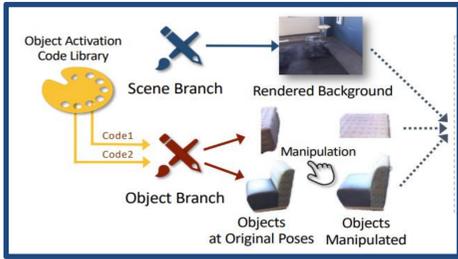


# Instruct-Object2Object: Object-Aware Editable Neural Radiance Fields With Prompts

Yuyang Tang, Xutong Jiang, Chenghao Gong  
Department of Computer Science, University of Toronto

## Motivation

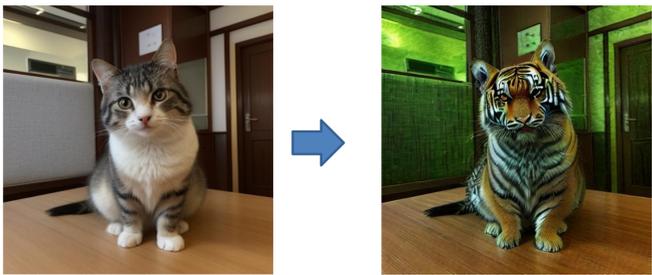
- Object-NeRF addresses object-compositional rendering but requires manual annotation and faces issues with occlusions in post-editing.



- Given prompt, Instruct-NeRF2NeRF edits each image from data set as a whole without separating the object of interest and the scene.



- This project aim to design a handy pipeline for object-aware neural rendering from customized prompts. Our model is designed to render the object and the scene separately, which enable specific editing on objects and scene, as well as object displacement like moving, rotating in the scene.



## Related Work

### Neural Radiance Fields (NeRF):

Neural rendering methods such as Neural Radiance Fields (NeRF) have demonstrated their capability for novel view synthesis, and various adaptations have shown their ability for editable scene rendering.

### Diffusion Model:

Simulating the process of data generation by gradually adding random noise to an initial distribution, diffusion models are widely used in generating image data.

### Zero-shot Segmentation:

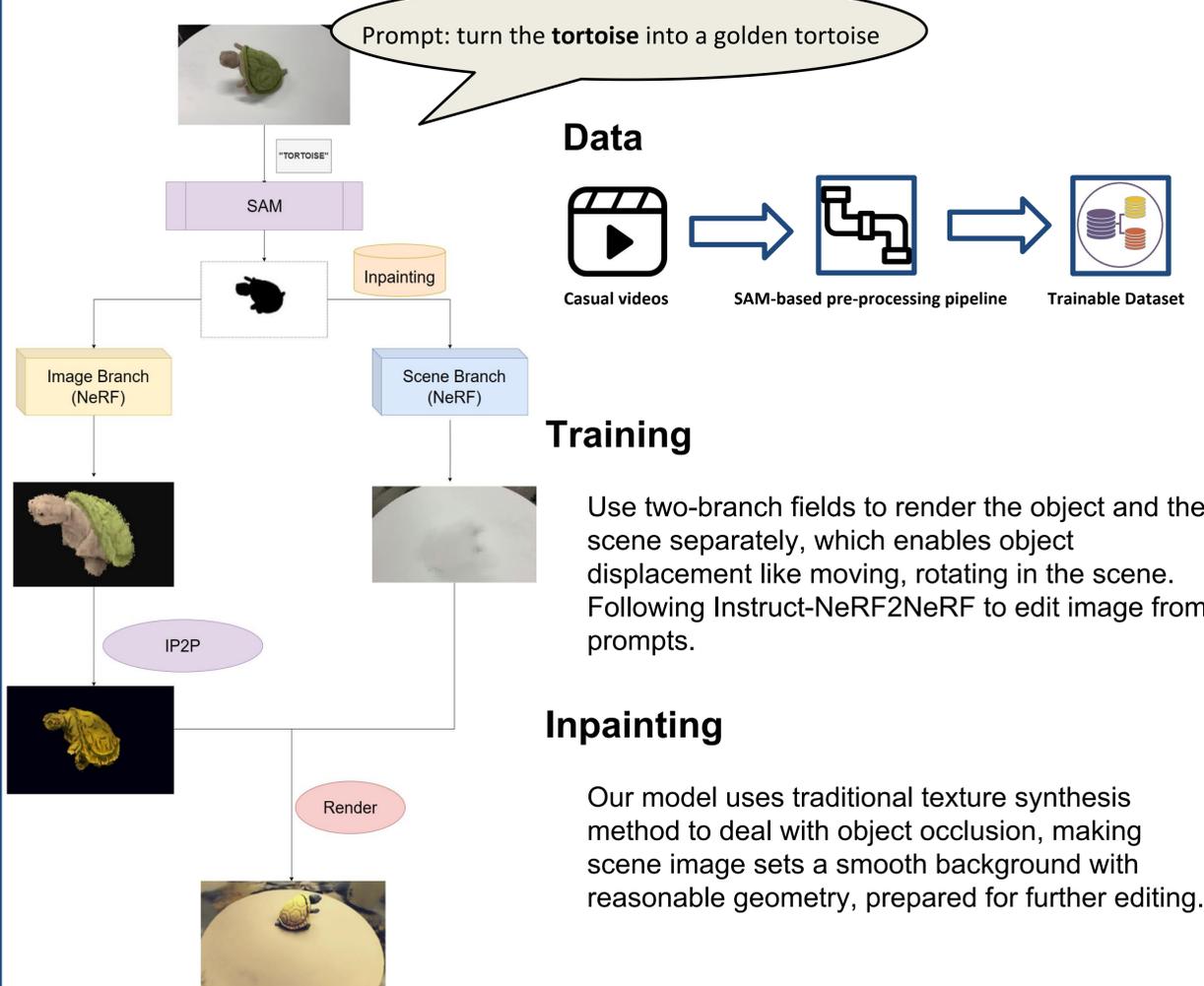
SAM facilitates zero-shot transfer to various tasks through prompt engineering, and it can be trained to be promptable for downstream applications.

## References

- [1] B. Mildenhall, P. P. Srinivasan, M. Tancik, J. T. Barron, R. Ramamoorthi, and R. Ng, "Nerf: Representing scenes as neural radiance fields for view synthesis," CoRR, vol. abs/2003.08934, 2020.
- [2] B. Yang, Y. Zhang, Y. Xu, Y. Li, H. Zhou, H. Bao, G. Zhang, and Z. Cui, "Learning object-compositional neural radiance field for editable scene rendering," CoRR, vol. abs/2109.01847, 2021.
- [3] A. Haque, M. Tancik, A. A. Efros, A. Holynski, and A. Kanazawa, "Instruct-nerf2nerf: Editing 3d scenes with instructions," 2023.
- [4] T. Brooks, A. Holynski, and A. A. Efros, "Instructpix2pix: Learning to follow image editing instructions," 2023.

## New Technique

### Flow of Work



### Training

Use two-branch fields to render the object and the scene separately, which enables object displacement like moving, rotating in the scene. Following Instruct-NeRF2NeRF to edit image from prompts.

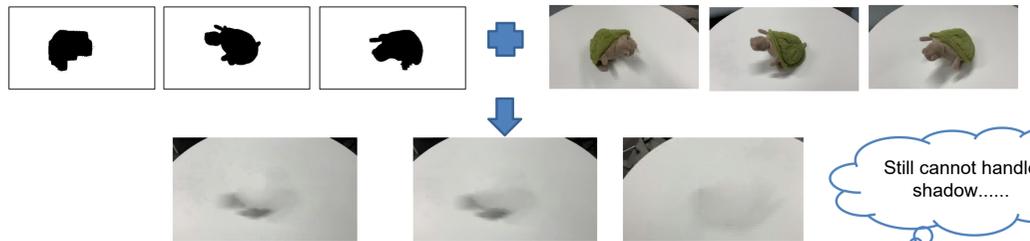
### Inpainting

Our model uses traditional texture synthesis method to deal with object occlusion, making scene image sets a smooth background with reasonable geometry, prepared for further editing.

## Experimental Results

We filmed our own datasets, named barry\_tortoise, toy\_sheep, and desk\_apple. Example results of the flow are listed below.

### Segmentation + Inpainting Results:

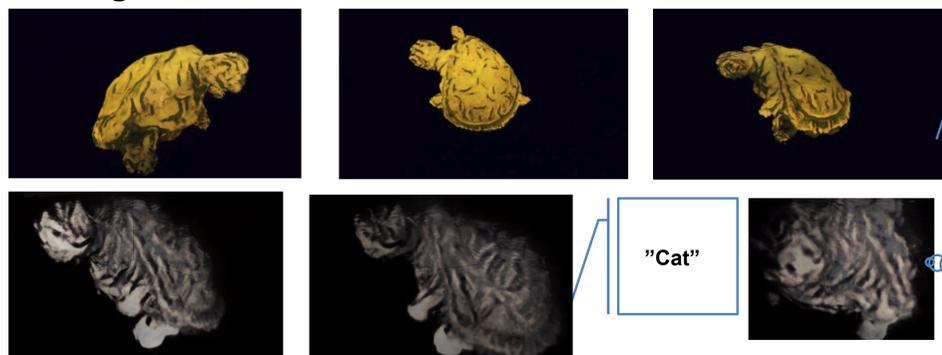


### Reconstruction Results:

- Our field of object branch reconstructs object for the first stage:



### Editing:



"Golden Tortoise"

"Cat"

Where is head?