Eigen Faces

Ali Abdollahian

Problem?

- **High Dimensionality of Image Data**
  Image data, especially face images, typically involve thousands of pixels. Each pixel represents a dimension, leading to very high-dimensional data.

- **Importance of Variance**
  In face images, not all pixels are equally important for distinction. Some remain constant across faces

Solution

- **Principal Component Analysis (PCA)**
  PCA is a statistical method used to reduce the dimensionality of a dataset while retaining those characteristics of the dataset that contribute most to its variance.
  This is achieved by transforming the original variables into a new set of variables, which are the principal components. These components are uncorrelated and ordered so that the first few retain most of the variation present in all of the original variables.

Eigen Faces

- **Eigenfaces**: In the context of face recognition, PCA is applied to a collection of face images to produce a set of basis features. These features are called "eigenfaces" because they are the eigenvectors of the data's covariance matrix. Essentially, eigenfaces are the "principal components" of the face dataset and represent the directions in which the images vary the most.

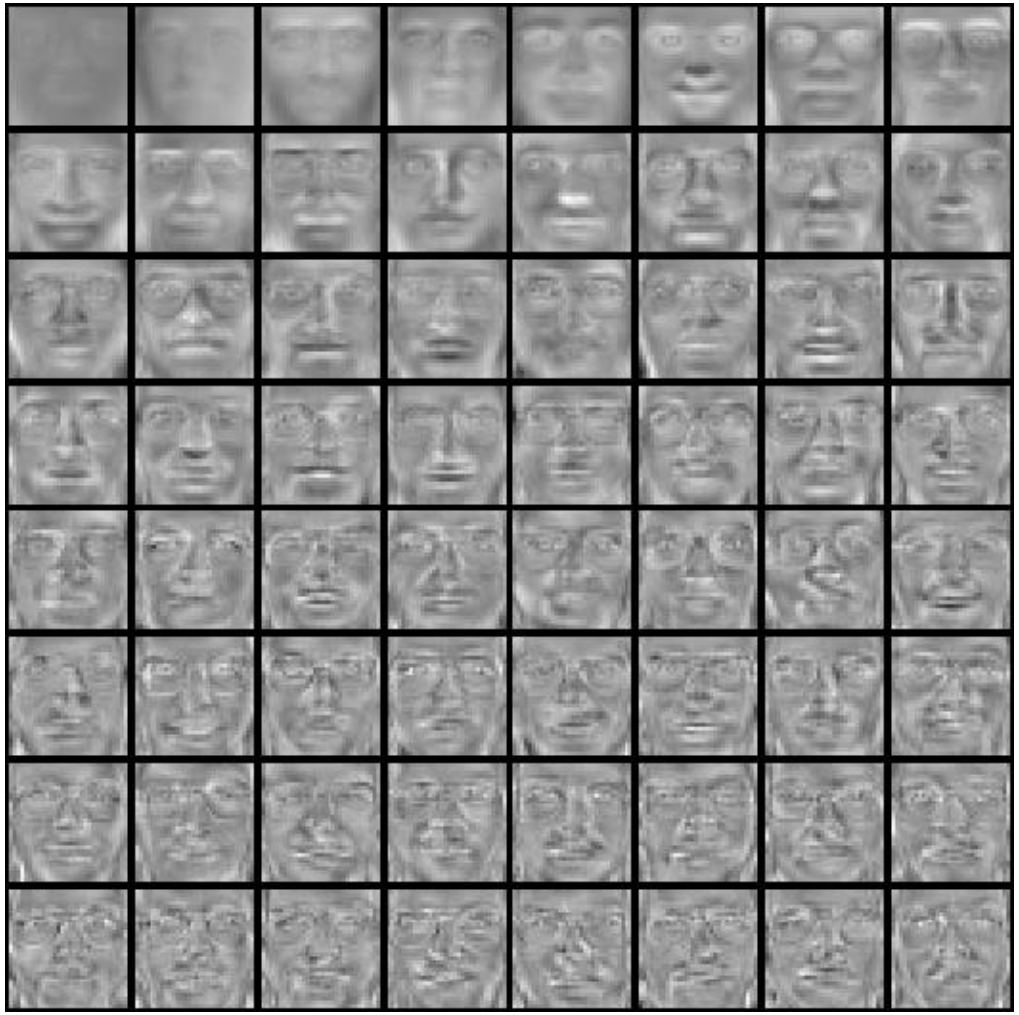
PCA Steps

- **Step 1: Gather Data**
  Each image is transformed into a one-dimensional vector. For instance, if an image is 100x100 pixels, it becomes a vector of 10,000 dimensions.

- **Step 2: Standardize the Data**
  involves subtracting the mean face (average image) from each face vector. The mean face is calculated by averaging each pixel across all images.

Mean: μ

Top eigenvectors: $u_1,\ldots u_k$



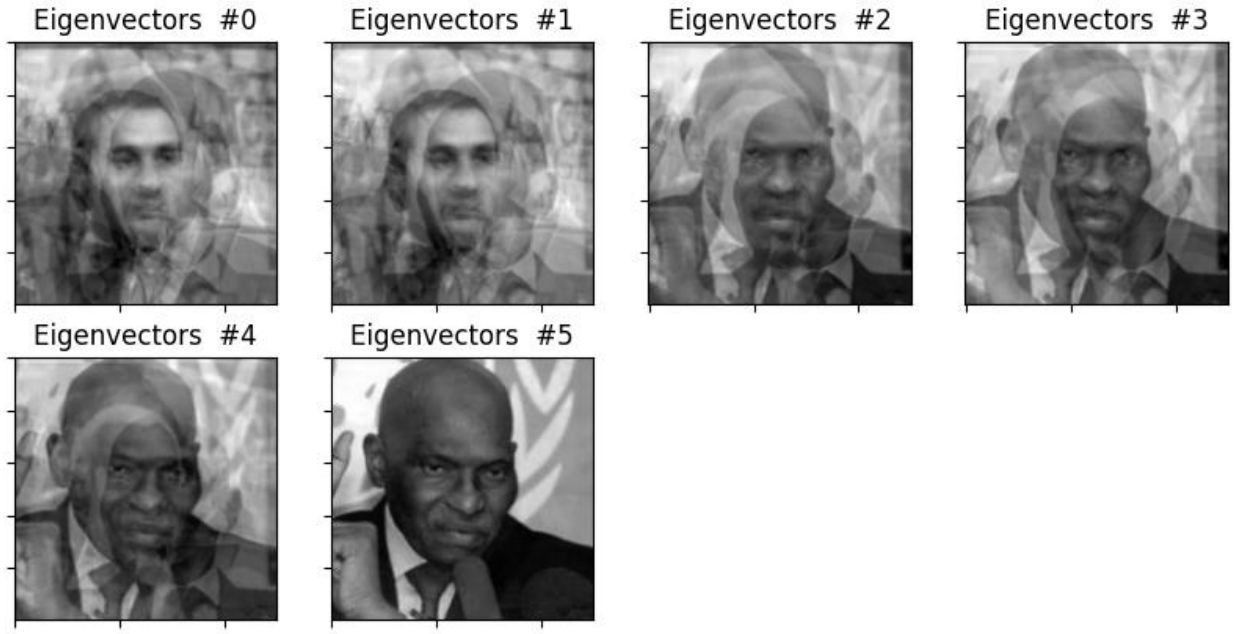
PCA Steps

- **Step 3: Calculate the Covariance Matrix**
  The covariance matrix captures how changes in one dimension correlate with changes in another. In the context of face images, it shows how changes in pixel values are related across all images.

- **Step 4: Calculate Eigenvalues and Eigenvectors**
  The eigenvectors of the covariance matrix are computed next. These eigenvectors represent the directions in the data that have the highest variance, which are the principal components.

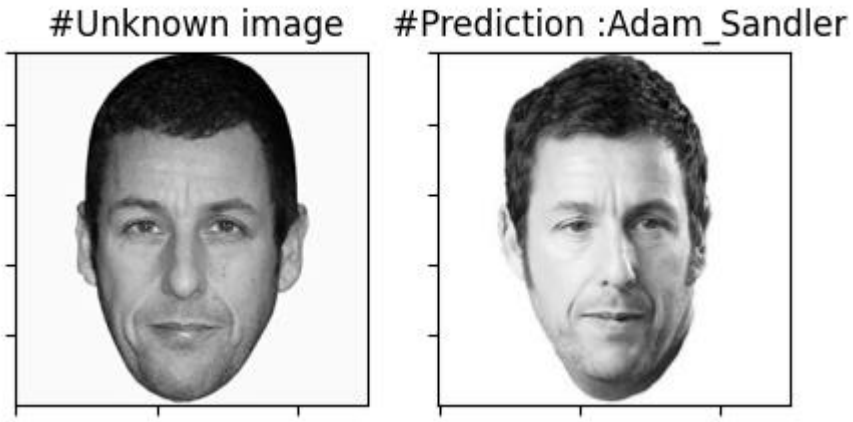
PCA Steps

- **Step 5: Select Principal Components**
  Typically, not all eigenvectors are used—only those corresponding to the largest eigenvalues. This is because the largest eigenvalues represent the dimensions where the data varies the most, and these are the most informative for recognizing different faces.



Face detection

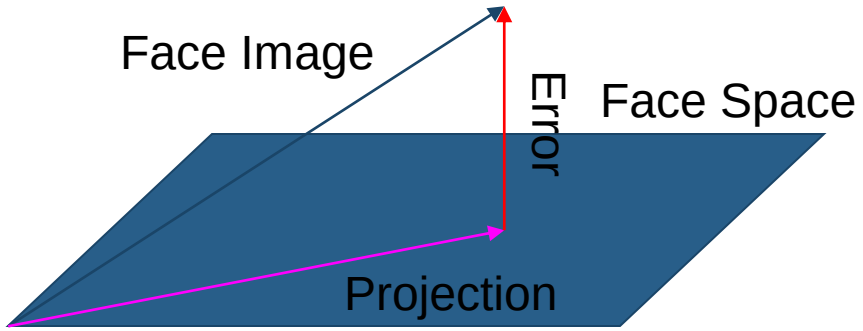
- **Project Faces onto the Eigenfaces**
  To identify or verify a face using the eigenfaces, one projects the standardized face vector (image minus mean face) onto the selected eigenfaces. This reduces the dimensionality of the original face vector from thousands of pixels to just a few eigenfaces, while retaining the most important features of the data.



Face detection and recognition

- **The error is the difference between the original image and its projection image onto eigenface space.**

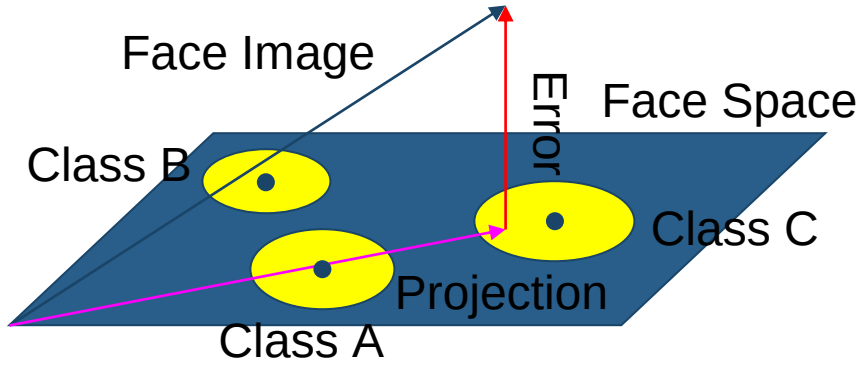
- **If the error is within a threshold, the image is detected as a face image.**
- **Efficient calculation available.**



Face detection and recognition

- **If the projection of face image onto eigenface space is close to one training face, it is identified as that training face.**

- **Distance measure can be Euclidian distance.**



Face detection and recognition

Detection

Recognition

“Sally”

Mathematical Overview

- Let X be a matrix where each column is a vectorized standardized face image.
- The covariance matrix C is calculated as $C = \frac{1}{n-1} X X^T$
- Solve the eigenvalue problem $Cv=\lambda v$, where v are eigenvectors and λ are eigenvalues.

Mathematical Overview

- Sort eigenvectors by decreasing eigenvalues, and select the top vectors.
- Each face image can then be approximated by a linear combination of these top eigenvectors (eigenfaces).