1. ***What did you learn after looking on our dataset?***
   First thing I notice those are the frames from videos taken in different angels. As videos are continuously recording, dataset consists images of different illumination. I also seen that images have different resolution. Also, we have two kinds of images firstly one has just illumination changes and have smaller size images and other have change in illumination changes and changing objects. Also, all images have the distortion.

2. ***How does you program work?***
   The program is simple. In the program I compare two images using the function already provided. after comparing if the score is less than the threshold, I will move the image into another folder. Need to run image_comparasion.py file with –dataset_path, --destination_path, --score_threshold, --min_countor_area, --and --image_extention. So, command my look like "**python image_comparision.py --dataset_path 'path of the dataset' –destination_path 'path to move the similar images –score_threshold 'integer' –min_countor_area 'integer' –image_extention '.png' "**. After running the program, the original dataset directory will have the images that are not so similar.

3. ***What values did you decide to use for input parameters and how did you find these values?***
   I choose min_countor_area = 5000 and score_threshold for the comparison as 10000. I choose them after series of trails. I have observed the max score is around 22000 to 25000 so thought 10000 can be a good choice for the threshold. I almost eliminated 60% of the similar images. So, if we have enough images, we can use high threshold so that our data have very versatile image selection. So, min_countor_area is threshold for the area of the contours we obtained from the function.

4. ***What would you suggest implementing to improve data collection of unique cases in future?***
   Now the dataset is more of different illuminations. I think if the camera is rotating, we can get different angle in the videos with that similar frames will be decreased in the dataset which increases the trainable images.

5. ***Any other comments about your solution?***
   I used the argsprase which is easy to play with parameters as we can give them in the command line. Even though I was a simple program I thought it is good have args. I have not uploaded the data as it is more for the GitHub repo.

Thank you so much for choosing me. I hope I did this task good enough to get chance for the interview. I am confident that I can convince that I am qualified for the position.