

多粒度特征融合的行人再识别研究

张良，车进，杨琦

引用本文：

张良，车进，杨琦. 多粒度特征融合的行人再识别研究[J]. 液晶与显示, 2020, 35(6): 555–563.

ZHANG Liang, CHE Jin, Yang Qi. Multi-granularity feature fusion for person re-identification[J]. *Chinese Journal of Liquid Crystals and Displays*, 2020, 35(6): 555–563.

在线阅读 View online: <https://doi.org/10.3788/YJYXS20203506.0555>

您可能感兴趣的其他文章

Articles you may be interested in

多尺度特征融合与极限学习机结合的交通标志识别

Traffic sign recognition based on multi-scale feature fusion and extreme learning machine

液晶与显示. 2020, 35(6): 572–582 <https://doi.org/10.3788/YJYXS20203506.0572>

基于深度卷积神经网络的多特征融合的手势识别

Multi-feature fusion gesture recognition based on deep convolutional neural network

液晶与显示. 2019, 34(4): 417–422 <https://doi.org/10.3788/YJYXS20193404.0417>

角度空间三元组损失微调的人脸识别

Face recognition of triple loss fine-tuning in angular space

液晶与显示. 2019, 34(1): 110–117 <https://doi.org/10.3788/YJYXS20193401.0110>

基于卷积神经网络的响应自适应跟踪

Response adaptive tracking based on convolution neural network

液晶与显示. 2018, 33(7): 596–605 <https://doi.org/10.3788/YJYXS20183307.0596>

基于多尺度残差网络的CT图像超分辨率重建

CT image super-resolution reconstruction based on multi-scale residual network

液晶与显示. 2019, 34(10): 1006–1012 <https://doi.org/10.3788/YJYXS20193410.1006>

文章编号:1007-2780(2020)06-0555-09

多粒度特征融合的行人再识别研究

张 良^{1,2}, 车 进^{1,2*}, 杨 琦^{1,2}

(1. 宁夏大学 物理与电子电气工程学院, 宁夏 银川 750021;
2. 宁夏大学 沙漠信息智能感知重点实验室, 宁夏 银川 750021)

摘要:结合全局特征和局部特征是提高行人再识别精度的一种途径。现有的算法通常从人体特定的语义区域提取特征, 由于没有将人体结构考虑在内, 增加了学习难度, 在差异较大的场景下效率和鲁棒性较差。为了较好地解决上述问题, 本文提出一种融合了全局特征、局部特征以及人体结构特征的多粒度特征融合的行人再识别算法。本算法不引入任何人体结构先验知识, 在特征提取方面, 采用均值池化和最大池化对特征图加权得到强辨识性的全局特征。对特征图切片得到局部特征, 在原有局部特征的基础上, 引入局部相对特征作为人体结构特征。在度量方面, 采用三元组损失与 ID 损失在不同尺度下的多级监督机制。在 Market1501、DukeMTMC-reID 的实验表明, 算法的 Rank-1 指标相比于部分卷积基线(PCB)方法提升了 1.3%、3.9%, 平均精度均值(mAP)提升了 5.1%、9.8%。

关 键 词:全局特征;局部特征;人体结构特征;特征融合;深度学习;卷积神经网络;行人再识别

中图分类号:TP391.4 **文献标识码:**A **doi:**10.3788/YJYXS20203506.0555

Multi-granularity feature fusion for person re-identification

ZHANG Liang^{1,2}, CHE Jin^{1,2*}, Yang Qi^{1,2}

(1. School of Physics and Electronic-Electrical Engineering,
Ningxia University, Yinchuan 750021, China;

2. Key Laboratory of Intelligent Sensing for Desert Information,
Ningxia University, Yinchuan 750021, China)

Abstract: Combining global features and local features is one way to improve person re-identification accuracy. Existing algorithms usually extract features from specific semantic regions of the human body. Since the human body structure is not taken into account, the learning difficulty is increased, and the efficiency and robustness are poor in the scenes with large differences. In order to solve the above problems, this paper proposes a person re-identification algorithm based on multi granularity feature fusion, which combines global feature, local feature and human structure feature. The algorithm does not introduce any prior knowledge of human body structure. In terms of feature extraction, the average pooling and maximum pooling are used to weight the feature map to obtain strong global features. Local features are obtained by slicing the feature map. Based on the original local features, local relative features are introduced as human structural features. In terms of metrics, a multi-level

收稿日期:2019-11-04; 修订日期:2019-11-25.

基金项目:国家自然科学基金(No. 61861037)

Supported by National Natural Science Foundation of China(No. 61861037)

* 通信联系人, E-mail:koalache@126.com

supervision mechanism with triple loss and ID loss at different scales is used. Experiments on Market1501 and DukeMTMC-reID show that the Rank-1 index of the algorithm is 1.3% and 3.9% higher than Part-based Convolutional Baseline (PCB) method, and the mean Average Precision (mAP) is 5.1% and 9.8% higher than PCB method.

Key words: global feature; local feature; human structural feature; feature fusion; deep learning; convolutional neural network; person re-identification

1 引言

行人再识别^[1-4] (Person ReID) 通常被视为一个图像检索问题, 即从非重叠的摄像机视角中实现行人匹配任务。给定一幅查询图像, 行人再识别旨在从行人图像库中找出该行人的所有图像。由于拍摄的图像受到光照、姿势、视角、遮挡、背景杂波等影响, 会造成同一行人在不同摄像机视角中表现出较大的类内差异, 给行人再识别任务造成了巨大的困难和挑战。

随着深度学习在计算机视觉领域的广泛应用, 提取的深度特征已被证实具有良好的判别性和鲁棒性, 近年来涌现出一批优秀的行人再识别算法。文献[5-7]考虑到样本空间中类内差异与类间差异的影响, 提出多损失联合训练的方式来优化网络。文献[8]提出一种改进的三元组损失算法, 在特征空间内将正负样本对分开, 而且优化了正样本对之间的距离。文献[9]提出一种难样本采样损失函数, 在每一个训练的批量样本中挑选一张图像, 从批量样本中选择一个最难的正样本以及负样本与该图像组成一个三元组, 实验结果显示, 使用该损失函数训练的网络模型鲁棒性更强。在此之后, 不少研究学者开始利用局部特征来提高行人再识别的精度。文献[10-12]提出将人物图像由上至下分为几个不同部分, 对不同的部分进行特征提取, 利用多损失共同监督网络, 融合不同部分的特征得到最终的特征表示。上述算法的缺点就是对图像对齐的要求较高, 为了解决上述的不对齐问题, 文献[13]提出一种基于 SP 距离的自动对齐模型, 在不需额外监督信息的条件下自动对齐局部特征。文献[14-17]提出利用姿态预估模型得到关键点, 利用人体关键点将人体划分为不同的区域, 得到不同尺度下的特征, 通过融合各部分特征进而提高精度。文献[18]也采用人体关键点作为先

验知识, 不同于文献[14]采用仿射变换对人体区域进行划分, 它直接利用关键点将人体分为 7 个 ROI 区域, 最后将全局特征与局部特征进行融合。文献[19]提出一种全局-局部对齐特征描述子, 不同的是将人体区域分为 3 部分, 每部分的输入对应每部分的损失, 并不是融合为一个特征计算损失。文献[20-21]提出同时学习全局与局部特征, 并通过权重子网络对不同部分加权融合局部与全局特征。然而以上的算法都必须有一个姿态预估模型, 增加了学习的难度。

本文提出一种用于 ReID 的深度学习算法, 在不引入任何人体框架先验知识的前提下, 融合人物图像的全局、局部、以及人体结构特征。采用三元组损失与 ID 损失在不同尺度下的多级监督机制优化网络。实验结果表明, 这种采用联合监督机制的多粒度特征融合算法有一定的优越性。

2 网络架构

文献[22-24]提出多尺度提取特征, 融合不同的特征得到最终的特征表示。考虑到网络的竞争性能以及相对简洁的架构, 本文采用 ResNet-50^[25]作为基准网络, 保留全均值池化层 (Global Average Pooling, GAP) 之前的结构, 之后的网络层均被抛弃。对修改的基准网络输出的特征图采用不同的池化策略, 得到加权后的全局特征、局部特征以及人体结构特征, 如图 1 所示。在训练阶段, 采用多级损失函数监督机制对网络进行训练。在测试阶段, 将全局特征、局部特征以及人体结构特征级联形成最终的行人特征描述符实现多粒度特征融合。

2.1 全局特征

行人图像经过基准网络得到特征图 F , F 经过一个双分支网络, 上层分支为全局均值池化层 (Global Average Pooling, GAP), 由于平均池化能够感知整幅图像中的全局信息, 对邻域内的特

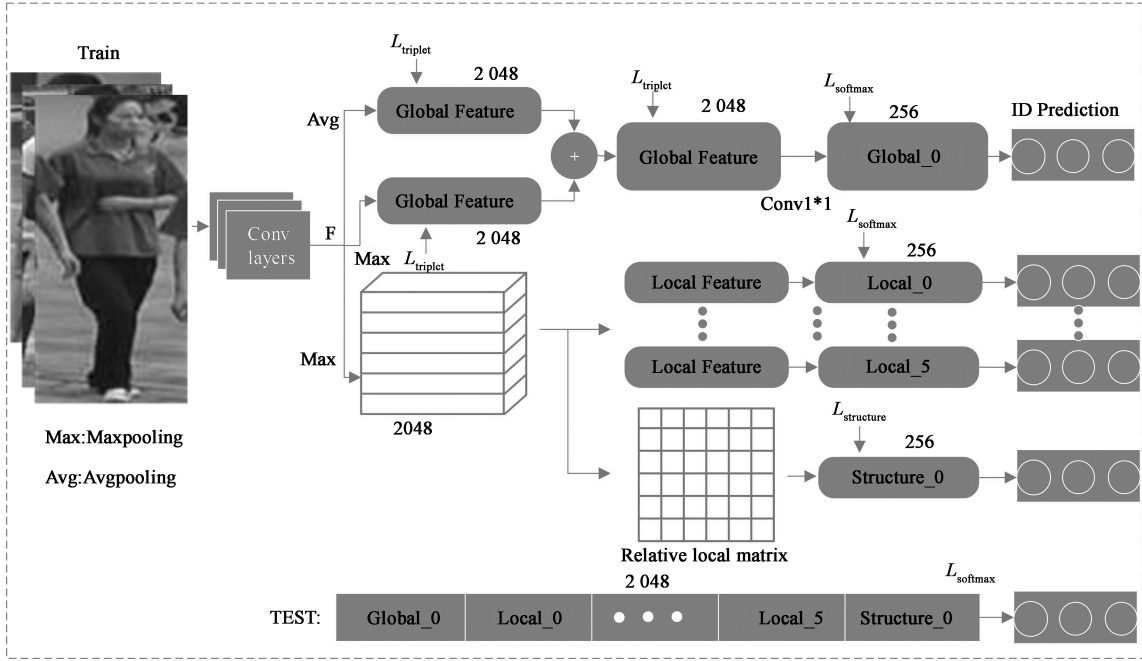


图1 网络架构

Fig.1 Network architecture

征点求平均,能够减小邻域大小受限造成的估计值方差增大,更多地保留图像的背景信息。下层分支为全局最大池化层(Global Max Pooling, GMP),最大池化的目标是提取最具有鉴别性的信息,对邻域内的特征点取最大值,能够减小卷积层参数误差造成估计均值的偏移。最大池化忽略了由于平均池化造成的干扰信息,保留更多的纹理信息。考虑到两种池化策略的不同优势,本文采用两种池化策略相结合得到更具有判别性的全局特征,可以表示为:

$$G = \text{avgpool}(F) + \text{maxpool}(F) \quad (1)$$

降维前的全局特征均采用批量硬三元损失^[26]优化网络,批量硬三元损失的核心内容是随机选择 P 类行人,每类随机选择 K 张图片,形成一个包含 PK 张图片的批量样本,对于该批量样本中某个确定的样本,选择批量样本中最难的正样本和最难的负样本来计算损失。优化该批量样本内最难的正样本以及最难的负样本,在特征空间内满足类内距离的最大化并且类间距离最小化,使得网络学习特征的鲁棒性变强。每条支路的损失函数都是独立且不共享网络参数,可以表示为:

$$L_{\text{bhtri}} = - \sum_{i=1}^P \sum_{a=1}^K [\alpha + \max_{p=1..k} \|f_a^{(i)} - f_p^{(i)}\|_2 - \min_{\substack{n=1..K \\ j=1..P \\ j \neq i}} \|f_a^{(i)} - f_n^{(j)}\|_2] \quad (2)$$

其中: f_a^i , f_p^i , f_n^i 分别表示从目标样本、正样本以及负样本提取的特征, α 表示一个常量。这里的正负样本指的是与目标样本相比是否为相同身份的行人。

2.2 局部特征

当关注的对象受到遮挡时,由于全局特征无法分辨出前景和背景这一劣势,全局特征的表达能力就会被削弱。此时需要局部特征来弥补全局特征的不足,局部特征是某些特定区域的稳定且具有良好的判别性的特征,即使发生了部分遮挡等情况,一些局部的强辨识性特征仍然存在。

最大池化的目标是提取最具有鉴别性的特征,提取邻域内最大的特征点,因此采用最大池化策略获取局部特征。特征图切片^[8]是一种常见的局部特征提取方式,本文对特征图 F 首先进行水平方向池化,将特征向量在水平方向的分量池化为单个部件级别的向量。由于人物图像是垂直的,在垂直方向上特征的鲁棒性较强,因此将经过

水平池化后的特征图 F 沿垂直方向均分为 6 个区域,每个区域特征降维得到局部特征,采用 Softmax 损失计算每部分的损失,局部特征中所有的 Softmax 损失共享参数,可以表示为:

$$L_{id} = \sum_{i=1}^N -q_i \log(p_i) \begin{cases} q_i = 0, y \neq i \\ q_i = 1, y = i \end{cases}, \quad (3)$$

其中: p_i 表示预测 i 属于类别 q_i 的概率, N 表示行人图像的总数量。

由于测试集中的人物图像从未在训练集中出现过,因此行人再识别可以看作是 One-Shot 任务,因此防止网络过拟合是十分重要的。Label Smoothing^[25] (LS) 是一种广泛的防止分类任务过度拟合方法,具体是改变公式(3)中的 q_i :

$$q_i = \begin{cases} 1 - \frac{N-1}{N}\epsilon & \text{if } i = y \\ \epsilon/N & \text{otherwise} \end{cases}, \quad (4)$$

其中: N 表示行人图像的总数量, ϵ 是一个常量,本文中设置为 0.1。

2.3 人体结构特征

本文在不引入任何人体姿态先验知识的前提下,对特征图进行切片操作得到局部特征。利用局部特征的相对距离得到局部特征相对矩阵 D ,进而得到人体结构特征,如图 2 所示(图 2 中仅仅列出 h_1 与其他块的关系)。局部特征相对矩阵 D 是由局部特征之间的余弦距离得到,余弦距离使用两个特征向量夹角的余弦值作为衡量两个特征之间的差异,相比于欧式距离,余弦距离更加注重两个向量在方向上的差异,使用余弦距离能够更好地显示两个特征向量之间的相似度。定义 d_{ij} 表示两个局部特征向量的余弦距离:

$$d_{ij} = h_i^T h_j, \quad (5)$$

以矩阵的表示形式为:

$$D = H^T H. \quad (6)$$

由于人物图像是垂直分布的,相对距离矩阵 D 对各种可变形的人体结构具有很强的鲁棒性。受到摄像机视角、可变的姿势等影响,采集到的人物图像在相同部位的大小、形状存在很大差异。因此将矩阵 D 重构为特征向量,用来表示人体结构特征。为了通过人体结构特征区分不同的行人,在重构的特征向量之后加入两层全连接层以及对应的人体结构损失使得人体结构特征是可以学习的。我们可以根据相对距离矩阵 D 区分不

同行人的 ID 信息,采用 Softmax 作为人体结构的损失函数,如式(3)所示。

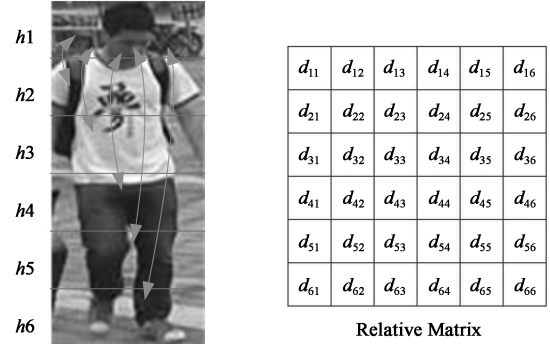


图 2 人体结构和局部特征相对矩阵

Fig. 2 Human body structure and local feature relative matrix

2.4 总损失

采用三元组损失与 ID 损失相结合的多级监督机制优化网络,网络的总损失包括 ID 损失、三元组损失、人体结构损失:

$$L = L_{id} + L_{bhtri} + L_{str}. \quad (7)$$

3 实 验

3.1 数据预处理

行人图像库中,行人图像有时会被其他物体遮挡,从而降低了模型的鲁棒性。为了解决该问题,提高再识别模型的泛化能力,Zhong^[28]提出一种随机擦除的数据扩充办法(Random Erasing Augmentation, REA)。不同于以往的水平翻转进行数据增强,在实际训练中,一张来源于批量样本的输入照片以概率 p_e 被擦除,保持不变的概率为 $1 - p_e$ 。REA 方法是从图片中随机的选取一个矩形区域(W_e, H_e),用一个随机的像素值擦除该矩形区域的原有的像素值。即给定一幅图像 $S = W \times H$,以及擦除区域 $S_e = W_e \times H_e$,用 $S_1 = S_e/S$ 表示擦除区域的面积比,其中 $S_1 \in (s_1, s_h)$,其中 s_1, s_h 表示擦除区域占原图像面积的最小值以及最大值。擦除区域的宽高比 r_e 在 r_1 和 r_2 之间进行随机初始化,其中 r_1, r_2 表示擦除区域宽高比的最小值以及最大值。其中 $H_e = \sqrt{S_e \times r_e}, W_e = \sqrt{S_e / r_e}$, REA 随机初始化一个点 $P = (X_e, Y_e)$,如果 $X_e + W_e < W$ 且 $Y_e + H_e < H$,我们则选定区域 $I_e = (X_e, Y_e, X_e + W_e, Y_e +$

H_c)作为选定的区域。否则,重复上述的进程,直到选定的区域满足上述条件。REA 效果图如图3所示,第一行为原始图像,第二行为经过 REA 的图像。



图3 REA 效果图
Fig.3 REA renderings

3.2 实验设置

为了验证本算法的有效性,实验在 pytorch

环境下搭建深度网络进行实验,实验设备为搭载4块 TITAN V GPU 的图形工作站。本文中,擦除概率 p_c 设置为 0.5,设置擦除面积最小值 $s_1 = 0.02$ 、最大值 $s_h = 0.4$,擦除区域宽高比最小值 $r_1 = 0.3$ 、最大值 $r_2 = 3.33$,常量 ϵ 设置为 0.1,基础的学习率设置为 $2E \sim 4$,迭代周期设置为 160,批量大小设置为 128,类别 $P = 32$, $K = 4$,常量 $\alpha = 0.3$,正负样本的初始比例为 1/3。

3.3 实验数据库

在两大公开的数据集 Market1501、DukeMTMC-reID 对模型的有效性进行实验验证,并与几种最新的行人再识别算法进行比较,使用累积匹配特征曲线(CMC)^[29]和平均精度(mAP)^[30-31]这两个指标来评价所提算法的性能,其中 CMC 曲线的指标为:首位命中率(Rank-1, R1)、前五位命中率(Rank-5, R5)、前十位命中率(Rank-10, R10)。表1列出了数据集的详细信息。

Market1501 数据集是行人再识别中最常用的数据集,该数据集包含了 1 501 个行人的 32 668 幅带标签的边界框。由 6 个不同的摄像头

表1 数据集的详细信息
Tab.1 Details of the dataset

数据集	ID	Train ID	Test ID	Bbox	Bbox/ID	Cam
Market1501	1 501	751	750	32 668	21.8	6
DukeMTMC-ReID	1 404	702	702	36 441	25.9	8

在不同时间、不同光照条件下拍摄。数据集被分为训练集和测试集,训练集包含 751 个行人的 12 936 个裁剪图像,测试集包含 750 个行人的 19 732 个裁剪图像,边界框直接由可变形零件模型(DPM)^[32]检测而不是使用手绘的边界框,这更接近于真实的场景。

DukeMTMC-reID 数据集由 8 个摄像头采集的视频序列经过等间隔采样得到,数据库包含 1 812 个行人,其中 1 404 个行人出现在大于两个摄像头视角下,这些出现在多摄像头的图片分为 702 个行人训练集 16 522 张和 702 个行人的测试集 19 889 张,剩下的 408 个行人仅出现在一个摄像头下,因此并没有采用该部分的行人图像。

3.4 实验数据

为了验证不同的池化策略下的全局特征对模

型的影响,在 Market1501 数据集上进行实验,数据如表 2 所示,由表中数据可得结合两种池化策略的网络性能比任何单一池化策略的性能要好,进一步验证了结合两种池化策略的可行性与有效性。

表2 不同池化策略下的全局特征对模型的影响
Tab.2 Effect of global features on different models under different pooling strategies

Model	R1/%	R5/%	R10/%	mAP/%
Avg	90.2	95.3	97.2	78.8
Max	90.3	94.7	97.4	79.0
Avg+Max	91.0	95.7	97.8	79.2

本文提出的特征融合策略,在 Market1501 数据集上对全局特征、局部特征、人体结构特征进行消融实验,实验数据如表 3、4 所示,其中 str 表示人体结构特征,由表中的数据可得融合人体结构特征能进一步提高网络模型的性能,进一步验证了可以利用局部特征之间的相对特征表征人体

表 3 全局特征与人体结构特征对模型的影响

Tab.3 Effect of global features and anatomical features on the model

Model	R1/%	R5/%	R10/%	mAP/%
Global	91.0	95.7	97.8	79.2
Global + str	92.7	97.3	98.3	80.8

表 5 全局特征、局部特征与人体结构特征对模型的影响

Tab.5 Effect of global features, local features and anatomical features on the model

Model	Market1501				DukeMTMC-ReID			
	R1/%	R5/%	R10/%	mAP/%	R1/%	R5/%	R10/%	mAP/%
Global+Local	92.4	96.4	97.9	80.7	84.5	92.5	93.8	73.7
Global+Local+str	93.7	97.6	98.3	82.4	85.8	93.3	95.1	75.1

本文所提出的算法与现有的算法进行比较如表 6 所示,可以看出所提出的算法性能优越,在 Market1501 上的 Rank-1 仅仅比 PCB+RPP 稍低,但 mAP 的性能要高于 PCB+RPP。在 DukeMTMC-ReID 的 Rank-1 的表现效果很好,mAP 仅次于 PSE+ECN。由实验数据可得,所提出的算法在两大公开数据集上的表现优越,Rank-1 以及 mAP 要高于现有的大多数算法,充分说明了所提算法的优越性。

表 6 本文算法与现有的算法比较

Tab.6 Comparison of the algorithm in this paper with existing algorithms

Model	Market1501		DukeMTMC-ReID	
	R1/%	mAP/%	R1/%	mAP/%
BoW+Kissme[33]	20.76	44.42	25.13	12.17
IDE[6]	72.5	46.0	67.7	47.1
LOMO+XQDA[1]	72.58	51.96	30.75	17.04
Part-Aligned[17]	81.0	63.4	84.4	69.3
SVDNet[34]	82.3	62.1	76.7	56.8

表 4 局部特征与人体结构特征对模型的影响

Tab.4 Effect of local features and human structural features on the model

Model	R1/%	R5/%	R10/%	mAP/%
Local	91.1	95.8	97.5	78.9
Local + str	92.2	97.1	98.1	79.6

结构特征。

本文提出的行人再识别算法在两大公开的数据集上进行实验,实验数据如表 5 所示,由表 3、4、5 中的数据可得,多粒度特征融合能进一步提高网络的性能。

续 表

Model	Market1501		DukeMTMC-ReID	
	R1/%	mAP/%	R1/%	mAP/%
PAN[15]	82.8	63.4	71.59	51.51
APR[7]	84.3	64.7	70.69	51.88
DPFL[22]	88.9	73.1	79.2	60.0
HA-CNN[21]	91.2	75.7	80.5	63.8
DaRe[20]	89.0	76.0	80.2	64.5
PSE+ECN[16]	90.3	84.0	85.2	79.8
GP-reid[24]	92.2	81.2	85.2	72.8
PCB[12]	92.4	77.3	81.9	65.3
PCB+RPP[12]	93.8	81.6	83.3	69.2
Ours	93.7	82.4	85.8	75.1
Ours+rerank	94.8	92.0	87.7	84.9

4 结 论

本文在不考虑任何人体结构先验知识的前提下,提出一种融合人物图像的全局特征、局部特征

以及人体结构特征的行人再识别算法。采用不同的池化策略加权捕获人物图像的全局特征,对特征图切片获取局部强辨识性特征,利用局部特征的相对距离得到人体结构特征。采用不同的损失函数在不同尺度下对网络进行监督,这种利用多

级损失函数进行监督的监督机制使得网络模型变得更加鲁棒。在公开数据集的实验结果显示,算法的 Rank-1 指标相比于 PCB 方法提升了 1.3%、3.9%,mAP 提升了 5.1%、9.8%。

参 考 文 献:

- [1] LIAO S C, HU Y, ZHU X Y, *et al.* Person re-identification by local maximal occurrence representation and metric learning [C]//*Proceedings of 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Boston, MA: IEEE, 2015: 2197-2206.
- [2] 陈亮雨,李卫疆. 多形状局部区域神经网络结构的行人再识别 [J]. 中国图象图形学报, 2019, 24(11):1932-1941.
CHEN L Y, LI W J. Multishape part network architecture for person re-identification [J]. *Journal of Image and Graphics*, 2019, 24(11):1932-1941. (in Chinese)
- [3] 蒋建国,杨宁,齐美彬,等. 区域块分割与融合的行人再识别 [J]. 中国图象图形学报, 2019, 24(4):513-522
JIANG J G, YANG N, QI M B, *et al.* Person re-identification with region block segmentation and fusion [J]. *Journal of Image and Graphics*, 2019, 24(4):513-522. (in Chinese)
- [4] 朱福庆,孔祥维,付海燕,等. 两路互补对称 CNN 结构的行人再识别 [J]. 中国图象图形学报, 2018, 23(7):1052-1060
ZHU F Q, KONG X W, FU H Y, *et al.* Two-stream complementary symmetrical CNN architecture for person re-identification [J]. *Journal of Image and Graphics*, 2018, 23(7):1052-1060. (in Chinese)
- [5] 陈兵,查宇飞,李运强,等. 基于卷积神经网络判别特征学习的行人重识别 [J]. 光学学报, 2018, 38(7):0720001.
CHEN B, ZHA Y F, LI Y Q, *et al.* Person re-identification based on convolutional neural network discriminative feature learning [J]. *Acta Optica Sinica*, 2018, 38(7): 0720001. (in Chinese)
- [6] ZHENG L, YANG Y, HAUPTMANN A G. Person re-identification: past, present and future [J]. *arXiv*, 2016: 1610.02984.
- [7] LIN Y T, ZHENG L, ZHENG Z D, *et al.* Improving person re-identification by attribute and identity learning [J]. *ArXiv*, 2017:1703.07220.
- [8] CHENG D, GONG Y H, ZHOU S P, *et al.* Person re-identification by multi-channel parts-based CNN with improved triplet loss function [C]//*Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Las Vegas, NV: IEEE, 2016: 1335-1344.
- [9] HERMANS A, BEYER L, LEIBE B. In defense of the triplet loss for person re-identification [J]. *ArXiv*, 2017: 1703.07737.
- [10] SUN Y F, XU Q, LI Y L, *et al.* Perceive where to focus: learning visibility-aware part-level features for partial person re-identification [J]. *ArXiv*, 2019:1904.00537.
- [11] VARIOR R R, SHUAI B, LU J W, *et al.* A Siamese long short-term memory architecture for human re-identification [C]//*Proceedings of the 14th European Conference on Computer Vision*. Amsterdam, The Netherlands: Springer, 2016: 135-153.
- [12] SUN Y F, ZHENG L, YANG Y *et al.* Beyond part models: person retrieval with refined part pooling (and a strong convolutional baseline) [C]//*Proceedings of the 15th European Conference on Computer Vision*. Munich, Germany: Springer, 2018: 501-518.
- [13] ZHANG X, LUO H, FAN X, *et al.* AlignedReID: surpassing human-level performance in person re - identification [J]. *ArXiv*, 2017:1711.08184.
- [14] ZHENG L, HUANG Y J, LU H C, *et al.* Pose invariant embedding for deep person re-identification [J]. *ArXiv*, 2017:1701.07732.
- [15] ZHENG Z D, ZHENG L, YANG Y. Pedestrian alignment network for large-scale person re-identification [J].

- IEEE Transactions on Circuits and Systems for Video Technology*, 2019, 29(10): 3037-3045.
- [16] SARFRAZ M S, SCHUMANN A, EBERLE A, *et al.* A pose-sensitive embedding for person re-identification with expanded cross neighborhood re-ranking [C]//*Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Salt Lake City, UT: IEEE, 2018: 420-429.
- [17] SU H Y, WANG J H, TANG S Y, *et al.* Part-aligned bilinear representations for person re-identification [C]//*Proceedings of the 15th European Conference on Computer Vision*. Munich, Germany: Springer, 2018: 418-437.
- [18] 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]//*Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Honolulu, HI: IEEE, 2017: 907-915.
- [19] WEI L H, ZHANG S L, YAO H T, *et al.* GLAD: global-local-alignment descriptor for scalable person re-identification [J]. *IEEE Transactions on Multimedia*, 2019, 21(4): 986-999.
- [20] SU C, LI J N, ZHANG S L, *et al.* Pose-driven deep convolutional model for person re-identification [C]//*Proceedings of 2017 IEEE International Conference on Computer Vision (ICCV)*. Venice: IEEE, 2017: 3980-3989.
- [21] LI W, ZHU X T, GONG S G. Harmonious attention network for person re-identification [C]//*Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Salt Lake City, UT: IEEE, 2018: 2285-2294.
- [22] CHEN Y B, ZHU X T, GONG S G, *et al.* Person re-identification by deep learning multi-scale representations [C]//*Proceedings of 2017 IEEE International Conference on Computer Vision Workshops (ICCVW)*. Venice: IEEE, 2017: 2590-2600.
- [23] WANG Y, WANG L Q, YOU Y R *et al.* Resource aware person re-identification across multiple resolutions [C]//*Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Salt Lake City, UT: IEEE, 2018: 8042-8051.
- [24] 贡卫国, 史其琦, 王民. 基于深度卷积神经网络的多特征融合的手势识别[J]. 液晶与显示, 2019, 34(4): 417-422.
- YUN W G, SHI Q Q, WANG M. Multi-feature fusion gesture recognition based on deep convolutional neural networks [J]. *Chinese Journal of Liquid Crystals and Displays*, 2019, 34(4): 417-422. (in Chinese)
- [25] HE K M, ZHANG X Y, REN S Q, *et al.* Deep residual learning for image recognition [C]//*Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Las Vegas, NV: IEEE, 2016: 770-778.
- [26] HERMANS A, BEYER L, LEIBE B. In defense of the triplet loss for person re-identification [J]. *ArXiv*, 2017: 1703.07737.
- [27] SZEGEDY C, VANHOUCKE V, IOFFE S, *et al.* Rethinking the inception architecture for computer vision [C]//*Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Las Vegas, NV: IEEE, 2016: 2818-2826.
- [28] ZHONG Z, ZHENG L, KANG G, *et al.* Random erasing data augmentation [J]. *ArXiv*, 2017: 1708.04896.
- [29] 赵泉华, 张洪云, 李玉. 采用非规则标识点过程的 LiDAR 点云数据目标提取 [J]. 光学 精密工程, 2018, 26(5): 1201-1210.
- ZHAO Q H, ZHANG H Y, LI Y. Target extraction from LiDAR point cloud data using irregular geometry marked point process [J]. *Optics and Precision Engineering*, 2018, 26(5): 1201-1210. (in Chinese)
- [30] 金志刚 李静昆. 基于对象性和多层线性模型的协同显著性检测 [J]. 光学 精密工程, 2019, 27(8): 1845-1853.
- JIN Z G, LI J K. Co-saliency detection based on objectness and multi-layer linear model [J]. *Optics and Precision Engineering*, 2019, 27(8): 1845-1853. (in Chinese)
- [31] 雷金周, 曾令斌, 叶南. 工业机器人单目视觉对准技术研究 [J]. 光学 精密工程, 2018, 26(3): 733-741.
- LEI J Z, ZENG L B, YE N. Research on industrial robot alignment technique with monocular vision [J]. *Optics and Precision Engineering*, 2018, 26(3): 733-741. (in Chinese)
- [32] LIN L, WANG X L, YANG W, *et al.* Discriminatively trained and-or graph models for object shape detection [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2015, 37(5): 959-972.
- [33] ZHENG L, SHEN L Y, TIAN L, *et al.* Scalable person re-identification: a benchmark [C]//*Proceedings of 2015 IEEE International Conference on Computer Vision (ICCV)*. Santiago: IEEE, 2015: 1116-1124.

- [34] SUN Y F, ZHENG L, DENG W J, *et al.* SVDNet for pedestrian retrieval [C]//*Proceedings of 2017 IEEE International Conference on Computer Vision (ICCV)*. Venice: IEEE, 2017: 3820-3828.

作者简介:



张 良(1994—),男,山东潍坊人,硕士研究生,2017 年于青岛农业大学获得学士学位,主要从事图像处理、智能视频方面的研究。E-mail: zhangliang_nxu@163.com



车 进(1973—),男,宁夏银川人,博士,教授,2014 年于天津大学获得博士学位,主要从事图像处理、智能视频方面的研究。E-mail: koalache@126.com