision **19**(1), e70011 (2025). https://doi.org/https://doi.org/10.1049/cvi2.70011
24. Lu, Y., Hao, X., Li, Y., Chai, W., Sun, S., Velipasalar, S.: Range-aware attention network for lidar-based 3d object detection with auxiliary point density level estimation. IEEE Transactions on Vehicular Technology **74**(1), 292–305 (2025). https://doi.org/10.1109/TVT.2024.3454607

25. Madry, A., Makelov, A., Schmidt, L., Tsipras, D., Vladu, A.: Towards deep learning models resistant to adversarial attacks (2019), https://arxiv.org/abs/1706.06083
26. Madry, A., Makelov, A., Schmidt, L., Tsipras, D., Vladu, A.: Towards deep learning models resistant to adversarial attacks. arXiv preprint arXiv:1706.06083 (2017). https://doi.org/10.48550/arXiv.1706.06083
27. Mao, W., Wang, T., Zhang, D., Yan, J., Yoshie, O.: Pillarnest: Embracing backbone scaling and pretraining for pillar-based 3d object detection. IEEE Transactions on Intelligent Vehicles **10**(4), 2905–2914 (2025). https://doi.org/10.1109/TIV.2024.3386576
28. Scarano, A., Aria, M., Mauriello, F., Riccardi, M.R., Montella, A.: Systematic literature review of 10 years of cyclist safety research. Accident Analysis Prevention **184**, 106996 (2023). https://doi.org/https://doi.org/10.1016/j.aap.2023.106996, https://www.sciencedirect.com/science/article/pii/S000145752300043X
29. Shi, G., Li, R., Ma, C.: Pillarnet: Real-time and high-performance pillar-based 3d object detection. In: European Conference on Computer Vision. pp. 35–52 (2022). https://doi.org/https://doi.org/10.1007/978-3-031-20080-9_3
30. Shi, S., Guo, C., Jiang, L., Wang, Z., Shi, J., Wang, X., Li, H.: Pv-rcnn: Point-voxel feature set abstraction for 3d object detection. In: 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). pp. 10526–10535 (2020). https://doi.org/10.1109/CVPR42600.2020.01054
31. Shi, S., Jiang, L., Deng, J., Wang, Z., Guo, C., Shi, J., Wang, X., Li, H.: Pv-rcnn++: Point-voxel feature set abstraction with local vector representation for 3d object detection (2022), https://doi.org/10.48550/arXiv.2102.00463
32. Shi, S., Wang, X., Li, H.: Pointrcnn: 3d object proposal generation and detection from point cloud. In: 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). pp. 770–779 (2019). https://doi.org/10.1109/CVPR.2019.00086
33. Shi, S., Wang, Z., Shi, J., Wang, X., Li, H.: From points to parts: 3d object detection from point cloud with part-aware and part-aggregation network (2020), https://doi.org/10.48550/arXiv.1907.03670
34. Song, Z., Liu, L., Jia, F., Luo, Y., Jia, C., Zhang, G., Yang, L., Wang, L.: Robustness-aware 3d object detection in autonomous driving: A review and outlook. IEEE Transactions on Intelligent Transportation Systems **25**(11), 15407–15436 (2024). https://doi.org/10.1109/TITS.2024.3439557
35. Sun, P., Kretzschmar, H., Dotiwalla, X., Chouard, A., Patnaik, V., Tsui, P., Guo, J., Zhou, Y., Chai, Y., Caine, B., Vasudevan, V., Han, W., Ngiam, J., Zhao, H., Timofeev, A., Ettinger, S., Krivokon, M., Gao, A., Joshi, A., Zhang, Y., Shlens, J., Chen, Z., Anguelov, D.: Scalability in perception for autonomous driving: Waymo open dataset. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) (June 2020). https://doi.org/http://ieeexplore.ieee.org/document/9156973/
36. Tu, J., Ren, M., Manivasagam, S., Liang, M., Yang, B., Du, R., Cheng, F., Urtasun, R.: Physically realizable adversarial examples for lidar object detection. In: 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). pp. 13713–13722 (2020). https://doi.org/10.1109/CVPR42600.2020.01373
37. Wang, J., Meng, Q., Liu, G., Yan, L., Wang, K., Cheng, M.M., Hou, Q.: Towards stable 3d object detection. In: European conference on computer vision. Springer (2024). https://doi.org/https://doi.org/10.1007/978-3-031-72973-7_12

38. Wang, Y., Wu, L., Jin, J., Wang, E., Zhang, Z., Zhao, Y.: An imperceptible adversarial attack against 3-d object detectors in autonomous driving. IEEE Internet of Things Journal **12**(12), 21404–21414 (2025). https://doi.org/10.1109/JIOT.2025.3547966
39. Wang, Z., Li, Y., Chen, X., Zhao, H., Wang, S.: Uni3detr: unified 3d detection transformer. In: Proceedings of the 37th International Conference on Neural Information Processing Systems. NIPS '23, Curran Associates Inc., Red Hook, NY, USA (2023). https://doi.org/https://dl.acm.org/doi/10.5555/3666122.3667855
40. Xiang, C., Qi, C.R., Li, B.: Generating 3d adversarial point clouds. In: 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). pp. 9128–9136 (2019). https://doi.org/10.1109/CVPR.2019.00935
41. Xiang, X., Zhang, J.: Fusionvit: Hierarchical 3d object detection via lidar-camera vision transformer fusion (2023), https://arxiv.org/abs/2311.03620
42. Xu, W., Zhou, S., Yuan, Z.: Pillartrack: Redesigning pillar-based transformer network for single object tracking on point clouds. arXiv preprint arXiv:2404.07495 (2024), https://doi.org/10.48550/arXiv.2404.07495
43. Yan, Y., Mao, Y., Li, B.: Second: Sparsely embedded convolutional detection. Sensors **18**(10), 3337 (2018). https://doi.org/10.3390/s18103337
44. Yan, Z., Lin, Y., Wang, K., Zheng, Y., Wang, Y., Zhang, Z., Li, J., Yang, J.: Tri-perspective view decomposition for geometry-aware depth completion. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). pp. 4874–4884 (June 2024). https://doi.org/https://doi.org/10.48550/arXiv.2403.15008
45. Yang, Z., Sun, Y., Liu, S., Jia, J.: 3dssd: Point-based 3d single stage object detector. In: 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). pp. 11037–11045 (2020). https://doi.org/10.1109/CVPR42600.2020.01105
46. Yin, T., Zhou, X., Krähenbühl, P.: Center-based 3d object detection and tracking. In: 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). pp. 11779–11788 (2021). https://doi.org/10.1109/CVPR46437.2021.01161
47. Zhang, P., Li, X., Lin, X., He, L.: A new literature review of 3d object detection on autonomous driving. J. Artif. Int. Res. **82** (Jun 2025), https://doi.org/10.1613/jair.1.15961
48. Zhang, Y., Hou, J., Yuan, Y.: A comprehensive study of the robustness for lidar-based 3d object detectors against adversarial attacks (2023), https://doi.org/10.1007/s11263-023-01934-3
49. Zhou, C., Zhang, Y., Chen, J., Huang, D.: Octr: Octree-based transformer for 3d object detection. In: 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). pp. 5166–5175 (2023). https://doi.org/10.1109/CVPR52729.2023.00500
50. Zhou, S., Tian, Z., Chu, X., Zhang, X., Zhang, B., Lu, X., Feng, C., Jie, Z., Chiang, P.Y., Ma, L.: Fastpillars: A deployment-friendly pillar-based 3d detector (2023), https://arxiv.org/abs/2302.02367
51. Zhou, Y., Tuzel, O.: Voxelnet: End-to-end learning for point cloud based 3d object detection. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 4490–4499 (2018). https://doi.org/10.1109/CVPR.2018.00472
52. Zhu, J., et al.: Vulnerability analysis of lidar-based 3d object detection to adversarial lidar points. IEEE Transactions on Intelligent Transportation Systems (2023). https://doi.org/10.1109/TITS.2023.3274635

53. Zhu, Z., Zhang, Y., Chen, H., Dong, Y., Zhao, S., Ding, W., Zhong, J., Zheng, S.: Understanding the robustness of 3d object detection with bird'view representations in autonomous driving. In: 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). pp. 21600–21610 (2023). https://doi.org/10.1109/CVPR52729.2023.02069