

引用格式:王林,杨乐. 嵌套池化三元组卷积神经网络的行人再识别[J]. 微电子学与计算机, 2019, 36(9): 73-78. [WANG L, YANG L. Pedestrian re-identification of nested pooling triple convolutional neural networks[J]. Microelectronics & Computer, 2019, 36(9): 73-78. ]

# 嵌套池化三元组卷积神经网络的行人再识别

王 林, 杨 乐

(西安理工大学 自动化与信息工程学院, 陕西 西安 710048)

**摘 要:** 针对行人再识别易受遮挡, 光照、视角等非理想条件变化的影响, 提出一种嵌套池化三元组卷积神经网络. 它在均方根池化后依次添加平均池化和最大池化提取全局特征, 并通过最短路径损失自动对齐局部特征, 然后采用改进型 Log-logistic 函数代替传统三元组损失函数训练网络, 得到与局部特征联合优化的全局特征. 在 Market-1501、CUHK03 和 VIPeR 数据集上的识别率都比基于传统方法的提高了 5% 以上. 实验结果表明, 本文提出的嵌套池化三元组卷积神经网络, 能有效解决非理想自然条件下存在的部分遮挡、分辨率低和旋转变换等问题, 同时具有良好的泛化能力和适用范围.

**关键词:** 行人再识别; 嵌套池化; 三元组损失函数; 局部特征; 间接度量

**中图分类号:** TP301.6

**文献标识码:** A

**文章编号:** 1000—7180(2019)09—0073—06

## Pedestrian re-identification of nested pooling triple convolutional neural networks

WANG Lin, YANG Le

(School of Automation and Information Engineering, Xi'an University of Technology, Xi'an 710048, China)

**Abstract:** A nested pooling tri-tuple convolution neural network is proposed for pedestrian re-identification, which is vulnerable to occlusion, illumination, viewing angle and other non-ideal conditions. It adds the average pooling and maximum pooling to extract global features in turn after the root mean square pooling, and automatically aligns the local features with the shortest path loss. Then the improved Log-logistic function is used instead of the traditional triple loss function to train the network, and get global features jointly optimized with local features. The recognition rates on the Market-1501, CUHK03, and VIPeR datasets are both more than 6% higher than those based on traditional methods. The experimental results show that nested pooling tri-tuple convolutional neural network proposed in this paper can effectively solve the problems of partial occlusion, low resolution and rotation change under non-ideal natural conditions, and has good generalization ability and scope of application.

**Key words:** person re-identification; nested pooling; ternary loss function; local features; indirect metric

### 1 引言

行人再识别不仅能为公安部门提供预防和侦破商场盗窃、聚众斗殴等重大刑事案件的线索, 而且能够高效准确地定位和追踪寻人启事中走失的老人和

小孩, 已成为计算机视觉领域一个热点问题. 目前行人再识别所使用的特征主要有: 底层视觉与手工特征、语义属性特征和基于深度学习的特征. 底层视觉特征复杂性较高, 手工特征需要研究人员具有相当坚实的专业知识和大量的实践经验, 特征提取过程

较为主观,可移植性差.与低层视觉特征相比,语义属性具有判别性好以及变换不变性的优点,但很难检测到语义属性.深度学习灵活地在训练数据的驱动下,对输入图像逐级根据需求自动地提取有用的特征,相比人工设计的特征,对目标具有更强的表达能力.

基于深度学习的传统行人再识别方法将行人外观作为整体进行建模和匹配,但研究发现基于局部区域的分块匹配有更高的识别精度,所以基于分块匹配的方法成为近年来的研究热点.文献[1]将行人图像分割为头、躯干和腿,然后训练多分支卷积神经网络自适应地提取各部分的特征.虽然可以得到精确的局部区域特征,但分割存在准确率的问题.文献[2]提出可变形块度量学习方法,使对应块在匹配时可在一定范围内改变它们的位置(变形),以此来模拟姿态变化,并通过计算块的特征差异和位置变形成本得到行人特征相似度.但这需要额外的监督和姿势估计步骤,其通常容易出错.为了自适应对齐局部特征,文献[3]提出 AlignedReID 方法,通过引入最短路径损失自动对齐局部特征,得到较准确的局部距离,可有效克服行人图像身体结构错位,遮挡以及不完整等问题.

另外,研究发现三元组损失函数在行人再识别中具有更大的应用潜力.三元组损失函数通过将特征学习和相似性度量合并到一个统一的框架中,使网络根据最终的任务学习有辨识力的特征,不需要附加特定的层而可达到与设计新颖的网络识别架构相提并论的效果<sup>[4]</sup>.

因此,本文以三元组卷积神经网络为基础,利用 AlignedReID 方法计算局部距离.但仅利用该方法将会存在以下两大问题:(1)行人图像可能存在平移、尺度、旋转等变化,若只利用最大池化提取显著性特征往往会导致误匹配;(2)特征学习过程中值域是变化的,而传统三元组损失函数中的间隔参数需要事先指定,缺乏灵活性,不利于网络训练.为了解决上面的问题,本文提出嵌套池化三元组卷积神经网络.主要做了以下两个贡献:(1)利用 NIP(nested invariance pooling,嵌套不变池)理论<sup>[5]</sup>,合理地在卷积层后依次添加均方根池化层、平均池化以及最大池化层,有效提高特征对平移、尺度、旋转变化的鲁棒性;(2)改进三元组损失函数,以 Log-logistic 函数代替传统的铰链函数,同时优化损失函数的约束条件来加速收敛并改善性能.

## 2 改进的行人再识别方法

首先描述嵌套池化三元组卷积神经网络的特征提取过程和相似性距离计算方法,然后针对传统损失函数的不足,提出改进型三元组损失函数.最后,详细阐述了用于间接度量的置信度测量机制.

### 2.1 网络模型

如图 1 所示,以 Resnet50 的卷积层为基础,在最后一个卷积层后添加全局特征分支与局部特征分支,并利用三元组损失函数联合优化全局特征和局部特征.下面从全局特征、局部特征的提取过程和相似性距离计算方法两个方面进行介绍.

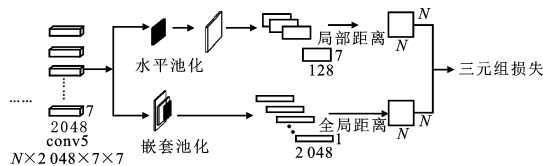


图 1 AlignedReID 模型框图

#### 2.1.1 全局特征提取过程

传统的网络结构往往只在卷积层后添加最大池化层来提取显著性特征,这样可能会使特征丢失很多有用的信息.然而文献[5]表明:均方根池化可以改善平移变化,平均池化可以改善尺度变化,最大池化可以改善旋转变换,利用嵌套不变池(Nested Invariance Pooling, NIP)理论,合理地在卷积层后依次添加均方根池化层、平均池化以及最大池化层,可有效提高特征对平移、尺度、旋转变化的鲁棒性.

因此,为了克服行人再识别易受遮挡,光照、旋转等非理想条件变化的影响,本文利用嵌套不变池理论改进传统的只利用最大池化层的网络结构.在提取全局特征的过程中,首先利用 Resnet50 的卷积层提取图像特征,输出大小为  $2\,048 \times 7 \times 7$  的特征图.2 048 是通道数,  $7 \times 7$  是空间大小.然后进行  $2 \times 2$  大小的均方根池化,移动步长为 1,得到  $2\,048 \times 6 \times 6$  大小的输出,再进行  $2 \times 2$  大小的平均池化,移动步长为 2,得到  $2\,048 \times 3 \times 3$  大小的输出,最后进行全局最大池化,得到 2 048 维的全局特征向量.

#### 2.1.2 局部特征提取及相似性距离计算

根据文献[6]的启发,为了提高特征在水平方向的鲁棒性,有效克服视角变化.在 Resnet50 最后一个卷积层的输出后执行水平池化,进行池化窗口为  $1 \times 7$  的最大池化,得到 7 个局部特征.然后利用  $1 \times$

1 的卷积将通道数从 2 048 减少到 128,最后每张图片得到  $128 \times 7$  的局部特征,局部特征(128 维向量)表示行人图像水平部分的信息。

在此基础上,利用 L2 距离计算全局特征间的相似度.局部距离来源于一个设想:对于身份相同的两个行人图像,第一个行人图像身体部位的局部特征与另一个图像语义对应身体部位的更相似,而他们在水平方向可能并不是一一对应的.因此,通过从上到下动态匹配局部块来对齐局部特征,从而得到两幅图像的最小局部距离.给定两幅行人图像的局部特征,  $F = \{f_1, \dots, f_7\}$  和  $G = \{g_1, \dots, g_7\}$ ,首先利用公式(1)将两幅图像局部特征之间的距离归一化到  $[0, 1)$ ,公式(1)如下所示:

$$d_{i,j} = \frac{e^{\|f_i - g_j\|_2} - 1}{e^{\|f_i - g_j\|_2} + 1} \quad i, j \in 1, 2, 3, \dots, 7 \quad (1)$$

式中,  $d_{i,j}$  是第一个图像第  $i$  个水平块与第二个图像第  $j$  个水平块之间的距离.这些距离形成距离矩阵  $\mathbf{D}$ ,它的第  $i$  行第  $j$  个元素是  $d_{i,j}$ .然后,定义行人图像间的局部距离为矩阵  $\mathbf{D}$  中从  $(1, 1)$  到  $(H, H)$  最短路径的距离总和.局部距离可以通过公式(2)动态计算,其中:  $S_{i,j}$  是距离矩阵  $\mathbf{D}$  中从  $(1, 1)$  到  $(i, j)$  最短路径的总和,  $S_{H,H}$  是距离矩阵  $\mathbf{D}$  中从  $(1, 1)$  到  $(H, H)$  最短路径的总和,即两幅图像的局部距离.

$$S_{i,j} = \begin{cases} d_{i,j} & i = 1, j = 1 \\ S_{i-1,j} + d_{i,j} & i \neq 1, j = 1 \\ S_{i,j-1} + d_{i,j} & i = 1, j \neq 1 \\ \min(S_{i-1,j}, S_{i,j-1}) + d_{i,j} & i \neq 1, j \neq 1 \end{cases} \quad (2)$$

非对应身体部位局部特征间 L2 距离相对较大,因此对最短路径的贡献很小,最短路径的绝大部分都来源于对应身体部位局部特征间的距离.因此,通过动态对齐计算局部特征间的距离,可以有效改善行人图像中身体结构错位、不完整,以及遮挡等缺陷的影响。

## 2.2 改进的三元组损失函数

在训练三元组卷积神经网络的过程中,不进行难样本挖掘会使网络相对快地学习较简单的(对训练网络提供很少价值)三元组样本,而在平均的过程中抵消少数有用样本组贡献,导致训练阻塞收敛效果不佳.但选择过难的三元组样本,又会引起训练不稳定收敛变难,而文献[4]介绍的小批量难样本挖掘策略,可有效克服上述缺陷,因此本文利用该策略选择三元组样本。

该策略的核心思想是:通过随机采样  $P$  个类(行人身份),并从每个类中随机选择  $K$  个图片,形成具有  $PK$  个图像的一个小批量.然后,对于小批量中的每一个样本  $X$ ,选择其在小批量中相似性距离最大的正样本和相似性距离最小的负样本,形成  $PK$  个有效的三元组样本来计算损失.称其为小批量难样本:

$$\begin{aligned} & \overbrace{\max_{p=1, \dots, K} D(f_\theta(x_a^i), f_\theta(x_p^i))}^{\text{hardest} \wedge \text{positive}} \\ & \underbrace{\min_{\substack{j=1, \dots, P \\ n=1, \dots, K \\ j \neq i}} D(f_\theta(x_a^i), f_\theta(x_n^j))}_{\text{hardest} \wedge \text{negative}} \end{aligned} \quad (3)$$

式中,  $X$  代表小批量中的任意样本,并且样本  $x_j^i$  对应于小批量中第  $i$  个行人的第  $j$  个图像.  $f_\theta(x)$  是经过卷积神经网络后样本图像的输出特征,  $D(f_\theta(x_a), f_\theta(x_p))$  和  $D(f_\theta(x_a), f_\theta(x_n))$  是样本特征间差异度.利用该策略选择的三元组样本不仅数量比传统方法增加了 3 倍,同时,因选择的三元组样本是数据集的一个小子集中最难样本,信息价值较高又不容易过拟合,因此可认为是最适合网络训练的样本。

其次,基于三元组卷积神经网络的传统行人再识别方法以铰链(hinge)函数为损失函数,但在特征自动学习的过程中其值域是变化的,而式中的间隔参数却是固定不变的,需要事先指定,训练中无法自适应调整,缺乏灵活性,同时可能因为间隔参数设定的不合理而影响网络训练.而文献[7]采用 Log-logistic 函数代替铰链函数,无需人工设定间隔参数,改进了特征与度量函数联合优化的效果.然而该方法只是让正样本对间相似度尽可能大,负样本对间相似度尽可能小,没有约束正样本对和负样本对之间的相似度差异,这可能会使三元组样本中出现正样本对间相似度比负样本对间更小的情况,显然,这不利于改善行人再识别性能。

为了克服上述缺陷,本文考虑到在正样本对差异度最小化的情况下,负样本对间的差异度越大,两者之间差异度的间隔也就越大,因此本文提出公式(4)所示的改进型三元组损失函数.该公式不仅能使正样本对特征间差异度最小化,负样本对特征间差异度最大化,而且使三元组样本中负样本对间差异度比正样本对间的差异度尽可能大.改进型损失函数旨在拉近属于同一行人身份的样本,同时在学习的特征空间中推远属于不同行人的样本.这与许多数据聚类 and 判别分析方法所使用的原理更加一致。

$$\text{loss} = -\frac{1}{PK} \sum_{i=1}^P \sum_{a=1}^K \left[ \lg(1 + \exp(\min_{\substack{j=1 \dots P \\ n=1 \dots K \\ j \neq i}} D(f_{\theta}(x_a^i), f_{\theta}(x_n^j)) - 1)) \max_{p=1 \dots K} D(f_{\theta}(x_a^i), f_{\theta}(x_p^i)) \right. \\ \left. + \lg(1 + \exp(-(\max_{p=1 \dots K} D(f_{\theta}(x_a^i), f_{\theta}(x_p^i)))) \right] \quad (4)$$

式中,  $PK$  是小批量中的样本个数,  $i$  指的是  $P$  个行人身份中的任意一个,  $a$  指的是  $k$  个相同身份行人图片中的任意一张图片. 从公式(4)可以看出, 为了优化该目标函数, 训练过程中会自动调整网络权重, 使正样本对间差异度最小化的同时, 负样本对间差异度比正样本对间的差异度更大, 而不再追求  $D(f_{\theta}(x_a^i), f_{\theta}(x_n^j)) - D(f_{\theta}(x_a^i), f_{\theta}(x_p^i)) \geq C$  这种硬性指标, 因此无需设置间隔参数  $C$ , 提高了目标函数优化的灵活性, 不仅能方便训练, 同时能有效提高特征与度量函数联合优化的能力.

### 2.3 置信度测量机制

对于表观特征较相似的不同行人, 如果仅利用直接度量进行两幅图像的相似度比较, 往往会出现负样本比正样本更匹配测试样本的情况. 为了有效避免误匹配, 充分利用与图像对相关的其他图像的判别信息, 本文利用置信度测量机制<sup>[8]</sup>进行间接度量.

置信度测量机制的主要思想: 画廊集图像中与测试行人图像匹配的目标行人, 需满足以下两个条件: (1) 与测试行人图像身份相同的画廊集图像应该位于相似性排名列表的前面, 因为其与测试行人图像的差异度较小; (2) 测试行人图像与目标图像之间的距离应该比目标图像与画廊集其余图像之间的距离都小, 因为目标图像与测试图像中的行人身份相同. 因此, 认为目标图像与其余画廊集图像之间的距离比其与测试图像之间的距离大的个数, 与该目标图像与测试图像身份相同的置信度得分成正相关. 具体计算过程如下.

首先, 根据卷积神经网络的输出特征, 计算测试图像和画廊集中所有行人图像之间的距离, 并根据距离大小计算排名列表, 差异度距离小的排名越靠前. 表示为:  $R = [r_1, r_2, \dots, r_N]$ ,  $r_1$  表示画廊集第一个行人图像的排名,  $r_2$  表示画廊集第二个行人图像的排名, 其余的则以此类推. 重新排序机制具体计算过程包括两个步骤: 第一步, 计算所有画廊集行人图像相互之间的距离. 第二步: 计算每个画廊集样本图像相对于测试图像的置信度得分. 对每一个画廊集行人  $g_i$ ,  $i = 1, \dots, C$ , 置信度得分  $sg_i$  取决于  $g_i$  与

其余画廊集行人(outliers) 图像之间的距离小于  $g_i$  和测试行人图像之间距离的总数. 如果  $g_i$  是完全正确的匹配, 它对应的  $sg_i$  必须等于零. 也就是说,  $sg_i$  越小意味着  $g_i$  与测试行人图像越相似,  $g_i$  为目标行人图像的概率越大. 计算完所有画廊集行人图像的置信度得分后, 得分被汇总为:  $S_g = [sg_1, sg_2, \dots, sg_c]$ .

最终的排名列表  $Rf$  是第一阶段输出  $R = [r_1, r_2, \dots, r_N]$  和第二阶段输出  $S_g = [sg_1, sg_2, \dots, sg_c]$  的融合, 所以  $Rf$  为:  $Rf = R \circ (1 + S_g)$ .  $\circ$  表示  $R$  和  $S_g$  之间逐元素哈达马积(hadamard product).

## 3 实验分析

### 3.1 数据集

为了验证嵌套池化三元组卷积神经网络的有效性, 本文在 Market-1501、CUHK03 和 VIPeR 这 3 个数据集上进行测试实验. 表 1 给出实验数据集的具体信息. 其中: ID 表示该数据集含有的行人数量; box 表示该数据集包含的行人图像数量; box/ID 表示每个行人平均含有的行人图像数量.

表 1 数据集的具体信息

| 数据集         | ID    | box    | box/ID |
|-------------|-------|--------|--------|
| Market-1501 | 1 501 | 32 217 | 21.5   |
| CUHK03      | 1 467 | 13 164 | 9      |
| VIPeR       | 632   | 1 264  | 2      |

### 3.2 实验结果及分析

为评估所用方法在解决行人再识别问题时的有效性, 本文使用 rank\_5 匹配率和平均准确率(mAP)来衡量所用方法的性能.

$$P = \frac{A}{A+B} \times 100\%$$

$$R = \frac{A}{A+C} \times 100\%$$

$$mAP = \int_0^1 P(R) dR$$

式中,  $P$  为准确率;  $R$  为召回率;  $A$  为系统检测到的相关样本;  $B$  为系统检测到的不相关样本;  $C$  为未检

测到的相关样本.

在 GPU@3.90 GHz,128 GB 显卡内存,64 位 centos7 的计算机系统下进行了相应的实验,并对结果进行分析.实验分为两个部分.

实验一、嵌套池化三元组卷积神经网络包含两个新的成分:(1)利用嵌套不变池理论改进网络结构.(2)利用 Log-logistic 函数代替传统三元组损失函数中的铰链函数,改进损失函数的优化策略.因此,为了评估每个成分对性能改善的影响,实验设置 4 个变量:

(1)(Ours)使用传统的网络结构和三元组损失函数;

(2)(OursT)使用改进后的网络结构和传统的三元组损失函数;

(3)(OursTI)使用改进后的网络结构和新型三元组损失函数;

(4)(OursTIR)使用改进后的网络结构和新型三元组损失函数,并利用置信度测量机制进行间接度量.

在 Market-1501、CUHK03 和 VIPeR 数据集上进行训练和测试,实验结果如表 2、3 和 4 所示,比较不同实验变量下行人再识别评价指标,得出以下结论:

(1)对比 Ours 和 OursN 的实验结果可知:与传统网络结构相比仅利用 NIP(Nested Invariance Pooling)理论改进网络结构就可将 Rank\_5 识别率提高 3%;

(2)对比 OursN 和 OursNI 的实验结果发现,与使用原三元组损失函数的相同深度卷积神经网络相比,仅使用改进的三元组损失函数就可以将行人再识别率提高约 2%;

(3)对比 OursNI 和 OursNIR 的结果,可知置信度测量机制在 VIPeR 数据集上相比于第一次排序准确率提升约 2%,而在 Market1501 和 CUHK03 数据集中准确率提升约 1%.这表明,置信度测量机制在 single-shot 数据集中比 multi-shot 表现更好.这是因为在 single-shot 数据集中,每个测试样本只有一个正确匹配,如果匹配样本与测试样本之间的距离足够小,很有可能利用置信度测量机制纠正误匹配.而在 multi-shot 数据集中,每个测试样本有多个正确匹配,重新排序将处理那些与测试行人之间距离很小,但不属于正确匹配的样本.相比于 single-shot 数据集,识别率提升很小.

表 2 Market1501 数据集实验结果 %

| Method  | Rank_5 | mAP   |
|---------|--------|-------|
| Ours    | 83.38  | 68.63 |
| OursN   | 86.56  | 71.92 |
| OursNI  | 88.92  | 72.32 |
| OursNIR | 91.01  | 83.09 |

表 3 CUHK03 数据集实验结果 %

| Method  | Rank_5 | mAP   |
|---------|--------|-------|
| Ours    | 70.79  | 58.64 |
| OursN   | 74.08  | 60.36 |
| OursNI  | 75.82  | 62.99 |
| OursNIR | 76.73  | 63.65 |

表 4 VIPeR 数据集实验结果 %

| Method  | Rank_5 | mAP   |
|---------|--------|-------|
| Ours    | 64.06  | 55.57 |
| OursN   | 67.78  | 57.39 |
| OursNI  | 69.92  | 58.68 |
| OursNIR | 72.06  | 60.06 |

实验二、在自建数据集中实验结果

为了验证本文方法在非理想自然条件下的行人再识别效率,从 Market-1501、CUHK03 和 VIPeR 数据集中选取具有亮度变化、部分遮挡、分辨率低和旋转、尺度等变化的行人图像样本作为测试集,从而建立专门针对非理想自然条件下的行人再识别数据集,其中测试数据集样本尺寸归一化为  $224 \times 224$  大小.

表 5 中的实验结果表明:Sakrapree 方法<sup>[9]</sup>通过组合几个不同的方法,提高了再识别率. DeepM 方法<sup>[1]</sup>将行人图像分割为头、躯干和腿,然后训练多分支卷积神经网络自适应地提取各部分的特征,取得了不错的效果. LOMO 方法<sup>[6]</sup>提取水平方向发生概率较大的特征,有效改善了视角变化的影响. Deep SP 方法<sup>[10]</sup>在输入层利用空间金字塔结构,分别在整体图像,4 个大小相同条纹以及 8 个大小相同小块上训练 3 个不同的卷积神经网络模型,提取不同分辨率的特征,并根据特征的精细程度加权求和不同分辨率特征的相似度. 该方法的再识别效果优于一般方法,而逊色于本文的行人再识别效果.

表 5 自建数据集中对比实验 %

| Method   | Top1 | Top10 | Top20 |
|----------|------|-------|-------|
| LOMO     | 50.7 | 68.2  | 77.3  |
| DeepM    | 52.3 | 77.7  | 89.1  |
| Deep SP  | 56.5 | 82.9  | 96.4  |
| Sakrapee | 54.8 | —     | —     |
| OursNIR  | 60.3 | 86.2  | 97.8  |

4 结束语

本文利用嵌套不变池理论在均方根池化后,依次添加平均池化和最大池化提取全局特征,提高了特征对平移、尺度、旋转变化的鲁棒性. 其次,利用 AlignedReID 方法计算局部距离,同时进行局部特征和全局特征的联合学习,有助于网络关注有用的图像区域,并区分具有细微差别的类似人物图像. 然后改进损失函数的优化策略,不仅方便训练,同时提高了特征与损失函数的联合优化效果. 但是,置信度测量机制在 multi-shot 数据集中避免误匹配的效果不是很好,有待进一步提升.

参考文献:

[1] SONG Z, CAI X, CHEN Y, et al. Deep convolutional neural networks with adaptive spatial feature for person re-identification [C] // Advanced Information Technology, Electronic and Automation Control Conference. Chongqing, China, IEEE, 2017:2020-2023.

[2] BaK S, CARR P. Deep deformable patch metric learning for person re-identification[J]. IEEE Transactions on Circuits & Systems for Video Technology, 2017 (99):1-3.

[3] ZHANG X, LUO H, FAN X, et al. Alignedreid: Surpassing human-level performance in person re-iden-

tification[J]. arXiv preprint arXiv:1711.08184, 2017, 31(2):1-10.

[4] HERMANS A, BEYER L, LEIBE B. In defense of the triplet loss for person re-identification[J]. arXiv preprint arXiv:1703.07737, 2017, 4(21):1-17.

[5] LOU Y, BAI Y, LIN J, et al. Compact deep invariant descriptors for video Rretrieval[C]//Data Compression Conference [s. l. ]. IEEE, 2017:420-429.

[6] LIAO S, HU Y, ZHU X, et al. Person re-identification by local maximal occurrence representation and metric learning[C] // Computer Vision and Pattern Recognition. Boston, Massachusetts, USA, IEEE, 2015:2197-2206.

[7] 朱建清, 曾焕强, 杜永兆, 等. 基于新型三元卷积神经网络的行人再辨识算法[J]. 电子与信息学报, 2018, 40(4):1012-1016.

[8] MIRMAHBOUB B, MEKHAFI M L, MURINO V, et al. Person re-identification by order-induced metric fusion[J]. Neurocomputing, 2018(275):667-676.

[9] PAISITKRIANGKRAI S, SHEN C, HENGEL A V D. Learning to rank in person re-identification with metric ensembles[C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Boston, MA, USA; IEEE, 2015:1846-1855.

[10] CARR S B P. Deep spatial pyramid for person re-identification[C]// IEEE International Conference on Advanced Video and Signal Based Surveillance. Lecce, Italy, IEEE, 2017:1-6.

作者简介:

王 林 男, (1962-), 博士, 教授. 研究方向为无线传感器网络、复杂网络社团发现、大数据、数据挖掘.

杨 乐(通讯作者) 女, (1994-), 硕士研究生. 研究方向为计算机视觉、深度学习. E-mail: m13572837504\_2@163. com.