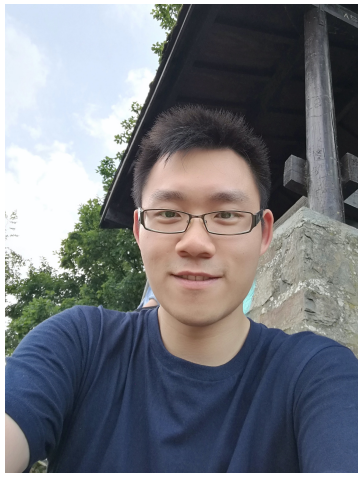
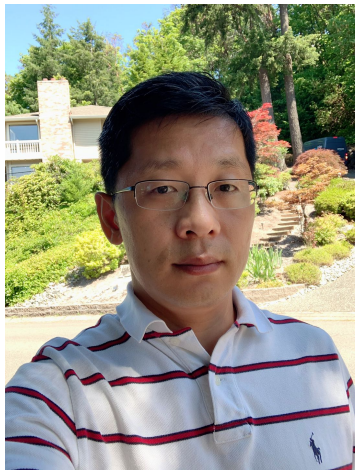
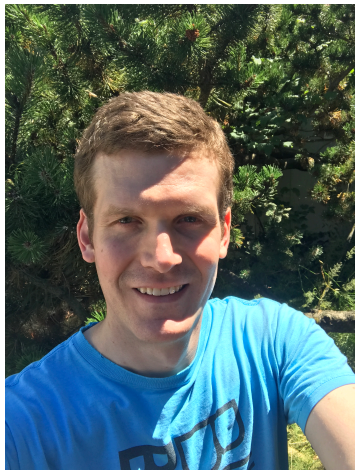
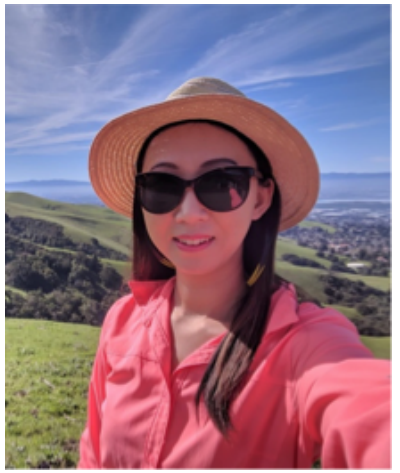
ECCV'20 ONLINE
23-28 AUGUST 2020

Unselfie: Translating Selfies to Neutral-pose Portraits in the Wild

Project page: http://charliememory.github.io/ECCV20_Unselfie/



Liqian Ma[1] Zhe Lin[2] Connelly Barnes[2] Alyosha Efros[2,3] Jingwan Lu[2]

[1]KU Leuven [2]Adobe Research [3]UC Berkeley

KU LEUVEN

Adobe

Berkeley
UNIVERSITY OF CALIFORNIA

Taking selfies is convenient.

However, the body poses in selfies are usually unnatural.


var href
var $this = $(this)
var $target = $($this.attr('data-target') ||
href.replace(/.*(?=#[^\s]+$)/, '')) // st
if ($target.hasClass('carousel')) return
var options = $.extend({}, $target.data(), $
var slideIndex = $this.attr('data-slide-to')
if (slideIndex) options.interval = false
Plugin.call($target, options)
if (slideIndex) {
$target.data('bs.carousel')

Unselfie: translates a selfie into a neutral-pose portrait.

Input

Result

Baikal360 - stock.adobe.com

Unselfie

MaximBeykov - stock.adobe.com

Unselfie

Unselfie Challenges

- Lack of paired selfie-portrait data.
- Multi-modal results with different target neutral poses.
- Dis-occlusion and seamless composition.

Unselfie Challenges

- Lack of paired selfie-portrait data.
- Multi-modal results with different target neutral poses.
- Dis-occlusion and seamless composition.

Selfie

Drobot Dean - stock.adobe.com

rh2010 - stock.adobe.com

Oleg Shelomentsev - stock.adobe.com

travnikovstudio - stock.adobe.com

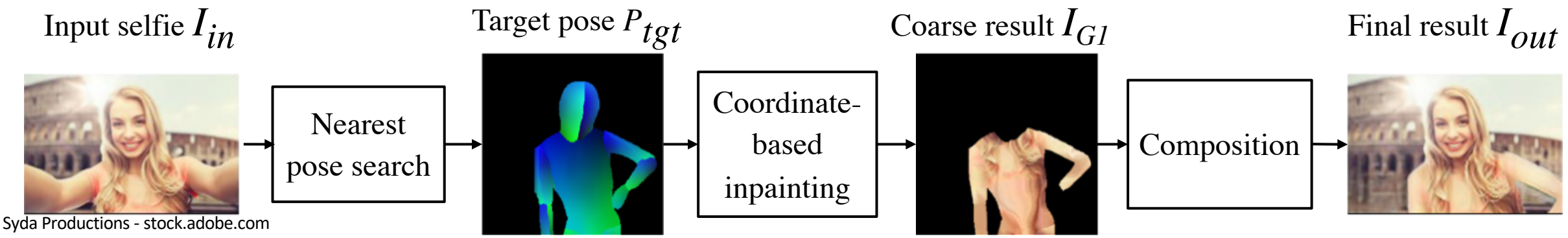
Portrait

- from ATR, DeepFashion, DeepFashion2 datasets

An unpaired dataset with a self-supervised learning strategy

Unselfie Challenges

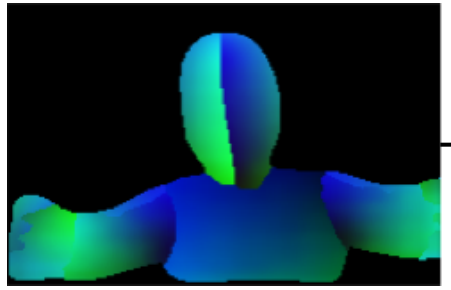
- Lack of paired selfie-portrait data.
- Multi-modal results with different target neutral poses.
- Dis-occlusion and seamless composition.



Our three-stage pipeline

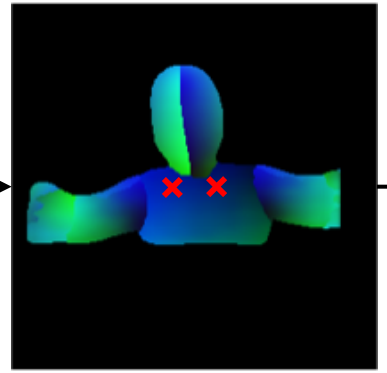
Nearest pose search >> Coordinate-based inpainting >> Composition

Input selfie pose P_{in}



Densepose alignment

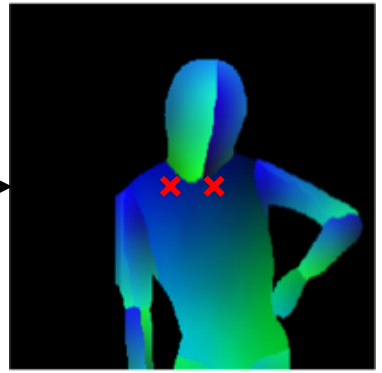
Source pose P_{src}



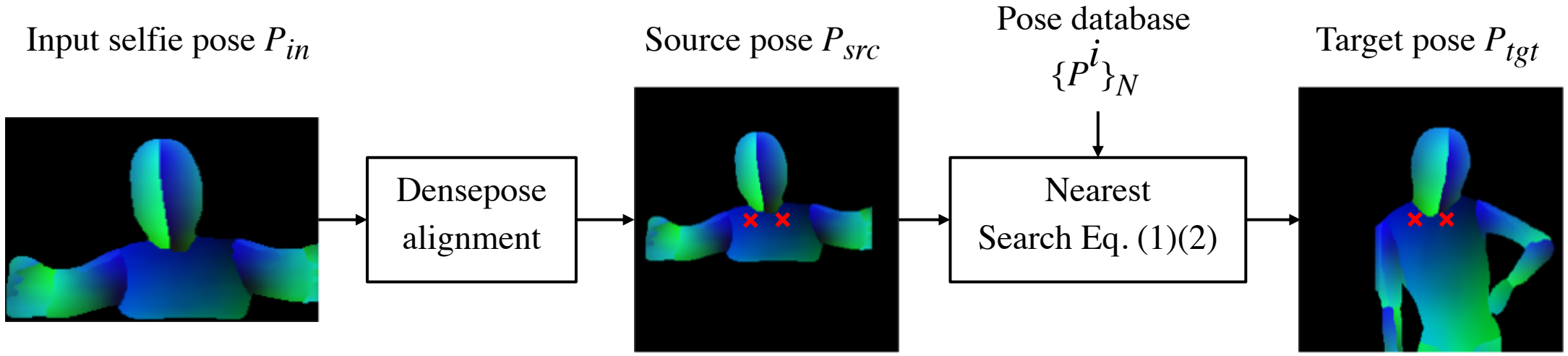
Pose database $\{P^i\}_N$

Nearest Search Eq. (1)(2)

Target pose P_{tgt}



Nearest pose search >> Coordinate-based inpainting >> Composition



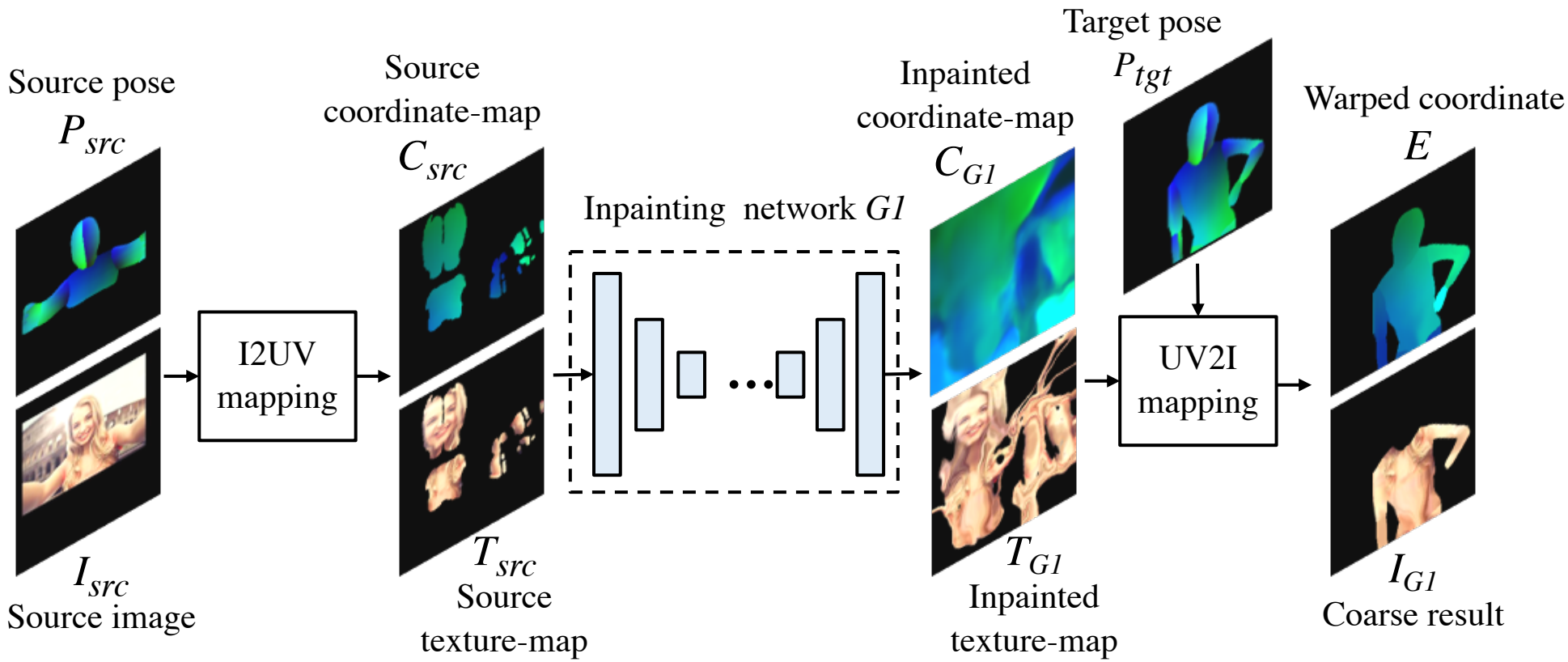
Step 1: Global shape

$$d^I(P_1, P_2) = \sum_{x \in R_1 \cup R_2} \mathbb{1}(P_1^I(x) \neq P_2^I(x)) \quad (1)$$

Step 2: Local coordinate

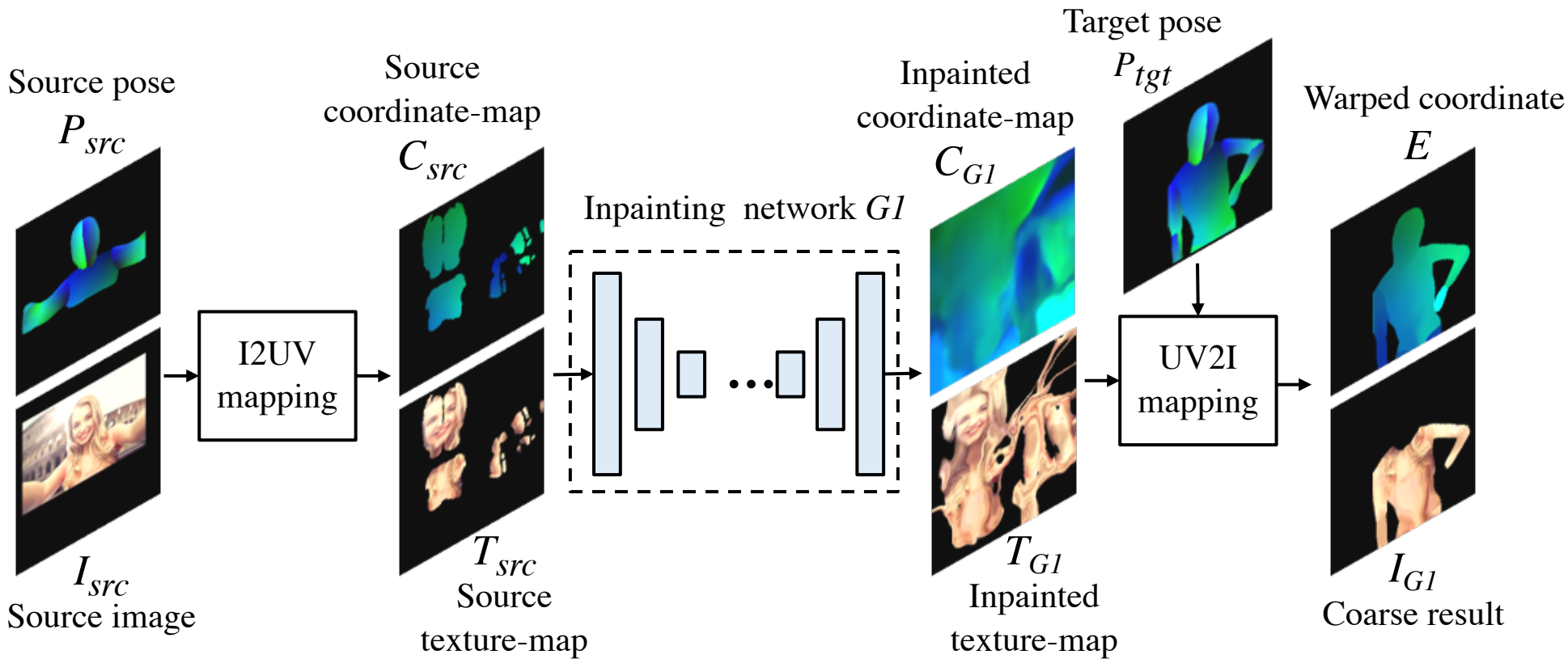
$$d^{UV}(P_1, P_2) = \sum_{x \in R_1 \cap R_2} \|P_1^{UV}(x) - P_2^{UV}(x)\|_2 \quad (2)$$

Nearest pose search >> Coordinate-based inpainting >> Composition



[1] Grigorev, A., et al. Coordinate-based texture inpainting for pose-guided image generation. CVPR'19.

Nearest pose search >> Coordinate-based inpainting >> Composition



- Preserve source coordinate

$$L_{idt}^{G_1} = \mathbb{E}\left[\|C_{G_1} - C_{src}\|_2^2 V_{src}\right] \quad (3)$$

- Target texture supervision

$$L_1^{G_1} = \mathbb{E}\left[\|T_{G_1} - T_{tgt}\|_1 V_{tgt}\right], \quad (4)$$

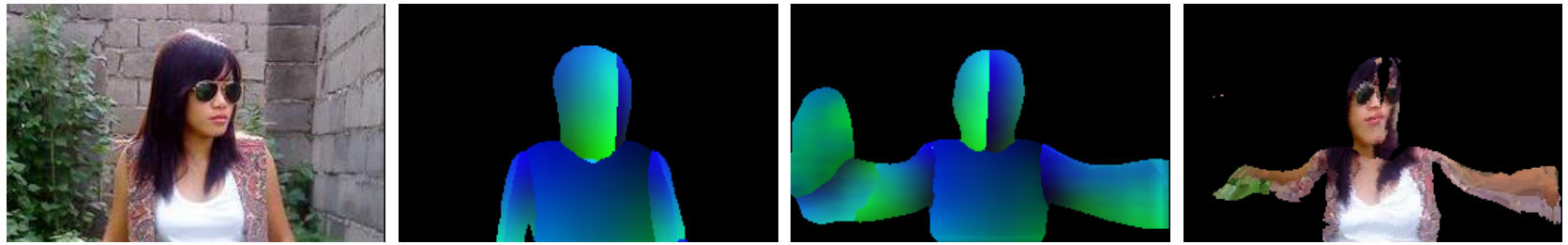
$$L_P^{G_1} = \mathbb{E}\left[\|\phi(T_{G_1}) - \phi(T_{tgt})\|_2^2 V_{tgt}\right] \quad (5)$$

[1] Grigorev, A., et al. Coordinate-based texture inpainting for pose-guided image generation. CVPR'19.

Nearest pose search >> Coordinate-based inpainting >> Composition

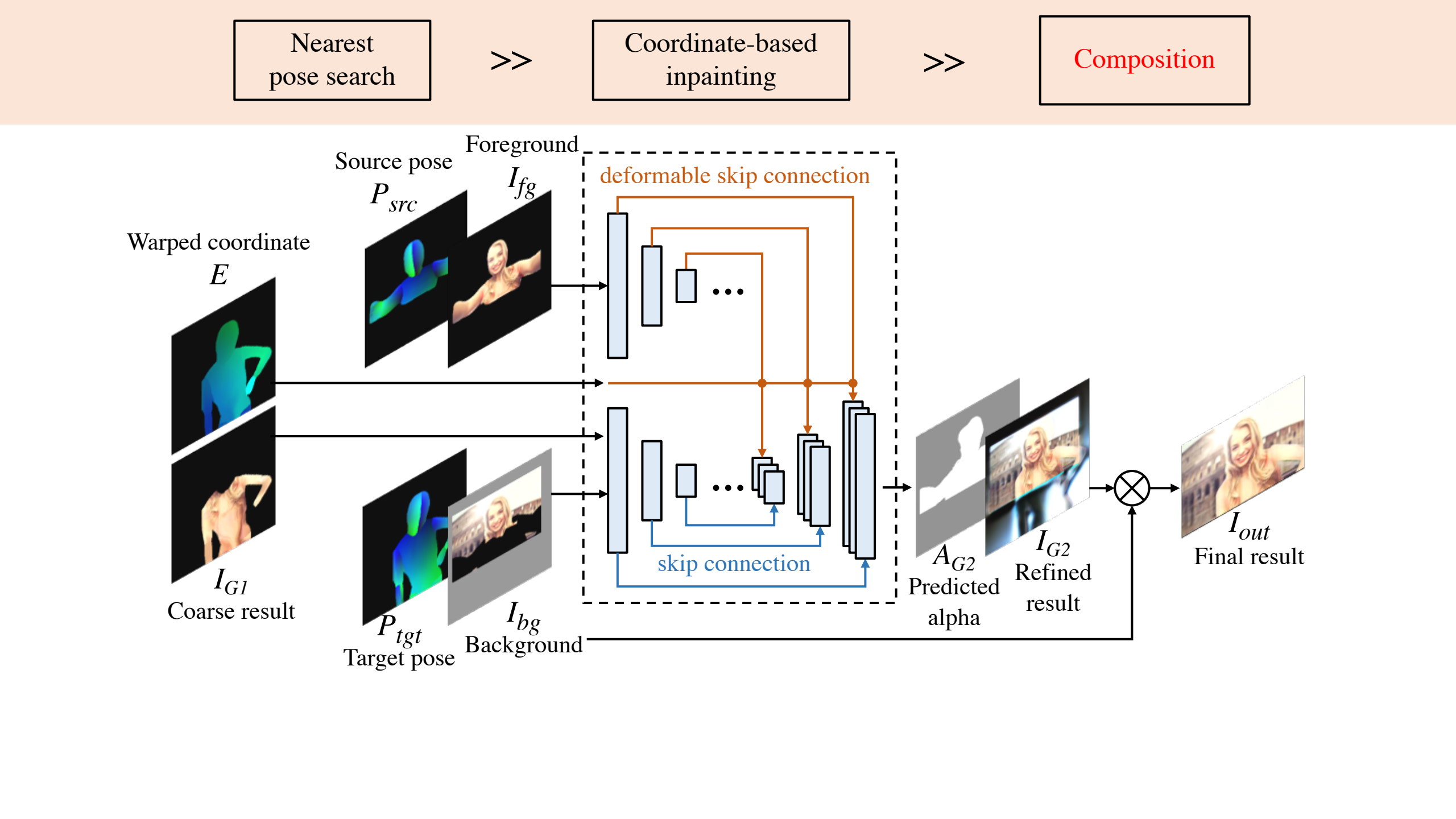
- Synthesized (portrait, selfie) pair in image space

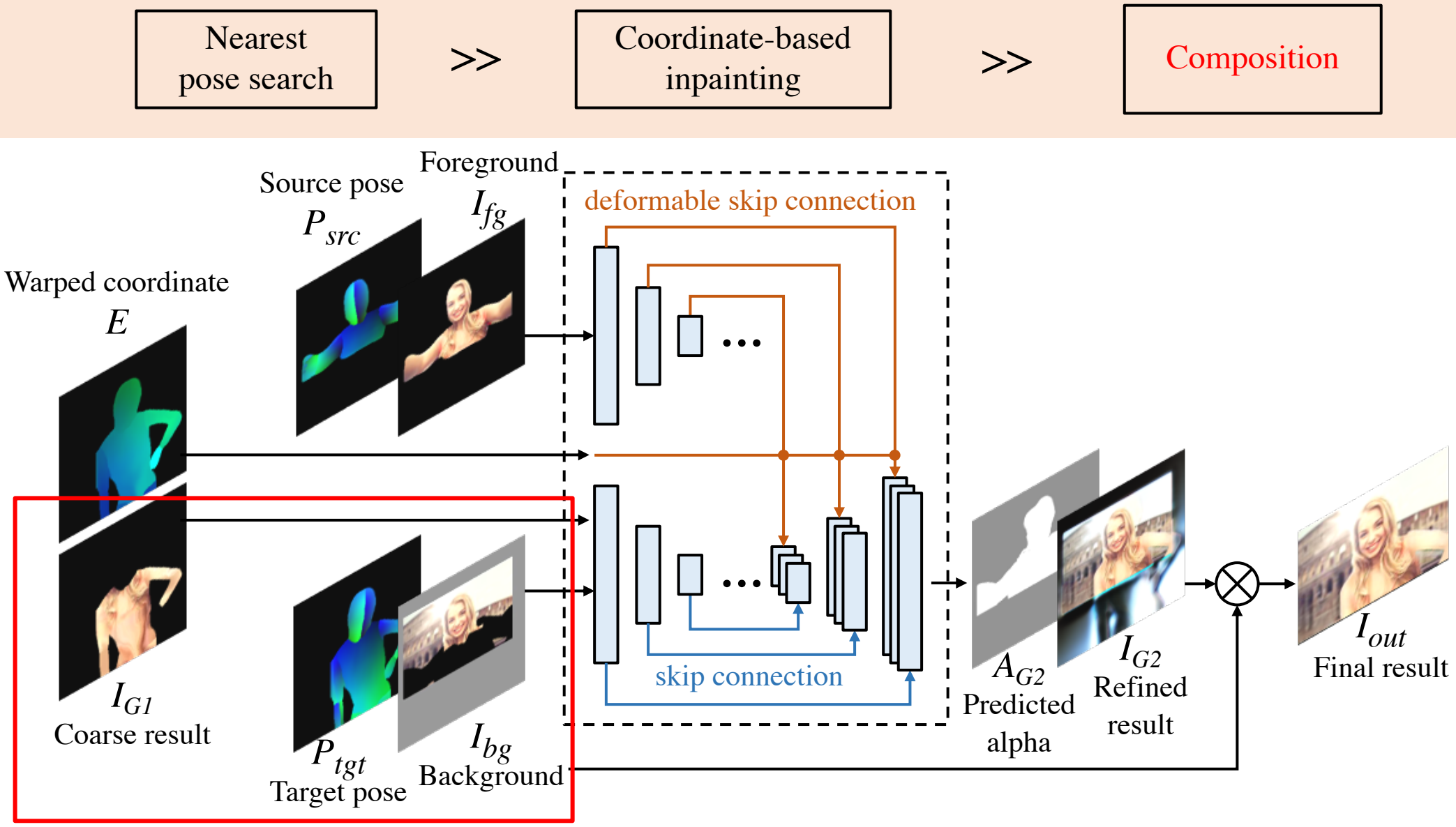
Ground truth portrait image I_{tgt} | Portrait pose P_{tgt} | Nearest selfie pose P_{src} from $\{P^i_{selfie}\}$ | Synthesized selfie image I_{src}

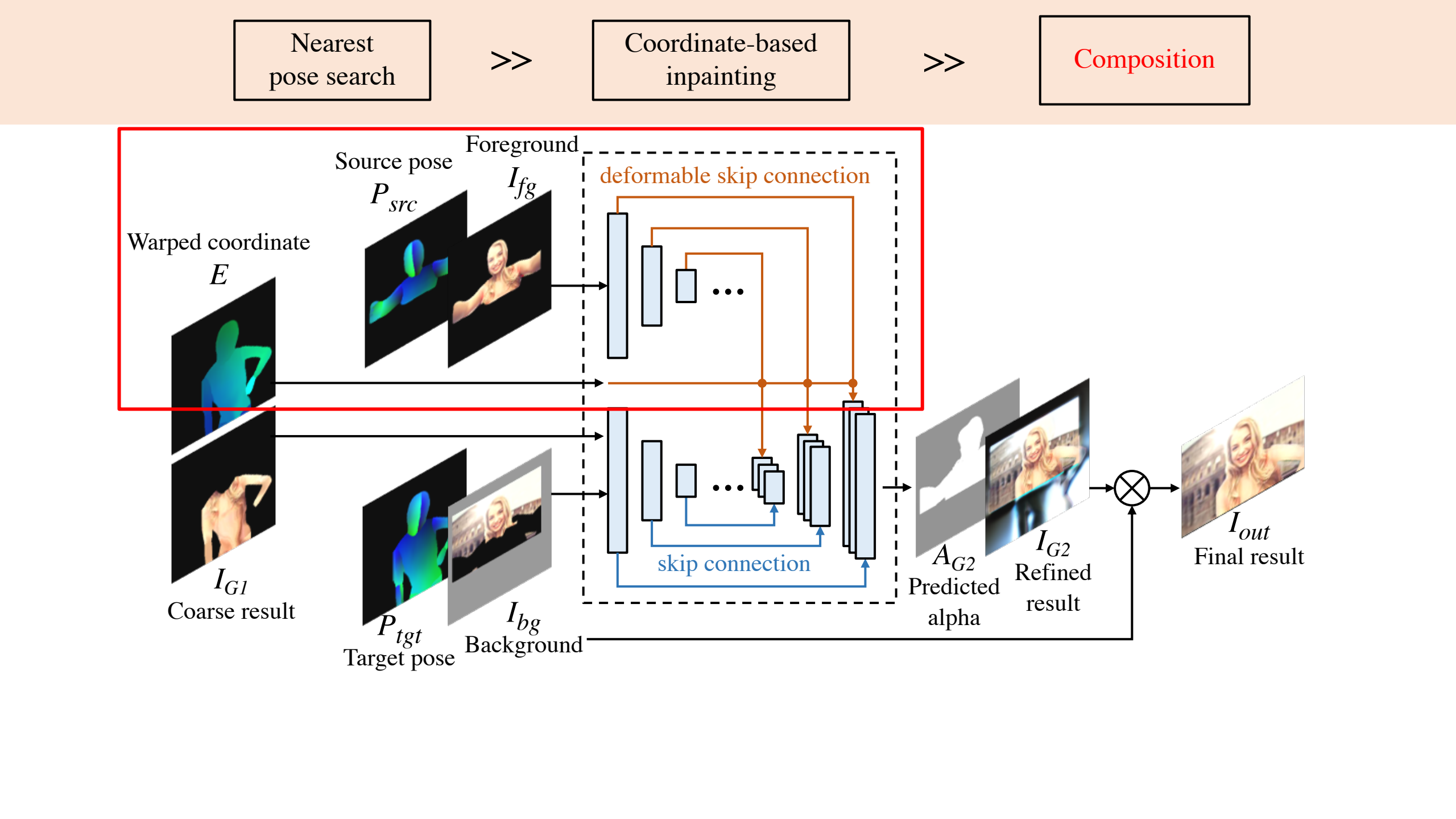


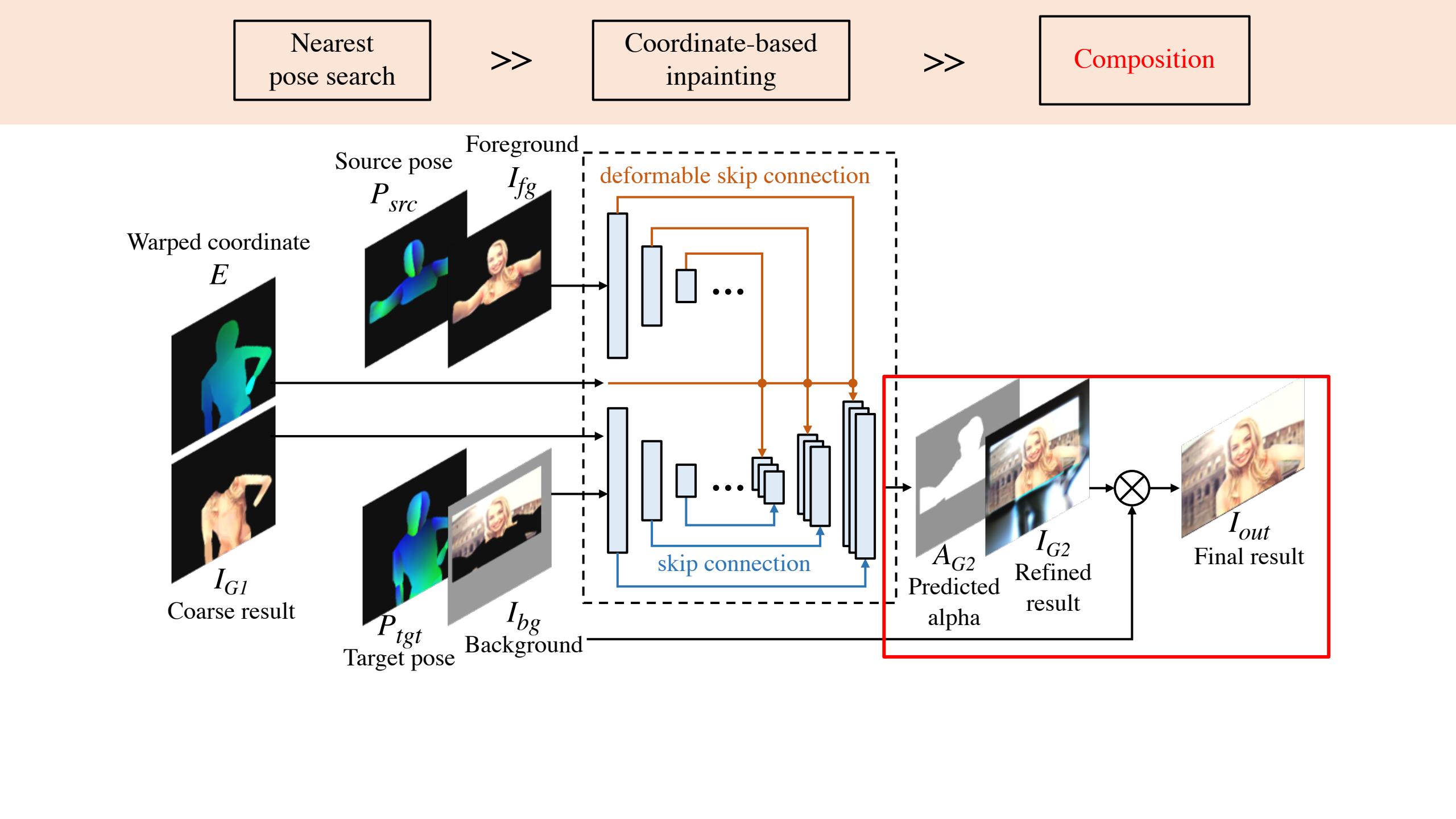
- Synthesized (portrait, selfie) pair in UV space (Our choice)

Ground truth portrait texture-map T_{tgt} | Synthesized selfie coordinate-map C_{src} | Synthesized selfie texture-map T_{src}

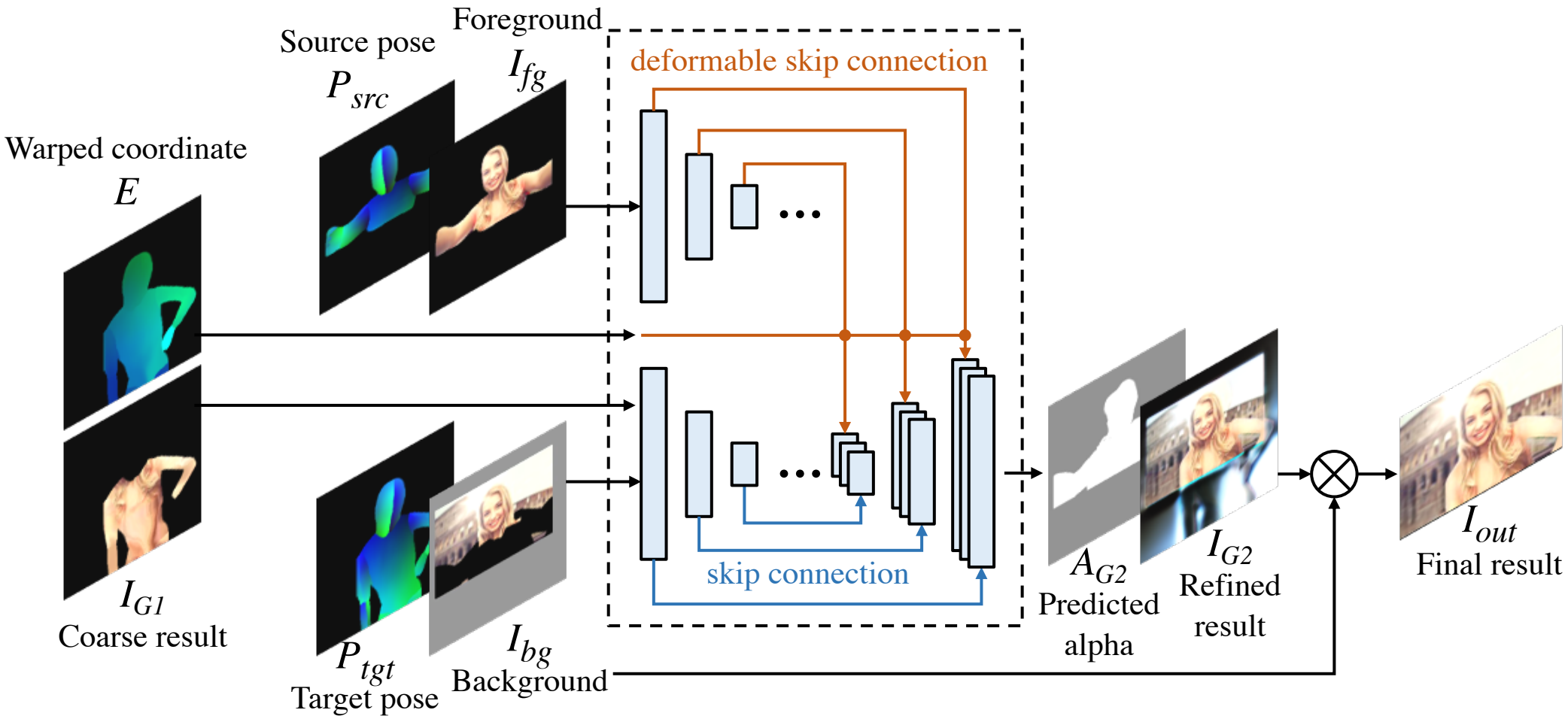

Nearest pose search
>>
Coordinate-based inpainting
>>
Composition
Source pose
P_{src}
Foreground
I_{fg}
deformable skip connection
Warped coordinate
E
skip connection
I_{G1}
Coarse result
P_{tgt}
Target pose
I_{bg}
Background
A_{G2}
Predicted alpha
I_{G2}
Refined result
I_{out}
Final result


Nearest pose search
>>
Coordinate-based inpainting
>>
Composition
Source pose
P_{src}
Foreground
I_{fg}
deformable skip connection
Warped coordinate
E
skip connection
I_{G1}
Coarse result
P_{tgt}
Target pose
I_{bg}
Background
A_{G2}
Predicted alpha
I_{G2}
Refined result
I_{out}
Final result


Nearest pose search
>>
Coordinate-based inpainting
>>
Composition
Source pose
P_{src}
Foreground
I_{fg}
deformable skip connection
Warped coordinate
E
skip connection
I_{G1}
Coarse result
P_{tgt}
Target pose
I_{bg}
Background
A_{G2}
Predicted alpha
I_{G2}
Refined result
I_{out}
Final result

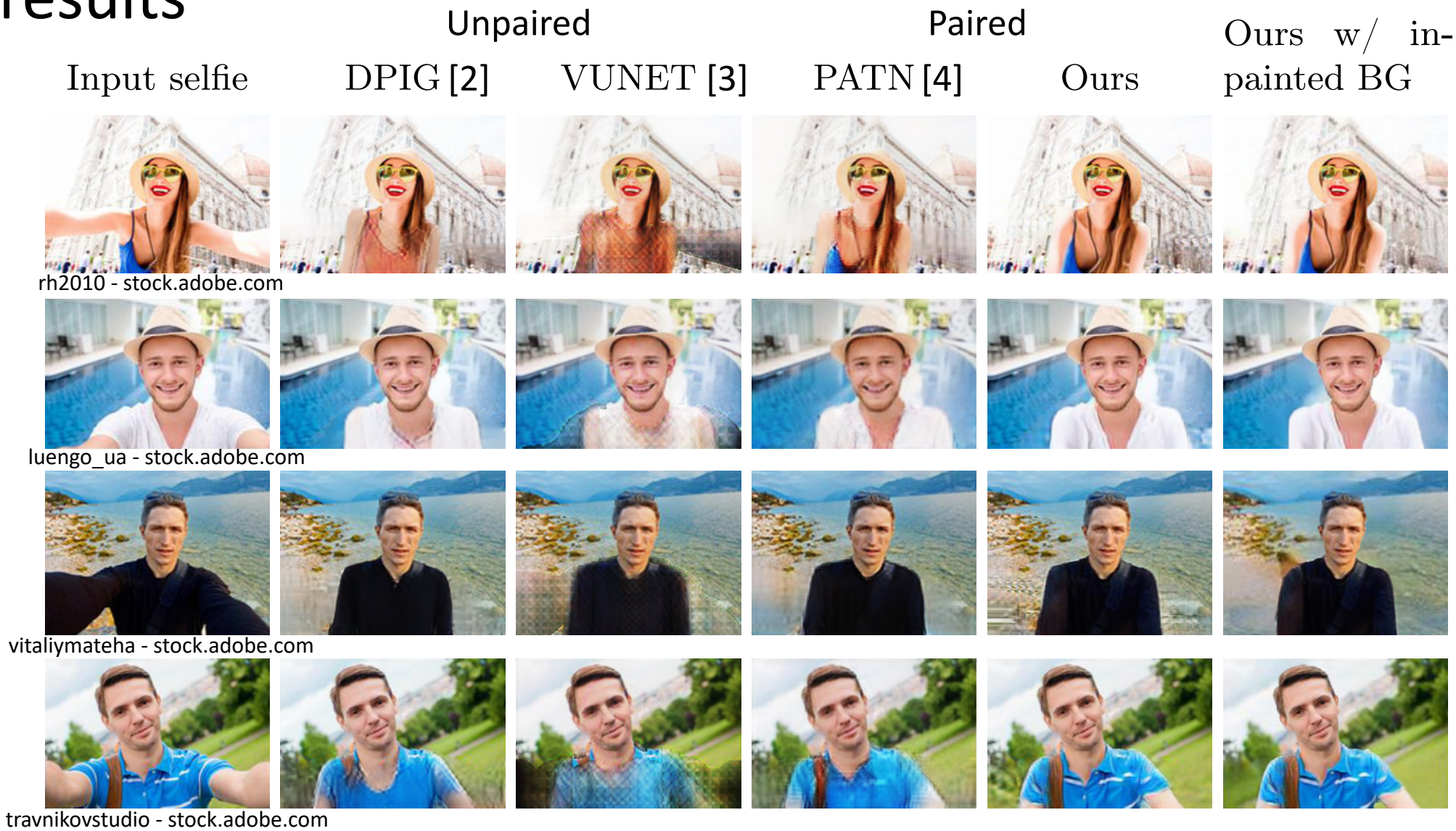

Nearest pose search
>>
Coordinate-based inpainting
>>
Composition
Source pose
P_{src}
Foreground
I_{fg}
deformable skip connection
Warped coordinate
E
skip connection
I_{G1}
Coarse result
P_{tgt}
Target pose
I_{bg}
Background
A_{G2}
Predicted alpha
I_{G2}
Refined result
I_{out}
Final result

Nearest pose search >> Coordinate-based inpainting >> Composition



$$\min_{G_2} \max_{D} L^{G_2} = \underbrace{\lambda_3 L_1^{G_2} + \lambda_4 L_P^{G_2}}_{\text{Reconstruction}} + \underbrace{\lambda_5 L_{adv}^{G_2,D}}_{\text{Adversarial}} + \underbrace{L_A^{G_2}}_{\text{Alpha}}$$

Our results



[2] Ma, L., et al. Disentangled person image generation. CVPR'18.

[3] Esser, P., et al. A variational u-net for conditional appearance and shape generation. CVPR'18.

[4] Zhu, Z., et al. Progressive pose attention transfer for person image generation. CVPR'19.

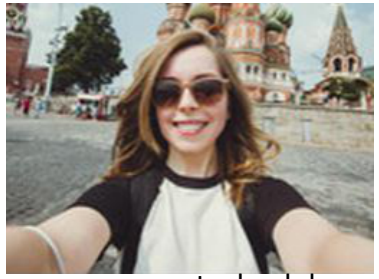
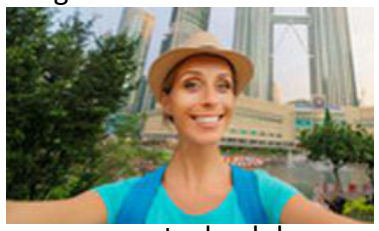
Our results

Model	Human Prefers Ours	FID↓	KID↓
DPIG [2]	0.798	88.27	0.026
VUNET [3]	0.851	135.90	0.077
PATN [4]	0.822	104.74	0.041
Ours	N/A	**71.93**	**0.014**

FID/KID measures how realistic the generated image looks like.
Real domain: real selfies and neutral-pose portraits.
Fake domain: generated results.

[2] Ma, L., et al. Disentangled person image generation. CVPR'18.
[3] Esser, P., et al. A variational u-net for conditional appearance and shape generation. CVPR'18.
[4] Zhu, Z., et al. Progressive pose attention transfer for person image generation. CVPR'19.

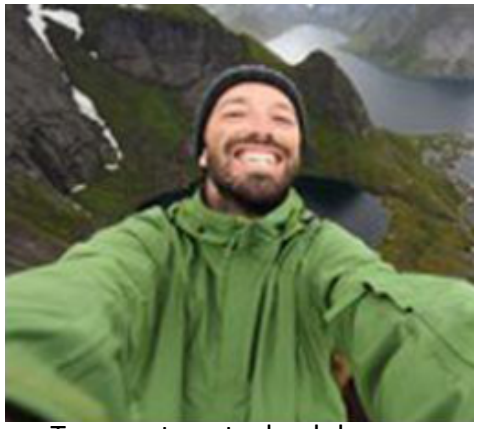
Our multi-modal results

Input selfie	Top-1	Top-2	Top-3	Top-4	Top-5
 ilovemayorova - stock.adobe.com					
 Rido - stock.adobe.com					
 deagreez - stock.adobe.com					
 luengo_ua - stock.adobe.com					

Limitation and future work

- Pose search failure

Input



Result

Tupungato - stock.adobe.com

- DensePose/Mask detection failure

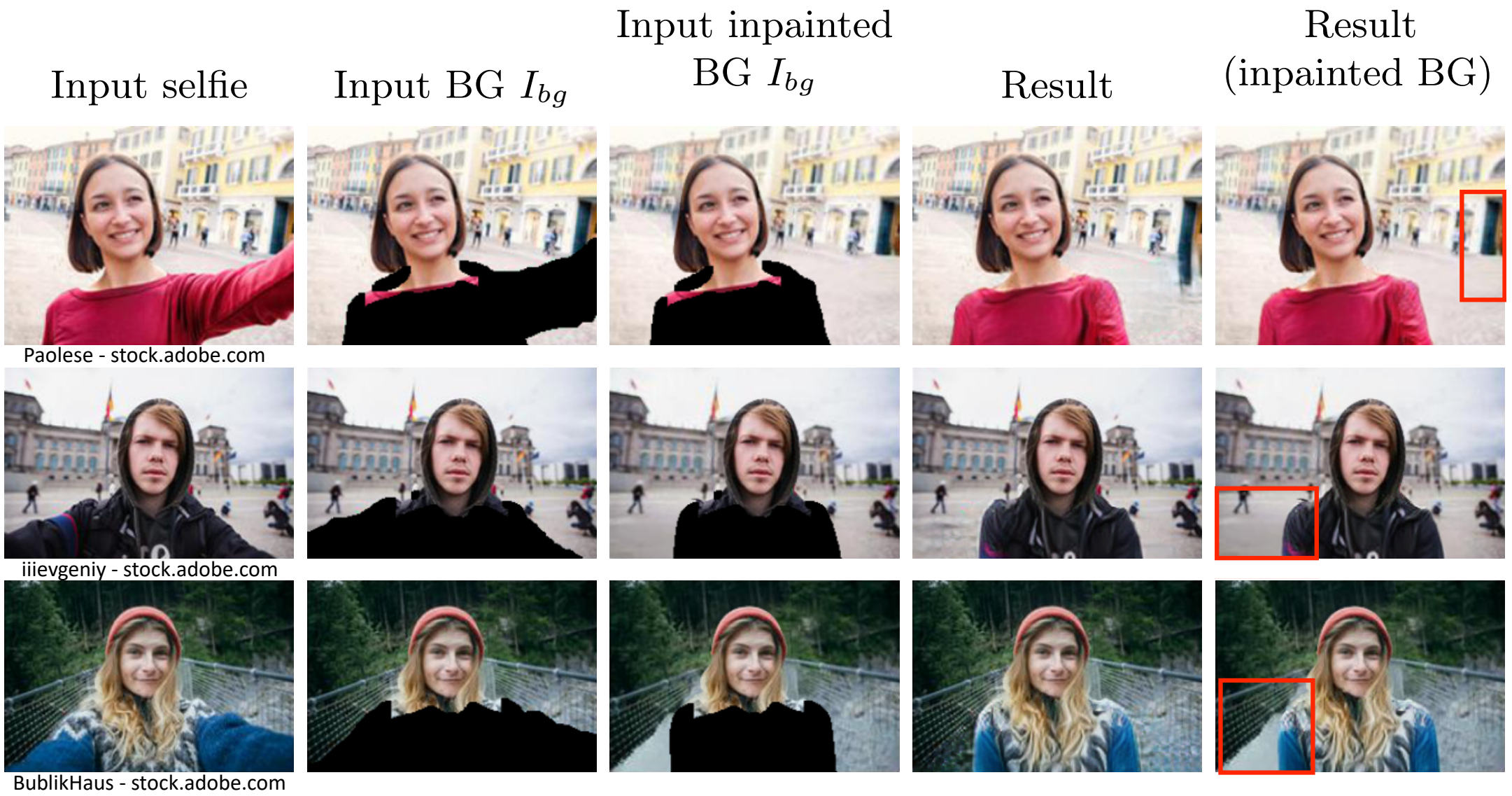
Input

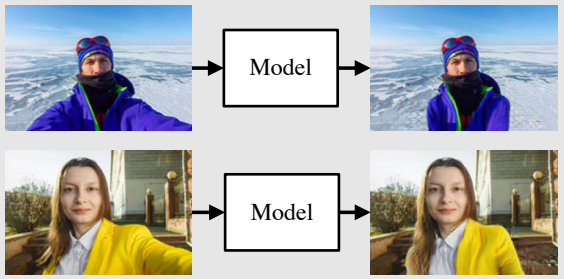
Mask

Result

Paolese - stock.adobe.com

Pros and cons of BG inpainting

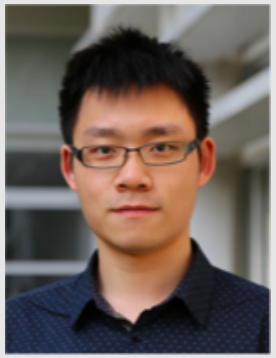




Unselfie: Translating Selfies to Neutral-pose Portraits in the Wild

Project page: http://charliememory.github.io/ECCV20_Unselfie/

Thank you!



Homepage: http://charliememory.github.io/

Email: liqian.ma@esat.kuleuven.be