CENG 483

Introduction to Computer Vision

Fall 2021-2022

Take Home Exam 2
Object Recognition
Student Random ID:

Please fill in the sections below only with the requested information. If you have additional things to mention, you can use the last section. Please note that all of the results in this report should be given for the **validation set**. Also, when you are expected to comment on the effect of a parameter, please make sure to fix other parameters.

1 Local Features (25 pts)

- Explain SIFT and Dense-SIFT in your own words. What is the main difference?

- Put your quantitative results (classification accuracy) regarding 5 values of SIFT and 3 values of Dense-SIFT parameters here. In SIFT change each parameter once while keeping others same and in Dense-SIFT change size of feature extraction region. Discuss the effect of these parameters by using 128 clusters in k-means and 8 nearest neighbors for classification.

Sift is a keypoint detector that is designed to capture features that are indepent from the scale of the image. Its scale invariance property stems from difference of gaussian operations with different kernel size resolutions which captures keypoints at different scales. Then we can quantize an orientation for each keypoint and when orientations of all features are described wrt the keypoint orientation, then it'll additionally have rotation invariance alongside scale invariance. These are all desired invariances that are needed to select interesting or descriptive points in images in order to identify, match, classify etc. Dense-Sift is a kind of modification on top of conventional Sift method. In Sift locations of keypoint descriptors are determined by the default algorithm. On the contrary, in Dense-Sift center-positions (x,y) can be fed to Sift algorithm to create descriptors around that point, or equivalently one can partition the image into grid like portions and feed them individually through default Sift which in a way has similar behaviour, resulting in specific sift vectors around that spatial region.

Sift: Parameters vs Accuracy					
Accuracy (1.0)	nfeatures	nOctaveLayers	contrastThreshold	edgeThreshold	sigma
0.188	0	3	0.04	10	1.6
24 x 24	1.0	1.0	0.36	0.22	1.0
16 x 16	1.0	1.0	0.3	0.255	1.0
12 x 12	1.0	1.0	0.31	0.3	1.0

2 Bag of Features (45 pts)

- How did you implement BoF? Briefly explain.

- Give pseudo-code for obtaining the dictionary.

- Give pseudo-code for obtaining BoF representation of an image once the dictionary is formed.

- Put your quantitative results (classification accuracy) regarding 3 different parameter configurations for the BoF pipeline here. Discuss possible reasons for each one's relatively better/worse accuracy. You are suggested to keep $k \leq 1024$ in k-means to keep experiment durations managable. You need to use the best feature extractor you obtained in the previous part together with the same classifier.

3 Classification (30 pts)

- Put your quantitative results regarding k-Nearest Neighbor Classifier for k values 16, 32 and 64 by using the best k-means representation and feature extractor. Discuss the effect of these briefly.

- What is the accuracy values, and how do you evaluate it? Briefly explain.

- Give confusion matrices for classification results of these combinations.

4 Additional Comments and References

(if there any)