Pyramid Diffusion for Fine 3D Large Scene Generation

Yuheng Liu[1,2*], Xinke Li[3*], Xueting Li[4], Lu Qi[5†], Chongshou Li[1], Ming-Hsuan Yang[5,6]

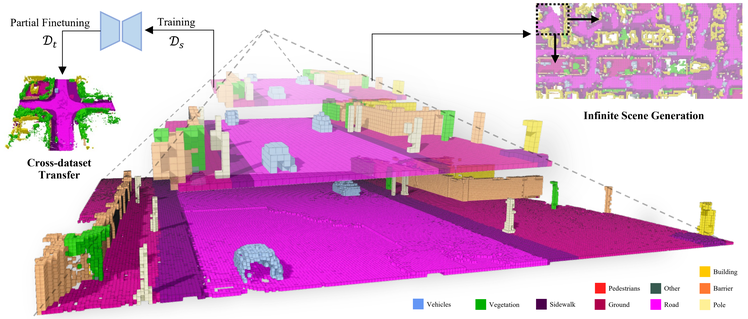
1 Southwest Jiaotong University 2 University of Leeds 3 City University of Hong Kong 4 NVIDIA 5 UC Merced 6 Yonsei University

ECCV EUROPEAN CONFERENCE ON COMPUTER VISION MILANO 2024

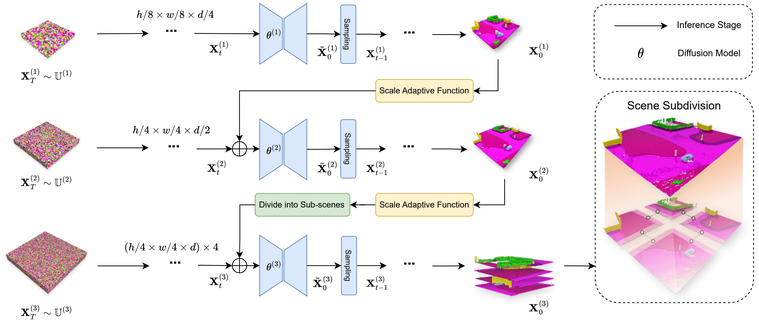
Code & Visual Results

Contributions

- Implement a **coarse-to-fine** strategy for 3D outdoor scene generation via designing a novel pryamid diffusion model.
- Conduct extensive experiments on PDD, demonstrating its generation of **high quality** 3D scenes.
- Introduce **new metrics** to evaluate the quality of 3D scene generation from various perspectives.
- Showcase broader applications: **cross dataset generation** and **infinite scene generation**.

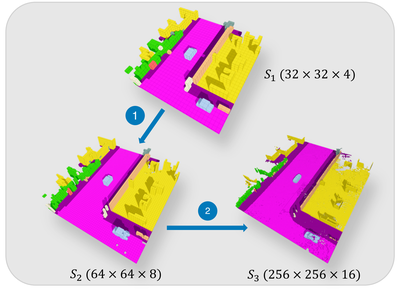


Pyramid Discrete Diffusion Structure



Unconditional Generation

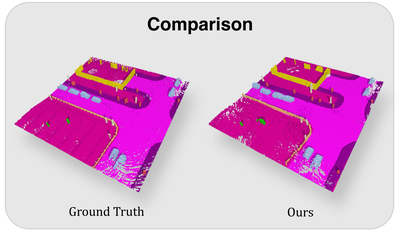
Comparasion of **PDD** with baseline models: **DiscreteDiff** and **LatentDiff**



- Compared to baseline models, our method generates scenes with **better semantic accuracy**.
- Our approach produces more **diverse** and **random** 3D scenes, containing **more objects and intricate details**.

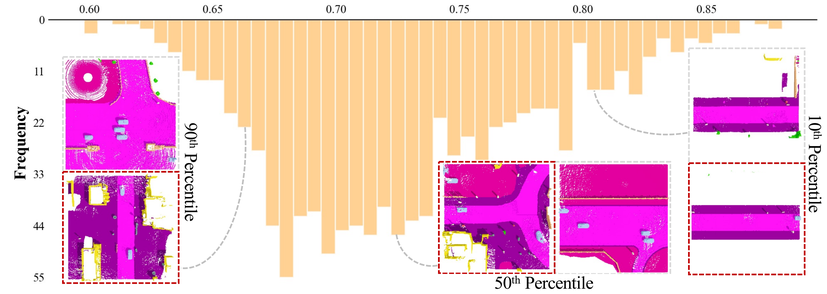
Conditional Generation

Baseline: GT and **Use Point Clouds** as structural guidance

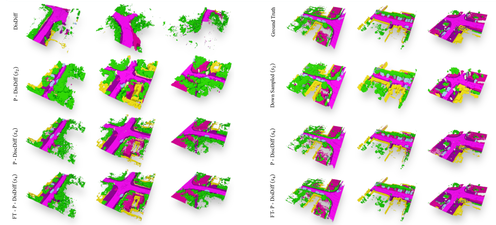


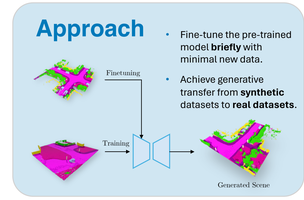
- Use point cloud as guidance can produce structurally precise scenes that align well with the point clouds. However, the generated scenes exhibit many **semantic inconsistencies** compared to the GT.
- Sences generated by our approach exhibit **high similarity** to the GT, indicating that our model has good **reconstruction capabilities**.

None-overfitting Verification

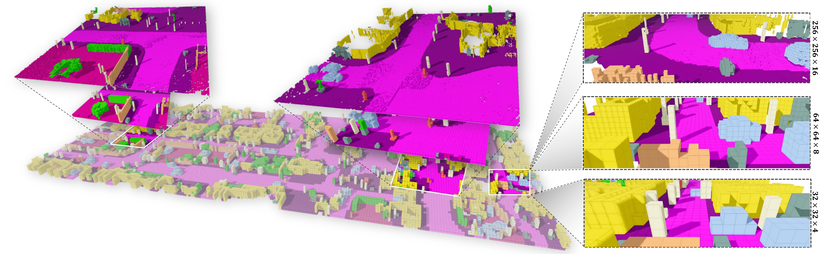


Cross-dataset Generation





Infinite-scene Generation



- We adopt the same approach as the Scene Subdivision Module, quickly **generating** infinite scenes at a low scale and then **refining** the generated scenes.
- Scan the QR code to view our infinite scenes.

Infinite Demo Video

Youtube