IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received October 21, 2017, accepted December 11, 2017, date of publication December 27, 2017,
date of current version February 28, 2018.

Digital Object Identifier 10.1109/ACCESS.2017.2786672

Multicriteria-Based Active Discriminative Dictionary Learning for Scene Recognition

CAIXIA ZHENG[1,2,3], YUGEN YI[4], MIAO QI[1], FUCONG LIU[1], CHAO BI[5], JIANZHONG WANG[1], AND JUN KONG[1]

[1]College of Information Sciences and Technology, Northeast Normal University, Changchun 130117, China
[2]School of Environment, Northeast Normal University, Changchun 130117, China
[3]Key Laboratory for Applied Statistics of MOE, Northeast Normal University, Changchun 130024, China
[4]School of Software, Jiangxi Normal University, Nanchang 330022, China
[5]School of Psychology, Northeast Normal University, Changchun 130024, China

Corresponding authors: Jianzhong Wang (wangjz019@nenu.edu.cn) and Jun Kong (kongjun@nenu.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 61702092, Grant 61672150, Grant 61471111, Grant 61403078, and Grant 61602221, in part by the China Postdoctoral Science Foundation under Grant 2017M621193, in part by the Fund of the Jilin Provincial Science and Technology Department under Grant 20180520215JH, Grant 20180201089GX, Grant 20170204018GX, and Grant 20160204047GX, in part by the Natural Science Foundation of Jiangxi Province under Grant 20171BAB212009, and in part by the Fundamental Research Funds for the Central Universities under Grant 2412017QD029.

ABSTRACT Scene recognition is a significant and challenging problem in the field of computer vision. One of the principal bottlenecks in applying machine learning techniques to scene recognition tasks is the requirement of a large number of labeled training data. However, labeling massive training data manually (especially labeling images and videos) is very expensive in terms of human time and effort. In this paper, we present a novel multicriteria-based active discriminative dictionary learning (M-ADDL) algorithm to reduce the human annotation effort and create a robust scene recognition model. The M-ADDL algorithm possesses three advantages. First, M-ADDL introduces an active learning strategy into the discriminative dictionary learning model so that the performance of discriminative dictionary learning can be improved when the number of labeled samples is small. Second, different from most existing active learning methods that measure either the informativeness or representativeness of unlabeled samples to select useful samples for expanding the training dataset, M-ADDL employs both informativeness and representativeness to query useful unlabeled samples and utilizes the manifold-preserving ability of unlabeled samples as an additional sample selection criterion. Finally, a more effective representativeness criterion is presented based on the reconstruction coefficients of the samples. The experimental results of four standard scene recognition databases demonstrate the feasibility and validity of the proposed M-ADDL algorithm.

INDEX TERMS Active learning, dictionary learning, multicriteria of sample selection, scene recognition.

I. INTRODUCTION

Scene recognition is an important issue in the field of computer vision since it helps reduce the semantic gap of scene understanding between human beings and computers. Moreover, scene recognition also plays a key role in the success of many application areas such as human-machine interaction, image retrieval, and autonomous driving [1]. However, scene recognition is still a challenging problem because of the high variability of scale, illumination, viewpoint and layout of objects in the images.

Numerous algorithms have been proposed to classify images into semantic categories of scenes [2]–[4]. Lu *et al.* [1] adopted Gaussian Mixture models to produce the probability density response maps used for feature extraction, and they applied a bagged LDA classifier to recognize different categories of scene images. Choi *et al.* [5] proposed a hypergraph-based modeling method to extract the higher-order relationship of semantic attributes for scene recognition. Xie *et al.* [6] presented orientational pyramid matching to model the orientational context of scene images for indoor scene classification. Fei-Fei and Perona [7] proposed a Bayesian hierarchical model, which can learn the intermediate-level and the distribution of codewords without supervision, to recognize natural scene categories. Zhang *et al.* [3] proposed an object-to-class distance to model scene images, and they further adopted the distance representation for the final classification. Recently, many scene recognition methods based on the deep learning mechanism have been proposed to obtain better recognition accuracy [8]–[11]. Although these methods achieve good performance in scene recognition, they need massive datasets to train the classifier. However, collecting the vast number of datasets, especially the labeled data, is very time-consuming. For example, 111 researchers spent 220+ hours labeling only 63 hours of Trecvid 2003 development corpus [12]. The human vision system has a fascinating characteristic: we categorize images with only a few labeled training samples. Is it possible for a computer to achieve this result with machine learning techniques? This possibility is the motivation of this paper. To relieve the tedious work of labeling the training data and to build a competitive classifier with a limited amount of labeled training data, we develop a novel active learning algorithm based on discriminative dictionary learning for scene recognition.

Active learning based on pool setting has been a topic of recent interest. This type of active learning actively selects the most useful samples from a candidate unlabeled dataset (usually referred to as an active pool) and then asks humans to label the samples for training [13]. The overall procedure of active learning based on pool setting is shown in Fig. 1. In recent years, researchers have proposed various active learning algorithms and have applied them to visual concept recognition [13]–[15]. The key issue in active learning is how to decide whether a sample is ‘‘useful’’. There are two main sample selection criteria, namely, informativeness and representativeness [16], [17].

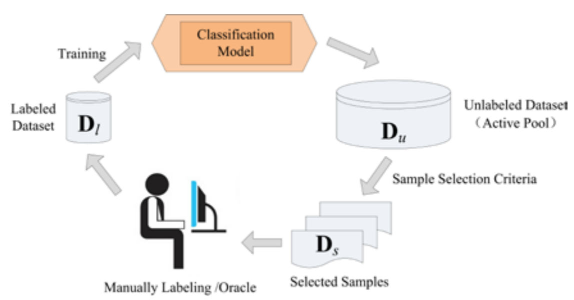


FIGURE 1. Schematic diagram of active learning based on pool setting.

The informativeness criterion is used to query the most informative samples that can reduce the uncertainty of the classifier. The most typical informativeness criteria include the following [14]: 1) query by committee, which chooses the instance for which several classifiers disagree the most [18], [19]; 2) large margin heuristic-based methods, which choose the most informative instance as the instance that is closest to the classification boundary [15], [20], [21]; and 3) posterior probability-based methods in which the posterior probability is used to measure the uncertainty of candidates [22], [23]. These approaches only focus on informative instances and usually do not consider the distribution information of unlabeled data, which may lead to serious sample bias and consequently, undesirable performance [24].

The representativeness criterion is used to query the most representative samples that can preserve the overall patterns or underlying distributions of the unlabeled dataset [17]. The representativeness criterion usually adopts clustering technologies [25], [26] to select samples from the highest density clusters. The approaches that deploy the representativeness criterion alone may require querying a relatively large amount of instances to label before converging to a good solution [24].

Most active learning algorithms only adopt one of the two criteria (informativeness or representativeness) for query selection [16], which can significantly deteriorate the performance of active learning because of the abovementioned drawbacks. Although several active learning algorithms [14], [16], [27], [28] have been proposed to query the unlabeled samples that are both informative and representative, they are usually heuristic in designing the specific query criterion based on the traditional classifiers, such as the support vector machine (SVM) [14], [16] and Gaussian Process classifiers(GP) [27], [28]. In addition, these algorithms do not consider the manifold-preserving ability of unlabeled samples when selecting useful unlabeled data. According to some researchers [29], [30], the manifold structure not only is important to characterize the intrinsic distribution of input data but also contributes to remove the outliers and noisy samples. Thus, adopting the manifold-preserving ability as a sample selection criterion is favorable.

In this paper, a novel multicriteria-based active discriminative dictionary learning (M-ADDL) algorithm is proposed for scene recognition. M-ADDL introduces an active learning mechanism into discriminative dictionary learning to improve the recognition performance of discriminative dictionary learning when the labeled samples are in smaller quantities. When querying the useful unlabeled samples from the dataset, it not only measures the informativeness and representativeness of unlabeled samples but also considers the manifold-preserving ability of unlabeled samples. Moreover, a more effective representativeness criterion is presented based on the reconstruction coefficients of the samples.

The rest of this paper is organized as follows. Section 2 briefly reviews some of the related works. Section 3 provides the details of the proposed algorithm. Many experiments and comparisons are conducted in Section 4, and Section 5 concludes the paper.

II. RELATED WORKS

A. DISCRIMINATIVE DICTIONARY LEARNING

Dictionaries play a crucial role in sparse coding or sparse representation-based image classification and reconstruction. Approaches to effectively learn dictionaries from training data have attracted significant attention in recent years. One representative dictionary learning method is k-singular value decomposition (KSVD) [31], which achieves satisfactory results in image restoration. However, KSVD is unsuitable for classification tasks because it learns a dictionary only by minimizing the residual error of reconstructing the original signals and neglects the class label of training data [32]. By exploring the class labels of training instances, supervised dictionary learning methods have been proposed to promote the discrimination capability of the learned dictionary. Supervised dictionary learning methods have achieved state-of-the-art performance in various tasks of pattern recognition [33]–[35].

Existing discriminative dictionary learning algorithms can be mainly divided into two categories [32]. The first type of algorithms learns a shared dictionary for all classes and enforces the representation coefficients to be discriminative [36], [37]. Jiang *et al.* [36] proposed a label-consistent KSVD (LC-KSVD) method by applying a binary class label sparse code matrix, which encourages samples from the same class to have similar sparse coding coefficients. Mairal *et al.* [37] proposed a task-driven dictionary learning (TDDL) method, which minimizes the different risk functions of the representation coefficients for different tasks. The second type of algorithm learns class-specific dictionaries and computes the class-specific representation residual for the classification [32], [38]. Yang *et al.* [32] developed a fisher discrimination dictionary learning (FDDL) framework in which both the representation residual and the representation coefficients are discriminative. Ramirez *et al.* [38] introduced an incoherence promoting term to force subdictionaries associated with different classes to be independent. However, most of the existing discriminative dictionary learning methods adopt l_0-norm or l_1-norm to regularize the representation coefficients, which often suffers heavy computational costs and makes both the training and testing phases inefficient. To avoid this problem, Gu *et al.* [39] proposed a projective dictionary pair learning (DPL) algorithm that learns a discriminative synthesis and analysis dictionary pair to reduce the time complexity in the training and testing phases.

The abovementioned discriminative dictionary learning methods achieve good performance for pattern recognition tasks, but all these methods need a massive number of labeled training samples to learn powerful dictionaries.

B. ACTIVE LEARNING

Active learning is an effective technique to reduce human labeling efforts in image and video annotation; it achieves better classification results when the number of labeled training instances is small [40]. Active learning iteratively selects the most useful instances to label in an interactive learning process. Thus, the redundant and unnecessary labeling of non-useful instances is avoided, which can significantly reduce the cost and time of manual annotations. Moreover, active learning can decrease the computational complexity of the training phase [14].

Informativeness and representativeness are two types of widely used sample selection criteria in active learning [41]. Since using either type of criterion alone is insufficient to achieve optimal performance, several researchers have attempted to query the unlabeled instances with both high informativeness and high representativeness [17]. Huang *et al.* [16] provided a systematic way to measure the informativeness and representativeness of an instance based on the min-max view of active learning. Freytag *et al.* [27], [28] proposed an active learning method based on the Gaussian process regression that automatically selects exploitative and explorative unlabeled examples for annotation. Donmez *et al.* [42] proposed a dynamic approach that combines the uncertainty and density information to query the unlabeled data; this approach adaptively updates the strategy selection parameters based on the estimated future residual error reduction. Wang *et al.* [14] and Wang and Ye [17] introduced an empirical risk minimization principle to active learning that employs the maximum mean discrepancy to measure the distribution difference and obtains an empirical upper bound for the active learning risk. Although these active learning methods combine informativeness and representativeness as sample selection criteria, they are restricted to a binary classification. Recently, several active learning algorithms for multi-class classification have been developed. Li *et al.* [43] proposed a serial active learning algorithm that first measures the informativeness of unlabeled samples based on the difference of probability and then queries representative samples from the selected informative sample set based on the expected error reduction. Li and Guo [13] presented an active learning approach that selects the representative and informative unlabeled samples based on mutual information and conditional entropy. Ebert *et al.* [24] analyzed different sampling criteria and formulated the criteria (informativeness and representativeness) selection as a Markov decision process based on reinforcement learning. Aodha *et al.* [45] proposed a hierarchical subquery evaluation for active learning on a graph that can balance exploration and exploitation to refine decision boundaries as needed within the time budget that is specified by the user. The abovementioned approaches have been proposed for multi-class classification problems. Nevertheless, because the manifold structure of the original data is ignored, the samples selected by these approaches may all be from a small region in the sample space, which decreases the generalization ability of the classifier [44], [46]. Furthermore, most of the instance selection criteria in active learning are designed based on traditional classifiers (e.g., SVM and GP). Very few studies have focused on developing an appropriate criterion for dictionary learning algorithms.

III. MULTICRITERIA-BASED ACTIVE DISCRIMINATIVE DICTIONARY LEARNING

This section presents the details of our proposed M-ADDL algorithm. In M-ADDL, DPL is employed as the classifier because DPL exhibits a highly competitive classification accuracy and a significantly higher efficiency. Other discriminative dictionary learning algorithms can also be used in our work. Specifically, the active learning mechanism proposed in this paper can be applied to most of the existing discriminative dictionary learning algorithms to improve their performance when the quantity of labeled data is insufficient. Fig. 2 shows a flow chart of M-ADDL. First, M-ADDL learns an initial dictionary by using DPL from the labeled training dataset. Second, M-ADDL selects samples from the unlabeled dataset to construct a subset that can preserve the manifold structure of the original unlabeled dataset. Finally, M-ADDL queries the highly informative and representative samples in the manifold-preserving subset to label to update the labeled training dataset and learns a refined dictionary based on the updated training dataset. This procedure of sample selection is iteratively performed to continuously improve the discriminative power of the learned dictionary.

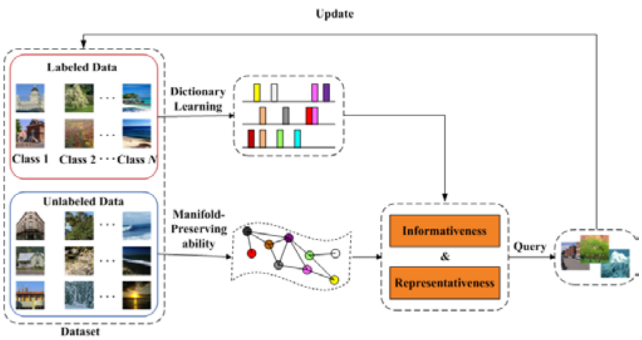


FIGURE 2. Flow chart of M-ADDL.

A. DPL ALGORITHM

The DPL algorithm [39] can learn discriminative dictionaries but without the costly l_0-norm or l_1-norm sparsity constraint on the representation coefficients. Therefore, DPL can achieve very competitive accuracies in visual recognition tasks, and it significantly reduces the time complexity in the training and testing phases. Because of the advantages of DPL, we used it as the classifier in this paper.

Denote $X = [X_1, \ldots, X_c, \ldots, X_C]$ as a dataset that includes m-dimensionality training samples from C classes, and denote $L = [L_1, \ldots, L_c, \ldots, L_C]$ as the corresponding label set, where $X_c \in R^{m \times n_r}$ represents samples in the c-th class, and L_c denotes the label of the c-th class. DPL simultaneously learns an analysis dictionary P and a synthesis dictionary D by using the following objective function:

$$\{\hat{P}, \hat{D}\} = \arg\min_{P,D} \sum_{c=1}^{C} \|X_c - D_c P_c X_c\|_F^2 + \lambda \left\|P_c \overline{X}_c\right\|_F^2, \quad s.t.\ \|d_i\|_2^2 \le 1, \tag{1}$$

where $P = [P_1; \ldots; P_c; \ldots; P_C] \in R^{kC \times m}$ is used for linear encoding the representation coefficients, $D = [D_1, \ldots, D_c, \ldots, D_C] \in R^{m \times kC}$ is used for the class-specific discriminative reconstructing samples, $D_c \in R^{m \times k}$ and $P_c \in R^{k \times m}$ are the sub-dictionary pairs that are learned from class c, $\overline{X}_c$ represents the complementary data matrix of X_c in the entire training set X, $\lambda \succ 0$ is the scalar constant to control the discriminative property of dictionary P, and d_i represents the i-th atom of dictionary D.

The objective function in (1) is generally non-convex. By introducing a variable matrix A, equation (1) can be relaxed to the following form:

$$\{\hat{P}, \hat{A}, \hat{D}\} = \arg\min_{P,A,D} \sum_{c=1}^{C} \|X_c - D_c A_c\|_F^2 + \tau \|P_c X_c - A_c\|_F^2 + \lambda \left\|P_c \overline{X}_c\right\|_F^2, \quad s.t.\ \|d_i\|_2^2 \le 1, \tag{2}$$

where τ is the scalar constant. The objective function in (2) can be solved using an alternatively updated manner; see [19] for details.

When the optimal D and P have been learned, the class-specific reconstruction residual is used to estimate the class label of the test sample x_t, as shown in the following formula:

$$label(x_t) = \arg\min_{c} \|x_t - D_c P_c x_t\|_2 . \tag{3}$$

In general, DPL requires a massive number of labeled training data to learn the discriminative dictionary pair to obtain good classification results. However, it is difficult and expensive to obtain a vast quantity of labeled training data. If we can fully utilize the information provided by the inexpensive unlabeled data, we would can learn a more discriminative dictionary than the dictionary learned by using only a limited number of labeled training data. To achieve this, we introduce an active learning technique to DPL and propose multicriteria of sample selection to actively query the beneficial unlabeled samples from the unlabeled dataset to improve DPL's classification performance.

B. MULTICRITERIA OF SAMPLE SELECTION

The criterion of sample selection is crucial for active learning methods. Different active learning methods have different strategies in identifying which sample should be queried for the current classifier. The proposed M-ADDL develops multicriteria to select useful samples, which has the three key components of a manifold-preserving ability criterion, informativeness criterion and representativeness criterion. We introduce each criterion below.

1) MANIFOLD-PRESERVING ABILITY CRITERION

In machine learning, the concept of manifold-preserving means the selection of a small number of samples to represent the original manifold structure of a dataset. Specifically, the samples with high space connectivity should be selected to construct a manifold-preserving subset (MPS); thus, the outliers and noisy samples can be removed. An MPS is beneficial to many machine learning tasks, for example, recognition problems [29], [30]. Therefore, we introduce the concept of manifold-preserving into active learning, which can avoid oversampling on dense regions to a large extent.

Sun *et al.* [30] proposed the technique of manifold-preserving graph reduction for sparse semi-supervised learning. Inspired by their work, we apply manifold-preserving graph reduction to construct an MPS to preserve the manifold structure of the unlabeled dataset, and then we query highly informative and representative samples from the MPS. Therefore, the samples queried in active learning are more beneficial for improving the generalization ability of the classifier.

When constructing the MPS, it must be ensured that the samples in the MPS have highly similar properties and labels to the samples outside of the MPS. Therefore, the classifier learned from an MPS can generalize well to unseen samples with a high probability. The graph $G(V, E, W)$ comprises unlabeled samples, where V represents the vertex set, E represents the edge set, and W represents the symmetric weight matrix. If sample x_i(i-th vertex) and sample x_j(j-th vertex) are k-neighbors, the weight $w_{i,j}$ is computed using the following Gaussian kernel function or otherwise, $w_{i,j} = 0$:

$$w_{ij} = \exp \frac{-\left\|x_i - x_j\right\|^2}{2\sigma}, \tag{4}$$

where σ is the parameter.

The connectivity of samples in an MPS to the samples outside the MPS is defined as follows:

$$\frac{1}{N-p}\sum_{i=p+1}^{N}\left(\max_{j=1,\ldots,p} w_{ij}\right), \tag{5}$$

where N is the number of all vertices, and p is the number of candidate vertices to construct the MPS. To construct an optimal MPS, equation (5) should be maximized. The problem of exactly seeking MPS graphs by using (5) is NP-hard [30]. Therefore, the connectivity degree $d(i)$ of sample x_i in graph G is defined in the following form to replace (5):

$$d(i) = \sum\nolimits_{i \sim j} w_{ij}. \tag{6}$$

A larger $d(i)$ indicates that sample x_i has a higher connectivity to other samples and has more useful information [30], [46]. Therefore, it is more likely to be added into the MPS.

2) INFORMATIVENESS CRITERION

The informativeness criterion is used to select informative samples to reduce the classification uncertainty of the classifier. The samples, which are not well-reconstructed by using the current learned dictionary, are likely to provide more information in further refining the dictionary. In M-ADDL, we employ the reconstruction error of sample x_j to compute its informativeness M_{infor}.

$$M_{infor} = \min_{c} \left\|x_j - D_c P_c x_j\right\|^2, \tag{7}$$

where D_c and P_c represent the sub-dictionary pairs learned by the DPL algorithm for class c. A larger M_{infor} means that the current learned dictionaries do not reconstruct sample x_j well and should thus be queried in active learning to refine the learned dictionaries.

3) REPRESENTATIVENESS CRITERION

The representativeness criterion is used to evaluate the relations between the selected samples and the remaining samples in the unlabeled dataset and aims to query highly representative samples to label for expanding the training dataset. Most of the existing active learning methods have developed the representativeness measure based on the probability distribution of the dataset [13] or have selected the samples near the center of clustering as highly representative samples [25]. The representativeness criteria in these methods are designed based on the Euclidean space. In this paper, a more appropriate representativeness criterion for the discriminative dictionary learning is proposed. Since discriminative dictionary learning adopts a class-specific reconstruction residual to assign the class label of samples, we propose a novel representativeness criterion based on the reconstruction coefficients of the samples. Given one of the unlabeled samples, it can be approximated using a linear combination of other samples in an unlabeled dataset [47]. The highly representative samples are defined as the samples with larger reconstruction coefficients. Given a dataset $X = \{x_1, \ldots, x_i, \ldots, x_N\}$, the reconstruction coefficients $B = \{b_1, \ldots, b_i, \ldots, b_N\}$ can be derived by solving the following object function:

$$\min \|X - XB\|_2 + \alpha \|B\|_{2,1}, \tag{8}$$

where $\|X - XB\|_2$ is the reconstruction error, $\|B\|_{2,1} = \sum_{i=1}^{N} \|b_i\|_2$ is the regularization term that enforces the group sparsity on variable B, and α is the parameter. When the value of α is larger, B has more zero rows.

Inspired by Wright *et al.* [48], the representativeness measure that is based on the reconstruction coefficients is defined as

$$M_{rep}(x_i) = -\frac{\max(b_i)}{||b_i||_1}. \tag{9}$$

The large $M_{rep}(x_i)$ indicates that the sparse coefficient b_i is spread evenly over all samples, which demonstrates that x_i is a representative sample.

C. IMPLEMENTATION SCHEME of M-ADDL

Based on the proposed manifold-preserving ability, informativeness and representativeness criteria, we provide the implementation scheme of M-ADDL as shown in Algorithm 1.

IV. EXPERIMENTS

In this section, we evaluate the proposed M-ADDL algorithm on different datasets and compare it with other algorithms.

A. DATASETS AND EXPERIMENTAL SETUP

1) Datasets

The proposed M-ADDL is evaluated on three small datasets that include an 8-Scene dataset [48], a UIUC-Sports dataset [49] and a 15-Scene dataset [50], which are the most frequently used scene recognition datasets in the literature thus far. The 8-Scene dataset contains 2,688 outdoor scene

Algorithm 1 Multicriteria-Based Active Discriminative Dictionary Learning (M-ADDL)

1. Inputs: Labeled training dataset X^l with a label set L^l, unlabeled dataset X^u, the number (I_t) of iteration in active learning, the total number (N_u) of unlabeled samples in each iteration, the number (N_s) of unlabeled samples to be queried for expanding the training dataset in each iteration, and the number (p) of samples selected for constructing the MPS (X^{mps});
2. Initialization: Learn an initial dictionary pair D and P through the DPL algorithm from X^l;
3. For $t = 1$ to I_t, do
4. For $j = 1$ to p, do
5. Compute the degree $d(i), i = 1, 2, \ldots, N_u - j + 1$ through (6);
6. Remove the samples in X^u with a larger $d(i)$ to X^{mps};
7. End for
8. Compute M_{infor} and M_{rep} for each sample in X^{mps} through (7) and (9);
9. Query N_s samples (noted by X^s) with a higher sum of M_{infor} and M_{rep} from X^{mps} to label and then add them to the training dataset X^l, update $X^u = X^u - X^s$ and $X^l = X^l \cup X^s$;
10. Learn the refined dictionary pair D^t_{new} and P^t_{new} from the updated dataset X^l;
11. End for
12. Outputs: Final learned dictionary pair $D^{I_t}_{new}$ and $P^{I_t}_{new}$.

images across 8 categories, and the size of each image is 256 × 256. The UIUC-Sports dataset consists of 1,579 images that are labeled into 8 complex sport scene categories, and the resolution of the images is from 800 × 600 to thousands of pixels per dimension. The 15-Scene dataset includes 4,485 gray scene images with 15 categories, and the average resolution of the images is 300 × 250. In addition, we evaluate the proposed M-ADDL on a larger scale dataset MIT-Indoor [51]. The MIT-Indoor dataset contains 15,620 images of 67 indoor scene classes, and all images have a minimum resolution of 200 pixels in the smallest axis. Fig. 3 shows example images from the different datasets. These datasets are very challenging for the scene recognition task because 1) the backgrounds of the images in the same category of a scene are highly diverse and 2) within the same class, the sizes, appearances and numbers of objects are different.

2) EXPERIMENTAL SETUP

Three types of descriptors that include GIST [48], PHOW [52] and bag-of-words [53] of LBP [54] are extracted as the feature representation of the images. In each dataset, the images are randomly split into a labeled training set, unlabeled set and testing set according to proportions of 10%, 60% and 30%, respectively. Random splitting is repeated 10 times, and the average accuracy and standard deviation are reported.

FIGURE 3. Example images from different datasets. (a) 8-Scene[48], (b) UIUC- Sports[49], (c) 15-Scene[50], (d) MIT-Indoor[51].

The proposed M-ADDL is compared with several state-of-the-art algorithms including: active learning by querying informative and representative examples (QURIE_SVM for short) [55], active learning for large multi-class problems (PKNN for short) [56] and active discriminative dictionary learning for weather recognition (ADDL for short) [57]. These algorithms are the general recognition algorithms that are mainly focused on developing an effective active learning mechanism. QURIE_SVM [55] is proposed based on the min-max view of active learning, which measures the informativeness and representativeness of a sample by using its prediction uncertainty. PKNN [56] learned an accurate kernel function over the input space and provided a natural notion of uncertainty over class labels for active learning. ADDL [57] actively queried the informative and representative samples based on the entropy of the probability distribution over the class-specific reconstruction error and the distribution of the unlabeled dataset. To verify that the proposed M-ADDL can effectively improve the performance of the original DPL algorithm when the labeled training samples are limited,

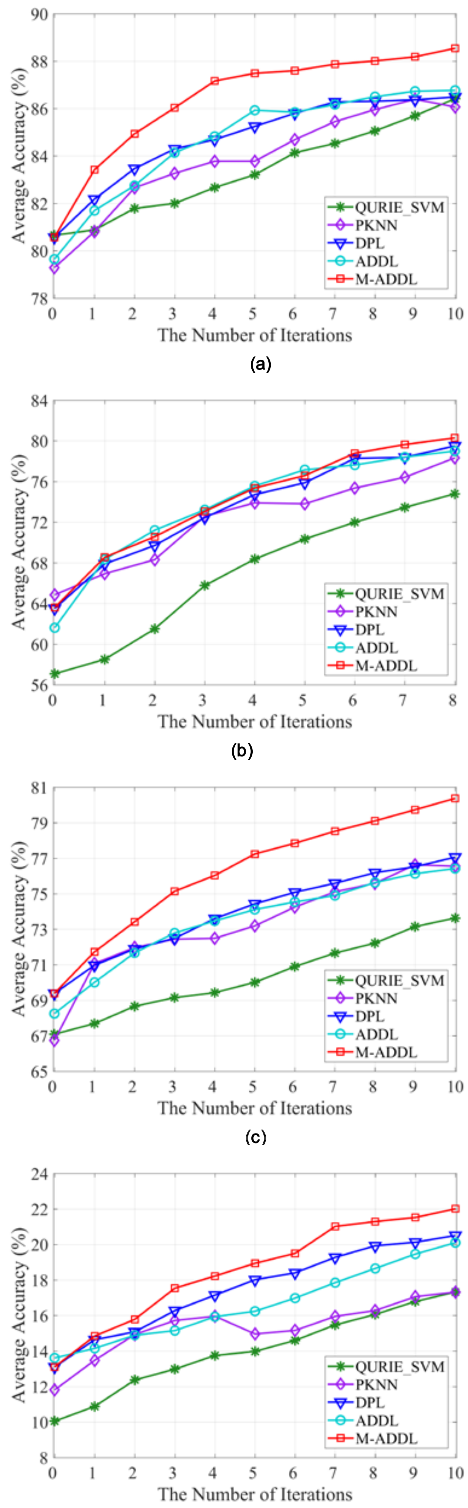


FIGURE 4. Comparison of the average recognition accuracy obtained by various methods on four different datasets. (a) 8-Scene. (b) UIUC-Sports. (c) 15-Scene. (d) MIT-Indoor.

we also compare M-ADDL with the original DPL algorithm, which randomly selects samples from the unlabeled dataset to expand the training dataset. Each algorithm starts with the labeled samples and iteratively selects samples from the unlabeled set to label, namely, 100 samples for the 8-Scene, UIUC-Sports and 15-Scene datasets and 200 samples for the MIT-Indoor dataset in each iteration. The maximum number of iteration is set to 10 for the 8-Scene, 15-Scene and MIT-Indoor dataset, while it is set to 8 for the UIUC-Sports dataset because the number of samples in this dataset is small.

B. EXPERIMENTAL RESULTS AND ANALYSIS

In this section, the experimental results are presented to demonstrate the effectiveness of the proposed M-ADDL framework in scene recognition problems. Fig. 4 shows the average recognition accuracy obtained using various methods on four different datasets. From Fig. 4, the following points can be observed. First, the performances of the methods that are based on discriminative dictionary learning (M-ADDL, ADDL and DPL) are better than the methods based on SVM (QURIE_SVM) and KNN (PKNN) in most cases. The reason for this phenomenon may be that the discriminative dictionary learning method is more effective for scene recognition tasks. Second, the proposed M-ADDL outperforms DPL, which demonstrates that the introduction of an active learning mechanism into DPL can effectively improve the performance of DPL. Finally, the proposed M-ADDL is generally superior to other active learning methods because M-ADDL measures not only the informativeness and representativeness of unlabeled samples but also the manifold-preserving ability of unlabeled samples. Therefore, M-ADDL can actively query more useful unlabeled samples to label to further improve the classifier performance.

From Fig. 4, we also note that on the UIUC-Sports dataset, the performance of M-ADDL is inferior to the performance of ADDL in the first couple of iterations. The reason to this phenomenon may due to that the number of training samples in the previous iterations is small in this dataset, so the manifold structure cannot be well captured by the MPS.

In order to demonstrate that the proposed M-ADDL can achieve better recognition performance than the discriminative dictionary learning algorithms when the labeled training data are in smaller quantities, we also compare M-ADDL with FDDL [32] algorithm. Since FDDL is a supervised algorithm which cannot use the information of the unlabeled training data, we only utilize the labeled samples for FDDL training in this experiment. Table 1 lists the scene recognition results of M-ADDL and FDDL. From Table 1, it can be seen that M-ADDL can gain higher accuracies by integrating active learning into discriminative dictionary learning.

TABLE 1. Average recognition rates (%) and standard deviations (%) of M-ADDL and FDDL.

Dataset / Method	8-Scene	UIUC-Sports	15-Scene Categories	MIT-Indoor
FDDL	80.42±1.59	64.68±1.83	67.71±1.53	12.11±0.93
M-ADDL	88.53±1.24	80.30±1.95	80.37±0.92	22.01±0.67

TABLE 2. Average recognition rates (%) and standard deviations (%) of M-ADDL with different k values.

Dataset / k	8-Scene	UIUC-Sports	15-Scene Categories	MIT-Indoor
25	**88.53±1.24**	78.14±1.39	80.12±0.77	**22.01±0.67**
50	87.24±1.12	78.98±1.47	79.92±1.02	21.74±0.77
75	87.18±1.23	**80.30±1.95**	79.82±1.05	21.55±0.86
100	87.40±1.03	78.31±0.93	**80.37±0.92**	21.72±0.86
125	87.25±1.12	78.35±1.03	79.84±1.09	21.69±0.66
150	86.94±0.95	77.82±1.14	80.15±0.79	21.85±0.64
175	87.13±1.27	77.52±0.94	80.13±0.86	21.69±0.73

TABLE 3. Average recognition rates (%) and standard deviations (%) of M-ADDL with different λ values.

Dataset / λ	8-Scene	UIUC-Sports	15-Scene Categories	MIT-Indoor
0.0005	85.09±1.10	77.67±0.58	68.32±1.36	18.37±1.07
0.001	85.93±0.99	78.20±1.25	69.58±1.51	18.19±0.54
0.005	86.43±0.72	77.48±0.84	72.78±1.07	17.73±0.71
0.01	86.62±1.16	76.78±0.99	75.13±0.64	18.01±0.67
0.05	88.18±1.23	76.93±1.85	79.28±0.89	18.37±0.84
0.1	**88.53±1.24**	77.31±1.66	**80.37±0.92**	19.02±0.73
0.5	88.10±1.24	79.22±1.68	80.01±1.14	**22.01±0.67**
1	87.08±1.40	**80.30±1.95**	79.22±1.00	21.10±0.91
10	83.77±1.49	79.17±2.03	73.10±1.28	17.33±0.98

TABLE 4. Average recognition rates (%) and standard deviations (%) of M-ADDL with different τ values.

Dataset / τ	8-Scene	UIUC-Sports	15-Scene Categories	MIT-Indoor
0.005	85.29±1.23	78.94±0.91	76.83±1.09	19.24±0.73
0.05	86.69±1.16	68.73±2.33	79.08±1.05	19.33±0.92
0.5	88.05±1.27	**80.30±1.95**	**80.37±0.92**	21.90±0.84
1	**88.53±1.24**	78.94±0.91	80.03±1.11	**22.01±0.67**
5	87.80±1.10	78.92±2.44	79.54±0.81	21.21±0.79
15	87.25±1.14	79.39±1.10	78.72±0.64	19.39±0.70
35	87.31±1.12	79.15±1.03	78.09±0.97	18.55±0.65
55	87.06±1.10	77.58±0.74	78.15±0.63	17.80±0.92
75	86.67±1.13	79.11±0.91	78.04±0.54	17.42±0.93

TABLE 5. Average recognition rates (%) and standard deviations (%) of M-ADDL with different α values.

Dataset / α	8-Scene	UIUC-Sports	15-Scene Categories	MIT-Indoor
0.0005	87.95±1.27	79.85±1.71	80.28±0.76	21.14±0.55
0.001	87.92±1.38	80.04±1.44	80.15±0.82	21.67±0.57
0.005	88.00±1.32	80.17±1.39	80.10±0.98	21.63±0.59
0.01	**88.53±1.24**	**80.30±1.95**	80.13±1.22	21.41±0.61
0.05	87.84±1.22	79.96±1.73	80.03±0.95	**22.01±0.67**
0.1	87.75±1.43	79.72±1.64	80.24±1.13	21.74±0.64
0.5	87.92±1.38	79.49±1.50	**80.37±0.92**	21.60±0.97
1	87.25±1.27	79.66±1.49	79.93±1.07	20.81±0.81
10	87.35±1.52	80.11±1.65	80.10±0.94	18.83±0.62

Then, the performance of our algorithm under different parameters was compared. There are three important parameters, k, λ, and τ, in (2) and α in (8), where k is the size of each sub-dictionary D_c, λ is the parameter that controls the discriminative property of dictionary P, α is the scalar constant in a DPL algorithm, and α is the parameter that controls the sparsity of the reconstruction coefficients B. As shown by the experimental results in Table 2, the highest average classification accuracy is obtained when $k = 25$, 75 or 100. This finding indicates that the proposed M-ADDL has a better performance when learning a relatively compact dictionary, which is highly beneficial for reducing the time during the testing phase. Tables 3 and 4 show that the values of parameters λ and τ have an important effect on the performance of the proposed M-ADDL. This importance is because a λ value that is too large will cause the reconstruction coefficient in M-ADDL to be too sparse. However, if the λ value is too small, this will lead to the reconstruction coefficient to be too dense. A very sparse or dense reconstruction coefficient will deteriorate the classification performance of M-ADDL. Likewise, if τ is too large, the influence of the reconstruction error constraint (the first term in (4)) and the sparse constraint (the third term in (4)) is weakened, which will decrease the discrimination ability of the learned dictionary. In contrast, if τ is too small, the second term in (4) will be ignored during dictionary learning, which will also reduce the performance of M-ADDL. Table 5 shows the classification results under different values of α. This table indicates that the value of α should not be too large or too small because it leads to the reconstruction coefficients to be too sparse or too dense, respectively, and this will degrade the performance of M-ADDL.

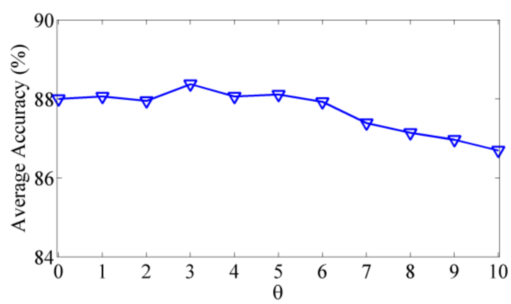


FIGURE 5. Average recognition rates of M-ADDL with different values of θ.

To identify the importance of the informativeness and representativeness criteria for discriminative dictionary learning-based active learning, we use the 8-Scene dataset as an example to analyze different combinations of the informativeness and representativeness criteria. The combination formula is $M_{sum} = \theta M_{infor} + (1 - \theta) M_{rep}$, and the unlabeled samples with a large M_{sum} are queried to expand the training dataset. Fig. 5 shows the average recognition rate under different values of θ. From Fig. 5, we can make the following observations. First, using the informativeness criterion alone works better than using the representativeness criterion alone in M-ADDL because the informativeness criterion is directly computed based on the recognition results of the learned dictionary. Second, combining the informativeness and representativeness criteria in M-ADDL with a proper weight can obtain the best recognition results, which indicates that both informativeness and representativeness are necessary in our proposed algorithm.

V. CONCLUSION

This paper presents an active learning algorithm (M-ADDL) for discriminative dictionary learning-based scene recognition. Because M-ADDL adopts multicriteria (informativeness, representativeness and a manifold-preserving ability) to query unlabeled samples to expand the training dataset, it achieves better performance than other algorithms. In the experiments, four public scene databases are utilized to evaluate M-ADDL. By comparing the performance of M-ADDL with other state-of-the-art algorithms, the effectiveness and advantages of M-ADDL are demonstrated.

REFERENCES

[1] X. Lu, X. Li, and L. Mou, "Semi-supervised multitask learning for scene recognition," *IEEE Trans. Cybern.*, vol. 45, no. 9, pp. 1967–1976, Sep. 2015.

[2] L. Zhang, R. Ji, Y. Xia, Y. Zhang, and X. Li, "Learning a probabilistic topology discovering model for scene categorization," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 26, no. 8, pp. 1622–1634, Aug. 2015.

[3] L. Zhang, X. Zhen, and L. Shao, "Learning object-to-class kernels for scene classification," *IEEE Trans. Image Process.*, vol. 23, no. 8, pp. 3241–3253, Aug. 2014.

[4] M. Dixit, S. Chen, D. Gao, N. Rasiwasia, and N. Vasconcelos, "Scene classification with semantic Fisher vectors," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2015, pp. 2974–2983.

[5] S.-W. Choi, C. H. Lee, and I. K. Park, "Scene classification via hypergraph-based semantic attributes subnetworks identification," in *Proc. Eur. Conf. Comput. Vis.*, 2014, pp. 361–376.

[6] L. Xie, J. Wang, B. Guo, B. Zhang, and Q. Tian, "Orientational pyramid matching for recognizing indoor scenes," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2014, pp. 3734–3741.

[7] L. Fei-Fei and P. Perona, "A Bayesian hierarchical model for learning natural scene categories," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2005, pp. 524–531.

[8] B. Zhou, A. Lapedriza, J. Xiao, A. Torralba, and A. Oliva, "Learning deep features for scene recognition using places database," in *Proc. Adv. Neural Inf. Process. Syst.*, 2014, pp. 487–495.

[9] Y. Yuan, L. Mou, and X. Lu, "Scene recognition by manifold regularized deep learning architecture," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 26, no. 10, pp. 2222–2233, Oct. 2015.

[10] L. Herranz, S. Jiang, and X. Li, "Scene recognition with CNNs: Objects, scales and dataset bias," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2016, pp. 571–579.

[11] K. Nogueira, O. A. B. Penatti, and J. A. dos Santos, "Towards better exploiting convolutional neural networks for remote sensing scene classification," *Pattern Recognit.*, vol. 61, pp. 539–556, Jan. 2017.

[12] C.-Y. Lin, B. L. Tseng, and J. R. Smith, "Video collaborative annotation forum: Establishing ground-truth labels on large multimedia datasets," in *Proc. TRECVID Workshop*, 2003, pp. 1–19.

[13] X. Li and Y. Guo, "Adaptive active learning for image classification," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2013, pp. 859–866.

[14] Z. Wang, B. Du, L. Zhang, and L. Zhang, "A batch-mode active learning framework by querying discriminative and representative samples for hyperspectral image classification," *Neurocomputing*, vol. 179, pp. 88–100, Feb. 2016.

[15] A. Kapoor, K. Grauman, R. Urtasun, and T. Darrell, "Active learning with Gaussian processes for object categorization," in *Proc. IEEE 11th Int. Conf. Comput. Vis. (ICCV)*, Oct. 2007, pp. 1–8.

[16] S.-J. Huang, R. Jin, and Z.-H. Zhou, "Active learning by querying informative and representative examples," in *Proc. Adv. Neural Inf. Process. Syst.*, 2010, pp. 892–900.

[17] Z. Wang and J. Ye, "Querying discriminative and representative samples for batch mode active learning," *ACM Trans. Knowl. Discovery Data*, vol. 9, no. 3, 2015, Art. no. 17.

[18] Y. Freund, H. S. Seung, E. Shamir, and N. Tishby, "Selective sampling using the query by committee algorithm," *Mach. Learn.*, vol. 28, no. 2, pp. 133–168, 1997.

[19] S. Argamon-Engelson and I. Dagan, "Committee-based sample selection for probabilistic classifiers," *J. Artif. Intell. Res.*, vol. 11, pp. 335–360, Nov. 1999.

[20] A. Freytag, E. Rodner, P. Bodesheim, and J. Denzler, "Rapid uncertainty computation with gaussian processes and histogram intersection kernels," in *Proc. Asian Conf. Comput. Vis.*, 2012, pp. 511–524.

[21] X. Li and Y. Guo, "Active learning with multi-label SVM classification," in *Proc. IJCAI*, 2013, pp. 1479–1485.

[22] T. Luo *et al.*, "Active learning to recognize multiple types of plankton," *J. Mach. Learn. Res.*, vol. 6, pp. 589–613, Apr. 2005.

[23] N. Roy and A. McCallum, "Toward optimal active learning through sampling estimation of error reduction," in *Proc. ICML*, Williamstown, MA, USA, 2001, pp. 441–448.

[24] S. Ebert, M. Fritz, and B. Schiele, "RALF: A reinforced active learning formulation for object class recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2012, pp. 3626–3633.

[25] H. T. Nguyen and A. Smeulders, "Active learning using pre-clustering," in *Proc. 21st Int. Conf. Mach. Learn.*, 2004, p. 79.

[26] S. Dasgupta and D. Hsu, "Hierarchical sampling for active learning," in *Proc. 25th Int. Conf. Mach. Learn.*, 2008, pp. 208–215.

[27] A. Freytag, E. Rodner, P. Bodesheim, and J. Denzler, "Labeling examples that matter: Relevance-based active learning with Gaussian processes," in *Proc. German Conf. Pattern Recognit.*, 2013, pp. 282–291.

[28] A. Freytag, E. Rodner, and J. Denzler, "Selecting influential examples: Active learning with expected model output changes," in *Proc. Eur. Conf. Comput. Vis.*, 2014, pp. 562–577.

[29] J. Zhu and S. Sun, "Sparse Gaussian processes with manifold-preserving graph reduction," *Neurocomputing*, vol. 138, pp. 99–105, Aug. 2014.

[30] S. Sun, Z. Hussain, and J. Shawe-Taylor, "Manifold-preserving graph reduction for sparse semi-supervised learning," *Neurocomputing*, vol. 124, pp. 13–21, Jan. 2014.

[31] M. Aharon, M. Elad, and A. Bruckstein, "K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation," *IEEE Trans. Signal Process.*, vol. 54, no. 11, pp. 4311–4322, Nov. 2006.

[32] M. Yang, L. Zhang, X. Feng, and D. Zhang, "Sparse representation based Fisher discrimination dictionary learning for image classification," *Int. J. Comput. Vis.*, vol. 109, no. 3, pp. 209–232, Sep. 2014.

[33] J. Dong, C. Sun, and W. Yang, "A supervised dictionary learning and discriminative weighting model for action recognition," *Neurocomputing*, vol. 158, pp. 246–256, Jun. 2015.

[34] A. Moeini, K. Faez, H. Moeini, and A. M. Safai, "Facial expression recognition using dual dictionary learning," *J. Vis. Commun. Image Represent.*, vol. 45, pp. 20–33, May 2017.

[35] X. Wu, Q. Li, L. Xu, K. Chen, and L. Yao, "Multi-feature kernel discriminant dictionary learning for face recognition," *Pattern Recognit.*, vol. 66, pp. 404–411, Jun. 2017.

[36] Z. Jiang, Z. Lin, and L. S. Davis, "Label consistent K-SVD: Learning a discriminative dictionary for recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 11, pp. 2651–2664, Nov. 2013.

[37] J. Mairal, F. Bach, and J. Ponce, "Task-driven dictionary learning," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 4, pp. 791–804, Apr. 2012.

[38] I. Ramirez, P. Sprechmann, and G. Sapiro, "Classification and clustering via dictionary learning with structured incoherence and shared features," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2010, pp. 3501–3508.

[39] S. Gu, L. Zhang, W. Zuo, and X. Feng, "Projective dictionary pair learning for pattern classification," in *Proc. Adv. Neural Inf. Process. Syst.*, 2014, pp. 793–801.

[40] X. Li and Y. Guo, "Multi-level adaptive active learning for scene classification," in *Proc. Eur. Conf. Comput. Vis.*, 2014, pp. 234–249.

[41] B. Settles, "Active learning literature survey," Dept. Comput. Sci., Univ. Wisconsin-Madison, Madison, WI, USA, Tech. Rep. 1648, 2010, p. 11, vol. 52.

[42] P. Donmez, J. G. Carbonell, and P. N. Bennett, "Dual strategy active learning," in *Proc. Eur. Conf. Mach. Learn.*, 2007, pp. 116–127.

[43] C. Li, P. Zhao, J. Wu, H. Xu, and Z. Cui, "A serial sample selection framework for active learning," in *Proc. Int. Conf. Adv. Data Mining Appl.*, 2014, pp. 435–446.

[44] L. Qiao, L. Zhang, and S. Chen, "Dimensionality reduction with adaptive graph," *Frontiers Comput. Sci.*, vol. 7, no. 5, pp. 745–753, 2013.

[45] O. M. Aodha, N. D. Campbell, J. Kautz, and G. J. Brostow, "Hierarchical subquery evaluation for active learning on a graph," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2014, pp. 564–571.

[46] J. Zhou and S. Sun, "Gaussian process versus margin sampling active learning," *Neurocomputing*, vol. 167, pp. 122–131, Nov. 2015.

[47] S. T. Roweis and L. K. Saul, "Nonlinear dimensionality reduction by locally linear embedding," *Science*, vol. 290, no. 5500, pp. 2323–2326, Dec. 2000.

[48] J. Wright, A. Y. Yang, A. Ganesh, S. S. Sastry, and Y. Ma, "Robust face recognition via sparse representation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 31, no. 2, pp. 210–227, Feb. 2009.

[49] L.-J. Li and L. Fei-Fei, "What, where and who? Classifying events by scene and object recognition," in *Proc. IEEE 11th Int. Conf. Comput. Vis. (ICCV)*, Oct. 2007, pp. 1–8.

[50] S. Lazebnik, C. Schmid, and J. Ponce, "Beyond bags of features: Spatial pyramid matching for recognizing natural scene categories," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2006, pp. 2169–2178.

[51] A. Quattoni and A. Torralba, "Recognizing indoor scenes," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2009, pp. 413–420.

[52] A. Bosch, A. Zisserman, and X. Munoz, "Image classification using random forests and ferns," in *Proc. IEEE 11th Int. Conf. Comput. Vis. (ICCV)*, Oct. 2007, pp. 1–8.

[53] G. Csurka, C. Dance, L. Fan, J. Willamowski, and C. Bray, "Visual categorization with bags of keypoints," in *Proc. Workshop Statist. Learn. Comput. Vis. (ECCV)*, 2004, pp. 1–2.

[54] T. Ojala, M. Pietikäinen, and D. Harwood, "A comparative study of texture measures with classification based on featured distributions," *Pattern Recognit.*, vol. 29, no. 1, pp. 51–59, 1996.

[55] S.-J. Huang, R. Jin, and Z.-H. Zhou, "Active learning by querying informative and representative examples," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 36, no. 10, pp. 1936–1949, Oct. 2014.

[56] P. Jain and A. Kapoor, "Active learning for large multi-class problems," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2009, pp. 762–769.

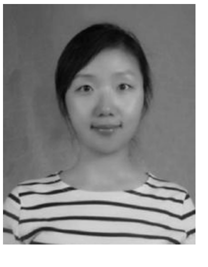
[57] C. Zheng, F. Zhang, H. Hou, C. Bi, M. Zhang, and B. Zhang, "Active discriminative dictionary learning for weather recognition," *Math. Problems Eng.*, vol. 2016, Apr. 2016, Art. no. 8272859.

CAIXIA ZHENG was born in Gongzhuling, China. She received the B.S. and M.S. degrees from the College of Computer Science and Information Technology, Northeast Normal University, China, in 2009 and 2012, respectively, and the Ph.D. degree from the School of Mathematic and Statistic, Northeast Normal University in 2016. She holds a post-doctoral position with the College of Information Science and Technology, Northeast Normal University. Her research interests include pattern recognition, machine learning, and image processing. From 2013 to 2014, she studied as a Visiting Ph.D. Student with the University of California, Davis, CA, USA.

YUGEN YI was born in Pingxiang, China. He received the B.S. degree from the College of Humanities and Sciences, Northeast Normal University, China, in 2009, the M.S. degree from the College of Computer Science and Information Technology, Northeast Normal University, in 2012, and the Ph.D. degree from the School of Mathematics and Statistics, Northeast Normal University, in 2015. He is currently a Lecturer with the School of Software, Jiangxi Normal University. His research interests include artificial intelligence, computer vision, and machine learning.

MIAO QI was born in Huludao, China. She received the B.S. degree from Anshan Normal University, China, in 2004, the M.S. degree from the College of Computer Science and information Technology, Northeast Normal University, China, in 2007, and the Ph.D. degree from the College of Chemistry, Northeast Normal University, in 2010.She is currently an Associate Professor with the College of Information Science and Technology, Northeast Normal University. Her research interests include feature extraction and pattern recognition.

FUCONG LIU was born in Changchun, China, in 1988. He received the B.S. degree from the College of Measurement and Control Technology and Instrument, Changchun University of Science and Technology, China, in 2011, and the M.S. degree from the School of mathematics, Jilin University, China, in 2015. He is currently pursuing the Ph.D. degree with the College of Information Science and Technology, Northeast Normal University, China. His research interests include dimensionality reduction and feature extraction.



CHAO BI was born in Harbin, China. She received the B.S. degree from the College of Computer Science and information Technology, Northeas Normal University, China, in 2009 and the Ph.D. degree from the School of Mathematic and Statistic, Northeast Normal University, in 2017. She currently holds a post-doctoral position with the School of Psychology, Northeast Normal University. Her research interests include computer vision and image processing.

JIANZHONG WANG was born in Changchun, China. He received the B.S. degree from the Computer School, Jilin University, China, in 2004, the M.S. degree from the Computer School, Northeast Normal University, China, in 2007, and the Ph.D. degree from the School of Mathematics and Statistics, Northeast Normal University, in 2010. He is currently a Lecturer with the College of Information Science and Technology, Northeast Normal University. His research interests focus on dimensionality reduction and image processing.

JUN KONG was born in Jilin, China. He received the B.S. and M.S. degrees from the Department of Mathematics, Northeast Normal University, China, in 1992 and 1997, respectively, and the Ph.D. degree from the College of Mathematics, Jilin University, in 2001. From 2003 to 2004, he was a Visiting Scholar with Edith Cowan University, Perth, WA, Australia. He is currently a Professor with the College of Information Science and Technology, Northeast Normal University. His research interests include artificial intelligence, digital image processing, pattern recognition, machine learning, biometrics, and information security.