



MonoNeRF: Learning a Generalizable Dynamic Radiance Field from Monocular Videos

ICCV23

PARIS

Fengrui Tian, Shaoyi Du, Yueqi Duan



Fengrui's website



Code

Video



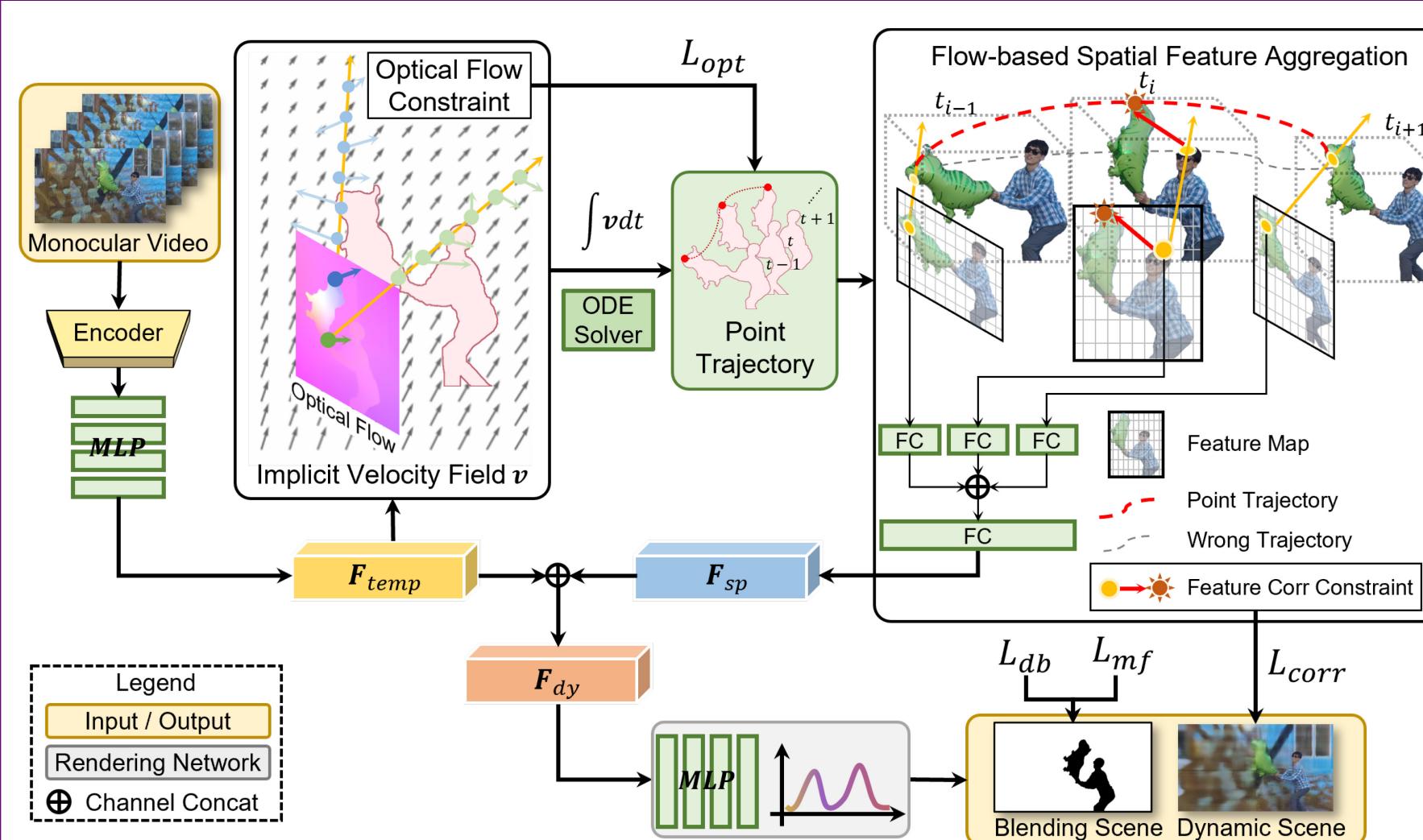
TL; DR: We propose a radiance field that could **generalize** to multiple dynamic scenes.

Fengrui Tian is also **looking for Ph.D. positions** in computer vision in Fall 2024!

Motivation

- Learning dynamic radiance fields from monocular videos suffers from 2D-to-3D ambiguity problem.
- Previous works (such as DynNeRF [1]) use positional embedding to break the ambiguity, but have limited transferable ability to multiple scenes.
- Learn a **generalizable** dynamic radiance field.

Overview



1. Learn an implicit velocity field from \mathbf{F}_{temp} .
2. Exploit the spatial feature \mathbf{F}_{sp} with flow-based spatial feature aggregation module.
3. Incorporate \mathbf{F}_{temp} and \mathbf{F}_{sp} as the point feature \mathbf{F}_{dy} for rendering dynamic scene.

Spatio-temporal Constraints

- Spatial constraint

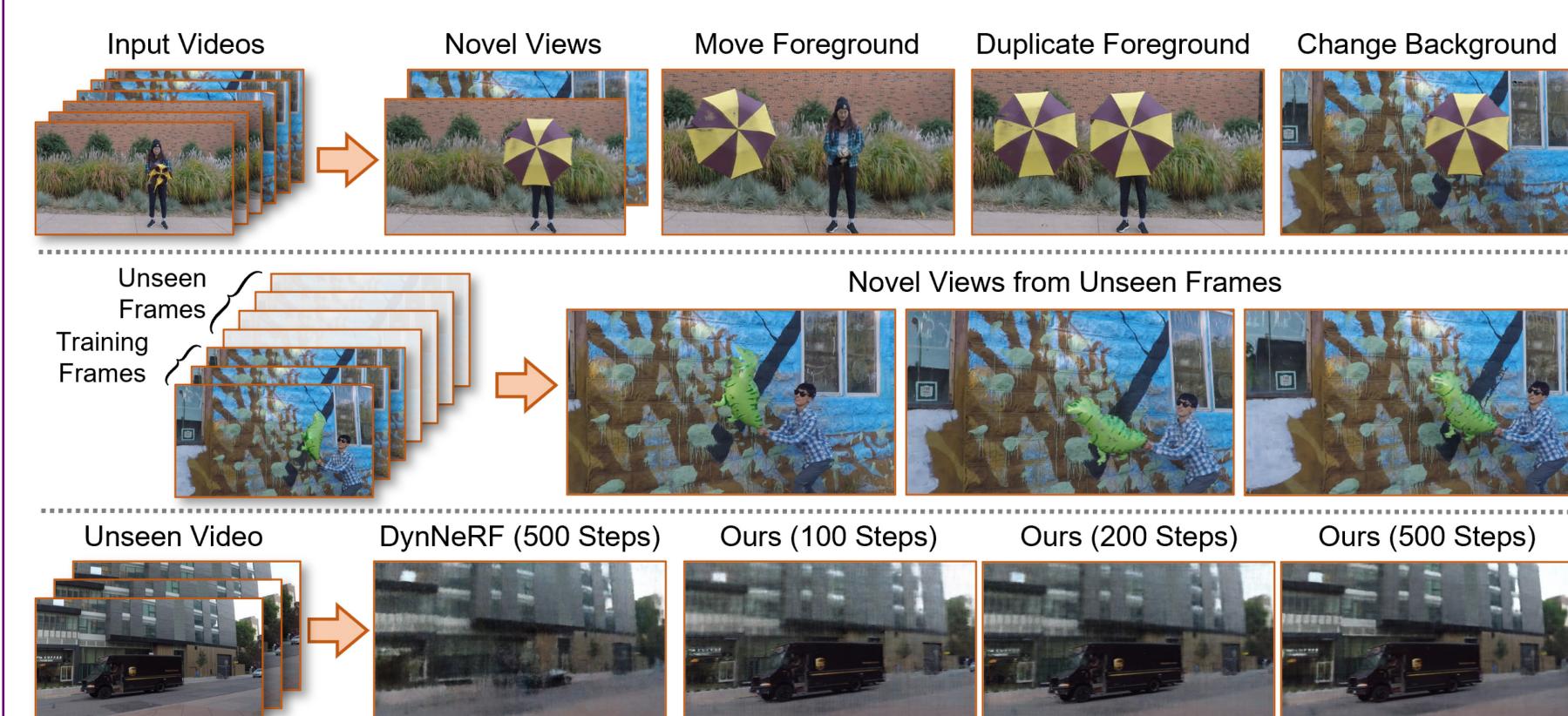
$$L_{\{bw, curr, fw\}} = \sum_r \|\mathbf{C}_{dy}(\mathbf{r}_{\{bw, curr, fw\}}) - \mathbf{C}_{dy}^{gt}(\mathbf{r})\|_2$$

$$L_{corr} = L_{bw} + L_{curr} + L_{fw}$$

- Temporal constraint

$$L_{opt} = \sum_r \left(\mathbf{f}_{\{bw, fw\}}(\mathbf{r}) - \mathbf{f}_{\{bw, fw\}}^{gt}(\mathbf{r}) \right)$$

New Applications



- Scene editing (**Top**)

- Unseen frame synthesis (**Middle**)

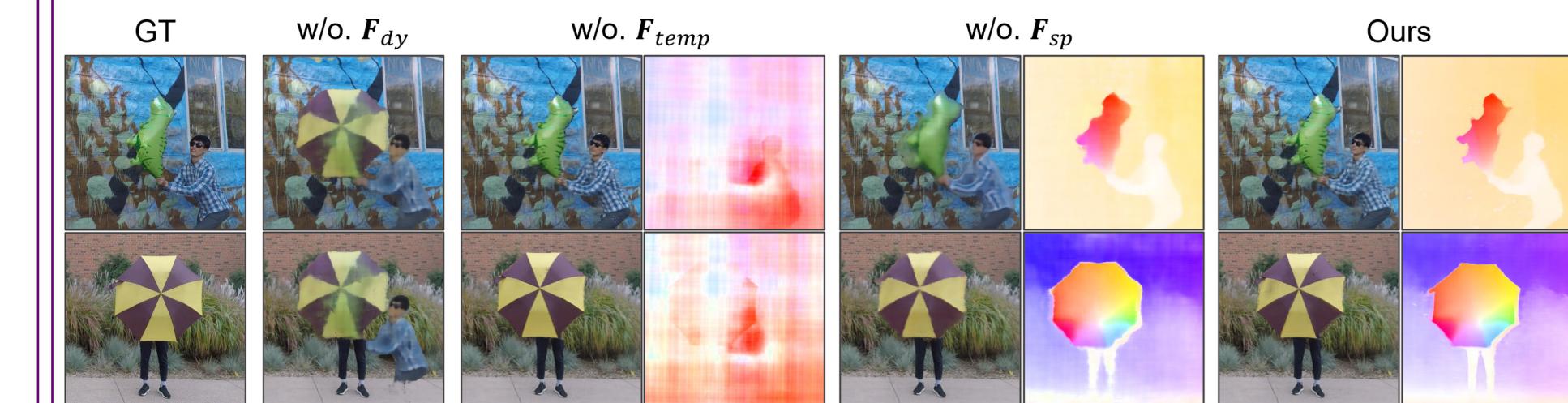
- Novel scene adaption (**Bottom**)

Reference: [1] Gao, C., et al. Dynamic view synthesis from dynamic monocular video. In ICCV, 2021.

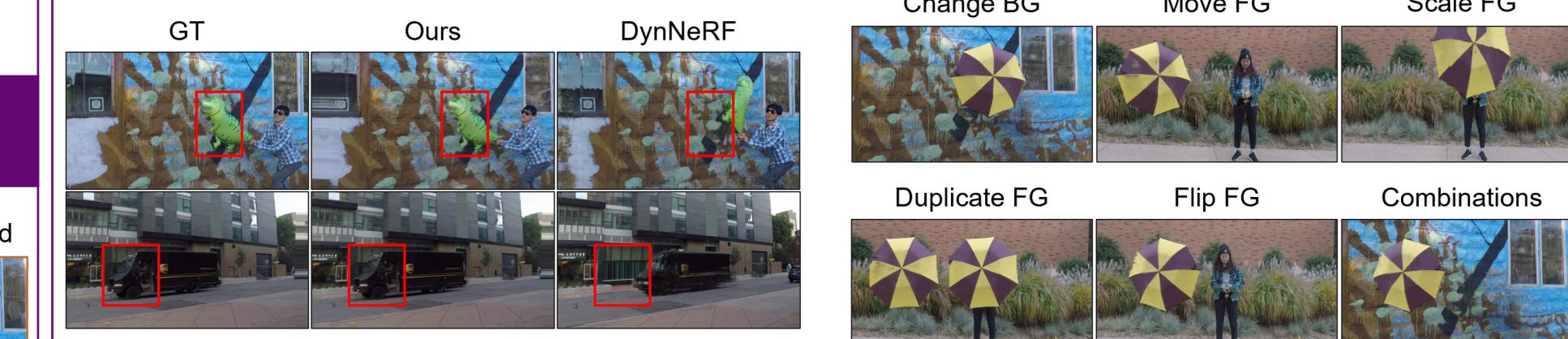
Experimental Results

We strongly recommend to watch our [Youtube Video](#).

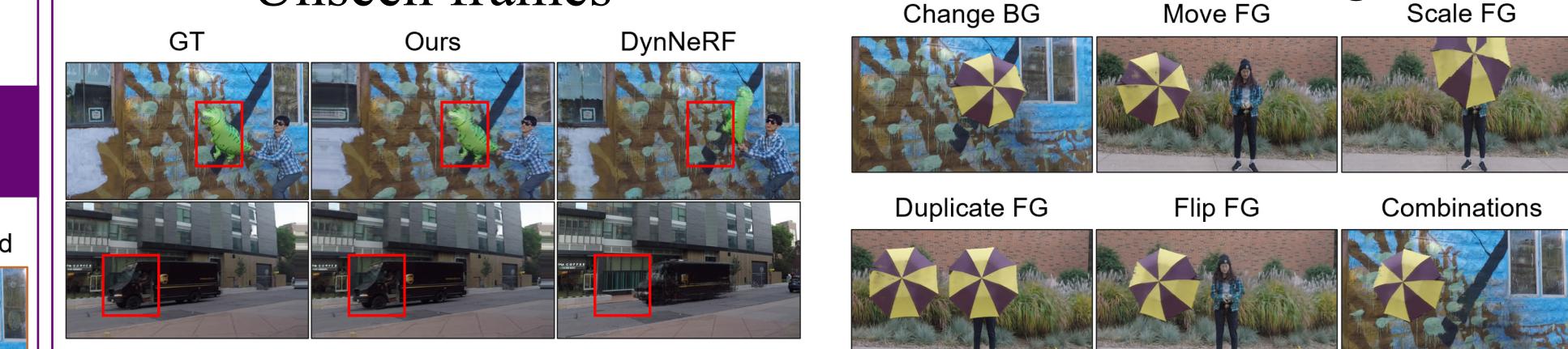
Novel view synthesis from multiple videos



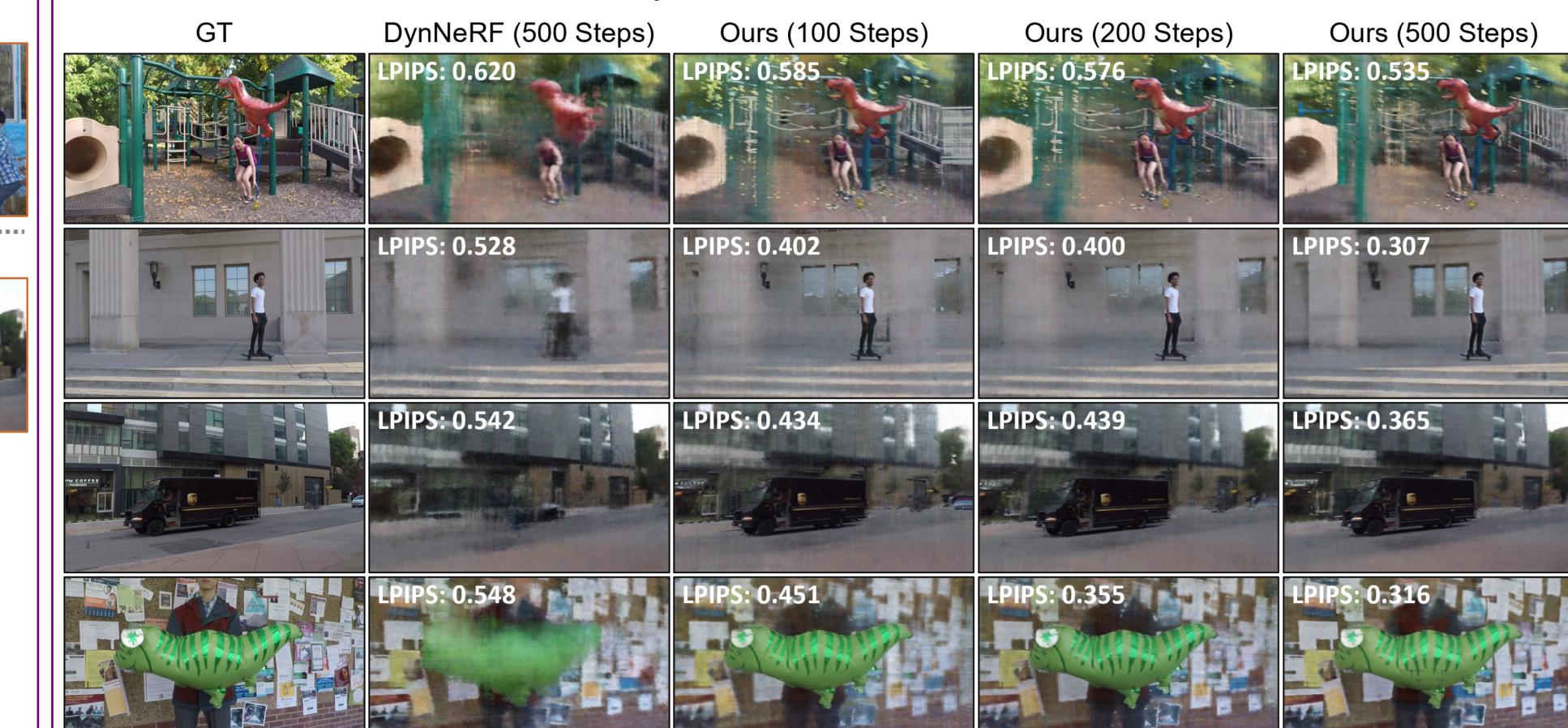
Unseen frames



Scene editing



Novel view synthesis from unseen videos



Contact Fengrui Tian: tianfr1999@gmail.com