

# Programing Assignment 2 : Image Recognition

Jeanne Hecquard

May 2021

## 1 Confusion matrices

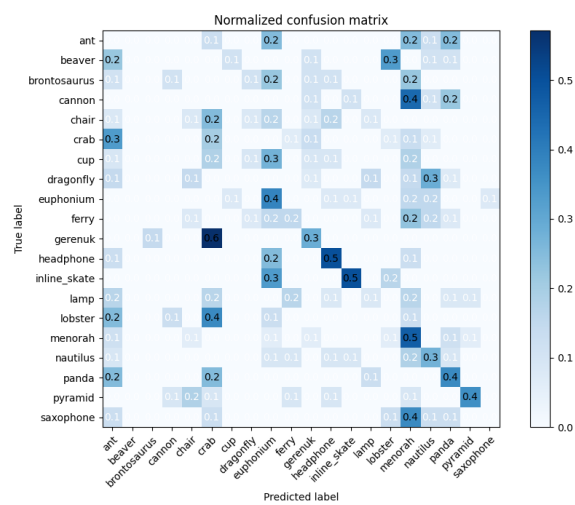


Figure 1: Confusion map of the Bag of Words method  
Average accuracy : 0.251

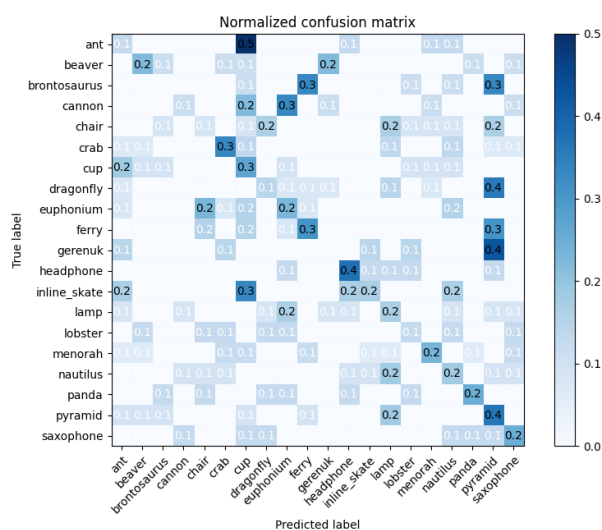


Figure 2: Confusion map of the Spatial Pyramid Matching method  
Average accuracy : 0.306

## 2 Bag of Visual Words versus Spatial Pyramid Matching

The Bags of Word array saves the number of descriptors with the most similarity of each cluster in our codebook by calculating their distances. Then we store the cluster id of the smallest distance and add 1 to the corresponding codebook cell. Therefore, the sum of each row of the Bag of Row matrix should be the number of features. Since the size of images are different, we need to normalize the bag of words matrix in the end.

We can also use pyramid representation which contains more spatial information. Instead of feeding only one image to the feature extraction function, we divide the image into smaller subimages and do the feature extraction on all of them. In the first level, we divide image into 4 parts, so the size of each image will be  $(0.5 \times \text{height}, 0.5 \times \text{width})$ . In the second level, we divide image into 16 parts, so the subimage size will be  $(0.25 \times \text{height}, 0.25 \times \text{width})$ . However, as the smaller the sub-image is the less feature points it has, we need to multiply their weights when accumulating the histograms together. This creates a balanced feature extraction.

As we can see in the figures 1 and 2, the Spatial Pyramid Matching method is more effective than only the Bag of Words one, as is it more precise. Indeed, the histograms of two images containing the same features will be the same, regardless of where those features are located in the images. The Spatial Pyramid Matching method allows us to discriminate based on the features' locations, which is very important for image classification.

## 3 Visualization example

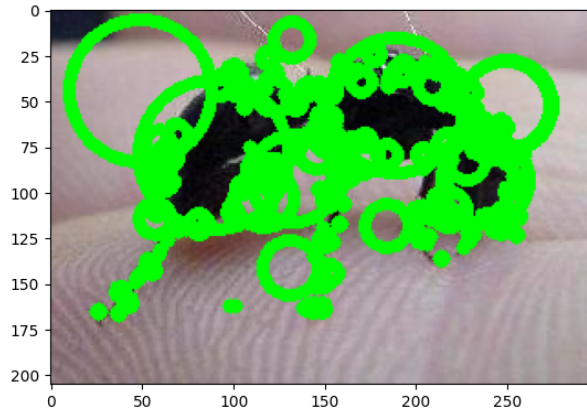


Figure 3: Visualization of the important visual words of a picture of an ant

## 4 Image retrieval

I did not manage to have a satisfactory image retrieval algorithm and have thus decided to not turn this part of my work in.