

平移变化性相似性学习的行人重识别算法

陈兵 查宇飞* 李运强 张胜杰 张园强

(空军工程大学航空航天工程学院 西安 710038)

摘要: 行人重识别的精确度主要取决于相似性度量方法和特征学习模型。现有的度量方法存在平移不变性的特点,会增加网络参数训练的难度。现有的几种特征学习模型只强调样本之间的绝对距离而忽略了正样本对和负样本对之间的相对距离,造成网络学习到的特征判别性不强。针对现有度量方法的缺点该文提出一种平移变化的距离度量方法,能够简化网络的优化并能高效度量图像之间的相似性。针对特征学习模型的不足,提出一种增大间隔的逻辑回归模型,模型通过增大正负样本对之间的相对距离,使得网络得到的特征判别性更强。实验中,在Market1501和CUHK03数据库上对所提度量方式和特征学习模型的有效性进行验证,实验结果表明,所提度量方式性能更好,其平均精确率超出马氏距离度量6.59%,且所提特征学习模型也取得了很好的性能,算法的平均精确率较现有的先进算法有显著提高。

关键词: 行人重识别; 平移变化; 相似性学习; 逻辑回归

中图分类号: TP391.41

文献标识码: A

文章编号: 1009-5896(2018)10-2381-07

DOI: 10.11999/JEIT180184

Shift-variant Similarity Learning for Person Re-identification

CHEN Bing ZHA Yufei LI Yunqiang ZHANG Shengjie ZHANG Yuanqiang

(Aeronautics Engineering College, Air Force Engineering University, Xi'an 710038, China)

Abstract: The accuracy of pedestrian re-recognition mainly depends on the similarity measure and the feature learning model. The existing measurement methods have the characteristics of translation invariance, which make the training of network parameters difficult. Several existing feature learning models only emphasize the absolute distance between sample pairs, but ignore the relative distance between positive sample pairs and negative sample pairs, resulting in a weak discriminant feature in network learning. In view of the shortcomings of existing measurement methods, a distance measurement method of translation change is presented, which can effectively measure the similarity between images. To overcome the shortcomings of the feature learning model, based on the proposed translation distance metric, a new logistic regression model with enlarged intervals is proposed. By increasing the relative distance between the positive and negative sample pairs, the network can get more discriminant features. In the experiment, the validity of the proposed measurement and the feature learning model is verified on the Market1501, CUHK03 database. Experimental results show that the proposed metric performs better than the Mahalanobis distance metric 6.59%, and the proposed feature learning algorithm also achieves good performance. The average precision of the algorithm is improved significantly compared with the existing advanced algorithms.

Key words: Person Re-identification; Shift-variant; Similarity learning; Logistic regression

1 引言

行人重识别任务旨在从不同摄像头拍摄的海量

图片数据库中找出与目标行人匹配的行人图片,其在人员跟踪、监视和安防方面都有着极其重要的应用。在这样的应用中,想要让训练集包含所有行人信息是不现实的,这就要求算法要能让测试集中两个相似的行人在相似性排序中靠近,不相似的行人排序远。

当前,行人重识别的研究方法主要有两种:基于特征描述的方法和基于距离度量的方法。基于特征描述的方法通常要找到一种鲁棒性良好的特征以

收稿日期: 2018-02-09; 改回日期: 2018-07-17; 网络出版: 2018-07-23

*通信作者: 查宇飞 735754591@qq.com

基金项目: 国家自然科学基金(61472442, 61773397, 61701524), 陕西省科技新星资助(2015kjxx-46)

Foundation Items: The National Natural Science Foundation of China (61472442, 61773397, 61701524), Shanxi Science and Technology New Star Fund (2015kjxx-46)

解决光照和视角变化带来的影响。基于距离度量的方法着重从训练数据中学习出一种鲁棒性的度量方法来解决复杂的匹配问题。在行人重识别领域中许多特征描述方法都得到了应用,包括尺度不变特征变换特征(SIFT)^[1]、局部二值模式特征(LBP)^[2]和颜色直方图特征^[3]等。这些手工设计特征对于光照变化、不同的视场和姿势的泛化能力还不够鲁棒。最近,卷积神经网络(Convolutional Neural Networks, CNN)提取的深度特征已经被证明有着良好的鲁棒性。基于卷积神经网络的深度特征学习^[4,5]和深度距离度量^[6,7]在行人重识别任务上都取得了很好的效果。文献^[7]提出使用孪生网络来学习输入图像对的特征描述。文献^[8]设计了一个特殊的网络,可以同时处理遮挡、光照、几何等变化。文献^[9]在传统的网络上设计了两个特殊的层,学习输入样本对之间的关系。文献^[10]设计了一种特殊的子网络插入到孪生网络之后来学习单个图像的特征描述并同时学习图像之间的图像之间的特征描述。

行人重识别领域中度量学习就是学习一种高效的距离度量方式。验证模型中的度量函数通常是在某一个度量约束条件下学习,例如在度量矩阵约束下的 L_1 归一化距离^[11]、 L_2 归一化欧氏距离^[4,11,12]和马氏距离(Mahalanobis distance)^[13],以及范围约束在 $[-1, 1]$ 之间的余弦距离^[8]和 $[0, 2\pi]$ 之间的角度距离^[14,15]。文献^[16]提出应用局部判别分析(LFDA)进行距离度量前的特征降维。文献^[17]对其使用核技巧来进一步处理非线性的数据。文献^[18,19]在对比损失函数(contrastive loss)中为负样本对设置一个间隔,用来惩罚距离小于这个间隔的负样本对。三元组损失函数(triplet loss)^[5,20]通过明确的为正样本对和负样本对距离之间设置一个间隔来约束正负样本之间相对距离关系。

本文从距离度量和特征学习两个方面着手,关注于去学习高效的相似性度量函数和鲁棒的特征学习模型。首先,本文提出了一种平移变化的度量方式,简化了网络的训练。本文提出的距离度量方式与之前的方法有很大的不同,既没有度量矩阵的约束,也没有将距离函数严格的约束在某一个范围之内,相似性的值域遍布整个实数域,比传统度量方式的优化空间更广泛。然后基于这种度量方式,本文将样本之间的距离与相似性概率联系起来提出了一种增大间隔的逻辑回归损失函数,并在正负样本的相似性学习中引入决策间隔来增强特征的判别性,使得正负样本分得更开。在Market1501, CUHK03这些主流的数据库上,本文算法能取得较好的性能。

2 本文算法

本文提出了一种平移变化的度量方式,简化了网络的优化,并基于该度量方式提出了一种增大间隔的逻辑回归损失函数。其将样本之间的距离与样本的相似性概率联系起来,约束距离小的样本相似的概率越大,距离大的样本相似的概率越小。然后将增大间隔的逻辑回归损失函数嵌入到卷积神经网络中进行端到端的优化,其网络结构如图1所示。本文采用双路的孪生网络,以ResNet-50作为孪生网络的主干网络,在训练的过程中两路网络共享参数。

2.1 平移变化的相似性度量

度量查询图片和候选图片之间的相似性是行人重识别任务的关键。现有的距离度量方法具有平移不变性的特点,典型的有欧氏距离和马氏距离。记两个特征向量 $\mathbf{p} \in R^d$ 和 $\mathbf{q} \in R^d$ 之间的欧氏距离 $\|\mathbf{p} - \mathbf{q}\|$,平移不变性是指对这两个向量同时加上

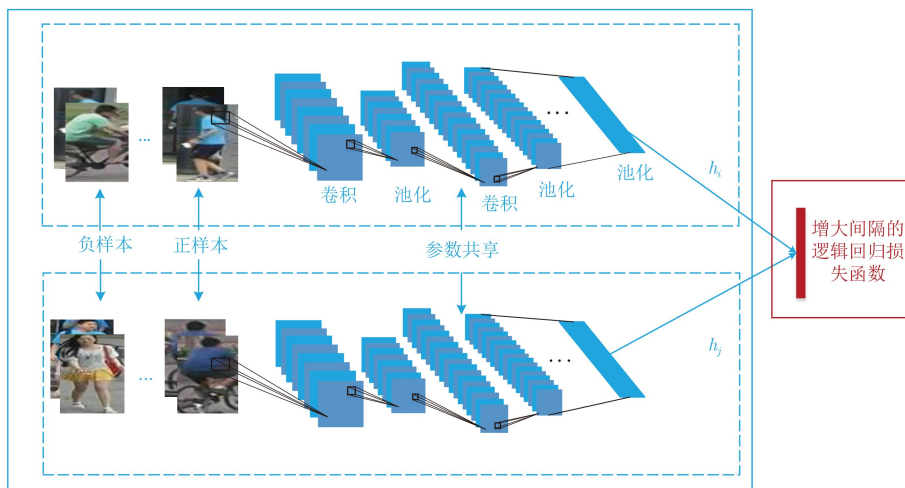


图1 端到端的行人重识别

一个向量 $\mathbf{b} \in R^d$ 的欧氏距离 $\|(\mathbf{p} + \mathbf{b}) - (\mathbf{q} + \mathbf{b})\|$ 和原来的距离是相等的。平移不变性使得距离不依赖于底层特征向量的位置，当深度网络的优化和度量学习结合时，平移不变性使得特征学习变得复杂，平移不变性为底层特征的表示提供了无穷多的解，这使得网络优化变得复杂。

针对传统度量方法的缺点，本文提出了一种平移变化的相似性度量方式，并称之为加权内积距离。当使用孪生网络学习相似性度量时，将输入图像对表示为 I_1 和 I_2 ，提取的深度特征分别记作 $\mathbf{h}_1 \in R^d$ 和 $\mathbf{h}_2 \in R^d$ ，则样本之间的马氏距离为

$$d^2 = (\mathbf{h}_1 - \mathbf{h}_2)^T \mathbf{M} (\mathbf{h}_1 - \mathbf{h}_2) \quad (1)$$

它允许学习距离的权重并通过优化 \mathbf{M} 矩阵来获得好的度量方式，式中 \mathbf{M} 为半正定矩阵， $\mathbf{M} = \mathbf{L}^T \mathbf{L}$ 。为了减少参数本文只保留 \mathbf{M} 的对角矩阵 \mathbf{w} 。式(1)可以写成

$$d^2 = \sum_{i=1}^N (\mathbf{h}_1^i - \mathbf{h}_2^i)^T \mathbf{w}^i (\mathbf{h}_1^i - \mathbf{h}_2^i) \quad (2)$$

式中， \mathbf{w}^i 为权重向量 \mathbf{w} 中的元素 i ，则把式(2)可以展开为：

$$d^2 = \sum_{i=1}^N \mathbf{w}^i \mathbf{h}_1^{i2} + \mathbf{w}^i \mathbf{h}_2^{i2} - 2\mathbf{w}^i \mathbf{h}_1^i \mathbf{h}_2^i \quad (3)$$

为了改进马氏距离的平移不变性，我们去除 \mathbf{h}_1^{i2} 和 \mathbf{h}_2^{i2} 项，没有了这些项，则变成了平移变化的相似性度量方式：

$$\hat{d}^2 = \sum_{i=1}^N -2\mathbf{w}^i \mathbf{h}_1^i \mathbf{h}_2^i \quad (4)$$

由于单调性，式(4)相当于

$$D(\mathbf{h}_1, \mathbf{h}_2) = \mathbf{w}^T (\mathbf{h}_1 \odot \mathbf{h}_2) \quad (5)$$

此处首先进行 \mathbf{h}_1 和 \mathbf{h}_2 之间的元素乘积，然后用权重向量 \mathbf{w} 对其进行线性变换。由式(5)可知，本文这种加权内积的度量方式具有平移变化的特性，可以简化网络的训练，且优化空间更广，为 $(-\infty, +\infty)$ 。更重要的是它允许学习距离的权重，考虑了不同特征维度的重要性。

2.2 增大间隔的逻辑回归损失函数

2.2.1 逻辑回归损失函数

为了得到判别性更强、鲁棒性更好的特征，本文利用逻辑回归模型把样本的相似性概率和样本之间的距离联系起来，重新定义条件概率函数，提出了一种增大间隔的逻辑回归损失函数。并通过上文加权内积的度量方法引导网络减小正样本对之间的距离和扩大负样本对之间的距离。

逻辑回归函数的表达式为： $\sigma(\theta) = \frac{1}{1 + e^{-\theta}}$ ，对任意一个输入样本 θ ，逻辑回归函数可以将它映射到一个 $[0, 1]$ 之间的值，式中输出 σ 可以作为概率值。本文记 $\mathcal{T} = \{x_i, x_j, s_{ij}\}$ 为训练数据集，其中 (x_i, x_j) 为输入样本对， $s_{ij} \in \{0, 1\}$ 为样本对之间的相似性标签，当输入为同类样本(正样本对)时 $s_{ij} = 1$ ，反之(负样本对) $s_{ij} = 0$ ，记样本对的深度特征为 $\{\mathbf{h}_i, \mathbf{h}_j\}_{i=1}^n$ 。然后定义条件概率：

$$p(s_{ij}|\mathbf{F}) = \begin{cases} \delta(D_{ij}), & s_{ij} = 1 \\ 1 - \delta(D_{ij}), & s_{ij} = 0 \end{cases} \quad (6)$$

其中， $\delta(D_{ij}) = 1 - \theta(D_{ij}) = \frac{1}{1 + e^{D_{ij}}}$ ， D_{ij} 为样本之间的距离，可以为欧氏距离、马氏距离以及加权内积距离等，在本文中相似性 $D_{ij} = \mathbf{w}^T (\mathbf{h}_i \odot \mathbf{h}_j)$ 为上文提出的加权内积距离。可以看出当样本之间的距离 D_{ij} 越小的时候，样本对被判为相同标签的概率值 $p(s_{ij} = 1|\mathbf{F})$ 越大，反之亦然，其中当 D_{ij} 趋于负无穷时 $p(s_{ij} = 1|\mathbf{F})$ 为 1，这些性质符合行人重识别任务的期望。则观测样本的似然函数可以表示为

$$P(S) = \prod_{(i,j) \in \mathcal{T}} p(s_{ij}) \quad (7)$$

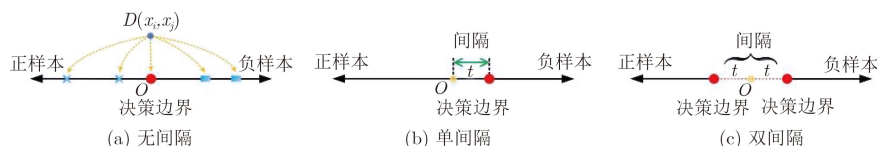
对式(7)取负对数可得

$$\mathcal{L} = - \sum_{(i,j) \in \mathcal{T}} ((1 - s_{ij}) \lg e^{D_{ij}} - \lg (e^{D_{ij}} + 1)) \quad (8)$$

优化该损失函数将把同类样本之间的距离压缩至负无穷 ($D_{ij}^+ \rightarrow -\infty$)，同时将不同类样本之间的距离扩大到正无穷 ($D_{ij}^- \rightarrow +\infty$)。

2.2.2 增大间隔的逻辑回归损失函数

为了进一步增强特征的判别性，本文引入了决策间隔来增大正负样本的相对距离，图2说明了间隔 t 对决策边界的影响(决策边界：使 $p^+ = p^- = 1/2$ 的点)。如图2(a)所示，以本文的加权内积距离为例，在没有加入间隔时，逻辑回归的决策边界在 $D_{ij} = 0$ 处，此时正负样本对之间没有安全距离。如图2(b)所示，而当对负样本的判决加入间隔 t 时，相当于负样本的决策边界向右平移了 t ，即只有当 $D_{ij} > t$ 的样本对才被判决为负样本对，而此时正样本的决策边界还在 $D_{ij} = 0$ 处，此时正负样本对之间的距离产生一个安全间隔 t ，有助于学习到判别性更强的特征。同理，如图2(c)所示，再对正样本的判决中加入间隔 t ，使正样本的判决点向左平移 t ，这样正负样本之间产生的安全间隔为 $2t$ 使得正负样本之间距离更大，被分得更开。加入间隔 t 的后验概率公式可以写为

图2 间隔 t 对决策边界的影响

$$\hat{p}(s_{ij}|\mathbf{F}) = \begin{cases} \frac{1}{1 + e^{D_{ij}+t}}, & s_{ij} = 1 \\ \frac{1}{1 + e^{t-D_{ij}}}, & s_{ij} = 0 \end{cases} \quad (9)$$

此时损失函数的表达式重写为

$$\mathcal{L} = \sum_{(i,j) \in \mathcal{T}} (s_{ij} \lg(e^{D_{ij}+t} + 1) + (1 - s_{ij}) \lg(e^{t-D_{ij}} + 1)) \quad (10)$$

在正负样本的判决中加入间隔 t 来扩大了正负样本的决策边界, 增大了正负样本对之间的相对距离, 增强了特征的判别性。

3 实验

本文基于Matconvnet深度学习框架, 在Market1501, CUHK03两个主流的行人重识别数据库上来验证本文算法的有效性。实验主要来验证以下3个问题: (1)本文提出的平移变化的加权内积距离是否优于其他的平移不变的距离度量方式。(2)本文引入的决策间隔 t 对性能的影响。(3)本文算法的性能是否优于其他先进算法。

3.1 实验数据集

Market1501数据库是行人重识别最大的数据库之一, 其由32668张来自不同摄像头拍摄得到的照片所构成, 总共有1501类, 其中训练集包含图片12936张, 类别数为751类、测试集有19732张图片, 共750类。

CUHK03收集于香港中文大学, 由6个不同摄像头在不同视场下拍摄得到。数据库总共包含14000多张、1467类行人图片, 其中训练集1367类, 测试集100类。本文用平均精确率(mAP)和累计匹配曲线(CMC)中的rank-1, rank-5和rank-10来作为评价指标。

3.2 实验设置

本文采用ResNet-50网络为孪生网络的骨干网络对本文模型进行验证, 网络首先在ImageNet数据库上进行预训练, 并在本文模型的约束下对参数进行精调。在进行精调时, 去掉网络中的最后一个全连接层, 提取出第5个池化层的值作为样本的特征。然后再利用这些特征进行相似性度量, 所有的输入图片在输入网络之前都裁剪到 224×224 以适合

ResNet50。本文采用梯度下降法就可以对模型进行端到端的优化。

3.3 实验结果

3.3.1 和其他相似性度量方法对比 为了验证本文提出的加权内积距离的有效性, 本文在Market1501数据库上进行实验, 把本文的加权内积距离和欧氏距离、马氏距离这两种平移不变的相似性度量方法嵌入到增大间隔的逻辑回归模型中进行性能对比。从表1数据来看加权内积距离的性能大大超过了马氏距离和欧氏距离这两种度量方法, 其中mAP分别提升了17.2%和6.59%。这是因为欧氏距离和马氏距离具有平移不变性, 使网络优化变得复杂, 且它们的范围都是从0到 $+\infty$ 。而本文提出的加权内积距离是平移变化的, 可以简化网络的训练而且优化范围更广, 为 $(-\infty, +\infty)$ 。实验说明了本文提出的加权内积距离在本文模型上可以有效地度量相似性。从表1中数据也可以看出, 基于本文的模型用马氏距离要比用欧氏距离的度量方式好, 因为马氏距离可以看成归一化(加权)的欧氏距离, 且它考虑了每个特征维度的尺度。

为了进行定性和直观的比较, 本文在Market1501数据库中随机选了6个行人, 并用t-SNE^[21]对其特征进行聚类。图3分别展示了在相同的数据上使用加权距离和马氏距离的t-SNE^[21]聚类图。这两种度量方式都是在本文提出的增大间隔的逻辑回归损失函数的约束下进行对比的, 且其中 t 都是选择的最优的值。图3(b)为使用马氏距离度量的结果, 其没有将第1, 3, 4类分得很开。使用本文提出的度量方式很好的将这6类行人分开了, 效果如图3(a)所示。图3说明了本文提出的加权距离度量方式要优于马氏距离这种度量方式。

3.3.2 决策间隔 t 对结果的影响 为了说明本文提出的间隔 t 对实验结果的影响, 本文在Market1501和

表1 与其他不同的距离度量方式相比

不同的度量方式	Market1501			
	rank-1	rank-5	rank-10	mAP
欧氏距离	55.08	75.03	82.63	37.53
马氏距离	65.68	83.82	88.93	48.14
加权内积距离(本文)	71.37	88.51	93.17	54.73

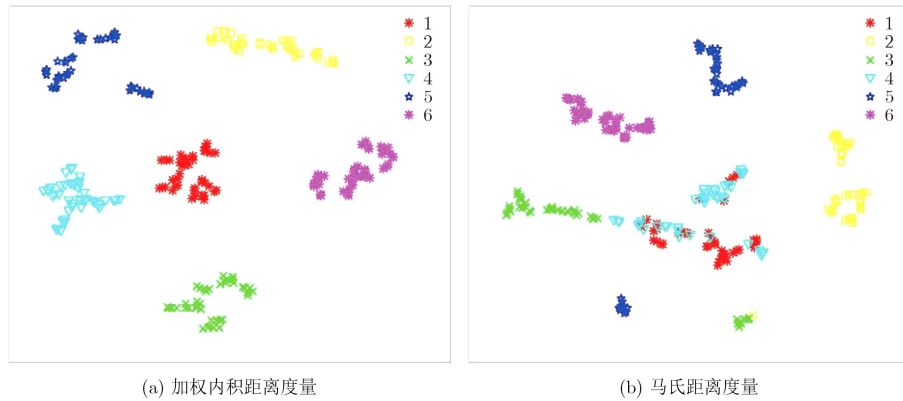


图 3 不同距离度量方式的t-SNE^[21]图

CUHK03两个数据库上分别使用马氏距离和加权内积距离两种度量方式进行实验说明。从图4可以看出，刚开始随着 t 的变化，两个数据库上的mAP均有很大的提升，随着 t 进一步的增加，性能均开始出现下降，这是因为大的间隔使得 $D=0$ 附近的点难以收敛。其中Market1501数据库上的实验结果如图4(a)所示，当 $t = 1.0$ 时用加权内积距离度量的性能取得最优且较之 $t = 0$ 时提高了4.83%，当 $t = 1.5$ 时用马氏距离度量的性能取得最优。图4(b)为CUHK03

数据库上的实验结果， $t = 0.5$ 时用加权内积距离度量的性能取得最优且较之 $t = 0$ 时提高了4.62%，可见正确地选择间隔 t 的值对性能的提升很关键，在不同的数据库上间隔 t 的最优值也会不同。

3.3.3 和其他先进算法的比较 本节在两个主流数据库上比较了本文模型和其他先进的行人重识别算法，实验中展示了本文算法在两种情况下的性能：(1)在没有间隔 t 的情况下本文的实验结果。(2)在最优的间隔 t 下本文的实验结果。

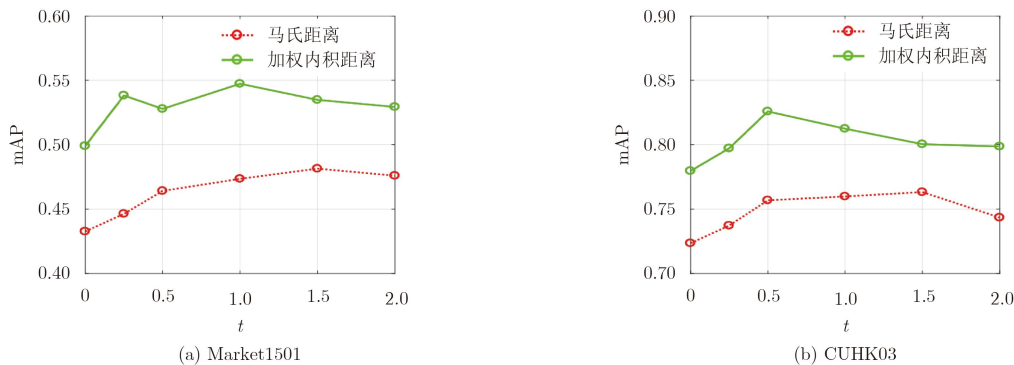


图 4 间隔 t 对结果的影响

在行人重识别领域，结合验证模型和类别分类模型已经是一种常用的方法。为了单纯地验证本文模型的有效性，本文没有结合分类模型，因而为了实验对比的公平性，本文的对比算法也都是没有结合分类模型的最新算法。在Market1501数据库上的结果实验如表2所示，粗体表示性能最好的算法。本文模型取得了54.73%的mAP, 71.37%的rank-1精确率，其中mAP, rank-5, rank-10在所有的对比算法里是最高的，rank-1为次高。本文算法的平均精确率(mAP)较之最新的相关算法DLGLE^[23]和SOMAnet^[25]算法分别提升了7.95%和6.84%。值得一提的是本文算法在间隔等于0的情况下，依然取得了较好的性能。

在CUHK03数据上的结果如表3所示，与DLCE^[6],

表 2 和先进算法在Market1501数据库上对比

方法	Market1501			
	rank-1	rank-5	rank-10	mAP
Null Space ^[22]	55.43	-	-	29.87
LSTM Siamese ^[19]	61.6	-	-	35.3
Gated Siamese ^[18]	65.88	-	-	39.55
DLCE (Pairwise) ^[6]	64.58	-	-	44.94
DLGLE(R) ^[23]	72.3	86.4	90.6	46.78
P2P ^[24]	61.31	-	-	35.71
P2S ^[24]	70.72	-	-	42.27
SOMAnet ^[25]	73.87	88.03	92.22	47.89
本文模型($t=0$)	64.96	85.78	91.3	49.9
本文模型	71.37	88.51	93.17	54.73

表3 和先进算法在CUHK03数据库上对比

方法	CUHK03			
	rank-1	rank-5	rank-10	mAP
BoW+KISSME	24.33	—	—	—
Deep ReID ^[8]	19.9	49.3	64.7	—
Null Space ^[22]	62.55	90.05	94.8	29.87
LSTM Siamese ^[19]	57.3	80.1	88.3	46.3
Gated Siamese ^[18]	68.1	88.1	94.6	58.84
DLCE (Pairwise) ^[6]	64.58	—	—	44.94
DLSGLE(R) ^[23]	73.2	93.7	97.2	—
SIR&CIR ^[10]	52.17	—	—	—
Triplet(Embedding) ^[26]	60.13	90.51	95.15	—
Triplet(Learned Metric)	61.60	92.41	97.47	—
SOMAnet ^[25]	68.9	91.1	95.6	—
本文模型($t=0$)	73.07	95.71	98.54	77.96
本文模型	78.81	96.72	99.02	82.58

DLSGLE^[23,24], P2P^[24], SOMAnet^[25]这些先进算法相比, 本文性能均有提升。其中SOMAnet算法的rank-1在Market1501上超过了本文算法2.5%, 但在CUHK03数据库上本文算法的rank-1较它提升了

9.91%。实验结果表明本文算法的性能较其他的算法有很大的提升, 说明了本文提出的增大间隔的逻辑回归模型较其他算法鲁棒性更好、泛化性更强, 得到的特征判别性更强。

3.3.4 可视化学习到的距离度量方式 为了直观地了解本文模型学习到的相似性度量方式, 本文在Market1501数据库的测试集选择了一部分样本用t-SNE^[21]可视化了样本图片的分布图。图5表示了从Market1501数据库的测试集随机选择了50类样本总共956张图片进行可视化的效果图。

4 总结

本文提出了一种平移变化的度量方式来度量样本之间的相似性, 减轻了传统平移不变的度量方式带来的特征解过多、网络参数难以训练的缺点。基于这种平移变化的相似性度量方法, 本文提出了一种增大间隔的逻辑回归模型来约束网络的学习特征, 约束了正负样本对之间的相对距离, 增强了特征的判别性。在Market1501和CUHK03这两种主流的数据集上进行了实验验证, 实验结果说明本文度量方式更加高效, 并且本文模型具有很好的鲁棒性, 得到的特征更具有判别性。

图5 本文模型在Market1501测试数据集上的t-SNE^[21]可视化图

参考文献

- [1] 周全, 魏昕, 陈建新, 等. 一种基于稠密SIFT特征对齐的稀疏表达人脸识别算法[J]. 电子与信息学报, 2015, 37(8): 1913–1919. doi: [10.11999/JEIT141194](https://doi.org/10.11999/JEIT141194).
ZHOU Quan, WEI Xin, CHEN Jianxin, et al. Improved sparse representation algorithm for face recognition via dense SIFT feature alignment[J]. *Journal of Electronics & Information Technology*, 2015, 37(8): 1913–1919. doi: [10.11999/JEIT141194](https://doi.org/10.11999/JEIT141194).
- [2] 张洁玉, 赵鸿萍, 陈曙. 自适应阈值及加权局部二值模式的人脸识别[J]. 电子与信息学报, 2014, 36(6): 1327–1333. doi: [10.3724/SP.J.1146.2013.01218](https://doi.org/10.3724/SP.J.1146.2013.01218).
ZHANG Jieyu, ZHAO Hongping, and CHEN Shu. Face recognition based on weighted local binary pattern with adaptive threshold[J]. *Journal of Electronics & Information Technology*, 2014, 36(6): 1327–1333. doi: [10.3724/SP.J.1146.2013.01218](https://doi.org/10.3724/SP.J.1146.2013.01218).
- [3] KOSTINGER M, HIRZER M, WOHLHART P, et al. Large scale metric learning from equivalence constraints[C]. IEEE Conference on Computer Vision and Pattern Recognition, Providence, USA, 2012: 2288–2295. doi: [10.1109/CVPR.2012.6247939](https://doi.org/10.1109/CVPR.2012.6247939).
- [4] XIAO Tong, LI Hongsheng, OUYANG Wanli, et al. Learning deep feature representations with domain guided dropout for person re-identification[C]. IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, USA, 2016: 1249–1258. doi: [10.1109/CVPR.2016.140](https://doi.org/10.1109/CVPR.2016.140).
- [5] DING Shengyong, LIN Liang, WANG Guangrun, et al. Deep feature learning with relative distance comparison for

- person re-identification[J]. *Pattern Recognition*, 2015, 48(10): 2993–3003. doi: [10.1016/j.patcog.2015.04.005](https://doi.org/10.1016/j.patcog.2015.04.005).
- [6] ZHENG Zhedong, ZHENG Liang, and YANG Yi. A discriminatively learned CNN embedding for person re-identification[J]. *ACM Transactions on Multimedia Computing, Communications, and Applications*, 2017, 14(1): 13–28. doi: [10.1145/3159171](https://doi.org/10.1145/3159171).
- [7] YI Dong, LEI Zhen, LIAO Shengcai, *et al.* Deep metric learning for person re-identification[C]. International Conference on Pattern Recognition, Stockholm, Sweden, 2014: 24–28. doi: [10.1109/ICPR.2014.16](https://doi.org/10.1109/ICPR.2014.16).
- [8] LI Wei, ZHAO Rui, XIAO Tong, *et al.* Deep reid: Deep filter pairing neural network for person re-identification[C]. IEEE Conference on Computer Vision and Pattern Recognition, Columbus, USA, 2014: 152–159. doi: [10.1109/CVPR.2014.27](https://doi.org/10.1109/CVPR.2014.27).
- [9] AHMED E, JONES M, and MARKS T K. An improved deep learning architecture for person re-identification[C]. IEEE Conference on Computer Vision and Pattern Recognition, Boston, USA, 2015: 3908–3916. doi: [10.1109/CVPR.2015.7299016](https://doi.org/10.1109/CVPR.2015.7299016).
- [10] WANG Faqiang, ZUO Wangmeng, LIN Liang, *et al.* Joint learning of single-image and cross-image representations for person re-identification[C]. IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, USA, 2016: 1288–1296. doi: [10.1109/CVPR.2016.144](https://doi.org/10.1109/CVPR.2016.144).
- [11] CHOPRA S, HADSELL R, and LECUN Y. Learning a similarity metric discriminatively, with application to face verification[C]. IEEE Conference on Computer Vision and Pattern Recognition, San Diego, USA, 2005: 539–546. doi: [10.1109/CVPR.2005.202](https://doi.org/10.1109/CVPR.2005.202).
- [12] SONG H O, XIANG Y, JEGELKA S, *et al.* Deep metric learning via lifted structured feature embedding[C]. IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, USA, 2016: 4004–4012. doi: [10.1109/CVPR.2016.434](https://doi.org/10.1109/CVPR.2016.434).
- [13] HIRZER M, ROTH P M, KOSTINGER M, *et al.* Relaxed pairwise learned metric for person re-identification[C]. European Conference on Computer Vision, Florence, Italy, 2012: 780–793. doi: [10.1007/978-3-642-33783-3_56](https://doi.org/10.1007/978-3-642-33783-3_56).
- [14] LIU Weiyang, WEN Yangdong, YU Zhiding, *et al.* SpheroFace: Deep hypersphere embedding for face recognition[C]. IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, USA, 2017: 6738–6746. doi: [10.1109/CVPR.2017.713](https://doi.org/10.1109/CVPR.2017.713).
- [15] WANG Jian, ZHOU Feng, WEN Shilei, *et al.* Deep metric learning with angular loss[C]. IEEE International Conference on Computer Vision, Honolulu, USA, 2017: 2593–2601. doi: [10.1109/ICCV.2017.283](https://doi.org/10.1109/ICCV.2017.283).
- [16] PEDAGADI S, ORWELL J, VELASTIN S A, *et al.* Local fisher discriminant analysis for pedestrian re-identification[C]. IEEE Conference on Computer Vision and Pattern Recognition, Portland, USA, 2013: 3318–3325. doi: [10.1109/CVPR.2013.426](https://doi.org/10.1109/CVPR.2013.426).
- [17] XIONG Fei, GOU Mengran, CAMPS O, *et al.* Person re-identification using kernel-based metric learning methods[C]. European Conference on Computer Vision, Zurich, Switzerland, 2014: 1–16. doi: [10.1007/978-3-319-10584-0_1](https://doi.org/10.1007/978-3-319-10584-0_1).
- [18] VARIOR R R, HALOI M, and WANG Gang. Gated siamese convolutional neural network architecture for human re-identification[C]. European Conference on Computer Vision, Amsterdam, Netherlands, 2016: 791–808. doi: [10.1007/978-3-319-46484-8_48](https://doi.org/10.1007/978-3-319-46484-8_48).
- [19] VARIOR R R, SHUAI B, LU Jiwen, *et al.* A siamese long short-Term memory architecture for human re-identification[C]. European Conference on Computer Vision, Amsterdam, Netherlands, 2016: 135–153. doi: [10.1007/978-3-319-46478-7_9](https://doi.org/10.1007/978-3-319-46478-7_9).
- [20] CHENG De, GONG Yihong, ZHOU Sanping, *et al.* Person re-identification by multi-channel parts-based CNN with improved triplet loss function[C]. IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, USA, 2016: 1335–1344. doi: [10.1109/CVPR.2016.149](https://doi.org/10.1109/CVPR.2016.149).
- [21] VAN DER MAATEN L J P and HINTON G E. Visualizing high dimensional data using t-SNE[J]. *Journal of Machine Learning Research*, 2008, 9(2): 2579–2605.
- [22] ZHANG Li, XIANG Tao, and GONG Shaogang. Learning a discriminative null space for person re-identification[C]. IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, USA, 2016: 1239–1248. doi: [10.1109/CVPR.2016.139](https://doi.org/10.1109/CVPR.2016.139).
- [23] CHENG De, GONG Yi, LI Zhe, *et al.* Deep feature learning via structured graph laplacian embedding for person re-identification[OL]. arXiv preprint arXiv: 1707.07791, 2017.
- [24] ZHOU Sanping, WANG Jinjun, WANG Jiayun, *et al.* Point to set similarity based deep feature learning for person re-identification[C]. IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, USA, 2017: 5028–5037. doi: [10.1109/CVPR.2017.534](https://doi.org/10.1109/CVPR.2017.534).
- [25] BARBOSA I B, CRISTANI M, CAPUTO B, *et al.* Looking beyond appearances: Synthetic training data for deep CNNs in re-identification[OL]. arXiv preprint arXiv: 1701.03153, 2017.
- [26] CHEN Weihua, CHEN Xiaotang, ZHANG Jianguo, *et al.* Beyond triplet loss: A deep quadruplet network for person re-identification[C]. IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, USA, 2017: 1320–1329. doi: [10.1109/CVPR.2017.145](https://doi.org/10.1109/CVPR.2017.145).
- 陈兵：男，1994年生，博士生，研究方向为计算机视觉、行人重识别。
- 查宇飞：男，1979年生，副教授，研究方向为计算机视觉及模式识别、目标检测、目标跟踪。
- 李运强：男，1992年生，博士生，研究方向为计算机视觉、二值图像检索及人脸识别。
- 张胜杰：男，1994年生，硕士生，研究方向为计算机视觉、二值图像检索。
- 张园强：男，1994年生，硕士生，研究方向为计算机视觉、视觉目标跟踪。