

文章编号: 1000-5862(2021)04-0390-08

基于低秩交叉重构的领域自适应算法

郭蔚颖¹, 房小兆², 吴宝昌^{3*}, 滕少华¹

(1. 广东工业大学计算机学院, 广东 广州 510006; 2. 广东工业大学自动化学院, 广东 广州 510006;

3. 广东金融学院公共管理学院, 广东 广州 510520)

摘要: 为了解决现有领域在自适应方法中忽略了整个数据域内部结构的信息和源域与目标域之间的差异问题, 提出了一种新的基于低秩交叉重构的领域自适应方法. 通过对源域和目标域的交叉重构来构造新的源域与目标域, 使得同类数据相互交织, 缩短了同类数据之间的距离; 通过对重构矩阵施加低秩约束, 将 2 个域的同类数据对齐, 以此来充分挖掘源域和目标域同类数据之间的内在结构信息, 并利用该结构信息来学习分类器, 从而取得更好的跨域识别效果. 在 5 个公开数据集上的实验结果表明: 该方法具有较高的跨域识别准确率.

关键词: 领域自适应; 交叉重构; 低秩约束; 跨域识别

中图分类号: TP 391 **文献标志码:** A **DOI:** 10.16357/j.cnki.issn1000-5862.2021.04.11

0 引言

随着在线自媒体和内容共享应用的快速发展, 对图像和其他多媒体数据的自动识别和分析的需求越来越大. 然而, 通过获得足够的标记数据来训练机器学习模型是非常昂贵并且效率较低的. 因此, 在利用现有领域中充足的标记样本来促进相关目标领域的模型学习是更经济、更高效的方法. 领域自适应是解决这一类跨域识别问题的一种十分有效的方法^[1-2].

在领域自适应中, 数据通常来自 2 个域, 分别是源域与目标域: (i) 源域, 有着充足标签样本的数据域; (ii) 目标域, 需要完成识别任务, 但缺乏足够标签样本的数据域. 这 2 个域通常共享相同的标签或任务, 但源域与目标域数据具有不同的特征分布或来自不同的特征空间. 因此, 实现领域自适应的关键是缩小源域和目标域之间的分布差异. 现有的领域自适应算法主要可以分为 2 大类: (i) 实例重加权方法, 通过加权技术重用来自源域的分类器以辅助目标域样本分类; (ii) 特征匹配方法, 通过将源域和目标域的特征投影到一个子空间中, 并在子空间中进

行不同域数据的对齐, 以此来减小 2 个域之间的分布差异. 然而, 现有的领域自适应方法仍存在以下不足: (i) 在特征学习的过程中, 这些方法忽略了数据主要信息的保留; (ii) 这些方法通常使用源域的总数据来重构目标域的总数据, 并且没有重视相同标签类数据的对齐; (iii) 若没有足够的代表性数据, 则这些方法的性能就会较差^[3-4].

本文的基于低秩交叉重构的领域自适应算法 (LRCR) 能够有效地解决以上问题. 它利用原始源域和目标域的交叉重构来实现提取数据潜在相似性, 再通过对重构矩阵施加低秩约束使得来自不同数据域的相同类样本对齐, 从而充分挖掘源域和目标域之间的潜在联系. LRCR 的主要贡献如下: (i) 结合改进的类 PCA 正则化项, 使最终特征表示能保留原始数据的主要信息; (ii) 利用来自不同域但标签相同的数据之间的相互重构, 使得同类数据做到局部对齐; (iii) 挖掘数据域的整体信息来完成识别任务. 这样, 最终得到的特征表示具有较强的鉴别能力, 无论来自不同域的数据是否有相同的特征分布, 只要它们共享相同的标签, 它们就都能被较好地对齐. 因此, 本文提出的方法能较好地完成跨域识别任务.

收稿日期: 2021-02-17

基金项目: 国家自然科学基金(61772141, U1911401, 61972102), 广东省科技计划(2019B020208001, 2019B110210002), 广东省重点领域研发计划(2019B010121001, 2019B010118001, 2019B010119001) 和广州市科技计划(201903010107) 资助项目.

通信作者: 吴宝昌(1978—), 男, 广东廉江人, 讲师, 主要从事电子政务、大数据与公共治理的研究. E-mail: 8426008@qq.com

1 相关工作

1.1 低秩表征

低秩表征(low-rank representation, LRR)^[5]方法假设数据是近似从多个低维子空间中近似采样的,若样本数据 $X \in \mathbf{R}^{m \times n}$ (m 和 n 分别是样本的数据维度和数量) 是从多个子空间的联合空间 $U_i^k S_i$ (S_i 表示低维子空间) 中选出来的,则低秩表征方法能够找到所有样本的最低秩表示来恢复数据. LRR 的目标函数为

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

其中矩阵 A 的列是一系列已知的基向量或字典项. 通常,将 X 本身作为字典, E 表示噪声项, $\text{rank}(\cdot)$ 表示矩阵的秩, $\|\cdot\|_0$ 表示矩阵的 l_0 范数. 求解式(1)是十分困难的,所以 LRR 的目标函数被改为

$$\min_{Z, E} \|Z\|_* + \lambda \|E\|_1, \text{ s. t. } X = XZ + E, \quad (2)$$

其中 $\|\cdot\|_*$ 表示矩阵的核范数,它的值是矩阵所有奇异值之和, $\|\cdot\|_1$ 表示矩阵的 l_1 范数,它的值为在矩阵中每个元素的绝对值之和.

当得到式(2)的低秩解(Z^*, E^*)后,通过 XZ^* 或 $X - E^*$ 来恢复数据 X 的主要成分^[6].

1.2 潜在低秩表征

潜在低秩表征(Latent low-rank representation, LatLRR)^[7]是一种基于低秩表征的子空间学习方法,它能利用大量未观测样本来更好地表示原样本,它的原始目标函数为

$$\min_Z \|Z\|_* \text{ s. t. } X = (X_O, X_H)Z, \quad (3)$$

其中 X_O 代表能够观察到的数据, X_H 代表不能被直接观察到的数据,通过贝叶斯引理^[5]可知 X 可以被表示为 $X = XZ + CX$, 则式(3)的目标函数为

$$\min_{Z, E} \|Z\|_* + \lambda \|C\|_1 \text{ s. t. } X = XZ + CX, \quad (4)$$

其中 Z 是低秩约束的重构矩阵, C 是投影矩阵,为了减少噪声的影响,在 LatLRR 中引入了噪声矩阵 E , 最终的目标函数为

$$\min_{Z, C, E} \|Z\|_* + \|C\|_1 + \lambda \|E\|_1, \text{ s. t. } X = XZ + CX + E. \quad (5)$$

2 基于低秩交叉重构的领域自适应方法

正如在上面所提到的, LatLRR 能较好地提取出主要特征,但是经过 LatLRR 所提取特征的维度是固定的,不利于提取灵活的特征. 为了解决这个问

题,本文提出用 2 个矩阵 PQ^T 来代替单一的投影矩阵 C ^[8], 即式(5)的目标函数可以被改写为

$$\min_{P, Q, Z, E} \|Z\|_* + \lambda_1 \|Q\|_F^2 + \lambda_2 \|E\|_1, \text{ s. t. } X = XZ + PQ^T X + E, \quad (6)$$

其中 $\|\cdot\|_F$ 表示矩阵的 Frobenius 范数,其值为在矩阵中每个元素的平方和再开平方的值. $P \in \mathbf{R}^{m \times n}$, $Q \in \mathbf{R}^{m \times d}$ ($d \ll m$), 因此 $C = PQ^T$ 的秩为 d . λ_1 和 λ_2 为惩罚系数. 通过投影矩阵 Q , 可以提取出样本的显著特征 $Q^T X$, 且 $Q^T X$ 的秩为 d . 为了避免问题的平凡解,对矩阵 P 施加正交约束, 则式(6)可以改写为

$$\min_{P, Q, Z, E} \|Z\|_* + \lambda_1 \|Q\|_F^2 + \lambda_2 \|E\|_1, \text{ s. t. } X = XZ + PQ^T X + E, P^T P = I. \quad (7)$$

从式(7)可以发现,若忽略 XZ 这一项,则剩下的约束就变成了 $X = PQ^T X + E$, 它是一个类 PCA 项,这确保 $Q^T X$ 能够保存数据的主要信息,以保证更好的识别结果^[9]. 这样 $PQ^T X$ 和 XZ 分别可以恢复数据的显著特征信息和主要特征信息,已经证明投影矩阵能够较好地提取数据的主要细节信息,所以特征表示 $Q^T X$ 能较好地用来进行样本分类^[10].

在传统的领域自适应方法中,一般把源域和目标域投影到一个子空间中,在这个子空间中,它们的分布差异较小^[11]. 然而,这些方法都没有充分地利用原始数据,对数据潜在的关系挖掘不够充分. 本文采取了一种新的数据重构方法来使得数据能够被有效利用. X_s 和 X_t 分别表示原始的源域和目标域,本文构造 2 个新的数据域 X 和 Y , 分别从原始数据域 X_s 和 X_t 的每一类中随机抽取相同数目的样本数据,然后交叉混合作为新域的每一类数据. 若 $X_s \in \mathbf{R}^{m \times n_s}$, 则 $X_t \in \mathbf{R}^{m \times n_t}$ (其中 m 表示样本数据的维度, n_s 和 n_t 分别表示在源域和目标域中样本的数量). $X_s^i \in \mathbf{R}^{m \times n}$ 和 $X_t^i \in \mathbf{R}^{m \times n}$ 分别表示从原始源域 X_s 和 X_t 的第 i 个类中抽取出的样本,则新构造的目标域可以被表示为

$$X = (X_s^1 X_t^1, X_s^2 X_t^2, \dots, X_s^c X_t^c) X \in \mathbf{R}^{m \times 2nc}, \\ Y = (X_t^1 X_s^1, X_t^2 X_s^2, \dots, X_t^c X_s^c) X \in \mathbf{R}^{m \times 2nc}.$$

例如,在 USPS 和 MNIST 数据集中包含 10 个类,在实验时从原始源域和目标域每个类中抽取 5 个样本构成新域的一个包含 10 个样本特征类. L 表示所构建的 2 个域的标签(它们的标签是一样的). 则 LRCR 的目标函数为

$$\min_{P, Q, Z, E, W} \|L - W^T Q^T Y\|_F^2 + \lambda_1 \|Z\|_* + \lambda_2 \|E\|_1 + \lambda_3 \|W^T Q^T\|_F^2, \text{ s. t. } X = YZ + PQ^T Y + E, P^T P = I, \quad (8)$$

其中 $\|\cdot\|_F$ 表示矩阵的 Frobenius 范数, $\|\cdot\|_*$ 表示矩阵的核范数, $\|\cdot\|_{2,1}$ 表示矩阵的 $2,1$ 范数, 它的值是矩阵每一行的 l_2 范数之和^[12]. $\lambda_1, \lambda_2, \lambda_3 \geq 0$ 是惩罚系数, 用于调整每一项的权重.

由于凸优化理论可知, 本文的目标函数整体是非凸的, 但是每个变量 P, Q, Z, E, W 的优化均为凸优化问题^[13]. 因此, 本文使用迭代更新方式来优化目标函数(8). 为方便优化, 引入辅助变量 A, H 替代 $W^T Q^T, Z$, 将式(8)中的目标函数重写为

$$\min_{P, Q, Z, E, W, A, H} \|L - AY\| + \lambda_1 \|H\|_* + \lambda_2 \|E\|_{2,1} + \lambda_3 \|A\|_F^2, \text{ s. t. } X = YZ + PQ^T Y + E, P^T P = I, Z = H, W^T Q^T = A. \quad (9)$$

式(9)的增广拉格朗日函数为

$$F(P, Q, Z, E, W, A, H) = \|L - AY\|_F^2 + \lambda_1 \|H\|_* + \lambda_2 \|E\|_{2,1} + \lambda_3 \|A\|_F^2 + \langle \tau_1, X - YZ - PQ^T Y - E \rangle + \langle \tau_2, Z - H \rangle + \langle \tau_3, W^T Q^T - A \rangle + \mu (\|X - YZ - PQ^T Y - E\|_F^2 + \|Z - H\|_F^2 + \|W^T Q^T - A\|_F^2) / 2, \text{ s. t. } P^T P = I, \quad (10)$$

其中 τ_1, τ_2, τ_3 是拉格朗日乘子, $\mu > 0$ 是惩罚系数. 变量通过交替方向乘子法 (Alternating Direction Method of Multipliers, ADMM)^[14] 更新, 在更新其中一个系数时会保证其他的系数固定不变. 迭代优化步骤如下:

步骤 1(更新 P) 固定 Q, Z, E, W, A, H , 并求解

$$\min_P \mu \|X - YZ - PQ^T Y - E + \tau_1/\mu\|_F^2 / 2. \quad (11)$$

这是一个典型的正交普鲁克问题 (orthogonal ProCrustes problem)^[15], 对 $(X - YZ - E + \tau_1/\mu) Y^T Q$ 进行奇异值分解, 令 $(X - YZ - E + \tau_1/\mu) Y^T Q = USV^T$, 则可以得到 P 的解为

$$P = UV^T.$$

步骤 2(更新 Q) 固定 P, Z, E, W, A, H , 并求解

$$\min_Q \mu (\|X - YZ - PQ^T Y - E + \tau_1/\mu\|_F^2 + \|W^T Q^T - A + \tau_3/\mu\|_F^2) / 2. \quad (12)$$

令 $F(Q) = \mu (\|X - YZ - PQ^T Y - E + \tau_1/\mu\|_F^2 + \|W^T Q^T - A + \tau_3/\mu\|_F^2) / 2$, 可以通过求偏导 $\partial F(Q) / \partial Q = 0$ 得到 Q 的解, 此时, 要解决一个李亚普洛夫问题^[16], 即

$$YY^T Q + QWW^T = Y(X - YZ - E + \tau_1/\mu)^T P - (\tau_3/\mu - A)^T - W^T.$$

步骤 3(更新 Z) 固定 P, Q, E, W, A, H , 并求解

$$\min_Q \mu (\|X - YZ - PQ^T Y - E + \tau_1/\mu\|_F^2 + \|Z - H + \tau_2/\mu\|_F^2) / 2. \quad (13)$$

令 $F(Z) = \min_Q \mu (\|X - YZ - PQ^T Y - E + \tau_1/\mu\|_F^2 + \|Z - H + \tau_2/\mu\|_F^2) / 2$, 通过求 $\partial F(Z) / \partial Z$ 得到 Z 的解为

$$Z = (Y^T Y + E)^{-1} (Y^T (X - PQ^T Y - E + \tau_1/\mu) - \tau_2/\mu + H).$$

步骤 4(更新 E) 固定 P, Q, Z, W, A, H , 并求解

$$\min_E \lambda_2 \|E\|_{2,1} + \mu \|X - YZ - PQ^T Y - E + \tau_1/\mu\|_F^2 / 2. \quad (14)$$

对于式(14), 可以根据以下引理来求解.

引理 1^[17] 对于问题

$$\min_W \alpha \|W\|_{2,1} + \|W - Q\|_F^2 / 2,$$

其中 Q 是一个已知矩阵, 若该问题的最优解是 W^* , 则 W^* 的第 i 列的值为

$$(W^*)_i = \begin{cases} \frac{\|Q_i\|_2 - \alpha}{\|Q_i\|_2} Q_i, & \|Q_i\|_2 > \alpha, \\ 0, & \text{其他情况} \end{cases}$$

用 τ_2/μ 替换其中的 α , 用 $X - YZ - PQ^T Y + \tau_1/\mu$ 替换其中的 Q 就能够得到 E 的解.

步骤 5(更新 W) 固定 P, Q, Z, E, A, H , 通过求解

$$\min_W \mu \|W^T Q^T - A + \tau_3/\mu\|_F^2 / 2. \quad (15)$$

令 $F(W) = \mu \|W^T Q^T - A + \tau_3/\mu\|_F^2 / 2$, 当求偏导 $\partial F(W) / \partial W = 0$ 时, 能够得到 W 的解为

$$W = Q^{-1} (A - \tau_3/\mu)^T.$$

步骤 6(更新 A) 固定 P, Q, Z, E, W, H , 通过求解

$$\min_A \|L - AY\|_F^2 + \lambda_3 \|A\|_F^2 + \mu \|W^T Q^T - A + \tau_3/\mu\|_F^2 / 2. \quad (16)$$

令 $F(A) = \|L - AY\|_F^2 + \lambda_3 \|A\|_F^2 + \mu \|W^T Q^T - A + \tau_3/\mu\|_F^2 / 2$. 通过令 $\partial F(A) / \partial A = 0$, 可以得到 A 的解为

$$A = (2LY^T + \mu(W^T Q^T + \tau_3/\mu))(2YY^T + (2\lambda_3 + \mu)E)^{-1}.$$

步骤 7(更新 H) 固定 P, Q, Z, E, W, A , 可以通过求解

$$\min_H \lambda_1 \|H\|_* + \mu \|Z - H + \tau_2/\mu\|_F^2 / 2. \quad (17)$$

可以得出 $H = \Theta_{\lambda_1/\mu}(Z + \tau_2/\mu)$.

算法 1 LRCR.

输入: 训练样本的特征矩阵 $X \in \mathbf{R}^{m \times 2nc}$, 测试样

本特征矩阵 $Y \in \mathbf{R}^{m \times 2nc}$, 训练样本标签矩阵 $L \in \mathbf{R}^{c \times 2nc}$, 超参数 $\lambda_1, \lambda_2, \lambda_3$.

初始化: $W = I(c)$ $Q = I(E) = I(H) = I(A) = W^T Q^T$. 其中 I 表示一个合适大小的矩阵, 它的元素都是 1. $P = \arg \min_W \text{tr}(P^T(-\Sigma)P)$ s. t. $P = I \Sigma$ 表示数据的协方差矩阵. $\lambda_1 = z_{eros}(m, 2nc)$ $\lambda_2 = z_{eros}(2nc, 2nc)$ $\lambda_3 = z_{eros}(c, m)$ $\mu = 0.1$ $\rho = 1.01$, $\mu_{\max} = 10^5$.

迭代更新:

(i) 通过式(11)更新 P ;

(ii) 通过式(12)更新 Q ;

(iii) 通过式(13)更新 Z ;

(iv) 通过式(14)更新 E ;

(v) 通过式(15)更新 W ;

(vi) 通过式(16)更新 A ;

(vii) 通过式(17)更新 H ;

(viii) 通过下式更新 τ_1, τ_2, τ_3 和 μ :

$$\begin{cases} \tau_1 = \tau_1 + \mu(X - PQ^T Y - YZ - E), \\ \tau_2 = \tau_2 + \mu(Z - H), \\ \tau_3 = \tau_3 + \mu(W^T Q^T - A), \\ \mu = \min(\rho\mu, \mu_{\max}); \end{cases}$$

(xi) 直至收敛.

输出: 分类矩阵 $A \in \mathbf{R}^{c \times m}$.

3 实验

为了验证基于低秩约束的交叉重构领域适应方法(LRCR)的有效性, 将 LRCR 在 COIL20、MNIST & USPS、MSRC & VOC2007、Office & Caltech 和 Office-Home 等 5 个基准数据集上分别进行实验. 从这些数据集的源域和目标域的每一类中随机抽取 5 个样本来构造新的数据域, 剩下的数据作为测试样本使用. 通过 LRCR 得到分类矩阵 A , 使用 AX 来作为最后的分类标签. 若第 i 个测试样本的特征向量为 $x_i \in \mathbf{R}^m$, 通过计算得到它的标签向量 $l_i = (l_i^1, l_i^2, \dots, l_i^c) = Ax_i \in \mathbf{R}^c$, 则在标签向量 l_i 中最大值所处位置为该样本所属标签类. 若到了第 i 个测试样本的标签向量 $l_i = (l_i^1, l_i^2, \dots, l_i^c) \in \mathbf{R}^c$, 并且 $\max\{l_i^1, l_i^2, \dots, l_i^c\} = l_i^k$, 其中 $1 \leq k \leq c$, 则把这个样本划分到第 k 类数据中. 为了保证实验的准确性, 选取了 20 次实验的平均值作为最终的实验结果. 实验在 Matlab2019b, Intel(R) Core(TM) i7-6700 CPU@3.40GHz 环境下进行.

3.1 数据集介绍

COIL 20 数据集: 该数据集包含 20 个不同对象

以 360° 旋转成像. 每旋转 5° 收集 1 幅物体的图像, 即每个物体有 72 幅图像, 均为不同角度, 共 1 440 幅图像. 所有的图像裁剪并转换为 32×32 像素的灰度图像^[18].

MSRC & VOC2007 数据集: MSRC 包含了 18 个层次的 4 323 幅图像, 而 VOC 2007 包含了 20 个概念的 5 011 幅图像, 它们分别共享了飞机、羊、汽车、牛、鸟、自行车等 6 个语义类, 并且将所有图像调整为 256 像素^[19].

MNIST & USPS 数据集: 该数据集有 7 291 幅训练图像和 2 007 幅测试图像; MNIST 数据集有 60 000 幅训练图像和 10 000 幅测试图像, 这 2 个数据集共有 10 个语义类, 每个语义类对应数字 0~9, 所有的图像都被转换为 16×16 像素的灰度图像^[20].

Office & Caltech256 数据集: Office 数据集是视觉对象识别的基准数据, 包括来自 3 个不同领域的常见对象类别, 即 A (Amazon)、W (Webcam) 和 D (DSLR), 每个领域共有 31 个对象类别. 例如笔记本电脑、键盘、显示器、自行车等, 一共含有 4 652 幅图像. 在 Amazon 域中, 每个类别平均有 90 幅图像, 而在 DSLR 或 Webcam 域中, 每个类别平均有 30 幅图像. Caltech-256 数据集是用于目标识别的标准数据集, 有 30 607 幅图像和 256 个类别^[21].

Office-Home 数据集: 该数据集由来自 4 个不同领域的图像组成, 它们是艺术图像、剪辑艺术、产品图像、现实世界图像, 对于每个领域, 数据集包含 65 个对象类别的图像, 通常在办公室和家庭中发现. 在实验中的数据特征是由预先训练好的 ResNet50 模型提取得到的^[22].

3.2 评估度量以及超参数设置

如式(8)所示, 本文的方法包含了 3 个超参数, 它们分别是 $\lambda_1, \lambda_2, \lambda_3$, 它们的取值都属于集合 $\{1 \times 10^{-8}, 1 \times 10^{-7}, \dots, 1\}$, 超参数的最后取值通过网格搜索策略确定. 对于数据集 COIL 20, 最终参数选择为 $\lambda_1 = 1 \times 10^{-6}$, $\lambda_2 = 1 \times 10^{-5}$, $\lambda_3 = 1 \times 10^{-1}$; 对于数据集 MSRC & VOC2007, 最终参数选择为 $\lambda_1 = 1 \times 10^{-8}$, $\lambda_2 = 1 \times 10^{-2}$, $\lambda_3 = 1$; 对于数据集 MNIST & USPS, 最终参数选择为 $\lambda_1 = 1 \times 10^{-1}$, $\lambda_2 = 1 \times 10^{-5}$, $\lambda_3 = 1$; 对于数据集 Office & Caltech, 最终参数的选择为 $\lambda_1 = 1 \times 10^{-2}$, $\lambda_2 = 1 \times 10^{-6}$, $\lambda_3 = 1$; 对于 Office-Home 数据集, 最终参数选择为 $\lambda_1 = 1 \times 10^{-6}$, $\lambda_2 = 1 \times 10^{-6}$, $\lambda_3 = 1 \times 10^{-7}$.

3.3 实验结果和分析

为了验证 LRCR 的有效性,选取了几个迁移学习和领域适应的方法做对比,它们分别是 geodesic flow kernel (GFK)^[22]、low-rank transfer subspace learning (LTSL)^[23]、fisher discrimination dictionary learning (FDDL)^[24]、joint geometrical and statistical alignment (JGSA)^[25]、weakly-supervised cross-domain dictionary learning for visual recognition (WSCDDL)^[26]、

visual domain adaptation with manifold embedded distribution alignment (MEDA)^[27].

3.3.1 在几个公开数据集上的实验结果 表 1 展示了 LRCR 与对比算法在 COIL20、MSRC & VOC2007、MNIST & USPS 这 3 个数据集上的识别准确率.表 2 展示了 LRCR 与对比算法在 Office + Caltech256 数据集上的识别准确率.表 3 展示了 LRCR 与对比算法在 Office-Home 数据集上的识别准确率.

表 1 在 3 个不同数据集上的实验结果

数据集	GFK	LTSL	FDDL	JGSA	WSCDDL	MEDA	LRCR
COIL1→COIL2	91.02	36.38	41.93	91.83	84.90	90.00	94.68
COIL2→COIL1	90.38	39.27	39.97	90.39	85.60	90.83	86.61
MSRC→VOC	29.04	24.62	29.97	29.98	30.40	36.08	35.01
VOC→MSRC	58.11	46.63	60.98	60.55	64.52	54.85	74.05
MNIST→USPS	72.03	36.32	75.85	72.90	74.05	39.94	80.40
USPS→MNIST	63.04	39.13	60.54	62.97	62.88	45.40	66.50
平均值	67.33	37.06	51.54	63.73	66.97	59.52	72.88

表 2 在 Office + Caltech 数据集上的实验结果

数据集	GFK	LTSL	FDDL	JGSA	WSCDDL	MEDA	LRCR
A→C	36.97	34.99	37.65	37.65	38.97	43.99	44.25
A→D	53.63	38.49	51.23	55.76	53.24	45.86	56.49
A→W	59.34	39.58	59.79	59.98	58.83	53.22	63.72
D→A	45.78	42.41	45.64	47.10	46.51	41.23	52.33
D→C	33.64	34.94	36.20	34.96	34.43	34.91	39.79
D→W	79.12	70.12	78.60	79.69	76.60	87.46	80.70
C→A	46.70	40.29	48.70	48.90	50.29	56.58	55.30
C→D	57.43	40.49	58.03	58.29	56.62	50.32	58.44
C→W	57.16	42.01	62.97	57.87	68.32	53.90	73.02
W→A	45.11	44.16	45.78	47.89	47.45	42.69	51.03
W→C	32.50	36.44	34.79	35.60	36.44	34.28	39.36
W→D	67.22	69.09	69.78	70.25	62.33	88.54	75.97
平均值	51.22	43.75	52.41	52.93	52.68	52.75	57.53

表 3 在 Office-Home 数据集上的实验结果

数据集	GFK	LTSL	FDDL	JGSA	WSCDDL	MEDA	LRCR
$A_r \rightarrow C_l$	39.74	33.30	43.89	44.51	45.56	39.62	49.51
$A_r \rightarrow P_r$	61.83	54.01	67.40	65.86	67.62	57.09	69.81
$A_r \rightarrow R_w$	61.28	36.25	65.50	65.11	65.48	65.43	69.46
$C_l \rightarrow A_r$	33.63	21.56	42.70	41.00	43.76	39.28	47.73
$C_l \rightarrow P_r$	54.16	42.52	57.63	56.31	59.72	50.45	62.38
$C_l \rightarrow R_w$	49.67	31.04	55.64	53.40	54.28	52.10	57.75
$P_r \rightarrow A_r$	35.60	32.48	47.80	46.23	48.10	42.09	50.48
$P_r \rightarrow C_l$	36.42	32.08	42.89	43.74	43.74	31.83	46.25
$P_r \rightarrow R_w$	60.46	40.50	63.64	64.46	65.82	63.92	68.94
$R_w \rightarrow A_r$	49.20	34.23	59.86	58.22	60.07	55.43	60.40
$R_w \rightarrow C_l$	41.56	39.99	48.89	47.23	49.27	41.27	50.70
$R_w \rightarrow P_r$	69.56	54.63	72.65	72.37	73.42	70.44	75.32
平均值	49.42	37.71	56.12	54.87	56.40	50.75	59.06

3.3.2 实验结果分析 对于上述实验结果,进行如下分析:

(i) LRCR 方法在大多数数据集上的性能都优

于与之比较的方法.这表明:使用改进的类 PCA 正则化项,LRCR 能较好地保存数据信息.此外,通过对重构矩阵施加低秩约束,可以将来自不同领域但

共享同一标签的数据较好地对齐,这就保证了新的特征表示的识别性,因此,该方法可以显著提高识别准确率。

(ii) 在高维数据集 Office-Home 上的实验结果表明 LRCR 在大多数高维数据集上表现良好,结果表明该方法能够较好地处理高维度数据。

(iii) 在半监督领域适应中,人工添加伪标签的对抗式方法是很常见的做法,本文使用原始数据重构代替伪标签来提高分类精度,通过实验发现 LRCR 的性能比 MEDA 更好,这表明了该方法的有效性。

3.3.3 参数敏感性分析 在本文的 LRCR 方法中存在 3 个超参数,在图 1 中,分别绘制了在不同数据集中,随着 λ_1 、 λ_2 、 λ_3 变化对识别准确率的影响。在

图 2 中,绘制了这 3 个超参数在 Office-Home 中对 6 个不同数据域的识别效果造成的影响。

从图 1 和图 2 可以发现:在 LRCR 中,随着 λ_1 的变化,识别准确率的变化不大,这是因为 λ_1 控制的是低秩矩阵的权重,而低秩矩阵本身就较小, λ_1 使得它更小,所以对识别准确率的影响不大。随着 λ_2 的增大,识别准确率会有一个先减小再增大再减小的趋势,这是因为 λ_2 控制的是噪声矩阵所占的比例,随着它的变化,分类器的比例会相应地变化,所以对识别准确率的影响是波动的。 λ_3 控制的是分类器权重,从图 1 和图 2 可知:对于不同的数据集,随着 λ_3 的增长,准确率的变化是不同的,这说明分类器矩阵和数据集是十分相关的。

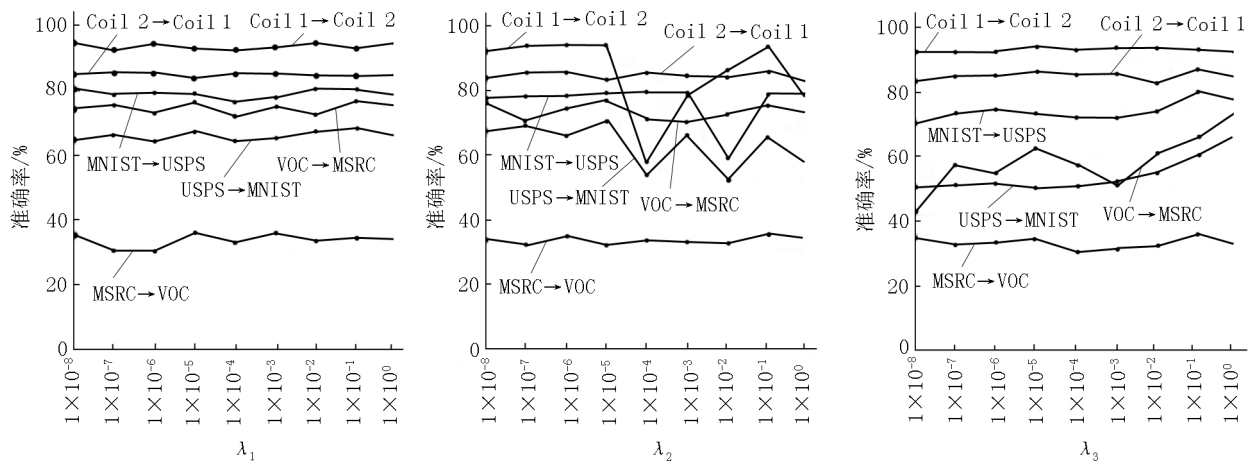


图 1 对 COIL20、MSRC & VOC2007、MNIST & USPS 识别准确率的影响

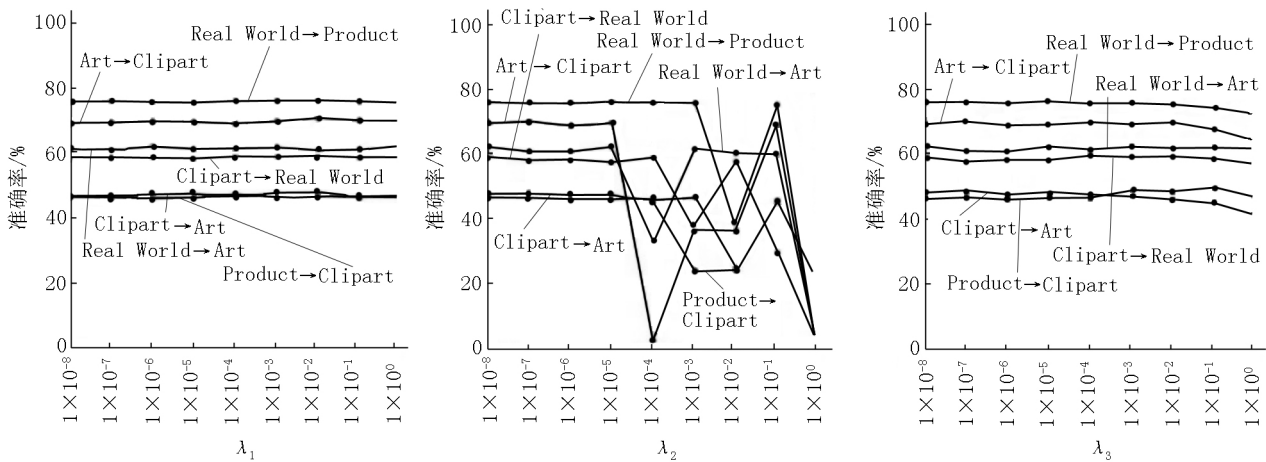


图 2 对在 Office-Home 中 6 个不同任务识别准确率的影响

3.3.4 模型收敛性分析 当 ADMM 只有 2 个或更少块时的收敛性已经被证明,算法 1 有 7 个分块,没有严格的理论收敛性支持,然而,一些理论可以扩展 ADMM 的收敛性的范围,例如, Jia Zehui 等^[28]在单个函数是强凸函数和线性函数复合的前提下证明了具有 2 个块的可分离凸规划的经典 ADMM 可以推广到 3 个或 3 个以上的块, Hong Mingyi 等^[29]证明了在增广拉格朗日函数中惩罚参数足够大的情况

下,经典 ADMM 收敛于平稳解集, Hong Mingyi 等^[30]指出:当线性约束的非光滑凸可分离函数的和最小时, ADMM 与 n ($n > 2$) 变量块也可以收敛。

图 3 是 LRCR 在 2 个数据集上的目标函数值随着迭代次数的结果,从图 3 可以看出随着迭代次数的增加,模型的目标函数值快速地下降,并在 5 ~ 10 次迭代后趋于平稳,实验结果表明本文提出的方法具有良好的收敛性。

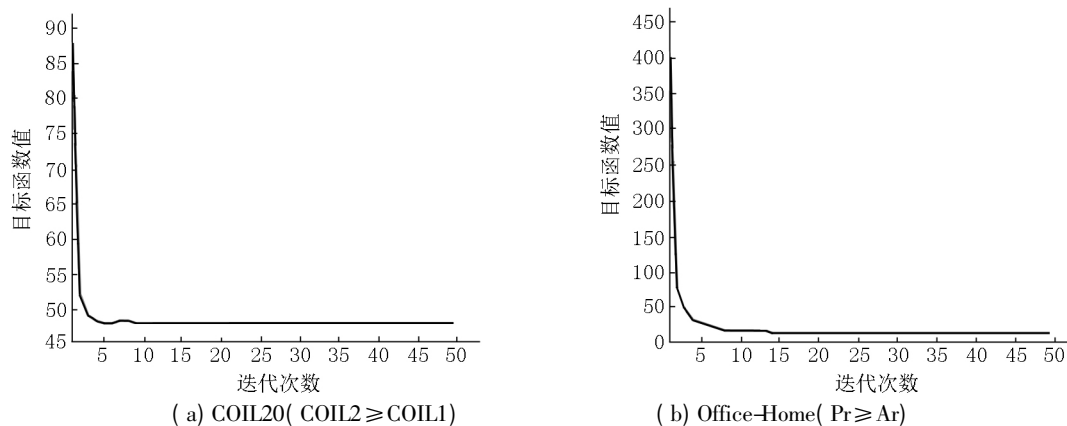


图3 模型收敛性分析

4 结论

本文提出了一种新颖的基于低秩约束的交叉重构领域自适应算法用于跨域识别. 该算法通过将源域和目标域数据进行交叉重构来使原始源域和目标域做到相互表示, 同时对重构矩阵施加低秩约束来对齐同类数据, 从而挖掘同类数据之间的相似性, 以最大程度保留数据主要信息. 最后, 对分类矩阵施加稀疏约束以去除数据冗余信息, 达到提高模型性能的目的. 在 5 个常用的数据集 COIL20、MNIST & USPS、MSRC&VOC2007、Office & Caltech、Office-Home 上验证模型的有效性. 实验结果显示: 与其他相关的传统领域自适应方法相比, LRRCR 算法能更好地保留投影数据信息, 取得更好的跨域识别效果.

5 参考文献

- [1] Weiss K, Khoshgoftaar T M, Wang Dingding. A survey of transfer learning [J]. Journal of Big Data, 2016, 3(1): 1-40.
- [2] Dai Wenyuan, Yang Qiang, Xue Guirong, et al. Boosting for transfer learning [EB/OL]. [2020-11-23]. <https://doi.org/10.1145/1273496.1273521>.
- [3] Pan Sinno Jialin, Yang Qiang. A survey on transfer learning [J]. IEEE Transactions on Knowledge and Data Engineering, 2010, 22(10): 1345-1359.
- [4] 庄福振, 罗平, 何清, 等. 迁移学习研究进展 [J]. 软件学报, 2015, 26(1): 26-39.
- [5] Lin Zhouchen, Zhang Chao, Zhou Pan. Integrated low-rank-based discriminative feature learning for recognition [J]. IEEE Transactions on Neural Networks and Learning Systems, 2016, 27(5): 1080-1093.
- [6] Liu Guangcan, Lin Zhouchen, Yan Shuicheng, et al. Robust recovery of subspace structures by low-rank representation [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2013, 35(1): 171-184.
- [7] Liu Guangcan, Yan Shuicheng. Latent low-rank representation for subspace segmentation and feature extraction [EB/OL]. [2020-06-17]. <https://dl.acm.org/doi/10.1109/ICCV.2011.6126422>.
- [8] Fang Xiaozhao, Han Na, Wu Jigang, et al. Approximate low-rank projection learning for feature extraction [J]. IEEE Transactions on Neural Networks and Learning Systems, 2018, 29(11): 5228-5241.
- [9] 靳洋. 基于 PCA/ICA 的图像特征提取算法研究 [D]. 西安: 西安电子科技大学, 2014.
- [10] Hervé Abdi, Williams L J. Principal component analysis [J]. Wiley Interdisciplinary Reviews Computational Statistics, 2010, 2(4): 433-459.
- [11] Pan Sinno Jialin, Tsang I W, Kwok J T, et al. Domain adaptation via transfer component analysis [J]. IEEE Transactions on Neural Networks, 2011, 22(2): 199-210.
- [12] 屈磊, 方怡, 熊友玲, 等. 基于 L2, L1 模和图正则化的低秩迁移子空间学习 [J]. 控制理论与应用, 2018, 35(12): 1738-1749.
- [13] Georgiades A S, Belhumeur P N, Kriegman D J. From few to many: illumination cone models for face recognition under variable lighting and pose [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2002, 23(6): 643-660.
- [14] 吴越, 曾向荣, 周典乐, 等. 图像恢复中的稳健交替方向乘子法 [J]. 国防科技大学学报, 2018, 40(2): 115-121.
- [15] Tai Ying, Yang Jian, Zhang Yigong, et al. Face recognition with pose variations and misalignment via orthogonal procrustes regression [J]. IEEE Transactions on Image Processing, 2016, 25(6): 2673-2683.
- [16] Fridman E. Descriptor discretized Lyapunov functional method [J]. IFAC Proceedings Volumes, 2006, 39(9): 638-643.
- [17] Yang Junfeng, Yin Wotao, Zhang Yin, et al. A fast algorithm for edge-preserving variational multichannel image restoration [J]. SIAM Journal on Imaging Sciences, 2009, 2(2): 569-592.
- [18] Rate C, Retrieval C. Columbia object image library (COIL-

- 20] [EB/OL]. [2020-06-17]. <https://www.igi-global.com/pdf.aspx?tid=104931&ptid=46172&ctid=17&t=index>.
- [19] Galleguillos C, Mcfee B, Belongie S, et al. From region similarity to category discovery [EB/OL]. [2020-06-17]. <https://dl.acm.org/doi/10.1109/CVPR.2011.5995527>.
- [20] Fernando B, Tommasi T, Tuytelaars T. Joint cross-domain classification and subspace learning for unsupervised adaptation [J]. *Pattern Recognition Letters* 2015 65(1): 60-66.
- [21] Liang Jian, He Ran, Sun Zhenan, et al. Aggregating randomized clustering-promoting invariant projections for domain adaptation [EB/OL]. [2020-12-11]. <https://ieeexplore.ieee.org/document/8353356>.
- [22] Gong Boqing, Shi Yuan, Sha Fei, et al. Geodesic flow kernel for unsupervised domain adaptation [EB/OL]. [2020-06-17]. <https://ieeexplore.ieee.org/document/6247911>.
- [23] Shao Ming, Kit D, Fu Yun. Generalized transfer subspace learning through low-rank constraint [J]. *International Journal of Computer Vision* 2014 109(1/2): 74-93.
- [24] Yang Meng, Zhang Lei, Feng Xiangchu, et al. Sparse representation based fisher discrimination dictionary learning for image classification [J]. *International Journal of Computer Vision* 2014 109(3): 209-232.
- [25] Zhang Jing, Li Wanqing, Ogunbona P. Joint geometrical and statistical alignment for visual domain adaptation [EB/OL]. [2020-06-17]. <http://www.doc88.com/p-2374903824918.html>.
- [26] Zhu Fan, Shao Ling. Weakly-supervised cross-domain dictionary learning for visual recognition [J]. *International Journal of Computer Vision* 2014 109(1/2): 42-59.
- [27] Wang Jindong, Feng Wenjie, Chen Yiqiang, et al. Visual domain adaptation with manifold embedded distribution alignment [EB/OL]. [2020-06-17]. <https://arxiv.org/abs/1807.07258>.
- [28] Jia Zehui, Guo Ke, Cai Xingju. Convergence analysis of alternating direction method of multipliers for a class of separable convex programming [J]. *Abstract and Applied Analysis* 2013(2): 205-215.
- [29] Hong Mingyi, Luo Zhijuan, Razaviyayn M. Convergence analysis of alternating direction method of multipliers for a family of nonconvex problems [EB/OL]. [2020-06-17]. <https://arxiv.org/abs/1410.1390>.
- [30] Hong Mingyi, Luo Zhijuan. On the linear convergence of the alternating direction method of multipliers [J]. *Mathematical Programming* 2012 162(1/2): 1-29.

The Low-Rank Constraint-Based Cross Reconstruction for Domain Adaptation

GUO Weiying¹, FANG Xiaozhao², WU Baochang^{3*}, TENG Shaohua¹

(1. School of Computers, Guangdong University of Technology, Guangzhou Guangdong 510006, China;

2. School of Automation, Guangdong University of Technology, Guangzhou Guangdong 510006, China;

3. School of Public Administration, Guangdong University of Finance, Guangzhou Guangdong 510520, China)

Abstract: In order to solve these problems, a new low-rank constraint-based cross reconstruction (LRCR) for domain adaptation is proposed. Through the cross reconstruction of the source domain and the target domain, the new source domain and the target domain are constructed so that the same label data are interweaved with each other. By applying the low-rank constraint, the same label data in the two domains are aligned to fully mine the internal structure information between the same label data in the source domain and the target domain. The structure information is used to learn the classifier to achieve better cross-domain recognition accuracy. The experimental results on five public datasets show that LRCR has high cross-domain recognition accuracy.

Key words: domain adaptation; cross reconstruction; low-rank constraint; cross-domain recognition

(责任编辑: 冉小晓)