

Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks (2017)

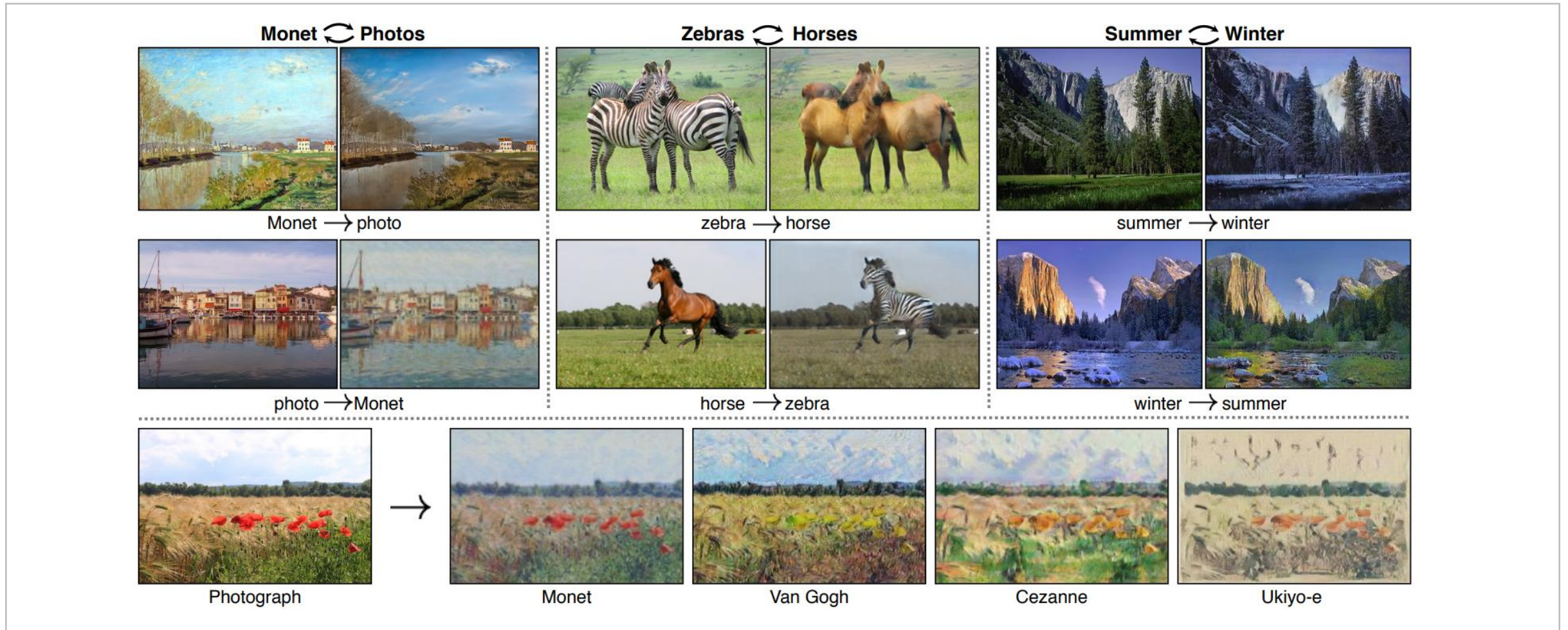
Jun-Yan Zhu*, Taesung Park*, Phillip Isola, Alexei A. Efros
Berkeley AI Research (BAIR) laboratory, UC Berkeley

II E8557-01 동적계획법과 강화학습

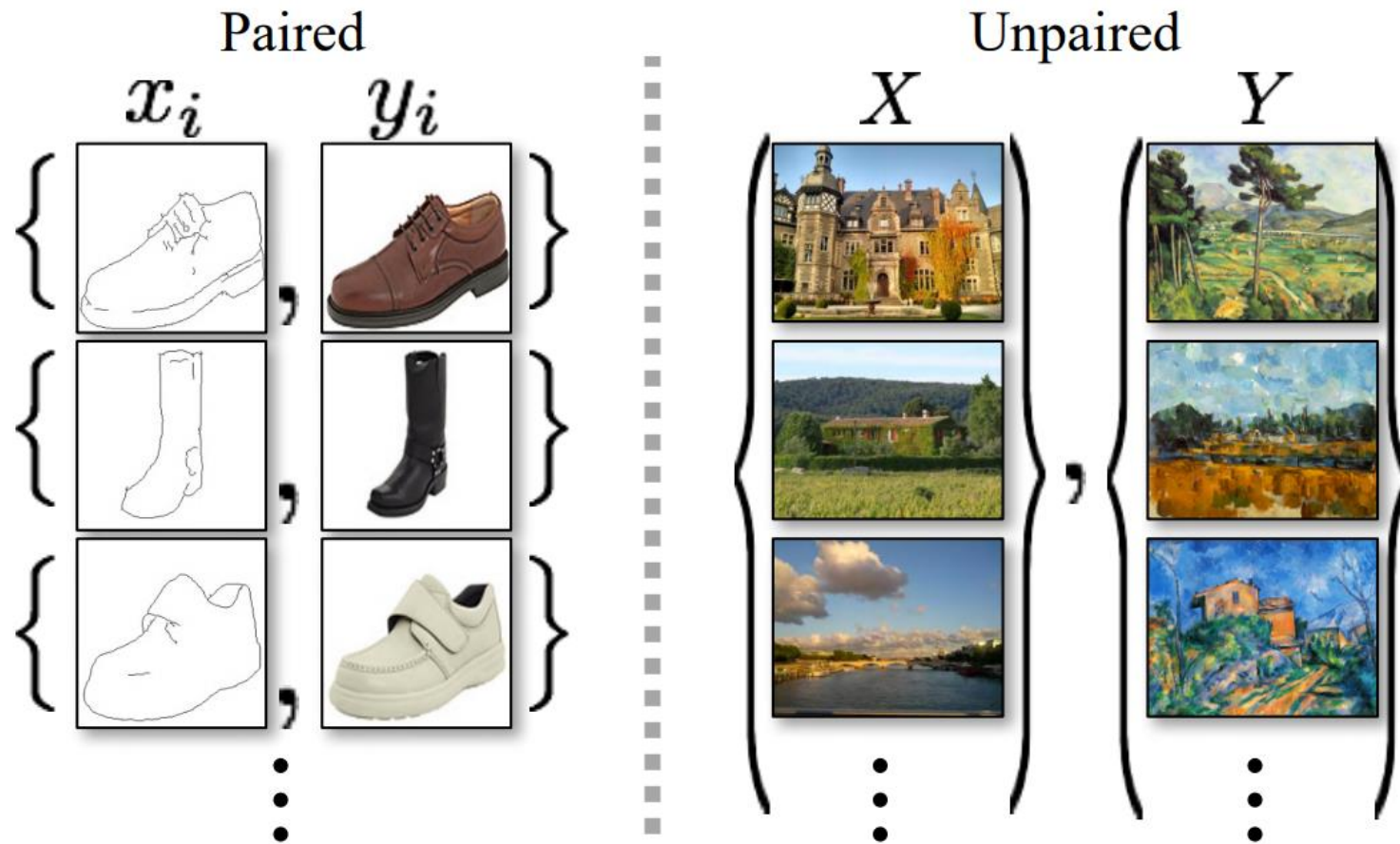
경영과학연구실 김윤석

Image to Image Translation

- Image to Image Translation은 이미지를 다른 이미지로 변환하는 작업임
- 예시 및 응용 분야에는 스타일 변환, 세그멘테이션, 스케치에서 사진으로, 낮에서 밤 사진으로, 얼굴 변환 등이 있음



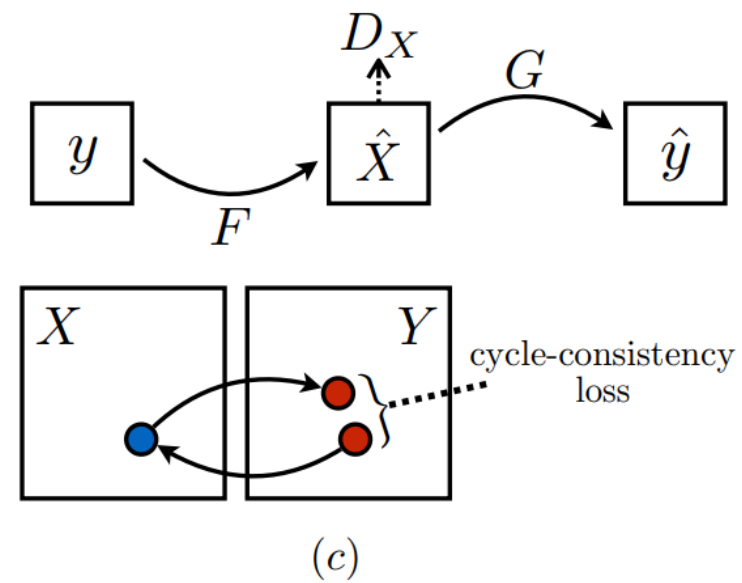
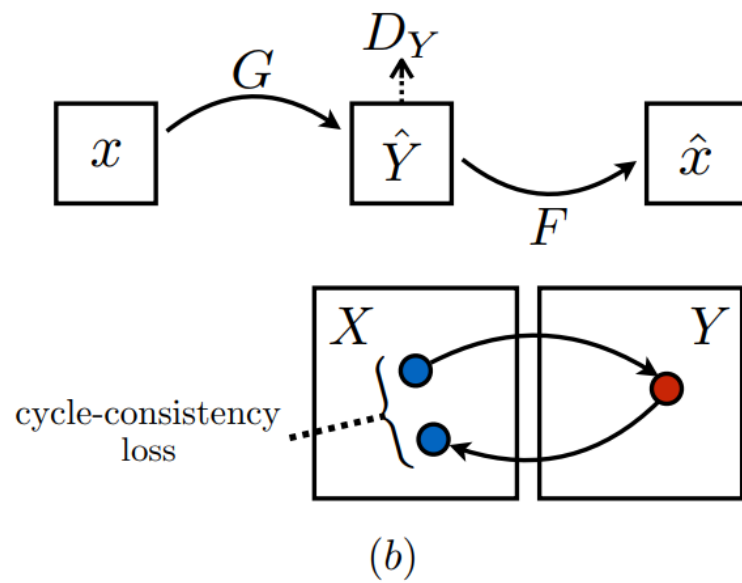
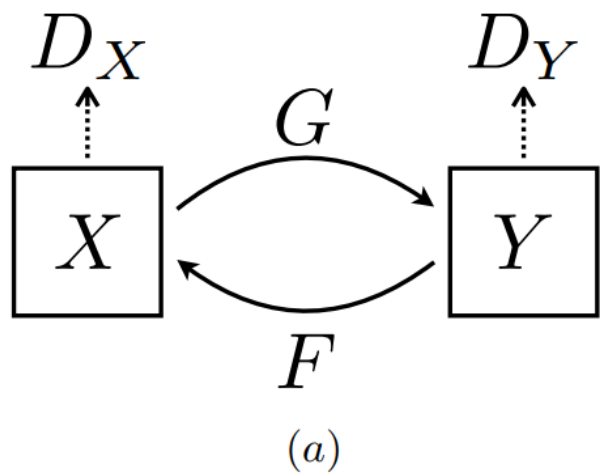
- Image to Image translation은 입력 이미지와 출력 이미지 간의 매핑을 학습하는 것이 목표임
- 학습을 위해 정렬된 이미지 쌍의 training set을 필요로 함
- 그러나 많은 작업에서는 쌍을 이루는 training set이 없음



CycleGAN은 쌍을 이루는 training set이 없는 상황에서 매핑을 학습하는
방법에 대해 연구함

- X 에서 Y 로 매핑하는 적대적 생성 신경망과 Y 에서 X 적대적 신경망 그리고 Cycle Consistency Loss를 제안함
- 제안한 방법을 통해 순환 일관성을 갖게 함

Cycle Consistency



- $F: Y \rightarrow X$ 와 $G: X \rightarrow Y$ 에 대한 Adversarial Loss와 Cycle Consistency Loss를 사용하여 Loss function을 구성
- Adversarial Loss는 GAN의 Loss로 손실 함수는 minimize G, maximize D를 목적으로 학습

1. Adversarial Loss

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

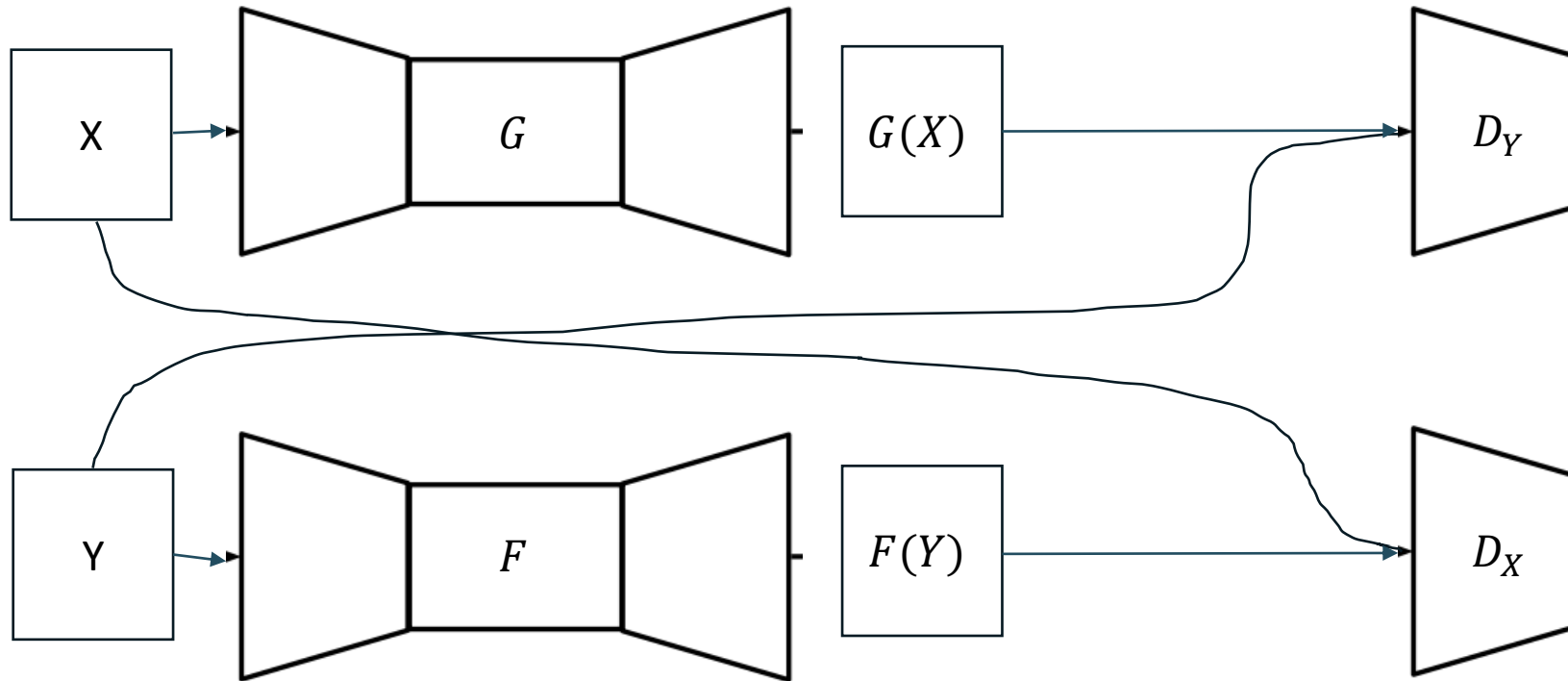
2. Cycle Consistency Loss

$$\begin{aligned}\mathcal{L}_{\text{cyc}}(G, F) &= \mathbb{E}_{x \sim p_{\text{data}}(x)} [\|F(G(x)) - x\|_1] \\ &\quad + \mathbb{E}_{y \sim p_{\text{data}}(y)} [\|G(F(y)) - y\|_1].\end{aligned}$$

3. Full Objective (Loss)

$$\begin{aligned}\mathcal{L}(G, F, D_X, D_Y) &= \mathcal{L}_{\text{GAN}}(G, D_Y, X, Y) \\ &\quad + \mathcal{L}_{\text{GAN}}(F, D_X, Y, X) \\ &\quad + \lambda \mathcal{L}_{\text{cyc}}(G, F),\end{aligned}$$

- CycleGAN은 두개의 생성자 F, G 와 두 개의 판별자 D_Y, D_X 로 구성되어 있음
- CycleGAN은 잠재 벡터를 사용하지 않는 특징을 갖고 있음



- 실험은 평가를 위해 입력-출력 쌍이 사용 가능한 쌍의 데이터 셋을 사용함
- Cityscapes 데이터 셋을 사용했으며, semantic label 이미지와 지도-항공 이미지를 사용함

Baseline

- **CoGAN:** 두 개의 연결된 생성적 적대 신경망을 사용하여 한 도메인의 데이터를 다른 도메인의 데이터로 변환하는 모델이
- **SimGAN:** 적대적 손실을 사용하여 도메인 x 에서 y 로의 이미지 변환을 학습하는 모델이며, 픽셀 레벨에서 큰 변화를 줄이기 위해 정규화 항을 사용하여 모델을 훈련함
- **Feature loss + GAN:** SimGAN의 변형으로, RGB 픽셀 값 대신 사전에 훈련된 네트워크(VGG-16 relu4 2)를 사용하여 깊은 이미지 특징 위에서 L1 손실을 계산함
- **BiGAN/ALI:** 무작위 노이즈를 이미지로 변환하는 생성자와 이미지를 무작위 노이즈로 역변환하는 인코더를 동시에 학습하여, 양방향 매핑을 가능하게 하는 모델임
- **pix2pix:** 쌍으로 이루어진 데이터(입력 이미지와 대응하는 출력 이미지)를 기반으로 이미지-대-이미지 변환을 학습하는 조건적 생성적 적대 네트워크(Conditional Generative Adversarial Network, cGAN) 모델임

- AMT test table은 사람들을 이용해 현실감을 평가한 결과임
- FCN-Score는 생성된 이미지가 완전 합성곱 네트워크(FCN)에 의해 얼마나 잘 해석되는지를 측정한 평가 지표임
- Segmentation image에 대한 성능 평가를 실험함
- CycleGAN은 전체적인 실험에서 좋은 성능을 보임, Table3에서 pix2pix가 더 좋은 이유는 학습 데이터 쌍이 잘 이루어진 데이터이기 때문임

Loss	Map → Photo		Photo → Map	
	% Turkers labeled <i>real</i>		% Turkers labeled <i>real</i>	
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.

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.

- CycleGAN은 두 개의 다른 도메인 간에 이미지를 변환하는 데 사용됨
- Transfer Learning은 한 작업에서 학습한 정보를 다른 작업에서 적용하는 기술로 한 도메인에서 훈련된 모델의 가중치를 다른 도메인의 모델 학습에 사용이 가능함
- CycleGAN은 도메인의 스타일 또는 특성을 다른 도메인으로 Transfer함, 이 점에서 Transfer Learning의 큰 범주 안에 들어감
- CycleGAN의 학습 방식에 대해서 잘 생각해 보면 Transfer Learning에 사용할 수 있을 것으로 생각됨

Q & A