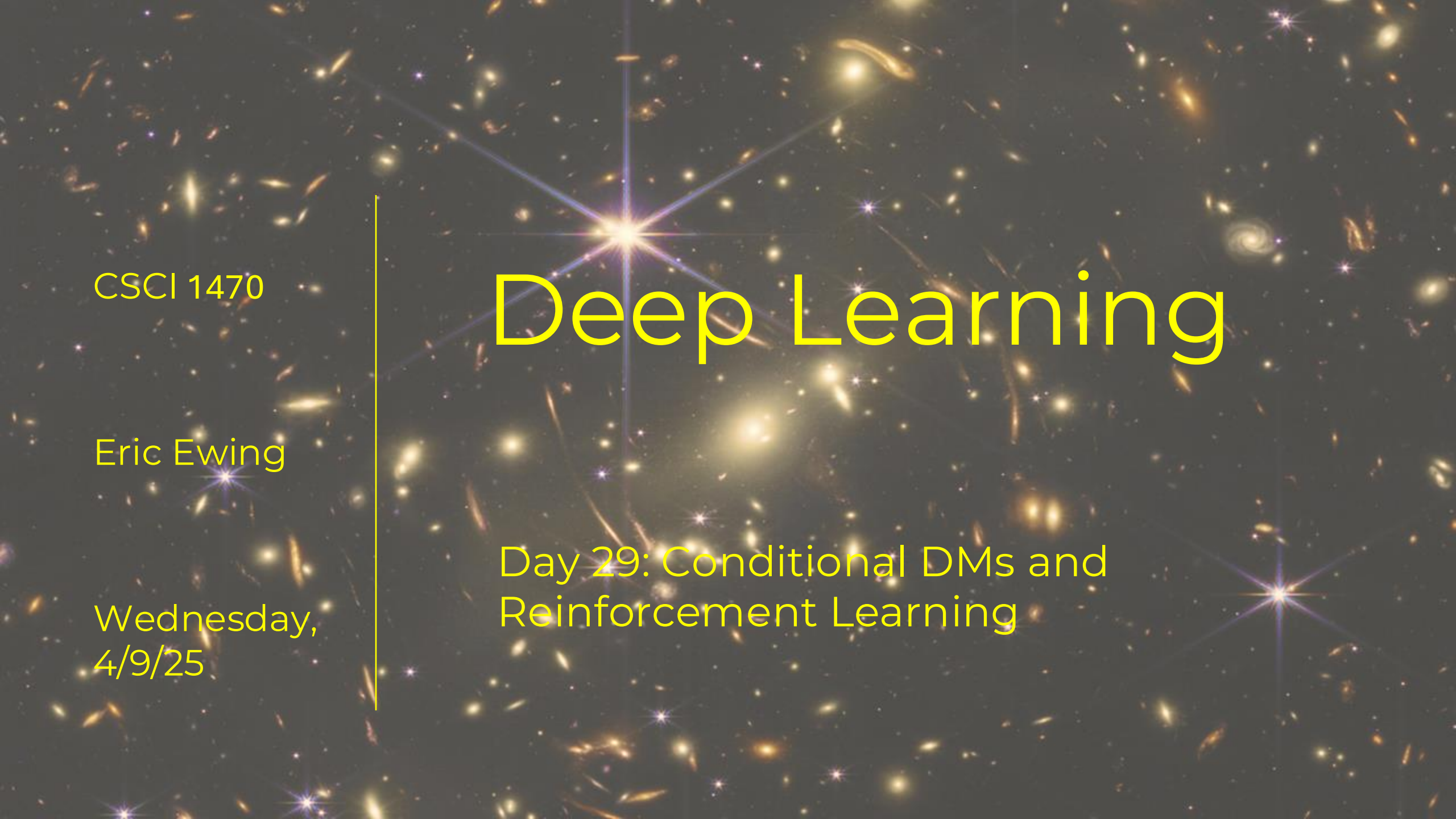


CSCI 1470

Eric Ewing

Wednesday, 4/9/25

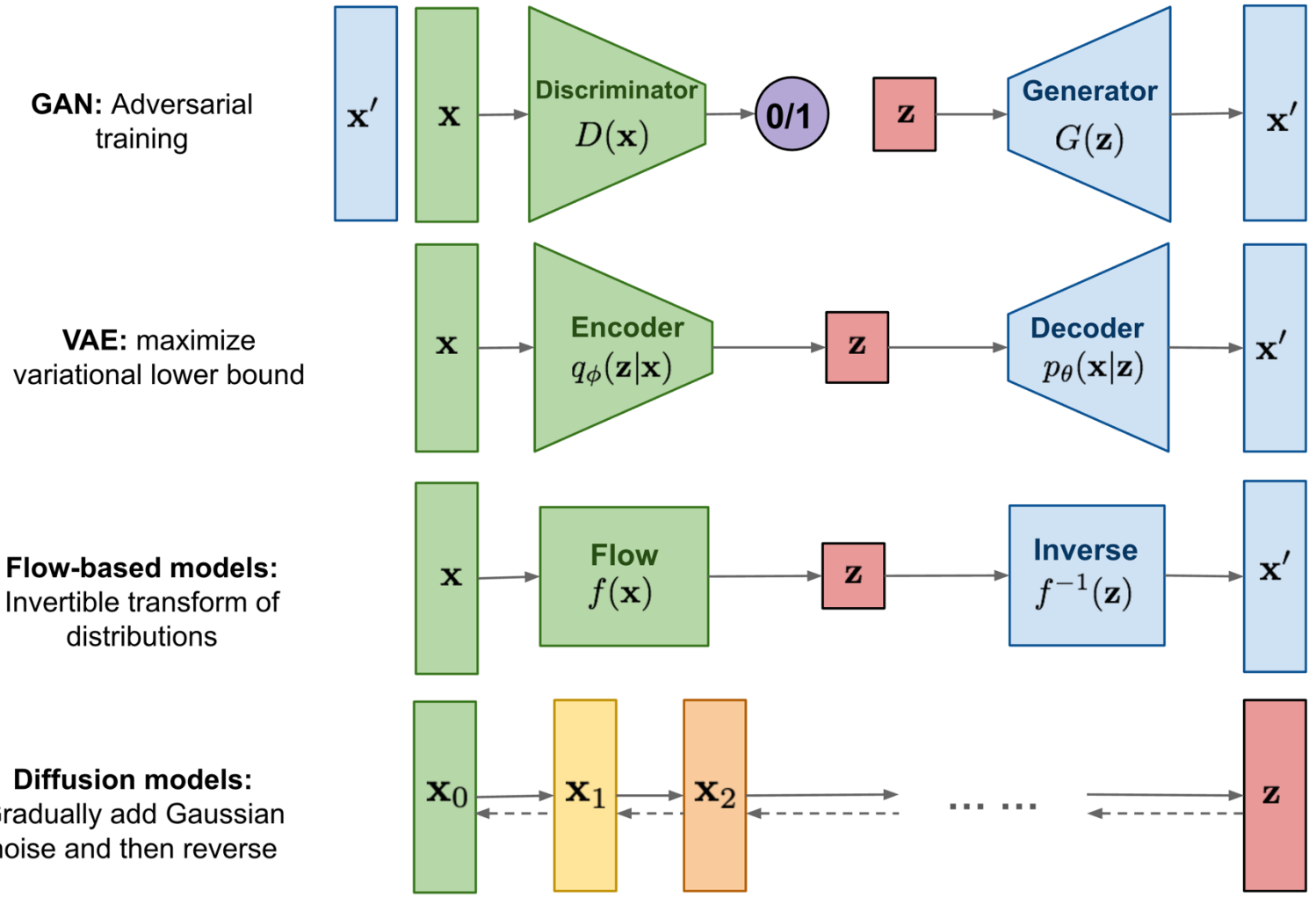
Deep Learning

Day 29: Conditional DMs and Reinforcement Learning

Logistics

- Weekly Quiz #9 is available now

Review of Generative Models



Source: https://lilianweng.github.io/posts/2021-07-11-diffusion-models/

But How do we specify what we want?

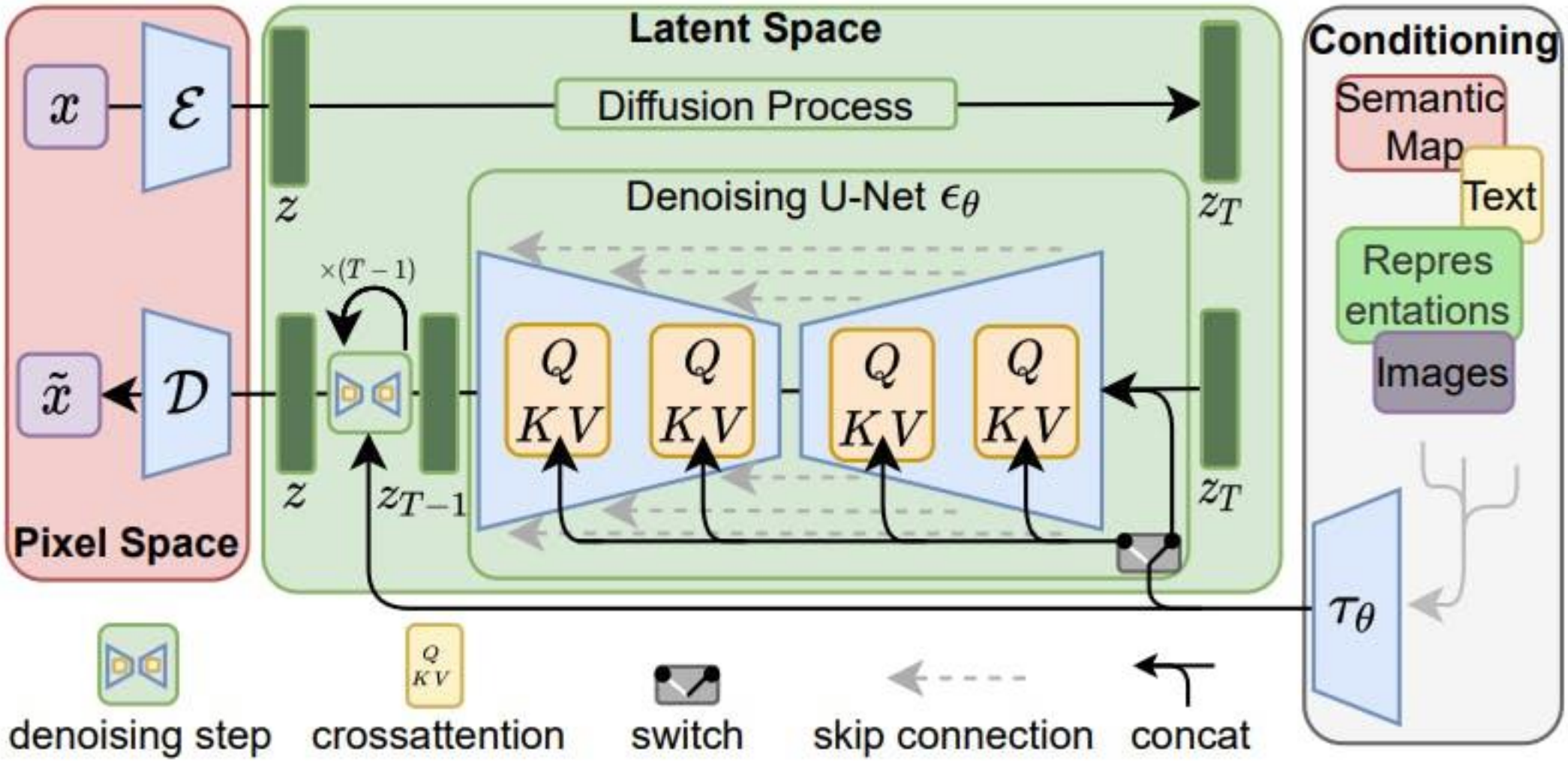
Can you generate an image of deep space for me?

Can you remove the people from this image?

Can you give me your best joke in image format?

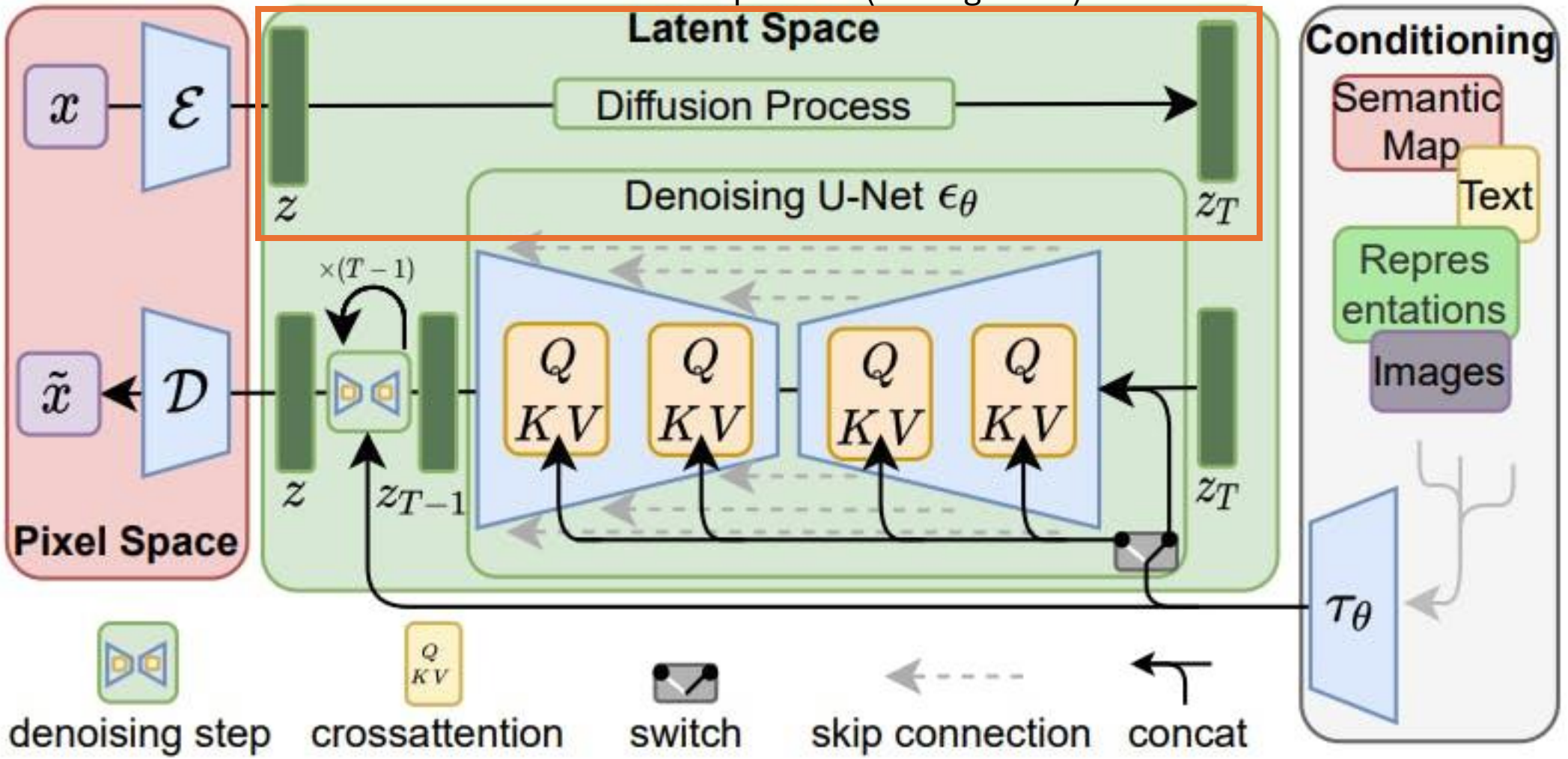
Source: chatgpt

This is just one specific implementation

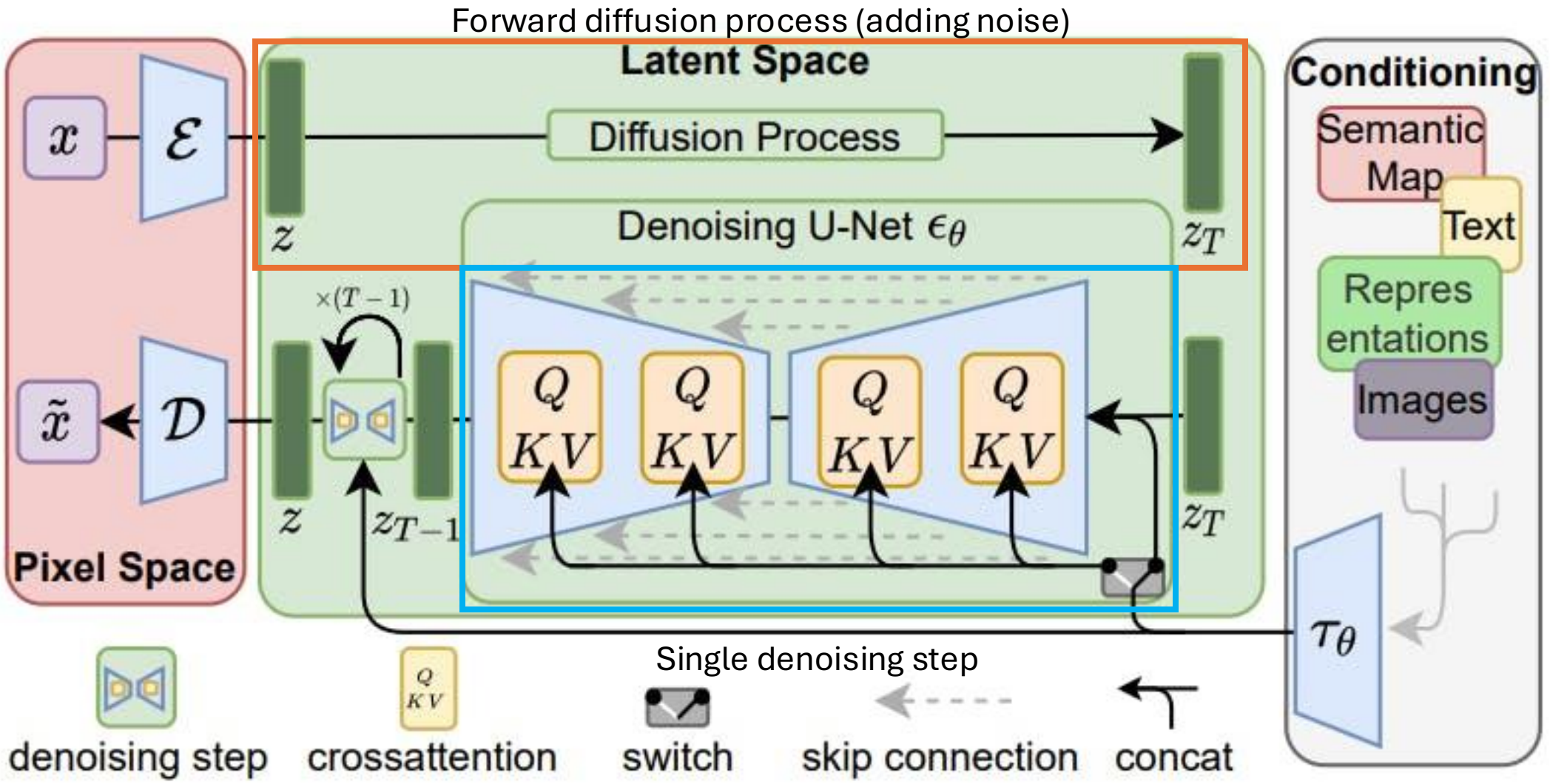


Source: High-Resolution Image Synthesis with Latent Diffusion Models https://arxiv.org/pdf/2112.10752

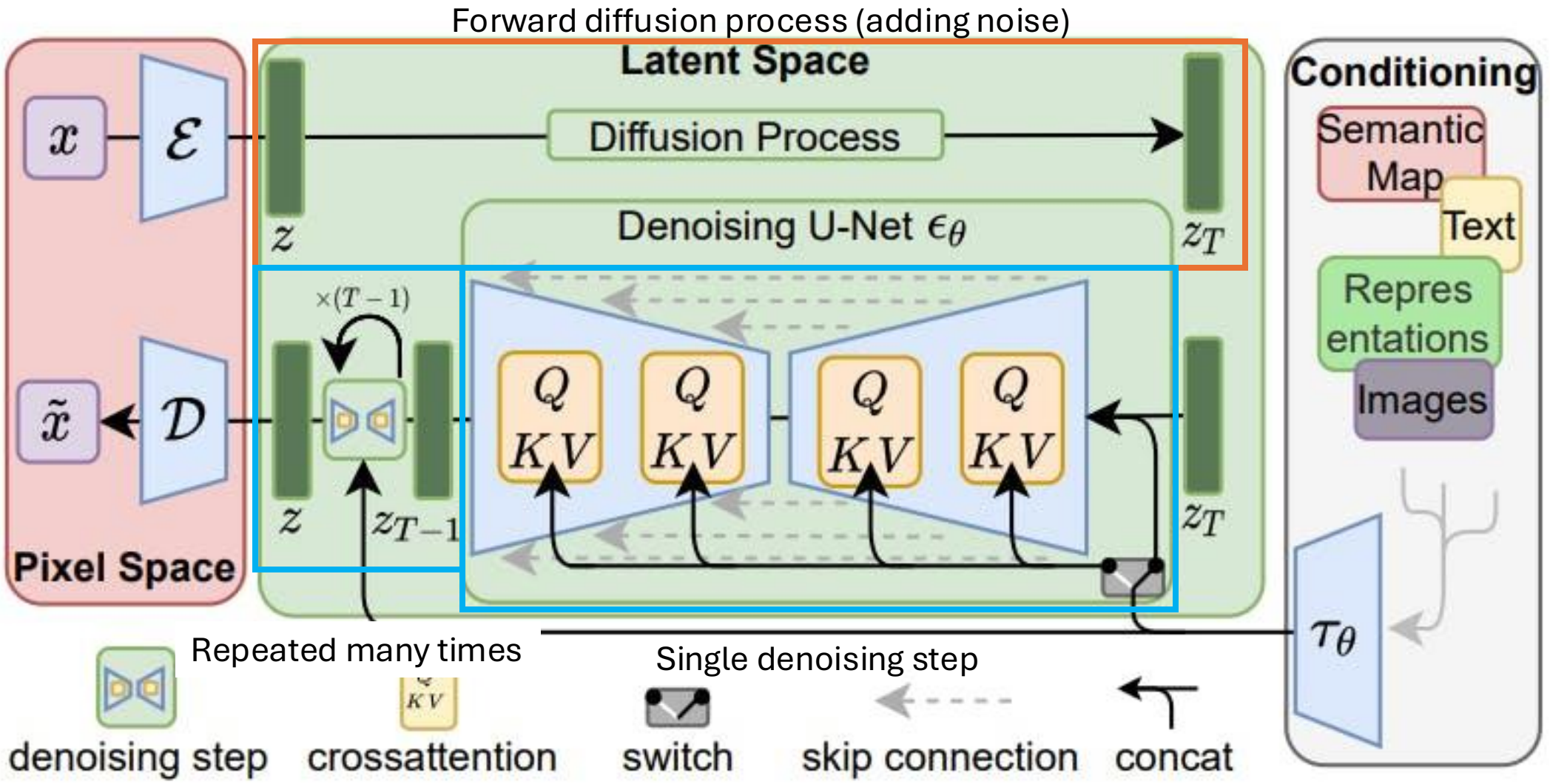
Forward diffusion process (adding noise)



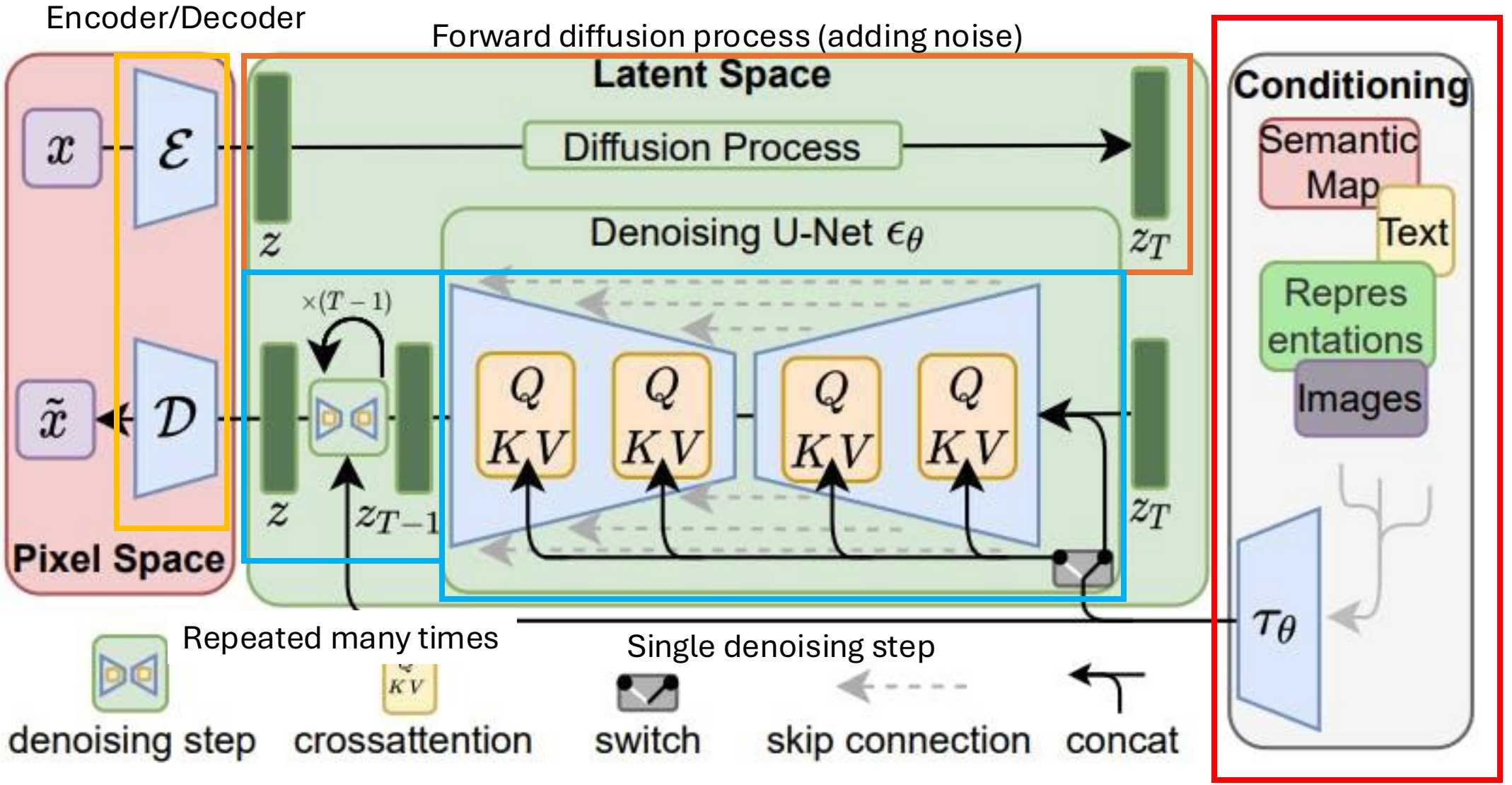
Source: High-Resolution Image Synthesis with Latent Diffusion Models https://arxiv.org/pdf/2112.10752



Source: High-Resolution Image Synthesis with Latent Diffusion Models https://arxiv.org/pdf/2112.10752



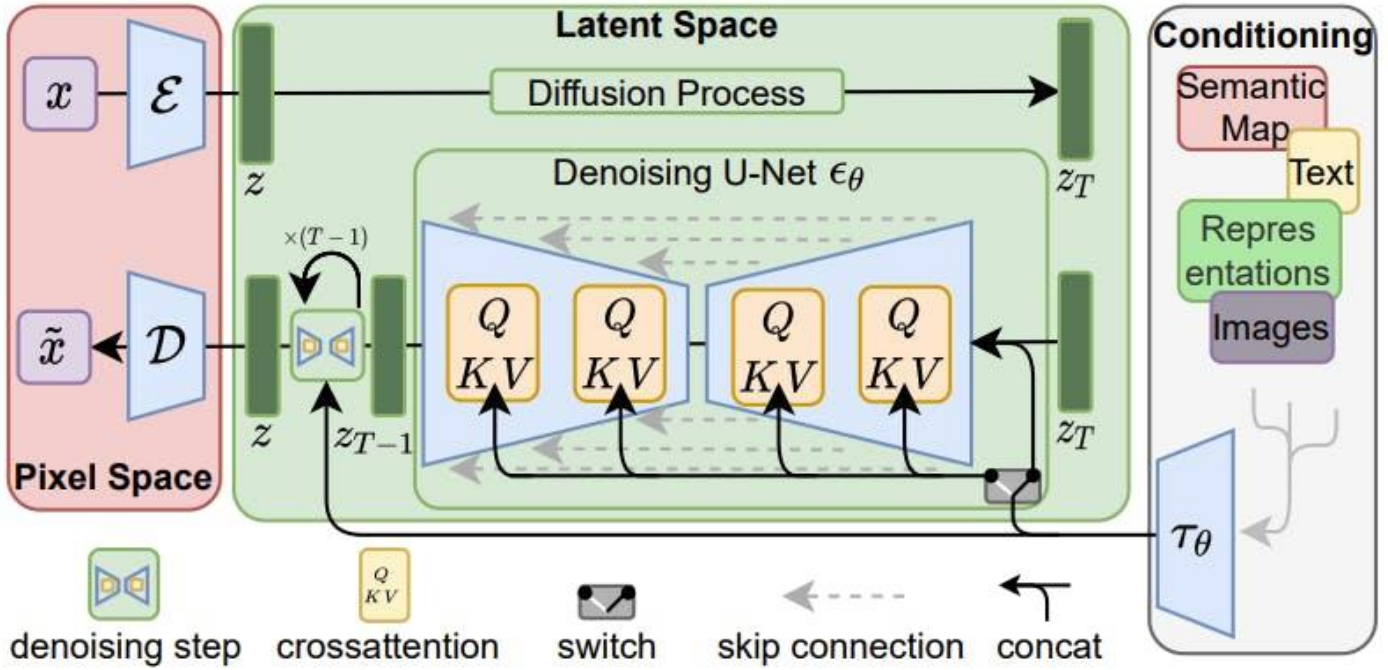
Source: High-Resolution Image Synthesis with Latent Diffusion Models https://arxiv.org/pdf/2112.10752



Source: High-Resolution Image Synthesis with Latent Diffusion Models https://arxiv.org/pdf/2112.10752

Encoders and Decoders are back!

Why might encoders and decoders be useful now?

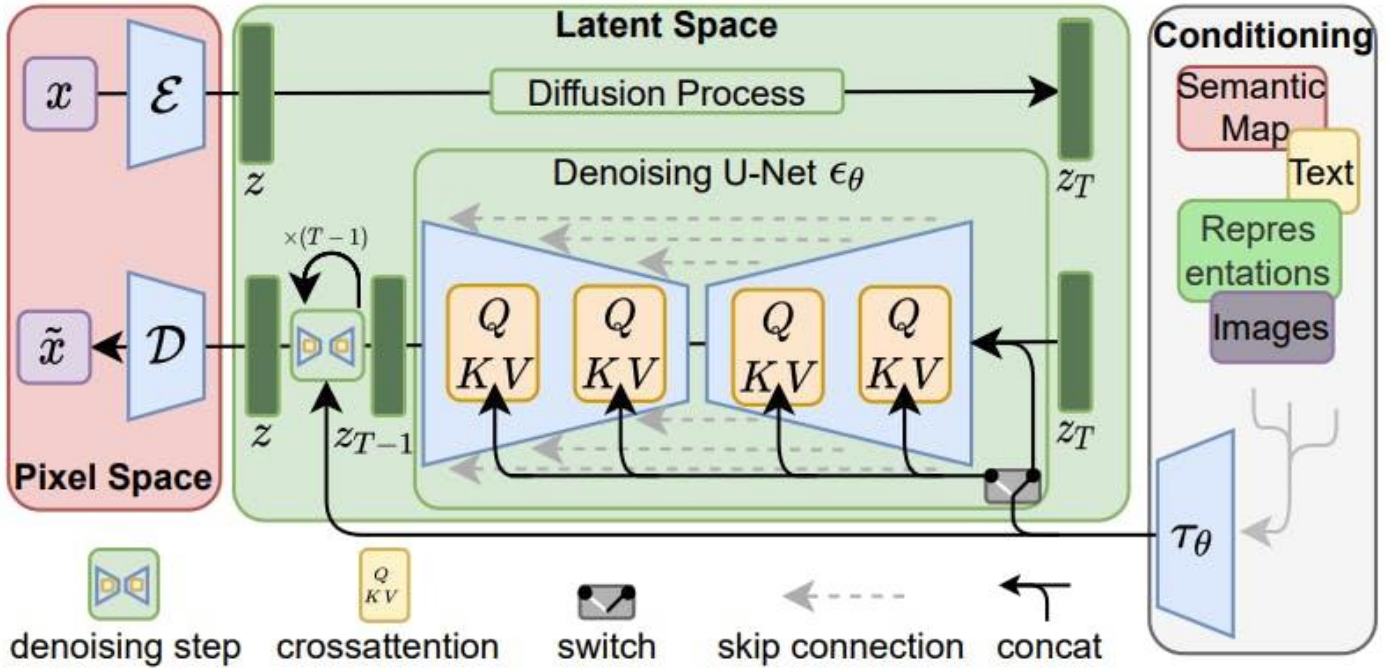


Encoders and Decoders are back!

These are intended to be very *general* models

Why might encoders and decoders be useful now?

- Work with language prompting
- Incorporate existing images
- And any other mode of input you can think of

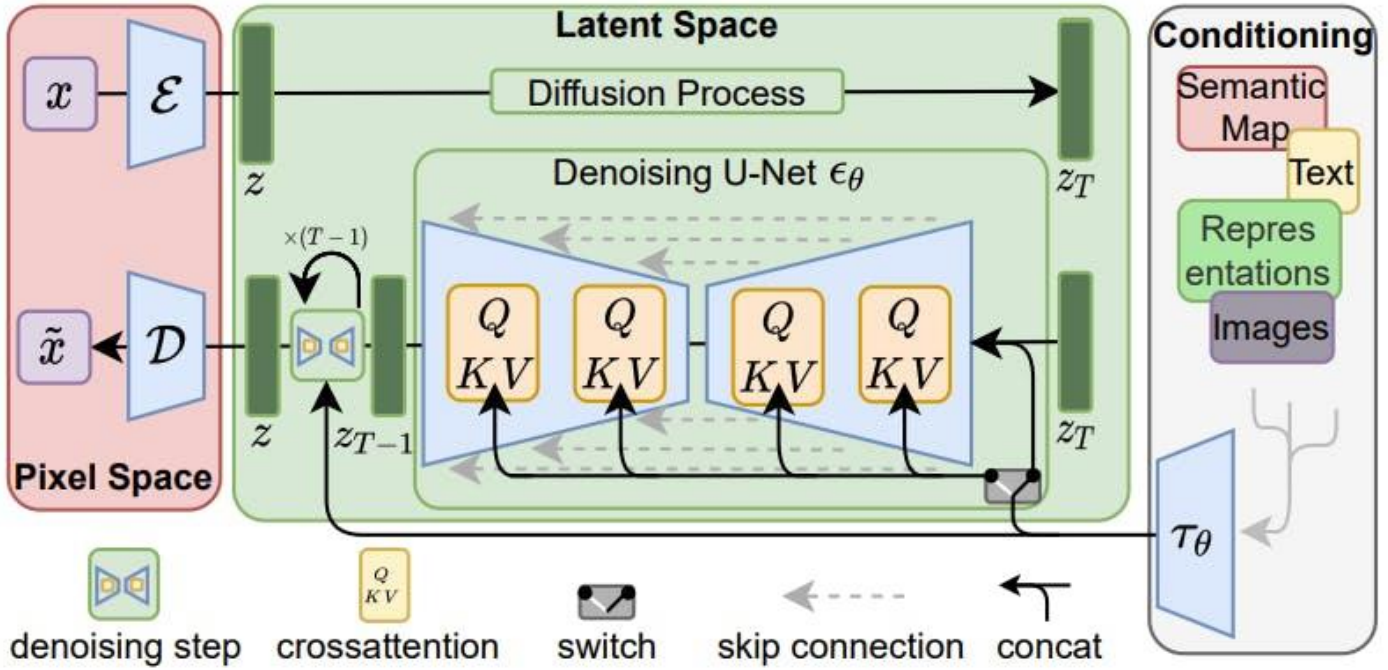


Encoders and Decoders are back!

These are intended to be very *general* models

Why might encoders and decoders be useful now?

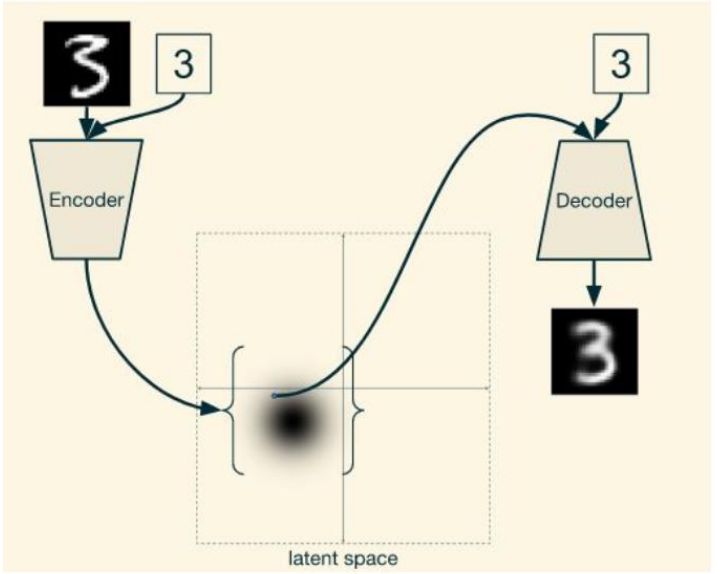
Embeddings are a sort of “common language” for all types of input

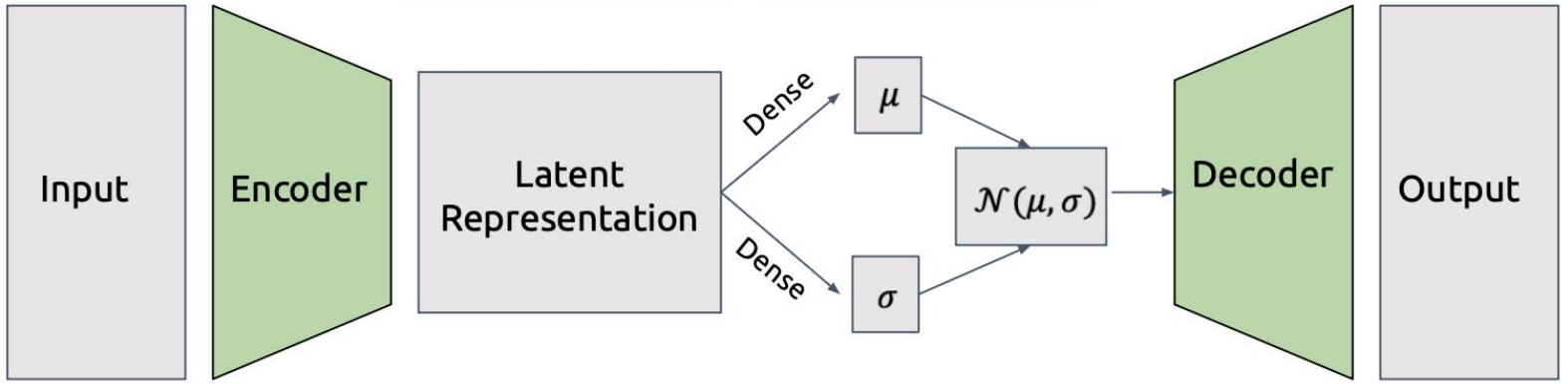


Conditional VAE

Label is concatenated to input of encoder and decoder (or an embedding based on the label)

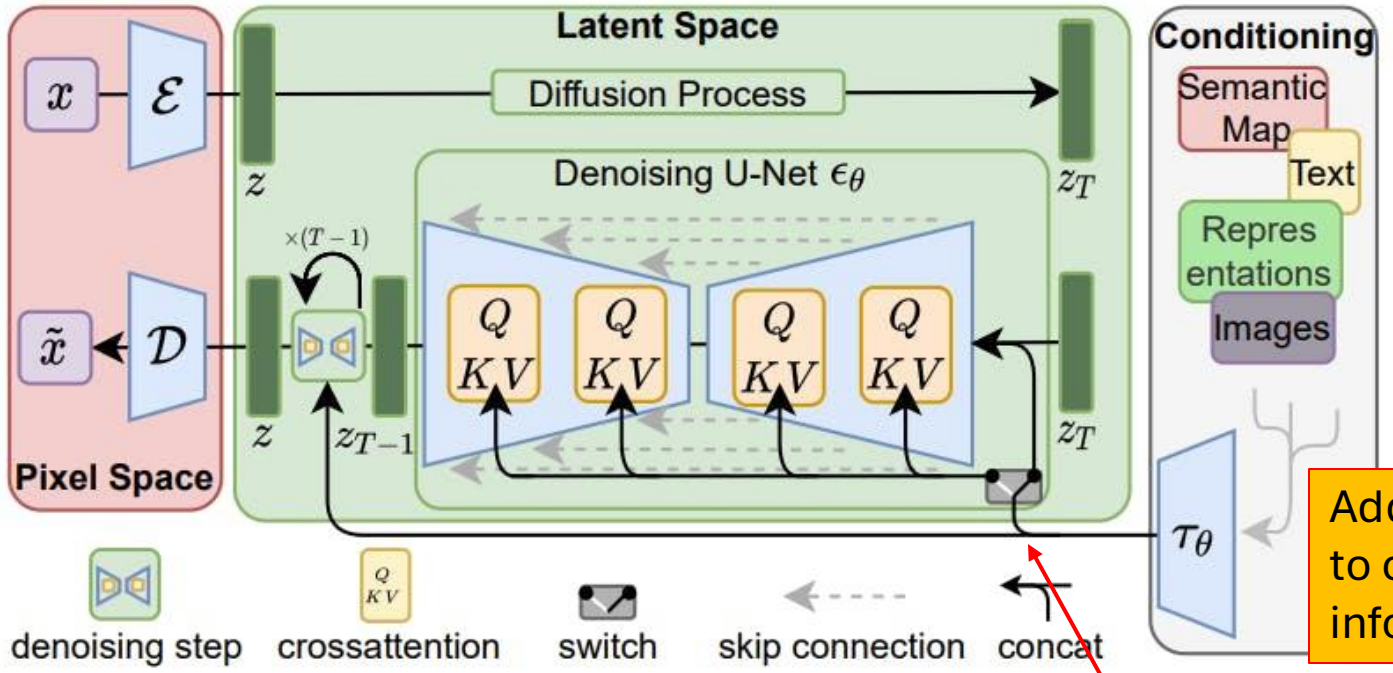
Label is just one more piece of useful information





https://towardsdatascience.com/understanding-conditional-variational-autoencoders-cd62b4f57bf8

Conditioning in Diffusion Models



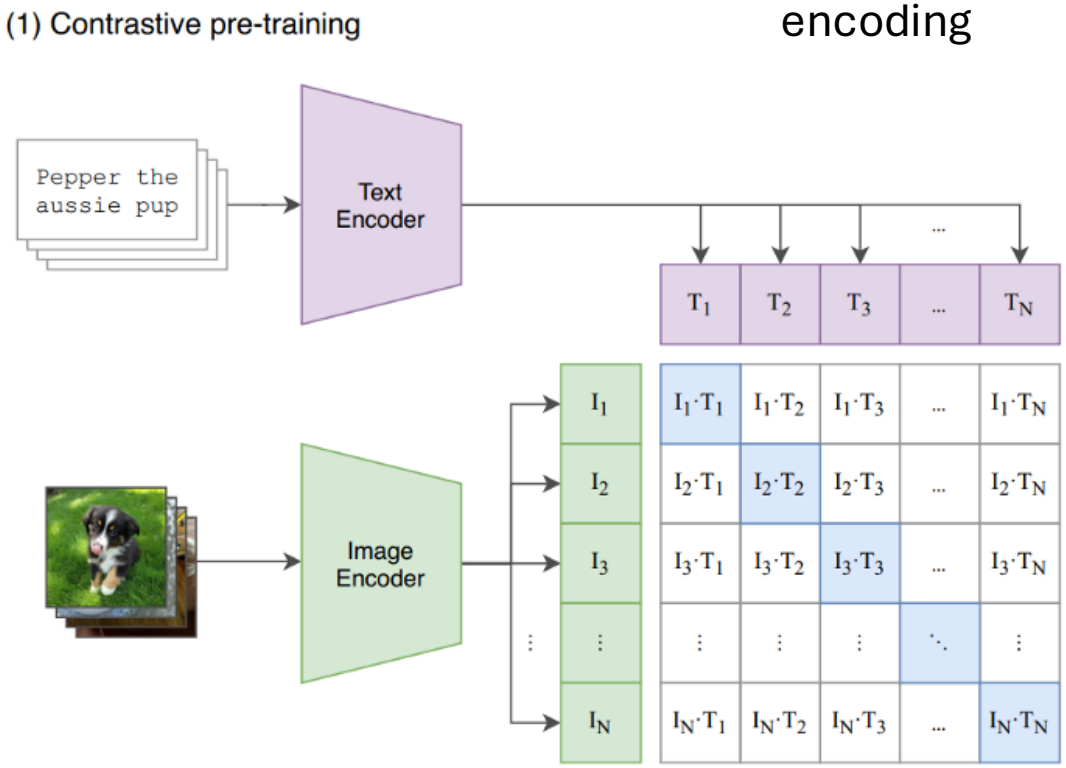
Additional model is trained to convert other information into encodings

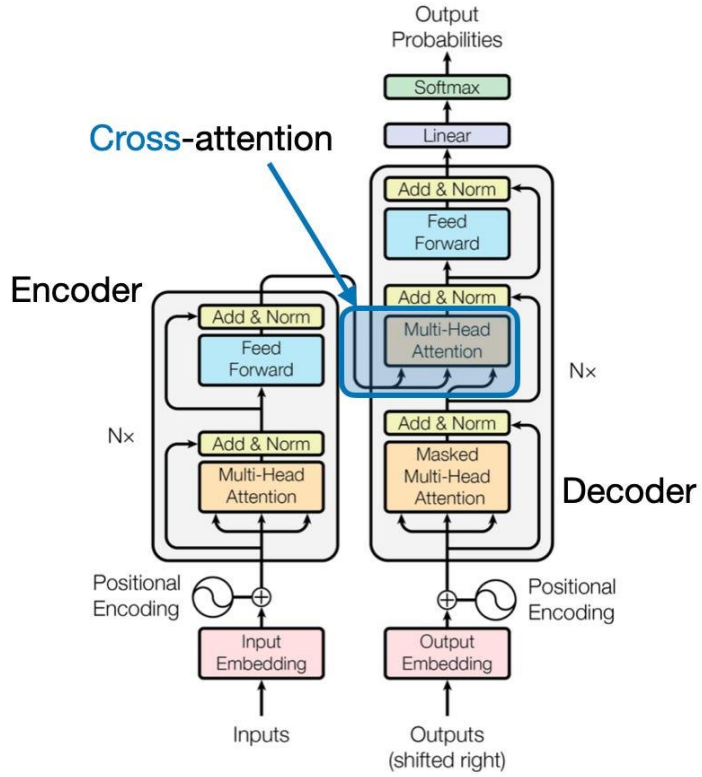
Why don't we add any extra info to the input to the encoder?

Additional information can be concatenated or used in cross attention

Cross attention in Language Diffusion Models

Cross-attention computes attention scores between text and image encoding

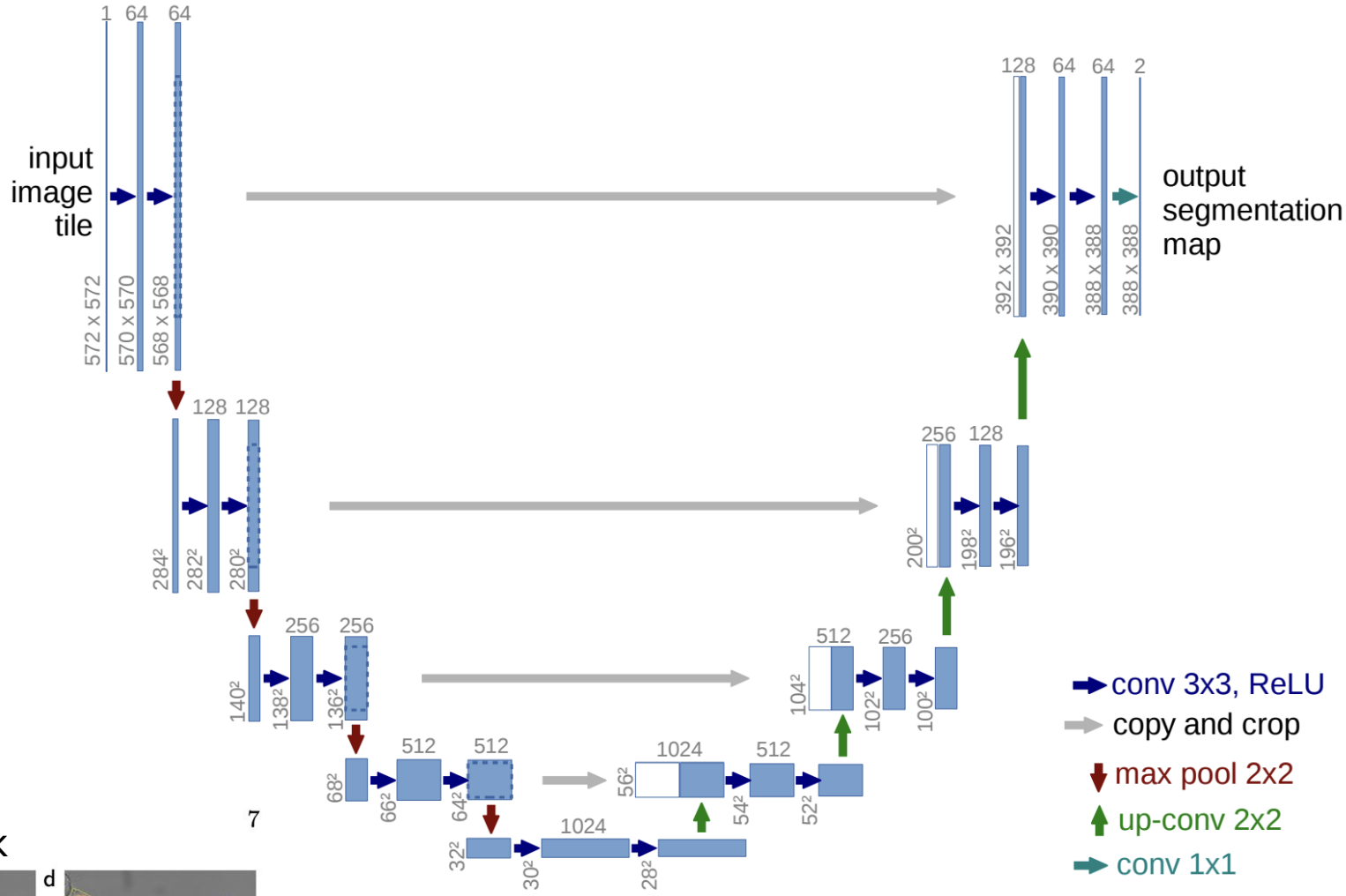




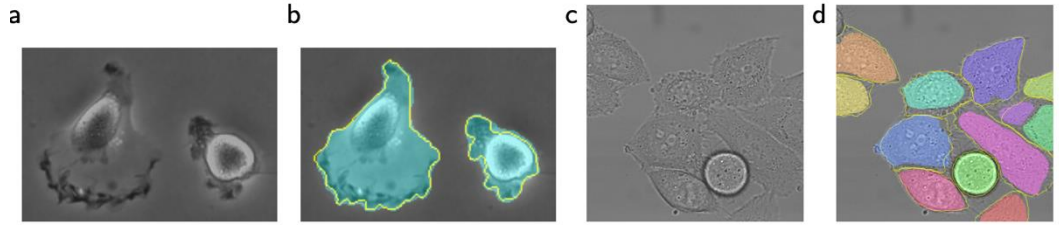
Source: “Attention Is All You Need” (https://arxiv.org/abs/1706.03762)

U-NET

Convolutional network originally designed for segmentation



Segmentation Task



Source: U-Net: Convolutional Networks for Biomedical Image Segmentation

Quick Detour into segmentation

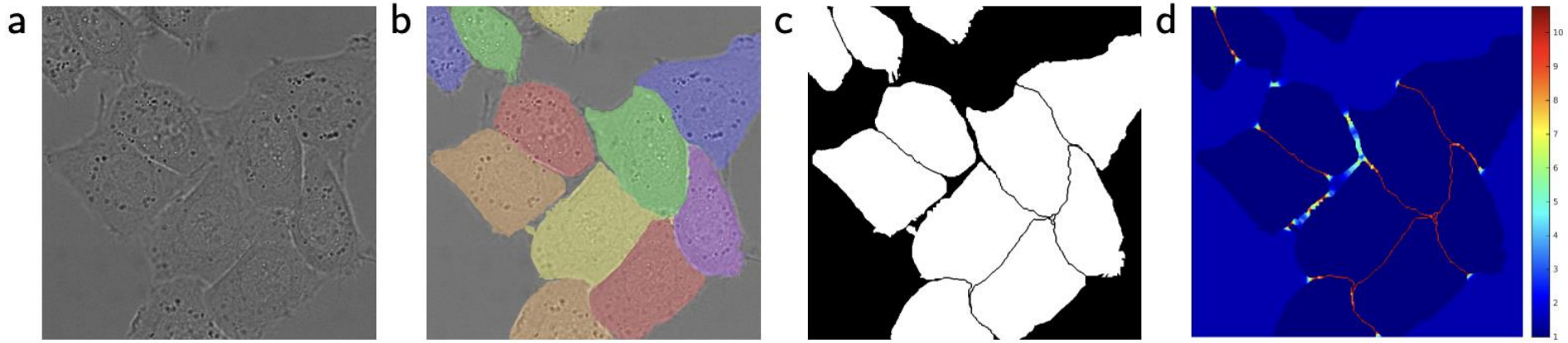


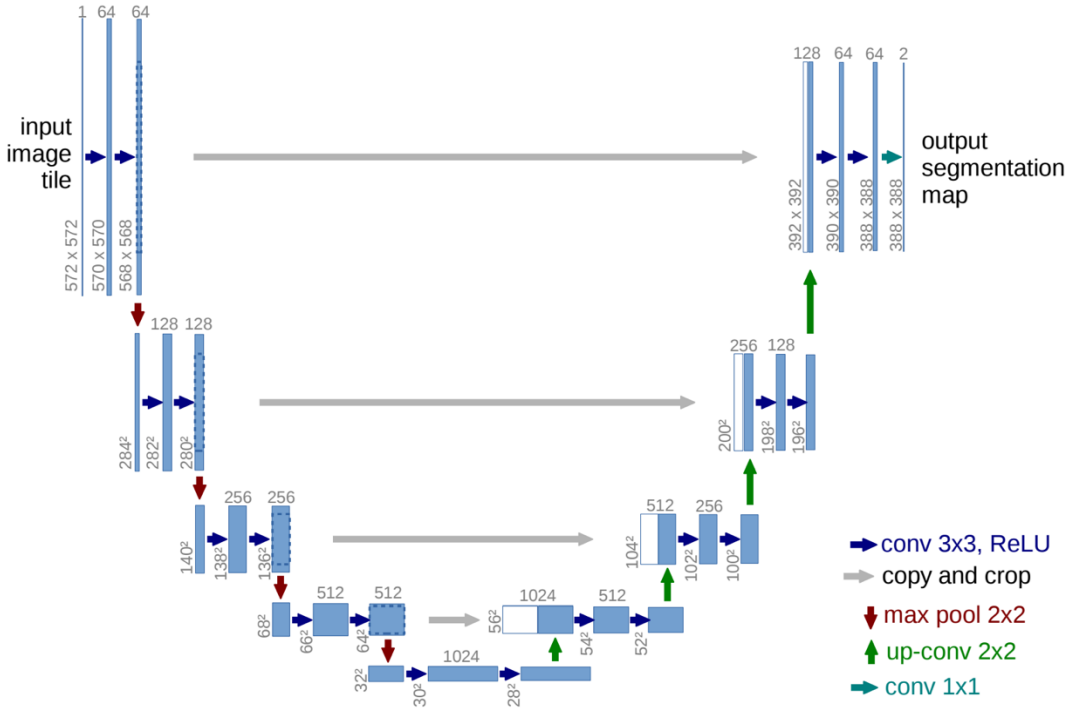
Fig. 3. HeLa cells on glass recorded with DIC (differential interference contrast) microscopy. (**a**) raw image. (**b**) overlay with ground truth segmentation. Different colors indicate different instances of the HeLa cells. (**c**) generated segmentation mask (white: foreground, black: background). (**d**) map with a pixel-wise loss weight to force the network to learn the border pixels.

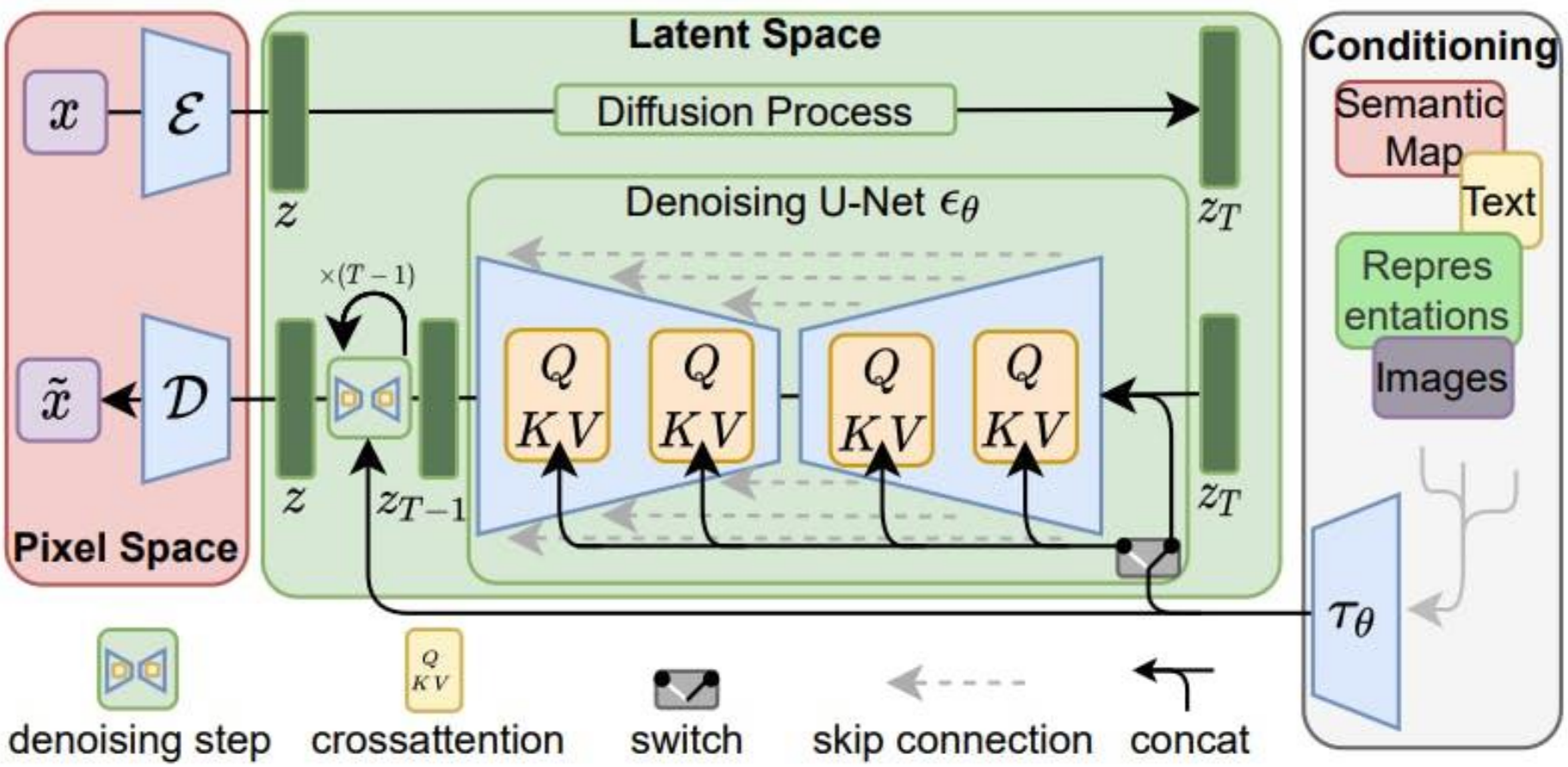
U-Nets

Last class, we were agnostic on what the neural network in a diffusion model actually is…

Why use a U-net?

1. Convolutions > MLPs
2. Learn hierarchical features (with skip/residual connections)
3. Good at handling multi-scale information





And how do we find the vector τ_θ?

Depends on the task. At least in the original paper that trains separate models for different tasks. If you had enough data you can train a single model for lots of different types of inputs and tasks.

Step 1: Train a language model

Step 2: Train diffusion model conditioned on language model

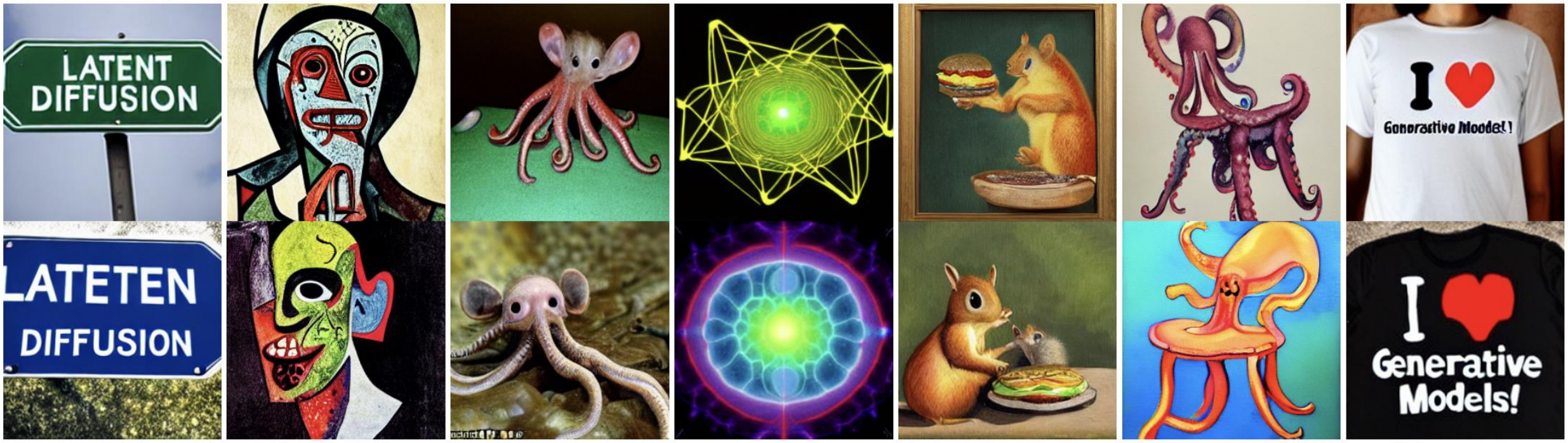
(requires dataset of captioned images)

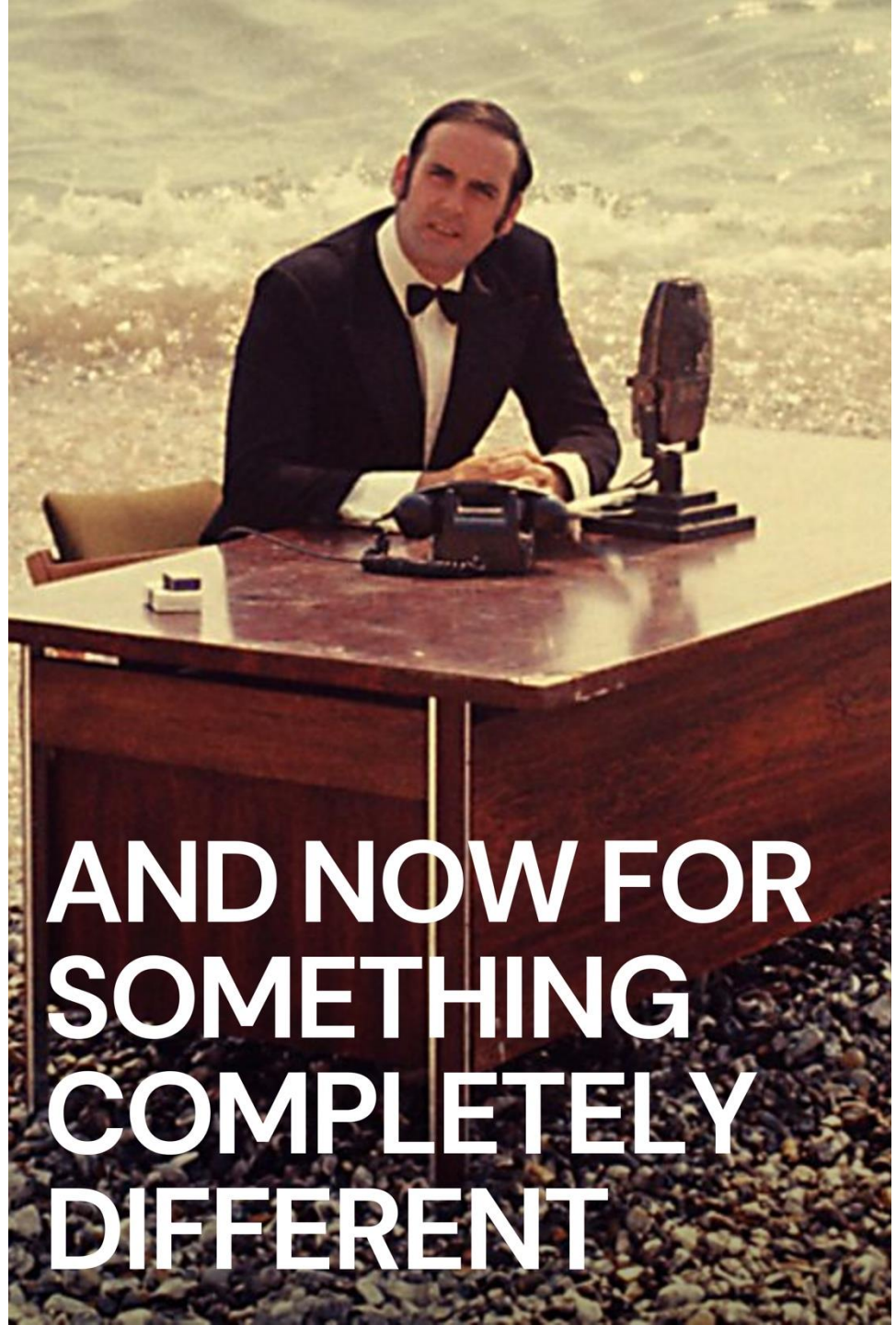
?

?

Text-to-Image Synthesis on LAION. 1.45B Model.

'A street sign that reads "Latent Diffusion" '	*'A zombie in the style of Picasso'*	*'An image of an animal half mouse half octopus'*	*'An illustration of a slightly conscious neural network'*	*'A painting of a squirrel eating a burger'*	*'A watercolor painting of a chair that looks like an octopus'*	*'A shirt with the inscription: "I love generative models!" '*

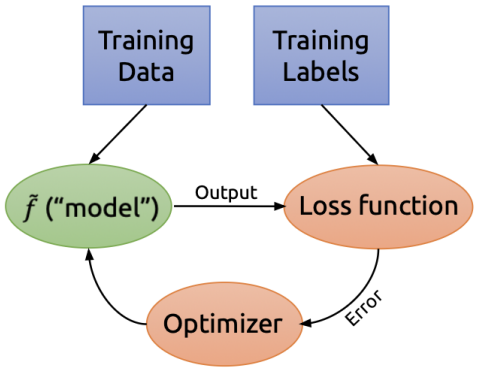



AND NOW FOR
SOMETHING
COMPLETELY
DIFFERENT

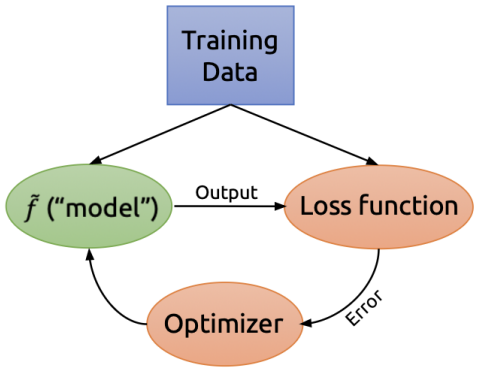
What we've done so far

Different Learning Paradigms

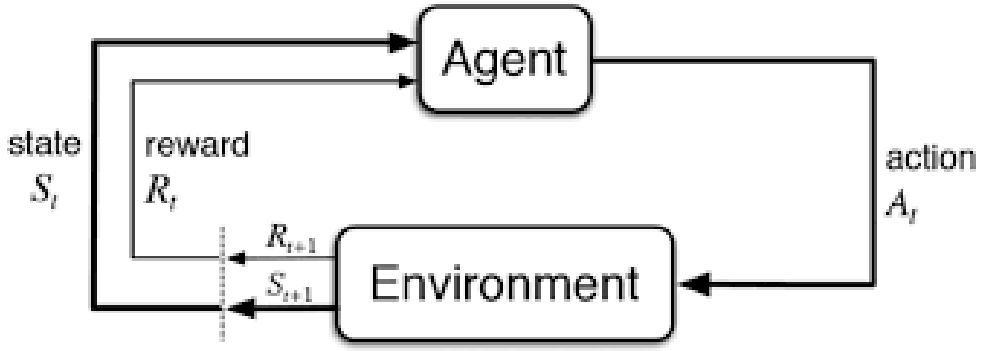
Supervised Learning



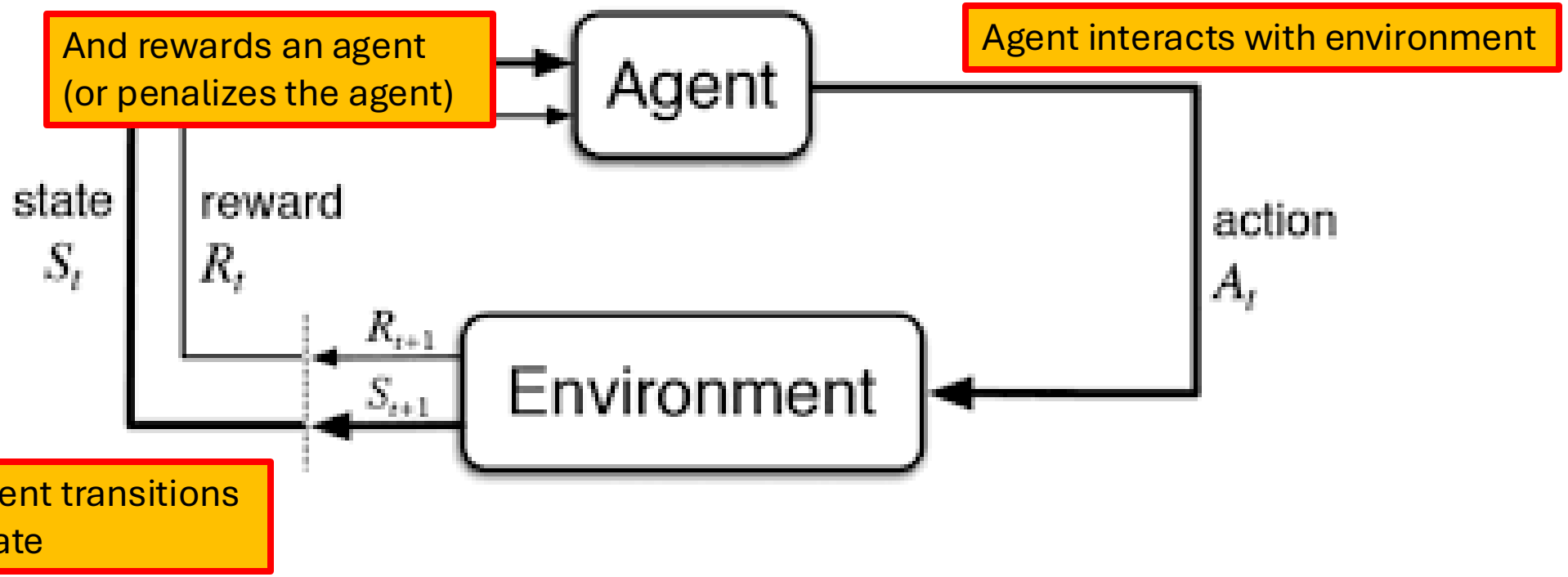
Unsupervised Learning



Reinforcement Learning



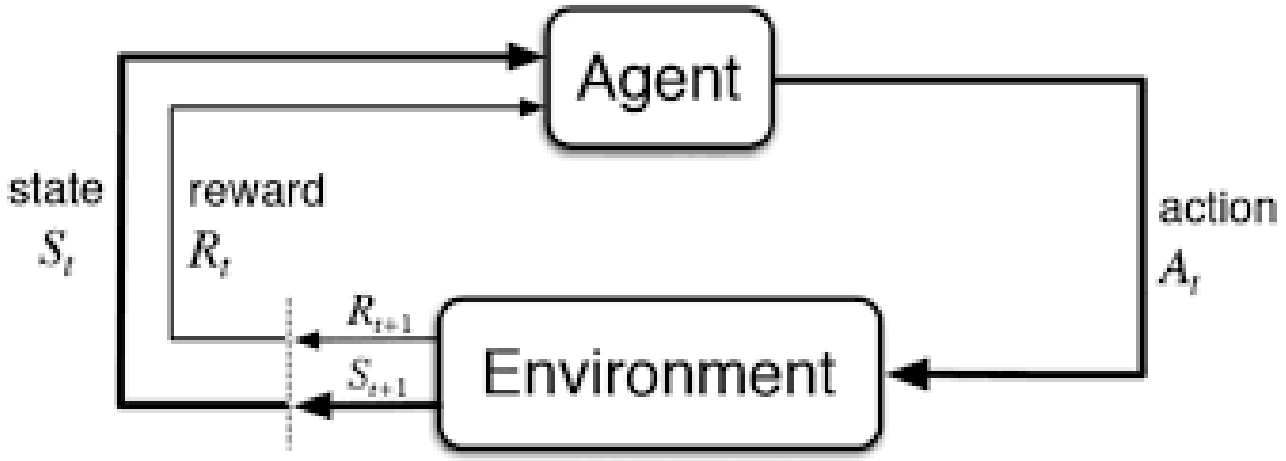
We've focused on this thus far...


And rewards an agent
(or penalizes the agent)
Agent
Agent interacts with environment
state
S_t
reward
R_t
action
A_t
R_{t+1}
S_{t+1}
Environment
Environment transitions
to new state

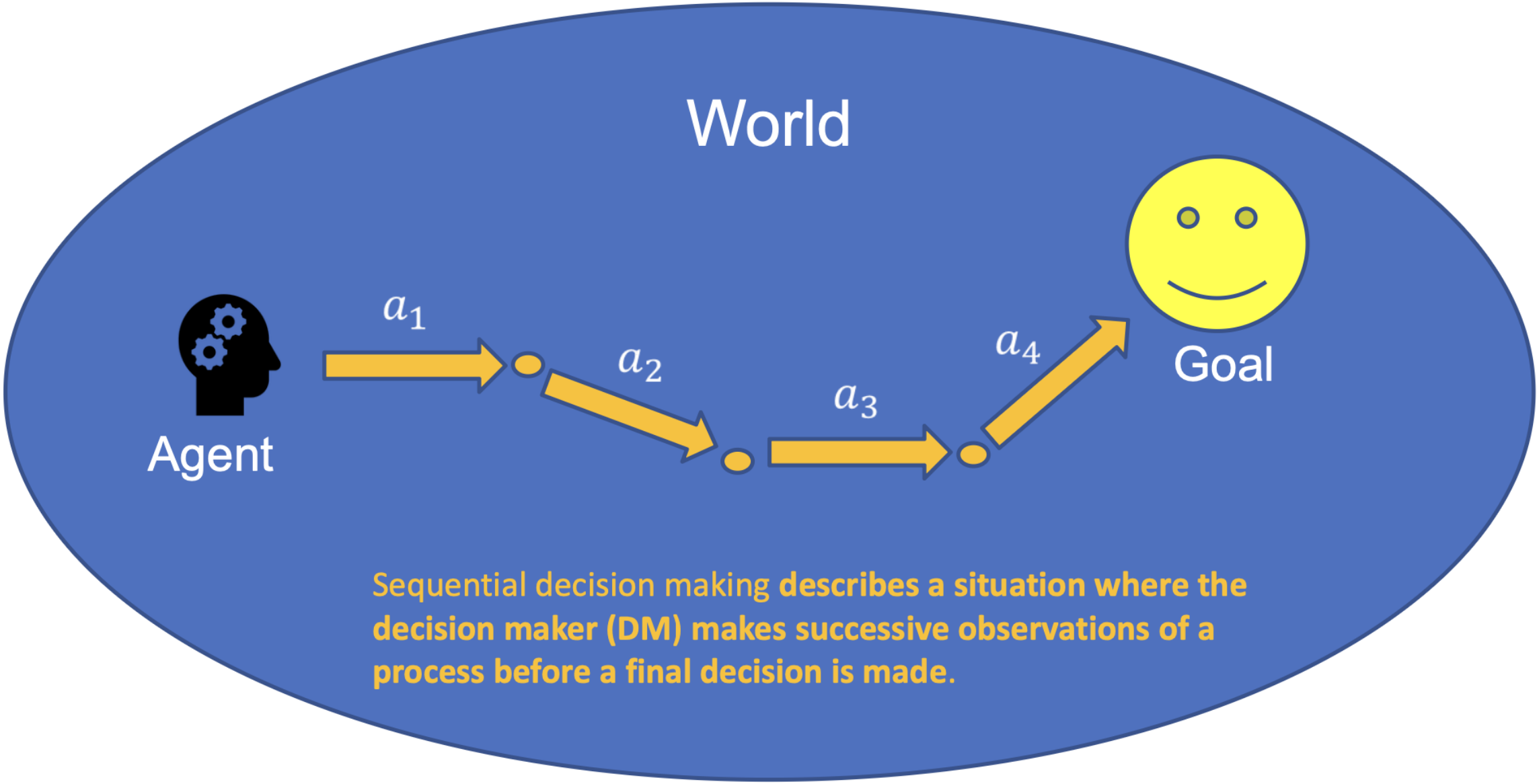
Why Reinforcement Learning?

- Reinforcement learning doesn't require data in the same way that supervised and unsupervised learning do
- There is no dataset *X* required, just a model of the environment
- Agents learn from interacting with the environment

This is how you got so smart…



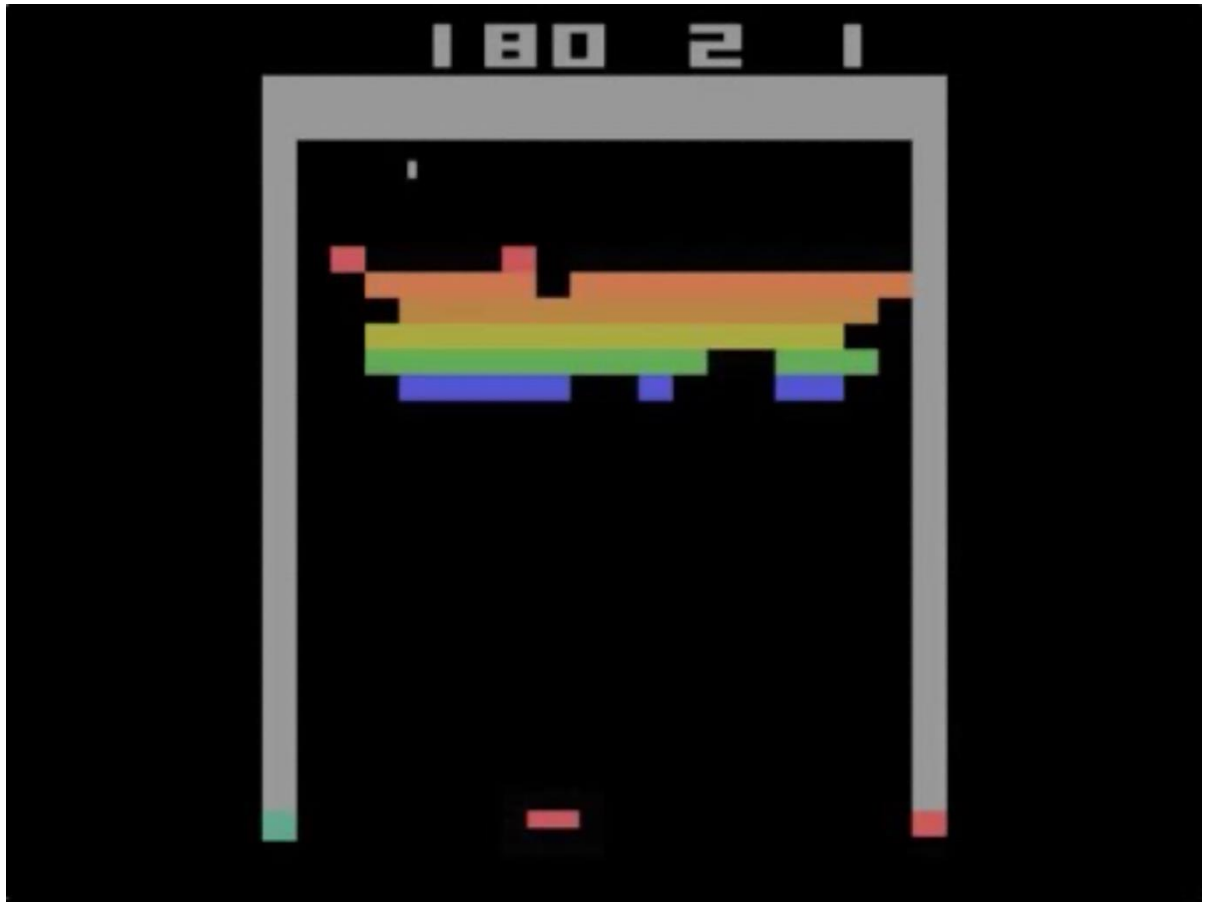
RL: Sequential Decision Making



What's a common example of a sequential decision making process?

- Playing games!
- Let’s look at a specific example…

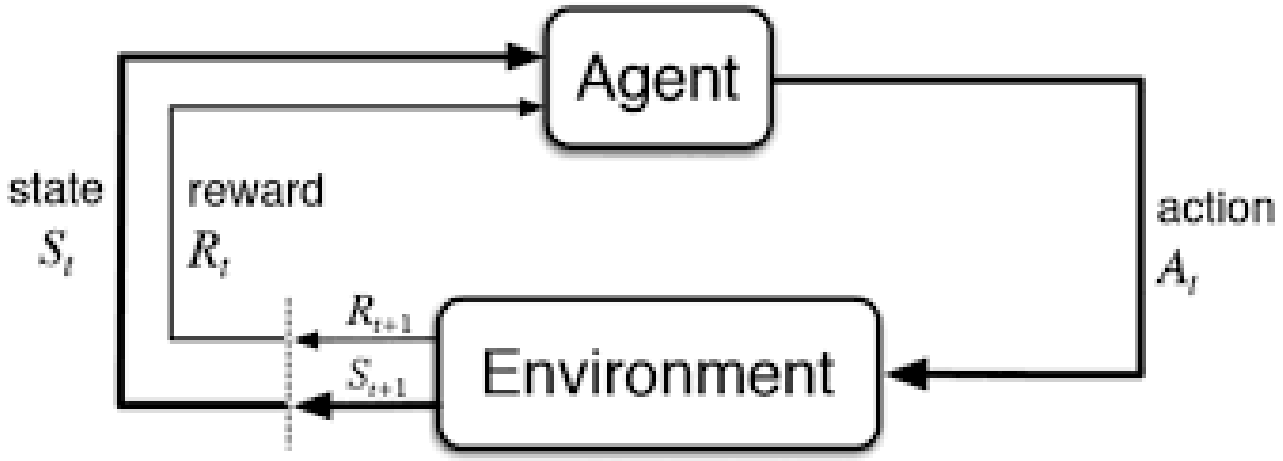
Atari!



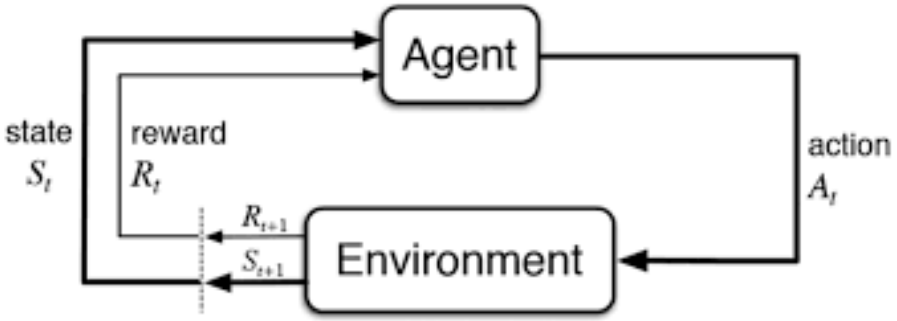
Mnih et al. Playing Atari with Deep Reinforcement Learning, 2013

Markov Decision Processes (MDPs)

How can we formalize the problem we are trying to solve? What components does it have?



Markov Decision Processes (MDPs)



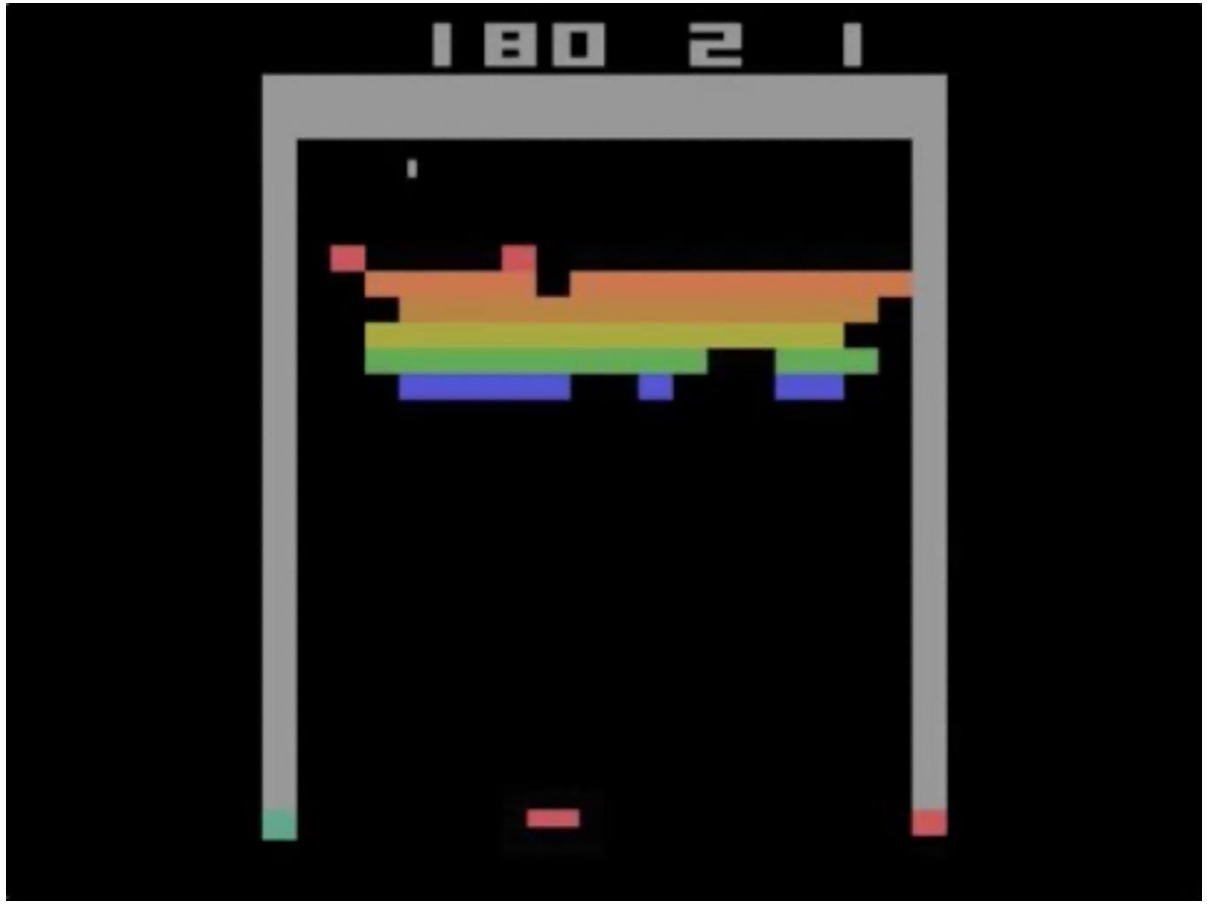
- Set of States: S
  - All possible configurations the world can be in
- Set of Actions: A
  - All possible actions the agent is able to take
- Reward Function: R: $S \rightarrow \mathbb{R}$
  - Reward function takes in a state and returns a number
- Transition Function: T: $S \times A \times S \rightarrow \mathbb{R}$
  - If you take an action in a specific state, what's the probability you transition to any other state?

States

What would the state for breakout be?

Option: Location of paddle, ball, and all breakable blocks

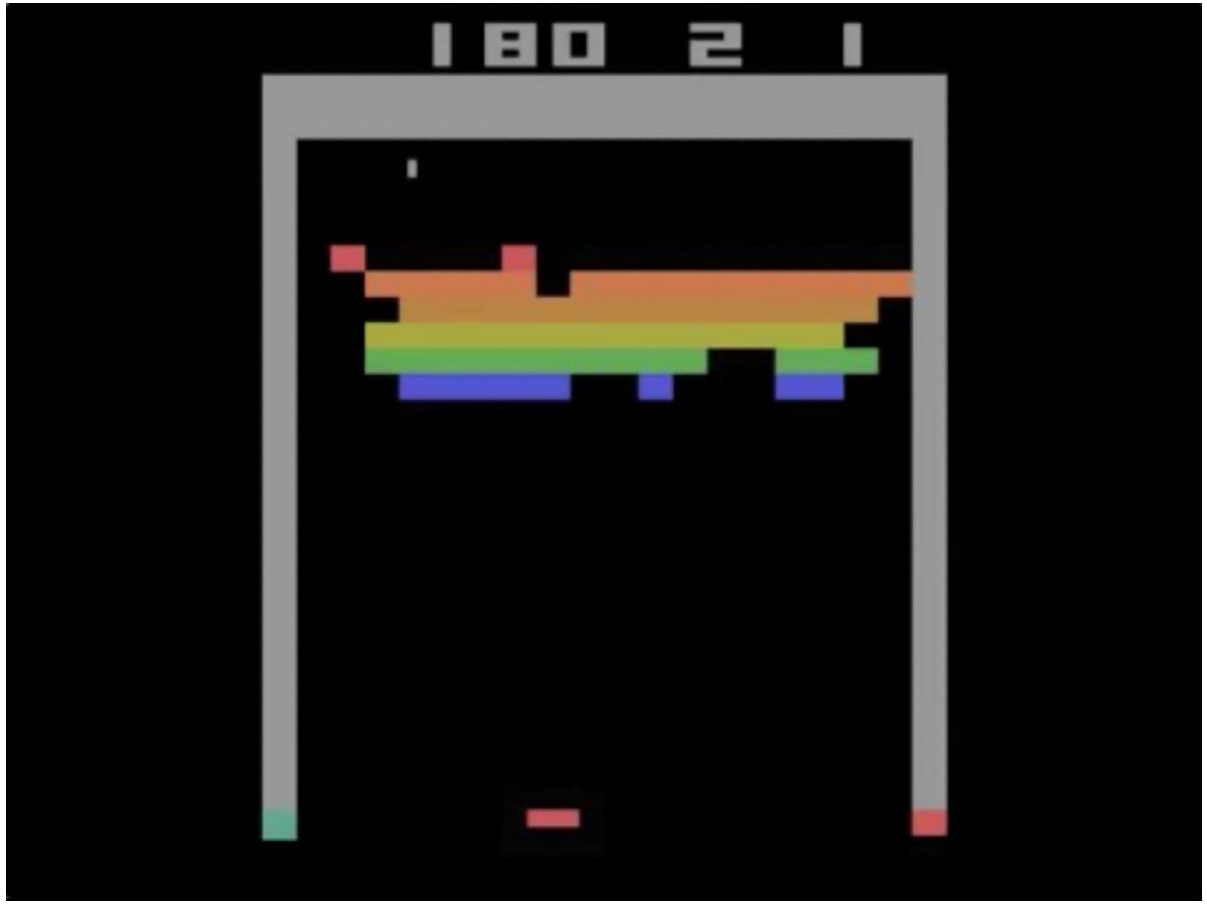
Option: The image of the game…



Actions

What actions can the agent take?

A = Left or Right



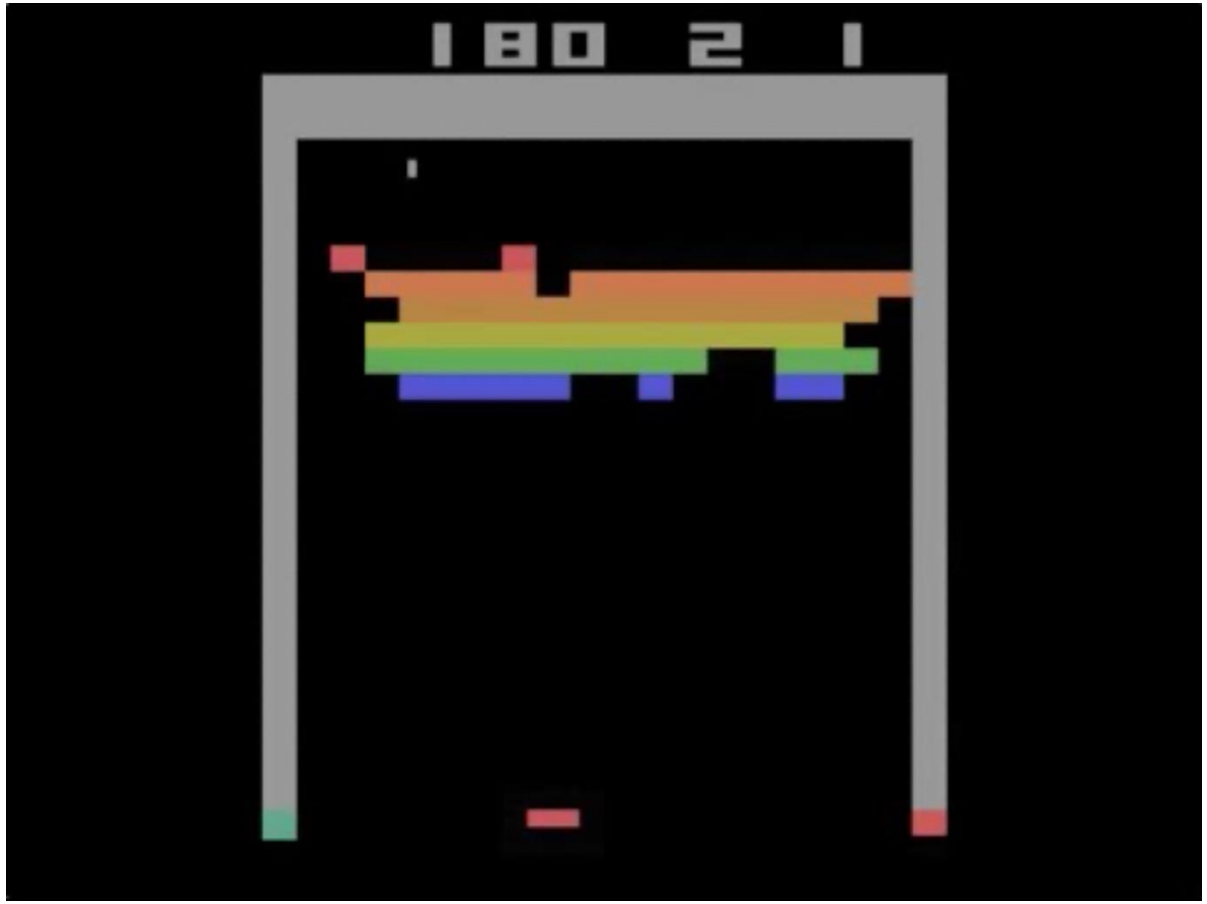
Reward Function

What is the reward function?

There is no predefined reward function necessarily

We can use:

1. The score (get reward when a block is broken)
2. Large penalty for losing, Large reward for winning
3. And many other combinations of things



Transition Function

In general, MDPs describe stochastic processes. There can be randomness in what happens.

Breakout is deterministic, the physics of the ball is known and when you tell your paddle to go left it goes left.

Solving MDPs

What would it mean to solve an MDP, like breakout?

Policy: A function $\pi: S \rightarrow A$, that takes in a state and returns an action

We seek the best possible policy π^*, that could tell us the best action to take in any state.

But how do we know one policy is better than another?
If we try to learn a policy, what would our loss function be?
And many many more remaining questions… for next time