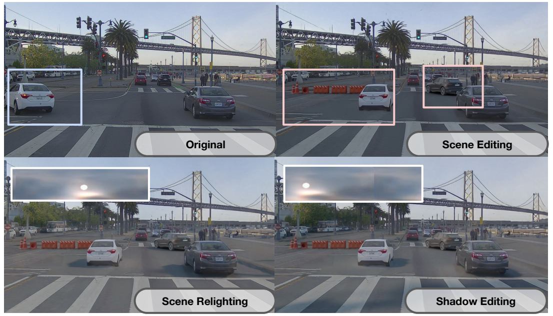
LightSim: Neural Lighting Simulation for Urban Scenes

Ava Pun*, Gary Sun*, Jingkang Wang*, Yun Chen, Ze Yang, Sivabalan Manivasagam, Wei-Chiu Ma, Raquel Urtasun

Waabi, University of Toronto, University of Waterloo, MIT

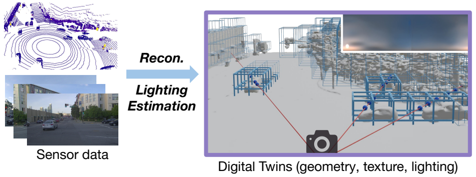
Motivation: Lighting Camera Simulation

- **Varying lighting conditions** may harm robot perception performance
- **Camera simulation** can scalably generate *diverse*, *controllable*, and *realistic* images under different lighting conditions
- **Perception performance is improved** by training on generated data



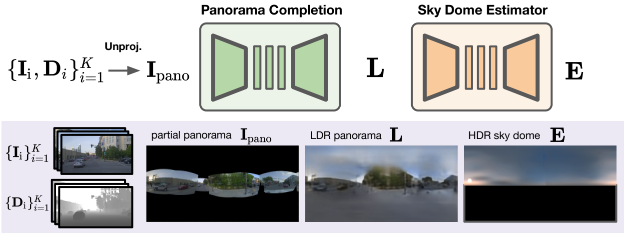
Building Lighting-Aware Digital Twins

- **Building digital twins from the real world:**
  - **diverse:** twins are reconstructed from large real-world dataset
  - **controllable:** PBR allows dynamic actor placement, modification of SDV location, and simulation of novel lighting conditions
  - **realistic:** neural deferred rendering enhances PBR
- **Existing methods:**
  - bake lighting into scene; cannot simulate new lighting conditions
  - have difficulty performing well on large outdoor scenes
  - are restricted to static scenes

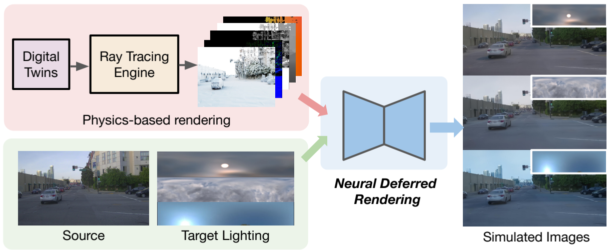


LightSim

- **LightSim:** (a) allows simulation of novel lighting conditions (b) reconstructs large, dynamic outdoor scenes (c) enables lighting-aware scene editing & actor insertion
- **Lighting Estimation**
  - Leverage multi-camera data to build panorama
  - Fill gaps in panorama with image inpainting to create LDR panorama
  - Lift LDR pano to HDR sky dome with estimator network



- **Neural Deferred Rendering**
  - PBR images capture lighting effects well but lack realism
  - Use neural rendering network with lighting-relevant PBR buffers for enhanced realism
  - Train on combo of synthetic & real-world data to ensure controllable lighting & realism



Enhance realism through edge-based content-preserving loss

$$\mathcal{L}_{\text{relight}} = \frac{1}{N}\sum_{i=1}^{N}\Bigg(\underbrace{\left\|\mathbf{I}_i^{\text{tgt}} - \hat{\mathbf{I}}_i^{\text{tgt}}\right\|_2}_{\mathcal{L}_{\text{color}}} + \underbrace{\lambda_{\text{lpips}}\sum_{j=1}^{M}\left\|V^j(\mathbf{I}_i^{\text{tgt}}) - V^j(\hat{\mathbf{I}}_i^{\text{tgt}})\right\|_2}_{\mathcal{L}_{\text{lpips}}} + \lambda_{\text{edge}}\underbrace{\left\|\nabla\mathbf{I}_i^{\text{tgt}} - \nabla\hat{\mathbf{I}}_i^{\text{tgt}}\right\|_2}_{\mathcal{L}_{\text{reg}}}\Bigg)$$

Results

- **Qualitative comparison with SOTA approaches:**

- **Quantitative comparison with SOTA approaches:**

Method	FID ↓	KID ($\times 10^3$) ↓
NeRF-OSR [Rudnev et al., 2022]	143.9	94.0 ± 7.5
EPE [Richter et al., 2021]	93.0	56.0 ± 5.0
LightSim (Ours)	**87.1**	$\mathbf{30.4 \pm 4.0}$

Model	mAP (%)
Real	32.1
Real + Sim (EPE)	32.5 (+0.4)
Real + Sim (Ours)	**36.6** (+4.5)

- **Generalization to nuScenes dataset:**

- **Limitations:** (a) no support for local light sources (b) difficulty removing strong shadows in sunny scenes (c) fixed materials may cause real-sim domain gap