E-mail: jig@aircas.ac.cn
Website: www.cjig.cn
Tel: 010-58887035
中国图象图形学报
JOURNAL OF IMAGE AND GRAPHICS

中图法分类号:TP391.4　　**文献标识码**:A　　**文章编号**:1006-8961(2023)05-1396-13

论文引用格式:Liu Y Q and Ma B P. 2023. Sketch images-guided clothes-changing person re-identification. Journal of Image and Graphics, 28(05):1396-1408(刘宇奇, 马丙鹏. 2023. 素描图像指导的换装行人重识别. 中国图象图形学报, 28(05):1396-1408)[DOI:10.11834/jig.220783]

素描图像指导的换装行人重识别

刘宇奇,马丙鹏*

中国科学院大学计算机科学与技术学院,北京　100049

摘　要:**目的**　传统行人重识别方法提取到的特征中包含大量的衣物信息,在换装行人重识别问题中,依靠衣物相关的信息难以准确判别行人身份,使模型性能显著下降;虽然一些方法从轮廓图像中提取行人的体型信息以增强行人特征,但轮廓图像的质量参差不齐,鲁棒性差。针对这些问题,本文提出一种素描图像指导的换装行人重识别方法。**方法**　首先,本文认为相对于轮廓图像,素描图像能够提供更鲁棒、更精准的行人体型信息,因此本文使用素描图像提取行人的体型信息,并将其融入表观特征以获取完备的行人特征。然后,提出一个基于素描图像的衣物无关权重指导模块,进一步使用素描图像中的衣物位置信息指导表观特征的提取过程,从而减少表观特征中的衣物信息,增强表观特征的判别力。**结果**　在LTCC(long-term cloth changing)和PRCC(person re-identification under moderate clothing change)两个常用换装行人数据集上,本文方法与最先进的方法进行了对比。相较于先进方法,在LTCC和PRCC数据集上,本文方法在Rank-1性能指标上分别提高了6.5%和3.9%。实验结果表明,素描图像在鲁棒性和准确性上均优于轮廓图像,能够更好地获取行人体型信息,并且能够为表观特征提供更多体型互补信息。**结论**　提出的衣物无关权重指导模块能有效减少行人表观特征中衣物信息的含量;提出的素描图像指导的换装行人重识别方法有效获取了包含衣物无关表观特征和体型特征在内的完备行人特征。

关键词:计算机视觉;换装行人重识别;素描图像;表观特征;体型特征;双流网络

Sketch images-guided clothes-changing person re-identification

Liu Yuqi, Ma Bingpeng*

School of Computer Science and Technology, University of Chinese Academy of Sciences, Beijing　100049, China

Abstract: **Objective**　Video surveillance systems have been widely used for public security such as tracking the suspect and looking for missing person. It is really expensive and time-consuming to analyze videos manually. Person re-identification (ReID) aims to match the same person appearing at different times and places under non-overlapping cameras. With the development of deep learning, ReID has gained significant performance increment on the benchmark. It is known that in ReID, the retrieval mainly depends on apparent cues such as clothes information. However, if the surveillance video is captured in a long-time span, people may change their clothes when appearing in the surveillance system. Besides, criminals may also change their clothes to evade surveillance cameras. In such cases, existing methods are likely to fail because they extract unreliable clothes-relevant features. Clothes-changing problem is inevitable in the real-scene application of ReID. Recently, clothes-changing ReID receives a lot of attention. In clothes-changing ReID, every person wears multiple outfits. The key point to clothes-changing ReID is to extract discriminative clothes-irrelevant features from

收稿日期:2022-08-08;**修回日期**:2023-01-03;**预印本日期**:2023-01-10

*通信作者:马丙鹏　bpma@ucas.ac.cn

基金项目:国家自然科学基金项目(62276246);中央高校基本科研业务费专项资金资助

Supported by: National Natural Science Foundation of China (62276246); Fundamental Research Funds for the Central Universities

images with different clothes. Usually, body shape can be used to identify people. Some existing methods extract body shape information from contour images but suffer from low image quality and poor robustness. To resolve these problems, we propose a sketch images-guided method for clothes-changing person re-identification method. There are two main approaches in existing methods: 1) extracting clothes-irrelevant information such as key points, pose and gait and fuse clothes-irrelevant information into person features. 2) Decoupling clothes-irrelevant and clothes-relevant feature using an encoder-decoder architecture. **Method** First, to improve the accuracy and robustness of body shape information, we propose to obtain more accurate and robust shape information from sketch images rather than contour images. Then we use an extra independently-trained network to extract the shape features of the person. Additionally, to reduce the clothes information in visual features and improve the discrimination of visual features, we propose a clothes-irrelevant weight guidance module based on sketch images. The module further uses the clothes position information in sketch images to guide the extraction process of visual features. With the guidance, the model can extract features with less clothes information. We use a two-stream network to fuse the shape feature and clothes-irrelevant apparent feature to get the complete person feature. We implement our method by using python and PyTorch. We train our network with one NVIDIA 3090 GPU device. We perform random horizontal flips and random erasing for augmentation. We use Adam optimizer and the learning rate is set to 0.000 35. The learning rate will decay per 20 batches. **Result** The performance of the proposed method is evaluated on two public clothes-changing dataset: long-term cloth changing (LTCC) and person re-identification under moderate clothing change (PRCC). Our method overperforms the state-of-the-art methods on both two datasets. The proposed method obtains 38.0% Rank-1 and 15.9% mean average precision (mAP) on LTCC dataset; 55.5% Rank-1 and 52.6% mAP on PRCC dataset. The results of the ablation experiment demonstrate that the sketch images have their priority in robustness and accuracy compared with contour images. Visible results show that our proposed method can effectively weaken the model's attention on the clothes area. **Conclusion** We propose a better way to extract body shape information and propose a sketch-based guidance module, which utilizes clothes-irrelevant information to wipe out clothing information in visual features. Experiments show that sketch images are superior to contour images in robustness and accuracy. Sketch images can provide more body shape information as a complement to visual features than contour images. The proposed clothes-irrelevant weight guidance module can effectively reduce clothing information in visual features. Our proposed sketch images-guided clothes-changing person re-identification method effectively extracts complete person features, which include clothing-irrelevant visual features and body shape features.

Key words: computer vision; clothes-changing person re-identification; sketch images; appearance feature; shape feature; two-stream network

0 引 言

行人重识别(person re-identification,ReID)旨在从不重叠的摄像机视角下匹配出不同时间、地点出现的同一行人,在智能安防、罪犯抓捕以及走失人员寻找等多种实际问题中有着重要作用。经过多年探索,行人重识别领域已经取得长足发展(Zhang等,2018;Zheng等,2019;史维东 等,2020;王新年 等,2020;吴岸聪 等,2022;Ye等,2022)。然而,目前绝大多数方法没有考虑行人的换装问题,而是默认一个行人一直穿着同一套衣物。因而这些方法提取到的行人的表观特征中包含较大比重的衣物信息。但是,在实际应用场景中,行人往往会更换衣物,此时由于行人所着衣物发生变化,表观特征中的衣物信息不能用于判别行人身份,这导致表观特征判别力显著下降,鲁棒性差,不同的行人因穿着类似衣物而被误认为是同一个人的情况时有发生。

为了解决上述行人换装问题,Yang等人(2019)首次提出了换装行人重识别问题。与常规行人重识别问题不同,该问题需要对不同摄像头下、穿着不同衣物的行人进行检索,因此行人的标注不仅包含身份与摄像头标签,还额外包含行人的衣物标签。

目前换装行人重识别方法大致可分为两类。一类方法对衣物无关特征和衣物相关特征进行解耦合(Lorenz等,2019),使用解耦合得到的衣物无关特征进行重识别,如Li等人(2021)提出分别使用两个编码器单独提取图像的衣物无关特征(即形状特征)和

衣物相关特征(即颜色特征),通过使用灰度图像的形状特征和彩色图像的颜色特征进行图像重建的方式对形状和颜色特征的生成过程进行监督,最后使用学习到的衣物无关特征进行判别。这类方法的缺点在于难以在解耦合过程中准确地去除所有衣物相关特征,例如衣物纹理特征。其原因在于衣物纹理特征分布位置广泛,受行人姿态影响大,很难将它们全部找出,且容易与背景产生混淆,这使得解耦得到的衣物无关特征中仍留存衣物信息,不完全的解耦合会导致解耦出的衣物无关特征判别力下降。

另一类方法通过使用其他任务中的模型获取衣物无关的辅助信息增强行人特征。辅助信息包括行人关键点位置、行人轮廓(Eitz等,2012)和行人步态等。Qian等人(2020)提出使用姿态检测模型对姿态关键点进行检测,将关键点信息映射为形状信息融入行人的表观特征中,使用注意力模块将融合特征解耦为衣物相关特征和衣物无关特征,分别使用衣物类别和行人类别计算两个交叉熵(cross entropy, CE)损失进行监督。Hong等人(2021)提出了一个密集互相学习模型,使得行人原图像和行人语义分割得到的轮廓图像在训练过程中能够密集地互相学习,从而使模型提取到的特征同时包含行人的表观特征和体型特征,此外,作者还加入了一个姿态预测模块将行人姿态聚类为3种类别,对其分别处理以降低姿态变化导致的类内偏差。Chen等人(2021)提出通过对行人图像进行3D重建挖掘行人的体型信息,由编码器生成行人的姿势与体型特征进行3D重建,使用训练好的姿态、轮廓预测模型对重建过程进行监督,在测试阶段使用重建过程中由编码器生成的形状特征作为辅助信息进行检索。相比于第1类方法,此类方法的优点在于可以直接从其他任务的模型中获取衣物无关的辅助信息以增强表观特征的辨识力。此类方法更容易获取衣物无关程度更高的行人特征,在性能表现上往往优于解耦合方法,因此本文也选择此类使用辅助信息的方法展开研究。

但是,使用辅助信息增强行人特征的方法仍然面临两个巨大挑战。1)如何获取较为准确的辅助信息。由于获取辅助信息所使用的模型来源于其他任务,在行人重识别的数据集上使用时会受数据集的域偏差影响,且模型的效果也受其本身性能上限的限制,此外数据集中图像质量参差不齐,从低质量图像中提取到的信息往往不够鲁棒。2)如何有效使用辅助信息。各类辅助信息中包含的衣物无关信息各不相同,针对每一种辅助信息,采取合适的使用方式,方能获取具有辨识力的衣物无关特征。

本文从上述两个关键点出发,对使用行人体型信息作为辅助信息的方法展开研究,提出一种高效、鲁棒的行人体型信息获取方式,并基于此方式进一步提出一个衣物无关权重指导模块,对行人表观特征的提取过程进行指导,以降低表观特征中衣物信息含量,提高表观特征在换装场景下的辨识力,最终显著提高换装行人重识别任务的性能。

1 一种鲁棒的体型信息获取方法

在使用辅助信息的换装行人重识别方法中,部分方法(Yu等,2020b;Hong等,2021)从人体解析模型生成的行人部分区域或全身轮廓中提取行人的体型信息,并将其作为辅助信息。这种方式产生的行人轮廓图像质量参差不齐,大量轮廓图像包含较多的缺陷,难以从中获取到准确的行人体型信息。针对该问题,本文提出一种改进方法,从信息量更丰富的素描图像中获取更鲁棒、更准确的行人体型信息。

1.1 轮廓图像分析

行人的轮廓图像如图1所示,通常可由人体解析方法获得。人体解析问题的研究目前取得了一定进展,但仍有一些难点未能很好地解决。例如,图像中人体会穿着各种样式、颜色和纹理的服装,容易与背景和其他部位发生混淆;图像中人体姿态变化很大,身体部位经常会发生重叠,导致难以精准解析出每一个部位的位置;多样且复杂的背景也会一定程度上影响解析的精度。

因此,使用人体解析模型生成轮廓图像容易受到行人重识别数据集中常见问题的影响,如光照条件过差、行人被遮挡以及行人姿势复杂等。当行人图像质量较高、背景较为简单时,人体解析模型能够准确定位人体所在位置,生成高质量的轮廓图像,如图1前两组图像(3幅同列图像为一组)。而当行人图像出现上述问题时,人体解析模型难以区分行人与背景的边界,不能精准地定位每个人体部件所在位置,最终生成低质量的轮廓图像,如图1后3组图像所示。这种低质量的轮廓图像无法准确描述行人的体型信息,使用此类图像难以有效改善在换装场景下行人的识别性能。


(a) 原始图像


(b) 轮廓图像

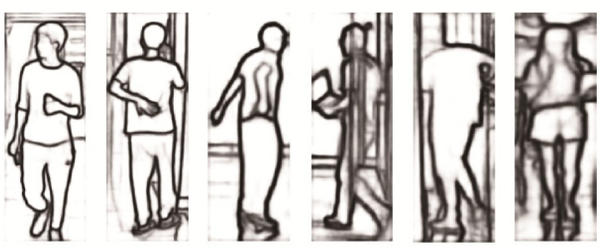

(c) 素描图像

图 1　行人图像及其对应的轮廓图和素描图

Fig. 1　Person images and their corresponding contour images and sketch images ((a) original images; (b) contour images; (c) sketch images)

1.2　素描图像介绍

1.2.1　素描图像的鲁棒性

本文中的素描图像由边缘检测方法生成。沿用 Bhattarai 等人(2020)的描述方式，本文使用素描图像描述行人图像经过边缘检测得到的图像。

如图 1 所示，相较于轮廓图像，素描图像对于行人重识别数据集中的一些问题，如光照变化和遮挡，具有更强的鲁棒性，这是二者的任务目标差异导致的。人体解析任务不仅要划分行人与背景的边界，同时要考虑行人内部的各个部位如何划分，因此它的任务目标相较于使用目的存在一定冗余。当原图像质量不高时，由于模型很难准确解析行人每个部位的位置，任务目标的冗余会加剧生成图像的低质量问题，影响边缘的分割效果。而边缘检测任务的目的是检测出图像所有的边缘(Yu 等，2016)，行人的轮廓是行人图像边缘的主要组成部分。边缘检测的任务目标与使用其生成的素描图像的目的间不存在差异，因此边缘检测模型生成的素描图像更为准确和鲁棒。

以行人重识别数据集中较常见的两种情况，光照条件较差、行人被物体遮挡为例，说明素描图像相较于轮廓图像在准确性和鲁棒性方面的优势：

1)当光照条件较差时，如图 1 第 5 组图像所示。图像中行人部分亮度较低，轮廓图像会因人体解析模型解析不到上半身的位置丢失上半部分的轮廓。而由于行人的身体边缘连贯，能在亮度不佳的情况下与背景及其他物体保持一定的区分度，边缘检测模型仍然可以检测到较准确的行人边缘。

2)当行人被物体遮挡时，如图 1 第 4 组图像所示。在轮廓图像中，因人体解析模型将被遮挡部分识别为背景区域，导致行人的下半身区域轮廓缺失。在素描图像中，边缘检测模型虽难以避免遮挡物体的干扰，但在遮挡区域外能够准确地找出行人边缘。

1.2.2　素描图像包含额外体型信息

根据 Hertzmann(2020，2021)提出的现实主义假说，特定条件下的素描图像包含一定量的 3D 信息。人的视觉系统能从素描中感知出物体的类别和 3D 形状，为了对这一现象做出解释，Hertzmann 提出了现实主义假说，将素描定义为勃朗白材质、单一光源的 3D 物体建模在某一视角下获得的 2D 图像，因此人可以从 2D 的素描中获取到 3D 物体的信息。同时，Hertzman 指出，街景以及室内拍摄的一些照片的素描能生成基本正确的深度图像，这意味着从某些场景下的素描图像中可以提取出 3D 信息，而行人重识别问题中的图像几乎都属于这两种情况。因此，根据现实主义假说的理论分析可知，素描图像在特定条件下会比轮廓图像多包含一定的 3D 信息，如行人和场景的深度信息，这些信息是轮廓图像中不具有的。而这些信息与行人体型相关，有助于换装场景下的行人身份判别。

2　素描图像指导的行人完备特征获取方法

在换装场景下，由于行人表观特征中占主导地位的衣物表观特征已经不足以作为准确判别身份的依据，且表观特征中缺乏体型信息，传统的行人重识别方法的性能显著下降。针对这一问题，本文提出一个素描图像指导的行人完备特征获取方法，利用素描图像中行人衣物的位置信息，减少行人表观特

征中的衣物信息;同时,从素描图像中提取体型信息补全行人表观特征。具体来说,该方法由一个基于素描图像的衣物无关权重指导模块和一个双流网络组成。

2.1 基于素描图像的衣物无关权重指导模块

基于素描图像的衣物无关权重指导模块使用素描图像中行人衣物的位置信息对行人图像的特征提取进行指导。模块的核心部分是一个衣物无关权重矩阵,与行人图像的衣物部分对应的位置,在权重矩阵中会被赋予一个较低的权重,反之亦然。权重矩阵可以有效降低对行人衣物部分的关注度,从而获得包含更少衣物表观信息的特征,增强特征的判别力。

在素描图像中,灰度值相对较低的地方主要是行人的边缘,行人的大部分衣物所在位置的灰度值较高,大部分背景区域的灰度值也较高。因此,对于任意一个像素点,如果其灰度值较大,那么这个像素点应当位于行人的衣物区域或背景区域,在特征提取过程中需要减少对这一区域的关注。基于这一思想,本文提出了一个衣物无关权重矩阵,用于表示行人图像中每一个位置的衣物无关度。在该矩阵中,图像中每一个像素点所在位置的权重由该点的灰度值计算得到。在图2中,左侧图像为行人图像,中间图像为素描图像,右侧图像为由素描图像得到的权重矩阵的可视化图像。对比素描图像和权重矩阵可视化图像可知,将素描图像转化为权重矩阵后,大部分行人的衣物部分和背景区域都被赋予了较低的权重。这表明权重矩阵能够较为准确地表达每一个位置的衣物无关度。

衣物无关权重矩阵可以用于生成更具有鉴别力的特征。具体来说,对于衣物无关权重矩阵中权重低的衣物部分,在特征提取过程中应当减少对其的关注度,反之亦然。通过权重矩阵,表观特征中衣物信息的含量会显著减少,因而在换装场景下判别力增强。

使用两个子模块构建基于素描图像的衣物无关权重指导模块,即生成子模块和指导子模块。其中,生成子模块实现从素描图像中提取衣物无关权重信息,指导子模块使用衣物无关权重信息指导行人表观特征的提取过程。基于素描图像的衣物无关权重指导模块可按下述方式实现。

1)生成子模块。素描图像是灰度图,用$\boldsymbol{I}_{\mathrm{gs}}$表示。由于素描图像中衣物或背景部分的灰度值较大,而其在权重矩阵中的权重较低,因此对$\boldsymbol{I}_{\mathrm{gs}}$进行灰度值反转,并进一步将灰度值归一化至[0,1]区间。通过这种方式,可得到生成素描图像的原图像在特征提取过程中所需要的衣物无关位置权重矩阵$\boldsymbol{W}$,$\boldsymbol{W}$中的值代表了其对应位置的衣物无关度。

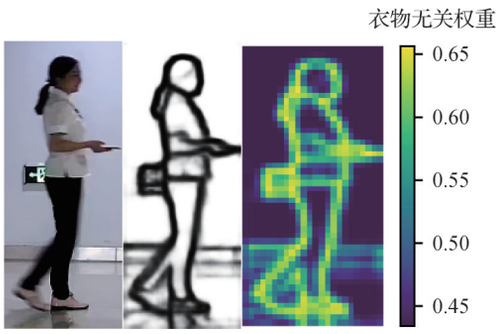


图2 衣物无关位置权重矩阵示意图

Fig. 2 Clothes-irrelevant position weight matrix schematic

2)指导子模块。该子模块通过衣物无关权重矩阵实现对行人表观特征提取的指导。具体来说,衣物无关权重矩阵包含每个位置的衣物无关度,由于素描图像和原图像在位置关系上完全对齐,因此权重矩阵可以直接指导行人图像的特征提取。将矩阵与网络在特定阶段后的输出特征加权求和即可有效降低特征中衣物所在区域的权重,最终降低表观特征中衣物信息的含量,即

$$\begin{aligned}\boldsymbol{f}_n' &= \boldsymbol{f}_n + \boldsymbol{f}_n \otimes \boldsymbol{W}_n \\ \boldsymbol{f}_{n+1} &= F_{n+1}(\boldsymbol{f}_n')\end{aligned} \tag{1}$$

式中,$\boldsymbol{f}_n$表示网络第n阶段输出的特征,⊗表示逐元素相乘,$\boldsymbol{W}_n$表示与$\boldsymbol{f}_n$对应的衣物无关位置权重矩阵,$\boldsymbol{f}_n'$表示经过加权后的特征,F_n表示网络的第n阶段。

在网络中,相对其输入,在宽和高的维度上每一个阶段的输出都会缩小。为保持空间位置语义上的一致性,衣物无关位置权重矩阵也进行相应的降采样,从而与特征保持相同大小的宽和高。降采样以平均池化的方式进行。衣物无关权重指导模块的位置及降采样过程使用的参数将在3.5节进行讨论。

2.2 模型整体结构及损失函数

2.2.1 模型整体结构

模型整体结构如图3所示。采用双分支特征融合结构作为网络的主体结构,两个分支的骨干网络均为ResNet-50(50-layer residual network)。双分支

结构通过在行人的表观特征中补足形状(体型)信息来降低卷积神经网络(convolution neural network，CNN)对纹理特征的归纳偏置的影响，进而使行人特征在更加鲁棒的同时，也能包含丰富的表观信息和体型信息，最终获取完备的行人特征。

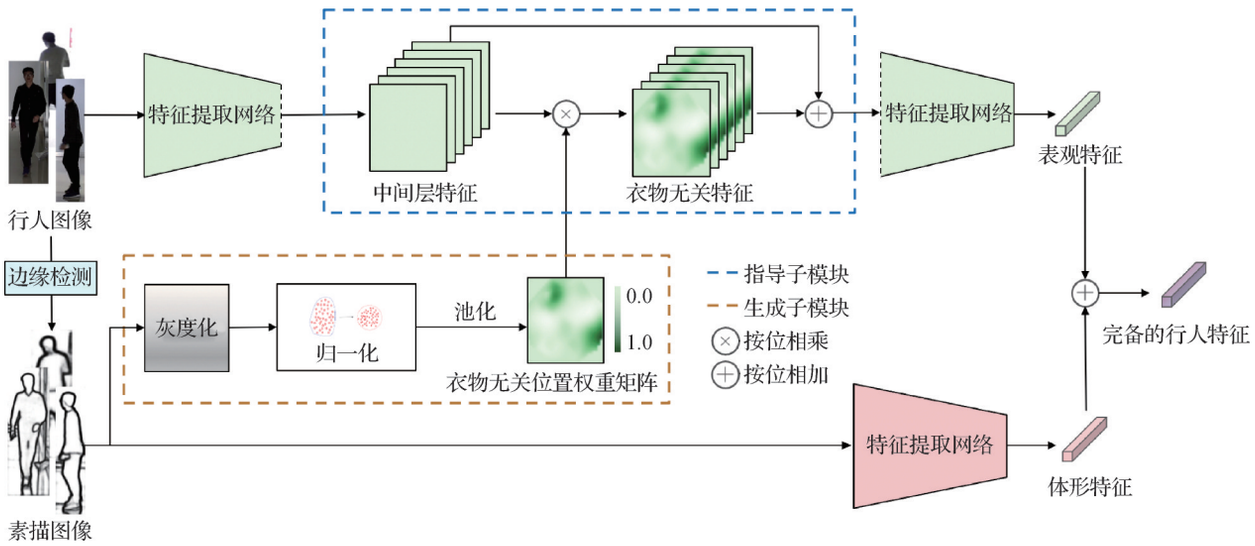


图3　模型整体结构

Fig. 3　Overall structure of the model

Geirhos 等人(2018)指出相较于形状，卷积神经网络对纹理具有更强的归纳偏置，这导致传统行人重识别方法提取到的表观特征中形状信息很少，在换装场景下性能下降显著。素描图像除能提供衣物无关的位置信息外，本身也包含了行人的体型信息，这些信息难以从行人图像中直接获得。通过双分支结构，可以将体型信息融入行人的特征中，增加特征的信息量，增强特征的鲁棒性，使其更适于换装行人重识别任务。

由于行人图像和素描图像之间存在一定的域差距，因此两个分支分别在两种图像上单独训练，从而获得表观和体型特征的最佳表示。衣物无关权重指导模块放置于行人图像分支的特定阶段后。在分别提取到两个分支的特征后，对二者进行融合以获取包含表观信息和体型信息在内的完整的衣物无关行人特征，即

$$\boldsymbol{f}_c = w_1 \times \boldsymbol{f}_a + w_2 \times \boldsymbol{f}_s \tag{2}$$

式中，$\boldsymbol{f}_a$ 与 $\boldsymbol{f}_s$ 分别表示表观特征与体型特征。w_1 和 w_2 均大于 0，且 $w_1 + w_2 = 1$。

2.2.2　损失函数

考虑到目前许多行人重识别方法都结合使用交叉熵损失和难采样三元组损失作为损失函数，并取得了较好的效果，本文同样结合使用这两种损失作为损失函数。

交叉熵损失函数通过最小化真实概率分布与预测概率分布之间的差异对模型进行优化，使模型能够学习到类别相关的特征，即

$$L_{\mathrm{ce}} = -\sum_{i=1}^{N} y_i \log\left(\frac{\mathrm{e}^{\boldsymbol{W}_k^{\mathrm{T}} \boldsymbol{f}_i}}{\sum_{k=1}^{C} \mathrm{e}^{\boldsymbol{W}_k^{\mathrm{T}} \boldsymbol{f}_i}}\right) \tag{3}$$

式中，y_i 和 $\boldsymbol{f}_i$ 分别表示第 i 个样本的真实标签和特征，$\boldsymbol{W}_k$ 表示类别 k 的权重向量。

难样本三元组损失挖掘一个训练批次中所有样本的最难三元组，即距离最近的负样本对和距离最远的正样本对，使正负样本对间的距离差大于阈值，有助于网络辨别表观特征相似而身份不同的样本对，即

$$\begin{aligned} L_{\mathrm{tri}} &= \sum_{i=1}^{P}\sum_{a=1}^{K}[\,D^*_{a,p} - D^*_{a,n} + m\,]_+ \\ D^*_{a,p} &= \max_{p=1,\cdots,K} D(f(x_a^i), f(x_p^i)) \\ D^*_{a,n} &= \min_{\substack{j=1,\cdots,P \\ n=1,\cdots,K \\ j\neq i}} D(f(x_a^i), f(x_n^j)) \end{aligned} \tag{4}$$

式中，P 表示一个批次中随机选取 P 个行人，K 表示一个行人选取 K 个样本，f 表示特征提取网络，D 表示欧氏距离。

模型的损失函数 L 可以表示为

$$L = w_1 \times L_{ce} + w_2 \times L_{tri} \tag{5}$$

式中，L_{ce}和L_{tri}分别表示交叉熵损失和难样本三元组损失。w_1和w_2均大于0，且$w_1+w_2=1$。

3 实　验

为验证素描图像相较于轮廓图像的优越性，以及提出的基于素描图像的衣物无关权重指导模块的有效性，对提出的方法进行测试，并与其他先进方法进行对比。

3.1 数据集和评价指标

本文在LTCC（long-term cloth changing）（Qian等，2020）和PRCC（person re-identification under moderate clothing change）（Yang等，2019）两个换装行人重识别数据集上对所提方法进行测试。LTCC数据集主要在室内场景下拍摄，包含152个行人，17 138幅图像，478套不同的服装，由12个摄像头拍摄得到，数据集中包含许多光照变化、姿势复杂以及发生遮挡的行人图像，因此具有一定的挑战性。PRCC数据集是一个大型的换装行人重识别数据集，包含221个行人，33 698幅图像，由3个摄像头拍摄得到。

本文使用的评价指标为行人重识别任务中的标准评价指标，即累计匹配特征（cumulative match characteristic，CMC）曲线中的首个匹配率Rank-1和平均精度均值（mean average precision，mAP）。

3.2 实验环境及网络参数设置

实验环境使用的CPU为Intel（R）Xeon（R）CPU E5-2620 v4@2.10 GHz，操作系统为Ubuntu16.04，GPU为NVIDIA GeForce RTX 3090。基于python 3.7.2和深度学习框架pytorch1.10.0完成模型的实现，模型在Deng等人（2009）提出的ImageNet数据集上进行预训练。

在数据的预处理阶段，对训练集的数据依次进行随机拉伸1/8长度并裁剪到原先的大小、调整大小至384×192像素、随机水平翻转、正则化以及随机擦除，调整测试集的数据大小至384×192像素并进行正则化。训练过程中，训练批次大小为32，每个行人随机挑选4个样本，使用Adam优化器优化网络参数，初始学习率（learning rate，LR）为0.000 35，权重衰减率为0.000 5，总共迭代训练60次，其中学习率分别在第20、40次迭代后降低至之前的0.1。

此外，在LTCC数据集上进行测试时，衣物无关权重模块放置在表观特征提取网络的第3层后，平均池化的核大小为5；在PRCC数据集上进行测试时，衣物无关权重模块放置在表观特征提取网络的第1层后，平均池化的核大小为7。

3.3 与现有方法的比较

为验证本文方法的有效性，与多种主流换装行人重识别方法在LTCC和PRCC两个数据集上进行对比，结果如表1所示。可以看出，在LTCC数据集上，本文方法比先进的同类方法UCAD（universal clothing attribute disentanglement）在Rank-1和mAP指标上分别提高了6.5%和0.8%；在PRCC数据集上，本文方法比先进的同类方法ADC（attentive decoupling）在Rank-1指标上提高了3.9%。这些结果说明本文方法有效地使用素描图像减少表观特征中的衣物信息，获取包含表观信息和体型信息的完整行人特征，验证了本文方法的有效性。

在本文对比的方法中，Qian等人（2020）提出的CESD（cloth-elimination shape-distillation）方法、Chen等人（2021）提出的3DSL（3D shape learning）方法、Jin等人（2022）提出的GI-ReID（gait recognition as an auxiliary task to drive the Image ReID）方法和Yan等人（2022）提出的UCAD方法与本文方法都属于使用额外信息辅助的方法。CESD提取行人的关键点信息，并将其映射为形状特征以辅助衣物相关和衣物无关特征的分离，但辅助过程缺乏有效的监督和可视化分析，难以保证辅助信息得到充分利用，此外，关键点信息的提取也会受行人图像质量参差不齐的影响。3DSL从行人图像中提取形状信息进行3D重建，并使用姿态预测和前景分割生成姿态信息和轮廓信息作为重建过程的监督，从而优化行人形状信息的提取过程。GI-ReID则从行人特征中提取步态相关的衣物无关信息，并使用一个步态预测网络监督提取到的步态相关信息，从而使网络学习到更多衣物无关特征。UCAD使用人体解析方法将行人的衣物部分单独分离出来，使用一个额外的特征提取网络从分离出的衣物部分中提取衣物特征，并将衣物特征蒸馏至提取行人特征的网络中，在行人特征的提取网络中使用衣物特征和行人特征的正交损失抑制行人特征中衣物信息的含量。

与这些方法的不同之处在于，本文使用素描图像作为辅助信息的来源，素描图像的鲁棒性使得辅

助信息的质量有所保障。此外对于辅助信息的使用也采用了较为简洁但有效的方式，将素描图像的辅助信息转化为行人图像在位置上的权重，显式地指导行人图像特征的提取过程，并进一步从素描图像中提取体型特征以补足表观特征，对辅助信息进行了充分使用。

表 1　与主流方法在 LTCC 和 PRCC 数据集上的结果对比

Table 1　Comparison with state-of-the-art methods on LTCC dataset and PRCC dataset

/%

方法	LTCC		PRCC	
	Rank-1	mAP	Rank-1	mAP
PCB	23.50	10.00	41.80	38.70
IANet	25.00	12.60	46.30	45.90
CASE-Net	–	–	39.51	–
ADC	–	–	51.60	56.20
CESD	26.15	–	–	–
GI-ReID	28.86	14.10	37.55	–
3DSL	31.20	14.80	51.30	–
UCAD	32.50	15.10	51.30	–
Pos-Neg	36.22	14.43	54.87	**65.77**
CAL	**40.1**0	**18.0**0	<u>55.20</u>	<u>55.80</u>
本文	<u>38.00</u>	<u>15.90</u>	**55.5**0	52.60

注：加粗字体表示各列最优结果，加下划线字体表示次优结果，“–”表示原文未提供数据。

为了进一步说明本文方法的有效性，与传统方法、特征解耦合方法以及数据增广方法进行比较。在传统方法中，本文沿用 Gu 等人（2022）的设定，选取 Sun 等人（2018）提出的 PCB（part-based convolutional baseline）和 Hou 等人（2019）提出的 IANet（interaction-and-aggregation network）进行比较。与传统方法相比，本文方法在 LTCC 和 PRCC 数据集上的性能都有明显优势。在解耦合的这一类方法中，Li 等人（2021）提出的 CASE-Net（clothing agnostic shape extraction network）使用两个编码器分别提取行人图像的形状特征和颜色特征，通过使用灰度图像的形状特征和彩色图像的颜色特征进行图像重建的方式对形状和颜色特征的生成过程进行监督。Yang 等人（2022）在以往基于编码器的解耦方法的基础上，提出了 ADC 方法，通过竞争注意力使模型能够在已经关注到的区域外持续学习其他区域的特征，从而获得衣服特征以外的身份相关特征。结果表明，本文在 PRCC 数据集上的性能相对于 CASE-Net 和 ADC 有比较明显的优势。Jia 等人（2022）提出一种数据增广方法 Pos-neg（positive and negative augmentations），借助人体解析模型的辅助，直接交换两幅图像中均为衣服区域的矩形区域，从而在增广出正样本的同时保留图像的身份信息。结果表明，本文方法在 LTCC 和 PRCC 数据集上的性能均优于 Pos-neg。Gu 等人（2022）提出的 CAL（clothes-based adversarial loss）使用对抗学习的思路，在特征提取网络中加入一个衣物分类器和衣物对抗学习损失，约束网络对身穿不同衣服的同一个人预测出相同的标签，从而使网络在学习到行人特征的同时尽可能消除衣物特征的干扰。虽然 CAL 在 LTCC 数据集上的性能优于本文方法，但 CAL 需要使用额外的衣物标签进行训练，这同样需要一定的标注成本。

3.4　轮廓图像与素描图像的消融实验

为验证本文提出的素描图像相较于轮廓图像更鲁棒、更适于作为换装行人重识别任务中行人体型信息的获取方式，分别进行轮廓图像和素描图像单独使用、与行人图像共同使用的消融实验。使用 SCHP（self-correction for human parsing）（Li 等，2019）生成 LTCC 和 PRCC 数据集中行人的轮廓图像，使用 pidinet（Su 等，2021）生成 LTCC 数据集中行人的素描图像。PRCC 数据集中包含行人的素描图像，可直接使用。

3.4.1　单独使用轮廓图像和素描图像

为了验证素描图像优于轮廓图像，本文首先对单独使用两种图像的情况进行对比实验。实验模型采用 ResNet-50 网络结构，分别在 LTCC 和 PRCC 两个数据集上进行实验，损失函数采用交叉熵损失和难采样三元组损失（后续消融实验若不作特殊说明，均为此设置）。对轮廓图像进行实验时，给定一个行人重识别数据集 $\boldsymbol{D}_r=\{(\boldsymbol{I}_i^r,y_i)\}_{i=1}^n$，其中 n 表示数据集中的图像数量，$\boldsymbol{I}_i^r$ 表示第 i 幅图像（上标用以标识图像种类，r 表示行人原图像），y_i 表示第 i 幅图像的标签，每一幅图像都使用人体解析模型生成对应的轮廓图像，组成新的数据集 $\boldsymbol{D}_c=\{(\boldsymbol{I}_i^c,y_i)\}_{i=1}^n$，其中 c 表示轮廓图像。对素描图像进行实验时，使用边缘检测模型可以得到新的素描图像数据集 $\boldsymbol{D}_s=$

$\{(\boldsymbol{I}_i^s, y_i)\}_{i=1}^n$，$s$表示素描图像。使用$\boldsymbol{D}_c$和$\boldsymbol{D}_s$进行实验的结果如表2所示。由表2可知，相较于轮廓图像，素描图像在各项指标上均有明显的性能提升。

表2的实验结果验证了素描图像优于轮廓图像。原因在于，相较于轮廓图像，素描图像能提供更鲁棒、更准确的行人体型信息。在鲁棒性方面，当行人图像出现遮挡或光照条件较差时，素描图像相较于轮廓图像受这些因素影响更小，行人体型更完整，如图4前两组图像所示（图4(a)(b)(c)与图中相同位置的3幅图为一组）；在准确性方面，素描图像的行人身体边界与行人图像更加吻合，行人体型更加准确，如图4后两组图像所示。图4中矩形框标记的区域为轮廓图像中有缺陷的区域。

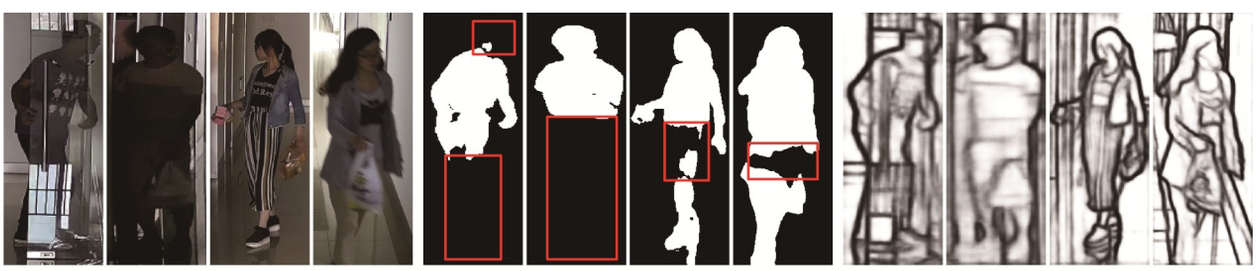
表2　LTCC和PRCC数据集上轮廓图像和素描图像的评估结果对比

Table 2　Comparison of evaluation results between contour images and sketch images on LTCC dataset and PRCC dataset

/%

图像类别	LTCC		PRCC	
	Rank-1	mAP	Rank-1	mAP
轮廓图像	11.2	4.2	35.9	28.3
素描图像	**15.6**	**6.4**	**39.5**	**32.5**

注：加粗字体表示各列最优结果。



(a) 原始图像　(b) 轮廓图像　(c) 素描图像

图4　低质量轮廓图像及其对应的行人、素描图像

Fig. 4　Low quality contour images and their corresponding person and sketch images ((a) origin images; (b) contour images; (c) sketch images)

3.4.2　轮廓图像和素描图像结合行人图像

除对单独使用轮廓图像和素描图像情况进行实验外，进行了轮廓图像和素描图像分别与行人图像共同使用的对比实验。实验模型结构为两个并行的ResNet-50网络，一个对行人图像进行特征提取，另一个对轮廓图像或素描图像特征进行提取，从两个网络中提取的特征以最佳比例进行融合。两个子网络分别在行人图像数据集和轮廓/素描图像数据集上进行训练，实验结果如表3所示。

由表3的实验结果对比可知，在各项评估指标上，素描图像都比轮廓图像有性能上的优势。原因在于行人表观特征中体型信息不足，而相较于轮廓图像，素描图像能提供更鲁棒、更准确的体型信息给表观特征作为补充，从而获得更完备的行人特征。在给表观特征提供体型信息作为补充的方面，素描图像也有明显的优势。实验结果表明，素描图像能比轮廓图像给行人图像提供更完整的互补信息。

表3　LTCC数据集和PRCC数据集上轮廓图像和素描图像结合行人图像的评估结果

Table 3　Evaluation results of contour images and sketch images combined with person images on LTCC dataset and PRCC dataset

/%

图像类别	LTCC		PRCC	
	Rank1	mAP	Rank-1	mAP
行人图像	35.2	15.7	47.2	46.6
行人图像+轮廓图像	37.0	15.6	51.9	46.1
行人图像+素描图像	**37.5**	**16.1**	**53.5**	**51.2**

注：加粗字体表示各列最优结果。

3.5　衣物无关权重指导模块的消融实验

为探究衣物无关权重模块的最佳位置，同时验证本文所提出的基于素描图像的衣物无关权重指导模块的有效性，对模块的位置和模块中平均池化的参数进行消融实验，结果如表4所示。在表4中，模块位置一列的Layer n表示将衣物无关权重指导模

块放置在行人图像特征提取分支的第 n 个阶段后，尺寸列表示衣物无关权重指导模块平均池化降采样时使用的核尺寸(kernel size)，在本文中取值范围为{2, 3, 5, 7}，其对应的步长分别为2,2,3,4。为了减少冗余，提高表格的可读性，将每一个模块位置对应的平均池化降采样的最佳参数直接列在模型性能之后。

由表4可知，在LTCC数据集上，衣物无关权重模块最佳位置为网络的第3层后，池化层的kernel size为5；在PRCC数据集上，衣物无关权重模块最佳位置为网络的第1层后，池化层的kernel size为7。此外，在加入了衣物无关权重指导模块后，在PRCC数据集上的评估结果对比没有加入时均有提升，在LTCC数据集上后两层的结果有提升，前两层没有变化，从而证明了其有效性。

在衣物无关权重指导模块中，池化层的kernel size用于控制降采样后边缘点周围的衣物无关权重分布，kernel size越小，每个位置的衣物无关权重的关联区域越小。在卷积神经网络中，网络的底层学习到的是图像的低级语义特征，且其感受野较小，此时应当设置一个较小的kernel size以减少网络对于衣物纹理和颜色信息的学习；随着网络的深入，网络能够学习到图像的高层语义特征，且感受野逐渐增加，此时应当增大kernel size使衣物无关权重的关联区域符合感受野的大小。另外，高层语义特征与最后的身份特征直接相关，因此在网络高层放置指导模块优于放置在底层。将指导模块放置于底层，其提供的权重信息在经过较多层的变换后会与原先信息产生较大出入。

表4　衣物无关权重指导模块放置于不同位置的评估结果对比

Table 4　Comparison of evaluation results of clothes-irrelevant weight guidance module at different locations

/%

模块位置	LTCC		PRCC	
	Rank-1	尺寸	Rank-1	尺寸
无	35.2	–	47.2	–
Layer1	35.2	3	**54.0**	7
Layer2	35.2	5	50.1	2
Layer3	**36.7**	5	51.3	5
Layer4	36.2	7	50.4	5

注：加粗字体表示各列最优结果，“–”表示没有该参数。

而在PRCC数据集中，绝大多数行人的姿势是正面朝向摄像头的，且脸部较为清晰，因此在网络的底层放置衣物无关权重模块，并设置一个较大的kernel size，使网络最大程度保留脸部信息，有利于换装场景中的识别。

3.6　表观特征与体型特征融合参数的消融实验

为了探究表观特征与体型特征按比例相加的最佳比例参数，本文对融合时使用的参数 w_1 和 w_2 进行实验，w_1 和 w_2 分别代表表观特征和体型特征的权重参数，$w_2 = 0$ 表示只使用表观特征，$w_1 = 0$ 表示只使用体型特征，实验结果如表5所示。由表5可知，在LTCC数据集上，当 $w_1/w_2 = 4$ 时，模型的性能达到最优；在PRCC数据集上，当 $w_1/w_2 = 2$ 时，模型的性能达到最优。融合时使用的权重参数的差异是数据集本身的差异导致的，LTCC数据集更具有挑战性，其素描图像的识别性能相对较低，因此在特征融合时其所占比例也要适当降低，才能在不影响表观特征的前提下补充体型特征，获取最佳行人特征。

表5　不同 w_1/w_2 的评估结果对比

Table 5　Comparison of evaluation results between different w_1/w_2 ratio

/%

w_1/w_2	LTCC		PRCC	
	Rank-1	mAP	Rank-1	mAP
1.0	31.9	13.1	54.4	47.8
2.0	37.2	15.6	**55.5**	52.6
3.0	37.2	**15.7**	55.0	**52.7**
4.0	**38.0**	15.6	54.8	52.6
5.0	36.7	15.4	54.5	52.5
0.5	21.4	9.1	46.5	36.3
0.33	18.1	7.9	41.5	32.1
0.25	16.6	7.5	39.4	30.4
0.2	16.6	7.2	39.4	30.5
$w_1 = 0$	15.6	6.4	39.5	32.5
$w_2 = 0$	35.2	15.7	49.0	49.1

注：加粗字体表示各列最优结果。

3.7　检索结果可视化分析

为了直观地展示本文提出的基于素描图像的行人完备特征获取方法在换装场景下的有效性，在LTCC数据集上对基线方法和本文所提出的方法进

行可视化对比。基线方法使用的骨架网络为ResNet-50，在最后的全局池化层和分类层间加入一个BN(batch normalization)层，使用行人图像数据集进行训练，损失函数为交叉熵损失和难采样三元组损失，训练策略与本文提出的方法相同。可视化结果如图5所示。

图5(a)是查询图像，图5(b)(c)分别是基线方法检索结果和本文方法检索结果的前5幅图像，红色框表示此检索结果与查询结果不是同一个行人；蓝色框表示此检索结果与查询结果是同一个行人。

从基线方法与本文方法检索结果对比可以看出，基线方法的检索结果受衣物特征影响很大，使用基线方法检索得到的前5个结果大多为穿着相似的不同行人。本文方法则能够有效使用行人的体型信息获取到完备的行人特征，减少穿着相似衣物的不同行人的干扰，从而准确地检索到穿着不同衣物的相同行人。

(a) 查询图像　(b) 基线方法的检索结果　(c) 本文方法的检索结果

图5　换装场景下重识别结果可视化对比

Fig. 5　Visual comparison of re-identification results in cloth-change setting((a)query images;(b)baseline method;(c)ours)

3.8　激活图可视化分析

为了直观地说明从素描图像中提取的体型特征能为表观特征提供互补信息，增强特征的辨识力，在LTCC数据集上分别单独使用行人图像集和素描图像集作为数据集训练并测试基线方法，并将二者在行人图像上的激活图进行可视化对比，结果如图6所示。

在图6中，左侧图像是使用行人图像训练得到的模型的激活图，可以看出，模型主要关注行人的衣物部分，如上半身衣物、鞋子等，这些衣物区域的特征在换装场景下不具有辨识力。右侧图像是使用素描图像训练得到的模型的激活图(为了便于观察和对比，将激活图放置于行人图像上)，可以看出，模型的关注区域分布在行人的全身各处，从这些区域中可以获取衣物无关的行人体型信息。此外，模型还

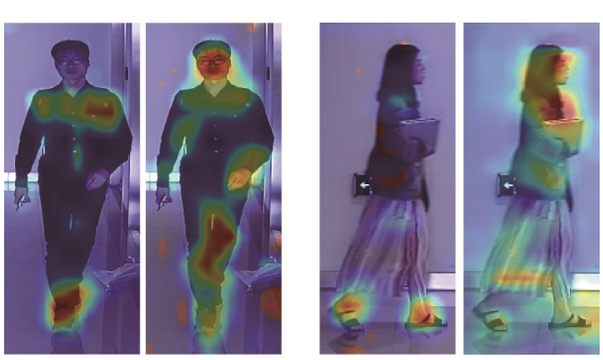


图6　行人图像和素描图像的激活图可视化对比

Fig. 6　Visual comparison of the activation map of two models

重点关注到行人的头部区域，头部区域的形状特征也是在换装场景下判别行人身份的重要依据之一。由左侧和右侧的图像对比可知，素描图像能够使模型关注到衣物以外的区域，降低特征中衣物信息含量，从而获得更有辨识力的特征。

4 结 论

行人更换服装对行人重识别模型的准确度有较大的影响，是行人重识别实际应用亟待解决的问题。为此，本文提出一种素描图像指导的换装行人重识别方法，针对轮廓图像鲁棒性和准确性不足的问题，使用素描图像提取行人的体型特征，并使用双流网络将体型特征融入表观特征以获取完备的行人特征；针对表观特征中包含大量不具有辨别力的衣物特征的问题，提出一个基于素描图像的衣物无关权重指导模块，进一步使用素描图像中行人衣物位置信息指导行人表观特征的提取过程，从而降低表观特征中衣物信息的含量，提高特征的判别力。在两个数据集上的实验结果表明，本文方法能提取到更准确、鲁棒的行人体型信息，通过衣物无关矩阵的指导降低表观特征中衣物信息的含量，提高了模型对不同着装行人的判别力。

但是，本文方法仍然存在一些问题值得进一步研究。例如，本文提出的衣物无关指导模块需要对特定数据集进行放置位置和池化层的参数设定，而参数选择对模型的训练具有较大影响，不同参数的模型性能有很大差距。另外，本文方法没有使用换装数据集中衣物的标签。未来将在不同模态信息对表观特征的指导和衣物标签的使用两个方面进行进一步研究。

参考文献(References)

Bhattarai M, Oyen D, Castorena J, Yang L P and Wohlberg B. 2020. Diagram image retrieval using sketch-based deep learning and transfer learning//Proceedings of 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle, USA: IEEE: 174-175 [DOI: 10.1109/CVPRW50498.2020.00095]

Chen J X, Jiang X Y, Wang F D, Zhang J, Zheng F, Sun X and Zheng W S. 2021. Learning 3D shape feature for texture-insensitive person re-identification//Proceedings of 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Nashville, USA: IEEE: 8142-8151 [DOI: 10.1109/CVPR46437.2021.00805]

Deng J, Dong W, Socher R, Li L J, Li K and Li F F. 2009. ImageNet: a large-scale hierarchical image database//Proceedings of 2009 IEEE Conference on Computer Vision and Pattern Recognition. Miami, USA: IEEE: 248-255 [DOI: 10.1109/CVPR. 2009. 5206848]

Eitz M, Hays J and Alexa M. 2012. How do humans sketch objects? ACM Transactions on Graphics, 31 (4): #44 [DOI: 10.1145/2185520.2185540]

Geirhos R, Rubisch P, Michaelis C, Bethge M, Wichmann F A and Brendel W. 2018. ImageNet-trained CNNs are biased towards texture; increasing shape bias improves accuracy and robustness//Proceedings of the 7th International Conference on Learning Representations (ICLR). New Orleans, USA: OpenReview. net.: 1-10 [DOI: 10.48550/arXiv.1811.12231]

Gu X Q, Chang H, Ma B P, Bai S T, Shan S G and Chen X L. 2022. Clothes-changing person re-identification with RGB modality only//Proceedings of 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition. New Orleans, USA: IEEE: 1060-1069 [DOI: 10.1109/CVPR52688.2022.00113]

Hertzmann A. 2020. Why do line drawings work? A realism hypothesis. Perception, 49 (4): 439-451 [DOI: 10.1177/0301006620908207]

Hertzmann A. 2021. The role of edges in line drawing perception. Perception, 50(3): 266-275 [DOI: 10.1177/0301006621994407]

Hong P X, Wu T, Wu A C, Han X T and Zheng W S. 2021. Fine-grained shape-appearance mutual learning for cloth-changing person re-identification//Proceedings of 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Nashville, USA: IEEE: 10508-10517 [DOI: 10.1109/CVPR46437.2021.01037]

Hou R B, Ma B P, Chang H, Gu X Q, Shan S G and Chen X L. 2019. Interaction-and-aggregation network for person re-identification//Proceedings of 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach, USA: IEEE: 9309-9318 [DOI: 10.1109/CVPR.2019.00954]

Jia X M, Zhong X, Ye M, Liu W X and Huang W X. 2022. Complementary data augmentation for cloth-changing person re-identification. IEEE Transactions on Image Processing, 31: 4227-4239 [DOI: 10.1109/TIP.2022.3183469]

Jin X, He T Y, Zheng K C, Yin Z H, Xu S, Zhen H, Feng R Y, Huang J Q, Chen Z B, Hua X S. 2022. Cloth-changing person re-identification from a single image with gait prediction and regularization//Proceedings of 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition. New Orleans, USA: IEEE: 14278-14287 [DOI: 10. 1109/CVPR52688. 2022. 01388]

Li P K, Xu Y Q, Wei Y C and Yang Y. 2019. Self-correction for human parsing. IEEE Transactions on Pattern Analysis and Machine Intelligence, 44 (6): 3260-3271 [DOI: 10.1109/TPAMI. 2020. 3048039]

Li Y J, Weng X S and Kitani K M. 2021. Learning shape representations for person re-identification under clothing change//Proceedings of 2021 IEEE/CVF Winter Conference on Applications of Computer Vision. Waikoloa, USA: IEEE: 2431-2440 [DOI: 10.1109/WACV48630.2021.00248]

Lorenz D, Bereska L, Milbich T and Ommer B. 2019. Unsupervised

part-based disentangling of object shape and appearance//Proceedings of 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach, USA: IEEE: 10947-10956 [DOI: 10.1109/CVPR.2019.01121]

Qian X L, Wang W X, Zhang L, Zhu F R, Fu Y W, Xiang T, Jiang Y G and Xue X Y. 2020. Long-term cloth-changing person re-identification//Proceedings of the 15th Asian Conference on Computer Vision. Kyoto, Japan: Springer: 71-88 [DOI: 10.1007/978-3-030-69535-4_5]

Shi W D, Zhang Y Z, Liu S W, Zhu S D and Pu J N. 2020. Person re-identification based on deformation and occlusion mechanisms. Journal of Image and Graphics, 25(12): 2530-2540 (史维东, 张云洲, 刘双伟, 朱尚栋, 暴吉宁. 2020. 针对形变与遮挡问题的行人再识别. 中国图象图形学报, 25(12): 2530-2540) [DOI: 10.11834/jig.200016]

Su Z, Liu W Z, Yu Z T, Hu D W, Liao Q, Tian Q, Pietikäinen M and Liu L. 2021. Pixel difference networks for efficient edge detection//Proceedings of 2021 IEEE/CVF International Conference on Computer Vision. Montreal, Canada: IEEE: 5117-5127 [DOI: 10.1109/ICCV48922.2021.00507]

Sun Y F, Zheng L, Yang Y, Tian Q and Wang S J. 2018. Beyond part models: person retrieval with refined part pooling (and a strong convolutional baseline)//Proceedings of the 15th European Conference on Computer Vision. Munich, Germany: Springer: 501-518 [DOI: 10.1007/978-3-030-01225-0_30]

Wang X N, Liu C H, Qi G Q and Zhang S Q. 2020. Person re-identification based on top-view depth head and shoulder sequence. Journal of Image and Graphics, 25(7): 1393-1407 (王新年, 刘春华, 齐国清, 张世强. 2020. 俯视深度头肩序列行人再识别. 中国图象图形学报, 25(7): 1393-1407) [DOI: 10.11834/jig.190608]

Wu A C, Lin C Z and Zheng W S. 2022. Single-modality self-supervised information mining for cross-modality person re-identification. Journal of Image and Graphics, 27(10): 2843-2859 (吴岸聪, 林城栝, 郑伟诗. 2022. 面向跨模态行人重识别的单模态自监督信息挖掘. 中国图象图形学报, 27(10): 2843-2859) [DOI: 10.11834/jig.211050]

Yan Y M, Yu H M, Li S Z, Lu Z H, He J F, Zhang H Z and Wang R F. 2022. Weakening the influence of clothing: universal clothing attribute disentanglement for person re-identification//Proceedings of the 31st International Joint Conference on Artificial Intelligence. Vienna, Austria: Morgan Kaufmann: 1523-1529 [DOI: 10.24963/ijcai.2022/212]

Yang Q Z, Wu A C and Zheng W S. 2019. Person re-identification by contour sketch under moderate clothing change. IEEE Transactions on Pattern Analysis and Machine Intelligence, 43(6): 2029-2046 [DOI: 10.1109/TPAMI.2019.2960509]

Yang Z W, Zhong X, Liu H, Zhong Z and Wang Z. 2022. Attentive decoupling network for cloth-changing re-identification//Proceedings of 2022 IEEE International Conference on Multimedia and Expo. Taipei, China: IEEE: 1-6 [DOI: 10.1109/ICME52920.2022.9859851]

Ye M, Shen J B, Lin G J, Xiang T, Shao L and Hoi S C H. 2022. Deep learning for person re-identification: a survey and outlook. IEEE Transactions on Pattern Analysis and Machine Intelligence, 44(6): 2872-2893 [DOI: 10.1109/TPAMI.2021.3054775]

Yu Q, Liu F, Song Y Z, Xiang T, Hospedales T M and Loy C C. 2016. Sketch me that shoe//Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas, USA: IEEE: 799-807 [DOI: 10.1109/CVPR.2016.93]

Yu Z X, Zhao Y L, Hong B, Jin Z M, Huang J Q, Cai D, He X F and Hua X S. 2020b. Apparel-invariant feature learning for apparel-changed person re-identification[EB/OL]. [2020-08-17]. https://arxiv.org/pdf/2008.06181.pdf

Zhang X, Luo H, Fan X, Xiang W L, Sun Y X, Xiao Q Q, Jiang W, Zhang C and Sun J. 2018. AlignedReID: surpassing human-level performance in person re-identification [EB/OL]. [2018-01-31]. https://arxiv.org/pdf/1711.08184.pdf

Zheng Z D, Yang X D, Yu Z D, Zheng L, Yang Y and Kautz J. 2019. Joint discriminative and generative learning for person re-identification//Proceedings of 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach, USA: IEEE: 2138-2147 [DOI: 10.1109/CVPR.2019.00224]

作者简介

刘宇奇,男,硕士研究生,主要研究方向为行人重识别。
E-mail: liuyuqi201@mails.ucas.ac.cn
马丙鹏,通信作者,男,教授,主要研究方向为计算机视觉和模式识别。E-mail: bpma@ucas.ac.cn