

Stable diffusion inpainting task clothes and background

Implementation and Experimental Approaches

Two main approaches were tested to achieve background alteration and clothes inpainting.

1. First Approach(failed significantly)

In this approach, I directly provided the input image and attempted to alter the background using two inpainting models:

- [**diffusers/stable-diffusion-xl-1.0-inpainting-0.1**](#)
- [**stabilityai/stable-diffusion-2-inpainting**](#)

For masking, I experimented with both **DeepLab** and **Segment Anything (SAM)**.

However, this approach **failed significantly**, as the models often **segmented the entire person including the face**, resulting in the **face being inpainted into an avatar-like or fake appearance** instead of preserving it.

Second Approach (Most Accurate)

In the second approach, I divided the pipeline into **four clear steps**, which yielded better results:

- 1. Person Segmentation — Used YOLOv8-seg (large)** to accurately predict the person mask and separate the

subject from the background.

2. **Background Masking** — Created a clean **background mask** using the predicted person mask to ensure proper isolation between subject and background.
3. **Clothing Segmentation** — Applied my **custom-trained U-Net model** to isolate clothing regions from the human body.
4. **Inpainting Process** — Used inpainting models to modify specific regions:

- The **stable-diffusion-xl-1.0-inpainting** model performed **extremely well for background inpainting** but **struggled with clothes**, often ignoring the prompt.
- The **stabilityai/stable-diffusion-2-inpainting** model, on the other hand, **handled clothes better**, especially with **simpler prompts**, but was less effective on complex [backgrounds](#). **so i input to it the clothes mask**

Why this approach helped as now when we segment clothes we dont include face so model stick to the

clothes inpainting , we needed the yolo model to accurately get the person mask so also face wont be consider or body part as the background and be inpainted

The model failed in cases

where the segmentation of the person or clothing was inaccurate. This typically occurred when objects were present beside the person or when the background was plain or uniformly colored,

Also person **with glasses** model fail to generate the face correct again causing the model to misidentify boundaries and blend regions incorrectly.

Although the background was masked correctly, the model also failed to inpaint complex prompts,

such as changing blouses or pants, where fine details and textures were not accurately generated.

Additionally, since the segmentation model was trained on **women's clothing with 13 specific classes**, both the model and the entire pipeline

performed well only within those categories and struggled to generalize beyond them.

Prompts that are detailed or hard model fail to make them such as background detailed prompts or clothes direct prompts only worked

Generated



Real



