

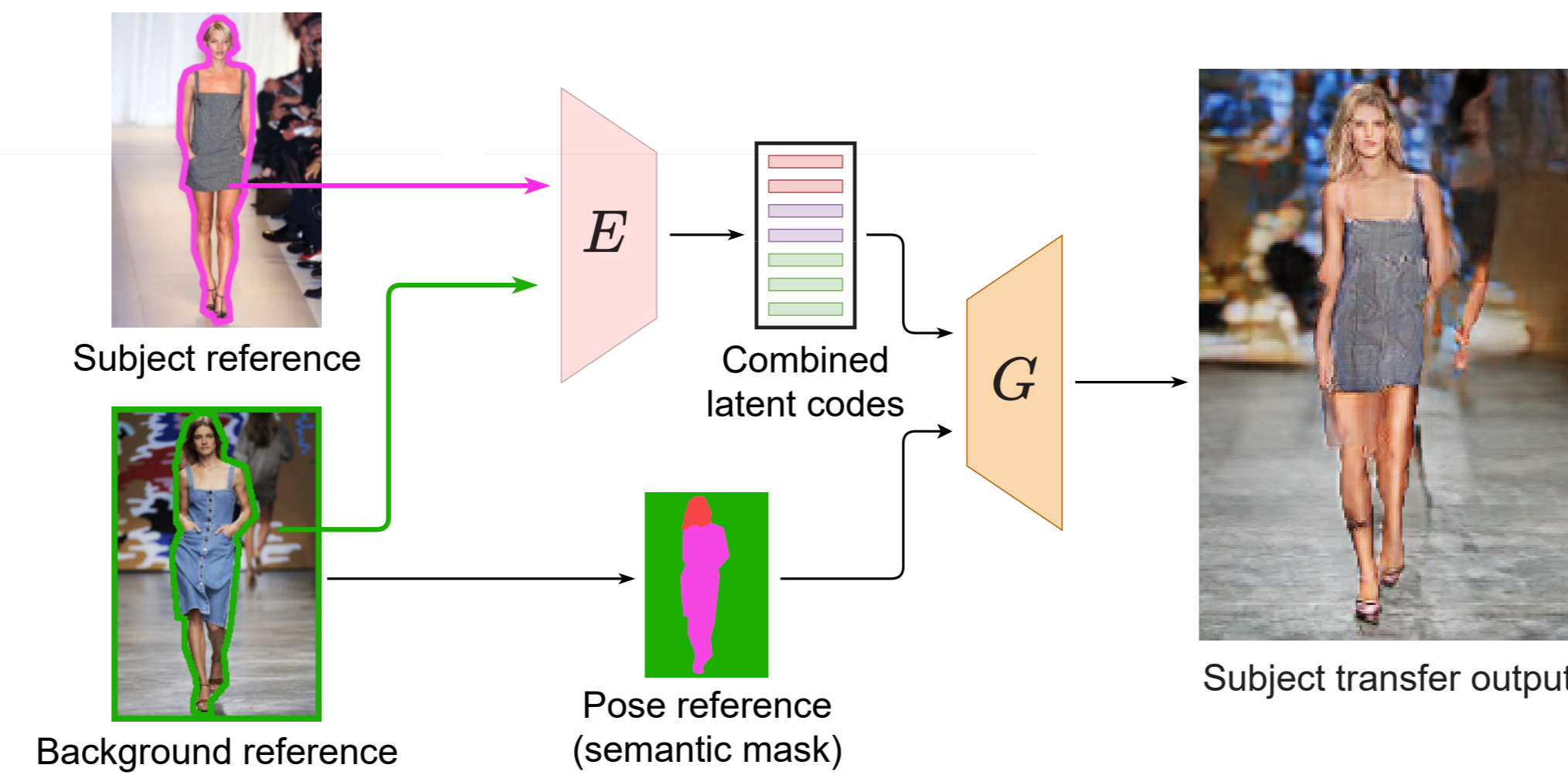
## Introduction

**Goal:** Easily *swap* subjects in an image.

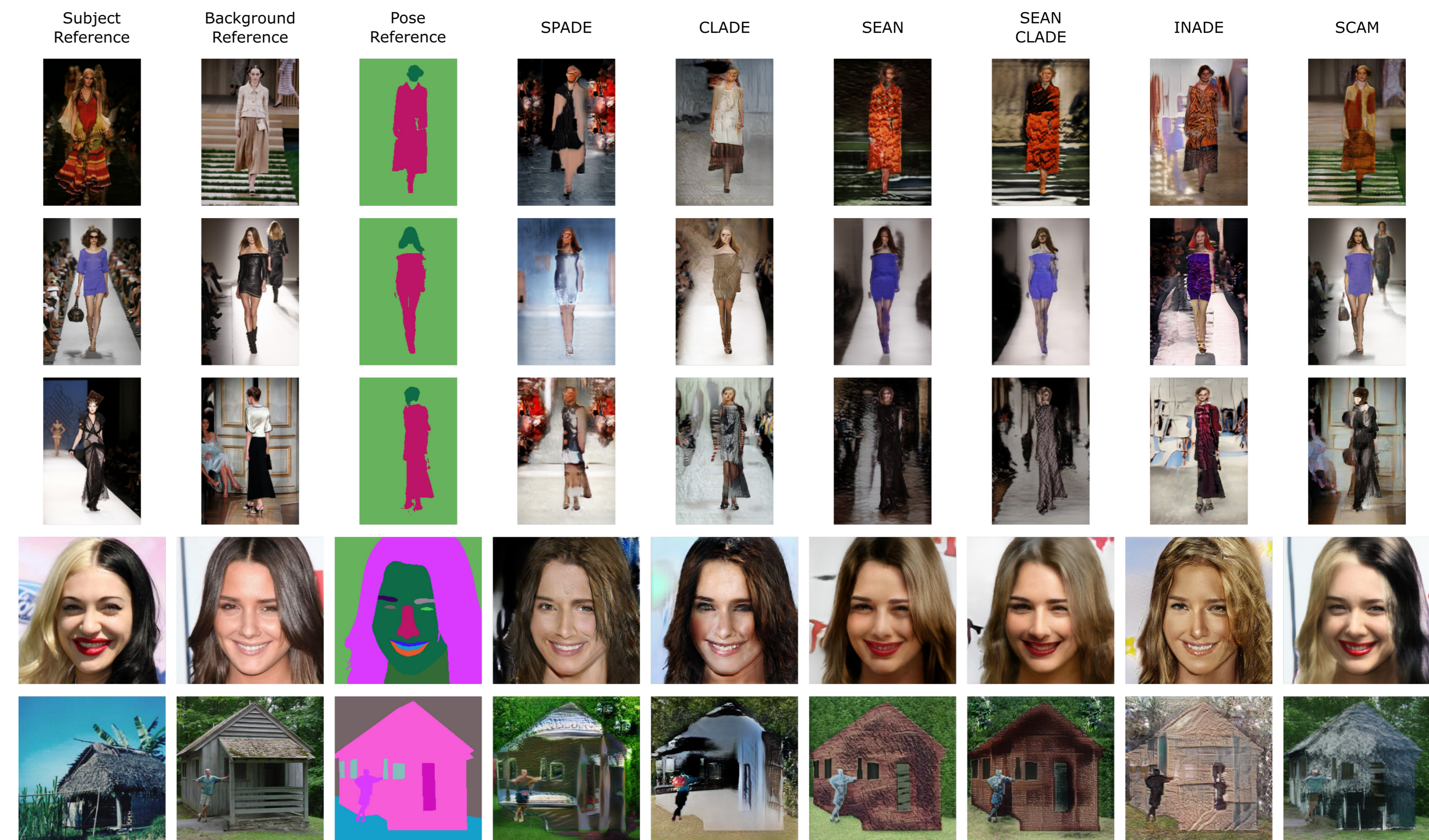
**Approach:** Separate the latent space with a *segmentation mask*.

**Prior work:** *Single style code per image [ParkCVPR19], Single style code per segmentation region [ZhuCVPR20].*

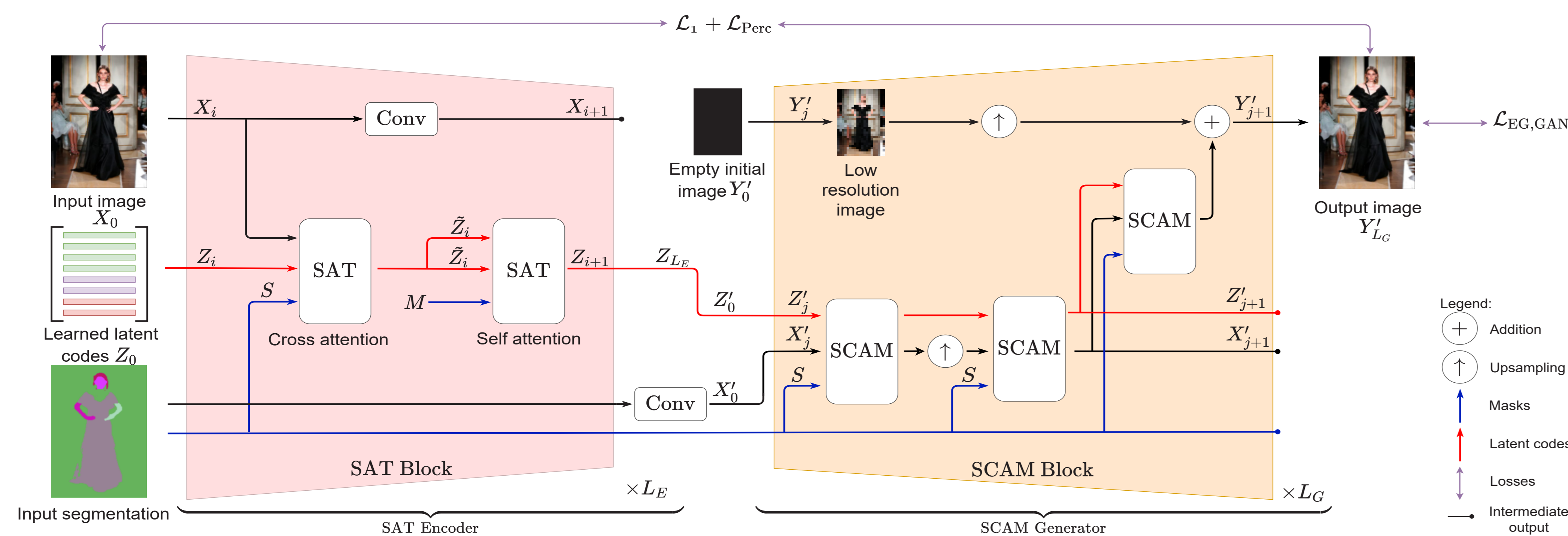
**Our method:** *Multiple style codes per semantic region.*



## Subject Transfer Results



## Method



Training setup of the proposed SCAM architecture.

### Semantic Cross Attention (SCA)

**SCA:** semantically constrained attention on pixels and latents.

- Pixel and group of latents have an *assigned* semantic label.
- Pixels or latents can only attend same semantic label latents.

$$SCA(I_1, I_2, I_3) = \sigma \left( \frac{QK^T \odot I_3 + \tau(1 - I_3)}{\sqrt{d_m}} \right) V$$

$$Q = W_Q I_1, K = W_K I_2 \text{ and } V = W_V I_2$$

### Semantic Cross Attention Modulation (SCAM)

**SCAM:** Multiple latents per semantic region.

- More representative power for *coarse* semantic regions.
- Discover *unsupervised semantic structures* inside the labels.

### Training setup

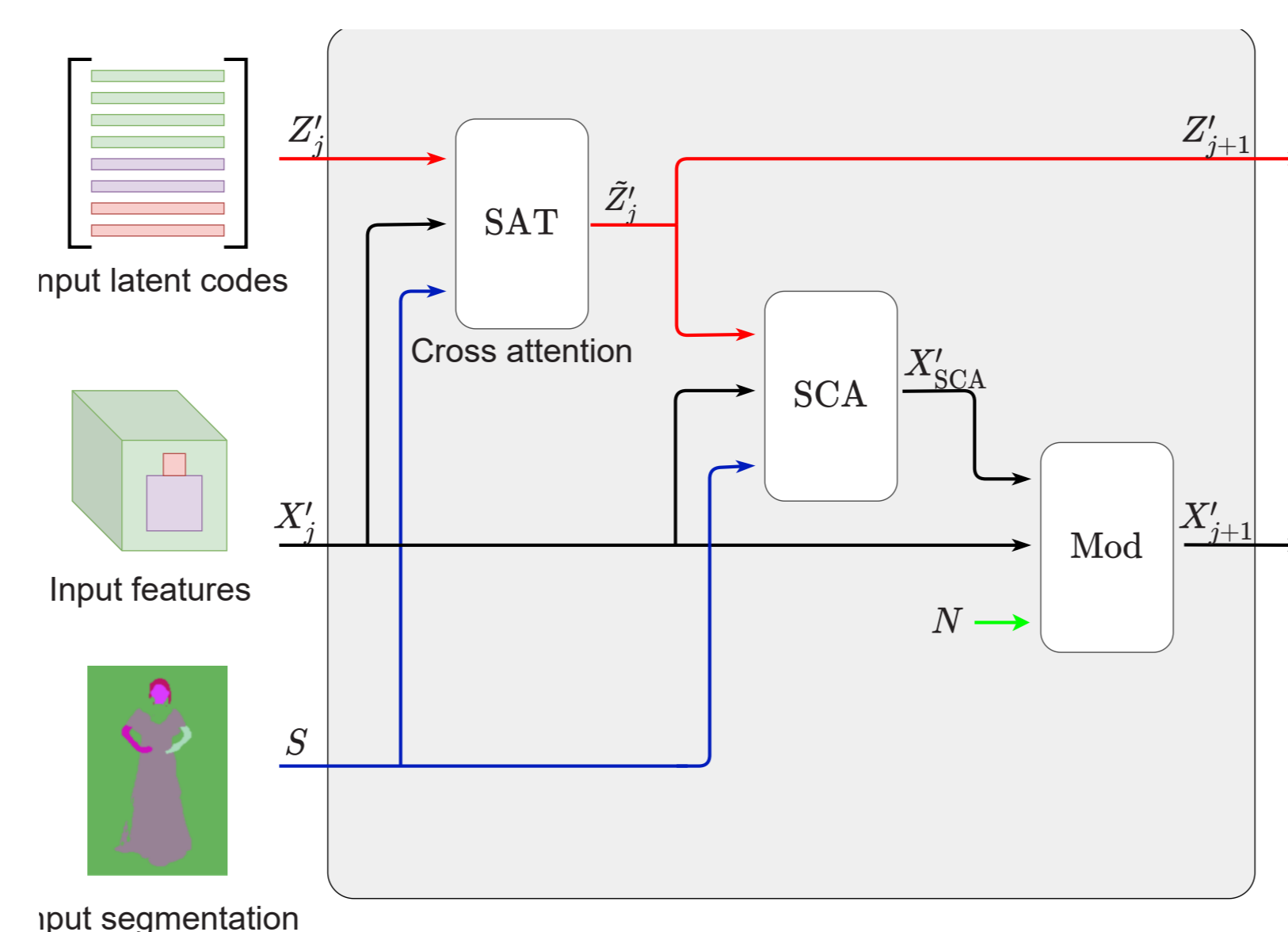
◇ Hinge GAN loss + L1 loss + Perceptual Loss.

◇ PatchGAN discriminator.

### Semantic Attention Transformer (SAT)

**SAT Operation:** *transformer-like* operation with SCA.

SAT encode/decode data in a *semantically constrained* fashion.



SCAM-Operation.

## Unsupervised Semantic Knowledge



## Quantitative Results

Method	iDesigner						CelebAMask-HQ			ADE20K	
	PSNR ↑	R-FID ↓	S-FID ↓	REIDSim ↑	REIDAcc ↑		PSNR ↑	R-FID ↓	S-FID ↓	PSNR ↑	R-FID ↓
SPADE [ParkCVPR19]	10.4	66.7	67.5	0.67	0.26		10.9	38.2	38.3	10.7	59.7
CLADE [TanTPAMI21]	11.3	45.4	46.1	0.68	0.29		10.8	41.8	42.0	10.4	53.7
SEAN-CLADE [TanTPAMI21]	15.3	48.4	56.1	0.75	0.31		16.2	19.8	24.3	14.0	38.7
INADE [TanCVPR21]	12.0	33.0	33.9	0.72	0.34		12.24	22.7	23.4	11.3	48.6
SEAN [ZhuCVPR20]	14.9	53.5	58.7	0.74	0.30		16.2	18.9	22.8	14.6	47.6
SCAM (Ours)	<b>21.4</b>	<b>13.2</b>	<b>26.9</b>	<b>0.81</b>	<b>0.56</b>		<b>21.9</b>	<b>15.5</b>	<b>19.8</b>	<b>20.0</b>	<b>27.5</b>