Stable Diffusion ile Lokal DALL-E Arayüzünü Yap

Yağmur YILDIZ
13 Temmuz 2024

GÖRSEL ÜRETME 101

ÇIKARIM (INFERENCE)

- Eğitilmiş yapay zeka modelini canlıda kullanarak çıktı elde etmek
- Bir model 30-40GB
- Bir cycle 5-20 sn
- Aynı anda birden fazla istek

POPÜLER MODELLER

Kapalı Kaynak

ChatGPT

Midjourney

DALL·E

runway

IIElevenLabs

ANTHROP\C

Gemini

Açık Kaynak

MISTRAL AI_

BLOOM

Whisper

LLaMA by Meta

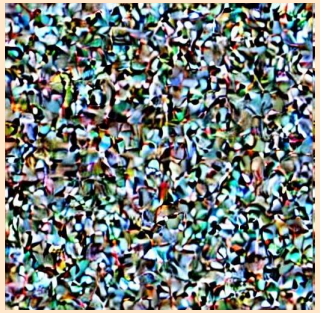
Cody

Meta AI AudioCraft

OpenAI Jukebox

GÖRSEL ÜRETME

Daha çok diffusion modelleri kullanılan bu alanda amaç inputun ögelerini anlamlandırarak bir gürültünün o ögeleri barındıran bir kompozisyona dönüşmesidir.



Latent Noise

Generated Image

- Text2Image
- Image2Image
- Audio2Image
- Video2Image

STABLE DIFFUSION MODELLERİ

Version	Release Date	Resolution	Parameters	Prompts Technology	Strengths	Weaknesses
1.4	08.2022	512x512	860M	CLIP	beginner-friendly, a little more artistic driven	long prompts, lower resolution
1.5	10. 2022	512x512	860M	CLIP	beginner-friendly, stronger portrait generation	long prompts, lower resolution
2.0	11.2022	768x768	-	OpenCLIP	shorter prompts, richer colors	aggressive NSFW filtering
2.1	12.2022	768x768	-	OpenCLIP	shorter prompts, richer colors	more “censored”, celeb filtered
XL 1.0	07.2023	1024x1024	3.5B	OpenCLIP & CLIP	shorter prompts, high resolution	requires GPU
XL Turbo	11.2023	512x512	3.5B	OpenCLIP & CLIP	Shorter prompts, less latency	cannot render text, faces and people are bad
3.0	02.2024	1024×1024	8B	Diffusion Transformer	best quality, best text render	requires GPU

SAMPLER, SCHEDULER, SEED

diversity, quality, speed, convergence

Sampler

Olasılık uzayında modelin nasıl çalışacağını belirler.
Euler, DDIM, DDPM

Scheduler

Modelin her sampleda nasıl yakınsayacağını belirler.
Linear, PNDM, Karras

Seed

Üretim sürecine başlarken kullanılan random değerdir.

SAMPLING STEP, CFG SCALE, SIZE

quality, creativity, computing power

Sampling Step

Samplerın kaç kez çalışacağını belirler. 30-75 arası tercih edilir.

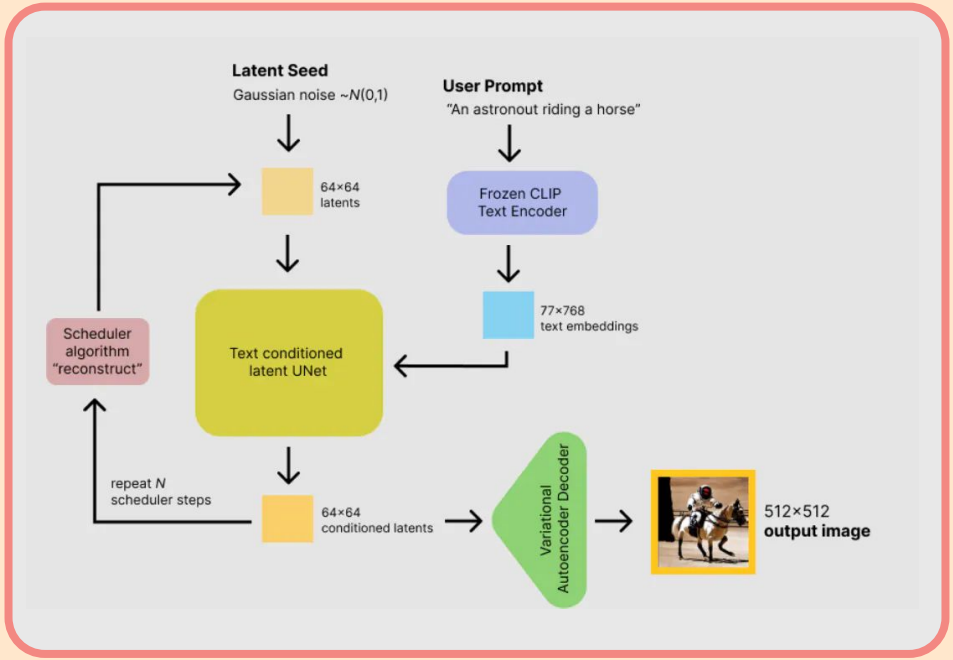
CFG Scale

Promptun generation üzerindeki etkisini belirler. 7-10 arası tercih edilir.

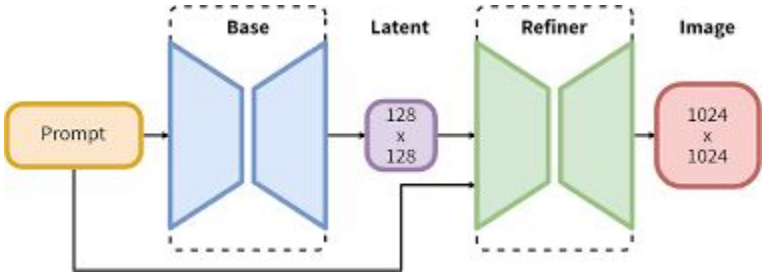
Size

Üretilecek resmin WxH boyutudur. Her modelin baz bir boyutu vardır.

STABLE DIFFUSION ALGORITHM



STABLE DIFFUSION XL DIFFERENCE

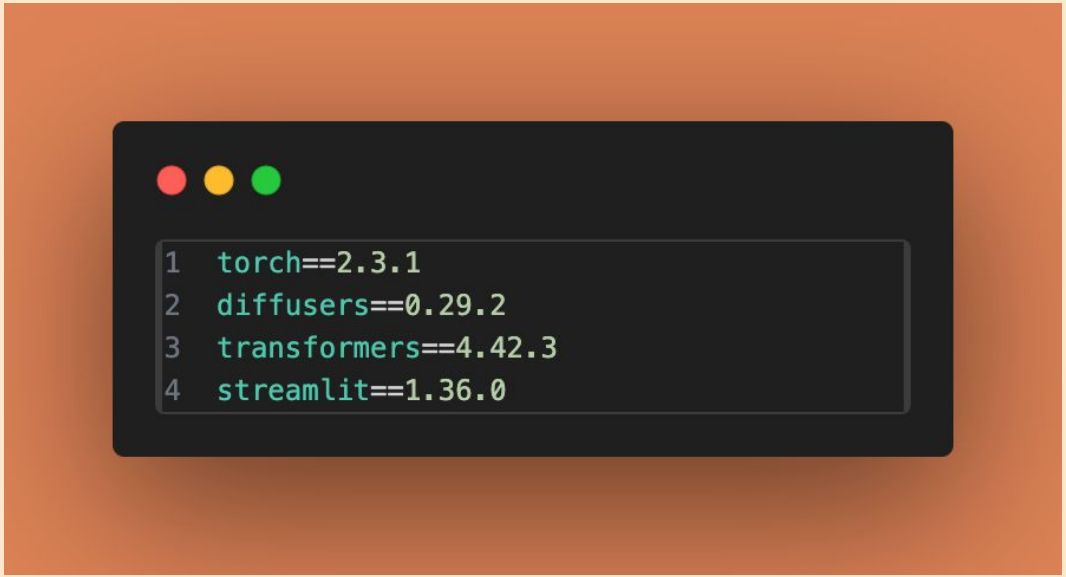


SDXL PIPELINE COMPONENTS

1. Variational Auto-Encoder (VAE) : AutoencoderKL
2. Text Encoder: CLIPTextModel (clip-vit-large-patch14)
3. Text Encoder 2: CLIPTextModelWithProjection (bigG-laion2B)
4. Tokenizer: CLIPTokenizer
5. Tokenizer 2: CLIPTokenizer
6. Unet: UNet2DConditionModel
7. Scheduler: DDIMScheduler from KarrasDiffusionSchedulers

HAMDİ

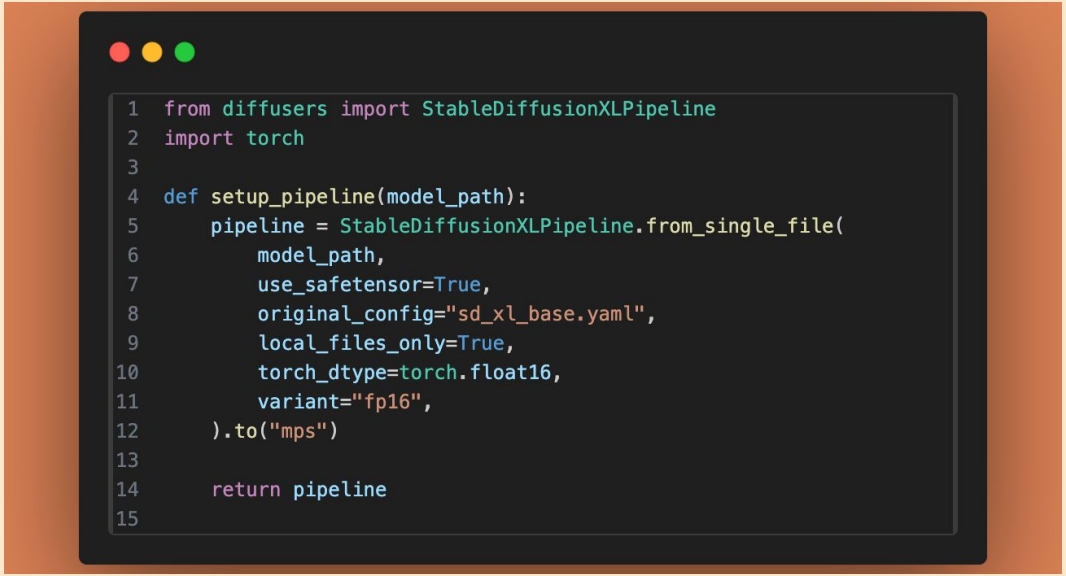
PYTHON PACKAGES



GENERATION SCHEMA

```python
class GEN_SCHEMA:
    def __init__(
        self,
        prompt: str,
        height: int,
        width: int,
        num_inference_steps: int,
        cfg_scale: float,
    ):
        self.prompt = prompt
        self.height = height
        self.width = width
        self.num_inference_steps = num_inference_steps
        self.cfg_scale = cfg_scale
```

SETUP PIPELINE



```python
from diffusers import StableDiffusionXLPipeline
import torch

def setup_pipeline(model_path):
    pipeline = StableDiffusionXLPipeline.from_single_file(
        model_path,
        use_safetensor=True,
        original_config="sd_xl_base.yaml",
        local_files_only=True,
        torch_dtype=torch.float16,
        variant="fp16",
    ).to("mps")

    return pipeline

```

GENERATE IMAGE FUNCTION

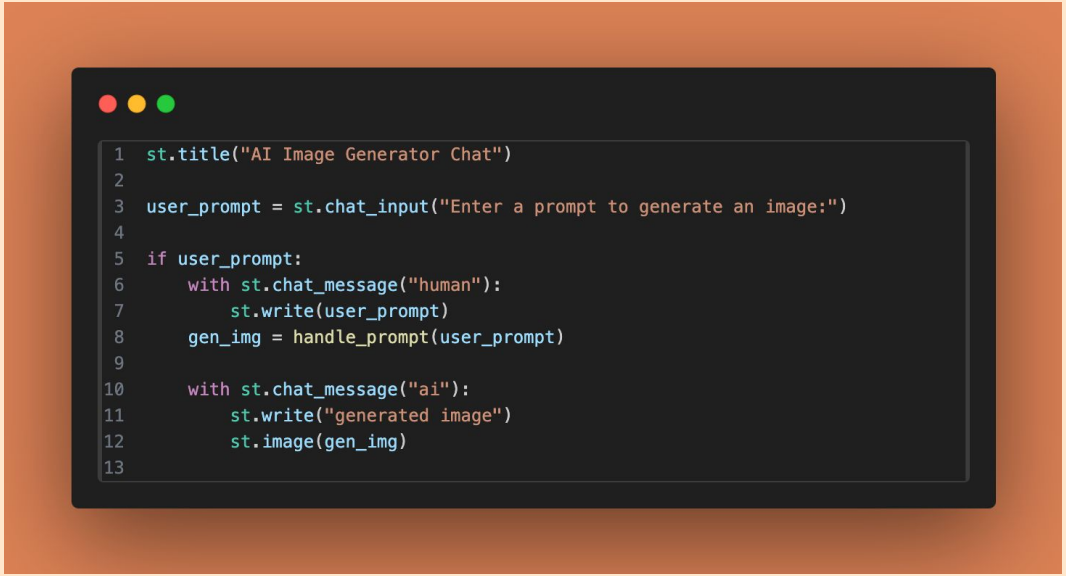
```python
from diffusers import StableDiffusionXLPipeline


def generate_image(
    components, prompt, height, width, num_inference_steps, guidance_scale
):
    pipeline = StableDiffusionXLPipeline(**components.components)

    image = pipeline(
        prompt=prompt,
        height=height,
        width=width,
        num_inference_steps=num_inference_steps,
        guidance_scale=guidance_scale,
    ).images[0]

    return image
```

CREATE PIPELINE OBJECT

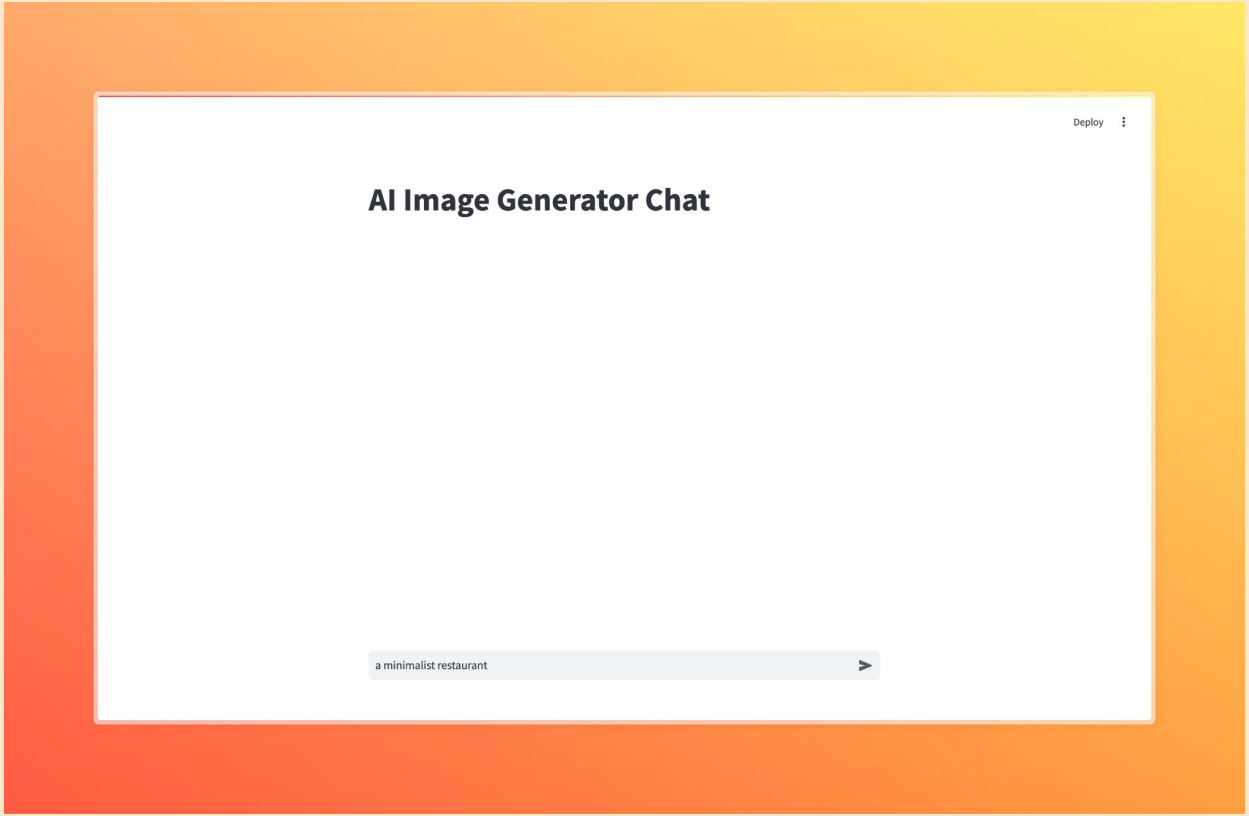
HANDLE PROMPT WITH STATIC PARAMS

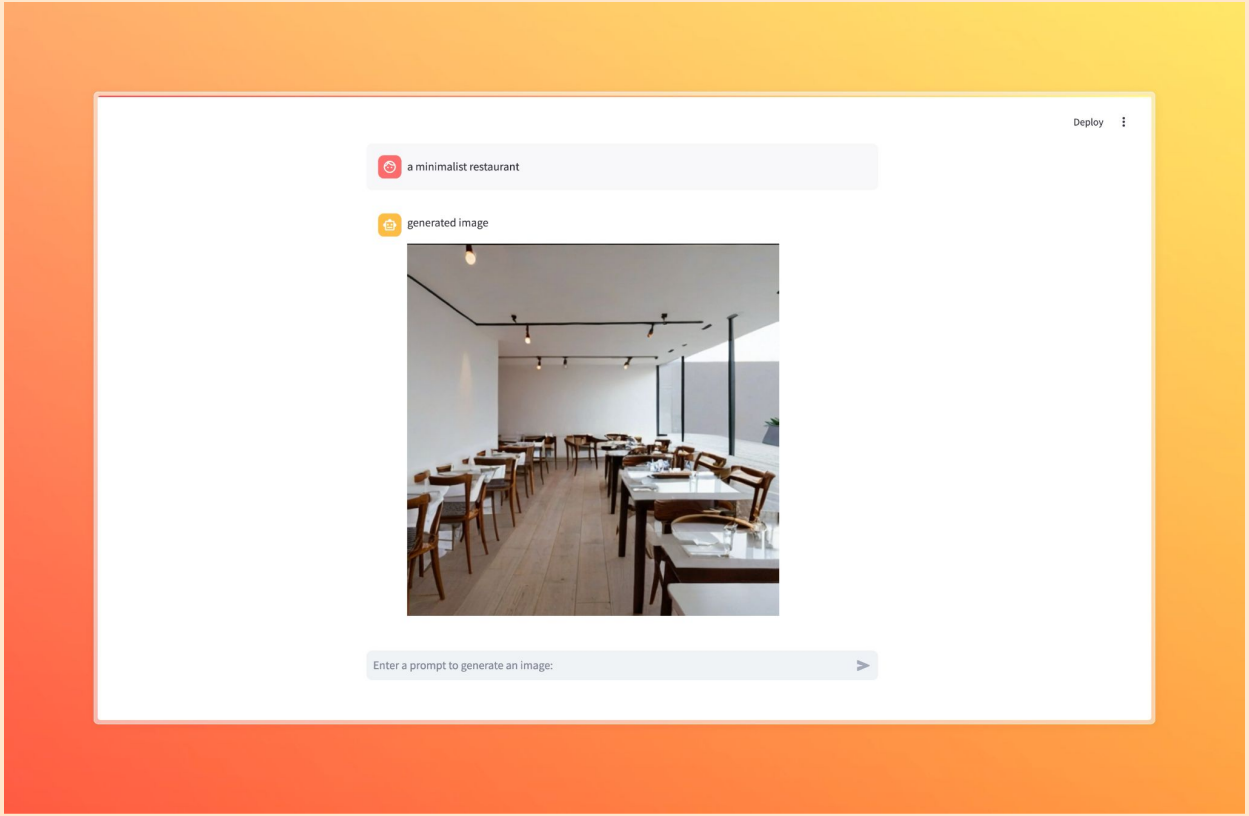
```python
def handle_prompt(prompt):
    if prompt:
        gen_params = GEN_SCHEMA(
            prompt=prompt,
            height=512,
            width=512,
            num_inference_steps=20,
            cfg_scale=7.5,
        )

        img = generate_image(
            components=pipeline,
            prompt=gen_params.prompt,
            height=gen_params.height,
            width=gen_params.width,
            num_inference_steps=gen_params.num_inference_steps,
            guidance_scale=gen_params.cfg_scale,
        )

        return img
```

UI



```python
st.title("AI Image Generator Chat")

user_prompt = st.chat_input("Enter a prompt to generate an image:")

if user_prompt:
    with st.chat_message("human"):
        st.write(user_prompt)
    gen_img = handle_prompt(user_prompt)

    with st.chat_message("ai"):
        st.write("generated image")
        st.image(gen_img)

```


Deploy
AI Image Generator Chat
a minimalist restaurant



TEŞEKKÜRLER

/in/ygryildiz

@yagmurudurdurabilirmisin

@yagmurxyildiz

@yagmurx

yagmur.cc