Unpaired Image-to-Image Translation using Cycle-Consistency Adversarial Networks

CycleGAN - 논문 리딩 GAN 2팀 박지호 이민재 최명헌

목차

1. *Image-to-Image Translation Task*
2. *About GAN*
3. *Cycle GAN*

Image-to-Image Translation Task

$$G: X \rightarrow Y$$

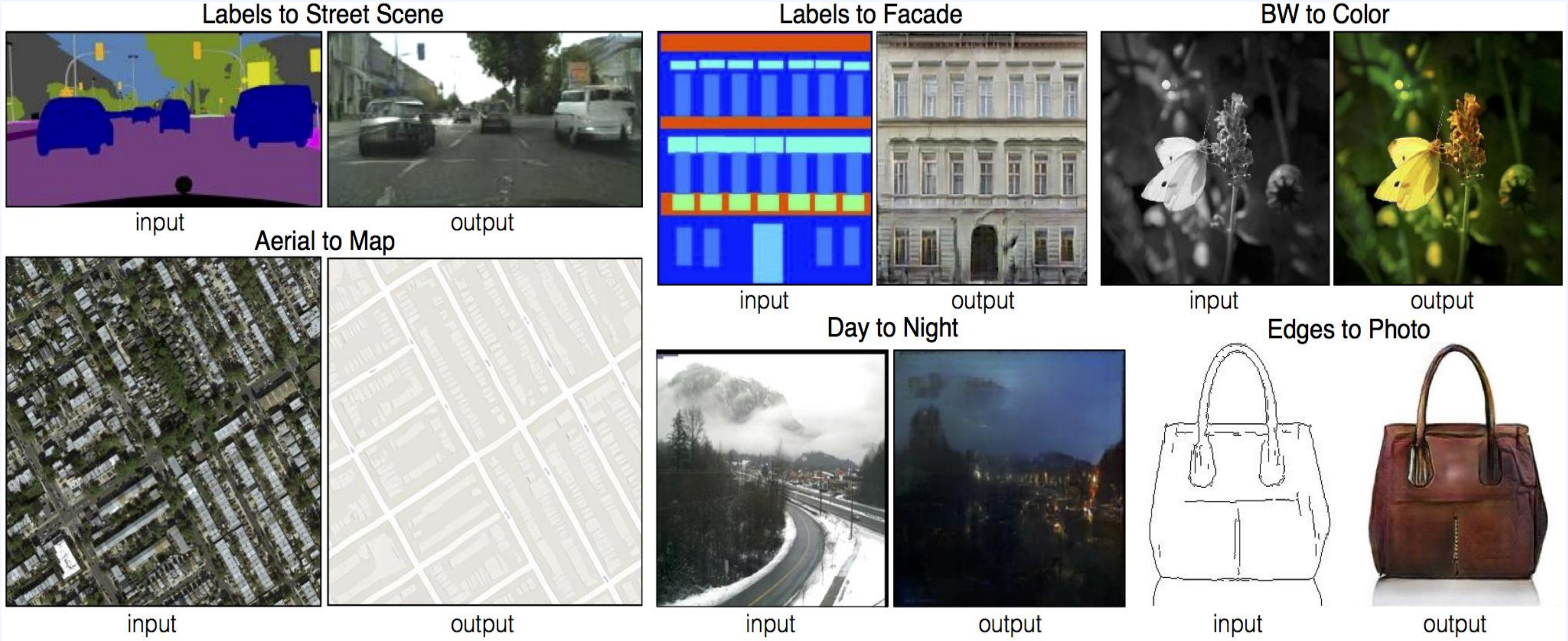
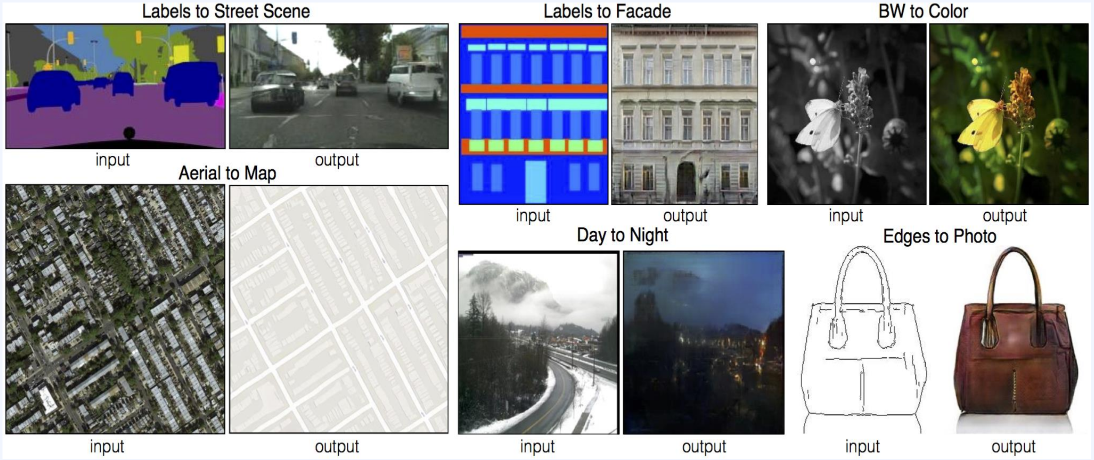


Image-to-Image Translation Task

$$G: X \rightarrow Y$$



Certain Solution

Using Paired Dataset(pix2pix)

$$\{x_{data}, y_{data}\}$$

Image-to-Image Translation Task

$G: X \to Y$

Certain Solution

Using Paired Dataset

$\{x_{data}, y_{data}\}$

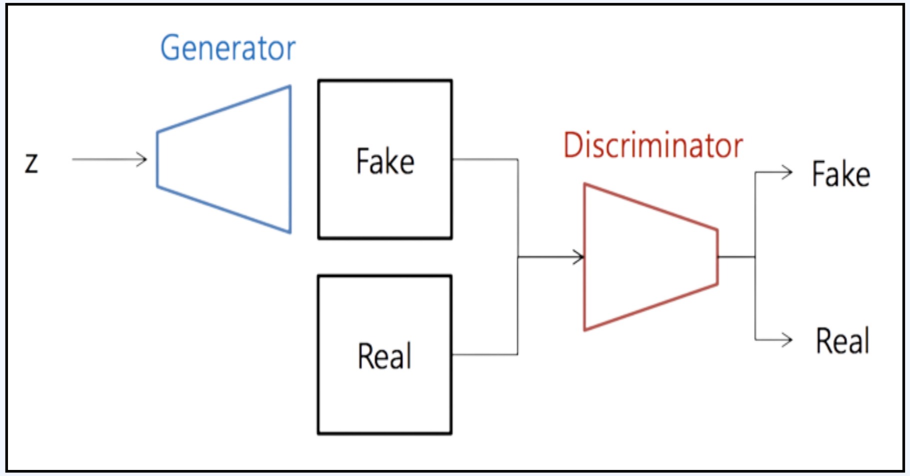
Paired Datasets are Rare and Expensive

Training with Unpaired Dataset

$$G: z \rightarrow y$$

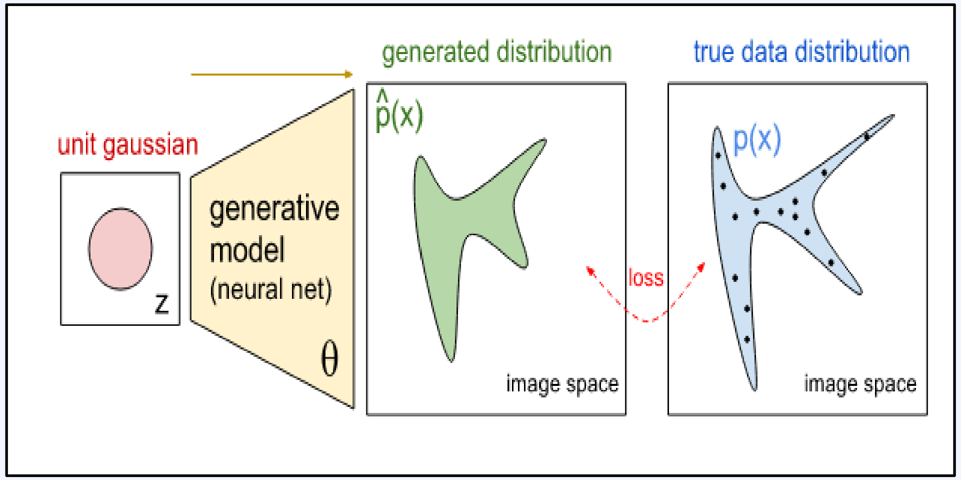
Objective(Loss) Function

$$\min_{G} \max_{D} V(G, D) = E_{y \sim p_{data}(y)}[logD(y)] + E_{z \sim p_{data}(z)}[\log(1 - D\,(G(z))]$$



Optimal Solution

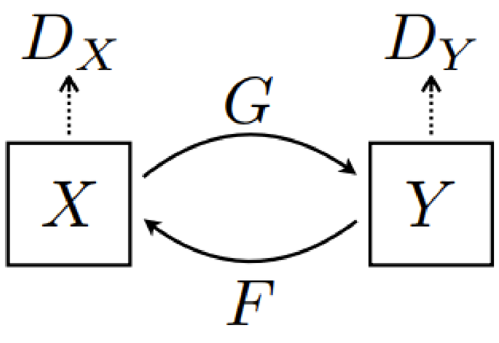
$$p_{generated} = p_{data}$$



GAN ⇒ Making an arbitrary Mapping Function G between two domains

Cycle Consistency

with 2 Translation(GAN) functions



(a)

$$G: X \rightarrow Y,\ D_Y$$

$$L_{GAN}(G, D_Y, X, Y) = E_{y \sim p_{data}(y)}[log D_Y(y)] + E_{x \sim p_{data}(x)}[\log(1 - D_Y(G(x))]$$

$$F: Y \rightarrow X,\ D_X$$

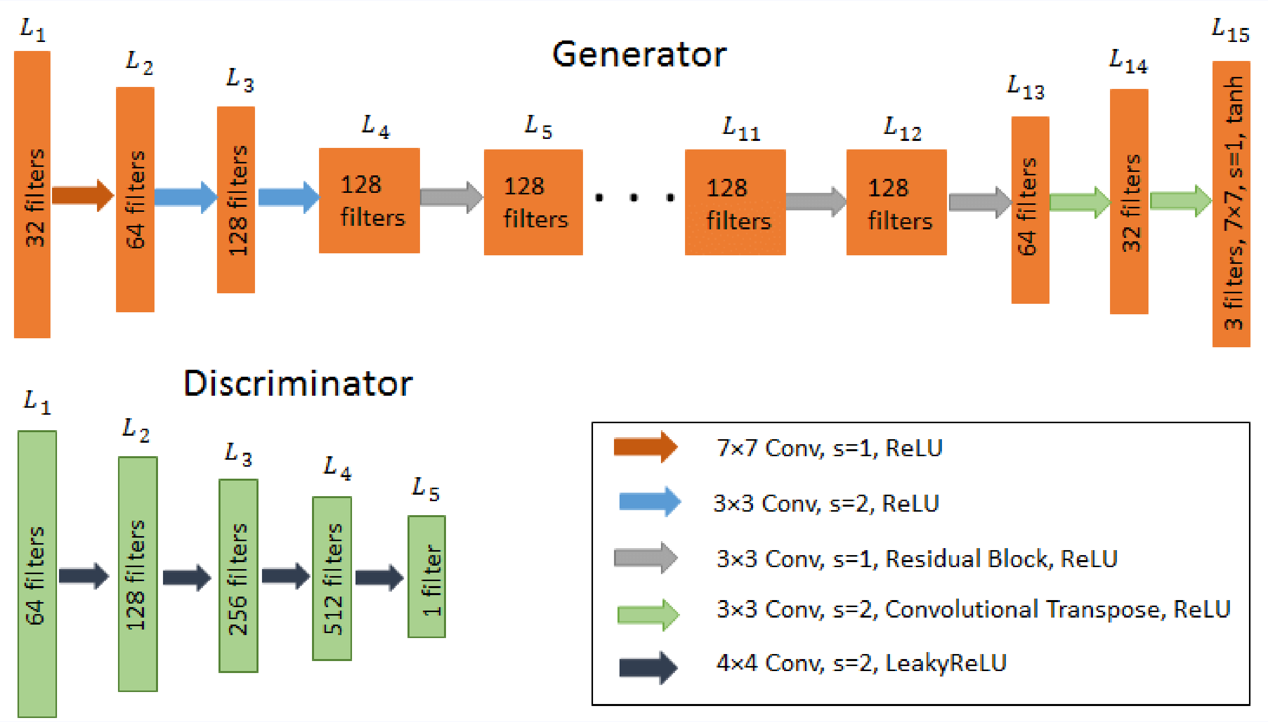
$$L_{GAN}(F, D_X, Y, X) = E_{x \sim p_{data}(x)}[log D_X(x)] + E_{y \sim p_{data}(y)}[\log(1 - D_X(F(y))]$$

Cycle Consistency

with 2 Translation(GAN) functions

Generator
- *Convolutions*
- *Residual Blocks*

Discriminator(from PatchGAN)
- *Convolutions*
- *Discriminating every 70x70 Patches*



Cycle Consistency

with 2 Translation(GAN) functions

$$G : X \rightarrow Y, \quad F : Y \rightarrow X$$

Cycle Consistency

$$F\big(G(x)\big) \approx x$$

$$G\big(F(y)\big) \approx y$$

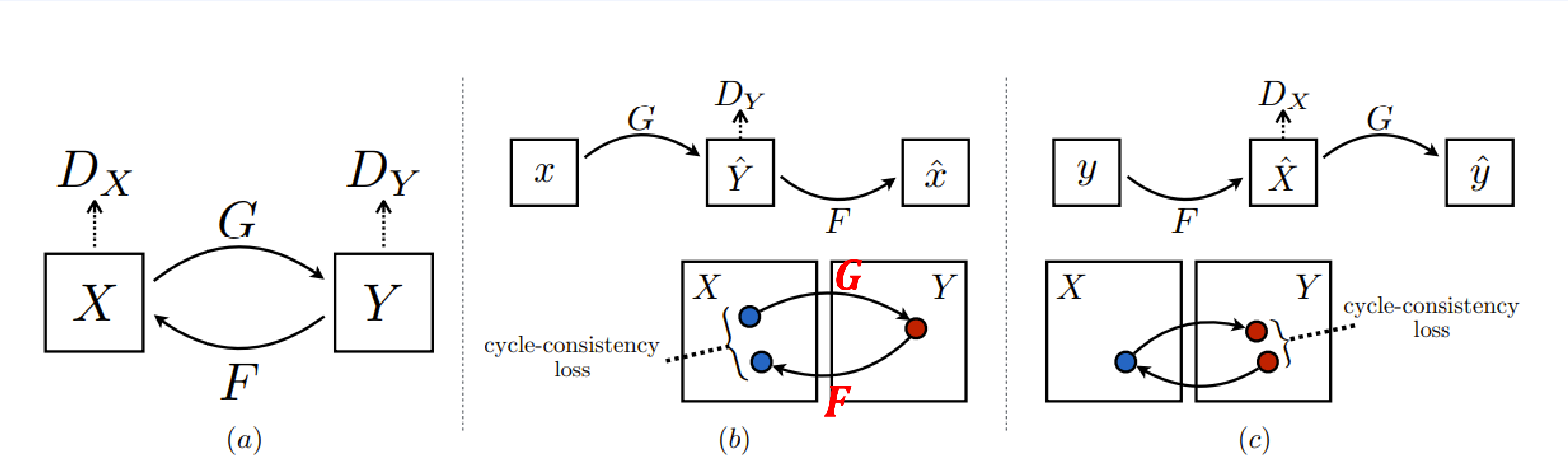
Input x Output $\mathrm{G}(x)$ Reconstruction $\mathrm{F}(\mathrm{G}(x))$

To make Reconstruction easy,
G remains the content of the input

Cycle Consistency Loss

with 2 Translation(GAN) functions

$$L_{cyc}(G,F) = E_{x\sim p_{data}(x)}\left[\left\|F\big(G(x)\big) - x\right\|_1\right] + E_{y\sim p_{data}(y)}\left[\left\|G\big(F(y)\big) - y\right\|_1\right]$$



3. CycleGAN: Loss Equation

Objective Function

$$G : X \rightarrow Y, \qquad L_{GAN}(G, D_Y, X, Y) = E_{y\sim p_{data}(y)}[log D_Y(y)] + E_{x\sim p_{data}(x)}[\log(1 - D_Y(G(x))]$$

$$F : Y \rightarrow X, \qquad L_{GAN}(F, D_X, Y, X) = E_{x\sim p_{data}(x)}[log D_X(x)] + E_{y\sim p_{data}(y)}[\log(1 - D_X(F(y))]$$

$$Cycle\ Consistency \quad L_{cyc}(G, F) = E_{x\sim p_{data}(x)}\left[\left\|F(G(x)) - x\right\|_1\right] + E_{y\sim p_{data}(y)}\left[\left\|G(F(y)) - y\right\|_1\right]$$

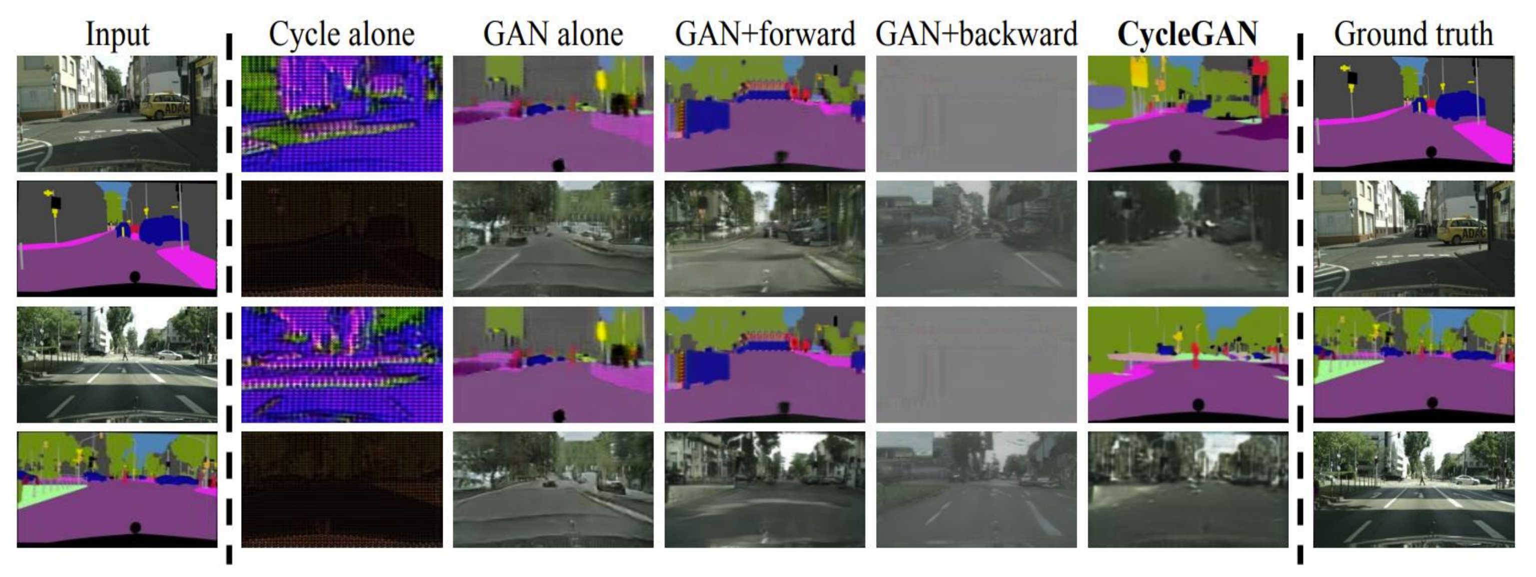
$$L(G, F, D_X, D_Y) = \underbrace{L_{GAN}(G, D_Y, X, Y) + L_{GAN}(F, D_X, Y, X)}_{\text{Remains the Domain of the generated image}} + \underbrace{\lambda L_{cyc}(G, F)}_{\text{Remains the content of input}}$$

$$G^*, F^* = \arg \min_{G, F} \max_{D_x, D_y} L(G, F, D_X, D_Y)$$

3. CycleGAN: Loss Analysis

Loss Analysis

- Gan+forward-cycle: Higher FCN-score but mode collapse, unstable
- Conclusion: GAN loss & Cycle loss are both critical



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

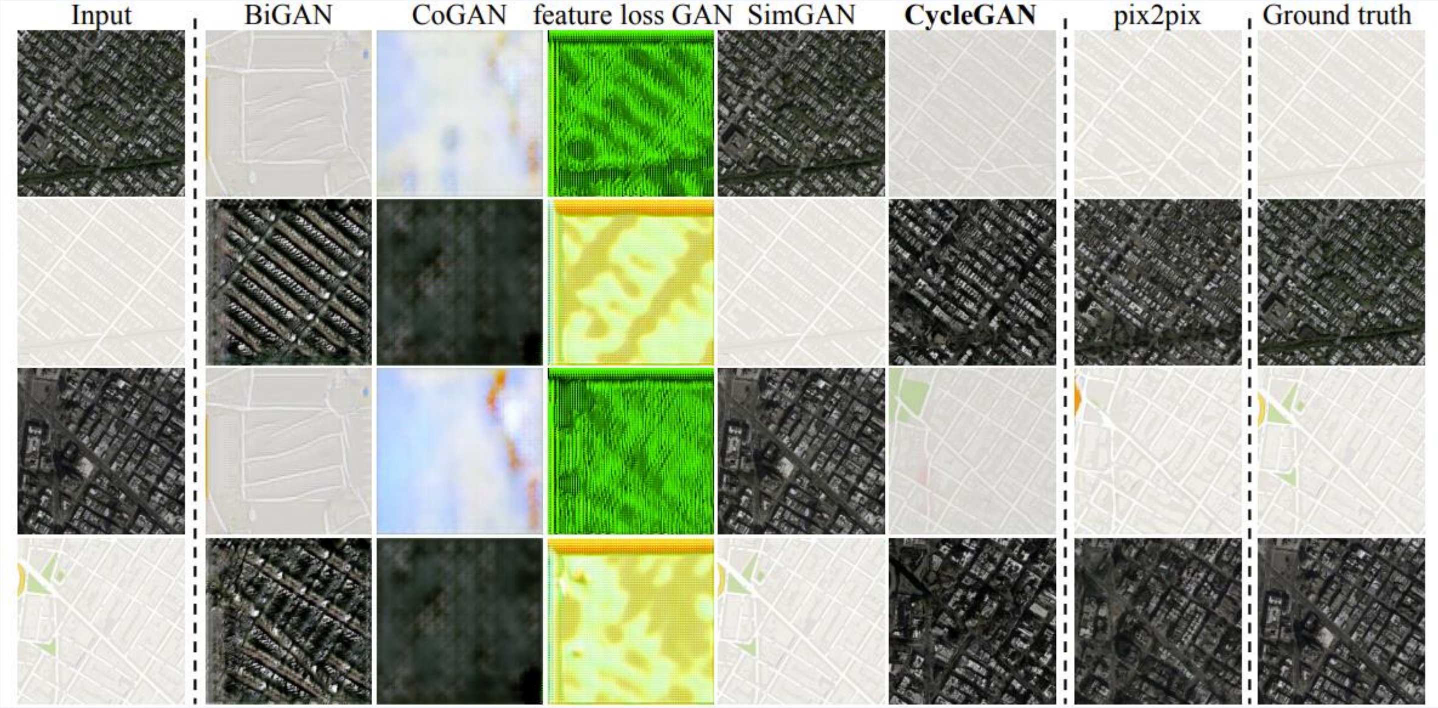
Table 4: Ablation study: FCN-scores for different variants of our method, evaluated on Cityscapes labels→photo.

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.

Advantages

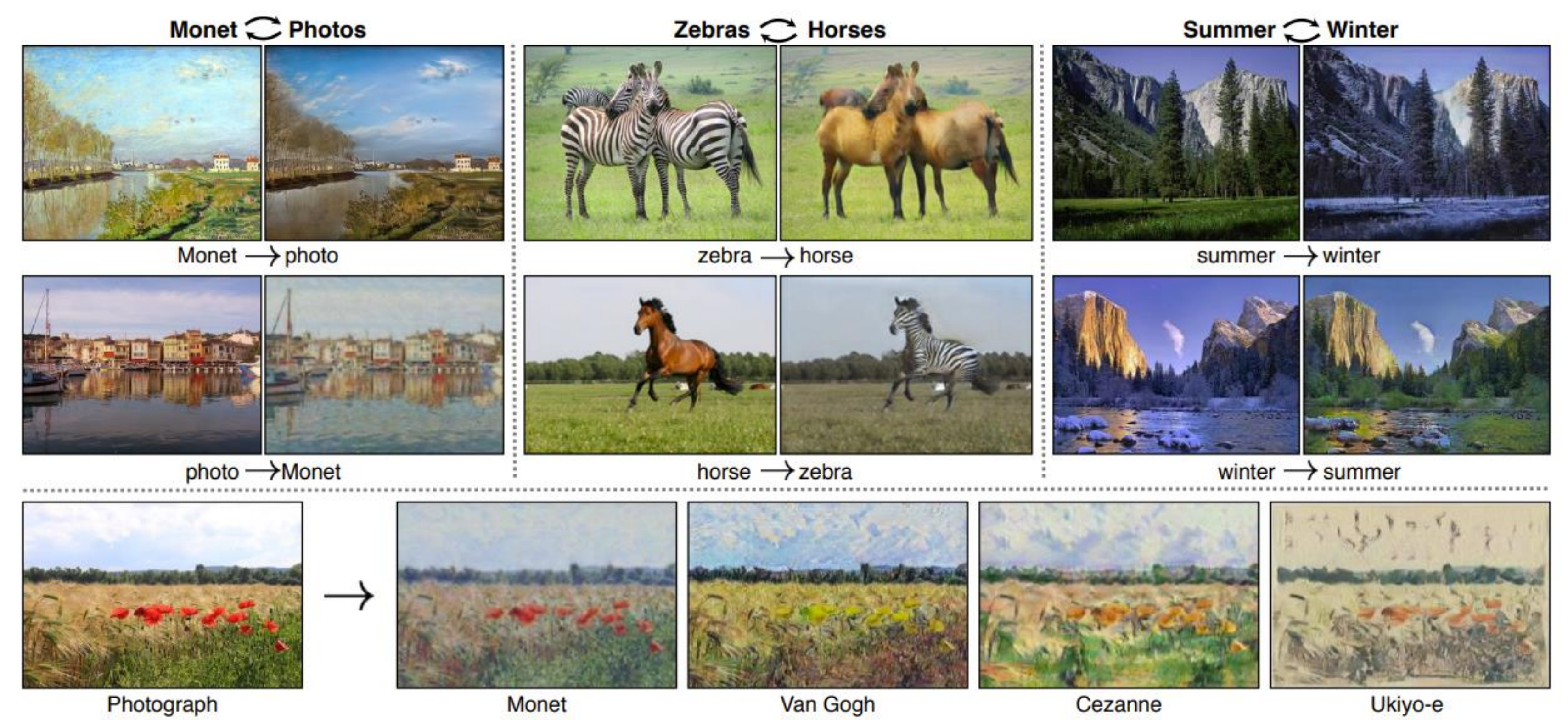
- Good Performance with Unpaired Training
- Appliable on Various Tasks
  (ex. style transfer, photo enhancement, …)



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

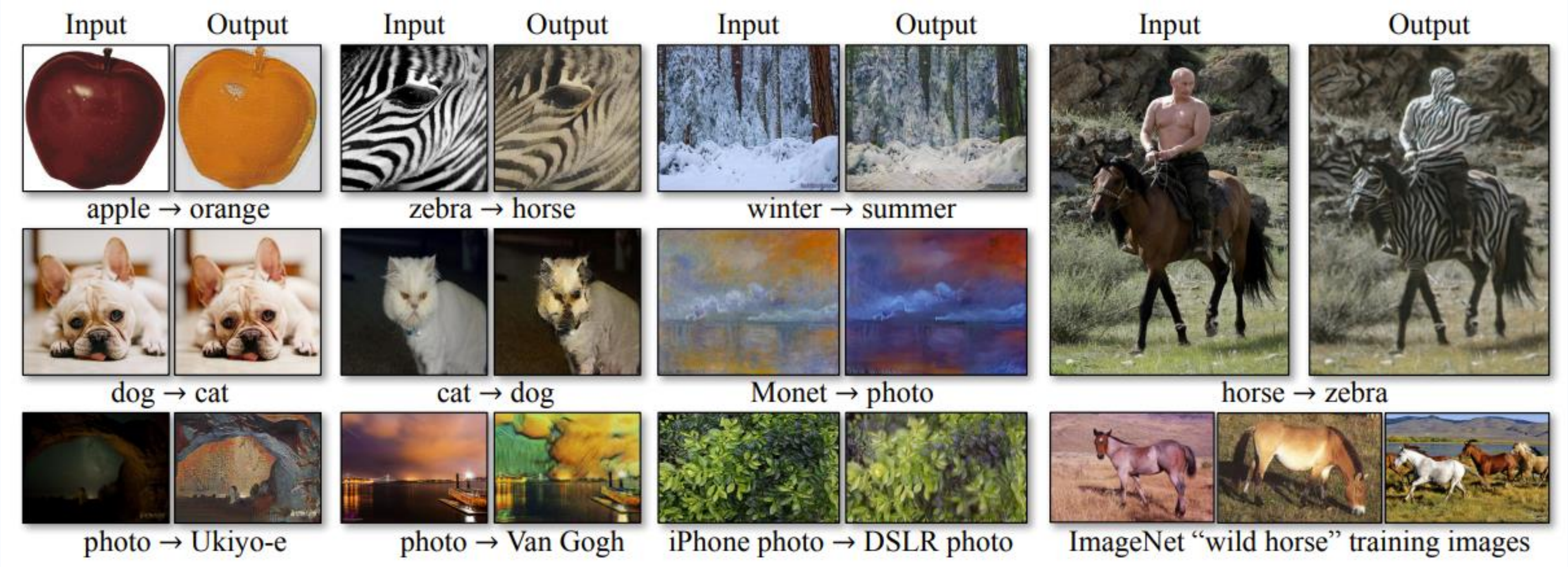
Advantages

- Good Performance with Unpaired Training
- Appliable on Various Tasks
  (ex. style transfer, photo enhancement, ...)



Limitation

- Bad with Geometric Changes
- Bad with Untrained Images
- Loss of Color Information

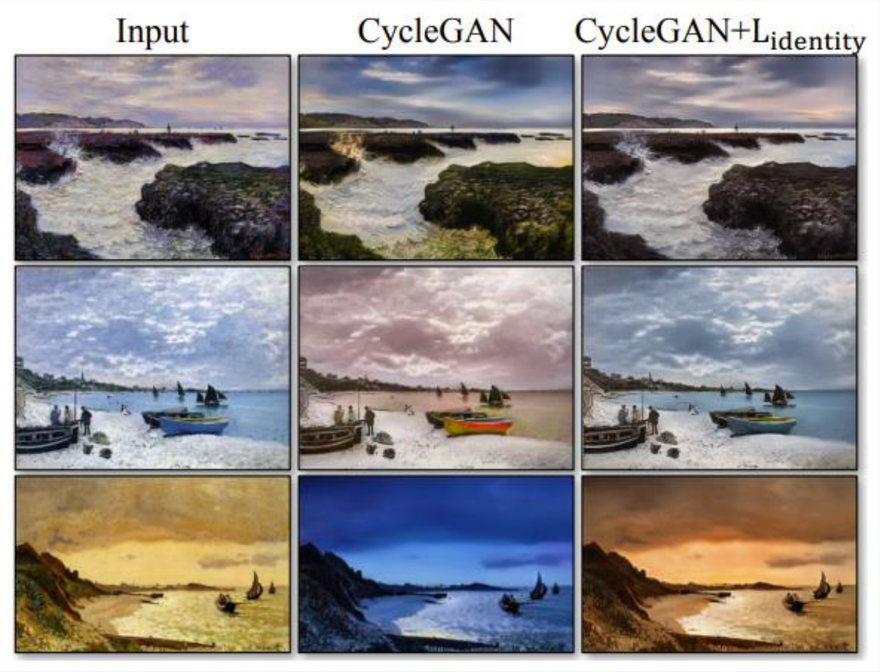


Limitation

- Bad with Geometric Changes
- Bad with Untrained Images
- Loss of Color Information(for painting→photo task)

Identity Loss

$$L_{identity}(G,F) = E_{x\sim p_{data}(x)}[\|G(x)-x\|_1] + E_{y\sim p_{data}(y)}[\|F(y)-y\|_1]$$



Thank You!

Figure 10: Collection style transfer I: we transfer input images into the artistic styles of Monet, Van Gogh, Cezanne, and Ukiyo-e. Please see our website for additional examples.

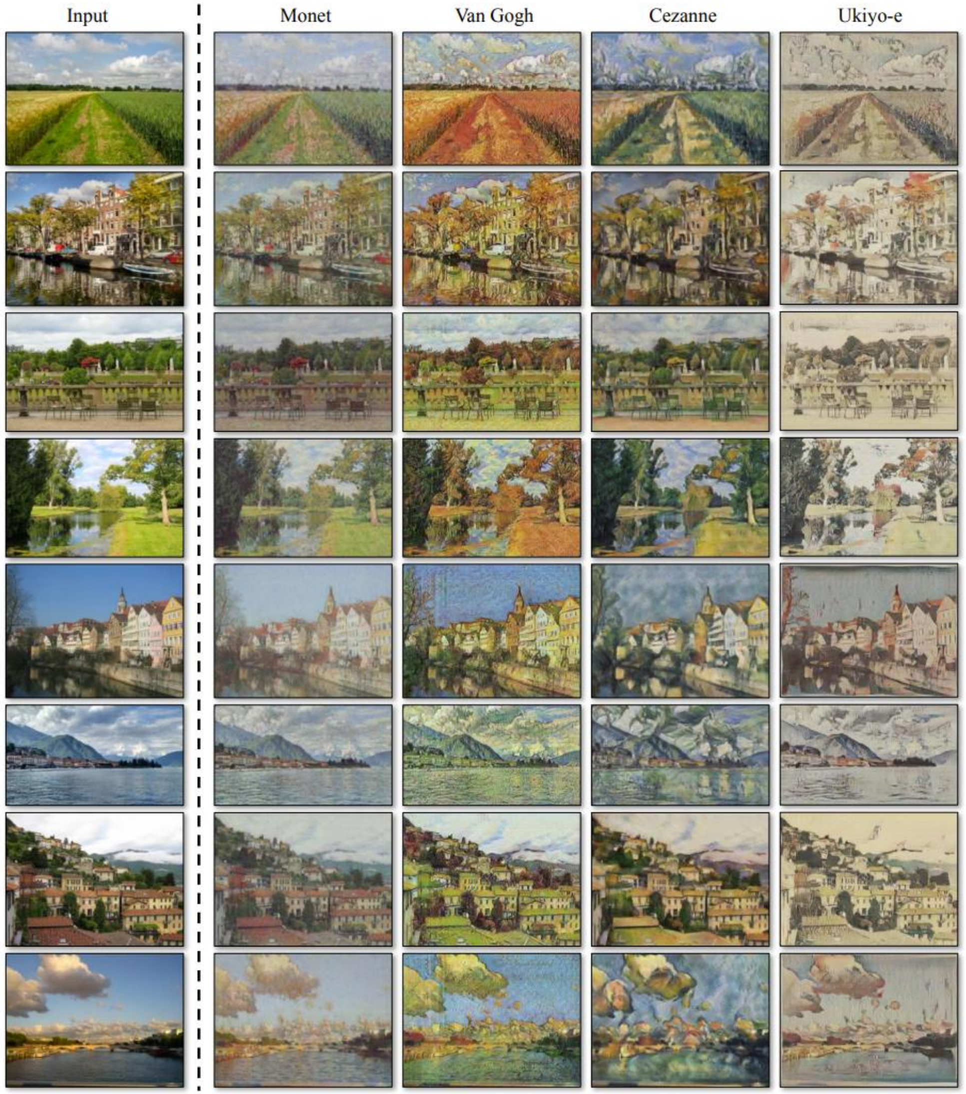


Figure 11: Collection style transfer II: we transfer input images into the artistic styles of Monet, Van Gogh, Cezanne, Ukiyo-e. Please see our website for additional examples.

Thank You!

Figure 12: Relatively successful results on mapping Monet's paintings to a photographic style. Please see our website for additional examples.