

基于深度学习的小目标检测研究与应用综述

刘 颖^{1,2,3}, 刘红燕³, 范九伦³, 公衍超^{1,3}, 李莹华³, 王富平^{1,3}, 卢 津^{1,3}

(1. 电子信息现场勘验应用技术公安部重点实验室, 陕西西安 710121; 2. 陕西省无线通信与信息处理技术国际合作研究中心, 陕西西安 710121; 3. 西安邮电大学图像与信息处理研究所, 陕西西安 710121)

摘 要: 目标检测在基于传统手工特征及深度学习算法上已经取得较大发展, 然而针对小目标检测的研究近几年才开始出现, 研究成果较少, 且大都是在已有目标检测算法基础上进行改进, 以提高小目标检测的检测精度. 小目标像素点少, 本身携带的特征少, 多次下采样后就更难进行特征提取, 因而小目标检测面临极大挑战. 小目标检测在自动驾驶、遥感图像检测、刑侦等领域都有广泛应用需求, 对于小目标检测技术的研究有重要的实用价值. 本文对小目标检测的现有研究成果进行了详细综述. 首先, 将现有算法按照检测需要的阶段数分为一阶段、二阶段、多阶段, 描述了 RetinaNet、CornerNet-Lite、特征金字塔网络 (Feature Pyramid Network, FPN) 等算法的原理并进行了对比分析. 其次, 本文描述了小目标检测技术在不同领域的应用情况, 并汇总了 MS COCO、PASCAL VOC、DOTA、KITTI 等数据集及算法性能评价指标. 最后, 总结了小目标检测面临的挑战, 并展望了未来的研究方向.

关键词: 小目标检测; 尺度变换; 特征金字塔; 深度学习; 特征提取; 卷积神经网络

中图分类号: TP391.41; TP183

文献标识码: A

文章编号: 0372-2112 (2020) 03-0590-12

电子学报 URL: <http://www.ejournal.org.cn>

DOI: 10.3969/j.issn.0372-2112.2020.03.024

A Survey of Research and Application of Small Object Detection Based on Deep Learning

LIU Ying^{1,2,3}, LIU Hong-yan³, FAN Jiu-lun³, GONG Yan-chao^{1,3}, LI Ying-hua³, WANG Fu-ping^{1,3}, LU Jin^{1,3}

(1. Key Laboratory of Electronic Information Application Technology for Scene Investigation, Ministry of Public Security, Xi'an, Shaanxi 710121, China;
2. International Joint Research Center for Wireless Communication and Information Processing, Shaanxi, Xi'an, Shaanxi 710121, China;
3. Center for Image and Information Processing, Xi'an University of Posts and Telecommunications, Xi'an, Shaanxi 710121, China)

Abstract: Object detection has been well studied based on the traditional manual features and deep learning algorithm. However, the research on small object detection has just begun in recent years and there are little research outcome available. Furthermore, most of the methods proposed are based on the traditional object detection algorithms with certain modifications so as to improve the accuracy of small object detection. Small object contains fewer pixels and has less features, and it is even harder to extract object features after down sampling. Hence, small object detection is a challenging task. Small object detection has a wide range of application requirements in the fields of automatic driving, remote sensing image detection, and criminal investigation. It has important practical value for the research of small object detection technology. In this paper, the existing research results of small object detection are summarized. Firstly, the existing algorithms are classified into one stage, two stages and multi-stages according to the number of stages for detection. The principles of RetinaNet, CornerNet-Lite, feature pyramid network (FPN) and other algorithms are described and compared. Secondly, this paper describes the application of small object detection technology in different fields, and summarizes the data sets such as MS COCO, PASCAL VOC, DOTA, KITTI and algorithm performance evaluation indicators. Finally, the challenges faced by small object detection are concluded, and the future research directions are prospected.

Key words: small object detection; scale variation; feature pyramid; deep learning; feature extraction; convolutional neural network

收稿日期: 2019-09-29; 修回日期: 2019-11-20; 责任编辑: 覃怀银

基金项目: 陕西省国际合作交流项目 (No. 2018KW-003); 国家自然科学基金 (No. 61671377); 国家重点研发计划 (No. 2017YFC0803805)

1 引言

小目标检测在许多领域中都有着广泛且重要的应用^[1]. 如在自动驾驶领域, 汽车采集到的高分辨率场景照片中行人目标或交通标志太小, 但这些小目标的准确检测是实现安全自动驾驶的重要前提; 在医学领域, 医学图像中微小肿块的成功检测是早期准确诊断肿瘤的重要前提; 自动工业检查定位材料表面上的小缺陷也说明小目标检测的重要性; 卫星图像分析中需要有效地注释诸如汽车、船和房屋之类的物体, 然而这些物体都太小而不易检测到; 在刑侦图像中, 异常的小包裹、小行人、车里面的小挂件、衣服上的小标志、室内的一些小摆设等都是破案的关键线索. 因为在现实世界中部署了更复杂的系统, 所以小目标检测有很大的价值. 小目标的定义主要有两种^[2]: 第一种是绝对小物体, COCO 数据集^[3]中指明, 当物体的像素点数小于 32×32 时, 此物体即可被看作是小物体. 第二种是相对小物体, 是相对于原图的长宽来看, 当目标尺寸小于原图尺寸的 0.1 倍时就可认为是相对小物体. 如图 1 所示, (a) 中黄色框圈出的为绝对小物

体, (b) 中红色框圈出来的是相对小物体. 对于小目标检测有很大的难度, 首先, 小物体本身的像素点数少, 即所带的特征信息少; 其次是因小物体的像素点数少, 多次下采样后就更难提取到特征^[4], 因此, 小目标检测面临极大挑战. 目前对于小目标检测的研究还很少, 现有算法大部分还是在目标检测的基础上做修改, 提高对于尺度变化的鲁棒性, 进而提高对于小目标的检测效果^[5-7]. 小目标检测是目标检测的一个难题, 目前只有一篇关于小目标检测的综述^[8]. 表 1 从算法和数据集两个方面对文献[8]和本文进行了对比分析. 在算法方面, 文[8]首先介绍了目标检测算法, 然后介绍了几种小目标检测算法. 本文将现有小目标检测算法按照检测需要的阶段数分为三类, 描述了各类算法原理并进行对比分析, 并且介绍了小目标检测在各个领域的应用算法, 本文对于小目标检测的算法介绍的更加全面. 在数据集方面, 文[8]介绍了 4 种航空遥感数据集, 本文不仅介绍了学者研究使用的公共数据集, 并且介绍了几类特定领域应用数据集, 对小目标检测做了一个更加全面的介绍.

表 1 综述的比较

	算法	数据集
文[8]	1. 目标检测算法: 二阶段算法: R-CNN、SPP-Net、Fast R-CNN、Faster R-CNN 一阶段算法: YOLO 系列、SSD 2. 小目标检测算法: Takeki A 等 ^[9] 、Kampffmeyer M 等 ^[10] 、Mundhenk T N 等 ^[11] 、FPN、DSSD、RRC、Ren Y 等 ^[12] 、Kb-RANN、YOLOv3	航空遥感数据集: DOTA VEDAI RSOD-Dataset UCAS-AOD
本文	1. 小目标检测公共数据集下的算法: 一阶段算法: YOLOv2、SSD513、DSSD513、RetinaNet、CornerNet-Lite 二阶段算法: Faster R-CNN + + +、Faster R-CNN w FPN、Faster R-CNN by G-RMI、Deformable R-FCN、Mask R-CNN、DetNet-59、SNIP、SNIPER 多阶段算法: Cascade R-CNN、Cascade R-CNN * * 2. 小目标检测各领域应用算法: Eggert C 等 ^[13] 、finding tiny faces、ALFnet、PGAN、Kb-RANN、Ren Y 等 ^[12]	公共数据集: MS COCO、PASCAL VOC 各领域应用数据集: DOTA、UCAS-AOD、NWPU、VHR-10、RSOD-Dataset、Tsinghua-Tencent100K、KITTI、Flickrlogos、CHIP-SO



图1 小目标样例

2 小目标检测算法研究现状

小目标检测的算法都是基于卷积神经网络^[8,14-16]. 目前流行的小目标检测算法可以分为三类^[17]: 第一类是使

用一个卷积神经网络直接预测不同目标类别与位置的一阶段算法. 一阶段算法不需要使用候选框, 将目标框定位的问题转换为回归的问题, 直接产生目标的类别概率和位置坐标值, 经过单次检测可直接得到最终的检测结果. 这类算法速度快, 但是准确性较低. 第二类是基于候选区域的算法, 这类算法是两阶段的, 需要先使用区域候选网络(Region Proposal Network, RPN)产生候选区域, 然后再在候选区域上进行分类与回归, 经过两个阶段得到最终的检测结果. 这一类方法准确度较高, 但速度慢. 第三类是多阶段算法, 这类算法是对二阶段检测器进行迭代, 迭代过程中使用不同的检测头和预选框. 迭代会大大增加模型的大小, 使得在训练过

程中需要更大的计算能力支持. 因为硬件条件限制, 目前有三阶段算法, 因实现过程更为复杂, 三阶段算法精度更高, 速度会减慢.

2.1 小目标检测算法网络结构

深度学习主要涉及三种方法: 自编码神经网络^[18]; 深度置信网络^[19]; 卷积神经网络. 卷积神经网络通过共享权值减少了网络各层之间的连接, 同时降低了过拟合的风险. 卷积和池化简化了模型的复杂度, 减少了模型的参数. 卷积神经网络由于避免了对图像的复杂前期预处理, 可以直接输入原始图像, 因而在图像处理领域得到了广泛的应用. 可应用于图像分类、目标检测、目标分割等方向.

小目标检测领域中, 广泛使用的网络是卷积神经网络, 小目标检测算法中常用的卷积神经网络模型有 ResNet(Residual Neural Network)、Hourglass. ResNet 是应用最为广泛的卷积神经网络模型^[20], 提出原因是网络到达一定深度时, 随着网络深度的增加并不能使性能进一步增加, 反而导致梯度消失或者梯度爆炸. ResNet 的主要思想是残差学习模块, 将输入信息传到后面层中, 采用跳跃连接, 这一层网络不用学习整个输出, 而是学习上一个网络输出的残差. Hourglass 模型的网络结构形似沙漏状^[21], 如图 2 所示, 使用多尺度特征, 捕捉人类各关节的空间位置信息, 可用于人类姿态估计. 在小目标检测中, Hourglass 模型是一种新的分支, 通过预测边界框的左上角和右下角坐标确定目标的位置, 打破了传统锚框的方法. 这种架构考虑到了多尺度问题, 有利于小目标的检测.

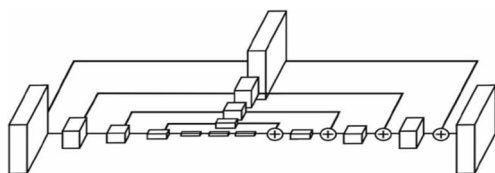


图2 Hourglass模型的结构

2.2 一阶段的小目标检测算法

2.2.1 YOLOv2 算法^[22]

YOLOv2 算法是端到端的目标检测算法, 在 YOLOv1 的基础上, 作者提出了改进的 YOLOv2, 添加了 BP 层, 用 448×448 的输入分辨率在 ImageNet 上进行参数微调, 并把全连接层取消, 使用锚框来预测目标方框, 用 k 均值来选择好的方框先验, 对于每个方格, 预测 5 个方框, 并且把相邻特征分到不同的通道中, 进行多尺度的训练.

2.2.2 SSD513 算法^[23]

SSD(Single Shot Multi-Box Detector) 算法也是端到端的目标检测算法. 低层特征具有较高的分辨率, 高分

辨率特征包含很强的位置信息, 这对于小目标的检测有重大意义. SSD 算法提取多层特征进行预测, 结合不同层次的特征进行更好的检测, 检测效果相对于 YOLO 算法有了大的提升. SSD513 算法是在 SSD 算法上做了修改, 它的网络模型是 ResNet-101, 区别于 SSD 的 VGG 网络.

2.2.3 DSSD513 算法^[24]

DSSD(Deconvolutional Single Shot Detector) 算法在改进的 SSD 算法中效果较为显著, DSSD513 算法的网络模型是 ResNet-101, 区别于 SSD 的 VGG 网络. DSSD 将 SSD 的 VGG 网络用 ResNet-101 进行了替换, 如图 3 所示, 在分类回归之前引入了残差模块, 在 SSD 添加的辅助卷积层后又添加了反卷积层形成“宽-窄-宽”的“沙漏”结构. DSSD 与 SSD 相比, 其小目标的检测效果有了明显的提高.

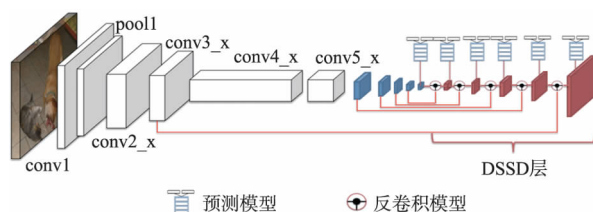


图3 DSSD算法

2.2.4 RetinaNet 算法^[25]

RetinaNet 是对原有 FPN 网络与 FCN(Fully Convolutional Networks) 网络的组合应用. Focal loss 的提出及在单阶段目标检测网络 RetinaNet 的成功应用是该算法的亮点. Focal loss 是一种改进了的交叉熵损失, 它通过在原有的交叉熵损失上与使易检测目标对模型训练贡献削弱的指数式相乘, 从而使得 Focal loss 成功地解决了在目标检测时正负样本区域极不平衡而目标检测损失易被大批量负样本所左右的问题.

2.2.5 CornerNet-Lite 算法^[26]

该算法提出了一种不需要锚框的基于关键点的目标检测方法. CornerNet 在单阶段检测器中的性能是顶级的, 但是需要大量的预处理才能得到较高的准确率, 消耗较多的计算资源. 文献[26]提出 CornerNet 的两种变形: 一是 CornerNet-Saccade, 如图 4 所示, 基于注意力机制, 从而不需要对图片中的每个像素做处理. 另一个是 CornerNet-Squeeze, 引入了更加紧凑的主干网络. CornerNet-Saccade 在追求高准确率的同时, 尽可能提高速度, 以准确率优先, 对标于 CornerNet. CornerNet-Squeeze 在追求高实时性的同时, 尽可能提高准确率, 以速度优先, 对标于 YOLOv3.

2.3 两阶段的小目标检测算法

2.3.1 Faster R-CNN +++ 算法^[20]

此算法是基于 Faster R-CNN 的目标检测方法, 不

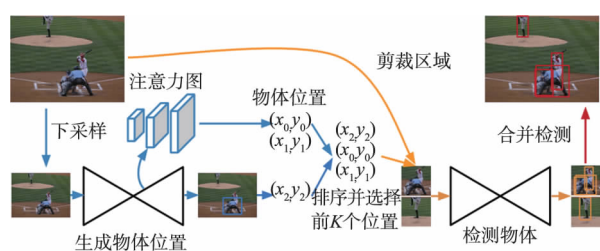


图4 CornerNet-Saccade的流程图

同点是采用了 ResNet-101 网络, 替换了原来的 VGG-16 网络, 效果得到提升. 该算法提供了一种残差学习框架, 能够使深的网络的训练过程得到简化, 在 COCO 数据集上有了一定的提升, 在 2015 年的 ILSVRC 分类任务上获得了第一名.

2.3.2 Faster R-CNN w FPN 算法^[27]

特征金字塔在检测多尺度物体的检测系统中是一个基本的构建. 通过使用图像金字塔提取特征, 构成特征金字塔来进行多尺度物体的检测. 但随着卷积神经网络的出现, 由于卷积神经网络固有的多尺度特性, 不同的卷积层可以看成是一个特征金字塔结构. FPN 算法利用卷积神经网络的多尺度, 金字塔结构构建了特征金字塔网络.

如图 5 所示, 卷积神经网络的不同卷积层与自上而下的结构进行横向连接, 使得不同层次的特征都具有很强的语义信息. 从图中可以看出, 语义信息越强, 特征层的边缘越深. 图 5 的左半部分是卷积神经网络的卷积层, 在卷积的过程中, 底层特征具有较高的分辨率, 但是语义信息较差; 上层特征具有高的语义信息, 但是分辨率较低. 右半部分是自上而下的结构, 上采样得到的特征与相对应的自下而上的特征图融合, 这个过程直到最后的特征金字塔. FPN 算法在 COCO 数据集上取得了很好的效果, 提高了对小目标的检测效果.

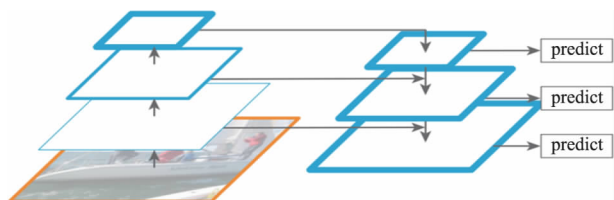


图5 特征金字塔网络

2.3.3 Faster R-CNN by G-RMI 算法^[28]

此算法的提出是为了探寻一种目标检测体系, 以实现在应用平台速度、记忆、准确率的平衡, 在文献 [28] 中, 提出了一种 Faster R-CNN、R-FCN 和 SSD 系统的统一实现, 这种系统也叫“meta architectures”通过使用可替换的特征提取器和变换这个体系结构里面的例如图像尺寸的参数, 来制定速度、精确度曲线, 从而确定对于检测速度、精度, 内存等的影响. 提出的算法实

现速度快, 可以部署在移动设备上, 且精度较高, 在 COCO 数据集上有较好的表现.

2.3.4 Deformable R-FCN 算法^[29]

卷积神经网络由于其构建模块中的几何结构是固定的, 所以只能处理模型几何变换的问题. 在这种算法中, 引入了可变形卷积和可变形感兴趣区域池化这两个新的模块. 它们用来增强卷积神经网络的转换建模能力, 这两种方法不需要额外的监督, 都是基于增加模块中的空间采样位置和从目标任务中学习偏移量的思想. 新的模块可以很容易地取代现有卷积神经网络中的普通模块, 并且可以通过标准的反向传播轻松地进行端到端的训练, 从而产生可变形的卷积网络.

2.3.5 Mask R-CNN 算法^[30]

Mask R-CNN 可以看做是一个通用实例分割架构. Mask RCNN 以 Faster RCNN 原型, 增加了一个分支用于分割任务. Mask RCNN 比 Faster RCNN 速度慢一些, 达到了 5fps. 可用于人的姿态估计等其他任务. Mask RCNN 主要是用来目标分割, 在小目标检测上效果提升明显.

2.3.6 DetNet-59 算法^[31]

目前的目标检测方法大都需先在 ImageNet 上做预训练, 然后在小目标检测数据集上微调. 目标检测和目标分类有一定的差别, 而目前的网络特征提取器都是面向目标分类的, 面向目标检测的网络仍是一片空白, 所以设计了 DetNet 网络. DetNet 是一个为了解决上述问题而设计的全新网络. DetNet 针对不同大小的物体而像 FPN 一样使用了更多的层. 和 ResNet 的 5 层结构相比, DetNet 保留了 ResNet50 中的前 4 层, 同时增加了额外的第六层, 且 4 到 6 层的步长都是 16, 即便如此, 它优于 ImageNet 预训练模型的地方在于成功保留了特征的空间分辨率, 但同时也会增加神经网络的计算和内存成本. 为保证 DetNet 的效率, 文献 [31] 引入了低复杂度的 dilated bottleneck, 为图 6 中的 A 和 B 两种, 由此, DetNet 做到了较高分辨率和较大感受野的兼得.

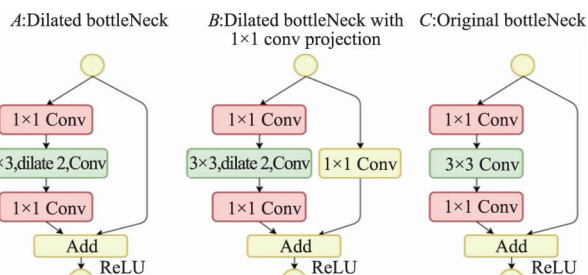


图6 DetNet网络

2.3.7 SNIP 算法^[32]

SNIP (Scale Normalization for Image Pyramid) 算法在图像金字塔的相同尺度下训练和测试检测器. 文献

[32]通过不同的实验引出了 SNIP 算法. 文献中的第一个实验说明, 训练图像的分辨率和验证图像的分辨率应一致, 提出了在小目标检测中上采样有很大的意义. SNIP 是在第二个实验多尺度训练(Multi-Scale Training, MST)基础上改进的, 限定了不同尺寸的物体在训练过程中的梯度回传, 削弱了训练数据中尺寸非常大或非常小的物体对训练效果的影响, SNIP 算法在针对小目标的检测上具有很好的效果.

图 7 给出了 SNIP 算法的系统图, 算法通过把测试图片转换为 (480×480) 、 (800×1200) 、 (1400×2000) 三种不同尺度的图片, 在不同的尺度变换图片中, 选择

相应尺度范围的候选区域进行反向传播, 舍弃无效的候选区域. SNIP 算法使得不同尺度的物体都能得到充分的训练, 提高了对于小目标的检测效果.

2.3.8 SNIPER 算法^[33]

SNIPER 摆脱了对较高分辨率图像的依赖, 训练时使用较低分辨率的 chip, 所以 batch 也可以调大, 节省了内存, 使得 SNIP 思想可以落地, 即所谓标题中的“实战版本”. 文献[33]证明了 SNIP 提出的“在高分辨率图片上进行训练这一常用技巧并不是提升精度必须要使用的策略”这一观点.

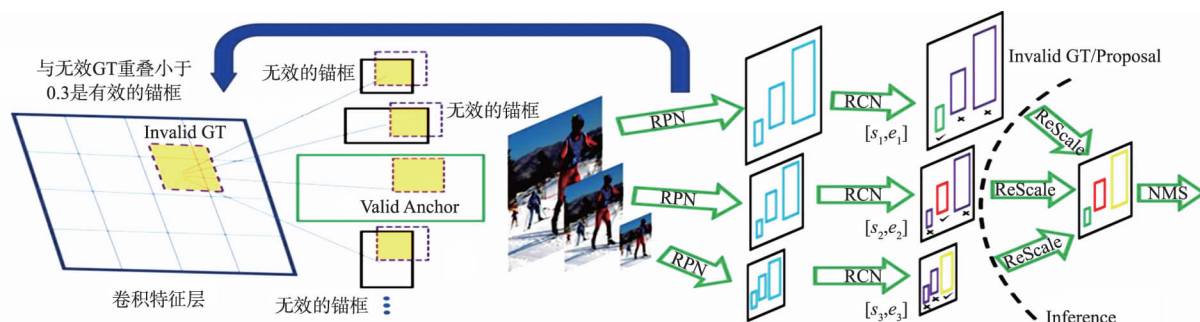


图7 SNIP算法的系统图

2.4 多阶段的小目标检测算法

2.4.1 Cascade R-CNN 算法^[34]

文献[34]和 SNIP 算法都是通过大量实验来探索对小目标检测的提升效果. Cascade R-CNN 算法主要是针对目标检测中的两阶段方法中的阈值进行改进. 我们都知道检测中的 IoU 阈值对于样本的选取是至关重要的. 如果 IoU 阈值过高, 会导致正样本质量很高, 但是数量会很少, 会出现样本比例不平衡的问题; 如果 IoU 阈值较低, 样本的数量就会增加, 但是样本的质量也会下降. 因此如何平衡样本数量与质量这一关系, 即如何选取好的 IoU, 对于检测结果来说很重要.

文献[34]为了解决这一问题, 提出了多阶段的结构, 通过不断提高 IoU 的阈值, 使得在保证样本数量的同时, 也能使得样本的质量不下降, 最后训练出高质量的检测器.

在图 8 中主要介绍了几种不同的类似于级联方法的比较. (a) 图中是通用的 Faster-RCNN 方法; (b) 图是在推理中使用迭代的方式不断地改进推理结果, 但是在检测时使用的是相同的检测头. 这会出现两个问题: 第一是单一阈值对于检测高质量的目标并不有利, 第二是每次检测出来的分布差别较大; (c) 图中虽然使用了多个检测头, 但是预选框只有一个分布, 以及只做了一个边界框回归; (d) 图即是文献[34]提出的方法, 融合了前面的几个方法, 在得到 B0 之后, 通过每一个不同的检测头, 该检测头的 IoU 阈值只逐渐提升的, 在通

过检测头之后, 得到的 B1 会和前一个 B0 结合生成新的预选框, 用作下一个检测头, 依次处理, 就可以得到最终的结果. 提高了小目标检测的效果.

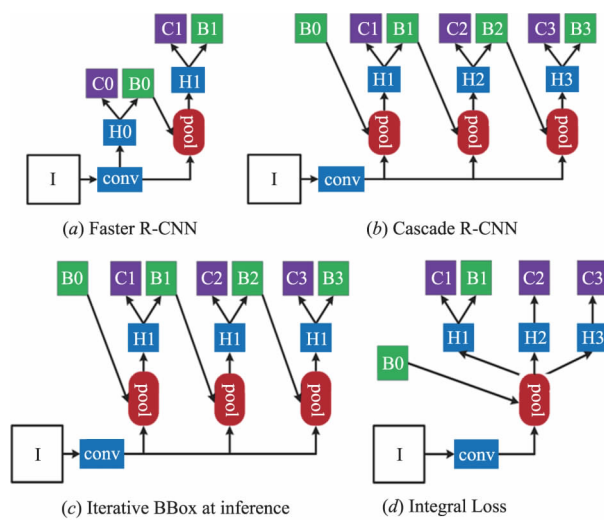


图8 不同的框架结构

2.4.2 Cascade R-CNN** 算法^[35]

Cascade R-CNN** 算法在物体检测方面和 Cascade R-CNN 类似, 使用了级联的 R-CNN 网络, 网络主干是 ResNeXt-452. 该算法在许多数据集上都实现了性能的提升, 用于通用和特定的目标检测. 该算法也在实例分割上也有很大的性能提升, 在未来, 级联的 R-CNN 结构在目标检测和实例分割领域会有很大的帮助.

2.5 算法优缺点对比和实验结果对比

本小节对通用目标下的小目标检测各个算法的优缺点和性能进行对比分析. 表 2 是不同小目标检测算法的优缺点对比, 表 3 是不同小目标检测算法性能的比较情况, 比较时各个算法都是把 COCO 数据集作为测试基准, 因为 COCO 中的图片背景比较复杂, 目标数量比较多, 目标尺寸更小, 所以 COCO 数据集上的任务更难, 是小目标检测公用数据集, 算法性能的比较参数是平均准确率(Average Precision, AP), APs 是针对小目标的评价指标, 小目标指的是像素小于 32×32 的目标.

表 2 和表 3 从一阶段、二阶段、多阶段算法三个方面对通用目标下的小目标算法进行分类总结. 一阶段

的小目标检测算法速度都很快, 但是精度普遍都没有很高, 在检测小目标时效果较差, 可以使用在对速度有较高要求的应用上. 在刑侦领域, 对于小目标检测的研究有很大的意义, 因为监控视频中目标检测的实时性和精准率都有很高的要求^[36], 所以可以从一阶段算法入手, 进一步提高检测精准率. 两阶段比一阶段的小目标检测算法的精度普遍高, 更适用于小目标检测. 多阶段的小目标检测算法较少, Cascade R-CNN^{**} 算法是三阶段算法, 检测精度为 33.4%, 是目前通用目标下小目标检测的最高精度, 未来对于小目标检测算法的研究可以从多阶段入手, 有很大的研究价值.

表 2 不同小目标检测算法的优缺点对比

	算法	算法逻辑	年份	优点	缺点
一阶段算法	YOLOv2	增加 BP 层, 提高图像分辨率, 取消全连接层等是 YOLOv1 的改进	2016	在 YOLOv1 速度的基础上, 提高了精度	使用预训练, 难迁移
	SSD513	多层特征预测, 采用 ResNet-101 网络模型	2016	结合多层特征信息	对于小目标检测底层信息用的仍不够
	DSSD513	多层特征预测, 采用 ResNet-101 网络模型, 引入了残差模块	2017	在小目标检测上精度比 SSD 有了提高	速度比 SSD 慢
	RetinaNet	FPN + FCN, Focal loss 的提出	2017	解决样本类别不平衡的问题, 精度超越了二阶段算法	—
	CornerNet-Lite	为了提高速度, 减少要处理的像素点数和每个像素点数的处理次数	2019	精度与速度兼得	—
两阶段算法	Faster R-CNN + + +	Faster R-CNN, ResNet-101 网络	2016	使深的网络训练得到简化	—
	Faster R-CNN w FPN	自下往上、自上往下、横向连接结构	2017	特征金字塔的使用, 结合多层特征	增加了时间成本
	Faster R-CNN by G-RMI	通过 Faster r-cnn、SSD、R-FCN 统一框架下的对比, 提高在实际平台上的检测速度和准确率	2016	速度有了很大提升, 可以在移动终端上使用	—
	Deformable R-FCN	可变形卷积模块和可变形感兴趣区域池化模块的提出	2017	提高了 CNN 建模几何变换的能力	—
	Mask R-CNN	Mask R-CNN 以 Faster R-CNN 原型, 增加了一个分支用于分割任务	2017	在小目标检测上效果提升明显	Mask RCNN 主要是用来目标分割的
	DetNet-59	设计了小目标检测的 DetNet 网络	2018	DetNet 做到了较高分辨率和较大感受野的兼得	会增加神经网络的计算和内存成本
	SNIP	在图像金字塔的相同尺度下训练和测试检测器	2018	不同尺度的物体都得到了充分的训练, 小目标检测效果很好	—
	SNIPER	使用了较低分辨率的 chip 进行训练	2018	SNIP 的实战版本	—
多阶段算法	Cascade R-CNN	提出了多阶段的结构, 通过不断提高 IoU 的阈值, 最后训练出高质量的检测器.	2017	三阶段结构	—
	Cascade R-CNN ^{**}	多阶段结构, 迭代方法	2019	网络采用了 ResNeXt-152, 精度是目前最高的	—

表 3 不同小目标检测算法的性能比较

	算法	网络	AP	APs
一阶段 算法	YOLOv2 ^[22]	DarkNet-19	21.6	5.0
	SSD513 ^[23]	ResNet-101	31.2	10.2
	DSSD513 ^[24]	ResNet-101	33.2	13.0
	RetinaNet ^[25]	ResNet-101	39.1	21.8
	CornerNet-Lite ^[26]	Hourglass-54	43.2	24.4
两阶段 算法	Faster R-CNN ++ ^[20]	ResNet-101	34.9	15.6
	Faster R-CNN w FPN ^[27]	ResNet-101	36.2	18.2
	Faster R-CNN by G-RMI ^[28]	Inception-ResNet-v2	34.7	13.5
	Deformable R-FCN ^[29]	Aligned-Inception-ResNet	37.5	19.4
	Mask R-CNN ^[30]	ResNet-101	38.2	20.1
	DetNet-59 ^[31]	DetNet-59	40.3	23.6
	SNIP ^[32]	ResNet-101(ResNet-101 proposals)	43.4	27.2
	SNIPER ^[33]	ResNet-101	46.1	29.6
多阶段 算法	Cascade R-CNN ^[34]	ResNet-101	42.8	23.7
	Cascade R-CNN ** ^[35]	ResNeXt-152	50.9	33.4

3 小目标检测在各领域应用研究

小目标检测算法目前在企业标识检测、小人脸检测、行人检测、交通标志检测领域、自动驾驶领域、遥感图像检测领域、刑侦等领域有广泛应用需求. 小目标检测算法在各个领域的应用都是处于起步阶段, 都是在通用小目标检测算法的基础上做一些改进^[37], 在未来具有重要研究意义和较大的研究空间.

Eggert C 等^[13]将 Faster R-CNN 应用于企业标识的检测, 分析了 Faster R-CNN 对于小目标检测的性能情况, 并从多尺度 RPN 和多尺度分类网络进行研究, 提出了一种改进的产生锚框方案, 并利用了小物体的高分辨率特征图对 faster R-CNN 改进, 在 Flickrlogos 数据集上检测性能取得了明显的提高.

Hu P 等^[38]提出了一个小人脸检测的算法, 该算法从尺度不变性、图像分辨率、上下文推理三个角度针对小人脸识别问题进行了探究, 在一张有 1000 个人脸的图片中可以检测的人脸数大约为 800 个, 该算法在 FD-DB 和 WIDER FACE 数据集上取得了很好的效果.

Liu W 等^[39]提出了一种针对于行人检测的算法, 该算法用到了级联网络, 文献[39]的主要思想是利用不断提升的 IoU 阈值训练多个定位模块, 来达到提升定位精度的目的, 该文献以 SSD 为框架, 以 ResNet-50 为基础网络, 该文献证明了多阶段检测是提升小目标检测的一个很有效的方法.

Li J 等^[40]提出的 PGAN 方法应用在交通标志检测上, 该方法对小目标生成高分辨率特征表示, 使小目标

的特征表示与大目标特征表示类似. 生成器网络通过较前层提取细粒度特征, 将小目标分辨率较低的特征转换为分辨率较高的特征, 判别器网络一方面分辨小物体生成的高分辨率特征与真实大物体特征, 同时证明带感知损失的生成高分辨率特征对检测准确率是有帮助的. 生成器和鉴别器以竞争的方式分辨特征.

准确的交通标志检测作为一个典型的小目标检测问题, 在先进的驾驶员辅助系统和自动驾驶系统中具有重要意义. Yi K 等^[41]提出了一个以大脑为灵感的网络, 命名为 Kb-RANN. 算法采用了一种新的注意力机制的复发性神经网络, 以提高检测精度的细粒度方式. 此外, 还结合了特定领域的知识和直观的知识来提高效率. 方法移植到了设计的嵌入式系统上, 并成功地部署到了自驾汽车上. 实验结果表明, 该方法比常用的几种目标检测方法具有更好的性能.

Ren Y 等^[12]研究了将 Faster R-CNN 模型应用在光学遥感图像中的小目标检测任务. PASCAL VOC 挑战的表现在 Faster R-CNN 已经很高了. 然而, 直接将 Faster R-CNN 应用于小型遥感对象通常会导致性能较差. 所以对 RPN 设置适当的锚框, 而且设计了一个采用自上而下和跳跃连架构. 此外, 还结合背景信息进一步提高了小型遥感目标的探测性能, 同时采用了一种简单的采样策略来解决不同类别之间图像数量不平衡的问题. 最后, 介绍了一种简单而有效的训练数据增强方法——“随机旋转”. 实验结果表明, 改进的 Faster R-CNN 算法大大提高了遥感小目标检测的平均精度.

在刑侦领域中, 小目标检测在公安监控视频中具有

研究价值监控视频中的小包裹,视频里面的小行人,衣服上的小标志,车里面的小挂件,室内的小摆设都是破案的关键线索. 目前对于监控视频中小目标的检测大多数还是使用人海战术,这样又费时费力,人有疲惫期,可能造成漏检的发生. 西安邮电大学图像与信息处理研究所借助与公安部门合作的实验室平台,在进行这方面的研究,并建立了刑侦应用领域的小目标数据集^[42].

对于小目标的检测是一个新的研究领域,通用目标下的小目标检测算法都很少,在各领域应用的研究还处于初始状态,但是小目标检测在各个领域都有很高的应用价值,以后有很大的发展前景.

4 小目标检测数据集和评价指标

4.1 数据集

在图像中小目标较小,具备的特征少,不易被检测,对于小目标的研究,需要针对于小目标的特定数据集,目前关于小目标检测的数据集包含公共数据集和各领域应用的数据集,公共数据集是指学者们用来学术研究的自然图像数据集,各领域应用的数据集是指在特定研究对象使用的数据集,比如航空遥感图像目标检测数据集、公司标志数据集、交通标志数据集、自动驾驶数据集. 其中常用的几个数据集及其主要参数如表 4 所示.

表 4 常用数据集及其参数

数据集	年份	单位	种类	图片数量
MS COCO	2014	微软公司	91	328000
PASCAL VOC	2005	Pattern Analysis, Statical Modeling and Computational Learning (一个由欧盟资助的网络组织)	20	17125
DOTA	2017	武大遥感国重实验室夏桂松与华科电信学院白翔	15	2806
UCAS-AOD	2015	中科大模式识别实验室	2	2420
NWPU VHR-10	2014	西北工业大学	10	800
RSOD-Dataset	2015	武汉大学	4	976
Tsinghua-Tencent100K	2016	清华大学与腾讯联合实验室	128	100000
KITTI	2011	德国卡尔斯鲁厄理工学院和丰田美国技术研究院	5	200000
Flickrlogos	2011	奥格斯堡大学多媒体计算和计算机视觉实验室	32	8240

4.1.1 公共数据集

MS COCO 数据集^[3]包含了自然图片以及生活中常见的目标图片,背景比较复杂,目标数量比较多,目标尺寸更小,图 9 是 COCO 数据集的样例. 因此 COCO 数据集上的任务更难,对于小目标检测任务来说,现在衡量一个模型好坏的标准更加倾向于使用 COCO 数据集上的检测结果.



图9 COCO数据集样例

PASCAL VOC^[43]作为基准数据集,在图像分类与检测、图像分割领域广泛使用. 对视觉任务中监督学习提供标签数据. 从 2005 年到 2012 年每年都会举行一场图像识别竞赛. 该竞赛的主要目的是识别真实场景中一些类别的物体.

4.1.2 各领域应用数据集

航空遥感图像具有尺度多样性,拍摄高度从几百

米到近千米,地面目标即使是同类目标也大小不一,且航空遥感图像的目标很多都是小目标,含有几十个甚至几个像素,所以航空遥感图像目标检测数据集大多数都是小目标检测数据集. 常用的数据集有以下几个. DOTA 数据集^[44]是一个用于航空图像中目标检测的大规模数据集. 对于 DOTA-v1.0,它包含来自不同传感器和平台的 2806 个航空图像. 每幅图像的大小在 800 × 800 到 4000 × 4000 像素之间,包含显示各种比例、方向和形状的对象. 完全注释的 DOTA 图像包含 188282 个实例. UCAS-AOD 数据集^[45]只包含两类目标:汽车、飞机,以及背景负样本. NWPU VHR-10 数据集^[46]共有 800 张图像,其中包含目标的 650 张,背景图像 150 张. RSOD-Dataset 数据集^[47]包含飞机、操场、立交桥、油桶四类目标,飞机有 446 张图片,操场有 189 张图片,立交桥有 176 张图片,油桶有 165 张图片.

Tsinghua-Tencent100K 是一个大规模交通标志基准数据集^[48],交通标志尺寸过小,成为交通标志小目标检测的经典数据集. 数据集图片涵盖大部分光照和天气条件,由腾讯实拍街景图分割而成. 所有图片通过手工标注,类别分为禁令、指示和警告标志三大类. Tsinghua-Tencent100K 是目前国内仅有的基于真实场景的公开

交通标志数据集,图片分辨率高且数量规模较大,填补了国内缺乏基准数据集的空白。

KITTI 数据集^[49]是目前国际上最大的自动驾驶场景下的计算机视觉算法评测数据集。KITTI 包含市区、乡村和高速公路等场景采集的真实图像数据,每张图像中最多达 15 辆车和 30 个行人,还有各种程度的遮挡与截断。

Flickrlogos 数据集^[50]由从网络相册收集的真实图像组成,这些图像描述了各种情况下的公司徽标。数据集共有 8240 张图像,分为 32 类。

在刑侦领域,西安邮电大学图像与信息处理研究所建立了小目标检测数据集 CIIP-SO^[42],图 10 为刑侦小目标检测数据集样例,把刑侦特色领域小目标检测数据集分了 17 类,有车内小挂件、衣服上的标志、行人、建筑上的商标牌、车辆上的划痕、包裹、自行车、摩托车、电瓶车、小汽车、客车、货车、双肩背包、斜挎包、手提包、皮箱、室内字画。

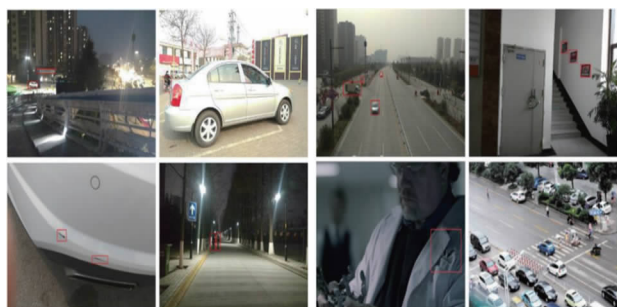


图10 刑侦小目标检测数据集样例

4.2 评价指标

平均准确率(Average Precision, AP)^[51]是小目标检测的一个常用评价指标,用于计算平均的检测精度和衡量检测器在每个类别上的性能好坏;而 MAP 则更多用于评价多目标的检测器性能,衡量检测器在所有类别上的性能好坏,即得到每个类别的 AP 值后再取所有类别的平均值。在 COCO 数据集中 AP 默认为 MAP, APs 是针对小目标的评价指标,小目标物体指的是像素小于 32×32 的目标,AP 计算公式为:

$$AP = \frac{TP}{TP + FP} \quad (1)$$

式(1)中 TP 是将正类预测为正类数,FP 是将负类预测为正类数误报。

除了检测准确率,检测速率也是小目标检测的一个重要评价指标,FPS(Frame Per Second)^[52]是小目标检测中速度的评价指标,指一秒内识别的图像数。小目标检测中速度是很重要的一个评价指标,只有速度够快,才能实现实时监测。

平均召回率(Average Recall, AR)表示检出正样本占实际正样本总数的比例,其计算公式为:

$$AR = \frac{TP}{TP + FN} \quad (2)$$

式(2)中 TP 是将正类预测为正类数, FN 是将正类预测为负类数漏报。

5 小目标检测的研究趋势

小目标检测是目标检测的一个难点,在现实生活中有重要应用价值,例如在刑侦领域,桌子上的小包裹,监控视频中角落的小行人,衣服上面的小标志等,都是破案的线索,小目标检测具有重要的研究意义。对于小目标检测的研究,因小目标本身携带的特征少等原因,小目标检测面临大的挑战。本节针对小目标研究的难点,指出了未来的研究方向。

(1) 充分利用网络的底层特征,进行多尺度检测。对于小目标检测的研究,小目标数据集中存在过大过小的物体,检测图片的尺度变换较大。卷积神经网络底层特征语义信息弱、分辨率高,高层特征语义信息强、分辨率低,但底层特征对检测小目标很重要。可以通过特征融合,使得卷积神经网络的所有尺度都有强语义信息,建立更加完善的特征金字塔网络^[53]。

(2) 提高小目标的分辨率。小目标的尺寸小,特征不明显,不容易提取,可以通过生成对抗网络等图像超分辨率方法,对小目标生成高分辨率表示。充分提取小目标的特征^[31],缩小小目标与大目标之间的特征表征,进而提高小目标检测的性能。

(3) 同时提高特征分辨率和感受野。特征分辨率越低就越难检测小目标,可以通过降低下采样率来提高特征的分辨率,但与此同时,会使感受野变小,不利于较大物体的检测。可以使用膨胀卷积等方法在提高特征分辨率的同时,提高感受野^[54]。

(4) 利用注意力机制合理化分配资源。资源指模型的大小、训练模型时占用的硬件、需要的训练时间等,小目标检测会分配较多资源用于候选框的生成。视觉注意力机制的使用会使资源更合理化分配,注意力机制的本质是快速找到感兴趣区域,忽略不重要的信息。在未来的小目标检测研究中,可以利用注意力机制有效地选择信息,对感兴趣区域分配更多资源进行更精确的检测。

(5) 提高小目标的训练数据量。数据扩充是训练深度学习模型的关键组成部分,注释图像对于小目标检测的成本较高,一个优化的数据扩充策略可以有效的提高小目标检测效果^[55,56]。目前小目标研究主要针对普通目标数据集,网络并不能深入学习小目标的特征。今后可建立针对于小目标检测的数据集,充分学习小目标特征,更有利于小目标的检测。

6 结束语

目标检测算法目前已经有了很大研究进展,而对于小目标检测的研究近几年才开始,算法还比较少.本文对小目标检测领域的现有研究成果进行了归类总结,对比了各类算法的性能并分析了其优缺点;其次,介绍了小目标检测在各个不同领域的应用;此外,总结了现有文献中小目标检测常用的数据集及算法性能评价指标.

参考文献

- [1] KISANTAL M, WOJNA Z, MURAWSKI J, et al. Augmentation for Small Object Detection [DB/OL]. arXiv: 1902.07296v1, 2019.
- [2] wq604887956. 小目标检测文章总结 [EB/OL]. <https://blog.csdn.net/wq604887956/article/details/83053927>, 2018-10-16/2019-09-02.
- [3] LIN T Y, MAIRE M, BELONGIE S, et al. Microsoft COCO: Common objects in context [A]. The 13th European Conference on Computer Vision (ECCV) [C]. Cham: Springer, 2014. 740 – 755.
- [4] BODLA N, SINGH B, CHELLAPPA R, et al. Soft-NMS—Improving object detection with one line of code [A]. 2017 IEEE International Conference on Computer Vision [C]. New York, USA: IEEE, 2017. 5561 – 5569.
- [5] LI Y H, CHEN Y T, WANG N Y, et al. Scale-aware Trident Networks for Object Detection [DB/OL]. arXiv: 1901.01892, 2019.
- [6] 王海屹. 基于特征融合的小物体目标检测研究与应用 [D]. 北京: 北京邮电大学, 2018.
- [7] EGGERT C, BREHM S, WINSCHER A, et al. A closer look: Small object detection in faster R-CNN [A]. 2017 IEEE International Conference on Multimedia and Expo (ICME) [C]. Hong Kong, China: IEEE Computer Society, 2017. 421 – 426.
- [8] 刘晓楠, 王正平, 贺云涛, 等. 基于深度学习的小目标检测研究综述 [J]. 战术导弹技术, 2019, 193(01): 106 – 113.
LIU X N, WANG Z P, HE Y T, et al. Review of small object detection based on deep learning [J]. Tactical Missile Technology, 2019, 193(01): 106 – 113. (in Chinese)
- [9] TAKEKI A, TU T T, YOSHIHASHI R, et al. Combining deep features for object detection at various scales: finding small birds in landscape images [J]. IPSJ Transactions on Computer Vision & Applications, 2016, 8(1): 5.
- [10] KAMPFFMEYER M, Salberg A B, Jenssen R. Semantic segmentation of small objects and modeling of uncertainty in urban remote sensing images using deep convolutional neural networks [A]. 2016 IEEE Conference on Computer Vision and Pattern Recognition Workshops [C]. Las Vegas, NV: IEEE, 2016. 680 – 688.
- [11] MUNDHENK T N, Konjevod G, Sakla W A, et al. A large contextual dataset for classification detection and counting of cars with deep learning [A]. The 14th European Conference on Computer Vision [C]. Cham: Springer, 2016. 785 – 800.
- [12] REN Y, ZHU C R, XIAO S P. Small object detection in optical remote sensing images via modified faster R-CNN [J]. Applied Sciences, 2018, 8(5): 813.
- [13] EGGERT C, ZECHA D, BREHM S, et al. Improving small object proposals for company logo detection [A]. Proceedings of The 2017 ACM on International Conference on Multimedia Retrieval [C]. New York, USA: ACM, 2017. 167 – 174.
- [14] 于进勇, 丁鹏程, 王超. 卷积神经网络在目标检测中的应用综述 [J]. 计算机科学, 2018, 45(S2): 27 – 36.
YU J Y, DING P C, WANG C. Summary of application of convolutional neural network in object detection [J]. Computer Science, 2018, 45(S2): 27 – 36. (in Chinese)
- [15] HARIHARAN B, Arbeláez P, GIRSHICK R, et al. Hypercolumns for object segmentation and fine-grained localization [A]. 2014 IEEE Conference on Computer Vision and Pattern Recognition [C]. Boston, MA, USA: IEEE, 2014. 447 – 456.
- [16] 李旭冬, 叶茂, 李涛. 基于卷积神经网络的目标检测研究综述 [J]. 计算机应用研究, 2017, 34(10): 2881 – 2886.
LI X D, YE M, LI T. Summary of research on object detection based on convolutional neural network [J]. Computer Applied Research, 2017, 34(10): 2881 – 2886. (in Chinese)
- [17] 周俊宇, 赵艳明. 卷积神经网络在图像分类和目标检测应用综述 [J]. 计算机工程与应用, 2017, 53(13): 34 – 41.
ZHOU J Y, ZHAO Y M. Summary of application of convolutional neural network in image classification and object detection [J]. Computer Engineering and Application, 2017, 53(13): 34 – 41. (in Chinese)
- [18] 崔广新, 李殿奎. 基于自编码算法的深度学习综述 [J]. 计算机系统应用, 2018, 27(09): 47 – 51.
CUI G X, LI D K. A survey of deep learning based on self coding algorithm [J]. Computer System Application, 2018, 27(09): 47 – 51. (in Chinese)
- [19] 徐毅, 李蓓蓓, 宋威. 改进的深度置信网络分类算法研究 [J]. 计算机科学与探索, 2019, 13(04): 596 – 607.
XU Y, LI B B, SONG W. Research on improved depth confidence network classification algorithm [J]. Computer Science and Exploration, 2019, 13(04): 596 – 607. (in Chinese)
- [20] HE K M, ZHANG X Y, REN S Q, et al. Deep residual

- learning for image recognition [A]. Proceedings of The 2016 IEEE Conference on Computer Vision and Pattern Recognition [C]. USA: IEEE Computer Society, 2016. 770 – 778.
- [21] NEWELL A, YANG K, DENG J. Stacked hourglass networks for human pose estimation [A]. The 14th European Conference on Computer Vision [C]. Cham: Springer, 2016. 483 – 499.
- [22] REDMON J, DIVVALA S, and GIRSHICK R, et al. You only look once: unified, real-time object detection [A]. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) [C]. Las Vegas, NV: IEEE, 2016. 779 – 788.
- [23] LIU W, ANGUELOV D, ERHAN D, et al. SSD: Single shot multiBox detector [A]. The 14th European Conference on Computer Vision [C]. Cham: Springer, 2016. 21 – 37.
- [24] FU C Y, LIU W, RANGA A, et al. DSSD: Deconvolutional Single Shot Detector [DB/OL]. arXiv: 1701.06659, 2017.
- [25] LIN T Y, GOYAL P, GIRSHICK R, et al. Focalloss for dense object detection [J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2017, PP(99) : 2999 – 3007.
- [26] LAW H, TENG Y, RUSSAKOVSKY O, et al. CornerNet-Lite: Efficient Keypoint Based Object Detection [DB/OL]. arXiv: 1904.08900, 2019.
- [27] LIN T Y, Dollár, and PIOTR, et al. Feature pyramid networks for object detection [A]. 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) [C]. Honolulu, HI: IEEE Computer Society, 2017. 2117 – 2125.
- [28] HUANG J, RATHOD V, SUN C, et al. Speed/accuracy trade-offs for modern convolutional object detectors [A]. 2016 IEEE Conference on Computer Vision and Pattern Recognition [C]. New York, USA: IEEE, 2016. 3296 – 3297.
- [29] DAI J F, QI H Z, XIONG Y W, et al. Deformable Convolutional Networks [DB/OL]. arXiv: 1703.06211, 2017.
- [30] HE K M, GKIOXARI G, Dollár P, et al. Mask R-CNN [A]. 2017 IEEE International Conference on Computer Vision [C]. Venice: IEEE Computer Society, 2017. 2980 – 2988.
- [31] LI Z M, PENG C, YU G, et al. DetNet: A Backbone Network for Object Detection [DB/OL]. arXiv: 1804.06215, 2018.
- [32] SINGH B and DAVIS L S. An analysis of scale invariance in object detection-SNIP [A]. 2018 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) [C]. New York, USA: IEEE Computer Society, 2018. 3578 – 3587.
- [33] SINGH B, NAJIBI M, DAVIS L S. SNIPER: Efficient-multi-scale Training [DB/OL]. arXiv: 1805.09300, 2018.
- [34] CAI Z W, VASCONCELOS N. Cascade R-CNN: Delving into high quality object detection [A]. 2018 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) [C]. New York, USA: IEEE Computer Society, 2018. 6154 – 6162.
- [35] CAI Z W, VASCONCELOS N. Cascade R-CNN: High Quality Object Detection and Instance Segmentation [DB/OL]. arXiv: 1906.09756, 2019.
- [36] 张亚须. 视频目标检测与跟踪算法研究及其在公安刑侦系统中的应用 [D]. 云南: 云南师范大学, 2018.
- [37] KRISHNA H, JAWAHAR C V. Improving small object detection [A]. 2017 4th IAPR Asian Conference on Pattern Recognition [C]. Nanjing: Nanjing University of Science and Technology (NJUST), 2017. 340 – 345.
- [38] HU P Y, RAMANAN D. Finding tiny faces [A]. 2017 IEEE Conference on Computer Vision and Pattern Recognition [C]. Honolulu, HI: IEEE, 2017. 1522 – 1530.
- [39] LIU W, LIAO S C, HU W D, et al. Learning efficient single-stage pedestrian detectors by asymptotic localization fitting [A]. European Conference on Computer Vision [C]. Munich, Germany: Springer, 2018. 618 – 634.
- [40] cLI J N, LIANG X D, WEI Y C, et al. Perceptual generative adversarial networks for small object detection [A]. 2017 IEEE Conference on Computer Vision and Pattern Recognition [C]. New York, USA: IEEE, 2017. 1951 – 1959.
- [41] YI K, JIAN Z Q, CHEN S T, et al. Knowledge-based Recurrent Attentive Neural Network for Traffic Sign Detection [DB/OL]. arXiv: 1803.05263, 2018.
- [42] 西安邮电大学图像处理团队小目标检测图像数据库 [DB/OL]. <http://www.xuptciip.com.cn/show.html?database=xksjk8>, 2018-11-01/2019-09-02.
- [43] EVERINGHAM M, WINN J. The PASCALvisual object classes challenge 2007 (VOC2007) development kit [J]. International Journal of Computer Vision, 2006, 111(1) : 98 – 136.
- [44] XIA G S, BAI X, DING J, et al. DOTA: A Large-scale Dataset for Object Detection in Aerial Images [DB/OL]. arXiv: 1711.10398, 2017.
- [45] ZHU H G, CHEN X G, DAI W Q, et al. Orientation robust object detection in aerial images using deep convolutional neural network [A]. 2015 IEEE International Conference on Image Processing (ICIP) [C]. Quebec City, QC, Canada : IEEE, 2015. 3735 – 3739.
- [46] CHENG G, HAN J W, ZHOU P C, et al. Multi-class geospatial object detection and geographic image classification based on collection of part detectors [J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2014, 98(1) : 119 – 132.
- [47] XIAO Z F, LIU Q, TANG G F, et al. Elliptic fourier transformation-based histograms of oriented gradients for rotationally invariant object detection in remote-sensing ima-

- ges [J]. International Journal of Remote Sensing, 2015, 36 (2): 618–644.
- [48] ZHU Z, LIANG D, ZHANG S H, et al. Traffic-sign detection and classification in the wild [A]. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) [C]. Las Vegas, NV, USA: IEEE Computer Society, 2016. 2110–2118.
- [49] LWJjing. Kitti 数据集简介 [EB/OL]. https://blog.csdn.net/weixin_43229119/article/details/89927707, 2019-05-07/2019-09-02.
- [50] ROMBERG S, PUEYO L G, LIENHART R, et al. Scalable logo recognition in real-world images [A]. Proceedings of the 1st International Conference on Multimedia Retrieval (ICMR) [C]. New York, USA: ACM, 2011. 25.
- [51] BELL S, ZITNICK C L, BALA K, et al. Inside-outside net: Detecting objects in context with skip pooling and recurrent neural networks [A]. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) [C]. Las Vegas, NV, USA: IEEE, 2016. 2874–2883.
- [52] LONG J, SHELHAMER E, DARRELL T. Fullyconvolutional networks for semantic segmentation [J]. 2014 IEEE Transactions on Pattern Analysis & Machine Intelligence, 2014, 39(4): 640–651.
- [53] SHEN Z Q, SHI H H, YU J H, et al. Learning Object Detectors from Scratch with Gated Recurrent Feature Pyramids [DB/OL]. arXiv: 1712.00886, 2017.
- [54] YU F, KOLUN V. Multi-scale Context Aggregation by Dilated Convolutions [DB/OL]. arXiv: 1511.07122, 2015.
- [55] ZOPH B, CUBUK E D, GHIASI G, et al. Learning Data Augmentation Strategies for Object Detection [DB/OL]. arXiv: 1906.11172, 2019.
- [56] KHOREVA A, BENENSON R, ILG E, et al. Lucid data dreaming for video object segmentation [J]. International Journal of Computer Vision, 2019, 127(9): 1175–1197.

作者简介



刘 颖(通信作者) 女, 1972 年 1 月生于陕西户县。2007 年获得澳大利亚莫纳什大学博士学位。现为西安邮电大学教授、研究生导师, 图像处理团队负责人及电子信息现场勘验应用技术公安部重点实验室总工程师, 主要研究方向为图像检索及图像清晰化等。
E-mail: liuying_ciiip@163.com



刘红燕 女, 1996 年 5 月生于山西省平遥县。现为西安邮电大学通信与信息工程学院硕士研究生, 主要研究方向为小目标检测。
E-mail: hy_liu1314@163.com



范九伦 男, 1964 年 11 月生于河南省温县。教授, 博士生导师。1998 年获西安电子科技大学博士学位, 现为西安邮电大学校长, 主要研究方向为模糊信息处理、图像处理、信息安全等。
E-mail: jiulunf@xupt.edu.cn



公衍超 男, 1986 年 1 月生于山东省蒙阴县。2017 年获西北工业大学博士学位, 现为西安邮电大学讲师, 图像与信息处理研究所科研成员。主要研究方向为视频编码及码率控制等。
E-mail: gongyanchao@xupt.edu.cn



李莹华 女, 1988 年 6 月生于陕西省宁强县。2019 年获西安电子科技大学博士学位, 现为西安邮电大学讲师, 图像与信息处理研究所科研成员。主要研究方向为图像安全与图像超分辨率重建等。
E-mail: liyinghua@xupt.edu.cn



王富平 男, 1987 年 7 月生于陕西省兴平县。2017 年获西安电子科技大学博士学位, 现为西安邮电大学讲师, 图像与信息处理研究所科研成员。主要研究方向为图像特征提取及图像检索等。
E-mail: wfp1608@163.com



卢 津 男, 1987 年 5 月生于陕西省礼泉县。2018 年获西北工业大学博士学位, 现为西安邮电大学讲师, 图像与信息处理研究所科研成员。主要研究方向为三维痕迹图像数据采集及处理等。
E-mail: lujin@xupt.edu.cn