# Face Recognition Based on PCA

Shen Yuanzhang
*1024041127*
*Nanjing University of Posts and Telecommunications*
School of Computer Science
Nanjing, China

***Abstract*—This paper investigates the application of Principal Component Analysis (PCA) in the context of face recognition. By employing PCA for dimensionality reduction on facial image datasets, the storage requirements are reduced to approximately 5% of the original data size, thereby achieving significant reductions in data redundancy and computational complexity. To evaluate the effectiveness of PCA in preserving critical information for recognition tasks, several classification methods are employed, including K-Nearest Neighbors (KNN), Support Vector Machines (SVM), and a combination of Linear Discriminant Analysis (LDA) with KNN. Experimental results indicate that, even after dimensionality reduction, PCA retains essential features of the original data, enabling recognition accuracies exceeding 90%. These findings demonstrate the robustness and practicality of PCA as a dimensionality reduction technique in face recognition systems, offering an efficient balance between data compression and recognition performance.**

***Index Terms*—PCA, Face Recognition, KNN, SVM,LDA.**

## I. Introduction

Principal Component Analysis (PCA) is a widely adopted and highly esteemed methodology in the domains of pattern recognition and feature extraction, particularly recognized for its superior capability in reducing dimensionality while retaining essential information [1], [2]. In the domain of face recognition, input data frequently exhibit high-dimensional characteristics, such as those associated with high-resolution facial images. These high-dimensional datasets pose significant computational and storage challenges, often resulting in elevated processing costs. Furthermore, such data typically contain a considerable amount of redundant information and noise, which further complicates the recognition process and reduces the efficiency of classification systems [3]. PCA effectively addresses these challenges by transforming the original high-dimensional data into a lower-dimensional space, where the critical information is encapsulated in a set of orthogonal principal components [4]. This transformation achieves a substantial reduction in storage and computational requirements while simultaneously filtering out noise and eliminating redundancy, thereby yielding a more compact and informative feature representation that enhances both the efficiency and accuracy of subsequent recognition algorithms [5].

In this study, PCA was employed to perform dimensionality reduction on facial image datasets, significantly compressing their dimensions. The compressed representations retain only the most salient features of the original data, thereby facilitating efficient processing. To evaluate the effectiveness of PCA in preserving critical information for recognition tasks, several established classification algorithms were utilized, including K-Nearest Neighbors (KNN), Support Vector Machines (SVM), and a combination of Linear Discriminant Analysis (LDA) with KNN. Experimental results demonstrate that PCA not only substantially reduces data dimensions but also retains the essential features necessary for accurate face recognition. Recognition accuracies achieved by the employed classification algorithms consistently exceeded expectations, even after dimensionality reduction, underscoring the capability of PCA to preserve discriminative information vital for effective classification.

The findings presented in this study affirm the practical value and robustness of PCA as a dimensionality reduction technique for face recognition systems. By achieving a significant reduction in data size while maintaining the integrity of essential features, PCA strikes an optimal balance between data compression and recognition performance. These results underscore the importance of PCA in addressing the challenges posed by high-dimensional data in face recognition applications. Moreover, the demonstrated ability of PCA to enhance computational efficiency and maintain high recognition accuracy highlights its role as a foundational tool in the development of scalable, efficient, and reliable recognition systems. This study contributes to the growing body of evidence supporting PCA's applicability in real-world face recognition scenarios, particularly in resource-constrained environments where computational and storage efficiency are paramount.

## II. Detailed Steps of PCA Processing

In this study, Principal Component Analysis (PCA) is applied to reduce the dimensionality of face images, aiming to minimize the data dimensions while preserving the most significant features. The following section explains the mathematical principles, computational steps, and specific implementation of PCA in processing face images.

### A. Mathematical Principles of PCA

PCA is a linear transformation technique that projects high-dimensional data into a lower-dimensional space by identifying the directions (principal components) that capture the maximum variance. The process involves the following steps:

1. **Data Centering**: The data matrix $\mathbf{X}$ (where each row represents a flattened face image) is centered by subtracting the mean vector $\boldsymbol{\mu}$ from each data point:

$$\tilde{\mathbf{X}} = \mathbf{X} - \boldsymbol{\mu} \tag{1}$$

where $\boldsymbol{\mu}$ is calculated as:

$$\boldsymbol{\mu} = \frac{1}{N}\sum_{i=1}^{N}\mathbf{x}_i \tag{2}$$

Here, $N$ represents the number of samples, and $\mathbf{x}_i$ is the $i$-th data sample.

2. **Covariance Matrix Computation**: The covariance matrix $\boldsymbol{\Sigma}$ of the centered data is computed as:

$$\boldsymbol{\Sigma} = \frac{1}{N-1}\tilde{\mathbf{X}}^\top\tilde{\mathbf{X}} \tag{3}$$

3. **Eigenvalue Decomposition**: The covariance matrix $\boldsymbol{\Sigma}$ is decomposed to find its eigenvalues $\lambda$ and eigenvectors $\mathbf{W}$:

$$\boldsymbol{\Sigma}\mathbf{w}_i = \lambda_i\mathbf{w}_i \tag{4}$$

where $\lambda_i$ represents the variance explained by the $i$-th principal component $\mathbf{w}_i$.

4. **Sorting and Projection**: The eigenvalues are sorted in descending order, and the top $l$ eigenvectors (corresponding to the largest eigenvalues) are selected to form the projection matrix $\mathbf{W}_l$. The reduced-dimensional data $\mathbf{X}_{\text{new}}$ is then obtained as:

$$\mathbf{X}_{\text{new}} = \tilde{\mathbf{X}}\mathbf{W}_l \tag{5}$$

### *B. Application of PCA in Face Image Processing*

PCA is applied to the Olivetti face dataset, and the following specific steps are performed:

1. **Dimensionality Reduction**: Each face image (originally a 64×64 pixel grid, flattened into a 4096-dimensional vector) is reduced to a 20-dimensional representation using PCA, significantly reducing storage and computational requirements.

2. **Visualization of Eigenfaces**: The eigenvectors $\mathbf{W}_l$, also known as eigenfaces, are reshaped into 64×64 matrices and visualized as grayscale images. Each eigenface represents a dominant pattern of variation in the dataset.

3. **Computation of the Mean Face**: The mean of all face images is computed as:

$$\text{Mean Face} = \frac{1}{N}\sum_{i=1}^{N}\mathbf{x}_i \tag{6}$$

This mean face provides a baseline representation of the dataset and is used in the reconstruction process.

4. **Face Reconstruction**: Each face image is reconstructed by projecting its reduced-dimensional representation back to the original space:

$$\text{Reconstructed Face}_i = \text{Mean Face} + \mathbf{W}_l\mathbf{x}_{\text{new},i} \tag{7}$$

The reconstructed images retain the main features of the original faces while reducing noise.

5. **Saving and Visualization**: The original images, reconstructed images, mean face, and eigenfaces are saved as image files for further analysis and comparison.

Through these steps, PCA achieves efficient dimensionality reduction while preserving key features of the face images, providing a robust foundation for subsequent recognition tasks. Figure 1 illustrates the comparison between the original images, the mean face, and the PCA-reconstructed faces. This comparison demonstrates that the PCA-reconstructed faces retain the primary characteristics of the original images to a large extent.

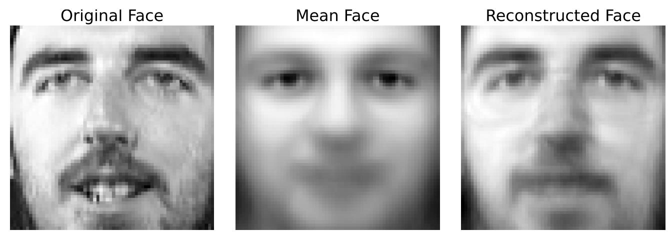


Fig. 1. Comparison of Original, Mean, and PCA-Reconstructed Faces

## III. Face Recognition: Detailed Steps

In this study, the Olivetti Faces dataset is used for face recognition experiments. This dataset consists of 400 grayscale face images, belonging to 40 individuals, with each individual contributing 10 images. Each image has a resolution of 64×64 pixels (a total of 4096-dimensional pixel features). To evaluate recognition performance, the dataset is divided into training and testing sets, with some images of each person used for training and the rest for testing.

### *A. Data Preprocessing and Dimensionality Reduction*

*1) Training and Testing Set Split:* The dataset is split such that a random selection of images (e.g., 7 images per person) is used as the training set, while the remaining images (e.g., 3 per person) are reserved for testing.

*2) PCA Dimensionality Reduction:* Before processing, the training images are reduced in dimensionality using Principal Component Analysis (PCA) to minimize data dimensions while preserving key features. For instance, if images are reduced to 20 dimensions, each image is represented by 20 projection values along the principal directions. The steps are as follows:

1) Compute the covariance matrix from the training set and extract its eigenvalues and eigenvectors.
2) Select the top 20 eigenvectors to form the dimensionality reduction matrix.
3) Project the training images onto the reduced 20-dimensional space using the selected eigenvectors.
4) Similarly, project the test images onto the same 20-dimensional space using the same eigenvectors.

After dimensionality reduction, both the training and testing images are represented as vectors of length 20.

### B. Face Recognition Methods

In the reduced feature space, three classification methods are used: KNN, SVM, and LDA+KNN. The detailed steps of each method are described below.

*1) K-Nearest Neighbors (KNN) Classification:* KNN is a non-parametric method based on distance measurement. The steps for KNN classification are:

1) **Training**: Store the reduced-dimensional training images and their corresponding class labels (one class per individual).
2) **Testing**: For each reduced-dimensional test image $\mathbf{x}_{\text{test}}$, calculate its Euclidean distance to all training samples:

$$d(\mathbf{x}_{\text{test}}, \mathbf{x}_i) = \sqrt{\sum_{j=1}^{l} (\mathbf{x}_{\text{test},j} - \mathbf{x}_{i,j})^2} \tag{8}$$

where $l$ is the dimensionality of the reduced space (e.g., 20), and $\mathbf{x}_i$ is the $i$-th training sample.
3) **Classification**: Assign the test image to the class that appears most frequently among the $k$ nearest neighbors.

*2) Support Vector Machines (SVM) Classification:* SVM is a linear classifier based on maximizing the margin between classes. The steps are:

1) **Training**: Use the reduced-dimensional training set to train a multi-class SVM model. For two classes, the separating hyperplane is given by:

$$\mathbf{w}^\top \mathbf{x} + b = 0 \tag{9}$$

The parameters $\mathbf{w}$ and $b$ are optimized to maximize the margin:

$$\text{Margin} = \frac{2}{\|\mathbf{w}\|} \tag{10}$$

2) **Testing**: For each test image $\mathbf{x}_{\text{test}}$, compute its distance to the hyperplane and use the SVM model to predict its class.

*3) Linear Discriminant Analysis (LDA) + KNN:* LDA is a supervised dimensionality reduction method that maximizes the class separability by optimizing between-class and within-class scatter matrices. The steps are:

1) **Training**:
   - Compute the within-class scatter matrix $S_w$ and between-class scatter matrix $S_b$:

$$S_w = \sum_{i=1}^{c} \sum_{j=1}^{n_i} (\mathbf{x}_j^{(i)} - \boldsymbol{\mu}_i)(\mathbf{x}_j^{(i)} - \boldsymbol{\mu}_i)^\top \tag{11}$$

$$S_b = \sum_{i=1}^{c} n_i (\boldsymbol{\mu}_i - \boldsymbol{\mu})(\boldsymbol{\mu}_i - \boldsymbol{\mu})^\top \tag{12}$$

   where $c$ is the number of classes, $n_i$ is the number of samples in the $i$-th class, $\boldsymbol{\mu}_i$ is the mean vector of class $i$, and $\boldsymbol{\mu}$ is the global mean vector.
   - Solve the generalized eigenvalue problem:

$$S_w^{-1} S_b \mathbf{v} = \lambda \mathbf{v} \tag{13}$$

   Select the top $d$ eigenvectors to form the projection matrix.
   - Project the reduced-dimensional training set into the LDA subspace.
2) **Testing**:
   - Project the reduced-dimensional test set into the same LDA subspace.
   - Use KNN in the LDA subspace to classify the test images, leveraging the discriminative features extracted through the dimensionality reduction process.

## IV. Experiments

This section presents the experimental setup and results, divided into two parts: (1) the impact of PCA dimensionality on recognition accuracy using KNN classification, and (2) the comparison of recognition accuracy across different classification methods (KNN, SVM, LDA+KNN) .

### A. Impact of PCA Dimensionality on Recognition Accuracy

In this experiment, the classification method is fixed as KNN (with $k = 3$), and the effect of varying PCA dimensionality on recognition accuracy is explored. The experimental setup is as follows:

1) **Dimensionality Settings**: The dimensionality of the PCA-reduced feature space is set to 10, 20, 30, and 40. The principal components are extracted accordingly for classification.
2) **Classification Process**: For each dimensionality setting, the PCA-reduced features are used to train a KNN classifier, which is then applied to the test set. The recognition accuracy is computed.
3) **Evaluation Metric**: Recognition accuracy is used as the evaluation metric, defined as:

$$\text{Accuracy} = \frac{\text{Number of correctly classified samples}}{\text{Total number of samples}} \times 100\% \tag{14}$$

4) **Results and Analysis**: The relationship between PCA dimensionality (10, 20, 30, 40) and recognition accuracy is analyzed and visualized through a line graph to illustrate how dimensionality affects classification performance. As presented in Table 1 and Figure 2, the recognition accuracy achieved using KNN is compared under varying numbers of principal components obtained through PCA. The results indicate that recognition accuracy improves as the number of training images increases, demonstrating a clear positive correlation. Under optimal conditions, the recognition accuracy exceeds 90%, highlighting the effectiveness of increasing training data for enhancing classification performance. In the graph, the value of $K$ represents the number of training images used for each subject.

TABLE I
RECOGNITION ACCURACY WITH DIFFERENT PCA COMPONENTS

| 2*Training Pictures | Accuracy | | | |
|---|---|---|---|---|
| | **Components:10** | **Components:20** | **Components:30** | **Components:40** |
| 1 | 0.1917 | 0.2056 | 0.1944 | 0.2028 |
| 2 | 0.3844 | 0.4312 | 0.4438 | 0.4562 |
| 3 | 0.4071 | 0.4929 | 0.4929 | 0.4893 |
| 4 | 0.475 | 0.5958 | 0.6125 | 0.625 |
| 5 | 0.475 | 0.645 | 0.66 | 0.65 |
| 6 | 0.6 | 0.6938 | 0.7438 | 0.7625 |
| 7 | 0.7 | 0.775 | 0.8 | 0.7917 |
| 8 | 0.7625 | 0.8625 | 0.8875 | 0.875 |
| 9 | 0.825 | 0.85 | 0.875 | 0.875 |

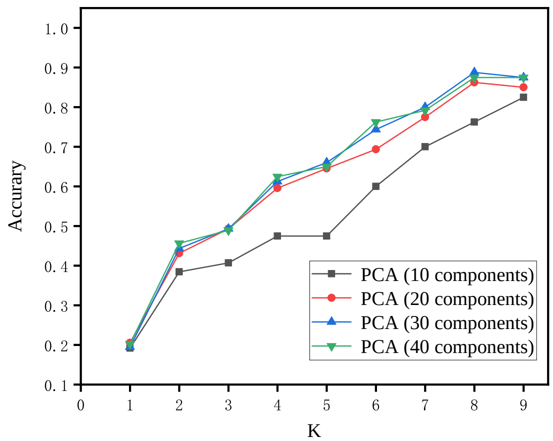


Fig. 2. Recognition Accuracy Comparison under Different PCA Principal Component Numbers

### *B. Impact of Classification Methods on Recognition Accuracy*

In this experiment, the PCA dimensionality is fixed at 30, and three classification methods—KNN, SVM, and LDA+KNN—are compared to investigate their impact on recognition accuracy. The experimental setup is as follows:

1) **Dimensionality Setting**: PCA is applied to reduce all face image data to 30 dimensions, retaining the most significant features.
2) **Classification Methods**:
   - **KNN**: A KNN classifier with $k = 3$ is used to classify the test samples.
   - **SVM**: A multi-class SVM classifier is trained using a linear kernel.
   - **LDA+KNN**: The PCA-reduced data is further reduced using LDA, and a KNN classifier is applied in the LDA subspace.
3) **Classification Process**: For each method, the classifier is trained on the PCA-reduced training set and tested on the PCA-reduced test set. Recognition accuracy is computed for each method.
4) **Results and Analysis**: The recognition accuracy of the three methods is compared and visualized through a bar chart, illustrating the performance differences among KNN, SVM, and LDA+KNN under the fixed PCA dimensionality condition. As shown in Table 2 and Figure 3, the recognition accuracy of facial images after PCA-based dimensionality reduction remains at a relatively high level across different recognition methods. Under optimal conditions, the accuracy can reach 90% or higher, demonstrating the robustness and effectiveness of PCA in preserving critical features for reliable face recognition performance.

TABLE II
RECOGNITION ACCURACY WITH DIFFERENT CLASSIFICATION METHODS

| **Training Pictures** | **Accuracy** | | |
|---|---|---|---|
| | **KNN** | **SVM** | **LDA+SVM** |
| 1 | 0.1944 | 0.2389 | 0.1306 |
| 2 | 0.4438 | 0.4875 | 0.6281 |
| 3 | 0.4929 | 0.6536 | 0.8 |
| 4 | 0.6125 | 0.6792 | 0.825 |
| 5 | 0.66 | 0.76 | 0.885 |
| 6 | 0.7438 | 0.8625 | 0.9125 |
| 7 | 0.8 | 0.8667 | 0.925 |
| 8 | 0.8875 | 0.9375 | 0.975 |
| 9 | 0.875 | 0.925 | 0.975 |

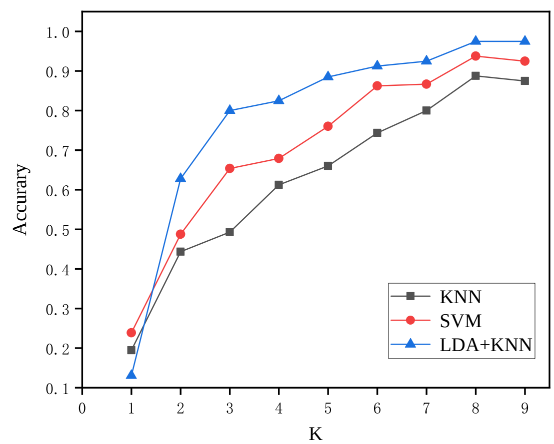


Fig. 3. Impact of Classification Methods on PCA-Reduced Face Recognition

## V. ANALYSIS OF EXPERIMENTAL

### *A. Impact of PCA Dimensionality Reduction on Storage Efficiency*

In this experiment, the Olivetti Faces dataset was used, which contains 400 facial images, each with a size of 64 × 64 pixels, resulting in 4096 dimensions for each image. The storage requirement for the original high-dimensional data is substantial, making direct processing both resource-intensive and computationally expensive. To address this, PCA was employed to reduce the dimensionality of each image from 4096 dimensions to 20 dimensions, significantly decreasing storage requirements. However, to support facial reconstruction, a mapping matrix from 20 dimensions to 4096 dimensions must also be stored. The detailed storage calculations are as follows.

After dimensionality reduction, each image is represented by only 20 dimensions, leading to a total storage requirement of $20 \times 400 = 8000$ dimensions for the entire dataset. Additionally, the PCA mapping matrix, which has a size of 4096 × 20, requires storage for $4096 \times 20 = 81,920$ dimensions. The total storage required after PCA is thus $8000 + 81,920 = 89,920$

dimensions. In comparison, the original dataset without dimensionality reduction requires $4096 \times 400 = 1,638,400$ dimensions. Therefore, the storage requirement after PCA is only about $89,920 \div 1,638,400 \approx 5.49\%$ of the original, effectively compressing the data to approximately 1/20 of its original size. This demonstrates that PCA dimensionality reduction can significantly optimize storage efficiency, reducing storage space by approximately 94.5%, while also simplifying subsequent computational tasks.

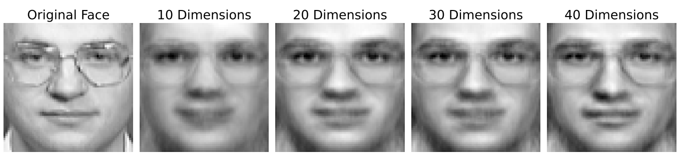


Fig. 4. Comparison of Original and PCA-Reconstructed Faces at Different Dimensionalities

### *B. Impact of PCA Dimensionality Reduction on Recognition Accuracy*

Although PCA significantly reduces data size and redundancy, experimental results show that the reduced-dimensional data retains the primary features necessary for effective facial recognition. In this experiment, three classification methods—KNN, SVM, and LDA+KNN—were used to evaluate the recognition performance on data reduced to 20 dimensions. The results indicate that KNN achieves approximately 80% recognition accuracy after dimensionality reduction, demonstrating its capability to distinguish facial classes effectively in a low-dimensional space. SVM, with its ability to construct more complex classification boundaries, further improves recognition accuracy to 85% 90%, showcasing superior performance in the reduced feature space. Moreover, the combination of LDA and KNN, where LDA extracts more discriminative features and KNN is used for classification, achieves recognition accuracy exceeding 90% under ideal conditions.

PCA excels in identifying and preserving the most representative features while discarding redundant and noisy information. Despite reducing the feature dimensions from 4096 to 20, a compression rate of over 99%, the retained features ensure that recognition accuracy is minimally affected. This highlights PCA's ability to balance storage efficiency and classification performance, making it a highly effective technique for facial recognition systems, especially in scenarios with limited storage and computational resources.

## CONCLUSION

In this study, Principal Component Analysis (PCA) was utilized to perform dimensionality reduction on facial image datasets, followed by experiments on face reconstruction and recognition. The results demonstrate that PCA can effectively reduce the redundancy in facial image storage, significantly lowering storage requirements. For example, when reducing the dimensionality to 20, 30, or 40 components, the storage demands were drastically reduced while retaining the critical feature information of the images. Comparative analyses revealed that even at lower dimensions, such as 20, the reconstructed facial images maintained high visual similarity to the original images.

In addition, the impact of PCA-based dimensionality reduction on facial recognition accuracy was evaluated. The results showed that when the dimensionality was reduced to 40 components, the recognition accuracy reached over 90%. At 30 and 20 components, the recognition accuracy remained above 88% and 85%, respectively. These findings indicate that PCA achieves a balance between reducing data dimensionality and maintaining high facial recognition performance, highlighting its practicality and reliability in high-dimensional data analysis and facial recognition tasks.

## REFERENCES

[1] M. Turk and A. Pentland, "Eigenfaces for Recognition," *Journal of Cognitive Neuroscience*, vol. 3, no. 1, pp. 71–86, 1991.
[2] M. Kirby and L. Sirovich, "Application of the Karhunen-Loève Procedure for the Characterization of Human Faces," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 12, no. 1, pp. 103–108, 1990.
[3] L. Sirovich and M. Kirby, "Low-Dimensional Procedure for the Characterization of Human Faces," *Journal of the Optical Society of America A*, vol. 4, no. 3, pp. 519–524, 1987.
[4] P. N. Belhumeur, J. P. Hespanha, and D. J. Kriegman, "Eigenfaces vs. Fisherfaces: Recognition Using Class Specific Linear Projection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, no. 7, pp. 711–720, 1997.
[5] W. Zhao, A. Krishnaswamy, and R. Chellappa, "Discriminant Analysis of Principal Components for Face Recognition," *Face Recognition: From Theory to Applications*, vol. 163, pp. 73–85, Springer, Berlin, Heidelberg, 1998.