

一种基于全局特征的行人重识别改进算法

张涛, 易争明*, 李璇, 孙星

天津大学电气自动化与信息工程学院, 天津 300072

摘要 基于全局特征的行人重识别算法主要使用交叉熵损失函数和三元组损失函数来监督网络的学习。然而, 原始三元组损失函数在增大类间距离的同时并未很好地优化类内距离, 为了解决这个问题, 提出一种基于全局特征的行人重识别改进算法。该算法是在三元组损失函数的基础上进行改进, 即在原始三元组损失函数中引入一项类内距离, 使改进后的三元组损失函数能够在增大类间距离的同时减小类内距离。在 Market1501、DukeMTMC-reID 和 CUHK03 数据集上进行大量实验, 实验结果表明所提算法得到的特征具有更强的判别性, 在基于全局特征的模型中可以取得最优的性能, 接近甚至超过一些基于局部特征的模型。

关键词 机器视觉; 光计算; 行人重识别; 全局特征; 三元组损失

中图分类号 TP391

文献标志码 A

doi: 10.3788/LOP57.241503

Improved Algorithm for Person Re-Identification Based on Global Features

Zhang Tao, Yi Zhengming*, Li Xuan, Sun Xing

School of Electrical and Information Engineering, Tianjin University, Tianjin 300072, China

Abstract Person re-identification algorithms based on global features primarily use the cross-entropy loss function and triplet loss function to supervise network learning. However, the original triplet loss function does not optimize an intraclass distance and increases an interclass distance. To solve this problem, an improved person re-identification algorithm based on global features is proposed. The algorithm is improved on the basis of the triplet loss function, that is, an intraclass distance is introduced into the original triplet loss function, so that the improved triplet loss can be reduced while increasing the interclass distance. A number of experiments have been conducted on the Market1501, DukeMTMC-reID, and CUHK03 datasets. The experimental results show that the proposed algorithm obtains discriminative features, and a model based on the global features can achieve a performance that approaches or even exceeds some local feature models.

Key words machine vision; optics in computing; person re-identification; global features; triplet loss

OCIS codes 150.0155; 200.4560; 100.4995; 100.4996; 100.4999

1 引言

行人重识别^[1]也称作行人再识别, 其是利用计算机视觉技术来判断图像或者视频中是否存在特定行人, 即给定一张行人图像, 检索跨设备下的该行人图像, 故行人重识别也可以看作图像检索的问题。行人重识别有着许多实际的应用, 例如刑事侦查、视频监控以及行为理解等。

近年来, 随着深度卷积神经网络的发展, 基于深度卷积神经网络的行人重识别算法的性能越来越好, 故吸引了越来越多的研究者的关注。当前基于

深度卷积神经网络的行人重识别算法一般使用 GoogleNet^[2]、ResNet^[3] 和 DensNet^[4] 等作为骨干网络来提取行人图像的特征, 再进行端到端的学习。

行人重识别算法使用的损失函数一般分为分类损失函数和度量损失函数两类。基于分类损失函数的网络称为 IDE(Identification Embedding)模型, 其将行人重识别看作一个分类任务, 将行人 ID 作为标签并用来训练行人重识别网络, 如文献[5-7]基于分类损失函数提出了不同的算法模型。不同于分类损失函数, 度量损失函数将行人重识别看作一个聚类或者排序任务。行人重识别网络中使用最广泛的度量损

失函数是三元组损失函数,其优化目标是使不同类图像之间的特征距离大于同类图像之间的特征距离。为了解决三元组损失函数中三元组数量过多的问题,Hermans等^[8]提出了一个在线挖掘困难样本三元组的方法,该方法大大减少了所需训练的三元组数量。

行人重识别算法的最理想的优化目标是最小化类内距离和最大化类间距离,然而原始的三元组损失函数只能将类间优化出来一个间隔,并未很好地优化类内距离,这使得基于原始三元组损失的全局特征行人重识别算法不能达到最优的性能。为了解决这个问题,本文改进原始的三元组损失函数。在原始三元组损失函数的基础上增加一项同类样本之间的距离损失,使改进后的算法学习出来的特征类内更紧密且类间更离散。

2 三元组损失函数及网络结构

2.1 三元组损失函数

三元组损失函数是由Schroff等^[9]在2015年提出的,其最初应用在人脸识别任务中,但人脸识别与行人重识别任务具有相似性,故也广泛应用在行人重识别任务中。三元组损失函数是一种度量损失函数,一个三元组包括一个锚点(a)、一个正样本(p)和一个负样本(n),其可以使用欧氏距离来表示样本之间的距离。三元组损失函数的学习过程如图1所示。

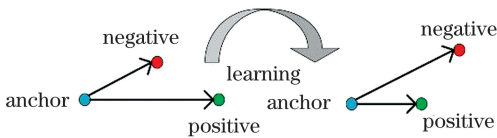


图1 三元组损失函数的学习过程

Fig. 1 Learning process of triplet loss function

在一个 d 维的欧氏距离空间 \mathbf{R}^d 中,定义 $f(x)$ 为图像 x 在 \mathbf{R}^d 空间中的映射,即 $f(x) \in \mathbf{R}^d$ 。因此,对于任意一个三元组 (x_i^a, x_i^p, x_i^n) ,其中 i 表示第 i 个三元组,总是希望 x_i^a 与 x_i^p 之间的距离比 x_i^a 与 x_i^n 之间的距离近得多,表达式为

$$\|f(x_i^a) - f(x_i^p)\|_2^2 + \alpha < \|f(x_i^a) - f(x_i^n)\|_2^2, \quad (1)$$

$$\forall \{f(x_i^a), f(x_i^p), f(x_i^n)\} \in T, \quad (2)$$

式中: α 表示一个强制间隔; T 表示训练集中所有三元组的集合。因此三元组损失函数可以表示为

$$L = \sum_i^N \max[\|f(x_i^a) - f(x_i^p)\|_2^2 - \|f(x_i^a) - f(x_i^n)\|_2^2 + \alpha, 0], \quad (3)$$

式中: N 表示所有需要训练的三元组个数。

2.2 改进的三元组损失函数

为了更好地描述三元组损失函数,使用 $d(a, p)$ 来表示 $\|f(x_i^a) - f(x_i^p)\|_2^2$,使用 $d(a, n)$ 来表示 $\|f(x_i^a) - f(x_i^n)\|_2^2$,因此可以将(3)式表示为

$$L = \sum_i^N \max[d(a, p) - d(a, n) + \alpha, 0]. \quad (4)$$

对于(4)式的原始三元组损失函数,经过样本采样的三元组损失有如下三种情况。

1)若 $d(a, p) - d(a, n) + \alpha < 0$,这属于简单样本三元组,损失为零,该类三元组不需要训练。

2)若 $d(a, p) > d(a, n)$,这属于困难样本三元组,损失不为零,该类三元组需要训练。

3)若 $d(a, p) < d(a, n) < d(a, p) + \alpha$,这属于半困难样本三元组,该类三元组需要训练。

对于所有经过训练的三元组,原始三元组都可以优化成第1)种情况。对于原始的三元组损失,满足第1)种情况的损失为零,则这个三元组就不需要训练。对于三元组损失,最理想的情况是 $d(a, p)$ 值无限接近于零, $d(a, n)$ 值无限大,即网络将相同ID的图像映射到空间中的一个点上,将不同ID的图像映射到无穷远处。分析第1)种情况,原始的三元组损失并未很好地约束相同ID图像之间的距离。对于第1)种情况,现考虑两种情形:① $d(a, p) = 10$, $d(a, n) = 11$, $\alpha = 0.3$,则有 $10 - 11 + 0.3 = -0.7 < 0$,这是满足第1)种情况;② $d(a, p) = 1$, $d(a, n) = 11$, $\alpha = 0.3$,则有 $1 - 11 + 0.3 = -9.7 < 0$,这同样满足第1)种情况。比较情形①和情形②,结果表明情形②更好,因为情形②中的 $d(a, p)$ 值比 $d(a, n)$ 值小得多,这两种情形都能使原始的三元组损失为零,并不会继续优化。针对这个问题,提出改进的三元组损失函数,表达式为

$$L = \sum_i^N \max[d(a, p) - d(a, n) + \alpha, 0] + \lambda \times d(a, p), \quad (5)$$

式中: λ 表示引入的一个参数。从(5)式可以看到,在原始三元组损失函数的基础上增加一项 $\lambda \times d(a, p)$,并引入一个参数 λ 。改进后的三元组损失函数当满足第1)种情况时,其损失值为类内距离,当不能优化类间距离时,则优化类内距离,使类内距离更小。 λ 可以控制对类内距离的优化程度, λ 值过大或者过小都可能对网络学习出来的特征造成影响。

2.3 基于全局特征的网络结构

所提模型采用与文献[10]相同的网络结构

ResNet50,并将其作为模型的骨干网络,结构如图2所示。将ResNet50中最后一层的全连接(FC)层去掉,并将ResNet50中每个块的最后一个卷积层的步长设置为1,以增大网络输出的特征图尺寸,从而获得更多的特征信息。ResNet50输出的特征图经过自适应平均池化处理得到2048维的特征;然后将其通过两个分支,一个分支是用来计算三元组损

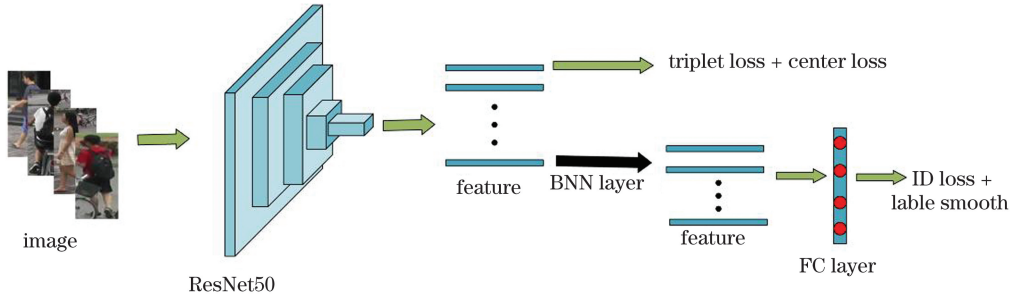


图2 网络模型的结构

Fig. 2 Structure of network model

3 实验与结果分析

3.1 数据集及评价指标

实验是在Market1501^[11]、DukeMTMC-reID^[12]和CUHK03^[13]三个主流的行人重识别数据集上进行的。分析实验结果前,首先对三个实验数据集进行介绍。

Market1501数据集是在清华大学校园内采集的,该数据集共包含1501个行人,通过6个摄像头采集而来,行人裁剪框是通过DPM(Deformable Parts Model)检测器自动检测而来的,整个数据集被分为训练集和测试集。训练集包含751个行人的12936张图片。测试集为剩下的750个行人的19732张图片,其中3368张图片作为检索图片,15913张图片作为被检索图片。根据每个行人的检索图片数量将评估方法分为两种:一种是每个行人只有一张图片用于检索;另外一种是每个行人有多张图片用于检索,实验采用的是第二种评估方法。

DukeMTMC-reID数据集是DukeMTMC数据集的子集,其是在杜克大学校园内采集的,该数据集共包含1404个行人,通过8个摄像头采集而来。不同于Market1501数据集,DukeMTMC-reID数据集的行人裁剪框是人工标注的,整个数据集被分为训练集和测试集。训练集包含702个行人,其中16522张图片作为训练数据。测试集为剩下的702个行人,其中2228张图片作为检索图片,17661张图片作为被检索图片。与Market1501数据集相

失和中心损失,另一个分支是经过一个去掉偏置的批标准化(BNN)层和一个FC层,以计算带标签平滑的交叉熵损失;最后将这三种损失值加起来,得到最终的损失。其中交叉熵损失用于分类,中心损失能够使网络学习一个类中心以优化类内距离,三元组损失可以同时优化类内和类间距离,三种损失联合使用能够更好地使网络优化类内和类间距离。

比,DukeMTMC-reID数据集对于行人重识别算法来说是一个更具有挑战性的数据集,因为其包含更多的遮挡和重叠现象。

CUHK03数据集是在香港中文大学校园内采集的,该数据集共包含1467个行人的14097张图片,每个行人均是由两个摄像头采集的。该数据集可以提供两种行人裁剪框,一种是由手工标注的,另一种是由DPM检测器自动检测的。对CUHK03数据集使用文献[14]提出的一个类似于Market1501数据集的简化评估方法,并将人工标注的数据集称为CUHK03-Labeled,检测标注的数据集称为CUHK03-Detected。

实验采用两个评价指标对模型性能进行评估,指标为人重识别算法最常用的累计匹配曲线(CMC)中的首位匹配率(rank1)和均值平均精度(mAP)。

3.2 实验环境和参数设置

实验采用PyTorch深度学习框架来搭建网络,使用配有2.20 GHz Xeon E5-2630 v4 CPU和GTX 2080TI GPU的设备来训练网络。网络使用ResNet50在ImageNet上的预训练模型进行初始化。输入图像的尺寸为256 pixel×128 pixel,数据增强的方法有随机翻转、随机裁剪和随机擦除。使用Adam优化器来优化网络参数,设置 $\epsilon = 10^{-8}$, $\beta_1 = 0.9$, $\beta_2 = 0.9$,初始学习率设置为 3.5×10^{-4} ,并采用文献[3]的学习率调整策略来调整网络的学习率,并在第40次和第70次迭代的时候将学习率

分别减小到 3.5×10^{-5} 和 3.5×10^{-6} , 每次训练共迭代 120 次。

3.3 参数 λ 对模型性能的影响

参数 λ 影响着模型对类间距离的优化结果, 为了分析不同的 λ 值对模型性能的影响, 首先选取 0.1、0.3、0.5、0.7、0.9 和 1.0 6 个值, 通过在 Market1501 数据集和 DukeMTMC-reID 数据集上的两组对比实验来说明参数 λ 对实验结果的影响。

表 1 和图 3 为在 Market1501 数据集和 DukeMTMC-reID 数据集上对所提算法进行实验得到的结果。从表 1 可以看到, 在 Market1501 数据集上, 当 $\lambda=0.1$ 和 $\lambda=0.3$ 时, 模型的性能相似, 当 $\lambda=1.0$ 时, 模型的性能最差; 在 DukeMTMC-reID 数据集上, 当 $\lambda=0.3$ 时, 模型的性能最好, 当 $\lambda=1.0$ 时, 模型的性能最差。从图 3 可以看到, 参数 λ 主要影响 mAP, 对 rank1 的影响明显小于 mAP, 对 rank1 影响较小的主要原因是当模型性能达到一定

表 1 不同参数的性能对比

Table 1 Performance comparison of different parameters

λ	Market1501		DukeMTMC-reID	
	mAP/%	rank1/%	mAP/%	rank1/%
0.1	87.4	94.8	78.2	88.0
0.3	87.5	94.7	78.4	88.2
0.5	87.1	94.7	77.8	87.7
0.7	86.8	94.7	77.3	87.8
0.9	86.4	94.6	76.7	87.6
1.0	86.3	94.7	76.2	87.5

程度时, 最容易匹配的图像已经全部被找到; 参数 λ 对模型性能的影响在两个数据集上具有一致性, 即当 $\lambda < 0.3$ 时, 模型的性能较好, 当 $\lambda > 0.3$ 时, 模型的性能随着 λ 值的增大逐渐变差。因此, 实验结果表明正确选择参数 λ 值对模型性能的提升有着很大的影响。

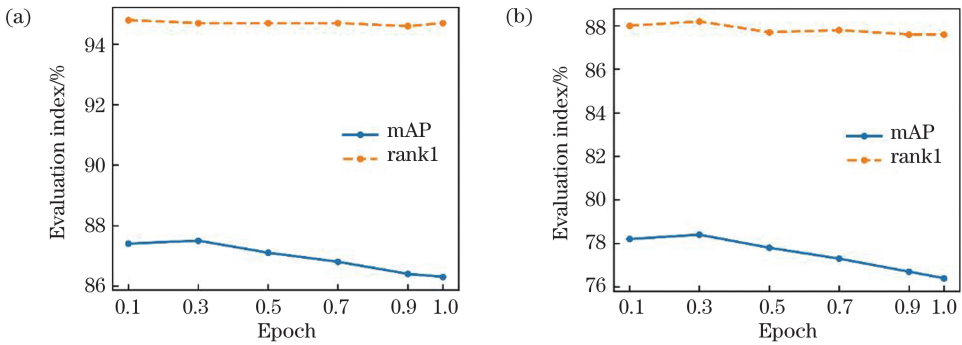


图 3 不同 λ 值在不同数据集上的性能曲线。(a) Market1501 数据集; (b) DukeMTMC-reID 数据集

Fig. 3 Performance curves for different λ values on different datasets. (a) Market1501 dataset; (b) DukeMTMC-reID dataset

3.4 与原始三元组损失比较

在 Market1501 数据集和 DukeMTMC-reID 数据集上分别采用原始三元组损失算法和改进后的算法进行训练, 实验过程中除了三元组损失不同外, 其

他设置均保持完全一致, 并联合中心损失和交叉熵损失一起训练, 每迭代 5 次计算算法的性能指标, 实验结果如图 4 所示。从图 4 可以看到, 在原始算法上, 两个数据集在迭代前 40 次的曲线都出现抖动的

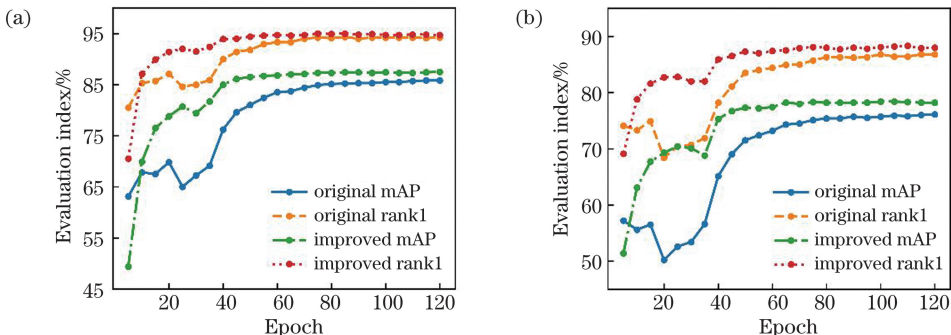


图 4 不同算法在不同数据集上的对比实验结果。(a) Market1501 数据集; (b) DukeMTMC-reID 数据集

Fig. 4 Comparative experimental results of different algorithms on different datasets. (a) Market1501 dataset; (b) DukeMTMC-reID dataset

现象,并且在迭代 80 次后才渐渐达到稳定的最佳性能;在改进后的算法上,训练曲线更加平滑且抖动较小,并且在迭代 60 次后就渐渐达到稳定的最佳性能。对比实验结果表明,改进后的算法的曲线在训练过程中抖动小,并且比原始算法能够在更快的时间内收敛到最佳值。

3.5 与其他主流算法的比较

为了进一步说明改进后算法的有效性,在

Market1501、DukeMTMC-reID 和 CUHK03 数据集上进行实验,并与基于全局特征和局部特征的两类模型进行对比,结果如表 2 和表 3 所示,其中“—”表示对应方法未在数据集上进行实验,基于局部特征的模型有 AlignedReID++^[15]、SCPNet^[16]、PCB+RPP^[17]、Pyramid^[18]、HPM^[19]和 MGN^[20],基于全局特征的模型有 IDE^[21]、SVDNet^[22]、TriNet^[9]、AWTL^[23]、Baseline^[7]、MLFN^[24]、HA-CNN^[25]和 PAN^[26]。

表 2 所提算法在 Market1501 数据集和 DukeMTMC-reID 数据集上与其他先进算法的性能对比

Table 2 Performance comparison of proposed algorithm with other advanced algorithms on Market1501 dataset and DukeMTMC-reID dataset

Algorithm	Market1501		DukeMTMC-reID	
	mAP/%	rank1/%	mAP/%	rank1/%
AlignedReID++ ^[15]	77.7	90.6	67.4	81.2
SCPNet ^[16]	75.2	91.2	62.6	80.3
PCB+RPP ^[17]	81.6	93.8	69.2	83.3
Pyramid ^[18]	88.2	95.7	79.0	89.9
HPM ^[19]	82.7	94.2	74.3	86.6
MGN ^[20]	86.9	95.7	78.4	88.7
IDE ^[21]	59.9	79.5	—	—
SVDNet ^[22]	62.1	82.3	56.8	76.7
TriNet ^[9]	69.1	84.9	—	—
AWTL ^[23]	75.7	89.5	63.4	79.8
Baseline ^[7]	85.9	94.5	76.4	86.4
Proposed algorithm	87.5	94.7	78.4	88.2

表 3 所提算法在 CUHK03 数据集上与其他先进算法的性能对比

Table 3 Performance comparison of proposed algorithm with other advanced algorithms on CUHK03 dataset

Algorithm	CUHK03-Labeled		CUHK03-Detected	
	mAP/%	rank1/%	mAP/%	rank1/%
MGN ^[20]	67.4	68.0	66.0	68.0
PCB+RPP ^[17]	—	—	57.5	63.7
MLFN ^[24]	49.2	54.7	47.8	52.8
HA-CNN ^[25]	41.0	44.4	38.6	41.7
SVDNet ^[22]	37.8	40.9	37.3	41.5
PAN ^[26]	35.0	36.9	34.0	36.3
IDE ^[21]	21.0	22.2	19.7	21.3
Proposed algorithm	67.3	69.9	63.8	65.4

从表 2 和表 3 可以看到,与基于全局特征的行人重识别算法相比,所提算法在三个数据集上都取得最优的性能;与基于局部特征的行人重识别算法

相比,在 Market1501 数据集上,所提算法的性能好于 MGN 算法^[20],略差于 Pyramid 算法^[18];在 DukeMTMC-reID 数据集上,所提算法的性能与

MGN 算法^[20]相似,略差于 Pyramid 算法^[18];在 CUHK03-Labeled 数据集上,所提算法的性能接近 MGN 算法^[20],但是在 CUHK03-Detected 数据集上,所提算法差于 MGN 算法^[20],原因在于 CUHK03-Detected 数据集是自动标注的,比 CUHK03-Labeled 手工标注有更多的遮挡可折叠现象,这对于算法来说更具有挑战性。

4 结 论

提出一种基于全局特征的行人重识别改进算法,其是在三元组损失函数的基础上进行改进,即在三元组损失函数中引入一项正样本之间的距离损失,使得改进后的算法能够更好地优化样本之间的距离,学习出来的特征更具有判别性,从而可以进一步提高重识别的精度。在 Market1501、DukeMTMC-reID 和 CUHK03 数据集上对所提算法进行大量实验。实验结果表明,改进算法在基于全局特征的模型中可以取得最优的性能。

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] Szegedy C, Liu W, Jia Y, et al. Going deeper with convolutions [C] // 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 7-12, 2015, Boston, MA, USA. New York: IEEE, 2015: 15523970.
- [3] 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.
- [4] Huang G, Liu Z, van der Maaten L, et al. Densely connected convolutional networks [C] // 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), July 21-26, 2017, Honolulu, HI, USA. New York: IEEE, 2017: 2261-2269.
- [5] Bi X J, Wang H. Person re-identification based on view information embedding[J]. *Acta Optica Sinica*, 2019, 39(6): 0615007.
毕晓君, 汪灏. 基于视角信息嵌入的行人重识别[J]. *光学学报*, 2019, 39(6): 0615007.
- [6] Pan T, Li W G. Person re-identification based on feature stitching [J]. *Laser & Optoelectronics Progress*, 2019, 56(16): 162001.
潘通, 李文国. 基于特征拼接的行人重识别[J]. *激光与光电子学进展*, 2019, 56(16): 162001.
- [7] Xu L Z, Peng L. Person reidentification based on multiscale convolutional feature fusion[J]. *Laser & Optoelectronics Progress*, 2019, 56(14): 141504.
徐龙壮, 彭力. 基于多尺度卷积特征融合的行人重识别[J]. *激光与光电子学进展*, 2019, 56(14): 141504.
- [8] Hermans A, Beyer L, Leibe B. In defense of the triplet loss for person re-identification [EB/OL]. (2017-11-21) [2020-05-25]. <https://arxiv.org/abs/1703.07737>.
- [9] Schroff F, Kalenichenko D, Philbin J. FaceNet: a unified embedding for face recognition and clustering [C] // 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 7-12, 2015, Boston, MA, USA. New York: IEEE, 2015: 815-823.
- [10] Luo H, Jiang W, Gu Y Z, et al. A strong baseline and batch normalization neck for deep person re-identification[J]. *IEEE Transactions on Multimedia*, 2020, 22(10): 2597-2609.
- [11] 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.
- [12] Ristani E, Solera F, Zou R, et al. Performance measures and a data set for multi-target, multi-camera tracking[M] // Hua G, Jégou H. *Computer vision-ECCV 2016. Lecture notes in computer science*. Cham: Springer, 2016, 9914: 17-35.
- [13] Li W, Zhao R, Xiao T, et al. DeepReID: deep filter pairing neural network for person re-identification[C] // 2014 IEEE Conference on Computer Vision and Pattern Recognition, June 23-28, 2014, Columbus, OH, USA. New York: IEEE, 2014: 152-159.
- [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] Luo H, Jiang W, Zhang X, et al. AlignedReID++: dynamically matching local information for person re-

- identification[J]. *Pattern Recognition*, 2019, 94: 53-61.
- [16] Fan X, Luo H, Zhang X, et al. SCPNet: spatial-channel parallelism network for joint holistic and partial person re-identification[M]//Jawahar C, Li H, Mori G, et al. *Computer vision-ACCV 2018. Lecture notes in computer science*. Cham: Springer, 2019, 11362: 19-34.
- [17] Sun Y F, Zheng L, Yang Y, et al. Beyond part models: person retrieval with refined part pooling (and a strong convolutional baseline)[M]//Ferrari V, Hebert M, Sminchisescu C, et al. *Computer vision-ECCV 2018. Lecture notes in computer science*. Cham: Springer, 2018, 11208: 501-518.
- [18] Zheng F, Deng C, Sun X, et al. Pyramidal person re-identification via multi-loss dynamic training[C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 15-20, 2019, Long Beach, CA, USA. New York: IEEE, 2019: 8506-8514.
- [19] Fu Y, Wei Y, Zhou Y, et al. Horizontal pyramid matching for person reidentification[C]//The Thirty-Third AAAI Conference on Artificial Intelligence (AAAI-19), January 27-February 1, 2019, Honolulu, Hawaii, USA. California: AAAI, 2019: 8295-8302.
- [20] Wang G S, Yuan Y F, Chen X, et al. Learning discriminative features with multiple granularities for person re-identification[C]//2018 ACM Multimedia Conference on Multimedia Conference-MM'18, October 15-26, 2018. Seoul, Republic of Korea. New York: ACM, 2018: 274-282.
- [21] Zheng Z D, Zheng L, Yang Y. A discriminatively learned CNN embedding for person reidentification[J]. *ACM Transactions on Multimedia Computing, Communications, and Applications*, 2017, 14(1): 13.
- [22] Sun Y F, Zheng L, Deng W J, et al. SVDNet for pedestrian retrieval[C]//2017 IEEE International Conference on Computer Vision (ICCV), October 22-29, 2017, Venice, Italy. New York: IEEE, 2017: 3820-3828.
- [23] Ristani E, Tomasi C. Features for multi-target multi-camera tracking and reidentification[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, June 18-23, 2018, Salt Lake City, UT, USA. New York: IEEE, 2018: 6036-6046.
- [24] Chang X B, Hospedales T M, Xiang T. Multi-level factorisation net for person re-identification[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, June 18-23, 2018, Salt Lake City, UT, USA. New York: IEEE, 2018: 2109-2118.
- [25] Li W, Zhu X, Gong S. Harmonious attention network for person re-identification[C]//IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 18-22, 2018, Salt Lake City, UT, USA. New York: IEEE, 2018: 2285-2294.
- [26] 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.