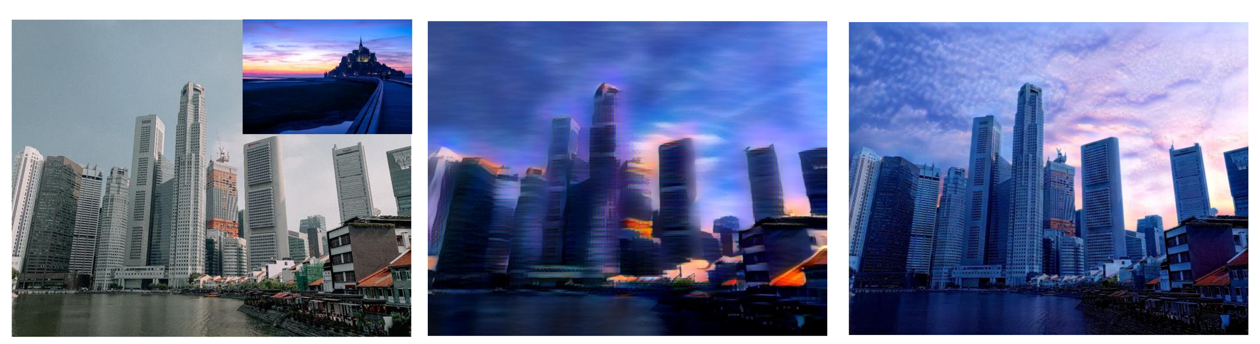
Time-of-Day Neural Style Transfer for Architectural Photographs

Yingshu Chen[1] Tuan-Anh Vu[1] Ka-Chun Shum[1] Binh-Son Hua[2] Sai-Kit Yeung[1]
[1]The Hong Kong University of Science and Technology [2]VinAI Research

香港科技大學 THE HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY
VinAI
INTERNATIONAL CONFERENCE ON COMPUTATIONAL PHOTOGRAPHY 2022
August 1-3, Caltech, Pasadena

Motivation and Problem

- Architectural photography style transfer is challenging due to its special composition of **dynamic sky and static foreground**.
- Generic neural style transfer and image-to-image translation treat the architectural image as a single entity without knowing the foreground and background, leading to the mismatched chrominance and destroyed geometric features of the original architecture.
- Task: given an architectural photo and the style reference, we **transfer styles of background and foreground separately while keeping the foreground geometry intact**.



content & style | mismatched color & destroyed geometry [1] | correct semantic style (Ours)

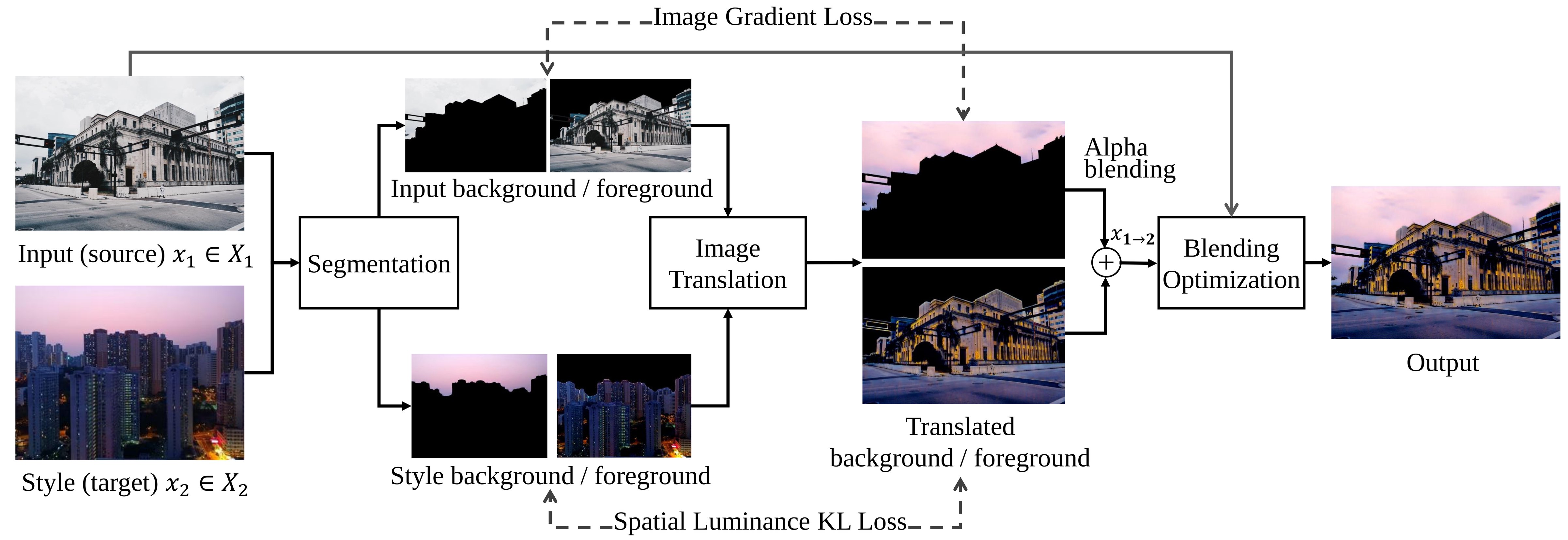
Contributions

1) A new problem setting for style transfer: **photorealistic style transfer for architectural photographs of different times of day**.

2) A two-branch image-to-image translation neural network with disentanglement representation that **separately considers style transfer for image foreground and background** respectively, accompanied with simple but effective **geometry losses** designed for image content preservation.

3) A **new dataset of architectural photographs** and an extensive benchmark for architectural style transfer.

Methodology



Architectural style transfer framework contains three main modules: segmentation, image translation and blending optimization. Segmented foreground and background images are fed into the translation network respectively. The translated and blended image $x_{1\to2}$ with input source x_1 can be further refined by blending optimization module.

High-frequency geometry losses

- Gradient loss:

$$\mathcal{L}_{gd} = \mathbb{E}_{x_1,x_2}[\|\nabla(Y(x_{1\to2}) - \nabla(Y(x_1)\|_1]$$

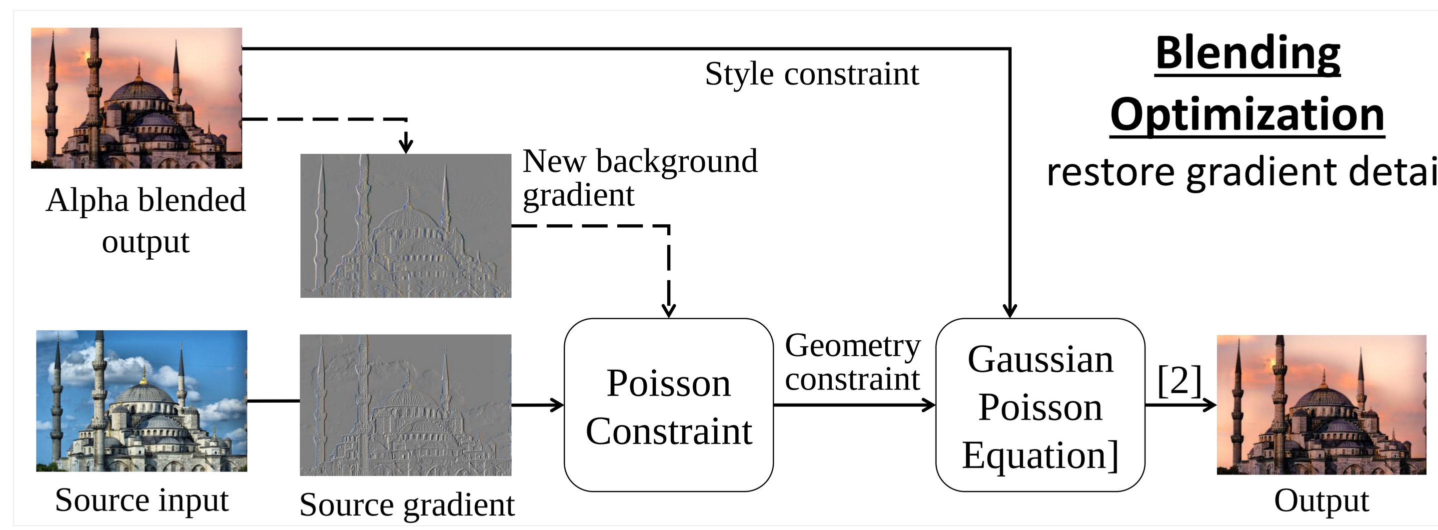
- Spatial luminance KL loss:

$$\mathcal{L}_{kl} = \mathbb{E}_{x_1,x_2}[KL(Y(x_{1\to2})\|Y(x_2))]$$

*$Y(\cdot)$ *is luminance channel.*

Blending Optimization

restore gradient detail



Experiments

Ablation Study

segmentation & blending optimization

	e-SSIM↑	Acc↑	IS↑	IoU↑
Ours-whole	0.6838	0.8282	2.5240	0.7410
Ours	0.6359	0.9486	2.7290	0.7257
Ours-opt	0.8094	0.9007	2.6127	0.7715

**whole:* w/o segmentation; *opt:* with blending optimization.

geometry losses

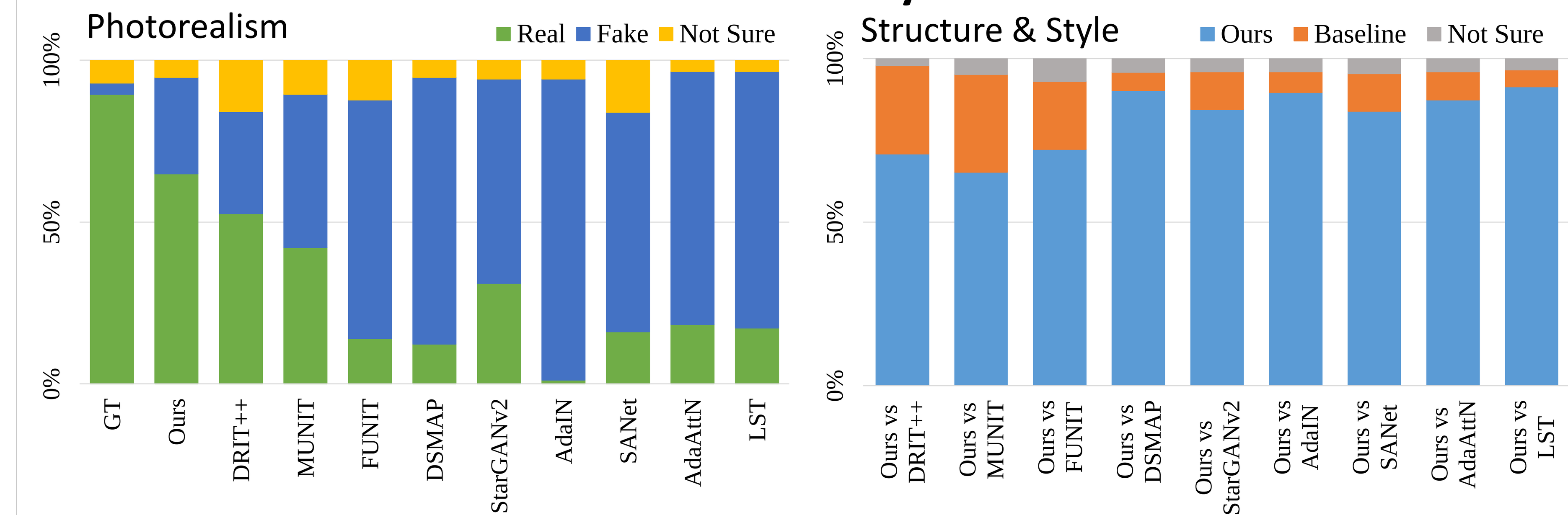
	w/o $\mathcal{L}_{kl} + \mathcal{L}_{gd}$	w/o $\mathcal{L}_{kl}$	w/o $\mathcal{L}_{gd}$	$\mathcal{L}_{total}$
e-SSIM↑	0.4800	0.5539	0.5159	**0.6359**
Acc↑	0.8934	0.9201	0.9265	**0.9486**
IS↑	2.6858	2.7183	2.7241	**2.7290**
IoU↑	0.6056	0.6536	0.6612	**0.7257**

Quantitative Evaluation

	DRIT++	MUNIT	FUNIT	DSMAP	StarGANv2	AdaIN	SANet	AdaAttN	LST	Ours
e-SSIM↑	0.5214	0.5653	0.4959	0.4790	0.4778	0.4962	0.4854	0.5194	0.4903	**0.6359**
Acc↑	0.8903	0.8678	0.77.14	0.9106	0.8788	0.7352	0.6193	0.6443	0.7071	**0.9486**
IS↑	2.6160	2.5916	2.5903	2.6580	2.6088	2.4082	2.1062	2.0928	1.7299	**2.7290**
IoU↑	0.6915	**0.7382**	0.5473	0.4975	0.4100	0.6642	0.7183	0.6532	0.6264	0.7257

Experiments

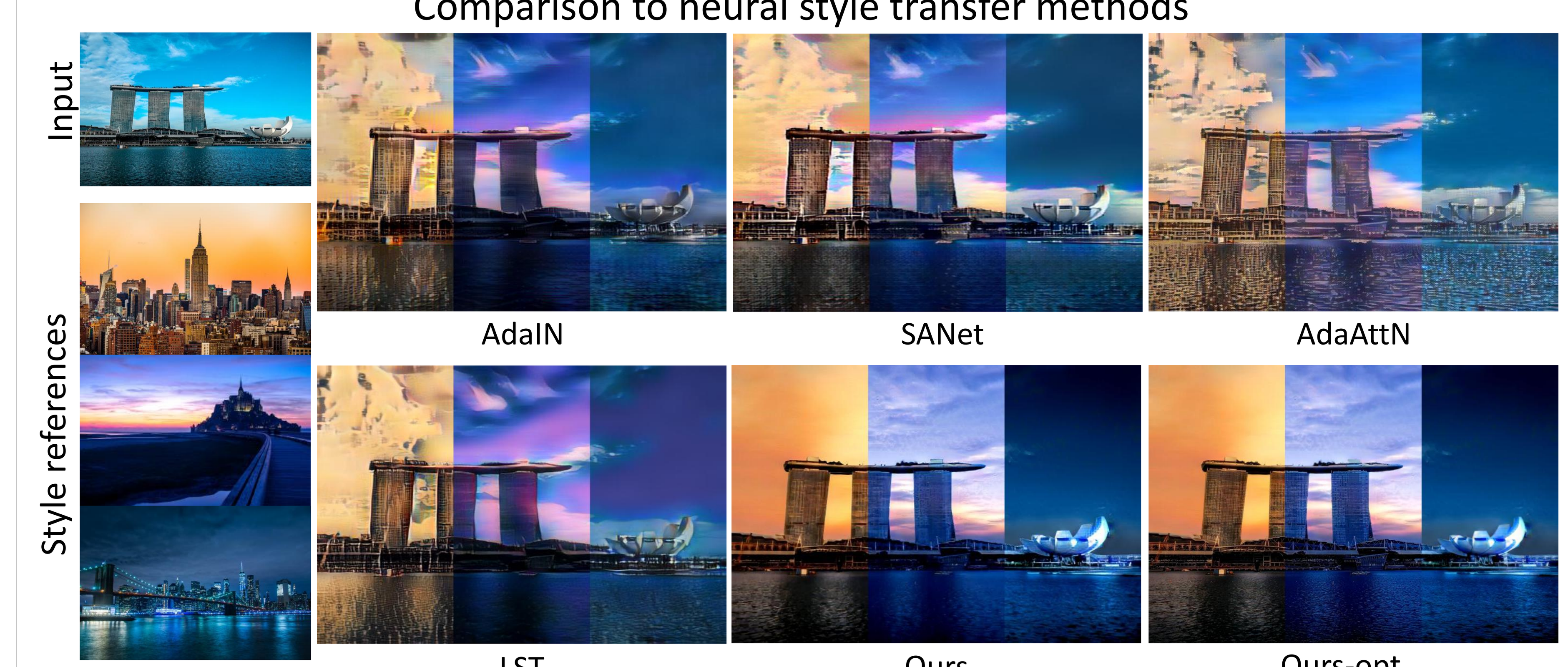
User Study



Qualitative Comparisons

Comparison to image-to-image translation methods

Comparison to neural style transfer methods



References

[1] X. Huang and S. Belongie, "Arbitrary style transfer in real-time with adaptive instance normalization", ICCV 2017.
[2] H. Wu, S. Zheng, J. Zhang, and K. Huang, "GP-GAN: Towards realistic high-resolution image blending,"ACMM 2019.