

DOI:10.11918/202006133

无监督域适应的表示学习算法

许亚雲, 严 华

(四川大学 电子信息学院, 成都 610065)

摘 要: 在传统模式识别的训练方法中, 拥有大量标签的有监督学习方法在识别准确率上取得了很好的效果. 然而在实际生活中样本常常缺失标签, 或现存有标签的样本与目标样本具有较大分布差异而不能直接使用. 为了解决这些问题, 无监督域自适应算法应运而生, 借助源域有标签但不同分布的样本去识别无标签的目的域样本. 针对目标识别样本与训练样本分布不一致的情况, 本文提出了一种探寻两个样本域之间的最优表示学习的无监督域适应算法. 通过在共同的子空间上引入两个表示矩阵去更好地减少两个域的分布差异, 同时对两个表示矩阵进行各自的最优化约束设计, 使得源域和目的域最优地相互表示, 缩小两个域之间的分布差异, 从而实现无监督跨域学习(即迁移学习). 最后在3个迁移学习常见的无监督域适应数据集上开展实验, 实验结果表明: 本文算法的识别准确率超过了目前很多优秀的传统迁移学习方法和一些深度方法. 实验结果验证了本文提出的无监督域适应的表示学习算法的有效性和鲁棒性.

关键词: 模式识别; 迁移学习; 无监督域适应; 表示学习; 分布差异

中图分类号: TP391.4

文献标志码: A

文章编号: 0367-6234(2021)02-0040-07

Representation learning for unsupervised domain adaptation

XU Yayun, YAN Hua

(College of Electronics and Information Engineering, Sichuan University, Chengdu 610065, China)

Abstract: Among the conventional training methods of pattern recognition, the supervised learning methods with abundant labels have achieved significant performance in recognition accuracy. However, in real life, there are problems that samples often lack labels, or existing labeled samples cannot be used directly due to the distributional divergences of the target samples. To resolve these issues, unsupervised domain adaptation is often applied to recognize samples of unlabeled target domain by taking advantage of the data of the source domain with sufficient labels but different distributions. Considering the situation that the distributions of the target recognition samples and the source training samples are different, an optimal representation learning method for unsupervised domain adaption was proposed. Two representation matrices were introduced to the common subspace of the domain samples to better reduce the distributional divergence of the domains. Then, optimization constraints were implemented on the two representation matrices, so as to make the source domain and the target domain optimally represent each other, thereby reducing the distributional divergence between the domains. In this way, the unlabeled target domain samples could be recognized by the fully labeled source domain samples (i.e., transfer learning). Experiments on three common unsupervised domain adaption datasets show that the proposed method outperformed the conventional transfer learning methods and deep learning methods in recognition accuracy, which verifies the validity and robustness of the proposed algorithm.

Keywords: pattern recognition; transfer learning; unsupervised domain adaptation; representation learning; distributional divergence

有监督的模式识别过于依赖有标签的训练样本的数量. 一方面, 在现实生活中为样本打标签十分耗费资源和时间; 另一方面, 这种完全由样本的数量和质量决定的学习方式非常容易导致过拟合. 由于监督学习的局限性和不便性, 大量的研究者开始探索无监督和半监督的学习方法. 如主动学习, 它不是

将所有样本打上标签, 而是提出一些标注请求, 将一些经过筛选的数据提交给相关领域专家进行标注.

域自适应或迁移学习, 因为不需要大量的数据标注, 近年来受到研究者的广泛关注. 在域自适应的问题中, 将需要进行识别的数据集称为测试集. 测试集中的数据全部没有标签时称为无监督域自适应^[1-2], 测试集中有少部分数据有标签时则称为半监督域自适应^[3-4]. 与主动学习不一样的是, 不需要采取人工辅助的方式为样本打上标签, 用于训练做训练集的源域中已经拥有大量有标签的数据, 但是

收稿日期: 2020-06-25

基金项目: 国家自然科学基金(11872069)

作者简介: 许亚雲(1994—), 女, 硕士研究生

通信作者: 严 华, yanhua@scu.edu.cn

源域的数据和真正需要进行分类的目的域数据并不是同分布的. 所以域自适应主要解决两个域分布适应的问题,从而借助源域中大量有标签的数据集对目的域的数据进行识别. 目前传统的域自适应算法的相关研究工作主要分为以下3种方式:1)数据分布的适应:即通过某种变换直接将两个域的分布拉近或者选择出分布相似的公共特征,主要通过最大平均差异(maximum mean discrepancy, MMD)^[5]度量变换后的两个域的相似性,例如 TCA^[6]、TJM^[7]、JDA^[8]和 JGSA^[9]. TCA 首先提出通过一个特征映射使得映射后的两个域的边缘分布接近, JDA 同时考虑了边缘分布适配和条件分布适配. 其他部分方法基于 TCA 和 JDA 做出了扩展,例如 TJM 在 TCA 中加入了源域样本选择, ARTL^[10]将 JDA 嵌入结构风险最小化框架. 2)子空间学习:包括低维统计特征子空间对齐(SA^[11]、SDA^[12]和 RTML^[13])和低维流形结构的子空间对齐(GFK^[14]和 DIP^[15]). 3)表示学习:在子空间学习的基础上通过源域样本表示目的域样本(DTSL^[16]和 LSDT^[17]).

1 域适应表示学习算法

本文的方法基于表示学习,与现有表示学习方法思想相同:都是在子空间学习的基础上引入表示矩阵以更好地减少两个域的分布差异. 但是现有方法通常只采用一个单一的表示矩阵来建立两个域之间的映射关系. 本文的方法与之不同:1)本文改进地采用两个不同的表示矩阵,分别用源域表示目的域和用目的域表示源域;2)同时提出两个表示矩阵各自的最优化约束设计,使得源域和目的域最优地相互表示,从而减少域差异,实现借助源域数据对目的域无标签数据的分类.

1.1 源域有监督子空间学习

域自适应问题最关键的是减小两个域的差异. 通常可以通过寻求一个两域的共同子空间去实现两域之间的迁移. 由于源域具有可靠的真实标签,首先在源域学习一个标签引导的子空间. 模型如式(1)所示

$$\begin{aligned} \min_{P, C} \quad & \frac{1}{2} \| P^T X_s - (Y_s + D \odot C) \|_F^2, \\ \text{s. t. } \quad & C \geq 0, \\ & D_{i,j} = \begin{cases} +1, Y_s\{i,j\} = 1; \\ -1, \text{其他.} \end{cases} \end{aligned} \quad (1)$$

式中: $\odot$ 是 Hadamard 乘积运算, $P \in \mathbf{R}^{m \times d}$ 是共同的子空间, $X_s \in \mathbf{R}^{m \times n_s}$ 是源域的数据, $Y_s \in \mathbf{R}^{d \times n_s}$ 是源

域标签矩阵, m 是样本特征的维度, n_s 代表源域样本的数目, d 是共同子空间的维度. $C \in \mathbf{R}^{d \times n_s}$ 是松弛标签矩阵,加入松弛标签 C 是为了更自由地获得共同子空间.

1.2 构建表示矩阵

通常引入表示矩阵可以促进共同子空间的学习. 常见的方法是使用一个表示矩阵,即单方向地用源域的数据表示目的域数据,或者目的域数据表示源域数据,例如 DTSL 和 LSDT. 但这样的表示方式下,两个域的有用信息很难被完全保留. 这是因为源域和目标域的特征分布不同,两个域需要保留的信息不同,而使用同一个表示矩阵不能很好地保留两个域特有的有用信息和结构特征. 尤其是,当带有标签的源域样本单方向地靠近目的域数据时,会对源域的基本结构造成一定的破坏.

于是本文提出在源域和目的域采用两个不同的表示矩阵来表示另一个域. 即在源域存在一个表示矩阵去表示目的域的特征,同时在目的域存在另一个表示矩阵去表示源域的特征. 该模型如式(2)(3)所示

$$P^T X_s Z_s = P^T X_t, \quad (2)$$

$$(P^T X_s)^T = Z_t (P^T X_t)^T. \quad (3)$$

式中: $Z_s \in \mathbf{R}^{n_s \times n_t}$ 是作用于源域的表示矩阵; $Z_t \in \mathbf{R}^{n_s \times n_t}$ 是作用于目的域的表示矩阵; $X_t \in \mathbf{R}^{m \times n_t}$ 是目的域的样本, n_t 代表目的域样本的数目. 对于分布不同的两个域,按照各自特征去学习不同的表示矩阵有助于对齐两个域的分布,同时保存自己特有的信息.

2 表示矩阵的约束设计

共同子空间和合理的两个表示矩阵可以减小两个域差异的同时尽可能地保存两个域的原始有用信息. 为了得到尽可能最优的两个表示矩阵,对上面提出的两个表示矩阵进行了相应的约束设计,进而借助它们学习到有利于域适应的共同子空间. 由于两个表示矩阵需要保留的特征不同,应该按条件对它们施加不同的约束. 在两个域分布差异尽可能小的公共子空间上,目的域数据通过表示矩阵可以被源域的数据线性表示,如式(2)所示. 也就是说,目的域中的每个样本都可以视作是源域样本的线性组合. 再则考虑到源域的数据具有可靠的标签,对源域表示矩阵进行按列稀疏约束,表示为

$$\min_{P, Z_t} \| Z_t \|_1, \text{ s. t. } P^T X_s Z_s = P^T X_t. \quad (4)$$

式中 $\| \cdot \|_1$ 表示 1-范数. 这样目的域的样本可以

由更少的源域样本线性组合,保留了数据局部结构的同时能够更确切地分类.

但是同一类别往往有很多样本,当某一个样本由同一类样本线性组合表示的时候并不会丢失可判别性,反而会提高可判别性并降低过拟合的风险.受这个思路和最优化理论^[18]的启发,本文采用 group-lasso 作为 $\mathbf{Z}_s$ 表示矩阵的稀疏约束,对表示矩阵按类别进行稀疏约束,同一类别采用 2-范数降低稀疏约束强度,表示为

$$\sum_j \sum_{cl} \|\mathbf{Z}_s(\tau_{cl}, j)\|_2,$$

式中, $\|\cdot\|_2$ 表示 2-范数, τ_{cl} 是源域第 cl 类数据对应表示矩阵 $\mathbf{Z}_s$ 中相应行的组合, j 代表样本特征维度的第 j 维. 综上,源域表示矩阵 $\mathbf{Z}_s$ 受到的稀疏约束由式(4)改进为式(5).

$$\min_{P, \mathbf{Z}_s} \|\mathbf{Z}_s\|_{\text{group-lasso}}, \text{ s. t. } \mathbf{P}^T \mathbf{X}_s \mathbf{Z}_s = \mathbf{P}^T \mathbf{X}_t, \quad (5)$$

其中, $\|\mathbf{Z}_s\|_{\text{group-lasso}} = \sum_j \sum_{cl} \|\mathbf{Z}_s(\tau_{cl}, j)\|_2$.

为了更好地保留数据的结构信息,最理想的情况是源域中相同类别的数据被目的域中的数据用同一种线性表示方式所表示. 例如作用在目的域中表示矩阵的秩降低到等于类别数,即说明同一类样本拥有同样的线性表示方式. 于是作用于目的域表示矩阵应该受到低秩约束,如式(6)所示

$$\min_{P, \mathbf{Z}_t} \|\mathbf{Z}_t\|_*, \text{ s. t. } (\mathbf{P}^T \mathbf{X}_s)^T = \mathbf{Z}_t (\mathbf{P}^T \mathbf{X}_t)^T. \quad (6)$$

将式(1)~(3)、(5)和(6)合并,得到最终的模型,总模型为

$$\begin{aligned} \min_{P, \mathbf{C}, \mathbf{Z}_s, \mathbf{Z}_t} & \frac{1}{2} \|\mathbf{P}^T \mathbf{X}_s - (\mathbf{Y}_s + \mathbf{D} \odot \mathbf{C})\|_F^2 + \\ & \alpha \|\mathbf{Z}_s\|_{\text{group-lasso}} + \beta \|\mathbf{Z}_t\|_*, \\ \text{ s. t. } & \mathbf{P}^T \mathbf{X}_s \mathbf{Z}_s = \mathbf{P}^T \mathbf{X}_t, \\ & (\mathbf{P}^T \mathbf{X}_s)^T = \mathbf{Z}_t (\mathbf{P}^T \mathbf{X}_t)^T, \mathbf{C} \geq 0. \end{aligned} \quad (7)$$

其中, α 和 β 是超参数. 总模型示意图如图 1 所示.

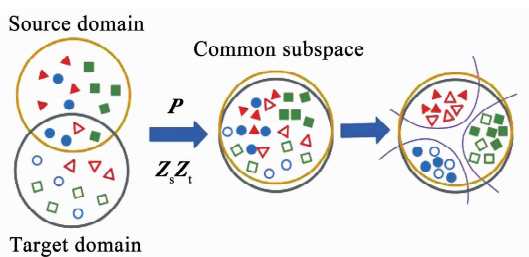


图 1 模型示意

Fig. 1 Schematic diagram of model

3 模型求解

式(7)的最优化问题是非凸的,求解只能保证局部最优而不是全局最优. 为了解决这个问题,利用不精确拉格朗日乘子法 (IALM) 将该问题转换成凸问题,转换如下:

$$\begin{aligned} \min_{P, \mathbf{C}, \mathbf{Z}_s, \mathbf{Z}_t} & \frac{1}{2} \|\mathbf{P}^T \mathbf{X}_s - (\mathbf{Y}_s + \mathbf{D} \odot \mathbf{C})\|_F^2 + \\ & \alpha \|\mathbf{Z}_s\|_{\text{group-lasso}} + \beta \|\mathbf{Z}_t\|_*, \\ \text{ s. t. } & \mathbf{P}^T \mathbf{X}_s \mathbf{Z}_s = \mathbf{P}^T \mathbf{X}_t, \\ & (\mathbf{P}^T \mathbf{X}_s)^T = \mathbf{Z}_t (\mathbf{P}^T \mathbf{X}_t)^T, \\ & \mathbf{Z}_s = \mathbf{Z}_2, \mathbf{Z}_t = \mathbf{Z}_1, \mathbf{C} \geq 0. \end{aligned} \quad (8)$$

可以通过增广拉格朗日乘子法,进一步将式(8)转换为

$$\begin{aligned} L = & \frac{1}{2} \|\mathbf{P}^T \mathbf{X}_s - (\mathbf{Y}_s + \mathbf{D} \odot \mathbf{C})\|_F^2 + \\ & \beta \|\mathbf{Z}_t\|_* + \alpha \|\mathbf{Z}_s\|_{\text{group-lasso}} + \\ & \langle \mathbf{Y}_1, \mathbf{P}^T \mathbf{X}_s \mathbf{Z}_s - \mathbf{P}^T \mathbf{X}_t \rangle + \\ & \langle \mathbf{Y}_2, (\mathbf{P}^T \mathbf{X}_s)^T - \mathbf{Z}_t (\mathbf{P}^T \mathbf{X}_t)^T \rangle + \\ & \langle \mathbf{Y}_3, \mathbf{Z}_t - \mathbf{Z}_1 \rangle + \langle \mathbf{Y}_4, \mathbf{Z}_s - \mathbf{Z}_2 \rangle + \\ & \frac{\mu}{2} \|\mathbf{P}^T \mathbf{X}_s \mathbf{Z}_s - \mathbf{P}^T \mathbf{X}_t\|_F^2 + \\ & \frac{\mu}{2} \|(\mathbf{P}^T \mathbf{X}_s)^T - \mathbf{Z}_t (\mathbf{P}^T \mathbf{X}_t)^T\|_F^2 + \\ & \frac{\mu}{2} \|\mathbf{Z}_t - \mathbf{Z}_1\|_F^2 + \frac{\mu}{2} \|\mathbf{Z}_s - \mathbf{Z}_2\|_F^2. \end{aligned} \quad (9)$$

式中, μ 是惩罚参数; $\mathbf{Y}_1, \mathbf{Y}_2, \mathbf{Y}_3$ 和 $\mathbf{Y}_4$ 代表拉格朗日乘子,在优化求解过程中,每次更新其中一个变量并固定其余变量迭代求解.

4 理论分析

Ben-David^[19]定理提出源域分类器在目的域中的误差上限,即借助源域有标签的样本训练得到的分类器来识别无标签的目的域样本的误差上限. 如式(10)所示

$$\begin{aligned} \epsilon_T(h) & \leq \epsilon_S(h) + d_1(\mathbf{D}_S, \mathbf{D}_T) + \\ & \min_{D_S} E_{D_S}[|f_S(X) - f_T(X)|], \\ & E_{D_T}[|f_S(X) - f_T(X)|]. \end{aligned} \quad (10)$$

式中: $\epsilon_T(h) = \epsilon_T(h, f) = E_{x \sim D_T}[|h(X) - f(X)|]$. $\epsilon_T(h)$ 表示源域分类器在目的域中的误差上限, $f(X)$ 是实际标签函数, $h(X)$ 是假设的标签函数. $\epsilon_S(h)$ 表示源域分类器在源域中的误差. $d_1(\mathbf{D}_S,$

D_T) 表示两个域之间的差异. $\min E_{D_S}[|f_S(X) - f_T(X)|], E_{D_T}[|f_S(X) - f_T(X)|]$ 表示标签函数在两个域之间的差异.

由于标签函数是已知的, 所以 $\min E_{D_S}[|f_S(X) - f_T(X)|], E_{D_T}[|f_S(X) - f_T(X)|]$ 是一个常量. 从式 (10) 可以推出, 若需减小目的域分类误差, 则需要减小 $\epsilon_S(h)$ 和 $d_1(D_S, D_T)$. 所以域自适应算法的关键是在保证源域误差尽可能小的同时减小两个域之间的差异. 本文模型中的式 (1) 就是利用源域的可靠真实标签来减小 $\epsilon_S(h)$. 同时通过式 (2) 和 (3) 构建了两个域之间的关系, 然后提出最优化约束设计来减小两域的差异, 减小 $d_1(D_S, D_T)$. 相较于目前其他域自适应算法, 所提算法使两域可以最优地相互表示, 从而降低过拟合以及破坏两域基本结构的风险.

5 实 验

在 3 个迁移学习常用的数据集上开展实验进行验证. 1) COIL20 数据集将 1 440 张灰度图片分成了两个域: COIL1 (C1) 和 COIL2 (C2). 该数据集的样本有 20 个类别, COIL20 数据集的部分示例见图 2. 2) Office - Caltech 10^[10] 是最广泛使用的数据集, 一共分为 4 个域: Caltech (C)、Amazon (A)、DSLR (D) 和 Webcam (W). 本文的实验使用了该数据集 800 维的 SURF 特征和 4 096 维的 DeCAF 特征. 3) ImageCLEF - DA^[16] 由 3 个数据集共同的 12 类数据组成, 3 个数据集分别是: Caltech - 256 (C)、ImageNet ILSVRC 2012 (I) 和 VOC 2012 (P). 可构建 6 个跨域任务: C→I、C→P、I→C、I→P、P→C 和 P→I.



图 2 COIL20 数据集部分示例

Fig. 2 Example images from COIL20

与以下 10 种传统迁移学习方法进行了对比, 包括: TCA^[6]、GFK^[14]、JDA^[8]、SA^[11]、DTSL^[16]、

CORAL^[20]、BDA^[21]、DICD^[22]、KOT^[23] 和 DST - ELM^[24], 同时还与以下 2 种深度方法也进行了对比: AlexNet^[25] 和 JDOT^[26]. 实验结果见表 1~5, 其中识别准确率的最好值与次好值分别通过加粗和下划线表示. 从实验结果来看, 本文的方法超过了很多的传统迁移学习方法和一些深度方法, 其中包括一些比较先进的传统迁移学习方法如 DST - ELM 和深度方法如 JDOT 等, 这体现了本文方法的有效性.

5.1 实验结果及分析

在 COIL20 数据集构建两个跨域任务: COIL1→COIL2 和 COIL2→COIL1. 其中, COIL1 由角度为 $[0^\circ, 85^\circ] \cup [180^\circ, 265^\circ]$ 的灰度图像组成, 而 COIL2 则由 $[90^\circ, 175^\circ] \cup [270^\circ, 355^\circ]$ 的灰度图像组成. 实验结果如表 1 所示, 本文算法的平均准确率达 89.3%, 比 CORAL 准确率高了 7.3%. JDA 简单地考虑边缘分布和条件分布适配反而比 DTSL 和 CORAL 效果好. 由于两个域的差异仅仅来自图片的拍摄角度不同, 所以该任务比其他数据集简单, 过度拟合和破坏源域的基本结构反而会降低在该数据集的识别准确性. 而本文采用 group - lasso 作为表示矩阵的稀疏约束, 可以降低过拟合风险, 从实验结果可以验证本文模型的有效性和鲁棒性.

表 1 在 COIL20 数据集上的准确率

Tab. 1 Accuracy on COIL20					%
任务	JDA	SA	DTSL	CORAL	本文
C1→C2	89.3	84.9	84.6	82.2	<u>86.9</u>
C2→C1	<u>88.5</u>	87.2	84.2	81.7	91.7
平均值	<u>88.9</u>	86.0	84.4	82.0	89.3

在 Office - Caltech 10 数据集上, 不论使用 SURF 特征还是 DeCAF 特征, 本文算法都超过了对比的算法. 采用 SURF 特征的平均准确率为 51.2%, 使用 DeCAF 特征, 本文模型的平均准确率达 90.6%, 分别如表 2 和表 3 所示. 与 DTSL 对比, 使用 SURF 特征, 本文模型的准确率提升了 4.5%, 如果使用 DeCAF 深度特征, 本文模型将准确率从 83.8% 提升到了 90.6%. 由于 DTSL 只使用了一个表示矩阵, 对源域的基本结构造成了一定的破坏, 很难保存所有目的域的有用信息. 所以本文模型采用两个表示矩阵解决了相关问题, 从而提升了准确率. DICD 通过 MMD 测量来减小两个域之间的差异, 虽然考虑了减小类内距离和扩大类间距离, 但是只通过 MMD 距离来判断差距也会对结构造成一定的破坏, 本文算法在这个数据集上的效果仍然高于 DICD, 这也验证了本文算法具有不错的可判别性.

表 2 在 Office – Caltech 10 (SURF) 数据集上的准确率

Tab.2 Accuracy on Office – Caltech 10(SURF) %					
任务	DTSL	CORAL	DICD	DST – ELM	本文
C→A	53.3	52.1	47.3	<u>53.7</u>	57.0
C→W	45.8	46.4	46.4	<u>51.5</u>	52.0
C→D	51.0	45.9	49.7	49.0	<u>50.3</u>
A→C	<u>44.7</u>	45.1	42.4	37.5	43.6
A→W	38.3	44.4	<u>45.1</u>	39.7	48.5
A→D	39.5	39.5	38.9	<u>51.0</u>	52.2
W→C	30.3	<u>33.7</u>	33.6	36.0	32.6
W→A	34.7	<u>36.0</u>	34.1	42.1	33.8
W→D	82.8	<u>86.6</u>	89.8	82.2	86.6
D→C	30.7	<u>33.8</u>	34.6	29.5	32.7
D→A	33.2	37.7	34.5	44.1	<u>38.3</u>
D→W	76.6	84.7	91.2	84.1	<u>86.8</u>
平均值	46.7	48.8	49.0	<u>50.0</u>	51.2

表 3 在 Office – Caltech 10 (DeCAF) 数据集上的准确率

Tab.3 Accuracy on Office – Caltech 10(DeCAF) %					
任务	DTSL	CORAL	KOT	DST – ELM	本文
C→A	91.5	91.5	<u>91.7</u>	91.2	92.1
C→W	76.6	80.7	88.8	90.4	<u>90.0</u>
C→D	88.0	88.0	91.1	<u>91.6</u>	94.5
A→C	85.8	84.0	<u>86.0</u>	86.3	85.7
A→W	73.6	75.6	<u>85.1</u>	84.2	86.1
A→D	82.2	84.7	<u>87.3</u>	86.3	88.7
W→C	72.8	67.8	82.5	85.1	<u>83.3</u>
W→A	75.5	72.4	89.4	91.1	<u>89.8</u>
W→D	100.0	100.0	<u>99.4</u>	100.0	100.0
D→C	75.2	72.0	83.5	86.5	<u>84.9</u>
D→A	85.0	84.2	90.4	<u>91.7</u>	92.1
D→W	99.3	99.0	97.3	99.9	<u>99.5</u>
平均值	83.8	83.3	89.4	<u>90.4</u>	90.6

传统的方法在 Office – Caltech 10 数据集使用 DeCAF 特征都取得了不错的结果. 虽然本文的方法与一些先进的传统方法(如 DST – ELM)对比只提升了一点,但是与某些先进的深度方法(如 JDOT)做对比时,本文的方法仍能显示出一定的优势,如表 4 所示.

在 ImageCLEF – DA 数据集,使用 ResNet50 网络提取的深度特征. 显然,本文算法的效果优于所有对比的算法. 实验结果如表 5 所示.

表 4 在 Office – Caltech 10 (DeCAF) 数据集上与深度方法对比准确率

Tab.4 Comparison of accuracy of proposed method and deep methods on Office – Caltech 10 %			
任务	AlexNet	JDOT	本文
C→A	<u>91.9</u>	91.5	92.1
C→W	83.7	<u>88.8</u>	90.0
C→D	87.1	<u>89.8</u>	94.5
A→C	83.0	<u>85.2</u>	85.7
A→W	79.5	<u>84.8</u>	86.1
A→D	87.4	<u>87.9</u>	88.7
W→C	73.0	<u>82.6</u>	83.3
W→A	83.8	90.7	<u>89.8</u>
W→D	100.0	<u>98.1</u>	100.0
D→C	79.0	<u>84.3</u>	84.9
D→A	87.1	<u>88.1</u>	92.1
D→W	<u>97.7</u>	96.6	99.5
平均值	86.1	<u>89.0</u>	90.6

表 5 在 ImageCLEF – DA 数据集上的准确率

Tab.5 Accuracy on ImageCLEF – DA %					
任务	TCA	GFK	CORAL	BDA	本文
C→I	89.3	86.3	83.0	90.8	<u>89.5</u>
C→P	74.5	73.3	71.5	<u>73.7</u>	71.8
I→C	93.2	93.0	88.7	<u>94.0</u>	95.2
I→P	77.5	75.5	73.7	75.3	<u>76.2</u>
P→C	<u>83.7</u>	82.3	72.0	83.5	90.8
P→I	<u>80.8</u>	78.0	71.3	77.8	90.5
平均值	<u>83.2</u>	81.4	76.7	82.5	85.7

5.2 消融分析

为了进一步验证对作用于源域 的表示矩阵施加 group – lasso 约束是否可以提高模型的识别准确率,进行了消融实验. $RL_{low-rank}$ 模型只对 Z_t 矩阵进行低秩约束,对 Z_s 矩阵不做约束. RL_{sparse} 模型对 Z_s 矩阵做 1 – 范数稀疏约束. $RL_{group-lasso}$ 模型对 Z_s 矩阵采用 group – lasso 约束. RL_{joint1} 模型对 Z_s 和 Z_t 两个表示矩阵分别做 1 – 范数和低秩约束,而 RL_{joint2} 模型对 Z_s 和 Z_t 两个表示矩阵分别做 group – lasso 和低秩约束. 实验结果如表 6 所示. 通过对比 $RL_{group-lasso}$ 和 RL_{sparse} 两个模型的实验结果,可以明显得出在域自适应表示学习中,稀疏约束采用 group – lasso 优于使用 1 – 范数. 通过 RL_{joint1} 的实验结果,仅仅对作用于源域的表示矩阵 Z_s 采用 group – lasso 约束的分类准确率甚至高于同时使用 1 – 范数和低秩约束. 对比 $RL_{group-lasso}$ 和 RL_{joint2} 模型的实验结果,对 Z_t 的低秩约束也是必不可少的,能够更好地保留样本特征的结构,从而提高准确率.

表 6 Office – Caltech 10 (SURF) 数据集上消融实验的结果

Tab.6 Results of ablation experiments on Office – Caltech 10 (SURF) %

任务	RL _{low-rank}	RL _{sparse}	RL _{group-lasso}	RL _{joint1}	RL _{joint2}
C→A	51.3	51.2	<u>56.9</u>	51.3	57.0
C→W	37.0	37.6	<u>41.0</u>	38.6	52.0
C→D	45.9	47.1	<u>47.8</u>	47.1	50.3
A→C	43.4	43.2	44.1	43.4	<u>43.6</u>
A→W	34.6	35.9	<u>40.7</u>	36.6	48.5
A→D	38.2	38.9	<u>44.6</u>	38.9	52.2
W→C	29.4	29.3	32.7	29.8	<u>32.6</u>
W→A	<u>34.0</u>	34.1	33.5	34.1	33.8
W→D	82.2	82.2	81.5	<u>82.8</u>	86.6
D→C	31.1	31.5	<u>31.8</u>	31.6	32.7
D→A	33.5	33.4	<u>34.6</u>	33.2	38.3
D→W	76.6	77.2	<u>83.1</u>	77.3	86.8
平均值	44.8	45.1	<u>47.7</u>	45.4	51.2

5.3 参数敏感度分析

由式(7)可知,本文的总模型含有两个超参数 α 和 β . 为了分析不同参数对模型识别准确率的影响,在 COIL20、Office – Caltech 10 和 ImageCLEF – DA 3 个数据集进行了参数敏感度实验:在离散数据集 $[0.001, 0.01, 0.1, 1, 5, 10]$ 的范围内改变两个超参数的值,观察识别准确率的变化,实验结果见图 3. 从图 3 可以观察到,参数 α 和 β 的取值在 $0.001 \sim 10$ 大范围变化,对识别准确率造成的影响依然很小. 所以虽然本文的模型需要调试两个超参数,但是两个超参数的选择却较简单,在相对大的范围内取值,都可以使本文的模型分类准确率达到一个很好的效果.

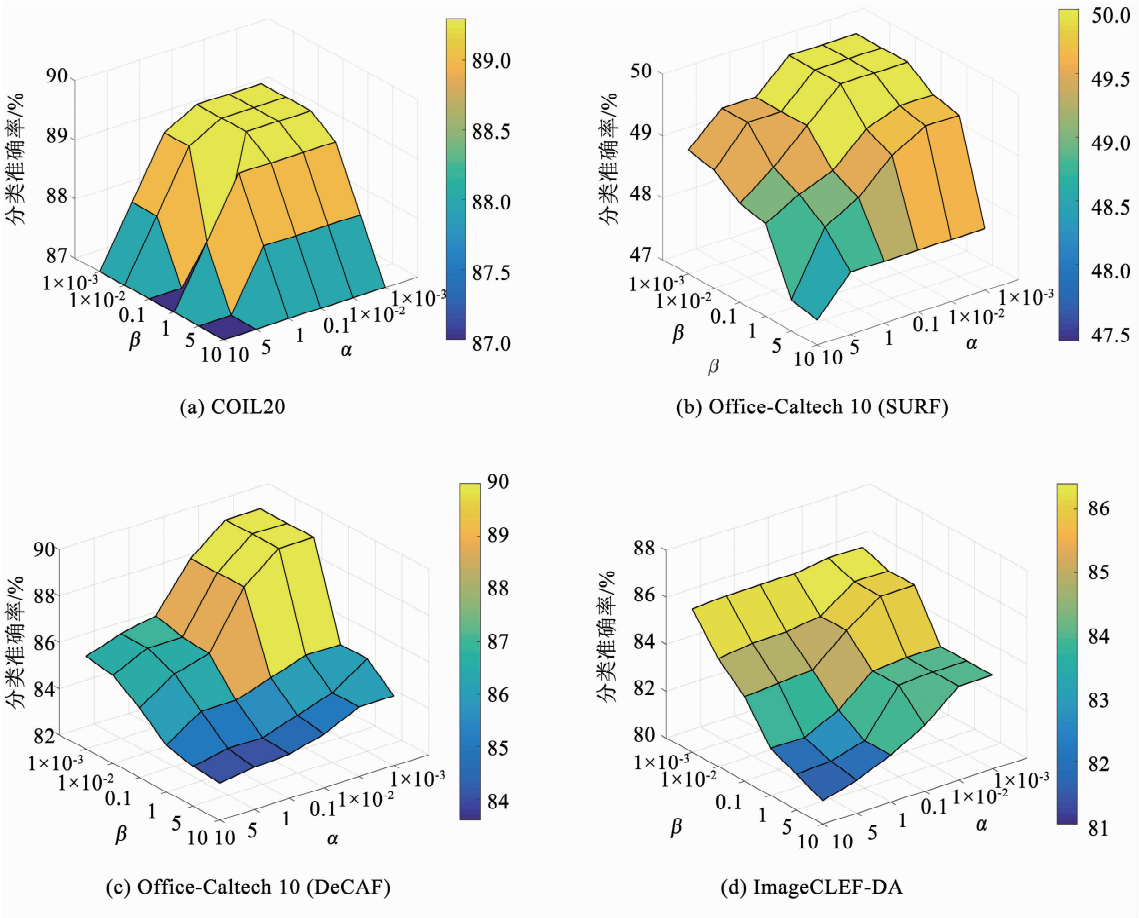


图 3 参数敏感度分析

Fig.3 Parameter sensitivity analysis

6 结 论

针对无监督域自适应问题,本文提出了一个新颖的表示学习算法. 为了在学得共同子空间下更好地保留源域和目的域样本的特有特征和结构,使

用两个不同的表示矩阵分别作用于两个域,同时基于线性表示和最优传输相关理论为这两个表示矩阵设计不同的约束. 大量的实验验证了本文模型的有效性和鲁棒性,在多个数据集上,本文方法的识别精度超过了很多先进的无监督域自适应方法.

参考文献

- [1] LIU Hong, LONG Mingsheng, WANG Jianmin, et al. Transferable adversarial training: a general approach to adapting deep classifiers [C]// International Conference on Machine Learning. Long Beach: PMLR, 2019: 1
- [2] CHEN Yu, YANG Chunling, ZHANG Yan, et al. Deep conditional adaptation networks and label correlation transfer for unsupervised domain adaptation [J]. Pattern Recognition, 2020, 98: 107072. DOI: 10.1016/j.patcog. 2019.107072
- [3] ZHANG Lei, WANG Shanshan, HUANG Guangbin, et al. Manifold criterion guided transfer learning via intermediate domain generation [J]. IEEE Transactions on Neural Networks and Learning Systems, 2019, 30(12): 59. DOI: 10.1109/TNNLS.2019.2899037
- [4] PEREIRA L A, TORRES R D S. Semi-supervised transfer subspace for domain adaptation [J]. Pattern Recognition, 2018, 75: 235. DOI: 10.1016/j.patcog. 2017.04.011
- [5] GRETTON A, BORWARDT K, RASCH M, et al. A kernel method for the two-sample-problem [C]// International Conference on Neural Information Processing Systems. Cambridge: MIT Press, 2006: 513
- [6] PAN S J, TSANG I W, KWOK J T, et al. Domain adaptation via transfer component analysis [J]. IEEE Transactions on Neural Networks, 2011, 22(2): 199. DOI: 10.1109/tnn.2010.2091281
- [7] LONG Mingsheng, WANG Jianmin, DING Guiguang, et al. Transfer joint matching for unsupervised domain adaptation [C]// 2014 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2014: 1. DOI: 10.1109/CVPR.2014.183
- [8] LONG Mingsheng, WANG Jianmin, DING Guiguang, et al. Transfer feature learning with joint distribution adaptation [C]// 2013 IEEE International Conference on Computer Vision. Piscataway: IEEE, 2013: 1. DOI: 10.1109/ICCV.2013.274
- [9] ZHANG Jing, LI Wanqing, OGUNBONA P. Joint geometrical and statistical alignment for visual domain adaptation [C]// 2017 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2017: 2. DOI: 10.1109/CVPR.2017.547
- [10] LONG Mingsheng, WANG Jianmin, DING Guiguang, et al. Adaptation regularization: a general framework for transfer learning [J]. IEEE Transactions on Knowledge and Data Engineering, 2014, 26(5): 1076. DOI: 10.1109/TKDE.2013.111
- [11] FERNANDO B, HABRARD A, SEBBAN M, et al. Unsupervised visual domain adaptation using subspace alignment [C]// 2013 IEEE International Conference on Computer Vision. Piscataway: IEEE, 2013: 1. DOI: 10.1109/ICCV.2013.368
- [12] SUN Baochen, SAENKO K. Subspace distribution alignment for unsupervised domain adaptation [C]// Proceedings of British Machine Vision Conference. Berlin: Springer, 2015: 24. DOI: 10.5244/C.29.24
- [13] DING Zhengming, FU Yun. Robust transfer metric learning for image classification [J]. IEEE Transactions on Image Processing, 2017, 26(2): 660. DOI: 10.1109/TIP.2016.2631887
- [14] GONG Boqing, SHI Yuan, SHA Fei, et al. Geodesic flow kernel for unsupervised domain adaptation [C]// 2012 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2012: 1. DOI: 10.1109/CVPR.2012.6247911
- [15] BAKTASHMOTLAGH M, HARANDI M T, LOVELL B C, et al. Unsupervised domain adaptation by domain invariant projection [C]// International Conference on Computer Vision. Piscataway: IEEE, 2013: 1. DOI: 10.1109/ICCV.2013.100
- [16] XU Yong, FANG Xiaozhao, WU Jian, et al. Discriminative transfer subspace learning via low-rank and sparse representation [J]. IEEE Transactions on Image Processing, 2016, 25(2): 850. DOI: 10.1109/TIP.2015.2510498
- [17] ZHANG Lei, ZUO Wangmeng, ZHANG D. LSST: latent sparse domain transfer learning for visual adaptation [J]. IEEE Transactions on Image Processing, 2016, 25(3): 1177. DOI: 10.1109/TIP.2016.2516952
- [18] COURTY N, FLAMARY R, TUIA D, et al. Optimal transport for domain adaptation [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2016, 39(9): 1853. DOI: 10.1109/TPAMI.2016.2615921
- [19] BEN-DAVID S, BLITZER J, CRAMMER K, et al. A theory of learning from different domains [J]. Machine Learning, 2010, 79(1/2): 151. DOI: 10.1007/s10994-009-5152-4
- [20] SUN Baochen, FENG Jiashi, SAENKO K. Return of frustratingly easy domain adaptation [C]// Thirtieth AAAI Conference on Artificial Intelligence. Palo Alto: AAAI, 2016: 2058
- [21] WANG Jindong, CHEN Yiqiang, HAO Shuji, et al. Balanced distribution adaptation for transfer learning [C]// 2017 IEEE International Conference on Data Mining. Piscataway: IEEE, 2017: 1129. DOI: 10.1109/ICDM.2017.150
- [22] LI Shuang, SONG Shiji, HUANG Gao, et al. Domain invariant and class discriminative feature learning for visual domain adaptation [J]. IEEE Transactions on Image Processing, 2018, 27(9): 4260. DOI: 10.1109/TIP.2018.2839528
- [23] ZHANG Zhen, WANG Mianzhi, HUANG Yan, et al. Aligning infinite-dimensional covariance matrices in reproducing kernel Hilbert spaces for domain adaptation [C]// 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2018: 3437. DOI: 10.1109/CVPR.2018.00362
- [24] CHEN Yiming, SONG Shiji, LI Shuang, et al. Domain space transfer extreme learning machine for domain adaptation [J]. IEEE Transactions on Cybernetics, 2019, 49(5): 1909. DOI: 10.1109/TCYB.2018.2816981
- [25] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. ImageNet classification with deep convolutional neural networks [J]. Communications of the ACM, 2017, 60(6): 84. DOI: 10.1145/3065386
- [26] COURTY N, FLAMARY R, HABRARD A, et al. Joint distribution optimal transportation for domain adaptation [Z]. arXiv:1705.08848

(编辑 苗秀芝)