Bulletin of Electrical Engineering and Informatics
Vol. 15, No. 3, June 2026, pp. 2212∼2226
ISSN: 2302-9285, DOI: 10.11591/eei.v15i3.11042

Enhanced face recognition with nuclear norm-based angle 2D-PCA using QR decomposition

Jamal Elalji, Driss Gretete, Khalid Chougdali
Engineering Science Laboratory, National School of Applied Sciences, Ibn Tofail University, Kenitra, Morocco

Article Info

Article history:

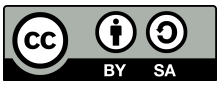
Received Jul 16, 2025
Revised Mar 14 2026
Accepted Mar 31, 2026

Keywords:

Dimension reduction
Face recognition
Nuclear norm
QR decomposition
Two-dimensional principal component analysis

ABSTRACT

Several approaches based on two-dimensional principal component analysis (2DPCA) have shown limitations in terms of classification performance. To enhance its robustness, an angular variant of 2DPCA has been proposed, establishing a relationship between reconstruction error and data variance through the Frobenius norm. However, this technique still encounters certain challenges. To overcome these shortcomings and further strengthen resilience to data variations, we propose a novel framework: nuclear norm-based angular 2DPCA using QR-decomposition (AN2DPCA-QR). This new formulation leverages the nuclear norm to optimize a variance-related criterion by maximizing the ratio of projected to original variance, aiming to improve the discriminative capacity of the projection space. The method employs a non-greedy iterative algorithm to solve the optimization problem, incorporating adaptive mean centralization for bias reduction, and QR decomposition instead of singular value decomposition (SVD) for numerical stability and reduced complexity. Compared to its predecessor, AN2DPCA-QR offers enhanced robustness, and interpretability. Results obtained on various public benchmark datasets clearly demonstrate the practical relevance and resilience of the proposed method.



Corresponding Author:

Jamal Elalji
Engineering Science Laboratory, National School of Applied Sciences, Ibn Tofail University
BP 241, 14 000 Kenitra, Morocco
Email: jamal.elalji@uit.ac.ma

1. INTRODUCTION

In various fields such as machine learning, computer vision, and pattern recognition, raw data often exhibits high dimensionality, noise, and outliers, making direct analysis difficult and computationally expensive. Feature extraction techniques address this challenge by transforming raw data into compact and informative representations while preserving essential information. This process is particularly important in face recognition tasks, where images must be analyzed under changes in lighting conditions, head orientation, facial expressions, and possible occlusions. Classical linear feature extraction techniques, such as principal component analysis (PCA) [1], linear discriminant analysis (LDA) [2], and independent component analysis (ICA) [3], are widely used for dimensionality reduction while maintaining the intrinsic structure of the data.

Classical PCA seeks projection directions that retain the largest variability in the data, thereby preserving the most significant structural characteristics while reducing information loss [1]. However, when PCA is applied to image data, the images must first be converted into vectors, which disrupts their inherent spatial organization and results in high-dimensional covariance matrices. To address this issue, Yang *et al*. [4]

introduced the two-dimensional PCA (2DPCA) method, which processes image matrices directly rather than transforming them into one-dimensional vectors. This strategy not only lowers the computational burden but also maintains important spatial relationships within the images. Nevertheless, conventional 2DPCA is still affected by noise, illumination variations, and the presence of outliers.

To enhance robustness, several variants of PCA and 2DPCA have been developed. Kwak [5] introduced L1-PCA, where the conventional L_2-norm is substituted with the L_1-norm in order to reduce the influence of outliers. Li *et al*. [6] later extended this concept to the two-dimensional case through the L1-2DPCA model.

Other works have explored different strategies to improve robustness. For example, Ding *et al*. [7] developed a rotationally invariant formulation of L1-PCA, while Nie *et al*. [8] proposed a non—greedy L1-RPCA method designed to mitigate the risk of convergence to local minima. Gao *et al*. [9] introduced a two-dimensional maximum local variation framework that highlights discriminative local structures. In addition, extensions based on L_p norms have been investigated [10], [11], providing a flexible compromise between robustness and numerical stability. However, their performance is highly dependent on the selection of the parameter p.

More recently, numerous enhanced versions of 2DPCA have been introduced to further improve robustness and recognition performance. Zhao *et al*. [12] provided a comprehensive review and unified several 2DPCA variants within a general ridge regression framework. In addition, Maafiri and Chougdali [13] developed a kernel-driven PCA approach integrated with RRQR decomposition to improve numerical stability. Huang *et al*. [14] presented Fp-2DPCA, which employs an F-norm criterion together with bilateral projections to enhance robustness against outliers. Wang *et al*. [15] proposed the DB2DPCA method, where bilateral projections derived from multiple PCA variants are fused to improve both reconstruction accuracy and recognition capability. Other related approaches include 2DPCA-Tℓ_1 [16], Cos-2DPCA [17], ROMCA-2DPCA [18], and 2DPCA-2-Lp [19], which attempt to increase robustness by adopting different optimization formulations.

Despite these developments, several limitations still persist. Many PCA-based models [20]–[23] remain vulnerable to severe outliers or structured disturbances such as occlusions. Moreover, certain robust formulations sacrifice rotational invariance, while others depend heavily on singular value decomposition (SVD), leading to increased computational cost and reduced scalability for large-scale datasets. In addition, the use of a fixed mean for data centering may introduce bias when outliers are present.

To address these limitations, we introduce a new method called nuclear norm-based angle 2DPCA with QR decomposition (AN2DPCA-QR). The proposed framework formulates an optimization objective that relies on the relationship between the variance of the projected data and that of the original observations, while incorporating nuclear norm regularization to enhance robustness. In addition, an adaptive mean estimation mechanism is employed to reduce the influence of outliers, and QR decomposition is adopted instead of SVD to decrease computational cost while preserving numerical stability. Table 1 provides a comparison of the main methods according to several properties relevant to facial recognition tasks. The main contributions of this work can be summarized as follows:

- A new variance-ratio formulation is introduced, which avoids the direct minimization of reconstruction error and leads to improved computational efficiency.
- Nuclear norm regularization together with an adaptive mean estimation strategy is incorporated to increase robustness against noise and outliers.
- A non-greedy iterative optimization algorithm is designed, accompanied by a theoretical study of its convergence behavior and rotational invariance properties.
- QR decomposition is adopted as an alternative to SVD in order to considerably reduce computational complexity.

Table 1. Comparison of PCA variants on key properties

Method	Norm	Model	Robust to outliers	Rot. invariant	Cnvg proof	Complexity
PCA [1]	L2	1D	Low	Yes	N/A	$O(Sl^2c^2 + (lc)^3)$
2DPCA [4]	Frobenius	2D	Medium	Yes	N/A	$O(Src^2 + c^3)$
LDA [2]	L2	1D	Low	No	N/A	$O(Sl^2c^2 + (lc)^3)$
2DPCA-L1 [6]	L1	2D	High	No	No	$O(iter \cdot Slc^2)$
N-2DPCA [24]	Nuclear	2D	High	Yes	Yes	$O(c^3)$
Angle-2DPCA [25]	Frobenius	2D	High	Yes	No	$O(iter \cdot Slc^2)$
Proposed model	Nuclear	2D	High	Yes	Yes	$O(iter \cdot lc^2)$

The remainder of this paper is structured as follows. Section 2 reviews related studies and presents the mathematical formulation of the proposed AN2DPCA-QR method. Section 3 describes the experimental evaluation and analyzes the comparative performance of the approach. Section 4 provides a discussion of the result. Finally, section 5 summarizes the key findings and discusses potential directions for future work.

2. METHOD

Given S image matrices $X_1, \ldots, X_S \in \mathbb{R}^{l \times c}$, we assume that the data are centered such that $C = \frac{1}{S}\sum_{s=1}^{S} X_s = 0$.

2.1. 2DPCA and 2DPCA-L1

2DPCA [4] generalizes PCA to operate directly on matrices, preserving structural information and reducing the dimensionality of covariance computations. The goal is to obtain low-dimensional representations $Y_s = X_s M$ by maximizing the total variance along the projected directions:

$$\arg\max_{M^\top M = I_d} \sum_{s=1}^{S} \|X_s M\|_F^2 = \arg\max_{M^\top M = I_d} \operatorname{tr} \sum_{s=1}^{S} M^\top X_s^\top X_s M, \tag{1}$$

where $M = [m_1, \ldots, m_d] \in \mathbb{R}^{c \times d}$ denotes a projection matrix with $d \ll c$. Here, $\operatorname{tr}(\cdot)$ refers to the trace operator, I_d denotes the $d \times d$ identity matrix, and $\|\cdot\|_F$ represents the Frobenius norm.

Solving this problem is equivalent to minimizing the sum of reconstruction errors across all samples: $\sum_{s=1}^{S} \|E_s\|_F^2 + \sum_{s=1}^{S} \|X_s M\|_F^2 = \sum_{s=1}^{S} \|X_s\|_F^2$, where $E_s = X_s - X_s M M^\top$ is the reconstruction residual for sample s. Consequently, one can also express the objective as (2):

$$\arg \min_{M^\top M = I_d} \sum_{s=1}^{S} \|E_s\|_F^2. \tag{2}$$

The optimal projection matrix M is formed by the eigenvectors associated with the d largest eigenvalues of $X_T = \sum_{s=1}^{S} X_s^\top X_s$. Although the Frobenius norm is commonly adopted, it is known to be sensitive to outliers. To improve robustness, the 2DPCA-L1 model [6] substitutes the Frobenius norm with the L_1-norm:

$$\arg \max_{M^\top M = I_d} \sum_{s=1}^{S} \|X_s M\|_{L_1}. \tag{3}$$

Although this approach mitigates the influence of extreme values, it introduces new limitations. The L_1-norm is not invariant to rotations, and maximizing the projected L_1-norm does not guarantee a minimal total reconstruction error: $\sum_{s=1}^{S} \|E_s\|_{L_1} + \sum_{s=1}^{S} \|X_s M\|_{L_1} \neq \sum_{s=1}^{S} \|X_s\|_{L_1}$. Directly minimizing the L_1-reconstruction error:

$$\arg \min_{M^\top M = I_d} \sum_{s=1}^{S} \|E_s\|_{L_1} \tag{4}$$

is computationally intensive. Hence, developing methods that are robust to outliers while effectively reducing reconstruction error remains an open research challenge.

2.2. Angle 2DPCA and N-2DPCA

N-2DPCA [24] introduces a subspace learning strategy that leverages the nuclear norm to assess the quality of data approximation. Specifically, it searches for a projection matrix that reduces the total nuclear norm of the residual matrices:

$$\arg \min_{M^\top M = I_d} \sum_{s=1}^{S} \|E_s\|_* , \tag{5}$$

where $\|\cdot\|_*$ denotes the nuclear norm. Using its representation in terms of a weighted Frobenius norm, the optimization problem can be equivalently expressed as (6):

$$\arg \min_{M^\top M=I_d} \sum_{s=1}^{S} \|Z_s E_s\|_F^2 , \tag{6}$$

with $Z_s = (E_s E_s^\top)^{-1/4}$. This formulation is solved using an iterative weighted least-squares procedure. Although Zhang *et al*. [24] did not explicitly mention it, N-2DPCA maintains rotational invariance due to the inherent properties of the nuclear norm. While this method effectively preserves the global structural information in image matrices, it incurs higher computational cost because of repeated SVD calculations. Angle 2DPCA [25] aims to produce a rotation-invariant low-dimensional representation by stabilizing the standard 2DPCA approach. The projection matrix M is obtained by minimizing the ratio between the reconstruction error and the projection magnitude:

$$\arg \min_{M^\top M=\mathbf{I}_d} \sum_{s=1}^{S} \frac{\|\mathbf{E}_s\|_F}{\|\mathbf{X}_s M\|_F} = \arg \min_{M^\top M=\mathbf{I}_d} \sum_{s=1}^{S} \frac{\left\|\mathbf{X}_s - \mathbf{X}_s M M^\top\right\|_F}{\|\mathbf{X}_s M\|_F}. \tag{7}$$

This approach simultaneously promotes larger projected variance and smaller reconstruction errors, enhancing robustness to noise and outliers. Nevertheless, several challenges remain, including the computational burden of the optimization, potential inaccuracies from using a fixed mean for centering, and the lack of a formal convergence guarantee in the original study. Motivated by these considerations, the present work proposes a new framework (see subsection 2.3) designed to boost the effectiveness of 2DPCA, 2DPCA-L1, N-2DPCA, and Angle-2DPCA for image recognition tasks intended to further enhance the overall performance of 2DPCA, 2DPCA-L1, N-2DPCA, and Angle 2DPCA in image recognition applications.

2.3. Nuclear norm-based angular 2DPCA using QR-decomposition

Angle-2DPCA incorporates the Frobenius norm as its distance measure, combining reconstruction error and projection magnitude within its objective function. Nevertheless, under the F-norm, maximizing the projection distance does not necessarily imply minimizing the cumulative reconstruction error, as illustrated by the inequality:

$$\sum_{s=1}^{S} \|E_s\|_F + \sum_{s=1}^{S} \|\mathbf{X}_s M\|_F \neq \sum_{s=1}^{S} \|X_s\|_F . \tag{8}$$

This discrepancy stems from the fact that the similarity measure operates at the matrix level rather than on individual vectors, making the method vulnerable to outliers distributed across matrix rows and columns. Furthermore, the objective function in Angle-2DPCA seeks to minimize the ratio between reconstruction error and projection distance, but the optimization is inherently constrained by the projection matrix M, which can reduce the method's robustness. To enhance robustness in facial image analysis, especially under challenging conditions such as occlusions or variable lighting,the nuclear norm has been adopted as an effective alternative to traditional distance measures. Unlike the L_2-norm, which can amplify the impact of outliers, the nuclear norm provides greater resilience by promoting low-rank approximations and capturing spatial dependencies through pixel correlations. Moreover, the nuclear norm acts as a convex surrogate for the rank operator, allowing for more stable optimization in noisy environments. This makes it particularly suitable for handling structured or non-Gaussian noise, where both the L_1- and L_2-norms often fail to accurately model reconstruction errors. Motivated by these advantages,we develop a novel nuclear norm–based model for face recognition, termed AN2DPCA-QR. In parallel, the AN2DPCA-QR model addresses bias in feature extraction by optimizing the mean calculation within the learning process itself. Rather than relying on a fixed, precomputed mean, typically affected by outliers induced by the use of the squared L_2-norm, AN2DPCA-QR incorporates this step into its optimization framework. This is achieved through a non-greedy iterative algorithm that yields an analytical solution at each step, thereby improving classification accuracy. The nuclear norm thus emerges as the most appropriate distance measure for this approach, and its application is defined as (9):

$$\underset{M^T M=I_d, C}{\arg\max} \left(\sum_{s=1}^{S} \frac{\|(X_s - C)M\|_*}{\|(X_s - C)\|_*} \right) \tag{9}$$

The choice of QR decomposition over SVD is justified by its lower complexity ($O(c^2 d)$ vs. $O(c^3)$) and numerical stability, especially for ill-conditioned matrices common in image data [13]. This reduces time

overhead in iterative optimization while maintaining accuracy. Table 1 summarizes key methods on properties relevant to face recognition. Before presenting the approach to derive the optimization algorithm, the following theorem provides a key result:

Theorem 1. [26] Let $X \in \mathbb{R}^{l\times c}$ be a matrix. The nuclear norm $\|X\|_*$ admits the following equivalent representation:

$$\|X\|_* = \|(XX^T)^{-\frac{1}{4}}X\|_F^2. \tag{10}$$

For a matrix X of rank r, the matrix power X^k is defined by $X^k = U\Sigma^k V^T$, where $\Sigma^k = \mathrm{diag}(\sigma_1^k, \ldots, \sigma_r^k)$. Here, $X = U\Sigma V^T$ denotes the SVD of X.

To solve for the optimal solution M and C in (9), we employ the above theorem to transform the nuclear norm optimization problem into an equivalent Frobenius norm problem:

$$\underset{M^TM=I_d,C}{\arg\max} \sum_{s=1}^{S} \frac{\|(X_s - C)M\|_*}{\|(X_s - C)\|_*} = \underset{M^TM=I_d,C}{\arg\max} \sum_{s=1}^{S} \frac{\|K_s(X_s - C)M\|_F^2}{\|(X_s - C)\|_*}, \tag{11}$$

where $K_s = \left[((X_s - C)M)((X_s - C)M)^T\right]^{-\frac{1}{4}}$. This can further be rewritten as (12):

$$\arg \max_{M^TM=I_d,C} \sum_{s=1}^{S} \mathrm{tr}\left((K_s(X_s - C)M)^T \Delta_s (K_s(X_s - C)M)\right), \tag{12}$$

with $\Delta_s = \frac{1}{\|(X_s - C)\|_*}$. In this context, the optimization problem, referred to as AN2DPCA-QR (nuclear norm-based angular 2DPCA using QR-decomposition), involves three interdependent unknown variables: M, C, and Δ_s, these variables are interrelated, making it infeasible to directly solve (12) in closed form. Their mutual dependence necessitates a non-greedy iterative algorithm to alternately update each variable until convergence. Below, we outline this procedure.

First, keeping M and Δ_s fixed, the optimal C is obtained by minimizing:

$$C = \arg\max_{C} \sum_{s=1}^{S} \mathrm{tr}\left((K_s(X_s - C)M)^T \Delta_s (K_s(X_s - C)M)\right). \tag{13}$$

$$= \arg\max_{C} tr\left(\sum_{s=1}^{S} (K_sCM)\Delta_s(K_sCM)^T\right) - 2tr\left(\sum_{s=1}^{S} (K_sX_sM)\Delta_s(K_sCM)^T\right). \tag{14}$$

Expanding the trace term and taking the derivative with respect to C:

$$\left(\sum_{s=1}^{S} K_sK_s^TC - K_sK_s^TX_s\right) MM^T\Delta_s = 0 \tag{15}$$

Assuming the projection matrix M satisfies the symmetry property, the optimal solution for C can be expressed as (16):

$$C = \left(\sum_{s=1}^{S} K_sK_s^T\right)^{-1} \left(\sum_{s=1}^{S} K_sK_s^TX_s\right). \tag{16}$$

Second, update M With C and Δ_s fixed, the problem (12) reduces to:

$$M = \arg \max_{M^TM=I_d} \mathrm{tr}\left(M^TH\right), \tag{17}$$

where $H = \sum_{s=1}^{S} (K_s\tilde{X}_s)^T\Delta_sK_s\tilde{X}_s$ and $\tilde{X}_s = X_s - C$. Now, to solve this sub problem, we introduce the following result:

Theorem 2. [27] Let $H \in \mathbb{R}^{c\times d}$ and assume that its QR factorization is written as $H = QR$, where $Q \in \mathbb{R}^{c\times d}$ has orthonormal columns and $R \in \mathbb{R}^{d\times d}$ is an upper triangular matrix. Let R admit the reduced SVD, defined as (18):

$$R = UDV^T, \tag{18}$$

where $U, V \in \mathbb{R}^{d\times d}$ are orthogonal matrices and $D \in \mathbb{R}^{d\times d}$ is a diagonal matrix whose entries correspond to the singular values of R. Under these conditions, the optimal solution of problem (17) can be written as (19):

$$M = QUV^T. \tag{19}$$

Proof. We aim to determine an orthogonal matrix M that maximizes : $M^* = \arg\max_{M^TM=I} \mathrm{tr}(M^TH)$. Consider the QR factorization of H, written as $H = QR$, where $Q \in \mathbb{R}^{c\times d}$ has orthonormal columns and $R \in \mathbb{R}^{d\times d}$ is an upper triangular matrix. Using this decomposition, the trace term becomes $\mathrm{tr}(M^TH) = \mathrm{tr}(M^TQR)$. Next, let the SVD of R be given by $R = UDV^T$, where U and V are orthogonal matrices and $D = \mathrm{diag}(\delta_1, \ldots, \delta_d)$ contains the singular values of R. Substituting this expression into the previous equation yields $\mathrm{tr}(M^TH) = \mathrm{tr}(M^TQUDV^T)$. By exploiting the cyclic invariance property of the trace operator, we obtain $\mathrm{tr}(M^TH) = \mathrm{tr}(V^TM^TQUD)$. Define the matrix $A = V^TM^TQU$. Since the matrices M, Q, U, and V are orthogonal, their product A is also orthogonal, implying $A^TA = I$. Consequently, the optimization problem becomes equivalent to maximizing $\mathrm{tr}(DA)$ over all orthogonal matrices A. Because D is diagonal with nonnegative singular values, the maximum value of $\mathrm{tr}(DA)$ is obtained when $A = I$. This condition implies $V^TM^TQU = I$.Rearranging the above relation gives $M^T = VU^TQ^T$. Taking the transpose of both sides, the optimal orthogonal matrix that maximizes the objective function is therefore: $M^* = QUV^T$.
Using the QR decomposition of H as $H = QR$, and letting $R = UDV^T$ (the SVD of R), the optimal M is:

$$M = QUV^T. \tag{20}$$

To avoid division by zero when $\|\tilde{X}_s\|_* = 0$, Δ_s is updated as (21):

$$\Delta_s = \begin{cases} \frac{1}{\|\tilde{X}_s\|_*} & \text{if } \|\tilde{X}_s\|_* \neq 0, \\ 0 & \text{otherwise.} \end{cases} \tag{21}$$

The above procedure is repeated iteratively until convergence, with updates to M, C, and Δ_s at each step. By leveraging QR decomposition instead of direct SVD, computational complexity is significantly reduced, making this approach efficient for large-scale applications. The overall complexity is $O(iter \cdot Slc^2)$, where iter is the number of iterations (typically < 20), lower than SVD-based methods due to QR's efficiency. Algorithm 1 provides a concise pseudo-code representation of the proposed procedure.

2.4. Convergence analysis

Theorem 3. The objective value in Algorithm 1 increases monotonically along the iterations. At iteration $(t + 1)$, the following inequality is obtained:

$$\sum_{s=1}^{S} \frac{\|\tilde{\mathbf{X}}_s M^{(t+1)}\|_*}{\|\tilde{\mathbf{X}}_s\|_*} \geq \sum_{s=1}^{S} \frac{\|\tilde{\mathbf{X}}_s M^{(t)}\|_*}{\|\tilde{\mathbf{X}}_s\|_*}. \tag{22}$$

Proof. From the update rule of Algorithm 1, the matrix $M^{(t+1)}$ is obtained by maximizing the linearized objective, which implies:

$$\mathrm{tr}\Big((M^{(t+1)})^TH^{(t)}\Big) \geq \mathrm{tr}\Big((M^{(t)})^TH^{(t)}\Big), \tag{23}$$

where:

$$H^{(t)} = \sum_{s=1}^{S} \frac{(K_s^{(t)}\tilde{X}_s)^T(K_s^{(t)}\tilde{X}_sM^{(t)})}{\|\tilde{\mathbf{X}}_s\|_*}. \tag{24}$$

Expanding the trace in terms of the rows $(K_s^{(t)}\tilde{X}_s)(j,:)$ gives:

$$\sum_{s=1}^{S}\sum_{j=1}^{l}\left[(K_s^{(t)}\tilde{X}_s)(j,:)M^{(t)}\right]\left[(K_s^{(t)}\tilde{X}_s)(j,:)M^{(t+1)}\right]^T \geq \sum_{s=1}^{S}\sum_{j=1}^{l}\|(K_s^{(t)}\tilde{X}_s)(j,:)M^{(t)}\|_2^2. \tag{25}$$

By the Cauchy Schwarz inequality, the inner product above is bounded by the product of the corresponding norms, which yields:

$$\|(K_s^{(t)}\tilde{X}_s)(j,:)M^{(t+1)}\|_2 \geq \|(K_s^{(t)}\tilde{X}_s)(j,:)M^{(t)}\|_2. \tag{26}$$

Summing over all rows and samples and using the definition of the nuclear norm leads to:

$$\sum_{s=1}^{S}\frac{\|\tilde{X}_sM^{(t+1)}\|_*}{\|\tilde{X}_s\|_*} \geq \sum_{s=1}^{S}\frac{\|\tilde{X}_sM^{(t)}\|_*}{\|\tilde{X}_s\|_*}. \tag{27}$$

Therefore, the objective function is non-decreasing throughout the iterations, which establishes the monotonic convergence of the algorithm.

Algorithm 1 Nuclear Norm-based Angular 2DPCA using QR-decomposition (AN2DPCA-QR)

Input: $X_s \in \mathbb{R}^{l\times c}$ $(s=1,...,S)$, desired dimension d.
Initialize: $M^{(0)} \in \mathbb{R}^{c\times d}, (M^{(0)})^TM^{(0)} = I_d, \Delta_s^{(0)} = 0, i = 0, C^{(0)} = \frac{1}{S}\sum_{s=1}^{S}X_s, K_s^{(0)}$.
while convergence criterion not satisfied **do**

1. Compute the mean matrix $C^{(t)}$ by solving: $C^{(t)} = \arg\min_C \left(\sum_{s=1}^{S}\frac{\|(X_s-C^{(t-1)})M^{(t)}\|_*}{\|X_s-C^{(t-1)}\|_*}\right)$. The optimal solution can be calculated by Equation (16),
2. For each X_s, calculate $K_s^{(t)}$ and $\Delta_s^{(t)}$ using the equations (11) and (21).
3. Update the projection matrix $M^{(t)}$ as the solution to:

$$M^{(t)} = \arg\min_{M^TM=I_d} \text{tr}(M^TH),$$

where $H = \sum_{s=1}^{S}\Delta_s^{(t)}(K_s^{(t)}X_s)^TM^{(t)}(K_s^{(t)}X_s)$. Using Theorem 2, the optimal solution is given by:

$$M^{(t)} = QUV^T,$$

4. Evaluate the objective function: $\Gamma(M^{(t)}) = \sum_{s=1}^{S}\frac{\|(X_s-C^{(t)})M^{(t)}\|_*}{\|X_s-C^{(t)}\|_*}$. If $\Gamma(M^{(t)}) \leqslant \Gamma(M^{(t-1)})$, proceed to the Step (6). Otherwise, go to next Step .
5. Use sub-gradient descent with Armijo line search to refine $M^{(t)}$, ensuring $\Gamma(M^{(t)}) \leqslant \Gamma(M^{(t-1)})$. If no such solution exists, terminate the loop. Otherwise, continue.
6. Update $\Delta_s^{(t)}$ for each s using the corresponding equation.
7. Increment t: $t \leftarrow t+1$.

end while
Output: $M^{(t)} \in \mathbb{R}^{c\times d}$.

2.5. Rotational invariance

Theorem 4. The AN2DPCA-QR algorithm is rotationally invariant. In particular, if the input matrices X_s undergo a rotation defined by an orthogonal matrix Ω satisfying $\Omega^T\Omega = I$, the value of the objective function remains unchanged.

Proof. Let $\Omega \in \mathbb{R}^{c\times c}$ be an orthogonal rotation matrix such that $\Omega^T\Omega = I$. Consider the rotated data matrices and the corresponding transformed projection matrix defined as (28):

$$\tilde{X}_s = X_s\Omega^T, \qquad \tilde{M} = \Omega M. \tag{28}$$

Substituting these expressions into the objective function yields:

$$\sum_{s=1}^{S}\frac{\|X_sM\|_*}{\|X_s\|_*} = \sum_{s=1}^{S}\frac{\|X_s(\Omega^T\Omega)M\|_*}{\|X_s\Omega^T\|_*}. \tag{29}$$

Since Ω is orthogonal, multiplication by Ω or Ω^T does not modify the nuclear norm. Consequently,

$$\sum_{s=1}^{S} \frac{\|X_s M\|_*}{\|X_s\|_*} = \sum_{s=1}^{S} \frac{\|\tilde{X}_s \tilde{M}\|_*}{\|\tilde{X}_s\|_*}. \tag{30}$$

Therefore, the value of the objective function is preserved under the rotation Ω. Moreover, if M is an optimal projection matrix for the original problem, then $\tilde{M} = \Omega M$ provides the corresponding solution for the rotated data. The resulting low-dimensional representation satisfies:

$$\tilde{X}_s \tilde{M} = X_s \Omega^T \Omega M = X_s M, \tag{31}$$

which shows that the extracted features remain unchanged. Hence, the proposed AN2DPCA-QR method possesses rotational invariance.

3. RESULTS

This section investigates the classification performance of the proposed AN2DPCA-QR algorithm through a comparative study with six representative unsupervised dimensionality reduction techniques, namely PCA, PCA-L1, Angle-PCA, 2DPCA, 2DPCA-L1, and N-2DPCA. The evaluation is conducted on both synthetic data and two widely used benchmark face recognition datasets: the CMU PIE and AR databases. All simulations are implemented in MATLAB (R2021a) on a computer equipped with an Intel Core i7-1365U processor operating at 1.80 GHz and 16 GB of RAM. For classification, a 1-nearest neighbor (1NN) rule is adopted, while the dimensionality of the projection subspace is systematically varied from 1 to 30 components. The 1NN classifier is chosen because it is simple, does not require a training phase, and directly reflects the discriminative ability of the projection method in the reduced subspace. Unlike more complex classifiers such as SVMs, which introduce additional parameters and training complexity, 1NN provides a neutral and standard benchmark that allows fair comparison across different dimensionality reduction methods. Results include mean accuracy ± SD over 10 runs, execution times, and normalized confusion matrices for further insight.

3.1. Synthetic data experiments

Synthetic datasets are first employed to assess the effectiveness of the proposed AN2DPCA-QR method. The aim is to examine the classification behavior of several dimensionality reduction approaches, including AN2DPCA-QR, when the data are corrupted by Gaussian noise. The synthetic data consists of 20 classes of matrices, denoted by $X_{i,j}$, where i represents the class index and j the matrix number within each class. Each class contains 10 matrices, and 20% of the matrices are corrupted with additive Gaussian noise. To generate the data, we defined $B_{i,j} \in \mathbb{R}^{m \times n}$ and $A_{i,j} \in \mathbb{R}^{d \times m}$ as element-wise i.i.d. Gaussian $N(0,1)$ matrices, such that $X_{i,j} = A_{i,j} B_{i,j} \in \mathbb{R}^{d \times n}$. The matrix dimensions were set to $n = 32$, $d = 56$, and $m = 52$. To simulate noise, 20% of the matrices in each class were corrupted with additive Gaussian noise $N(0,5)$. The classification performance of various dimensionality reduction methods, including PCA, PCA-L1, 2DPCA, 2DPCA-L1, N-2PCA and AN2DPCA-QR, was evaluated on this dataset, with results averaged over 10 independent runs. Furthermore, we evaluated the performance under two distinct initialization approaches: random subspace initialization and PCA for 1D methods (Angle-PCA and PCA-L1) and 2DPCA for 2D methods (2DPCA-L1, N-2PCA, and AN2DPCA-QR). The initial experiment (random subspace initialization), summarized in Table 2, revealed that AN2DPCA-QR achieved the highest classification accuracy (41.00% ± 0.0083), significantly outperforming PCA, PCA-L1, and 2DPCA, which showed accuracies of 32.67% ± 0.0119, 34.00% ± 0.0099, and 33.67% ± 0.0352, respectively.

Table 2. Average recognition accuracy and associated standard deviation on synthetic data

Method	Average accuracy ±SD
PCA	0.3267±0.0119
PCA-L1	0.3400±0.0099
2DPCA	0.3367±0.0352
2DPCA-L1	0.3300±0.0284
N-2DPCA	0.3301±0.0287
AN2DPCA-QR	**0.4100±0.0083**

In a subsequent experiment (PCA for 1D methods and 2DPCA for 2D methods), we included Angle-PCA and tested different initialization methods, as shown in Table 3. AN2DPCA-QR remains the best-performing method, showing a slight improvement over the initial experiment by achieving 44.50%±0.0276, which further highlights its robustness in noisy environments. 2DPCA, N-2DPCA, and 2DPCA-L1 yield comparable results to those of the first experiment but remain less effective than angle-PCA and AN2DPCA-QR.

Table 3. Average recognition accuracy and standard deviation on synthetic data, including angle-PCA

Method	Average accuracy ±SD
PCA	0.3250±0.0119
PCA-L1	0.3400±0.0099
Angle-PCA	0.4333±0.0207
2DPCA	0.3367±0.0352
2DPCA-L1	0.3300±0.0284
N-2DPCA	0.3301±0.0287
AN2DPCA-QR	**0.4450±0.0276**

The results confirm that the classification methods are generally insensitive to different initialization approaches. Among these, AN2DPCA-QR consistently demonstrates superior performance across both experiments, showcasing its robustness to noise and achieving the highest classification accuracy.

3.2. The public face databases experiment

3.2.1. The CMU PIE C29 subset dataset

The CMU PIE face database [28] contains images of 68 individuals, with a total of 41,368 facial samples collected under different conditions such as pose, illumination, and expression variations. In this study, the C29 subset is used, which includes 1,632 images corresponding to the 68 subjects (24 images per subject). This subset exhibits variations in pose, facial expressions, and lighting conditions. All images were automatically aligned according to eye locations and then cropped to a resolution of 32×32 pixels. Two experimental scenarios were considered. The first uses the original images without corruption, while the second introduces synthetic noise by adding block noise and dot artifacts of at least 20×20 pixels to 20% of the 1,632 images at random positions. Examples of clean and corrupted images are illustrated in Figure 1. Figure 2 illustrates the algorithm pipeline.

Figure 1. Examples of clean and corrupted images from the CMU PIE database

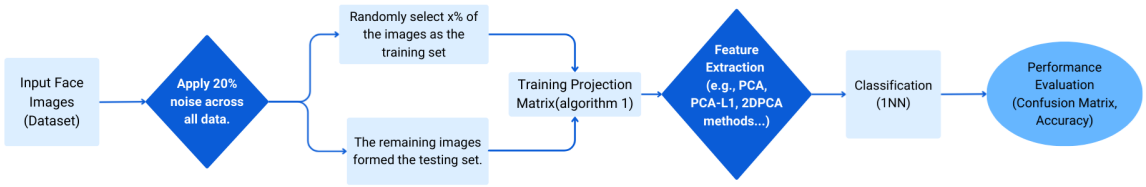


Figure 2. The flowchart of experiment 2

The proposed AN2DPCA-QR approach was evaluated against six representative methods—PCA, PCA-L1, Angle-PCA, 2DPCA, 2DPCA-L1, and N-2DPCA using a nearest neighbor classifier. The evaluation protocol randomly selected 10% of the images for testing and repeated the experiment ten times. Recognition accuracy was averaged over feature dimensions ranging from 1 to 30 (Figures 3(a) and (b)). On noise-free data, AN2DPCA-QR achieves the best performance with an accuracy of about 94%, clearly outperforming PCA, PCA-L1, and Angle-PCA (approximately 88.71%), while remaining competitive with 2DPCA-based approaches (Figure 3(a)). Under noisy conditions, the proposed method maintains the highest accuracy (around

80%), whereas PCA, Angle-PCA, and PCA-L1 drop significantly (55–60%), and the remaining 2DPCA variants show lower robustness (Figure 3(b)).

The boxplots in Figure 4, obtained from 10 independent runs, further highlight the superior accuracy and stability of AN2DPCA-QR compared with PCA-based techniques, while confirming the strong performance of 2D-based models, particularly in the presence of noise.

Table 4 reports the average execution times on the CMU PIE dataset. The comparison focuses on N-2DPCA because both methods rely on the nuclear norm, allowing the impact of the QR-based optimization to be clearly observed. Unlike N-2DPCA, which requires SVD, the proposed method employs QR decomposition. As a result, AN2DPCA-QR achieves significantly lower computational times (20.04 s for clean data and 21.69 s for noisy data), corresponding to nearly a 60% reduction compared with the 54.10 s and 54.37 s required by N-2DPCA. Overall, these findings highlight the computational efficiency and stability of the introduced method across different experimental conditions.

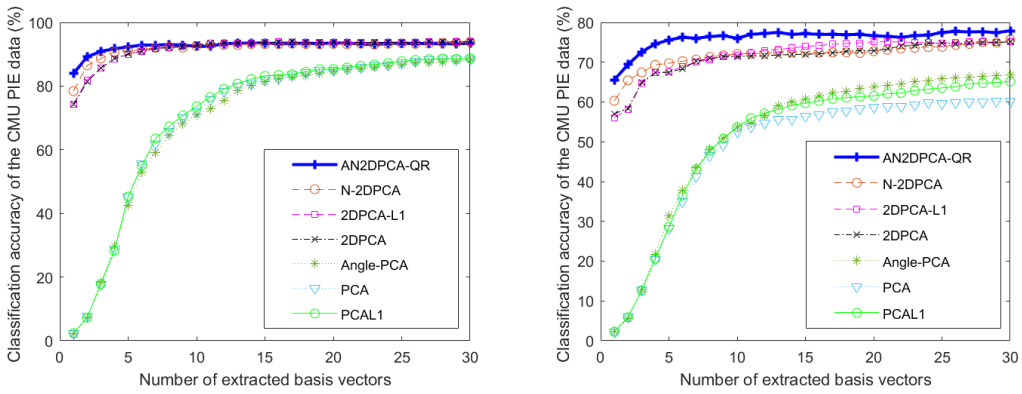


Figure 3. Recognition accuracy versus dimensional variation on the CMU PIE database; (a) the CMU PIE dataset without noise and (b) the noisy CMU PIE dataset

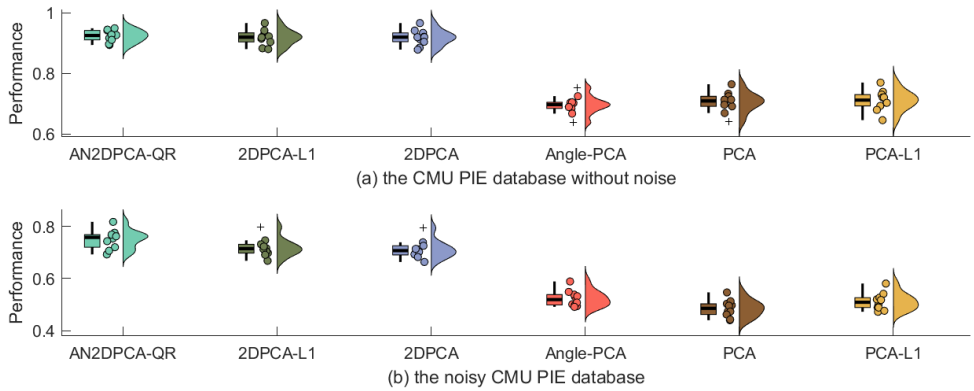


Figure 4. Boxplots of recognition accuracy on the CMU PIE database over 10 runs; (a) the CMU PIE dataset without noise and (b) the noisy CMU PIE dataset

Table 4. Average execution times on CMU PIE (in seconds) of N-2DPCA and AN2DPCA-QR

Method	Without noise	Noisy CMU PIE
N-2DPCA	54.1041	54.3690
AN2DPCA-QR	**20.0413**	**21.6907**

3.2.2. The AR face dataset

The AR face database [29] contains facial images of 120 individuals collected during two distinct sessions. In each session, 13 color images are recorded for every subject, including 6 images affected by occlusions (e.g., sunglasses or scarves) and 7 unobstructed facial images displaying variations in facial expression and illumination. Representative examples from the dataset are shown in Figure 5. For this study, each facial region was manually delineated and normalized to dimensions of 50×40 pixels. Two experimental setups were implemented: the first utilized the original, noise-free images, while the second introduced artificial corruption by adding block noise and dot patterns (minimum size 20×20 pixels) to 20% of the total 3,120 images. In both cases, following standard protocol, 10% of the images from each individual were randomly chosen for testing, with the remaining images used for training. This procedure was repeated ten times, and the mean classification accuracy across runs was computed. Finally, the classification results were averaged over feature dimensions ranging from 1 to 30, and the outcomes are depicted in Figures 6(a) and (b). AN2DPCA-QR was compared with six methods (2DPCA-L1, N-2DPCA, 2DPCA, Angle-PCA, PCA, and PCA-L1) for feature extraction on the AR database, using a nearest neighbor classifier. On noise-free data, AN2DPCA-QR stands out with an accuracy of approximately 90% starting from 15 basis vectors, stabilizing thereafter, and outperforms all other models (Figure 6(a)). In noisy conditions, it maintains the lead with 80% accuracy, showing rapid gains in the last 15-30 vectors, surpassing 2DPCA-L1, N-2DPCA, and 2DPCA (70-75%), as well as Angle-PCA and PCA/PCA-L1 (50-55%) (Figure 6(b)). The boxplots in Figure 7, which illustrate the recognition accuracy on the AR dataset over 10 runs—(a) without noise and (b) with noise—together with the execution times reported in Table 5, further confirm the superior accuracy, stability, and computational efficiency of AN2DPCA-QR. Specifically, AN2DPCA-QR records execution times of 60.17 s (clean) and 67.68 s (noisy), approximately 50% faster than N-2DPCA's 120.50 s and 143.72 s, reinforcing its robustness and computational advantage in challenging conditions.

Figure 5. Examples of clean and corrupted images from the AR face dataset

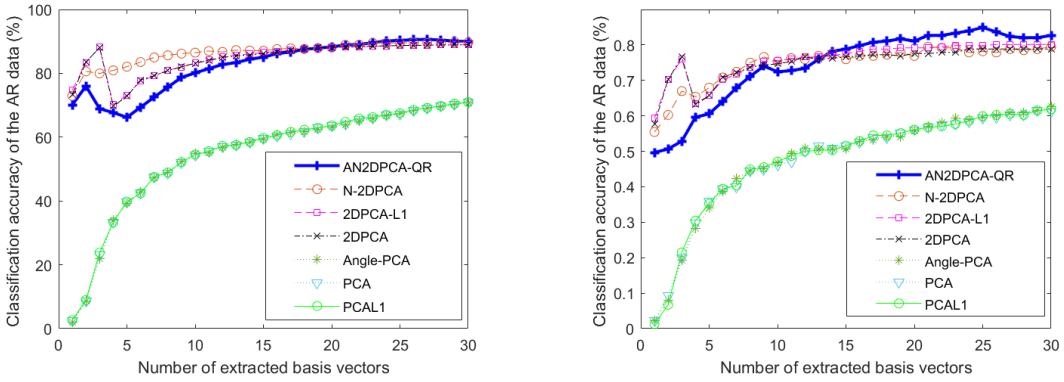


Figure 6. Recognition accuracy versus dimensional variation on the AR database; (a) the AR dataset without noise and (b) the noisy AR dataset

To provide a deeper assessment of the classification performance of the proposed AN2DPCA-QR algorithm, normalized confusion matrices are computed for the CMU PIE and AR datasets under clean conditions. For each experiment, half of the samples from every class were randomly assigned to the training set,

while the remaining samples were reserved for testing. This process was repeated ten times, and the obtained confusion matrices were averaged to ensure reliable evaluation. Figure 8 illustrates the resulting normalized confusion matrices for both datasets. The boxplots of recognition accuracy on the AR dataset over 10 runs Figure 8(a) without noise and Figure 8(b) with noise are also presented. The matrices exhibit a clear concentration of entries along the main diagonal, indicating that most instances are correctly classified. This strong diagonal dominance reflects high intra-class recognition performance with only marginal inter-class errors.

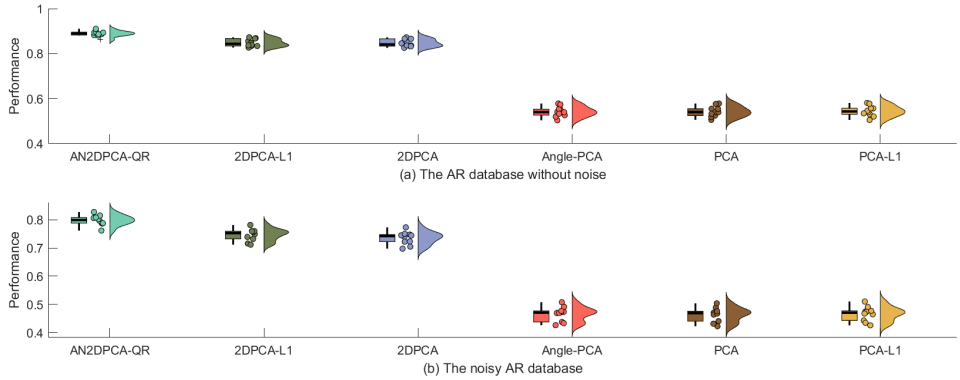
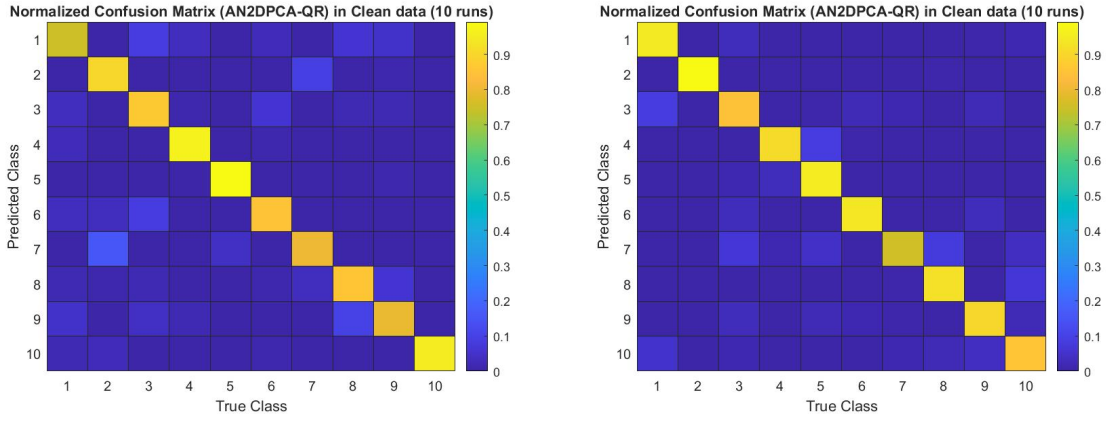


Figure 7. Boxplots of recognition accuracy on noisy AR over 10 runs; (a) the AR dataset without noise and (b) the noisy AR dataset

Table 5. Average execution times on AR

Method	Without noise	Noisy CMU PIE
N-2DPCA	120.497	143.719
AN2DPCA-QR	**60.171**	**67.680**



(a) (b)

Figure 8. Normalized confusion matrices for AN2DPCA-QR; (a) the CMU PIE database and (b) the AR database

4. DISCUSSION

Despite advances in PCA-based methods for face recognition, significant gaps remain. Traditional 2DPCA [4] and its variants [30]–[32], such as L1-2DPCA [6] and N-2DPCA [24], struggle with extreme outliers and structured noise, such as occlusions. The reliance on SVD in methods like N-2DPCA introduces high computational complexity [12], while fixed mean centralization exacerbates bias from outliers [9]. Ad-

ditionally, rotational invariance and convergence guarantees are often lacking, limiting practical applicability in real-time scenarios. The proposed AN2DPCA-QR model demonstrates superior performance, achieving higher recognition accuracy and robustness across benchmark datasets compared to angle-2DPCA [17] and nuclear norm-based approaches [24]. Experimental results (Tables 2 and 3) show improved standard deviations and reduced computation times (Tables 4 and 5), attributed to the use of QR decomposition over SVD. The non-greedy iterative algorithm ensures convergence, validated by Theorem 3, enhancing reliability.

The nuclear norm optimization in AN2DPCA-QR effectively captures spatial dependencies, outperforming Frobenius norm-based angle-2DPCA [17] in noisy conditions. Compared to 2DPCA-L1 [6], which lacks rotational invariance, AN2DPCA-QR maintains this property while reducing complexity from $O(c^3)$ to $O(iter.Slc^2)$ [13]. Boxplots (Figures 4 and 7) and confusion matrices (Figure 8) further confirm its discriminative capacity, though it aligns closely with ROMCA-2DPCA [18] in controlled settings.

The current model is tailored to grayscale images, limiting its applicability to color datasets. Sensitivity to non-Gaussian noise remains a challenge, as the nuclear norm assumes structured noise patterns [19]. Additionally, the iterative nature of the algorithm, while efficient, may still pose scalability issues for extremely large datasets, a concern noted in prior works [12].

AN2DPCA-QR offers significant implications for biometric systems, particularly in real-time face recognition under varying conditions. The use of QR decomposition suggests potential extensions to tensor-based methods for color images or integration with deep learning frameworks [15]. Future work could explore adaptive noise modeling and parallel computing to address scalability, enhancing its utility in large-scale applications. In summary, AN2DPCA-QR addresses critical shortcomings in existing PCA variants by leveraging nuclear norm optimization, adaptive mean centralization, and QR decomposition. It provides a robust, efficient, and convergent solution, outperforming predecessors while opening avenues for further research in multi-dimensional data analysis and real-world deployment.

5. CONCLUSION

This paper introduces AN2DPCA-QR, an alternative framework derived from Angle 2DPCA that incorporates the nuclear norm as a similarity measure in combination with QR decomposition. The proposed approach combines projection variance with the original input information within the objective function, which results in enhanced classification performance compared with traditional 1D and standard 2D models. Furthermore, an efficient non-greedy optimization strategy is developed to compute the AN2DPCA-QR solution, and theoretical analysis demonstrates both its convergence behavior and rotation-invariant property.

Although the proposed approach demonstrates promising results, several limitations should be acknowledged. The experimental evaluation is primarily conducted on grayscale image datasets under particular noise settings, which may limit the applicability of the method to broader visual scenarios. Moreover, the current implementation involves a relatively high computational cost when applied to large-scale datasets.Future research will focus on addressing these issues by extending the experimental validation to a wider variety of datasets, including color image collections. In addition, further efforts will be devoted to developing more efficient optimization schemes in order to decrease computational overhead. Another potential direction is to combine the proposed framework with deep learning techniques to enhance its robustness, especially when dealing with structured outliers appearing across both rows and columns.

FUNDING INFORMATION

Authors state no funding involved.

AUTHOR CONTRIBUTIONS STATEMENT

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Jamal Elalji	✓	✓	✓	✓	✓	✓		✓	✓	✓			✓	✓
Driss Gretete		✓				✓		✓	✓	✓	✓	✓		
Khalid Chougdali	✓	✓	✓	✓	✓	✓			✓		✓	✓	✓	

C : **C**onceptualization
M : **M**ethodology
So : **So**ftware
Va : **Va**lidation
Fo : **Fo**rmal Analysis
I : **I**nvestigation
R : **R**esources
D : **D**ata Curation
O : Writing - **O**riginal Draft
E : Writing - Review & **E**diting
Vi : **Vi**sualization
Su : **Su**pervision
P : **P**roject Administration
Fu : **Fu**nding Acquisition

CONFLICT OF INTEREST STATEMENT

Authors state no conflict of interest.

DATA AVAILABILITY

In this study, we used a dataset available at https://cmu.flintbox.com/technologies.

REFERENCES

[1] I. T. Jolliffe, "Principal component analysis," *Technometrics*, vol. 45, no. 3, p. 276, 2003, doi: 10.1002/0470013192.bsa501.

[2] 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, Jul. 1997, doi: 10.1109/34.598228.

[3] M. S. Bartlett, J. R. Movellan, and T. J. Sejnowski, "Face recognition by independent component analysis," *IEEE Transactions on Neural Networks*, vol. 13, no. 6, pp. 1450-1464, Nov. 2002, doi: 10.1109/TNN.2002.804287.

[4] J. Yang, D. Zhang, A. F. Frangi, and J. Yang, "Two-dimensional PCA: A new approach to appearance-based face representation and recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 26, no. 1, pp. 131-137, Jan. 2004, doi: 10.1109/TPAMI.2004.1261097.

[5] N. Kwak, "Principal component analysis based on l1-norm maximization," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 30, no. 9, pp. 1672-1680, Sep. 2008, doi: 10.1109/TPAMI.2008.114.

[6] X. Li, Y. Pang, and Y. Yuan, "L1-Norm-Based 2DPCA," *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 40, no. 4, pp. 1170-1175, Aug. 2010, doi: 10.1109/TSMCB.2009.2035629.

[7] C. Ding, D. Zhou, X. He, and H. Zha, "R1-pca: Rotational invariant l1-norm principal component analysis for robust subspace factorization," in *Proceedings of the 23rd International Conference on Machine Learning*, 2006, pp. 281–288, doi: 10.1145/1143844.1143880.

[8] F. Nie, H. Huang, C. Ding, D. Luo, and H. Wang, "Robust principal component analysis with non-greedy l1-norm maximization," in *Proceedings-International Joint Conference on Artificial Intelligence (IJCAI)*, vol. 22, no. 1, p. 1433, 2011.

[9] Q. Gao, F. Gao, H. Zhang, X.-J. Hao, and X. Wang, "Two-dimensional maximum local variation based on image euclidean distance for face recognition," *IEEE Transactions on Image Processing*, vol. 22, no. 10, pp. 3807–3817, 2013, doi: 10.1109/TIP.2013.2262286.

[10] J. Wang, "Generalized 2-d principal component analysis by lp-norm for image analysis," *IEEE Transactions on Cybernetics*, vol. 46, no. 3, pp. 792–803, 2015, doi: 10.1109/TCYB.2015.2416274.

[11] P. Bi and X. Du, "Arbitrary triangle structure adaptive mean PCA and image recognition," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 34, no. 2, pp. 754–769, 2023, doi: 10.1109/TCSVT.2023.3289716.

[12] M. Zhao, Z. Jia, Y. Cai, X. Chen, and D. Gong, "Advanced variations of two-dimensional principal component analysis for face recognition," *Neurocomputing*, vol. 452, pp. 653–664, 2021, doi: 10.1016/j.neucom.2020.08.083.

[13] A. Maafiri and K. Chougdali, "Robust face recognition based on a new kernel-pca using rrqr factorization," *Intelligent Data Analysis*, vol. 25, no. 5, pp. 1233–1245, 2021, doi: 10.3233/IDA-205377.

[14] P. Huang, Q. Ye, F. Zhang, G. Yang, W. Zhu, and Z. Yang, "Double l2, p-norm based pca for feature extraction," *Information Sciences*, vol. 573, pp. 345–359, 2021, doi: 10.1016/j.ins.2021.05.079.

[15] J. Wang, M. Zhao, X. Xie, L. Zhang, and W. Zhu, "Fusion of bilateral 2dpca information for image reconstruction and recognition," *Applied Sciences*, vol. 12, no. 24, p. 12913, 2022, doi: 10.3390/app122412913.

[16] X. Yang, W. Wang, L. Liu, Y. Shao, L. Zhang, and N. Deng, "Robust 2dpca by $t\ell_1$ criterion maximization for image recognition," *IEEE Access*, vol. 9, pp. 7690–7700, 2021, doi: 10.1109/ACCESS.2021.3049535.

[17] X. Wang, L. Shi, J. Liu, and M. Zhang, "Cosine 2dpca with weighted projection maximization," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 34, no. 12, pp. 9643–9656, 2022, doi: 10.1109/TNNLS.2022.3159011.

[18] P. Bi, Y. Deng, and X. Du, "A robust optimal mean cosine angle 2dpca for image feature extraction," *Neural Computing and Applications*, vol. 34, no. 22, pp. 20 117–20 134, 2022, doi: 10.1007/s00521-022- 07572-z.

[19] H. Zhang, H. Bi, X. Wang, and P. Zhang, "A joint-norm distance metric 2dpca for robust dimensionality reduction," *Information Sciences*, vol. 640, p. 119036, 2023, doi: 10.1016/j.ins.2023.119036.

[20] Z. Tan and H. Yang, "Reweighted angle two-dimensional principal component analysis for feature extraction," *Pattern Recognition*, vol. 168, p. 111817, 2025, doi: 10.1016/j.patcog.2025.111817.

[21] P. Bi, M. Chen, X. Du, F. Sohrab, and M. Gabbouj, "Double layer robust 2dpca: Joint flexible cut norm and adaptive weighted

learning for small sample size image recognition," *IEEE Transactions on Consumer Electronics*, vol. 71, no. 4, pp. 10491-10506, Nov. 2025, doi: 10.1109/TCE.2025.3602258.
[22] L. Chen, L. Ge, X. Jiang, and H. Li, "Boosting RPCA by Prior Subspace," *IEEE Transactions on Signal Processing*, vol. 73, pp. 2170-2186, 2025, doi: 10.1109/TSP.2025.3569861.
[23] B. Bazatbekov, C. Turan, S. Kadyrov, and A. Aitimov, "2d face recognition using pca and triplet similarity embedding," *Bulletin of Electrical Engineering and Informatics*, vol. 12, no. 1, pp. 580–586, 2023, doi: 10.11591/eei.v12i1.4162.
[24] F. Zhang, J. Yang, J. Qian, and Y. Xu, "Nuclear norm-based 2-dpca for extracting features from images," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 26, no. 10, pp. 2247–2260, Oct. 2015, doi: 10.1109/TNNLS.2014.2376530.
[25] Q. Gao, L. Ma, Y. Liu, X. Gao, and F. Nie, "Angle 2DPCA: A new formulation for 2DPCA," *IEEE Transactions on Cybernetics*, vol. 48, no. 5, pp. 1672–1678, 2017, doi: 10.1109/TCYB.2017.2712740.
[26] M. Fornasier, H. Rauhut, and R. Ward, "Low-rank matrix recovery via iteratively reweighted least squares minimization," *SIAM Journal on Optimization*, vol. 21, no. 4, pp. 1614–1640, 2011, doi: 10.1137/100811404.
[27] A. Sharma, K. K. Paliwal, S. Imoto, and S. Miyano, "Principal component analysis using qr decomposition," *International Journal of Machine Learning and Cybernetics*, vol. 4, no. 6, pp. 679–683, 2013, doi: 10.1007/s13042-012-0131-7.
[28] T. Sim, S. Baker, and M. Bsat, "The cmu pose, illumination, and expression database," in *Proceedings of Fifth IEEE International Conference on Automatic Face Gesture Recognition*, Washington, DC, USA, 2002, pp. 53-58, doi: 10.1109/AFGR.2002.1004130.
[29] A. Martinez and R. Benavente, "The AR face database, CVC," *Copyright of Informatica*, vol. 96, 1998.
[30] Y. Xie, T. Wang, J. Kim, K. Lee, and M. K. Jeong, "Least angle sparse principal component analysis for ultrahigh dimensional data," *Annals of Operations Research*, pp. 1–27, 2024, doi: 10.1007/s10479-024- 06428-0.
[31] H. Zheng, L. Fu, and Q. Ye, "Flexible capped principal component analysis with applications in image recognition," *Information Sciences*,vol. 614, pp. 289–310, 2022, doi: 10.1016/j.ins.2022.06.038.
[32] J. Wang, X. Xie, L. Zhang, J. Li, H. Cai, and Y. Feng, "Robust sparse smooth principal component analysis for face reconstruction and recognition," *PLoS One*, vol. 20, no. 5, p. e0323281, 2025, doi: 10.1371/journal.pone.0323281.

BIOGRAPHIES OF AUTHORS

Jamal Elalji obtained the M.Sc. degree in Applied Mathematics from Ibn Tofail University, Kenitra, Morocco, in 2015. He is currently pursuing the Ph.D. degree with the Engineering Science Laboratory, affiliated with the CED-ST Doctoral Center of Sciences and Technologies at the National School of Applied Sciences (ENSA), Ibn Tofail University, Kenitra, Morocco. His research activities focus on the design and analysis of machine learning methodologies for bio-metric technologies, particularly in the area of automatic face recognition. His broader scientific interests also encompass computer vision, digital image processing, and statistical pattern recognition, with the objective of developing reliable and intelligent biometric identification frameworks. He can be contacted at email: jamal.elalji@uit.ac.ma.

Driss Gretete received his Ph.D. in Probability Theory from Aix-Marseille University, France, in 2005. He is a full Professor of Mathematics at Ibn Tofail University, Kenitra, Morocco, and currently directs the Smart Communications Research Team (ERSC), formerly LEC. His academic interests span probability theory, stochastic processes, fuzzy logic, and logical philosophy. Through his research and teaching activities, he has made extensive contributions to advancing knowledge in mathematical sciences. He can be contacted at email: drissgretete@hotmail.com.

Khalid Chougdali earned his Ph.D. in Computer Science from Mohammed V University–Agdal, Rabat, Morocco, in 2010. He is currently an Associate Professor in the Department of Computer Science at the National School of Applied Sciences (ENSA), Ibn Tofail University, Kenitra, Morocco. His research work covers several areas in computer science, particularly information security, pattern recognition, and biometric technologies, with an emphasis on the development of robust and secure computational solutions. He can be contacted at email: chougdali@gmail.com.