

Unsupervised Person Image Generation with Semantic Parsing Transformation

Sijie Song¹, Wei Zhang², Jiaying Liu¹, and Tao Mei²

¹Institute of Computer Science and Technology, Peking University

²JD AI Research



Condition Image



Pose-Guided Image Generation



Clothing Texture Transfer

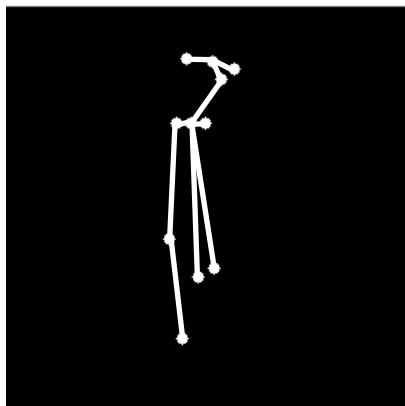
Controlled Image Manipulation

■ Pose Guided Image Generation:

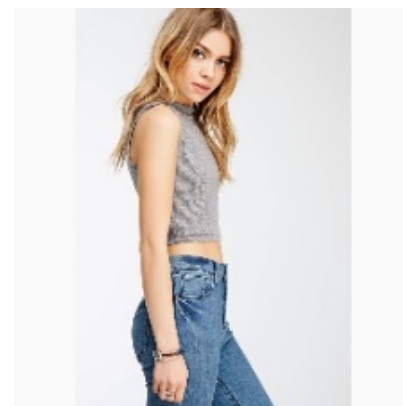


Condition Image

+



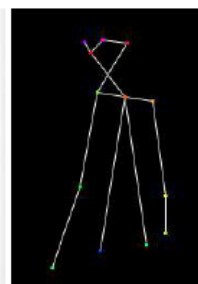
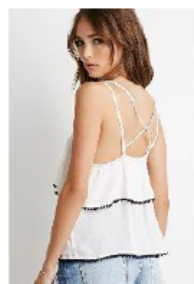
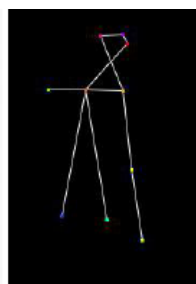
Target Pose



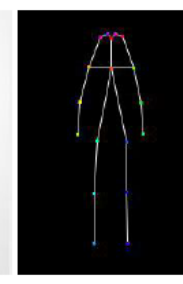
Output Image

■ Unsupervised Setting:

Training data



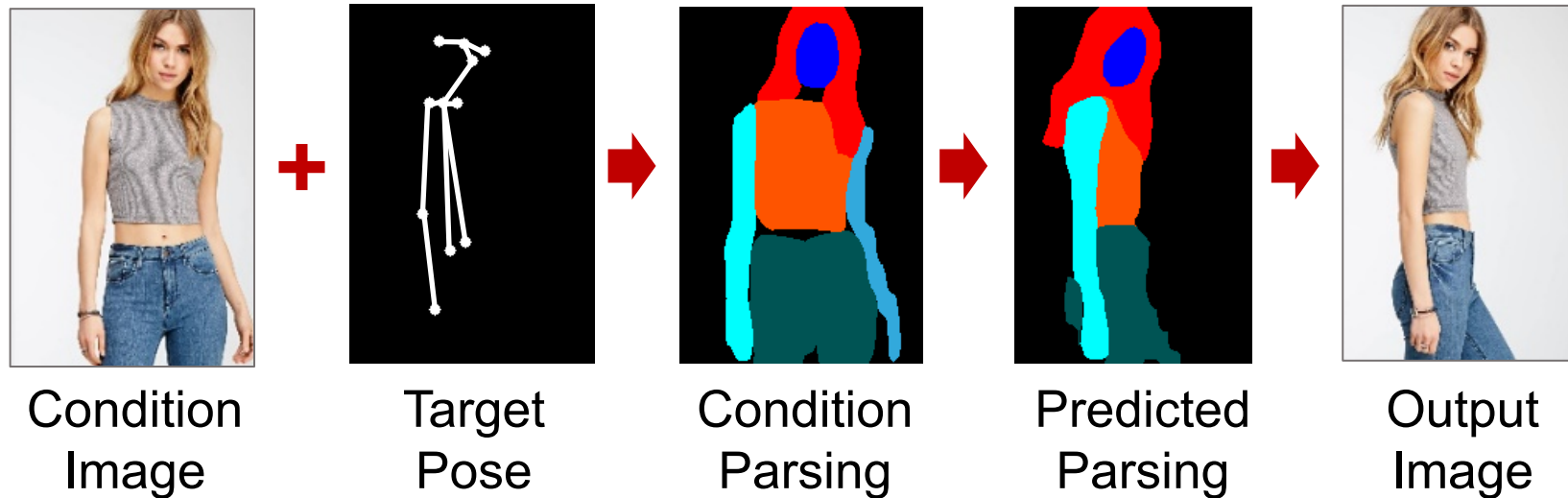
...



■ Challenges:

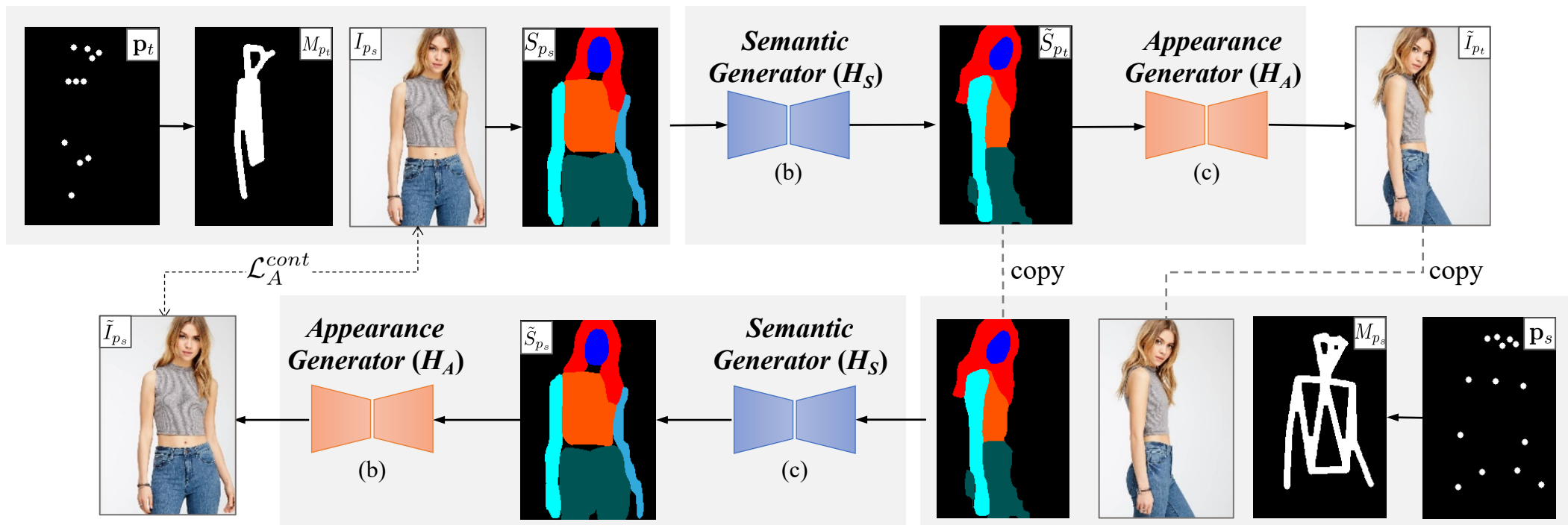
- Non-rigid deformation
- Preserving clothing attributes, *i.e.*, sleeve lengths, textures, ...
- Little clue with unpaired data

■ Solution:

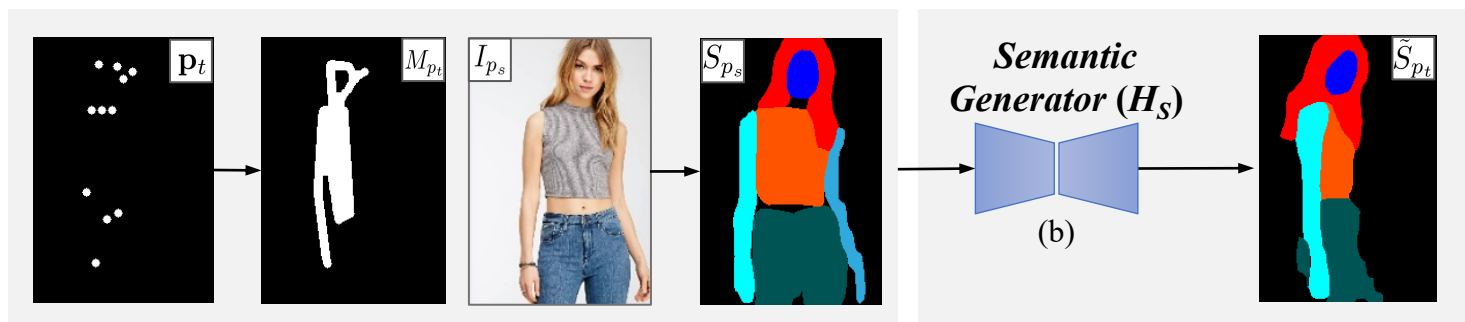


Overall Network Architecture

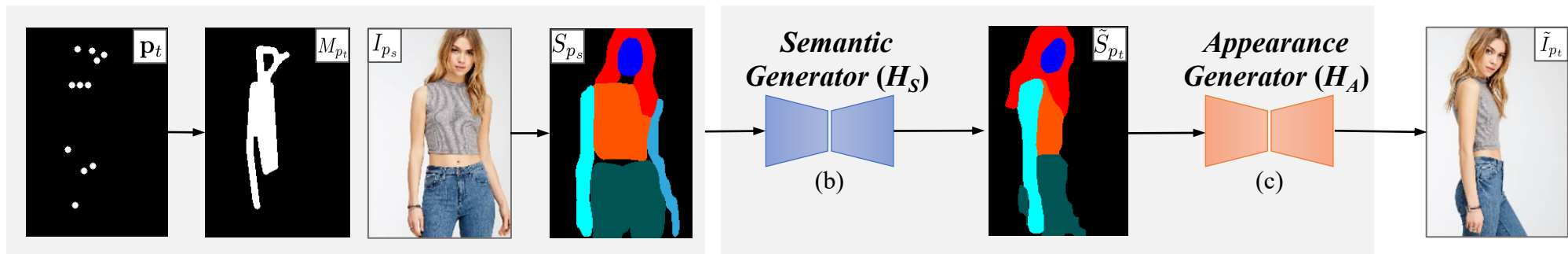
- Semantic parsing transformation
- Appearance generation
- End-to-end training



- **Overall Network Architecture**
 - **Semantic parsing transformation**
 - Appearance generation
 - End-to-end training

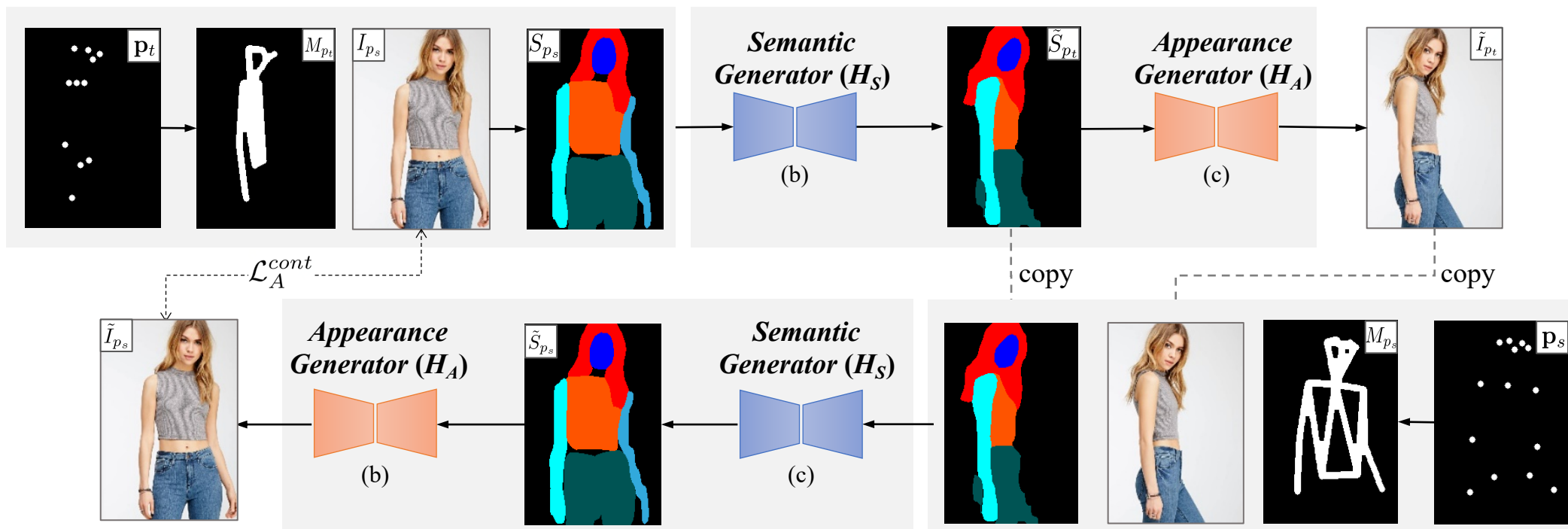


- **Overall Network Architecture**
 - Semantic parsing transformation
 - **Appearance generation**
 - End-to-end training



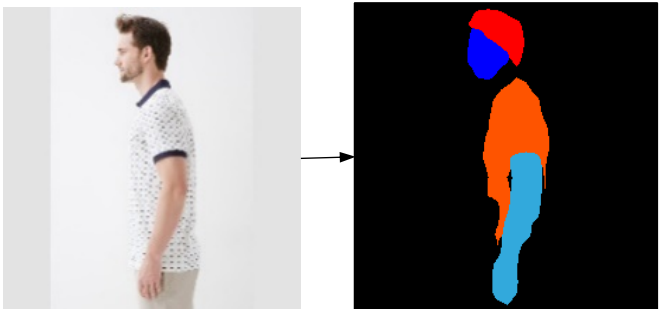
Overall Network Architecture

- Semantic parsing transformation
- Appearance generation
- End-to-end training**



■ Semantic Parsing Transformation (Pre-train)

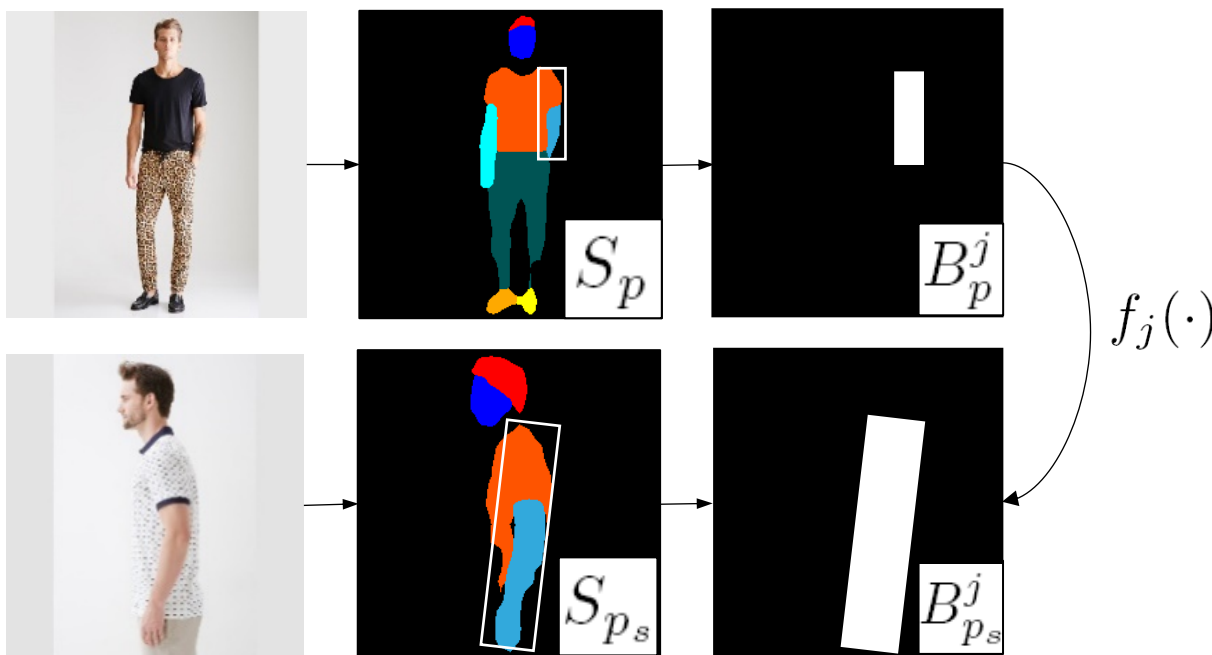
- Different clothing → Similar semantic maps.



■ Semantic Parsing Transformation (Pre-train)

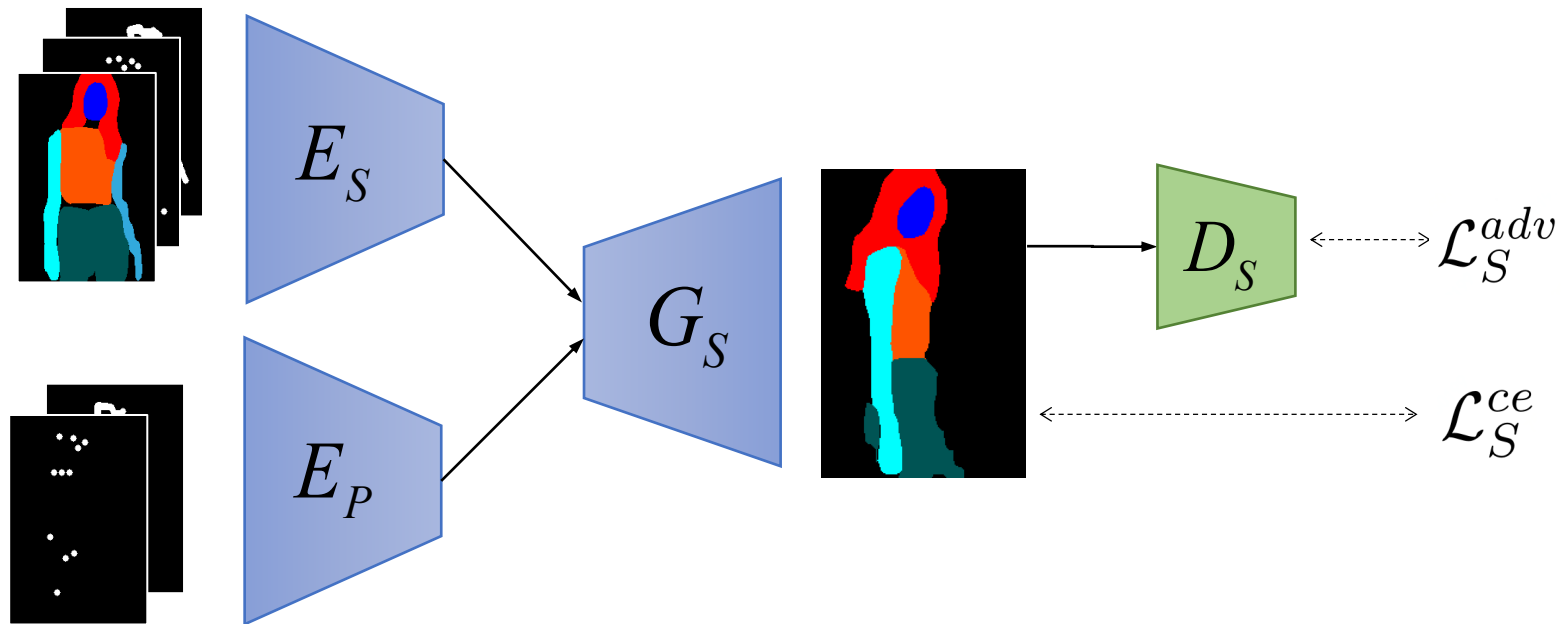
- Different clothing \rightarrow Similar semantic maps
- Pseudo-label generation

$$S_{p_t}^* = \arg \min_{S_p} \sum_{j=1}^{10} \|B_p^j \otimes S_p - f_j(B_{p_s}^j \otimes S_{p_s})\|_2^2$$

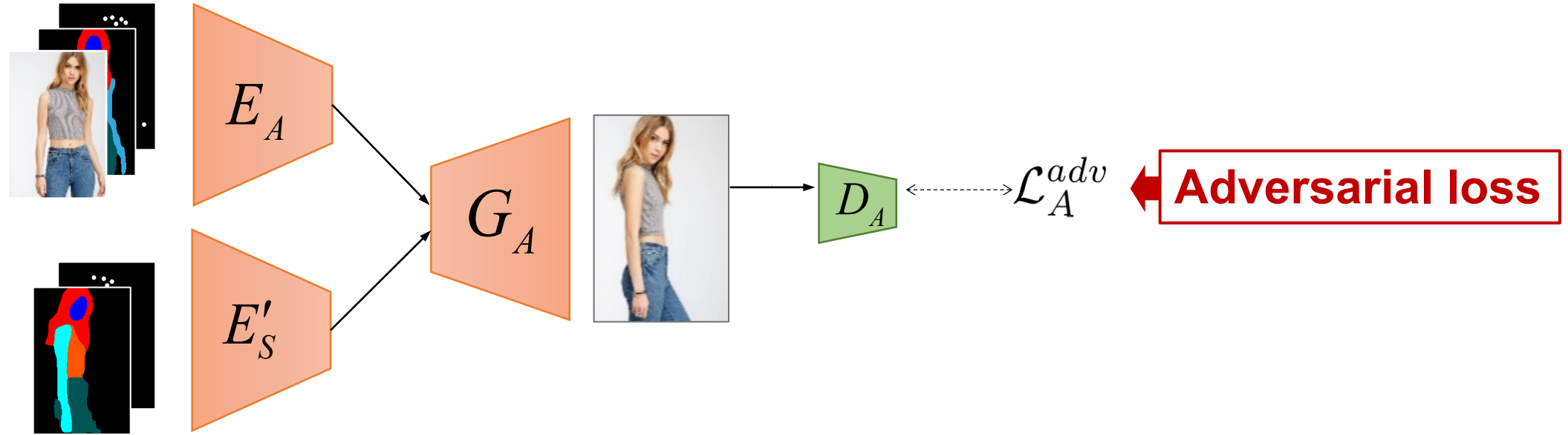


■ Semantic Parsing Transformation (Pre-train)

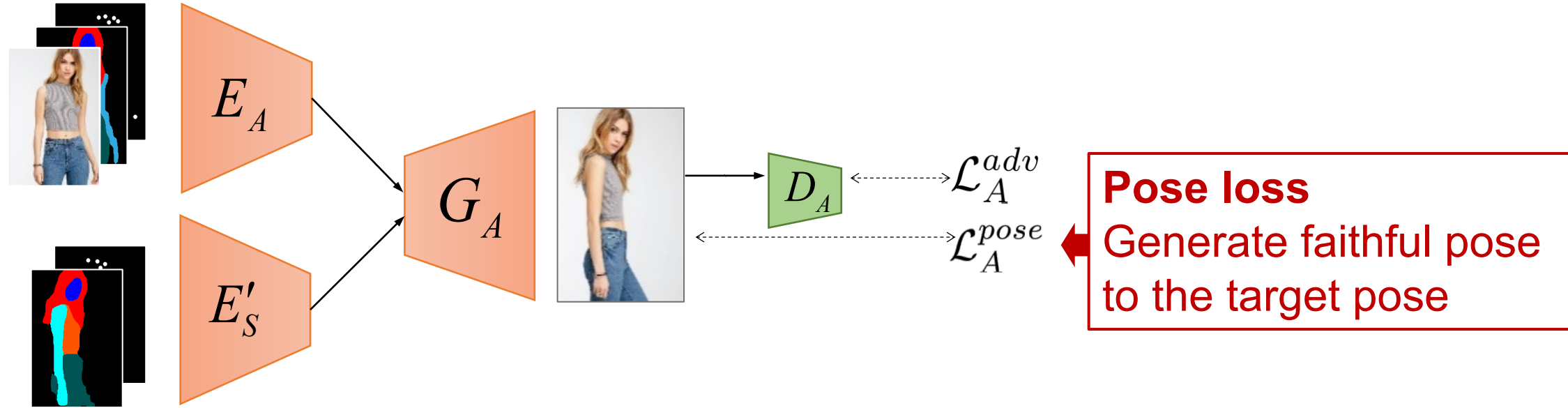
- Cross-entropy loss \mathcal{L}_S^{ce}
- Adversarial loss \mathcal{L}_S^{adv}



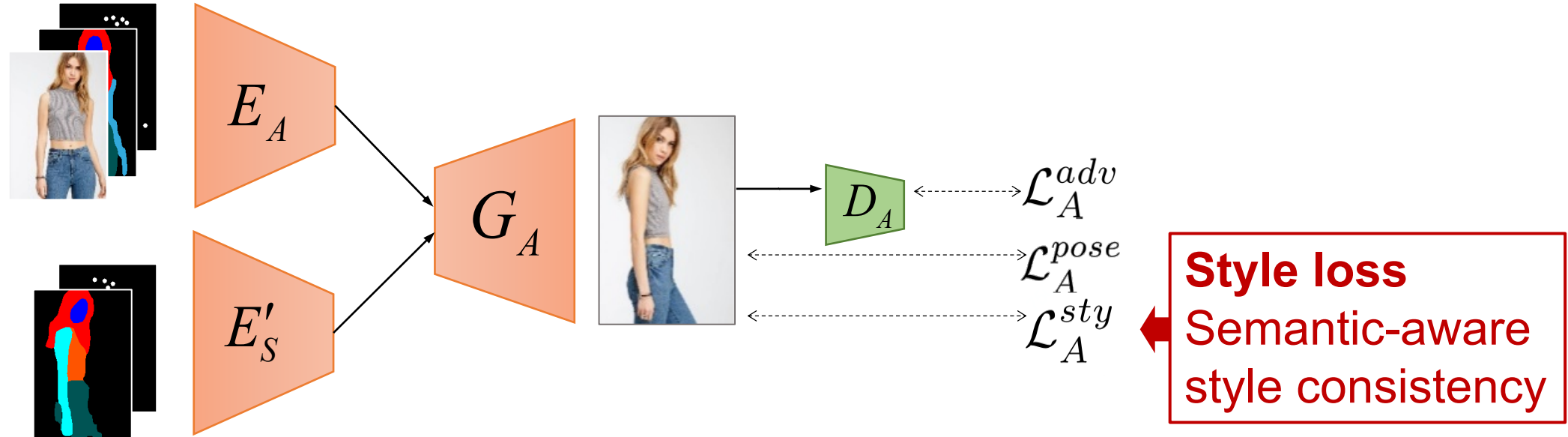
■ Appearance Generation



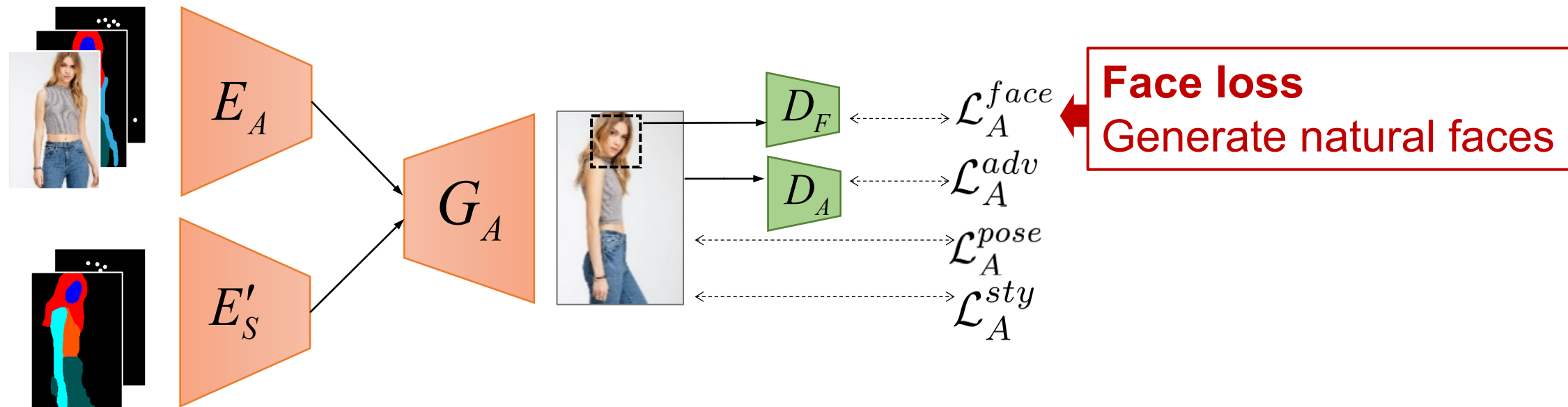
■ Appearance Generation



■ Appearance Generation



■ Appearance Generation



■ Instabilities in Two-Stage Training

- Searching error: inaccurate pseudo semantic label
- Parsing error: no semantic label to finetune human parser



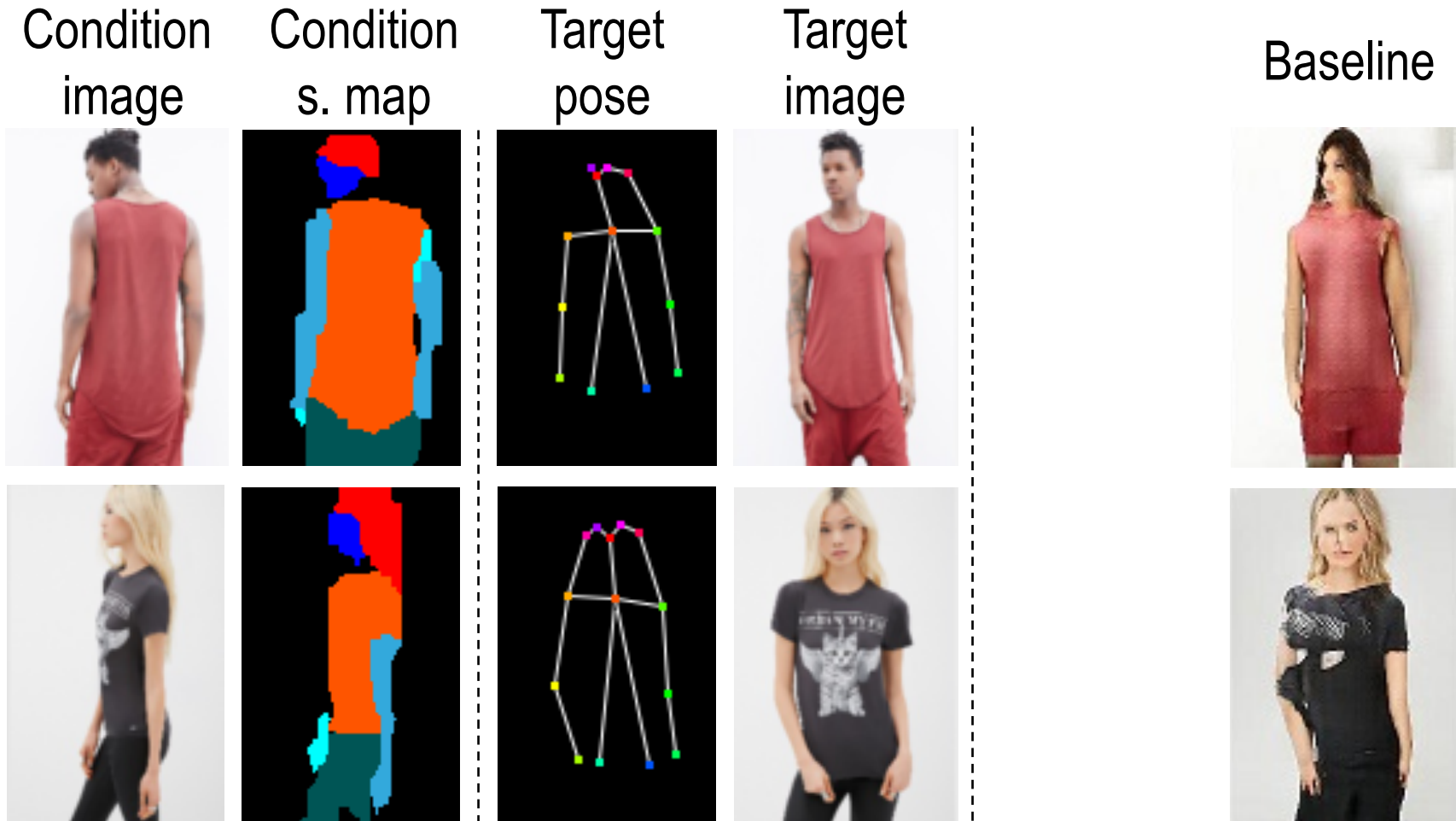
■ Instabilities in Two-Stage Training

- Searching error: inaccurate pseudo semantic label
- Parsing error: no semantic label to finetune human parser

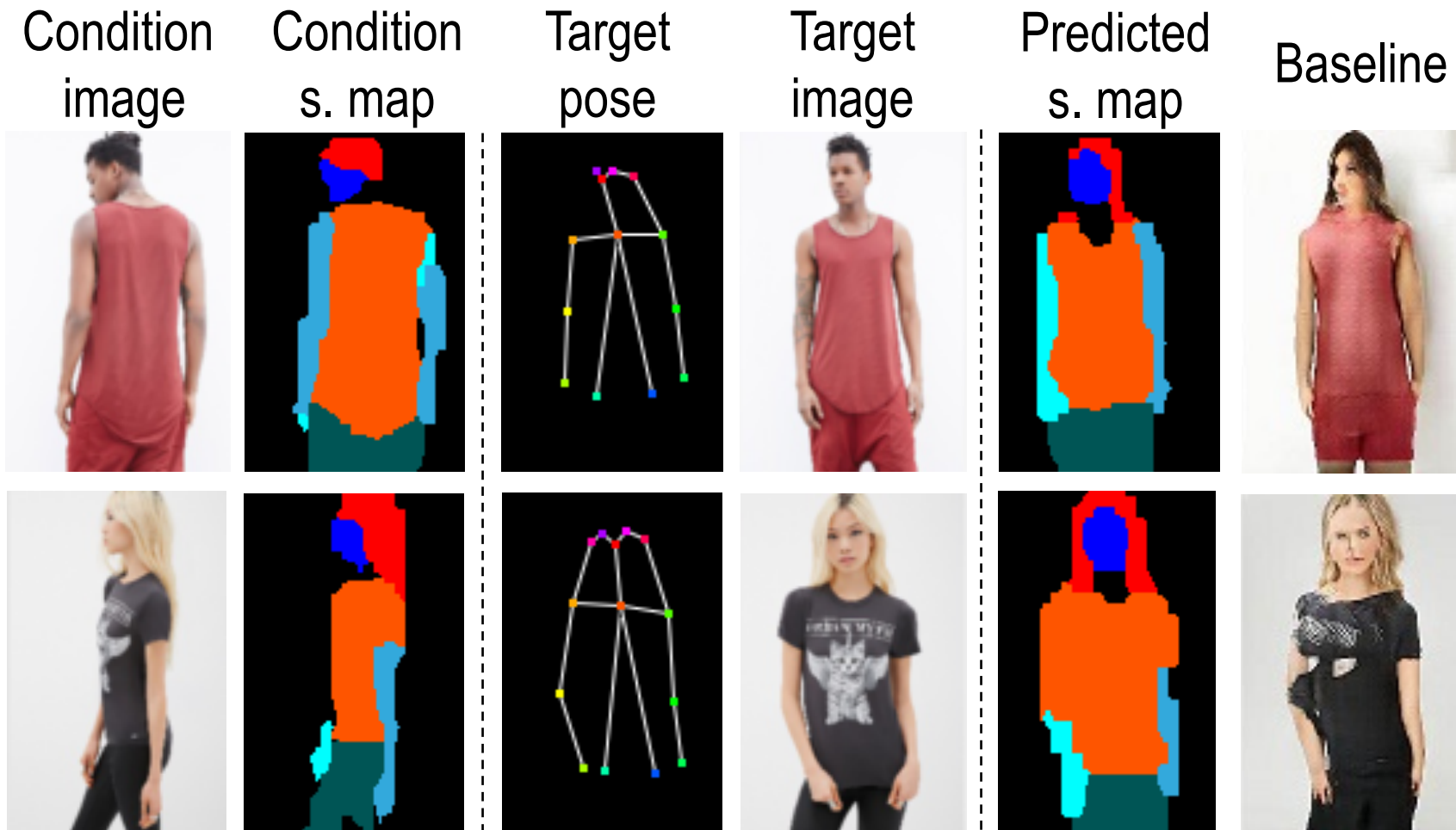


➔ End-to-end training

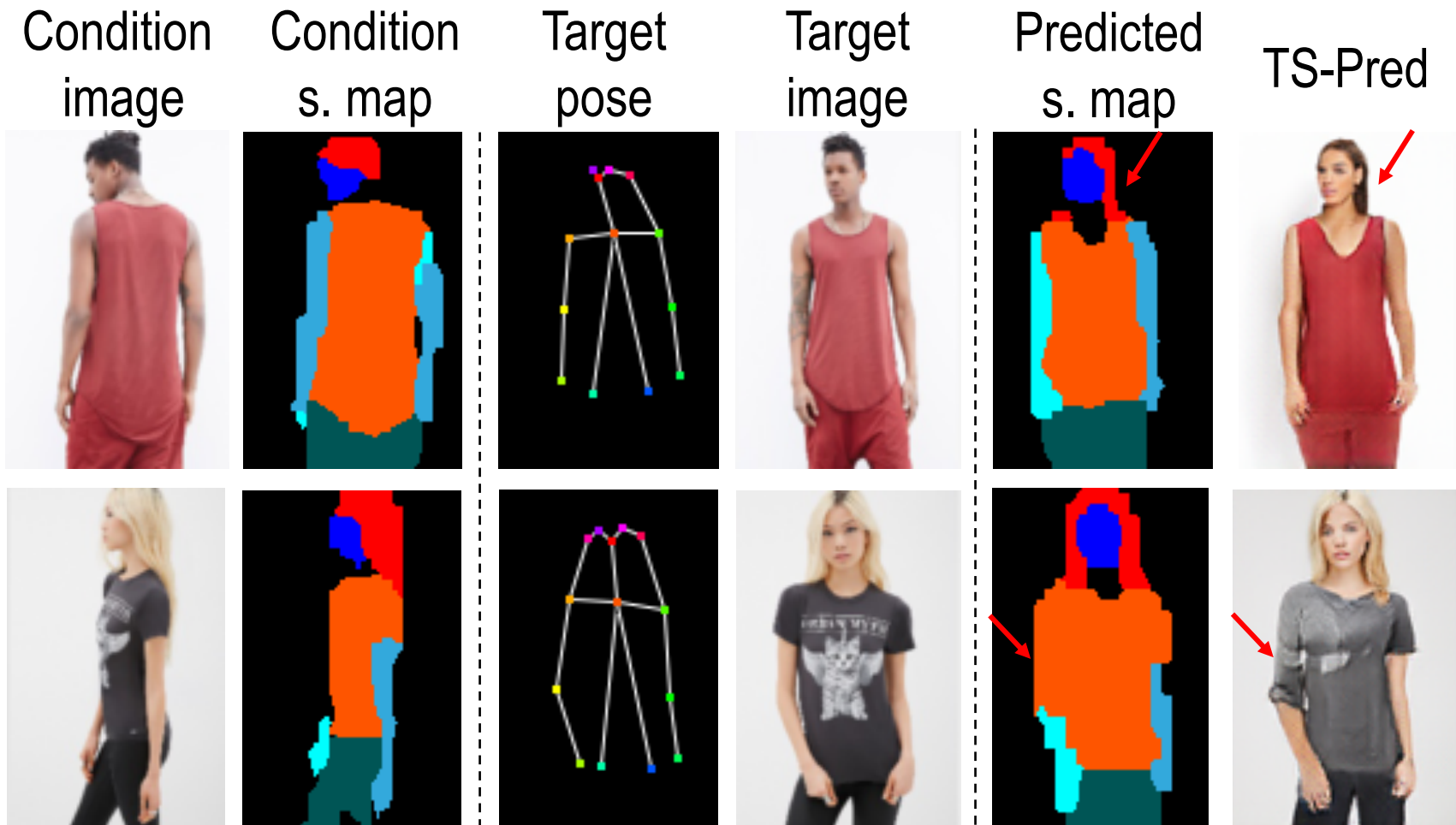
■ Effectiveness of Semantic Transformation



■ Effectiveness of Semantic Transformation



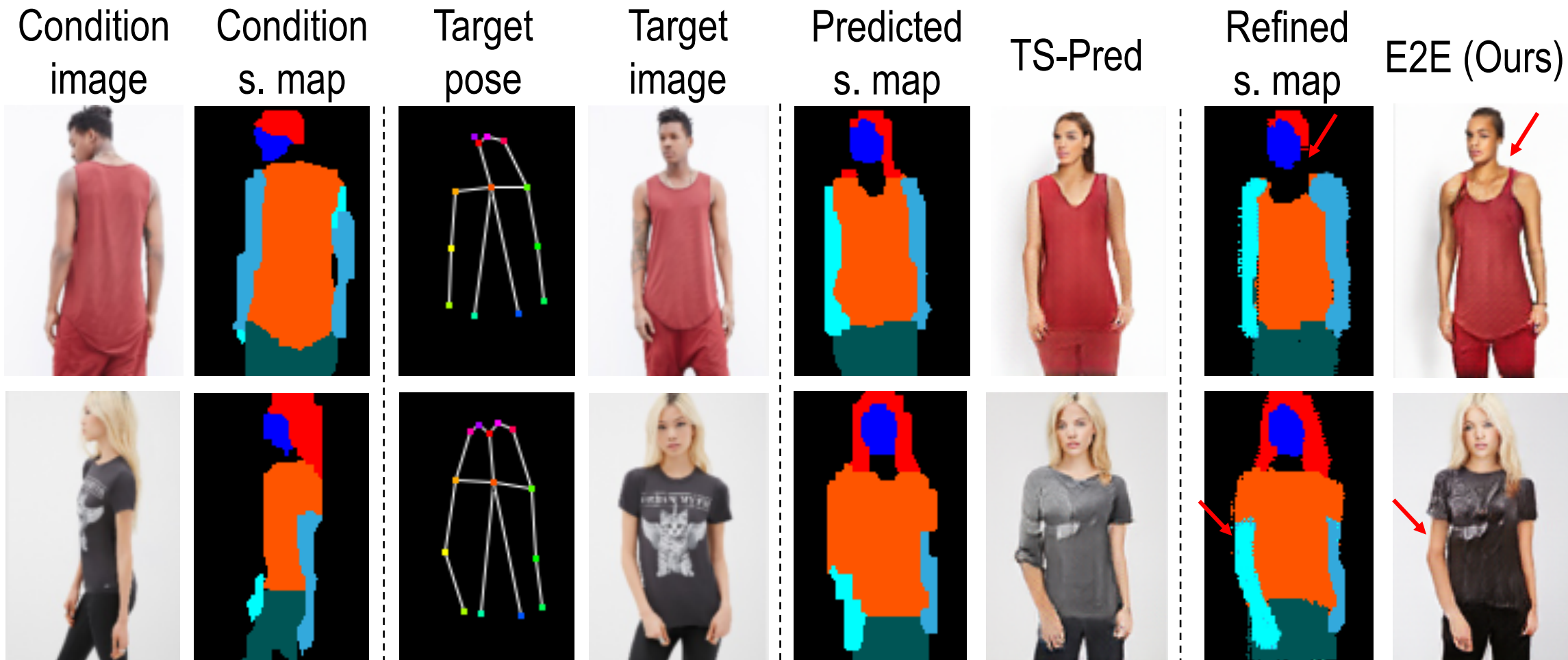
■ Effectiveness of Semantic Transformation



■ Effectiveness of End-to-End Training



■ Effectiveness of End-to-End Training



■ Pose-Guided Image Generation

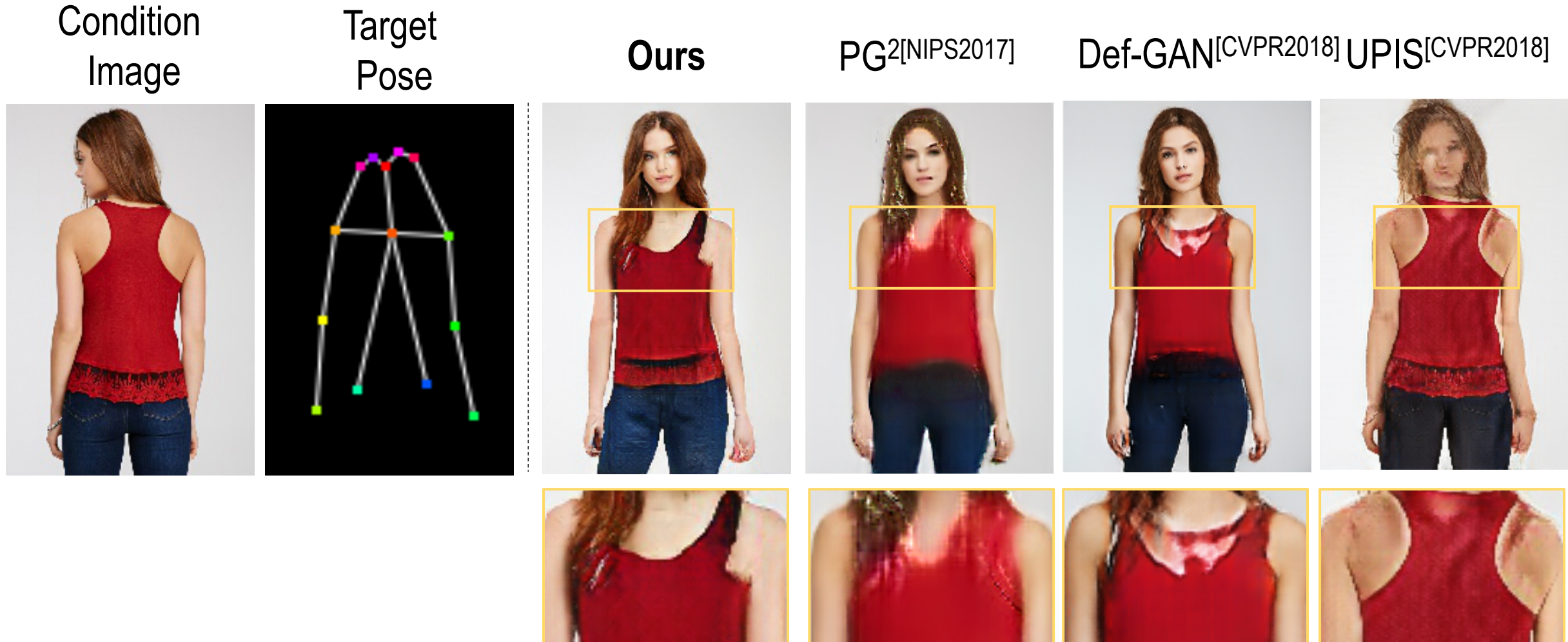


Ma L, Jia X, Sun Q, et al. Pose guided person image generation. NIPS 2017.

Siarohin A, Sangineto E, Lathuilière S, et al. Deformable gans for pose-based human image generation. CVPR 2018.

Pumarola A, Agudo A, Sanfeliu A, et al. Unsupervised Person Image Synthesis in Arbitrary. CVPR 2018.

■ Pose-Guided Image Generation



Ma L, Jia X, Sun Q, et al. Pose guided person image generation. NIPS 2017.

Siarohin A, Sangineto E, Lathuilière S, et al. Deformable gans for pose-based human image generation. CVPR 2018.

Pumarola A, Agudo A, Sanfeliu A, et al. Unsupervised Person Image Synthesis in Arbitrary. CVPR 2018.

■ Pose-Guided Image Generation

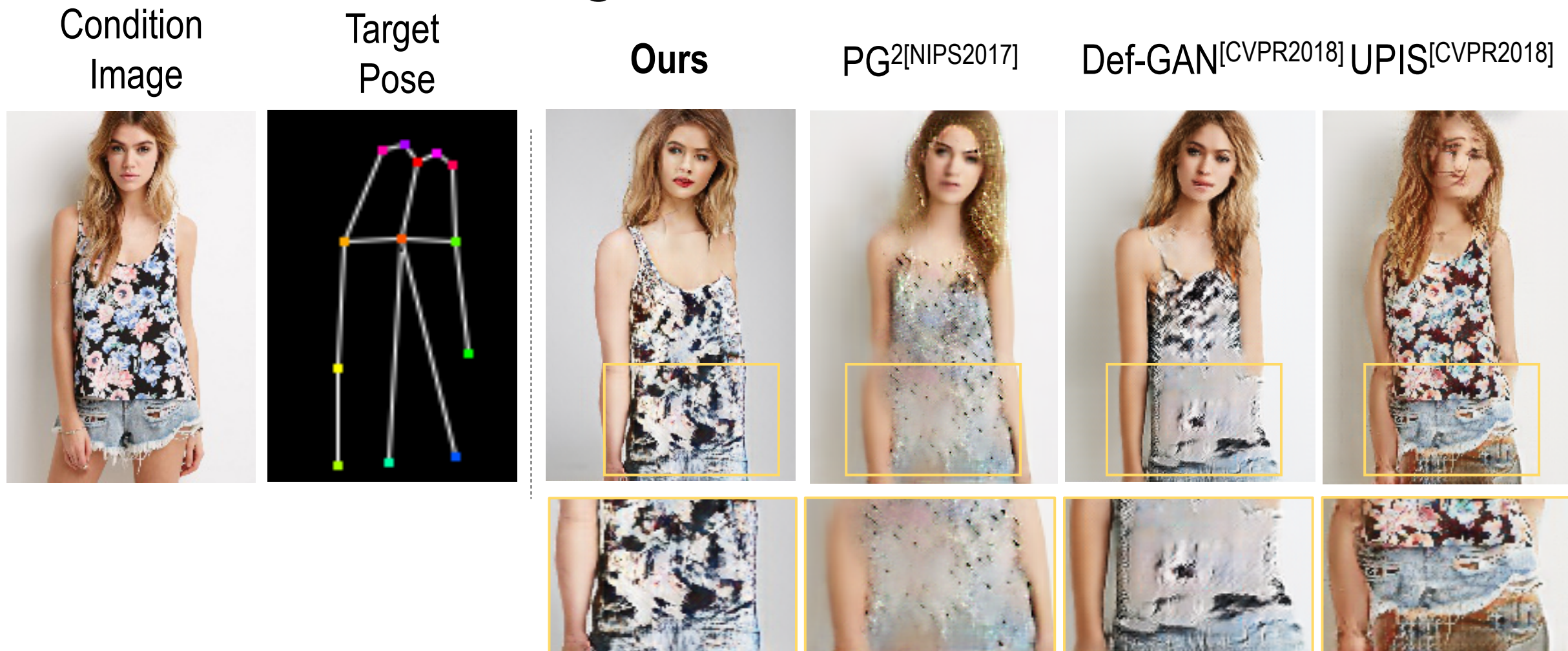


Ma L, Jia X, Sun Q, et al. Pose guided person image generation. NIPS 2017.

Siarohin A, Sangineto E, Lathuilière S, et al. Deformable gans for pose-based human image generation. CVPR 2018.

Pumarola A, Agudo A, Sanfeliu A, et al. Unsupervised Person Image Synthesis in Arbitrary. CVPR 2018.

■ Pose-Guided Image Generation



Ma L, Jia X, Sun Q, et al. Pose guided person image generation. NIPS 2017.

Siarohin A, Sangineto E, Lathuilière S, et al. Deformable gans for pose-based human image generation. CVPR 2018.

Pumarola A, Agudo A, Sanfeliu A, et al. Unsupervised Person Image Synthesis in Arbitrary. CVPR 2018.

■ Clothing Texture Transfer

Image A

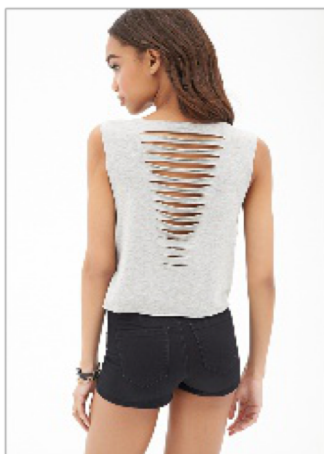


Image B



A → B



B → A



■ Controlled Image Manipulation

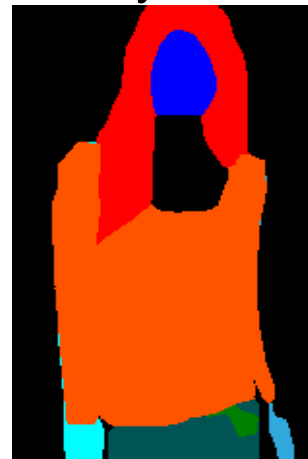
Condition
image



Original
layout



Desired
layout



Ours



■ **Unsupervised Person Image Generation**

- Semantic parsing transformation
- Appearance generation
- End-to-end training

■ **Experimental Results**

- Impressive results compared with other state-of-the-arts
- Generalizable to other tasks, i.e., clothing texture transfer and controlled image manipulation



Thank You!



Project Page

Poster Info:

Poster Tag : J122

Tuesday 3:20-6:00

Sijie Song: <https://sijiesong.github.io/>