Lab 3: PCA-Based Face Recognition

In this lab, we will use PCA to extract eigen-faces as features from a dataset of face images. Using the dataset, we will construct a subspace with dimensionality (k) less than or equal to the dimensionality (d) of the dataset such that this subspace has the maximum dispersion for the d projections. We will do that by taking the eigenvectors with k largest eigenvalues.

We will use two datasets:
(1) ORL database, available to download on AT&T's website. This database contains photographs showing the faces of 40 people. Each one of them was photographed 10 times. These photos are stored as grayscale images with 112 x 92 pixels.
(2) The Pokemon datafile represents a 2d array where each row is a 64 by 64 pixel greyscale picture. The entries are floats between 0 and 1, where 0 is white and 1 is black. Note that while the images are 64 by 64 entries, the dataset you load has rows of size 4096 which is 64 x 64 to allow the data to be saved as a 2D array.

The first database is kept in a folder called 'orlfaces' in the compressed file provided for this lab. For every person, 9 of the 10 images are used in the training set and 1 image is kept for the test set. The pictures belonging to the 40 people are kept in 40 different folders called s1, s2, …, s40 for both training and testing. You can load the second dataset from the file ``pokemon.csv`` using ``np.loadtxt``.

For this lab, first we will perform PCA on the training data and observe some of the eigenfaces. Then we will reconstruct a training image using $k = \{10, 20, 30, 40\}$ eigenfaces to see how taking more eigenfaces affect the reconstruction. Finally, we will predict the identity of the test images using k eigenfaces by comparing the projections of the test images with the projections of the training images.

All the instructions are provided in the following Google Colab template-

https://colab.research.google.com/drive/1fo-8xq6TILh0dVg7hrWdcHxdat-PEqYZ?usp=sharing

Submission instructions:

Deliverables:

- Lab3_PCA_based_Face_Recognition.ipynb (Notebook with your code, answers/comments and figures)
- Lab3_PCA_based_Face_Recognition.py (download from Colab)

Compress the deliverables into one zip file and name it 'lab3_<your JHED id>'. Submit the zip file on Canvas.