Generative AI

Diving deeper into the rabbit hole of AI

Objectives

Understand how does generative AI work

Understand the terminology used in Generative AI

How to create more control to generate images

How generative AI works - Text to Image

Input Data → Latent Diffusion Model → Generated Example

Realistic concept art, space giant rabbit civilization, sci-fi, high tech society, full of rabbit-like species

sd_xl_base_1.0_0.9vae.safetensors

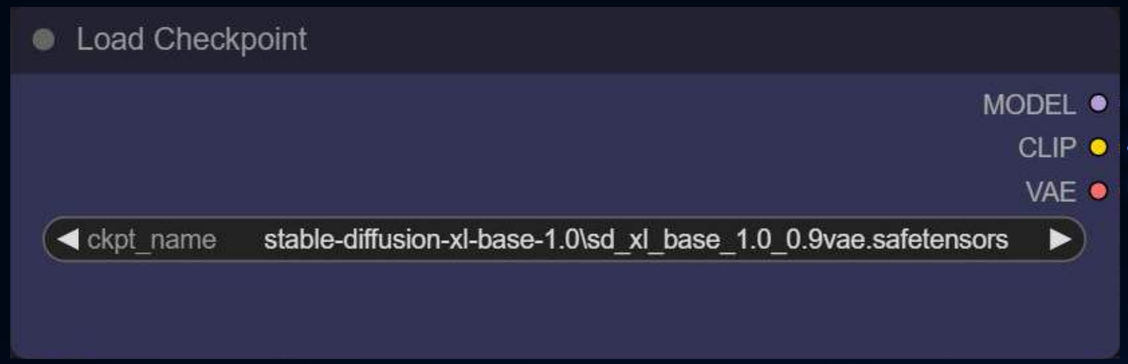
--

Diffusion-based text-to-image generative model

*Super duper simplified flow

Common Terminologies

While picking up image generative AI software, there are some terms that often show up while using them. Having a better understanding of these term will help facilitate more control over the images we want to generate



Checkpoint(ckpt) and safetensors are model weights in machine learning. We need these to learn and make predictions.

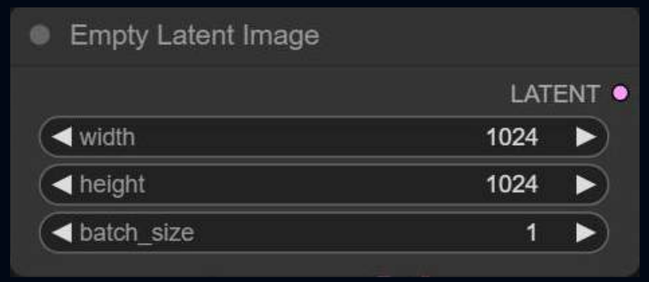
Think of weights as ingredients in a recipe. The right amount of each ingredient (weight) makes the recipe (model) successful.

CLIP stands for **Contrastive Language–Image Pre-Training**. It combines text and image understanding in a new way.

CLIP can be used for a wide range of tasks, such as image classification, generating captions for images, and finding images based on text queries.

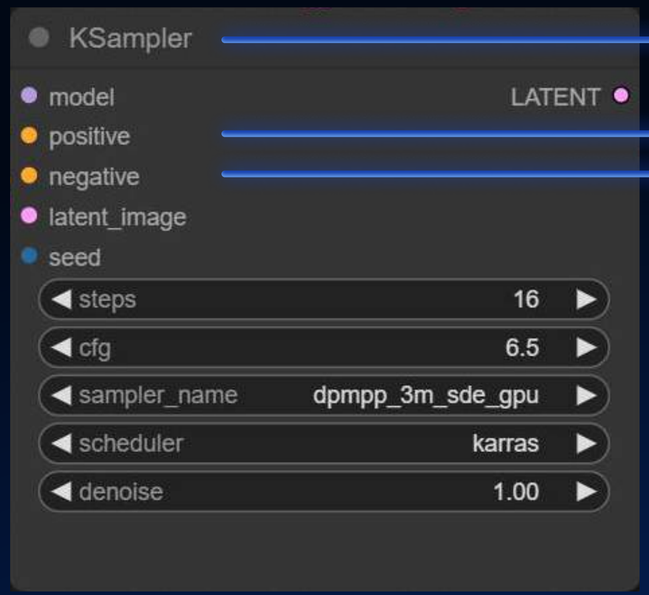
VAE stands for **Variational Autoencoder**.
It's a type of neural network that generates new data that is similar to the data it was trained on.

Common Terminologies



Latent refers to the **hidden, compressed representation** of an image that a model, like a Variational Autoencoder (VAE) model would use.

Basically an empty template



Sampling is the process of **generating new data from a model**

It’s used to create new images by selecting points from the **latent space** and decoding them into complete images.

Positive Prompts shows what you want to see

Negative Prompts hide what you don’t want to see

Common Terminologies

ControlNet is a specialised neural network architecture designed to add more control and precision to image generation processes.

Think of is as extra instructions to a drawing robot to make sure it draws exactly what you want.

LoRAs stands for **Low-Rank Adaptation.** It's a technique to fine-tune large pre-trained models efficiently without having to change or retrain the whole model

Safetensors are related to saving and loading model weights in machine learning. It's similar to checkpoints but it's a newer format specifically for storing tensors in a safe and efficient manner.

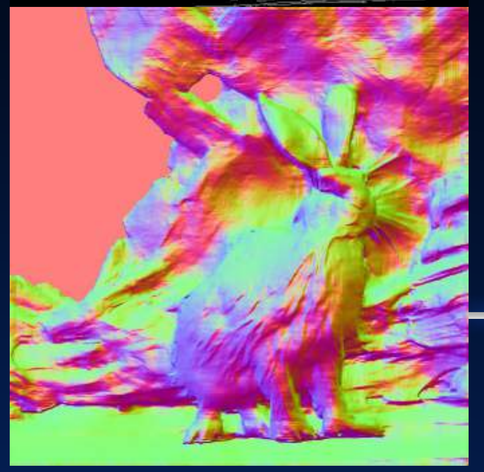
*Tensors are like multi-dimensional grids of numbers

How ControlNet Works

Depth Detection

Line Detection

Light Detection



How ControlNet Works

Depth Detection

New Prompt

New Prompt

Line Detection

How ControlNet Works

Controlling AI generation is possible
just by giving a different input

How ControlNet Works

Use cases for evyd?

As a health-tech company, text to image generation might not have impactful use cases. However, we can use it for internal sharing, deck building and at most for marketing purposes.

Thoughts

After deep diving into the rabbit hole of Generative AI, the possibilities are endless.

Anyone can create beautiful images, songs and art.

AI in UXUI will soon become a norm when UI screen designs can be generated, UX flows can be create with just a click of a button from a prompt.

How will we set ourselves apart as UXUI designers?