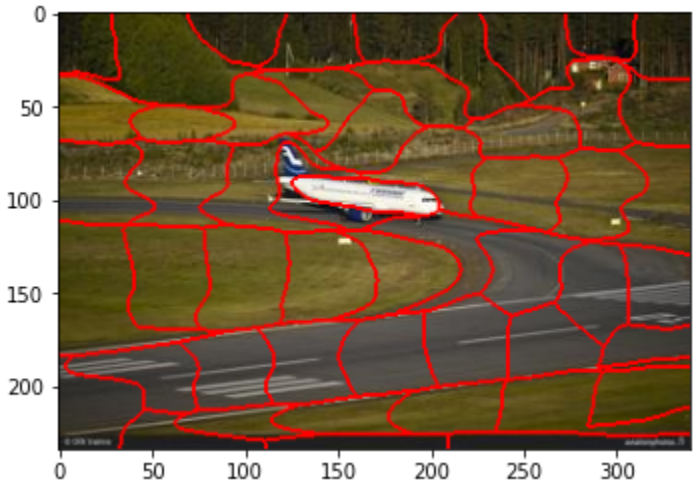
CV Assignment 3 Report

Aryan GD Singh
2019459

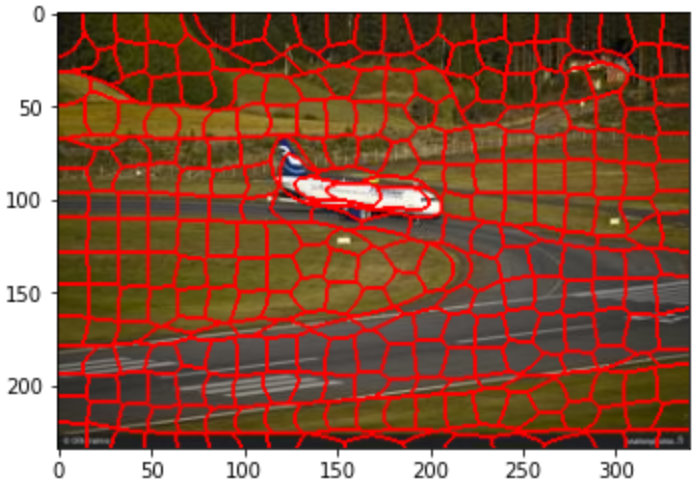
- Check included ipynb files for outputs of q1, q3, q4

Q1.

- The image is divided into superpixels using skimage's implementation of SLIC.
- Different number of superpixels are tested: 50, 100, 200, 300, 500, 800
- As the number of superpixels increases, the superpixels start becoming more gridlike.

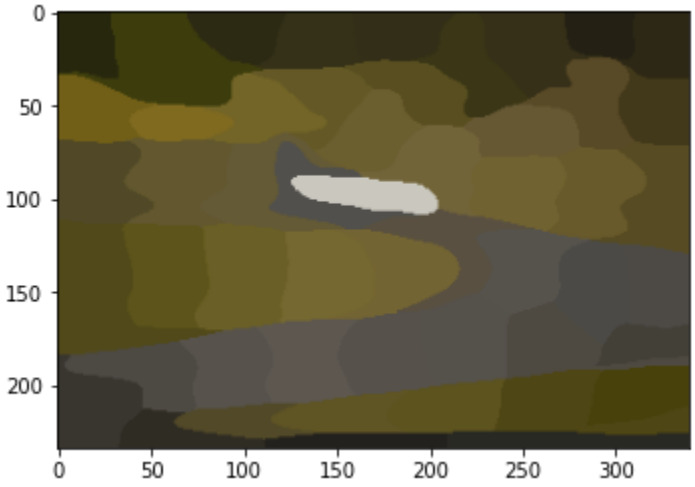


SP = 50

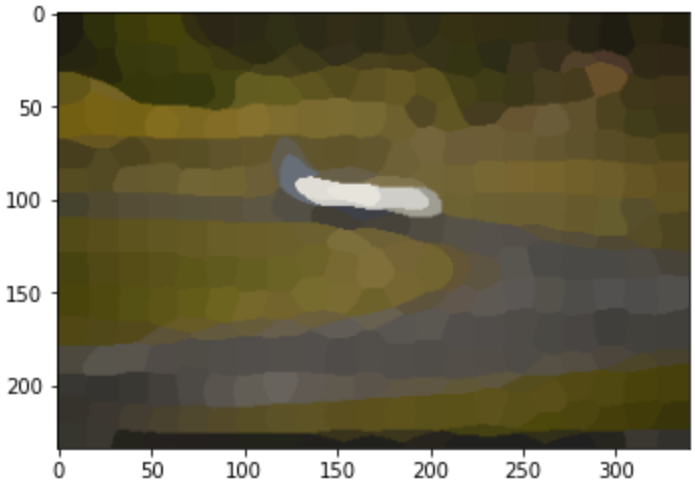


SP = 300

- Coloring the superpixels with their mean values leads to the segmented images.

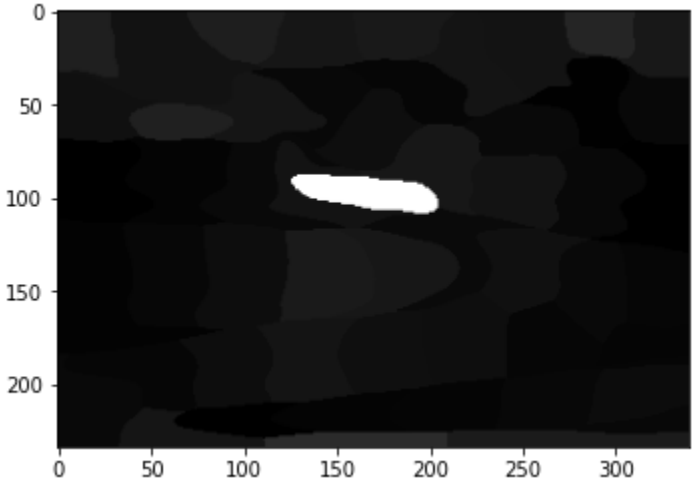


SP = 50

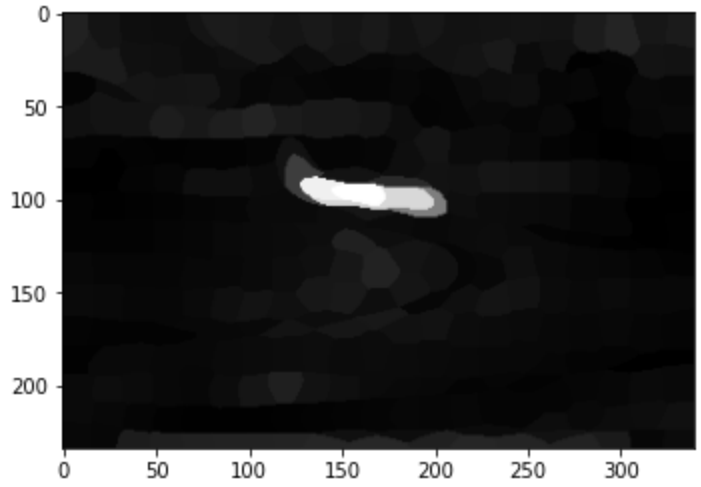


SP = 300

- Color and Location of each superpixel is calculated. Color is the mean value and location is the centroid of the pixels belonging to the superpixel.
- Saliency maps are computed using the given formula.



SP = 50



SP = 300

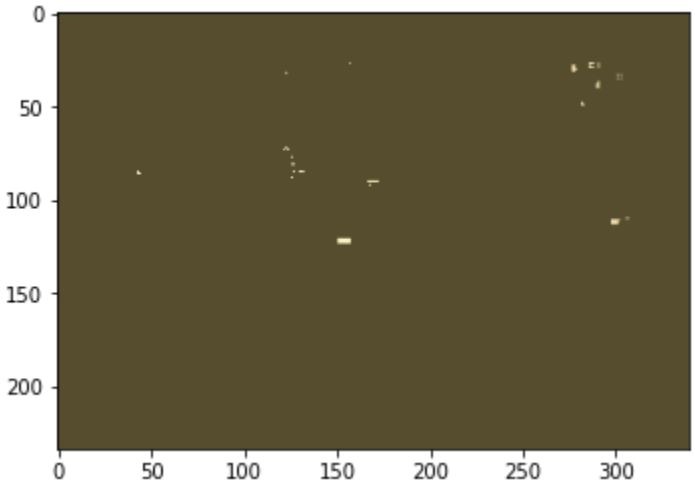
Overall, as the number of superpixels is increased the background noise starts affecting the saliency map more, but the map also can capture the shape of the foreground more clearly. From the eye-test, 50 superpixels saliency map was the best.

Q2.

- Following mean shift methodology, the coordinates of the pixels are added to the RGB values.
- Image is flattened so that we have 79560 values for the pixels to be clustered using DBSCAN.
- Clustering is performed and labels are acquired.
- -1 labels, ie, noise labels are set to belong to the cluster which they have minimum euclidean distance with.
- DBSCAN is performed for a wide range of parameters.
    a. sample = [10, 20, 50, 100]
    b. eps = [5, 10, 15, 25]

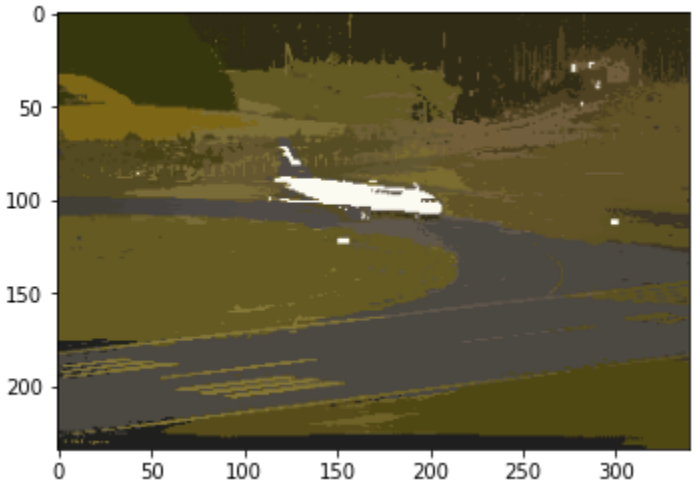
One way to choose the best performing parameters can be to choose the ones which lead to the minimum number of noise labels.

This was given by sample = 10 and eps = 25



As we can see, this is not good segmentation, as almost the whole image is segmented into 1 big segment.

From the eye-test, best segmentation was performed with sample = 50 and eps = 10



Q3.

SIFT vs SURF

SIFT has a faster version called SURF. Lowe used Difference of Gaussian to estimate the Laplacian of Gaussian in SIFT to determine scale-space. With Box Filter, SURF goes a step farther and approximates LoG. Convolution with a box filter can be simply estimated using integral pictures, which is a huge advantage of this approach. It's also possible to perform it simultaneously for different scales. In addition, the SURF uses a Hessian matrix determinant for both scale and location. SURF uses wavelet responses in the horizontal and vertical directions for a neighborhood of size 6s to assign orientation. It is also given appropriate gaussian weights. Calculating the sum of all replies inside a sliding orientation window of angle 60 yields the dominant orientation. In general, while SURF is faster, SIFT is slightly more robust.