생성 모델과 시각 지능

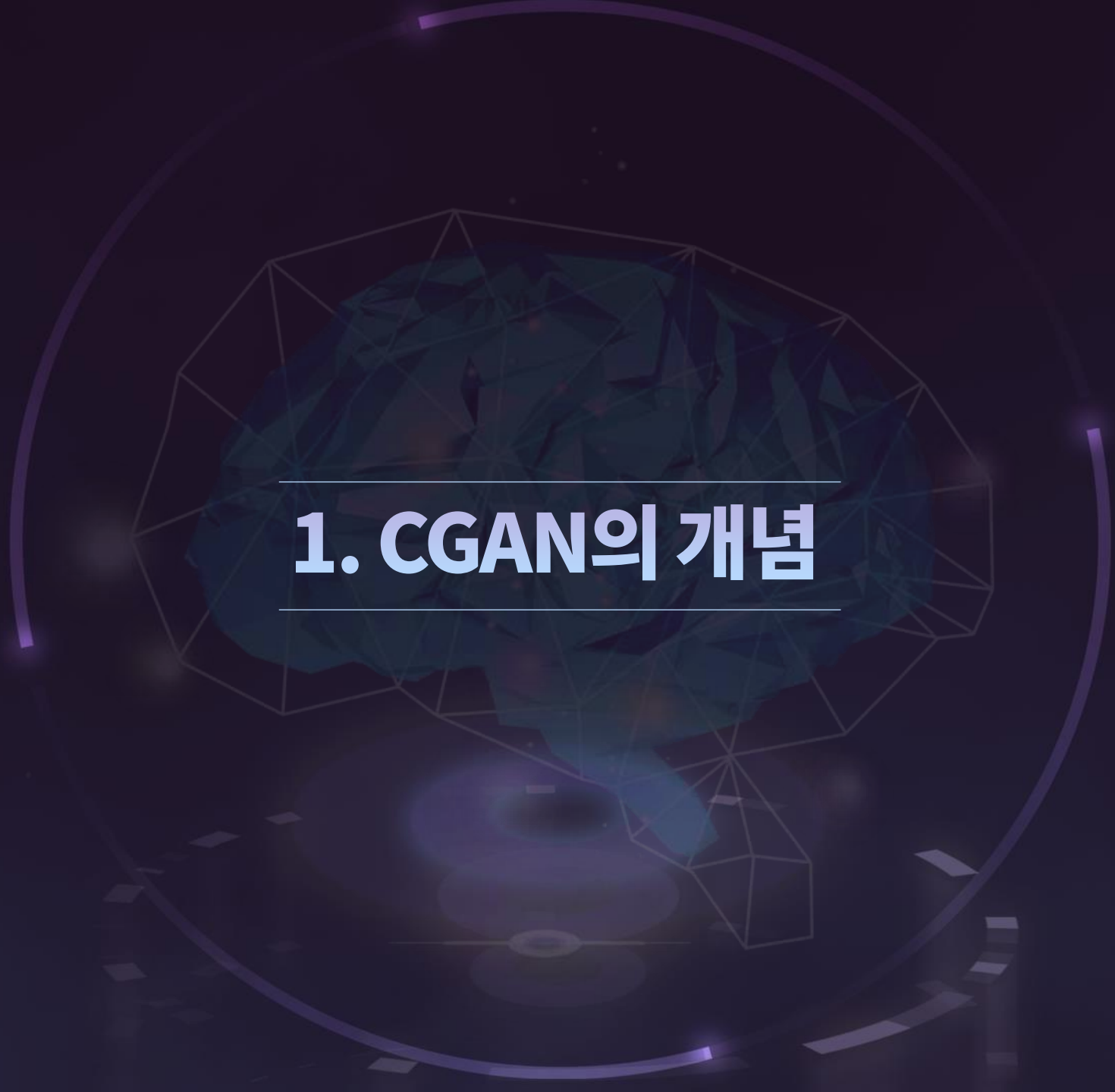
Generative Model and Visual Intelligence

07 주차 | GAN 발전 3

상명대학교 컴퓨터과학과
민 경 하

상명대학교 SANGMYUNG UNIVERSITY

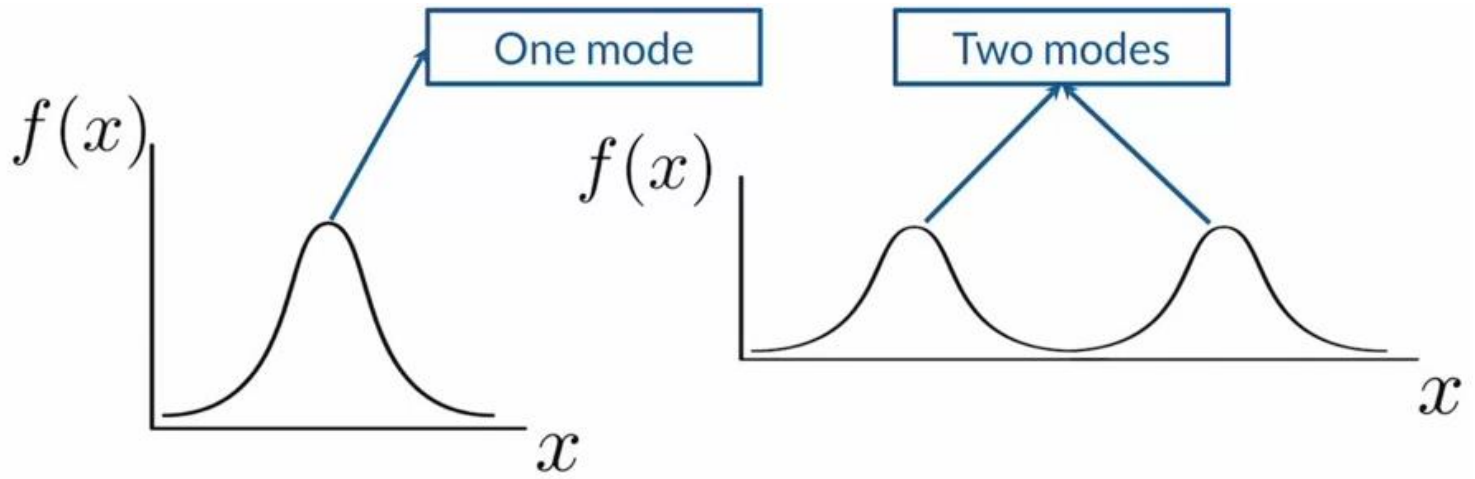
학습목차

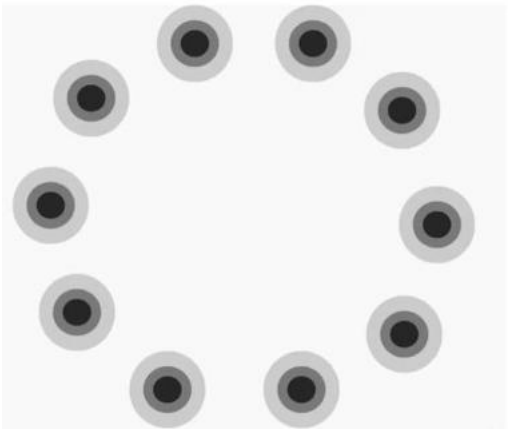


1. CGAN의 개념

CGAN(Conditional GAN)

- Control the results of GAN by assigning conditions



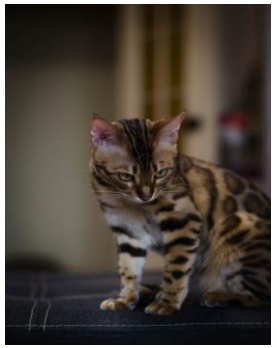


출처: Mirza, M. and Osindero, S., Conditional GAN, https://arxiv.org/abs/1411.1784, 2014.

CGAN(Conditional GAN)

- Key idea
  - Input: latent vector (z) + condition vector (y)
- Example: 7 cat classes


Korean short hair
[1]

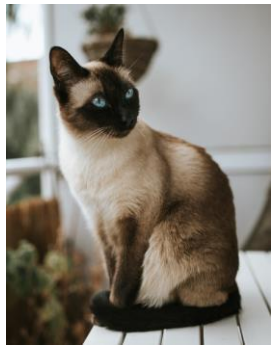

bengal
[2]


munchkin
[3]


persian
[4]


Russian blue
[5]


Scottishfold
[6]

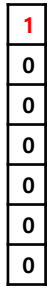

Siamese
[7]

출처: https://unsplash.com/photos/AGehl6k8xVo
출처: https://unsplash.com/photos/KGiQFgF7dkc
출처: https://unsplash.com/photos/NJQkNgL7cLY
출처: https://unsplash.com/photos/LvrfTxNE-WQ
출처: https://www.mk.co.kr/news/culture/view/2021/07/701143/
출처: http://catfriends.co.kr/product/%EC%8A%A4%EC%BD%94%ED%8B%B0%EC%89%AC%ED%8F%B4%EB%93%9C-%EC%99%84%EB%8B%A4/14/
출처: https://ko.wikipedia.org/wiki/%EB%A8%BC%EC%B9%98%ED%82%A8_(%EA%B3%A0%EC%96%91%EC%9D%B4)

CGAN(Conditional GAN)

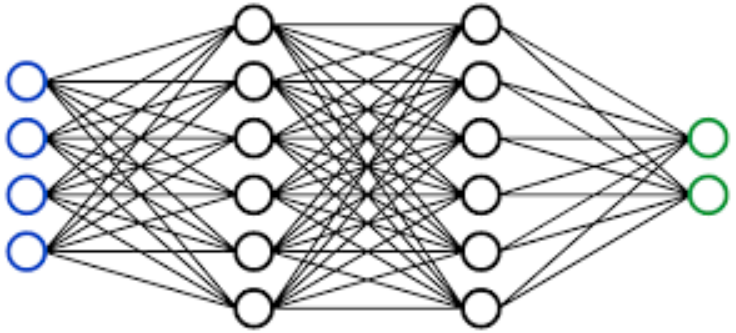
- Input: latent vector (z) + condition vector (y)

입력: Condition vector (y)



Latent vector (z)

$G: p_{z/y} \rightarrow p_g$

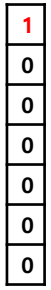


Generator (G)
(θ_g)

Korean short hair
[1]

CGAN(Conditional GAN)

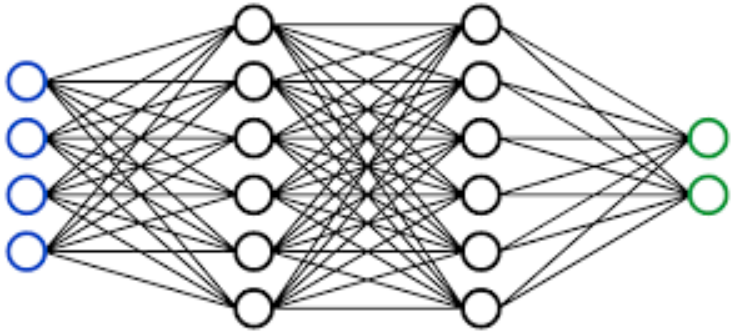
- Input: latent vector (z) + condition vector (y)



입력: Condition vector (y)

Latent vector (z)

$G: p_{z/y} \rightarrow p_g$



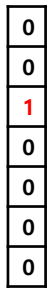
Generator (G)
(θ_g)

Korean short hair
[1]

출처: https://pixabay.com/ko/photos/%EA%B3%A0%EC%96%91%EC%9D%B4-%EC%83%88%EB%81%BC%EA%B3%A0%EC%96%91%EC%9D%B4-%EA%B8%B8%EA%B3%A0%EC%96%91%EC%9D%B4-5270321/

CGAN(Conditional GAN)

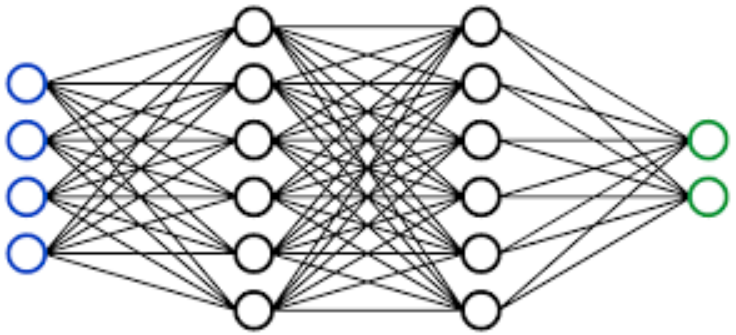
- Input: latent vector (z) + condition vector (y)



입력: Condition vector (y)

Latent vector (z)

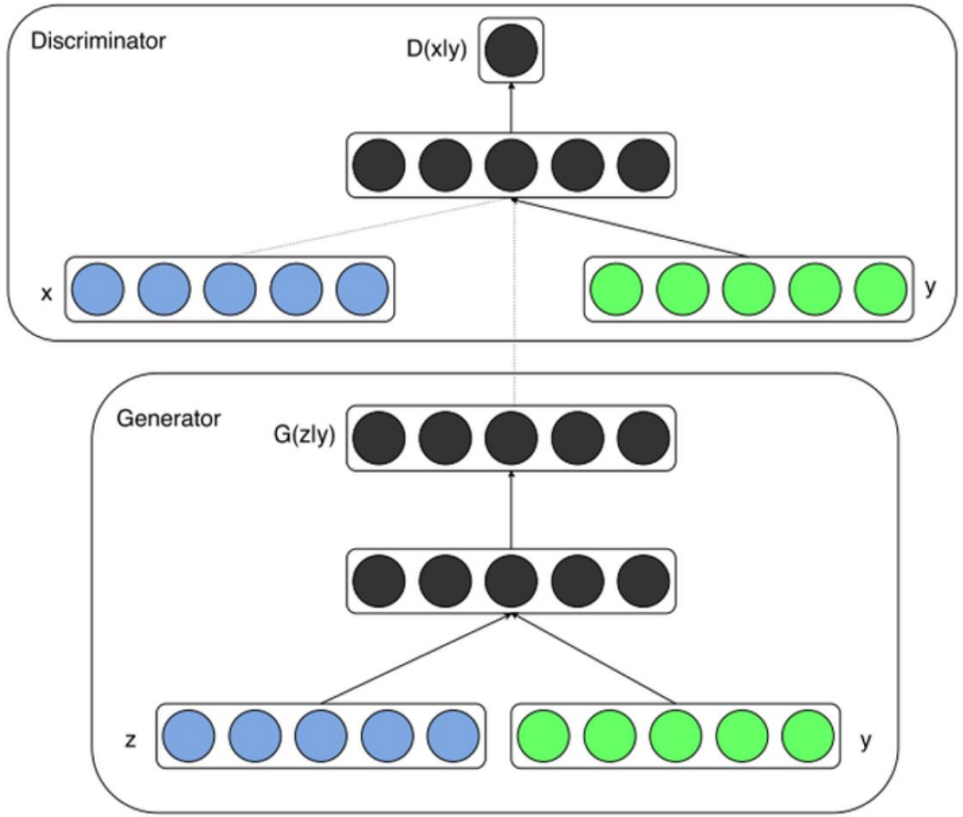
$G: p_{z/y} \rightarrow p_g$



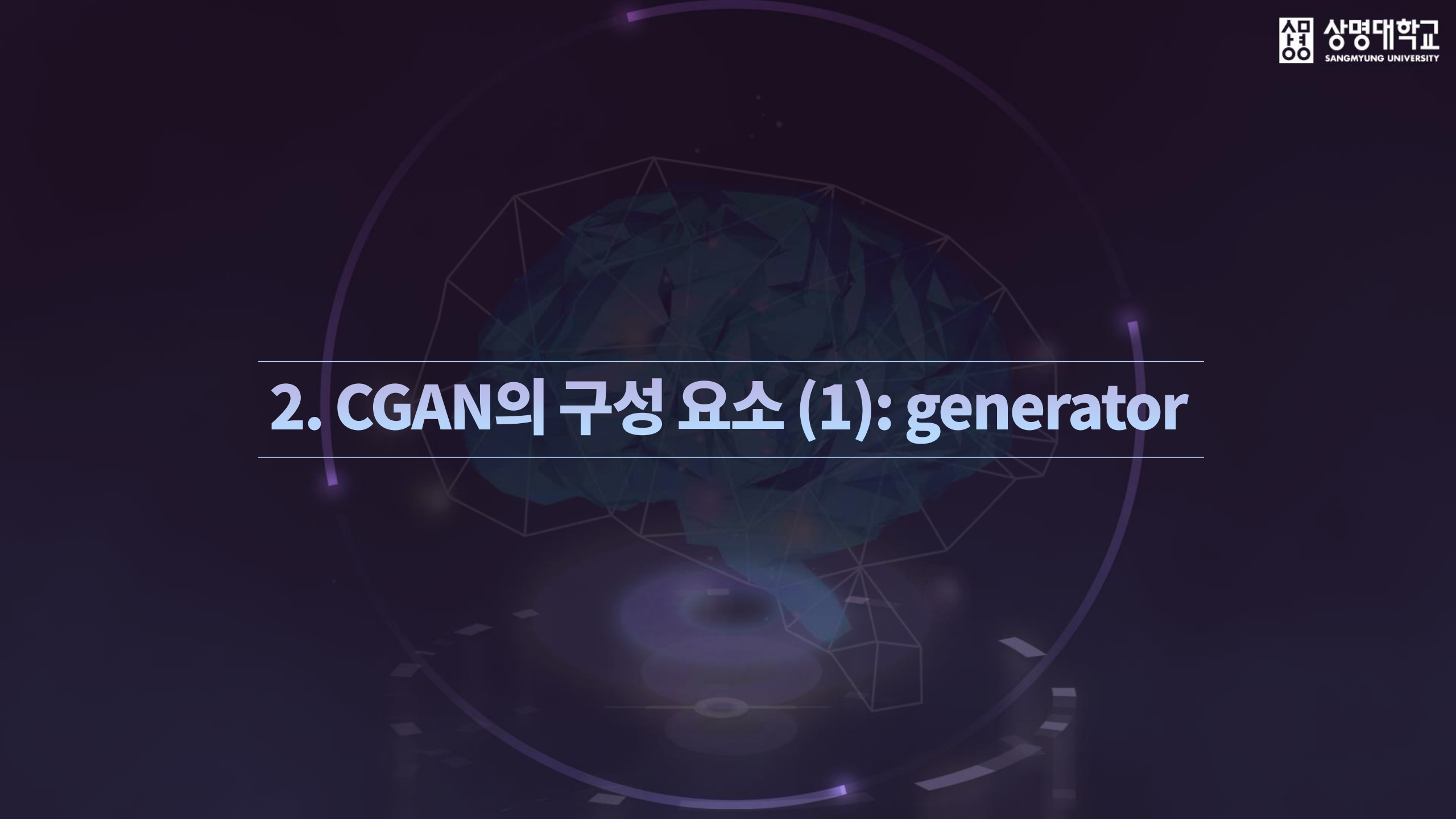
Generator (G)
(θ_g)

munchkin
[3]

CGAN (Conditional GAN)의 구조



출처: CGAN(2014), Conditional Generative Adversarial Nets

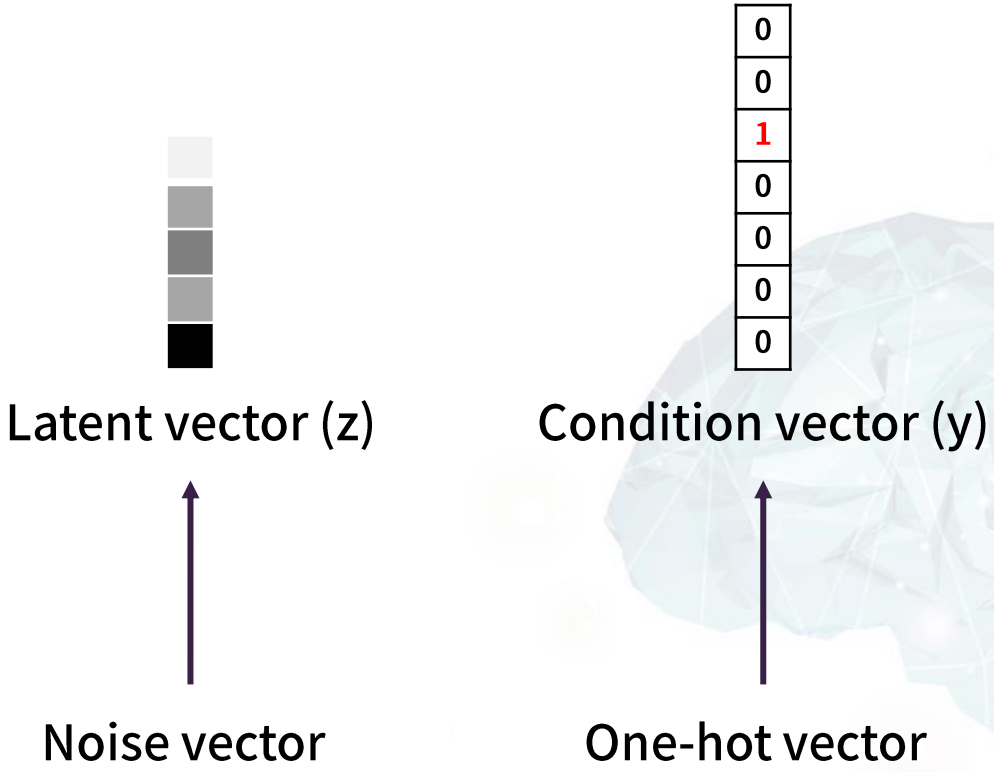


2. CGAN의 구성 요소 (1): generator

2. CGAN의 구성 요소 (1): generator

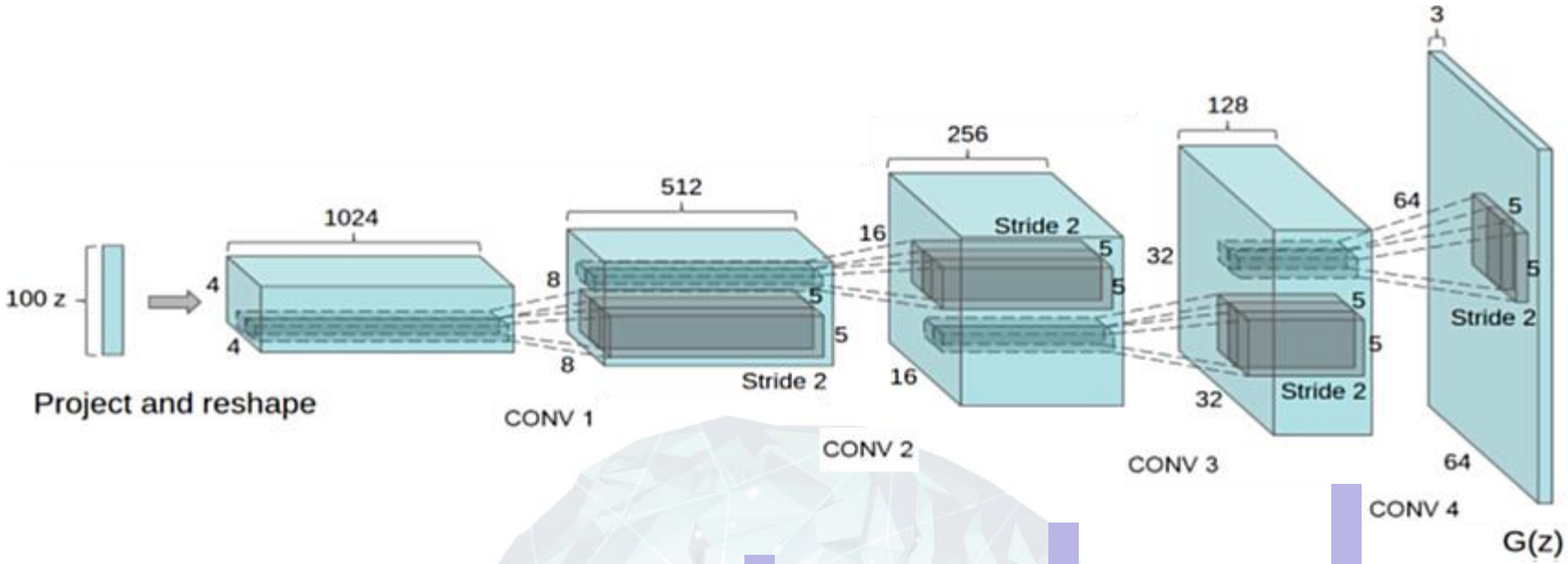
CGAN (Conditional GAN)

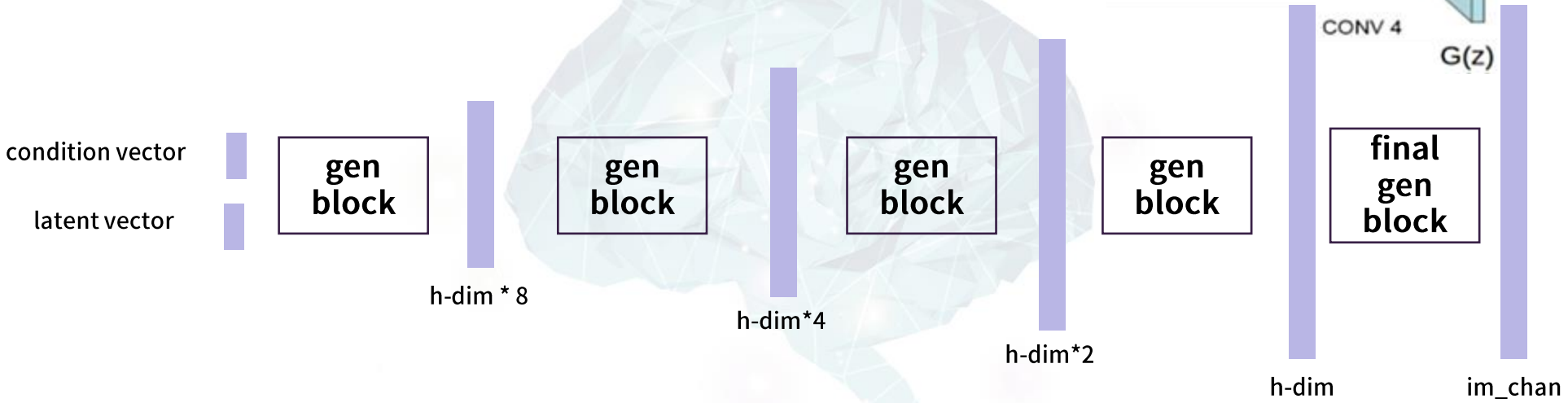
- Input: latent vector (z) + condition vector (y)



2. CGAN의 구성 요소 (1): generator

CGAN (Conditional GAN)





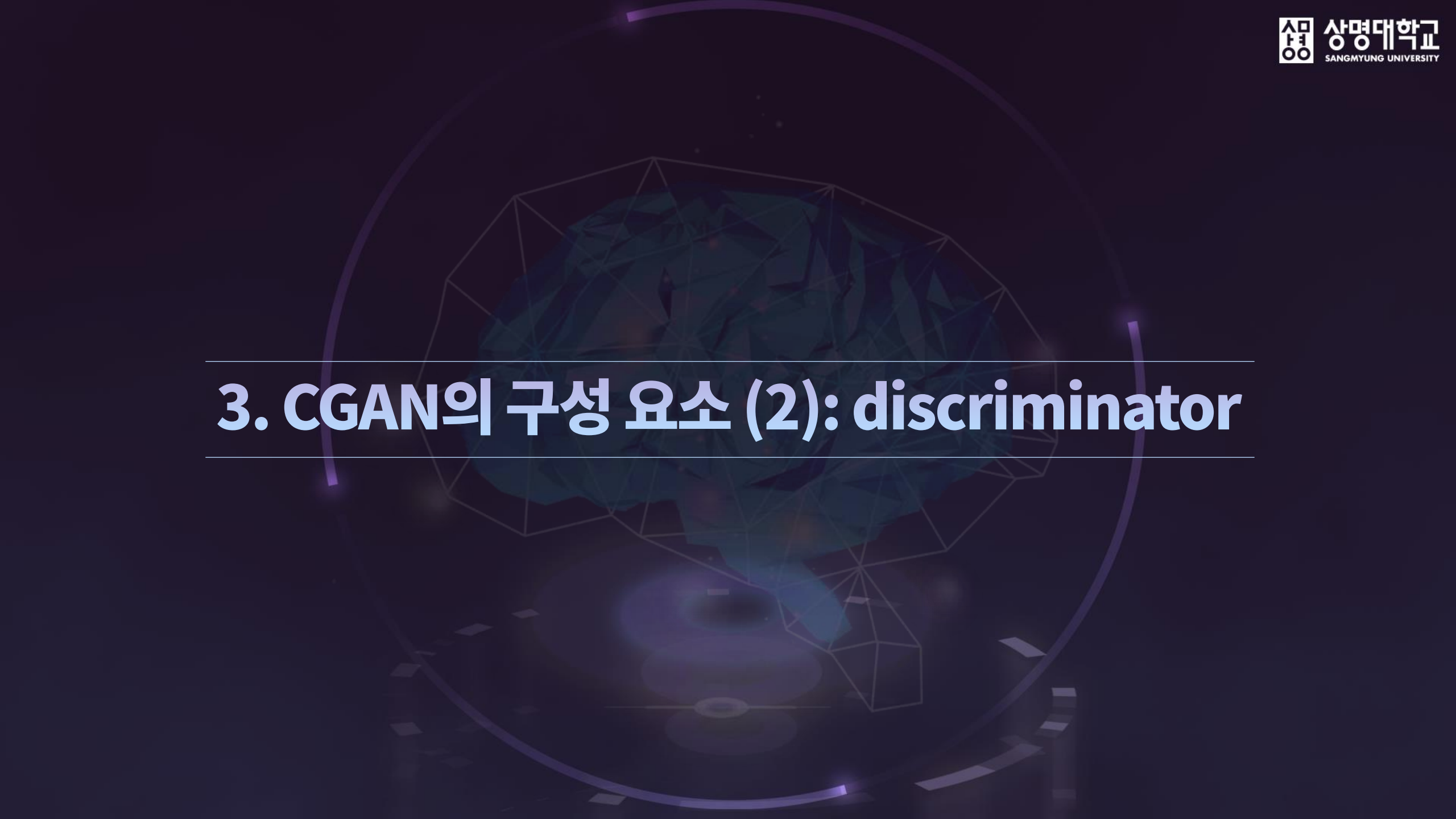
출처: https://ydy8989.github.io/2021-02-05-gan2/

Gen block

- Parameter
  : input_channels, output_channels, kernel, stride, final_layer

- Components for internal
  - Transposed convolution + batch norm + ReLU

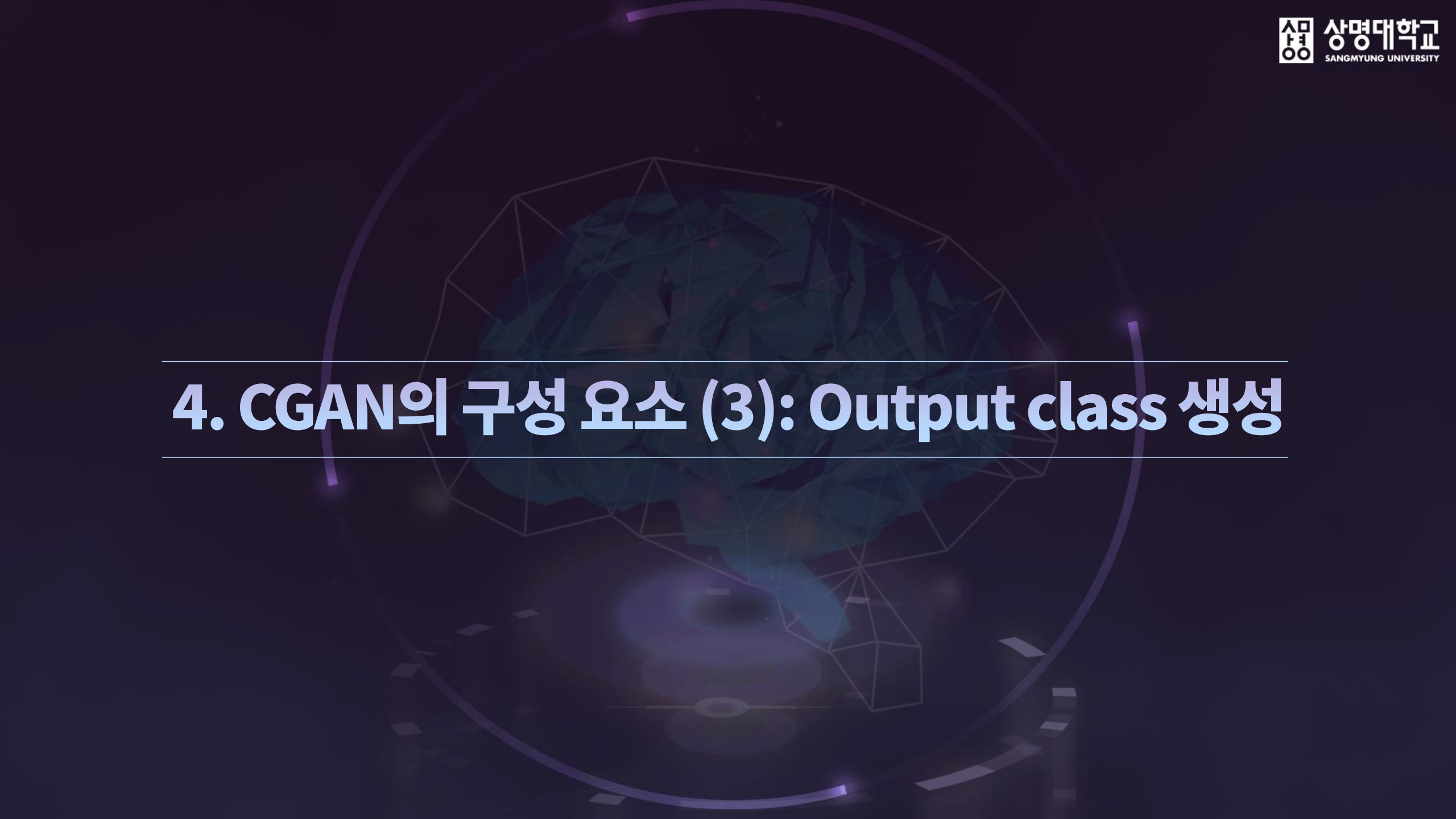
- Components for final
  - Transposed convolution + tanh



3. CGAN의 구성 요소 (2): discriminator

Discriminator

- Disc block
  - Parameter: input_channels, output_channels, kernel, stride, final_layer)
  - Components for internal
    - » convolution + batch norm + LeakyReLU (0.2)
  - Components for final
    - » convolution



4. CGAN의 구성 요소 (3): Output class 생성

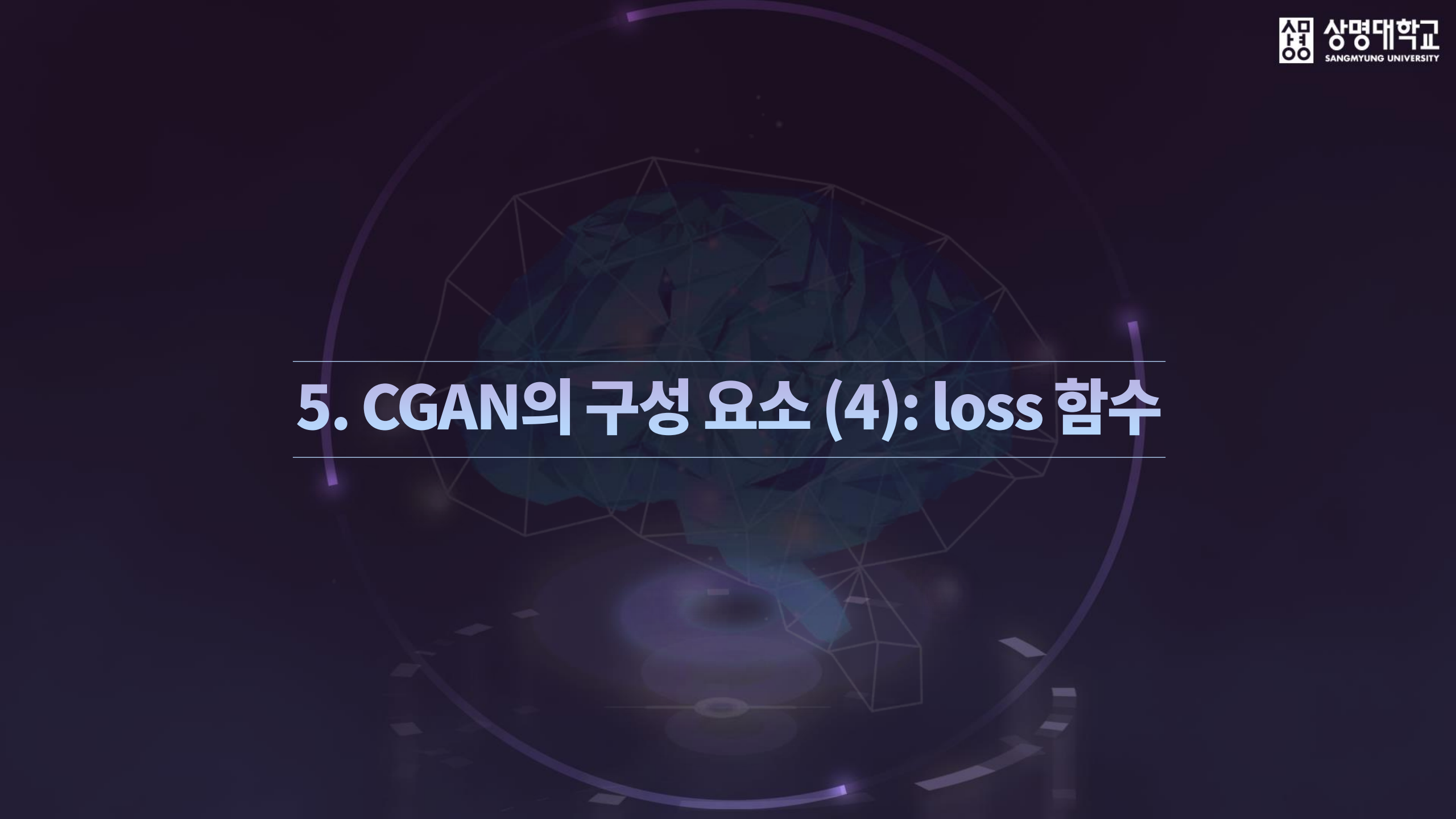
4. CGAN의 구성 요소 (3): Output class 생성

One-hot vector 생성

```
import torch.nn.functional as F
def get_one_hot_labels(labels, n_classes):
    return F.one_hot(labels, n_classes)
```

Latent vector와 one-hot vector의 concatenation

```
def combine_vectors(x, y):
    combined = torch.cat((x.float(), y.float()), 1)
    return combined
```



5. CGAN의 구성 요소 (4): loss 함수

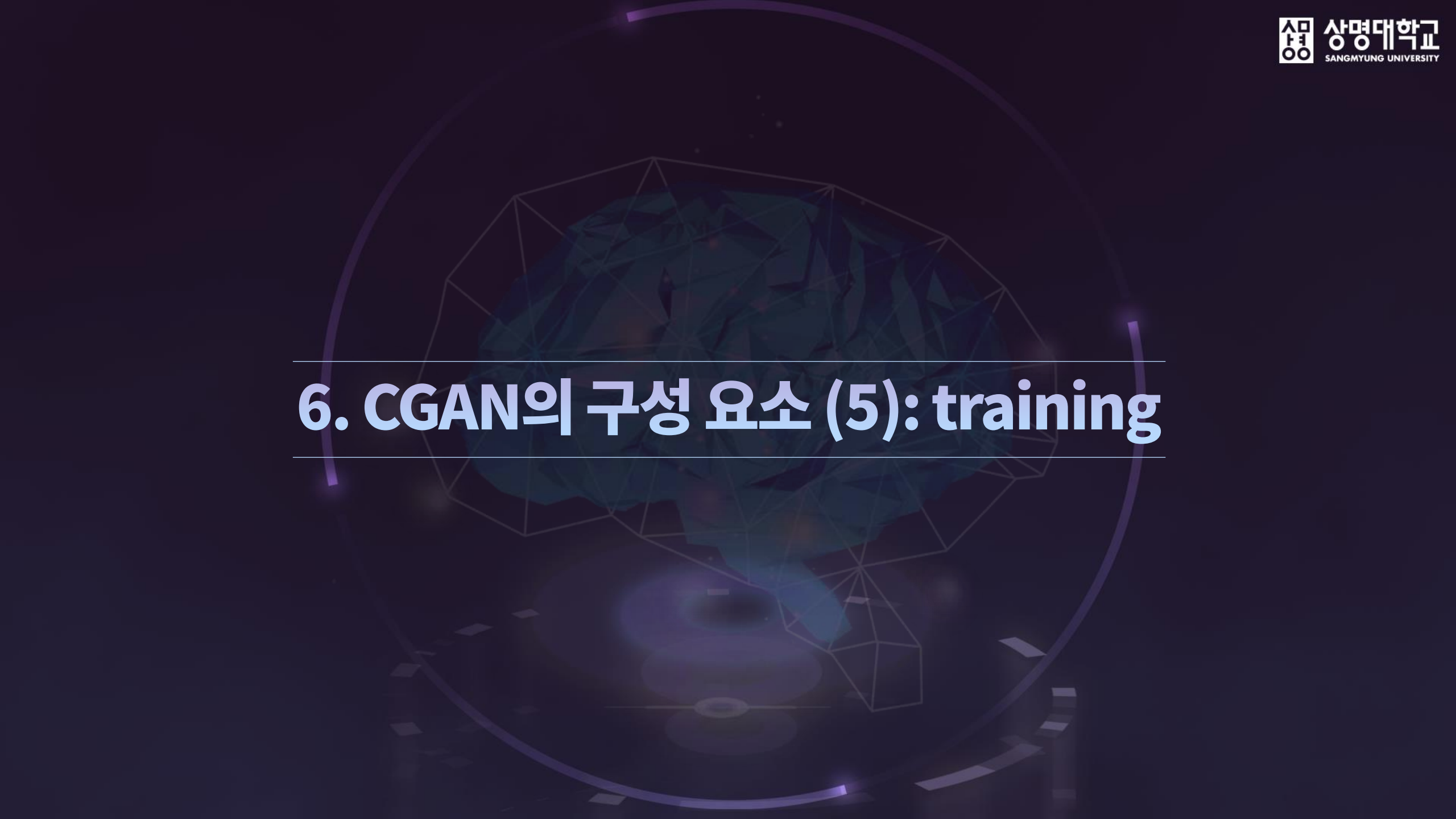
loss 함수

Vanilla GAN의 loss 함수

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

CGAN의 loss 함수

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



6. CGAN의 구성 요소 (5): training

6. CGAN의 구성 요소 (5): training

DCGAN의 training

update discriminator

```
disc_opt.zero_grad ()

fake_noise = get_noise(cur_batch_size, z_dim, device=device) # z
fake = gen(fake_noise) # G(z)

disc_fake_pred = disc(fake.detach())
disc_real_pred = disc(real)

disc_fake_loss = criterion(disc_fake_pred, torch.zeros_like(disc_fake_pred))
disc_real_loss = criterion(disc_real_pred, torch.ones_like(disc_real_pred))
disc_loss = (disc_fake_loss + disc_real_loss) / 2
```

6. CGAN의 구성 요소 (5): training

CGAN의 training

update discriminator

```
disc_opt.zero_grad()

    one_hot_labels = get_one_hot_labels(labels.to(device), n_classes)
    image_one_hot_labels = one_hot_labels[:, :, None, None]
    image_one_hot_labels = image_one_hot_labels.repeat(1, 1, mnist_shape[1], mnist_shape[2])

    fake_noise = get_noise ( cur_batch_size, z_dim, device=device)

    noise_and_labels = combine_vectors(fake_noise, one_hot_labels)
    fake = gen(noise_and_labels)

    fake_image_and_labels = combine_vectors(fake, image_one_hot_labels)
    real_image_and_labels = combine_vectors(real, image_one_hot_labels)
    disc_fake_pred = disc(fake_image_and_labels.detach())
    disc_real_pred = disc(real_image_and_labels)

    disc_fake_loss = criterion(disc_fake_pred, torch.zeros_like(disc_fake_pred))
    disc_real_loss = criterion(disc_real_pred, torch.ones_like(disc_real_pred))
    disc_loss = (disc_fake_loss + disc_real_loss)/2
```

6. CGAN의 구성 요소 (5): training

DCGAN의 training

update generator

```
gen_opt.zero_grad ()

   fake_noise_2 = get_noise (cur_batch_size, z_dim, device=device)
   fake_2 = gen(fake_noise_2)

   disc_fake_pred = disc(fake_2)
   gen_loss = criterion ( disc_fake_pred, torch.ones_like(disc_fake_pred))
   gen_loss.backward()
   gen_opt.step()
```

6. CGAN의 구성 요소 (5): training

CGAN의 training

update generator

```
gen_opt.zero_grad()

    fake_image_and_labels = combine_vectors(fake, image_one_hot_labels)

    disc_fake_pred = disc(fake_image_and_labels)

    gen_loss = criterion(disc_fake_pred, torch.ones_like(disc_fake_pred))
    gen_loss.backward()
    gen_opt.step()
```