Textual Inversion Training using AUTOMATIC1111 Stable Diffusion Web UI

1. Requirements

- AUTOMATIC1111's Stable Diffusion Web UI installed (locally or in Colab).
- At least 6–8 GB VRAM (higher is better).
- A folder with 5–30 training images (JPG/PNG).
- Python & PyTorch environment ready.

2. Prepare Dataset

- Create folder: textual_inversion/mytoken/
- Place your training images inside it (512x512 recommended).
- (Optional) Add captions:

  001.jpg
  001.txt → "a photo of a vrtkls person"

3. Create Embedding

1. Open the Web UI in browser.
2. Go to: Train > Create Embedding
3. Fill in:
   - Name: vrtkls
   - Initialization text: person (or similar concept)
   - Vectors per token: 1
4. Click Create embedding

This creates vrtkls.pt inside the /embeddings/ folder.

4. Train the Embedding

Go to: Train > Train Embedding
Fill out the form:

Field	Value
Embedding	vrtkls
Dataset directory	textual_inversion/vrtkls/
Learning rate	0.005
Max steps	3000 (or 5000)
Vectors per token	1
Prompt Template	prompt.txt (optional)

Example of prompt.txt:
a photo of *person*
Then click Train Embedding

5. Monitor Progress

- Loss will appear on the UI.
- Save interval can be adjusted.
- After training, the final .pt file is saved in /embeddings/.

6. Test the Trained Token

Go to txt2img tab and enter a prompt:

a portrait of vrtkls person wearing sunglasses

Use various modifiers to test quality.

7. Tips

- Use clean, consistent images.
- 10–20 images usually give good results.
- Try different prompt styles.
- For faces, DreamBooth may be better.

vrtkls close-up headshot, natural light, confident expression, DSLR photo, sharp focus, by vrtkls

4/75

blurry, double body

Generation | Textual Inversion | Hypernetworks | Checkpoints | Lora

Generate

Sampling method
DPM++ 2M Karras

Sampling steps 20

Hires. fix ◀

Refiner ◀

Width 512

Batch count 1

Height 512

Batch size 1

CFG Scale 7

Seed
-1

Extra

ControlNet v1.1.455 ◀

36°C
Smoke

Search

ENG
IN

16:35
12-05-2025

A man and a <drashti> woman gazing into each other's eyes, in a romantic setting with sunset in the background, soft lighting, warm tones

31/75

blurry

1/75

Generate

| Generation | Textual Inversion | Hypernetworks | Checkpoints | Lora |

Sampling method

DPM++ 2M Karras

Sampling steps 20

Hires. fix ◄

Refiner ◄

Width 512

Batch count 1

Height 512

Batch size 1

CFG Scale 7

Seed

-1

☐ Extra

ControlNet v1.1.455 ◄