

Stable Diffusion の代表的なモデル

2024-11-16

Name: ADACHI Yuya

E-mail: s2120001@jaist.ac.jp

- Stability AI (<https://stability.ai/>) が作ってるモデル
- SD1.4 や SD1.5, SDXL など色々なバージョンがある
- 執筆時点で最新のモデルは SD3.5
- 個人的な肌感としては以下のような感じ
 - SD1.5 は派生モデルや開発環境が成熟している
 - SDXL は派生モデルが出だして SD1.5 から移行が進んでいる
 - SD3.5 をベースとしたモデルが日進月歩で開発中

- 一応、コマンドラインで画像生成することは可能 (開発者向け)
 - 必要なライブラリ (PyTorch や transformers など) がインストールして
 - モデルファイル (.ckptまたは.safetensors) をダウンロードする

```
$ git clone https://github.com/CompVis/stable-diffusion.git
$ cd stable-diffusion
$ pip install -r requirements.txt
$ python scripts/txt2img.py \
    --prompt "a fantasy castle in the clouds" \
    --ckpt "path_to_model/model.ckpt" \
    --config "configs/stable-diffusion/v1-inference.yaml" \
    --H 512 --W 512 --seed 42 --n_iter 1 --n_samples 1 --ddim_steps 50
```

- GUI で画像生成できるツールがサードパーティで公開されている
 - AUTOMATIC1111: <https://github.com/AUTOMATIC1111/stable-diffusion-webui>
 - Forge: <https://github.com/lllyasviel/stable-diffusion-webui-forge>
 - ComfyUI: <https://github.com/comfyanonymous/ComfyUI>

```
$ sudo apt install wget git python3 python3-venv libgl1 libglu1-mesa-dev -y
$ git clone https://github.com/AUTOMATIC1111/stable-diffusion-webui
$ cd stable-diffusion-webui
$ ./webui.sh
```

- Stability AI が 2022-10 に公開したモデル
- 512 × 512 に最適化されている
- RTX 3050 (8GB) でも普通に生成できる
- SD1.5 ベースの Counterfeit-V3.0 を使用

1girl,
Negative prompt: EasyNegativeV2,
Steps: 20, Sampler: DPM++ 2M, Schedule type: Karras, CFG scale: 7, Seed: 1548741274,
Size: 512x512, Model hash: a54c944e4c, Model: Counterfeit-V3.0_fix_fp16, Version: v1.10.1
経過時間1.0 sec. A: 2.57 GB, R: 2.72 GB, Sys: 2.8/15.6953 GB (17.6%)

Config Presets




- 最近は txt2img や img2img から派生して txt2mov が開発されている
- 5sec (8fps) の GIF だと約 1min で生成可能
- SD1.5 ベースなら VRAM 8GB でも生成可能


1girl,
Negative prompt: EasyNegativeV2,
Steps: 20, Sampler: DPM++ 2M, Schedule type: Karras, CFG scale: 7, Seed: 2995563498,
Size: 512x512, Model hash: a54c944e4c, Model: Counterfeit-V3.0_fix_fp16, AnimateDiff:
"model: mm_sd15_v3.safetensors, video_length: 40, fps: 8, loop_number: 0, closed_loop:
A, batch_size: 16, stride: 1, overlap: 4, interp: Off, interp_x: 10, freeinit_enable: False,
mm_hash: a8ca0ead, version: v2.0.1-a-1-ga88e889", Version: v1.10.1


経過時間 1 min. 17.1 sec.

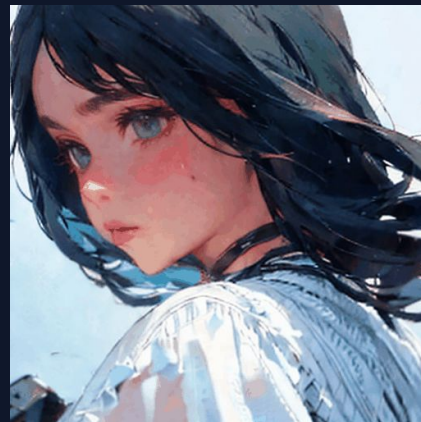
A: 5.00 GB, R: 6.36 GB, Sys: 6.7/15.6953 GB (42.5%)

Config Presets









- Stability AI が 2023-07 に公開したモデル
- 1024 × 1024 に最適化されている
- VRAM 8GB だとチューニングしないと厳しい

1girl,blue archive,general,masterpiece,best quality,good quality,very aesthetic,absurdres,
Negative prompt: normal quality,worst quality,lowres,(bad),text,error,fewer,extra,missing,worst
quality,jpeg artifacts,low quality,unfinished,displeasing,oldest,early,chromatic
aberration,signature,extra digits,artistic error,username,scan,abstract,monochrome,speech
bubble,comic,
Steps: 25, Sampler: Euler a, Schedule type: Automatic, CFG scale: 7, Seed: 1479553355, Size:
1024x1024, Model hash: ff827fc345, Model: NoobAI-XL, VAE hash: 235745af8d, VAE:
sdxl_vae.safetensors, Version: v1.10.1

経過時間24.4 sec.

A: 9.50 GB, R: 12.11 GB, Sys: 12.4/15.6953 GB (79.1%)

Config Presets

None



RTX 4090 をくれ！！

- Stable Diffusion とは Stability AI が作ってるモデル
- SD1.4 や SD1.5, SDXL など色々なバージョンがある
- 世間では SD1.5 から SDXL に移行中
- 開発者は SD3.5 をこねくり回してる