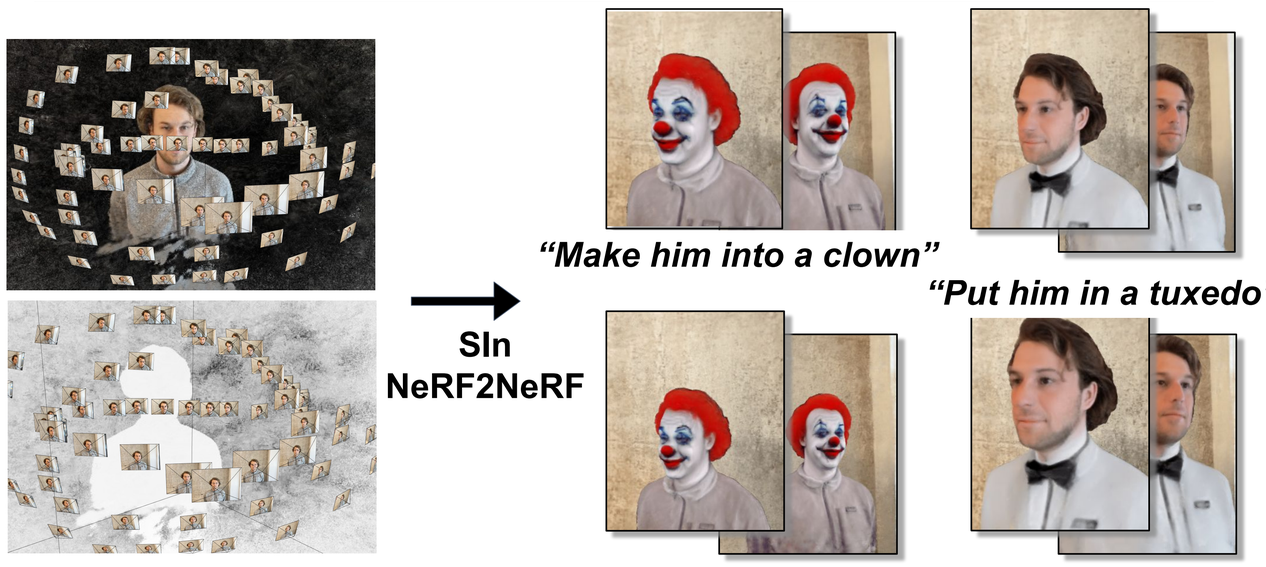
SIn-NeRF2NeRF: Editing 3D Scenes with Instructions through Segmentation and Inpainting

Performance Improvement

Jiseung Hong, Changmin Lee, Gyusang Yu

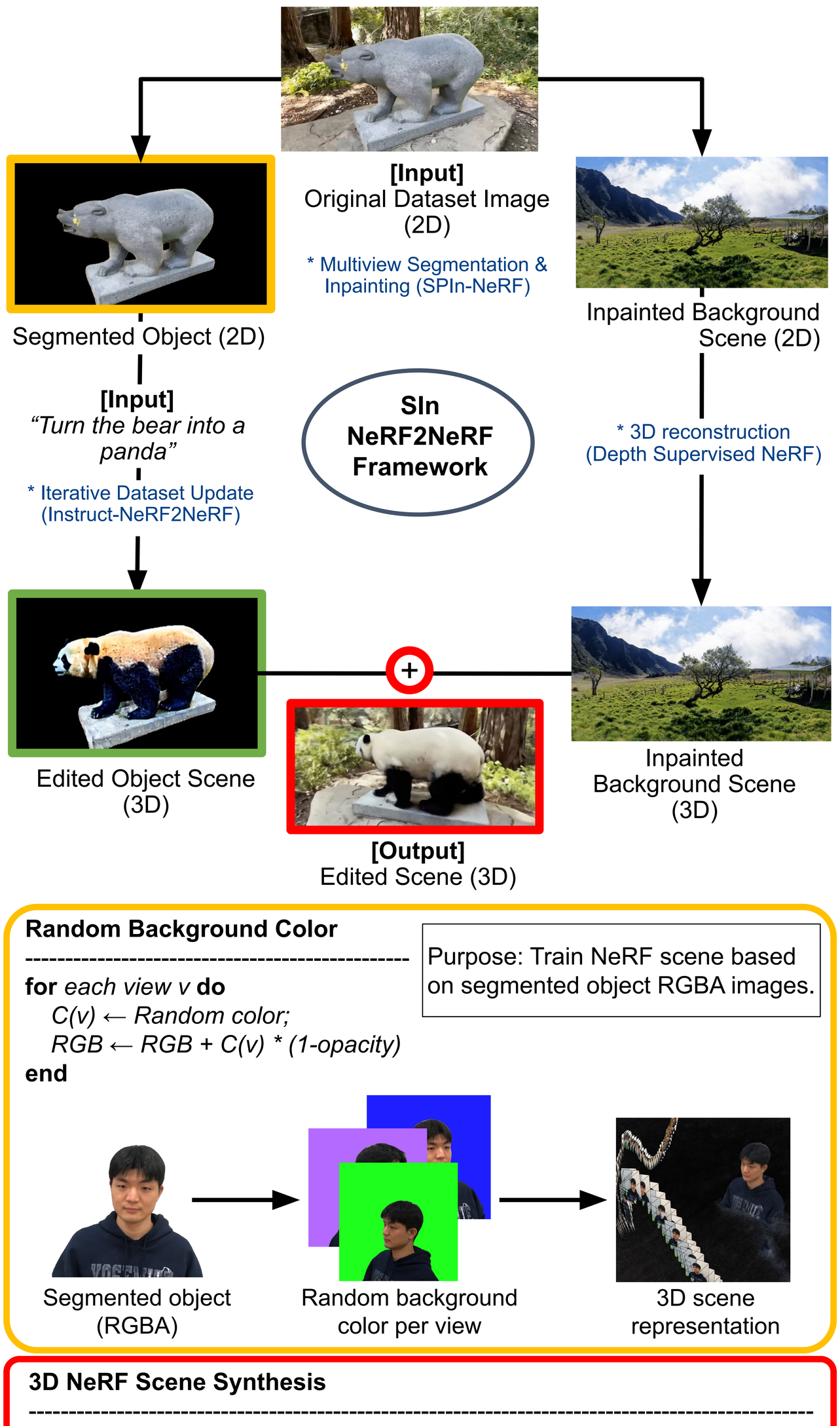
KAIST CS479 Project Fall 2023

Overview



TL;DR Perform 3D object editing selectively by disentangling it from the background scene.

Main Idea



Random Background Color

---

for *each view v* **do**
  C(v) ← Random color;
  *RGB ← RGB + C(v) * (1-opacity)*
end

Purpose: Train NeRF scene based on segmented object RGBA images.

3D NeRF Scene Synthesis

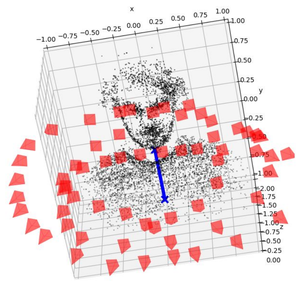
- Object and background scene share the same camera parameters.
- Sort the sampled points for the same rays by depth values.

Main Results

Object Transformation

- We made the object transformation (**scaling, translation, rotation**) possible by disentangling the object from the background scene.

- We use COLMAP to acquire the **coordinates** and the **centroid** of the 3D object.



Baseline Comparison

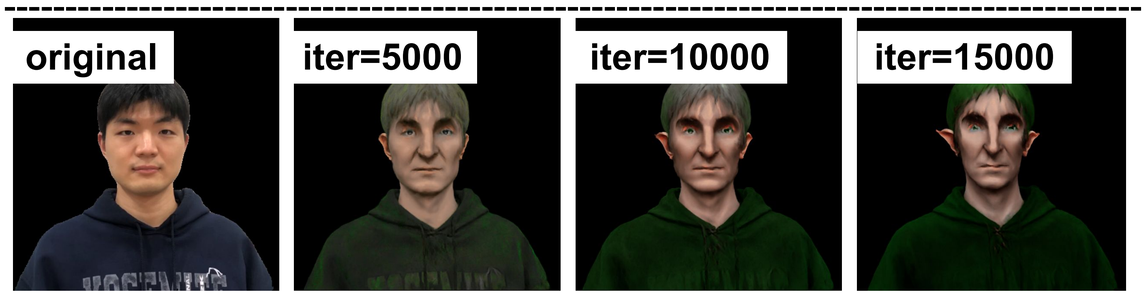
Quantitative Results

Scene (face) \ CLIP	Text-Image Similarity		Direction Consistency	
	in2n	sn2n	in2n	sn2n
Clown scene	**0.2372**	0.2081	0.9071	**0.9117**
Tuxedo scene	0.0251	**0.0481**	0.8451	**0.8599**

Iterative Dataset Update (for RGBA)

---

for *each iteration* **do**
  for *each viewpoints v* **do**
    1. Alpha blend RGBA image w. black background;
    2. Update image using ip2p;
    3. Segment the object;
  end
end



"Turn him into a tolkien elf"

References

[1] Haque et al., Instruct-NeRF2NeRF: Editing 3D Scenes with Instructions , ICCV 2023 (Oral).

[2] Mirzaei et al., SPIn-NeRF: Multiview Segmentation and Perceptual Inpainting with Neural Radiance Fields, CVPR 2023

Acknowledgements

Changmin Lee has implemented the main pipeline. Jiseung Hong and Gyusang Yu revised the code and performed qualitative/quantitative analysis. Data acquisition, methodology discussion and report conduction were equally done.