

INCOHERENT DICTIONARY LEARNING WITH LOCALITY CONSTRAINED LOW-RANK REPRESENTATION FOR IMAGE CLASSIFICATION

Nguyen Hoang Vu*, Tran Quoc Cuong

Tien Giang University

ARTICLE INFO	ABSTRACT
Received: 20/02/2024	Low-rank representation (LRR) plays a significant role in image classification tasks due to its ability to capture the underlying structure and variations in image data. However, traditional low-rank representation-based dictionary learning methods struggle to leverage discriminative information effectively. To tackle this issue, we propose an incoherent dictionary learning approach with locality-constrained low-rank representation (LCLRR-IDL) for image classification. Firstly, we introduce low-rank representation to handle potential data contamination in both training and test sets. Secondly, we integrate a locality constraint to recognize the intrinsic structure of the training data, ensuring similar samples have similar representations. Thirdly, we develop a compact incoherent dictionary with local constraints in the low-rank representation to classify images, even in the presence of corruption. Experimental results on public databases validate the effectiveness of our approach.
Revised: 28/3/2024	
Published: 29/3/2024	
KEYWORDS	
Image Classification	
Low Rank Representation	
Locality Constraint	
Dictionary Learning	
Incoherent Dictionary	

HỌC TỪ ĐIỂN KHÔNG MẠCH LẠC VỚI RÀNG BUỘC CỤC BỘ ĐẠI DIỆN HẠNG THẤP TRONG PHÂN LOẠI HÌNH ẢNH

Nguyễn Hoàng Vũ*, Trần Quốc Cường

Trường Đại học Tiền Giang

THÔNG TIN BÀI BÁO	TÓM TẮT
Ngày nhận bài: 20/02/2024	Biểu diễn hạng thấp (LRR) đóng một vai trò quan trọng trong các nhiệm vụ phân loại hình ảnh do khả năng nắm bắt cấu trúc cơ bản và các biến thể trong dữ liệu hình ảnh. Tuy nhiên, các phương pháp học từ điển dựa trên biểu diễn hạng thấp thường gặp khó khăn trong việc tận dụng thông tin phân biệt trong hình ảnh một cách hiệu quả. Để giải quyết vấn đề này, chúng tôi đề xuất một phương pháp học từ điển không mạch lạc với ràng buộc cục bộ trong đại diện hạng thấp (LCLRR-IDL) để phân loại hình ảnh. Đầu tiên, chúng tôi giới thiệu cách biểu diễn hạng thấp để xử lý nhiễu trong dữ liệu huấn luyện và kiểm tra. Thứ hai, chúng tôi kết hợp ràng buộc cục bộ để nhận biết cấu trúc đa dạng nội tại của dữ liệu huấn luyện, đảm bảo các mẫu tương tự có cách biểu diễn tương tự nhau. Thứ ba, chúng tôi huấn luyện một từ điển không mạch lạc nhỏ gọn với các ràng buộc cục bộ trong biểu diễn hạng thấp để phân loại hình ảnh. Kết quả thử nghiệm trên cơ sở dữ liệu tiêu chuẩn đã xác nhận tính hiệu quả của phương pháp đề xuất.
Ngày hoàn thiện: 28/3/2024	
Ngày đăng: 29/3/2024	
TỪ KHÓA	
Phân loại hình ảnh	
Đại diện hạng thấp	
Ràng buộc cục bộ	
Học từ điển	
Từ điển không mạch lạc	

DOI: <https://doi.org/10.34238/tnu-jst.9733>* Corresponding author. Email: nguyenhoangvu@tgu.edu.vn

1. Introduction

In recent years, dictionary learning (DL) has garnered increasing attention within the image classification and signal processing communities. Dictionary learning algorithms have emerged as prominent methods in the field of image processing, showcasing the superiority of learned dictionaries over pre-designed ones like wavelets in image reconstruction tasks. This is because dictionaries learned directly from training images can capture more complex and diverse image structures, leading to more accurate and faithful reconstructions compared to fixed, pre-designed dictionaries. The effective application of DL in image restoration tasks has consequently spurred its adoption in image classification problems. Recent research has seen a rise in the development of discriminative dictionary learning (DDL) methods that make use of label information to achieve enhanced classification performance. Unlike standard dictionary learning, these DDL methods aim to learn a dictionary by optimizing an objective function that integrates both reconstructive and discriminative terms. There are several great DDL methods such as LC-KSVD [1], FDDL [2], ODFDL [3], SMLFDL [4]. These methods simultaneously learn the dictionary and classifier by incorporating different classification losses on the coding vectors. While conventional dictionary learning (DL) methods excel in various classification and recognition tasks, their performance diminishes significantly when the training data is heavily contaminated by factors such as occlusion, lighting/viewpoint variations, or corruption. To overcome these limitations, several dictionary learning (DL) methods have emerged, integrating rank minimization into the sparse representation framework. These approaches have shown promising outcomes, particularly in scenarios with significant noise presence [5], [6]. In addition, the importance of local information in data has been widely recognized in various practical applications, especially in sparse coding and dictionary learning. Many studies have built a dictionary learning algorithm with local constraints to fully exploit local information in training data [7] - [9].

Previous studies have demonstrated the pivotal role of dictionary quality in bolstering the efficacy of sparse representation methods. Researchers have dedicated significant efforts towards learning high-quality dictionaries. Dictionary learning techniques typically prioritize enhancing both the reconstruction accuracy and discriminative power of the dictionary. This paper introduces a novel algorithm called Locality Constrained Low Rank Representation - based Incoherent Dictionary Learning (LCLRR-IDL) for image classification. The proposed algorithm incorporates a locality constraint term to effectively capture the intrinsic manifold structure of the training data. By enforcing this constraint, similar samples are encouraged to have similar representations. Compared to other dictionary learning methods, we explicitly incorporate an incoherent penalty in the dictionary learning model to make the dictionary more discerning. Contaminated images can be recovered during our dictionary learning process. The main contributions of this paper are summarized as follows:

(1) We adopt low-rank and sparse representation to handle potential data contamination in both training and test images.

(2) By integrating the locality constrained term, this approach promotes the similarity between samples in terms of their representations. This ensures that similar samples are given comparable representations. The resulting learned representations can be used directly for classification purposes, without the need for further processing or feature extraction.

(3) We learn a compact incoherent dictionary with local constraints in the low-rank representation to classify images even in the presence of corruption.

The rest of this paper is outlined as follows: Section 2 introduces the proposed methodology. Section 3 illustrates the Experimental Results and discussions. Finally, Section 4 concludes this paper.

2. Methodology

2.1. Low-Rank and Sparse Representation

Liu *et al.* [10] presented a low-rank representation (LRR) method to recover the subspace structures from data that reside on a union of multiple subspaces. LRR is based on the assumption that all data are sufficiently sampled from multiple low dimensional subspaces embedded in a high-dimensional space. Assume that data samples X are drawn from a union of many subspaces, the LRR model aims to seek the low-rank representation Z and the sparse noises E based on the given dictionary D . Specifically, LRR is formulated as the following rank minimization problem:

$$\min_{Z,E} \text{rank}(Z) + \lambda \|E\|_0, \quad s.t. X = DZ + E \quad (1)$$

Leveraging both low-rank and sparse representation, the method known as Learning Low-Rank and Sparse Representation (LRSR) [11] is formulated as follows:

$$\min_{Z,E} \text{rank}(Z) + \lambda \|E\|_0 + \beta \|Z\|_0, \quad s.t. X = DZ + E \quad (2)$$

Given the non-convex nature of the problem above, a common approach is to replace the l_0 -norm with the l_1 -norm and the rank function with the nuclear norm. Consequently, the optimization problem (2) is relaxed into the following formulation:

$$\min_{Z,E} \|Z\|_* + \lambda \|E\|_1 + \beta \|Z\|_1, \quad s.t. X = DZ + E \quad (3)$$

The nuclear norm $\|Z\|_*$ defined as the sum of all singular values of Z , while $\|\cdot\|_1$ represents the l_1 -norm, which is the sum of the absolute values of entries in the given matrix. Parameters λ and β control the sparsity of sparse noise E and sparse representation Z , respectively, and D is the dictionary that linearly spans the data space.

2.2. LCLRR-IDL Method

2.2.1. Definition of LCLRR-IDL

Given a training dataset $X = [X_1, X_2, \dots, X_c] \in \mathbb{R}^{d \times n}$, where c is the number of classes, d denotes the feature dimension, and $X_i \in \mathbb{R}^{d \times n_i}$ is the samples from class i which has n_i samples. From X , we learn a discriminative dictionary D and the coding coefficient Z , which is utilized to future classification task. The objective function is designed to facilitate learning an incoherent dictionary and low-rank representation while imposing locality constraints. This is achieved by explicitly integrating a correlation penalty into the dictionary learning model:

$$\min_{D,Z,E} \|Z\|_* + \lambda \|E\|_{21} + \beta \|Z\|_1 + \alpha \|M \odot Z\|_1 + \gamma \|D^T D - I\|_F^2 \quad s.t. X = DZ + E \quad (4)$$

The term $\|M \odot Z\|_1$ represents a locality-constrained term aimed at uncovering local geometric information [12]. It serves to enhance the sparsity of the objective function, which proves advantageous for classification tasks, particularly when employing a sparse graph that delineates locality relationships. \odot denotes the Hadamard product and $M = \|x_i - x_j\|_2^2$; $\|D^T D - I\|_F^2$ is the incoherent term. Minimizing the incoherent term guarantees that the dictionary can efficiently represent the input samples and achieve higher accuracies for classification tasks. The l_{21} -norm of E is used to model sample-specific corruptions and outliers. There are four regularization parameters ($\lambda > 0$, $\beta > 0$, $\alpha > 0$ and $\gamma > 0$) that reflect relative contributions of each term.

2.2.2. The Optimization of LCLRR-IDL

In this section, we adopt the high efficiency Augmented Lagrange Multiplier (ALM) method [13] to solve the proposed LCLRR-IDL. Firstly, we introduce three auxiliary variables J , L and Q to make the problem (4) become easily solvable. Thus, we convert problem (4) into the following equivalent optimization problem:

$$\begin{aligned} \min_{D,Z,E,J,L,Q} \|J\|_* + \lambda \|E\|_{2,1} + \beta \|L\|_1 + \alpha \|M \circ Q\|_1 + \gamma \|D^T D - I\|_F^2 \\ \text{s. t. } X = DZ + E, Z = J, Z = L, Z = Q \end{aligned} \quad (5)$$

The corresponding augmented Lagrangian function for equation (5) can be written as:

$$\begin{aligned} \mathcal{L}(J, Z, L, Q, E, D, Y_1, Y_2, Y_3, Y_4, \mu) = \|J\|_* + \lambda \|E\|_{2,1} + \beta \|L\|_1 + \alpha \|M \circ Q\|_1 + \gamma \|D^T D - I\|_F^2 \\ - I\|_F^2 + \langle Y_1, X - DZ - E \rangle + \langle Y_2, Z - J \rangle + \langle Y_3, Z - L \rangle + \langle Y_4, Z - Q \rangle + \frac{\mu}{2} (\|X - DZ - E\|_F^2 + \\ \|Z - J\|_F^2 + \|Z - L\|_F^2 + \|Z - Q\|_F^2) \end{aligned} \quad (6)$$

where Y_1, Y_2, Y_3, Y_4 are Lagrange multipliers, $\mu > 0$ is the penalty parameter, and $\langle A, B \rangle$ is trace of $A^T B$. We iteratively solve problem (6) by updating J, Z, L, Q, E and D once at a time. The detailed procedures are as follows:

(1) Updating J : Fix the other variables and solve the following problem:

$$\begin{aligned} J^{k+1} &= \arg \min_J \|J\|_* + \langle Y_2^k, Z^k - J \rangle + \frac{\mu^k}{2} \|Z^k - J\|_F^2 \\ &= \arg \min_J \frac{1}{\mu^k} \|J\|_* + \frac{1}{2} \left\| J - \left(Z^k + \frac{Y_2^k}{\mu^k} \right) \right\|_F^2 \\ &= US_{\frac{1}{\mu^k}}[\Sigma]V^T \end{aligned} \quad (7)$$

where $(U, \Sigma, V^T) = SVD\left(Z^k + \frac{Y_2^k}{\mu^k}\right)$, and $S_\epsilon[\cdot] = \text{sign}(x)(|x| - \epsilon)$ is the soft - thresholding (shrinkage) operator.

(2) Updating Z : Fix the other variables and solve the following problem:

$$\begin{aligned} Z^{k+1} &= \arg \min_Z \langle Y_1^k, X - D^k Z - E^k \rangle + \langle Y_2^k, Z - J^{k+1} \rangle + \langle Y_3^k, Z - L^k \rangle + \langle Y_4^k, Z - Q^k \rangle + \\ &\quad \frac{\mu}{2} (\|X - D^k Z - E^k\|_F^2 + \|Z - J^{k+1}\|_F^2 + \|Z - L^k\|_F^2 + \|Z - Q^k\|_F^2) \\ &= \left[(D^k)^T D^k + 2I \right]^{-1} \left[(D^k)^T (X - E^k) + J^{k+1} + L^k + Q^k + \frac{(D^k)^T (Y_1^k - Y_2^k - Y_3^k)}{\mu^k} \right] \end{aligned} \quad (8)$$

(3) Updating L : Fix the other variables and solve the following problem:

$$\begin{aligned} L^{k+1} &= \arg \min_L \beta \|L\|_1 + \langle Y_3^k, Z^{k+1} - L \rangle + \frac{\mu^k}{2} (\|Z^{k+1} - L\|_F^2) \\ &= \arg \min_L \frac{\beta}{\mu^k} \|L\|_1 - \frac{1}{2} \left\| L - \left(Z^{k+1} + \frac{Y_3^k}{\mu^k} \right) \right\|_F^2 \\ &= S_{\frac{\beta}{\mu^k}} \left[Z^{k+1} + \frac{Y_3^k}{\mu^k} \right] \end{aligned} \quad (9)$$

(4) Updating Q : Fix the other variables and solve the following problem,

$$\begin{aligned} Q^{k+1} &= \arg \min_Q \alpha \|M \circ Q\|_1 + \langle Y_4^k, Z^{k+1} - Q \rangle + \frac{\mu^k}{2} (\|Z^{k+1} - Q\|_F^2) \\ &= \arg \min_Q \alpha \|M \circ Q\|_1 + \frac{1}{2} \left\| Q - \left(Z^{k+1} + \frac{Y_4^k}{\mu^k} \right) \right\|_F^2 \end{aligned} \quad (10)$$

The elementwise strategy allows for the updating of each element individually. For the (i, j) element Q_{ij} , the optimal solution of problem (10) is:

$$\begin{aligned} Q_{ij}^{k+1} &= \arg \min_{Q_{ij}} \alpha M_{ij} |Q_{ij}| + \frac{\mu^k}{2} \|Q_{ij} - R_{ij}\|_F^2 \\ &= S_{\frac{\alpha M_{ij}}{\mu^k}} \left(Z_{ij}^{k+1} + \left(\frac{Y_4^k}{\mu^k} \right)_{ij} \right) \end{aligned} \quad (11)$$

(5) Updating E : Fix the other variables and solve the following problem:

$$\begin{aligned}
E^{k+1} &= \arg \min_E \lambda \|E\|_{2,1} + \langle Y_1^k, X - D^k Z^{k+1} - E \rangle + \frac{\mu^k}{2} (\|X - D^k Z^{k+1} - E\|_F^2) \\
&= \arg \min_E \frac{\lambda}{\mu^k} \|E\|_{2,1} + \frac{1}{2} \left\| E - \left(X - D^k Z^{k+1} + \frac{Y_1^k}{\mu^k} \right) \right\|_F^2 \\
&= S_{\frac{\lambda}{\mu^k}} \left[X - D^k Z^{k+1} + \frac{Y_1^k}{\mu^k} \right] \tag{12}
\end{aligned}$$

(6) Updating D : Fix the other variables and solve the following problem:

$$D^{k+1} = \arg \min_D \gamma \|D^T D - I\|_F^2 + \langle Y_1^k, X - D Z^{k+1} - E^{k+1} \rangle + \frac{\mu^k}{2} (\|X - D Z^{k+1} - E^{k+1}\|_F^2) \tag{13}$$

The optimization problem (13) can be solved via the first-order gradient descent method [14]:

$$D_{q+1} = \Pi_D \{ D_q - \eta \nabla F(D_q) \} \tag{14}$$

where $F(D) = \gamma \|D^T D - I\|_F^2 + \langle Y_1^k, X - D Z^{k+1} - E^{k+1} \rangle + \frac{\mu^k}{2} (\|X - D Z^{k+1} - E^{k+1}\|_F^2)$, $\nabla F(D_q)$ represents the gradient of $F(D_q)$ with respect to D_q , parameter η denotes the step size, and Π_D represents the projection function that maps each column to the l_2 -norm unit ball. The optimization process of (4) is summarized in Algorithm 1.

Algorithm 1: Optimization procedure of LCLRR-IDL

Input: Training sample $X = [X_1, X_2, \dots, X_C]$, parameter λ, β, α and γ .

Initialize: $Z^0 = 0, J^0 = 0, L^0 = 0, Q^0 = 0, Y_1^0 = 0, Y_2^0 = 0, Y_3^0 = 0, \mu^0 = 10^{-5}, \mu_{max} = 10^5, \rho = 1.1, \varepsilon = 10^{-5}$ and D^0 is initialized by randomly selecting training samples.

While not converged do

1. Compute the optimal solution of J, Z, L, Q, E and D according to (7), (8), (9), (10), (12) and (13) respectively.

2. Update the Lagrange multipliers by

$$\begin{aligned}
Y_1^{k+1} &= Y_1^k + \mu^k (X - D^{k+1} Z^{k+1} - E^{k+1}); & Y_2^{k+1} &= Y_2^k + \mu^k (X^{k+1} - J^{k+1}); \\
Y_3^{k+1} &= Y_3^k + \mu^k (X^{k+1} - L^{k+1}); & Y_4^{k+1} &= Y_4^k + \mu^k (X^{k+1} - Q^{k+1}).
\end{aligned}$$

3. Update the parameter μ by $\mu^{k+1} = \min(\rho \mu^k, \mu_{max})$

4. Check the convergence conditions

$$\begin{aligned}
&\|X^{k+1} - J^{k+1}\|_\infty < \varepsilon, \|X^{k+1} - L^{k+1}\|_\infty < \varepsilon, \|X^{k+1} - Q^{k+1}\|_\infty < \varepsilon, \text{ and} \\
&\|X^{k+1} - D^{k+1} Z^{k+1} - E^{k+1}\|_\infty < \varepsilon.
\end{aligned}$$

End while

Output: Z, D , and E .

2.2.3. Classification

To classify the test data, we utilize a linear classifier in the testing phase, following the approach outlined in Ref. [11]. The coefficients Z for training samples and \hat{Z} for testing samples are obtained through Algorithm 1. Subsequently, we can learn a linear classifier W based on the coefficients of training samples and their corresponding labels:

$$\begin{aligned}
W &= \arg \min_W \|H - WZ\|_F^2 + \eta \|W\|_F^2 \\
&= HZ^T (ZZ^T + \eta I)^{-1} \tag{15}
\end{aligned}$$

where H represents the label matrix of training samples, η is the weight of the regularization term. The label for test sample i is determined by:

$$k = \arg \max_k W^* \hat{Z}_i \tag{16}$$

where k is corresponding to the classifier with the largest output.

3. Experimental Results and Discussion

In this section, we evaluate our algorithm on two datasets: AR face databases [15], and Caltech 101 object category database [16]. Our algorithm (LCLRR-IDL) is compared with related approaches including SRC [17], SLRR [18], LC-KSVD [1], FDDL [2], and LCDL-SV [7].

3.1. The AR database

The AR database comprises over 4000 images representing 126 individuals. Each individual has 26 images captured in two separate sessions, featuring different illumination and expression variations. Each session includes thirteen images, with three images featuring sunglasses, another three with scarves, and the remaining seven representing clean/neutral images with various illumination and expression changes. Fig. 1 showcases example images from this database, with each image sized at 165×120 pixels. In our experiments, we select a subset of the AR database comprising 50 men and 50 women. We convert the color images to grayscale and down-sample them by a factor of $1/3$, resulting in a feature vector dimension of 2200. Experiments are conducted in three scenarios: Sunglasses, Scarf, and Mixed, as outlined in [18]. To achieve optimal results across all scenarios, it is recommended to employ a compact dictionary with 5 items for each class. Table 1 summarizes the comparison of different methods on the AR database. We use $\lambda = 0.07, \beta = 0.15, \alpha = 0.01$ and $\gamma = 0.005$ in this experiment.



Figure 1. Example images from the AR database

Table 1. Recognition accuracy (%) on the AR database

Methods	Sunglasses	Scarf	Mixed
SRC	86.8	83.2	79.2
SLRR	87.3	83.4	82.4
LC-KSVD	85.4	78.5	83.7
FDDL	88.7	82.2	87.4
LCDL-SV	89.2	85.3	89.2
LCLRR-IDL	92.6	90.8	90.6

Table 2. Average recognition rates with different numbers of atoms on the AR face database

The number of atoms	120	240	360	480	600	720
LC-KSVD	68.8	72.4	73.1	75.4	76.2	76.8
FDDL	69.5	73.3	73.9	75.8	76.5	77.6
LCDL-SV	70.1	73.6	75.1	76.3	76.9	78.0
LCLRR-IDL	73.2	74.3	75.8	76.5	78.1	78.8

From Table 1, our approach demonstrates the highest recognition performance, surpassing SLRR by 5.3% in the sunglasses scenario, 7.4% in the scarf scenario, and 8.2% in the mixed scenario. Our method exhibits robustness against severe occlusions such as sunglasses and scarves. In contrast, LC-KSVD and FDDL show subpar performance when both training and test images suffer from significant corruption. The quality of the dictionary proves critical for learning discriminative representations in the presence of severe corruption in both training and test images. Furthermore, we compare our results with other methods using varying numbers of atoms. For the training samples, we select seven neutral images from session 1 and one corrupted image per person. This results in a total of 16 test images available for testing. Table 2 displays the average recognition rates of LC-KSVD, FDDL, LCDL-SV, and LCLRR-IDL algorithms. It's

evident that as the number of atoms increases, the average recognition rates of these algorithms also increase. Notably, the LCLRR-IDL algorithm consistently outperforms the other three dictionary learning algorithms in terms of average recognition rates.

3.2. Caltech 101 database

The Caltech 101 database is a popular dataset for object classification, featuring 102 classes (comprising 101 object classes and one background class). The number of images per category is unbalanced, ranging from 31 to 800, with a total of 9144 images in the dataset. To ensure a fair comparison, we utilize the 3000-dimensional SIFT-based features employed in LC-KSVD. Following the standard experimental protocol, we randomly select 5, 10, 15, 20, 25, and 30 samples per category for training, and test on the remaining images. The parameters are set as follows, $\lambda = 0.1, \beta = 0.1, \alpha = 0.05$ and $\gamma = 0.005$. Table 3 summarizes the classification results. As shown in Table 3, LCLRR-DL consistently outperforms other competing approaches across all cases. This highlights relying solely on the locality constraint may not ensure optimal performance. However, when combined with the proposed classification scheme, LCLRR-IDL demonstrates its superiority over other dictionary learning approaches. Examples from classes achieving high classification accuracy with 30 training images per category are illustrated in Fig. 2. Furthermore, Table 4 presents the average recognition rates of LC-KSVD, FDDL, LCDL-SV, and LCLRR-IDL algorithms using different numbers of atoms ($K = 102, 204, \dots, 612, 714$). It is evident that the LCLRR-IDL algorithm consistently achieves higher average recognition rates compared to LC-KSVD, FDDL, and LCDL-SV algorithms.

Table 3. Recognition accuracy (%) on the Caltech101 database

Methods	5	10	15	20	25	30
SRC	48.8	60.1	64.9	67.7	69.2	70.7
SLRR	49.6	59.5	66.1	68.6	71.1	73.6
LC-KSVD	54.0	63.1	67.7	70.5	72.3	73.6
FDDL	53.6	63.6	66.8	69.8	71.7	73.1
LCDL-SV	56.9	65.6	69.2	72.6	74.5	76.5
LCLRR-IDL	57.7	65.9	69.6	73.4	75.9	77.3

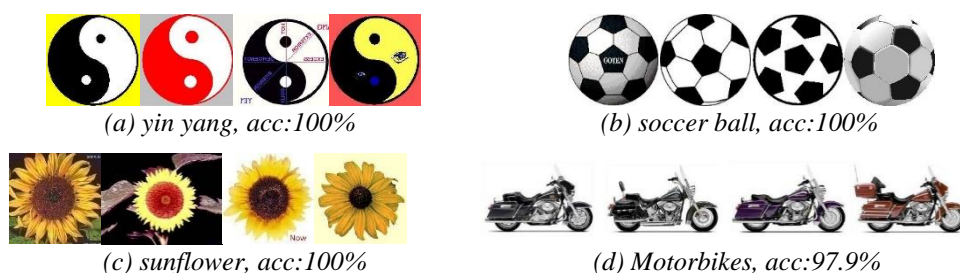


Figure 2. Example images from classes with high classification accuracy of the Caltech101

Table 4. Average recognition rates with different numbers of atoms on the Caltech 101 database

The number of atoms	102	204	306	408	510	612	714
LC-KSVD	51.2	53.0	54.4	54.6	55.3	56.4	56.9
FDDL	52.1	53.2	54.3	55.0	55.7	57.2	57.6
LCDL-SV	54.3	54.2	54.8	55.1	56.5	57.5	57.7
LCLRR-IDL	54.5	54.3	54.9	55.4	57.1	58.1	58.6

4. Conclusion

In this paper, we introduced a novel image classification method based on the dictionary learning framework utilizing the LCLRR-IDL model. By incorporating a locality constraint term

into the low-rank and sparse representation model, our proposed approach effectively leverages the intrinsic manifold structure of the training data. Additionally, we imposed incoherence constraints on the dictionary to minimize similarity between atoms associated with different classes. Consequently, class-specific dictionaries are learned from the optimization process, ensuring their independence as much as possible. Our method introduces an incoherent dictionary to simultaneously learn discriminative low-rank and sparse representations with locality constraints for both training and testing images. Experimental results validate the effectiveness and robustness of our proposed method, achieving state-of-the-art performance, particularly in scenarios where both training and testing samples are contaminated.

REFERENCES

- [1] Z. Jiang, Z. Lin, and L. S. Davis, "Label consistent K-SVD: Learning a discriminative dictionary for recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 11, pp. 2651–2664, 2013, doi: 10.1109/TPAMI.2013.88.
- [2] M. Yang, L. Zhang, X. Feng, and D. Zhang, "Fisher Discrimination Dictionary Learning for sparse representation," in *Proceedings of the IEEE International Conference on Computer Vision*, 2011, pp. 543–550, doi: 10.1109/ICCV.2011.6126286.
- [3] G. Zhang, J. Yang, Y. Zheng, Z. Luo, and J. Zhang, "Optimal discriminative feature and dictionary learning for image set classification," *Inf. Sci. (Ny)*, vol. 547, pp. 498–513, Feb. 2021, doi: 10.1016/j.ins.2020.08.066.
- [4] B.-Q. Yang, X.-P. Guan, J.-W. Zhu, C.-C. Gu, K.-J. Wu, and J.-J. Xu, "SVMs multi-class loss feedback based discriminative dictionary learning for image classification," *Pattern Recognit.*, vol. 112, Apr. 2021, Art. no. 107690, doi: 10.1016/j.patcog.2020.107690.
- [5] H. Nguyen, W. Yang, B. Sheng, and C. Sun, "Discriminative low-rank dictionary learning for face recognition," *Neurocomputing*, vol. 173, pp. 541–551, Jan. 2016, doi: 10.1016/j.neucom.2015.07.031.
- [6] H. V. Nguyen, Q. C. Tran, and T. P. Tran "Discriminative Dictionary Pair Learning for Image Classification," *J. Comput. Sci. Cybern.*, vol. 36, no. 4, pp. 347–363, 2020, doi: 10.15625/1813-9663/36/4/15105.
- [7] H. F. Yin, X. J. Wu, and S. G. Chen, "Locality Constraint Dictionary Learning with Support Vector for Pattern Classification," *IEEE Access*, vol. 7, pp. 175071–175082, 2019, doi: 10.1109/ACCESS.2019.2957417.
- [8] Y. Peng, S. Liu, X. Wang, and X. Wu, "Joint local constraint and fisher discrimination based dictionary learning for image classification," *Neurocomputing*, vol. 398, pp. 505–519, Jul. 2020, doi: 10.1016/j.neucom.2019.05.103.
- [9] J. Huang, K. Liu, and X. Li, "Locality Constrained Low Rank Representation and Automatic Dictionary Learning for Hyperspectral Anomaly Detection," *Remote Sens.*, vol. 14, no. 6, Mar. 2022, Art. no. 1327, doi: 10.3390/rs14061327.
- [10] G. Liu, Z. Lin, S. Yan, J. Sun, Y. Yu, and Y. Ma, "Robust recovery of subspace structures by low-rank representation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 1, pp. 171–184, 2013, doi: 10.1109/TPAMI.2012.88.
- [11] Y. Li, J. Liu, H. Lu, and S. Ma, "Learning robust face representation with classwise block-diagonal structure," *IEEE Trans. Inf. Forensics Secur.*, vol. 9, no. 12, pp. 2051–2062, 2014, doi: 10.1109/TIFS.2014.2361936.
- [12] L. Wei, A. Wu, and J. Yin, "Latent space robust subspace segmentation based on low-rank and locality constraints," *Expert Syst. Appl.*, vol. 42, no. 19, pp. 6598–6608, 2015, doi: 10.1016/j.eswa.2015.04.041.
- [13] Z. Lin, M. Chen, and Y. Ma, "The Augmented Lagrange Multiplier Method for Exact Recovery of Corrupted Low-Rank Matrices," *Journal of Structural Biology*, Sep. 2010, doi: 10.1016/j.jsb.2012.10.010.
- [14] T.-Y. Hung, J. Lu, Y.-P. Tan, and S. Gao, "Efficient Sparsity Estimation via Marginal-Lasso Coding," in *Proceedings of Computer Vision—ECCV 2014: 13th European Conference*, Zurich, Switzerland, September 6–12, 2014, Part IV 13, Springer, 2014, pp. 578–592.
- [15] A. M. Martinez and A. C. Kak, "PCA versus LDA," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 23, no. 2, pp. 228–233, Feb. 2001, doi: 10.1109/34.908974.

- [16] L. Fei-Fei, R. Fergus, and P. Perona, "Learning generative visual models from few training examples: An incremental bayesian approach tested on 101 object categories," in *2004 conference on computer vision and pattern recognition workshop*, IEEE, 2004, p. 178.
- [17] J. Wright, A. Y. Yang, A. Ganesh, S. S. Sastry, and Y. Ma, "Robust face recognition via sparse representation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 31, no. 2, pp. 210–227, 2009, doi: 10.1109/TPAMI.2008.79.
- [18] Y. Zhang, Z. Jiang, and L. S. Davis, "Learning structured low-rank representations for image classification," *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, 2013, pp. 676–683, doi: 10.1109/CVPR.2013.93.