16-720 Computer Vision
Spring 2020

Homework 1

Zongwen Mu, Andrew ID: zongwenm

1 Representing the World with Visual Words

1.1 Extracting Filter Responses

1.1.1

The Gaussian filter reduces the noises and high frequency responses in images, which would make it easier for classifier to find similarities between images.

The derivative of Gaussian filter in X direction picks up distinct vertical features.

The derivative of Gaussian filter in Y direction picks up distinct horizontal features.

The Gaussian Laplacian filter picks up the contours in the images.

We need multiple scales of filter response so that the classifier could work with images with different sizes.

1.1.2

Set default scales as $[1\ 2\ 4]$, the collage of images is shown below:

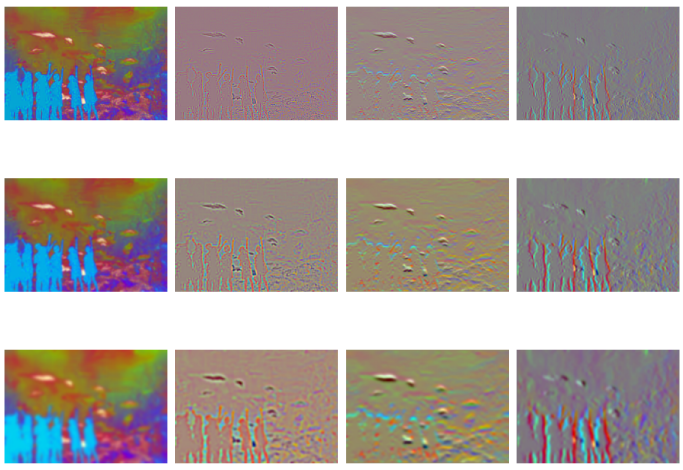
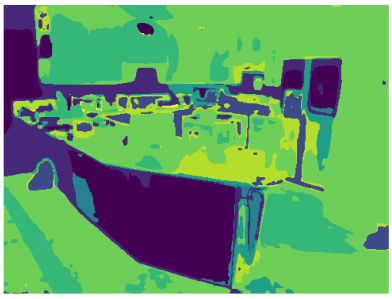


Figure 1: collage of images

1.3 Computing Visual Words

Image1: /kitchen/sun_aasmevtpkslccptd.jpg

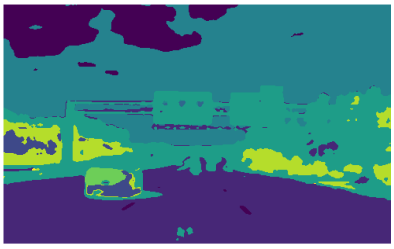
(a) RGB image



(b) wordmap

Image2: /highway/sun_aqpxkrnzgghnpmvu.jpg

(c) RGB image



(d) wordmap

Image3: /waterfall/sun_bjgujnjeakwxtjtz.jpg

(e) RGB image

(f) wordmap

Comment: The "word" boundaries could be matched with the highlights, shadows, contours and blobs in RGB images easily.

2 Building a Recognition System

2.5 Quantitative Evaluation

With default parameters $K = 10$, $L = 2$, the confusion matrix is shown below:

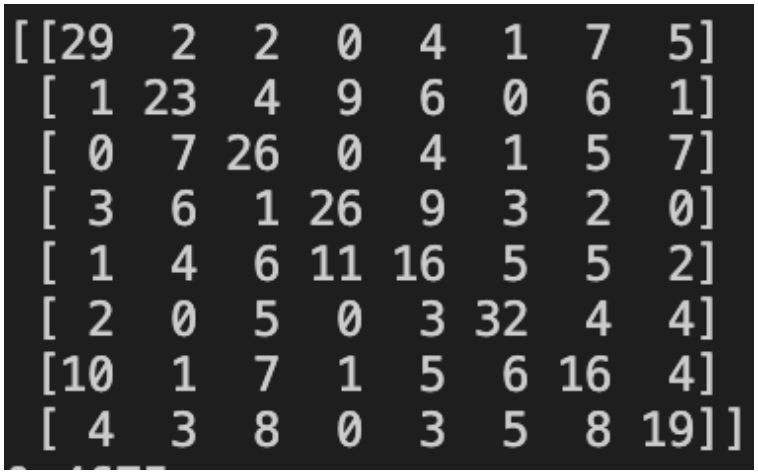


Figure 2: confusion matrix

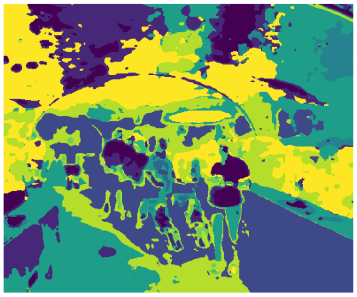
The accuracy is 46.75%

2.6 Find the Failures

Many aquarium images are classified as waterfall.

Sample: sun_aqzlijeqktetgrtd.jpg

(a) RGB image



(b) wordmap

The wordmap contains large amount of green and yellow which resembles the nature sight of waterfall.

3 Improving Performance

3.1 Hyperparameter Tuning

At first step I tuned K from 10 to 30 and tuned L from 2 to 3, and using scales $[1\ 2\ 4]$ the accuracy increased from 46.75% to 56%.

At second step I changed L to 4 and changed *alpha* from 25 to 50, then I achieved accuracy 59.25%.

With *alpha* increased to 75 and all other parameters remain the same, the accuracy dropped to 58.5%.

With $alpha = 50$, $K = 100$, $L = 4$, scales $= [1\ 2\ 4\ 8]$, I achieved accuracy 60%.