

基于改进 Center Loss 函数的行人再辨识

张璐¹, 刘阿建²

(1. 武汉大学, 湖北 武汉 430061;

2. 太原理工大学 信息与计算机学院, 山西 晋中 030600)

摘要:通过改进距离度量函数,在开集测试协议的基础上,对行人再辨识相关问题进行研究,使测量的行人特征满足以下两点:类间最小距离较大和类内最大距离较小。目前还没有存在的算法能够满足这个条件。文中采用 Center Loss 函数和分类损失函数相结合,使网络在分类损失与 Center Loss 函数的联合监督下,可以学习出更具判别性的行人特征。其中,行人特征分辨性问题分类损失函数能很好地解决,但常规的 Center Loss 函数只能使类内最大距离较小,但未能解决类间最小距离较大的问题。因此对 Center Loss 函数进行改进,在 Center Loss 函数中加入类间距离变量,使类间中心最小距离较大。最后通过几组再辨识数据集的实验证明了提出的网络与改进 Center Loss 函数的优越性。

关键词:行人再辨识;深度学习;Center Loss;分类损失函数

中图分类号:TP391

文献标识码:A

文章编号:1673-629X(2019)09-0045-06

doi:10.3969/j.issn.1673-629X.2019.09.009

Pedestrian Re-identification Based on Improved Center Loss

ZHANG Lu¹, LIU A-jian²

(1. Wuhan University, Wuhan 430061, China;

2. School of Information and Computer, Taiyuan University of Technology, Jinzhong 030600, China)

Abstract:By improving the distance measurement function and on the basis of the open-collection test protocol, pedestrian re-identification related problems are studied to make the measured pedestrian characteristics meet the following two requirements: the minimum distance between classes is large and the maximum distance within classes is small. At present, there is no algorithm can satisfy the two goals. In this paper, Center Loss function and classification loss function are combined to make the network learn more discriminative pedestrian characteristics under the joint supervision. Among them, the classification loss function can solve problem of pedestrian feature resolution very well. The conventional Center Loss function can only make the maximum distance within the class smaller, but it can't solve the problem that the minimum distance between classes is large. Therefore, we improve the Center Loss function by adding the variable of distance between different classes, which can make the minimum distance between different classes larger. Finally, the experiment shows that the improved classification loss function have a better realization.

Key words: pedestrian re-identification; deep learning; Center Loss; classification loss function

0 引言

常用的行人再辨识网络模型有两种:一种是基于样本对或三元组的验证网络模型。该模型由于早期公开的行人再辨识数据集较小被广泛采用,如 VIPeR^[1], 仅为每个身份的行人提供两幅图片;另一种极具潜力的是基于分类损失函数的分类模型^[2]。该模型最初应用在图像分类中,之后受到行人再辨识研究者的青睐得益于几个大数据集的公布,如 CUHK03^[3], Market1501^[4]和 DukeMTMC-reID^[5]。这使得训练一

个更深的网络而不出现过拟合成为可能。尽管分类损失函数不仅可以充分利用数据集的标注信息,而且无需将样本进行配对,但其主要缺陷是训练目标与测试过程不一致。如训练出来的分类网络模型在进行测试时,需要将分类损失函数层移除掉(开集测试过程)。一部分研究者^[6-7]结合两种损失函数的优点使行人特征具有较好的判别性。此方法与人脸深层验证网络^[8]有相类的地方,不同类的特征被分类损失函数分隔出来,同时类内距离小且类间距离大的问题被验证损失

收稿日期:2018-09-03

修回日期:2019-01-08

网络出版时间:2019-04-24

基金项目:湖北省自然科学基金(2018CFB136)

作者简介:张璐(1989-),女,硕士,工程师,研究方向为基于医院数据的云计算、大数据应用。

网络出版地址: <http://kns.cnki.net/kcms/detail/61.1450.TP.20190424.1004.008.html>

函数解决。这些方法跟目前常用的一些方法相比,在几组常用的行人再辨识数据集上有较优的表现,但是在网络训练之前,需要配对训练样本,样本种类越多,配对方式和数量越多。即使对样本进行严格的控制,但网络依然存在稳定性差、收敛慢等缺陷。

为了避免样本配对,并且使学习到的行人特征依然具有较好的判别能力,文中将 Center Loss 函数和分类损失函数 Softmax Loss 相结合,使网络在分类损失 Softmax Loss 与 Center Loss 函数的联合监督下,可以学习出更具判别性的行人特征。Softmax Loss 函数将不同类分隔开,同时 Center Loss 函数具有和验证函数类似的功能,能使类内距离变小,使行人特征更具判别性。文中将人脸识别函数 Center Loss 用于行人再辨识研究中,作用是使同类之间的差别变小,但是该函数不能解决不同类之间差距不大的问题。在同一个空间中,两个不同的类分布的很靠近或者部分重叠时,虽然每个类大部分都集中在中心点或距中心点距离较近的地方,但是判别性依旧不太理想。因此文中对 Center Loss 函数进行改进,在 Center Loss 函数中加入类间距离变量,使类间中心最小距离变大,使其具有较好的判别性。

1 Center Loss 函数

Center Loss 函数在进行网络构建的过程中,在分类网络 Softmax Loss 的基础上,在其特征提取层中加入损失函数 Center Loss,用来计算每个类中心的欧氏距离。这样对网络进行收敛后,同一类型的样本,其行人再辨识的特征就集中在此类别的中心。这种与不加入损失函数 Center Loss 的分类网络相比,同类别的行人特征类内的距离更小,其中 Center Loss 函数的定义如下:

$$L_C = \frac{1}{2M} \sum_{i=1}^M \|x_i - c_{y_i}\|_2^2 \quad (1)$$

其中, $c_{y_i} \in R^D$ 表示所属类别为 y_i 的特征 x_i 的中心点,并且 c_{y_i} 随训练批次的更新不断更新, y_i 样本数及特征的变化时, c_{y_i} 随之更新; M 表示样本数量。

另外,为了排除个别特殊样本给训练带来波动,定义中心学习率 α 。当网络的传播方向为反向时,其梯度 L_C 的定义见式 2, c_{y_i} 的更新见式 3:

$$\frac{\partial L_C}{\partial x_i} = x_i - c_{y_i} \quad (2)$$

$$\Delta c_j^t = \frac{\sum_{i=1}^M \delta(y_i = j) \cdot (c_j^t - x_i)}{1 + \sum_{i=1}^M \delta(y_i = j)} \quad (3)$$

$$c_j^{t+1} = c_j^t - \alpha \cdot \Delta c_j^t$$

其中, $\alpha \in [0, 1]$, 如果 $y_i = j$, 则 $\delta(\text{条件}) = 1$, 否则为 0。

文中是将 Softmax 函数和 Center Loss 函数相结合,因此总损失函数表示为:

$$L = L_1 + \lambda L_C = -\frac{1}{M} \sum_{i=1}^M \log \frac{e^{w_i^T x_i + b_{y_i}}}{\sum_{j=1}^N e^{w_j^T x_i + b_j}} + \frac{\lambda}{2M} \sum_{i=1}^M \|x_i - c_{y_i}\|_2^2 \quad (4)$$

其中 Softmax 函数和 Center Loss 函数的平衡系数用 λ 表示。当 $\lambda = 0$ 时,则该式表示仅使用 Softmax Loss 函数的分类网络。

2 Center Loss 函数的改进

标准的 Center Loss 函数在行人特征空间中,为每个类在训练过程中计算一个中心,并且使该类的特征向计算的中心点集中,其目标是减小类内距离,但是此函数不能解决不同类之间差距不大的问题。在同一个空间中,两个不同的类分布的很靠近或者部分重叠时,虽然每个类大部分都集中在中心点或距中心点距离较近的地方,但是判别性依旧不太理想。因此在训练过程中,既考虑类内距离又考虑类间距离,通过对 Center Loss 函数进行改进,在 Center Loss 函数中加入类间距离变量,使类间中心最小距离变大,使其具有较好的判别性。改进 Center Loss 函数的定义为:

$$L_{IC} = \frac{1}{2M} \sum_{i=1}^M \|x_i - c_{y_i}\|_2^2 - \frac{1}{2 \cdot N^2} \sum_{j=1}^N \sum_{k=1}^N \|c_{y^j} - c_{y^k}\|_2^2 \quad (5)$$

因为类内中心距离矩阵具有对称性,文中将式 5 进行简化,即:

$$L_{IC} = \frac{1}{2M} \sum_{i=1}^M \|x_i - c_{y_i}\|_2^2 - \frac{1}{2 \cdot C_N^2} \sum_{j=1}^N \sum_{k=j+1}^N \|c_{y^j} - c_{y^k}\|_2^2 \quad (6)$$

其中 N 表示一次训练样本中不同类的个数,即样本中有 N 个类中心,如果将不同类的所有类间距离加入损失函数,一共会有 C_N^2 个。那么随着 N 的增加,计算复杂程度会急剧增加,特别是当网络传播方向反转时,梯度函数的计算会更加困难,并且不利于网络收敛。针对以上问题,文中只将 C_N^2 个类间距离中最小距离加入到损失函数中,也就是使每一次迭代,将距离最小的两个中心之间的距离进行扩大,同时也使梯度函数的计算难度降低。另外在网络传播方向反向时,网络参数的力度随网络梯度的更新而更新,在损失函数中加入参数 β 对不同类间距离在函数中比重进行控制。其中 $\beta \in [0, 1]$, 当 $\beta = 0$ 时,该式表示没有改进

的、标准的 Center Loss 函数的损失函数。因此改进的 Center Loss 函数为:

$$L_{IC} = L_C + L_{IMPR} = \frac{1}{2M} \sum_{i=1}^M \|x_i - c_{y_i}\|_2^2 - \beta \cdot \frac{1}{2} \|c_{y^A} - c_{y^B}\|_2^2 \tag{7}$$

$$\text{s. t. } \forall j, k \in [1, N], j \neq k, \text{不等式 } \|c_{y^A} - c_{y^B}\|_2^2 \leq \|c_{y^j} - c_{y^k}\|_2^2 \text{ 恒成立}$$

在反向传播中, L_{IMPR} 的梯度计算为:

$$\Delta L_{IMPR} = -\beta(c_{y^A} - c_{y^B}) \cdot (\Delta c_{y^A} - \Delta c_{y^B}) = -\beta(c_{y^A} - c_{y^B}) \cdot \left\{ \sum_{i=1}^M [\delta(y_i = y^A) \cdot (c_{y^A} - x_i) - \delta(y_i = y^B) \cdot (c_{y^B} - x_i)] \right\} \tag{8}$$

3 实验与分析

3.1 Mnist 数据集上的实验

本节通过实验展示网络学习行人再辨识特征分布的图像,直观感受改进的网络模型是否具有较优的表现,同时通过定性定量分析确定改进算法是否更加适合行人再辨识的应用。实验参考文献[9]中的 Mnist 数据集^[10]和 LeNets++网络。为了便于观察结果,将选定网络的输出维度设置为 2,即学习到的行人再辨识的特征为 2 维,可以通过二维平面显示实验结果。实

验中参数设置与文献[9]相同,除特殊说明外,实验随机选取样本图像。文中实验样本数量设置为 10k。

采取四组网络模型进行行人再辨识实验,四组网络模型分别是:仅使用 Softmax Loss 函数作为损失函数;仅使用 Center Loss 函数作为损失函数;Softmax 与标准 Center Loss 相结合作为损失函数;Softmax 与改进的 Center Loss 相结合作为损失函数。把参数 λ 和 β 分别设置为 0.1 和 1。图 1 一共四组图,分别对应上述 4 种损失函数。

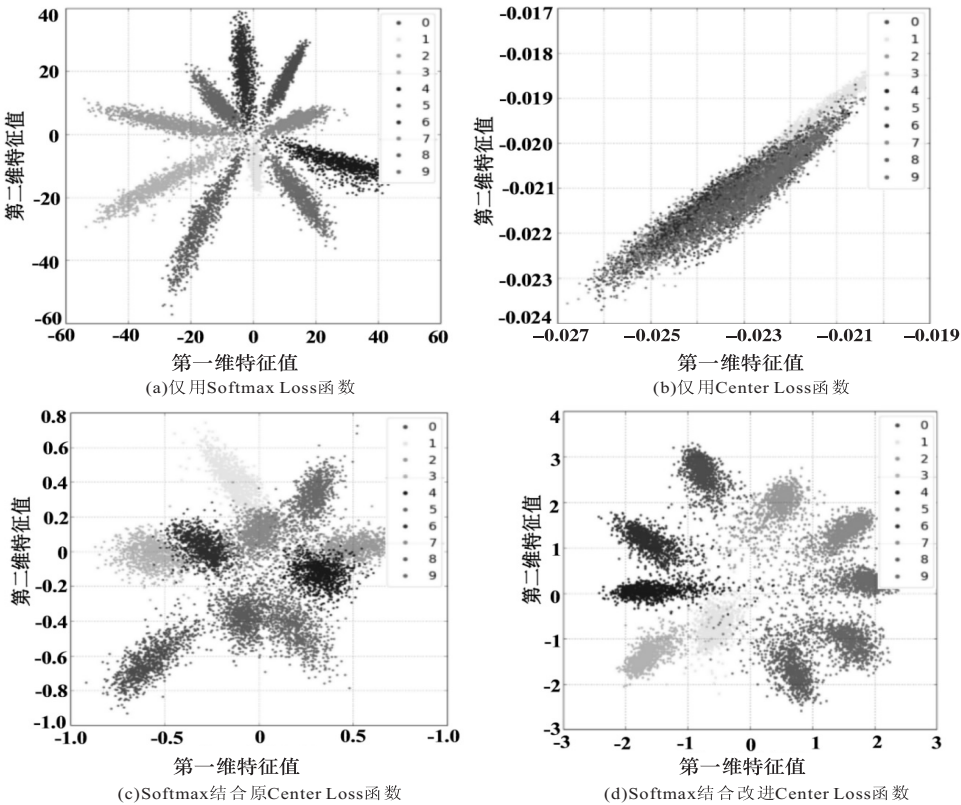


图 1 Mnist 数据集图像二维特征的分布

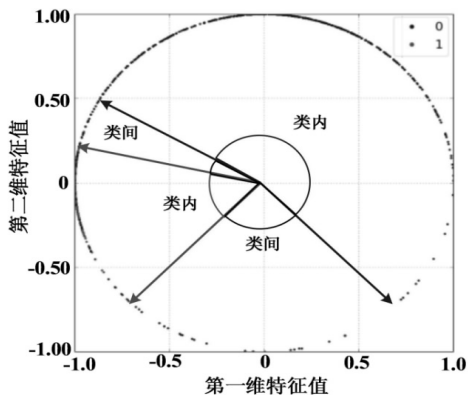
从图 1(a)可知,仅使用 Softmax Loss 函数作为损失函数,不同类的特征之间能很好地区分开,但是同类之间的类内距离不能很好的聚集;从图 1(b)可知,仅使用 Center Loss 函数作为损失函数,同类之间的类内距离控制在-0.03 ~ -0.02 之间,但是不同类也重叠在一起,不能很好地区别开;从图 1(c)可知,Softmax 与标准 Center Loss 相结合作为损失函数,同类之间的类

内距离控制在-0.75 ~ 0.75 之间,并且不同类之间没有重叠,部分已能区分开,但是部分依旧有重叠,例如类别 1 和类别 7 有部分特征重叠,这可能影响经过该网络进行行人再辨识特征的提取分类结果的准确性。从图 1(d)可知,以 Softmax 与改进的 Center Loss 相结合作为损失函数,同类之间的类内距离控制在-2 ~ 2 之间,同时所有不同类别的类能很好地分离开,并且类

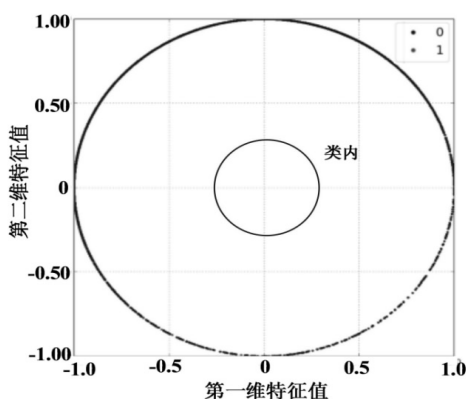
间距离比较大,在 $-3 \sim 4$ 之间,使行人再辨识特征的提取更加容易。实验结果表明,在 Mnist 数据集和 LeNets++网络基础上,改进的 Center Loss 和 Softmax 函数相结合的方法能使类内距离较小,类间距离较大,能够很好地满足行为再辨识的要求。

样本数设置为 10k 和 50k 进行实验,再对实验结果进行归一化处理,处理结果如图 2 和图 3 所示。从图 2(a)可以看出,10k 样本数时,提取的特征值基本

能分开,但样本数达到 50k 时,图 2(b)显示提取的特征值基本无法区分。从图 3(a)和图 3(b)可以看出,样本数不管是 10k 还是 50k,都能满足类间距离较大类内距离较小的要求。由实验结果可知,Softmax 与改进的 Center Loss 相结合作为损失函数的方法比 Softmax 与标准 Center Loss 相结合作为损失函数的方法更优,尤其是在样本数较多的情况下,改进的方法优势更加明显。

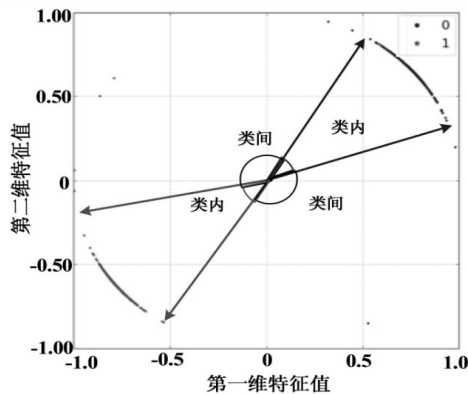


(a) 10k 幅图像特征分布

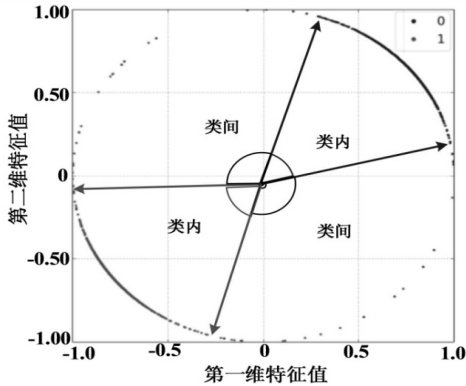


(b) 50k 幅图像特征分布

图 2 Softmax 结合原 Center Loss 函数的归一化特征



(a) 10k 幅图像特征分布



(b) 50k 幅图像特征分布

图 3 Softmax 结合改进 Center Loss 函数的归一化特征

3.2 行人再辨识实验

3.2.1 数据集

选取了 4 个常用于行人再辨识研究的数据集进行实验,即 CUHK03 与 Market1501、VIPeR 与 CUHK01^[11],其中前两个是大数据集,后两个是小数据集。CUHK03 与 Market1501 通过 DPM 检测器对行人图像进行裁剪,这样的数据集更接近实际应用。另外,该实验除了 Market1501 均采用 single-shot 的测试方法,而 Market1501 采用的是 multi-shot 的测试方法。single-shot 的测试方法是指在数据集中有且仅有一幅图像与目标图像来源于同一个行人, multi-shot 的测试方法是指在数据集中有多幅图像与目标图像来源于同一个行人。实验的数据集均根据各自的测试协议对样本进行设置和数据集进行训练,测试报告按测试协

议对测试结果取平均值。

3.2.2 数据输入准备

实验第一步是准备数据集图像,将所有图像设置成 128×48 的大小;然后通过所有图像计算得到训练集的均值,并把均值图像剔除出来;再通过二维变换把剩下的图像进行扩充,增长 7 倍;最后把幅度范围控制在 -6 到 6 之间,将图像中心设置为左边原点,随机向坐标轴方向平行移动。

3.2.3 网络框架与基本网络

文中使用的基础网络是 CNN^[12] 与 Attention-39^[13](网络的详细介绍参见文献[12-13]),这种网络的主要优点是在反向传播的过程中,防止出现梯度爆炸式增长或消失不见等情况,同时有利于网络快速收敛;基本网络的最后一个全连接层输出 512 维的特征

FC ;除了全连接层,其他所有层后面都连接激活函数,该函数是线性单元且带有修正参数。

3.2.4 测试方法

因为行人再辨识是一个开集测试过程,即测试样本与训练样本的类别没有交集。因此分类子网络在测试过程中是无效的。首先通过基础网络对准备的目标图像进行特征提取,同时提取待测图像的特征;然后计算每一个待测图像特征与目标图像特征的余弦值,构建一个 $M \times N$ 的矩阵,其中 M 表示选取的目标图像数量, N 表示选取的数据集图像数量。实验将特征值进行归一化处理,最后得到的余弦值就是欧氏距离。

表 1 CUHK03 数据集上的实验结果比较

CUHK03 (SQ)	Rank-1	Rank-5	Rank-10
LOMO+XQDA ^[14]	46.3	78.9	88.6
Gate-SCNN ^[15]	68.1	80.9	88.3
CNN-FRW-I-C ^[16]	82.1	96.2	98.2
ResNet-50 ^[6]	83.4	97.1	98.7
Deep Transfer ^[7]	84.1	—	—
CNN-I-V(E) ^[16]	80.2	94.9	97.3
CNN-I-C	80.2	96.1	97.0
CNN-I-IC	82.8	97.5	98.6
Attention-39-I-C	82.6	97.1	98.0
Attention-39-I-IC	84.0	98.2	98.7

从表中结果可以看出,两种基础网络“I-C”的结果要优于“I”的结果,表明 Center Loss 函数能有效减小类内距离;“I-IC”的结果除了 Rank-1 项,其他均优于“I-C”的结果,表明文中的改进算法可以有效减小类内距离,同时增加了类间间距;“Attention-39-I-IC”算法跟当前最好的几种算法相比较,也显示出了优势。

在 Market1501 数据集上,文中算法同样优于目前很多算法,如表 2 所示。然而,在两个小数据集上的结果表明(见表 3),DNS 网络模型更适用于小数据集,有较好的表现。

表 2 Market1501 数据集上的实验结果比较

Method	Single-shot		Multi-shot	
	Rank-1	mAP	Rank-1	mAP
LOMO+XQDA ^[14]	43.8	22.2	54.1	28.4
Gate-SCNN ^[15]	65.88	39.55	76.04	48.45
ResNet-50 ^[6]	79.51	59.87	85.84	70.33
Deep Transfer ^[7]	83.7	65.5	89.6	73.8
CNN-I-IC	79.2	60.5	85.2	69.7
Attention-39-I-IC	82.3	63.1	87.4	71.3

从实验结果可知,Deep transfer 算法表现要优于文中改进算法,原因是 Deep transfer 算法使用分步转

实验采用 Rank-1-5-10 和 mAP (mean average precision),此方法操作方便且节省空间。并且将实验结果与当前最好的同类实验结果进行比较分析,

3.2.5 实验结果与分析

基于 CUHK03 数据集的实验结果见表 1,表中不特殊说明,“I”表示仅使用 Softmax Loss 为损失函数,“I-C”表示使用 Softmax Loss 结合标准的 Center Loss 为损失函数,“I-IC”表示使用 Softmax Loss 结合改进的 Center Loss 为损失函数,CNN 与 Attention-39 是文中选定的两个基础网络。

化的方法,首先根据训练集精调网络模型,再进行测试。另外文改进算法在小数据集上表现一般,原因是文中算法要求充足的样本对每一个类进行训练,训练不够会引起网络的过拟合。虽然该算法在小数据集上性能不好,但是改进的 Center Loss 函数相比原函数,依然有一个点左右的提升。

表 3 VIPeR 与 CUHK01 数据集上的实验结果比较

Method	VIPeR		CUHK01	
	Rank-1	Rank-5	Rank-1	Rank-5
IDLA ^[17]	34.8	63.6	47.5	71.6
$\ell 1$ GL ^[18]	41.5	—	50.1	—
MCP-CNN ^[19]	47.8	74.7	53.7	84.3
DNS ^[20]	51.2	82.1	69.1	86.9
CNN-FRW-I-C ^[16]	50.4	77.6	70.5	90.0
Deep Transfer ^[7]	56.3	—	77.0	—
CNN-I-C ^[16]	49.3	77.3	70.1	90.5
CNN-I-IC	50.2	77.6	70.5	91.3
Attention-39-I-IC	48.8	76.9	68.6	85.8

4 结束语

文中对行为再辨识算法进行研究,Center Loss 函数能使同类类内距离较小,加入 Center Loss 函数的方

法能使不同类之间被区分开,但由于类间间距较小,影响特征分类的准确性。所以文中对 Center Loss 函数进行改进,该算法满足了减小类内最大距离和增加类间最小距离的要求。实验结果表明,文中改进算法优于标准函数算法,同时也优于当前很多比较好的算法,特别是在大数据集上表现更加明显。

参考文献:

- [1] GRAY D, TAO Hai. Viewpoint invariant pedestrian recognition with an ensemble of localized features[C]//Proceedings of the 10th European conference on computer vision. Marseille, France; Springer-Verlag, 2008: 262–275.
- [2] ZHENG Liang, YANG Yi, HAUPTMANN A G. Person re-identification: past, present and future[J]. Journal of Latex Class Files, 2015, 14(8): 110–130.
- [3] LI Wei, ZHAO Rui, XIAO Tong, et al. DeepReID: deep filter pairing neural network for person re-identification[C]//IEEE conference on computer vision and pattern recognition. Columbus, OH, USA; IEEE, 2014: 152–159.
- [4] ZHENG Liang, SHEN Liyue, TIAN Lu, et al. Scalable person re-identification: a benchmark[C]//IEEE international conference on computer vision. Santiago, Chile; IEEE, 2015: 1116–1124.
- [5] ZHENG Zhedong, ZHENG Liang, YANG Yi. Unlabeled samples generated by GAN improve the person re-identification baseline in vitro[C]//IEEE international conference on computer vision. [s. l.]: IEEE, 2017: 3774–3782.
- [6] ZHENG Zhedong, ZHENG Liang, YANG Yi. A discriminatively learned CNN embedding for person re-identification[J]. ACM Transactions on Multimedia Computing Communications & Applications, 2016, 14(1): 375–385.
- [7] GENG Mengyue, WANG Yaowei, XIANG Tao. Deep transfer learning for person re-identification[C]//Computer vision and pattern recognition. [s. l.]: [s. n.], 2016: 1771–1783.
- [8] 杨红玲, 宣士斌, 莫愿斌. 基于卷积神经网络的手势识别[J]. 计算机技术与发展, 2018, 28(7): 11–14.
- [9] WEN Yandong, ZHANG Kaipeng, LI Zhifeng, et al. A discriminative feature learning approach for deep face recognition[C]//European conference on computer vision. [s. l.]: Springer International Publishing, 2016: 499–515.
- [10] LI Deng. The MNIST database of handwritten digit images for machine learning research[J]. IEEE Signal Processing Magazine, 2012, 29(6): 141–142.
- [11] LI Wei, ZHAO Rui, WANG Xiaogang. Human reidentification with transferred metric learning[C]//Asian conference on computer vision. Berlin; Springer, 2012: 31–44.
- [12] JIN Haibo, WANG Xiaobo, LIAO Shengcai, et al. Deep person re-identification with improved embedding and efficient training[C]//Computer vision and pattern recognition. [s. l.]: [s. n.], 2017: 554–568.
- [13] 周飞燕, 金林鹏, 董军. 卷积神经网络研究综述[J]. 计算机学报, 2017, 40(6): 1229–1251.
- [14] LIAO Shengcai, HU Yang, ZHU Xiangyu, et al. Person re-identification by local maximal occurrence representation and metric learning[C]//IEEE conference on computer vision and pattern recognition. Boston, MA, USA; IEEE, 2015: 2197–2206.
- [15] VARIOR R, HALOI M, WANG Gang. Gated siamese convolutional neural network architecture for human re-identification[C]//European conference on computer vision. Amsterdam; Springer International Publishing, 2016: 791–808.
- [16] JIN Haibo, WANG Xiaobo, LIAO Shengcai, et al. Deep person re-identification with improved embedding and efficient training[C]//IEEE international joint conference on biometrics. Denver, Colorado; IEEE, 2017: 261–267.
- [17] AHMED E, JONES M, MARKS T K. An improved deep learning architecture for person re-identification[C]//IEEE conference on computer vision and pattern recognition. Boston, MA, USA; IEEE, 2015: 3908–3916.
- [18] KODIROV E, XIANG Tao, FU Zhenyong, et al. Person re-identification by unsupervised, graph learning[C]//European conference on computer vision. Amsterdam; Springer International Publishing, 2016: 178–195.
- [19] 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, NV, USA; IEEE, 2016: 1335–1344.
- [20] ZHANG LI, XIANG Tao, GONG Shaogang. Learning a discriminative null space for person re-identification[C]//IEEE conference on computer vision and pattern recognition. Las Vegas, NV, USA; IEEE, 2016: 1239–1248.