

基于小数据的钢筋检测方法的研究

Research on Rebar Detection Method Based on Small Data

王志丹(中国联通智能城市研究院,北京 100048)

Wang Zhidan(China Unicom Smart City Research Institute, Beijing 100048, China)

摘要:

在工地建筑领域,钢筋已然成为不可或缺的基本材料之一,在我国的绝大多数工地,现有的钢筋计数方法是工人逐一统计已计数和未计数的钢筋,不仅耗时、耗力,而且由于工人长时间计数钢筋,人眼和人脑容易疲劳,计数误差将增加^[2]。由于工地钢筋图像数据不便于大量采集和标注,提出一种基于小数据的钢筋检测模型,将该模型应用于智慧工地中的钢筋计数应用,工人只需将拍摄的钢筋图像通过终端传入模型,便可以快速准确地得到图像中钢筋的数量。

关键词:

智慧工地;钢筋计数;计算机视觉

doi:10.12045/j.issn.1007-3043.2020.02.009

文章编号:1007-3043(2020)02-0039-06

中图分类号:TP39

文献标识码:A

开放科学(资源服务)标识码(OSID):



Abstract:

In the field of construction site, rebar has become one of the indispensable basic materials. At present, in the vast number of construction sites in China, the existing rebar counting method is that workers count the counted and uncounted rebar one by one. This process is not only time-consuming and labor-consuming, but also worker counts rebars for a long time, worker's eyes and brain are prone to fatigue, and the counting error will increase^[2]. Because the image data of rebar in construction site is not easy to collect and label, it proposes a rebar detection model based on small data. This model is applied to the rebar counting application in intelligent construction site. Workers only need to upload the captured rebar image into the model through the terminal, then they can quickly and accurately get the number of rebar in the image.

Keywords:

Smart construction site; Rebar counting; Computer vision

引用格式:王志丹. 基于小数据的钢筋检测方法的研究[J]. 邮电设计技术, 2020(2):39-44.

0 引言

近年来,国家大力推动钢结构建筑发展,数据表明,2017年以来中国钢筋年产量已超过2亿吨,2017年建筑业的总产值超过21万亿元,而其中钢铁在建筑材料中所占的造价达到了20%~30%^[3]。钢筋广泛应用于各种建筑结构,特别是大型、重型以及高层建筑结构,但在实际运用中,钢筋计数问题一直困扰着从业者。自上世纪70年代便有好多工程人员投入至钢筋计数的研究中,对于成捆钢筋,曾有研究利用重量

传感器的方法来进行计数,该研究理论为同类型钢筋有着相同的重量和长度的标准,根据总重量和单根钢筋的重量便可以得出一捆钢筋的数量。但是由于不同钢材厂的工艺不同,以及部分钢材厂存在负公差生产,故重量传感器方法无法得出精准的结果。随着计算机视觉技术的发展,有研究将神经网络应用于钢筋计数方法,该研究认为虽然钢筋端面形状不规则,但是面积是相近的,故可以通过将一捆钢筋端面图片进行二值化,通过判断每一个端面的面积与平均值的差异便可以得出成捆钢筋的数量^[4]。但该方法存在一定缺陷,钢筋端面平均值和目标物体有较大的依赖性,自适应性较差,同时,若多个钢筋的端面大面积重叠,

收稿日期:2020-01-13

也无法判断具体数量。随着双目视觉和激光技术的发展,一些大型钢材制造商将“双目视觉+线激光”技术应用于钢筋计数,该技术利用激光测距的测量结果对双目视觉的测量结果进行修正,有效地提高双目立体视觉远距离测量精度^[5],但是该技术成本高,需要使用双目立体视觉仪和激光测距仪等大型设备,在工地这种特殊而复杂的场景中无法便利地使用。为解决以上问题,同时工地钢筋图像数据不便于大量采集和标注,本文提出基于小数据目标检测算法并将其应用于钢筋计数应用中,工地工作人员只需将钢材照片通过终端上传至云服务器中的钢筋检测模型,便可以快速准确地得到照片中钢筋的数量。

1 目标检测算法

目前,深度学习已经应用到了各个领域,应用场景主要为3类:物体识别、目标检测、自然语言处理。其中目标检测作为计算机视觉的重要任务之一,不仅要识别出物体属于哪个分类,还要得到物体在图片中的位置。近年来随着深度学习技术和算力的高速发展,目标检测算法也从基于手工特征的传统算法转向了基于深度神经网络的检测技术,深度神经网络的目标检测算法主要分为2类:one-stage和two-stage,one-stage算法主要包含SSD^[14]、YOLO^[11]等算法,two-stage算法主要包含R-CNN^[13]、Fast R-CNN^[12]、Faster R-CNN^[6]等算法,两者的主要区别在于two-stage算法需要先生成可能包含物体的预选框,其次再进行物体检测;而one-stage算法会直接在网络中提取特征来预测物体的分类和位置,one-stage相关的算法计算速度比较快,但是准确率略低于two-stage算法。

本文使用的算法主要为two-stage中的Faster R-CNN算法,算法模型结构如图1所示,Faster R-CNN算法实现图像中的目标检测需经过如下4个步骤:第1步将预训练卷积神经网络作为特征提取器,对输入的图片进行特征提取得到图像的特征图;第2步利用RPN(Region Propose Network)网络对提取的卷积特征

图进行处理,得到一定数量的包含目标物体的候选区域;第3步通过ROI Pooling对卷积神经网络提取的特征和包含目标物体的边界框进行处理,同时提取每个候选区域的特征,得到一个新的向量;第4步基于R-CNN模块对前3步得到的特征进行分类和边框回归。在深度目标检测算法中,最难的问题是生成长度不定的边界框列表,边界框是具有不同尺寸和长宽比的矩形,但是构建深度神经网络时,最后的网络输出一般是固定尺寸的张量输出,Faster R-CNN网络在第2步的RPN中,通过采用锚(Anchors)来解决边界框列表长度不定的问题,Anchors是固定尺寸的边界框,即在原始图像中统一放置固定大小的参考边界框,不同于直接检测目标位置,这里将问题转化为求解每个Anchor是否包含目标和如何调整Anchor以更好地拟合目标对象^[15],故RPN网络最终为每个Anchor提供2个输出结果:一是Anchor目标物体的概率和背景的概率,二是4个预测值(Δx_{center} 、 Δy_{center} 、 $\Delta width$ 、 $\Delta height$),用来调整Anchor的边界框,使其更好地拟合预测的目标对象。RPN处理后,将得到约2000个候选区域,若直接进行分类和回归计算,则会出现效率低、速度慢等问题,Faster R-CNN采用RoI Pooling即对每个候选区域提取固定尺寸的特征图,继而对固定尺寸的特征图进行分类和回归,这样便可以提高计算的效率。以上便是Faster R-CNN算法实现图像目标检测的流程。

2 基于小数据的目标检测模型

本文通过将优化后的目标检测算法模型应用于钢筋识别应用中,在工地复杂环境中,工人只需将拍摄后的钢筋照片通过终端上传至模型,便可以快速准确地得出照片中钢筋的数量。为了训练出钢筋检测模型,本文首先收集了约250张成捆钢筋图像,钢筋图像如图2所示,每张图像的标记为 $[N, 4]$,其中 N 为一张图像中的目标框的数量,4维数据分别为 $[X_1, Y_1, W, H]$, X_1 和 Y_1 分别为目标中心点的 (X, Y) 坐标, W 和 H 分

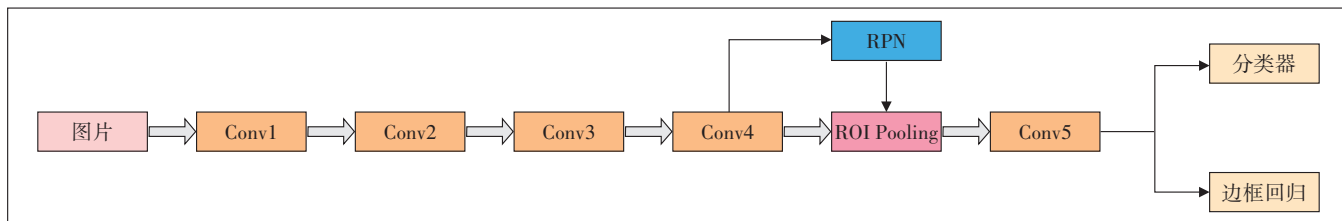


图1 Faster R-CNN结构图



图2 钢筋图像

别为目标的宽度和高度值,将这250张图像随机分为训练集(200张图像)和测试集(50张图像),训练集图像用于训练钢筋检测模型,测试集图像用于检测模型的准确率。由于工地环境复杂,钢筋图像采集较困难,同时单个钢筋目标较小,标注需要大量的人力,故很难达到目标检测算法所需的图像数量级(通常为100 000数量级),为了在小数据集的图像上得到更高的准确率,本文做了2处优化,一是从数据层面对训练集图像进行数据增强,二是利用在ImageNet数据集上的预训练模型对钢筋检测模型进行微调。

数据增强是一种通过让有限的数据产生更多等价数据的人工扩展训练数据集的技术,它是克服训练数据集不足的有效手段,目前在深度学习的各个领域应用广泛,对于图像分类问题,已有大量实验通过对图像进行剪裁、平移等操作来增强数据集,实验结果表明数据增强可以极大的改进图像分类的表现。而对于图像目标检测问题,若对原始图像进行随机剪裁和平移操作,对应图像内目标的位置坐标信息也会发生变化,Quoc Le等研究者表示^[8],尽管数据增强能极大的改进图像分类表现,但它在目标检测任务上的效果还未被透彻研究过,此外,目标检测所用图像的注释会造成大量的成本,所以数据增强对此计算机视觉任务的影响可能会更大。在此研究中,作者们研究了数据增强在目标检测上的影响,在COCO数据集上的实验结果说明,优化后的数据增强策略将检测准确率提升了2.3 mAP,使单推理模型能够达到最佳效果——50.7 mAP。

计算机视觉中的图像增强包括剪裁、翻转、旋转

一定角度、扭曲等几何变换以及像素扰动、添加噪声、光照调节、对比度调节、样本加和或插值、分割补丁、基于GAN网络^[9]的图像生成等的数据增强方法。但是并不是所有的数据增强方法都适合于数据集不足的应用,对于不同的应用场景需要具体分析,数据增强要保证原始数据、增强后的数据以及测试数据满足同一分布。本文的钢筋计数应用中的单个钢筋具有旋转不变性,故可以使用几何变换的数据增强方法,同时,考虑由于实际应用场景受光照、拍照角度等因素的影响,为了让模型训练学习的更全面,我们还使用了噪声扰动的数据增强方法。对钢筋图像训练集使用旋转和随机剪裁增强方法强制目标检测模型学习这些变换方式,一方面使模型学习的数据更全面,同时数据增强起到了正则化的作用,这样可以减小模型的结构风险,提高模型的鲁棒性,防止发生过拟合。本文使用imgaug方法对钢筋图像做了旋转、随机裁剪以及噪声扰动这3种方式的数据增强,增强后训练集约有10 000张图像,imgaug方法支持图像中关键点和目标边框一起变换,故对图像进行旋转和剪裁操作图像中目标的坐标也会随之发生变化,减少了大量的人工标注的工作量。图3为随机剪裁增强后的钢筋图像,图4为扰动增强后钢筋图像。

本文使用了在ImageNet数据集上的预训练模型对钢筋检测模型进行微调。ImageNet数据集由李飞飞教授团队在2009年发布,包含了超过2万类物体,共计1 400多万张图片,为整个人工智能领域奠定数据



图3 随机剪裁后的钢筋图像



图4 噪声扰动后的钢筋图像

基础,诸多计算机视觉任务都是在 ImageNet 数据集上进行预训练,将预训练得到的参数应用在相应的目标任务上作为初始参数,在新模型训练过程中慢慢调整参数,以取得较好的效果。卷积神经网络底层提取的图像特征更为一般化,包括边缘特征、纹理特征等特征,而顶层才会提取到图像特有的高级特征,在目标检测算法中加入微调,即使算法中的卷积神经网络只需侧重于提取高级特征,这样效果会比原始的网络提升很多。本文使用了同类深度学习模型中精度较高的深度残差网络 ResNet101 的预训练模型^[7], ResNet101 包含 100 个卷积层,一个全局平均值池化层和一个 1 000 维的全连接层,去掉最后的全局平均值池化层和全连接层,最后加上一个 1 024 维的 1×1 卷积层用以降维。

本文提出的基于小数据的目标检测系统架构如图 5 所示,其中底层为基础的数据层,数据层主要是产生高质量的数据以供服务层的算法模型使用,数据对于一个算法模型至关重要,需要为算法提供足够多、高质量的数据,除先验知识外,算法能达到的最佳性能受限于数据所提供的有用信息容量,当算法性能接近某个极限时,不管如何改进算法,基本没有意义,但是通过增加有用的数据集却可以提高算法准确性。架构的第二层为服务层,在服务层封装基于小数据的目标检测算法,对外提供调用接口。第三层为接口

层,接口层介于应用层与服务层之间,应用层通过接口层的接口服务模块调用钢筋识别能力服务,接口层兼有任务分配和记录功能,支持 HTTP 协议。顶层为应用层,对接终端,用户通过发送请求便可以快捷地使用算法模型,图 6 为终端调用钢筋识别算法进行钢筋计数,工地用户只需将待计数的成捆钢筋进行拍照,将拍摄后的照片上传即可快速地得到图片中钢筋的数量,该方法相比于传统的人工计数更准确、更便捷。

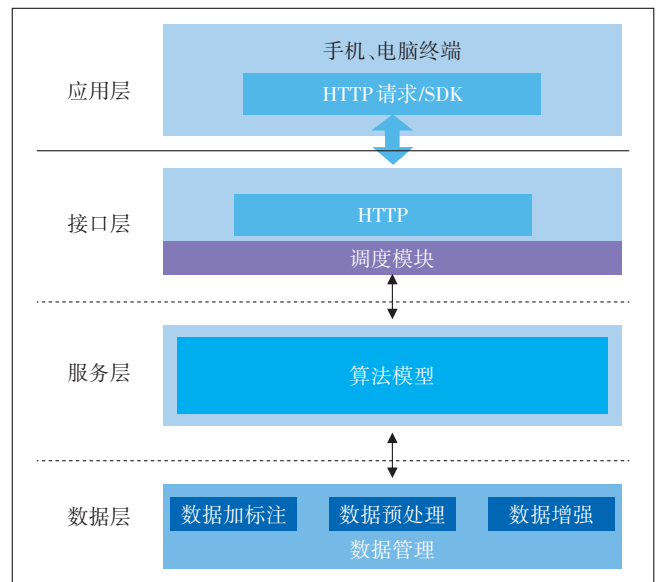


图5 基于小数据目标检测算法的架构

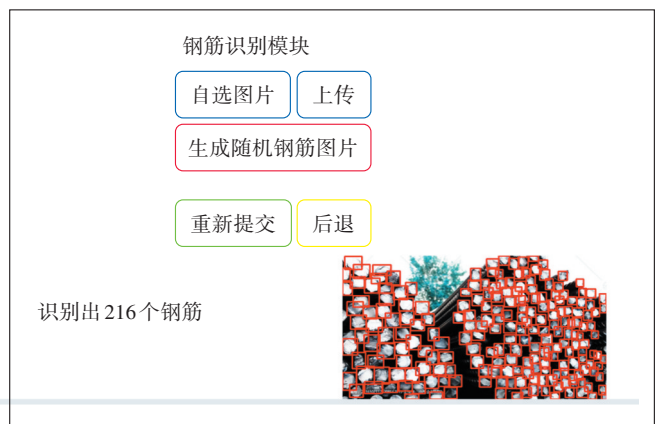


图6 终端调用钢筋识别算法进行钢筋计数

3 实验

本文的实验在 NVIDIA TITAN XP 12G 显卡、64G 内存、Ubuntu 操作系统的计算机上运行,实验中使用的编程语言为 python3,深度学习框架为 PyTorch,Py-

Torch是一个实现机器学习算法的接口,同时也是执行机器学习算法的框架,具有简洁、快速、易用的特点。本文使用的250张原始数据集为网上公开的钢筋图像。

在目标检测问题中,目标的分类和定位都需要进行评估,每张图像都可能具有不同类别的不同目标,因此,在图像分类问题中所使用的标准度量不能直接应用于目标检测问题,目标检测问题中常用的度量指标为均值平均精度(MAP——Mean Average Precision),即将所有类别的平均精度求和除以所有类别,如式(1)所示。

$$\text{MeanAveragePrecision} = \frac{\sum \text{AveragePrecision}C}{N(\text{classes})} \quad (1)$$

本文使用的IoU(目标框和预测框的交并比)阈值为0.5,若IoU>0.5则认为正确的检测,否则认为是错误的检测,给定一张图像类别C,计算其Precision如公式(2)所示,图像正确预测的数量除以在这张图像上这类目标的总数量。

$$\text{Precision}C = \frac{N(\text{TruePositives})C}{N(\text{TotalObjects})C} \quad (2)$$

C类目标的平均精度为在验证集上所有图像对于C类的精度值的总和除以有C类目标的所有图像的数量,如公式(3)所示。

$$\text{AveragePrecision}C = \frac{\sum \text{precision}C}{N(\text{TotalImages})C} \quad (3)$$

Faster R-CNN 目标检测算法的损失主要分为RPN网络的损失和Faster R-CNN的损失,计算方法如公式(4),并且两部分都包含分类损失和回归损失。计算分类损失用的是交叉熵损失,而计算回归损失用的是Smooth_l1_loss, Smooth_l1_loss 损失和L₁损失基本相同,只是当L₁误差值非常小时,表示为一个确定值σ,在计算回归损失的时候,只计算正样本(前景)的损失,不计算负样本的位置损失。

$$L(\{p_i\}, \{t_i\}) = \frac{1}{N_{\text{cls}}} \sum_i L_{\text{cls}}(p_i, p_i^*) + \lambda \frac{1}{N_{\text{reg}}} \sum_i p_i^* L_{\text{reg}}(t_i, t_i^*) \quad (4)$$

图7为本文提出的基于小数据的目标检测算法和Faster R-CNN算法训练损失的变化,基于小数据的目标检测算法的分类损失和回归损失均比原始Faster R-CNN算法低很多,同时有效地防止了模型过拟合,加速收敛。此外,基于小数据的目标检测算法的mAP为98.2%,比原始Faster R-CNN提高了2.1%,图8为基于小数据的目标检测算法的识别结果,图9为基于Faster R-CNN算法识别的结果。通过以上结果证明本文提出的小数据钢筋检测模型在钢筋图像数量不足的情况下显著地提高了钢筋计数的准确率和泛化能力。

4 总结与展望

本文提出了基于钢筋训练数据集不足情况下的

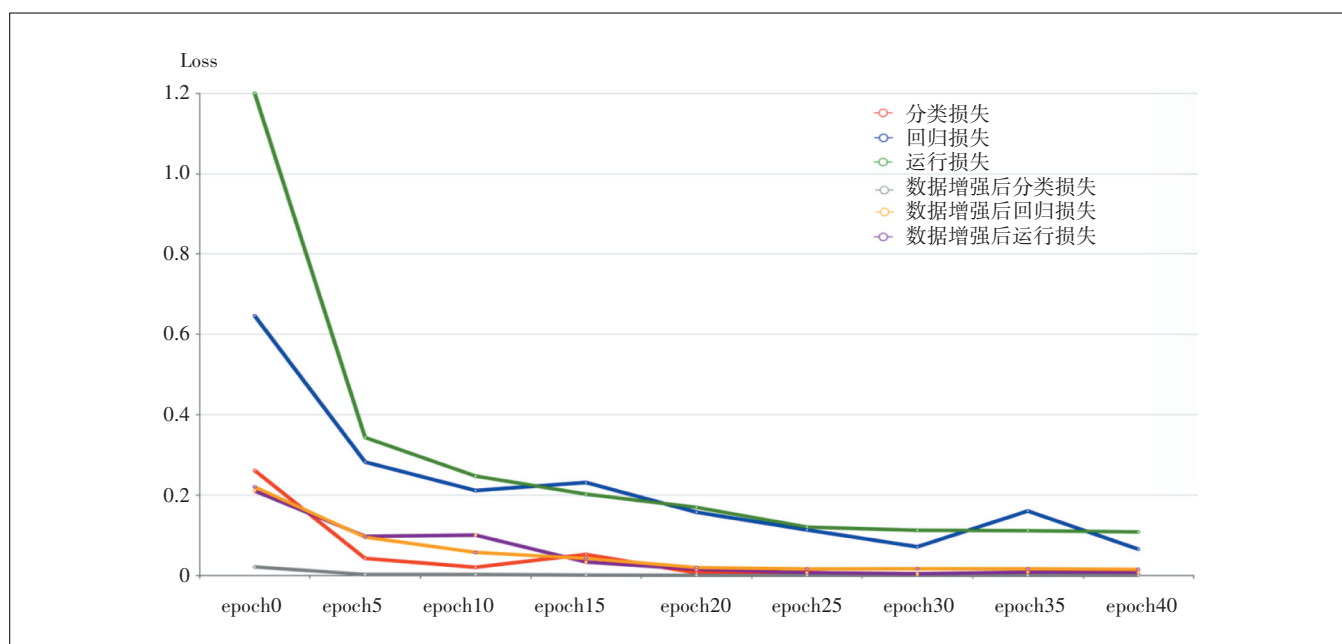


图7 2种算法的训练损失变化

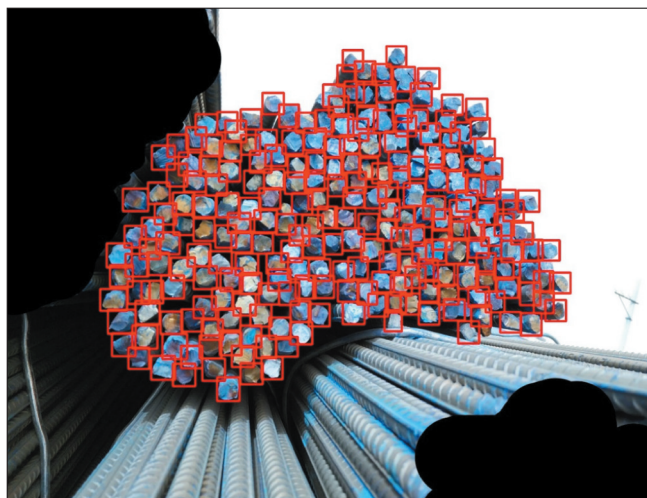


图8 基于小数据目标检测算法的检测结果

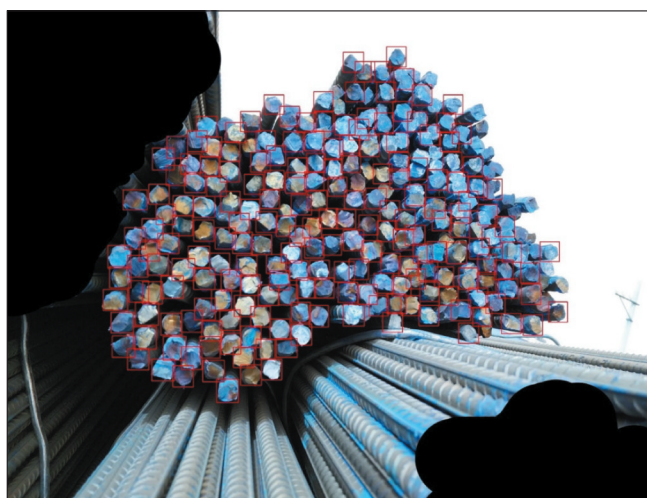


图9 Faster R-CNN算法的检测结果

小数据目标检测模型,该模型可用于工地复杂环境中,工地工作人员只需将拍摄的钢筋图像上传至该模型,便可以快速得出钢筋的数量,相比于传统的人工计数方法,更高效、准确。同时,该模型在一些特殊情况下识别效果有待提高,例如钢筋长短不齐、存在遮挡等情况,后续需根据实际应用场景出现的各种情况进行针对性的算法优化。

参考文献:

- [1] 李正,许前江,张峰,等. 智慧工地系统在建筑施工过程中的应用[J]. 建筑电气,2017(9):65-68.
- [2] 彭俊善. 基于数字图像处理技术建筑工地钢筋计数的方法[D]. 深圳:深圳大学,2017.
- [3] 郝际平. 把握机遇与挑战 携手共创装配式钢结构建筑新辉煌 在2017年全国建筑钢结构行业大会上的讲话[J]. 中国建筑金属结构,2017(6).

- [4] 陈至坤,潘晓娣,王福斌,等. 基于神经网络的钢筋计数方法研究[J]. 传感器与微系统,2010(8):44-47.
- [5] 林俊义,黄常标,刘斌,等. 一种基于双目立体视觉的激光线扫描技术[J]. 机械设计与制造,2011(8):205-207.
- [6] REN S, HE K, GIRSHICK R, et al. Faster r-cnn: Towards real-time object detection with region proposal networks [C]//Advances in neural information processing systems. 2015:91-99.
- [7] 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.
- [8] ZOPH B, CUBUK E D, GHIASI G, et al. Learning Data Augmentation Strategies for Object Detection [J]. arXiv preprint arXiv:1906.11172,2019.
- [9] RADFORD, ALEC, LUKE M, et al. Unsupervised representation learning with deep convolutional generative adversarial networks [J]. arXiv preprint arXiv:1511.06434 (2015).
- [10] LONG, M, CAO Y, WANG J, et al. Learning transferable features with deep adaptation networks [J]. arXiv preprint arXiv:1502.02791, 2015.
- [11] REDMON, J, DIVVALA, S, GIRSHICK, R, et al. You only look once: Unified, real-time object detection [C]// In Proceedings of the IEEE conference on computer vision and pattern recognition, 2015.
- [12] ROSS GIRSHICK. Fast R-CNN [C]// 2015 IEEE International Conference on Computer Vision (ICCV). IEEE, 2016..
- [13] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation [C]// In Proceedings of the IEEE conference on computer vision and pattern recognition.2014.
- [14] LIU W, ANGELOV D, ERHAN D, et al. SSD: Single Shot MultiBox Detector [C]// In European conference on computer vision. 2015.
- [15] Faster R-CNN: Down the rabbit hole of modern object detection [EB/OL]. [2019-11-22]. http://www.sohu.com/a/217946113_642762.
- [16] NING C C, ZHOU H J, SONG Y, et al. Inception Single Shot Multi-Box Detector for object detection [C]//2017 IEEE International Conference on Multimedia & Expo Workshops (ICMEW). IEEE, 2017.
- [17] FERNANDO N, WICKRAMASINGHE M, ZOYSA K D, et al. A quality metric for object detection and focus for low-cost UAVs [C]// 2016 Sixteenth International Conference on Advances in ICT for Emerging Regions (ICTer). IEEE, 2016.

作者简介:

王志丹,毕业于北京交通大学,工程师,硕士,主要从事智慧城市、深度学习等相关工作。

