

<복제물에 대한 경고>

본 저작물은 **저작권법 제25조 수업목적 저작물 이용 보상금제도**에 의거, **한국복제전송저작권협회**와 약정을 체결하고  
적법하게 이용하고 있습니다. 약정범위를 초과하는 사용은 저작권법에 저촉될 수 있으므로

**저작물의 재 복제 및 수업 목적 외의 사용을 금지합니다.**

2020. 03. 30.

건국대학교(서울)한국복제전송저작권협회

<전송에 대한 경고>

본 사이트에서 수업 자료로 이용되는 저작물은 **저작권법 제25조 수업목적 저작물 이용 보상금제도**에 의거,

**한국복제전송저작권협회**와 약정을 체결하고 적법하게 이용하고 있습니다.

약정범위를 초과하는 사용은 저작권법에 저촉될 수 있으므로

**수업자료의 대중 공개·공유 및 수업 목적 외의 사용을 금지합니다.**

2020. 03. 30.

건국대학교(서울)한국복제전송저작권협회

# Generative Adversarial Network

# Generative Adversarial Network (GAN)

- 적대적 생성 신경망

- 고품질의 이미지를 생성하기 위해서 생성자(generator)와 판별자(discriminator)를 경쟁적 학습시키는 인공 신경망 구조

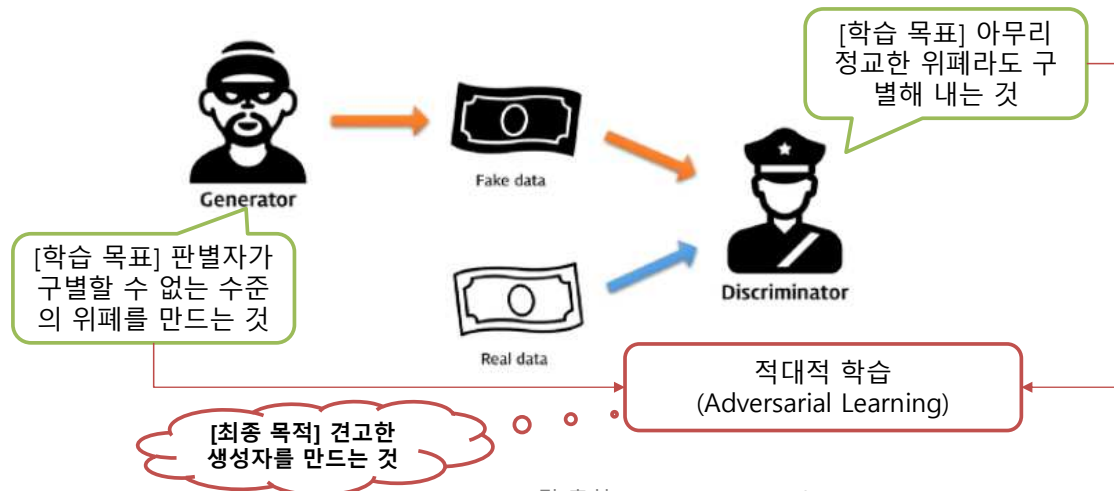
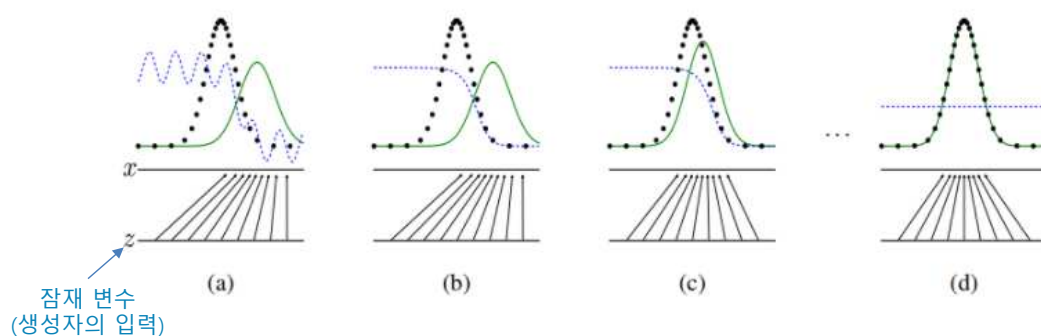


그림 출처: <https://dreamgonfly.github.io/blog/gan-explained/>



Edited by Harksoo Kim

## Adversarial Learning Process



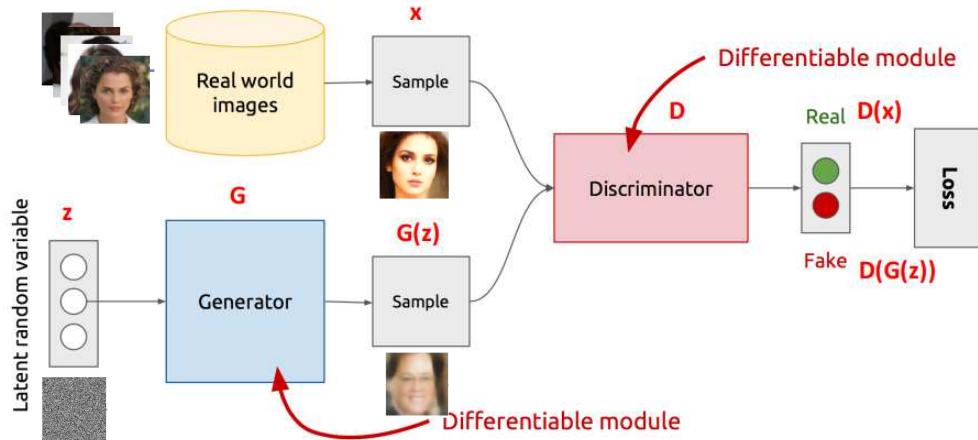
- 검은 점선: 학습 데이터의 분포  $\rightarrow P_{data}(X)$
- 녹색 실선: 생성자에 의해 생성된 데이터의 분포  $\rightarrow P_{Gen}(X)$
- 파란 점선: 판별자의 출력  $\rightarrow 0 < P_{Dis}(X) < 1$

그림 출처: Ian Goodfellow의 GAN 논문



Edited by Harksoo Kim

# Architecture of GAN



- $z$  is some random noise (Gaussian/Uniform).
- $z$  can be thought as the latent representation of the image.

그림 출처: 스탠포드 대학교 Fei-Fei 교수 강의자료



Edited by Harksoo Kim

## Training Discriminator

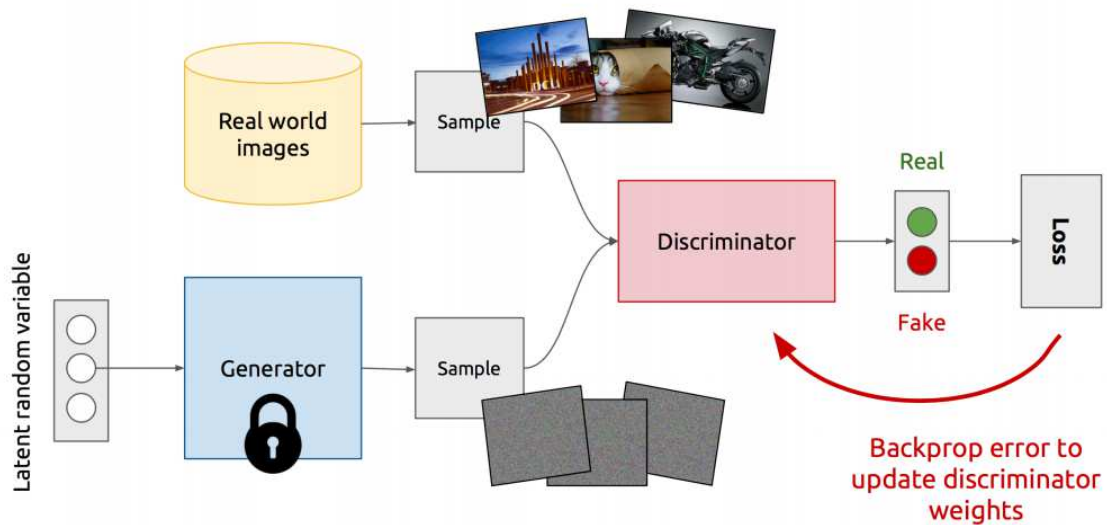


그림 출처: 스탠포드 대학교 Fei-Fei 교수 강의자료



Edited by Harksoo Kim

# Training Generator

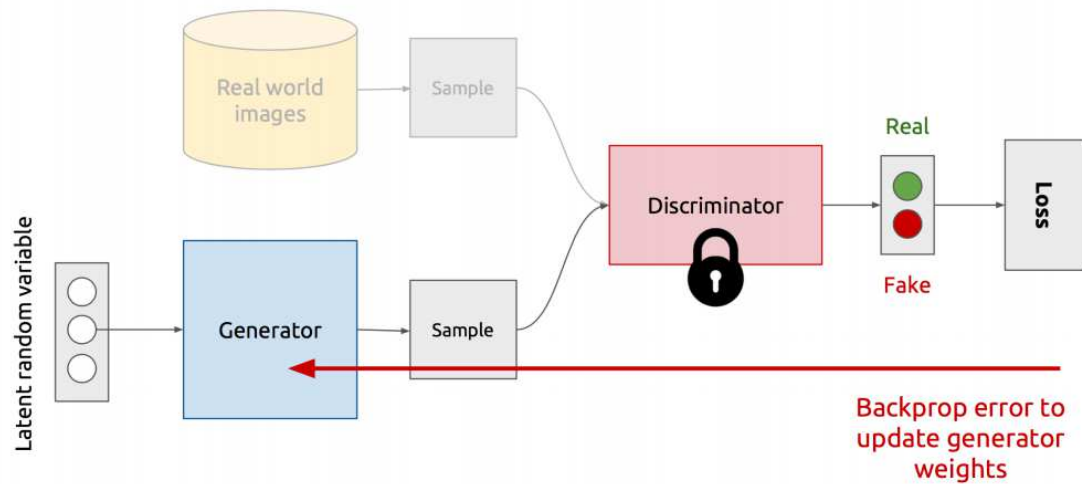


그림 출처: 스탠포드 대학교 Fei-Fei 교수 강의자료



Edited by Harksoo Kim

## Object of GAN Training

### • 목적 함수

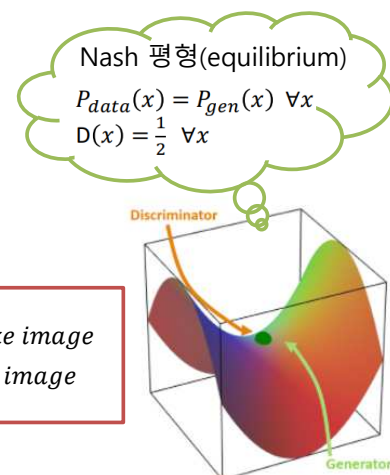
$$\min_G \max_D V(D, G) = \mathbb{E}_{\mathbf{x} \sim p_{\text{data}}(\mathbf{x})} [\log D(\mathbf{x})] + \mathbb{E}_{\mathbf{z} \sim p_{\mathbf{z}}(\mathbf{z})} [\log(1 - D(G(\mathbf{z})))]$$

D(x): 본인의 리워드 최대화

$$\begin{cases} \log(D(x)) = \log(1) = 0, \text{ if } x \text{ is a real image} \\ \log(1 - D(G(z))) = \log(1) = 0, \text{ if } G(z) \text{ is a fake image} \end{cases}$$

G(z): D(x)의 리워드 최소화

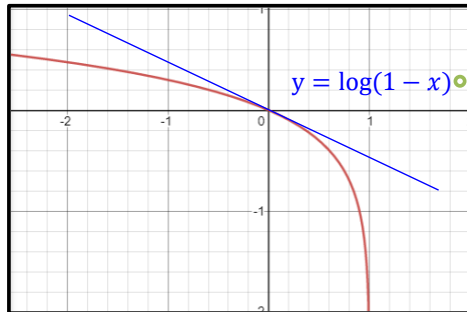
$$\begin{cases} \log(1 - D(G(z))) = \log(1 - 1) = -\infty, \text{ if } G(z) \text{ is a real-like image} \\ \log(1 - D(G(z))) = \log(1 - 0) = 0, \text{ if } G(z) \text{ is a fake-like image} \end{cases}$$



Edited by Harksoo Kim

# Changing Object Function

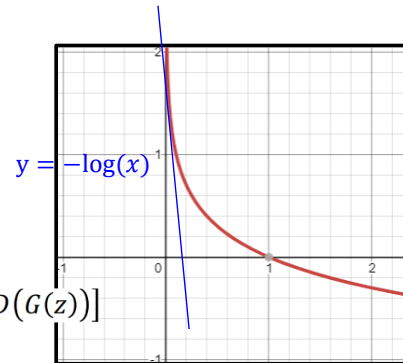
$$\min_G \max_D V(D, G) = \mathbb{E}_{\mathbf{x} \sim p_{\text{data}}(\mathbf{x})} [\log D(\mathbf{x})] + \mathbb{E}_{\mathbf{z} \sim p_{\mathbf{z}}(\mathbf{z})} [\log(1 - D(G(\mathbf{z})))]$$



초기 G(z)는 판별하기 너무 쉬운 이미지 생성 →  $x=0$ , Small gradient!



Minimize  $-\mathbb{E}_{\mathbf{z} \sim q(\mathbf{z})} [\log D(G(\mathbf{z}))]$



Edited by Harksoo Kim

## Limitation of Vanilla GAN

GANs are still evolving!

- Non-Convergence Problem
  - 양쪽이 조금씩 발전해야 내쉬 평형 상태가 되는데 한쪽이 강력해지면 다른 쪽은 발전하지 못하는 문제
- Mode-Collapse Problem
  - 제대로 학습되지 못한 경우에 다양한 이미지를 만드는 것이 아니라 비슷한 이미지만 계속 생성하는 문제

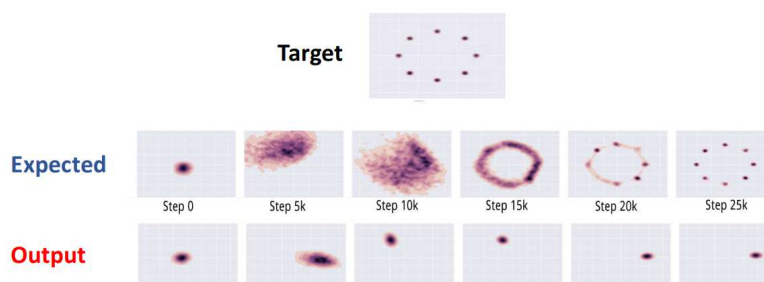


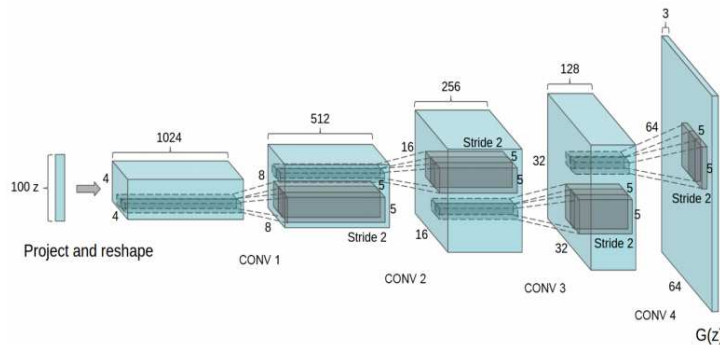
그림 출처: Metz et al. "Unrolled Generative Adversarial Networks", arXiv, 2016



Edited by Harksoo Kim

# Deep Convolutional GAN (DCGAN)

## Generator Architecture



### Key ideas:

- Replace FC hidden layers with Convolutions
  - **Generator:** Fractional-Strided convolutions
- Use Batch Normalization after each layer
- **Inside Generator**
  - Use ReLU for hidden layers
  - Use Tanh for the output layer

그림 출처: 스탠포드 대학교 Fei-Fei 교수 강의자료



Edited by Harksoo Kim

# Semi-Supervised GAN (SGAN)

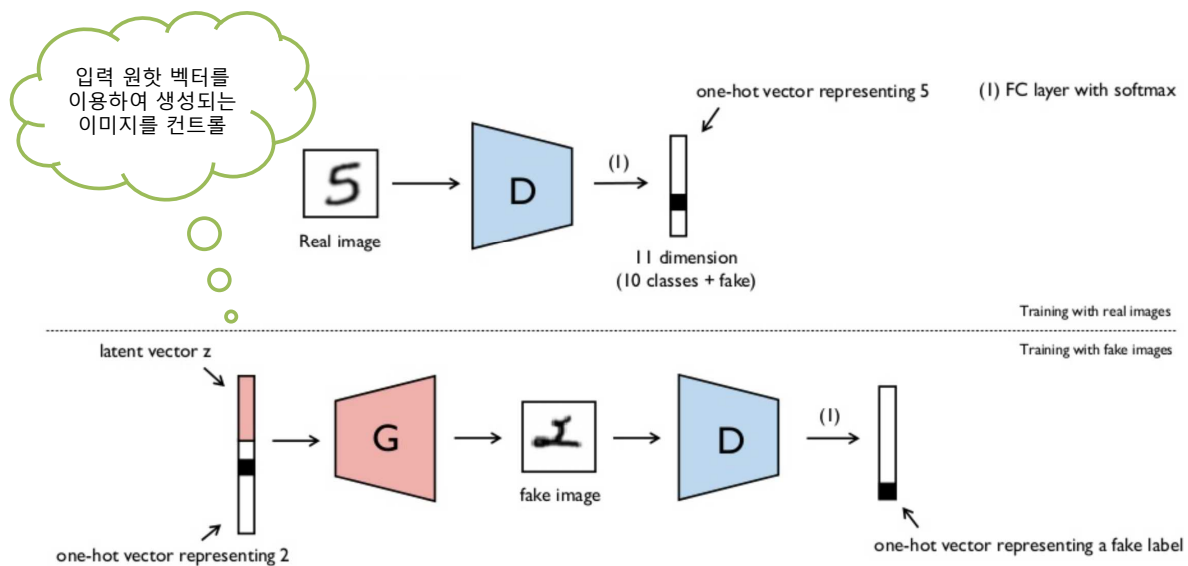
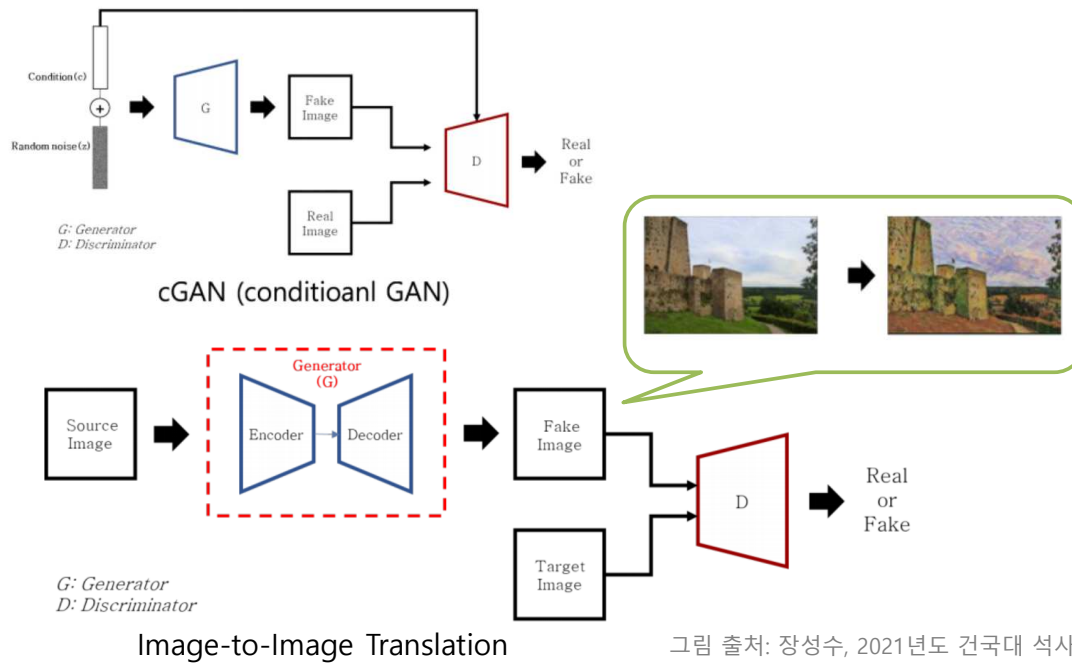


그림 출처: Augustus Odena et al. "Semi-Supervised Learning with Generative Adversarial Networks", 2016



Edited by Harksoo Kim

# Other GANs



Edited by Harksoo Kim

## 질의응답

# Q&A

Homepage: <http://nlp.konkuk.ac.kr>  
E-mail: [nlpdrkim@konkuk.ac.kr](mailto:nlpdrkim@konkuk.ac.kr)



Edited by Harksoo Kim