Nanditha Umasankar
Andrew ID: numasank

16-720 Computer Vision - Homework 1

1.1 Extracting Filter Responses

Q 1.1.1

1. Gaussian Filters smoothen images and remove high-frequency noise. This helps avoid picking up unnecessary details while detecting features.
2. The Laplacian of Gaussian filter is the second derivative of the Gaussian function. It detects contours in images using the zero-crossings of the filtered response.
3. Derivative of Gaussian in x detects the change in x-axis which helps to pick the vertical edges in the image
4. Derivative of Gaussian in y detect the horizontal edges by virtue of picking changes in the y-axis.

Nanditha Umasankar
Andrew ID: numasank

Q 1.1.2

The filter responses for the image **'aquarium/sun aztvjgubyrgvirup.jpg'** are displayed below. The scales are [1 2 4 8 8√2].

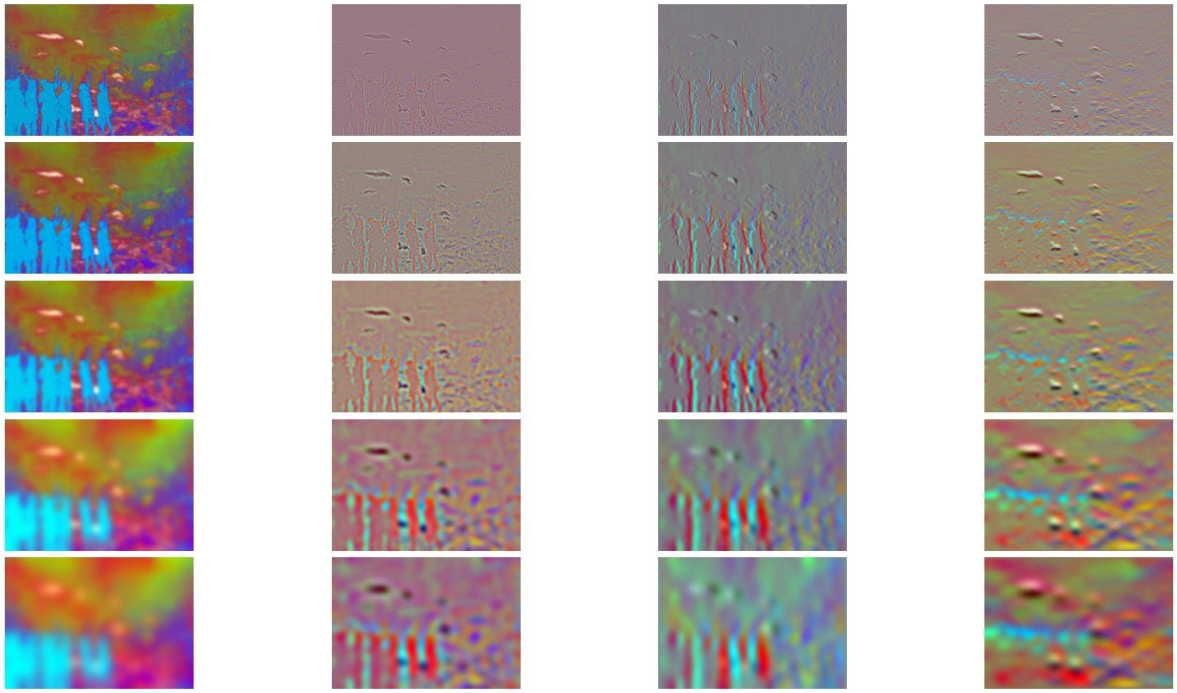


Image: 'aquarium/sun aztvjgubyrgvirup.jpg'

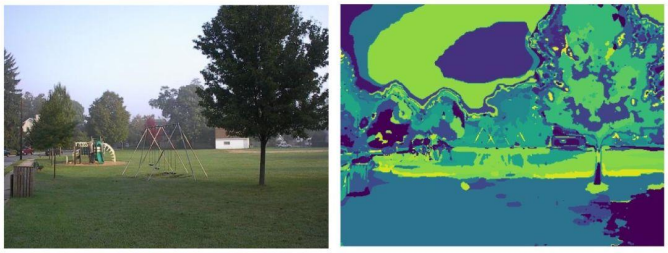
Nanditha Umasankar
Andrew ID: numasank

Q 1.3 Visual words

1. ‘desert/sun_adpbjcrpyetqykvt.jpg

2. ‘windmill/sun_bhkjsgaqdxnikeej.jpg’

3. ‘park/sun_bfruscgbtatjlcjw.jpg’.



- The wordmap is able to map certain objects with great distinction - like the windmill. These images are likely to have great detail with sharp changes in color and/or texture.
- The light and shadow in the scene also impact the wordmap, for example, the desert image has fewer details but there is a clear distinction of the sand dune due the sharp change in pixel values at the edge.
- The gradation in the color of the sky in the park image has been picked up by the wordmap as distinct features. Since the image is mostly monotonous, lesser details can be differentiated except for any distinct shadows or changes in textures.

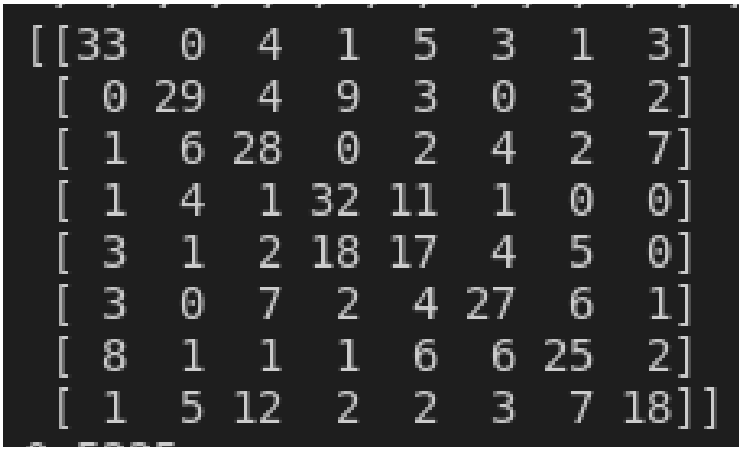
Nanditha Umasankar
Andrew ID: numasank

Q 2.5 Evaluation

With the default parameters K = 10, alpha = 25, and L = 2, the trained system gives the following results.

Accuracy: 52.25%

Confusion Matrix:

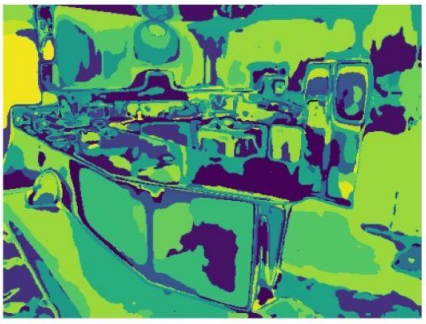


Nanditha Umasankar
Andrew ID: numasank

Q 2.6 Finding failures

From the confusion matrix, we can observe that the highway, kitchen, and laundromat images have the highest error in prediction.
On closer look, we can see that the laundromat and kitchen images have a lot of details that have been picked but due to the similar structures, they have been wrongly classified.



The highway images contain fewer features, and hence many of them have been incorrectly categorized.


Nanditha Umasankar
Andrew ID: numasank

Q3.1 Hyperparameter Tuning

To improve the accuracy of the evaluation system, the system has been trained using different hyperparameters - K, alpha, and L. The accuracy for each case has been tabulated below.

	Default values			
filter_scales	5	5	5	5
K	10	30	100	150
alpha	25	70	125	175
L	1	3	2	2
Accuracy	52.25%	47.25%	54%	61.25%

The L parameter doesn't affect the trained system a lot. Keeping it at L = 2 provides fair results.

Increasing K has an improved accuracy but the value cannot be higher than a certain value since it may classify closely comparable features as completely distinct features. Maintaining an optimum relationship between alpha and K is pertinent to not overfit the features.