NEURAL INFORMATION PROCESSING SYSTEMS

Neural Lighting Simulation for Urban Scenes

Ava Pun*[1,3], Gary Sun*[1,3], Jingkang Wang*[1,2], Yun Chen[1,2], Ze Yang[1,2]
Siva Manivasagam[1,2], Wei-Chiu Ma[1,4], Raquel Urtasun[1,2]

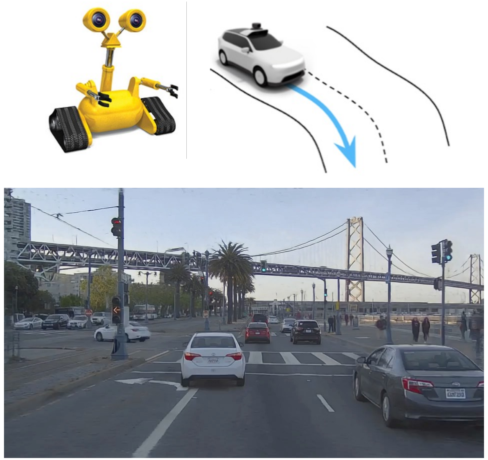
https://waabi.ai/lightsim

waabi UNIVERSITY OF TORONTO UNIVERSITY OF WATERLOO MIT

37th Conference on Neural Information Processing Systems (2023), New Orleans, USA

Simulation for Robust Image Perception in Robots

- Modern camera-based perception systems are not robust under different lighting
- Collecting data under various lighting are expensive and time-consuming
- We need scalable and affordable way to generate experiences - Simulation!



Real Data Collection

Simulation

Simulation Variations (Actor cut in + lighting changes)

Existing Simulators Lack Scale and Diversity

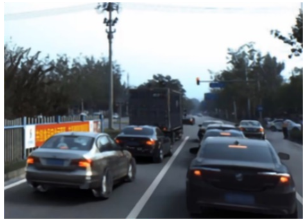
- Standard game engines for simulation such as CARLA [1]:
  - not scalable, lacking diversity, unrealistic
- Limited number of manually designed assets and lighting conditions
- Trained perception system generalizes poorly to the real world [2]

[1] CARLA: An open urban driving simulator. [Dosovitskiy, et al., CoRL 2017]
[2] Enhancing photorealism enhancement. [Richter et al., PAMI 2021]

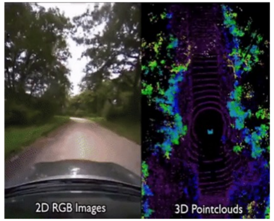
Existing Simulators Bake the Lighting

- Data-driven simulators build digital twins with baked lighting
  - Simulation limited to one single scene and cannot generalize
  - No lighting simulation (shadows, inter-object lighting effects)

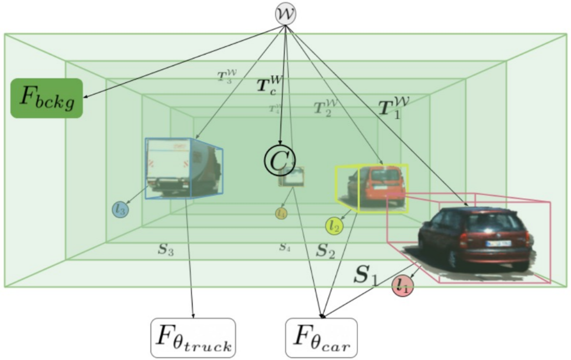
GeoSim [1]



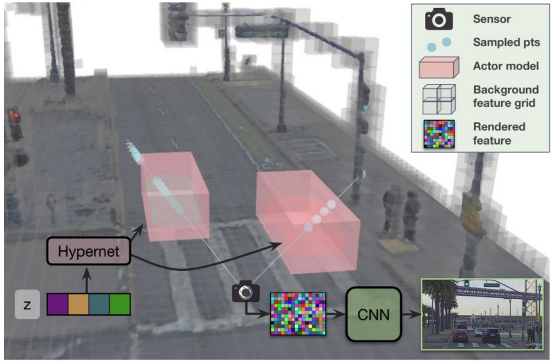
AADS [2]



VISTA 2.0 [3]



Neural Scene Graph [4]

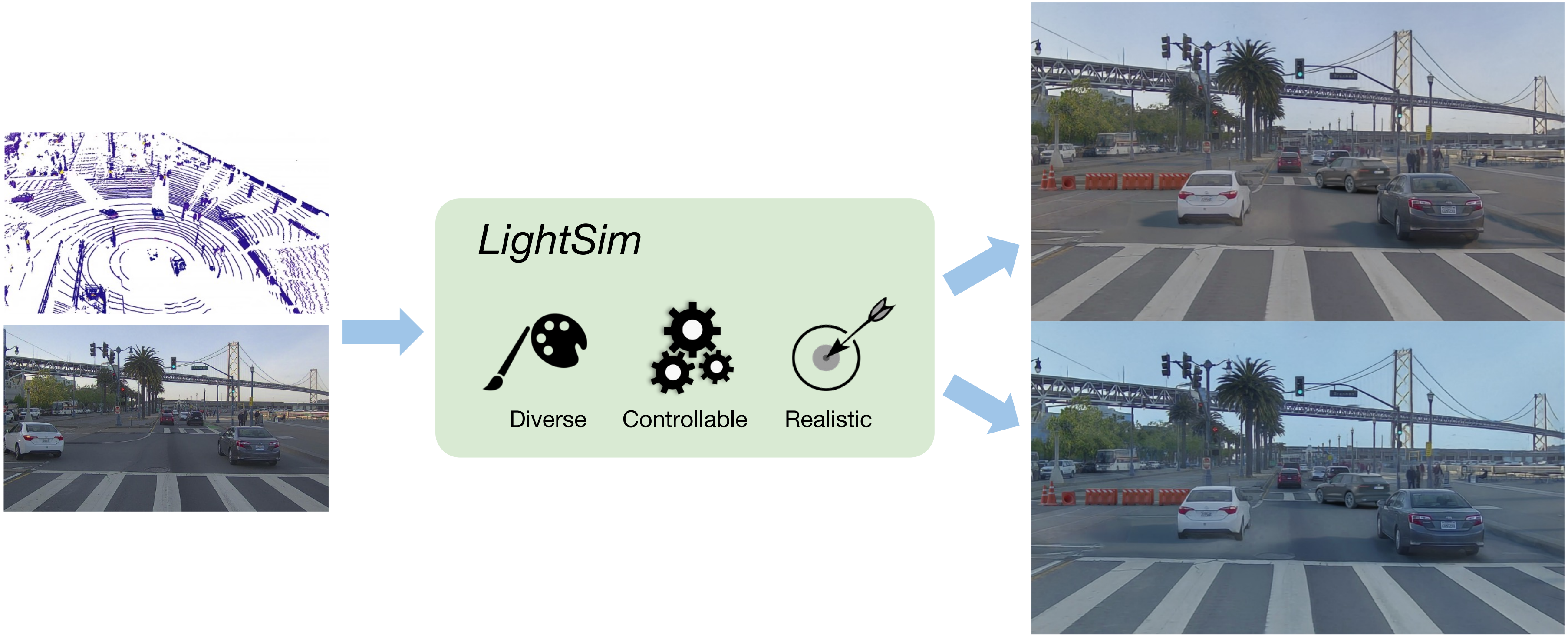


UniSim [5]

[1] GeoSim: Realistic Video Simulation via Geometry-Aware Composition for Self-Driving. [Chen at al., CVPR 2021]
[2] AADS: Augmented autonomous driving simulation using data-driven algorithms. [Li et al., Sci. Robotics. 2021]
[3] VISTA 2.0: An Open, Data-driven Simulator for Multimodal Sensing and Policy Learning for Autonomous Vehicles [Amini et al., ICRA 2022]
[4] Neural Scene Graphs for Dynamic Scenes. [Ost et al., CVPR 2021]
[5] UniSim: A Neural Closed-Loop Sensor Simulator. [Yang, et al. CVPR 2023]

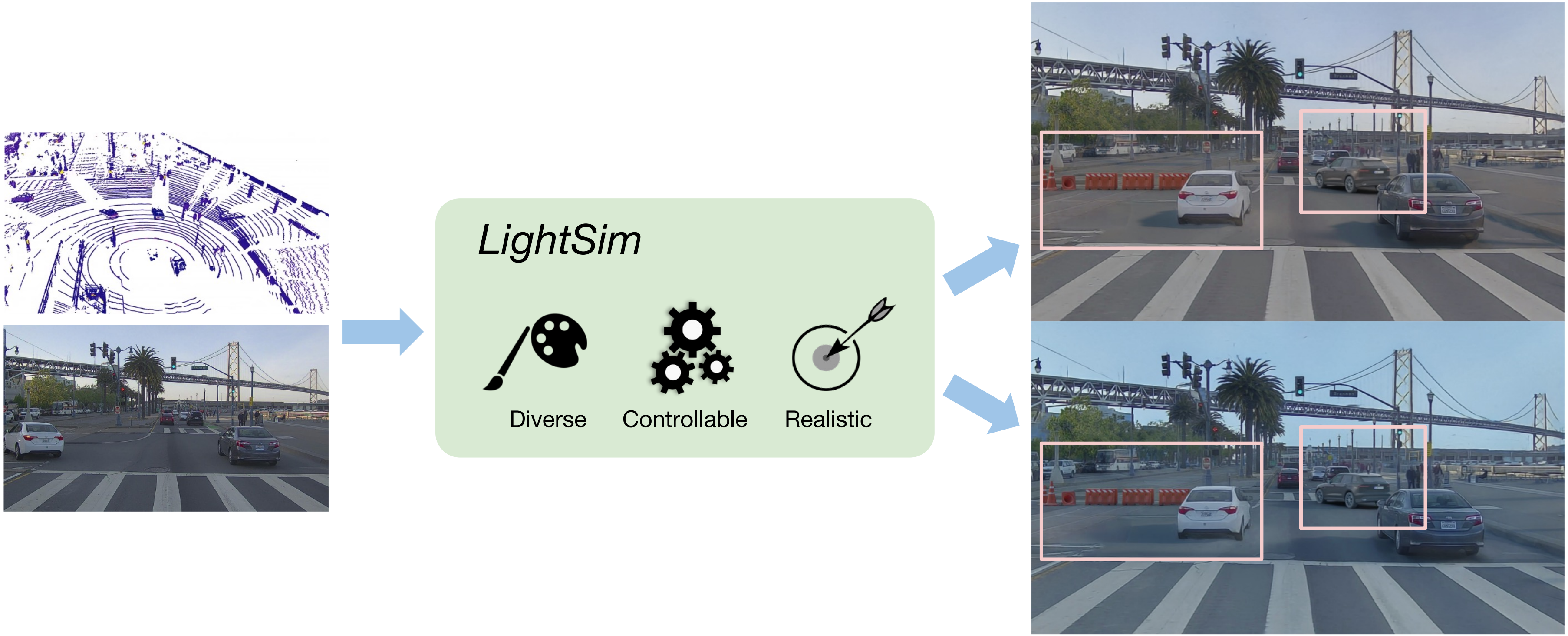
Our Goal

- Create a diverse, controllable, and realistic simulator that can generate camera data of scenes at scale under diverse lighting conditions



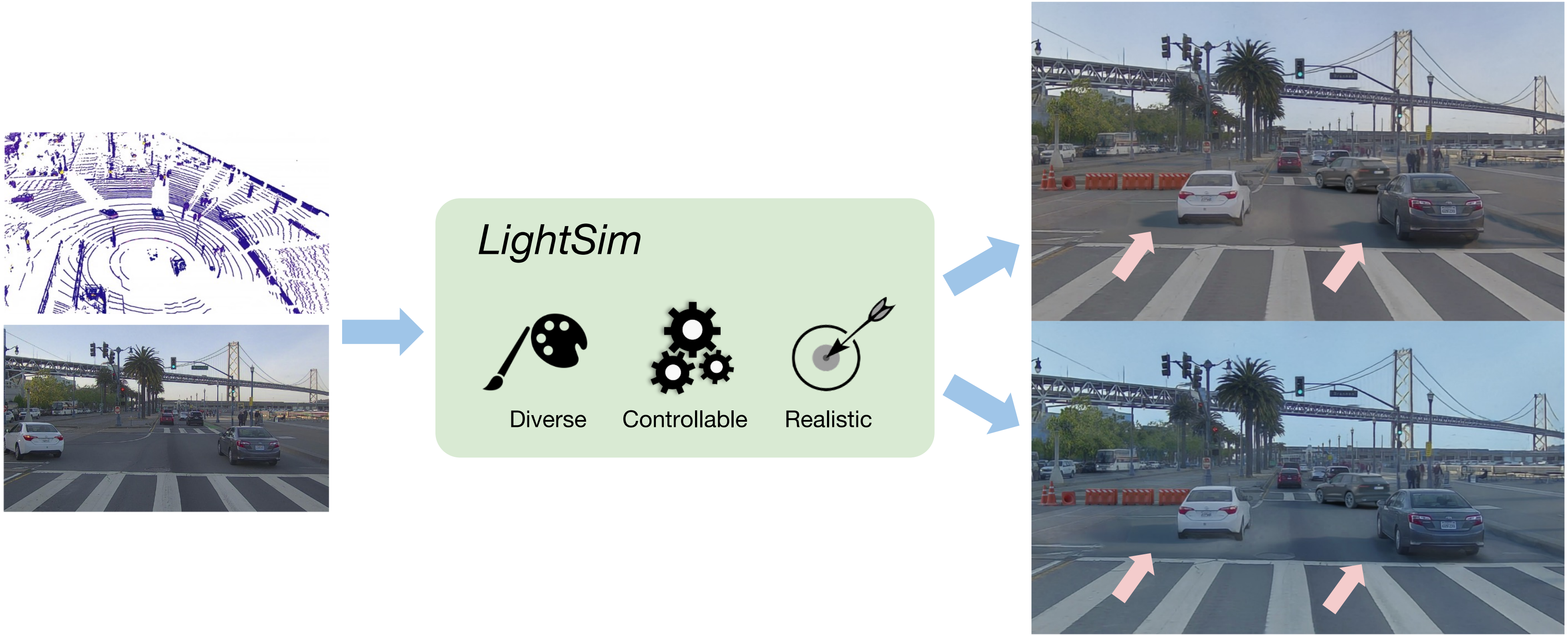
Our Goal

- Create a diverse, controllable, and realistic simulator that can generate camera data of scenes at scale under diverse lighting conditions



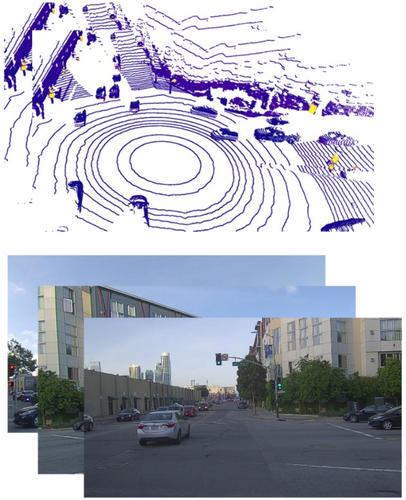
Our Goal

- Create a diverse, controllable, and realistic simulator that can generate camera data of scenes at scale under diverse lighting conditions



Step 1: Building Relightable Digital Twins of the Real World

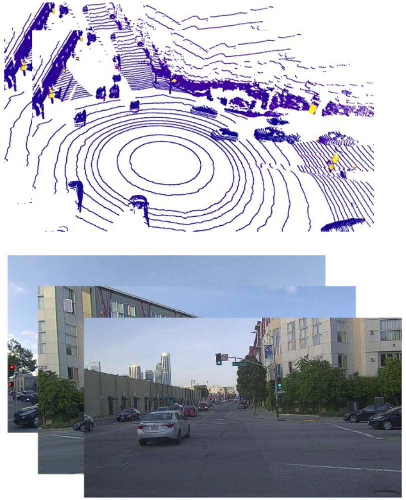
- Neural scene reconstruction to recover scene geometry and texture



Sensor data

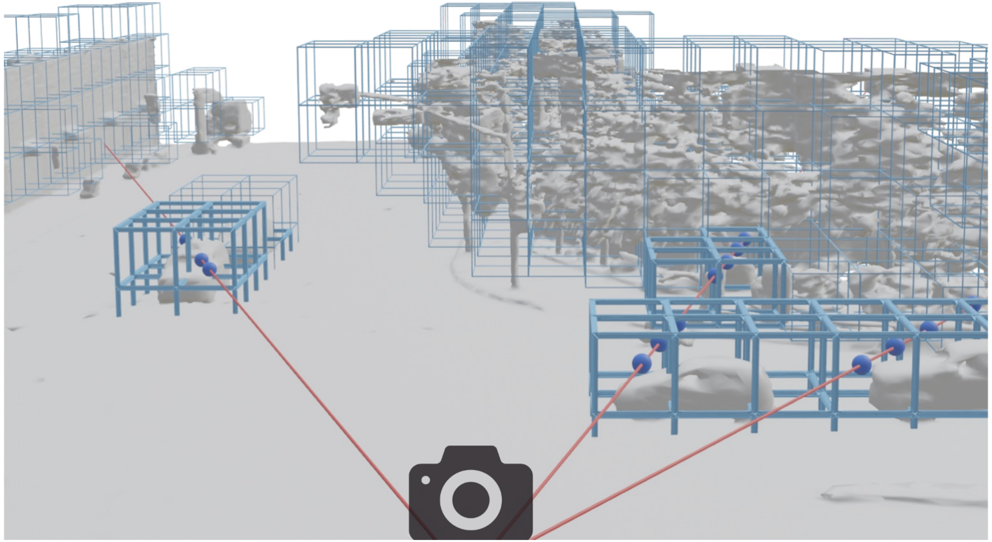
Step 1: Building Relightable Digital Twins of the Real World

- Neural scene reconstruction to recover scene geometry and texture



Sensor data

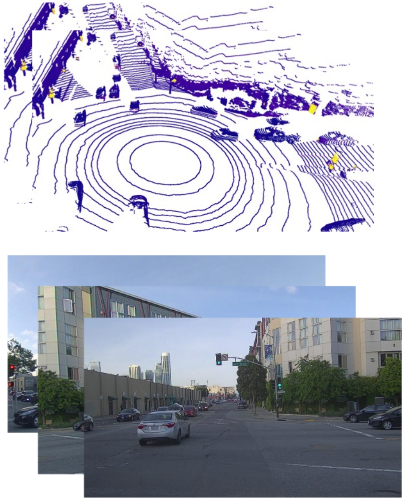
Neural Scene Reconstruction



Compositional neural radiance field (background + actors)

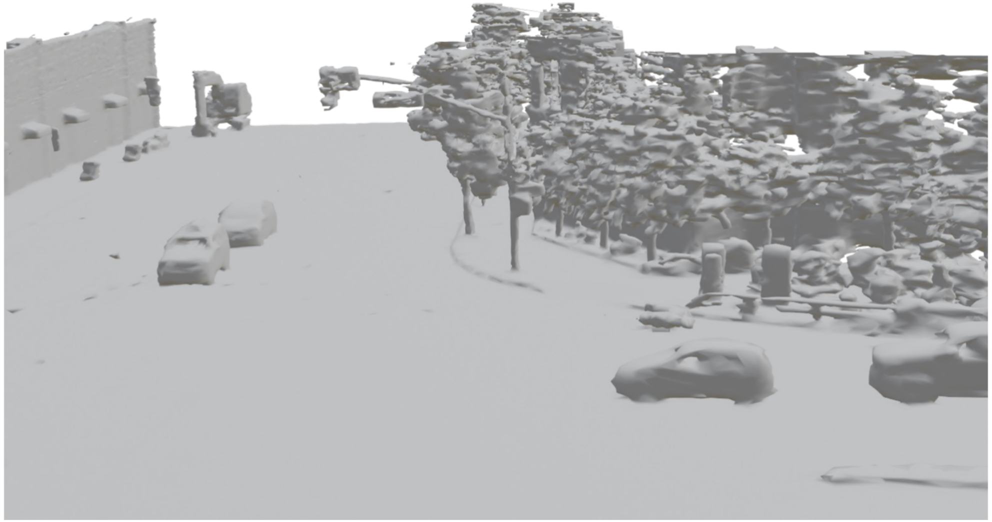
Step 1: Building Relightable Digital Twins of the Real World

- Neural scene reconstruction to recover scene geometry and texture



Sensor data

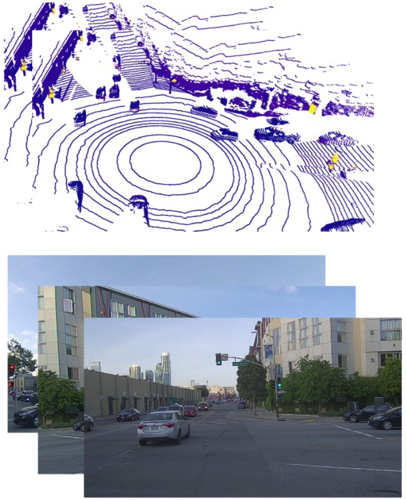
Neural Scene Reconstruction



Explicit digital twins (geometry, texture)

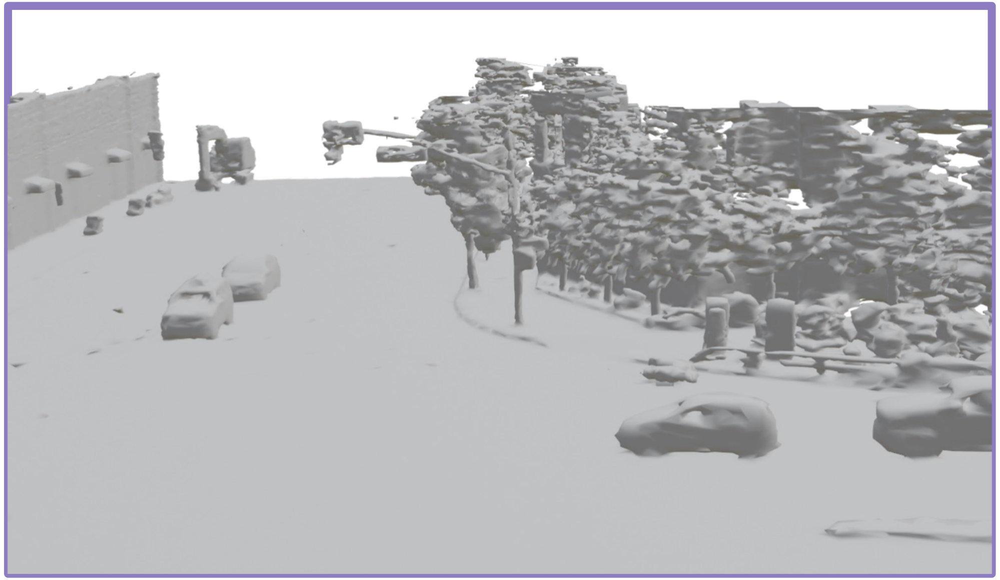
Step 1: Building Relightable Digital Twins of the Real World

- Neural scene reconstruction to recover scene geometry and texture



Sensor data

Neural Scene Reconstruction



Explicit digital twins (geometry, texture)

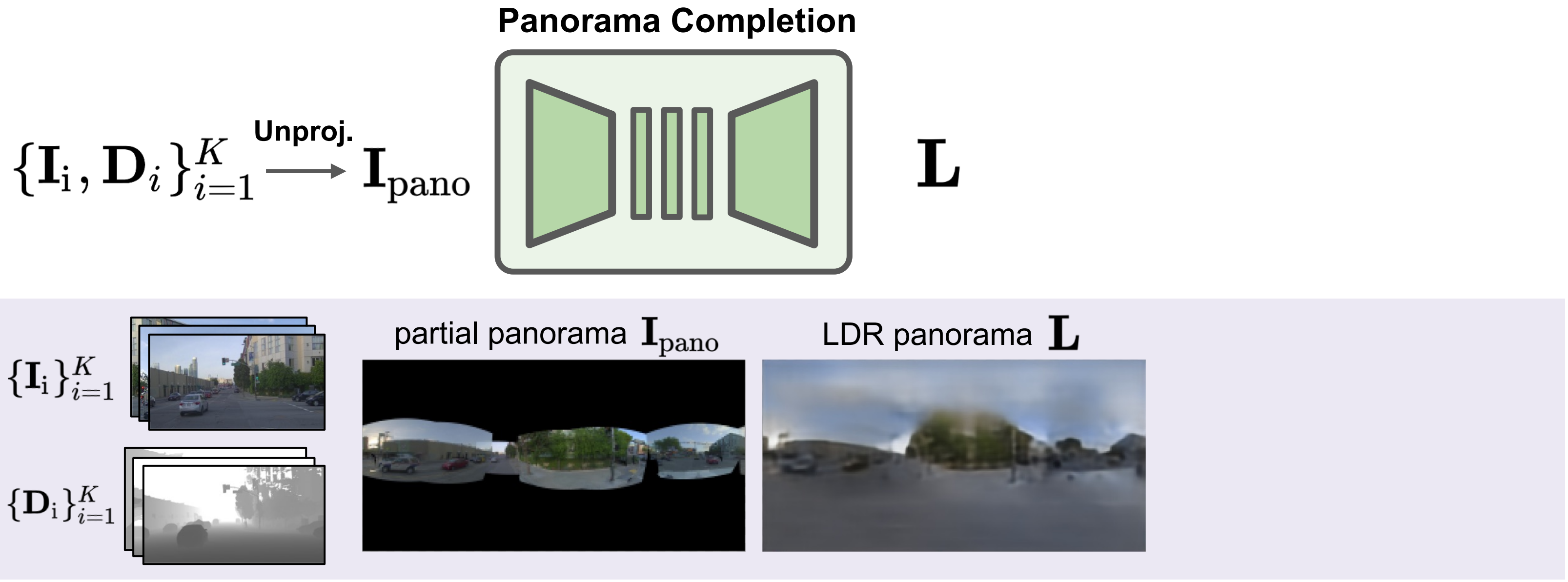
Step 1: Building Relightable Digital Twins of the Real World

- Neural lighting simulation to recover the HDR sky dome

$$\{\mathbf{I}_{\mathrm{i}}, \mathbf{D}_i\}_{i=1}^{K} \xrightarrow{\textbf{Unproj.}} \mathbf{I}_{\mathrm{pano}}$$

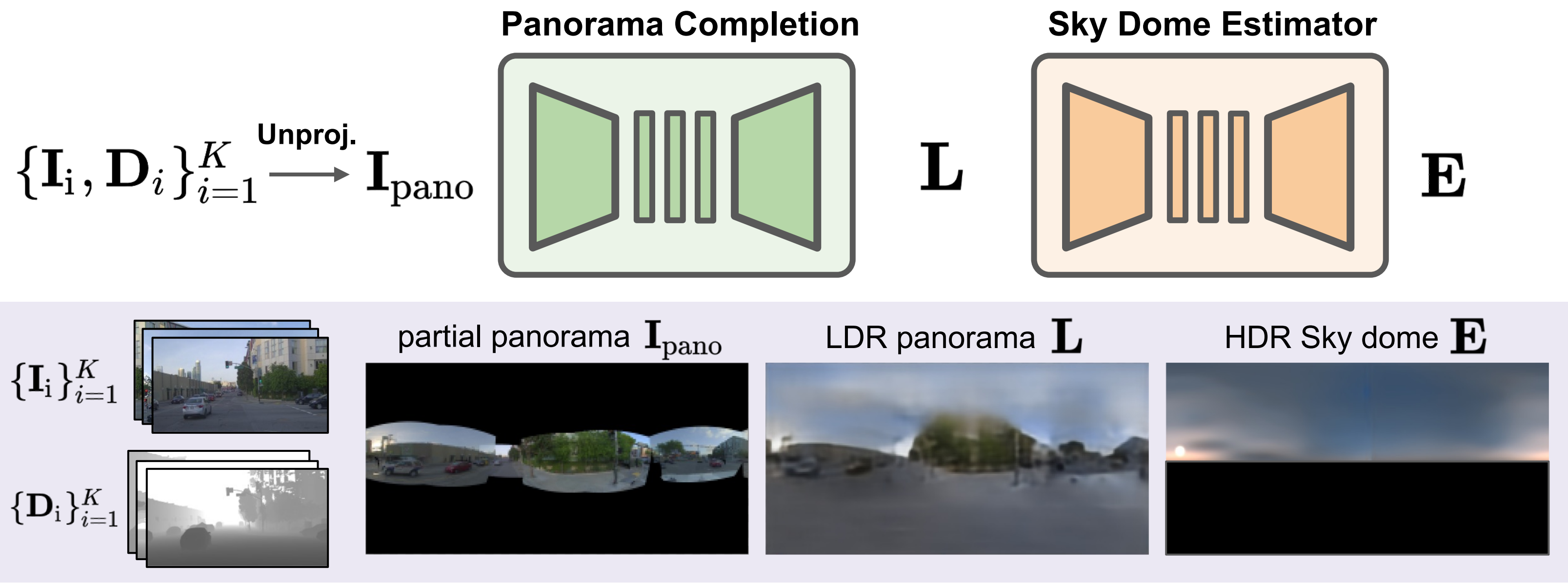
Step 1: Building Relightable Digital Twins of the Real World

- Neural lighting simulation to recover the HDR sky dome



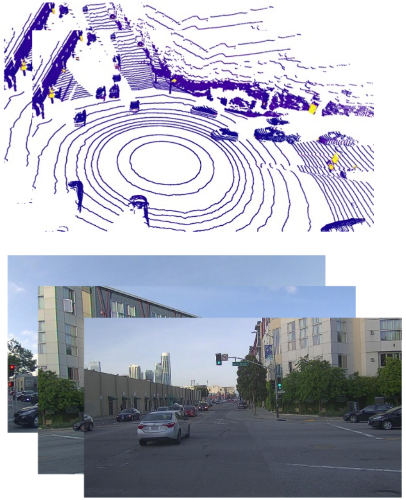
Step 1: Building Relightable Digital Twins of the Real World

- Neural lighting simulation to recover the HDR sky dome



Step 1: Building Relightable Digital Twins of the Real World

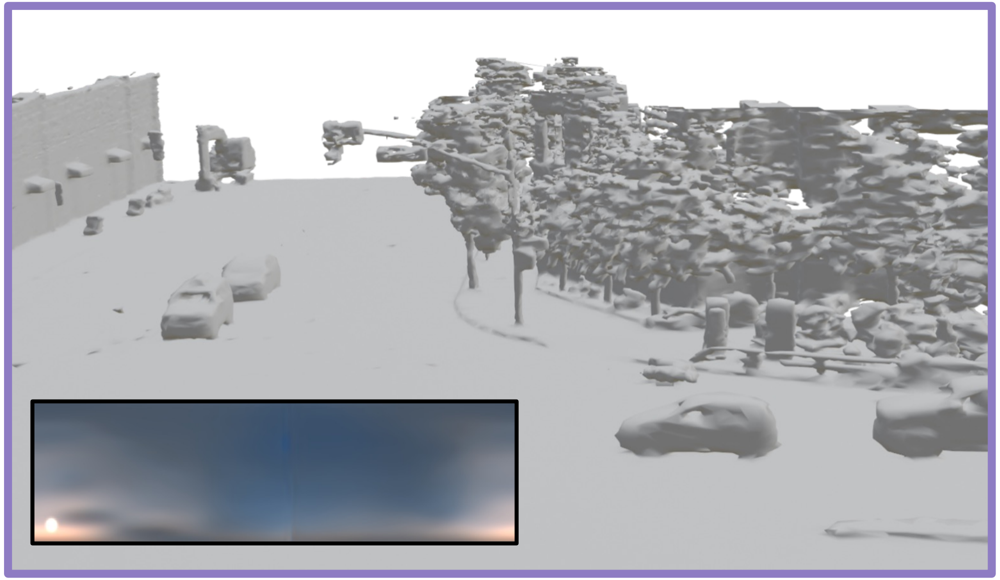
- Neural scene reconstruction to recover scene geometry and texture
- Neural lighting simulation to recover the HDR sky dome



Sensor data

Neural Scene Reconstruction

Neural Lighting Estimation



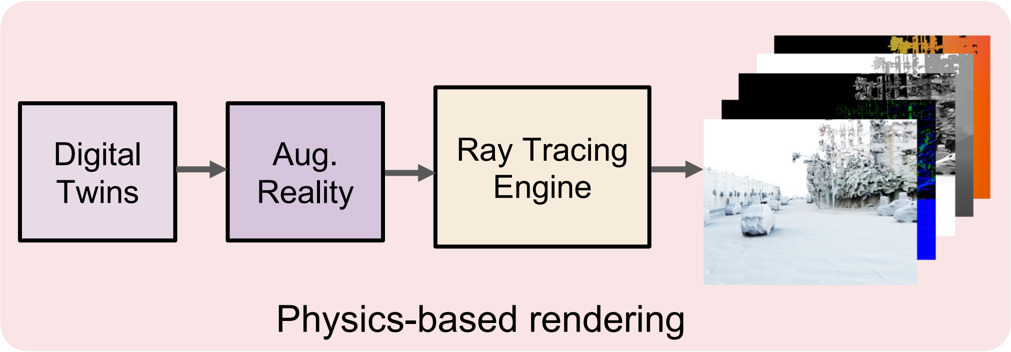
Relightable digital twins (geometry, texture, lighting)

Step 2: Neural Lighting Simulation of Dynamic Urban Scenes

- Derive augmented reality representation from digital twins
- Generate lighting-relevant data with physically-based rendering
- Neural deferred rendering for lighting simulation

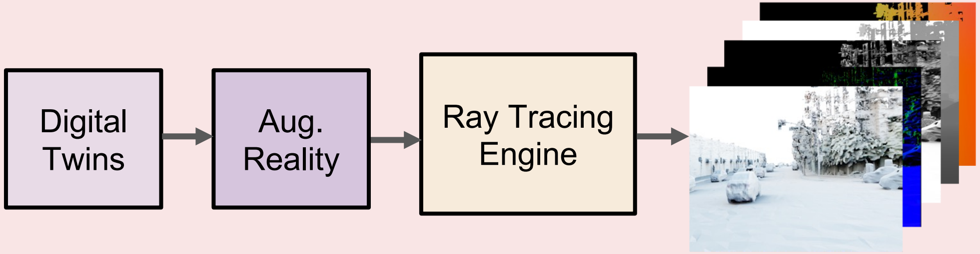
Step 2: Neural Lighting Simulation of Dynamic Urban Scenes

- Derive augmented reality representation from digital twins
- Generate lighting-relevant data with physically-based rendering
- Neural deferred rendering for lighting simulation

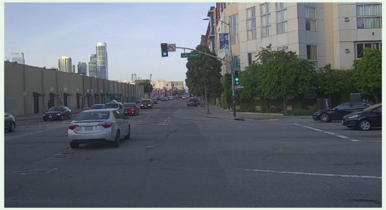


Step 2: Neural Lighting Simulation of Dynamic Urban Scenes

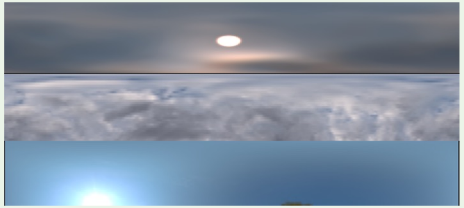
- Derive augmented reality representation from digital twins
- Generate lighting-relevant data with physically-based rendering
- Neural deferred rendering for lighting simulation



Physics-based rendering



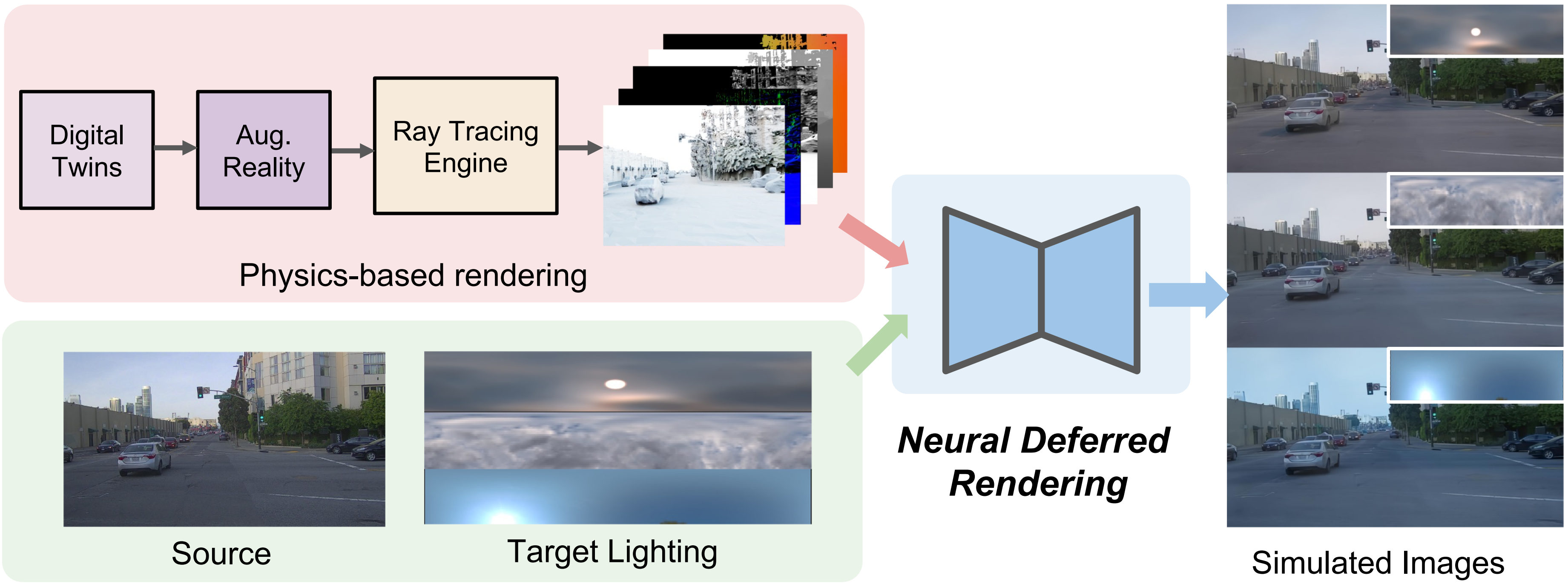
Source



Target Lighting

Step 2: Neural Lighting Simulation of Dynamic Urban Scenes

- Derive augmented reality representation from digital twins
- Generate lighting-relevant data with physically-based rendering
- Neural deferred rendering for lighting simulation

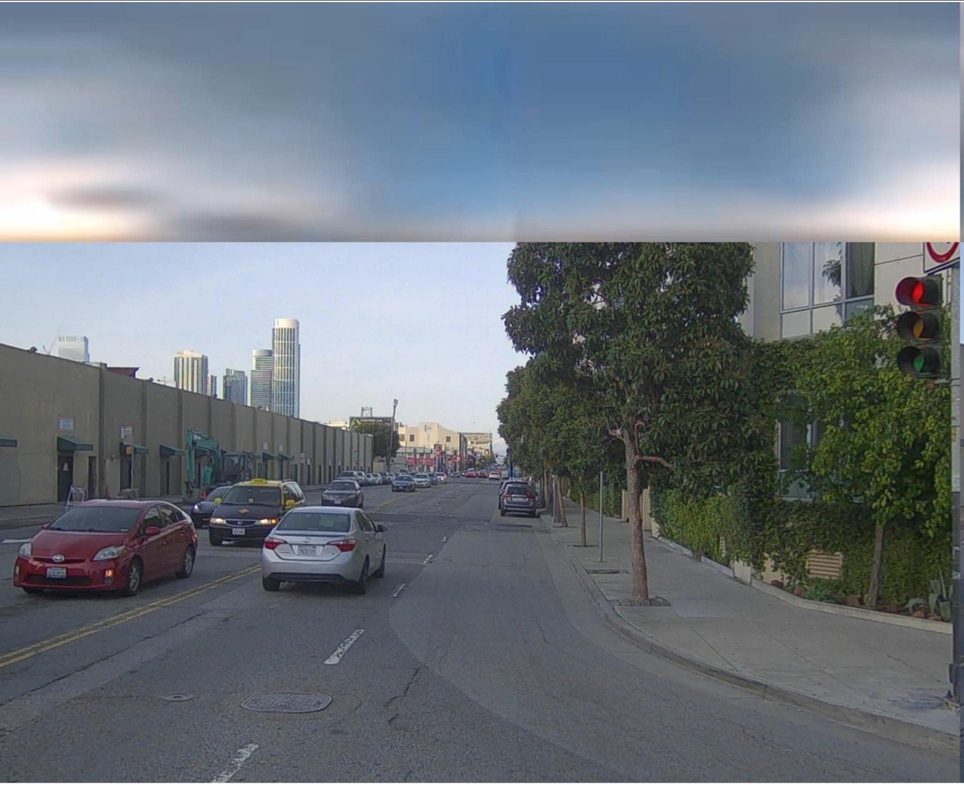


Scene Relighting

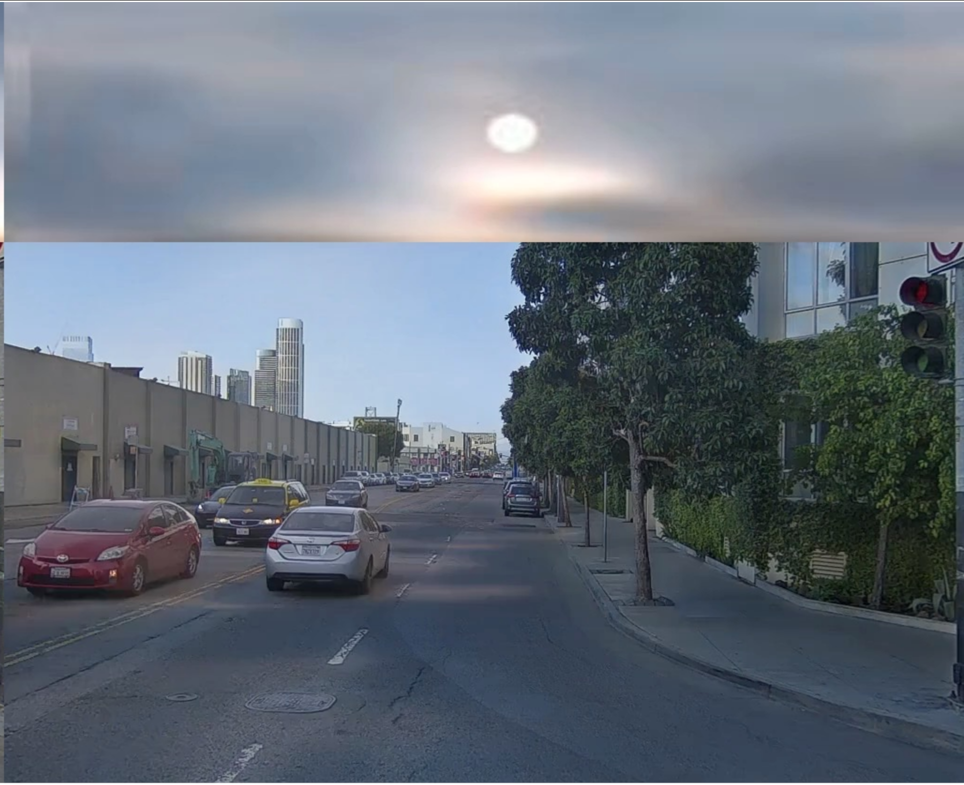
Real Video and Estimated Source Lighting

Simulated Video with Target Lighting

Shadow Editing



Real Image and Estimated Source Lighting



Simulated Video with Rotated Lighting

Lighting-aware Actor Insertion

Controllable Camera Simulation

Controllable Camera Simulation – variation 1

Original

Simulated Variations

Controllable Camera Simulation – variation 2

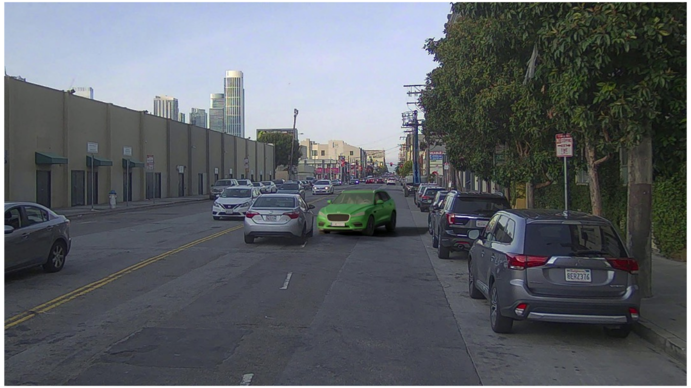
Original

Simulated Variations

Lighting Estimation Evaluation via Actor Insertion

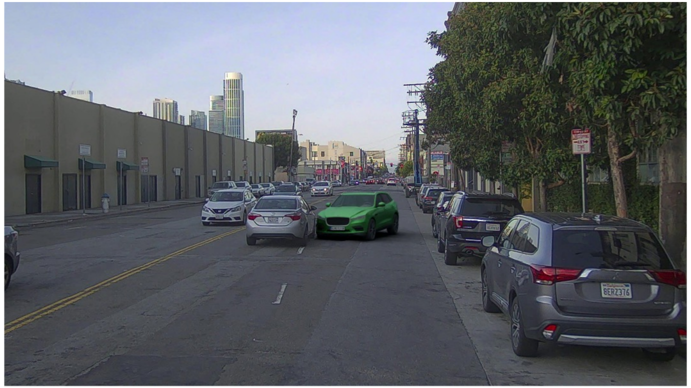
Original

SOLD-Net



NLFE* (Panorama)

Ours

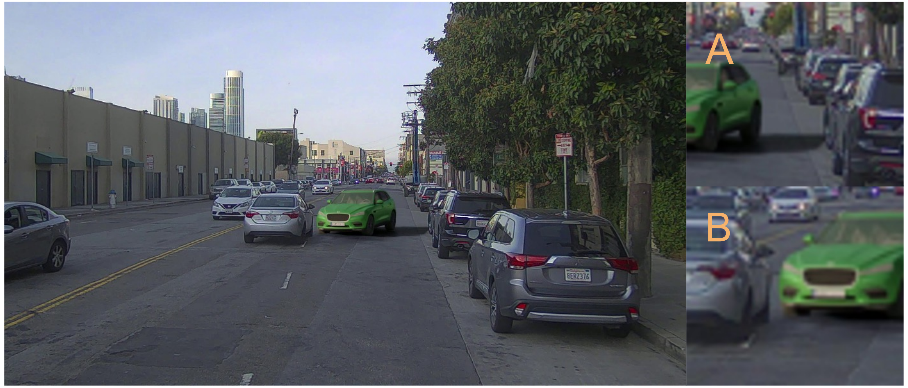


*HDR skydome only

Lighting Estimation Evaluation via Actor Insertion

Original

SOLD-Net



NLFE* (Panorama)

Ours

*HDR skydome only

Generalization on nuScenes

Real

LightSim Simulated

Downstream Perception Training

- Realistic lighting simulation can help improve the performance of downstream object detection task under unseen lighting conditions

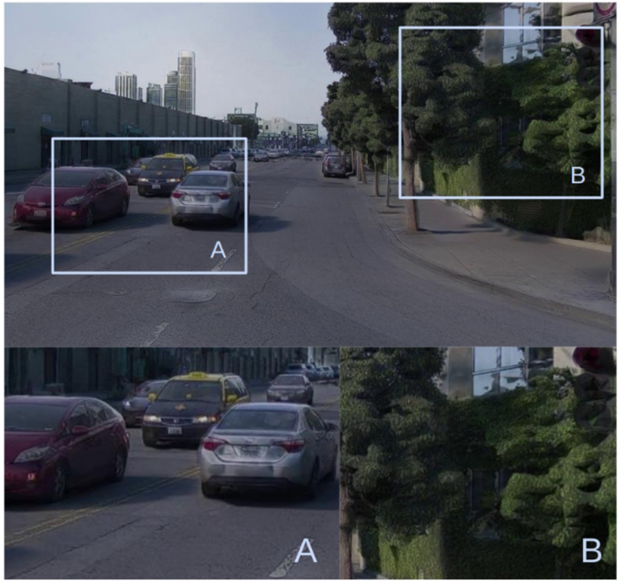
Model	mAP (%)
Real	32.1
Real + Color aug. [41]	33.8 (+1.7)
Real + Sim (Self-OSR)	30.3 (−1.8)
Real + Sim (EPE)	32.5 (+0.4)
Real + Sim (Color Transfer)	35.1 (+3.0)
Real + Sim (Ours)	**36.6** (**+4.5**)

Comparison in Scene Relighting

Ablation Study

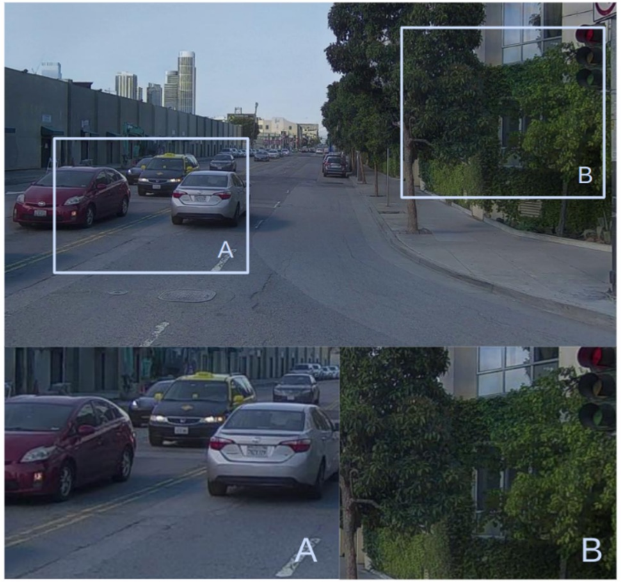
- Content-preserving loss

$\lambda_{\text{edge}} = 0$



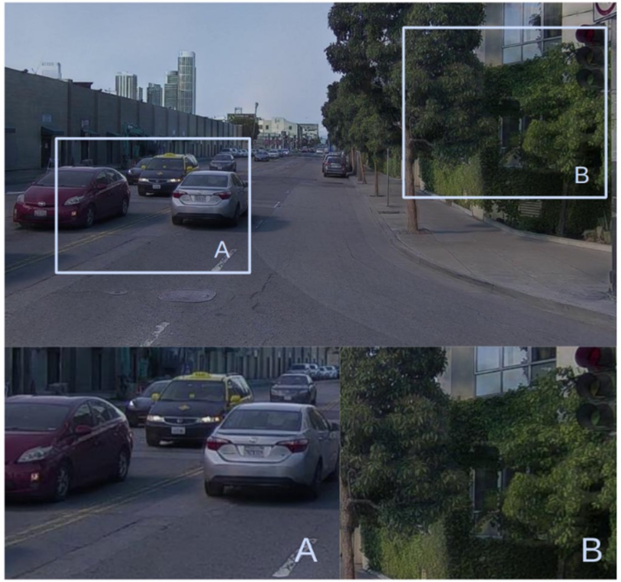
FID = 109.8

$\lambda_{\text{edge}} = 800$



FID = 57.3

Ours ($\lambda_{\text{edge}} = 400$ **)**



FID = 55.4

Ablation Study

- *sim-to-real* and *identity* pairs

w/o sim-real pairs

FID = 60.9

w/o identity pairs

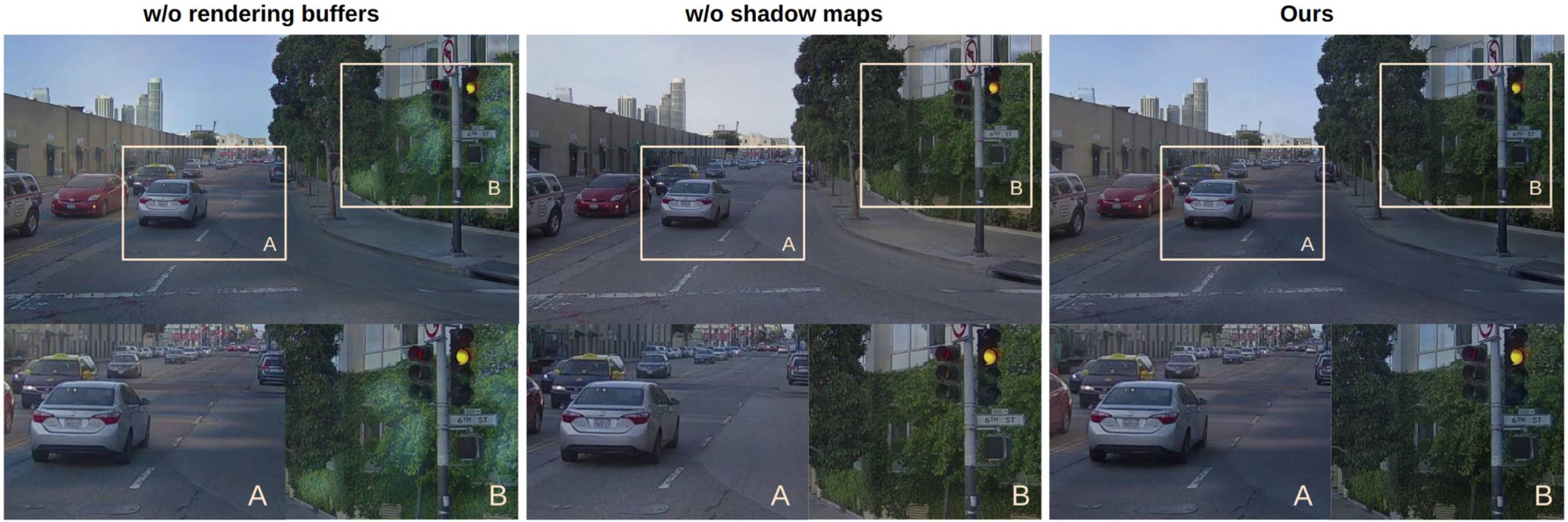
FID = 62.5

Ours

FID = 55.4

Ablation Study

- Rendering buffers and shadow maps



NEURAL INFORMATION PROCESSING SYSTEMS

Thank you!

Neural Lighting Simulation for Urban Scenes

Ava Pun*[1,3], Gary Sun*[1,3], Jingkang Wang*[1,2], Yun Chen[1,2], Ze Yang[1,2]

Siva Manivasagam[1,2], Wei-Chiu Ma[1,4], Raquel Urtasun[1,2]

https://waabi.ai/lightsim

waabi UNIVERSITY OF TORONTO UNIVERSITY OF WATERLOO MIT