Computer Vision
Assignment 3 - Rubric

Q1,

You need to verify the following for question 1 through their code/report-

1. SLIC super pixels are calculated. **(0.25 mark)** (if this is not done, give 0 in the entire question)
2. Aggregation (any method of aggregation accepted) of colour values in a super pixel to represent a single RGB value for each super pixel. **(0.25 mark)**
3. Aggregation (any method of aggregation accepted) of pixel locations to represent a single location for each super pixel. (**0.25 marks)**
4. The saliency value of each super pixel is calculated as per the given formula in the question. The formula is given below for reference. **(1 mark)**

$$Sal(SP_i) = \sum_{j \in \{1,\cdots,|SP|\}} ||SP_i(c) - SP_j(c)||_2 \times e^{-\left(\frac{||SP_i(l)-SP_j(l)||_2}{\sqrt{width^2+height^2}}\right)}$$

where:
SP_i denotes $i-th$ super-pixel.
$SP_i(c)$ denotes color of the $i-th$ super-pixel.
$SP_i(l)$ denotes location of the $i-th$ super-pixel.
$||\cdot||_2$ denotes L2-norm (Euclidean Distance).
e denotes the base of natural algorithm.
$|SP|$ denotes number of super-pixels.
$width$ denotes the width of the image.
$height$ denotes the height of the image.
Sal denotes the saliency value

Q2,

You need to verify the following for question 2 through their code/report-

1. They have used varying values of min points/samples and epsilon. **(0.25 mark)**

2. Must have shown, with image examples, the difference between using different min points/samples and epsilon values. **(0.25 + 0.25 mark)**
    a. It should be shown for min points/samples, that as min points increase, the minimum number of points for a cluster to be formed is increased and thus small clusters are not formed. (0.25 mark)

    b. It should be shown for epsilon, that as epsilon increases, far away points can now belong to the same cluster. Thus, it reduces the formation of smaller clusters. (0.25 mark)

3. Must have shown their best Clustering output along with the parameters used to obtain it. **(1 mark)**

<u>Q3,</u>

You need to verify the following for question 3 through their code/report-

1. Ran the code on any sample image and compared speeds **(0.75 mark)**

2. For the first part of the question, there should be atleast three points that mention the improvements made by SURF. Below are some of the most prominent improvements made by SURF. (This is not an exhaustive list, it is possible that some improvements aren't mentioned below). If the TA is satisfied with the points mentioned they can award marks for the same.

    The following are some of the key differences (or advantages) that SURF has over SIFT. **(0.75 mark)**
    a. SURF uses a Hessian Matrix based measure which uses a very basic approximation technique like DoG (Difference of Gaussian) for the detector. This method helps speed up the computations significantly as compared to the SIFT Method.

    b. The SURF Method makes use of *Integral Images*. They allow for the fast implementation of box type convolution filters. The entry of an integral image $I\Sigma(x)$ at a location x = (x, y) represents the sum of all pixels in the input image I of a rectangular region formed by the point x and the origin. In simpler terms an Integral Image can be thought of as a way to store the information of all pixel values from origin to the current point. This again offers significant improvements in computational time.

c. SURF makes use of *box type convolutional filters*. Due to the use of box filters and integral images, the same filter is not required to be applied iteratively to the output of a previously filtered layer, but instead could be applied to filters of any size at exactly the same speed and directly on the original image.

d. SURF works with a considerably lesser number of dimensions (64) as compared to the SIFT method, thereby reducing the time for feature computation and matching, and increasing simultaneously the robustness. SURF also presents a new indexing step based on the sign of the Laplacian, which increases not only the matching speed, but also the robustness of the descriptor.

e. The SURF descriptor describes a distribution of *Haar-wavelet responses* within the interest points in the neighbourhood. This increases the accuracy of the SURF descriptor in feature extraction.

Each point is for **(0.25 marks).** Atleast any three points should be mentioned in the report.