

基于改进 Yolo v3 算法的遥感建筑物检测研究

董彪¹, 熊风光¹, 韩燮¹, 况立群¹, 徐清宇²

1. 中北大学 大数据学院, 太原 030051

2. 北方自动控制技术研究所 仿真装备研发部, 太原 030051

摘要:针对遥感图像中建筑物检测存在小型建筑物检测难度大、检测过程中无法满足实时性等问题, 提出将基于深度学习的目标检测算法 Yolo v3 应用到建筑物检测场景中。以实时性及泛用性良好的 Yolo v3 为基本算法, 满足实时性的要求; 通过改进 Yolo v3 的网络结构, 以修改特征图分辨率、调整先验框维度为方向加强对小型建筑物的检测能力。实验结果表明, 改进的 Yolo v3 目标检测算法既满足了实时性的要求, 且检测精度和召回率达到了 91.29% 和 95.61%, 较原算法分别提高了 5.35% 和 2.34%。因此提出的改进方法有效解决了遥感领域小型建筑物的检测问题。

关键词:目标检测; 遥感; 深度学习; Yolo v3; 图像处理

文献标志码:A **中图分类号:**TP391 **doi:**10.3778/j.issn.1002-8331.2004-0410

董彪, 熊风光, 韩燮, 等. 基于改进 Yolo v3 算法的遥感建筑物检测研究. 计算机工程与应用, 2020, 56(18): 209-213.

DONG Biao, XIONG Fengguang, HAN Xie, et al. Research on remote sensing building detection based on improved Yolo v3 algorithm. Computer Engineering and Applications, 2020, 56(18): 209-213.

Research on Remote Sensing Building Detection Based on Improved Yolo v3 Algorithm

DONG Biao¹, XIONG Fengguang¹, HAN Xie¹, KUANG Liqueun¹, XU Qingyu²

1. College of Big Data, North University of China, Taiyuan 030051, China

2. Simulation Equipment R&D Department, North Automatic Control Technology Institute, Taiyuan 030051, China

Abstract: In order to solve the problems in building detection for remote sensing images, such as the difficulty of detecting small buildings and the lack of real-time performance in the detection process, the deep learn-based target detection algorithm Yolo v3 is applied to the building detection scene. The basic algorithm of Yolo v3 with good real-time and universality is adopted to meet the real-time requirements. By improving the network structure of Yolo v3, the detection ability of small buildings is strengthened by modifying the resolution of feature map and adjusting the dimension of prior frame. The experimental results show that the improved Yolo v3 target detection method meets the real-time requirements and the detection accuracy and recall rate reaches 91.29% and 95.61%, which are 5.35% and 2.34% higher than the original algorithm. Therefore, the improved method effectively solves the detection problem of small buildings in remote sensing field.

Key words: target detection; remote sensing; deep learning; Yolo v3; image processing

1 引言

随着遥感技术的发展, 遥感图像幅员更加辽阔, 包含的信息更加复杂, 因此其已被广泛应用于图像识别、语义识别、视频分析、文本分析和大数据分析等领域^[1]。遥感图像中对固定地物与移动地物的检测一直是研究

热点, 其中固定地表建筑物是人类主要活动场所, 因此对该类目标实现检测显得尤为重要。考虑到实际工程应用中实时性是一个极为重要的问题, 且检测目标(建筑物)小, 对建筑物检测也造成了极大的困难。故而, 如何实现对该类建筑物的检测是目标检测领

基金项目:国家重点研发计划(No.2018YFB2101504); 国家部委基金; 山西省重点研发计划(No.201803D121081, No.201903D121147); 山西省自然科学基金(No.201901D111150)。

作者简介:董彪(1995—), 男, 硕士研究生, 研究方向为仿真与可视化; 熊风光(1979—), 男, 博士, 副教授, 研究方向为软件技术、虚拟现实; 韩燮(1964—), 女, 博士, 教授, 研究方向为仿真与可视化、智能信息处理; 况立群(1976—), 男, 博士研究生, 副教授, 研究方向为仿真与可视化; 徐清宇(1976—), 研究员级高级工程师, 研究方向为武器装备仿真。

收稿日期:2020-04-28 **修回日期:**2020-06-12 **文章编号:**1002-8331(2020)18-0209-05

CNKI 网络出版:2020-06-28, <https://kns.cnki.net/kcms/detail/11.2127.TP.20200624.1708.006.html>

域的一个研究难点^[2]。

相比传统的根据建筑物颜色、形状、纹理等提取简单特征的目标检测算法,基于深度学习的目标检测利用深度神经网络强大的特征学习能力来解决复杂的遥感图像检测问题,更能满足今天遥感图像中目标检测的要求^[3]。本文的研究重点为固定地表中的建筑物检测,以基于深度学习的目标检测算法 Yolo v3 为研究基础,并以修改特征图分辨率、调整先验框维度为改进方向,加强对小型建筑物的检测,生成更适合遥感建筑物检测的网络模型。

2 相关工作

从遥感图像中检测建筑物有传统算法和基于深度学习的算法两大类。传统目标检测算法是一种基于人工设计特征的分类方法,利用了建筑物某一方面的特征,如颜色、形状等,虽然具有一定的检测效果,但是在检测精度和检测速度上具有较大的局限性。如文献[4]提出一种阴影验证的建筑物检测方法,该方法通过建筑物的阴影、形状以及颜色特征结合起来检测建筑物,对特征简单的建筑物检测效果良好,但对于复杂特征的建筑物检测,该方法无法满足工程应用中对检测精度及实时性的要求。与之相比,基于深度学习的检测算法以其适应能力强、检测效率高而得到了广泛应用,该类检测算法以卷积神经网络为基础,具有强大的特征表示能力,在背景模糊、地理环境复杂的情况下,仍然能够学习到有用的特征,大幅提高了物体检测精度,适用于当今遥感图像复杂多变的场景,具有非常好的泛化能力和鲁棒性^[5]。如文献[6]提出一种端到端的密集反卷积神经网络,将图像中的抽象特征与细节特征结合,能够捕捉到目标建筑物的细微复杂特征;文献[7]中使用金字塔技术构建多级训练的样本,提出候选建筑区域检测建筑物,在检测精度及实时性方面有显著提升。

目前基于深度学习的检测算法可以分为两类:基于区域的目标检测算法和基于端到端的目标检测算法。前者先由算法生成一系列作为样本的候选框,再通过卷积神经网络进行样本分类,这类算法在检测精度和定位精度上占优,代表算法有:Faster RCNN^[8-9]和 Fast RCNN等;后者不生成候选框,直接将目标边框定位的问题转化为回归问题,这类算法在检测速度上占优,代表算法有:Yolo^[10-11]系列和 SSD^[12]等。

考虑到目前遥感技术的发展,需要对上千平方公里图像进行处理,而且目标小、背景复杂度高,因此检测速度及泛用性是首先考虑的问题,所以本文将基于深度学习的 Yolo v3 检测算法应用于建筑物的检测,并在检测小型建筑物方面提出改进。

3 改进 Yolo v3 算法

3.1 Yolo v3 原理

Yolo v3 是基于端到端的检测算法,网络结构分为骨干网络(Darknet-53)和检测网络。Darknet-53 由 52 个卷积层和 1 个全连接层组成,构成了一个 53 层的卷积神经网络(CNN),并输出 13×13 、 26×26 和 52×52 三种尺度的特征,送入检测网络。检测网络对三种尺度的特征回归,预测出多个预测框,并使用非极大抑制(NMS)算法删除置信得分较低的预测框,保留置信得分较高的预测框为目标检测框,从而得到目标的类别和位置,检测流程如图 1。

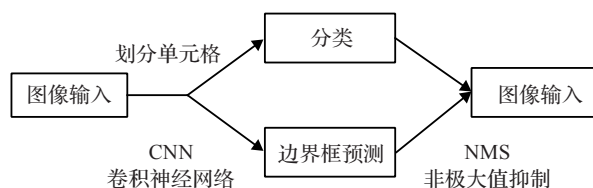


图1 Yolo v3 检测流程

(1)边界框预测:检测网络对 13×13 、 26×26 和 52×52 三种尺度的特征回归,预测出多个预测框,在检测的过程中,每个网格会产生多个边界框,每个边界框都要预测边界框位置信息 (x, y, w, h) 和置信度(confidence)。置信度是边界框所含目标的可能性 $\Pr(object)$ 与边界框准确度 IOU 的乘积,计算公式如式(1)所示:

$$confidence = \Pr(object) \times IOU_{pred}^{truth} \quad (1)$$

若网格中包含某个目标则 $\Pr(object)=1$, 否则 $\Pr(object)=0$; IOU_{pred}^{truth} 表示真实标注框与预测框的交并比,用来衡量预测边界框的准确性,当 $IOU=1$ 时,说明真实标注框与预测框重合。

(2)分类:对包含目标的网格预测多个类别概率,用 $\Pr(class_i|object)$ 表示。为了得到每个边界框的分类信息,将边界框所属网格的类别概率与边界框的置信度相乘得到置信度评分,如公式(2)所示,其代表了边界框中目标属于某个类的概率。

$$\Pr(class_i) \times IOU_{pred}^{truth} = \Pr(class_i|object) \times \Pr(object) \times IOU_{pred}^{truth} \quad (2)$$

在模型训练中,将边界框每个类别的置信度评分按高低排序,将置信度得分小于设定阈值的边界框的置信度设为 0,然后再用非极大抑制算法(NMS)去掉重复率较大的边界框,最后每一个边界框只保留置信度评分大于 0 且最高的类,从而检测出各类目标。

3.2 改进 Yolo v3

Yolo v3 算法的网络权重是基于 COCO 数据集训练测试所得,COCO 数据集包括 80 种类别目标,且该数据集目标与本文检测的遥感建筑物尺度差异较大,故而原 Yolo v3 算法不适用本文应用场景,尤其是对小型建筑物的检测表现不好。

目前随着 Yolo v3 的广泛应用,已有相关文献针对具体场景做出了相应改进。文献[13]提出一种基于 Yolo v2 的船舶检测算法 Yolt,该算法网络结构与 Yolo v2 相比,最大的修改是最后输出特征尺寸为 26×26 ,加强了对小目标的检测,结果表明在稀疏的小目标场景下检测效果良好,但在目标密集的场景下表现不好^[13];文献[14]提出了一种基于 Yolo v3 改进的电气元件检测算法,通过对 PASCAL VOC 和 KITTI 数据集进行 k -means 聚类分析,调整先验框维度,并在原来三种尺度的特征图上增加 104×104 的特征图,达到加强小目标检测的能力,结果表明在文献提供的电气元件的数据集中表现良好,但该应用场景不具有泛用性^[14]。

故本文在以下两方面进行改进:一是修改特征图分辨率,扩大感受野,减少语义特征丢失;二是使用 k -means 算法对本文的遥感建筑物数据集进行聚类分析,得到对应的先验框维度^[15]。

3.2.1 修改特征图分辨率

Yolo v3 为加强对小目标的检测,借鉴了特征金字塔网络 FPN(Feature Pyramid Network),将高层特征与浅层特征信息融合,采用多个尺度融合的方式,在多个尺度的特征图上进行位置和类别预测,但 Yolo v3 网络结构采取的三个尺度的特征融合方法对遥感图像中检测较小的目标有不利的影响, 13×13 的特征图语义丢失严重,易造成小目标丢失。考虑到特征的分辨率将直接影响小目标的检测和整体性能指标,所以在 Darknet-53 基础上对原特征图三个尺度的分辨率 13×13 、 26×26 和 52×52 修改为两个较大尺度的分辨率 26×26 和 52×52 。

修改后的网络模型结构如图 2 所示,将原网络结构 26×26 的特征图分辨率作为第一尺度;将 L61(61 层)结果进行五次卷积操作(Convolutional Set),首先为了提升计算效率进行 1×1 卷积操作降维,然后上采样(Up Sample),再与 L36(36 层)融合,最后融合之后采用 3×3 的卷积核对融合结果进行卷积,目的是消除上采样的混

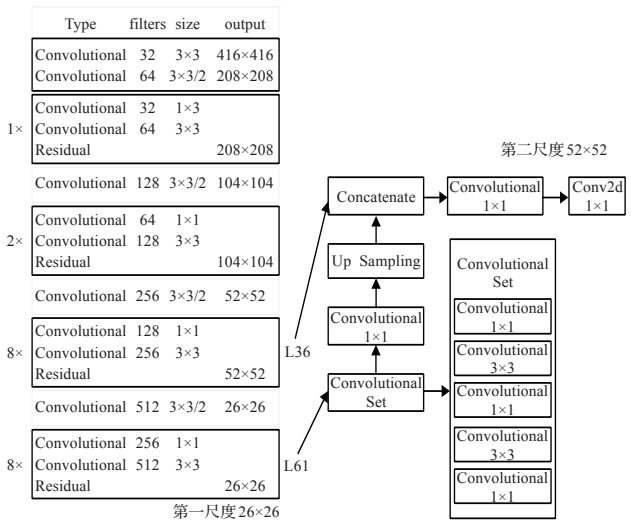


图2 修改后的网络结构

叠效应,如此就得到了一个新的特征图 52×52 ,作为第二尺度特征。

经修改后,本文对 Yolo v3 改进的网络结构能够在深层网络中保持较高分辨率和较大的特征图的感受野,以增强对小目标的检测能力。

3.2.2 k -means 聚类分析调整先验框维度

Yolo v3 借鉴先验框来预测边界框坐标,使用 k -means 算法预测得到 9 个先验框,并将其均分到 3 个尺度特征图,以此来获得更多的目标边缘信息。Yolo v3 中计算得到的 9 组先验框维度分别为 $(10, 13)$, $(16, 30)$, $(33, 23)$, $(30, 61)$, $(62, 45)$, $(59, 119)$, $(116, 90)$, $(156, 198)$, $(373, 326)$,但是在本文应用场景中,Yolo v3 算法中计算得到的先验框维度不适合建筑物的检测。因此,在遥感建筑物检测场景下,使用 k -means 算法对自制的建筑物数据集进行聚类分析,得到先验框维度,从小到大分别为: $(49, 47)$, $(92, 50)$, $(52, 97)$, $(175, 51)$, $(142, 95)$, $(74, 200)$,分辨率较小的 26×26 特征图有较大的感受野,故采用较大的先验框 $(175, 51)$, $(142, 95)$, $(74, 200)$,分辨率较大的 52×52 的特征图有较小的感受野,故采用较小的先验框 $(49, 47)$, $(92, 50)$, $(52, 97)$,如表 1 所示。

表1 数据集的先验框

特征图	先验框
26×26	(175, 51) (142, 95) (74, 200)
52×52	(49, 47) (92, 50) (52, 97)

4 实验结果及分析

4.1 数据集制作

本文使用的数据集来自我国北方某城市的卫星图像,使用谷歌地图下载,图像尺寸为 $25\text{ km}\times 25\text{ km}$,对图像裁剪及筛选,使用标注工具 labellmg 进行标注,制作本文的数据集。

labellmg 是专门为 Yolo 系列算法的网络训练准备数据的工具,利用该标注工具,手动完成矩形框标注,最后得到 xml 文件,操作界面如图 3 所示。一共得到 1 000 张训练集和 200 张测试集。

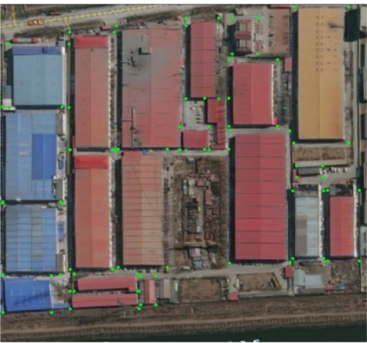


图3 标注界面

4.2 实验过程及平台

本文采用计算机为笔记本电脑,配置为 Intel i7-

9750H, 32 GB 内存, 64 位 Windows 10 操作系统; 开发环境为 tensorflow1.14, opencv4.1, python3.7。

原 Yolo v3 及本文改进算法都是在 keras 环境下运行, 以 1 000 张遥感图像为训练集, 200 张做测试集, 初始阶段学习率为 0.001, 衰减系数为 0.000 5, 为使损失函数进一步收敛, 在训练迭代次数为 25 000 次时, 学习率降低为 0.000 1。最终, loss 值稳定在 0.5 左右, 迭代至 30 000 次停止, 如图 4。

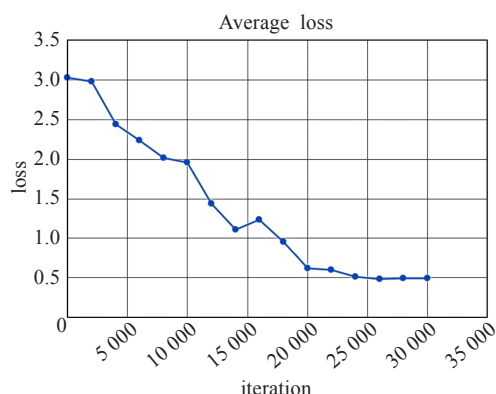


图4 损失值曲线

4.3 实验结果与分析

4.3.1 本文网络模型检测结果

本文网络模型基于 Yolo v3 算法, 通过修改特征图分辨率及使用 k -means 算法对本文遥感建筑物数据集进行聚类分析, 加强对小型建筑物的检测。经过本文网络模型检测, 选取尺寸相差较大、排列密集的建筑物检测场景做样例分析, 结果如图 5。

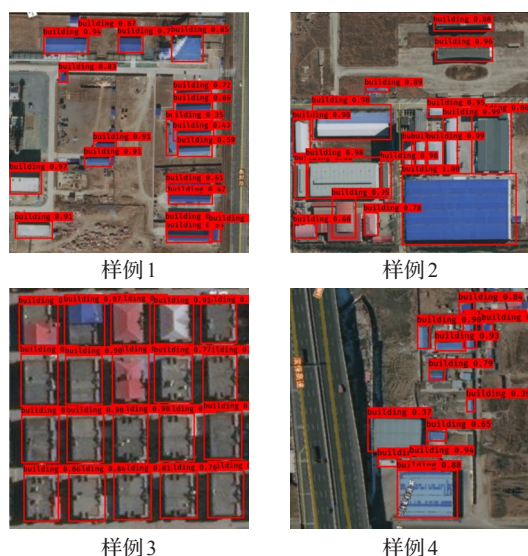


图5 网络模型检测结果

从检测结果可以看出, 本文对 Yolo v3 进行改进, 通过选择合适的特征图分辨率及调整先验框维度提高了对小型建筑物的检测能力, 对场景中排列密集建筑物(样例1)、尺度差距大(样例2)、背景复杂度高(样例3)及在有道路下的建筑物(样例4)做到了很好的识别效果, 尤其对于小型建筑物的检测效果良好。

4.3.2 评价指标

本文采用平均检测精度(mAp)、召回率(recall)、平均检测时间(time)作为检测评价指标。其中精度和召回率分别定义为:

$$Precision = \frac{TP}{FP + TP} \quad (3)$$

$$Recall = \frac{TP}{TP + FN} \quad (4)$$

以遥感建筑物为例, TP 、 FP 分别表示检测模型将建筑物检测正确与错误的数量, FN 表示误检的数量。召回率(recall)表示能够正确识别建筑物的概率; precision 表示已识别为建筑物的检测精度, 对 precision 取平均值, 以平均检测精度(mAp)作为评价指标; time 表示平均每张检测时间。

4.3.3 本文网络模型效率对比分析

本文针对遥感图像中的小型建筑物展开研究, 选取建筑物排列密集(a组)和相对尺寸差距大(b组)两种场景做对比分析, 经过本文网络模型(Article method)与原 Yolo v3 算法检测结果对比如图 6 所示。

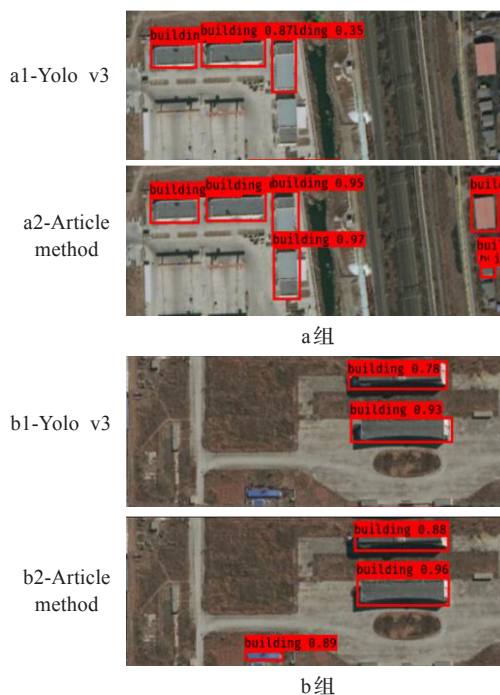


图6 检测结果对比

对比实验结果可以看出, a组的小型目标建筑物排列密集, b组目标建筑物尺度差距相对较大。原 Yolo v3 算法对小型建筑物(a1、b1)的检测效果表现不好, 在场景 a1 中右方排列密集的小型建筑物和在场景 b1 中下方的小型建筑物都出现了漏检; 本文网络结构(Article method)的检测结果(a2、b2)很好地实现了小型目标建筑物的检测, 在排列密集和相对尺度大的两种场景中都能对小型建筑物做到了很好的检测效果, 且对于相同目标的检测, 本文改进的网络结构相比原 Yolo v3 算法的检测精度也有提高。

基于深度学习的目标检测算法分为基于区域的目

标检测算法和基于端到端的目标检测算法。Yolo v3 算法是一种基于端到端的检测算法,为验证本文改进网络的优越性,故与主流算法进行对比分析。因此,本文设计实验:以本文自制遥感建筑物数据集为实验数据集,200 张作为测试集,通过本文改进网络结构(Article method)与原 Yolo v3 算法、同样基于端到端的 SSD 算法及基于区域的 Faster RCNN 算法在平均检测精度(mAp)、平均每张检测时间(time)及召回率(recall)等方面做效率比较,效率分析如表 2 所示。

表 2 效率分析

算法	mAp/%	time/s	recall/%
Yolo v3	85.94	0.041	93.27
SSD	80.49	0.253	90.84
Faster RCNN	93.81	2.370	97.33
Article method	91.29	0.042	95.61

从实验结果可以看出,在平均检测精度(mAp)方面,原 Yolo v3 算法的平均精度为 85.94%,SSD 算法的平均精度为 80.49%,Faster RCNN 算法的平均精度为 93.81%,本文改进网络结构(Article method)平均精度为 91.29%,比 Faster RCNN 算法略低,较原 Yolo v3 算法和 SSD 算法提高了 5.35%和 10.8%;而在检测时间(time)方面,原 Yolo v3 算法的检测时间为 0.041 s,SSD 算法的检测时间为 0.253 s,Faster RCNN 算法的检测时间为 2.37 s,本文改进网络结构(Article method)的检测时间为 0.042 s,仅比原 Yolo v3 算法的检测时间多 0.001 s,较 SSD 算法和 Faster RCNN 算法快了 0.211 s 和 2.328 s,很好地满足了实时性;在召回率(recall)方面,原 Yolo v3 算法的召回率为 93.27%,SSD 算法的召回率为 90.84%,Faster RCNN 算法的召回率为 97.33%,本文改进网络结构(Article method)的召回率为 95.61%,比 Faster RCNN 算法略低,较原 Yolo v3 算法及 SSD 算法提高了 2.34%和 4.77%。

故本文改进网络结构(Article method)在保持了原算法速度优势的基础上,提高了平均检测精度和召回率,所以本文提出的基于 Yolo v3 的改进网络模型对遥感建筑物检测效果更好。

5 结束语

本文针对遥感建筑物的检测,对小型建筑物检测难度大的问题进行研究,提出了基于深度学习的 Yolo v3 改进算法,以修改特征图分辨率、调整先验框维度为方向得到一种适合遥感建筑物检测的网络模型,加强了对小型建筑物的检测能力。实验结果表明,本文提出的方法在满足了实时性的基础上,有效解决了遥感图像中固定地物建筑物检测过程中存在的小型建筑物检测难度大的问题。接下来将针对遥感图像中移动地物做研究。

参考文献:

[1] Arsar G R.Advances in quantitative earth remote sensing:

past, present and future.[J].Sensors (Basel, Switzerland) , 2019,19(24).

[2] Wang Min,Chen Jinyong,Wang Gang,et al.High resolution remote sensing image ship target detection technology based on deep learning[J].Optoelectronics Letters, 2019,15(5).

[3] Chen Junde,Zhang Defu,Nanehkaran Yaser A,et al.Detection of rice plant diseases based on deep transfer learning[J].Journal of the Science of Food and Agriculture, 2020,100(7).

[4] Lin T Y,Dollae P,Girshick R,et al.Feature pyramid networks for object detection[C]//IEEE Conference on Computer Vision and Pattern Recognition,2017:936-944.

[5] Liu Qiangwei,Xiang Xiuqiao,Wang Yuanfang,et al.Aircraft detection in remote sensing image based on corner clustering and deep learning[J].Engineering Applications of Artificial Intelligence,2020,87.

[6] Dai J F,Li Y,He K M,et al.R-FCN:object detection via refion-based fully convolutional networks[C]//IEEE Conference on Computer Vision and Pattern Recognition, 2016:379-387.

[7] Bodla N,Singh B,Chellappa R,et al.Soft-NMS—improving object detection with one line of code[C]//Proceedings of the IEEE International Conference on Computer Vision,2017:5561-5569.

[8] Xie Han,Chen Yunfan,Shin H.Context-aware pedestrian detection especially for small-sized instances with Deconvolution Integrated Faster RCNN(DIF R-CNN)[J].Applied Intelligence,2019,49(3).

[9] Han Cen,Gao Guangyu,Zhang Yu.Real-time small traffic sign detection with revised faster-RCNN[J].Multimedia Tools and Applications,2019,78(10).

[10] Dow C R,Ngo H H,Lee L H,et al.A crosswalk pedestrian recognition system by using deep learning and zebra-crossing recognition techniques[J].Software: Practice and Experience,2020,50(5).

[11] Chen Binghuang,Miao Xiren.Correction to: distribution line pole detection and counting based on YOLO using UAV inspection line video[J].Journal of Electrical Engineering & Technology,2020,15(2).

[12] Auradkar P,Prashanth T,Aralihalli S,et al.Performance tuning analysis of spatial operations on Spatial Hadoop cluster with SSD[J].Procedia Computer Science,2020, 167:2253-2266.

[13] Van Etten A.You only look twice:rapid multi-scale object detection in satellite imagery[J].arXiv:1805.09512,2018.

[14] Chen Haipeng,He Zhentao,Shi Bowen,et al.Research on recognition method of electrical components based on YOLO V3[J].IEEE Access,2019:157818-157829.

[15] Tumas P,Serackis A.Automated image annotation based on YOLOv3[C]//2018 IEEE 6th Workshop on Advances in Information,Electronic and Electrical Engineering (AIEEE),2018:1-3.