

引用格式: 党婉丽, 程建, 罗谦, 等. 航站楼场景下行人重识别方法研究 [J]. 电子科技大学学报, 2025, 54(5): 755-762.
DANG W L, CHENG J, LUO Q, et al. Research on pedestrian re-identification method in terminal scenarios[J]. Journal of University of Electronic Science and Technology of China, 2025, 54(5): 755-762.

航站楼场景下行人重识别方法研究



党婉丽^{1,2}, 程建^{1*}, 罗谦², 邵杰², 郑怀宇²

(1. 电子科技大学 信息与通信工程学院, 成都 611731; 2. 中国民用航空总局 第二研究所, 成都 610041)

摘要: 聚焦使用重识别方法解决航站楼旅客跨摄像头追踪问题。航站楼行人密集、灯光变化引起行人遮挡、颜色特征不断变化。现有重识别模型过于依赖行人外观及边缘变化信息, 而忽略了颜色本身特征信息, 难以适应航站楼场景。对重识别网络全连接层进行特征自相关分析, 获取特征向量中的颜色向量。定义行人动态变化区域, 构建航站楼旅客匹配特征函数, 解决由颜色、遮挡引起的匹配不准确问题。最后, 以广州白云国际机场、成都双流国际机场数据进行实验, 实验表明, 提出方法比现有 DeepSort、SOLTDER 更优 (Top1 准确率提升 5.05%、召回率提升 2.62%), 更适应航站楼场景。

关键词: 航站楼; 重识别; 特征自相关; 颜色特征向量

中图分类号: TP391

文献标志码: A

DOI: 10.12178/1001-0548.2024183

Research on pedestrian re-identification method in terminal scenarios

DANG Wanli^{1,2}, CHENG Jian^{1*}, LUO Qian², SHAO Jie², and ZHENG Huaiyu²

(1. School of Information and Communication Engineering, University of Electronic Science and Technology of China, Chengdu 611731, China;

2. The Second Research Institute, Civil Aviation Administration of China, Chengdu 610041, China)

Abstract: This paper focuses on using re-identification method to solve the problem of passenger cross-camera tracking in terminal. The terminal is usually crowded with pedestrians and the lighting keeps changing, this can cause pedestrians to be partially obscured and their color characteristics to constantly change. The existing re-identification models overly rely on the information of pedestrian appearance and edge change, and ignore the color itself, difficult to adapt to the terminal scene. In this paper, we perform feature auto-correlation analysis on the fully-connected layer of the re-identification network and extract color vectors from feature representations. By defining the dynamic changing areas of pedestrians and constructing the matching feature functions for passengers in the terminal, the problem of inaccurate matching caused by color variations and occlusions is solved. Experiments at Guangzhou Baiyun and Chengdu Shuangliu International Airports show the method outperforms DeepSort and SOLTDER, increasing Top1 Accuracy by 5.05% and Recall by 2.62%, demonstrating strong adaptability in terminal environments.

Key words: terminal; re-identification; feature auto-correlation; color feature vector

行人重识别方法为在多个摄像机之间建立一致的标签, 在跟踪时对同一行人重新标识, 是实现跨摄像头下特定行人跟踪的有效方法。可解决航站楼旅客跨摄像头追踪问题, 在重点旅客寻找、旅客轨迹异常检测等方面具有广泛应用。

现有研究中, 行人重识别方法被广泛视为检索或识别问题, 分为特征提取和特征匹配两个子任务。特征提取包括外观、动作步态以及时间信息等特征。其中行人外观信息包含行人的着装颜色、外

观纹理等特征。特征匹配通过对行人特征表示向量的距离进行度量, 反映跨摄像头下行人特征差异。两个子任务可以进行独立学习和优化。

行人重识别算法主要包含两类: 1) 基于传统方法的特征提取匹配; 2) 基于深度学习的行人重识别方法。传统算法使用 HOG^[1]、SIFT^[2]、LOMO^[3] 等方法来实现行人特征的特征, 利用 XQDA^[3] (cross-view quadratic discriminant analysis) 或 KISSME^[4] (keep it simple and straightforward metric learning)

收稿日期: 2024-07-18

基金项目: 国家自然科学基金民航联合重点基金 (U2133211)

作者简介: 党婉丽, 博士, 副研究员, 主要从事视频分析、图像处理等方面的研究。

*通信作者 E-mail: chengjian@uestc.edu.cn

学习相似度量函数。传统方法对复杂场景任务难以适用,在数据量巨大情况下,传统度量学习方法求解变得异常困难。

随着深度学习的快速发展,卷积神经网络在计算机视觉领域取得良好效果。一系列的研究工作专注于改进特征提取器、相似性度量或同时对这两者进行优化。文献[5]提出的行人重识别框架 HOReID,通过聚合多个层级特征图的细粒度特征实现语义姿势对齐,增强学习姿势特征的鲁棒性。文献[6]提出卷积块注意力模型,与自注意机制集成提取不同关注域特征,引入不同粒度局部特征,联合进行行人重识别。文献[7]提出一种渐进式特征增强算法,基于来自不同卷积神经网络层的多尺度判别线索,将其融合到单个特征向量中来获取更具判别性的特征表示。文献[8]提出了全尺度网络学习特征表示,捕获不同尺度的特征,通过多个尺度的不同组合形成全尺度特征。文献[9]设计了姿势及掩码引导层,生成显著性热图以进一步指导判别特征的学习。文献[10]将网络分为身份识别和属性识别两个分支以得到更全面的行人特征表示,通过通道注意力机制学习不同属性间的关联,以交叉验证的方式确定每个属性的权重。

在相似性度量方面,文献[11]采用交叉熵损失和三元组损失函数联合方式训练网络模型,获得更有效的特征度量。文献[12]利用交叉熵损失、三元组损失和中心损失函数作为联合损失函数,使得类间和类内具有更好的区别性。文献[13]提出融合关系学习的行人重识别方法,利用注意力机制使最重要的特征块更显著,采用融合相似性度量,度量方式更全面。文献[14]提出了一种有效的小批量采样方法,用于大规模深度度量学习。文献[15]构建细粒度特征注意力模块,引入区域特征补全模块,使遮挡区域特征能够通过相同聚类的特征恢复,使用身份、加权正则化三元组和中心损失对模型进行训练优化。

外观信息是匹配的重要特征。颜色特征是一种直观的有辨识度的属性,在行人重识别方法中应用广泛。文献[16]使用颜色随机干扰检索图像颜色,实现抑制颜色偏差引起的类内变化。文献[17]通过惩罚重识别模型对服装的预测能力,从原始图像挖掘与服装无关特征。文献[18]引入步态识别作为辅助任务驱动重识别模型,主要加入步态特征进行辅助判断。文献[19]在 ResNet50 基础网络中引入实例归一化层,利用 Transformer 的多头自注意力模块与卷积结合,增强特征表征能力,解决网络

对场景图像风格、颜色等鲁棒性不高问题。文献[20]提出了一种结合特征融合和通道注意力的多分支换装行人重识别方法,提取行人细粒度特征的同时,使用对抗损失和交叉熵标签平滑损失多种损失函数引导模型学习与服装无关的信息,减少服装颜色的影响。这些方法在初始识别中均取得了较好的效果,但在行人遮挡情况下,难以获取有效信息。

为有效应对遮挡及颜色特征不断变化,更好适应航站楼场景,本文提出一种航站楼场景下行人重识别方法。首先使用 yolox^[21]对行人进行检测,对图像行人框进行裁剪,获得行人区域图像块。其次使用 OSNet^[8]提取行人图像块的特征。在上述基础,设计一种增加颜色特征权重方案,进行特征自相关性分析,获取特征中颜色相关向量,使模型在关注外观、边缘、纹理等信息同时关注颜色特征。针对遮挡问题,定义行人动态变化观测区域作为特征匹配区域。与较为流行行人重识别方法相比,本文提出的方法可获取关注区域更具判别性的行人特征表示,适应遮挡及颜色特征变化场景。最后,在广州白云国际机场、成都双流国际机场进行实验验证,取得了较好的效果。

1 研究动机

1.1 机场场景特点

相较于公开数据集,机场航站楼场景具有复杂、人流密集、灯光变化影响大等特点。现有方法应用于航站楼具有明显的瓶颈:1) 人流密集,遮挡严重,造成提取匹配特征无效;2) 不同灯光或环境,同一行人外观特征表征不同。

1.2 现有方法实验分析

本节使用已有行人重识别方法对机场航站楼场景进行实验分析(数据来源于某机场授权拍摄视频,邀请工作人员模拟旅客)。具体从两个安检通道(CAMERA 1、CAMERA 2)获取寻找旅客的全身照,共计 231 人,将标注的 231 人作为待寻找行人数据集合,寻找行人数据集合为通过通道中 6 个摄像头(CAMERA3、CAMERA4、CAMERA5、CAMERA6、CAMERA7、CAMERA8)的旅客。针对实验场景,使用 Yolo 系列进行检测,获取行人图像块,使用 OSNet^[8]网络提取行人全连接层特征并进行匹配。同时,使用 DeepSort^[22]、SOLIDER^[23]进行了实验验证,实验中,摄像头的位置、时间等信息均为已知条件,如通过 CAMERA1 摄像头旅客完成安检工作后,仅可到达 CAMERA5 覆盖区域,且捕获时间不超过 2 min。

使用召回率及准确率来衡量3种算法的有效性, 通过统计, 3种方法的平均召回率为89.5%, Top5的准确率为85.6%。表1展示了使用3种方法在召回率以及Top5准确率上的具体表现。

表1 3种算法实验结果分析表

方法	召回率	Top5准确率
Yolo+OSNet ^[8]	88.5	84.3
DeepSort ^[22]	90.1	87.0
SOLTDER ^[23]	89.9	85.6
平均值	89.5	85.6

3种算法均可较好适应姿势等变化的检测。实验中发现, 外观颜色差异明显的场景依然识别错误, 如穿黑色衣服人员与橘色衣服人员进行了匹配。经统计, 此类错误在整个识别错误中占40.3%。通过分析在外形颜色差异大的场景中, 识别旅客边缘特征类似。因此可做出如下假设: 现有模型过于依赖行人外观及边缘变化信息, 而忽略了颜色本身的特征表征。本文基于上述假设, 分析颜色特征表征的影响, 通过特征自相关性分析获取颜色特征向量。为保证颜色表征特征的有效性, 定义行人动态变化观测区域, 来实现匹配特征的有效表达。

2 方法设计

针对以上的分析假设, 本节首先进行特征自相关性分析, 其次定义行人动态变化观测区域, 最后设计适应机场航站楼场景的行人重识别网络架构, 图1为本文方法具体流程。

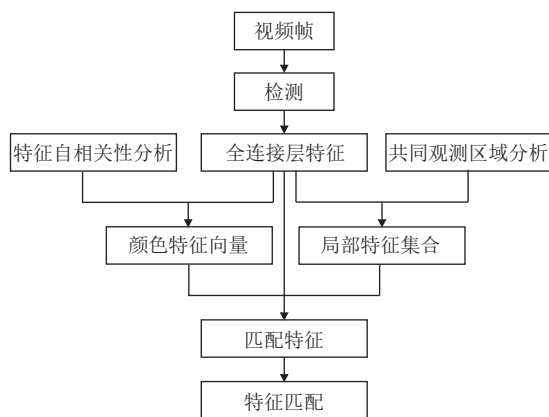


图1 本文方法流程图

2.1 特征自相关分析

针对现有模型过于依赖行人外观边缘变化信息, 而忽略颜色本身信息问题, 使用控制变量分析颜色相关特征向量。在形态、外观等条件一致情况下, 提取穿衣不同的人物特征。为减少背景干扰, 选用

纯白背景场景下的某一人物, 使用CP-VTON+^[24]对选取人物进行处理, 获得同一人物在同一场景下着不同颜色同一风格服装。

选取 Q 个人, K 种衣服风格, 每种衣服包含 M 种颜色, 每一个人、每一种衣服类型、每一种颜色生成一张行人图像 x 。将图像 x 输入OSNet网络中, 提取全连接层特征 $f \in \mathcal{R}^{512}$, Φ 为特征提取函数:

$$f = \Phi(x) \quad (1)$$

针对 $q(1 \leq q \leq Q)$ 个人的第 $k(1 \leq k \leq K)$ 种衣服类型, 获取 M 种颜色的特征向量集合 $\{f_1, f_2, \dots, f_M\}$ 。对 f_1, f_2, \dots, f_M 进行相关性分析, 获取特征向量中与颜色相关性最高各维度向量下标。

将特征 f_1, f_2, \dots, f_M 沿着同一维度连接为特征矩阵 $F \in \mathcal{R}^{P \times 512}$, 记为:

$$F = \text{Concat}[f_1, f_2, \dots, f_M] \quad (2)$$

式中, Concat为矩阵的连接。求解特征矩阵的最大矩阵 F_{MAX} , $F_{\text{MAX}} \in \mathcal{R}^{P \times P \times 512}$ 和最小矩阵 F_{MIN} , $F_{\text{MIN}} \in \mathcal{R}^{P \times P \times 512}$ 。

求解特征局部相似矩阵 S_{simi} , 其中 e^x 代表指数运算, 具体为:

$$S_{\text{simi}} = e^{F_{\text{MIN}}} / e^{F_{\text{MAX}}} \quad (3)$$

将 S_{simi} 转化为二维矩阵 $S_{\text{simi-samples}}$, $S_{\text{simi-samples}} \in \mathcal{R}^{P^2 \times 512}$, 计算每个特征向量中低于相似度阈值 α 的属性比率向量 r , 其中 $r \in \mathcal{R}^{512}$:

$$r = \frac{\sum S_{\text{simi-samples}} < \alpha}{M^2} \quad (4)$$

定义特征阈值 β , 计算 r 中高于 β 的特征向量, 获取 f_1, f_2, \dots, f_M 中与颜色相关性最高的各维度向量下标集合 $V_{q,k}$ 。计算获得 $Q \times K$ 个向量集合 $\{V_{q,k}; q = 1, 2, \dots, Q; k = 1, 2, \dots, K\}$, 取所获得向量的交集作为最后特征集合 V :

$$V = \bigcap_{q=1,2,\dots,Q; k=1,2,\dots,K} V_{q,k} \quad (5)$$

使用获取的下标对 f 进行特征选择, 构成新的行人特征向量 $f^V \in \mathcal{R}^{\text{card}(V)}$ 。

2.2 行人动态共同变化观测区域定义

不同摄像头下, 由于遮挡、角度、姿态及灯光原因, 造成同一行人特征表现不同。本节通过定义行人动态变化观测区域, 获取同一行人在不同摄像头固有的一致性特征。

姿态估计可获取行人有效姿态关节点, 根据关

节点信息, 构建不同行人姿态关节有效集合, 确定两个不同行人图像共同动态观测区域。动态变化观测区域由不同行人图像姿态有效集合交集确定。具体地, 采用 Openpose^[25] 获取行人二维姿态信息 $\{P_j, j \in \{1, 2, 3, \dots, 18\}\}$, j 代表与人体关节对应序号, 图 2 为二维姿态信息与关节对应序号表示。当存在遮挡时, 被遮挡部分关键点二维坐标为无效值。获取的不同图像不同行人姿势估计有效集合可确定两个不同行人图像共同被观测区域 (common seen area, CSA)。通过姿势引导注意力 (pose-guided attention, PGA) 捕获共同区域局部特征。将获取的有效集合按照人体逻辑结构进行连接, 确定两个行人共同被观测到的关节线段。

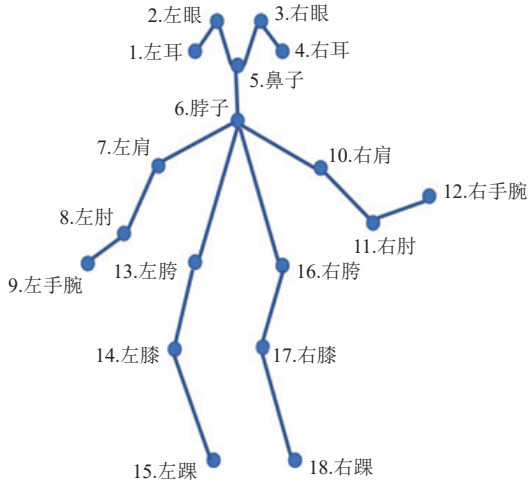


图 2 二维姿态信息与关节对应序号表示

通过外观特征与姿势引导的注意力图融合来获取被观测关节线段相关身体部位特征。使用 Openpose^[25] 提取 P_j 以及局部亲和区域 L_p 。姿势编码器将 $P = P_j \oplus L_p$ 作为输入将姿势信息嵌入高级姿势特征 $F_{\text{pose}} = PE(P; \theta_e)$, 其中 θ_e 是姿势估计器的参数。采用 1×1 卷积层及 Sigmoid 函数获取关节线段局部二维特征 B :

$$B = \sigma(\text{PGA}(F_{\text{pose}}; \theta_e)) \in \mathfrak{R}^{N_p \times H \times W} \quad (6)$$

式中, N_p 为定义二维姿态信息数量; θ_e 是卷积层的参数; σ 代表 Sigmoid 激活函数。结合获取的关键线段, 获取 CSA 的局部特征集合 f^B 。

2.3 网络结构

图 3 为本文方案的网络架构。在行人检测部分, 使用 Yolox^[21] 网络, 检测出行人区域并进行提取。将提取的行人图像输入 OSNet^[8] 网络, 获取行人特征, 输入 2.1 节描述方法中设计的特征选择器, 实现对需要设置权重特征向量的选择。同时, 利用 2.2 节中定义共同观测区域方法获取局部特征信息。

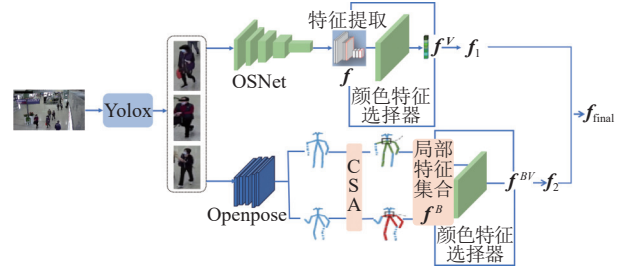


图 3 网络结构图

特征选择器使用下标对获取的行人特征向量进行选择, 构成新的行人自相关特征向量 $f^V \in \mathfrak{R}^{\text{card}(V)}$, 获取行人完整特征 f_1 , 表示如下:

$$f_1 = \text{norm}(f + \gamma f^V), 0 < \gamma < 1 \quad (7)$$

共同被观测区域特征向量 $f^B \in \mathfrak{R}^{\text{card}(V)}$ 通过颜色特征选择器, 获取具有颜色特征权重的共同观测区域特征 $f^{BV} \in \mathfrak{R}^{\text{card}(V)}$ 。共同观测区域特征 f_2 表示如下:

$$f_2 = \text{norm}(f^V + \lambda f^{BV}), 0 < \lambda < 1 \quad (8)$$

最终获取行人的完整特征 f_{final} , 表示如下:

$$f_{\text{final}} = f_1 + f_2 \quad (9)$$

式中, norm 表示特征归一化处理, $0 < \gamma < 1$ 和 $0 < \lambda < 1$ 分别表示自相关特征 f^V 和共同观测特征 f^B 的权重平衡参数。最后, 使用余弦相似度计算行人之间的相似程度, 实现对目标行人匹配。

2.4 方案优化

在实验中, 本文方案对视频取流、模型选择和超参数调优进行了优化。具体地, 选择 3 s 一帧的间隔模式进行视频流图像抽帧。检测及特征提取模型, 根据不同机场不同场景进行选择。特别地, 相同场景下选择不同检测模型, 对结果影响较小, 因此, 在实际应用中, 更倾向选择较为轻量化模型。针对 γ 及 λ 值的设置, 计算 OSNet 网络中匹配行人的相似度, 相似度越低, 则 γ 及 λ 值表现最优。经实验, 当 γ 和 λ 的值为 0.6 和 0.7, 可实现在实际场景中最优性能, 图 4 为不同 γ 、 λ 值与相似度之间的变化关系。同样, 实验验证中融入路径以及时间因素。

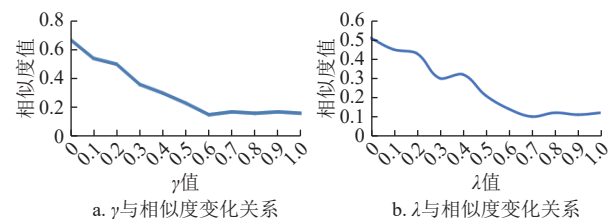


图 4 不同 γ 和 λ 与相似度变化关系

3 实验

3.1 实验数据

实验数据包含用于颜色自相关特征分析的数据。颜色自相关特征分析数据是通过 CP-VTON+^[24] 算法得到的换装图像。实验数据来源于两类场景: 一类为模拟场景, 具体地, 以一名工作人员为寻找人员, 设置路线, 共收集 19 路摄像头的视频, 用以验证本文方法的可行性; 一类收集整理了机场航站楼实际场景中的 8 个摄像头监控数据, 其中包括 2 个安检口, 以及 6 个通道摄像头数据。共标注 231 人作为实验的待寻找行人数据集合, 寻找行人数据集为所有通过 6 个通道摄像头的旅客, 共计 7 789 人。

3.2 实验设置

实验使用 PyTorch 深度学习框架。模型训练阶段, 基于 Linux 平台使用显存为 24 GB 的 RTX3090 的 GPU 显卡进行加速计算。在 ImageNet^[26] 数据集进行预训练参数初始化。行人重识别模块在机场航站楼数据集上进行训练, 损失函数为交叉熵损失函数。在训练过程中, 随机选择 M 个身份的 N 张图片, 形成一个训练批次。实验中 $M = 16$, $N = 4$ 。批处理中图像为 256×128 像素大小。实验中使用 Adam (adaptive moment estimation) 优化器优化模型。全局共设置 240 次迭代训练模型, 采用预学习率的策略使网络发挥更好性能。在前 20 个迭代将学习率从 3.5×10^{-5} 线性提升到 3.5×10^{-4} , 在第 80 次和第 120 次迭代学习率分别衰减到 3.5×10^{-5} 和 3.5×10^{-6} 。

针对路径以及时间设置, 在模拟场景下, 设置的摄像头路线为: $1 \rightarrow 2 \rightarrow \dots \rightarrow 19$, 人员上一摄像头出现与下一摄像头出现的时间不超过 2 min。在某机场实际场景, 设置摄像头路线如表 2 所示, 人员上一摄像头出现与下一摄像头出现的时间不超过 1.5 min。

表 2 实际场景人员路线

安检口摄像头	路线1	路线2
CAMERA1	CAMERA3	CAMERA4
	CAMERA5	CAMERA6
	CAMERA7	CAMERA8
CAMERA2	CAMERA3	CAMERA4
	CAMERA5	CAMERA6
	CAMERA7	CAMERA8

3.3 评价指标

方案的有效性通过定义 3 类指标进行衡量。召

回率 R_{recall} , 准确率 A_{ar} , 匹配行人 Top1 置信度 P 。

$$R_{\text{recall}} = \frac{\text{检测匹配正确目标数}}{\text{总目标数}}$$

$$A_{\text{ar}} = \frac{\text{正确匹配的检测目标数}}{\text{检测匹配正确的目标数}}$$

式中, 检测匹配正确目标数指在模型设置阈值后, 检测出大于阈值目标数。正确匹配的检测目标数是指模型识别出的目标为正确的样本目标数量。

3.4 实验结果

3.4.1 模拟实验结果分析

模拟实验采集某一工作人员的图像, 选定 19 段视频覆盖的路径, 让其随机行走, 验证本文方法的有效性。

表 3 中, Y 分别代表目标真实出现、模型检测到目标、检测到目标匹配正确。N 分别代表未出现、未检测到、检测不准确。在模拟数据场景下, 模型的召回率为 91%, 准确率为 86.7%。

表 3 模拟实验结果分析表

序号	是否真实出现	是否检测到	检测是否正确
1	Y	Y	Y
2	N	Y	N
3	N	N	Y
4	Y	N	N
5	Y	Y	Y
6	Y	Y	Y
7	Y	Y	Y
8	Y	Y	Y
9	Y	Y	Y
10	Y	Y	Y
11	Y	Y	Y
12	Y	Y	Y
13	N	Y	N
14	N	N	Y
15	N	Y	N
16	Y	Y	Y
17	N	N	Y
18	N	Y	N
19	N	Y	N

根据测试结果, 在单个摄像头中, 存在较高召回率 (>10%), 在多个摄像头下, 同一旅客漏检率与摄像头呈强烈负相关关系。如某旅客经过了 3 个摄像头, 则该旅客的漏检率为 $0.1^3 = 0.001$, 即 1%, 实验表明本文方法在航站楼场景下的有效性及可行性。

图 5 展示了在模拟场景中, Top1 的置信度数值 (蓝色分布点), 可以看出, Top1 的置信度 P 平均达到 0.75 (红色分布点)。

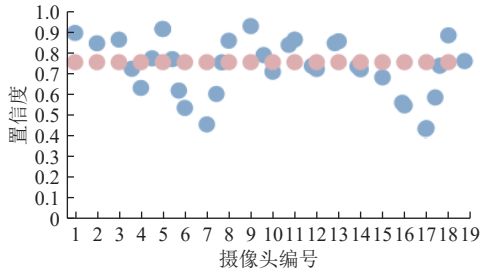


图 5 Top1 置信度分布图

3.4.2 航站楼实际场景实验结果分析

以机场实际场景 2 个安检通道为例, 获取通过安检通道的实际旅客 (2 个小时内通过安检口的旅客), 231 名旅客共测试 241 人次 (同一人在不同摄像头出现, 则按照人次统计)。在 6 个通道中摄像头中, 经统计检测匹配正确的目标数为 220 人次, R_{recall} 平均为 91.12%, A_{ar} 平均为 88.25%, 表 4 展示了测试摄像头的 R_{recall} 以及 A_{ar} 。

表 4 单个摄像头下检测旅客的 R_{recall} 以及 A_{ar}

摄像头名称	实际出现旅客数	检测匹配正确目标数	正确匹配检测目标数	$R_{\text{recall}}/\%$	$A_{\text{ar}}/\%$
CAMERA3	21	19	17	89.5	89.5
CAMERA4	24	22	20	91.7	91.0
CAMERA5	42	38	33	90.5	87.0
CAMERA6	39	36	32	92.4	88.0
CAMERA7	27	25	22	92.6	88.0
CAMERA8	88	80	69	90.0	86.0

表 4 正确匹配检测目标数为匹配行人 Top1 的值, 在航站楼实际场景数据集中, Top1 的置信度 P 平均达到了 0.72。同时, 计算了在 Top5 之内一定存在正确匹配检测目标数, 如表 5, 通过计算, 本文方法的准确率平均达到了 93.2%, 与 1.2 节中的结果对比, 本文方法优于现有方法。

表 5 检测旅客 Top5 中的 A_{ar}

摄像头名称	实际出现旅客数	检测匹配正确目标数	正确匹配检测目标数	$A_{\text{ar}}/\%$
CAMERA3	21	19	18	94
CAMERA4	24	22	21	95
CAMERA5	42	38	35	92
CAMERA6	39	36	34	94
CAMERA7	27	25	23	92
CAMERA8	88	80	74	92

值得注意的是, 文章中的方法可适应不同姿态变化场景 (正面、侧面、背面), 并适应局部颜色的变化 (脱衣、敞衣、包遮挡), 均可引起形态, 尤其外观的变化, 本文方法均可以较好的适应。

3.5 实验结果对比分析

将本文方法与现有方法进行比较。除了与 OSNet

网络进行比较, 同时选择了解决行人颜色特征变化引起的匹配不准相关方法。文献 [27] 提出一种基于身体部位的重识别模型 BPBreID, 通过预测身体部位注意力图生成基于目标身体部位特征, 该方法对遮挡和非判别性局部外观具有鲁棒性。文献 [28] 提出一个有效框架 LRe-ID 来共享一致特征并为每个样本提取判别性特征。文献 [29] 设计了一个分层掩模生成器 DPM, 利用分层语义选择整体和被遮挡输入图像的特征表示之间的可见模式空间。文献 [30] 设计互补数据增强策略 CDA, 通过随机拼补不同衣服的人物图像来满足 ID 空间, 模拟丰富的外观以增强对衣服变化的鲁棒性。文献 [31] 提出 MBUNet, 引入一个多生物特征分支, 提取生物各种特征, 如头部、颈部, 应对更换衣服而带来的特征变化。

分别从召回率、准确率以及 Top1 的置信度 P 进行指标的比较, 表 6 展示了模拟数据在各类方法性能展示 (调整了各个方法的 Top1 置信度 P 值, 使各个方法的 A_{ar} 以及 R_{recall} 均达到最优)。表 7 展示了实际场景数据上各类方法性能展示, 同样, 对于 Top1 置信度 P 值, 调整了各个方法的 Top1 置信度 P 值, 使各个方法的 A_{ar} 以及 R_{recall} 均达到最优。

表 6 针对模拟数据集的指标对比

方法	$A_{\text{ar}}/\%$	$R_{\text{recall}}/\%$	P
OSNet ^[8]	82.2	90.1	0.54
BPBreID ^[27]	84.2	88.3	0.46
LRe-ID ^[28]	85.7	89.8	0.58
DPM ^[29]	81.3	86.5	0.41
CDA ^[30]	85.6	91.2	0.65
MBUNet ^[31]	84.7	89.3	0.68
本文方案	86.7	91.0	0.75

表 7 针对实际场景数据集的指标对比

方法	$A_{\text{ar}}/\%$	$R_{\text{recall}}/\%$	P
OSNet ^[8]	83.2	88.5	0.54
BPBreID ^[27]	85.1	87.0	0.48
LRe-ID ^[28]	84.6	88.1	0.62
DPM ^[29]	82.7	87.2	0.40
CDA ^[30]	86.1	90.1	0.60
MBUNet ^[31]	85.9	91.3	0.71
本文方案	88.3	91.1	0.72

可以看出, 在模拟数据集上, 本文方法的 Top1 置信度值优于现有方法, 在高于平均召回率的同时, 准确率达到最优。在实际场景数据集上, 同样表现出了优秀的性能。相比较于 Yolo+OSNet 方法, 本文方法不仅可适应姿态的变化, 同样适应

局部颜色变化场景。但由于分支的增加, 模型的训练时间较 Yolo+OSNet 方法增长, 实际中, 训练时间长短不影响实际使用, 更关注召回率以及准确率。

3.6 模型性能分析

为探究特征自相关性和共同动态观测区域引入方法的有效性, 增强本文方法的可解释性, 进行了单一限制条件下的实验分析。

确定共同被观测区域后, 对图像块进行颜色信息统计分析以获取区域图像颜色表征。对同一行人及不同行人进行分析, 图 6 展示了旅客共同观测区域灰度图像颜色直方图, 实验采用颜色数值统计量进行分析, 以分析灰度图像灰度值的统计直方图信息为例进行分析。虚线上方为同一行人在直方图数量为 32 和 64 时的统计折线图, 虚线下方为不同行人在直方图数量为 32 和 64 时的统计折线图。可以看出, 对行人共同观测区域颜色统计量分析, 相同行人情况下, 在直方图数量为 32 和 64 的时候, 均保持高度一致性, 在不同行人情况下, 两种尺度统计直方图中, 相差很大, 即其颜色特征距离较远。

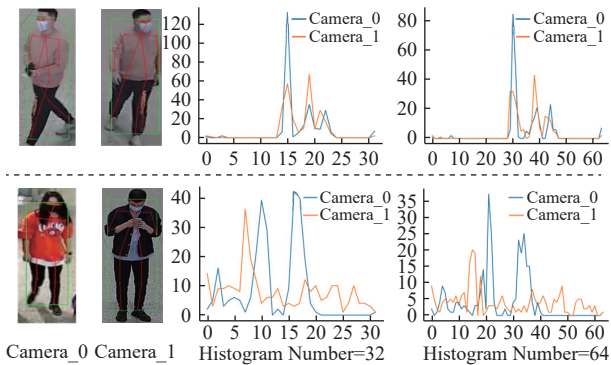


图 6 旅客共同观测区域灰度图像颜色直方图统计

同时, 以 OSNet 网络为基准, 分别增加特征自相关分析以及共同被观测区域。1) 仅增加特征自相关分析, 可以减少 1.2 节中由着装颜色差异巨大、外形边缘相似度高引起的匹配错误的情形, 但对于遮挡及脱衣等引起的局部颜色的变化难以适应, 同时对于较为相近着装表现效果极差。2) 仅增加共同观测区域, 在未考虑颜色特征自分析的情况下, 共同观测区域可以解决由敞衣、姿态变化等引起的颜色特征局部变化问题, 但不能解决 1.2 节中提到的颜色差异大, 边缘信息相似导致的匹配错误。3) 两者共同考虑时, 其表现出了优秀性能。

表 8 分别就仅加入颜色特征自相关或者共同被观测区域分别与整体方案进行了比较 (实际场景数据集), 其中 Top1 的置信度 P 设置为 0.5, 可以看

出, 无论是准确率还是召回率, 颜色自相关性特征以及共同被观测区域特征对其均有一定的影响。

表 8 特征自相关分析和共同被观测区域对网络的影响

算法	准确率	召回率
特征自相关性分析	87.41	90.16
共同被观测区域	85.96	89.68
两者结合	88.25	91.12

4 结束语

本文基于机场航站楼真实场景, 优化已有方法, 解决此场景行人重识别任务中, 现有算法在边缘特征相似的基础上, 导致颜色特征错误匹配的问题。特别地, 本文提出共同可观测区域概念, 通过定义行人动态变化区域, 对共同观测区域颜色统计分析, 获得鲁棒性较好的行人颜色特征, 用于对行人进行匹配识别, 解决由颜色、遮挡引起的匹配不准确问题。以机场模拟数据、实际场景数据为基础数据集进行实验, 表明本文方法比现有的重识别方法更优, 更适应航站楼场景。

参考文献

- [1] DALAL N, TRIGGS B. Histograms of oriented gradients for human detection[C]//Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition. New York: IEEE, 2005: 886-893.
- [2] LOWE D G. Distinctive image features from scale-invariant keypoints[J]. *International Journal of Computer Vision*, 2004, 60(2): 91-110.
- [3] LIAO S C, HU Y, ZHU X Y. Person re-identification by Local Maximal Occurrence representation and metric learning[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. New York: IEEE, 2015: 2197-2206.
- [4] XIONG F, GOU M R, CAMPS O, et al. Person re-identification using kernel-based metric learning methods[C]//Computer Vision - ECCV 2014. Cham: Springer International Publishing, 2014: 1-16.
- [5] WANG P Y, ZHAO Z C, SU F. HOREID: Deep high-order mapping enhances pose alignment for person re-identification[J]. *IEEE Transactions on Image Processing*, 2021, 30: 2908-2922.
- [6] 王鹏, 宋晓宁, 吴小俊, 等. 用于行人重识别的多类型特征网络[J]. *模式识别与人工智能*, 2020, 33(10): 879-888. WANG P, SONG X N, WU X J, et al. Multi-type features network for person re-identification[J]. *Pattern Recognition and Artificial Intelligence*, 2020, 33(10): 879-888.
- [7] ZHONG Y J, WANG Y W, ZHANG S L. Progressive feature enhancement for person re-identification[J]. *IEEE Transactions on Image Processing*, 2021, 30: 8384-8395.
- [8] ZHOU K Y, YANG Y X, CAVALLARO A. Learning

- generalisable omni-scale representations for person re-identification[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2021, 44(9): 5056-5069.
- [9] HE L X, LIU W. Guided saliency feature learning for person re-identification in crowded scenes[C]//*Computer Vision – ECCV 2020*. Cham: Springer International Publishing, 2020: 357-373.
- [10] 王继民. 融合行人属性信息的行人重识别研究[J]. *集成电路应用*, 2023, 40(12): 420-424.
WANG J M. Study on pedestrian re-identification by integrating pedestrian attribute information[J]. *Application of IC*, 2023, 40(12): 420-424.
- [11] CHEN T L, DING S J, XIE J Y. ABD-net: Attentive but diverse person re-identification[C]//*Proceedings of the IEEE/CVF International Conference on Computer Vision*. New York: IEEE, 2019: 8350-8360.
- [12] LUO H, GU Y Z, LIAO X Y. Bag of tricks and a strong baseline for deep person re-identification[C]//*Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*. Long Beach: IEEE, 2019: 4321-4329.
- [13] 伍子强, 常虹, 马丙鹏. 基于融合关系学习网络的行人重识别[J]. *模式识别与人工智能*, 2021, 34(9): 798-808.
WU Z Q, CHANG H, MA B P. Person re-identification based on fusion relationship learning network[J]. *Pattern Recognition and Artificial Intelligence*, 2021, 34(9): 798-808.
- [14] LIAO S C, SHAO L. Graph sampling based deep metric learning for generalizable person re-identification[C]//*Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. New York: IEEE, 2022: 7349-7358.
- [15] 孙志伟, 吴广群, 马永军. 基于区域特征补全和细粒度特征注意力的行人重识别方法[J]. *天津科技大学学报*, 2024, 39(5): 63-71.
SUN Z W, WU G Q, MA Y J. A pedestrian re-identification method based on regional feature completion and fine-grained feature attention[J]. *Journal of Tianjin University of Science & Technology*, 2024, 39(5): 63-71.
- [16] GONG Y P, HUANG L Q, CHEN L F. Person re-identification method based on color attack and joint defence[C]//*Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*. New York: IEEE, 2022: 4312-4321.
- [17] GU X Q, CHANG H, MA B P, et al. Clothes-changing person re-identification with RGB modality only[C]//*Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. New York: IEEE, 2022: 1050-1059.
- [18] JIN X, HE T Y, ZHENG K C, et al. Cloth-changing person re-identification from A single image with gait prediction and regularization[C]//*Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. New York: IEEE, 2022: 14258-14267.
- [19] 杨萍, 吴晓红, 何小海. 逐点特征匹配的跨域行人重识别方法[J]. *模式识别与人工智能*, 2022, 35(6): 516-525.
YANG P, WU X H, HE X H, et al. Cross-domain person re-identification method based on point-by-point feature matching[J]. *Pattern Recognition and Artificial Intelligence*, 2022, 35(6): 516-525.
- [20] 胡涌涛, 黄洪琼. 结合特征融合和通道注意力的多分支换装行人重识别[J]. *计算机工程*, 2025, 51(1): 225-234.
HU Y T, HUANG H Q. Multi-branch clothes-changing person re-identification with feature fusion and channel attention[J]. *Computer Engineering*, 2025, 51(1): 225-234.
- [21] GE Z, LIU S T, WANG F, et al. YOLOX: Exceeding YOLO series in 2021[EB/OL]. (2021-07-18)[2024-07-04]. <https://arxiv.org/abs/2107.08430>.
- [22] WOJKE N, BEWLEY A, PAULUS D. Simple online and realtime tracking with a deep association metric[C]//*Proceedings of the IEEE International Conference on Image Processing*. New York: IEEE, 2017: 3645-3649.
- [23] CHEN W H, XU X Z, JIA J, et al. Beyond appearance: A semantic controllable self-supervised learning framework for human-centric visual tasks[C]//*Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. New York: IEEE, 2023: 15050-15061.
- [24] MINAR M R, TUAN T T, AHN H, et al. Cp-vton+: Clothing shape and texture preserving image-based virtual try-on[C]//*Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. New York: IEEE, 2020: 10-14.
- [25] CAO Z, SIMON T, WEI S H, et al. Realtime multi-person 2D pose estimation using part affinity fields[C]//*Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. New York: IEEE, 2017: 1302-1310.
- [26] DENG J, DONG W, SOCHER R, et al. ImageNet: A large-scale hierarchical image database[C]//*Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. New York: IEEE, 2009: 248-255.
- [27] SOMERS V, DE VLEESCHOUWER C, ALAHI A. Body part-based representation learning for occluded person re-identification[C]//*Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*. New York: IEEE, 2023: 1613-1623.
- [28] HUANG J Z, YU X H, AN D, et al. Learning consistent region features for lifelong person re-identification[J]. *Pattern Recognition*, 2023, 144: 109837.
- [29] TAN L, DAI P Y, JI R R, et al. Dynamic prototype mask for occluded person re-identification[C]//*Proceedings of the 30th ACM International Conference on Multimedia*. New York: ACM, 2022: 531-540.
- [30] JIA X M, ZHONG X, YE M, et al. Complementary data augmentation for cloth-changing person re-identification[J]. *IEEE Transactions on Image Processing*, 2022, 31: 4227-4239.
- [31] ZHANG G, LIU J, CHEN Y, et al. Multi-biometric unified network for cloth-changing person re-identification[J]. *IEEE Transactions on Image Processing*, 2023, 32: 4555-4566.