

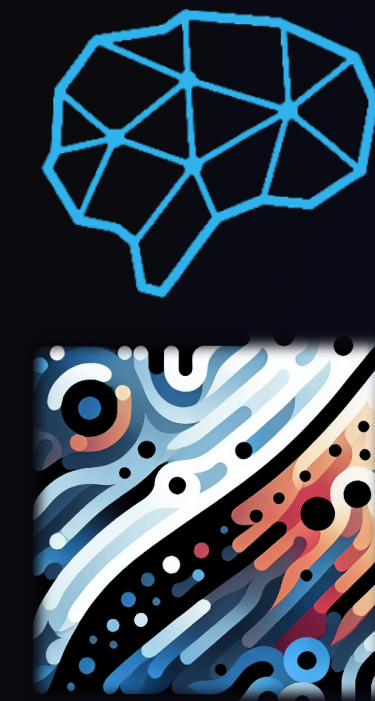
Diffusion Model

Cogito Project 2024 Spring

Marijan, Thomas, Mauritz, Amanda and Baris

30/04/2024

Introduction



GANs

Diffusion
Models



2014

2021

The CLIP

 DALL·E

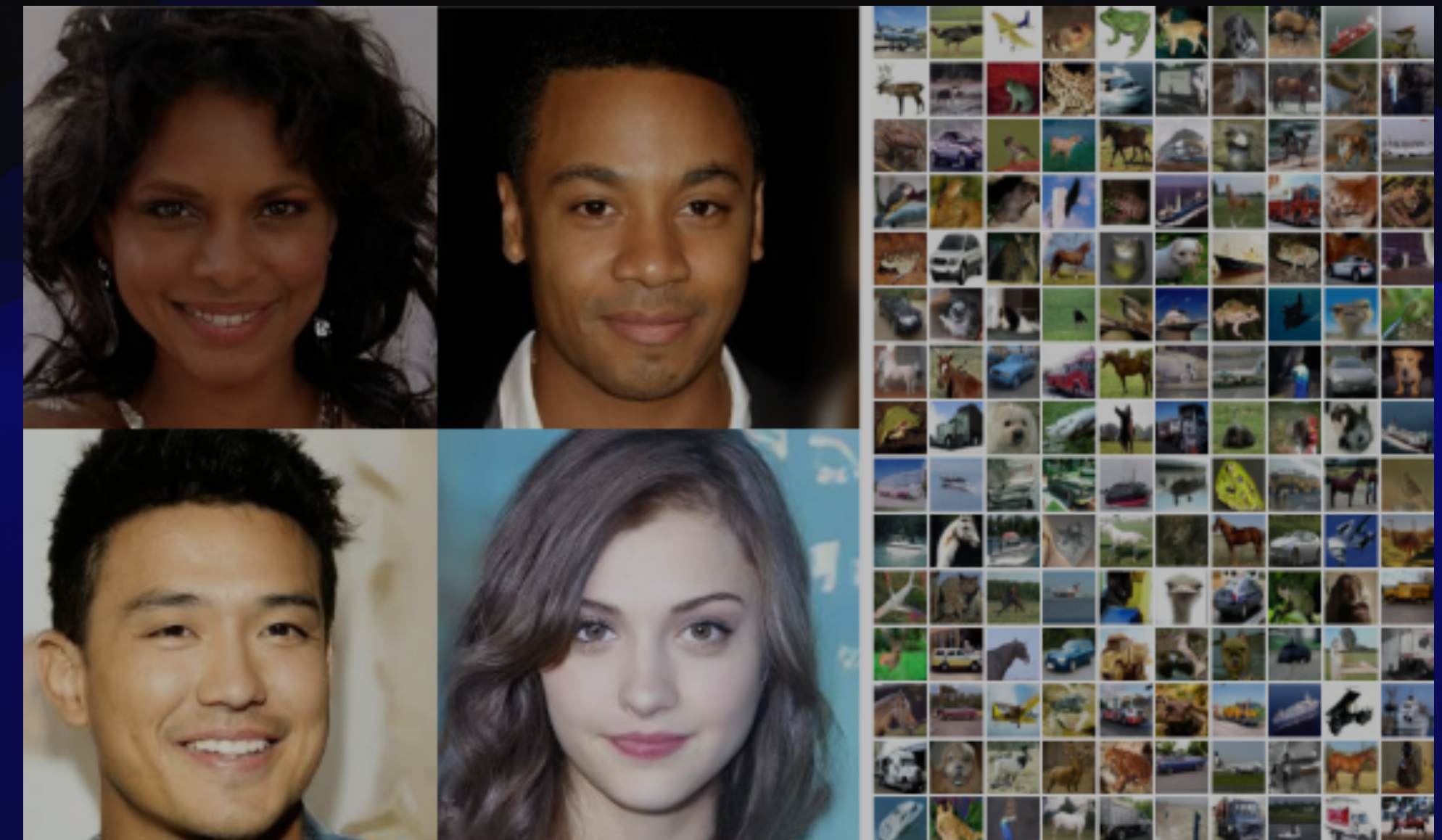


S.

Background

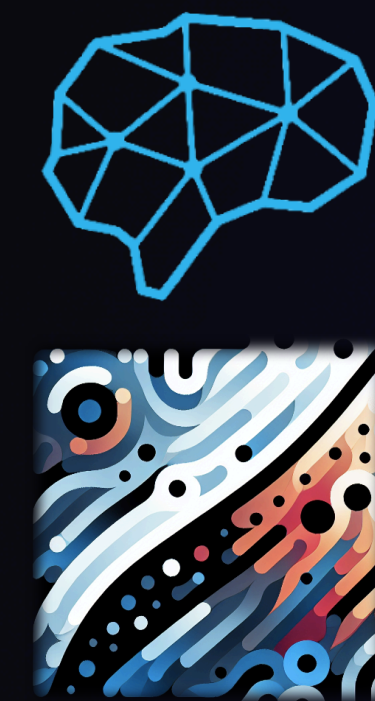


- Team
- Workflow
- “*Denoising Diffusion Probabilistic Models*” by Jonathan Ho, Ajay Jain and Pieter Abbeel



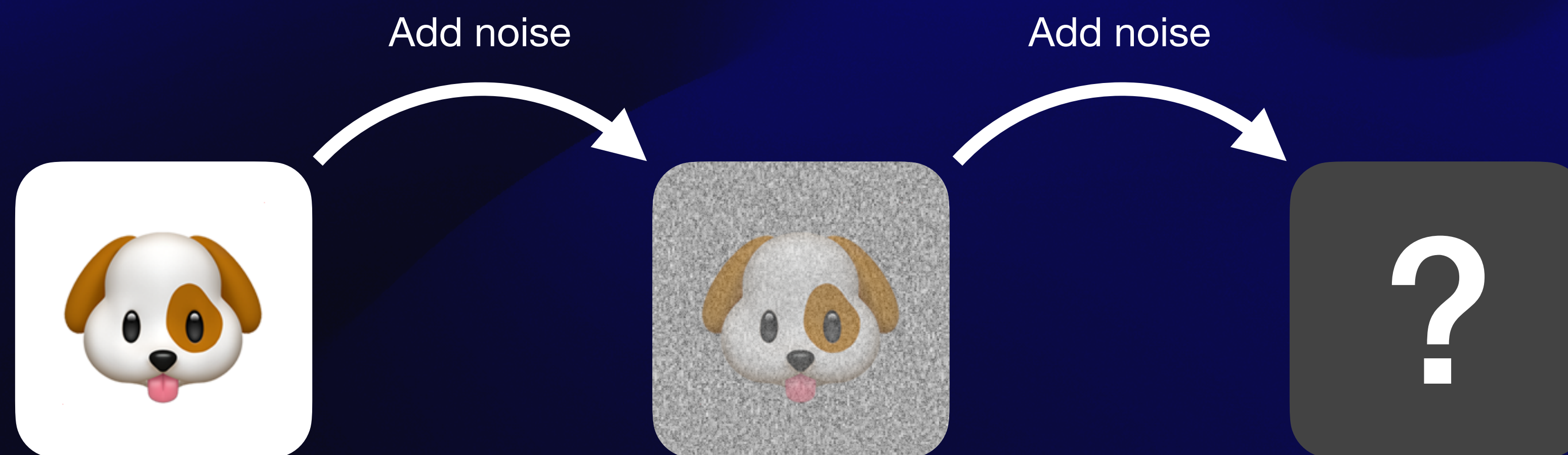
$$q(X_1, \dots, X_T | X_0) \triangleq \prod_{t=1}^T q(X_t | X_{t-1}) \quad q(X_t | X_{t-1}) \sim \mathcal{N}(\sqrt{1 - \beta_t} X_{t-1}, \beta_t I_n)$$

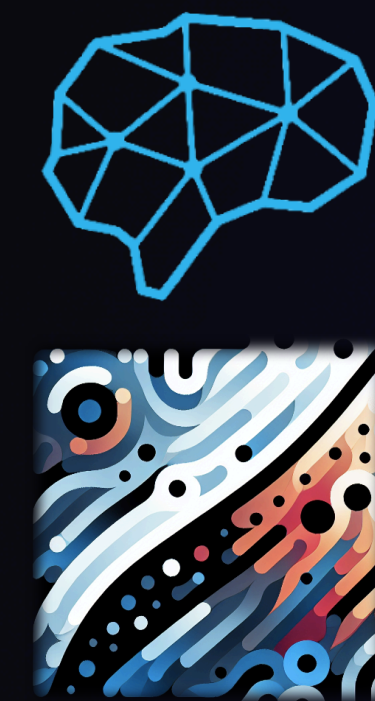
$$p_\theta(X_0, \dots, X_T) \triangleq \prod_{t=1}^T p_\theta(X_{t-1} | X_t) \quad p_\theta(X_{t-1} | X_t) \sim \mathcal{N}(\mu_\theta(X_t, t), \Sigma_\theta(X_t, t))$$



Forward Process

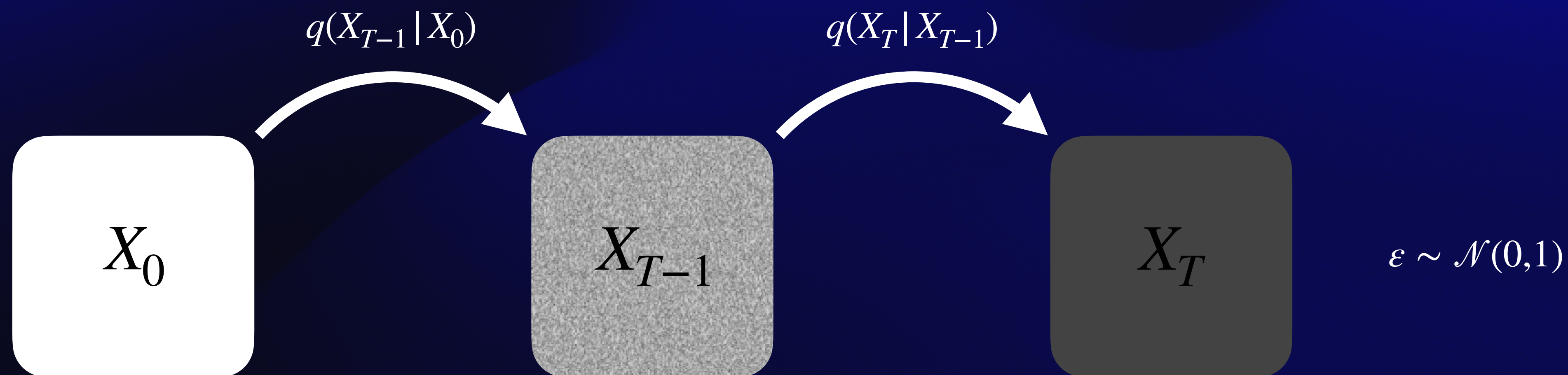
- Gradually adds Gaussian noise at each time step
- Markov chain: memory-less progression
- Generated states provide training data for the reverse process

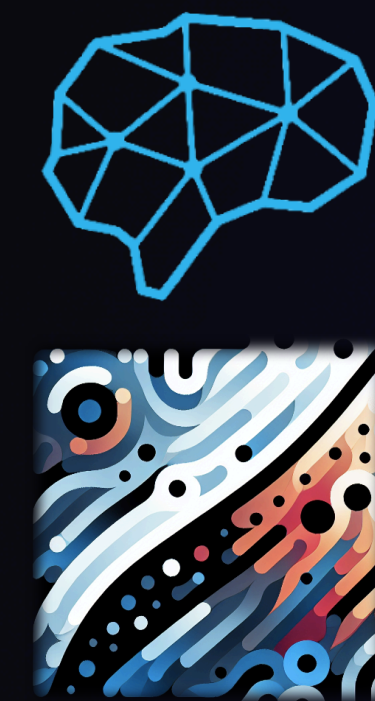




Forward Process

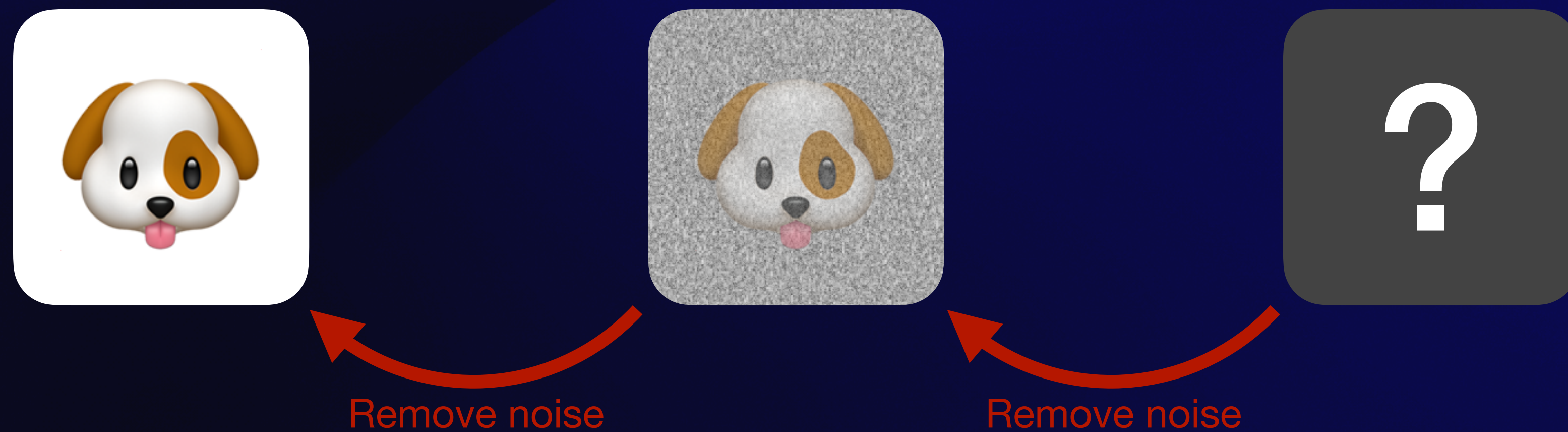
- Gradually adds Gaussian noise at each time step
- Markov chain: memory-less progression
- Generated states provide training data for the reverse process

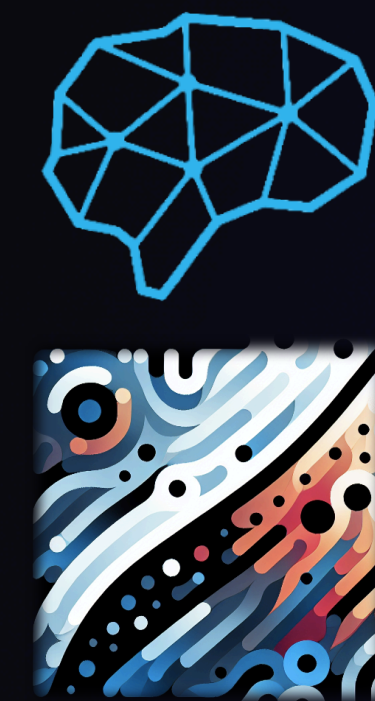




Backward Process

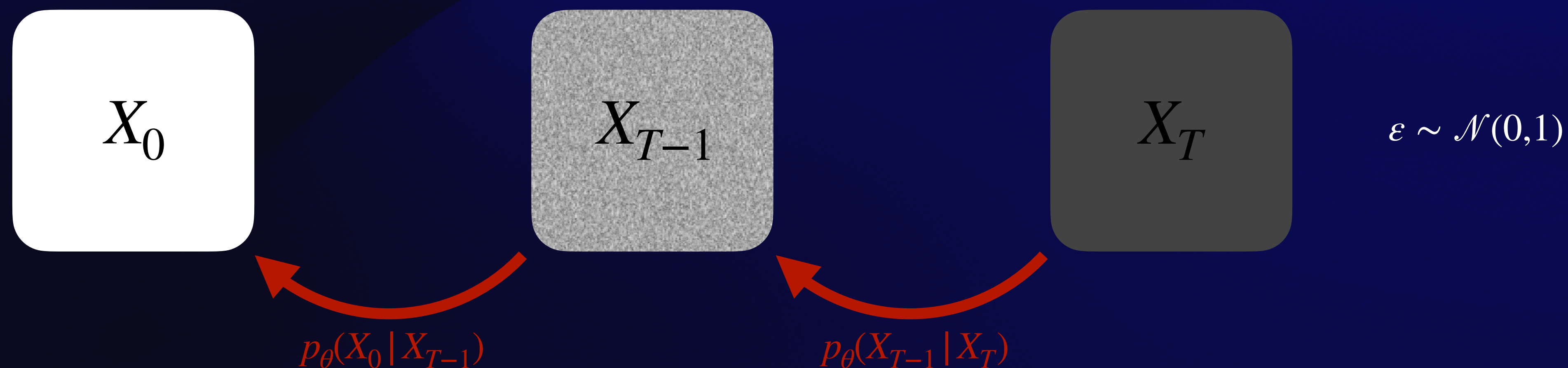
- Inverse of forward process
- Removes noise from image
- Predicts X_{t-1} by knowing X_t





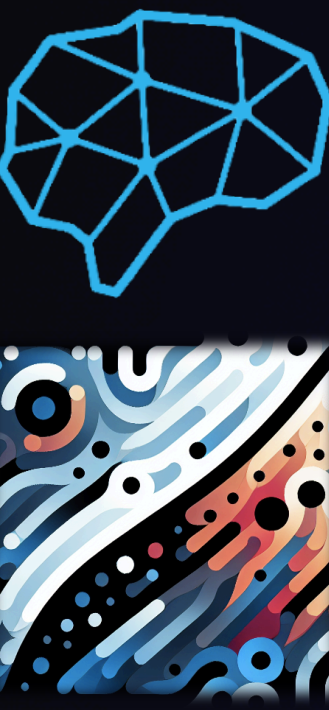
Backward Process

- Inverse of forward process
- Removes noise from image
- Predicts X_{t-1} by knowing X_t



Model Overview

Sum up

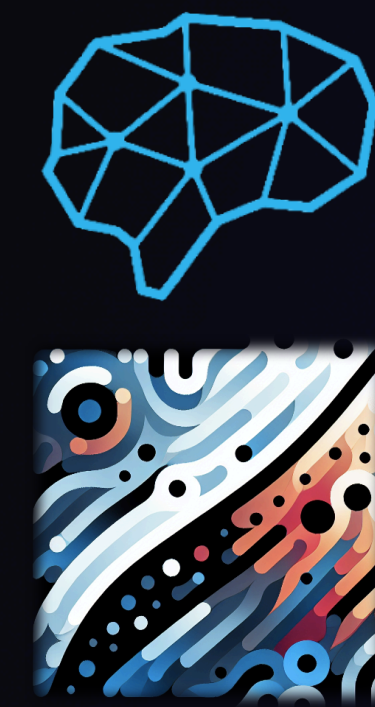


$t=0$

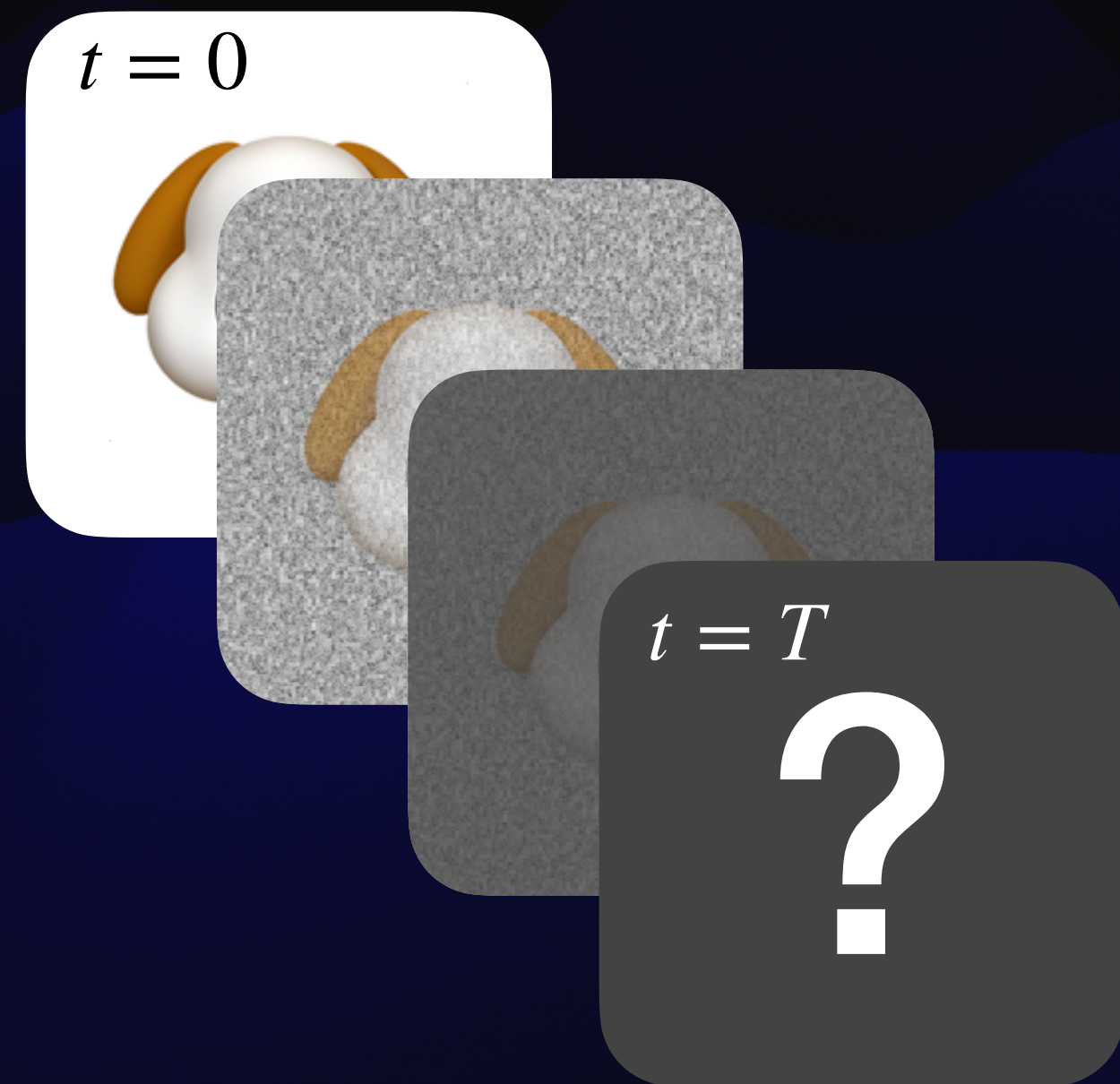


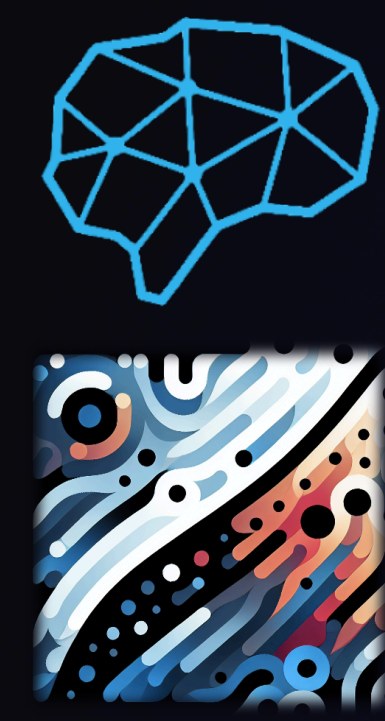
Model Overview

Data “augmentation” with Forward Process



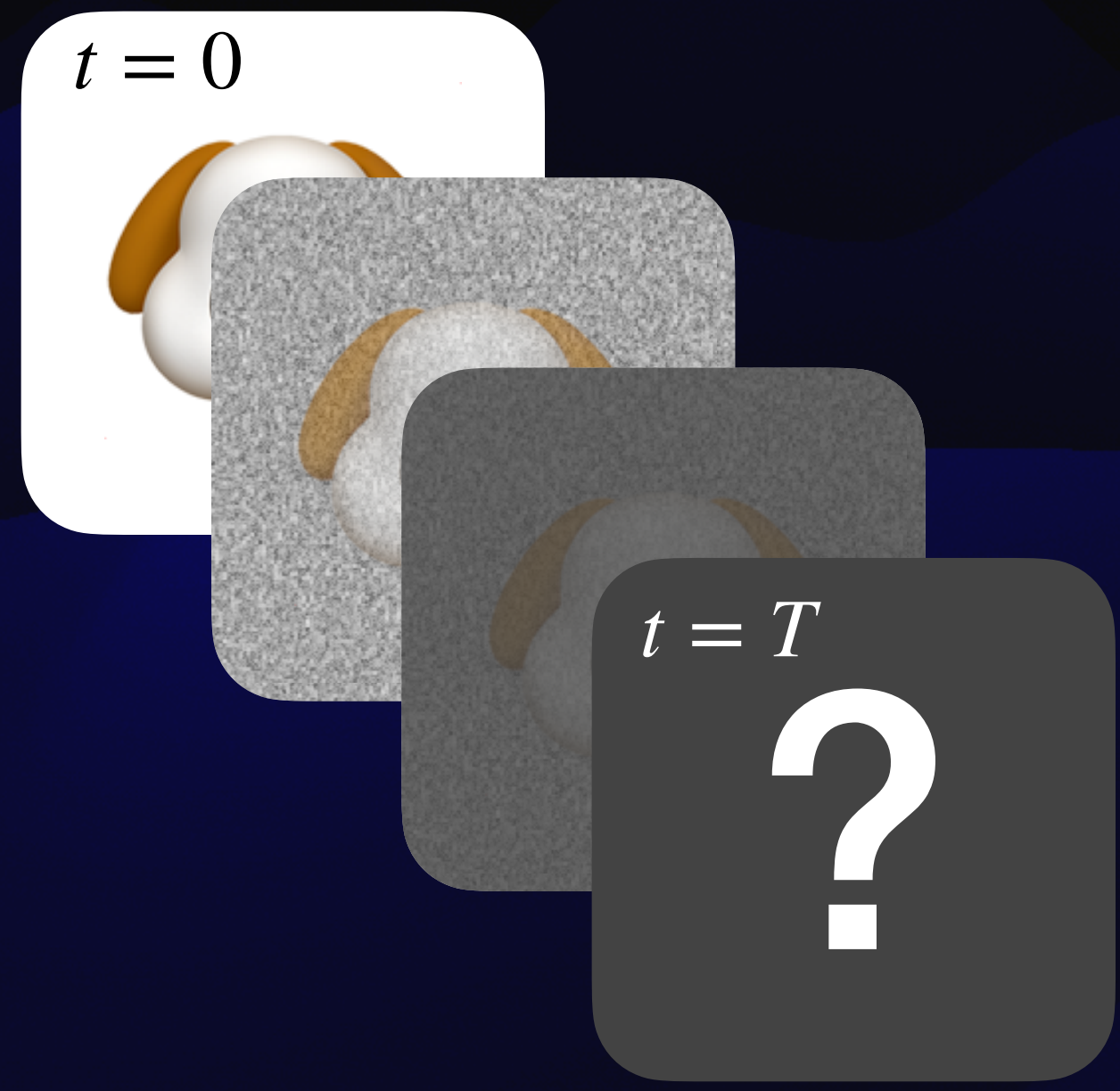
$t = 0$





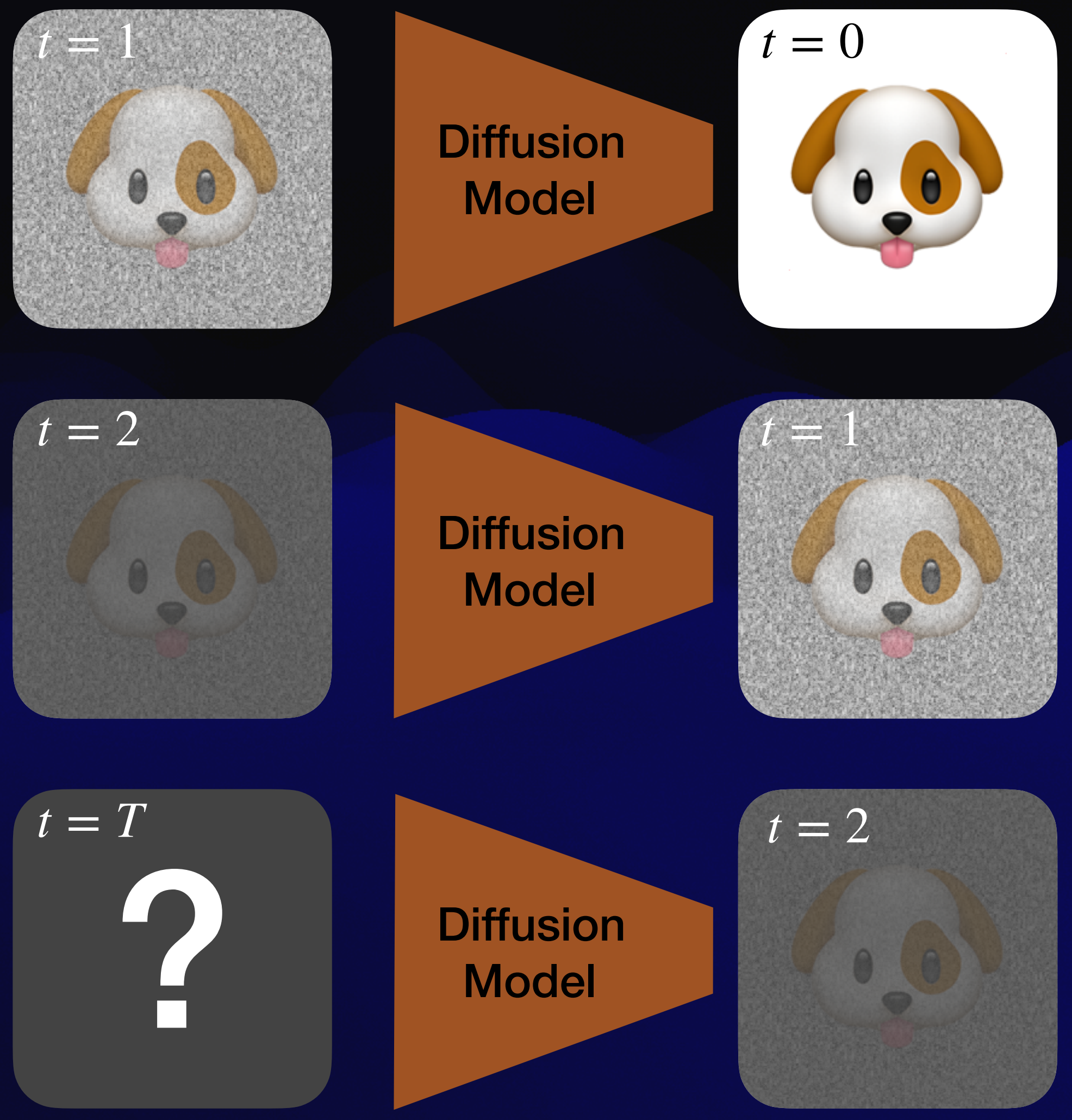
Model Overview

Training process



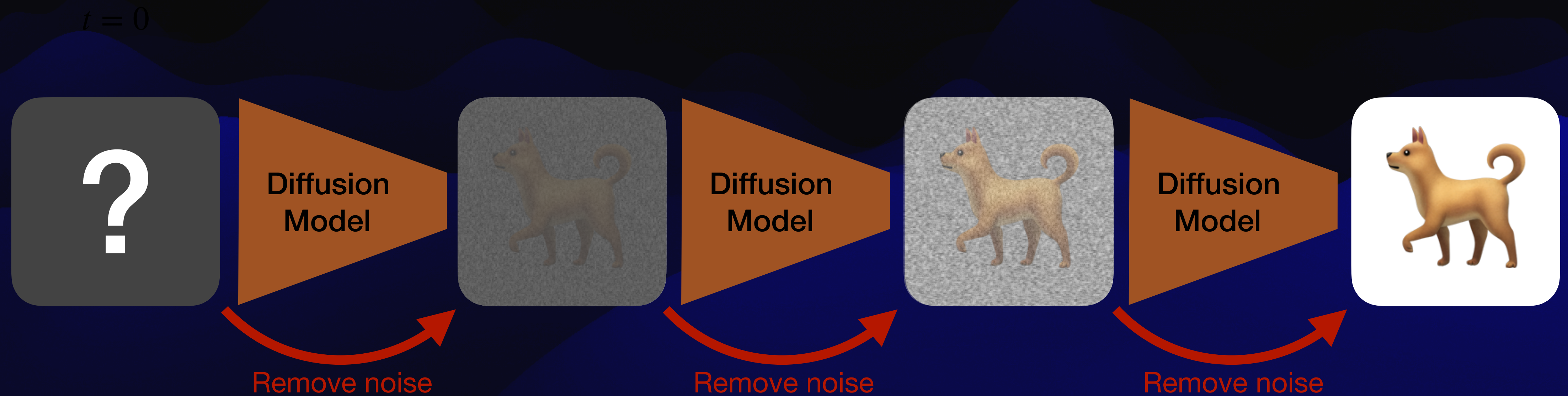
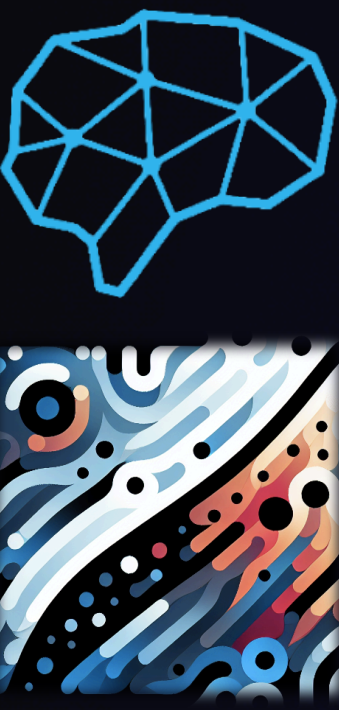
$$(X_t, t) \mapsto X_{t-1}$$

(Input, Output)

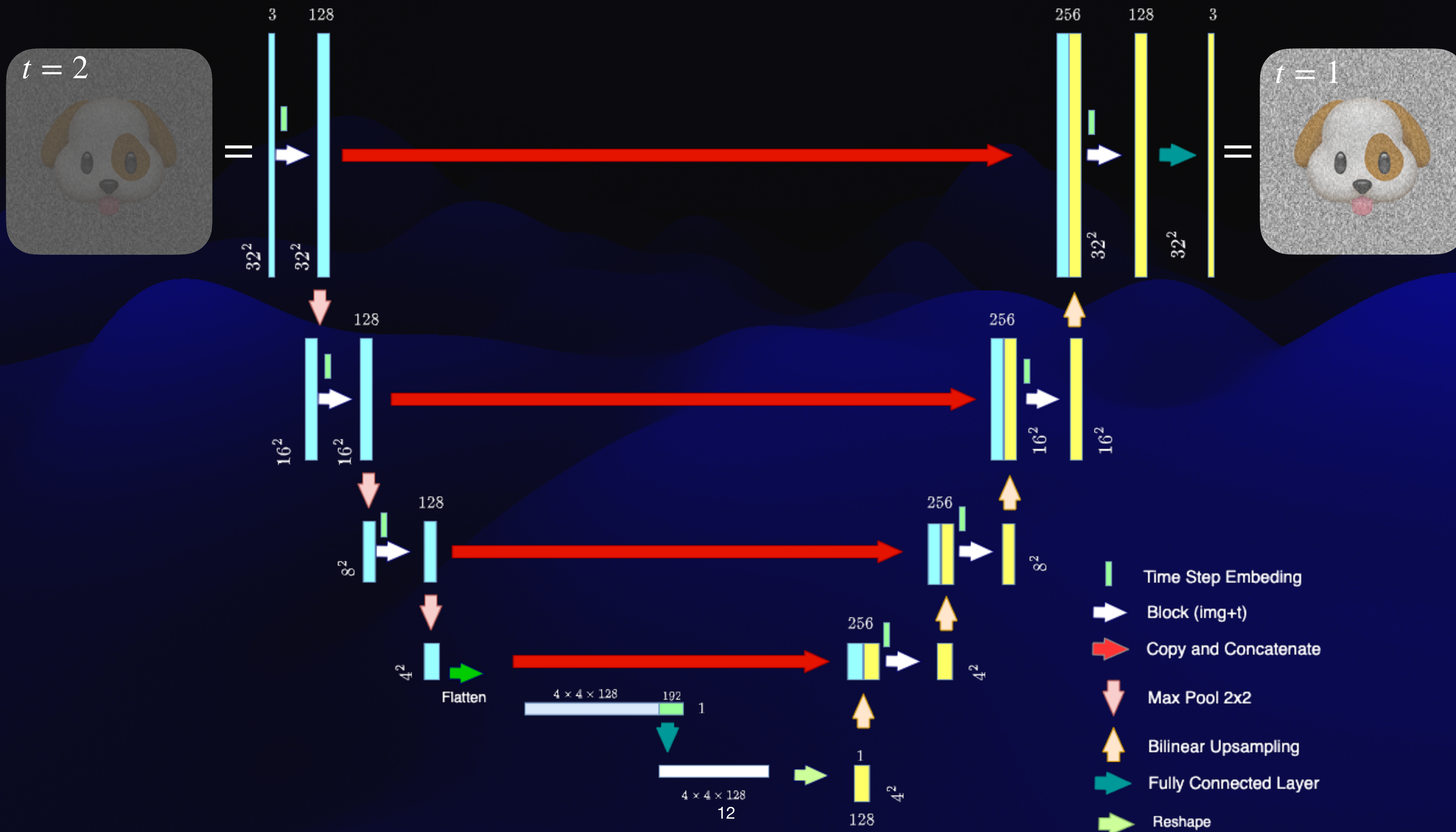
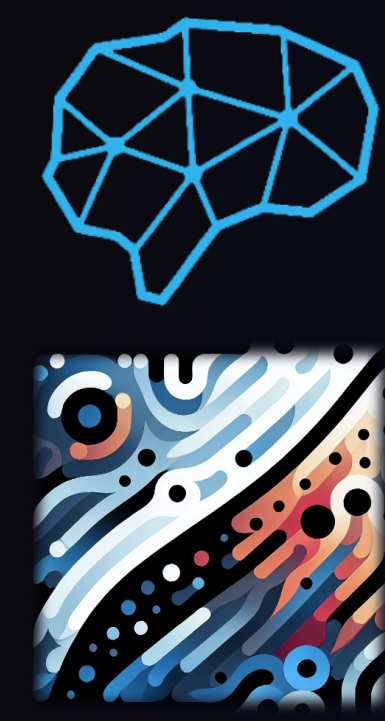


Model Overview

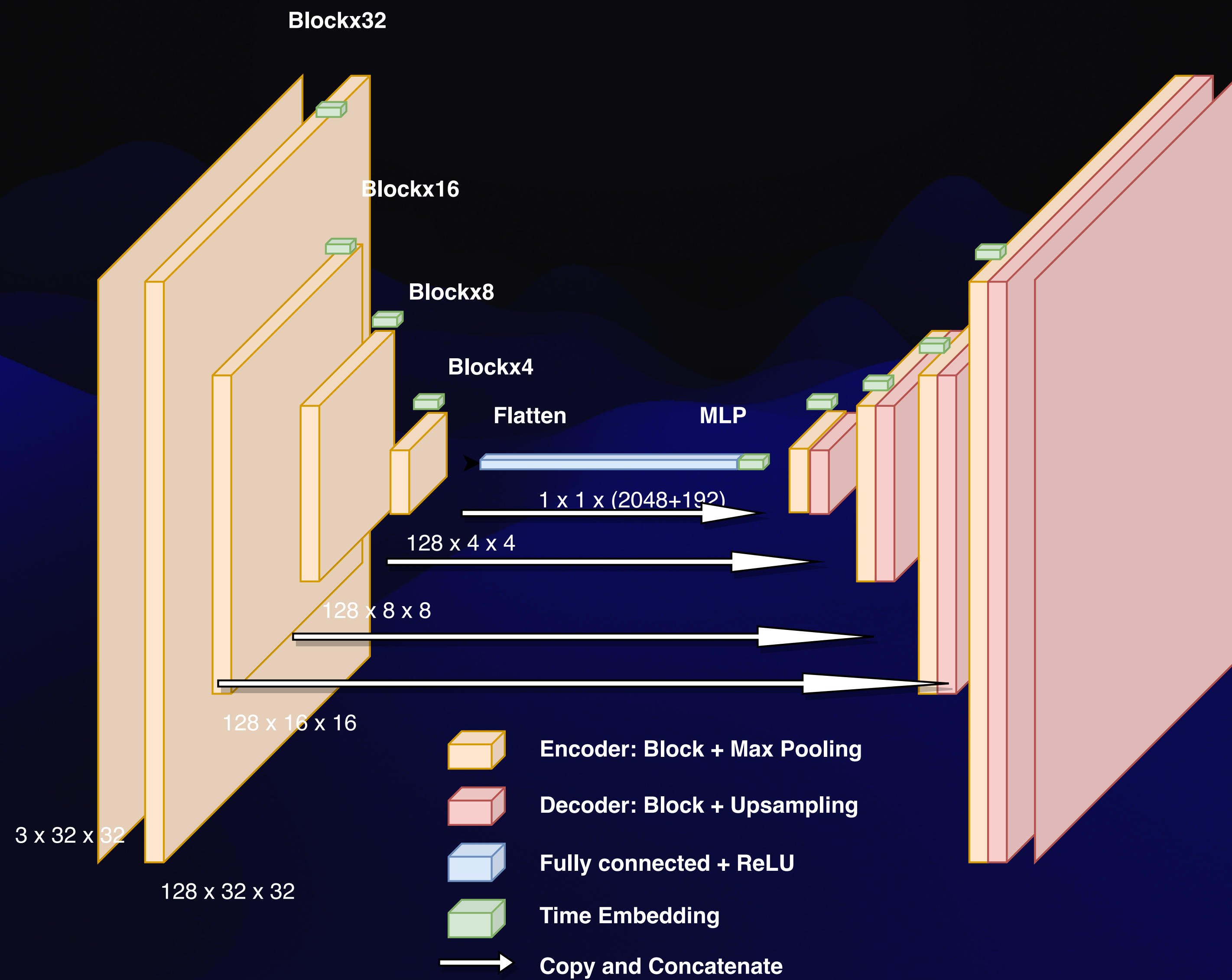
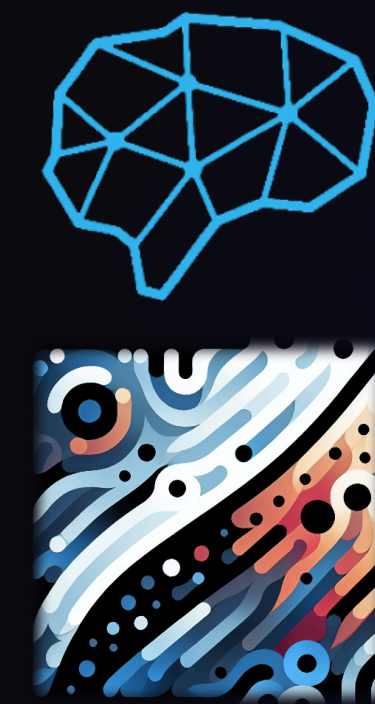
Inference: Image generator



Implementation



Implementation



Conclusion



- **Our model**
 - Simplification of the model
 - Low resolution image 32x32
 - Training takes 1 hour (GPU)
 - Available on:
 - Hugging Face: “**Cogito-Diffusion-Model**”
 - *Soon on Cogito’s website*
- **To improve**
 - Embedding time step
 - Attention mechanism
 - Fixed variance in forward process
 - Conditional generative (text-to-image)

Diffusion Model

<https://huggingface.co/spaces/marijanic/Cogito-Diffusion-Model>