ECS 174: Intro to Computer Vision, Spring 2020
Problem Set 4

Instructor: Yong Jae Lee (yongjaelee@ucdavis.edu)
Instructor: Krishna Kumar Singh (krsingh@ucdavis.edu)
TA: Haotian Liu (lhtliu@ucdavis.edu)
TA: Utkarsh Ojha (uojha@ucdavis.edu)
TA: Yuheng Li (yhnli@ucdavis.edu)

Due: Wednesday, June 10th, 11:59 PM

Instructions

1. Answer sheets must be submitted on Canvas. Hard copies will not be accepted.

2. Please submit your answer sheet containing the written answers in a file named: FirstName_LastName_PS4.pdf.

3. You may complete the assignment individually or with a partner (i.e., maximum group of 2 people). If you worked with a partner, provide the name of your partner. **We will be checking for instances of plagiarism/cheating.**

1   Short answer problems [60 points]

1. In the boosting algorithm AdaBoost, an ensemble of classifiers are selected in sequence, and the weight wi on each labeled training example xi is adjusted from iteration to iteration. What is the purpose of these weights? How do they influence which feature+classifier combination is selected in the next round? [10 points]

2. Invariant interest point detection and geometric verification using a parametric transformation are both commonly used in "instance" recognition (e.g., to recognize a landmark building in a tourist photo), but not in object category recognition (e.g., to recognize any dog). Explain why. [10 points]

3. How does a k-nearest neighbor classifier use the k nearest neighbors to make a label prediction? [10 points]

4. A deep neural network has multiple layers with non-linear activation functions (e.g., ReLU) in between each layer, which allows it to learn a complex non-linear function. Suppose instead we had a deep neural network without any non-linear activation functions. Concisely describe what effect this would have on the network. (Hint: can it still be considered a deep network?) [10 points]

5. One module of a standard convolutional neural network is the max-pooling operation. Given the 4x4 image below, perform max-pooling with a stride of 2 and a pooling window of 2x2. [10 points]

2	5	3	56
20	5	1	32
3	3	7	46
4	3	12	23

6. Determine whether each statement is true or false (no need to explain your answer) [10 points]:

   a. To detect profile (side) views of faces with the boosting-based Viola-Jones face detector, one should re-train and select a new set of discriminative features using profile face images.

   b. A key idea of deep learning for visual recognition is to learn a feature hierarchy all the way from pixels to classifier.

2 Backpropagation [40 points]

Consider the following function:

$$f(w, x, y, z) = (w + xy)z \qquad \text{where} \quad w = 1, x = -1, y = 2, z = 1$$

Draw the computational graph for this function, and write all the intermediate variable output values as well as their partial derivatives (i.e., $\partial f/\partial(variable)$). Clearly define any intermediate variables that you use (*Hint:* there should be two). Ultimately, we are interested in computing the following four quantities:

$$\partial f/\partial w, \partial f/\partial x, \partial f/\partial y, \partial f/\partial z$$