

Laboratory 6 - Textons and classifiers

Ricardo MENDOZA-LEON

Universidad de los Andes

Bogotá Colombia

ra.mendoza35@uniandes.edu.co

Abstract

Image content classification is an important task with several applications in computer vision problems. Results of several experiments in multiclass image classification using textons to classify textured images are reported. Two different classifiers: nearest-neighborhood (NN) and random forest, were trained with k-means quantized histograms of texton responses on 750 train images, and tested in a dataset of 250 images. Results showed better scores for random forest classifiers, specially when the number of trees was increased. However, NN classifier performance was comparable, confirming that a suitable representation for the problem at hand is key.

1. Dataset

The dataset used, obtained from Ponce Group, comprised 25 natural (grayscale) texture categories, each one with 30 images for training and 10 for testing. Texture orientation accounted for most of the intraclass variability observed in this dataset.

2. Methods

The classification pipeline comprised the following steps.

A) Apply a filter bank with 8 oriented second-derivative even-filters and 8 oriented odd-filters (Hilbert transform of the later) in two scales, for a total of 32 dimensions to represent each pixel.

B) Build a texton dictionary to quantize pixel representations using k-means clustering. To build this dictionary a fixed number of pixels in the filter bank response space are sampled in each image. Tests were performed using 50 and 5000 samples (sp) per image and also different dictionary sizes (k) were explored using $k = 50$ and 100.

C) Create a histogram of texton representation for image. This is accomplished by applying the filter bank to the image, labeling each pixel in the response space by assign-

ing the label of the closest texton (centroid) in the dictionary and finally, creating the texton histogram by aggregating the labeled image pixels.

D) Train a classifier, using as examples 750 histogram and labels corresponding to the train set images. Two kinds of classifiers were tested. The first, was a Nearest-Neighborhood (NN) classifier, whose model is given directly by the train histograms, without the need for an explicit training procedure. Classification using NN is carried-out by looking for the closest model-histogram to the histogram to classify. Given that images are represented as histograms, the histogram interception distance was used as a measure of proximity between histogram pairs. The second classifier used was the Random Forest classifier, which trains a number of binary-tree structures and assigns simple separation criteria at each node to maximize the information gain on the input example distribution. Many parameters such as tree-depth and separation criteria space, among others, define the behavior of these models. For the experiments, models with 20 and 80 random trees were trained.

E) Finally, each image in the test set is represented as a texton histogram following the steps in C, and its class predicted by the classifier. Results are collected in a confusion matrix, and after normalization, the global performance is computed as the mean value of the diagonal terms in the matrix.

3. Results and discussion

Figure 1 summarizes the global performance for all the experiments over the two dictionary sizes k . First it can be observed an performance increase related with a bigger dictionary size. Also, increasing the sampling (Sample) and, in the case of the random forest classifier (RF) increasing the number of trees, also improved the performance. Figures for confusion matrices are shown in 2 - 3 and 4 - 5 for nearest neighborhood with $sp=50$ and $sp=5000$ respectively, 6 - 7 and 8 - 9 for random forest classifier with 20 trees and $sp=50$ and $sp=5000$ respectively, and 10 - 11 and 12 - 13 for random forest classifier with 80 trees and $sp=50$ and $sp=5000$ respectively.

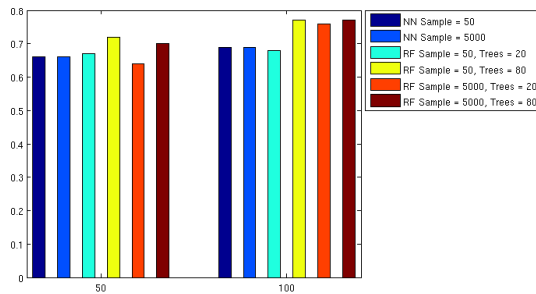


Figure 1. summary results for all tests

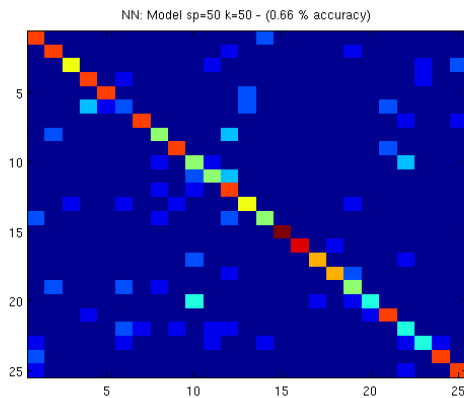


Figure 2.

The random forest classifier was consistently the best classifier in the tests. Begin a richer model, it can define separation regions that generalize better the distribution of representation histograms in the train set. However, the difference between the two classifiers was smaller than expected (best performance was NN=0.69 vs RF=0.77).

An important difference between the classifiers was training and evaluation time. In the case of nearest neighborhood, the train time is 0 (given the features are already extracted). However, the difference for classification time is wide; NN has $O(n^2)$ order and RF is $O(\log(n))$.

Finally, confusion matrices showed that class 23 (knot) had the lowest classification scores in all the experiments, because it is often confused with classes 22 (fur) and 24 (corduroy) that look very close. Given that the intraclass variability is small, this classification task can be considered as easy, however, its small size increased the challenge.

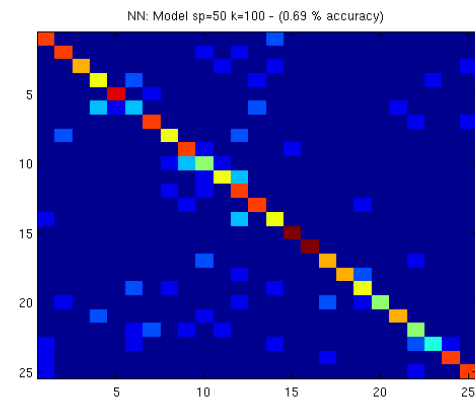


Figure 3.

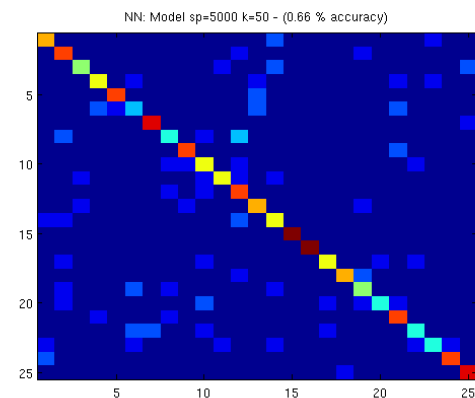


Figure 4.

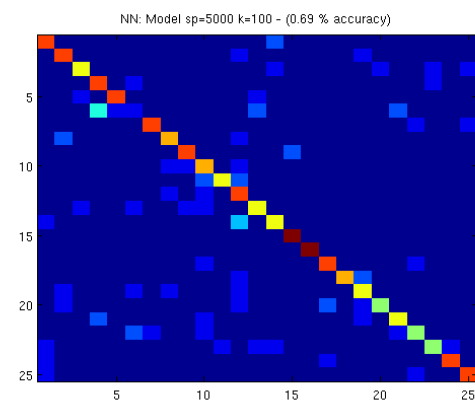


Figure 5.

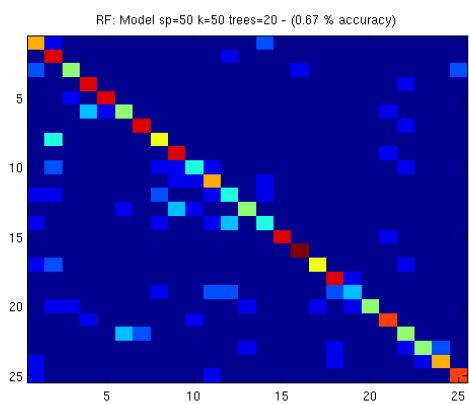


Figure 6.

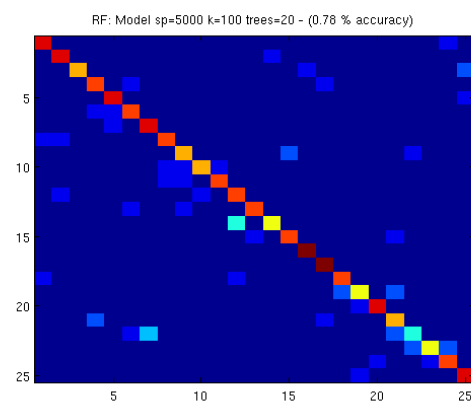


Figure 9.

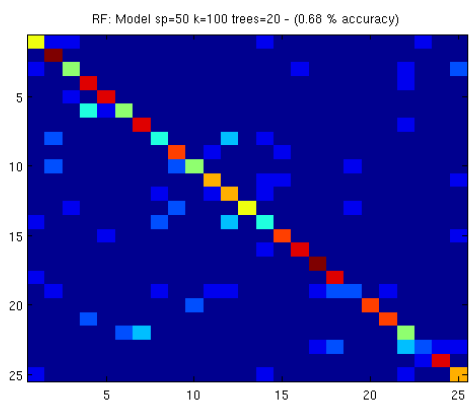


Figure 7.

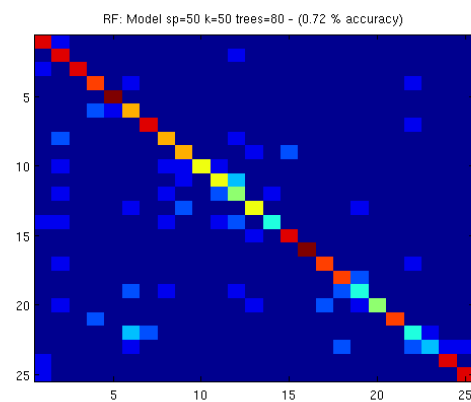


Figure 10.

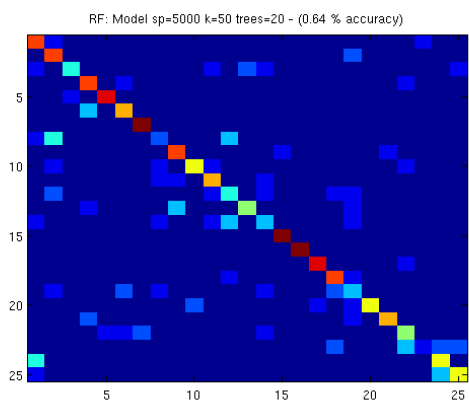


Figure 8.

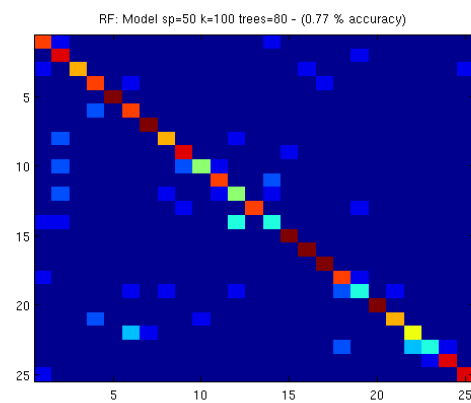


Figure 11.

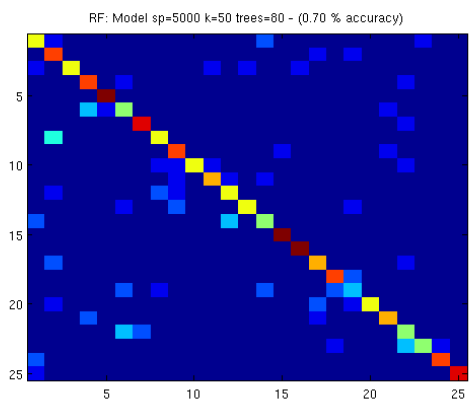


Figure 12.

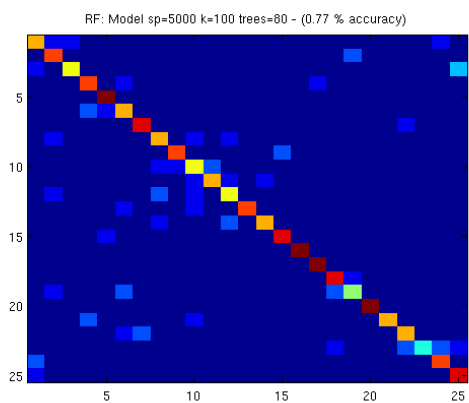


Figure 13.