16-720 Computer Vision: Homework 1

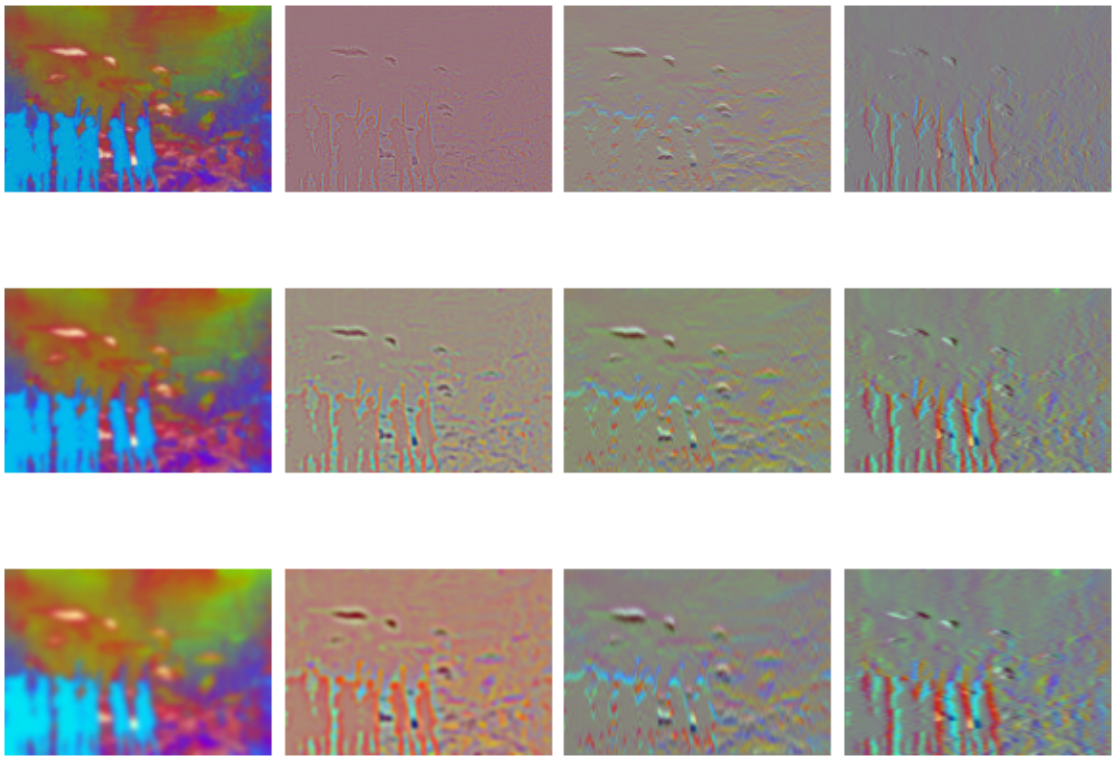
Thomas Zhang

September 2023

1.1.1

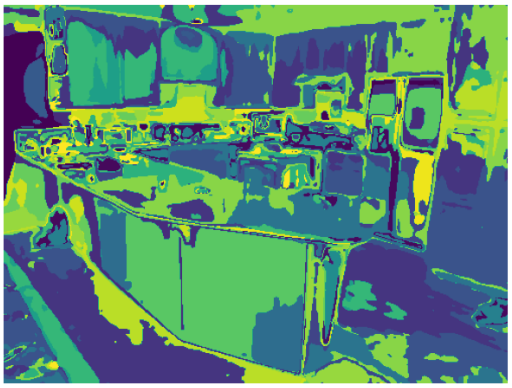
The Gaussian filter smooths the image causing edges to become blurred. The Laplacian filter makes all edges more distinct while also causing some blurring. The derivative of the Gaussian in the x direction causes vertical edges to become more apparent while also blurring the image somewhat in the horizontal direction. The derivative of the Gaussian in the y direction does the same as the derivative of the Gaussian in the x direction except it effects the horizontal edges and causes vertical blurring.

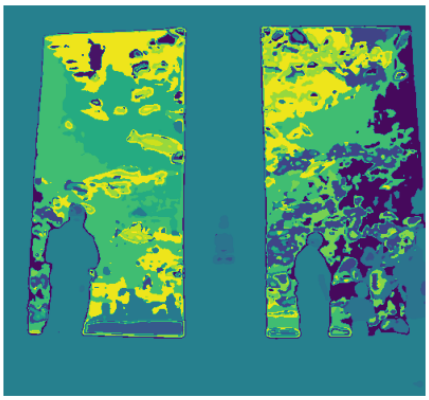
1.1.2

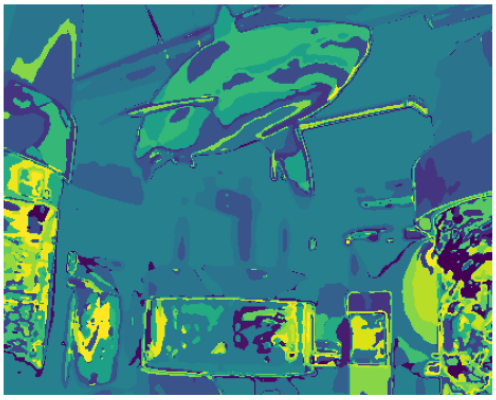


Filters used from left to right: Gaussian, Laplacian, Derivative of Gaussian in y-direction, Derivative of Gaussian in x-direction. Scales used were 1,2, and 3.

1.1.3







2.5

Confusion Matrix and Accuracy.

```
[[42.  1.  2.  0.  0.  3.  4.  2.]
 [ 0. 33.  4.  2.  0.  0.  0.  2.]
 [ 0.  3. 34.  0.  2.  4.  1. 11.]
 [ 1.  5.  0. 34. 12.  0.  1.  0.]
 [ 3.  1.  1.  9. 30.  1.  6.  0.]
 [ 1.  1.  2.  2.  1. 36.  5.  2.]
 [ 1.  1.  3.  2.  5.  2. 30.  2.]
 [ 2.  5.  4.  1.  0.  4.  3. 31.]]
0.675
```

For the Confusion Matrix. Predicted label is row, true label is column.

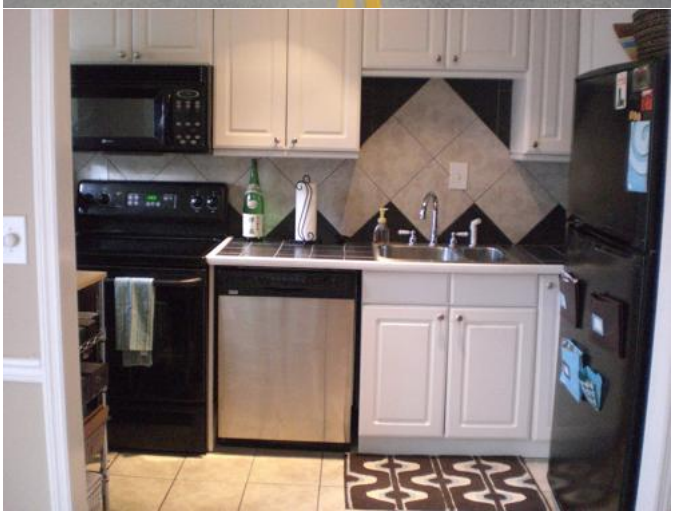
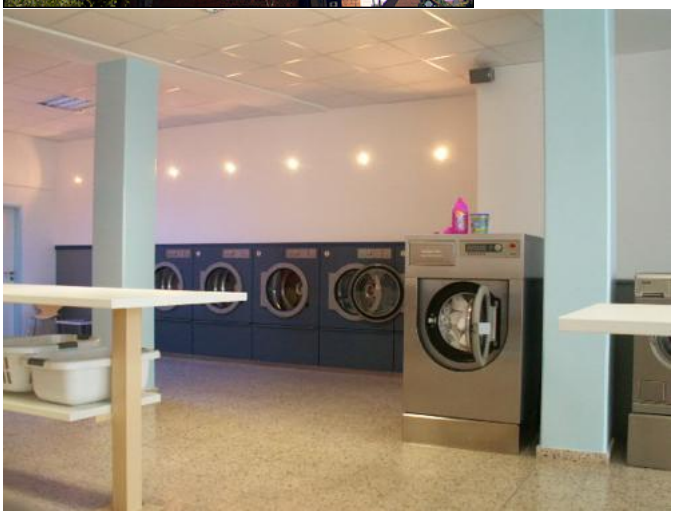
0	1	2	3	4	5	6	7
aquarium	desert	highway	kitchen	laundromat	park	waterfall	windmill

Table 1: Numerical label to descriptive label.

2.6

Hard Cases are highways [2] getting mixed up with windmills [7] and kitchens [3] getting mixed up with laundromats [4]. One possibility for why highways and windmills are more easily confused by the model is that the images for highways and windmills both tend to have straight lines towards the bottom of the pictures (roads or windmill tower) and towards the top of the picture there is a large region of sky.
For kitchens and laundromats, there are some similarities because there's a lot of solid colors from cabinets and laundry machines. Then there are a lot of horizontal and vertical lines interspersed throughout the image from the edges of laundry machines for laundromat images and tables/cabinets for kitchen images.



Filter Scales	K	Alpha	L	Accuracy
1,2	10	25	1	0.515
1,2	10	25	2	0.56
1,2	10	50	2	0.5775
1,2	10	100	2	0.575
1,2,3	10	25	2	0.5675
1,2	20	50	2	0.6025
1,2	30	50	2	0.675
1,2	40	100	2	0.6475

Table 2: Ablation table

3.1

Increasing the number of layers from 1 to 2 increased the accuracy from 0.515 to 0.56. Increasing L, increases the number of sub-sections and lets the model identify finer features in the image. This most likely contributes the most to the relatively large jump in accuracy from increasing L.
Increasing the value of alpha from 25 to 50 or 100 slightly increase the accuracy (from 0.56 to 0.57). This makes sense because increasing alpha, increases the number of sampled points so the model is more likely to pick up points containing important features that can be used to distinguish the image. However, there are diminishing returns since the distinguishing features of an image take up relatively small portions of the image.
Increasing the value of K from 10 to 20 and 30 caused a relatively large jump in accuracy (from 0.5675 for k = 10 to 0.6025 for K = 20 and 0.575 for k = 30). This makes sense because K is the number of words in our "dictionary" and the more words available to describe an image the easier it is to distinguish between different types of images.