Dimensionality Reduction & Classification

CISC-820 Project 4

Gobinda Pandey and Justin Cole

December 17, 2024

Objective

- Explore the use of Principal Component Analysis (PCA) for reducing the dimensions of facial images.
- Dataset: 400 images from 40 people with 10 images per person.
- Use PCA to identify eigenfaces representing key facial features.
- Investigate reconstruction quality and classification accuracy.

Using PCA to Extract Eigenfaces

- PCA using Singular Value Decomposition (SVD).
- Eigenfaces: Principal components representing key facial features.
- Visual appearance of leading eigenfaces.
- Decreasing contribution of eigenfaces as their number increases.

Leading Eigenfaces

- First few eigenfaces show recognizable facial features.
- Example: First eigenface is a blurry outline of a face.
- Higher eigenfaces show more details like cheekbones, eyebrows, and ears.

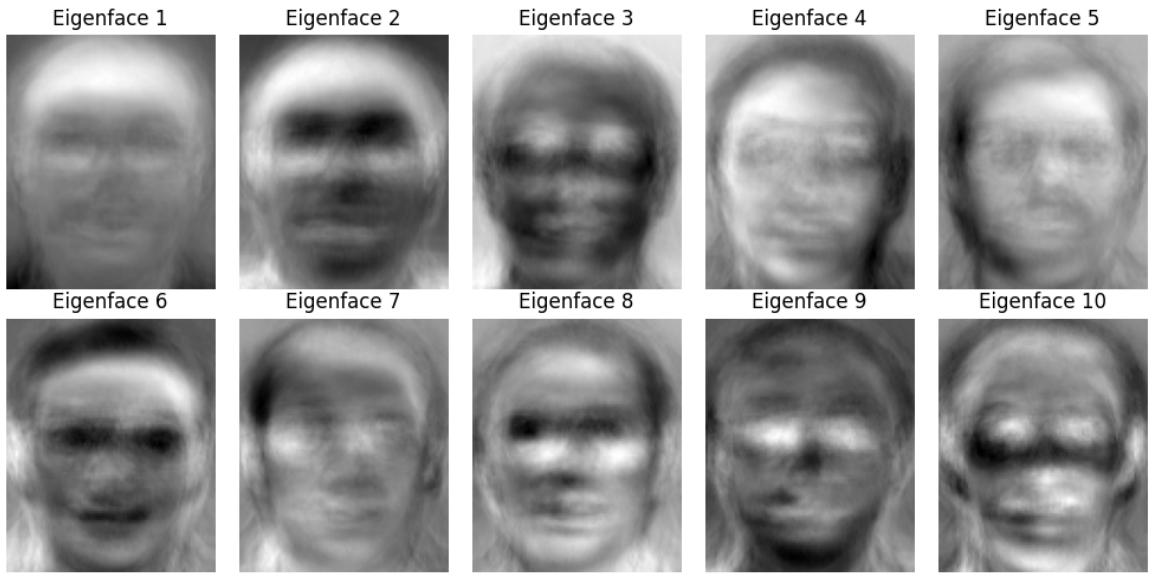

Figure: Example Image of Eigenfaces

Importance of Eigenfaces

- First 50 eigenfaces capture 81.61% of total variance.
- First 100 eigenfaces capture 89.06% of total variance.
- First 400 eigenfaces capture 99.42% of total variance.
- Diminishing returns as more eigenfaces are added.

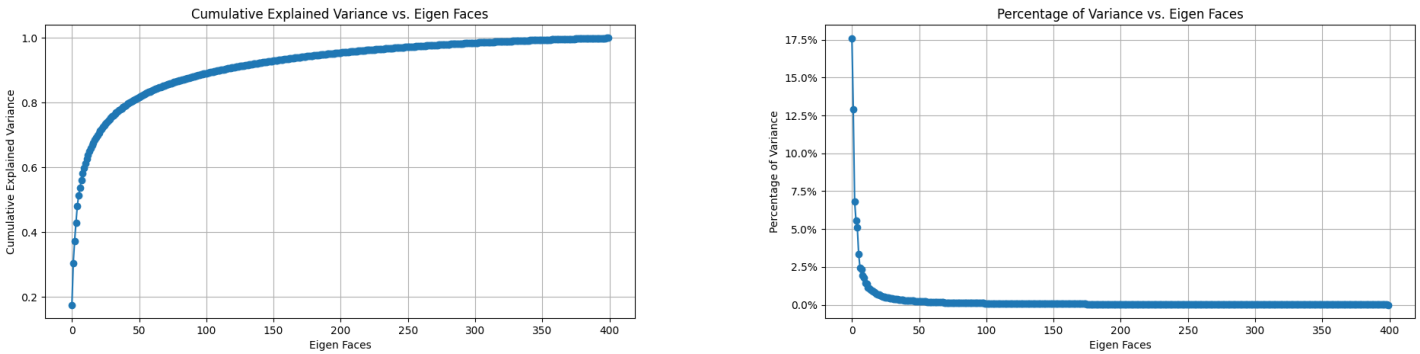

Figure: Variance Captured by Eigenfaces

Explanation of Variance

- **Cumulative Explained Variance:** Shows how many eigenfaces are needed to capture significant variance.
- **Percentage of Variance:** Identifies most important eigenfaces.
- First few eigenfaces capture most variance, subsequent eigenfaces contribute less.

Facial Image Reconstruction with PCA

- Quality of reconstructed images improves with more eigenfaces.
- Fewer eigenfaces (e.g., 10, 50) result in blurry images.
- More eigenfaces (e.g., 150, 250, 350, 400) result in clearer images.
- Higher-dimensional representations preserve fine details.

Facial Image Reconstruction with PCA (Cont'd)

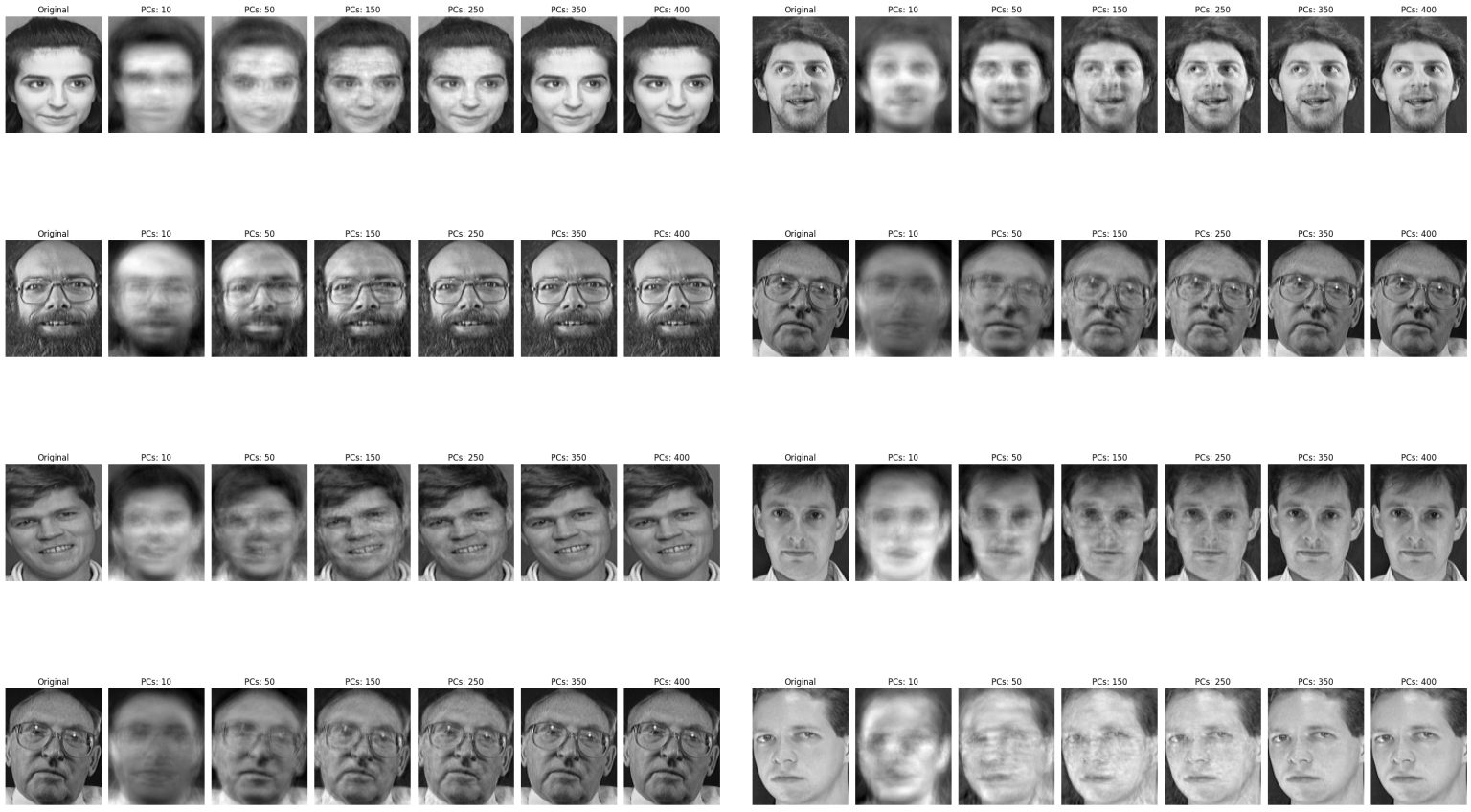

Figure: Reconstructed Images Using Different Numbers of Eigenfaces

Number of Eigenfaces for Reconstruction

- Mean Squared Error (MSE) used to assess reconstruction quality.
- 350 eigenfaces result in low MSE of 8.9894.
- Around 350 eigenfaces are sufficient for accurate reconstruction with minimal error.

Classification Results

- Images divided into 35 classes, 8 images per class, 2 test images.
- Methods: KNN, Logistic Regression, Custom Linear Regression.
- KNN Accuracy: 99.25%
- Logistic Regression Accuracy: 100.00%
- Custom Linear Regression Accuracy: 7.00%
- Lower accuracy for Custom Linear Regression due to continuous predictions.

Conclusion

- PCA using SVD effectively reduces dimensionality of facial images.
- 400 eigenfaces provide near-perfect reconstruction, capturing 99.99% of total variance.
- Eigenfaces represent key facial features and can reconstruct images with varying accuracy.
- KNN and Logistic Regression perform exceptionally well for classification.
- Custom Linear Regression not suitable for this classification task.

Thank You!

Questions?