The Surprising Effectiveness of Linear Unsupervised Image-to-Image Translation

האוניברסיטה העברית בירושלים
THE HEBREW UNIVERSITY OF JERUSALEM

Eitan Richardson and Yair Weiss

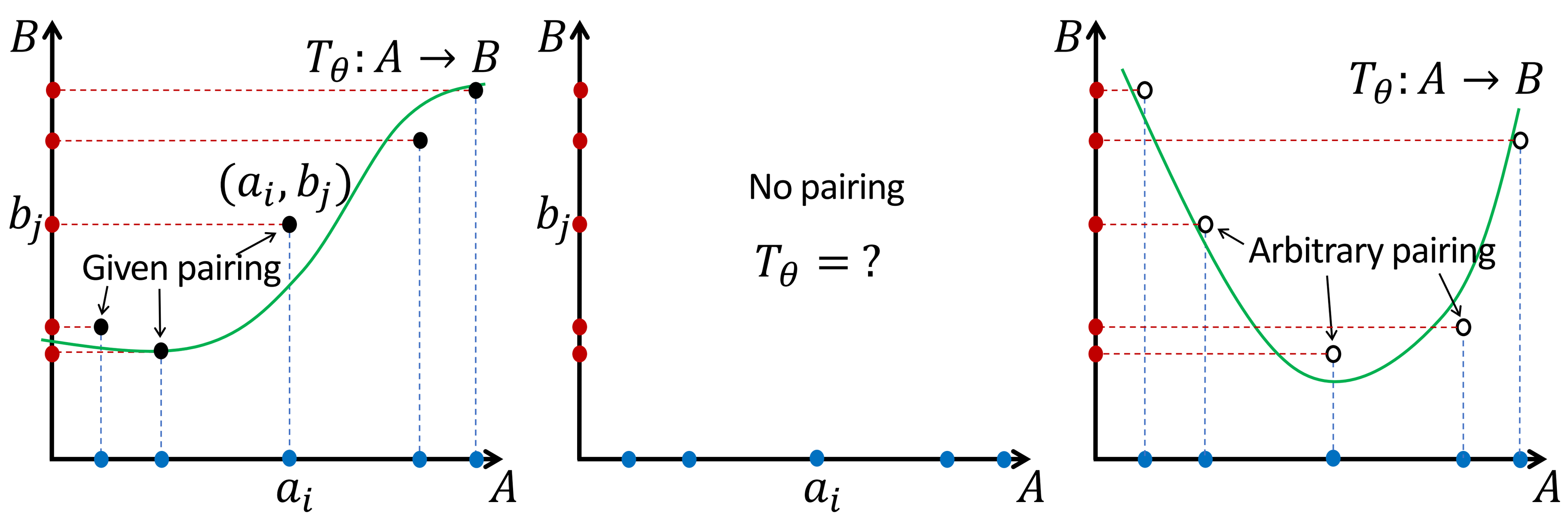
ICPR 2020
25th INTERNATIONAL CONFERENCE ON PATTERN RECOGNITION
Milan, Italy 10 | 15 January 2021

Unsupervised Domain Translation (UDT)

Input: datasets D_A, D_B sampled from the *marginals* $P(A), P(B)$ of some joint distribution $P(A,B)$.

Task: Learn $P(B|A)$

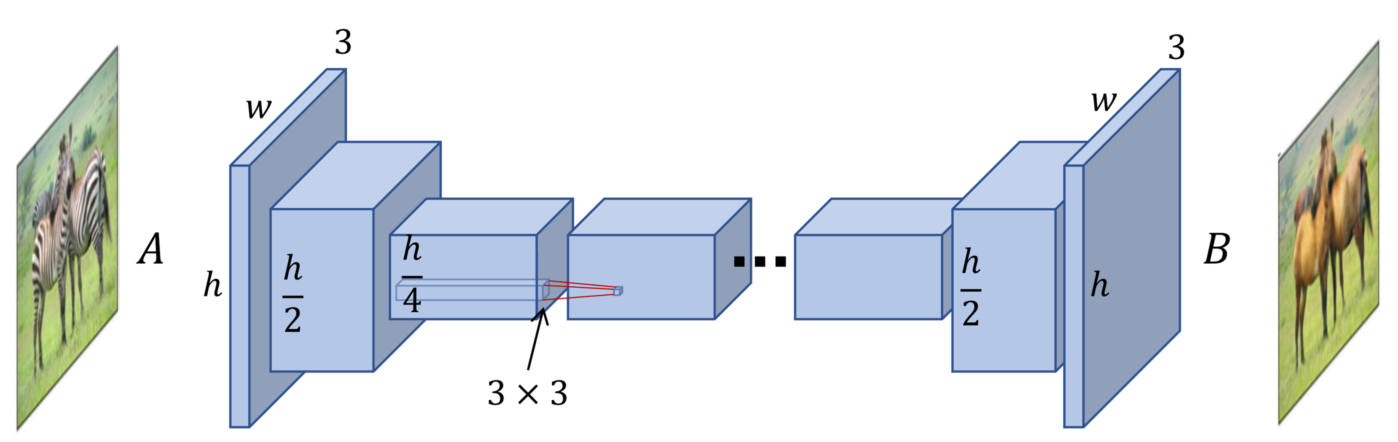
UDT is ill-posed !



Toy example: $A, B \subset \mathbb{R}$. Without matching pairs, an arbitrary pairing defines a valid transformation.

So how does CycleGAN, MUNIT, … work?

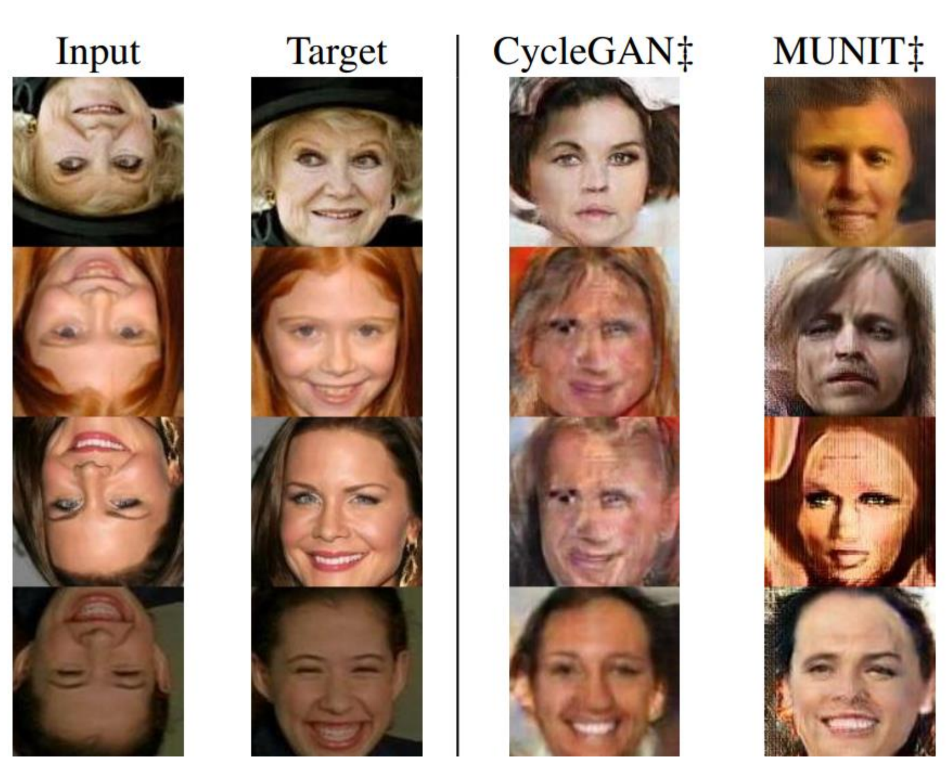
Locality bias: Problem + Architecture



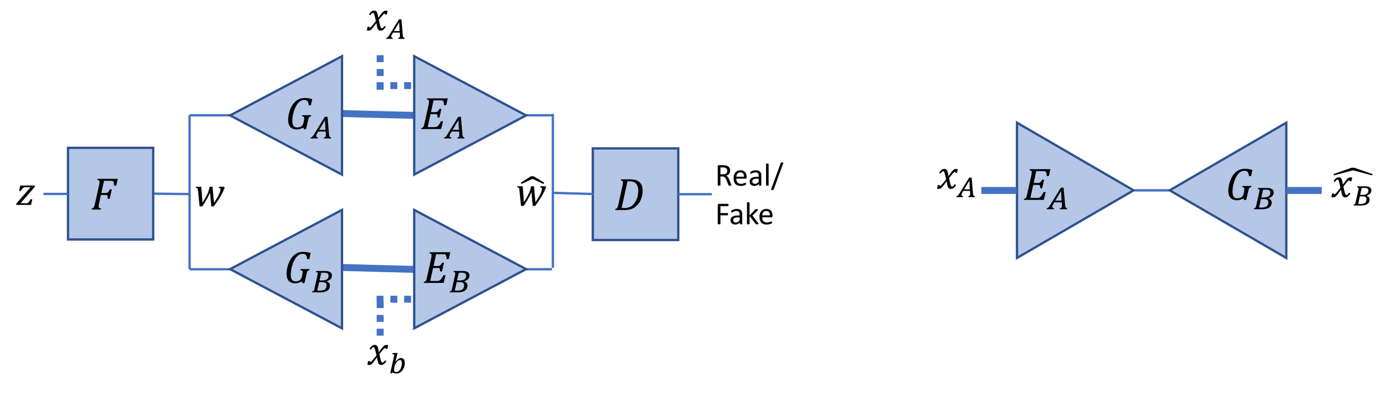
CycleGAN, … use an autoencoder bottleneck with a large spatial size and small convolution kernels.

Nonlocal problem – failure:

CycleGAN and MUNIT fail to learn a simple nonlocal problem like *vertical flip*.

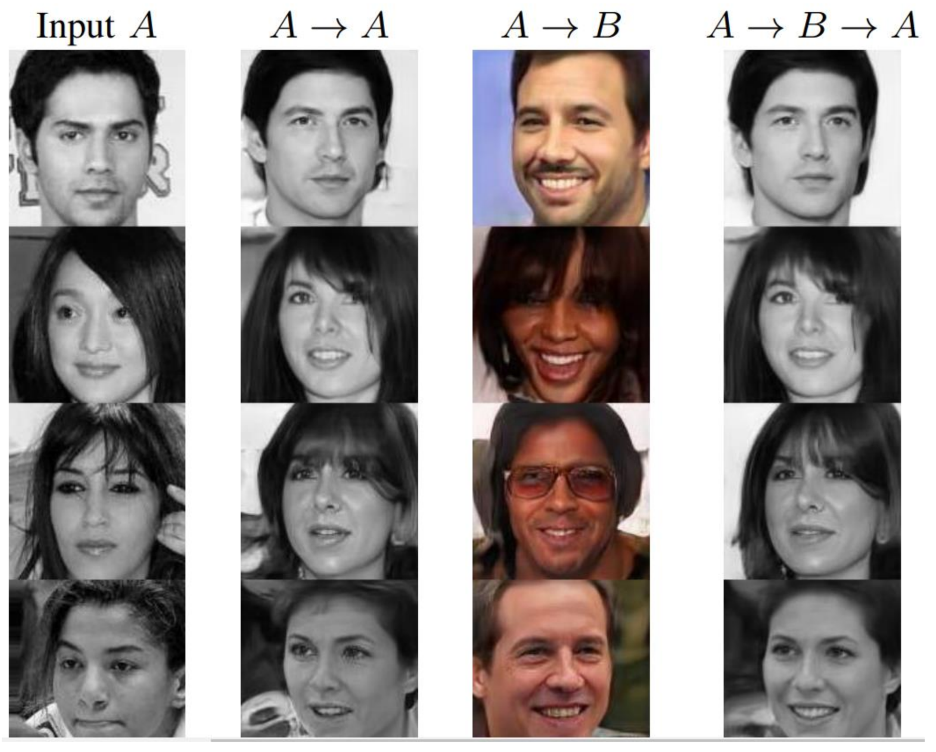


Nonlocal architecture – failure:



StyleGAN encoder-decoder with flat bottleneck, based on ALAE. Left: training, Right: inference.

A deep UDT architecture w/o locality bias converges to an ***arbitrary*** solution.



Linear, orthogonal image-to-image translation

Find $T: A \to B$ ($b = Ta$) such that $TT' = T'T = I$

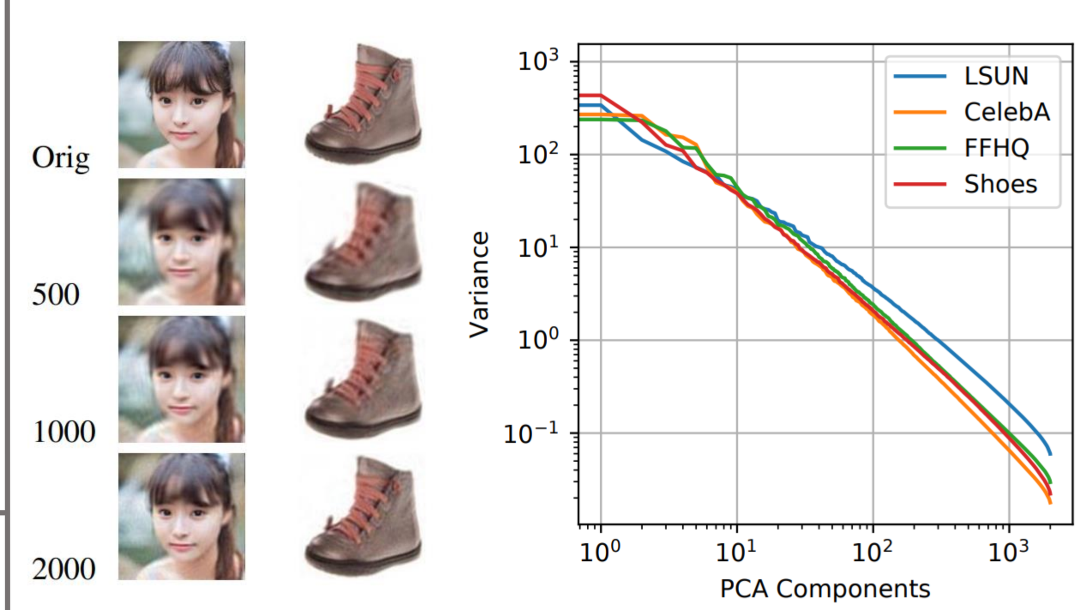
Challenges:

- T is very large
- Unsupervised learning scheme?
- Expressiveness?

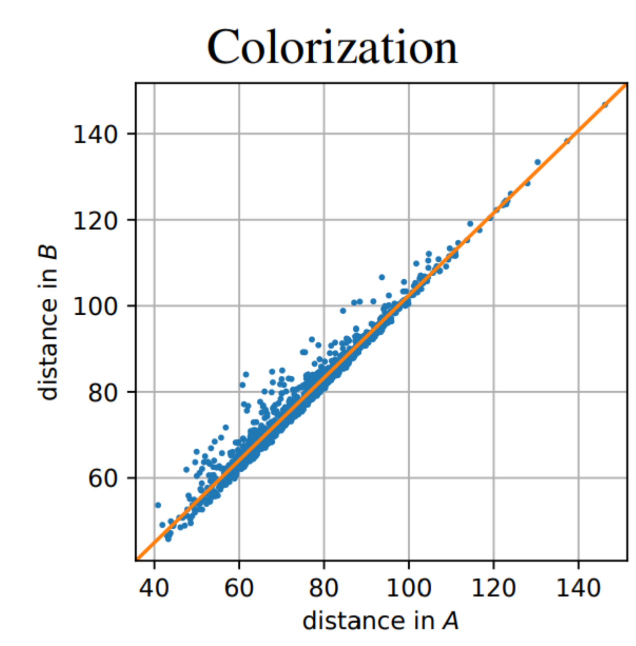
Solution: Learn a linear transformation in *PCA space*

$$T = W_B Q W_A'$$

r principal components of D_B — learned $r \times r$ orthogonal matrix — r principal components of D_A

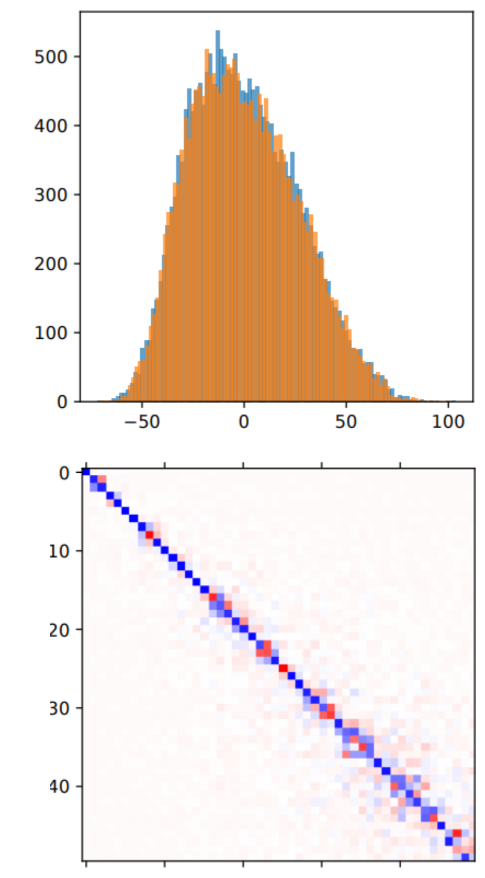


Natural images are well represented by relatively few PCA components ($r \ll d$).



Some real tasks like colorization are close to being distance-preserving.

Learning method: Procrustes + ICP in PCA space



PCA ambiguity is resolved via skewness. Target Q is close to identity.

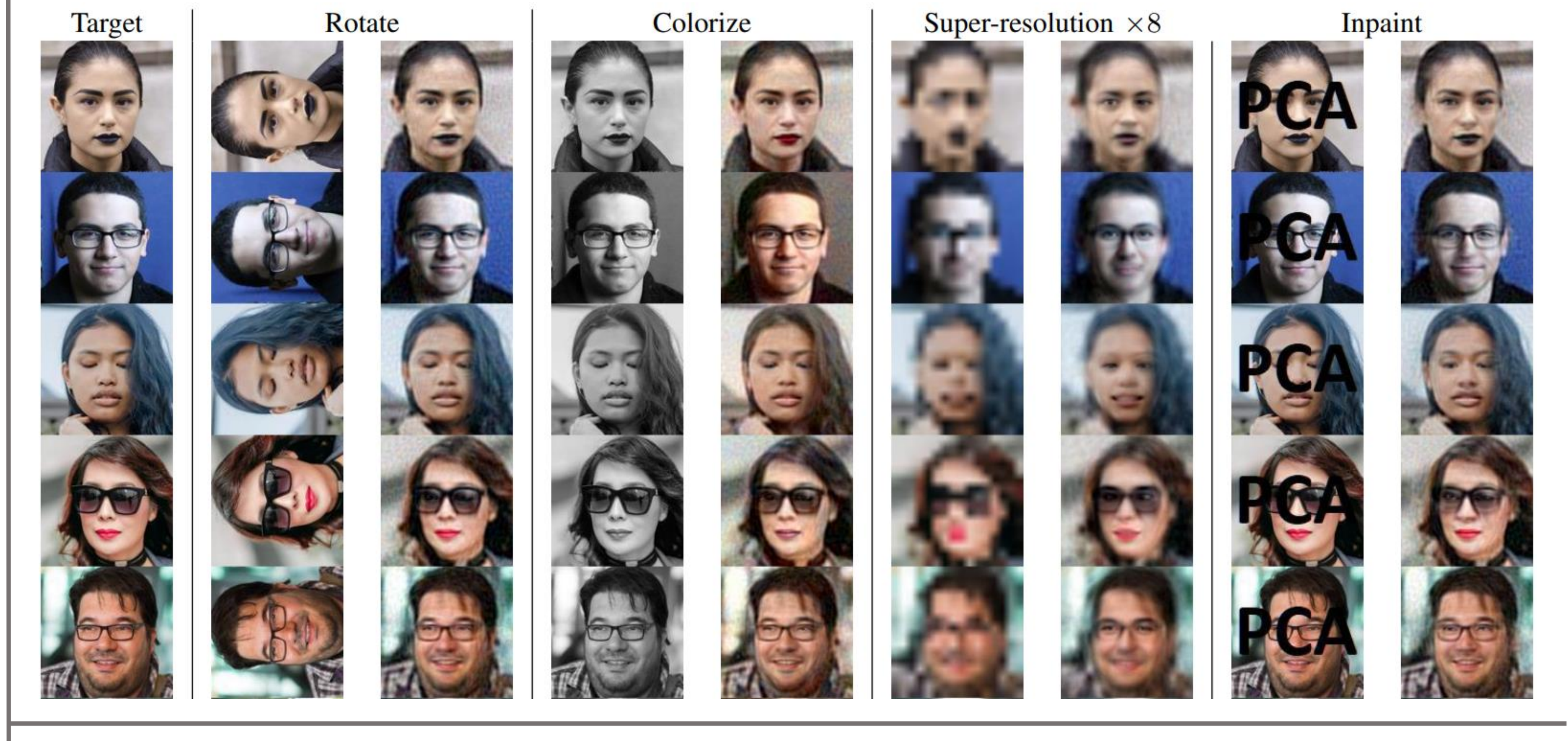
Algorithm 1: Orthogonal UDT in PCA subspace

Input: $\mathcal{D}_A = \{x_1^A, \dots x_n^A\}$, $\mathcal{D}_B = \{x_1^B, \dots x_m^B\}$, r
Result: Orthogonal transformation $T: A \to B$

Compute W_A, W_B: r principal components of $\mathcal{D}_A, \mathcal{D}_B$
Fix eigenvectors sign for positive skew
Compute PCA embedding $\{z_1^A, \dots z_n^A\}$, $\{z_1^B, \dots z_m^B\}$
Initialize $Q \leftarrow I$
while *not converged* **do**
  $A \leftarrow \emptyset$, $B \leftarrow \emptyset$
  for $i \leftarrow 1$ ***to*** n **do**
    $j \leftarrow \arg\min_{j'} ||z_i^A Q - z_{j'}^B||$
    $k \leftarrow \arg\min_{k'} ||z_{k'}^A Q - z_j^B||$
    if $k = i$ **then**
      A.insert-row(z_i^A), B.insert-row(z_j^B)
  $U, S, V \leftarrow \mathrm{SVD}(A^T B)$
  $Q \leftarrow UV$
return $T \leftarrow W_A^T Q W_B$

Results

Task	CycleGAN‡			MUNIT‡			Ours‡			Ours†		
	MSE	SSIM	T[h]	MSE	SSIM	T[h]	MSE	SSIM	T[h]	MSE	SSIM	T[h]
CelebA-colorize	0.0066	**0.914**	49	0.0256	0.750	52	**0.0043**	0.883	**0.04**	0.0071	0.761	0.04
CelebA-vflip	0.1167	0.358	43	0.1084	0.333	48	**0.0012**	**0.917**	**0.04**	0.0041	0.780	0.04
FFHQ-rot90	0.1267	0.302	39	0.1220	0.268	39	**0.0023**	**0.870**	**0.05**	0.0335	0.381	0.05



Conclusion

- UDT is in general ill-posed.
- SOTA unsupervised im2im methods rely on locality bias.
- Our approach - linear orthogonal transformations - can be learned in a few seconds and works well for many true relations.

eitanrich@cs.huji.ac.il https://github.com/eitanrich/lin-im2im yweiss@cs.huji.ac.il