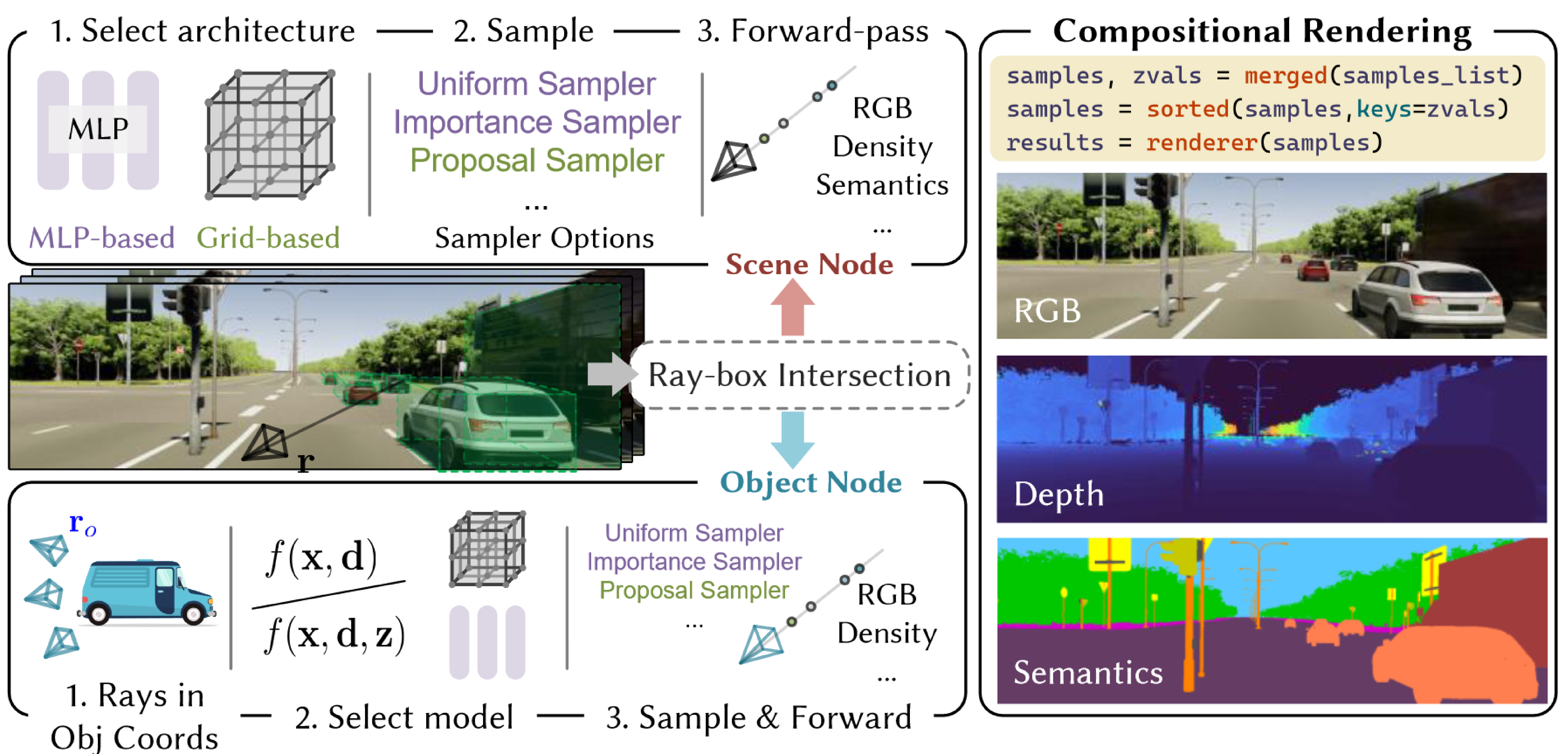


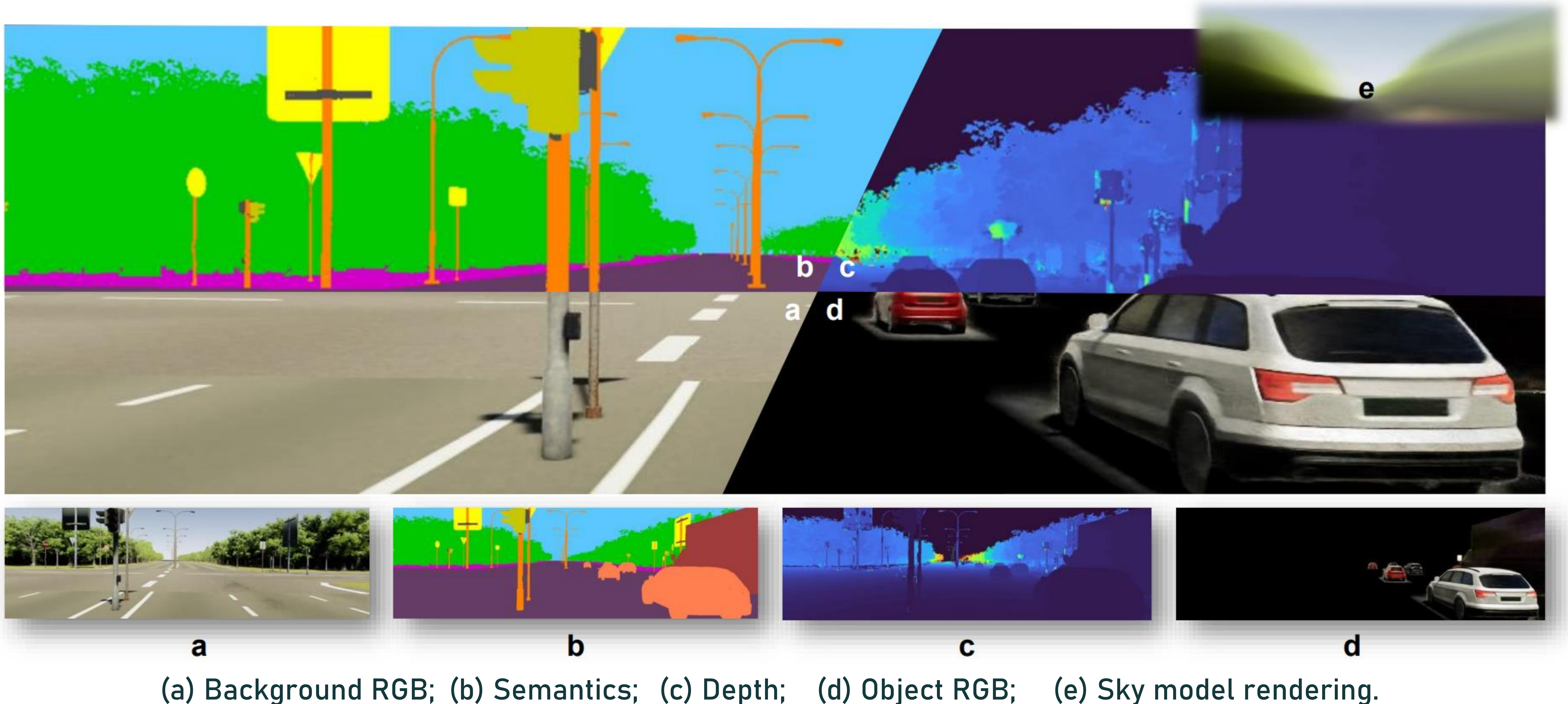
MARS: An Instance-aware, Modular and Realistic Simulator for Autonomous Driving

Zirui Wu, Tianyu Liu, Liyi Luo, Zhide Zhong, Jianteng Chen, Hongmin Xiao, Chao Hou, Haozhe Lou, Yuantao Chen, Runyi Yang, Yuxin Huang, Xiaoyu Ye, Zike Yan, Yongliang Shi, Yiyi Liao, and Hao Zhao*

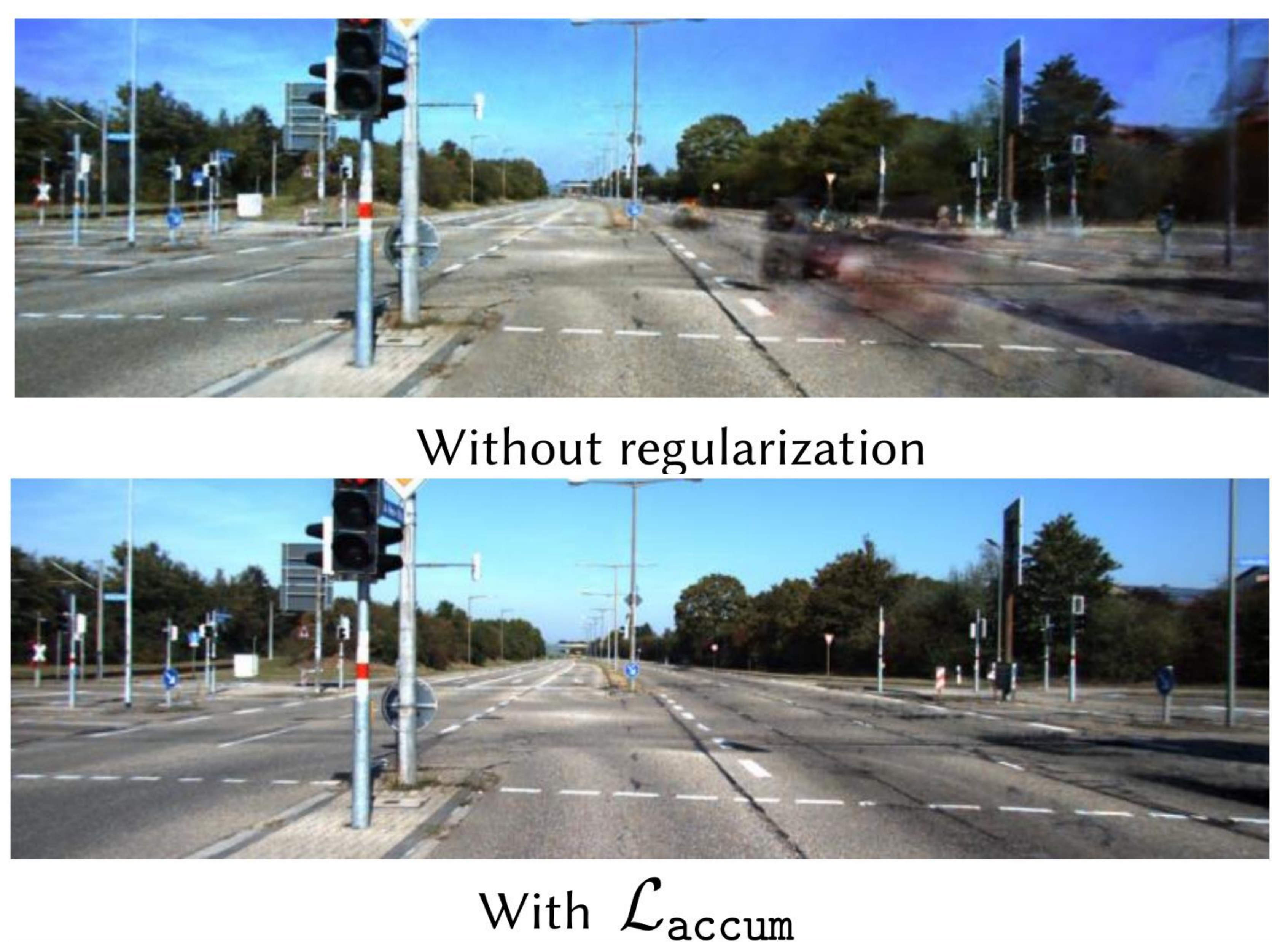
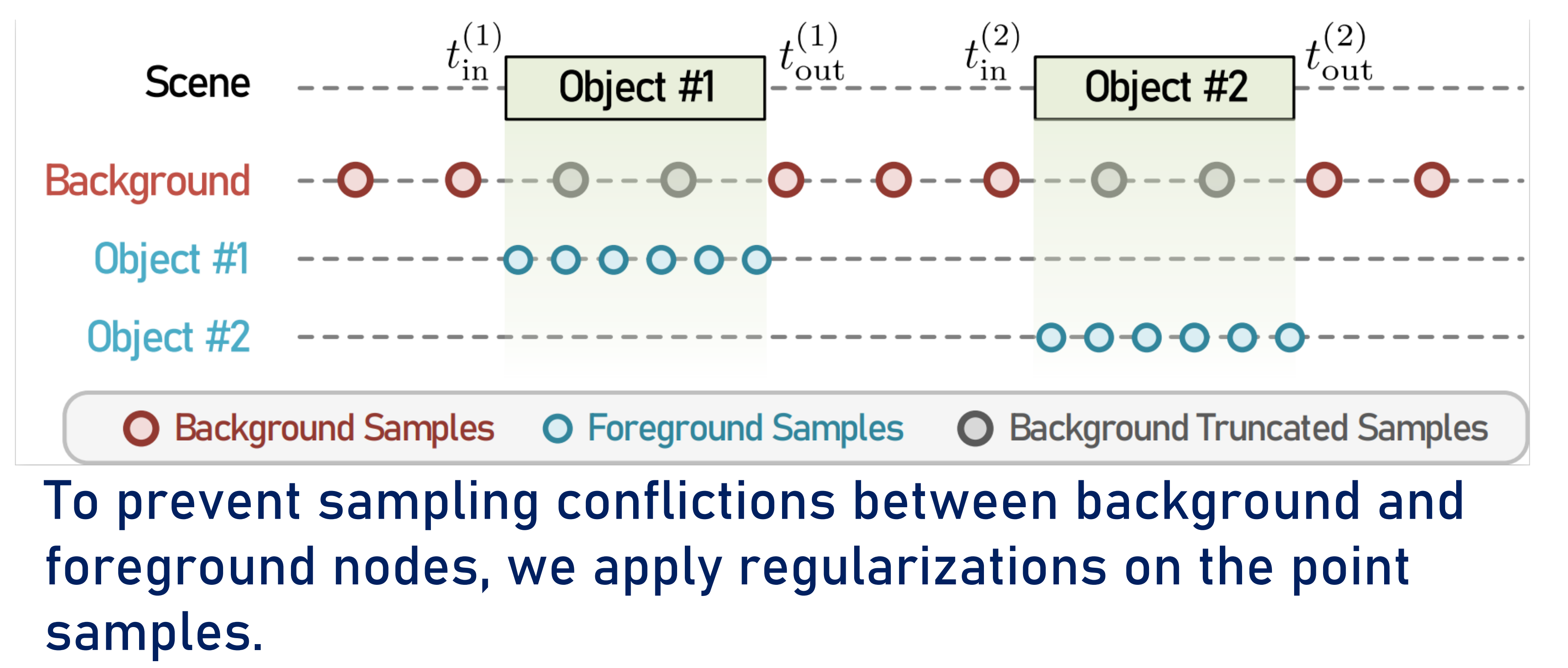
Abstract



We propose an autonomous driving simulator based upon neural radiance fields (NeRFs). Compared with existing works, ours has three notable features: (1) Instance-aware. (2) Modular. (3) Realistic. Our simulator achieves state-of-the-art photorealism results given the best module selection. Our simulator will be open-sourced while most of our counterparts are not.



Sample



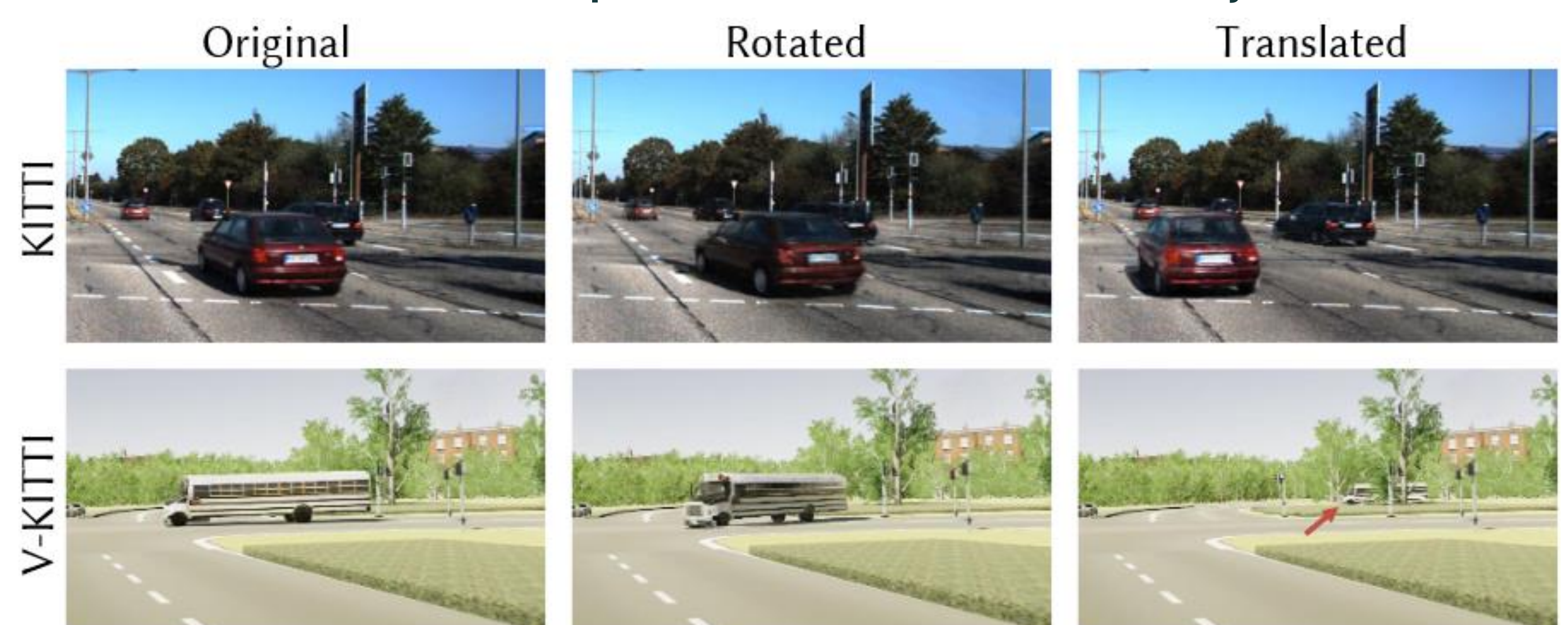
Ablation study on the sampling regularization.

	NeRF [16]	NeRF+Time	NSG [19]	PNF [13]	SUDS [24]	Ours
PSNR \uparrow	23.34	24.18	26.66	27.48	28.31	29.06
SSIM \uparrow	0.662	0.677	0.806	0.870	0.876	0.885
Instance-aware	×	×	✓	×	×	✓
Modular	×	×	×	×	×	✓
Open-sourced	✓	-	✓	×	✓	✓

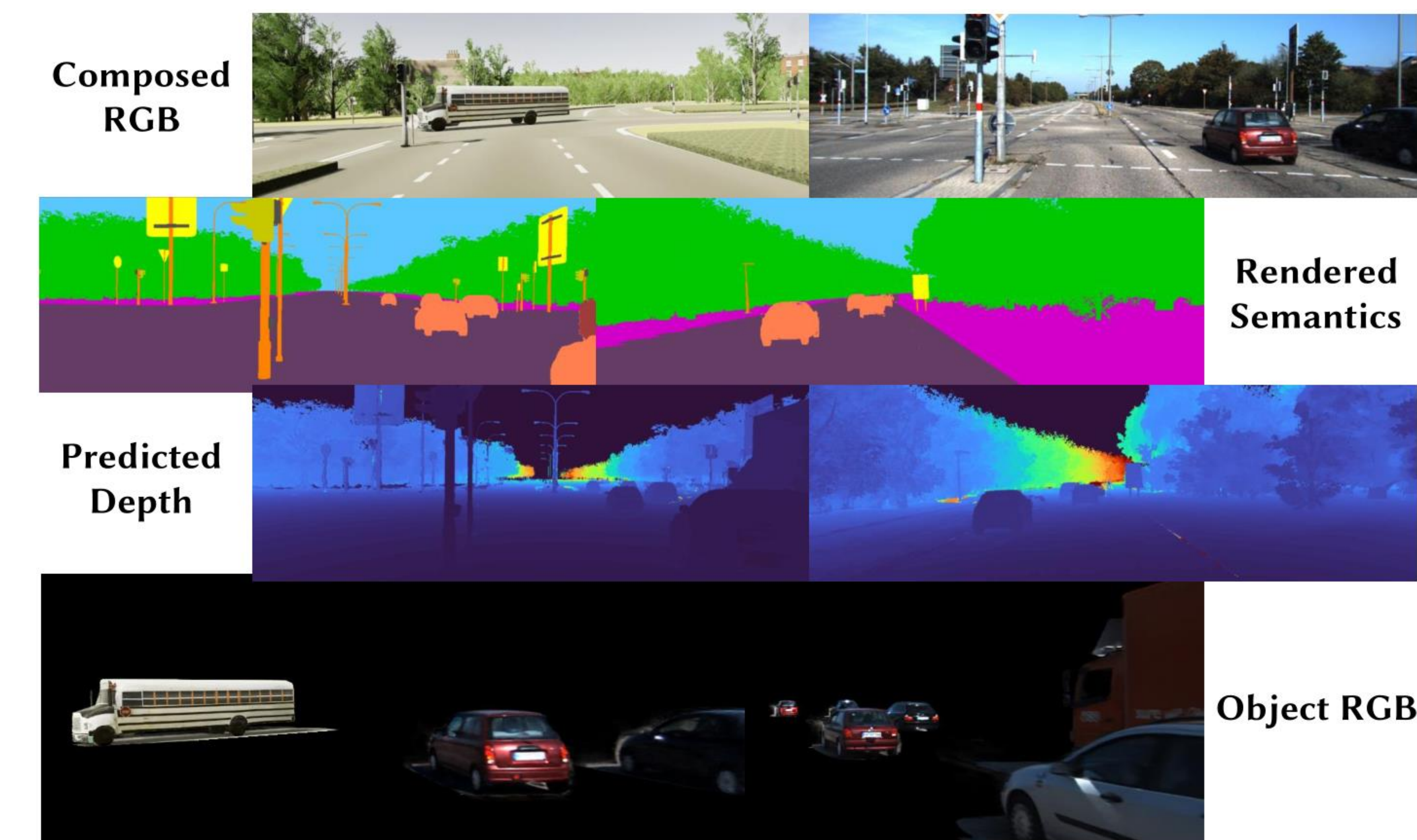
Quantitative comparisons on KITTI image reconstruction & Setting differences with baseline methods.

	KITTI-75%			KITTI-50%			KITTI-25%		
	PSNR \uparrow	SSIM \uparrow	LPIPS \downarrow	PSNR \uparrow	SSIM \uparrow	LPIPS \downarrow	PSNR \uparrow	SSIM \uparrow	LPIPS \downarrow
NeRF [16]	18.56	0.557	0.554	19.12	0.587	0.497	18.61	0.570	0.510
NeRF+Time	21.01	0.612	0.492	21.34	0.635	0.448	19.55	0.586	0.505
NSG [19]	21.53	0.673	0.254	21.26	0.659	0.266	20.00	0.632	0.281
SUDS [24]	22.77	0.797	0.171	23.12	0.821	0.135	20.76	0.747	0.198
Ours	24.23	0.845	0.160	24.00	0.801	0.164	23.23	0.756	0.177
	+1.46	+0.048	-0.011	+0.88	-0.020	+0.029	+2.47	+0.009	-0.021
	VKITTI-75%			VKITTI-50%			VKITTI-25%		
	PSNR \uparrow	SSIM \uparrow	LPIPS \downarrow	PSNR \uparrow	SSIM \uparrow	LPIPS \downarrow	PSNR \uparrow	SSIM \uparrow	LPIPS \downarrow
NeRF [16]	18.67	0.548	0.634	18.58	0.544	0.635	18.17	0.537	0.644
NeRF+Time	19.03	0.574	0.587	18.90	0.565	0.610	18.04	0.545	0.626
NSG [19]	23.41	0.689	0.317	23.23	0.679	0.325	21.29	0.666	0.317
SUDS [24]	23.87	0.846	0.150	23.78	0.851	0.142	22.18	0.829	0.160
Ours	29.79	0.917	0.088	29.63	0.916	0.087	27.01	0.887	0.104
	+5.92	+0.071	-0.062	+5.85	+0.065	-0.055	+4.83	+0.058	-0.056

Quantitative comparisons on novel-view synthesis.



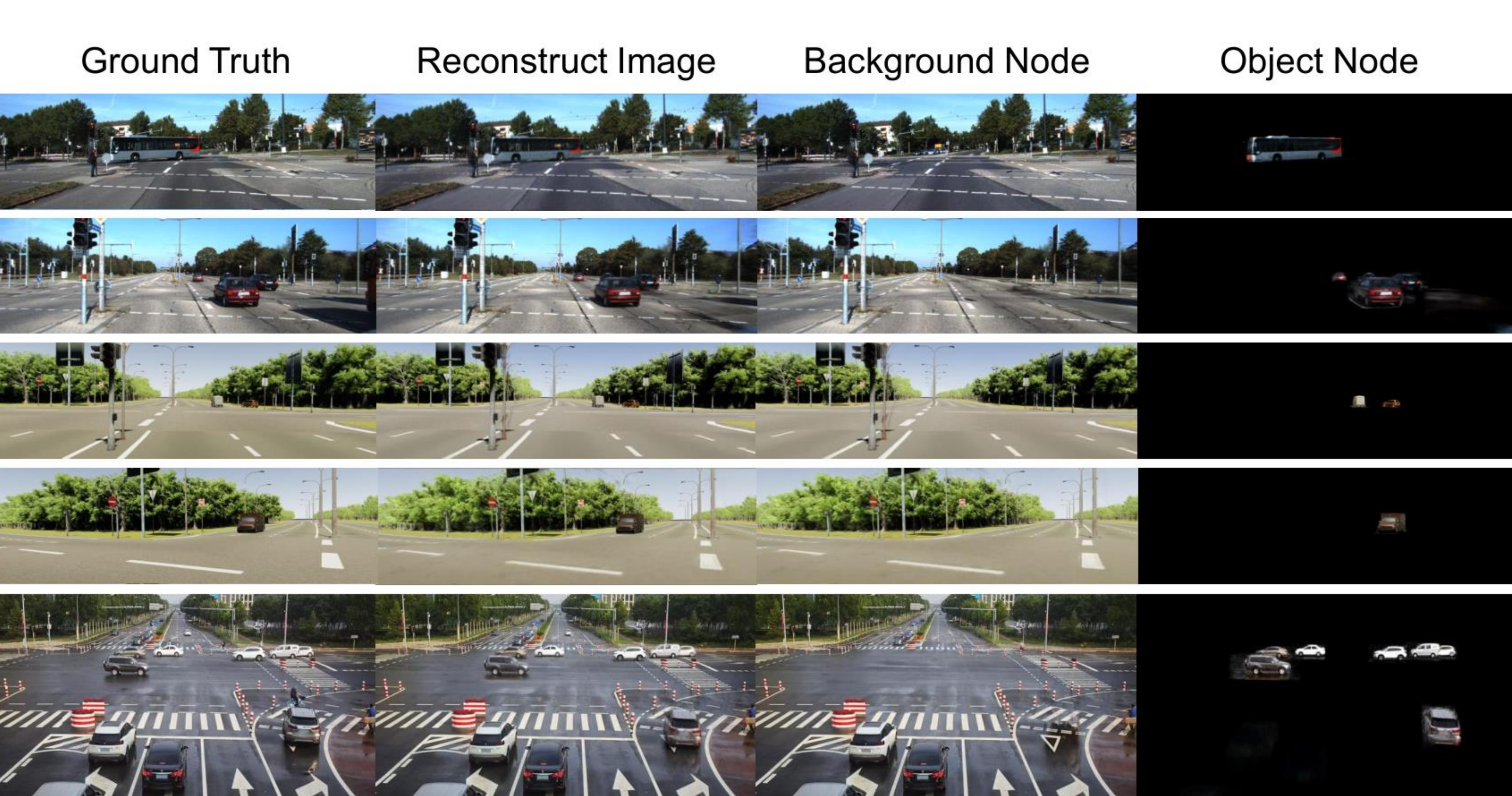
Editing results.



Rendering results on each channel.

Rendering & Editing

Our framework separately models background and foreground nodes, which allows us to separately render and edit the scene in an instance-aware manner.



Qualitative results on KITTI / V-KITTI / DAIR-V2X datasets.

