

# 基于特征拼接的行人重识别

潘通, 李文国\*

昆明理工大学机电工程学院, 云南 昆明 650500

**摘要** 基于卷积神经网络算法对行人进行多特征提取, 并使用拼接后的多特征对行人进行特征表征。使用全局池化和多个卷积构建多分支结构, 利用多分支结构来弥补丢失的信息。为了减小过拟合, 采用自行设计的瓶颈层代替模型中的分类层。实验时, 分别在 Market1501、CUHK03、DukeMTMC-Reid 数据集上对本文所提算法进行验证。在 Market1501 数据集上, 本文所提算法预测正确的概率(Rank1)为 95.2%, 平均预测均值(mAP)为 86.0%。实验结果表明, 本文所提算法提取的行人特征具有较强的辨别力, 识别准确率明显高于其他先进的算法。

**关键词** 光计算; 卷积神经网络; 行人重识别; 多特征; 特征拼接

中图分类号 TP391

文献标识码 A

doi: 10.3788/LOP56.162001

## Person Re-Identification Based on Feature Stitching

Pan Tong, Li Wenguo\*

*Faculty of Mechanical & Electrical Engineering, Kunming University of Science and Technology,  
Kunming, Yunnan 650500, China*

**Abstract** A convolutional neural networks based algorithm is proposed to extract multiple features from a single person. Further, the feature representation of a person using spliced multi-features is also proposed. Initially, the multi-branch structure is constructed using global pooling and multiple convolution; this multi-branch structure is used to offset the information loss. Subsequently, the bottleneck layer is designed to replace the classification layer in the model to reduce overfitting. In the experiment, the proposed algorithm is verified using the Market1501, CUHK03, and DukeMTMC-Reid datasets. In Market1501, the proposed algorithm achieves the first correct prediction probability (Rank1) of 95.2% and mean average precision (mAP) of 86.0%. The experimental results indicate that the proposed algorithm can extract discriminative features. Furthermore, the recognition accuracy of the proposed algorithm is significantly better than that of other advanced algorithms.

**Key words** optics in computing; convolutional neural network; person re-identification; multiple features; feature stitching

**OCIS codes** 200.4560; 150.0155; 100.4996; 100.4999

## 1 引言

行人重识别是一种给出感兴趣的行人图像的技术, 其主要目的是在由不同摄像头或同一摄像头不同时段拍摄的图像数据库中找出与目标行人匹配的行人图像。行人重识别的具体做法是: 1) 要构建一个提取行人图像特征的模型; 2) 利用距离度量给出相似性排序。行人重识别在很多领域均有应用, 如视频监控、人机交互、机器人、基于内容的视频检索

等。亮度、光照等因素的变化会给行人图像带来巨大的影响, 因此行人重识别仍然面临着挑战。

最早研究者是对手工设计的特征进行识别<sup>[1-2]</sup>。随着深度学习的发展, 目前行人重识别的主要做法是基于深度学习进行识别。基于深度学习的行人重识别研究方法主要包括基于度量学习的方法和基于卷积神经网络特征学习的方法。基于度量学习方法的主要思想是利用行人之间的特征距离来解决识别问题, 通过三元组函数<sup>[3]</sup>来使相同行人的类内距离

收稿日期: 2019-02-19; 修回日期: 2019-03-07; 录用日期: 2019-03-21

基金项目: 云南省自然科学基金(KKSY201301070)

\* E-mail: 475284191@qq.com

尽可能小,使不同行人的类间距离尽可能大。Hermans 等<sup>[4]</sup>发现三元组损失函数的正、负样本对的选择很随机,于是提出了基于难样本挖掘的三元组。Hermans 等通过在一个批次中选择最不像的相同行人来构建正样本对,选择同一个批次中最像的不同行人来构建负样本对。为了进一步减小类内方差并增大类间方差,Chen 等<sup>[5]</sup>提出四元组损失函数。基于度量学习的方法有一个比较大的缺点,即训练容易过拟合。目前,基于度量学习的方法主要用于多任务学习<sup>[6-7]</sup>,很少单独使用。基于卷积神经网络特征学习的方法是找到能够较好地表征行人特征的模型。陈兵等<sup>[8]</sup>通过加强对样本监督信息的利用和提升对样本间距离关系的约束,使神经网络能够得到判别性强的特征。由于单一特征对行人特征描述不够完整,很多研究者把重点由单一特征转移到多特征<sup>[9-12]</sup>。文献<sup>[10]</sup>中将描述行人的特征变成全局特征和局部特征融合后的特征,利用行人的全局特征和局部特征进行行人重识别。Sun 等<sup>[11]</sup>把行人图片水平切块,然后再利用 RPP (Refined Part Pooling)<sup>[11]</sup>算法使每个相似的块对齐,每个相似的块均接入损失函数进行训练,通过拼接多个局部特征的方法来得到一个判别性强的行人特征。文献<sup>[12]</sup>中通过语义分割把行人图片分成头部、背景、上半身、下半身和鞋子 5 个部分,然后通过拼接 5 个部分的特征对行人图像进行特征表征。上述文献中的模型都是通过增加行人的特征来提高识别准确率。

为了提高模型对行人特征表征的效果,本文也利用多特征对行人特征进行描述。上述文献主要关注全局特征和局部特征或者全局特征和属性特征<sup>[9-12]</sup>,本文则是通过多次提取全局特征来提高模型对行人特征的描述效果,通过对行人特征进行多次提取,减少了信息的缺失。本文所提出的模型在 Market1501、CUHK03、DukeMTMC-Reid 数据集上进行了实验,实验结果表明本文所提出的模型能够很好地对行人特征进行描述。

## 2 基本原理

行人重识别在测试时不同于其他的图像识别,行人重识别是用卷积神经网络提取图像中行人的特征,依据特征值的相似程度给出相似性排序。目前研究者都在寻找一种能够对行人特征具有较强辨别力的模型。受集成学习和 Inception<sup>[13]</sup>网络的启发,本文提出利用全局池化和多个卷积对行人进行多次全局特征提取的方法。

### 2.1 网络结构

当分类效果不好时,研究者普遍会采用集成学习的方法。集成学习是将多个单一模型集成在一起使用,即先单独训练多个单一模型,测试时再将训练的多个单一模型进行组合使用,形成一个先进的组合模型<sup>[14]</sup>。如利用 Resnet、Inception 等不同的卷积神经网络,在测试预测结果时利用某种策略将单一模型的预测结果组合在一起,组合方式一般有平均法和学习法。Inception 网络成功的原因在于其扩充了模型的宽度<sup>[13]</sup>,主要思想是对同一个输入运用 Inception 块进行池化、多个卷积,然后拼接在一起作为输出,如图 1 所示。这样的操作相比单独一个卷积操作能够对输入进行更好的特征描述。本文在 Resnet50 conv5\_x 后运用 Inception 块,且受集成学习思想的启发,没有在输出时进行拼接,而是将每个卷积和池化操作分别接入损失函数。由于  $3 \times 3$  和  $5 \times 5$  卷积操作计算量比较大,后期的 Inception 块用  $1 \times 3$  和  $3 \times 1$  代替。为了加快测试的速度,依据图 1,本文只设计了步长为 1 的  $1 \times 1$  卷积、全局池化和步长为 2 的  $1 \times 1$  卷积三个操作,并将这三个操作中的每个操作看成是集成学习中的单一模型,但不是单独训练,而是将每个操作接入损失函数,进行联合优化。

行人重识别利用模型提取特征,并依据特征值进行相似性排序。设步长为 1 的  $1 \times 1$  卷积、全局池化和步长为 2 的  $1 \times 1$  卷积三个操作来提取特征,特征值分别为  $n_1$ 、 $n_2$  和  $n_3$ 。目前在测试时,研究者普遍使用全局池化后的  $n_2$  进行相似性排序,本文则是把  $n_1$ 、 $n_2$  和  $n_3$  拼接在一起对行人的相似性进行排序,即行人的特征描述变成  $n = [n_1, n_2, n_3]$ ,这使得行人的特征描述更加丰富。

本文模型如图 2 所示,分支 1、2、3 分别对应上面的全局池化、步长为 1 的  $1 \times 1$  卷积、步长为 2 的  $1 \times 1$  卷积三个操作,分支 1、2、3 联合优化不但能够使它们优势互补,而且能够避免模型训练过程中信息的丢失。将分支 1、2、3 提取到的特征拼接起来对行人特征进行描述,使用拼接后的特征作为判断依据更加可靠。Szegedy 等<sup>[13]</sup>在 Inception 网络的中间层加入辅助分类器,减小了梯度消失,故本文也在模型中间层加入辅助分类器,即图 2 中的 Auxiliary classifier。分支 1 用瓶颈层 1 代替分类层,辅助分类器也是使用瓶颈层 1 代替分类层。Hermans 等<sup>[4]</sup>指出目前 Resnet50 的分类层不能对行人进行很好的描述,且过拟合比较严重,为减小过拟合,

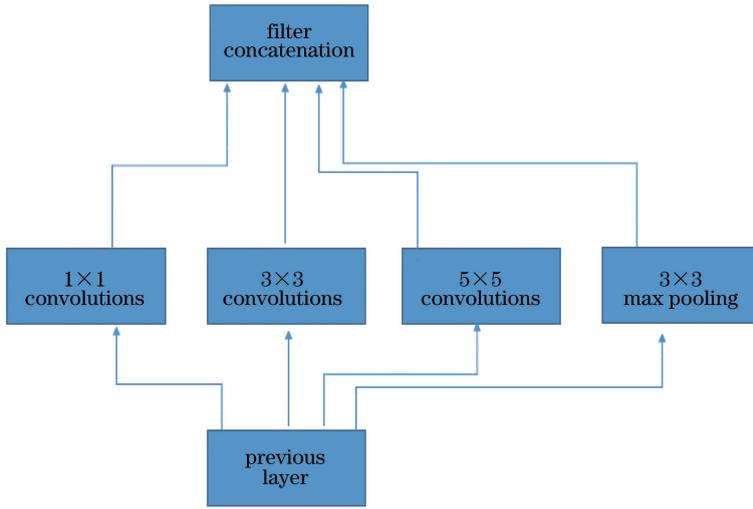


图 1 Inception 块

Fig. 1 Inception block

他们重新设计了瓶颈层,大幅提高了行人重识别的准确率。鉴于此,本文也设计了瓶颈层,即瓶颈层1,具体为:全局平均池化→归一化层→激活层→随机失活→全连接层→归一化层→全连接层。步长为1的1×1卷积和步长为2的1×1卷积提取的行人特征经过全局平均池化和分类层后分别集中到瓶颈

层2和3,最后接入损失函数,如图2所示。分支2中的瓶颈层2为:步长为1的1×1卷积层→归一化层→全连接层。分支3中的瓶颈层3为:步长为2的1×1卷积层→归一化层→全连接层。文献[15]中指出分类层中全连接层的偏置项会影响分类任务的效果,故需将瓶颈层中全连接层的偏置项去除。

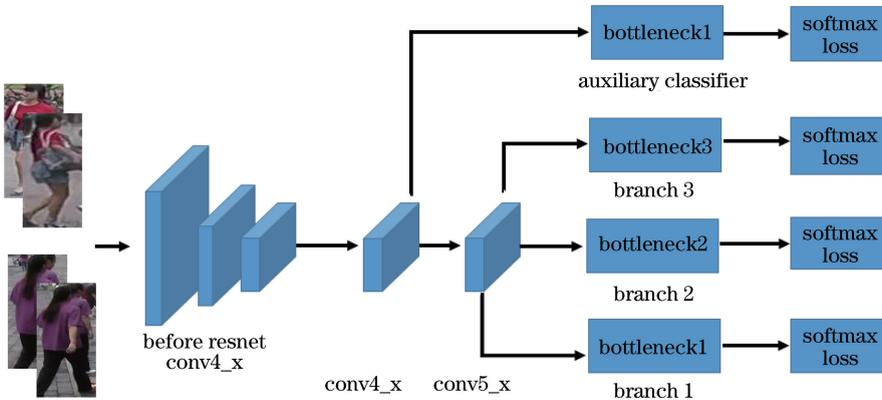


图 2 网络结构

Fig. 2 Network structure

## 2.2 损失函数

目前,Softmax 损失在各种行人重识别模型中被广泛使用。不同于传统的 Softmax 损失,根据文献[15],本文所采用的 Softmax 损失的形式去除了线性多类分类器中分类层的偏置项,这有助于提高算法的识别能力。Softmax 损失的定义为

$$L_{\text{softmax}} = - \sum_{i=1}^n \log \frac{\exp[\mathbf{W}_{y_i}^T f(\mathbf{b}_i)]}{\sum_{j=1}^c \exp[\mathbf{W}_j^T f(\mathbf{b}_j)]}, \quad (1)$$

式中: $n$  为批量的大小; $c$  为类别数; $f(\mathbf{b}_i) \in \mathbf{R}^d$  为

第  $i$  个样本的深度特征,其类别为  $y_i$ , 维度为  $d$ ;  $\mathbf{W}_j$  为第  $j$  层的权重向量;  $\mathbf{b}_j$  为第  $j$  层的偏置向量;  $T$  为转置符号。

## 3 实验结果与分析

### 3.1 数据集

本文所提算法在 Market1501、CUHK03、DukeMTMC-Reid 数据集上进行了实验,给出每个数据集上的首次预测正确的概率(Rank1)、平均预测均值(mAP)。Market1501 数据集包括由 6 个摄

像头(5个高清摄像头和1个低清摄像头)拍摄到的1501个行人和32668个检测到的行人矩形框。训练集有751个行人,包含12936张图像,平均每个人有17.2张训练图片;测试集有750个行人,包含19732张图像,平均每个人有26.3张测试图片。CUHK03数据集含有由5对摄像头采集得到的1467个行人。数据集在使用时一般不采用旧版的测试方式,本文将数据集分为包含767个行人的训练集和包含700个行人的测试集。DukeMTMC-Reid数据集是由8个不同摄像头采集的视频序列经过等间隔采样得到的,共有36411张图片,包含1812类行人。将该数据集中的16522张图片作为本文的训练集,19889张图片作为测试集。

### 3.2 实验设置

所提算法是在Pytorch深度学习框架上实施的。卷积神经网络采用的是Resnet50,初始学习率为0.01,Bottleneck学习率为0.1,epochs为260,训练批次大小为 $n=C \times D$ ,其中 $n$ 为32, $C$ 为该批次中行人类别数,其值为8, $D$ 为该批次中每个行人的图片数,其值为4。分别在第60个epoch、160个epoch和240个epoch时,学习率均减小90%。优化器选择的是带动量的SGD,动量系数为0.9,正则化系数为0.0005。训练损失是三分支损失之和,然后再加上带有加权系数的辅助分类器损失。训练数据增强方式为旋转、裁剪。

### 3.3 参数对实验的影响

本文所提出的模型用瓶颈层代替分类层以尽可能地避免过拟合现象。瓶颈层中的参数需要合理设置,这些参数包括瓶颈层1中第一个全连接层神经元个数以及瓶颈层2、瓶颈层3中的卷积核个数。在本文中,全连接层神经元个数和卷积核个数取相同值,这主要是为了保持测试进行特征拼接时每个分支的特征值的维数相同。将全连接层神经元个数和卷积核个数统称为参数 $N$ 。表1中给出当 $N$ 取值不同时,本文所提出的算法在Market1501数据集上的实验结果。通过对比表1中的数据,发现当 $N$ 为512时的实验结果最好,即当瓶颈层1中第一个全连接层神经元个数为512以及瓶颈层2、瓶颈层3中卷积层的卷积核个数也为512时,实验结果最好。

Inception网络对辅助分类器设置了损失加权系数,故本文在使用时也设置了加权系数 $M$ 。Inception网络中的加权系数是在 $0 \sim 1$ 之间取值,本文也采取相同的取值方式。表2给出了加权系数

表1 不同 $N$ 的实验结果

Table 1 Experimental results corresponding to different  $N$  %

$N$	Rank1	mAP
2048	94.2	84.9
1024	94.8	85.7
512	95.2	86.0
256	94.6	85.5

表2 不同 $M$ 时的实验结果

Table 2 Experimental results corresponding to different  $M$  %

$M$	Rank1	mAP
1.0	95.2	86.0
0.8	94.5	85.6
0.6	94.6	85.8
0.4	94.3	85.3
0.2	94.0	85.4

取不同值时的实验结果,该实验是在Market1501数据集上进行。由表2可以看出,当加权系数为1时实验效果最好。

### 3.4 多特征拼接模型

为了验证本文所提出的模型的有效性,在Market1501数据集上对其进行实验验证。表3中方法1是基于Resnet50的原始模型的方法,方法2是基于Resnet50的原始模型并用本文所提出的瓶颈层代替分类层单独训练的方法,方法3是基于Resnet50的原始模型,不但用瓶颈层代替分类层单独训练而且去除全连接层中偏置项的方法。方法2和3的准确率比方法1均有显著的提升,从而证明本文所提出的瓶颈层代替分类层、去除全连接层偏置项两项改进是有效的。

表3 基础模型实验结果

Table 3 Experiment results of basic model %

Method	Rank1	mAP
1 (Resnet50)	87.2	72.3
2 (Resnet50+bottleneck layer)	89.6	73.4
3 (Resnet50+bottleneck layer+bias removing)	90.5	75.1

表4中方法E代表只使用分支1训练,方法F代表使用分支1、分支2一起训练,方法H代表分支1、分支2、分支3联合优化,方法J代表分支1、分支2、分支3、辅助分类层一起训练。本来辅助分类层只在训练时使用,在测试时不使用,后来研究者发现在测试时加上辅助分类层提取的特征后Rank1提高了0.7%,mAP提高了1.3%,故研究者在后面的实验测试时均加上辅助分类层提取的特征。

表4 不同方法的实验结果

Table 4 Experimental results of different methods %

Method	Rank1	mAP
E(branch 1)	90.5	75.1
F(branches 1,2)	93.8	83.3
H(branches 1,2,3)	93.9	84.2
J(branches 1,2,3 and auxiliary classifier)	95.2	86.0

表4中方法F的Rank1和mAP比方法E分别提高了3.3%和8.2%；方法H的Rank1和mAP比方法E分别提高了3.4%和9.1%。方法J是基于本文所提出的完整模型的方法，方法J的Rank1和mAP比方法E分别提高了4.7%和10.9%。这是因为方法J的模型能够对行人提取更丰富的特征，减少了训练过程中的信息丢失，比方法E(单一的全局特征)更有辨别力。表4中基于辅助分类器模型的方法J比方法H的Rank1和mAP分别提高了1.3%和1.8%。可以看出，每个分支对提高实验结果都是有效的。上述实验均是基于Market1501数据集进行的。

### 3.5 测试运行时间

目前研究者普遍使用全局池化后的2048维特征对算法进行运行时间测试。本文在测试时使用3个分支及辅助分类器各自提取512维特征，并将其拼接在一起，最后得到2048维特征。分支1使用瓶颈层中第一个全连接层后的512维特征，分支2和3均使用瓶颈层中卷积后的512维特征，辅助分类器也是使用瓶颈层中第一全连接层后的512维特征。表5列出了最基础的模型Resnet50、本文模型和PCB+RPP<sup>[11]</sup>在Market1501测试集上运行时所花费的时间，Market1501测试集有750类行人，包含19732张图像。在Market1501测试集上，本文模型相比基础模型Resnet50慢了27s，但是Rank1提高了8%；本文模型比PCB+RPP<sup>[11]</sup>快了6s，Rank1也提高了1.4%。

表5 Market1501测试集上测试时间

Table 5 Test time on Market1501 dataset

Method	Test time /s
Resnet50	179
Ours	206
PCB+RPP <sup>[11]</sup>	212

### 3.6 实验结果与主流算法比较

为了验证本文所提模型的有效性，在三个主流数据集上对基于本文所提模型的行人重识别算法进行实验，并与最近比较先进的算法进行对比。表6、7和8分别给出本文所提算法和其他先进算法

表6 在Market1501数据集上的结果对比

Table 6 Comparison of results on Market1501 %

Method	Rank1	Rank5	Rank10	mAP
MLFN <sup>[16]</sup>	90.0	—	—	74.3
PSE+ECN <sup>[17]</sup>	90.4	94.5	96.8	80.5
HA-CNN <sup>[18]</sup>	91.2	—	—	75.7
MACS <sup>[7]</sup>	93.1	—	—	82.3
DNN_CRF <sup>[19]</sup>	93.5	—	—	81.6
HSP <sup>[12]</sup>	93.7	97.6	98.4	83.4
PCB+RPP <sup>[11]</sup>	93.8	97.5	98.5	81.6
Ours	95.2	98.0	98.8	86.0

表7 在DukeMTMC-Reid数据集上的结果对比

Table 7 Comparison of results on DukeMTMC-Reid %

Method	Rank1	Rank5	Rank10	mAP
AOS <sup>[20]</sup>	79.2	—	—	62.1
HA-CNN <sup>[18]</sup>	80.5	—	—	63.8
GSRW <sup>[21]</sup>	80.7	88.5	90.8	66.4
MLFN <sup>[16]</sup>	81.0	—	—	62.8
DuATM <sup>[22]</sup>	81.8	90.2	—	64.6
PCB+RPP <sup>[11]</sup>	83.3	90.5	92.5	69.2
Ours	86.7	93.5	95.5	74.0

表8 在CUHK03数据集上的结果对比

Table 8 Comparison of results on CUHK03 %

Method	Rank1	Rank5	Rank10	mAP
HA-CNN <sup>[18]</sup>	41.7	—	—	38.6
MLFN <sup>[16]</sup>	52.8	—	—	47.8
PCB+RPP <sup>[11]</sup>	63.7	80.6	86.9	57.5
Ours	68.3	84.7	90.5	64.6

在Market1501、DukeMTMC-Reid和CUHK03的对比结果。

本文所提模型在Market1501的Rank1和mAP分别为95.2%和86.0%。从表6可以看出本文所提模型的Rank1和mAP比PCB+RPP<sup>[11]</sup>、HSP<sup>[12]</sup>等先进算法均有显著的提高。PCB+RPP<sup>[11]</sup>使用的多判别性特征是局部特征，而局部特征在对齐时存在很大的挑战。本文所提模型只关注全局特征，通过多个分支提取全局特征弥补了训练过程中的信息丢失，而且通过使用瓶颈层代替分类层减小了过拟合，因此能够提取行人更丰富的特征，使行人的特征描述变得更有辨别力。从表7可以看出，本文所提模型在DukeMTMC-Reid的Rank1和mAP分别达到了86.7%和74.0%，相比其他算法有明显提高。从表8可以看出，本文所提模型在CUHK03的Rank1和mAP分别达到了68.3%和64.6%，同样也超过了其他的先进算法。实验结果表明本文通过拼接多分支提取的行人特征能够很好

地对行人进行重识别。

图3是本文所提模型实际测试时的效果图。在图3中,左侧是查询图像,右侧绿色外框代表正确识别

别,红色外框代表错误识别。从图3可以看出,本文所提模型在 Rank1 到 Rank5(预测5次)均能做到正确的识别。



图3 识别效果

Fig. 3 Recognition effects

## 4 结 论

提出一种提取多特征的行人重识别模型,利用拼接后的多特征对行人进行特征表征。基于集成学习的思想设计了多分支来提取多特征。用所设计的瓶颈层代替原先卷积神经网络的分类层。本文所提模型不仅能提取判别力强的特征,还避免了特征学习过程中的信息丢失,从而提高了行人重识别的准确率。在Market1501数据集上使用RK<sup>[23]</sup>以后,本文算法的 Rank1 可以达到 95.6%,mAP 可以达到 93.1%,实验结果表明本文所提模型能够对行人提取判别性强的特征。本文所提出的模型相比其他模型加入 LSTM、图像水平分割等操作,可以做到端到端的训练。

## 参 考 文 献

- [1] Sun X W, Xu Q S, Cai Y, *et al.* Sea sky line detection based on edge phase encoding in complicated background[J]. *Acta Optica Sinica*, 2017, 37(11): 1110002.  
孙熊伟, 徐青山, 蔡熠, 等. 基于边缘相位编码的复杂背景下海天检测[J]. *光学学报*, 2017, 37(11): 1110002.
- [2] Wang B, Su Y M, Wan L, *et al.* Sea sky line detection method of unmanned surface vehicle based on gradient saliency[J]. *Acta Optica Sinica*, 2016, 36(5): 0511002.

王博, 苏玉民, 万磊, 等. 基于梯度显著性的水面无人艇的海天线检测方法[J]. *光学学报*, 2016, 36(5): 0511002.

- [3] 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] // 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE, 2016: 1335-1344.
- [4] Hermans A, Beyer L, Leibe B. In defense of the triplet loss for person re-identification[J/OL]. (2017-11-21) [2019-01-05]. <https://arxiv.org/abs/1703.07737>.
- [5] Chen W H, Chen X T, Zhang J G, *et al.* Beyond triplet loss: a deep quadruplet network for person re-identification [C] // 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), July 21-26, 2017, Honolulu, HI, USA. New York: IEEE, 2017: 403-412.
- [6] Chen W H, Chen X T, Zhang J G, *et al.* A multi-task deep network for person re-identification [J/OL]. (2016-11-25) [2019-01-06]. <https://arxiv.org/abs/1607.05369>.
- [7] Wang C, Zhang Q, Huang C, *et al.* Mancs: a multi-task attentional network with curriculum sampling for person re-identification[M] // Ferrari V, Hebert M, Sminchisescu C, *et al.* Computer vision-ECCV 2018. Lecture notes in computer science. Cham: Springer, 2018, 11208: 384-400.

- [8] 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.  
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.
- [9] Liu H, Feng J S, Qi M B, *et al.* End-to-end comparative attention networks for person re-identification [J]. *IEEE Transactions on Image Processing*, 2017, 26(7): 3492-3506.
- [10] Wei L H, Zhang S L, Yao H T, *et al.* Glad: global-local-alignment descriptor for pedestrian retrieval[C]// *Proceedings of the 25th ACM International Conference on Multimedia*, October 23-27, 2017, Mountain View, California, USA. New York: ACM, 2017: 420-480.
- [11] 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. *Computer vision-ECCV 2018. Lecture notes in computer science*. Cham: Springer, 2018, 11208: 501-518.
- [12] Kalayeh M M, Basaran E, Gökmen M, *et al.* Human semantic parsing 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: 1062-1071.
- [13] Szegedy C, Liu W, Jia Y Q, *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.
- [14] Schapire R E. The boosting approach to machine learning: an overview[M]// Denison D D, Hansen M H, Holmes C C, *et al.* *Nonlinear estimation and classification*. New York, NY: Springer, 2003, 171: 149-171.
- [15] Wang F, Xiang X, Cheng J, *et al.* NormFace: L<sub>2</sub> hypersphere embedding for face verification [C] // *Proceedings of the 2017 ACM on Multimedia*, October 23-27, 2017, Mountain View, California, USA. New York: ACM, 2017: 1041-1049.
- [16] 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.
- [17] Sarfraz M S, Schumann A, Eberle A, *et al.* A pose-sensitive embedding for person re-identification with expanded cross neighborhood re-ranking [C] // 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, June 18-23, 2018, Salt Lake City, UT, USA. New York: IEEE, 2018: 420-429.
- [18] Li W, Zhu X T, Gong S G. Harmonious attention network 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: 2285-2294.
- [19] Chen D P, Xu D, Li H S, *et al.* Group consistent similarity learning via deep CRF 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: 8649-8058.
- [20] Huang H J, Li D W, Zhang Z, *et al.* Adversarially occluded samples 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: 5098-5107.
- [21] Shen Y T, Li H S, Xiao T, *et al.* Deep group-shuffling random walk 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: 2265-2274.
- [22] Si J L, Zhang H G, Li C G, *et al.* Dual attention matching network for context-aware feature sequence based 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: 5363-5372.
- [23] 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.