CV Assignment 2 Report

Aryan GD Singh
2019459

- Check included ipynb files for outputs of q1, q3, q4

Q1.

- Non-DL saliency maps are created using the StaticSaliencyFineGrained method in the saliency module of opencv.
- DL saliency maps are created using the VGG16 model.
  Code used - DL

- Based on eye-test, non DL saliency maps are better than DL based saliency maps
- Otsu thresholding is performed on the saliency maps.
- separation measure(ϕ) and concentration measure(ψ) values are calculated and stored.
- For calculation of ψ, the biggest component in image is assumed to be the background, and rest of the components are treated as foreground.
- Quality measure is the product of the above two mentioned quantities, this is also calculated.
- A csv file containing the quality measures for both DL and non-DL based saliency maps is created.

Mean scores are shown below

	phi(DL)	**phi(non-DL)**	**psi(DL)**	**psi(non-DL)**	**quality(DL)**	**quality(non-DL)**
mean	0.401517	0.420798	0.516099	0.353491	0.209271	0.152256

Overall, the DL based saliency maps performed better as they were much better than the non-DL maps at concentrating the foregrounds. But the DL model used may not have been the correct choice as it struggled in separating the foreground and background, performing slightly worse than the non-DL maps.

For the DL models, concentrating the foreground meant lower accuracy for the separation between foreground and background.

Q2.

- The images are converted to grayscale and resized to 256x256 for easier processing and patch division.
- Modified LBP as written in the assignment pdf is applied, for LBP, if the central pixel and adjacent pixel are both 0, then code 0 is used.
- SPP is applied using the image as -
    1. full image - 256x256
    2. 4 patches - 128x128
    3. 16 patches - 64x64
- The histograms for the patches are concatenated, finally giving us a 21x256 feature for each image.
- Fuzzy c-means clustering is performed with value of k = 5(variable)
- The labels are used to write the images into different subfolders.

Q3.

- The images are converted to grayscale and resized to 256x256 for easier processing and patch division.
- Images are divided into 256 16x16 patches
- HOG is calculated for each patch of each image in 8x8 cells normalized over 16x16 cells.

- This gives us a descriptors array of shape 50*256*36, and 256*36 feature for search image
- BoW model is trained using K-means clustering, with k chosen as 5(variable).
- This provides us with 5 centers of length 36 for each of the 256 patches, ie, vocabulary of shape 256*5*36
- For each patch in the search image, Euclidean distance with the corresponding 5 centers is calculated, and center with minimum distance is selected.
- Histogram for the search image is created that lists the number of patches belonging to each cluster.

Q4.

- Shi Tomasi corner detection is performed on the search image and 40 best corners are selected.
- Same is performed for the dataset images.
- Out of these 40, 25 viable(with patches lying inside the image) corners are selected.
- Patches of size 11x11 are created around these corners, and then LBP is applied for the 9x9 pixels inside the 11x11 border.
- Histograms are calculated for these patches, and are concatenated to form the feature for an image.
- Euclidean distance for the search image with the other images is calculated, and the k=5(variable) dataset images with least values are displayed.