Training-Free Scanning Robustness Guided Diffusion Model for Aesthetic QR Code Generation

Jia-Wei Liao

National Taiwan University[1], Academia Sinica[2]

Co-work with Winston Wang[2], Tzu-Sian Wang[2], Li-Xuan Peng[2],
Ju-Hsuan Weng[2], Cheng-Fu Chou[1], Jun-Cheng Chen[2]

June 3, 2024

Outline

Outline

1. Introduction

2. Diffusion Models

3. Iterative Refinement Algorithm

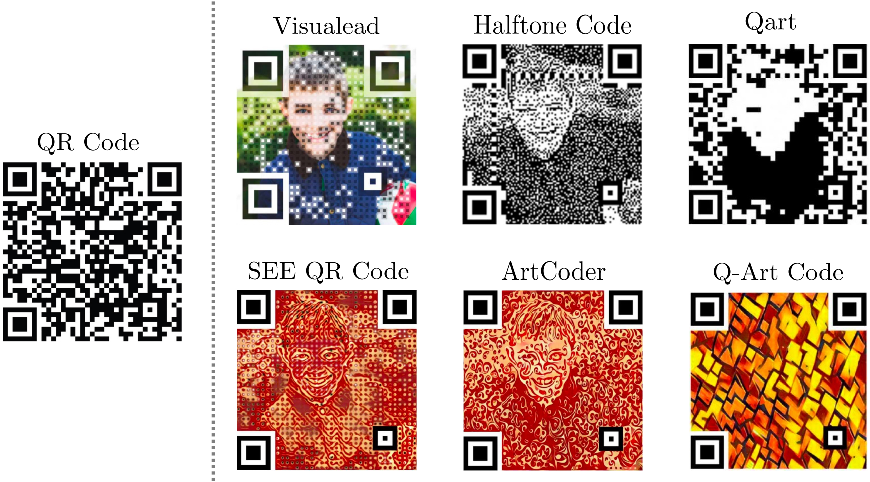
4. Experiments

5. Conclusion

Motivation

詹記麻辣火鍋
\徵/
內外場夥伴
3或4月新展店招募
啦
地點在西門町
!?
詹仆『掃我額頭上1111打com』
祖籍 山東
嘿!～呦!
嘿!～呦!
表!哥!台!北!
門!店!缺!人!
大!家!快!來!
嘿!～呦!
月均薪47,000 ~50,000
*$ 50,000含夜班津貼
*入職滿一年之待遇
*含四大獎金. 全勤獎金
【 四大獎金 】
1端午 & 2中秋獎金
＼依年資發放, 滿1年每節$7,500, 滿4年後每節$15,000／
3年終獎金 ＼每年發放1次, $30,000／
4紅利獎金 ＼每年發放1次, $ 50,000／
(*發放金額依公司當年度盈餘狀況、職等、評鑑並以面試現場說明為主)

Aesthetic QR Code



Our Work

QR Code

Winter wonderland, fresh snowfall, evergreen trees, cozy log cabin, smoke rising from chimney, aurora borealis in night sky.

Text Prompt

Aesthetic QR Code

Our Work

QR Code

Winter wonderland, fresh snowfall, evergreen trees, cozy log cabin, smoke rising from chimney, aurora borealis in night sky.

Text Prompt

Aesthetic QR Code

Our Work

QR Code

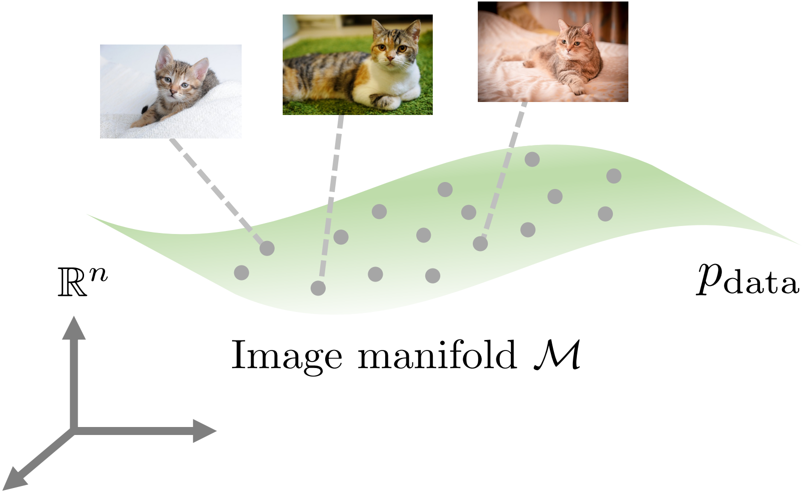
Winter wonderland, fresh snowfall, evergreen trees, cozy log cabin, smoke rising from chimney, aurora borealis in night sky.

Text Prompt

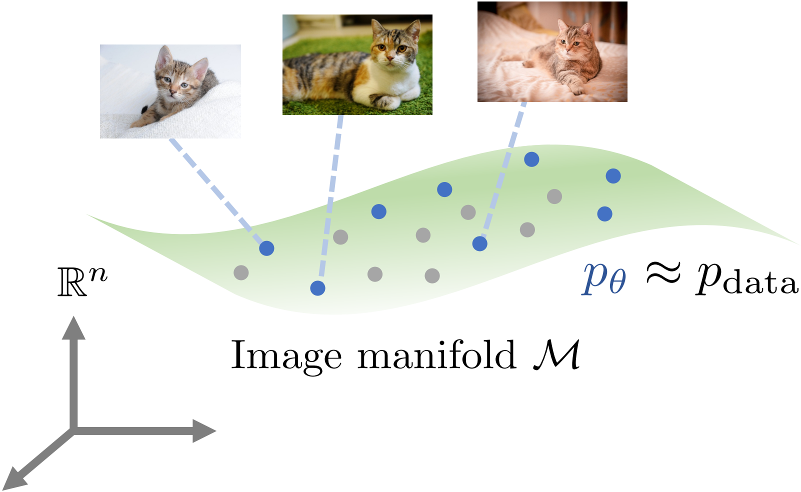
Aesthetic QR Code

Outline

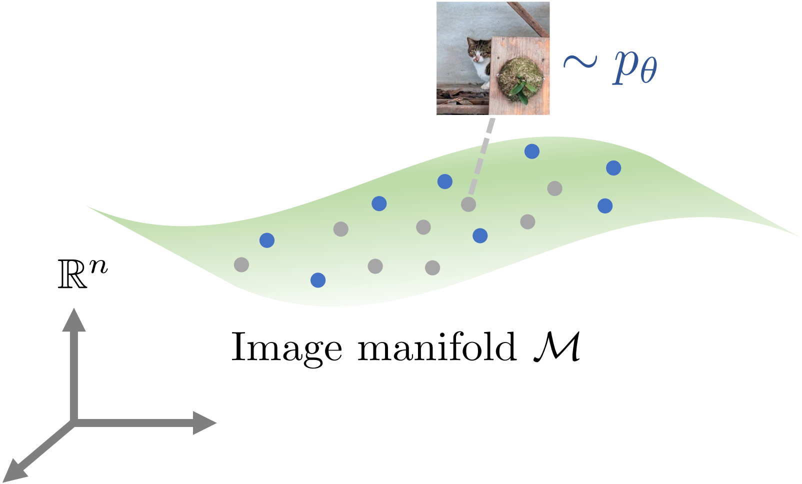
What is Generative Model Learning?



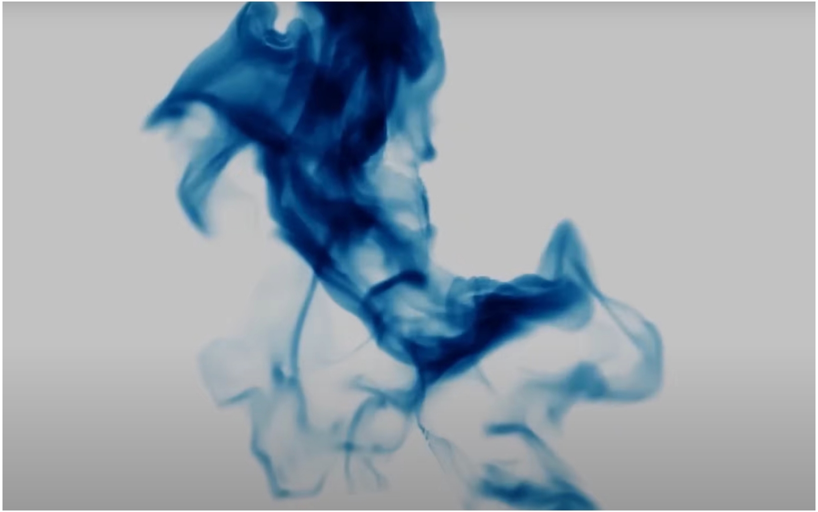
What is Generative Model Learning?



What is Generative Model Learning?



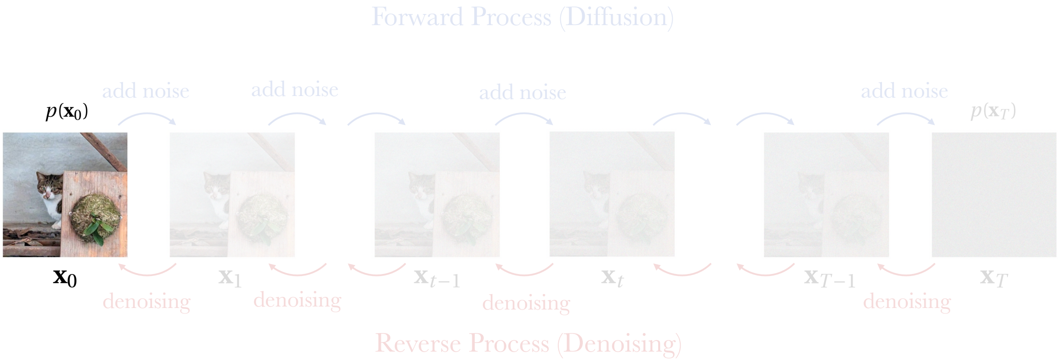
Diffusion Phenomenon



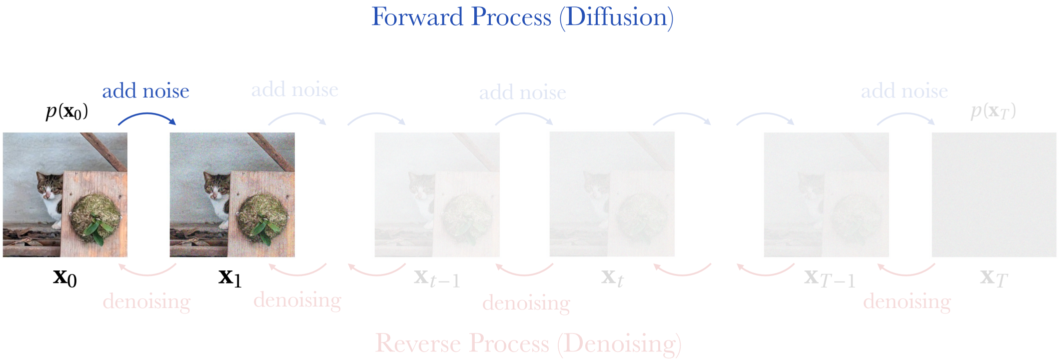
Development of Diffusion Models

- Sohl-Dickstein, Jascha, et al. "Deep unsupervised learning using nonequilibrium thermodynamics." International Conference on Machine Learning (ICML). 2015.
- Song, Yang, et al. "Score-Based Generative Modeling through Stochastic Differential Equations." International Conference on Learning Representations (ICLR). 2020.
- Ho, Jonathan, et al. "Denoising diffusion probabilistic models." Proceedings of the 34th International Conference on Neural Information Processing Systems (NeurIPS). 2020.

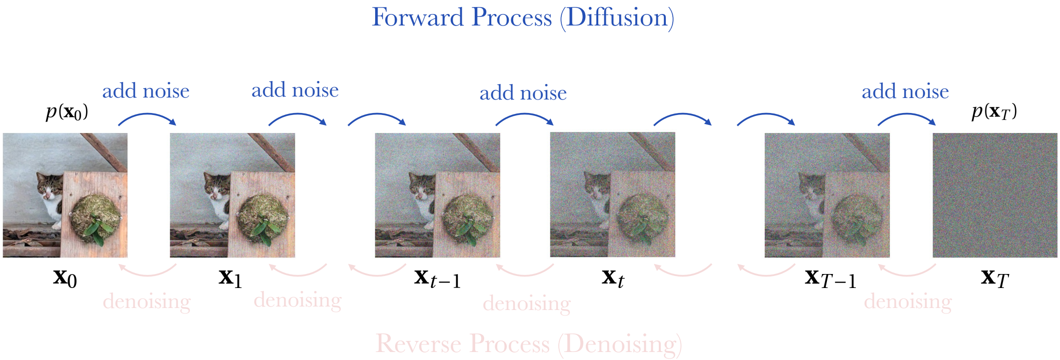
Diffusion Models: Forward Process



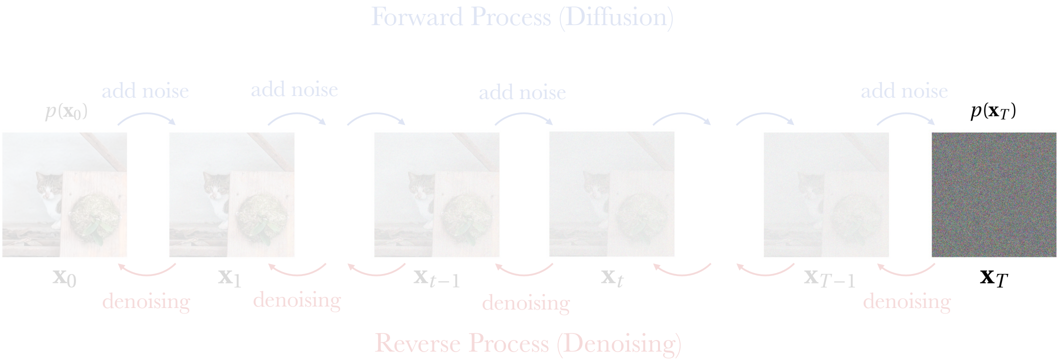
Diffusion Models: Forward Process



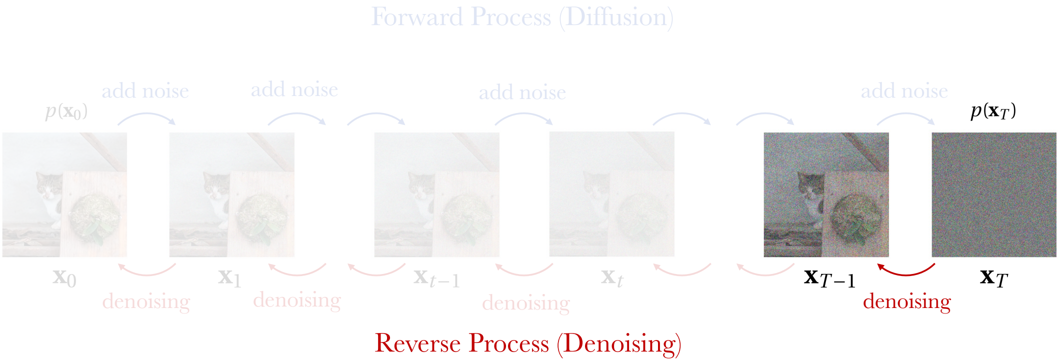
Diffusion Models: Forward Process



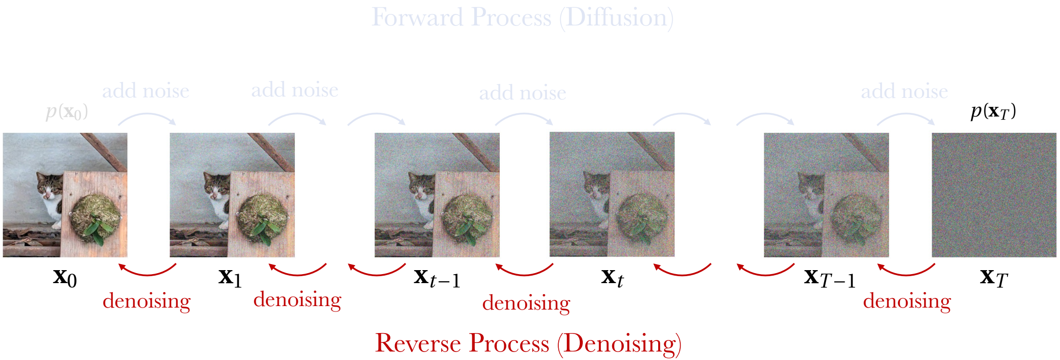
Diffusion Models: Reverse Process



Diffusion Models: Reverse Process



Diffusion Models: Reverse Process

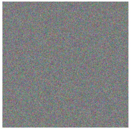


Diffusion Models

Given $1 > \alpha_1 > \alpha_2 > \ldots > \alpha_T > 0$,

$= \sqrt{\alpha_t}$

$+ \sqrt{1 - \alpha_t}$



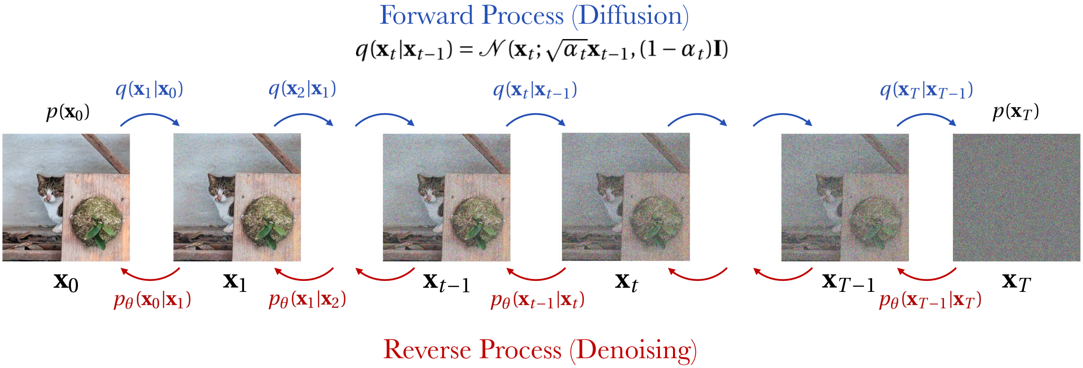
$\mathbf{x}_t$ $\quad$ $\mathbf{x}_{t-1}$ $\quad$ $\mathbf{z}_t \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$

$$\mathbf{x}_t \sim \mathcal{N}(\underbrace{\sqrt{\alpha_t}\mathbf{x}_{t-1}}_{\text{mean}}, \underbrace{(1 - \alpha_t)\mathbf{I}}_{\text{variance}})$$

$$\mathbf{x} \sim \mathcal{N}(\mu, \sigma^2) \iff \mathbf{x} = \mu + \sigma\mathbf{z} \quad \text{with} \quad \mathbf{z} \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$$

Diffusion Models

Given $1 > \alpha_1 > \alpha_2 > ... > \alpha_T > 0$,

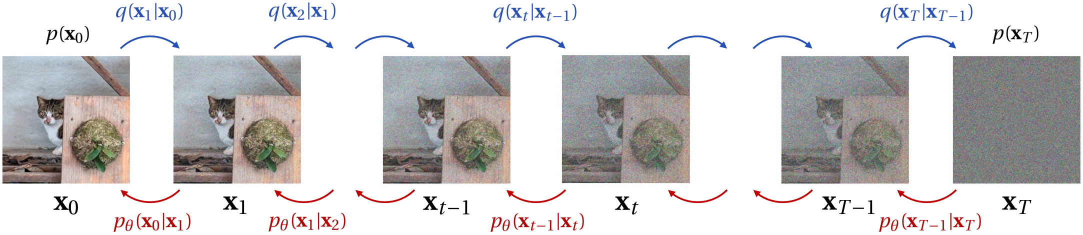


Diffusion Models

Given $1 > \alpha_1 > \alpha_2 > ... > \alpha_T > 0$,

Forward Process (Diffusion)

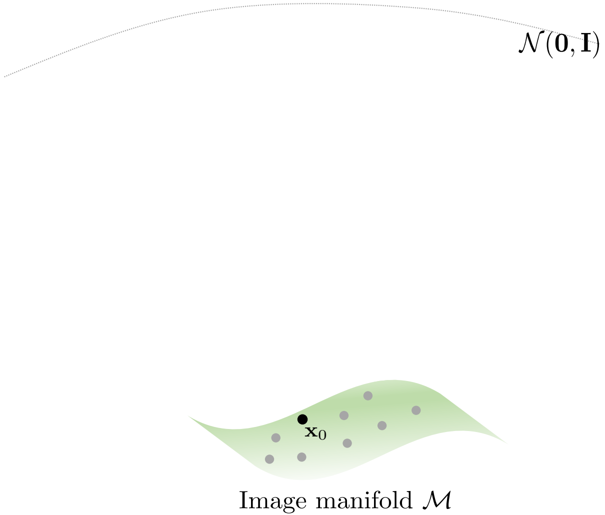
$$q(\mathbf{x}_t|\mathbf{x}_{t-1}) = \mathcal{N}(\mathbf{x}_t; \sqrt{\alpha_t}\mathbf{x}_{t-1}, (1-\alpha_t)\mathbf{I})$$



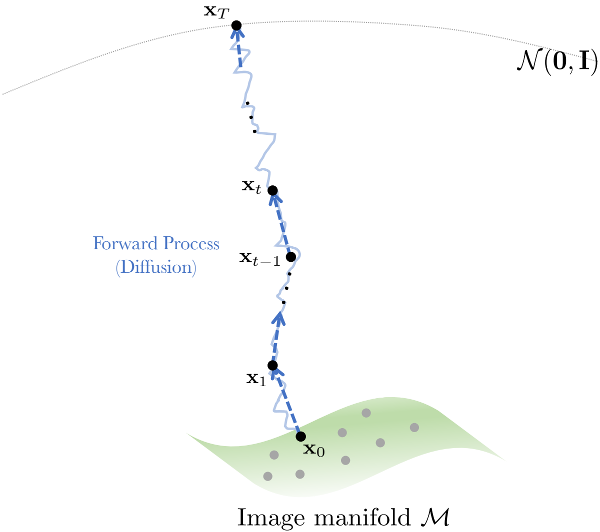
Reverse Process (Denoising)

1. $p(\mathbf{x}_T) = \mathcal{N}(\mathbf{0}, \mathbf{I})$
2. $p_\theta(\mathbf{x}_{t-1}|\mathbf{x}_t) = \mathcal{N}(\mathbf{x}_{t-1}; \mu_\theta(\mathbf{x}_t, t), \sigma_t^2\mathbf{I})$

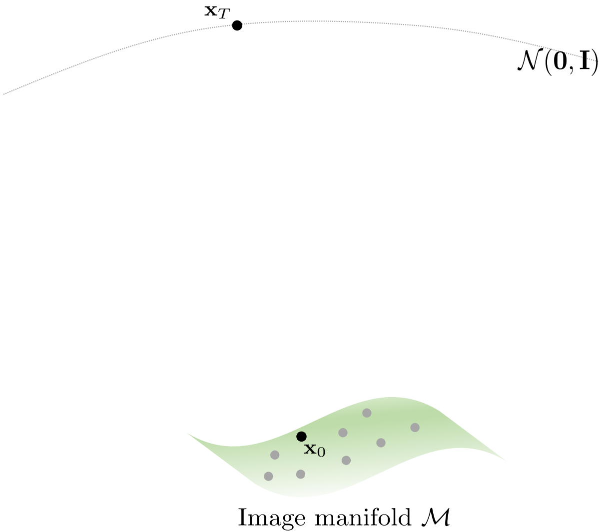
Diffusion Process of Image Manifold



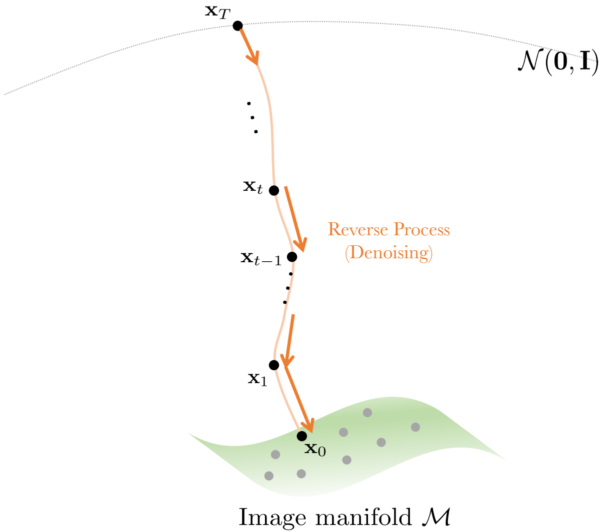
Diffusion Process of Image Manifold



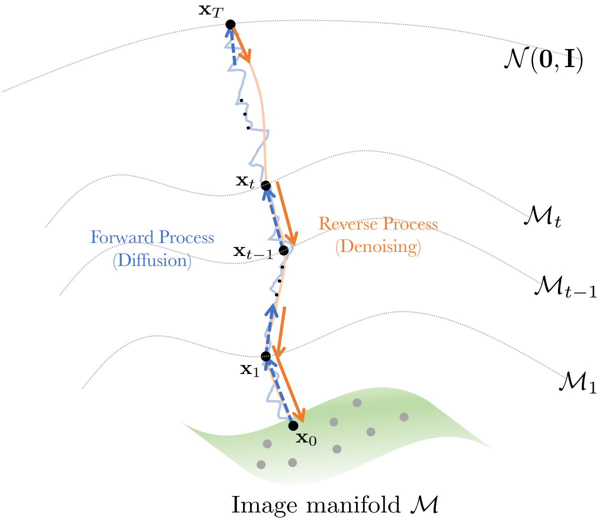
Reverse Process back to Image Manifold



Reverse Process back to Image Manifold



Reverse Process back to Image Manifold



Derivation of Distribution at Last Timestep

Since $p(\mathbf{x}_t|\mathbf{x}_{t-1}) = \mathcal{N}(\mathbf{x}_t; \sqrt{\alpha_t}\mathbf{x}_{t-1}, (1-\alpha_t)\mathbf{I})$,

$$p(\mathbf{x}_t|\mathbf{x}_0) = \mathcal{N}(\mathbf{x}_t; \sqrt{\bar{\alpha}_t}\mathbf{x}_0, (1-\bar{\alpha}_t)\mathbf{I}).$$

where $\bar{\alpha}_t = \Pi_{i=1}^t \alpha_i$. Then

$$p(\mathbf{x}_t) = \int p(\mathbf{x}_t|\mathbf{x}_0)p(\mathbf{x}_0)\mathrm{d}\mathbf{x}_0 \to \mathcal{N}(\mathbf{0}, \mathbf{I}).$$

by letting $t \to \infty$. We assume the final step distribution $p(\mathbf{x}_T)$ is standard normal distribution, i.e., $p(\mathbf{x}_T) = \mathcal{N}(\mathbf{0}, \mathbf{I})$.

Derivation of Reverse Process

Notice that

$$p(\mathbf{x}_t|\mathbf{x}_{t-1}) = \mathcal{N}(\mathbf{x}_t; \sqrt{\alpha_t}\mathbf{x}_{t-1}, (1-\alpha_t)\mathbf{I}), \quad p(\mathbf{x}_t|\mathbf{x}_0) = \mathcal{N}(\mathbf{x}_t; \sqrt{\bar{\alpha}_t}\mathbf{x}_0, (1-\bar{\alpha}_t)\mathbf{I}).$$

By Bayes' theorem, we can derive the conditional distribution in reverse process

$$p(\mathbf{x}_{t-1}|\mathbf{x}_t, \mathbf{x}_0) = \frac{p(\mathbf{x}_t|\mathbf{x}_{t-1})\, p(\mathbf{x}_{t-1}|\mathbf{x}_0)}{p(\mathbf{x}_t|\mathbf{x}_0)}.$$

Therefore, we have $p(\mathbf{x}_{t-1}|\mathbf{x}_t, \mathbf{x}_0) = \mathcal{N}(\mathbf{x}_{t-1}; \mu_t(\mathbf{x}_t, \mathbf{x}_0), \sigma_t^2\mathbf{I})$ with

$$\mu_t(\mathbf{x}_t, \mathbf{x}_0) = \frac{\sqrt{\alpha_t}(1-\bar{\alpha}_{t-1})}{1-\bar{\alpha}_t}\mathbf{x}_t + \frac{\sqrt{\bar{\alpha}_{t-1}}(1-\alpha_t)}{1-\bar{\alpha}_t}\mathbf{x}_0 \quad \text{and} \quad \sigma_t^2 = \frac{(1-\bar{\alpha}_{t-1})(1-\alpha_t)}{1-\bar{\alpha}_t}.$$

Diffusion Models

- **Forward Process:**

$$p(\mathbf{x}_t|\mathbf{x}_{t-1}) = \mathcal{N}(\mathbf{x}_t; \sqrt{\alpha_t}\mathbf{x}_{t-1}, (1-\alpha_t)\mathbf{I})$$

- **Reverse Process:**

$$p(\mathbf{x}_{t-1}|\mathbf{x}_t, \mathbf{x}_0) = \mathcal{N}(\mathbf{x}_{t-1}; \mu_t(\mathbf{x}_t, \mathbf{x}_0), \sigma_t^2\mathbf{I})$$

Diffusion Models

- **Forward Process:**

$$p(\mathbf{x}_t|\mathbf{x}_{t-1}) = \mathcal{N}(\mathbf{x}_t; \sqrt{\alpha_t}\mathbf{x}_{t-1}, (1-\alpha_t)\mathbf{I})$$

- **Reverse Process:**

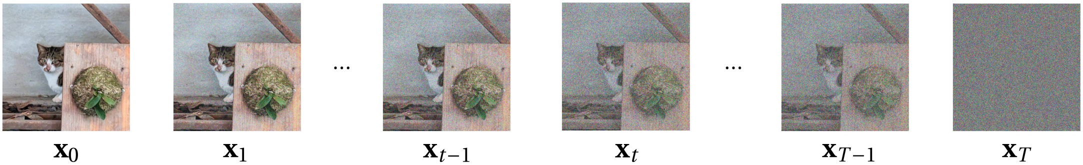
$$p(\mathbf{x}_{t-1}|\mathbf{x}_t, \mathbf{x}_0) = \mathcal{N}(\mathbf{x}_{t-1}; \mu_t(\mathbf{x}_t, \mathbf{x}_0), \sigma_t^2\mathbf{I})$$

In practice, we don't have $\mathbf{x}_0$. Thus, our goal is to train the deep learning model to reconstruct $\mathbf{x}_0$ such that

$$p_\theta(\mathbf{x}_{t-1}|\mathbf{x}_t) := p(\mathbf{x}_{t-1}|\mathbf{x}_t, \hat{\mathbf{x}}_{0|t}) \approx p(\mathbf{x}_{t-1}|\mathbf{x}_t, \mathbf{x}_0).$$

Estimating $\mathbf{x}_0$ and Training Objective

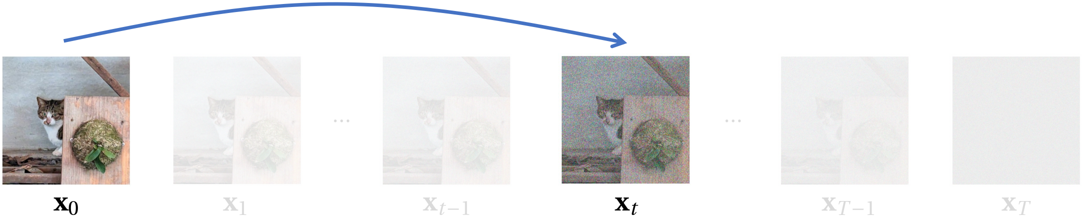
$$\mathcal{L}(\theta) = \mathbb{E}_{\mathbf{x}_0 \sim \mathcal{D}^N, t \sim U(1,T), \epsilon_t \sim \mathcal{N}(\mathbf{0},\mathbf{I})} \| \epsilon_\theta (\underbrace{\sqrt{\bar{\alpha}_t}\mathbf{x}_0 + \sqrt{1-\bar{\alpha}_t}\epsilon_t}_{\mathbf{x}_t}, t) - \epsilon_t \|_2^2.$$



Estimating $\mathbf{x}_0$ and Training Objective

$$\mathcal{L}(\theta) = \mathbb{E}_{\mathbf{x}_0 \sim \mathcal{D}^N, t \sim U(1,T), \epsilon_t \sim \mathcal{N}(\mathbf{0},\mathbf{I})} \| \epsilon_\theta(\underbrace{\sqrt{\bar{\alpha}_t}\mathbf{x}_0 + \sqrt{1-\bar{\alpha}_t}\epsilon_t}_{\mathbf{x}_t}, t) - \epsilon_t \|_2^2.$$

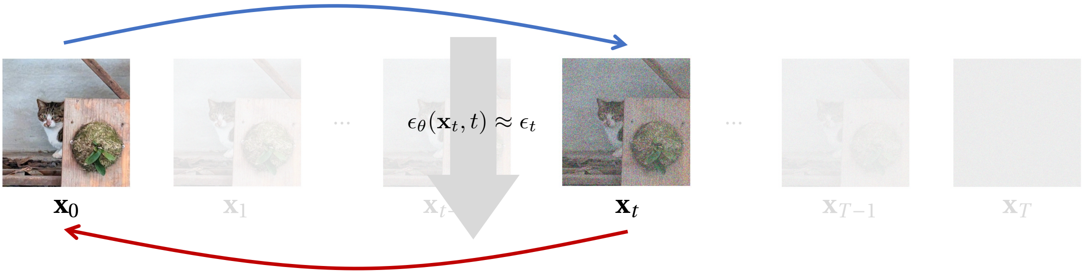
$$\mathbf{x}_t = \sqrt{\bar{\alpha}_t}\mathbf{x}_0 + \sqrt{1-\bar{\alpha}_t}\epsilon_t$$



Estimating $\mathbf{x}_0$ and Training Objective

$$\mathcal{L}(\theta) = \mathbb{E}_{\mathbf{x}_0 \sim \mathcal{D}^N, t \sim U(1,T), \epsilon_t \sim \mathcal{N}(\mathbf{0},\mathbf{I})} \| \epsilon_\theta(\underbrace{\sqrt{\bar{\alpha}_t}\mathbf{x}_0 + \sqrt{1-\bar{\alpha}_t}\epsilon_t}_{\mathbf{x}_t}, t) - \epsilon_t \|_2^2.$$

$$\mathbf{x}_t = \sqrt{\bar{\alpha}_t}\mathbf{x}_0 + \sqrt{1-\bar{\alpha}_t}\epsilon_t$$



$$\hat{\mathbf{x}}_{0|t} = \frac{1}{\sqrt{\bar{\alpha}_t}}\left(\mathbf{x}_t - \sqrt{1-\bar{\alpha}_t}\epsilon_\theta(\mathbf{x}_t, t)\right)$$

Sampling Algorithm

1. $\mathbf{x}_T \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$.
2. $\mathbf{x}_{t-1} \sim p_\theta(\mathbf{x}_{t-1}|\mathbf{x}_t) = \mathcal{N}(\mu_\theta(\mathbf{x}_t, t), \sigma_t^2\mathbf{I}), t = T, T-1, \ldots, 1$.

Sampling Algorithm

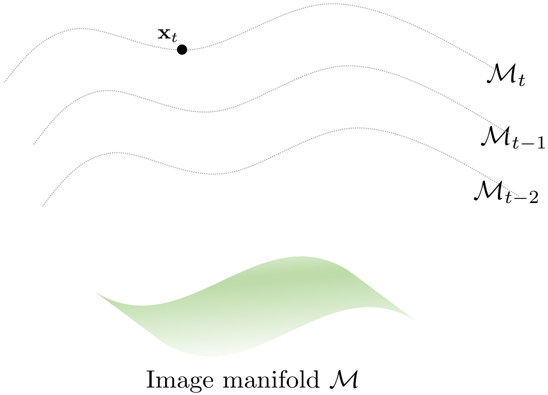
1. $\mathbf{x}_T \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$.
2. $\mathbf{x}_{t-1} \sim p_\theta(\mathbf{x}_{t-1}|\mathbf{x}_t) = \mathcal{N}(\mu_\theta(\mathbf{x}_t, t), \sigma_t^2 \mathbf{I}), t = T, T-1, ..., 1$.

$$
\begin{aligned}
\mathbf{x}_{t-1} &= \mu_\theta(\mathbf{x}_t, t) + \sigma_t \mathbf{z}_t \\
&= \frac{\sqrt{\alpha_t}(1-\bar{\alpha}_{t-1})}{1-\bar{\alpha}_t}\mathbf{x}_t + \frac{\sqrt{\bar{\alpha}_{t-1}}(1-\alpha_t)}{1-\bar{\alpha}_t}\hat{\mathbf{x}}_{0|t} + \sqrt{\frac{(1-\bar{\alpha}_{t-1})(1-\alpha_t)}{1-\bar{\alpha}_t}}\mathbf{z}_t,
\end{aligned}
$$

where $\mathbf{z}_t \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$ and $\hat{\mathbf{x}}_{0|t} = \frac{1}{\sqrt{\bar{\alpha}_t}}\left(\mathbf{x}_t - \sqrt{1-\bar{\alpha}_t}\epsilon_\theta(\mathbf{x}_t, t)\right)$.

Sampling Algorithm with Reverse Process

$$\mathbf{x}_{t-1} = \frac{\sqrt{\alpha_t}(1-\bar{\alpha}_{t-1})}{1-\bar{\alpha}_t}\mathbf{x}_t + \frac{\sqrt{\bar{\alpha}_{t-1}}(1-\alpha_t)}{1-\bar{\alpha}_t}\hat{\mathbf{x}}_{0|t} + \sqrt{\frac{(1-\bar{\alpha}_{t-1})(1-\alpha_t)}{1-\bar{\alpha}_t}}\mathbf{z}_t$$



Sampling Algorithm with Reverse Process

$$\mathbf{x}_{t-1} = \frac{\sqrt{\alpha_t}(1-\bar{\alpha}_{t-1})}{1-\bar{\alpha}_t}\mathbf{x}_t + \frac{\sqrt{\bar{\alpha}_{t-1}}(1-\alpha_t)}{1-\bar{\alpha}_t}\hat{\mathbf{x}}_{0|t} + \sqrt{\frac{(1-\bar{\alpha}_{t-1})(1-\alpha_t)}{1-\bar{\alpha}_t}}\mathbf{z}_t$$

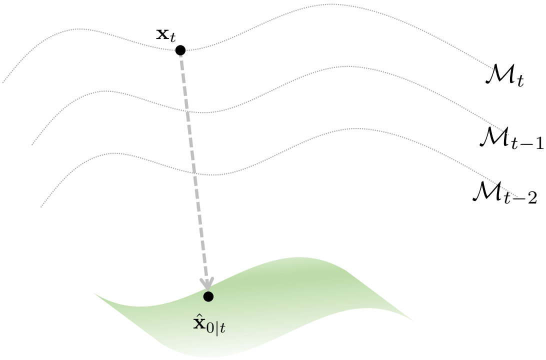
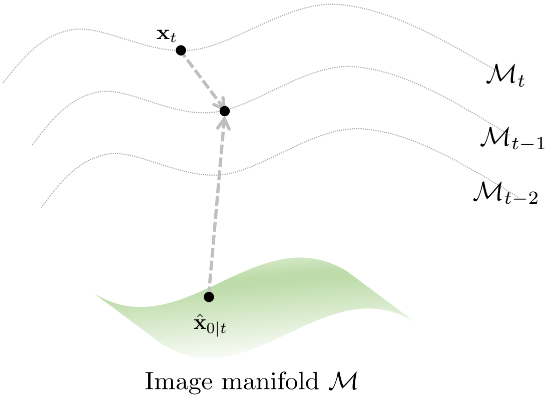


Image manifold $\mathcal{M}$

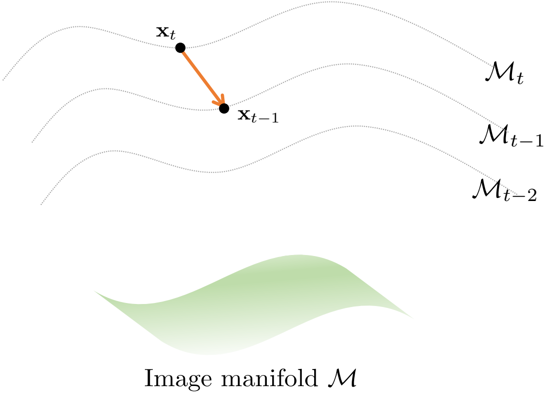
Sampling Algorithm with Reverse Process

$$\mathbf{x}_{t-1} = \frac{\sqrt{\alpha_t}(1-\bar{\alpha}_{t-1})}{1-\bar{\alpha}_t}\mathbf{x}_t + \frac{\sqrt{\bar{\alpha}_{t-1}}(1-\alpha_t)}{1-\bar{\alpha}_t}\hat{\mathbf{x}}_{0|t} + \sqrt{\frac{(1-\bar{\alpha}_{t-1})(1-\alpha_t)}{1-\bar{\alpha}_t}}\mathbf{z}_t$$



Sampling Algorithm with Reverse Process

$$\mathbf{x}_{t-1} = \frac{\sqrt{\alpha_t}(1-\bar{\alpha}_{t-1})}{1-\bar{\alpha}_t}\mathbf{x}_t + \frac{\sqrt{\bar{\alpha}_{t-1}}(1-\alpha_t)}{1-\bar{\alpha}_t}\hat{\mathbf{x}}_{0|t} + \sqrt{\frac{(1-\bar{\alpha}_{t-1})(1-\alpha_t)}{1-\bar{\alpha}_t}}\mathbf{z}_t$$



Sampling Algorithm with Reverse Process

$$\mathbf{x}_{t-1} = \frac{\sqrt{\alpha_t}(1-\bar{\alpha}_{t-1})}{1-\bar{\alpha}_t}\mathbf{x}_t + \frac{\sqrt{\bar{\alpha}_{t-1}}(1-\alpha_t)}{1-\bar{\alpha}_t}\hat{\mathbf{x}}_{0|t} + \sqrt{\frac{(1-\bar{\alpha}_{t-1})(1-\alpha_t)}{1-\bar{\alpha}_t}}\mathbf{z}_t$$

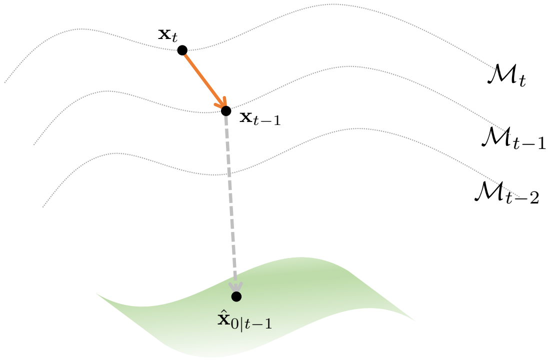


Image manifold $\mathcal{M}$

Sampling Algorithm with Reverse Process

$$\mathbf{x}_{t-1} = \frac{\sqrt{\alpha_t}(1-\bar{\alpha}_{t-1})}{1-\bar{\alpha}_t}\mathbf{x}_t + \frac{\sqrt{\bar{\alpha}_{t-1}}(1-\alpha_t)}{1-\bar{\alpha}_t}\hat{\mathbf{x}}_{0|t} + \sqrt{\frac{(1-\bar{\alpha}_{t-1})(1-\alpha_t)}{1-\bar{\alpha}_t}}\mathbf{z}_t$$

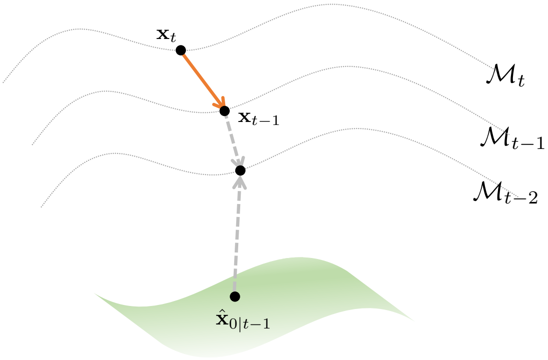
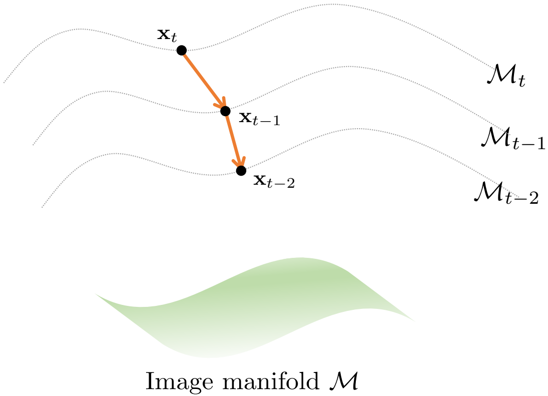


Image manifold $\mathcal{M}$

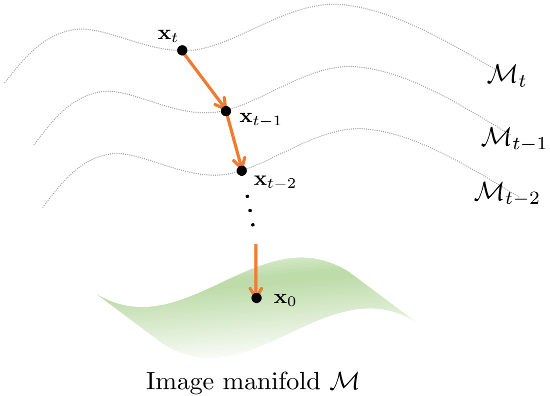
Sampling Algorithm with Reverse Process

$$\mathbf{x}_{t-1} = \frac{\sqrt{\alpha_t}(1-\bar{\alpha}_{t-1})}{1-\bar{\alpha}_t}\mathbf{x}_t + \frac{\sqrt{\bar{\alpha}_{t-1}}(1-\alpha_t)}{1-\bar{\alpha}_t}\hat{\mathbf{x}}_{0|t} + \sqrt{\frac{(1-\bar{\alpha}_{t-1})(1-\alpha_t)}{1-\bar{\alpha}_t}}\mathbf{z}_t$$



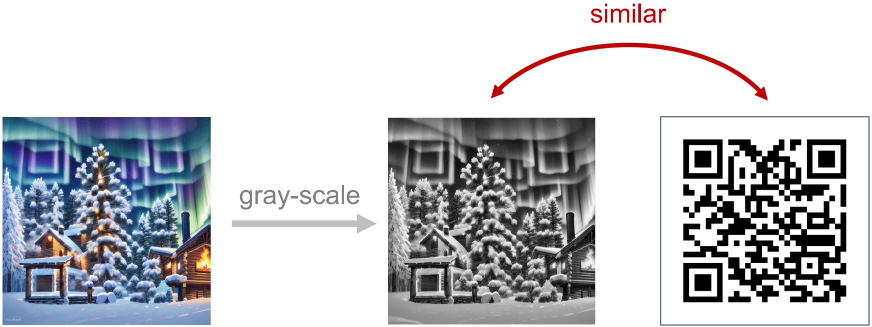
Sampling Algorithm with Reverse Process

$$\mathbf{x}_{t-1} = \frac{\sqrt{\alpha_t}(1-\bar{\alpha}_{t-1})}{1-\bar{\alpha}_t}\mathbf{x}_t + \frac{\sqrt{\bar{\alpha}_{t-1}}(1-\alpha_t)}{1-\bar{\alpha}_t}\hat{\mathbf{x}}_{0|t} + \sqrt{\frac{(1-\bar{\alpha}_{t-1})(1-\alpha_t)}{1-\bar{\alpha}_t}}\mathbf{z}_t$$



Outline

Similarity Between Image and QR Code



$\mathbf{x}$ $\mathcal{G}(\mathbf{x})$ $\mathbf{y}$

Our goal is to define a smooth loss function to measure the similarity between image and QR code.

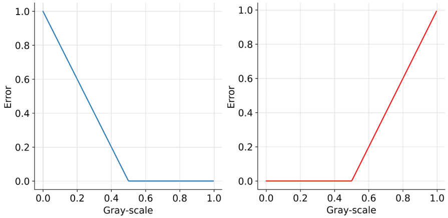
Scanning-Robust Loss

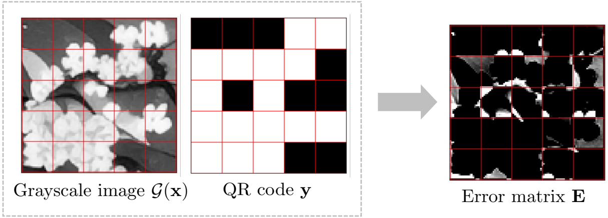
We define the Scanning Robust Loss (SRL) as

$$\mathscr{L}_{\mathrm{SR}}(\mathbf{x}, \mathbf{y}) = \|\mathrm{vec}(\mathbf{E})\|_1,$$

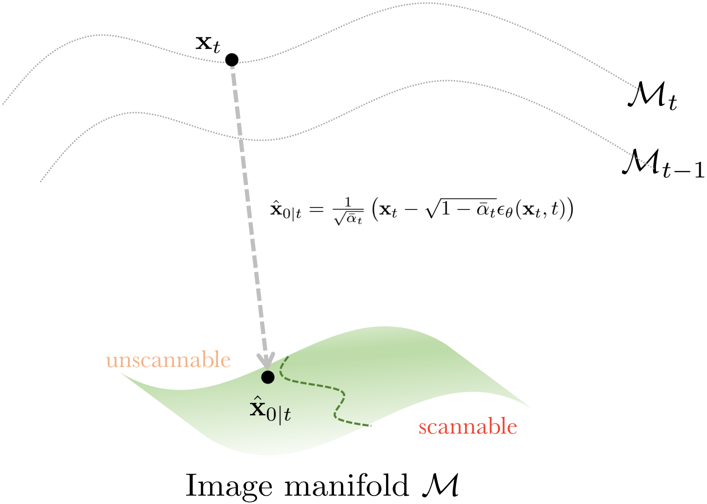
where the error matrix $\mathbf{E}$ is

$$\mathbf{E} = \max(1 - 2\mathscr{G}(\mathbf{x}), 0) \odot \mathbf{y} + \max(2\mathscr{G}(\mathbf{x}) - 1, 0) \odot (1 - \mathbf{y}).$$

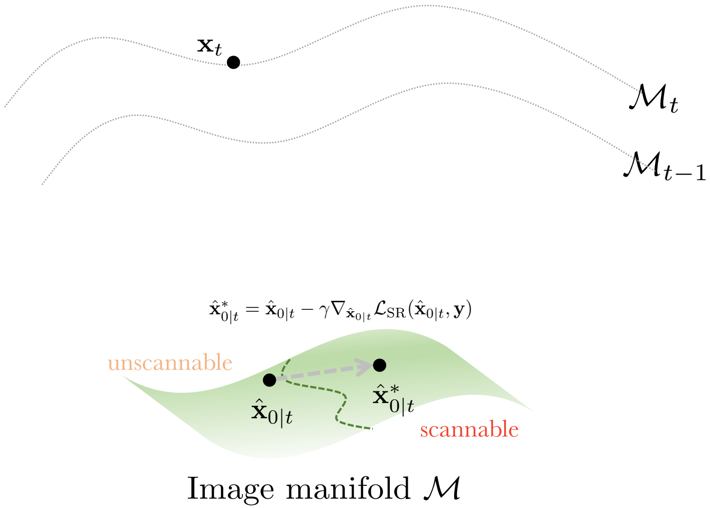




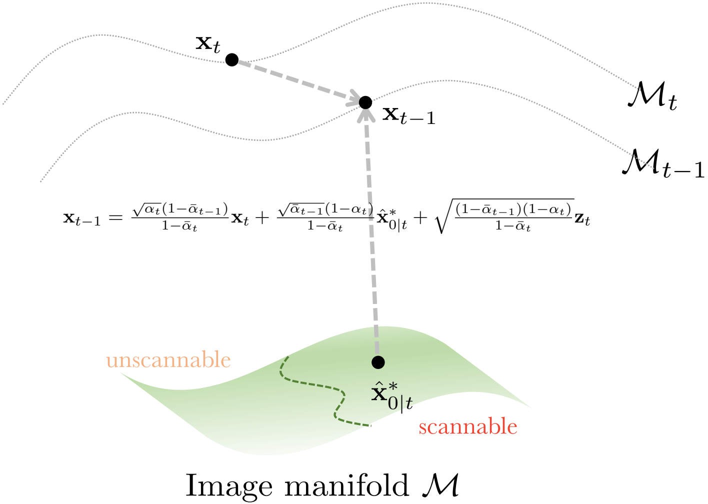
Sampling with Iterative Refinement Algorithm



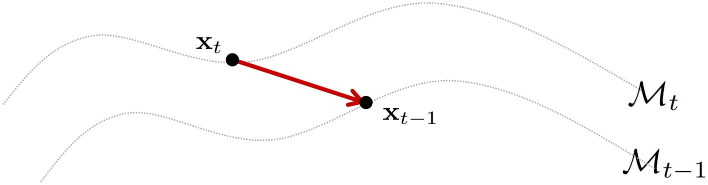
Sampling with Iterative Refinement Algorithm



Sampling with Iterative Refinement Algorithm



Sampling with Iterative Refinement Algorithm



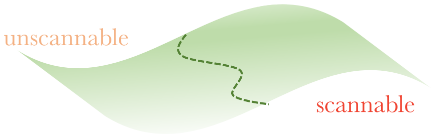


Image manifold $\mathcal{M}$

Sampling with Iterative Refinement Algorithm

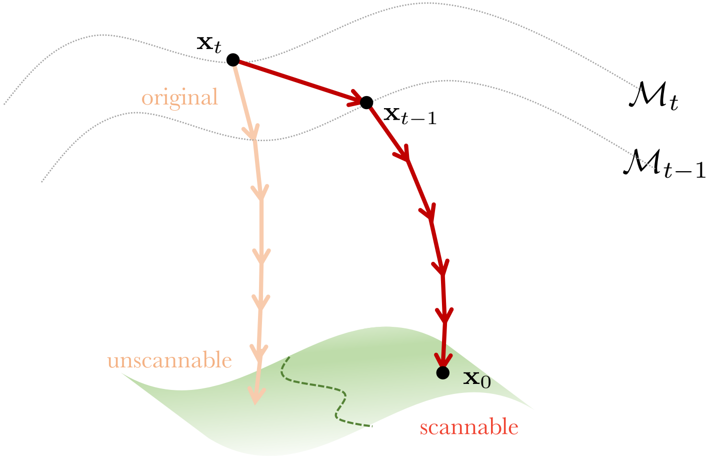


Image manifold $\mathcal{M}$

Sampling with Iterative Refinement Algorithm

For $t = T, T-1, \ldots, 1$,

$$\hat{\mathbf{x}}_{0|t} = \frac{1}{\sqrt{\bar{\alpha}_t}}\left(\mathbf{x}_t - \sqrt{1-\bar{\alpha}_t}\epsilon_\theta(\mathbf{x}_t, t)\right)$$

$$\hat{\mathbf{x}}^*_{0|t} = \hat{\mathbf{x}}_{0|t} - \gamma \nabla_{\hat{\mathbf{x}}_{0|t}} \mathscr{L}_{\mathrm{SR}}(\hat{\mathbf{x}}_{0|t}, \mathbf{y})$$

$$\mathbf{x}_{t-1} = \frac{\sqrt{\alpha_t}(1-\bar{\alpha}_{t-1})}{1-\bar{\alpha}_t}\mathbf{x}_t + \frac{\sqrt{\bar{\alpha}_{t-1}}(1-\alpha_t)}{1-\bar{\alpha}_t}\hat{\mathbf{x}}^*_{0|t} + \sqrt{\frac{(1-\bar{\alpha}_{t-1})(1-\alpha_t)}{1-\bar{\alpha}_t}}\mathbf{z}_t$$

Outline

Qualitative Results

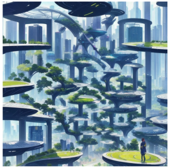
Original QR Code

Winter wonderland, fresh snowfall, evergreen trees, cozy log cabin, smoke rising from chimney, aurora borealis in night sky.

Cherry blossom festival, pink petals floating in the air, traditional lanterns, peaceful river, people in kimonos, sunny day.

Majestic waterfall, lush rainforest, rainbow in the mist, exotic birds, vibrant flowers, serene pool below.

Abandoned amusement park, overgrown rides, haunting beauty, sense of nostalgia, sunset lighting.



Futuristic urban park, green spaces amid skyscrapers, eco-friendly design, people enjoying outdoors, advanced city life.

Old European town square, cobblestone streets, café terraces, flowering balconies, gothic cathedral, bustling morning.

Lost city of Atlantis, underwater ruins, mythical creatures, ancient mysteries, ocean exploration.

Old Western saloon at night, lively music, dancing, vintage decor, sense of time travel.

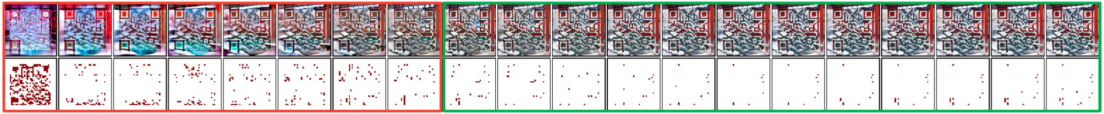
Quantitative Comparison with Other Methods

- SSR: Scanning Success Rate
- LAS: LAION Aesthetic Score

Method	**SSR** $\uparrow$	**LAS** $\uparrow$
QR Diffusion	96%	5.5150
QR Code AI Art	90%	5.7003
QRBTF	56%	7.0817
QR Code Monster	60%	7.0661
Ours	99%	6.8233

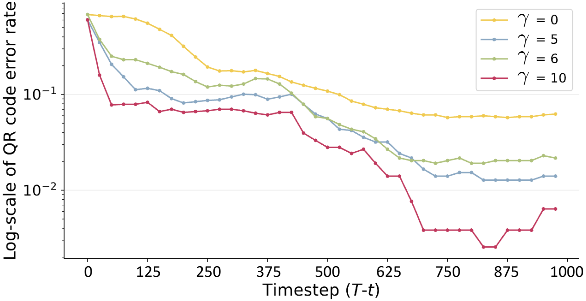
Table: Quantitative results.

Error Analysis



Unscannable
(#1000 ~ #625 timestep)

Scannable
(#625 ~ #1 timestep)



Outline

Conclusion

- We develop a training-free iterative refinement algorithm for diffusion model with the development of the Scanning-Robust Loss (SRL), significantly enhancing QR code scannability without compromising visual appeal.
- We demonstrated a higher scanning success rate compared to commercial alternatives, maintaining visual quality and confirming the suitability of these QR codes for real-world applications.

Thank you!