Exploring Face Recognition Capabilities through Eigenfaces and Principal Component Analysis

Group Number: 2
Jingjuan Huang, Qiang Hu, Wang Ma, Linjie Tang

June 15, 2024

Exploring Face Recognition Capabilities through Eigenfaces and Principal Component Analysis

Jingjuan Huang, Qiang Hu, Wang Ma, Linjie Tang

June 15, 2024

Abstract

This project investigates the efficacy of face recognition techniques using Principal Component Analysis (PCA) and the Eigenface method. By transforming facial images into a reduced set of characteristic feature images (eigenfaces), we aim to identify or verify individual identities efficiently. We conduct extensive experiments on a real-world dataset, evaluate the performance of various classifiers, and provide a comprehensive analysis of the results. Our findings demonstrate high recognition accuracy, particularly when using Support Vector Machines (SVM) and k-Nearest Neighbors (KNN) as classifiers, showcasing the practicality and robustness of the Eigenface method.

1 Introduction

Face recognition technology has evolved significantly over the decades, transitioning from early geometric techniques to modern machine learning approaches. This study revisits one of the classical statistical approaches, the Eigenface method, which leverages Principal Component Analysis (PCA) for dimensionality reduction and feature extraction. We explore its application in face recognition, assess its performance across various dimensions and training sample sizes, and compare it with contemporary classifiers.

1.1 Background: Face Recognition

Face recognition aims to either identify or verify an individual's identity based on their facial features. The evolution of face recognition can be categorized into three main phases. The early geometric techniques of the 1990s, such as Elastic Bunch Graph Matching (EBGM) and Hausdorff Distance, laid the groundwork by focusing on spatial relationships and shape similarity. The classical statistical approaches of the 1990s introduced significant advancements with methods like Eigenfaces (PCA), Fisherfaces (LDA), and Support Vector Machines (SVM), which leveraged statistical models for more robust recognition. In the modern era, machine learning methods, including Convolutional Neural Networks (CNNs) from 2014, Metric Learning from 2015, and Generative Methods from the late 2010s, have revolutionized face recognition by enabling automated feature learning, improved metric spaces, and realistic face synthesis.

1.2 Literature Review

1.2.1 Early Geometric Techniques

Elastic Bunch Graph Matching (EBGM): Introduced in the 1990s, EBGM employs graph matching techniques for robust face identification. This method constructs a graph where nodes represent facial landmarks, and edges are labeled with jets, which are feature vectors derived from Gabor wavelet transformations. By leveraging the elastic properties of the graph, EBGM can effectively accommodate variations in facial expressions and orientations, making it a seminal technique in early face recognition research [WFKvdM97].

Hausdorff Distance: This technique measures the similarity between two point sets, which in the context of face recognition, typically represent the spatial distribution of facial features. The

Hausdorff distance is particularly valuable for its robustness to outliers and partial occlusions, offering a geometric perspective on the matching of facial structures [HKR93].

1.2.2 Classical Statistical Approaches

Eigenfaces (PCA): Proposed by Turk and Pentland in 1991, the Eigenfaces method applies Principal Component Analysis (PCA) to facial images, capturing the most significant features in a lower-dimensional subspace. This dimensionality reduction not only enhances computational efficiency but also highlights the intrinsic variance in facial data, facilitating effective recognition [TP91].

Fisherfaces (LDA): Building on the foundations of Linear Discriminant Analysis (LDA), Fisherfaces, developed in the mid-1990s, aim to maximize the ratio of between-class variance to within-class variance in facial data. This method improves discriminative power, particularly in scenarios with varying lighting conditions and facial expressions, by focusing on the most relevant features for classification [BHK97].

Support Vector Machines (SVM): Utilized in the late 1990s, SVMs have been pivotal in classification tasks within face recognition. By finding the optimal hyperplane that maximizes the margin between different facial classes, SVMs offer a robust framework for distinguishing between individual faces. Their effectiveness in high-dimensional spaces has made them a cornerstone in classical face recognition methods [OFG97].

1.2.3 Modern Machine Learning Methods

Convolutional Neural Networks (CNNs): Since their emergence in 2014, CNNs have revolutionized face recognition with their ability to automatically learn hierarchical feature representations from raw pixel data. Architectures such as VGG-Face, DeepFace, and FaceNet have demonstrated remarkable improvements in accuracy, leveraging large-scale datasets and deep learning techniques to achieve near-human performance in face recognition tasks [TYRW14, PVZ15, SKP15].

Metric Learning: Introduced around 2015, metric learning focuses on learning distance metrics that optimize the similarity between faces of the same identity and the dissimilarity between faces of different identities. Techniques like Triplet Loss and Contrastive Loss have been instrumental in fine-tuning the feature space, enhancing the discriminative capability of face recognition systems [SKP15, HA15].

Generative Methods: Emerging in the late 2010s, generative models, particularly Generative Adversarial Networks (GANs) and Variational Autoencoders (VAEs), have been employed for face synthesis and recognition. These models not only generate high-quality synthetic faces but also improve recognition performance by augmenting training datasets and learning robust feature representations [GPAM+14, KW13].

1.3 Eigenface Method

The Eigenface method represents a pivotal application of PCA in face recognition. This technique transforms high-dimensional facial images into a smaller set of characteristic features known as eigenfaces, which are derived from the principal components of the training images. The recognition process involves projecting a new image into the subspace defined by these eigenfaces and classifying it based on its coordinates in this reduced-dimensional space. This approach not only simplifies the computational complexity but also enhances the focus on the most significant facial features, thereby improving recognition performance while mitigating the impact of minor variations in the face and background.

1.4 Structure of the Report

This report is structured as follows:

- **Section 1: Introduction** - Provides a background introduction to face recognition technology, reviews relevant literature, and outlines the structure of the report.
- **Section 2: Principal Component Analysis (PCA) in Image Recognition** - Discusses the challenges of image recognition, explains the steps of PCA, and details how PCA is applied to face recognition.

- **Section 3: Empirical Analysis** - Describes the datasets used in this study, including the ORL database and custom datasets, and summarizes the key findings from the analysis.
- **Section 4: Experiments and Visualization** - Presents the experimental setup, explores the relationship between dimensionality and accuracy, examines different classifiers, and visualizes the results.
- **Section 5: Conclusion** - Summarizes the main findings, highlights the effectiveness of PCA and various classifiers in face recognition, and discusses the generalization of the model to real-world scenarios.

2 Principle Component Analysis (PCA) in Image Recognition

2.1 The Difficulties in Image Recognition

1. Face images are typically high-dimensional data. Assuming a standard grayscale face image of 64×64 pixels, this means each image has 4096 pixel (feature) points. If such high-dimensional data is used in a model, not only is the computational cost high, but it may also lead to overfitting.
2. Pixel points in face images often have strong correlations. For example, features such as eyes, noses, and mouths have similar positions and shapes across different face images, so identifying the main features of the image is an important step in image recognition.
3. In general, image photos have some small variations even consecutive images of the same person. The lighting conditions, facial expressions, and head positions may be different in some photos. It is important to ignore some of these minor changes, like slight facial expression variations, to improve the robustness of recognition.

2.2 Steps of PCA in Image Recognition

Principal Component Analysis (PCA) can effectively address the mentioned issues. It is highly effective in the field of image processing and recognition, especially when dealing with high-dimensional image data. The basic idea of PCA is to project high-dimensional data onto a lower-dimensional space through linear transformation for dimensionality reduction It can also preserve as much variance in the data as possible. Additionally, it can ignore some minor variations. Below is a detailed explanation of the steps of PCA in image recognition.

1. Image Data Representation: Image data can be represented as a matrix, where each element corresponds to a pixel's grayscale or color value. For grayscale images here, assuming the size is $m \times n$, it can be flattened into a one-dimensional vector of length $d = m \times n$.
2. Data Matrix Construction: Collect a set of images for training, flatten each image into a one-dimensional vector, and combine these vectors into a data matrix X, where each row represents an image. Assuming there are k images, the data matrix X has dimensions $k \times d$.
3. Mean Centering: Since all image pixel values are within the same range and the grayscale values are consistent (0 - 255), we only need to perform mean centering. First, calculate the mean vector of all image vectors $\mu = \frac{1}{N}\sum_{i=1}^{N} x_i$, where, x_i is the flattened vector of the i-th image. μ is also called the mean face. Subtract the mean vector μ from each image vector in the data matrix X to obtain the mean-centered matrix $A = X - \mu$

Figure 1: Mean Face

4. Compute the covariance matrix and perform eigenvalue decomposition: Calculate the covariance matrix C of the mean-centered matrix A:

$$C = \frac{1}{N} A A^T$$

Next, perform eigenvalue decomposition on the covariance matrix C to obtain the eigenvalues and corresponding eigenvectors. Let the eigenvalues of C be $\lambda_1 > \lambda_2 > \cdots > \lambda_d$ and the eigenvectors be $v_1, v_2, \ldots, v_d$, satisfying: $Cv_i = \lambda v_i$

5. Selecting Principal Components: Select the top r eigenvectors corresponding to the largest eigenvalues to form a projection matrix V_r. These eigenvectors represent the principal components, capturing the most significant variations in the data. The resulting projection matrix V_r represents the eigenfaces in the eigenface method.

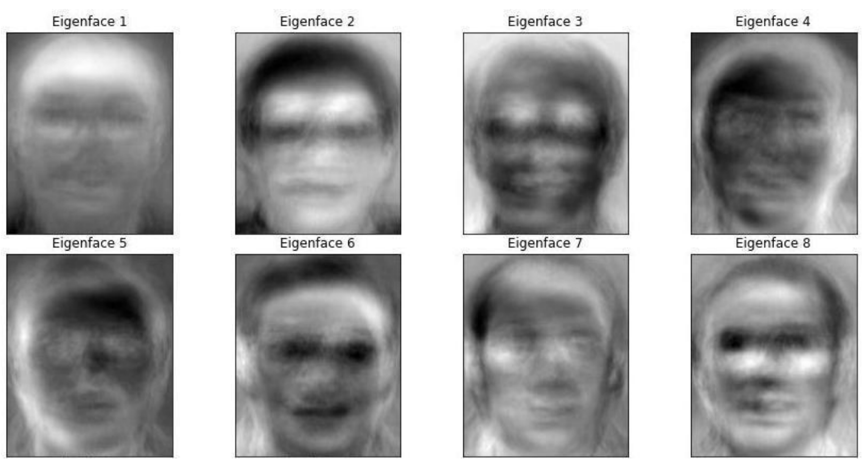


Figure 2: Eigenfaces

6. Projecting to Lower-Dimensional Space: Project the original image data onto the lower-dimensional space formed by the principal components:

$$Y = A * V_r$$

where Y is the representation in the lower-dimensional space.

7. Recognition and Classification: The low-dimensional feature vectors Y can be used to build classifiers for recognition.

Figure 3: Dimension Reduction to 20

3 Empirical analysis

3.1 Data description

1. Olivetti Research Laboratory (ORL) database of faces:

   Our work utilize the for experimental research. This database is a classic facial image dataset provided by the Olivetti Research Laboratory in Cambridge.

   The database contains 10 different images of 40 distinct subjects, totaling 400 images. These subjects are either employees of Olivetti or students at the University of Cambridge. The age range of the subjects spans from 18 to 81 years, with most subjects being between 20 and 35 years old. Among them, there are 4 female and 36 male subjects. Each image is 92×112 pixels in size and is a grayscale image. The images are taken at different times and under varying lighting conditions, but the background is always dark. The subjects face the camera, with only limited side movement and tilt allowed. The facial expressions in the images are diverse, including open eyes, closed eyes, smiling, and not smiling, with some individuals also wearing glasses.For a more intuitive display of the dataset, we present here a few sample images:

Figure 4: Sample Images

2. Custom data sets:

   These data sets include photos of several group members, ten of them each. We process it and convert it to a format consistent with the ORL data set so that the model can be trained more easily and better. Use cv2 to recognize the faces in the picture, rotate and expand to align, and finally crop and resize to the specified specifications (112*92)

3.2 Analysis steps and results summary

We conducted empirical analysis using the Olivetti Research Laboratory (ORL) database and a custom dataset to evaluate the performance of PCA-based face recognition.

For the ORL database, consisting of 400 grayscale images of 40 subjects, we explored the impact of dimensionality reduction on classification accuracy. We projected images onto 1, 5, 10, and 20 dimensions and varied the training sample size from 1 to 9. Our findings indicate that higher dimensions generally improve accuracy, with significant performance gains observed beyond 5 dimensions. Additionally, increasing the number of training samples consistently enhanced recognition accuracy across all dimensions except 1.

Furthermore, we experimented with various classifiers including Support Vector Machines (SVM) and k-Nearest Neighbors (KNN). SVM with a linear kernel and KNN with cosine distance metric outperformed other methods, achieving approximately 93% accuracy with optimal parameter settings.

In parallel, we extended our analysis to a custom dataset augmented with 20 additional photos of team members. Despite the small dataset augmentation, our model demonstrated robust performance, affirming its generalizability to new facial images.

Overall, PCA combined with discriminant analysis methods proved effective in facial recognition tasks, showcasing high accuracy and resilience across different datasets and experimental conditions.

4 Experiments and Visualization

This section presents a comprehensive exploration of the relationship between image classification accuracy and various factors including dimensionality reduction and training sample size. The experiments conducted provide insights into the effectiveness of different dimensionalities and sample sizes in accurately capturing and classifying image features.

4.1 Overall Exploration

We project the images onto 1, 5, 10, and 20 dimensions respectively, and train the model with training sample sizes k ranging from 1 to 9.

1. **Accuracy vs. Dimensionality**

   By investigating the relationship between image classification accuracy and the number of dimensions, we observe the following: Higher dimensionality leads to higher accuracy. Specifically, the accuracy is significantly low when the dimensionality is reduced to 1. However, when the dimensionality is increased to 5, it sufficiently captures the image characteristics. The accuracy obtained by projecting onto 10 and 20 dimensions is similar, indicating that extremely high dimensions are not necessarily required for effective feature extraction.

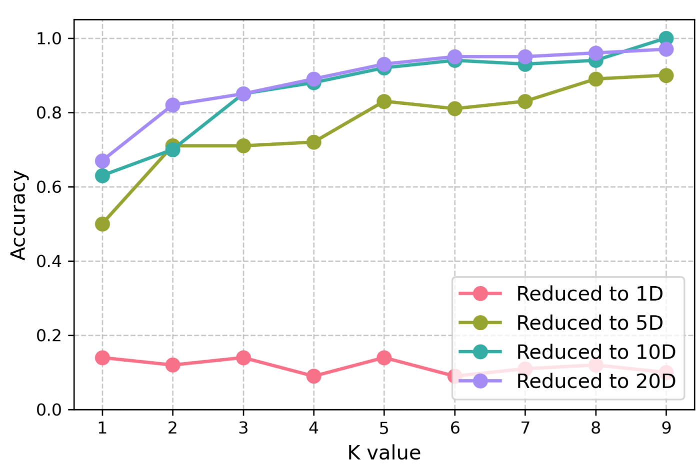


Figure 5: Comparison of Recognition Accuracy at All Dimensions

2. **Accuracy vs. Training Sample Size (k)**

   Combining the results from the previous four graphs, we analyze the variation in accuracy with respect to the training sample size. It is evident that as the number of training samples increases, the accuracy improves. This trend is not observed in the case of 1-dimensional projection due to insufficient feature extraction.

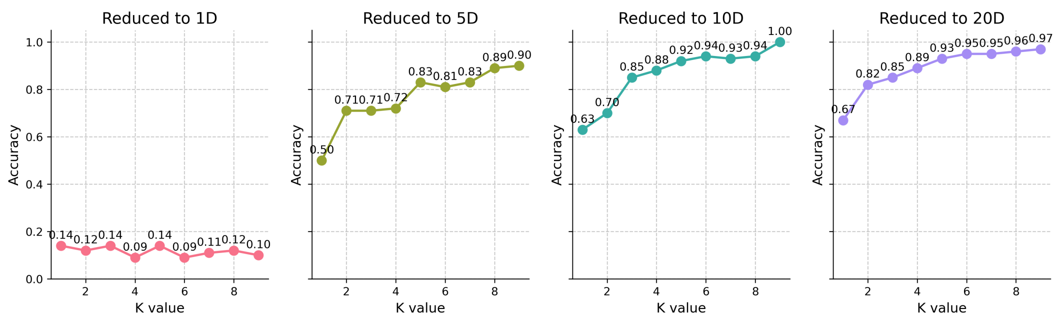


Figure 6: Comparison of Recognition Accuracy at Each Dimension

3. **Confusion Matrix for 10D Projection (k=7)**

   Next, we examine the confusion matrix for the 10-dimensional projection with 7 training samples. The high values along the diagonal and low values along the borders indicate that the model performs well in its predictions.

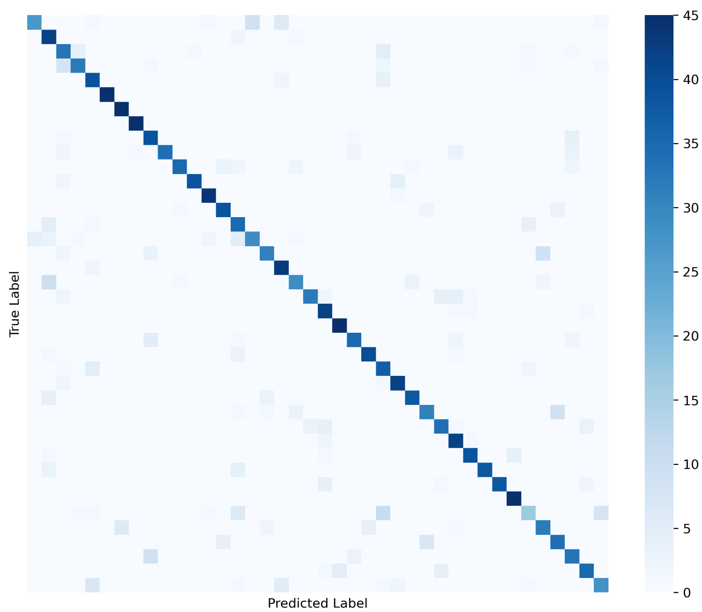


Figure 7: Confusion Matrix for 10D Projection (k=7)

4. **Eigenfaces for 10D Projection (k=7)**

   We then look at the eigenfaces for the 10-dimensional projection with 7 training samples. The first two images account for a significant proportion of variance, specifically 17.18% and 13.3% respectively. The top 10 features collectively explain over 60% of the variance.

(a) Eigenfaces

```
Explained variance ratios:
Feature 1: 17.18%
Feature 2: 13.30%
Feature 3: 6.77%
Feature 4: 5.90%
Feature 5: 5.20%
Feature 6: 3.35%
Feature 7: 2.63%
Feature 8: 2.29%
Feature 9: 2.04%
Feature 10: 1.69%
Total explained variance: 60.33 %
```

(b) Explained Variance

Figure 8: Eigenfaces and Explained Variance for 10D Projection (k=7)

5. **Reconstructed Images**

   Finally, we plot the reconstructed images obtained from the dimensional reduction. These images are linear combinations of the eigenfaces. When the dimensionality is reduced to 1, primarily the contour information is extracted. As the dimensionality increases, more granular details such as eyes, nose, and mouth features are progressively captured.

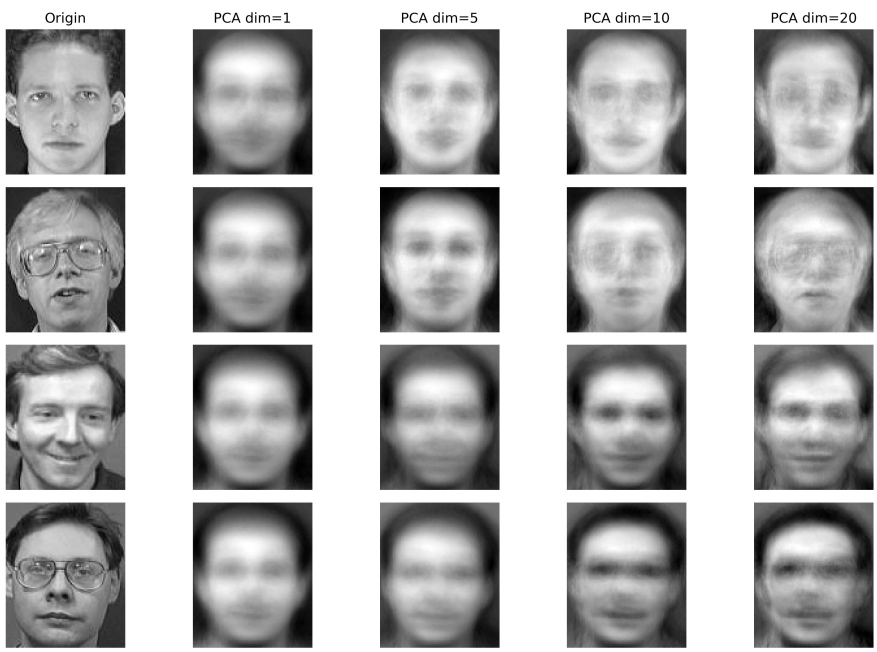
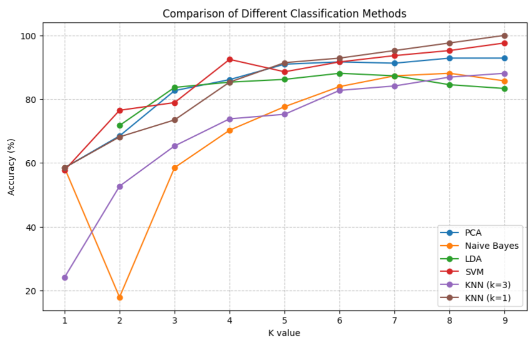


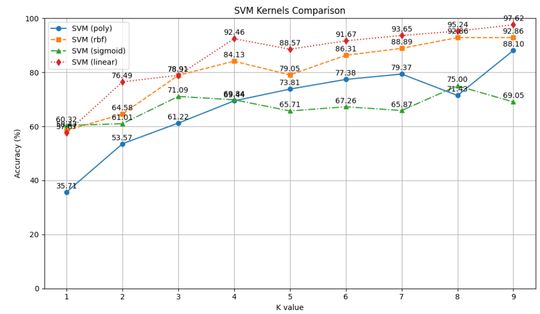
Figure 9: Reconstructed Images for Different Dimensions

4.2 Different discriminant analysis methods

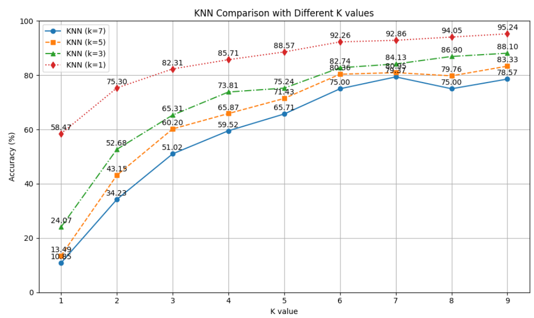
As previously mentioned, the face matching process consists of two main components: PCA (Principal Component Analysis) for dimensionality reduction of the original photos, and discriminant analysis for classifying the reduced-dimension vectors. For discriminant analysis, we experimented with distance discriminant, Naive Bayes, LDA (Linear Discriminant Analysis), SVM (Support Vector Machine), and KNN (k-Nearest Neighbors) with k=1 and k=3. Based on the classification performance curves across various training sample sizes (ranging from 1 to 9 photos), KNN (k=1) and SVM demonstrated the best performance. Specifically, KNN (k=1) outperformed slightly when the sample size was larger, whereas SVM showed a slight advantage with smaller sample sizes. Additionally, the trend of improved classification performance with increasing sample size, as noted earlier, holds true for all the discriminant methods evaluated in this study.



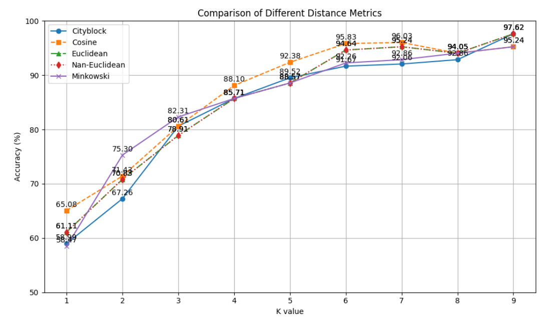
(a) different classifiers



(b) SVM with different kernels



(c) KNN with different Ks



(d) KNN with different distances

Figure 10: Experiments with different classifiers

We conducted further experiments on the SVM and KNN methods, which showed the best performance in our initial experiments:

4.2.1 kernel of SVM

For SVM, we investigated the impact of different kernel selections on classification performance. The results indicate that the linear kernel achieves the best performance, eliminating the need for more complex nonlinear kernel functions. Consistent with the previous experimental results, increasing the number of training samples improves classification performance for each kernel function.

4.2.2 K& Distance metrics for KNN

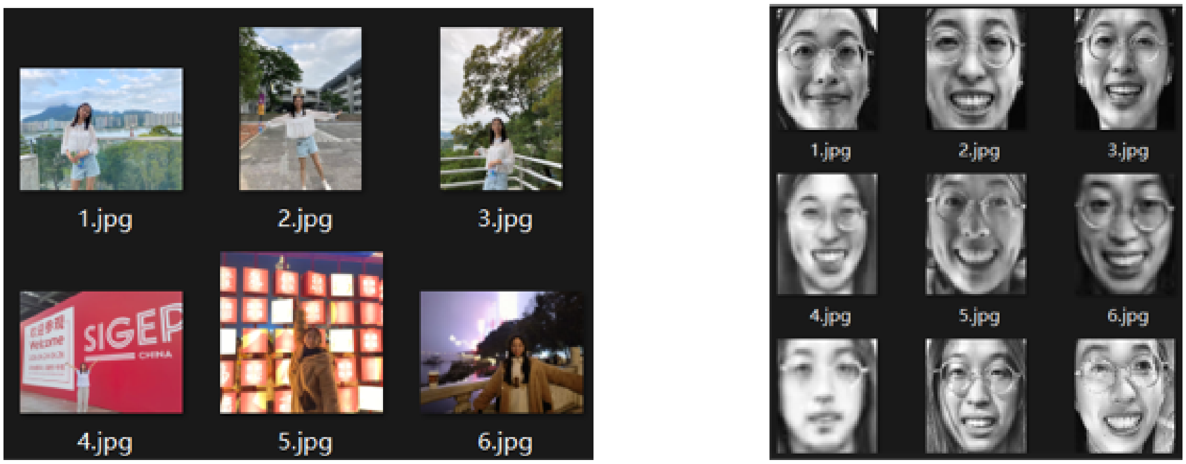
We further explored the KNN method by examining the effects of different values of k and various distance metrics on classification performance.

Selection of k: From the results, it is evident that the classification performance is best when $k = 1$. As k increases, the classification performance gradually deteriorates. This decline may be due to the introduction of noise, which interferes with discrimination. The superiority of $k = 1$ is more pronounced when the number of training samples is small, supporting the idea that larger k values introduce noise. For example, when each person has only one training sample, using $k > 1$ inevitably considers other people's faces in the classification process. Even if the test sample is very similar to that training sample, this inclusion of other faces disrupts the classification results.

Distance Metrics: Regarding distance metrics, the results indicate that when the sample size is sufficiently large (i.e., each person has at least four training photos), the cosine distance metric yields the best classification performance. This superior performance of cosine distance can be attributed to its focus on the direction of vectors rather than their magnitude. This mechanism effectively disregards variations in brightness and other distracting factors between photos. Generally, the directional features of vectors are more valuable than their magnitudes for facial recognition (i.e., the shape of the face is more important than overall skin tone).

4.3 Exploration on Custom Dataset

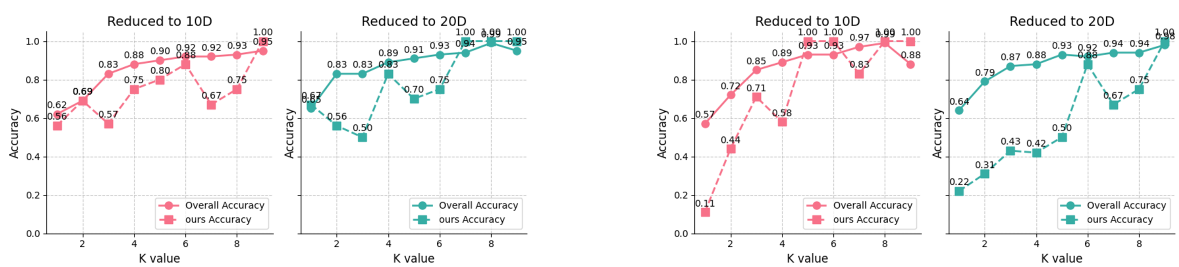
This subsection will be dedicated to examining the same factors using a custom dataset. In addition to experiments on the existing dataset, we also augmented the dataset with 20 photos of two team members. These photos were processed to match the format of the original dataset using various cv2 operations such as scaling, rotation, alignment, cropping, and adjustments to contrast and illumination.



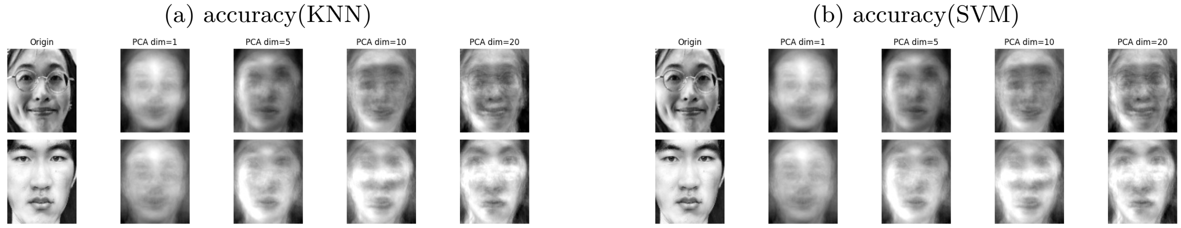
(a) Raw photos (b) Processed photos

Figure 11: Photo Preprocessing

We trained the model on the expanded dataset and evaluated the overall classification performance as well as the performance specifically on our own photos. The results showed that the classification performance remained consistent overall, with some minor fluctuations due to the small number of additional photos, which is unavoidable. This consistency in performance indicates that the model generalizes well.



(a) accuracy(KNN) (b) accuracy(SVM)



(c) reconstructed faces(KNN) (d) reconstructed faces(SVM)

Figure 12: Performance on our photos

5 Conclusion

In our study, we explored the efficacy of face matching using PCA for dimensionality reduction and various discriminant analysis methods for classification. Our key findings are summarized as follows:

- The first 10 principal components accounted for approximately 60% of the variance when $k = 7$.
- The classification accuracy consistently remained around 93%.
- The optimal performance was achieved using SVM with a linear kernel or KNN with cosine distance as classifiers.
- The model demonstrated excellent generalization to real-world scenarios, as evidenced by its consistent performance when applied to additional photos of our own faces.

These results highlight the robustness and effectiveness of using PCA combined with discriminant analysis methods for face matching tasks.

References

[BHK97] Peter N Belhumeur, Joao P Hespanha, and David J Kriegman. Eigenfaces vs. fisherfaces: Recognition using class specific linear projection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(7):711–720, 1997.

[GPAM$^+$14] Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative adversarial nets. In *Advances in Neural Information Processing Systems (NeurIPS)*, pages 2672–2680, 2014.

[HA15] Elad Hoffer and Nir Ailon. Deep metric learning using triplet network. In *Proceedings of the International Workshop on Similarity-Based Pattern Recognition*, pages 84–92. Springer, 2015.

[HKR93] Daniel P Huttenlocher, Gregory A Klanderman, and William J Rucklidge. Comparing images using the hausdorff distance. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 15(9):850–863, 1993.

[KW13] Diederik P Kingma and Max Welling. Auto-encoding variational bayes. In *Proceedings of the International Conference on Learning Representations (ICLR)*, 2013.

[OFG97] Edgar Osuna, Robert Freund, and Federico Girosi. Support vector machines: Training and applications. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 130–136. IEEE, 1997.

[PVZ15] Omkar M Parkhi, Andrea Vedaldi, and Andrew Zisserman. Deep face recognition. In *Proceedings of the British Machine Vision Conference (BMVC)*, volume 1, page 6, 2015.

[SKP15] Florian Schroff, Dmitry Kalenichenko, and James Philbin. Facenet: A unified embedding for face recognition and clustering. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 815–823. IEEE, 2015.

[TP91] Matthew A Turk and Alex P Pentland. Eigenfaces for recognition. *Journal of Cognitive Neuroscience*, 3(1):71–86, 1991.

[TYRW14] Yaniv Taigman, Ming Yang, Marc'Aurelio Ranzato, and Lior Wolf. Deepface: Closing the gap to human-level performance in face verification. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1701–1708. IEEE, 2014.

[WFKvdM97] Laurenz Wiskott, Jean-Marc Fellous, Norbert Krüger, and Christoph von der Malsburg. Face recognition by elastic bunch graph matching. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(7):775–779, 1997.