

Deep Learning I Period 4 - Week 5

Diffusion Models

Stijn Oomes

Lecturer Applied Mathematics
Stijn.Oomes@Inholland.nl

Overview

- Generative AI → Veo3
- Deep Learning versus Generative AI
- Diffusion Models
- → training & sampling a model to generate novel fashion

Generative AI

Generative AI

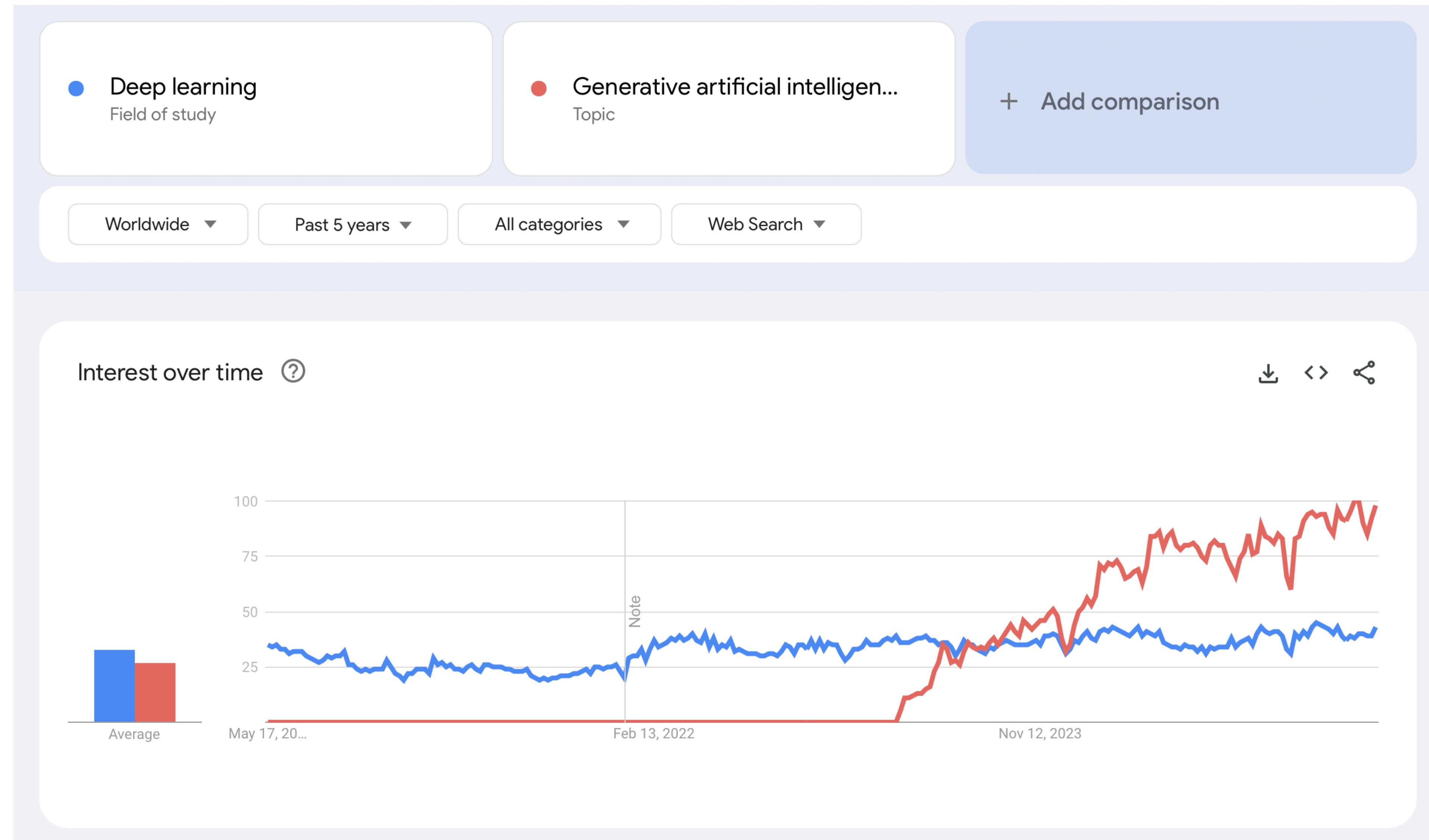
- Google DeepMind Veo3
 - <https://deepmind.google/models/veo/>





Deep Learning versus Generative AI

Deep Learning versus Generative AI



Deep Learning versus Generative AI

- Deep Learning
 - mostly about perception
 - what objects are in the image?
 - where are those objects?
- Generative AI
 - mostly about “mimicry”
 - does this image look cool / convincing?

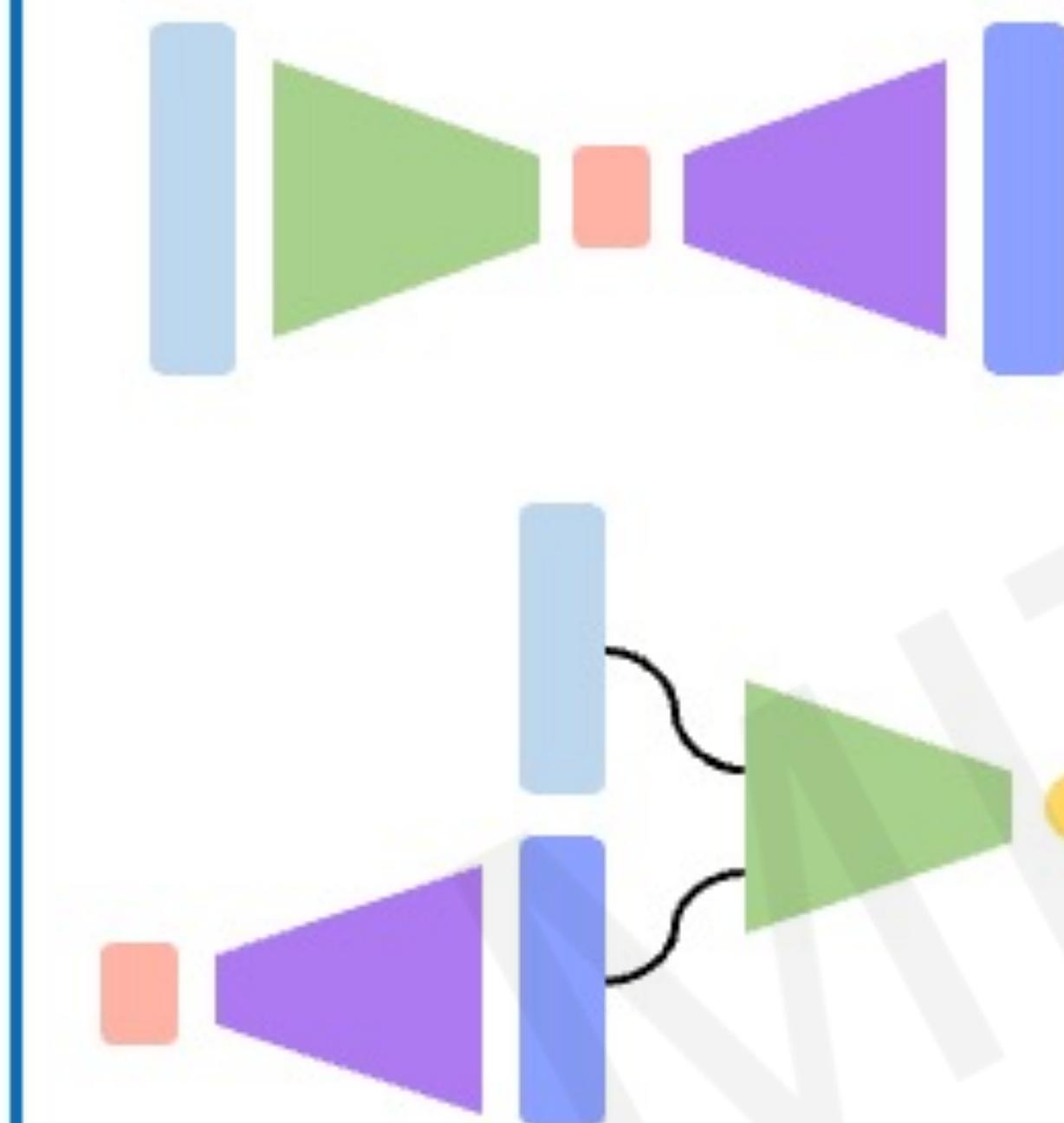
Deep Learning versus Generative AI

- Deep Learning
 - mostly about perception
 - **we can check if it is correct**
- Generative AI
 - mostly about “mimicry”
 - **we can only check if we “appreciate” it**

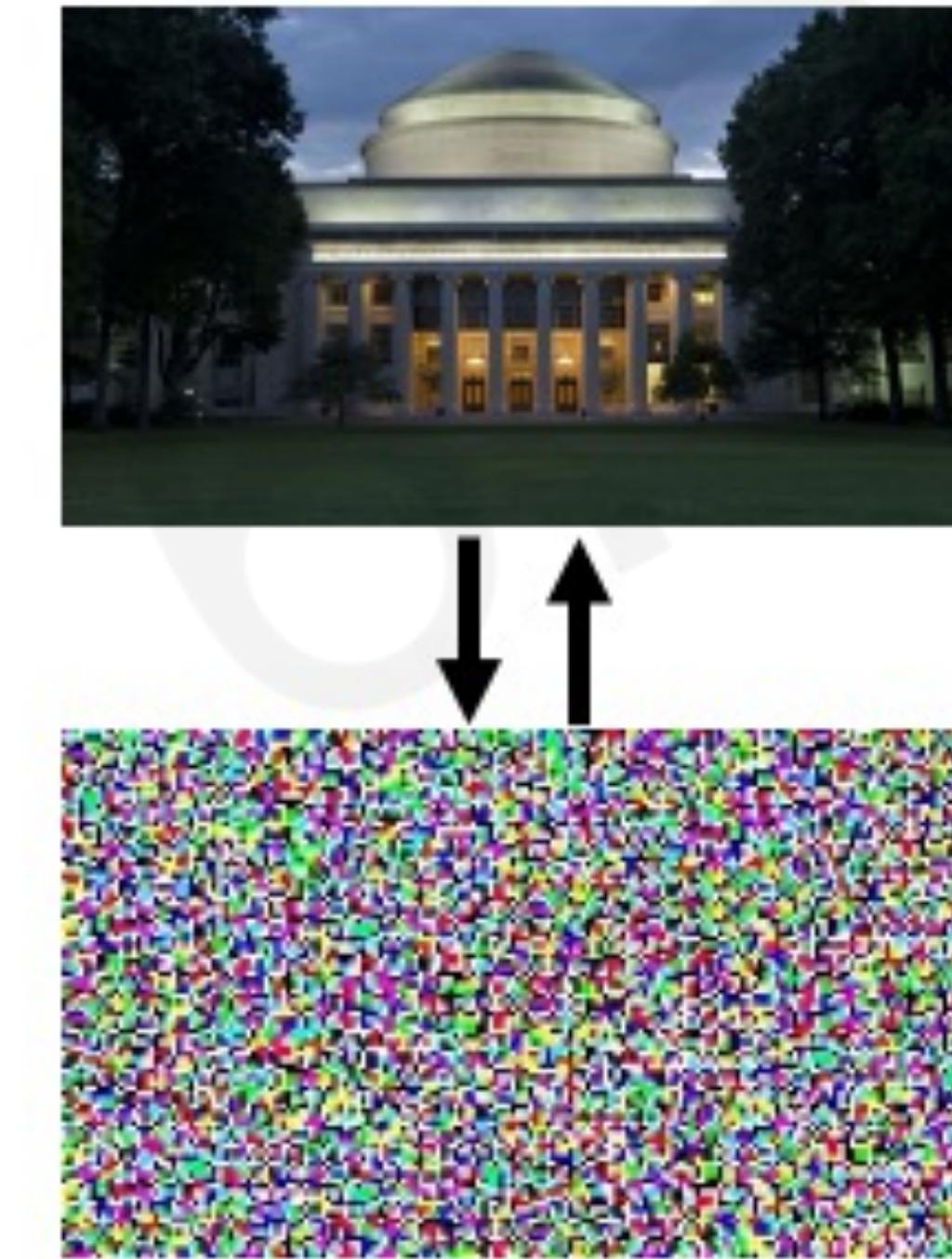
Diffusion Models

The Landscape of Generative Modeling

Lecture 4: VAEs and GANs



Diffusion Models

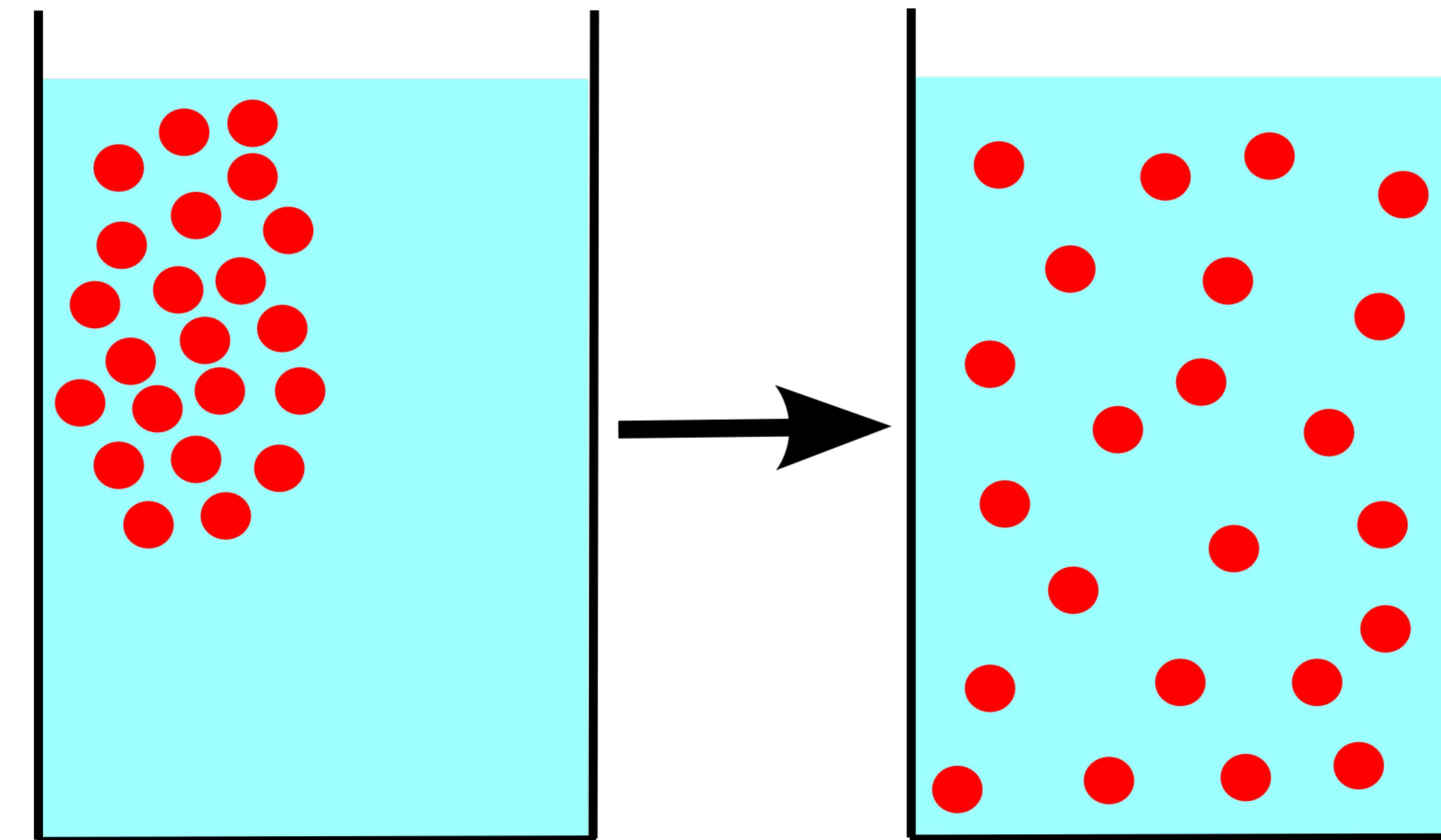


Text-to-Image

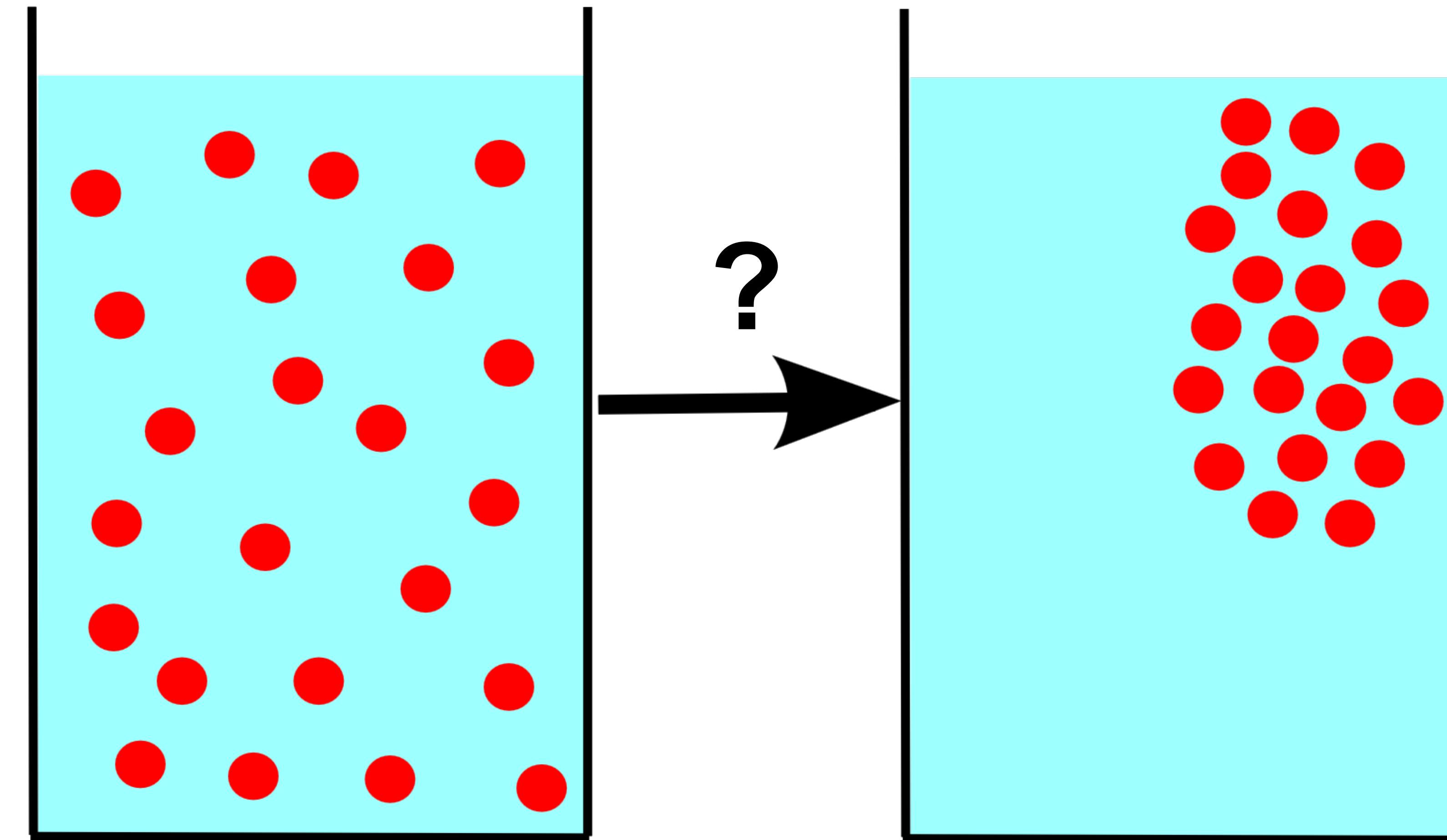


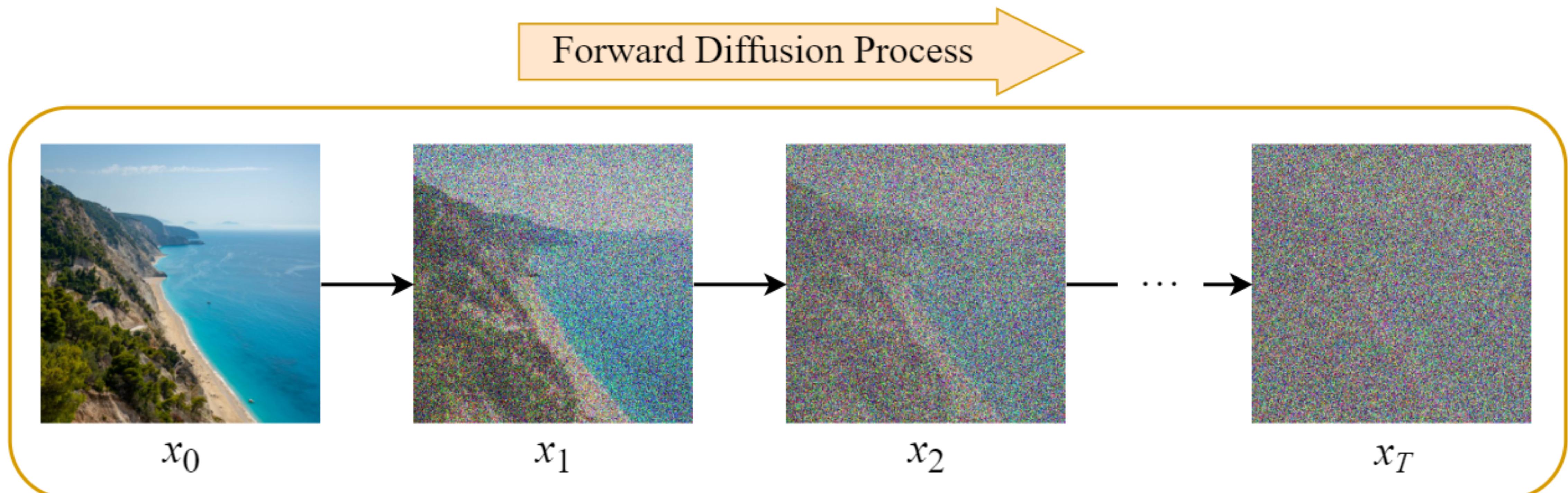
"Two cats doing research"

Diffusion in a liquid

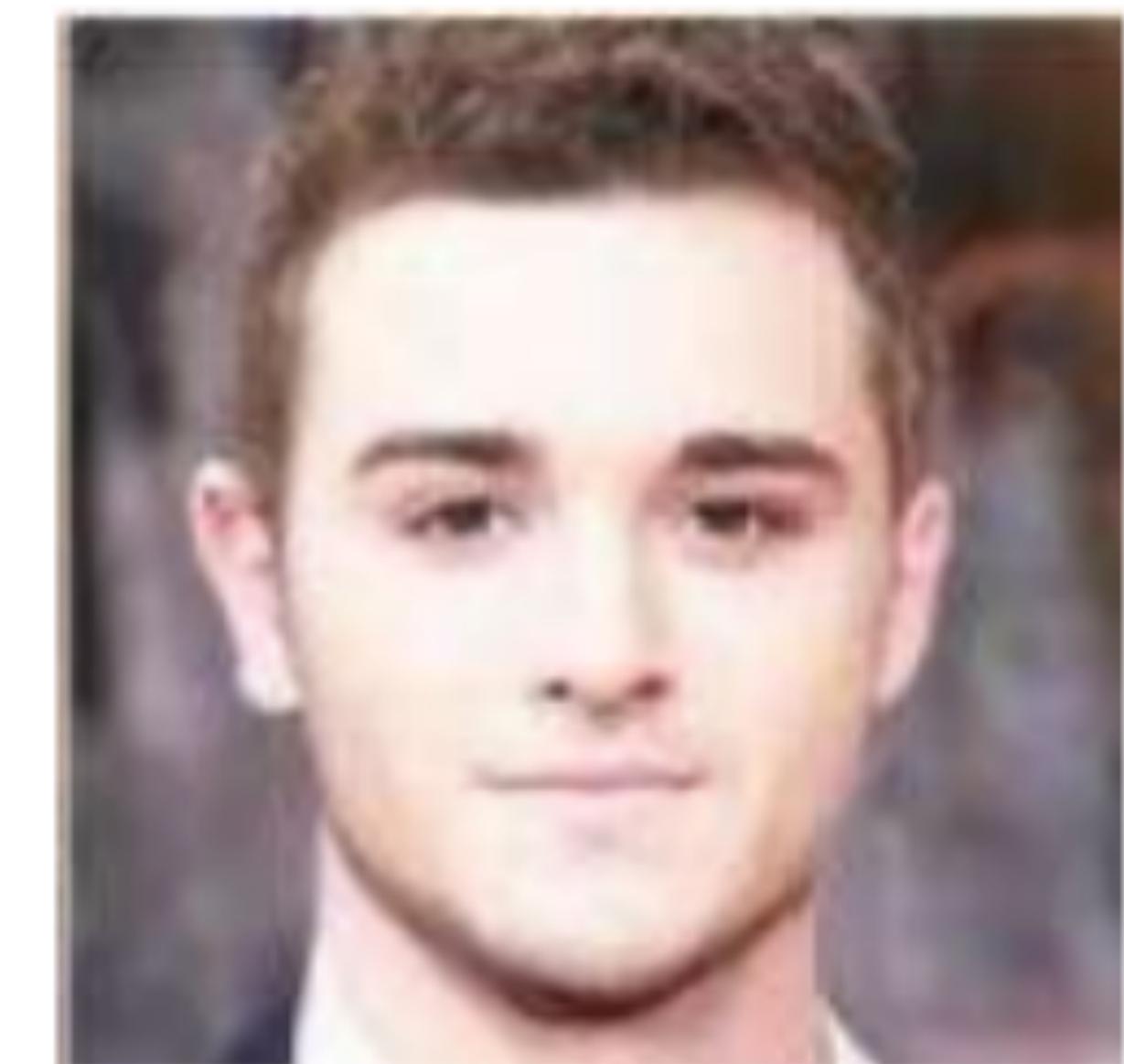


Is this possible?





Reversing the diffusion

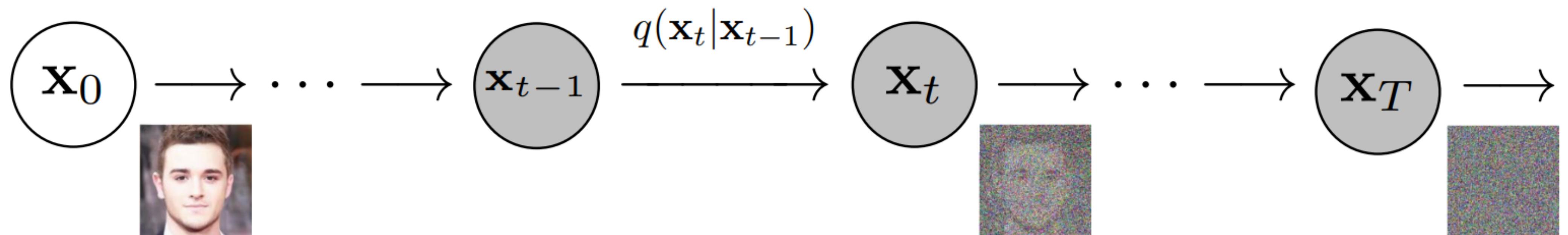


[https://www\[assemblyai.com/blog/diffusion-models-for-machine-learning-introduction](https://www[assemblyai.com/blog/diffusion-models-for-machine-learning-introduction)

Diffusion Models

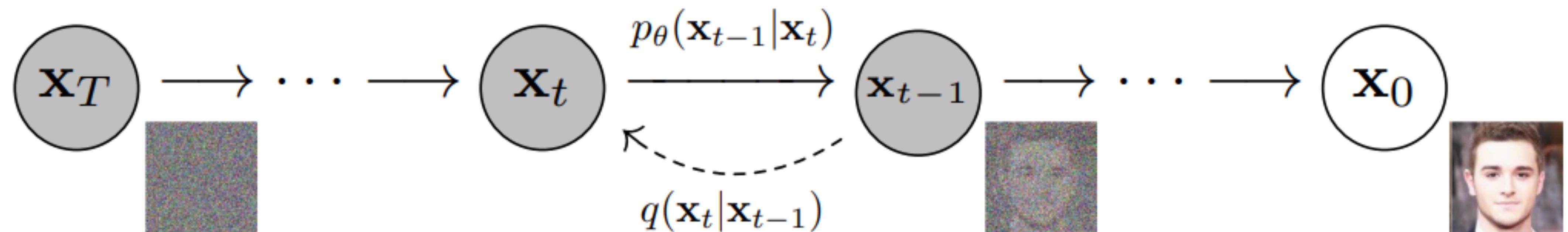
- forward → training / “noising”
 - add noise to training images in T steps
 - learn *the noise* at each step
- backward → sampling / “denoising”
 - start with random image
 - remove the noise at each step of T steps

Forward - training

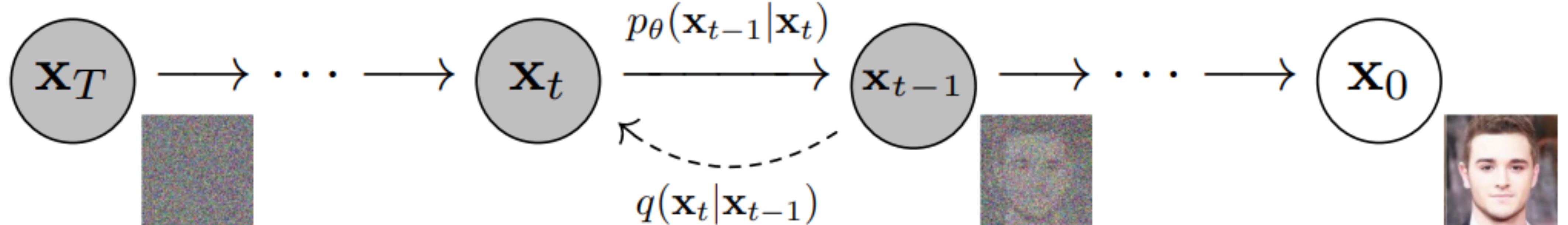
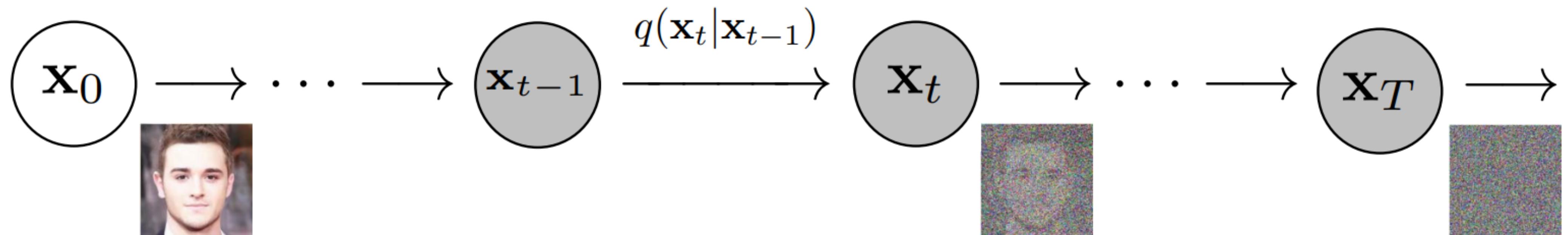


about 1.000 steps

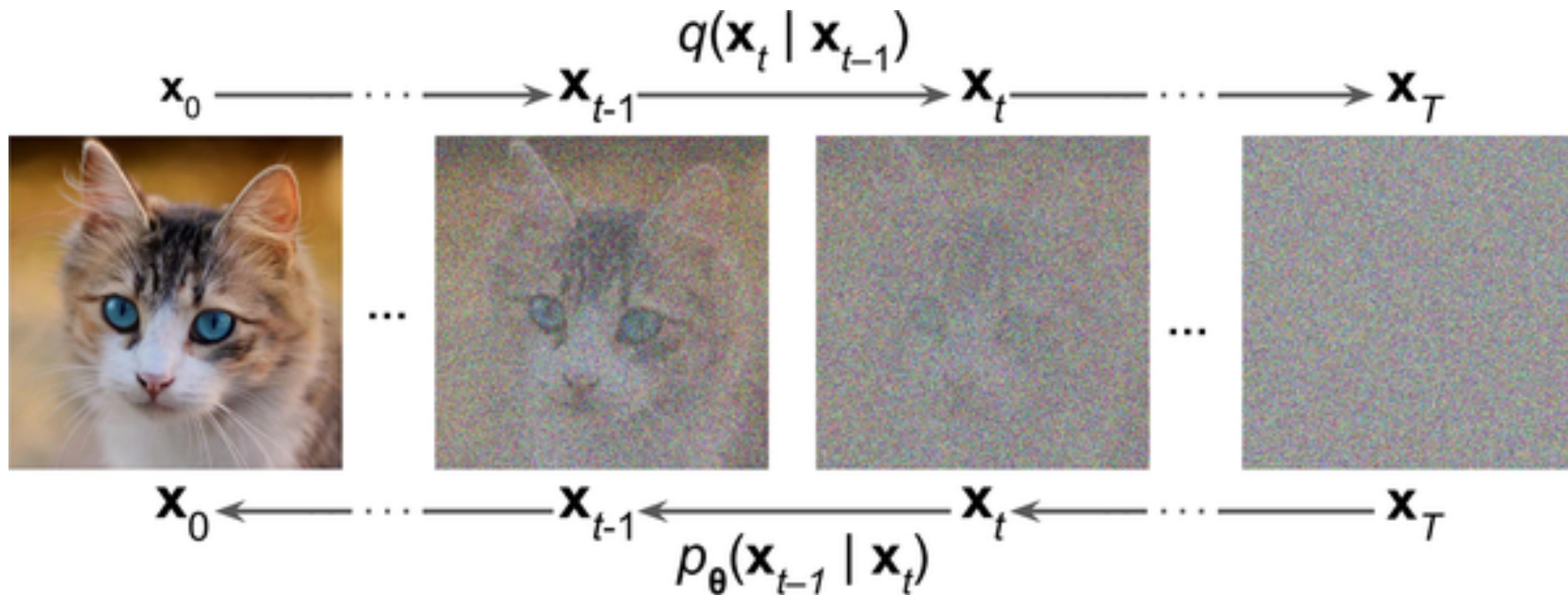
Backward - sampling



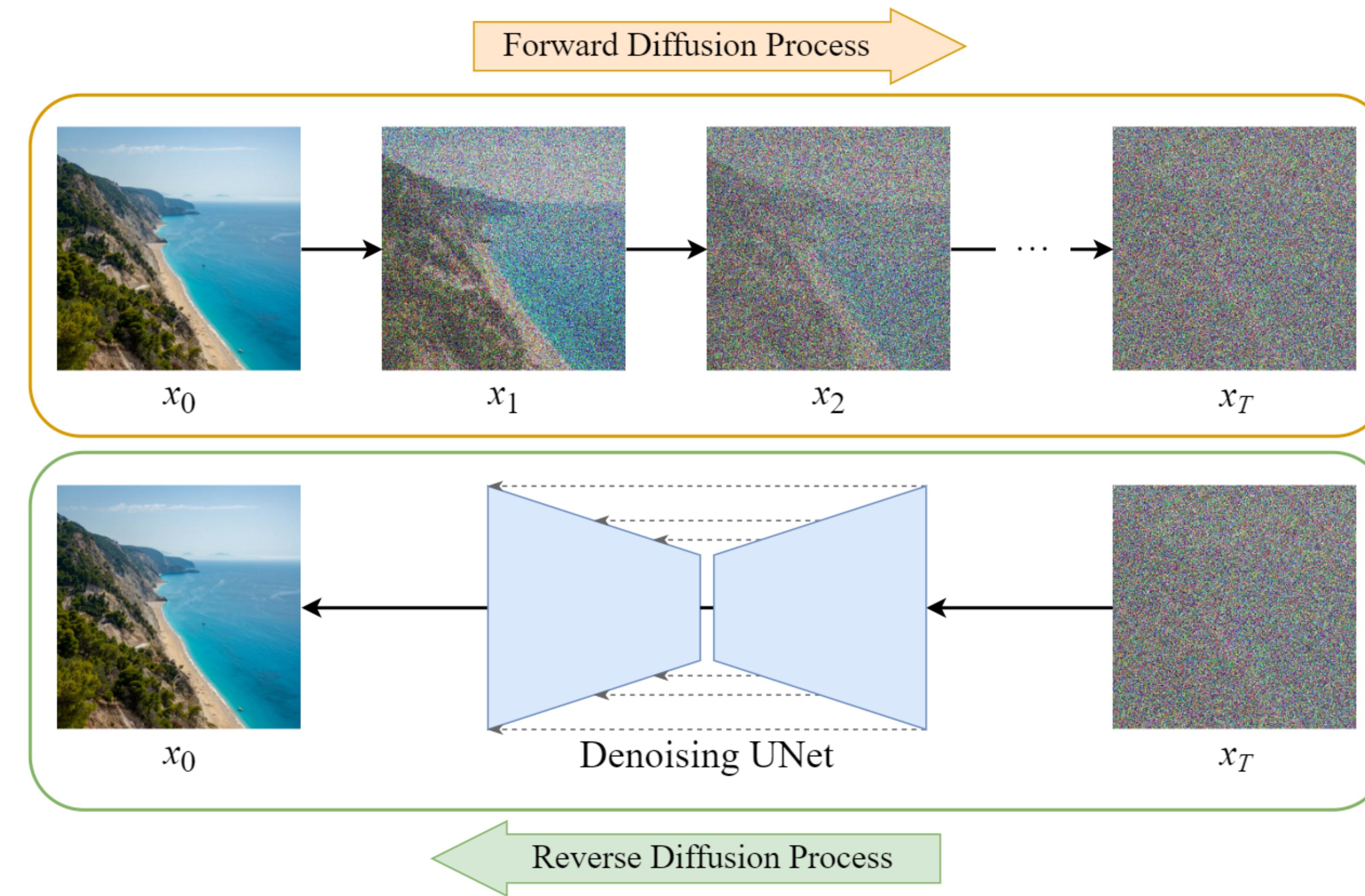
Noising and Denoising



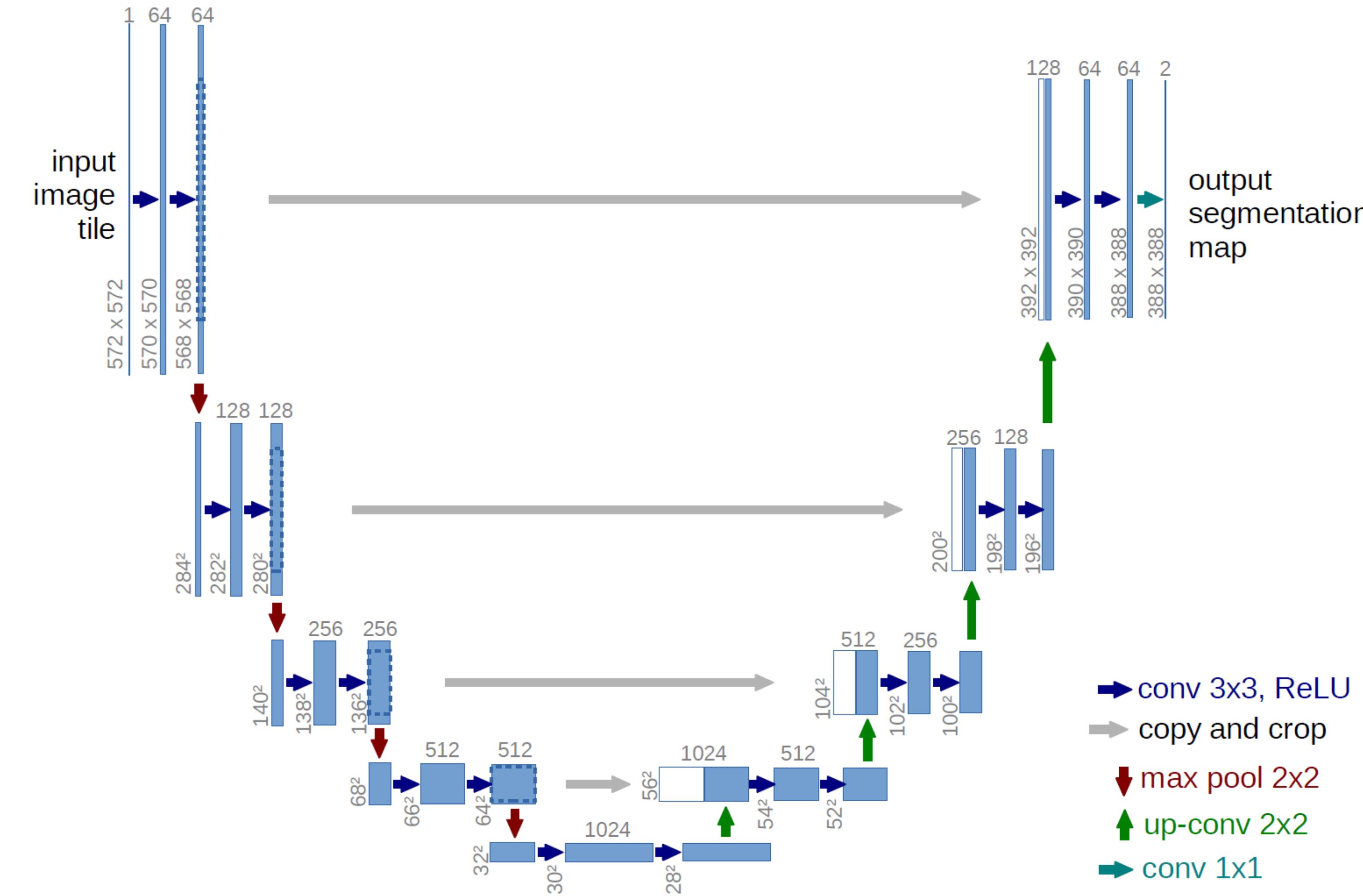
Reversing the diffusion



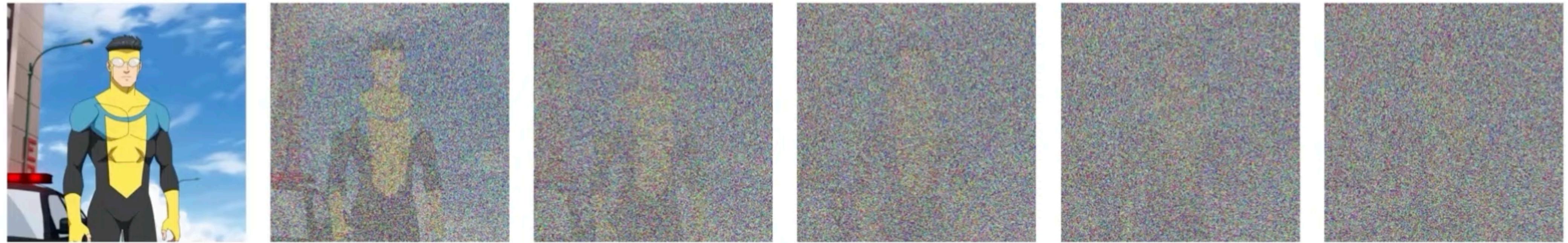
Reversing the diffusion



U-Net model



Denoising Diffusion Probabilistic Models (DDPM)



Algorithm 1 Training

```
1: repeat
2:    $\mathbf{x}_0 \sim q(\mathbf{x}_0)$ 
3:    $t \sim \text{Uniform}(\{1, \dots, T\})$ 
4:    $\boldsymbol{\epsilon} \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$ 
5:   Take gradient descent step on
      
$$\nabla_{\theta} \left\| \boldsymbol{\epsilon} - \boldsymbol{\epsilon}_{\theta}(\sqrt{\bar{\alpha}_t} \mathbf{x}_0 + \sqrt{1 - \bar{\alpha}_t} \boldsymbol{\epsilon}, t) \right\|^2$$

6: until converged
```

Algorithm 2 Sampling

```
1:  $\mathbf{x}_T \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$ 
2: for  $t = T, \dots, 1$  do
3:    $\mathbf{z} \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$  if  $t > 1$ , else  $\mathbf{z} = \mathbf{0}$ 
4:    $\mathbf{x}_{t-1} = \frac{1}{\sqrt{\alpha_t}} \left( \mathbf{x}_t - \frac{1 - \alpha_t}{\sqrt{1 - \bar{\alpha}_t}} \boldsymbol{\epsilon}_{\theta}(\mathbf{x}_t, t) \right) + \sigma_t \mathbf{z}$ 
5: end for
6: return  $\mathbf{x}_0$ 
```

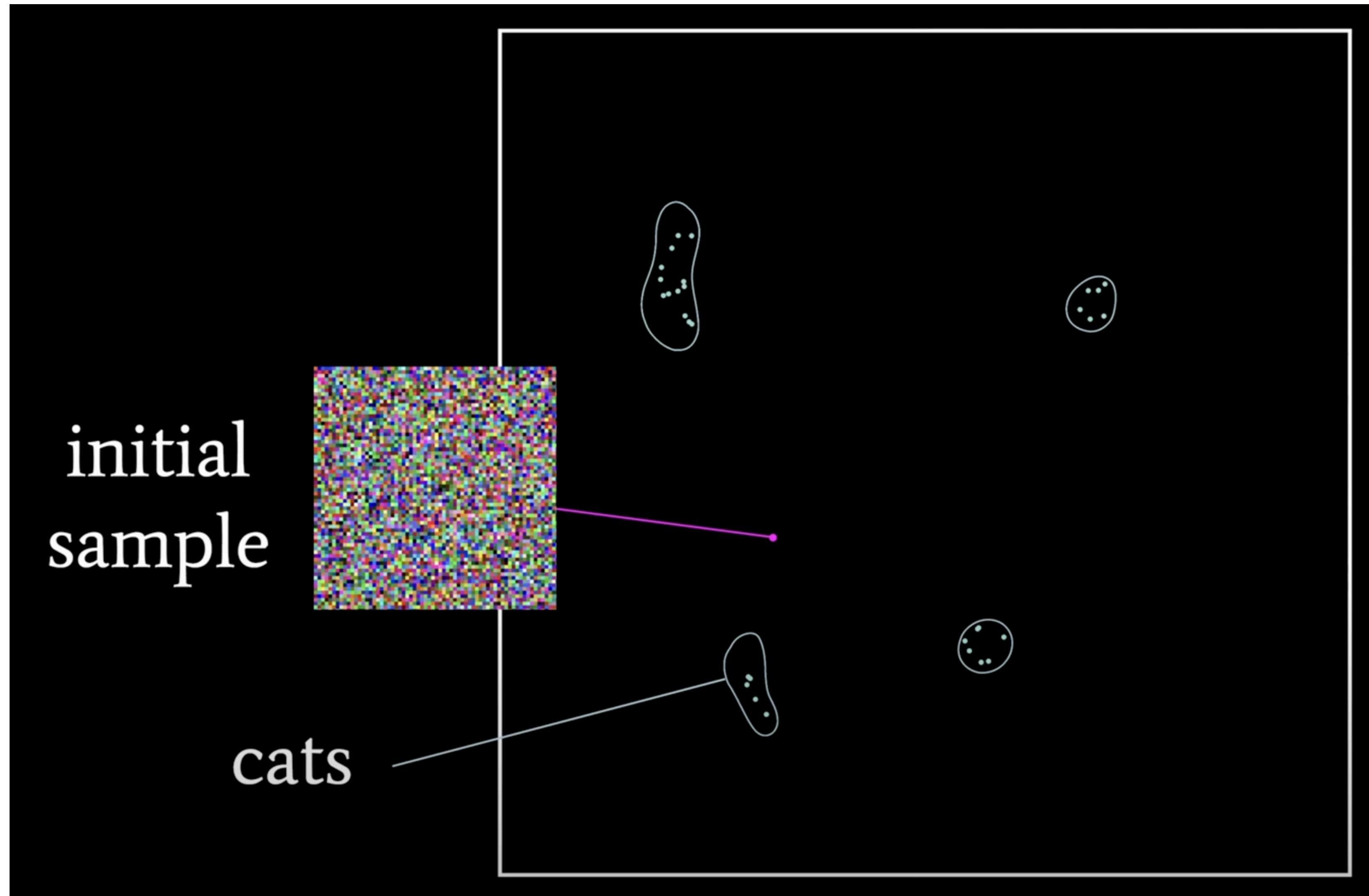
Do I understand these models?

- mechanics of the training and sampling
 - Yes, it is quite complicated but doable with some effort
- *why* does it work?
 - No, I do not understand it (yet)

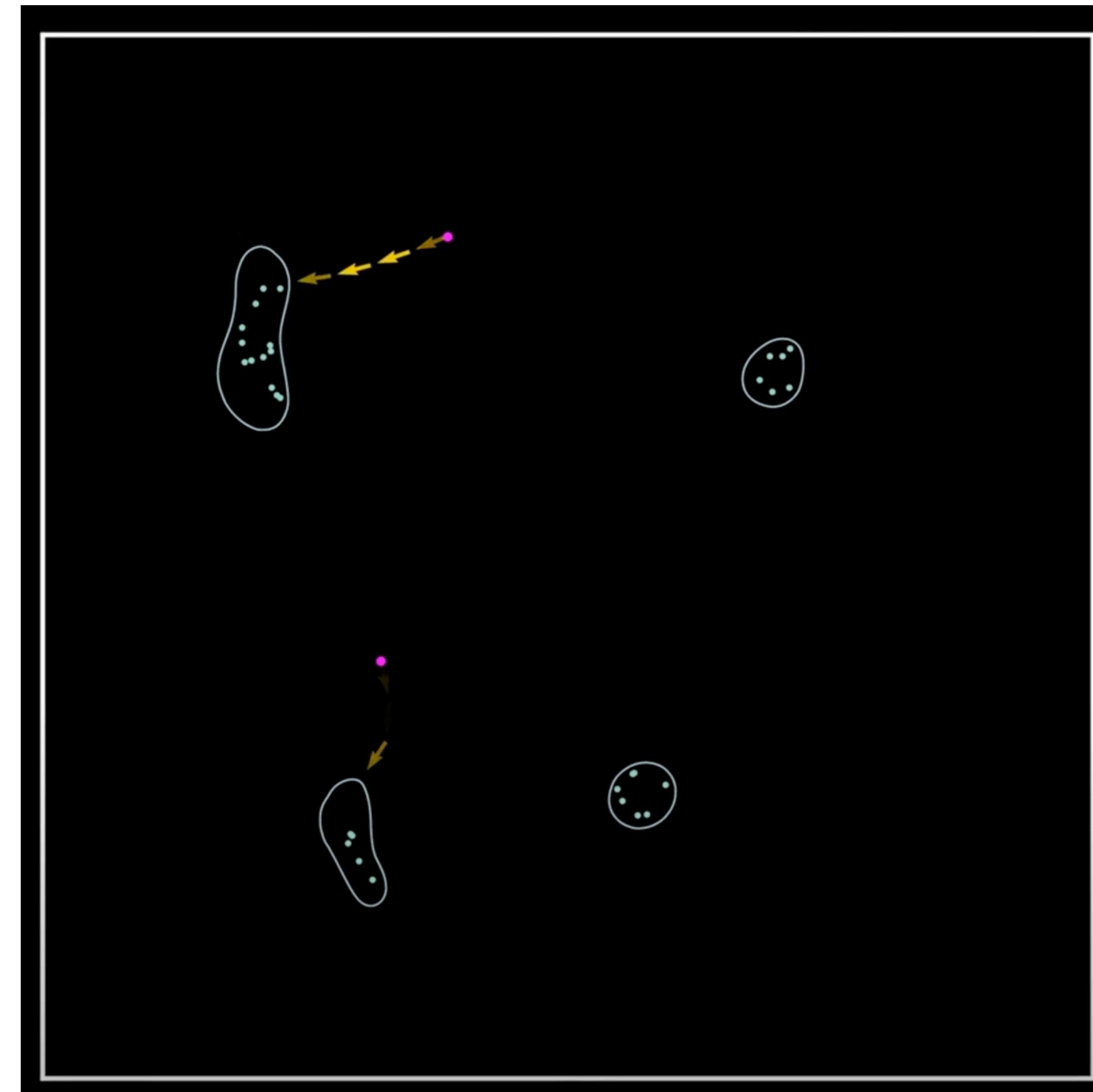
Why do diffusion models work?

- beginning of an answer:
 - the space of possible images is *huge*
 - the training images create clusters of images in an almost empty space
 - <https://youtu.be/1pgiu--4W3I?si=xkgTgOPCOmLyYrPT>

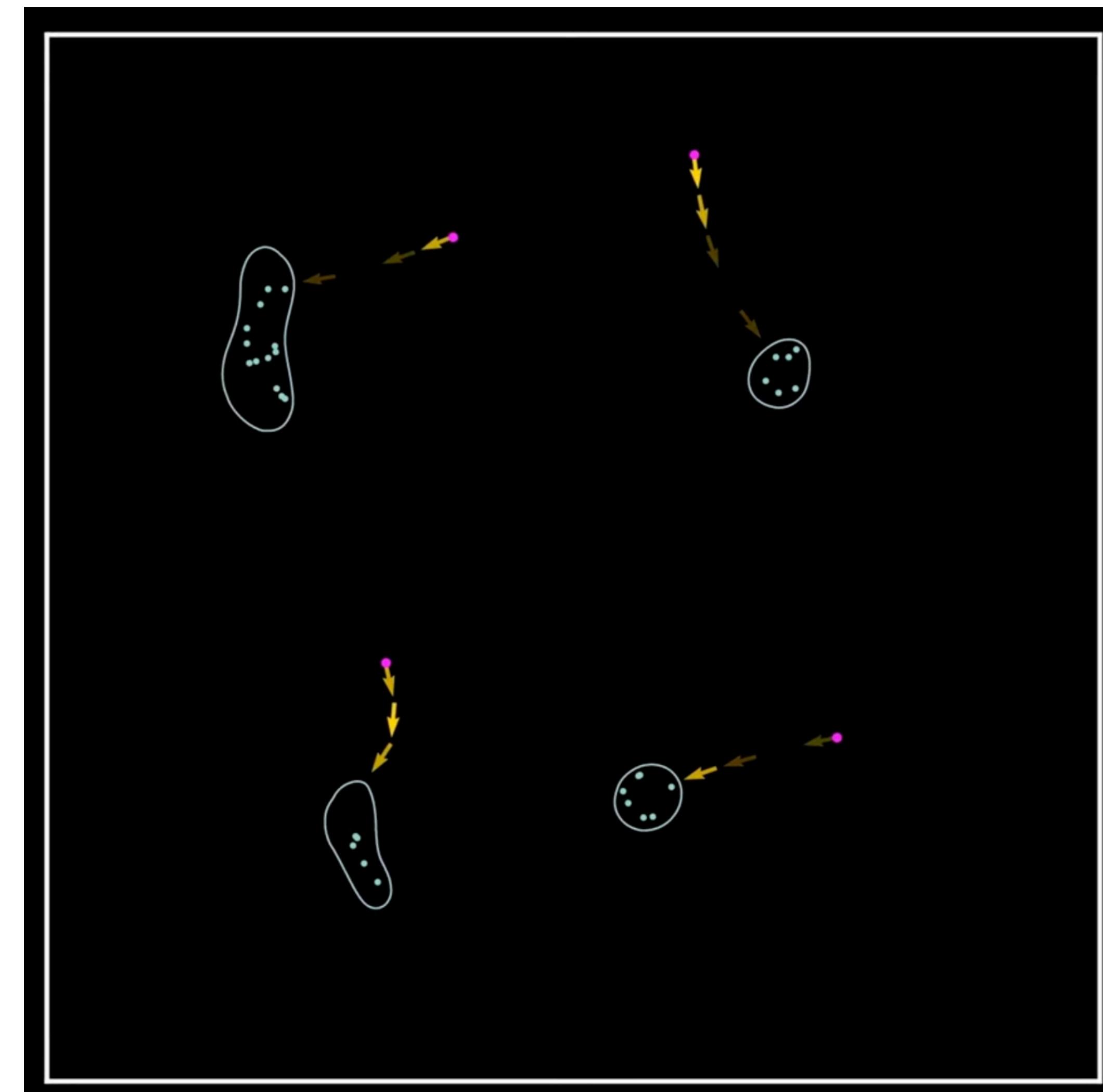
Why do diffusion models work?



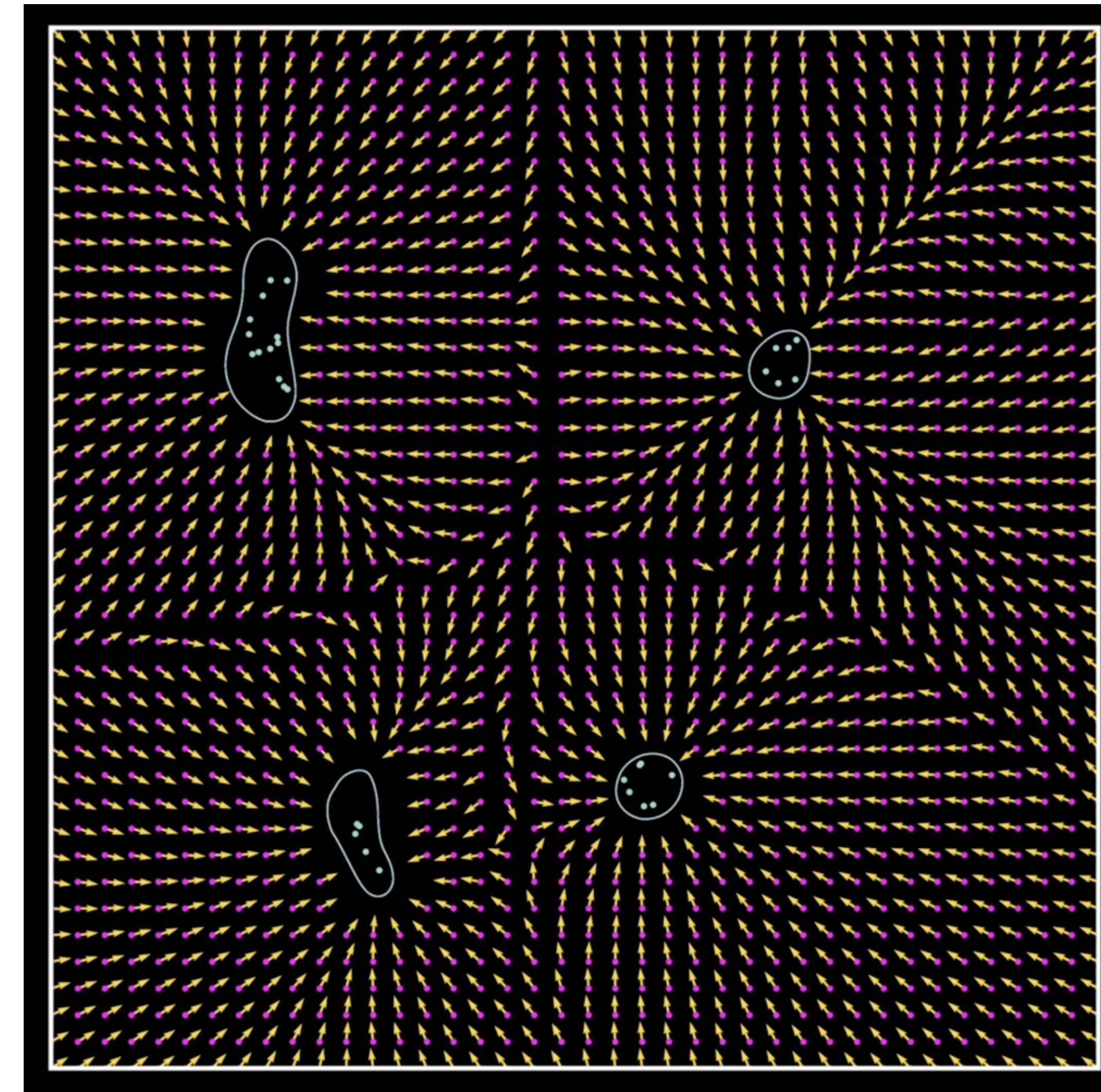
Why do diffusion models work?



Why do diffusion models work?



Why do diffusion models work?



Why do diffusion models work?

- beginning of an answer:
 - the space of possible images is *huge*
 - the training images create clusters of images in an almost empty space
 - <https://youtu.be/1pgiu--4W3I?si=xkgTgOPCOmLyYrPT>

Notebook

MNIST Fashion dataset

- Hands-on Machine Learning with Scikit-Learn, Keras and TensorFlow (3rd edition), Aurélien Géron
- https://github.com/ageron/handson-ml3/blob/main/17_autoencoders_gans_and_diffusion_models.ipynb



12 epochs



24 epochs



50 epochs



inholland
hogeschool