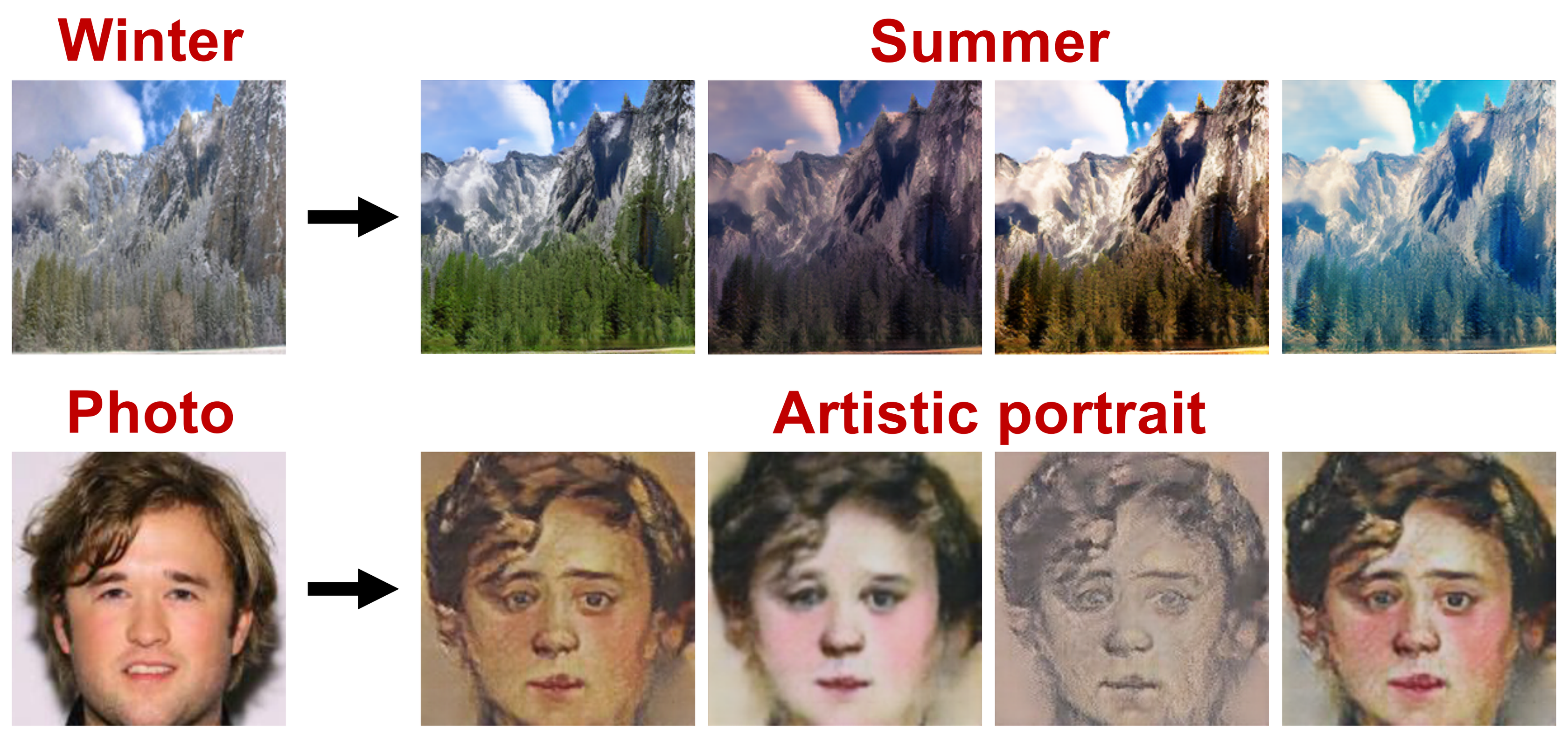
Diverse Image-to-Image Translation via Disentangled Representations

Code available!
http://bit.ly/DRIT-ECCV18

Hsin-Ying Lee*[1] Hung-Yu Tseng*[1] Jia-Bin Huang[2] Maneesh Singh[3] Ming-Hsuan Yang[1,4]

[1]University of California, Merced [2]Virginia Tech [3]Verisk Analytics [4]Google Cloud AI

Image-to-image translation



Challenges

1. Lack of aligned training pairs
2. Multiple possible outputs given singe input image

	Paired data	Unpaired data
One-to-one	Pix2pix [Isola *et al.*]	DiscoGAN [Kim *et al.*] CycleGAN [Zhu *et al.*] UNIT [Liu *et al.*]
One-to-many	BicycleGAN [Zhu *et al.*] Pix2pixHD [Wang *et al.*]	DRIT (Ours)

Contributions

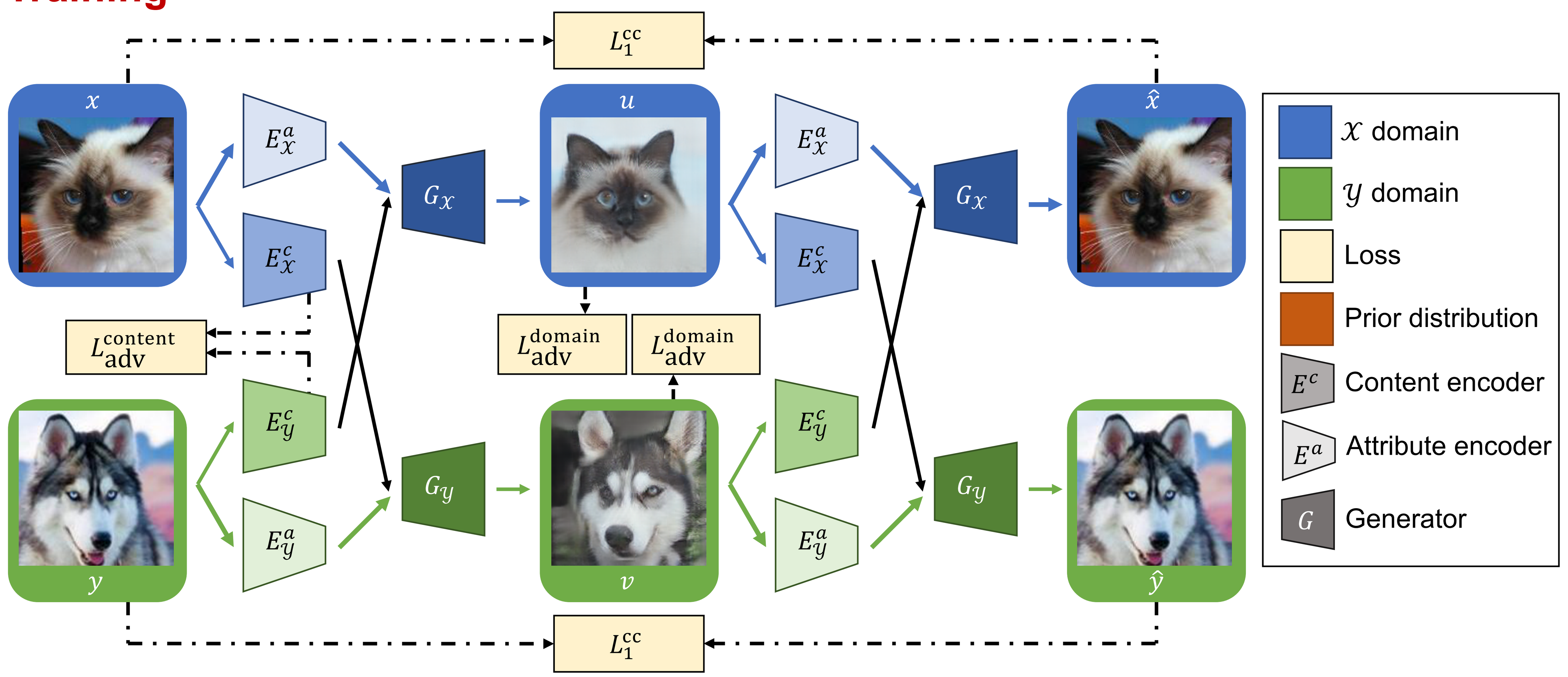
1. Disentangled representation & cross cycle consistency
2. Diverse translation from unpaired data
3. Competitive performance on domain adaptation

References

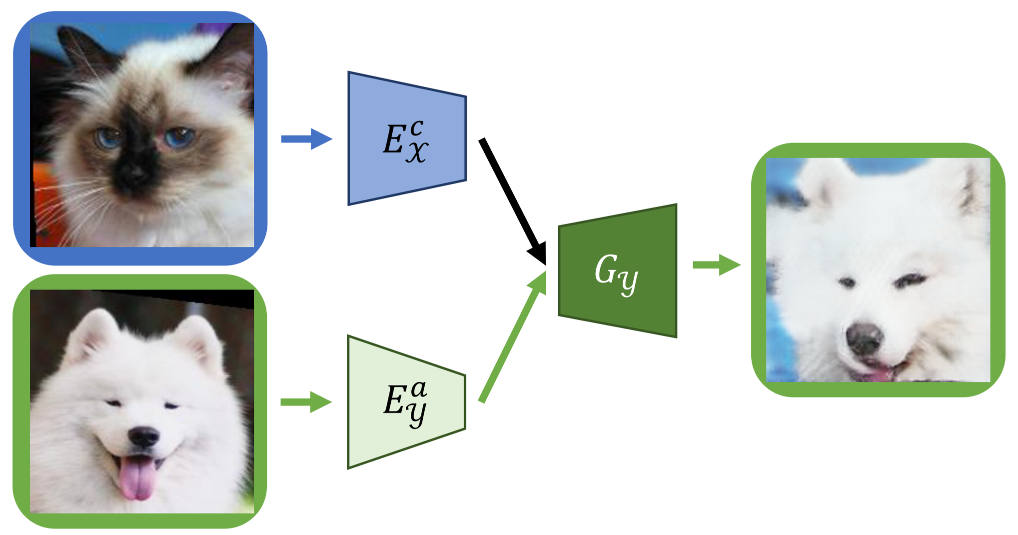
[1] Liu *et al.* Unsupervised Image-to-Image Translation Networks. In *NIPS*, 2017
[2] Zhu *et al.* Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks. In *ICCV*, 201
[3] Zhu *et al.* Toward Multimodal Image-to-Image Translation. In *NIPS*, 2017

Disentangled representation for image-to-image translation (DRIT)

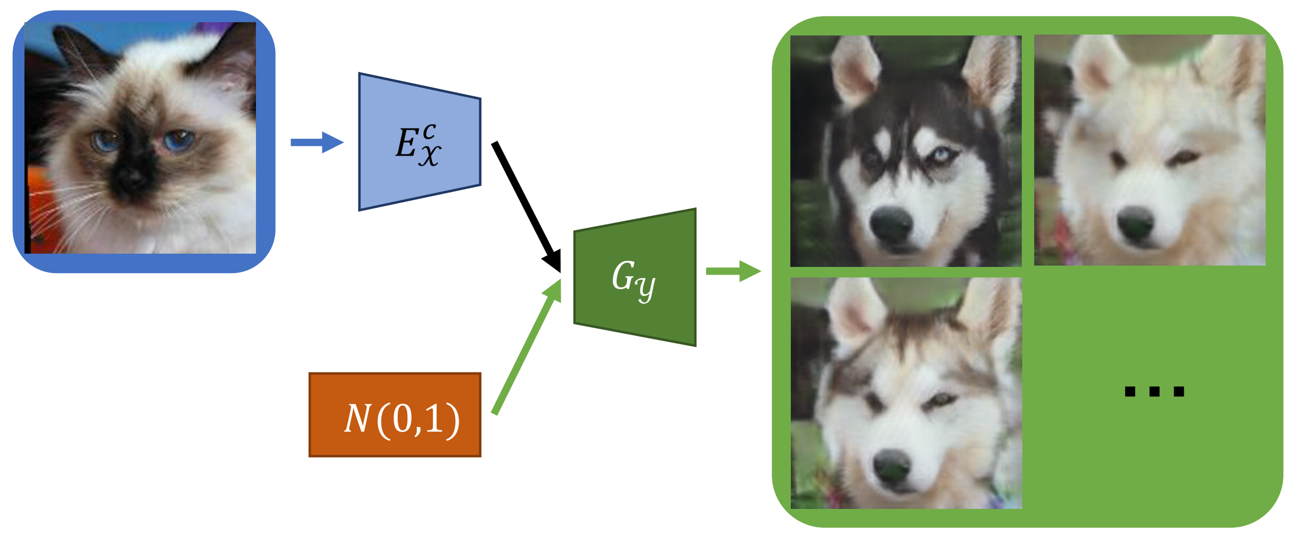
Training



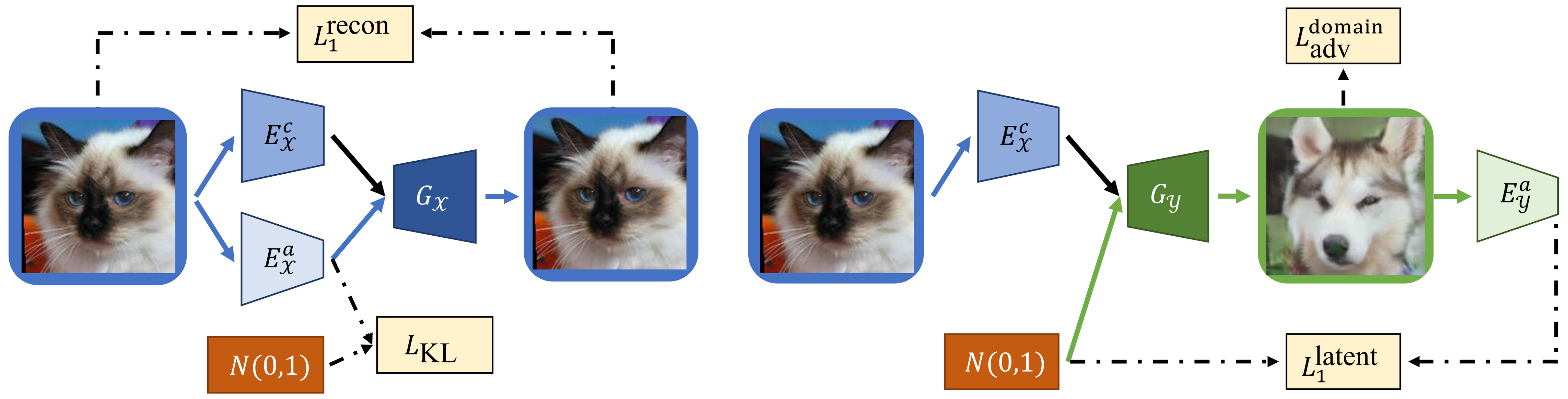
Example guided translation



Randomly sampled translation



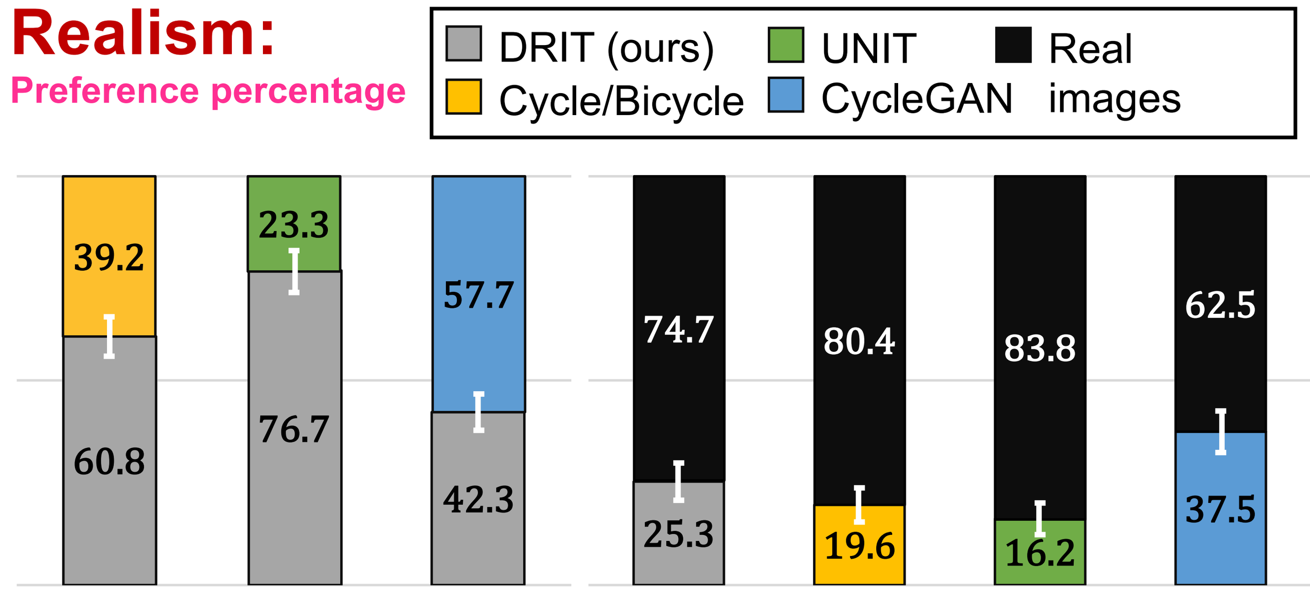
Other training loss



Experimental results

Visual results

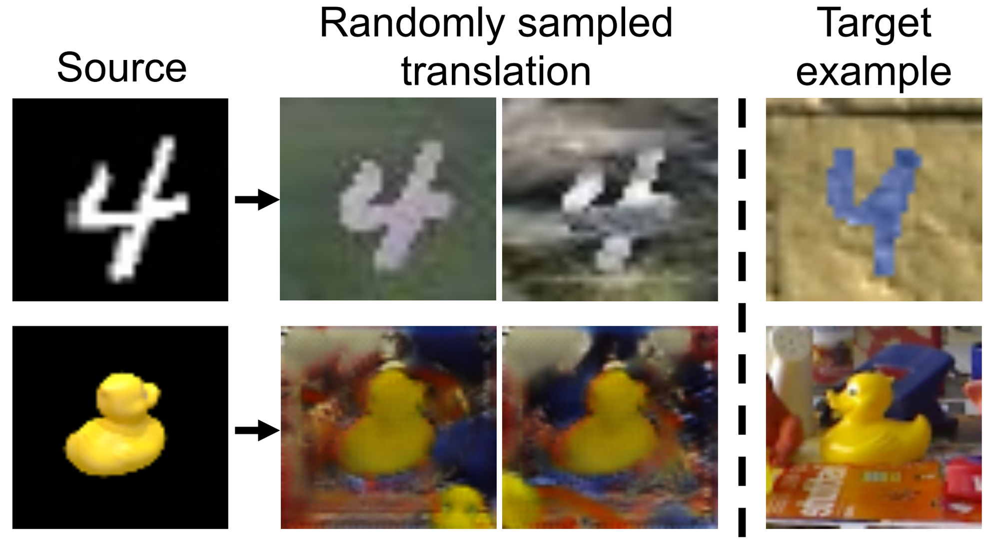
Realism: Preference percentage



Diversity: LPIPS score

Method	Diversity↑
Real images	.448 ± .012
DRIT	**.424** ± .010
UNIT [1]	.406 ± .022
CycleGAN [2]	<u>.413</u> ± .008
Cycle/Bicycle	.399 ± .009

Domain adaptation



Model	(a) MNIST-M Classification Accuracy (%)	(b) LineMOD Mean Angle Error (°)
Source-only	56.6	73.7 (89.2)
CycleGAN [2]	74.5	47.45
Ours, ×1	86.93	42.06
Ours, ×3	<u>90.21</u>	<u>37.35</u>
Ours, ×5	**91.54**	**34.4**
Target-only	96.5	12.3 (6.47)