



信号处理
Journal of Signal Processing
ISSN 1003-0530, CN 11-2406/TN

《信号处理》网络首发论文

题目：基于 PCB 特征加权的行人重识别算法
作者：赵师亮，吴晓富，张索非
网络首发日期：2020-06-12
引用格式：赵师亮，吴晓富，张索非. 基于 PCB 特征加权的行人重识别算法. 信号处理.
<https://kns.cnki.net/kcms/detail/11.2406.TN.20200612.1346.044.html>



网络首发：在编辑部工作流程中，稿件从录用到出版要经历录用定稿、排版定稿、整期汇编定稿等阶段。录用定稿指内容已经确定，且通过同行评议、主编终审同意刊用的稿件。排版定稿指录用定稿按照期刊特定版式（包括网络呈现版式）排版后的稿件，可暂不确定出版年、卷、期和页码。整期汇编定稿指出版年、卷、期、页码均已确定的印刷或数字出版的整期汇编稿件。录用定稿网络首发稿件内容必须符合《出版管理条例》和《期刊出版管理规定》的有关规定；学术研究成果具有创新性、科学性和先进性，符合编辑部对刊文的录用要求，不存在学术不端行为及其他侵权行为；稿件内容应基本符合国家有关书刊编辑、出版的技术标准，正确使用和统一规范语言文字、符号、数字、外文字母、法定计量单位及地图标注等。为确保录用定稿网络首发的严肃性，录用定稿一经发布，不得修改论文题目、作者、机构名称和学术内容，只可基于编辑规范进行少量文字的修改。

出版确认：纸质期刊编辑部通过与《中国学术期刊（光盘版）》电子杂志社有限公司签约，在《中国学术期刊（网络版）》出版传播平台上创办与纸质期刊内容一致的网络版，以单篇或整期出版形式，在印刷出版之前刊发论文的录用定稿、排版定稿、整期汇编定稿。因为《中国学术期刊（网络版）》是国家新闻出版广电总局批准的网络连续型出版物（ISSN 2096-4188，CN 11-6037/Z），所以签约期刊的网络版上网络首发论文视为正式出版。

基于 PCB 特征加权的行人重识别算法

赵师亮¹ 吴晓富¹ 张索非²

(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: 为充分挖掘行人重识别特征, 最近流行的 PCB 算法给出了一种特征均匀分块并通过 RPP 网络对齐特征的方法。PCB 算法充分发挥了局部特征的作用, 有效提高了行人重识别的准确率。为进一步提高行人重识别的性能, 本文基于全局特征与局部特征对网络性能的影响差异提出了一种特征加权的 PCB 行人重识别算法。在典型的行人识别数据库 Market1501、DukeMTMC-Reid 上的实验结果表明: 所提算法具有更好的首中准确率 (Rank1) 和平均准确率 (mAP); 相比与经典的 PCB+RPP 算法, 所提算法在 Market1501 数据集上 Rank1 提高了 0.8%, mAP 提高了 4.5%; 在 DukeMTMC-Reid 数据集上 Rank1 提高了 5.5%, mAP 提高了约 7%。

关键词: 行人重识别; 特征加权; 深度神经网络; 特征多样性

中图分类号: TP391 **文献标识码:** A **DOI:**

Weighted PCB for Person Re-Identification

ZHAO Shi-liang¹ WU Xiao-fu¹ ZHANG Suo-fei²

(1. College of Communication and Information Engineering, Nanjing University of Posts and Telecommunication, Nanjing, Jiangsu 210003, China; 2. Internet of Things College, Nanjing University of Posts and Telecommunication, Nanjing, Jiangsu 210003, China)

Abstract: In order to fully exploit the features of any input image for person re-identification (Re-ID), part-based convolutional baseline (PCB) was proposed to employ a uniform partition for any input image and a refined part pooling (RPP) method is followed for enhanced within-part consistency. In order to further improve the performance of Person Re-ID, this paper proposes a weighted PCB algorithm, which combines the global feature and local part-based features in a weighted form. Experiments show that the proposed algorithm is better than other weighted methods. Experiments over Market1501 and DukeMTMC-Reid show that the proposed algorithm can achieve better performance in both the Rank1 accuracy and the mean average accuracy (mAP). Compared with the PCB+RPP algorithm, the proposed algorithm provides a margin of 0.8% and 4.5% over Market1501 for Rank1 and mAP, respectively. For the dataset of DukeMTMC-Reid, it improves PCB by 5.5% in Rank1 accuracy and by about 7% in mAP.

Key words: person re-identification; weighted feature; deep neural network; feature diversity

1 引言

行人重识别^[1]技术主要旨在解决跨摄像头、跨场景下行人的识别与检索。当前行人重识别主要包括行人的特征表示和相似度计算两个阶段。由于姿态差异、视角变化、光线和遮挡等问题使行人图像变化巨大, 给检索任务带来巨大的困难和挑战。如何更好地提取行人的特征成为行人重识别技术的关键。近年来, 随着深度学习的迅速发展, 手工设

计特征的方式被逐步取代。相比与传统的手工提取的特征, 深度学习^[2]提取的特征表示具有更好判别性。

在特征表示方面, 早期关于深度学习的行人重识别算法^[1]主要关注行人图像的全局特征, 将行人图像经过神经网络, 提取其全局特征, 然后利用该特征进行相似度计算。随着全局特征研究的成熟, 行人重识别性能的提升遇到了瓶颈, 为了进一步提升行人重识别的性能, 开始挖掘细粒度特征, 渐渐

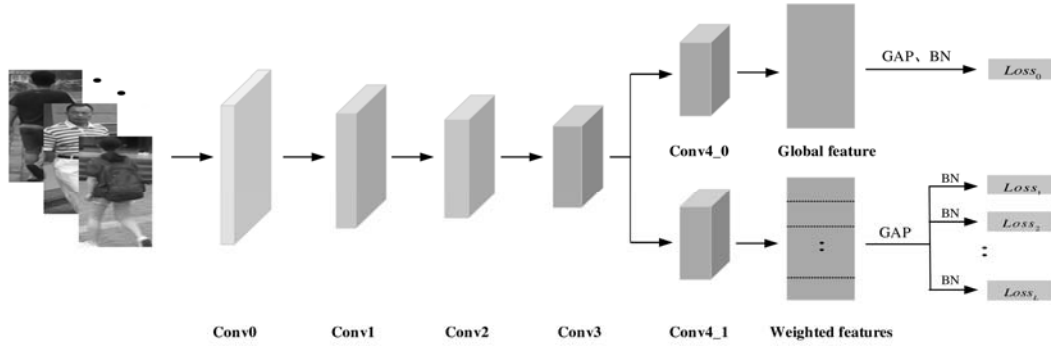


图 1 网络结构

Fig 1 The structure of network

研究起图像的局部特征^[3,4]。常用的提取局部特征的思想主要有图像或特征切块、利用骨骼关键点定位以及姿态矫正等^[3-5]。其中切块是局部特征研究中最常见的一种提取特征的方式。PCB^[3]算法是局部特征研究中的代表，其充分挖掘了图像的局部特征，有效提高了行人重识别的准确率。

考虑到仅用局部特征的局限性，以及垂直分割特征可能带来信息损失，本文在 PCB 算法的基础上，进一步融入图像的全局特征；考虑到全局特征与局部特征的关联性与差异性，提出了一种基于全局特征与局部特征的加权算法。该算法通过有效融合全局特征和局部特征，可以更好地提高模型的准确率和泛化性能。

2 基本原理

2.1 PCB 算法原理介绍

PCB^[3]算法主要贡献是，一是设计了 PCB (Part-based Convolutional Baseline) 网络，二是提出了 RPP (Refined part pooling) 方法

2.1.1 PCB 网络结构

PCB 选用 ResNet50^[6]作为主干网络，PCB 在 ResNet50 的基础上做了一些修改。ResNet50 中 GAP (global average pooling) 之前的层保持不变，PCB 修改的是 GAP 层以及之后的层。

首先，将输入的图片经过主干网络得到特征 T ，定义 channel 维的列向量为 f 。然后，PCB 算法将 T 水平均分成 p 份，经过 average pooling 得到 p 个列向量 g ，再使用 1×1 的卷积核降维得到 p 个降维后的特征 h 。最后，每个 h 输入到由 FC (fully-connected layer) 和 Softmax function 构成的分类器 classifier 中，从而预测输入的 ID，这里的 FC 参数不共享，即共有 p 个 classifier，实验发现参数不共享的效果更好。

在训练时，PCB 使用 p 个 classifier 的交叉熵 (Cross Entropy) 的和作为损失函数。在测试时，PCB 可以选择把 p 个 g 或者 p 个 h 进行拼接得到最后用于判别特征 G 或者 H ，其中， $G = [g_1, g_2, \dots, g_p]$ ， $H = [h_1, h_2, \dots, h_p]$ 。

2.1.2 RPP 网络

PCB 的均分策略是简单、有效的方法，但是可能存在特征对齐的问题，即行人图像是有明显的语义信息的，而简单的划分特征，可能破坏这些关联的特征。RPP 即 Refined part pooling，其是在训练好的 PCB 网络的基础上，更改 PCB 网络的 average pooling，用 RPP 网络代替，用来修正由于 PCB 均分策略而产生的不良影响，进一步提升 PCB 的性能。RPP 本质上就是注意力模块，目标是把 p 个部分对应的空间分布进行软权值分配，进而对齐各个部分。

2.2 基于 PCB 特征加权网络

本文采用的网络架构类似与 PCB 网络，其中主干网络也采用 ResNet50，详见图 1，其中 ResNet50 中 Conv0 - Conv3 层的参数共享。从 Conv4 后网络分为两个分支，即 Global 分支和 Local 分支，其中，Conv4_0 与 Conv4_1 初始参数相同。这里主要是参考文献^[7]多分支训练方式，让不同分支的参数学习不同的特征，增强不同分支学习特征差异性。与 PCB 网络采用同样的技巧，将 Conv4 层中的下采样步长由 2 设置为 1，这样可以增大最终模型提取特征的感受野，提高行人识别的性能。

Global 分支中，Conv4 后直接采用 GAP 和 BN (Batch Norm)，用以提取图像的全局特征。Local 分支中，与 PCB 网络类似，经过 Conv4 后的特征 T ，被水平均匀分割成 p 份，经过 Average Pooling 得到 p 个向量，每个特征向量经过 BN，最后分别输入到 p 个 FC (fully-connected layer) 用以分类。

2.3 网络训练的技巧及数据增强

较人脸识别的算法而言,行人重识别领域有更多可以提升网络性能的训练技巧。本文在训练基于PCB特征加权网络时,也采用了其中一些常用的技巧。

(1)warmup^[8]学习率调整:其是一种在训练时,随着迭代次数的增加,学习率先由小慢慢增大,然后再衰减的训练过程。

(2)随机翻转(random flip):即在训练阶段以一定概率,随机水平翻转输入的图像,用以扩充训练数据。它是一种常用的数据增强的手段。

(3)随机擦除数据增强^[8,9](random erase):训练时,通过随机擦除归一化后的图片输入的某一块区域,并用归一化后的图像均值填充该区域,起到数据增强,从而提高模型的泛化性能。

(4)Last Stride^[8]:ResNet50中最后一层的Conv4层有一个下采样步长为2的步骤,用以进一步缩小特征图的尺度。而Last Stride即是把该Conv4层的下采样步长为2设置为1,从而可以使得最终得到的特征图尺度更大。

(5)BN^[8]:即网络输出的特征在送入分类器前,先经过一个BN(Batch Norm)层,再送入分类器。

2.4 相似性度量

在行人重识别任务中,主要分为两个阶段,第一阶段即网络的训练,第二个阶段则是用训练好的模型提取输入图像的特征信息,进而用提取的特征进行相似性度量。这两个阶段都至关重要,相辅相成。基于表征学习的行人重识别算法中,常用的相似性度量的方式一般有两种,一是基于欧氏距离,简称L2距离,二是基于余弦距离,即Cosine距离。考虑到本文在训练时,在特征送入分类层前,经过了BN(Batch Norm)层,对特征进行了归一化操作,考虑到后续的加权算法,余弦距离是更适合本文网络的度量方式。测试时,所采用的特征均为经过BN后,即送入分类层之前的特征,每一分支的特征维度为1 x 2048。

2.5 多分支训练

与PCB算法相同,本文网络也采用的是多分支训练方式。不同的是在PCB算法的基础上,本文还加入了全局特征的信息,全局特征可以很好的补充局部特征分割时带来的特征对齐问题,从而不需要额外的RPP网络进行特征的矫正。从图1可以看出,网络从ResNet50-Conv3后分为上下两分支,其中上分支为global分支,用于提取全局特征的特征图

(feature map),然后经过全局池化层(GAP),得到最终的全局特征,下分支主要是对模型提取的feature map进行垂直方向的等距离分割,将分割后的每个子feature map经过GAP,得到多个局部特征。其中之所以ResNet-Conv4的参数训练时不再共享,是因为这样可以让相应的Conv4_0层更好的学习全局特征参数,Conv4_1层更好的学习局部特征的参数。若下分支将feature map均分成L块,则最终加上全局特征,共有L+1个特征,其中每一维度均为1 x 2048。在训练阶段,将它们分别送入2048 x classes的全连接层,其中classes为训练集类别数,加上global分支,共有L+1个独立的分类层,损失函数采用Cross Entropy Loss,分别对应 $loss_0, loss_1, \dots, loss_L$,训练时的最终loss为所有子loss的和,即:

$$loss_{all} = \sum_{i=0}^L loss_i \quad (1.1)$$

2.6 特征加权融合算法

基于多分支的训练后,便是如何融合全局特征和多个局部特征来提升模型准确率的问题。PCB算法因为只应用了局部特征,故其可以直接拼接所有局部特征,进而用RPP网络来矫正所提取局部特征的不足,进而衡量行人重识别网络的性能。而本文所提网络加入了全局特征,用来补充局部特征的局限性,故不能简单的拼接全局特征和局部特征。单个局部特征与全局特征信息对模型性能的影响性不同,直接拼接全局特征与全部局部特征来进行最终的相似性度量显然是不合适的。此外在网络训练的过程中,每一分支都通过各自分类器单独监督各个特征的,即每一分支的特征是相对对立的,通过加权各个分支的特征,可以提高模型的鲁棒性。加权因子的选取逻辑如下:设两张输入图片的全局特征为 f_g, h_g (列矢量),将其分解为L个局部特征,也即, $f_g = [f_1^T, \dots, f_L^T]^T$, $h_g = [h_1^T, \dots, h_L^T]^T$,则有:

$$\begin{aligned} |f_g - h_g|^2 &= |[f_1^T, \dots, f_L^T]^T - [h_1^T, \dots, h_L^T]^T|^2 \\ &= \sum_{i=1}^L |f_i - h_i|^2 \end{aligned} \quad (1.2)$$

由此可见:如果每个分支提取的局部特征贡献是相等的,则全局特征的贡献大致是每一局部特征贡献的L倍。如果全局特征保持不变,则每一局部特征与全局特征融合时的权重系数应该是 $1/\sqrt{L}$ 。因此,最终多支路融合的相似性度量公式为

$$f = f_{global} + \sum_{i=1}^L \frac{1}{\sqrt{L}} f_{local_i} \quad (1.3)$$

这里‘+’表示特征拼接的含义,即将所有的局部特征先乘以一个加权因子,全局特征的加权因子为

1, 然后拼接所有的特征, 用于行人重识别的度量。之所以对每一局部特征均采用均值加权的方式, 而不是采用其他概率分布的形式, 主要是考虑到了算法的泛化性能, 不同的输入图片其自上而下的特征分布可能位置有很大的不同, 而且不同的数据集也可能差异巨大。故采用均值加权局部特征的方式可以使模型具有更好的泛化性能。

3 实验结果分析与讨论

3.1 数据集与评价评价指标

本文分别在数据级 Market1501^[10] 和 DukeMTMC-Reid^[11] 两大数据集上进行了实验。其中, Market1501 数据集是由清华大学所收集, 包含从 6 个不同的摄像头拍摄到的 1501 个行人, 其中训练集有 751 个行人, 共 12936 张图片; 测试集有 750 个行人, 共 19732 张图片。Duke 数据集是由 Duke 大学所收集, 包含 1404 个行人, 其中训练集含 702 个行人, 共 16522 张行人图片, 测试集包含 702 个行人, 共 17661 张行人图片。这两个数据是目前行人重识别中较大的两个数据集, 包含有姿态变化、光照变化、遮挡等问题, 符合实际应用场景中常见的情况。

目前行人重识别方法常用的评价指标为 CMC 曲线 (Cumulated Matching Characteristic) 和平均准确率 mAP (Mean Average Precision,) 两种评价指标评估方法的性能。

3.2 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

Table 1 Experimental platform parameters

Configuration	Parameter
Operating System	Ubuntu 14.04
CPU	Intel(R) Xeon(R) CPU E5-2680 v4@ 2.40GHz
GPU	Tesla P40
Software Platform	Pytorch 1.1

在训练过程中, 骨干网络采用在 ImageNet 上预训练过的 ResNet50 模型, 在网络初始化时已完成, 以此减少训练时间。为了加大最后网络提取特征的尺度, 行人图片被调整为 384*128 分辨率输入, 最后模型提取的 *feature map* 为 24*8, *local* 分支在在垂直切分来提取局部特征时, 可以有更多的切割选择。

训练批次大小为 64, 采用随机采样方式。共进行 120 个 epoch 训练, 学习率采用线性的 warmup 方式, 这样可以更好的发挥预训练模型参数的作用, 从 1-10 epoch, 学习率从 3e-6 逐渐增加到 3e-4, 11-50 epoch, 学习率保持 3e-4, 51-90 epoch, 学习率变为十分之一, 即 3e-5, 91-120 epoch, 学习率再衰减为原来的十分之一, 变为 3e-6。优化器选择的是带动量系数的 Adam, 一阶动量系数为 0.9, 二阶动量系数为 0.999。训练数据增强方式为随机水平翻转 (random flip)、随机擦除 (random erase)。

3.3 不同局部特征分割方式的影响

表 2 不同 L 情况下 Market1501 实验

Table 2 Market1501 with different L

L	Rank 1	Rank 5	Rank 10	mAP
0	92.9	97.4	98.5	81.4
1	93.2	97.5	98.7	83.4
2	94.6	97.6	98.4	86.1
3	94.2	97.8	98.7	85.8
4	94.0	97.8	98.6	84.4
6	93.9	97.6	98.5	83.7
8	91.9	96.9	98.1	78.9

表 3 不同 L 情况下 Duke 实验

Table 3 Duke with different L

L	Rank 1	Rank 5	Rank 10	mAP
0	84.8	92.9	94.7	71.4
1	86.8	93.6	95.3	74.4
2	88.4	94.5	96.1	77.0
3	88.2	94.8	96.4	77.1
4	88.8	94.8	96.2	76.5
6	87.7	93.7	95.2	76.0
8	83.8	91.2	94.8	68.5

因为输入图片的尺寸为 384*128, 且 ResNet50 最后一层中的步长被修改为 1, 故经过此 ResNet50 模型后的 *feature map* 尺寸为 24*8。显然, 1、2、3、4、6、8 为 24 的因子, 即网络下分支的 *feature map* 可以有多种划分的方式。

这里不考虑 12、24 等划分情况, 因为划分太细存在严重的特征对齐问题, 过分的划分只会扩大特征对齐问题, 反而不利于网络模型的学习。若因子为 L, 即表示模型的下分支有 L 个局部特征, 每个局部特征的加权因子为 $1/\sqrt{L}$ 。下面是基于不同因子时在 Market1501 与 DukeMTMC-Reid 上的实验结果, 见表 2、3。

表中 $L=0$ 表示, 仅用 *global* 分支, 没有 *local* 时, 即仅用全局特征时, 网络的性能。对网络中 *local* 分支提取的 *feature map* 进行不同的划分, 网络的性能有所差异。综合分析不同的分割方式, 网络一般在对 *feature map* 分割数为 2、3、4 情况下, 取得较好的性能。

若不采用式 (1.3) 的特征加权融合方式, 直接

拼接所有特征, 即直接拼接全局特征和所有的局部特征, 不考虑全局特征和局部特征的差异性, 将最终的拼接的特征作为融合后的特征, 用于最终网络性能的衡量。网络的性能见图 2、3, 其中 “with weighted” 表示经过加权网络的性能, “without weighted” 表示直接拼接特征网络的性能。

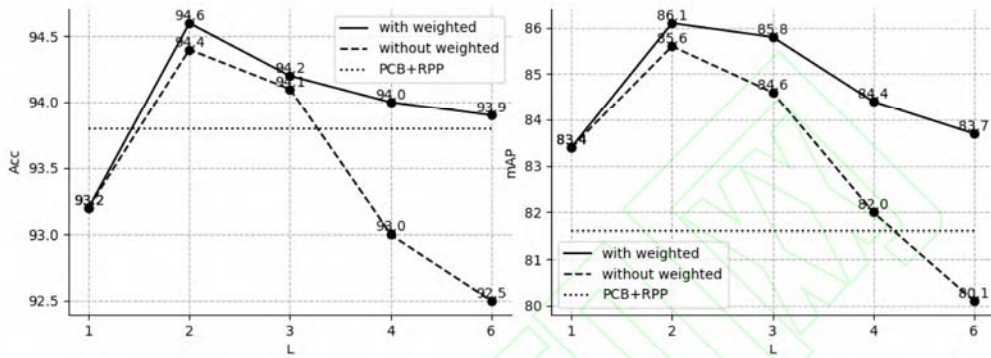


Fig 2 Comparison of results on Market1501 with and without weighting method

图 2 Market1501 数据集上特征加权与直接拼接性能比较

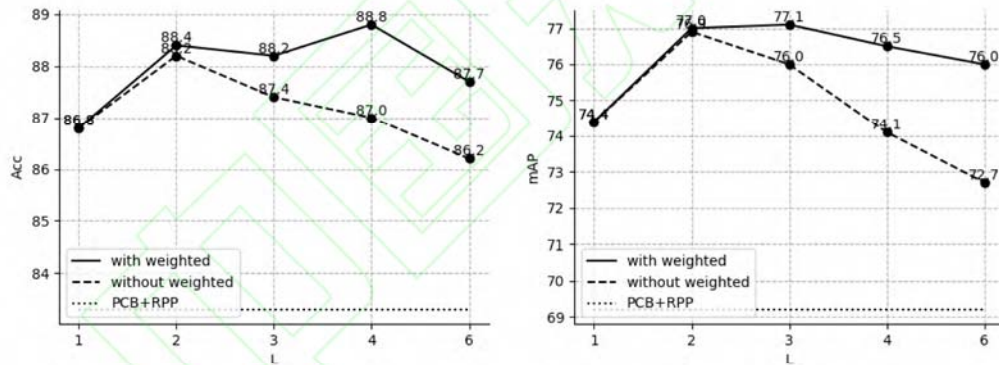


Fig 3 Comparison of results on DukeMTMC with and without weighting method

图 3 DukeMTMC 数据集上特征加权与直接拼接性能比较

显然, 经过式 (1.3) 方式进行加权的特征性能比全局特征和局部特征未经加权、直接拼接后的性能无论在 rank1 方面还是 mAP 方面的性能都要好很多。

3.4 特征加权系数消融研究

对于上文提及的特征加权方式, 即 (1.3) 式: 对局部分支全局 *feature map* 进行等分成 L 份, 每一份在特征测试时予以 $1/\sqrt{L}$ 的加权因子, 为进一步说明其合理性, 我们适当调节了每一局部特征的加权系数。对于局部 *feature map* 进行 L 均匀划分后的特征在测试时, 予以 $1/\sqrt{\alpha}$ 的加权系数, 其中 $L-1 \leq \alpha \leq L+1$ 。即新的加权公式如下:

$$f = f_{global} + \sum_{i=1}^L \frac{1}{\sqrt{\alpha}} f_{local_i} \quad (1.4)$$

此外还比较了以 $1/L$ 、 $1/L^2$ 的形式对特征进行

加权。即:

$$f = f_{global} + \frac{1}{L} \sum_{i=1}^L f_{local_i} \quad (1.5)$$

$$f = f_{global} + \frac{1}{L^2} \sum_{i=1}^L f_{local_i} \quad (1.6)$$

对于不同的划分 L 及性能比较, 实验结果见图 4、5。

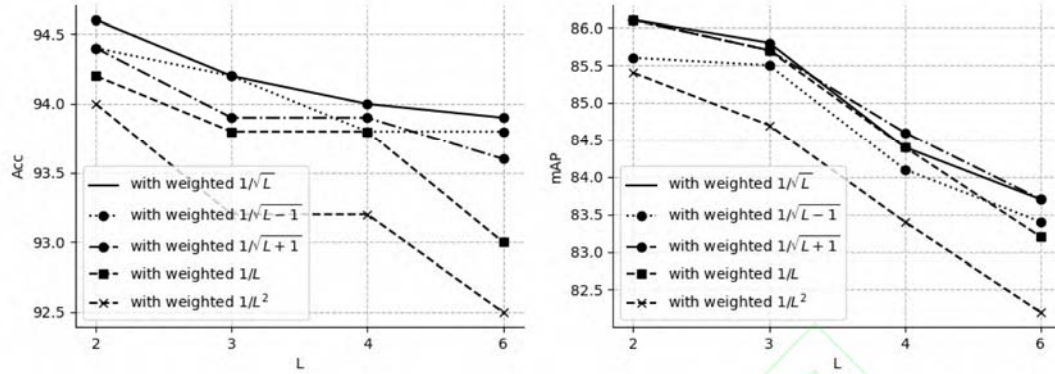


Fig 4 Comparison of results on Market1501 with different weighting method
图 4 Market1501 数据集上各种加权方式性能对比

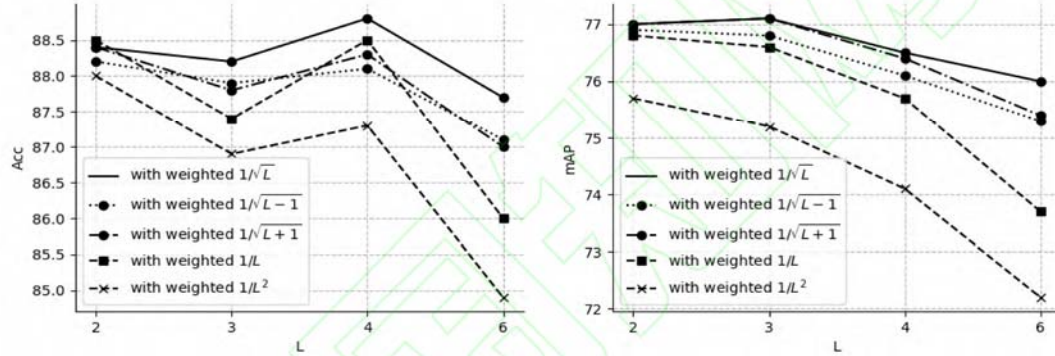


Fig 5 Comparison of results on DukeMTMC with different weighting method
图 5 DukeMTMC 数据集上各种加权方式性能对比

在 Market1501、DukeMTMC-Reid 上的实验结果表明:对网络下分支提取的 feature map 进行 L 份的均匀分割,并以 $1/\sqrt{L}$ 的系数进行加权是合理的,提高了模型的准确率,各种加权方式比较,可以看出,这种加权方式有利于提高模型的泛化性能。

3.5 其他特征融合方式探讨

本文算法是用于 PCB 的后处理阶段,即对 PCB 网络提取后的各分路特征进行加权处理。该加权处理的逻辑在于: global 分支特征矢量与 local 分支 L 个分部特征矢量 (part-level features) 之间的信息表达能力有显著差异,存在通过加权提升性能的空间。本文所考虑 PCB 网络的两个分支处于同一网络层级,通过这样的加权后处理方式,可以提高模型的性能。目前文献报道的其他多分支网络,典型如 MGN^[17] 算法,它们各个分支是都包含 global 特征,因此多分支融合时并没有对各分支进行加权融合,而是直接拼接各个分支的特征矢量用于最终性能评估。Pyramid^[18] 则采用了多路金字塔式的多尺度特征,最终也是采用直接拼接方式来融合各路特征矢量。由于 MGN 和 Pyramid 的特征提取方式以及融合条件与基本的 PCB 网络差距加大,本文的加权融合方式很难用于 MGN 和 Pyramid 算法。本论文的

分析与仿真表明:在论文的 PCB 网络框架下,各支路直接拼接融合的方案性能要明显差于本文提出的加权融合方式。

3.6 实验结果与主流方法的比较

为了验证本文所提模型的有效性,在两个主流数据集上对基于本身所提模型的行人重识别算法进行实验,并于最近比较先进的算法进行比较。表 4 和 5 分别给出本文所提算法与其他先进算法在 Market1501 和 DukeMTMC-Reid 上的对比结果。

从表 4、5 中可以看出本文所提的模型在 Market1501 数据集上的 rank1 和 mAP 分别为 94.6% 和 86.1%。在 DukeMTMC-Reid 数据集上的 Rank1 和 mAP 分别达到了 88.8% 和 77.0%, 本文所提模型的 Rank1 和 mAP 比 PCB 等先进算法均有显著的提高。本文所提模型不仅关注局部特征,而且加入了全局特征,考虑了局部特征与全局特征对最终网络模型的性能,对它们分别进行了加权操作,全局特征弥补了局部特征在划分及对齐问题,相比与 PCB 算法,提高了模型的性能。加权后的特征具有更加丰富的信息,使行人的特征描述变得更有辨别力。

对于基于多特征融合的 MGN 和 Pyramid 算法，它们除了与 PCB 类似的特征分割方式，还都引入了基于度量学习的三元组损失，以及动态训练等其他技术，故其性能要更优。本文主要验证加权算法，没有更多的加入基于度量学习等其他技术。

表 4 在 Market1501 数据集上的结果对比

Method	R1	R5	R10	mAP
HSP ^[12]	93.7	97.6	98.4	83.4
AANet ^[13]	93.9	—	—	83.4
IANet ^[14]	94.4	—	—	83.1
CASN ^[15]	94.4	—	—	82.8
AutoReid ^[16]	94.5	—	—	85.1
Bag of tricks ^[8]	94.5	—	—	85.9
CAMA ^[7]	94.7	—	—	84.5
MGN ^[17]	95.7	—	—	86.9
Pyramid ^[18]	95.7	98.4	99.0	88.2
PCB ^[3]	92.3	97.2	98.2	77.4
PCB+RPP ^[3]	93.8	97.5	98.5	81.6
Ours	94.6	97.6	98.4	86.1

表 5 在 Duke 数据集上的结果对比

Method	R1	R5	R10	mAP
AANet ^[13]	87.7	—	—	74.3
IANet ^[14]	87.1	—	—	73.4
CASN ^[15]	87.7	—	—	73.7
AutoReid ^[16]	88.5	—	—	75.1
Bag of tricks ^[8]	86.4	—	—	76.4
CAMA ^[7]	85.8	—	—	72.9
MGN ^[17]	88.7	—	—	78.4
Pyramid ^[18]	89.0	—	—	79.0
PCB ^[3]	81.7	89.7	91.9	66.1
PCB+RPP ^[3]	83.3	90.5	92.5	69.2
Ours	88.8	94.8	96.2	76.5

4 结论

本文在经典 PCB 行人重识别算法的基础上，融入了图像的全局特征，用以弥补局部特征划分时带来的特征对齐问题，因此不需要额外的 RPP 网络进行特征对齐。在性能测试阶段，采用一种特征加权的方式融合全局特征与局部特征信息。本文在两个不同的行人重识别数据集上进行了实验，实验表明，融入全局特征、对全局特征与局部特征进行加权融合，模型提取的行人特征具有更强的判别性，改善了 PCB 网络的性能。

参考文献

[1] LIAO S, HU Y, ZHU X, et al. Person re-identification by local maximal occurrence representation and metric

learning[C]//Proceedings of the IEEE conference on computer vision and pattern recognition.

- [2] 甄俊杰, 应自炉, 赵毅鸿, 等. 深度学习和迭代量化在图像检索中的应用研究[J]. 信号处理, 2019(5): 919-925.
Zhen Junjie, Ying Zilu, Zhao Yihong, et al. Application research of deep learning and iterative quantization in image retrieval. Journal of Signal Processing, 2019, 35(5): 919-925.(in Chinese)
- [3] SUN Y, 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).
- [4] VARIOR R R, SHUAI B, LU J, et al. A siamese long short-term memory architecture for human re-identification[C]//European conference on computer vision. Springer, 2016: 135-153.
- [5] WEI L, ZHANG S, YAO H, et al. Glad: Global-local-alignment descriptor for pedestrian retrieval[C]//Proceedings of the 25th ACM international conference on Multimedia.
- [6] 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.
- [7] YANG W, HUANG H, ZHANG Z, et al. Towards rich feature discovery with class activation maps augmentation for person re-identification[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition.
- [8] LUO H, GU Y, LIAO X, et al. Bag of tricks and a strong baseline for deep person re-identification[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops.
- [9] ZHONG Z, ZHENG L, KANG G, et al. Random erasing data augmentation[J]. arXiv preprint arXiv:1708.04896, 2017.
- [10] ZHENG L, SHEN L, TIAN L, et al. Scalable person re-identification: A benchmark[C]//Proceedings of the IEEE international conference on computer vision.
- [11] RISTANI E, SOLERA F, ZOU R, et al. Performance measures and a data set for multi-target, multi-camera tracking[C]//European Conference on Computer Vision. Springer, 2016: 17-35.
- [12] KALAYEH M M, BASARAN E, GÖKMEN M, et al. Human semantic parsing for person

- re-identification[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition.
- [13] 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.
- [14] HOU R, MA B, CHANG H, et al. Interaction-and-aggregation network for person re-identification[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition.
- [15] ZHENG M, KARANAM S, WU Z, et al. Re-identification with consistent attentive siamese networks[C]//Proceedings of the IEEE conference on computer vision and pattern recognition.
- [16] QUAN R, DONG X, WU Y, et al. Auto-ReID: Searching for a part-aware ConvNet for person re-identification[C]//Proceedings of the IEEE International Conference on Computer Vision.
- [17] WANG G, YUAN Y, CHEN X, et al. Learning Discriminative Features with Multiple Granularities for Person Re-Identification[J]. 2018 ACM Multimedia Conference on Multimedia Conference - MM '18, 2018: 274–282. DOI:10.1145/3240508.3240552.

- [18] ZHENG F, DENG C, SUN X, et al. Pyramidal person re-identification via multi-loss dynamic training[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition.

作者简介



赵师亮 男，1994 年生，江苏连云港人。南京邮电大学通信与信息工程学院硕士研究生。主要研究方向为行人重识别
E-mail: 1018010632@njupt.edu.cn



吴晓富 男，1975 年生，江苏南京人。博士，南京邮电大学教授。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。
E-mail: xfuwu@njupt.edu.cn



张索非 男，1982 年生，江苏镇江人。博士，南京邮电大学讲师。研究方向包括图像与视频信号处理、机器学习、物联网技术等。
E-mail: zhangsuofei@njupt.edu.cn