

Kopernikus Challenge for Machine Learning / Computer Vision Engineer role

1. What parameters you decided to use for the provided example dataset

In this dataset, I used very few parameters from my side.

One of them is the minCountourArea and other one is the black_mask whose values I modified from the original code.

2. How you found these values

To find the first parameter, the minContourArea, I made a separate method in which I thresholded the image using an auto threshold function and looked for contours in the image. After the contours got detected, I extracted the area of each of those contours. I stored the values of all contour areas into a list ignoring all small contours. After this I found the average of all the contour areas in the list and returned that value to the main function.

The black_mask value was found by playing around with the values a little so that the mask covers the top portion of the image which is a concrete bar or roof. I believed that the application of the project/ data collection is to detect and look for vehicles and the surroundings, which actually made the roof irrelevant. Hence, I made up a value with which the mask actually covers this roof region and not covering the parking region in the images.

3. What amount of duplicates script found with these parameters

All the duplicates were found using this script keeping a single copy and deleting all others.

4. What you would suggest improving to make data collection of unique cases better

Sometimes there can be abnormalities or noises happening in the images due to lack of light or aberration of lenses towards the edges or even some subjects blocking the field of view. Here, one can make use of multiple cameras for image collection so that the same script with differently calibrated parameters can be used in these cameras so as to improve accuracy.

5. Any other comments about `imaging_interview.py` or your solution

This script is designed for this provided dataset only. Any possible application to different images may require calibration to the values.