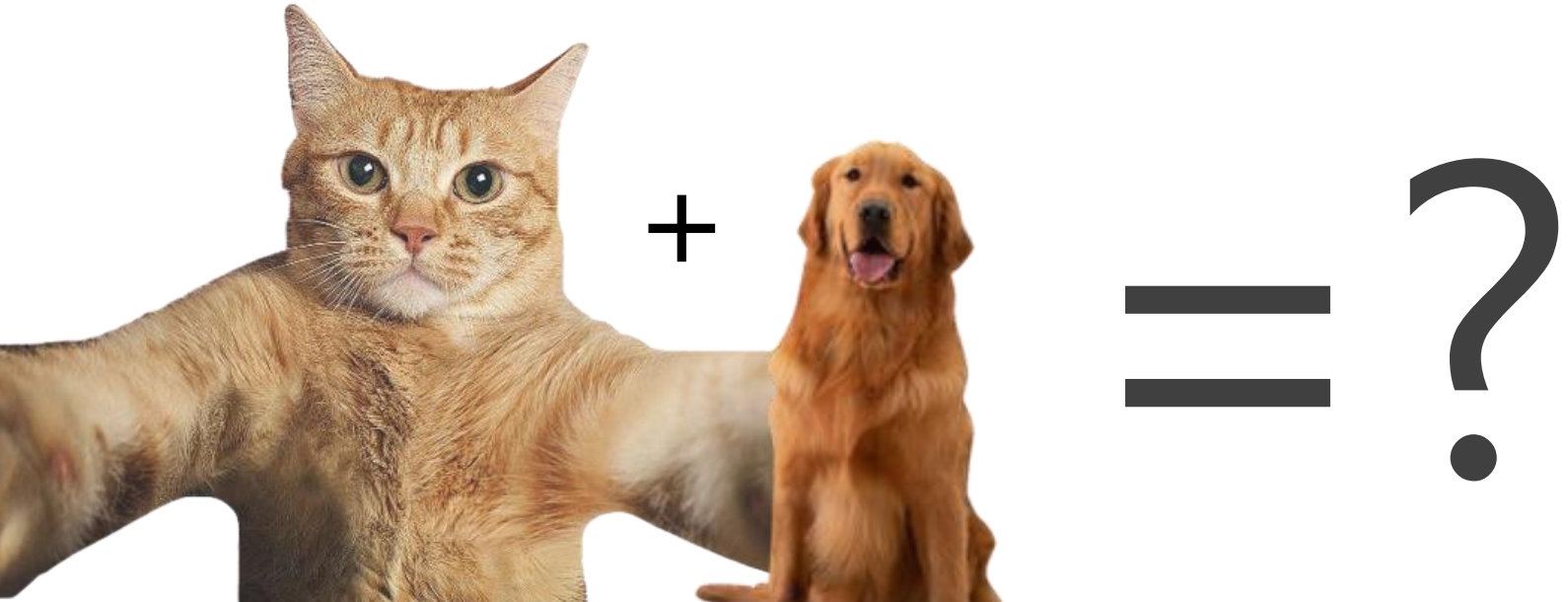
Convolutional-Conditional GAN을 사용한 개냥이 만들기

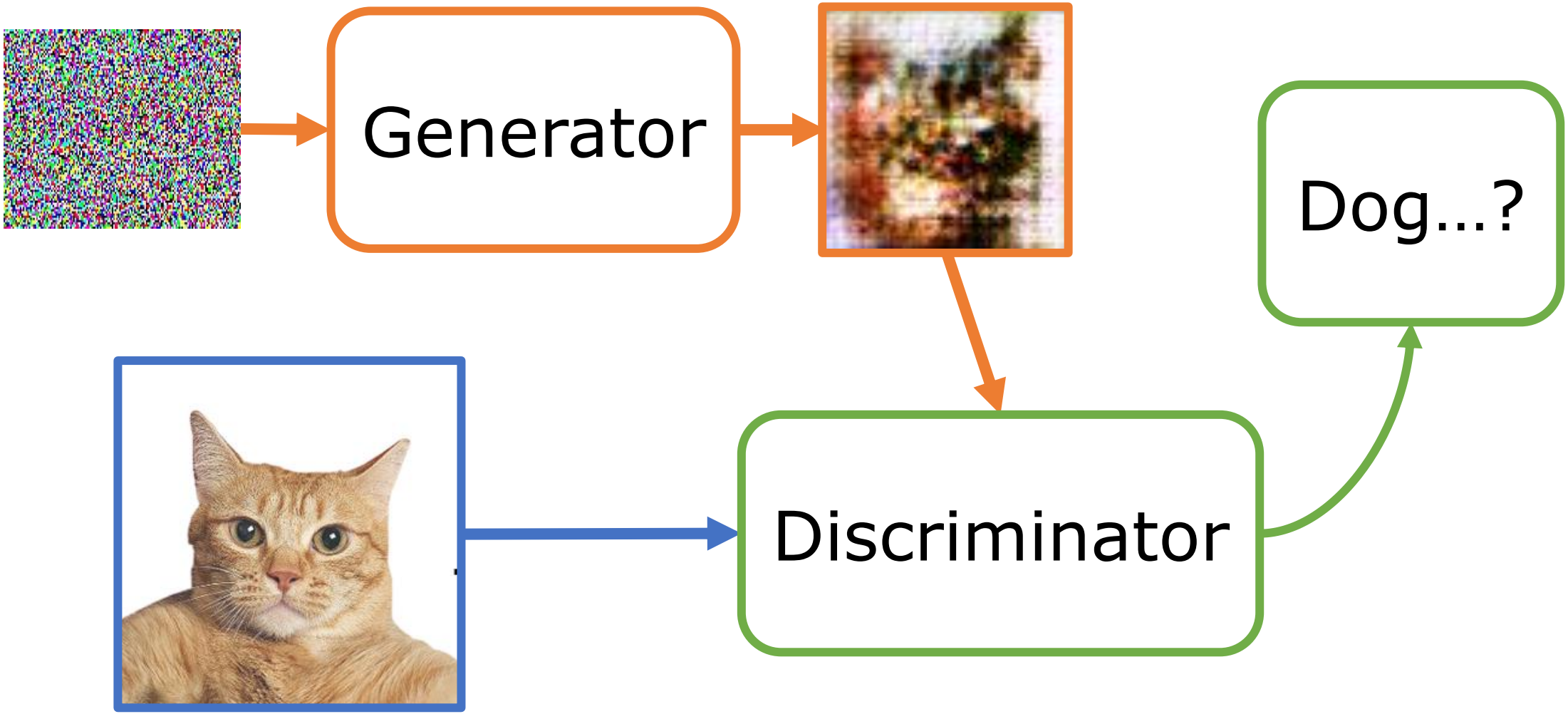
2019320030 컴퓨터학과 황수인



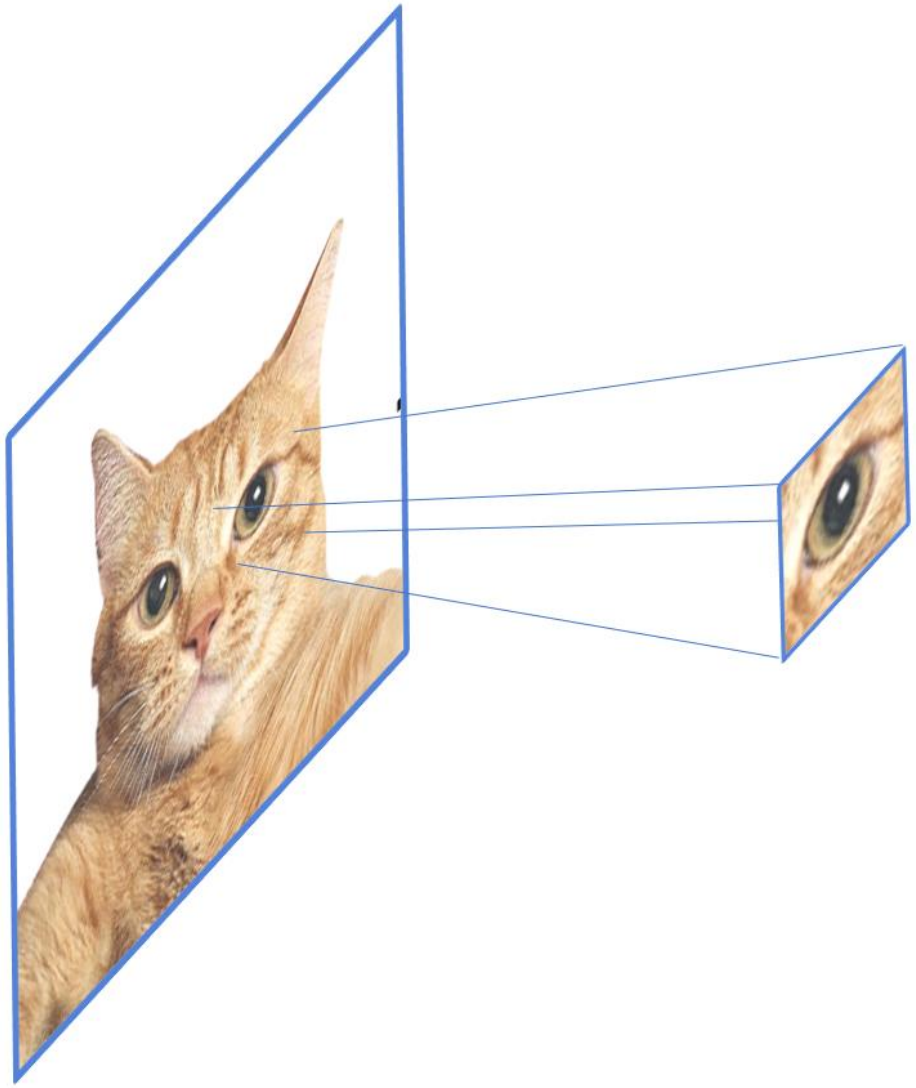
+

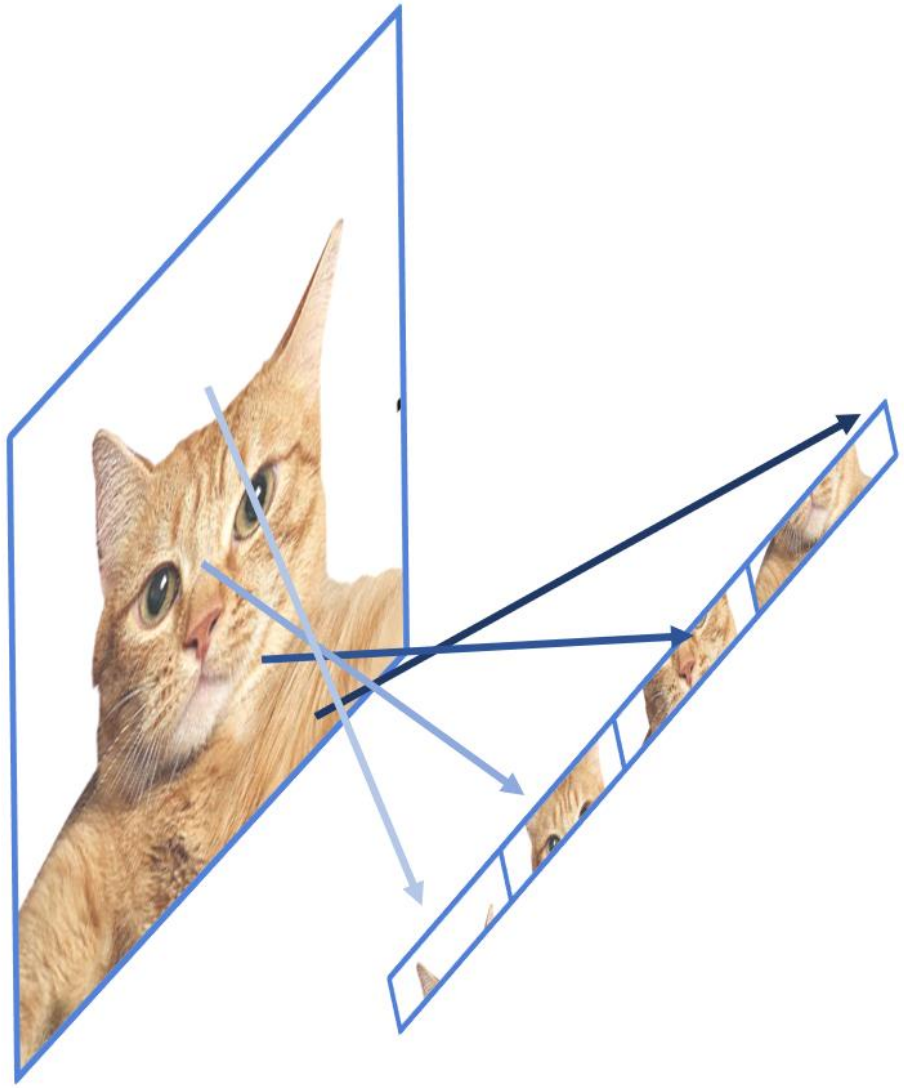
=?

GAN, Generative Adversarial Networks

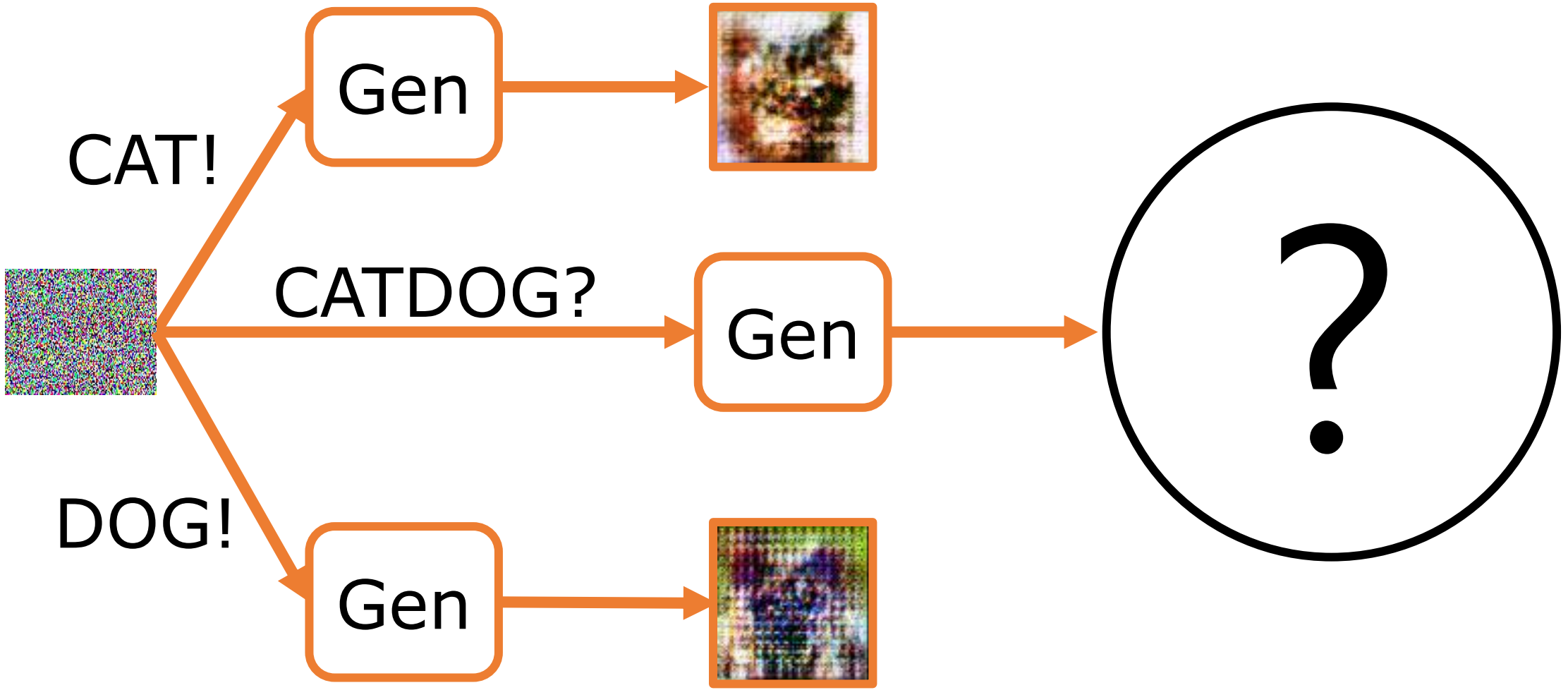


Why Convolutional?





Why Conditional?



Datasets: **Animal Faces**

16,130 images belonging to 3 classes.

Cat
5,653

Dog
5,239

Wild
unused

kaggle www.kaggle.com/datasets/andrewmvd/animal-faces

Yunjey Choi, Youngjung Uh, Jaejun Yoo, & Jung-Woo Ha (2020). StarGAN v2: Diverse Image Synthesis for Multiple Domains. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*.

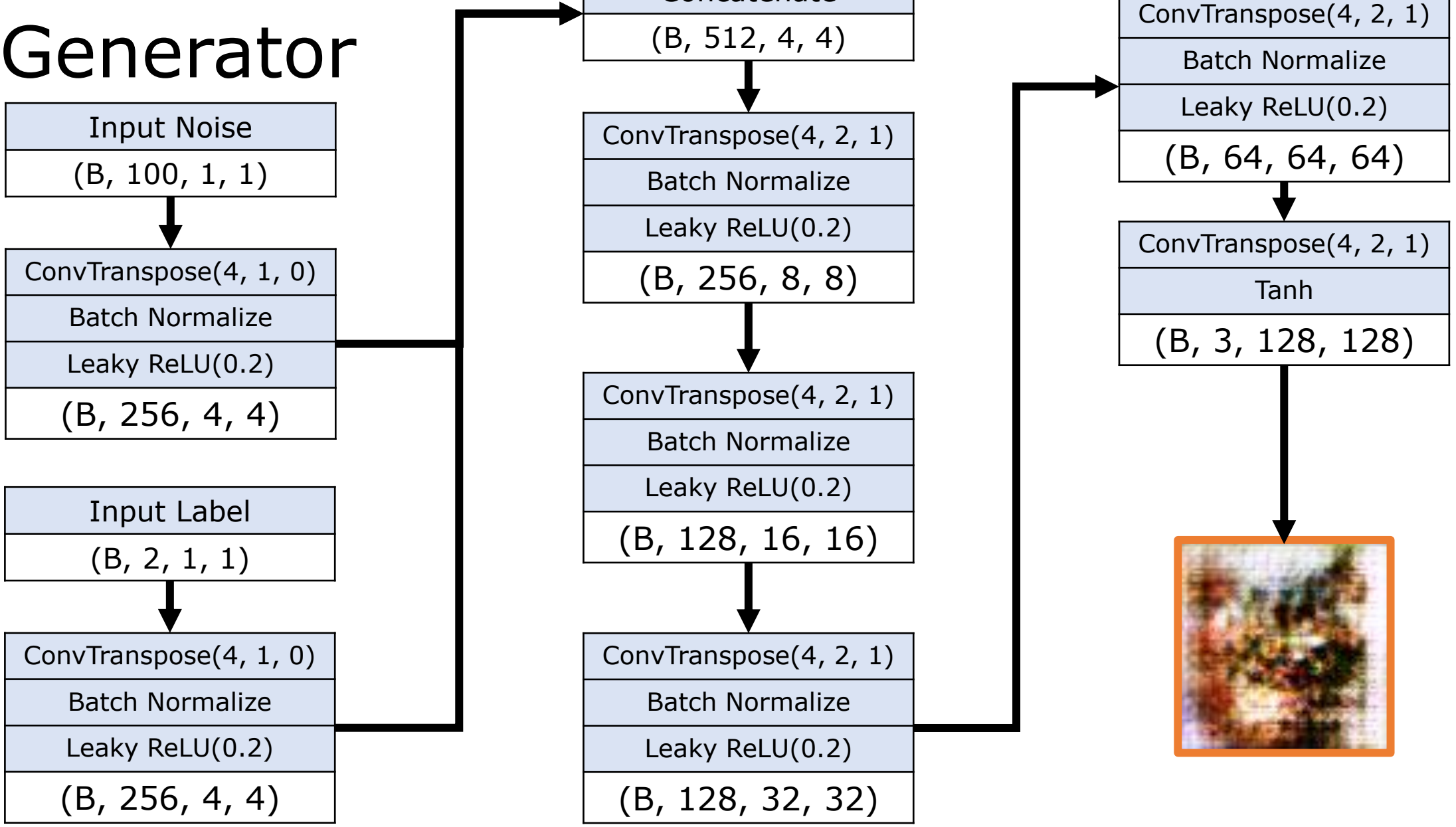
Data Preprocess

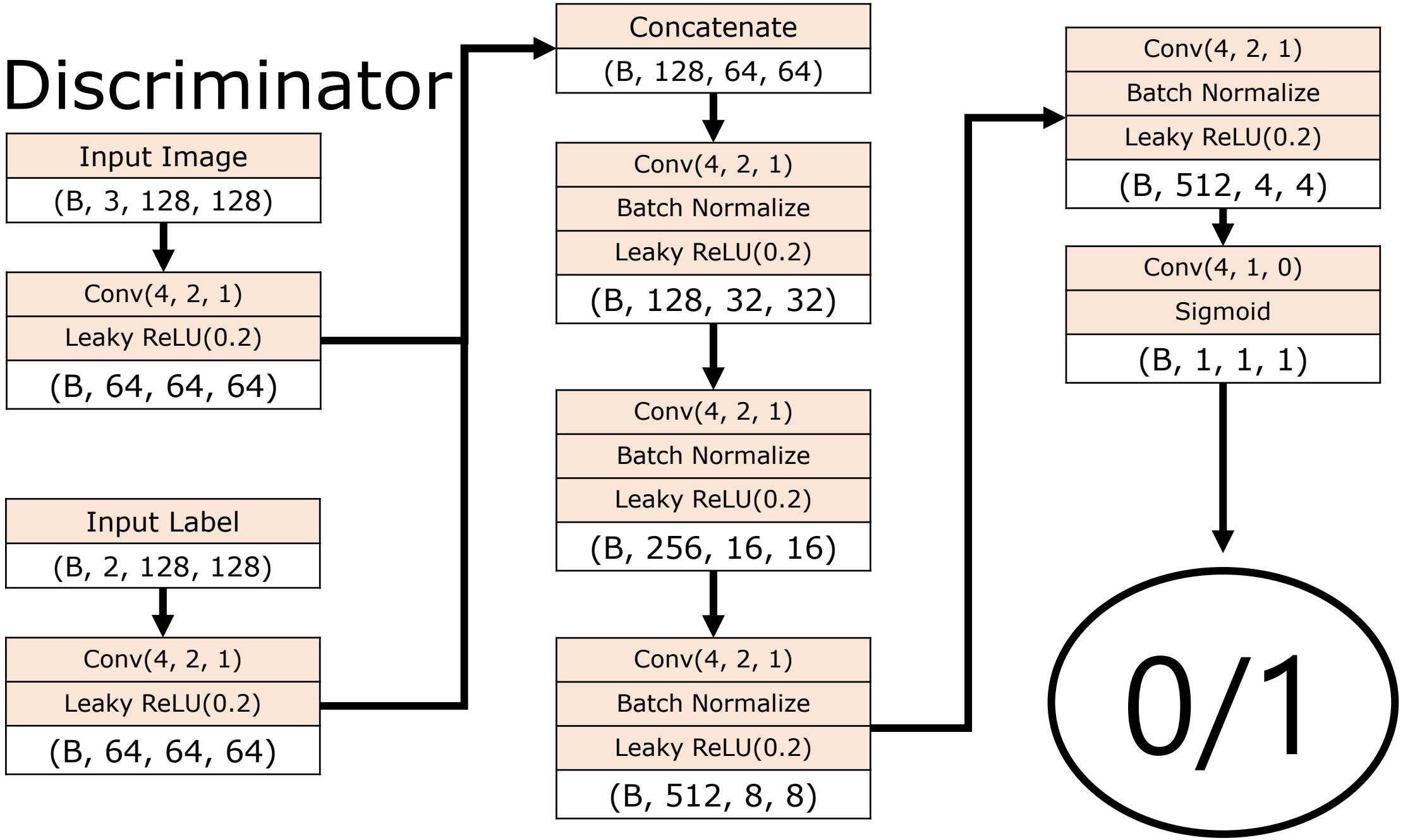
```
img_size = 128
batch_size = 128

dataset = dsets.ImageFolder(root='./afhq',
                            transform=transforms.Compose([
                                transforms.Resize(img_size), # (512 x 512) => (128 x 128)
                                transforms.ToTensor(),
                                transforms.Normalize((0.5, 0.5, 0.5), (0.5, 0.5, 0.5)),
                            ]))

dataloader = torch.utils.data.DataLoader(dataset, batch_size=batch_size, shuffle=True, num_workers=2)
```

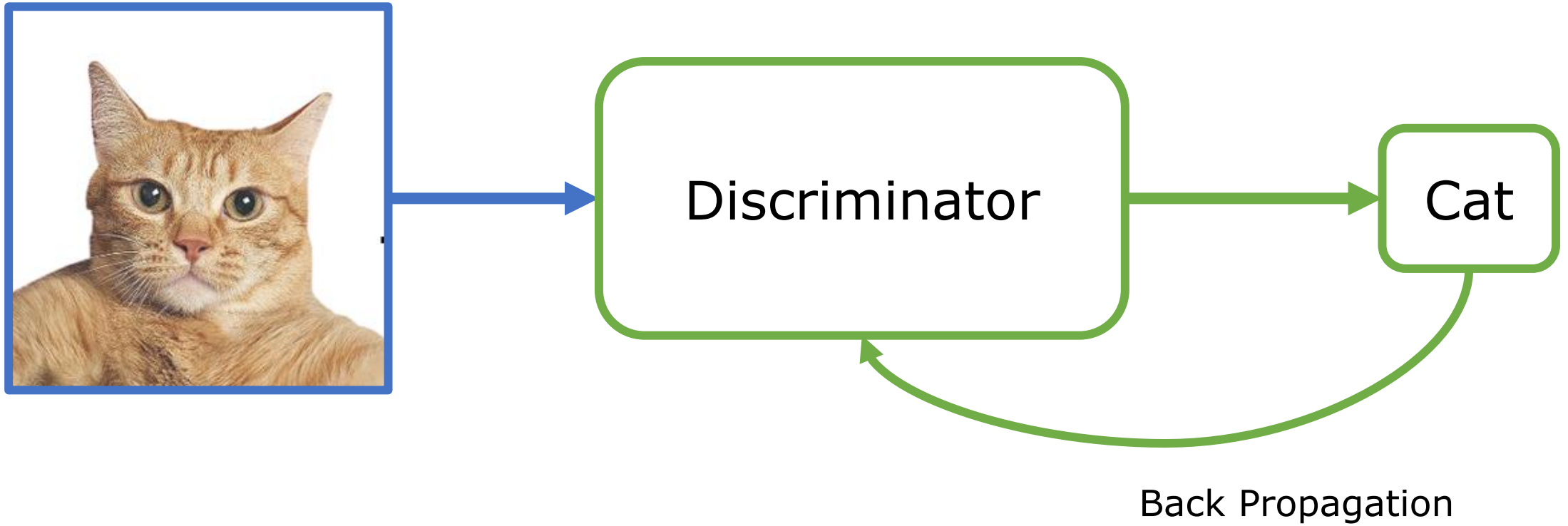
Generator


Input Noise
(B, 100, 1, 1)
ConvTranspose(4, 1, 0)
Batch Normalize
Leaky ReLU(0.2)
(B, 256, 4, 4)
Input Label
(B, 2, 1, 1)
ConvTranspose(4, 1, 0)
Batch Normalize
Leaky ReLU(0.2)
(B, 256, 4, 4)
Concatenate
(B, 512, 4, 4)
ConvTranspose(4, 2, 1)
Batch Normalize
Leaky ReLU(0.2)
(B, 256, 8, 8)
ConvTranspose(4, 2, 1)
Batch Normalize
Leaky ReLU(0.2)
(B, 128, 16, 16)
ConvTranspose(4, 2, 1)
Batch Normalize
Leaky ReLU(0.2)
(B, 128, 32, 32)
ConvTranspose(4, 2, 1)
Batch Normalize
Leaky ReLU(0.2)
(B, 64, 64, 64)
ConvTranspose(4, 2, 1)
Tanh
(B, 3, 128, 128)


Discriminator
Input Image
(B, 3, 128, 128)
Conv(4, 2, 1)
Leaky ReLU(0.2)
(B, 64, 64, 64)
Input Label
(B, 2, 128, 128)
Conv(4, 2, 1)
Leaky ReLU(0.2)
(B, 64, 64, 64)
Concatenate
(B, 128, 64, 64)
Conv(4, 2, 1)
Batch Normalize
Leaky ReLU(0.2)
(B, 128, 32, 32)
Conv(4, 2, 1)
Batch Normalize
Leaky ReLU(0.2)
(B, 256, 16, 16)
Conv(4, 2, 1)
Batch Normalize
Leaky ReLU(0.2)
(B, 512, 8, 8)
Conv(4, 2, 1)
Batch Normalize
Leaky ReLU(0.2)
(B, 512, 4, 4)
Conv(4, 1, 0)
Sigmoid
(B, 1, 1, 1)
0/1

Training process

1. Training Discriminator with Real Data



Training process

2. Training Discriminator with Fake Data

Generator

Back Propagation

Not Cat

Discriminator

Training process

3. Training Generator

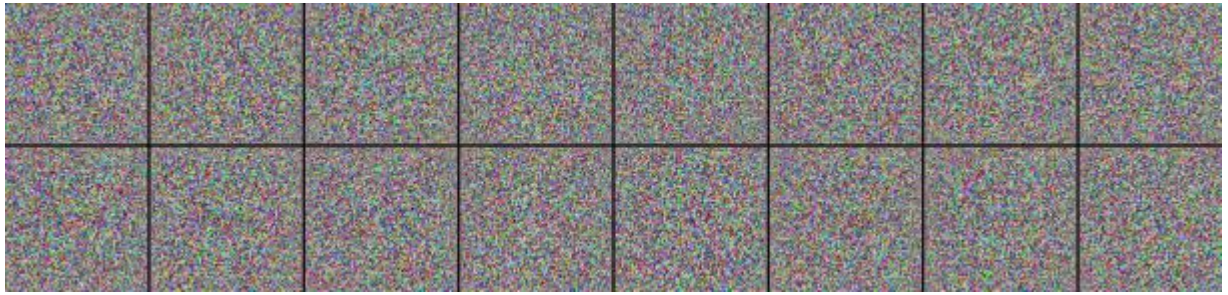
Generator

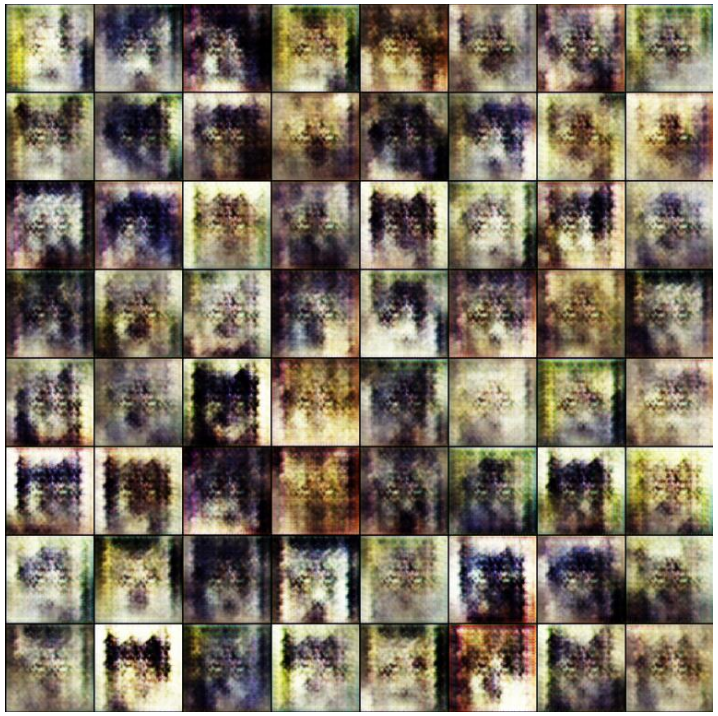
Back Propagation

Not Cat

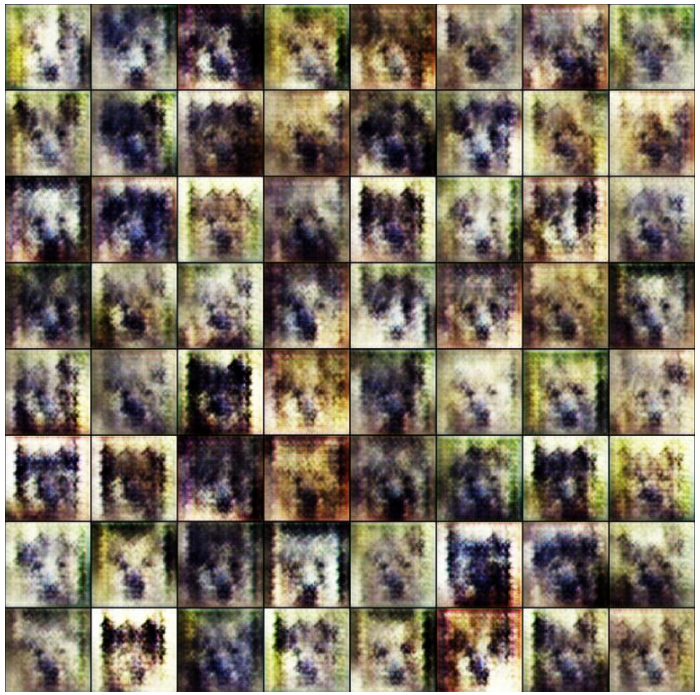
Discriminator

Results: 10 epochs

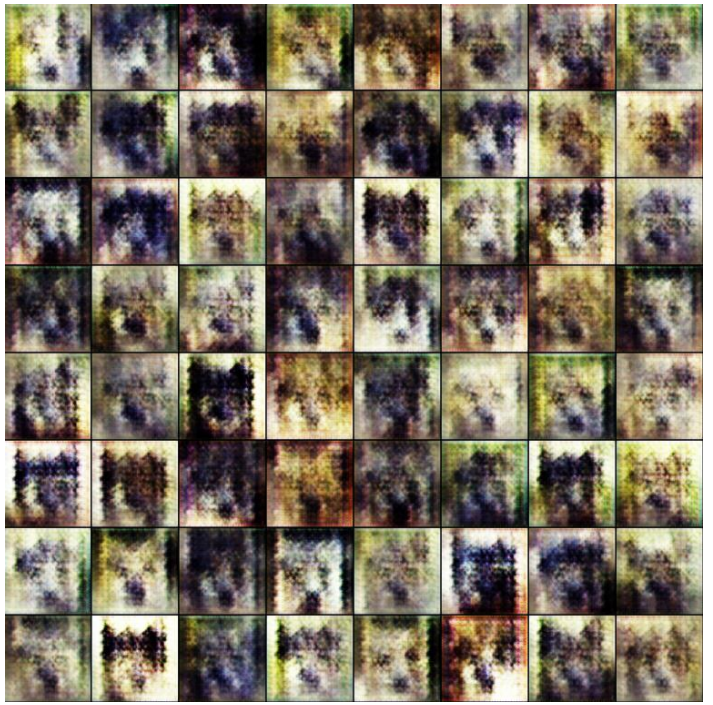




Cat



Dog



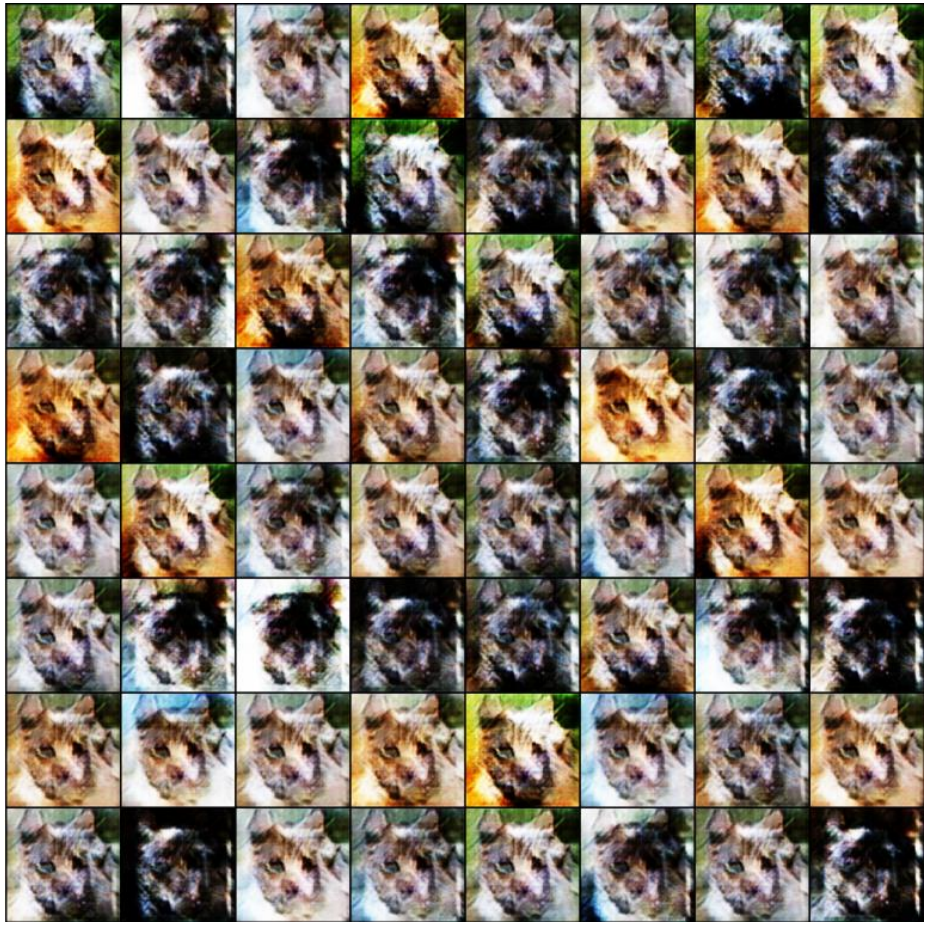
Catdog

Results: 100 epochs

Cat

Dog

Catdog



Results: 500 epochs

Cat

+

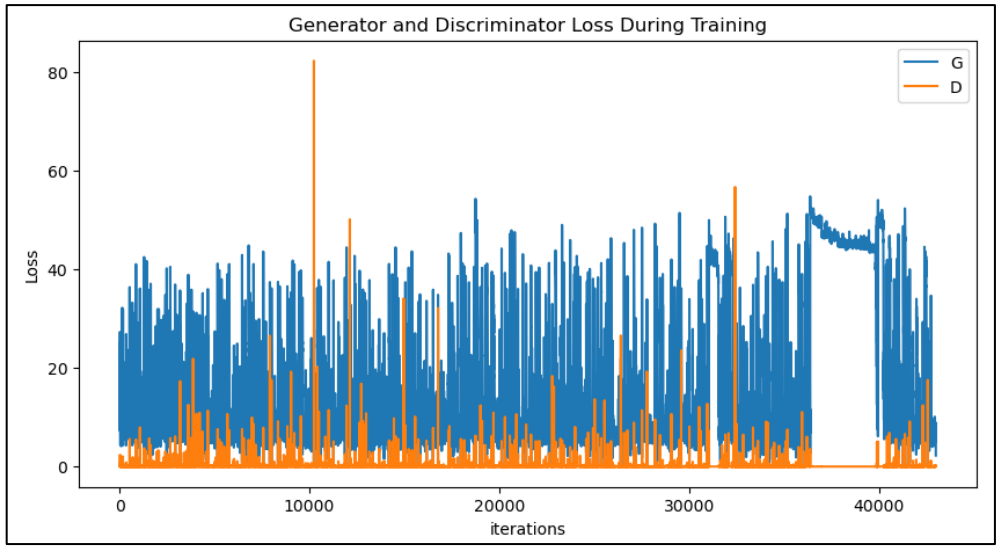
Dog

=

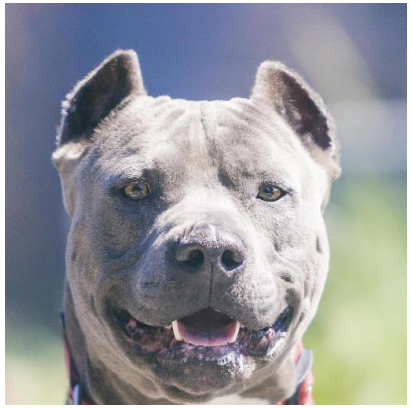
“Catdog”

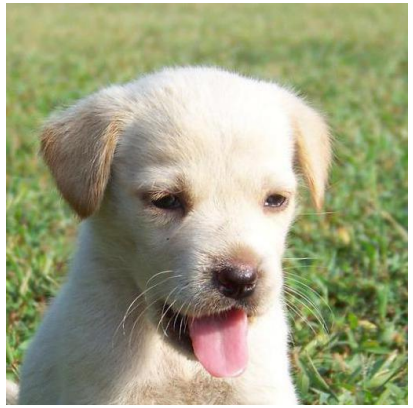
Further improvements

1. Overtrained?



2. More attributes





- More details
- More animals