

## Overview of Image Generation

---

AI-generated images are changing the way people create art, design, and digital content. With the help of artificial intelligence, computers can now make realistic pictures, improve creative work, and even help businesses. In this chapter, we'll explore how AI creates images, different types of image-making models, and how they are used in real life.

### How AI Creates Images

AI image generation works by learning from a huge collection of pictures. The AI studies patterns in the images and then creates new ones that look similar. This technology has improved a lot over the years, making images that are more realistic and creative. It is now used in video games, movies, advertising, and even fashion.

### Early Methods: PixelRNN and PixelCNN

Before today's advanced AI models, researchers created early image-generation methods like PixelRNN and PixelCNN. These models created images by predicting one pixel at a time.

- **PixelRNN:** uses a system called a recurrent neural network (RNN) to predict pixel colors one after another. While it worked well, it was very slow;
- **PixelCNN:** improved on PixelRNN by using a different type of network, called convolutional layers, which made image creation faster.

Even though these models were a good start, they weren't great at making high-quality images. This led to the development of better techniques.

### Autoregressive Models

Autoregressive models also create images one pixel at a time, using past pixels to guess what comes next. These models were useful but slow, which made them less popular over time. However, they helped inspire newer, faster models.

## How AI Understands Text for Image Creation

Some AI models can turn written words into pictures. These models use **Large Language Models (LLMs)** to understand descriptions and generate matching images. For example, if you type “a cat sitting on a beach at sunset,” the AI will create a picture based on that description.

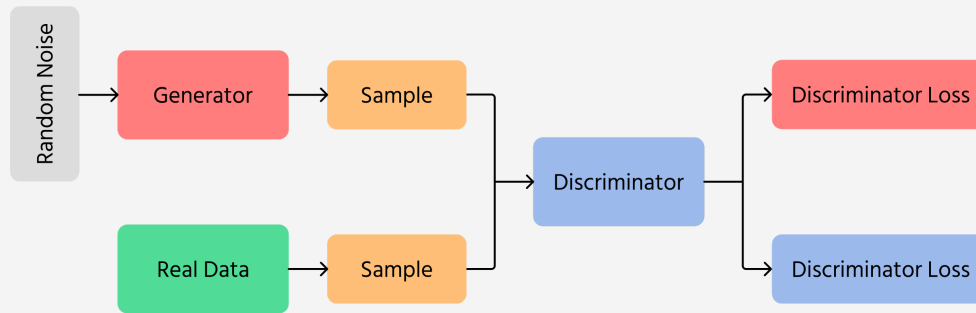
AI models like OpenAI's **DALL-E** and Google's **Imagen** use advanced language understanding to improve how well text descriptions match the images they generate. This is possible through **Natural Language Processing (NLP)**, which helps AI break down words into numbers that guide image creation.

## Generative Adversarial Networks (GANs)

One of the most important breakthroughs in AI image generation was **Generative Adversarial Networks (GANs)**. GANs work by using two different neural networks:

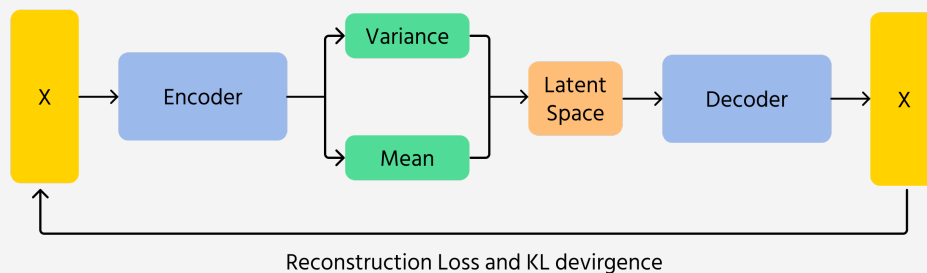
- **Generator:** creates new images from scratch;
- **Discriminator:** checks if the images look real or fake.

The generator tries to make images so realistic that the discriminator can't tell they are fake. Over time, the images improve and look more like real photographs. GANs are used in deepfake technology, artwork creation, and improving image quality.



## Variational Autoencoders (VAEs)

VAEs are another way AI can generate images. Instead of using competition like GANs, VAEs encode and decode images using probability. They work by learning the underlying patterns in an image and then reconstructing it with slight variations. The probabilistic element in VAEs ensures that each generated image is slightly different, adding variety and creativity.



A key concept in VAEs is **Kullback-Leibler (KL) divergence**, which measures the difference between the learned distribution and a standard normal distribution. By minimizing KL divergence, VAEs ensure that generated images remain realistic while still allowing creative variations.

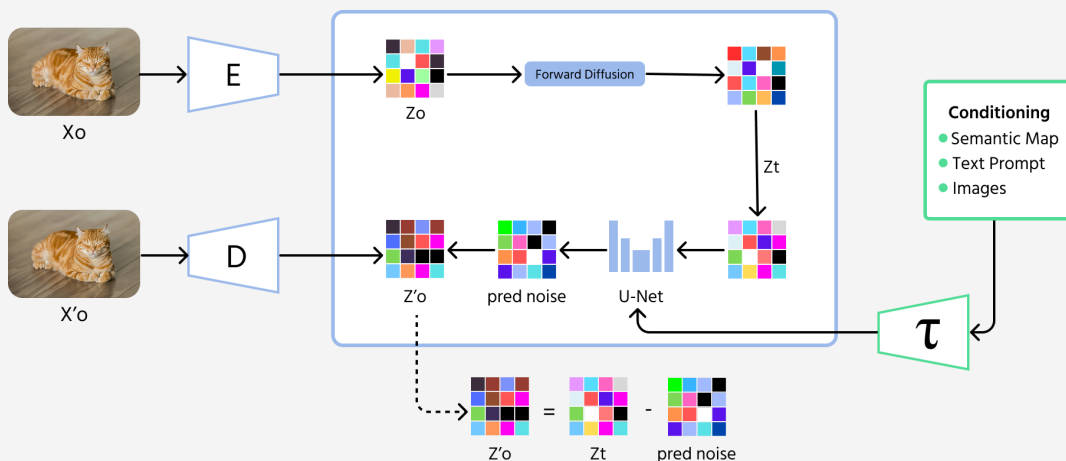
## How VAEs Work

1. **Encoding:** the input data  $x$  is fed into the encoder, which outputs the parameters of the latent space distribution  $q(z|x)$  (mean  $\mu$  and variance  $\sigma^2$ );
2. **Latent space sampling:** latent variables  $z$  are sampled from the distribution  $q(z|x)$  using techniques like the reparameterization trick;
3. **Decoding & reconstruction:** the sampled  $z$  is passed through the decoder to produce the reconstructed data  $x'$ , which should be similar to the original input  $x$ .

VAEs are useful for tasks like reconstructing faces, generating new versions of existing images, and even making smooth transitions between different pictures.

## Diffusion Models

Diffusion models are the latest breakthrough in AI-generated images. These models start with random noise and gradually improve the picture step by step, like erasing static from a blurry photo. Unlike GANs, which sometimes create limited variations, diffusion models can produce a wider range of high-quality images.



## How Diffusion Models Work

1. **Forward process (noise addition):** the model starts by adding random noise to an image over many steps until it becomes completely unrecognizable;
2. **Reverse process (denoising):** the model then learns how to reverse this process, gradually removing the noise step by step to recover a meaningful image;
3. **Training:** diffusion models are trained to predict and remove noise at each step, helping them generate clear and high-quality images from random noise.

A popular example is MidJourney, DALL-E and Stable Diffusion, which is known for making realistic and artistic images. Diffusion models are widely used for AI-generated art, high-resolution image synthesis, and creative design applications.

## Examples of Images Generated by Diffusion Models



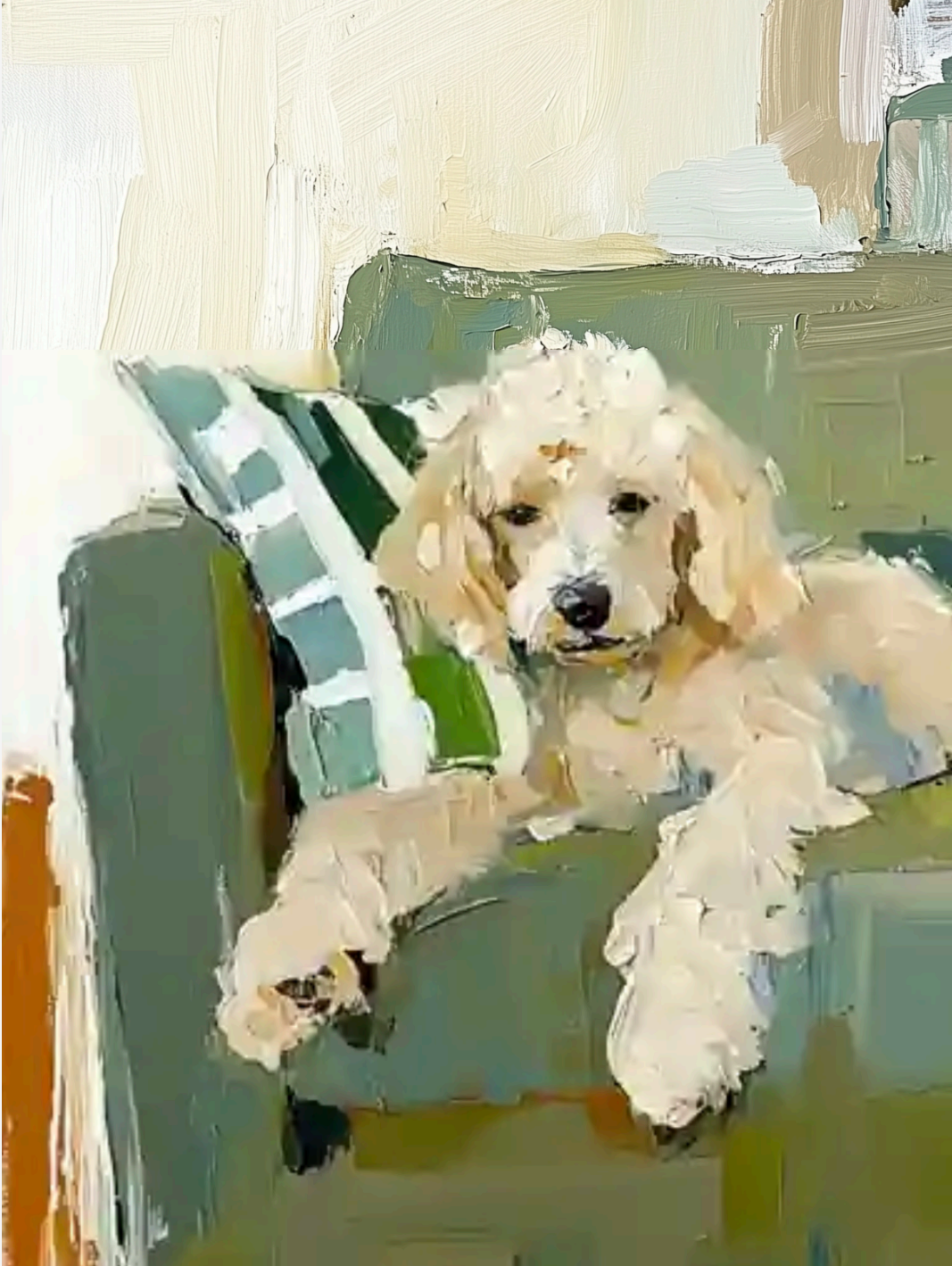
Prompt





Prompt





Prompt





Prompt



## Challenges and Ethical Concerns

Even though AI-generated images are impressive, they come with challenges:

- **Lack of control:** AI might not always generate exactly what the user wants;
- **Computing power:** creating high-quality AI images requires expensive and powerful computers;
- **Bias in AI models:** since AI learns from existing images, it can sometimes repeat biases found in the data.

There are also **ethical concerns**:

- **Who owns AI art?:** if an AI creates an artwork, does the person who used the AI own it, or does it belong to the AI company?
- **Fake images and deepfakes:** GANs can be used to create fake images that look real, which can lead to misinformation and privacy issues.

## How AI Image Generation is Used Today

AI-generated images are already making a big impact in different industries:

- **Entertainment:** video games, movies, and animation use AI to create backgrounds, characters, and effects;
- **Fashion:** designers use AI to create new clothing styles, and online stores use virtual try-ons for customers;
- **Graphic design:** AI helps artists and designers quickly make logos, posters, and marketing materials.

## The Future of AI Image Generation

As AI image generation keeps improving, it will continue to change the way people create and use images. Whether in art, business, or entertainment, AI is opening new possibilities and making creative work easier and more exciting.