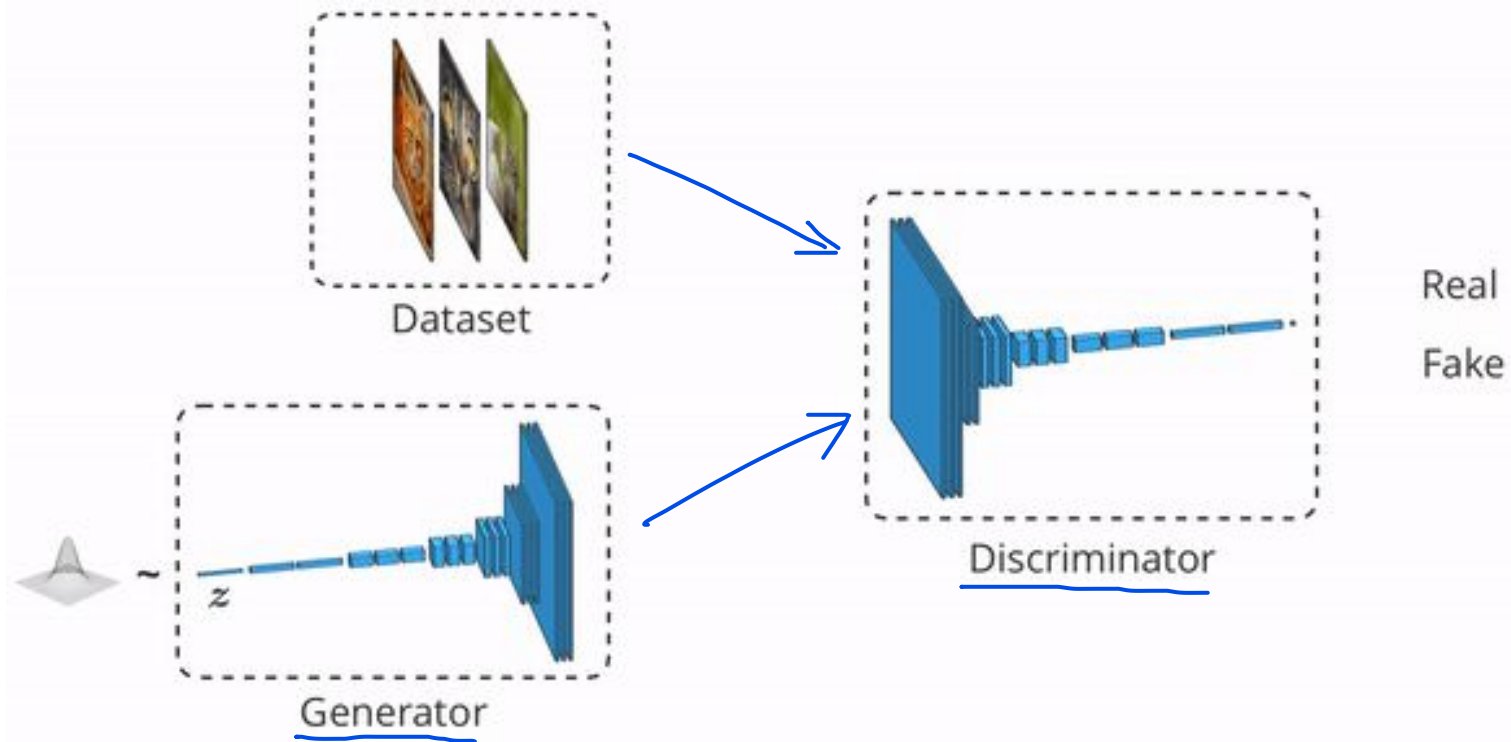


# Cool GAN Applications

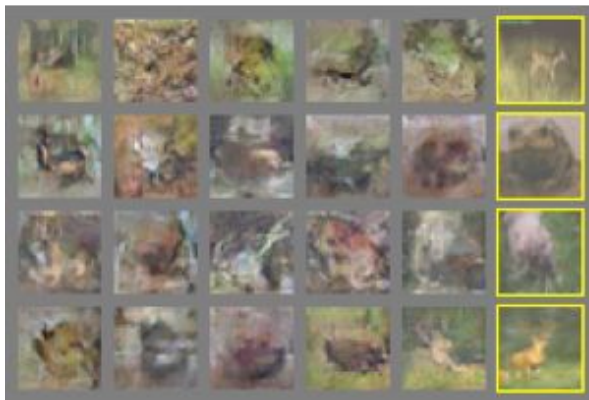
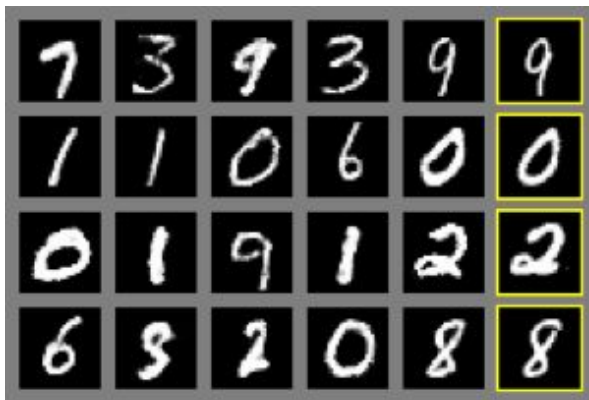


# The original GAN



노랑: 훈련자료

나머지: 결과



Goodfellow's GAN(2014)은 MNIST, Toronto Face Dataset(TFD) 및 [CIFAR-10](#) 데이터 세트를 모방합니다. 마지막 열은 픽셀 단위 유클리드 거리로 마지막에서 두 번째 열에 가장 가까운 훈련 이미지를 보여줍니다.

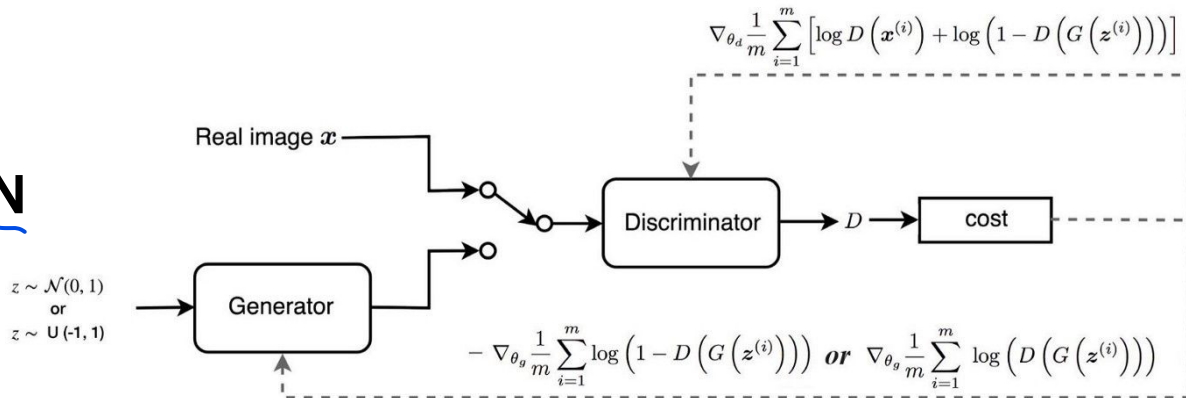
생성적 적대 신경망은 어렵고 불안정한 훈련으로 인해 다소 악명이 높습니다. 한 가지 문제는 훈련 중 생성자가 판별자보다 너무 앞서 나아갈 경우, 판별자를 가장 잘 속일 수 있는 몇 개만 생성하여, 이미지가 다양하게 생성되지 않는다는 것입니다. 실제로, 이 문제를 모드 축소/붕괴(mode collapse) 라고합니다.



이미지 생성을 위한 강력한 유도 바이어스가 부족하도록 구성된 4 개의 레이어, 512 개 유닛 및 ReLU 활성화 기능을 갖춘 MLP 네트워크로 훈련된 결과로, 상당한 수준의 모드 축소/붕괴 (mode collapse)를 보여줍니다. 이 침실 이미지 중 많은 부분이 서로 매우 비슷하게 보입니다.

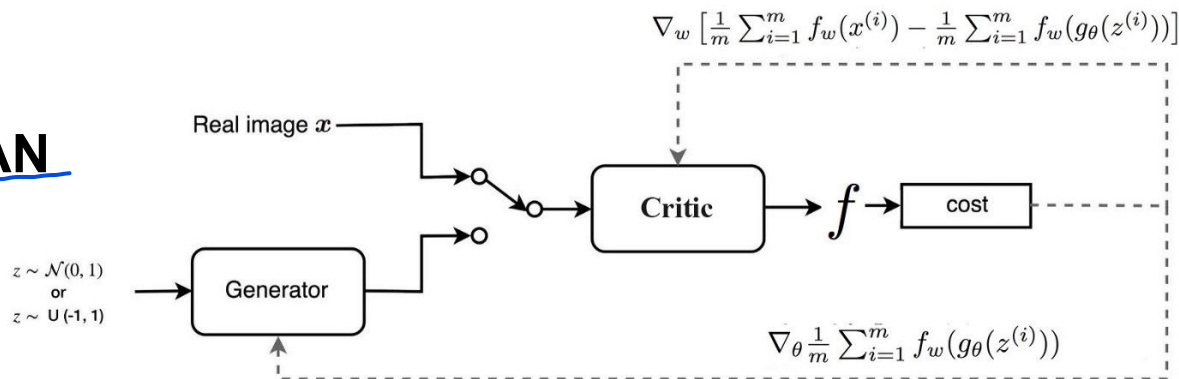
# WGAN (Wasserstein GAN) 해결책

GAN



VS.

WGAN



WGAN

**cost function**을 재정의하여 안정성을 높입니다.



다음과 같은 손실함수  $f()$ 를 사용:

$f$  : a 1-Lipschitz function as follows

$$w \leftarrow w + \alpha \cdot \text{RMSProp}(w, g_w)$$

$$w \leftarrow \text{clip}(w, -c, c)$$





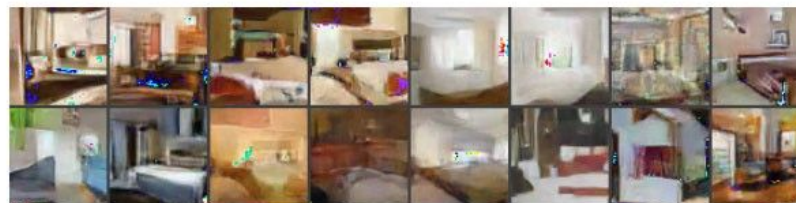
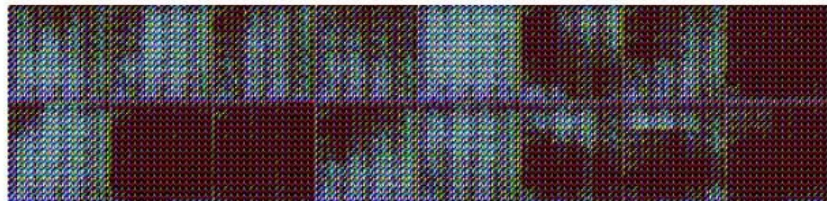
WGAN with DCGAN generator



GAN with DCGAN generator



**Without batch normalization &  
constant number of filters at each layer**



**Using a MLP as the generator**

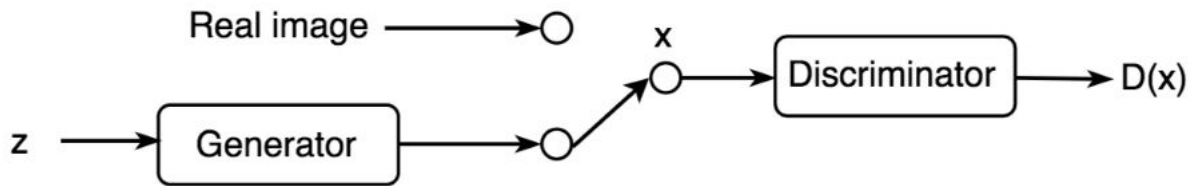


**All critics and discriminators follow the same discriminator design in DCGAN**

# CGAN (Conditional GAN)

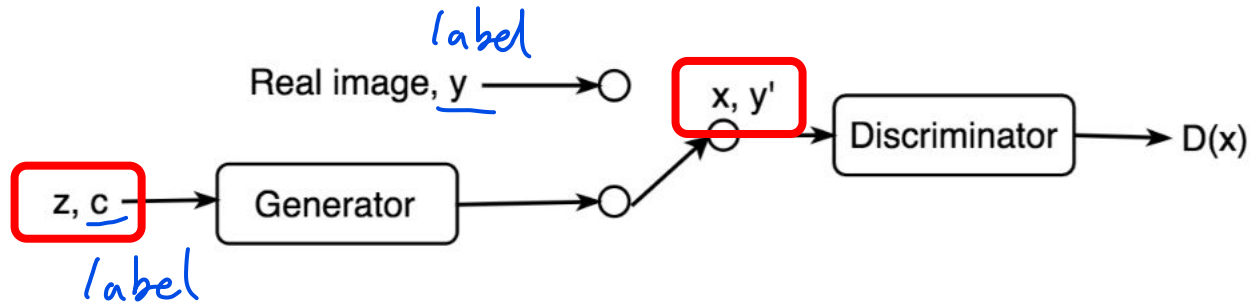
CGAN (Conditional GAN)에서 레이블은 이미지를 보다 잘 생성하고 식별하기 위해 잠재 공간  $z$ 의 확장으로 사용됩니다. 랜덤하게 Generation하는 것이 아닌, 내가 구체적으로 원하는 것을 generation할 수 있게 됩니다.

GAN



vs.

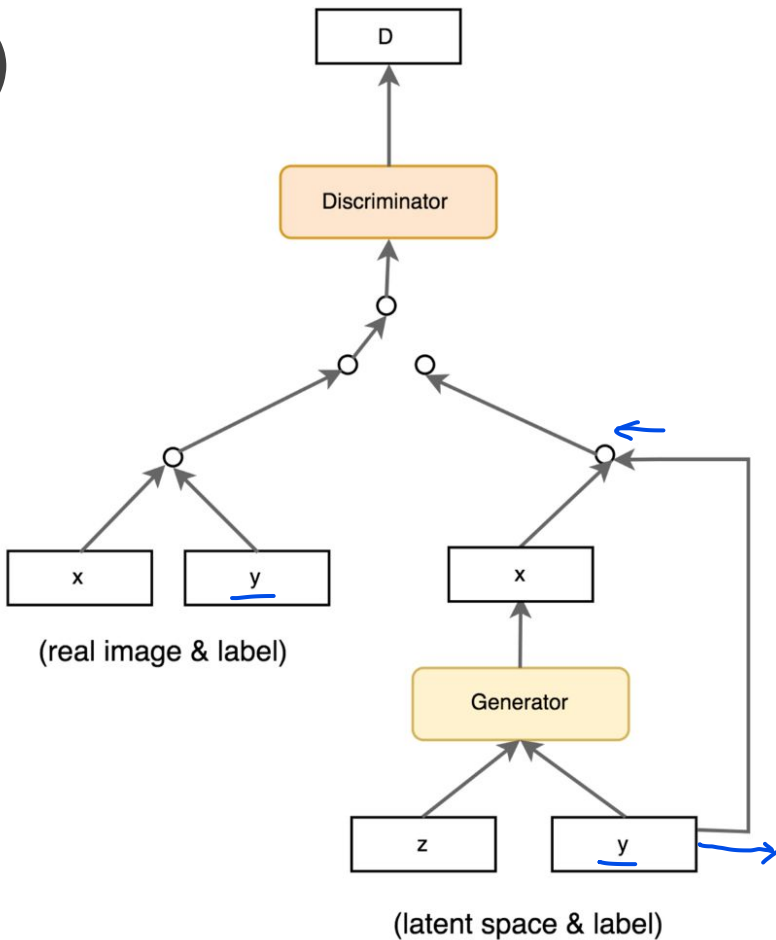
CGAN



# CGAN (Conditional GAN)

GAN에서는 생성 할 데이터 모드를 제어 할 수 없습니다. 조건부 GAN은 레이블  $y$  를 생성기에 추가 매개 변수로 추가하여 변경하고, 해당 이미지가 생성되기를 희망합니다. 또한 실제 이미지를 더 잘 구별하기 위해 판별기 입력에 레이블을 추가합니다.

MNIST에서는 균일 분포에서 레이블  $y$  를 샘플링하여 0에서 9까지의 숫자를 생성합니다. 이 값을 1-hot-vector 로 인코딩합니다. 예를 들어, 값 3은 (0, 0, 0, 1, 0, 0, 0, 0, 0) 으로 인코딩됩니다.





# the Result of CGAN for MNIST



# the Result of CGAN for Fashion MNIST

무작위



Unconditional GAN

생략됨



Conditional GAN

# ProGAN (Progressive growing of GANs) [[demo](#)]

고해상도 이미지를  
생성하기 위해서,

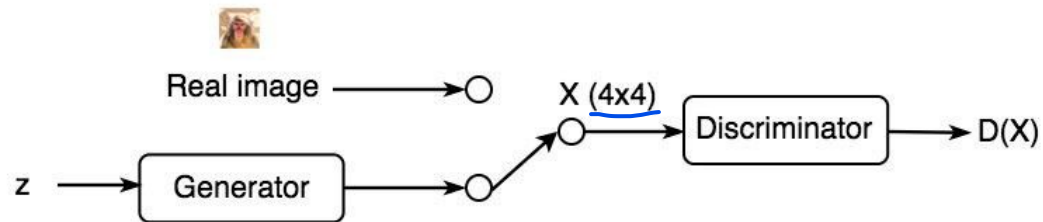
ProGAN은

저해상도에서  
네트워크 교육을  
시작하고 원하는  
해상도에 도달 할  
때까지 점점 더 많은  
레이어를  
추가합니다.

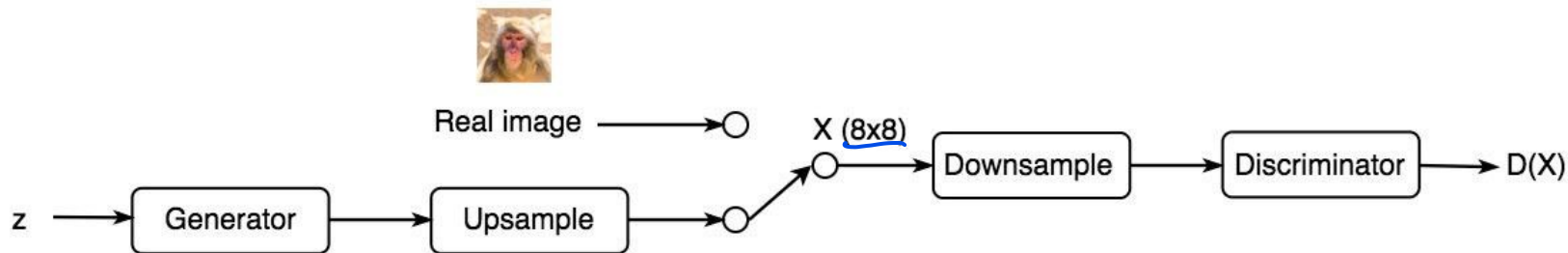


Figure 5:  $1024 \times 1024$  images generated using the CELEBA-HQ dataset. See Appendix F for a larger set of results, and the accompanying video for latent space interpolations.

Phase 1



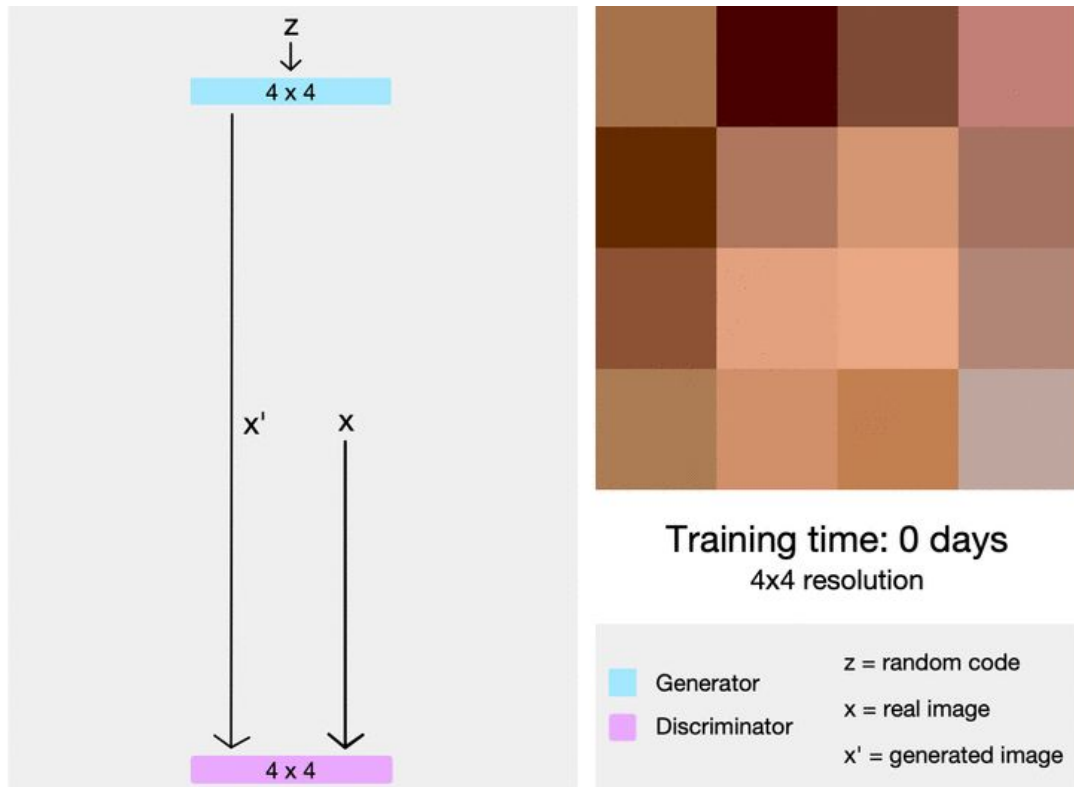
Phase 2



new generator

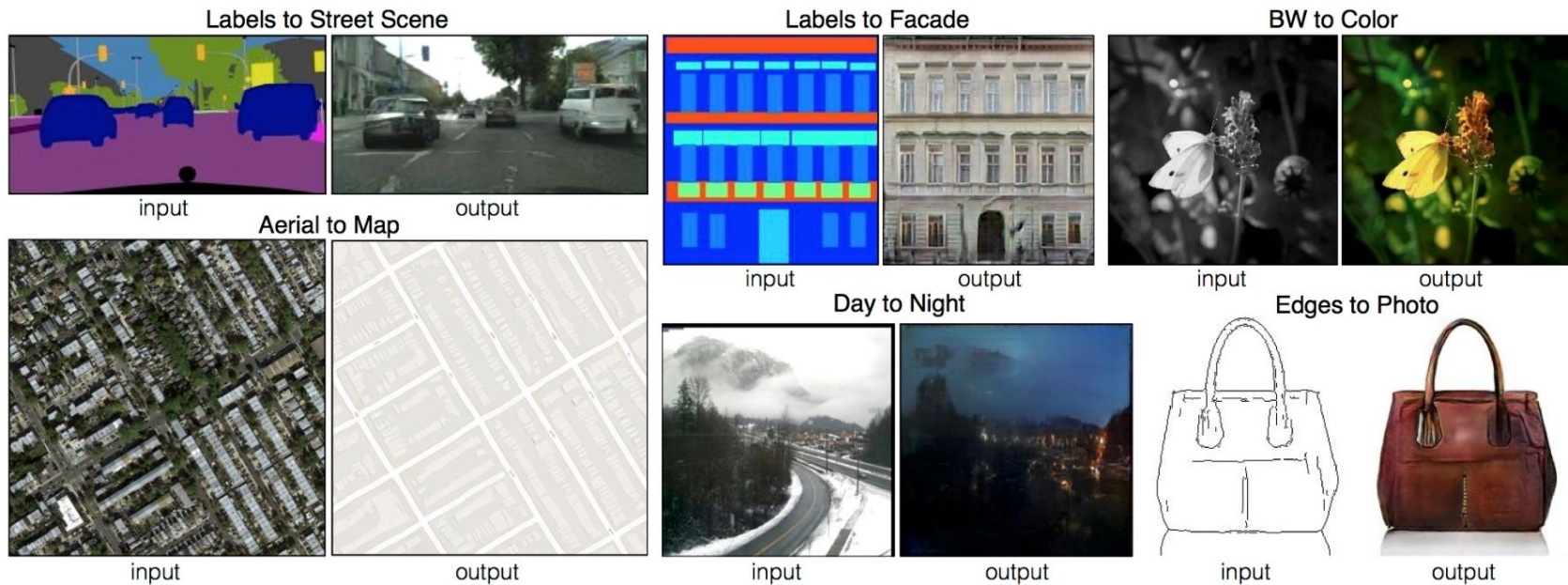
new discriminator

Phase K



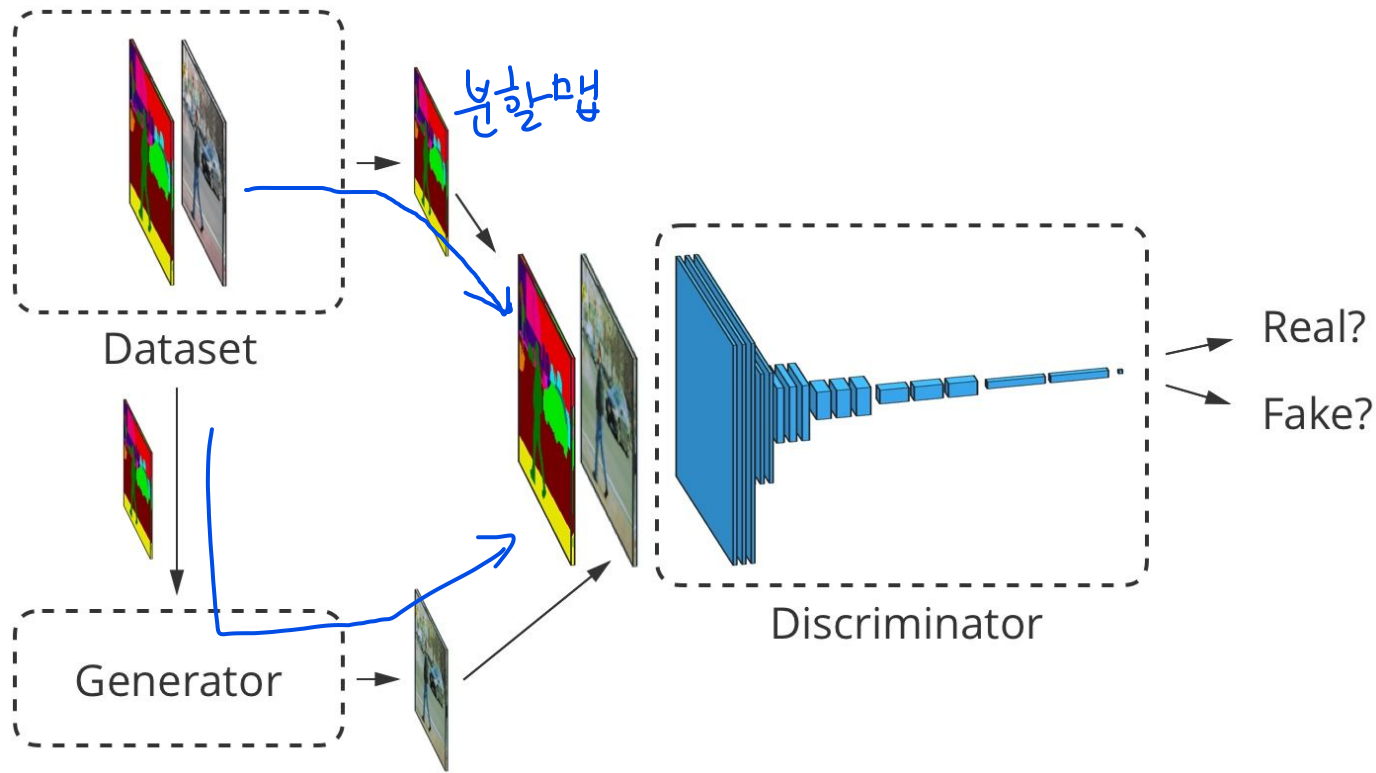
GAN 훈련 중에 고해상도 레이어를 점진적으로 추가하면 **훈련 안정성, 속도 및 결과 이미지 품질이 크게 향상**됩니다.

# Image-to-Image (aka Pix2Pix)



지금까지 무작위 입력에서 이미지를 생성하는 방법에 대해 이야기했습니다. 그러나 각 픽셀에 물건 클래스를 할당하는 **분할 맵(segmentation map)**이라는 이미지를 사용함으로써, 그 이미지에서 그 분할 맵에 맞는 사진처럼 보이는 것을 만들어냅니다. 이는 이미지 간 변환입니다.

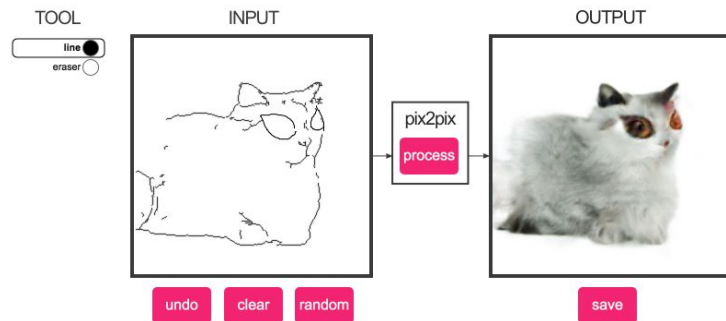




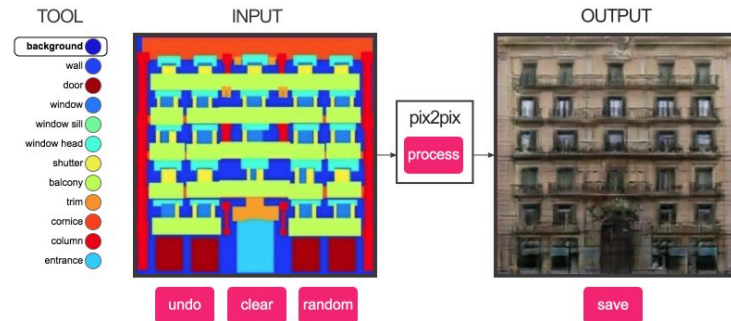
생성자가 매핑을 학습하려면 **분할 맵**과 해당 사진의 데이터 세트가 필요합니다. **GAN** 아키텍처를 수정하여 생성자와 판별자가 둘 다 분할 맵을 수신합니다. 물론 생성기는 어디에서 그려야 할지를 알기 위해 지도가 필요하지만, 판별자는 또한 생성자가 올바른 장소에 올바른 종류의 물건을 넣었는지 확인해야 합니다.

# Online [ [demo](#) ]

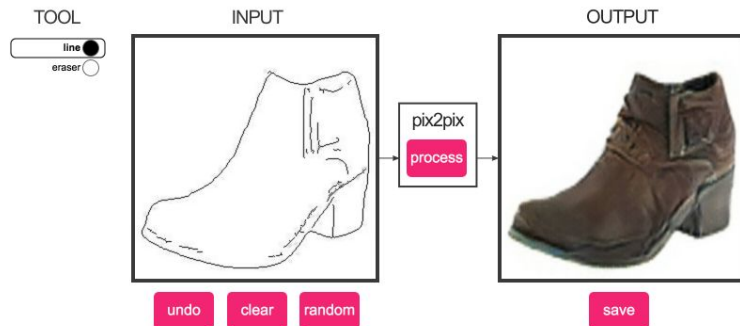
edges2cats



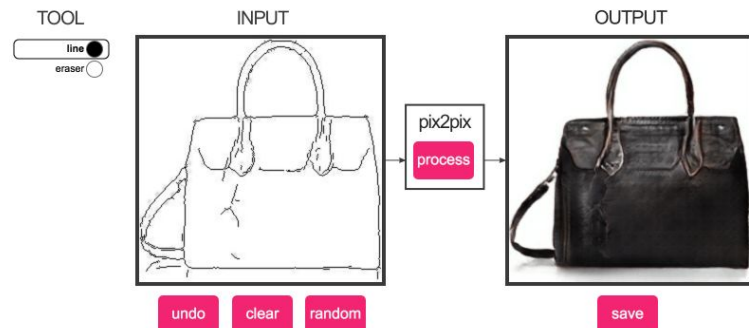
facades



edges2shoes

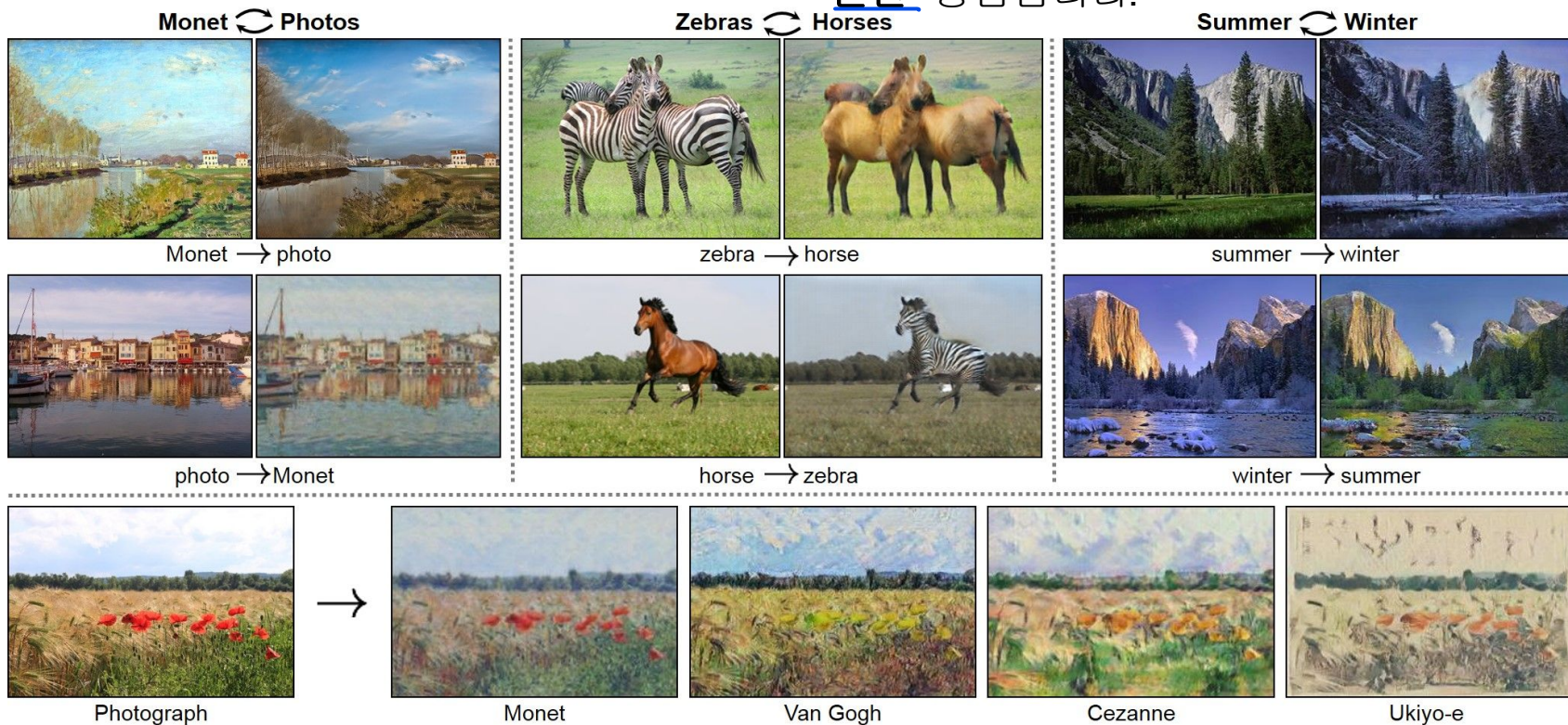


edges2handbags



# CycleGAN (aka Style Transfer)

CycleGAN은 소스 도메인에서 대상 도메인으로 이미지를 변환하는 “스타일 변환” 방법입니다.



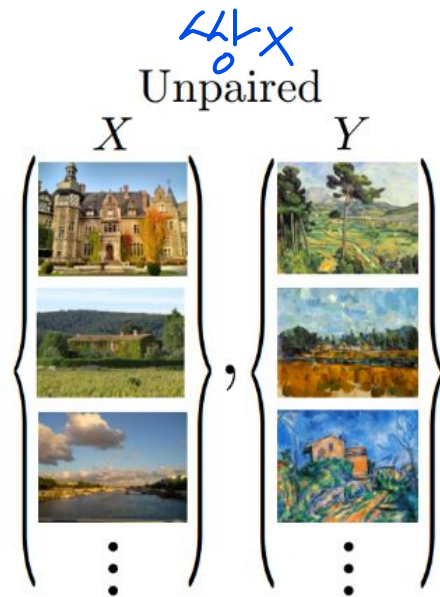
# Pix2Pix vs CycleGAN [Link to CycleGAN project page](#)

오른쪽 사진에서 Pix2Pix 모델에서 신발의 외곽선 이미지로 완전한 신발 이미지를 생성하고 싶다면 학습 단계에서 특정한 신발 이미지와 그에 해당하는 외곽선 이미지 쌍이 존재해야 합니다. 이미지에서 기존의 외곽선 추출 기법을 통해 외곽선 이미지를 얻어내는 것은 쉬운 일이다.

하지만 **Style Transfer**와 같은 일은 이러한 쌍을 얻기가 불가능합니다. 그러나 우리는 고풍 그림, 풍경 사진, 조랑말, 그리고 얼룩말 사진을 모을 수는 있습니다. 따라서 이와 같은 상황에서는 ‘쌍을 이루는’(Paired) 데이터가 아닌, ‘쌍을 이루지 않는’(Unpaired) 데이터로 학습하는 수 밖에 없습니다.

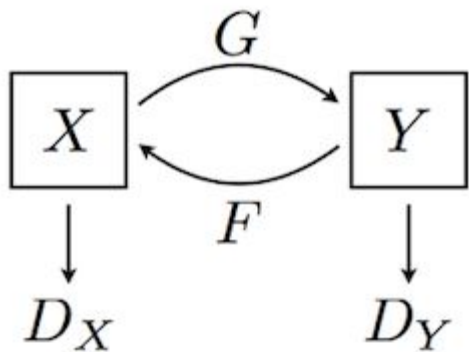


Pix2Pix



CycleGAN

# CycleGAN Architecture



CycleGAN은 두 가지 매핑을 학습하는 생성기 네트워크를 학습합니다. CycleGAN은 2개의 생성기 및 2개의 판별기 네트워크를 훈련시킵니다.

- 두 개의 매핑  $G: X \rightarrow Y$  및  $F: Y \rightarrow X$
- 해당 적대적 감별자  $D_Y$ 와  $D_X$

**G의 역할** :  $G$ 는  $X$ 를 출력으로 변환하려고 합니다. 이 출력은 도메인  $Y$ 에 따라  $D_Y$ 를 통해 공급되어 실제 또는 가짜인지 확인합니다.

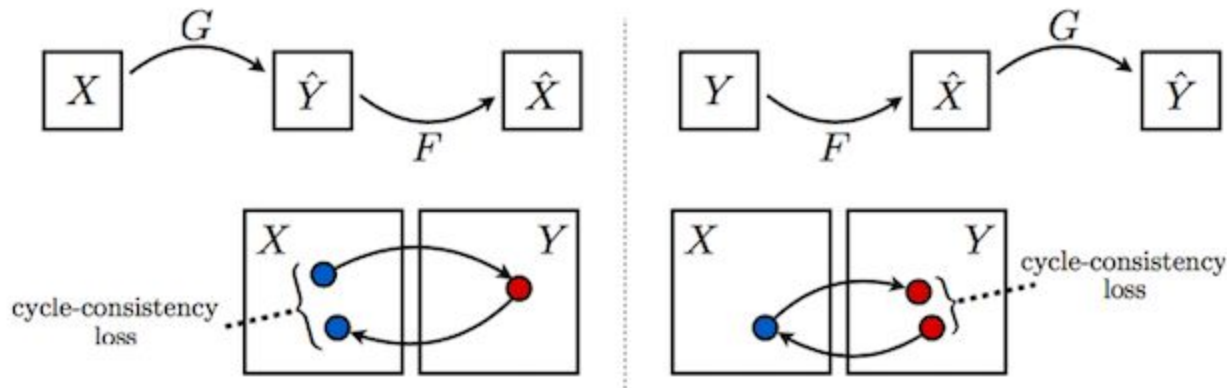
**F의 역할** :  $F$ 는  $Y$ 를 출력으로 변환하려고 시도합니다. 이 출력은  $D_X$ 를 통해 공급되어 도메인  $X$ 와 구별 할 수 없는지 확인합니다.



# Loss Function 손실 함수

CycleGAN의 진정한 힘은 사용하는 손실 함수에 있습니다. 생성기 및 감별기 손실 외에도 다음과 같이 순환-일관성 손실(cycle-consistency loss)을 사용합니다.

이런 종류의 손실은 매핑 함수  $G$ 를 사용하여 도메인  $X$ 에서  $Y$ 로 샘플을 변환한 다음, 함수  $F$ 를 사용하여  $X$ 로 다시 매핑하면, 원래 샘플에 얼마나 근접한지를 측정하는데 사용합니다. 마찬가지로, 샘플을  $Y$ 에서  $X$ 로 변환한 다음 다시  $Y$ 로 다시 변환하여 발생하는 손실을 계산합니다. 이 순환 손실을 최소화하도록  $F$ 와  $G$ 를 구해야 합니다.

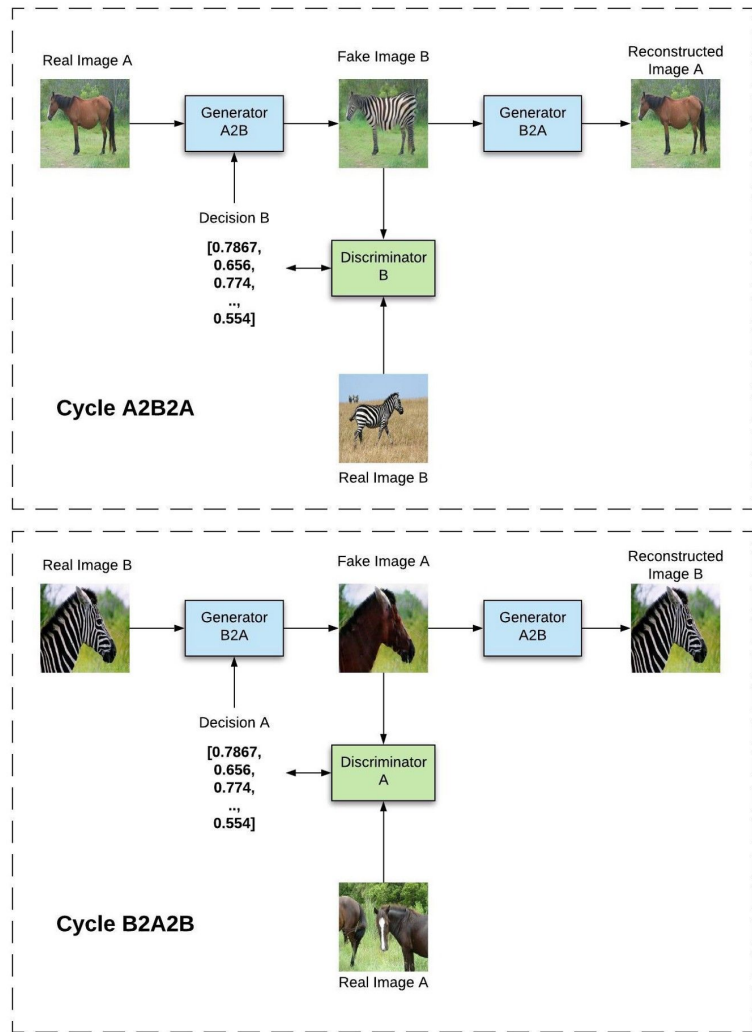


$$\underline{x \rightarrow G(x) \rightarrow F(G(x)) \approx x}$$

$$\underline{y \rightarrow F(y) \rightarrow G(F(y)) \approx y}$$



실제로 여기에는 두 개의 사이클이 필요합니다. 우리는 생성기 A2B와 생성기 B2A를 함께 훈련시키기 때문에 두 생성기가 시간이 지남에 따라 개선되고 있는지 확인해야 합니다. 그렇지 않으면 좋은 이미지를 재구성하는 데 여전히 문제가 있습니다. 또한, 위에서 논의한 바와 같이, 생성기를 개선한다는 것은 그 동안 감별기를 개선해야 한다는 것을 의미합니다. 사이클 A2B2A (A→B→A)에서 우리는 감별기 B를 사용하여 생성된 이미지가 도메인 B에 있는지 여부를 결정합니다. 따라서 감별기 B가 학습됩니다. 마찬가지로, 감별기 A도 훈련 할 수 있도록 주기 B2A2B가 필요합니다. 감별기 A와 감별기 B가 모두 잘 훈련되면 생성기 A2B와 B2A도 개선 될 수 있습니다!



# discoGAN (Discover Cross-Domain Relations GAN)

쌍을 이루지 않는 자료(Unpaired Dataset)와 순환-일관성(Cycle Consistency) 비용을 사용하는 개념적 내용으로는 CycleGAN과 완전히 동일하다. 하지만, Domain Transfer를 목적으로 간단한 Network 구조를 사용하여, 형태 변화가 자유롭지만 해상도는 낮음 단점과, 네트워크 규모가 단순해서 학습이 상대적으로 빠른 장점도 갖는다.



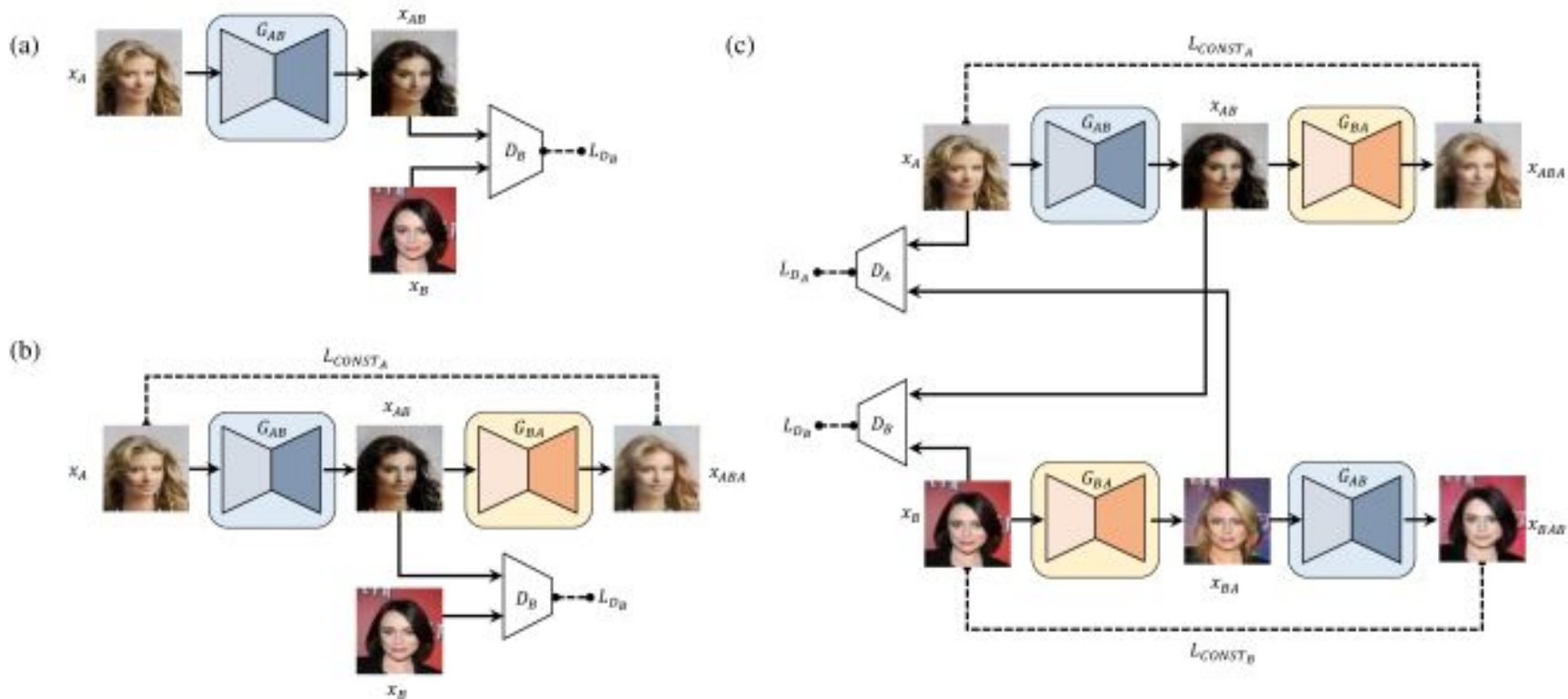


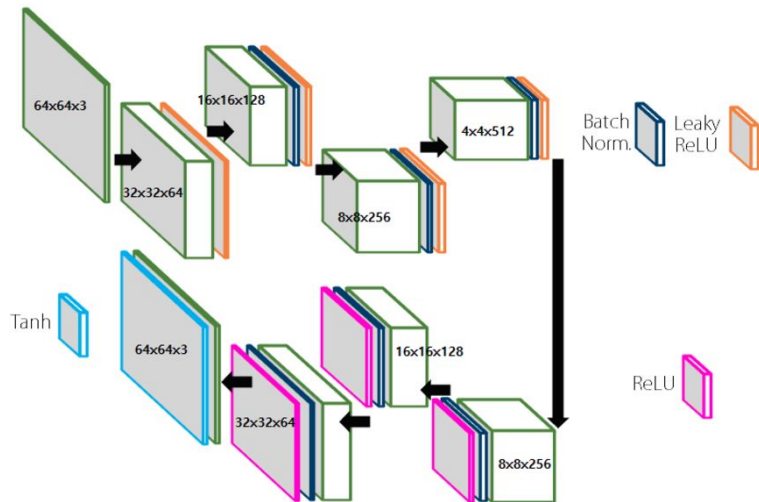
그림 2 : 세 가지 조사 모델. (a) 표준 GAN (Goodfellow et al., 2014), (b) 재구성 손실이 있는 GAN, (c) 두 개의 짝이 없고 레이블이 지정되지 않은 두 데이터 세트 간의 관계를 발견하도록 설계된 제안 모델 (DiscoGAN).

CycleGAN이 Style Transfer에 초점을 맞추어 형태를 변형시키는 것이 어렵다면 DiscoGAN의 경우에는 **형태에 변화가 큰 자료**에 주로 실험을 진행한 것을 알 수 있다. (신발, 핸드백, 자동차, 얼굴 등)

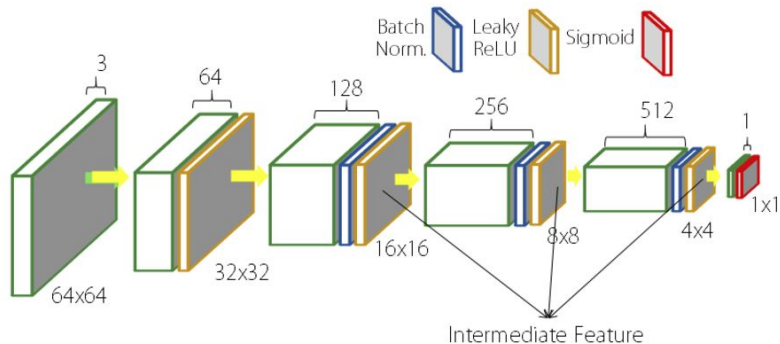
이것이 가능했던 이유는 DiscoGAN의 Network 구조를 먼저 살펴 보면 알 수 있는데, DiscoGAN에서의 각 Generator는 Resnet도 U-net도 아닌 **단순한 Encoder-decoder 구조**를 가져간다.

이렇게 구조를 가져가면 **정보의 손실**이 발생하는데, 학습이 살짝 어려울 수 있고, 높은 해상도를 기대하긴 힘들 수 있어도 역설적으로 **형태의 변환**에서는 자유롭다.

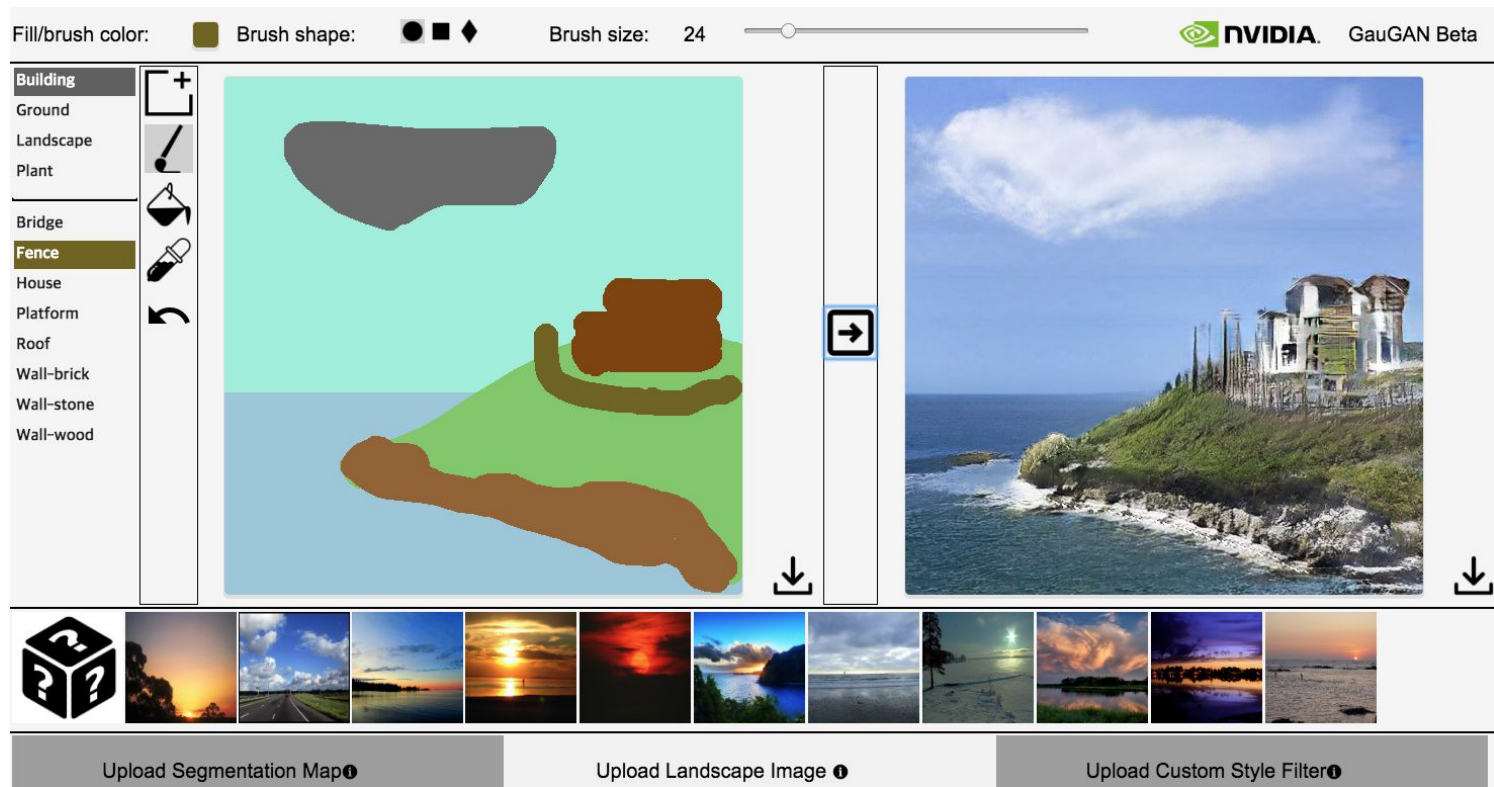
3.2 Generator Networks (network.py)



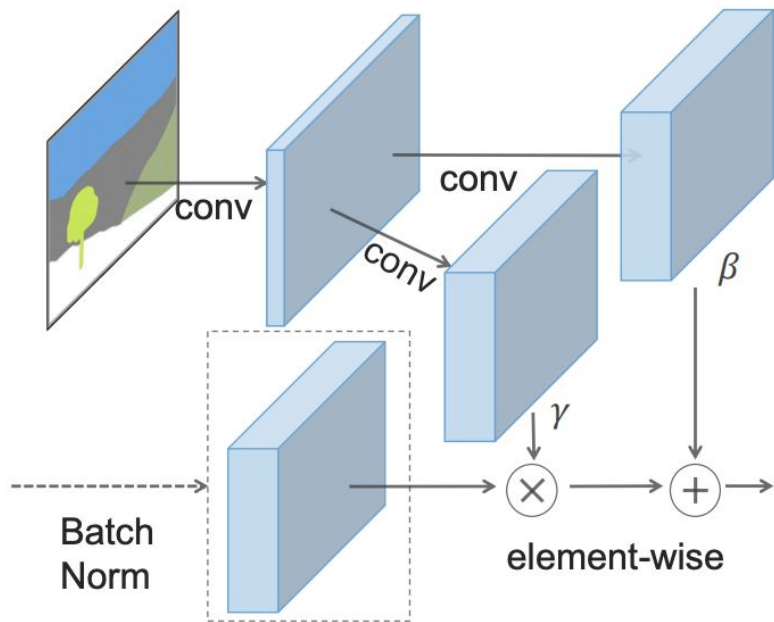
3.3 Discriminator Networks (network.py)



# GauGAN2 painting by Ndivia ([link](#)) ([youtube](#))



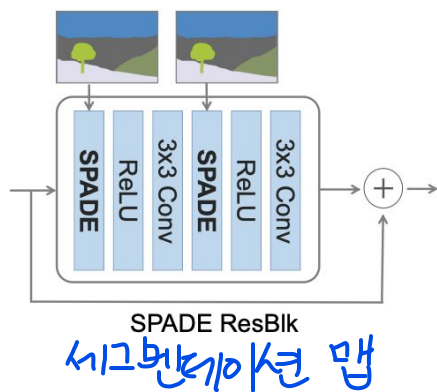
분할 맵들로부터 이미지들을 생성 : 공간-적응적 비정규화 (Spatially-adaptive (de)normalization : SPADE) 방법.



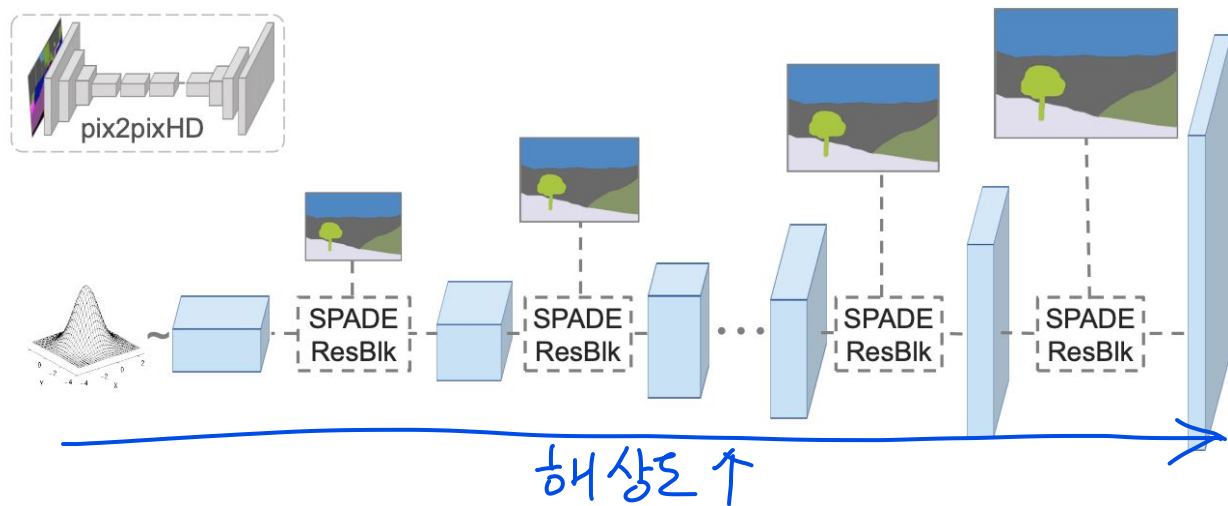
SPADE는 분할 맵을 첫 번째 레이어에 입력하는 대신 다운 샘플링 된 버전을 사용하여 모든 레이어의 배치-정규화된 출력을 변조합니다.



SPADE 생성기는 이 디자인을 작은 "잔여 블록(residual block)"으로 통합하여 업-샘플링된 레이어 (전치 된 컨볼루션) 사이에 끼웁니다.



SPADE의 "잔여 블록"에는 컨볼루션 레이어 및 건너뛰기 연결이 포함됩니다.



pix2pixHD의 생성기와 비교한 SPADE의 생성기에 대한 거시적 형태

# Additional Examples : need to be included in the lecture

- <https://this-person-does-not-exist.com/en>
- <https://thiscitydoesnotexist.com>
- <https://thiseyedoesnotexist.com>
- <https://thischairdoesnotexist.com>
- <https://www.artbreeder.com/>