PHOTOSWAP: Personalized Subject Swapping in Images

Jing Gu[1], Yilin Wang[2], Nanxuan Zhao[2], Tsu-Jui Fu[3], Wei Xiong[2], Qing Liu[2], Zhifei Zhang[2], He Zhang[2], Jianming Zhang[2], HyunJoon Jung[2], Xin Eric Wang[1]
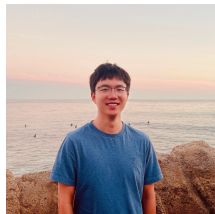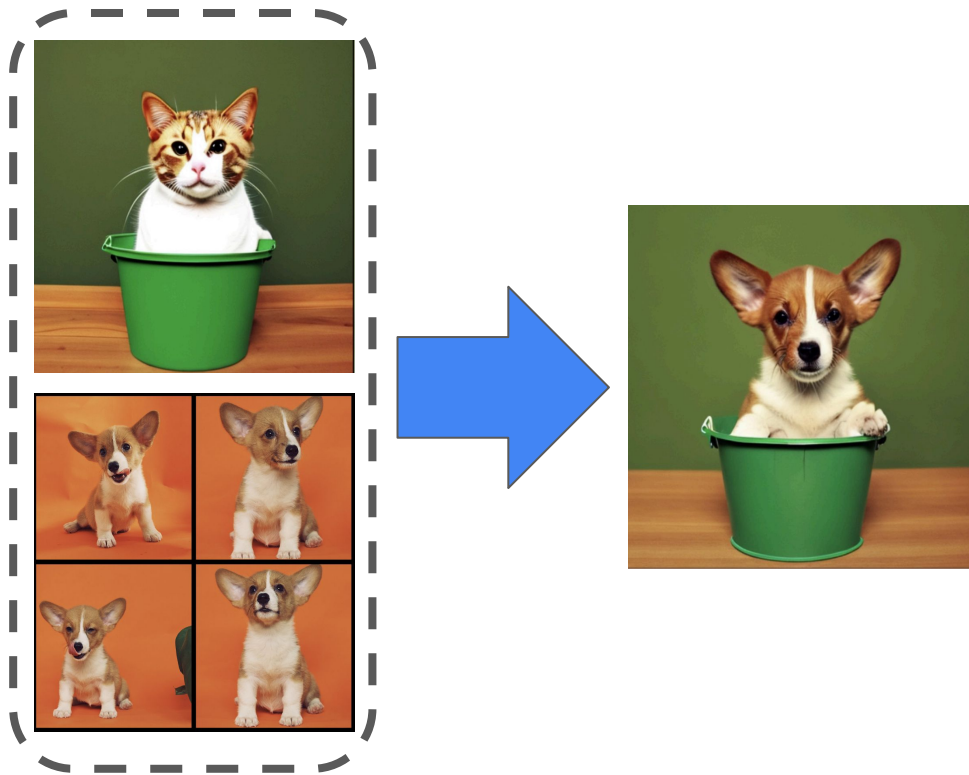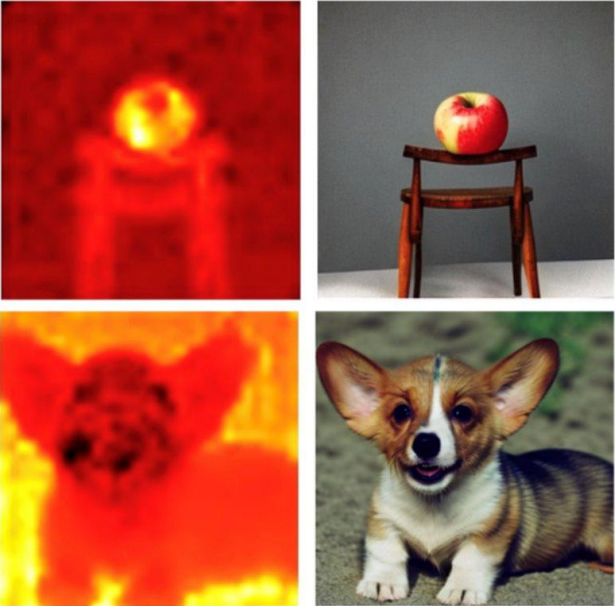
[1]University of California, Santa Cruz, [2]Adobe, [3]University of California, Santa Barbara

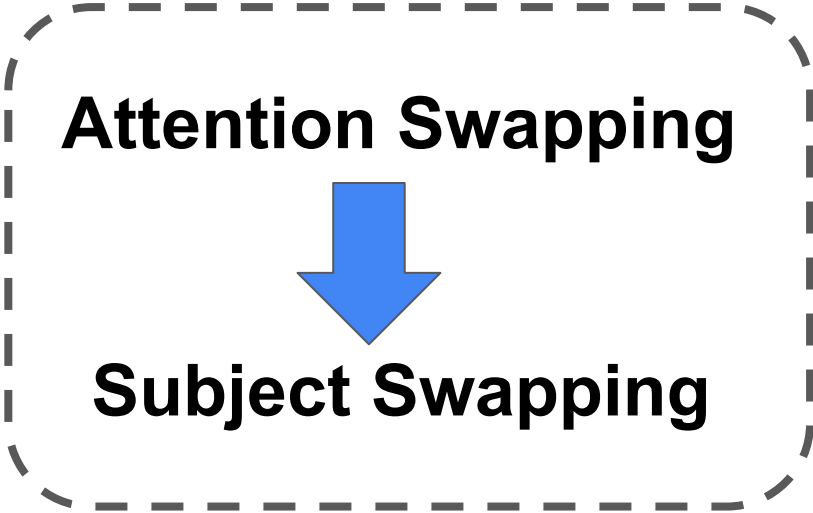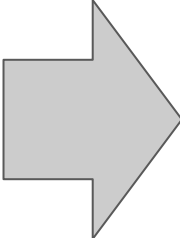➢ How to do personalized swapping?

➢ Could it be training-free?

Attention Variable in Diffusion Process correlates with the generated image
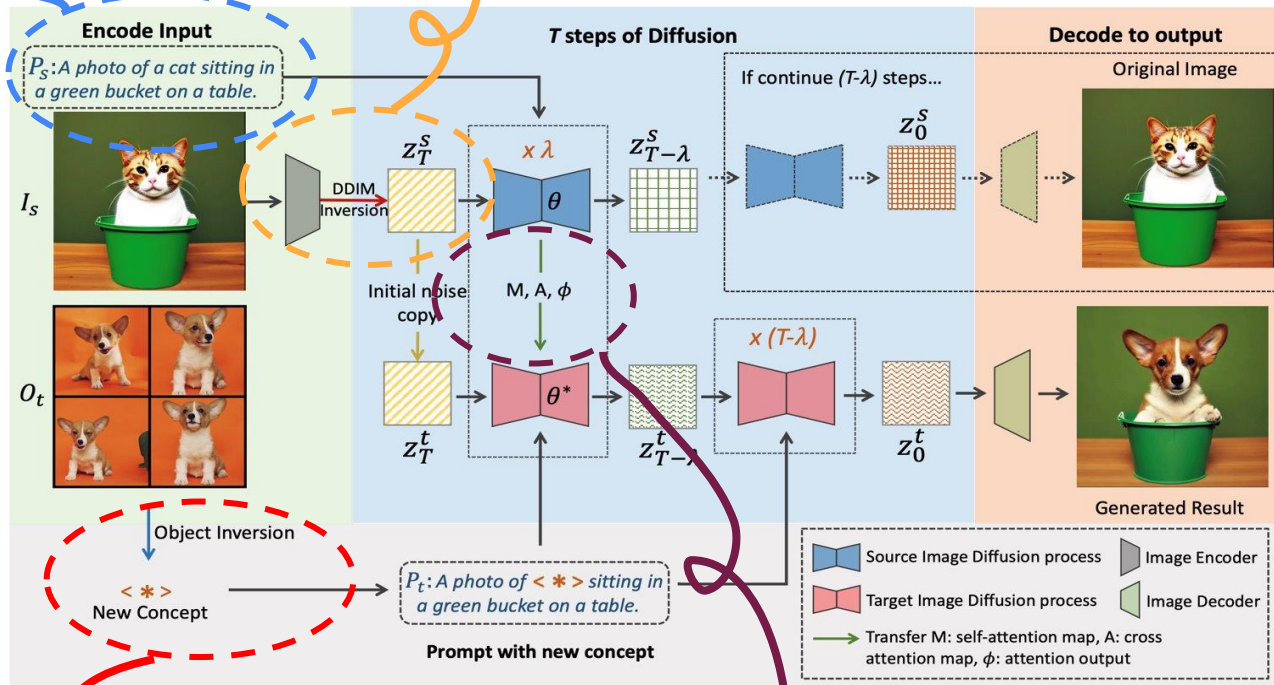

Attention Visualization   Generated Image

Attention Swapping

⬇

Subject Swapping

Photoswap: Swapping subject without training

2. Generate caption for source image

3. DDIM inversion the image into Initial noise

Encode Input

P_s: A photo of a cat sitting in a green bucket on a table.

I_s

O_t

Object Inversion

< * >
New Concept

Prompt with new concept

P_t: A photo of < * > sitting in a green bucket on a table.

T steps of Diffusion

DDIM Inversion

z_T^s

$x \lambda$

θ

$z_{T-\lambda}^s$

Initial noise copy

M, A, ϕ

z_T^t

θ^*

$z_{T-\lambda}^t$

$x (T-\lambda)$

z_0^t

If continue (T-λ) steps...

z_0^s

Decode to output

Original Image

Generated Result

Source Image Diffusion process — Image Encoder

Target Image Diffusion process — Image Decoder

Transfer M: self-attention map, A: cross attention map, ϕ: attention output
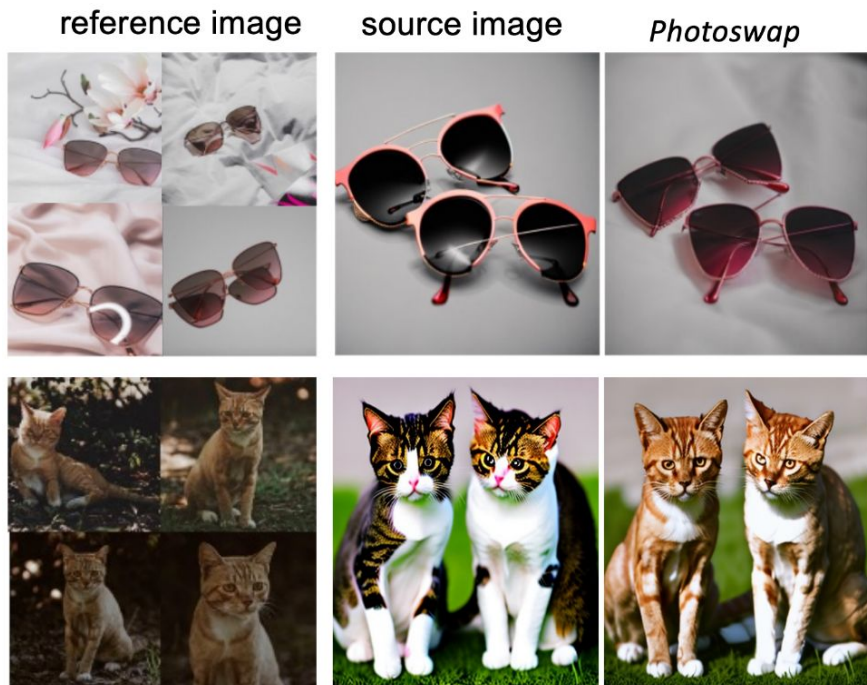
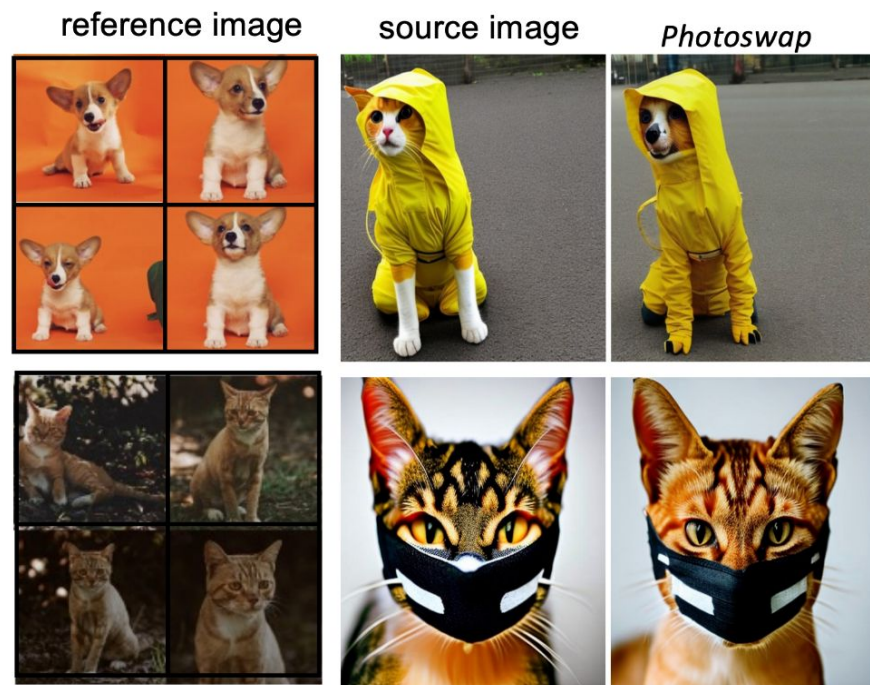**1. Convert subject into new token**

4. Swap attention variable during diffusion process

Swapping Results on different subjects
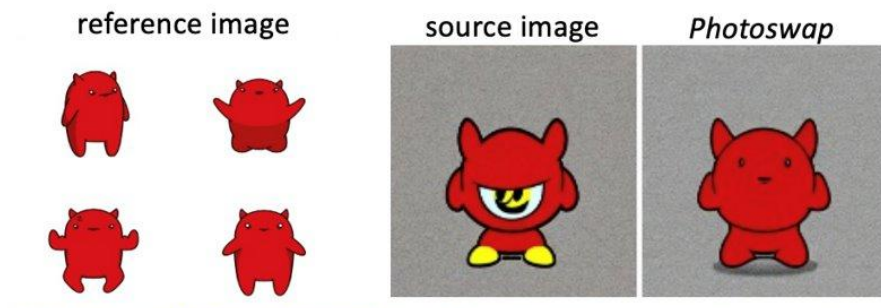
Results on Complex Images

(a) Multi-subject swap.

(b) Occluded subject swap.

https://photoswap.github.io/

Contact: Jing Gu, jgu110@ucsc.edu

Check our poster at **Great Hall & Hall B1+B2 #304**