Texture Analysis Classification

Luisa F. Borbn R.
Universidad de los Andes
Departamento de Ingeniera Biomdica
lf.borbon@uniandes.edu.co

Juan Camilo Martinez
Universidad de los Andes
Departamento de Ingeniera Industrial
jc.martinez10@uniandes.edu.co

Abstract

Texture is fundamental in human perception as it gives valuable information to recognize an image. In this sense, a texture analysis technique was applied on the CIFAR-10 database, in order to classify 10000 images into 10 different categories. Textons, the basic unit of texture perception, were computed

Keywords: textons, classification, histogram

1. Introduction

Texture is an important visual feature, that is a basic component of the human perception and its necessary to recognize an scene or image. In this sense, for computer scientists its fundamental to understand the texture in order to emulate human visuals and solve image problems using advanced computational methods. The first idea of texture analysis was introduced in 1981 by Julesz *et al*. He contributed to this field by declaring *textons*, that are the "putative units of pre-attentive human texture perception"[4], which provide spatial information or local patterns that make up an image.

Texture analysis has different applications, including classification, segmentation (finding texture boundaries), automated inspection, medical image analysis and remote sensing. Specifically, texture classification is done by assuming each class will be associated to an specific set or arrange of textons [3]. Generally, the two main classification methods are: supervised, in which the computer is provided with examples of each texture class as a training set to learn, or unsupervised, that doesn't require prior training data [5].

In this work, the problem of classifying the images of the CIFAR-10 dataset was - using a texture analysis technique. The methodology used is described on section 2, while results and discussion can be found on sections 3 and 4, respectively.

2. Materials and methods

2.1. Database and Images

As we mentioned before, the dataset used to apply the analysis texture classification method was CIFAR-10 by Alex Krizhevsky [1]. This dataset consists of 60.000 colour images of size 32x32, with 6000 images for each of the 10 categories: airplane, automobile, bird, cat, deer, dog, frog, horse, ship and truck. The images are split in five training batches and one test batch, each containing 1000 images. For this project, 10000 of the 50000 training images and 10000 test images were used. [2]

First, considering the analysis was only going to be made on the different textures of the images and not on other features like the color, we decided to transform the images into gray scale. This was also done in order to speed up the computer processing time and make better use of the resources. Second, we realized that the images and labels were not sorted and this would complicate the analysis, so we ordered them to obtain balanced data without running all the batch.

2.2. Texture Analysis

A texton dictionary was generated from a subset of the training images, balanced across all categories. The images were convolved and clustered according to similitude using a K-Means algorithm. Then texton histograms (the frequency with which each texton occurs in a class) were computed for both the training and the test images. Classifiers were subsequently used to label each image according to the class of the model that most closely resembles the image. The texton representation of an image tells us the most common repeating patterns that are visible on it.

In order to obtain the texture of the images, we used specific functions adapted from provided matlab scripts to the python programming language. The images were filtered using a filter bank of 32 instances with different orientations and configurations. To generate the texton dictionary we used 1000 sample images, 100 of each category. The selection of the neighbour number was done varying it between 10, 50, 100, 150, 200 and 500 neighbours and selecting the one that gave the best accuracy score, in this case 500.

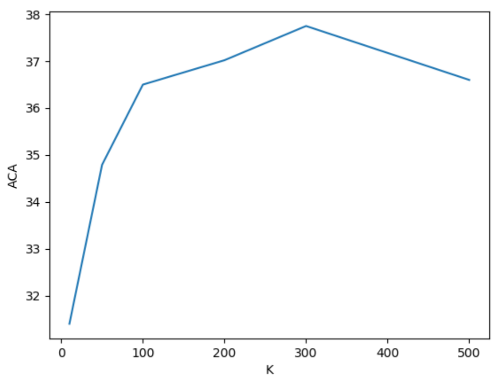

Figure 1. K vs ACA using default classifier parameters

2.3. Classification

Once we had the different texton histograms, we proceeded to train two classifiers: nearest neighbors (NN) and Random Forest (RndF). The hyperparameters considered were number of neighbors for NN, number of estimators and ratio of number of features to consider for RndF.

3. Results

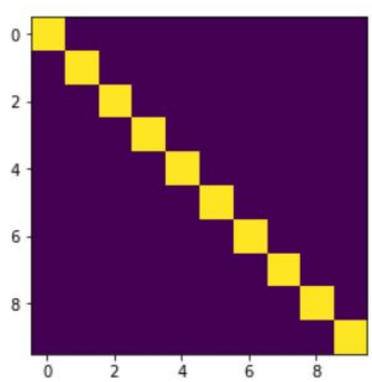

Figure 2. Train Result Confusion Matrix

The confusion matrix presents not only how many images were correctly classified, it also provides insight into the behavior of the incorrectly classified classes, including but not being limited to which classes were considered to be similar enough that the classifier encountered trouble discriminating among them. Average classification accuracy (ACA) was used as a shorthand classification metric to

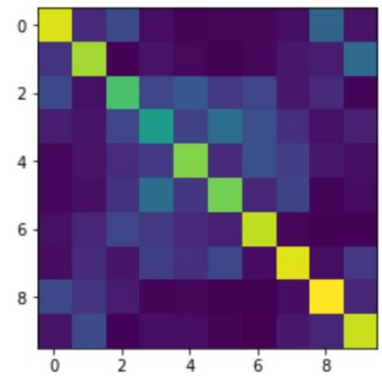

Figure 3. Test Result Confusion Matrix

quickly evaluate the effects of hyperparameter variations in the classifier.

The creation of the texton dictionary takes roughly 6 hours. The texture dictionary is created through K-Means clustering, which increases in complexity with the number of input samples. Moreover, inputs contain thousands of 32x32 arrays, obtained from an sizable matrix that needs to be pre-processed by transposing it.

Classifiers are fitted to a training batch in a relatively short span. In less than 20 minutes the generated histograms are used to train the models and these in turn take very little time to be applied to the test images.

The categories that present the most confusion are 2 and 5 ("Cat" and "Dog"), but especially the animal classes and the reason behind this can be inferred from the limitations of the method: different classes can present similar texture patterns (eyes, fur, feathers, etc.).

As mentioned before, similar texture patterns can be found in several classes for certain classification problems. In some cases, segmentation can provide the missing information to differentiate between two similarly textured images.

To lower the complexity of the texton dictionary, we converted the RGB images to a single grayscale channel. Defining the dictionary by incorporating every channel could improve the differentiation between classes like "Truck" and "Automobile" where textons could be similar in grayscale but color histograms might be significantly different.

4. Conclusions

A classifier based on texture analysis and random forest was implemented to classify 10000 test images according to 10 categories. Tuning of hyperparameters allowed us to increase the average classification accuracy from an initial 0.103 to a much more accurate 0.46. This result is considerably bellow the "state of the art" achieved for this problem (0.89), but we consider it can theoretically be improved by an even larger texton dictionary including color channels and finer adjustment of hyperparameters.

References

[1] A. Krizhevsky. Learning multiple layers of features from tiny images. [online] https://www.cs.toronto.edu/ kriz/learning-features-2009-TR.pdf.

[2] A. Krizhevsky and G. Hinton. The cifar-10 dataset. [online] https://www.cs.toronto.edu/ kriz/cifar.html.

[3] MathWorks. Texture analysis. [online] https://la.mathworks.com/help/images/texture-analysis-1.html?lang=en.

[4] D. Zhou. Julesz conjecture. [online] https://www.cs.auckland.ac.nz/ georgy/research/texture/thesis-html/node6.html.

[5] D. Zhou. Texture analysis and its applications. [online] https://www.cs.auckland.ac.nz/ georgy/research/texture/thesis-html/node7.htmlSECTION00431000000000000000.