CC3D: Layout-Conditioned Generation of Compositional 3D Scenes

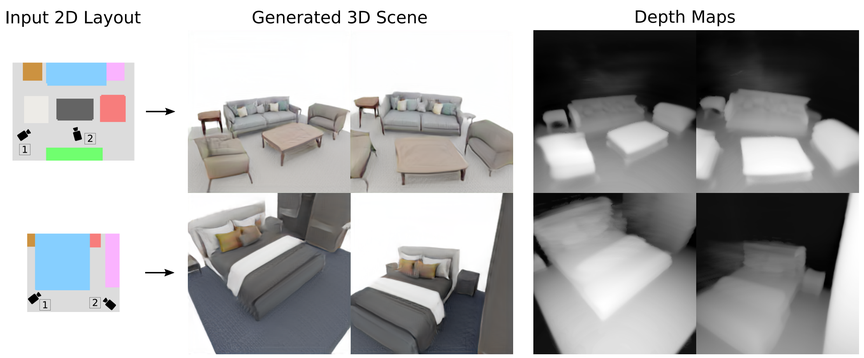
Sherwin Bahmani*[1], Jeong Joon Park*[2], Despoina Paschalidou[2], Xingguang Yan[4],
Gordon Wetzstein[2], Leonidas Guibas[2], Andrea Tagliasacchi[1,3,4]

* Equal contribution

[1] University of Toronto [2] Stanford University [3]Google Research [4]Simon Fraser University

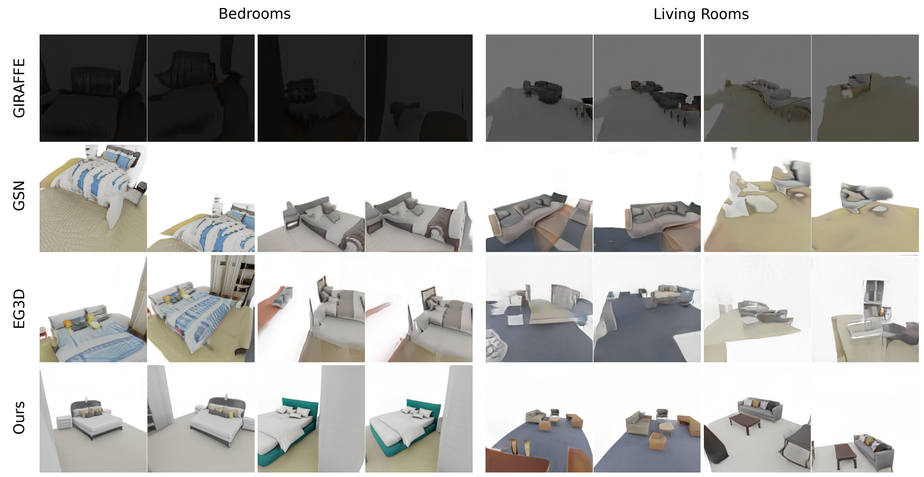
Motivation

- Recent progress mainly on 3D object generation
- Current single object methods fail when applied on unaligned complex scenes
- We tackle the task of compositional 3D scene generation from 2D layout priors



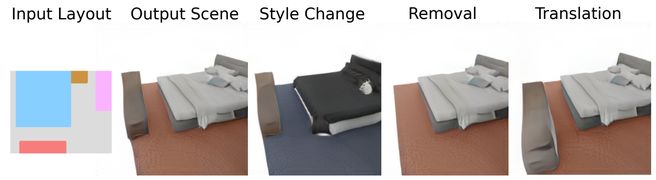
3D-FRONT

We show realistic 3D indoor scene synthesis in comparison to previous approaches



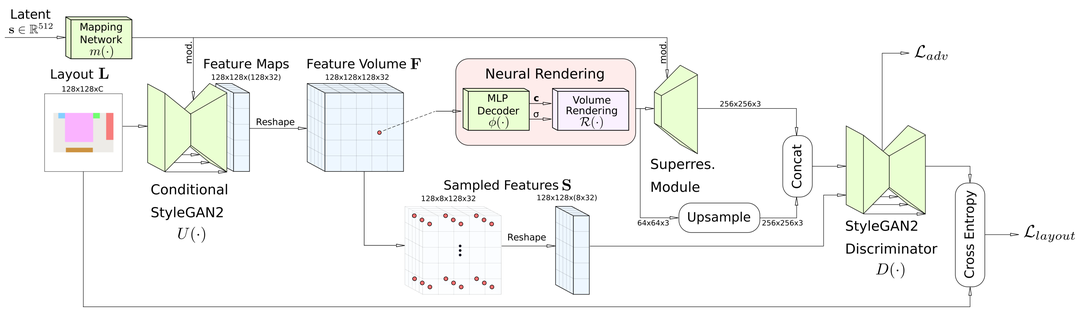
3D-FRONT editing

We can edit generated scenes by adjusting the input layout or latent code



Method

- Given a semantic 2D scene layout, we condition a style-based 2D generator
- We extrude a feature volume that can be queried with trilinear interpolation
- Following existing 3D GANs, we render an image for a given camera pose
- We train our pipeline on a combination of adversarial and layout consistency loss
- There is no multi-view supervision required



KITTI-360

Our method generalizes to outdoor scenes and achieves higher quality trajectories

3D-FRONT depth

Compositionality leads to high-quality depth maps