Computer Vision Assignment Ten @ ETH Zurich

Image Categorization

Ossama Ahmed 18-936-872

December 21, 2018

1 Local Feature Extraction

At first 100 points are chosen in the image that we are trying to extract local features from. The 100 points were chosen to be internal to the image neglecting a border of size 8 as instructed. Then HIG feature desriptor was created out of these grid points by taking a window of 4x4 cells around them. The histogram of oriented gradients (HOG) is a feature descriptor used in computer vision and image processing for the purpose of object detection. The technique counts occurrences of gradient orientation in localized portions of an image. Also, the image patch was saved for every descriptor to visualize the codebook created later on.

2 Codebook Construction

After creating 100 feature descriptor for every image in our dataset. we cluster them in k cluster that the center of each cluster will represent a codeword in out codebook. After visualization of the codebook we get the following representation.

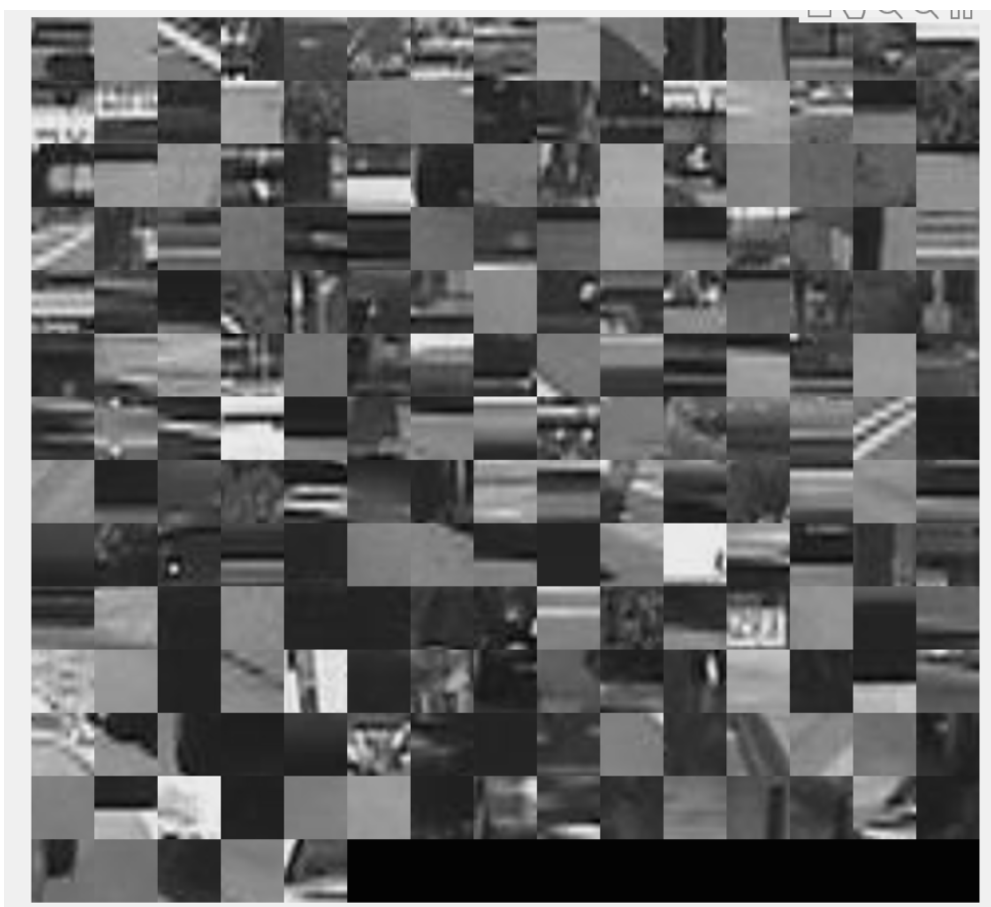

Figure 1: The codebook result.

3 Bag of words Image representation

After training phase and getting the codebook. A histogram is needed to represent each image in our testing dataset. Each feature extracted from the testing image will be assigned to a one of the words in a our codebook and then a histogram of length of our codebook will be used to represent this image.

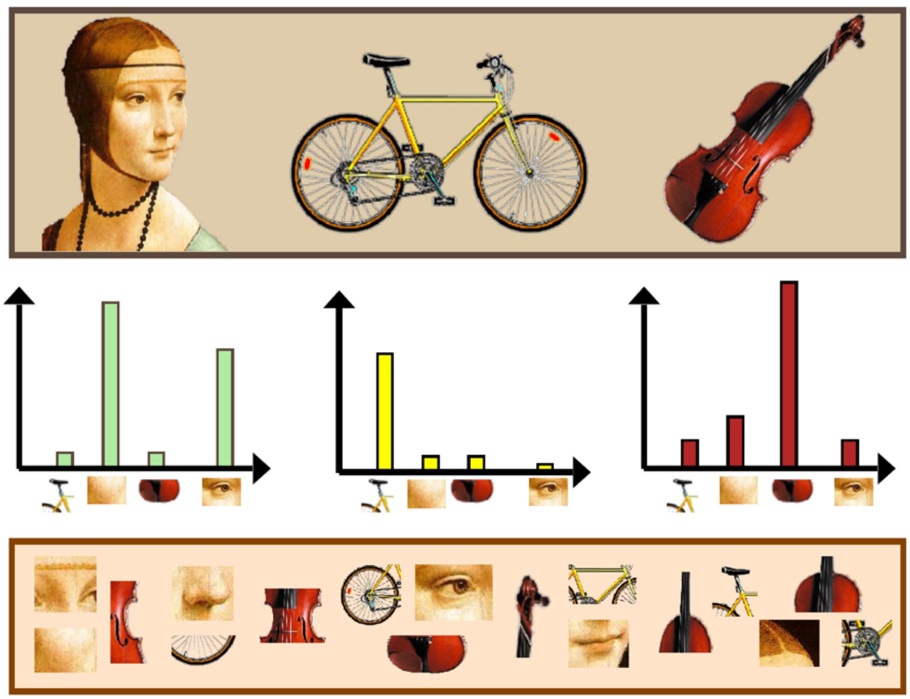

Figure 2: Bag of words image representation.

4 Nearest Neighbor Classification

For the nearest neighbor classifier, we represent each image in our trianing dataset by a histogram of words and then during testing phase we find the nearest neighbor to the histogram of the testing image. The method has high complexity since for every testing image, we iterate over all the histograms of the training images to get the closest neigbor. I used here euclidean distance to define the metric of judging how two representations are close to each other.

5 Bayesian Classification

For bayesian classification we create a normal probability density function out of the mean and standard deviation of all the bins in the histograms of the training dataset. We calculate the posterior using bayes rule $P(Car|hist) = P(Hist|Car) * P(Car)/P(hist|Car) * P(Car) + P(hist|!Car) * P(!Car)$

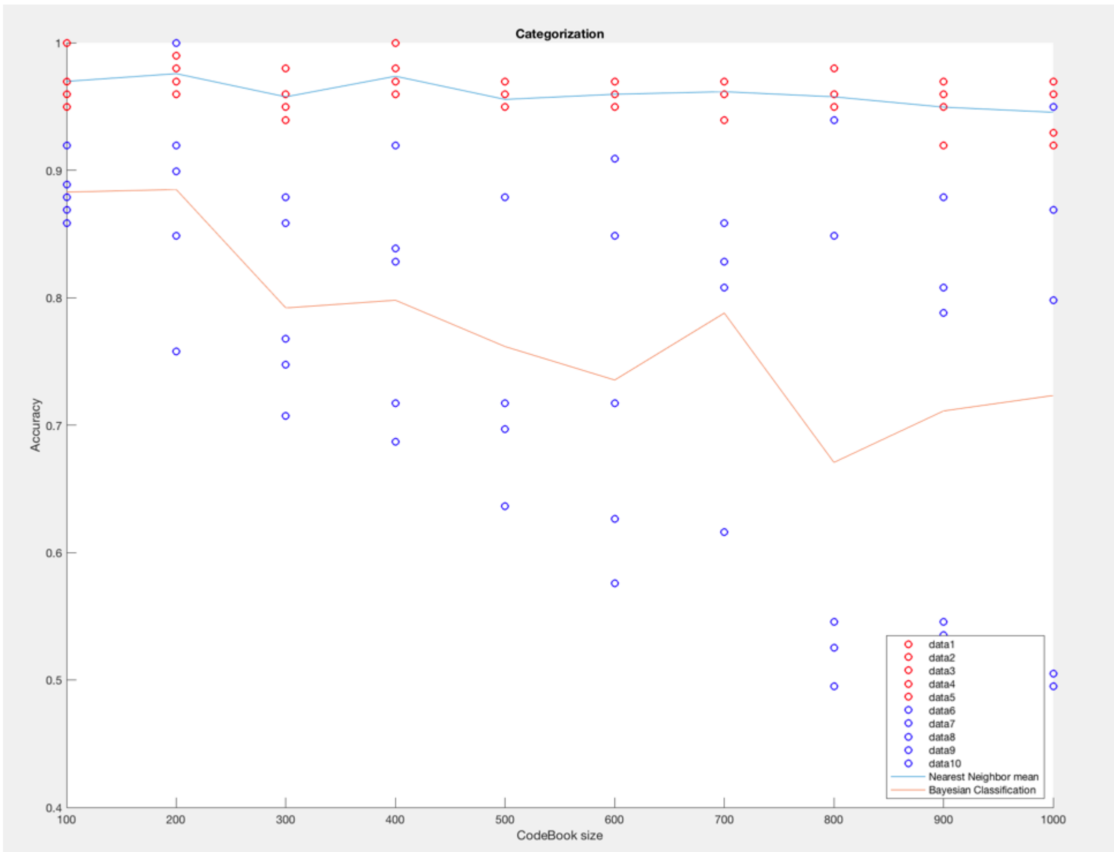

Figure 3: Classification methods comparison.

Bayesian Classification would decrease with the size of the codebook as we can see in the plot but also it has a very high variance. Given more training data, bayesian classification is expected to perform better than nearest neighbor. Nearest neighbor has a very low variance compared to the bayesian classifier but it has high computational complexity. With larger codebook, each codeword won't be representing an essential feature so the histograms will be really sparse which will make comparing features harder and less accurate.

6 Bonus

I ran the same code and classifier on the dogs vs cats dataset and I got the following results. Dogs were labelled as the positive images that the code book was created from. Note: these two datasets have very similar features obviously, so the accuracy was way lower then car vs no car dataset as shown below.

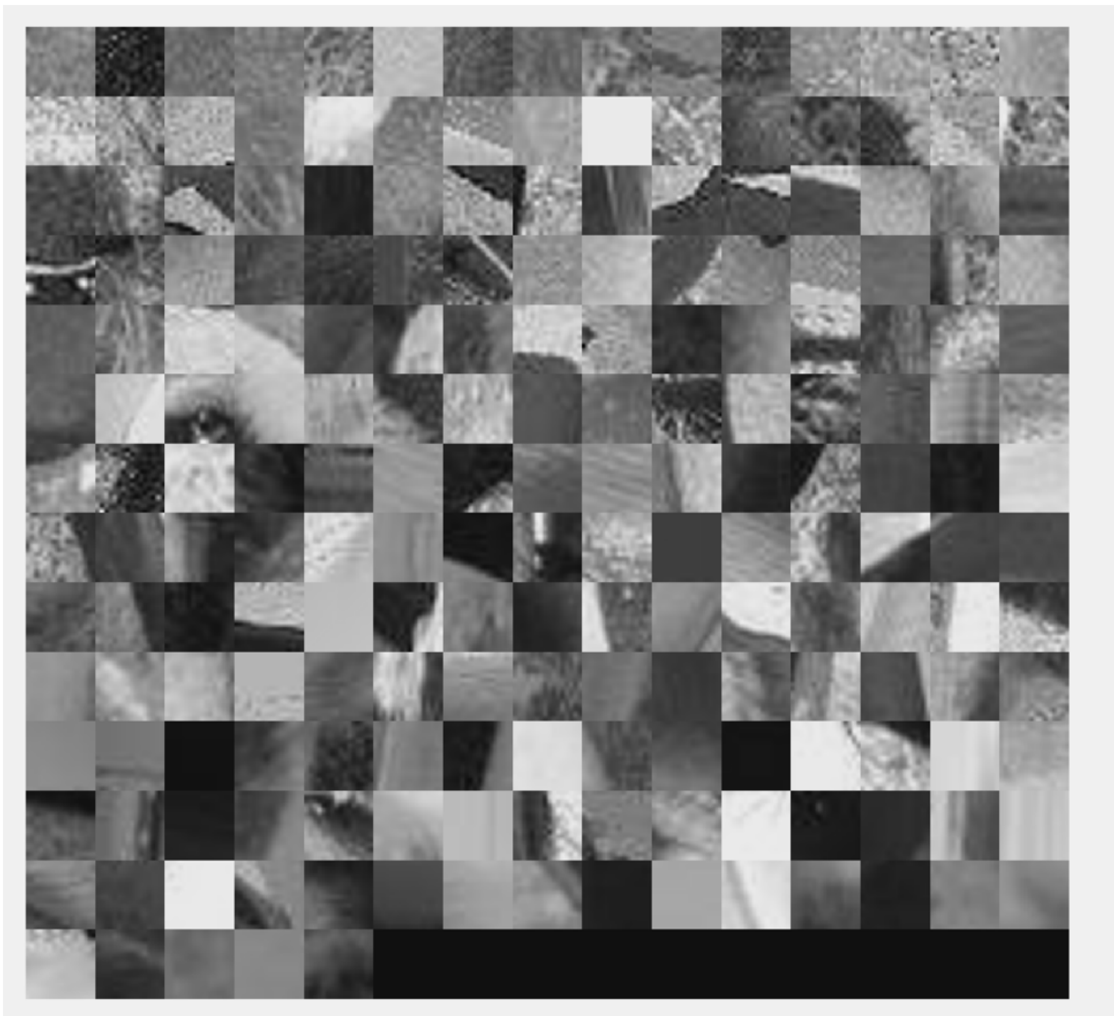

Figure 4: Cats code book result.

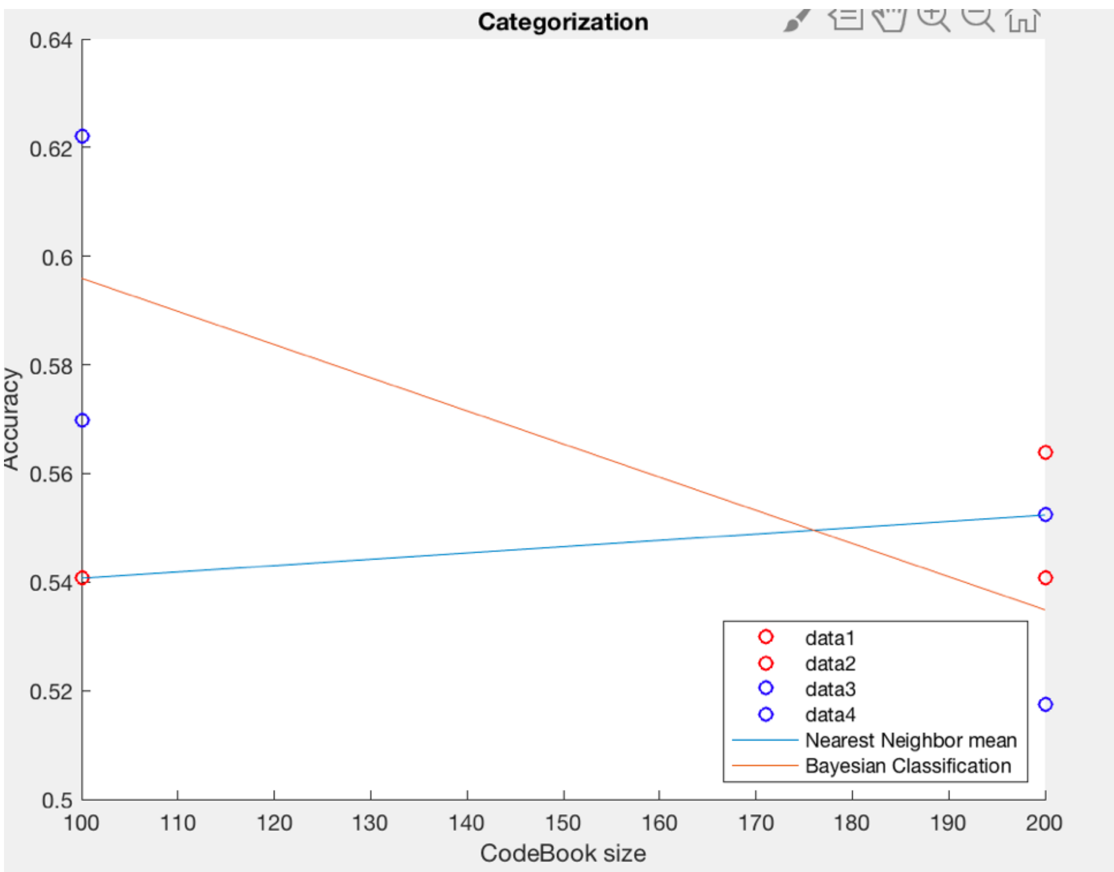

Figure 5: Cats code book result.