Hindawi
Complexity
Volume 2019, Article ID 3251026, 14 pages
https://doi.org/10.1155/2019/3251026

Research Article

Discriminative, Competitive, and Collaborative Representation-Based Classification with l_2-Norm Regularizations

Jianping Gou,[1] **Junyu Lu**,[1] **Heping Song**,[1] **Hongxing Ma**,[2] **Weihua Ou**,[3] **and Jia Ke**[1]

[1]*School of Computer Science and Communication Engineering and Jiangsu Key Laboratory of Security Tech. for Industrial Cyberspace, Jiangsu University, Zhenjiang, Jiangsu 212013, China*
[2]*College of Electrical and Information Engineering, North Minzu University, Yinchuan, Ningxia 750021, China*
[3]*School of Big Data and Computer Science, Guizhou Normal University, Guiyang, Guizhou 550025, China*

Correspondence should be addressed to Jianping Gou; goujianping@ujs.edu.cn

Received 26 September 2019; Accepted 14 November 2019; Published 7 December 2019

Academic Editor: Ludovico Minati

Recently, collaborative representation-based classification (CRC) and its many variations have been widely applied for various classification tasks in pattern recognition. To further enhance the pattern discrimination of CRC, in this article we propose a novel extension of CRC, entitled discriminative, competitive, and collaborative representation-based classification (DCCRC). In the proposed DCCRC, the class discrimination information is fully utilized for promoting the true class of each testing sample to dominantly represent the testing sample during collaborative representation. The class discrimination information is well considered in the newly designed discriminative l_2-norm regularization that can decrease the ability of representation from the interclasses of each testing sample. Simultaneously, a competitive l_2-norm regularization is introduced to the DCCRC model with the class discrimination information with the aim of enhancing the competitive ability of representation from the true class of each testing sample. The effectiveness of the proposed DCCRC is explored by extensive experiments on the several public face databases and some real numerical UCI data sets. The experimental results demonstrate that the proposed DCCRC achieves the superior performance over the state-of-the-art representation-based classification methods.

1. Introduction

Nowadays, the linear representation-based classification (RBC) often including sparse representation-based classification (SRC) [1] and collaborative representation-based classification (CRC) [2] has attracted more and more attention in pattern recognition. In both SRC and CRC, each testing sample is linearly represented by all the training samples and always classified by the class-specific representation residuals. Due to the excellent representation-based classification performance, the RBC methods have been widely used in many classification tasks, such as image classification [3–10] and face recognition [11–19].

It has been well known that SRC with the l_1-norm regularization of representation coefficients is a very promising kind of RBC owing to its good property of sparsity and natural discrimination [1, 20, 21]. However, it has been argued that the representation-based pattern discrimination originated from the l_2-norm collaborative representation of all the training samples instead of the l_1-norm sparse representation of a few training samples, and then the standard CRC was first proposed as a general extension of SRC [2]. Specifically, using the l_2-norm regularization of representation coefficients, the effective discrimination benefits from the collaborative representation from all the class-specific training samples. Because of the efficient closed-form solution of CRC for the effective classification performance, a great many CRC extensions have been developed in recent years [6, 15, 17–19, 22–35]. Moreover, the possible reasons of the natural discrimination from CRC were detailedly analyzed from the perspective of class separability of data [18] and the probability [22]. Among the CRC methods, the general extensions are the weighted CRC using the localities of data as the weights that constrained the collaborative representation coefficients [17, 27–29, 31]. Since collaborative representation has the efficient and effective classification performance, several two-phase collaborative representation-based classification methods have been designed in [30–33, 36]. Moreover, such two-phase collaborative representation-based classification also has the property of sparsity for enhancing the ability of pattern discrimination [30]. Using the superiorities of sparse representation and collaborative representation, the extensions of combining both were proposed for classification in [34, 35, 37, 38]. Besides, due to good latent discrimination contained in the representation, sparse representation and collaborative representation were utilized to design the effective nearest neighbor classification [39–41].

In many latest extensions of CRC, the class discrimination information of data in fact was fully employed for strengthening the power of the pattern classification [42–47]. From the point of view of probability, a probabilistic CRC (ProCRC) was developed by using the discriminative regularization of the representations between all the classes and each class [22]. Using the prior information of data the extended ProCRC (EProCRC) was proposed in [43], and using the coarse to fine representation the two-phased ProCRC was proposed in [33]. Through designing the discriminative regularization of pairs of the representations of any two classes, the new discriminative sparse representation method for classification (DSRC) was proposed in [44]. On the basis of DSRC and ProCRC, a novel discriminative CRC method was proposed to extend DSRC [45]. To overcome the issue that the phases of representation and classification in the most CRC variations are not integrated into a unified model, a collaborative and competitive representation-based classifier (CCRC) was proposed in [46]. CCRC directly includes the classification decision in its model and can enhance the training sample from each class to competitively represent each testing sample. With the aim of obtaining the similar competitive representations among all the classes, the discriminative l_2-norm regularization of the representations of all the classes except any one class was designed for proposing the competitive and collaborative representation classification method (Co-CRC) [47]. As argued in these discriminative CRC extensions above, the discriminative representation was achieved for favorable classification.

Based on the fact that the discrimination information of data can be explored for enhancing the power of pattern discrimination in collaborative representation, in this article we proposed a novel discriminative competitive and collaborative representation-based classification method (DCCRC) by using the discriminative representation among all the classes. The proposed DCCRC assumes that each class can discriminatively and competitively represent the testing samples. The discriminative and competitive collaborative representations among all the classes can be realized by two l_2-norm regularizations in the DCCRC model. One is the newly designed l_2-norm regularization of the pairs of representation from all the classes and representations from all the classes excluding any one class. The other is the competitive l_2-norm regularization of representations from all the classes excluding any one class [47]. To experimentally verify the classification performance of the proposed DCCRC, we compare it to the state-of-the-art RBC methods on several face databases and some real numerical UCI data sets. The conducted experiments show that the proposed method is effective with better classification results than the competing RBC methods. In summary, our main contributions in this article are given as follows:

(1) A new discriminative l_2-norm regularization is designed by using the representations from all the classes excluding any one class

(2) A novel discriminative, competitive, and collaborative representation is proposed for classification by considering the discrimination information of data

(3) The experimental analyses are reported for well demonstrating the effectiveness of the proposed DCCRC

The rest of this article is organized as follows. Section 2 briefly describes the related work. Section 3 detailedly presents the proposed DCCRC and then analyzes it. Section 4 reports extensive experiments to evaluate the effectiveness of the proposed DCCRC. Finally, the conclusions of this article are given in Section 5.

2. The Related Work

In this section, we briefly review some related RBC models. First of all, some commonly used notations are denoted here. We suppose that the set of all the training samples from C classes is denoted as $X = [x_1, x_2, \ldots, x_N] = [X_1, \ldots, X_C] \in R^{d\times N}$, where d is the dimensionality of the feature space and N and N_i are the numbers of all the training samples from all the classes and class i, respectively. Note that the ith column vector of X represents the training sample x_i and the subset of the training samples from class i is $X_i \in R^{d\times N_i}$. Besides, we also assume $y \in R^d$ is a given testing sample used for classification. In the liner representation-based classification, the testing sample y is approximately represented as $y \approx x_1 s_1 + x_2 s_2 + \cdots + x_N s_N = XS$, where $S = [s_1, s_2, \ldots, s_N]^T = [S_1^T, \ldots, S_C^T]^T \in R^N$ is the vector of all the representation coefficients corresponding to all the training samples of X and S_i is the subvector of the representation coefficients from class i.

2.1. CRC.

CRC is a typical linear representation-based classifier proposed recently [2]. In the CRC, a given testing sample y is collaboratively represented by all the training samples for classification. The CRC model is defined as

$$\min_S \left\{ \|y - XS\|_2^2 + \lambda \|S\|_2^2 \right\}, \quad (1)$$

where λ is a positive regularization parameter. Clearly, CRC can learn the closed-form solution of S as

$$S = H_1 y, \tag{2}$$

where $H_1 = (X^T X + \lambda I)^{-1} X^T$ with an identity matrix $\mathbf{I}$. Using the learned $S = (X^T X + \lambda I)^{-1} X^T y$, the class-specific representation residuals are determined as $\|y - X_i S_i\|_2 / \|S_i\|_2$. Finally, the given testing sample y is classified into the class with the minimum representation residual among all the classes.

2.2. DSRC. DSRC [44] is a discriminative sparse representation method with a l_2-norm regularization of the pairs of any two class-specific representations. It can achieve the good pattern discrimination among the different classes with sparsity. The DSRC model is defined as

$$\min_S \left\{ \|y - XS\|_2^2 + \gamma \sum_{i=1}^{C} \sum_{j=1}^{C} \left\| X_i S_i + X_j S_j \right\|_2^2 \right\}, \tag{3}$$

where γ is a positive regularization parameter. Through some algebra operations, the efficient solution of S can be obtained as

$$S = H_2 y, \tag{4}$$

where $H_2 = ((1 + 2\gamma) X^T X + 2\gamma LM)^{-1} X^T$ and $M = \begin{bmatrix} X_1^T X_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & X_C^T X_C \end{bmatrix}$. Using the learned $S = ((1 + 2\gamma) X^T X + 2\gamma LM)^{-1} X^T y$, we compute the class-specific representation residuals with $X_i S_i - y_2^2$ and classify y into the class with the minimum representation residual among all the classes.

2.3. Co-CRC. Co-CRC [47] is a new extension of CRC that can induce each training class to discriminatively and competitively represent each testing sample. The Co-CRC model is defined as

$$\min_S \left\{ \|y - XS\|_2^2 + \beta \sum_{i=1}^{C} \left\| X_{-i} S_{-i} \right\|_2^2 \right\}, \tag{5}$$

where β is a positive regularization parameter. The second term in equation (5) is the competitive representation constraint. According to the way of solving S [47], the learned solution of S is achieved as

$$S = H_3 y, \tag{6}$$

where $H_3 = (X^T X + \beta \sum_{i=1}^{C} P_{-i}^T X_{-i}^T X_{-i} P_{-i})^{-1} X^T$ and $P_{-i}^T = \begin{bmatrix} I_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & I_n \end{bmatrix}$. Using $S = (X^T X + \beta \sum_{i=1}^{C} P_{-i}^T X_{-i}^T X_{-i} P_{-i})^{-1} X^T y$, the class-specific representation residuals are calculated as $\|X_i S_i - y\|_2^2$ and the testing sample y is classified into the class with the minimum representation residual among all the classes.

3. The Proposed DCCRC

In this section, we detailedly present the proposed DCCRC method. The basic idea of DCCRC is first given, and then the DCCRC model and its solving procedure are described, finally the essential properties that DCCRC holds are analyzed.

3.1. Idea of DCCRC. The proposed DCCRC contains two assumptions that are originally inspired by the competitive and collaborative representation [47]. For clear descriptions, the collaborative representation of the given testing sample y using all training samples is rewritten as $y \approx XS = X_1 S_1 + X_2 S_2 + \cdots + X_C S_C = X_i S_i + X_{-i} S_{-i}$, where X_{-i} represents the training samples excluding samples from class i with the corresponding vector S_{-i} of the representation coefficients. The one assumption originates from the expectation that the true class of given testing sample y can dominantly represent y and the other classes have little contribution to representing it (i.e., $\|X_{-i} S_{-i}\|_2^2 = 0$). Unfortunately, the true class of the testing sample y is not known and any one of all the training classes could be chosen as the true class of y. In fact, we only make the training samples from one class to competitively represent the testing sample y as soon as possible and the contribution to representing y from other classes is as poor as possible in ideal case. Accordingly, with this good expectation, the testing sample y is well represented as $X_i S_i$ from class i by simultaneously minimizing the representation $X_{-i} S_{-i}$ from other classes. Thus, in the proposed method we introduce the competitive constraint $\|X_{-i} S_{-i}\|_2^2$ that was first designed in [47].

In collaborative representation, all the training samples approximately represent the testing sample y as soon as possible, i.e., $y \approx XS$. During the process of representation, if y belongs to class i with dominant representation $X_i S_i$, the representation $X_{-i} S_{-i}$ from the other classes tends to be very small. In specific, $X_i S_i$ tends to be equivalent to the representation XS in some degree. In such an ideal case, the approximate equalities can be learned, $y \approx X_i S_i \approx XS$. Borrowing the idea of degrading the correlations among classes by minimizing the discriminative constraint $\|X_i S_i + X_j S_j\|_2^2$ [44], we also assume that the correlation between the representation $X_i S_i$ from class i and the representation $X_{-i} S_{-i}$ from the other classes is as small as possible. That is to say, if class i can dominantly represent y with $X_i S_i$ and all the training samples can well represent y with XS, the correlation between XS and $X_{-i} S_{-i}$ should be small. Similar to the definition of $\|X_i S_i + X_j S_j\|_2^2$ [44], we design the another new discriminative constraint $\|XS + X_{-i} S_{-i}\|_2^2$. It is obvious that to minimize $\|XS + X_{-i} S_{-i}\|_2^2$ can minimize $\|XS\|_2^2$, $\|X_{-i} S_{-i}\|_2^2$, and $(XS)^T X_{-i} S_{-i}$. Minimizing $\|X_{-i} S_{-i}\|_2^2$ satisfies the first assumption. If $y \approx X_i S_i \approx XS$, minimizing $(XS)^T X_{-i} S_{-i}$ approximately equals to minimize $(X_i S_i)^T X_{-i} S_{-i}$ that can well degrade the correlation between the representation from one class and the representation from the other classes.

3.2. Model of DCCRC. In this section, we first introduce the objective function of the proposed DCCRC model and then present the procedures of solving it in details. The given testing sample y is represented by collaborative representation of all the training samples, and the DCCRC model on the basis of its idea is defined as follows:

$$\min_{S} \left\{ \|y - XS\|_2^2 + \lambda_1 \sum_{i=1}^{C} \left\|X_{-i}S_{-i}\right\|_2^2 + \lambda_2 \cdot \sum_{i=1}^{C} \left\|XS + X_{-i}S_{-i}\right\|_2^2 + \lambda_3 \|S\|_2^2 \right\}, \tag{7}$$

where λ_1, λ_2, and λ_3 is the positive regularization parameters. In equation (7), the second term $\sum_{i=1}^{C}\|X_{-i}S_{-i}\|_2^2$, first designed in [47], is the competitive constraint that can make each class competitively and discriminatively represent the testing sample y among all the classes. The third term $\sum_{i=1}^{C}\|XS + X_{-i}S_{-i}\|_2^2$ is the discriminative constraint that not only makes each class competitively represent the testing sample y but also degrades the representation correlations between one class and the other classes for more discrimination. Note that when $\lambda_1 = \lambda_2 = 0$, DCCRC is the same as CRC, and when $\lambda_2 = \lambda_3 = 0$, DCCRC is the same as Co-CRC.

In order to achieve the solution of the representation coefficient vector S, equation (7) should be further reformulated as

$$\min_{S} \left\{ \|y - XS\|_2^2 + \lambda_1 \sum_{i=1}^{C} \left\|\bar{X}_{-i}S\right\|_2^2 + \lambda_2 \cdot \sum_{i=1}^{C} \left\|XS + \bar{X}_{-i}S\right\|_2^2 + \lambda_3 \|S\|_2^2 \right\}, \tag{8}$$

where $\widehat{X}_i = [0, 0 \ldots 0, X_i, 0 \ldots 0]$ and $\bar{X}_{-i} = X - \widehat{X}_i = [X_1, X_2, \ldots, X_{i-1}, 0, X_{i+1}, \ldots, X_C]$. To simply solve S, let $F_1(S) = \|y - XS\|_2^2 + \lambda_1 \sum_{i=1}^{C}\|\bar{X}_{-i}S\|_2^2 + \lambda_3\|S\|_2^2$ and $F_2(S) = \lambda_2 \sum_{i=1}^{C} \|XS + \bar{X}_{-i}S\|_2^2$. Firstly, the derivative of $F_1(S)$ with respect to **S** is calculated as

$$\begin{aligned} \frac{\partial F_1(S)}{\partial S} &= \frac{\left(\|y - XS\|_2^2 + \lambda_1 \sum_{i=1}^{C}\left\|\bar{X}_{-i}S\right\|_2^2 + \lambda_3\|S\|_2^2\right)}{\partial S} \\ &= -2X^T(y - XS) + 2\lambda_1 \sum_{i=1}^{C}\left(\bar{X}_{-i}^T\bar{X}_{-i}S\right) + 2\lambda_3 S. \end{aligned} \tag{9}$$

Since $\sum_{i=1}^{C}(\bar{X}_{-i}^T\bar{X}_{-i}S)$ can be rewritten as

$$\begin{aligned} \sum_{i=1}^{C}\left(\bar{X}_{-i}^T\bar{X}_{-i}S\right) &= \sum_{i=1}^{C}\left(X - \widehat{X}_i\right)^T\left(X - \widehat{X}_i\right)S \\ &= \sum_{i=1}^{C}\left(X^TX - 2X^T\widehat{X}_i + \widehat{X}_i^T\widehat{X}_i\right)S \\ &= \left(CX^TX - \sum_{i=1}^{C}\widehat{X}_i^T\widehat{X}_i\right)S \\ &= \left(CX^TX - G\right)S, \end{aligned} \tag{10}$$

where $X^T\widehat{X}_i = \widehat{X}_i^T\widehat{X}_i$, and G is defined as

$$G = \begin{bmatrix} X_1^TX_1 & \ldots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ldots & X_C^TX_C \end{bmatrix}. \tag{11}$$

Using equations (9) and (10), $\partial F_1(S)/\partial S$ is reformulated as

$$\frac{\partial F_1(S)}{\partial S} = -2X^T(y - XS) + 2\lambda_1\left(CX^TX - G\right)S + 2\lambda_3 S. \tag{12}$$

Then, the derivative of $F_2(S)$ with respect to S is calculated as

$$\begin{aligned} \frac{\partial F_2(S)}{\partial S} &= \frac{\partial\left(\lambda_2\sum_{i=1}^{C}\left\|XS + \bar{X}_{-i}S\right\|_2^2\right)}{\partial S} \\ &= 2\lambda_2 \sum_{i=1}^{C}\left(\left(X + \bar{X}_{-i}\right)^T\left(X + \bar{X}_{-i}\right)\right)S \\ &= 2\lambda_2\left(\sum_{i=1}^{C}X^TX + 2\sum_{i=1}^{C}X^T\bar{X}_{-i} + \sum_{i=1}^{C}\left(\bar{X}_{-i}^T\bar{X}_{-i}\right)\right)S. \end{aligned} \tag{13}$$

In equation (13), using $\bar{X}_{-i} = X - \widehat{X}_i$, $\sum_{i=1}^{C}X^T\bar{X}_{-i}$ can be reformulated as

$$\begin{aligned} \sum_{i=1}^{C}X^T\bar{X}_{-i} &= X^T\left(CX - \left(\widehat{X}_1 + \widehat{X}_2 + \cdots + \widehat{X}_C\right)\right) \\ &= (C - 1)X^TX. \end{aligned} \tag{14}$$

Using equations (10) and (14), equation (13) can be finally rewritten as

$$\begin{aligned} \frac{\partial F_2(S)}{\partial S} &= \frac{\partial\left(\lambda_2\sum_{i=1}^{C}\left\|XS + \bar{X}_{-i}S\right\|_2^2\right)}{\partial S} \\ &= 2\lambda_2\left(2(2C - 2)X^TX - G\right)S. \end{aligned} \tag{15}$$

Clearly, the objective function of DCCRC is $F_1(S) + F_2(S)$. Using equations (10) and (15), the derivative of the proposed function with respect to S is

$$\begin{aligned} \frac{\partial(F_1(S) + F_2(S))}{\partial S} &= -2X^T(y - XS) + 2\lambda_1\left(CX^TX - G\right)S \\ &\quad + 2\lambda_2\left(2(2C - 2)X^TX - G\right)S \\ &\quad + 2\lambda_3 S. \end{aligned} \tag{16}$$

Finally, we set $\partial(F_1(S) + F_2(S))/\partial S = 0$, and the solution of the representation coefficient vector S in equation (7) is obtained as

$$S = \left((1 - (\lambda_1 + 4\lambda_2)C + 4\lambda_2)X^TX + (\lambda_1 + \lambda_2)G + \lambda_3 I\right)^{-1}X^Ty. \tag{17}$$

After obtaining the representation coefficient vector S, we calculate the class-specific representation residuals and determine the class label c_y of the testing sample y as

$$c_y = \underset{c_i}{\operatorname{argmin}} \left\| y - X_i S_i \right\|_2, \quad i = 1, 2, \ldots, C. \tag{18}$$

That is to say, the given testing sample y is classified into the class with the minimal representation residuals among all the classes. According to the descriptions of the proposed DCCRC model above, the proposed DCCRC is briefly summarized in Algorithm 1.

3.3. Analysis of DCCRC. In this section, we first further analyze the terms $\sum_{i=1}^{C} \|X_{-i}S_{-i}\|_2^2$ and $\sum_{i=1}^{C} \|XS + X_{-i}S_{-i}\|_2^2$ in the proposed DCCRC method, in order to explain the more power of pattern discrimination. And then, the analyses of differences between the proposed DCCRC and Co-CRC, DSRC are emphasized.

Using the way of analyzing the competitive representation [47], $\|X_{-i}S_{-i}\|_2^2$ in term $\sum_{i=1}^{C} \|X_{-i}S_{-i}\|_2^2$ can be rewritten as $\|XS - X_iS_i\|_2^2$, and we can obtain the equality as

$$\begin{aligned} \left\| XS - X_iS_i \right\|_2^2 = \|XS\|_2^2 - 2(XS)^T (X_iS_i) + \left\| X_iS_i \right\|_2^2 \geq \|XS\|_2^2 \\ + \left\| X_iS_i \right\|_2^2 - 2\|XS\|_2 \left\| X_iS_i \right\|_2. \end{aligned} \tag{19}$$

Assume the angle between XS and X_iS_i is α. Using equation (19), $\cos(\alpha)$ can be obtained as

$$\cos(\alpha) = \frac{(XS)^T (X_iS_i)}{\|XS\|_2 \left\| X_iS_i \right\|_2} \leq 1. \tag{20}$$

According to equation (20), when $\cos(\alpha) = 1$, XS approaches X_iS_i with the same direction. In this ideal case, the given testing sample y is dominantly represented by X_iS_i from class i that y truly belongs to. Thus, to minimize $\sum_{i=1}^{C} \|X_{-i}S_{-i}\|_2^2$ could have two advantages. One is that each class competitively represents the testing sample y. Another one is that the true class of y could competitively represent it and the other classes poorly represent it.

Moreover, $\|XS + X_{-i}S_{-i}\|_2^2$ in term $\sum_{i=1}^{C} \|XS + X_{-i}S_{-i}\|_2^2$ can be reformulated as $\|XS - 2X_iS_i\|_2^2$. Through simple algebra of $\|XS - 2X_iS_i\|_2^2$, we can also achieve equation (20). This fact means that to minimize $\sum_{i=1}^{C} \|XS + X_{-i}S_{-i}\|_2^2$ has the very similar superiorities of minimizing $\sum_{i=1}^{C} \|X_{-i}S_{-i}\|_2^2$. And also, we can rewrite $\|XS + X_{-i}S_{-i}\|_2^2$ as $\|X_iS_i + 2X_{-i}S_{-i}\|_2^2 = \|X_iS_i\|_2^2 + 4\|X_{-i}S_{-i}\|_2^2 + 4(X_iS_i)^T (X_{-i}S_{-i})$. To minimize $\|XS + X_{-i}S_{-i}\|_2^2$ is to simultaneously minimize $\|X_iS_i\|_2^2$, $\|X_{-i}S_{-i}\|_2^2$, and $(X_iS_i)^T (X_{-i}S_{-i})$. We can see that except minimizing $\|X_{-i}S_{-i}\|_2^2$, minimizing $(X_iS_i)^T (X_{-i}S_{-i})$ can degrade the correlation between X_iS_i and $X_{-i}S_{-i}$ [44]. That is to say, to minimize $\sum_{i=1}^{C} \|XS + X_{-i}S_{-i}\|_2^2$ could degrade the correlation between one class and the other classes, in order to enhance the power of pattern discrimination and competitive representations among all the classes. Thus, the terms $\sum_{i=1}^{C} \|X_{-i}S_{-i}\|_2^2$ and $\sum_{i=1}^{C} \|XS + X_{-i}S_{-i}\|_2^2$ can obtain the competitive and discriminative collaborative representation for favorable classification. Besides, the pattern discrimination among all the classes can be intuitively verified in the next section.

The differences between the proposed DCCRC and Co-CRC, DSRC can be analyzed by comparing their corresponding models (i.e., equation (7) for DCCRC, equation (5) for Co-CRC, and equation (3) for DSRC). According to equations (3) and (7), DCCRC is very different from DSRC, but both have similar discriminative terms. The term $\|X_iS_i + X_jS_j\|_2^2$ in DSRC can degrade the correlations between any two classes for favorable pattern discrimination, but the term $\|XS + X_{-i}S_{-i}\|_2^2$ in DCCRC can degrade the correlations between any one class and the other classes for competitively enhancing the discriminative representation from each class for classification. Besides, compared to DSRC, the proposed DCCRC also has the competitive constraint and the regularization of the representation coefficients. Furthermore, the proposed DCCRC is the extension of Co-CRC because DCCRC and Co-CRC have the same competitive constraint $\sum_{i=1}^{C} \|X_{-i}S_{-i}\|_2^2$. In contrast with Co-CRC, the proposed DCCRC also has the designed discriminative constraint $\|XS + X_{-i}S_{-i}\|_2^2$ and the regularization of the representation coefficients, in order that DCCRC further enhances the competitive representations among all the classes. Thus, the proposed DCCRC has more pattern discrimination than DSRC and Co-CRC that can be experimentally verified in the next experimental section.

4. Experiments

In this section, the extensive experiments on several face databases and some real numerical UCI data sets are conducted. In the experiments, we compare the proposed DCCRC with the state-of-the-art RBC methods including SRC [1], CRC [2], CCRC [46], Co-CRC [47], DSRC [44], ProCRC [22], and EProCRC [43]. It should be noted that all regularized parameters in the competing methods are preset as the range $[10^{-3}, 10^{-2}, \ldots, 1, 10, 100]$ for fair comparisons in the experiments. The optimal classification results of each competing method are obtained among the range of its parameters.

4.1. Data Sets. In this section, we briefly describe the used data sets including the AR, YaleB, IMM, Yale, and PIE29 face databases and the real UCI data sets. The YaleB database (http://vision.ucsd.edu/leekc/ExtYaleDatabase/ExtYaleB.html) was taken under different poses and uncontrolled illumination conditions. The Yale database (http://cvc.yale.edu/projects/yalefaces/yalefaces.html) was taken by different facial expressions. The AR database (http://www2.ece.ohio-state.edu/aleix/ARdatabase.html) was taken by various facial expressions and illumination conditions, and we use a subset of AR with 1400 image from 100 subjects. The IMM database (http://www.imm.dtu.dk/~aam/datasets/datasets.html) contains 240 annotated monocular images from 40 subjects. The PIE29 database (http://www.intbox.com/public/project/4742/) was taken in different conditions including 13 postures, 43 lights, and 4

Input: The given training set $X = [X_1, X_2, \ldots, X_C]$ and testing sample y with three regularized parameters λ_1, λ_2, and λ_3.
Output: Determine the class label of y.
(1) Normalize each sample of X and y.
(2) Calculate $X^T X$ and G with equation (11).
(3) Solve the representation vector with equation (17).
(4) Calculate the class-specific representation residuals with $\|y - X_i S\|_{i2}$.
(5) Predict the label of y with equation (18).

Algorithm 1: The proposed DCCRC method.

expressions. In the experiments, each image is cropped and resized into 32×32 with 256 gray levels per pixel and also the gray level values are normalized to [0, 1]. The numbers of total samples, classes, samples per class, and chosen training samples per class are shown in Table 1. As an example, the image samples of one subject from each face data base are shown in Figure 1.

The real used eight UCI data sets were downloaded from UC Irvine Machine Learning Repository (UCI) (http://archive.ics.uci.edu/ml). They are "Wine," "Vehicle," "Auto MPG," "Statlog (Heart)," "Statlog (Australian Credit Approval)," "Credit Approval," "Isolet," and "Ionosphere." Note that "Auto MPG," "Statlog (Heart)," "Statlog (Australian Credit Approval)," "Credit Approval," and "Ionosphere" are abbreviated as "Auto," "Heart," "SCredit," "Credit," and "Iono," respectively. The numbers of total samples, classes, attributes, and training samples per class are displayed in Table 2. In the experiments, each sample on these UCI data sets is also normalized to [0, 1]. Furthermore, on these face and UCI data sets, they are randomly divided into the sets of the training and testing samples ten times, and the training samples chosen from each class are shown in Tables 1 and 2.

4.2. Experiment 1. In this section, we first conduct the experiments to analyze the competitive term $\lambda_1\|X_{-i}S_{-i}\|_2^2$ and the discriminative term $\lambda_2\sum_{i=1}^{C}\|XS + X_{-i}S_{-i}\|_2^2$ by varying the values of the parameters λ_1 and λ_2 in the proposed DCCRC on the five face databases. The values of the parameters λ_1, λ_2, and λ_3 are preset as $[10^{-3}, 10^{-2}, \ldots, 1, 10, 100]$, and the numbers of training samples per class are chosen as $l = 4$ on AR, $l = 4$ on IMM, $l = 18$ on YaleB, $l = 2$ on Yale, and $l = 6$ on PIE29. For visual comparisons, the model $\min_S \{\|y - XS\|_2^2 + \lambda_1\sum_{i=1}^{C}\|X_{-i}S_{-i}\|_2^2 + \lambda_3 S_2^2\}$ without $\lambda_2\sum_{i=1}^{C}\|XS + X_{-i}S_{-i}\|_2^2$ is denoted as DCCRC$_1$, and the model $\min_S \{\|y - XS\|_2^2 + \lambda_2\sum_{i=1}^{C}\|XS + X_{-i}S_{-i}\|_2^2 + \lambda_3\|S\|_2^2\}$ without $\lambda_1\|X_{-i}S_{-i}\|_2^2$ is denoted as DCCRC$_2$. Accordingly, we compare DCCRC$_1$ with DCCRC to demonstrate the discrimination of the term $\lambda_2\sum_{i=1}^{C}\|XS + X_{-i}S_{-i}\|_2^2$ by varying the values of the parameter λ_1. And we compare DCCRC$_2$ with DCCRC to demonstrate the discrimination of the term $\lambda_1 X_{-i}S_{-i2}^2$ by varying the values of the parameter λ_2. It should be noted that the values of the parameters λ_2 and λ_3 are optimal with best classification accuracies when DCCRC$_1$ is compared with DCCRC, and the values of the parameters λ_1 and λ_3 are optimal with best classification accuracies when DCCRC$_2$ is compared with DCCRC. For conveniently presenting the values of λ_1 and λ_2 in the figures, we use $p_1 = \lg(\lambda_1)$ and $p_2 = \lg(\lambda_2)$ (i.e., the values of p_1 and p_2 correspond to that of λ_1 and λ_2, respectively).

The classification accuracies of DCCRC$_1$ and DCCRC with varying λ_1 are shown in Figure 2, and the ones of DCCRC$_2$ and DCCRC with varying λ_2 are shown in Figure 3. From the experimental results in Figure 2, we can see that DCCRC with $\lambda_2\sum_{i=1}^{C}\|XS + X_{-i}S_{-i}\|_2^2$ significantly performs better than DCCRC$_1$ without $\lambda_2\sum_{i=1}^{C}\|XS + X_{-i}S_{-i}\|_2^2$, and DCCRC is more robust to the variations of λ_1 than DCCRC$_1$. As shown in Figure 3, we can also observe that DCCRC with $\lambda_1\|X_{-i}S_{-i}\|_2^2$ significantly performs better than DCCRC$_2$ without $\lambda_1\|X_{-i}S_{-i}\|_2^2$ and DCCRC is more robust to the variations of λ_2 than DCCRC$_2$. In addition, the classification performance of DCCRC$_1$ with variations of λ_1 and DCCRC$_2$ with variations of λ_2 shows that the terms $\lambda_1\|X_{-i}S_{-i}\|_2^2$ and $\lambda_2\sum_{i=1}^{C}\|XS + X_{-i}S_{-i}\|_2^2$ can improve the power of the pattern discrimination. The experimental results in two figures imply that the proposed DCCRC has effective and robust classification performance. As a consequence, the more pattern discrimination of the proposed DCCRC originated from the competitive and discriminative terms is well verified.

And then, we visually verify the discriminative ability of the proposed DCCRC method in comparison with the competitive CRC method (i.e., Co-CRC). As discussed in Section 3.3, we define the class-specific representation contribution for the given testing sample y as

$$\mathrm{con}_y^i = \frac{\|X_i S_i\|_2^2}{\sum_{i=1}^{C}\|X_i S_i\|_2^2}. \tag{21}$$

Clearly, both DCCRC and Co-CRC classify each testing sample into the class with the largest con_y^i among all the classes. Then, the pattern discrimination ability of both is intuitively represented by the representation reconstructive images for the given testing samples from class 26 in IMM and class 9 in AR. The first five representation reconstructive images of the testing samples corresponding to the top five largest representation contributions are illustrated in Figure 4. Note that the numbers in the bracket under each reconstructive image are the class and its representation contribution con_y^i. For example, (26, 10.74) under the reconstructive image means the class 26 has the representation contribution $\mathrm{con}_y^{26} = 10.74$. As can be seen

Table 1: The main information about the used face databases.

Data	Total samples	Classes	Samples per class	Training samples per class
YaleB	2432	38	64	12, 18, 24, 30
Yale	165	15	11	2, 3, 4, 5
AR	1400	100	14	2, 4, 6, 8
IMM	240	40	6	2, 3, 4, 5
PIE29	1632	68	24	2, 4, 6, 8

(a) (b) (c) (d) (e)

Figure 1: The example of the facial images of one individual from five face database. (a) YaleB. (b) PIE29. (c) IMM. (d) Yale. (e) AR.

Table 2: The main information about the used UCI data set.

Data	Total samples	Classes	Attributes	Training samples per class
Wine	178	3	13	7, 8
Vehicle	846	4	18	10, 11
Auto	392	3	8	10, 14
Credit	690	2	15	8, 10
Heart	270	2	13	6, 8
SCredit	690	2	14	35, 50
Isolet	1560	2	617	21, 26
Iono	351	2	34	15, 20

in Figure 4, the proposed DCCRC correctly represents and classifies the testing samples, but Co-CRC wrongly represents and classifies them. Moreover, we can observe that the first reconstructive image that is reconstructed by the class with the largest representation contribution in Co-CRC is very similar to the testing image on each face database. Through the experimental illustrations in Figure 4, the proposed DCCRC is more discriminative than Co-CRC for classification. This means the designed term $\lambda_2 \sum_{i=1}^{C} \|XS + X_{-i}S_{-i}\|_2^2$ is discriminative. Therefore, it can be concluded that the proposed DCCRC has the effective and robust classification due to the competitive and discriminative constraints.

4.3. Experiment 2. In this section, we compare the proposed DCCRC to the competing methods on the face databases and the UCI data sets. The experimental results of each competing method are the averages of the classification accuracies on ten division of each data set. The best classification accuracies of each method are achieved among the range of its parameter, and the preset class-specific training samples on each data set are shown in Tables 1 and 2.

The classification accuracies of all the competing methods are shown in Table 3 on face databases and Table 4 on the UCI data sets. Note that the best classification performance among all the methods on each data set is indicated in bold face. As shown in two tables, the classification accuracies of each competing method almost ascend with the increase of the class-specific training samples on all the data sets. On the face databases, we can see that the proposed DCCRC nearly achieves the best classification accuracies among all the competing methods, but it could not significantly improve very much in comparison with some methods. As displayed in Table 4, the proposed

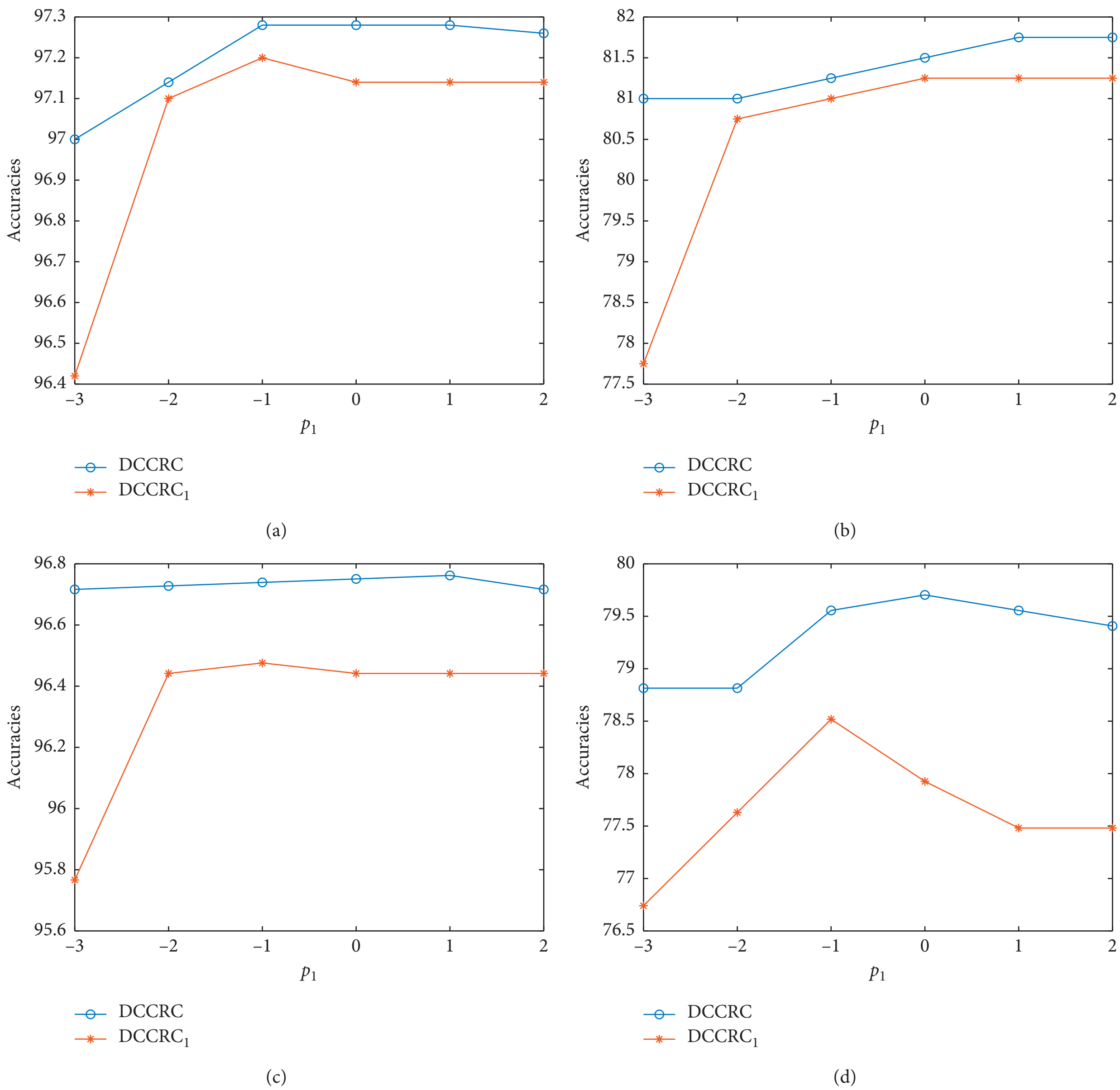
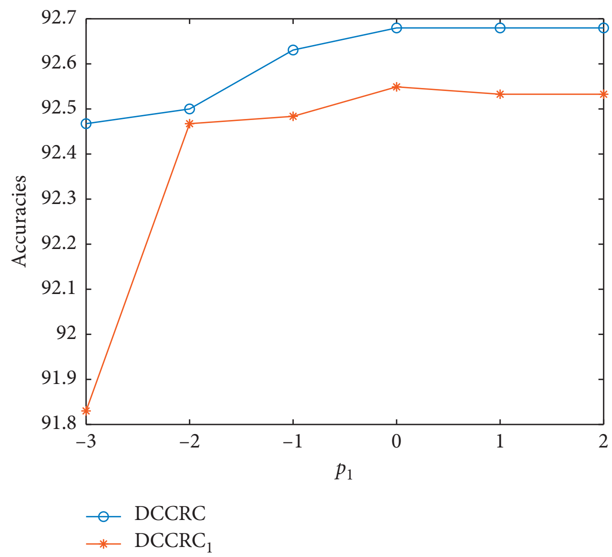
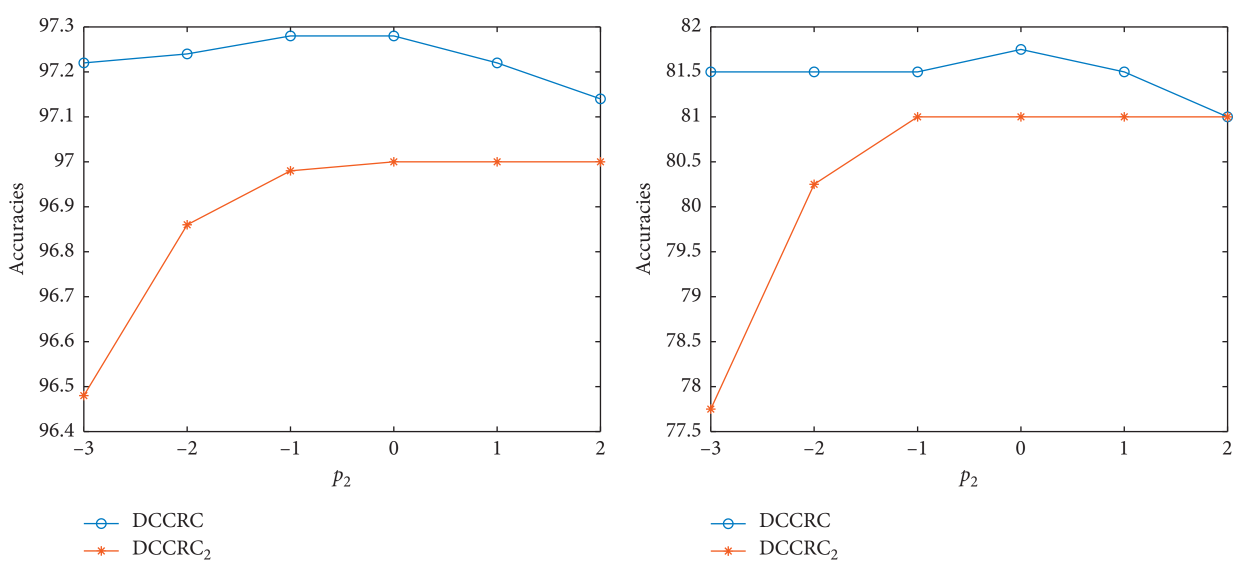


Figure 2: Continued.



(e)

Figure 2: The comparisons of DCCRC and DCCRC$_1$ with varying values of parameter λ_1 on each face database. (a) AR ($l = 4$). (b) IMM ($l = 4$). (c) YaleB ($l = 18$). (d) Yale ($l = 2$). (e) PIE29 ($l = 6$).



(a) (b)

Figure 3: Continued.

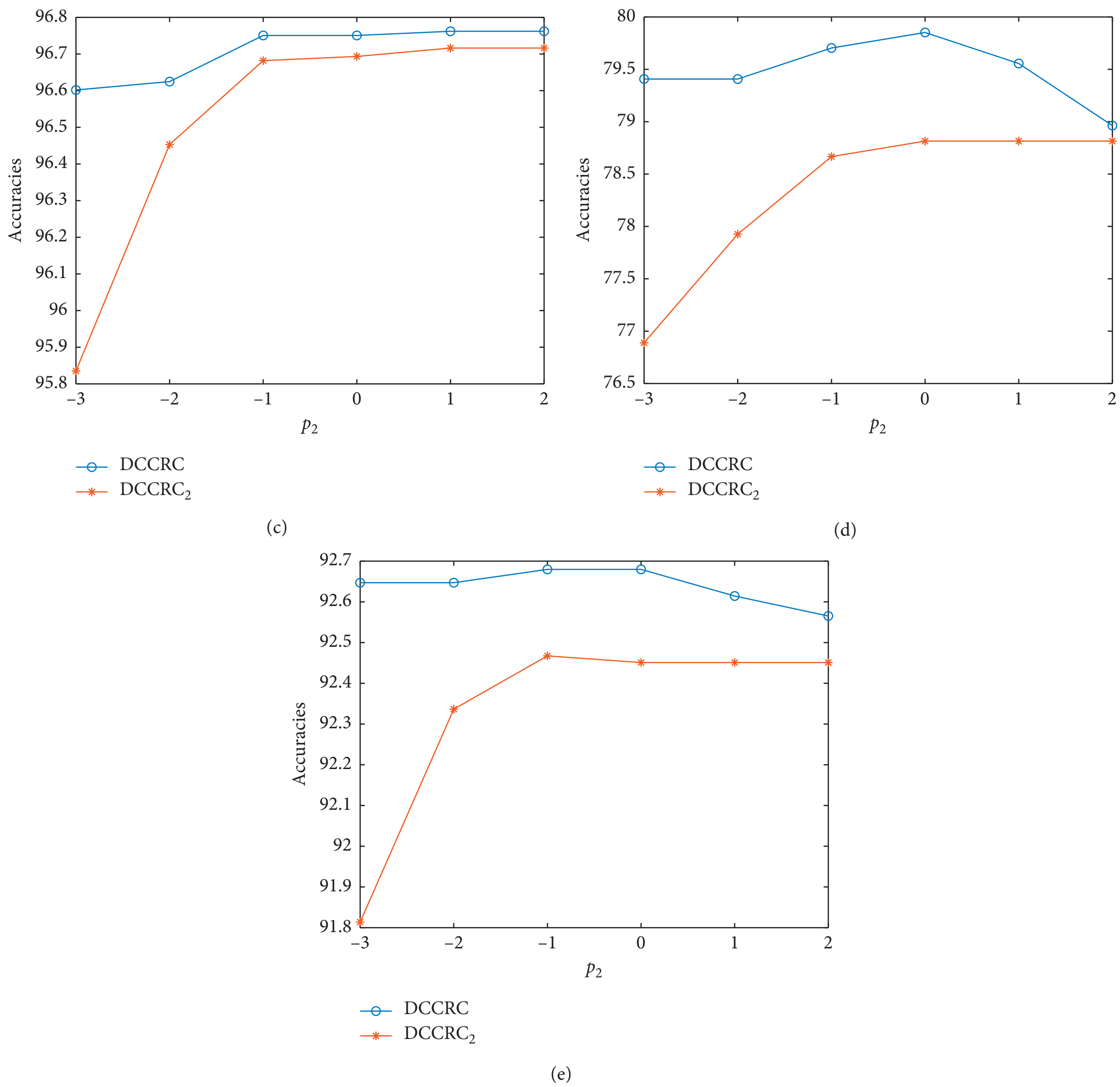


FIGURE 3: The comparisons of DCCRC and $DCCRC_2$ with varying values of parameter λ_2 on each face database. (a) AR ($l = 4$). (b) IMM ($l = 4$). (c) YaleB ($l = 18$). (d) Yale ($l = 2$). (e) PIE29 ($l = 6$).

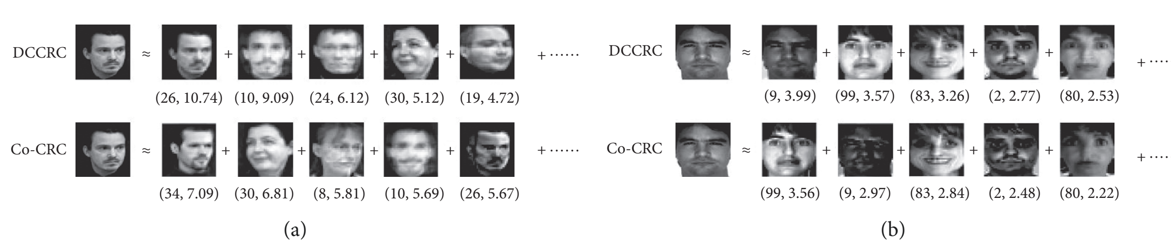


FIGURE 4: The reconstructive images of the given testing samples by the classes with the top five largest representation contributions via DCCRC and Co-CRC on the face databases (a) IMM and (b) AR.

DCCRC significantly performs better than the other competing methods. In addition, from the classification results in two tables, we can observe that CCRC, Co-CRC, DSRC, ProCRC, and EProCRC obtain the similar competitive classification performance. The possible reason is that these methods can fully employ the class-specific representations

Table 3: The classification accuracies (%) of the competing methods with the corresponding standard deviations with different class-specific training samples on each face database.

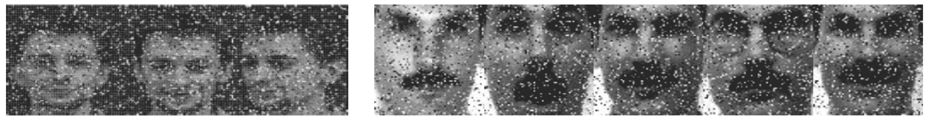
Data	l	SRC	CRC	CCRC	ProCRC	EProCRC	Co-CRC	DSRC	DCCRC
YaleB	12	91.770.95	93.58 ± 0.58	93.78 ± 0.63	94.34 ± 0.51	93.72 ± 0.67	94.53 ± 0.40	94.05 ± 0.53	**94.61 ± 0.38**
	18	95.48 ± 0.45	96.21 ± 0.41	96.55 ± 0.46	96.63 ± 0.51	96.51 ± 0.49	96.52 ± 0.46	96.09 ± 0.54	**96.91 ± 0.38**
	24	96.81 ± 0.43	97.31 ± 0.42	97.62 ± 0.23	97.69 ± 0.23	97.59 ± 0.22	97.54 ± 0.29	97.13 ± 0.30	**97.93 ± 0.24**
	30	97.57 ± 0.47	97.64 ± 0.38	98.12 ± 0.46	98.13 ± 0.48	98.12 ± 0.46	97.63 ± 0.43	97.46 ± 0.47	**98.25 ± 0.45**
AR	2	84.99 ± 0.90	86.71 ± 0.77	89.92 ± 0.70	90.15 ± 0.77	89.89 ± 0.68	90.10 ± 0.79	85.46 ± 0.80	**90.18 ± 0.77**
	4	94.18 ± 0.49	94.48 ± 1.16	96.55 ± 0.78	96.60 ± 0.73	96.50 ± 0.80	96.48 ± 0.50	94.93 ± 0.63	**96.72 ± 0.78**
	6	96.67 ± 0.94	96.25 ± 0.75	98.04 ± 0.87	98.02 ± 0.89	97.85 ± 0.90	97.78 ± 0.58	97.35 ± 0.46	**98.08 ± 0.80**
	8	98.00 ± 0.87	98.00 ± 0.88	99.08 ± 0.55	**99.17 ± 0.51**	99.06 ± 0.64	98.38 ± 0.34	98.22 ± 0.62	**99.17 ± 0.43**
Yale	2	73.21 ± 4.23	71.60 ± 4.68	77.53 ± 4.36	77.53 ± 4.36	77.04 ± 4.39	77.65 ± 4.53	76.54 ± 3.59	**78.52 ± 4.16**
	3	82.92 ± 1.56	80.69 ± 1.11	85.97 ± 1.44	86.39 ± 1.95	85.97 ± 1.44	85.56 ± 2.51	85.14 ± 2.91	**86.81 ± 1.93**
	4	85.87 ± 2.12	81.75 ± 2.98	88.41 ± 2.37	88.89 ± 2.23	88.57 ± 2.17	87.78 ± 2.65	87.94 ± 2.93	**89.37 ± 2.37**
	5	90.74 ± 3.95	88.33 ± 4.65	92.04 ± 2.57	92.41 ± 2.37	92.22 ± 2.43	88.89 ± 1.57	88.89 ± 6.29	**92.96 ± 2.78**
IMM	2	63.54 ± 2.76	60.00 ± 1.58	66.67 ± 2.67	66.77 ± 2.54	66.67 ± 2.52	66.15 ± 2.66	65.83 ± 2.19	**66.88 ± 2.40**
	3	69.72 ± 2.34	63.89 ± 2.15	73.06 ± 2.34	**73.61 ± 3.48**	72.50 ± 2.47	73.33 ± 3.21	72.08 ± 2.09	**73.61 ± 3.15**
	4	74.38 ± 4.24	68.96 ± 5.09	77.08 ± 3.42	78.33 ± 3.59	77.08 ± 3.68	77.08 ± 2.92	75.83 ± 3.23	**78.54 ± 3.57**
	5	82.92 ± 4.01	81.67 ± 4.08	85.00 ± 2.24	85.42 ± 3.68	84.58 ± 2.92	83.33 ± 3.42	83.33 ± 3.42	**86.25 ± 3.45**
PIE29	4	89.90 ± 0.70	90.94 ± 0.57	91.01 ± 0.71	91.37 ± 0.55	90.97 ± 0.74	91.56 ± 0.48	90.50 ± 0.60	**91.68 ± 0.58**
	6	92.12 ± 0.77	92.29 ± 0.40	92.68 ± 0.64	92.81 ± 0.65	92.60 ± 0.70	92.76 ± 0.67	92.45 ± 0.41	**92.91 ± 0.70**
	8	92.94 ± 0.48	92.74 ± 0.75	93.49 ± 0.54	93.49 ± 0.60	93.51 ± 0.54	93.22 ± 0.66	93.14 ± 0.56	**93.77 ± 0.60**
	10	93.68 ± 0.61	94.03 ± 0.78	94.47 ± 0.53	94.54 ± 0.54	94.39 ± 0.47	93.99 ± 0.85	93.89 ± 0.92	**94.60 ± 0.54**

Table 4: The classification accuracies (%) of the competing methods with the corresponding standard deviations with different class-specific training samples on each UCI data set.

Data	l	SRC	CRC	CCRC	ProCRC	EProCRC	Co-CRC	DSRC	DCCRC
Vehicle	10	56.382.00	62.11 ± 3.27	60.02 ± 3.60	60.07 ± 3.62	60.02 ± 3.62	53.55 ± 4.14	57.94 ± 2.21	**63.03 ± 3.44**
	11	57.03 ± 1.13	64.56 ± 1.77	63.22 ± 1.69	63.24 ± 1.69	63.22 ± 1.69	54.34 ± 7.00	61.22 ± 4.23	**65.01 ± 2.70**
Auto	10	70.54 ± 2.24	73.02 ± 2.09	73.47 ± 1.66	73.67 ± 1.93	73.68 ± 1.90	69.18 ± 5.38	70.02 ± 2.86	**73.89 ± 1.59**
	14	73.80 ± 4.06	75.28 ± 3.03	75.66 ± 3.14	75.63 ± 3.18	75.62 ± 3.17	71.16 ± 7.51	74.16 ± 4.13	**75.93 ± 3.60**
Credit	8	61.25 ± 5.80	67.98 ± 2.44	68.72 ± 1.04	68.55 ± 0.84	68.55 ± 0.82	59.91 ± 5.39	62.85 ± 6.82	**69.50 ± 3.05**
	10	64.90 ± 4.28	68.78 ± 1.12	69.25 ± 2.19	69.31 ± 2.02	69.22 ± 2.01	61.88 ± 3.92	63.82 ± 6.59	**72.42 ± 2.79**
Wine	7	66.88 ± 4.79	77.58 ± 6.10	87.13 ± 3.17	88.41 ± 2.64	72.10 ± 3.67	62.80 ± 3.62	76.56 ± 6.17	**89.81 ± 2.06**
	8	70.52 ± 5.39	82.8 ± 1.93	89.22 ± 4.17	90.00 ± 3.00	77.79 ± 3.96	66.88 ± 7.01	80.91 ± 6.17	**91.04 ± 2.57**
Heart	6	58.76 ± 5.69	61.94 ± 5.16	64.65 ± 2.89	64.42 ± 2.77	64.42 ± 2.77	61.24 ± 2.2	63.64 ± 8.63	**67.13 ± 4.04**
	8	63.54 ± 4.11	70.87 ± 3.32	71.42 ± 2.65	71.42 ± 2.65	71.42 ± 2.65	69.53 ± 5.49	66.14 ± 4.40	**73.23 ± 2.83**
SCredit	35	63.10 ± 2.77	71.97 ± 2.87	72.00 ± 2.77	72.03 ± 2.71	72.23 ± 2.77	57.58 ± 3.35	63.74 ± 4.81	**75.84 ± 3.81**
	50	66.07 ± 3.51	74.64 ± 2.83	74.41 ± 1.38	74.44 ± 1.34	74.47 ± 1.52	59.32 ± 2.60	65.53 ± 4.88	**76.95 ± 4.32**
Isolet	21	71.41 ± 3.02	71.38 ± 1.48	69.83 ± 0.94	69.47 ± 1.10	69.47 ± 1.10	66.36 ± 2.69	72.89 ± 2.39	**73.19 ± 2.50**
	26	78.10 ± 1.70	79.22 ± 2.02	77.80 ± 2.20	77.53 ± 2.29	77.55 ± 2.32	75.81 ± 3.01	79.50 ± 2.27	**79.88 ± 2.28**
Iono	15	83.49 ± 4.25	88.66 ± 3.88	87.54 ± 3.50	87.35 ± 3.39	87.35 ± 3.43	82.06 ± 3.84	81.99 ± 3.18	**89.72 ± 4.29**
	20	86.37 ± 1.94	91.38 ± 1.27	88.49 ± 2.05	88.04 ± 1.55	87.97 ± 1.45	82.06 ± 1.45	83.09 ± 1.91	**92.99 ± 1.12**

in the collaborative representation to improve the pattern discrimination among all the classes. As a consequence, we can conclude that our DCCRC method is a promising representation-based classifier in pattern classification with effectiveness and robustness.

4.4. Experiment 3. In this section, we conduct the experiments on IMM and Yale to compare the proposed DCCRC with the competing methods under the situations of the testing samples with the corruptions. In the experiments, the numbers l of the class-specific training samples are preset as $l = 3$ on IMM and $l = 6$ on Yale, and the remaining samples per class are regarded as the testing samples. And the classification results of each competing method are the averages of recognition accuracies on ten training and testing divisions of data. Moreover, the testing samples per class are randomly corrupted by randomly adding the pixels and the block occlusion with a panda. That is to say, the pixel corruptions are that some pixels of each testing image are randomly replaced by the uncertain gray scale values between 0 and 255, and some part of each testing image is randomly occluded by the panda. The ratios of the corrupted size to the original size of each testing image are from 0.1 to



(a) (b)

Figure 5: The testing images with the random pixel corruptions from one class. (a) IMM. (b) Yale.

(a) (b)

Figure 6: The testing images with the random block occlusions from one class. (a) IMM. (b) Yale.

Table 5: The classification accuracies (%) of the competing methods with the corresponding standard deviations under the situations of the random block occlusions.

Data	Ratios	SRC	CRC	CCRC	ProCRC	EProCRC	Co-CRC	DSRC	DCCRC
IMM	0.1	62.17 ± 1.80	63.00 ± 1.00	65.17 ± 1.33	65.50 ± 0.85	62.33 ± 0.97	63.67 ± 1.35	62.67 ± 0.97	**66.00 ± 0.82**
	0.2	54.33 ± 1.43	56.17 ± 1.80	59.50 ± 1.87	58.67 ± 1.63	55.50 ± 3.27	55.33 ± 1.55	57.50 ± 1.18	**59.67 ± 1.45**
	0.3	45.67 ± 2.44	47.67 ± 1.62	50.00 ± 2.36	50.83 ± 2.47	48.17 ± 2.44	49.50 ± 2.51	49.83 ± 1.22	**51.50 ± 2.26**
	0.4	37.33 ± 3.14	35.83 ± 1.90	40.50 ± 2.15	**41.00 ± 2.44**	38.33 ± 2.36	39.67 ± 1.72	39.33 ± 2.20	**41.00 ± 2.44**
Yale	0.1	82.67 ± 2.83	86.67 ± 2.11	86.50 ± 2.25	86.50 ± 1.94	85.50 ± 2.53	85.00 ± 3.18	84.17 ± 3.23	**87.33 ± 2.31**
	0.2	76.00 ± 2.83	79.17 ± 3.46	79.00 ± 1.73	79.17 ± 1.76	78.67 ± 1.76	75.33 ± 1.15	75.00 ± 1.60	**79.33 ± 1.89**
	0.3	64.83 ± 3.52	68.83 ± 2.49	70.33 ± 2.56	69.50 ± 3.01	69.17 ± 2.94	62.50 ± 3.23	65.50 ± 3.49	**70.83 ± 1.94**
	0.4	53.07 ± 2.13	56.80 ± 1.36	57.87 ± 2.32	56.27 ± 4.49	56.53 ± 4.81	45.33 ± 2.53	50.67 ± 2.39	**58.40 ± 2.29**

Table 6: The classification accuracies (%) of the competing methods with the corresponding standard deviations under the situations of the random pixel corruptions.

Data	Ratios	SRC	CRC	CCRC	ProCRC	EProCRC	Co-CRC	DSRC	DCCRC
IMM	0.1	65.17 ± 1.11	64.67 ± 0.67	65.33 ± 1.55	65.67 ± 1.86	63.17 ± 2.32	62.17 ± 2.39	63.50 ± 1.11	**66.00 ± 2.07**
	0.2	58.67 ± 1.35	59.33 ± 1.33	58.17 ± 1.22	58.17 ± 1.22	57.33 ± 1.33	53.00 ± 1.63	57.17 ± 0.85	**59.67 ± 1.80**
	0.3	51.33 ± 1.35	52.17 ± 1.55	49.33 ± 1.62	49.67 ± 1.80	48.67 ± 2.27	43.33 ± 2.74	48.172.07	**53.17 ± 3.00**
	0.4	40.50 ± 3.64	40.50 ± 3.10	39.83 ± 2.26	40.00 ± 2.11	39.83 ± 2.66	32.50 ± 2.04	32.17 ± 3.60	**40.83 ± 2.30**
Yale	0.1	87.73 ± 2.29	89.87 ± 1.81	90.67 ± 1.69	90.67 ± 1.69	89.33 ± 1.89	**90.93 ± 1.96**	89.87 ± 2.17	**90.93 ± 1.96**
	0.2	83.47 ± 1.81	85.87 ± 1.36	86.93 ± 1.31	87.47 ± 1.07	85.07 ± 1.00	86.13 ± 1.81	86.40 ± 1.77	**88.00 ± 1.19**
	0.3	79.73 ± 4.25	80.27 ± 2.59	83.73 ± 2.72	83.20 ± 2.61	79.73 ± 1.96	83.20 ± 2.87	83.20 ± 3.44	**84.53 ± 2.32**
	0.4	69.07 ± 2.44	74.67 ± 2.23	75.73 ± 1.55	74.67 ± 2.53	70.13 ± 3.22	72.00 ± 1.19	76.00 ± 1.89	**76.27 ± 1.55**

0.4 with a step 0.1. As an example, the testing samples with the random pixels from one class are shown in Figure 5 and with the random occlusions shown in Figure 6.

The classification results of the competing methods on IMM and Yale with varying ratios of the random corruptions are shown in Table 5 for random occlusions and in Table 6 for random corrected pixels. Note that the best classification performance among all the methods in two tables is indicated in bold face. As listed in two tables, the classification accuracies of each competing method descend with the increase of the ratios of the corrupted size of each testing image. From these experimental results, we can see that the proposed DCCRC is nearly the most robust among all the competing methods, since it outperforms the other competing methods. Thus, the proposed method has more effectiveness and robustness under the situations of data with noises.

5. Conclusions

Collaborative representation-based classification is a typical technique for pattern recognition. To further improve pattern discrimination in collaborative representation, we design a new discriminative, competitive, and collaborative representation-based classification method (DCCRC) in this article. The proposed DCCRC extends the Co-CRC method and mainly designs a discriminative regularization of the collaborative representation from all the classes and the ones from all the classes excluding any one class. The proposed method fully utilizes the class-specific representations in collaborative representation and can competitively and discriminatively enhance the class-specific representations for good classification. The extensive experiments on some face databases and UCI data sets are conducted for verifying the effectiveness and robustness of the proposed DCCRC method. Through comparing DCCRC with the state-of-the-art representation-based classification methods, the proposed DCCRC outperforms the competing methods. Thus, the proposed DCCRC is an effective and robust classifier in pattern recognition. In the future work, we will employ the idea of the competitive and collaborative representation among all the classes into the other kinds of classifiers.

Data Availability

The UCI and face data used in our article to support the findings of this study have been deposited in their corresponding public repository. The authors have given the websites, from which the used data can be downloaded.

Conflicts of Interest

All the authors declare that there are no conflicts of interest regarding the publication of this article.

Acknowledgments

This work was supported in part by the National Natural Science Foundation of China (Grant nos. 61976107, 61962010, and 61502208), Natural Science Foundation of Jiangsu Province of China (Grant nos. BK20150522 and BK20170558), International Postdoctoral Exchange Fellowship Program of China Postdoctoral Council (no. 20180051), Research Foundation for Talented Scholars of Jiangsu University (Grant no. 14JDG037), China Postdoctoral Science Foundation (Grant no. 2015M570411), Open Foundation of Artificial Intelligence Key Laboratory of Sichuan Province (Grant no. 2017RYJ04), Natural Science Foundation of Guizhou Province (nos. [2017]1130 and [2017]5726-32), Natural Science Foundation of Ningxia of China (no. 2019AAC03122), and Key Science and Research Project of North Minzu University (no. 2019KJ43).

References

[1] J. Wright, A. Y. Yang, A. Ganesh, S. S. Sastry, and Y. Ma, "Robust face recognition via sparse representation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 31, no. 2, pp. 210–227, 2009.

[2] L. Zhang, M. Yang, and X. Feng, "Sparse representation or collaborative representation: which helps face recognition?," in *Proceedings of the International Conference on Computer Vision*, pp. 471–478, Tokyo, Japan, May 2011.

[3] Y. Chen, N. M. Nasrabadi, T. D. Tran et al., "Hyperspectral image classification using dictionary-based sparse representation," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 49, no. 10, pp. 3973–3985, 2011.

[4] W. Li, J. Liu, and Q. Du, "Sparse and low-rank graph for discriminant analysis of hyperspectral imagery," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 54, no. 7, pp. 4094–4105, 2016.

[5] B. Tu, S. Huang, L. Fang et al., "Hyperspectral image classification via weighted joint nearest neighbor and sparse representation," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 11, no. 11, pp. 4063–4075, 2018.

[6] J. Gou, B. Hou, W. Ou, Q. Mao, H. Yang, and Y. Liu, "Several robust extensions of collaborative representation for image classification," *Neurocomputing*, vol. 348, pp. 120–133, 2019.

[7] X. Peng, J. Feng, S. Xiao, W.-Y. Yau, J. T. Zhou, and S. Yang, "Structured autoencoders for subspace clustering," *IEEE Transactions on Image Processing*, vol. 27, no. 10, pp. 5076–5086, 2018.

[8] C. Chen, N. Chen, J. T. Peng et al., "Nearest regularized joint sparse representation for hyperspectral image classification," *IEEE Geoscience and Remote Sensing Letters*, vol. 13, no. 3, pp. 424–428, 2016.

[9] S. Hao, W. Wang, Y. Yan, and L. Bruzzone, "Class-wise dictionary learning for hyperspectral image classification," *Neurocomputing*, vol. 220, pp. 121–129, 2017.

[10] W. Li and Q. Du, "Laplacian regularized collaborative graph for discriminant analysis of hyperspectral imagery," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 54, no. 12, pp. 7066–7076, 2016.

[11] Y. Xu, Z. Li, J. Yang, and D. Zhang, "A survey of dictionary learning algorithms for face recognition," *IEEE Access*, vol. 5, pp. 8502–8514, 2017.

[12] T. Shu, B. Zhang, and Y. Y. Tang, "Sparse supervised representation-based classifier for uncontrolled and imbalanced classification," *IEEE Transactions on Neural Networks and Learning Systems*, pp. 1–10, 2018.

[13] M. Yang, L. Zhang, X. Feng, and D. Zhang, "Sparse representation based Fisher discrimination dictionary learning for image classification," *International Journal of Computer Vision*, vol. 109, no. 3, pp. 209–232, 2014.

[14] W. Deng, J. Hu, and J. Guo, "Extended SRC: undersampled face recognition via intraclass variant dictionary," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 34, no. 9, pp. 1864–1870, 2012.

[15] Z. Zhao, G. Feng, L. Zhang, J. Zhu, and Q. Shen, "Novel orthogonal based collaborative dictionary learning for efficient face recognition," *Knowledge-Based Systems*, vol. 163, pp. 533–545, 2019.

[16] X. Peng, J. Lu, Z. Yi, and R. Yan, "Automatic subspace learning via principal coefficients embedding," *IEEE Transactions on Cybernetics*, vol. 47, no. 11, pp. 3583–3596, 2016.

[17] J. Gou, L. Wang, Z. Yi, J. Lv, Q. Mao, and Y.-H. Yuan, "A new discriminative collaborative neighbor representation method for robust face recognition," *IEEE Access*, vol. 6, pp. 74713–74727, 2018.

[18] W. Deng, J. Hu, and J. Guo, "Face recognition via collaborative representation: its discriminant nature and superposed representation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 40, no. 10, pp. 2513–2521, 2017.

[19] S. Zeng, J. Gou, and L. Deng, "An antinoise sparse representation method for robust face recognition via joint l_1 and l_2 regularization," *Expert Systems with Applications*, vol. 82, pp. 1–9, 2017.

[20] Z. Zhang, Y. Xu, J. Yang, X. Li, and D. Zhang, "A survey of sparse representation: algorithms and applications," *IEEE Access*, vol. 3, pp. 490–530, 2015.

[21] X. Peng, C. Lu, Z. Yi, and H. Tang, "Connections between nuclear-norm and frobenius-norm-based representations," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 29, no. 1, pp. 218–224, 2016.

[22] S. Cai, L. Zhang, W. Zuo et al., "A probabilistic collaborative representation based approach for pattern classification," in *Proceedings of the IEEE Conference on Computer Vision and Pattern*, pp. 2950–2959, Las Vegas, NV, USA, June 2016.

[23] J. Xu, W. An, L. Zhang, and D. Zhang, “Sparse, collaborative, or nonnegative representation: which helps pattern classification?,” *Pattern Recognition*, vol. 88, pp. 679–688, 2019.

[24] Z. Liu, G. Xie, L. Zhang, and J. Pu, “Fusion linear representation-based classification,” *Soft Computing*, vol. 23, no. 6, pp. 1891–1899, 2019.

[25] Y. Peng, L. Li, S. Liu, T. Lei, and J. Wu, “A new virtual samples-based CRC method for face recognition,” *Neural Processing Letters*, vol. 48, no. 1, pp. 313–327, 2018.

[26] Y. Chi and F. Porikli, “Classification and boosting with multiple collaborative representations,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 36, no. 8, pp. 1519–1531, 2013.

[27] R. Timofte and L. Van Gool, “Adaptive and weighted collaborative representations for image classification,” *Pattern Recognition Letters*, vol. 43, pp. 127–135, 2014.

[28] X. Peng, L. Zhang, Z. Yi, and K. K. Tan, “Learning locality-constrained collaborative representation for robust face recognition,” *Pattern Recognition*, vol. 47, no. 9, pp. 2794–2806, 2014.

[29] X. Dong, H. Zhang, L. Zhu et al., “Weighted locality collaborative representation based on sparse subspace,” *Journal of Visual Communication and Image Representation*, vol. 58, pp. 187–194, 2019.

[30] Y. Xu, D. Zhang, J. Yang, and J. Y. Yang, “A two-phase test sample sparse representation method for use with face recognition,” *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 21, no. 9, pp. 1255–1262, 2011.

[31] Z. Liu, J. Pu, M. Xu, and Y. Qiu, “Face recognition via weighted two phase test sample sparse representation,” *Neural Processing Letters*, vol. 41, no. 1, pp. 43–53, 2015.

[32] J. Gou, Y. Xu, D. Zhang, Q. Mao, L. Du, and Y. Zhan, “Two-phase linear reconstruction measure-based classification for face recognition,” *Information Sciences*, vol. 433-434, pp. 17–36, 2018.

[33] J. Gou, L. Wang, B. Hou, J. Lv, Y. Yuan, and Q. Mao, “Two-phase probabilistic collaborative representation-based classification,” *Expert Systems With Applications*, vol. 133, pp. 9–20, 2019.

[34] N. Akhtar, F. Shafait, and A. Mian, “Efficient classification with sparsity augmented collaborative representation,” *Pattern Recognition*, vol. 65, pp. 136–145, 2017.

[35] Z. Liu, X. J. Wu, and Z. Shu, “Sparsity augmented discriminative sparse representation for face recognition,” *Pattern Analysis and Applications*, vol. 22, no. 4, pp. 1527–1535, 2019.

[36] Y. E. Traboulsi, F. Dornaika, Y. Ruichek et al., “Semi-supervised two phase test sample sparse representation classifier,” *Knowledge-Based Systems*, vol. 160, pp. 16–27, 2018.

[37] S. Zeng, X. Yang, and J. Gou, “Multiplication fusion of sparse and collaborative representation for robust face recognition,” *Multimedia Tools and Applications*, vol. 76, no. 20, pp. 20889–20907, 2017.

[38] S. Zeng, J. Gou, and X. Yang, “Improving sparsity of coefficients for robust sparse and collaborative representation-based image classification,” *Neural Computing and Applications*, vol. 30, no. 10, pp. 2965–2978, 2018.

[39] J. Gou, W. Qiu, Z. Yi, Y. Xu, Q. Mao, and Y. Zhan, “A local mean representation-based K-nearest neighbor classifier,” *ACM Transactions on Intelligent Systems and Technology (TIST)*, vol. 10, no. 3, p. 29, 2019.

[40] H. Ma, J. Gou, X. Wang, J. Ke, and S. Zeng, “Sparse coefficient-based K-nearest neighbor classification,” *IEEE Access*, vol. 5, pp. 16618–16634, 2017.

[41] J. Gou, W. Qiu, Z. Yi, X. Shen, Y. Zhan, and W. Ou, “Locality constrained representation-based K-nearest neighbor classification,” *Knowledge-Based Systems*, vol. 167, pp. 38–52, 2019.

[42] R. S. Lan and Y. C. Zhou, “An extended probabilistic collaborative representation based classifier for image classification,” in *Proceedings of the 2017 IEEE International Conference on Multimedia and Expo (ICME)*, pp. 1392–1397, Hong Kong, China, July 2017.

[43] R. S. Lan, Y. Zhou, Z. Liu, and X. Luo, “Prior knowledge-based probabilistic collaborative representation for visual recognition,” *IEEE Transactions on Cybernetics*, pp. 1–11, 2018.

[44] Y. Xu, Z. Zhong, J. Yang, J. You, and D. Zhang, “A new discriminative sparse representation method for robust face recognition via l_2 regularization,” *IEEE Transactions on Neural Networks and Learning Systems*, vol. 28, no. 10, pp. 2233–2242, 2017.

[45] J. Gou, B. Hou, Y. Yuan, W. Ou, and S. Zeng, “A new discriminative collaborative representation-based classification method via l_2 regularizations,” *Neural Computing and Applications*, pp. 1–15, 2019.

[46] H. Yuan, X. Li, F. Xu, Y. Wang, L. L. Lai, and Y. Y. Tang, “A collaborative-competitive representation based classifier model,” *Neurocomputing*, vol. 275, pp. 627–635, 2018.

[47] H. Chi, H. Xia, L. Zhang, C. Zhang, and X. Tang, “Competitive and collaborative representation for classification,” *Pattern Recognition Letters*, 2018.