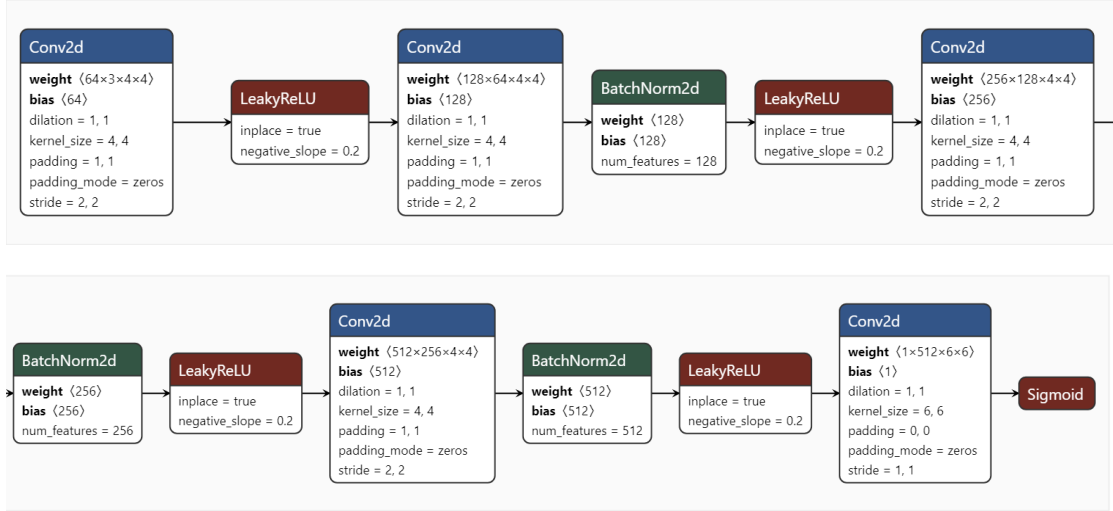


HW03 Report

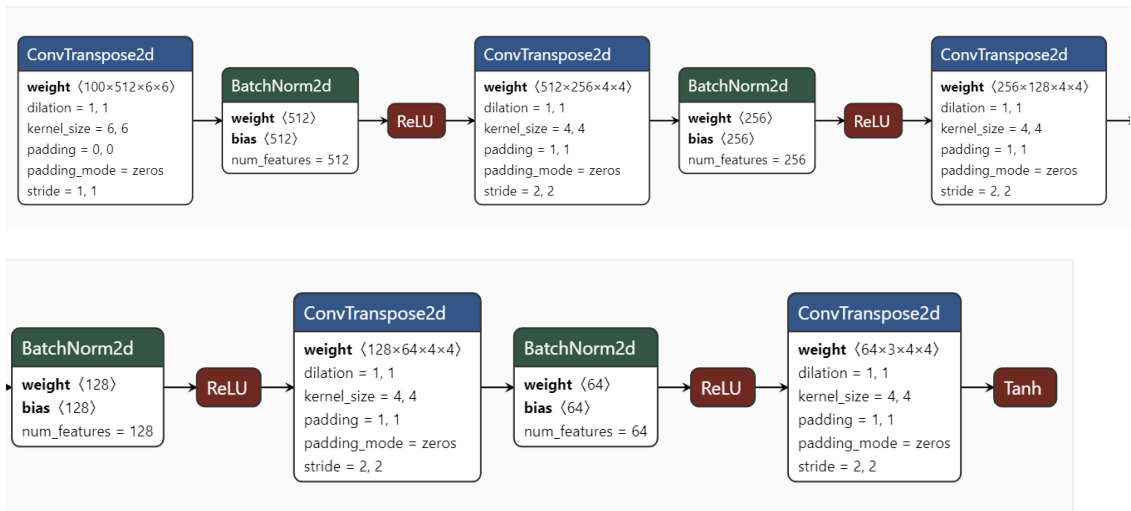
201850050 徐培宾 生命科学学院

1. DCGAN:

1) Discriminator architecture:



2) Generator architecture:



3) Loss function:

交叉熵 BCELoss

$$loss_D = \frac{BCELoss_{data \sim 1} + BCELoss_{g \sim 0}}{2}$$
$$loss_G = BCELoss_{g \sim 1}$$

4) Optimizer:

a) D_optimizer:

- Adam
- learning rate = 0.0002
- betas=(0.5, 0.999)

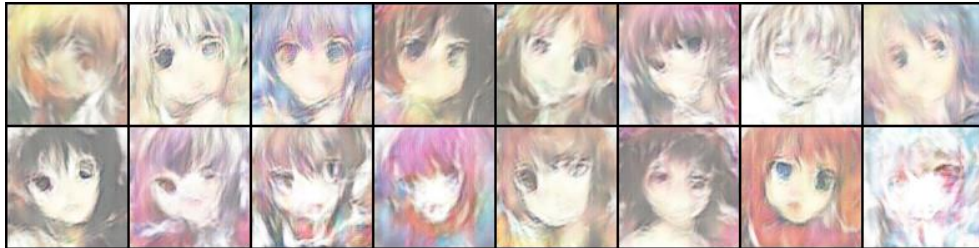
b) G_optimizer:

- Adam
- learning rate = 0.0002
- betas=(0.5, 0.999)

5) Epoch:

Total epoch = 100

6) Result:



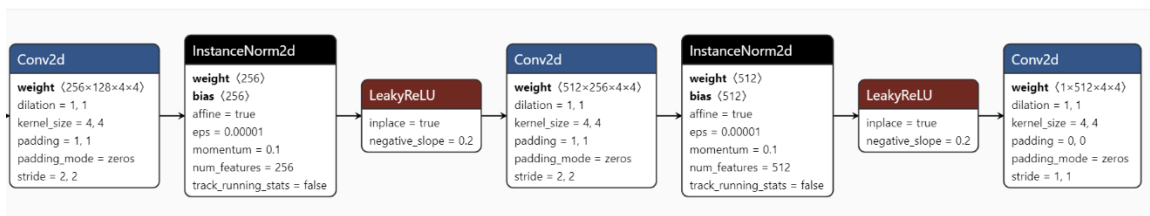
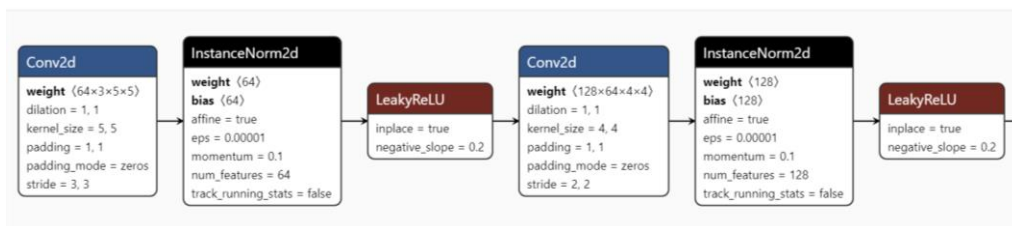
Epoch = 50



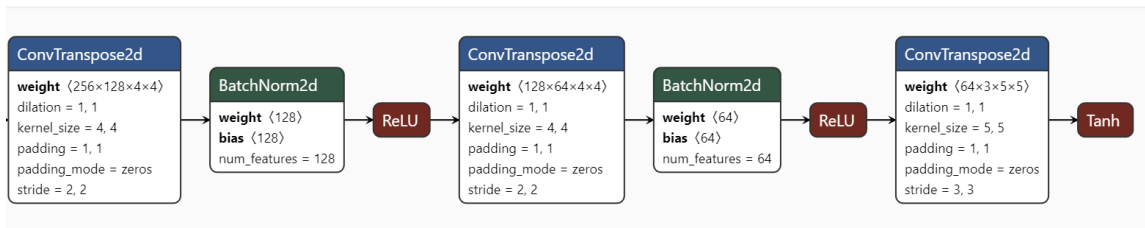
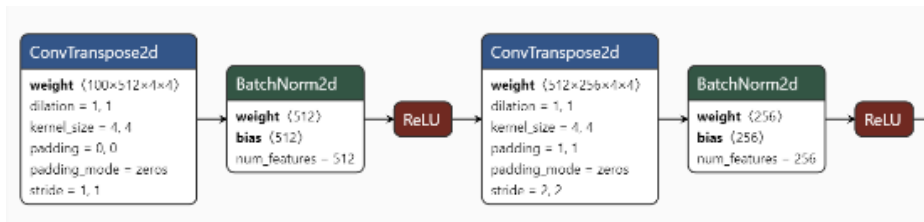
Epoch = 100

2. WGAN-GP:

1) Discriminator architecture:



2) Generator architecture:



3) Loss function:

$$loss_D = loss_{data} - loss_g + \lambda * gradient\ penalty$$

$$loss_g = -mean(D(g))$$

其中:

$$loss_{data} = mean(D(data))$$

$$loss_g = mean(D(g))$$

4) Optimizer:

c) D_optimizer:

- Adam
- learning rate = 0.0002
- betas=(0.5, 0.999)

d) G_optimizer:

- Adam
- learning rate = 0.0002
- betas=(0.5, 0.999)

5) Epoch:

Total epoch = 100

6) Result:



Epoch = 50



Epoch = 100

3. 讨论：

DCGAN model 与 WGAN-GP model 两者在 Generator 和 Discriminator 网络上几乎没有差别，区别在于损失函数，DCGAN 使用的是交叉熵 BCE loss，WGAN-GP 使用的是推土机距离 Wasserstein loss，即目标值与预测值乘积的均值，因此 WGAN-GP 的 Discriminator 末端不需要激活函数 Sigmoid，直接输出全连接网络的值。

另外，对于 WGAN-GP 而言还需要进行权值的 clamp，防止网络更新太快。

训练结果而言，DCGAN 容易出现梯度消失的问题，也可能产生 Mode collapse 的问题，前者会使训练无法进行，后者会使 Generator 产生重复的图片，而 WGAN-GP 在加入了 Gradient penalty 后，一定程度上解决了这两种问题，但是 GAN 网络难以训练的问题仍然存在。