

一种基于 CNN 的足迹图像检索与匹配方法

陈 扬^{1,2}, 曾 诚³, 程 成¹, 邹恩岑¹, 顾建伟^{1,2}, 陆 悠¹, 奚雪峰^{1,2}

(1.苏州科技大学电子与信息工程学院,江苏 苏州 215009)

(2.苏州科技大学苏州市虚拟现实智能交互及应用技术重点实验室,江苏 苏州 215009)

(3.昆山市公安局指挥中心,江苏 苏州 215300)

[摘要] 足迹图像作为犯罪现场的重要痕迹物证之一,在破解串并案上有着不可忽视的作用.传统的足迹图像检索与匹配,需要耗费大量的时间与人力,极大地影响了破案进展.卷积神经网络(CNN)在图像识别与检索上表现出很好的效果.面向公安足迹图像比对实战需求,提出了一种基于卷积神经网络的足迹图像检索与匹配方法,对检索结果设置不同检索区,可以满足不同业务需求.初步实验表明该方法的有效性和实用性.

[关键词] 深度学习,卷积神经网络,足迹检索,图像处理

[中图分类号] TP183 [文献标志码] A [文章编号] 1672-1292(2018)03-0039-07

A CNN-based Approach to Footprint Image Retrieval and Matching

Chen Yang^{1,2}, Zeng Cheng³, Cheng Cheng¹, Zou Encen¹, Gu Jianwei^{1,2}, Lu You¹, Xi Xuefeng^{1,2}

(1.School of Electronic and Information Engineering, Suzhou University of Science and Technology, Suzhou 215009, China)

(2.Suzhou Key Laboratory of Virtual Reality and Intelligent Interaction, Suzhou University of Science and Technology, Suzhou 215009, China)

(3.Command Center of Kunshan Public Security Bureau, Suzhou 215300, China)

Abstract: Footprint images, as one of the important evidences of crime scenes, can't be ignored in the cracking of serial cases. Traditional footprint comparison and retrieval require a lot of time and manpower, greatly affecting the progress of the case. Convolutional Neural Network (CNN) has shown good results in image recognition and retrieval. In order to meet the actual needs of public security footprint image retrieval, this paper proposes an approach to footprint image retrieving and matching based on convolutional neural network, and sets different search areas for search results to meet different business requirements. Preliminary experiments show that the proposed approach is effective and practical.

Key words: deep learning, convolutional neural network, footprint searching, image processing

图像检索研究始于 20 世纪 70 年代,起初都是基于文本的图像检索,通过使用文字,对图像的特征进行描述;20 世纪 90 年代开始出现基于内容的图像检索,即对图像颜色和纹理等内容进行分析,并使用一些浅层的分类器如 SVM 等技术进行图像检索,在一定程度上提高了检索的准确率,但仍无法解决语义鸿沟的问题^[1].

随着深度学习的研究与发展,卷积神经网络(convolutional neural network, CNN)近年来在图像检索及识别方面有着不俗的表现^[2]. LeCun 等人成功地将 CNN 用于手写体内容识别上,同时伴随着如 ImageNet、Kaggle 等各类世界图像识别大赛的举办,各种基于 CNN 改变而来的模型如 AlexNet、VGG、GoogLeNet 及 ResNet 等^[3]脱颖而出,表现优异.这些模型凭借着其超高的识别准确率,已在图像检索与识别领域占据了主导地位^[4].由此可见,基于 CNN 技术的图像检索与识别研究具有可靠性.

在科技发展的同时,公安信息化建设也在如火如荼地进行中,将深度学习技术与公安工作相结合,可有效提高公安办公效率,在相关案件侦查方面有很大的应用前景.目前,人脸识别技术以及指纹、掌纹自动识别检索技术在公安的刑事案件侦查中的应用已十分成熟.足迹作为犯罪现场的一个主要痕迹也有着不可忽视的作用^[5].传统的足迹检索与识别往往需要耗费大量的人力资源与时间,且大数据量下的人工

收稿日期:2018-04-18.

基金项目:国家自然科学基金(61750110534, 61728205)、苏州市科技发展计划(重点实验室 SZS201609/产业前瞻性项目 SYG201707).

通讯联系人:奚雪峰,博士,副教授,研究方向:机器学习、数据挖掘. E-mail:xfxi@mail.usts.edu.cn

检索也容易出错^[6]. 足迹自动匹配系统虽然可以省去不少力,但目前的研究大多是通过传统浅层机器学习方法对足迹图像进行处理,过程耗时太长,准确率有待提升. 本文提出了一种基于 CNN 的足迹图像检索与匹配方法,相关实验表明了该方法的有效性和实用性.

1 基于 CNN 的足迹图像检索与匹配方法

CNN 作为一种前馈式神经网络,是基于全连接神经网络的一种改进版,大大减少了需要计算的参数^[7]. 其网络结构主要由卷积层、池化层、全连接层及一些激活函数构成. 在卷积层中,通过不同的卷积核在原图像窗口上滑动进行卷积操作,提取特征,得到不同的特征图,将卷积后的结果做非线性变化,得到输出作为下一层的输入^[8]. 经过卷积层后,若得到的特征图比较大,可通过池化操作来进行维度的下降,可选择最大值池化操作和均值池化操作,经过池化层后,输出的图像深度不变,仍为特征图的个数. 池化层可在一定程度上防止模型的过拟合,方便池化层优化后将所有的神经元进行权重连接,与传统的神经网络连接方式一样,最后得到输出^[9]. 图 1 所示为一个一般的 CNN 架构图.

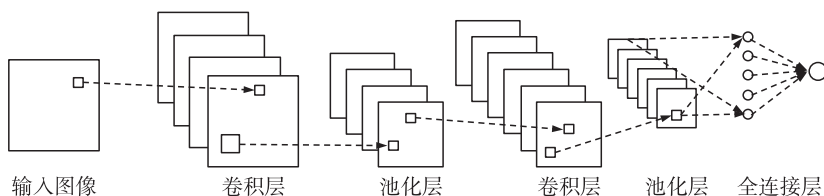


图 1 CNN 结构图

Fig. 1 The structure of CNN

本文采用 VGG16 模型,使用 Keras 深度学习框架构建了一个足迹图像检索与匹配方法,结合足迹图像数据集开展模型训练及迁移学习. 方法主要包括 3 个步骤:数据预处理;通过 VGG16 模型提取足迹数据集特征,建立特征向量索引;输入待检索图片,设置不同的数量检索候选区域,实现足迹图像检索.

足迹图像检索与匹配方法总体框架如图 2 所示.

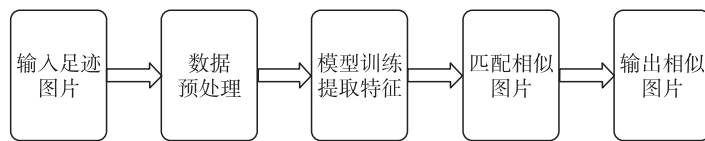


图 2 总体框架图

Fig. 2 Overall framework diagram

1.1 数据预处理

为方便特征参数的提取,将大小约为 $2\,000 \times 1\,000$ 像素的输入足迹图像调整为 224×224 像素,并对数据进行归一化操作. 针对本数据集,采用逐样本均值消减方法进行归一化操作,具体操作方法为将输入图片转化为矩阵后,在每个样本上减去数据的统计平均值(即分别计算各图像样本的均值,然后各样本减去其对应的均值). 归一化操作可移除图像的平均亮度值,从而减少图像背景对实验的干扰. 由于 Sobel 算子能较好地保留图像的细节信息及内容的完整性,故本文采用 Sobel 算子进行对比实验分析. 预处理图片如图 3 所示.

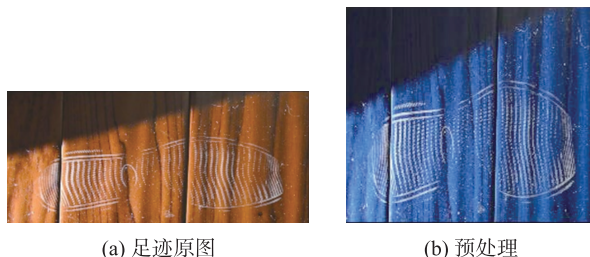


图 3 图片的预处理

Fig. 3 Preprocessed image

1.2 模型训练特征提取

基于深度学习的图像检索中,特征向量的提取是关键一步,本文使用 VGG16 模型提取足迹数据集特征. 首先,使用 ImageNet 上的海量数据预训练 VGG16 模型,借助其预训练好的参数,将除全连接层以外的其余网络结构保留到本文的模型中. 将足迹数据库中每张图像经该网络模型提取特征,生成具有 4 096 维的向量,提取所有图像特征后,建立针对足迹图像数据库的特征向量索引库. VGG16 模型能较好地保留图像的颜色、纹理及轮廓方面的特征^[10].

提取特征时,每一层的特征提取都是经过卷积、池化及激活函数的操作后得到下一层的输出^[11]. 输入图像为矩阵 X ,卷积核参数为 W ,偏置 bias 为 b ,采用 ReLU 激活函数,经式(1)计算后得到输出 $h_{i,j}^k$,再进行池化操作后得到下一个卷积层的输入:

$$h_{i,j} = \text{Relu}((W^k * X)_{ij} + b), \quad (1)$$

式中, i, j 表示矩阵中的行与列.

设 i 为一个卷积层, f^i 为第 i 层卷积核的大小, p^i 为第 i 层的 padding 数量, s^i 为第 i 层的步长大小, n_c^i 为第 i 层卷积核的个数, $n_H^{i-1} \times n_W^{i-1} \times n_C^{i-1}$ 为第 $i-1$ 层输入图像矩阵的长、宽及通道数,则第 i 层的图像输出大小为:

$$n_H^i = \frac{n_H^{i-1} + 2 * p^i - f^i}{s^i} + 1, \quad (2)$$

$$n_W^i = \frac{n_W^{i-1} + 2 * p^i - f^i}{s^i} + 1. \quad (3)$$

输出图像的通道数即为卷积核的个数. 因此,第 i 层图像矩阵的输出为 $n_H^i \times n_W^i \times n_C^i$.

最终根据上述公式可求得最后的特征向量维度为 4 096.

1.3 图像的相似度学习与度量

计算样本之间的距离是一个研究热点,目前被广泛运用的距离度量有欧式距离、余弦相似度距离及汉明距离等^[12]. 基于深度学习的图像检索中,使用 CNN 提取图像特征,建立特征向量,并基于图像的特征向量来表示对应的图像,通过计算图像特征向量之间的距离来判断图像间的相似性^[13]. 本文采用欧式距离作为相似图片特征向量的计算方法:

$$d(x, y) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \dots + (x_m - y_m)^2} = \sqrt{\sum_{i=1}^m (x_i - y_i)^2}, \quad (4)$$

式中, m 为图像的特征向量维度. 通过比较待检索图像的特征向量与图像检索库中的各个特征向量之间的距离^[14],即图像之间每个维度上的距离,当距离越小时,则证明图像之间的相似度越高.

$$D_i = q - T_i. \quad (5)$$

假设图片库中一共有 n 张图片, q 表示待检索图像的特征向量, T 为图片库中第 i 张足迹图的特征向量, D 表示特征向量间的距离差,其中, $i \in [1, n]$; 当 D 的值越小时,说明两张图片的相似性越高. 通过排序算法比较 D 值,查找并输出最为相似的 N 张图片. 本文利用预训练好的 VGG16 模型对图像库的每一张图片进行特征提取并建立特征向量索引库,将待检索图片通过模型提取特征向量^[15],通过上述距离算法进行比较,计算出相似性的匹配程度. 具体的检索流程图如图 4 所示.

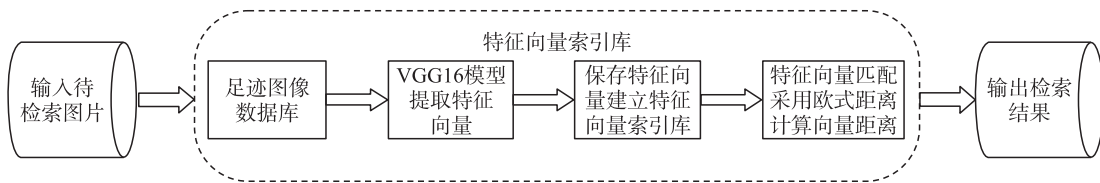


图 4 检索流程图

Fig. 4 The flow chart of image retrieval

2 实验分析与讨论

2.1 实验数据

本课题的数据集由合作方昆山市公安局提供. 针对犯罪现场的足迹图像的实际情况,首先通过数据清洗,人工挑选出质量较高的便于研究的足迹图像,再通过旋转角度、翻转等方式扩充数据集,共计 600 余张. 同时根据所提供的足迹图像,分析影响检索精度的若干问题,设置对比实验探索研究.

2.2 实验评估方法

查全率和查准率是评判准确率的两个重要准则,同样也可以作为图像检索的评判标准. 本文定义查准率为检索结果中的相似图像数量与检索结果的全部图像数量之比:

$$\gamma = N_i / N_l, \quad (6)$$

式中, N_i 为检索结果中相似的图像数量; N_l 为检索结果输出图像总量。

定义查全率为检索结果中相似图像数量与数据库中相似图像数量总量之比:

$$\varphi = N_i / N_s, \quad (7)$$

式中, N_i 为检索结果中相似的图像数量; N_s 为数据库中相似图像数量总量。

通常查准率和查全率很难同时达到最优。一般情况下, 查准率高时, 查全率往往会较低; 查全率较高时, 查准率会较低。查全率体现了检索结果的全面性, 查准率则侧重表达的准确性。

本课题借助于查全率和查全率的思想, 提出命中准确率。若 n 表示一张输入图像, 且关于 n 的相似图像总量集合为 M , 检索输出图像总量为 S , 其中输出相似图像为 T , 则命中准确率为:

$$\rho = (T/S) \times 100\%. \quad (8)$$

2.3 实验分析与讨论

2.3.1 粗检

为便于量取犯罪现场足迹尺寸, 原足迹图中每个脚印旁都带有比例尺。先建立第 1 组数据库, 将相似脚印拆分到两个数据集中, 一个数据集作为待检索数据集, 另一个数据集作为检索图片库。首先将检索图片库中的所有图片作为输入, 通过 VGG16 模型训练提取特征, 建立图片特征向量索引, 然后在待检索数据集中挑选图片进行输入, 输出最为相似的图片结果。部分检索结果如图 5。

从检索结果可以发现, 部分较明显的足迹图片能够较准确地检索匹配到对应的相似图片, 而不清晰的足迹图片在检索匹配时会产生误差。分析误差产生原因, 可能是足迹图中的比例尺对检索结果产生影响。

2.3.2 去除比例尺后的检索实验

在第 2 组实验中, 截取原足迹图中的脚印部分, 去除比例尺, 重复步骤, 通过 VGG16 模型训练提取另一组特征向量索引, 并输入图片进行实验。部分检索结果如图 6 所示。

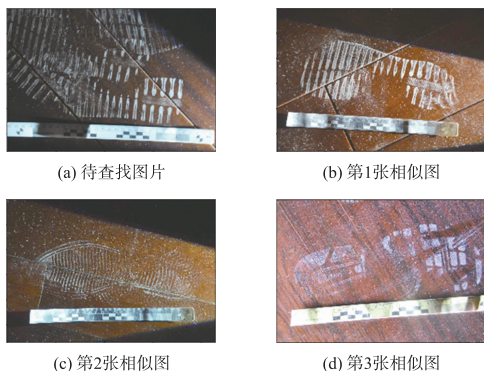


图 5 实验 1 的输出

Fig. 5 The output of the first experiment

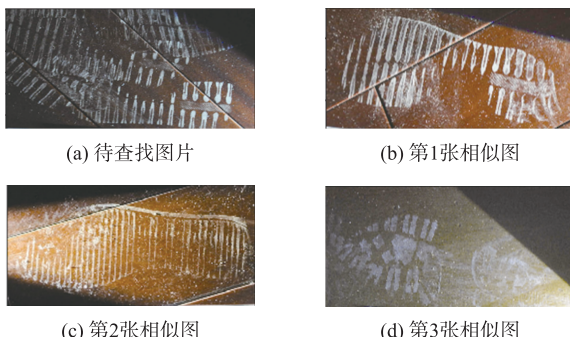


图 6 实验 2 的输出

Fig. 6 The output of the second experiment

第 2 组实验中, 较明显的足迹图仍能较准确地检索匹配到相似图; 实验 1 中未能准确检索到的部分足迹图在本次试验中匹配到了相似图, 而之前实验 1 中部分能够检索到相似图片的足迹却未能匹配到相似图片。

2.3.3 背景对检索的影响

第 3 组实验研究背景是否会在足迹图的相似性检索上产生干扰。本文选择使用 PS 技术和 Sobel 算子对足迹图像进行处理, 并重复检索流程中的步骤进行实验。检索输出如图 7、图 8 所示。

实验发现, 经 PS 处理后的图片检索效果较好, 准确率较高; 而 Sobel 算子处理的图片检索效果较差。

2.3.4 填充操作对检索的影响

由于本实验数据集中脚印图片几乎都是矩形, 在缩放成 224×224 时由于挤压图片易导致脚印变形。为保证脚印的形状, 对足迹图片进行填充操作, 将原图脚印补成接近正方形的比例, 使其通过图片缩放后保证脚印的正常形状, 然后再输入进入 VGG16 模型中训练提取特征, 检索结果如图 9 所示。

同时, 为了制造更多的数据, 并检验翻转操作是否对检索结果有影响, 将足迹图片进行旋转或者水平映射操作。部分检索结果如图 10 所示。由图 10 可知, 待查找足迹图与第 1 张和第 2 张的相似图为一对

足迹图像,而第 3 张输出结果则与输入图像有一定的误差.可见,对于翻转的足迹图像,仍能较为准确地识别出结果.

由以上实验可见,进行填充操作后的足迹图片在检索性能上有可见的效果提升.

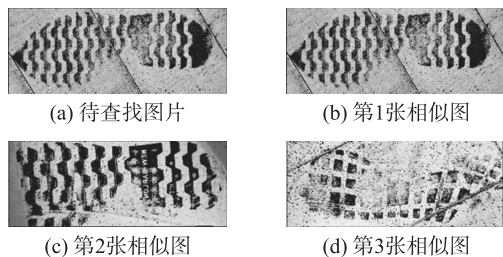


图 7 实验 3(1) PS 处理的足迹图

Fig. 7 Footprint of PS processed

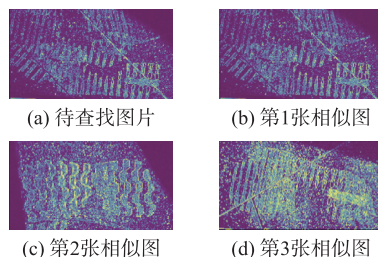


图 8 实验 3(2) Sobel 算子处理的足迹图

Fig. 8 Footprint processed by Sobel operator

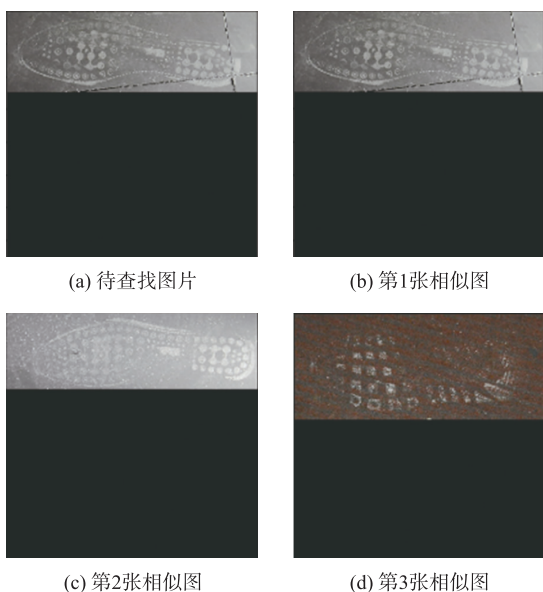


图 9 填充操作后的检索结果

Fig. 9 Retrieval results of padding operation

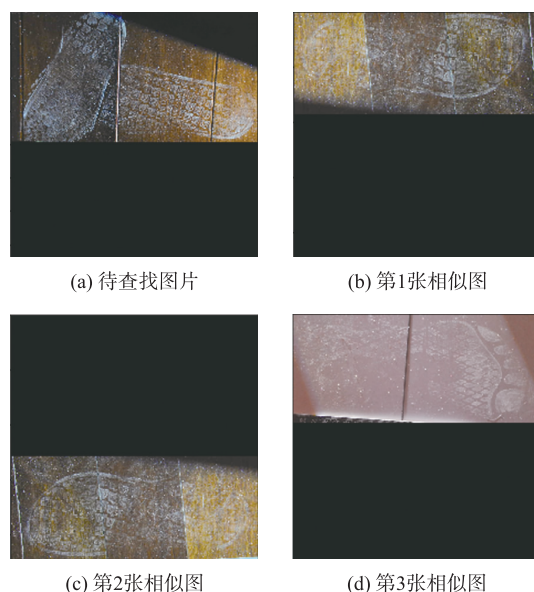


图 10 图片翻转后的检索结果

Fig. 10 Retrieval results of flipped image

2.3.5 设备不同检索候选区对检索结果的影响

在 Padding 操作的基础上,设置不同的检索结果候选区,即设置检索结果输出图片个数,进行进一步实验.

如图 11~图 13 所示,随着检索候选区图像数量的增大,越来越多的相似图片出现在检索结果中.

分析上述 5 组实验,并多次求平均值,同时扩大检索候选区,进行输出 10 张检索相似图的实验,可得当输出 3 张、5 张、10 张相似图时,检索匹配的准确率如表 1 所示.由表 1 可见,随着相似图输出数量的增大,检索匹配的精度也随之提高.同时,针对不同的数据量,可以合理设置不同数量的相似图输出,这有助于提高最终的检索质量与精度.

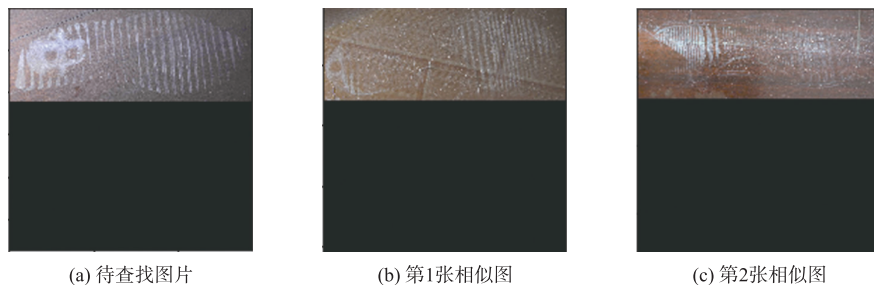


图 11 输出 2 张相似足迹图

Fig. 11 The output of 2 similar footprints

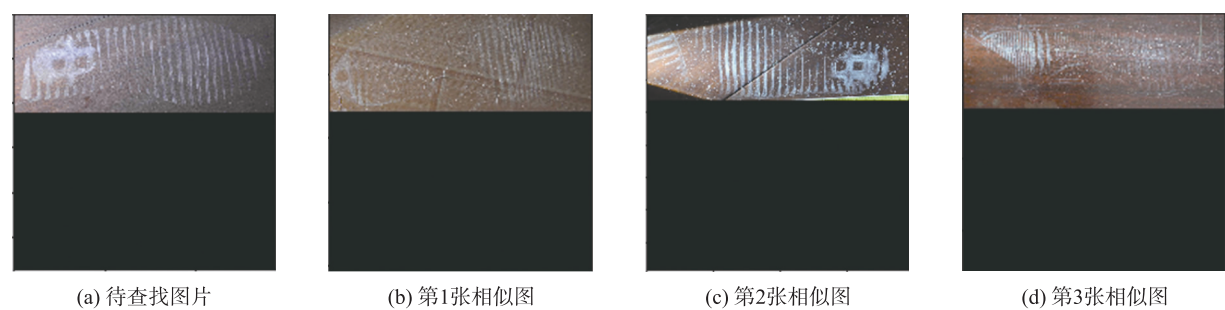


图 12 输出 3 张相似足迹图

Fig. 12 The output of 3 similar footprints

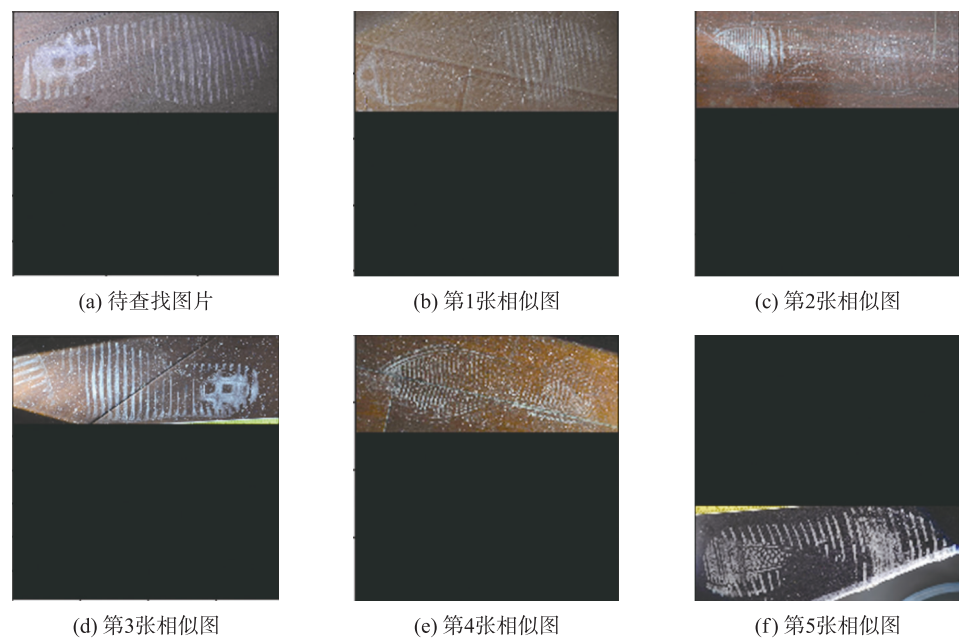


图 13 输出 5 张相似足迹图

Fig. 13 The output of 5 similar footprints

表 1 各组对比实验评估准确率

Table 1 The accuracy of comparison experimens in each group									
实验	输出 3 张相似图			输出 5 张相似图			输出 10 张相似图		
	$\gamma/\%$	$\varphi/\%$	$\rho/\%$	$\gamma/\%$	$\varphi/\%$	$\rho/\%$	$\gamma/\%$	$\varphi/\%$	$\rho/\%$
实验 1	50	20	50	54	28	54	62	34	62
实验 2	48	12	48	52	26	52	55	38	55
实验 3(1)	60	25	60	65	35	65	73	45	73
实验 3(2)	20	13	20	26	16	26	32	20	32
实验 4	50	22	50	60	28	60	75	35	75
实验 5	53	20	53	75	30	75	85	58	85

3 结语

本文所提出的一种基于 CNN 的足迹图像检索匹配方法,使用预训练好的 VGG16 模型提取特征,灵活设置相似图候选区数量,将海量图片匹配问题减少为仅需从若干限定数量的候选图片中检索,能够高效查询得到相似足迹,从而为进一步人工判断提供了可行性. 从实验结果来看,该方法大幅度减小了人工检索中所消耗的精力,具有实际应用价值.

本文的研究工作尚属初步,存在较多不足,如数据量的规模还不够大,这可能影响到基于 CNN 的深度学习模型预训练性能,未来将结合公安实战场景,重点关注样本数据集建设. 同时,图片背景处理及足迹增强算法对检索性能有着较为关键的影响,这也将是下一步研究的重点.

[参考文献] (References)

- [1] 李钊,卢苇,邢薇薇,等. CNN 视觉特征的图像检索[J]. 北京邮电大学学报,2015,38(增刊1):103-106.
LI Z,LU W,XING W W,et al. Image retrieval based on CNN visual features[J]. Journal of Beijing university of posts and telecommunications,2015,38(Suppl. 1):103-106.(in Chinese)
- [2] 胡二雷,冯瑞. 基于深度学习的图像检索系统[J]. 计算机系统应用,2017,26(3):8-19.
HU E L,FENG R. Image retrieval system based on deep learning[J]. Computer systems and applications,2017,26(3):8-19.(in Chinese)
- [3] 朱煜,赵江坤,王逸宁,等. 基于深度学习的人体行为识别算法综述[J]. 自动化学报,2016,42(6):848-857.
ZHU Y,ZHAO J K,WANG Y N,et al. A review of human action recognition based on deep learning[J]. Acta automatica sinica,2016,42(6):848-857.(in Chinese)
- [4] 周晔,张军平. 基于多尺度深度学习的商品图像检索[J]. 计算机研究与发展,2017,54(8):1824-1832.
ZHOU Y,ZHANG J P. Multi-scale deep learning for product image search[J]. Journal of computer research and development,2017,54(8):1824-1832.(in Chinese)
- [5] 韩科. 应用模糊 BP 神经网络对足迹图像识别方法的研究[D]. 沈阳:东北大学,2006.
HAN K. Research on an automatic toe shape recognition method based on fuzzy comprehensive evaluation model and BP NN clustering[D]. Shenyang:Northeastern university,2006.(in Chinese)
- [6] 牛瑞娟. 足迹图像的特征提取与分类[D]. 青岛:山东科技大学,2007.
NIU R J. Study of the feature extraction and classification of footprint images[D]. Qingdao:Shandong university of science and technology,2007.(in Chinese)
- [7] 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.
ZHOU F Y,JIN L P,DONG J. Review of convolutional neural network[J]. Chinese journal of computers,2017,40(6):1229-1251.(in Chinese)
- [8] FU R,LI B,GAO Y,et al. Content-based image retrieval based on CNN and SVM[C]//IEEE International Conference on Computer and Communications. Chengdu,China,2016.
- [9] SEDDATI O,DUPONT S,MAHMOUDI S,et al. Towards good practices for image retrieval based on CNN features[C]//IEEE International Conference on Computer Vision Workshop. Venice,Italy,2017.
- [10] 陈宏彩,程煜,张常有. 卷积神经网络在车辆目标快速检测中的应用[J]. 软件学报,2017,28(增刊1):107-114.
CHEN H C,CHENG Y,ZHANG C Y. Convolutional neural network applied on fast vehicle objects detection[J]. Journal of software,2017,28(Suppl. 1):107-114.(in Chinese)
- [11] SCHROFF F,KALENICHENKO D,PHILBIN J. Facenet:a unified embedding for face recognition and clustering[C]//IEEE Conference on Computer Vision and Pattern Recognition. Boston,USA,2015.
- [12] WANG J Y,QIAN Y,YE Q Q,et al. Image retrieval method based on metric learning for convolutional neural network[C]//International Seminar on Advances in Materials Science and Engineering. Singapore,2017.
- [13] RADENOVIC F,TOLIAS G,CHUM O. Fine-tuning CNN image retrieval with no human annotation[DB/OL].(2018-06-08). <https://arxiv.org/abs/1711.02512v1>.
- [14] 金连文,钟卓耀,杨钊,等. 深度学习在手写汉字识别中的应用综述[J]. 自动化学报,2016,42(8):1125-1141.
JIN L W,ZHONG Z Y,YANG Z,et al. Applications of deep learning for handwritten Chinese character recognition;a review[J]. Acta automatica sinica,2016,42(8):1125-1141.(in Chinese)
- [15] BABENKO A,SLESAREV A,CHIGORIN A,et al. Neural codes for image retrieval[C]//European conference on computer vision. Zurich,Switzerland,2014.

[责任编辑:严海琳]