

跨模态行人重识别研究与展望^①



陈丹¹, 李永忠¹, 于沛泽², 邵长斌^{1,2}

¹(江苏科技大学 计算机科学与技术系, 镇江 212003)

²(南京大学 计算机学院, 南京 210023)

通讯作者: 李永忠, E-mail: 443482301@qq.com

摘要: 行人重识别是计算机视觉的热门研究方向, 其对智能安防、视频监控的发展有着重大意义. 目前大部分工作主要集中在研究基于可见光的行人重识别, 然而可见光摄像头无法在光线不足的黑夜中正常使用, 而新型摄像头能够随机切换红外模式进行 24 小时视频监控, 因此最近有一些工作对 RGB-IR 跨模态行人重识别问题进行了研究. 本文分别从定义、研究难点和发展现状介绍了跨模态行人重识别问题, 并根据不同的技术类型将目前存在的方法分为三类, 即基于统一特征模型的方法; 基于度量学习的方法; 基于模态转换的方法. 本文也详细介绍了该任务的数据集和评价准则, 并对现有算法的性能进行分析与归纳. 最后, 总结了跨模态行人重识别的未来发展方向.

关键词: 跨模态行人重识别; 红外图像; 统一特征模型; 度量学习; 模态转换

引用格式: 陈丹, 李永忠, 于沛泽, 邵长斌. 跨模态行人重识别研究与展望. 计算机系统应用, 2020, 29(10): 20-28. <http://www.c-s-a.org.cn/1003-3254/7621.html>

Research and Prospect of Cross Modality Person Re-Identification

CHEN Dan¹, LI Yong-Zhong¹, YU Pei-Zei², SHAO Chang-Bin^{1,2}

¹(School of Computer Science, Jiangsu University of Science and Technology, Zhenjiang 212003, China)

²(School of Computer Science, Nanjing University, Nanjing 210023, China)

Abstract: Person re-identification (Re-ID) has attract lots of attention in computer vision, which is of great significance to the development of intelligent security and video surveillance. Currently, most existing methods focus on the person re-identification based on visible light, and have achieved good performance. However, the visible light camera cannot be used normally in the dark night, and the new generation of cameras can automatically switch the mode between infrared and visible settings for 24 hours monitoring. Therefore, some scholars have started to study the RGB-IR cross-modality pedestrian re-identification. This paper introduces the Re-ID and cross-modality Re-ID respectively from the definition, research difficulties, and development status. For RGB-IR cross-modality Re-ID, according to the types of methods, they are divided into three categories: methods based on unified feature models; methods based on metric learning; and methods based on modal transformation. We also describe the corresponding datasets and evaluation protocol. Besides, we analyze and summarize the performance of existing algorithms. Finally, the future development directions of RGB-IR cross-modality Re-ID are summarized.

Key words: cross modality person re-identification; infrared image; unified feature model; metric learning; modality transformation

① 基金项目: 国家自然科学基金 (61471182); 江苏省研究生创新工程 (CXLX13_70, KYCX17_1845)

Foundation item: National Natural Science Foundation of China (61471182); Graduates Innovation Program of Jiangsu Province (CXLX13_70, KYCX17_1845)

收稿时间: 2020-03-14; 采用时间: 2020-04-12; csa 在线出版时间: 2020-09-30

1 引言

随着社会的进步与发展,人们的安全意识逐渐加强,对公共安全的要求也越来越高,大量的摄像头被安装在街道、商场、高铁站、电影院等公共场所进行监控.由于大型摄像网络的建立,视频信息数据量十分庞大,使得仅仅依靠人力来准确有效地定位人的行踪或通过摄像机跟踪一个人变得极其困难,传统的人工处理和识别监控视频的方法已经无法适应现在网络发展的趋势.随着人工智能和深度学习的兴起,出现了利用机器处理视频数据,在跨摄像镜头下识别行人的技术,称为行人重识别 (person Re-Identification, 简称 Re-ID).行人重识别技术能够快速有效的寻找特定的行人,在实际生活中应用性很强,由此得到了学术界的广泛关注,成为计算机视觉领域的一个重要的研究热点.

先前大部分的研究工作主要关注可见光的行人重识别问题,并且取得了很大的成功.然而,在现实生活中,犯罪分子往往都是夜间作案,可见光摄像机无法捕捉到有效的行人外观特征.随着科技的进步,大量的红外摄像头在视频监控系统中投入使用,利用红外线捕捉行人的外貌特征,帮助执行监察工作,大大地提高了破案的效率.因此,RGB-IR跨模态行人重识别技术被提出,解决该问题对公共安全和刑侦有着非常重要的现实意义^[1],在加强社会管理、预防犯罪行为发生、维护国家安全等方面具有广阔的应用前景.

2 介绍

2.1 行人重识别

2.1.1 定义

行人重识别是图像检索的一个子任务,目的是在一系列由独立监控摄像头拍摄的图像中寻找特定的人^[2],即判断跨镜头下是否是同一个行人,如图1所示.大部分的识别工作关注于RGB-RGB的图像匹配,所以行人重识别也叫做单模态下行行人重识别.

2.1.2 面临的问题

行人重识别技术在现实生活中受到图像分辨率低、不同的光照条件及视角、行人姿态变化以及外界遮挡等许多挑战.在这些因素下,即使是同一个行人,在不同的摄像头下也会造成很大的外观差异,难以区分.

2.1.3 发展现状

单模态行人重识别的研究主要有两个关键点,一是特征提取,即对目标行人图像和候选行人图像进行

学习,提取出具有鲁棒性的行人特征;二是度量学习,即计算两者特征向量之间的距离,比较它们的相似性.早期的工作主要利用颜色直方图、Gabor特征^[3]、HOG特征^[4]、LBP^[5]、颜色域^[6]、SIFT特征^[7]等方法,以手工方式提取行人特征,再利用LMNN^[8]、PRDC^[9]、KISSME^[10]、RDC^[11]、LFDA^[12]、XQDA^[13]等算法进行相似性度量学习.但是由于人工方式的局限性,难以适用当今社会的大数据任务,取得的成果也不太理想.直到2012年,卷积神经网络在ImageNet^[14](ILSVRC)大型视觉识别大赛上获得了冠军,由此以卷积神经网络为代表的深度学习开始流行起来.Li等首次将深度学习应用到行人重识别中,取得了惊人的效果.从此,越来越多的学者将深度学习与行人重识别技术相结合,通过提取鲁棒的局部特征^[15,16],设计不同的损失函数^[17,18]等方法提高了模型的泛化能力,在公开数据集上达到了非常高的准确率.



图1 行人重识别示意图

2.2 跨模态行人重识别

2.2.1 定义

当白天外界环境光线不足或者夜晚时,可见光摄像头无法拍到清晰的行人图像,而红外摄像头可以利用红外线采集到行人图像,实现24小时的监控.与单模态行人重识别不同,RGB-IR跨模态行人重识别主要研究红外图像与可见光图像之间的匹配,即给定一个特定人物的可见(红外)图像,尝试从由其他光谱相机采集的图库中搜索相应的红外(可见)图像,如图2所示.

2.2.2 面临的困难

RGB-IR跨模态行人重识别在现实世界具有很强的实用性,但它很少被研究.直到近几年来,才受到学术界的关注.在研究的过程中存在很大的困难,具体问题:

(1) 两个模态之间的巨大差异.从本质上来讲,RGB

图像和 IR 图像有很大的不同, RGB 图像有 3 个包含可见光颜色信息的通道, 而 IR 图像只有一个包含不可见光信息的通道; 从图像成像的原理来讲, RGB 图像和 IR 图像的波长范围也不同, IR 图像丢失了颜色, 曝光等重要信息, 使的识别更加困难。

(2) 传统行人重识别的模态内差异, 例如视角变化, 姿态变化等问题仍然存在. 以上情况都给 RGB-IR 跨模态行人重识别的研究带来了巨大的挑战, 导致在现实生活中无法应用。

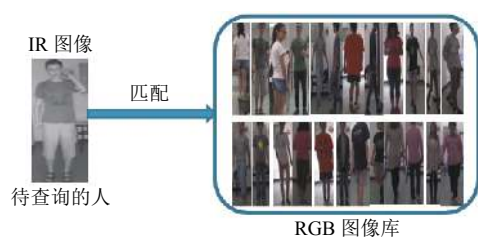


图 2 RGB-IR 跨模态行人重识别

2.2.3 发展现状

早期, Jungling 等^[19] 使用红外图像进行匹配, 但是只考虑了 IR-IR 图像之间的识别. 后来有学者关注文本与图像之间的检索, Zhao 等^[20] 提出了新颖的端到端的深度学习框架, 首次将多视图问题转换为单视图哈希问题. Peng 等^[21] 首次使用 GAN 学习文本和图像之间的共享特征, 解决了它们之间的差异. 由于文本和红外图像之间的不同, 这些方法无法在跨模态行人重识别中直接使用. 直到 2017 年, Wu 等^[22] 首次为 RGB-IR 行人重识别的研究提供了一个公开的基准 SYSU Multiple Modality Re-ID (SYSU-MM01) 数据集, 与常用的行人重识别数据集^[23-27] 比较, 如表 1 所示, 有很大的不同. 随后, 越来越多的人开始投入到 RGB-IR 跨模态行人重识别的研究中去, 由此开启了跨模态行人重识别的研究大门。

表 1 SYSU-MM01 与传统行人重识别数据集的比较

数据集	发表时间 (年)	ID数	图片数	相机数	有无 IR图像
ViPeR ^[23]	2007	632	1264 RGB	2	无
CUHK03 ^[24]	2014	1467	13 164 RGB	10	无
Market1501 ^[25]	2015	1501	32 217 RGB	6	无
DukeMTMC-reID ^[26]	2017	1812	36 441 RGB	8	无
MSMT17 ^[27]	2018	4101	126 441 RGB	15	无
SYSU-MM01 ^[22]	2017	491	30 071 RGB+ 15 792 IR	6	有

3 跨模态行人重识别问题

跨模态行人重识别是近几年新兴的研究方向, 相比于其他领域起步较晚, 发表在顶级计算机视觉会议上的文章相对较少. 解决跨模态行人重识别的关键在于学习两种模态的共享特征, 减小不同模态之间的差异. 起初的方法一般同时考虑特征学习和度量学习, 先用双路的卷积网络分别提取 RGB 图像和红外图像的特征, 然后将两个模态的特征输入共享参数的网络. 随着对跨模态行人重识别的深入研究, 出现了越来越多的优秀算法, 并逐渐取得不错的效果, 识别率得到了很大的提高。

下面对目前的 RGB-IR 跨模态行人重识别的方法进行归纳总结, 跨模态行人重识别可分为 3 类。

3.1 基于统一特征模型的方法

基于统一特征模型法, 即将不同模态的信息映射到相同的特征空间后学习出具有鉴别性和鲁棒性的特征模型, 从而减小模态间数据的差异。

最初, Wu 等^[22] 分析了 3 种不同的网络框架 (单路网络、双路网络和非对称的全连接网络) 之间的关系, 发现所有的结构最终都可以用单路网络结构表示, 并且提出深度补零操作, 将 RGB 图像转换为单通道的灰色图像放置在第一通道, 其补零图像放置在第二通道, 将 IR 图像直接放置在第二通道, 其补零图像放置在第一通道, 这样可以灵活的学到特定域的信息, 最终提出深度补零优化单路网络结构的方法解决跨模态行人重识别问题。

后来, 许多工作运用双路网络结构学习共享特征. Ye 等^[28] 利用双路网络结构学习 RGB 和 IR 图像的共有特征, 并提出了一种分层跨模态学习方法 (Hierarchical Cross-modality Matching Model, HCML) 融合特征损失和对比损失进行相似度学习. Dai 等^[29] 首次将 GAN 应用到跨模态行人重识别中, 提出了一种跨模态生成对抗网络 (cross-modality Generative Adversarial Network, cmGAN), 利用生成器学习不同模态下的特征, 利用鉴别器进行模态分类, 结合识别损失和跨模态三重损失训练, 减少了跨模态的差异和模态间的变化. 考虑到不同的 CNN 结构对应着不同的语义特征, Liu 等^[30] 提出了一种增强鉴别特征学习 (Enhancing the Discriminative Feature Learning, EDFL), 采用端到端的双路网络结构, 融合中层特征提取出更具有鲁棒性的特征。

在传统的双路结构的基础上, Zhang 等^[31]设计了一种双路径空间结构保持的公共空间网络 (DSCSN) 和一个对比相关网络 (CCN), 采用三维张量表示特征空间而不是传统的一维向量, 增加对比特征的学习有利于区别不同的行人. Hao 等^[32]考虑了空间和模态的一致性, 采用局部特性来提取模态的不变信息, 并设计了一个类内分布损失函数来减小可见图像和红外图像之间的间隙以及一个类内相关损失来对齐可见图像和红外图像的特征空间. Xiang 等^[33]利用 RGB 和 IR 图像之间的内在联系, 提出了一个端到端的双路多分支交叉模态网络, 并引入 MGN 架构学习具有鉴别性的跨模态特征, 通过结合图像的局部和全局信息来提取鲁棒性的特征.

3.2 基于度量学习的方法

基于度量学习的跨模态行人重识别, 目前的工作主要集中在采用不同的度量方法或者设计不同的损失函数提高模型的泛化能力, 目的是缩小两个模态同 ID 各个图像之间的距离和跨模态同 ID 各个图像之间的距离, 增大跨模态不同 ID 各个图像之间的距离. Ye 等^[34]同时考虑模态间和模态内的变化, 在双路网络结构的基础上设计了一种基于双向约束高阶损失 (Bi-directional Dual-constrained Top-Ranking loss, BDTR) 对行人特征进行约束. Hao 等^[35]提出了超球面流行嵌入网络 (Hyper-Sphere Manifold Embedding network, HSME), 该方法主要通过 Sphere Softmax 函数将学习到的共享特征映射到超球面上, 结合身份损失和排序损失训练模型, 再使用 KL 散度衡量两个领域预测的匹配性, 最终通过单矢量分解 (SVD) 方法修正 Sphere Softmax 最大值权矩阵. Lin 等^[36]首次将单模态下行人重识别的网络迁移到跨模态行人重识别中, 提出了新的特征学习框架 (Hard Pentaplet and Identity Loss Network, HPILN), 设计了新的硬五态损失结合特征损失提高模型的准确性. 通过引入协同学习, Ye 等^[37]提出了一种基于双流网络的模式感知协同学习方法 (Modality-Aware Collaborative, MAC) 同时处理特征级和分类器级的模态差异, 并提出协同学习方案来规范共享模式和特定模式的身份分类器. Zhu 等^[38]设计了双流局部特征网络 (Two-Stream Local Feature Network, TSLFN), 为了改进类内跨模态相似性, 提出异质中心损失 (HC loss) 限制两个异质模态中心之间的距离. Ye 等^[39]在双路网络的基础上

提出了一个双向中心约束顶级排序 (eBDTR), 将前两个约束合并到一个公式中, 同时解决了跨模态和模态内变化.

3.3 基于模态转换的方法

区别于一般的解决方法, 考虑到将不同模态的图像数据转换成统一的模态数据可以在很大程度上减小两种模态的差异.

随着生成对抗网络 (GAN) 的发展, CycleGAN^[40]、PNGAN^[41]、FDGAN^[42] 等方法的提出可以实现图片风格的转换, 有效地缓解了模态差异这一难点. 大部分工作主要使用 GAN 进行图像转换, 主要思想是将 RGB 图像转换为对应的 IR 图像或者将 IR 图像转换为对应的 RGB 图像, 之后再行单模态下的行人重识别的一般操作, 可以有效地提高识别率. Wang 等^[43]提出一种双级差异减少方法 (Dual-level Discrepancy Reduction Learning, D2RL), 具体来讲, 图像级差减子网络 T_I 利用 GAN 将 RGB (IR) 图像生成其对应的 IR (RGB) 图像, 形成统一的多光谱图像, 减少了模态间差异; 在统一的基础上, 特征级差减子网络 T_F 利用传统的 Re-ID 方法减小外观差异, 两个子网络 T_I and T_F 以端到端的方式进行联合训练. Wang 等^[44]提出一种对齐生成对抗网络 (Alignment Generative Adversarial Network, AlignGAN), 包含像素对齐模块 (P), 特征对齐模块 (F), 联合判别模块 (Dj) 3 个模块, P 模块利用 CycleGAN 模型将 RGB 图像训练生成伪 IR 图像, 并通过 cycle-consistency loss 和 identity loss 进行训练, G_p 减少跨模态间差异, F 模块用特征生成器 G_f 将伪红外图像和真红外图像编码到一个共享的特征空间中以减少模式内的差异, Dj 使得 G_p 和 G_f 相互学习, 最终学习到鲁棒的特征.

不同于上述利用 GAN 的思想, 利用 CycleGAN 等方法会产生噪声图像, 影响最终的图像匹配效果. Tekeli 等^[45]将 RGB 图像转换为灰度图像后, 提出了基于距离的分数层, 利用距离度量对网络进行训练. Basaran 等^[46]提出了四流网络结构学习有区别性的特征, 将图像进行转换后作为输入图像, 在每个流中利用 CNN 单独训练, 从每个流中学习不同且互补的特征. Wang 等^[47]提出了生成跨模态配对图像, 并执行全局集合级和细粒度实例级对齐, 这种方法可以通过解开特定于模态和模态不变的特征来执行集合级对齐, 同时可以从交换的图像生成跨模态成对图像, 最小化每

对图像的距离直接执行实例级对齐。

4 数据集和评价标准

为了评估跨模态行人重识别的相关方法,一般在公开数据集上进行实验,并通过统一的评价标准来评估所提出方法的性能。下面介绍了跨模态行人重识别的相关数据集和评价标准。

4.1 数据集

目前只有两个公开数据集,如表2所示,用于跨模态行人重识别的实验。

SYSU-MM01^[22](图3)是在2017年首次公开跨模态行人重识别的数据集,也是目前最具有挑战性的数据集。它由6个摄像头采集的图像组成,分别是2个红外摄像头和4个可见光摄像头。红外摄像头与可见光摄像头不同,即使在黑暗环境下,它也能正常的工

作,捕捉到行人的特征。该数据集包含491个不同身份行人,其中296个用于训练,99个用于验证,96个用于测试,总共有30071张RGB图片和15792张IR图片。

RegDB数据集^[48](图4)同时使用可见光摄像头和红外摄像头拍摄。总共有412个不同的行人,其中女性254人,男性158人,每个人分别对应10张可见光图像和10张红外图像,其中拍摄到156个行人的正面,256个行人的背面。该数据集总共有4120张可见光图像和4120张的热图像。

表2 跨模态行人重识别数据集

数据集	ID	RGB摄像机数	IR摄像机数	RGB图片数	IR图片数
SYSU-MM01 ^[22]	491	4	2	30071	15792
RegDB ^[48]	412	1	1	4120	4120

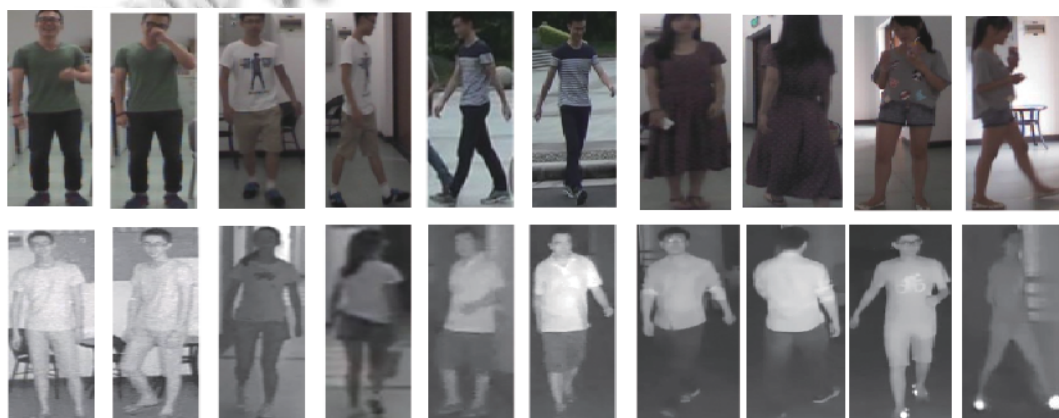


图3 SYSU-MM01数据集行人实例



图4 RegDB数据集行人实例

4.2 评价标准

(1) CMC 曲线

CMC曲线全称是Cumulative Match Characteristic (CMC) Curve,即累计匹配曲线,是行人重识别重要的评测指标,它可以综合反映分类器的性能。具体来说,在候选行人库(gallery)中检索待查询(probe)的行人,前 k 个检索结果中包含正确匹配结果的比率,通常用

Rank- k 的形式表示。Rank-1识别率就是表示按照某种相似度匹配规则匹配后,第一次就能返回正确匹配的概率,即最匹配候选目标刚好为待查询图片目标的概率,Rank-5识别率就是指前5个匹配候选目标中存在待查询图片目标的概率。

(2) mAP 均值平准精度

目前大部分的研究都是跨多个摄像头,而CMC曲

线只适用于两个摄像头之间的检索,因此 Zheng 等人^[49]提出了均值平均精度 (mean Average Precision, mAP) 对算法进行评估. mAP 的具体操作是,分别求出每个类别的 AP 值后取平均值. AP 值是求 PR 曲线下的面积,综合考虑了 P(准确率)和 R(召回率),是衡量一个模型好坏的标准.

5 现有方法分析

本节对近些年具有代表性算法进行分析,基于统一特征模型和度量学习的方法包括 Deep Zero-Padding^[22]、HCML^[28]、cmGAN^[29]、EDFL^[30]、DSCSN+CCN^[31]、DFE^[32]、BDTR^[34]、HSME^[35]、HPILN^[36]、MAC^[37]、TSLFN+HC^[38]、eBDTR^[39]、IPVT-1 and MSR^[50]; 基于模态转换的方法包括 D2RL^[43]、AlignGAN^[44]、Dist.based^[45]、4-stream framework+LZM^[46]. 表 3 总结比较了这十余种方法在跨模态行人重识别数据集 RegDB 和 SYSU-MM01 的识别率以及发表状况. 采取准确率 (Rank-1) 和平均准确率 (mAP) 作为评价标准, — 表示没有实验结果.

总的看来,跨模态行人重识别方法发展迅速,最优的与最初的算法相比,准确率大概增长 50%. 在 RegDB 数据集上, DSCSN+CCN^[31] 取得了最高识别率, Rank-1 达到 60.80%, mAP 达到 60.00%. 在 SYSU-MM01 数据集上, 4-stream framework+LZM^[46] 效果最好, Rank-1 达到 63.05%, mAP 达到 67.13%. 在上述算法中, HSME^[35]、IPVT-1 and MSR^[50]、EDFL^[30]、DSCSN+CCN^[31] 和 AlignGAN^[44] 在 RegDB 数据集上识别率超过 50%, 而在 SYSU-MM01 数据集上, 只有 TSLFN+HC^[38] 和 4-stream framework+LZM^[46] 准确率超过了 50%. 大部分方法在 SYSU-MM01 数据集上的识别率都比在 RegDB 数据集上高, 表明数据集 SYSU-MM01 比 RegDB 更具有挑战性. 在 2019 年发表在期刊上的优秀论文迅猛增长, 跨模态行人重识别逐渐得到学术界的重视. 经分析, 我们发现网络结构越来越复杂, 起初采用单流和双流结构, 到目前有人提出四流网络结构, 学习到更具鉴别性的特征; 同时发现基于模态转换的方法相比于其他方法, 识别率提升较高, 存在巨大的优势.

表 3 跨模态行人重识别方法在数据集 RegDB 和 SYSU-MM01 的识别结果

方法	RegDB		SYSU-MM01 (All)		SYSU-MM01 (Indoor)		出版时间
	Rank-1(%)	mAP(%)	Rank-1(%)	mAP(%)	Rank-1(%)	mAP(%)	
Deep Zero-Padding ^[22]	—	—	14.80	15.95	20.58	26.92	ICCV2017
HCML ^[28]	24.44	20.80	—	—	—	—	AAAI2018
BDTR ^[34]	33.47	31.83	17.01	19.66	—	—	IJCAI2018
cmGAN ^[29]	—	—	26.97	27.80	31.63	42.19	IJCAI2018
HSME ^[35]	50.85	47.00	20.68	23.12	—	—	AAAI2019
MAC ^[37]	36.43	37.03	33.26	36.22	33.37	44.95	ACM2019
IPVT-1 and MSR ^[50]	58.76	47.85	23.18	22.49	—	—	IEEEAccess2019
EDFL ^[30]	52.58	52.98	36.94	40.77	—	—	Neurocomputing2019
HPILN ^[36]	—	—	41.36	42.95	45.77	56.52	IET Image Processing 13 (2019)
TSLFN+HC ^[38]	—	—	56.96	54.95	59.74	64.91	Neurocomputing2019
eBDTR ^[39]	34.62	33.46	27.82	28.42	—	—	TIFS2019
DSCSN+CCN ^[31]	60.80	60.00	35.10	37.40	—	—	ArXiv2019
DFE ^[32]	—	—	48.71	48.59	—	—	ACM2019
D2RL ^[43]	43.40	44.10	28.90	29.20	—	—	CVPR2019
AlignGAN ^[44]	56.30	53.40	42.40	40.70	45.90	54.30	ICCV2019
Dist.based ^[45]	38.64	38.08	29.05	30.94	32.74	44.26	IEEE Computer Vision2019
4-stream framework+LZM ^[46]	—	—	63.05	67.13	69.06	76.95	ArXiv2019

6 总结展望

跨模态行人重识别是行人重识别的一个新的发展趋势, 对智能化社会有着重要的研究意义和应用价值. 虽然目前取得了一定的研究成果, 但跨模态行人重识

别的发展仍处在初级阶段. 想要取得更大的突破, 未来的发展方向可以从以下方面考虑.

(1) 构建高质量的数据集. 现有跨模态行人重识别的数据集数量少, 并且其规模也很小只包含了几百个

行人的ID,可供训练的图片非常有限,影响跨模态行人匹配的效果.同时目前数据集的场景不够丰富^[51],但是现实会遇到多样的环境,不同的环境,不同的光线等因素都会影响跨模态图像之间的匹配,造成很大的差异.

(2) 关注模态转换的研究.研究者们通常采用一般的方法,结合特征提取和度量学习解决模态间和模态内变化.根据对现有方法的分析,发现采用模态转换的方法,识别率明显优于传统的方法.其中GAN、风格迁移等方法可以有效地实现两个域之间的转换,有效缓解模态间的差异.

(3) 结合局部特征学习.在行人重识别中颜色是区别行人的有效信息,由于红外图像特殊性,无法在跨模态行人重识别中使用.因此,其他的信息变得异常关键,我们可以结合局部特征,学习出具有鲁棒性特征,从而提高行人识别率.

参考文献

- 1 宋婉茹,赵晴晴,陈昌红,等.行人重识别研究综述.智能系统学报,2017,12(6):770-780.[doi:10.11992/tis.201706084]
- 2 Zheng L, Yang Y, Hauptmann AG. Person re-identification: Past, present and future. arXiv: 1610.02984, 2016.
- 3 Zheng L, Bie Z, Sun YF, *et al.* Mars: A video benchmark for large-scale person re-identification. Proceedings of the 14th European Conference on Computer Vision. Amsterdam, The Netherlands. 2016. 868-884.
- 4 Yi D, Lei Z, Liao SC, *et al.* Deep metric learning for person re-identification. Proceedings of the 22nd International Conference on Pattern Recognition. Stockholm, Sweden. 2014. 34-39.
- 5 Felzenszwalb P, McAllester D, Ramanan D. A discriminatively trained, multiscale, deformable part model. Proceedings of 2008 IEEE Conference on Computer Vision and Pattern Recognition. Anchorage, AK, USA. 2008. 1-8.
- 6 Yang Y, Yang JM, Yan JJ, *et al.* Salient color names for person re-identification. Proceedings of the 13th European Conference on Computer Vision. Zurich, Switzerland. 2014. 536-551.
- 7 Jüngling K, Bodensteiner C, Arens M. Person re-identification in multi-camera networks. Proceedings of CVPR 2011 Workshops. Colorado Springs, CO, USA. 2011. 55-61.
- 8 Weinberger KQ, Saul KL. Distance metric learning for large margin nearest neighbor classification. The Journal of Machine Learning Research, 2009, 10: 207-244.
- 9 Zheng WS, Gong SG, Xiang T. Person re-identification by probabilistic relative distance comparison. Proceedings of the CVPR 2011. Providence, RI, USA. 2011. 649-656.
- 10 Köstinger M, Hirzer M, Wohlhart P, *et al.* Large scale metric learning from equivalence constraints. Proceedings of 2012 IEEE Conference on Computer Vision and Pattern Recognition. Providence, RI, USA. 2012. 2288-2295.
- 11 Zheng WS, Gong SG, Xiang T. Reidentification by relative distance comparison. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2013, 35(3): 653-668. [doi: 10.1109/TPAMI.2012.138]
- 12 Pedagadi S, Orwell J, Velastin S, *et al.* Local fisher discriminant analysis for pedestrian re-identification. Proceedings of 2013 IEEE Conference on Computer Vision and Pattern Recognition. Portland, OR, USA. 2013. 3318-3325.
- 13 Liao SC, Hu Y, Zhu XY, *et al.* Person re-identification by local maximal occurrence representation and metric learning. Proceedings of 2015 IEEE Conference on Computer Vision and Pattern Recognition. Boston, MA, USA. 2015. 2197-2206.
- 14 Deng J, Dong W, Socher R, *et al.* ImageNet: A large-scale hierarchical image database. Proceedings of 2009 IEEE Conference on Computer Vision and Pattern Recognition. Miami, FL, USA. 2009. 248-255.
- 15 Sun YF, Zheng L, Yang Y, *et al.* Beyond part models: Person retrieval with refined part pooling (and a strong convolutional baseline). Proceedings of the 15th European Conference on Computer Vision (ECCV). Munich, Germany. 2018. 480-496.
- 16 Zhao HY, Tian MQ, Sun SY, *et al.* Spindle net: Person re-identification with human body region guided feature decomposition and fusion. Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, HI, USA. 2017. 1077-1085.
- 17 Hermans A, Beyer L, Leibe B. In defense of the triplet loss for person re-identification. arXiv: 1703.07737, 2017.
- 18 Chen WH, Chen XT, Zhang JG, *et al.* Beyond triplet loss: A deep quadruplet network for person re-identification. Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, HI, USA. 2017. 403-412.
- 19 Jüngling KK, Arens M. Local feature based person reidentification in infrared image sequences. Proceedings of the 7th IEEE International Conference on Advanced Video and Signal Based Surveillance. Boston, MA, USA. 2010.

- 448–455.
- 20 Zhao X, Ding G, Guo Y, Han J, & Gao Y. TUCH: Turning cross-view hashing into single-view hashing via generative adversarial nets. Proceedings of the 26th International Joint Conference on Artificial Intelligence. Melbourne, VIC, Australia. 2017. 3511–3517.
- 21 Peng YX, Qi JW. CM-GANs: Cross-modal generative adversarial networks for common representation learning. ACM Transactions on Multimedia Computing, Communications, and Applications, 2019, 15(1): 1–24.
- 22 Wu AC, Zheng WS, Yu HX, *et al.* RGB-infrared cross-modality person re-identification. Proceedings of 2017 IEEE International Conference on Computer Vision. Venice, Italy. 2017. 5380–5389.
- 23 Gray D, Tao H. Viewpoint invariant pedestrian recognition with an ensemble of localized features. Proceedings of the 10th European Conference on Computer Vision. Marseille, France. 2008. 262–275.
- 24 Li W, Zhao R, Xiao T, *et al.* Deepreid: Deep filter pairing neural network for person re-identification. Proceedings of 2014 IEEE Conference on Computer Vision and Pattern Recognition. Columbus, OH, USA. 2014. 152–159.
- 25 Zheng L, Shen LY, Tian L, *et al.* Scalable person re-identification: A benchmark. Proceedings of 2015 IEEE International Conference on Computer Vision. Santiago, Chile. 2015. 1116–1124.
- 26 Zheng ZD, Zheng L, Yang Y. Unlabeled samples generated by GAN improve the person re-identification baseline in vitro. Proceedings of 2017 IEEE International Conference on Computer Vision. Venice, Italy. 2017. 3754–3762.
- 27 Wei LH, Zhang SL, Gao W, *et al.* Person transfer GAN to bridge domain gap for person re-identification. Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, UT, USA. 2018. 79–88.
- 28 Ye M, Lan XY, Li JW, *et al.* Hierarchical discriminative learning for visible thermal person re-identification. Proceedings of the 32nd AAAI Conference on Artificial Intelligence, the 30th Innovative Applications of Artificial Intelligence (IAAI-18), and the 8th AAAI Symposium on Educational Advances in Artificial Intelligence (EAAI-18). New Orleans, LA, USA. 2018. 7501–7508.
- 29 Dai PY, Ji RR, Wang HB, *et al.* Cross-modality person re-identification with generative adversarial training. Proceedings of the 27th International Joint Conference on Artificial Intelligence. Stockholm, Sweden. 2018. 2.
- 30 Liu HJ, Cheng J, Wang W, *et al.* Enhancing the discriminative feature learning for visible-thermal cross-modality person re-identification. Neurocomputing, 2020, 398: 11–19. [doi: [10.1016/j.neucom.2020.01.089](https://doi.org/10.1016/j.neucom.2020.01.089)]
- 31 Zhang SZ, Yang YF, Wang P, *et al.* Attend to the difference: Cross-modality person re-identification via contrastive correlation. arXiv: 1910.11656, 2019.
- 32 Hao Y, Wang NN, Gao XB, *et al.* Dual-alignment feature embedding for cross-modality person re-identification. Proceedings of the 27th ACM International Conference on Multimedia. New York, NY, USA. 2019. 57–65.
- 33 Xiang XZ, Lv N, Yu ZT, *et al.* Cross-modality person re-identification based on dual-path multi-branch network. IEEE Sensors Journal, 2019, 19(23): 11706–11713. [doi: [10.1109/JSEN.2019.2936916](https://doi.org/10.1109/JSEN.2019.2936916)]
- 34 Ye M, Wang Z, Lan XY, *et al.* Visible thermal person re-identification via dual-constrained top-ranking. Proceedings of the 27th International Joint Conference on Artificial Intelligence. Stockholm, Sweden. 2018. 2. [doi: [10.24963/ijcai.2018/152](https://doi.org/10.24963/ijcai.2018/152)]
- 35 Hao Y, Wang NN, Li J, *et al.* HSME: Hypersphere manifold embedding for visible thermal person re-identification. Proceedings of the AAAI Conference on Artificial Intelligence, 2019, 33(1): 8385–8392.
- 36 Zhao YB, Lin JW, Xuan Q, *et al.* HPILN: A feature learning framework for cross-modality person re-identification. IET Image Processing, 2019, 13(14): 2897–2904. [doi: [10.1049/iet-ipr.2019.0699](https://doi.org/10.1049/iet-ipr.2019.0699)]
- 37 Ye M, Lan XY, Leng QM. Modality-aware collaborative learning for visible thermal person re-identification. Proceedings of the 27th ACM International Conference on Multimedia. Nice, France. 2019. 347–355.
- 38 Zhu YX, Yang Z, Wang L, *et al.* Hetero-center loss for cross-modality person re-identification. Neurocomputing, 2020, 386: 97–109. [doi: [10.1016/j.neucom.2019.12.100](https://doi.org/10.1016/j.neucom.2019.12.100)]
- 39 Ye M, Lan XY, Wang Z, *et al.* Bi-directional center-constrained top-ranking for visible thermal person re-identification. IEEE Transactions on Information Forensics and Security, 2019, 15: 407–419.
- 40 Zhu JY, Park T, Isola P, *et al.* Unpaired image-to-image translation using cycle-consistent adversarial networks. Proceedings of 2017 IEEE International Conference on Computer Vision. Venice, Italy. 2017. 2223–2232.
- 41 Qian XL, Fu YW, Xiang T, *et al.* Pose-normalized image generation for person re-identification. Proceedings of the 15th European Conference on Computer Vision (ECCV). Munich, Germany. 2018. 650–667.

- 42 Ge YX, Li ZW, Zhao HY, *et al.* FD-GAN: Pose-guided feature distilling GAN for robust person re-identification. Proceedings of the 32nd International Conference on Neural Information Processing Systems. Montréal, QQC, Canada. 2018. 1230–1241.
- 43 Wang ZX, Wang Z, Zheng YQ, *et al.* Learning to reduce dual-level discrepancy for infrared-visible person re-identification. Proceedings of 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach, CA, USA. 2019. 618–626.
- 44 Wang GA, Zhang TZ, Cheng J, *et al.* RGB-infrared cross-modality person re-identification via joint pixel and feature alignment. Proceedings of 2019 IEEE/CVF International Conference on Computer Vision. Seoul, Republic of Korea. 2019. 3623–3632.
- 45 Tekeli N, Burak Can A. Distance based training for cross-modality person re-identification. Proceedings of 2019 IEEE/CVF International Conference on Computer Vision Workshops. Seoul, Republic of Korea. 2019. 4540–4549.
- 46 Basaran E, Gokmen M, Kamasak ME. An efficient framework for visible-infrared cross modality person re-identification. arXiv: 1907.06498, 2019.
- 47 Wang GA, Yang TZ, Yang Y, *et al.* Cross-modality paired-images generation for RGB-infrared person re-identification. arXiv: 2002.04114, 2020.
- 48 Nguyen DT, Hong HG, Kim KW, *et al.* Person recognition system based on a combination of body images from visible light and thermal cameras. Sensors, 2017, 17(3): 605. [doi: [10.3390/s17030605](https://doi.org/10.3390/s17030605)]
- 49 Chen YC, Zheng WS, Lai JH, *et al.* An asymmetric distance model for cross-view feature mapping in person reidentification. IEEE Transactions on Circuits and Systems for Video Technology, 2017, 27(8): 1661–1675. [doi: [10.1109/TCSVT.2016.2515309](https://doi.org/10.1109/TCSVT.2016.2515309)]
- 50 Kang JK, Hoang TM, Park KR. Person re-identification between visible and thermal camera images based on deep residual CNN using single input. IEEE Access, 2019, 7: 57972–57984. [doi: [10.1109/ACCESS.2019.2914670](https://doi.org/10.1109/ACCESS.2019.2914670)]
- 51 罗浩, 姜伟, 范星, 等. 基于深度学习的行人重识别研究进展. 自动化学报, 2019, 45(11): 2032–2049.