Pattern Recognition / NCTU Spring 2020 Program Assignment #2 (due: 5/26/2020)

The goal of this program assignment is to let you experiment a little more with classifiers, especially with dimensionality reduction. You can use Matlab, python, or C++. You are not allowed to use existing libraries/modules/toolboxes for classification/evaluation, but libraries for linear algebra operations are allowed. You can continue to use the datasets mentioned in the first assignment, or any other datasets that interest you.

Task 1: FLD/LDA

Apply FLD/LDA to find the projection that preserves the most class separability. Compute the separability measures before and after the projection. For two-class datasets, also plot the resulting ROC curves as well as AUC. For datasets also used in the first assignment, be sure to compare with the results there from the linear classifier implemented then.

Task 2: PCA and classification

Take the datasets used in the first assignment and/or the SVM experiments, apply PCA to them, and then repeat the experiments. The objective is for you to see how PCA affects classification performance. Pay attention to how you choose the retained the dimensions as this is an important factor here. You do not need to repeat every experiments you did previously, but you should do enough for you to make meaningful observations about the effect of PCA. Also note whether the effects of PCA are different for different classifier types. Here you need to split your data into training and validation subsets, and the projection matrices are only derived from the training subset. The same splitting should be used when comparing different classifiers or settings.

Task 3: Eigenface and classification

Here your goal is to use the provided face image datasets for gender classification (fP1.bmp and mP1.bmp) and face recognition (facesP1.bmp, 5 images for each class), respectively, after PCA (the "eigenface" version) is applied to the images.

When working with face images, it is a common practice to include their horizontally flipped versions as well (training subset only). This is considered a way of training data augmentation.

Write a report (limited to 10 pages, single-spaced) describing

- Methods you have implemented.
- Experiments you have done, and the results.
- Analysis - Are the results what you expect? Why?

The report (PDF format) is to be submitted electronically through e3. Also include the program listing as an appendix to your report (in the same file, not counted toward the page limit).

Late submission policy: 10% credit deduction for each day late; up to 7 days late accepted.