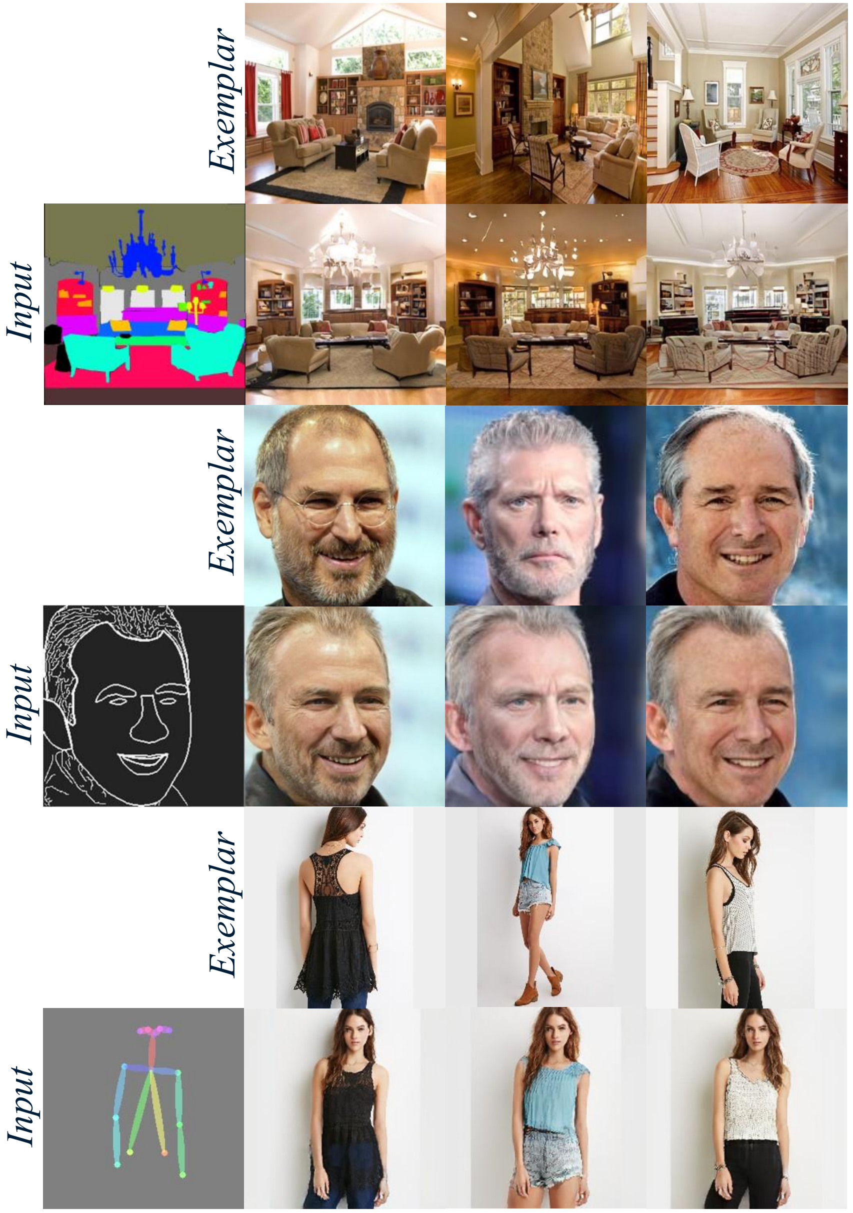
Cross-domain Correspondence Learning for Exemplar-based Image Translation (CoCosNet)

CVPR 2020 oral

Pan Zhang, Bo Zhang, Dong Chen, Lu Yuan, Fang Wen



Prior image translation methods

- Lack of fine-grain controllability

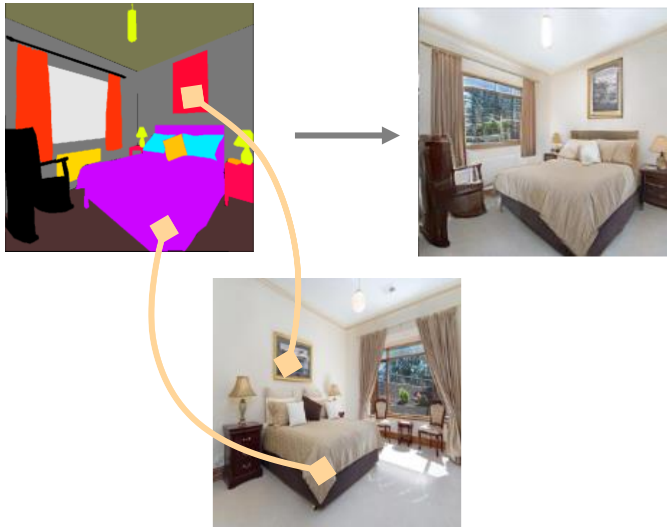
Edge → shoes

(MUNIT, ECCV 2018)

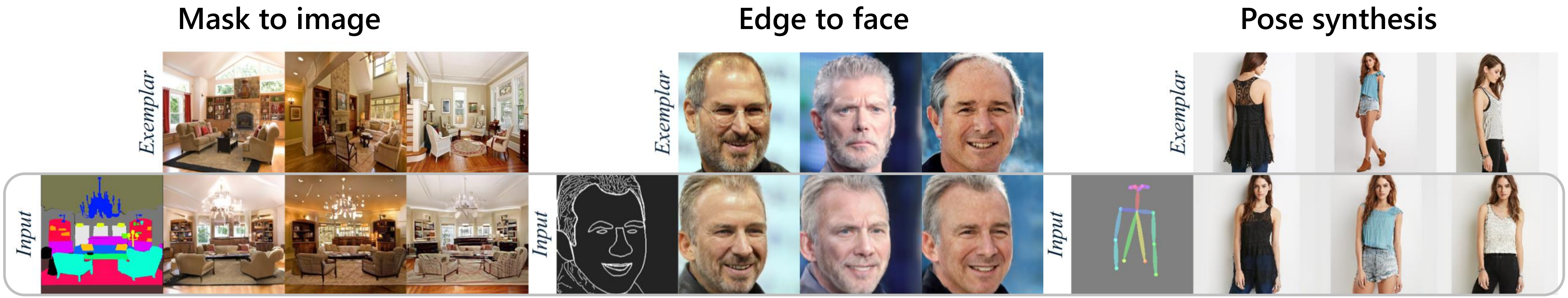
- Significant artifacts in complex scenes

SPADE, CVPR 2019

Proposed exemplar-based solution

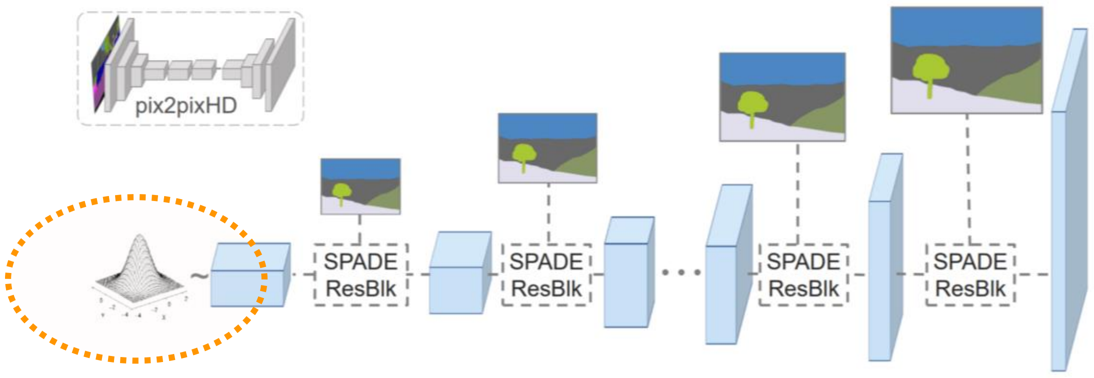


- Instance-level style control
- Significantly improve the image quality
- General translation solution

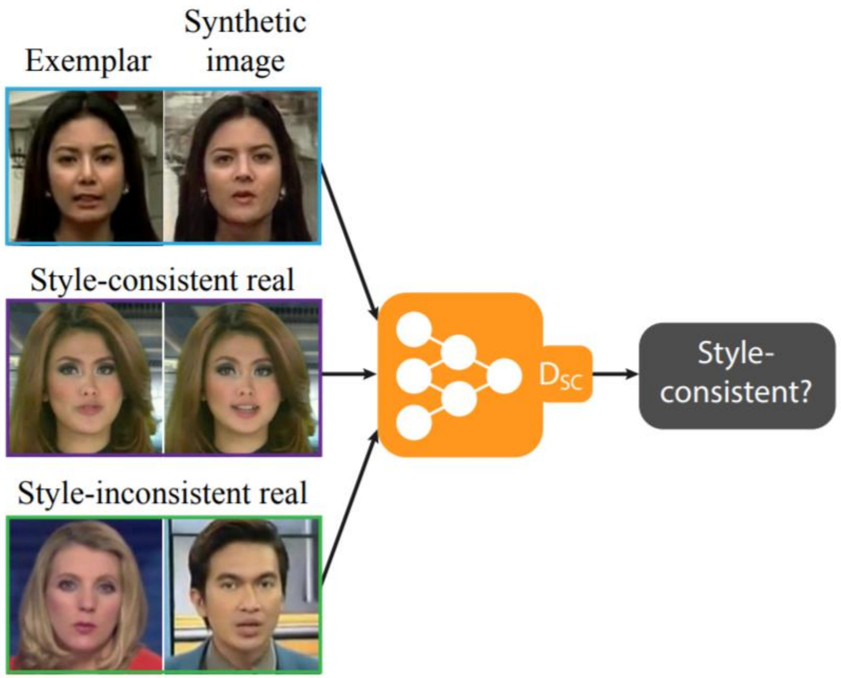


Relation with exemplar-based methods

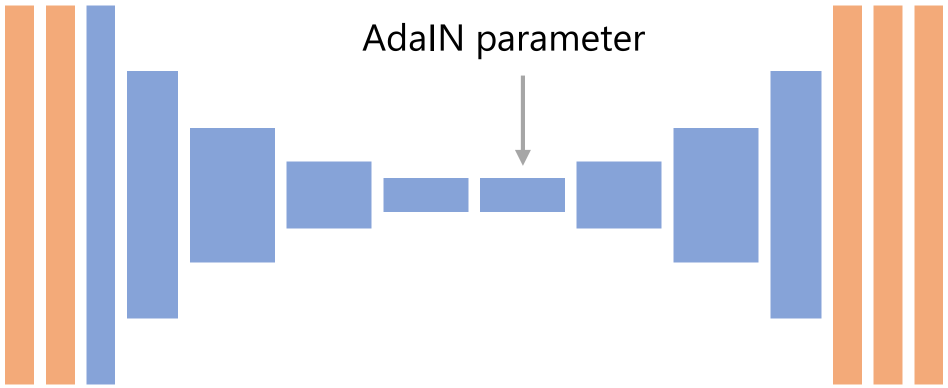
Embed style with latent space



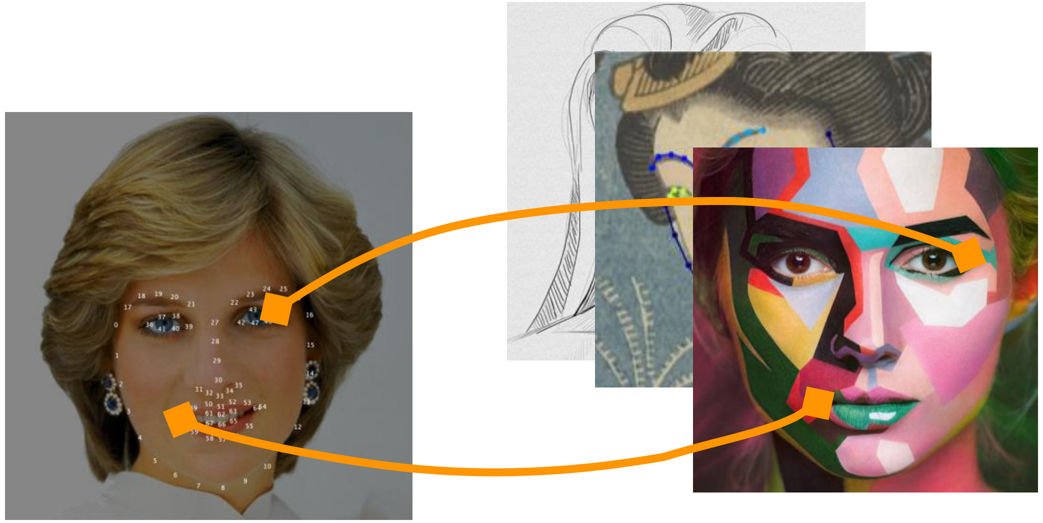
Style discriminator

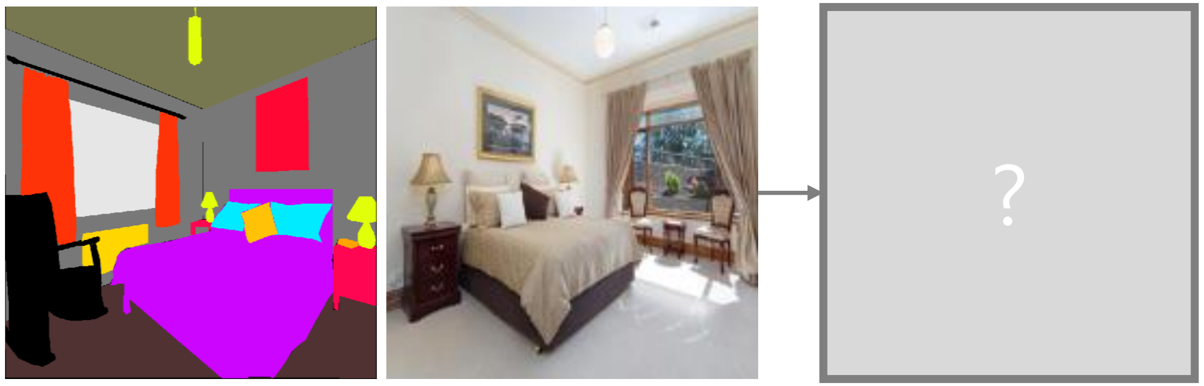


Style injection during AdaIN



Motivation



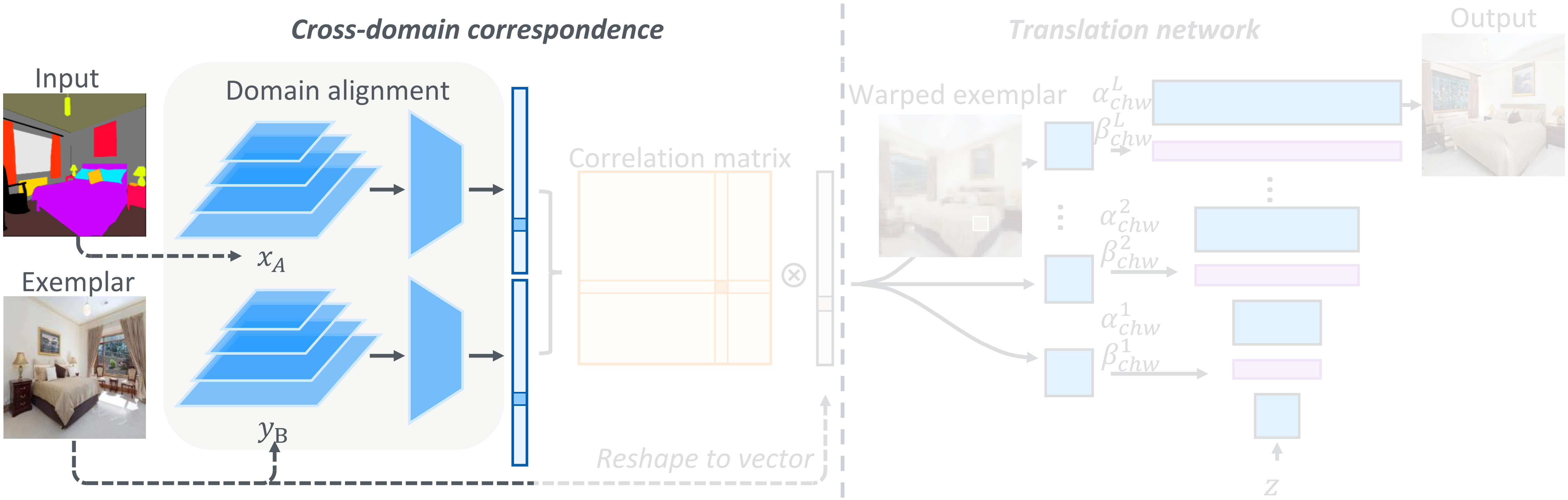


How to establish correspondence for heterogeneous images?

What is the desired translation output given an exemplar?

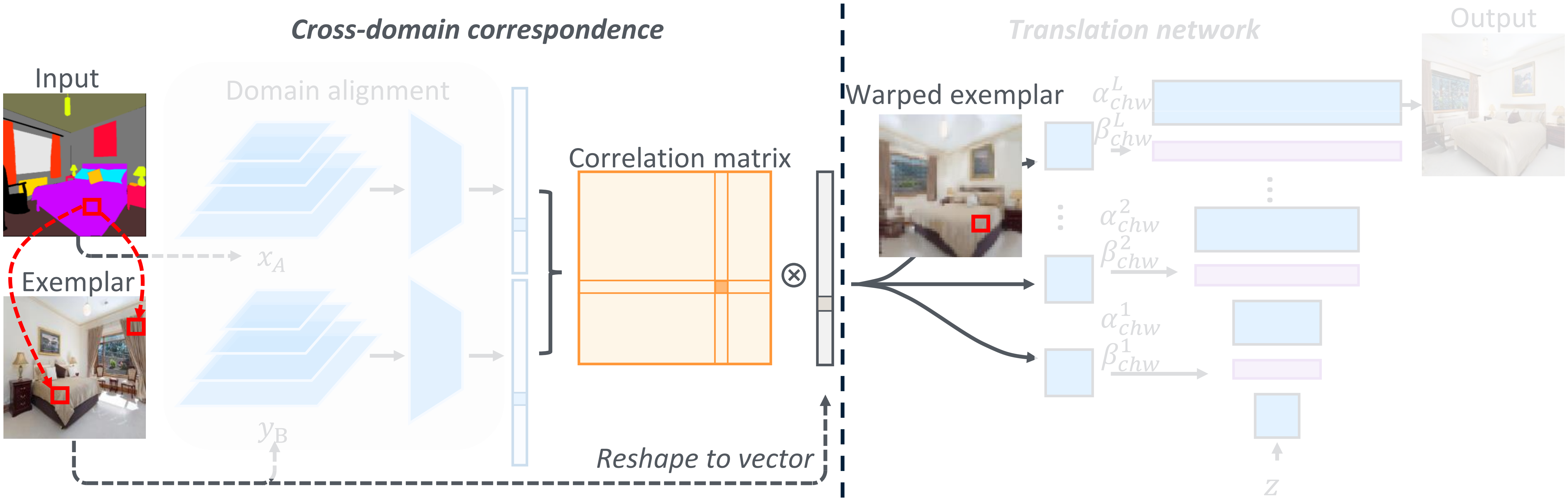
Facilitate each other

Framework



Cross-domain correspondence
Translation network
Output
Input
Domain alignment
Warped exemplar
Correlation matrix
x_A
Exemplar
y_B
Reshape to vector
α^L_{chw}
β^L_{chw}
α^2_{chw}
β^2_{chw}
α^1_{chw}
β^1_{chw}
z

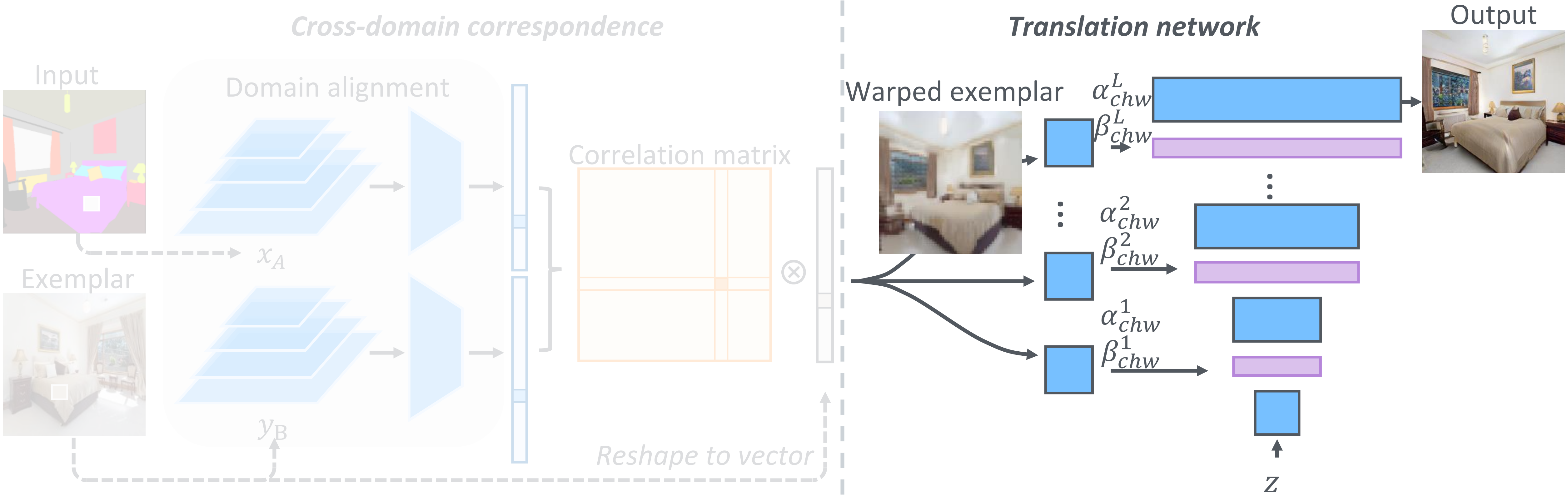
Framework



Cross-domain correspondence
Translation network
Output
Input
Domain alignment
x_A
Exemplar
y_B
Correlation matrix
Warped exemplar
Reshape to vector
α^L_{chw}
β^L_{chw}
α^2_{chw}
β^2_{chw}
α^1_{chw}
β^1_{chw}
z

Framework

$$\alpha^i_{h,w}(r_{y\to x}) \times \boxed{\frac{F^i_{c,h,w} - \mu^i_{h,w}}{\sigma^i_{h,w}}} + \beta^i_{h,w}(r_{y\to x}),$$



Loss

Translation network

Pseudo exemplar loss:

$$\mathcal{L}_{feat} = \sum_{l} \lambda_l \left\| \phi_l(\mathcal{G}(x_A, x'_B)) - \phi_l(x_B) \right\|_1 ,$$

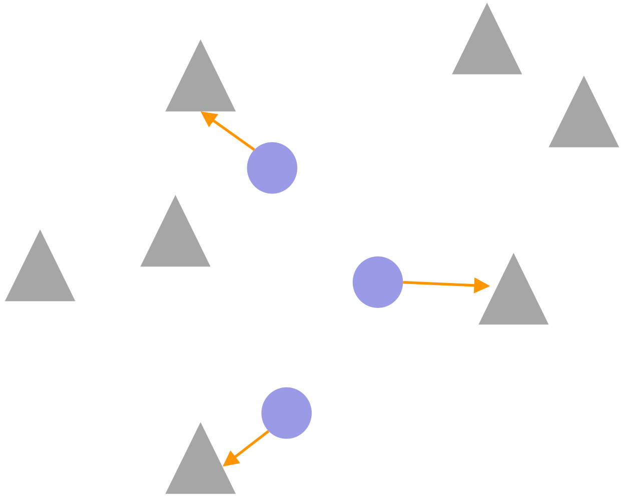
Pseudo exemplar pairs

Loss
Translation network

Contextual loss: let the output to mimic the appearance of the semantically corresponding patches of the exemplar.

$$\mathcal{L}_{context} = \sum_l \omega_l \left[-\log \left(\frac{1}{n_l} \sum_i \max_j A^l(\phi_i^l(\hat{x}_B), \phi_j^l(y_B)) \right) \right],$$

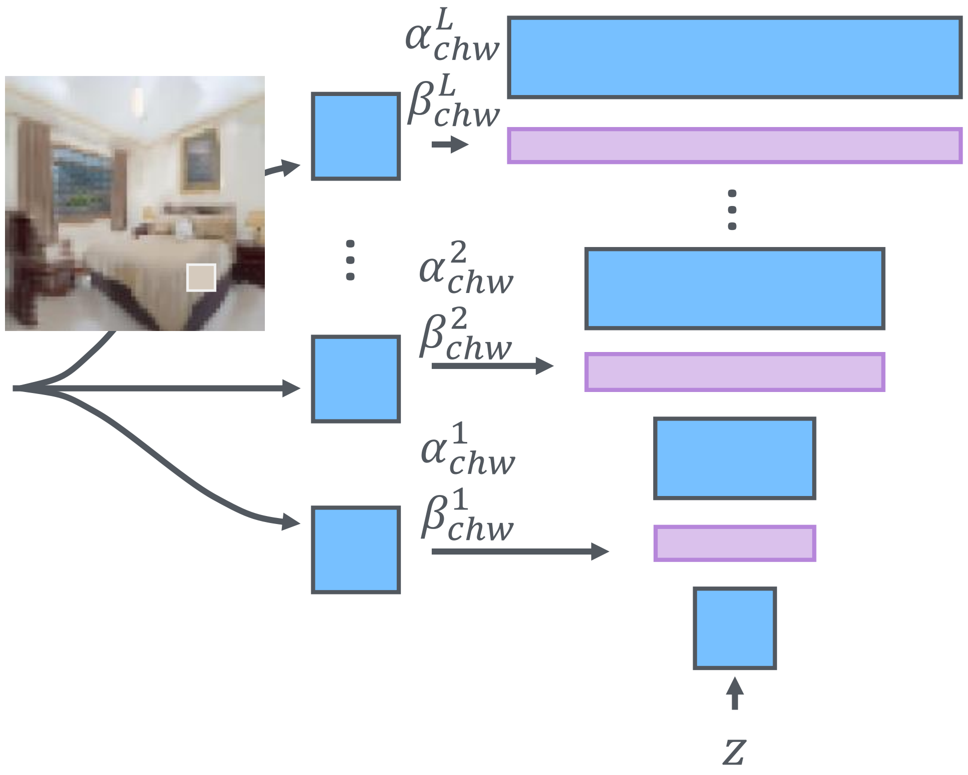
pairwise affinities in the VGG feature space



Loss

Translation network

Translation network



Perceptual loss: the output should maintain the semantics as the input

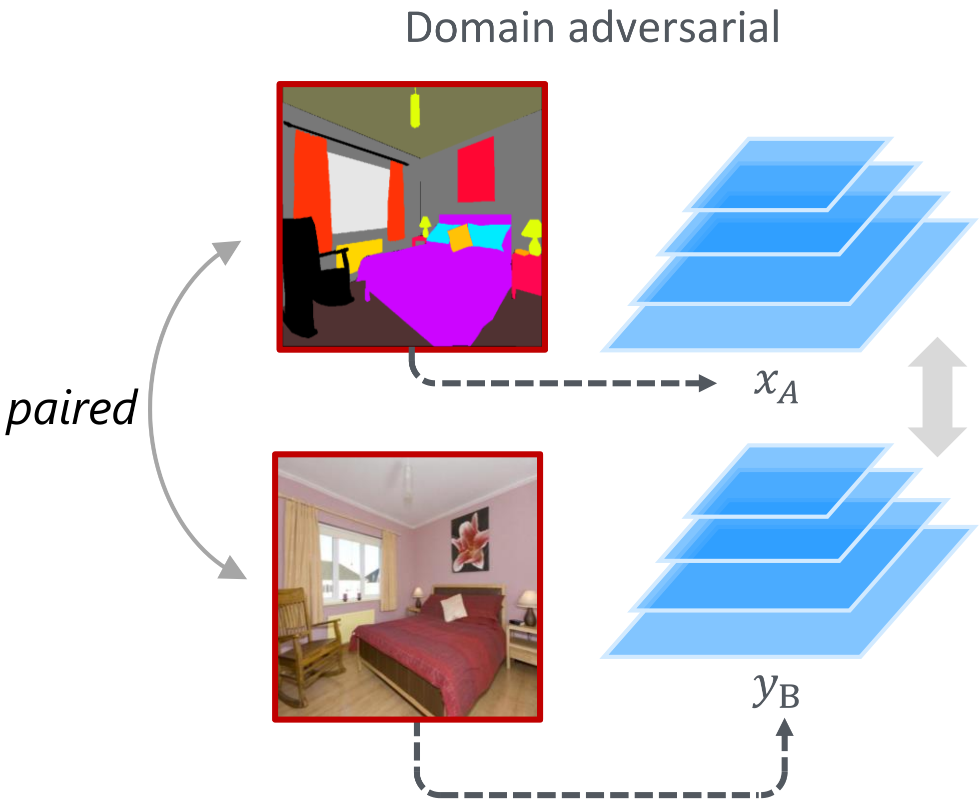
$$\mathcal{L}_{perc} = \left\| \phi_l(\hat{x}_B) - \phi_l(x_B) \right\|_1 .$$

Adversarial loss: make the output as realistic as possible

$$\mathcal{L}^{\mathcal{D}}_{adv} = -\mathbb{E}[h(\mathcal{D}(y_B))] - \mathbb{E}[h(-\mathcal{D}(\mathcal{G}(x_A, y_B)))]$$
$$\mathcal{L}^{\mathcal{G}}_{adv} = -\mathbb{E}[\mathcal{D}(\mathcal{G}(x_A, y_B))],$$

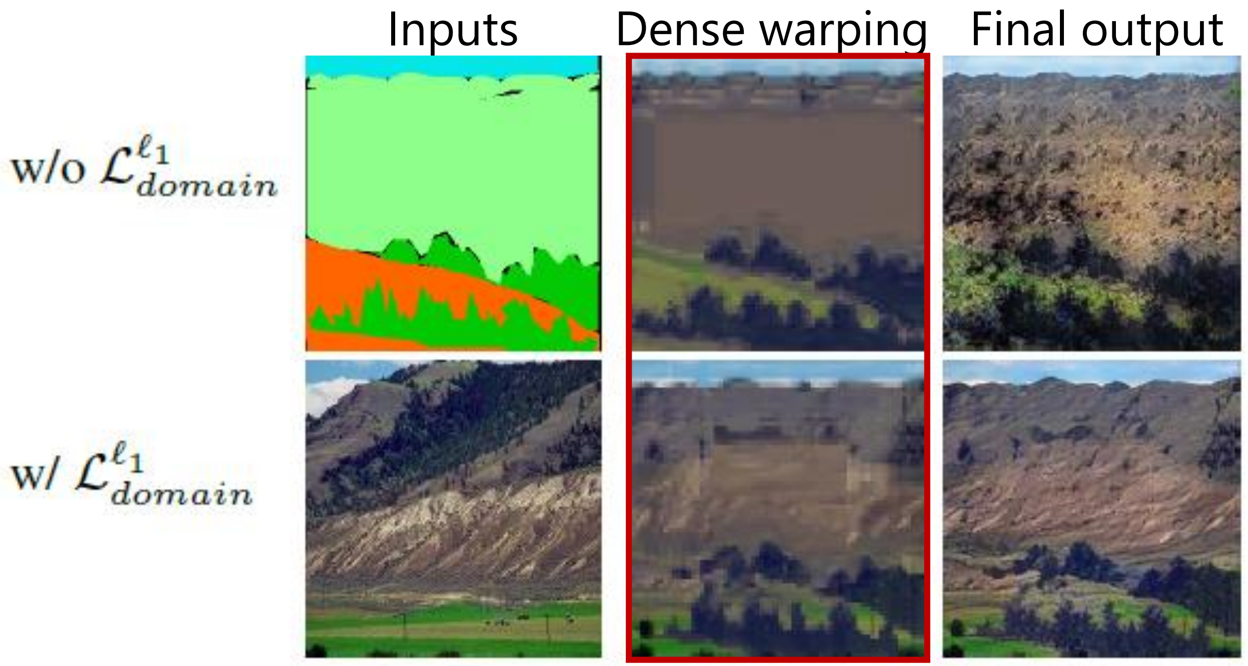
Loss

Cross-domain correspondence network



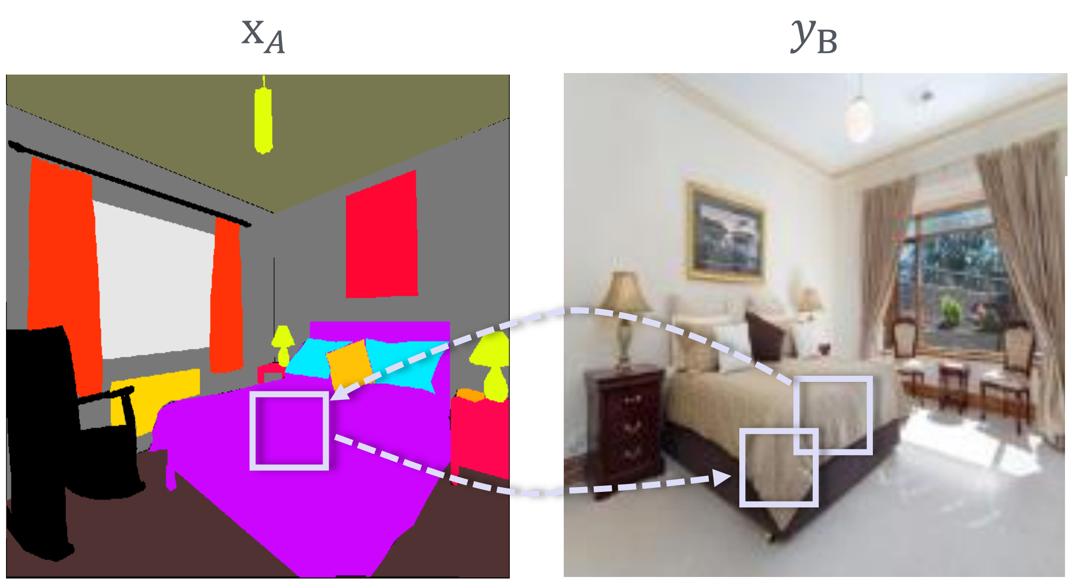
Domain alignment loss: the embeddings of inputs should lie in the same domain

$$\mathcal{L}_{domain}^{\ell_1} = \|\mathcal{F}_{A\to S}(x_A) - \mathcal{F}_{B\to S}(x_B)\|_1$$



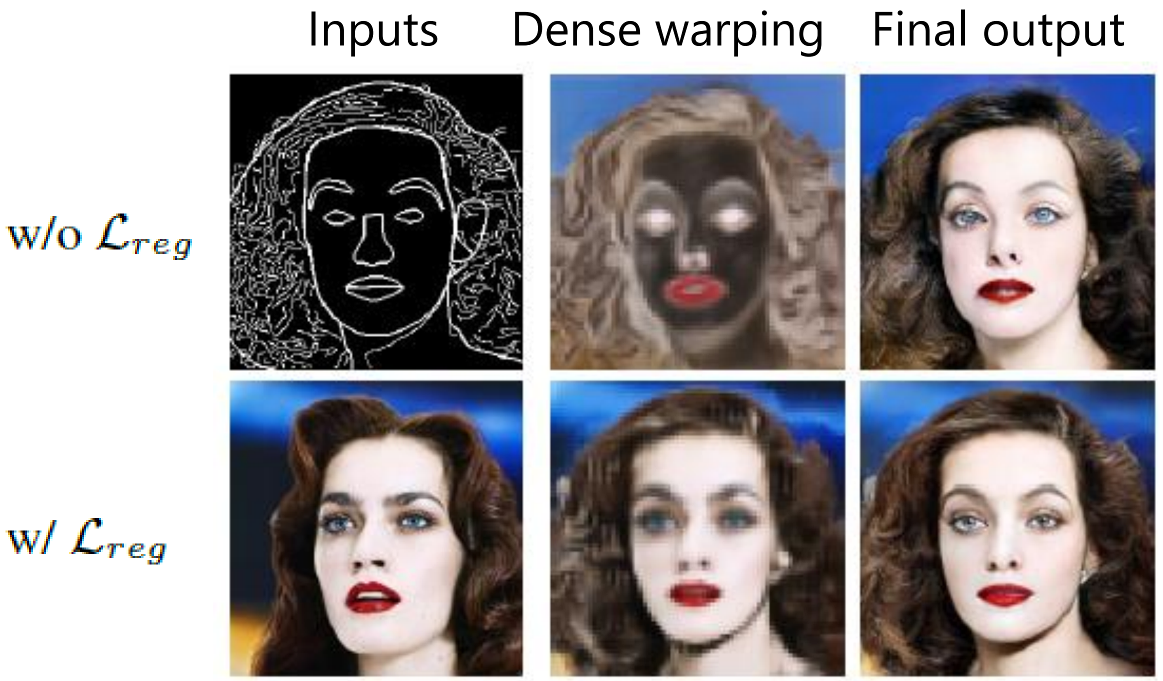
Loss

Cross-domain correspondence network



Cycle warping regularization:

$$\mathcal{L}_{reg} = \left\| r_{y \to x \to y} - y_B \right\|_1 ,$$

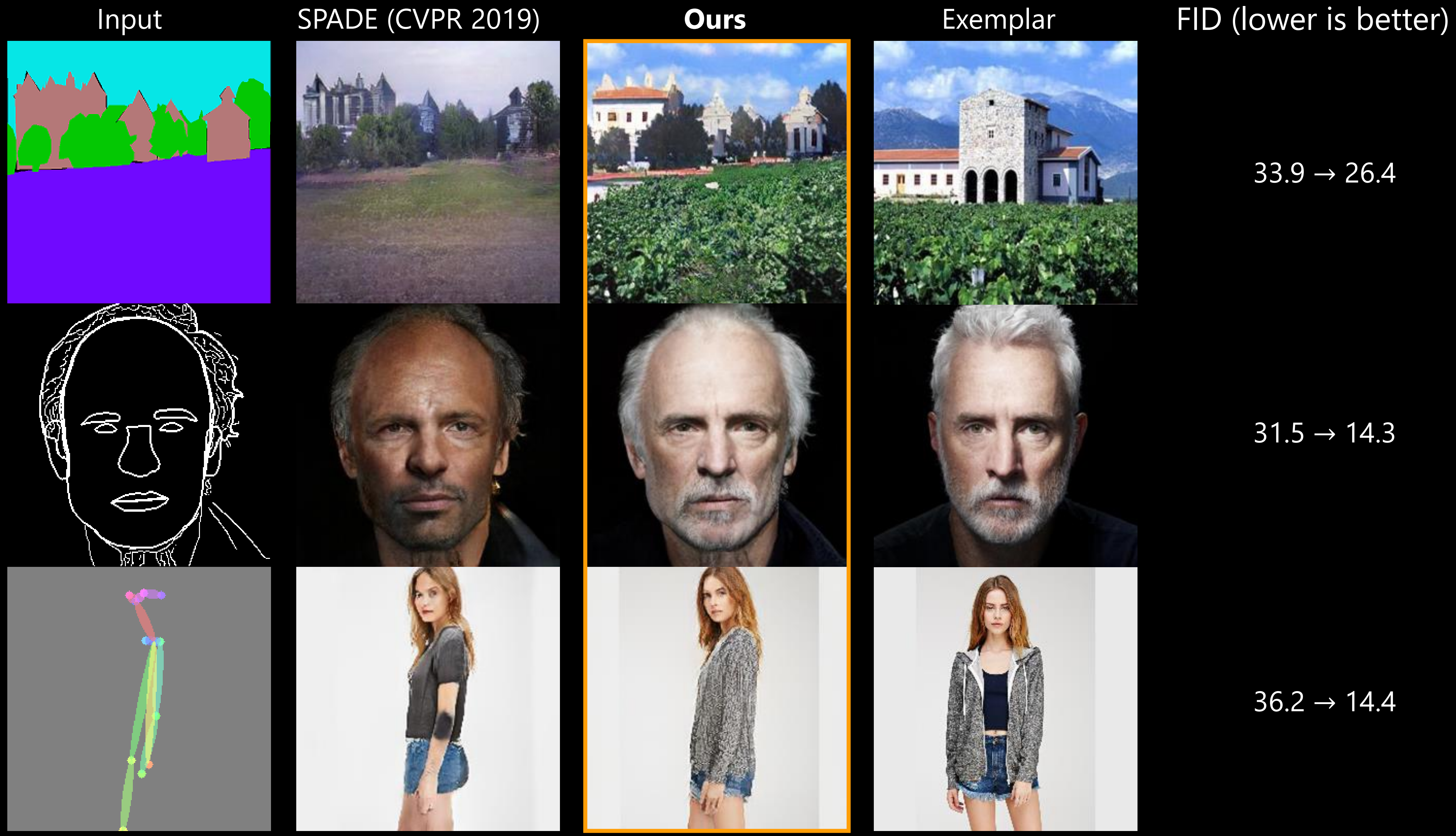


Total loss

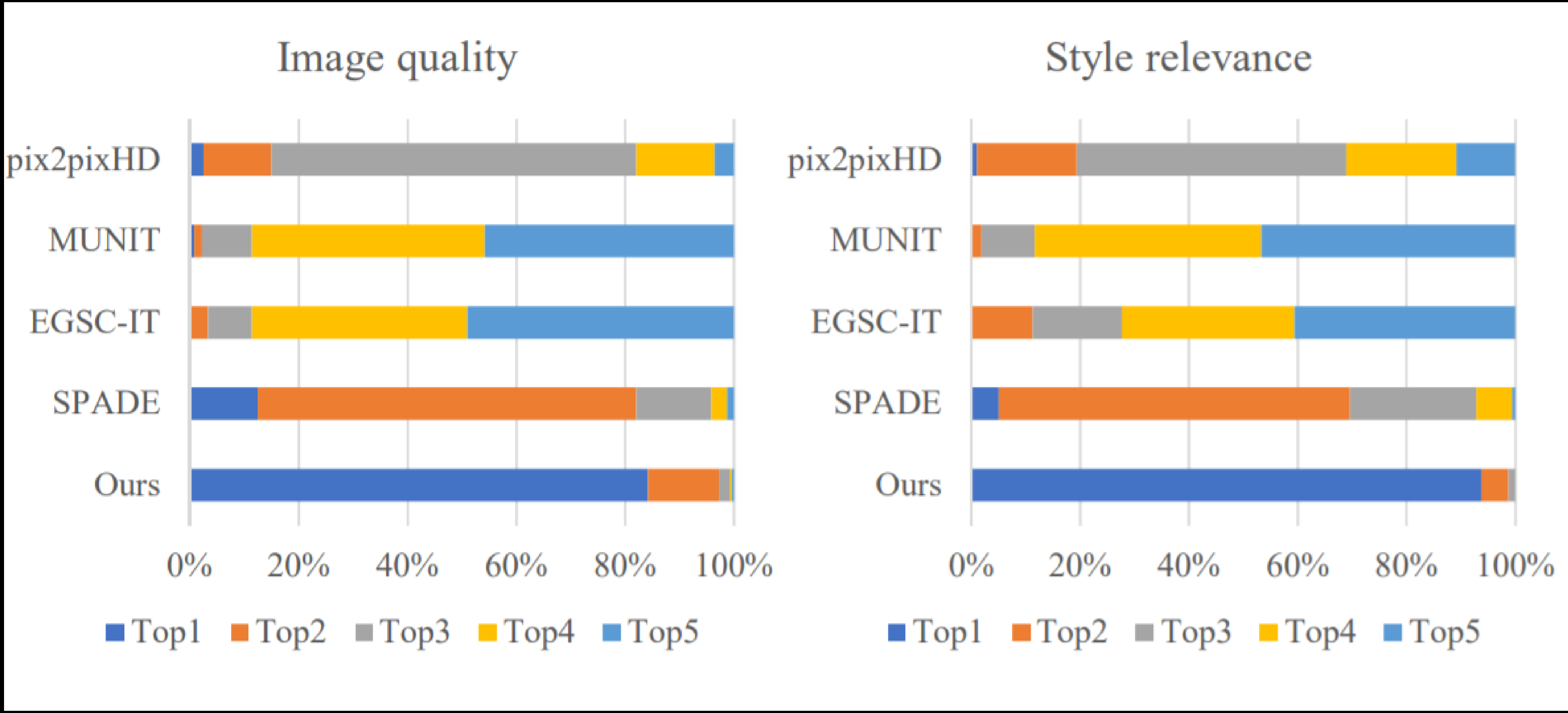
$$\mathcal{L}_\theta = \min_{\mathcal{F},\mathcal{T},\mathcal{G}} \max_{\mathcal{D}} \psi_1 \mathcal{L}_{feat} + \psi_2 \mathcal{L}_{perc} + \psi_3 \mathcal{L}_{context} + \psi_4 \mathcal{L}_{adv}^{\mathcal{G}} + \psi_5 \mathcal{L}_{domain}^{\ell_1} + \psi_6 \mathcal{L}_{reg},$$

- Pseudo exemplar pairs:
  - VGG feature matching
- Real exemplar pairs:
  - Perceptual loss
  - Contextual loss
- Adversarial loss:
  - hinge loss
  - Discriminator feature matching
- Domain alignment loss
  - Domain l1 loss
- Correspondence regularization
  - Cyclic warping loss

Comparison with state-of-the-art

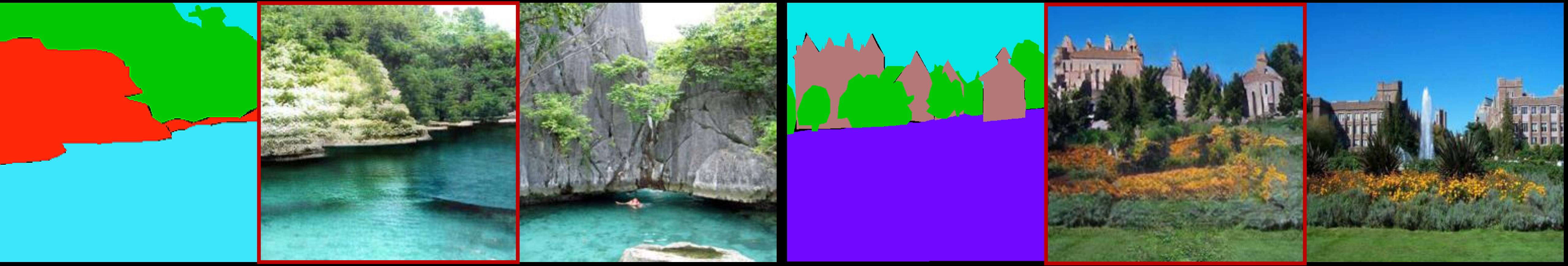


User preference

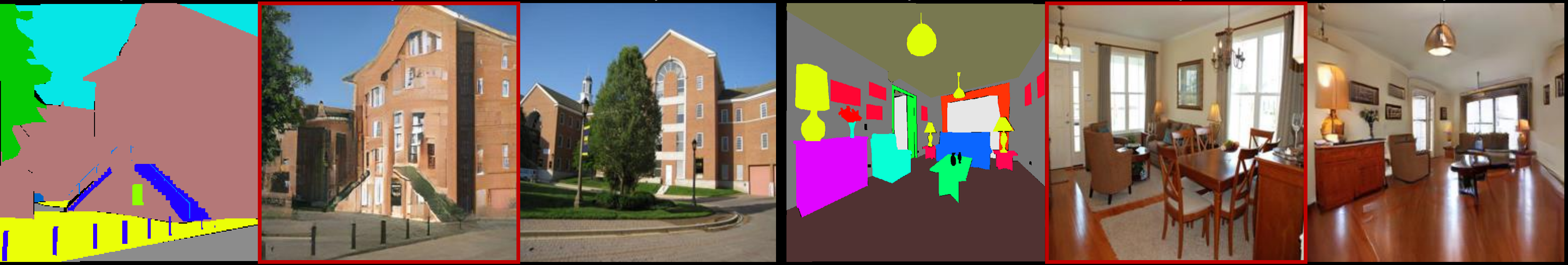


Results on ADE20k

Input Output Exemplar Input Output Exemplar

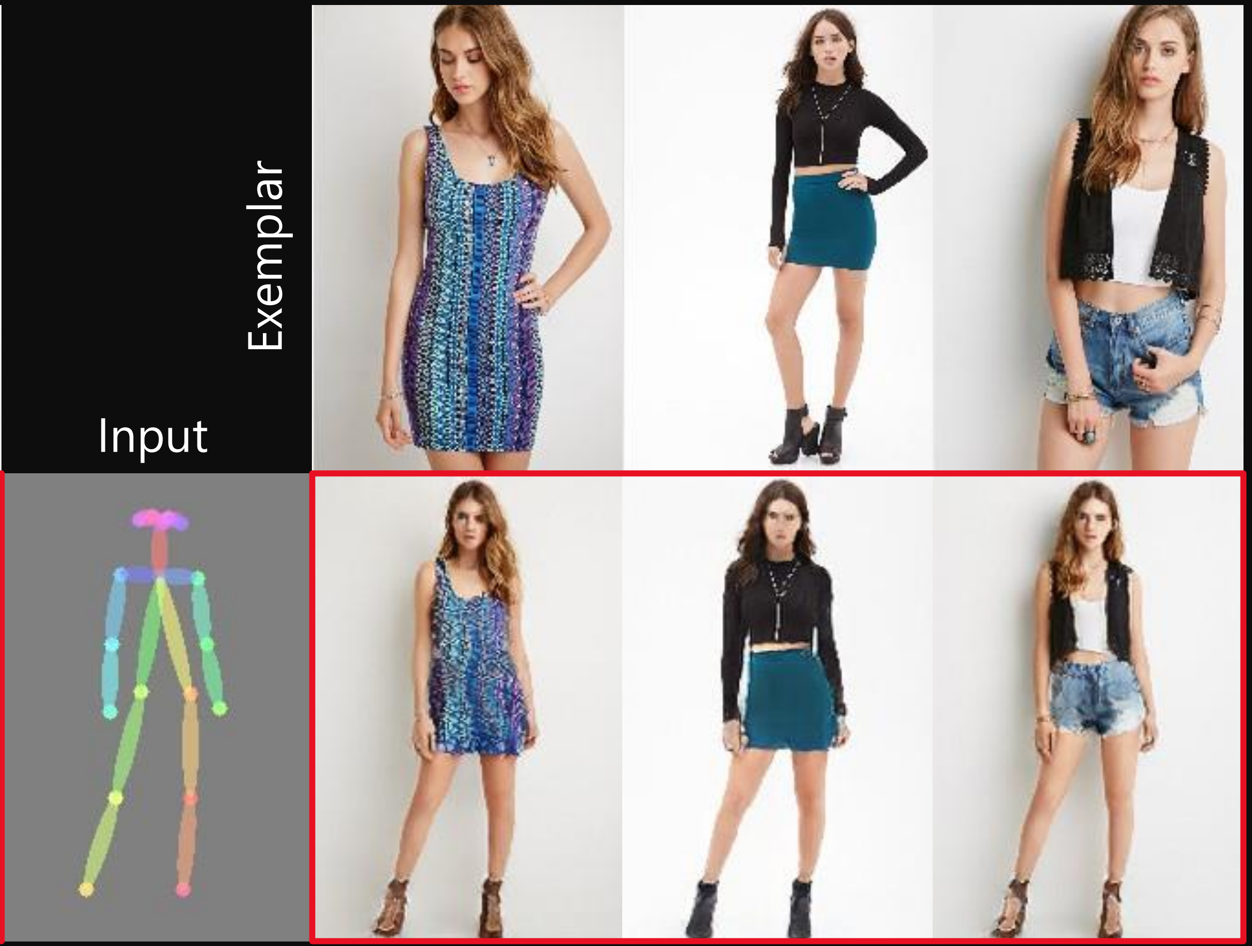


Input Output Exemplar Input Output Exemplar

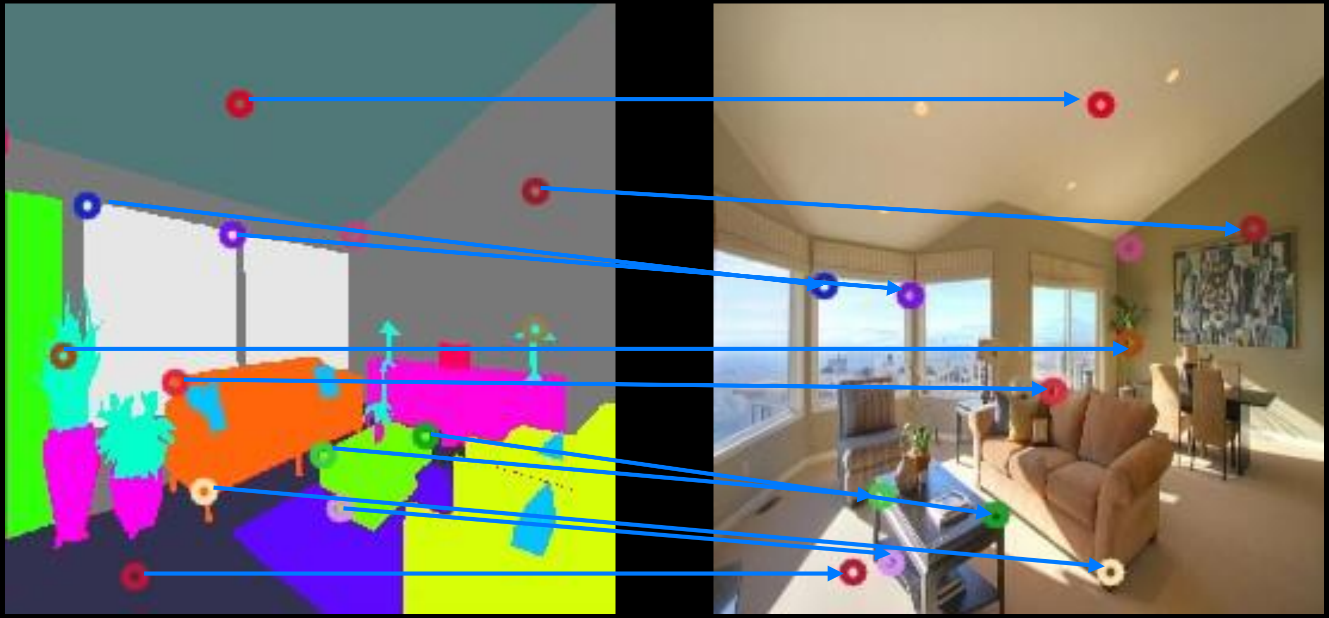


Results on Celeb-A

Results on Deepfashion

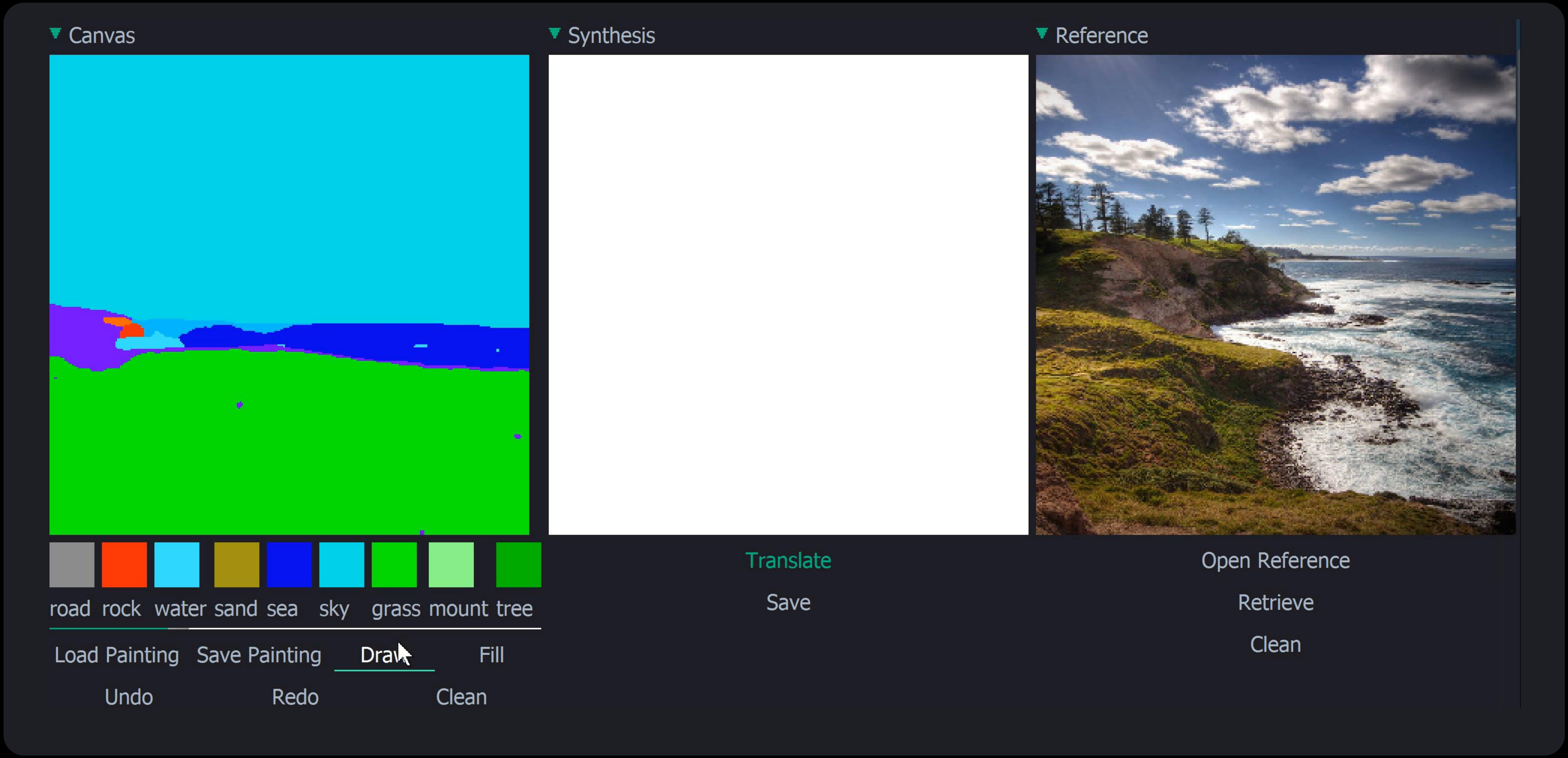


Cross-domain correspondence

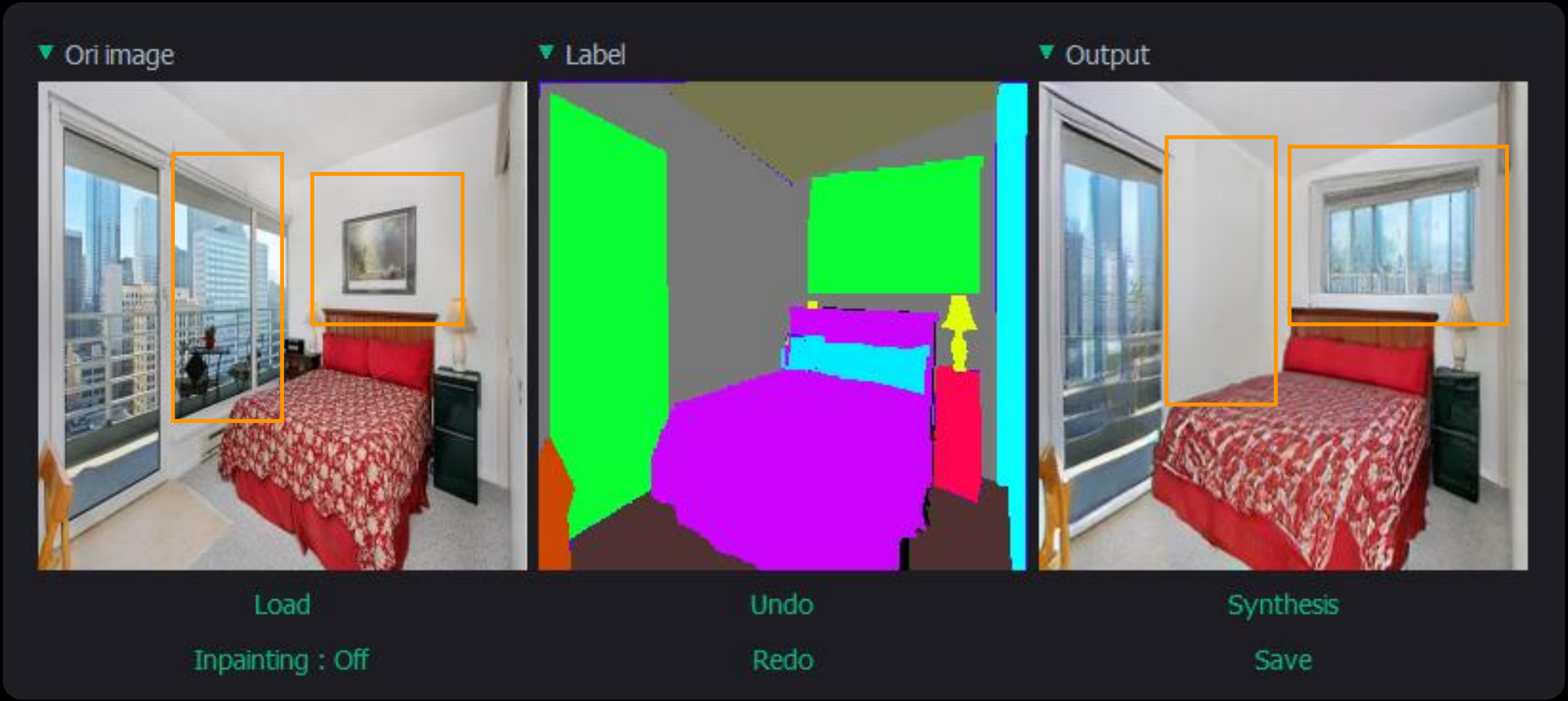


Weak supervised learning !

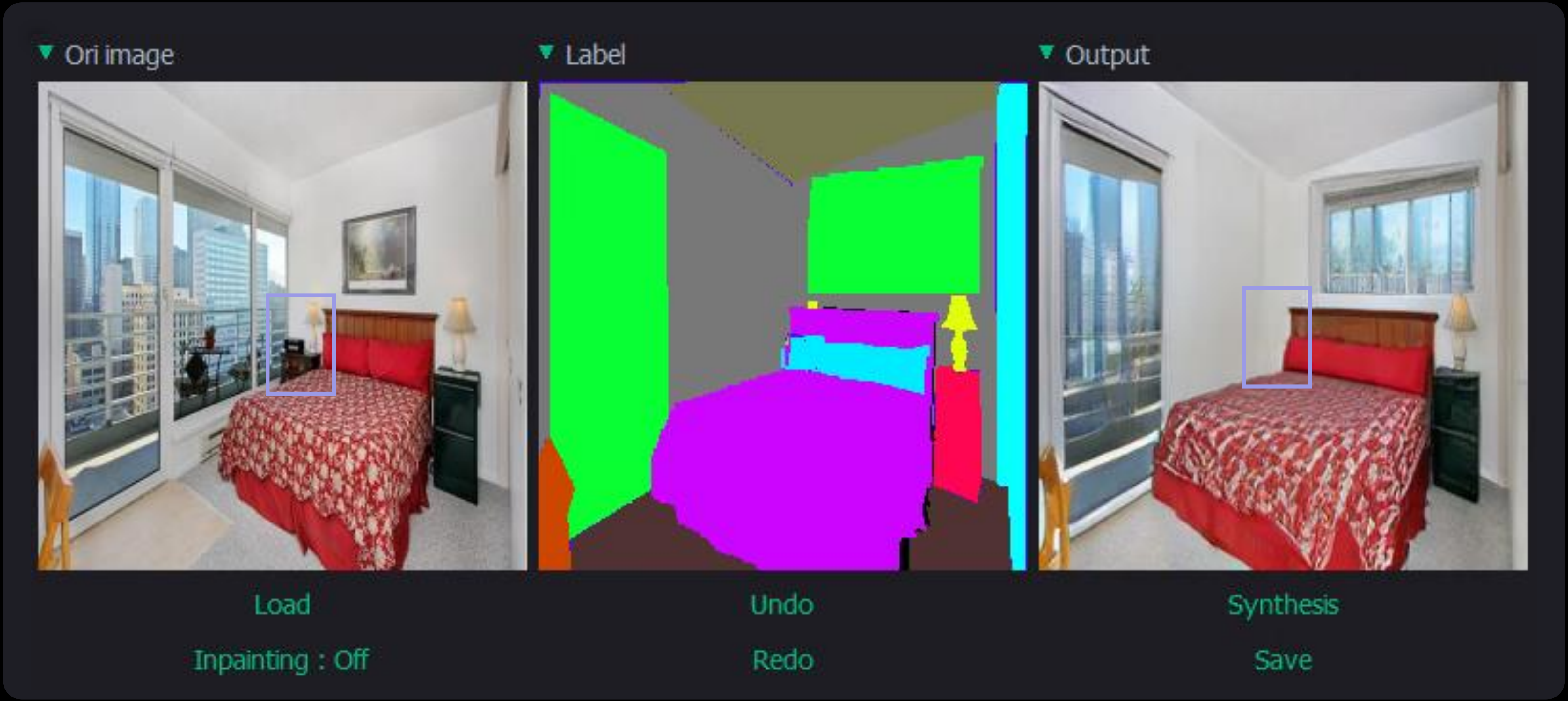
Application: interactive painter



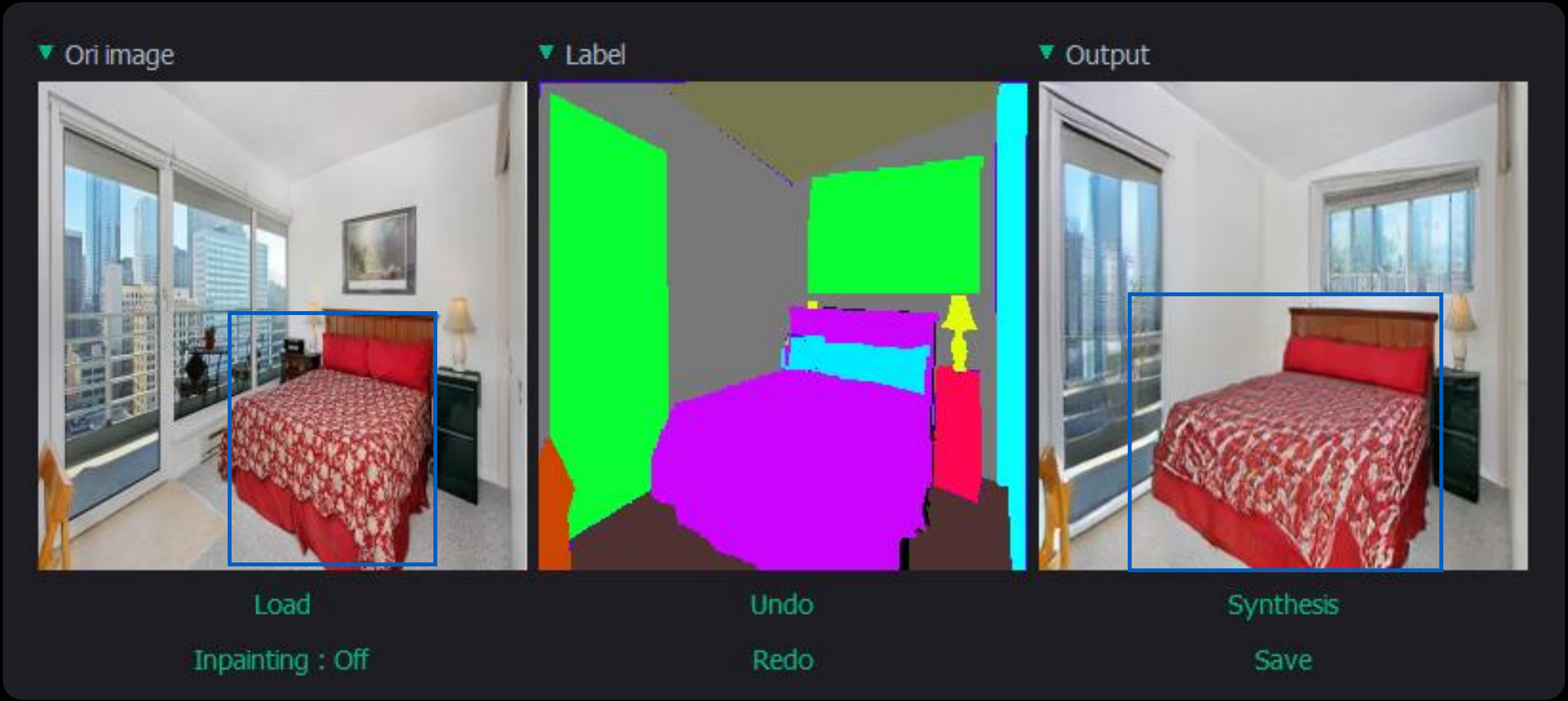
Application: image editing



Application: image editing



Application: image editing



Application: makeup transfer

Microsoft

Thank you!

Project webpage: https://panzhang0212.github.io/CoCosNet/
Code will be released soon.

Bringing photo back to life, CVPR 2020 oral

Project page: http://raywzy.com/Old_Photo/