Direct Discriminative Optimization

Your Likelihood-Based Visual Generative Model is Secretly a GAN Discriminator

Kaiwen Zheng[1,2], Yongxin Chen[1], Huayu Chen[2], Guande He[3], Ming-Yu Liu[1], Jun Zhu[2], Qinsheng Zhang[1]

[1] NVIDIA
[2] Tsinghua University
[3] The University of Texas at Austin

ICML 2025 Spotlight

Issue of Maximum Likelihood Estimation

Diffusion Probabilistic Models (DPMs) and Autoregressive Models (ARs) based on **Maximum Likelihood Estimation (MLE)** have achieved dominance in both continuous and discrete data generation. Its

are better than Generative Adversarial Networks (GANs).

Issue of Maximum Likelihood Estimation

Diffusion Probabilistic Models (DPMs) and Autoregressive Models (ARs) based on **Maximum Likelihood Estimation (MLE)** have achieved dominance in both continuous and discrete data generation. Its **Stability**,

are better than Generative Adversarial Networks (GANs).

Issue of Maximum Likelihood Estimation

Diffusion Probabilistic Models (DPMs) and Autoregressive Models (ARs) based on **Maximum Likelihood Estimation (MLE)** have achieved dominance in both continuous and discrete data generation. Its **Stability**,

Scalability and

are better than Generative Adversarial Networks (GANs).

Issue of Maximum Likelihood Estimation

Diffusion Probabilistic Models (DPMs) and Autoregressive Models (ARs) based on **Maximum Likelihood Estimation (MLE)** have achieved dominance in both continuous and discrete data generation. Its **Stability**,

Scalability and

Generalization are better than Generative Adversarial Networks (GANs).

Issue of Maximum Likelihood Estimation

However, when the model capacity is limited, the training method of MLE will cause the generation of blur in order to cover a wider range of samples,

Issue of Maximum Likelihood Estimation

However, when the model capacity is limited, the training method of MLE will cause the generation of blur in order to cover a wider range of samples, also known as **"MODE-COVERING."**

Issue of Maximum Likelihood Estimation

However, when the model capacity is limited, the training method of MLE will cause the generation of blur in order to cover a wider range of samples, also known as **"MODE-COVERING."**

In order to solve this problem, introducing the adversarial loss of GANs is one of the solutions, but

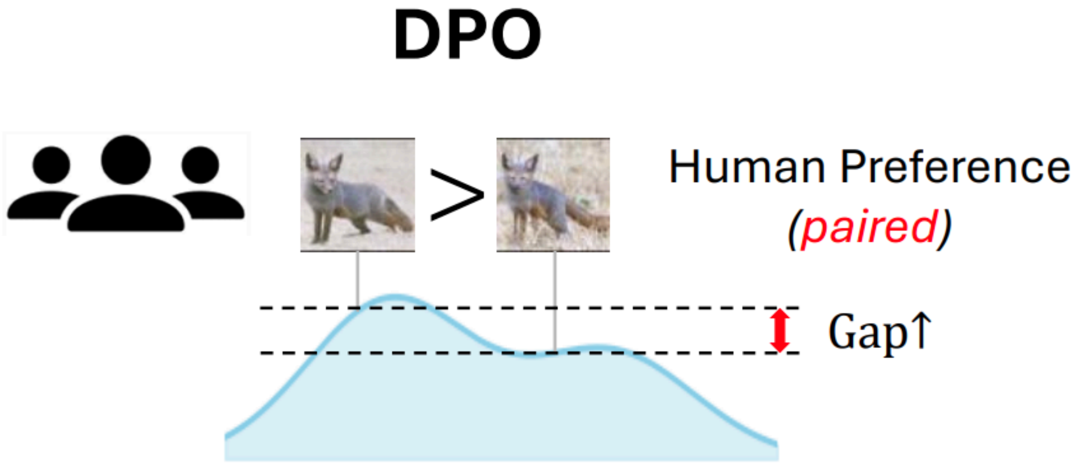
Issue of Maximum Likelihood Estimation

However, when the model capacity is limited, the training method of MLE will cause the generation of blur in order to cover a wider range of samples, also known as **"MODE-COVERING."**

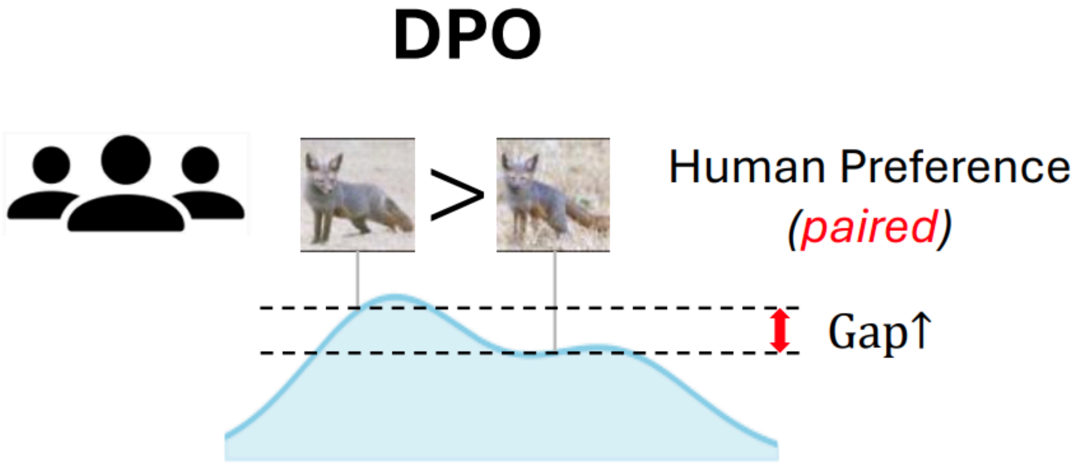
In order to solve this problem, introducing the adversarial loss of GANs is one of the solutions, but

- GANs need to construct additional discriminators for alternating optimization, which increases engineering complexity.
- The iterative sampling of DPMs and ARs makes it difficult to easily optimize through GANs.

Direct Preference Optimization (DPO)

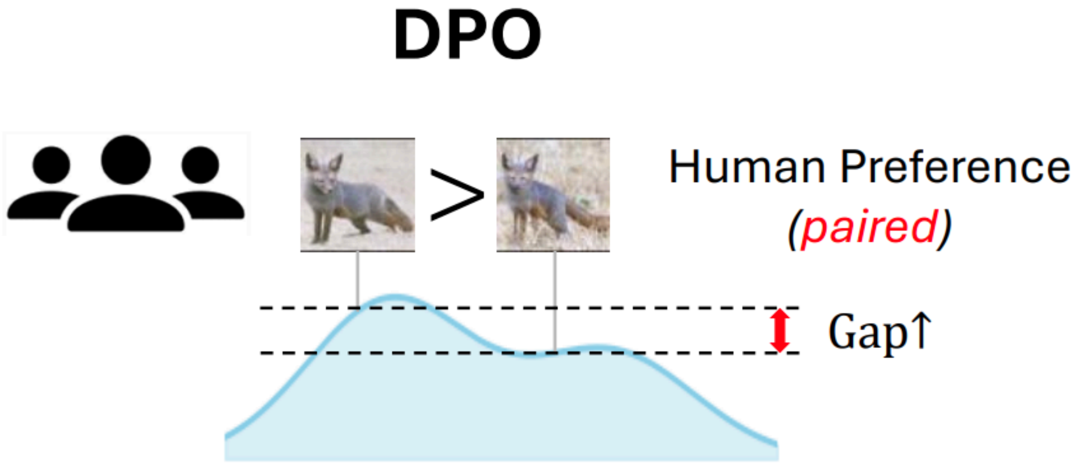


Direct Preference Optimization (DPO)



- Paired Data

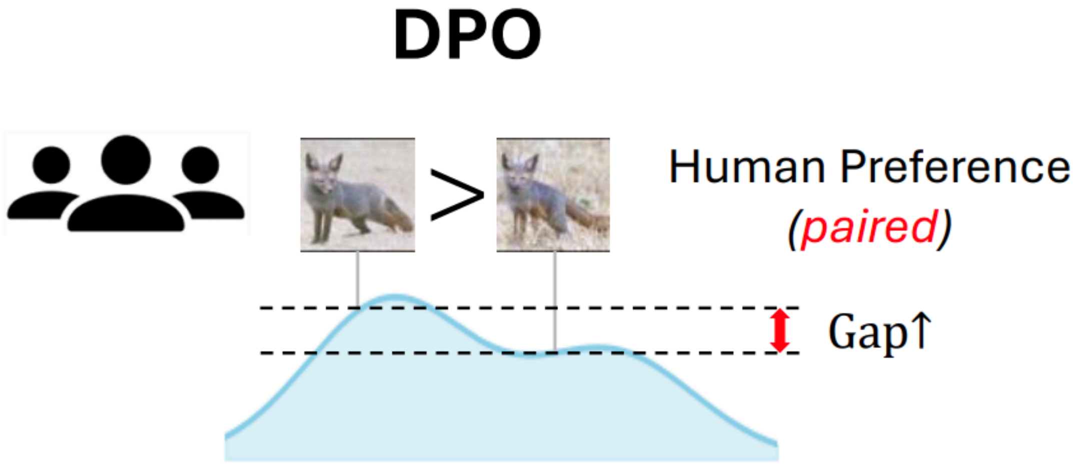
Direct Preference Optimization (DPO)



- Paired Data
- Human-Annotated

preferred/non-preferred

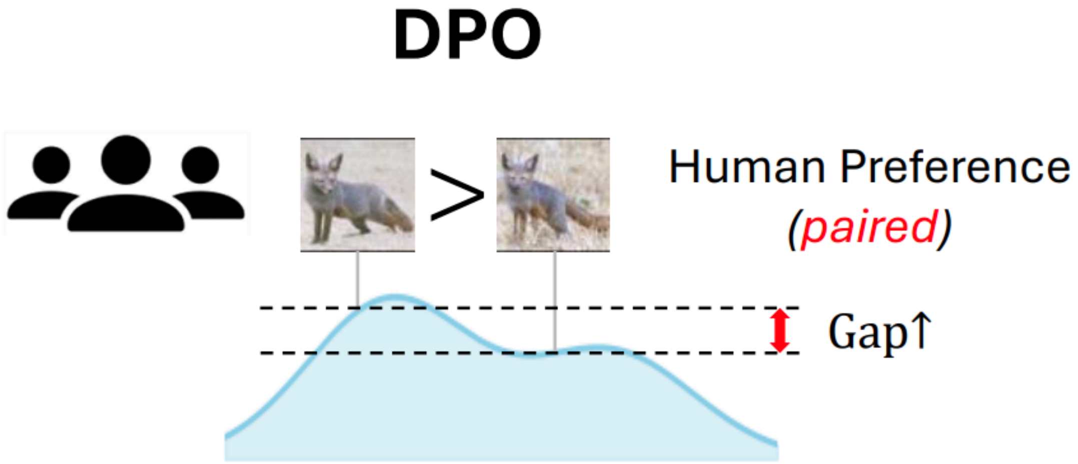
Direct Preference Optimization (DPO)



- Paired Data
- Human-Annotated

  preferred/non-preferred
- Supervised Learning

Direct Preference Optimization (DPO)



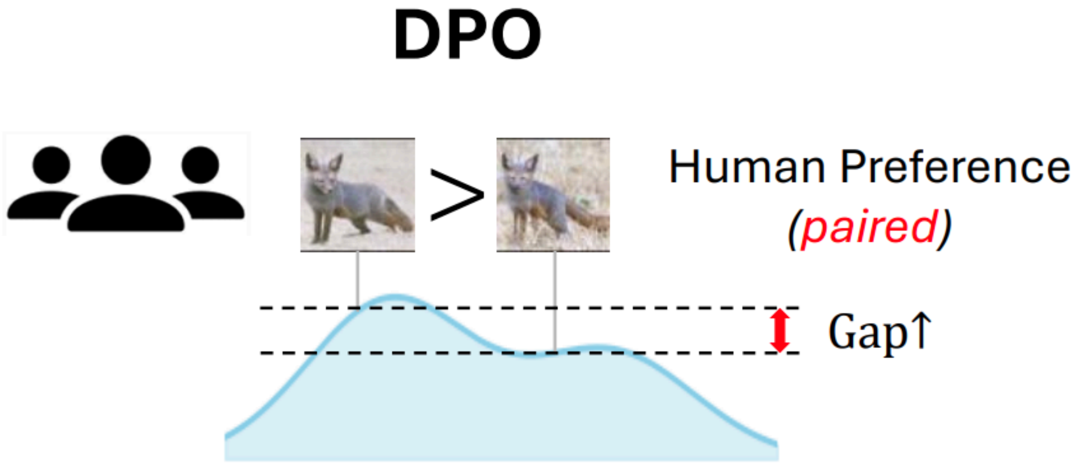
- Paired Data
- Human-Annotated

  preferred/non-preferred
- Supervised Learning

Optimizing the reward function

$$p(y_w \succ y_l|x) := \frac{e^{r(x,y_w)}}{e^{r(x,y_l)} + e^{r(x,y_w)}} = \sigma(r(x,y_w) - r(x,y_l))$$

Direct Preference Optimization (DPO)



- Paired Data
- Human-Annotated

  preferred/non-preferred
- Supervised Learning

Your Likelihood-Based Generative Model is Secretly a Discriminator

$$\mathcal{L}^{\text{DPO}}(\theta) = -\mathbb{E}_{(x,y_w,y_l)\sim\mathcal{D}} \log\sigma\left(\beta\log\frac{\pi_\theta(y_w|x)}{\pi_{\theta_{\text{ref}}}(y_w|x)} - \beta\log\frac{\pi_\theta(y_l|x)}{\pi_{\theta_{\text{ref}}}(y_l|x)}\right)$$

Can we use **Secret Discriminator** to enhance MLE-based Models?

Can we use **Secret Discriminator** to enhance MLE-based Models?

Direct Preference Optimization

- Paired Data
- Preferred / Non-preferred

Can we use **Secret Discriminator** to enhance MLE-based Models?

Direct Preference Optimization	>>	**Direct Discriminative Optimization**
▪ Paired Data		▪ Real Data / Generated Data
▪ Preferred / Non-preferred		▪ Truth / Fake

Direct Discriminative Optimization

Direct Discriminative Optimization

Direct Discriminative Optimization

Direct Discriminative Optimization

Practical Implementation

Generalized Objective with Extra Coeffecients

To address gradient vanishing and numerical precision issues.

$$\begin{aligned}\mathcal{L}_{\alpha,\beta}(\theta) = &-\mathbb{E}_{p_{\mathrm{data}}(\mathrm{x})}\left[\log\sigma\left(\beta\log\frac{p_\theta(\mathrm{x})}{p_{\theta_{\mathrm{ref}}}(\mathrm{x})}\right)\right]\\ &-\alpha\mathbb{E}_{p_{\theta_{\mathrm{ref}}(\mathrm{x})}}\left[\log\left(1-\sigma\left(\beta\log\frac{p_\theta(\mathrm{x})}{p_{\theta_{\mathrm{ref}}}(\mathrm{x})}\right)\right)\right],\ \alpha\in[0.5,50]\text{ and }\beta\in[0.01,0.1]\end{aligned}$$

Direct Discriminative Optimization

Practical Implementation

Handling Compute-Intensive Likelihood

$$\log \frac{p_\theta(\mathbf{x})}{p_{\theta_{\text{ref}}}(\mathbf{x})} \approx \mathbb{E}_{t,\epsilon}\left[\Delta_{\mathbf{x}_t,t,\epsilon}\right]$$

where $\mathbf{x}_t = \alpha_t \mathbf{x} + \sigma_t \epsilon,\ \ \epsilon \sim \mathcal{N}(0, I)$ and
$\Delta_{\mathbf{x}_t,t,\epsilon} = -w(t)\left(\|\epsilon_\theta(\mathbf{x}_t, t) - \epsilon\|_2^2 - \|\epsilon_{\theta_{\text{ref}}}(\mathbf{x}_t, t) - \epsilon\|_2^2\right)$

Direct Discriminative Optimization

Practical Implementation

Handling Compute-Intensive Likelihood

$$\log \frac{p_\theta(\mathrm{x})}{p_{\theta_{\text{ref}}}(\mathrm{x})} \approx \mathbb{E}_{t,\epsilon}\left[\Delta_{\mathrm{x}_t,t,\epsilon}\right]$$

where $\mathrm{x}_t = \alpha_t \mathrm{x} + \sigma_t \epsilon,\ \ \epsilon \sim \mathcal{N}(0, I)$ and

$$\Delta_{\mathrm{x}_t,t,\epsilon} = -w(t)\left(\|\epsilon_\theta(\mathrm{x}_t, t) - \epsilon\|_2^2 - \|\epsilon_{\theta_{\text{ref}}}(\mathrm{x}_t, t) - \epsilon\|_2^2\right)$$

Apply Jensen's inequalit: $\psi(\mathbb{E}[X]) \leq \mathbb{E}[\psi(X)]$, if ψ is a convex function e.g. $-\log \sigma(\cdot)$.

Direct Discriminative Optimization

Practical Implementation

Handling Compute-Intensive Likelihood

$$\log \frac{p_\theta(\mathrm{x})}{p_{\theta_{\mathrm{ref}}}(\mathrm{x})} \approx \mathbb{E}_{t,\epsilon}\left[\Delta_{\mathrm{x}_t,t,\epsilon}\right]$$

where $\mathrm{x}_t = \alpha_t \mathrm{x} + \sigma_t \epsilon, \ \epsilon \sim \mathcal{N}(0, I)$ and

$$\Delta_{\mathrm{x}_t,t,\epsilon} = -w(t)\left(\|\epsilon_\theta(\mathrm{x}_t, t) - \epsilon\|_2^2 - \|\epsilon_{\theta_{\mathrm{ref}}}(\mathrm{x}_t, t) - \epsilon\|_2^2\right)$$

Apply Jensen's inequalit: $\psi(\mathbb{E}[X]) \le \mathbb{E}[\psi(X)]$, if ψ is a convex function e.g. $-\log \sigma(\cdot)$.

$$\begin{aligned}
&\mathcal{L}(\theta) \\
&\approx -\mathbb{E}_{p_{\mathrm{data}}(\mathrm{x})} \log \sigma\left(\mathbb{E}_{t,\epsilon}[\Delta]\right) - \mathbb{E}_{p_{\theta\mathrm{ref}}(\mathrm{x})} \log\left(1 - \sigma\left(\mathbb{E}_{t,\epsilon}[\Delta]\right)\right) \\
&\le \underbrace{-\mathbb{E}_{t,\epsilon}\left[\mathbb{E}_{p_{\mathrm{data}}(\mathrm{x})} \log \sigma(\Delta) + \mathbb{E}_{p_{\theta_{\mathrm{ref}}}(\mathrm{x})} \log(1 - \sigma(\Delta))\right]}_{\text{upper bound}}
\end{aligned}$$

Direct Discriminative Optimization

Amazing Feature: Multi-Round Refinement via Self-Play

$$\text{Round } n: \quad \ldots \quad \rightarrow \quad \underbrace{p_{\theta^*_{n-1}}}_{\text{Reference}} \quad \rightarrow \quad \underbrace{\sigma\left(\beta \log \frac{p_{\theta_n}}{p_{\theta^*_{n-1}}}\right)}_{\text{Discriminator}}$$

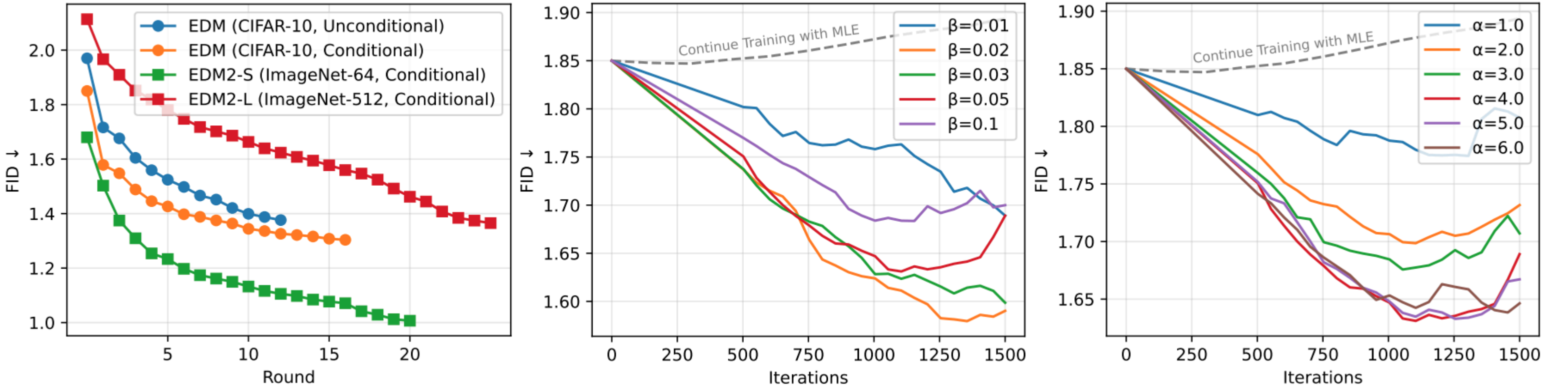
$$\text{Round } n+1: \quad \rightarrow \quad \underbrace{p_{\theta^*_{n}}}_{\text{Reference}} \quad \rightarrow \quad \ldots$$

Direct Discriminative Optimization

Amazing Feature: Multi-Round Refinement via Self-Play

$$\text{Round } n: \quad \dots \to \underbrace{p_{\theta^*_{n-1}}}_{\text{Reference}} \to \underbrace{\sigma\left(\beta \log \frac{p_{\theta_n}}{p_{\theta^*_{n-1}}}\right)}_{\text{Discriminator}}$$

$$\text{Round } n+1: \quad \to \underbrace{p_{\theta^*_{n}}}_{\text{Reference}} \to \dots$$



(a) FID-Round

(b) $\alpha = 4.0$

(c) $\beta = 0.05$

Experiment

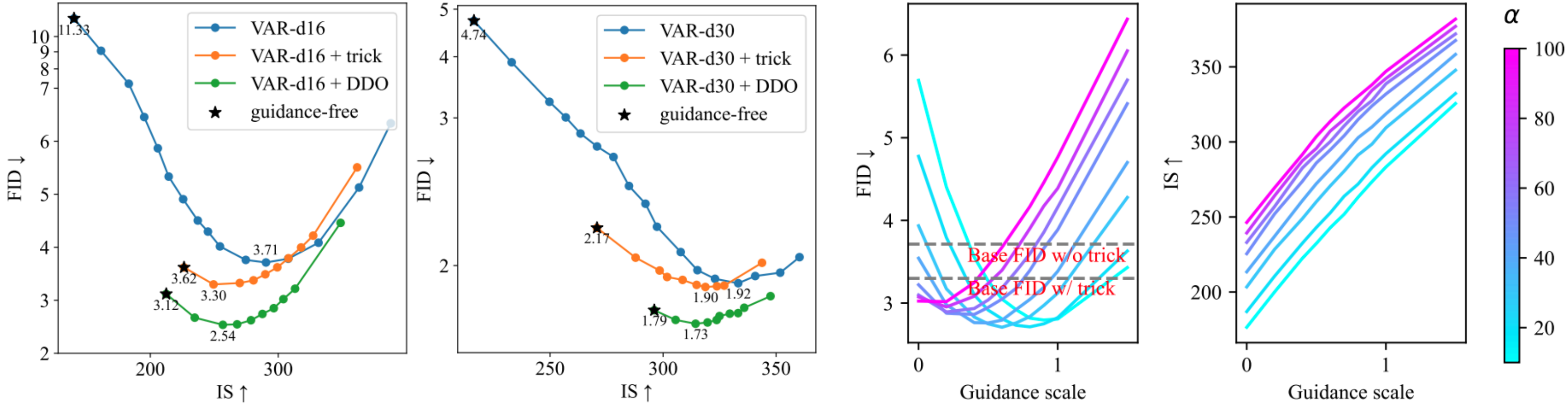
ImageNet 64x64

Type	Model	#Params	NFE	FID↓
GAN	StyleGAN-XL (Sauer et al., 2022)	135M	1	1.51
	R3GAN (Huang et al., 2025)	104M	1	2.09
	CTM (Kim et al., 2023b)	324M	1	1.92
	DMD2 (Yin et al., 2024)	296M	1	1.28
	PaGoDA (Kim et al., 2024)	296M	1	1.21
	SiD^{2}A (Zhou et al., 2024b)	296M	1	1.11
Diffusion	iDDPM (Nichol & Dhariwal, 2021)	270M	250	2.92
	ADM (Dhariwal & Nichol, 2021)	296M	250	2.07
	RIN (Jabri et al., 2022)	281M	1000	1.23
	EDM (Karras et al., 2022)	296M	511	1.36
	VDM++ (Kingma & Gao, 2024)	296M	511	1.43
	DisCo-Diff (Xu et al., 2024)	-	623	1.22
	EDM2-S (Karras et al., 2024b)	280M	63	1.58
	+ CFG (Ho & Salimans, 2021)	560M	126	1.48
	+ AG (Karras et al., 2024a)	405M	126	1.01
	EDM2-M	498M	63	1.43
	EDM2-L	777M	63	1.33
	EDM2-XL	1.1B	63	1.33
Ours	EDM2-S (retested)	280M	63	1.60
	+ DDO	280M	63	**0.97**

ImageNet 512x512

Type	Model	w/o G			w/ G
		#Params	NFE	FID↓	FID↓
GAN	BigGAN-deep (Brock, 2018)	112M	1	8.43	-
	StyleGAN-XL (Sauer et al., 2022)	168M	1	2.41	-
	SiD^{2}A (Zhou et al., 2024b)	1.5B	1	1.37	-
Diffusion	ADM (Dhariwal & Nichol, 2021)	559M	250	23.24	7.72†
	ADM-U	731M	500	9.96	3.85†
	DiT-XL/2 (Peebles & Xie, 2023)	675M	250	12.03	3.04
	SiT-XL (Ma et al., 2024)	675M	250	-	2.62
	+ REPA (Yu et al., 2024b)	675M	250	-	2.08
	RIN (Jabri et al., 2022)	410M	1000	3.95	-
	U-ViT, L (Hoogeboom et al., 2023)	2B	512	3.54	3.02
	VDM++ (Kingma & Gao, 2024)	2B	512	2.99	2.65
	USiT-2B (Chen et al., 2024b)	2B	-	2.90	1.72
	EDM2-XS (Karras et al., 2024b)	125M	63	3.53	2.91
	EDM2-S	280M	63	2.56	2.23
	+ AG				1.34‡
	EDM2-L	777M	63	2.06	1.88
	EDM2-XXL	1.5B	63	1.91	1.81
	+ AG				1.25‡
Masked	MaskGIT (Chang et al., 2022)	227M	12	7.32	-
	MAGVIT-v2 (Yu et al., 2023)	307M	64	3.07	1.91
	MAR-L (Li et al., 2024)	481M	1024	2.74	1.73
AR	VAR-d36-s (Tian et al., 2024)	2.3B	10	-	2.63
Ours	EDM2-L (retested)	777M	63	1.96	1.77
	+ DDO	777M	63	**1.26**	**1.21**‡
	+ DPM-Solver-v3 (Zheng et al., 2023a)	777M	25	1.29	**1.21**‡

Experiment



(a) VAR-d16 (b) VAR-d30 (c) FID and IS w.r.t. guidance scale w

Figure 7. Illustrations of DDO on autoregressive models. (a)(b) FID-IS trade-off curves and (c) the impact of α under $\beta = 0.02$.

Conclusions

Advantages

- Hight efficiency and effectiveness.
- No additional redundant structure required.

Weaknesses

- The need for hyperparameter searching.
- It takes time to iteratively sample fake samples as training data.

Conclusions

Advantages

- Hight efficiency and effectiveness.
- No additional redundant structure required.

Weaknesses

- The need for hyperparameter searching.
- It takes time to iteratively sample fake samples as training data.
  Maybe it can be combined with a single-step generation method, e.g. Consistency Models [Song et al., 2023; Song et al., 2023; Lu et al., 2025]

END