

Homework #12

- 1. Mode collapsing이 무엇인지 설명하고 왜 발생하는 것인가? (서술형)
- 2. Mode collapsing 현상을 극복하기 위해서 방안 3가지가 무엇인지 설명하시오.
- 3. Cycle GAN의 Loss를 제시하고 이에 대하여 설명하시오.

1. GAN의 optimal function으로  $\min_{\theta_g} \max_{\theta_d} [E_{z \sim p_{data}} \log D_{\theta_d}(x) + E_{z \sim p_z} \log (1 - D_{\theta_d}(G_{\theta_g}(z)))]$ 을 사용한다.  
즉, Generator는 Discriminator를 속이는 것을 목표로 하며 D의 출력을 최대로 하는 방식으로 학습이 이뤄진다.  
이때 Generator가 Discriminator를 속이기 위해 distribution의 특정 mode만을 학습하고 비롯한 output만을 출력하는 문제가 발생할 수 있는데, 이런 문제를 Mode collapse라고 한다.

2. ① feature matching: generator의 출력값들의 feature 평균과 실제 데이터들의 feature 평균의 차이의 절대값을 최소화하여 generator의 출력값의 특징과 실제 데이터의 특징이 평균적으로 같아지도록 함.  
$$\min_{\theta_g} \| E_{z \sim p_{data}} f(x) - E_{z \sim p_z} f(G(z)) \|^2$$
  
② historical averaging: 학습데이터의 mode가 바뀌었을 때  $\theta_g$ 가 급격하게 변하는 것을 방지함  
$$\| \theta - \frac{1}{t} \sum_{i=1}^t \theta_i \|^2$$
  
③ minibatch discrimination: discriminator의 실제 데이터 또는 generator에 의해 생성된 데이터를 입력할 때, 그 데이터가 속한 batch의 similarity score를 계산하여 함께 입력함  
mode collapsing이 발생하면, batch의 score가 실제 데이터 batch보다 높아지게 되므로 discriminator가 거짓데이터로 판단할 확률이 높아짐!  
style 외의 다른 content

3. Cycle GAN은 두 개의 generator를 사용하는 구조이다.  
CycleGAN의 total loss는 discriminator loss + reconstruction loss이다.  
첫번째 generator로 데이터의 source style을 target style에 맞춰 변형한 후 discriminator에 입력하면 이 변형된 데이터에 대한 discriminator loss를 구할 수 있다.  
두번째 generator는 첫번째 generator에 의해 변형된 데이터를 다시 원본 데이터로 reconstruction 한다.  
원본데이터와 두번째 generator를 거친 변형된 데이터의 차를 제곱한 값이 reconstruction loss인데, 이를 최소화하기 위해서는 첫번째 generator가 style을 제외한 contents의 변형을 최소화해야한다.