

CS 423/523 Computer Vision Project 3 and 4 Bag of Words - Scene Categorisation Report

Implemented Algorithms :

1. Tiny Image representations and K Nearest Neighbours(KNN) classification.
2. Bag of Sift representations and KNN classification.
3. Bag of Sift representations and KNN classification.

Tiny Image Representations :

In order to obtain Tiny Images, training data set is processed in three steps. First, mean value is extracted from each training image in the test in order to obtain zero mean images. Next, obtained her mean images are resized to 16x16 images and images are converted to 1x256 vector. Lastly, these vectorised training images are normalised via element wise division with mean. After the training set is obtained, same process is computed for each test image and vectorised test images are obtained. In order to classify data NN and KNN algorithms are implemented, performance results for various K values are given in Table 1.

Tiny Images and KNN performance

K values	Performance
1 (NN)	0.21
5	0.20
10	0.19
20	

Table 1.

Bag of Sift :

Bag of Sift is a Bag of Words where the descriptors that are to be clustered to obtain visual words are Sift descriptors. Therefore, a point cloud of Sift descriptors have to be extracted from the training set in a manner that the extracted features are unified samples from training set. In order the satisfy this statement, Sift descriptors are sampled from each training image by scanning the image with a design parameter Step Size. Step Size governs our sampling rate, as step size increases the algorithm will take less samples from each image where our samples are Sift descriptors at the sample position on the image. After the sift descriptors are obtained, visual words are computed by KMeans algorithm. KMeans algorithm determine k number of clusters given the point cloud data where k is an input of KMeans. Bag of Sift algorithm's second design parameter Dictionary Size determines the k value which is to be used in Means algorithm.

Bag of Sift Descriptors of Images :

In order to make a classification in Bag of Sift space, all images have to be represented by the dictionary that is created according to Step Size and Dictionary Size. This representation is obtained in three steps for a given image. First, Sift descriptors are extracted in the same manner the dictionary is built, where only sampling rate is doubled in contrast to Bag of Sift creation which means that Step Size is halved. Sample rate is doubled in order to represent the data with more detail. Secondly, extracted Histogram of the extracted Sift descriptors are obtained by calculating how many descriptors belong to each visual word in our Bag of Words.

Bag of Sift and KNN Results :

Results for KNN classification is evaluated with various range of k values and dictionary and step sizes. Confusion matrix , true positives, false positives and false negatives are evaluated for the best performance configuration.

Confusion Matrix for best case of



K Value	Dictionary Size	Step Size	Performance
10	50	20	0.62
10	100	20	0.62
10	100	10	0.61
10	300	20	0.62
10	300	15	0.61
10	300	10	0.62
20	300	10	0.62
50	300	10	0.60

Class Name	True Positives	False Positives	False Negatives
Bedroom	35	64	81
Coast	174	82	86
Forest	215	78	13
Highway	127	100	33
Industrial	35	28	176
Inside City	145	59	63
Kitchen	64	162	46
Living Room	78	113	111
Mountain	195	55	79
Office	81	105	34
Open Country	136	39	174
Store	54	16	161
Street	168	117	24
Suburb	125	66	16
Tall Building	211	58	45

Bag of Sift and SVM :

Classification is made by training 1vsALL SVM's for each class and classifying a test image according to maximum confidence among all SVM's responses. Since there is 100 positive and 1400 negative samples for each class, class weights are determined in order to reduce bias before the training phase of 1vsALL SVM's. Results for 1vsALL SVM classification is evaluated with various range of C values and dictionary and step sizes. Confusion matrix , true positives, false positives and false negatives are evaluated for the best performance configuration.

Confusion Matrix for best Case of SVM



C Value	Dictionary Size	Step Size	Performance
100	50	20	0.64
100	100	20	0.68
100	100	10	0.66
100	300	20	0.71
100	300	15	0.70
100	300	10	0.70
10	300	10	0.67
0.1	300	10	0.14

Class Name	True Positives	False Positives	False Negatives
Bedroom	53	80	80
Coast	189	73	71
Forest	207	37	21
Highway	123	45	37
Industrial	98	47	113
Inside City	162	55	46
Kitchen	66	70	44
Living Room	92	82	97
Mountain	209	71	65
Office	73	47	42
Open Country	203	91	107
Store	133	55	82
Street	156	43	36
Suburb	112	47	29
Tall Building	222	44	34