

Project 2

Instructions

- You need to mark which question you are answering clearly.
- You need to clearly explain all the steps of your solution for this project, including the code you wrote or code you referenced.
- GenAI tools can be used for coding, but not for writing the report.

Project Description

You have to teach a robot a sequence of actions with the following scene as our target:



The start scene will look a bit messy:



You need to complete the following.

A. Video Recording

On February 4, at 5:25 pm, we will collect data in **SCDI 4004!** For those in person, please go directly to SCDI 4004. Those online can watch the process via Zoom and potentially provide feedback (mainly because you all need to work with the recorded videos!). Please note that there are the following rules for the recording:

- 1) Only use one hand and manipulate one object at a time.
- 2) Do not try to do the same order as the other people! Every person can pick a different sequence of actions to complete the task.
- 3) Try to make your hand gestures obvious when grasping or releasing an object!

B. Object and Human Action Recognize [50%]

- 1) Identify a pretrained model or develop your own method that can detect objects in all videos. Provide details of the model you used or the method you developed and screenshots of the code and the results. [25%]
- 2) Identify a pretrained model or develop your own method that can recognize the actions performed in the videos (e.g., grasping, moving, releasing, etc.). Provide details of the model you used or the method you developed and screenshots of the code and the results. [25%]

C. Automatic Generation of Sequence of Actions [50%]

- 1) Identify a method or develop your own method that can generate a sequence of actions when applied to all the videos. Explain the method you used/developed and provide screenshots of the code and the results. Discuss the limitations of your approach if extended to other scenarios. [30%]
- 2) Design a Markov Decision Process based on all the videos. Define states, actions, probabilities, and rewards based on the videos. Show a diagram of the Markov Decision Process. [20%]