

# Assignment

## Problem Statement: *Edit pose of an object in a scene*

Recent advancement in generative AI has led to a development of a lot of creative workflows. One such workflow is to use generative AI techniques for editing product photographs after they have been shot at a studio, for example to polish the scene for displaying the product on an e-commerce website. One such post-production editing requirement could be editing the pose of the object by rotating it within the same scene.

This problem statement involves two tasks - for the eventual goal of developing technology for a user-friendly pose edit functionality. The first task is to segment an object (defined by a user given class prompt) in a given scene. This enables the 'user-friendly' part of the problem statement. The second task is to edit the pose of the object by taking user poses (e.g. Azimuth +10 degrees, Polar -5 degrees). The final generated scene should look realistic and composite.

Note: We do not expect you to train or fine-tune your own models, but use existing models (including the amazing one you have in your skull) and creatively find solutions to the two tasks.

You can go through the references for additional guidance and feel free to write at [harsh.maheshwari@avataar.ai](mailto:harsh.maheshwari@avataar.ai) for any clarifications.

## Tasks:

1. Task1. This task is to write an executable code that takes the input scene and the text prompt from the command line argument and outputs an image with a red mask on all pixels where the object (denoted in the text prompt) was present.  
(e.g. `python run.py --image ./example.jpg --class "chair" --output ./generated.png`)
2. Task2. The second task is to change the pose of the segmented object by the relative angles given by the user. You can use a consistent direction as positive azimuth and polar angle change and mention what you used.  
(e.g. `python run.py --image ./example.jpg --class "chair" --azimuth +72 --polar +0 --output ./generated.png`)

### The generated image:

- a. Should preserve the scene (background)
- b. Should adhere to the relative angles given by the user

## Examples:



Figure 1. The input scene,  
Class name to detect the object: "chair"  
Rotation: azimuth +72 degrees; polar +0 degrees



Figure 2. The output scene

## Attachments:

Example input images that you can try your algorithm with suitable text-prompts (file names are the reference to the object to be segmented) are attached [here](#).

## Deliverables:

As part of the submission, provide a link of the github repository with the code and a readme file with the results on the attached images (and other examples that you may want to show). Also describe how to execute the code, how did you approach the problem and any visual results from failed and successful experiments, if any. In your report, also analyse the successes and failures, and present ideas on how to fix failures/improve performance. You can create examples with small and major pose changes with easy or complex text prompts. Be creative and enjoy the exercise!

## Reference:

<https://segment-anything.com/>  
<https://zero123.cs.columbia.edu/>  
<runwayml/stable-diffusion-inpainting>  
<stabilityai/stable-diffusion-2-inpainting>