Pattern Recognition and Machine Learning

James L. Crowley

ENSIMAG 3 Final Exam - Feb 2018

Conditions: You have the right to use any notes or written material. You may answer questions in English or in French. When appropriate, illustrate your answer with mathematics. Your written answers must be clear and legible. Illegible text will not be graded. Duration: 3 hours.

You are asked to use Principal Components Analysis to learn an orthogonal set of local feature detectors for handwritten digits, and then to construct a Bayesian Pattern Recognizer for handwritten digits using Gaussian Mixture Models. Each local feature detector, $f_d(i,j)$, is an 8 x 8 window of coefficients that can be convolved with an image. Convolution with a set of D feature detectors will provide a D dimensional feature vector, $\vec{X}(i,j)$ for each pixel.

You will use a training image to learn a Gaussian Mixture Model $p(\vec{X} \mid C_k)$ of the local features for each class of digit, C_k. This Gaussian Mixture Model can then be used to compute the probability that a pixel is within a bounding box of a digit $p(C_k \mid \vec{X})$ using Bayes rule. Assume that there is an 11th class that is not within the bounding box of any digit.

1) (4 points) Principal Components Analysis: You are provided with 1000 gray-scale imagettes of handwritten digits of size 8 by 8 pixels. Present and describe the calculations involved in using Principal Components Analysis to compute the feature detectors, $f_d(i,j)$. How many feature detectors are computed? How can you select a subset of D feature detectors that will capture 90% of the energy for reconstructing the dataset.

2) (4 points) Feature Detection: You are given a labeled training image of handwritten digits. Explain how to use convolution with the D feature detectors to build a D dimensional feature vector, $\vec{X}(i,j)$, for each image position of the training image. Assume a pixel value of 0 for image positions outside the image.

3) (4 points) Expectation Maximization: Ground truth for the training image is in the form of function $y(i,j)$ that returns a class label C_k for each pixel. For pixels within a bounding box of a digit, k is the digit (0 through 9). For pixels that are not within a bounding box (the null class), k=10. The training image contains M_k examples for each class C_k. You have a total M image positions for training.

Explain how to use Expectation Maximization to learn 11 Gaussian Mixture models with N components that represents $p(\vec{X} \mid C_k)$ for each of the 11 classes from the labeled training image.

4) (4 points) Bayesian Classification: Explain how to use Bayes rule and the 11 Gaussian Mixture models to determine the most probably class, $\hat{C}_k$, at each pixel in an image.

5) (4 points) Performance Evaluation: Explain how to use a new labeled image, with ground truth, to compare the performance of your detector, as a function of the numbers of feature detectors, D, and the numbers of principal components, N. What performance metrics do you propose? How are they calculated? How are they interpreted?