# DreamBooth Training Guide: Cinematic two Person Conversion

- **Goal**

Train a Stable Diffusion DreamBooth model using the notebook to generate cinematic-style images of a specific person, suitable for character conversations or movie-like storytelling.

- **Prerequisites**

- Google Account
- Basic knowledge of Colab & Hugging Face
- A few high-quality images (3–10) of the target person
- Hugging Face token (for downloading base models)

- **Steps to Train:**

1. Open the Notebook:
   Use the provided link to open in Google Colab.:
   https://colab.research.google.com/drive/1GW-j1_5nSM7YjByBgOdXtJMk9b_Yt2vv?authuser=1
2. Setup Environment:
   Run the initial cells to install dependencies (diffusers, accelerate, etc.).
3. Install Requirements:

```
!wget -q
https://gist githubusercontent com/FurkanGozukara/be7be5f9f7820d0bb85a3052
874f184e/raw/d8d179da6cab0735bd5832029c2dec5163db87b4/train_dreamboo
th py
!wget -q
https://github com/ShivamShrirao/diffusers/raw/main/scripts/convert_diffusers_
to_original_stable_diffusion py
%pip uninstall torchtext --yes
%pip install -qq git+https //github.com/ShivamShrirao/diffusers
%pip install torch==2.2.0 torchvision torchaudio --index-url
https //download.pytorch.org/whl/cu121 --upgrade
%pip install -q -U --pre triton --upgrade
%pip install -q accelerate transformers ftfy gradio natsort safetensors
%pip install bitsandbytes==0.41.3  --upgrade
%pip install xformers==0.0.24  --upgrade
```

```
%pip install triton==2.2.0 --upgrade
%pip install diffusers==0.27.0 --upgrade
%pip install huggingface_hub==0.25.2
%pip install numpy==1.26.4
%pip install transformers==4.43
```

4. Connect drive, download stable diffusion model
5. Login to Hugging Face:
   Enter your Hugging Face token when prompted (sign up at https://huggingface.co).
6. Upload Your Images:
   Prepare 3–10 photos of the person in cinematic lighting (portrait, different angles).
   Upload to a folder named something like 'person_cinematic/'.
7. Define Your Prompt Template:
   Use a custom token like 'cinematic [name]' or 'photo of cinematic man' in your training prompt. Example: 'a cinematic portrait of sks person'.
8. Set Training Parameters:
   - Pretrained model: runwayml/stable-diffusion-v1-5
   - Resolution: 512x512
   - Steps: ~800–2000 (depending on overfitting)
   - Learning Rate: ~5e-6
   - Save checkpoints every 500 steps
9. Run Training Cell:
   It will train the model using the uploaded images and prompt.
10. Run Gradio UI for generating images.
11. Test Generation:
    Use the inference cell to try prompts like: 'a cinematic close-up of sks person, dramatic lighting, 35mm film'
12. Optional) Delete diffuser and old weights and only keep the ckpt to free up drive space.

- **Tips**

- Use cinematic tags: 'cinematic lighting', 'moody', 'film grain', 'anamorphic', 'bokeh'.
- Maintain 1:1 face-to-frame size for input images.
- Use prompt editing later for expressions or different camera angles.

- **Links**

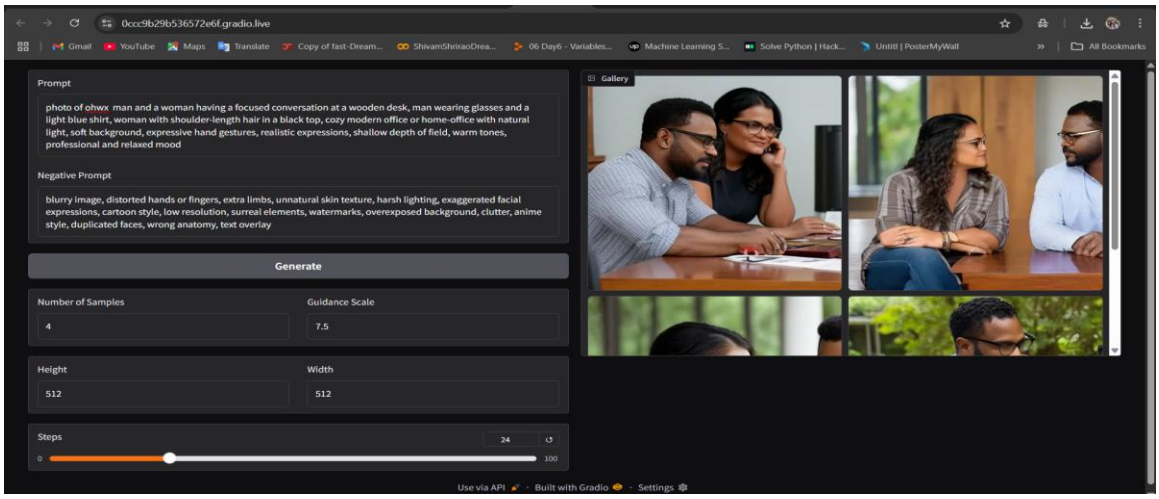
Hugging Face: https://huggingface.co

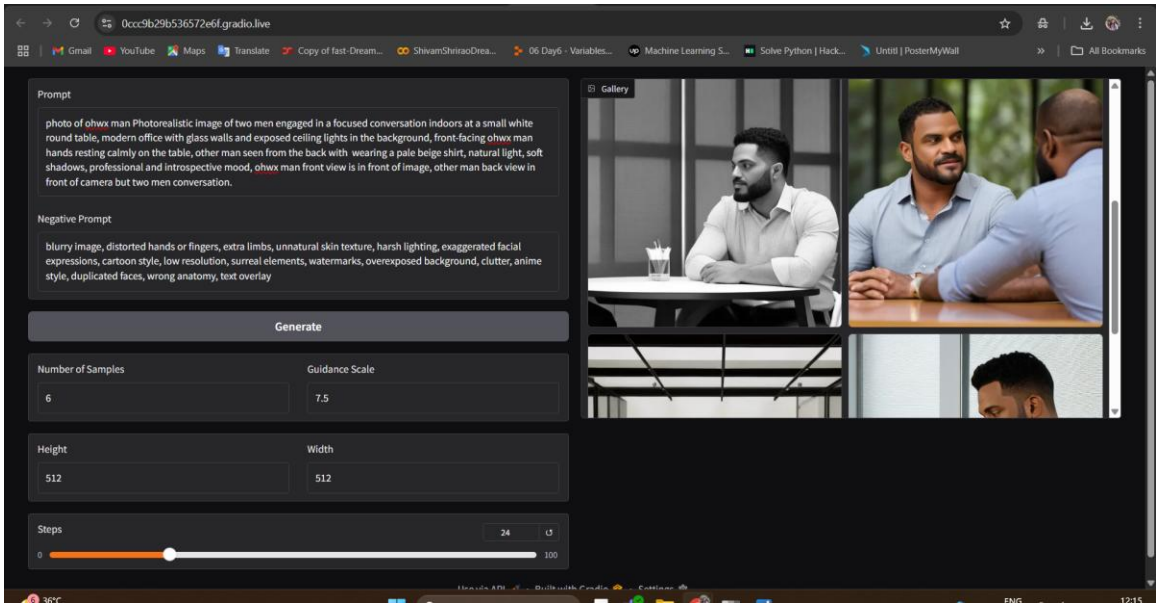
Stable Diffusion Model: stable-diffusion-v1-5/stable-diffusion-v1-5:
https://huggingface.co/stable-diffusion-v1-5/stable-diffusion-v1-5

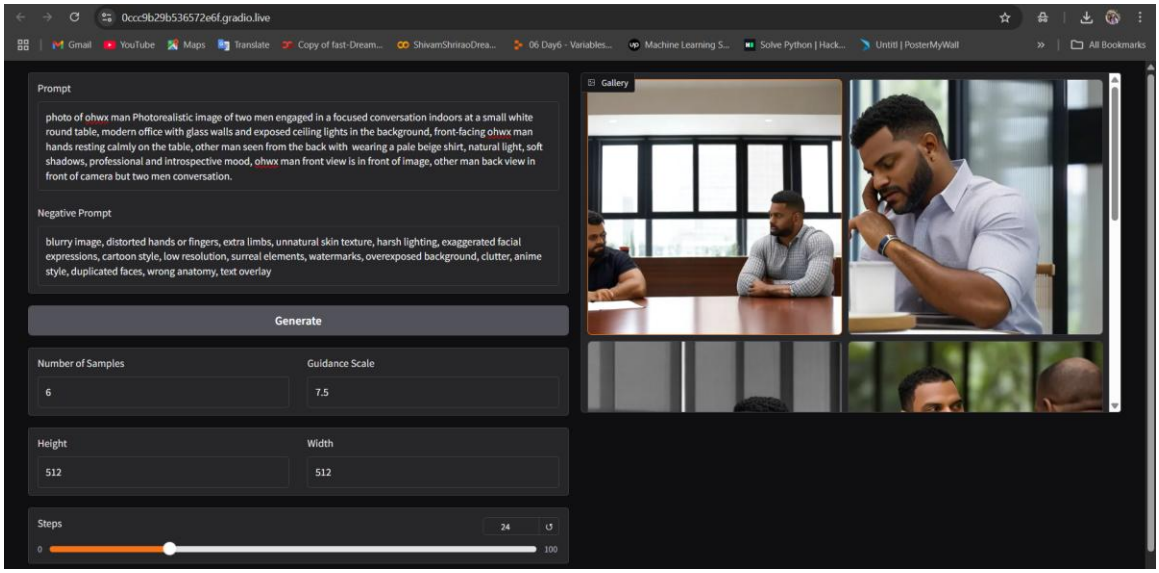
DreamBooth Colab Notebook:
https://colab.research.google.com/github/FurkanGozukara/Stable-Diffusion/blob/main/DreamBooth/ShivamShriraoDreamBooth.ipynb?authuser=1

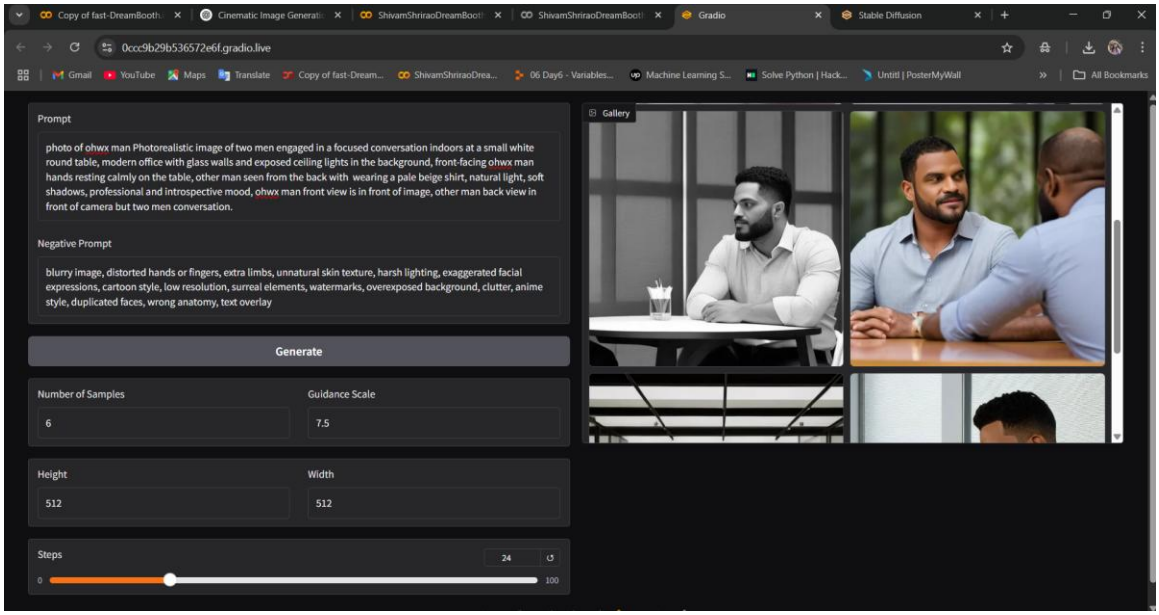
- **Outputs of model**

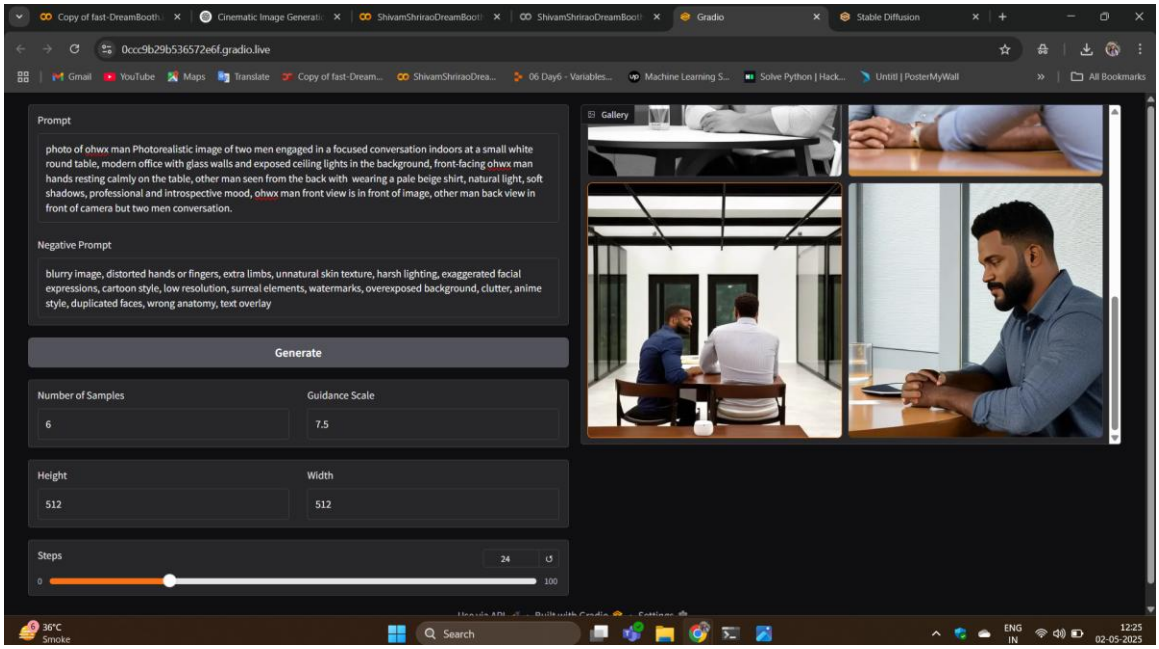
**Ohwx man:**



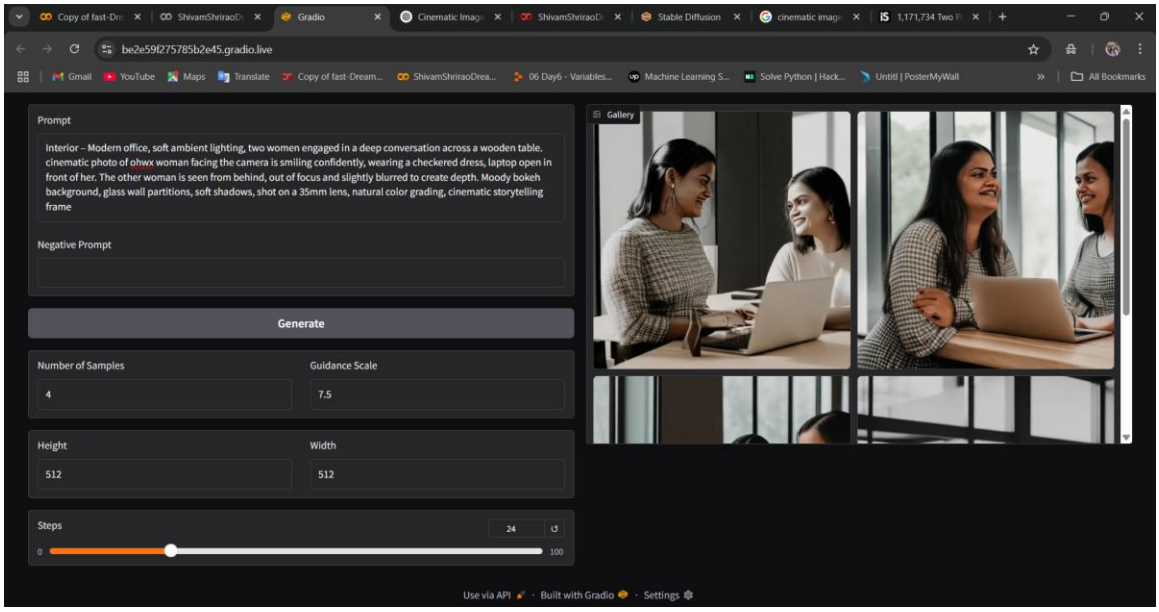




0ccc9b29b536572e6f.gradio.live
Prompt
photo of ohwx man Photorealistic image of two men engaged in a focused conversation indoors at a small white round table, modern office with glass walls and exposed ceiling lights in the background, front-facing ohwx man hands resting calmly on the table, other man seen from the back with wearing a pale beige shirt, natural light, soft shadows, professional and introspective mood, ohwx man front view is in front of image, other man back view in front of camera but two men conversation.
Negative Prompt
blurry image, distorted hands or fingers, extra limbs, unnatural skin texture, harsh lighting, exaggerated facial expressions, cartoon style, low resolution, surreal elements, watermarks, overexposed background, clutter, anime style, duplicated faces, wrong anatomy, text overlay
Generate
Number of Samples
6
Guidance Scale
7.5
Height
512
Width
512
Steps
24
0
100
Gallery


0ccc9b29b536572e6f.gradio.live
Prompt
photo of ohwx man Photorealistic image of two men engaged in a focused conversation indoors at a small white round table, modern office with glass walls and exposed ceiling lights in the background, front-facing ohwx man hands resting calmly on the table, other man seen from the back with wearing a pale beige shirt, natural light, soft shadows, professional and introspective mood, ohwx man front view is in front of image, other man back view in front of camera but two men conversation.
Negative Prompt
blurry image, distorted hands or fingers, extra limbs, unnatural skin texture, harsh lighting, exaggerated facial expressions, cartoon style, low resolution, surreal elements, watermarks, overexposed background, clutter, anime style, duplicated faces, wrong anatomy, text overlay
Generate
Number of Samples
6
Guidance Scale
7.5
Height
512
Width
512
Steps
24
0
100
Gallery


Gradio
Stable Diffusion
0ccc9b29b536572e6f.gradio.live
Prompt
photo of ohwx man Photorealistic image of two men engaged in a focused conversation indoors at a small white round table, modern office with glass walls and exposed ceiling lights in the background, front-facing ohwx man hands resting calmly on the table, other man seen from the back with wearing a pale beige shirt, natural light, soft shadows, professional and introspective mood, ohwx man front view is in front of image, other man back view in front of camera but two men conversation.
Negative Prompt
blurry image, distorted hands or fingers, extra limbs, unnatural skin texture, harsh lighting, exaggerated facial expressions, cartoon style, low resolution, surreal elements, watermarks, overexposed background, clutter, anime style, duplicated faces, wrong anatomy, text overlay
Generate
Number of Samples
6
Guidance Scale
7.5
Height
512
Width
512
Steps
24
Gallery

## Ohwx woman:


be2e59f275785b2e45.gradio.live
Prompt
Interior – Modern office, soft ambient lighting, two women engaged in a deep conversation across a wooden table. cinematic photo of ohwx woman facing the camera is smiling confidently, wearing a checkered dress, laptop open in front of her. The other woman is seen from behind, out of focus and slightly blurred to create depth. Moody bokeh background, glass wall partitions, soft shadows, shot on a 35mm lens, natural color grading, cinematic storytelling frame
Negative Prompt
Generate
Number of Samples
4
Guidance Scale
7.5
Height
512
Width
512
Steps
24
Gallery
Use via API · Built with Gradio · Settings

**vrtk man:**