Bag of visual words coursework

Q2

4. For this part we iterate over each feature point (SIFT descriptor), computing the index of the closest dictionary words learned previously and then assigning correctly. This is reflected in code **lines 119 to 132.**

Q3

In order to complete this task, it was necessary to use the loaded features associated with each particular image, for each one of these many features we record the closest vector word as in 2 and then update the row for the histogram of the image appropriately. **Lines 163-175 contain our code, to illustrate the format of the output, careful consideration of the initialisation the BoW array is needed.**

Q4

As defined in slides the histogram intersection is computed by:

hinst(H1,H2) = $S_{1..k}$ Min(H1(K) , H2(K)), but since we are looking for a distance metric for KNN between histogram we suppose that we want histogram similarity which could be defined as:

1 – hinst = 1 – $S_{1..k}$ Min(H1(K), H2(K))

This has been reflected in line 74-90 in knnsearch.m, method has been changed to 2 and K has been set to 5 since we know the number of classes we're looking for. this has let to a total error **err_all** of **0.27 whereas using the l1 norm gives 0.38 classification error. The resuling confusion matrices are supplied below:**

It seems like the L1 norm is predicting airplane for car a lot of the time and keyboards instead of faces.. the histogram similarity metric is better across the board but seems to trip up and confuse faces for keyboards.

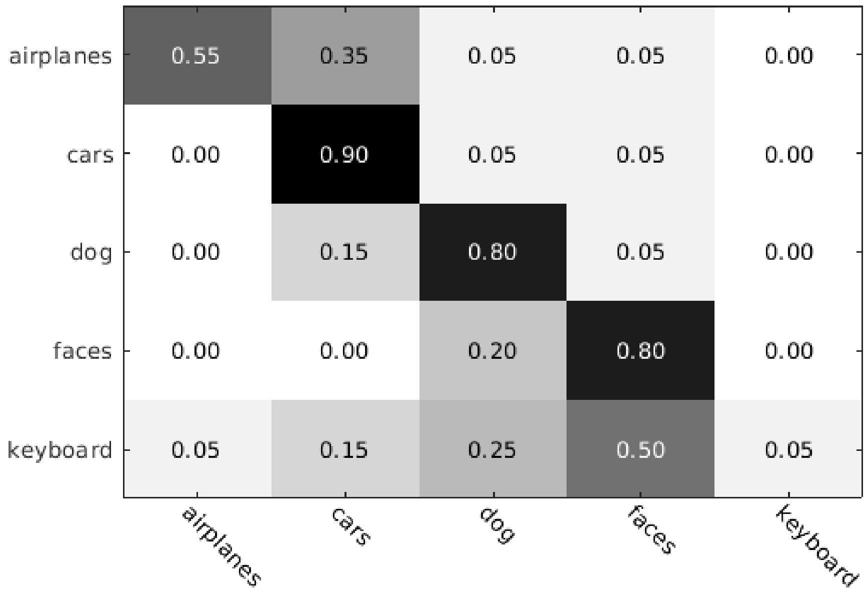

Illustration 1: Confusion matrix for KNN with l1 norm distance metric

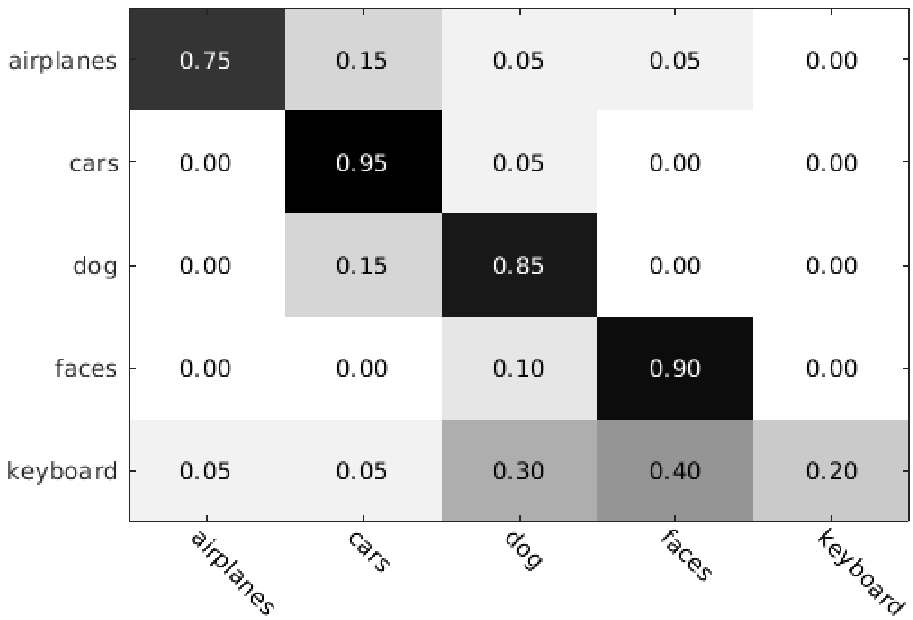

Illustration 2: confusion matrix for KNN with histogram similarity metric