Cycle-GAN

Zhu, et al. 2017

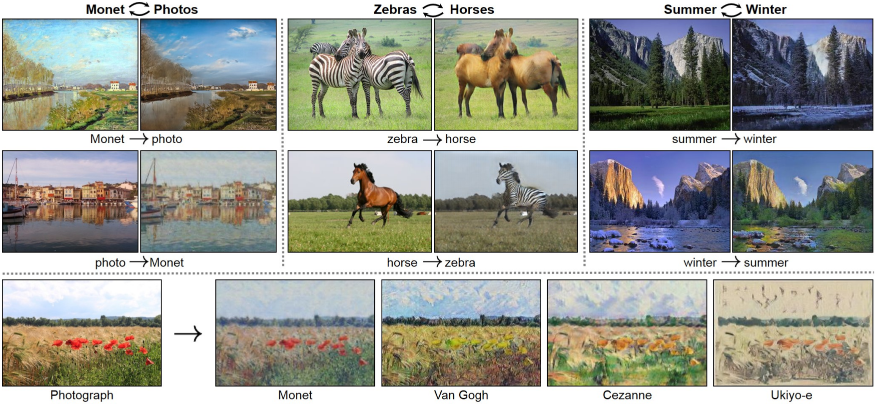
CS598 GAN 2

Outline

Outline

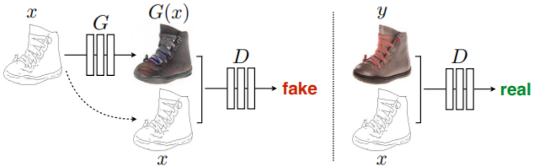
Introduction

- Convert an image from one representation to another.
- Capture characteristics of one image domain and figure out how these characteristics could be translated into the other domain.

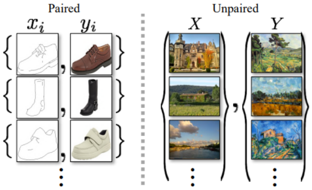


Related Works

pix2pix:



(a)



(b)

- Limitation: require supervision of paired images.

Related Works

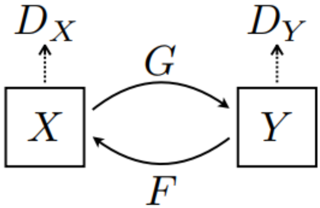
Neural Style Transfer Methods:

- Given a style image, we disentangle the style information from the content information.
- Transfer the style information to the target image.
- Limitation: only learn mapping between two specific images.

Style | Content | Ours | Chen and Schmidt | Ulyanov *et al.* | Gatys *et al.*

Formulations

- Training samples $\{x_i\}_{i=1}^N$ where $x_i \in X$ and $\{y_i\}_{i=1}^N$ where $y_i \in Y$.
- Two mapping $G : X \to Y$ and $F : Y \to X$.
- Two discriminators: D_X and D_Y.



Outline

Adversarial Loss

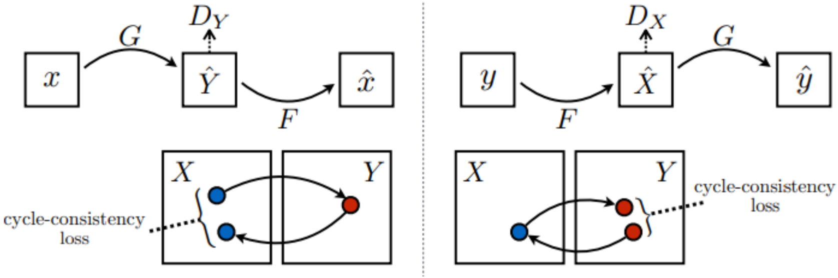
- Encourage G to generate images similar to images in domain Y and D_Y to distinguish $G(x)$ from y.

$$\mathcal{L}_{\mathrm{GAN}}(G, D_Y, X, Y) = \mathbb{E}_{y \sim p_{\text{data}}(y)}[\log D_Y(y)]$$
$$+\mathbb{E}_{x \sim p_{\text{data}}(x)}[\log(1 - D_Y(G(x))]$$

- Similarly, for F, we also have $\mathcal{L}_{\mathrm{GAN}}(F, D_X, X, Y)$.

Consistency Loss

- A neural network is capable of mapping the input images to any subsets of images in the target domain, so we need extra regularization other than the adversarial loss.
- Mapping function should be cycle-consistent, i.e., $x \rightarrow G(x) \rightarrow F(G(x)) \approx x$.



Consistency Loss

- Cycle consistency loss:

$$\mathcal{L}_{\mathsf{cyc}}(G, F) = \mathbb{E}_{x \sim p_{\mathsf{data}}(x)}[\|F(G(x)) - x\|_1] \\ +\mathbb{E}_{y \sim p_{\mathsf{data}}(y)}[\|G(F(y)) - y\|_1]$$

- Therefore, the full objective is:

$$\mathcal{L}(G, F, D_X, D_Y) = \mathcal{L}_{\mathrm{GAN}}(G, D_Y, X, Y) \\ +\mathcal{L}_{\mathrm{GAN}}(F, D_X, Y, X) \\ +\lambda\mathcal{L}_{\mathrm{cyc}}(G, F)$$

Outline

Quantitative Results

Loss	Map → Photo % Turkers labeled *real*	Photo → Map % Turkers labeled *real*
CoGAN [32]	0.6% ± 0.5%	0.9% ± 0.5%
BiGAN/ALI [9, 7]	2.1% ± 1.0%	1.9% ± 0.9%
SimGAN [46]	0.7% ± 0.5%	2.6% ± 1.1%
Feature loss + GAN	1.2% ± 0.6%	0.3% ± 0.2%
CycleGAN (ours)	**26.8% ± 2.8%**	**23.2% ± 3.4%**

Table 1: AMT "real vs fake" test on maps↔aerial photos at 256 × 256 resolution.

Loss	Per-pixel acc.	Per-class acc.	Class IOU
CoGAN [32]	0.40	0.10	0.06
BiGAN/ALI [9, 7]	0.19	0.06	0.02
SimGAN [46]	0.20	0.10	0.04
Feature loss + GAN	0.06	0.04	0.01
CycleGAN (ours)	**0.52**	**0.17**	**0.11**
pix2pix [22]	0.71	0.25	0.18

Table 2: FCN-scores for different methods, evaluated on Cityscapes labels→photo.

Loss	Per-pixel acc.	Per-class acc.	Class IOU
CoGAN [32]	0.45	0.11	0.08
BiGAN/ALI [9, 7]	0.41	0.13	0.07
SimGAN [46]	0.47	0.11	0.07
Feature loss + GAN	0.50	0.10	0.06
CycleGAN (ours)	**0.58**	**0.22**	**0.16**
pix2pix [22]	0.85	0.40	0.32

Table 3: Classification performance of photo→labels for different methods on cityscapes.

Quantitative Results

Loss	Per-pixel acc.	Per-class acc.	Class IOU
Cycle alone	0.22	0.07	0.02
GAN alone	0.51	0.11	0.08
GAN + forward cycle	**0.55**	**0.18**	**0.12**
GAN + backward cycle	0.39	0.14	0.06
CycleGAN (ours)	0.52	0.17	0.11

Table 4: Ablation study: FCN-scores for different variants of our method, evaluated on Cityscapes labels→photo.

Loss	Per-pixel acc.	Per-class acc.	Class IOU
Cycle alone	0.10	0.05	0.02
GAN alone	0.53	0.11	0.07
GAN + forward cycle	0.49	0.11	0.07
GAN + backward cycle	0.01	0.06	0.01
CycleGAN (ours)	**0.58**	**0.22**	**0.16**

Table 5: Ablation study: classification performance of photo→labels for different losses, evaluated on Cityscapes.

Qualitative Results

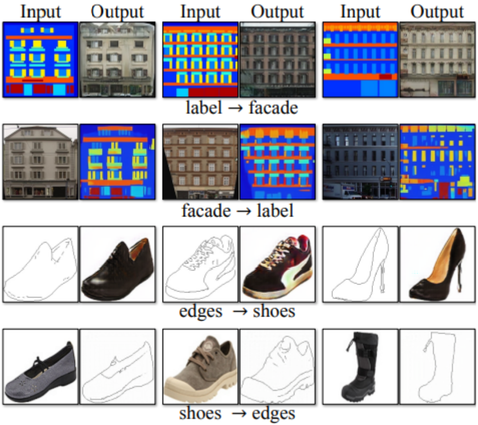


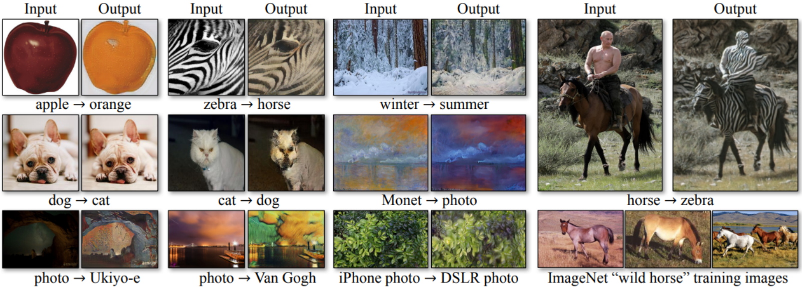
Figure 8: Example results of CycleGAN on paired datasets used in "pix2pix" [22] such as architectural labels↔photos and edges↔shoes.

Qualitative Results

Qualitative Results

Qualitative Results

Qualitative Results



References

- Jun-Yan Zhu*, Taesung Park*, Phillip Isola, and Alexei A. Efros. "Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks", in IEEE International Conference on Computer Vision (ICCV), 2017
- Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, Alexei A. Efros Image-to-Image Translation with Conditional Adversarial Networks, CVPR, 2017
- Arbitrary Style Transfer in Real-time with Adaptive Instance Normalization Xun Huang, Serge Belongie, ICCV 2017