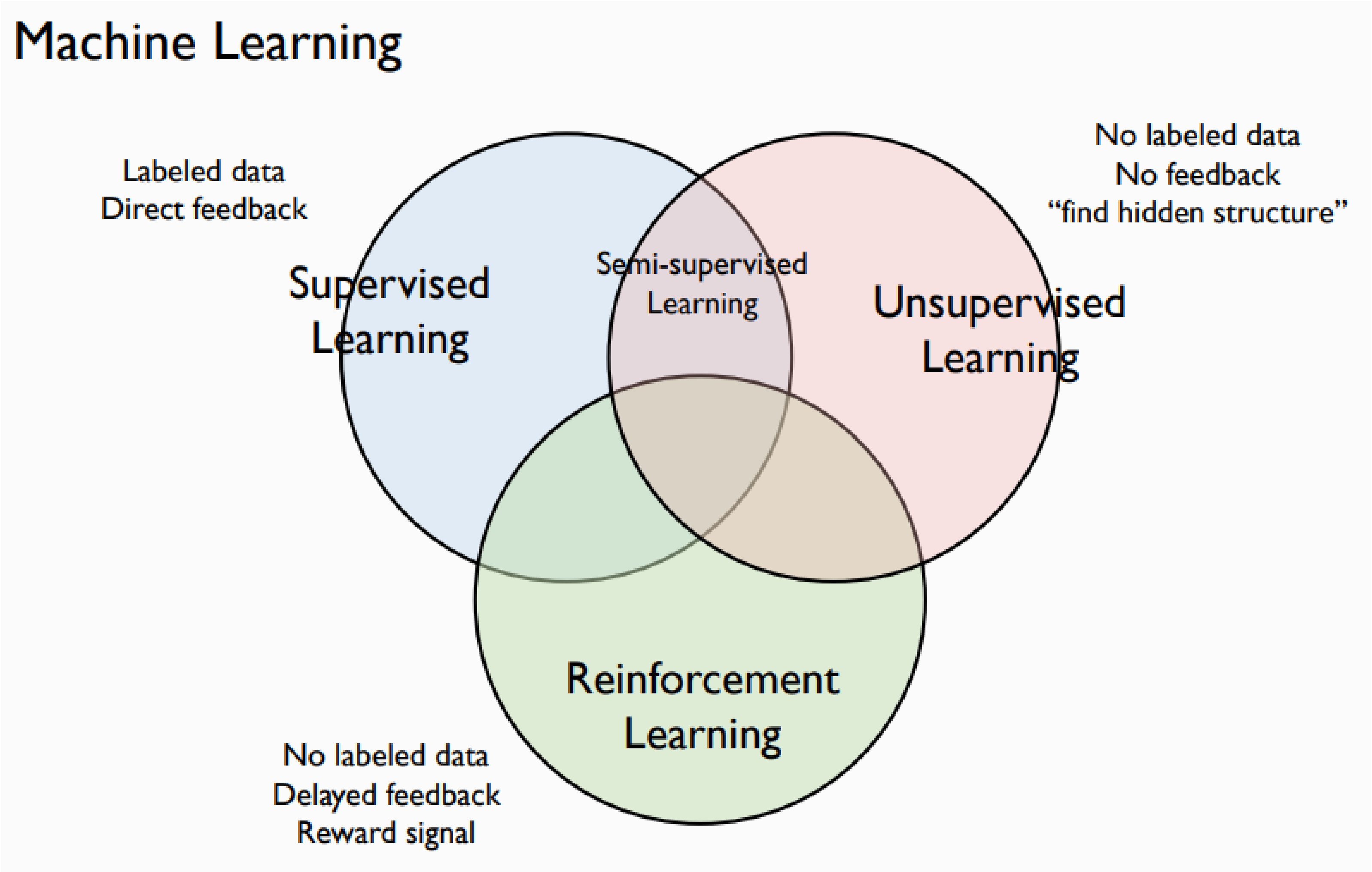
Generative
Adversarial
Network

20.03.30 세미나#11

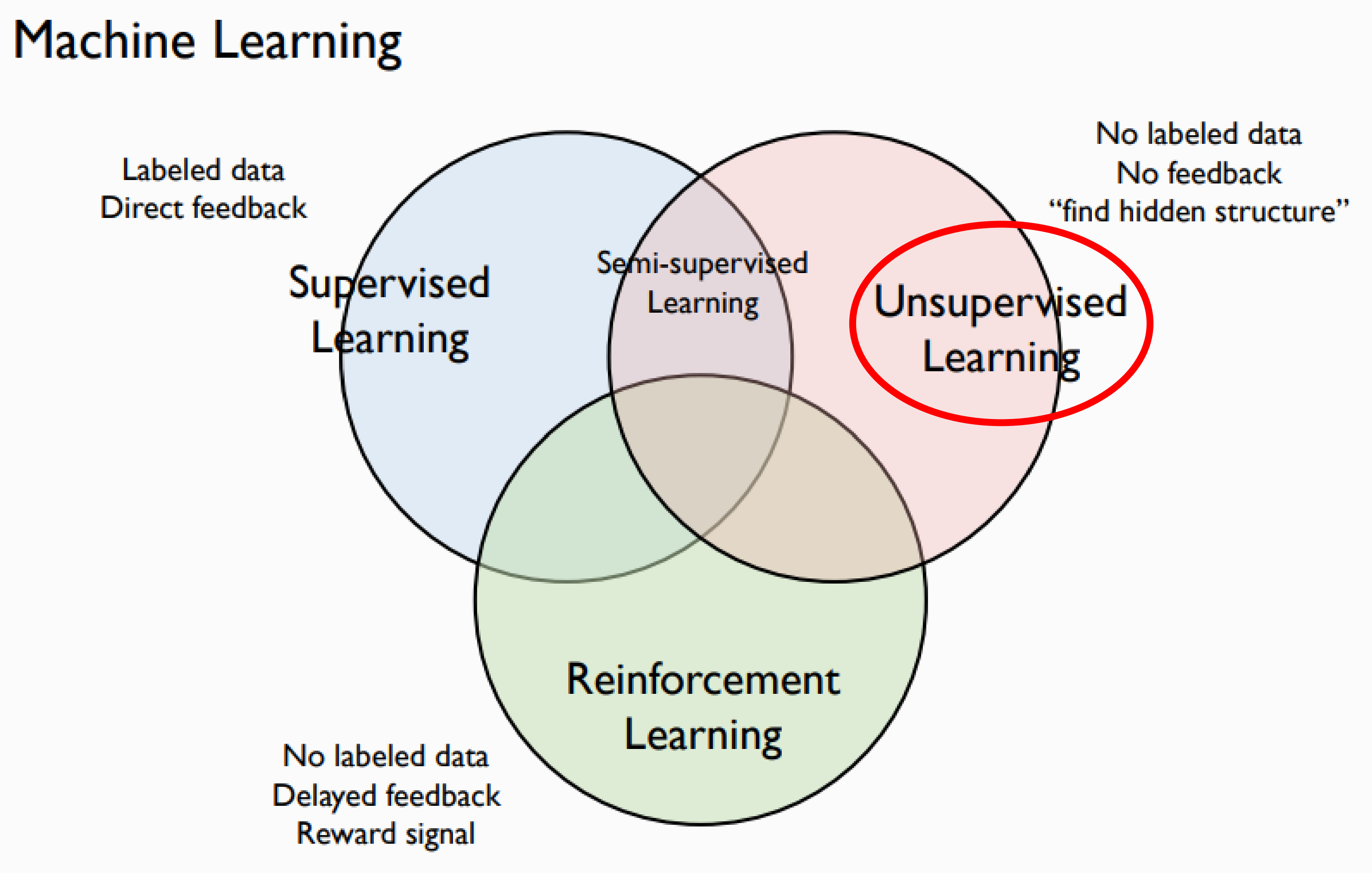
AI Lab 1기 김필성

Introduction

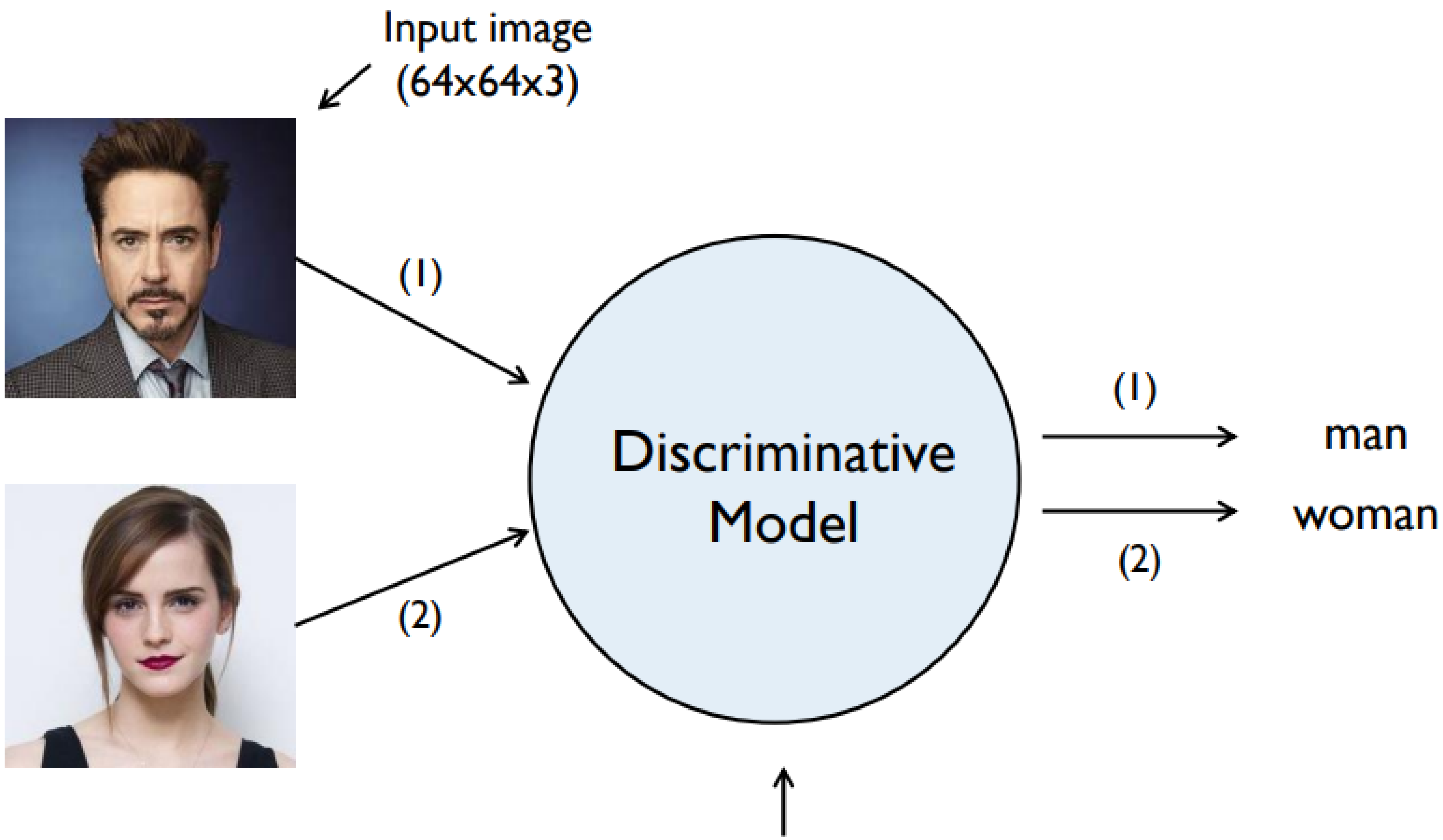
Introduction



Introduction

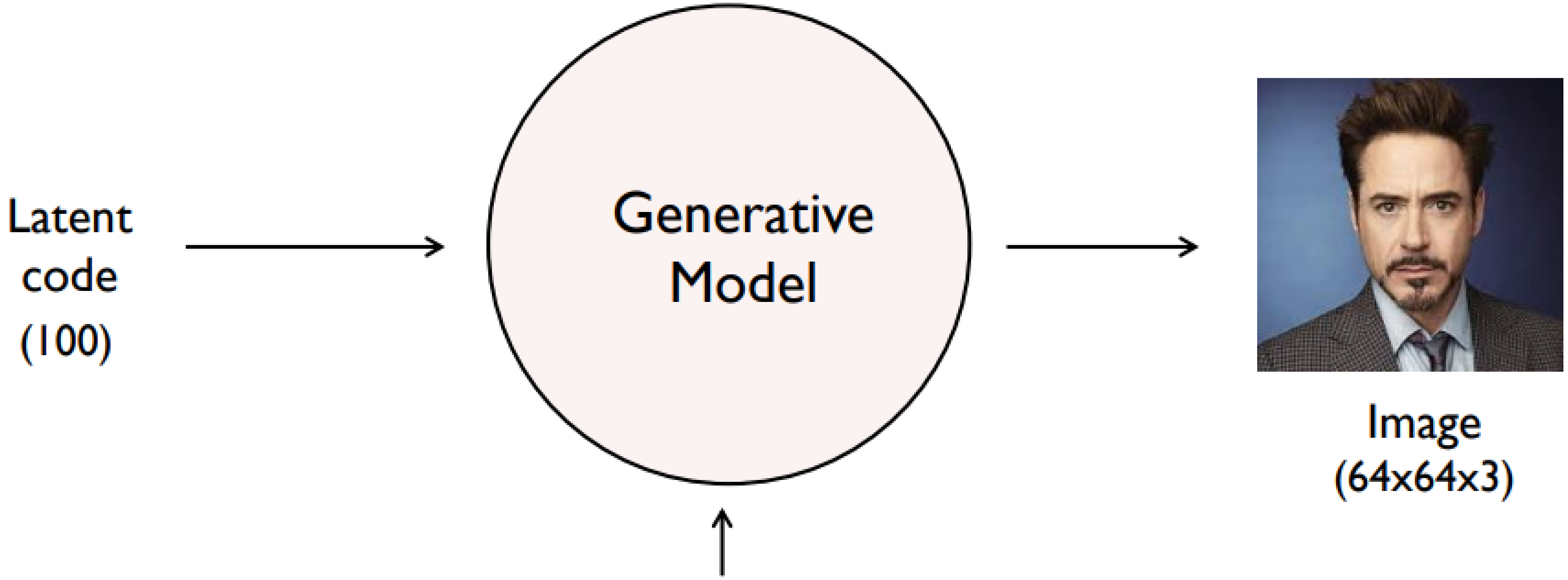


Introduction – Supervised Learning



The discriminative model learns how to classify input to its class.

Introduction – Unsupervised Learning



The generative model learns the distribution of training data.

Introduction – Unsupervised Learning

- supervised learning 보다 더 challeng하다 :
  - label이 없다 ⟶ self learning

- Some NN solutions :
  - Boltzmann machine
  - AE or VAE
  - GAN

Introduction – Unsupervised Learning

- supervised learning 보다 더 challenge하다 :
  - label이 없다 ⟶ self learning
- Some NN solutions :
  - Boltzmann machine
  - AE or VAE
  - GAN

Generative Adversarial Network

GAN – Generative model

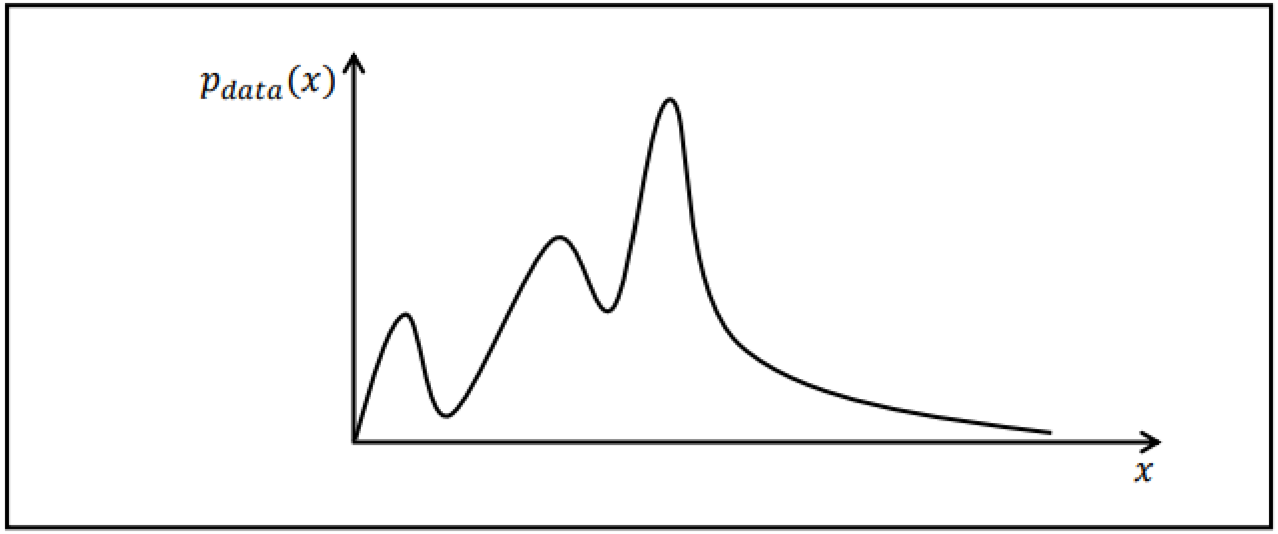
What if x is actual images in the training data?

At this point, x can be represented as a (for example) 64x64x3 dimensional vector.

GAN – Generative model

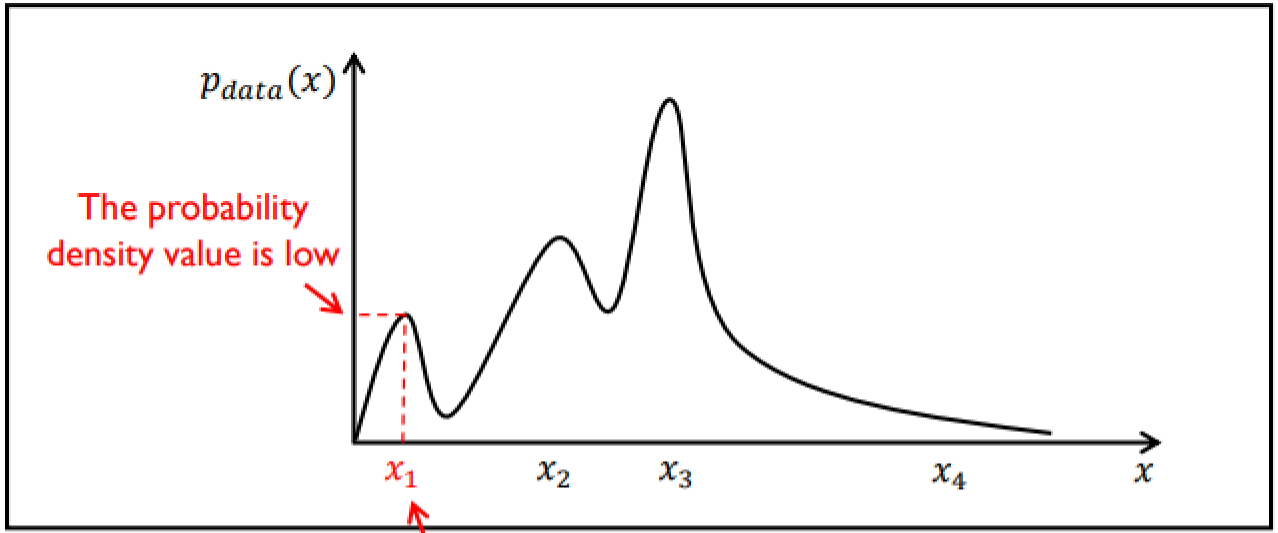
Probability density function

There is a $p_{data}(x)$ that represents the distribution of actual images.



GAN – Generative model

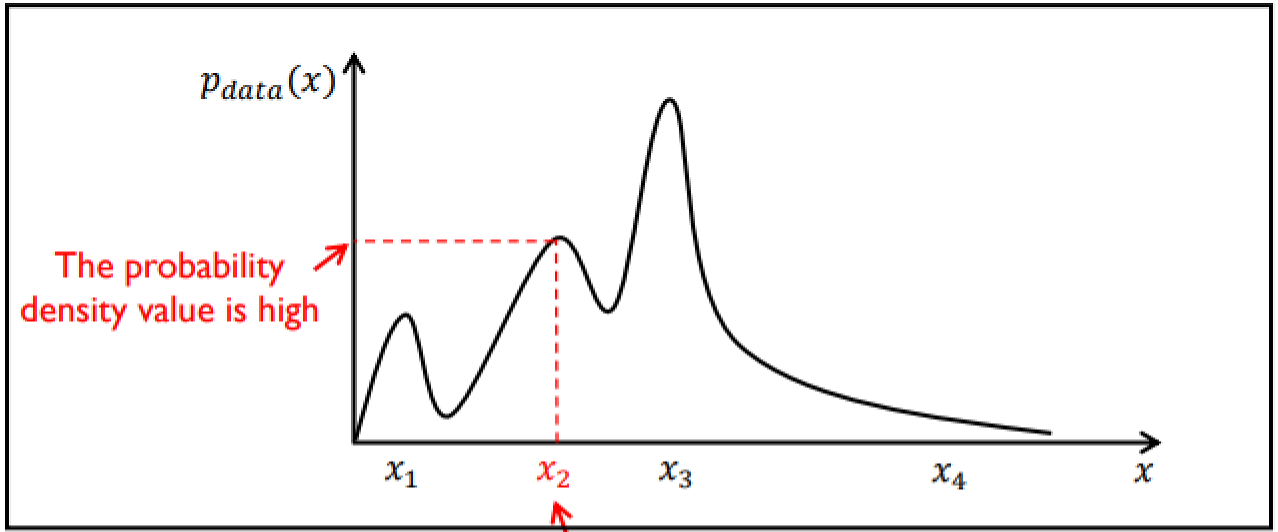
Let's take an example with human face image dataset.
Our dataset may contain few images of men with glasses.



x_1 is a 64x64x3 high dimensional vector representing a man with glasses.

GAN – Generative model

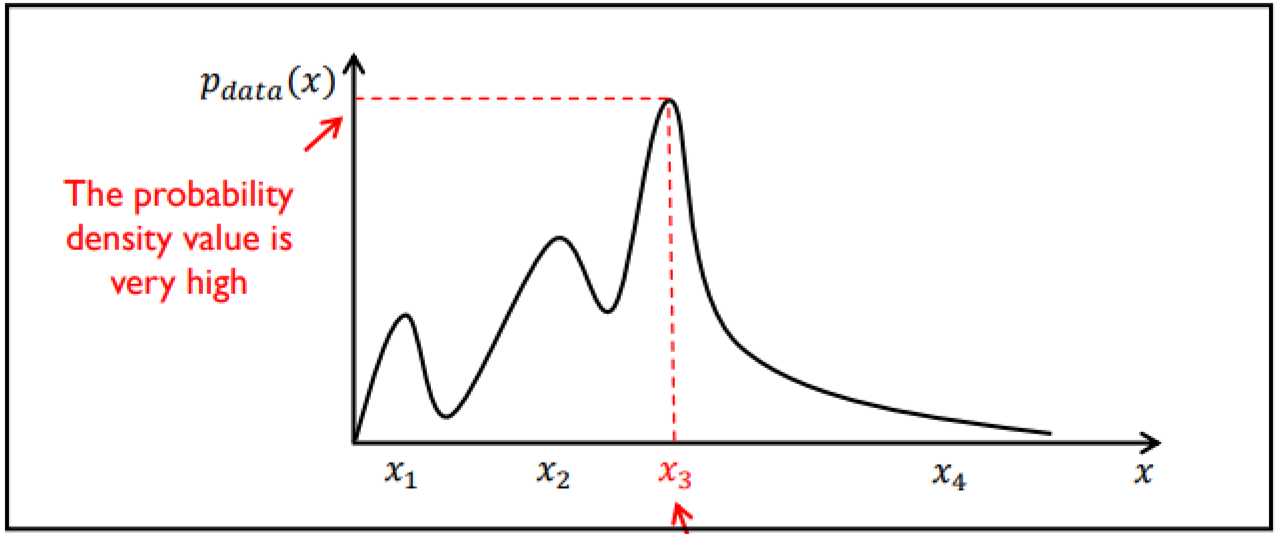
Our dataset may contain many images of women with black hair.



x_2 is a 64x64x3 high dimensional vector representing a woman with black hair.

GAN – Generative model

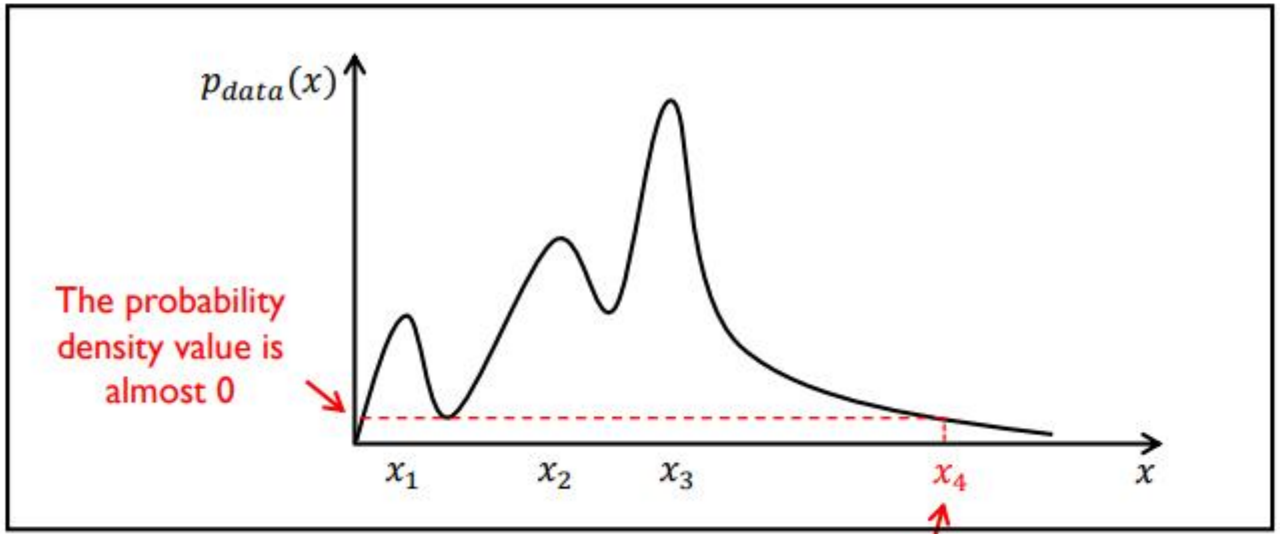
Our dataset may contain very many images of women with blonde hair.



x_3 is a 64x64x3 high dimensional vector representing a woman with blonde hair.

GAN – Generative model

Our dataset may not contain these strange images.



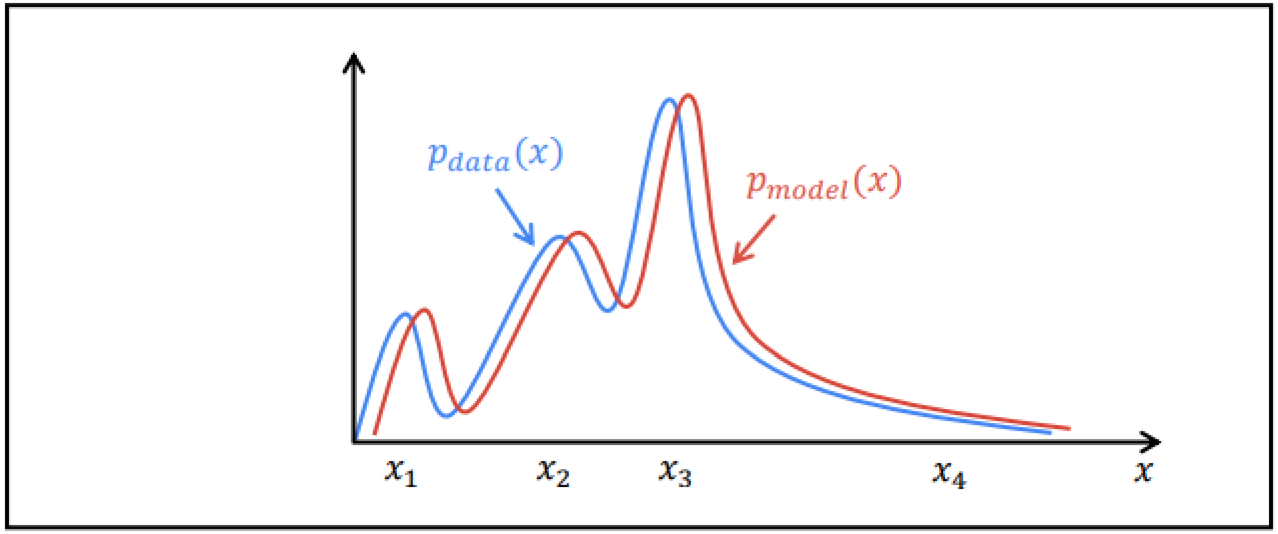
x_4 is an 64x64x3 high dimensional vector representing very strange images.

GAN – Generative model

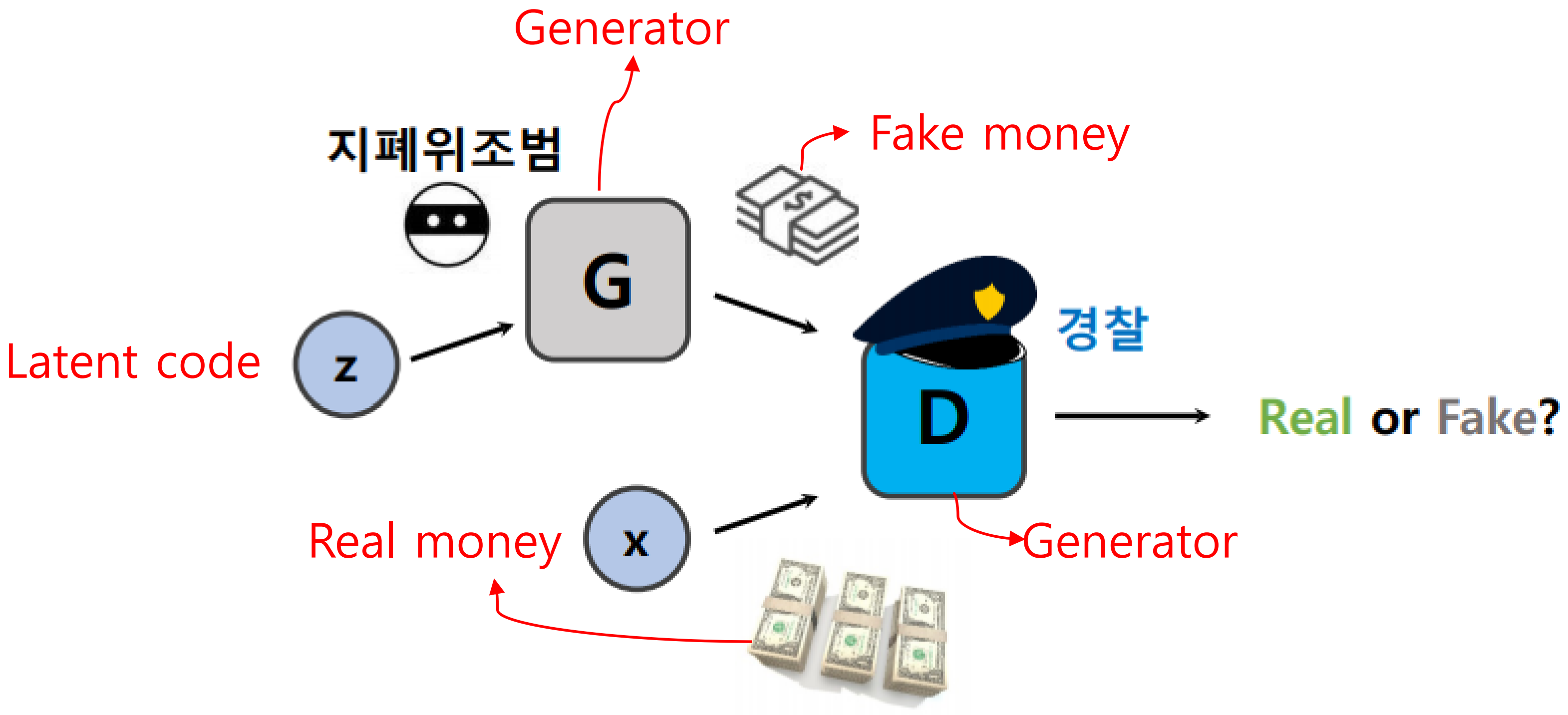
$\curvearrowright$ *Distribution of images generated by the model*

The goal of the generative model is to find a $p_{model}(x)$ that approximates $p_{data}(x)$ well.

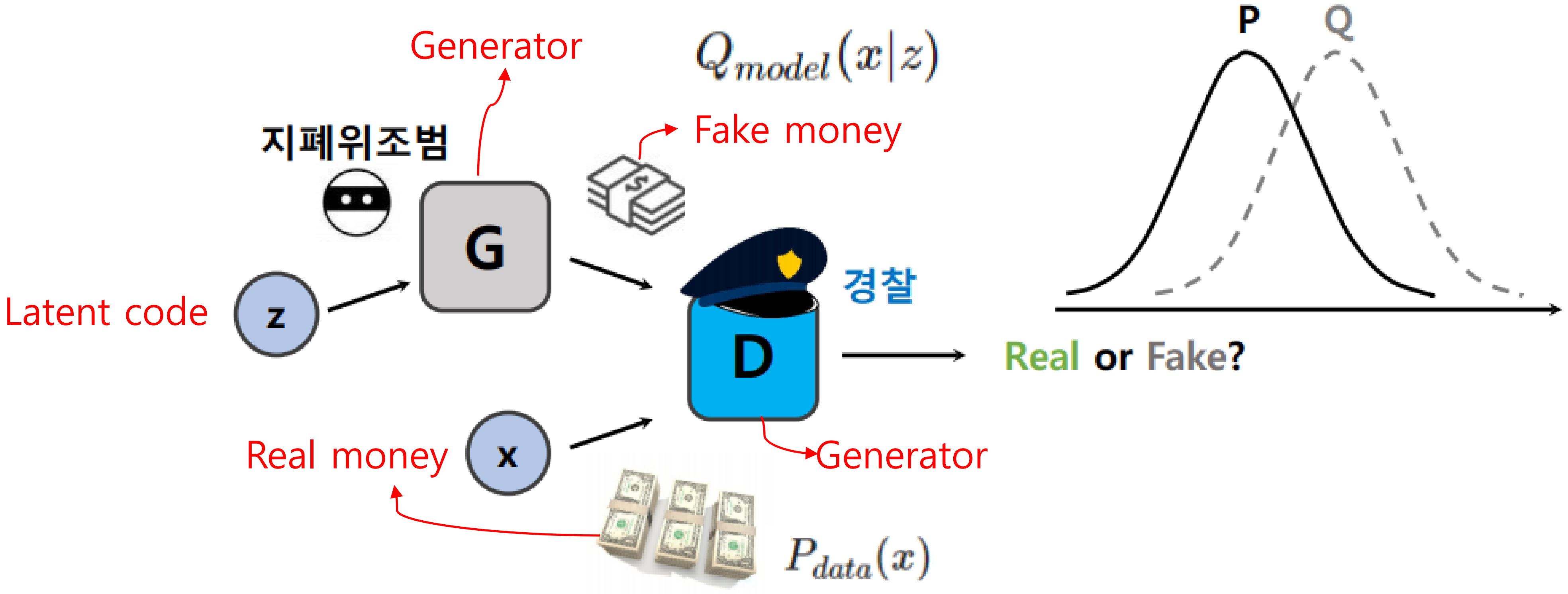
Distribution of actual images



GAN – Schematic overview

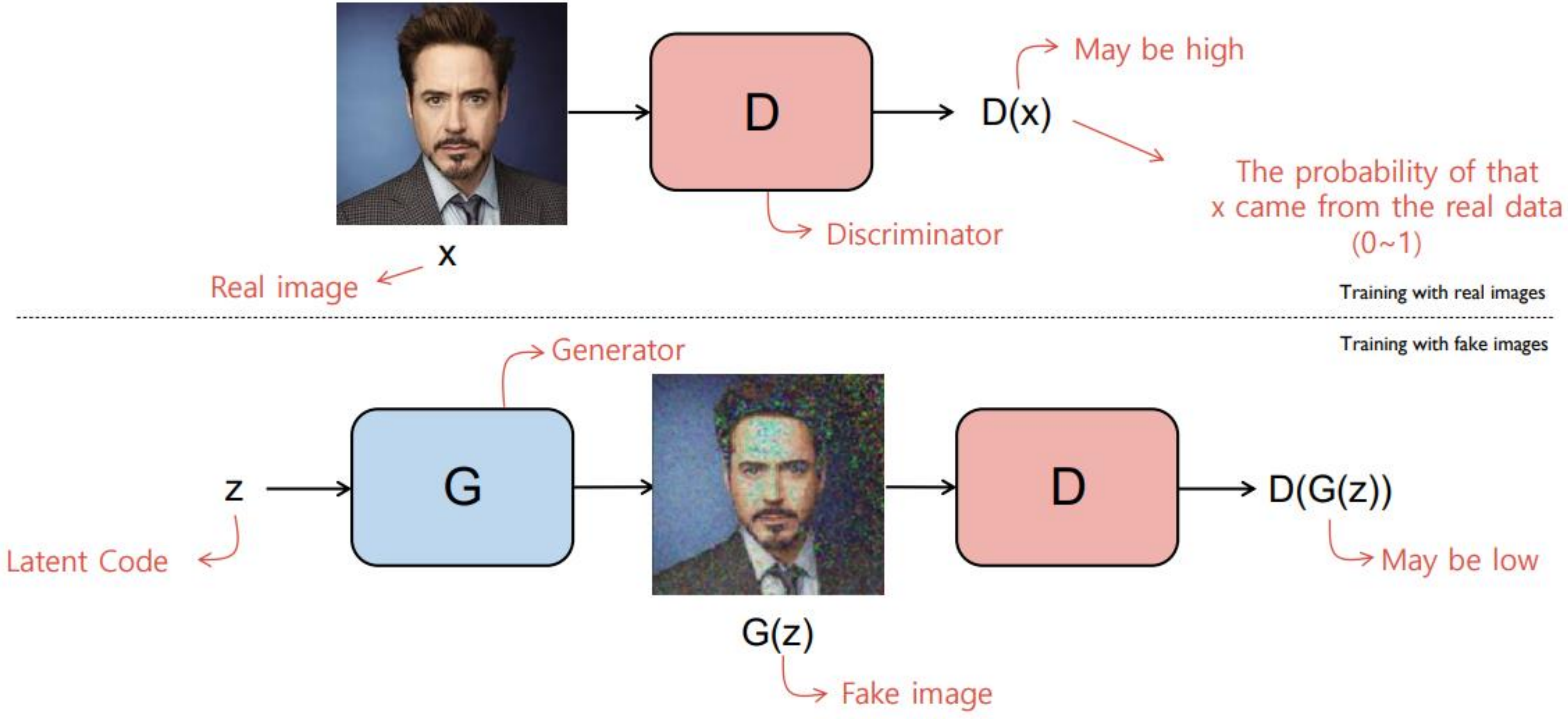


GAN – Schematic overview

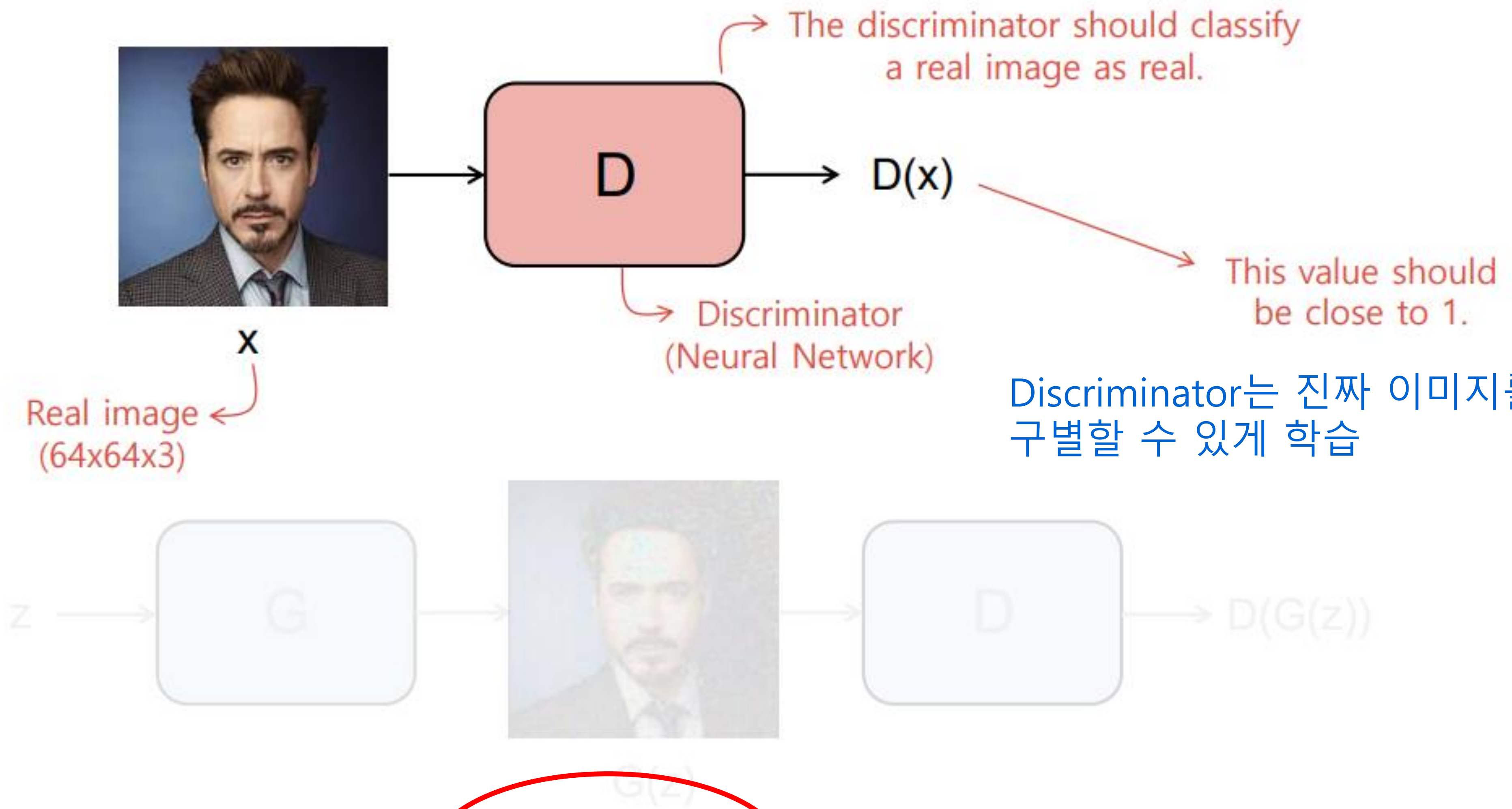


GAN – Schematic overview

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



GAN – Schematic overview



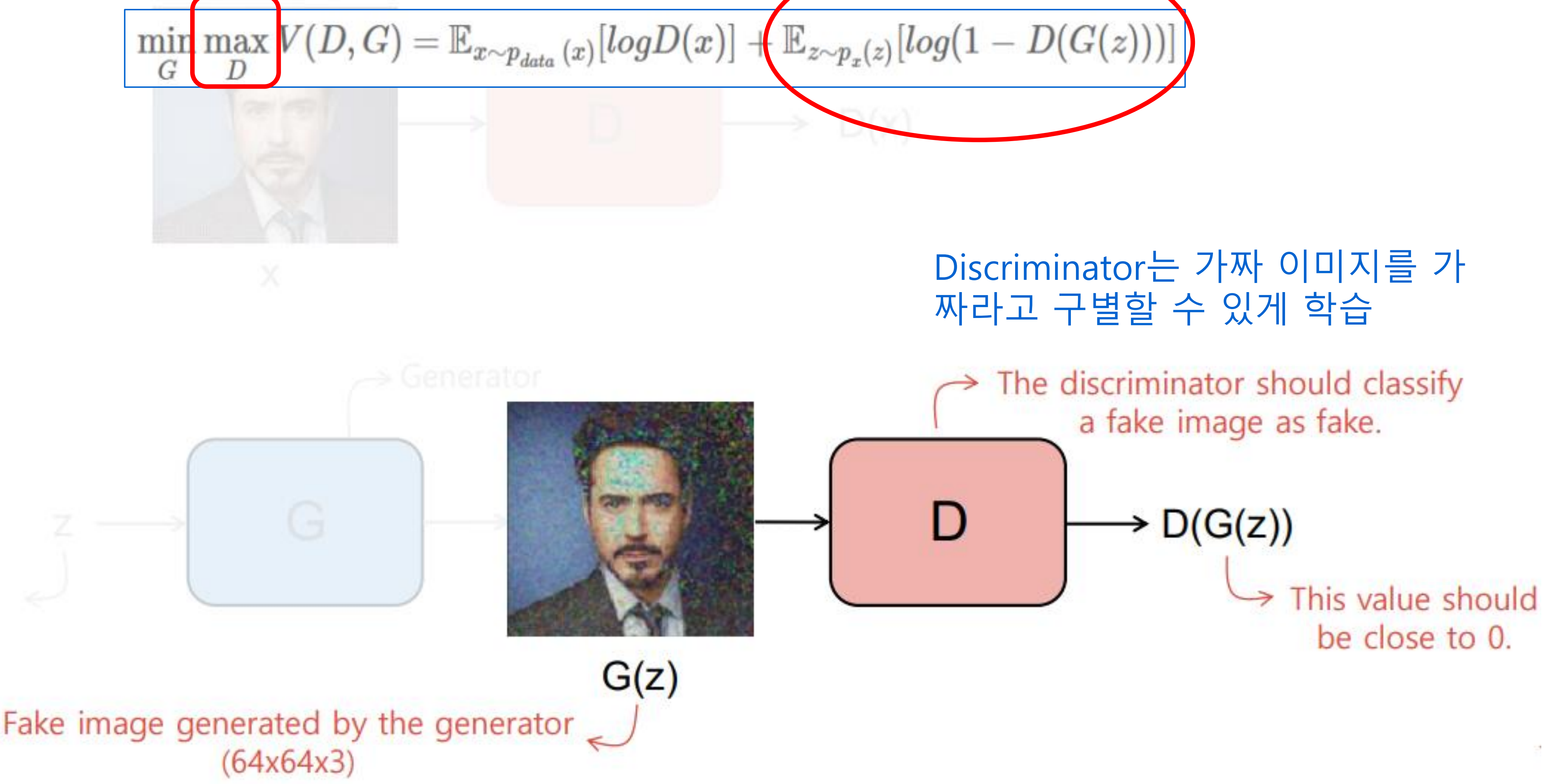
Discriminator는 진짜 이미지를 잘 구별할 수 있게 학습

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

GAN – Schematic overview

$$\min_{G} \boxed{\max_{D}} V(D,G) = \mathbb{E}_{x\sim p_{data}(x)}[logD(x)] + \mathbb{E}_{z\sim p_z(z)}[log(1-D(G(z)))]$$

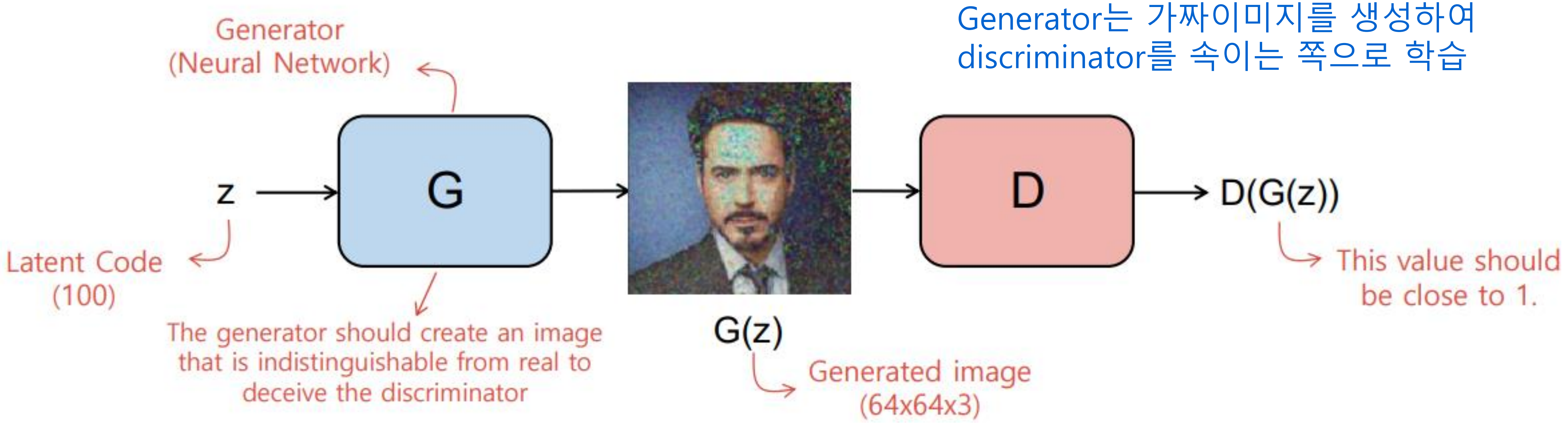
Discriminator는 가짜 이미지를 가짜라고 구별할 수 있게 학습



GAN – Schematic overview

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

Generator는 가짜이미지를 생성하여 discriminator를 속이는 쪽으로 학습



GAN – Objective Function

- Minimax problem of GAN

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

- 2단계의 증명이 필요

1. GAN의 Minimax problem은 $\boldsymbol{p_g = p_{data}}$ 일때 global optimum을 가짐
2. global optimum 일 때 적절한 알고리즘을 찾을 수 있다

GAN – Objective Function(optimal D)

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

에서 G가 고정되어 있을 때, 최적화된 D(목적함수의 미분값이 0)는

$$D_G^*(x) = \frac{p_{data}(x)}{p_{data}(x) + p_g(x)}.$$

$$\begin{aligned}
C(G) &= \max_D V(G,D) \\
&= \mathbb{E}_{x\sim p_{data}}\left[logD_G^*(x)\right] + \mathbb{E}_{z\sim p_z}\left[log(1 - D_G^*(G(z)))\right] \\
&= \mathbb{E}_{x\sim p_{data}}\left[logD_G^*(x)\right] + \mathbb{E}_{x\sim p_g}\left[log(1 - D_G^*(x))\right] \\
&= \mathbb{E}_{x\sim p_{data}}\left[log\frac{p_{data}(x)}{p_{data}(x) + p_g(x)}\right] + \mathbb{E}_{x\sim p_g}\left[log\frac{p_g(x)}{p_{data}(x) + p_g(x)}\right]
\end{aligned}$$

GAN – Objective Function(global minimum)

global minimum은 C(G)가 $p_g = p_{data}$ 가 되면 된다.

이 때 $C(G) = -log4$

For $p_g = p_{data}$, $D_G^*(x) = \frac{1}{2}$ and

$$C(G) = \mathbb{E}_{x \sim p_{data}}[-log(2)] + \mathbb{E}_{x \sim p_g}[-log(2)] = -log(4).$$

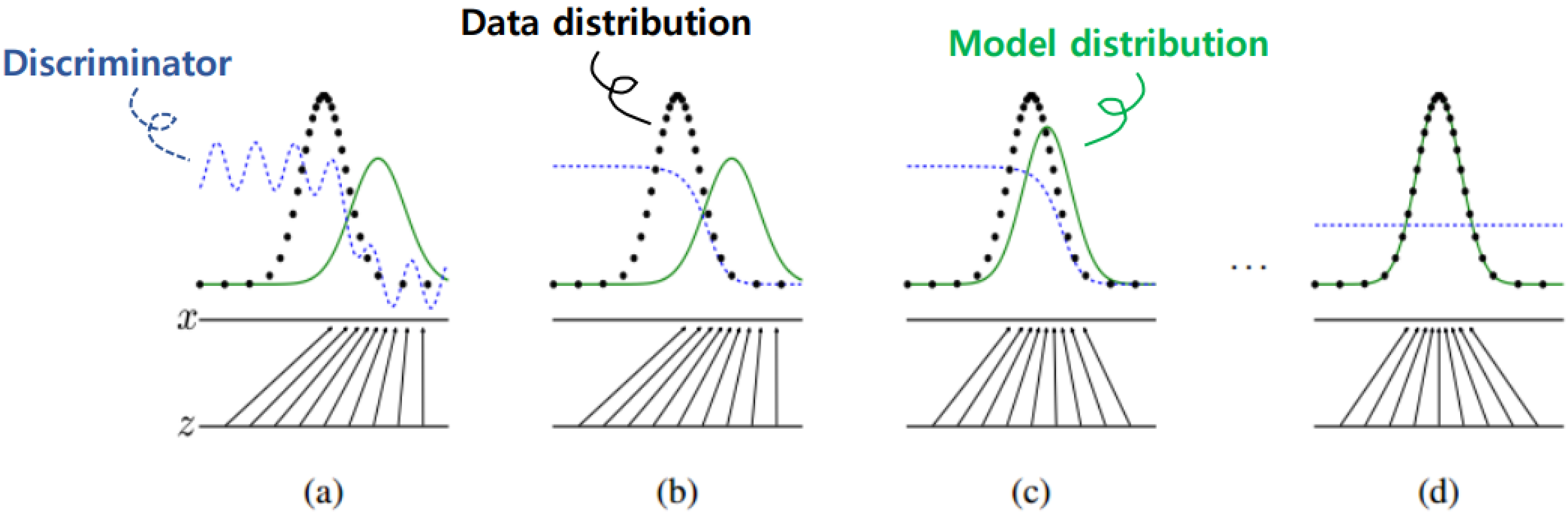
To show that this is the best possible value of $C(G)$:

$$\begin{aligned} C(G) &= -log(4) + KL\left(p_{data}\middle\|\frac{p_{data} + p_g}{2}\right) + KL\left(p_g\middle\|\frac{p_{data} + p_g}{2}\right) \\ &= -log(4) + 2 \cdot JSD(p_{data}\|p_g). \end{aligned}$$

$C(G)$가 앞의 수식을 변형한 식이 되고 $JSD(Pdata||Pg) \geq 0$이므로

C(G)의 global minimum은 –log4가 되고 이것은 $p_g = p_{data}$일때만 가능

GAN – Objective Function



GAN 의 한계

- 항상 optimal한 Discriminator를 만들어 내기가 힘들다.

  - Discriminator가 optimal 하지 않은 경우 generator가 discriminator가 구분하지 못하는 data를 계속 만들어 낼 수 있다.(mode collapse)

- 큰 수의 data set을 필요로 한다.

- Diminished gradient : discriminator가 너무 완벽하면 generator의 gradient사라짐

- 학습이 어렵다

- 텍스트를 생성하는데 적용이 어렵다.

- GAN의 결과물 자체가 새롭게 만들어진 sample이라서 기존 sample과 비교하여 얼마나 비슷한 지를 확인할 수 있는 정량적 척도가 없고, 어떤 형태로 그 결과가 나오게 되었는지 그 과정을 알 수 없다.

Thank you