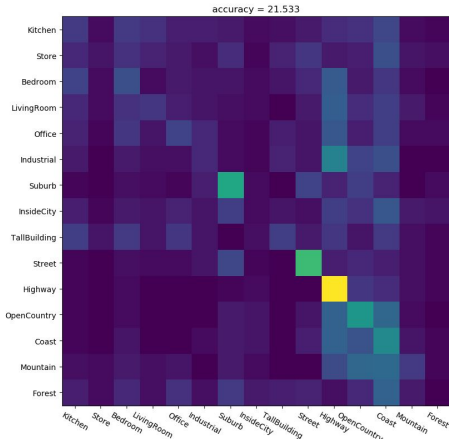
Shriya Rai | rai00016@umn.edu | 5462955

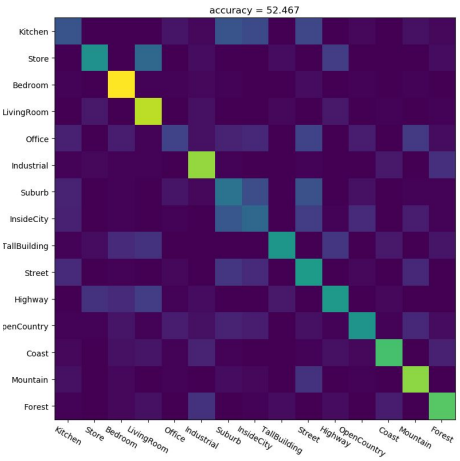
CSCI 5561: Assignment 3 - Scene Recognition

1.



Above is the output for predicting class using kNN in combination with tiny images. Tiny image function resizes the image into a 256 (16x16 flatten) data and normalized this resized image, a pure clean image feature will be collected as opposed to the complicated features for the original image. This is given as input to the kNN model where its target/response i.e. is the class here for test samples is predicted utilising its ‘k’ closest to the members. Model with k = 3 yields an accuracy achieved 21% and k=2 achieves higher accuracy of 22%.

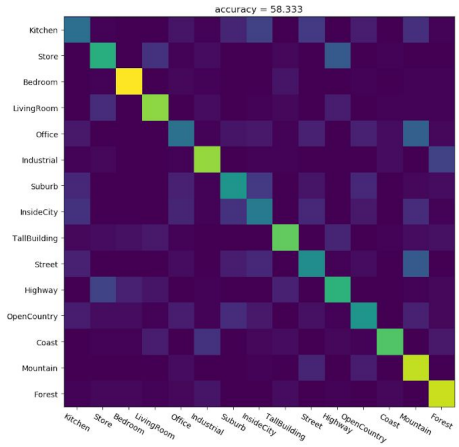
2.



This is the output for predicting class using kNN in combination with bag of word or in this case quantized sift features actually. Before representing images as bag of feature histograms, a vocabulary of visual words of created. The vocabulary is created by sampling many local features from the training samples and then clustering them with k means. In order to create the histograms, dense sift descriptors are computed. This is followed by counting of how many SIFT

descriptors fall into each cluster(bin of histogram) in the vocabulary. This is done by finding the nearest neighbor k means centroid for every SIFT feature. This is then given as input features to the kNN model. The value for ‘k’ here is set to 7 and the stride and size for dense sift computation is set to 10 while the vocabulary size is set to 50. This model yields an accuracy of 52.467%.

3.



This is the output for predicting class using SVM in combination with bag of words. The bag of words is computed in exactly the same manner as stated above with stride and size for dense sift computation set to 10 and vocabulary size set to 200. SVM model is implemented by training and combining 15 one vs all Linear SVM classifiers with lambda value set to 0.7. The training data here is the bag of words representation for the original training images. Accuracy achieved with this configuration is 58.333%. SVM proves to be a better classifier here when compared to kNN for the same set of training features/ representation.