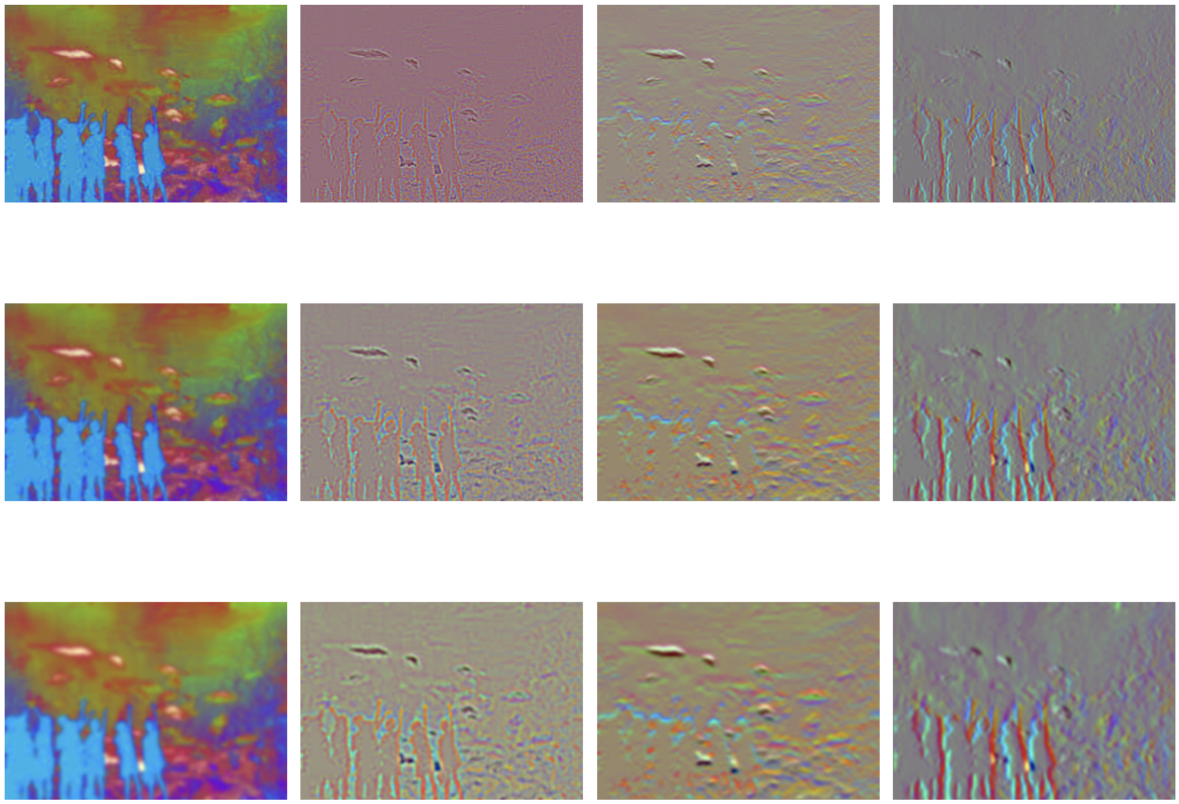
1.1.1

- Gaussian filter - The gaussian filter blurs the image, softening the sharp edges which makes the image more uniform, removing the noise, and false detections.
- Laplacian of Gaussian filter - The laplacian of gaussian filter picks up the edges in the image, since it is sensitive to rapid changes in intensity.
- Derivative of Gaussian in the X direction - highlights the colour variations in the X direction and, finds the verticle edges in the image
- Derivative of Gaussian in the Y direction - highlights the colour variations in the Y direction and, finds the horizontal edges in the image
- Scales - since different images have varying levels of noise, and size of features, we need to try different scales to account for these variations.

1.1.2



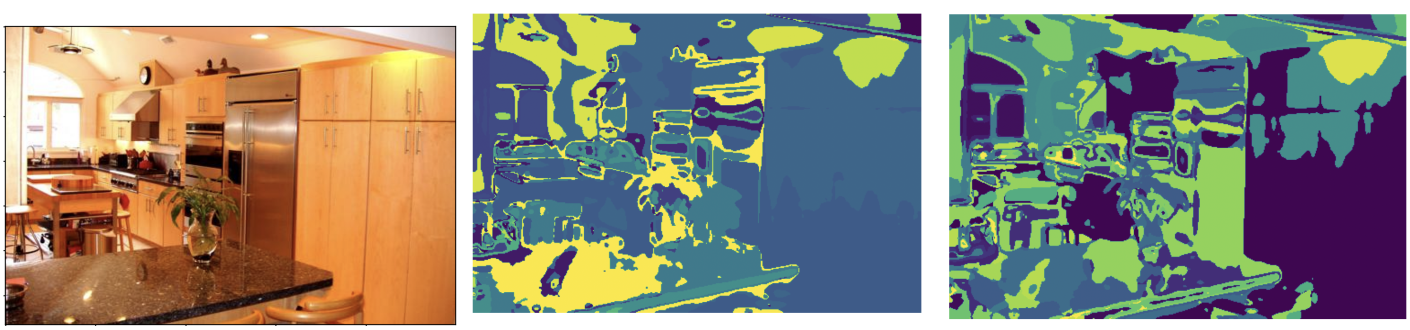
Visualization of the filter responses for 4 filters (Gaussian, Laplacian of Gaussian, derivative of Gaussian in the X direction, and derivative of Gaussian in the X direction) and 3 scales (1,2,3)

1.3

The wordmaps make sense since different features of the image are highlighted by different colors (words). As expected, different objects are colored differently: the colour assigned to the road, and the sky are different, that of the people and the aquarium are different.
As the number of clusters increase, the wordmap becomes more and more detailed, highlighting almost every edge in the image.
Below are the wordmap visualizations and their corresponding RGB images.

Original RGB images

Wordmap visualization for the default values (K=10, α = 25, scales = (1, 2))



Wordmap visualization for higher values of K and alpha, for the RGB image

2.5

For the default values as mentioned in the question,

```
[[27.  6.  4.  1.  4.  4.  7.  5.]
 [ 0. 13.  4.  2.  2.  0.  0.  3.]
 [ 1.  9. 15.  2.  0.  4.  1.  5.]
 [ 0. 10.  2. 28. 13.  2.  2.  2.]
 [ 8.  9.  3. 16. 21.  4. 10.  4.]
 [ 2.  1.  4.  1.  3. 25.  8.  5.]
 [ 5.  0.  9.  0.  6.  6. 20.  3.]
 [ 7.  2.  9.  0.  1.  5.  2. 23.]]
```

Screenshot of the confusion matrix

```
accuracy  0.43
```

Screenshot of the accuracy

2.6

Hard examples

As seen in the screenshot, we can see that the first pair with the most false positives are kitchen and laundromat (classes 3,4). In the next pair, highway is often mixed with windmill.

```
[[24.  0.  0.  1.  0.  1.  3.  1.]
 [ 8. 39.  8.  6.  4.  1.  8.  7.]
 [ 0.  4. 25.  2.  4.  4.  1.  8.]
 [ 2.  4.  2. 26. 13.  1.  1.  0.]
 [ 2.  0.  2. 10. 23.  3.  3.  3.]
 [ 4.  0.  2.  2.  1. 34.  5.  4.]
 [ 4.  2.  0.  2.  4.  3. 27.  0.]
 [ 6.  1. 11.  1.  1.  3.  2. 27.]]
0.5625
```

Confusion matrix and accuracy for final training

Examples from the kitchen-laundromat pair. As we can see, the colors, and objects are very similar across both pictures and can be easily confused by a vision model.

3.1

Ablation study

L	K	Alpha	Scales	Accuracy
1	10	25	1,2	43%
3	100	100	1,2,4,8,12	52.5%
3	50	100	1,2,4,8,12	55.5%
3	150	250	1,2,4,8,12	56.25%

We can see that as the number of alpha is increased, we can see the accuracy going up, since the number of points being considered for comparison are increased. We can also see that as the number of layers in the pyramid (L) increases, the accuracy increases, since more features are considered.