BMI/CS 567 Medical Image Analysis
University of Wisconsin-Madison
Assignment #5

Instructor: Jeanette Mumford
Due May 1, 2018 *before* 2PM

For this assignment you should turn in a single pdf file containing your code and output requested.

(1) (5 points) For this problem you will do the second additional task associated with Practical Lesson 7.6.2. Specifically, using the nearest neighbor interpolation example that we went over in class (in the book-related file: `NNInterpolation_7.m`), which applied a rotation image transformation with Nearest Neighbor interpolation, adapt it to instead use a bilinear interpolation. Plot the nearest neighbor result along with the bilinear interpolation result, side-by-side.

(2) (5 points) This problem is similar to part of the first additional exercise for Practical Lesson 9.9.2. You may not use any built in MATLAB functions for computing histograms, entropy or mutual information. Note, we went through the code for this practical lesson in class. Compute and display entropy for each image (the CT and T1), which was defined in class as

$$H = \sum_i p_i log\left(\frac{1}{p_i}\right).$$

Be sure to leave out any probabilities of 0 in the computation, since this will give you a `Nan` or `inf` result. Use the natural log, which is simply `log` in MATLAB. Next, for angles between -90 to 100 (in increments of 10, as we did in class) compute the joint entropy, which was defined in class as

$$H(CT, T1) = \sum_{i,j} p_{i,j} log\left(\frac{1}{p_{i,j}}\right),$$

were CT and $T1$ refer to the CT and T1 images used and p_{ij} are the probabilities in the joint histogram. *Be sure you use the probabilities in your calculation.* Change the figure heading of each joint histogram to show the joint entropy value as well as the normalized mutual information,

$$\frac{H(CT) + H(T1)}{H(CT, T1)}.$$

You may need to use the newline command (`\n`) and `%.4g` in your `sprintf` command to make the figure headings readable.

(3) (5 points) Use the test.txt and train.txt data that were distributed with the homework assignment (read them in using `load('path/to/file', '-ascii')`). These data consist of normalized handwritten digits scanned from envelopes by the U.S. Postal Service where each image is a 16×16 grayscale image. The files contain 257 columns, where the first column indicates the written digit and the following 256 columns represent the grayscale values of the written digit. If you read test.txt with the variable name `test`, then `test(1,1)` shows the digit represented in the image and `imagesc(reshape(test(1,2:end), 16,16)')` will plot the first image, which you will see is an image of the digit that matches `test(1,1)`. Using all 256 features, write the code for K-nearest neighbors to perform this 10 category classification. You must write this code yourself and cannot use any KNN algorithms either built into Matlab or supplied online. Since there are more than 2 categories, you cannot directly use the code from lecture, but must adapt it to use Euclidean distance for the identification of neighbors and classification is determined by majority vote, randomly selecting in the case of ties. Using 5 nearest neighbors, classify the test data according to the training data. Create a bar plot that shows the classification accuracy for each digit.