

基于特征自动提取的足迹图像聚类方法

陈扬^{1,2} 王金亮³ 夏炜³ 杨颢^{1,2} 朱润³ 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 江苏 昆山 215300

(mango@post.usts.edu.cn)

摘要 足迹图像是公安在串并案的侦破过程中最为重要的线索,且每年各处公安都会收集很多犯罪现场的足迹,如何自动化地整理和归类这些足迹图像成为当前公安信息化的一个难点。面向公安实战需求,文中结合卷积神经网络和 DBSCAN 算法,提出了一种对足迹图像聚类的方法。首先,对足迹图像进行预处理以便满足模型训练要求;接着,通过模型预训练改进了 Resnet50 和 Densenet121 两类卷积神经网络模型结构,提取足迹图像特征并建立特征向量库;随后,基于 DBSCAN 聚类算法,利用上述特征向量库实现对足迹图像的整理归类。实验结果表明,该方法具有良好的实用性和有效性。

关键词: 足迹图像;卷积神经网络;图像聚类;特征自动抽取

中图分类号 TP181

Footprint Image Clustering Method Based on Automatic Feature Extraction

CHEN Yang^{1,2}, WANG Jin-liang³, XIA Wei³, YANG Hao^{1,2}, ZHU Run³ and XI Xue-feng^{1,2}

1 School of Electronic & Information Engineering, Suzhou University of Science and Technology, Suzhou, Jiangsu 215009, China

2 Suzhou Virtual Reality Intelligent Interaction and Application Technology Key Laboratory, Suzhou, Jiangsu 215009, China

3 Kunshan Public Security Bureau, Kunshan, Jiangsu 215300, China

Abstract Footprint images are the most important clues in the detection process of public security cases. Every year, public security agencies collect many crime scene footprints. How to automatically organize and categorize these footprint images has become a difficulty for public security informatization. To meet the actual needs of public security, this paper combines a convolutional neural network and DBSCAN algorithm to propose a method for clustering footprint images. First, the footprint image is preprocessed to meet the model training requirements. Then, through model pre-training, the two types of Resnet50 and Densenet121 convolutional neural network model structures are improved to extract footprint image features and establish a feature vector library. Based on DBSCAN Similar algorithms, we use the above feature vector library to organize and classify footprint images. Experimental results show that the method has good practicability and effectiveness.

Keywords Footprint image, Convolutional neural network, Image clustering, Automatic feature extraction

1 引言

随着信息化时代社会建设发展,“互联网+”和“智能+”等计划的深入推进,各个领域都在悄然发生改变,朝着愈发智能的方向前进。近年来,得益于各类算法研究的突破、硬件计算能力的提升以及互联网下大数据的处理与积累,各项实验室中的技术逐渐走向产业化落地。普通人眼中看起来传统严肃的公安领域也在迫切地进行信息化、智能化建设,越来越多的技术正与公安系统相结合,提高警务工作的科技含量。

图像技术是公安在案件侦破过程中最频繁接触到的内容。对于公安而言,图像中有着巨大的信息且这些信息包含重要价值,如人脸、指纹、足迹图像等,这些图像信息为公安提

供了非常重要的线索,也常常是案件侦破的关键所在。除了已经被成熟运用在公安日常案件侦破中的人脸识别与指纹识别之外,足迹作为犯罪现场的重要物证之一,也有着不可忽视的作用。在绝大多数犯罪现场,可以将足迹作为案件的重要物证,这也为警方在串并案的侦破中提供了重要的线索。因此,对于足迹图像的研究有着极大的社会价值^[1]。

与此同时,每年各处公安都会收集很多犯罪现场的足迹,这些足迹信息多以图像形式保存,且这些犯罪现场的足迹图像库愈发壮大。但这些足迹图像中有价值的信息往往难以直接获取,并且在大量数据中快速得到足迹图片的有效信息一直是个难点,如何自动化地整理和归类这些足迹图像成为当前公安的迫切需求。

基金项目:国家自然科学基金(61876217);江苏省“六大人才高峰”高层次人才项目资助(XYDXX-086);苏州科技大学校创新科研基金(SKSJ18_010)

This work was supported by the National Natural Science Foundation of China(61876217), Innovative Team of Jiangsu Province(XYDXX-086) and Postgraduate Research & Practice Innovation Program of Suzhou University of Science and Technology(SKSJ18_010).

通信作者:奚雪峰(xfxi@usts.edu.cn)

以往受限于现场足迹图像的提取设备和拍照设备的性能,多数关于足迹图像的研究集中于如何提升现场采集到的足迹图像清晰度^[2]。不过随着硬件设备的性能提升,获得清晰的高分辨率的犯罪现场的足迹图像已不再是难点^[3]。近些年针对足迹图像方面的研究逐渐转向足迹图像识别与检索^[4],但关于足迹图像的自动聚类的研究还较少。传统足迹图像的存档归类大多依靠人工来实现,需要人工对大量图像数据进行标注,并按照足迹图像的相似性有序进行归档,耗时耗力,且在人工情况下容易出错。

针对上述问题,本文结合足迹图像的特点,使用卷积神经网络的方法提取足迹图像的特征,并借助 DBSCAN(Density-Based Spatial Clustering of Applications with Noise)聚类算法实现足迹图像的自动聚类。本文主要的贡献度在于:1)利用卷积神经网络自动提取足迹图像特征,并借助 DBSCAB 方法实现足迹图像自动聚类,为公安串并案提供信息化手段。据我们所知,这在公安足迹图像研究应用中尚属首次。2)探究了不同卷积网络模型对足迹图像特征提取的影响,并结合多种指标更为全面地评估聚类效果。

2 相关工作

2.1 卷积神经网络简介

卷积神经网络是一种前馈神经网络,它的人工神经元可以响应一部分覆盖范围内的周围单元,对于大型图像处理有出色表现。20世纪60年代,Hubel等通过对猫视觉皮层细胞的研究,提出了感受野这个概念。到80年代,Fukushima在感受野概念的基础之上提出了神经认知机的概念,其可以看作卷积神经网络的第一个实现网络^[5]。在这个基础上,Lecun等^[6]在1988年提出了真正意义上的当代卷积神经网络的雏形。直到现在,各种卷积神经网络模型层出不穷,这些模型也在图像识别领域大放异彩。卷积神经网络主要由卷积层、池化层以及全连接层构成。卷积层通过滑动窗口机制,使用多个卷积核对原图像提取特征,可使得原始信号的某些特征增强,并且降低噪声。池化层主要对经过卷积操作后得到的特征图像进行下采样,可以减少数据处理量,同时保留有用信息。全连接层则对提取的特征进行非线性组合以得到输出。

卷积神经网络通过使用大量样本数据对自身非线性模型结构中的参数进行训练,学习样本数据中的特征,再采用梯度下降法最小化损失函数对网络中的权重参数逐层反向调节,通过频繁的迭代训练提高网络的精度^[7]。相比于一般神经网络而言,卷积神经网络能够较好地适应图像的结构,同时进行特征提取和分类,使得特征提取有助于特征分类,其中权值共享可以减少网络的训练参数,使得神经网络结构变得简单,适应性更强。

2.2 卷积神经网络模型

随着卷积神经网络的飞速发展,以及相关硬件设备计算能力的提升,涌现出一批优秀的深度卷积神经网络模型,如 LeNet-5 model, AlexNet model, VGGNet model, GoogleNet model, ResNet model and DenseNet 等。本文选取了 ResNet 和 DenseNet 两种典型的网络用于实验。

研究者发现较深的神经网络可以提取更丰富的图像特征,但与此同时,梯度消失和爆炸问题变得更加突出,该问题将使网络无法训练。He等^[8]提出的 ResNet 在一定程度上解决了这个问题,且该模型在 ILSVRC2015 中获得了第一名。

ResNet 中的恒等快捷连接(identity shortcut connection)结构使前层的图像信息可以直接传输到更深的层,从而保护了图像特征信息的完整性。图1显示了 ResNet 中使用的典型残差块。

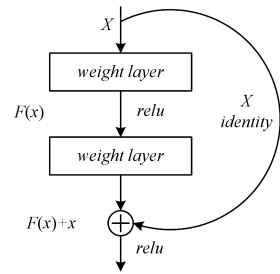


图1 ResNet 残差结构图

Fig. 1 Residual structure diagram of ResNet

DenseNet^[9]的基本思想类似于 ResNet,它们都在上一层和下一层之间建立连接。与 ResNet 相比,DenseNet^[22]提出了密集连接,所有层都可以相互连接。具体来说,每一层都接收其所有先前的层作为额外输入。图2给出了 DenseNet 的密集连接。DenseNet 的另一个主要特征是它通过通道上要素的连接实现了要素重用。这些特性使 DenseNet 的参数和计算成本低于 ResNet,但性能却优于 ResNet。图2给出了 DenseNet 网络结构。本文选择使用 DenseNet-121 进行实验。

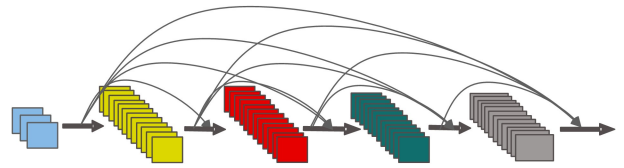


图2 DenseNet 网络结构图

Fig. 2 Network structure diagram of DenseNet

2.3 DBSCAN 聚类算法

聚类算法^[10]是机器学习中涉及对数据进行分组的一种算法。在给定的数据集中,可以通过聚类算法将其分成一些不同的组。在理论上,相同的组的数据之间有相同的属性或者特征,不同组数据之间的属性或者特征相差比较大。聚类算法是一种非监督学习算法,并且作为一种常用的数据分析算法在很多领域上得到了应用。常用的比较出名的聚类算法有 k-means 和 Hierarchical Agglomerative Clustering,它们需要提前指定簇的数量,而在实际的应用中无法知道具体数据集中的类别数量,因此,k-means 和 Hierarchical Agglomerative Clustering 聚类算法不适用,我们需要使用基于密度或基于图的聚类算法,这类算法不仅可以聚类数据点,还可以根据数据密度确定聚类数量,符合实际应用要求。本文具体采用了 Density-based spatial clustering of applications with noise (DBSCAN)聚类算法^[11]。

DBSCAN 聚类算法使用一组关于“邻域”的参数来描述样本分布的紧密程度。针对具体的算法描述,先定义以下几个相关的概念。

邻域:对于任意样本 i 和给定距离 e ,样本 i 的 e 邻域是指所有与样本 i 距离不大于 e 的样本集合。

核心对象:若样本 i 的 e 邻域中至少包含 MinPts 个样本,则 i 是一个核心对象。

密度直达:若样本 j 在样本 i 的 e 邻域中,且 i 是核心对象,则称样本 j 由样本 i 密度直达。

密度可达:对于样本 i 和样本 j ,如果存在样本序列 $p_1, p_2, \dots, p_n (p_1=i, p_n=j)$, 并且 p_m 由 p_{m-1} 密度直达, 则称样本 i 与样本 j 密度可达。

密度相连:对于样本 i 和样本 j ,若存在样本 k 使得 i 与 j 均由 k 密度可达, 则称 i 与 j 密度相连。

对于给定的邻域距离 e 和邻域最小样本个数 MinPts , DBSCAN 算法的具体步骤如下:

(1) 遍历所有样本, 找出所有满足邻域距离 e 的核心对象的集合;

(2) 任意选择一个核心对象, 找出其所有密度可达的样本并生成聚类簇;

(3) 从剩余的核心对象中移除步骤(2)中找到的密度可达的样本;

(4) 从更新后的核心对象集合重复执行步骤(2)一步骤(3), 直到核心对象都被遍历或移除。

3 本文方法

结合一些深度神经网络模型在人脸识别上的杰出表现, 本文基于 ResNet50, DenseNet121 网络开展足迹图像的特征提取, 并采用 DBSCAN 聚类算法对足迹图像进行聚类。与此同时, 在合作方昆山市公安局提供的足迹数据集上结合迁移学习对深度卷积神经网络模型进行了训练。

本文的方法主要包含如下 4 个部分: 1) 数据清洗及预处理; 2) 微调模型结构并训练模型; 3) 使用深度卷积神经网络提取足迹图像特征并建立特征向量库; 4) 使用 DBSCAN 算法对提取好的足迹图像特征进行聚类计算。具体的流程如图 3 所示。

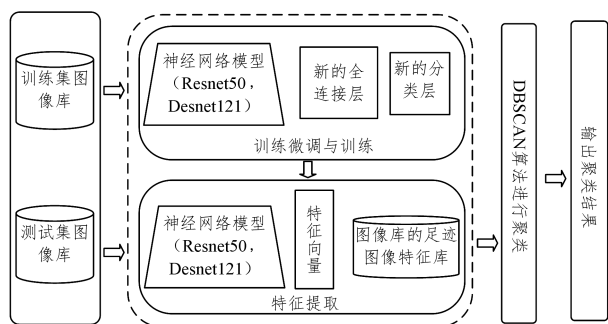


图 3 总体框架图

Fig. 3 Overall framework

3.1 预处理

数据预处理在计算机视觉中起着重要作用。为了更好的实验结果, 结合本项目的实验数据集, 需要对足迹图像进行预处理。首先, 为了便于提取图像的特征参数, 将大约 1000×2000 像素的输入覆盖区调整为 128×256 。然后填充图像并将图像随机裁剪为 128×256 。与此同时, 数据将被归一化, 针对实验数据集, 采用逐样本均值减少方法进行归一化。具体的操作方法是: 将输入图像转换为矩阵, 从每个样本中减去数据的统计平均值, 即计算每个图像样本的平均值, 然后以平均像素为中心, 减去每个样本的对应平均值。对于图像数据, 此归一化方法可以去除图像的平均亮度值, 从而减少图像的背景效果对实验的干扰。最后, 图像被转换为张量并归一化为 0-1。

3.2 模型的结构微调与训练

本文中, 对数据预处理修改了两个模型的全连接层以及

全连接层的后续部分。按照 Linear, Batch normalization, ReLU, Linear 的顺序定义模型新的全连接层和分类层。在 Resnet50 和 Densenet121 中, 我们将全连接层修改为 512 维, 并添加了新的分类层。对于增加的层, 使用 Kaiming Normal 初始化参数, 其他参数使用在 ImageNet 上预训练的参数, 再通过足迹图像数据集对模型进行再一次训练。

我们还使用了自适应平均池化 (Adaptive Average Pool), 因为足迹图像的高度大于宽度, 我们需要指定池化内核, 而自适应平均池化更易于实现。在训练过程中, 我们设置了 40 个 $epoch$, 并通过 ImageNet 上预先训练的参数使用 SGD 进行反向传播, 如下式所示:

$$\theta = \theta - \eta \cdot \nabla_{\theta} J(\theta; x^i, y^i) \quad (1)$$

其中, J 是 SGD 需要优化的目标函数, x^i 和 y^i 表示训练样本和训练标签, θ 表示模型的参数如权重、偏参数。

实验中使用了动态学习率, 具体公式如下:

$$lr = lr_0 \times \lambda^{\frac{epoch}{step_size}} \quad (2)$$

其中, lr 是目前的学习率, lr_0 是初始学习率, λ 是衰减系数, $step_size$ 是学习率变化步数, $epoch$ 是指当前训练的轮数。

与此同时, 在模型的训练中, 样本数据训练的批处理 (batch size) 大小为 32, 初始学习率为 0.1 且学习率每隔 40 次迭代变化一次。最终模型在训练集上的准确率可达 98%。

3.3 特征提取

在基于深度学习的图像处理中, 图像特征向量的提取是至关重要的一步。通过对当前各种主流卷积神经网络模型的分析, 结合实际情况, 本文使用了两种不同的模型进行实验。在足迹图像训练集上训练这两个模型之后, 我们保留了网络结构中除去分类层以外的部分。足迹图像库中的每张图像将由这两个模型提取特征, 并建立足迹图像特征向量库。其中通过 ResNet50 和 DenseNet121 模型提取的每张图片的特征向量维度为 512。

3.4 图像聚类

考虑到我们的数据量较小, 实验采用 DBSCAN 算法以实现图像聚类, 该算法可在 N 维空间中对紧密排列的点进行分组, 其中靠在一起的点将被分为一组。DBSCAN 聚类算法可将之前提取好的足迹图像特征向量拉到高维空间中, 相似的足迹图像将自动归为一类, 进而实验足迹图像的聚类。当然 DBSCAN 算法还可以轻易处理异常值, 如果一些数据散落在低密度区域, 则会被标记。

4 实验

4.1 实验设备及数据集介绍

本文中的实验均在 Ubuntu16.04 系统上运行, 并且使用 RTX2080ti 显卡加速运算。其中实验数据由合作方昆山市公安局提供。针对犯罪现场足迹图像的实际情况, 首先通过人工挑选, 去除其中质量较差的图像, 并对剩余的数据进行标记。同时, 进行了数据扩充和增强, 例如旋转、随机裁剪、灰度值和对比度变换等操作。最终本文采用的实验数据集共包含 51 个类别, 有超过 5000 张足迹图像。

4.2 实验评估指标

V-measure^[12] 是聚类评价指标同质性 (homogeneity) 和完整性 (completeness) 的均衡平均。其中同质性指标表示一个集群内是否仅包含有单一类别分子; 完整性指标表示已知在

同一类别中的分子经过聚类后是否都分配在同一个集群。V-measure 取值在 0 到 1 之间且值越大则表明聚类结果越好。

轮廓系数 (Silhouette Coefficient)^[13] 体现了同一集群内样本间的距离, 对于集群内某一样本, 设 a 是与它同类别中其他样本的平均距离, b 是与它距离最近的不同类别中样本的平均距离, 其轮廓系数 s 为:

$$s = \frac{b-a}{\max(a,b)} \quad (3)$$

轮廓系数的取值范围是 $[-1, 1]$, 且分数越高表明聚类效果越好, 此时表明同类别样本间距离越近, 不同类别样本间距离越远。

互信息 (Mutual Information)^[14] 用来衡量已知数据标签 X 和预测数据标签 Y 之间分布的一致程度, 评判两者是否有关系以及关系的强弱。如果 $(X, Y) \sim p(x, y)$, 则 X, Y 之间的互信息 $I(X; Y)$ 定义为:

$$I = (X; Y) = \sum_{x \in X} \sum_{y \in Y} p(x, y) \log \frac{p(x, y)}{p(x)p(y)} \quad (4)$$

其取值范围在 $[0, 1]$ 之间, 且值越大意味着聚类结果与真实情况越吻合, 此时聚类效果更好。

调整兰德系数 (Adjusted rand index) 中, 兰德指数 (Rand index)^[15] 衡量两个数据分布的吻合程度, 需要给定实际类别信息 C , 假设 K 是聚类结果, a 表示 C 与 K 中都是同类别的元素对数, b 表示 C 与 K 中都是不同类别的元素对数, 则兰德指数为:

$$RI = \frac{a+b}{C_2^{n_{samples}}} \quad (5)$$

其中, $C_2^{n_{samples}}$ 表示数据集中可以组成的总元素对数。但是对于随机结果 RI 并不能保证分数接近零, 针对此现象调整兰德系数被提出, 其也具有更高的区分度, 公式如下:

$$ARI = \frac{RI - E[RI]}{\max(RI) - E[RI]} \quad (6)$$

ARI 取值范围为 $[-1, 1]$, 值越大意味着聚类结果与真实情况越吻合。从广义的角度来讲, ARI 衡量的是两个数据分布的吻合程度。

4.3 实验结果与分析

本文分别对 ResNet50 和 DenseNet121 提取的图像特征, 使用 DBSCAN 方法进行聚类。其中 DBSCAN 算法参数 $min_samples$ 即样本点要成为核心对象所需要的 ϵ -邻域的样本数阈值, 设置为 5; 邻距离度量参数 $metric$ 采用默认的欧氏距离公式。图 4 和图 5 分别展示了在两个图像特征上的 DBSCAN 聚类效果, 其中 X 轴为 Eps 值, 取值范围为 $[0.1, 0.9]$, Y 轴为对应 Eps 的评价指标分数, 绿色曲线代表 V-measure, 红色曲线代表调整兰德系数分数, 天蓝色曲线代表互信息分数, 蓝色曲线代表轮廓系数分数。

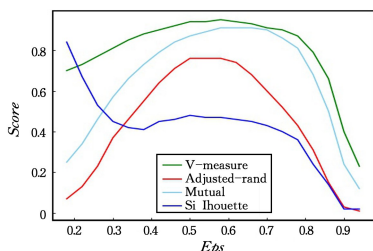


图 4 ResNet 提取的图片特征聚类结果 (电子版为彩色)

Fig. 4 Image feature clustering results extracted by ResNet

从图 4 可以发现, 互信息、调整兰德系数及 V-measure 均随着 DBSCAN 算法参数 Eps , 即 ϵ -邻域的距离阈值的增加而增加, 在 Eps 为 0.5 时趋于平缓, 而后随着 Eps 的增大变小。轮廓系数则在 Eps 小于 0.4 前, 随 Eps 值的增大而减小, 同时在 Eps 为 0.5 附近时趋于平缓随后逐渐减小。

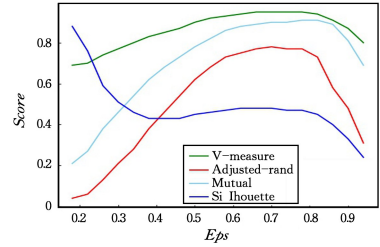


图 5 DenseNet121 提取的图片特征聚类结果 (电子版为彩色)

Fig. 5 Image feature clustering results extracted by DenseNet121

从图 5 可以发现, 互信息、调整兰德系数及 V-measure 均随着 DBSCAN 算法参数 Eps , 即 ϵ -邻域的距离阈值的增加而增加, 在 Eps 为 0.6 时取得最大值, 而后趋于平缓随后下降。轮廓系数则在 Eps 小于 0.5 前, 随 Eps 值的增大而减小, 同时在 Eps 为 0.6 附近时趋于平缓随后逐渐减小。

如表 1 所列, 我们对四大评价指标取得的最优值以及取得最优值时对应的 Eps 值进行了统计对比。其中 ResNet 提取的图像特征聚类结果在 $Eps=0.58$ 时, V-measure 取得最大值 0.95; 在 $Eps=0.54$ 时, 调整兰德系数取得最大值 0.76; 在 $Eps=0.62$ 时, 互信息取得最大值 0.91; 在 $Eps=0.62$ 时, 轮廓系数取得稳定最大值 0.48。而 DenseNet121 提取的图像特征聚类结果在 $Eps=0.7$ 时, V-measure 取得最大值 0.95; 在 $Eps=0.7$ 时, 调整兰德系数取得最大值 0.78; 在 $Eps=0.74$ 时, 互信息取得最大值 0.91; 在 $Eps=0.66$ 时, 轮廓系数取得稳定最大值 0.48。

表 1 评价指标 Max_score 及 Eps 取值

Table 1 Evaluation index Max_score and Eps value

评估指标	ResNet		DenseNet	
	Eps	Max_score	Eps	Max_score
V-measure	0.58	0.95	0.70	0.95
Adjusted rand index	0.54	0.76	0.70	0.78
Mutual Information	0.62	0.91	0.74	0.91
Silhouette	0.62	0.48	0.66	0.48

如图 6 和图 7 所示, 我们对 ResNet 提取的图像特征及 DenseNet121 提取的图像特征的部分聚类结果进行展示, 由于整个数据集图片种类众多, 为了更加直观简单的展示, 我们只选择了聚类结果中的 3 类进行展示, 可见不同类之间并没有很明显的掺杂现象。

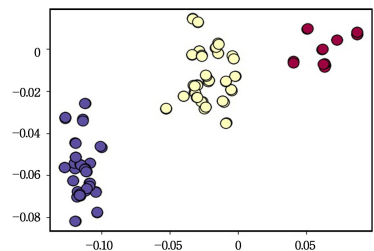


图 6 ResNet 部分聚类结果

Fig. 6 Partial clustering results of ResNet

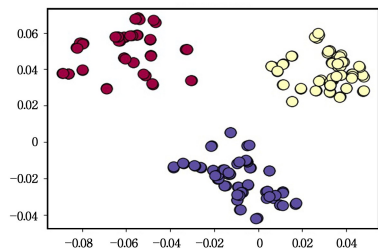


图7 DenseNet121 部分聚类结果

Fig. 7 Partial clustering results of DenseNet121

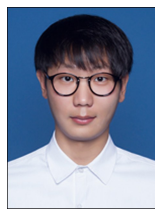
综合以上结果,结合对 ResNet 和 DenseNet 取得的图像特征的聚类结果横向分析可得,尽管两者在互信息、调整兰德系数、V-measure 及轮廓系数四大聚类评价指标取得最优值时的参数 Eps 设置不同,但都取得相当不错的结果,且各指标的数据曲线基本一致。此外,结合表 1 纵向对比可发现,当各指标均取得最优值时,参数 Eps 也稳定在同一个值附近。以 DenseNet121 提取的图像特征聚类结果为例,四大指标在 Eps 值为 0.7 附近时均取得了最优值,这也与 DBSCAN 聚类算法的特点相吻合。同时,四大指标整体表现较好,其中 V-measure 和 Mutual Information 的分数更是达到了 0.9 以上。最后,结合图 6 和图 7 中两者部分的聚类结果可知,两者聚类结果中类与类之间边界清晰,无明显的相互掺杂现象,聚类效果良好。

结束语 本文通过两个 CNN 模型提取了足迹图像的特征,并分别结合 DBSCAN 聚类算法实现了足迹图像聚类。通过实验表明,该方法能够有效帮助公安实现智能化的足迹图像归类,大大减少人力和时间消耗,且在一定程度上辅助公安进行串并案的侦破,具有实际应用价值。

本文的研究工作也存在一些不足,如存在数据量较小,且数据分布不均等问题。下一步将重点研究大数据量下的足迹图像聚类。

参 考 文 献

- [1] DICKENSON M, GUEGUEN L. Rotated Rectangles for Symbolized Building Footprint Extraction[C]// IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops. 2018:225-228.
- [2] SUN Y, WANG Y Y, ZHU X X. Automatic Registration of SAR Image and GIS Building Footprints Data in Dense Urban Area [C]// 2019 IEEE International Geoscience and Remote Sensing Symposium. 2019:927-930.
- [3] LI Z S. Simulation Study on the Rapid Identification of Footprint Image of Aerobics Athletes [J]. Computer Simulation, 2017, 34(3):221-224.
- [4] CHEN Y, ZENG C, CHENG C, et al. A CNN-based Approach to Footprint Image Retrieval and Matching[J]. Journal of Nanjing Normal University(Engineering and Technology Edition), 2018, 18(3):45-51.
- [5] ZHOU F Y, JIN L P, DONG J. Review of Convolutional Neural Network[J]. Chinese Journal of Computers, 2017, 40(6):1229-1251.
- [6] LECUN Y, BOTTOU L, BENGIO Y, et al. Gradient-based learning applied to document recognition[J]. Proceedings of the IEEE, 1998, 86(11):2278-2324.
- [7] CHEN S D, WEI W, HE B Q, et al. Action recognition base on improved deep convolutional neural network[J]. Application Research of Computers, 2019(3):62.
- [8] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2016:770-778.
- [9] HUANG G, LIU Z, VAN DER MAATEN L, et al. Densely connected convolutional networks[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2017:4700-4708.
- [10] WANG J H, ADRIAN H, JIANG J M. Spectral Analysis Network for Deep Representation Learning and Image Clustering [C]// 2019 IEEE International Conference on Multimedia and Expo (ICME). 2019:1540-1545.
- [11] SCHUBERT E, SANDER J, ESTER M, et al. DBSCAN revisited, revisited: why and how you should (still) use DBSCAN[J]. ACM Transactions on Database Systems (TODS), 2017, 42(3):1-21.
- [12] NEWMAN M E J, CANTWELL G T, YOUNG J G. Improved mutual information measure for clustering, classification, and community detection[J]. Physical Review E, 2020, 101(4):23-34.
- [13] SUN S L, WANG C, ZHAO Y D. Parameter independent clustering of air traffic trajectory based on silhouette coefficient[J]. Journal of Computer Applications, 2019, 39(11):3293-3297.
- [14] QIAN X D, LUO Y F. Incomplete Data Clustering Algorithm Based on Mutual Information Attributes Ranking[J]. Information and Control, 2019, 48(1):84-91.
- [15] XIE J Y, ZHOU Y, WANG M Z, et al. CAAI Transactions on Intelligent Systems[J]. New Criteria for Evaluating the Validity of Clustering, 2017, 12(6):873-882.



CHEN Yang, born in 1994, postgraduate. His main research interests include deep learning and computer vision.



XI Xue-feng, born in 1978, Ph.D, associate professor, master supervisor. His main research interests include natural language processing and machine learning.