Introduction to Deep Learning for Computer Vision

Adhyayan '23 - ACA Summer School
Department of Computer Science and Engineering
Indian Institute of Technology Kanpur

Lecture 7

Generative Adversarial Network (GAN)

Generative Adversarial Networks

Generative Adversarial Networks

Which one is fake?

GANs: What if we just want to sample?

GANs: What if we just want to sample?

Idea: Don't explicitly model density, instead just sample to generate new instances.

GANs: What if we just want to sample?

Idea: Don't explicitly model density, instead just sample to generate new instances.

Problem: Want to sample from a complex distribution - can't do this directly!

GANs: What if we just want to sample?

Idea: Don't explicitly model density, instead just sample to generate new instances.

Problem: Want to sample from a complex distribution - can't do this directly!

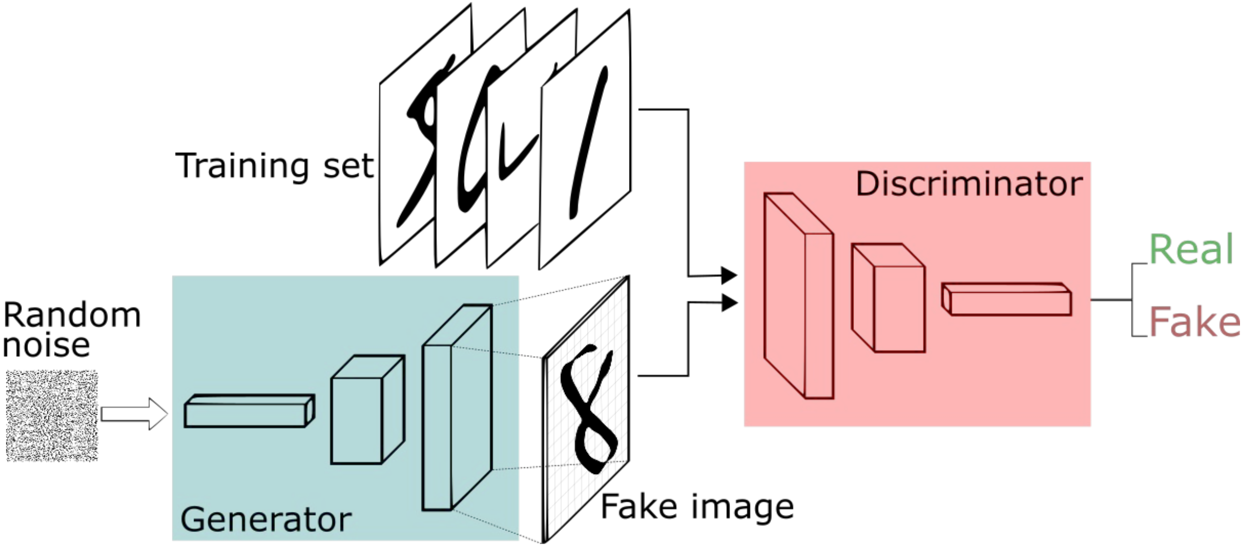
Solution: sample from something simple. learn a transformation to the data distribution.

GANs: What if we just want to sample?

Idea: Don't explicitly model density, instead just sample to generate new instances.

Problem: Want to sample from a complex distribution - can't do this directly!

Solution: sample from something simple. learn a transformation to the data distribution.

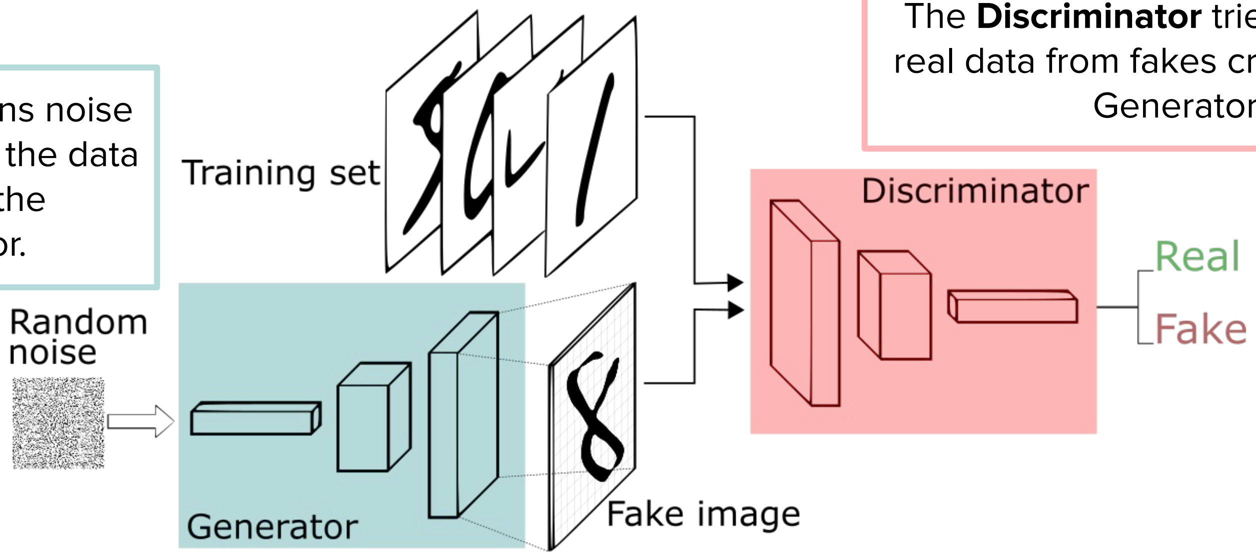


Generative Adversarial Networks (GANs)

GANs are a way to make a generative model by making two neural networks compete with each other.

The **Discriminator** tries to identify real data from fakes created by the Generator.

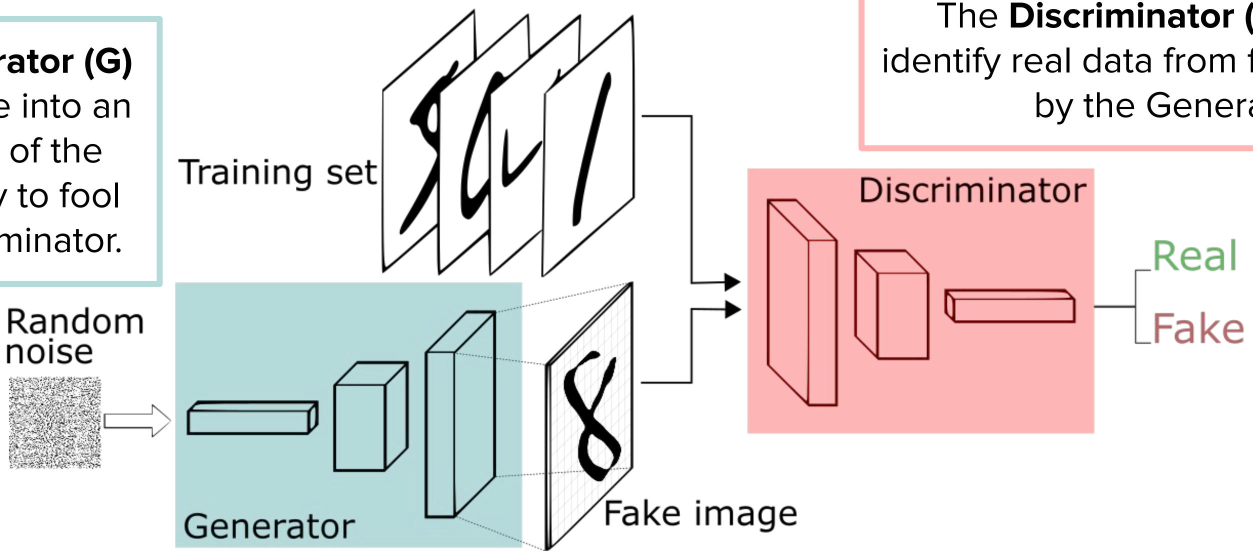
The **Generator** turns noise into an imitation of the data to try to fool the Discriminator.



Training GANs

The **Generator (G)** turns noise into an imitation of the data to try to fool the Discriminator.

The **Discriminator (D)** tries to identify real data from fakes created by the Generator.

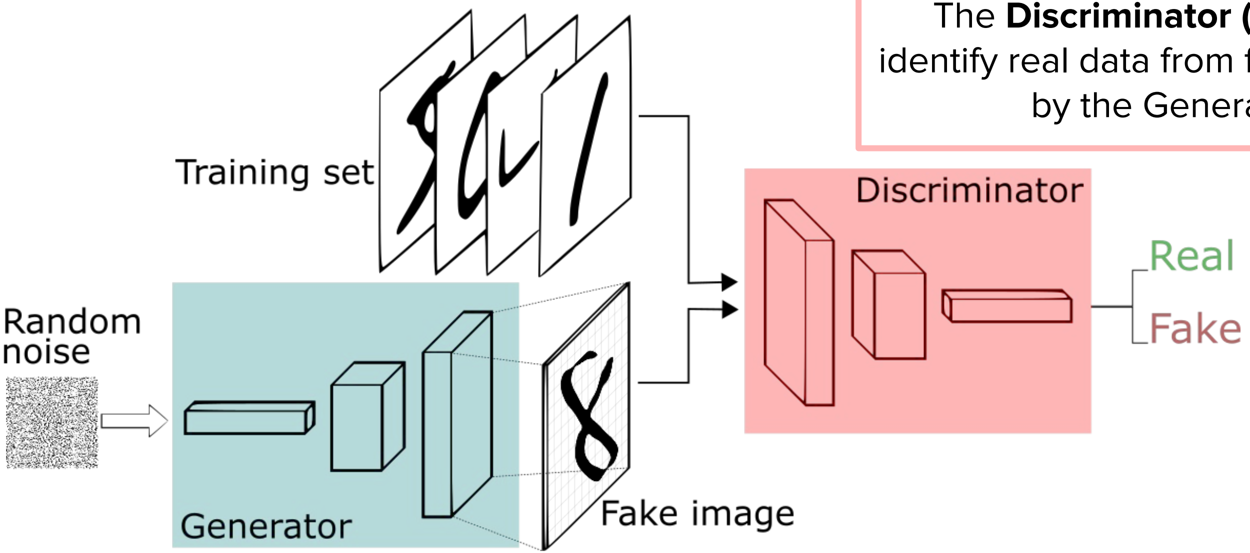


Training: adversarial objectives for **D** and **G**.

Global Optimum: **G** reproduces the true distribution.

Training GANs

The **Discriminator (D)** tries to identify real data from fakes created by the Generator.



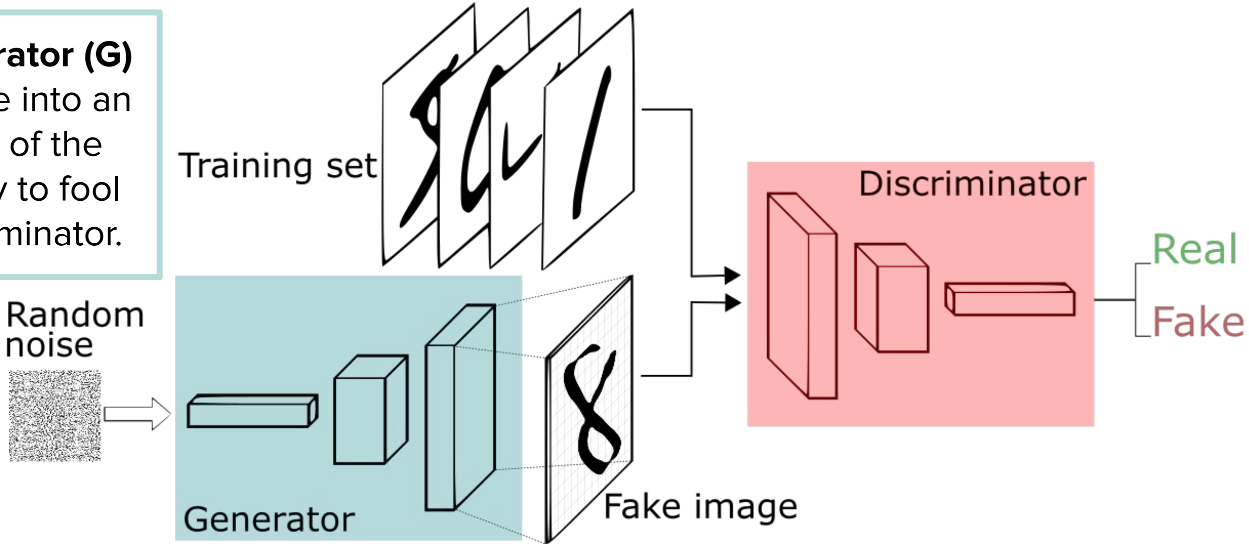
$$\text{argmax}_D \mathbb{E}_{\mathbf{z},\mathbf{x}}[\log D(G(\mathbf{z})) + \log(1 - D(\mathbf{x}))]$$

Fake

Real

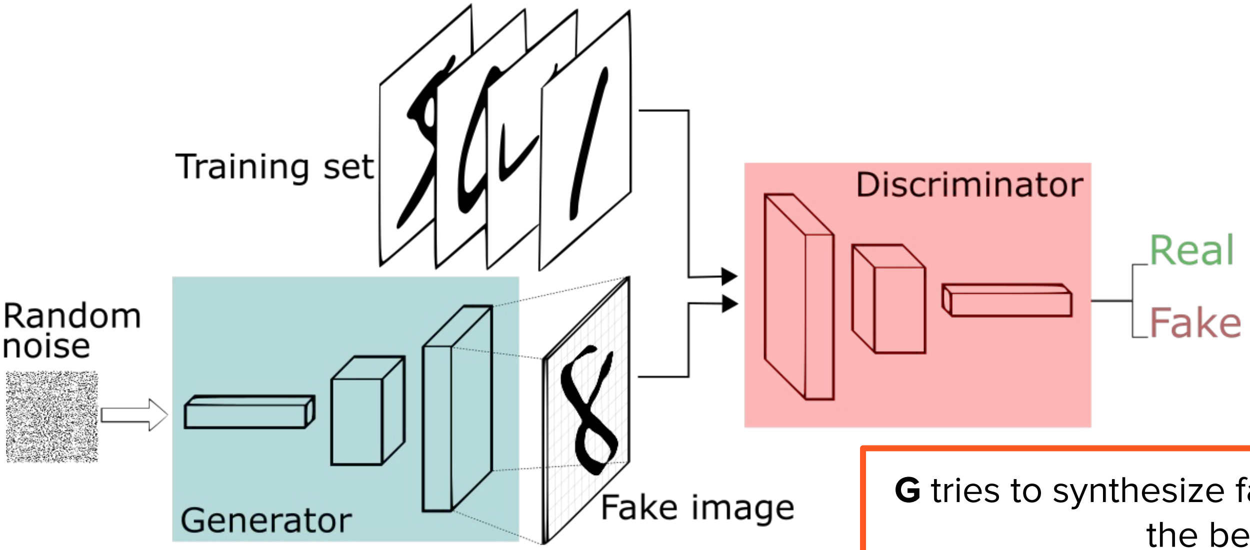
Training GANs

The **Generator (G)** turns noise into an imitation of the data to try to fool the Discriminator.



$$\mathrm{argmin}_G \mathbb{E}_{\mathbf{z},\mathbf{x}}[\log D(G(\mathbf{z})) + \log(1 - D(\mathbf{x}))]$$

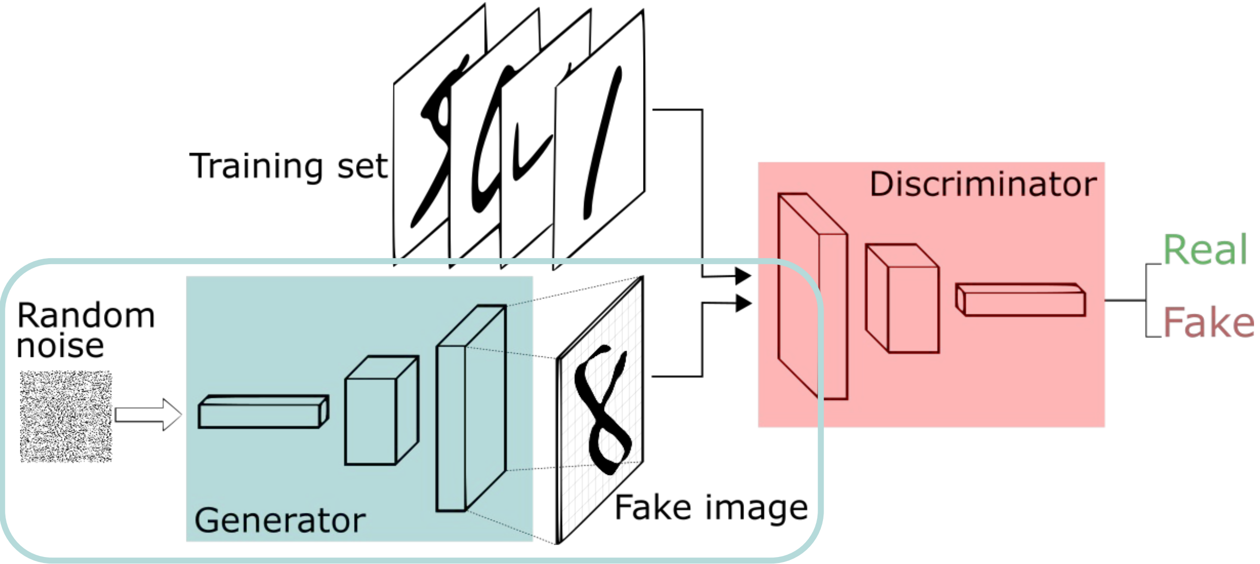
Training GANs



G tries to synthesize fake images that fool the best **D**.

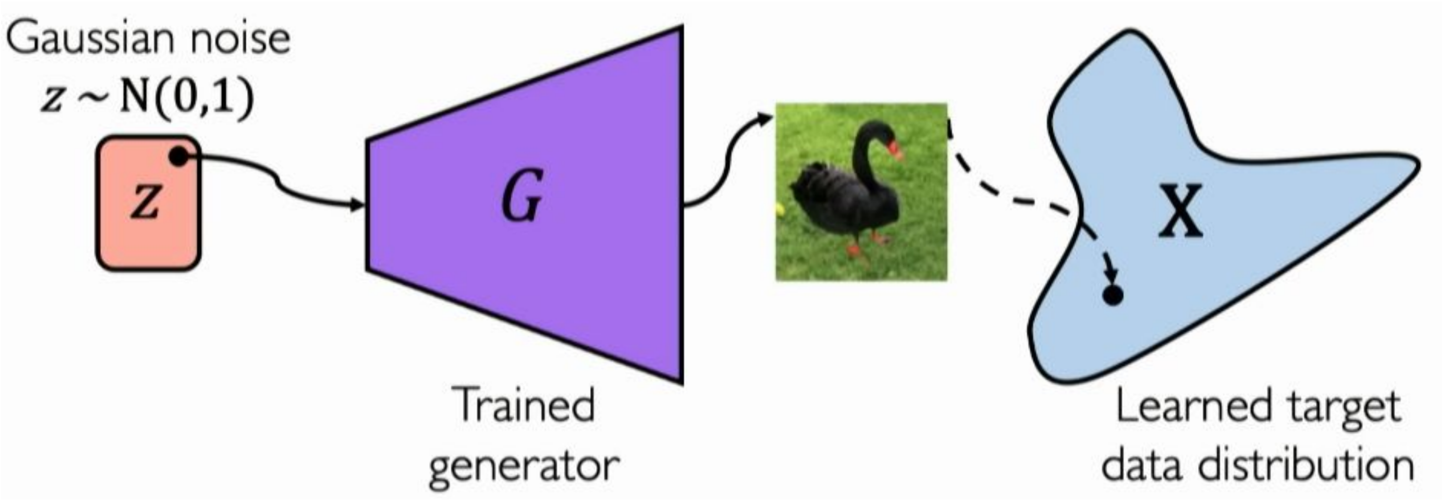
$$\text{argmin}_G \text{max}_D \mathbb{E}_{\mathbf{z},\mathbf{x}}[\log D(G(\mathbf{z})) + \log(1 - D(\mathbf{x}))]$$

Generating New Data from GANs

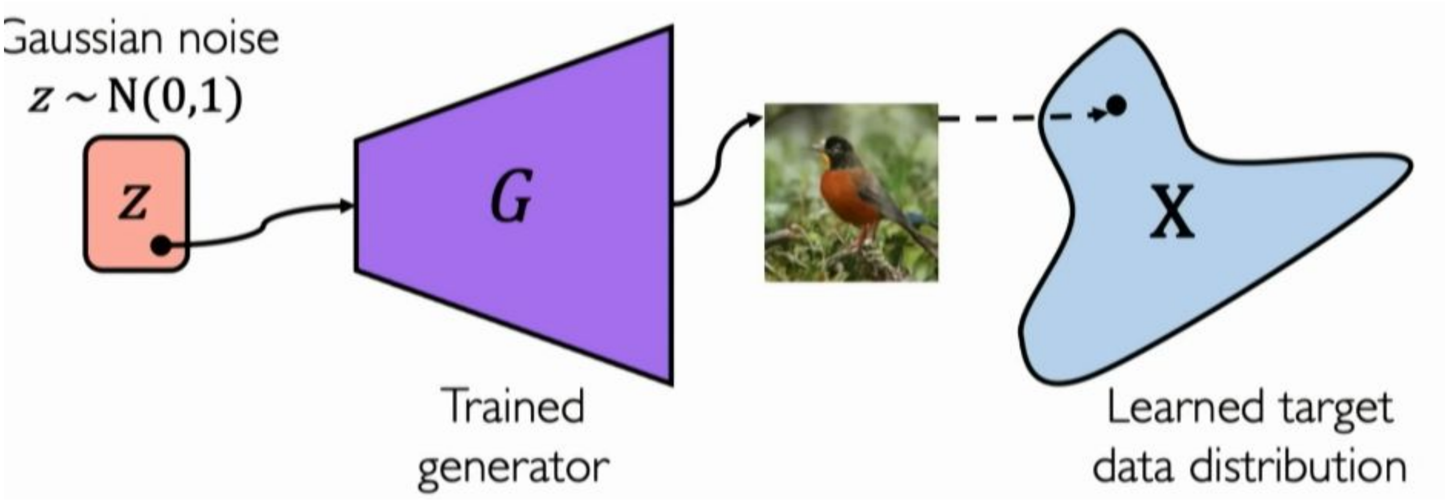


After training, use the generator network to generate **new data** that has never been seen before!

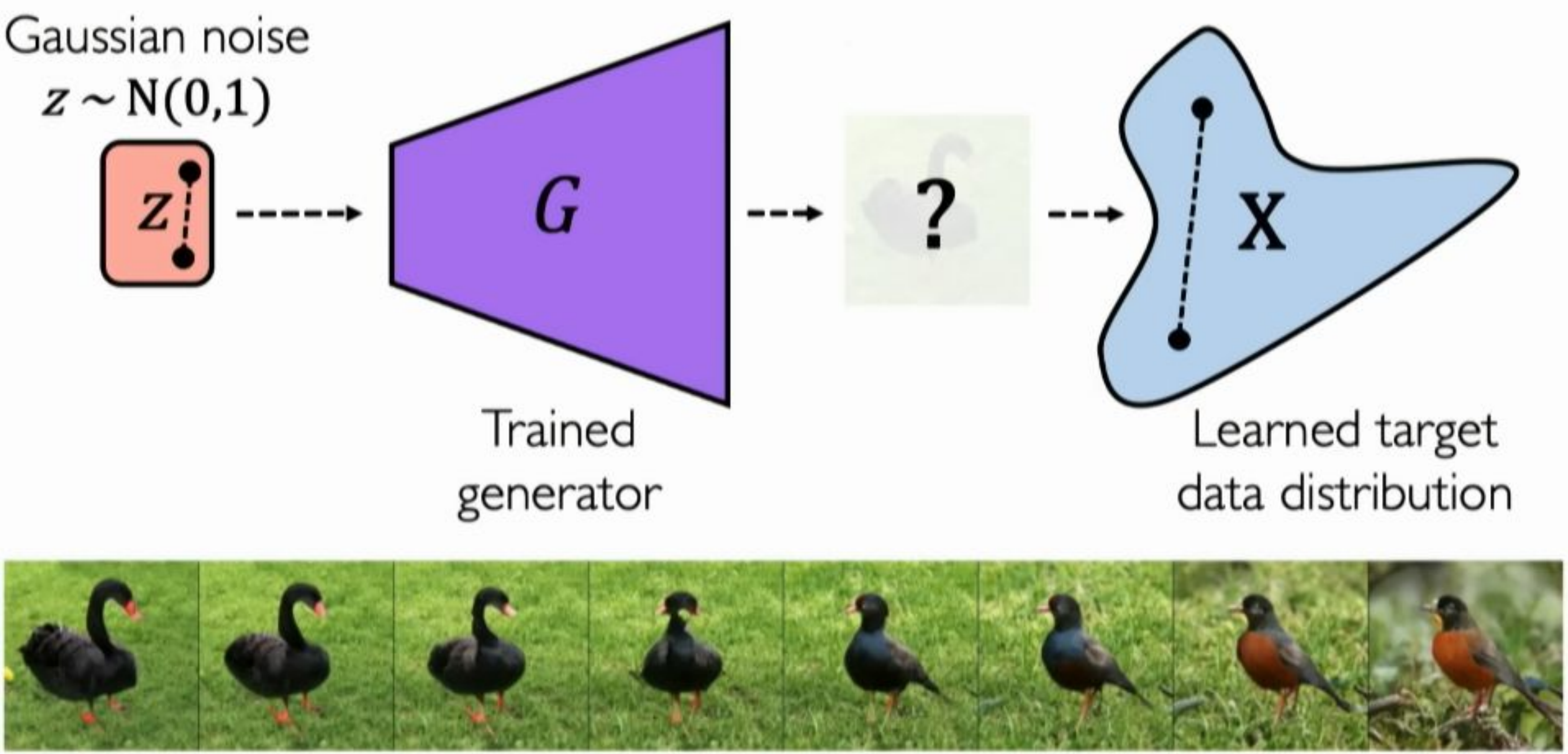
GANs are Distribution Transformers



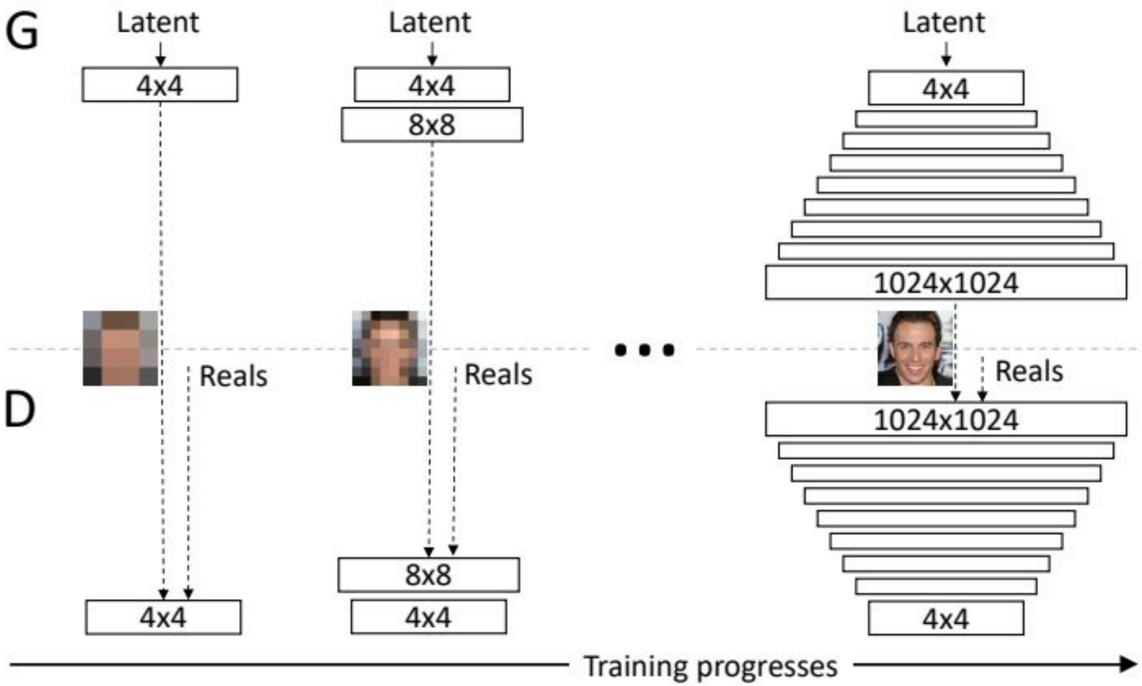
GANs are Distribution Transformers



GANs are Distribution Transformers

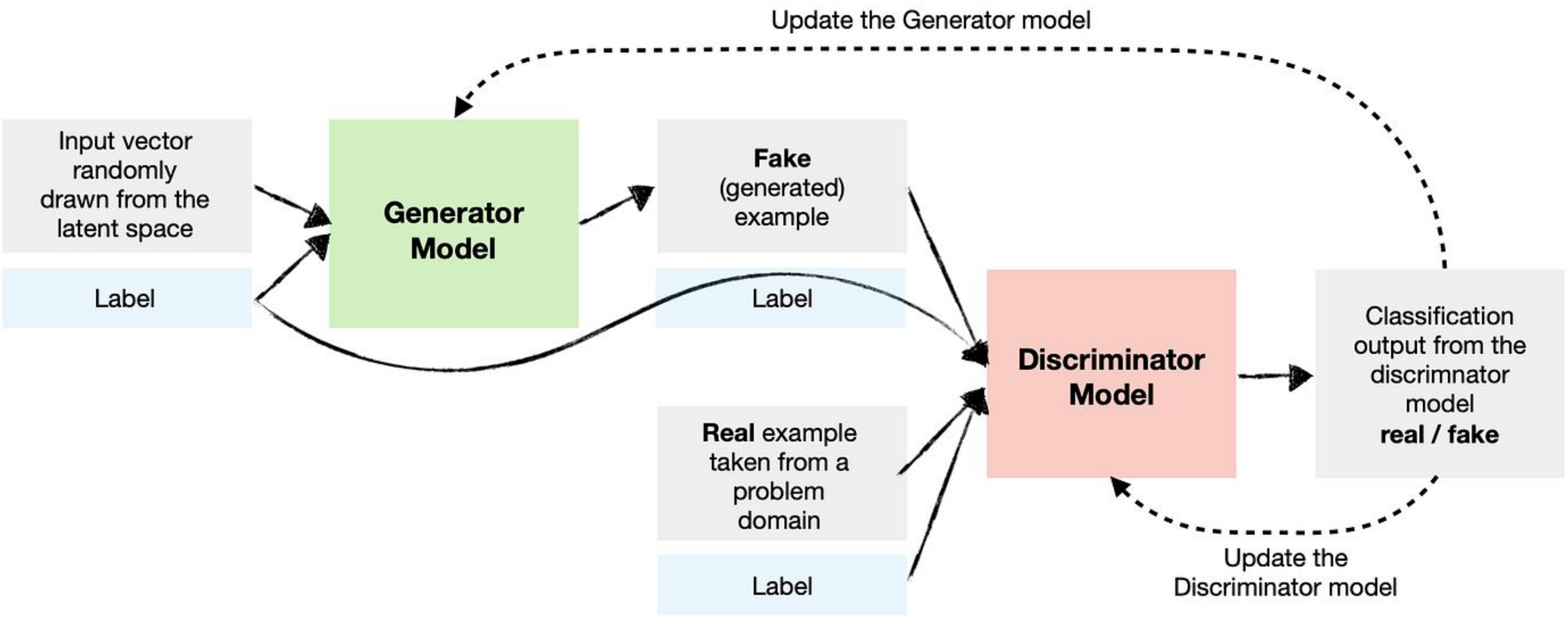


Progressively Growing GANs

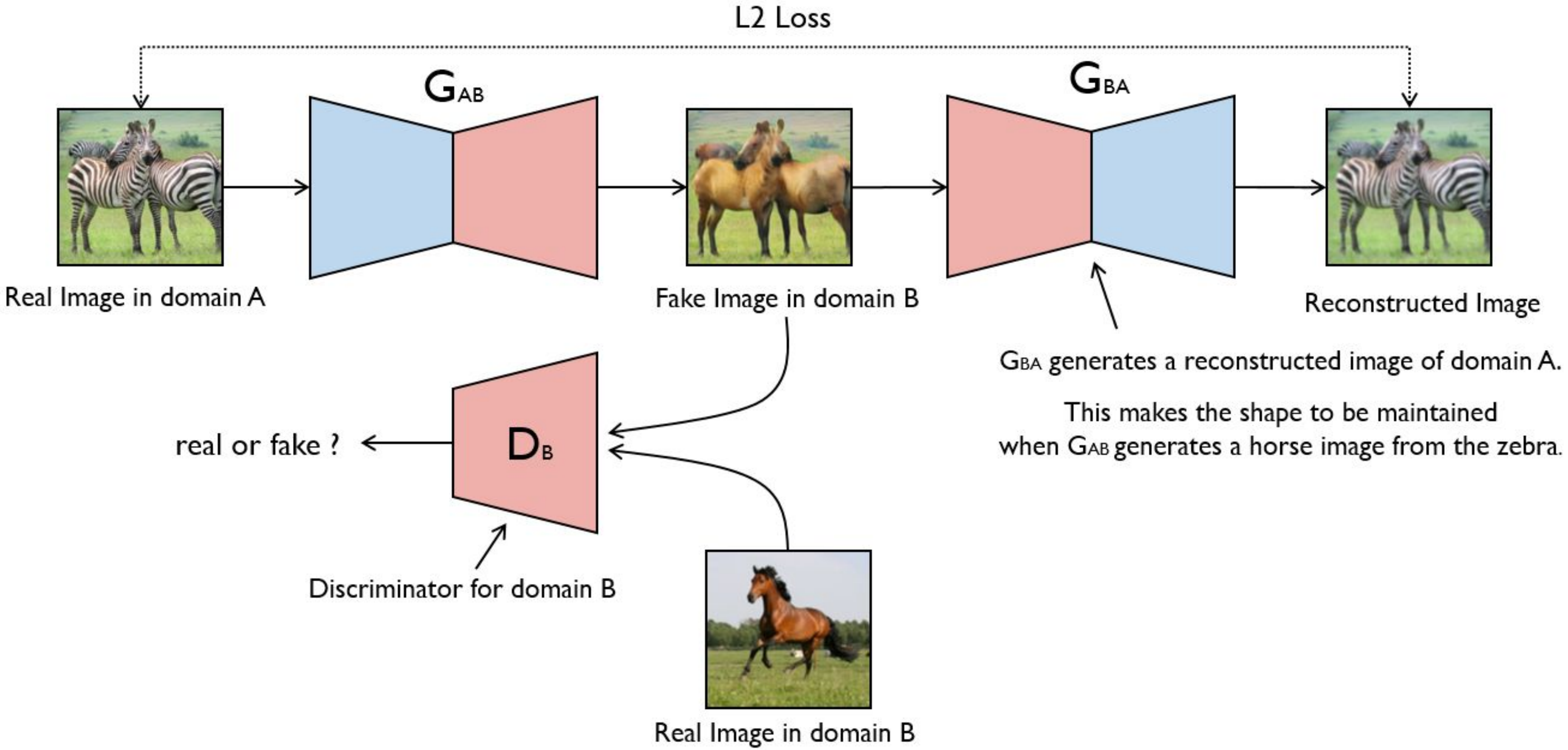


Progressively Growing GANs: Results

Conditional GANs



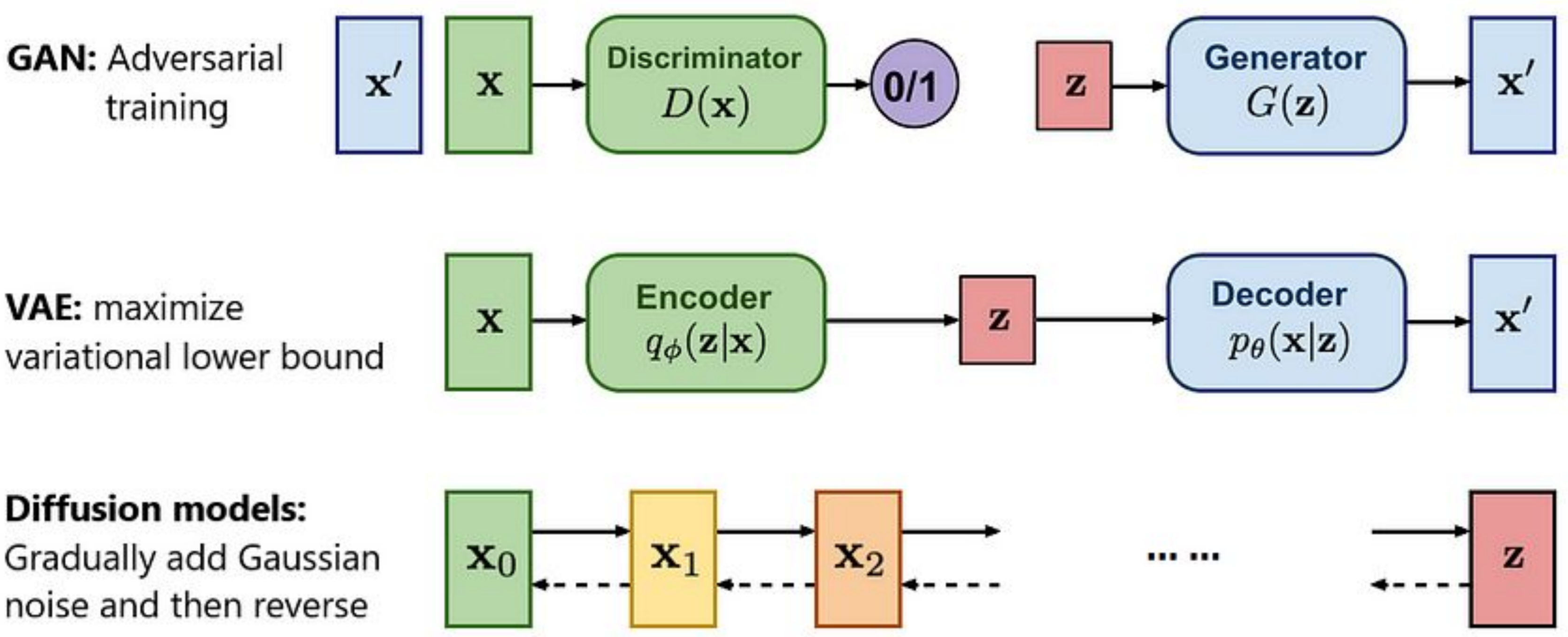
Cycle GANs



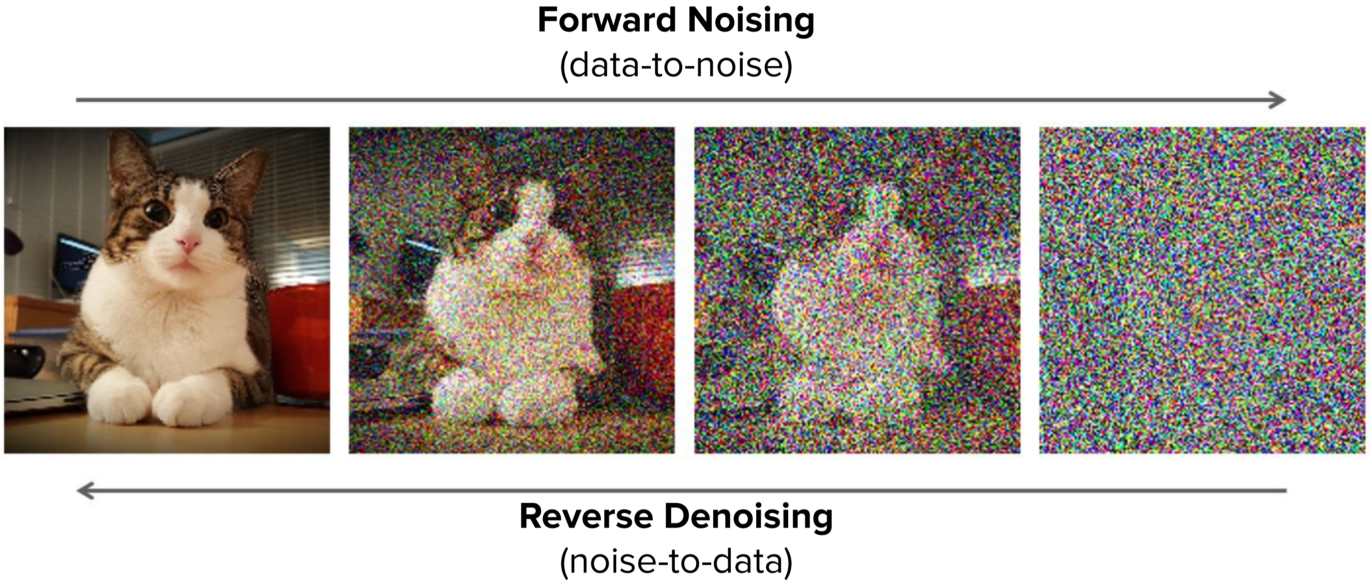
L2 Loss
G_{AB}
G_{BA}
Real Image in domain A
Fake Image in domain B
Reconstructed Image
G_{BA} generates a reconstructed image of domain A.
This makes the shape to be maintained
when G_{AB} generates a horse image from the zebra.
real or fake ?
D_B
Discriminator for domain B
Real Image in domain B

Diffusion Models!

Diffusion Models

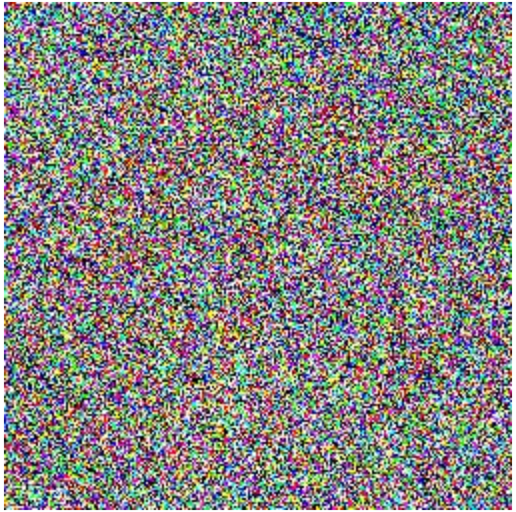


The Diffusion Process



Forward Noising

Step 1: Given an image (left), randomly sample a random noise pattern (right)



Forward Noising

Step 2: Progressively add more and more of the noise to your image

T = 0

100% Image
0% Noise

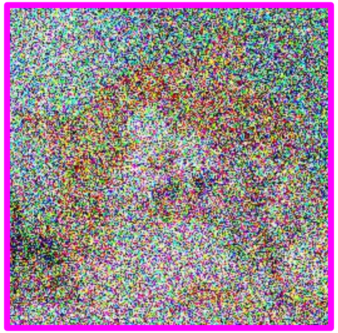
T = 1

75% Image
25% Noise

T = 2

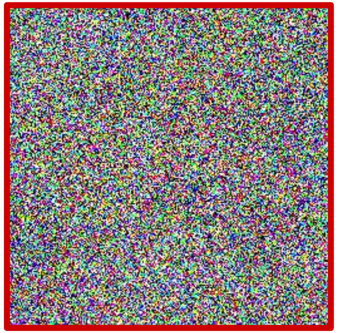
50% Image
50% Noise

T = 3



25% Image
75% Noise

T = 4



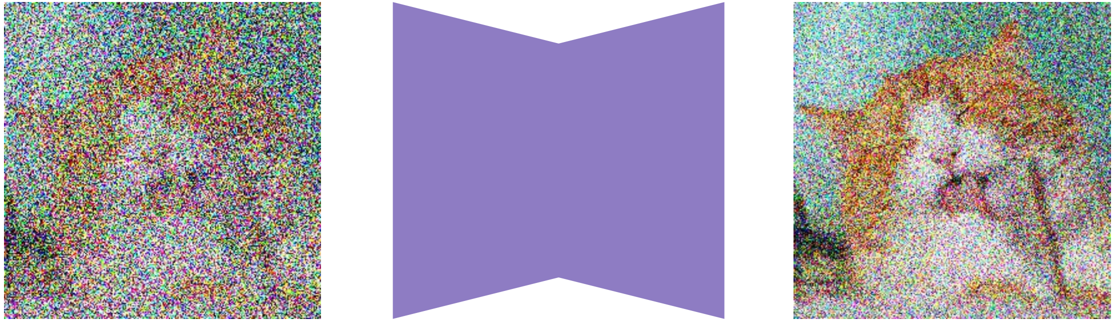
0% Image
100% Noise

Reverse Denoising

T = 0 ... T = 4

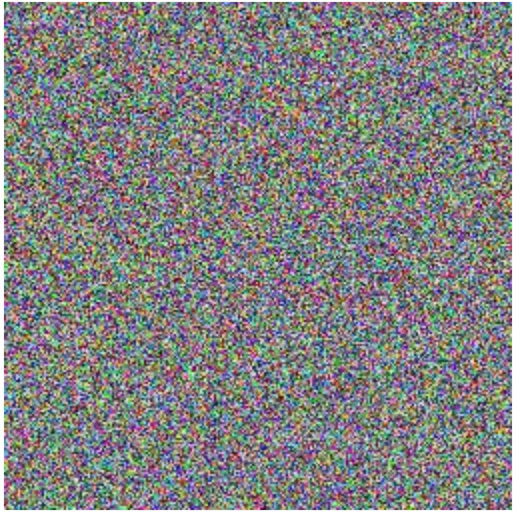
Goal: Given image at **T**, can we learn to estimate image at **T - 1**?

T = 3 **T = 2**



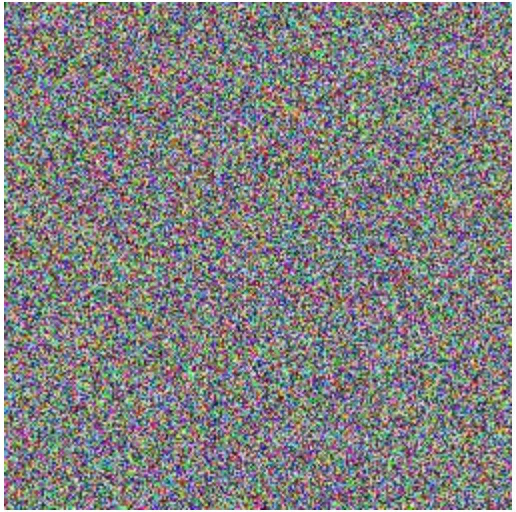
How can we train this network?

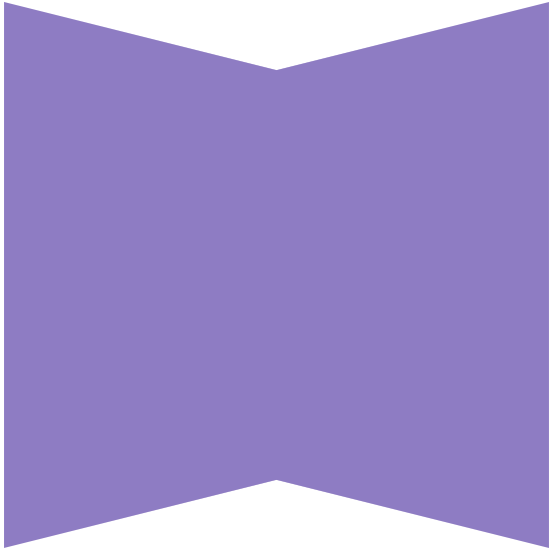
Sampling Brand New Generations

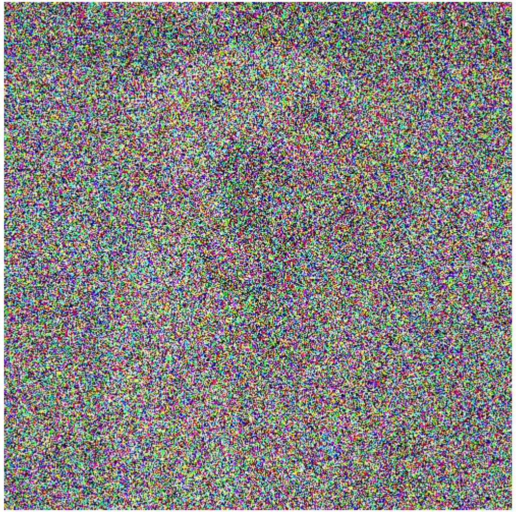


T

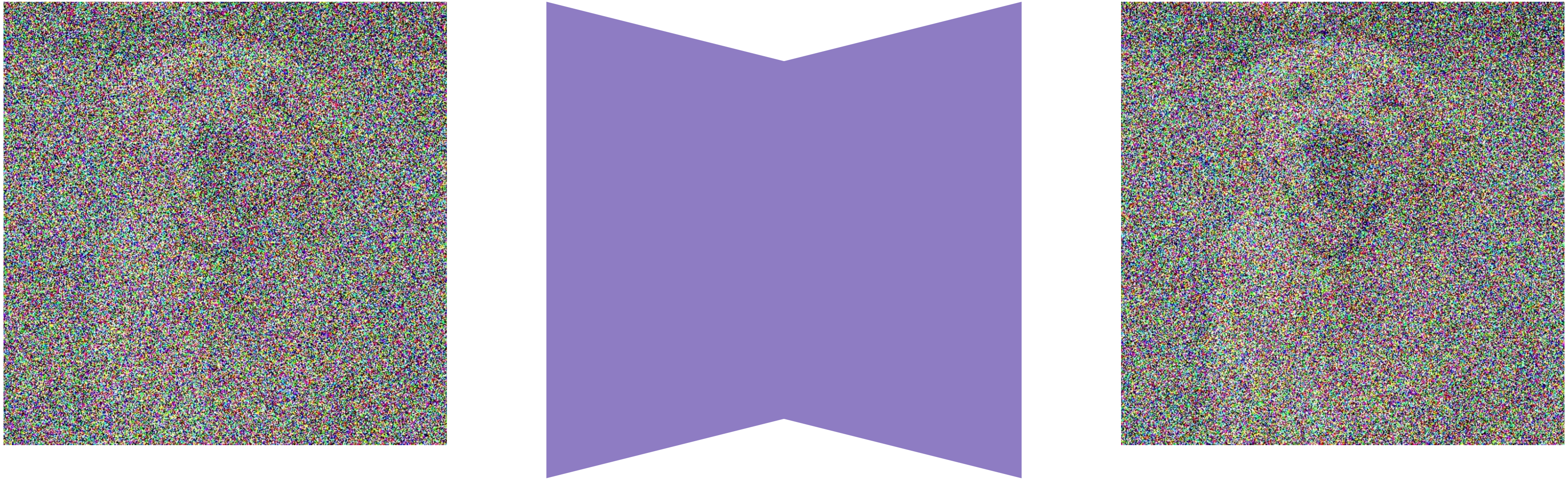
Sampling Brand New Generations



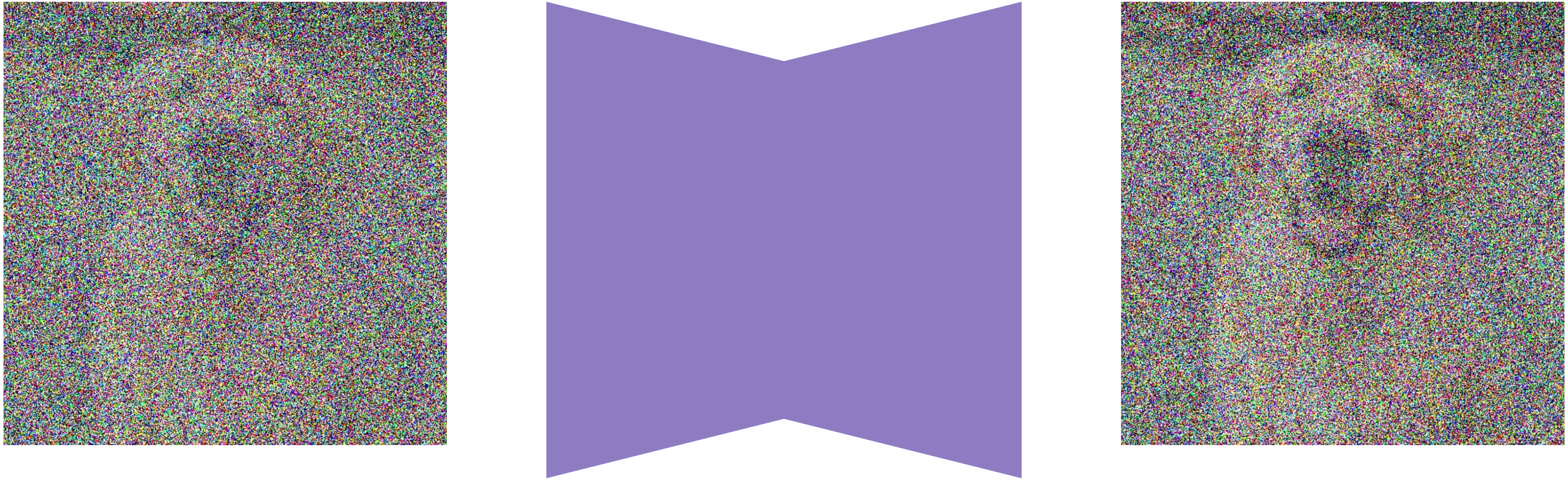




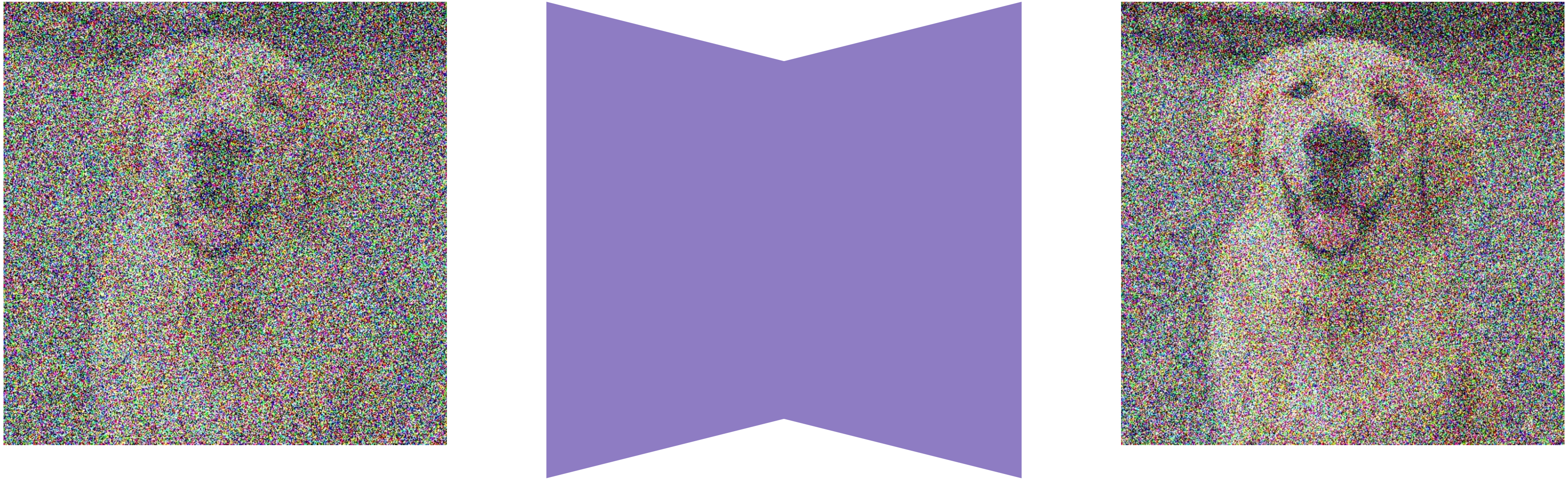
Sampling Brand New Generations



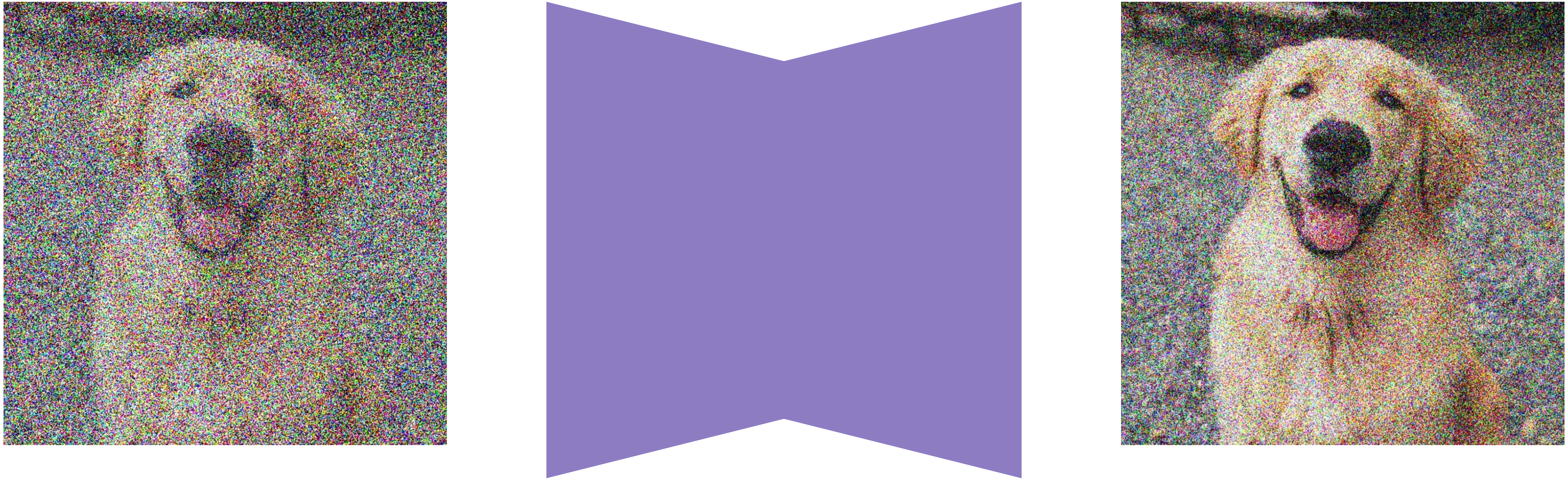
Sampling Brand New Generations



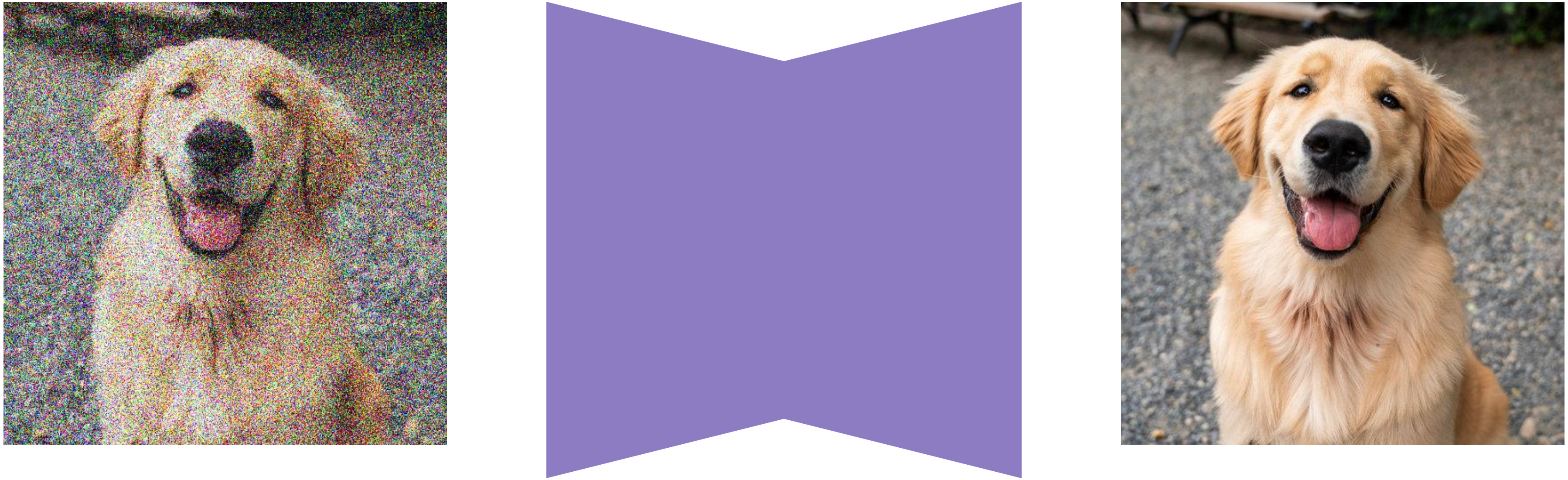
Sampling Brand New Generations



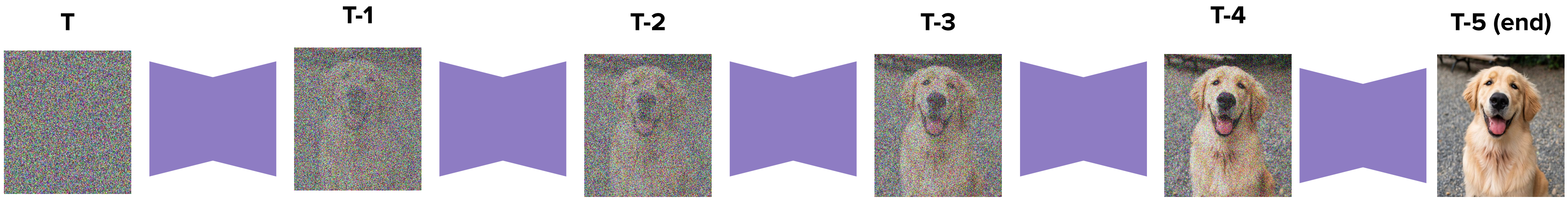
Sampling Brand New Generations



Sampling Brand New Generations

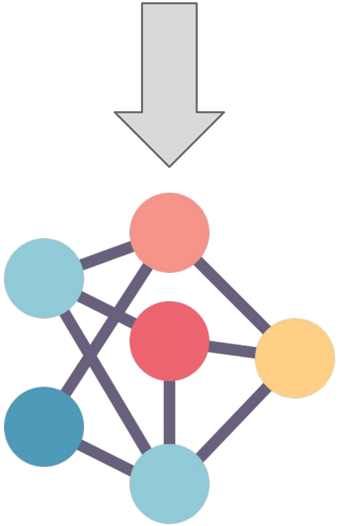


Sampling Brand New Generations



Generating Images from Natural Language

“A Photo of an astronaut riding a horse in Mars”



Generating Images from Natural Language

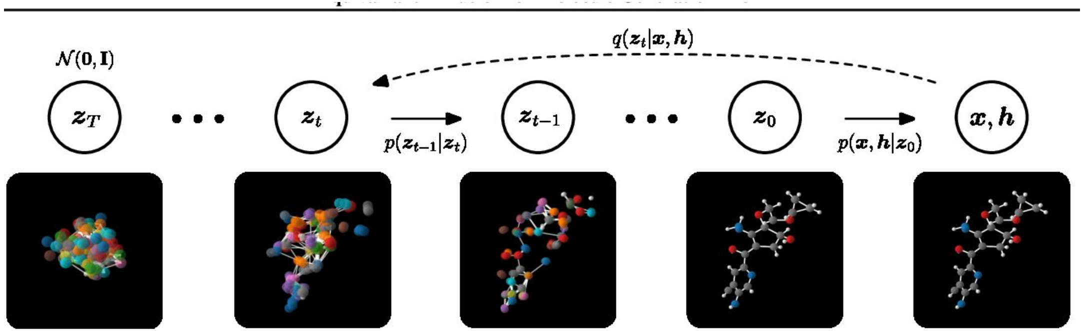
"a painting of a fox sitting in a field at sunrise in the style of Claude Monet"

"An ibis in the wild, painted in the style of John Audubon"

"close-up of a snow leopard in the snow hunting, rack focus, nature photography"

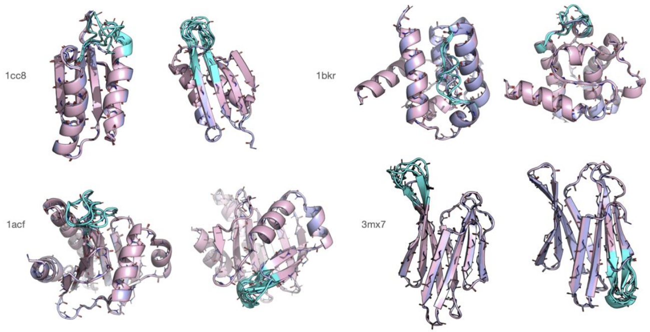
Beyond Images: Molecular Design

Chemistry: Generating Molecules in 3D



Equivariant Diffusion for Molecule Generation in 3D

Biology: Generating Novel Proteins



Protein Structure and Sequence Generation with Equivariant Denoising Diffusion Probabilistic Models

Next Lecture: Object Detection And Segmentation!