

doi:10.11835/j.issn.1000-582X.2021.11.008

低秩先验引导的无监督域自适应行人重识别

李玲莉^{a,b}, 谢明鸿^a, 李凡^{a,b}, 张亚飞^{a,b}, 李华锋^{a,b}, 谭婷婷^a

(昆明理工大学 a.信息工程与自动化学院; b.云南省人工智能重点实验室 昆明 650500)

摘要:无监督域自适应行人重识别在智能监控中发挥着重要作用,并引起了研究者的广泛关注。尽管目前的研究已经取得了较大进步,但不同数据集之间的域偏移问题给行人重识别带来极大挑战。研究发现,在连续时间里,同一摄像机视角下的行人图像具有相同的风格,如果将这种风格信息从行人图像中分离出去,将有效缓解由图像风格差异引起的域偏移问题。为此,提出一种低秩先验引导的域不变信息分离的字典学习方案。根据风格信息的低秩先验性,将行人图像特征中的风格信息和行人身份信息分离开来,根据同一身份行人属性的域不变性建立视觉特征与属性之间的联系,缓解域偏移所带来的影响,通过自训练策略来调整学习参数。实验表明,方法的性能在很多数据集上超过了传统的无监督域自适应行人重识别方法以及部分基于深度学习的无监督域自适应行人重识别方法。

关键词:行人重识别;低秩先验;域偏移;特征分离。

中图分类号:TP311

文献标志码:A

文章编号:1000-582X(2021)11-057-15

Unsupervised domain adaptive person re-identification guided by low-rank priori

LI Lingli^{a,b}, XIE Minghong^a, LI Fan^{a,b}, ZHANG Yafei^{a,b}, LI Huafeng^{a,b}, TAN Tingting^a

(a. Faculty of Information Engineering and Automation; b. Key Laboratory of Artificial Intelligence in Yunnan Province, Kunming University of Science and Technology, Kunming 650500, P. R. China)

Abstract: Unsupervised domain adaptive person re-identification plays an important role in intelligent monitoring, but the domain shift among different datasets brings great challenges to person re-identification. Studies have reported that the pedestrian images captured from the same camera view have same style in continuous time. If this style information is separated from the pedestrian image, the domain shift problem caused by image style difference will be effectively alleviated. In this paper, a low rank prior guided dictionary learning scheme with domain invariant information separation was proposed. Firstly, according to the low rank priori of the style information, style information and pedestrian identity information in the pedestrian image features were separated. Secondly, according to the domain invariance of the pedestrian attributes of the same identity, the relationship between the visual features and the

收稿日期:2020-08-12

基金项目:国家自然科学基金资助项目(61966021,61562053),大学生创新创业训练计划项目(202010674098)。

Supported by National Natural Science Foundation of China (61966021, 61562053); College Students' Innovative Entrepreneurial Training Plan Program(202010674098).

作者简介:李玲莉(1994—),女,硕士研究生,主要从事计算机视觉,图像处理方向研究,(Tel)15601527650;(E-mail)996137628@qq.com。

通讯作者:谢明鸿(1976—),男,高级工程师,硕士生导师,主要从事计算机视觉,图像处理方向研究。

attributes was established to alleviate the impact of domain shift. Finally, self-training strategy was used to adjust the learning parameters. Experimental results show that the proposed method outperforms the traditional unsupervised domain adaptive person re-identification methods and some unsupervised domain adaptive person re-identification methods based on deep learning in many datasets.

Keywords: person re-identification; low-rank priori; domain shift; feature separation

行人重识别是一种从不同相机视角下在采集的行人图像中搜索出相同行人的技术,在智能监控中发挥重要作用,因此受到研究者的广泛关注。在现实监控环境中,同一行人的不同图像常常因相机视角差异、光照变化等影响呈现出不同的外貌特征,这给行人重识别带来了极大挑战。由于不需要进行大量人工样本标记,近年来无监督域自适应行人重识别方法已成为研究热点之一。这些方法主要是通过无监督域不变特征提取、对抗学习和图像风格迁移等手段来减小源域与目标域之间的差异。然而,无监督域不变特征提取的方法往往很难挖掘具有判别性的特征信息;基于对抗学习的无监督方法提取到的特征往往是来自不同域样本的共有信息,丢失了不同域样本之间的特有信息,容易加剧样本特征之间的歧义性;基于图像风格迁移的方法虽然有效,但容易导致迁移图像标签信息改变。

不同数据集之间产生域偏移的主要原因之一是相机的成像风格不同。从图 1 可发现来自同一视角下的行人图像,呈现出相似的风格信息,如果将这些图像的视觉特征由矩阵表示,这个矩阵在数学上就会呈现出低秩性。因此根据这种低秩先验性,设计一种域不变特征分离模型,将风格信息从图像特征中分离出去,这种策略不仅能缓解不同视角之间域偏移给行人重识别带来的挑战,还能保持行人身份特征信息不发生改变,避免传统无监督域自适应方法的标签迁移。整体思路如图 2 所示。



图 1 挑选自不同数据集同一摄像头的行人图像

Fig. 1 Pedestrian images selected from the same camera view of different datasets

为解决单个数据集中多个视角下行人图像风格不一致问题,提出低秩成分分解的字典学习模型,将不同相机视角下的行人图像特征分解成具有低秩特性的风格信息和具有判别性的行人身份信息,通过去除分解出来的风格信息,利用剩下的行人身份信息来训练判别字典学习模型,并将行人身份信息在其对应字典下的判别系数作为行人的潜在身份特征,作为最后行人身份相似性度量。为进一步缓解域偏移所带来的影响,利用同一身份行人属性的域不变性,将属性与特征关联模块嵌入到字典学习框架中,构建从特征到属性的映射,进一步挖掘属性与特征之间的关系。为了保证潜在视觉特征迁移到语义属性空间的过程中行人身份信息不发生改变,在字典学习模型中引入自重约束以及身份判别特征与标签一致性约束,提升字典的判别性。通过构建标记样本身份判别特征和标签之间的关系,拉近同一行人潜在特征之间的距离,拉远不同行人潜在特征之间的距离。最后,通过选择置信度高的伪标签来调整模型参数,使其更适应目标数据的识别。

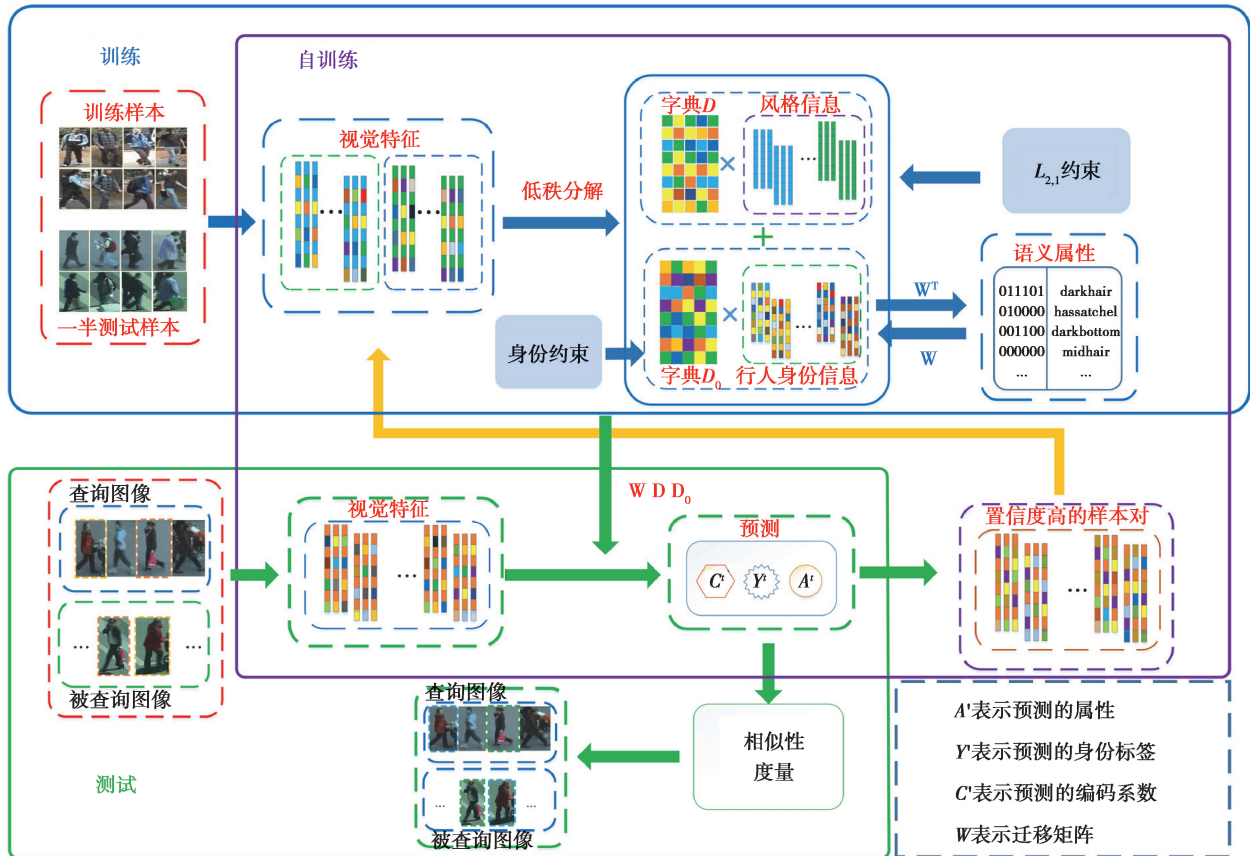


图 2 研究算法的整体框架

Fig. 2 The overall framework of the algorithm in this paper

1 相关研究

1.1 无监督域自适应行人重识别

为解决有监督行人重识别方法和传统无监督域自适应行人重识别方法推广应用方面的不足,近 2 年研究者提出了一些新的无监督域自适应的行人重识别方法。其中包括基于样本存储的域自适应方法^[1],基于块判别性特征学习的无监督行人重识别方法^[2],基于自相似聚类的域自适应方法^[3],基于元学习的域生成方法^[4],基于对抗学习的域自适应行人重识别方法^[5-7]以及基于域不变特征提取方法^[8]。

基于样本存储^[1]的方法利用样本存储实现了 3 种基本不变性,即样本不变性,相机不变性和邻域不变性,从而减小了源域与目标域之间的差异。虽然基于块的判别性特征学习^[2]也是较为有效的算法,但这类算法没有考虑块与块之间的关系,导致识别性能稍弱。基于自相似聚类方法^[3]可获得较高的识别性能,但这种方法克服域偏移问题的本质是自训练,与现实场景不符。基于元学习的域自适应方法是最近提出的一种有

效方法^[4],该方法使用元学习来更新模型中的分类器参数,并使用多个数据集来训练模型,使模型具有域不变性。基于对抗学习^[5-7]的域自适应方法可以实现不同数据集行人图像的迁移,但在迁移过程中容易造成不同样本的特有信息丢失,不利于行人身份信息保留。基于域不变特征提取的方法^[8]通过提取不同数据集间的共有信息用作行人身份识别,该方法虽然有效,但未能充分考虑域信息的低秩先验性,从而导致提取的特征判别性较弱,限制了识别性能的进一步提升。这些方法虽然能取得不错的实验结果,但不能有效挖掘无标记样本的判别信息,导致限制识别性能的提升空间,且对标记的训练数据的数量和规模都要求较大。为实现域自适应的行人重识别,利用同一视角下行人图像风格的低秩先验性,来实现风格信息与行人身份信息的分离,该方法在不需要大规模有标记训练样本的情况下便能接近甚至达到基于深度学习的无监督域自适应方法识别性能。

1.2 基于字典学习的行人重识别

字典学习因其在模式识别和图像处理方面的优异表现,已成为解决行人重识别问题的常用方法。这类方法通常利用字典和表示系数的乘积来表示特征,并把表示系数作为最终的底层视觉特征,用于最终行人图像间的相似性度量;这类方法还可以利用模型的正则项约束来提升表示系数的判别能力。近年来,一些基于字典学习的行人重识别方法被提出。文献[9]提出了一种多任务字典学习的跨数据集迁移学习方法,该方法通过将字典分解成任务共享字典和视角特定字典,分别获取不同视角下行人的潜在属性和行人外貌差异表示。文献[10]提出一种基于卡方核的正则化线性判别分析行人重识别算法,使数据在低维空间能够保持高维空间的可分性,提高行人重识别算法性能。文献[11]提出一种面向无监督域自适应的联合属性身份嵌入字典学习方法,该方法通过视觉特征空间和语义属性空间的对齐来学习语义属性字典和视觉特征字典。文献[12]将多特征字典学习和自适应多特征图整合到一个联合学习模型中,使得学习到的字典具有判别性。虽然上述方法在行人重识别中能达到不错的识别性能,但无法有效克服相机风格差异带来的域偏移问题。基于成像风格的低秩先验性,提出了一种判别字典学习的图像风格分离模型,降低成像风格差异所导致的不同视角行人图像之间的域偏移。

1.3 基于属性学习的行人重识别

由于语义属性具有域不变性,因此在行人重识别中得到了广泛应用。但与视觉特征相比,属性通常具有较弱的判别性。为解决该问题,文献[13]提出联合属性-身份的深度学习方法,通过联合提取属性和标记数据的身份特征来提取具有判别性的身份属性信息。文献[14]提出一种从粗到细的学习框架来分别预测行人身份和行人性别,实现身份标签与属性标签的综合利用。文献[15]提出一种属性互补信息的学习方法,它通过在单个数据集上训练一个属性分类器,并将属性信息整合到卷积神经网络模型的训练中,使网络学习到与语义属性互补的行人身份信息。文献[16]提出一种基于属性的行人重识别方法,通过同时学习行人身份嵌入和属性映射,使行人身份信息与属性相互影响。上述这些方法都是将属性作为辅助信息对行人身份信息进行补充。而将属性作为不同域之间的桥梁,在目标域上建立视觉潜在特征与身份标签之间的联系,以便能对目标域内的样本属性进行正确预测,用于后续识别。

2 方法论

2.1 问题定义

假设有标签源数据集为 $\Phi_s = \{\mathbf{X}^s, \mathbf{Y}^s, \mathbf{A}^s\}$,无标签目标数据集为 $\Phi_t = \{\mathbf{X}^t, \mathbf{Y}^t\}$,其中 $\mathbf{X}^s = \{\mathbf{X}_1^s, \mathbf{X}_2^s, \dots, \mathbf{X}_M^s\}$ 和 $\mathbf{X}^t = \{\mathbf{X}_1^t, \mathbf{X}_2^t, \dots, \mathbf{X}_L^t\}$ 分别为源域和目标域内的样本集合, M 和 L 分别表示相机个数。源域中,第 v 个视角下的样本集合表示为 $\mathbf{X}_v^s = [x_1^{sv}, x_2^{sv}, \dots, x_{n_v}^{sv}] \in R^{m \times n_v}$,目标域中第 h 个视角下的未标记样本集合表示为 $\mathbf{X}_h^t = [x_1^{th}, x_2^{th}, \dots, x_{n_h}^{th}] \in R^{m \times n_h}$,其中 $x_i^{sv} \in R^{m \times 1}$, $i \in [1, 2, \dots, n_v]$ 表示在源域中第 v 个摄像头下拍摄到的第 i 张图像的特征向量, $x_j^{th} \in R^{m \times 1}$, $j \in [1, 2, \dots, n_h]$ 表示在目标域中第 h 个摄像头下拍摄到的第 j 张图像的特征向量。 $\mathbf{A}^s = \{\mathbf{A}_1^s, \mathbf{A}_2^s, \dots, \mathbf{A}_M^s\}$ 表示源域的语义属性,其中 $\mathbf{A}_i^s = [a_1^{sv}, a_2^{sv}, \dots, a_{n_v}^{sv}] \in R^{c \times n_v}$ 表示源域中第 v

个视角下行人样本所对应的语义属性。 \mathbf{Y}^s 表示源域行人样本的身份矩阵, $\mathbf{Y}_v = \{y_1^v, y_2^v, \dots, y_{m_v}^v\} \in R^{m_v \times n_v}$ 表示第 v 个视角下的行人样本标签, 其中, m_v 表示第 v 个摄像头捕捉到的行人身份数量。 \mathbf{Y}^t 表示需要进行预测的目标数据集标签, 在源域 Φ_s 中, $\mathbf{X}^s, \mathbf{Y}^s, \mathbf{A}^s$ 是已知的, 而在目标域 Φ_t 中, 只有视觉特征 \mathbf{X}^t 是已知的。所提出方法的主要任务是利用在源域中训练的模型来预测目标域的行人身份。

2.2 判别性字典学习

在无监督的行人重识别任务中, 由于不同数据集间存在域偏移问题, 在源域上训练的模型常常不能直接应用到目标域上。为解决这个问题, 基于行人图像风格信息的低秩先验性设计了一个域不变信息分离模型

$$\begin{aligned} \min_{\mathbf{D}, \mathbf{D}_0, \mathbf{C}_v^{sc}} \sum_{v=1}^M \|\mathbf{X}_v^s - \mathbf{D}_0 \mathbf{C}_v^s\|_F^2 + \|\mathbf{X}_v^s - \mathbf{D} \mathbf{C}_v^{sc} - \mathbf{D}_0 \mathbf{C}_v^s\|_F^2 + \lambda_1 \|\mathbf{X}_v^s - \mathbf{D}_0 \mathbf{C}_v^s\|_* + \alpha_1 \|\mathbf{C}_v^{sc}\|_{2,1} + \gamma \|\mathbf{D}\|_* \\ \text{s.t. } \|\mathbf{d}_i\|_2 \leq 1, \|\mathbf{d}_{0,j}\|_2 \leq 1, \forall i, j, \end{aligned} \quad (1)$$

其中, $\mathbf{C}_v^{sc} \in R^{d \times n_v}$ 和 $\mathbf{C}_v^s \in R^{d_0 \times n_v}$ 分别表示风格字典 $\mathbf{D} \in R^{m \times d}$ 和行人身份字典 $\mathbf{D}_0 \in R^{m \times d_0}$ 对应的风格信息和行人身份信息的编码系数。由于来自同一摄像头视角下的行人风格呈现很强的相关性, 所以用 $\|\cdot\|_*$ (矩阵的核范数) 来约束风格字典。 λ_1, α_1 和 γ 是平衡参数, $\|\mathbf{C}_v^{sc}\|_{2,1} = \sum_i \sqrt{\sum_j (\mathbf{C}_v^{sc})_{i,j}^2}$ 。在公式(1)中, 第一项用 \mathbf{D}_0 来重构行人图像特征, 第二项利用 \mathbf{D} 来重构剩余的風格特征, 第三项和第五项是低秩约束正则项。为了将行人图像中不同的成分进行分离, 表示域信息的风格特征在字典 \mathbf{D} 下有相同的编码系数, 而且能使用相同的原子来表示, 因此使用 $\|\mathbf{C}_v^{sc}\|_{2,1}$ 正则化编码系数 \mathbf{C}_v^{sc} , 这有利于行人图像中不同成分的分离。因原始视觉特征是由编码系数和其对应的字典重构得到, 将编码系数称之为潜在特征。

由于不同视角下同一行人的属性具有域不变性, 如果将训练样本的属性与潜在特征建立一种映射关系, 挖掘潜在信息, 有利于实现目标域样本属性的预测。除此之外, 利用属性作为辅助信息也能提高编码系数的判别性。因此, 将属性引入到行人重识别的模型中, 公式如下

$$\begin{aligned} \min_{\substack{\mathbf{D}, \mathbf{D}_0, \mathbf{C}_v^{sc} \\ \mathbf{C}_v^s, \mathbf{W}}} \sum_{v=1}^M \|\mathbf{X}_v^s - \mathbf{D}_0 \mathbf{C}_v^s\|_F^2 + \|\mathbf{X}_v^s - \mathbf{D} \mathbf{C}_v^{sc} - \mathbf{D}_0 \mathbf{C}_v^s\|_F^2 + \lambda_1 \|\mathbf{X}_v^s - \mathbf{D}_0 \mathbf{C}_v^s\|_* + \\ \beta (\|\mathbf{C}_v^s - \mathbf{W} \mathbf{A}_v^s\|_F^2 + \|\mathbf{W}^T \mathbf{C}_v^s - \mathbf{A}_v^s\|_F^2) + \alpha_1 \|\mathbf{C}_v^{sc}\|_{2,1} + \gamma \|\mathbf{D}\|_* \\ \text{(s.t. } \|\mathbf{d}_i\|_2 \leq 1, \|\mathbf{d}_{0,j}\|_2 \leq 1, \forall i, j), \end{aligned} \quad (2)$$

其中, \mathbf{W} 是转换矩阵, 其作用是建立行人潜在特征空间与属性空间的联系, β 是平衡参数。在公式(2)中, 利用自重构思想来保证行人属性和潜在特征的一致性。

为提高字典的判别性和减少行人特征之间的歧义性, 将标签信息应用到模型当中, 公式如下

$$\begin{aligned} \min_{\substack{\mathbf{D}, \mathbf{D}_0, \mathbf{C}_v^{sc}, \mathbf{C}_v^s \\ \mathbf{C}_v^t, \mathbf{W}}} \sum_{v=1}^M \|\mathbf{X}_v^s - \mathbf{D}_0 \mathbf{C}_v^s\|_F^2 + \|\mathbf{X}_v^s - \mathbf{D} \mathbf{C}_v^{sc} - \mathbf{D}_0 \mathbf{C}_v^s\|_F^2 + \delta \|\mathbf{C}_v^{sc T} \mathbf{W} \mathbf{A}_v^s - \mathbf{Y}_v^s\|_F^2 + \\ \beta (\|\mathbf{C}_v^s - \mathbf{W} \mathbf{A}_v^s\|_F^2 + \|\mathbf{W}^T \mathbf{C}_v^s - \mathbf{A}_v^s\|_F^2) + \lambda_1 \|\mathbf{X}_v^s - \mathbf{D}_0 \mathbf{C}_v^s\|_* + \\ \delta_1 \|\mathbf{W} \mathbf{A}_v^s\|_F^2 + \alpha_1 \|\mathbf{C}_v^{sc}\|_{2,1} + \gamma \|\mathbf{D}\|_* \text{ s.t. } \|\mathbf{d}_i\|_2 \leq 1, \|\mathbf{d}_{0,j}\|_2 \leq 1, \forall i, j, \end{aligned} \quad (3)$$

其中, δ 和 δ_1 是平衡参数, $\mathbf{Y}_v^s = \mathbf{Y}_v^T \mathbf{Y}_v$ 。在公式(3)中, 引入 $\|\mathbf{C}_v^{sc T} \mathbf{W} \mathbf{A}_v^s - \mathbf{Y}_v^s\|_F^2$ 来挖掘标记样本的身份信息, $\|\mathbf{W} \mathbf{A}_v^s\|_F^2$ 来防止过拟合。为了提高所学字典在目标域中的表示能力, 加入一部分目标域样本到源域中训练模型, 公式如下

$$\begin{aligned} \min_{\substack{\mathbf{D}, \mathbf{D}_0, \mathbf{C}_v^{sc}, \mathbf{C}_v^s \\ \mathbf{C}_h^{tc}, \mathbf{C}_h^t, \mathbf{W}}} \sum_{v=1}^M \|\mathbf{X}_v^s - \mathbf{D}_0 \mathbf{C}_v^s\|_F^2 + \|\mathbf{X}_v^s - \mathbf{D} \mathbf{C}_v^{sc} - \mathbf{D}_0 \mathbf{C}_v^s\|_F^2 + \beta (\|\mathbf{C}_v^s - \mathbf{W} \mathbf{A}_v^s\|_F^2 + \|\mathbf{W}^T \mathbf{C}_v^s - \mathbf{A}_v^s\|_F^2) + \\ \delta \|\mathbf{C}_v^{sc T} \mathbf{W} \mathbf{A}_v^s - \mathbf{Y}_v^s\|_F^2 + \lambda_1 \|\mathbf{X}_v^s - \mathbf{D}_0 \mathbf{C}_v^s\|_* + \alpha_1 \|\mathbf{C}_v^{sc}\|_{2,1} + \delta_1 \|\mathbf{W} \mathbf{A}_v^s\|_F^2 + \gamma \|\mathbf{D}\|_* + \\ \sum_{h=1}^L \|\mathbf{X}_h^t - \mathbf{D}_0 \mathbf{C}_h^t\|_F^2 + \|\mathbf{X}_h^t - \mathbf{D} \mathbf{C}_h^{tc} - \mathbf{D}_0 \mathbf{C}_h^t\|_F^2 + \lambda_2 \|\mathbf{X}_h^t - \mathbf{D}_0 \mathbf{C}_h^t\|_* + \alpha_2 \|\mathbf{C}_h^{tc}\|_{2,1} \\ \text{s.t. } \|\mathbf{d}_i\|_2 \leq 1, \|\mathbf{d}_{0,j}\|_2 \leq 1, \forall i, j, \end{aligned} \quad (4)$$

其中, α_2 和 λ_2 是平衡参数, $\mathbf{C}_h^{tc} \in R^{d \times n_h}$ 和 $\mathbf{C}_h^t \in R^{d_0 \times n_h}$ 分别是目标域样本中字典 \mathbf{D} 和 \mathbf{D}_0 对应的编码系数。

2.3 模型优化

对于所有变量来说,公式(4)是非凸函数。然而,固定其它变量而更新其中某一个变量时,模型具有全局最优解。因此采用交替迭代法来获得每个变量的最优解。

1) 固定 $\mathbf{D}, \mathbf{D}_0, \mathbf{C}_v^s, \mathbf{C}_h^t, \mathbf{C}_h^{tc}, \mathbf{W}$, 更新 \mathbf{C}_v^{sc} , 目标函数可以表示为

$$\mathbf{C}_v^{sc} = \arg \min_{\mathbf{C}_v^{sc}} \|\mathbf{X}_v^s - \mathbf{D}\mathbf{C}_v^{sc} - \mathbf{D}_0 \mathbf{C}_v^s\|_F^2 + \alpha_1 \|\mathbf{C}_v^{sc}\|_{2,1}, \quad (5)$$

式(5)具有如下形式的解析解

$$\mathbf{C}_v^{sc} = (\alpha_1 \sum + 2\mathbf{D}^T \mathbf{D})^{-1} (2\mathbf{D}^T \mathbf{X}_v^s - 2\mathbf{D}^T \mathbf{D}_0 \mathbf{C}_v^s), \quad (6)$$

其中 \sum 是第 i 个对角线元素为 $\sigma_{ii} = \frac{1}{\|(\mathbf{C}_v^{sc})_{i,:}\|_2}$ 的对角矩阵。同理求得

$$\mathbf{C}_h^{tc} = (\alpha_2 \mathbf{\Omega} + 2\mathbf{D}^T \mathbf{D})^{-1} (2\mathbf{D}^T \mathbf{X}_h^t - 2\mathbf{D}^T \mathbf{D}_0 \mathbf{C}_h^t), \quad (7)$$

其中 $\mathbf{\Omega}$ 是第 i 个对角线元素为 $\omega_{ii} = \frac{1}{\|(\mathbf{C}_h^{tc})_{i,:}\|_2}$ 的对角矩阵。

2) 固定 $\mathbf{D}, \mathbf{D}_0, \mathbf{C}_v^{sc}, \mathbf{C}_h^t, \mathbf{C}_h^{tc}, \mathbf{W}$, 更新 \mathbf{C}_v^s , 目标函数可表示为

$$\begin{aligned} \mathbf{C}_v^s = \arg \min_{\mathbf{C}_v^s} & \|\mathbf{X}_v^s - \mathbf{D}_0 \mathbf{C}_v^s\|_F^2 + \|\mathbf{X}_v^s - \mathbf{D}\mathbf{C}_v^{sc} - \mathbf{D}_0 \mathbf{C}_v^s\|_F^2 + \lambda_1 \|\mathbf{X}_v^s - \mathbf{D}_0 \mathbf{C}_v^s\|_* + \\ & \beta (\|\mathbf{C}_v^s - \mathbf{W}\mathbf{A}_v^s\|_F^2 + \|\mathbf{W}^T \mathbf{C}_v^s - \mathbf{A}_v^s\|_F^2) + \delta \|\mathbf{C}_v^{scT} \mathbf{W}\mathbf{A}_v^s - \mathbf{Y}_v^s\|_F^2, \end{aligned} \quad (8)$$

为便于求解,引入松弛变量 \mathbf{F}_v 和 \mathbf{F}_v^s , 公式(8)可写为

$$\begin{aligned} [\mathbf{F}_v, \mathbf{F}_v^s, \mathbf{C}_v^s] = \arg \min_{\mathbf{F}_v, \mathbf{F}_v^s, \mathbf{C}_v^s} & \|\mathbf{X}_v^s - \mathbf{D}_0 \mathbf{F}_v^s\|_F^2 + \|\mathbf{X}_v^s - \mathbf{D}\mathbf{C}_v^{sc} - \mathbf{D}_0 \mathbf{F}_v^s\|_F^2 + \lambda_1 \|\mathbf{F}_v^s\|_* + \\ & \beta (\|\mathbf{F}_v^s - \mathbf{W}\mathbf{A}_v^s\|_F^2 + \|\mathbf{W}^T \mathbf{F}_v^s - \mathbf{A}_v^s\|_F^2) + \delta \|\mathbf{C}_v^{scT} \mathbf{W}\mathbf{A}_v^s - \mathbf{Y}_v^s\|_F^2 + \\ & \|\mathbf{X}_v^s - \mathbf{D}_0 \mathbf{F}_v^s - \mathbf{F}_v\|_F^2 + \|\mathbf{C}_v^s - \mathbf{F}_v^s\|_F^2, \end{aligned} \quad (9)$$

其中, \mathbf{F}_v 可由奇异值分解 (SVT) 算法^[17] 求解公式(10)得到

$$\mathbf{F}_v = \arg \min_{\mathbf{F}_v} \|\mathbf{X}_v^s - \mathbf{D}_0 \mathbf{C}_v^s - \mathbf{F}_v\|_F^2 + \lambda_1 \|\mathbf{F}_v\|_*, \quad (10)$$

而 \mathbf{F}_v^s 有解析解

$$\mathbf{F}_v^s = (\bar{\mathbf{D}}_0^T \bar{\mathbf{D}}_0)^{-1} (\bar{\mathbf{D}}_0^T \bar{\mathbf{X}}), \quad (11)$$

其中, $\bar{\mathbf{X}} = [\mathbf{X}_v^s; \mathbf{X}_v^s - \mathbf{D}\mathbf{C}_v^{sc}; \mathbf{X}_v^s - \mathbf{F}_v; \beta \mathbf{W}\mathbf{A}_v^s; \beta \mathbf{A}_v^s; \mathbf{C}_v^s]$, $\bar{\mathbf{D}}_0 = [\mathbf{D}_0; \mathbf{D}_0; \mathbf{D}_0; \beta \mathbf{I}; \beta \mathbf{W}^T; \mathbf{I}]$, \mathbf{I} 为单位阵。获得 $\mathbf{F}_v, \mathbf{F}_v^s$ 后, 可得 \mathbf{C}_v^s 的解析解为

$$\mathbf{C}_v^s = (\delta \mathbf{W}\mathbf{A}_v^s \mathbf{A}_v^{sT} \mathbf{W}^T + \mathbf{I})^{-1} (\delta \mathbf{W}\mathbf{A}_v^s \mathbf{Y}_v^{sT} + \mathbf{F}_v^s), \quad (12)$$

3) 固定 $\mathbf{D}, \mathbf{D}_0, \mathbf{C}_v^{sc}, \mathbf{C}_v^s, \mathbf{C}_h^t, \mathbf{W}$ 更新 \mathbf{C}_h^t , 目标函数可表示为

$$\mathbf{C}_h^t = \arg \min_{\mathbf{C}_h^t} \|\mathbf{X}_h^t - \mathbf{D}_0 \mathbf{C}_h^t\|_F^2 + \|\mathbf{X}_h^t - \mathbf{D}\mathbf{C}_h^{tc} - \mathbf{D}_0 \mathbf{C}_h^t\|_F^2 + \lambda_2 \|\mathbf{X}_h^t - \mathbf{D}_0 \mathbf{C}_h^t\|_*. \quad (13)$$

为便于求解,引入辅助变量 \mathbf{P}_h , 公式(13)可写为

$$\begin{aligned} [\mathbf{C}_h^t, \mathbf{P}_h] = \arg \min_{\mathbf{C}_h^t, \mathbf{P}_h} & \|\mathbf{X}_h^t - \mathbf{D}_0 \mathbf{C}_h^t\|_F^2 + \|\mathbf{X}_h^t - \mathbf{D}\mathbf{C}_h^{tc} - \mathbf{D}_0 \mathbf{C}_h^t\|_F^2 + \\ & \lambda_2 \|\mathbf{P}_h\|_* + \|\mathbf{X}_h^t - \mathbf{D}_0 \mathbf{C}_h^t - \mathbf{P}_h\|_F^2, \end{aligned} \quad (14)$$

其中 \mathbf{P}_h 可通过求解式(15)来得到

$$\mathbf{P}_h = \arg \min_{\mathbf{P}_h} \|\mathbf{X}_h^t - \mathbf{D}_0 \mathbf{C}_h^t - \mathbf{P}_h\|_F^2 + \lambda_2 \|\mathbf{P}_h\|_*, \quad (15)$$

在 \mathbf{P}_h 更新后, 可以求得 \mathbf{C}_h^t 的解析式为

$$\mathbf{C}_h^t = (\bar{\bar{\mathbf{D}}}_0^T \bar{\bar{\mathbf{D}}}_0)^{-1} (\bar{\bar{\mathbf{D}}}_0^T \bar{\bar{\mathbf{X}}}), \quad (16)$$

其中, $\bar{\bar{\mathbf{X}}} = [\mathbf{X}_h^t; \mathbf{X}_h^t - \mathbf{D}\mathbf{C}_h^{tc}; \mathbf{X}_h^t - \mathbf{P}_h]$, $\bar{\bar{\mathbf{D}}}_0 = [\mathbf{D}_0; \mathbf{D}_0; \mathbf{D}_0]$ 。

4) 固定 $\mathbf{D}_0, \mathbf{C}_v^{sc}, \mathbf{C}_v^s, \mathbf{C}_h^t, \mathbf{C}_h^{tc}, \mathbf{W}$, 更新 \mathbf{D} , 此时目标函数可表示为

$$\mathbf{D} = \arg \min_{\mathbf{D}} \|\mathbf{X}^s - \mathbf{D}\mathbf{C}^{sc} - \mathbf{D}_0 \mathbf{C}^s\|_F^2 + \|\mathbf{X}^t - \mathbf{D}\mathbf{C}^{tc} - \mathbf{D}_0 \mathbf{C}^t\|_F^2 + \gamma \|\mathbf{D}\|_*$$

$$\text{s.t. } \|\mathbf{d}_i\|_{\frac{2}{2}} \leq 1, \|\mathbf{d}_{0,j}\|_{\frac{2}{2}} \leq 1, \forall i, j, \quad (17)$$

公式(17)可直接用 SVT 算法来求解,为此对其进行松弛化处理

$$[\mathbf{D}_1, \mathbf{D}] = \arg \min_{\mathbf{D}_1, \mathbf{D}} \|\hat{\mathbf{X}} - \mathbf{D}\mathbf{C}\|_{\frac{2}{F}} + \|\mathbf{D} - \mathbf{D}_1\|_{\frac{2}{F}} + \gamma \|\mathbf{d}_1\|_*, \quad (18)$$

其中 $\hat{\mathbf{X}} = [\mathbf{X}^s - \mathbf{D}_0 \mathbf{C}^s, \mathbf{X}^t - \mathbf{D}_0 \mathbf{C}^t, \mathbf{D}_1]$, $\mathbf{C} = [\mathbf{C}^{sc}, \mathbf{C}^{tc}, \mathbf{I}]$ 。 \mathbf{D}_1 可以直接采用 SVT 算法求解,然后,再对 \mathbf{D} 进行更新,目标函数表达式如下

$$\mathbf{D} = \arg \min_{\mathbf{D}} \|\hat{\mathbf{X}} - \mathbf{D}\mathbf{C}\|_{\frac{2}{F}} \quad \text{s.t. } \|\mathbf{d}_i\|_{\frac{2}{2}} \leq 1, \forall i, \quad (19)$$

公式(19)可使用拉格朗日对偶法^[18]进行求解。

5) 固定 \mathbf{D} , \mathbf{C}_v^{sc} , \mathbf{C}_v^s , \mathbf{C}_h^{tc} , \mathbf{C}_h^t , \mathbf{W} , 更新 \mathbf{D}_0 , 此时目标函数可表示为

$$\begin{aligned} \mathbf{D}_0 = \arg \min_{\mathbf{D}_0} & \|\mathbf{X}^s - \mathbf{D}_0 \mathbf{C}^s\|_{\frac{2}{F}} + \|\mathbf{X}^s - \mathbf{D}_0 \mathbf{C}^{sc} - \mathbf{D}_0 \mathbf{C}^s\|_{\frac{2}{F}} + \lambda_1 \|\mathbf{X}^s - \mathbf{D}_0 \mathbf{C}^s\|_* + \\ & \|\mathbf{X}^t - \mathbf{D}_0 \mathbf{C}^t\|_{\frac{2}{F}} + \|\mathbf{X}^t - \mathbf{D}_0 \mathbf{C}^{tc} - \mathbf{D}_0 \mathbf{C}^t\|_{\frac{2}{F}} + \lambda_2 \|\mathbf{X}^t - \mathbf{D}_0 \mathbf{C}^t\|_* , \\ \text{s.t. } & \|\mathbf{d}_i\|_{\frac{2}{2}} \leq 1, \|\mathbf{d}_{0,j}\|_{\frac{2}{2}} \leq 1, \forall i, j, \end{aligned} \quad (20)$$

为便于求解,引入 2 个变量 \mathbf{H} 和 \mathbf{T} , 公式(20)可表示为

$$\begin{aligned} [\mathbf{D}_0, \mathbf{H}, \mathbf{T}] = \arg \min_{\mathbf{D}_0, \mathbf{H}, \mathbf{T}} & \|\mathbf{X}^s - \mathbf{D}_0 \mathbf{C}^s\|_{\frac{2}{F}} + \|\mathbf{X}^s - \mathbf{D}_0 \mathbf{C}^{sc} - \mathbf{D}_0 \mathbf{C}^s\|_{\frac{2}{F}} + \lambda_1 \|\mathbf{H}\|_* + \\ & \|\mathbf{X}^t - \mathbf{D}_0 \mathbf{C}^t\|_{\frac{2}{F}} + \|\mathbf{X}^t - \mathbf{D}_0 \mathbf{C}^{tc} - \mathbf{D}_0 \mathbf{C}^t\|_{\frac{2}{F}} + \lambda_2 \|\mathbf{T}\|_* + \\ & \|\mathbf{X}^s - \mathbf{D}_0 \mathbf{C}^s - \mathbf{H}\|_{\frac{2}{F}} + \|\mathbf{X}^t - \mathbf{D}_0 \mathbf{C}^t - \mathbf{T}\|_{\frac{2}{F}}, \end{aligned} \quad (21)$$

其中,更新 \mathbf{H} 和 \mathbf{T} 的目标函数分别表示为

$$\mathbf{H} = \arg \min_{\mathbf{H}} \|\mathbf{X}^s - \mathbf{D}_0 \mathbf{C}^s - \mathbf{H}\|_{\frac{2}{F}} + \lambda_1 \|\mathbf{H}\|_* , \quad (22)$$

$$\mathbf{T} = \arg \min_{\mathbf{T}} \|\mathbf{X}^t - \mathbf{D}_0 \mathbf{C}^t - \mathbf{T}\|_{\frac{2}{F}} + \lambda_2 \|\mathbf{T}\|_* , \quad (23)$$

公式(22)和(23)均可采用奇异值分解(SVT)算法来求解。然后,更新 \mathbf{D}_0 , 其目标函数为

$$\begin{aligned} \mathbf{D}_0 = \arg \min_{\mathbf{D}_0} & \|\mathbf{X}^s - \mathbf{D}_0 \mathbf{C}^s\|_{\frac{2}{F}} + \|\mathbf{X}^s - \mathbf{D}_0 \mathbf{C}^{sc} - \mathbf{D}_0 \mathbf{C}^s\|_{\frac{2}{F}} + \|\mathbf{X}^t - \mathbf{D}_0 \mathbf{C}^t\|_{\frac{2}{F}} + \\ & \|\mathbf{X}^t - \mathbf{D}_0 \mathbf{C}^{tc} - \mathbf{D}_0 \mathbf{C}^t\|_{\frac{2}{F}} + \|\mathbf{X}^s - \mathbf{D}_0 \mathbf{C}^s - \mathbf{H}\|_{\frac{2}{F}} + \|\mathbf{X}^t - \mathbf{D}_0 \mathbf{C}^t - \mathbf{T}\|_{\frac{2}{F}}, \\ \text{s.t. } & \|\mathbf{d}_{0,j}\|_{\frac{2}{2}} \leq 1, \end{aligned} \quad (24)$$

可将公式(24)简化为

$$\mathbf{D}_0 = \arg \min_{\mathbf{D}_0} \|\tilde{\mathbf{X}} - \mathbf{D}_0 \tilde{\mathbf{C}}_0\|_{\frac{2}{F}} \quad \text{s.t. } \|\mathbf{d}_{0,j}\|_{\frac{2}{2}} \leq 1, \quad (25)$$

公式(25)可直接由拉格朗日对偶法来求解。其中, $\tilde{\mathbf{X}} = [\mathbf{X}^s, \mathbf{X}^s - \mathbf{D}_0 \mathbf{C}^{sc}, \mathbf{X}^s - \mathbf{H}, \mathbf{X}^t, \mathbf{X}^t - \mathbf{D}_0 \mathbf{C}^{tc}, \mathbf{X}^t - \mathbf{T}]$,

$\tilde{\mathbf{C}}_0 = [\mathbf{C}^s, \mathbf{C}^s, \mathbf{C}^s, \mathbf{C}^t, \mathbf{C}^t, \mathbf{C}^t]$ 。

6) 固定 $\mathbf{D}, \mathbf{D}_0, \mathbf{C}_v^{sc}, \mathbf{C}_v^s, \mathbf{C}_h^{tc}, \mathbf{C}_h^t$, 更新 \mathbf{W}

$$\begin{aligned} \mathbf{W} = \arg \min_{\mathbf{W}} & \beta (\|\mathbf{C}^s - \mathbf{W}\mathbf{A}^s\|_{\frac{2}{F}} + \|\mathbf{W}^T \mathbf{C}^s - \mathbf{A}^s\|_{\frac{2}{F}}) + \\ & \delta \|\mathbf{C}^{sT} \mathbf{W}\mathbf{A}^s - \mathbf{Y}^s\|_{\frac{2}{F}} + \delta_1 \|\mathbf{W}\mathbf{A}^s\|_{\frac{2}{F}}. \end{aligned} \quad (26)$$

为便于求解,公式(26)可松弛化为

$$\begin{aligned} [\mathbf{W}, \mathbf{L}] = \arg \min_{\mathbf{W}, \mathbf{L}} & \beta (\|\mathbf{C}^s - \mathbf{W}\mathbf{A}^s\|_{\frac{2}{F}} + \|\mathbf{W}^T \mathbf{C}^s - \mathbf{A}^s\|_{\frac{2}{F}}) + \\ & \delta \|\mathbf{C}^{sT} \mathbf{W}\mathbf{A}^s - \mathbf{Y}^s\|_{\frac{2}{F}} + \delta_1 \|\mathbf{L}\|_{\frac{2}{F}} + \|\mathbf{L} - \mathbf{W}\mathbf{A}^s\|_{\frac{2}{F}}, \end{aligned} \quad (27)$$

从而得到更新 \mathbf{L} 的解析表达式

$$\mathbf{L} = (\tilde{\mathbf{C}}^T \tilde{\mathbf{C}})^{-1} (\tilde{\mathbf{C}}^T \tilde{\mathbf{C}}^s), \quad (28)$$

其中: $\tilde{\mathbf{C}}^s = [\beta \mathbf{C}^s; \delta \mathbf{Y}^s; 0; \mathbf{W}\mathbf{A}^s]$; $\tilde{\mathbf{C}} = [\beta \mathbf{I}; \delta \mathbf{C}^{sT}; \delta_1 \mathbf{I}; \mathbf{I}]$, \mathbf{I} 是单位矩阵。最后得到更新 \mathbf{W} 的公式为

$$\beta \mathbf{C}^s \mathbf{C}^{sT} \mathbf{W} + \mathbf{W}\mathbf{A}^s \mathbf{A}^{sT} = \beta \mathbf{C}^s \mathbf{A}^{sT} + \mathbf{L} \mathbf{A}^{sT}. \quad (29)$$

公式(29)是 Sylvester 方程,利用 Bartels-stewart 算法^[19]求解。为便于理解,将求解公式(4)的优化算法

总结在表 1 中。

表 1 域自适应行人重识别低秩分解字典学习算法
Table 1 Low rank decomposition dictionary learning algorithm for domain adaptation person re-identification

算法 1: 域自适应行人重识别低秩分解字典学习算法	
输入: X^s, X^t, Y^s, A^s	
1) 随机初始化 $D, D_0, C_v^{sc}, C_v^s, C_h^{tc}, C_h^t, W$ 。	
循环迭代	
2) 通过公式 (6), 更新 C_v^{sc} ,	3) 通过公式 (7), 更新 C_h^{tc}
4) 通过公式 (12), 更新 C_v^s ,	5) 通过公式 (16), 更新 C_h^t
6) 通过公式 (19), 更新 D ,	7) 通过公式 (25), 更新 D_0
8) 通过公式 (29), 更新 W 。	
结束迭代	
输出: D, D_0, W	

2.4 相似性度量

得到风格字典 D 和行人身份特征字典 D_0 后, 通过公式 (30), 可以得到目标域风格编码系数 C_h^{tc} 和行人身份特征的编码系数 C_h^t

$$[C_h^t, C_h^{tc}] = \arg \min_{C_h^t, C_h^{tc}} \|X_h^t - D_0 C_h^t\|_F^2 + \|X_h^t - D C_h^{tc} - D_0 C_h^t\|_F^2 + \lambda_2 \|X_h^t - D_0 C_h^t\|_*, \quad (30)$$

同时, 得到映射矩阵 W 后, 可预测目标域的属性特征 A^t

$$A^t = (W^T W)^{-1} (W^T C^t), \quad (31)$$

在得到测试样本行人身份信息编码系数 C^t , 行人属性表示 A^t 后, 通过余弦距离^[20] 得到行人图像对的相似性分数

$$\text{sim}(z_a, z_b) = \frac{z_a^T \cdot z_b}{\|z_a\|_2 \cdot \|z_b\|_2 \cdot (\|z_a - z_b + \epsilon\|_2^2)}, \quad (32)$$

其中, $z_l (l=a, b)$ 表示语义属性空间或视觉特征空间中的一个向量, $\epsilon > 0$ 是一个常数。为综合潜在视觉特征空间和语义属性空间的互补性, 最终的相似分数可由下式来求得

$$\text{sim}_{\text{final}} = \tau \text{sim}_C + (1 - \tau) \text{sim}_A, \quad (33)$$

其中 τ 为权值, sim_C 和 sim_A 分别表示潜在特征相似性分数和属性相似性分数。由于属性的判别性较弱, 因此取 $\tau = 0.8$ 。

3 实验

3.1 数据集和实验的设置

为了验证算法的有效性, 选择了 4 个具有挑战性的数据集: VIPeR^[21], PRID2011^[22], CUHK01^[23], GRID^[24]。其中, VIPeR, PRID2011 和 GRID 数据集有 2 个摄像头且每个行人身份在一个摄像头视角中只有一张图像。PRID2011 和 GRID 都包含干扰图像, 因此, 更接近真实场景。CUHK01 也有 2 个摄像头, 但每个身份在一个摄像头下有 2 张图像。具体的数据集描述如表 2 所示。

在上述的数据集中, 只有 VIPeR, PRID2011, GRID 标注了语义属性。在算法中, 这些被标注了语义属性的数据集可以用作训练集, 而未标注语义属性的数据集可以用作测试集。具体来说, 将上述的单个数据集的全部样本用来训练模型, 而目标数据集中的样本按行人身份分成互不重叠的 2 个子集, 一个参与模型的训练, 另一个用于模型的测试, 随机取 10 组数据的平均值作为实验结果。测试时, 所有的干扰图像都会被用作测试集样本。该模型一共有 10 个参数, 其中包括字典 D 和 D_0 的大小 d 和 d_0 , 和 8 个超参数 $\alpha_1, \alpha_2, \lambda_1, \lambda_2, \delta, \delta_1, \beta, \gamma$ 。设定 $d = 650, d_0 = 800, \alpha_1 = 0.1, \alpha_2 = 0.1, \lambda_1 = 0.6, \lambda_2 = 0.5, \delta = 5, \delta_1 = 6, \beta = 0.15, \gamma = 0.5$ 。在特征的选择上, 使用 Gaussian of Gaussian (GOG) 描述符^[25] 来描述行人的外貌特征。在属性表示上, 使用的 VIPeR,

PRID2011, GRID 数据集上的属性定义是由 R.Layne 等人^[26]所标注的。

表 2 实验数据集的详细设置

Table 2 Details of dataset settings for experiment

数据集	摄像头数	行人身份数	训练的行人身份数	测试的行人身份数	干扰图像的身份数	图像数量
VIPeR	2	632	316	316	0	1 264
PRID2011	2	749	100	100	549	949
GRID	2	1025	125	125	775	1 275
CUHK01	2	971	486	485	0	3 884

3.2 对比实验结果

在 VIPeR 上的实验:将 PRID2011 作为源数据集, VIPeR 作为目标数据集,实验结果与 UMDL^[9], SAAVF^[11], MFFAG^[12], ADV^[27], GL^[28], SDC^[29], CAMEL^[30], DECAMEL^[31], AIESL^[32], SNR^[33] 比较,对比结果如表 3 所示。由此可以看出,在 VIPeR 上,研究方法的性能超过了深度学习的方法 DECAMEL^[31] 和传统的无监督域自适应行人重识别方法 UMDL^[9], SAAVF^[11], MFFAG^[12], ADV^[27], GL^[28], SDC^[29], CAMEL^[30], AIESL^[32], SNR^[33]。证明了方法的有效性和优越性。

表 3 在 VIPeR 数据集上的识别性能

Table 3 Performance on the VIPeR dataset

方法	Rank1	rank5	rank10	rank20
UMDL ^[9]	31.50	—	—	—
SAAVF ^[11]	26.84	39.72	49.27	60.38
MFFAG ^[12]	33.70	—	—	—
ADV ^[27]	22.80	38.60	50.30	63.90
GL ^[28]	33.50	—	—	—
SDC ^[29]	25.80	—	—	—
CAMEL ^[30]	30.90	—	—	—
DECAMEL ^[31]	20.30	38.04	49.11	60.38
AIESL ^[32]	28.92	40.41	46.58	52.44
SNR ^[33]	32.60	—	—	—
Ours	34.50	49.18	65.73	76.61

在 GRID 上的实验:将 VIPeR 作为源数据集, GRID 作为目标数据集,实验结果与 DIMN^[4], AIESL^[32], SNR^[33], TLSTP^[34] 进行了比较,对比结果如表 4 所示。由于 GRID 数据集携带了大量干扰图像,因此在该数据集上进行行人身份的匹配充满了挑战。由表 4 中的数据可以看出,提出的方法在性能超过了传统无监督域自适应行人重识别方法 AIESL^[32], SNR^[33], TLSTP^[34] 和深度学习方法 DIMN^[4]。

在 CUHK01 上的实验:为进一步验证算法的有效性,将 VIPeR 用作源数据集, CUHK01 用作目标数据集,实验结果与 UDML^[9], MFFAG^[12], CAMEL^[30], DECAMEL^[31], AIESL^[32], TSR^[35], DAS^[36], UJSDL^[37] 进行比较,对比结果如表 5 所示。由此可以看出,提出的方法在性能上超过了传统的无监督域自适应行人重识别方法 UDML^[9], MFFAG^[12], CAMEL^[30], AIESL^[32], TSR^[35], DAS^[36], UJSDL^[37] 和深度学习方法 DECAMEL^[31]。

表 4 在 GRID 数据集上的识别性能

Table 4 Performance on the GRID dataset

方法	rank1	rank5	rank10	rank20
DIMN ^[4]	29.28	53.28	65.84	—
AIESL ^[32]	16.80	29.60	37.20	46.80
SNR ^[33]	26.00	—	—	—
TLSTP ^[34]	28.40	65.60	80.40	—
Ours	38.56	57.68	67.68	78.16

表 5 在 CUHK01 数据集上的识别性能

Table 5 Performance on the CUHK01 dataset

方法	rank1	rank5	rank10	rank20
UDML ^[9]	27.10	—	—	—
MFAG ^[12]	58.10	—	—	—
CAMEL ^[30]	57.30	—	—	—
DECAMEL ^[31]	65.81	—	—	—
AIESL ^[32]	63.26	81.63	87.36	91.84
TSR ^[35]	22.40	35.90	47.90	64.50
DAS ^[36]	54.90	—	—	—
UJSDL ^[37]	27.74	48.81	57.71	66.85
Ours	67.34	86.00	91.75	95.20

3.3 算法分析

3.3.1 收敛性分析

目标函数(4)对变量 $\{\mathbf{D}, \mathbf{D}_0, \mathbf{W}\}$ 来说是一个非凸函数,然而固定其它变量而更新其中某个变量时,模型具有全局最优解。在 VIPeR 数据集上对算法的收敛性进行了验证,其结果如图 3 所示。由此可发现,对于变量 $\{\mathbf{D}, \mathbf{D}_0, \mathbf{W}\}$ 而言,算法是收敛的,而且当迭代次数达到 10 次的时候,算法能达到稳定解。因此,将迭代次数设置为 10。

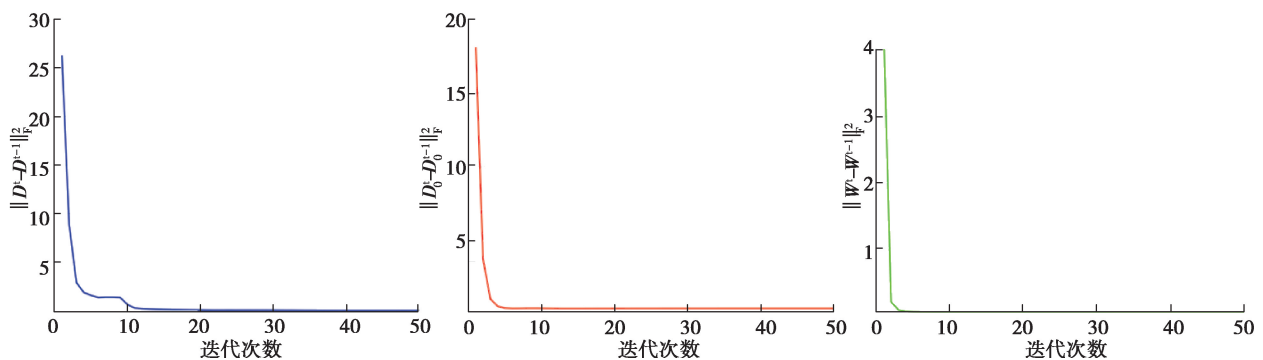


图 3 算法在 VIPeR 数据集上的收敛性分析

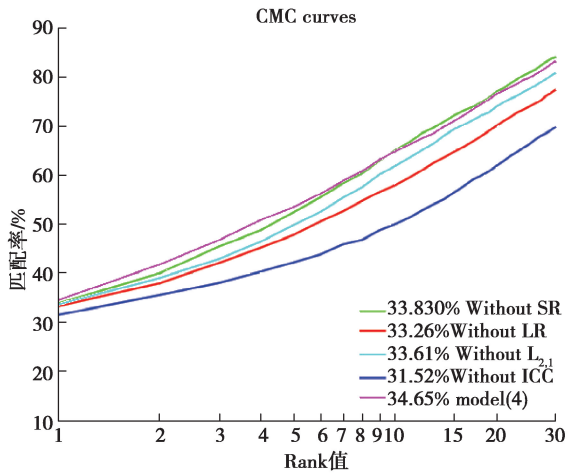
Fig. 3 Convergence analysis of the algorithm on VIPeR dataset

3.3.2 算法的复杂度分析

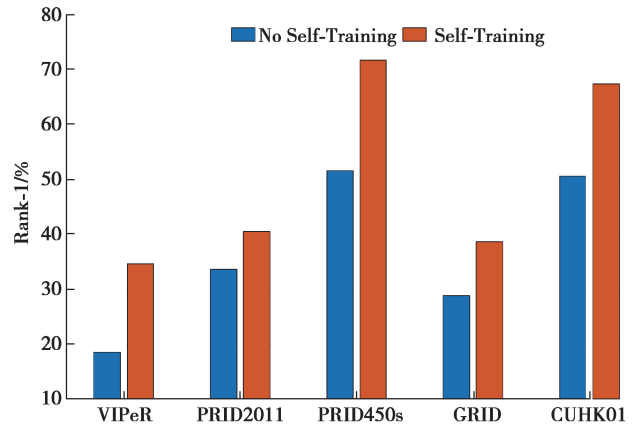
算法主要包含训练、测试、自训练 3 个阶段。在训练中,算法的复杂度主要来源于核范数以及 Sylvester 方程的求解。在迭代 k 次的情况下,求解 Sylvester 方程的复杂度为 $kO(d_0^3)$ 。在求解公式(10)时,SVT 算法的复杂度为 $O(mn_v^2)$,求解公式(19)的算法复杂度为 $O(mn_h^2)$,求解公式(23)的算法复杂度为 $O(md^2)$ 。在算法 1 迭代了 k 次的情况下,训练阶段总复杂度为 $kO(d_0^3) + kO(mn_v^2) + kO(mn_h^2) + kO(md^2)$ 。在测试阶段,算法复杂度的主要来源仍然是核范数的求解,具体为 $O(mn_h^2)$ 。在自训练中,假设有 M 轮训练阶段,每轮需要迭代 K 次,且每轮训练都包含训练阶段和测试阶段,所以算法总的复杂度为 $O(MKk(d_0^3 + mn_v^2 + mn_h^2 + md^2) + Mmn_h^2)$ 。

3.3.3 消融性分析

公式(4)中包含自重构约束(SR) $\|C_v^s - WA_v^s\|_F^2 + \|W^T C_v^s - A_v^s\|_F^2$,低秩约束(LR) $\|X_v^s - D_0 C_v^s\|_* + \|X_h^s - D_0 C_h^s\|_*$ 和 $\|D\|_*$, $L_{2,1}$ 范数约束 $\|C_v^{sc}\|_{2,1} + \|C_h^{sc}\|_{2,1}$ 和身份一致性约束(ICC) $\|C_v^{sT} WA_v^s - Y_v^s\|_F^2$ 。为说明这 4 个约束项的作用,交替将其从公式(4)中移除,然后再与公式(4)的性能进行对比。图 4(a)表示缺少各部分约束项的情况下,模型识别性能与最终模型在 VIPeR 数据集上的实验结果。从中可以看出,去除低秩约束后,rank1 为 33.26%,去除自重构约束项后,rank1 为 33.83%,去除 $L_{2,1}$ 范数约束后,rank1 值为 33.61%,去除身份一致性约束后,rank1 值为 31.52%,由此可以证明研究方法的每一项约束对识别性能的提高都是有积极作用。图 4(b)表示在有自训练和没有自训练情况下模型的识别性能,从中可以看出自训练对实验性能的提升起到很大的帮助。



(a) 正则项的有效性分析



(b) 自训练的有效性分析

图 4 算法的消融性分析

Fig. 4 Ablation analysis of the algorithm

3.4 参数的选择

在算法中共有 8 个超参数 $\alpha_1, \alpha_2, \beta, \lambda_1, \lambda_2, \delta, \delta_1, \gamma$ 和字典的尺寸大小 d, d_0 需要选择,根据实验经验,设置 $d=650, d_0=800, \alpha_1=0.1, \alpha_2=0.1, \delta_1=6, \gamma=0.5$ 。其余主要模块参数 $\beta, \lambda_1, \lambda_2, \delta$ 的设置采用交叉验证方法来进行选择。在此过程中,改变其中某个参数,其余参数固定,研究该参数对识别性能的影响,找到最优参数值。对于 $\beta, \delta, \lambda_1, \lambda_2$ 来说,其变化范围从 10^{-6} 到 10^3 ,每次以 10 的倍数递增。图 5(a)—(d)给出了这 4 个参数取不同数值时的识别性能,由此可以看出,当 $\lambda_1=0.6, \lambda_2=0.5, \delta=5, \beta=0.15$ 时,算法能取得较为满意的结果。

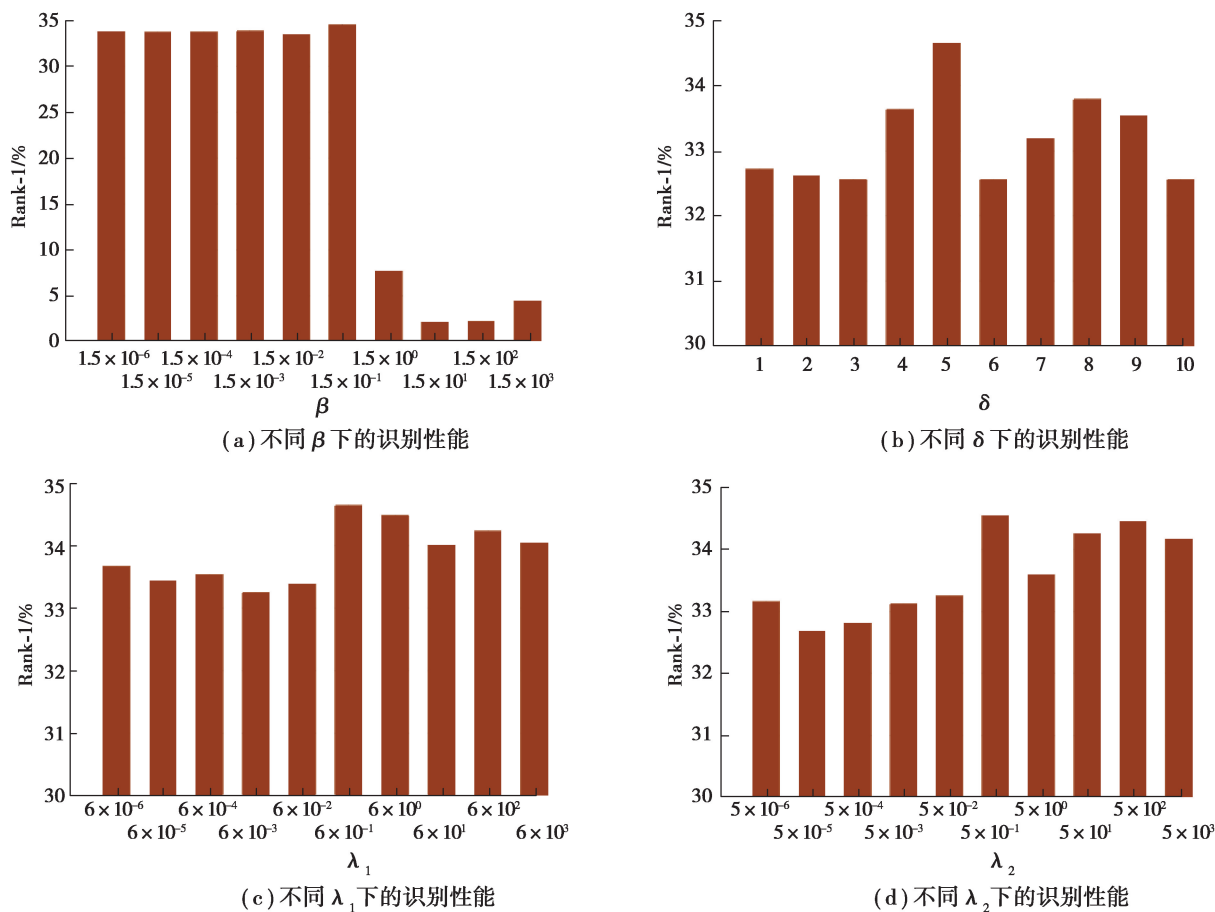


图 5 算法在 VIPeR 数据集上的参数分析

Fig. 5 Parameters sensitivity analysis of the algorithm on the VIPeR dataset

4 结 论

根据同一视角下行人图像所表现的低秩先验性,提出对风格鲁棒的无监督域自适应行人重识别方法。该方法将反映图像相机风格的信息从图像特征中分离出去,极大缓解数据集之间由风格差异引起的域偏移给跨数据集行人重识别带来的影响。在算法中,为建立目标数据集与源数据集之间的联系,提出联合学习算法,将行人身份特征信息、语义属性和标签信息嵌入到一个字典学习模型中。为了有效利用目标数据集中的信息,将目标数据集中被分配伪标签且置信度较高的样本选出来调整已训练模型参数。在域自适应行人重识别任务中,算法比传统的无监督域自适应行人重识别方法及部分基于深度学习的无监督域自适应行人重识别方法表现出较强的竞争力。

参考文献:

- [1] Zhong Z, Zheng L, Luo Z M, et al. Invariance matters: exemplar memory for domain adaptive person Re-identification [C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 15-20, 2019, Long Beach, CA, USA; IEEE, 2019; 598-607.
- [2] Yang Q Z, Yu H X, Wu A C, et al. Patch-based discriminative feature learning for unsupervised person Re-identification [C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 15-20, 2019, Long Beach, CA, USA; IEEE, 2019; 3628-3637.
- [3] Fu Y, Wei Y C, Wang G S, et al. Self-similarity grouping: a simple unsupervised cross domain adaptation approach for person Re-identification [C]//2019 IEEE/CVF International Conference on Computer Vision (ICCV), October 27 -

- November 2, 2019, Seoul, Korea (South); IEEE, 2019: 6111-6120.
- [4] Song J F, Yang Y X, Song Y Z, et al. Generalizable person re-identification by domain-invariant mapping network[C]// 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). June 15-20, 2019, Long Beach, CA, USA; IEEE, 2019: 719-728.
- [5] Wei L H, Zhang S L, Gao W, et al. Person transfer GAN to bridge domain gap for person Re-identification[C]// 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. June 18-23, 2018, Salt Lake City, UT, USA; IEEE, 2018: 79-88.
- [6] Lv J, Wang X T. Cross-dataset person Re-identification using similarity preserved generative adversarial networks[C]// Conference on Knowledge Science, Engineering and Management. Berlin: Springer 2018: 171-183. DOI:10.1007/978-3-319-99247-1_15.
- [7] Deng W J, Zheng L, Ye Q X, et al. Image-image domain adaptation with preserved self-similarity and domain-dissimilarity for person Re-identification [C] // 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. June 18-23, 2018, Salt Lake City, UT, USA; IEEE, 2018: 994-1003.
- [8] Li Y J, Yang F E, Liu Y C, et al. Adaptation and Re-identification network: an unsupervised deep transfer learning approach to person Re-identification [C] // 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). June 18-22, 2018, Salt Lake City, UT, USA; IEEE, 2018: 285-2856.
- [9] Peng P X, Xiang T, Wang Y W, et al. Unsupervised cross-dataset transfer learning for person Re-identification[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). June 27-30, 2016, Las Vegas, NV, USA; IEEE, 2016: 1306-1315.
- [10] 雷大江, 滕君, 王明达, 等. 基于卡方核的正则化线性判别行人再识别算法[J]. 重庆大学学报, 2018, 41(09), 41: 66-76.
Lei D J, Teng J, Wang M D, et al. Chi square kernel regularized linear discriminant analysis for person reidentification [J]. Journal of Chongqing University, 2018, 41(09), 41: 66-76. (in Chinese)
- [11] Li H F, Kuang Z Y, Yu Z T, et al. Structure alignment of attributes and visual features for cross-dataset person re-identification[J]. Pattern Recognition, 2020, 106: 107414.
- [12] Zhou R W, Chang X J, Shi L, et al. Person reidentification via multi-feature fusion with adaptive graph learning[J]. IEEE Transactions on Neural Networks and Learning Systems, 2020, 31(5): 1592-1601.
- [13] Wang J Y, Zhu X T, Gong S G, et al. Transferable joint attribute-identity deep learning for unsupervised person Re-identification[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. June 18-23, 2018, Salt Lake City, UT, USA; IEEE, 2018: 2275-2284.
- [14] Franco A, Oliveira L. Convolutional covariance features: conception, integration and performance in person re-identification[J]. Pattern Recognition, 2017, 61: 593-609.
- [15] Schumann A, Stiefelhagen R. Person Re-identification by deep learning attribute-complementary information[C]// 2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). July 21-26, 2017, Honolulu, HI, USA; IEEE, 2017: 1435-1443.
- [16] Lin Y T, Zheng L, Zheng Z D, et al. Improving person re-identification by attribute and identity learning[J]. Pattern Recognition, 2019, 95: 151-161.
- [17] Lei Y W, Zhou D X. Analysis of singular value thresholding algorithm for matrix completion[J]. Journal of Fourier Analysis and Applications, 2019, 25(6): 2957-2972.
- [18] Lee H, Battle A, Raina R, et al. Efficient sparse coding algorithms[C]// Advances in Neural Information Processing Systems (NIPS), Canada; MIT Press, 2006: 801-808.
- [19] Bartels R H, Stewart G W. Solution of the matrix equation $ax + xb = C$ [J]. ACM Communications, 1972, 15(9): 820-826.
- [20] Li H F, Zhu J T, Tao D P. Asymmetric projection and dictionary learning with listwise and identity consistency constraints for person Re-identification[J]. IEEE Access, 2018, 6: 37977-37990.
- [21] Gray D, Tao H. Viewpoint invariant pedestrian recognition with an ensemble of localized features [C] // European Conference on Computer Vision - ECCV 2008, Berlin; Springer, 2008: 262-275. DOI:10.1007/978-3-540-88682-2_21.
- [22] Hirzer M, Belezni C, Roth P M, et al. Person Re-identification by descriptive and discriminative classification [C] // Scandinavian Conference on Image Analysis, Berlin; Springer, 2011: 91-102. DOI:10.1007/978-3-642-21227-7_9.

- [23] Li W, Zhao R, Wang X G. Human reidentification with transferred metric learning[C]// Asian Conference on Computer Vision - ACCV 2012, Berlin: Springer, 2013: 31-44. DOI:10.1007/978-3-642-37331-2_3.
- [24] Loy C C, Liu C X, Gong S G. Person re-identification by manifold ranking[C]// 2013 IEEE International Conference on Image Processing. September 15-18, 2013, Melbourne, VIC, Australia: IEEE, 2013: 3567-3571.
- [25] Matsukawa T, Okabe T, Suzuki E, et al. Hierarchical Gaussian descriptor for person re-identification[C]// 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). June 27-30, 2016, Las Vegas, NV, USA: IEEE, 2016: 1363-1372.
- [26] Layne R, Hospedales T, Gong S. Person re-identification by attributes[C]// British Machine Vision Conference (BMVC), Guildford: BMVA, 2012, 2(3): 1-11.
- [27] Ganin Y, Ustinova E, Ajakan H, et al. Domain-adversarial training of neural networks[J]. The Journal of Machine Learning Research, 2016, 17(1): 2096-2030.
- [28] Kodirov E, Xiang T, Fu Z Y, et al. Person re-identification by unsupervised l_1 graph learning[M]// Computer Vision - ECCV 2016. Cham: Springer International Publishing, 2016: 178-195.
- [29] Zhao R, Oyang W, Wang X G. Person re-identification by saliency learning[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(2): 356-370.
- [30] Yu H X, Wu A C, Zheng W S. Cross-view asymmetric metric learning for unsupervised person Re-identification[C]// 2017 IEEE International Conference on Computer Vision (ICCV). October 22-29, 2017, Venice, Italy: IEEE, 2017: 994-1002.
- [31] Yu H X, Wu A C, Zheng W S. Unsupervised person re-identification by deep asymmetric metric embedding[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020, 42(4): 956-973.
- [32] Li H F, Yan S L, Yu Z T, et al. Attribute-identity embedding and self-supervised learning for scalable person re-identification[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2020, 30(10): 3472-3485.
- [33] Jin X, Lan C L, Zeng W J, et al. Style normalization and restitution for generalizable person re-identification[C]// 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). June 13-19, 2020, Seattle, WA, USA: IEEE, 2020: 3140-3149.
- [34] Lv J, Chen W H, Li Q, et al. Unsupervised cross-dataset person re-identification by transfer learning of spatial-temporal patterns[C]// 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. June 18-23, 2018, Salt Lake City, UT, USA: IEEE, 2018: 7948-7956.
- [35] Shi Z Y, Hospedales T M, Xiang T. Transferring a semantic representation for person re-identification and search[C]// 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). June 7-12, 2015, Boston, MA, USA: IEEE, 2015: 4184-4193.
- [36] Bakk S, Carr P, Lalonde J F. Domain adaptation through synthesis for unsupervised person Re-identification[C]// Computer Vision - ECCV 2018, Berlin: Springer, 2018: 193-209. DOI:10.1007/978-3-030-01261-8_12.
- [37] Qi L, Huo J, Fan X C, et al. Unsupervised joint subspace and dictionary learning for enhanced cross-domain person Re-identification[J]. IEEE Journal of Selected Topics in Signal Processing, 2018, 12(6): 1263-1275.

(编辑 侯 湘)