Multiclass classification using random forest and textons as classifiers

Daniel Viteri Noguera
Universidad de los Andes
d.viteri@uniandes.edu.co

Juan Felipe Nieto
Universidad de lo Andes
jf.nieto@uniandes.edu.co

1. Introduction

Textons are a basic structure for both images and video [1]. They can be used to represent texture or geometry. In order to use textons for image analysis, it is necessary to define a dictionary containing the different textons that will be applied. Once the dictionary is defined, it is necessary to apply them to the different images, this can be made in the same way a filter is used, therefore returning a map representing the coincidence between images and the textons. By having the map, it is possible to analyze the data, by either clustering, or classification methods such as k-means or random forest [1]. Given the variety of textons than can be designed, it is plausible to use them for different applications such as semantic segmentation or classification.
n For the classification problem, it is necessary to create a histogram representation for the image after applying the textons, then fitting a classificatory to organize future data in the classes present in the annotations. In the clustering case, the difference is the object to analyze. Instate of analyzing the complete image, the analysis is made for every pixel in it, giving every pixel a label that represents a category.
n The current paper is focused in the classification problem using the CIFAR-10 data base using textons and different hyperparameters for the classifier, to achieve the best possible outcome.

2. Dataset

As mentioned before, we are using the CIFAR-10 database for image classification. This data set consist in 60,000 RGB images with 32x32 size. For these data there are 10 classes, each with 5,000 images and 10,000 as test group. By downloading the data set, it can be realised tha it is a tar compressed file with 5 different batches, each one with balanced train and test division and enabled for c, MATLAB, and python.

Methods In order to achieve the results different procedures are required, each one was performed separately and each one is presented in the current subsections.

2.1. Image Reading

For loading the images the pickle library was used in order to manage the images, once the images are load into the work space it was possible to change them form RGB to gray scale, then in order to organize the data set and ussing a while cycle the dataset was balanced allowing to create a new file containing the images and anotations to facilitate every further process.

2.2. Generation of Textons

to generate the textons is necesary to create a filter bank and computting it with the images therefore generating a respose for a given filter, and by ussing k-means in dis case is possible to generate a cluster of the given data allowing the future classification. In this case, this procedure was made two times for different k values, 16 and 128 in the spectation of showing significative differences between a big diffetence unthe number of clusters.

2.3. Classification

Once the textons are generated is necessary to apply them to the images in order to get the response to the dictionary then we get the histogram for the image by defining a struct with field the same size as number of bins ang assigning everyone to the corresponding site in the struct. Then is necessary to define the forest that will be used in this case, we are varying the depth and number of trees between 5 and 10, and 10 to 15 respectively.
After the definition a cross-validation is used in order to determine the best possible model, once the classifier is defined, the data are fit into the model to generate the classifier.

2.4. Validation

For validation, the textons are assigned to the test group and the category is predicted with the model used, and by using the annotations is possible to calculate the confusion matrix and the ACA index.

Table 1. cross-validation score for different K and and hyperparrameters conbinations

2*		**Hyper Parameters Depth/Trees**			
		5-10	**5-15**	**10-10**	**10-15**
2***number of K**	**16**	0,15	0,155	0,147	0,153
	128	0,163	0,162	0,141	0,14

3. Results and Discussion

By using the cross-validation, we could show that the pest possible outcome comes from 128 Clusters and 5 depth and 10 trees for the hyperpatameters in the classificator, by observing this relationship we can observe there is no relationship between depth and number os trees as it is shown in table 1, this can be explained due to the randomness of the methos it self.
after determined the best possible classificator and fitting the model into it, the confusion matrix is obtained, with the test data, then the ACA is calculated and it is posible to observe the reduction of the value compared to the cross-validation in the training group, this can be explained by the fact that the training group was used to create the texston dictionary therefore producing some perfect mattes in the algorithm, in the other hand, the test group does not have this perfect or near perfect data and so the metric value is reduced.
this las procidures failed during the last run of the code therefore was not possible to present the data supporting the spected resultd mentioned before.

Besides the problem mentioned before, is necessary to show different problems such as the image size, since the data base used consist only of 32x32 images some details are lost due to the resolution making more difficult to distinguish the different textures in the image and also having background noise due to the proximity between the interest section of the image and the background.
Finally as a possible solution to the noise problem it will be possible to add a prepossessing stage in which the images is segmented given only the interest section of the image, although it seems as a perfect solution, this represent a difficulty given that not all images are the same therefore having different characteristics to compare.

4. Conclussions

In conclussin the data set is not the best due to the lack of resolution in the images, but the outcome can be easyly imporved by creating a pre prossesing algorithm reducing the noise therefore improving the cuality of the textonsmap and clasificator, as a positive aspect the small images have the advantage of reducing the required resources to create a classification method.
Random forest show the ability to clasify this database with a 0.16 acurracy wich given the information in disposition is not a bad ressult al though, the results cant be easyly recoverd due to the randomness of the method.

References

[1] Zhu, Song-Chun, et al. ”What are textons?.” International Journal of Computer Vision 62.1-2 (2005): 121-143. 1