





# Locally Stylized Neural Radiance Fields

Hong-Wing Pang<sup>1</sup>, Binh-Son Hua<sup>2,3</sup>, Sai-Kit Yeung<sup>1</sup> <sup>1</sup> Hong Kong University of Science and Technology <sup>2</sup> Trinity College Dublin



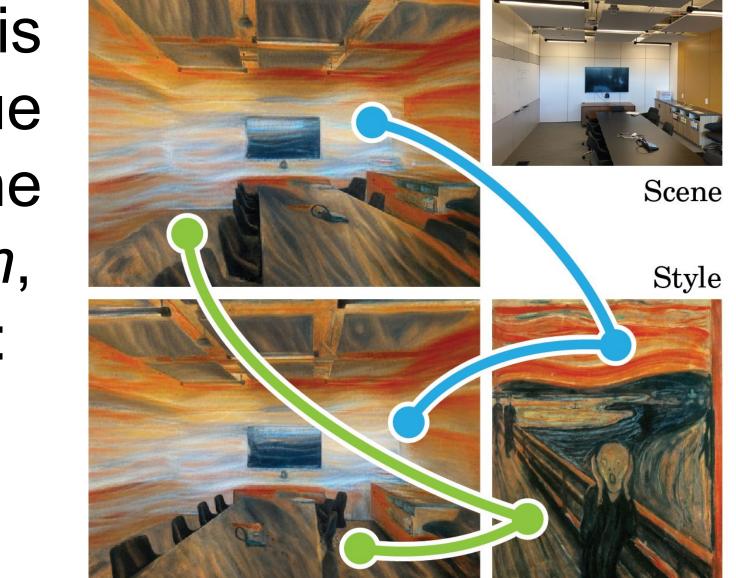
#### Motivation

NeRF Style transfer - Transferring the style of a reference image onto a NeRF scene in a consistent way across different novel views.

- Previous stylization methods suffer from repetitive patterns and lack of controllability.
- We propose a novel method with improved and customizable stylization results.

#### Methodology Overview

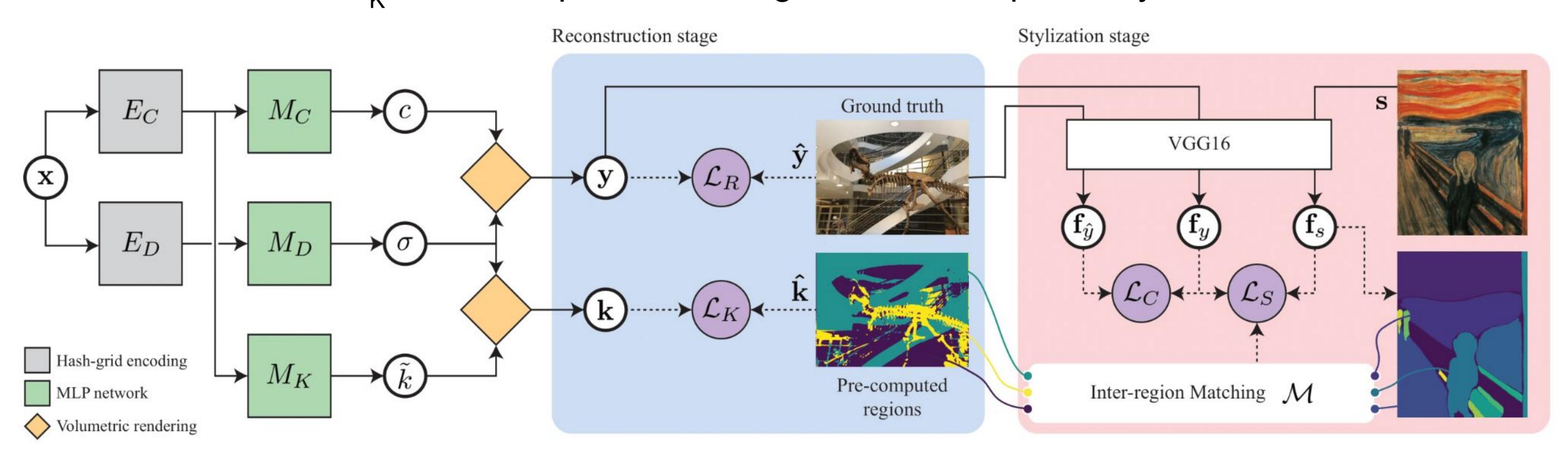
- Style image is segmented into S style regions
- Each ground truth image is segmented into C scene regions
- Reconstruction stage: NeRF model is trained to learn the following simultaneously:
  - Render ground truth images via regular NeRF training with MSE loss
  - Predict the scene region segmentation map for any arbitrary novel view
- Each scene region is matched with a unique style region via the Hungarian algorithm, based on the following:



- Position of region
- Feature similarity
- Stylization stage: NeRF model is fine-tuned using the following losses:
- Style loss (L<sub>s</sub>): for each VGG feature from rendered image, compute distance with nearest style image VGG feature in matched style region
- Content loss (L<sub>c</sub>): L2 distance between content and style feature maps

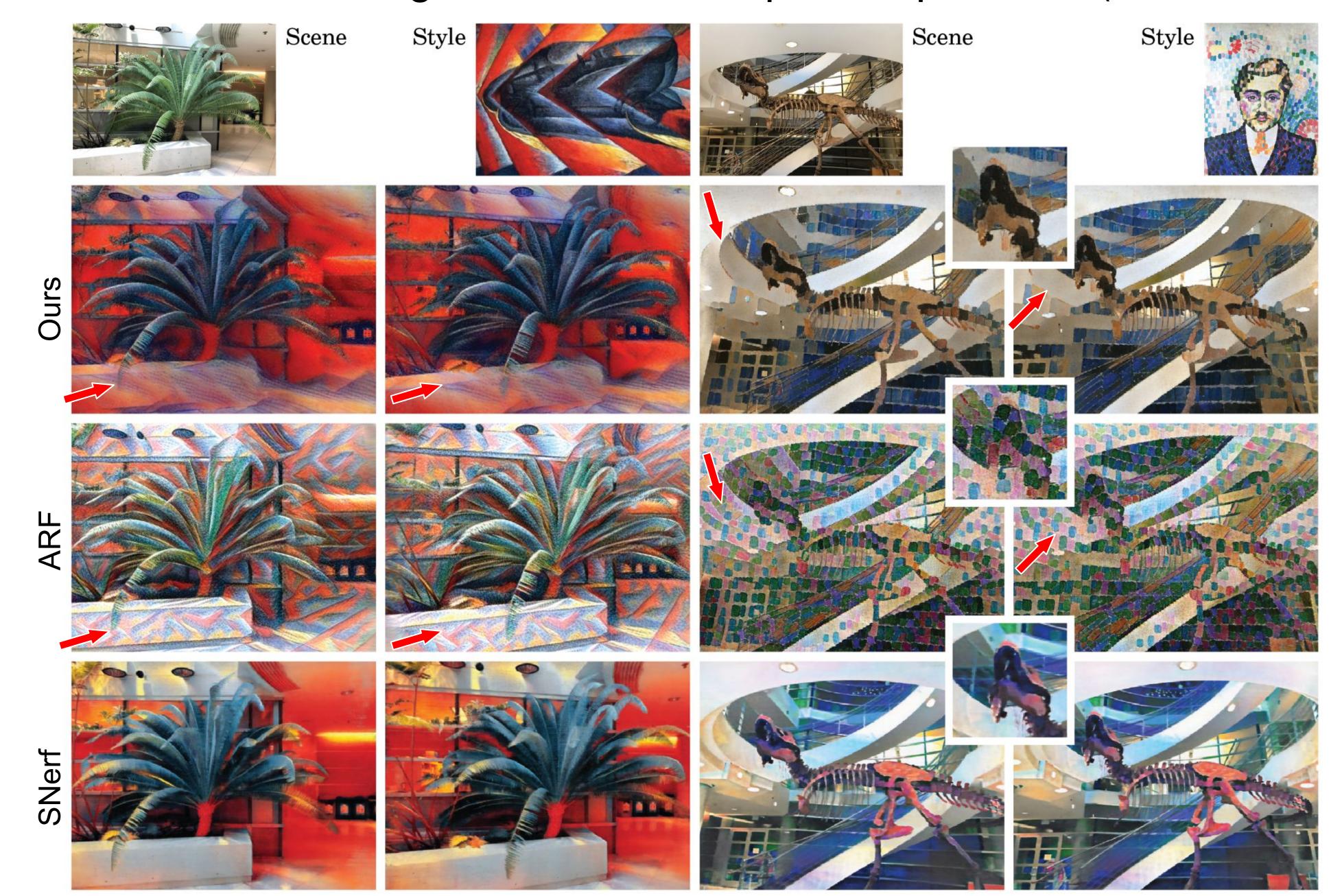
## Training Pipeline

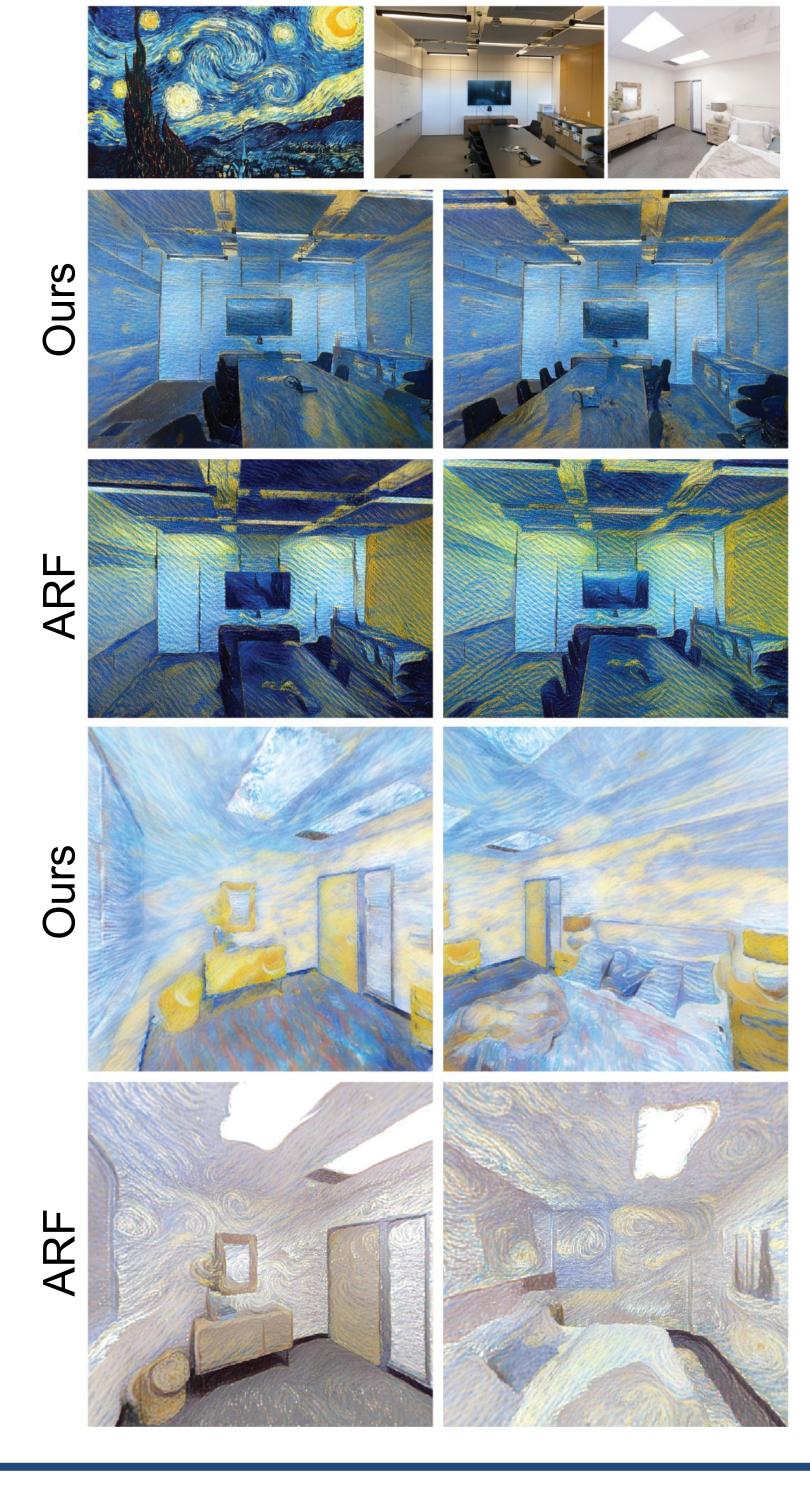
- We use a variant of Instant-NGP [1] as the NeRF backbone to facilitate fast training.
- Two separate hash grids (for color and density) are used to store the positional encodings of any arbitrary input position x. During the stylization stage, only  $E_c$  is fine-tuned.
- The MLP network M<sub>k</sub> is used to predict the segmentation maps of any rendered novel view.



#### Baseline Comparisons with SNerf [2] and ARF [3]

Under our method, different local styles across the style image can be transferred, reducing the amount of repetitive patterns. (see red arrows)





#### Additional Functionalities

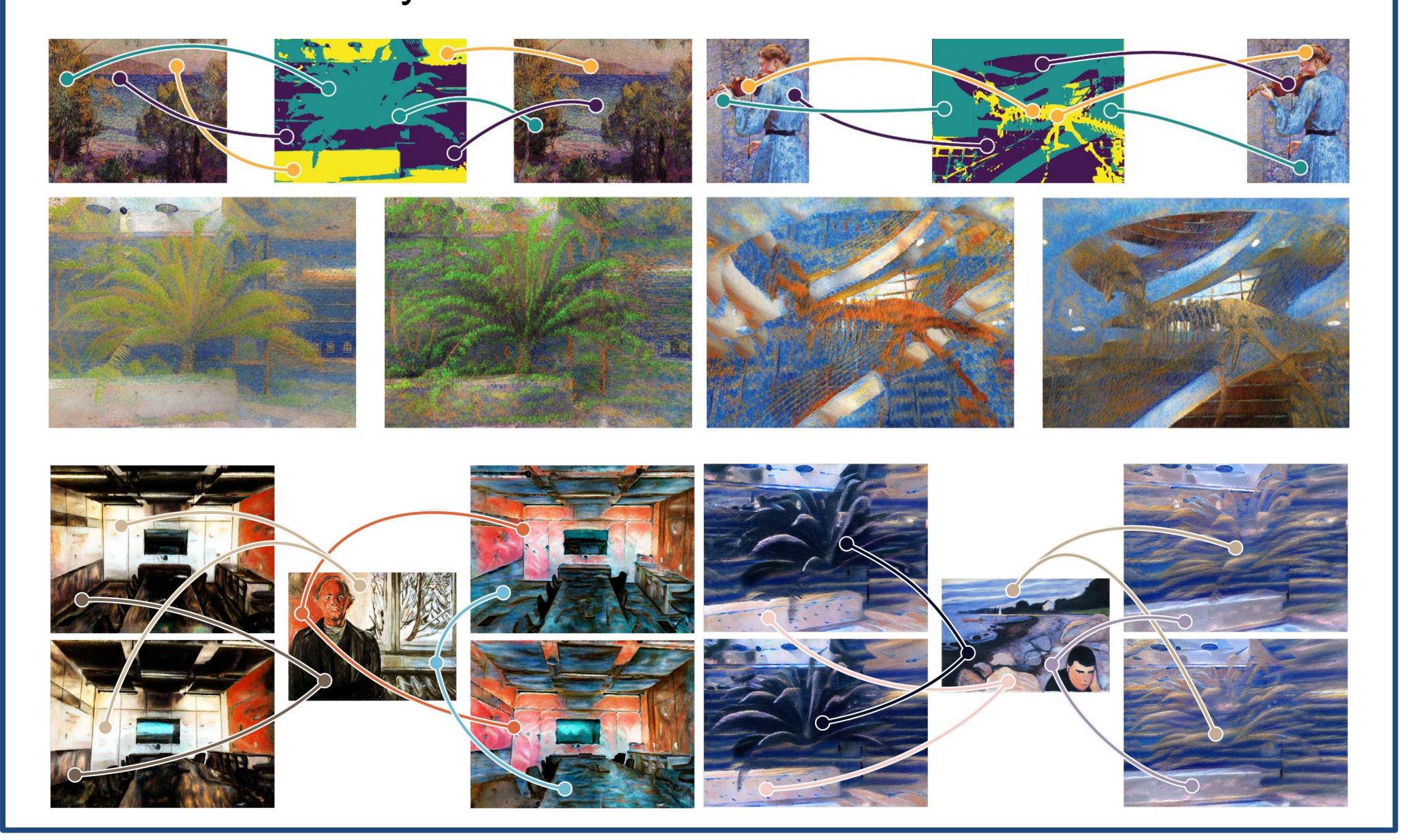
We modify the hash function used in the positional encoding hash grid to provide the following features:

- Modify hash coefficients to get different stylization results.
- Simultaneously train multiple styles within a single model.



### Customization of stylization

The matching between scene / style regions can be altered manually to obtain diverse stylization results.



#### References

- [1] Thomas Muller, Alex Evans, Christoph Schied, and Alexander Keller. Instant neural graphics primitives with a multiresolution hash encoding. ACM Trans. Graph., 2022.
- [2] Thu Nguyen-Phuoc, Feng Liu, and Lei Xiao. Snerf: Stylized neural implicit representations for 3d scenes. ACM Trans. Graph., 2022.
- [3] Kai Zhang, Nick Kolkin, Sai Bi, Fujun Luan, Zexiang Xu, Eli Shechtman, and Noah Snavely. Arf: Artistic radiance fields. ECCV 2022.

