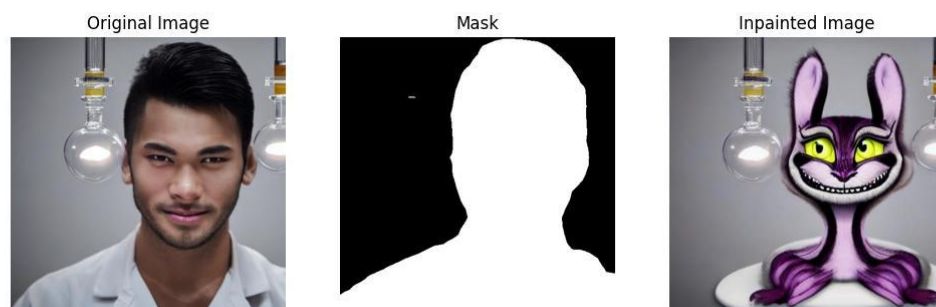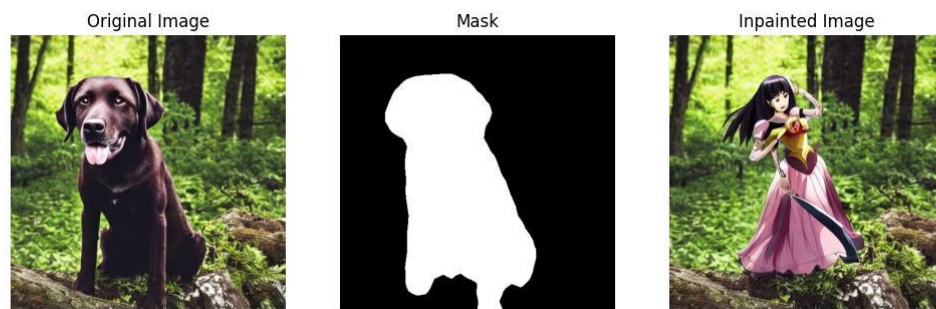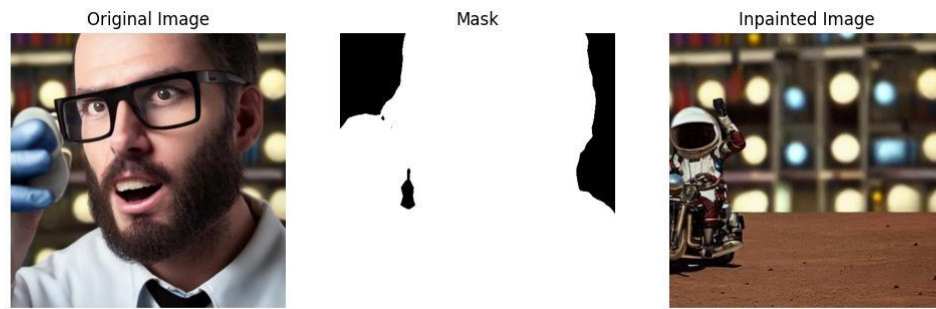# StoryCraft

Description pipeline:

1. The pipeline generates a photo-realistic image given any text input.
2. The image generated by the diffusion model is segmented automatically. Two types of segmentation are available: hair segmentation or segmenting the image into the following categories: background, person, cat, dog, and potted plant. The user needs to choose the type of segmentation.
3. Finally, an inpainting diffusion model takes the first generated image, its mask, and a new prompt. The inpainting model replaces the portion of the original image segmented with another image based on a textual prompt.

Limitations pipeline:

● The Diffusion models may not generate accurate faces or limbs because the training data doesn't include sufficient images with these features.
● The inpainting diffusion model was not specifically trained to inpaint hair. However, it was trained using samples that include cats and dogs, along with descriptions of their hair. Therefore, it is possible to modify the hair of a person using a textual description of animal hair.
● The Diffusion models don't perform well on more difficult tasks which involve compositionality, such as rendering an image corresponding to "A red cube on top of a blue sphere".
● The inpainting diffusion model was not trained
● Stable Diffusion inpainting typically works best with images of lower resolutions, such as 256×256 or 512×512 pixels. When working with high-resolution images (768×768 or higher), the method might struggle to maintain the desired level of quality and detail.
● Only two models for segmentation are available in the pipeline: DeepLab-v3 and hair_segmenter. Selecting the first model allows a change in the character while preserving the background (context). The second one allows you to keep the context and character, only performing a change in hair color.
● DeepLab-v3 segmentation model identifies segments for a number of categories, including background, person, cat, dog, and potted plant.
● Hair_segmenter model was designed to segment a single person's hair within images or videos captured by a smartphone camera. Hairstyles with thin and long pieces, such as mohawks and hairstyles with elongated braids, may not be reliably segmented.
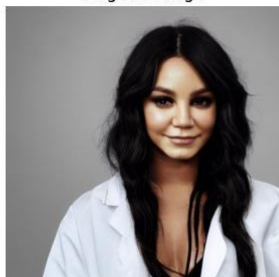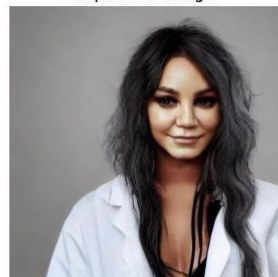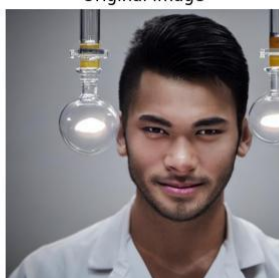
Example character inpainting:

Original Image	Mask	Inpainted Image

Original Image	Mask	Inpainted Image

Original Image	Mask	Inpainted Image

Examples hair inpainting:

Original Image	Mask	Inpainted Image

Original Image	Mask	Inpainted Image

Original Image	Mask	Inpainted Image