

王月明,王辉辉,张宝华.面向监控视频的行人检索系统设计与实现[J].湖南科技大学学报(自然科学版),2022,37(2): 104-109. doi:10.13582/j.cnki.1672-9102.2022.02.015

WANG Y M, WANG H H, ZHANG B H. Design and Implementation of Pedestrian Retrieval System for Surveillance Video [J]. Journal of Hunan University of Science and Technology (Natural Science Edition), 2022,37(2): 104-109. doi:10.13582/j.cnki.1672-9102.2022.02.015

面向监控视频的行人检索系统设计与实现

王月明*,王辉辉,张宝华

(内蒙古科技大学 信息工程学院, 内蒙古 包头 014010)

摘要:为了提高面向监控视频的行人检索的准确率和鲁棒性,提出了一种基于 HSV 颜色直方图与 ResNet50 的两级检索算法.首先利用 HSV 颜色直方图对行人进行初筛,再采用 ResNet50 预训练模型对目标行人进行二级检索.在自制数据库上对不同模型进行行人检索实验,结合准确率、召回率、F-Measure 和检索时间这 4 个指标进行评价,验证了该算法的有效性.最后基于该算法设计了一个检索特定目标行人监控视频的系统,为监控视频的快速检索提供了解决方案.

关键词: 监控视频;行人检索;HSV 颜色直方图;ResNet50

中图分类号:TP391 文献标志码:A 文章编号:1672-9102(2022)02-0104-06

Design and Implementation of Pedestrian Retrieval System for Surveillance Video

WANG Yueming, WANG Huihui, ZHANG Baohua

(School of Information Engineering, Inner Mongolia University of Science and Technology, Baotou 014010, China)

Abstract: In order to improve the accuracy and robustness of pedestrian retrieval for surveillance video, a two-level retrieval algorithm based on HSV color histogram and ResNet50 was proposed. Firstly, HSV color histogram was used to screen pedestrians, and then the pre-training model of ResNet50 was used to secondly retrieve the target. On the self-made database, the pedestrian retrieval experiments were carried out on different models, and carried the evaluation out with four metrics of accuracy, recall, F-measure and retrieval time, which proved that the performance of the method was the best. Finally, a system of retrieving the surveillance video for specific target was designed based on the algorithm, which provided a solution for efficient video retrieval.

Keywords: surveillance video; pedestrian retrieval; HSV color histogram; ResNet50

随着数字多媒体技术和监控设备的发展,近年来,监控摄像头已遍布城市的大街小巷,密集的摄像头带来了海量的监控视频数据,如何从中快速便捷地检索到包含特定目标行人的视频已成为亟待解决的问题.面向监控视频的行人检索在刑侦、安防和社会治安等领域有着广泛的应用^[1].目前,针对行人检索的研究主要集中在行人表征学习方面,例如,Gheissari 等^[2]基于颜色、显著边缘直方图和 Hessian-Affine 兴趣点运算符表征行人;Gray 等^[3]通过融合 8 种颜色空间特征和 19 种纹理特征描述行人;赵瑞等^[4]提出将行人

收稿日期:2020-02-21

基金项目:国家自然科学基金资助项目(61962046)

*通信作者,E-mail:2407763303@qq.com

按人体结构划分为不同区域的特征表示方法;Krizhevsky 等^[5]通过卷积神经网络 AlexNet 在 ImageNet 图像识别竞赛中夺冠^[5],之后深度学习技术被应用到行人检索领域;Ouyang 等^[6]提出联合深度网络 JointDeep 模型用于表征行人;许方洁等^[7]提出基于深度学习与属性学习相结合的行人检索方法。

针对目前面向监控视频的行人检索提出的准确率高、响应速度快和智能化的要求,在两级检索算法的启发下^[8-9],从用户实际需求以及系统快速落地的角度出发,提出了基于 HSV 颜色直方图与 ResNet50 的两级检索算法.在自制行人数据库上,分别对基于 HSV 颜色直方图、SIFT 特征、VGG16 预训练模型、ResNet50 预训练模型和两级检索算法进行了行人检索实验,结合准确率、召回率、F-Measure 和检索时间这 4 个指标进行性能评价.实验结果表明:两级检索算法在检索时间和准确率方面有突出表现,该算法为目标行人快速准确检索提供了手段.最后,基于该算法设计了一个检索特定目标行人监控视频的系统,为解决面向监控视频的行人检索问题提供了新的方案。

1 基于 HSV 颜色直方图与 ResNet50 的两级检索算法

为了适应监控视频中场景复杂、行人姿态多变的实际情况,使检索算法更具鲁棒性的同时尽可能地提高系统检索的准确率和响应时间,提出基于 HSV 颜色直方图与 ResNet50 的两级检索算法,算法思路见图 1.先利用 HSV 颜色直方图和卡方距离度量方法在原始数据库中进行预检索,得到初筛数据库以缩小二级检索规模和提高检索速度;然后利用 Keras 提供的 ResNet50 预训练模型和余弦距离度量方法在初筛数据库进行二级检索,得到最终的行人检索结果,实现提高检索目标的准确率。

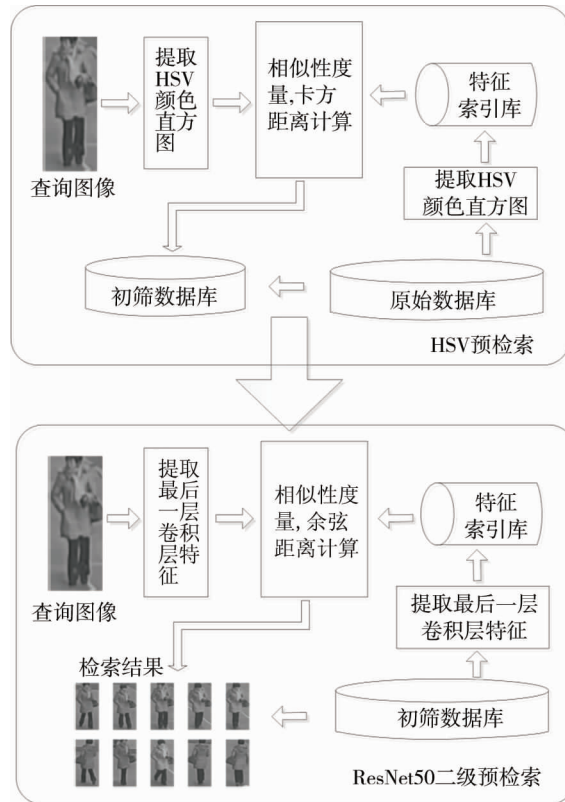


图 1 两级检索算法思路

1.1 HSV 颜色直方图预检索

颜色是最直观显著的一种底层视觉特征^[10],HSV 颜色空间是根据人对色泽、亮度和色调的直观感受而定义的,更符合人类的视觉感知^[11-12].颜色直方图描述的是不同色彩在整幅图像中所占的比列,并不关注每种颜色在图像中的空间分布^[13].行人图像划分如图 2 所示.为了兼顾颜色的空间分布,本算法将图像划分为 5 个部分,分别进行 HSV 直方图的提取。



图2 行人图像划分

预检索过程如下:

步骤 1:量化 HSV 各通道.H 通道量化为 8 个 bin, S 通道量化为 12 个 bin, V 通道量化为 3 个 bin.

步骤 2:将图像转换到 HSV 颜色空间并获取图像尺寸,将其划分为左上、右上、右下、左下和中心 5 部分.

步骤 3:利用掩码思想迭代每个区域,用 0 填充整幅图像,用 255 遮罩掩码区,提取掩码区的颜色直方图后再标准化直方图.

步骤 4:索引化特征向量.提取原始数据库中的图像特征,将图像名和对应的特征向量输出到 CSV 文件中,实现索引化.

步骤 5:相似性度量.计算查询图像特征与索引表中值的卡方距离并保存在字典中,卡方距离越小,图像越相近,返回预检索结果,得到初筛数据库.

卡方距离的计算公式为

$$d_{\text{chi_square}}(X, Y) = \sum_{i=1}^n \frac{(x_i - y_i)^2}{(x_i + y_i)}.$$

式中: $d_{\text{chi_square}}(X, Y)$ 为图像 X 与 Y 之间的卡方距离; x_i, y_i 为图像 X 和 Y 特征向量中的第 i 个值.

1.2 ResNet50 二级检索

深度学习可以自主完成图像从底层视觉特征到高层语义特征的学习^[14], ResNet 提出的恒等映射解决了深度神经网络中模型退化的问题^[15], 其训练结果的准确度高于 VGG 和 GoogLeNet^[16]. Keras 提供的带有预训练权重的 ResNet50 模型, 其学习到的空间上的特征层次结构可以作为视觉世界的通用模型, 具有较强的泛化能力, 多用于预测、特征提取和迁移学习等. 因此, 利用 ResNet50 预训练模型在初筛数据库中进行二级检索, 其检索过程如下:

步骤 1:添加 ResNet50 自定义网络.网络结构不包含 3 个全连接层,指定全局最大值池化.

步骤 2:提取特征.加载图像并将其数字化,对数据进行预处理(归一化、白化等),提取 ResNet50 最后一个卷积层的输出作为图像特征.

步骤 3:索引化特征向量.提取初筛数据库中的图像特征,将图像名和对应的特征向量保存到 HDF5 文件中,实现索引化.

步骤 4:相似性度量.计算查询图像特征与索引表中值的余弦距离,并重新排序距离结果,余弦距离越大,图像越相近,返回前 10 张图像作为最终检索结果.

余弦距离的计算公式为

$$d_{\cos \theta} = \frac{\sum_1^n (a_i b_i)}{\sqrt{\sum_1^n a_i^2} \sqrt{\sum_1^n b_i^2}}.$$

式中: $d_{\cos \theta}$ 为图像 a 与 b 之间的余弦距离; a_i, b_i 为图像 a 和 b 特征向量中的第 i 个值.

2 实验结果分析

2.1 实验环境

为了验证算法在普通硬件条件下即可取得较好的效果,所有实验都是在一台普通电脑上进行,其基本配置参数和相关软件环境如表 1 所示.

表 1 实验环境

类别	名称	类型及参数
硬件	CPU	Intel(R) Core(TM) i5-3230M CPU @ 2.60 GHz
	内存	4.00 GB RAM
软件	操作系统	64 位 Windows 7
	深度学习框架	Tensorflow, Keras

2.2 实验结果

为了验证两级检索算法的有效性,制作一个包含 14 个行人共 100 张行人图像的原始数据库,并在此基础上对不同模型分别进行行人检索实验.图 3 为两级检索算法的实验示例,其中,图 3a 为 HSV 颜色直方图在原始数据库中的预检索结果,返回 15 张图像,得到初筛数据库,其数据量是原始数据库的 0.15 倍;图 3b 为 ResNet50 预训练模型在初筛数据库中的二级检索结果,返回 10 张图像,得到最终检索结果.

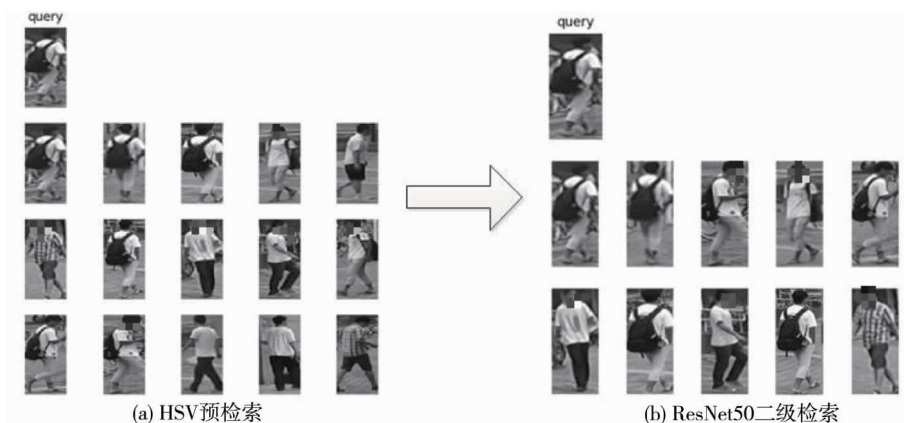


图 3 两级检索实验结果

为了更好地表现两级检索算法的优秀性能,在自制行人数据库上分别对基于 HSV 颜色直方图、SIFT 特征、VGG16 预训练模型、ResNet50 预训练模型和两级检索算法进行行人检索实验,并结合准确率、召回率、F-Measure 和检索时间这 4 个指标进行评价.表 2 为 5 种模型的行人检索实验的各个指标结果.

表 2 算法性能评价指标

评价指标	HSV	SIFT	VGG16	ResNet50	两级检索算法(本文)
准确率	0.70	0.42	0.58	0.66	0.84
召回率	0.58	0.36	0.47	0.55	0.71
F-Measure	0.63	0.38	0.52	0.60	0.77
检索时间/s	1.901	5.137	64.423	68.448	17.095

表 2 的实验结果表明:基于 HSV 颜色直方图与 ResNet50 的两级检索算法对行人检索的准确率为 0.84,召回率为 0.71,F-Measure 为 0.77,检索时间为 17.095 s,与 HSV 颜色直方图、SIFT 特征、VGG16 预训练模型、ResNet50 预训练模型 4 种检索算法的最好评价指标相比,其准确率提升了 0.14,召回率提升了 0.13,F-Measure 提升了 0.14,检索时间与 ResNet50 模型相比缩短了 51.353 s.两级检索算法通过预检索提取行人底层视觉特征,通过二级检索提取行人高层语义特征,双重检索使得其各项指标的结果均达到预期效果,实现了准确率高、检索时间短的目标.

3 检索系统设计与实现

考虑到刑侦领域中嫌疑人活动路线的还原、交通领域中行人轨迹的确定以及安防领域中目标视频的检索等问题,设计了一个可以快速准确地查找包含特定目标行人视频的检索系统,该系统共包含人机交互模块、视频预处理模块、行人检测模块、行人检索模块和获取有效视频模块这5个模块。人机交互模块提供可视化窗口,用于输入查询图像和待检索视频,以及显示检索结果;视频预处理模块完成视频到图像的剪辑并根据时间线索命名图像;行人检测模块执行人行检测和提取任务;行人检索模块通过两级检索算法查找目标行人,并确定行人在监控视频中出现和消失的时刻;获取有效视频模块根据检索结果截取包含目标行人的视频片段。检索系统设计的框图如图4所示。

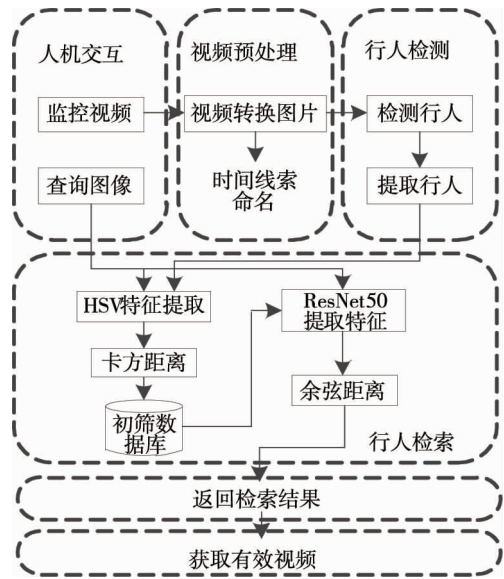


图4 检索系统设计

使用 PyQt5 工具包构建系统的可视化界面,系统可视化界面如图5所示。用户点击“检索图”按钮,系统加载查询图像;点击“视频路径”按钮,选择待检索视频;点击“视频预处理”按钮,运用 OpenCV 计算机视觉库完成视频到图片的剪辑;点击“行人检测”按钮,使用 ImageAI 库调用 RetinaNet 算法完成对行人的检测和提取;依次点击“HSV 预检索”“ResNet50 二级检索”按钮,系统分别执行预检索和二级检索功能;点击“确定有效时间”按钮,用于确定查询目标在监控视频中出现和消失的时刻;点击“保存有效视频”按钮,系统调用 MoviePy 库剪切包含目标行人的视频片段。

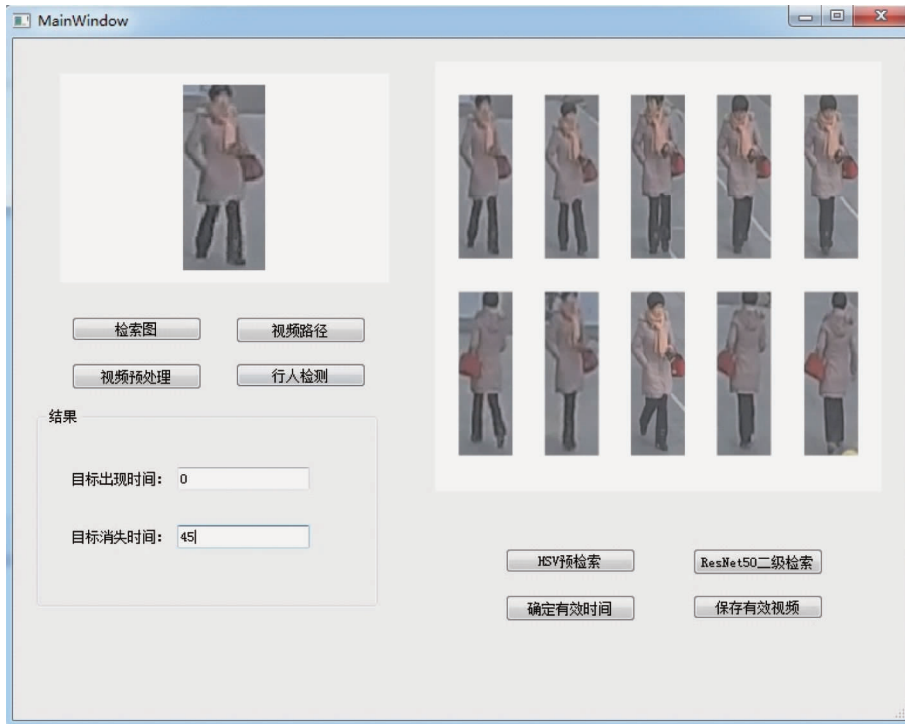


图5 系统可视化界面

4 结论

1) 基于 HSV 颜色直方图和 ResNet50 的两级检索算法,结合不同模型进行一系列行人检索实验,通过对准确率、召回率、F-Measure 和检索时间这4个指标的对比分析,表明两级检索算法有较高的准确率和

较快的检索速度.

2)使用该两级检索算法构建的面向监控视频的行人检索系统,为目标行人监控视频的检索提供了切实可行的解决方案.

参考文献:

- [1] 简佳雁,方志军,高永彬.面向视频监控的距离度量行人再识别[J].传感器与微系统,2019,38(9):129-133.
- [2] Gheissari N, Sebastian T B, Hartley R. Person Reidentification Using Spatiotemporal Appearance[C]//IEEE Computer Society Conference on Computer Vision and Pattern Recognition. IEEE, 2006:1528-1535.
- [3] Gray D, Tao H. Viewpoint Invariant Pedestrian Recognition with an Ensemble of Localized Features[C]//European Conference on Computer Vision. Berlin, Springer, 2008:262-275.
- [4] Zhao R, Ouyang W, Wang X. Unsupervised salience learning for person re-identification[C]// IEEE Conference on Computer Vision and Pattern Recognition. IEEE, 2013:3586-3593.
- [5] Krizhevsky A, Sutskever I, Hinton G. ImageNet Classification with Deep Convolutional Neural Networks[C]// Proceedings of the 25th International Conference on Neural Information Processing Systems. Curran Associates Inc, 2012:1097-1105.
- [6] Ouyang W, Wang X. Joint Deep Learning for Pedestrian Detection[C]// IEEE International Conference on Computer Vision. IEEE, 2013:2056-2063.
- [7] 许方洁,成科扬,张建明.基于深度学习与属性学习相结合的行人再识别[J].数据采集与处理,2018,33(4):751-757.
- [8] 蔡国华.基于内容的医学影像检索技术研究[D].沈阳:东北大学,2012.
- [9] 陈光喜,康春生,黄勇,等.基于深度特征编码的两级车辆检索方法[J].计算机应用研究,2018,35(9):2823-2826.
- [10] 邹煜.基于轮廓提取和颜色直方图的图像检索[D].重庆:西南大学,2011.
- [11] 李国辉,柳伟,曹莉华.一种基于颜色特征的图象检索方法[J].中国图象图形学报,1999(3):72-75.
- [12] 王红雨,尹午荣,汪梁,等.基于 HSV 颜色空间的快速边缘提取算法[J].上海交通大学学报,2019,53(7):765-772.
- [13] Swain M J, Ballard D H. Color indexing[J]. International Journal of Computer Vision, 1991, 7(1):11-32.
- [14] 张顺,龚怡宏,王进军.深度卷积神经网络的发展及其在计算机视觉领域的应用[J].计算机学报,2019,42(3):453-482.
- [15] He K, Zhang X, Ren S, et al. Deep residual learning for image recognition[C]//IEEE Conference on Computer Vision and Pattern Recognition, 2006:770-778.
- [16] Sharma N, Jain V, Mishra A. An Analysis of Convolutional Neural Networks for Image Classification[J]. Procedia Computer Science, 2018, 132:377-384.