Computer Vision
ETH Zürich
Fall 2020

Lab Assignment 10 - Image Categorization

Mokhles Bouzaien

1 Local Feature Extraction

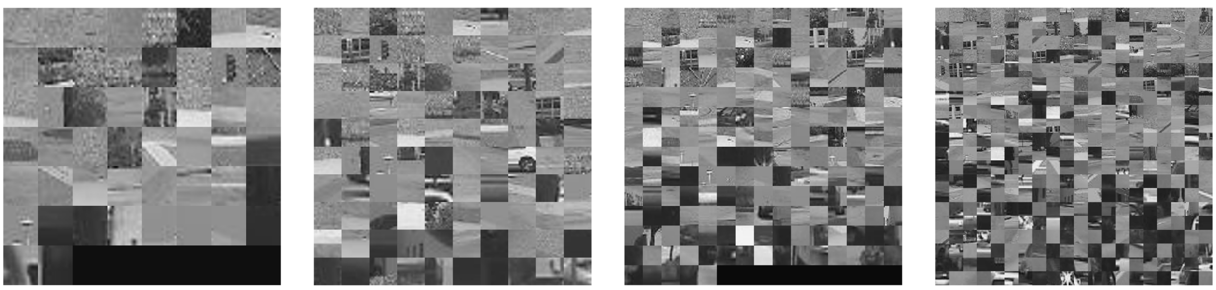
Feature Detection I started by creating a 10×10 grid using the `grid_points` leaving a border of 8 pixels in each dimension. The output of this function is an $N \times 2$ matrix containing the coordinates of the detected features where $N = \texttt{nPointsX} \times \texttt{nPointsY}$.

Feature Description A local descriptor is created and associated to each extracted feature using the `descriptors_hog` function. A histogram of oriented gradient is calculated for each cell of size `cellWidth` $\times$ `cellHeight`, then all histograms are concatenated to form 128-dimension vector.

2 Codebook Construction

Create Codebook Once the descriptors are extracted from all images, they are classified with K-means into k classes. The predefined MATLAB `kmeans` function was used at this step.

Visualize Codebook The visualization of codebooks for different values of k are shown in figure 1.



(a) $k = 50$ (b) $k = 100$ (c) $k = 200$ (d) $k = 400$

Figure 1: Codebooks for different values of k.

3 Bag-of-words Image Representation

Bag-of-Words Histogram After obtaining the clusters' centers, we can apply bag-of-words to each image. This consists in matching every feature of the image to the closest cluster (lowest Euclidean distance), and then calculate the histogram of matches for this images.

4 Classification

To classify unseen images, 2 different classifiers are used.

Nearest Neighbor Classification The first one is the nearest neighbor classifier. It consists of comparing the bag-of-words histogram of the test image to the training histograms and then assign it to the nearest one using the `knnsearch` function. i.e., the label of the test image is the same as the "nearest image" label.

Bayesian Classification The second one is the Bayesian classifier. This is a probabilistic classifier based on the Bayes rule.

$$P(C|h) = \frac{P(h|C)P(C)}{P(h)}$$

where C and h are the class (presence or not of car) and the histogram of the image. $P(C)$ is the prior of each class (fixed to 0.5). Since $P(h)$ is common for both classes, it's sufficient to compare $P(h|C_1)$ and $P(h|C_2)$.

Results After executing both classifiers for different values of $k \in \{10, 50, 100, 200, 400\}$, I got the results of the table 1. We can notice that

- both classifiers have comparable results with a slightly higher performance on average of the Bayesian classifier.
- the NN classifier accuracy decreases for higher values of k. This is due to the fact that given a high number of features, it's more likely to have similarities between them, so it's more difficult to compare with new feature.

The NN classifier is slower especially for high values of k because every feature is compared to all other existing ones.

k	10	50	100	200	400
NN classifier	93.94%	89.90%	89.90%	88.89%	81.82%
Bayesian classifier	88.89%	93.94%	89.90%	85.86%	87.88%

Table 1: Table of performance of each classifier.