Flux4D: Flow-based Unsupervised 4D Reconstruction

Jingkang Wang*, Henry Che*, Yun Chen*, Ze Yang, Lily Goli, Sivabalan Manivasagam, Raquel Urtasun

Waabi, University of Toronto

https://waabi.ai/flux4d

Motivation: Unsupervised Reconstruction

+ **Why**: Scalable 4D reconstruction is important for simulation!
+ **Existing approaches**:
  + human labels to separate foreground and background
  + costly per-scene optimization and overfits to training views
+ *Flux4D* is an **unsupervised** and **generalizable** reconstruction approach that enables accurate and efficient driving scene reconstruction at scale.
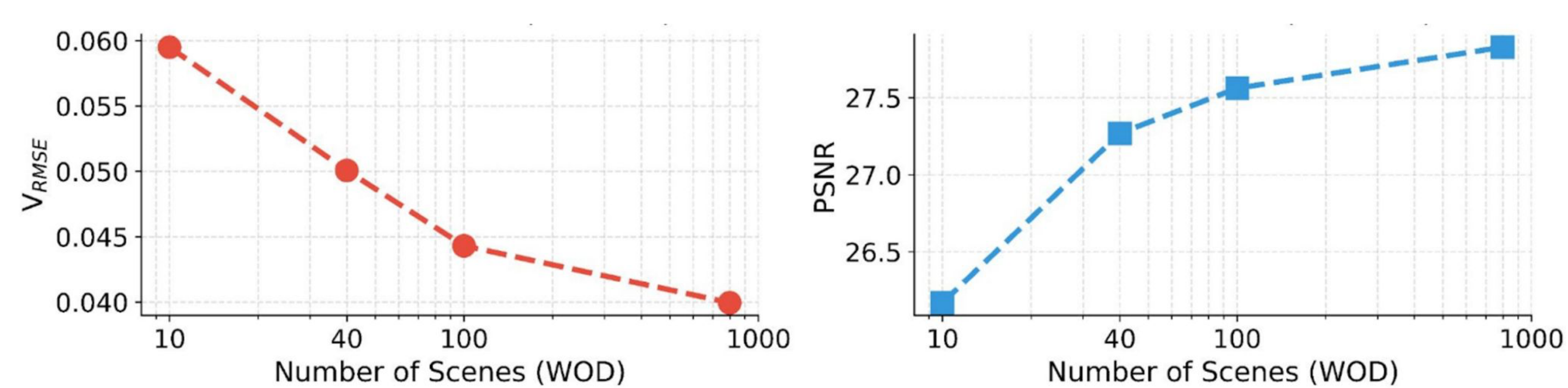
Per-scene 4D scene reconstruction with labels
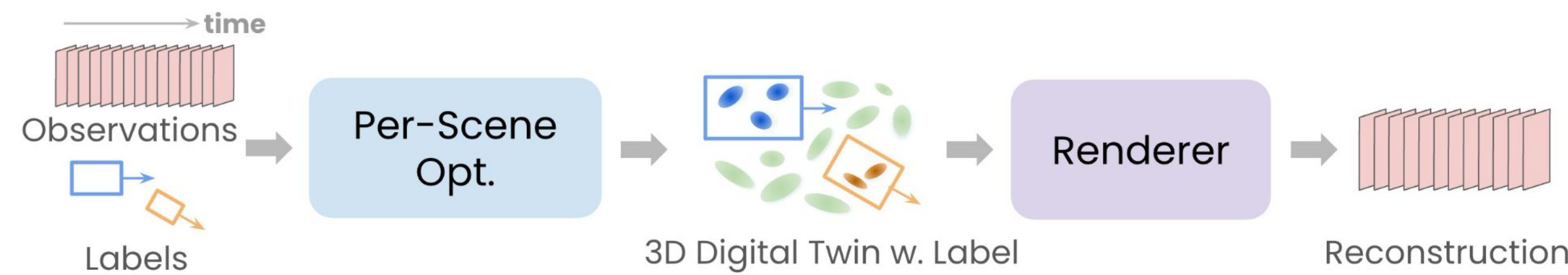
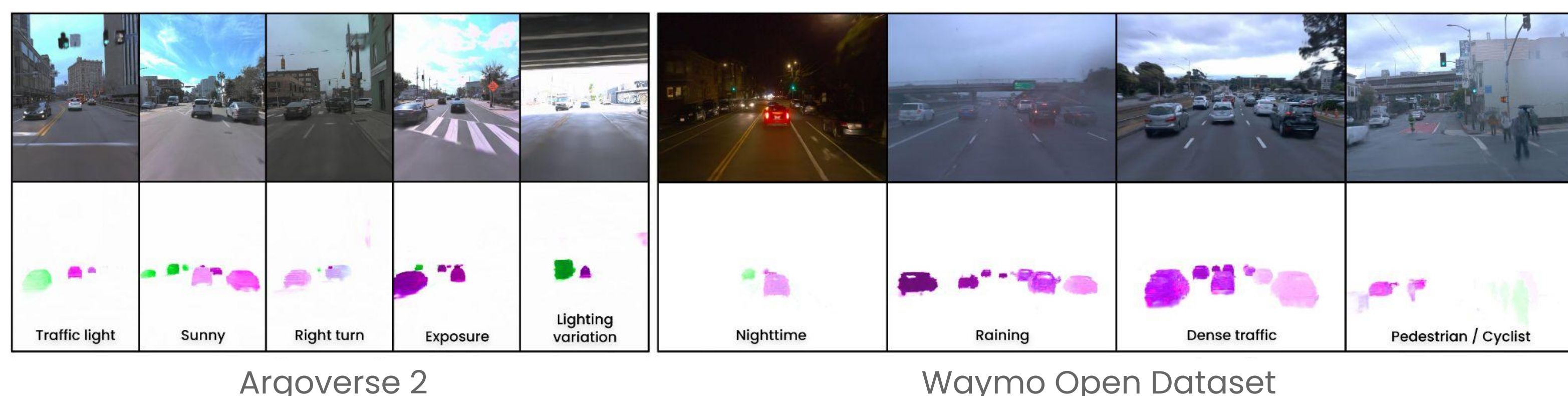
Ours: unsupervised generalizable 4D scene reconstruction
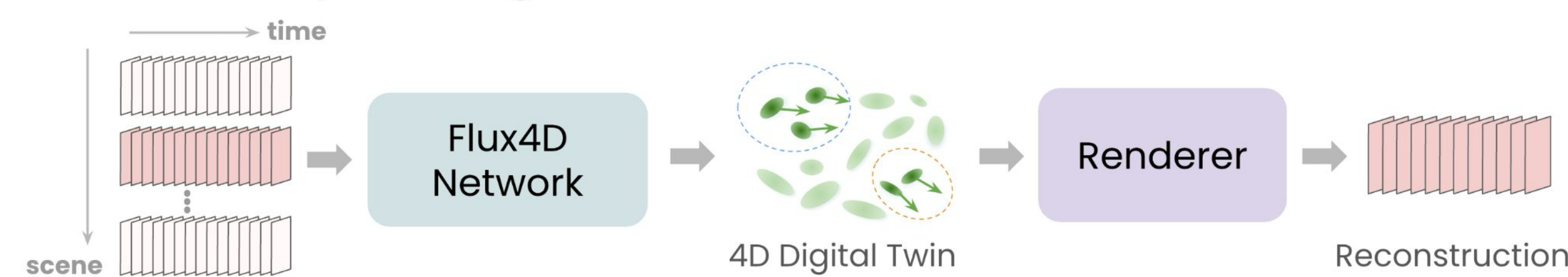
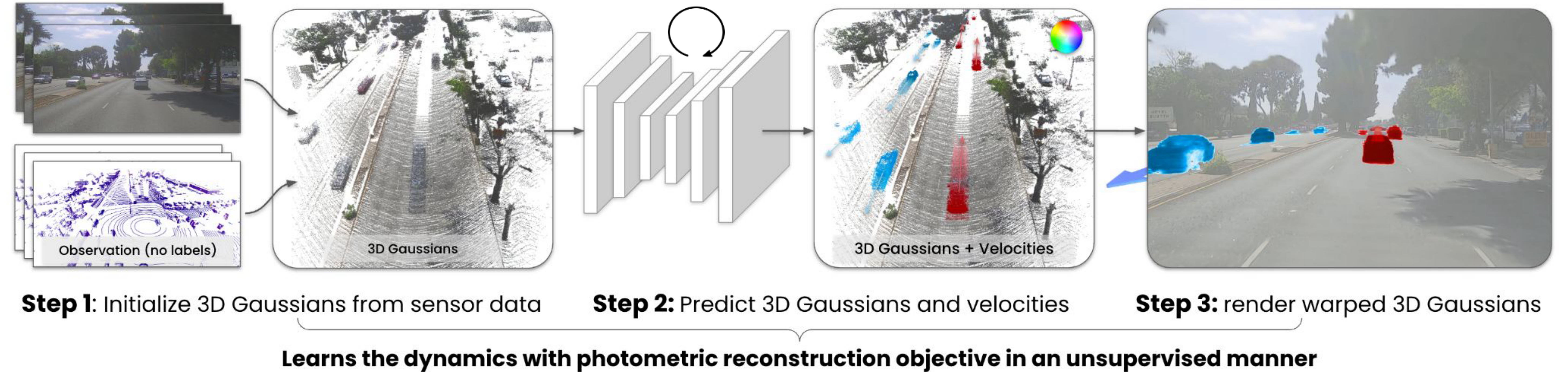## Scaling Laws & Generalizability

+ *Flux4D* scales with more data

+ *Flux4D* works on diverse and large-scale datasets

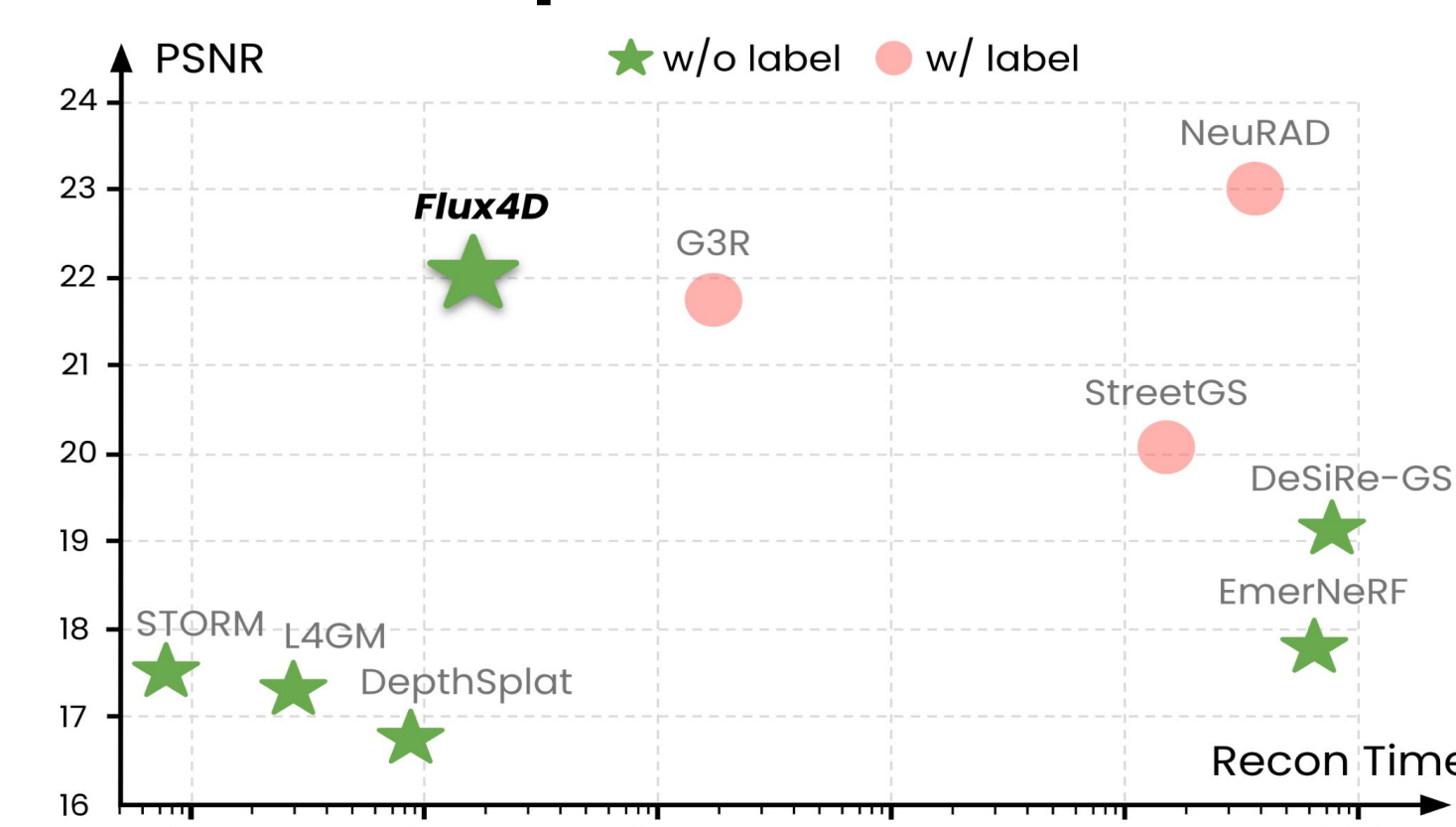

Argoverse 2 — Waymo Open Dataset

Flux4D: Flow-based Unsupervised 4D Reconstruction

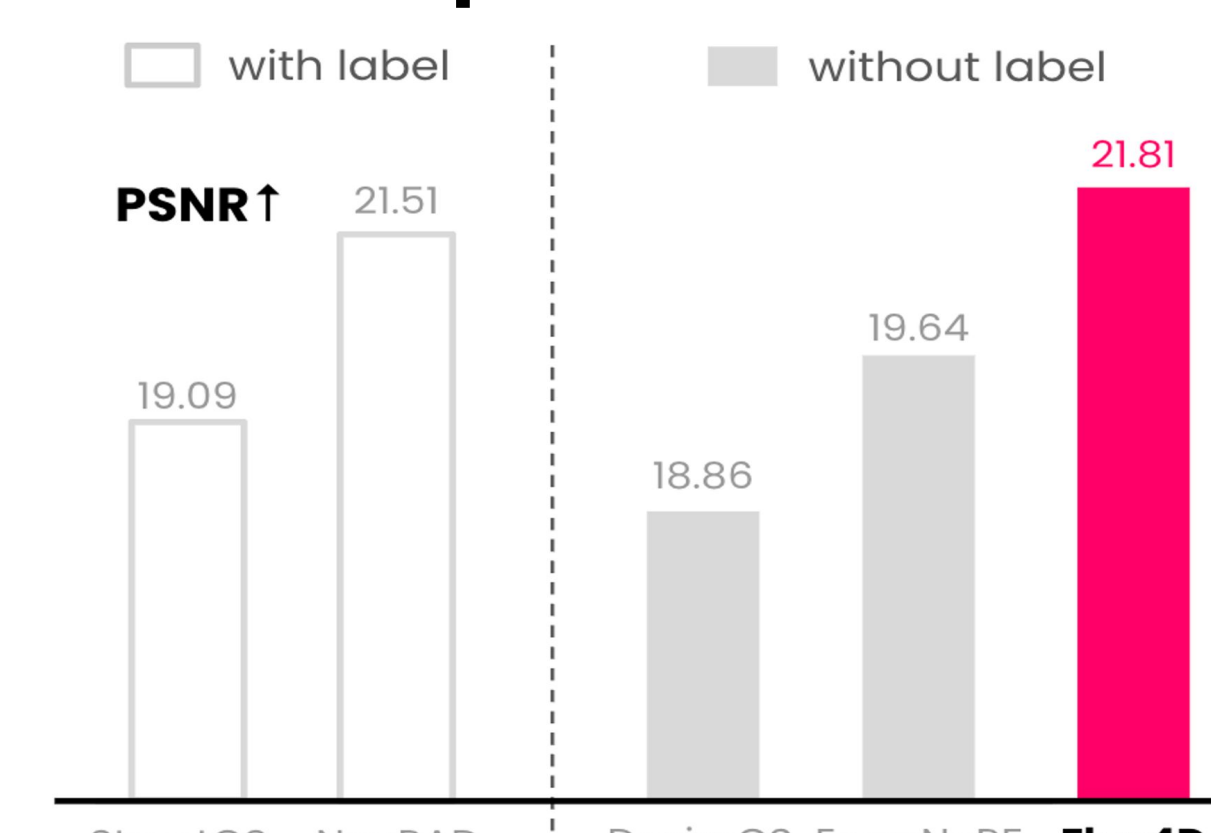+ *Flux4D learns object concepts and physical dynamics purely from observation by training on many scenes*

Step 1: Initialize 3D Gaussians from sensor data **Step 2**: Predict 3D Gaussians and velocities **Step 3**: render warped 3D Gaussians

Learns the dynamics with photometric reconstruction objective in an unsupervised manner

Results

+ **NVS comparison with SOTA**

+ **Future prediction**

L4GM — EmerNeRF — DeSiRe-GS — Ours

+ **Ablation study**

(a) Flux4D-base   (b) Flux4D-base + iterative refinement   (c) Flux4D-base + iterative + poly velocity

+ **Limitations:** (a) flow estimation for actors with complex motion; (b) generalization to unseen motion; (c) handle longer-horizon scenes

Applications

+ **Scalable 4D reconstruction**

RGB

Flow

+ **Realistic and controllable camera simulation**

Original — Velocity — Moving Instances (No Labels) — Remove Dynamics

Move Sensor Up — Move Sensor Down — Move Sensor Left — Move Sensor Forward

Move Actor Right — Move Actor Back — Move Actor Forward — Insert Actor