Stable Diffusion

Stable Diffusion is an AI model for text to image generation. There are various GUIs available to tweak settings, add specific styles like Disney or photorealistic (Lora models) and help adjusting existing images via picture to picture or selection infill.

The guides below focus on the Automatic1111 Stable Diffusion WebUI.

Linux Docker install

This docker requires docker as well as the GPU drivers installed on the host system.

https://github.com/AbdBarho/stable-diffusion-webui-docker

```
git clone https://github.com/AbdBarho/stable-diffusion-webui-docker
cd stable-diffusion-webui-docker
vi docker-compose.yaml
#adjust name to stable-diffusion if desired
#adjust data and output dir location to /opt/stable-diffusion/data and
output respectively if desired
mkdir -p /opt/stable-diffusion/{data,output}

#Build the ai and download the models.
#Note, first start will take 15-60 minutes to download models to data folder
as cache depending on internet connection
#size approximately 10GB

docker compose --profile download up --build

# if download errors occur, just repeat the command.

# build the desired interface:
# docker compose --profile [ui] up --build
# where [ui] is one of: invoke | auto | auto-cpu | comfy | comfy-cpu

# use auto-cpu or comfy-cpu for cpu only interface

#docker compose --profile auto-cpu up --build
docker compose --profile auto up --build

# later:
docker compose --profile auto up -d
```

access the app on:

http://localhost:7860/

Windows install for AMD GPU

Windows AMD AUTOMATIC1111 stable diffusion webui using Microsoft DirectML for GPU acceleration:

1. Download and run installers for Python 3.10.6 (ticking Add to PATH) and git
2. Clone the fork as AMD is not officially supported
   Paste into command prompt in a directory where it should live

   ```
   git clone
   https://github.com/lshqqytiger/stable-diffusion-webui-directml
   cd stable-diffusion-webui-directml
   git submodule init
   git submodule update

   venv\Scripts\python.exe -m pip install --upgrade pip
   venv\Scripts\pip.exe install torch-directml
   ```

3. Right-click webui-user.bat and adjust the command line args to

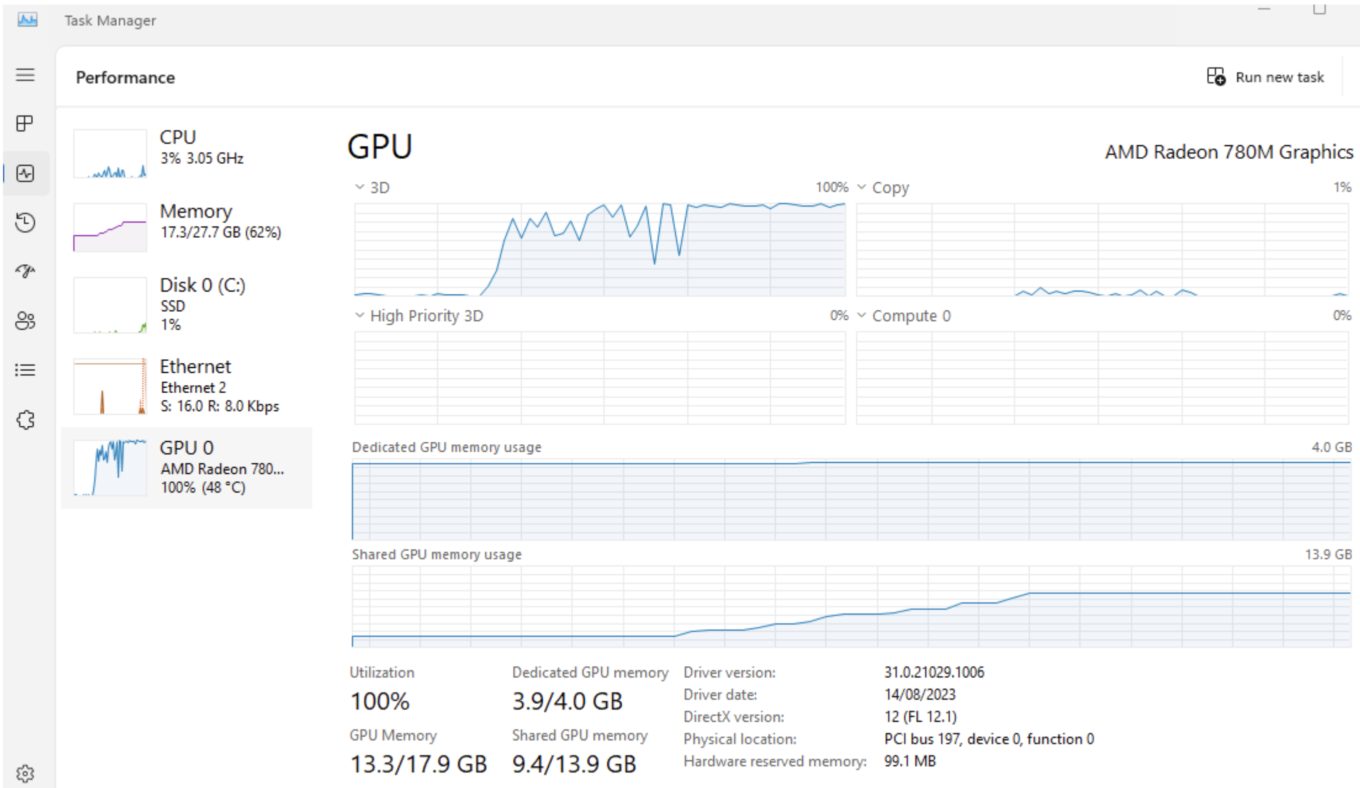
   ```
   COMMANDLINE_ARGS=--use-directml --skip-torch-cuda-test
   ```

4. If you have only 4-6gb vram, try adding these additional flags to webui-user.bat's command line argslike so:

   ```
   --opt-sub-quad-attention --lowvram --disable-nan-check
   ```

5. Double-click webui-user.bat in that directory, this will download ~2.7GB torch and other python modules, followed by the 4GB Stable-Diffusion 1.5 pruned model.
6. If it looks like it is stuck when installing or running, press enter in the terminal and it should continue.
7. once done, a browser opens to http://localhost:7860/

Check/Test if GPU is used in Task Manager's GPU tab. Dedicated GPU Memory should show large usage for the loaded Model and when generating a test image in the web interface, the GPU usage should ramp up like here:



Gallery Addon

As an extension for SD-webui: Open the Extensions tab in SD-webui.

1. Select the Install from URL option.
2. Enter https://github.com/zanllp/sd-webui-infinite-image-browsing
3. Click on the Install button.
4. Wait for the installation to complete and click on Apply and restart UI.

general advice

Some general Stable Diffusion prompt advice

Models / Settings

LoRA models/styles/tools as well as base models can be searched and downloaded from CivitAI

Realistic Vision + Lora

Description for photorealistic images:

1. Download model Realistic Vision V6.0 B1 Base Model from

https://civitai.com/models/4201/realistic-vision-v11 and save it into the `models/stable-diffusion` directory
2. create `models/Lora` directory
3. download LORA Style hyperrealism art and save to `models/Lora` directory
4. download LORA epi_noiseoffset and save to `models/Lora` directory
5. in web interface go to extensions → install from url and paste to install:

```
https://github.com/mcmonkeyprojects/sd-dynamic-thresholding
https://github.com/opparco/stable-diffusion-webui-composable-lora
```

6. then go to installed tab and click apply and quit
7. restart stable diffusion web ui
8. go to Lora tab, configure entries (which will create json files, then the entries disappear)

Base settings:

```
set sampling steps to 20
set sampling method to DPM++ SDE Karras
set width to 768 and height to 512
set CFG scale to 6
set seed to -1
enable Dynamic Thresholding (CFG Scale Fix)
enable all composable Lora
```

Stable Diffusion SDXL Model

```
cd /opt/stable-diffusion/data/models/Stable-diffusion
wget
https://huggingface.co/stabilityai/stable-diffusion-xl-base-1.0/raw/main/sd_xl_base_1.0_0.9vae.safetensors
```

Stable Diffusion XL Turbo model

This is for general objects/scenes/animals/etc, faster than SDXL

download SDXL Turbo from https://huggingface.co/stabilityai/sdxl-turbo and copy into `Models/Stable-diffusion` directory.

Image Style/Prompt libraries

Lexica PromptHero

Some test image prompts

```
3d, 8k, masterpiece, (best quality), Japanese garden, waterfall, cherry
blossoms in full bloom, conifer, koi pond, (pagoda:1.1), stone lantern,
water basin, water reflections, winding path, (night sky:1.:1), hokusai
inspiration, ultra-realistic, pastel color scheme, soft lighting, golden
hour, tranquil atmosphere, landscape orientation

EasyNegative, (worst quality:1.2), (low quality:1.2), (lowres:1.1),
(monochrome:1.1), (greyscape), multiple views, comic, sketch, watermark

25 steps, 2m karras, 1024x768 seed 1192013237
```

```
greg rutkowski, highly detailed, dark, surreal scary swamp, terrifying,
horror, poorly lit, trending on artstation, incredible composition,
masterpiece

with settings:
 LMS Karras
 50 steps
 CFG scale 8
 random seed
```

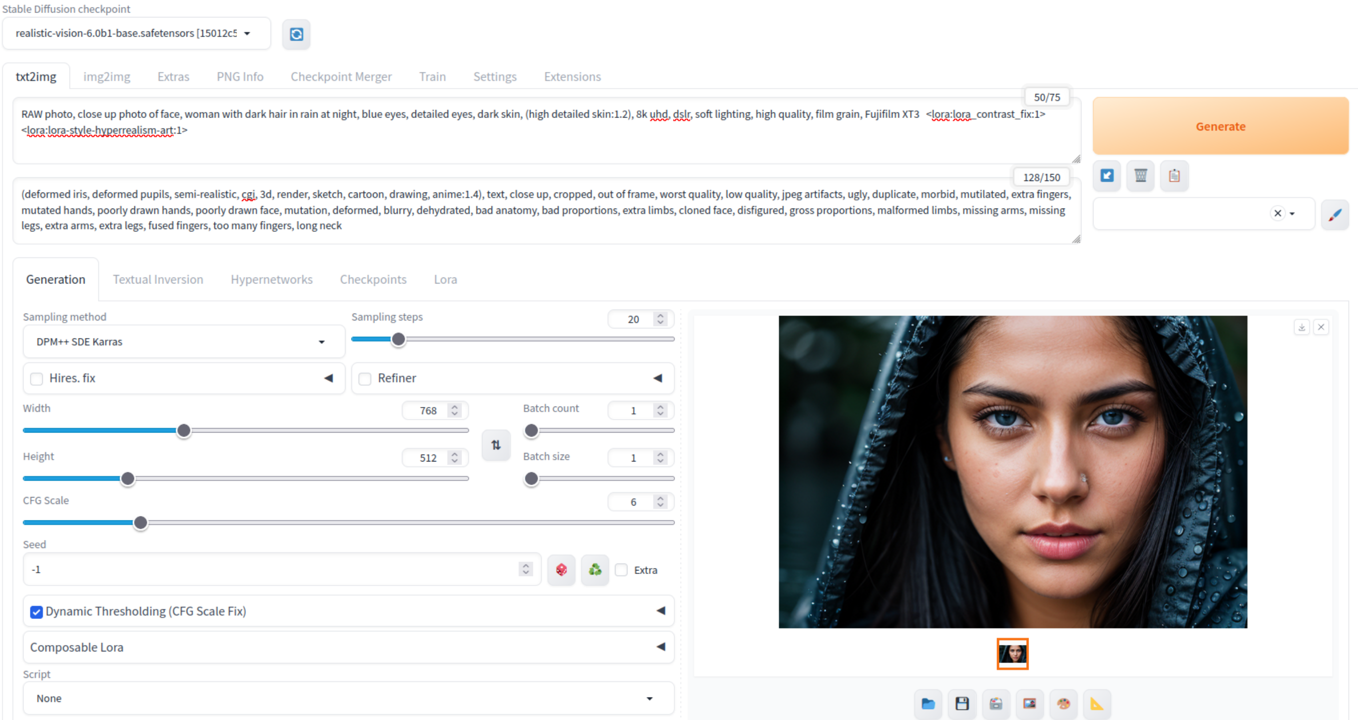
Test. use in txt2img the following prompt:

```
RAW photo, close up photo of face, woman with dark hair in rain at night,
blue eyes, detailed eyes, dark skin, (high detailed skin:1.2), 8k uhd, dslr,
soft lighting, high quality, film grain, Fujifilm XT3
<lora:lora_contrast_fix:1>  <lora:lora-style-hyperrealism-art:1>
```

and as negative prompt below:

```
(deformed iris, deformed pupils, semi-realistic, cgi, 3d, render, sketch,
cartoon, drawing, anime:1.4), text, close up, cropped, out of frame, worst
quality, low quality, jpeg artifacts, ugly, duplicate, morbid, mutilated,
extra fingers, mutated hands, poorly drawn hands, poorly drawn face,
mutation, deformed, blurry, dehydrated, bad anatomy, bad proportions, extra
limbs, cloned face, disfigured, gross proportions, malformed limbs, missing
arms, missing legs, extra arms, extra legs, fused fingers, too many fingers,
long neck
```

then click generate as often as you like


Stable Diffusion checkpoint
realistic-vision-6.0b1-base.safetensors [15012c5
txt2img
img2img
Extras
PNG Info
Checkpoint Merger
Train
Settings
Extensions
50/75
RAW photo, close up photo of face, woman with dark hair in rain at night, blue eyes, detailed eyes, dark skin, (high detailed skin:1.2), 8k uhd, dslr, soft lighting, high quality, film grain, Fujifilm XT3 <lora:lora_contrast_fix:1> <lora:lora-style-hyperrealism-art:1>
Generate
128/150
(deformed iris, deformed pupils, semi-realistic, cgi, 3d, render, sketch, cartoon, drawing, anime:1.4), text, close up, cropped, out of frame, worst quality, low quality, jpeg artifacts, ugly, duplicate, morbid, mutilated, extra fingers, mutated hands, poorly drawn hands, poorly drawn face, mutation, deformed, blurry, dehydrated, bad anatomy, bad proportions, extra limbs, cloned face, disfigured, gross proportions, malformed limbs, missing arms, missing legs, extra arms, extra legs, fused fingers, too many fingers, long neck
Generation
Textual Inversion
Hypernetworks
Checkpoints
Lora
Sampling method
DPM++ SDE Karras
Sampling steps
20
Hires. fix
Refiner
Width
768
Batch count
1
Height
512
Batch size
1
CFG Scale
6
Seed
-1
Extra
Dynamic Thresholding (CFG Scale Fix)
Composable Lora
Script
None

00002-859497391.png 551.0 KB

RAW photo, close up photo of face, woman with dark hair in rain at night, blue eyes, detailed eyes, dark skin, (high detailed skin:1.2), 8k uhd, dslr, soft lighting, high quality, film grain, Fujifilm XT3 <lora:lora_contrast_fix:1> <lora:lora-style-hyperrealism-art:1>
Negative prompt: (deformed iris, deformed pupils, semi-realistic, cgi, 3d, render, sketch, cartoon, drawing, anime:1.4), text, close up, cropped, out of frame, worst quality, low quality, jpeg artifacts, ugly, duplicate, morbid, mutilated, extra fingers, mutated hands, poorly drawn hands, poorly drawn face, mutation, deformed, blurry, dehydrated, bad anatomy, bad proportions, extra limbs, cloned face, disfigured, gross proportions, malformed limbs, missing arms, missing legs, extra arms, extra legs, fused fingers, too many fingers, long neck
Steps: 20, Sampler: DPM++ SDE Karras, CFG scale: 6, Seed: 859497391, Size: 768x512, Model hash: cc6cb27103, Model: v1-5-pruned-emaonly, Lora hashes: "lora_contrast_fix: d1131f7207d6, lora-style-hyperrealism-art: 6d4ea812a2d2", Dynamic thresholding enabled: True, Mimic scale: 7, Separate Feature Channels: True, Scaling Startpoint: MEAN, Variability Measure: AD, Interpolate Phi: 1, Threshold percentile: 100, Version: v1.7.0