1. **Introduction**

In this homework, we are given 20 different classes and each of them having 20 to 50 instances. For each class, 20 images are randomly selected. Using descriptors, we create a codebook with varying codebook size based on the whole training data and encode each image in the training set according to the codebook. The encoding is histogram of codebook assignments of the image. The same procedure is done for test images and once their encoding is complete, a NN classifier based on histogram intersection kernel is applied.

Each image is normalised to 128x128 size.

The descriptors are sparse and dense SIFT. For descriptors, VLFeat library is used.

Codebook generation is done via kmeans clustering. Kmeans is applied to the whole training set. The resulting cluster centres are codebooks. Kmeans is applied with 100 and 500 cluster numbers.

Having generated the codebook, each test image is encoded. A histogram showing cluster centre assignments for a given test image descriptor is the histogram encoding. There is also a binary encoding, which is nothing but the binary version of the histogram encoding. For test images, histogram and binary encoding is obtained as well.

Lastly, test images are compared with train images in terms of histogram kernel matching metric. This metric is:

Using this metric, we search for the nearest neighbour by finding the maximum intersection; and assign the label.







asd