

DiscoScene: Spatially Disentangled Generative Radiance Fields for Controllable 3D-aware Scene Synthesis

CVPR 2023 Highlight, TUE-PM-026

Yinghao Xu, Menglei Chai, Zifan Shi, Sida Peng,
Ivan Skorokhodov, Aliaksandr Siarohin, Ceyuan Yang, Yujun Shen,
Hsin-Ying Lee, Bolei Zhou, Sergey Tulyakov

CUHK



Snap Inc.



HKUST



ZJU



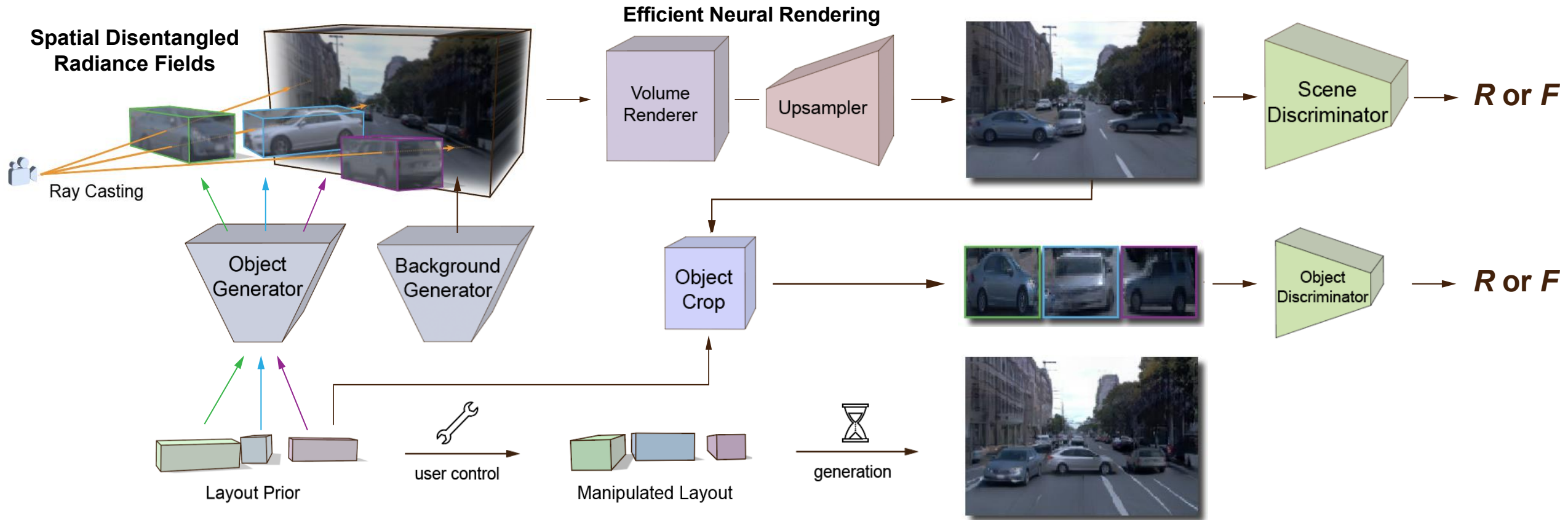
KAUST



UCLA

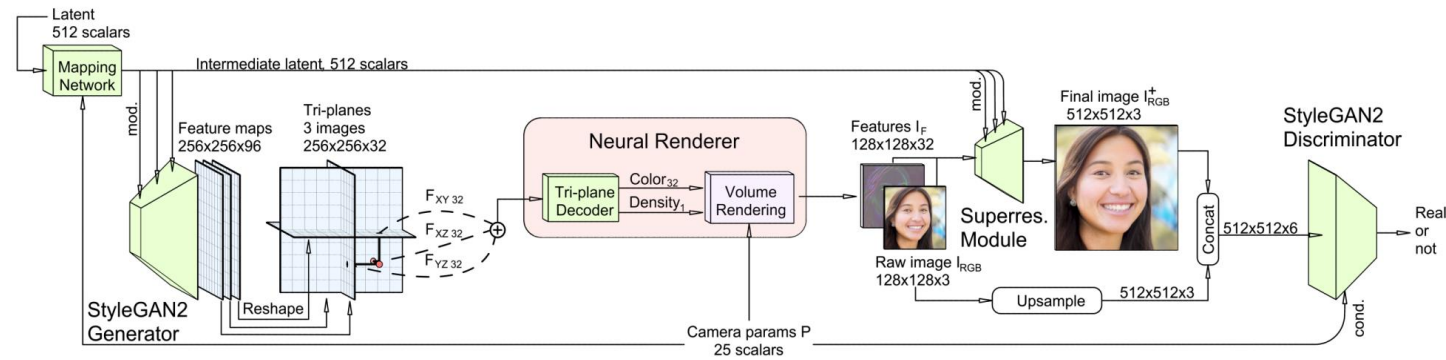


Overview



3D GANs for Object-Centric Dataset

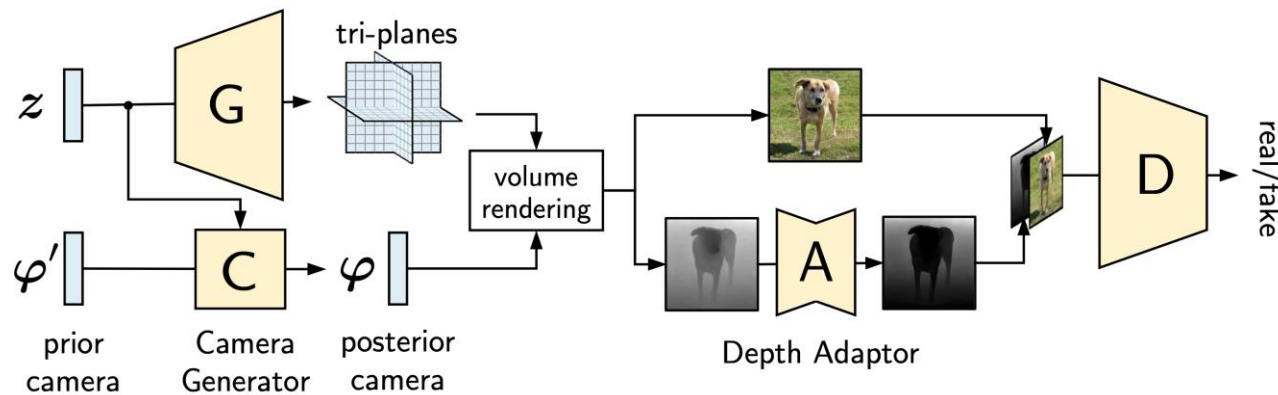
EG3D [1]



well-curated, aligned datasets



3DGP [2]



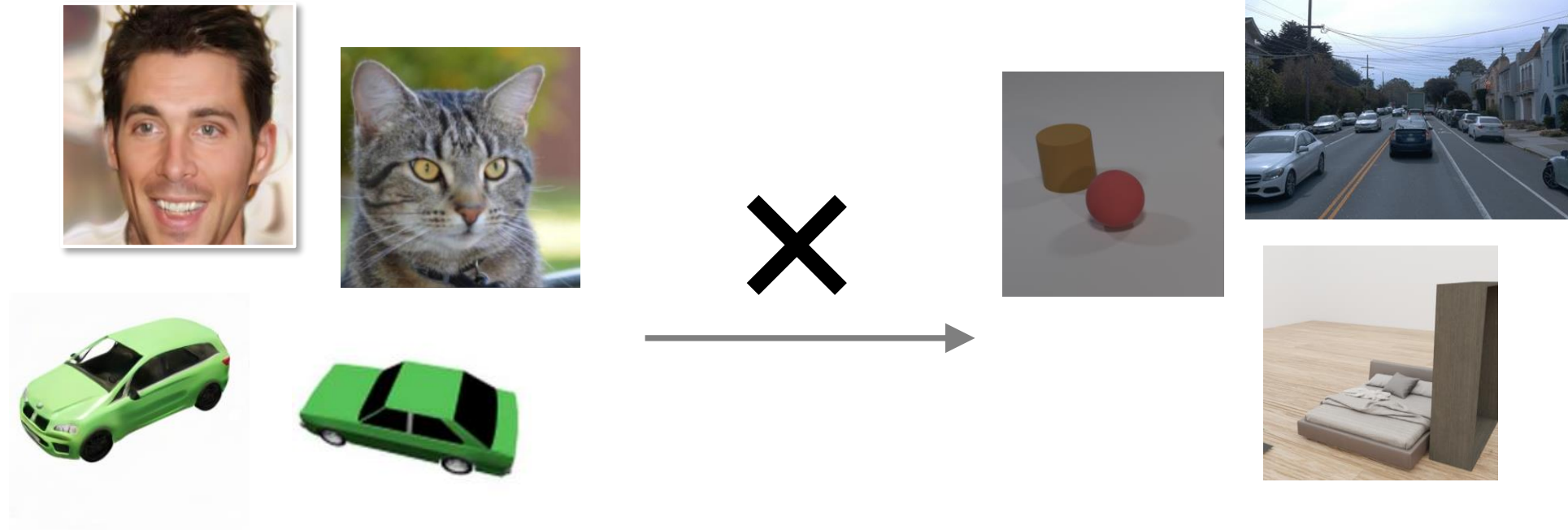
in-the-wild, non-aligned datasets



[1] Eric et al., "Efficient Geometry-aware 3D Generative Adversarial Networks", CVPR 2022

[2] Ivan et al., "Learning to Recover 3D Scene Shape from a Single Image", CVPR 2021

Extending 3D GAN to Complex Scenes



Single Object

Complex Scenes

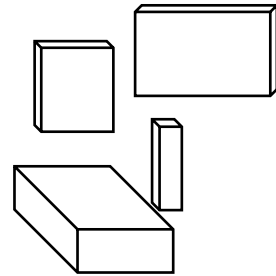
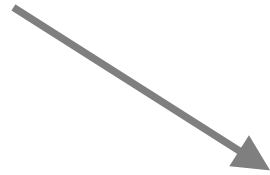
How can we scale up 3D GANs to **scenes** with complex layout and multiple object?

Scene Representation

Empty room



Furniture

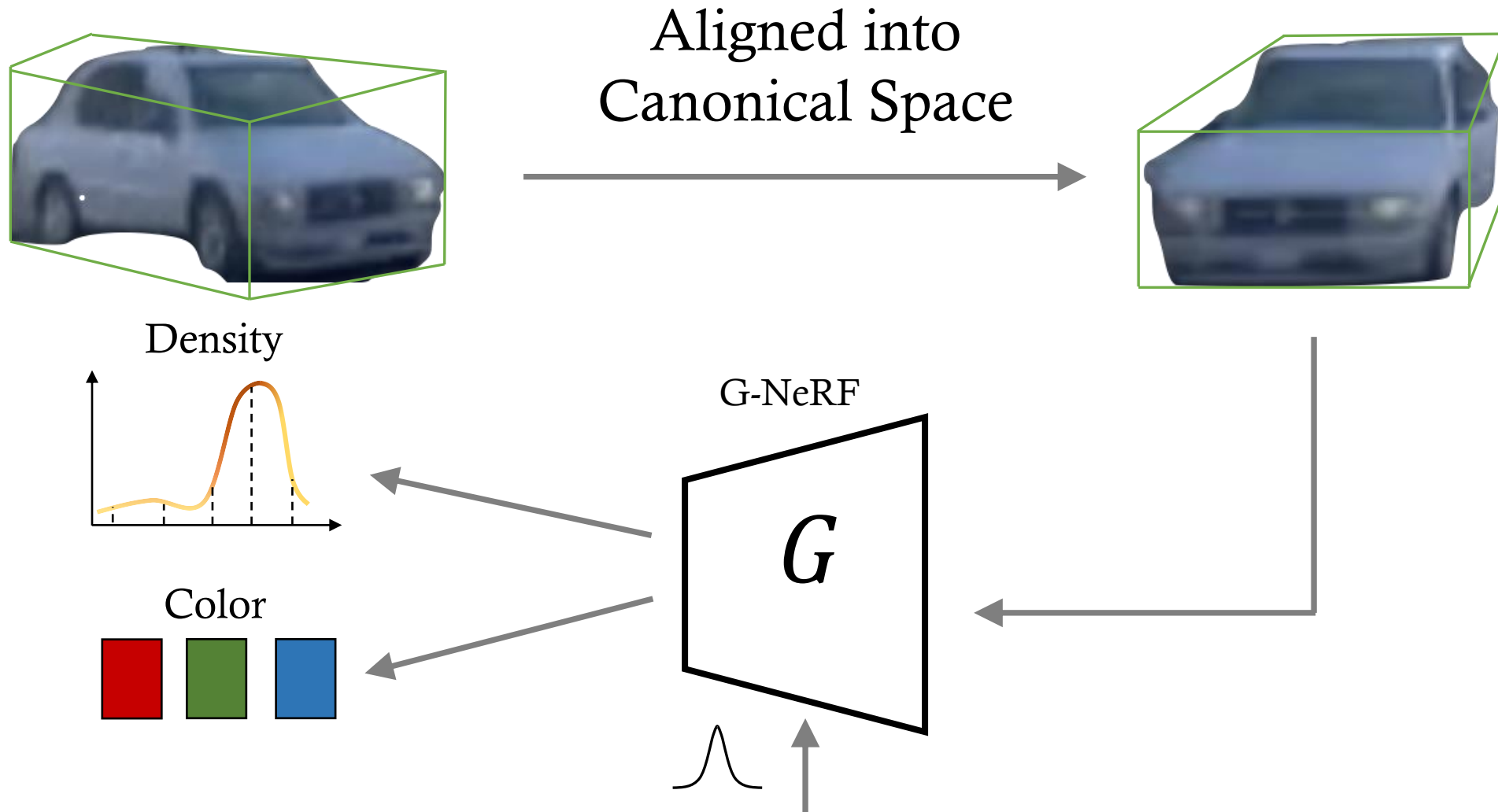


Layout



We should leverage **Layout Prior!**

Object Modeling

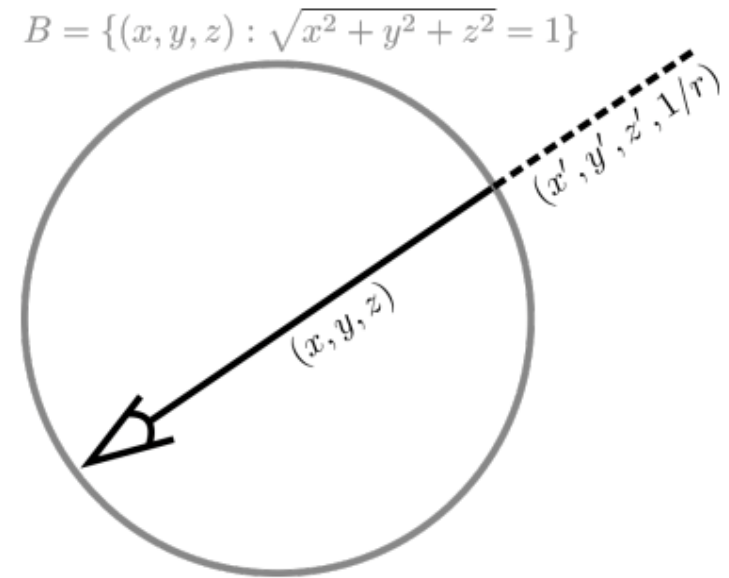


Background Modeling



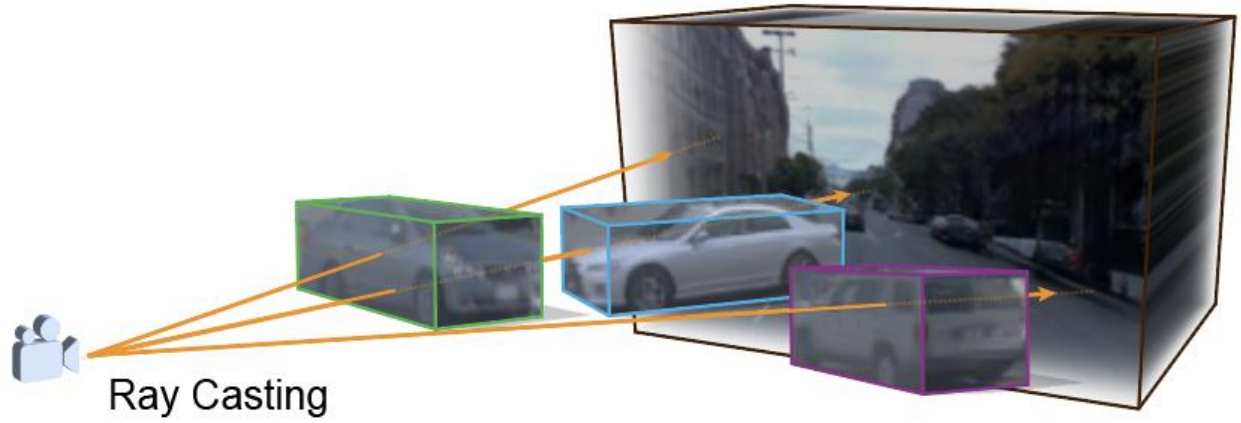
Background

Evaluated in
World Space

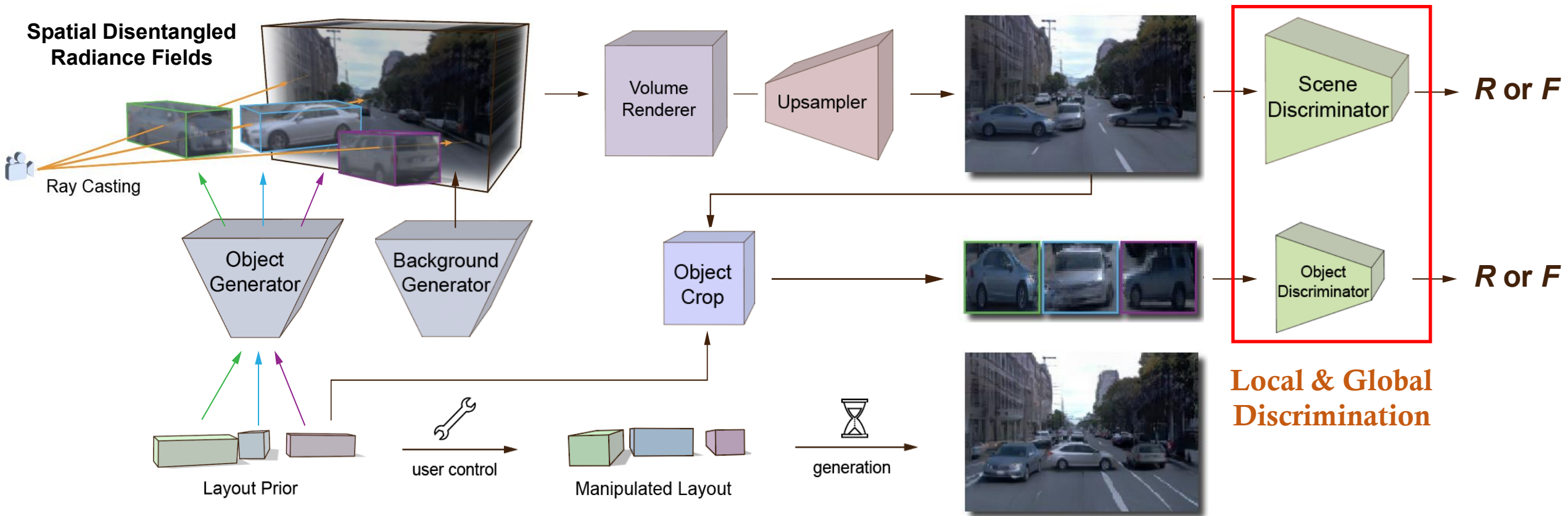


Efficient Rendering

- ▶ Points Sampling:
 - ▶ Object: Ray-Box intersection
 - ▶ Background: NeRF++
- ▶ Composition
 - ▶ Sorting depth for occlusion
 - ▶ Background is behind objects



DiscoScene

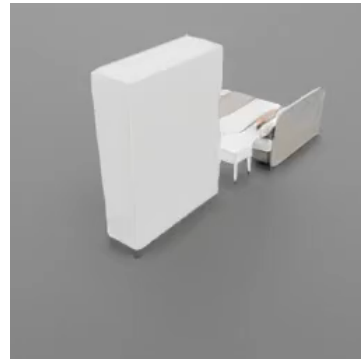
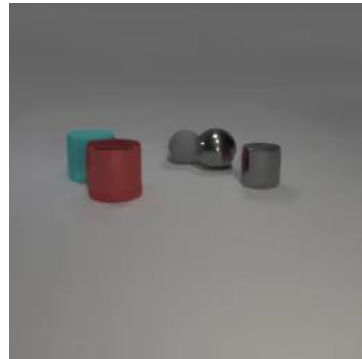
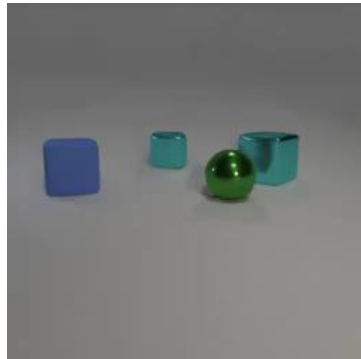
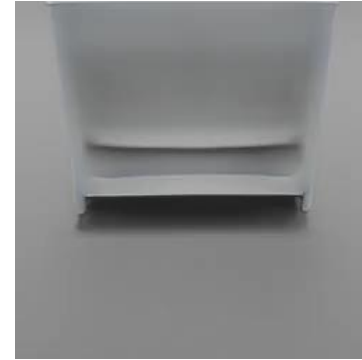
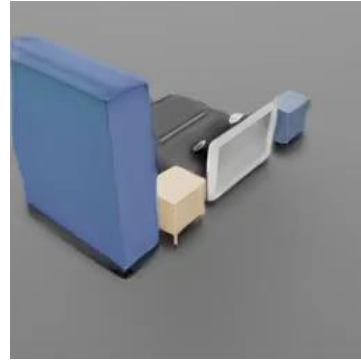
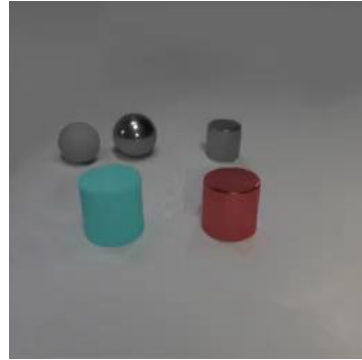
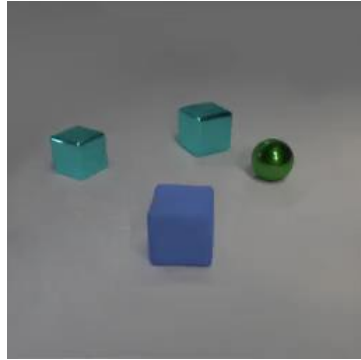


Quantitative Evaluation

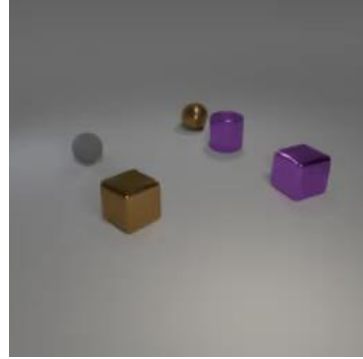
Model	CLEVR				3D-FRONT		WAYMO	
	FID ↓	KID ↓	TR. ↓	INF. ↓	FID ↓	KID ↓	FID ↓	KID ↓
StyleGAN2 [25]	4.5	3.0	13.3	44	12.5	4.3	15.1	8.3
EpiGRAF [45]	10.4	8.3	16.0	114	107.2	102.3	27.0	26.1
VolumeGAN [59]	7.5	5.1	15.2	90	52.7	38.7	29.9	18.2
EG3D [4]	4.1	12.7	25.8	55	19.7	13.5	26.0	45.4
GIRAFFE [31]	78.5	61.5	5.2	62	56.5	46.8	175.7	212.1
GSN [8]	–	–	–	–	130.7	87.5	–	–
DisCoScene	3.5	2.1	18.1	95	13.8	7.4	16.0	8.4

TR. and INF. denote training and inference costs evaluated in *V100days* and *ms/image*.

Explicit Camera Control



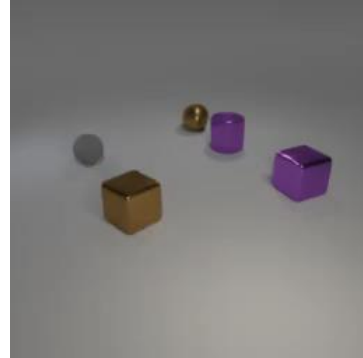
Object Arrangement



Rotation



Translation



Object Removal / Insertion

Removal

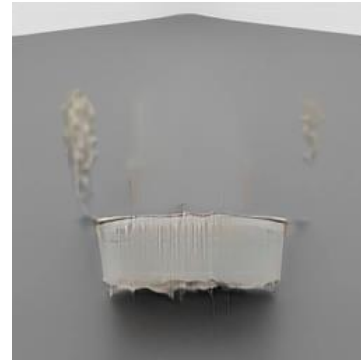
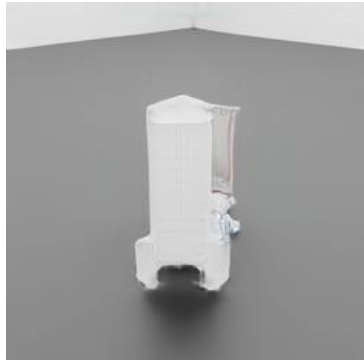


Insertion

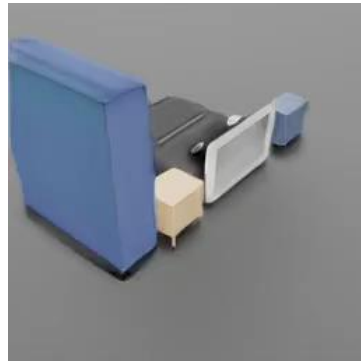


Comparison with EG3D

EG3D

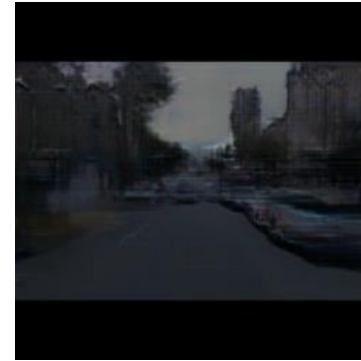
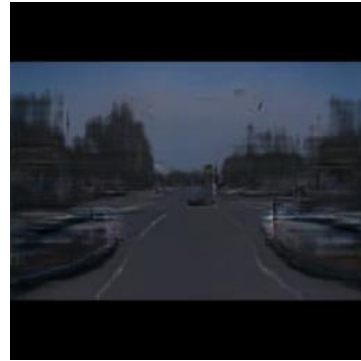
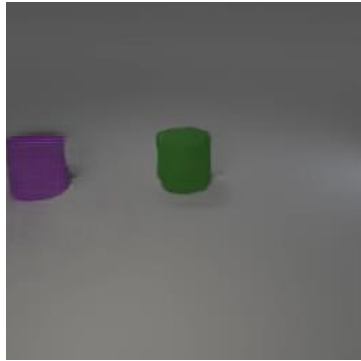
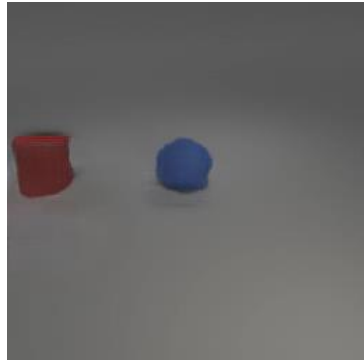


Ours

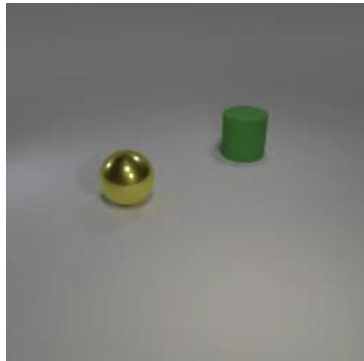


Comparison with GIRAFFE

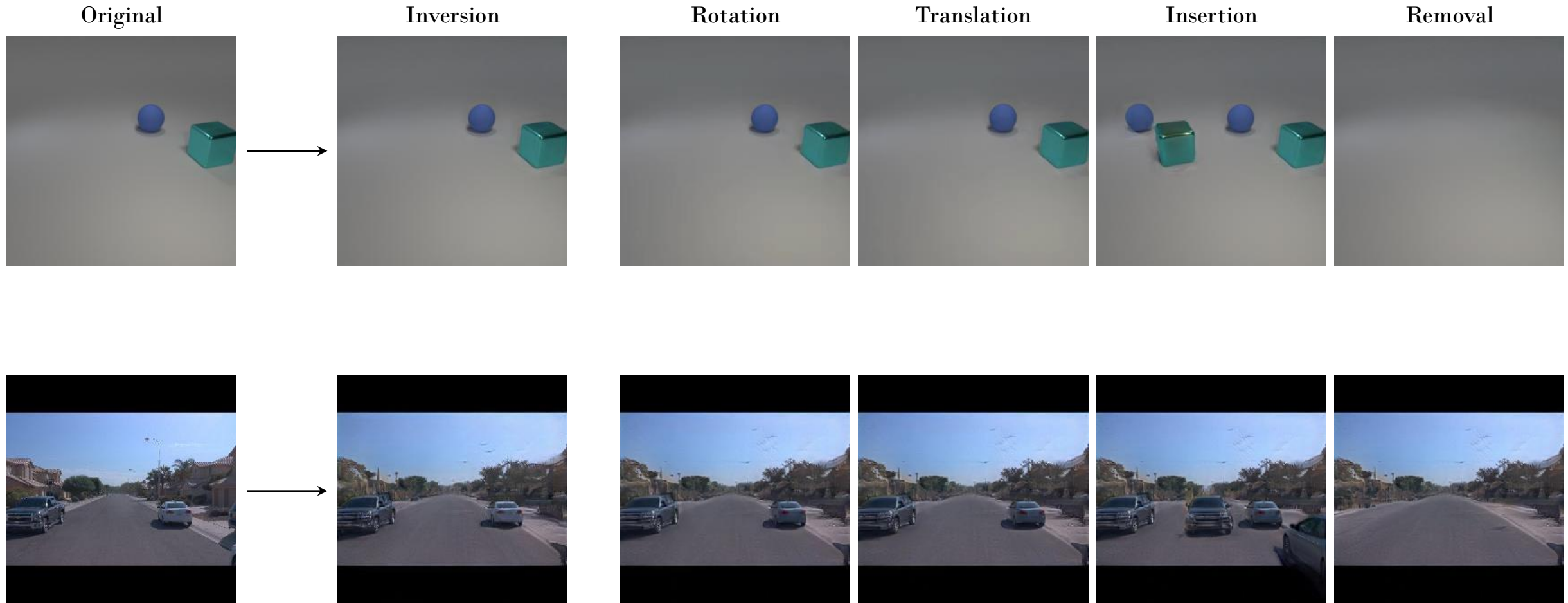
GIRAFFE



Ours



Real Scene Editing



Thanks!