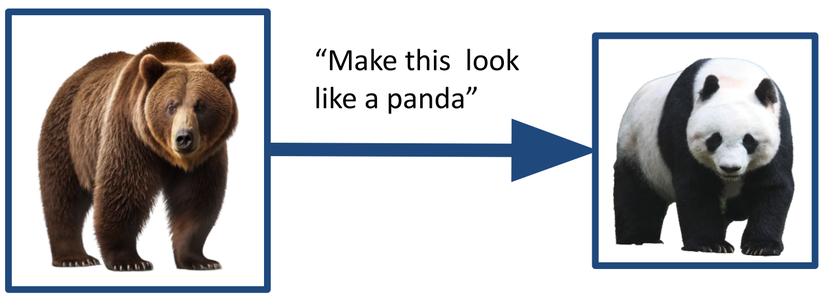
Text-Based Style Transfer on 3D Objects with Gaussian Splatting

Noé Artru, Adam Barroso
Department of Computer Science, University of Toronto

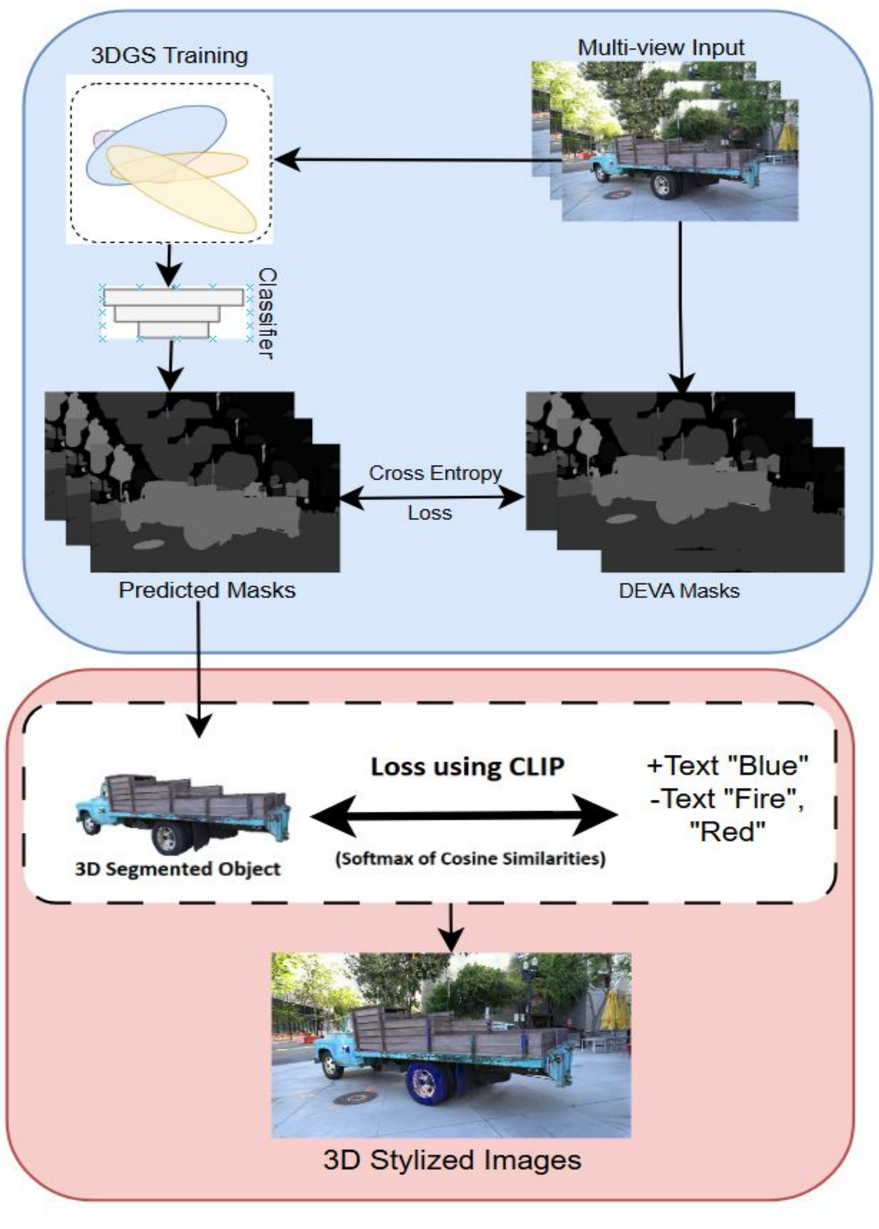
Motivation

- Enable intuitive 3D scene editing using text-based prompts



- Descriptive text is simpler to create than Image-based inputs, and can leverage tools such as speech-to-text for easier integration.
- Leverage advancements in language encoders, which more closely align with desired results compared to image encoders, making text-to-image pipelines a more effective approach for scene editing.
- Lay groundwork for more accessible and flexible 3D editing tools, moving closer to integrating seamless text-based scene transformation into practical applications.

New Technique



Related Work

Gaussian Splatting
- Efficient representation and rendering of 3D scenes using gaussian shaped objects

NeRF-Art
- NeRF based 3D scene stylization with text
- NeRF architecture demands significant compute and is difficult to interpret

Style-Splat
- Uses images for style transfer of objects in 3D scenes using Gaussian Splatting
- Doesn't support textual input

Contrastive Language-Image Pre-training
- Pre-trained model which encodes images and text into the same feature space
- Allows us to define a loss based on similarity between text and image

References

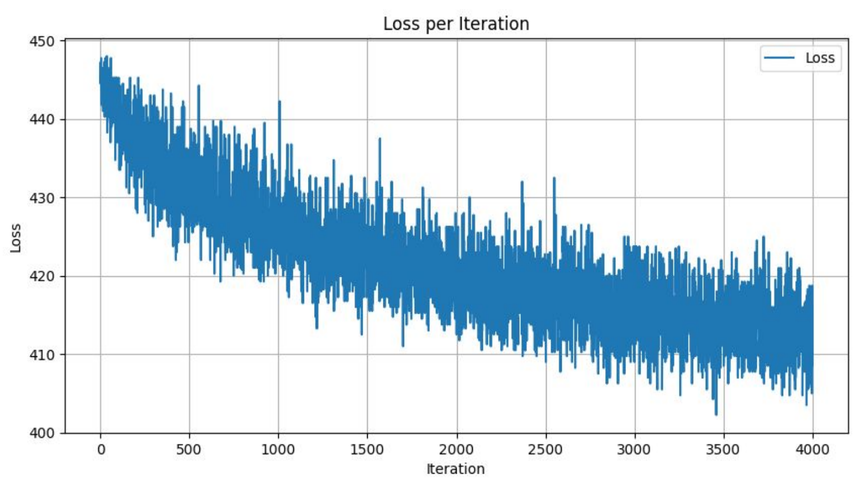
[1] Kerbl, Kopanas, 3D Gaussian Splatting for Real-Time Radiance Field Rendering, SIGGRAPH,2023
[2] NeRF-Art: Text-Driven Neural Radiance Fields Stylization, Wang, 2022
[3] Jain, Kuthiala, StyleSplat: 3D Object Style Transfer with Gaussian Splatting, 2024
[4] Tracking Anything with Decoupled Video Segmentation, Cheng, ICCV 2023
[5] Ke, Gaussian Grouping: Segment and Edit Anything in 3D Scenes, ECCV, 2024
[6] Learning Transferable Visual Models From Natural Language Supervision, Radford, Kim, 2021

Experimental Results

Original render

"Fire" input using image generator

"Blue" +encoding
"Fire" -encoding
Using our CLIP-based loss



Training loss evolution using our CLIP-based loss