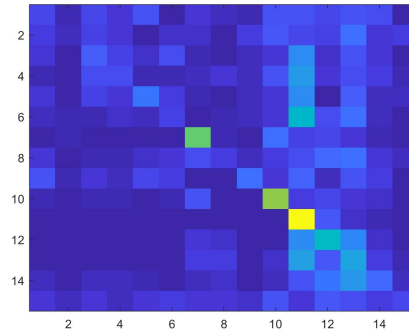[ALL SYMBOLS USED ARE SAME AS IN LECTURE SLIDES]

The task of Scene Recognition is done using three algorithms:

1) Resize the image to a fixed resolution, normalize and use knn to predict the label of the test images. For ClassifyKNN_Tiny, the accuracy is 23.8667%.


(a)

Figure 1: Confusion matrix for KNN and Tiny Image Representation with accuracy 23.8667%

2) Build a bag of visual dictionary using vl_sift and clustering them using into dic_size clusters using kmeans. Extract features from each training image and compute their bag-of-words. Create a model using a. KNN and b. SVM and calculate the accuracy.

a. ClassifyKNN_BoW: KNN model is used for the prediction. The vocabulary is computed using vl_dsift using parameters Step=5 and dic_size=50. The optimum results is given by the fast version of the function, accuracy is 51.2% for k=9.
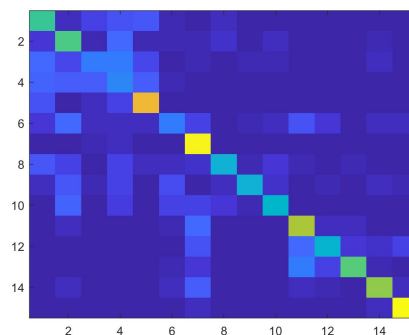

(a)

Figure 2: Confusion matrix for KNN and BoW representation using dic_size=50 and vl_dsift step=5 (fast) and accuracy 51.2%

b. ClassifySVM_BoW: SVM model is used for the prediction. The vocabulary is computed

Homework 3

using vl_dsift using parameters Step=5. The optimum results is given by the fast version of the function, the accuracy is 62.800% for dic_size=200 and lamba=0.00005 (the maxIter paramter is by default 10/lambda).
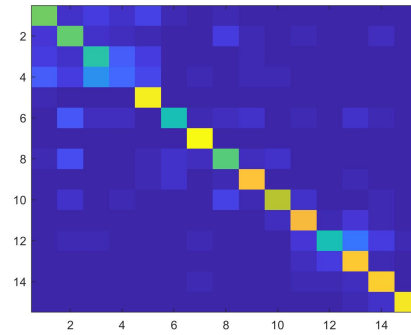
(a)

Figure 3: Confusion matrix for SVM using BoW representation using vl_dsift step=5 dic_size=200, lambda=0.00005 and accuracy=62.800%