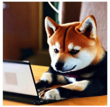
Denoising Diffusion Probabilistic Models

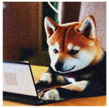
Jonathan Ho Ajay Jain Pieter Abbeel

UC Berkeley

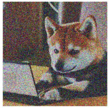
November 9, 2022



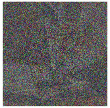
Yujin



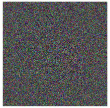
Xinke-Wang



Changfeng-Duan



JieLi



Yuhong-Sun

Content

Outline

What is generative model?

Regardless of precise definition, the terminology is constitutional because a generative model can be used to "generate" random instances

Deep generative models

With the rise of deep learning, a new family of methods, called deep generative models (DGMs)

- Generative adversarial networks (GANs)
- Variational autoencoders (VAEs)
- Flow Based Model

Outline

What is Diffusion Model?

We can say...

In machine learning, diffusion models, also known as diffusion probabilistic models, are a class of latent variable models.

Intuitive understanding

Outline

DDPM Features

Name.	Likelihood	Speed	Methods	Stability	Others
GAN	None	fast	One-step	Unstable	high quality
VAE	Uncertain	fast	One-step	Stable	-
Flow Model	Exactly	fast	Multi-step	Stable	-
DDPM	Uncertain	slow	Multi-step	Stable	beat GANs

Table: Comparision of generative model

Beat GANs

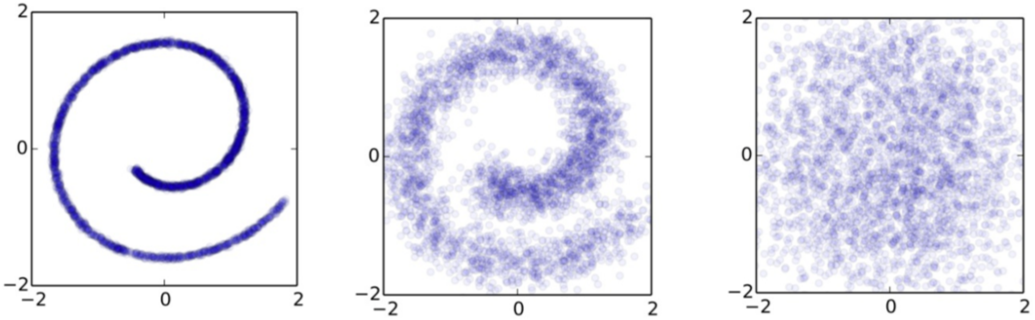
Model	FID	sFID	Prec	Rec
LSUN Bedrooms 256×256				
DCTransformer† [42]	6.40	6.66	0.44	**0.56**
DDPM [25]	4.89	9.07	0.60	0.45
IDDPM [43]	4.24	8.21	0.62	0.46
StyleGAN [27]	2.35	6.62	0.59	0.48
ADM (dropout)	**1.90**	**5.59**	**0.66**	0.51
LSUN Horses 256×256				
StyleGAN2 [28]	3.84	6.46	0.63	0.48
ADM	2.95	**5.94**	0.69	**0.55**
ADM (dropout)	**2.57**	6.81	**0.71**	**0.55**
LSUN Cats 256×256				
DDPM [25]	17.1	12.4	0.53	0.48
StyleGAN2 [28]	7.25	**6.33**	0.58	0.43
ADM (dropout)	**5.57**	6.69	**0.63**	**0.52**
ImageNet 64×64				
BigGAN-deep* [5]	4.06	3.96	**0.79**	0.48
IDDPM [43]	2.92	**3.79**	0.74	0.62
ADM	2.61	**3.77**	0.73	0.63
ADM (dropout)	**2.07**	4.29	0.74	**0.63**

Model	FID	sFID	Prec	Rec
ImageNet 128×128				
BigGAN-deep [5]	6.02	7.18	**0.86**	0.35
LOGAN† [68]	3.36			
ADM	5.91	**5.09**	0.70	**0.65**
ADM-G (25 steps)	5.98	7.04	0.78	0.51
ADM-G	**2.97**	**5.09**	0.78	0.59
ImageNet 256×256				
DCTransformer† [42]	36.51	8.24	0.36	**0.67**
VQ-VAE-2†‡ [51]	31.11	17.38	0.36	0.57
IDDPM‡ [43]	12.26	5.42	0.70	0.62
SR3†‡ [53]	11.30			
BigGAN-deep [5]	6.95	7.36	**0.87**	0.28
ADM	10.94	6.02	0.69	0.63
ADM-G (25 steps)	5.44	5.32	0.81	0.49
ADM-G	**4.59**	**5.25**	0.82	0.52
ImageNet 512×512				
BigGAN-deep [5]	8.43	8.13	**0.88**	0.29
ADM	23.24	10.19	0.73	**0.60**
ADM-G (25 steps)	8.41	9.67	0.83	0.47
ADM-G	**7.72**	**6.57**	0.87	0.42

Figure: Some data about beat GANs

Outline

The Process of DDPM



(a) Initial (b) Add noise (c) Final

Figure: how the figure tranform to a Gaussian Distribution

Forward Equation

Forward

$$q(x_t|x_{t-1}) = N(x_t; x_{t-1}\sqrt{1-\beta t}, I\beta_t) \tag{1}$$

$$x_t = x_{t-1}\sqrt{1-\beta_t} + z_t\sqrt{\beta_t} \tag{2}$$

$$\begin{aligned}
x_t &= x_{t-1}\sqrt{1-\beta_t} + z_t\sqrt{\beta_t} \quad (\text{let } 1-\beta_t = \alpha_t) && \text{(1a)}\\
&= x_{t-2}\sqrt{\alpha_t}(\sqrt{\alpha_{t-1}} + z_{t-2}\sqrt{1-\alpha_{t-1}}) + z_{t-1}\sqrt{1-\alpha_t} && \text{(2a)}\\
&= x_{t-2}\sqrt{\alpha_t\alpha_{t-1}} + z_{t-2}\sqrt{\alpha_t - \alpha_t\alpha_{t-1}} + z_{t-1}\sqrt{1-\alpha_t} && \text{(3a)}\\
&= x_{t-2}\sqrt{\alpha_t\alpha_{t-1}} + z\sqrt{1-\alpha_t\alpha_{t-1}} && \text{(4a)}\\
&= \cdots\\
&= x_0\sqrt{\bar{\alpha}_t} + z\sqrt{1-\bar{\alpha}_t} && \text{(5a)}
\end{aligned}$$

Backward Equation

$$
\begin{aligned}
q(x_{t-1}|x_t, x_0) &= q(x_t|x_{t-1}, x_0)\frac{q(x_{t-1}|x_0)}{q(x_t|x_0)} \\
&\propto \exp\left(-\frac{1}{2}\left(\left(\frac{\alpha_t}{\beta_t} + \frac{1}{1-\bar{\alpha}_{t-1}}\right)x_{t-1}^2 - \left(\frac{2\sqrt{\alpha_t}}{\beta_t}x_t + \frac{2\sqrt{\bar{\alpha}_{t-1}}}{1-\bar{\alpha}_{t-1}}\right)x_{t-1} + C(x_t, x_0)\right)\right) \\
&= \exp\left(-\frac{1}{2}\left(Ax_{t-1}^2 + Bx_{t-1} + C\right)\right)
\end{aligned}
$$

$$
\mu = -\frac{B}{2A}, \sigma^2 = \frac{1}{A} \tag{3}
$$

Backward

$$
p(x_t|x_{t-1}) = N(x_{t-1}; f_\mu(x_t, t), f_\sigma^2(x_t, t)) \tag{4}
$$

Outline

Text2Image

Generate images that match the description from the text, give some descriptions, and generate pictures that meet the description.

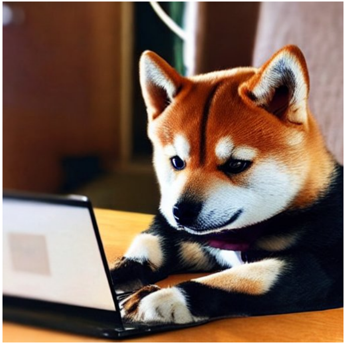


Figure: A dog is reading book

Image Refinement

Refine the blurry image to give the picture more detail.

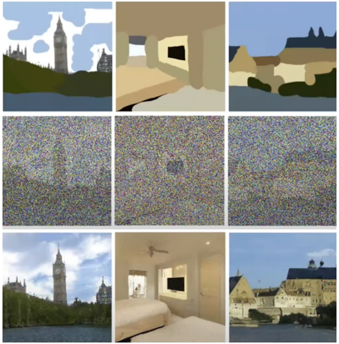


Figure: Image Refinement with Diffusion

Inpainting

Padding the missing image can complement the details and content of the image.

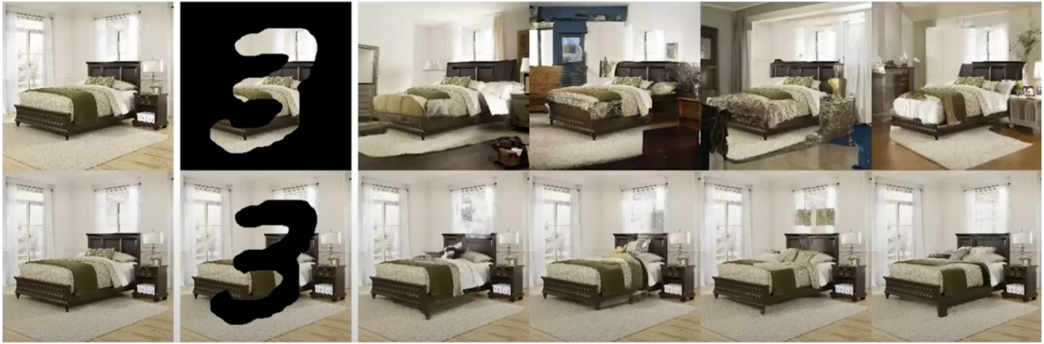


Figure: inpainting with Diffusion

Colorization

Color fill images with missing colors.

Figure: Colorization with Diffusion

Outline

Worries and Prospect

Worries

malicious uses for political purposes

Prospect

uses in art, photography, and music

References

Thank you for watching!

Here is the **References**

- Ho J, Jain A, Abbeel P. Denoising diffusion probabilistic models[J]. Advances in Neural Information Processing Systems, 2020, 33: 6840-6851.
- Dhariwal P, Nichol A. Diffusion models beat gans on image synthesis[J]. Advances in Neural Information Processing Systems, 2021, 34: 8780-8794.

Edit with LaTeX