



Stable Diffusion: ChatGPT in Art

KYOSS
Liang Yan

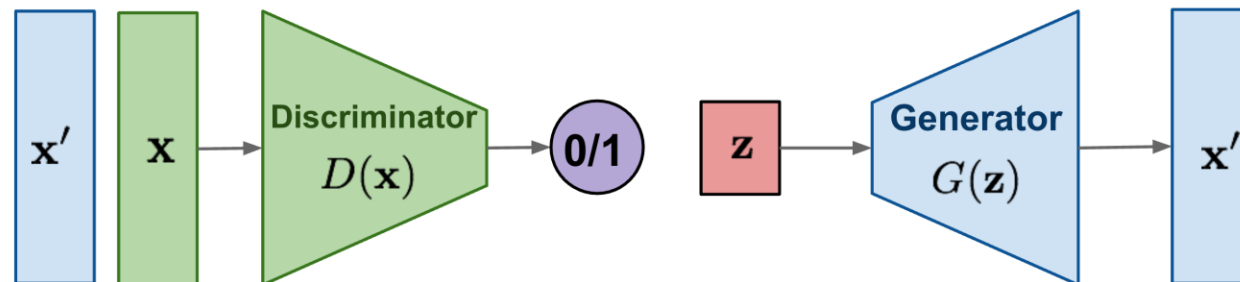


Stable Diffusion

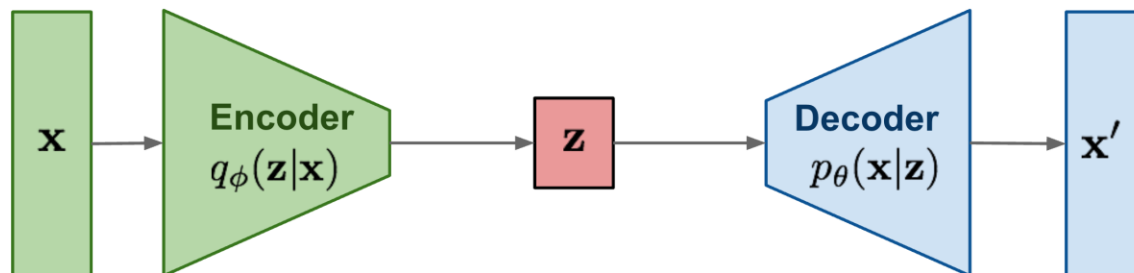
- An image generation AI
- Diffusion Model, with billion parameters
- Dataset LAION-5B
- Stability AI, weights is hosted by Hugging Face, research by Runway, CompVis lab
- AGI: Artificial generative intelligence
 - GAN: Generative Adversarial Network
 - VAE: Variational AutoEncoder
 - DDPM: Denoising Diffusion Probabilistic Models

AGI Model

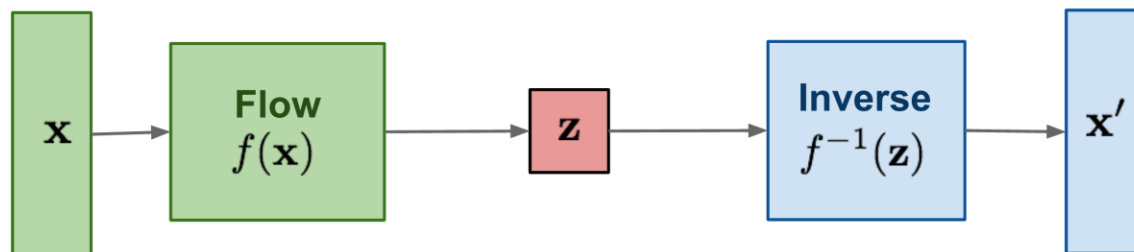
GAN: Adversarial training



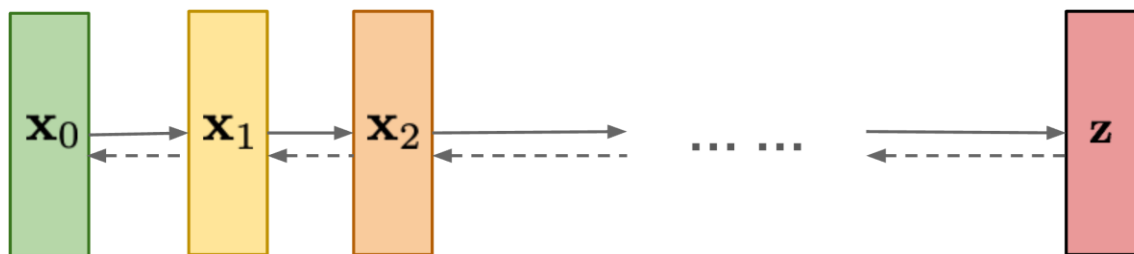
VAE: maximize variational lower bound



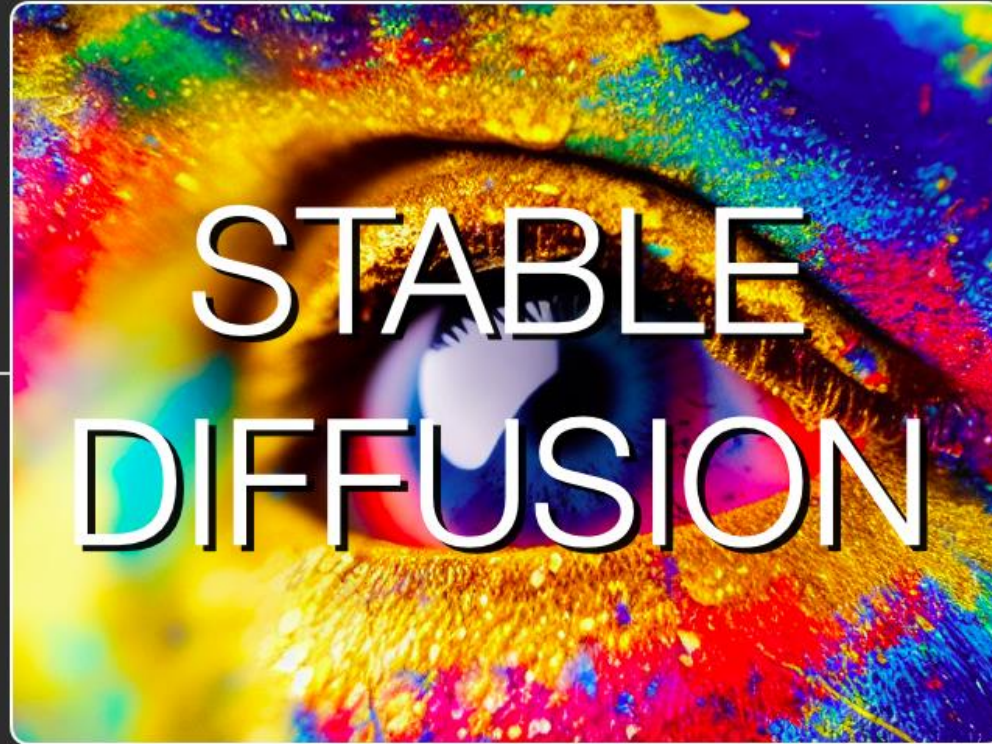
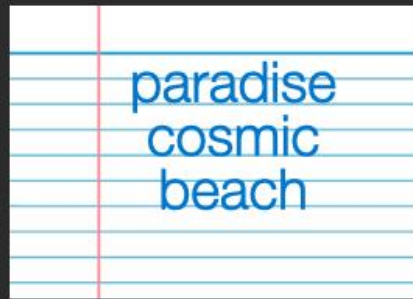
Flow-based models:
Invertible transform of distributions



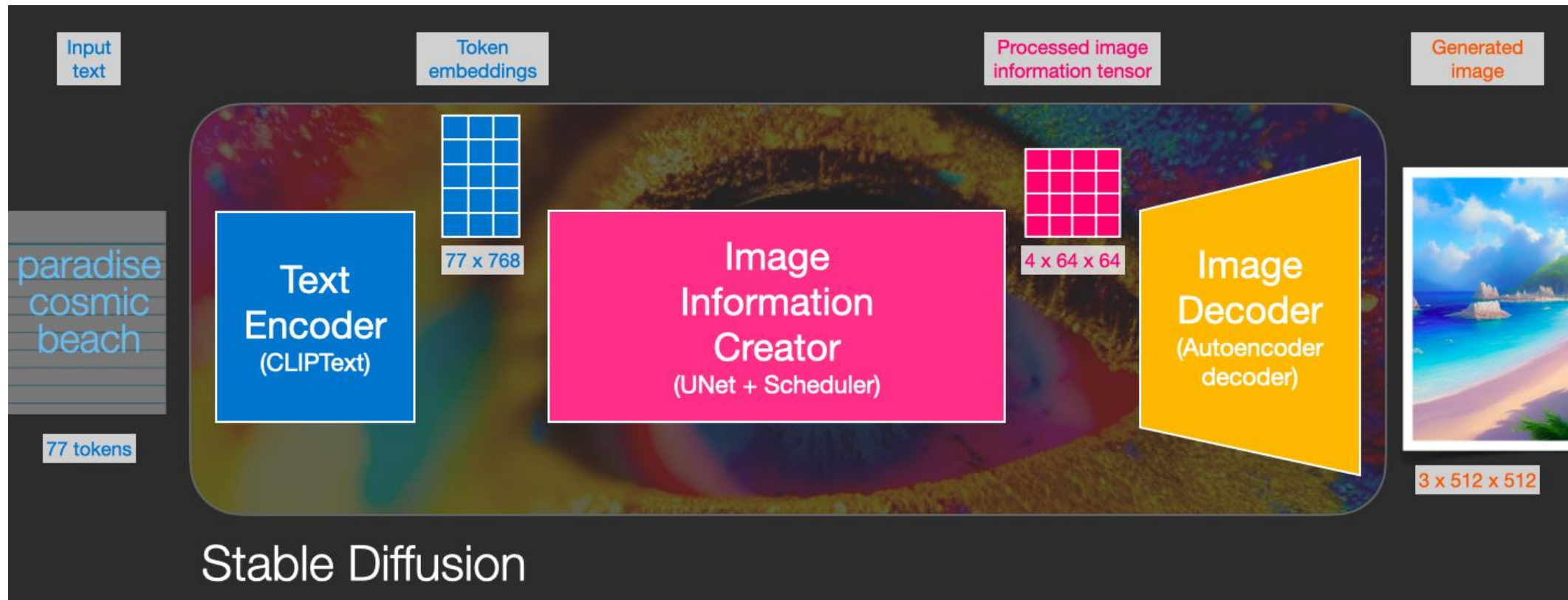
Diffusion models:
Gradually add Gaussian noise and then reverse



Stable Diffusion



Stable Diffusion



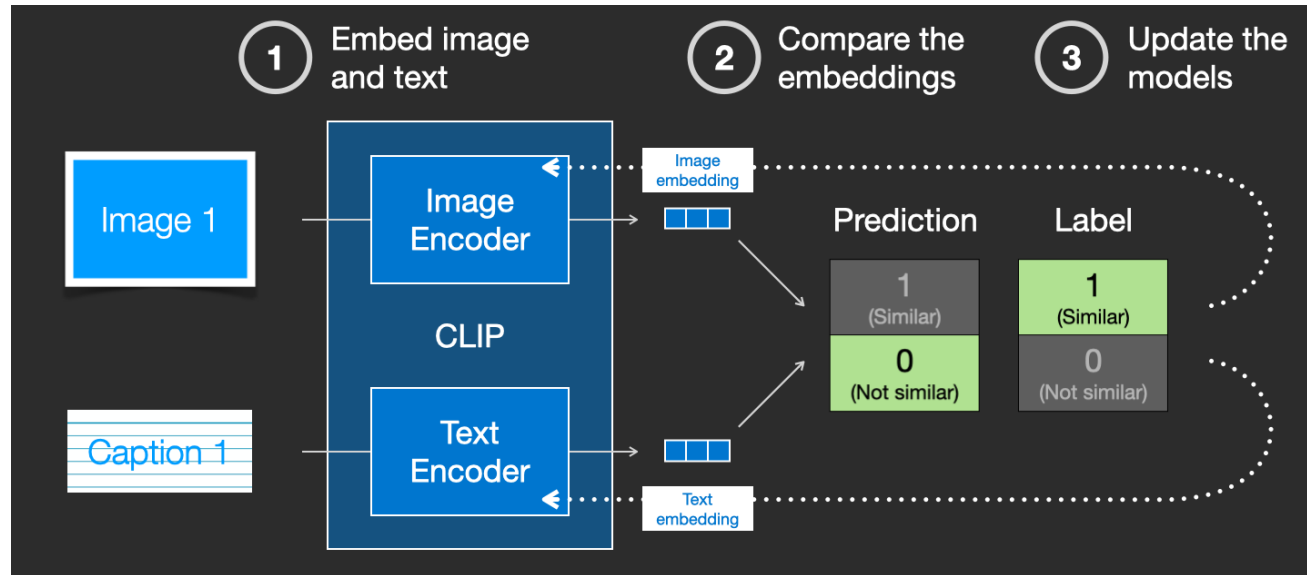
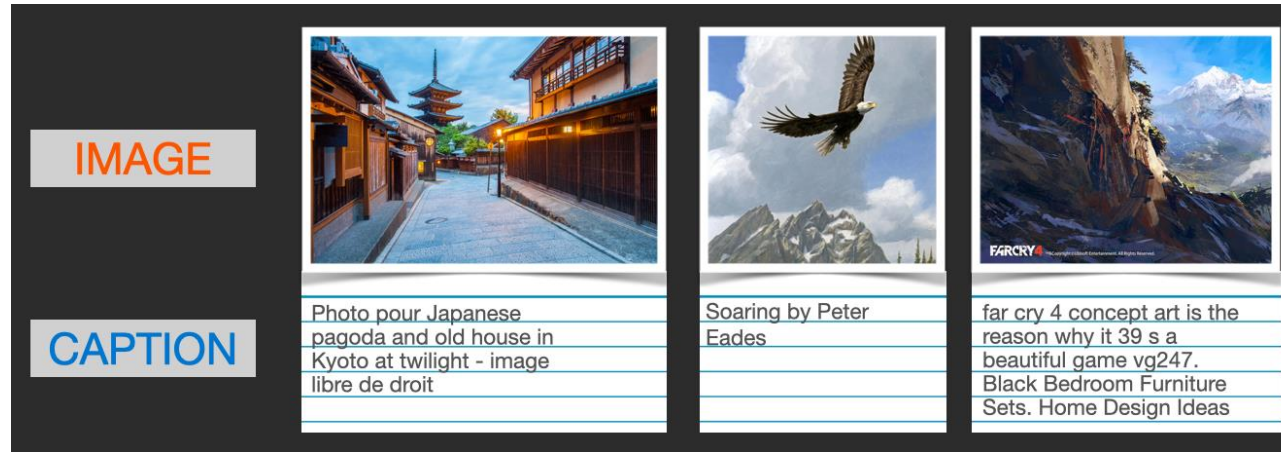
<https://jalammar.github.io/illustrated-stable-diffusion/>

Stable Diffusion

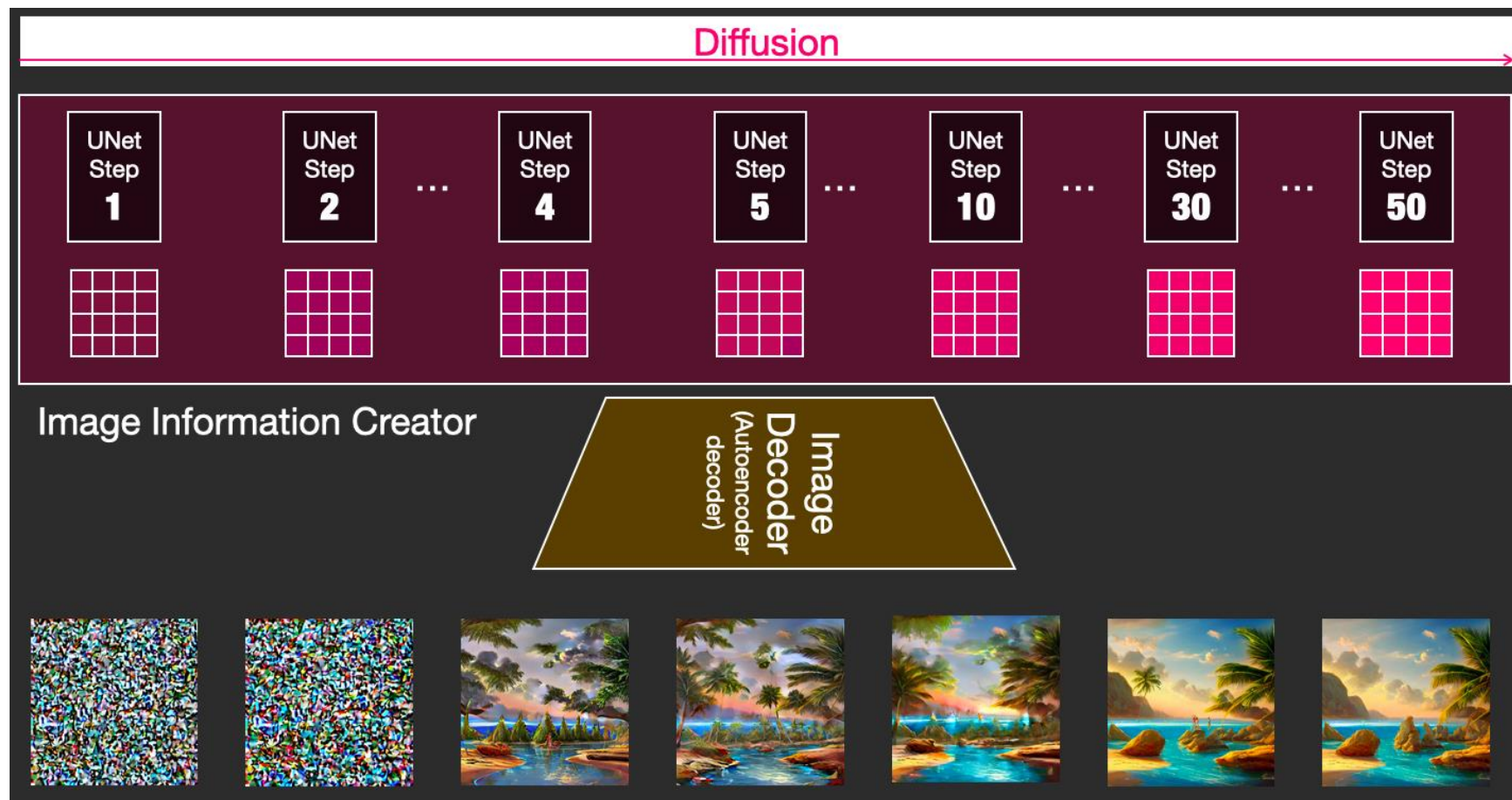
- CLIP based on GPT: Transformer
- Image information Creator:
 - Unet(Noise Predictor)
 - Scheduler
- Image Decoder: VAE

<https://jalammar.github.io/illustrated-stable-diffusion/>

CLIP

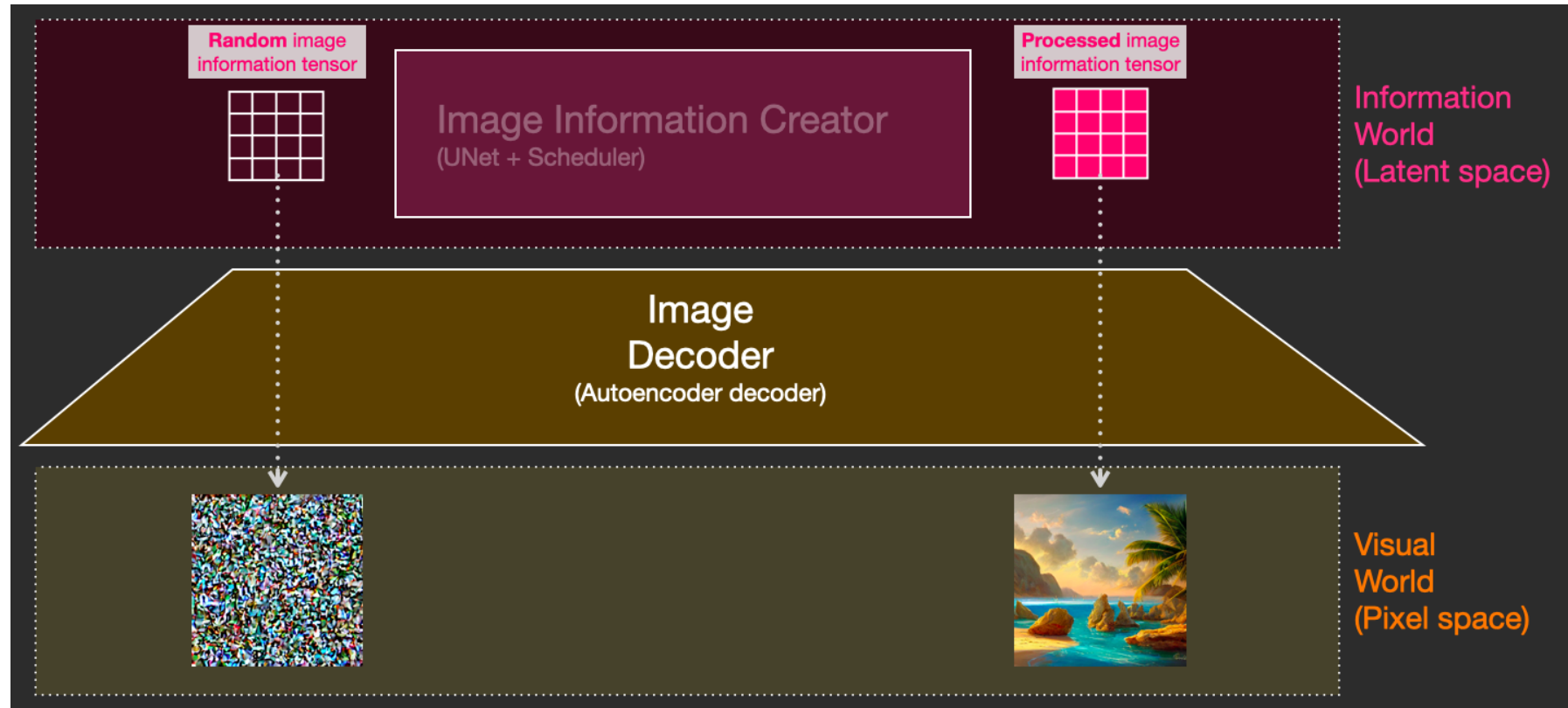


Diffusion



<https://jalammar.github.io/illustrated-stable-diffusion/>

Decoder



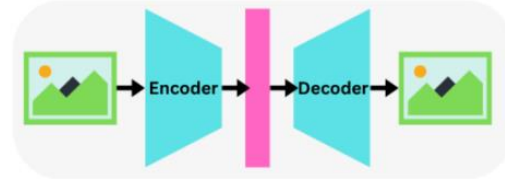
<https://jalammar.github.io/illustrated-stable-diffusion/>

Stable Diffusion

Step1

Train a Variational Autoencoder

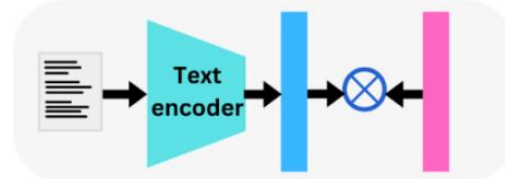
Create a latent space representation of the image



Step 2

Train a CLIP model to encode Text

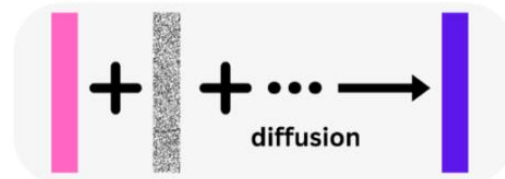
Create a text embedding semantically aligned with the image representation



Step 3

Apply forward diffusion

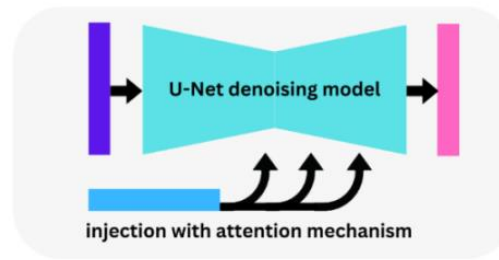
Iteratively apply random noise to the latent representation of the image



Step 4

Train a Denoising model with text prompt

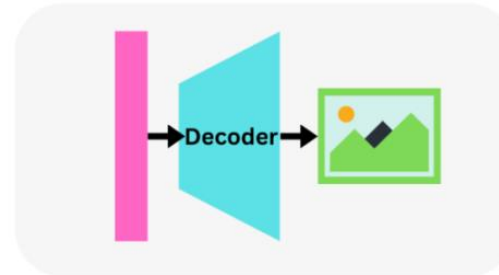
Learn to denoise the diffused image embedding and inject the text embedding with attention mechanism



Step 5

Decode the image

Use the trained autoencoder to decode the embedding into an image



Similar projects

- Stable Diffusion
- Imagen
- DELL-E2(Open AI)
- Disco Diffusion(Google)
- https://colab.research.google.com/github/alembics/disco-diffusion/blob/main/Disco_Diffusion.ipynb
- Midjourney
- NovelAI

Local Setup

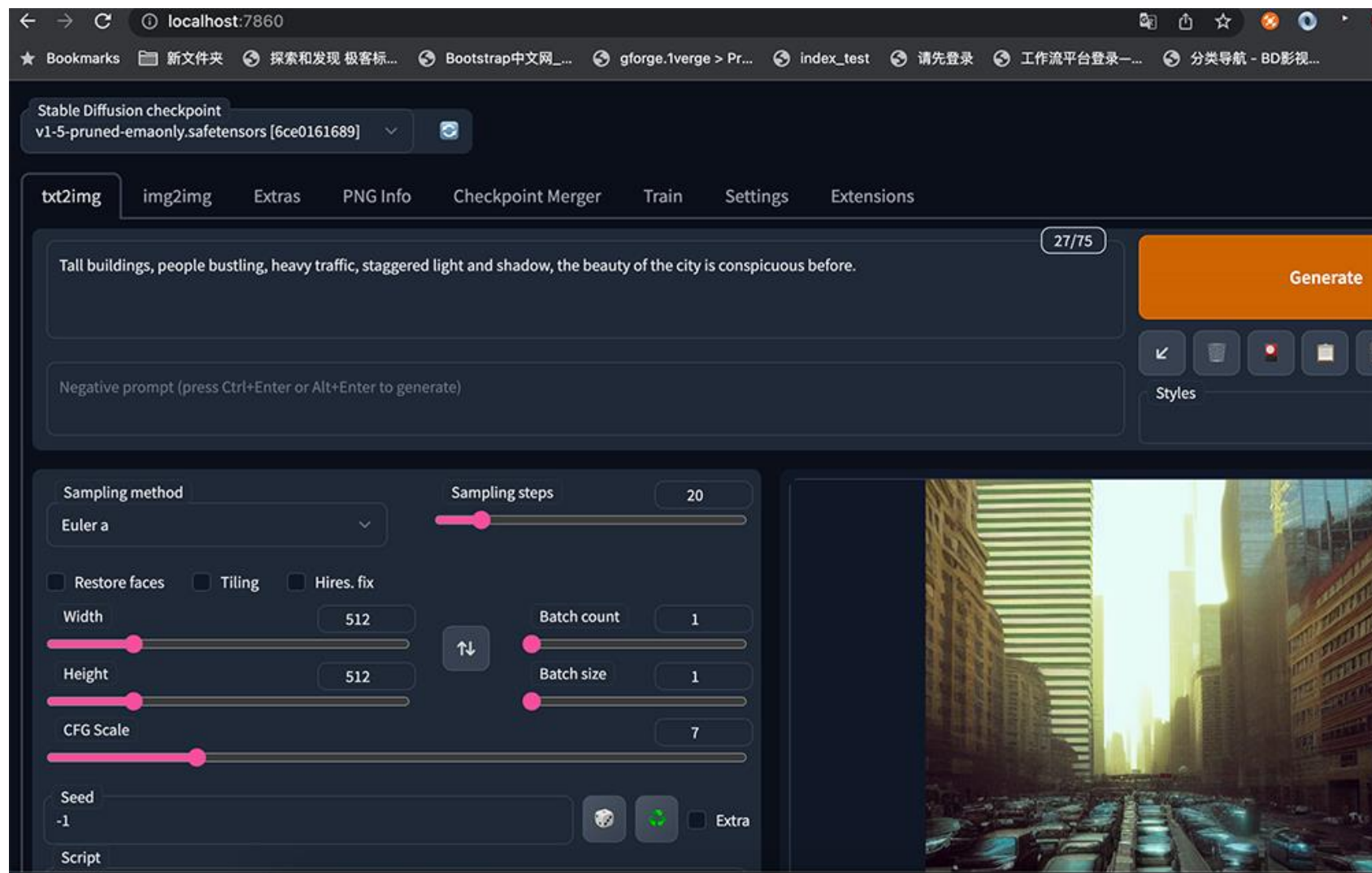
- **Stable-Diffusion-Webui**

```
git clone https://github.com/AUTOMATIC1111/stable-diffusion-webui.git  
cd stable-diffusion-webui  
./webui.sh
```

- **Stable-Diffusion-Webui-Docker**

- **DiffusionBee**

Local Setup



Online

- <https://civitai.com/>
- <https://huggingface.co/>
- <https://beta.dreamstudio.ai/dream>
- <https://stablediffusionweb.com/#demo>
- <https://huggingface.co/spaces/stabilityai/stable-diffusion>
- <https://civitai.com/models/5414/pastel-mix-stylized-anime-model-fantasyai>

Inference Demo

A dream of a beautiful girl, lovely, by Caspar David Friedrich, matte painting trending on artstation HQ

illustration of close-up street view of gothic town, night, by peter mohrbacher, by alex andreev, by jacek yerka, large depth of field, super detailed, digital art, trending on artstation, minimalism

winter winter winter winter a chinese temple building on hill and cliff with many peach trees blossoms, on himalayas, fantasy scene, fairy particle, illustration comic manga painting of road environment : 6 fantasy environment, digital painting, volumetric lighting by feng zhu, 3d alejandro alvarez alena aenami artworks in 4 k beeper, by thomas kinkade hearstone league of legends dofus overwatch

<https://hub.baai.ac.cn/view/19979>

Inference Demo

military base, cybernetic baroque centipede cyborg urban epic scene + wall is berlin iron, inside organic robotic tubes and parts + translucent surface, German Third Reich army cyberpunk 2077+ Intricate Details, Hyperrealism, Wet, unreal engine, 3d, reflections + by Arsen Asyrankulov and Liudmila Kirdiashkina

Hyper detailed painting of a cyberpunk cyberpunk cyberpunk TRAIN STATION, epic angle , ribcage skeleton symmetrical artwork, Hell of fire, 3d with depth of field, blurred background, translucent, nautilus, energy flows of water and fire. a highly detailed epic cinematic concept art CG render. made in Maya, Blender and Photoshop, octane render, excellent composition, cinematic dystopian brutalist atmosphere, dynamic dramatic cinematic lighting, aesthetic, very inspirational, arthouse. y Greg Rutkowski, Ilya Kuvshinov, by James Jean and dan mumford

close-up maximalist illustration of black panther, by makoto shinkai, akihiko yoshida, yoshitaka amano, super detailed

<https://hub.baai.ac.cn/view/19979>

Inference Demo

Prompt

solo, 1girl, portrait, looking at viewer, masterpiece, best quality, 4k, 8k, black_background, Simple background, 130mm f1.4 lens, (shiny skin:1.3), (vibrant colors), solo, ([dalcefo](#), sakimistyle:1.1), (stare, realistic, lips, [eyelashes](#):0.75), looking up, [portrait](#), solo, upper body, good hand, 4k, high-res, [masterpiece](#), best quality, head:1.3, ((Hasselblad photography)), finely detailed skin, sharp focus, (cinematic lighting), collarbone, morning, soft lighting, dynamic angle, [:(detailed face:1.2):0.2], medium breasts, cleavage, <lora:stLouisLuxuriousWheels_v1:1>, st. louis \luxurious wheels\ \([azur lane](#)\), hair ornament, earrings, [necklace](#), portrait, silver dress, revealing clothes, looking at viewer, solo

Negative prompt

(worst quality, low quality, extra digits:1.2), ([mutated hands and fingers](#), missing fingers:1.3), dark skin, monochrome, naked, (bad_prompt_version2:1), distant view, lip, Pregnant, maternity, pointy ears, realistic, tan, muscular, [greyscale](#), monochrome, lineart, 2koma, 3koma, 4koma, manga, 3D, 3Dcubism, pablo picasso, disney, marvel, [mutated breasts](#), mutated nipple, cropped, lowres, [bad anatomy](#), bad hands, text, error, [missing fingers](#), extra digit, fewer digits, cropped, worst quality, low quality, normal quality, jpeg artifacts, signature, watermark, username, blurry, artist name, lowres, trademark, watermark, title, text, deformed, bad anatomy, disfigured, mutated, extra limbs, ugly, missing limb, floating limbs, disconnected limbs, out of frame, mutated hands and fingers, poorly drawn hands, malformed hands, poorly drawn face, poorly drawn asymmetrical eyes, (blurry:1.4), duplicate (loli, petite, child, infant, [toddlers](#), chibi, sd character, teen age:1.4), tsurime, helmet hair, evil smile, smug_face, naughty smile, multiple view, Reference sheet, (worst quality, low quality:1.4),

Parameters:

Steps: 40, Sampler: DPM++ SDE Karras, CFG scale: 7, Seed: 2555331127, Size: 512x768, Model hash: b4201c7c37, Denoising strength: 0.6, ENSD: 31337, Hires upscale: 1.75, Hires steps: 10, Hires upscaler: R-ESRGAN 4x+

Inference Demo

- Prompt :
- (8k, RAW photo, best quality, masterpiece:1.2), (realistic, photo-realistic:1.37), <lora:koreanDollLikeness_v10:0.5>, omertosa, 1 lady, (aegyo sal:1), city, , professional lighting, photon mapping, [radiosity](#), physically-based rendering, blue backlight bokeh, smile, mature, [head portrait](#), sunlight>, blonde, short jacket, [school uniform](#), hoodie, short hair, floating hair
- Negative prompt:
- EasyNegative, paintings, sketches, (worst quality:2), (low quality:2), (normal quality:2), lowres, normal quality, ((monochrome)), ((grayscale)), skin spots, acnes, skin blemishes, age spot, glans, extra fingers, fewer fingers, strange fingers, bad hand, [high heel shoes](#), fat ass, hole, kid, teen, cute, naked, Ripped jeans, hole jeans, fat thigh, 6 fingers, bikini, underwear, nsfw, nude, briefs, knickers, [underpants](#), panties, swim suit, leg open
- Steps: 20, Sampler: DPM++ 2M Karras, CFG scale: 7, Seed: 3534095324, Size: 768x512, Model hash: 95afa0d9ea, Model: chilloutmix_NiPrunedFp32, Denoising strength: 0.5, Hires upscale: 2, Hires steps: 20, Hires upscaler: Latent
- [Korean Doll Likeness](#) | [Stable Diffusion LORA](#) | [Civitai](#)



Q & A

Thanks!

