CS376: Computer Vision: Assignment 3
Due: April 26th, 11:59 PM

1 Short answer problems [20 pts]

1. In stereo matching, what is the search space for each pixel of one image in the other image?

2. In stereo matching, what is the difference between determining the target pixels independently and determining the target pixels jointly? Please list three criteria for determining the target pixels jointly.

3. In k-nearest neighbor classifier, what is impact of varying k on the classification accuracy?

4. List two differences between non-parametric methods and parametric methods for visual recognition.

2 Image Classification (80 points)

Our task is to perform image classification using the CIFAR-10 dataset, which consists of 50K training images and 10K testing images. We will experiment with two categories of methods: non-parametric methods and parametric methods. Our goal is to understand the performance of each method on this dataset. Our specific aims are:

- Compare the performance between k-nearest neighbor classifier and support vector machine classifier;
- Analyze the performance of these two methods when varying the size of the training set and the size of the testing set;
- Perform cross-validation to study the hyper-parameters of each method;
- Try different feature representations.

Step 0: Download the CIFAR-10 data set. The dataset can be downloaded from `https://www.cs.toronto.edu/~kriz/cifar.html`.

Step 1 (10 points): K-nearest neighbor classifier. Implement the k-nearest neighbor classifier with $k = 10$. Each image of size $32 \times 32 \times 3$ is represented as a vector with dimension 3072. Please report an overall accuracy as well as a confusion matrix.

Step 2 (15 points): Support vector machine (or SVM) classifier. Try a 1-versus-all SVM classifier. For one class, the instances in that class are considered positive instances, while the instances in other classes are considered negative instances. To turn the SVM classifier. Please use the libsvm built in for Matlab, which can be found at `https://www.csie.ntu.edu.tw/~cjlin/libsvm/#download`. Note that to use libsvm, please read the reference found at `https://www.csie.ntu.edu.tw/~cjlin/papers/guide/guide.pdf`. Libsvm essentially solves the following optimization problem:

$$\begin{aligned} \min_{\boldsymbol{w},b,\psi} \quad & \frac{1}{2}\boldsymbol{w}^T\boldsymbol{w} + C\sum_{i=1}^{l}\psi_i \\ \text{subject to} \quad & y_i(\boldsymbol{w}_i^T\boldsymbol{x}_i + b) \geq 1 - \psi_i \\ & \psi_i \geq 0 \end{aligned}$$

Note that you will have to solve one such program for each class. During the testing time, the trained linear classifier gives you a score for each class. The class with the maximum score is the associated label. We may choose a particular λ for this step. Again, please report a confusion matrix.

Step 3 (5 points): Comparison. Please compare the performance of K-nearest neighbor classifier and SVM classifier. Which classes does K-nearest neighbor do better and which classes does SVM do better, and why? How about running time?

Step 4 (10 points): Cross-validation I. Perform 5-fold cross-validation for the K-nearest neighbor classifier. Report the optimized hyper-parameter K and the corresponding confusion matrix.

Step 5 (10 points): Cross-validation II. Perform 5-fold cross-validation for the SVM classifier. Report the optimized hyper-parameter C and the corresponding confusion matrix.

Step 6 (10 points): Varying training dataset I. Apply K-nearest neighbor classifier to reduced training datasets of size 10K and 30K images. Apply 5-fold cross-validation to optimize the parameter C. Report the confusion matrix in each case.

Step 7 (10 points): Varying training dataset II. Apply SVM classifier to reduced training datasets of size 10K and 30K images. Apply 5-fold cross-validation to optimize the hyper-parameter C. Report the confusion matrix in each case.

Step 8 (10 points): An alternative feature representation. Apply K-nearest neighbor or SVM classifier on an alternative feature representation, e.g., Histogram of gradients (or HOG)`https://www.mathworks.com/help/vision/ref/extracthogfeatures.html`. Again, please perform 5-fold cross validation to optimize the hyper-parameter. Compare the performance gain with respect to the image pixel representation.

Submission instructions:

Create a single zip file so submit on Canvas that includes

- Your well-commented code, including the files and functions named as specified above.
- A pdf writeup of your results with embedded figures where relevant.

Please do not include any saved matrices or images etc. within your zip file.