

"AI-Enhanced Product Photoshoot Visuals and Filter"

Key Objectives-

1. **Generative AI for Visuals:**

- Implement a generative AI model capable of creating realistic and visually appealing product photoshoot visuals.
- The model should be able to understand and simulate various lighting conditions, backgrounds, and angles commonly used in product photography.

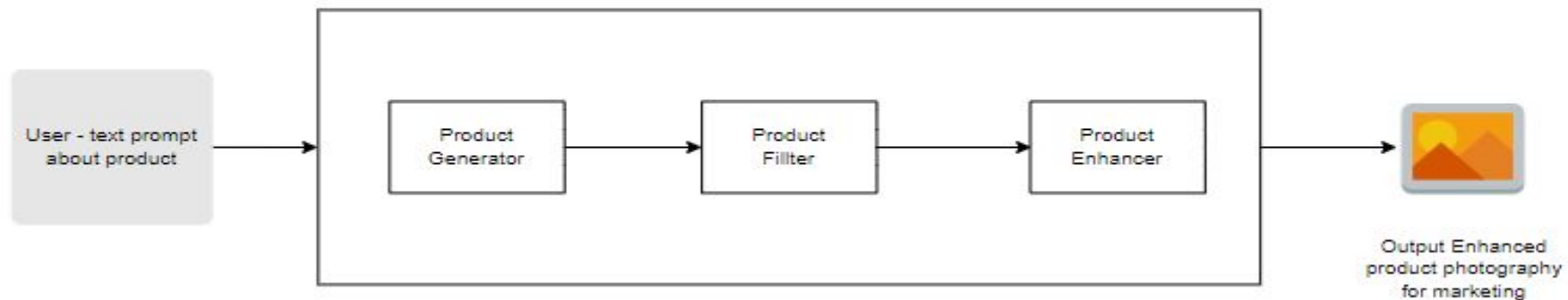
2. **Product Recognition Filter:**

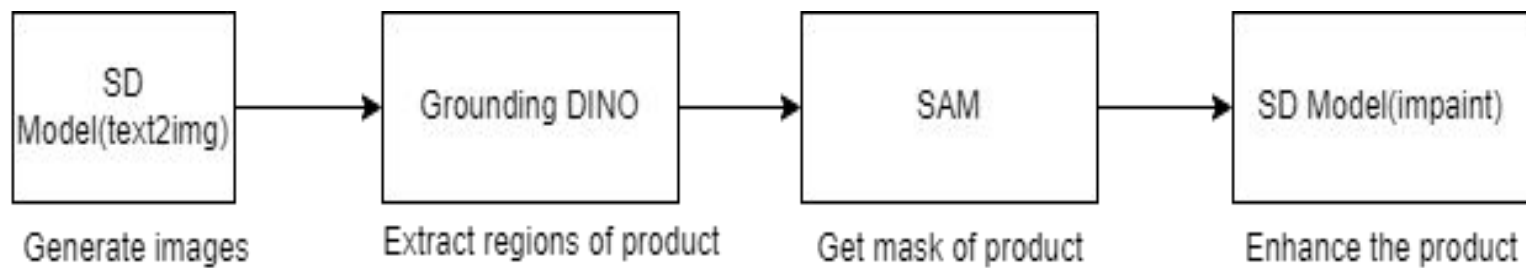
- Develop an AI filter that can identify and isolate specific products within a given image dataset.
- The filter should be trained to recognize a predefined set of products (e.g., apparel, electronics, accessories) and enhance their features while maintaining the overall image integrity.

3. **Exclusion of Non-Relevant Images:**

- Implement a mechanism to filter out images that do not contain any of the specified products.

Solution Approach





SAMPLE DEMO





- **Generative AI for Visuals:** SD XL based model
- **Product Recognition Filter:** SAM + SD based inpainting model
- **Exclusion of Non-Relevant Images:** We use grounding dino which takes image and object text as input, if the object is not present in image —> it will not go to Product Recognition Filter step

How to improve? Other approaches

- For the image generation model, we can improve the results by tweaking with the prompts
- By playing around with num_inference_steps, guidance_scale, negative prompts
- For the segmentation model, we can use other variants of SAM - HQ-SAM.
- For the enhancement module, we can tweak the text prompts, try different inpainting models

Where do the system fail?

As the overall system is a chain of pipeline, there is high chances of compounding error where can result in bad output result.

If the initial prompt for image generation is not optimal, it can result in artifacts in the images, due to which the detection and segmentation model can cause more errors which can result in bad image inpainting.

Further improvements:

As the task is focused on generating images for creating creative Ads, we can also extend the generated image to video Ads using img2vid SD models.

A Gradio based web ui to try the demo in Live

References:

- 1) <https://huggingface.co/docs/diffusers/en/using-diffusers/inpaint>
- 2) https://huggingface.co/docs/transformers/en/model_doc/grounding-dino
- 3) https://huggingface.co/docs/transformers/main/en/model_doc/sam