NeRF in the Wild
Neural Radiance Fields for Unconstrained Photo Collections

Tabisheva Anastasia, 171

What is NeRF-W?

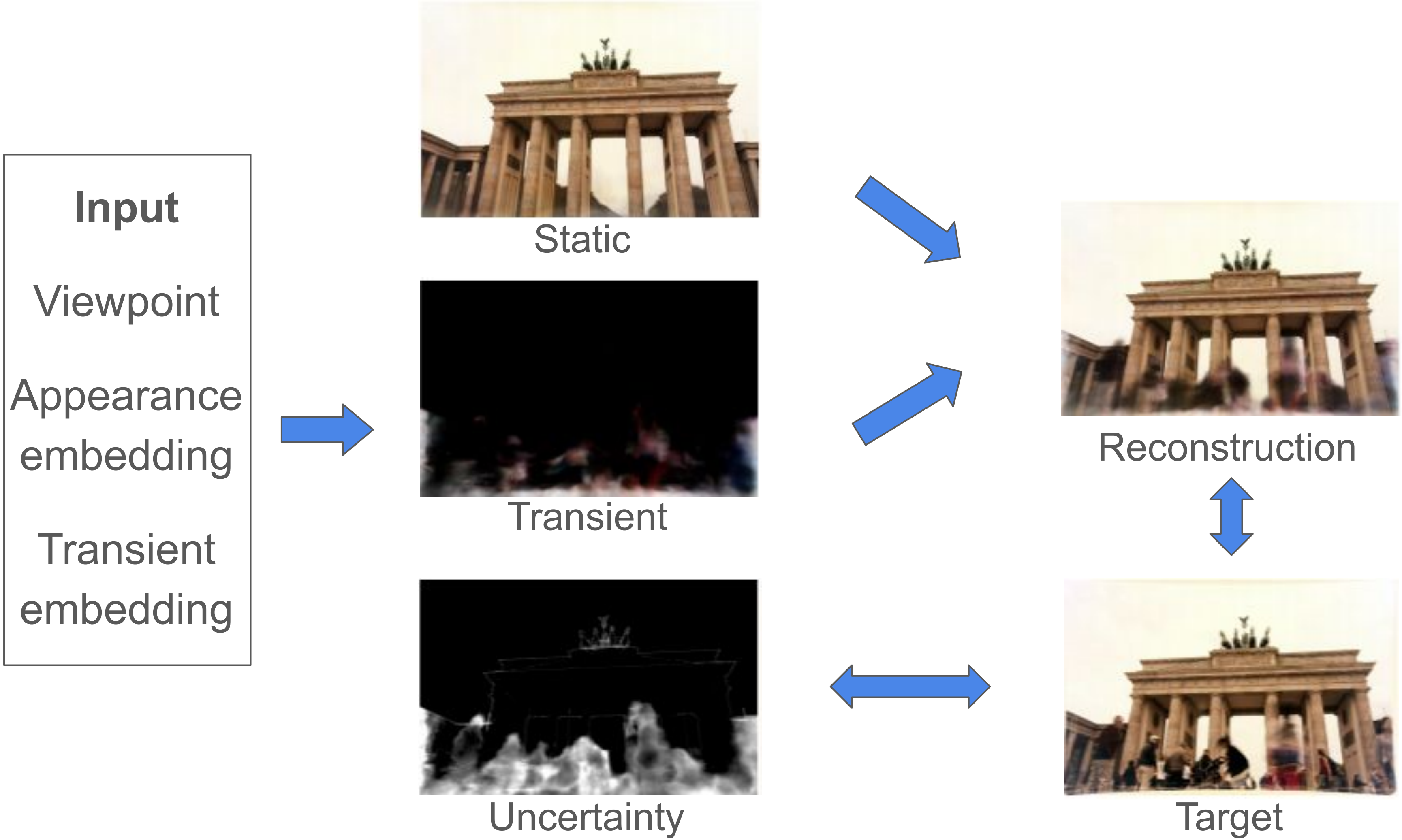
NeRF-W (Neural Radiance Fields in the Wild) - an extension of NeRF for synthesizing novel views of complex outdoor scenes using only unstructured collections of in-the-wild photographs

Input

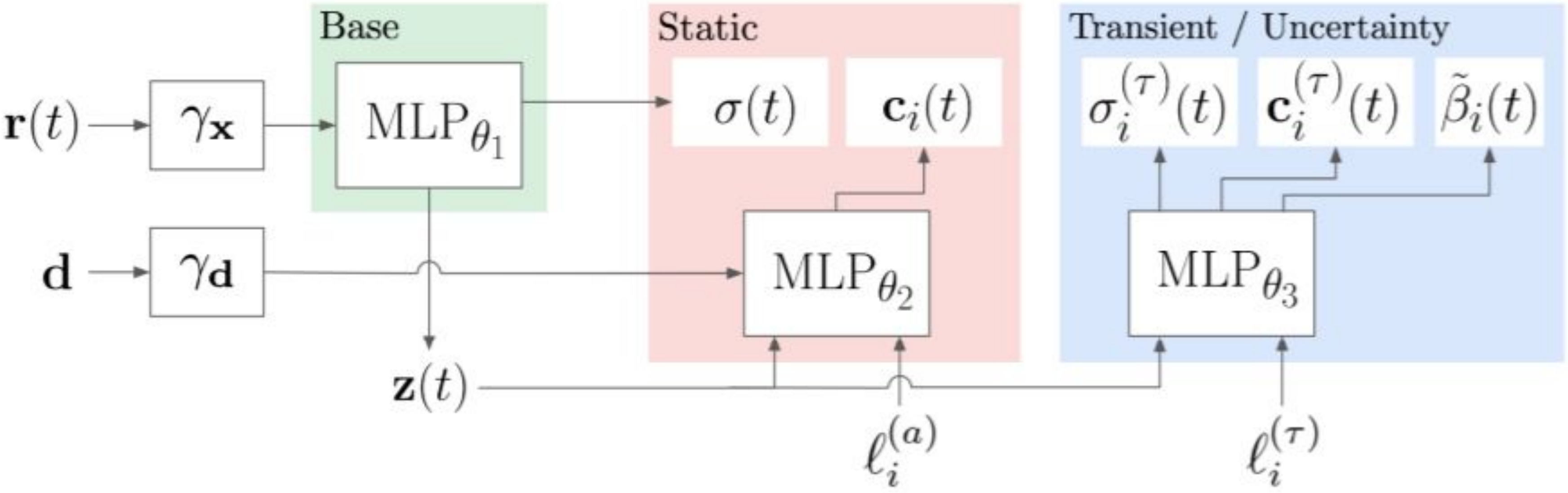
Output

Why not NeRF?

- In outdoor photography, time of day and camera settings directly impact the illumination
- Real-world landmarks are rarely captured in isolation, without moving objects or distractors around them


Input
Viewpoint
Appearance embedding
Transient embedding
Static
Transient
Uncertainty
Reconstruction
Target

NeRF-W model architecture



Adaptation NeRF to variable lighting

NeRF

$$\bar{\mathbf{C}}(\mathbf{r}) = \int_{t_n}^{t_f} T(t)\sigma(t)\mathbf{c}(t)dt$$

$$T(t) = \exp\left(-\int_{t_n}^{t} \sigma(s)ds\right)$$

$$\mathbf{c}(t) = \mathrm{MLP}_{\theta_2}(\mathbf{z}(t), \gamma_{\mathbf{d}}(\mathbf{d}))$$

NeRF-W

$$\bar{\mathbf{C}}_i(\mathbf{r}) = \int_{t_n}^{t_f} T(t)\sigma(t)\mathbf{c}_i(t)dt$$

$$T(t) = \exp\left(-\int_{t_n}^{t} \sigma(s)ds\right)$$

$$\boxed{\mathbf{c}_i(t)} = \mathrm{MLP}_{\theta_2}\left(\mathbf{z}(t), \gamma_{\mathbf{d}}(\mathbf{d}), \boxed{\ell_i^{(a)}}\right)$$

Adaptation NeRF to transient phenomena

$$\bar{\mathbf{C}}_i(\mathbf{r}) = \int_{t_n}^{t_f} T_i(t)\left(\sigma(t)\,\mathbf{c}_i(t) + \boxed{\sigma_i^{(\tau)}(t)\,\mathbf{c}_i^{(\tau)}(t)}\right)dt \qquad T_i(t) = \exp\left(-\int_{t_n}^{t}\left(\sigma(s) + \boxed{\sigma_i^{(\tau)}(s)}\right)ds\right)$$

(we augment static density and radiance with transient counterparts)

$$\left[\sigma_i^{(\tau)}(t), \mathbf{c}_i^{(\tau)}(t), \tilde{\beta}_i(t)\right] = \mathrm{MLP}_{\theta_3}\left(\mathbf{z}(t), \boldsymbol{\ell}_i^{(\tau)}\right)$$

(new MLP in NeRF-W for transient parts)

NeRF-W's loss function

$$Loss = \sum_{ij} L_i(\mathbf{r}_{ij}) + \frac{1}{2}\left\|\mathbf{C}(\mathbf{r}_{ij}) - \hat{\mathbf{C}}_c(\mathbf{r}_{ij})\right\|_2^2$$

$$L_i(\mathbf{r}) = \frac{1}{2\beta_i(\mathbf{r})^2}\left\|\mathbf{C}_i(\mathbf{r}) - \hat{\mathbf{C}}_i(\mathbf{r})\right\|_2^2 + \frac{1}{2}\log\beta_i(\mathbf{r})^2 + \frac{\lambda_u}{K}\sum_{k=1}^{K}\sigma_i^{(\tau)}(t_k)$$

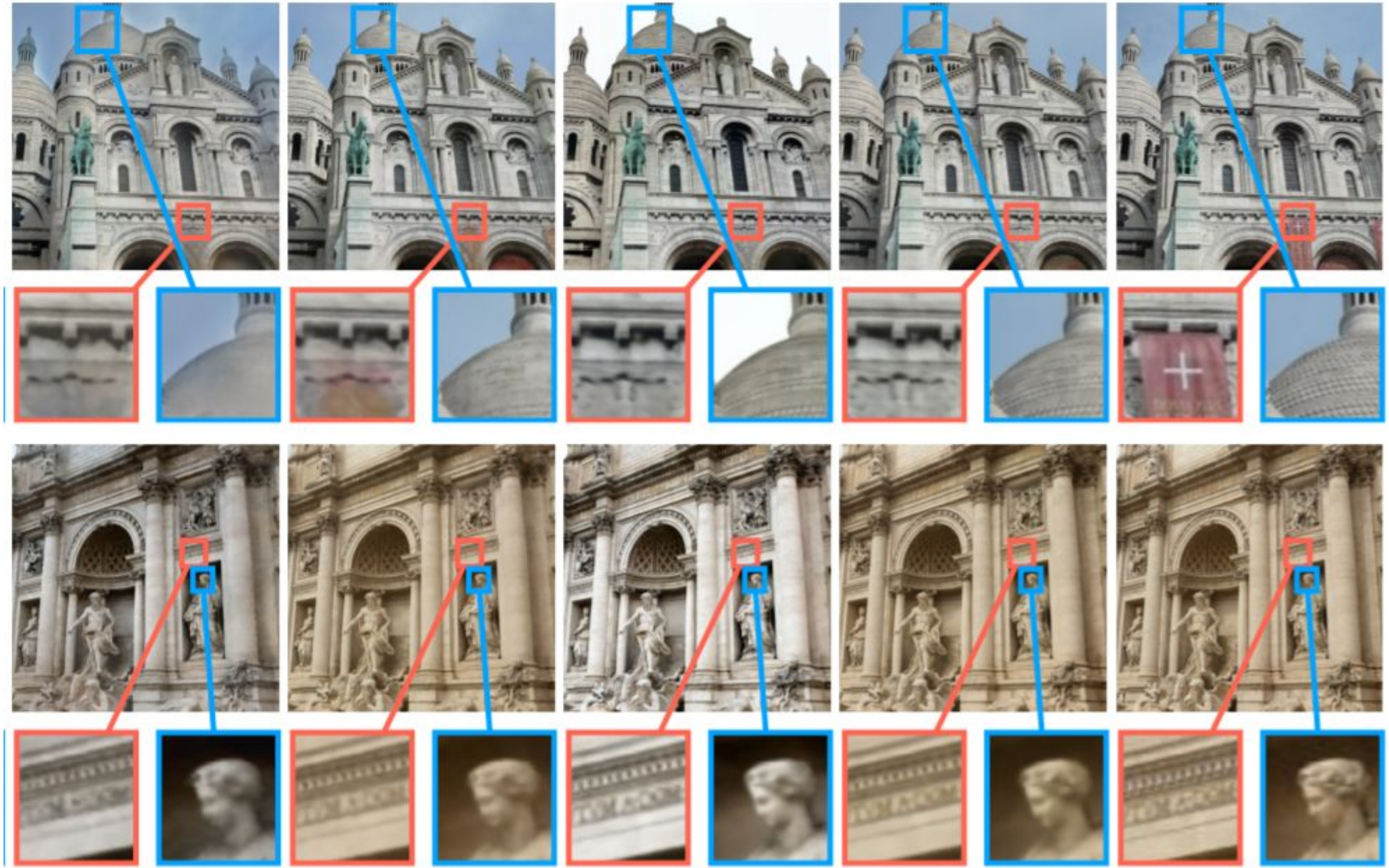
(only for fine model with extensions)

Models for experiments

- NeRF, NeRF-W
- NeRF-A (appearance), wherein the ‘transient’ head is eliminated
- NeRF-U (uncertainty), wherein appearance embedding is eliminated

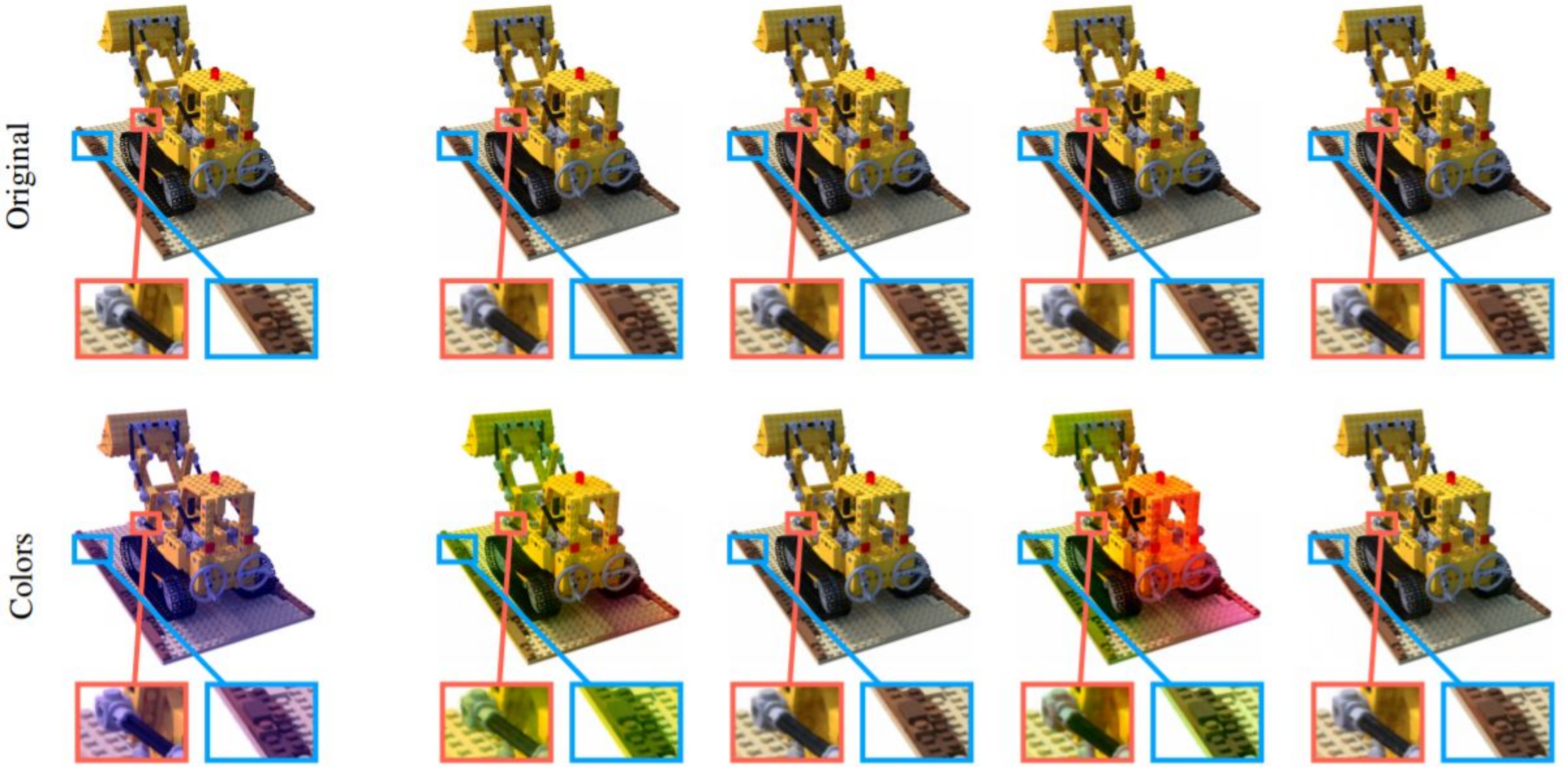
Datasets

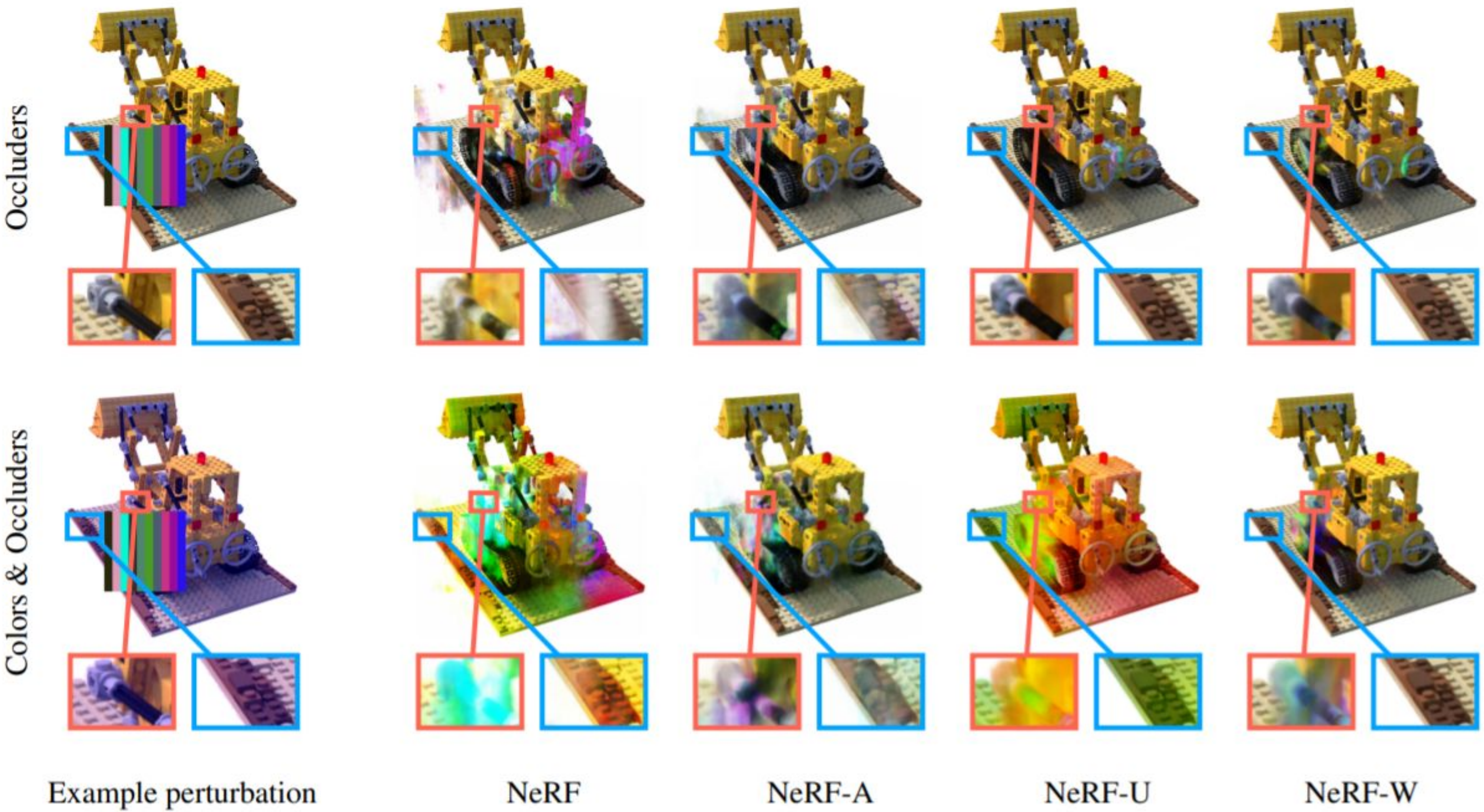
- Phototourism Dataset
- Lego Dataset



NeRF NeRF-A NeRF-U NeRF-W Ground-truth

Controllable Appearance:


Original
Colors


Occluders
Colors & Occluders
Example perturbation
NeRF
NeRF-A
NeRF-U
NeRF-W

Quantitative results on the Lego dataset

	Original				Color Perturbations			
	↑ PSNR	↑ MSSSIM	↓ CT	↓ LPIPS	↑ PSNR	↑ MSSSIM	↓ CT	↓ LPIPS
NeRF	**33.35**±0.05	**0.989**±0.000	**0.033**±0.000	**0.019**±0.000	23.38±0.05	0.964±0.001	0.039±0.000	0.076±0.001
NeRF-A	33.04±0.06	**0.989**±0.000	**0.033**±0.000	0.020±0.000	30.66±1.38	0.983±0.007	0.038±0.006	0.031±0.015
NeRF-U	33.07±0.27	**0.989**±0.001	**0.033**±0.000	**0.019**±0.001	24.87±0.52	0.968±0.000	0.039±0.001	0.063±0.007
NeRF-W	32.89±0.14	**0.989**±0.000	**0.033**±0.000	0.020±0.001	**31.51**±0.28	**0.987**±0.001	**0.034**±0.000	**0.022**±0.001

	Occluders				Colors Perturbations & Occluders			
	↑ PSNR	↑ MSSSIM	↓ CT	↓ LPIPS	↑ PSNR	↑ MSSSIM	↓ CT	↓ LPIPS
NeRF	19.35±0.11	0.891±0.001	0.057±0.000	0.112±0.001	15.73±3.13	0.804±0.109	0.061±0.003	0.217±0.100
NeRF-A	22.71±0.63	0.922±0.005	0.051±0.001	0.086±0.003	21.08±0.41	0.903±0.007	0.057±0.004	0.116±0.016
NeRF-U	23.47±0.50	0.944±0.004	**0.045**±0.001	**0.059**±0.004	17.65±4.10	0.846±0.130	0.053±0.007	0.183±0.117
NeRF-W	**25.03**±1.00	**0.946**±0.009	0.046±0.002	0.063±0.009	**22.19**±0.30	**0.927**±0.003	**0.050**±0.001	**0.087**±0.004

Вопросы

- Какие проблемы возникают при использовании NeRF на реальных датасетах?
- Какой лосс оптимизирует NeRF-W?
- Опишите архитектуру модели NeRF-W, какие входы и выходы