Diverse Image-to-Image Translation via Disentangled Representations

Hsin-Ying Lee, Hung-Yu Tseng, Jia-Bin Huang, Maneesh Singh, Ming-Hsuan Yang

In ECCV, 2018 (oral paper)

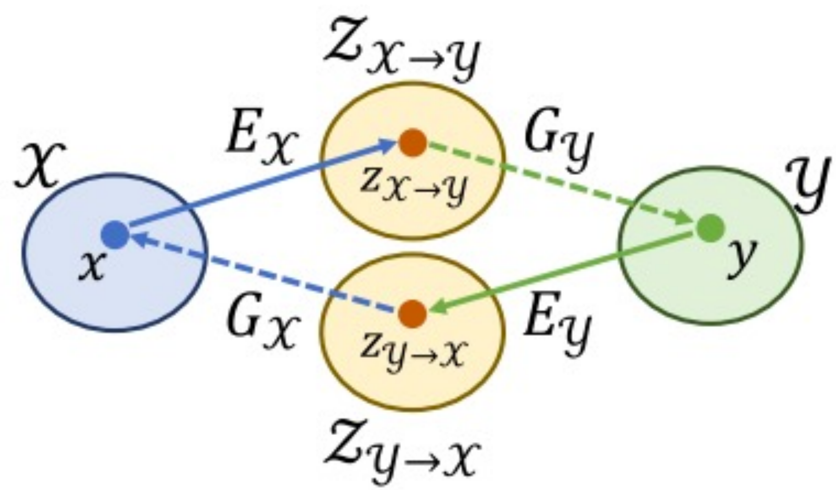
Sharer: Du Ang
2018.08.15

Image-to-Image (I2I) Challenges

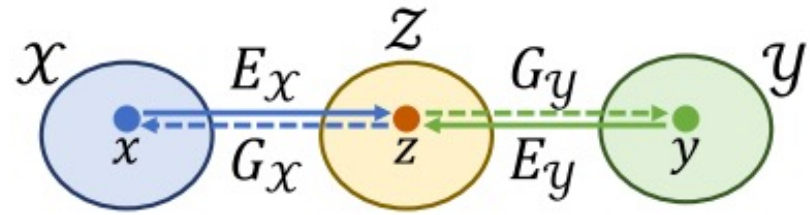
(1) the lack of aligned training image pairs

(2) multiple possible outputs from a single input image

Related Work

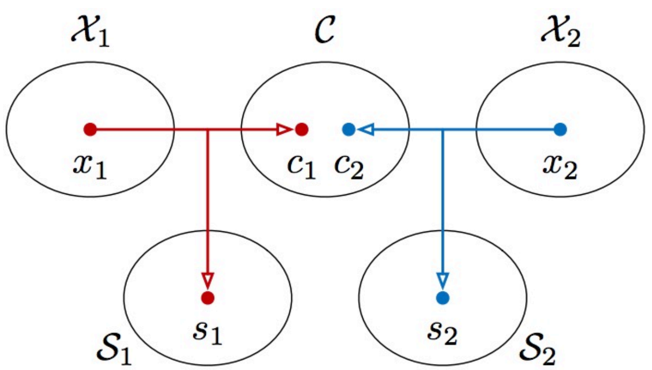


CycleGAN

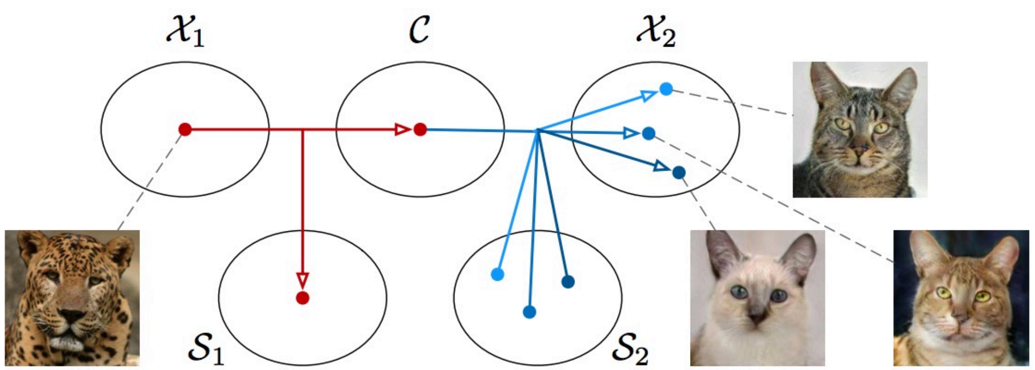


UNIT

Related Work



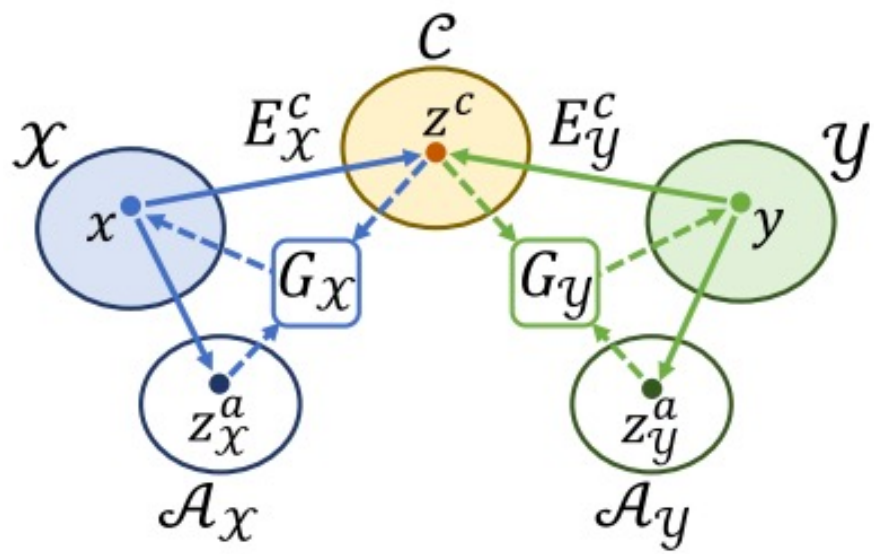
(a) Auto-encoding



(b) Translation

MUNIT

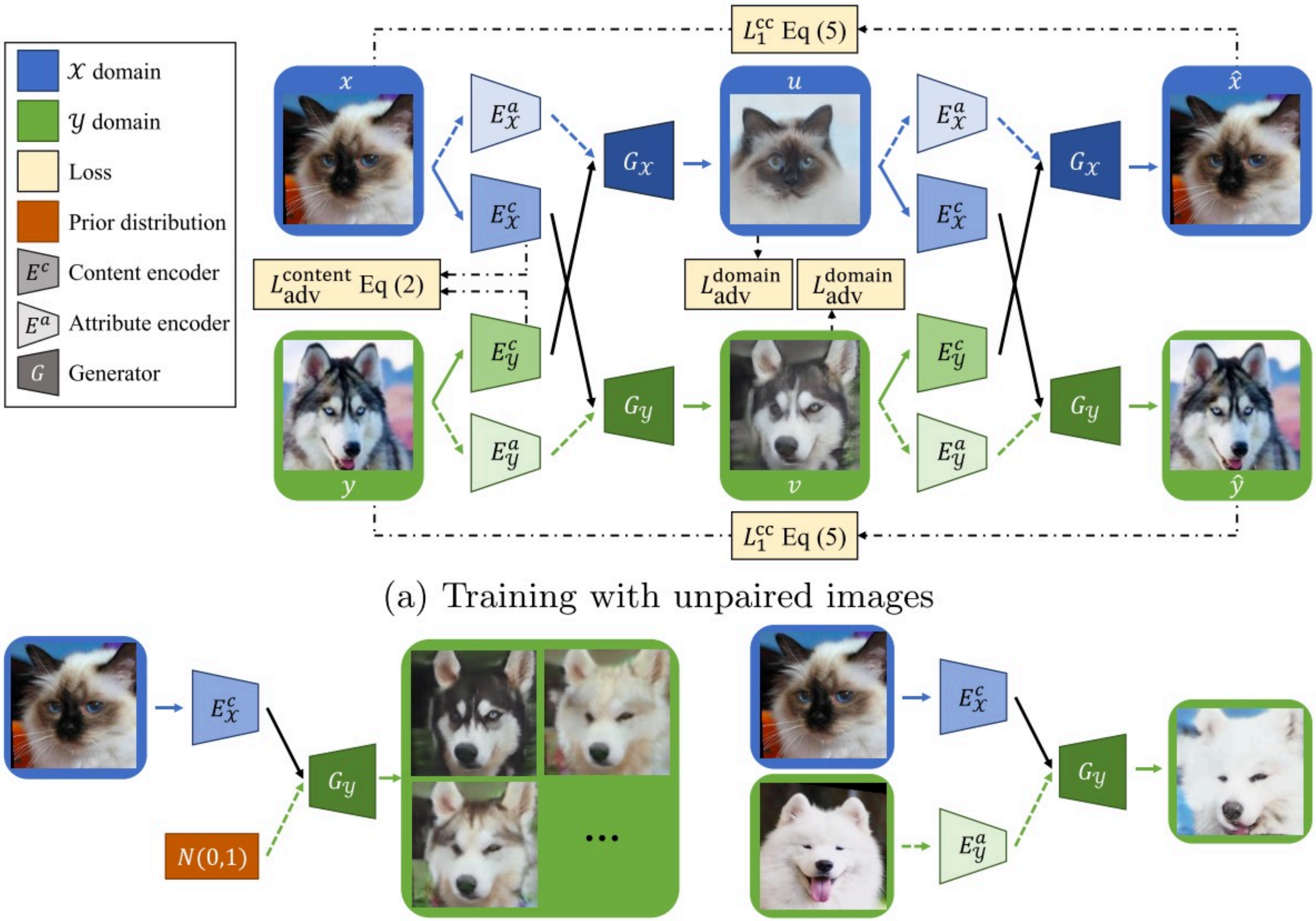
Method



$\mathcal{C}$
$E^c_{\mathcal{X}}$
z^c
$E^c_{\mathcal{Y}}$
$\mathcal{X}$
$\mathcal{Y}$
x
y
$G_{\mathcal{X}}$
$G_{\mathcal{Y}}$
$z^a_{\mathcal{X}}$
$z^a_{\mathcal{Y}}$
$\mathcal{A}_{\mathcal{X}}$
$\mathcal{A}_{\mathcal{Y}}$

DRIT

Method



(a) Training with unpaired images

(b) Testing with random attributes

(c) Testing with a given attribute

DRIT

Method

Content adversarial loss of the content discriminator

$$L_{\text{adv}}^{\text{content}}(E_{\mathcal{X}}^c, E_{\mathcal{Y}}^c, D^c) = \mathbb{E}_x[\frac{1}{2}\log D^c(E_{\mathcal{X}}^c(x)) + \frac{1}{2}\log(1 - D^c(E_{\mathcal{X}}^c(x)))]$$
$$+ \mathbb{E}_y[\frac{1}{2}\log D^c(E_{\mathcal{Y}}^c(y)) + \frac{1}{2}\log(1 - D^c(E_{\mathcal{Y}}^c(y)))]$$

Cross-cycle consistency loss

$$L_1^{\text{cc}}(G_{\mathcal{X}}, G_{\mathcal{Y}}, E_{\mathcal{X}}^c, E_{\mathcal{Y}}^c, E_{\mathcal{X}}^a, E_{\mathcal{Y}}^a) = \mathbb{E}_{x,y}[\|G_{\mathcal{X}}(E_{\mathcal{Y}}^c(v), E_{\mathcal{X}}^a(u)) - x\|_1$$
$$+\|G_{\mathcal{Y}}(E_{\mathcal{X}}^c(u), E_{\mathcal{Y}}^a(v)) - y\|_1],$$

where $u = G_{\mathcal{X}}(E_{\mathcal{Y}}^c(y)), E_{\mathcal{X}}^a(x))$ and $v = G_{\mathcal{Y}}(E_{\mathcal{X}}^c(x)), E_{\mathcal{Y}}^a(y))$.

Method

Other loss functions

- Domain adversarial loss
- Self-reconstruction loss
- KL loss
- Latent regression loss

Full objective function:

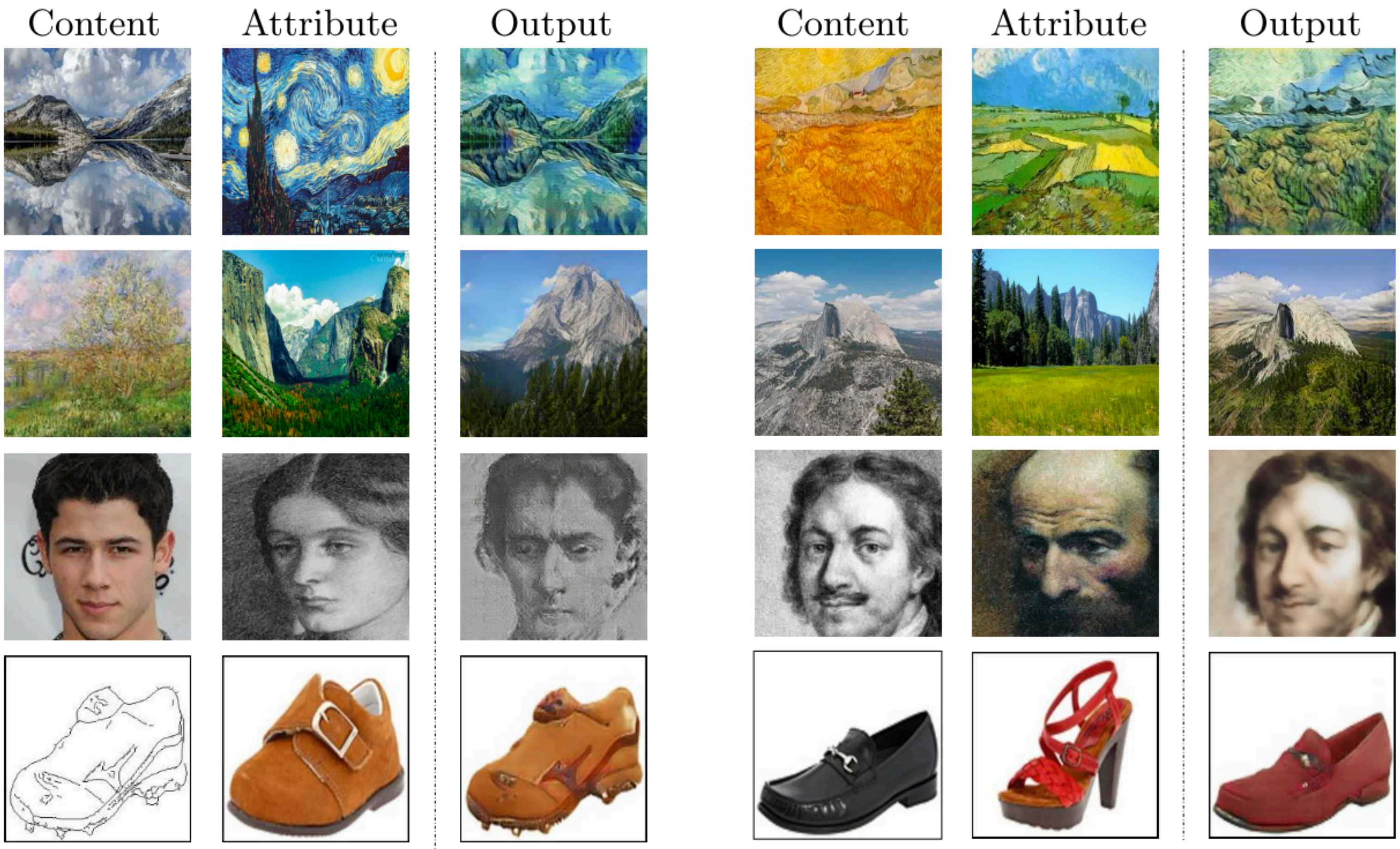
$$\min_{G,E^c,E^a} \max_{D,D^c} \quad \lambda_{\text{adv}}^{\text{content}} L_{\text{adv}}^{\text{c}} + \lambda_1^{\text{cc}} L_1^{\text{cc}} + \lambda_{\text{adv}}^{\text{domain}} L_{\text{adv}}^{\text{domain}} + \lambda_1^{\text{recon}} L_1^{\text{recon}}$$
$$+\lambda_1^{\text{latent}} L_1^{\text{latent}} + \lambda_{\text{KL}} L_{\text{KL}}$$

where the hyper-parameters λs control the importance of each term.

Experimental Results (Qualitative)

Diversity

Experimental Results (Qualitative)

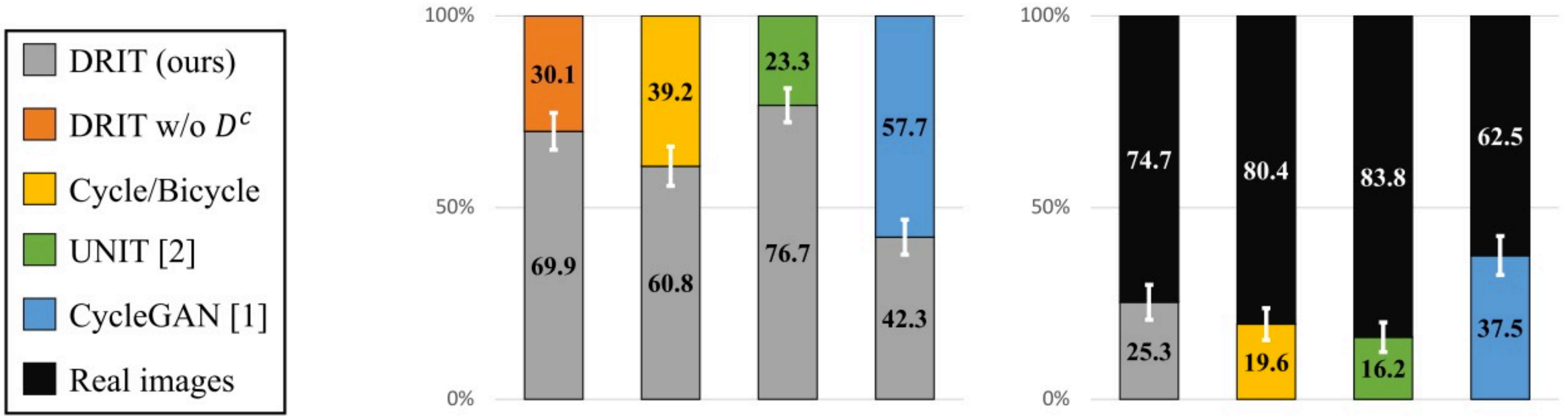


(a) Inter-domain attribute transfer

(b) Intra-domain attribute transfer

Attribute transfer

Experimental Results (Quantitative)



Realism

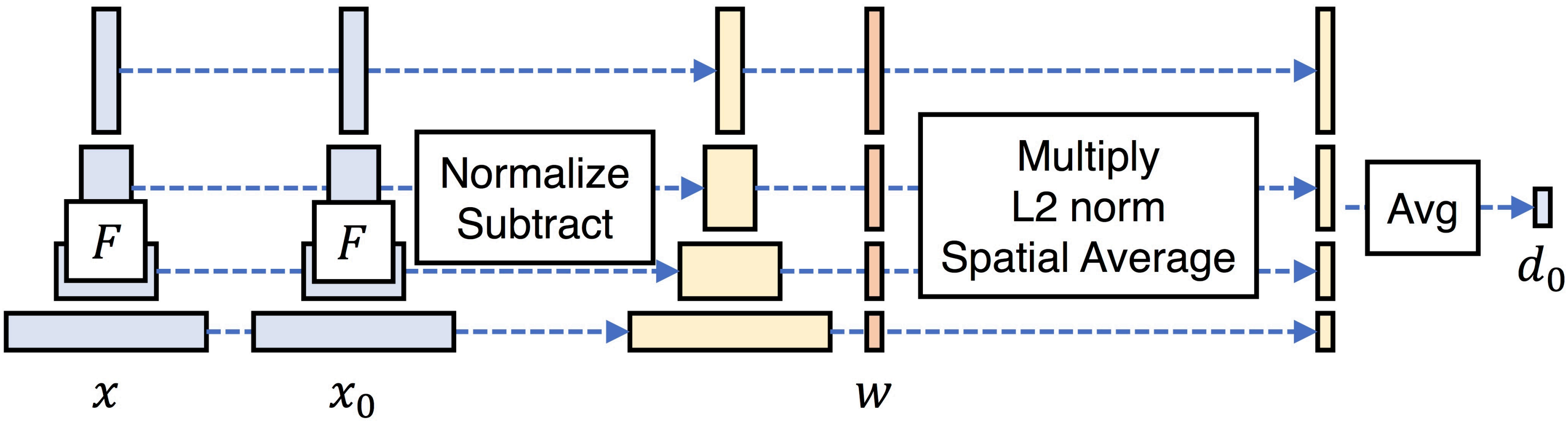
Experimental Results (Quantitative)

Method	Diversity
real images	$.448 \pm .012$
DRIT	$\mathbf{.424} \pm .010$
DRIT w/o D^c	$.410 \pm .016$
UNIT [27]	$.406 \pm .022$
CycleGAN [48]	$\underline{.413} \pm .008$
Cycle/Bicycle	$.399 \pm .009$

Diversity (LPIPS metric)

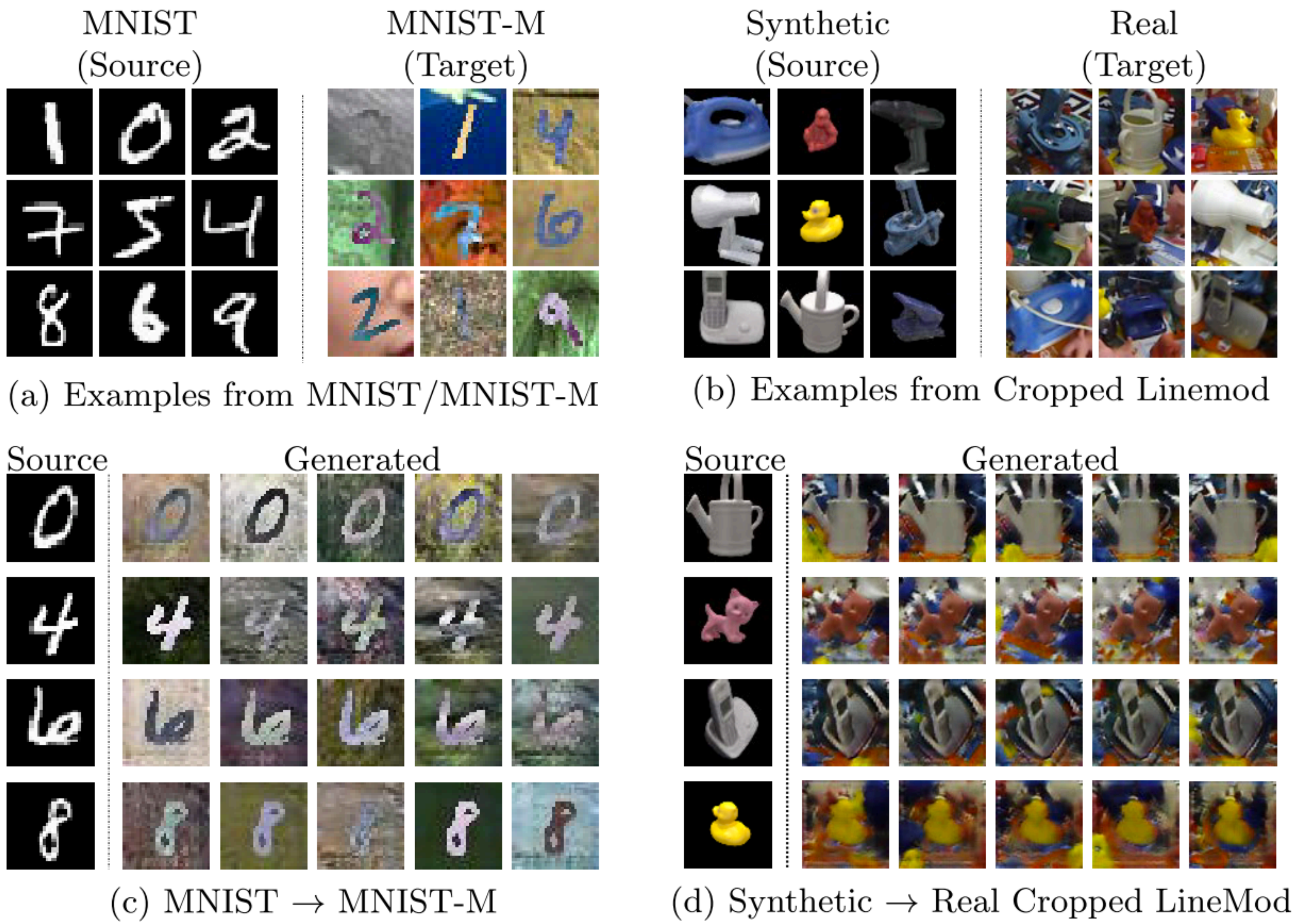
Experimental Results

LPIPS Metric



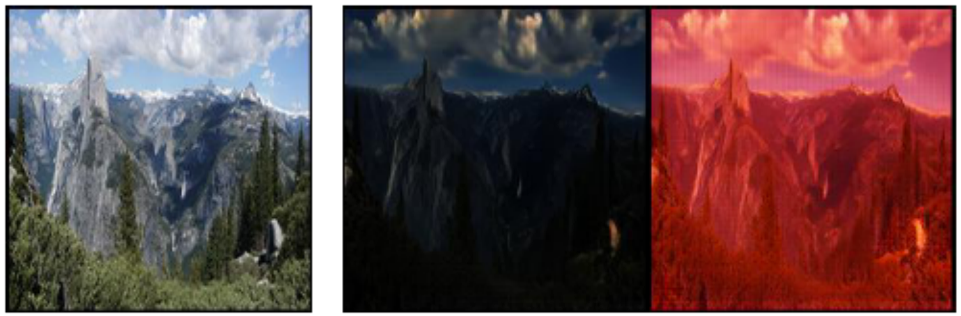
R. Zhang, P. Isola, A. A. Efros, E. Shechtman, O. Wang. The Unreasonable Effectiveness of Deep Features as a Perceptual Metric. In CVPR, 2018

Domain Adaptation Experiments



Limitations

1. Attribute space is not fully exploited



(a) Summer $\rightarrow$ Winter

2. Distribution characteristic difference

(b) van Gogh $\rightarrow$ Monet

Q & A