Panoptic Neural Fields: A Semantic Object-Aware Neural Scene Representation

Abhijit Kundu, Kyle Genova, Xiaoqi Yin, Alireza Fathi, Caroline Pantofaru, Leonidas Guibas, Andrea Tagliasacchi, Frank Dellaert, Thomas Funkhouser.
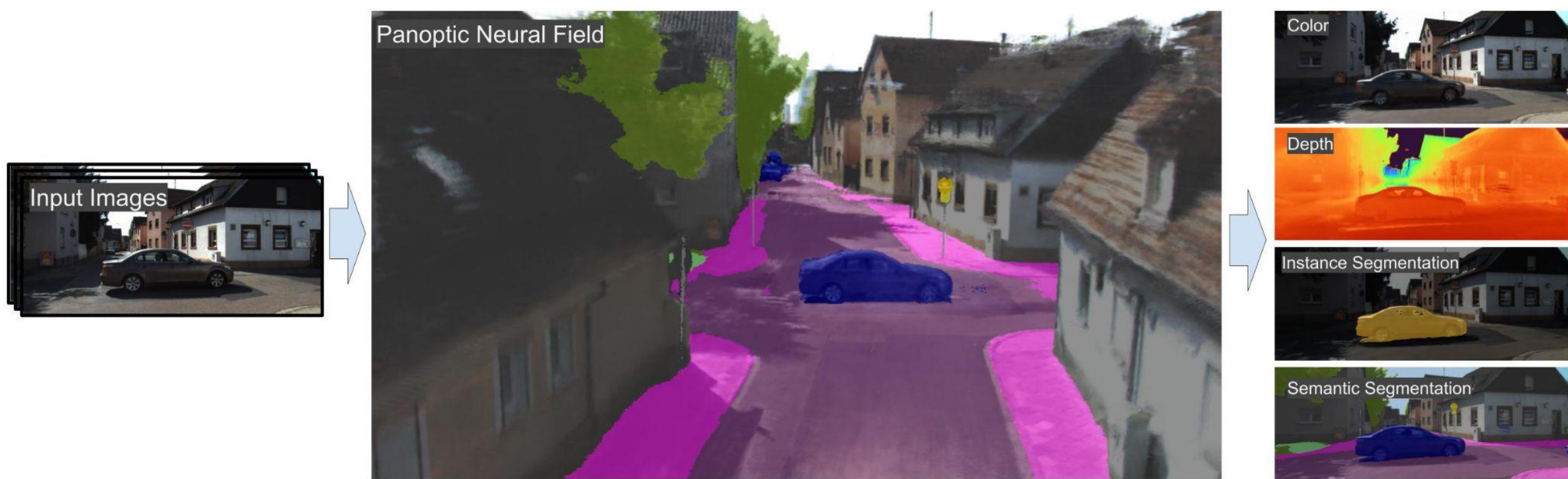
Google Research
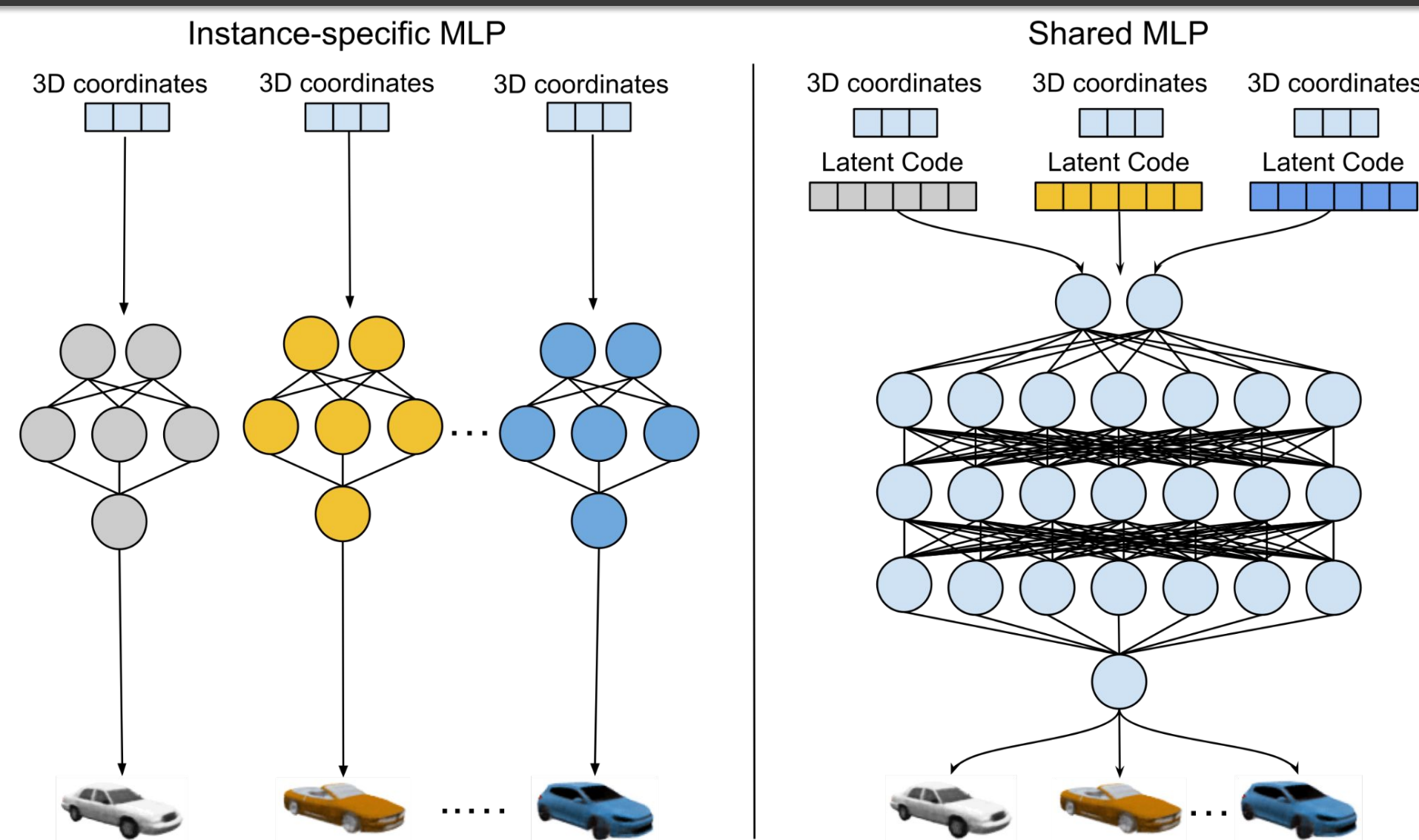
Introduction

Problem Statement: Given an image sequence, create a semantic object-aware neural 3D scene representation.
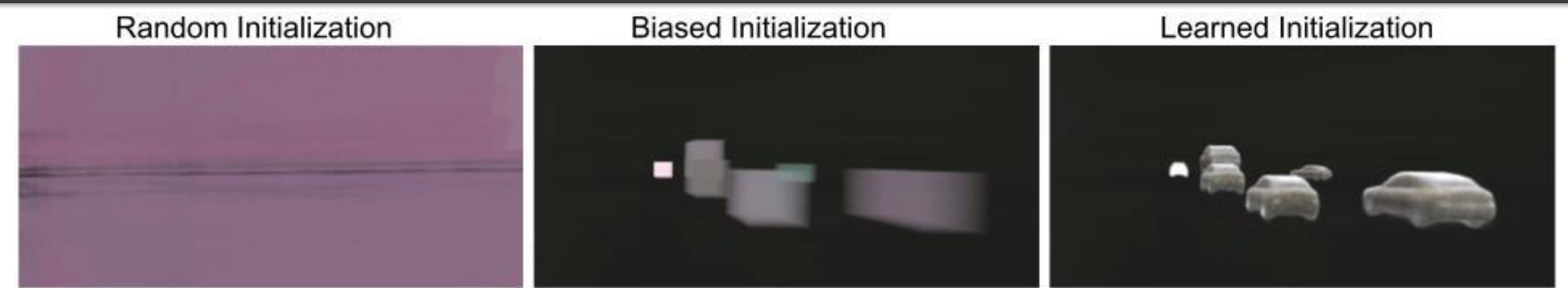
Key Points:

- Separate small instance-specific MLPs for each object instance (**things**) in the scene and a large MLP for static **stuff** background.
- Creates a panoptic-radiance field that can be queried at any 3D point over time for the semantic label, instance label, color, and density.
- Single unified model for multiple tasks like semantic segmentation, panoptic segmentations, view synthesis, scene editing.
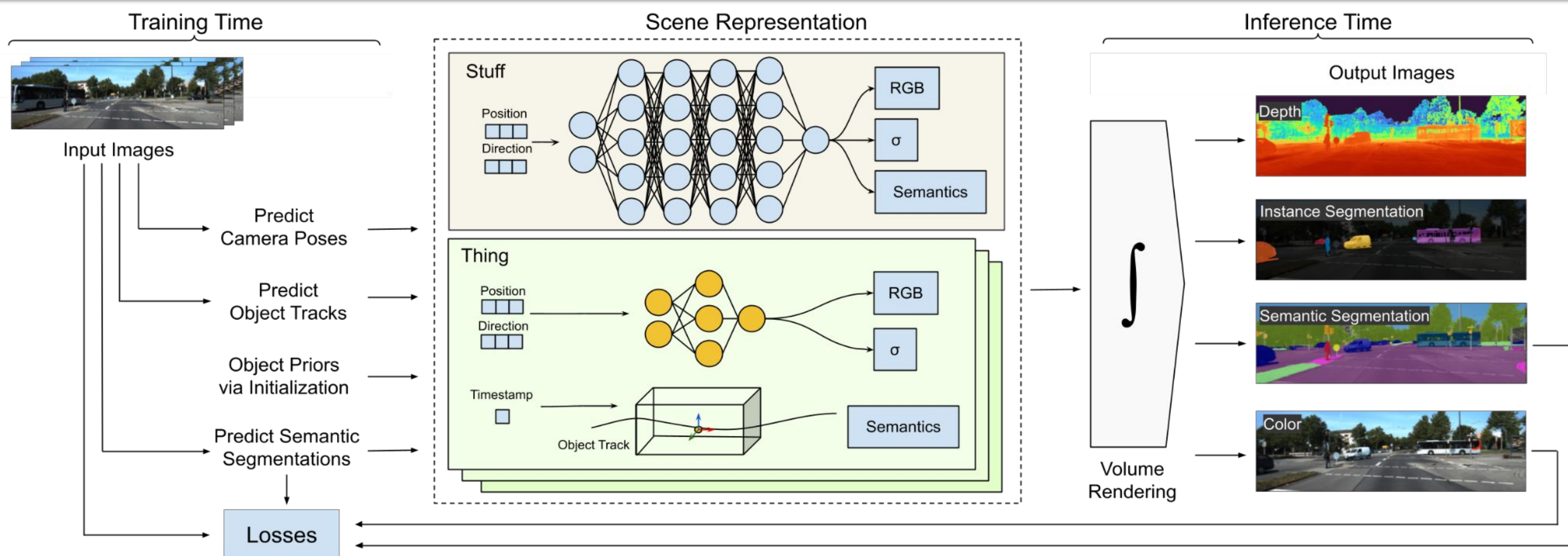- Handles dynamic 3D scenes with multiple moving (rigid) objects.

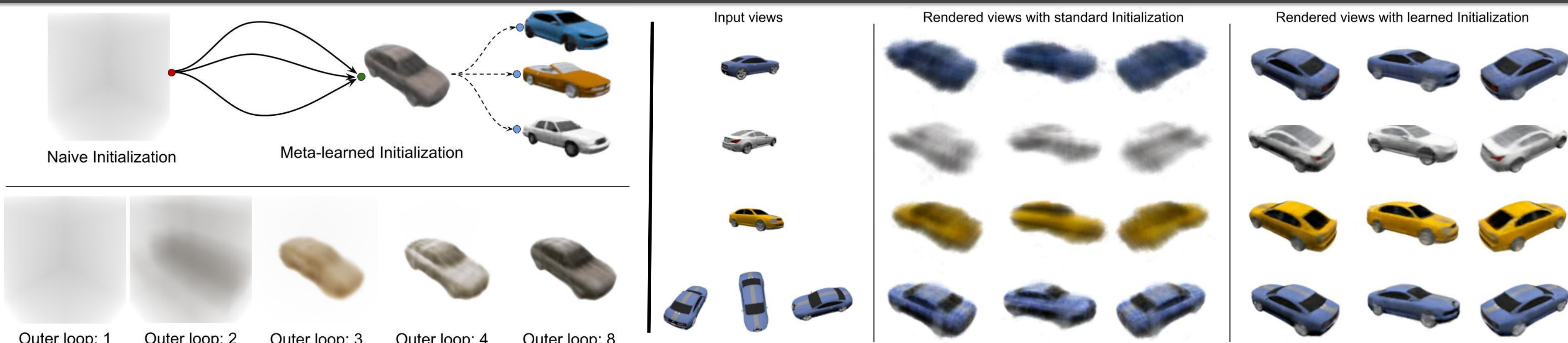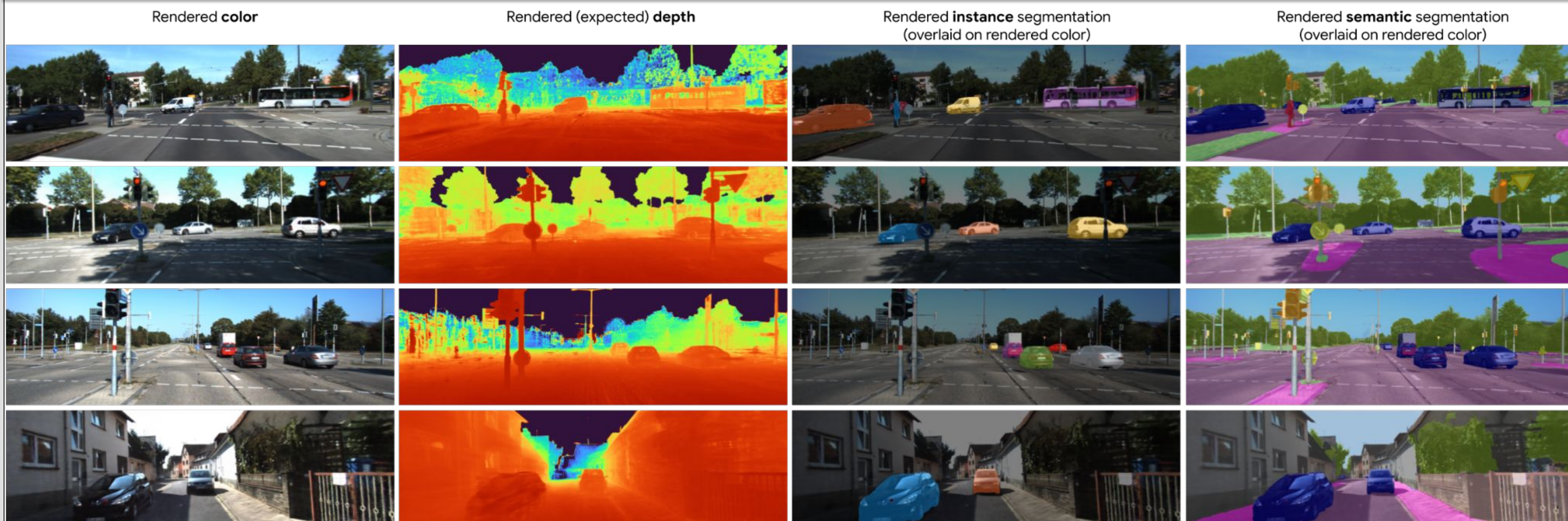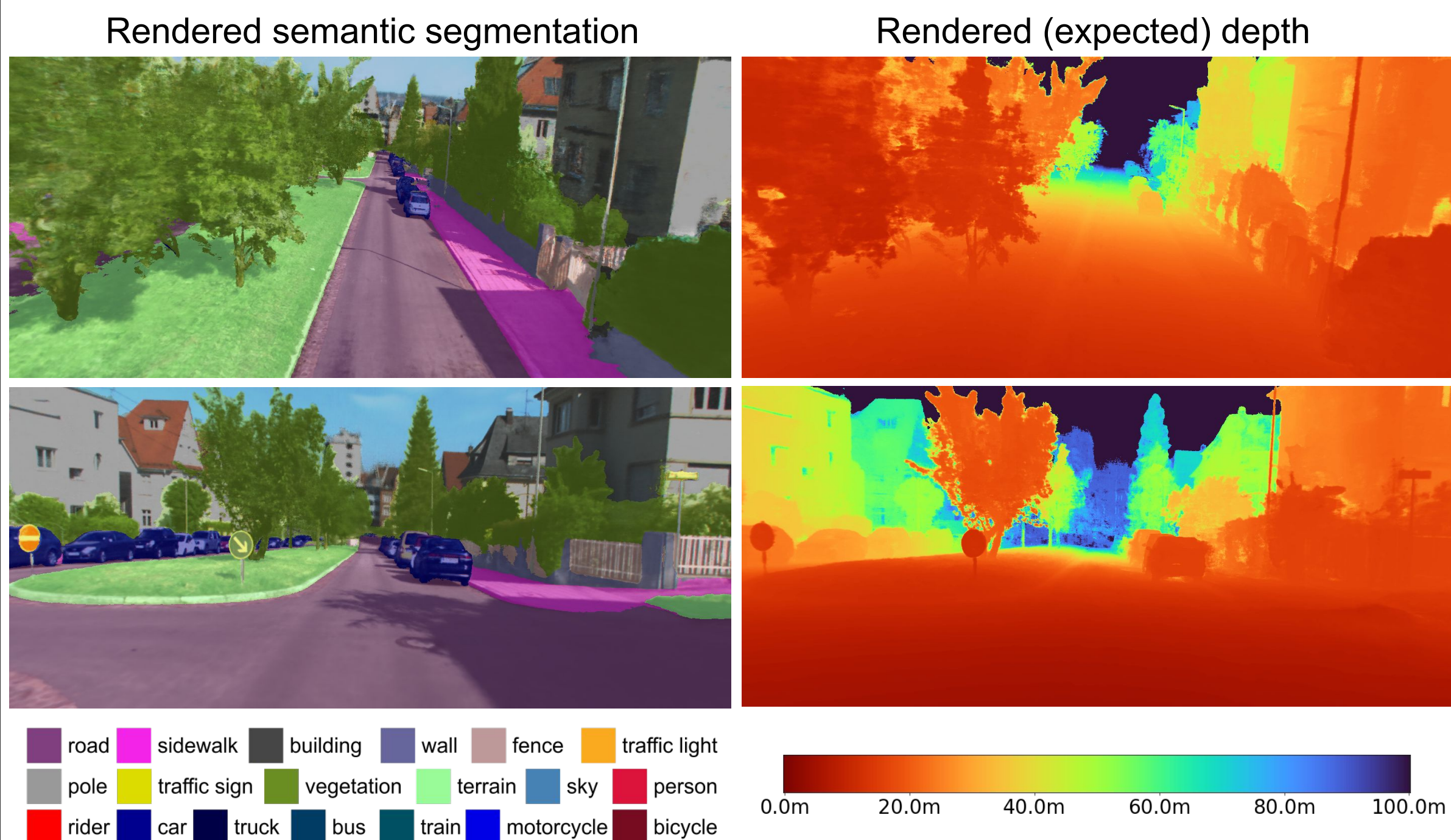## Instance-specific MLP vs shared MLP

Initialization schemes

Random Initialization · Biased Initialization · Learned Initialization

System Overview

Meta Learned Initialization with Federated Averaging

Naive Initialization · Meta-learned Initialization

Input views · Rendered views with standard Initialization · Rendered views with learned Initialization

Outer loop: 1 · Outer loop: 2 · Outer loop: 3 · Outer loop: 4 · Outer loop: 8

Experiments on KITTI

Rendered **color** · Rendered (expected) **depth** · Rendered **instance** segmentation (overlaid on rendered color) · Rendered **semantic** segmentation (overlaid on rendered color)

Experiments on KITTI-360

Rendered semantic segmentation · Rendered (expected) depth

road · sidewalk · building · wall · fence · traffic light · pole · traffic sign · vegetation · terrain · sky · person · rider · car · truck · bus · train · motorcycle · bicycle

0.0m · 20.0m · 40.0m · 60.0m · 80.0m · 100.0m

Method	Appearance PSNR
NeRF	21.18
FVS	20.00
PBNR	19.91
Mip-NeRF	21.54
Ours	**21.91**

Method	Semantic IoU
NeRF + PSPNet	53.01
FVS + PSPNet	67.08
PBNR + PSPNet	65.07
Mip-NeRF + PSPNet	51.15
Ours	**74.28**

Scene Editing

Rendered views **without** scene editing · Rendered views **with** scene editing