

# 面向行人重识别的多域批归一化问题研究

张誉馨<sup>1</sup>, 张索非<sup>2</sup>, 王文龙<sup>1</sup>, 吴晓富<sup>1</sup>

(1. 南京邮电大学 通信与信息工程学院, 江苏 南京 210003;

2. 南京邮电大学 物联网学院, 江苏 南京 210003)

**摘要:**近年来基于深度神经网络的行人重识别算法取得了长足的进步,被广泛应用于网络中的批归一化(batch normalization)模块发挥着重要作用。批归一化模块在多数情况下可有效提高网络收敛速度和训练稳定性,然而当多个独立标注的数据库混合在一块进行跨域或者多域训练时,数据之间的分布差异使得目前的批归一化算法工作逻辑存疑。由于不同批次下训练数据的分布差异较大,归一化过程中的统计参数不稳定导致批归一化效果恶化。该文聚焦于多数据集合并下的行人重识别模型训练问题,通过对多数据集分布差异导致的多域模型批归一化存在的问题进行分析。然后针对模型批量归一化算法面对的多域差异,提出了一种解决策略,在多个数据集并行训练下提高了模型的泛化能力。实验结果表明:所提出的多域归一化方法在多域训练下能有效提高模型最终的泛化能力,获得更高的识别准确度,并且可应用于其他行人重识别网络以进一步提升模型性能。

**关键词:**计算机视觉;深度学习;行人重识别;多域训练;批归一化

中图分类号:TP391.41

文献标识码:A

文章编号:1673-629X(2022)01-0091-07

doi:10.3969/j.issn.1673-629X.2022.01.016

## Research on Batch Normalization for Multi-domain Person Re-identification

ZHANG Yu-xin<sup>1</sup>, ZHANG Suo-fei<sup>2</sup>, WANG Wen-long<sup>1</sup>, WU Xiao-fu<sup>1</sup>

(1. School of Telecommunications & Information Engineering, Nanjing University of

Posts and Telecommunications, Nanjing 210003, China;

2. School of Internet of Things, Nanjing University of Posts and Telecommunications, Nanjing 210003, China)

**Abstract:** In recent years, the pedestrian re-identification (Re-ID) algorithms based on deep neural networks (DNNs) have made considerable progress. The use of batch normalization modules in constructing DNNs is key to the success of deep Re-ID, which can effectively improve the convergence speed and training stability in most cases. However, when multiple independently labeled datasets are mixed together for cross-domain or multi-domain training, the domain gap makes the use of the current batch normalization module questionable. The distribution difference among training batches may result in unstable estimates for various statistical parameters, which may lead to possible performance deterioration for the employed batch normalization. We focus on the training of pedestrian re-recognition models under the combination of multiple data sets and analyze the problems in the batch normalization of multi-domain models caused by the distribution differences of multiple data sets. Aiming at the multi-domain difference of the batch model normalization algorithm, a strategy is proposed to improve the generalization ability of the model under the parallel training of multiple data sets. Experiment shows that the proposed domain-specific batch normalization algorithm can effectively improve the final generalization ability under multi-domain training and obtain higher recognition accuracy. Furthermore, it can be applied to existing Re-ID networks for further improving their performance.

**Key words:** computer vision; deep learning; pedestrian Re-ID; multi-domain training; batch normalization

## 0 引言

行人重识别(pedestrian re-identification, Re-ID)<sup>[1-3]</sup>是利用计算机视觉<sup>[4-5]</sup>技术判断图像或者视频

序列中是否存在特定行人的技术。给定一个待检行人图像,Re-ID任务一般要求检索出跨监控摄像头下的该行人图像。随着安防需求的不断增长以及计算机视

收稿日期:2021-02-04

修回日期:2021-06-08

基金项目:国家自然科学基金(61701252)

作者简介:张誉馨(1996-),男,硕士研究生,研究方向为智能信号处理、机器学习、计算机视觉;张索非,博士,讲师,研究方向为图像与视频信号处理、机器学习、物联网技术等;吴晓富,博士,教授,研究方向为信息论与编码、机器学习与计算机视觉、密码学与信息安全。

觉技术的迅猛发展,行人重识别技术已成为当前计算视觉研究的一个热点方向。

传统的行人重识别主要关注设计手工视觉特征<sup>[6]</sup>与度量相似性两个问题。随着深度学习的发展,用深度神经网络来有效提取行人重识别特征<sup>[7]</sup>的技术取得了飞速发展。文献[8]提出一种基于分块的卷积基准网络(part based convolutional baseline, PCB),通过将行人均等分为六分块引导网络更加关注局部粒度信息;文献[9]中提出的(batch drop block, BDB)算法,通过对数据随机局部遮挡,使网络获得更为全面的空间特征;文献[10]提出一种将判别信息与各粒度相结合的多粒度网络(multiple granularity network, MGN)。这些算法可以有效提升单一数据集下训练的模型泛化能力,但是对于现实中的复杂场景,往往训练集中包含两种甚至更多种风格的图像(例如不同背景和不同光照下拍摄的照片)。当数据集中出现明显的风格分歧时,大部分算法并没有基于多域数据的风格差异问题进行训练优化。

如何处理训练数据之间的分布差异并利用多个不同风格的数据集来训练提升模型的泛化能力是行人重识别领域需要认真考虑的一个问题。利用多个风格不同的数据集来训练深度神经网络的一个主要障碍来自于深度神经网络模型中广泛应用的批归一化(batch normalization, BN)模块。BN模块最初由Sergey等提出<sup>[11]</sup>,用于有效控制层间信号的取值范围,其已被证明能大幅提高深度神经网络的训练收敛以及最终性能。然而,BN模块应用于多域数据却存在先天的逻辑缺陷。例如,当一个批次的训练数据中呈现两种不同分布时,以该批次计算的均值与方差会大幅震荡,导致训练过程中归一化效果恶化。因此,多域行人重识别面临的一个重要问题是如何解决多域数据分布差异导致的BN模块设计问题。

为讨论行人重识别使用多个风格数据库的多域训练问题,该文尝试将两个典型的行人数据集Market1501<sup>[12]</sup>和DukeMTMC-reID<sup>[13]</sup>进行合并训练,以此分析多域批归一化问题并提供解决问题的思路。通过将不同数据集下的图片进行分域归一化处理,从而规避了数据分布差异导致的批归一化应用困境,基于此思路该文尝试了一种并行训练模式下的多域归一化方案。实验分析表明:采用该方案可以有效提升模型的泛化能力,在Market1501和DukeMTMC-reID数据集上获得了明显的性能提升。

## 1 问题研究

### 1.1 BN算法

深度神经网络的BN层原理如下:设一个训练批

次的输入数据为 $B = \{x_{1,2}, \dots, x_m\}$ ,输出为 $Y = \{y_{1,2}, \dots, y_m\}$ ,则BN层进行归一化处理的关键是计算这一批次样本的均值和方差。

$$\mu_B = \frac{1}{m} \sum_{i=1}^m x_i \quad (1)$$

$$\sigma_B^2 = \frac{1}{m} \sum_{i=1}^m (x_i - \mu_B)^2 \quad (2)$$

根据该批次样本统计得到的均值和方差对数据进行归一化后再缩放和平移,具体过程为:

$$\hat{x}_i \leftarrow \frac{x_i - \mu_B}{\sqrt{\sigma_B^2 + \varepsilon}} \quad (3)$$

$$y_i \leftarrow \gamma \hat{x}_i + \beta \equiv \text{BN}(x_i) \quad (4)$$

其中, $\gamma$ 和 $\beta$ 分别为缩放和平移参数。由式(1)和式(2)可知:当训练集中出现两种甚至多种呈现明显分布差异的数据时,短时间窗口所得到的均值与方差必然不稳定,表现为不同训练批次计算出的均值和方差存在振荡现象,导致模型的归一化效果恶化。

### 1.2 多域归一化问题

多域数据分布差异广泛存在于行人重识别研究中,例如不同摄像头<sup>[14]</sup>、不同光照下采集得到的行人数据集。鉴于目前公开的不同行人数据集的固有分布差异,该文以多个行人数据集合并训练存在的多域归一化问题作为主要研究案例。以Market1501和DukeMTMC-reID为例,Market1501于2015年夏采集自清华大学校园内,包含1501个行人ID,共有32668张图片;DukeMTMC-reID于2015年冬采集自杜克大学校园内,包含1812个行人ID,共有36411张图片。由于数据采集的地理位置和所处的季节不同,这两种行人数据集在行人着装、背景颜色、镜头风格方面都存在明显差异。图1分别展示了来自Market1501和DukeMTMC-reID的图片样例。



图1 图片样例

通过对比图片样例可发现两种数据集中采集的行人图片存在显著差异,例如在Market1501中行人多为身着夏装的亚洲人群,而DukeMTMC中多为身着冬装的欧美人群。在模型训练中,图片的差异将导致模型BN层中的统计值也呈现出两种不同的分布,图2

中展示了采用 ResNet50<sup>[15]</sup> 模型分别在 Market1501 和 DukeMTMC-reID 训练下各 BN 层的均值分布差异。

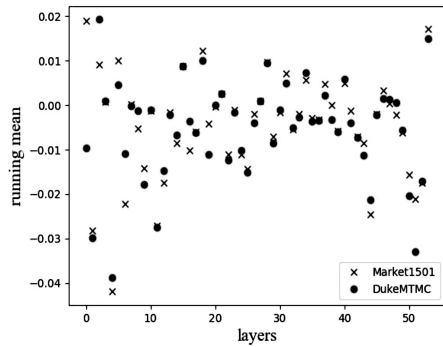


图 2 Market1501 和 DukeMTMC-reID 训练下的 BN 层均值分布差异

为评估由于数据集之间的分布差异导致的多域归

表 1 Market1501 与 DukeMTMC-reID 数据集交叉测试结果

%

训练集	测试集			
	Market1501		DukeMTMC-reID	
	Rank-1	mAP	Rank-1	mAP
Market1501	94.5	85.4	27.4	15.0
DukeMTMC-reID	50.3	21.6	85.9	74.7
Market+DukeMTMC	93.7	84.2	85.7	74.6

## 2 文中方案

### 2.1 多域归一化方法

针对多数据集分布差异导致的多域归一化问题,该文尝试一种简单的解决方案。在模型的批归一化处理模块中,将数据流根据所属域进行分离,对不同域下

一化问题的影响,将两种不同的行人数据集进行合并,再用于训练行人重识别模型,最终将合并训练的模型与单一数据集训练的模型进行性能比较。表 1 展示了使用 ResNet50 网络作为基线模型训练后在 Market1501 和 DukeMTMC-reID 测试集上进行交叉测试的性能对比。

将多数据集进行合并训练在理论上可提升模型的整体泛化性能,但表 1 中的数据显示将 Market1501 和 DukeMTMC 合并后训练的模型性能却在首位配准率 (Rank-1) 和平均精度均值 (mean average precision, mAP)<sup>[16]</sup> 两项测试指标下均弱于单一数据集下训练的结果。这表明在多个数据库合并训练下,多个数据集之间的分布差异导致的多域归一化问题很大程度影响了模型的训练效果,造成模型最终的性能下降。

的数据进行单独的归一化,以此有效缓解归一化过程中出现的统计值波动困扰。根据此思路,提出一种针对特定域数据的批归一化方法 (domain-specific batch normalization, DSBN),图 3 展示了 DSBN 模块的基本流程。

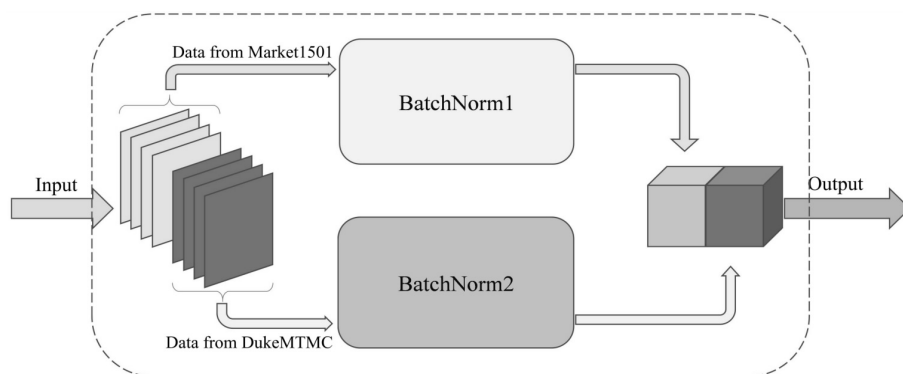


图 3 DSBN 模块数据处理流程

如图 3 所示,DSBN 模块的每个训练批次是由两个子批次 (mini-batch) 组成,每个子批次的样本分别取自于两种不同的数据库。一个完整的训练批次 (super-batch) 可表示为  $B_{\text{super}} = 2 \times B_{\text{mini}}$ ,子批次样本分别送入两个独立的 BN 模块中进行批归一化处理。为使子批次中的图片来自不同数据库,该方案采用多数据集并行训练方式,整体的训练流程如图 4 所示。

使用 ResNet50 作为主干网络,并将其中所有的

BN 层替换为 DSBN 模块,其余层的参数设置不变。在并行训练中,模型中所有的归一化层对来自不同数据集的样本进行独立的归一化,同时模型中其他层的参数则共享两个数据集的信息。

采用此方案的优势是避免了由于多域数据分布差异导致的归一化问题,同时可以有效利用多数据集训练来提升模型的整体泛化能力。

但是在处理数据集更大、面对更复杂的多域问题



时,该方案的实现复杂度相对较高,如何改进还有待进一步研究。

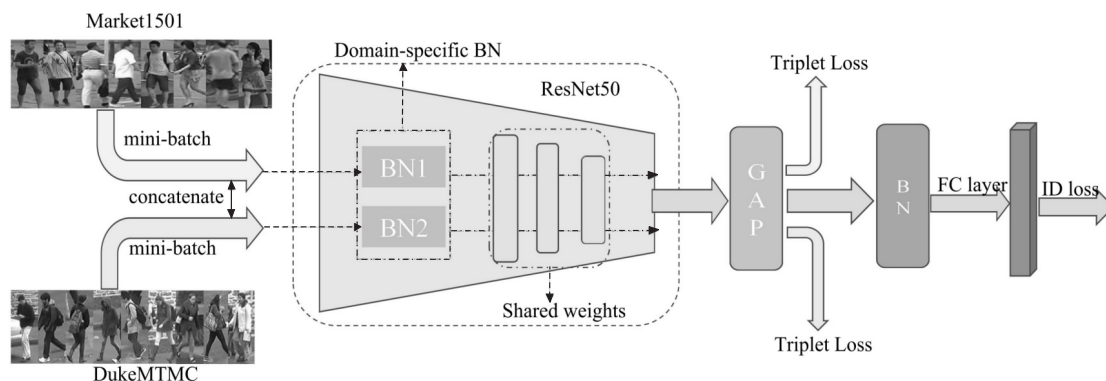


图 4 多数据集并行训练流程

## 2.2 损失函数设置

采用了两种损失函数对网络进行优化,分别为交叉熵损失函数(cross entropy loss)<sup>[17]</sup>和三元组损失函数(triplet loss)<sup>[18]</sup>,交叉熵函数用于计算行人 ID 的分类误差,记为 ID loss;triplet loss 优化特征之间的度量距离,使正样本之间的间距更近,负样本之间的间距更远。

具体而言,该方案主干网络输出的特征图经过全局均值池化层(global average pooling, GAP)<sup>[19]</sup>后获得通道维数为 2 048 的全局特征向量,分别计算各子批次的 triplet loss,即:

$$L_{Tri}^i = [d_p^i - d_n^i + \alpha]_+, i \in \text{domain}\{A, B\} \quad (5)$$

其中,  $d_p$  和  $d_n$  分别表示正样本对和负样本对的特征距离,  $\{A, B\}$  为数据所属的两个不同域,  $[z]_+$  表示  $\max(z, 0)$ ,  $\alpha$  表示距离余量,该方案中将其设为 0.3。经过 GAP 得到的特征向量依次经过 BNNeck<sup>[20]</sup>和全连接层(fully connected layer, FC)后输出行人的分类预测,总类别为两个数据集的行人 ID 数量总和。采用交叉熵损失函数直接计算总的 ID loss,表示为:

$$L_{ID} = - \sum_{i=1}^N q_i \log(p_i) \quad (6)$$

其中,  $N$  表示总的行人 ID 数目,  $p_i$  表示 FC 层输出的对应各 ID 的预测概率值,当类别预测与行人实际标签相符时  $q_i = 1$ ,否则  $q_i = 0$ 。最终的 loss 表示为:

$$L = L_{ID} + L_{Tri}^A + L_{Tri}^B \quad (7)$$

## 3 实验及分析

### 3.1 实验环境及参数配置

实验所用的行人数据集为 Market1501 和 DukeMTMC-reID。在并行训练方案中,每个子批次(mini-batch)的 batch size 设为 64,合并后的 super-batch 大小为 128。由于两种数据集的大小存在差异,因此以其中图片数量较多的数据集完全训练一次作为一个 epoch,数量较少的数据集存在少量图片的重复

遍历。

在训练策略上,采用 Adam<sup>[21]</sup>优化算法,一阶动量系数设置为 0.9,二阶动量系数设为 0.999,权重衰减系数为 0.000 5。初始学习率为 0.000 035。采用的 WarmUp<sup>[22]</sup>策略——经过 20 个 epoch 将学习率从 0.0 线性提升到 0.000 35;再每隔 30 个 epoch 将学习率降低为 0.2 倍,共训练 140 个 epoch。

在图片预处理上,对两个域的数据均采用了随机擦除(random erasing)<sup>[23]</sup>和随机翻转(random flip)<sup>[24]</sup>算法进行处理,并且将输入图像的尺寸统一设置宽 128 像素、高 256 像素。

该方案采用 ResNet50 作为基础主干网络,将其中的 BN 层替换为 DSBN 层,并且载入了 ImageNet<sup>[25]</sup>预训练参数。训练过程中同时对 DSBN 中的两套归一化模块参数进行优化,在测试过程中,模型载入对应数据集的归一化参数对各个测试集进行单独性能测试。

实验测试集包含待检索图片集 Query 和检索图片库 Gallery 两部分。在测试过程中,先通过网络推理得到测试集所有测试图片对应的特征,然后根据特征间的欧氏距离计算出待检索图 Query 与检索库 Gallery 中所有图片的相似度并进行排序,最终计算首位准确度(Rank-1)和平均精度均值(mean average precision, mAP)作为模型性能评价指标。其中首位准确度主要用于模型对简单样本配准性能的评估,而平均精度均值则兼顾了模型对困难样本配准性能的评估。

实验运行环境为 Ubuntu16.04 操作系统,使用 Pytorch1.1 深度学习框架实现,并在 NVIDIA Tesla P40 GPU 上进行模型训练。

### 3.2 实验结果对比

将采用 DSBN 模块的模型性能与优化之前的结果进行比较,所对比的网络结构区别仅在于改变了其中一个网络的归一化模块,所采用的训练方式包括单一数据集训练和多数据集并行方式两种。对比结果如

表2所示。

表2 采用DSBN的并行训练与普通训练方式性能对比 %

训练集/训练方式	归一化模块	测试集			
		Market1501		DukeMTMC-reID	
		Rank-1	mAP	Rank-1	mAP
Market1501	BN	94.4	85.4	27.4	15.0
DukeMTMC	BN	50.3	21.6	85.9	74.7
(Market + DukeMTMC)(混合)	BN	93.7	84.2	85.7	74.6
(Market + DukeMTMC)(并行)	BN	94.3	86.6	86.4	75.5
(Market + DukeMTMC)(并行)	DSBN(ours)	95.0	87.9	88.2	76.9

由表2可知:相较于单一数据集下的性能,最终采用DSBN的并行训练方案在两个数据集上分别在Rank-1和mAP上提高了0.6、2.5、2.3、2.2个百分点。由于并行训练模式下同批次中两种数据集的图片数量相同,训练时BN层的输入均值、方差也相对更均衡,因此采用并行训练方式得到的模型性能优于多数

数据集混合后训练所得性能。

DSBN模块针对两个域的数据分别进行归一化,因此DSBN中的两个BN模块分别对应了两种数据集。使用DSBN模块中两套对应不同数据集的BN参数分别在两个数据集上进行测试,结果如表3所示。

表3 DSBN中对应不同数据集的BN参数测试结果 %

训练集/训练方式	归一化模块	测试集			
		Market1501		DukeMTMC-reID	
		Rank-1	mAP	Rank-1	mAP
(Market + DukeMTMC)(并行)	BN(Market1501)	95.0	87.9	51.4	33.4
(Market + DukeMTMC)(并行)	BN(DukeMTMC)	67.6	40.7	88.2	76.9

在测试过程中采用与测试集相对应的BN模块参数时,模型识别的准确率较高;而使用对应另一测试集的BN参数则只能得到较低的识别准确率。这表明DSBN模块中的各BN只适于与其对应的一个数据域,从而避免了多域数据输入到同一BN模块中造成的归一化问题。表3中的对比数据同时也印证了前文提出的多数据域之间同步分布差异问题的存在,DSBN模块可在一定程度上缓解其带来的性能问题。

同时考察了在两种数据集上采用增量学习模式进行训练的效果。增量学习<sup>[26]</sup>模式是指模型依次在多个数据集上进行训练,使网络在不断增加的训练数据集上持续学习以提升最终性能。这里采用原始ResNet50模型依次在两种数据集上训练,采用的训练配置与文中方案保持完全一致,并和提出的并行学习方式进行比较,得出的测试结果如表4所示。

表4 与增量学习方案对比结果 %

训练集/训练方式	训练总时长	测试集			
		Market1501		DukeMTMC-reID	
		Rank-1	mAP	Rank-1	mAP
Market(先) + DukeMTMC(后)	约8h	54.7	27.1	84.5	73.4
Market(后) + DukeMTMC(前)	约8h	94.1	84.5	33.5	20.0
(Market + DukeMTMC)(并行)	约6h	95.0	87.9	88.2	76.9

由对比可以发现:该文所采用的DSBN方案在准确率上优于增量学习方案,分别在两个数据集Rank-1和mAP指标上相较于增量学习方案提高了0.9、3.4、3.7、3.5个百分点,并且在总训练时长上相较增量学习方案也具有一定优势。

### 3.3 与其他算法比较

将文中方法在Market1501和DukeMTMC上与其他行人重识别算法进行比较,并尝试将文中方案迁移到其他公开报道的算法之上,提升其在多数据训练下的性能。

实验对比结果如表 5 所示,其中 Ours + MGN、法相结合后得到的算法。  
Ours+PLR-OSNet 为文中方法与 MGNPLR-OSNet 算

表 5 与其他算法性能比较 %

Method	测试集			
	Market1501		DukeMTMC-reID	
	Rank-1	mAP	Rank-1	mAP
PCB+RPP <sup>[8]</sup> (2018)	93.8	81.6	83.3	69.2
OSNet <sup>[27]</sup> (2019)	94.8	84.9	88.6	73.5
BOT <sup>[20]</sup> (2019)	94.5	85.9	86.4	76.4
MGN * <sup>[10]</sup> (2019)	94.9	86.1	88.2	77.5
MHN <sup>[28]</sup> (2019)	95.1	85.0	89.1	77.2
AANet <sup>[29]</sup> (2019)	93.9	83.4	87.7	74.3
SNR <sup>[30]</sup> (2020)	94.4	84.7	84.4	72.9
PLR-OSNet * <sup>[31]</sup> (2020)	95.5	88.9	89.4	80.2
Baseline	94.5	85.4	85.9	74.7
Ours	95.0	87.9	88.2	76.9
Ours+MGN	95.4	86.9	88.6	77.9
Ours+PLR-OSNet	95.6	89.2	89.8	80.3

表 5 中标有 \* 的结果为实际复现结果。由于本方案采用的网络架构 (Ours) 仅利用了单分支网络的全局特征,因此性能仍有较大的提升空间,相较于未采用文中方案的基线模型 (Baseline) 有可观的性能提升,且与其他采用单分支网络的算法例如 OSNet<sup>[27]</sup>、BOT<sup>[20]</sup> 相比也具有一定优势。另外,Ours+MGN、Ours+PLR-OSNet 的结果表明该方案可应用于已报道的其他行人重识别网络 (如 MGN、PLR-OSNet) 并进一步提升其性能。

#### 4 结束语

该文聚焦于多数据集合并下的行人重识别模型训练问题。针对多域混合训练,重点分析了多域数据分布差异引起的模型批归一化模块的不稳定性问题。在此基础上,提出了一种 DSBN 模块来替代传统的批归一化模块。初步实验表明:所提方案可在多数据集合并训练下有效缓解多域归一化的波动问题,提升模型的泛化能力。当然,该方案只是多域训练方向的一个简单尝试,多域训练问题的彻底解决还有待更为深入和细致的研究。

#### 参考文献:

- [1] 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] 郑伟诗,吴岸聪.非对称行人重识别:跨摄像机持续行人追踪[J].中国科学:信息科学,2018,48:545-563.
- [3] 叶钰,王正,梁超,等.多源数据行人重识别研究综述[J].自动化学报,2020,46(9):1869-1884.

- [4] 张伟,钱云涛.引入全局约束的精简人脸关键点检测网络[J].信号处理,2019,35(3):507-515.
- [5] 范春奇,任坤,孟丽莎,等.基于深度学习的数字图像修复算法最新进展[J].信号处理,2020,36(1):102-109.
- [6] DALAL N, TRIGGS B. Histograms of oriented gradients for human detection[C]//2005 IEEE computer society conference on computer vision and pattern recognition (CVPR'05). San Diego, CA:IEEE,2005:886-893.
- [7] 赵师亮,吴晓富,张索非.基于 PCB 特征加权的行人重识别算法[J].信号处理,2020,36(8):1300-1307.
- [8] SUN Y, ZHENG L, YANG Y, et al. Beyond part models: person retrieval with refined part pooling (and a strong convolutional baseline)[C]//Computer vision - ECCV 2018. Munich, Germany:Springer,2018:510-518.
- [9] DAI Z, CHEN M, GU X, et al. Batch DropBlock network for person re-identification and beyond[C]//Proceedings of the IEEE international conference on computer vision. Los Alamitos, CA:IEEE,2019:3691-3701.
- [10] WANG G, YUAN Y, CHEN X, et al. Learning discriminative features with multiple granularities for person re-identification[C]//Proceedings of the 26th ACM international conference on Multimedia. New York, NY:ACM,2018:274-282.
- [11] IOFFE S, SZEGEDY C. Batch normalization: accelerating deep network training by reducing internal covariate shift[C]//Proceedings of the 32nd international conference on machine learning. Lille, France:JMLR.org,2015:448-456.
- [12] ZHENG L, SHEN L, TIAN L, et al. Scalable person re-identification:a benchmark[C]//Proceedings of the IEEE international conference on computer vision. Santiago, Chile:IEEE,2015:1116-1124.

- [13] RISTANI E, SOLERA F, ZOU R, et al. Performance measures and a data set for multi-target, multi-camera tracking [C]//Computer vision - ECCV 2016. Amsterdam, The Netherlands: Springer, 2016: 17-35.
- [14] ZHUANG Z, WEI L, XIE L, et al. Rethinking the distribution gap of person re-identification with camera-based batch normalization [C]//Computer vision - ECCV 2020. Glasgow, UK: Springer, 2020: 140-157.
- [15] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition [C]//Proceedings of the IEEE conference on computer vision and pattern recognition. Las Vegas, NV, USA: IEEE, 2016: 770-778.
- [16] ZHONG Z, ZHENG L, CAO D, et al. Re-ranking person re-identification with k-reciprocal encoding [C]//Proceedings of the IEEE conference on computer vision and pattern recognition. Honolulu, HI, USA: IEEE, 2017: 3652-3661.
- [17] ZHANG Z, SABUNCU M. Generalized cross entropy loss for training deep neural networks with noisy labels [J]. Advances in Neural Information Processing Systems, 2018, 31: 8778-8788.
- [18] CHENG D, GONG Y, ZHOU S, et al. Person re-identification by multi-channel parts-based CNN with improved triplet loss function [C]//Proceedings of the IEEE conference on computer vision and pattern recognition. Las Vegas, NV, USA: IEEE, 2016: 1335-1344.
- [19] LIN M, CHEN Q, YAN S. Network in network [J]. arXiv: 1312. 4400, 2013.
- [20] LUO H, GU Y, LIAO X, et al. Bag of tricks and a strong baseline for deep person re-identification [J]. arXiv: 1903. 07071, 2019.
- [21] KINGMA D P, BA J. Adam: a method for stochastic optimization [J]. arXiv: 1412. 6980, 2014.
- [22] FAN X, JIANG W, LUO H, et al. Spheredid: deep hypersphere manifold embedding for person re-identification [J]. Journal of Visual Communication and Image Representation, 2019, 60: 51-58.
- [23] ZHONG Z, ZHENG L, KANG G, et al. Random erasing data augmentation [C]//Proceedings of the AAAI conference on artificial intelligence. Palo Alto, CA: AAAI Press, 2020: 13001-13008.
- [24] ZHONG Z, ZHENG L, ZHENG Z, et al. Camstyle: a novel data augmentation method for person re-identification [J]. IEEE Transactions on Image Processing, 2018, 28(3): 1176-1190.
- [25] DENG J, DONG W, SOCHER R, et al. Imagenet: a large-scale hierarchical image database [C]//2009 IEEE conference on computer vision and pattern recognition. Miami, FL: IEEE, 2009: 248-255.
- [26] GIRAUD-CARRIER C. A note on the utility of incremental learning [J]. AI Communications, 2000, 13(4): 215-223.
- [27] ZHOU K, YANG Y, CAVALLARO A, et al. Omni-scale feature learning for person re-identification [C]//Proceedings of the IEEE international conference on computer vision. Seoul (KR): IEEE, 2019: 3701-3711.
- [28] CHEN B, DENG W, HU J. Mixed high-order attention network for person re-identification [C]//Proceedings of the IEEE international conference on computer vision. Seoul (KR): IEEE, 2019: 371-381.
- [29] TAY C P, ROY S, YAP K H. Aanet: attribute attention network for person re-identifications [C]//Proceedings of the IEEE conference on computer vision and pattern recognition. [s. l.]: IEEE, 2019: 7134-7143.
- [30] JIN X, LAN C, ZENG W, et al. Style normalization and restitution for generalizable person re-identification [C]//Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. [s. l.]: IEEE, 2020: 3143-3152.
- [31] XIE B, WU X, ZHANG S, et al. Learning diverse features with part-level resolution for person re-identification [J]. arXiv: 2001. 07442, 2020.