16720-A Computer Vision: Homework 1

Name: Harsh Dhruva

Andrew ID: hdhruva

Q 1.1.1

The Gaussian filter blurs the image to pick up the larger regions and spatial features in the image. The Laplacian of Gaussian filter picks up the edge that are present in the image. Derivative of Gaussians in x and y pick up the edges with vertical component and horizontal component in the image respectively.

We need multiple scales to extract spatial features of different sizes, which will make our extracted feature responses more comprehensive, improving the performance of classification.

Q 1.1.2
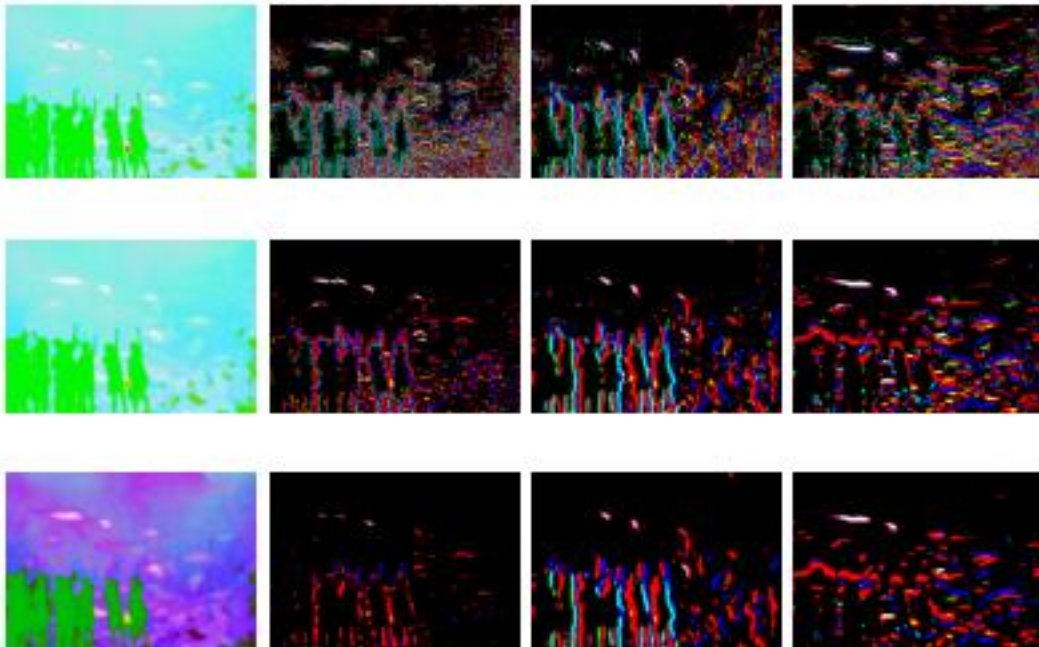
Q 1.3
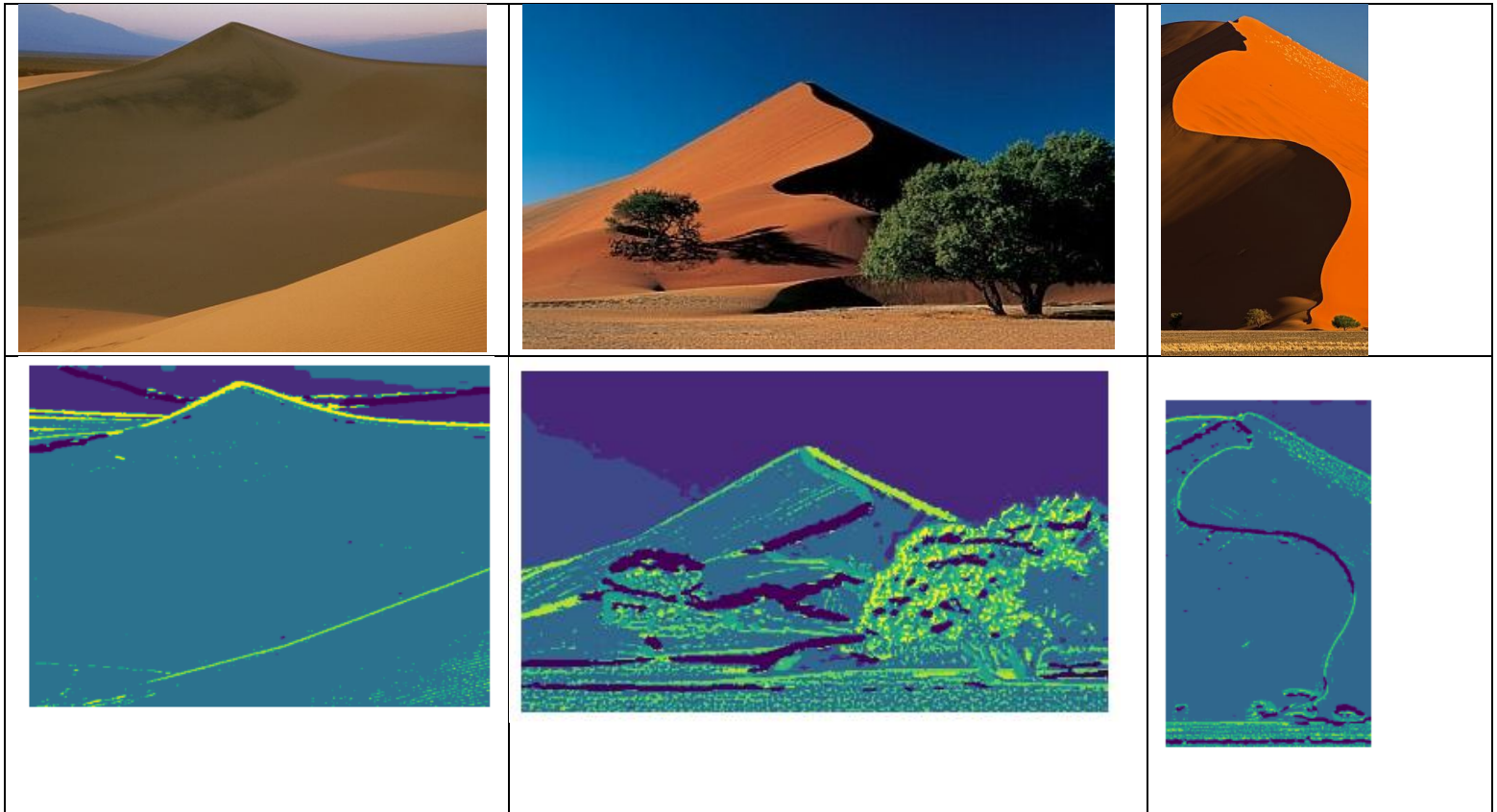
The word boundaries make sense to me. In the wordmaps you can clearly see the difference in the particular regions through the colors associated with them, as the region is mapped to particular visual words. The sky and sand can be distinctly recognized in the wordmap. Even the distinct boundaries are visible within the images.

Q 2.5

Confusion Matrix:

[[25. 0. 5. 0. 4. 3. 7. 6.]

 [0. 30. 5. 6. 1. 1. 2. 5.]

 [5. 4. 22. 2. 1. 1. 7. 8.]

 [1. 2. 2. 34. 7. 0. 3. 1.]

 [4. 1. 1. 17. 19. 2. 3. 3.]

 [4. 0. 1. 3. 1. 26. 11. 4.]

 [5. 1. 4. 0. 6. 6. 18. 10.]

 [4. 1. 11. 3. 6. 3. 4. 18.]]

Accuracy:

48% with **default parameters**.

Q 2.6

Some of the hard classes and samples were the "kitchen" and the "laundromat", since the two categories are quite similar, as both being rooms with floors and ceilings shared a lot of features. Other challenging classification was between the windmill and the highway, since they both contained large regions of sky, and ground, and some windmills were of large size which could be mistaken as the road in the highway.

Q 3.1

- filter_scales: Increased number of filter scales from 2(default) to 3 achieves accuracy 51%
- K: Increased K from 10(default) to 20 achieves accuracy 49%
- Alpha: increased value from 25(default) to 100 achieves accuracy of 53.75%

[[29. 1. 0. 2. 2. 3. 8. 5.]

[0. 38. 2. 7. 0. 0. 1. 2.]

[2. 5. 27. 1. 2. 2. 3. 8.]

[3. 0. 2. 31. 12. 0. 2. 0.]

[5. 0. 1. 14. 23. 1. 4. 2.]

[1. 0. 1. 3. 0. 26. 16. 3.]

[9. 1. 3. 1. 2. 6. 20. 8.]

[2. 1. 12. 4. 6. 3. 1. 21.]]

Accuracy: 0.5375

I increased the filter_scales which improved my accuracy slightly, potentially due to a better visualization of larger features. When I increased the K value keeping the filter scales the same, the accuracy went back down a little. This could be due to the over classification of visual words, potentially increasing the error in mapping. However, when the alpha value was increased the accuracy went up. This is because there were larger number of pixels sampled, thus providing a larger data set to map to. Therefore, with the increased value of both K and alpha the accuracy went up.