FreeCompose: Generic Zero-Shot Image Composition with Diffusion Prior

Zhekai Chen[1]*, Wen Wang[1,2]*, Zhen Yang[1], Zeqing Yuan[1],
Hao Chen[1], and Chunhua Shen[1,2]

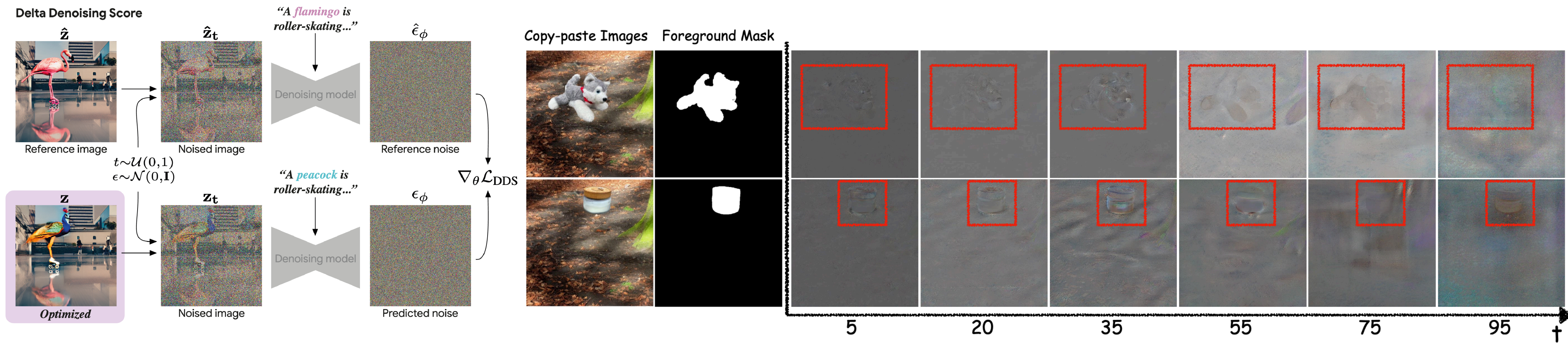[1] Zhejiang University, China   [2] Ant Group

Introduction

Previous

- Training consumption
- Requirement of data for specific
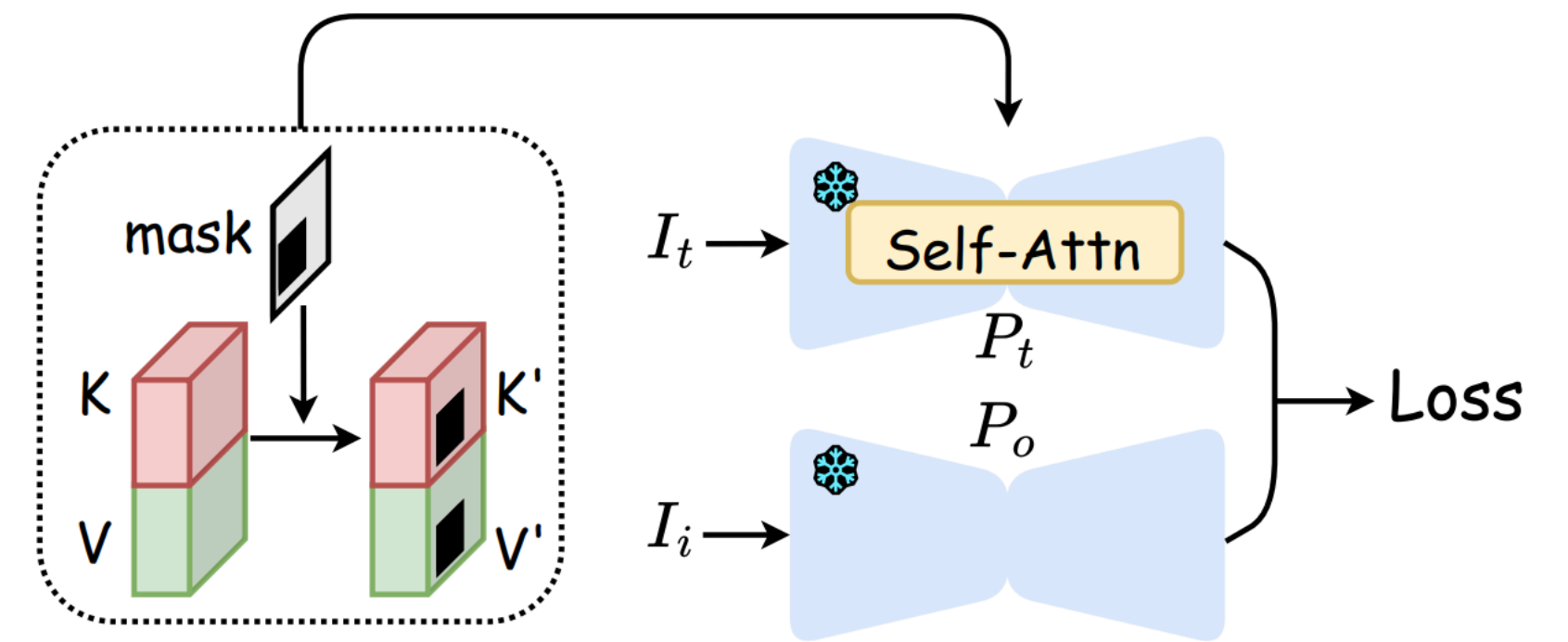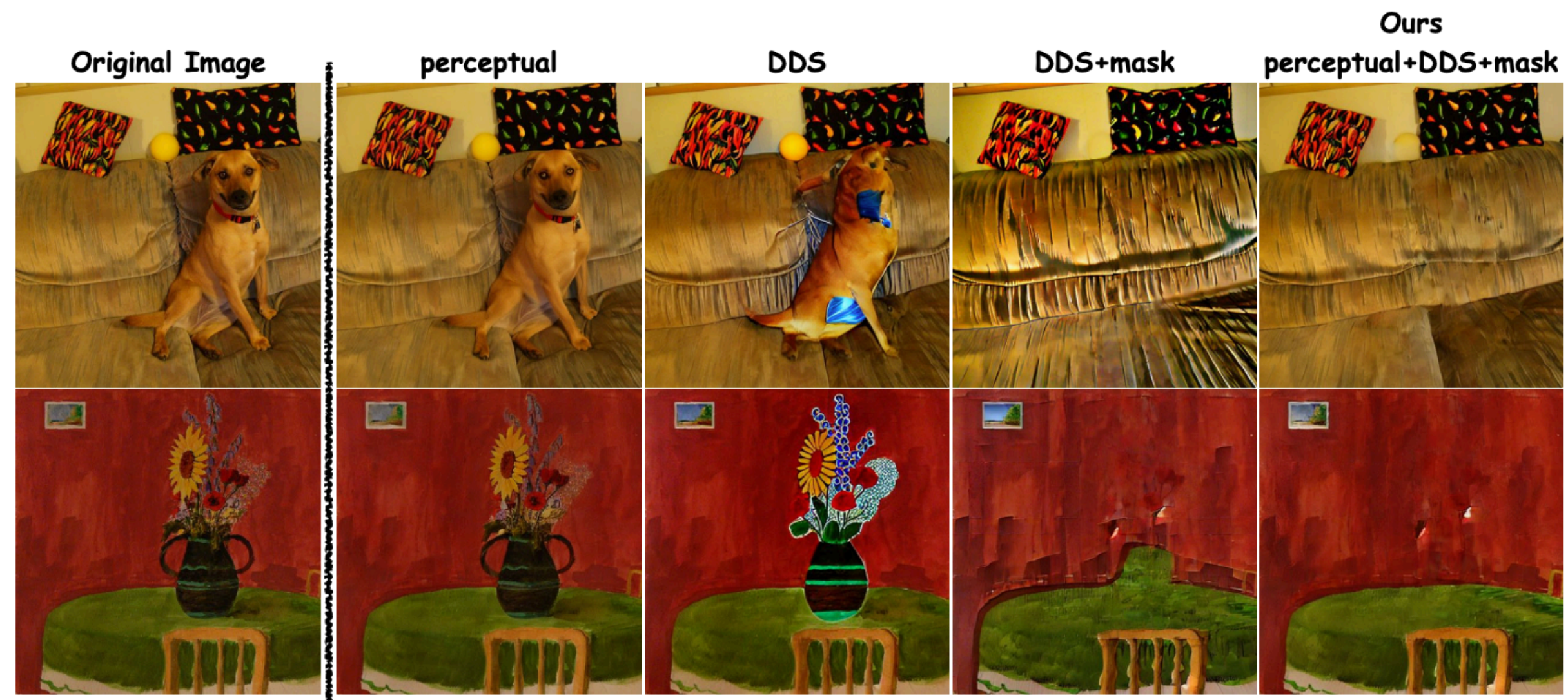
SDS/DDS loss

Harmonization

VGG-16 Perceptual loss

| Copy-paste Image | perceptual background+foreground | DDS | DDS+background | Ours DDS+perceptual |
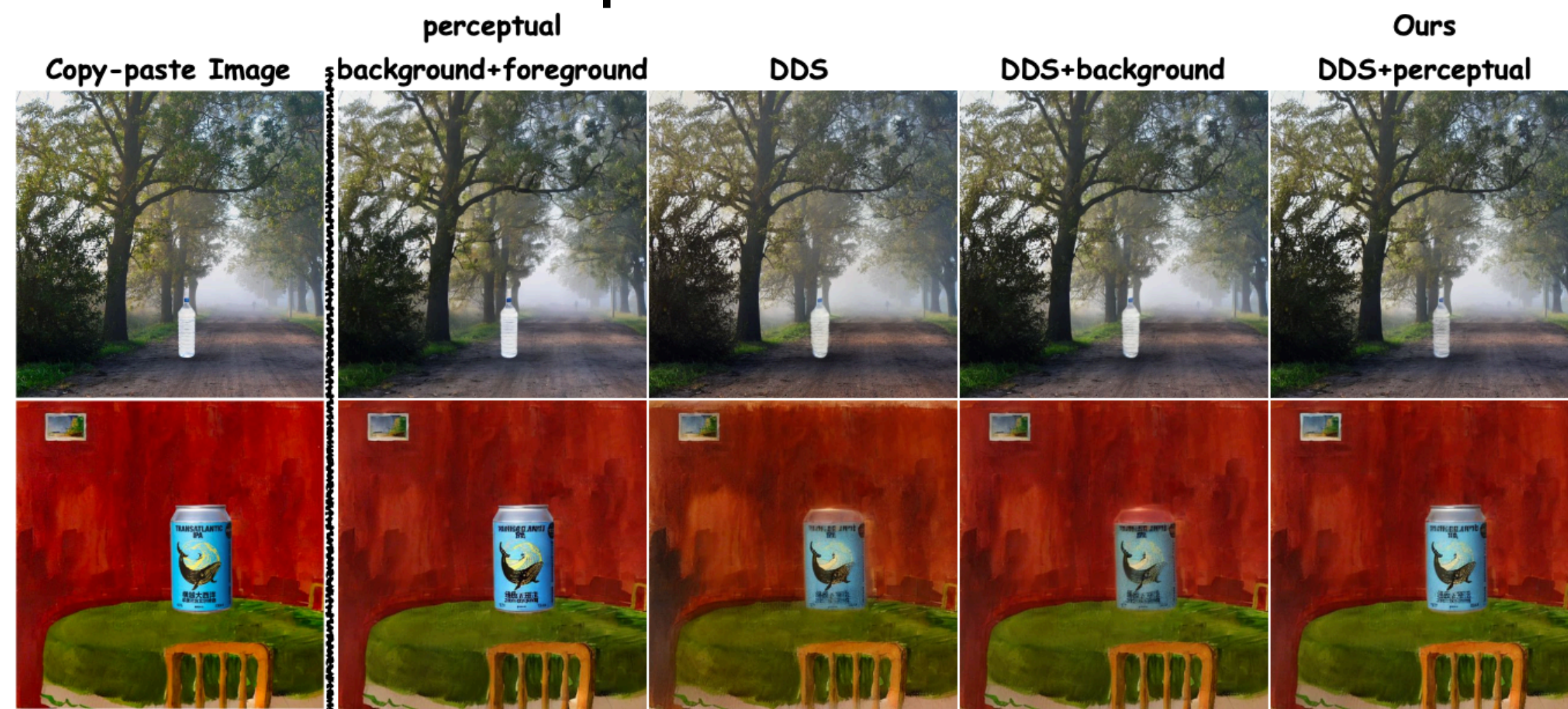
Results

Not limited to harmonization

Removal

DDS loss cannot remove object

Semantic information vs **Structure information**

Original Image	perceptual	DDS	DDS+mask	Ours perceptual+DDS+mask

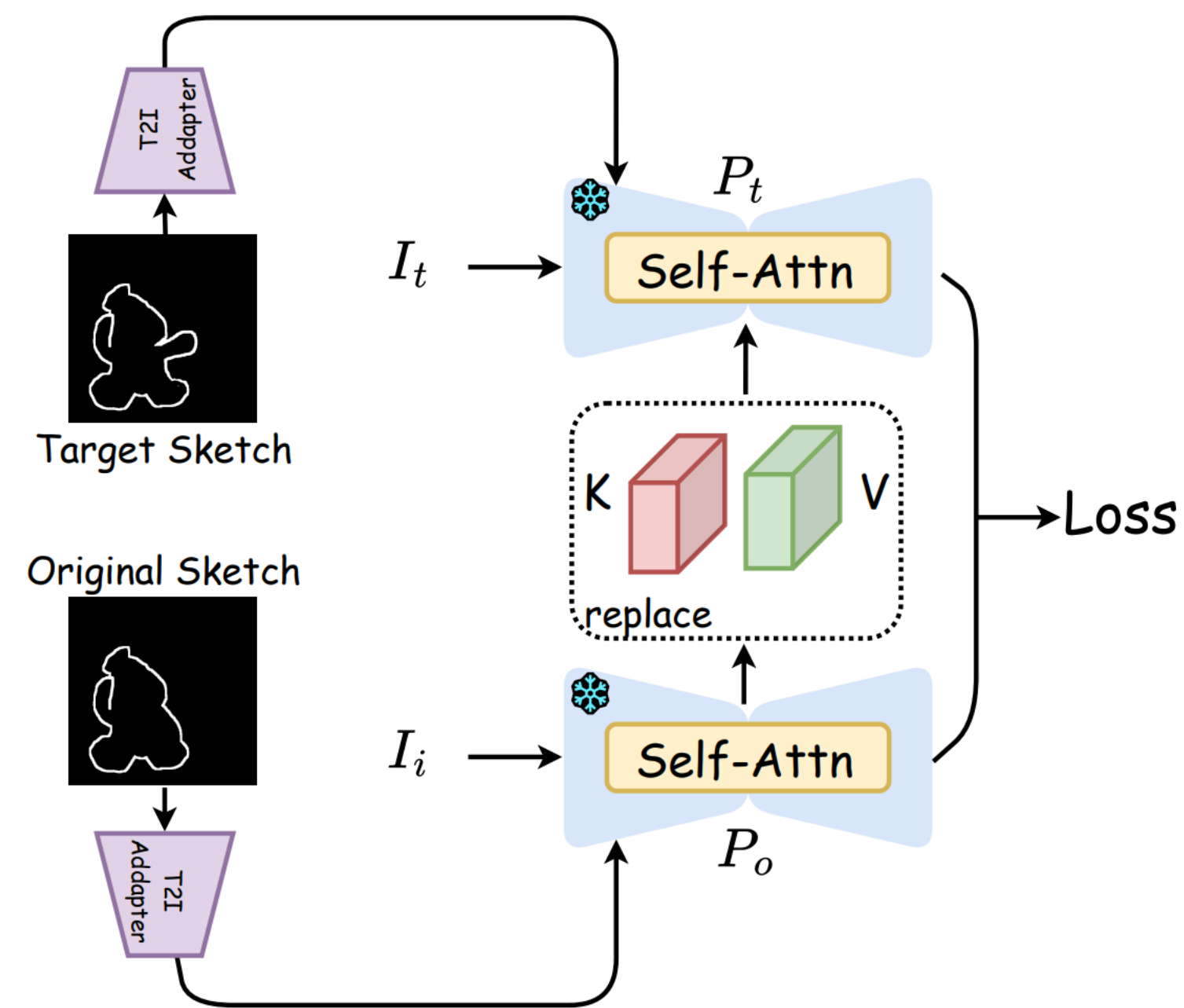(b) Loss of removal for one optimization step
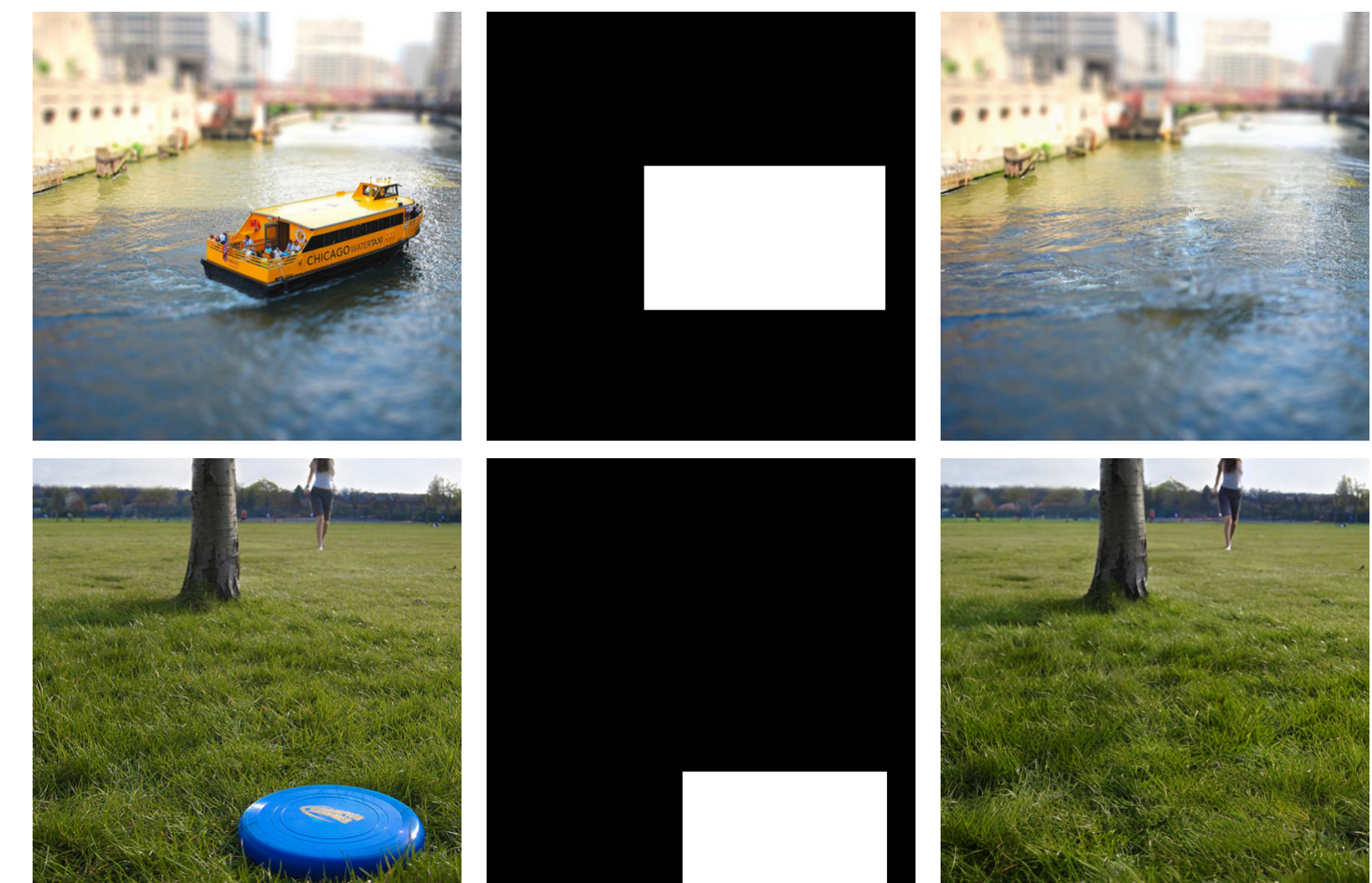
Results

Semantic Composition

Different from image harmonization

- Have regions need to be in-painted $\longrightarrow$ Mask region is **uncertain**
- Have regions need to be covered

(c) Loss of composition for one optimization step

Results

Pipeline

(a) Overview for object removal/image harmonization/semantic image composition

(b) Loss of removal for one optimization step

(c) Loss of composition for one optimization step

(a) Object stylization

(b) Zero-shot multi-character customization
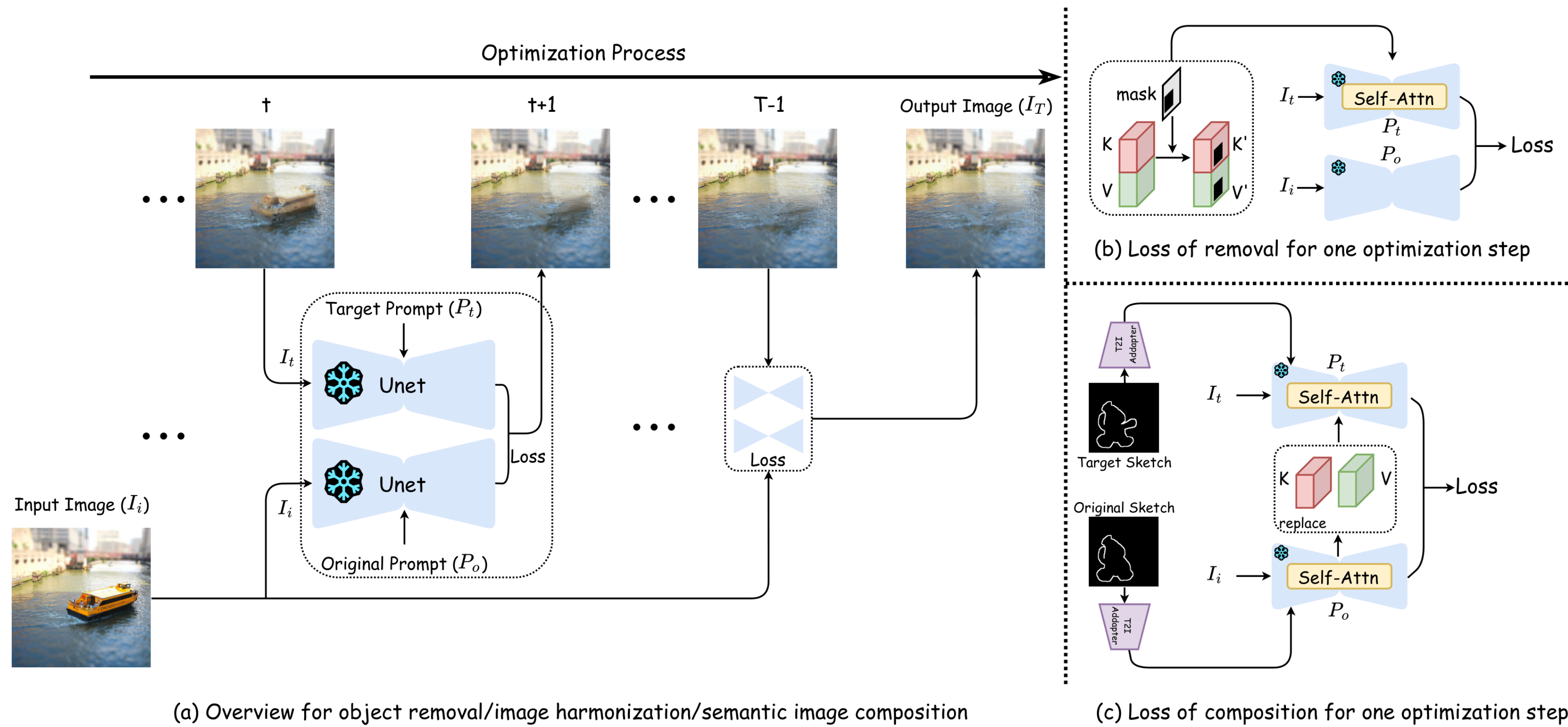
Quantitative Results

Removal

Methods	LPIPS↓	SSIM↑	MSE↓
LaMa [6]	0.0133	0.9849	37.73
SD-inpainting [4]	0.1733	0.7639	372.01
Ours	0.1120	0.7882	293.46

	Image Harmony ↑	Object Removal ↑
Repaint [LDR+22]	3.24 ± 1.23	3.82 ± 1.35
SD Inpainting [KZZ+23]	2.99 ± 1.37	3.55 ± 1.34
Lama [SLM+22]	3.47 ± 1.16	4.14 ± 0.94
FreeCompose (ours)	**3.85 ± 1.01**	**4.47 ± 0.73**

Composition

Methods	$CLIP_{fg}$ ↑	$Dino_{fg}$ ↑	$CLIP_{bg}$ ↑	$Dino_{bg}$ ↑	FID↓	QS↑
ObjectStitch [5]	72.13	67.14	79.05	87.99	34.71	35.82
ControlCom [8] (har)	**73.35**	**68.59**	<u>79.63</u>	<u>88.25</u>	<u>33.41</u>	<u>37.10</u>
ControlCom(com)	70.45	66.37	**81.65**	**89.01**	**30.05**	**42.28**
Diff-harmonization [1]	71.96	<u>68.20</u>	75.31	87.82	38.13	20.82
Ours	<u>72.52</u>	67.93	78.40	87.31	33.42	37.08

	Image Harmony↑	Object Identity Preserving↑
Diff Harmonization [HHL+24]	3.11 ± 1.04	3.83 ± 1.10
DucoNet [TLNZ23]	3.14 ± 1.17	**4.16 ± 1.04**
FreeCompose (ours)	**3.69 ± 1.07**	4.11 ± 0.92

Thank you for watching