Journal of
Electronic Imaging

JElectronicImaging.org

Sparsity augmented weighted collaborative representation for image classification

Zi-Qi Li
Jun Sun
Xiao-Jun Wu
He-Feng Yin

SPIE. IS&T imaging.org

Zi-Qi Li, Jun Sun, Xiao-Jun Wu, He-Feng Yin, "Sparsity augmented weighted collaborative representation for image classification," *J. Electron. Imaging* **28**(5), 053032 (2019), doi: 10.1117/1.JEI.28.5.053032.

Journal of Electronic Imaging 28(5), 053032 (Sep/Oct 2019)

Sparsity augmented weighted collaborative representation for image classification

Zi-Qi Li,[a] Jun Sun,[a,b,*] Xiao-Jun Wu,[a,b] and He-Feng Yin[a]
[a]Jiangnan University, School of Internet of Things Engineering, Lihu Avenue, Wuxi, China
[b]Jiangsu Provincial Laboratory of Pattern Recognition and Computational Intelligence, Lihu Avenue, Wuxi, China

Abstract. Sparse representation-based classification (SRC) and collaborative representation-based classification (CRC) have garnered significant attention recently. In CRC, it is argued that it is the collaborative representation mechanism but not the ℓ_1-norm sparsity that makes SRC successful for classification tasks. However, recent studies reveal that sparsity does play a critical role in accurate classification, thus it should not be totally overlooked due to relatively high computational cost. Inspired by these findings, we propose a method called sparsity augmented weighted collaborative representation-based classification (SA-WCRC) for image classification. First, the representation coefficients of the test sample are obtained via weighted collaborative representation and sparse representation, respectively. Second, we augment the coefficient obtained by weighted collaborative representation with the sparse representation. Finally, the test sample is classified based on the augmented coefficient and the label matrix of the training samples. Both the augmented coefficient and classification scheme make SA-WCRC efficient for classification. Experiments on three face databases and one scene dataset demonstrate the superiority of SA-WCRC over its counterparts. [DOI: 10.1117/1.JEI.28.5.053032]

Keywords: sparse representation; weighted collaborative representation; image classification.

Paper 190577 received Jun. 19, 2019; accepted for publication Oct. 4, 2019; published online Oct. 26, 2019.

1 Introduction

In recent years, image classification has aroused intense interest and remarkable progress has been made in this field. Meanwhile, numerous representation and classification approaches have been designed. As a mainstream technique in pattern classification, representation-based classification method (RBCM) is widely applied to various classification tasks. One of the representative methods is sparse representation-based classification (SRC),[1] which achieves promising results in face recognition. In SRC, there are two stages: coding and classification. First, the input test sample is represented as a sparse linear superposition of atoms in the dictionary, then the classification is performed by checking which class leads to the least residual. SRC attains impressive recognition performance even when the test image is badly corrupted or occluded. During the past few years, sparse representation has also been applied to other classification tasks, such as hyperspectral image classification,[2] text classification,[3] and image set classification.[4]

SRC directly uses all the training data as the dictionary. When the size of the dictionary is big, the subsequent sparse decomposition process can be very slow. Additionally, using the original training data as the dictionary could not fully exploit the discriminative information hidden in the training data. Dictionary learning (DL) is proposed to learn a compact and discriminative dictionary from the original training data. K-singular value decomposition (KSVD)[5] is one representative DL technique that is tailored for image processing. In order to make KSVD applicable to classification tasks, Zhang and Li[6] introduced the classification error term into the objective function of KSVD and proposed discriminative K-SVD (D-KSVD). Jiang et al.[7] developed a label consistent K-SVD (LC-KSVD) method that employs a label consistency term to encourage samples from the same class to have similar representations. To fully exploit the discriminative information in the coding coefficients, Yang et al.[8] developed a Fisher discrimination dictionary learning (FDDL) method, which imposes the Fisher discrimination criterion on the coding coefficients. Inspired by KSVD and FDDL, Zheng and Tao[9] presented a DL method called Fisher discriminative KSVD. FDDL only imposes Fisher discrimination criterion on the coding coefficients, it does not take the discrimination of atoms into consideration. Very recently, Li et al.[10] designed a discriminative Fisher embedding dictionary learning algorithm, which simultaneously establishes Fisher embedding models on the learned atoms and coefficients.

Another representative RBCM is collaborative representation-based classification (CRC). Zhang et al.[11] argued that it is the collaborative representation mechanism rather than the ℓ_1-norm sparsity that plays the crucial role for classification in SRC. Therefore, they proposed CRC, which replaces the ℓ_1-norm with the ℓ_2-norm. Compared with SRC, CRC can attain competitive performance but the computational cost is significantly reduced. Cai et al.[12] analyzed the classification mechanism of CRC from a probabilistic perspective and proposed a probabilistic collaborative representation-based classifier for pattern classification. Zheng and Wang[13] pointed out that it is unnecessary to use all the training data to represent a test sample. Thus they developed a k-nearest classes-based CRC scheme.

Though RBCM and its extensions achieve promising results in a variety of classification tasks, they emphasize too much on the role of ℓ_1-norm sparsity or collaborative representation. Recently, Akhtar et al.[14] revealed that sparseness

*Address all correspondence to Jun Sun, E-mail: Junsun@jiangnan.edu.cn

1017-9909/2019/$28.00 © 2019 SPIE and IS&T

of collaborative representation explicitly contributes to accurate classification. And they presented a sparsity augmented collaborative representation-based classification (SA-CRC), which utilizes both dense and sparse collaborative representations to classify a test sample. Nevertheless, locality information is ignored in SA-CRC, and Yu et al.[15] argued that under certain assumptions locality is more essential than sparsity. To overcome the drawback of SA-CRC, first we incorporate the locality information into the formulation of CRC and present weighted CRC (WCRC), then we augment the representation of WCRC with a sparse representation to further promote the sparsity of WCRC. Moreover, different from the classification rule used by CRC and WCRC, we exploit an efficient classification scheme that avoids calculating the reconstruction error class by class. The proposed method is termed as sparsity augmented weighted collaborative representation-based classification (SA-WCRC). Experimental results on several publicly available datasets demonstrate the effectiveness of our proposed SA-WCRC. The demo code of our proposed SA-WCRC is available at https://github.com/li-zi-qi/SA-WCRC.

The main contributions in this paper are summarized as follows.

- We present an effective SA-WCRC approach, which jointly takes sparsity of the coefficient and locality information of samples into account.
- We enhance the sparsity of WCRC by augmenting the representation of WCRC with a sparse representation.
- The classification stage of our proposed SA-WCRC only involves the multiplication of the label matrix of training data and the augmented coefficient of a test sample, which is very efficient to recognize the test sample.

The remainder of this paper is organized as follows. Section 2 outlines the related work. Section 3 presents our SA-WCRC approach. Section 4 analyzes our proposed SA-WCRC. Experimental evaluation on several benchmark databases is provided in Sec. 5. Finally, Sec. 6 concludes this paper.

2 Related Work

Suppose there are n training samples from C classes. Each training sample is represented as a vector corresponding to the i'th column of the dictionary. Thus all training samples constitute the matrix $\mathbf{X} = (\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_C) \in \mathbb{R}^{m\times n}$ and $\mathbf{X}_i = (\boldsymbol{x}_{i,1}, \boldsymbol{x}_{i,2}, \dots, \boldsymbol{x}_{i,n_i}) \in \mathbb{R}^{m\times n_i}$, $i = 1, 2, \dots, C$, where m is the dimensionality of each sample and n_i denotes the number of training samples in the i'th class.

2.1 *Sparse Representation-based Classification*

In SRC, all training data are used as the dictionary and an input test sample $\boldsymbol{y} \in \mathbb{R}^m$ is expressed as a linear combination of atoms in the dictionary, i.e.,

$$\boldsymbol{y} \approx \mathbf{X}\boldsymbol{\alpha}, \tag{1}$$

where $\boldsymbol{\alpha} = (\boldsymbol{\alpha}_1; \dots; \boldsymbol{\alpha}_i; \dots; \boldsymbol{\alpha}_C) \in \mathbb{R}^{n\times 1}$ and $\boldsymbol{\alpha}_i$ is the coefficient corresponds to the i'th class. If $\boldsymbol{y}$ belongs to the i'th class, then $\boldsymbol{y} \approx \mathbf{X}_i\alpha_i$ holds well, indicating that most coefficients in α_k, $k \neq i$, tend to zeros and only α_i has nonzeros elements. As a result, the sparse coefficient vector α contains the identity information of $\boldsymbol{y}$. The objective function of SRC is formulated as

$$\min_{\alpha} \|\boldsymbol{\alpha}\|_1, \quad \text{s.t. } \|\boldsymbol{y} - \mathbf{X}\boldsymbol{\alpha}\|_2^2 \leq \varepsilon, \tag{2}$$

where ε is a given error tolerance. Equation (2) can be rewritten as the following unconstrained formulation:

$$\min_{\alpha} \|\boldsymbol{y} - \mathbf{X}\boldsymbol{\alpha}\|_2^2 + \lambda\|\boldsymbol{\alpha}\|_1, \tag{3}$$

where λ is a balancing parameter. Once we obtain the coefficient vector α, the classification can be performed by checking which class leads to the least residual:

$$\text{identity}(\boldsymbol{y}) = \arg\min_i \|\boldsymbol{y} - \mathbf{X}_i\alpha_i\|_2. \tag{4}$$

2.2 *Collaborative Representation-based Classification*

The major difference between SRC and CRC is that CRC replaces the ℓ_1-norm constraint on the coefficient with the ℓ_2-norm, the objective function of CRC is as follows:

$$\min_{\boldsymbol{\alpha}} \|\boldsymbol{y} - \mathbf{X}\boldsymbol{\alpha}\|_2^2 + \lambda\|\boldsymbol{\alpha}\|_2^2. \tag{5}$$

Equation (5) has a closed form solution, which is formulated as $\boldsymbol{\alpha} = \mathbf{P}\boldsymbol{y}$, where $\mathbf{P} = (\mathbf{X}^{\mathrm{T}}\mathbf{X} + \lambda\mathbf{I})^{-1}\mathbf{X}^{\mathrm{T}}$.

In the classification stage, apart from the classwise residual, Zhang et al.[11] pointed out that the coefficients of each class $\boldsymbol{\alpha}_i$ contain some discriminative information for classification. Thus both the classwise residual and coefficients are exploited in classification, leading to the following classification rule:

$$\text{identity}(\boldsymbol{y}) = \arg\min_i \frac{\|\boldsymbol{y} - \mathbf{X}_i\alpha_i\|_2}{\|\boldsymbol{\alpha}_i\|_2}. \tag{6}$$

3 Proposed Approach

Though CRC achieves impressive results for various tasks, it does not take the locality information into account. Researches reveal that the local structure of data contains meaningful information than that of global structure. For instance, a sample and its neighbors may have the same class label. Therefore, similar samples (neighbors) should have similar representations. This goal can be attained by the following WCRC problem:

$$\min_{\breve{\boldsymbol{\alpha}}} \|\boldsymbol{y} - \mathbf{X}\breve{\boldsymbol{\alpha}}\|_2^2 + \lambda\|\mathbf{W}\breve{\boldsymbol{\alpha}}\|_2^2, \tag{7}$$

where $\mathbf{W}$ is the locality constraint matrix and it is a diagonal matrix, $\mathbf{W}_{ii} = \exp(\frac{d_i}{\sigma})$ and $d_i = \|\boldsymbol{y} - \boldsymbol{x}_i\|_2$. σ is employed to adjust the weight decay speed for the locality adaptor. Equation (7) has a closed-form solution, which is formulated as

$$\breve{\alpha} = (\mathbf{X}^{\mathrm{T}}\mathbf{X} + \lambda\mathbf{W}^{\mathrm{T}}\mathbf{W})^{-1}\mathbf{X}^{\mathrm{T}}\boldsymbol{y}. \tag{8}$$

In this paper, similar to SA-CRC,[14] we obtain the sparse representation via OMP,[16] and the objective function is:

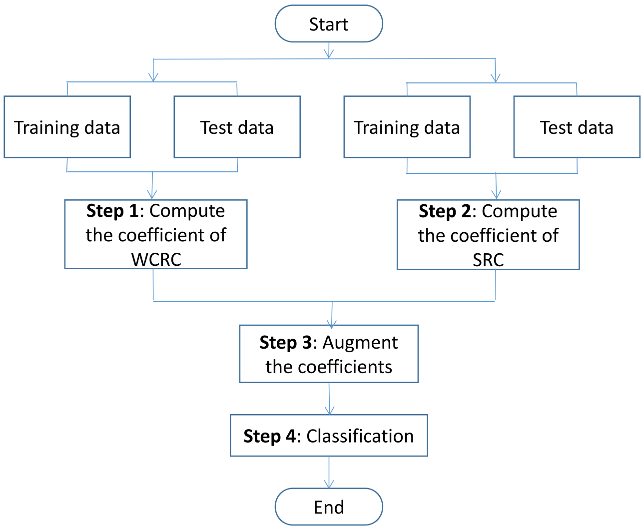

Fig. 1 Block diagram of proposed method.

$$\min_{\hat{\boldsymbol{\alpha}}} \|\boldsymbol{y} - \mathbf{X}\hat{\boldsymbol{\alpha}}\|_2, \quad \text{s.t. } \|\hat{\boldsymbol{\alpha}}\|_0 \leq k, \tag{9}$$

where k is the sparsity level.

When the coefficients of WCRC and SRC are obtained, the augmented coefficient $\mathring{\boldsymbol{\alpha}}$ is derived by employing the following formulation:

$$\mathring{\boldsymbol{\alpha}} = \frac{\hat{\boldsymbol{\alpha}} + \check{\boldsymbol{\alpha}}}{\|\hat{\boldsymbol{\alpha}} + \check{\boldsymbol{\alpha}}\|_2}. \tag{10}$$

Suppose $\mathbf{L} = (\boldsymbol{l}_1, \boldsymbol{l}_2, \ldots, \boldsymbol{l}_n) \in \mathbb{R}^{C\times n}$ is the label matrix of training data, and $\boldsymbol{l}_j = (0, 0, \ldots, 1, \ldots, 0, 0)^{\mathrm{T}} \in \mathbb{R}^{C\times 1}$ is the label vector of the j'th training sample. For the i'th class, $\mathbf{L}$ has n_i nonzero entries in its i'th row, at the indices associated with the columns of $\mathbf{X}_i$. It should be noted that $\mathbf{X}_i$ represents the subset of dictionary atoms stemming from the i'th class. Therefore, the i'th element of the vector $\boldsymbol{q} = \mathbf{L}\mathring{\boldsymbol{\alpha}}$ denotes the sum of coefficients in $\mathring{\boldsymbol{\alpha}}$ that correspond to the atoms in $\mathbf{X}_i$, and $\boldsymbol{q}$ is called the score of each class. As a result, the identity of the test sample can be determined by maximizing the coefficients of $\boldsymbol{q}$.

Our proposed SA-WCRC has the following steps. First, the weighted collaborative coefficient and sparse coefficient are obtained by solving Eqs. (8) and (9), respectively. Second, the augmented coefficient is attained by adding the above two coefficient vectors and normalize the result. Finally, the test sample is classified on the basis of the augmented coefficient vector and the label matrix of the training data. Figure 1 depicts the block diagram of SA-WCRC and Algorithm 1 summarizes the procedure of our proposed scheme.

Algorithm 1 SA-WCRC

Input: Training data matrix $\mathbf{X} = (\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_C) \in \mathbb{R}^{m\times n}$ and label matrix $\boldsymbol{L}$, test data $\boldsymbol{y} \in \mathbb{R}^m$, parameters λ and σ for WCRC, sparsity level k for SRC.

Output: label$(\boldsymbol{y}) = \arg\max_i(\boldsymbol{q}_i)$

1. Compute the coefficient $\check{\boldsymbol{\alpha}}$ of WCRC via Eq. (8)
2. Obtain the sparse coefficient $\hat{\boldsymbol{\alpha}}$ of SRC by Eq. (9)
3. Compute the augmented coefficient $\mathring{\boldsymbol{\alpha}} = \frac{\hat{\boldsymbol{\alpha}}+\check{\boldsymbol{\alpha}}}{\|\hat{\boldsymbol{\alpha}}+\check{\boldsymbol{\alpha}}\|_2}$
4. Compute $\boldsymbol{q} = \mathbf{L}\mathring{\boldsymbol{\alpha}}$

4 Analysis of SA-WCRC

In this section, we analyze the effectiveness of SA-WCRC from some experimental results. The experiments are conducted on the Yale database, and this database consists of 15 subjects and each subject has 11 images. Six images per person are used to form the dictionary, so the dictionary has 90 atoms. We choose a test sample from the first class, and the sparse coefficients and corresponding residual for each class are plotted in Figs. 2 and 3. We can see from Fig. 2 that coefficients belong to the first class are all positive. Though the 10th (red rectangle) and 15th classes have relatively large positive coefficients, the negative coefficients in them are prominent. Therefore, the residual for the 10th or

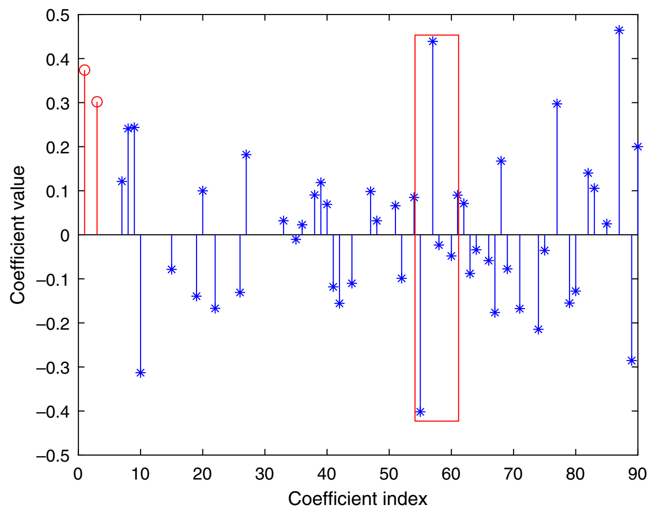

Fig. 2 Coefficients obtained by SRC. The coefficients in the red rectangle correspond to the 10th class.

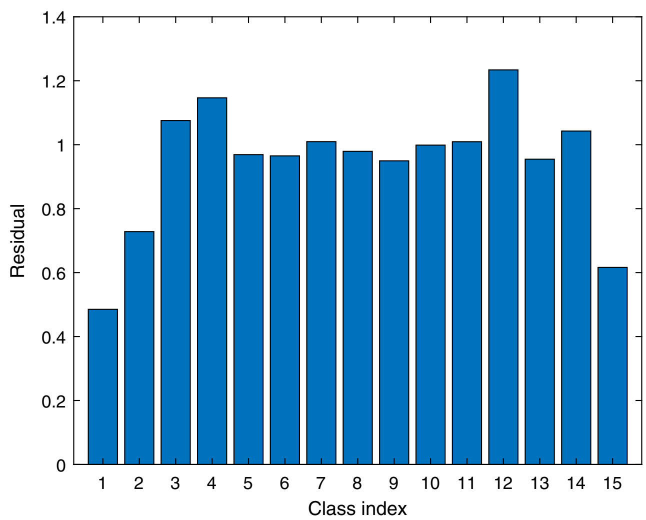

Fig. 3 The residual of SRC for each class and the first class has the minimal residual.

15th will not be the least. From Fig. 3, we can clearly see that the first class leads to the minimal residual, which indicates that SRC can correctly classify the test sample. Figure 4 presents the coefficients obtained by WCRC, and the 15th class has dominant coefficients. Figure 5 shows the residual of WCRC, one can see that the 15th class has the least residual, thus the test sample is wrongly classified to the 15th class. Coefficients computed by SA-WCRC are shown in Fig. 6, we can see that coefficients from the first class are all positive. Figure 7 illustrates the score of SA-WCRC for each class, it can be seen that the first class has the largest value. As a result, the test sample is designated to the first class by SA-WCRC. From the above experimental results, we can find that the dense representation of WCRC may lead to wrong classification. By augmenting the dense representation with a sparse representation, the wrong classification can be avoided. This demonstrates the efficacy of our proposed SA-WCRC.

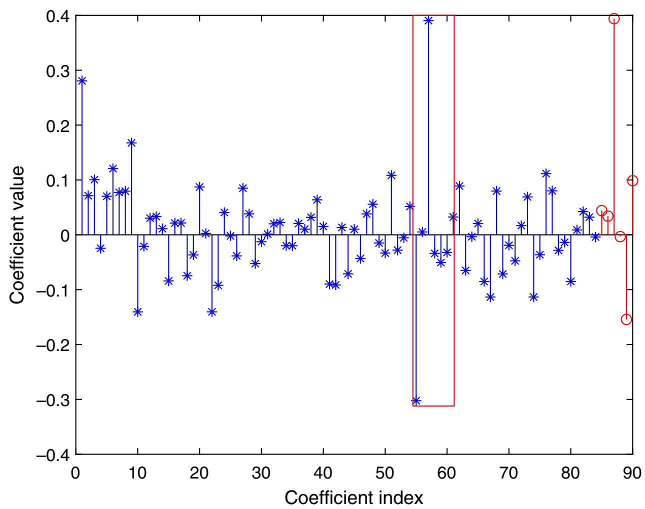

Fig. 4 Coefficients computed by WCRC. The coefficients in the red rectangle correspond to the 10th class.

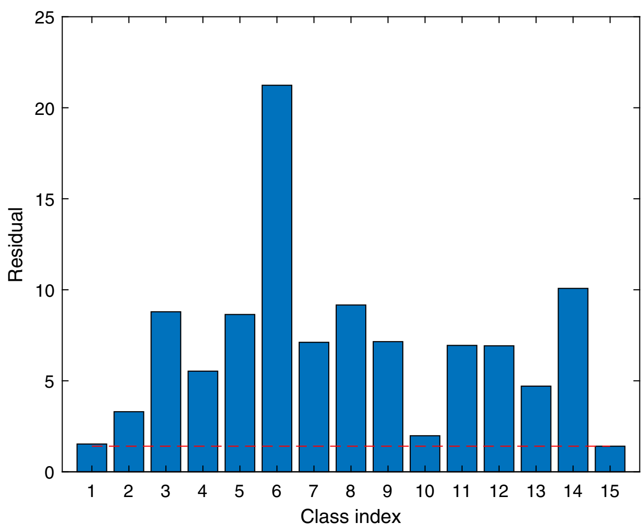

Fig. 5 The residual of WCRC for each class and the 15th class has the minimal residual.

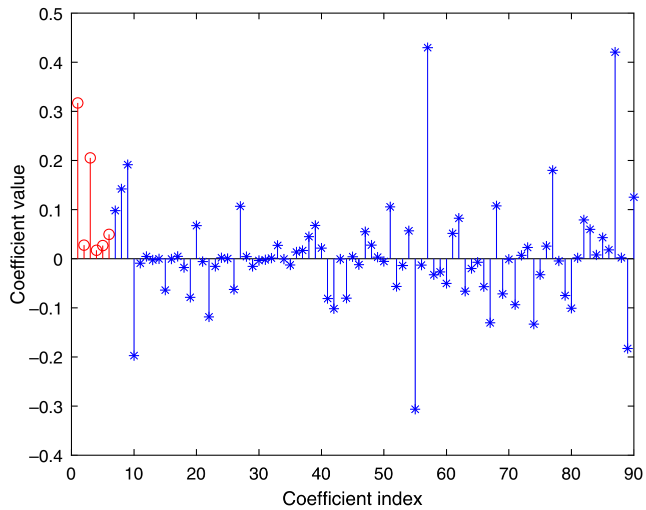

Fig. 6 Coefficients obtained by SA-WCRC.

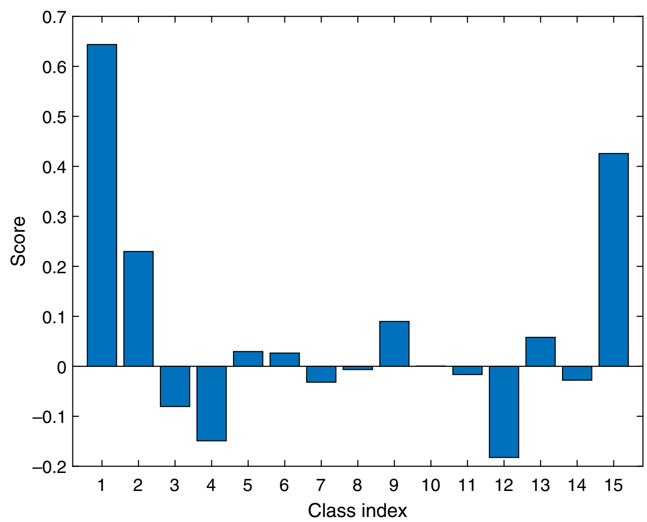

Fig. 7 The score of SA-WCRC for each class and the first class has the largest value.

5 Experiments

In this section, we evaluate the classification performance of our proposed SA-WCRC on four publicly available datasets: the Yale database,[17] the extended Yale B database,[18] the AR database,[19] and the scene 15 dataset,[20] the details of these datasets are summarized in Table 1. To illustrate the superiority of SA-WCRC, we mainly compare our method with the following approaches: SRC,[1] CRC,[11] WCRC, D-KSVD,[6] LC-KSVD,[7] FDDL,[8] COPAR,[21] JBDC,[22] and SA-CRC.[14] For SRC, we solve the problem in Eq. (2) as in Ref. 1. For CRC, LC-KSVD, FDDL, COPAR, JBDC, and SA-CRC, we use the codes released by their authors. WCRC employs the same classification rule as CRC, which is shown in Eq. (6). We modified the code of LC-KSVD to implement D-KSVD. For SA-CRC and our proposed method, OMP is employed to obtain the sparse representation. We use the same value of sparsity level ($k = 50$) as in SA-CRC,[14] for the other two parameters, we choose their optimal values by 10-fold cross-validation, and the optimal values of parameters on each

Table 1 Details of datasets used in our experiments. The columns from left to right are the names of datasets, total number of samples, number of classes, number of training samples, number of test samples, and the dimensionality of features.

Dataset	# Sample	# Class	# Training	# Test	# Feature
Yale	165	15	90	75	576
Ext. Yale B	2414	38	760	1654	504
AR	2600	100	1000	1600	540
Scene 15	4485	15	750	3735	3000

Table 2 The parameters of proposed SA-WCRC on each dataset.

Parameters	Yale	Ext. Yale B	AR	Scene 15
λ	1×10^{-3}	1×10^{-3}	1×10^{-3}	1×10^{-3}
σ	0.2	0.4	0.5	0.5
k	50	50	50	50

dataset are recorded in Table 2. All experiments are run with MATLAB R2019a under Windows 10 on PC equipped with 3.60 GHz CPU and 16 GB RAM.

5.1 *Yale Database*

Yale database[17,23] consists of 165 images for 15 subjects, each has 11 images. These images have illumination and expression variations, Fig. 8 shows some example images from this database. All the images are resized to 24×24, resulting in a 576-dimensional vector. Following the experimental settings in Ref. 24, six images per subject are randomly selected for training and the remaining for testing, and the training–testing ratio is 6:5. The error tolerance ε of SRC is 0.05, and the balancing parameter λ of CRC is 0.001. The sparsity level and number of atoms for D-KSVD and LC-KSVD are 30 and 60, respectively. Four label-particular atoms for each class and five common atoms in COPAR. Sparsity levels k and λ of SA-CRC are set to be 50 and 0.002, respectively. Experimental results are summarized in Table 3. Additionally, the average time (in milliseconds) for each competing method to classify a test sample is presented. One can see that SA-WCRC achieves the highest recognition accuracy, with a 25% reduction in the error rate of WCRC, and about 3% reduction in that of SA-CRC. Due to the joint dictionary and classifier learning framework, D-KSVD, LC-KSVD, and JBDC take less time than FDDL

Fig. 8 Example images from the Yale database.

Table 3 Recognition accuracy (%) and average time (in milliseconds) to classify one test sample of competing approaches on the Yale database.

Methods	Accuracy	Time (ms)
SRC[1]	95.06 ± 3.32	0.14
CRC[11]	94.53 ± 2.97	0.14
WCRC	94.27 ± 2.60	1.44
D-KSVD[6]	94.26 ± 2.88	0.05
LC-KSVD[7]	94.53 ± 3.17	0.04
FDDL[8]	95.73 ± 3.00	21.94
COPAR[21]	91.33 ± 4.23	5.40
JBDC[22]	94.93 ± 2.72	0.01
SA-CRC[14]	95.60 ± 2.59	0.01
SA-WCRC	$\mathbf{95.73}\pm\mathbf{2.72}$	0.01

Fig. 9 Example images from the extended Yale B database.

Table 4 Recognition accuracy (%) and average time (in milliseconds) to classify one test sample of competing approaches on the extended Yale B database

Methods	Accuracy	Time (ms)
SRC[1]	93.18 ± 0.55	0.54
CRC[11]	94.77 ± 0.48	0.01
WCRC	95.91 ± 0.54	20.03
D-KSVD[6]	90.79 ± 0.51	0.01
LC-KSVD[7]	91.48 ± 0.69	0.01
FDDL[8]	92.32 ± 0.68	0.46
COPAR[21]	90.81 ± 0.55	0.26
JBDC[22]	94.74 ± 0.83	0.01
SA-CRC[14]	95.52 ± 0.73	0.01
SA-WCRC	$\mathbf{96.00}\pm\mathbf{0.67}$	0.01

and COPAR. Our proposed SA-WCRC is more efficient than SRC and CRC.

5.2 *Extended Yale B Database*

The extended Yale B face database[18,25] contains 2414 images of 38 individuals. Each individual has 59 to 64 images taken under different illumination conditions, example images from this dataset are shown in Fig. 9. In our experiments, each 192×168 image is projected onto a 504-dimensional space via random projection. Following the experimental settings in Ref. 14, 20 images per person are selected for training and the remaining for testing, and the training–testing ratio is 20:44. We use the error tolerance of 0.05 for SRC, and the regularization parameter $\lambda = 0.001$ for CRC. The sparsity level and number of atoms for D-KSVD and LC-KSVD are 50 and 400, respectively. We use seven label-particular atoms for each class and five shared atoms in COPAR. Sparsity levels k and λ of SA-CRC are set to be 50 and 0.005, respectively. Table 4 lists the recognition accuracy and average time (in milliseconds) to recognize a test sample of the comparison methods. It can be seen that our proposed SA-WCRC is superior to its competing approaches.

Fig. 10 Example images from the AR database.

Table 5 Recognition accuracy (%) and average time (in milliseconds) to classify one test sample of competing approaches on the AR database.

Methods	Accuracy	Time (ms)
SRC[1]	91.25 ± 1.17	0.33
CRC[11]	92.04 ± 0.83	0.02
WCRC	92.06 ± 0.75	36.54
D-KSVD[6]	90.31 ± 1.13	0.01
LC-KSVD[7]	89.31 ± 1.27	0.01
FDDL[8]	91.01 ± 0.99	0.24
COPAR[21]	89.06 ± 1.54	0.19
JBDC[22]	90.97 ± 0.79	0.01
SA-CRC[14]	93.74 ± 0.84	0.01
SA-WCRC	$\mathbf{93.90} \pm \mathbf{0.80}$	0.01

5.3 *AR Database*

The AR database[19,26] has more than 4000 face images of 126 subjects with variations in facial expression, illumination conditions and occlusions, Fig. 10 shows example images from this database. We use a subset of 2600 images of 50 male and 50 female subjects from the database. Each 165×120 face image is projected onto a 540-dimensional vector by random projection. Following the experimental settings in Ref. 14, 10 images per person are randomly selected for training and the remaining for testing, and the training–testing ratio is 10:16. The error tolerance of SRC is 0.05, and the balancing parameter of CRC is 0.0014. The sparsity level and number of atoms for D-KSVD and LC-KSVD are 50 and 600, respectively. Seven label-particular atoms for each class and five common atoms in COPAR. Sparsity levels k and λ of SA-CRC are set to be 50 and 0.002, respectively. Experimental results are shown in Table 5. Moreover, the average time (in milliseconds) for each compared approach to classify a test sample is given. We can see that the best classification result is achieved by our proposed SA-WCRC, with a 23% reduction in the error rate of WCRC. In the classification stage, WCRC computes residuals class by class, whereas SA-WCRC only involves the multiplication of the label matrix of training data and the augmented coefficient of a test sample. Therefore, it is very efficient to classify a test sample by SA-WCRC. Meanwhile, from Table 5, one can see that SA-WCRC is faster than WCRC.

Fig. 11 Example images from the scene 15 dataset.

5.4 *Scene 15*

This dataset[27] contains 15 natural scene categories introduced in Ref. 20 and contains a wide range of indoor and outdoor scenes, such as bedroom, office, and mountain; example images from this dataset are shown in Fig. 11.

Table 6 Recognition accuracy (%) and average time (in milliseconds) to classify one test sample of competing approaches on the scene 15 dataset.

Methods	Accuracy	Time (ms)
SRC[1]	95.41 ± 0.13	0.23
CRC[11]	96.15 ± 0.33	0.09
WCRC	92.83 ± 0.70	77.10
D-KSVD[6]	95.12 ± 0.18	0.01
LC-KSVD[7]	96.37 ± 0.28	0.01
FDDL[8]	94.08 ± 0.43	0.18
COPAR[21]	96.02 ± 0.28	0.15
JBDC[22]	97.36 ± 0.32	0.01
SA-CRC[14]	97.18 ± 0.25	0.01
SA-WCRC	$\mathbf{97.42 \pm 0.27}$	0.01

For fair comparison, we employ the 3000-dimensional SIFT-based features used in LC-KSVD.[7] Following the experimental setting in Ref. 14, we randomly select 50 images per category as training data and use the rest for testing, and the training–testing ratio is 50∶249. The error tolerance of SRC is 1×10^{-6}, and the balancing parameter of CRC is 1. Fifty atoms are used for D-KSVD and LC-KSVD. Sparsity levels k and λ of SA-CRC are set to be 50 and 1, respectively. Recognition accuracy and the average time (in milliseconds) to classify a test sample of different approaches on this dataset are presented in Table 6. Again, SA-WCRC outperforms the comparison methods, and it is nine times faster than CRC. The confusion matrix for SA-WCRC is shown in Fig. 12, in which diagonal elements are well-marked. It can be seen that SA-WCRC attains more than 99.0% recognition accuracy for suburb, coast, inside-city, opencountry, and street.

5.5 *Parameter Sensitiveness Analysis*

Parameter selection is a critical step for various tasks in pattern recognition. To evaluate the influence of parameters on the classification results, we conduct some experiments on the Yale database. We vary the value of one parameter while fixing the value of the other. Figures 13 and 14 plot the recognition accuracy of SA-WCRC with varying λ and k. For comparison, we also present the results of WCRC. In Fig. 13, we set k to be 50 and vary λ. We can clearly see that the proposed SA-WCRC always achieves better results than WCRC and SA-WCRC performs stable in quite a wide range of λ when k is fixed to an optimal value. In Fig. 14, we vary the value of sparsity level k whereas λ is fixed to 1×10^{-3} for

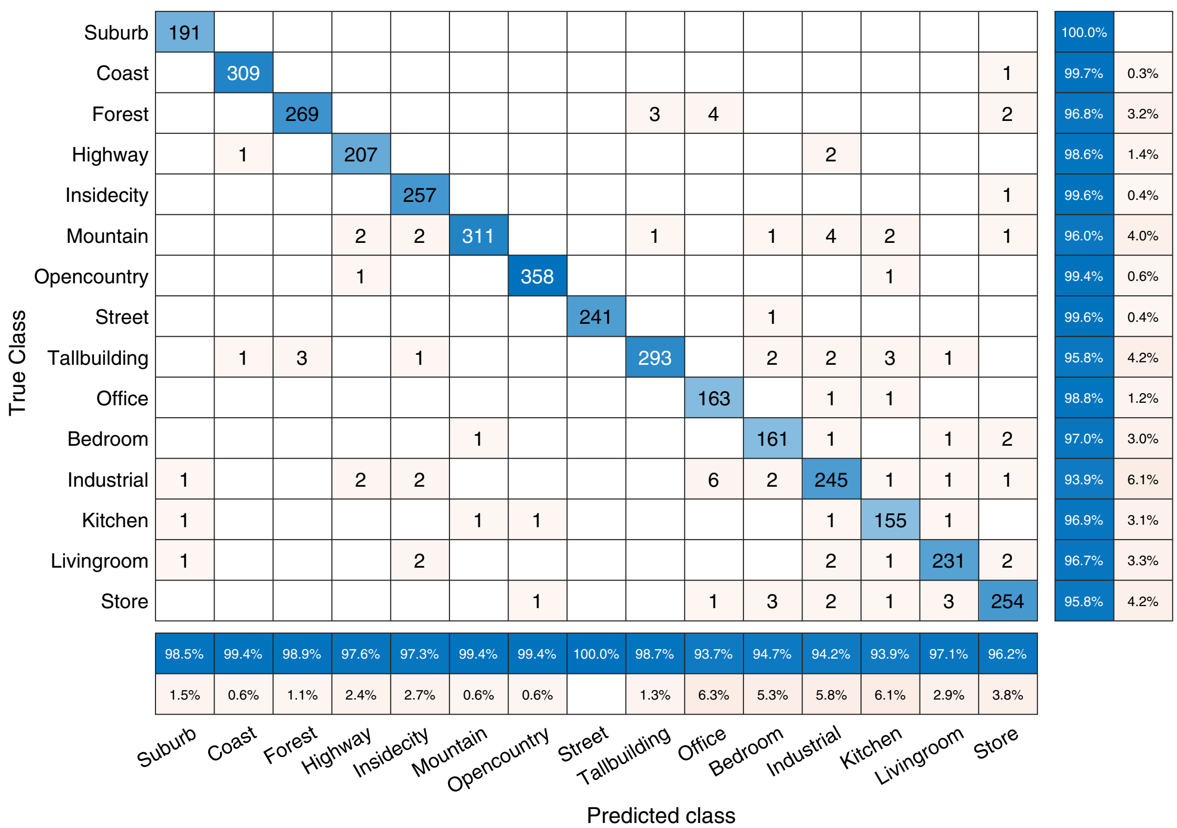

Fig. 12 Confusion matrix on the scene 15 dataset.

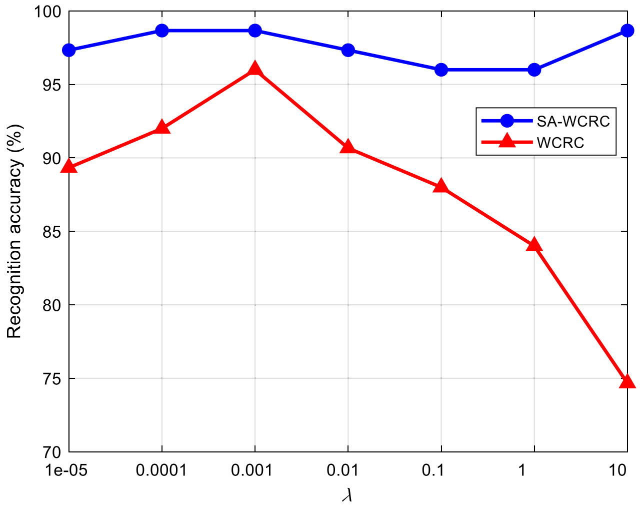

Fig. 13 The effect of λ on the recognition accuracy of SA-WCRC and WCRC for the Yale database, and k is fixed to 50.

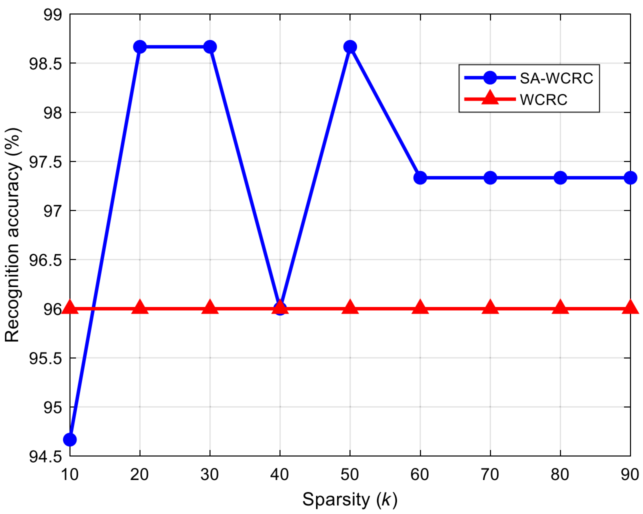

Fig. 14 The effect of k on the recognition accuracy of SA-WCRC for the Yale database, the results of WCRC are also plotted for comparison, and λ is fixed to 1×10^{-3}.

both SA-WCRC and WCRC. It can be seen that when $k \geq 20$, SA-WCRC consistently outperforms WCRC.

6 Conclusions

Though CRC and its variants claim that it is the collaborative representation that boost the performance of image classification, sparsity should not be completely ignored. Recent studies indicate sparsity does contribute to accurate classification. Therefore, we present an SA-WCRC approach that augments a dense weighted collaborative representation with a sparse representation for image classification. SA-WCRC is computationally efficient with the augmented coefficient and the classification scheme. Experimental results on the Yale, extended Yale B, AR, and scene 15 datasets demonstrate our proposed approach is superior to SA-CRC and some dictionary learning methods, like D-KSVD, LC-KSVD, FDDL, COPAR, and JBDC. In essence, SA-WCRC is a classifier belonging to RBCM, thus it is expected that it can be applied in other pattern classification tasks. Deep learning has become one of the most exciting topics in the communities of computer vision and pattern recognition. In future, we will exploit deep learning techniques to further improve the performance of RBCM for image classification.

Acknowledgments

This work was supported in part by the National Natural Science Foundation of China (Projects Nos. 61673194, 61672263, 61672265, and 61876072) and the National First-Class Discipline Program of Light Industry Technology and Engineering (Project No. LITE2018-25).

References

1. J. Wright et al., "Robust face recognition via sparse representation," *IEEE Trans. Pattern Anal. Mach. Intell.* **31**(2), 210–227 (2009).
2. J. Peng et al., "Local adaptive joint sparse representation for hyperspectral image classification," *Neurocomputing* **334**, 239–248 (2019).
3. P. Unnikrishnan, V. Govindan, and S. M. Kumar, "Enhanced sparse representation classifier for text classification," *Expert Syst. Appl.* **129**, 260–272 (2019).
4. P. Zheng et al., "Image set classification based on cooperative sparse representation," *Pattern Recognit.* **63**, 206–217 (2017).
5. M. Aharon et al., "K-SVD: an algorithm for designing overcomplete dictionaries for sparse representation," *IEEE Trans. Signal Process.* **54**(11), 4311–4322 (2006).
6. Q. Zhang and B. Li, "Discriminative K-SVD for dictionary learning in face recognition," in *IEEE Comput. Soc. Conf. Comput. Vision and Pattern Recognit.*, IEEE, pp. 2691–2698 (2010).
7. Z. Jiang, Z. Lin, and L. S. Davis, "Label consistent K-SVD: learning a discriminative dictionary for recognition," *IEEE Trans. Pattern Anal. Mach. Intell.* **35**(11), 2651–2664 (2013).
8. M. Yang et al., "Fisher discrimination dictionary learning for sparse representation," in *Int. Conf. Comput. Vision*, IEEE, pp. 543–550 (2011).
9. H. Zheng and D. Tao, "Discriminative dictionary learning via fisher discrimination K-SVD algorithm," *Neurocomputing* **162**, 9–15 (2015).
10. Z. Li et al., "Discriminative fisher embedding dictionary learning algorithm for object recognition," *IEEE Trans. Neural Networks Learn. Syst.*, 1–15 (2019).
11. L. Zhang, M. Yang, and X. Feng, "Sparse representation or collaborative representation: which helps face recognition?" in *IEEE Int. Conf. Comput. Vision*, IEEE, pp. 471–478 (2011).
12. S. Cai et al., "A probabilistic collaborative representation based approach for pattern classification," in *Proc. IEEE Conf. Comput. Vision and Pattern Recognit.*, pp. 2950–2959 (2016).
13. C. Zheng and N. Wang, "Collaborative representation with k-nearest classes for classification," *Pattern Recognit. Lett.* **117**, 30–36 (2019).
14. N. Akhtar, F. Shafait, and A. Mian, "Efficient classification with sparsity augmented collaborative representation," *Pattern Recognit.* **65**, 136–145 (2017).
15. K. Yu, T. Zhang, and Y. Gong, "Nonlinear learning using local coordinate coding," in *Adv. Neural Inf. Process. Syst.*, pp. 2223–2231 (2009).
16. J. A. Tropp and A. C. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," *IEEE Trans. Inf. Theory* **53**(12), 4655–4666 (2007).
17. K.-C. Lee, J. Ho, and D. J. Kriegman, "Acquiring linear subspaces for face recognition under variable lighting," *IEEE Trans. Pattern Anal. Mach. Intell.* **27**(5), 684–698 (2005).
18. A. S. Georghiades, P. N. Belhumeur, and D. J. Kriegman, "From few to many: illumination cone models for face recognition under variable lighting and pose," *IEEE Trans. Pattern Anal. Mach. Intell.* **23**(6), 643–660 (2001).
19. A. M. Martinez, "The AR face database," CVC Technical Report24 (1998).
20. S. Lazebnik, C. Schmid, and J. Ponce, "Beyond bags of features: spatial pyramid matching for recognizing natural scene categories," in *IEEE Comput. Soc. Conf. Comput. Vision and Pattern Recognit.*, IEEE, Vol. 2, pp. 2169–2178 (2006).
21. S. Kong and D. Wang, "A dictionary learning approach for classification: separating the particularity and the commonality," *Lect. Notes Comput. Sci.* **7572**, 186–199 (2012).
22. N. Akhtar, A. Mian, and F. Porikli, "Joint discriminative Bayesian dictionary and classifier learning," in *IEEE Conf. Comput. Vision and Pattern Recognit.*, pp. 1193–1202 (2017).
23. K.-C. Lee, "Yale face database," 1997, http://cvc.cs.yale.edu/cvc/projects/yalefaces/yalefaces.html (accessed 15 October 2019).
24. X. Wang and Y. Gu, "Cross-label suppression: a discriminative and fast dictionary learning with group regularization," *IEEE Trans. Image Process.* **26**(8), 3859–3873 (2017).

25. A. S. Georghiades, "The extended yale B face database," 1998, http://vision.ucsd.edu/leekc/ExtYaleDatabase/ExtYaleB.html (accessed 15 October 2019).
26. A. M. Martinez, "AR face database webpage," 2001, http://www2.ece.ohio-state.edu/~aleix/ARdatabase.html (accessed 15 October 2019).
27. S. Lazebnik, "Scene category dataset," 2006, http://slazebni.cs.illinois.edu/research/scene_categories.zip (accessed 15 October 2019).

Zi-Qi Li received his BE degree from the School of Information Engineering, Yangzhou University, Yangzhou, China, in 2012. Currently, he is pursuing his PhD at the School of Internet of Things Engineering, Jiangnan University, Wuxi, China. His research interests include feature representation, information fusion, dictionary learning, and low-rank representation.

Jun Sun received his MSc degree in computer science and technology and his PhD in control theory and engineering from Jiangnan University, China, in 2003 and 2009, respectively. He is currently working as a full professor in the Department of Computer Science and Technology, Jiangnan University, China. He is also a vice director of Jiangsu Provincial Engineering Laboratory of Pattern Recognition and Computational Intelligence, China. His major research areas and work are related to computational intelligence, machine learning, bioinformatics, and among others. He published more than 150 papers in journals, conference proceedings, and several books in the above areas.

Xiao-Jun Wu received his BS degree in mathematics from Nanjing Normal University, Nanjing, China, in 1991, his MS degree in 1996, and his PhD in pattern recognition and intelligent systems in 2002, both from Nanjing University of Science and Technology, Nanjing, China. He joined Jiangnan University in 2006, where he is currently a professor. He has published more than 200 papers in his fields of research. He was a visiting researcher at the Centre for Vision, Speech, and Signal Processing, University of Surrey, UK, from 2003 to 2004. His current research interests include pattern recognition, computer vision, fuzzy systems, neural networks, and intelligent systems.

He-Feng Yin received his BS degree from the School of Computer Science and Technology from Xuchang University, Xuchang, China, in 2011. Currently, he is a PhD candidate at the School of IoT Engineering, Jiangnan University, Wuxi, China. His research interests include representation-based classification methods, dictionary learning, and low-rank representation.