

# 基于多信息流动卷积神经网络的行人再识别

桑海峰<sup>1</sup>, 王传正<sup>1</sup>, 吕应宇<sup>1</sup>, 何大阔<sup>2</sup>, 刘 晴<sup>2</sup>

(1. 沈阳工业大学信息科学与工程学院, 辽宁沈阳 110870; 2. 东北大学信息科学与工程学院, 辽宁沈阳 110819)

**摘要:** 行人再识别问题中, 由于视角、光照和行人姿态等因素的变化, 导致难以提取有效的行人特征, 降低识别精度. 而神经网络在训练样本较少的情况下较难训练, 易出现过拟合现象. 针对上述问题, 本文提出一种多信息流动卷积神经网络(Multi-information Flow Convolutional Neural Network, MiF-CNN)模型, 模型中包含一个特殊的卷积结构, 该结构中每层卷积层提取到的特征与后续所有卷积层的输入相连接, 增强了网络的特征信息流动性和梯度的反向传播效率, 使得模型提取到的行人特征更具判别力. 采用多损失函数组合方式训练网络模型, 更好的区分行人类别. 最后利用欧氏距离对行人特征相似性进行排序. 在标准行人再识别数据集 VIPeR 和 CUHK01 上的实验表明, 本文方法进一步提高了行人再识别精度, 并有效改善了深度神经网络的过拟合现象.

**关键词:** 行人再识别; 多信息流动; 特征提取; 卷积神经网络

**中图分类号:** TP391 **文献标识码:** A **文章编号:** 0372-2112(2019)02-0351-07

**电子学报 URL:** <http://www.ejournal.org.cn> **DOI:** 10.3969/j.issn.0372-2112.2019.02.014

## Person Re-identification Based on Multi-information Flow Convolutional Neural Network

SANG Hai-feng<sup>1</sup>, WANG Chuan-zheng<sup>1</sup>, LÜ Ying-yu<sup>1</sup>, HE Da-kuo<sup>2</sup>, LIU Qing<sup>2</sup>

(1. School of Information Science & Engineering, Shenyang University of Technology, Shenyang, Liaoning 110870, China;

2. College of Information Science & Engineering, Northeastern University, Shenyang, Liaoning 110819, China)

**Abstract:** It is challenging to learn efficient features in person re-identification task due to complex variations of viewpoints, illumination, pose etc. In addition, deep neural network still suffers from overfitting with a small training set. To solve these problems, a Multi-information Flow Convolutional Neural Network (MiF-CNN) is designed for person re-identification which contains a specific convolutional architecture. In this architecture, features from each convolution layer are concatenated with all subsequent convolution layers so that it can improve the flowability of feature information and the efficiency of backpropagation gradient. In this way, the proposed network can learn more discriminative features. Moreover, a combination of three loss functions is used to train proposed network. Finally, the identification accuracy is obtained by ranking the similarity of extracted features measured by Euclidean distance. Experimental results on the VIPeR and CUHK01 datasets demonstrate that MiF-CNN outperform most of existing methods of person re-identification and reduce overfitting more effectively.

**Key words:** person re-identification; multi-information flow; feature extraction; convolutional neural network

## 1 引言

行人再识别是指在不同背景的非重叠摄像头视角下进行行人身份识别, 即确认在不同场景和不同时间出现的行人是否为同一个人<sup>[1]</sup>, 被认为是图像检索的一个子问题. 行人再识别在视频监控和公共安全等领域具有重要意义, 近年来引起了研究人员的广泛关注.

早期的行人再识别研究主要侧重特征提取和相似性度量两个方面. 常用的图像特征包括 HSV 颜色直方图、LBP 纹理特征和 Gabor 特征等, 基于特征提取的行人再识别方法通过提取这些人工设计的特征来表征不同的行人. 文献[2]使用重叠的滑动窗口提取行人图像中的 HSV 颜色特征和两个尺度不同的纹理特征, 然后对所有滑动窗口中的相应特征计算最大值, 提出 LOMO

特征. 文献[3]提出 CoPE 特征选择方法来提取具有判别力和有效性的行人特征子集,用于行人再识别问题. 文献[4]提出基于颜色空间的 SCNCD 颜色特征描述子,所提取到行人特征对于光照变化具有很好的鲁棒性.

相似性度量方法同样在行人再识别问题中得到了广泛的应用. 文献[5]运用基于马氏距离的度量学习方法解决行人再识别问题. 文献[6]提出了局部自适应决策函数(LADF),将距离度量和局部自适应阈值约束结合起来进行建模. 文献[7]提出排序支持向量机(RankSVM)方法来学习排序最高的正确匹配的行人特征子空间.

近年来,深度学习在图像识别领域取得了优异表现. 与传统方法相比,深度学习有更强的特征学习能力,且学习到的特征对原数据样本具有更本质的代表性,因此在行人再识别问题上有更好的性能. 文献[8]提出 FPNN 深度学习用于提取行人特征. 文献[9]提出一个新的多通道卷积神经网络,对行人的全身特征与局部身体特征联合学习,并使用三元组损失函数对网络进行训练.

以上基于深度学习的方法在网络结构上大致相似,输入图像经过卷积层学习到的特征图直接作为下一层卷积层的输入,这种网络结构往往忽略了每层产生的特征图对后续每一层的影响,在一定程度上降低了特征信息的流动性. 在梯度反向传播的过程中,随着神经网络层数的加深,梯度更新信息可能会以指数形式衰减,发生梯度消失现象. 此外,训练一个层数较深的神经网络,往往需要大量的数据样本,一旦数据样

本量过少,网络难以收敛且极易产生过拟合的现象,影响最终的识别精度.

针对以上问题,本文提出一种多信息流动卷积神经网络(MiF-CNN),用于解决行人再识别问题. 网络模型包含一系列由多层卷积层采用不同连接方式组成的多信息流动卷积结构,该结构将每层卷积层的输出与其后续所有卷积层的输入相连接,实现了特征的再次利用,增强了整个网络的特征信息流动和梯度反向传播. 采用多损失函数组合的方式训练网络模型,其中包括分类损失函数、中心损失函数和辅助损失函数. 最后使用欧氏距离对网络模型提取到的行人特征进行相似性度量并得到识别结果. 与在用大数据集预训练好的模型上微调小数据集的方法不同,本文直接在小规模数据集 CUHK01 和 VIPeR 上分别进行训练. 实验结果表明,多信息流动卷积神经网络模型能够取得较高的识别精度,且有效地改善了网络过拟合现象.

## 2 多信息流动卷积神经网络的设计

本文提出的 MiF-CNN 模型以多分类任务的思想解决行人再识别问题. 整体的网络模型结构如图 1 所示. 该模型主要包括:2 层堆叠的浅层卷积层、3 个采用新型连接方式的多信息流动卷积结构、全连接层、最大池化层和分类输出层. 不同于用两张或更多张图片作为输入的方法<sup>[9-11]</sup>,该网络以单张行人图像作为输入,经过两层卷积层提取浅层特征,然后经过层数更深的多信息流动卷积结构提取更高级的行人特征,再经过池化层降维之后,传递给全连接层进行信息整合,获得具有判别力的行人特征向量.

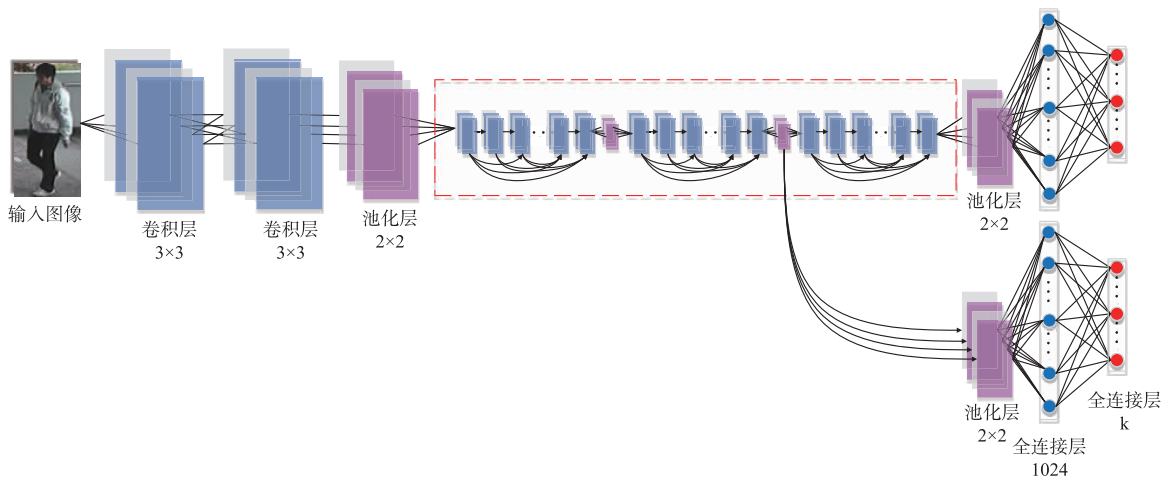


图1 MiF-CNN结构图,其中k为训练集中的行人类别数目.红色虚线框区域为多信息流动卷积结构

### 2.1 卷积神经网络的特征提取

网络模型中,各卷积层的卷积核大小为  $3 \times 3$ ,卷积步长为 1,每个卷积层之后都会接一个批标准化层

(Batch Normalization)和 ReLU 激活函数. 卷积层的运算过程可表示为:

$$\begin{cases} z_j^{(\ell)} = \sum_i x_i^{(\ell-1)} \otimes w_j^{(\ell)} \\ x_j^{(\ell)} = \sigma(z_j^{(\ell)}) \end{cases}, \ell > 1 \quad (1)$$

其中,  $x_i^{(\ell-1)}$  表示第  $\ell-1$  层卷积层输出的第  $i$  个特征图;  $z_j^{(\ell)}$  表示卷积核对  $x_i^{(\ell-1)}$  的卷积运算结果;  $x_j^{(\ell)}$  表示第  $\ell$  层卷积层输出的第  $j$  个特征图;  $w_j^{(\ell)}$  表示第  $\ell$  层的第  $j$  个特征图所对应的卷积核;  $\otimes$  符号表示卷积操作. 卷积提取特征的过程可理解为, 第  $\ell$  层的第  $j$  个特征图上的神经元通过卷积核  $w_j^{(\ell)}$  对第  $\ell-1$  层输出的每张特征图连接、卷积后进行求和, 并将提取到的特征映射

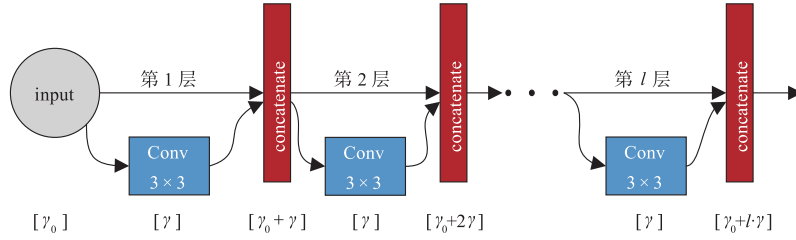


图2 多信息流动卷积结构示意图 ([ ]表示各层特征图的通道数目)

由图2可以看出, 每层卷积层的输入都来自其前面所有层的输出, 这样的连接方式使得在前向传递过程中, 每层输出的特征图都可以被后面的所有层重新利用, 增强了整个网络的信息流动. 而在反向传播的过程中, 每层输入的梯度包含了损失函数对输入的直接求导, 使得梯度的传播更加有效, 网络更易训练. 多信息流动卷积结构中, 每层卷积层产生的特征图数目为一个固定值  $\gamma$ , 则第  $\ell$  层卷积层的输入为  $\gamma_0 + \gamma(\ell-1)$  个特征图, 其中  $\gamma_0$  为多信息流动卷积结构输入层的特征图通道数目. 设输入层的第  $i$  个通道上的特征图为  $x_i^{(0)}$ , 其中  $i \in (1, \gamma_0)$ , 则多信息流动卷积结构的第一层卷积层输出的第  $j$  个通道的特征图为:

$$z_j^{(1)} = \sum_i x_i^{(0)} \otimes w_j^{(1)} \quad (2)$$

其中  $w_j^{(1)}$  为第一层卷积层的权重, 经过激活函数之后的输出为:

$$a_j^{(1)} = \sigma(z_j^{(1)}) \quad (3)$$

其中  $j \in (1, \gamma)$ . 所以, 多信息流动卷积结构中的第  $\ell$  层卷积层产生的特征图可表示为:

$$\begin{cases} z_m^{(\ell)} = \sum_k x_k^{(\ell-1)} \otimes w_m^{(\ell)} \\ a_m^{(\ell)} = \sigma(z_m^{(\ell)}) \end{cases}, \ell > 0 \quad (4)$$

其中,  $k \in (1, \gamma_0 + \gamma(\ell-1))$ ,  $m \in (1, \gamma)$ . 经连接操作之后, 第  $\ell$  层卷积层的输出为:

$$x_n^{(\ell)} = [x_k^{(\ell-1)}, a_m^{(\ell)}], \ell > 0 \quad (5)$$

其中,  $n \in (1, \gamma_0 + \gamma \cdot \ell)$ ,  $[\cdot, \cdot]$  表示在通道数维度上的连接操作.

在梯度的反向传播过程中, 设  $\Delta x_n^{(\ell)}$  为损失函数对于输出  $x_n^{(\ell)}$  的误差, 因为  $x_n^{(\ell)}$  由  $x_k^{(\ell-1)}$  和  $a_m^{(\ell)}$  两部分连

到第  $\ell$  层的第  $j$  个特征图上.  $\sigma(\cdot)$  函数为 ReLU 激活函数, 可表示为  $\sigma(x) = \max(x, 0)$ ; 因为加入了批标准化操作, 所以这里忽略了偏置对输出的影响.

## 2.2 多信息流动卷积结构

在该网络结构中, 每一层卷积层的输出与该卷积层的输入相连接后馈送进下一层卷积层, 即每一层卷积层的输入都是其前面所有层输出的连接组合, 具体形式如图2所示.

接而成, 所以在  $\Delta x_n^{(\ell)}$  反向传播的过程中会产生两部分梯度, 分别为:

$$\begin{cases} \Delta a_m^{(\ell)} = \Delta x_n^{(\ell)} \cdot \frac{\partial x_n^{(\ell)}}{\partial a_m^{(\ell)}} \\ \Delta x_k^{(\ell-1)} = \Delta x_n^{(\ell)} \cdot \frac{\partial x_n^{(\ell)}}{\partial x_k^{(\ell-1)}} \end{cases} \quad (6)$$

其中  $\Delta a_m^{(\ell)}$  为第  $\ell$  层卷积层经激活函数后的输出的反向传播误差,  $\Delta x_k^{(\ell-1)}$  为第  $\ell-1$  层卷积层的输出的反向传播误差. 第  $\ell$  层卷积层权重的反向传播误差可表示为:

$$\begin{cases} \Delta z_m^{(\ell)} = \Delta a_m^{(\ell)} \cdot \frac{\partial a_m^{(\ell)}}{\partial z_m^{(\ell)}} = \Delta x_n^{(\ell)} \cdot \frac{\partial x_n^{(\ell)}}{\partial a_m^{(\ell)}} \cdot \sigma'(z_m^{(\ell)}) \\ \Delta w_m^{(\ell)} = \Delta z_m^{(\ell)} \cdot \frac{\partial z_m^{(\ell)}}{\partial w_m^{(\ell)}} = \Delta z_m^{(\ell)} \cdot x_k^{(\ell-1)} \end{cases} \quad (7)$$

其中,  $\Delta z_m^{(\ell)}$  为第  $\ell$  层卷积层的卷积结果的反向传播误差,  $\sigma'(z_m^{(\ell)})$  为激活函数对  $z_m^{(\ell)}$  的导数,  $\Delta w_m^{(\ell)}$  为第  $\ell$  层卷积层权重的反向传播误差, 神经网络即使用该误差对每层的权重进行更新, 具体形式为:

$$(w_m^{(\ell)})_{update} = w_m^{(\ell)} - \eta \cdot \Delta w_m^{(\ell)} \quad (8)$$

其中,  $\eta$  为网络模型训练的学习率. 误差继续反向传播到第  $\ell-1$  层的输出层, 对于输出  $x_k^{(\ell-1)}$  的误差为:

$$\Delta x_k^{(\ell-1)} = \Delta z_m^{(\ell)} \cdot \frac{\partial z_m^{(\ell)}}{\partial x_k^{(\ell-1)}} = \Delta z_m^{(\ell)} \cdot w_m^{(\ell)} \quad (9)$$

由式(9)和式(6)可以看出, 本文的网络模型中, 损失函数对于多信息流动卷积结构中的每层卷积层的输出产生了两个梯度, 使得误差更加有效地在网络中反向传播, 且一定程度上抑制了“梯度消失”, 增强了误差信息在网络中的流动性.

本文网络结构中包含 3 个多信息流动卷积结构, 在每 2 个多信息流动卷积结构之间, 添加一个中间池化层用于压缩每个多信息流动卷积结构输出的冗余特征, 提高最终的识别精度.

实际训练时, 超参数  $\gamma$  取值较小, 且每次进入到一个新的多信息流动卷积结构,  $\gamma$  的值增加一倍. 这样的设计使得每层卷积层仅学习到少量特征, 减少了特征向量中的冗余特征, 降低了网络的计算量. 而结合较深的网络层数, 使得网络模型能够学习到更高级更复杂的行人特征, 提高最终的识别精度.

### 2.3 损失函数

现有的深度学习算法常采用分类损失函数作为目标函数, 其公式如式 (10) 所示.

$$L_s = -\frac{1}{m} \sum_{i=1}^m y^{(i)} \log \frac{e^{\theta_{y^{(i)}}^T x^{(i)}}}{\sum_{k=1}^K e^{\theta_k^T x^{(i)}}} \quad (10)$$

其中,  $y^{(i)}$  表示真实的行人类别标签,  $\theta$  是全连接层的参数,  $x^{(i)}$  为训练样本的特征向量,  $k$  为行人类别数目,  $m$  为批大小 (batch size).

而在实际应用的过程中发现, 当仅使用分类损失函数训练网络模型时, 如果网络模型提取的特征欠佳, 会导致类内距离大于类间距离的情况发生, 从而影响最终的分类和识别效果. 针对此问题, Wen<sup>[12]</sup> 等人在 2016 年提出了中心损失函数 (center loss), 结合分类损失函数使用, 使模型学习到的特征具有更强的判别力和泛化能力. 中心损失函数公式如下:

$$L_c = \frac{1}{2m} \sum_{i=1}^m \|x_i - c_j\|_2^2 \quad (11)$$

其中,  $c_j$  表示属于第  $j$  类行人特征的中心,  $x_i$  表示行人样本的特征向量. 由式 (11) 可以看出, 中心损失函数使得每一类样本的特征与该类中心的距离最小, 从而达到减小类内距离的目的. 每个类别的中心  $c_j$  根据式 (12) 进行更新:

$$\begin{cases} \Delta c_j^t = \frac{\sum_{i=1}^m \delta(y_i = j) \cdot (c_j^t - x_i)}{1 + \sum_{i=1}^m \delta(y_i = j)} \\ c_j^{t+1} = c_j^t - \alpha \cdot \Delta c_j^t \end{cases} \quad (12)$$

其中,  $\alpha$  为中心更新率,  $\delta(y_i = j)$  表示如果模型预测类别与真实标签相同, 则该项等于 1, 否则该项等于 0. 因此, 只有当预测类别与真实标签一致时, 该类别的中心才会更新. 在训练过程中, 该损失函数乘以一个权重  $\beta$  后, 加到分类损失函数中.

结合文献 [13] 思想, 本文将第二个多信息流动卷积结构的输出特征拿出用作辅助分类特征, 该部分使用与式 (10) 类似的分类损失函数, 其公式为:

$$L_a = -\frac{1}{m} \sum_{i=1}^m y^{(i)} \log \frac{e^{\theta_{y^{(i)}}^T x^{(i)}}}{\sum_{k=1}^K e^{\theta_k^T x^{(i)}}} \quad (13)$$

其中,  $\theta_a$  为辅助分类部分的全连接层参数.

综上所述, 本文所采用的损失函数由分类损失函数、中心损失函数和辅助分类损失函数三部分组成. 总的损失函数如式 (14) 所示.

$$Loss = L_s + \beta \cdot L_c + L_a \cdot 0.3 \quad (14)$$

其中, 辅助分类损失函数乘以一个较小的权重 0.3 后, 加到总的损失函数中.

## 3 实验结果与分析

### 3.1 实验设置与评估

在 VIPeR<sup>[14]</sup> 和 CUHK01<sup>[15]</sup> 两个数据集上对本文方法进行训练和测试. VIPeR 数据集是使用最为广泛且难度较高的行人再识别数据集之一, 包含 632 个行人的图像样本, 图像由两个视角不同的相机采集, 每个行人在每个不同视角下有一张图像. CUHK01 数据集包含了 971 个行人的 3884 张图像, 同样由两个不同视角的相机采集, 每个行人对应每个相机有两张图像. 两个数据集中的行人图像样本如图 3 所示.



图3 行人再识别数据集图像样本

实际训练在配有 8G 显存的 GTX 1080 图形处理器、Intel Core i7-7700K 中央处理器和 8G 内存的台式电脑上进行. CUHK01 的训练集包含 9700 张行人图像, 一次完整的训练需要约 2 小时 20 分钟, VIPeR 的训练集包含 3160 张行人图像, 一次完整的训练需要约 50 分钟.

测试分别在两个数据集上进行 10 次, 取 10 次测试的平均 rank-k 精度记录在累计匹配曲线 (CMC) 上作为评估指标.

### 3.2 VIPeR 数据集上的实验结果

将 VIPeR 数据集中的 632 类行人的图像样本平均分成两份, 316 类行人的图像作为训练样本, 另外 316 类行人的图像作为测试样本. MiF-CNN 的识别精度与其他经典行人再识别算法的识别精度对比如表 1

所示.

表 1 VIPeR 数据集上各算法的识别精度(%)

方法	Rank1	Rank5	Rank10
MSCAN	38.08	64.14	73.52
SIR-CIR	35.76	67.40	<b>83.50</b>
DeepFeature	40.50	60.80	70.40
Rank-CNN	38.37	<b>69.22</b>	81.33
JSTL + DGD	37.70	-	-
CSBT	36.66	66.20	78.64
S-CNN	37.80	66.90	77.40
MiF-CNN	<b>40.82</b>	68.70	81.60

由表 1 可以看出,本文方法在难度较大的 VIPeR 数据集上的性能较以往的经典算法有一定提升,rank1 精度比所有经典算法中最好的结果提高了约 3%. 其中, MSCAN<sup>[16]</sup>、SIR-CIR<sup>[17]</sup>、JSTL + DGD<sup>[18]</sup>、S-CNN<sup>[19]</sup>、DeepFeature<sup>[20]</sup>、Rank-CNN<sup>[21]</sup> 均为最新的基于深度学习的行人再识别算法. 各算法在 VIPeR 数据集上的 CMC 曲线如图 4(a) 所示.

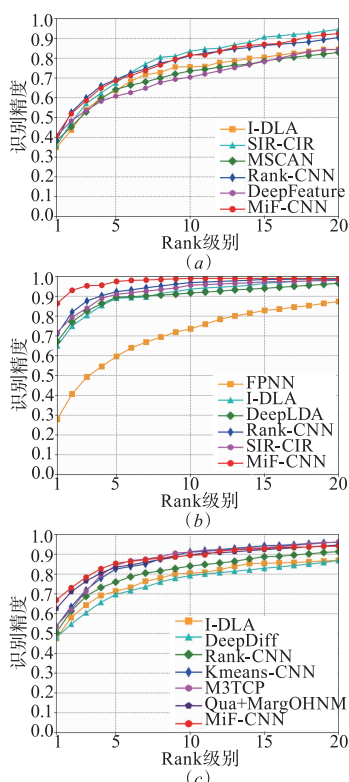


图 4 不同算法在两个数据集上的 CMC 曲线

### 3.3 CUHK01 数据集上的实验结果

本文在 CUHK01 数据集上按照两种不同的测试集规模进行测试,并与其它经典算法进行了横向比较.

(a) 训练集行人类别数 871, 测试集行人类别数 100: 在这种实验条件下,训练样本较多,使得深度卷积

神经网络更易训练,识别精度更高. 各算法在该实验条件下的识别精度如表 2 所示. CMC 曲线如图 4(b) 所示.

表 2 CUHK01 ( $k=100$ ) 数据集上各算法的识别精度(%)

方法	Rank1	Rank5	Rank10
FPNN	27.87	59.64	73.53
Rank-CNN	70.94	92.30	96.90
I-DLA	65.00	89.00	93.60
SIR-CIR	71.80	91.00	95.60
DeepLDA	67.12	89.45	91.68
kLFDA	42.76	69.01	79.63
MiF-CNN	<b>86.40</b>	<b>97.50</b>	<b>99.00</b>

(b) 训练集行人类别数 485, 测试集行人类别数 486: 该实验条件下的各算法识别精度如表 3 所示. 训练样本明显减少的情况下,训练的深度卷积神经网络更易发生过拟合现象,但是通过对比可以看出, MiF-CNN 在训练样本较少的情况下依然可以获得较高的识别精度,rank1 精度和 rank5 精度在所有算法中均为最高,表明了与 I-DLA<sup>[10]</sup>、JSTL + DGD<sup>[18]</sup> 和 DeepDiff<sup>[22]</sup> 等基于深度学习的算法相比,本文的 MiF-CNN 网络结构在改善过拟合方面有着较大优势. CMC 曲线如图 4(c) 所示.

表 3 CUHK01 ( $k=486$ ) 数据集上各算法的识别精度(%)

方法	Rank1	Rank5	Rank10
JSTL + DGD	63.00	-	-
M3TCP	53.70	84.30	91.00
Rank-CNN	50.41	75.93	84.07
DeepDiff	47.90	69.67	79.20
I-DLA	47.53	71.35	80.50
Qua + MargOHNM	62.55	83.44	89.71
Kmeans-CNN	53.50	82.50	<b>91.20</b>
MiF-CNN	<b>66.87</b>	<b>85.39</b>	89.51

### 3.4 网络模型深度对识别精度的影响

为了验证每个多信息流动卷积结构中卷积层层数的不同对识别精度的影响,图 5 给出了采用不同层数设计的多信息流动卷积结构进行行人再识别的实验对比,所有实验在 VIPeR 数据集上进行.

由图 5 可以看出,随着每个多信息流动卷积结构中的层数逐渐增大,识别精度逐渐增高,当层数为 9 层时识别精度最高,而当层数超过 9 层时,识别精确度有所下降,并逐渐趋于稳定,表明如果设计的深度卷积神经网络层数过深,会增加网络的训练难度,加剧了网络过拟合的程度,影响最终识别精度.

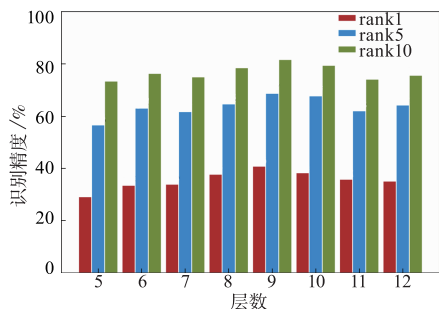


图5 不同层数的多信息流动卷积结构对识别精度的影响

### 3.5 实验结果分析

在众多对比算法中, MSCAN、SIR-CIR、JSTL + DGD、S-CNN、DeepFeature、Rank-CNN 和 I-DLA 等方法与本文相似,均为基于卷积神经网络的深度学习方法,但是本文方法在两个数据集上的 rank1 精度和 rank5 精度均超过了上述方法,表明 MiF-CNN 能够有效地提高行人再识别的识别精度。

值得注意的是,由于数据集 CUHK01 和 VIPeR 属于小规模数据集,训练样本较少,给网络模型的训练增加了难度,容易产生过拟合现象。为降低过拟合,基于深度学习的 JSTL + DGD 算法在多个数据集混合组成的大数据集上进行联合学习,并在训练后的网络模型上分别对数据集 CUHK01 和 VIPeR 进行微调;而 M3TCP<sup>[11]</sup>算法在实际训练过程中,网络模型的结构和超参数需要手动调节,以适应不同数据集的规模的变化,对于样本较少的小数据集,则需要减少网络模型的层数来降低网络过拟合的现象。

相比而言,本文设计的 MiF-CNN 结构固定,且分别使用两个数据集的样本直接训练,即使网络层数较深,网络模型也能够快速收敛,在 CUHK01 数据集上的 rank1 精度比 JSTL + DGD 算法的 rank1 精度高出 0.2%,比 M3TCP 算法的 rank1 精度高出 13.1%,在 VIPeR 数据集上的 rank1 精度比 JSTL + DGD 算法的 rank1 精度高出 0.27%,再次表明 MiF-CNN 具有良好的泛化能力,使得网络更加容易训练,有效地改善了因训练样本较少而产生的过拟合情况,在行人再识别问题上有着优异性能。

## 4 结论

针对行人再识别问题,本文设计一个多信息流动卷积神经网络模型,模型中包含的多信息流动卷积结构有效地实现了特征的重利用,在增强网络特征流动性的同时,提高了梯度反向传播的效率,使得网络更易训练,提取到更具判别力的行人特征。同时采取多损失函数组合的方式训练网络,更大程度上的增大类间距离,减小类内距离。大量的对比实验表明,本文方法的

识别精度较以往的经典算法有较明显的提升,尤其在数据集训练样本较少的情况下能够有效改善过拟合现象,提高网络的泛化能力。未来计划针对特征选择方法进行研究,减少特征冗余,进一步提高算法效率和识别精度。

### 参考文献

- [1] 杨钰源. 基于度量学习和深度学习的行人重识别研究[D]. 重庆:重庆大学,2016.
- [2] Liao S, Hu Y, Zhu X, et al. Person re-identification by local maximal occurrence representation and metric learning [A]. Computer Vision and Pattern Recognition [C]. IEEE, 2015. 2197 - 2206.
- [3] Tahir S F, Cavallaro A. Cost-effective features for reidentification in camera networks [J]. IEEE Transactions on Circuits & Systems for Video Technology, 2014, 24(8): 1362 - 1374.
- [4] Yang Y, Yang J, Yan J, et al. Salient color names for person re-identification [A]. European Conference on Computer Vision [C]. Springer, Cham, 2014. 536 - 551.
- [5] Zheng W S, Gong S, Xiang T. Person re-identification by probabilistic relative distance comparison [A]. Computer Vision and Pattern Recognition [C]. IEEE, 2011. 649 - 656.
- [6] Li Z, Chang S, Liang F, et al. Learning locally-adaptive decision functions for person verification [A]. Computer Vision and Pattern Recognition [C]. IEEE, 2013. 3610 - 3617.
- [7] Engel C, Baumgartner P, Holzmann M, et al. Person re-identification by support vector ranking [A]. British Machine Vision Conference, BMVC 2010 [C]. Aberystwyth, UK, August 31-September 3, 2010. Proceedings. DBLP, 2010. 1 - 11.
- [8] Li W, Zhao R, Xiao T, et al. DeepReID: Deep filter pairing neural network for person re-identification [A]. Computer Vision and Pattern Recognition [C]. IEEE, 2014. 152 - 159.
- [9] Cheng D, Gong Y, Zhou S, et al. Person re-identification by multi-channel parts-based CNN with improved triplet loss function [A]. Computer Vision and Pattern Recognition [C]. IEEE, 2016. 1335 - 1344.
- [10] Ahmed E, Jones M, Marks T K. An improved deep learning architecture for person re-identification [A]. Computer Vision and Pattern Recognition [C]. IEEE, 2015. 3908 - 3916.
- [11] Chen W, Chen X, Zhang J, et al. Beyond triplet loss: a deep quadruplet network for person re-identification [A]. The Conference on Computer Vision and Pattern Recognition [C]. 2017. 1320 - 1329.
- [12] Wen Y, Zhang K, Li Z, et al. A Discriminative Feature Learning Approach for Deep Face Recognition [M]. Computer Vision-ECCV 2016. Springer International Pub-

- lishing, 2016. 499 – 515.
- [13] Szegedy C, Liu W, Jia Y, et al. Going deeper with convolutions [A]. IEEE Conference on Computer Vision and Pattern Recognition [C]. IEEE Computer Society, 2015. 1 – 9.
- [14] Gray D, Tao H. Viewpoint invariant pedestrian recognition with an ensemble of localized features [A]. Computer Vision-ECCV 2008, European Conference on Computer Vision [C]. Marseille, France, October 12-18, Proceedings. DBLP, 2008. 262 – 275.
- [15] Li W, Zhao R, Wang X. Human reidentification with transferred metric learning [A]. Asian Conference on Computer Vision [C]. Springer, Berlin, Heidelberg, 2012. 31 – 44.
- [16] Li D, Chen X, Zhang Z, et al. Learning deep context-aware features over body and latent parts for person re-identification [A]. Computer Vision and Pattern Recognition [C]. IEEE, 2017. 384 – 393.
- [17] Wang F, Zuo W, Lin L, et al. Joint learning of single-image, age and cross-image representations for person re-identification [A]. Computer Vision and Pattern Recognition [C]. IEEE, 2016. 1288 – 1296.
- [18] Xiao T, Li H, Ouyang W, et al. Learning deep feature representations with domain guided dropout for person re-identification [A]. Computer Vision and Pattern Recognition [C]. IEEE, 2016. 1249 – 1258.
- [19] Varior R R, Haloi M, Wang G. Gated siamese convolutional neural network architecture for human re-identification [A]. European Conference on Computer Vision [C]. Springer International Publishing, 2016. 79 – 808.
- [20] Ding S, Lin L, Wang G, et al. Deep feature learning with relative distance comparison for person re-identification [J]. Pattern Recognition, 2015, 48 (10): 2993 – 3003.
- [21] Chen S Z, Guo C C, Lai J H. Deep ranking for person re-identification via joint representation learning [J]. IEEE Transactions on Image Processing, 2015, 25 (5): 2353 – 2367.
- [22] Huang Y, Sheng H, Zheng Y, et al. DeepDiff: Learning deep difference features on human body parts for person re-identification [J]. Neurocomputing, 2017, 241: 191 – 203.

#### 作者简介



**桑海峰** 男, 1978 年生于辽宁沈阳, 博士, 沈阳工业大学教授, 主要研究方向为视觉检测技术与图像处理。  
E-mail: sanghaif@163.com



**王传正 (通信作者)** 男, 1994 年生于辽宁大连, 沈阳工业大学硕士研究生, 主要研究方向为计算机视觉与图像处理。  
E-mail: qing\_0304@outlook.com