

Paper summary

AI VISION Lab

1. 공부한 논문의 제목, 게재된 학회 혹은 저널 등 논문 기본 정보를 적으세요.
 - A. 이름: **Classifier-Free Diffusion Guidance**
 - B. 저널: **NeurIPS**
 - C. 도메인: **Diffusion**
 - D. 출판연도: **2022**
 - E. 저자: **Jonathan Ho**
2. 논문에서 제안한 알고리즘 및 프레임워크에 대해 본인이 이해한대로 다이어그램을 그려보세요. 논문 **Figure**를 그대로 따라 그리면 안됩니다.
 - A. [선행연구] Classifier Guidance의 한계: Classifier의 학습이 필요함(이 때, Pretrained된 Classifier를 사용할 수 없음 - x_t , Noise 가 더해진 상태에서 $\text{Class}(y)$ 를 구해야 하는 구조라서 추가적인 학습이 필요함.) 또한, Classifier의 Gradient를 추가로 더해주는 구조임(이를 Adversarial Attack의 구조로 판단 -> GAN을 그대로 가져왔기 때문에 성능이 좋게 나오는 것이 아닌가?)
 - B. 따라서, 본 논문에서는 Classifier가 없는 상황에서도, Fidelity-Diversity간의 Trade-off가 가능한 구조를 제안하고자 함.
 - C. Classifier Guidance는 Unconditional: $\epsilon_\theta(z_t)$ 혹은 Conditional: $\epsilon_\theta(z_t, c)$ 의 Score Estimation에 Gradient of Classifier를 빼는 구조로 구성됨.
 - D. Classifier Free Guidance의 경우는, Classifier를 추가적으로 학습할 필요 없는 구조를 제안함. $\epsilon_\theta(z_t, c) = (1 + w) * \epsilon_\theta(z_\lambda, c) - w * \epsilon_\theta(z_\lambda)$ 이때, 두 모델에 대해서, 따로 학습하는 구조가 아닌, 한 번에 학습을 진행하는 구조인데, 이는 Unconditional한 구조를 $\epsilon_\theta(z_\lambda, c = \emptyset)$ 로 볼 수 있다는 점을 반영하여, Random(Hyperparameter) 하게 Condition을 주거나 주지 않음을 결정하는 식으로 해결함.
 - E. 이때, 수식을 전개하면, Conditional - $W * (\text{Unconditional} - \text{Conditional})$ 로 작성할 수 있는데, 이때 문제점은 Non-Conservative한 구조로, Closed Form형태의 수식적 설명이 불가능하다는 점임.

- F. 그러나 해당 논문에서는 제안한 구조를 기반으로 실험을 진행했고, 결과적으로 Classifier Guidance와 비슷한 성능을 보임.
3. 본인이 생각하는 이 논문의 장점이 무엇이라고 생각하나요? **논문 Contribution bullet을 그대로 따라 적으면 안됩니다.**
- A. Classifier의 추가적인 학습이 필요하다는 점을 새로운 구조를 제안함으로써 해결했다는 점
4. 이 논문을 읽으면서 느낀 점, 혹은 배운 점이 있으면 적어보세요.
- A. 비판적 읽기의 필요성. Classifier Guidance 논문을 너무 신뢰하며 읽었던 것이 아닌가 싶은 반성
5. 이 논문의 한계점이 있다면 무엇이라고 생각하나요?
- A. Class label이 없는 경우 사용불가능 함.
- B. 수치상 Trade-off를 입증했으며, 성능향상을 보였으나, 수식적인 증명이 부족함. 즉, Closed-form 계산이 없다는 점에서 다른 연구에 적용했을 경우에 성능이 잘 나올지에 대한 의문이 존재함.
6. 본인의 연구에 접목시켜볼 점이 있을지 생각하고 적어보세요.
- A. Diffusion 선행연구
7. 본 Summary를 작성하는 과정에서 생성형AI를 사용했나요?
- A. 아니요

날짜: 2025-07-11

이름: 신준원

Classifier Free Diffusion Guidance

01/11

Introduction

low temperature sampling

temperature \rightarrow entropy $\rightarrow t \rightarrow 0$ sharpen distribution \Rightarrow deterministic
 $t \rightarrow \infty$ flatten distribution \Rightarrow randomness

Classifier Guidance

GAN (BigGAN) - truncation trick
low temperature sampling $\sim \mathcal{N}(0, T^2 I)$

\downarrow \rightarrow Variance \uparrow , noise \uparrow \rightarrow \downarrow

Diffusion이 \rightarrow image quality \downarrow (blur)

\downarrow Diffusion에 (diversity, fidelity) \rightarrow \rightarrow \rightarrow Idea = Classifier Guidance

Classifier Guidance의 개념

- ① Classifier의 학습 필요 (이때, Classifier: Res to Inception...)
- \rightarrow noisy data \rightarrow pre-trained 사용 불가
 - \rightarrow noisy input \rightarrow Classifier 필요
 - \rightarrow $P(y|t)$, Guidance는 $\nabla_{t_t} P(y|t_t)$

- ② Adversarial attack \rightarrow Classifier Guidance $\nabla_{t_t} P(y|t_t)$

Score of diffusion + Gradient of Classifier
 \rightarrow Gradient를 주자 \Rightarrow Classifier를 \rightarrow \rightarrow \rightarrow GAN과 유사한 것.

\Rightarrow Q) GAN과 동일한 것 아니냐?
(아니)

① $(z_x) = x_{t_0}$ \triangleright paper notation

Classifier free Guidance

① Classifier $\frac{1}{\sigma^2} x$

② Adversarial attack

\rightarrow Classifier free Guidance

$$\tilde{\xi}_{\theta}(z_x, c) = (1+w) \underbrace{\xi_{\theta}(z_x, c)}_{\text{Conditional Score Estimates}} - w \underbrace{\xi_{\theta}(z_x)}_{\text{Unconditional score estimate}}$$

Unconditional score estimate

Conditional Score Estimates

Classifier Guidance of DDIM

① ~~diffusion~~ diffusion Score

$$\Rightarrow \xi_{\theta}(z_x, c) \approx -\sigma_x \nabla_{z_x} \log p(z_x | c)$$

\rightarrow Conditional Diffusion model

$$\textcircled{2} \Rightarrow \tilde{\xi}_{\theta}(z_x, c) = \xi_{\theta}(z_x, c) - w \sigma_x \nabla_{z_x} \log p(c | z_x)$$

$$\approx -\sigma_x \nabla_{z_x} [\log p(z_x | c) + w \log p(c | z_x)]$$

Conditional Diffusion model's Score

Classifier Guidance Control $2H$

\rightarrow Gradient of Classifier

③ $\tilde{\xi}_{\theta}$ of DDIM \Rightarrow Sampling by $\frac{1}{\sigma^2}$

$$\tilde{p}_{\theta}(z_x | c) \propto p_{\theta}(z_x | c) p_{\theta}(c | z_x)^w$$

Classifier free Guidance

문제: Classifier Guidance 식. (+) Classifier Guidance를 통한 성능 향상.

→ $p_\theta(z)$: Unconditional diffusion model

→ $p_\theta(z|c)$: Conditional diffusion model

→ 두 가지를 동시에 학습 ($p_\theta(z) \Rightarrow \epsilon_\theta(z), p_\theta(z|c) \Rightarrow \epsilon_\theta(z, c)$)

→ $\epsilon_\theta(z, c = \emptyset)$, $\frac{\gamma}{1+\gamma}$ 정도 Conditional Guidance 관련

Unconditional → Condition = \emptyset

→ hyper-parameter

→ random prob (Poncond)에 따라.

→ \emptyset or c return

즉, Classifier를 추가하지 않고도 비슷한 성능 가능.

$$\tilde{\epsilon}_\theta(z, c) = (1+\gamma)\epsilon_\theta(z, c) - \gamma\epsilon_\theta(z)$$

Classifier-free Guidance의 장점

① Gradient of classifier 식 \Rightarrow Adversarial attack X

② Classifier가 여러를 학습 (추가학습) 필요 X.

Classifier free Guidance의 원리

① non-conservative vector field 기반.

→ Classifier Guidance \Rightarrow Classifier의 Score \propto conservative \Rightarrow scalar potential

→ 원리인 이유: 학습 과정 \rightarrow 기대 \rightarrow 실제 \rightarrow 결과

본문 주장: \rightarrow Gradient of Implicit classifier 기반의 것.

$$\rightarrow p'(c|x) \propto \frac{p(x|c)}{p(x)}$$

Scores: \downarrow $\epsilon^*(x, c)$ \downarrow $\epsilon^*(x)$

$$\nabla_x \log p'(c|x) = -\frac{1}{\sigma^2} [\epsilon^*(x, c) - \epsilon^*(x)]$$

$$\epsilon^*(x, c) = \frac{(1+w)\epsilon^*(x, c) - w\epsilon^*(x)}{\text{Classifier Guidance}}$$

그런데, classifier-free Guidance의 경우와는 다름. (분명!!)

\Rightarrow 실제 기대는 성능 보충 (trade-off)