VR-Drive: Viewpoint-Robust End-to-End Driving with Feed-Forward 3D Gaussian Splatting

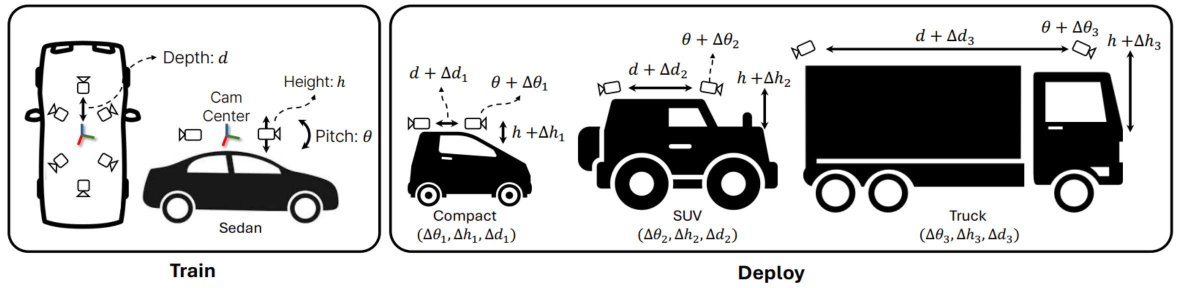
Hoonhee Cho*, Jae-Young Kang*, Giwon Lee*, Hyemin Yang*, Heejun Park, Seokwoo Jung, Kuk—Jin Yoon†

† Corresponding author

* Equal contribution

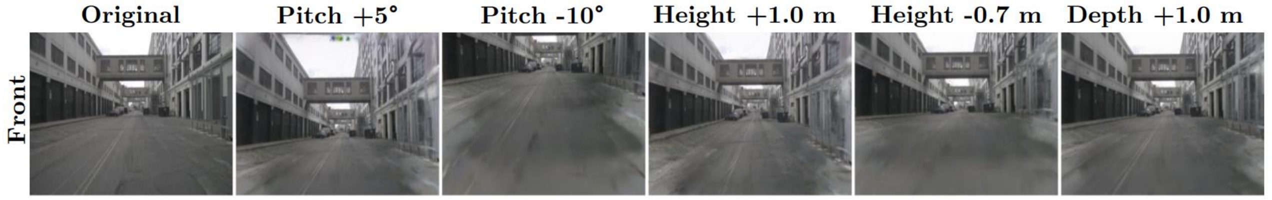
Project Page

Research Goal & Motivation



Viewpoint generalization for **end-to-end driving**, enabling robustness across diverse real-world camera viewpoints

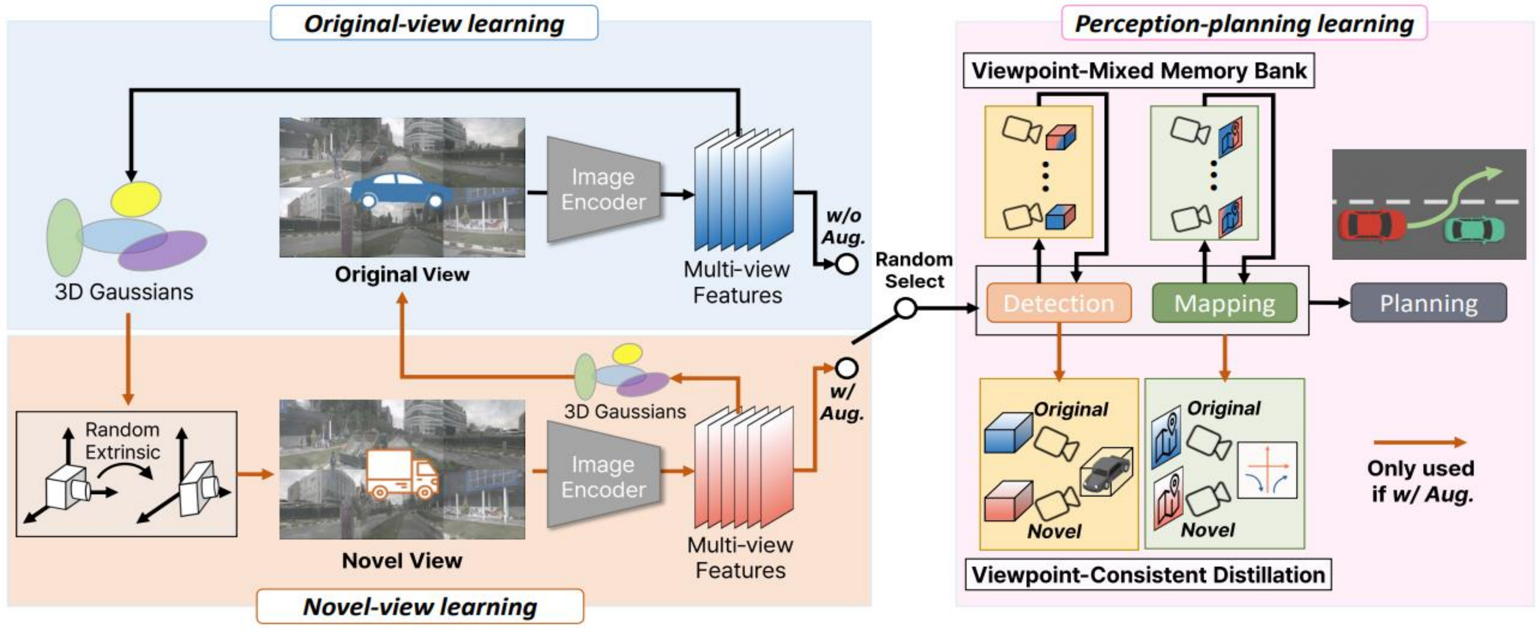
Benchmark with Viewpoint Variations



By optimizing 3D Gaussians offline at the scene level on the nuScenes dataset, we generate high-quality images from novel viewpoints.

Method

- Overall Framework

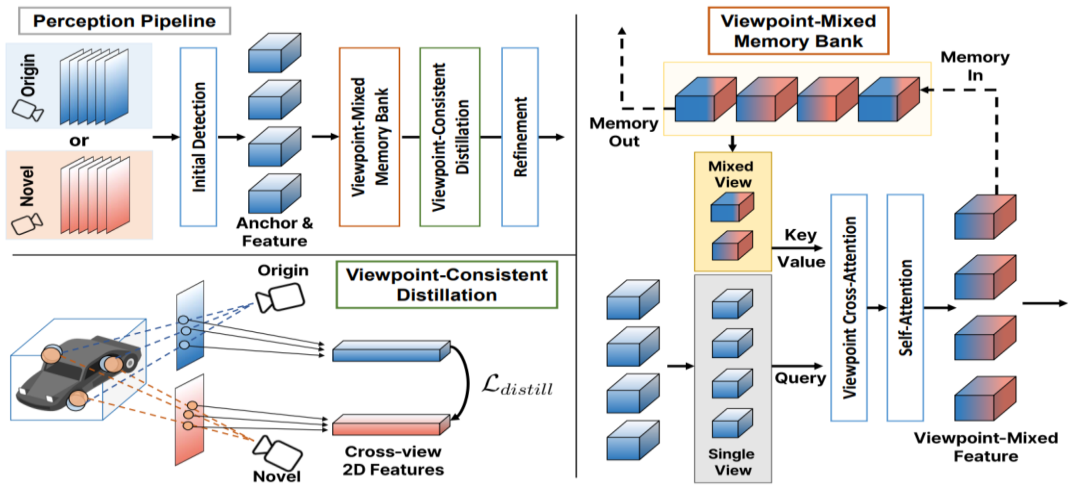


- Overall framework has three parts: (1) original-view learning, (2) novel-view learning, and (3) perception–planning learning.
- During novel-view learning, the perception–planning head is randomly assigned to the original or a novel view to promote viewpoint generalization.

- Proposed Modules

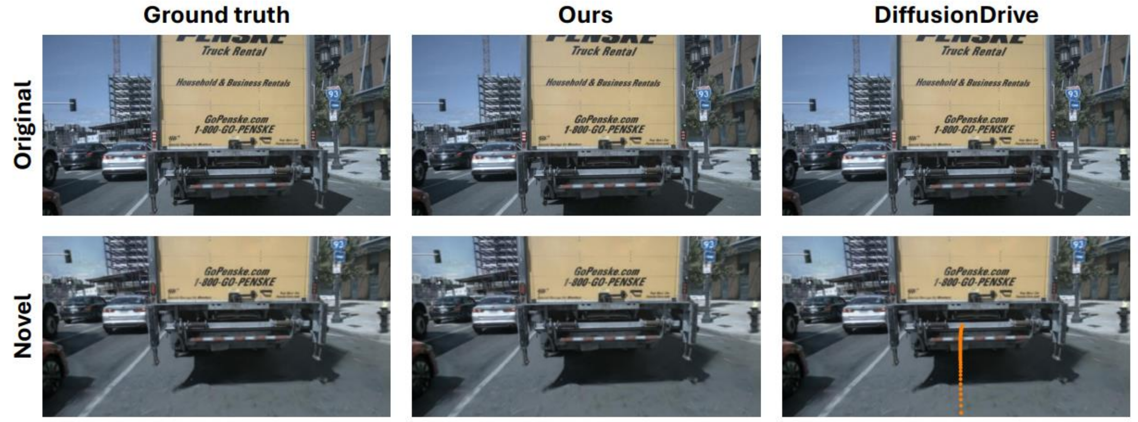
‣ **Viewpoint-Mixed Memory Bank:** Maintains diverse-view features and fuses them via cross-attention for stronger viewpoint generalization.

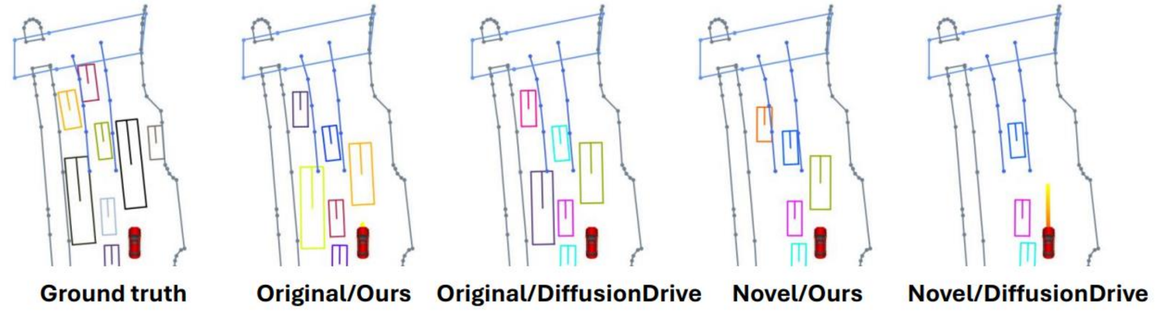
‣ **Viewpoint-Consistent Distillation:** Improves robustness to synthesized view artifacts by distilling reliable information from original views to novel views using anchor-based selective feature alignment



Experiment Results

- Qualitative Results





- Quantitative Results

Camera Setting		Methods	L2 (m) ↓ 1s	2s	3s	Avg.	Collision Rate (%) ↓ 1s	2s	3s	Avg.
Original		AD-MLP* [66]	**0.20**	**0.26**	**0.41**	**0.29**	0.17	0.18	0.24	0.19
		BEV-Planner* [33]	0.28	0.52	0.84	0.55	0.13	0.17	0.36	0.22
		VAD [27]	0.41	0.70	1.05	0.72	0.07	0.17	0.41	0.22
		SparseDrive [51]	0.29	0.58	0.96	0.61	**0.01**	0.05	0.18	0.08
		DiffusionDrive [35]	0.27	0.54	0.90	0.57	0.03	0.05	0.16	0.08
		VR-Drive (Ours)	0.29	0.57	0.95	0.60	**0.01**	**0.03**	**0.14**	**0.06**
Unseen	Average	AD-MLP* [66]	**0.20**	**0.26**	**0.41**	**0.29**	0.17	0.18	**0.24**	0.19
		BEV-Planner* [33]	0.28	0.54	0.88	0.57	0.36	0.42	0.63	0.47
		VAD [27]	0.46	0.75	1.11	0.78	0.20	0.37	0.69	0.42
		SparseDrive [51]	0.47	0.91	1.47	0.95	0.04	0.23	0.65	0.31
		DiffusionDrive [35]	0.62	1.14	1.76	1.17	0.07	0.36	0.80	0.41
		VR-Drive (Ours)	0.34	0.65	1.06	0.68	**0.01**	**0.07**	**0.24**	**0.11**

Open-loop Evaluation

Methods	Original		Unseen pitch +5°		pitch -10°		height +1.0 m		height -0.7 m		depth +1.0 m		**Average**	
	DS	RC	DS	RC	DS	RC	DS	RC	DS	RC	DS	RC	DS	RC
Town05-*Nov*														
ST-P3 [20]	44.24	100.00	41.00	100.00	23.85	100.00	25.83	100.00	28.60	100.00	32.06	100.00	30.27	100.00
TCP [62]	92.73	92.73	70.33	80.33	4.65	4.65	88.51	88.51	0.00	0.00	91.11	91.11	50.92	52.92
AD-MLP [66]	13.59	32.83	13.59	32.83	13.59	32.83	13.59	32.83	13.59	32.83	13.59	32.83	13.59	32.83
BEV-Planner [33]	17.25	28.70	7.30	28.89	7.74	28.83	8.51	28.95	7.69	28.70	7.75	28.95	7.80	28.86
Baseline	76.47	99.20	69.41	89.60	45.65	99.38	48.67	100.00	41.59	86.76	35.95	98.60	48.25	94.87
VR-Drive (Ours)	84.04	99.04	75.00	100.00	91.26	98.76	98.44	98.99	80.67	97.32	95.88	96.35	88.25	98.28

Closed -loop Evaluation