Learning to Generate 4D LiDAR Sequences

Anonymous ICCV Wild3D Workshop Submission

Paper ID 11

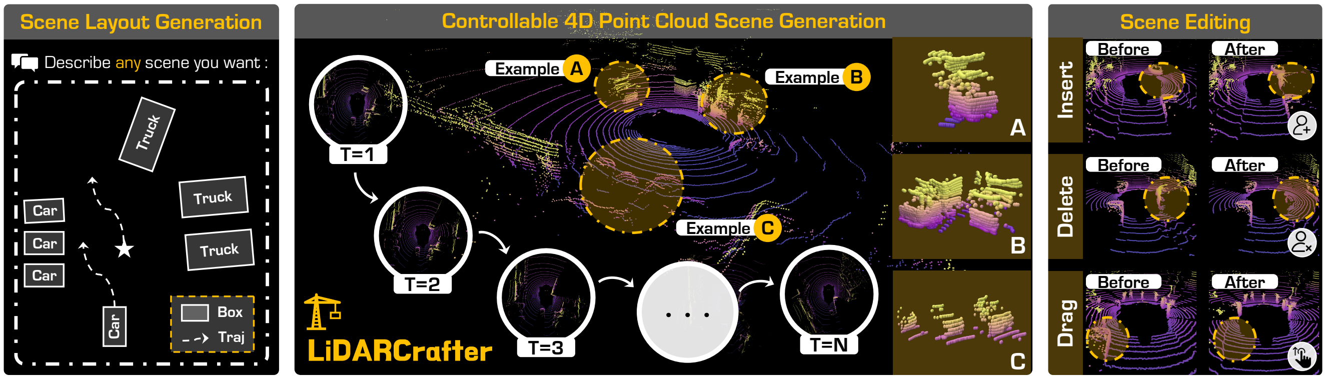


Figure 1. We propose **LiDARCrafter**, a 4D LiDAR-based generative world model that supports controllable point cloud layout generation (**left**), dynamic sequential scene generation (**center**), and rich scene editing applications (**right**). Our framework enables intuitive *"what you describe is what you get"* LiDAR-based 4D world modeling.

Abstract

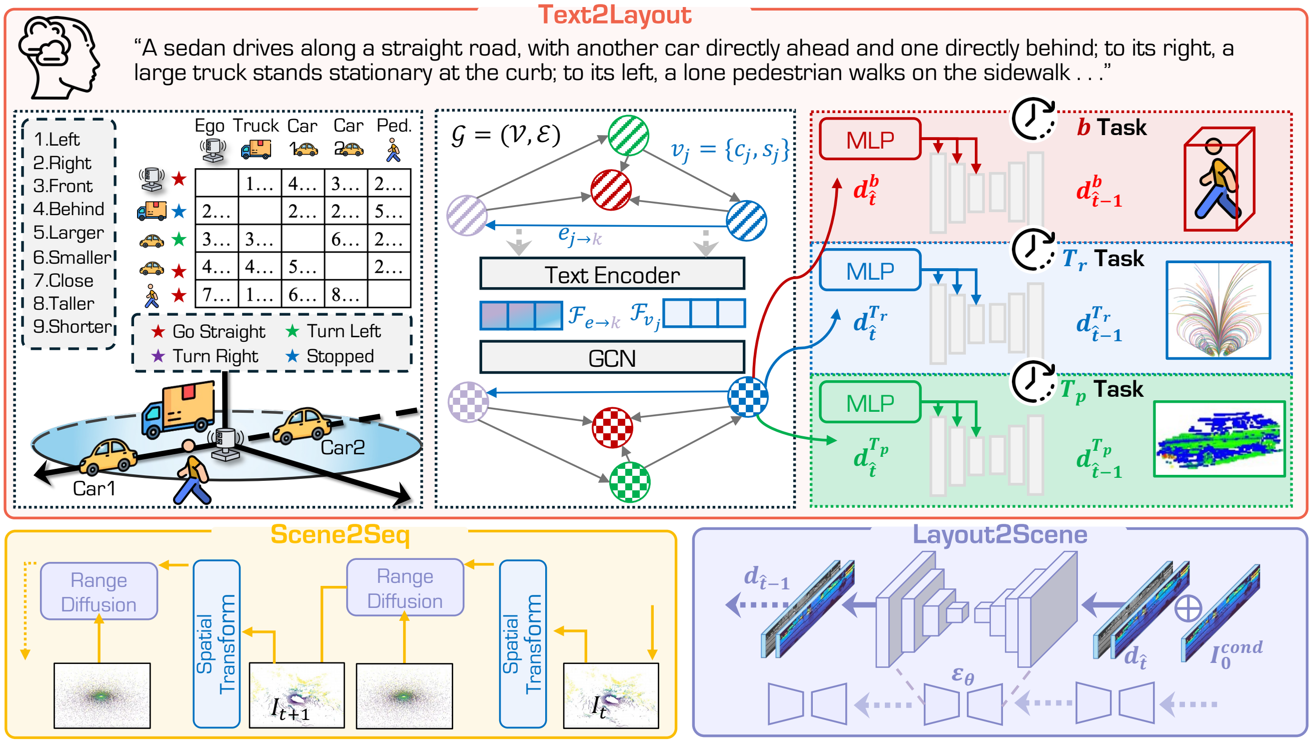
*While generative world models have advanced video and occupancy-based data synthesis, LiDAR generation remains underexplored despite its importance for accurate 3D perception. Extending generation to 4D LiDAR data introduces challenges in controllability, temporal stability, and evaluation. We present **LiDARCrafter**, a unified framework that converts free-form language into editable LiDAR sequences. Instructions are parsed into ego-centric scene graphs, which a tri-branch diffusion model transforms into object layouts, trajectories, and shapes. A range-image diffusion model generates the initial scan, and an autoregressive module extends it into a temporally coherent sequence. The explicit layout design further supports object-level editing, such as insertion or relocation. To enable fair assessment, we provide **EvalSuite**, a benchmark spanning scene-, object-, and sequence-level metrics. On nuScenes, LiDARCrafter achieves state-of-the-art fidelity, controllability, and temporal consistency, offering a foundation for LiDAR-based simulation and data augmentation.*

1. Introduction

Generative world models are reshaping autonomous driving by enabling scalable simulation and interpretation of sensor-rich environments [6, 16]. Recent advances have focused on structured modalities such as videos and occupancy grids, whose dense and regular representations align naturally with image or voxel pipelines. Video-based methods [6, 16, 32] leverage autoregression and richer conditioning, while BEV-based approaches such as MagicDrive [4] enforce temporal consistency across frames [23, 33]. Occupancy-based works [1, 22, 34, 35], capture fine spatial structure for downstream tasks. Multimodal frameworks [5, 13] further align cross-modal signals for consistency. Despite this progress, LiDAR, a core modality for precise 3D geometry and all weather robustness remains comparatively underexplored.

LiDAR point clouds present unique challenges. They are sparse, irregular, and unordered [12, 14, 15, 27], making direct application of image- or voxel-based techniques ineffective. Early efforts such as LiDARGen [36] project 360° scans to range images and adapt pixel-based diffusion, while subsequent works like RangeLDM [7], R2DM [17], and R2Flow [18] improved single-frame fidelity. Other methods such as Text2LiDAR [24], WeatherGen [25], and UltraLiDAR [26] introduced diverse conditioning or editing capabilities. Yet, most are restricted to static scans or lack temporal modeling, leaving 4D sequence generation and fine-grained control unresolved.

A central obstacle is **spatial controllability**. Existing



	Ego	Truck	Car 1	Car 2	Ped.
★		1…	4…	3…	2…
★	2…		2…	2…	5…
★	3…	3…		6…	2…
★	4…	4…	5…		2…
★	7…	1…	6…	8…	

Figure 2. **Overview of the LiDARCrafter framework.** In the **Text2Layout** stage (*cf.* Sec. 2.1), the natural-language instruction is parsed into an ego-centric scene graph, and a tri-branch diffusion network generates 4D conditions for bounding boxes, future trajectories, and object point clouds. In the **Layout2Scene** stage (*cf.* Sec. 2.2), a range-image diffusion model uses these conditions to generate a static LiDAR frame. In the **Scene2Seq** stage (*cf.* Sec. 2.3), an autoregressive module warps historical points with ego and object motion priors to generate each subsequent frame, producing a temporally coherent LiDAR sequence.

models often require costly inputs such as HD maps [21] or 3D bounding boxes [28, 31], while text-only methods [6, 24] are more accessible but lack spatial precision. Indoor scene synthesis has addressed this trade-off with intermediate scene graphs [29, 30], but such strategies are not yet established for outdoor, dynamic LiDAR streams. Beyond controllability, LiDAR world models also lack **temporal coherence**: single-frame synthesis cannot capture occlusions or motion patterns, and naïve cross-frame attention overlooks the geometric continuity of point clouds. Finally, unlike video models that benefit from benchmarks such as VBench [8], LiDAR has no standardized protocols to evaluate scene-, object-, and sequence-level quality.

We introduce **LiDARCrafter**, the first unified framework for controllable 4D LiDAR sequence generation. At its core is an explicit, object-centric 4D layout that bridges free-form language instructions with LiDAR geometry and motion. In the **Text2Layout** stage, a large language model parses descriptions into an ego-centric scene graph, which a tri-branch diffusion network expands into object boxes, trajectories, and shapes. **Layout2Scene** converts this layout into a high-fidelity initial scan using a range-image diffusion backbone, enabling precise editing such as object insertion or relocation. **Scene2Seq** autoregressively extends the sequence by warping foreground and background points with motion priors to maintain long-term temporal consistency. To close the evaluation gap, we release **EvalSuite**, the first benchmark that jointly scores semantic correctness, layout validity, and sequence smoothness.

In summary, LiDARCrafter establishes a new paradigm for LiDAR-based world modeling by combining intuitive language-driven control, explicit layout conditioning, and temporally stable generation. It offers both a practical synthesis tool and a standardized benchmark, providing the community with a foundation for controllable and consistent 4D LiDAR simulation.

2. LiDARCrafter: 4D LiDAR World Model

We introduce **LiDARCrafter**, the first generative world model dedicated to LiDAR, which transforms free-form instructions into temporally coherent 4D point cloud sequences with object-level control. The core idea is to maintain an explicit 4D layout that bridges language descriptions with LiDAR geometry. As shown in Fig. 2, the framework follows three stages: **Text2Layout**, which lifts language into a structured 4D layout; **Layout2Scene**, which generates a controllable first scan; and **Scene2Seq**, which autoregressively extends the sequence. We further introduce **EvalSuite**, a protocol for standardized evaluations.

2.1. Text2Layout: Language to 4D Layout

Since natural language lacks the spatial precision needed for LiDAR synthesis, we construct an intermediate scene graph. An LLM parses the user prompt into an ego-centric graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where nodes represent the ego vehicle and dynamic objects with semantic labels and motion states, while edges encode spatial relations. This explicit graph captures both semantic and relational cues.

Each node is lifted into a layout tuple $\mathcal{O}_i = (\mathbf{b}_i, \boldsymbol{\delta}_i, \mathbf{p}_i)$, including a 3D bounding box, a trajectory of planar offsets, and canonical shape points. To enrich these with contextual semantics, we process the graph with a TripletGCN [9], embedding nodes and edges with a frozen CLIP encoder [20]. The resulting features condition a tri-branch diffusion decoder: one branch denoises boxes, another predicts trajectories, and a third synthesizes coarse shapes. This structured 4D layout provides the foundation for later LiDAR synthesis while supporting explicit object-level control.

2.2. Layout2Scene: Controlled LiDAR Generation

Given the 4D layout, LiDARCrafter generates the initial LiDAR frame using a range-image diffusion model, which preserves LiDAR geometry while leveraging efficient convolutional backbones [11, 17]. To address sparsity, particularly for small or distant objects, we condition the network on compact object representations encoding category, pose, and shape priors, projected onto the range view [10]. A lightweight attention layer propagates context across objects, and scene-level embeddings provide global conditioning. During denoising, the noisy map is combined with this conditioning to yield a coherent first scan.

The explicit layout also enables fine-grained editing. By modifying layout tuples (e.g., inserting, deleting, or dragging objects), users can re-synthesize scenes with only local changes, preserving the rest of the scan. This makes LiDARCrafter suitable for interactive scenario design in simulation and planning research.

2.3. Scene2Seq: Autoregressive Sequence Synthesis

To extend a single scan into a full 4D sequence, we adopt an autoregressive strategy. Unlike video, where textures change every frame, LiDAR scenes are largely static except for moving agents and ego motion. LiDARCrafter exploits this by warping background points with the ego pose and foreground objects with their predicted trajectories. This warp provides a strong geometric prior at each timestep, which the diffusion model refines into a clean range map. To prevent accumulated drift, we include a warp from the first frame to every later frame, ensuring long-term stability. The result is a temporally consistent sequence where motion and occlusion patterns remain realistic.

Table 1. Evaluations of **scene-level fidelity** for LiDAR generation on the *nuScenes* dataset. MMD values are reported in 10^{-4} and JSD in 10^{-2}. Lower is better for all metrics (↓).

#	Method	Venue	Range		Points		BEV	
			FRD↓	MMD↓	FPD↓	MMD↓	JSD↓	MMD↓
Voxel	UniScene	CVPR'25	–	–	976.47	29.06	31.55	13.61
	OpenDWM	CVPR'25	–	–	714.19	21.95	20.17	5.61
	OpenDWM-DiT	CVPR'25	–	–	381.91	12.46	19.90	5.73
Range	LiDARGen	ECCV'22	759.65	1.71	159.35	35.52	5.74	2.39
	LiDM	CVPR'24	495.54	0.18	210.20	8.45	5.86	0.73
	RangeLDM	ECCV'24	–	–	–	–	5.47	1.92
	R2DM	ICRA'24	243.35	1.40	33.97	1.62	3.51	0.71
	LiDARCrafter	**Ours**	**194.37**	**0.08**	**8.64**	**0.90**	**3.11**	**0.42**

Table 2. Comparison of **foreground object quality** using FDC (↑), which reflects detector confidence on generated scenes. #Box is the average number of boxes per frame.

#	Method	Venue	Car↑	Ped↑	Truck↑	Bus↑	#Box
Uncond.	LiDARGen	ECCV'22	0.57	0.29	0.42	0.38	0.364
	LiDM	CVPR'24	0.65	0.22	0.45	0.31	0.28
	R2DM	ICRA'24	0.54	0.29	0.39	0.35	0.53
Cond.	UniScene	CVPR'25	0.53	0.28	0.35	0.25	0.98
	OpenDWM	CVPR'25	0.74	0.30	0.51	0.44	0.54
	OpenDWM-DiT	CVPR'25	0.78	0.32	**0.56**	0.51	0.64
	LiDARCrafter	**Ours**	**0.83**	**0.34**	0.55	**0.54**	**1.84**

Table 3. Evaluation of **object-level fidelity** for LiDAR generation. MMD is reported in 10^{-4}, and JSD in 10^{-2}.

#	Method	Venue	FPD↓	P-MMD↓	JSD↓	MMD↓
Uncond.	LiDARGen	ECCV'22	1.39	0.15	0.20	16.22
	LiDM	CVPR'24	1.41	0.15	0.19	13.49
	R2DM	ICRA'24	1.40	0.15	0.17	12.76
Cond.	UniScene	CVPR'25	1.19	0.18	0.23	16.65
	OpenDWM	CVPR'25	1.49	0.19	0.16	9.11
	OpenDWM-DiT	CVPR'25	1.48	0.18	**0.15**	9.02
	LiDARCrafter	**Ours**	**1.03**	**0.13**	**0.15**	**5.48**

2.4. EvalSuite: Comprehensive Evaluation

LiDAR generation requires metrics beyond static fidelity. Our **EvalSuite** measures quality across three levels: **Scene-level**: evaluating global realism and distributional fidelity of entire scenes. **Object-level**: verifying semantic correctness, bounding box geometry, and detection confidence. **Temporal-level**: assessing motion smoothness and frame-to-frame transform accuracy.

Together, these metrics provide the first standardized benchmark for 4D LiDAR sequence generation, enabling fair and holistic evaluation of future methods.

3. Experiments

We evaluate **LiDARCrafter** on nuScenes [2] using both classical LiDAR generation metrics (FRD, FPD, JSD, MMD) and our EvalSuite for measuring object-, layout- and sequence-level generation quality. Comparisons are made against recent LiDAR generative models, including R2DM

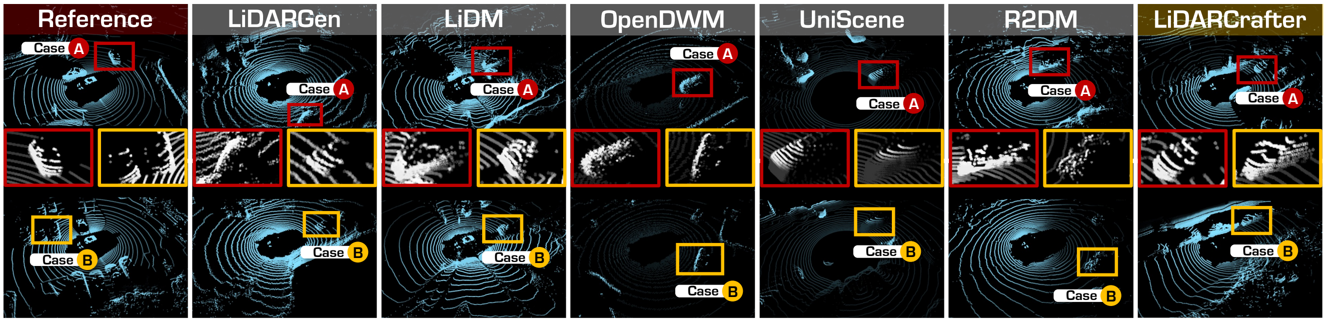


Figure 3. Single-frame LiDAR point cloud generation results. LiDARCrafter produces the pattern closest to the ground truth, with notably superior foreground quality compared to other methods. Best viewed at high resolution.

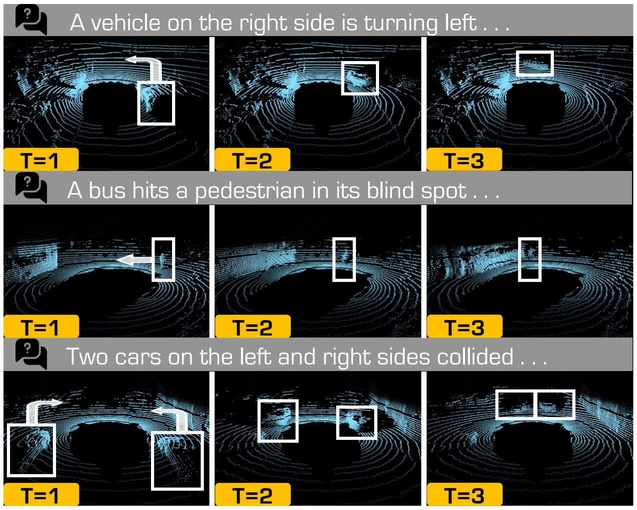

A vehicle on the right side is turning left . . .
T=1
T=2
T=3
A bus hits a pedestrian in its blind spot . . .
T=1
T=2
T=3
Two cars on the left and right sides collided . . .
T=1
T=2
T=3

Figure 4. **Diverse corner cases** generated by LiDARCrafter with object-centric controllability. Best viewed at high resolution. Frames are arranged sequentially from left to right.

Table 4. Comparison of **temporal consistency** in 4D LiDAR generation. TTCE measures transformation error from point cloud registration; CTC computes Chamfer Distance between adjacent frames. Numbers indicate frame intervals.

Method	Venue	TTCE↓		CTC↓			
		3	4	1	2	3	4
UniScene	CVPR'25	2.74	3.69	0.90	1.84	3.64	**3.90**
OpenDWM	CVPR'25	2.68	3.65	1.02	2.02	3.37	5.05
OpenDWM-DiT	CVPR'25	2.71	3.66	**0.89**	**1.79**	3.06	4.64
LiDARCrafter	**Ours**	**2.65**	**3.56**	1.12	2.38	**3.02**	4.81

[17], R2Flow [18], and OpenDWM [19].

3.1. Scene-Level Generation

At the scene level, we assess both whole-scan fidelity and the accuracy of synthesized foregrounds. As shown in Tab. 1, LiDARCrafter attains the lowest FRD and FPD, outperforming prior methods by a notable margin. Qualitative comparisons in Fig. 3 confirm that our model reconstructs realistic global structure while preserving sharp object geometry, whereas alternatives often suffer from noise or blurred backgrounds.

Foreground quality is further validated using a pre-trained VoxelRCNN detector [3]. We report Foreground Detection Confidence (FDC). LiDARCrafter achieves the highest scores across most categories (Tab. 2), demonstrating that generated objects align closely with ground-truth semantics and boxes.

3.2. Object-Level Generation

At the object level, we benchmark fidelity and consistency under box-level conditioning. Using 2,000 car instances, LiDARCrafter achieves the best scores in FPD and MMD (Tab. 3), showing superior reconstruction of fine-grained geometry compared to OpenDWM and R2Flow.

3.3. Temporal-Level Generation

We evaluate temporal consistency in 4D LiDAR generation in Table 4. Temporal Transformation Consistency Error (TTCE) measures the error between the predicted and ground-truth transformation matrices obtained via point cloud registration, while Chamfer Temporal Consistency (CTC) computes the Chamfer Distance between consecutive frames. Our approach achieves the lowest TTCE scores across both frame intervals and maintains competitive CTC performance at all intervals, demonstrating strong temporal coherence.

4. Conclusion

We presented **LiDARCrafter**, a unified framework for controllable 4D LiDAR sequence generation and editing. By leveraging scene graph descriptors, the multi-branch diffusion model, and an autoregressive generation strategy, our approach achieves fine-grained controllability and strong temporal consistency. Experiments on nuScenes demonstrate clear improvements over existing methods in fidelity, coherence, and controllability. Beyond high-quality data synthesis, LiDARCrafter enables the creation of safety-critical scenarios for robust evaluation of downstream autonomous driving systems. Future work will explore multi-modal extensions and further efficiency improvements.

References

[1] Hengwei Bian, Lingdong Kong, Haozhe Xie, Liang Pan, Yu Qiao, and Ziwei Liu. DynamicCity: Large-scale 4D occupancy generation from dynamic scenes. In *ICLR*, 2025.

[2] Holger Caesar, Varun Bankiti, Alex H Lang, Sourabh Vora, Venice Erin Liong, Qiang Xu, Anush Krishnan, Yu Pan, Giancarlo Baldan, and Oscar Beijbom. nuScenes: A multimodal dataset for autonomous driving. In *CVPR*, pages 11621–11631, 2020.

[3] Jiajun Deng, Shaoshuai Shi, Peiwei Li, Wengang Zhou, Yanyong Zhang, and Houqiang Li. Voxel R-CNN: Towards high performance voxel-based 3D object detection. In *AAAI*, pages 1201–1209, 2021.

[4] Ruiyuan Gao, Kai Chen, Enze Xie, Lanqing Hong, Zhenguo Li, Dit-Yan Yeung, and Qiang Xu. MagicDrive: Street view generation with diverse 3D geometry control. In *ICLR*, 2023.

[5] Xiangyu Guo, Zhanqian Wu, Kaixin Xiong, Ziyang Xu, Lijun Zhou, Gangwei Xu, Shaoqing Xu, Haiyang Sun, Bing Wang, Guang Chen, et al. Genesis: Multimodal driving scene generation with spatio-temporal and cross-modal consistency. *arXiv preprint arXiv:2506.07497*, 2025.

[6] Anthony Hu, Lloyd Russell, Hudson Yeo, Zak Murez, George Fedoseev, Alex Kendall, Jamie Shotton, and Gianluca Corrado. GAIA-1: A generative world model for autonomous driving. *arXiv preprint arXiv:2309.17080*, 2023.

[7] Qianjiang Hu, Zhimin Zhang, and Wei Hu. RangeLDM: Fast realistic LiDAR point cloud generation. In *ECCV*, pages 115–135. Springer, 2024.

[8] Ziqi Huang, Yinan He, Jiashuo Yu, Fan Zhang, Chenyang Si, Yuming Jiang, Yuanhan Zhang, Tianxing Wu, Qingyang Jin, Nattapol Chanpaisit, et al. VBench: Comprehensive benchmark suite for video generative models. In *CVPR*, pages 21807–21818, 2024.

[9] Justin Johnson, Agrim Gupta, and Li Fei-Fei. Image generation from scene graphs. In *CVPR*, pages 1219–1228, 2018.

[10] Ellington Kirby, Mickael Chen, Renaud Marlet, and Nermin Samet. LOGen: Toward LiDAR object generation by point diffusion. *arXiv preprint arXiv:2412.07385*, 2024.

[11] Lingdong Kong, Youquan Liu, Runnan Chen, Yuexin Ma, Xinge Zhu, Yikang Li, Yuenan Hou, Yu Qiao, and Ziwei Liu. Rethinking range view representation for LiDAR segmentation. In *ICCV*, pages 228–240, 2023.

[12] Lingdong Kong, Youquan Liu, Xin Li, Runnan Chen, Wenwei Zhang, Jiawei Ren, Liang Pan, Kai Chen, and Ziwei Liu. Robo3D: Towards robust and reliable 3D perception against corruptions. In *ICCV*, pages 19994–20006, 2023.

[13] Bohan Li, Jiazhe Guo, Hongsi Liu, Yingshuang Zou, Yikang Ding, Xiwu Chen, Hu Zhu, Feiyang Tan, Chi Zhang, Tiancai Wang, et al. UniScene: Unified occupancy-centric driving scene generation. In *CVPR*, pages 11971–11981, 2025.

[14] Ao Liang, Lingdong Kong, Dongyue Lu, Youquan Liu, Jian Fang, Huaici Zhao, and Wei Tsang Ooi. Perspective-invariant 3d object detection. In *ICCV*, 2025.

[15] Youquan Liu, Lingdong Kong, Jun Cen, Runnan Chen, Wenwei Zhang, Liang Pan, Kai Chen, and Ziwei Liu. Segment any point cloud sequences by distilling vision foundation models. In *NeurIPS*, pages 37193–37229, 2023.

[16] Jianbiao Mei, Tao Hu, Xuemeng Yang, Licheng Wen, Yu Yang, Tiantian Wei, Yukai Ma, Min Dou, Botian Shi, and Yong Liu. DreamForge: Motion-aware autoregressive video generation for multi-view driving scenes. *arXiv preprint arXiv:2409.04003*, 2024.

[17] Kazuto Nakashima and Ryo Kurazume. LiDAR data synthesis with denoising diffusion probabilistic models. In *ICRA*, pages 14724–14731, 2024.

[18] Kazuto Nakashima, Xiaowen Liu, Tomoya Miyawaki, Yumi Iwashita, and Ryo Kurazume. Fast LiDAR data generation with rectified flows. *arXiv preprint arXiv:2412.02241*, 2024.

[19] Jingcheng Ni, Yuxin Guo, Yichen Liu, Rui Chen, Lewei Lu, and Zehuan Wu. Opendwm: Open driving world models, 2025. https://github.com/SenseTime-FVG/OpenDWM.

[20] Alec Radford, Jong Wook Kim, Chris Hallacy, Aditya Ramesh, Gabriel Goh, Sandhini Agarwal, Girish Sastry, Amanda Askell, Pamela Mishkin, Jack Clark, et al. Learning transferable visual models from natural language supervision. In *ICML*, pages 8748–8763. PmLR, 2021.

[21] Alexander Swerdlow, Runsheng Xu, and Bolei Zhou. Street-view image generation from a bird's-eye view layout. *RA-L*, 9(4):3578–3585, 2024.

[22] Lening Wang, Wenzhao Zheng, Yilong Ren, Han Jiang, Zhiyong Cui, Haiyang Yu, and Jiwen Lu. OccSora: 4D occupancy generation models as world simulators for autonomous driving. *arXiv preprint arXiv:2405.20337*, 2024.

[23] Xiaofeng Wang, Zheng Zhu, Guan Huang, Xinze Chen, Jiagang Zhu, and Jiwen Lu. DriveDreamer: Towards real-world-drive world models for autonomous driving. In *ECCV*, pages 55–72. Springer, 2024.

[24] Yang Wu, Kaihua Zhang, Jianjun Qian, Jin Xie, and Jian Yang. Text2LiDAR: Text-guided LiDAR point cloud generation via equirectangular transformer. In *ECCV*, pages 291–310. Springer, 2024.

[25] Yang Wu, Yun Zhu, Kaihua Zhang, Jianjun Qian, Jin Xie, and Jian Yang. WeatherGen—: A unified diverse weather generator for LiDAR point clouds via spider mamba diffusion. In *CVPR*, pages 17019–17028, 2025.

[26] Yuwen Xiong, Wei-Chiu Ma, Jingkang Wang, and Raquel Urtasun. UltraLiDAR: Learning compact representations for LiDAR completion and generation. *arXiv preprint arXiv:2311.01448*, 2023.

[27] Xiang Xu, Lingdong Kong, Hui Shuai, Wenwei Zhang, Liang Pan, Kai Chen, Ziwei Liu, and Qingshan Liu. 4D contrastive superflows are dense 3D representation learners. In *ECCV*, pages 58–80. Springer, 2024.

[28] Xuemeng Yang, Licheng Wen, Yukai Ma, Jianbiao Mei, Xin Li, Tiantian Wei, Wenjie Lei, Daocheng Fu, Pinlong Cai, Min Dou, et al. DriveArena: A closed-loop generative simulation platform for autonomous driving. *arXiv preprint arXiv:2408.00415*, 2024.

[29] Zhifei Yang, Keyang Lu, Chao Zhang, Jiaxing Qi, Hanqi Jiang, Ruifei Ma, Shenglin Yin, Yifan Xu, Mingzhe Xing, Zhen Xiao, et al. MMGDreamer: Mixed-modality graph for geometry-controllable 3D indoor scene generation. In *AAAI*, pages 9391–9399, 2025.

[30] Guangyao Zhai, Evin Pınar Örnek, Dave Zhenyu Chen, Ruotong Liao, Yan Di, Nassir Navab, Federico Tombari, and

Benjamin Busam. EchoScene: Indoor scene generation via information echo over scene graph diffusion. In *ECCV*, pages 167–184. Springer, 2024.

[31] Jinhua Zhang, Hualian Sheng, Sijia Cai, Bing Deng, Qiao Liang, Wen Li, Ying Fu, Jieping Ye, and Shuhang Gu. PerlDiff: Controllable street view synthesis using perspective-layout diffusion models. *arXiv preprint arXiv:2407.06109*, 2024.

[32] Kaiwen Zhang, Zhenyu Tang, Xiaotao Hu, Xingang Pan, Xiaoyang Guo, Yuan Liu, Jingwei Huang, Li Yuan, Qian Zhang, Xiao-Xiao Long, et al. Epona: Autoregressive diffusion world model for autonomous driving. *arXiv preprint arXiv:2506.24113*, 2025.

[33] Guosheng Zhao, Xiaofeng Wang, Zheng Zhu, Xinze Chen, Guan Huang, Xiaoyi Bao, and Xingang Wang. DriveDreamer-2: LLM-enhanced world models for diverse driving video generation. In *AAAI*, pages 10412–10420, 2025.

[34] Wenzhao Zheng, Weiliang Chen, Yuanhui Huang, Borui Zhang, Yueqi Duan, and Jiwen Lu. OccWorld: Learning a 3D occupancy world model for autonomous driving. In *ECCV*, pages 55–72. Springer, 2024.

[35] Sicheng Zuo, Wenzhao Zheng, Yuanhui Huang, Jie Zhou, and Jiwen Lu. GaussianWorld: Gaussian world model for streaming 3D occupancy prediction. In *CVPR*, pages 6772–6781, 2025.

[36] Vlas Zyrianov, Xiyue Zhu, and Shenlong Wang. Learning to generate realistic LiDAR point clouds. In *ECCV*, pages 17–35. Springer, 2022.