3D Gaussian Splatting Scene Reconstruction for Autonomous Driving

Hao-Yu Chan, Peng-Chen Chen, Yung-Ching Sun

Department of Electrical Engineering and Computer Science, University of Michigan

Introduction & Motivation

Importance of 3D Driving Scene Reconstruction

- Accurate geometric and appearance for **BEV perception**, **scene understanding**, **localization**, and **planning**.
- **Novel view synthesis** or **realistic and controllable simulation** for system testing and validation.

Challenges and Current Gaps

- Sparse, multi-timestep, multi-view sensor data.
- High-speed, complex, and diverse dynamic objects.

Our Goals

- Compare three SOTA 3DGS reconstruction methods: **Street Gaussians [1]**, **OmniRe [2]**, **STORM [3]**.
- Identify strengths and limitations for each approach.

Methods Overview

Novel Strategies for Each Method

- **Street Gaussians:** 4D spherical harmonics appearance model and tracked pose optimization.
- **OmniRe:** Model diverse, non-rigid dynamic actors from occlusions and cluttered environments.
- **STORM:** Feed-forward, self-supervised method. Learns 3D Gaussians and scene flow jointly.

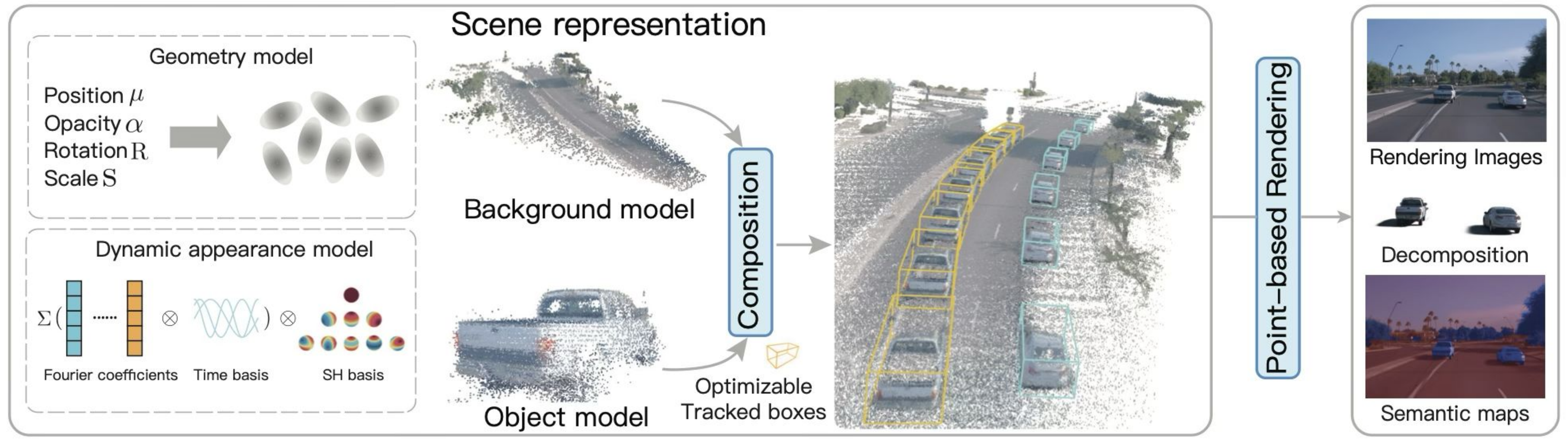
Overall Comparison

	Input	Decomp.	Rendering
Street	I + PC	3D BBox	I + D + Veh. P + Sem
OmniRe	I + P + PC	3D BBox	I + D
STORM	I + P	Self	I + D + SFlow

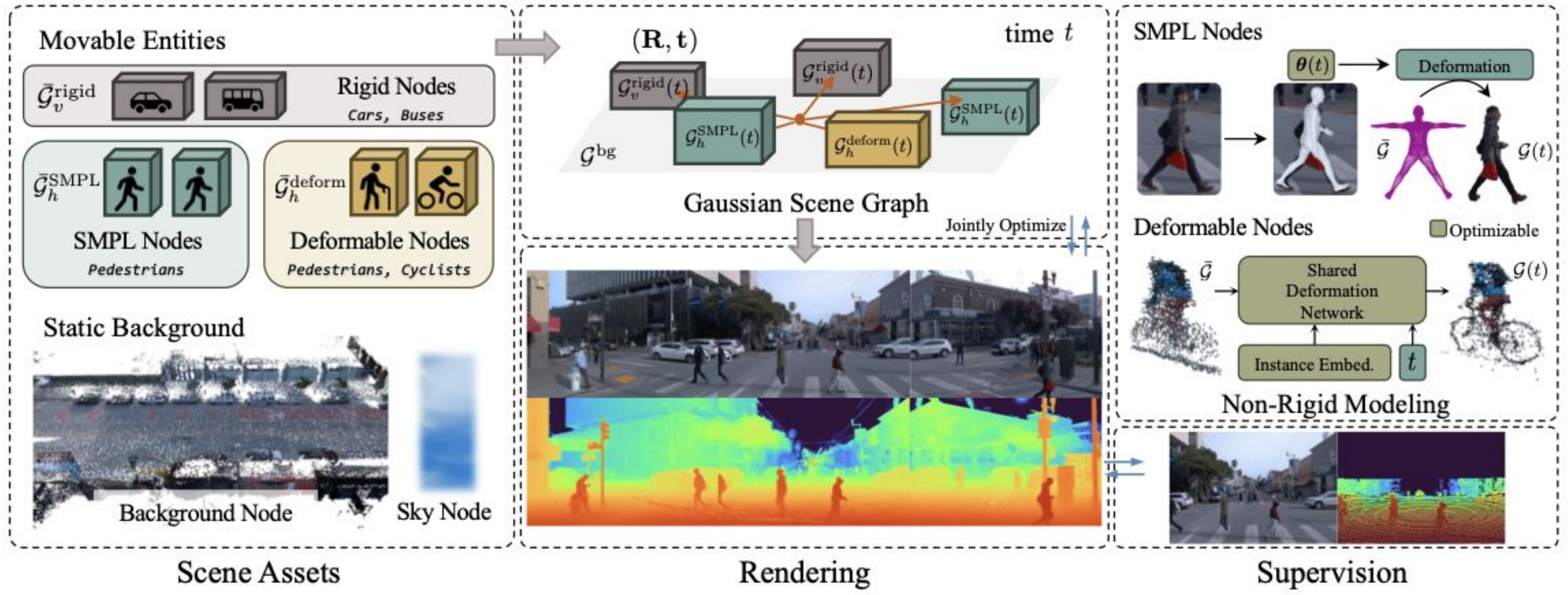
- **Input: I**: Image, **P**: Pose, **PC**: Point Cloud
- **Decomp: BBox:** Bounding Box, **Self**: Self-Supervised
- **Rendering: D**: Depth, **Sem**: Semantic, **SFlow:** 3D Scene Flow

Methodology

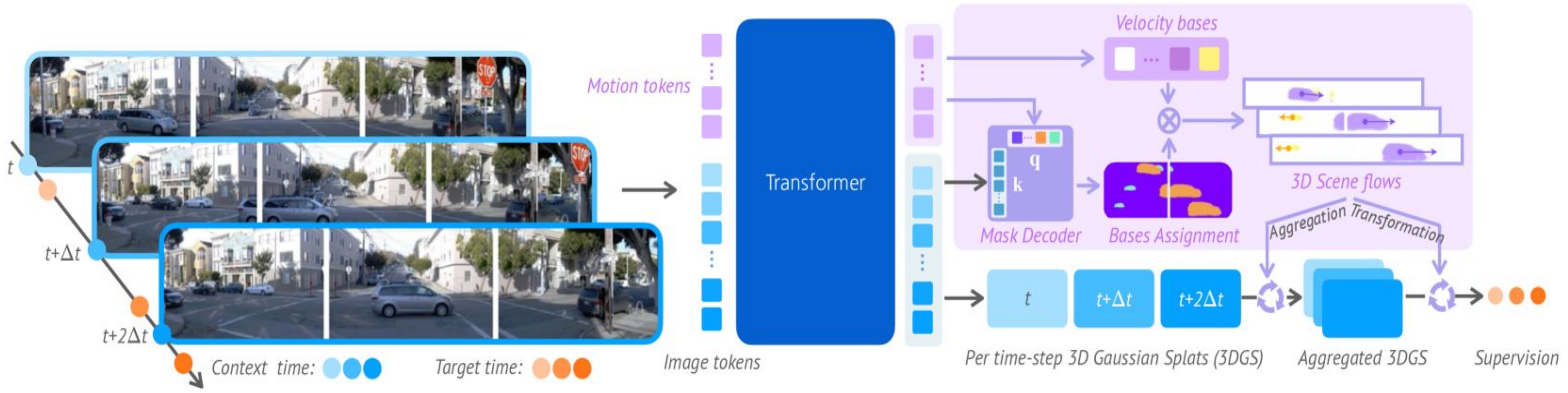
Street Gaussian: Modeling Dynamic Urban Scenes with Gaussian Splatting (ECCV 2024)



OmniRe: Omni Urban Scene Reconstruction (ICLR 2025)



STORM: Spatio-Temporal Reconstruction Model for Large-Scale Outdoor Scenes (ICLR 2025)



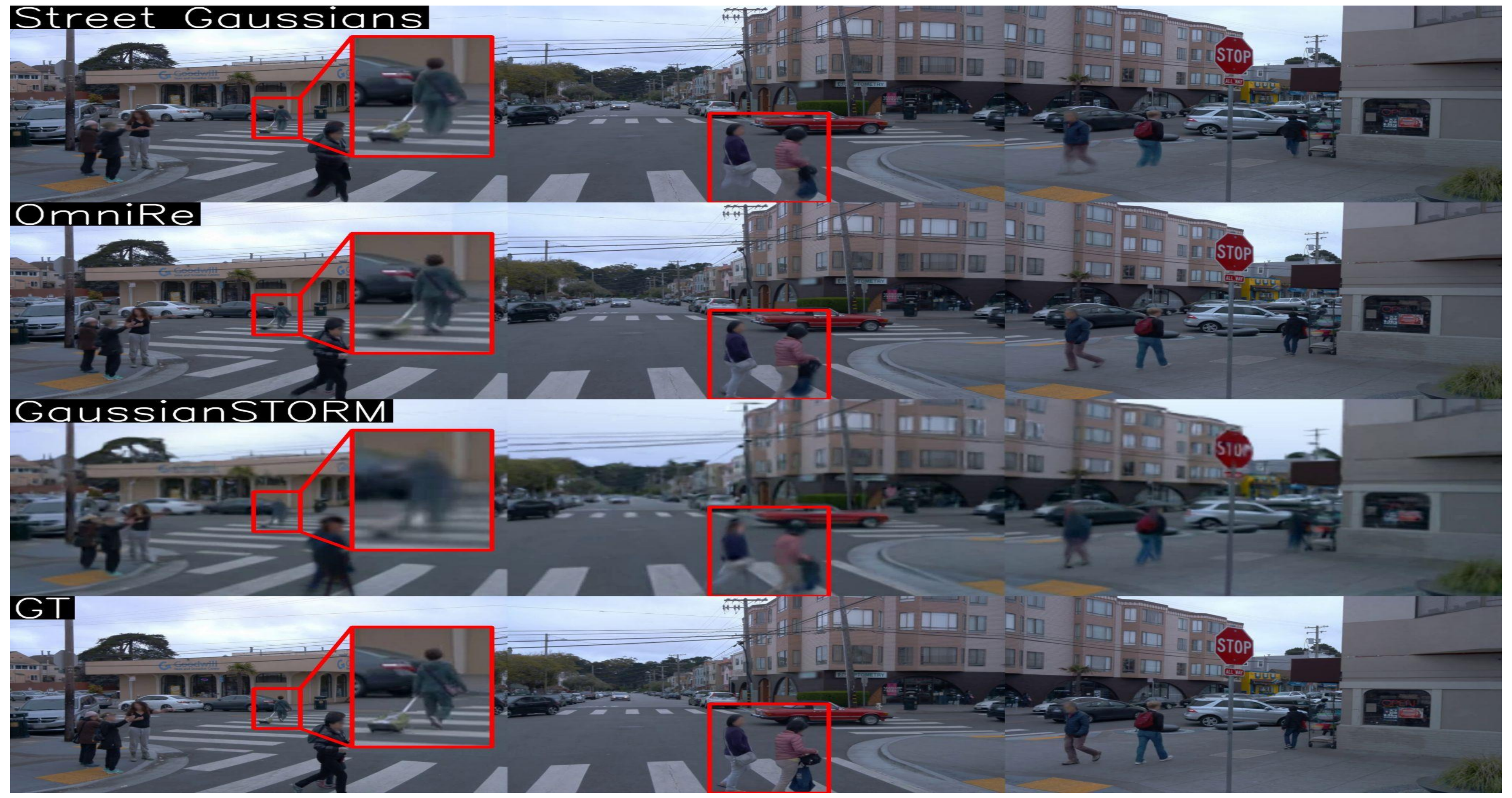
Results

On Waymo Dataset

	Scene_id 023			Scene_id 552		
	PNSR(↑)	SSIM(↑)	LPIPS(↓)	PNSR(↑)	SSIM(↑)	LPIPS(↓)
Street	31.49	0.9443	0.0576	30.64	0.9181	0.0965
OmniRe	**35.32**	**0.9599**	0.0494	**33.36**	**0.9462**	0.0749
STORM	30.95	0.9274	**0.0455**	29.92	0.9008	**0.0574**

On nuScenes Dataset

	Scene_id 000			Scene_id 003		
	PNSR(↑)	SSIM(↑)	LPIPS(↓)	PNSR(↑)	SSIM(↑)	LPIPS(↓)
Street	26.20	0.8101	0.2716	27.73	0.8585	0.2401
OmniRe	**28.28**	**0.8755**	**0.1894**	**29.35**	**0.8981**	**0.1852**
STORM	-	-	-	-	-	-



Conclusions & Potential Future Work

- SOTA methods mask dynamics using **fixed** labels (e.g., humans, cars); this can be improved by using **bag-of-words** and **SAM3** to capture all dynamic masks.
- Unhandled **lighting variations** may lead to visual harmony problems; This can be solved by **building a light model**.
- Novel view synthesis may fail under **large camera trajectory deviations**; **video generative models** can be used to address this.

References:

[1] Yan, Yunzhi, Haotong Lin, Chenxu Zhou, Weijie Wang, Haiyang Sun, Kun Zhan, Xianpeng Lang, Xiaowei Zhou, and Sida Peng. "Street gaussians: Modeling dynamic urban scenes with gaussian splatting." In European Conference on Computer Vision, pp. 156-173. Cham: Springer Nature Switzerland, 2024.

[2] Chen, Ziyu, Jiawei Yang, Jiahui Huang, Riccardo de Lutio, Janick Martinez Esturo, Boris Ivanovic, Or Litany et al. "OmniRe: Omni Urban Scene Reconstruction." In The Thirteenth International Conference on Learning Representations, 2025.

[3] Yang, Jiawei, Jiahui Huang, Boris Ivanovic, Yuxiao Chen, Yan Wang, Boyi Li, Yurong You et al. "STORM: Spatio-TempOral Reconstruction Model For Large-Scale Outdoor Scenes." In The Thirteenth International Conference on Learning Representations, 2025.