

Q1.1.1

- a. Gaussian filter

Pick up the low frequency features.

- b. Laplacian of Gaussian

Pick up the high frequency features such as edges with all directions.

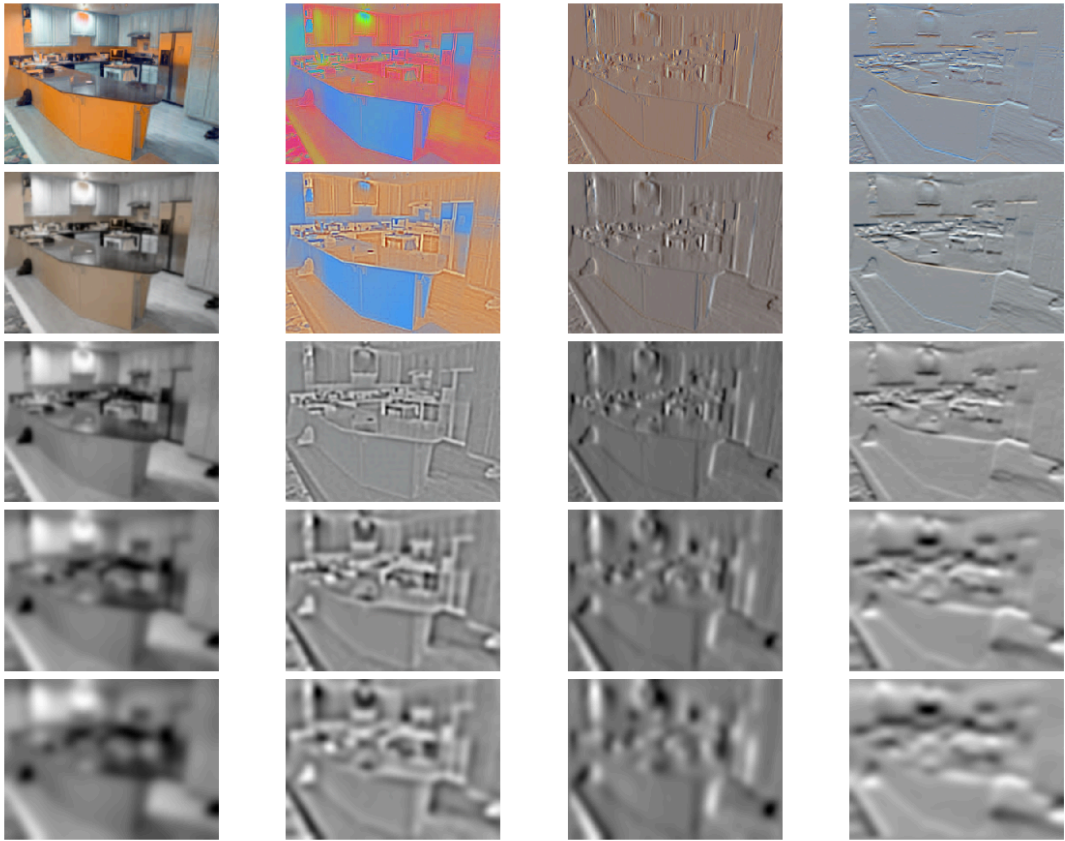
- c. Derivative of Gaussian in the x direction

Pick up the high frequency features such as vertical edges.

- d. Derivative of Gaussian in the y direction

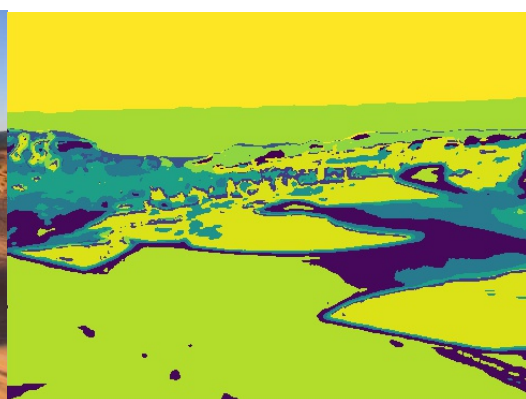
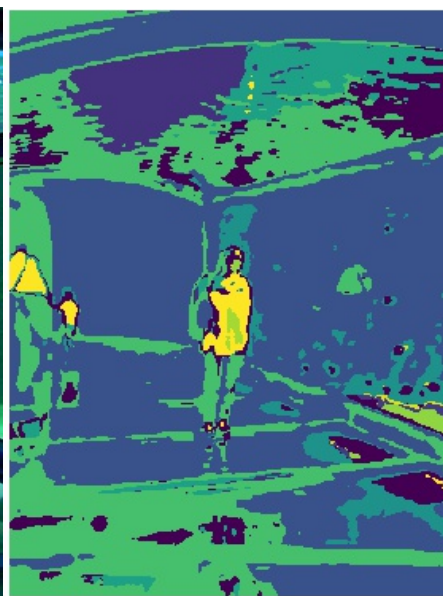
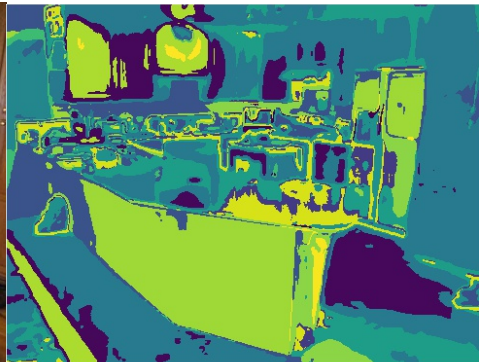
Pick up the high frequency features such as horizontal edges

Q1.1.2



Q1.3

These “word” boundaries make more sense to me on the flat regions since since these regions can represent the features scene classification better.



Q2.5

K = 120 and alpha = 150

Accuracy = 62.50%

```
0.625
[[14.  0.  2.  1.  2.  0.  0.  0.]
 [ 0. 13.  0.  2.  0.  2.  6.  1.]
 [ 0.  0. 11.  0.  0.  1.  0.  0.]
 [ 0.  1.  4. 17.  0.  1.  1.  2.]
 [ 0.  0.  4.  0.  9.  6.  1.  2.]
 [ 0.  1.  1.  0.  2. 12.  1.  0.]
 [ 0.  2.  0.  2.  0.  2. 12.  2.]
 [ 0.  1.  3.  4.  0.  0.  0. 12.]]
```

Q2.6

Here are some examples that are classified wrongly. I think the reason of these failure examples is that most part of these example are not the target scene to be classified. For example, first picture contains a man in the middle which contains no information of laundromat.



Q3.2

Accuracy = 93.13%

```
0.93125
[[13.  0.  0.  0.  0.  0.  0.  0.]
 [ 0. 16.  1.  1.  0.  0.  0.  1.]
 [ 0.  0. 23.  0.  0.  0.  1.  0.]
 [ 0.  1.  1. 25.  0.  0.  0.  0.]
 [ 0.  0.  0.  0. 13.  3.  0.  0.]
 [ 0.  0.  0.  0.  0. 21.  0.  0.]
 [ 1.  1.  0.  0.  0.  0. 20.  0.]
 [ 0.  0.  0.  0.  0.  0.  0. 18.]]
```

Yes, the pre-trained VGG16 model perform much better than the classical BoW. I think the way VGG16 extract features is done by the CNN model whose weights are well-trained through a large training set.