Laboratory exercise 5: Textons

Santiago Enrique Cortés
Bárbara & Frick
Av. Calle 82 # 10 18A - 12, Bogotá, Colombia
santiago.cortes@barbara.net.co

Juan Felipe Cerón Uribe
Universidad de los Andes
Cra 1 # 18A - 12, Bogotá, Colombia
jf.ceron10@uniandes.edu.co

Abstract

In this paper we present the results obtained in fifth lab of the 2019 IBIO4490 (Computer Vision) course. In this lab we use images' texton representations as features for Knn and random forest classifiers. We test this approach on the cyfar-10 dataset, which consists of simple pictures of objects belonging to 10 different categories.

1.. Introduction

Textons, in principle, represent units of texture. Furthermore, an object's textures are usually far more informative than its colors or light intensity. Thus, for an image, it makes sense to consider the textures surrounding each pixel, rather than its colors or gray-scale intensity, when facing a classification problem.

2.. Dataset

The images used were the cyfar-10 data set. In this data base there are 10 categories, the training set consists in 5000 images per class and a training set of 1000 images for class. We only considered gray-scale versions of the images since it is safe to assume that the textures in them will be retained by this simpler representation.

3.. Methodolgy

3.1.. Image representation by textons

The first challenge we faced during this exercise was to achieve the representation of an image by the texture surrounding each pixel. In order to define these textures, we created a *filter bank*; a set of filters meant to uncover the basic constituents of different textures. Examples of such filters are directional derivative filters and filters shaped as black or white spots. Specifically, we used directional derivative filters at 16 uniformly distributed angles.

Each filter in the size 32 filter bank was applied to each image in the training set, obtaining, for each pixel in each image, a 32 dimensional vector of its neighbourhood's response to the basic texture elements. These vectors were then clustered via k-means, with $k = 16 * 8$. We claim that the resulting centroids are good representatives of the different textures in the images because the basic-texture vectors of other pixels gather around them. Finally, we can represent a gray-scale image by replacing each pixel with the label of its closest centroid when represented as the aforementioned 32-dimensional basic-texture vector.

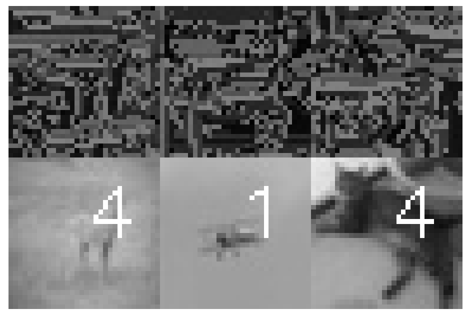

Figura 1. An example of three test sample images with their respective texton representation.

3.2.. Classification algorithms

Using the images' texton[1] representations, we implemented K-nearest neighbours and a random forest to attempt to classify the images in the 10 categories available in the dataset. For the K-nn algorithms we fixed $K =$. For the random forest we allowed 1000 trees to grow up to single-element leaves, with each split being able to choose from $\sqrt{1024} = 32$ features; a standart choice of parameters for this classifier.

We trained each model with 500 images (50 from each category) and tested the algorithms' performance on all 5000 of the available test images.

4.. Results

Figures 4 and 4 present the confusion matrices on the test set for each algorithm. Even though the training set was balanced, we observe that categories 1 and 3 were quite over

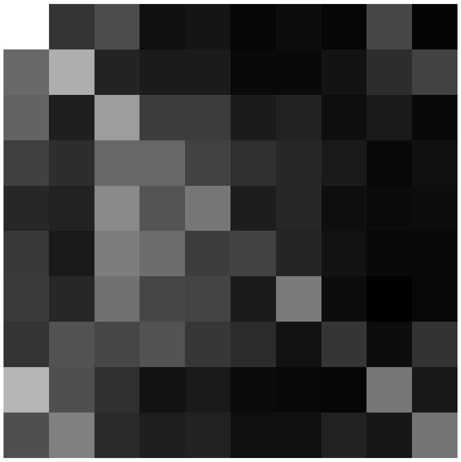

Figura 2. Confusion matrix on the test set for the K-nn classifier.

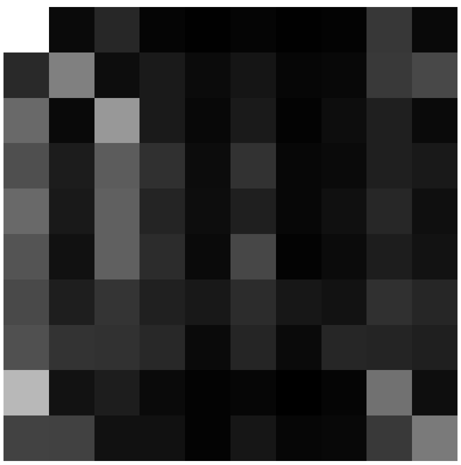

Figura 3. Confusion matrix on the test set for the random forest classifier.

represented in the classifications made by both algorithms. Category 8, on the other hand, received a small number of classifications in both cases. This might just be a coincidence, or it might tell us something about the distribution of the textures present in each category. Suppose, for example, that categories 1 and 3 have textures that are very common in the whole dataset, whereas category 8 has very unique, hard-to-tell textures. This could result in a higher probability of misclassifying any image in categories 1 and 3, and a reduced probability of being classified in category 8. This would even be the case for true category 8 images because the textures are not successfully described by the set of discovered textons.

Referencias

[1] J. Malik, S. Belongie, J. Shi, and T. Leung. Textons, contours and regions: Cue integration in image segmentation. In *Proceedings of the Seventh IEEE International Conference on Computer Vision*, volume 2, pages 918–925. IEEE, 1999. 1