

一种基于ResNet模型及重排序策略的行人重识别算法

黄海新, 陶文博

沈阳理工大学自动化与电气工程学院, 辽宁沈阳 110159

摘要: 基于深度学习的发展, 提出了一种行人重识别算法。算法的流程是先使用ResNet模型提取图像中的深层特征, 然后在低相似度约束下选择一组关联图像, 最后使用关联图像对检索结果进行重新排序。该算法具有三个优点: 与其他方法相比, 迁移学习方法降低了训练模型所需的成本; 采用自适应稀疏约束下的遗传算法检测相关图像, 降低了计算的复杂度; 可以将检测结果应用于改善初始检索结果。在Market1501数据集上的实验证明了算法是有价值的, 并且算法对于背景、姿态、视角的变化具有一定的鲁棒性。

关键词: 行人重识别; 残差网络; 重排序; 迁移学习

中图分类号: TP391.4

文献标识码: A

文章编号: 1672-0164 (2023) 02-0025-04

1 引言

行人重识别 (ReID, e-Identification) 旨在从图库中查找包含相同人物的图像, 达到定位目标人物的目的。这些图像在非重叠环境下拍摄的。行人重识别技术是监控系统中的重要任务, 具有多种应用场景, 重识别因其巨大的潜力在过去几十年中得到了迅速发展。

经过多年的研究, 重识别技术仍然面临许多严峻挑战。行人的图像在背景、姿态、照明、视角上有很大的差异, 因为这些图像是在非重叠环境下拍摄的。为了解决这些问题, 传统方法主要从两个方面进行研究: 设计鲁棒的特征描述符和学习距离度量^[1]。一些有代表性的方法, 如对称驱动特征^[2]、概率颜色直方图^[3]和PSD约束的非对称度量学习^[4], 促进了实验结果。采用深度学习可以减少人为干

预对模型的影响。Ahmed^[5]提出了一个二值分类网络, 以确定两幅图像是否包含相同的行人。Xiao^[6]混合了所有数据库, 以预先训练模型并对目标数据库进行微调。目前的主要研究思路是研究查询图片与参考集中每个图片之间的距离, 获取一个初始排序列表。在此基础上添加一个重排序步骤, 可以改善搜索结果。

文中提出一种使用ResNet模型和关联图片的新的重排序算法。主要工作如下:

①对预先训练好的ResNet模型进行微调, 并利用训练好的模型从图像中提取特征, 得到初始排名表。

②使用自适应稀疏约束下的遗传算法检测相互关联的相关图像。将原始距离与相关图片和参考集图片的距离相加得到最终结果。算法整体框架如图1。在Market1501上的实验证明了算法的有效性。

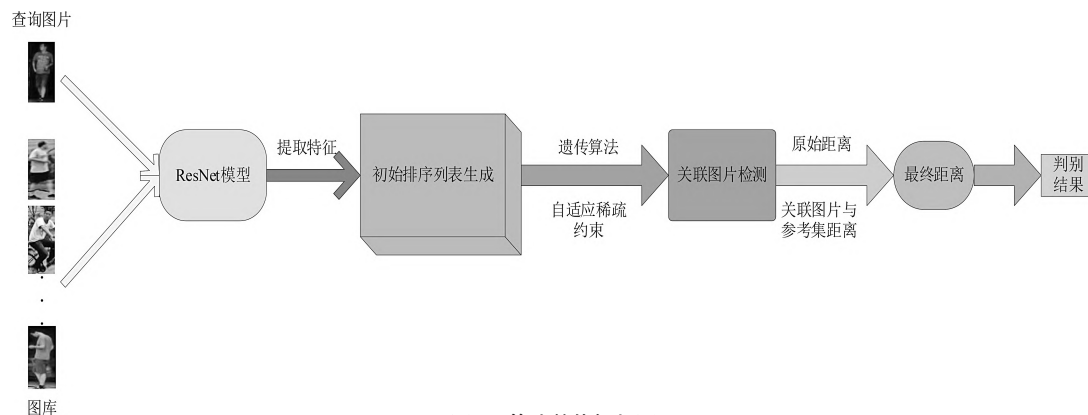


图1 算法整体框架图

2 使用的方法

2.1 问题定义

给定一个查询图片以及包含 N 张图片的参考集 $G = \{g_i | i = 1, 2, \dots, N\}$, 获取一个初始排序列表 $\bar{L}(p, G) = \{g_i^0 | i = 1, 2, \dots, N\}$, 简化为 \bar{L} , 度量两张图片 p 和 g_i 的距离 $d(p, g_i) = (x_p - x_{g_i})^T M (x_p - x_{g_i})$, 其中 x_p 与 x_{g_i} 分别代表查询图像 p 和参考集图片 g_i 的特征, M 是半正定矩阵。算法的目标是在初始排序列表中检测一组图片, 使用关联图像进行重排序, 期望得到正样本排名更高的排序结果。

2.2 初始排序列表生成

为了保证特征提取的有效性并且降低计算成本, 对目标行人重识别数据集上的预处理神经网络模型进行微调。迁移学习由于具有节约时间与计算量的优点, 在实际应用上十分普遍。选择 ResNet 进行特征提取。ResNet 采用跳跃连接解决了深度网络的梯度消失和梯度爆炸问题。网络基本结构如图 2 所示。

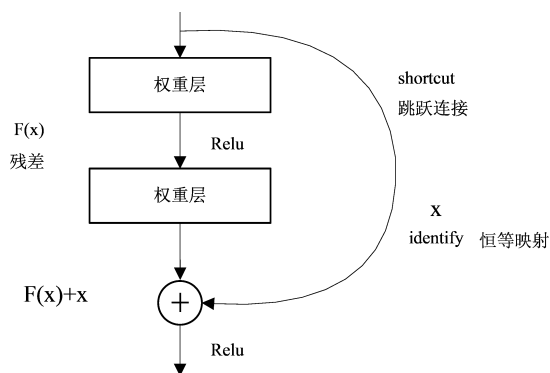


图2 残差网络基本结构

对网络模型进行微调后, 将模型的输出作为图片中提取到的特征。查询图片 p 以及参考集图片 g_i 之间的距离定义如公式 (1):

$$d(x_p, x_{g_i}) = x_p \odot x_p + x_{g_i} \odot x_{g_i} - 2 * x_p \odot x_{g_i} \quad (1)$$

初始排序列表是根据参考集图片与查询图片的距离从大到小排序获得的。令 \bar{L} 中的图片为 $\bar{G} = \{\bar{g}_i | i = 1, 2, \dots, N\}$ 。

2.3 关联图片

如前所述, 与查询图库的关联图片应该具有高度的相似性, 因为预期的个人图片在不同的相机中共用相同的视觉模式。这意味着正样本到初始排序列表中所有其他图片的总距离应小于负样本的总距离, 因为负样本与排序列表中的其他图片不具有类似的视觉模式。我们在初始排序列表中命名一组图片, 这些图片与其他关联图片非常相似。为了从初始序列中的前 M 项中检测相关图片, 命名一个矩阵 $K \in R^{M \times M}$, 其中 K_{ij} 是 $x_{\bar{g}_i}$ 与 $x_{\bar{g}_j}$ 的距离, 令 $s = \{s_i | i = 1, 2, \dots, M\}$ 为参考图库图片 $x_{\bar{g}_i}$ 和前 M 项每张图片的总

和, 见公式 (2)。

$$s_i = \sum_{j=1}^M d(\bar{g}_i, \bar{g}_j) \quad (2)$$

令 $c \in \{0, 1\}^M$ 为初始排序列表中检测相关图片的索引, $c_i = 1$ 代表图片 \bar{g}_i 被选为关联图片, $c_i = 0$ 代表图片未被选为关联图片。如果一个预期图像与其他图像的总距离很短, 该预期图像很有可能与其他图片很相似, 并且可能是一个关联图片。为了找到彼此最相似的图片集, 我们的目标是使 $\|s - Kc\|$ 的结果最大化。最后为 c 添加一个稀疏性约束来保证所选图片的精度。

$$\max \|s - Kc\|_p^q + a \|c\|_2 \quad (3)$$

$$\text{s.t. } c \in \{0, 1\}^M$$

其中 a 是 c 的稀疏性约束。为了简化问题, 我们对原问题的对偶问题进行解决。如公式 3。

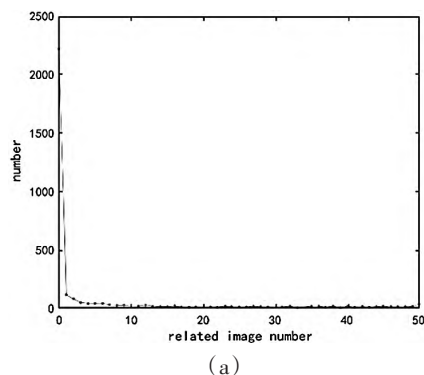
$$\max \|s - Kc\|_p^q + b \|\bar{c}\|_2 \quad (4)$$

$$\text{s.t. } \bar{c} \in \{0, 1\}^M$$

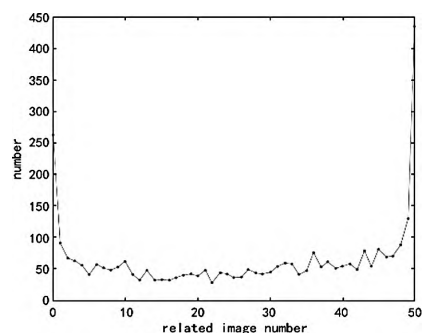
其中 c 是非关联图像的索引, c 等于 \bar{c} 的翻转。 b 是约束, 确保选中大多数的非关联图片。

前 50 项中检测关联图像。横轴表示初始排序列表中关联图像的数量, 纵轴表示初始排序列表的数量。(a) 是 $b=20$ 时的结果, (b) 是 $b=30$ 时的结果。

原问题和对偶问题都是 NP-hard (Non-deterministic Polynomial hard problem) 问题, 由于计算量巨大, 穷举搜索法并不可行。采用优化方法获得次优解。无法检测在相



(a)



(b)

图3 b 对检测关联图片的影响

同 b 值下不同初始排序列表的关联图像, 因为无法保证从不同初始排序列表检测到的所有关联图像都满足稀疏性约束, 如图 3 所示。提出了一种基于自适应稀疏参数的关联图像检测算法。为每个初始排序列表设置稀疏参数的上下限, 在当前 b 下, 通过二进制搜索适当的 b 值, 迭代计算特定初始排序列表的关联图像检测结果, 直到关联图像的数量符合稀疏约束。使用遗传算法检测相关图像, 因为 c 是不可微的, 关联图像的数量取决于 b 的值, 从不同权重的期望初始排序列表中检测关联图像的过程可能需要多次迭代, 直到获得最终结果。选择遗传算法的原因是它的计算量较小, 且可以解决非微分问题。

2.4 最终距离

原始距离的重要性不容忽视, 将原始距离和关联图片与参考集图片之间的距离联合起来, 最终距离 $d(g_i)$ 定义如公式 5。

$$\begin{aligned} d(g_i) &= d(x_p, x_{gi}) + vd(g_i, c) \\ d(g_i, c) &= \sum_{j=1}^N c_j * d(g_i, \bar{L}_j) \end{aligned} \quad (5)$$

其中 $v \in [0, 1]$ 表示关联图片与参考集图片之间距离的权重因子, $d(g_i, c)$ 是参考集图片到所有关联图片的总距离。关联图片的最终距离应低于非关联图片的总距离。当 $v=0$ 时, 仅考虑原始距离。

3 实验结果

在 Market1501 数据集上进行实验。Market1501 是当前最大的行人重识别数据集, 其中查询集包括 3368 张图片, 包含 750 个身份; 训练集包括 12936 张图片, 包含 751 个身份; 测试集包括 19732 个图像, 包含 750 个身份。使用 RANK-N 与 mAP 这两个评价指标与已有的一些算法进行比较, 以此评估提出的基于 ResNet 模型与重排序策略的行人重识别算法。Rank-N 指标衡量在所有查询中, 在前 N 项中查询到的正样本所占的比例。较高的秩 N 意味着该方法可以更快地在前 N 项中找到至少一个正样本。mAP 是平均精度, 反映了全局性能。

通过与以下算法的比较, 证明了该方法的有效性。传统方法, 如 NFST^[7] 是一种典型的度量学习算法, 它通过投影矩阵将高维向量投影到低维向量。BoW^[8] 是一种编码特征的方法, 旨在减少类内距离, 增加类间距离。WARCA^[9] 是一种基于马氏距离的度量算法。采用深度学习算法提高了模型效果。PersonNet^[10] 和 S-CNN^[11] 都是二元分类模型。CAN^[12] 提出了一种基于软注意力的模型。结果如表 1 所示。

结果表明文中的算法优于已有的部分传统及深度学习算法, 具有一定的价值。由 ResNet 模型生成的初始排序列表比上述传统方法和深度学习方法获得了更好的效果。基于关联图像的重排序阶段对排序结果进行了改进。主要原

表 1 在前 5 项中检测 2~3 张关联图片的结果

	Rank1	Rank5	mAP
NFST [7]	56.6	-	-
PersonNet [10]	49.7	-	-
S-CNN [11]	67.5	-	-
BoWbest [8]	46.2	62.8	-
WARCA [9]	47.23	-	-
CAN [12]	49.15	-	23.28
ResNet	72.8	87.66	50.95
ResNet+re-rank	73.29	88.29	52.47

因是 ResNet 从图片中提取到的特征对外界的复杂干扰具有鲁棒性, 并且特征可以很好地表示图片。同时, 从初始排序列表中检测到的图片彼此相似度很高, 关联图片与查询图片相关的概率也更高。在理想情况下, 最终排序结果是多摄像头中预期个体的总距离, 这可以缓解在非重叠环境下采集的同类样本之间的差异。

4 结 语

一种基于 ResNet 模型及重排序策略的行人重识别算法使用 ResNet 模型生成初始排序列表, 然后在列表中检测关联图像, 并将原始距离和参考集图片与关联图片之间的距离聚合, 以获得最终的排序列表。在 Market1501 上进行的实验表明, 与一些常用的算法相比, 文中的算法取得了更好的排序结果。

参 考 文 献

- [1] LIU C., EN C L., C, N C L. Person re-identification; hat features are important? [C]//International Conference on Computer Vision. 2012:391-401.
- [2] FARENZENA M, BA, ANI L and PERINA A. Person re-identification by symmetry-driven accumulation of local features [C]//Computer Vision and Pattern Recognition. 2010,60:2367.
- [3] D'ANGELO A, DUG, AY J L. People re-identification in camera networks based on probabilistic color histograms [C]//Proceedings of SPIE-The International Society for Optical Engineering. 2011,2(5):2157.
- [4] LIAO S, LI S., Efficient PSD Constrained Asymmetric Metric Learning for Person Re-Identification [C]//IEEE International Conference on Computer Vision. 2015,36:3693.
- [5] AHMED E, JONES, MARKS, K. An improved deep learning architecture for person re-identification [C]//Computer Vision and Pattern Recognition. 2015,390:3916.
- [6] XIAO T, LI H, OU, NG W, Learning Deep Feature Representations with Domain Guided Dropout for Person Re-identification [C]//Computer Vision and Pattern Recognition. 2016,1249:258.

(下转第 58 页)

Research and implementation of multi terminal cooperative line survey and design platform

TIAN Jun, WANG Haibo, GU Juan

China Information Consulting & Designing Institute Co., Ltd., Nanjing 210019, China

Abstract: In combination with the current trend of survey and design informatization, this paper discusses and summarizes a set of digital solutions for line survey and design based on the multi terminal collaboration of mobile terminals, desktop clients and web terminals. With the help of standardized data specifications and intelligent tools, the delivery quality of survey and design results can be improved, and the efficiency of data sharing and processing can be guaranteed, while the efficiency of survey and design can be greatly improved, at the same time, it also ensures the security of data, and it is also the discussion and practice of the working mode of future communication survey and design.

Keywords: Line route, Survey and design, Integrated, Coordination

○
(上接第27页)

[7]ZHANG L,XIANG T,G,G S G.,Discriminative Null Space for Person Re-identification[C]//Computer Vision and Pattern Recognition Learning 2016:48.

[8] ZHENG, LIANG, SHEN, Y, et al., Scalable Person Re-identification: A Benchmark [C]// IEEE International Conference on Computer Vision. 2015:1116.

[9]JOSE, CIJO, FLEURET F, et al., Scalable Metric Learning via Weighted Approximate Rank Component Analysis [C]//European Conference on Computer Vision. 2016:875-890.

[10]WU L, SHEN C H, HENGEL, V D. PersonNet: Person Re-identification with Deep Convolutional Neural Networks[J]. 2016, 1601: 07255.

[11]VA, OR, RAHUL R, HALOI, et al, Gated Squeeze-and-Excitation

Neural Network Architecture for Human Re-identification [C]//Euro-pean Conference on Computer Vision. 2016:791-808.

[12]LIU H, FENG J, Qi M B, et al. E, -to-En, Comparative Attention Networks for Person Re-identification [J]. IEEE Transactions on Image Processing A Publication of the IEEE Signal Processing Society 2016, 99:1.

作者简介

黄海新 (1973-), 女, 博士, 副教授, 主要研究方向: 机器学习, 人工智能, 智能电网。

陶文博 (1998-), 男, 硕士研究生, 主要研究方向: 计算机视觉与机器学习。

A person re-identification algorithm based on ResNet model and re-ranking strategy

HUANG Haixin, TAO Wenbo

School of Automation and Engineering, Shenyang Ligong University, Shenyang 110159, China

Abstract: Based on the development of deep learning, a person re-identification algorithm is proposed. The process of algorithm is to use ResNet model to extract deep features from the image, then select a set of images under the low similarity constraint, finally, re-ranking the searching results with the related images. The algorithm has three advantages: comparing with other alternative methods, transfer learning decreased the costs of the model training; detecting related images using genetic algorithm with adaptive sparse constraint, it decreased the complexity of computation; the detected results can be used to improve the initial searching result. The experiments on Market1501 data set prove that the algorithm is valuable, and the algorithm has a certain robustness to the changes of background, attitude and perspective.

Keywords: Person re-identification, ResNet, Re-ranking, Transfer learning