huya虎牙

GAMES Webinar
Generative Models for Clothed Human

Xintong Han, Huya Inc

虎牙直播 | NiMO TV

2019. 11. 28

01

Compatible and Diverse Fashion Image Inpainting

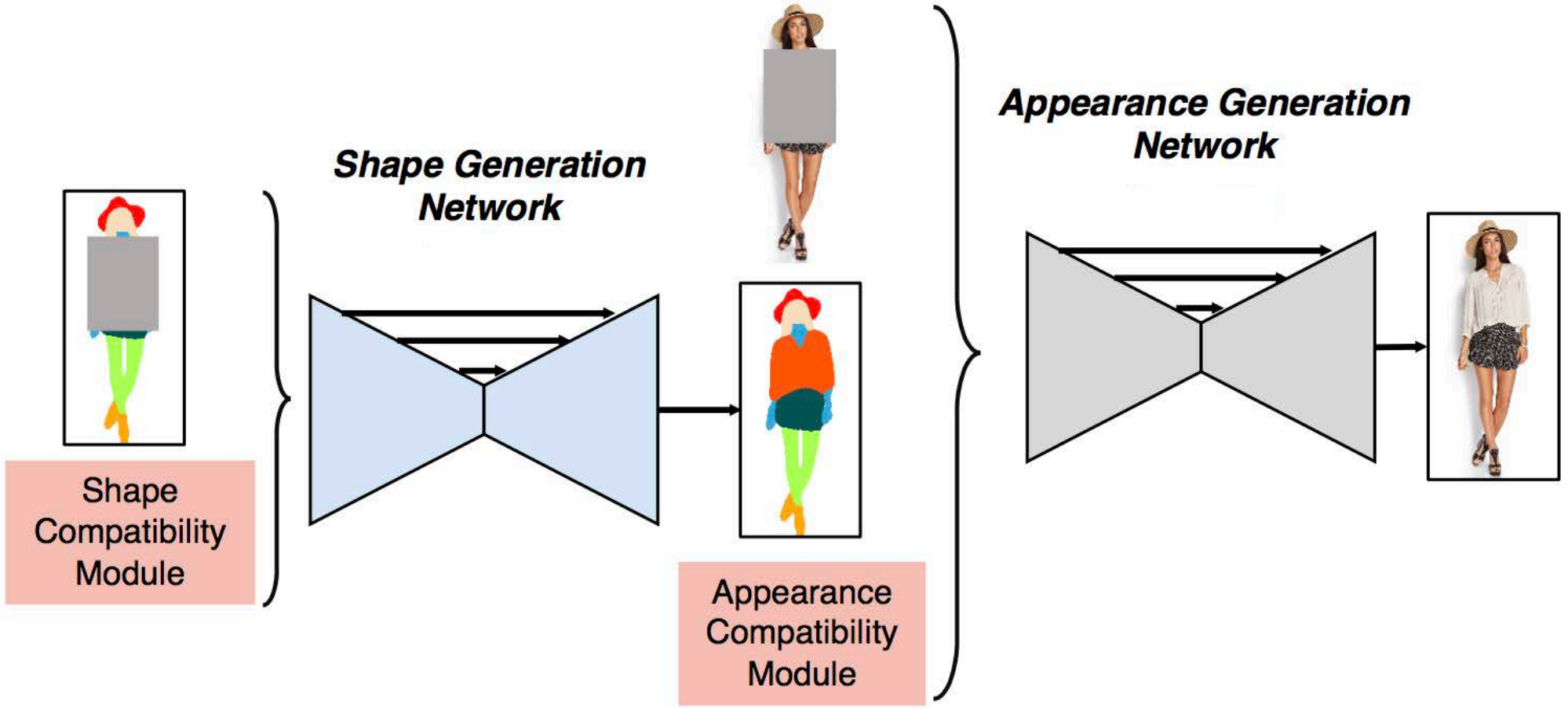
Fashion Recommendation as Inpainting

Compatibility: Matching style of existing garments

Diversity: Multiple results should be generated

Shape and **Appearance**

Framework



Two-stage framework that disentangles the generation of shape and appearance.

Shape Generation Network

Image-to-image Translation (U-Net):

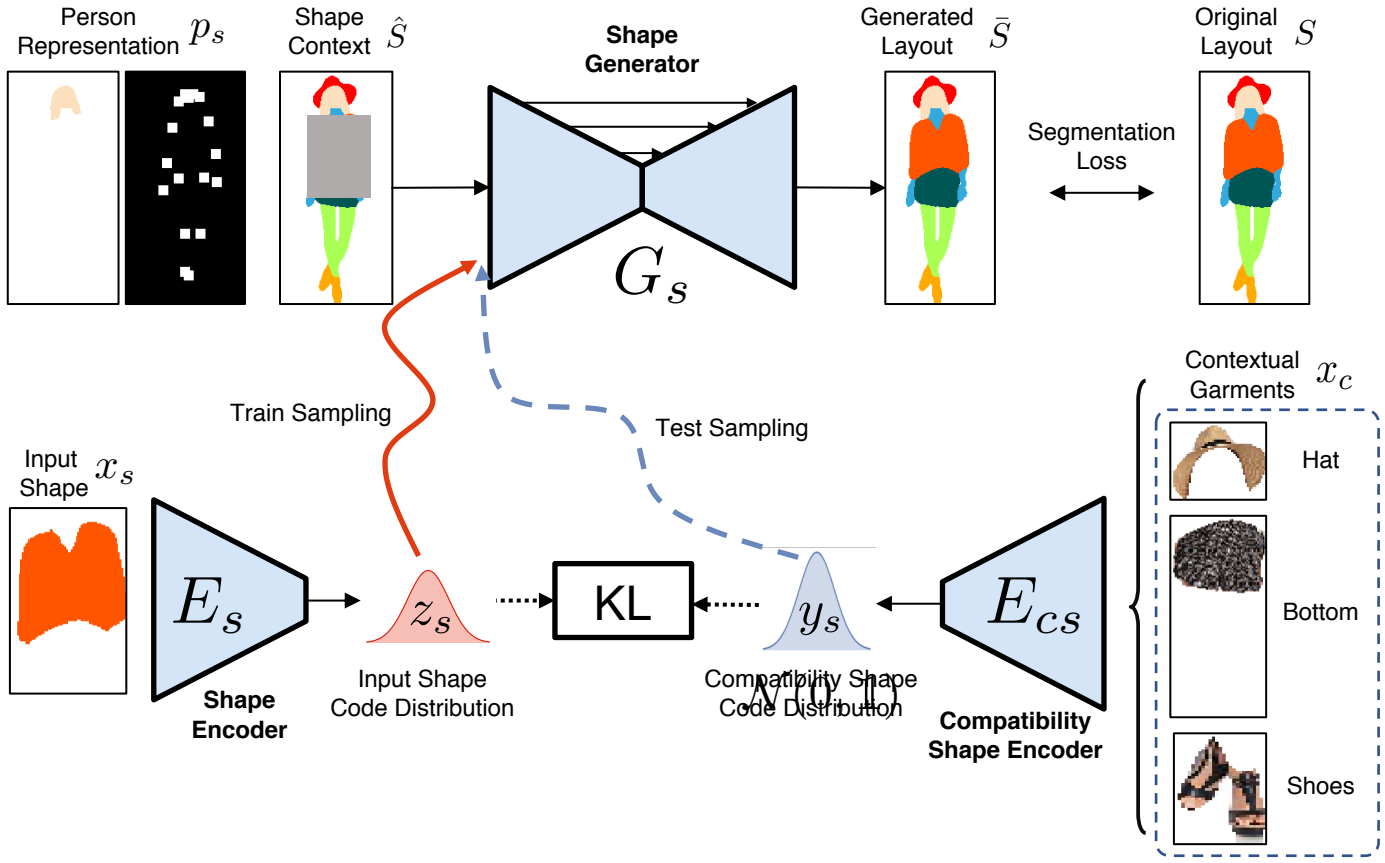
Person representation + Shape context -> Layout

+Variational Autoencoder (VAE):

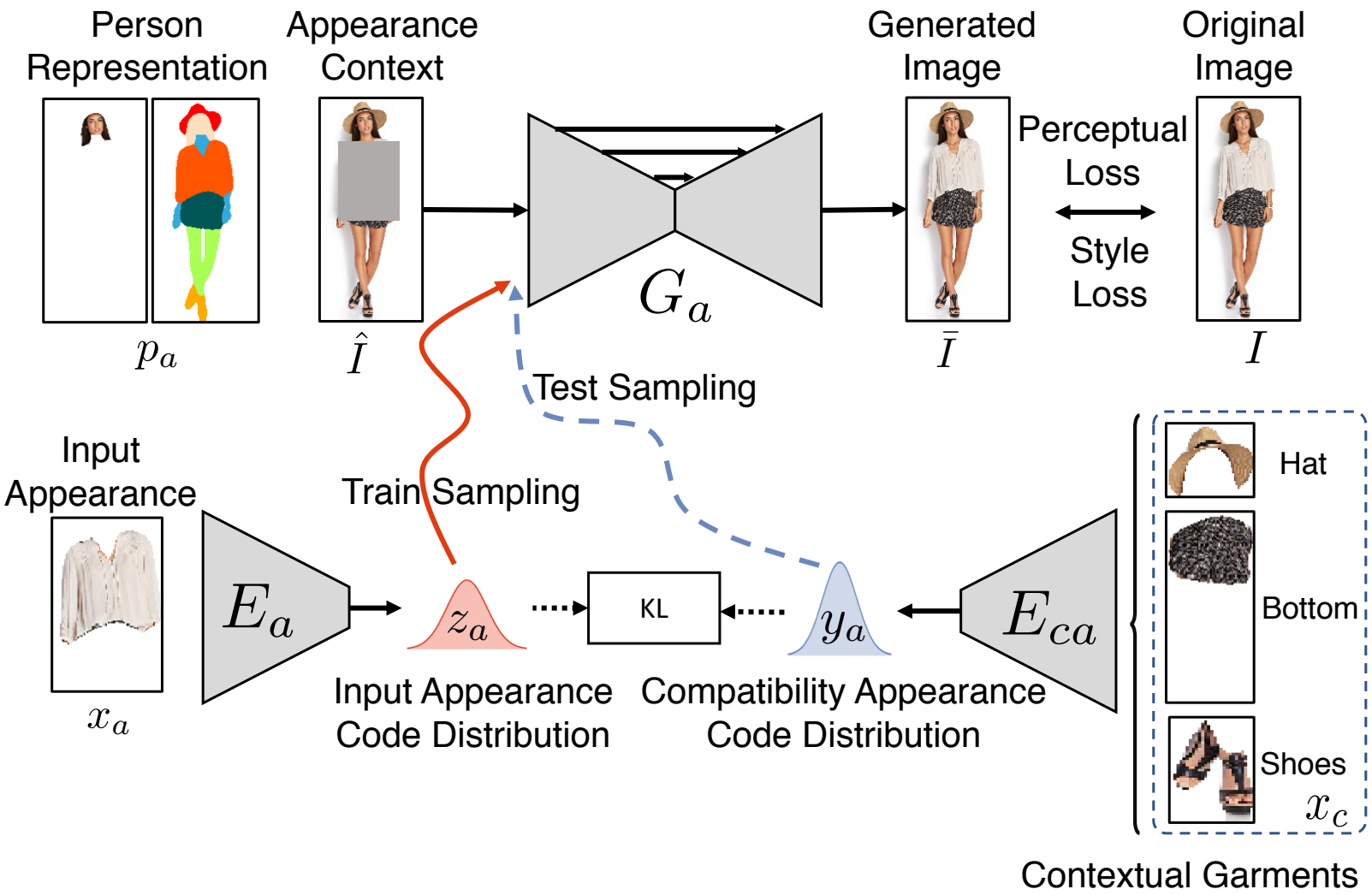
Enable diversity at test time

+Compatibility Encoder:

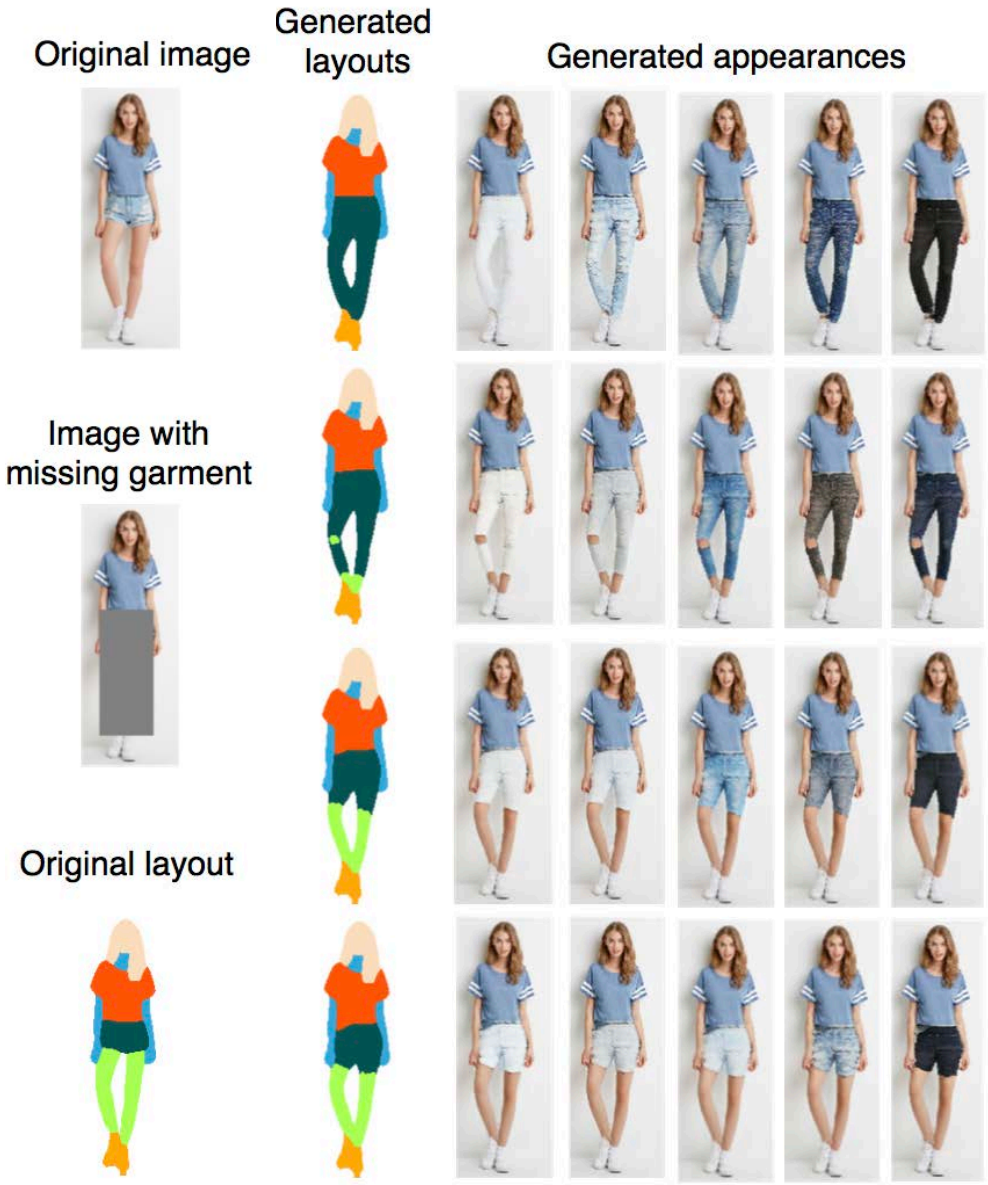
Inject compatibility (co-occurrence) using contextual garments

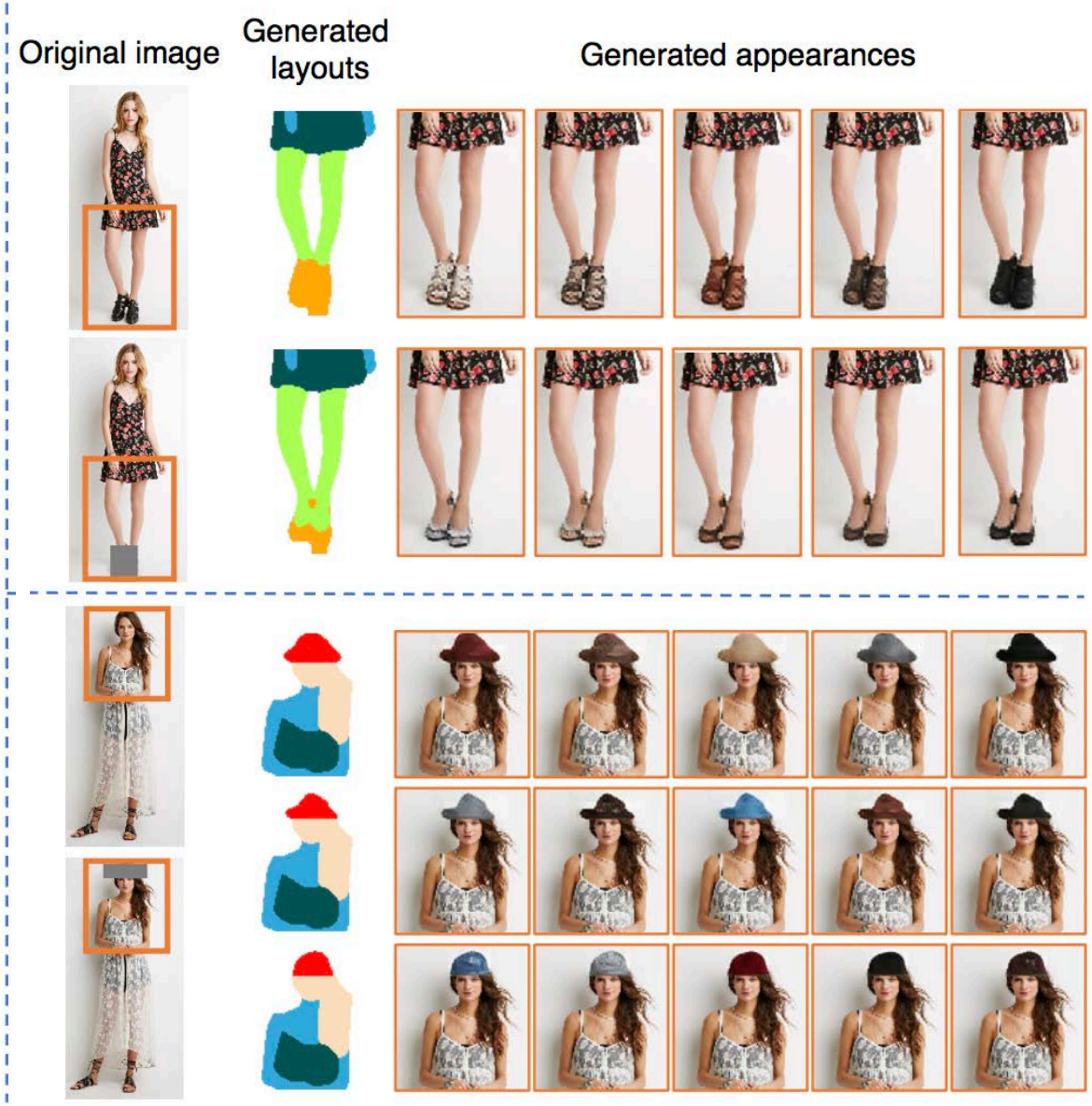


Appearance Generation Network



Inpainting Results





Comparison with Other Methods

Original image and input

FiNet

FiNet w/o two-stage

FiNet w/o two-stage w/o comp

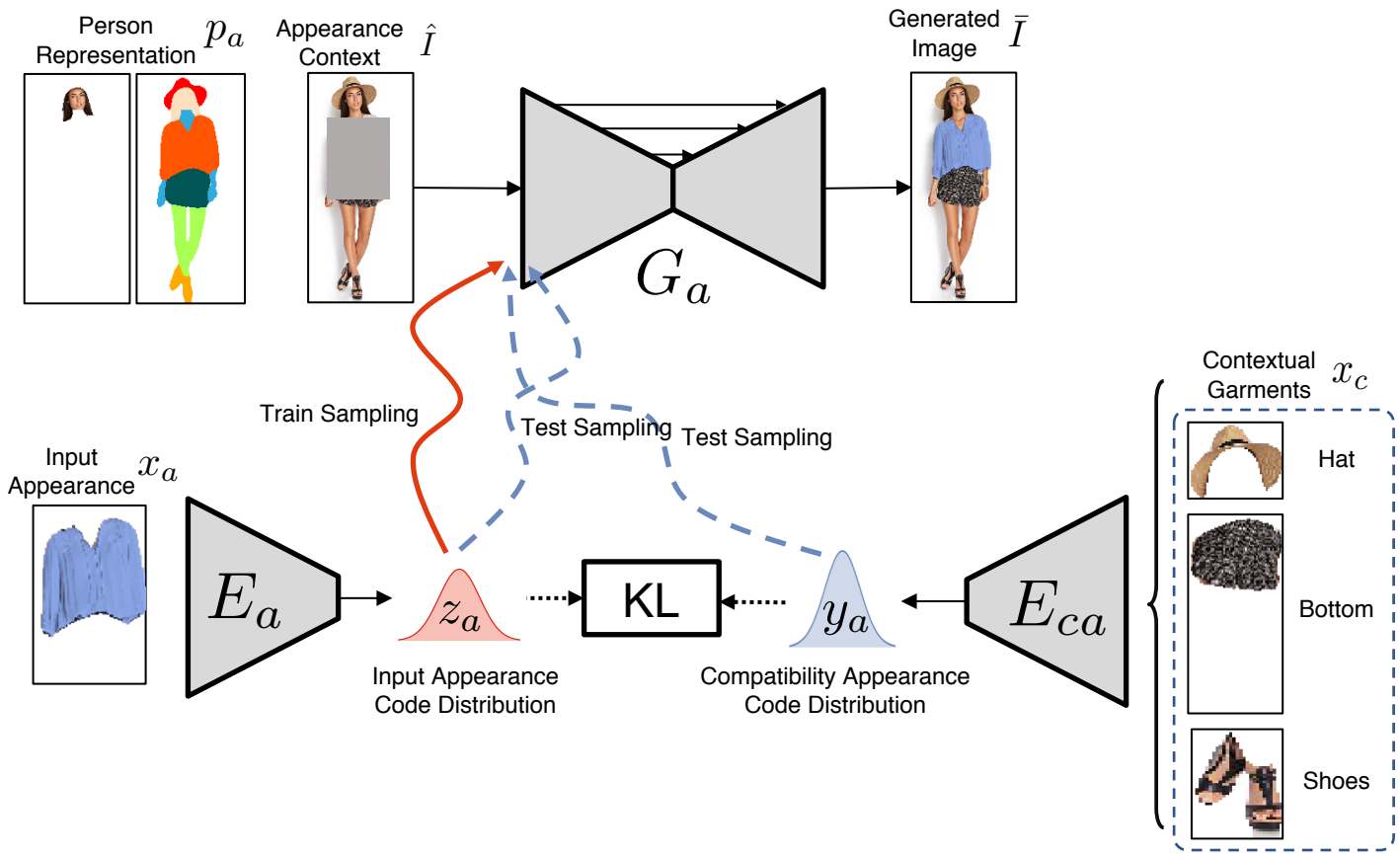
Pix2Pix + noise

BicyleGAN

VUNET

ClothNet

Virtual Try-on Applications

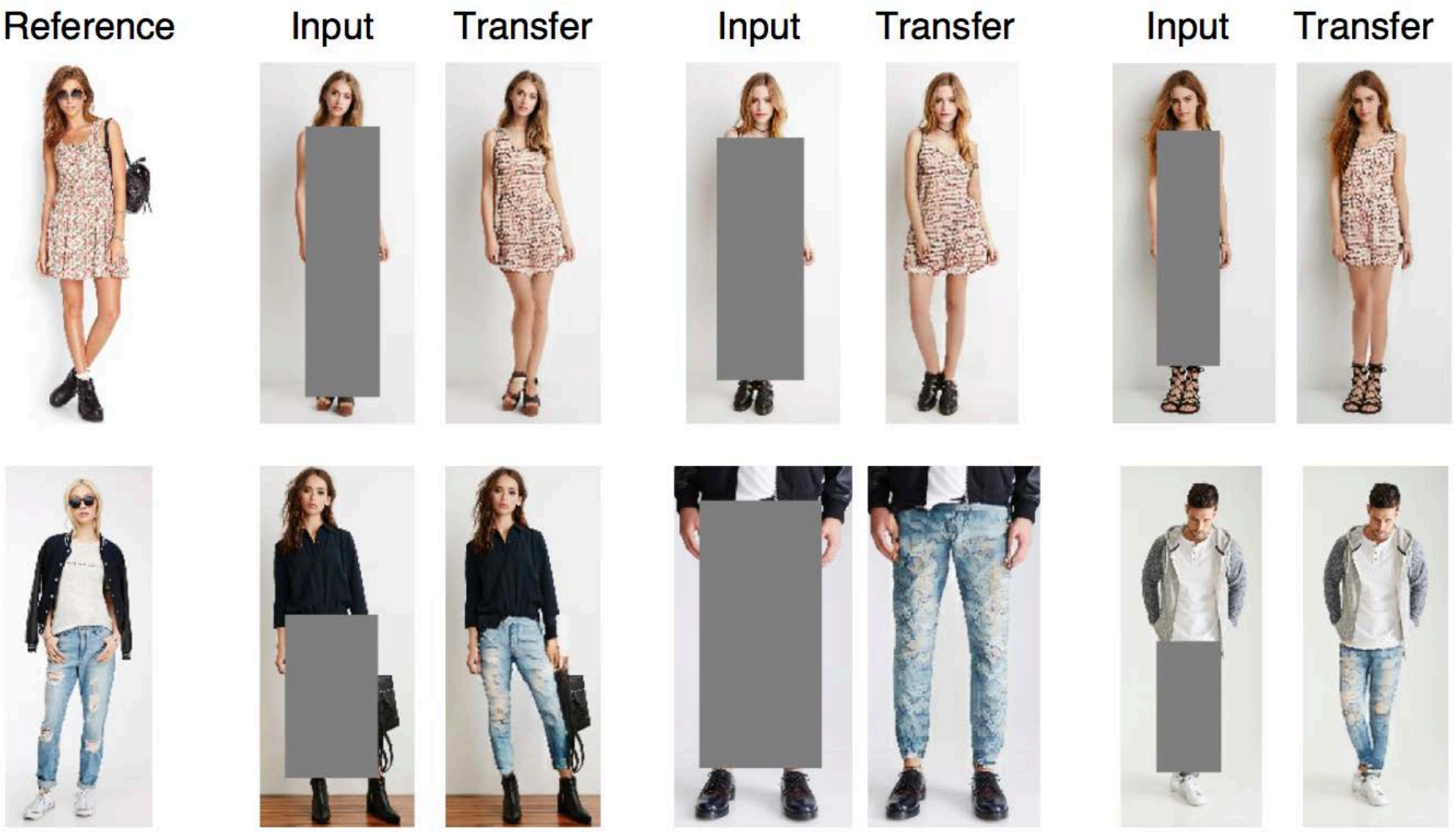


Disable compatibility encoder.

Test sampling from shape/appearance encoder.

Input a target clothing for reconstruction.

Virtual Try-on Results



Virtual Try-on Results

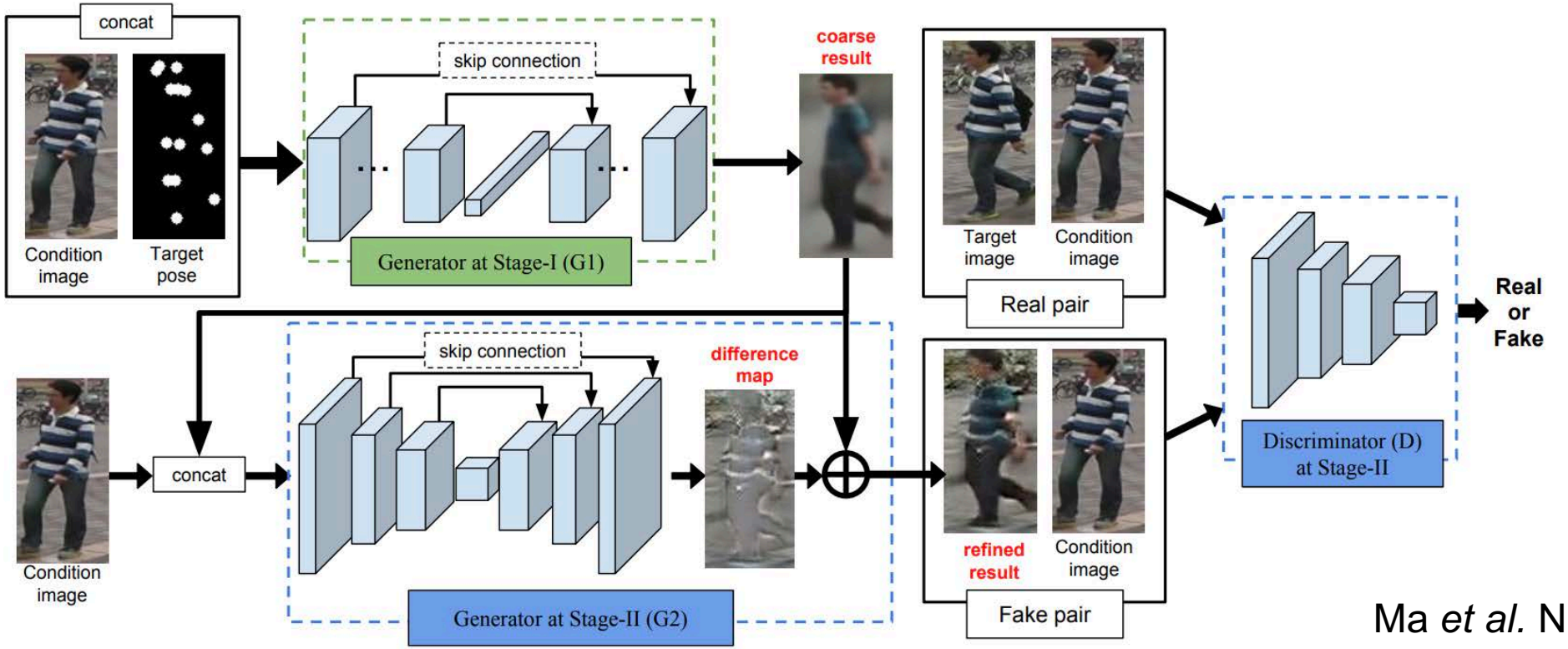
02

ClothFlow: A Flow-Based Model for Clothed Person Generation

huya虎牙

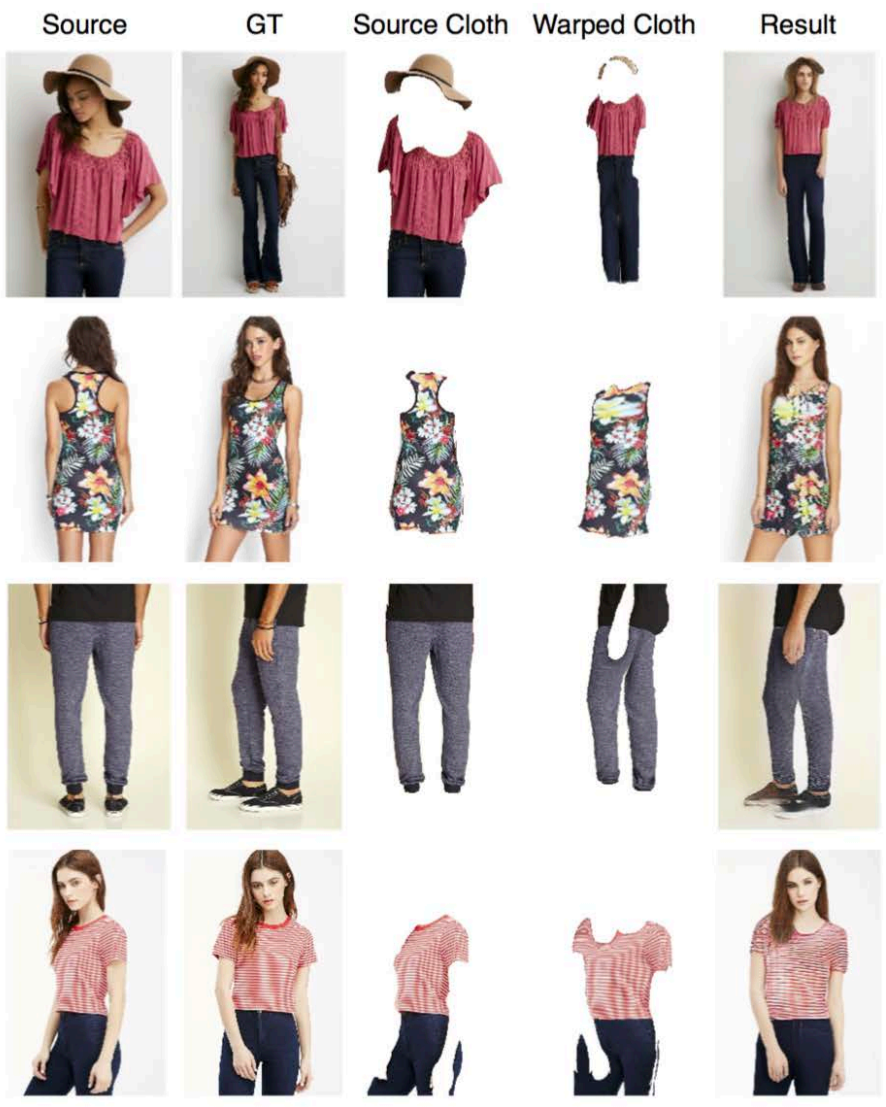
Pose-guided Person Image Generation

- **Does not model geometric transformations.**
- **Cannot preserve appearance details (textures, logos, graphics).**

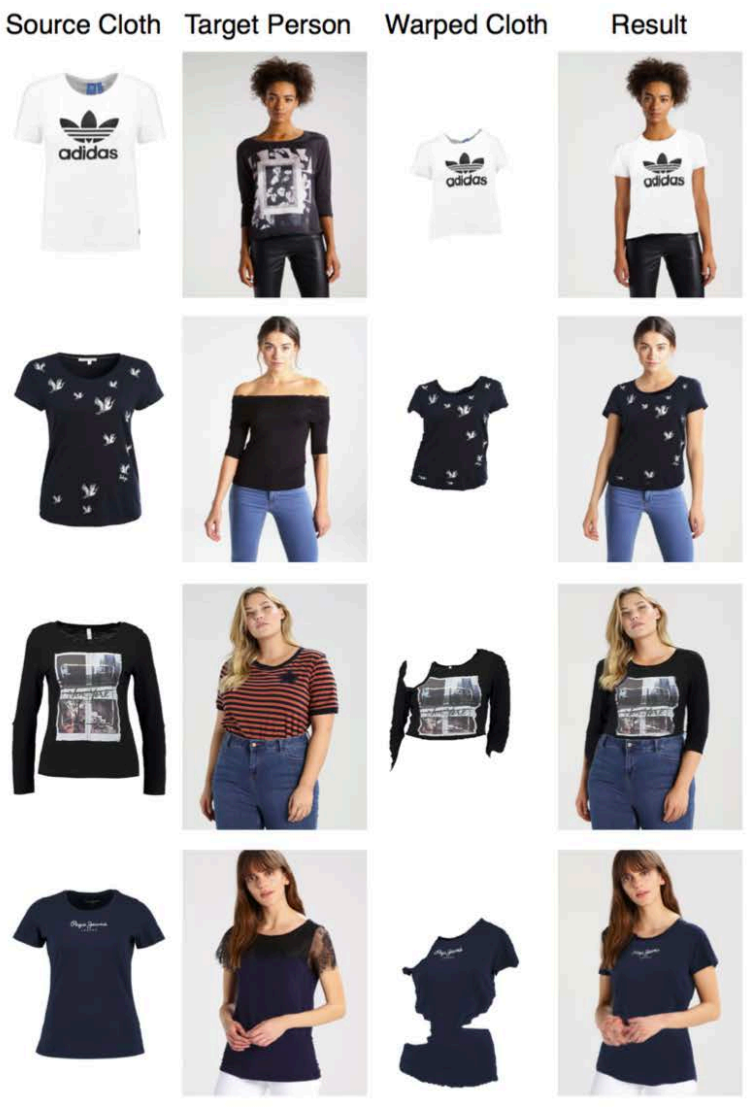


Ma *et al.* NIPS 2017

Our Results



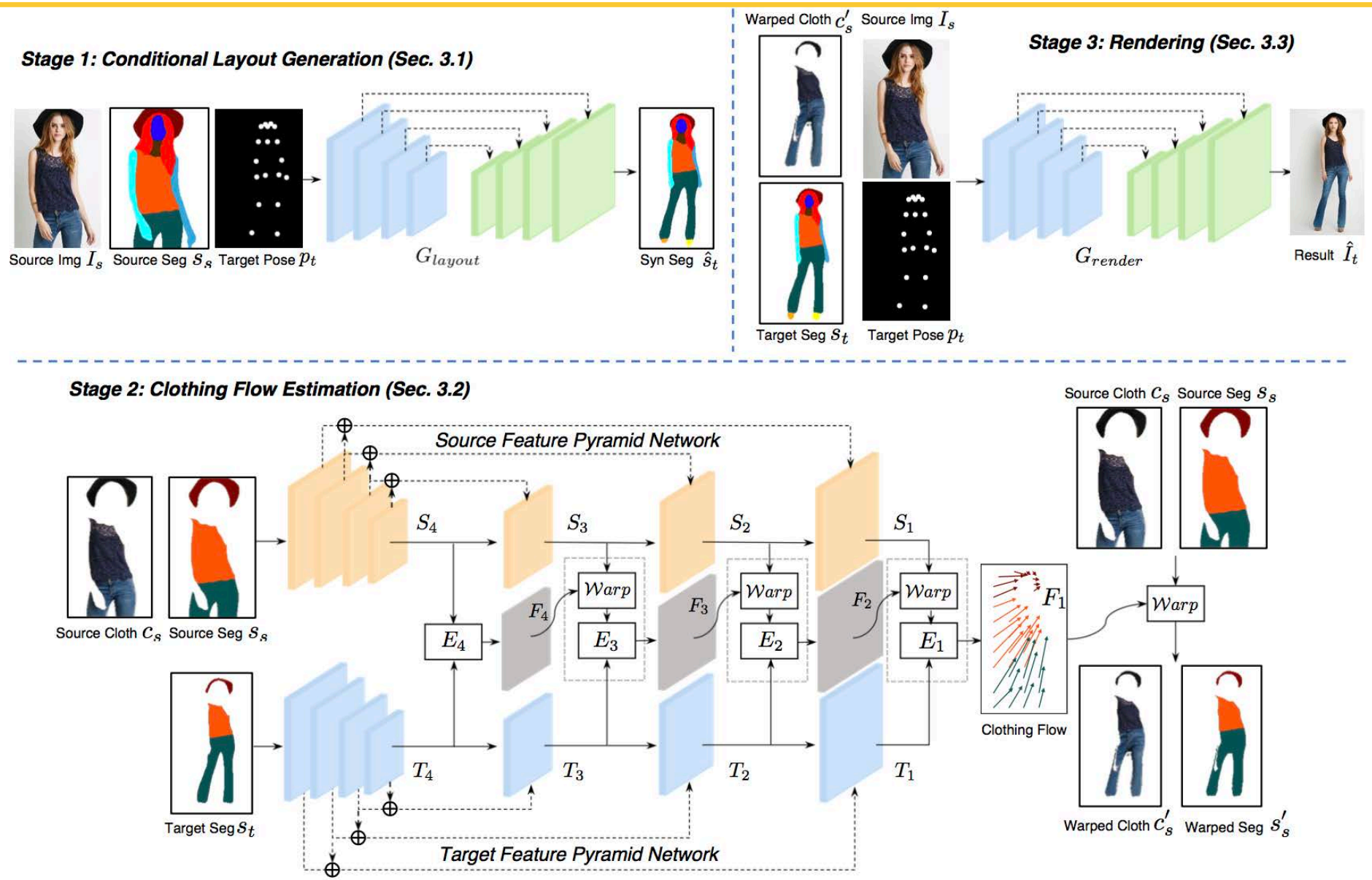
Pose guided person generation



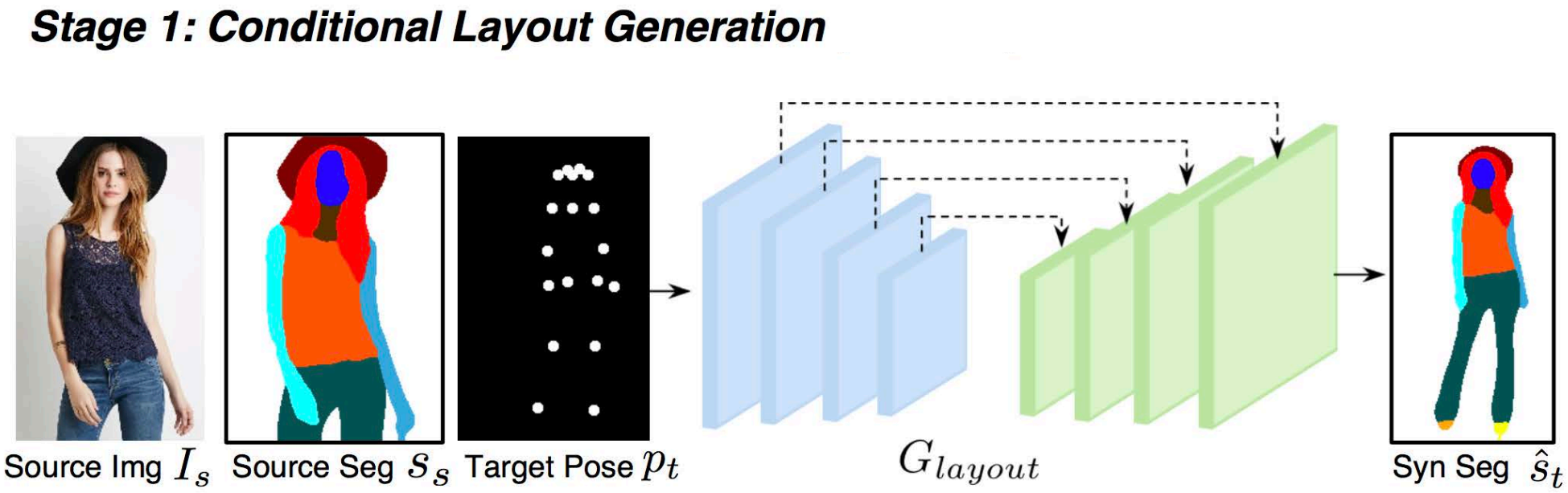
Virtual try-on

- Warp clothing regions by estimating **appearance flow** between two images.
- Appearance flow: a 2D **dense flow field** specifying which pixels in the source image could be **redirected** to reconstruct the target image.

ClothFlow Framework

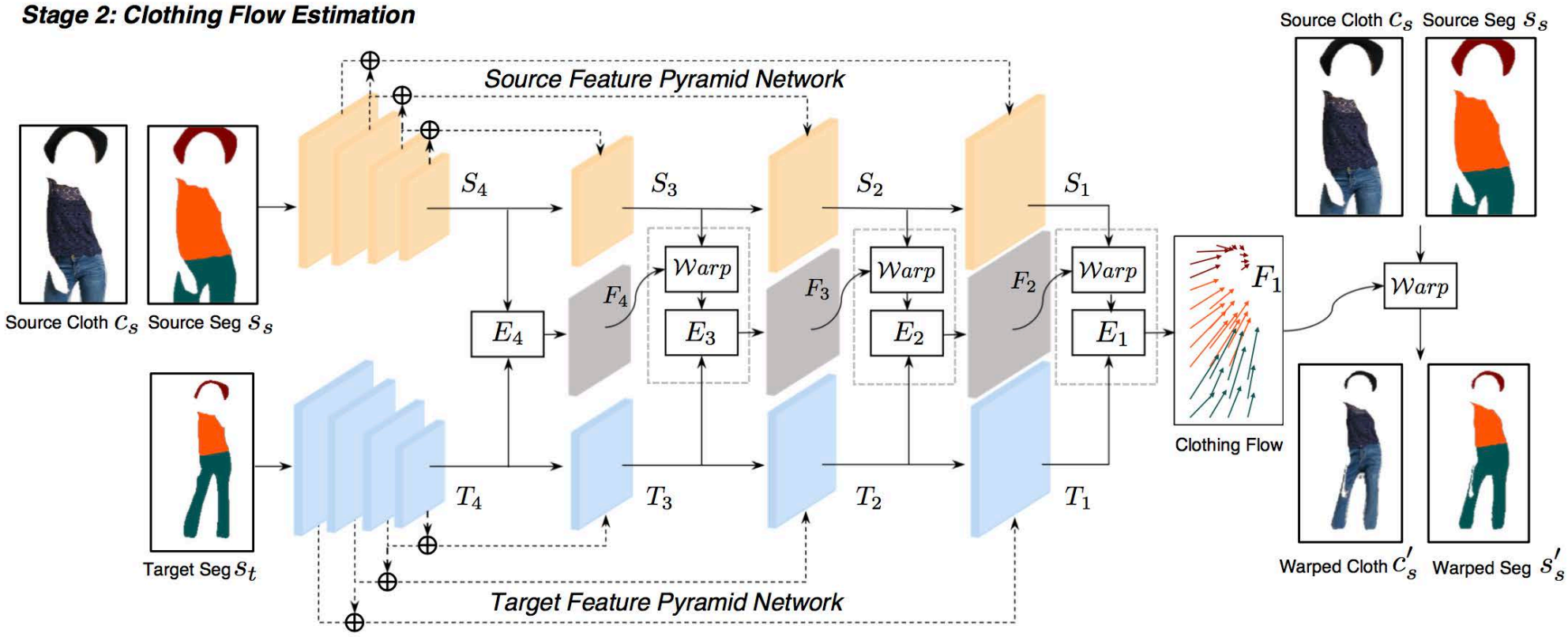


Hallucinate Target Segmentation Map



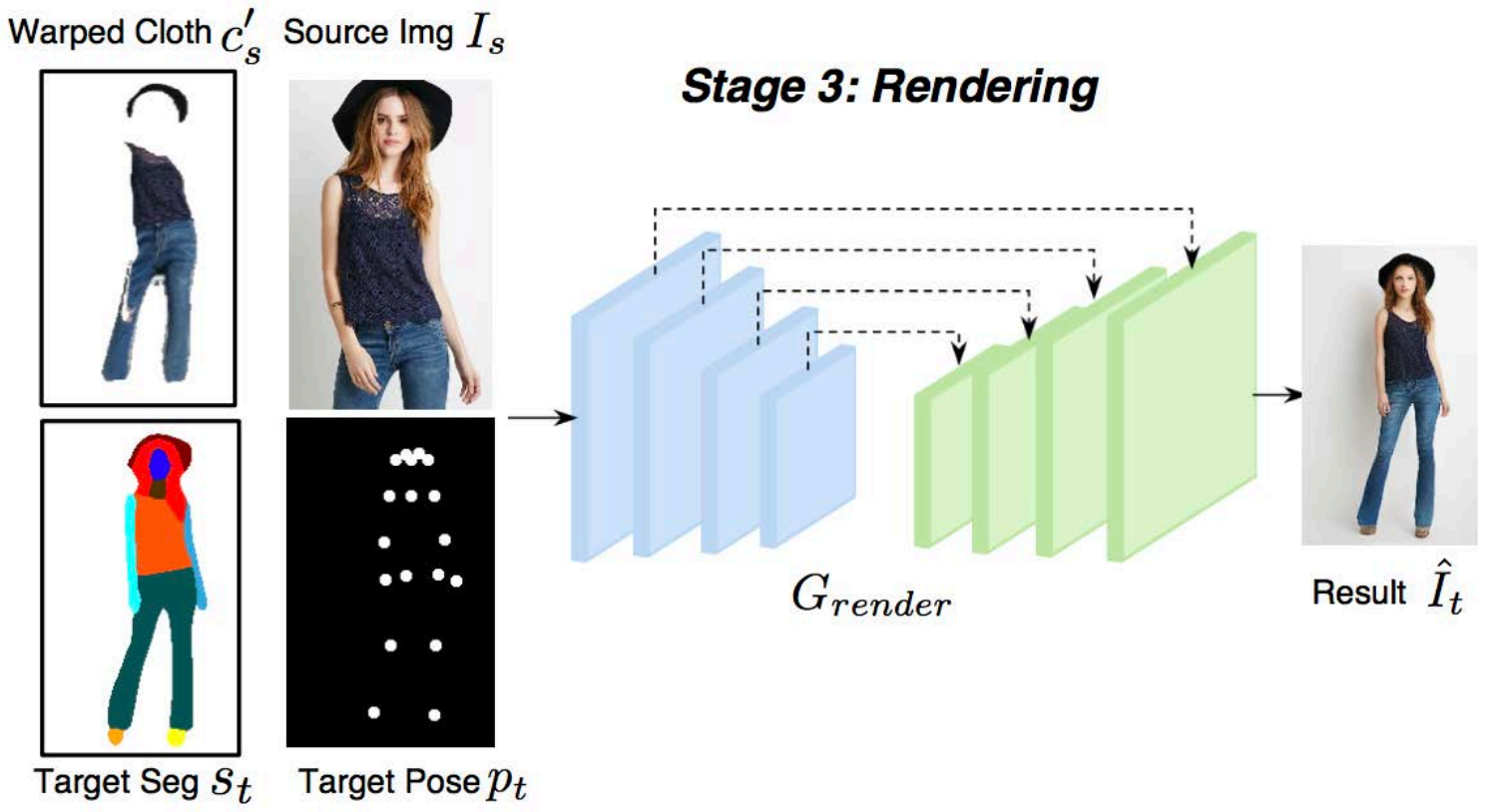
- Extract poses and segmentation maps with off-the-shelf models.
- U-Net predicts the target segmentation map with segmentation loss.
- Provide structural constraints for generating appearance.

Estimate Flow between Source and Target



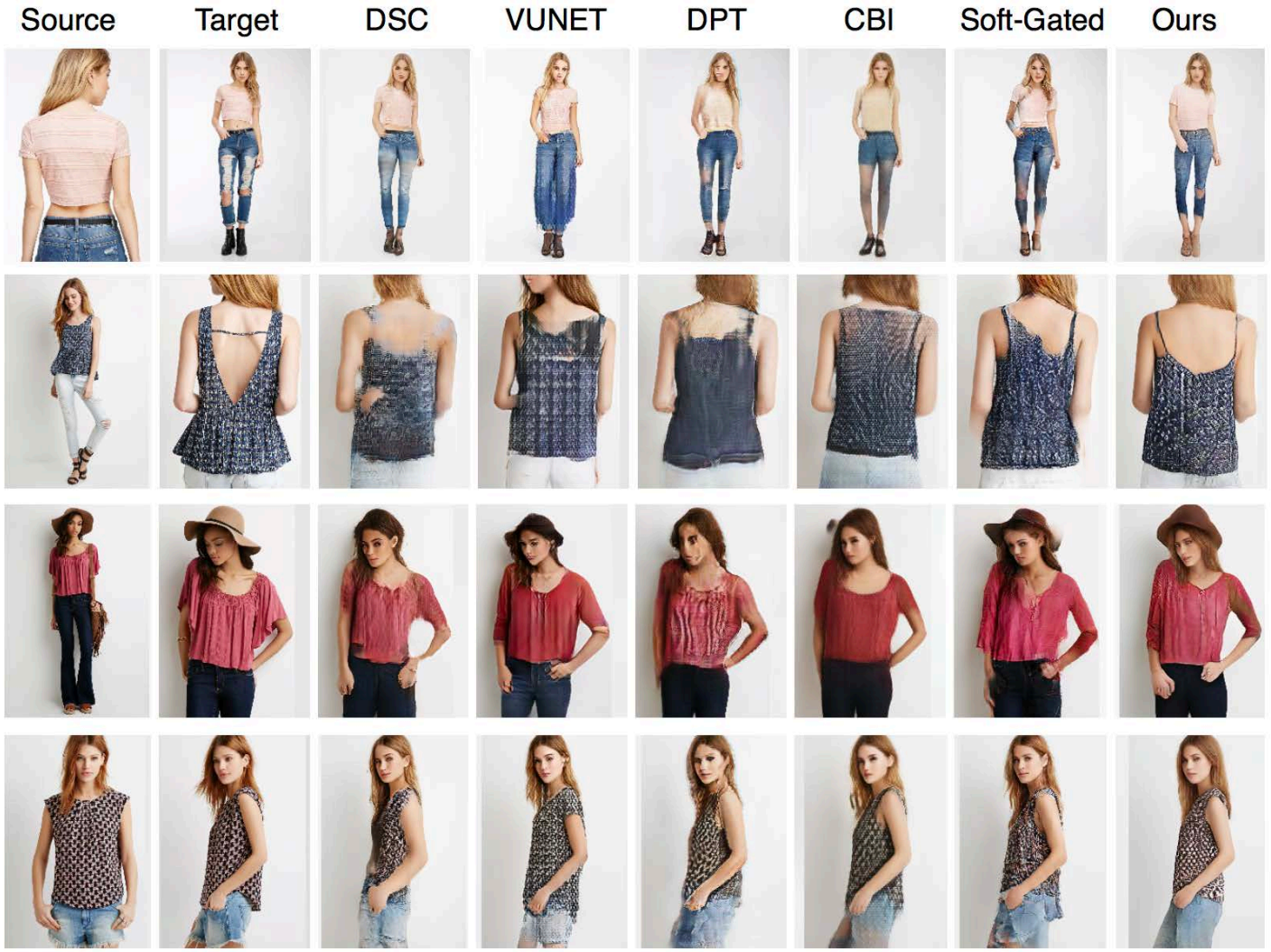
- Borrow idea from optical flow estimation.
- Dual feature pyramid networks to encode source and target information.
- Gradually estimate the appearance flow in a cascaded fashion.
- Warped clothing: perceptual loss; Warped segmentation: L1 loss.

Render the Final Result



- **U-Net taking all guidance to synthesize the final result.**
- **Train: use GT target segmentation. Testing: Synthesized segmentation from stage 1.**

Pose-guided Person Generation Results



Quantitative comparison

Methods	Deform	Dense	SSIM	IS
PG^2 [26]	✗	✗	0.762	3.09
DSC [34]	✓	✗	0.761	3.35
VUNET [7]	✗	✗	0.786	3.09
BodyROI7 [27]	✗	✗	0.614	3.23
DPT [28]	✗	✓	0.785	3.61
Soft-Gated [4]	✓	✗	0.793	3.31
CBI [10]	✓	✓	**0.835**	2.92
w/o Layout	✓	✗	0.758	3.63
w/o Flow	✗	✗	0.757	3.71
w/o Flow + TPS	✓	✗	0.758	3.74
w/o Cascade	✓	✗	0.759	3.74
w/o Style	✓	✗	0.756	3.56
ClothFlow	✓	✗	0.760	3.75
ClothFlowDense	✓	✓	0.771	**3.88**

Pair-wise user preference

DSC [34]	VUNET [7]	DPT [28]	CBI [10]	Soft-Gated [4]
80.9%	63.4%	90.2%	69.7%	57.9%

Virtual Try-on: Replace the source image by a clothing image

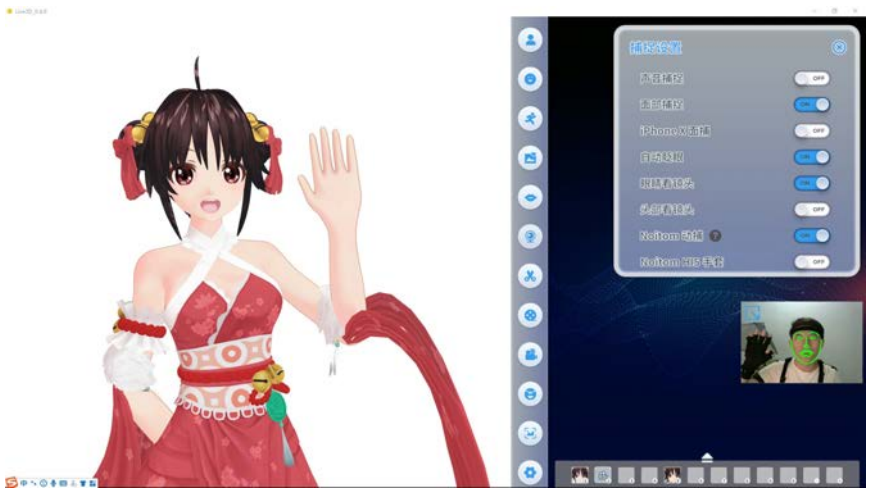
03

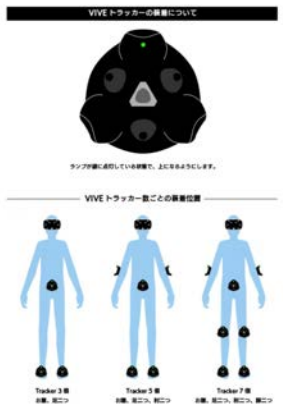
Content Creation at Huya

huya虎牙

虚拟直播风声水起

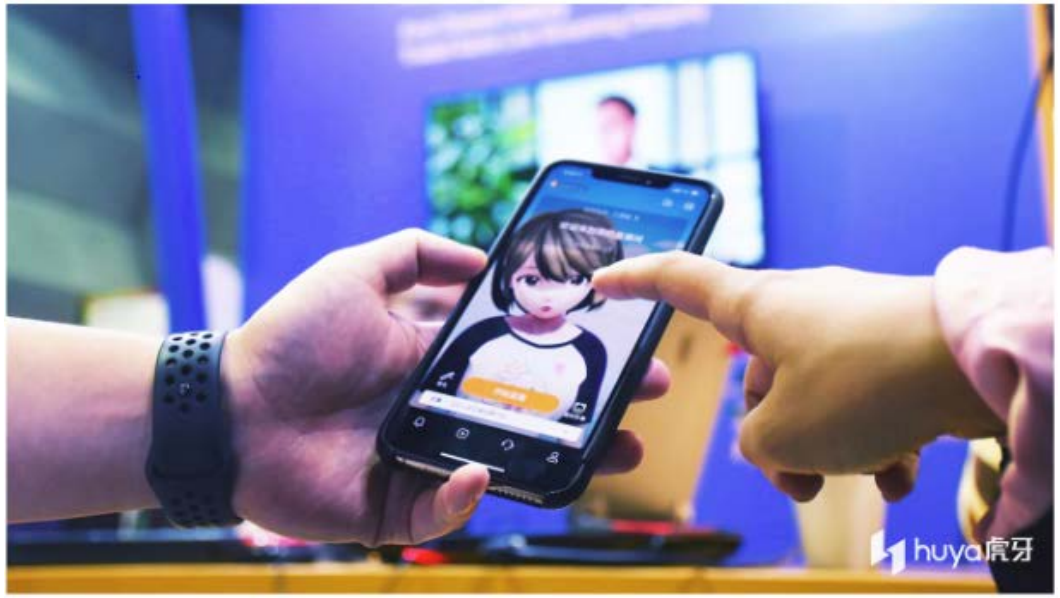

Virtual Cast™ (β)
バーチャルキャスト
青亭网



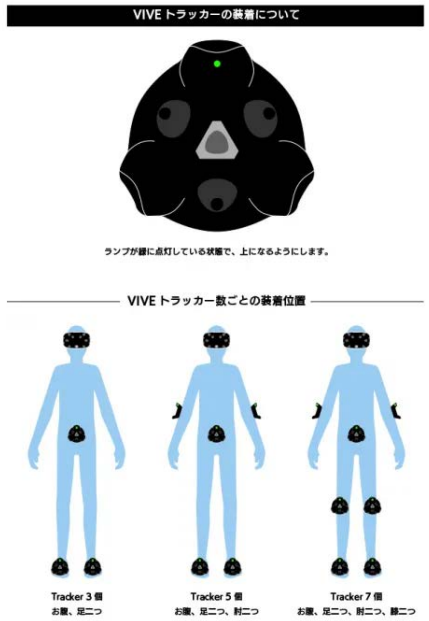

VIVE トラッカーの装着について
VIVE トラッカー装着位置


facerig
SOFTWARE
WEBCAM
facerig

虎牙的虚拟直播



存在问题



依赖动捕设备

虚拟形象自由度低

限制小主播，缺少新玩法

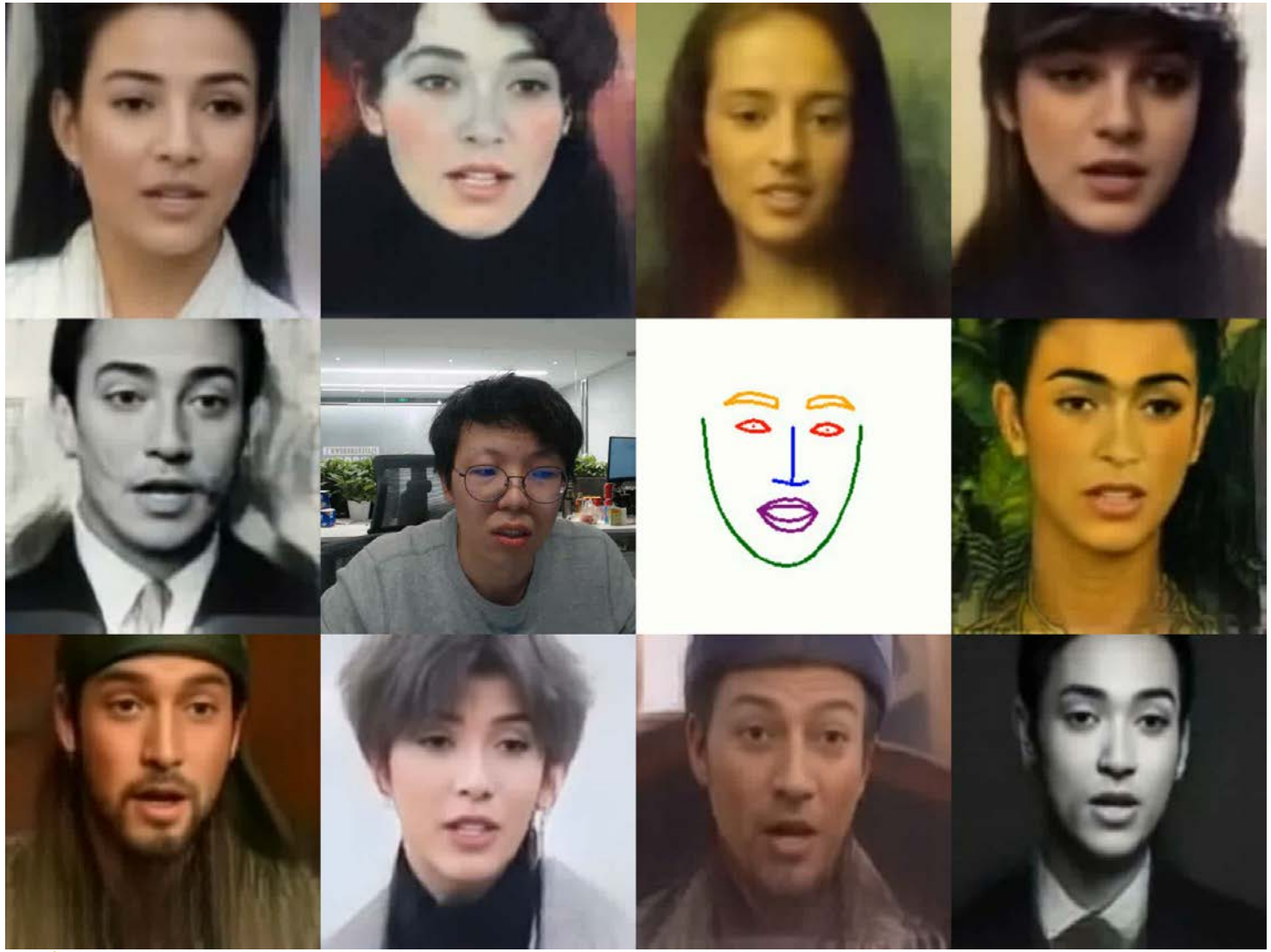
低成本方案：基于单张图片的人脸驱动虚拟直播

第一步：上传单张图片

第二步：单目摄像头直播

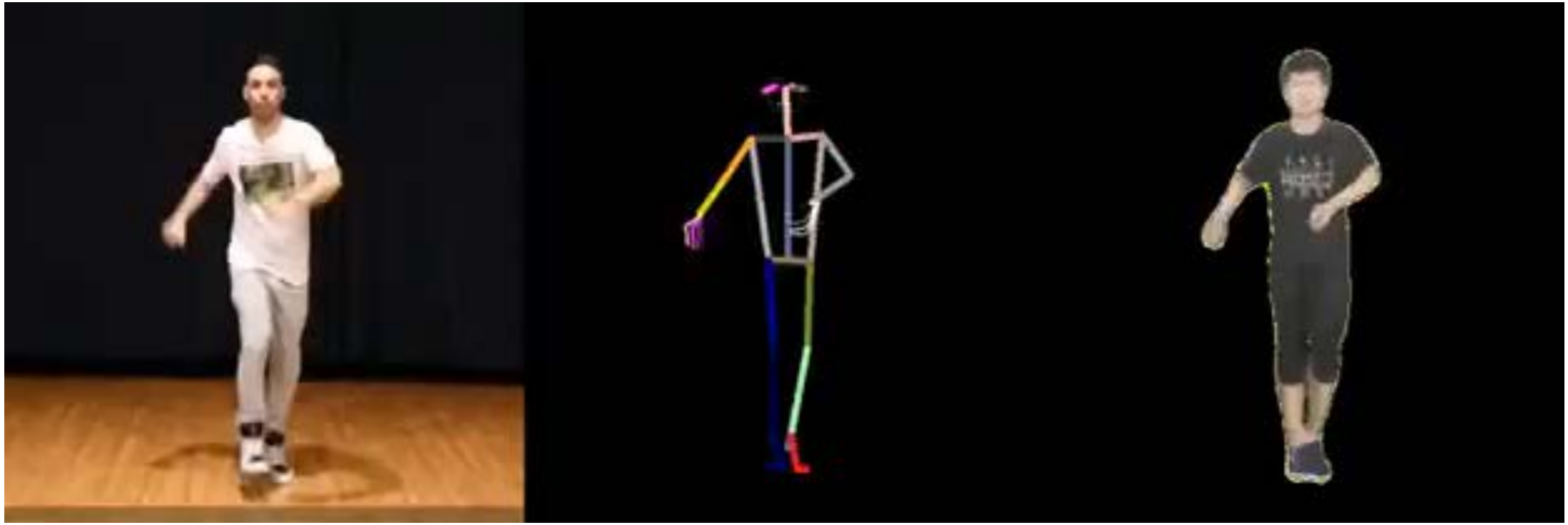
第三步：动起来

更多结果



该算法有两种模式：
左：无法很好的保持身份，头可以自由转动
右：头不能自由转动，但是身份保持得很好

肢体驱动



招聘啦

当前是视觉+图形学最好的时代！

地点：深圳/广州

职位：全职/实习

视觉/图形学：人脸/人体 建模/驱动/生成

语音：识别/合成

联系方式： hanxintong@huya.com

huya虎牙

THANK YOU

虎牙直播 | NiMO TV