Panoptic NeRF: 3D-to-2D Label Transfer for Panoptic Urban Scene Segmentation

Motivation

How?

Panoptic Annotation

1.5h / frame
Expensive
Per-frame 2D Annotation
Labeled Scene

0.75min / frame
Cheap
Coarse 3D Annotation

Introduction

Idea: Infer in 3D space to generate 2D panoptic labels via volume rendering

Improve Underlying Geometry

Method Overview

Fixed Semantic Field $s: x \in \mathbb{R}^3 \rightarrow s \in \mathbb{R}^M$.

Learned Semantic Field $s: x \in \mathbb{R}^3 \rightarrow s \in \mathbb{R}^M$.

Radiance Field $f_r: (x, d \in \mathbb{R}^3) \rightarrow (r \in \mathbb{R}^3, e \in \mathbb{R}^M)$.

Instance Field $f_l: x \in \mathbb{R}^3 \rightarrow I \in \mathbb{R}^M$.

Resolve Label Ambiguity

Challenges

• Poor Geometry - With sparse inputs views in driving scenario, vanilla NeRF suffers from poor geometry reconstruction that leads to wrong panoptic maps
• Label Ambiguity - At overlapping regions of the 3D bounding primitives, label ambiguity yields inaccurate 2D labels

Contributions

• End-to-end 3D-to-2D label transfer in terms of volume rendering
• State-of-the-art performance compared to existing label transfer methods
• Enable rendering semantic/instance labels at novel viewpoints

Experiments on KITTI-360

Panoptic NeRF

Based on improved geometry, learned semantic field further resolves label ambiguity at overlapping regions of 3D bounding primitives via 2D and 3D weak supervisions

• 2D pseudo-label usually performs well on frequently occurring classes, 3D supervision further suppresses noise of 2D predictions

Panoptic NeRF*