Research Article

Recursive Sample Scaling Low-Rank Representation

Wenyun Gao [1,2] **Xiaoyun Li,[1] Sheng Dai,[1] Xinghui Yin,[2]**
and Stanley Ebhohimhen Abhadiomhen [3]

[1]*Nanjing LES Information Technology Co., LTD, Nanjing, China*
[2]*College of Computer and Information, Hohai University, Nanjing 211100, China*
[3]*Department of Computer Science, University of Nigeria, Nsukka, Nigeria*

Correspondence should be addressed to Stanley Ebhohimhen Abhadiomhen; stanley.abhadiomhen@unn.edu.ng

The low-rank representation (LRR) method has recently gained enormous popularity due to its robust approach in solving the subspace segmentation problem, particularly those concerning corrupted data. In this paper, the recursive sample scaling low-rank representation (RSS-LRR) method is proposed. The advantage of RSS-LRR over traditional LRR is that a cosine scaling factor is further introduced, which imposes a penalty on each sample to minimize noise and outlier influence better. Specifically, the cosine scaling factor is a similarity measure learned to extract each sample's relationship with the low-rank representation's principal components in the feature space. In order words, the smaller the angle between an individual data sample and the low-rank representation's principal components, the more likely it is that the data sample is clean. Thus, the proposed method can then effectively obtain a good low-rank representation influenced mainly by clean data. Several experiments are performed with varying levels of corruption on ORL, CMU PIE, COIL20, COIL100, and LFW in order to evaluate RSS-LRR's effectiveness over state-of-the-art low-rank methods. The experimental results show that RSS-LRR consistently performs better than the compared methods in image clustering and classification tasks.

1. Introduction

The limitations of classical feature learning techniques such as PCA [1] easily made the robust principal component analysis (RPCA) method an efficient choice for dealing with noise and outliers. Specifically, RPCA is focused on learning a low-rank subspace directly from the original high-dimensional data to preserve its geometric structure in a low-dimensional subspace. And this strategy has shown tremendous improvements in several applications [2–5]. However, as RPCA only seeks a single low-rank subspace, it may still be limited with noise damage since high-dimensional data are known to reside in multiple low-dimensional subspaces [6]. Thus, extending RPCA's idea, Liu et al. [4, 7] proposed a method named "low-rank representation" (LRR). LRR's main advantage over RPCA lies in its aim to learn data's multiple low-dimensional subspaces and their membership. This approach makes LRR very robust to the negative effect of noise and outliers [8].

Therefore, considering LRR's robustness mentioned above, several attempts were made in the literature such as references [9–13] to improve its performance. For example, to deal with the data from nonlinear subspaces, Tang et al. [13] proposed robust kernel LRR (RKLRR). Liu et al. [11] then adopted a fixed-rank strategy to accelerate LRR's computation process. However, their performances could be reduced with insufficient or heavily corrupted samples. That is why Xiao et al. [14] had previously proposed the latent LRR (LatLRR) for joint subspace segmentation and feature selection. The idea behind LatLRR is to include hidden data in constructing the dictionary to improve robustness further. On the contrary, how to handle gross data damage remains unsolved. It is observed that the more the data corrupted by noise, the larger the degradation in the classification and clustering performance.

To address this issue, a recursive sample scaling low-rank representation (RSS-LRR) method is proposed in this paper. Since some data samples will be more damaged in gross data

corruption, we estimate each data sample's importance using a cosine scaling factor. This scaling factor measures the angle connecting each data sample and the low-rank representation's principal components in feature space. In this way, we then iteratively subdue noisy data samples to overcome their effect. Thus, the proposed RSS-LRR can effectively obtain a good low-rank representation than existing methods. Our main contributions are summarized as follows:

(1) We propose a novel method named "RSS-LRR," which measures each data sample's importance using a cosine scaling factor. This scaling factor is used in our model to extract each data sample's relationship with the low-rank matrix's principal components.

(2) The proposed RSS-LRR method can effectively handle noisy data samples by iteratively restricting noisy data samples using the sample scaling factor to suppress noise so as to obtain a good low-rank representation.

(3) Several experiments are performed with varying levels of corruption to evaluate RSS-LRR's effectiveness. The experimental results show that RSS-LRR consistently outperforms other state-of-the-art methods in image clustering and classification tasks.

2. Related Work

This section presents a brief review of the baseline methods RPCA and LRR. First, in this paper, matrices are written in uppercase, e.g., X. Thus, $\|X\|_F$ and $\|X\|_*$ denote the Frobenius norm and nuclear norm, respectively. $\|X\|_2$ and $\|X\|_{2,1}$ denote the vector norm and $L_{2,1}$-norm, which are defined by $\|X\|_1 = \sum_{i,j}|x_{ij}|$ and $\|X\|_{2,1} = \sum_i [\sum_i (x_{ij})^2]^{1/2}$.

2.1. Robust Principal Component Analysis. To recover a subspace structure from corrupted data, RPCA was proposed in [2]. Its strategy is to decompose a given data matrix into two components matrices by solving the following optimization problem:

$$
\begin{aligned}
\min_{D,E} \quad & \|D\|_* + \lambda\|E\|_1 \\
\text{s.t.} \quad & X = D + E,
\end{aligned}
\tag{1}
$$

where the data matrix $X = [x_1, x_2, \ldots, x_n] \in \mathfrak{R}^{m \times n}$ is a matrix of n samples in the m-dimensional space, D is the low-rank matrix, E is a sparse error matrix, and λ is the regularized parameter to balance the effects of the two terms. Thus, RPCA's main objective from the above formula is to obtain low-rank and sparse elements by combining the nuclear norm and L_1-norm. This approach is proven to be possible under some assumptions [16]. However, as RPCA assumes a single low-rank subspace, its performance can degrade easily.

2.2. Low-Rank Representation. Liu et al. [5] proposed LRR to tackle RPCA's limitations. Specifically, LRR is focused on pursuing a data representation matrix with the lowest rank.

It can achieve that by using data's self-expressiveness property such that the given data itself are utilized as a self-dictionary. This way, each data sample is then represented as a linear combination of similar samples belonging to the same class. The optimal low-rank matrix obtained by LRR is defined as follows:

$$
\begin{aligned}
\min_{Z,E} \quad & \|Z\|_* + \lambda\|E\|_p \\
\text{s.t.} \quad & X = XZ + E,
\end{aligned}
\tag{2}
$$

where the data matrix $X = [x_1, x_2, \ldots, x_n] \in \mathfrak{R}^{m \times n}$ denotes the self-dictionary E is used to capture the error components where $\|.\|_p$ denotes a certain norm, which can be determined based on the type of noise corruption. For example, while $\|.\|_F$ is a suitable candidate for data damaged by Gaussian noise, $\|.\|_1$ is good for random noise. Besides, $\|.\|_{2,1}$ is an efficient choice when only a part of data are contaminated.

Although LRR's approach is shown to be very effective, particularly in noisy settings, its performance may degrade with insufficient samples. For this reason, Liu et al. [15] also proposed latent LRR that exploits both the observed and the hidden data to construct the self-dictionary. This strategy is most useful for image restoration [17]. Consequently, the work of [18] proposed a method for line pattern noise removal to address contaminated instances. It is realized in a transform domain by using a line pattern's directional property. Besides, other efforts were made in references [11, 12, 19–31] to improve LRR's discriminative capability.

Notably, Bing-Kun Bao et al. [23] used a fixed-rank approach so as to reduce LRR's singular value decomposition cost. Zhang et al. [24] proposed two instantaneous methods: the first can reasonably handle noise interference by decomposing given data into two parts, namely, the low-rank sparse principal feature part and a noise-fitting error part. In [25], Tang et al. introduced a diversity regularization and a rank constraint to suppress the redundancy in different data views. In [26], Zhang et al. presented a method that adaptively preserves local information of salient features, thus guaranteeing a block-diagonal coefficient structure. Meanwhile, a compressive robust subspace clustering method was proposed in [27] for dimensionality reduction. However, because the subspace techniques, including LRR, do not provide linear dimensionality reduction (LDR) functionality, the feature selective projection (FSP) [28] was proposed. FSP combines feature extraction, feature selection, and LRR into a unified model to promote robust LDR. Likewise, a method was introduced in [29], which exploits a robust dictionary learning strategy to discover hybrid salient low-rank and sparse representations in a factorized compressed space. Furthermore, in an attempt to keep both similarity and local structures, the hierarchical weighted low-rank representation (HWLRR) [30] was proposed. Similarly, the more recent study [31] was focused on capturing cross-view information through an approach that preserves both diversity and consensus information of each data view.

3. The Proposed Method

3.1. RSS-LRR. Since real-world data are not always perfect and inevitably corrupted by noise in practice, most existing low-rank methods cannot guarantee robust performance. Therefore, noise interference must be carefully handled to resolve the present drawback. As a result, a rational solution is pursued in this paper, which uses a cosine scaling factor to estimate the importance of each data sample. Essentially, we suppose that the clean data samples will give high significant values, while the noisy ones would differ from the principal component of the data. Thus, the cosine scaling factor D is introduced into the LRR formulation in equation (2) using the constraint $XD = XDZ + E$. The motivation behind our approach is straightforward: according to reference [32], Z can be decomposed as $U\Sigma V^T$, where U and V are left and right low-rank singular vectors. In other words, U or V becomes the pursued projection vector such that $XU \epsilon R^m$ is the data projection in feature space. Therefore, XU_1 is chosen as our maximum projection direction where U_1, the maximum eigenvalue of Σ, is U's column vectors. So, for an outlier data sample x_j, the angle between it and the principal component vector XU_1 would differ more than that of clean data x_i, as described in Figure 1(a). Hence, d_i expressed in the following is used to estimate the importance of each data sample:

$$d_i = |\cos \theta_i| + \varepsilon = \left| \frac{x_i^T (XU_1)}{\|x\|_i XU_1} \right| + \varepsilon, \tag{3}$$

where ε is a constant that stops d_i from 0. Thus, using the significant factor d, a given data matrix can be scaled to minimize the effect of noisy data samples, allowing the low-rank structure to be realized with clean data as shown in Figure 1(b). As such, the proposed sample scaling low-rank model is obtained as follows:

$$\min_{Z,E} \quad \|Z\|_* + \lambda\|E\|_{2,1}$$
$$\text{s.t.} \quad XD = XDZ + E, \tag{4}$$

where $\mathbf{Z}$ denotes the low-rank matrix and $\mathbf{E}$ is used to capture the noise elements, similar to that in equation (2).

Then, $D = \begin{bmatrix} d_1 \\ & \ddots & d_i \\ & & \ddots & d_n \end{bmatrix}$ denotes the scaling factor of n samples.

From the above formulation, one can easily check that the scaling factor $\pi/2$ suits our goal, as we can then detect the noisy points from closer angles to $\pi/2$, with lower values being assigned to such points. Illustratively, let us assume $\theta_i = \pi/2$, and d_i is almost 0, so x_i is subdued with $\hat{x}_i = x_i d_i$. Thus, $\hat{X} = XD$ is used to obtain new training data. Besides, using SVD, $\hat{X} = XD = U_x \Sigma_x V_x^T D = \widehat{U}_x \widehat{\Sigma}_x \widehat{V}_x^T$, meaning that both new singular vectors $\widehat{U}_x$ and new projection vectors of sample space are obtained by suppressing the noisy data samples. As a result, our proposed method can then learn an optimal low-rank structure using new data $\hat{X}$ where the points closer to the principal component vector are enhanced.

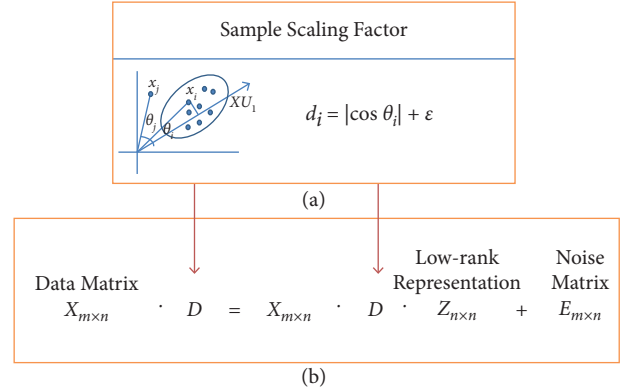We give the summary of our model's main characteristics as follows:

Figure 1: Our proposed model's framework.

(i) Unlike the existing low-rank methods, which use the input data X itself as the dictionary, a new dictionary is presented with $\hat{X} = XD$ by imposing a recursive scaling factor D on X to suppress the effect of noisy samples.

(ii) Specifically, our recursive modeling is very useful for learning a good low-rank representation, most especially when the data are heavily contaminated. As shown in equation (4), our focus is to find Z by minimizing equation (4) with the constraint of $\hat{X} = \hat{X}\hat{Z}$, thus allowing Z to preserve a better low-rank structure using only data samples with huge cosine similarity (referred to as clean data samples as they would have the smaller angle with the principal component). In other words, Z is obtained by equation (4) using clean data $\hat{X}$.

3.2. Optimization. In this section, we propose an optimization algorithm to solve equation (4). First, following standard practice, we introduce a variable $J = Z$ to relax equation (4) further. Thus, equation (4) can be recast as

$$\text{argmin}_{J,Z,E}\|J\|_* + \lambda\|E\|_{2,1}$$
$$\text{s.t.} \quad XD = XDZ + E, Z = J. \tag{5}$$

The augmented Lagrangian function equation (5) is given as

$$L(J, E, Z, \mu) = J_* + \lambda E_{2,1} + \text{tr}\left(M_1^T (XD - XDZ - E)\right)$$
$$+ \text{tr}\left(M_2^T (Z - J)\right) \tag{6}$$
$$+ \frac{\mu}{2}\left(XD - XDZ - E_F^2 + Z - J_F^2\right),$$

where M_1 and M_2 are Lagrange multipliers, $\mu > 0$ is a penalty parameter, and $\|\cdot\|_F$ denotes the Frobenius norm of a matrix. Many concepts for convex optimization have been developed [33–35], which rely on nuclear-norm regularization. And the optimization problem can be solved via the method in [36].

3.2.1. Computation of J. According to references [37, 38], nuclear-norm minimization methods have a stable performance. For computing J, we rewrite equation (6) as

$$J^* = \arg\min_J \|J\|_* + \frac{1}{2}\left\|J - \left(Z + \frac{M_2}{\mu}\right)\right\|_F^2. \qquad (7)$$

3.2.2. Computation of Z. By fixing J and E and substituting $\widehat{X} = X\,D$, Z can be updated using the following formula:

$$Z^* = \left(I + \widehat{X}^T\widehat{X}\right)^{-1}\left(\widehat{X}^T\widehat{X} - \widehat{X}^T E + J + \frac{\left(\widehat{X}^T M_1 - M_2\right)}{\mu}\right). \qquad (8)$$

3.2.3. Computation of E. With $\widehat{X}$, Z, and J fixed, E can be solved as follows:

$$E^* = \arg\min \frac{\lambda}{\mu}\|E\|_{2,1} + \frac{1}{2}\left\|E - \left(\widehat{X} - \widehat{X}Z + \frac{M_1}{\mu}\right)\right\|_F^2. \qquad (9)$$

Following reference [39], equation (9) can be solved by the following lemma.

Lemma 1 (see [40]). *Let* $Q = [q_1, q_2, \ldots, q_n]$ *be a given matrix. If* W^* *is the optimal solution, then*

$$W^* = \min_w \lambda\|W\|_{2,1} + \frac{1}{2}\|W - Q\|_F^2. \qquad (10)$$

Thus, the ith column of W^* is

$$W^*(:,i) = \begin{cases} \dfrac{q_i - \lambda}{q_i} q_i, & \text{if } \lambda < \|q_i\|, \\ \\ 0, & \text{otherwise.} \end{cases} \qquad (11)$$

Based on Lemma 1, supposing we have a matrix Q, W^* can be reached directly using the above formula, making the E computation process very efficient.

The complete solution is given in Algorithm 1.

3.3. Complexity Analysis. This section gives an analysis of the computational cost of Algorithm 1. In Step 6 of Algorithm 1, the values of J_k, Z_k, E_k are the same as J_{k-1}, Z_{k-1}, E_{k-1} in the inner loop. Therefore, as the number of iterations k increases, the time complexity in pursuing the low-rank variable Z_k decreases faster. From the above discussion, the computing time in t_k is far shorter than that in t_1. The cost of SVD is $O(n^3)$, where n denotes the number of data vectors. Besides, Z costs $O(n^2m + n^3)$ for computation, where m is the data dimension. E in Step 10 then costs $O(dn)$. Furthermore, supposing that the D subblock requires k times until convergence, the J, Z, E subproblems will then be calculated for t_k iterations. Therefore, based on the number of iterations, the combined cost is $O(t(n^3 + n^2m + dn))$, such that t is a representation of the number of iterations. Thus,

when $n \geq m$, the cost's upper bound would be $O(tn^3)$. Accordingly, the overall computational cost of the proposed method is $O(ktn^3)$.

4. Experiments

In this section, our proposed RSS-LRR method's effectiveness is evaluated by comparing it with similar methods such as LRR [5], latent LRR [15], LRRLC, GLRR [8], and GODEC+ [41] in background modeling from video [42] and image denoising [43].

4.1. Background Modeling from Surveillance Video

4.1.1. Experimental Settings. This experiment is performed using surveillance video with various illumination settings. It is composed of a chain of 200 grayscale frames of 32×32 dimensions. Thus, each algorithm's effectiveness is evaluated using precision, recall, and F-score metrics, and their parameters are tuned according to the corresponding literature. Precisely, background modeling [44] is measured by manually quoting out the activities. In this experiment, 50% of frames are randomly selected as the training set, while the remaining are treated as the testing set.

4.1.2. Experimental Results. In Figure 2, we show each algorithm's background recovery and activity segmentation performance. Additionally, it can be observed from Table 1 that RSS-LRR outperforms other methods in activity segmentation as it can better generalize to the testing frames. This result further substantiates our sample scaling factor approach's effectiveness such that a more reliable low-rank object is obtained than that from the compared methods.

4.2. Image Clustering with Varying Levels of Noise

4.2.1. Experimental Settings. Here, several experiments are performed on three well-known image datasets, namely, ORL, CMU PIE, and COIL20, to evaluate the effectiveness of the proposed algorithm on image clustering. Each algorithm's parameters are tuned according to the corresponding literature using the grid search strategy. We give a brief description of each one of these datasets as follows:

> ORL: it contains face images of ten individuals, with each of them contributing forty distinct images under various conditions such as facial details and different facial expressions.
>
> CMU PIE: it is a face image repository with images of sixty-eight individuals with different settings. It includes thirteen different poses, four different expressions, and forty-two different illuminations.
>
> COIL20: it is an object image dataset consisting of 20 separate objects. Each object contributes 72 grayscale images, amounting to a total of 1440 images.

For each dataset, the images are resized to $32 * 32$ dimensions in our experiments. As illustrated in Figure 3,

(1) **Input**: training dataset **X**, regulation parameter λ
(2) Initialize: $t = 0$, $k = 0$, $J_{k=0} = 0$, $Z_{k=0} = 0$, $E_{k=0} = 0$, $\epsilon_1 = 10^{-6}$, $\epsilon_2 = 10^{-6}$, $\varepsilon = 0.0001$, $D = \text{Diag}(\text{ones}(n,1))$.
(3) **While** not converged **do**
(4) $\quad$ Update $\widehat{X}$ by $\widehat{X} = X\,D$.
(5) $\quad$ Initialize: $M_1 = 0$, $M_2 = 0$, $\mu > 0$, $\rho > 0$.
(6) $\quad$ Set: $J_k = J_{k-1}$, $Z_k = Z_{k-1}$, $E_k = E_{k-1}$.
(7) $\quad$ **While** not converged **do**
(8) $\quad\quad$ Update J while fixing others by equation (7).
$$J^* = \arg\min{}_J J_* + (1/2)J - (Z + (M_2/\mu))_F^2$$
(9) $\quad\quad$ Update Z while fixing others by equation (8).
$$Z^* = (I + \widehat{X}^T \widehat{X})^{-1}(\widehat{X}^T \widehat{X} - \widehat{X}^T E + J + (\widehat{X}^T M_1 - M_2)/\mu)$$
(10) $\quad\quad$ Update E while fixing others by equation (9).
$$E^* = \arg\min (\lambda/\mu)\|E\|_{2,1} + (1/2)\|E - (\widehat{X} - \widehat{X}Z + (M_1/\mu))\|_F^2$$
(11) $\quad\quad$ Update the multipliers.
$$M_1 = M_1 + \mu(\widehat{X} - \widehat{X}Z - E),$$
$$M_2 = M_2 + \mu(Z - J)$$
(12) $\quad\quad$ Update μ by $\mu = \min(\rho\mu, \max(\mu))$
(13) $\quad\quad$ Update $t_k = t_k + 1$
(14) $\quad\quad$ Check convergence conditions
$$\|\widehat{X} - \widehat{X}Z_k - E_k\|_\infty < \epsilon_1, \|J_k - E_k\|_\infty < \epsilon_1$$
(15) $\quad$ **end while**
(16) $\quad$ $[U, S, V] = sv\,d(Z), U_1 = U(:,1);$
(17) $\quad$ Update **D** by $d_i = |\cos\theta_i| + \varepsilon$ and $\cos\theta_i = x_i^T(XU_1)/x_i XU_1$.
(18) $\quad$ Update $k = k + 1$
(19) $\quad$ Check convergence conditions
(20) $\quad\quad$ $\|\widehat{X} - \widehat{X}Z - E\|_\infty < \epsilon_2, \|J - E\|_\infty < \epsilon_2$
(21) **end while**
(22) **Output**: **Z**, **E**

ALGORITHM 1: Algorithm of our proposed RSS-LRR.

these datasets are then corrupted with 5%, 10%, 15%, and 20% random pixel noise to demonstrate each algorithm's robustness to noise. Thus, a spectral clustering algorithm is applied to the similarity matrix of each algorithm to obtain the clustering results with ten multiple tries to ensure fairness [45].

4.2.2. Experimental Results. Tables 2–4 display the clustering results of each algorithm concerning the accuracy evaluation metric on ORL, CMU PIE, and COIL20, respectively. Therefore, it is obvious to see that the accuracy of our proposed RSS-LRR method consistently beats those of the compared methods in all three datasets. For example, on the ORL dataset, the accuracy of our proposed method is about 1% higher than that of its closest competitor on clean data. Then, gradually increasing the noise level, one may notice that all the algorithms had reduced performance. However, our proposed method shows more robustness than the other methods, especially with 20% noise damage. Let us take, for instance, the clustering result of LRR, which moved from 0.7505 to 0.3497, while that of the proposed method had a lower drop from 0.7609 to 0.5650.

Similarly, in Tables 3 and 4, we can also see that the proposed method maintained its performance over other methods on the CMU PIE and COIL20 datasets. Particularly, RSS-LRR results with a 20% corruption level show that its accuracy is about 4% better than that of GODEC+, which is 0.5069 on the COIL20 dataset. Accordingly, we present the

clustering variation graph of each method in Figure 4 that further reveals the robustness of the proposed method to noise. Thus, it is safe to conclude that the proposed method's performance is steadier than that of the other algorithms, especially at a high level of corruption. We attribute that to our scaling factor approach in iteratively overcoming the noise effect.

4.3. Image Recognition with Contiguous Occlusion

4.3.1. Experimental Settings. In order to evaluate the robustness of RSS-LRR on image recognition under different levels of contiguous occlusions in images [19], we randomly add 6×6 and 8×8 46 block occlusions to each dataset, as illustrated in Figure 5. Therefore, 50% of samples in each dataset are selected as the training set and the rest as the testing set. Besides, we compare our proposed method's performance with that of similar ones: LRR, LRRLC, latent LRR (LLRR), GLRR, and GODEC+, by adopting relevant experimental settings in Section 4.2. Thus, each algorithm's classification accuracy is evaluated using the K nearest neighbor (KNN) classifier.

4.3.2. Experimental Results. In Table 5, the average classification accuracies are obtained on the ORL dataset in two levels of contiguous occlusions. From Table 5, we can see that the accuracy of RSS-LRR is only slightly higher than that
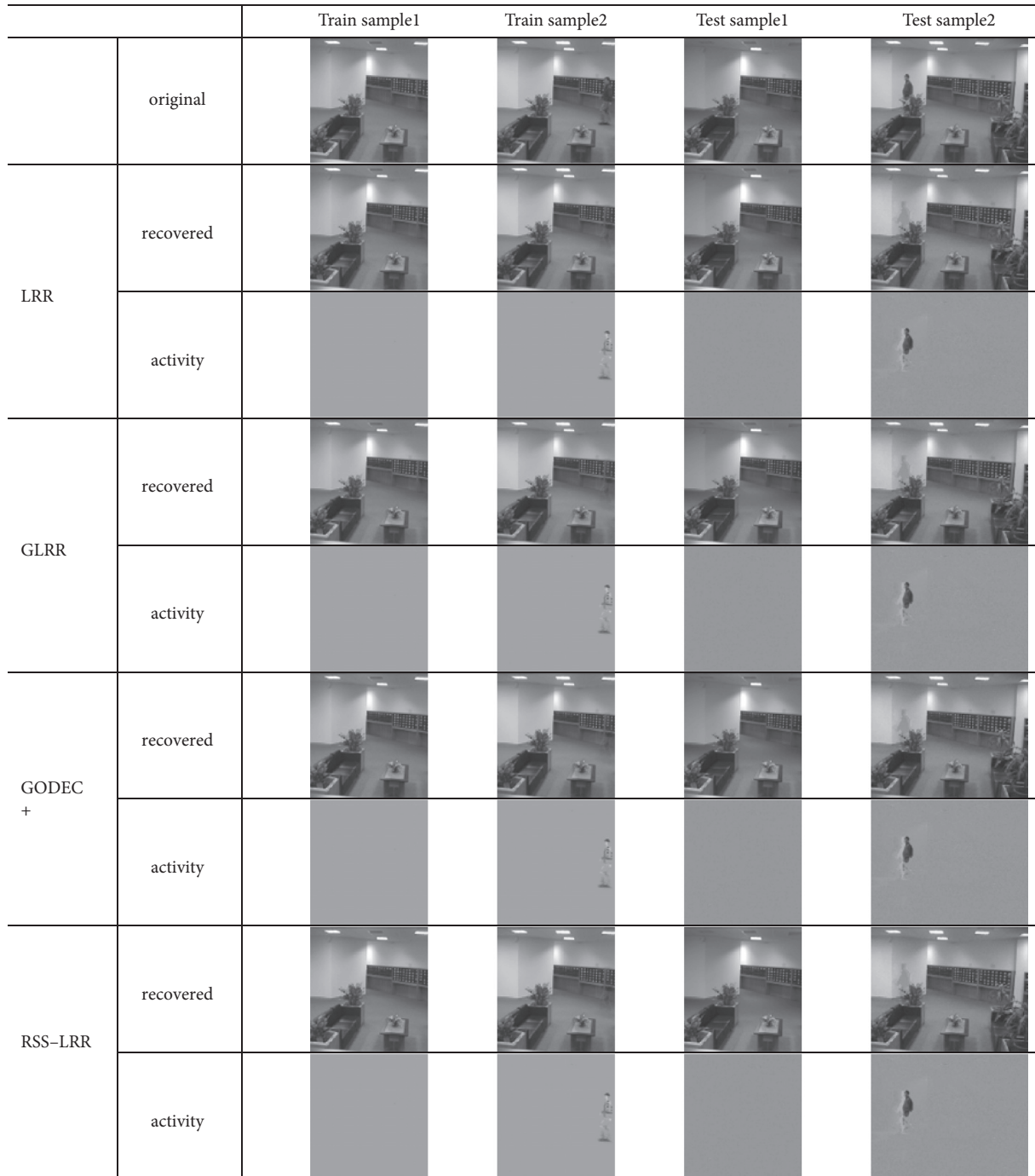
		Train sample1	Train sample2	Test sample1	Test sample2
	original				
LRR	recovered				
	activity				
GLRR	recovered				
	activity				
GODEC +	recovered				
	activity				
RSS–LRR	recovered				
	activity				

Figure 2: Results obtained by all algorithms on surveillance video.

Table 1: Activity results of all algorithms on surveillance video.

	LRR		GLRR		GODEC+		RSS-LRR	
	Train	Test	Train	Test	Train	Test	Train	Test
Precision	0.9153	0.4872	0.9204	0.8859	0.9208	0.8840	0.9388	0.8919
Recall	0.7897	0.7358	0.7937	0.7478	0.7940	0.7492	0.7999	0.7594
F-score	0.8478	0.5862	0.8523	0.8110	0.8527	0.8110	0.8638	0.8213

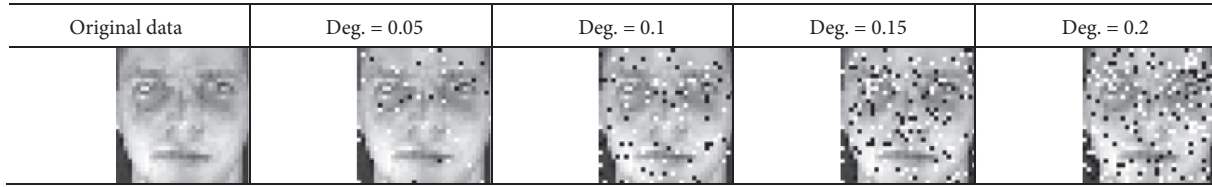| Original data | Deg. = 0.05 | Deg. = 0.1 | Deg. = 0.15 | Deg. = 0.2 |

FIGURE 3: Example of the original image and corrupted one with different degree of noise on the ORL dataset.

TABLE 2: Clustering results concerning the accuracy of different algorithms on the ORL dataset.

	0%	5%	10%	15%	20%
LRR	0.7505 ± 1.65	0.7335 ± 1.82	0.6955 ± 1.76	0.6247 ± 1.75	0.3497 ± 1.95
LRRLC	0.7512 ± 1.17	0.7248 ± 1.69	0.7053 ± 1.72	0.6260 ± 1.27	0.3604 ± 2.34
LLRR	0.7518 ± 1.70	0.7371 ± 1.31	0.7169 ± 1.18	0.6397 ± 1.67	0.3953 ± 1.98
GLRR	0.7460 ± 1.03	0.7402 ± 0.95	0.7098 ± 1.48	0.6650 ± 1.65	0.5638 ± 1.22
GODEC+	0.7469 ± 1.27	0.7344 ± 1.03	0.7156 ± 1.44	0.6726 ± 1.76	0.5413 ± 1.75
RSS-LRR	0.7609 ± 1.09	0.7487 ± 1.38	0.7182 ± 1.12	0.6825 ± 1.23	0.5650 ± 1.50

TABLE 3: Clustering results concerning the accuracy of different algorithms on the CMU PIE dataset.

	0%	5%	10%	15%	20%
LRR	0.6964 ± 1.03	0.6906 ± 1.54	0.6631 ± 1.80	0.5924 ± 1.81	0.4703 ± 1.75
LRRLC	0.7086 ± 1.41	0.7017 ± 1.84	0.6777 ± 1.70	0.6075 ± 1.95	0.4919 ± 2.07
LLRR	0.7188 ± 1.27	0.7083 ± 1.42	0.6926 ± 1.53	0.6149 ± 1.96	0.5118 ± 1.83
GLRR	0.7212 ± 1.57	0.7109 ± 1.67	0.6947 ± 1.03	0.6776 ± 1.71	0.5657 ± 1.92
GODEC+	0.7297 ± 0.88	0.7147 ± 1.21	0.7058 ± 1.75	0.6406 ± 1.97	0.5422 ± 1.86
RSS-LRR	0.7308 ± 1.27	0.7286 ± 1.39	0.7125 ± 1.48	0.6834 ± 1.65	0.5725 ± 1.61

TABLE 4: Clustering results concerning the accuracy of different algorithms on the COIL20 dataset.

	0%	5%	10%	15%	20%
LRR	0.7047 ± 1.25	0.6869 ± 1.55	0.6203 ± 1.47	0.4757 ± 1.97	0.3955 ± 1.91
LRRLC	0.7158 ± 1.48	0.6875 ± 1.52	0.6519 ± 1.60	0.5291 ± 2.04	0.4112 ± 2.09
LLRR	0.7170 ± 1.71	0.6949 ± 1.69	0.6618 ± 1.97	0.5340 ± 1.93	0.4260 ± 1.99
GLRR	0.7234 ± 2.02	0.6976 ± 1.98	0.6757 ± 1.51	0.5687 ± 2.38	0.4718 ± 2.12
GODEC+	0.7224 ± 1.44	0.7006 ± 1.56	0.6822 ± 1.72	0.5977 ± 2.13	0.5069 ± 1.86
RSS-LRR	0.7285 ± 1.66	0.7103 ± 1.42	0.7055 ± 1.55	0.6123 ± 1.89	0.5417 ± 2.35

of GODEC+. On the contrary, RSS-LRR's accuracy is significantly better than that of LRRLC, LLRR, and LRR by about 5%, 4%, and 12%, respectively, under 6×6 occlusion. For example, the classification accuracy of RSS-LRR is 0.6015 and 0.4891 in LRR. This indicates that our proposed RSS-LRR can effectively obtain a good low-rank representation than existing methods.

From Table 6, we can see that the accuracy of RSS-LRR is better than that of LRR by about 6%. Under $6 * 6$ occlusion, the accuracy of RSS-LRR is better than that of GODEC+ and GLRR by about 2%. Thus, our proposed RSS-LRR demonstrates approximately outstanding effectiveness on classification accuracies among all algorithms. In Table 7, the accuracy of RSS-LRR is higher than that of LRRLC and LRR by about 3% and 6% under $6 * 6$ occlusion. For example, the classification accuracy of GLRR is 0.6985 and 0.7130 in RSS-LRR, while the result of LRR is 0.6522 under $6 * 6$ occlusion.

Figures 6(a)–6(f) illustrate the variations of classification accuracies with increasing feature dimensionalities in different occlusions on ORL, CMU PIE, and COIL20 datasets.

From Figures 6(a) and 6(b), we can see that RSS-LRR achieves the highest accuracies among all algorithms. In Figures 6(c) and 6(d), we can see that the performance of RSS-LRR becomes better than that of others on the CMU PIE dataset when the feature dimensionality is over 70. From Figures 6(e) and 6(f), on the COIL20 dataset, we can see that the accuracies of RSS-LRR gradually show its superiority when the feature dimensionalities are more than 70.

From the above discussion, it can be seen that our proposed RSS-LRR shows better performance in classification accuracy compared with the other algorithms.

4.4. Experiments on Large-Scale Dataset

4.4.1. Experimental Settings. In this experiment, RSS-LRR effectiveness is further evaluated on two larger datasets, namely, COIL100 and LFW. Also, the relevant settings from previous sections are adopted to perform the large-scale experiment. We give a brief description of each dataset as follows:
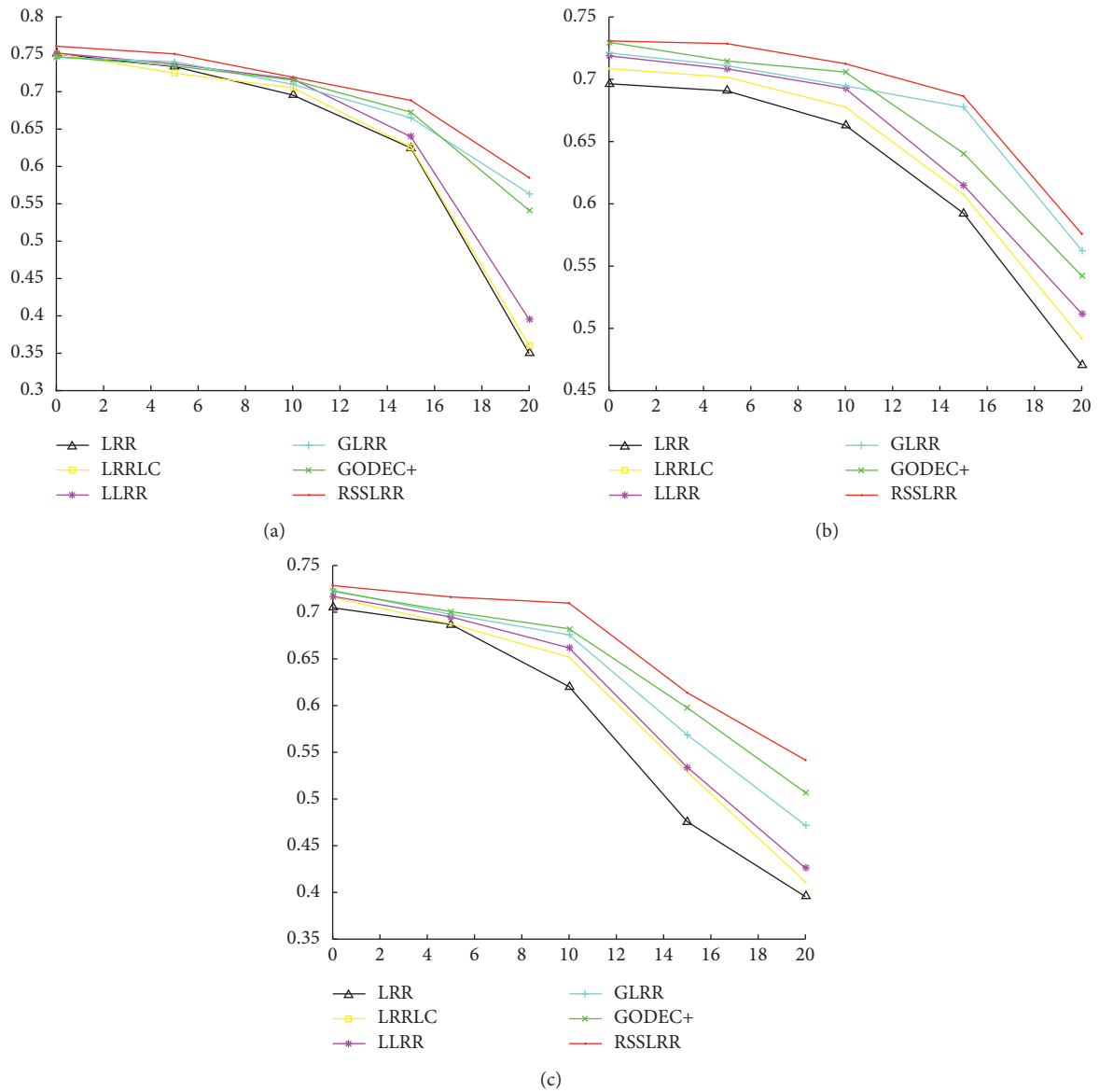
(a)

(b)

(c)

Figure 4: Clustering variation of different algorithms with varying noise levels on (a) ORL, (b) PIE, and (c) COIL20 datasets.

Figure 5: Example of the original image and corrupted one with different occlusion sizes on the COIL20 dataset.

Table 5: Classification accuracies on ORL under 6×6 and 8×8 occlusions.

	LRR	LRRLC	LLRR	GLRR	GODEC+	RSS-LRR
6×6	0.4891	0.5488	0.5534	0.5721	0.5953	0.6015
8×8	0.3899	0.4046	0.4337	0.4473	0.4620	0.4762

Table 6: Classification accuracies on CMU PIE under 6×6 and 8×8 occlusions.

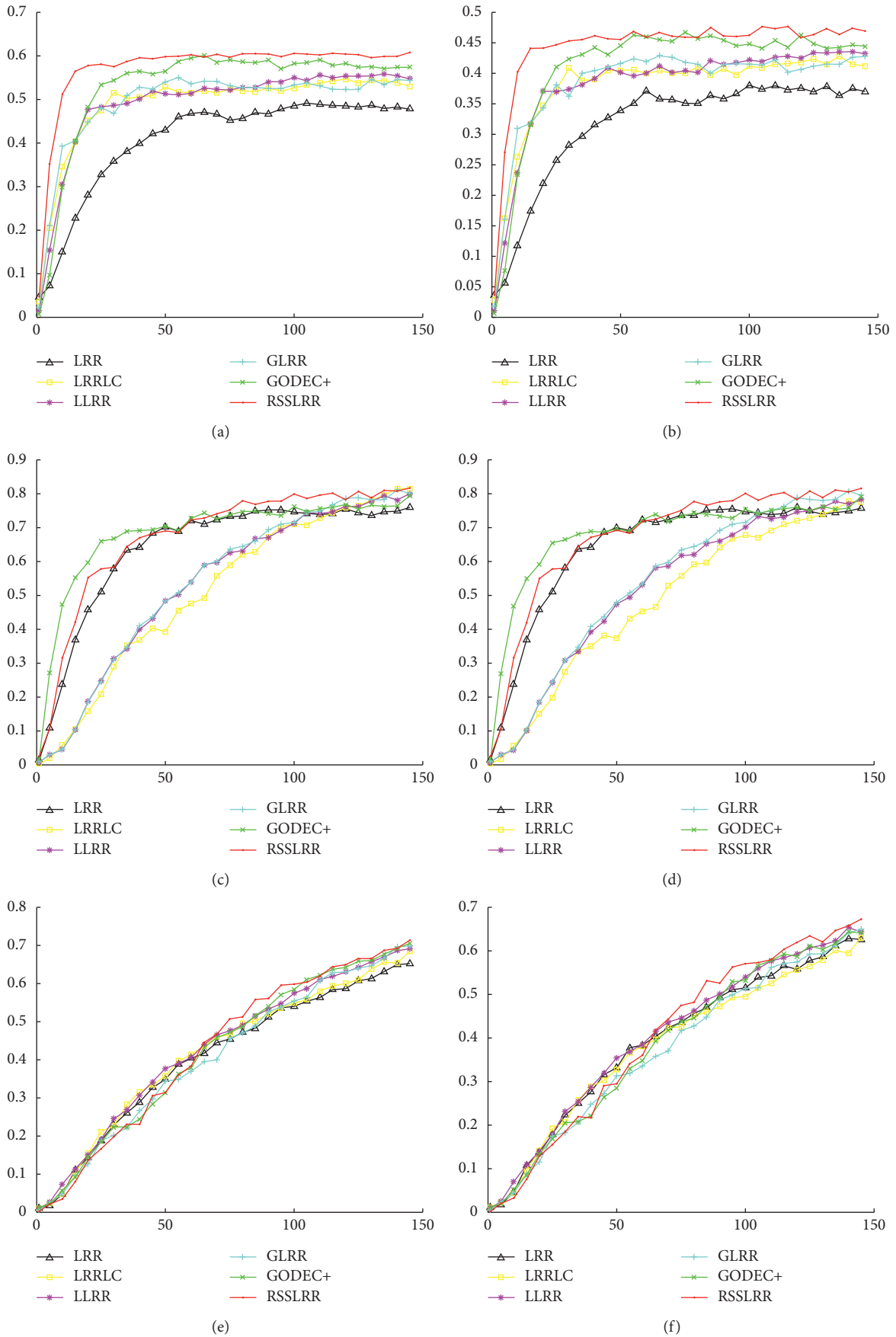|	LRR	LRRLC	LLRR	GLRR	GODEC+	RSS-LRR
6×6	0.7835	0.7980	0.8054	0.8247	0.8249	0.8415
8×8	0.7589	0.7773	0.7821	0.8081	0.7919	0.8112

Figure 6: Variations of classification accuracies with increasing feature dimensionalities of 6×6 and 8×8 occlusions on (a, b) ORL, (c, d) CMU PIE, and (e, f) COIL20 datasets.

TABLE 7: Classification accuracies on COIL20 under 6×6 and 8×8 occlusions.

	LRR	LRRLC	LLRR	GLRR	GODEC+	RSS-LRR
6×6	0.6522	0.6843	0.6912	0.6985	0.7056	0.7130
8×8	0.6282	0.6305	0.6538	0.6490	0.6433	0.6638

TABLE 8: Classification accuracies on COIL100 under 6×6 and 8×8 occlusions.

	LRR	LRRLC	LLRR	GLRR	GODEC+	RSS-LRR
6×6	0.4765	0.5369	0.5501	0.5689	0.5912	0.6052
8×8	0.3812	0.3991	0.4287	0.4454	0.4588	0.4803

TABLE 9: Classification accuracies on LFW under 6×6 and 8×8 occlusions.

	LRR	LRRLC	LLRR	GLRR	GODEC+	RSS-LRR
6×6	0.5106	0.5562	0.5732	0.5899	0.6021	0.6457
8×8	0.4031	0.4321	0.4416	0.4541	0.4953	0.5276

TABLE 10: Clustering results concerning the accuracy of different algorithms on the COIL100 dataset.

	0%	5%	10%	15%	20%
LRR	0.6854 ± 1.53	0.6741 ± 1.28	0.6376 ± 1.89	0.4964 ± 2.61	0.4236 ± 1.42
LRRLC	0.6931 ± 1.37	0.6823 ± 1.43	0.6423 ± 1.59	0.5191 ± 1.87	0.4243 ± 1.87
LLRR	0.7043 ± 1.46	0.6812 ± 1.53	0.6442 ± 1.64	0.5230 ± 2.01	0.4195 ± 1.78
GLRR	0.7102 ± 1.67	0.6854 ± 1.85	0.6532 ± 1.49	0.5694 ± 2.38	0.4832 ± 2.05
GODEC+	0.7156 ± 1.87	0.7001 ± 1.47	0.6588 ± 1.84	0.5854 ± 2.13	0.5154 ± 1.21
RSS-LRR	0.7264 ± 1.31	0.7113 ± 1.44	0.7015 ± 1.18	0.6103 ± 1.42	0.5417 ± 1.61

COIL100: it has 7200 images of 100 objects, which amount to 72 images for each object with each image taken at pose intervals of 5 degrees.

LFW: the Labeled Faces in the Wild (LFW) dataset originally contains more than 13000 face images, mainly from Internet sources. However, 2484 face images were extracted from 38 classes in our experiments due to fewer samples in some categories. Each image was resized to 64×64 pixels, yielding 4096 features per image.

4.4.2. Experimental Results. From the classification results in Tables 8 and 9, it can be noticed that RSS-LRR performance is consistently better than that of the compared methods. For instance, while RSS-LRR's performance of **0**.6052 on the COIL100 dataset corrupted with 6×6 block occlusion (Table 8) is slightly better than that of the second-best GODEC+ by over 1%, it is far better by over 2% under 8×8 block occlusion. Similar results are obtained on the LFW dataset (Table 9), where the proposed method's performance is also better than that of the GODEC+ method, which follows closely, except that more margin of over 4% is obtained under 6×6 block occlusion.

From the clustering results displayed in Tables 10 and 11, the following can be observed:

(i) Although all methods obtained comparative performances on both datasets, their accuracies degrade significantly by increasing noise. This, however, is not unexpected because more corruption levels would mean that more discriminative data features are destroyed, making it difficult to accurately group similar data samples in the same cluster.

(ii) Furthermore, while the relatively newer method GODEC+ shows more robustness to noise than the older methods, its clustering accuracy on clean data is slightly lower than that of the other methods on the LFW dataset, perhaps due to class imbalance in this dataset.

(iii) Overall, RSS-LRR's robustness to noise grows stronger with the increase in noise level. For example, its performance on COIL100 under 0% noise is merely 1% better than that of its closest competitor GODEC+, but it is over 2% better under 20% noise. The same can be said on the LFW dataset, where RSS-LRR's clustering accuracy is only about 2% better than that of LRR on clean data, whereas it is more than 4% better than that of GODEC+, which is the closest result under the 20% noise level.

Additionally, in order to demonstrate more novelty, RSS-LRR's clustering and classification performances on large-scale datasets are further compared with those of two more recent state-of-the-art (SOTA) methods, namely, nonnegative sparse discriminative low-rank representation (NSDLRR) [47] and low-rank and collaborative representations for hyperspectral anomaly detection (LRCRD) [48]. From the classification results displayed in Figure 7, it can be observed that all three methods obtain correlative results on both datasets. However, RSS-LRR shows more robustness, especially on the LFW dataset. Similarly, the clustering results are shown in Figure 8, and the results are also close with those of the proposed method displaying the best overall performance.

TABLE 11: Clustering results concerning the accuracy of different algorithms on the LFW dataset.

	0%	5%	10%	15%	20%
LRR	0.7605 ± 1.53	0.7305 ± 1.84	0.7122 ± 1.76	0.6263 ± 1.75	0.3547 ± 1.65
LRRLC	0.7512 ± 1.53	0.7298 ± 1.26	0.7099 ± 1.72	0.6310 ± 1.27	0.3864 ± 2.14
LLRR	0.7518 ± 1.32	0.7431 ± 2.17	0.7169 ± 2.18	0.6397 ± 1.74	0.4311 ± 1.98
GLRR	0.7560 ± 1.03	0.7457 ± 1.97	0.7098 ± 1.53	0.6780 ± 1.85	0.5256 ± 1.87
GODEC+	0.7491 ± 1.27	0.7434 ± 1.92	0.7156 ± 1.44	0.6856 ± 1.79	0.5413 ± 1.34
RSS-LRR	0.7832 ± 1.36	0.7567 ± 1.72	0.7269 ± 1.84	0.6952 ± 1.68	0.5893 ± 1.76

(a)

(b)
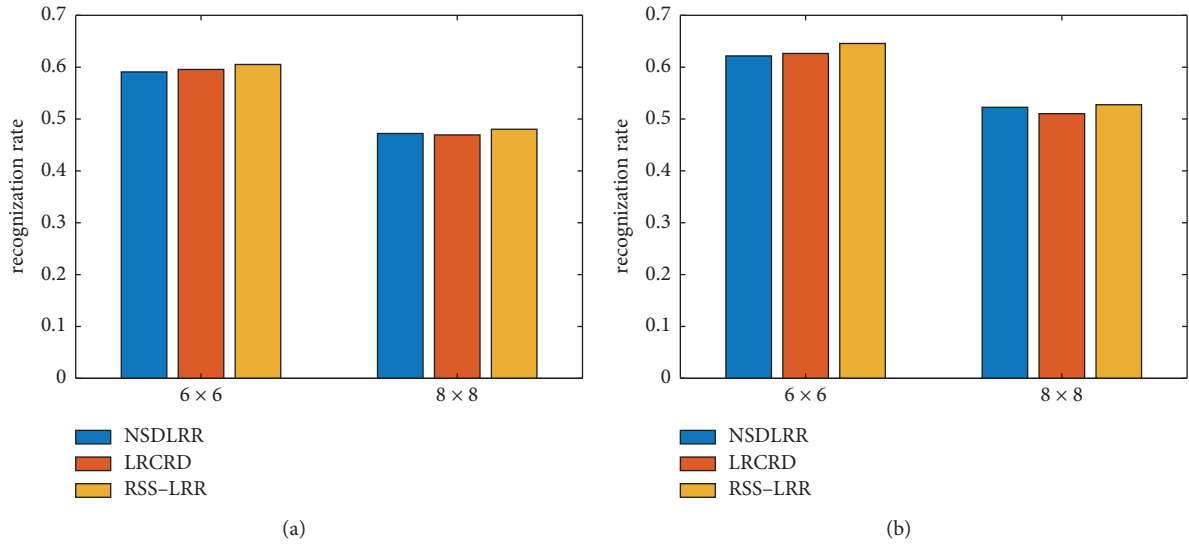
FIGURE 7: Classification results concerning the accuracy of different SOTA methods on (a) COIL100 and (b) LFW datasets under various block occlusion levels.

(a)

(b)
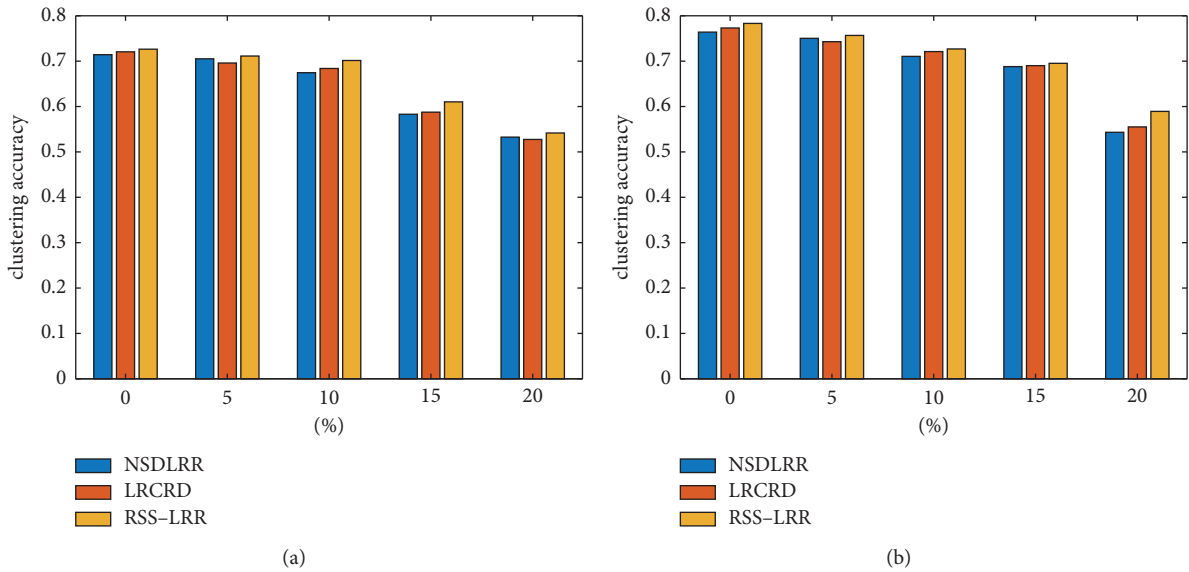
FIGURE 8: Clustering results concerning the accuracy of different SOTA methods on (a) COIL100 and (b) LFW datasets under various noise levels.

4.5. Convergence Study.

Based on established standards of inexact augmented Lagrange multiplier (IALM) optimization strategy in several pieces of literature [5, 49–51], the objective function values of the J, Z, and E subproblems were expected to decrease monotonically in each iteration until convergence. Besides, the Z subproblem is known to have a closed-form solution [51] since it has a convex function. Also, the convergence of the J and E subproblems is confirmed by references [36, 52–54], respectively. Thus, a slow convergence is supposed when everything is put
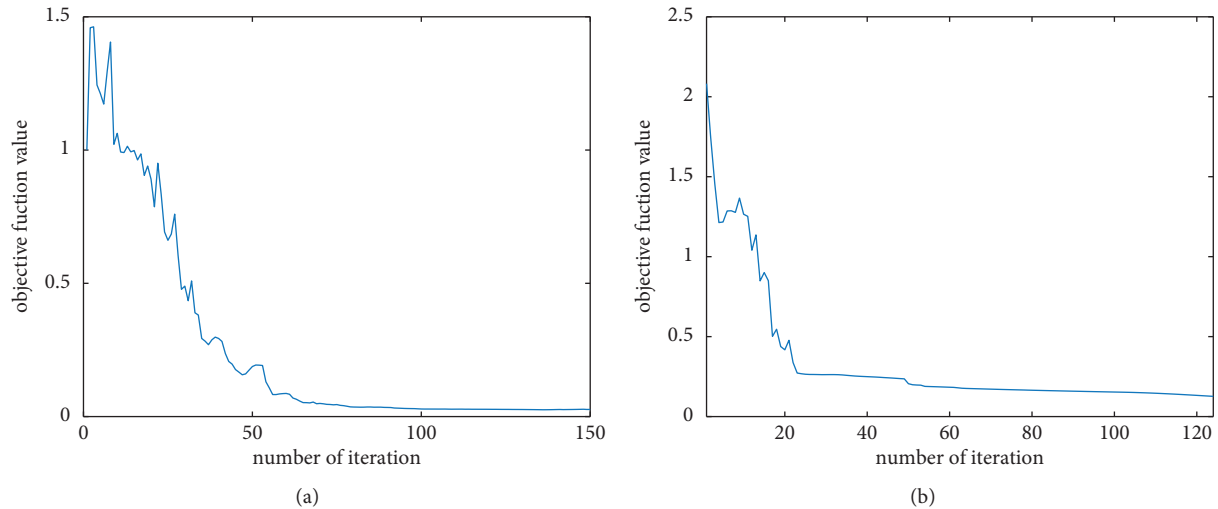
(a)

(b)

Figure 9: Convergence curve of the proposed algorithm on (a) COIL100 and (b) ORL corrupted with 6×6 block occlusion.

together. Luckily, as shown in Figure 9, the proposed algorithm has a strong convergence property, as it converges within 150 iterations on COIL100 and ORL datasets.

5. Conclusion

In this paper, we propose a recursive sample scaling low-rank representation method named "RSS-LRR." Different from the existing methods, each data sample's importance is estimated by introducing a cosine scaling factor. This scaling factor is used to extract each sample's relationship with the low-rank representation's principal components in the feature space. Thus, our proposed model can effectively reduce the noise effect by iteratively reducing the importance of noisy samples in learning the robust low-rank matrix. Several experimental results on well-known benchmark datasets demonstrate that RSS-LRR performs better in clustering and classification tasks than the state-of-the-art methods. It includes various experiments conducted on gross corrupted data. Therefore, we will extend RSS-LRR's idea to multiview data in future work.

Data Availability

The datasets used in this study are open benchmark datasets that are allowed for use in research. The following is a description and links to each one of them. Each can be accessed using the corresponding link. ORL contains face images of ten individuals, with each of them contributing forty distinct images under various conditions such as facial details and different facial expressions: http://cam-orl.co.uk/facedatabase.html. CMU PIE is a face image repository with images of sixty-eight individuals with different settings. It includes thirteen different poses, four different expressions, and forty-two different illuminations: https://www.cs.cmu.edu/afs/cs/project/PIE/MultiPie/Multi-Pie/Home.html. COIL20 is an object image dataset consisting of 20 separate objects. Each object contributes 72 grayscale images, amounting to a total of 1440 images: https://www.cs.

columbia.edu/CAVE/software/softlib/coil-20.php. COIL100 has 7200 images of 100 objects, which amount to 72 images for each object with each image taken at pose intervals of 5 degrees: https://www.kaggle.com/jessicali9530/coil100. The Labeled Faces in the Wild (LFW) dataset originally contains more than 13000 face images, mainly from Internet sources. However, 2484 face images were extracted from 38 classes in our experiments due to fewer samples in some categories. Each image was resized to 64×64 pixels, yielding 4096 features per image: http://vis-www.cs.umass.edu/lfw/.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this study.

Authors' Contributions

Wenyun Gao and Stanley Abhadiomhen conceptualized the research idea and reviewed and edited the paper. Xiaoyun Li and Sheng Dai were involved in data investigation. Sheng Dai ran the software. Stanley Abhadiomhen wrote the original draft. Wenyun Gao, Xinghui Yin, and Stanley Abhadiomhen supervised the project. Wenyun Gao was involved in funding acquisition.

Acknowledgments

References

[1] J. Jian Yang, D. Zhang, A. F. Frangi, and J. Y. Jing-Yu 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, 2004.

[2] E. J. Cand`es, X. Li, Y. Ma, and J. Wright, "Robust principal component analysis?" *Journal of the ACM*, vol. 58, no. 3, p. 11, 2011.

[3] J. Wang, J. Yang, K. Yu, F. Lv, T. Huang, and Y. Gong, "Locality-constrained linear coding for image classification," in *Proceedings of the Computer Vision on Pattern Recognition*, pp. 3360–3367, San Francisco, CA, USA, June 2010.

[4] E. Elhamifar and R. Vidal, "Sparse subspace clustering: algorithm, theory, and applications," *TPAMI*, vol. 35, pp. 2765–2781, 2012.

[5] G. Liu, Z. Lin, S. Yan, J. Sun, Y. Yu, and Y. Ma, "Robust recovery of subspace structures by low-rank representation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 1, pp. 171–184, 2013.

[6] F. Nie and H. Huang, "Subspace clustering via new discrete group structure constrained low-rank model," in *In Proceedings of the IJCAI*, New York, NY, USA, July 2016.

[7] L. Parsons, E. Haque, and H. Liu, "Subspace clustering for high dimensional data," *ACM Sigkdd Explorations Newsletter*, vol. 6, no. 1, pp. 90–105, 2004.

[8] X. Lu, Y. Wang, and Y. Yuan, "Graph-regularized low-rank representation for destriping of hyperspectral images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 51, no. 7, pp. 4009–4018, 2013.

[9] S. E. Abhadiomhen, Z. Wang, and X. Shen, "Coupled low rank representation and subspace clustering," *Applied Intelligence*, vol. 1-17, 2021.

[10] S. Du and Y. Ma, "Graph regularized compact low rank representation for subspace clustering," *Knowledge-Based Systems*, vol. 118, 2016.

[11] R. Liu, Z. Lin, F. De la Torre, and Z. Su, "Fixed-rank representation for unsupervised visual learning," in *Proceedings of the Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 598–605, Providence, RI, USA, June 2012.

[12] L. Wei, X. Wang, and J. Yin, "Self-regularized fixed-rank representation for subspace Segmentation," *Information Sciences*, vol. 412, 2017.

[13] K. Tang, R. Liu, Z. Su, and J. Zhang, "Structure-constrained low-rank representation," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 25, no. 12, pp. 2167–2179, 2014.

[14] S. Xiao, M. Tan, D. Xu, and Z. Y. Dong, "Robust kernel low-rank representation," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 27, no. 11, pp. 2268–2281, 2015.

[15] G. Liu and S. Yan, "Latent low-rank representation for subspace segmentation and feature extraction," in *Proceedings of the International Conference on Computer Vision*, pp. 1615–1622, Barcelona, Spain, November 2011.

[16] J. Wright, A. Ganesh, S. Rao, Y. Peng, and Y. Ma, "Robust principal component analysis: exact recovery of corrupted low-rank matrices via convex optimization," in *Advances in Neural Information Processing Systems*, pp. 2080–2088, Hyatt Regency Vancouver, Vancouver, Canada, December 2009.

[17] J. J. Gerbrands, "On the relationships between SVD, KLT and PCA," *Pattern Recognition*, vol. 14, no. 1, pp. 375–381, 1981.

[18] Y Chang, L. Yan, and S. Zhong, "Transformed low-rank model for line pattern noise removal," in *Proceedings of the International Conference on Computer Vision*, pp. 1726–1734, Venice, Italy, October 2017.

[19] Z. Gao and D. Li, "Face recognition with contiguous occlusion based on image segmentation," in *Proceedings of the IEEE Conference on Image Process*, pp. 156–159, Shanghai, China, July 2015.

[20] X. Peng, R. Yan, B. Zhao, H. Tang, and Z. Yi, "Fast low rank representation based spatial pyramid matching for image classification," *Knowledge-Based Systems*, vol. 90, pp. 14–22, 2015.

[21] Z. Zhang, Y. Xu, and L. Shao, "Discriminative block-diagonal representation learning for image recognition," *IEEE*, vol. 29, no. 7, pp. 3111–3125, 2017.

[22] Y. Peng, J. Suo, Q. Dai, and W. Xu, "Reweighted low-rank matrix recovery and its application in image restoration," *IEEE Transactions on Cybernetics*, vol. 44, no. 12, pp. 2418–2430, 2014.

[23] B. Bing-Kun Bao, G. Guangcan Liu, C. Changsheng Xu, and S. Shuicheng Yan, "Inductive robust principal component analysis," *IEEE Transactions on Image Processing*, vol. 21, no. 8, pp. 3794–3800, 2012.

[24] Z. Zhang, F. Li, M. Zhao, L. Zhang, and S. Yan, "Joint low-rank and sparse principal feature coding for enhanced robust representation and visual classification," *IEEE Transactions on Image Processing*, vol. 25, no. 6, pp. 2429–2443, 2016.

[25] C. Tang, X. Zhu, X. Liu et al., "Learning a joint affinity graph for multiview subspace clustering," *IEEE Transactions on Multimedia*, vol. 21, no. 7, pp. 1724–1736, 2019.

[26] Z. Zhang, J. Ren, S. Li, R. Hong, Z. Zha, and M. Wang, "Robust subspace discovery by block-diagonal adaptive locality-constrained representation," in *Proceedings of the 27th ACM International Conference on Multimedia*, pp. 1569–1577, Nice, France, October 2019.

[27] G. Liu, Z. Zhang, Q. Liu, and H. Xiong, "Robust subspace clustering with compressed data," *IEEE Transactions on Image Processing*, vol. 28, no. 10, pp. 5161–5170, Oct. 2019.

[28] C. Tang, X. Liu, X. Zhu et al., "Feature selective projection with low-rank embedding and dual laplacian regularization," *IEEE Transactions on Knowledge and Data Engineering*, vol. 32, no. 9, p. 1, 2019.

[29] J. Ren, Z. Zhang, S. Li et al., "Learning hybrid representation by robust dictionary learning in factorized compressed space," *IEEE Transactions on Image Processing*, vol. 29, pp. 3941–3956, 2020.

[30] Z. Fu, Y. Zhao, D. Chang, and Y. Wang, "A hierarchical weighted low-rank representation for image clustering and classification," *Pattern Recognition*, vol. 112, Article ID 107736, 2021.

[31] C. Tang, X. Zheng, X. Liu et al., "Cross-view locality preserved diversity and consensus learning for multi-view unsupervised feature selection," *IEEE Transactions on Knowledge and Data Engineering*, p. 1, 2021.

[32] Y. Li, J. Liu, Z. C. Li, Y. Zhang, H. Q. Lu, and S. D. Ma, "Learning low-rank representations with classwise block-diagonal structure for robust face recognition," in *Proceedings of the AAAI Conference on Artificial Intelligence*, pp. 2810–2816, Quebec, Canada, July 2014.

[33] M. Fazel, H. Hindi, and S. P. Boyd, "A rank minimization heuristic with application to minimum order system approximation," *Proceedings-American Control Conference*, vol. 6, pp. 4734–4739, 2001.

[34] V. Chandrasekaran, B. Recht, P. A. Parrilo, and A. S. Willsky, "The convex geometry of linear inverse problems," *Foundations of Computational Mathematics*, vol. 12, no. 6, pp. 805–849, 2012.

[35] V. Larsson and C. Olsson, "Convex low rank approximation," *International Journal of Computer Vision*, vol. 120, no. 2, pp. 194–214, 2016.

[36] Z. Lin, M. Chen, L. Wu, and Y. Ma, "The augmented lagrange multiplier method for exact recovery of corrupted low-rank matrices," Technical Report UILU-ENG-09-2215, University of Illinois Urbana-Champaign, Champaign, IL, 2009.

[37] B. Recht, M. Fazel, and P. A. Parrilo, "Guaranteed minimum-rank solutions of linear matrix equations via nuclear norm minimization," *SIAM Review*, vol. 52, no. 3, pp. 471–501, 2010.

[38] X. Xiaowei Zhou, C. Can Yang, and W. Weichuan Yu, "Moving object detection by detecting contiguous outliers in the low-rank representation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 3, pp. 597–610, 2013.

[39] J. Yang, W. Yin, Y. Zhang, and Y. Wang, "A fast algorithm for edge-preserving variational multichannel image restoration," *SIAM Journal on Imaging Sciences*, vol. 2, no. 2, pp. 569–592, 2009.

[40] G. Liu, Z. Lin, and Y. Yu, "Robust subspace segmentation by low-rank representation," in *Proceedings of the International Conference on Machine Learning*, pp. 663–670, Haifa, Israel, June 2010.

[41] K. Guo, X. Xu, and D. Tao, "Discriminative GoDec+ for classification," *IEEE Transactions on Signal Processing*, vol. 65, no. 13, pp. 3414–3429, 2017.

[42] Y. Yigang Peng, A. Ganesh, J. Wright, W. Wenli Xu, and Y. Yi Ma, "RASL: robust alignment by sparse and low-rank decomposition for linearly correlated images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 34, no. 11, pp. 2233–2246, 2012.

[43] S. Gu, L. Zhang, W. Zuo, and X. Feng, "Weighted nuclear norm minimization with application to image denoising," in *Proceedings of the Computer Vision on Pattern Recognition*, pp. 2862–2869, Columbus, OH, USA, June 2014.

[44] Z. Ding, M. Shao, and Y. Fu, "Low-rank embedded ensemble semantic dictionary for zero-shot learning," in *Proceedings of the Computer Vision on Pattern Recognition*, pp. 6005–6013, Honolulu, HI, USA, July 2017.

[45] S. E. Abhadiomhen, Z. Wang, X. Shen, and J. Fan, "Multiview common subspace clustering via coupled low rank representation," *ACM Transactions on Intelligent Systems and Technology*, vol. 12, no. 4, pp. 1–25, 2021.

[46] Y. Lu, Z. Lai, Y. Xu, X. Li, D. Zhang, and C. Yuan, "Low-rank preserving projections," *IEEE Transactions on Cybernetics*, vol. 46, no. 8, pp. 1900–1913, 2016.

[47] X. Hou, C. Chen, S. Zhou, and J. Li, "Robust face recognition based on non-negative sparse discriminative low-rank representation," in *Proceedings of the 2018 Chinese Control and Decision Conference (CCDC)*, pp. 5579–5583, CCDC), Shenyang, China, July 2018.

[48] Z. Wu, H. Su, and Q. Du, "Low-rank and collaborative representation for hyperspectral anomaly detection," in *Proceedings of the IGARSS 2019–2019 IEEE International Geoscience and Remote Sensing Symposium*, pp. 1394–1397, 2019.

[49] W. Gao, S. Dai, S. E. Abhadiomhen, W. He, and X. Yin, "Low rank correlation representation and clustering," *Scientific Programming*, vol. 2021, Article ID 6639582, 12 pages, 2021.

[50] Z. Y. Wang, S. E. Abhadiomhen, Z. F. Liu, X. J. Shen, W. Y. Gao, and S. Li, "Multi-view Intrinsic low-rank representation for robust face recognition and clustering," *IET Image Processing*, vol. 15, no. 14, pp. 3573–3584, 2021.

[51] Z. Yu, G. Zhang, J. Chen et al., "Toward noise-resistant graph embedding with subspace clustering information," *IEEE Transactions on Cybernetics*, pp. 1–13, 2021.

[52] Z. Lin, M. Chen, and Y. Ma, "The augmented lagrange multiplier method for exact recovery of corrupted low-rank matrices," 2010, https://arxiv.org/abs/1009.5055.

[53] J. Cai, E. J. Candès, and Z. Shen, "A singular value thresholding algorithm for matrix completion," *SIAM Journal on Optimization*, vol. 20, pp. 1956–1982, 2010.

[54] J. Yang, W. Yin, Y. Zhang, and Y. Wang, "A fast algorithm for edge-preserving variational multichannel image restoration," *SIAM Journal on Imaging Sciences*, vol. 2, no. 2, pp. 569–592, 2009.