doi: 10.6046/gtzyyg.2020.04.09

引用格式:陈朋弟,黄亮,夏炎,等. 基于 Mask R – CNN 的无人机影像路面交通标志检测与识别[J]. 国土资源遥感,2020,32 (4):61 – 67. (Chen PD, Huang L, Xia Y, et al. Detection and recognition of road traffic signs in UAV images based on Mask R – CNN [J]. Remote Sensing for Land and Resources, 2020,32(4):61 – 67.)

基于 Mask R - CNN 的无人机影像 路面交通标志检测与识别

陈朋弟¹,黄亮^{1,2},夏炎¹,余晓娜³,高霞霞¹ (1.昆明理工大学国土资源工程学院,昆明 650093;2.云南省高校高原山区空间信息测绘 技术应用工程研究中心,昆明 650093;3.昆明工业职业技术学院,昆明 650093)

摘要:交通标志的检测与识别是智能驾驶导航系统的重要组成部分,但传统方法的处理过程由于精度低、时间复杂 度高以及鲁棒性差等缺点,不能满足当前智能驾驶的需求。为此,提出了一种基于 Mask R - CNN 的无人机影像路 面交通标志检测与识别方法。首先,制作了一套高质量的无人机影像路面交通标志数据集;然后,根据统计的 200 个标记路标特征,对 Mask R - CNN 中区域候选网络(region proposal network, RPN)结构的锚框宽高比及初始参数进 行了改进,使其更好地应用于无人机影像路标场景;最后,采用精确度 - 召回率(precision - recall, PR)曲线和平均 精度值(mean average precision, mAP)进行精度评价。实验结果表明,锚框宽高比为1:1,1:2,1:3 时效果更好;该方 法得到的识别结果平均检测精度为 98.33%,高于 Faster R - CNN 和 YOLOv3 方法,具有较好的有效性。 关键词;无人机影像;交通标志检测;交通标志识别; Mask R - CNN; RPN;锚框

中图法分类号: P 231 文献标志码: A 文章编号: 1001 - 070X(2020)04 - 0061 - 07

0 引言

随着科技的进步,智能驾驶技术得到了空前发展,驾驶辅助系统受到了极大的关注,人们对驾驶辅助系统的要求也更加精细化,从之前的道路级导航发展到精细单车道级的需求。无人机影像由于获取方便、视野广且分辨率高等优点,为大面积交通标志的获取提供了条件,通过将检测和识别到的路面交通标志与各大地图软件相结合,可以实时而准确地为驾驶员提供交通信息。因此如何快速、高质量地从无人机影像中检测和识别路面交通标志(后文简称为路标)便成为了亟待解决的问题。

目前对于路标检测与识别的研究较少, Chira 等^[1]提出了一种模板匹配和边缘检测相结合的路 标检测方法,可轻松部署在专用计算机上,并用于驾 驶辅助系统中; Schreiber 等^[2]利用矢量图来训练光 学字符识别(optical character recognition, OCR)的对 象,采用消失点并利用逆透视投影变换(inverse perspective mapping, IPM)得到俯视图, 最后使用 TES-SERACT 对符号进行分类; Wu 等^[3] 通过训练数据 得到道路的标志模板,然后使用区域特征提取 (maximally stable extremal regions, MSER),执行模板 匹配来检测多个道路标志: Maier 等^[4] 通过提取车 辆前方的感兴趣区域,然后将其编码为圆弧样条,再 与提取的目标候选轮廓进行比较从而实现箭头的检 测与分类。此外,一些学者利用机器学习和深度学 习对路标检测与识别也开展了研究^[5-8]。Danescu 等^[9]利用水平线区域的暗光暗转换检测提取被绘 制对象的特征,然后利用高斯混合灰度分割对其进 行细化,并利用透视几何法重建三维边界框,最后使 用决策树约束对对象进行分类; Wang 等^[10] 通过 IPM 变换得到俯视图,然后利用改进的 Haar 小波提 取箭头标志特征,最后通过支持向量机对箭头进行 识别; Liu 等^[11]首先通过 IPM 变化抑制透视效果.

通信作者:黄亮(1985-),男,博士,硕士生导师,主要研究方向为遥感影像变化检测。Email: kmhuangliang@163.com。

收稿日期: 2020-01-15;修订日期: 2020-03-06

基金项目:国家自然科学基金项目"南方山地城镇建设用地与变化的坡度梯度效应研究"(编号:41961039)、云南省应用基础研究计划面上项目"基于全卷积神经网络的多源遥感影像变化检测"(编号:2018FB078)、云南省高校工程中心建设计划项目、自然资源部地球观测与时空信息科学重点实验室项目"基于直觉模糊集理论的多源遥感影像变化检测方法研究"(编号:201911)和昆明理工大学学生课外学术科技创新基金项目"基于 Mask Grid R - CNN 的无人机影像路面交通标志检测与识别系统"(编号:2020YB002)共同资助。

第一作者:陈朋弟(1993-),男,硕士研究生,主要研究方向为目标检测与遥感影像分割。Email: cpdhn1058475189@163.com。

• 62•

并利用滤波得到带道路标线的图像切片,然后通过 自适应增强(adaptive boosting, Adaboost)分类器和 极限学习机(extreme learning machine, ELM)分类器 来识别标志的类型;李强^[12]通过提取纹理信息并采 用局部软投票的方法分割出道路区域,然后利用显著 性融合和改进的 LeNet -5 神经网络对交通标志进行 检测和识别; Husan 等^[13]设计了一种卷积神经网络 (convolutional neural network, CNN)用于损坏路标的 检测和识别,该方法在一定程度上可对损坏路标进 行有效识别。

对于路边交通标志牌检测与识别是当前该领域 研究最多的方面,方法基本上都是基于神经网络 的^[14-19], Zhu 等^[20]提出了一种全卷积神经网络 (fully convolutional networks, FCN)和 CNN 相结合的 目标分类框架,先通过 CNN 对交通标志进行分类, 然后使用 FCN 建立边界框进行识别;伍锡如等^[21] 提出了一种基于图像聚类的交通标志 CNN 快速识 别算法,该算法将聚类思想和 CNN 网络结构相结 合,有效提高了路面交通标志牌的检测与识别;卢 飞宇[22]提出了一种基于改进的多任务级联卷积神 经网络(multi - task convolutional neural network, MTCNN) 交通标志检测与识别算法: Kryvinska 等^[23]设计了一种神经网络用于识别交通标志牌,该 网络有定位和识别2部分组成,并在瑞典交通标志 数据库(swedish traffic sign database, STSG)上进行 了测试,其精度只有63%,实用性较差。

上述方法的提出弥补了交通标志检测与识别的空

缺,但也有诸多不足之处,除对硬件的要求较高外,时间上也有很大的限制,同时也难以达到较高的准确率, 且当遇到复杂环境如阴影区域、遮挡路标以及磨损路标等情况时很难对交通标志进行有效的检测与识别。 为了实现无人机影像路标的高质量检测和识别,本文选取了当前较为先进的目标检测网络 Mask R - CNN 对路标进行检测和识别,并通过重新调整算法内部参数提高了路标识别准确率,克服了对阴影区路标、遮挡路标以及磨损路标等无法识别或识别精度低的问题。

1 数据集

因为当前并没有公开的路面标志数据集,而所 有已公开的数据集都为路边标志牌数据集,为此在 实验过程中手动进行了数据集制作。实验所用的原 始无人机影像是通过大疆 FC6310 和哈瓦 MEGA – V8 拍摄的,地点位于昆明市。采集的影像大小分别 为5 472 像素 ×3 648 像素和7 146 像素 ×5 364 像 素,共计 589 张。由于采集的影像数据中大部分没 有路标且多为重复拍摄区域,所以只选取了其中 200 张作为样本数据。为防止网络出现过拟合,且 能够识别路标的细节部分,实验过程中采用了旋转、 裁剪以及拉伸的方式将影像数量扩展到 800 张,然 后调整为实验所需的大小格式,再通过图像标注工 具 Labelme 手动进行路标标记,路标统一命名为 traffic。其中,80%为训练集,20% 为测试集,示例数 据集如表 1 所示。



2 Mask R – CNN 方法

Mask R – CNN 是 He 等^[24]提出的一种基于对 象的实例分割方法,它是在 Faster R – CNN^[25]的基 础上演化而来的网络模型。针对 Faster R – CNN 特 征图与原始图像对不准的问题, Mask R – CNN 取消 了 Faster R – CNN 中的感兴趣区域(region of interest, RoI)池化(Pooling),提出了 RoI 对齐(Align),并 增加了 Mask 预测分支,这样既增加了对小目标的 识别效果,又可以分割目标对象。Mask R – CNN 由 特征金字塔网络(feature pyramid networks, FPN)、区 域侯选网络(region proposal network, RPN)、Rol Align、Mask 预测分支以及原始分支组成。Mask 预测 分支是一个小型的 $FCN^{[26]}$,可对每个 Rol 进行语义 分割,而原始分支由回归和分类 2 部分组成,用来预 测路标。完整网络框架如图 1 所示。与 Faster R – CNN 相比, Mask R – CNN 的损失函数也因增加了分 割分支而产生了变化,新的损失函数 Loss 由 3 部分 组成,分别是分类损失 L_{els} 、边界框的回归损失 L_{box} 以及预测分割的交叉熵损失 L_{mask} ,即

$$Loss = L_{cls} + L_{box} + L_{mask} \quad (1)$$



图 1 Mask R - CNN 网络框架 Fig. 1 Mask R - CNN network framework

2.1 FPN

Faster R – CNN 对特征图的提取主要是通过 CNN 来实现的,它是将最后一层提取的特征图输入 到 RPN 中,但由于 CNN 的中间隐层网络会将小目 标当作噪声过滤掉或输入到 RPN 时某些锚点内的 小目标会被当成背景而处理掉,所以对小目标的识 别效果并不好。为解决该问题,Mask R – CNN 中引 入了 FPN 的方法,FPN 使用的是具有横向连接的自 上而下体系结构,其融合的特征图更具语义信息和 空间信息,准确度和速度上也获得了大幅度提升。 本文中使用的骨干网络为 ResNet101 + FPN。FPN 结构如图 2 所示^[27]。



Fig. 2 FPN architecture

该结构包括3个部分组成,分别为自底向上 (图2左上部分)、自顶向下(图2右上部分)和横向 连接(图2底部)。自底向上可视为网络的向前过 程,该过程中特征图的大小在经过某些层时会改变, 而在经过其他一些层时不会改变,这些不变的层定义 为一个阶段,每次提取的特征图都是每个阶段最后一 个层的输出。自顶向下实质为一个上采样过程,而横 向连接则是将上采样的结果和自底向上生成的相同 大小的特征图进行融合。在 FPN 结构的5 层特征金 字塔中,需要用到低纬度的4 个特征图,从中需要选 出一个作为 Rol 的获取,选取公式为:

$$k = k_0 + lb(\sqrt{wh/224})$$
, (2)

式中: *k* 为取整后需要切取的 RoI 特征图层号; *k*₀ 为 基准值,代表 P5 层的输出; *w* 和 *h* 分别为 RoI 的宽 和高; 224 为 ImageNet 数据集中图片标准输入 大小。

2.2 RPN

RPN 第一次出现是在 Faster R - CNN 结构中, 是用来提取候选框的,而 Mask R - CNN 中的 RPN 结构与 Faster R - CNN 相同。RPN 采用的是滑动窗 口机制,特征图上每个滑动窗口位置可预测 *k* 个候 选框,把每个候选框的中心点作为锚点,则会产生 *k* 个锚框,Mask R - CNN 使用了 3 种不同比例的锚框 宽高比。同时滑动窗口在中间层经过 3 × 3 卷积后 分别到达分类层和回归层,再经过 2 次 1 × 1 卷积后 得到 2 *k* 个分数和 4 *k* 个坐标。然后根据这些得到 的分数和坐标对所有锚框进行排序和过滤。最后再 通过非极大值抑制算法(non maximum suppression, NMS)预测剩余候选框的区域建议区,RPN 结构如 图 3 所示^[28]。



Fig. 3 RPN network architecture

实际中,路标的形状不同于其他物体,且大多为 小目标检测与识别。在原始 Mask R – CNN 的 RPN 中,锚框的宽高比分别取 1:1,1:2 和 2:1,因此,为 考虑模型更好地适用于应用场景,提高路标的检测 与识别效率,通过统计 200 个标记路标,将宽高比修 改为 1:1,1:2 和 1:3,其他结构保持不变。

2.3 RoI Align

RoI Align 消除了 RoI Pooling 带来的局限性,取 消了量化操作,在处理过程中使用双线性插值^[29]方

$$\frac{\partial L}{\partial x_i} = \sum_i \sum_j \left\{ d[i, i(r, j)] < 1 \right\} (1 - \Delta h) (1 - \Delta w) \frac{\partial L}{\partial y_{ij}} , \qquad (3)$$

式中: L 为代价函数; x_i 为池化前特征图上的点; y_{ij} 为池化后第r 个子区域中的第j 个点; i(r,j) 为一个 浮点数坐标; d[i,i(r,j)] 为像素点距离; Δh 和 Δw 分别为 x_i 与 i(r,j) 横纵坐标的差值。

3 结果与分析

3.1 Mask R - CNN 训练与测试

修改后的 Mask R - CNN 使用图形处理器 (graphics processing unit, GPU)进行训练和测试。 平台为 TensorFlow,计算机的配置为 Intel(R)i7 -9700k CPU,NVIDIA GeForce GTX1070 Ti 显卡,8 GB 显存。训练过程中初始学习率设置为0.001,权重衰 减系数设置为0.000 5,学习动量(MOMENTUM)设 置为0.9,IoU 阈值设置为0.7,迭代次数为4 000 次。图5 给出了 Mask R - CNN 在训练过程中损失 率的衰减情况。从图中可以看出,损失率随着迭代 次数的增加,先急剧下降,后逐渐趋于平缓。图6 表 明模型训练后可以达到对路标足够的检测精度。 法计算像素值,如图4^[24]所示,虚线框表示特征图, 实线框表示 RoI,每个子区域中包含4个采样点。



Fig. 4 RoI Align bilinear interpolation

通过特征图上距离最近的4个像素得到其像素 值,整个过程中没有对 RoI、均分 RoI 产生的子区域 和采样点进行量化,每个子区域中采样点的数目和 位置存在一定的规则,若采样点数为1,则该点位于 子区域的中心位置;若采样点数为4,则采样点的位 置为均分该子区域的4个小矩形后各自的中心点。 通常这些采样点的坐标为浮点数,所以需要用到插 值的方法获得其像素值。最后对每个单元格内的4 个采样点进行最大池化就可以得到最终的 RoI Align 的结果。RoI Align 的这一过程需进行反向传播 计算,反向传播公式为:



根据图 6 中精确度(Precision)和召回率(Recall)计算各类别的平均精度值(mean average precision, mAP)对网络进行评价。Precision, Recall 和 mAP 的计算公式分别为:

$$Precision = \frac{TP}{TP + FP} \quad , \tag{4}$$

$$Recall = \frac{TP}{TP + FN} , \qquad (5)$$

$$mAP = \frac{\sum \left(\frac{\sum Precision_c}{N_c}\right)}{N} , \qquad (6)$$

式中: *TP* 为路标被正确检测的数量; *FP* 为误检测 为路标的其他对象数量; *FN* 为错误检测的路标数 量; *Precisionc* 为类别*C* 的精确度; *Nc* 为含有类别*C* 的图片数目; *N* 为图片总数。

3.2 路标检测与识别方法比较

为验证本文方法的有效性,将修改后的 Mask R – CNN 分别与 Faster R – CNN 和 YOLOv3 方法^[30]的 实验结果进行了比较,这 3 种方法大致可分为 2 类, Faster R – CNN 与 Mask R – CNN 相似,都是基于目标候选区域的模型,而 YOLOv3 是基于回归的模型 方法。实验选取了 2 种无人机拍摄的 4 张不同环境下的图(序号分别为 a,b,c,d)进行示例分析,从表 2—4 可以看出:①Mask R – CNN 的检测与识别精度都高于其他 2 种方法,且对磨损路标(如表 4 中

a)、阴影区路标(如表4中b)以及只有箭头部分的 路标(如表4中c)都可以很好地识别,说明 Mask R – CNN 在像素级别的路标检测比采用矩形框定位对象 的 YOLOv3 和 Faster R – CNN 方法具有更高的定位 精度;②Faster R – CNN 对于磨损路标几乎无法检测 和识别,这主要是因为该模型中 CNN 和 RoI Pooling 结构对小目标的检测效果不佳;③从运行时间上可以 看出,Mask R – CNN 所用的时间都高于其他2种方法, 原因是因为 Mask R – CNN 中增加了分割分支,这使得 在整个处理过程中速度下降,由此造成时间增加。

表 2 示例图统计 Tab. 2 Sample graph statistics

1 0 1							
指标	方法	a	b	с	d		
识别数 量/个	Faster R – CNN	3	2	2	1		
	YOLOv3	5	2	2	1		
	Mask R – CNN	5	2	2	1		
平均识别 准确度/%	Faster R – CNN	60	99.5	100	100		
	YOLOv3	94.2	99.5	99	100		
	Mask R – CNN	98.3	99.9	100	100		

表3 3 种方法的检测结果

Tab. 3 Test results of three methods

方法	mAP/%	平均运行时间/s
Faster R – CNN	91.66	2.9
YOLOv3	97.84	0.3
Mask R – CNN	98.33	3.8

表4 Mask R – CNN 检测结果 Tab.4 Mask R – CNN test results

序号	Faster R – CNN	YOLOv3	Mask R – CNN
a			
Ь			
с		traffic traffic	
d		traffic	

4 结论

本文为解决智能驾驶辅助系统对路标的应用需

求,尝试利用基于像素水平检测的 Mask R - CNN 方 法对路标进行检测和识别。实验结果表明,相比于 传统方法,该方法在速度和精度上都有了大幅度提 升,并且具有较好的应用预期。通过实验主要得出 以下几点结论:

1)为满足实验需求,制作了一套高质量的无人 机影像路标数据集,以适应不同场景下对路标的检 测和识别。

2)实验过程中考虑到路标外形的特殊性,在对 Mask R - CNN 整体框架不变的情况下,对 RPN 的 锚框参数做了相应的修改,使之适合复杂环境下路 标的检测与识别。因为文中 Faster R - CNN 和 YOLOv3 方法都是对路标进行识别,所以本文方法 在示例图中隐藏了分割部分,也只显示识别部分。

3) 从测试结果中可以看出, Mask R - CNN 的 mAP 为98.33%, 都高于其他2种方法, 但速度上略 显落后, 这主要是因为增加了分割分支。后续将从 网络的整体结构入手, 寻找更好的方法对网络进行 改进, 在提高精度的同时着重提升路标的检测与识 別速度, 使之更好地应用于智慧交通。

参考文献(References):

- [1] Chira I M, Chibulcutean A, Danescu R G. Real time detection of road markings for driving assistance applications [C]// International Conference on Computer Engineering and Systems. IEEE, 2010:158 – 163.
- [2] Schreiber M, Poggenhans F, Stiller C. Detecting symbols on road surface for mapping and localization using OCR [C]//IEEE International Conference on Intelligent Transportation Systems. IEEE, 2014:597-602.
- [3] Wu T, Ranganathan A. A practical system for road marking detection and recognition [C]//Intelligent Vehicles Symposium. IEEE, 2012,7(2272):25 - 30.
- [4] Maier G, Pangerl S, Schindler A. Real time detection and classification of arrow markings using curve – based prototype fitting [C]//Intelligent Vehicles Symposium. IEEE, 2011:442 – 447.
- [5] Ding D, Yoo J, Jekyo J, et al. Efficient road sign detection based on machine learning [J]. Bulletin of Networking Computing Systems and Software, 2015, 4(1):15 – 17.
- [6] 胡忠闯. 基于深度学习的车道线检测算法研究[D]. 杭州:浙 江大学,2018.

Hu Z C. Deep learning based lane markings detection algorithm [D]. Hangzhou: Zhejiang University, 2018.

 [7] 任洪梅.基于深度学习的路面交通标志识别[J].信息通信, 2017(4):26-27.

Ren H M. Road traffic sign recognition based on deep learning [J]. Information and Communications ,2017(4):26-27.

[8] 张永涛. 基于机器视觉的路面交通标志识别的应用研究[D]. 广州:广东工业大学,2015.

Zhang Y T. Research on the application of pavement traffic signs recognition based on machine vision [D]. Guangzhou: Guangdong University of Technology, 2015.

[9] Danescu R, Nedevschi S. Detection and classification of painted road objects for intersection assistance applications [C]//International IEEE Conference on Intelligent Transportation Systems. IEEE, 2010:433-438.

- [10] Wang N, Liu W, Zhang C M, et al. The detection and recognition of arrow markings recognition based on monocular vision [C]//IEEE International Symposium on Industrial Electronics. IEEE, 2009: 4380-4386.
- [11] Liu W, Lyu J, Yu B, et al. Multi type road marking recognition using Adaboost detection and extreme learning machine classification[C]//Intelligent Vehicles Symposium (IV). IEEE, 2015:41 – 46.
- [12] 李强.基于多层次融合与卷积神经网络的道路交通标线检测 与识别[D].西安:长安大学,2018.
 Li Q. Road traffic marking and recognition based on multi - level fusion and convolutional neural networks [D]. Xi'an; Chang'an University,2018.
- [13] Husan V, Hyung H, Jin K, et al. Recognition of damaged arrow road markings by visible light camera sensor based on convolutional neural network[J]. Sensors, 2016, 16(12):2160.
- [14] Yang T, Long X, Sangaiah A K, et al. Deep detection network for real - life traffic sign in vehicular networks [J]. Computer Networks, 2018, 136:95 - 104.
- [15] 谢 锦,蔡自兴,邓海涛,等. 基于图像不变特征深度学习的交通标志分类[J]. 计算机辅助设计与图形学学报,2017,29(4):
 632-640.
 Xie J,Cai Z X,Deng H T, et al. Traffic sign classification based on

deep learning of image invariant feature [J]. Journal of Computer – Aided Design and Computer Graphics ,2017 ,29(4) :632 - 640.

- [16] Balducci F, Impedovo D, Pirlo G. Detection and validation of tow away road sign licenses through deep learning methods [J]. Sensors, 2018, 18(12):4147.
- [17] Filatov D M, Ignatiev K V, Serykh E V. Neural network system of traffic signs recognition [C]//IEEE International Conference on Soft Computing and Measurements (SCM). IEEE, 2017:1488 – 1491.
- [18] 杨梦梦. 基于 Faster R CNN 的交通标志检测方法[D]. 武汉: 湖北工业大学,2019.

Yang M M. Method for traffic sign detection based on Faster R – CNN[D]. Wuhan: Hubei University of Technology, 2019.

- [19] 何洪亮. 基于卷积神经网络的交通标志检测研究[D]. 南京: 南京理工大学,2018.
 He H L. Traffic sign detection based on convolutional neural network[D]. Nanjing: Nanjing University of Science and Technology, 2018.
- [20] Zhu Y, Zhang C, Zhou D, et al. Traffic sign detection and recognition using fully convolutional network guided proposals [J]. Neurocomputing, 2016, 214:758-766.
- [21] 伍锡如,雪刚刚. 基于图像聚类的交通标志 CNN 快速识别算法[J]. 智能系统学报,2019,14(4):670-678.
 Wu X R, Xue G G. CNN based image clustering algorithm for fast recognition of traffic signs [J]. CAAI Transactions on Intelligent Systems,2019,14(4):670-678.
- [22] 卢飞宇.基于改进的卷积神经网络的交通标志检测与识别算法[J].工业控制计算机,2019,32(6):54-56.
 Lu F Y. Traffic sign detection and recognition algorithm based on improved multi task convolutional neural network[J]. Industrial

Control Computer, 2019, 32(6):54 - 56.

- [23] Kryvinska N, Poniszewska Maranda A, Greguš M. An approach towards service system building for road traffic signs detection and recognition [J]. Procedia Computer Science, 2018, 141:64 – 71.
- [24] He K M, Georgia G, Piotr D, et al. Mask R CNN[C]//Proceedings of the IEEE Computer Society Conference on Computer Vision, 2017;2961 - 2969.
- [25] Ren S, He K M, Girshick R, et al. Faster R CNN: Towards real time object detection with region proposal networks [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 39(6):1137 - 1149.
- [26] Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation [C]//Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2015:3431-3440.

- [27] Lin T Y, Piotr D, Ross G, et al. Feature pyramid networks for object detection [C]//Proceedings of the IEEE Computer Society Conference on Computer Vision. IEEE, 2017;2117-2125.
- [28] 石杰,周亚丽,张奇志.基于改进Mask R CNN和 Kinect的服务机器人物品识别系统[J]. 仪器仪表学报,2019,40(4):216-228.
 Shi J,Zhou Y L,Zhang Q Z. Service robot item recognition system based on improved Mask R CNN and Kinect[J]. Chinese Journal of Scientific Instrument,2019,40(4):216-228.
- [29] He K M, Zhang X Y, Ren S Q, et al. Spatial pyramid pooling in deep convolutional networks for visual recognition [C]//European Conference on Computer Vision. Springer, 2014;346 - 361.
- [30] Redmon J, Farhadi A. YOLOv3: An incremental improvement [EB/OL]. (2018 04 08) [2020 01 07]. https://arxiv.org/abs/1804.02767.

Detection and recognition of road traffic signs in UAV images based on Mask R – CNN

CHEN Pengdi¹, HUANG Liang^{1,2}, XIA Yan¹, YU Xiaona³, GAO Xiaxia¹

(1. Faculty of Land Resource Engineering, Kunming University of Science and Technology, Kunming 650093, China;

 Surveying and Mapping Geo – Informatics Technology Research Center on Plateau Mountains of Yunnan Higher Education, Kunming 650093, China; 3. Kunming Vocational and Technical College of Industry, Kunming 650093, China)

Abstract: The detection and recognition of traffic signs is an important part of the intelligent driving navigation system. However, due to the shortcomings of low accuracy, high time complexity and poor robustness, the traditional method cannot meet the current needs of intelligent driving. Therefore, a method for detecting and recognizing road traffic signs of UAV images based on Mask R – CNN is proposed. Firstly, a set of high – quality UAV images road traffic sign data sets are produced. Then, based on the statistics of 200 labeled landmarks features, the region proposal network (RPN) structure anchor boxes width – to – height ratio and initial parameters in Mask R – CNN are improved to make it better applied to UAV images road sign scenes. Finally, the precision – recall (PR) curve and mean average precision (mAP) are used for accuracy evaluation. The experimental results show that the anchor boxes width – to – height ratio is better when the ratio is 1:1, 1:2, 1:3; and that the average detection accuracy obtained by this method is 98.33%, which is higher than the accuracy of Faster R – CNN and YOLOv3, indicating better effectiveness.

Keywords: UAV images; traffic sign detection; traffic sign recognition; Mask R - CNN; RPN; anchor boxes (责任编辑:张 仙)