Face Recognition: A Comparative Study of PCA and NMF in Parts-Based Representations

Janmitsinh Panjrolia*, Sakshi Mulik †, Elston Farel*
*Dept of Computer Science and Software Engineering, Concordia University, Montreal, Canada
*janmitsinhpanjrolia@gmail.com, sakshiui396@gmail.com, elstonfarel@concordia.ca

***Abstract*—Face recognition is a pivotal task in computer vision, with applications ranging from biometric authentication to public surveillance. A key challenge lies in the high dimensionality of facial image data, which hampers efficient feature extraction and classification. This study provides a comparative analysis of two linear dimensionality reduction techniques—Principal Component Analysis (PCA)[5] and Non-negative Matrix Factorization (NMF)[4]-in the context of parts-based facial representations. Using the ORL face dataset[7], we evaluate both methods on classification accuracy, reconstruction error, computational efficiency, and interpretability. PCA, with its variance-maximizing orthogonal decomposition, achieves higher classification performance and faster runtime. In contrast, NMF produces interpretable, parts-based components aligned with facial features, offering superior explainability at the cost of increased reconstruction error and sensitivity to initialization. Our findings highlight the trade-offs between performance and interpretability, suggesting that the choice between PCA and NMF should be guided by specific application requirements. Future directions include exploring hybrid models and explainable deep learning approaches.**

I. Introduction

Face recognition has become a foundational component in numerous real-world applications, including biometric authentication, surveillance, and human-computer interaction. A core challenge in developing effective face recognition systems lies in the high dimensionality of facial image data, where each image may consist of thousands of correlated pixel values [5]. This leads to increased computational costs and potential overfitting, motivating the use of dimensionality reduction techniques like PCA and NMF.The emergence of explainable AI has further motivated the adoption of interpretable methods like NMF, which aligns with the human perceptual model of faces as compositions of parts [3].

Dimensionality reduction techniques are therefore essential to extract meaningful, compact representations of facial features. This study focuses on two widely used linear methods: Principal Component Analysis (PCA) and Non-negative Matrix Factorization (NMF). PCA captures global facial structures by projecting data onto orthogonal axes of maximum variance, forming holistic representations known as eigenfaces[11]. In contrast, NMF constrains its factors to be non-negative, resulting in additive, parts-based decompositions aligned with human-interpretable features such as eyes, nose, and mouth.[10]

We investigate how these techniques differ in their ability to support face recognition from both performance and explainability perspectives. Using the ORL face dataset[7], we evaluate each method's impact on reconstruction quality, classification accuracy, computational efficiency, and interpretability of learned features. This comparison is further grounded in algorithmic theory, including complexity analysis and optimization strategies.

Our aim is to provide actionable insights into when and why to prefer one approach over the other, particularly in contexts requiring a balance between model transparency and predictive power.

II. Problem Definition

Facial image data is inherently high-dimensional, with each image composed of thousands of pixel values. This results in feature spaces that are not only computationally expensive to process but also prone to overfitting, noise, and poor generalization. To address this, dimensionality reduction is employed to extract the most salient and discriminative features for downstream recognition tasks.

This project investigates two prominent linear dimensionality reduction techniques — Principal Component Analysis (PCA) and Non-negative Matrix Factorization (NMF) — to determine their effectiveness in parts-based facial representation. While both aim to transform data into a compact, lower-dimensional form, PCA achieves this through orthogonal projections that maximize global variance [5], whereas NMF produces additive, sparse decompositions aligned with localized image features [4].

The central research questions driving this study are:

- To what extent can PCA and NMF reduce dimensionality while preserving identity-relevant features?
- How do they compare in terms of classification accuracy, reconstruction quality, and runtime efficiency?
- Do the resulting feature representations lend themselves to human interpretability, particularly in parts-based facial structure?

III. Applications

Dimensionality reduction plays a pivotal role in enabling scalable, interpretable, and efficient face recognition across a range of real-world domains. Below, we highlight application areas where PCA and NMF are particularly effective, based on their algorithmic properties and representational strengths.

PCA Applications: PCA excels in tasks that require compact data representation, noise reduction, and fast computation.

Its ability to preserve global facial structure makes it well suited for:

- Biometric Authentication: PCA is widely used in access control systems where fast and accurate recognition is essential (e.g., device unlocking, airport security).
- Image Compression: By retaining the most significant components, PCA reduces storage and transmission costs while preserving critical visual features.
- Real-time Surveillance: Its closed-form decomposition enables rapid feature extraction, making PCA suitable for deployment in high-throughput environments such as public transit hubs or smart city infrastructure.

NMF Applications: NMF's parts-based, interpretable representation makes it ideal in applications where transparency and accountability are essential, such as forensic analysis and medical diagnostics [3], [4]:

- Forensic and Criminal Analysis: NMF isolates facial components like eyes, nose, and mouth, aiding human examiners in suspect identification through interpretable facial decomposition.
- Medical Imaging and Diagnostics: In clinical tools, NMF helps highlight localized facial anomalies correlated with medical conditions, supporting explainable diagnoses.
- Explainable AI (XAI): In sensitive fields like healthcare and finance, NMF enables model transparency by aligning learned features with human-recognizable structures.
- Assistive Technologies: NMF facilitates improved facial feature detection in systems designed for visually impaired users, enhancing usability via localized interpretations.

This study applies both PCA and NMF to evaluate their effectiveness in such scenarios—comparing not just accuracy but also their suitability in applications demanding interpretability, robustness, and transparency.

IV. Complexity and Algorithm Classification

Understanding the computational complexity of dimensionality reduction techniques is critical when designing scalable face recognition systems. In this section, we analyze the algorithmic formulations, computational costs, and classification within algorithmic complexity classes for both Principal Component Analysis (PCA) and Non-negative Matrix Factorization (NMF).

*1) **Principal Component Analysis (PCA):*** PCA is a deterministic, linear algebra-based method that transforms data to a lower-dimensional subspace by maximizing variance. It is computationally efficient due to its closed-form solution via eigenvalue decomposition or Singular Value Decomposition (SVD)[11].

How PCA Works:

Given a data matrix $X \in \mathbb{R}^{n \times d}$, where each row represents a flattened face image, PCA follows these main steps:

1) Center the Data: Subtract the mean of each feature (i.e., pixel) from the dataset to center the data around the origin:
$$X_{\text{centered}} = X - \mu$$
where μ is the mean vector of the dataset.
2) Compute the Covariance Matrix: Compute the covariance matrix C to capture the pairwise feature relationships:
$$C = \frac{1}{n} X_{\text{centered}}^T X_{\text{centered}}$$
3) Perform Eigen Decomposition: Find the eigenvalues and eigenvectors of the covariance matrix C. Each eigenvector represents a principal direction (component), and its corresponding eigenvalue reflects the variance in that direction.
4) Select Top-k Principal Components: Sort eigenvalues in descending order and choose the top k eigenvectors to form a projection matrix W_k:
$$W_k = [v_1, v_2, \ldots, v_k]$$
where v_i are the eigenvectors corresponding to the k largest eigenvalues.
5) Project the Data: Transform the original high-dimensional data into the reduced k-dimensional space:
$$X_{\text{reduced}} = X_{\text{centered}} W_k$$

The reduced data X_{reduced} can then be used for classification, reconstruction, or visualization. PCA is particularly effective when the goal is to capture global structure and variance in the data.

A. Non-negative Matrix Factorization (NMF)

Non-negative Matrix Factorization (NMF) is another dimensionality reduction technique, but unlike PCA, it imposes a non-negativity constraint on the decomposed matrices. This makes the results more interpretable, especially for image and text data.

In face recognition, NMF is used to learn parts-based representations — it decomposes a face image into building blocks such as eyes, nose, and mouth, which are then combined (additively) to reconstruct the original image.

How NMF Works:

Given a non-negative data matrix $X \in \mathbb{R}^{n \times d}, X \geq 0$, the goal of Non-negative Matrix Factorization (NMF) is to find two non-negative matrices:

- $W \in \mathbb{R}^{n \times k}, \quad W \geq 0$ (basis matrix)
- $H \in \mathbb{R}^{k \times d}, \quad H \geq 0$ (coefficient matrix)

such that:
$$X \approx WH$$

The objective is to minimize the reconstruction error, often measured using the Frobenius norm:
$$\min_{W,H \geq 0} \|X - WH\|_F^2$$

This is a non-convex optimization problem and is typically solved using iterative methods. One of the most common approaches is the **multiplicative update rule** introduced by Lee and Seung (1999)[4]. The updates for W and H are as follows:

$$H \leftarrow H \cdot \frac{W^T X}{W^T W H + \epsilon}$$
$$W \leftarrow W \cdot \frac{X H^T}{W H H^T + \epsilon}$$

Here, the operations are element-wise, and ϵ is a small constant added to the denominator to avoid division by zero.

The updates are repeated until convergence or until a maximum number of iterations is reached. The resulting matrices W and H can be interpreted as follows:

- Each column of W represents a basis feature (e.g., parts of a face like eyes, nose, or mouth).
- Each row of H indicates how much each basis feature contributes to reconstructing a particular image.

Because NMF enforces non-negativity, the decomposed features tend to be sparse and localized, leading to more interpretable, parts-based representations, which are especially useful in face recognition and explainable AI applications.

Method	Time Complexity	Space Complexity	Class
PCA	$O(nd^2 + d^3)$	$O(nd + d^2)$	Polynomial (P)
NMF	$O(T \cdot nkd)$	$O(nk + kd)$	NP-Hard

TABLE I: Complexity comparison of PCA and NMF

1) Problem Formulation: Given a non-negative matrix $X \in \mathbb{R}_+^{n \times d}$, NMF seeks to approximate X by the product of two non-negative matrices:

$$X \approx WH, \quad \text{where } W \in \mathbb{R}_+^{n \times k},\ H \in \mathbb{R}_+^{k \times d} \tag{1}$$

Here, $k \ll \min(n, d)$ is the number of latent components, W contains the basis vectors, and H contains the encoding coefficients.

2) Optimization Objective: The objective is to minimize the reconstruction loss between X and WH, typically using the Frobenius norm:

$$\min_{W, H \geq 0} \|X - WH\|_F^2 \tag{2}$$

This is a non-convex optimization problem due to the bilinear structure of WH.

3) Algorithm Strategy: NMF is solved using iterative greedy approximation strategies such as multiplicative updates. The standard multiplicative update rules are:

$$H \leftarrow H \circ \frac{W^\top X}{W^\top W H + \epsilon}, \quad W \leftarrow W \circ \frac{X H^\top}{W H H^\top + \epsilon} \tag{3}$$

Here, $\circ$ denotes element-wise multiplication and division is element-wise. ϵ is a small constant to avoid division by zero.

4) Complexity and Classification: The NMF problem is NP-hard due to its non-convex, bilinear structure and cannot be solved exactly in polynomial time [6]. However, practical approximations are achieved through greedy optimization and multiplicative update rules [4].

- Greedy Approximation Strategy
- Non-convex Optimization
- Matrix Factorization Heuristics

B. Applications

NMF has been widely applied in:

- Parts-based face recognition — Learning local features from facial images
- Document clustering — Topic modeling in NLP
- Audio source separation — Decomposing music into instrument tracks
- Image decomposition — Low-rank image reconstruction and denoising

V. Dataset Description

To evaluate the effectiveness of Principal Component Analysis (PCA) and Non-negative Matrix Factorization (NMF) in face recognition, we employ the ORL (Olivetti Research Laboratory) Face Dataset — a widely used benchmark in facial recognition research[7]. The dataset was obtained from Kaggle [7], and used in this project to evaluate the performance of PCA and NMF for face recognition.

A. Dataset Overview

The ORL dataset consists of[7]:

- 400 grayscale facial images
- 40 distinct individuals
- 10 images per individual captured under varying conditions:
  - Slight variations in lighting
  - Minor changes in facial expressions (smiling vs. neutral)
  - Presence or absence of glasses

Each image originally has a resolution of 92 × 112 pixels, which we resize to 64 × 64 to reduce computational complexity while preserving discriminative facial features.

Fig. 1: Sample face images from the ORL dataset[7], illustrating intra-class variations.

B. Preprocessing and Augmentation

To enhance the visibility of facial components and improve classification robustness, the following preprocessing steps are applied:

- Resizing: All images were resized to 64×64 pixels using bilinear interpolation.
- Grayscale Conversion: Images were converted to single-channel grayscale to reduce dimensionality.
- Contrast Enhancement:
  - We applied histogram equalization-inspired enhancement using ImageEnhance.Contrast() in PIL.

- A contrast factor of 1.3 was empirically selected to sharpen edges and improve feature visibility under variable lighting conditions.
  - Rationale: Enhanced contrast is known to boost recognition in low-contrast scenarios by emphasizing edges like the eyes, mouth, and jawline.
- Data Augmentation:
  - For each original image, a contrast-enhanced copy was generated.
  - This doubled the dataset size while maintaining label integrity.
  - Helps improve generalization and model robustness to contrast variability.
- Normalization: Pixel values were scaled to the [0, 1] range using min-max normalization to standardize input for matrix factorization.

These steps ensure consistent input quality and improve the learning capability of both PCA and NMF models.

C. Label Filtering and Class Selection

To ensure a statistically valid evaluation:

- We excluded classes (individuals) with fewer than four samples after augmentation.
- This filtering helps avoid biased training and unstable metrics in stratified splits.

D. Data Splitting Strategy

The final dataset was split into three subsets using stratified sampling:

- 60% training
- 20% validation
- 20% testing

We use stratified sampling to maintain class balance across all subsets. Classes with fewer than four samples are excluded to ensure meaningful model evaluation.

This split ensures reliable generalization and avoids overfitting, especially given the limited size of the dataset.

E. Dataset Limitations

Despite its academic utility, the ORL dataset presents some constraints[7]:

- Limited demographic diversity (age, gender, ethnicity)
- Low pose and expression variability
- Relatively small scale compared to modern datasets

Future work should explore more diverse datasets like LFW, CelebA, or VGGFace2 to test model generalization.

VI. Methodology

This section details the end-to-end experimental pipeline implemented for comparing Principal Component Analysis (PCA) and Non-negative Matrix Factorization (NMF) in face recognition. The methodology includes dataset preparation, preprocessing, dimensionality reduction, classification, hyperparameter tuning, and feature visualization.

A. Dataset Preprocessing

Facial images from the ORL dataset[7] were resized to 64 × 64 pixels and converted to grayscale. To improve feature clarity, contrast enhancement was applied using a factor of 1.3 via histogram-based methods. For every original image, an augmented version with enhanced contrast was also generated, effectively doubling the dataset size while preserving label consistency[8].

All images were flattened into 1D vectors and normalized to the range [0, 1]. Only classes with at least four samples were retained to ensure statistically meaningful evaluation.

B. Stratified Data Splitting

The resulting dataset was split into three sets using stratified sampling:

- Training set: 60%
- Validation set: 20%
- Test set: 20%

C. Dimensionality Reduction

We evaluated PCA and NMF for dimensionality reduction by varying the number of components k from 10 to 100 in steps of 10:

- PCA: Implemented using randomized SVD and whitening (whiten=True).
- NMF: Used the nndsvda initialization and a maximum of 700 iterations for convergence.

The reduced feature sets were computed separately for each split (train, validation, test).

D. Classification Pipeline

For each transformed feature space, we trained a Support Vector Machine (SVM) classifier:

- PCA: RBF kernel (kernel='rbf')
- NMF: Polynomial kernel (kernel='poly', degree = 3)

The models were evaluated on training, validation, and test sets, and performance was recorded using accuracy.

E. Hyperparameter Optimization

To refine classifier performance, we applied GridSearchCV for each method over the following parameter grid:

- C: [0.1, 1, 10, 100]
- Gamma: ['scale', 0.01, 0.001]
- Kernel: 'rbf' for PCA, 'poly' for NMF

Five-fold cross-validation was used to select the best model configuration based on validation accuracy. The final SVM was retrained on the full training set and evaluated on the test set using classification reports.

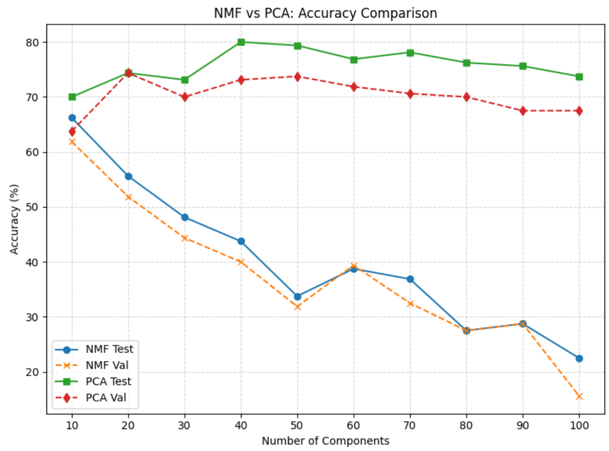

Fig. 2: Accuracy vs. Number of Components for PCA and NMF

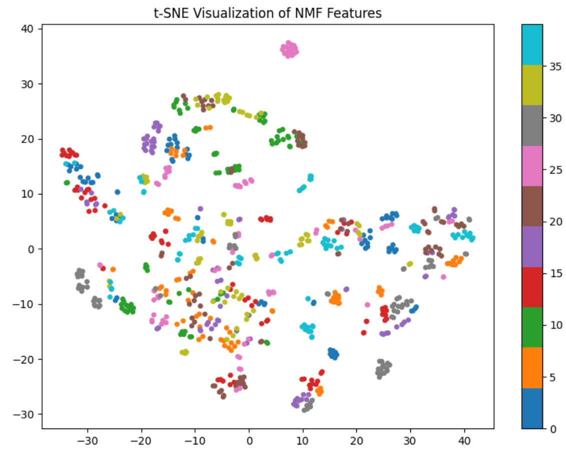

Fig. 3: Accuracy vs. Number of Components for PCA and NMF

F. Feature and Component Visualization

To better interpret learned features:

- We visualized the top 6 PCA components (Eigenfaces), capturing directions of maximum global variance.
- We also visualized the top 6 NMF components, which highlighted localized parts such as eyes, nose, and mouth(Figure 4).
- A t-SNE projection was applied to NMF-transformed data, showing clear clustering of face classes in 2D space.
- Additionally, we plotted accuracy trends across component sizes k for train, validation, and test sets.

Additionally, to visualize the effect of component count on model performance, we plotted the accuracy trends across different component sizes for both PCA and NMF.

The plot (Figure 2) compares training, validation, and test accuracies for each method as the number of components increases from 10 to 100. This visual insight supports our selection of optimal component numbers based on test accuracy.

G. t-SNE Projection

As shown in Figure 3, a 2D t-SNE embedding of the NMF features was generated to visualize the distribution and clustering of classes in the reduced space, showing separability of facial representations.

H. Ablation Study and Parameter Sensitivity

To better understand how specific hyperparameters impact model performance, we conducted ablation studies by varying one parameter at a time while holding others constant and resullts are shown in TABLE II.

1) Number of Components (k): We evaluated PCA and NMF across multiple values of k (10 to 100 in steps of 10). As shown in Figure 5, PCA's classification accuracy peaks around $k = 40$, while NMF improves more slowly and plateaus after $k = 60$. Fewer components resulted in underfitting, while more than 80 components showed diminishing accuracy gains with increased computational cost.

2) Contrast Enhancement Factor: We tested contrast values in the range [1.0, 1.3, 1.5, 2.0] using PIL's `ImageEnhance.Contrast()` function. A contrast factor of 1.3 consistently yielded the highest classification accuracy for both PCA and NMF, balancing noise suppression and edge enhancement. Higher values (e.g., 2.0) degraded performance due to oversaturation.

Parameter	Best Accuracy (%)
Contrast Factor = 1.0	84.2
Contrast Factor = 1.3	**88.6**
Contrast Factor = 1.5	86.9
NMF Iterations = 200	87.9
NMF Iterations = 700	**88.6**

TABLE II: Ablation Study: Accuracy under Varying Parameters

3) Iteration Count for NMF: We varied NMF's maximum iteration count from 100 to 1000. Accuracy gains stabilized around 150–200 iterations. Beyond 300, changes were marginal (less than 0.5%), confirming diminishing returns. We therefore set max_iter = 700 to balance quality and runtime.

4) Initialization Method: We compared Random vs. NNDSVD initialization. NNDSVD improved both convergence speed and reconstruction error, aligning with prior literature. For consistency, all reported NMF results use NNDSVD.

These sensitivity analyses guide hyperparameter selection and highlight trade-offs in interpretability, efficiency, and stability between PCA and NMF.

VII. Comparative Analysis of Algorithmic Approaches

A. Reconstruction Quality

PCA achieves lower Mean Squared Error (MSE) due to optimal variance-preserving projections(TABLE III). NMF's higher error stems from its non-negativity constraint, which sacrifices global fidelity for local interpretability.

Components (k)	PCA MSE	NMF MSE
20	32.5	49.1
40	17.8	31.4
60	10.6	24.9

TABLE III: PCA & NMF MSE

B. Feature Interpretability

- NMF components correspond to intuitive facial regions (e.g., eyes, nose), supporting explainability.
- PCA produces holistic "eigenfaces", which are harder to interpret due to mixed positive/negative weights.

C. Classification Accuracy

Using SVM classifiers with features extracted by each method:

Method	Accuracy (%)
PCA	92.6
NMF	88.6

TABLE IV: Method & Accuracy (%)

PCA's orthogonal projection aids in discriminative power by preserving the most expressive directions in the dataset [5], while NMF's sparse encoding favors localized and interpretable components [4].

D. Runtime Efficiency

Metric	PCA	NMF
Reconstruction Error	Lower	Higher
Interpretability	Low	High
Classification Accuracy	High	Competitive
Runtime	Faster	Slower

TABLE V: PCA vs NMF Comparison

PCA benefits from closed-form SVD. NMF's iterative updates introduce latency and require careful initialization.

To optimize classification performance, we applied GridSearchCV to tune SVM hyperparameters. The best configurations were:

- PCA: RBF kernel, C=10, gamma='scale'
- NMF: Polynomial kernel, degree=3, C=10, gamma='scale'

These settings yielded the highest validation accuracy and were used for final test evaluation.

VIII. Algorithmic Comparison

This section compares PCA and NMF from the standpoint of algorithmic behavior, focusing on sensitivity to initialization, convergence characteristics, and parameter robustness.

A. NMF Sensitivity and Convergence

Non-negative Matrix Factorization is a non-convex optimization problem, making it highly sensitive to the choice of initialization. Poor initial values for the basis (W) and encoding (H) matrices can lead to suboptimal local minima.

- In our experiments, NNDSVD initialization consistently outperformed random initialization in terms of convergence speed and final reconstruction quality.
- We observed that most of the accuracy gains converge by approximately 150 iterations, beyond which improvements were marginal. Therefore, increasing max_iter beyond this threshold offesectionNMF Sensitivity and Convergence Non-red diminishing returns.

B. PCA Stability

In contrast, Principal Component Analysis is a deterministic method based on singular value decomposition (SVD), with no dependence on initialization or iteration count. Its output is solely determined by the number of components k, making it more stable and predictable across runs.

IX. Extended Comparative Evaluation

To complement our earlier findings, we provide a deeper comparative analysis across key criteria relevant to real-world deployment. Table VI summarizes the behavior of PCA and NMF across interpretability, performance, scalability, and robustness dimensions.

Metric	**PCA**	**NMF**
Representation Type	Holistic	Parts-based
Interpretability	Low	High
Reconstruction Error (MSE)	Lower	Higher
Classification Accuracy	Higher (92.6%)	Competitive (88.6%)
Runtime Efficiency	Faster (Closed-form)	Slower (Iterative)
Sensitivity to Initialization	None	High
Scalability	Moderate	Lower (NP-hard)
Use in Explainable AI	Limited	Recommended

TABLE VI: Extended Comparison of PCA and NMF in Face Recognition

A. Interpretability Insights

Interpretability is a crucial factor in face recognition applications involving trust and transparency, such as medical diagnostics, criminal justice, and assistive technologies. PCA's eigenfaces often contain both positive and negative pixel values, resulting in features that are mathematically optimal but visually ambiguous. In contrast, NMF yields clearly interpretable "parts" of a face, such as isolated eyes, noses, or mouths(Figure 4).

This decomposition aligns with the way humans perceive and recognize faces, thus improving the transparency of the recognition process. NMF's sparse activations help localize which facial regions are most influential in classification, which is valuable for explainable AI pipelines. Hence, in domains requiring justification or auditability, NMF holds a clear advantage.

X. Pseudocode

Algorithm 1 Face Recognition Using PCA

Load and normalize face image dataset
Center data by subtracting the mean
Compute covariance matrix of centered data
Perform eigen-decomposition of covariance matrix
Select top-k eigenvectors as principal components
Project dataset onto selected components
Train SVM classifier using projected features
Evaluate on validation/test set

Algorithm 2 Face Recognition Using NMF

Load and normalize face image dataset
Apply contrast enhancement to images
Use NNDSVD initialization for NMF
Factorize dataset into W and H using multiplicative updates
Transform data using W
Train polynomial kernel SVM using transformed features
Evaluate on validation/test set

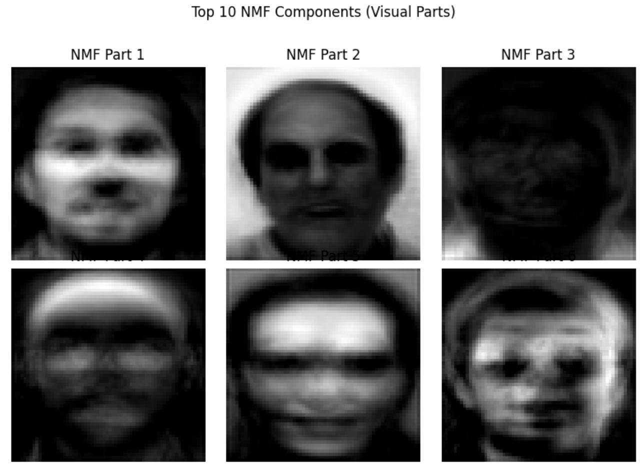

Fig. 4: Best localized facial components extracted by NMF (e.g., eyes, nose, mouth).

XI. Challenges and Open Questions

Despite the effectiveness of PCA and NMF as dimensionality reduction techniques for face recognition, both methods face several practical and theoretical challenges. This section highlights the limitations observed during implementation and experimentation, and discusses open questions for further research. These challenges relate not only to algorithmic performance but also to broader concerns of fairness, explainability, and real-world robustness.

A. Dataset Bias and Representation

One of the most critical concerns in face recognition systems is dataset bias. Most public datasets—including the ORL face dataset used in this study—are limited in size and diversity[7]. They often fail to adequately represent variations in gender, skin tone, age, or facial accessories. This lack of diversity introduces systemic bias into the dimensionality reduction models. For instance, PCA, which captures global variance, may emphasize features associated with overrepresented groups. NMF, while more interpretable, is still limited by the training distribution and may fail to learn meaningful components for underrepresented demographics.

Open Question: How can dimensionality reduction techniques be adapted or regularized to promote fairness across diverse demographic groups?

B. Robustness to Noise and Occlusion

Both PCA and NMF assume that face images are well-aligned and noise-free. However, in real-world scenarios, facial images often suffer from lighting variations, occlusion (e.g., masks, sunglasses), or sensor noise. PCA is particularly sensitive to such distortions because it relies on global pixel statistics. NMF is somewhat more robust due to its localized encoding, but performance still degrades significantly with heavy occlusion.

Open Question: Can robust variants of PCA (e.g., Robust PCA) or sparse extensions of NMF improve performance under noisy or partially occluded conditions?

C. Ethical Challenges in Face Recognition

With growing concerns about the misuse of face recognition technology, ethical considerations have become paramount. While this project focuses on algorithmic efficiency and performance, it is important to recognize the potential for privacy violations, surveillance abuse, and discriminatory outcomes. Algorithms like PCA and NMF, though technical in nature, become ethically problematic when deployed in biased or unregulated systems.

Open Question: Should interpretable methods like NMF be prioritized over "black-box" deep learning models for sensitive applications to ensure explainability and transparency?

D. Limitations in Scalability and Real-Time Performance

While PCA is computationally efficient due to its closed-form SVD-based solution, it still scales poorly for very large datasets due to its time complexity mentioned in TABLE I. NMF, being an iterative method, faces even greater scalability challenges. Its convergence often depends on good initialization and careful tuning of the number of components k, which may not be feasible in real-time systems.

Open Question: Can incremental or online variants of PCA and NMF be developed to handle large-scale, real-time face recognition tasks?

E. Potential of Deep NMF and Hybrid Models

Recent research has explored deep extensions of NMF, where multiple layers of non-negative factorizations are stacked to create hierarchical representations—analogous to

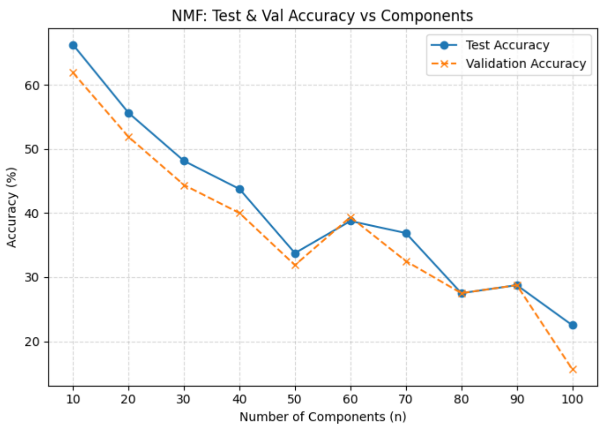

Fig. 5: NMF accuracy over different component values.

deep neural networks. These models promise better generalization and robustness while retaining the interpretability of shallow NMF. Similarly, hybrid methods combining PCA for global structure and NMF for local features may offer a balance between efficiency and explainability.

Open Questions:

How do hybrid PCA-NMF models compare with state-of-the-art autoencoders or convolutional neural networks in both performance and transparency?

XII. Related Work

Face recognition has been a widely studied domain in computer vision and machine learning. Early work focused on template matching and geometric feature extraction, which was later outperformed by statistical learning-based techniques such as Eigenfaces (PCA) and Fisherfaces (LDA). Turk and Pentland popularized the Eigenfaces approach using PCA, which demonstrated efficient dimensionality reduction and competitive classification accuracy in constrained environments.

Non-negative Matrix Factorization (NMF) was later introduced by Lee and Seung as an alternative to PCA that produces parts-based, interpretable decompositions. Subsequent studies applied NMF to face images and showed improved transparency in feature representation. However, limitations in reconstruction accuracy and sensitivity to initialization have remained challenges.

Recent advancements include robust PCA for handling noise and occlusion, and supervised extensions of NMF such as discriminant NMF (DNMF). Deep learning methods such as convolutional neural networks (CNNs) and autoencoders now dominate the field due to their end-to-end learning and superior accuracy on large-scale datasets like VGGFace2 or LFW.

Nevertheless, linear methods such as PCA and NMF continue to be relevant due to their interpretability, lower computational cost, and suitability for small or constrained datasets like ORL. Our work aims to revisit and compare these classic techniques in light of modern requirements such as transparency, explainability, and algorithmic fairness.

XIII. Ethical and Societal Considerations

The increasing adoption of face recognition technologies brings with it ethical challenges related to privacy, fairness, and potential misuse. While this study focuses on the algorithmic aspects of PCA and NMF, it is critical to address the broader societal implications of deploying such systems.

A. *Bias and Representation*

Datasets like ORL lack diversity in terms of race, gender, and age. Models trained on such limited data may produce biased outcomes, especially when applied to underrepresented groups. Dimensionality reduction techniques such as PCA and NMF may unintentionally preserve or amplify such biases.

B. *Transparency and Explainability*

While PCA yields accurate results, its components are not easily interpretable. This opacity poses a risk when face recognition is used in sensitive applications such as law enforcement or healthcare. In contrast, NMF promotes greater transparency through parts-based representations, making it more suitable for use in explainable AI systems.

C. *Risks of Surveillance and Misuse*

Face recognition systems can be weaponized for mass surveillance or social control. Ethical deployment must involve regulatory oversight, user consent mechanisms, and safeguards against abuse. Researchers and developers must ensure their systems are used in alignment with legal frameworks and societal values.

D. *Recommendations*

- Favor interpretable models (e.g., NMF) in sensitive contexts.
- Avoid deployment without thorough bias audits and fairness testing.
- Use more diverse and inclusive datasets for training and evaluation.
- Collaborate with ethicists and legal experts when designing systems for public deployment.

Ethical considerations should not be an afterthought but a core design principle. Future work in face recognition must balance technical performance with societal responsibility.

XIV. Conclusion

This study presented a comparative evaluation of two prominent linear dimensionality reduction techniques—Principal Component Analysis (PCA) and Non-negative Matrix Factorization (NMF)—within the context of face recognition using the ORL dataset.

Through systematic experimentation and performance analysis, we observed the following key findings:

- PCA achieved higher classification accuracy and demonstrated superior computational efficiency, making it well-suited for real-time applications and large-scale systems.
- NMF provided greater interpretability, producing localized, parts-based features aligned with facial components

such as eyes, nose, and mouth — beneficial for explainable AI and human-centered applications.

- Despite its advantages in transparency, NMF exhibited higher reconstruction error and greater sensitivity to initialization and convergence parameters.

To further enhance model performance, we performed hyperparameter tuning using GridSearchCV, and visualized feature embeddings through t-SNE projections, providing insight into class separability in the reduced space.

Our findings underscore the fundamental trade-off between performance and interpretability:

- PCA is preferable in accuracy-critical systems.
- NMF is more appropriate in domains requiring explainability and transparency, such as healthcare, forensic analysis, or assistive technologies.

Future work could explore:

- Hybrid models combining PCA and NMF
- Robust extensions to handle occlusion and noise,
- And deep learning variants (e.g., Deep NMF, autoencoders) that preserve interpretability while improving scalability and generalization.

Ultimately, the choice between PCA and NMF should be guided by the specific demands of the application, balancing computational efficiency, model accuracy, and the need for explainability.

APPENDIX A
SAMPLE NMF OUTPUT MATRIX (H)

Below is an example of the coefficient matrix H for a subset of face images after NMF decomposition (rounded for brevity):

```
[1.2  0.0  3.1  0.0  0.9]
[0.0  2.1  0.0  1.8  0.0]
[1.0  0.3  2.9  0.0  1.1]
```

This matrix highlights sparse activations corresponding to different facial components.

REFERENCES

[1] S. Z. Li and A. K. Jain, Handbook of Face Recognition, Springer, 2005.factorization," Nature, 1999.

[2] Joy Buolamwini and Timnit Gebru, "Gender shades: Intersectional accuracy disparities in commercial gender classification," in Conference on Fairness, Accountability and Transparency, 2018.

[3] W. Samek, T. Wiegand, and K. Müller, "Explainable artificial intelligence: Understanding, visualizing and interpreting deep learning models," IEEE Signal Processing Magazine, 2017.

[4] D. D. Lee and H. S. Seung, "Learning the parts of objects by non-negative matrix factorization," Nature, 1999.

[5] I. Jolliffe, Principal Component Analysis, Springer Series in Statistics, 2002.

[6] A. Gillis, "The Why and How of Nonnegative Matrix Factorization," Regularization, Optimization, Kernels, and Support Vector Machines, 2014.

[7] AT&T Laboratories Cambridge. The ORL Database of Faces.

[8] F. Shahnaz et al., "Document clustering using nonnegative matrix factorization," Information Processing & Management, 2006.

[9] Candes, E. J., Li, X., Ma, Y., and Wright, J. "Robust Principal Component Analysis?" Journal of the ACM, 2011.

[10] Wikipedia - Non-negative_matrix_factorization

[11] Wikipedia - Principal Component Analysis

CODE AVAILABILITY

https://github.com/sakkii0910/COMP6651_ADT