

DOI:10.16136/j.joel.2022.09.0873

基于注意力机制与局部关联特征的行人重识别

张晓涵*

(中国石油大学(华东) 计算机科学与技术学院, 山东 青岛 266580)

摘要:由于行人在真实场景下易受到背景、遮挡、姿态等问题的影响,为获取行人图像中更具辨别能力的特征,提出一种基于注意力机制和局部关联特征的行人重识别方法。首先,在网络框架中嵌入注意力模块以关注图像中表达能力强的特征;然后,利用图像中相邻区域的关联得到局部关联特征,并结合全局特征。本文方法在 Market1501 和 DukeMTMC-ReID 数据集上进行实验, Rank-1 指标分别达到了 95.3% 和 90.1%。结果证明,本文方法能充分获取判别力强的特征信息,使模型具有较强的识别能力。

关键词:行人重识别;注意力机制;局部关联特征;多粒度

中图分类号: TP391.41 **文献标识码:** A **文章编号:** 1005-0086(2022)09-0984-08

Person re-identification based on attention mechanism and local association feature

ZHANG Xiaohan*

(College of Computer Science and Technology, China University of Petroleum (East China), Qingdao, Shandong 266580, China)

Abstract: Because pedestrians are easily affected by background, occlusion, posture and other issues in real scenes, in order to obtain more discriminative features in pedestrian images, a person re-identification method based on attention mechanism and local association feature is proposed. Firstly, the attention module is embedded in the network framework to pay attention to the features with strong expressive ability in the image. Then, the local association features are obtained by using the association of adjacent parts in the image, and combined with the global features. The experiments on Market1501 and DukeMTMC-ReID datasets show that the Rank-1 index reaches 95.3% and 90.1%, respectively. The results show that the proposed method can fully obtain the feature information with strong discrimination and make the model have strong recognition ability.

Key words: person re-identification; attention mechanism; local association feature; multi-granularity

1 引言

近年来,公共安全问题引起人们越来越多的关注,加强对行人身份的识别尤为重要。行人重识别旨在给定一张目标行人图像,检索出不同摄像头下的该目标行人,该技术可与人脸识别、行人检测等技术相结合,满足当前智能安防的需要。行人重识别技术具有广阔的应用前景,除了智能安防领域的嫌疑犯追踪、走失救助,还有智慧商业领域的无人超市、人机交互等。但由于不同摄像

头拍摄的行人图像其光照、姿态、分辨率等属性差异巨大以及遮挡问题的干扰,使得该技术在实际应用场景中面临诸多挑战^[1]。因此,如何提取更具判别力的视觉特征成为行人重识别问题的研究热点。

传统方法通过手工设计 SIFT (scale invariant feature transform)、HOG (histogram of oriented gradient) 等颜色或纹理特征,效率低下且精度不高。随着深度学习的发展,用于提取深度特征的卷积神经网络被应用于行人重识别领域。由于在

* E-mail: 1241021052@qq.com

收稿日期: 2021-12-24 修订日期: 2021-01-28

实际场景中,受背景冗余、摄像头视角、光照、行人姿态、遮挡等客观因素的影响,导致依赖图像整体信息的全局特征不足以对行人进行精确匹配。所以,当前行人重识别研究通常结合局部特征对行人进行精细化表达,弥补全局特征的不足。为获取局部信息,一些研究利用人体姿态信息,如SpindleNet^[2]根据人体关键点产生7个语义区域(头、四肢等)。此类方法需要额外的标注信息,且容易引入姿态估计误差影响行人重识别的结果。为解决此问题,一些研究提出将特征图在空间维度上均匀分割为多个区域,对各区域的特征进行匹配。SUN等^[3]提出基于部件的PCB模型,将行人特征图水平平均分为6个部件,但该方法会造成各个部件无法有效地对齐,需要后续的精确部件池化来解决,无法实现端到端。并且该方法忽视了全局特征以及相邻部件之间的相关性,导致丢失部分重要信息。之后,ZHENG等^[4]提出了由粗到细的金字塔模型,该模型整合了不同尺度且包含重叠信息的局部特征和全局特征,能够缓解行人部件无法对齐的问题,并充分利用了相邻部件的关联信息与边缘信息。然而,该模型包含21个特征,一些特征对模型性能的提升影响较小,这大大增加了模型复杂度与计算量。

针对以上问题,本文提出一种基于注意力机制与局部关联特征的行人重识别方法。该方法不需要引入额外的姿态估计模型,有效实现相邻局部部件的相关性学习,同时保证计算量较小。由于轻量级注意力模块可以增强对图像中重要区域的关注度从而提升特征表达能力,并且可以无缝嵌入到卷积神经网络中。为获取行人图像中的关键信息用于后续的特征处理,本文在特征提取网络中嵌入注意力模块。鉴于空间分割获取局部特征的简单有效性,本文同样采用该方法并进行改进。为增强特征判别性与鲁棒性,同时减轻行人部件语义信息不对齐的影响,本文利用相邻部件之间的语义关联,提出局部关联特征。通过结合不同粒度的局部关联特征与全局特征,对行人图像各区域的信息充分挖掘与利用。联合分类损失和难样本采样三元组损失实现端到端训练。在Market1501数据集与DukeMTMC-reID数据集上进行实验,本文方法取得了较高的准确率。

2 相关工作

2.1 注意力机制

注意力机制被广泛应用于图像识别、图像分割等计算机视觉任务中。其基本思想是对输入图像的各个部分赋予不同权值,从而抽取出对当前任务更

有意义的特征。HU等^[5]提出一种通道域注意力机制SE-Net,通过挤压-激励操作使模型可以学习到不同特征通道的重要程度,依据重要度提升有价值的特征。WOO等^[6]提出了卷积注意力模块(convolutional block attention module,CBAM),将注意力应用到通道和空间两个维度上,取得了比仅应用通道注意力更好的效果。近两年,基于自注意力机制的Transformer在各大任务中取得了重大突破,其中HE等^[7]提出TransReID用于行人重识别,取得了与卷积神经网络相当的性能。本文在骨干网络中嵌入注意力模块,旨在获取对模型性能影响较大的特征描述子。

2.2 特征表示

近年来,深度学习在行人重识别中取得了一系列重大进展。ZHENG等^[8]把行人重识别看作一个分类问题,提出了基于ResNet50的IDE模型。基于局部特征的方法能大大增强对行人身体关键区域的学习能力,成为近年来的研究趋势。WEI等^[9]利用姿态估计算法估计人体的关键点,有效获取人体的头部、上半身、下半身区域特征,然而该方法很大程度上依赖于姿态估计模型的性能,且增加了运算成本。LI等^[10]将空间注意力与通道注意力相结合引入到行人重识别中,联合学习注意力选择和特征表示。WANG等^[11]设计了一种融合行人全局特征和局部特征的多粒度网络MGN,相比只采用局部特征的PCB模型^[3],效果显著提升。2020年,PARK等^[12]提出了一种关系网络,每个部件特征融合了相应部件本身与剩余部件特征,使得局部特征更具判别力。在以上工作的基础上,本文设计包含全局分支与局部分支的网络结构,卷积网络提取特征后,全局分支获取来自行人的整体信息。局部分支中,首先将特征图在水平方向均分成6个部件,将每个部件与其相邻部件组合再进行空间切割组成局部关联特征,该特征包含了行人上下文信息和重叠信息。

3 本文方法

本文在卷积层之间嵌入注意力模块以有效抑制行人图像中的无关噪声;设计新型的局部关联特征以充分利用人体空间结构的相关性与边缘信息,缓解人体部件语义信息不对齐的问题。本文方法网络结构如图1所示。

本文方法整体框架包含3部分:嵌入CBAM的残差网络、全局和局部分支。首先使用ImageNet预训练过的ResNet50^[13]提取行人特征,移除最后的平均池化层与全连接层,并将最后一个卷积层的步长设为1,在conv3_x与conv4_x之间插入CBAM^[6],

将提取的特征图分别经过全局分支与局部分支,得到不同粒度的行人特征。采用分类损失与难样本采

样三元组损失训练网络。测试阶段,将全局特征与局部关联特征融合作为行人检索特征。

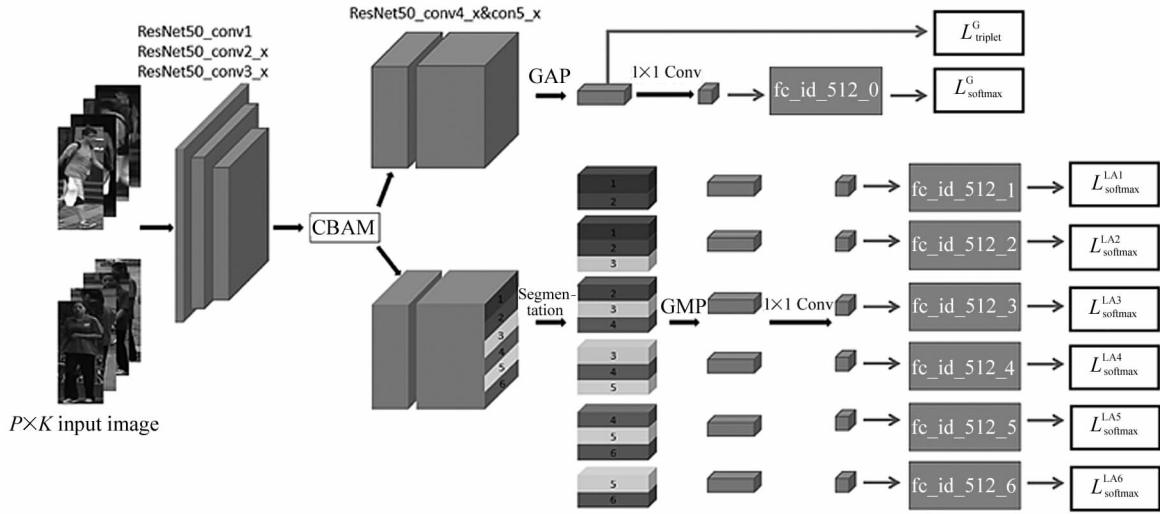


图1 本文网络结构

Fig. 1 The network structure of this article

3.1 ResNet_CBAM

本文选用在行人重识别领域被广泛使用的 ResNet50 作为骨干网络。为了使网络关注图像中的关键区域减轻背景等噪声区域的负面影响,在 ResNet50 中嵌入通用的轻量级的 CBAM^[6]。如图 1 所示,网络在经过 conv3_x 层提取特征图后,将特征图送入 CBAM。CBAM 包含通道注意力模块和空间注意力模块,从两个维度对特征图进行特征加权。如图 2 所示,特征图 F 首先经过通道注意力模块,为了提高网络的表示能力,同时应用最大池化与平均池化进行空间维度压缩,得到两个一维的通道特征描述。然后将之分别输入权重共享的多层感知机 (multi-layer perceptron, MLP),将输出的两个特征

逐元素相加后通过 sigmoid 激活操作得到通道注意力权重。将权重系数与输入特征图逐元素相乘即可得到通道注意力特征图 F^1 。将通道注意力模块的输出作为空间注意力模块的输入,应用最大池化与平均池化进行通道维度压缩,将生成的两个二维的空间特征进行拼接。然后经过卷积操作和 sigmoid 激活操作,将得到的空间注意力权重与输入特征图逐元素相乘即可得到空间注意力特征图 F^2 。CBAM 的两个子模块互补计算,从通道轴与空间轴关注了“什么”的特征与“哪里”的特征有意义。在本文方法中,将 CBAM 插入到卷积层之间,有效提取到对行人重识别任务更有用的行人特征,并在后续的多分支网络中得以利用。

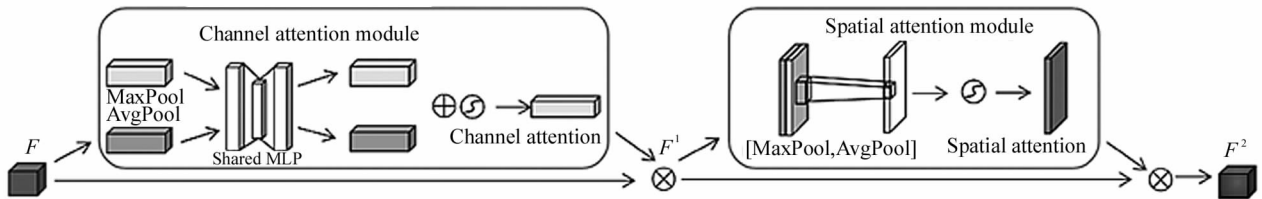


图2 卷积注意力模块示意图

Fig. 2 Schematic diagram of convolutional block attention module

3.2 多粒度特征联合

如图 1 所示,输入行人图像使用 ResNet_CBAM 提取到特征图,接下来分为全局与局部分支。全局

分支中,特征图经过全局平均池化(global average pooling,GAP)得到 2048 维的特征向量,随后进行降维操作降至 512 维。由于大型场景下,行人的外

观受到各种客观因素的影响,使用粗粒度的全局特征虽然能捕获行人外观最显著的特征来进行区分,但也会忽视一些不显著的细节信息,很容易把外观相似的不同行人误判成同一个人,这大大降低了行人重识别的准确率。因此,本文方法在充分利用行人全局信息的基础上,采用水平切块的方法获取行人的细粒度局部特征。PCB方法^[3]将特征图水平均分,采用6个无交叠局部区域的特征单独预测行人身份,由于划分的尺寸较小且各个区域没有关联,很容易造成行人局部区域语义不对齐。如图3所示,采用PCB方法,同一行人的两张图像经过水平均等切分后,最上层的部件分别表示行人头部信息与背景信息,直接进行分类预测将影响行人重识别性能。



(a)



(b)

图3 切分效果图:(a) PCB方法;(b) 本文方法

Fig. 3 Segmentation effect diagram: (a) PCB method; (b) Method of this article

如图3所示,采用局部关联特征,两张空间语义不对齐的图像在最上层的部件中均包含了行人头部信息,可以有效减轻行人部件不对齐造成误判的影响。本文方法获取的行人特征由粗粒度全局特征与

为解决此问题,本文在PCB的基础上,考虑了相邻部件之间的相关性。如图1所示,在局部分支中,将特征图水平划分为6个部件,需要注意的是,此时并不将各部件进行切割,而是将每个部件与其相邻部件组合,由此得到了部件12、123、234、345、456、56合并的部件,之后再进行切割。将6个合并后的部件分别经过最大池化层(global max pooling, GMP)得到2 048维的局部关联特征向量,之后降低维度至512维。本文方法获取的局部特征包含了重叠区域,各个区域的特征能得到多次充分利用,能有效学习到行人相邻区域的相关性。此外,本文方法切割的部件相比PCB尺寸更大,所以不同图像对应部件之间包含了更多的相似信息。

不同尺寸的细粒度局部特征组成,包含行人的整体信息与非显著性细节信息。

3.3 损失函数

为了提升网络学习特征表示的能力,本文实验

采用交叉熵损失与难样本采样三元组损失联合训练。采用交叉熵损失进行分类,采用难样本采样三元组损失进行度量学习。这两种损失都被经常应用于行人重识别网络中。

首先把行人重识别看作一个分类问题,即不同的行人属于不同的类别。利用行人的 ID 作为标签来训练网络,分类类别数即为训练集中行人 ID 数。对于第 i 个学习到的特征 f_i ,交叉熵损失表示如下:

$$L_{\text{softmax}} = - \sum_{i=1}^N \log \frac{e^{w_{yi}^T f_i}}{\sum_{k=1}^C e^{w_k^T f_i}}, \quad (1)$$

式中, N 表示实验批次数量, w_k 表示对应类别的权重, C 表示训练集中类别数。在本文实验中,对 512 维的全局特征和局部特征应用交叉熵损失函数。通过优化分类模型学习行人特征,预测行人标签。

度量学习旨在学习两张图像的相似性。三元组损失作为最常见的度量学习函数,可以有效用于行人重识别任务中,实现拉近正样本对、推开负样本对的效果。为提升网络对相似外观不同身份行人的辨别能力,本文对 2048 维的全局特征应用难样本采样三元组损失。难样本采样三元组损失表示如下:

$$L_{\text{triplet}} = - \sum_{i=1}^P \sum_{a=1}^D [\alpha + \max_{p=1, \dots, K} \|f_a^{(i)} - f_p^{(i)}\|_2 - \min_{\substack{n=1, \dots, K \\ j=1, \dots, P \\ j \neq i}} \|f_a^{(i)} - \hat{f}_n^{(j)}\|_2]_+. \quad (2)$$

对于一个训练批次,选取 P 个行人,每个行人选取 K 张图像。选取距离最远的标签、相同的样本作为正样本、选取距离最近的标签、不同的样本作为负样本。其中 $f_a^{(i)}$ 、 $f_p^{(i)}$ 、 $\hat{f}_n^{(j)}$ 分别表示选取的锚点图像、正样本图像、负样本图像的特征, α 是控制类内类间距离的超参数。

4 实验结果与分析

4.1 实验数据集

在 Market-1501 和 DukeMTMC-ReID 两大主流行人重识别数据集上进行本文方法的实验对比与分析。Market-1501 数据集图像由 5 个高分辨率摄像头和 1 个低分辨率摄像头采集,包含 1501 个行人的 32 668 张图像。其中训练集由 751 个不同身份行人的 19 732 张图像组成,测试集由来自剩余 750 个不同身份行人的 3 368 张查询图像与 19 732 张底库图像组成。DukeMTMC-ReID 数据集为行人跟随数据集 DukeMTMC 数据集的子集,使用 8 个高分辨率摄像头采集,包含 1401 个行人的 36 411 张图像。随机采样 702 个不同身份行人的 16 522 张图像作为训练

集,测试集包括来自剩余 702 个不同身份行人的 2 228 张查询图像和 17 661 张底库图像。

4.2 评价指标

本文实验使用 Rank-1 以及平均准确率(mean average precision, mAP)作为评价指标对行人重识别模型进行评估。

4.3 实验设置

本文实验使用在 ImageNet 数据集上预训练的 ResNet50 作为提取图像特征的骨干网络,并将最后一个卷积块的步长设为 1。初始化行人图像尺寸为 384×128 ,并采用水平翻转、随机擦除做数据增强。实验设置批量大小为 16, $P=4$, $K=4$, $\alpha=1.2$ 。模型优化器采用 Adam,共训练 500 个 epoch,初始学习率为 2×10^{-4} ,在 320 个 epoch 后衰减至 0.2×10^{-4} ,380 个 epoch 后衰减至 0.02×10^{-4} 。在测试阶段,融合 512 维的全局特征向量与局部特征向量作为行人的整体特征表示进行检索。本文实验环境为 NVIDIA Tesla P100,深度学习框架为 PyTorch。

4.4 实验分析

4.4.1 消融实验

为了验证本文方法的有效性,在 Market-1501 和 DukeMTMC-ReID 数据集上设计一系列消融实验,实验结果如表 1 和表 2 所示。其中 G 表示全局分支;L 表示局部分支中,舍弃局部关联特征,而是像 PCB 方法^[3]将 6 个水平平均分的部件直接进行切割、

表 1 Market-1501 数据集上不同方法的结果 (%)

Tab. 1 Results of different methods on the Market-1501 dataset (%)

Methods	Rank-1	mAP
ResNet50+G	92.6	81.7
ResNet50_CBAM+G	93.4	83.7
ResNet50_CBAM+G+L	95.0	86.0
ResNet50+G+LA	94.9	87.9
ResNet50_CBAM+G+LA	95.3	88.7
ResNet50_CBAM+G+LA+RK	96.3	94.8

表 2 DukeMTMC-ReID 数据集上不同方法的结果 (%)

Tab. 2 Results of different methods on the DukeMTMC-ReID dataset (%)

Methods	Rank-1	mAP
ResNet50+G	86.2	72.7
ResNet50_CBAM+G	87.7	74.7
ResNet50_CBAM+G+L	89.0	78.4
ResNet50+G+LA	89.8	79.6
ResNet50_CBAM+G+LA	90.1	80.7
ResNet50_CBAM+G+LA+RK	92.8	91.2

平均池化作为局部特征;LA 表示采用本文方法的局部关联特征作为局部特征;RK 表示在本文算法中加入重排序算法^[14],进一步优化检索结果。

从表中可以看出,使用 ResNet50 提取特征并只保留全局分支时,效果最差,在 Market-1501 数据集上 Rank-1 和 mAP 分别为 92.6%和 81.7%,DukeMTMC-ReID 数据集上 Rank-1 和 mAP 分别为 86.2%和 72.7%。将 CBAM 嵌入 ResNet50 后,在 Market-1501 和 DukeMTMC-ReID 两个数据集上的 mAP 指标都提升了 2.0%。同时,相比于 ResNet50+G+LA 使用 ResNet50 网络获取全局特征与局部关联特征,ResNet50_CBAM+G+LA 在 Market-1501 数据集上与 DukeMTMC-ReID 数据集上的 mAP 指标分别提高了 0.8%和 1.1%,这证明 CBAM 能有效关注行人图像中的关键区域,显著提升网络性能。为了进一步验证局部关联特征的效果,将 ResNet_CBAM+G+L 与 ResNet_CBAM+G+LA 进行比较,采用局部关联特征后在 Market-1501 数据集上 Rank-1 与 mAP 分别提升了 0.3%和 2.7%,在 DukeMTMC-ReID 数据集上 Rank-1 与 mAP 分别提升了 1.1%和 2.3%。此外,在引入重排序算法进行优化后,本文方法在 Market-1501 数据集和 DukeMTMC-ReID 数据集上的 mAP 指标进一步提高,分别达到了 94.8%和 91.2%。

4.4.2 与现有方法的对比

表 3 和表 4 分别展示了本文方法在 Market-1501 数据集和 DukeMTMC-ReID 数据集上与其他经典行人重识别方法的对比,包括传统方法和深度方法。分析表中数据,可以看出,本文方法在两个主流数据集上均取得了较好的效果。

表 3 Market-1501 数据集上与其他方法的对比(%)

Tab. 3 Comparison with other methods on the Market-1501 dataset(%)

Methods	Rank-1	mAP
LOMO+XQDA ^[15]	43.8	22.0
BoW+kissme ^[16]	43.2	22.0
IDE ^[8]	72.5	46.0
DPEL ^[17]	88.6	72.6
HA-CNN ^[10]	91.2	75.7
PCB ^[3]	92.3	77.4
PCB+RPP ^[3]	93.8	81.6
MGN ^[11]	95.7	86.9
Pyramid ^[4]	95.7	88.2
BoT ^[18]	94.5	85.9
RRID ^[12]	95.2	88.9
RGA ^[19]	95.8	88.1
Method of this article	95.3	88.7
Method of this article+RK	96.3	94.8

表 4 DukeMTMC-ReID 数据集上与其他方法的对比(%)

Tab. 4 Comparison with other methods on the DukeMTMC-ReID dataset(%)

Methods	Rank-1	mAP
LOMO+XQDA ^[15]	30.8	17.0
BoW+kissme ^[16]	25.1	12.2
IDE ^[8]	65.2	44.9
DPEL ^[17]	79.2	60.0
HA-CNN ^[10]	80.5	63.8
PCB ^[3]	81.7	66.1
PCB+RPP ^[3]	83.3	69.2
MGN ^[11]	88.7	78.4
Pyramid ^[4]	89.0	79.0
BoT ^[18]	86.4	76.4
RRID ^[12]	89.7	78.6
RGA ^[19]	86.1	74.9
Method of this article	90.1	80.7
Method of this article+RK	92.8	91.2

相比于早期的传统手工方法 XQDA^[15]等,基于深度学习的方法通过设计有效提取特征的卷积网络和优化网络的损失函数大大提升了行人重识别的精度。与经典的深度方法 PCB+RPP^[3]相比,本文方法在 Market-1501 数据集上的 Rank-1 指标和 mAP 指标分别高出了 1.5%和 7.1%,在 DukeMTMC-ReID 数据集上的 Rank-1 指标和 mAP 指标分别高出了 6.8%和 11.5%。同时,与现有方法 RRID^[12]、RGA^[19]对比,本文方法也取得了具有竞争力的结果。为了更直观地展示本文方法的优越性,分别选择姿态变化、背景干扰、物体遮挡情况下的行人图像进行检索,将检索结果可视化,如图 4 所示,其中绿色框表示结果匹配正确,红色框表示结果匹配错误。第 1 排行人的查询图像中,该行人骑着自行车。在姿态变化巨大的情况下,仍能被正确匹配。第 2 排中检索目标行人存在背景行人的干扰,本模型能够抑制噪声干扰,关注有用信息,提取判别力的行人特征进行检索。第 3 排中,目标行人的下半身被其他行人遮挡,仍取得了不错的检索结果。



图 4 检索结果可视化

Fig. 4 Visualization of retrieval results

5 结 论

本文提出一种基于注意力机制与局部关联特征的行人重识别方法。通过在骨干网络中嵌入卷积注意力模块,使网络关注行人图像中对重识别任务影响更大的区域。同时,本文设计局部关联特征,与全局特征相结合,使网络既能关注到行人的整体信息,又能关注到具备关联性的细节局部信息,同时减轻了行人部件不对齐问题的影响。实验表明,本文方法在两大行人重识别数据集上都获得了较高的准确率。下一步的工作将针对无监督行人重识别如何有效利用行人局部特征,在不依赖标注数据的情况下,提高行人重识别的性能。

参考文献:

- [1] LUO H, JIANG W, FAN X, et al. A survey on deep learning based person re-identification[J]. Acta Automatica Sinica, 2019, 45(11): 2032-2049.
罗浩, 姜伟, 范星, 等. 基于深度学习的行人重识别研究进展[J]. 自动化学报, 2019, 45(11): 2032-2049.
- [2] ZHAO H Y, TIAN M Q, SUN S Y, et al. Spindle net: person re-identification with human body region guided feature decomposition and fusion[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), July 21-26, 2017, Honolulu, HI, USA. New York: IEEE, 2017: 907-915.
- [3] SUN Y F, ZHENG L, YANG Y, et al. Beyond part models: person retrieval with refined part pooling (and a strong convolutional baseline)[C]//2018 European Conference on Computer Vision (ECCV), September 10-13, 2018, Munich, Germany. Heidelberg: Springer, 2018: 501-518.
- [4] ZHENG F, DENG C, SUN X, et al. Pyramidal person re-identification via multi-loss dynamic training[C]//2019 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 16-20, 2019, Long Beach, CA, USA. New York: IEEE, 2019: 8514-8522.
- [5] HU J, SHEN L, SUN G. Squeeze-and-excitation networks[C]//2018 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 18-22, 2018, Salt Lake City, UT, USA. New York: IEEE, 2018: 7132-7141.
- [6] WOO S, PARK J, LEE J, et al. CBAM: convolutional block attention module[C]//2018 European Conference on Computer Vision (ECCV), September 10-13, 2018, Munich, Germany. Heidelberg: Springer, 2018: 3-19.
- [7] HE S T, LUO H, WANG P C, et al. TransReID: Transformer-based object re-identification[EB/OL]. (2021-02-08) [2021-12-24]. <https://arxiv.org/abs/2102.04378>.
- [8] ZHENG L, YANG Y, HAUPTMANN A G. Person re-identification: past, present and future[EB/OL]. (2016-10-10) [2021-12-24]. <https://arxiv.org/abs/1610.02984>.
- [9] WEI L H, ZHANG S L, YAO H T, et al. GLAD: global-local-alignment descriptor for pedestrian retrieval[C]//2017 ACM International Conference on Multimedia (ACM MM), October 23-27, 2017, Mountain View City, CA, USA. New York: ACM, 2017: 420-428.
- [10] LI W, ZHU X T, GONG S G, et al. Harmonious attention

- network for person re-identification[C]//2018 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 18-22, 2018, Salt Lake City, UT, USA. New York: IEEE, 2018: 2285-2294.
- [11] WANG G S, YUAN Y F, CHEN X, et al. Learning discriminative features with multiple granularities for person re-identification[C]//2018 ACM International Conference on Multimedia (ACM MM), October 22-26, 2018, Seoul, Korea. New York: ACM, 2018: 274-282.
- [12] PARK H, HAM B. Relation network for person re-identification[C]//2020 AAAI Conference on Artificial Intelligence (AAAI), February, 7-12, 2020, New York, NY, USA. Palo Alto: AAAI, 2020: 11839-11847.
- [13] HE K M, ZHANG X Y, REN S Q, et al. Deep residual learning for image recognition[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE, 2016: 770-778.
- [14] ZHONG Z, ZHENG L, CAO D L, et al. Re-ranking person re-identification with k-reciprocal encoding [C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), July 21-26, 2017, Honolulu, HI, USA. New York: IEEE, 2017: 3652-3661.
- [15] LIAO S C, HU Y, ZHU X Y, et al. Person re-identification by local maximal occurrence representation and metric learning[C]//2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 7-12, 2015, Boston, MA, USA. New York: IEEE, 2015: 2197-2206.
- [16] ZHENG L, SHEN L Y, TIAN L, et al. Scalable person re-identification: a benchmark [C]//2015 IEEE International Conference on Computer Vision (ICCV), December 7-13, 2015, Santiago, Chile. New York: IEEE, 2015: 1116-1124.
- [17] CHEN Y B, ZHU X T, GONG S G. Person re-identification by deep learning multi-scale representations [C]//2017 IEEE International Conference on Computer Vision (ICCV), October 22-29, 2017, Venice, Italy. New York: IEEE, 2017: 2590-2600.
- [18] LUO H, GU Y Z, LIAO X Y, et al. Bag of tricks and a strong baseline for deep person re-identification [C]//2019 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 16-20, 2019, Long Beach, CA, USA. New York: IEEE, 2019: 4321-4329.
- [19] ZHANG Z, LAN C, ZENG W, et al. Relation-aware global attention for person re-identification [C]//2020 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), November 23-25, 2020, Seattle, WA, USA. New York: IEEE, 2020: 3186-3195.

作者简介:

张晓涵 (1997—), 女, 硕士研究生, 主要从事行人重识别、图像检索方面的研究。