

智慧城市管理案件图像自动识别应用分析^①



郭磊, 李进强, 党磊

(正元地理信息集团股份有限公司, 北京 101300)
通讯作者: 郭磊, E-mail: pujianf5411740@163.com

摘要: 智慧城市就是运用信息通信等技术手段感测、分析城市运行核心系统的各项关键信息, 需要对城市安全、常见案件等做出及时有效的智能响应. 为了提高城市常见案件处理的有效性和精确度, 本次研究提出了一种城市常见违规事件的自动识别算法. 利用改进的卷积神经网络对图像进行特征提取, 使用 BP 神经网络作为评价网络. 在 VOC 数据集上与 YOLO、SSD 等算法的性能进行对比. 结果表明, 改进的卷积神经网络的检测 mAP 值可达 76.5%, 对各种类型的案件识别的准确率均在 72% 以上, 识别“乱涂乱画张贴广告”类型图像的准确率达到 83.4%. 此次研究构建的智慧城市案件图像识别技术能够有效提高案件处理效率、节约人力物力资源, 可用于协助城管监察行政执法.

关键词: 目标检测算法; 卷积神经网络; BP 神经网络; 案件管理; 识别技术

引用格式: 郭磊, 李进强, 党磊. 智慧城市管理案件图像自动识别应用分析. 计算机系统应用, 2021, 30(5): 241-246. <http://www.c-s-a.org.cn/1003-3254/7923.html>

Application Analysis of Image Automatic Recognition in Smart City Management Cases

GUO Lei, LI Jin-Qiang, DANG Lei

(Zhengyuan Geographic Information Group Co. Ltd., Beijing 101300, China)

Abstract: Smart cities depend on information and communication technology to sense and analyze the key information in the core system of urban operation. It needs to make timely and effective intelligent response to urban security threats and common cases. In order to improve the effectiveness and accuracy of identifying common urban cases, this study proposes an automatic identification algorithm for common urban violations. The improved convolution neural network extracts image features, and BP neural network is used for evaluation. On the VOC data set, the algorithm is compared with YOLO and SSD in performance. The results show that the mAP of the improved convolutional neural network can reach 76.5%, and the accuracy of identifying various types of cases is more than 72%, and the accuracy of identifying “graffiti and posted advertisement” is 83.4%. The image recognition technology of cases in a smart city developed in this study can enhance the efficiency of case processing, save human and material resources, and thus can be used to assist urban management, supervision, and administrative law enforcement.

Key words: object detection algorithm; Convolution Neural Network (CNN); BP neural network; case management; recognition technology

城市监控视频采集和处理技术是智慧城市建设的
关键技术. 目标检测算法是深度学习算法中重要的一
部分, 已经在人脸识别、机械设备故障检测等领域有

了较好地应用^[1]. 在智慧城市案件图像识别领域, 目标
检测算法能够根据事物的不同特征在一定重复学习的
基础上, 精确判断图像内容的分类, 在一般城市案件识

^① 收稿时间: 2020-09-23; 修改时间: 2020-10-21; 采用时间: 2020-10-30; csa 在线出版时间: 2021-04-28

别中能够发挥重要作用,如车辆违章停放、违规张贴广告等.该算法还可以通过群体行为判断案件的严重性,能够及时向系统反馈异常情况.有效的城市案件图像识别技术对提高城市案件发现和处理效率,节省城市人力物力资源有极大的影响^[2].目标检测算法在目前城市管理案件图像自动识别的实际应用上还存在一定问题,识别精度不高、检测速度较慢等.有鉴于此,本次研究以提高检测城市案件图像识别的速度和精度为目标,构建了基于改进卷积神经网络的目标检测算法和图像评价网络模型.

1 基于卷积神经网络的目标检测算法

1.1 改进的卷积神经网络

卷积神经网络训练模型会涉及到大量且复杂的计算,其实质是通过卷积层将初始特征图像中提取出有效特征,实现图像滤波^[3].假设每个卷积层的初始特征图是 x_i ,则卷积运算表达式为:

$$x'_j = f \left(\sum_{i \in M_j} x_i^{l-1} * k_{ij}^l + b_j \right) \quad (1)$$

式(1)中, M_j 是初始特征图像的集合, $f(\cdot)$ 是激活函数,*是卷积运算的数学符号.将第*i*和第*j*个初始特征图像的卷积层卷积核标记为 K_{ij} .

对第*l*层的神经元的权重及更新值进行求解,首先需要计算节点处的灵敏度 δ .

$$\delta_j^l = \delta_j^{l+1} W_j^{l+1} \circ f'(u^l) = \beta_j^{l+1} u_p^{(l+1)} \circ f'(u^l) \quad (2)$$

式(2)是位于第*l*层的*j*神经元的灵敏度 δ_j^l 的计算公式.其中连接层*l*至连接层*l+1*之间结点的灵敏度的和为 δ_j^{l+1} ,其权重值为 W . β_j^{l+1} 为一常数,是第*l+1*下采样层的初始特征图像的权重. $f'(u^l)$ 是第*l*层神经元在输入为*u*的情况下的激活函数的倒数.

Bias 基的梯度的计算表达式为^[4]:

$$\frac{\partial E}{\partial b_j} = \sum_{u,v} (\delta_j^l)_{uv} \quad (3)$$

卷积核梯度运算公式为:

$$\frac{\partial E}{\partial k_{i,j}^l} = \sum_{u,v} (\delta_j^l)_{uv} (p_i^{l-1})_{uv} \quad (4)$$

此次研究使用卷积神经网络对智慧城市案件图像进行自动识别,检测网络的基础网络为 VGGNet,并增加了辅助层,共有 conv5_3、conv4_3 等 8 个卷积层,其

中 conv5_3、conv4_3 卷积层提供小目标特征.conv5_3、conv4_3 两个卷积层可将图像中的某些具体特征进行提取分析,从而提高识别精度.小目标特征提取之后的图像将通过仪器处理最终输出尺寸相等的图像.

检测网络在进行训练模型时,各网络层可能由于相互连接造成输入数据偏移或增大,影响训练效果^[5,6].为了加快训练速度,防止偏移情况发生,本次研究构建的卷积神经网络在传统模型的激活函数前加入批次规范化 (Batch Normalization, BN) 层.将输入激活函数的所有数据进行均值和标准差归一化处理.

$$\begin{cases} \mu_B \leftarrow \frac{1}{m} \sum_{i=1}^m x_i \\ \sigma_B \leftarrow \frac{1}{m} \sum_{i=1}^m (x_i - \mu_B)^2 \\ \hat{x}_i \leftarrow \frac{x_i - \mu_B}{\sqrt{\sigma_B^2 + \epsilon}} \\ y_i \leftarrow \gamma \hat{x}_i + \beta \equiv BN_{\gamma, \beta}(x_i) \end{cases} \quad (5)$$

式(5)是BN算法的向前步骤,其中*B*是单批次图像特征, γ 和 β 是学习重构参数, μ_B 是BN的平均值, σ_B 是BN的方差, \hat{x}_i 代表规范化.在BN算法运算时,输入对象是 γ 和 β ,经过批次规范化标准差和方差处理之后,输出为 y_i .

传统的神经卷积层在处理非线性和存在一定抽象化特征的图像时,精确率和速度都有所下降.针对智慧城市案件处理的特殊性,本次研究构建的网络使用多层感知层.针对城市案件图像特征复杂导致的网络参数过多的情况,此次研究在原来网络的基础上加入一个函数逼近器.研究构建的带逼近器的多感知层网络与传统线性卷积层对比情况如图1所示.

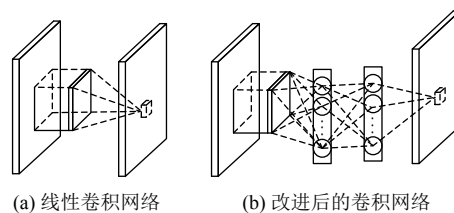


图1 典型和改进之后的卷积层对比

图1(a)是典型的线性卷积层,图1(b)是加入1×1卷积之后的卷积运算网络.函数逼近器的加入可以改变感知野中的网络结构进行抽象和非线性特征的提取.相比于典型的卷积层,改进之后的卷积层添加了一个

1×1 的卷积, 并加入了一个经过修正的线性激活函数, 表达式为:

$$\varphi(x) = \begin{cases} x, & \text{if } x > 0 \\ 0.1x, & \text{otherwise} \end{cases} \quad (6)$$

式(6)中, x 是输入数据. 当输入数据为正数时, 经过线性激活函数处理之后得到的数值为原数值. 若输入数据小于零, 则核函数将输入数值缩小十倍之后输出. 使用该线性激活函数可以降低特征图的维度, 扩大网络的应用规模, 进而增加卷积神经网络的深度和宽度.

本次研究构建的卷积神经网络算法流程如图2所示. 在图像信息的前向传播阶段, 算法会对初始权值进行随机赋值. 经过多层感知层、函数逼近器和池化层处理, 可以得到待测图像最为全面的特征. 之后全连接层会对多个隐含层的信息进行变换并传输至输出层. 输出层的输出结果与预期结果比较, 若误差函数满足要求; 则直接输出结果, 若不满足要求, 误差和权值信息反向传播, 算法对其进行重新赋值.

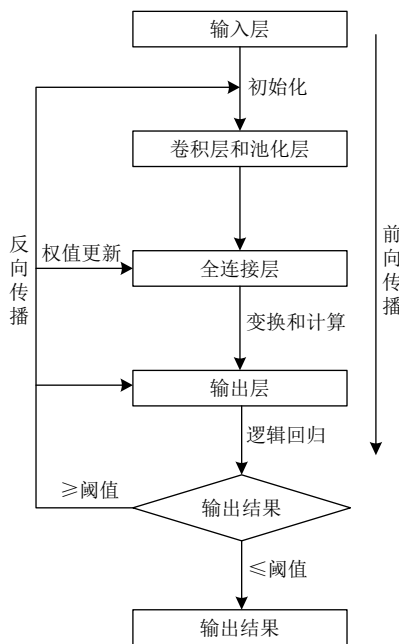


图2 改进的卷积神经网络算法流程图

1.2 评价网络结构

BP神经网络的分布式存储和并行协同处理机制能够快速对的大量数据进行分析^[7]. 典型的BP神经网络的拓扑图包含三层前馈网, 并且各层神经元只与相邻神经元之间全连接, 同层神经元之间无连接, 各层神

经元之间无反馈连接.

智慧城市案件识别模型的评价网络结构示意图如图3所示, 共有输入层、隐含层和输出层3部分构成. 输入层的输入数据是案件图像中关键内容的中心偏移量 x, y 和尺寸 w, h . 通过输入数据的正向传播和出现误差后实施的反向传播, 不断缩小误差函数值, 从而实现提高输出数据精确度的目的.

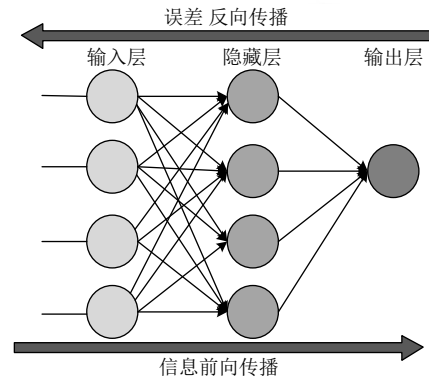


图3 城市案件图像的评价网络图

在对评价网络进行训练时, 需要将案件图像进行处理, 获取关键内容的主要数据. 这些数据输入训练网络, 通过图像分类和不同图像关键内容的区别进行BP神经网络的训练, 从而对案件类型做出准确地判断. 如果训练样本集的内容较少, 很容易出现训练结果拟合的现象^[8,9]. 针对这一情况, 本次研究使用数据集扩充的方式增加样本数据, 对原始图片设置0°、90°、180°和270°的旋转角度, 从而扩充数据, 提升网络学习图像不变性特征的能力. 本次研究的损失网络可以分为两个部分: 回归误差和类别误差.

$$L(x, c, l, g) = \frac{1}{N} L_{\text{conf}}(x, c) + \alpha L_{\text{loc}}(x, l, g) \quad (7)$$

式(7)是本次研究的网络损失函数表达式, 其中 c 是置信度, l 是预测框, g 是真实的框. N 表示匹配到的默认框的数量, α 反映定位损失和置信度损失权重的比例情况.

本次研究中目标物的位置损失表示如下:

$$\text{smooth}_{L1}(x) = \begin{cases} 0.5x^2, & \text{if } |x| < 1 \\ |x| - 0.5, & \text{otherwise} \end{cases} \quad (8)$$

1.3 图像处理技术

目标检测算法在对彩色图像进行处理时所消耗的时间和数据容量远超于处理灰色图像, 而智慧城市的

管理系统要求目标检测算法能够快速对案件图像进行识别^[10]。由此,灰度化图像的过程必不可少。

本次研究采用加权平均的方法对图像进行处理,公式具体内容如下:

$$F(i, j) = aR(i, j) + bG(i, j) + cB(i, j) \quad (9)$$

式(2)中,灰度化加权系数 a, b, c 分别取值为 0.3、0.59、0.11。

案件图像的像素灰度值的公式为:

$$\begin{cases} Gray = 0.3R(i, j) + 0.59G(i, j) + 0.11B(i, j) \\ R = G = B = Gray \end{cases} \quad (10)$$

式(10)中, R 为案件图像红色分量, G 为绿色分量, B 为蓝色分量。

城市案件图像灰度化处理结果如图4所示。图4(a)是未经处理的城市车辆违章停靠图像;图4(b)是经过灰度变换之后的图像。灰度化处理之后图像的关键信息的清晰度并未有下降,并且明暗、阴影、对比等信息均可以通过对比不同部分的灰度值获取。



图4 城市案件图像灰度化处理结果

从案件图像中提取关键目标信息还需要克服背景和噪声因素影响,本次研究采用灰度阈值化的方法提取图像的关键内容,将复杂的背景设置为0,前景目标值设置为255^[11,12]。由此经过处理之后得到的图像便只有黑色和白色两种颜色,更加有利于对案件图像进行分析,提高案件处理的速效率。此次研究采用二值化方法,通过最大类间差法计算得到:

$$\begin{cases} u = w_0 \times u_0 + w_1 \times u_1 \\ g = w_0 \times (u_0 - u)^2 + w_1 \times (u_1 - u)^2 \end{cases} \quad (11)$$

式(11)中, w_0, w_1 分别是图像前景、图像点数占据整张图的比例值, u_0, u_1 分别是前景和背景的平均灰度值, g 是类间方差值。

$$g = w_0 \times w_1 \times (u_0 - u_1)^2 \quad (12)$$

式(12)是对式(11)的进一步推导。当类间方

差值取最大值时,图像的前景和背景的差异度达到最大,此时分割前景和图像背景的阈值为 T 处于最佳状态。

图5(a)是经过灰度化处理之后得到的案件图像,图5(b)是经过灰度阈值化之后的图像。经过处理之后,图像仅有黑白两种颜色,并且图像特征更加鲜明,目标检测算法对图像的处理效率和精确度也会因图像特征的鲜明而显著提高。



图5 案件图像的二值化处理结果

2 目标检测算法模型识别效果检验

2.1 智慧城市案件图像数据集构建

目前在城市案件图像检测领域并没有明确的标准数据集,本次研究以实际城市管理系统常见的案件为分类依据,以网络上公开来源的案件图片为数据集来源,以城市管理监控视频中提取的关键帧为补充,构建了8个典型城市案件的图像数据库,8种分类包括非机动车违章停放、共享单车违规停放、雨水算子损坏、井盖损坏、乱涂乱画张贴广告、暴露性垃圾、交通护栏破损和机动车违章停放等8类。将不清晰、内容真实性存疑的图片进行删除,最终得到的数据集中共有9345张图片,每类城市案件约有1168张图片。每类图片的60%为训练集,其余图片为测试集。

图6是数据集中4类案件类型的示例图,每张图片的案件类型具有鲜明的特征,并无车辆、绿化、建筑等遮蔽物的遮挡。8类典型城市案件类型的编号分别为 event 0551、event 0552、unite 0550、unite 0100、event 0201、event 0102、unite 0211、event 0550。

2.2 目标检测算法识别效果分析

本次研究对构建的基于卷积神经网络的目标检测算法的测试在 Ubuntu18.04 操作系统上进行,采用 TensorFlow 深度学习框架。研究提出的算法与 Fast R-CNN、Faster R-CNN、YOLO、SDD 算法的测试结果对比情况如图7所示。

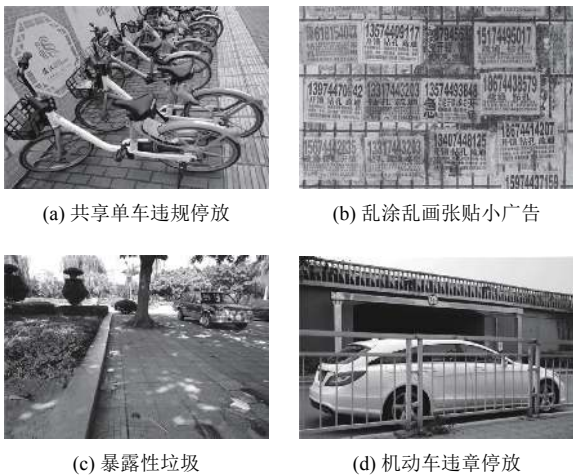


图6 智慧城市案件图像数据集示例图(部分)

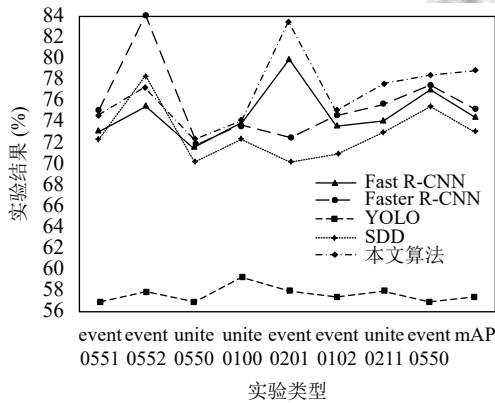


图7 5种算法的案件图像识别结果

图7是几种算法的对比实验结果情况,本次研究构建的目标检测算法在8类典型城市案件图像的识别上具有较好的识别效果,对各种类型的案件识别的准确率均在72%以上,识别“乱涂乱画张贴广告”类型图像的准确率达到83.4%。Faster R-CNN算法的识别效果略逊于本文算法,但是相较于YOLO算法等具有较高的识别准确率。这是由于本次研究构建的目标检测算法在卷积神经网络中增加了提取小目标特征的层数。此次研究构建的自动识别技术的mAP值为76.8%,高于其他技术。

图8是图像帧数为10 fps的情况下,不同自动识别技术对测试集中所有图像的处理情况。图像直观显示,各类算法检测速度的情况为: SDD算法(48 s)>YOLO算法(47 s)>本文算法(40 s)>Faster R-CNN算法(32 s)>Fast R-CNN算法(25 s)。本次研究构建的算法处理图像速度处于中等水平,并且与最高水平相差

不大。而检测速度最高的SDD算法的识别精确度最低,实际应用价值不大。本次研究构建的算法在图像识别精度和检测速度中有一定均衡,在保证识别效率的同时,也确保了识别的准确性。

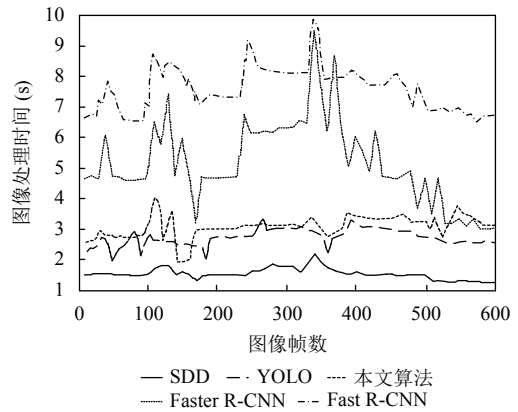


图8 5种算法识别案件图像类型耗时对比

以VOC2007+2012为数据集对100 Hz DPM、30 Hz DPM、YOLO、SDD以及本文算法进行训练和检测,检测结果如图9所示。从图中可以看出,虽然100 Hz DPM的检测速度最高,但是其检测精度却很低,并不能满足城市案件处理中对识别算法精确度和处理速度的要求。而本次研究构建的算法虽然在检测速度上处于中等水平,但是精确度最高。由此证明该算法具有较优的有效性和鲁棒性。

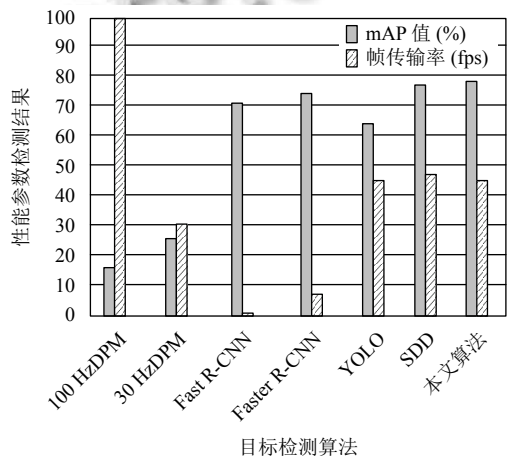


图9 VOC2007+2012环境下的实验结果

3 结论

城市案件的发现和识别速度是体现智慧城市管理水平的重要标准之一,传统的人工检测方式处理速度

较慢、识别精度较低。将深度学习与智能检测技术相结合是目前学术界研究的热点方向。有鉴于此,本次研究构造了基于改进卷积神经网络的目标检测算法,运用多层网络结构提取小目标特征。研究将改进之后的算法与YOLO算法等典型目标检测算法进行对比,结果表明,改进之后的算法相比于其他算法,鲁棒性和精确度有了明显提升,在检测准确率上可达83.4%。本次研究提出的识别技术能够为建筑智慧城市、提升城市安全提供一定的保障。但是目前缺乏标准、权威且成熟的城市案件数据集,因此后续还需要进行数据扩充,进一步检验自动识别技术的性能,改进其中可能存在的问题。

参考文献

- 1 刘颖,刘红燕,范九伦,等.基于深度学习的小目标检测研究与应用综述.电子学报,2020,48(3):590-624.[doi:10.3969/j.issn.0372-2112.2020.03.024]
- 2 郭金虎,顾曦文.基于图像缺损和遮挡的目标识别算法研究.光电技术应用,2016,31(2):54-56.[doi:10.3969/j.issn.1673-1255.2016.02.014]
- 3 赖特.基于模板匹配及神经网络算法的图像识别应用——MATLAB实现机动车牌号码辨识.智能建筑与智慧城市,2017,(11):45-48,52.[doi:10.13655/j.cnki.ibci.2017.11.015]
- 4 姬晓飞,秦宁丽.光学遥感图像多目标检测及识别算法设计与实现.计算机应用,2015,35(11):3302-3307.[doi:10.11772/j.issn.1001-9081.2015.11.3302]
- 5 罗荣辉,袁航,钟发海,等.基于卷积神经网络的道路拥堵识别研究.郑州大学学报(工学版),2019,40(2):18-22.
- 6 申扬,李巍,刚毅凝,等.基于背景先验与低秩恢复的显著性目标检测方法.计算机与现代化,2019,(1):33-39.[doi:10.3969/j.issn.1006-2475.2019.01.007]
- 7 杨辉辉,张天宇,李灵巧,等.基于MobileNet的移动端城管案件目标识别算法.计算机应用,2019,39(8):2475-2479.[doi:10.11772/j.issn.1001-9081.2019010232]
- 8 李隽,王伟.遗传算法优化神经网络在图像目标识别中的应用研究.现代电子技术,2017,40(20):111-113.[doi:10.16652/j.issn.1004-373x.2017.20.031]
- 9 张叶,胡玉兰,片兆宇.基于局部空间冗余视觉信息抑制的目标识别算法研究.微型机与应用,2016,35(1):42-44,48.[doi:10.3969/j.issn.1674-7720.2016.01.013]
- 10 单琳娜.基于运动目标的图像识别及追踪技术应用研究.佳木斯大学学报(自然科学版),2018,36(6):880-882.
- 11 王婷婷,潘祥.基于卷积神经网络的目标检测算法研究.长春师范大学学报,2020,39(6):42-48.
- 12 张艳艳,娄莉,梁硕.基于改进光流算法的运动目标检测技术研究.智能计算机与应用,2018,8(1):55-58.[doi:10.3969/j.issn.2095-2163.2018.01.013]