

doi:10.14132/j.cnki.1673-5439.2021.01.010

面向行人重识别分集特征挖掘的 多分支深度神经网络构造进展

吴晓富¹, 尹梓睿¹, 宋越², 张磊¹, 谢奔¹, 赵师亮¹, 张索非³

(1.南京邮电大学 通信与信息工程学院, 江苏 南京 210003)
2.95958 部队, 上海 200120
3.南京邮电大学 物联网学院, 江苏 南京 210003)

摘要:随着深度学习的兴起,近年来行人重识别性能提高的关键是如何构建具有分集特征挖掘的深度神经网络以获得行人图像的紧凑表达,而多分支深度神经网络是实现分集特征挖掘的主要途径。文中分析了近年来用于行人重识别多分支深度神经网络的几种主流构建方法,包括局部分割分集、注意力分集、特征遮挡分集、异构支路分集等,并在3个行人重识别数据集 Market1501, Duke, CUHK03 上对不同算法性能进行了比较分析。基于当前进展,最后还讨论了行人重识别分集特征挖掘的未来研究趋势。

关键词:行人重识别;多分支网络;分集特征;特征嵌入

中图分类号:TP391 **文献标志码:**A **文章编号:**1673-5439(2021)01-0078-08

Recent advances on construction of multi-branch neural networks for person re-identification

WU Xiaofu¹, YIN Zirui¹, SONG Yue², ZHANG Lei¹, XIE Ben¹, ZHAO Shiliang¹, ZHANG Suofei³

(1.College of Telecommunications & Information Engineering, Nanjing University of Posts and Telecommunications, Nanjing 210003, China)
2.Unit 95958 of PLA, Shanghai 200120, China
3.School of Internet of Things, Nanjing University of Posts and Telecommunications, Nanjing 210003, China)

Abstract: With the surge of deep learning, there is increased interest on the design of multi-branch neural networks for extracting rich features in person re-identification. This paper reviews some recent advances on this topic, including part-based approaches, attention-based approaches, dropping approaches, and heterogeneous approaches. Various state-of-the-art approaches on the popular pedestrian re-identification datasets, including Market1501, Duke and CUHK03, are compared and analyzed, respectively. Finally, future directions of the person re-identification with diverse features are discussed.

Keywords: person re-identification; multi-branch neural networks; diverse features; feature embedding

行人重识别(Person Re-Identification, ReID)是计算视觉应用的主要领域之一,其核心任务是寻找出现在不同摄像头视图中的特定行人^[1-2]。作为一个跨摄像机的特定行人检索问题,由于人工智能理论和应用的相互推动,该问题近年来得到了广泛的

关注和长足的发展。对于给定的待查询行人,需要匹配出在不同地点、时间段甚至不同摄像头捕捉到的同一行人图像^[3]。作为一个复杂检索问题,ReID面临着:不同摄像机捕捉的行人图像具有包括视角、图像分辨率、光照、遮挡等一系列差异,这些差异对

收稿日期:2020-08-18;修回日期:2020-09-07 本刊网址:<http://nyzr.njupt.edu.cn>

基金项目:国家自然科学基金(61671253)资助项目

作者简介:吴晓富,男,博士,研究员,博士生导师,xfuwu@njupt.edu.cn

引用本文:吴晓富,尹梓睿,宋越,等.面向行人重识别分集特征挖掘的多分支深度神经网络构造进展[J].南京邮电大学学报(自然科学版),2021,41(1):78-85.

于检索精度的提高构成了巨大的挑战^[2]。

为了提高行人重识别的性能,如何构建具有分集特征挖掘能力的深度神经网络近年来成为行人重识别的热点问题之一。鉴于多分支深度神经网络的不同分支具有表达不同特征的便利性,因而成为实现分集特征挖掘的流行网络模型。近年来,通过多分支深度神经网络来获得分集特征的方法得到了广泛的关注,主要包括 4 大类:(1) 局部分割分集网络,典型如 PCB^[4]、PLR-OSNet^[5];(2) 注意力分集网络,典型如 ABD-Net^[6]、MHN^[7];(3) 遮蔽分集网络,典型如 BDB^[8]、SDB^[9];(4) 异构分集网络,如 BC-OSNet^[10]、ASNet^[11]。

本文主要针对行人重识别多分支分集技术的进展进行回顾与梳理,通过实验给出在典型行人 ReID 数据集的性能评估,并对未来的相关研究方向进行展望。

1 多分支深度神经网络与分集特征挖掘

一般地,假设具有 L 个分支的多分支深度神经网络由公共网络 $f_c(\cdot)$ 以及 L 个子分支网络 $f_l(\cdot)$, $l=1,2,\dots,L$ 构成。这样,给定输入的图像 x ,多分支深度神经网络输出 L 个支路特征矢量: $F_l=f_l(f_c(x))$, $l=1,2,\dots,L$ 。对 L 个支路特征矢量进行拼接 (Concatenation), 得到 x 的最终特征矢量

$$\mathbf{F}=[\mathbf{F}_1^T,\mathbf{F}_2^T,\dots,\mathbf{F}_L^T]^T \quad (1)$$

鉴于不同支路特征矢量的相同位置分量(如第 i 个分量)可以代表完全不同的特征,如何衡量不同分支特征矢量之间的分集性是个非常困难的问题。例如 2 个分支特征矢量 \mathbf{F}_1 与 \mathbf{F}_2 之间的平方欧式距离 $\|\mathbf{F}_1-\mathbf{F}_2\|_2^2$ 越大并不能说明第一分支与第二分支的分集性能越好。事实上,如果 \mathbf{F}_2 是 \mathbf{F}_1 的置换形式,虽然实质上两矢量表达的特征完全是一样的,但其欧式距离非但不为 0,而是可以相当大(取决于具体置换)。为此,一个可能的方式是引入一个对抗网络 (Adversary Network)^[12] \mathbf{g} ,使得^[7]

$$\text{div}(\mathbf{F})=\min_{\mathbf{g}}\sum_{i \neq j}\|\mathbf{g}(\mathbf{F}_i)-\mathbf{g}(\mathbf{F}_j)\|_2^2 \quad (2)$$

虽然以上方法在文献[7]中取得了一定的效果,但在其他实验中的结果并不理想,对抗网络的收敛性及其性能在理论和实践上都无法得到严格的保证。

为此,近几年来计算视觉方向的会议如 CVPR、ICCV 对行人重识别的特征分集问题进行了广泛的

探讨,下面分 4 大类进行介绍。

2 局部分割分集网络

局部分割的思路在于将图像全局的特征图进行局部分割,每一个分支对应到图像的一个局部,通过各分支有效学习局部细粒度特征,最终提高特征的分集表达能力。

2.1 PCB

相较于单分支的全局特征,PCB (Part-based Convolutional Baseline) 网络将特征水平均匀分块,实现对局部特征的挖掘;同时提出了 RPP (Refined-Part Pooling)^[4]方法来对齐局部特征,修正由于 PCB 均分策略而产生的对齐缺陷,进一步提升 PCB 的性能。

PCB 网络选用 ResNet50^[13] 为主干网络,将最后一个 Stride 变为 1,使得主干网络获得的特征图尺寸从 12×4 扩大到 24×8 ,并将特征图垂直分割为 L 个部分,经过平均池化和特征压缩后获得 L 个列向量 \mathbf{F} ,然后将其分别送入到全连接层,最后计算 L 组交叉熵 (Cross Entropy) 的和作为损失函数。测试时,PCB 可以将 L 个列向量 \mathbf{F}_l , $l=1,2,\dots,L$ 进行拼接得到特征 \mathbf{F} (见式(1))用于最后的判别。

相比基本的 IDE^[4]网络,PCB 的 mAP 在 Market1501^[14]、Duke^[15]、CUHK03^[16] 上分别提高了 8.9%、13.5%、15.3%,在行人重识别的分集特征挖掘领域是具有里程碑意义的算法(见图 1)。

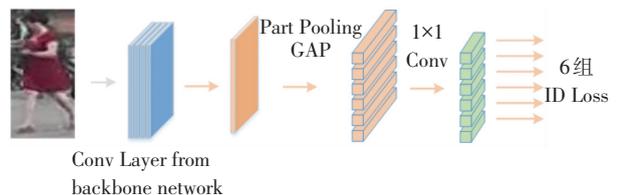


图 1 PCB 网络结构

2.2 PLR-OSNet

结合 PCB 结构局部分割的思想,PLR-OSNet^[5]是一种二分支深度神经网络模型,该分集网络由全局分支 (Global Branch) 和局部分支 (Local Branch) 组成。其中的局部分支采用统一分割策略进行局部特征分解。与 PCB 所不同的是,PLR-OSNet 并没有将 L 个局部特征单独分解为 L 个支路,而是提出将 L 个部分级特征向量拼接成单个列向量用于单独的一个 Local Branch,最终在该 Local Branch 上仅使用一个损失函数训练,这与现有基于部分分割的方法形成了鲜明的对比,如图 2 所示。测试时可以将全局分支特征和局部分支特征拼接后作为最后判别的特征。

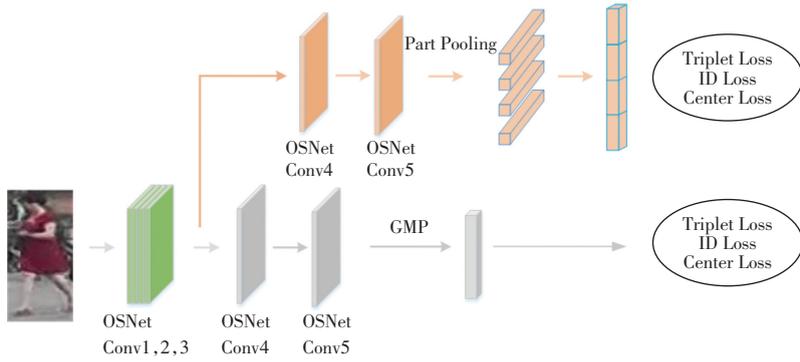


图 2 PLR-OSNet 网络架构

PLR-OSNet 全局特征结合局部特征,很好地实现了分集特征挖掘,对比 PCB+RPP 算法,其 mAP 在 Market1501 上继续提高了 7.3%。

3 注意力分集网络

类似于人类的注意力机制,利用注意力机制能有效提高神经网络的表达能力,在行人重识别领域得到了广泛的应用。为获取分集增益,加入注意力机制的多分支网络也被证明是行之有效的手段之一。

给定输入图像的某中间特征张量 $X \in \mathbf{R}^{C \times H \times W}$, 注意力机制可理解为对输入特征张量的所有元素进行加权,也即

$$Y = A(X) \odot X \quad (3)$$

其中, $A(X) \in \mathbf{R}^{C \times H \times W}$ 表示注意力模块的输出权重。 \odot 表示 Hadmard 乘法(逐元素相乘)。

3.1 ABD-Net

ABD-Net 是一种融入注意力机制以及正交正则

化技术的双分支网络。通过注意力机制,能够使网络注意力集中于前景的行人,从而淡化环境的影响,同时加入了正交正则化来保证不同支路特征具有分集多样性。

具体而言,ABD-Net 分别针对通道融合和位置感知,加入了一对互补的注意力模块:通道注意力模块(Channel Attention Module) $A(X) = \text{rep}[V] \mid^{H,W}$, 以及位置感知模块(Position Attention Module) $A(X) = \text{rep}[M] \mid^C$, 其中 $\text{rep}[V] \mid^{H,W}$ 表示信道掩模矢量在空间维度 H, W 上的重复扩展,同理 $\text{rep}[M] \mid^C$ 表示空间掩模在信道维度 C 上的重复扩展。ABD-Net 最终获得注意力加强的特征,如图 3 所示。为了增强输出特征的多样性,该网络提出了特征值差异正交约束(Spectral Value Difference Orthogonality, SVDO),并将其用于网络各个层的输入和输出端;通过同时正则化卷积、全连接层的权重和输出的特征张量,能够有效降低不同分支网络特征之间的相关度。

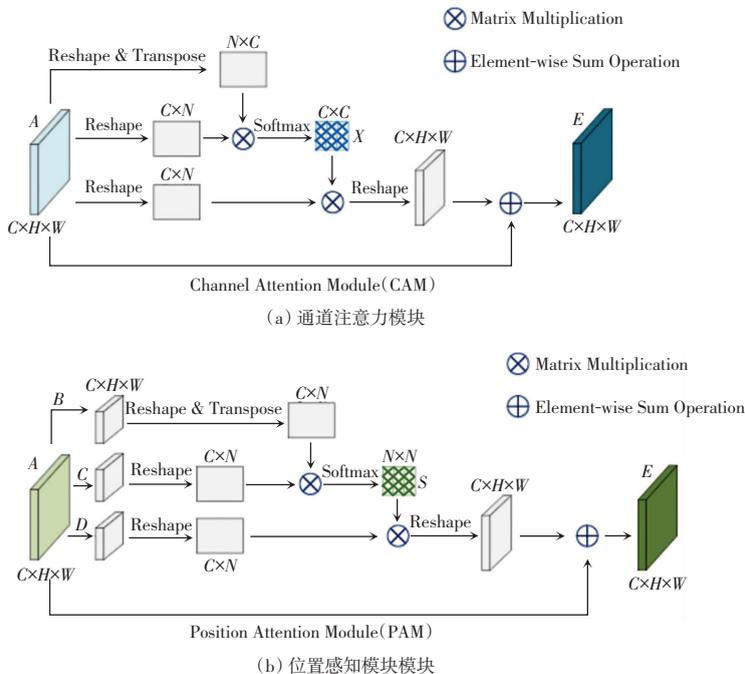


图 3 通道注意力模块和位置感知模块

3.2 MHN

ABD-Net 的注意力机制相对于输入张量 \mathbf{X} 而言是一阶的。为挖掘注意力机制的分集机理, MHN 首次提出高阶注意力机制 (High-Order Attention, HOA)。高阶注意力机制的核心是定义给定输入张量 \mathbf{X} 的 R 阶多项式表达

$$a(\mathbf{X}) = \sum_{r=1}^R (W^r, \otimes_r \mathbf{X}) \quad (4)$$

其中, $\otimes_r \mathbf{X}$ 表示 \mathbf{X} 分量的 r 阶多项式表达式, 以体现 \mathbf{X} 各分量之间的高阶关系。最终, 高阶注意力权重可以表示为

$$A(\mathbf{X}) = \text{sigmoid}(a(\mathbf{X})) \quad (5)$$

实验中, MHN 以 ResNet50 为骨架网络, 使用基于 PCB 的多分支网络, 每一个分支利用不同阶数的 HOA 模块获得的差异化分集特征, 如图 4 所示。

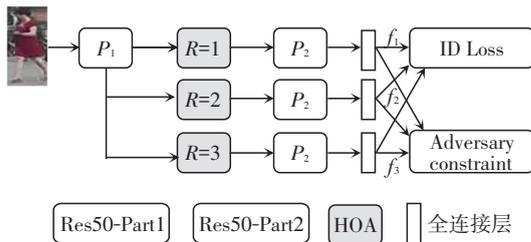


图 4 混合高阶注意力网络 (MHN)

其中 $R=1, 2, 3$ 分别代表一阶、二阶、三阶注意力模型, 通过不同的 HOA 模块实现具有分集能力的多分支网络, 然后连接到全连接层。为了使不同阶注意力模型输出特征更好地实现分集性能, 如式 (2) 所示, MHN 还通过训练对抗网络来分辨不同分支的特征, 使得不同分支的特征实现更高的分集度。

4 遮蔽分集网络

遮蔽分集网络主要是通过分支网络中添加遮挡部分, 以此强化未遮蔽部分特征表达能力的学习, 提取更富表达力的特征, 从而提高 ReID 模型的泛化能力。

4.1 BDB

针对深度卷积神经网络训练时易出现过拟合的问题, 各种数据增强技术获得了广泛的运用。例如, Random Erasing^[17] 直接对数据库中的图像进行随机的擦除, 获得增强的数据集, 一定程度上改善了过拟合现象。BDB 网络首次提出在网络中间层的特征层面进行随机的遮蔽, 以迫使神经网络学习更富有表达力的特征。如图 5 所示, BDB 网络包含两个分支, 也即全局分支与特征丢弃分支。其中, 全局分支

提取全局特征; 而特征丢弃分支则通过随机去除一个区域 (该区域图案在一个批次内保持不变), 区域内的所有像素均被清零, 再经过最大值池化和全连接层, 最后计算各种损失函数。

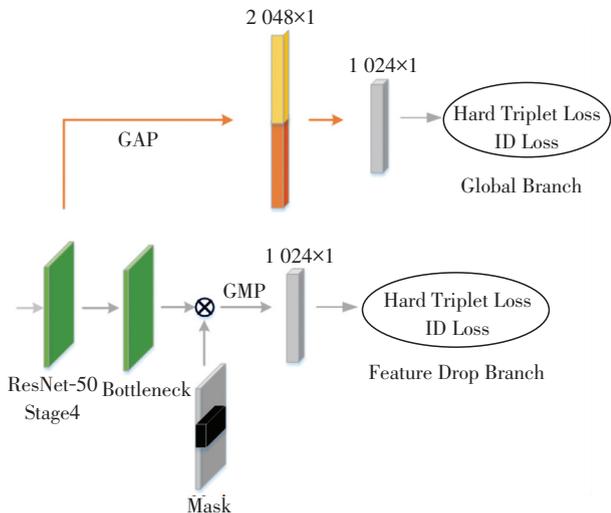


图 5 BDB (Batch DropBlock) 结构

通过加入遮蔽分支的方式, BDB 可以实现全局分支和特征丢弃分支之间的多样性, 有效实现了分集特征的挖掘。

4.2 SDB

BDB 作为一种典型的特征擦除方法, 已经被证明对 ReID 是有效的。然而, BDB 的丢弃模式在不同批次的擦除区域都不相同, 其网络可能难以学习到稳定的结构。基于该缺陷, SDB (Slow DropBlock) 提出了随机擦除图案在 Q 个批次 (Batch) 中保持不变, 并将擦除操作直接移至图像输入端。如图 6 所示, SDB 表示生成的擦除块至少在 $Q>1$ 个批次中保持不变, 当 $Q=1$ 时, SDB 等同于 BDB, 擦除的区域同样可以置零或者设为擦除区域值的平均。

由于删除了大量的输入图像, SDB 采用通常的训练方式难以收敛。为了解决这一问题, 论文提出了一种新的具有保证收敛性的双批次输入联合训练方法。该方法在训练过程中采用超级批次样本进行训练, 一个超级批次由两个批次组成 (一个常规批次和一个遮蔽批次), 因此其输入图像批次大小可表示为 $2B \times H \times W \times 3$ 的图像张量 (Tensor)。每个批次的样本数量均为 B 。超级批次首先输入到共享网络, 生成的特征图被分成两个子批次, 一个用于全局分支, 一个用于局部分支。

SDB 通过对不同局部分支输入批次的擦除区域高度设置, 获得了不同特征的多样性, 并将不同比

值下的 SDB 分支与全局分支拼接,进一步提升了网络性能。实现结果显示,四分支网络 SDB-4 对比 BDB 网络,mAP 在 Market1501,Duke,CUHK03 上分别提高了 2.0%、3.5%、4.0%。

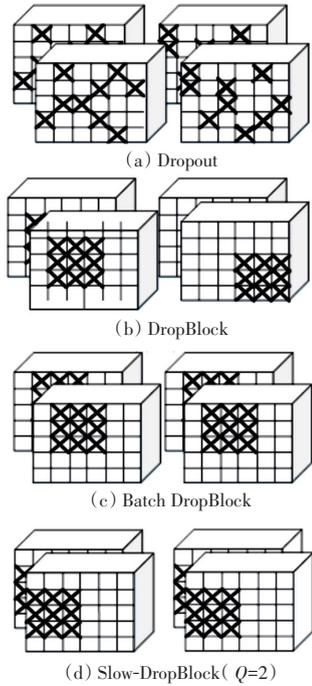


图 6 不同的 Dropout 方法的比较 (Batch Size=2)

5 异构支路分集网络

除了以上 3 大类分集网络外,还有一些工作无法归入以上 3 类,如 ASNet^[11] 和 BC-OSNet,本文将 其归入异构支路分集网络,其关键特征在于各支路 通过各种异构技术获得分集。下面以 BC-OSNet 为 例。

如图 7 所示,BC-OSNet 以 OSNet 为骨架网 络,结合了 PLR-OSNet 的部分分级方式,以及 Re- lation Network^[18] 的 GCP 和 One-vs-Rest 模块,再 加上全局分支,组成四分支网络。其中,One-vs- Rest^[18] 关系模块考虑了身体各个部位与身体其 他部位之间的关系,使得每个部分级别特征都包 含了相应部位本身以及身体其他部位的信息,使 其更具识别性。全局对比池化(GCP)^[18] 则通 过对比全局特征和局部特征,以获得更加突出的 全局特征表达。

最终 BC-OSNet 在 Market1501、Duke、CUHK03 三大数据集上都获得了性能的明显提升。

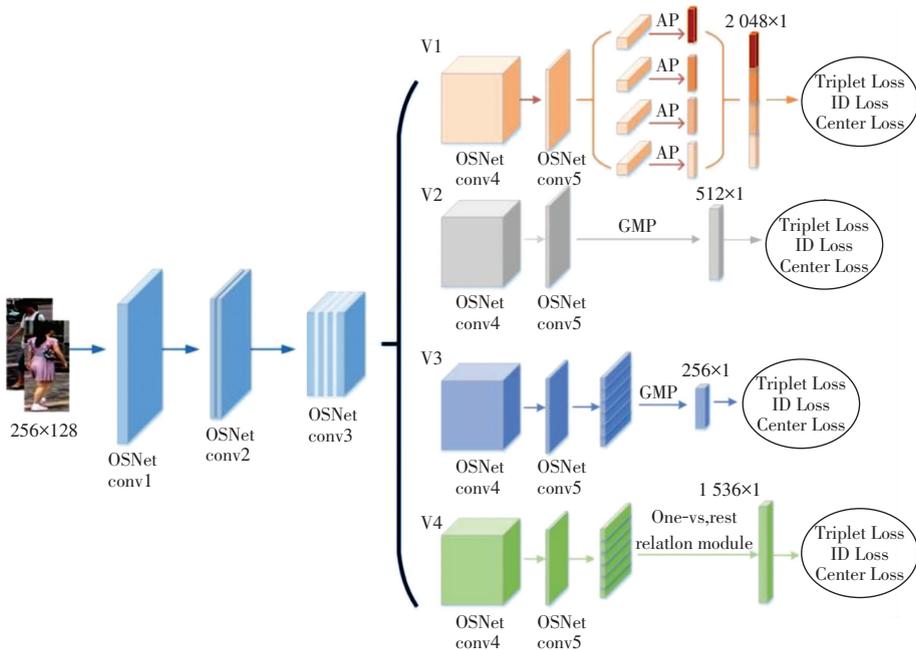


图 7 BC-OSNet 框图

6 实验结果分析与讨论

6.1 数据集

实验部分主要围绕 ReID 领域 3 个比较流行的数

据集上展开,分别是 Market1501、Duke 和 CUHK03。

6.2 实验环境及参数配置

本文实验采用的硬件和软件仿真环境如表 1 所示。

表 1 实验平台参数

配置	参数
Operating System	Ubuntu 16.04
CPU	Intel(R) Xeon(R) CPU E5-2620 v4@ 2.10 GHz
GPU	Tesla K80
Software Platform	Pytorch 1.1

6.3 实验复现与性能比较

在本次实验中,复现了部分分集网络,主要包括 PCB、PLR-OSNet、SDB,并加入了更多的训练技巧。在训练和测试中,输入图像尺寸被统一到尺度 384×128 ;数据增强方法采用了随机翻转、随机擦除、随机裁剪;优化器则选用了 Adam^[19] 优化器,动量设置为 0.9,衰减为 $5e-04$;在训练期间,批次大小设置为 64,每批包含 16 个行人身份,每个行人有 4 张样本图像。训练的 Epoch 数量为 120,学习率初始化为 $3.5e-5$;训练过程中采用 Warmup 策略使学习率呈线性增长,经过 20 个 Epoch,学习率变化为 $3.5e-5$;到达 60 个 Epoch,学习率变化为 $3.5e-5$;再到 90 个 Epoch,学习率变化为 $3.5e-6$ 。损失函数使用行人 ID Loss^[5], Soft-Margin Triplet Loss^[20] 和 Center Loss^[21],即

$$L_{\text{sum}} = \lambda_1 L_{\text{ID}} + \lambda_2 L_{\text{Triplet}} + \lambda_3 L_{\text{Center}}$$

最终通过 PyTorch 编程实现神经网络端到端的训练,表 2 给出了基于局部分割分集算法的实验结果。从实验结果可以看出,通过局部分割实现支路分集的特征挖掘技术在 ResNet、OSNet 等网络架构上都有很好的表现。

表 2 局部分割分集网络复现实验

方法	Market1501		Duke	
	mAP	Rank-1	mAP	Rank-1
IDE(ResNet50)	68.5	85.3	66.1	81.7
PCB ^[20]	77.3	92.4	65.3	81.9
OSNet	84.9	94.8	73.5	88.6
PLR-OSNet	88.9	95.6	81.2	91.6
PCB(复现)	78.4	93.9	70.4	82.8

基于以上系统配置,给出了遮蔽分集网络 SDB 的实验结果,骨架网络选用 ResNet-50,使用 $Q=5$ 的默认设置。对于 SDB-2 的其中一个分支使用遮蔽占宽比 $r_h=0.3$,对于 SDB-4 中的 3 个分支分别使用 $r_h=0.2, 0.3, 0.4$,表 3 给出了实验结果。

表 3 SDB 性能复现对比

方法	Market1501		Duke		CUHK03_Labeled		CUHK03_Detected	
	mAP	Rank-1	mAP	Rank-1	mAP	Rank-1	mAP	Rank-1
BDB	86.7	95.3	76.0	89.0	76.7	79.4	73.5	76.4
SDB-2	88.3	95.7	78.8	89.7	78.0	80.6	74.9	77.2
SDB-4	88.7	95.9	79.5	90.2	80.7	82.6	77.4	79.5

由表 3 可以看出,遮蔽分集的方式使得多分支网络学习到了更多的有效特征,能有效提高行人重识别性能。

最后,在统一的实验平台上给出了 BC-OSNet 实验结果,其设置如下:Epoch 的数量改为 160,学习率初始化为 $3.5e-5$,在训练过程中使用 Warmup,使学习率呈线性增长,经过 20 个 Epoch,学习率变化为 $3.5e-4$;到达 60 个 Epoch,学习率变化为 $3.5e-5$;再到 130 个 Epoch,学习率变化为 $3.5e-6$ 。表 4 中显示了与不同方法的性能比较,可以看出,BC-OSNet 以 4 个分支网络的规模相比于其他流行算法表现出几乎最好的性能。

表 4 在 Market1501 和 Duke 数据集上的各种网络性能对比

方法	Market1501		Duke	
	mAP	Rank-1	mAP	Rank-1
HA-CNN ^[22]	75.70	95.60	63.80	80.50
PCB ^[4]	77.30	92.40	65.30	81.90
Aligned ^[23]	79.10	91.80	69.70	82.10
PCB+RPP ^[4]	81.00	93.10	68.50	82.90
HPM ^[24]	82.70	94.20	74.30	86.60
OSNet ^[25]	84.90	94.80	73.50	88.60
MGN ^[26]	86.90	95.70	78.40	88.70
ABD ^[6]	88.28	95.60	78.59	89.00
GCP ^[18]	88.90	95.20	78.60	89.70
BDB ^[8]	86.70	95.30	76.00	89.00
SONA ^[27]	88.67	95.68	78.05	89.25
SDB-4 ^[9]	88.70	95.90	79.50	90.20
MHN-6(PCB) ^[7]	85.00	95.10	77.20	89.10
PLR-OSNet ^[5]	88.90	95.60	81.20	91.60
BC-OSNet ^[27]	89.90	95.70	82.10	91.30

6.4 网络参量大小比较

表5中给出了不同网络的参数量大小,一些网络由于没有公布代码,参量大小不好确定。由表5可知,由于OSNet网络的参量规模较小,以它为骨架搭建的多分支网络参量规模都较小。如BC-OSNet虽然有4个分支,最终参量规模为 8.2×10^6 ,仍大幅小于ResNet50,因此适合资源受限的应用场景。

表5 部分网络的参数量

方法	参数量/ 10^6
IDE(ResNet50)	26.0
PCB	27.0
PLR-OSNet	3.4
MHN	26.8
SDB	60.0
BC-OSNet	8.2

7 结束语

分集特征挖掘一直是提高行人重识别准确率的重要手段,本文简要总结了近年来基于多分支深度神经网络实现分集特征挖掘的各类算法,分别是局部分割分集、注意力分集、特征遮挡分集、异构分支分集。可以看出,使用多分支网络充分挖掘行人信息来提高检测性能是性能提高的主流手段之一,是近年计算视觉会议CVPR、ICCV行人重识别有监督学习领域的关注热点,取得了丰硕的成果。

然而,目前多分支神经网络提取特征之间的分集度数学上仍然缺乏严格的定义,式(2)虽然给出了一种可能,但该定义仍存在以下明显缺陷:(1)依赖对抗网络的定义本身不直接;(2)对抗网络本身难以收敛;(3)实验效果不突出。因此,如何给出特征分集的严格数学定义并设计有效的多分支网络来挖掘行人的细粒度信息是未来继续提升行人重识别性能的关键。

参考文献:

- [1] GONG S G, CRISTANI M, LOY C C, et al. The Re-identification challenge [M] // Person Re-Identification. London: Springer London, 2014: 1-20.
- [2] ZHENG L, YANG Y, HAUPTMANN A G. Person Re-identification: past, present and future [EB/OL]. [2020-08-06]. <https://arxiv.org/abs/1610.02984>.
- [3] GHEISSARI N, SEBASTIAN T B, HARTLEY R. Person re-identification using spatiotemporal appearance [C] // IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2006: 1528-1535.
- [4] 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 European Conference on Computer Vision (ECCV). 2018: 480-496.
- [5] XIE B, WU X F, ZHANG S F, et al. Learning diverse features with part-level resolution for person Re-identification [EB/OL]. [2020-08-06]. <https://arxiv.org/abs/2001.07442>.
- [6] CHEN T L, DING S J, XIE J Y, et al. ABD-net: Attentive but diverse person Re-identification [EB/OL]. [2020-08-06]. <https://arxiv.org/abs/1908.01114>.
- [7] CHEN B H, DENG W H, HU J N. Mixed high-order attention network for person Re-identification [C] // IEEE/CVF International Conference on Computer Vision (ICCV). 2019: 371-381.
- [8] DAI Z Z, CHEN M Q, GU X D, et al. Batch DropBlock network for person Re-identification and beyond [C] // IEEE/CVF International Conference on Computer Vision (ICCV). 2019: 3690-3700.
- [9] WU X F, XIE B, ZHAO S L, et al. Diversity-achieving Slow-DropBlock network for person Re-identification [EB/OL]. [2020-08-06]. <https://arxiv.org/abs/2002.04414>.
- [10] ZHANG L, WU X F, ZHANG S L, et al. Branch-cooperative OSNet for person Re-identification [EB/OL]. [2020-08-06]. <https://arxiv.org/abs/2006.07206>.
- [11] ZHANG S F, ZHANG L, WANG W L, et al. AsNet: asymmetrical network for learning rich features in person Re-identification [J]. IEEE Signal Processing Letters, 2020, 27: 850-854.
- [12] GOODFELLOW I, POUGET-ABADIE J, MIRZA M, et al. Generative adversarial nets [C] // Advances in Neural Information Processing Systems. 2014: 2672-2680.
- [13] HE K M, ZHANG X Y, REN S Q, et al. Deep residual learning for image recognition [C] // IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2016: 770-778.
- [14] ZHENG L, SHEN L, TIAN L, et al. Scalable person Re-identification: a benchmark [C] // IEEE/CVF International Conference on Computer Vision (ICCV). 2015: 1116-1124.
- [15] RISTANI E, SOLERA F, ZOU R, et al. Performance measures and a data set for multi-target, multi-camera tracking [C] // Proceedings of the European Conference on Computer Vision (ECCV). 2016: 17-35.
- [16] LI W, ZHAO R, XIAO T, et al. Deepreid: deep filter pairing neural network for person re-identification [C] // IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2014: 152-159.
- [17] ZHONG Z, ZHENG L, KANG G L, et al. Random erasing data augmentation [C] // Proceedings of the AAAI Confer-

- ence on Artificial Intelligence.2020;13001–13008.
- [18] PARK H, HAM B. Relation network for person Re-identification [C] // Proceedings of the AAAI Conference on Artificial Intelligence. 2020; 11839–11847.
- [19] KINGMA D P, BA J. Adam: a method for stochastic optimization [EB/OL]. [2020-08-06]. <https://arxiv.org/abs/1412.6980>.
- [20] HERMANS A, BEYER L, LEIBE B. In defense of the triplet loss for person Re-identification [EB/OL]. [2020-08-06]. <https://arxiv.org/abs/1703.07737>.
- [21] WEN Y D, ZHANG K P, LI Z F, et al. A discriminative feature learning approach for deep face recognition [C] // Proceedings of the European Conference on Computer Vision (ECCV). 2016.
- [22] LI W, ZHU X T, GONG S G. Harmonious attention network for person Re-identification [C] // IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2018; 2285–2294.
- [23] LUO H, JIANG W, ZHANG X, et al. AlignedReID++: dynamically matching local information for person Re-identification [J]. Pattern Recognition, 2019, 94: 53–61.
- [24] FU Y, WEI Y C, ZHOU Y Q, et al. Horizontal pyramid matching for person Re-identification [C] // Proceedings of the AAAI Conference on Artificial Intelligence. 2019; 8295–8302.
- [25] ZHOU K Y, YANG Y X, CAVALLARO A, et al. Omni-scale feature learning for person Re-identification [C] // IEEE/CVF International Conference on Computer Vision (ICCV). 2019; 3701–3711.
- [26] WANG G S, YUAN Y F, CHEN X, et al. Learning discriminative features with multiple granularities for person Re-identification [C] // Proceedings of the 26th ACM International Conference on Multimedia. 2018; 274–282.
- [27] BRYAN B, GONG Y, ZHANG Y Z, et al. Second-order non-local attention networks for person Re-identification [C] // IEEE/CVF International Conference on Computer Vision (ICCV). 2019; 3759–3768.