

Received April 10, 2020, accepted May 14, 2020, date of publication May 22, 2020, date of current version June 5, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2996796

Active Discriminative Cross-Domain Alignment for Low-Resolution Face Recognition

DONGDONG ZHENG¹, KAIBING ZHANG¹, JIAN LU¹, JUNFENG JING¹,
AND ZENGGANG XIONG²

¹School of Electronics and Information, Xi'an Polytechnic University, Xi'an 710048, China

²School of Computer and Information Science, Hubei Engineering University, Xiaogan 432000, China

Corresponding author: Kaibing Zhang (xihua_0169@163.com)

This work was supported in part by the National Natural Science Foundation of China under Grant 61971339, Grant 61972136, and Grant 61471161, in part by the Key Project of the Natural Science Foundation of Shaanxi Province under Grant 2018JZ6002 and Grant 2018GY-173, in part by the Hubei Provincial Department of Education Outstanding Youth Scientific Innovation Team Support Foundation under Grant T201410, in part by the Graduate Innovation Foundation of Xi'an Polytechnic University under Grant CHX2020016, and in part by the Doctoral Startup Foundation of Xi'an Polytechnic University under Grant BS1616.

ABSTRACT In real application scenarios, the face images captured by cameras often incur blur, illumination variation, occlusion, and low-resolution (LR), which leads to a challenging problem for many real-time face recognition systems due to a big distribution difference between the captured degraded images and the high-resolution (HR) gallery images. As widespread application of transfer learning in across-visual recognition, we propose a novel active discriminative cross-domain alignment (ADCDA) technique for LR face recognition method by jointly exploring both geometrical and statistical properties of the source domain and the target domain in a unique way. Specifically, the proposed ADCDA-based method contains three key components: 1) it simultaneously reduces the domain shift in both marginal distribution and conditional distribution between the source domain and the target domain; 2) it aligns the data of two domains in the common latent subspace by discriminant locality alignment (DLA); 3) it selects the representative and the diverse samples with an active learning strategy to further improve classification performance. Extensive experiments on six benchmark databases verify that the proposed method significantly outperforms other state-of-the-art predecessors.

INDEX TERMS Transfer learning, domain adaptation, discriminant locality alignment (DLA), active learning, low-resolution (LR) face recognition.

I. INTRODUCTION

Face recognition is one of the most activate research topics in the field of computer vision. Under a controlled imaging condition, most face recognition systems under HR face images able to achieve satisfying recognition performance. Unfortunately, in many practical scenarios, the performance of face recognition systems tends to be degraded dramatically due to the negative influence caused by the captured LR face images. Therefore, how to improve the recognition performance of LR faces has gained much attention in the domain of cross-resolution face recognition by many researchers.

Generally speaking, for many traditional machine learning-based image classification tasks, a fundamental

The associate editor coordinating the review of this manuscript and approving it for publication was Wei Wei¹.

assumption is that the samples from the source domain and the target domain have same data distribution. Based upon this assumption, the classifier learned from the source domain is accurate enough for the images in the target domain. In fact, there exists a very big distribution gap between the gallery images and the probe images due to many degradation factors such as illumination variation, pose change, occlusion, and LR, resulting in a poor performance for face recognition systems. Fortunately, transfer learning and domain adaptation have gained huge success in cross-modality and cross-domain image classification, which provides an enlightenment and idea to LR face recognition.

For many visual recognition tasks, it is crucial to obtain sufficient labeled data. Generally, based on the availability of labeled data in the target domain, domain adaptation can be generally divided into semi-supervised [1]–[3] and

unsupervised [4]–[8] domain adaptation. Since the unlabeled images are easier to obtain in practice, it is more significant to study unsupervised domain adaptation. In this paper, we mainly focus on the field of unsupervised domain adaptation.

According to a survey in [9], existing transfer learning approaches can be roughly categorized into four groups, namely instance-based, feature-based, parameter-based, and relational-based approaches, respectively. Instance-based [10]–[14] transfer learning approaches are mainly based on the instance weighting strategy. Feature transformation-based [15]–[19] approaches elaborate on transforming the source domain feature subspace into the target domain feature subspace, or transforming the feature spaces of both source and target domains into a common latent subspace to adapt two domains. Nevertheless, instance-based approaches are not fit for unsupervised visual recognition tasks due to the need of a large number of labeled source domain samples to train an accurate classifier for target domain. To mitigate this problem, in this paper, we will follow feature transformation-based approaches.

Pan *et al.* [20] proposed transfer component analysis (TCA) to learn the transfer components from different domains in Reproducing Kernel Hilbert Space (RKHS) with Maximum Mean Discrepancy (MMD), which aims to minimize the domain shift across domains distribution. To take advantage of low-dimensional structures that are intrinsic to many vision datasets, the work in [21] proposed a geodesic flow kernel (GFK) model to reduce domain shift by integrating an infinite number of subspaces that characterize changes in geometrical and statistical properties from the source domain to the target domain. JDA [22] improved the TCA by combining both marginal and conditional distributions to decrease the discrepancy between source domain and unlabeled target domain. Fernando *et al.* [23] introduced a subspace alignment (SA) framework to improve the matching of cross-domain image, which builds a subspace spanned by both source and target domains and learns a linear mapping to align the source subspace with the target subspace. In [24], TJM adapts both the distribution difference and the irrelevant instances by jointly matching the features and reweighting the instances across domains in a principled dimensionality reduction procedure. Considering that the existing unsupervised methods failed to take into account the difference of the two distributions in the subspace, Sun and Saenko [25] incorporates distribution alignment into subspace adaptation to address the problem. To improve alignment ability of the learned common subspace, the method in [26] proposed to impose low-rank and sparse constraints on the reconstruction coefficient matrix, so that the global and local structures of data can be well preserved. Tuia and Camps-Valls [27] proposed to transfer domain adaptation in a kernel-based feature space, and then manifold alignment is performed to preserve manifold geometric structures of both source and target domains. Ghifary *et al.* [28] presented scatter component analysis (SCA) to improve both domain adaptation

and domain generalization by simultaneously maximizing the separability of classes and minimizing the mismatch between two domains. In [29], two coupled projections are learned to project the source domain and target domain data into a low-dimensional subspace by simultaneously reducing the geometrical shift and distribution shift. Wang *et al.* [30] proposed class-specific reconstruction transfer learning (CRTL) model to exploit the intra-class dependency and inter-class independency of the reconstructed transfer matrix. In the particular method low-rank and sparse constraints are imposed on the class-specific reconstruction coefficient matrix such that the global and local data structures that contribute to domain correlation can be effectively preserved.

Recently, utilizing deep adaptation network [31]–[33] to deal with domain adaptation has gained much attention by many researchers. For example, Tzeng *et al.* [34] proposed to use CNN to solve deep domain adaptation, where an adaptation layer and an additional domain confusion loss are introduced to learn a semantic and domain invariant representation. In [35], Sun *et al.* presented to extend CORAL [36] to learn a nonlinear transformation to align the correlations of layer activations in deep neural networks (called Deep CORAL). [37] and [38] proposed to unite adversarial learning and domain adaptation to narrow the distribution difference between source and target domains data by combining discriminative modeling, weight sharing, and a GAN loss. Kang *et al.* [39] utilized a contrastive adaptation network (CAN) to optimize a new metric which explicitly models the intra-class domain discrepancy and the inter-class domain discrepancy.

In this paper, we develop a novel active discriminative cross-domain alignment (ADCDA) approach to target LR face recognition. Fundamentally, our method is related to JDA [22] and JGSA [29] but has many unique features. First, the newly proposed method joints both geometrical and statistical properties to adapt data distribution between labeled source and unlabeled target domains. Different from the two previous methods, we simultaneously minimize the difference of both the marginal distribution and conditional distribution between two different image domains, where Maximum Mean Discrepancy (MMD) [40] is employed to measure the difference in both marginal and conditional distributions. Moreover, inspired by the philosophy of subspace alignment proposed in [41], we explore both global and local manifold geometric structures in the source domain for obtaining more discriminative latent common subspace. In addition, we select more representative and diverse samples in the source domain to train the latent common subspace so as to benefit more accurate classification for the target domain. Taking the aforementioned factors into consideration, we propose an Active Discriminative Cross-Domain Alignment (ADCDA) by jointing active learning and domain adaptation to target LR face recognition task.

In summary, the major contributions of this paper are outlined as follows.

- 1) We reduce the domain shift between the labeled source and the unlabeled target domains by jointing both geometrical and statistical properties.
- 2) We align the feature subspaces spanned by both labeled source and unlabeled target domains in a learned common latent subspace by integrating subspace learning with DLA.
- 3) To further improve the classification accuracy for target domain, we employ the philosophy of active learning to select the representative and the diverse source samples to train our domain adaptation model.
- 4) Our proposed ADCDA approach can achieve the best recognition performance on six standard face datasets among eight state-of-the-art predecessors.

The rest of this paper is organized as follows. We briefly review the related work in Section II. Section III details the proposed Active Discriminative Cross-Domain Alignment (ADCDA). The validating experiments on six benchmark face datasets are demonstrated in Section IV. Finally Section V concludes this paper.

II. RELATED WORK

A. DISCRIMINANT LOCALITY ALIGNMENT

In the literature of machine learning, DLA [41] is an effective nonlinear dimensionality reduction algorithm. Unlike other classical dimensionality reduction algorithms such as principal component analysis (PCA) and linear discriminant analysis (LDA), the DLA performs a discriminant dimensionality reduction by using both local alignment and global alignment. It shows powerful capability of solving dimensionality reduction towards the samples with nonlinear distribution.

Generally, the DLA algorithm consists of three major stages: the part optimization, the sample weighting, and the whole alignment. In the phase of part optimization, the local patch constructed from the source domain samples contains rich discriminant information, where the local patch of each sample is associated with itself and its neighborhood samples belonging to the same class and different classes. In the phase of sample weighting, each partly optimized local patch is weighted by the margin degree to measure the importance of a given sample for classification. In the phase of whole alignment, the weighted and optimized local patches regarding each sample are globally aligned to further optimize the alignment matrix. Finally, a standard eigenvalue problem is solved to obtain the desired projection matrix for dimensionality reduction. The DLA shows good performance for high dimensional data with complicated nonlinear distribution.

B. ACTIVE LEARNING

In the machine learning domain, active learning sampling [42] is popular due to its capacity of finding a small number of representative and diverse labeled samples from a large number of training samples to train a more effective classifier. Usually, two important criteria, namely

representativeness and diversity, are widely considered in the active learning sampling. Given a sample, its representativeness is evaluated by a Gaussian kernel as

$$R(x_i) = \frac{1}{N_i} \sum_{x_i, x_j \in X, j \neq i} \exp\left(-\|x_i - x_j\|_2^2 / 2\sigma_R^2\right), \quad (1)$$

where $R(x_i)$ is the representativeness of sample x_i ; N_i is the neighbors of sample x_i ; $X = [x_1, x_2, \dots, x_{n_s}]$ is the set of original source domain samples; σ_R is the bandwidth of the Gaussian kernel, which is adaptively determined by the following Eq. (2), i.e.,

$$\sigma_R = \sqrt{\rho \left(\text{median}_{x_i, x_j \in X, i \neq j} \|x_i - x_j\|_2^2 \right)}, \quad (2)$$

where ρ is a scale coefficient. Fundamentally, a sample with higher representativeness will share more common information than those with lower representativeness.

The other criterion is diversity, which is evaluated by the following expression, i.e.,

$$D(x_i) = \min_{x_j \in S} \left[-\exp\left(-\|x_i - x_j\|_2^2 / 2\sigma_R^2\right) \right], \quad (3)$$

where $D(x_i)$ represents the diversity of sample x_i , S is the set of selected informative samples. For face images, the regions of eyes, nose, and mouth show diverse characteristics. Hence, we can utilize the diversity criterion to probe those samples with different appearances to improve the generality of proposed method.

In order to balance the function of both representativeness and diversity, Hoi *et al.* [42] proposed to find the most informative sample repeatedly in the remaining candidate sample set: $U = X - S$ by using the convex combination as below

$$\arg \max_{x_i \in U} (\lambda R(x_i) + (1 - \lambda) D(x_i)), \quad (4)$$

where λ is a trade-off parameter between the representativeness and the diversity of samples.

III. ACTIVE DISCRIMINATIVE CROSS-DOMAIN ALIGNMENT

In this section, we first describe the problem statement of proposed method. Next the details about ADCDA-based method, including its objective function and its optimization approach are described. Finally, a summarized algorithm for ADCDA-based LR face recognition is demonstrated.

A. PROBLEM DEFINITION

At first, we introduce the terminologies about transfer learning. The source domain data is indicated as $X^s \in \mathfrak{R}^{D \times n_s}$ and its marginal probability distribution as $P_s(X^s)$, the target domain data is denoted as $X^t \in \mathfrak{R}^{D \times n_t}$ and its marginal probability distribution as $P_t(X^t)$, where D is the dimensional size of image feature, n_s and n_t are the numbers of samples in the source and target domains, respectively. In this paper, we address an unsupervised domain adaptation problem of learning an accurate classifier with the help of sufficient labeled samples from the source domain

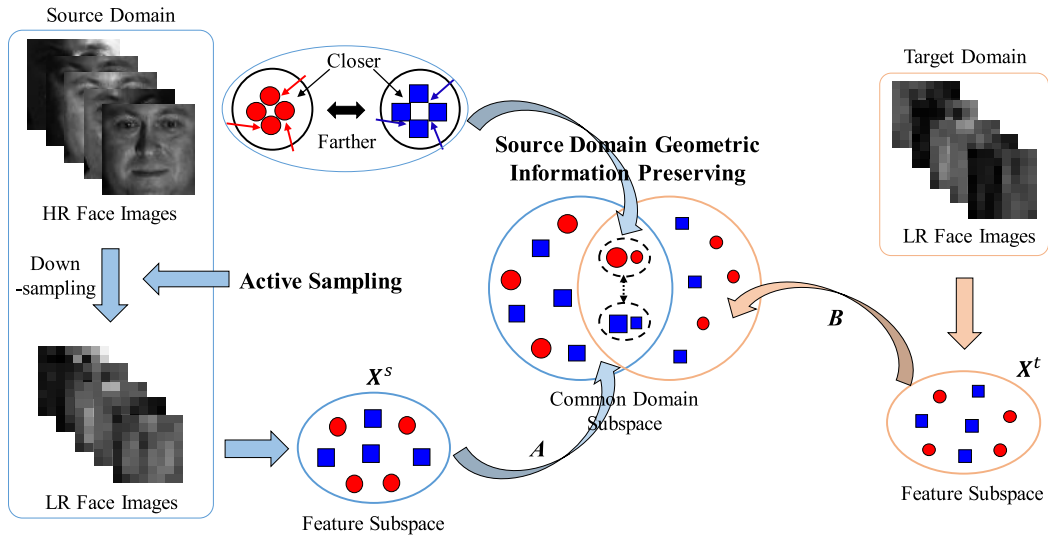


FIGURE 1. Illustration of the proposed method.

$D_s = \{(x_1, y_1), (x_2, y_2), \dots, (x_{n_s}, y_{n_s})\}$, and then apply it to classify the unlabeled samples in the target domain $D_t = \{x_1, x_2, \dots, x_{n_t}\}$. Assuming that the feature subspace and label subspace of the source and target domains are identical such that $X^s = X^t$, $Y^s = Y^t$. Due to domain shift of two domains, we have $P_s(X^s) \neq P_t(X^t)$ and $Q_s(Y^s | X^s) \neq Q_t(Y^t | X^t)$. As a result, our aim is to reduce the distribution difference by assuming $P_s(X^s) \approx P_t(X^t)$ and $Q_s(Y^s | X^s) \approx Q_t(Y^t | X^t)$. However, the conditional probability distribution $Q_t(Y^t | X^t)$ of target domain data is unknown. We address this problem by assuming $Q_t(Y^t | X^t) \approx Q_s(Y^t | X^t)$ [22].

B. ADCDA FRAMEWORK

The systematic flow chart of the proposed method is demonstrated in Fig.1. As illustrated, the newly proposed ADCDA-based LR face recognition method joints both geometrical alignment and statistical distributions to lessen the domain shift gap between the labeled source and unlabeled target domains. Moreover, to further improve the generality capability of the proposed method, we are in favor of using active sampling rather than random sampling, to choose more informative samples to train a more accurate classifier for the target domain. To accomplish LR face recognition, we conduct the domain adaptation by finding two coupled mappings, namely A for source domain and B for target domain to project the feature spaces of the source and target domains into a common latent subspace.

1) SOURCE DOMAIN GEOMETRIC INFORMATION PRESERVING

The success of DLA [41] motivated us to explore discriminative subspace spanned by the source domain. To this end, we combine local alignment, sample weighting, and global alignment to optimize a transformation matrix such that the

discriminative information of source domain can be well preserved. The local alignment is denoted as

$$\begin{aligned} & \arg \min_{X_i^s} \left(\sum_{j=1}^{k_w} \|x_i^s - x_{ij}^s\|^2 - \mu \sum_{p=1}^{k_b} \|x_i^s - x_{ip}^s\|^2 \right) \\ &= \arg \min_{X_i^s} \left(\sum_{j=1}^{k_w} \|x_i^s - x_{ij}^s\|^2 (\phi_i)_j + \sum_{p=1}^{k_b} \|x_i^s - x_{ip}^s\|^2 (\phi_i)_{p+k_w} \right) \\ &= \arg \min_{X_i^s} \text{tr} \left[X_i^s W_i (X_i^s)^T \right], \end{aligned} \tag{5}$$

where μ is a scaling factor ranged in $[0, 1]$ to balance the importance of intra-class and inter-class samples with respect to x_i^s ; k_w and k_b are the nearest neighborhood samples belonging to the same class and the different classes with respect to x_i^s , respectively;

$$\phi_i = \left[\overbrace{1, \dots, 1}^{k_w} \quad \overbrace{-\mu, \dots, -\mu}^{k_b} \right]^T$$

indicates a coefficient vector of the i -th local patch and

$$W_i = \begin{bmatrix} \sum_{j=1}^{k_w+k_b} (\phi_i)_j & -\phi_i^T \\ -\phi_i & \text{diag}(\phi_i) \end{bmatrix}$$

is the corresponding local alignment matrix.

To obtain better alignment, a margin degree function m_i is introduced to punish the samples nearby the classification boundary. The margin degree function m_i regarding to the i -th sample x_i^s is denoted as

$$m_i = \exp \left(-\frac{1}{(n_i + \delta)t} \right) \quad i = 1, \dots, n_s, \tag{6}$$

where n_i is the number of samples x_j^s whose labels are different from the label of x_i^s around a neighborhood circle of the sample x_i^s , δ is a regularization parameter, and t denotes a scaling factor [41]. By this way, the part optimization towards the i -th local patch can be weighted by a margin degree function of the i -th sample x_i^s , i.e.,

$$\begin{aligned} & \arg \min_{X_i^s} m_i \operatorname{tr} \left(X_i^s \mathbf{W}_i (X_i^s)^T \right) \\ & = \arg \min_{X_i^s} \operatorname{tr} \left(X_i^s m_i \mathbf{W}_i (X_i^s)^T \right). \end{aligned} \quad (7)$$

For each local patch X_i^s , Eq. (7) is used to weight the part optimization to obtain the discriminative alignment matrix. Finally, the part optimizations of all the local patches are unified as a whole one by assuming that the coordinate for the i -th patch $X_i^s = [x_{i_1}^s, x_{i_2}^s, \dots, x_{i_{k_w}}^s, x_{i_1}^s, \dots, x_{i_{k_b}}^s]$ is selected from the global coordinate $X^s = [X_1^s, \dots, X_{n_s}^s]$ [41], such that

$$X_i^s = X^s S_i, \quad (8)$$

where $S_i \in \mathfrak{R}^{n_s \times (k_w + k_b + 1)}$ indicates the selection matrix, which is defined as

$$(S_i)_{pq} = \begin{cases} 1 & \text{if } p = F_i\{q\} \\ 0 & \text{else} \end{cases}, \quad (9)$$

where $F_i = \{i, i^1, \dots, i^{k_w}, i_1, \dots, i_{k_b}\}$ is the index set for the i -th local patch.

By incorporating the selection matrix Eq. (9) into Eq. (7), the objective of alignment can be further rewritten as

$$\arg \min_{X^s} \operatorname{tr} \left(X^s S_i m_i \mathbf{W}_i S_i^T (X^s)^T \right). \quad (10)$$

By summing over all the part optimization described in Eq. (10) together, we can obtain the whole alignment as

$$\begin{aligned} & \arg \min_{X^s} \sum_{i=1}^{n_s} \operatorname{tr} \left(X^s S_i m_i \mathbf{W}_i S_i^T (X^s)^T \right) \\ & = \arg \min_{X^s} \operatorname{tr} \left(X^s \left(\sum_{i=1}^{n_s} S_i m_i \mathbf{W}_i S_i^T \right) (X^s)^T \right) \\ & = \arg \min_{X^s} \operatorname{tr} \left(X^s \mathbf{W} (X^s)^T \right), \end{aligned} \quad (11)$$

where $\mathbf{W} = \sum_{i=1}^{n_s} S_i m_i \mathbf{W}_i S_i^T \in \mathfrak{R}^{n_s \times n_s}$ represents the discriminative alignment matrix. Finally, we can obtain the updated discriminative alignment matrix by an iterative procedure, which contains the local manifold structures, the discriminant information, and the label information of source domain samples. The update procedure is represented as below

$$\mathbf{W}(F_i, F_i) \leftarrow \mathbf{W}(F_i, F_i) + m_i \mathbf{W}_i. \quad (12)$$

Finally, the constructed discriminative alignment matrix for preserving the geometrical discriminative information of labeled source domain samples is formulated as below

$$\min_A \operatorname{Tr} (A^T \mathbf{W} A). \quad (13)$$

2) DOMAIN SHIFT MINIMIZATION

We utilize the Maximum Mean Discrepancy (MMD) [40] criteria to reduce the statistical probability distribution of both source and target domains. The distribution distance between two domains can be measured by MMD as follow

$$\min_{A, B} \left\| \frac{1}{n_s} \sum_{x_i \in X^s} A^T x_i - \frac{1}{n_t} \sum_{x_j \in X^t} B^T x_j \right\|_F^2. \quad (14)$$

In the JDA [22], the authors proposed to apply a certain base classifiers trained on the labeled source data to find *pseudo* labels of the target data by employing an iterative pseudo label refinement strategy to minimize the difference between the conditional distributions of two domains. In the proposed method, we also follow JDA [22] to reduce the conditional distribution shift across two domains as below

$$\min_{A, B} \left\| \frac{1}{n_s^{(c)}} \sum_{x_i \in X^{s,(c)}} A^T x_i - \frac{1}{n_t^{(c)}} \sum_{x_j \in X^{t,(c)}} B^T x_j \right\|_F^2. \quad (15)$$

To achieve effective and sufficient domain adaptation, we simultaneously reduce the shift in both the marginal distribution and the conditional distribution between two domains. As such, the objective of minimizing domain shifts can be reformulated as below

$$\begin{aligned} & \min_{A, B} \left(\left\| \frac{1}{n_s} \sum_{x_i \in X^s} A^T x_i - \frac{1}{n_t} \sum_{x_j \in X^t} B^T x_j \right\|_F^2 \right. \\ & \left. + \left\| \frac{1}{n_s^{(c)}} \sum_{x_i \in X^{s,(c)}} A^T x_i - \frac{1}{n_t^{(c)}} \sum_{x_j \in X^{t,(c)}} B^T x_j \right\|_F^2 \right). \end{aligned} \quad (16)$$

Like JGSA [29], Eq. (16) can be further transformed into a more concise matrix form by the knowledge of linear algebra, i.e.,

$$\min_{A, B} \operatorname{Tr} \left([A^T \ B^T] \begin{bmatrix} \mathbf{M}_s & \mathbf{M}_{st} \\ \mathbf{M}_{ts} & \mathbf{M}_t \end{bmatrix} \begin{bmatrix} \mathbf{A} \\ \mathbf{B} \end{bmatrix} \right), \quad (17)$$

where

$$\begin{aligned} \mathbf{M}_s &= X^s \left(\mathbf{H}_s + \sum_{c=1}^C \mathbf{H}_s^{(c)} \right) (X^s)^T, \quad \mathbf{H}_s = \frac{1}{n_s n_s} \mathbf{1}_s \mathbf{1}_s^T, \\ (\mathbf{H}_s^{(c)})_{ij} &= \begin{cases} \frac{1}{n_s^{(c)} n_s^{(c)}} & \text{if } x_i, x_j \in X^{s,(c)} \\ 0 & \text{otherwise,} \end{cases} \end{aligned} \quad (18)$$

$$\begin{aligned} \mathbf{M}_{st} &= X^s \left(\mathbf{H}_{st} + \sum_{c=1}^C \mathbf{H}_{st}^{(c)} \right) (X^t)^T, \quad \mathbf{H}_{st} = -\frac{1}{n_s n_t} \mathbf{1}_s \mathbf{1}_t^T, \\ (\mathbf{H}_{st}^{(c)})_{ij} &= \begin{cases} -\frac{1}{n_s^{(c)} n_t^{(c)}} & \text{if } x_i \in X^{s,(c)}, x_j \in X^{t,(c)} \\ 0 & \text{otherwise,} \end{cases} \end{aligned} \quad (19)$$

$$\mathbf{M}_{ts} = \mathbf{X}^t \left(\mathbf{H}_{ts} + \sum_{c=1}^C \mathbf{H}_{ts}^{(c)} \right) (\mathbf{X}^s)^T, \quad \mathbf{H}_{ts} = -\frac{1}{n_t n_s} \mathbf{1}_t \mathbf{1}_s^T,$$

$$\left(\mathbf{H}_{ts}^{(c)} \right)_{ij} = \begin{cases} -\frac{1}{n_t^{(c)} n_s^{(c)}} & \text{if } x_i \in \mathbf{X}^{t,(c)}, x_j \in \mathbf{X}^{s,(c)} \\ 0 & \text{otherwise,} \end{cases} \quad (20)$$

$$\mathbf{M}_t = \mathbf{X}^t \left(\mathbf{H}_t + \sum_{c=1}^C \mathbf{H}_t^{(c)} \right) (\mathbf{X}^t)^T, \quad \mathbf{H}_t = \frac{1}{n_t n_t} \mathbf{1}_t \mathbf{1}_t^T,$$

$$\left(\mathbf{H}_t^{(c)} \right)_{ij} = \begin{cases} \frac{1}{n_t^{(c)} n_t^{(c)}} & \text{if } x_i, x_j \in \mathbf{X}^{t,(c)} \\ 0 & \text{otherwise.} \end{cases} \quad (21)$$

3) SUBSPACE ADAPTATION

To simultaneously minimize the distribution difference and achieve domain adaptation match between two different domains, similar to SA [23] and SDA [25], we make the feature subspaces of both source and target domains as closer as possible such that the domain shift can be effectively reduced. In order to reduce the distribution discrepancy across domains, the SA-based method [23] formulates an additional transformation matrix \mathbf{M} to project the source feature subspace into the target feature subspace. Different from SA [23], we minimize distribution difference of cross-domain adaptation by learning two coupled mappings (\mathbf{A} and \mathbf{B}). With the two mappings, we can convert the feature spaces of two different domains into a common latent subspace. The minimization of adaptation towards the two coupled mappings is represented as

$$\min_{\mathbf{A}, \mathbf{B}} \|\mathbf{A} - \mathbf{B}\|_F^2. \quad (22)$$

To simultaneously explore the shared and domain specific features across domains, we incorporate Eq. (22) into Eq. (17) to adapt the geometrical and statistical distributions of source and target domains.

4) ACTIVE SAMPLING

To further boost up the performance of the proposed DCDA method, we employ an active sampling, which has been reviewed in Section II-B, to select more informative labeled samples in the source domain to train an accurate classifier for classification of the unlabeled samples in the target domain. The effectiveness between DCDA and ADCDA will be proved in the following experimental section.

C. OBJECTIVE FUNCTION AND OPTIMIZATION

By incorporating Eqs (13), (17), and (22) together, the final objective function of the proposed method is described as follow

$$\min_{\mathbf{A}, \mathbf{B}} \left(\text{Tr} \left(\begin{bmatrix} \mathbf{A}^T & \mathbf{B}^T \end{bmatrix} \begin{bmatrix} \mathbf{M}_s & \mathbf{M}_{st} \\ \mathbf{M}_{ts} & \mathbf{M}_t \end{bmatrix} \begin{bmatrix} \mathbf{A} \\ \mathbf{B} \end{bmatrix} + \alpha \mathbf{A}^T \mathbf{W} \mathbf{A} \right) + \beta \|\mathbf{A} - \mathbf{B}\|_F^2 \right)$$

$$\text{s.t. } \mathbf{A}^T \mathbf{X}^s \mathbf{C}^s (\mathbf{X}^s)^T \mathbf{A} = \mathbf{I} \text{ and } \mathbf{B}^T \mathbf{X}^t \mathbf{C}^t (\mathbf{X}^t)^T \mathbf{B} = \mathbf{I}, \quad (23)$$

where α and β are two trade-off parameters to balance the importance between different terms; $\mathbf{C}^s = \mathbf{I}_s - \frac{1}{n_s} \mathbf{1}_s \mathbf{1}_s^T$ and

$\mathbf{C}^t = \mathbf{I}_t - \frac{1}{n_t} \mathbf{1}_t \mathbf{1}_t^T$ are the centering matrixes; $\mathbf{1}_s \in \mathbb{R}^{n_s}$ and $\mathbf{1}_t \in \mathbb{R}^{n_t}$ are the column vectors with all ones.

Afterwards, two coupled mappings are constructed to match the source and target domains in a learned common latent subspace, which leads to an optimization problem as below

$$\min_{\mathbf{A}, \mathbf{B}} \text{Tr} \left(\begin{bmatrix} \mathbf{A}^T & \mathbf{B}^T \end{bmatrix} \begin{bmatrix} \mathbf{M}_s + \alpha \mathbf{W} + \beta \mathbf{I} & \mathbf{M}_{st} - \beta \mathbf{I} \\ \mathbf{M}_{ts} - \beta \mathbf{I} & \mathbf{M}_t + \beta \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{A} \\ \mathbf{B} \end{bmatrix} \right)$$

$$\text{s.t. } \mathbf{A}^T \mathbf{X}^s \mathbf{C}^s (\mathbf{X}^s)^T \mathbf{A} = \mathbf{I} \text{ and } \mathbf{B}^T \mathbf{X}^t \mathbf{C}^t (\mathbf{X}^t)^T \mathbf{B} = \mathbf{I}. \quad (24)$$

Similar to JDA [22], to obtain a more concise expression, let $\mathbf{G}^T = [\mathbf{A}^T \ \mathbf{B}^T]$. Thus, the objective function in Eq. (24) can be rewritten as

$$\min_{\mathbf{A}, \mathbf{B}} \text{Tr} \left(\mathbf{G}^T \begin{bmatrix} \mathbf{M}_s + \alpha \mathbf{W} + \beta \mathbf{I} & \mathbf{M}_{st} - \beta \mathbf{I} \\ \mathbf{M}_{ts} - \beta \mathbf{I} & \mathbf{M}_t + \beta \mathbf{I} \end{bmatrix} \mathbf{G} \right)$$

$$\text{s.t. } \mathbf{A}^T \mathbf{X}^s \mathbf{C}^s (\mathbf{X}^s)^T \mathbf{A} = \mathbf{I} \text{ and } \mathbf{B}^T \mathbf{X}^t \mathbf{C}^t (\mathbf{X}^t)^T \mathbf{B} = \mathbf{I}. \quad (25)$$

Next, according to the constrained optimization theory, we let $\Phi = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_k) \in \mathbb{R}^{k \times k}$ be the Lagrange multiplier, and the Lagrange function for problem (25) is derived as

$$L = \text{Tr} \left(\mathbf{G}^T \begin{bmatrix} \mathbf{M}_s + \alpha \mathbf{W} + \beta \mathbf{I} & \mathbf{M}_{st} - \beta \mathbf{I} \\ \mathbf{M}_{ts} - \beta \mathbf{I} & \mathbf{M}_t + \beta \mathbf{I} \end{bmatrix} \mathbf{G} \right)$$

$$+ \text{Tr} \left(\left(\mathbf{A}^T \mathbf{X}^s \mathbf{C}^s (\mathbf{X}^s)^T \mathbf{A} \right) \Phi \right)$$

$$+ \text{Tr} \left(\left(\mathbf{B}^T \mathbf{X}^t \mathbf{C}^t (\mathbf{X}^t)^T \mathbf{B} \right) \Phi \right). \quad (26)$$

By setting $\frac{\partial L}{\partial \mathbf{G}} = 0$, we obtain a generalized eigenvalue decomposition as below

$$\begin{bmatrix} \mathbf{M}_s + \alpha \mathbf{W} + \beta \mathbf{I} & \mathbf{M}_{st} - \beta \mathbf{I} \\ \mathbf{M}_{ts} - \beta \mathbf{I} & \mathbf{M}_t + \beta \mathbf{I} \end{bmatrix} \mathbf{G} = \begin{bmatrix} \mathbf{X}^s \mathbf{C}^s (\mathbf{X}^s)^T \\ \mathbf{X}^t \mathbf{C}^t (\mathbf{X}^t)^T \end{bmatrix} \mathbf{G} \Phi. \quad (27)$$

Finally, the optimal adaptation matrix \mathbf{G} is boiled down to solving Eq. (27) for the k smallest eigenvectors $\mathbf{G} = [g_1, g_2, \dots, g_k]$. A complete procedure of ADCDA-based method for LR face recognition is summarized in **Algorithm 1**.

IV. EXPERIMENTS

In this section, we conduct extensive experiments on six benchmark face databases to verify the effectiveness of proposed ADCDA-based method. The compared approaches include two baseline methods (i.e., HR-PCA and HR-LDA), four state-of-the-art transfer learning based approaches (i.e., transfer component analysis (TCA) [20], joint distribution analysis (JDA) [22], transfer joint matching (TJM) [24], and joint geometrical and statistical alignment (JGSA) [29]), and two coupled mapping based methods (i.e., coupled locality preserving mappings (CLPMs) [43] and coupled discriminant multi-manifold analysis (CDMMA) [44]).

A. BENCHMARK FACE DATABASES

In the experiment, we employ six benchmark face databases including YALE-B, UMIST, ORL, FERET, CMU-PIE, and

Algorithm 1 (Active) Discriminative Cross-Domain Alignment (DCDA and ADCDA)

Input:

The source domain face images $X^s \in \mathbb{R}^{D \times n_s}$, the target domain face images $X^t \in \mathbb{R}^{D \times n_t}$, the class labels of source domain face images $Y^s \in \mathbb{R}^{1 \times n_s}$, the dimensional size of subspace k , and the model parameters λ , α , and β .

Output:

Adaptation transformation matrices: A and B , embeddings: Z^s and Z^t , and adaptive classifier f .

- 1: Select informative samples from the source domain via Eqs. (1-4) (ADCDA);
 - 2: Construct discriminative alignment matrix via Eqs. (5-12);
 - 3: Build M_s , M_{st} , M_{ts} , and M_t via Eqs. (18), (19), (20), and (21), respectively;
 - 4: **repeat**
 - 5: Solve the generalized eigenvalue decomposition problem in Eqs. (27);
 - 6: Obtain the adaptation transformation matrices A and B from G ;
 - 7: Project the face features of source domain and target domain into the embedded subspaces via $Z^s = A^T X^s$ and $Z^t = B^T X^t$;
 - 8: Train a base classifier f on $\{Z^s, Y^s\}$ and update the pseudo labels of the target domain via $\hat{Y}^t = f(Z^t)$;
 - 9: Update M_s , M_{st} , M_{ts} , and M_t via Eqs. (18), (19), (20), and (21);
 - 10: **until** Convergence
 - 11: Obtain the final adaptive classifier f trained on $\{Z^s, Y^s\}$.
-

AR to assess the performance of compared methods. **Fig.2** shows some example samples selected from six benchmark face databases.

The Extended YALE-B database contains 16,128 face images of 38 subjects with 9 different poses and 64 different illumination conditions. Each individual in the database has about 64 face images with frontal pose and different illumination conditions, a total of 2,414 face images. The CMU-PIE face database contains 41,368 face images of 68 subjects with 13 different poses, 43 different lighting conditions, and 4 different expressions. In this database, each person has about 49 face images with 27 positive postures and all lighting conditions, a total of 3,329 face images. All the face images are manually aligned and cropped, and the size of HR face images is set to 32×28 .

The UMIST face database contains 564 face images of 20 subjects with different angles and postures. The ORL face database contains 400 face images of 40 individuals with 10 different postures, expressions, and occlusions for each person. All the face images are manually aligned and cropped, and the size of HR face images is set to 90×80 .

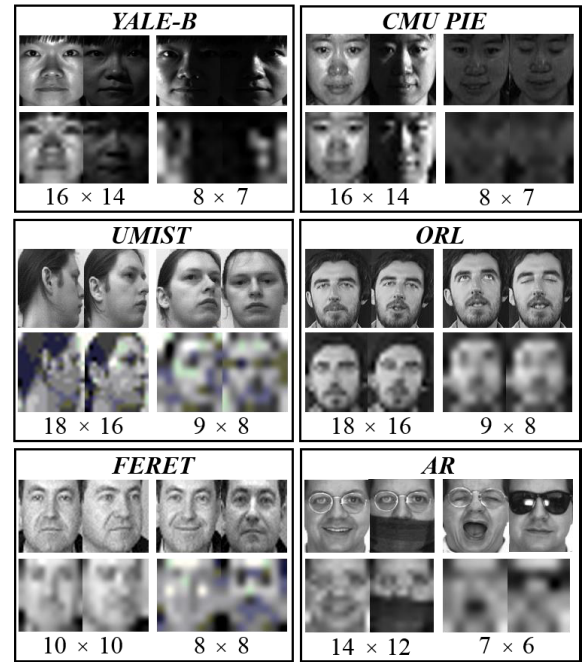


FIGURE 2. Some example samples of six benchmark face databases. (The first row is HR images, and the second row corresponds to LR images.)

The FERET database contains 14,126 face images of 1199 subjects with different poses, expressions, and illumination conditions. In our experiments, we randomly choose 200 individuals from 1199 subjects, each person has about 7 face images with frontal pose, expressions, and different illumination conditions, a total of 1,400 face images. All the face images are manually aligned and cropped, and the size of HR face images is set to 40×40 .

The AR database contains 3,288 face images of 116 subjects with different poses, expressions, and illumination conditions. In our experiments, we randomly choose 100 individuals from 116 subjects, in which each person has about 26 face images with frontal pose, expressions, and different illumination conditions, a total of 2,600 face images. All the face images are manually aligned and cropped. The size of all HR face images is set to 140×120 .

For each face database, we randomly select one half of each person as the source domain, and the remaining half as the target domain. To mimic the LR face images, the original HR face images in the YALE-B and CMU-PIE face databases are smoothly down-sampled to the size 16×14 and 8×7 by the factors of 2 and 4, respectively. Similarly, the HR face images in the UMIST and ORL face databases are smoothly down-sampled to the size 18×16 and 9×8 by the factors of 5 and 10, respectively. The HR face images in the FERET face database are smoothly down-sampled to the size 10×10 and 8×8 by the factors of 4 and 5, respectively. The HR face images in the AR face database are smoothly down-sampled to the size 14×12 and 7×6 by the factors of 10 and 20, respectively.

TABLE 1. Accuracy (%) of Eight State-of-the-Art Approaches and the Proposed Method with Two Resolutions on Six Standard Face Databases.

Face Database	Resolution	HR-PCA	HR-LDA	TCA [20]	JDA [22]	TJM [24]	JGSA [29]	CLPMs [43]	CDMMA [44]	DCDA	ADCDA
YALE-B	8 × 7	81.47	87.48	43.36	83.22	15.68	51.30	86.48	90.26	88.98	99.74
	16 × 14	81.47	87.48	61.79	85.27	17.81	68.86	92.40	95.09	94.53	99.92
UMIST	9 × 8	90.09	90.33	93.07	95.98	90.65	92.26	90.12	91.33	96.59	100
	18 × 16	90.09	90.33	93.37	96.59	91.33	93.50	88.00	91.87	97.21	100
ORL	9 × 8	83.50	84.50	86.90	90.60	78.90	85.50	82.50	82.80	92.50	97.50
	18 × 16	83.50	84.50	88.20	91.00	80.67	89.00	80.00	82.20	92.00	97.00
FERET	8 × 8	39.67	91.67	43.12	70.23	28.30	42.70	52.67	71.38	80.39	91.23
	10 × 10	39.67	91.67	42.33	73.92	27.95	44.61	59.33	72.83	82.67	93.00
CMU-PIE	8 × 7	94.66	97.22	86.72	94.78	52.23	77.43	96.22	97.81	98.53	99.98
	16 × 14	94.66	97.22	91.69	95.00	56.33	87.82	96.89	97.54	98.75	100
AR	7 × 6	52.69	94.69	36.87	55.08	26.29	30.55	66.31	72.94	61.72	82.53
	14 × 12	52.69	94.69	41.48	57.81	27.58	39.08	85.08	90.00	83.53	94.84

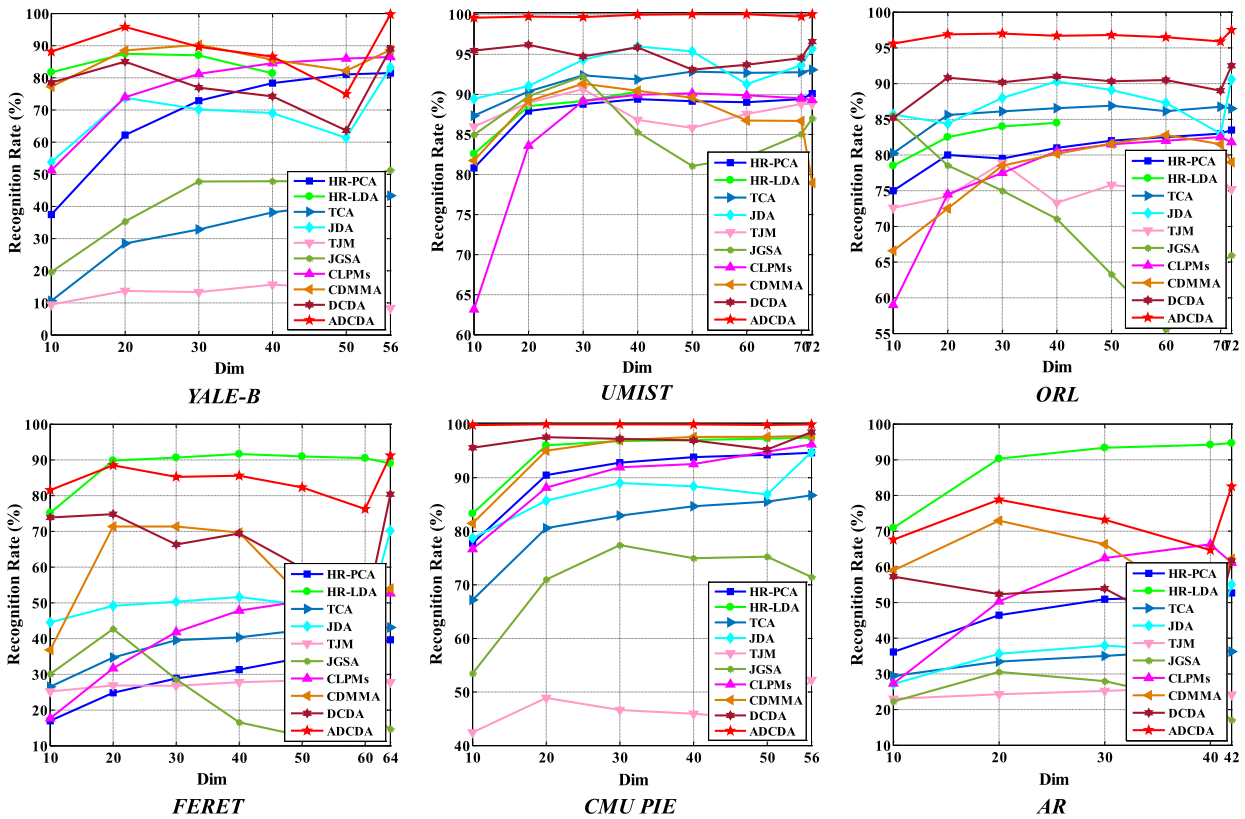


FIGURE 3. Recognition results of different methods with different dimensions on six face databases.

B. EXPERIMENTAL RESULTS

The recognition performance of the proposed method and the eight state-of-the-art approaches on six face datasets are illustrated in Table 1. As reported in Table 1, except slightly lower recognition than CDMMA on the YALE-B and AR face databases, the proposed DCDA (randomly divides source and target domains) method constantly achieves the best recognition performance on the rest of face databases with two resolutions. Moreover, when active learning is used to select source domain samples, the proposed ADCDA method significantly outperforms other competitors for all the face databases at different resolutions. More specifically,

the improvement of ADCDA increases by 9.48 percent and 4.83 percent on YALE-B face database, and 4.02 percent and 3.41 percent on UMIST face database, and 6.9 percent and 6.0 percent on ORL face database, and 19.85 percent and 19.08 percent on FERET face database (except for HR-LDA), and 2.17 percent and 2.46 percent on CMU-PIE face database, and 9.59 percent and 4.84 percent on AR face database, respectively. Particularly, the recognition rate of our ADCDA-based method surprisingly achieves 100 percent on UMIST and CMU-PIE face databases, respectively. The perfect performance mainly contributes to jointly applying both geometrical and statistical properties, which benefits

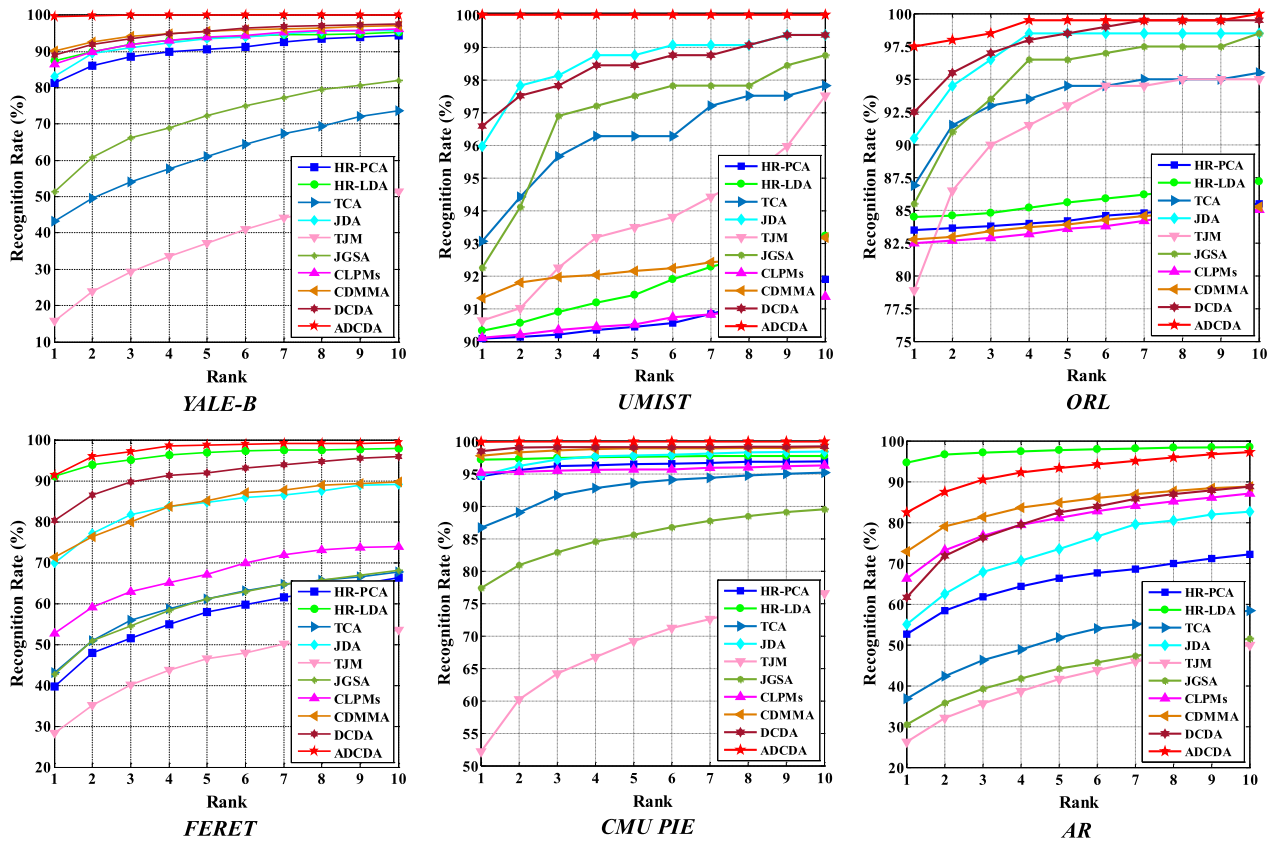


FIGURE 4. Recognition results of different methods with different ranks on six face databases.

to not only reduce the geometrical shift of subspace but also minimize the distribution shifts between two domains. Moreover, active sampling selects more informative samples from the source domain and therefore it is propitious to enhance the generality performance of the proposed DCDA-based method. In the following experiments, we only demonstrate the experimental results obtained by ADCDA-based method as comparison.

C. EXPERIMENTAL ANALYSIS ON DIMENSIONALITY k OF FEATURE SUBSPACE

In this subsection, we study how the dimensional size k (Dim) of feature subspace affects the recognition performance. In this experiment, the rank is set to 1 for all databases. The size of LR face images of the YALE-B and CMU-PIE face databases is set to 8×7 , that of UMIST and ORL face databases is set to 9×8 , that of FERET face database is set to 8×8 , and that of AR face database is set to 7×6 , respectively. As shown in Fig.3, the proposed ADCDA can achieve the best performance when the highest feature dimension is used for matching. Except the AR face database, the performance of the newly proposed approach constantly performs better than other methods. Moreover, our method also exceeds the four transfer learning-based approaches.

D. EXPERIMENTAL ANALYSIS ON RANK

To examine how the rank influences the recognition rate, we range the rank from 1 to 10 to verify the advantage of the proposed method. In this experiment, the experimental settings are same as Section IV-C. As shown in Fig.4, for all benchmark face databases, the recognition rates of ADCDA as well as other approaches steadily increase as the rank increases within the range from 1 to 10. Particularly, the performance of our method is always better than other compared methods at each case except for AR face database.

E. EXPERIMENTAL ANALYSIS ON DIFFERENT RESOLUTIONS

In this subsection, we further evaluate the performance of different methods at two different resolution levels, namely 8×7 and 16×14 for the YALE-B and the CMU-PIE face databases, and 9×8 and 18×16 for the UMIST and the ORL face databases, and 8×8 and 10×10 for the FERET face database, and 7×6 and 14×12 for the AR face database, respectively. Fig.5 demonstrates the compared results on corresponding to different methods and two resolution levels. In terms of the results, we can see that except for AR face database, the proposed ADCDA method exceeds other methods at the all cases.

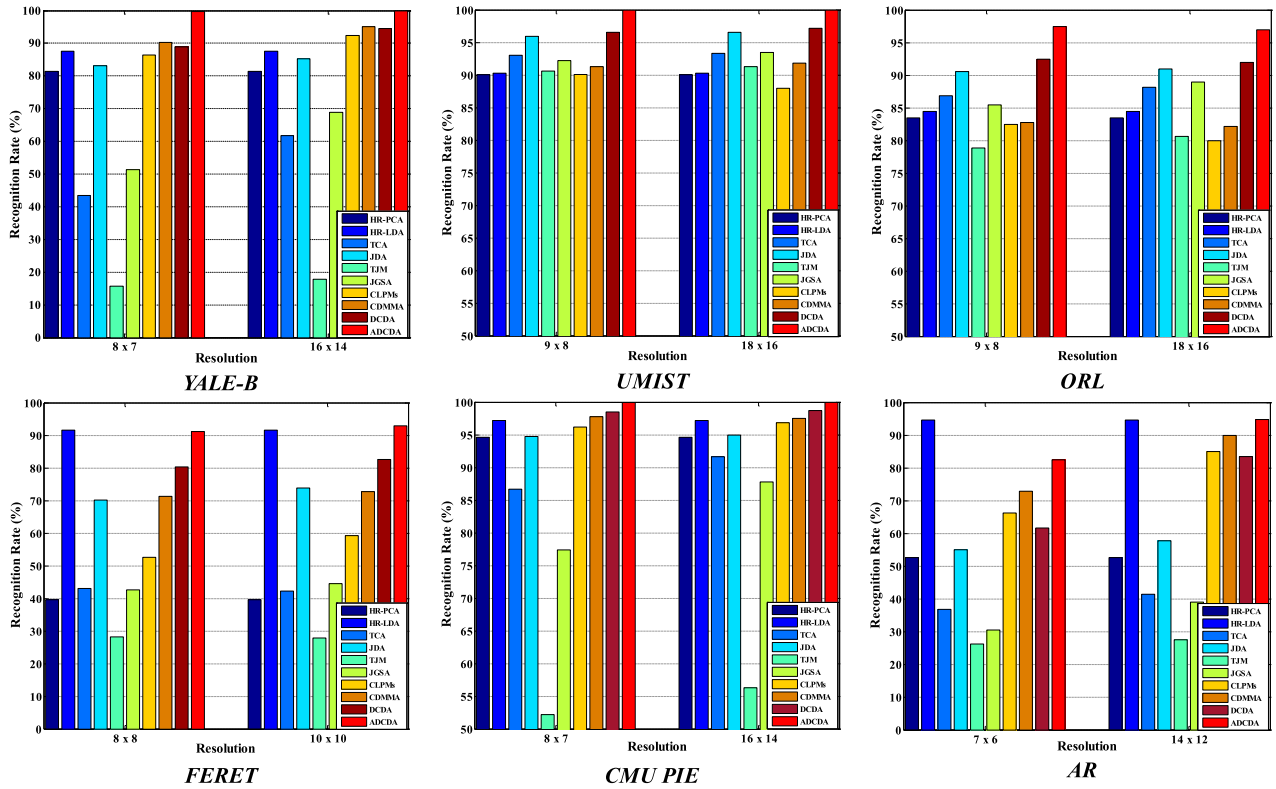


FIGURE 5. Recognition results of different methods with two different resolutions on six face databases.

F. EFFECTIVENESS ANALYSIS OF FEATURE SUBSPACE

In this subsection, we further examine the effectiveness of our cross-domain alignment method by projecting the learnt common features of six benchmark face databases into a 2-D space via PCA-based dimensionality reduction. The visualization results of compared approaches are illustrated in Fig.6. In terms of the results, we can find that the transformed features of source and target domains by the proposed ADCDA method in the common feature subspace can be aligned better than other predecessors. The major reasons lie in the following interpretations. The classical TCA method only considers marginal distribution of data without utilizing label information. The JDA method considers both marginal distribution and conditional distribution of data into consideration, but fails to reduce the geometrical shift during cross-domain adaptation. The TJM method minimizes the domain shift by jointly matching the features and reweighting the instances across domains but cannot consider any discriminative information, leading to the worst discrimination and separability. Although the JGSA method can effectively reduce both geometrical shift and distribution shift, it only explores the global discriminative information by using LDA. The CLPMs method elaborates on preserving the local geometric structures of source domain samples into the latent common subspace. Nevertheless, the method ignores the discriminative information of samples in the source domain. The CDMMA approach explores both the local geometric structure information and the label information to improve

the discrimination of mappings. Unfortunately, many noticeably aggregated samples remain in the feature subspace. In contrast to the above competitors, the newly proposed ADCDA method can significantly promote the discriminative capability and separability of the samples in the common latent subspace. Particularly, the counterparts of face images within the same class are well aligned and aggregated while those from different classes are obviously separated from each other in the learnt common latent subspace. This is due to that the proposed method jointly frames the local patches optimization, the sample weighting, and the global alignment of samples into a unified discriminative alignment matrix, which benefits to adapt the nonlinear distribution of face images in the source and target domains, leading to a more discriminative feature subspace than other competitors.

G. PARAMETER SENSITIVITY

In the proposed ADCDA, three important parameters, namely λ , α , and β , fundamentally affect the performance of proposed algorithm. To investigate how the three parameters have influence on the recognition performance, we take the YALE-B face database as test and repeat the experiments five times to explore the influence. To this end, we vary λ from 0 to 1 at an interval of 0.1, and empirically set $\alpha \in [0, 0.1]$ and $\beta \in [0.001, 100]$ to search the optimal parameters. It is worth noticing that λ is a trade-off parameter between the representativeness and the diversity, where $\lambda \rightarrow 0$ indicates the diversity of samples while $\lambda \rightarrow 1$ represents the

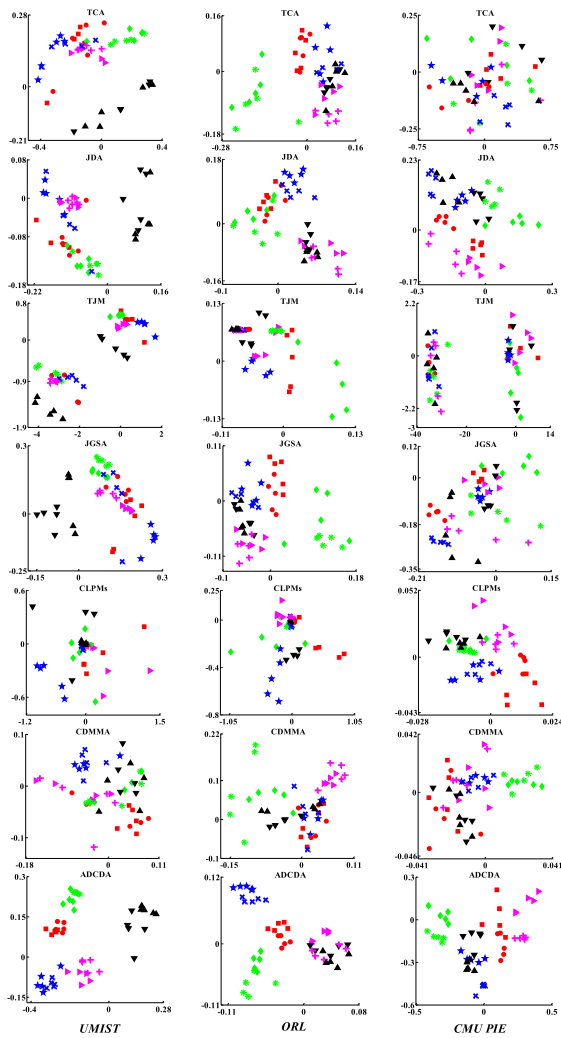


FIGURE 6. Comparison of the projection feature visualization for different approaches on the UMIST, ORL, and CMU PIE face databases. For visualization, we display the first two principal components of learned common features on a 2-D feature space, where the points of the same color and different shapes represent the source and target domains face images belonging to the same person, and those of different colors indicate different persons.

representativeness. **Fig.7a** illustrates the changes of recognition rate varying with λ . As illustrated from **Fig.7a**, the proposed ADCDA-based method tends to obtain

better performance with a smaller value of λ . It reflects the importance of diversity of samples to classifier. α is another trade-off parameter for balancing the importance of geometric structures of samples. As shown in **Fig.7b**, we can find that the recognition performance will drop greatly when the geometric structures of samples are not considered. By contrast, the recognition rate rises rapidly when the geometric structures of samples can be fully utilized. The fact proves that it is essential to explore both the intra-class and the inter-class information for improving the performance of transfer learning based visual recognition tasks. In addition, β is used to adapt the geometrical and statistical distributions between the source and target domains. As shown in **Fig.7c**, the recognition rate of the proposed method descends rapidly as β increases, which means that a smaller value of β is more beneficial to improve the recognition performance.

V. CONCLUSION

In this paper, we have proposed a novel ADCDA-based transfer subspace learning approach for LR face recognition. In the proposed method, both marginal distribution and conditional distribution between the source domain and the target domain are simultaneously explored to reduce the discrepancy of the two domains. To further improve the generality performance of proposed method, we employ active sampling to select more representative as well as diverse samples for subspace learning. Comprehensive experiments carried out upon six benchmark face databases have verified the effectiveness of our newly proposed method.

In the future, we will extend our model from two aspects. On one hand, to enhance the robustness of estimated pseudo labels of target domain samples, we can frame our method under a co-transfer ensemble learning framework [45]. On the other hand, to estimate more accurate conditional distribution of target domain, other semi-supervised manifold learning-based method [46] can be utilized to optimize the pseudo labels.

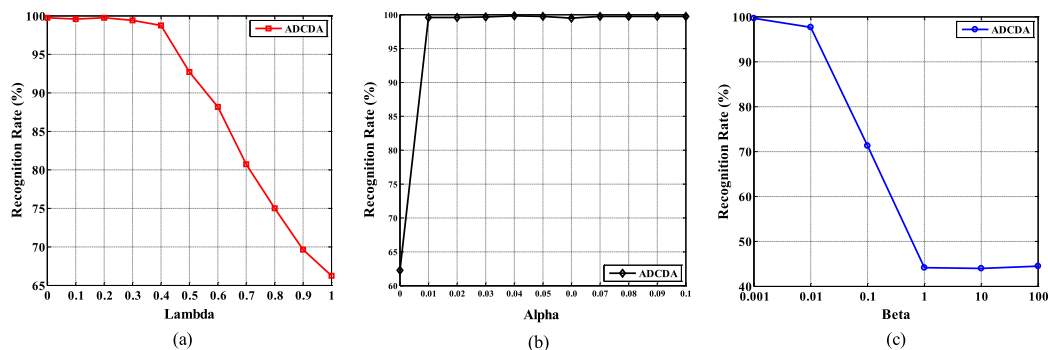


FIGURE 7. Recognition results of the proposed method with different parameters based on the YALE-B face database. (a), (b), and (c) indicate λ , α , and β respectively.

REFERENCES

- [1] W. Li, L. Duan, D. Xu, and I. W. Tsang, "Learning with augmented features for supervised and semi-supervised heterogeneous domain adaptation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 36, no. 6, pp. 1134–1148, Jun. 2014.
- [2] K. Yan, L. Kou, and D. Zhang, "Learning domain-invariant subspace using domain features and independence maximization," *IEEE Trans. Cybern.*, vol. 48, no. 1, pp. 288–299, Jan. 2018.
- [3] R. K. Sanodiya, J. Mathew, S. Saha, and M. D. Thalakkottur, "A new transfer learning algorithm in semi-supervised setting," *IEEE Access*, vol. 7, pp. 42956–42967, 2019.
- [4] M. Long, G. Ding, J. Wang, J. Sun, Y. Guo, and P. S. Yu, "Transfer sparse coding for robust image representation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2013, pp. 407–414.
- [5] M. Long, J. Wang, G. Ding, S. J. Pan, and P. S. Yu, "Adaptation regularization: A general framework for transfer learning," *IEEE Trans. Knowl. Data Eng.*, vol. 26, no. 5, pp. 1076–1089, May 2014.
- [6] J. Tahmoresnezhad and S. Hashemi, "Visual domain adaptation via transfer feature learning," *Knowl. Inf. Syst.*, vol. 50, no. 2, pp. 585–605, Feb. 2017.
- [7] N. Courty, R. Flamary, D. Tuia, and A. Rakotomamonjy, "Optimal transport for domain adaptation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 9, pp. 1853–1865, Sep. 2017, doi: [10.1109/TPAMI.2016.2615921](https://doi.org/10.1109/TPAMI.2016.2615921).
- [8] X. Zhong, S. Guo, H. Shan, L. Gao, D. Xue, and N. Zhao, "Feature-based transfer learning based on distribution similarity," *IEEE Access*, vol. 6, pp. 35551–35557, 2018.
- [9] F. Zhuang, Z. Qi, K. Duan, D. Xi, Y. Zhu, H. Zhu, H. Xiong, and Q. He, "A comprehensive survey on transfer learning," 2019, *arXiv:1911.02685*. [Online]. Available: <http://arxiv.org/abs/1911.02685>
- [10] L. Duan, D. Xu, and I. W. Tsang, "Domain adaptation from multiple sources: A domain-dependent regularization approach," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 23, no. 3, pp. 504–518, Mar. 2012.
- [11] L. Duan, I. W. Tsang, and D. Xu, "Domain transfer multiple kernel learning," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 3, pp. 465–479, Mar. 2012.
- [12] C.-X. Ren, D.-Q. Dai, K.-K. Huang, and Z.-R. Lai, "Transfer learning of structured representation for face recognition," *IEEE Trans. Image Process.*, vol. 23, no. 12, pp. 5440–5454, Dec. 2014.
- [13] J. Wang, Y. Chen, S. Hao, W. Feng, and Z. Shen, "Balanced distribution adaptation for transfer learning," in *Proc. IEEE Int. Conf. Data Mining (ICDM)*, New Orleans, LA, USA, Nov. 2017, pp. 1129–1134.
- [14] X. Liu, Z. Liu, G. Wang, Z. Cai, and H. Zhang, "Ensemble transfer learning algorithm," *IEEE Access*, vol. 6, pp. 2389–2396, 2018.
- [15] M. Shao, D. Kit, and Y. Fu, "Generalized transfer subspace learning through low-rank constraint," *Int. J. Comput. Vis.*, vol. 109, nos. 1–2, pp. 74–93, Aug. 2014.
- [16] B. Fernando, T. Tommasi, and T. Tuytelaars, "Joint cross-domain classification and subspace learning for unsupervised adaptation," *Pattern Recognit. Lett.*, vol. 65, pp. 60–66, Nov. 2015.
- [17] H. Sun, S. Liu, and S. Zhou, "Discriminative subspace alignment for unsupervised visual domain adaptation," *Neural Process. Lett.*, vol. 44, no. 3, pp. 779–793, Dec. 2016.
- [18] L. Zhang, S. Wang, G.-B. Huang, W. Zuo, J. Yang, and D. Zhang, "Manifold criterion guided transfer learning via intermediate domain generation," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 30, no. 12, pp. 3759–3773, Dec. 2019.
- [19] L. Zhang, J. Fu, S. Wang, D. Zhang, Z. Dong, and C. L. P. Chen, "Guide subspace learning for unsupervised domain adaptation," *IEEE Trans. Neural Netw. Learn. Syst.*, early access, Nov. 1, 2019, doi: [10.1109/TNNLS.2019.2944455](https://doi.org/10.1109/TNNLS.2019.2944455).
- [20] S. J. Pan, I. W. Tsang, J. T. Kwok, and Q. Yang, "Domain adaptation via transfer component analysis," *IEEE Trans. Neural Netw.*, vol. 22, no. 2, pp. 199–210, Feb. 2011.
- [21] B. Gong, Y. Shi, F. Sha, and K. Grauman, "Geodesic flow kernel for unsupervised domain adaptation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Providence, RI, USA, Jun. 2012, pp. 2066–2073.
- [22] M. Long, J. Wang, G. Ding, J. Sun, and P. S. Yu, "Transfer feature learning with joint distribution adaptation," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Sydney, NSW, Australia, Dec. 2013, pp. 2200–2207.
- [23] B. Fernando, A. Habrard, M. Sebban, and T. Tuytelaars, "Unsupervised visual domain adaptation using subspace alignment," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Sydney, NSW, Australia, Dec. 2013, pp. 2960–2967.
- [24] M. Long, J. Wang, G. Ding, J. Sun, and P. S. Yu, "Transfer joint matching for unsupervised domain adaptation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2014, pp. 1410–1417.
- [25] B. Sun and K. Saenko, "Subspace distribution alignment for unsupervised domain adaptation," in *Proc. Brit. Mach. Vis. Conf.*, Swansea, Wales, 2015, pp. 24.1–24.10.
- [26] Y. Xu, X. Fang, J. Wu, X. Li, and D. Zhang, "Discriminative transfer subspace learning via low-rank and sparse representation," *IEEE Trans. Image Process.*, vol. 25, no. 2, pp. 850–863, Feb. 2016.
- [27] D. Tuia and G. Camps-Valls, "Kernel manifold alignment for domain adaptation," *PLoS ONE*, vol. 11, no. 2, Feb. 2016, Art. no. e0148655.
- [28] M. Ghifary, D. Balduzzi, W. B. Kleijn, and M. Zhang, "Scatter component analysis: A unified framework for domain adaptation and domain generalization," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 7, pp. 1414–1430, Jul. 2017.
- [29] J. Zhang, W. Li, and P. Ogunbona, "Joint geometrical and statistical alignment for visual domain adaptation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Honolulu, HI, USA, Jul. 2017, pp. 5150–5158.
- [30] S. Wang, L. Zhang, W. Zuo, and B. Zhang, "Class-specific reconstruction transfer learning for visual recognition across domains," *IEEE Trans. Image Process.*, vol. 29, pp. 2424–2438, Nov. 2019, doi: [10.1109/TIP.2019.2948480](https://doi.org/10.1109/TIP.2019.2948480).
- [31] M. Ghifary, W. B. Kleijn, and M. Zhang, "Domain adaptive neural networks for object recognition," in *Proc. Pacific Rim Int. Conf. Artif. Intell.*, Gold Coast, QLD, Australia, 2014, pp. 898–904.
- [32] M. Long, Y. Cao, Z. Cao, J. Wang, and M. I. Jordan, "Transferable representation learning with deep adaptation networks," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 41, no. 12, pp. 3071–3085, Dec. 2019.
- [33] C. Chen, Z. Chen, B. Jiang, and X. Jin, "Joint domain alignment and discriminative feature learning for unsupervised deep domain adaptation," in *Proc. AAAI Conf. Artif. Intell.*, Honolulu, HI, USA, 2019, pp. 3296–3303.
- [34] E. Tzeng, J. Hoffman, N. Zhang, K. Saenko, and T. Darrell, "Deep domain confusion: Maximizing for domain invariance," 2014, *arXiv:1412.3474*. [Online]. Available: <http://arxiv.org/abs/1412.3474>
- [35] B. Sun and K. Saenko, "Deep CORAL: Correlation alignment for deep domain adaptation," in *Proc. Eur. Conf. Comput. Vis. Workshops (ECCVW)*, Amsterdam, The Netherlands, 2016, pp. 443–450.
- [36] B. Sun, J. Feng, and K. Saenko, "Return of frustratingly easy domain adaptation," in *Proc. AAAI Conf. Artif. Intell.*, Phoenix, AZ, USA, 2016, pp. 2058–2065.
- [37] E. Tzeng, J. Hoffman, K. Saenko, and T. Darrell, "Adversarial discriminative domain adaptation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Honolulu, HI, USA, Jul. 2017, pp. 2962–2971.
- [38] M. Long, Z. Cao, J. Wang, and M. I. Jordan, "Conditional adversarial domain adaptation," in *Proc. Annu. Conf. Neural Inf. Process. Syst.*, Montreal, QC, Canada, 2018, pp. 1640–1650.
- [39] G. Kang, L. Jiang, Y. Yang, and A. G. Hauptmann, "Contrastive adaptation network for unsupervised domain adaptation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Long Beach, CA, USA, Jun. 2019, pp. 4893–4902.
- [40] A. Gretton, K. Borgwardt, M. J. Rasch, B. Scholkopf, and A. J. Smola, "A kernel method for the two-sample problem," 2008, *arXiv:0805.2368*. [Online]. Available: <http://arxiv.org/abs/0805.2368>
- [41] T. Zhang, D. Tao, and J. Yang, "Discriminative locality alignment," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, 2008, pp. 725–738.
- [42] P. Bouboulis, K. Slavakis, and S. Theodoridis, "Adaptive learning in complex reproducing kernel Hilbert spaces employing Wirtinger's subgradients," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 23, no. 3, pp. 425–438, Mar. 2012.
- [43] B. Li, H. Chang, S. Shan, and X. Chen, "Low-resolution face recognition via coupled locality preserving mappings," *IEEE Signal Process. Lett.*, vol. 17, no. 1, pp. 20–23, Jan. 2010.
- [44] J. Jiang, R. Hu, Z. Wang, and Z. Cai, "CDMMA: Coupled discriminant multi-manifold analysis for matching low-resolution face images," *Signal Process.*, vol. 124, pp. 162–172, Jul. 2016.
- [45] H. S. Bhatt, R. Singh, M. Vatsa, and N. K. Ratha, "Improving cross-resolution face matching using ensemble-based co-transfer learning," *IEEE Trans. Image Process.*, vol. 23, no. 12, pp. 5654–5669, Dec. 2014.

[46] F. Nie, D. Xu, I. W.-H. Tsang, and C. Zhang, "Flexible manifold embedding: A framework for semi-supervised and unsupervised dimension reduction," *IEEE Trans. Image Process.*, vol. 19, no. 7, pp. 1921–1932, Jul. 2010.



JUNFENG JING is currently a Professor with the College of Electronics and Information, Xi'an Polytechnic University, Xi'an, China.

His main research interests include artificial intelligence, machine vision, image processing, and pattern recognition.



DONGDONG ZHENG received the B.Eng. degree in automation from Xi'an Polytechnic University, Xi'an, China, in 2018, where he is currently pursuing the M.S. degree with the College of Electronics and Information.

His main research interests include computer vision and low-resolution face recognition.



KAIBING ZHANG received the M.Sc. degree in computer software and theory from Xihua University, Chengdu, China, in 2005, and the Ph.D. degree in pattern recognition and intelligent system from Xidian University, Xi'an, China, in 2012.

He is currently a Distinguished Professor with the School of Electronics and Information, Xi'an Polytechnic University, Xi'an. His main research interests include pattern recognition, computer vision, and image super-resolution reconstruction. In these areas, he has published around 30 technical articles in refereed journals and proceedings, including the *IEEE TRANSACTIONS ON IMAGE PROCESSING*, the *IEEE TRANSACTIONS ON NEURAL NETWORKS AND LEARNING SYSTEMS*, *Signal Processing* (Elsevier), *Neurocomputing*, *CVPR*, and *ICIP*.



JIAN LU is currently an Associate Professor with the School of Electronics and Information, Xi'an Polytechnic University, Xi'an, China. His main research interests include artificial intelligence, machine learning, and pattern recognition.



ZENGGANG XIONG received the Ph.D. degree from the University of Science and Technology Beijing, Beijing, China, in 2009.

He is currently a Professor with the School of Computer and Information Science, Hubei Engineering University, Xiaogan, China. His main research interests include cloud computing, pattern recognition, and computer vision.

...