Voldy Retard Guide

(a.k.a Automatic1111)

1,000,000 VIEWS!

The definitive Stable Diffusion experience ™
(Special thanks to all anons who contributed)
---FEATURE SHOWCASE & HOWTO---

SD News
Japanese guide here 日本語ガイド (JP Resources)
NovelAI FAQ

Minimum requirements:
-16gb ram
-Nvidia Maxwell (GTX 7xx) or newer GPU with **at least** 2gb vram
-Linux or Windows 7/8/10+
-20gb disk space

--GUIDE--

Step 1: Install Git (page)
-When installing, make sure to check the box for 'Windows Explorer integration -> Git Bash'

Step 2: Clone the WebUI repo to your desired location:
-Right-click anywhere and select 'Git Bash here'
-Enter `git clone https://github.com/AUTOMATIC1111/stable-diffusion-webui`
(Note: to **update**, all you need to do is is type `git pull` within the newly made webui folder)

Step 3: Download your preferred model(s):

- Official Stable Diffusion 1.4: Huggingface (requires signup) or **HERE** | (magnet)
- Waifu Diffusion Final: (SD 1.4 + **680k** Danbooru images, heavy anime bias) **HERE** | (Float32 Ver.)
- Waifu Diffusion Alpha: (SD 1.4 + **56k** Danbooru images, slight anime bias) **HERE** | **(mirror)** | (magnet)
- Trinart2: (SD 1.4 trained for Anime concept art/'Pixiv style') **HERE**
- See **This page** for additional models (Most in beta)

Step 4: Rename your .ckpt file to "model.ckpt", and place it in the `/models/Stable-diffusion` folder
-You can have as many models as you want in the folder, "model.ckpt" is just the one it will load by default

Step 5: Install Python 3.10 (Windows 7 ver) (page)

-Make sure to choose "**add to PATH**" when installing

Step 6 (Optional):
This reduces VRAM, and allows you to generate at larger resolutions or batch sizes for a <10% loss in raw generation speed
(For me, singular results were slightly slower, but generating with a batch size of 4 made each result **25% faster** on average)
This is recommended for *most* users
-Edit `webui-user.bat`
-Change `COMMANDLINE_ARGS=` to `COMMANDLINE_ARGS=--medvram`

Step 7: Launch `webui-user.bat`, Open it as normal user, **not** as administrator.

- Wait patiently while it installs dependencies and does a first time run.
  It may *seem* "stuck" but it isn't. It may take up to 10-15 minutes.
  And you're done!

> ℹ️ **Usage**

- Launch webui-user.bat
- After loading the model, it should give you a LAN address such as **'127.0.0.1:7860'**
- Enter the address into your browser to enter the GUI environment
  Tip: Hover your mouse over UI elements for tooltips about what they do
- To exit, close the CMD window

--NOVELAI SETUP--

Provided in service to freedom of information, testing and research
If you enjoy your results, consider **subscribing**

Step 1: Follow the main guide above
-(Ignore steps **3** and **4** if you *only* plan on using the NovelAI model)
Open a git bash by right-clicking inside your main stable diffusion webui folder and type `git pull` to make sure you're updated

Step 2: Download a Torrent Client if you don't have one already
-Add the following Magnet link:
magnet:?xt=urn:btih:5bde442da86265b670a3e5ea3163afad2c6f8ecc
-**Deselect** everything except for the the `/animefull-final-pruned` subfolder (under /stableckpt), **and** `animevae.pt`
-(Optional: select desired **.pt** hypernetworks in the `/modules` subfolder (*These are **not** used by default in NAI but can provide unique results*)

Step 3: Once finished, **rename** the following files you torrented like so:
`animevae.pt` **>>** `nai.vae.pt`
`config.yaml` **>>** `nai.yaml`
`model.ckpt` **>>** `nai.ckpt`

Step 4: Place **all 3** files into `/stable-diffusion-webui/models/Stable-diffusion`

Step 5 (Optional): If you want to launch NovelAI by default, do the following
-Edit webui-user.bat in notepad and add the NAI .ckpt:
`COMMANDLINE_ARGS=--ckpt nai.ckpt`
Otherwise, you can just select it within the WebUI dropdown

Hypernetworks (Optional):
Using Hypernetwork .pt files can provide unique changes to your NAI output depending on which one you use
To enable them:
-Create `/stable-diffusion-webui/models/hypernetworks` folder if it doesn't exist already
-Paste your .pt modules (anime, aini, furry etc) into the `hypernetworks` folder
-Reload the WebUI
-In the 'Settings' tab, choose your .pt under **Stable Diffusion finetune hypernetwork** and hit apply

> 💡 **Replicating NovelAI defaults**

It is possible to create outputs **identical** to NovelAI's current subscription service by doing the following:

Standard (Euler)
-Set the sampler to **Euler** (Not Euler A)

-Use **28** Steps

-Set CFG Scale to **11**

-Use **masterpiece, best quality** at the beginning of all positive prompts

-Use **nsfw, 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** as the negative prompt
(nsfw is optional, and toggled on the site)

-In the Settings tab, change **Ignore last layers of CLIP model** to **2** and apply

Euler-A Defaults
If you want to replicate NovelAI Euler_a results, you must do the following alongside the defaults above:
-Under **Sampler Parameters** in settings, set 'ETA Noise Seed Delta' to **31337**

And you should be ready to prompt!

💡 **Testing accuracy**

The easiest way to test if your NovelAI setup is functioning properly is to attempt replicating a result from NovelAI itself:
For this example we'll be using best girl:

Sampler: `Euler`
Seed: `2870305590`
CFG: `12`
Resolution: `512x512`
Prompt: `masterpiece, best quality, masterpiece, asuka langley sitting cross legged on a chair`
Negative Prompt: `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`
(Double check you copied the prompts exactly!)
Note: You may see very slight variation if you are using an optimization such as lowvram or xformers, but it should still be **95%** similar
Troubleshooting incorrect Asukas

💡 **NovelAI resources**

- Artstyle tags
- NAI Expressions/Hair/Clothing (some nsfw)
- Prompt converter (Convert NovelAI prompt Syntax to WebUI)
- Emoji Prompts

--RUNNING ON 4GB (And under!)--

These parameters are also useful for regular users who want to make larger images or batch sizes!
It is possible to drastically reduce VRAM usage with **some modifications:**

- **Step 1:** Edit `webui-user.bat`
- **Step 2:** After `COMMANDLINE_ARGS=` , enter your desired parameters:
  Example: `COMMANDLINE_ARGS=--medvram`

- If you have 4GB VRAM and want to make 512x512 (or maybe larger) images,
  use `--medvram`.
- If you have **2GB** VRAM,
  use `--lowvram`

If you are getting 'Out of memory' errors on either of these, add `--always-batch-cond-uncond` to the other arguments

NOTES:

- If you get a **green/black screen** instead of generated pictures, you have a card that doesn't support half precision floating point numbers (known problem on 16xx cards):
  You must use `--precision full --no-half` in addition to other flags, and the model will take much more space in VRAM
  If you are using a **.vae** file, you must also add `--no-half-vae`
- Make sure to **disable** hardware acceleration in your browser and close anything which might be occupying VRAM if you are getting out-of-memory errors, and possibly remove GFPGAN (if you have it)

--ALTERNATE GUIDE (Conda)--

(You can try this method if the traditional install isn't working)

- **Follow Steps 1-4 on the main guide above**
- Download Miniconda HERE. Download Miniconda 3
- Install Miniconda in the default location. Install for **all users**.
  Uncheck "Register Miniconda as the system Python 3.9" unless you want to
- Open Anaconda Prompt (miniconda3)
- In Miniconda, navigate to the `/stable-diffusion-webui` folder wherever you downloaded using "cd" to jump folders.
  (Or just type "cd" followed by a space, and then *drag the folder into* the Anaconda prompt.)
- Type the following commands to make an environment and install the necessary dependencies:
- `conda create --name qwe`
  (You can name it whatever you want instead of qwe)
- `conda activate qwe`
- `conda install python`
- `conda install git`
- `webui-user.bat`
  (Note: it may seem like it's stuck on "Installing torch" in the beginning. This is normal and should take 10-15 minutes)
  It should now be ready to use

> ℹ️ **Usage**

- Navigate to `/stable-diffusion-webui` in Miniconda
- Type `conda activate qwe`
  (You will need to type 'conda activate qwe' every time you wish to run webui)
- Type `webui-user.bat`
- After loading the model it should give you a LAN address such as '**127.0.0.1:7860**'
  Enter the address into your browser to enter the GUI environment

--LINUX INSTALLATION--

Step 1: Install dependencies

- Debian-based:

```
sudo apt install wget git python3 python3-venv
```
- Red Hat-based:
```
sudo dnf install wget git python3
```
- Arch-based:
```
sudo pacman -S wget git python3
```

Step 2: To install in `/home/$(whoami)/stable-diffusion-webui/`, run:
```
bash <(wget -qO- https://raw.githubusercontent.com/AUTOMATIC1111/stable-diffusion-webui/master/webui.sh)
```

--LINKS--

- Danbooru Tag Groups (NAI/WD)
- Embedding list
- mbedding Collection (MEGA)
- Hypernetwork Training for dummies
- Textual Inversion guide (colab)
- SD 1.4 Artist list with pictures
- Top 500 artists comparison
- Search the training database
- Inpainting guide(nsfw)
- SD wiki (on hiatus)
- Voldy's mini-wiki
- CPU Guide (Openvino)
- Inpainting Tips
- Anon's guide for anime vectors (Waifu Diffusion)
- Prebuilt Xformers
- Remacri Upscaler(Landscapes) Lollypop Upscaler(Anthropomorphic Figures)
- Other Upscaler Models
  (Place upscaler models in ESRGAN folder)
- Prompts from China
- Build great aesthetic prompts using the prompt builder
- Japanese Keywords
- Use Darkreader to change your Gradio theme to dark mode
- Informal Training guide (30gb vram+)
- Debugging guide
- Python for Windows 7
- Stable diffusion WebUI repo
- Waifu Diffusion huggingface page

--TROUBLESHOOTING--

- Make sure your folder paths do **not have spaces**
- If you get **Git is not recognized as an internal or external command** when launching, your Git is not set to PATH (Common issue on W7)
  -**1:** Open the start menu and search for "Edit system environment variables", open it
  -**2:** Click "Environment Variables"
  -**3:** Find `Path` under **System Variables** and double click it
  -**4:** Add your Git path:
  -**W10**: With the "New" button in the PATH editor, add `C:\Program Files\Git\bin\` and `C:\Program Files\Git\cmd\` to the end of the list.
  -**W7**: At the end of "Variable value", insert a `;` followed by your Git path:
  eg. `;C:\Program Files\Git\bin\;C:\Program Files\Git\cmd\`. Do not put a space between `;` and the entry.
  -Git should now be recognized inside CMD!
- If you get **launch.py error, unrecognized arguments** when trying to use the `--ckpt` flag, try using the full relative path for it in webui-user.bat
  eg. `COMMANDLINE_ARGS= --ckpt ./models/Stable-diffusion/nai.ckpt`
- "I git pulled to update and something broke!"
  Naturally since many new features are introduced, there may be some instability. Follow **These instructions** to revert to a previous build

- If you are getting a **winerror** when installing, or you feel you broke something and want to reinstall from scratch, delete these directories: `venv`, `repositories` and try again
- If you are getting **Python not found** as an error, you may need to manually set your PATH in webui-user.bat: see "Setting Python PATH" under 'Howto Extras' below for instructions
- (img2img) if you get **RuntimeError: Sizes of tensors must match**, you need to change the resolution of your input image
- Make sure you have the latest CUDA toolkit and GPU drivers you can run
- If you get **Torch is not able to use GPU**, you may have to download Python 3.7 instead
- if your version of Python is not in PATH (or if another version is) create or modify webui.settings.bat in `/stable-diffusion-webui` folder add the line set PYTHON=python to say the full path to your python executable: `set PYTHON=B:\soft\Python310\python.exe` You can do this for python, but not for git.
- The installer creates a python virtual environment, so none of installed modules will affect your system installation of python if you had one prior to installing this.
- To prevent the creation of virtual environment and use your system python, edit webui.bat replacing set `VENV_DIR=venv` with `set VENV_DIR=`
- webui.bat installs requirements from files `requirements_versions.txt`, which lists versions for modules specifically compatible with Python 3.10.6. If you choose to install for a different version of python, editing webui.bat to have set REQS_FILE=requirements.txt instead of set REQS_FILE=requirements_versions.txt may help (but I still reccomend you to just use the recommended version of python).
- If you get a **green/black** output instead of generated pictures, you have a card that doesn't support half precision floating point numbers (known problem on 16xx cards): -edit webui-user.bat -Modify line 6 to `COMMANDLINE_ARGS=--precision full --no-half` along with any other flags you wish to add If you are using a **.vae** file, you must also add `--no-half-vae` Unfortunately, the model will take much more space in VRAM- So it is recommended to use flags like `--medvram` in combination with it
- If your output is a jumbled rainbow mess your image resolution is set TOO LOW
- Having too high of a CFG level will also introduce color distortion, your CFG should be between 5-15
- On older systems, you may have to change `cudatoolkit=11.3` to `cudatoolkit=9.0`
- Make sure your installation is on the C: drive
- This guide is designed for NVIDIA GPUs *only*, as stable diffusion requires cuda cores. AMD users should try THIS GUIDE

--TIPS--

- You can quickly switch between your downloaded .ckpt files within the WebUI
- Unlike base SD, NovelAI can generate images natively up to **768x768** without any distortion
- Otherwise, if you are generating images significantly **larger** than 512x512 in SD, make sure to check **Highres, fix** for the best results. if you don't, 'cloning' distortion may begin appearing (multiple faces, arms, etc) +Lower denoising strength seems to work best (I used 0.5)
- Even with the available fix, it is still recommended to generate in 512x512 for the most accurate results, as the non-NovelAI models were trained on 512x images
- The Waifu model and normal .ckpt have their own pros and cons; Non-anime promps done with the waifu .ckpt will be biased toward anime stylization, making realistic faces and people more difficult unless you merge it with another model
- Do **not** use the `--medvram` or `--lowvram` argument when training! Results will be highly inferior
- nai.yaml may **not** be necessary to achieve 1:1 with NovelAI results
- Use ((( ))) around keywords to increase their strength and [[[ ]]] to decrease their strength
- Unlike other samplers, **k_euler_a** can generate high quality results from low steps.
- **Save prompt as style** allows you to save a prompt as an easily selectable output. A box to select will appear to the left of Roll after you save your first style, allowing you to make a selection. Prompts can be deleted by accessing `styles.csv` (This can be helpful if you find a combination that generations really good images and want to repeatedly use it with varied subjects.)
- You can drag your favorite result from the output tab on the right **back into** img2img for further iteration
- The **k_euler_a** and **k_dpm_2_a** samplers give vastly different, more intricate results from the same seed & prompt However their results are not consistent throughout steps. Other samplers provide more predictable, linear refinement with more steps
- The seed for each generated result is in the output filename if you want to revisit it
- Using the same keywords as a generated image in img2img produces interesting variants
- It's recommended to have your prompts be at least 512 pixels in *one* dimension, or a 384x384 square at the absolute smallest Anything smaller will have heavy distortion

- CLIP interrogator takes up a lot of space (8gb), you might not want to select it if you don't plan on using it frequently
- Try Low strength (0.3-0.4) + High CFG in img2img for interesting outputs
- You can use Japanese Unicode characters in prompts

--HOWTO EXTRAS--

It is possible to run the full WebUI on a CPU instead of a GPU, but it is very sluggish
If you want a faster but very basic CPU experience, try the Openvino *version*
Otherwise:

- Follow the main Guide
- Modify these 2 lines under **def prepare_environment** like so:
- ```
  torch_command = os.environ.get('TORCH_COMMAND', "pip install torch torchvision torchaudio --extra-index-url https://download.pytorch.org/whl/cpu")
  ```
- ```
  commandline_args = os.environ.get('COMMANDLINE_ARGS', "--skip-torch-cuda-test --precision full --no-half")
  ```
- Launch webui-user.bat like normal

If you have multiple models installed and you want to launch from another conveniently, you can make another .bat

- Make a copy of webui-user.bat and name it whatever you want
- After `COMMANDLINE_ARGS=`, add `--ckpt` followed by the desired model to your launch parameters:
  eg: `COMMANDLINE_ARGS=--ckpt wd-v1-2-full-emma.ckpt`
  You can also select a different model while in the webUI, under the **settings** tab

After running once, a `ui-config.json` file appears in the webui directory:
Edit values to your liking and the next time you launch the program they will be applied.

Edit webui-user.bat and add the necessary parameters after `COMMANDLINE_ARGS=`, along with any existing parameters you have

- Use `--listen` to make the server listen to network connections. This will allow devices on your **local network** to access the UI, and if you configure port forwarding, also computers on the internet.
- Use `--share` option to run online. You will get a xxx.app.gradio link. This is the intended way to use the program in collabs.
- Use `--port xxxx` to make the server listen on a specific port, `xxxx` being the wanted port. Remember that all ports below 1024 needs root/admin rights, for this reason it is advised to use a port above 1024. Defaults to port 7860 if available.
- Use `--share --gradio-auth username:password` to add shared authentication
  Optionally, you can provide multiple sets of usernames and passwords separated by commas (user1:pw1, user2:pw2).

Sometimes, you may need to manually add your Python directory to PATH:

- Find the location of your Python install
  it should be something along the lines of `C:\Users\you\AppData\Local\Programs\Python`
- Open the folder and **shift+right click** `python.exe`
  then click 'copy as path'
- Edit `webui-user.bat`
  paste the path you copied after **PYTHON=** (keeping quotes), and save

Note: This only applies to those who used `git clone` to install the repo
You can set your script to automatically update by editing `webui-user.bat`
Add `git pull` one line above `call webui.bat`
Save

Although most features are relatively self explanatory to use, the X/Y plot script can be particularly obtuse to understand, notably the "S/R" option

- The S/R prompt Searches the whole prompt for the first entry in the the comma separated values field and Replaces all occurrences of the first word with with one entry from the values Prompt S/R field on every iteration.
- The iterations of course also happen for every value of the other type field.
- The keyword will *also* be iterated, so using something like "red, white, blue" will result in issues when your prompt also features "reddit

gayfurs".

-Copy the text Here and save it as `output.bat`, Move it to wherever you want your images to output to.
Run it, and it will create the appropriate sub-folders. You can delete the .bat after this is complete.
-Go to the Settings tab of the UI and assign your new file locations accordingly. Once you've assigned the locations, make sure to hit **Apply Settings**
It is also recommended to enable the following setting if you want your outputs to be organized
-[x] When writing images/grids, create a directory with name derived from the prompt

-----**GFPGAN**-----
GFPGAN is used for correcting realistic faces, it was replaced with CodeFormer face correction which comes with the main install and is generally better.
To install GFPGAN, download and place **GFPGANv1.3.pth** into the main webUI directory

--PRUNING A .CKPT--

'Unpruned' models can be up to 7gb due to redundant training data,
but it can be reduced to 3.6gb without any loss of quality, reducing ram usage and loading time
(The original model is not lost, a new pruned copy is made)
NOTE: You should only do this after running webui-user.bat at least once

- Place the .ckpt file you wish to prune in your main `/stable-diffusion-webui` folder
- Copy https://raw.githubusercontent.com/harubaru/waifu-diffusion/main/scripts/prune.py
  Delete line **6** and **8**
  Save as `prune.py`
  Save as 'all files' in your main `/stable-diffusion-webui` folder
- Edit the last line in prune.py to the name of your ckpt:
  eg. `prune_it('wd-v1-2-full-emma.ckpt')` and save
- Copy and save the script for launching prune.py **HERE**
  Save it as `prune.bat` in your main `/stable-diffusion-webui` folder
  Save as 'all files'
  (This loads venv and torch dependencies into memory before running prune.py)
- Run prune.bat
  The pruning process may take a few minutes
  Afterward, you should now have a pruned .ckpt alongside your normal one

--SAFE UPDATING--

The AUTOMATIC1111 repo is what we call *bleeding edge*, which means that although new features are added rapidly, it's not uncommon for new bugs to come along with them.
The good thing is, git pull updates are **reversible** if something ever breaks
Here's how you can update without taking a risk:

- Open a git bash in your main Stable Diffusion folder
- Enter `git rev-parse --short HEAD` to get the hash of your current version. Write it down
  It should be a 7 character string like 'cc58036'
- Update normally by typing `git pull`

If the newest version works fine, great!
If you need to revert, all you have to do is type `git checkout` followed by your previous version hash
eg. `git checkout 7edd58d`

--OUTPAINTING--

(9/17) A new and improved outpainting script has been added to the webUI!
Make sure to use `git pull` so you can update to the latest version
To use, go to img2img and select "Outpainting mk2" from the Script dropdown menu
~~**Recommended parameters**~~ (Further testing is needed)
Steps: 85-100
Sampler: Euler a
CFG Scale: 7.5

Denoising Strength: 0.8
Width: Same as source image
Height: Same as source image
Pixels to expand: 128-256
Mask Blur: 10-25
Fall-off exponent: 1-2
Color variation: 0.05
Tips

- Make sure your width and height are the same or close to the source image resolution,
  otherwise your outpainting results will be incoherent
- Don't feel locked into these parameters, tweaking is highly encouraged
  they are just a rough approximation of what seemed to work best for me through a few minutes of testing
- The higher the mask blur, the more 'seamless' results tend to be (to an extent)
  but if it's too high, deformed distortions occur

--W7 HELP--

On Windows 7, you **may** get "**api-ms-win-core-path-l1-1-0.dll is missing**" as an error when trying to follow this guide.
This is because many modern programs and frameworks require a system file only present in newer versions of Windows.
Luckily, it has been backported to be compatible with W7, and can be downloaded **HERE** (Github page)
Upzip and copy the **x86** .dll into `C:\Windows\SysWOW64` and the **x64** .dll into `C:\Windows\System32` and reboot, then you should be good to go
If you do not get that error, then you do not need to do this.

--MISC--

--OLD MODEL--
The original v1.3 leaked model from June can be downloaded here:
https://drinkordiecdn.lol/sd-v1-3-full-ema.ckpt
Backup Download: https://download1980.mediafire.com/3nu6nlhy92ag/wnlyj8vikn2kpzn/sd-v1-3-full-ema.ckpt
Torrent Magnet: https://rentry.co/6gocs

--OLD GUIDE--
Voldy guide pre-Table of Contents (9/15) https://rentry.org/voldyold
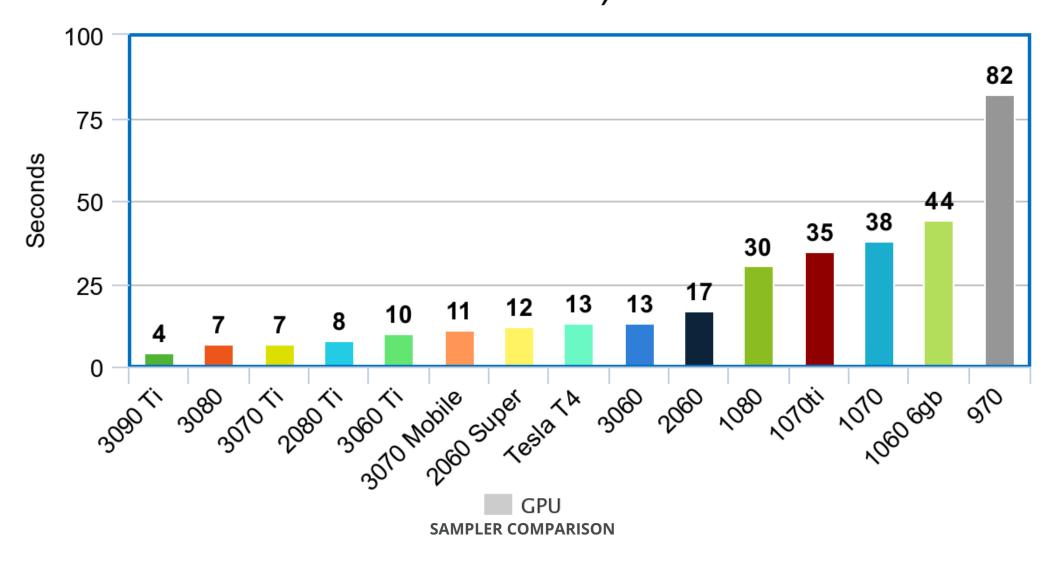The original hlky guide (replaced as of 9/8/22) is here: https://rentry.org/GUItard
Japanese hlky guide https://news.livedoor.com/article/detail/22794512/
The original guide (replaced as of 8/25/22) is here: https://rentry.org/kretard

APPROXIMATE RENDER TIME BY GPU (50 steps)

Time spent generating 512x512 sample (Stable Diffusion)

GPU	Seconds
3090 Ti	4
3080	7
3070 Ti	7
2080 Ti	8
3060 Ti	10
3070 Mobile	11
2060 Super	12
Tesla T4	13
3060	13
2060	17
1080	30
1070ti	35
1070	38
1060 6gb	44
970	82

GPU

SAMPLER COMPARISON

Sampler vs. Steps

(Scale = 15.0, W = 512, H = 512)