

##GAN 모델 문제점

1.모델 훈련의 불안정성 (instability)

- 수렴하지 못하는 경우가 생김
- 생성기와 판별기가 고루 학습이 되어야하는데, 편향적으로 되어버릴 수 있음

2.성능 평가의 어려움

- 사람의 눈으로 파악할 수 밖에 없음
- overfitting을 판단할 수학적 근거 부족

3.Mode collapsing 문제

- 훈련된 multi-modal 데이터의 일부 mode가 누락된 문제
- Ex) 0~9의 숫자로 구성된 데이터셋(MNIST)로 학습한 경우에 특정 숫자(5)를 생성해내지 못함
- 강아지의 제한된 색상과 특징을 생성하는 법을 학습할 수도 있음

3.발생하는 이유?

- 생성기가 훈련 데이터의 분포를 충분히 모델링하지 못했다는 것

<GAN 평가지표>

① IS (Inception Score)

$$IS(p_g) = \mathbb{E}_{x \sim p_g} [KL(P(y|x) || P(y))]$$

- conditional label distribution $p(y|x)$: 생성된 image의 품질 측정.
 - 생성된 image x 를 inception 모델을 통해 나온 각 label에 대한 확률분포
 - 가장 이상적인 conditional label distribution = 특정 label이 높게 나온 분포
- Marginal distribution $p(y)$: 생성된 image의 다양성을 측정하는 부분
 - 여러개의 생성된 이미지들을 사용하여 label의 분포를 합산한 새로운 분포
 - 가장 이상적인 marginal distribution == 균등한 분포.
- Kullback-Leibler divergence: $p(y|x)$ 과 $p(y)$ 가 얼마나 유사하거나 다른지에 대한 척도

② FID (Fréchet Inception Distance)

- Inception Network 사용.
 - pretrained된 inception V3에서 출력 레이어를 제외한 출력 마지막 pooling layer의 Activation을 사용.
- Inception Score는 생성된 이미지만 사용하여 성능을 평가하는 반면, FID는 대량 도메인의 실제 이미지 모음 통계와 생성된 이미지 모음 통계를 비교하여 평가를 진행함.

$$d^2((m, C), (m_w, C_w)) = \|m - m_w\|_2^2 + \text{Tr}(C + C_w - 2(CC_w)^{1/2})$$

m : 실제 데이터의 특징 평균

C : 실제 데이터의 특징 공분산 행렬

m_w : 생성한 데이터의 특징 평균

C_w : 생성한 데이터의 공분산 행렬

③ CID index.

$$CID = \text{Creativity} \times \text{Inheritance} \times \text{Diversity}$$

$R = \{I_r\}$: 실제 image들의 집합

$G = \{I_g\}$: GAN으로 생성된 image들의 집합

$G^{\text{rem}} = \{I_g \mid \text{SSIM}(I_g, I_r) < 0.8\}$: G에서 실제 image와 유사도가 80%보다 낮은 image들의 집합

g_r : GLCM-contrast를 적용한 R 의 평균

g_g : GLCM-contrast를 적용한 G^{rem} 의 평균

C_1, \dots, C_m : 유사한 image들로 구성된 각 cluster들의 entropy.