1. PCA based data distribution visualizations

   Let the data matrix $\mathbf{X}$ be of $n \times p$ size.
   First, perform SVD:

$$\mathbf{X} = \mathbf{U}\mathbf{S}\mathbf{V}^\intercal \tag{1}$$

   The columns of $\mathbf{U}\mathbf{S}$ are the principal components and the columns of $\mathbf{V}$ are the eigenfaces.
   For the principal components, only calculation for the highest dimension is necessary. For reduction to dimension k, the first k principal components can be obtained from the first k columns of $\mathbf{U}$ and the diagonal matrix formed by the first k values along the diagonal of $\mathbf{S}$.
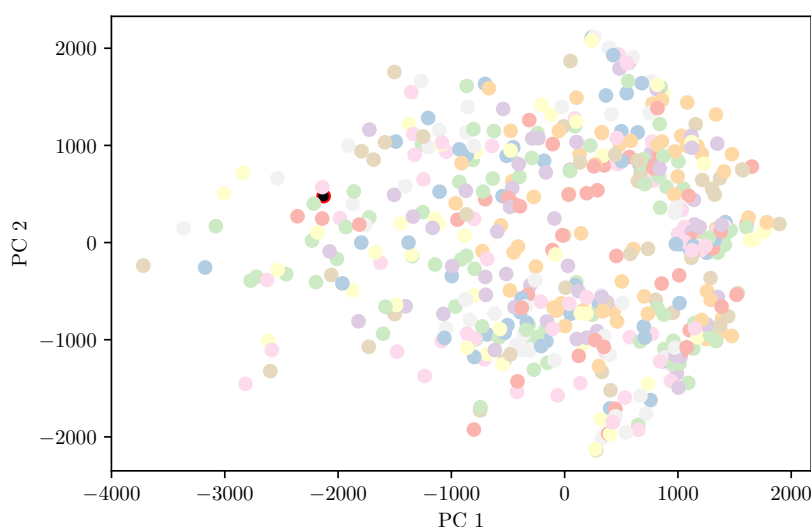

Figure 1: PCA projected data vector in 2D.

2. PCA plus nearest neighbor classification results

   For classification with PCA, obtain $\mathbf{X}\mathbf{V}$, projection of the faces in the test data matrix $\mathbf{X}$ on the eigenface space. Each column of $\mathbf{V}$ and therefore each column of $\mathbf{X}\mathbf{V}$ corresponds to a principal component.
   Again, only calculation for the highest dimension is necessary. The first k projections can be obtained with the test data matrix and the first k columns of $\mathbf{V}$.

   It is expected that classfication of selfies is unsuccessful. There are only three selfies in the test set and random sampling of 500 images for training included only one selfie which is an outlier in 3D as can be seen from the scatter plots. Moreover, my setup and lighting condition differs from those of the images in the PIE data set.
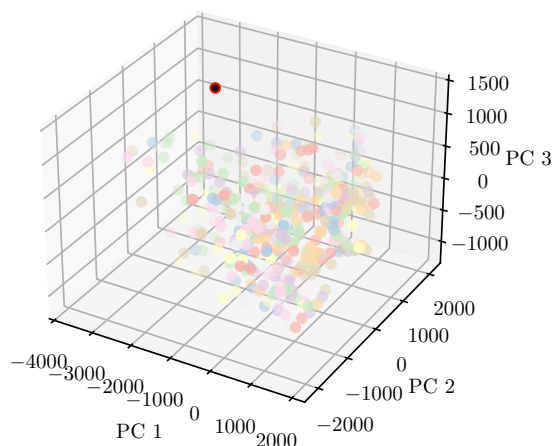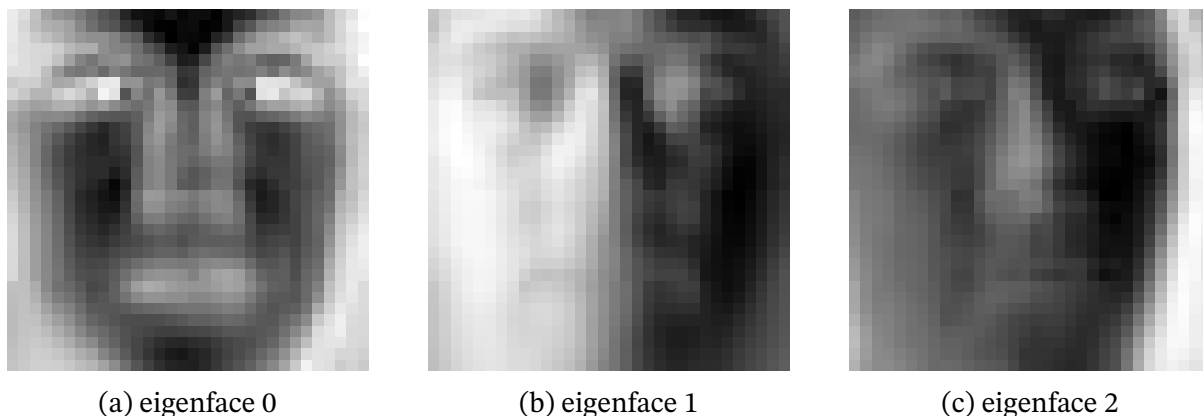
Figure 2: PCA projected data vector in 3D.

| (a) eigenface 0 | (b) eigenface 1 | (c) eigenface 2 |

Figure 3: Corresponding 3 eigenfaces used for PCA dimensionality reduction.

Test Images	Dimensionality	Accuracy
CMU PIE	40	49.02 %
	80	52.47 %
	200	55.22 %
Selfies	40	0.00 %
	80	0.00 %
	200	0.00 %

3. LDA based data distribution visualization

4. LDA plus nearest neighbor classification results

5. SVM classification results with different parameter values

6. CNN classification results with different network architectures