

基于深度学习的光学遥感图像目标检测研究进展

刘小波^{1,2,3} 刘鹏^{1,2} 蔡之华^{4,5} 乔禹霖^{1,2} 王凌^{4,6} 汪敏^{1,2}

摘要 光学遥感图像的目标检测 (Optical remote sensing images object detection, ORSIOD) 是航空和卫星图像分析领域的一个基本但具有挑战性的问题, 近年来受到广泛关注. 本文从如下几个方面介绍了基于深度学习的光学遥感图像目标检测的研究现状. 首先对光学遥感图像目标检测的主要难点进行了介绍, 接着对现有基于深度学习的目标检测算法进行概括, 并以光学遥感图像目标检测的难点为驱动分析对比了不同的基于深度学习的光学遥感图像目标检测方法的优缺点, 最后对未来的发展趋势进行了详细的分析.

关键词 深度学习, 光学遥感, 目标检测, 卷积神经网络

引用格式 刘小波, 刘鹏, 蔡之华, 乔禹霖, 王凌, 汪敏. 基于深度学习的光学遥感图像目标检测研究进展. 自动化学报, 2021, 47(9): 2078–2089

DOI 10.16383/j.aas.c190455

Research Progress of Optical Remote Sensing Image Object Detection Based on Deep Learning

LIU Xiao-Bo^{1,2,3} LIU Peng^{1,2} CAI Zhi-Hua^{4,5} QIAO Yu-Lin^{1,2} WANG Ling^{4,6} WANG Min^{1,2}

Abstract Optical remote sensing images object detection (ORSIOD) is a basic but challenging question in the field of aerial and satellite image analysis. In recent years, it has received extensive attention. In this paper we introduces The current research status of the deep learning based ORSIOD from the following aspects. Firstly, the main difficulties of optical remote sensing image object detection are introduced, and then the existing deep learning based object detection algorithm is summarized, and according to different difficulties the deep learning based ORSIOD algorithm faced. The advantages and disadvantages of the deep learning based ORSIOD method are compared and analyzed, and finally the future development trend is detailed analysis.

Key words Deep learning, optical remote sensing images, object detection, convolutional neural network (CNN)

Citation Liu Xiao-Bo, Liu Peng, Cai Zhi-Hua, Qiao Yu-Lin, Wang Ling, Wang Min. Research progress of optical remote sensing image object detection based on deep learning. *Acta Automatica Sinica*, 2021, 47(9): 2078–2089

收稿日期 2019-06-12 录用日期 2019-10-11

Manuscript received June 12, 2019; accepted October 11, 2019

国家自然科学基金 (61973285, 61873249, 61773355, 61603355), 湖北省自然科学基金 (2018CFB528), 地质探测与评估教育部重点实验室主任基金 (CUG2019ZR10), 中央高校基本科研业务费 (CUGL17022) 资助

Supported by National Natural Science Foundation of China (61973285, 61873249, 61773355, 61603355), Natural Science Foundation of Hubei Province (2018CFB528), Opening Fund of Ministry of Education Key Laboratory of Geological Survey and Evaluation (CUG2019ZR10), and Fundamental Research Funds for the Central Universities (CUGL17022)

本文责任编辑 桑农

Recommended by Associate Editor SANG Nong

1. 中国地质大学 (武汉) 自动化学院 武汉 430074 2. 复杂系统先进控制与智能自动化湖北省重点实验室 武汉 430074 3. 地质探测与评估教育部重点实验室 武汉 430074 4. 中国地质大学 (武汉) 计算机学院 武汉 430074 5. 北部湾大学北部湾大数据资源利用实验室 钦州 535011 6. 清华大学自动化系 北京 100084

1. School of Automation, China University of Geosciences, Wuhan 430074 2. Hubei Key Laboratory of Advanced Control and Intelligent Automation for Complex Systems, Wuhan 430074

3. Ministry of Education Key Laboratory of Geological Survey and Evaluation, Wuhan 430074 4. School of Computer Science, China University of Geosciences, Wuhan 430074 5. Beibu Gulf Big Data Resources Utilisation Laboratory, Beibu Gulf University, Qinzhou 535011 6. Department of Automation, Tsinghua University, Beijing 100084

光学遥感^[1], 是指传感器工作波段在可见光波段, 也就是 $0.38 \sim 0.76 \mu\text{m}$ 范围的遥感技术, 这个范围的波段是传统航空摄影侦查和航空摄影测绘中最常用的工作波段. 随着卫星数量的增多、重访周期的缩短、图像分辨率的提高, 遥感图像数据量越来越大, 且随着新卫星的升空, 数据量将会不断增加. 剧增的数据量为科学研究提供了很多可利用的信息, 但是要利用好遥感图像还有一定的难度, 特别是受天气、光照、海况、传感器参数等多种因素影响, 使得光学遥感图像的特征提取较为困难.

光学遥感图像目标检测是确定在给定的航空或卫星图像中是否包含一个或多个感兴趣的物对象, 并对对象进行定位的任务. 作为航空和卫星图像分析领域的基本问题, 光学遥感图像中的目标检测在环境检测等领域有着广泛的应用, 如检测、地质灾害探测、土地利用和土地覆盖 (Land-use and land-cover, LULC) 制图、地理信息系统 (Geographic information system, GIS) 更新、精准农业、城市规划等. 常见的光学遥感图像的检测目标有人

造物体, 如船只、车辆、房屋等, 以及非人造的自然环境中的物体, 如动物、树木、江河等。

传统的光学遥感图像目标检测算法都是以区域选择-特征提取-分类器为主线的方法^[2], 通过滑动窗口的方法来遍历整幅图像, 滑动窗口设置有不同的长宽比以增强对不同形状的目标的识别能力。接着对每一个滑动窗口进行特征提取, 这个阶段常用的特征有 Haar 特征、尺度不变特征变换 (Scale-invariant feature transform, SIFT) 匹配算法、方向梯度直方图 (Histogram of oriented gradient, HOG) 等, 都是基于图像的纹理、色彩、尺度等特征^[3-4], 特征的可解释性较强。最后通过传统机器学习分类器 (如支持向量机 (Support vector machine, SVM)、Adaboost 等) 对滑动窗口进行分类来达到目标检测的目的。传统的遥感图像目标检测算法有着严重的弊端, 由于滑动窗口法使用的是穷举法, 导致时间复杂度过高, 并且产生了很多冗余的窗口, 这也严重影响后续特征提取以及检测的速度和精度^[5]。并且由于目标的形态多样性、光照变化多样性^[6]、背景多样性等因素使得设计一个鲁棒性好的特征存在一定的难度, 而且所提取特征的好坏直接影响到分类的准确性^[7]。基于深度学习的光学遥感图像目标检测在 2016 年 Girshick 等^[8]首次将神经网络用于目标检测之后开始快速发展起来, 神经网络由于其强大的自动提取特征的能力, 使得其在遥感图像目标检测任务中被广泛运用。相比于传统光学遥感图像目标检测算法, 基于深度学习的方法解决了传统检测算法适应性不高、对背景模型的更新要求高、提取特征鲁棒性差和检测的实时性差等缺点^[9], 使检测模型在精度和速度方面都有了很大的提升。基于深度学习的光学遥感图像目标检测方法, 是计算机视觉检测领域的一个重大创新^[10]。这些算法的创新点都是把传统的计算机视觉和深度学习方法进行结合, 并取得了良好的检测效果。准确性和实时性是衡量目标检测系统性能的重要指标^[11], 也是一对矛盾体, 如何更好地平衡它们一直是目标检测算法研究的一个重要方向。随着深度学习的不断发展, 检测的精度和实效性也逐渐提升。因此, 基于深度学习的光学遥感图像目标检测算法得到了广大研究者的关注, 成为机器学习领域的热点话题之一。

光学遥感图像目标识别技术在表征学习以及检测方面取得了稳定进展, 特征提取由原来的手动提取特征转变成深度卷积神经网络 (Convolutional neural network, CNN) 来自动提取特征, 就提取效率而言是一种长远的进步。近年来的算法框架大部分都是基于深度学习的通用目标检测算法框架, 并试图通过在一个或多个方面进行改进来达到在特定

数据集实现目标检测。大致上, 这些算法可以分为两大类, 第 1 类是基于区域的算法 (Two-stage), 其将检测问题划分为两个阶段, 首先产生区域提议 (Region proposal), 然后对候选区域分类 (一般还需要对位置精修), 该类算法的特点是精度高但速度相对慢。第 2 类是端到端的算法 (One-stage), 这是一种整体性更高的方法, 不需要区域提议阶段, 直接产生目标的类别概率和位置坐标值, 该类算法的特点是速度快但精度相对低。基于深度学习的光学遥感图像目标检测方法都是基于这两类方法而演变得来的。

本文的其余部分安排如下: 第 1 节对光学遥感图像目标检测目前存在的难点进行介绍, 第 2 节对基于深度学习的目标检测方法进行了概括介绍。第 3 节以不同难点为驱动, 分析比较了基于深度学习的光学遥感图像目标检测不同算法的优缺点。第 4 节进行总结与展望。

1 光学遥感图像目标检测面临的挑战

光学遥感图像不同于普通光学图像, 其差异性主要体现在传感器和拍摄角度上。随着遥感技术的发展, 遥感设备所搭载的传感器分辨率以每 10 年一个数量级的速度提升, 高分辨率甚至超高分辨率成为了新一代遥感卫星的发展趋势, 同时, 高的空间覆盖能力也在同步加强。高分辨率和高覆盖范围在带来了巨大信息量的同时, 也给遥感图像目标检测领域带来了全新的挑战。一方面, 尽管基于深度学习的方法与传统方法相比有着很强的特征提取能力, 但是其良好的表现基于大量被标记样本的重复迭代学习以及较高的计算消耗, 面对如今超高分辨率的遥感图像 (如 GF-2 数据集图像的分辨率为 27 620×29 200 像素), 基于深度学习的方法难以高效地进行检测。另一方面, 相对于巨幅遥感图像, 部分目标类别 (如车辆、房屋等) 在整幅图像所占像素则偏小, 仅有十几个像素, 即使性能优异的深度学习方法也难以达到理想的检测效果。普通光学图像所拍摄的事物大多分布在垂直方向, 并且目标往往会占据整幅图的大面积区域, 整幅图像的语义信息较为简单。而光学遥感图像的语义信息较为复杂, 地表上空的拍摄角度所带来的问题有: 目标方向以及尺寸多变、目标密集排列出现、复杂的背景区域占据整幅图的较大区域等。图 1 分别展示了光学遥感图像目标检测所遇到的几个问题的典型案例。

2 基于深度学习的目标检测方法

2.1 基于区域的深度学习目标检测方法

对于传统目标检测算法中滑动窗口造成大量信

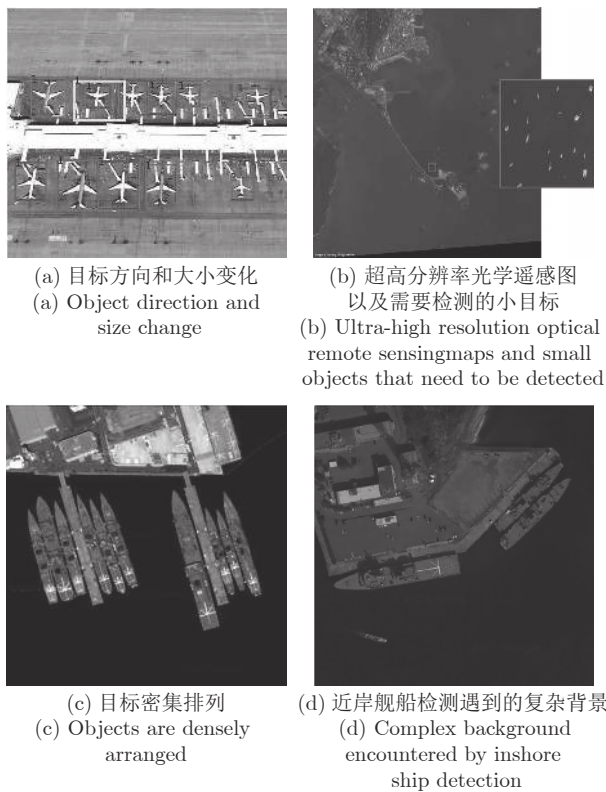


图 1 光学遥感图像目标检测所遇到的几个问题的典型案例图

Fig.1 A typical case diagram showing several problems encountered in optical remote sensing image target detection

息冗余的问题, 候选区域提供了很好的解决方案^[12]. 区域提议利用了图像中的纹理、边缘、颜色等信息预先找出图中目标可能出现的位置, 可以保证在选取较少窗口(几千个甚至几百个)的情况下保持较高的召回率^[13]. 这大大降低了后续操作的时间复杂度, 并且获取的候选窗口要比滑动窗口的质量更高. 这类算法的典型代表是基于候选区域的 R-CNN (Region-based convolutional neural network)^[8] 系列算法, 如 R-CNN、Fast R-CNN^[14]、Faster R-CNN^[15] 等, 图 2 为 Two-stage 网络的算法流程图.

目前, Two-stage 通用目标检测算法有如下几种: 1) R-CNN, 由 3 个部分组成, 第 1 部分使用一种称为“选择性搜索”的算法扫描可能是目标的输入图像, 生成约 2 000 个区域推荐; 第 2 部分在这些区域推荐上使用 CNN; 第 3 部分取每个 CNN 的输出使用支持向量机进行分类和线性回归来进行目标检测. 2) Fast R-CNN, 通过两个方面的增强提高了检测速度, 第 1 个是在推荐区域之前在图像上执行特征提取, 因此在整个图像上仅运行一个 CNN

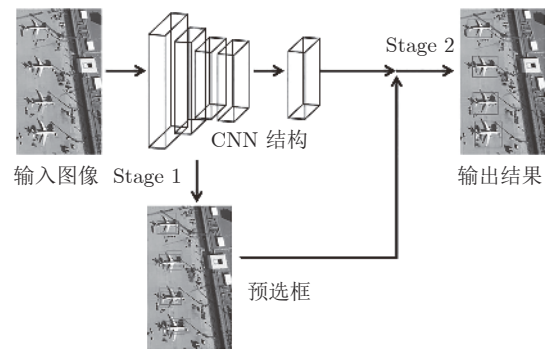


图 2 Two-stage 算法流程图

Fig.2 Two-stage algorithm flow chart

而不是在超过 2 000 个重叠区域运行 2 000 个 CNN; 第 2 个是用 Softmax 层代替 SVM, 从而扩展神经网络进行预测, 而不是创建一个新的模型. 3) Faster R-CNN, 使用快速神经网络来代替慢速的选择性搜索算法, 提出了区域提议网络 (Region proposal network, RPN) 来获取候选区域, 取得了更快的检测速度和更好的检测精度. 4) R-FCN (Region fully convolutional neural network), 采用 100% 共享计算的全卷积设计, 并提出了位置敏感分数图 (Position-sensitive score maps), 通过将感兴趣的区域 (Region of interest, ROI) 进行划分, 并给每个区域进行类别和位置的评估来最终确定整个目标的类别和位置, 最终在速度和精度上具有比 Faster-RCNN 更好的效果. 5) Mask-RCNN^[16], 在网络的第 2 阶段, Mask-RCNN 加入了一个全卷积神经网络的分支, 它能对每个感兴趣区域输出一个二进制的掩膜 (Mask) 从而提取精细的空间布局, 实现了像素级别的预测.

2.2 基于端到端的深度学习目标检测方法

自 R-CNN 以来, 基于区域的两阶段网络在检测问题上一直占有优势. 近期, 随着研究者的共同努力, R-CNN 系列算法在精度和效率上具有很好的表现, 但是检测速度仍有待提高. 很多研究学者已经开始从研究两阶段网络效率问题转向了整体性更高的单阶段网络的精度问题. One-stage 网络泛指直接进行分类预测和边界框回归的算法结构, 没有像两阶段网络一样的区域提议阶段和分类阶段. One-stage 网络与 Two-stage 网络相比结构更简单, 它将所有计算封装在一个网络中. 整个检测流程是一个单一的网络, 可以在检测性能上进行端到端的优化. 图 3 为 One-stage 算法的流程图.

目前, One-stage 通用目标检测算法有如下几

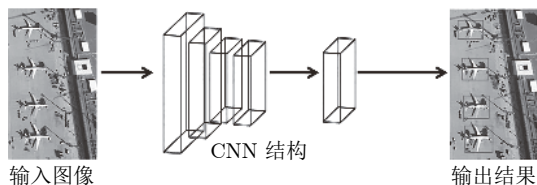


图 3 One-stage 算法流程图

Fig.3 One-stage algorithm flow chart

种: 1) YOLOv1 (You only look once v1)^[17], 由于区域提议生成阶段被完全取消, YOLOv1 直接使用一小部分候选区域来预测检测. 与 Two-stage 的方法不同, 例如 Faster R-CNN, 从局部区域的特征来预测目标, YOLOv1 全局地使用来自整个图像的特征. 特别是, YOLOv1 将图像划分为 $S \times S$ 个网格. 每个网格预测 C 类概率、 B 个边界框位置和这些框的置信度得分. 这些预测被编码为 $S \times S \times (5B + C)$ 的张量. 通过完全放弃区域推荐生成步骤, YOLOv1 在设计上是运行效率高的, 可以在 45 帧/s 上实时运行. 2) YOLOv2^[18], YOLOv1 的改进版本, 将 YOLOv1 中的 GoogleNet^[19] 换成了更简单的 DarkNet-19, 此外, 还利用了许多从现有工作中得出的策略, 如批归一化、删除全连接层、以及使用从 K 均值和多尺度中学到好的锚框. 3) YOLOv3, 相比 V2 版本, 主要进行了如下几方面的改进, 首先是网络结构的改进, 网络的结构由 Darknet-19 变为 Darknet-53, 跳跃连接的现象越来越普遍. 接着是末端激活函数的改进: 尾部的激活函数 (类别预测) 由 Softmax 改为 Sigmoid. 最后是锚框数量的改变: Anchor box 的数量由 5 个改为 3 个. 4) SSD (Single shot multiBox detector)^[20], 其结合了 Faster R-CNN、YOLO 以及多尺度卷积特征的思想. 与 YOLO 类似, SSD 对在边界框中的目标实例进行预测时得出可能的目标的边界框以及评价, 再通过非最大值抑制算法 (Non-maximum suppression, NMS) 算法得出最可能的一个目标类. SSD 采用全卷积神经网络, 通过在多个卷积特征映射上操作来执行多个尺度上的检测, 每个映射都预测适当大小的边界框的类别分数和边界框偏移量.

注 1. YOLOv3 是一个未正式发表的算法, 可参见文献: Adam V. You only look twice: Rapid multi-scale object detection in satellite imagery. arXiv: 1805.09512v1[cs.CV]. 2018.

3 基于深度学习的光学遥感图像目标检测方法研究现状

深度学习因其更强大的泛化能力, 在光学遥感

图像目标检测领域已经得到广泛运用. 光学遥感图像由于其传感器以及拍摄角度的特殊性, 与普通图像相比有诸多差异. 尽管有着强大的分歧特征提取能力, 受限于光学遥感图像的超高分辨率和超大覆盖面积所带来的新问题, 传统基于深度学习的目标检测算法难以取得理想的效果.

目前已经有许多学者正在尝试使用深度学习的方法来解决光学遥感图像目标检测, 这些创新性的算法结构都以传统的深度学习方法为基础. 本节将以深度学习方法的光学遥感图像目标检测所面临的问题为导向, 进行不同算法的归纳和总结.

3.1 面向光学遥感图像分辨率过高问题的深度学习目标检测方法

深度学习方法的强大性能是以牺牲计算速度为代价的, 其训练过程需要反复迭代才能达到较好效果. 面对超高分辨所带来的超大数据量, 传统深度学习的方法很难在保证检测质量的同时快速的学习其中的重要信息. 表 1 对当前解决图像分辨率过高问题的部分算法进行了比较.

文献 [21-27] 将原始尺寸的大图切割为一系列的小图进行处理, 不仅使得网络有了处理这些数据的能力, 同时也放大了图片, 提高了检测效果, 但是该方法容易使得位于小图片边缘区域的目标被切割. 文献 [28] 借鉴了卷积神经网络重叠区块的思想, 以一定重叠率对原图进行滑动切割, 避免了小图边缘区域目标被切割的问题. 但是滑动取块使得整幅大图的处理时间变得更加冗余, 文献 [29] 发现滑动取块法会造成信息冗余的原因是大图中部分的滑动窗口区域并没有包含有需要被检测的目标, 因此, 利用 RCNN 网络的两阶段检测特点, 在使用 20% 重叠率的大图取块后, 首先使用第一阶段的定位器对不包含有目标的凸显块进行过滤, 从而减小第二阶段的分类器的计算负担, 该方法在保证检测效果的同时, 减少了多余的计算消耗. 但是对于重叠部分可能会出现多次检测的问题, 因此算法 YOLT (You only look twice) 在使用 15% 重叠率的滑动取块之后, 利用非极大值抑制来防止重叠部分的多次检测, 进一步减小了计算消耗. 但是将大图切割为小图进行计算始终不是一个高效的处理方式.

注 2. YOLT 是一个未正式发表的算法, 可参见文献: Adam V. You only look twice: Rapid multi-scale object detection in satellite imagery. arXiv: 1805.09512v1[cs.CV]. 2018.

文献 [30] 直接对原图进行处理, 使用全卷积神经网络进行候选区域映射 (每个像素预测是否为飞

机,如果是,则以此为中心,取 196×196 像素的框作为候选区域,最后使用非最大值抑制来减少重复框),提取可能存在目标的区域,然后训练一个分类器来验证结果,有效减少搜索区域,但是候选框的大小是基于先验信息决定的,面对尺寸多变的检测任务则难以有效处理。

3.2 面向光学遥感图像目标过小问题的深度学习目标检测方法

基于深度学习的目标检测算法大多以卷积神经网络作为主干网络,因为卷积神经网络在图像处理方面有着优异的表现。但是面对仅有几十或者几百个像素的小目标时,卷积神经网络反而成了劣势,这是由其下采样层造成的,小目标往往会在经过多层下采样层的深层网络后消失。表 2 对当前解决图像目标过小问题的部分算法进行了比较。

最直观的解决办法是重新扩大特征图尺寸,大多数文献通过加入反卷积层来扩大特征图,算法 YOLT 使用类似 ResNet (Residual neural network) 恒等映射的通过层 (Passthrough layer) 将最后一层网络输出层与中间层拼接,浅层网络的加入提高了小目标的检测能力,同时作者增加了 YOLO 的网格个数,使得每个网格能预测更小的目标,但是对小目标提取能力的提高有限,考虑到计算消耗,也不能无限增加网格个数。文献 [31] 以 Faster R-CNN 为基础网络,在进行区域推荐后加入反卷积层,重新放大小目标的尺寸,改善了小目标的检测效果。但是反卷积的作用体现在当深层网络特征图中依然存在少数小目标的像素点时,对于在深层网络中已经消失的小目标特征无法起到放大作用,并且该网络结构对于不同大小的目标检测鲁棒性差。文献 [25] 则利用反卷积层进行浅层和深层网络的多尺度的特征融合,不仅为特征提取引入了上下文

信息,同时还能兼顾大、小目标的检测。但是浅层网络对于小目标来说会引入过多的背景噪声,因此文献 [24] 在使用多尺度信息进行特征融合的同时,使用平衡系数来平衡加入的上下文信息所引入的噪声。但是平衡因子的选择没有很好的理论依据,只能凭经验选择,对不同数据集的鲁棒性较差。

反卷积的加入同时也会引入过多的参数,文献 [32] 在 YOLOv2 模型中加入了膨胀卷积,在扩大感受野的同时,相比反卷积减少了参数。但是膨胀卷积会造成局部信息丢失的问题,文献 [33] 使用像素级注意力机制来融合局部信息和全局信息,弥补了膨胀卷积的不足,同时,全局注意力将浅层和深层特征结合,改善了小目标的检测效果。但是其全局注意力是通过全局池化结合平均池化,然后再进行下采样获得的,同样面临 CNN 网络下采样所面临的问题。

文献 [27] 提供了不同的思路,采用由粗到精的检测方法,结合两个 CNN 网络,第 1 阶段使用 CNN 进行范围较大的粗糙的目标检测(检测结果里只包含一个目标),第 2 阶段网络对第 1 阶段网络的粗预测结果再进行仔细的检测,缩小范围可以解决直接检测难以检测小目标的问题。但是两步检测都存在检测效率低的问题。

如何在保证大目标检测效果的同时提高小目标的检测效果,并且保证较快的检测速度,还需要更进一步研究。

3.3 面向光学遥感图像目标方向变化问题的深度学习目标检测方法

光学遥感图像由于其特殊的拍照角度,会更加凸显目标的方向性,而 CNN 的深度学习模型是不具有旋转不变性的,多变的角度会给检测带来困扰。目前解决目标方向变化的思路主要有两种:数据增强以及改变网络结构。表 3 对当前解决图像目标方

表 1 解决图像分辨率过高问题的不同方法对比

Table 1 Comparison of different methods to solve the problem of high image resolution

解决问题	论文来源	方法	优点	缺点/难点
图像分辨率过高	[21-27]	将大图切割为小图	使得网络能够对每个图像块进行处理提高检测效果	图像块边缘的目标被切割
	[28]	以一定重叠率切割	避免图像块边缘目标被切割	引入过多冗余信息
	[29]	在以一定重叠了切割后使用RCNN检测	RCNN 的第 1 阶段网络可以过滤掉冗余的背景信息	第 1 阶段网络仍然会受到冗余信息的影响
	YOLT	滑动取图像块之后,利用非极大值抑制来防止重叠部分的多次检测	减小重叠区域的重复检测	未能解决使用滑块方法造成的检测速度过慢问题
	[30]	使用全卷积神经网络将大图映射为较小特征图,每个像素对应固定尺寸的边界框	能够直接对整个大图进行处理	只能应对较大尺寸的目标

注: YOLT 是一个未正式发表的算法,参见文献: Adam V. You only look twice: rapid multi-scale object detection in satellite imagery. arXiv:1805.09512v1 [cs.CV]. 2018.

表 2 解决目标像素过少问题的不同方法对比

Table 2 Comparison of different methods to solve the problem of too few object pixels

解决问题	论文来源	方法	优点	缺点/难点
目标过小	YOLT	增加YOLOv3网络的网格数	提高了小目标的检测能力	降低检测速度
	[31]	加入反卷积层	对小目标进行放大	无法放大在反卷积层前已经消失的小目标
	[25]	用反卷积层来结合不同层的信息	兼顾大小目标的检测	浅层网络会引入过多噪声
	[24]	使用平衡系数来减小浅层网络中的噪声	提高小目标检测能力的同时减少背景影响	平衡系数不易定义
	[32]	在YOLOv2网络中加入膨胀卷积	扩大感受野同时减少参数	膨胀卷积会丢失局部信息
	[33]	使用像素级注意力机制	弥补了膨胀卷积的不足	其注意力机制经过多个池化层得来, 对于小目标不敏感

注: YOLT是一个未正式发表的算法, 参见文献: Adam V. You only look twice: Rapid multi-scale object detection in satellite imagery. arXiv:1805.09512v1 [cs.CV]. 2018.

表 3 解决目标方向变化问题的不同方法对比

Table 3 Comparison of different methods to solve the problem of object direction change

解决问题	论文来源	方法	优点	缺点/难点
目标方向变化	[22-26, 34]	使用几个不同角度进行旋转数据增强	易于实现	效果有限
	[35]	使用单应性变换来进行数据增强	比普通选择增强效果要好	属于数据增强范畴, 效果有限
	[36]	加入旋转不变层, 引入正则化约束	从增强网络能力上解决问题	正则化约束项难以定义
	[37]	预测阶段加入多角度锚框进行预测	对方向的变化目标可以更好的定位	锚框的角度固定, 不能很好适应实际情况
	[38]	采用全连接层增强旋转不变性	能够解决方向变化问题	采用全连接层会固定网络输入的尺寸
	[25]	加入方向预测分支	边界框可以以一定角度定位目标	没有解决网络对于方向变化的处理问题
	[39]	使用旋转ROI池化	解决普通ROI池化会引入背景噪声的问题	没有解决网络对于方向变化的处理问题

向变化问题的部分算法进行了比较。

文献 [22-26, 34] 分别使用几个不同的旋转角度, 对训练数据进行翻转扩充增强, 扩充不同角度的样本数量, 改善目标旋转对检测效果的影响。不同于普通的旋转变换, 文献 [35] 使用单应性变换 (Homography transform) 来进行数据增强, 其能够准确捕捉平面场景的范围变化, 达到更好的增强效果。

数据增强中的旋转增强是一种简单的改善目标方向变化问题的方法, 有着一定的效果, 但是旋转增强方式的旋转角度是固定的, 无法有效地解决网络对于目标方向变化的处理能力。文献 [36] 在框架的基础网络后加入旋转不变层, 通过引入正则化约束项来强制训练样本在旋转前后共享相似的特征, 从而达到旋转不变性。文献 [37] 在 Faster R-CNN 模型的 RPN 网络中加入多角度锚框来解决目标方向变化问题。文献 [38] 采用多尺度信息融合的全连接层来增强对于旋转和尺度变化的鲁棒性。

目标方向变化不仅给网络特征提取带来障碍, 同时给目标检测过程中的定位任务造成了困扰。尤其对于舰船这样细长的目标, 水平检测框难以达到精确的定位。为此, 文献 [25] 在网络检测阶段加入了方向预测分支, 使得目标检测框能按照目标不同

的方向进行精确定位。文献 [39] 在加入方向边界框的基础上, 使用基于旋转的 ROI 池化, 解决了普通的 ROI 池化对于舰船这样细长的目标会引入很多噪声的问题。文献 [22] 在使用方向边界框的基础上创新性地提出了船头方向预测。

3.4 面向光学遥感图像目标尺寸变化问题的深度学习目标检测方法

如今光学遥感图像目标检测的任务大多为多类别的检测任务, 这些类别中有相对小的目标如舰船、飞机、车辆、房屋建筑、储油罐等, 也有相对大的目标如篮球场、田径场、飞机场、网球场等。当大目标和小目标同时出现在检测任务当中时, 网络对于较大的尺寸变化难以做出很好的处理。表 4 对当前解决图像目标尺寸变化问题的部分算法进行了比较。

目前的研究主要通过加入多尺度信息来解决不同大小目标的检测问题, 文献 [30] 目标推荐区域调整为 3 个不同尺寸图, 并将 3 个图分别使用 3 个网络进行预测, 最终通过投票决定最后的目标类别, 使得不同尺寸目标在同一标准下分类预测。文献 [22] 使用稠密特征金字塔网络, 对不同层的特征图使用 RPN 网络, 最后通过卷积层映射到同一维度进行

融合, 结合了浅层网络的位置信息以及深层网络的语义信息, 解决了不同大小的目标检测问题. 文献 [40] 在 RCNN 基础网络的两个不同尺度特征图中加入了两个 RPN 网络, 来应对不同大小的目标提案. 文献 [28] 通过使用跳跃链接的自编码器多尺度结构来适应不同尺度的变化, 对不同尺度使用不同锚框进行预测, 解决目标尺度变化问题.

3.5 面向光学遥感图像目标密集排列问题的深度学习目标检测方法

部分目标类别如舰船、车辆、飞机等交通工具, 因人为因素会出现密集排列的问题, 这增加了目标定位和目标分类的难度. 表 5 对当前解决图像目标密集排列问题的部分算法进行了比较.

对于密集排列目标的定位问题, 文献 [22, 39] 通过带有方向的边界框检测解决了这个问题. 对于密集排列目标的区分问题, 文献 [41] 使用了一种局

部再识别机制, 利用网络中的一个尺度来重新识别每个目标块, 然后使用非最大值抑制过滤, 达到更精细的目标框定位, 从而减小边界框重叠. 算法 YOLT 对图像进行上采样, 以确保感兴趣的目标大且足够分散, 在提高小目标检测能力的同时改善了目标的密集排列检测问题.

3.6 面向光学遥感图像复杂背景问题的深度学习目标检测方法

光学遥感图像目标检测除了常见的目标类别 (标准数据集 NWPU VHR-10 包含有 10 类目标) 以外, 其他没有被检测的地物类别都被划分为背景, 而且由于图像覆盖范围广, 与普通图像相比背景所占比例较高, 这也给检测任务带来了困难. 表 6 对当前解决图像复杂背景问题的部分算法进行了比较.

靠岸舰船检测是受遥感图像复杂背景影响的一个典型问题. 靠岸船只的特点是紧靠船坞并且与船

表 4 解决目标尺寸变化问题的不同方法对比

Table 4 Comparison of different methods to solve the problem of object size change

解决问题	论文来源	方法	优点	缺点/难点
尺寸变化	[30]	使用三个网络预测不同尺寸的区域推荐	解决不同尺寸目标的分类问题	没有解决不同尺寸目标定位问题
	[22]	使用特征金字塔网络的不同层来获得区域推荐	解决不同尺寸目标的定位问题	层的选取需要视实际数据集而定
	[40]	使用网络的两个层分别使用RPN网络	解决不同尺寸目标的定位问题	两层不一定能够满足不同数据集的需要
	[28]	使用自编码器的多个层预测, 针对不同尺度使用不同尺寸的锚框	解决不同尺寸目标的检测问题	锚框的不同尺寸选取需要根据实际数据集而定

表 5 解决目标密集排列问题的不同方法对比

Table 5 Comparison of different methods to solve the problem that objects are densely arranged

解决问题	论文来源	方法	优点	缺点/难点
目标密集排列	[22, 39]	使用带有方向的边界框预测	解决密集目标定位问题	没有解决网络难以提取密集排列目标特征的问题
	[41]	使用局部再识别机制, 使用网络进行重复检测	减小因密集排列漏检的目标数	对密集排列严重的目标检测效果不够好
	YoIT	通过上采用来放大密集目标之间的间隙	将密集排列目标区分开	会增大图片分辨率

注: YOLT是一个未正式发表的算法, 参见文献: Adam V. You only look twice: Rapid multi-scale object detection in satellite imagery. arXiv:1805.09512v1 [cs.CV]. 2018.

表 6 解决复杂背景问题的不同方法对比

Table 6 Comparison of different methods for solving complex background problems

解决问题	论文来源	方法	优点	缺点/难点
背景复杂	[42]	使用语义分割区分海洋和陆地	实现海陆分离, 避免陆地背景的影响	密集停靠的舰船与陆地特征相近, 不易分割
	[23]	将大量不包含有目标的陆地信息作为负样本进行训练, 使网络在检测过程中自动实现海陆分离	无需海陆分离步骤, 减少陆地背景对于近岸船只的检测	海陆分离的效果易受选取的背景样本的影响
	[43]	使用多尺度的视觉注意力机制	减少不同尺寸目标预测过程中复杂背景的影响	多尺度的引入增加了网络计算量
	[44]	在Faster R-CNN 网络中加入位置敏感分数图预测	通过不同局部位置的预测综合确定目标类别	对于相似多较高的背景类难以区分

坞平行. 从上空角度来看, 船坞的结构与舰船的结构混淆, 这也是造成靠岸舰船难检测的一个重要原因. 文献 [42] 使用全卷积网络来训练提取靠岸舰船、海洋以及陆地三类别的语义特征, 然后作为注意力机制来给各个类别事物加权, 提升了网络海陆分离的能力. 但是利用先验信息不具有普适性, 当船坞的舰船密集停靠的时候, 其特征与陆地特征相近, 难以进行海陆分割. 文献 [23] 将大量不包含目标的图片用于负样本进行训练, 同时训练一种综合性的损失, 该损失在当陆地上类似舰船的目标被识别为舰船时最大, 从而有效抑制陆地背景中与舰船有相似特征的地物的影响, 实现无需海陆分离的靠岸舰船识别.

除靠岸舰船检测之外, 还有其他受复杂背景影响的例子. 文献 [43] 使用多尺度特征提取方法, 将各层特征图单独使用视觉注意力机制增强并检测, 消除了不同大小目标检测过程中的背景噪声影响, 但多尺度分别预测造成了网络计算量的增大. 文献 [44] 结合 R-FCN 网络的位置敏感分数图以及 Faster R-CNN 网络的 RPN, 在使用 RPN 网络获得区域推荐之后, 在目标提案上使用位置敏感分数图, 实现目标局部位置的预测, 只有达到一定的局部特征相似度才能被判别为某一目标类, 消除了背景的影响, 但是该方法对于较为相似的背景类, 可能会出现误判, 对复杂背景的分辨能力不够高.

3.7 面向光学遥感图像样本不足问题的深度学习目标检测方法

样本不足问题是所有目标检测任务都要面临的问题. 面对高昂的图片获取的成本以及人工标记成本, 光学遥感图像目标检测数据集更加显得珍贵. 并且深度学习的方法对数据集的要求较高, 现有数据集的样本分布难以有效地训练模型. 目前解决训练样本不足的问题主要的思路有: 数据增强、弱监督学习以及迁移学习. 表 7 对当前解决图像样本不足问题的部分算法进行了比较.

文献 [24–27, 36, 44–46] 分别使用了包括平移、

缩放、旋转、镜像翻转、水平翻转、对比度增强、亮度调节、仿射变换等方法对训练数据进行扩充. 这些数据增强方法不仅成倍扩充了样本数量, 而且一定程度上解决了部分前面提到的问题 (如方向和尺度变化). 但是数据增强并没有从本质上解决网络本身性能不足的问题, 并且这些变换都是基于原始训练数据, 经过变换之后的样本相似度太高, 本质上讲仍然是原始样本, 因此对检测效果的提升有限.

弱监督学习分为不完全监督 (即关注于现有标记数据集不足以训练出好的模型)、不确切监督 (即关注于现有标记数据提供的信息不够准确)、不准确监督 (即关注于现有标记数据集的信息不总是真值). 光学遥感图像目标检测主要的问题是不完全监督. 文献 [47] 使用生成式对抗网络 (Generative adversarial network, GAN) 通过少量训练样本来进行虚拟样本生成. 相比数据增强方法, 其生成的样本与原始训练样本并不相同, 因此能更好的解决样本不足问题. 但是单独使用 GAN 进行预处理降低了网络的运行速度. 文献 [30] 在对模型迭代训练过程中通过挖掘难分负样本来迭代扩充包含不同场景类别的负样本, 改善由于复杂背景导致的误检问题, 但是仅鉴于负样本的扩充, 没有解决有标记正样本的不足问题.

深度学习方法因为需要对数据进行反复迭代学习, 使得样本数量的问题在深度学习方法上显得格外重要. 迁移学习作为解决训练集不足问题广泛运用在多个领域, 目前对于光学遥感图像目标检测任务, 其主要运用是使用源域数据对模型进行预训练, 然后将预训练模型的知识进行迁移, 再在目标域数据上进行进一步训练. 例如文献 [16] 使用 ImageNet 数据集对 YOLOv3 网络进行预训练, 然后再迁移至目标域数据进行微调. 文献 [41] 使用 ImageNet 数据集对主干网络进行预训练, 然后迁移至目标域数据进行微调. 这些方法的理论基础是深度神经网络 (Deep neural network, DNN) 的浅层是低级语义信息 (如边缘信息、颜色信息等), DNN 的深层是高级语义信息 (不同目标类别之间的分歧信息).

表 7 解决样本不足问题的不同方法对比

Table 7 Comparison of different methods for solving sample insufficient

解决问题	论文来源	方法	优点	缺点/难点
样本不足	[24–27, 36, 44–46]	使用包括平移、缩放、旋转、镜像翻转、水平翻转、对比度增强、亮度调节等方法的数据增强	易于实现样本扩充方法	效果有限, 扩充的数据本质上讲还是原来的数据
	[47]	使用 GAN 来进行虚拟样本生成	有效扩充样本数	时间消耗较大
	[30]	使用难分负样本挖掘来增强扩充包含有不同场景类别的负样本	减少复杂背景的影响	没有扩充正样本
	[16, 41]	使用 ImageNet 数据集对主干网络进行预训练	减少深度学习模型训练过程中对样本的依赖	源域数据与目标域数据相差较大

4 总结与展望

本文简要回顾了光学遥感图像目标检测的研究背景,介绍了传统的检测方法,并概括了基于深度学习的传统目标检测方法,接着对基于深度学习的光学遥感图像目标检测所面临的困难进行了归纳,并以此对目前基于深度学习的光学遥感图像检测方法进行了分类对比,并总结了目前的研究现状.随着遥感技术的发展,光学遥感图像的质量会有进一步提升,因此光学遥感图像的应用前景依然十分广阔,目前基于深度学习的光学遥感图像目标检测发展迅速,潜力巨大.使用基于深度学习的模型可以有效解放人力,将特征工程交于算法完成,极大提高网络对于不同数据的鲁棒性.

目前基于深度学习的光学遥感图像目标检测主要问题在第3节已经列出.尽管有很多研究致力于解决这些问题,并且取得了一定的效果,但是在寻找更好和更高效的解决办法上仍然留有很大的发挥空间.因此,在后续的研究应该考虑以下几个研究方向.

1) 在解决因图像分辨率过高造成网络计算负担增加的问题上,目前较多的方案是将大图切割成小图进行训练.这些方法在提高检测精度的同时引入了过多冗余的计算,因为原始图像中的某些图像块不包含目标或者包含有少量的目标,但是目前的网络会对这些块等同处理.如何选取有效的图像块使得目标尽可能多地落入该区域,如何优化网络使得网络能够将更多的注意力放在包含有目标的图像块,可以作为后续的一个研究方向.

2) 在解决因目标方向变化影响检测效果的问题上,数据增强可以作为后续研究的一个优化技巧被嵌入网络中,在算法优化上,考虑使用一种更加有效的旋转不变映射机制,改善基于CNN的深度神经网络对于旋转的鲁棒性.

3) 在目标尺寸变化问题上,网络的浅层对应更小的感受野,适用于提取小目标特征,网络的深层对应更大的感受野,适用于提取大目标特征.目前不同算法在解决目标尺寸变化问题上,都是基于这一理论使用多尺度信息来改善不同尺寸目标的检测效果.但是多尺度的选取仅仅对大和小进行了区分,不能适应不同数据集中的目标尺寸变化问题.后续的研究考虑加入多尺度自适应机制,针对不同数据集中目标大和小的具体尺寸情况来选取适合的多尺度信息.

4) 在解决因目标密集排列造成目标难定位问题上,目前的解决方法大多集中在引入方向边界框预测,对密集排列目标的定位问题进行了很好的解

决,但是同时目标密集排列还会造成单个目标信息难以提取的问题.考虑在网络中加入更加精准的目标分割技术,首先对密集排列目标进行图像分割,再使用网络对分割的目标进行特征提取.

5) 在解决因复杂背景影响检测效果的问题上,大多方法利用分割技术,将背景和目标区分开,从而改善网络对目标特征的提取能力,也有部分方法是通过提高网络对于背景和目标的分歧能力来减小复杂背景的影响.这些解决方法作用在算法的不同阶段,可以考虑使用一种有效的结合方法,共同用来解决复杂背景问题.

6) 在解决样本不足的问题上,数据增强作为一种易于实现的小技巧,可以继续用于后续新算法的研究.考虑针对不同类样本的分布情况以及不同的背景情况,使用弱监督学习方法来改善样本不平衡问题.在迁移学习的应用上,目前的研究都是使用通用目标检测数据集进行迁移.目标域与源域的数据有着很大差别,使得只能进行浅层低级语义信息的迁移.光学遥感图像有着其特殊拍摄角度,如果使用同为高空角度拍摄图像(如其他光学传感器拍摄的光学遥感图像、合成孔径雷达影像、高光谱影像等),其与目标域的高层语义特征会更加相似,因此可以考虑使用同为高空拍摄的数据进行更深层的迁移以减少网络对于源域数据的依赖.

References

- 1 Lin Yu-Dong. Target Detection in Optical Remote Sensing Images with Complicated Background [Ph.D. dissertation], Southwest Jiaotong University, China, 2017.
(林煜东. 复杂背景下的光学遥感图像目标检测算法研究 [博士学位论文], 西南交通大学, 中国, 2017.)
- 2 Liu Huan. Research on Detection Algorithm for Specific Object in High-Resolution Optical Remote Sensing Images [Master thesis], Harbin Institute of Technology, China, 2016.
(刘欢. 基于高分辨率光学遥感影像的特定目标检测算法研究 [硕士学位论文], 哈尔滨工业大学, 中国, 2016.)
- 3 Cao Xiao-Ming. Vehicle Detection based on a Multi-Channel Image Feature Pyramid [Master thesis], Beijing Jiaotong University, China, 2016.
(曹晓明. 基于多图像特征金字塔的车辆检测 [硕士学位论文], 北京交通大学, 中国, 2016.)
- 4 Zhang Gui-Mei, Zhang Song, Chu Jun. A new object detection algorithm using local contour features. *Acta Automatica Sinica*, 2014, **40**(10): 2346-2355
(张桂梅, 张松, 储珺. 一种新的基于局部轮廓特征的目标检测方法. *自动化学报*, 2014, **40**(10): 2346-2355)
- 5 Yin Hong-Peng, Chen Bo, Chai Yi, Liu Zhao-Dong. Vision-based object detection and tracking: A review. *Acta Automatica Sinica*, 2016, **42**(10): 1466-1489
(尹宏鹏, 陈波, 柴毅, 刘兆栋. 基于视觉的目标检测与跟踪综述. *自动化学报*, 2016, **42**(10): 1466-1489)
- 6 Ju Yu-Cui. Research on Key Algorithm of Object Detection and

- Tracking based on Visual [Master thesis], Tianjin University of Technology, China, 2014.
(鞠玉翠. 基于视觉的目标检测和跟踪关键算法的研究 [硕士学位论文], 天津理工大学, 中国, 2014.)
- 7 Wang Yan-Qing, Ma Lei, Tian Yuan. State-of-the-art of ship detection and recognition in optical remotely sensed imagery. *Acta Automatica Sinica*, 2011, **37**(9): 1029–1039
(王彦清, 马雷, 田原. 光学遥感图像舰船目标检测与识别综述. 自动化学报, 2011, **37**(9): 1029–1039)
 - 8 Girshick R, Donahue J, Darrell T, Malik J. Region-based convolutional networks for accurate object detection and segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2016, **38**(1): 142–158
 - 9 Zhang Hao-Kui, Li Ying, Jiang Ye-Nan. Deep learning for hyperspectral imagery classification: The state of the art and prospects. *Acta Automatica Sinica*, 2018, **44**(6): 961–977
(张号逵, 李映, 姜晔楠. 深度学习在高光谱图像分类领域的研究现状与展望. 自动化学报, 2018, **44**(6): 961–977)
 - 10 Chen Yu-Si. Research on Moving Objects Detection and Tracking based on Background Subtraction with Illumination Robustness [Master thesis], Southwest Jiaotong University, China, 2011.
(陈雨丝. 基于背景差分的光照鲁棒性运动目标检测与跟踪技术研究 [硕士学位论文], 西南交通大学, 中国, 2011.)
 - 11 Wang Bin. Research and Implementation of Object Detection System Based on Optimized ViBE and HOG [Master thesis], Shenyang University of Technology, China, 2016.
(王彬. 基于改进的ViBE和HOG的运动目标检测系统研究与实现 [硕士学位论文], 沈阳工业大学, 中国, 2016.)
 - 12 Felzenszwalb P, McAllester D, Ramanan D. A discriminatively trained, multiscale, deformable part model. In: Proceedings of the 2008 IEEE Conference on Computer Vision and Pattern Recognition. Anchorage, AK, USA: IEEE, 2008. 1–8
 - 13 Sarikaya D, Corso J J, Guru K A. Detection and localization of robotic tools in robot-assisted surgery videos using deep neural networks for region proposal and detection. *IEEE Transactions on Medical Imaging*, 2017, **36**(7): 1542–1549
 - 14 Girshick R. Fast R-CNN. In: Proceedings of the 2015 IEEE International Conference on Computer Vision (ICCV). Santiago, Chile: IEEE, 2015. 1440–1448
 - 15 Ren S Q, He K M, Girshick R, Sun J. Faster R-CNN: Towards real-time object detection with region proposal networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, **39**(6): 1137–1149
 - 16 He K M, Gkioxari G, Dollár P, Girshick R. Mask R-CNN. In: Proceedings of the 2017 IEEE International Conference on Computer Vision (ICCV). Venice, Italy: IEEE, 2017. 2980–2988
 - 17 Redmon J, Divvala S, Girshick R, Farhadi A. You only look once: Unified, real-time object detection. In: Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas, NV, USA: IEEE, 2016. 779–788
 - 18 Redmon J, Farhadi A. YOLO9000: Better, faster, stronger. In: Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu, HI, USA: IEEE, 2017. 6517–6525
 - 19 Szegedy C, Liu W, Jia Y Q, Sermanet P, Reed S, Anguelov D, et al. Going deeper with convolutions. In: Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Boston, MA, USA: IEEE, 2015. 1–9
 - 20 Liu W, Anguelov D, Erhan D, Szegedy C, Reed S, Fu C Y, et al. SSD: Single shot MultiBox detector. In: Proceedings of Computer Vision–ECCV 2016. Amsterdam, The Netherlands: Springer, 2016. 21–37
 - 21 Yan Z G, Song X, Zhong H Y, Zhu X Z. Object detection in optical remote sensing images based on transfer learning convolutional neural networks. In: Proceedings of the 5th IEEE International Conference on Cloud Computing and Intelligence Systems (CCIS). Nanjing, China: IEEE, 2018. 935–942
 - 22 Yang X, Sun H, Sun X, Yan M L, Guo Z, Fu K. Position detection and direction prediction for arbitrary-oriented ships via multitask rotation region convolutional neural network. *IEEE Access*, 2018, **6**: 50839–50849
 - 23 He Y Q, Sun X, Gao L R, Zhang B. Ship detection without sea-land segmentation for large-scale high-resolution optical satellite images. In: Proceedings of the 2018 IEEE International Geoscience and Remote Sensing Symposium. Valencia, Spain: IEEE, 2018. 717–720
 - 24 Fu Y M, Wu F G, Zhao J S. Context-aware and depthwise-based detection on orbit for remote sensing image. In: Proceedings of the 24th International Conference on Pattern Recognition (ICPR). Beijing, China: IEEE, 2018. 1725–1730
 - 25 Li M J, Guo W W, Zhang Z H, Yu W X, Zhang T. Rotated region based fully convolutional network for ship detection. In: Proceedings of the 2018 IEEE International Geoscience and Remote Sensing Symposium. Valencia, Spain: IEEE, 2018. 673–676
 - 26 Schilling H, Bulatov D, Niessner R, Middelmann W, Soergel U. Detection of vehicles in multisensor data via multibranch convolutional neural networks. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2018, **11**(11): 4299–4316
 - 27 Li X B, Wang S J. Object detection using convolutional neural networks in a coarse-to-fine manner. *IEEE Geoscience and Remote Sensing Letters*, 2017, **14**(11): 2037–2041
 - 28 Wang C, Bai X, Wang S, Zhou J, Ren P. Multiscale visual attention networks for object detection in VHR remote sensing images. *IEEE Geoscience and Remote Sensing Letters*, 2019, **16**(2): 310–314
 - 29 Pang J M, Li C, Shi J P, Xu Z H, Feng H J. R²-CNN: Fast tiny object detection in large-scale remote sensing images. *IEEE Transactions on Geoscience and Remote Sensing*, 2019, **57**(8): 5512–5524
 - 30 Zhang F, Du B, Zhang L P, Xu M Z. Weakly supervised learning based on coupled convolutional neural networks for aircraft detection. *IEEE Transactions on Geoscience and Remote Sensing*, 2016, **54**(9): 5553–5563
 - 31 Zhang W, Wang S H, Thachan S, Chen J Z, Qian Y T. Deconv R-CNN for small object detection on remote sensing images. In: Proceedings of the 2018 IEEE International Geoscience and Remote Sensing Symposium. Valencia, Spain: IEEE, 2018. 2483–2486
 - 32 Liu W C, Ma L, Wang J, Chen H. Detection of multiclass objects in optical remote sensing images. *IEEE Geoscience and Remote Sensing Letters*, 2019, **16**(5): 791–795
 - 33 Ying X, Wang Q, Li X W, Yu M, Jiang H, Gao J, et al. Multi-attention object detection model in remote sensing images based on multi-scale. *IEEE Access*, 2019, **7**: 94508–94519
 - 34 Deng Z P, Sun H, Zhou S L, Zhao J P, Lei L, Zou H X. Fast multiclass object detection in optical remote sensing images us-

- ing region based convolutional neural networks. In: Proceedings of the 2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS). Fort Worth, TX, USA: IEEE, 2017. 858–861
- 35 Ji H, Gao Z, Mei T C, Li Y F. Improved faster R-CNN with multiscale feature fusion and homography augmentation for vehicle detection in remote sensing images. *IEEE Geoscience and Remote Sensing Letters*, 2019, **16**(11): 1761–1765
- 36 Cheng G, Zhou P C, Han J W. Learning rotation-invariant convolutional neural networks for object detection in VHR optical remote sensing images. *IEEE Transactions on Geoscience and Remote Sensing*, 2016, **54**(12): 7405–7415
- 37 Li K, Cheng G, Bu S H, You X. Rotation-insensitive and context-augmented object detection in remote sensing images. *IEEE Transactions on Geoscience and Remote Sensing*, 2018, **56**(4): 2337–2348
- 38 Zhang Y L, Yuan Y, Feng Y C, Lu X Q. Hierarchical and robust convolutional neural network for very high-resolution remote sensing object detection. *IEEE Transactions on Geoscience and Remote Sensing*, 2019, **57**(8): 5535–5548
- 39 Liu Z K, Hu J G, Weng L B, Yang Y P. Rotated region based CNN for ship detection. In: Proceedings of the 2017 IEEE International Conference on Image Processing (ICIP). Beijing, China: IEEE, 2017. 900–904
- 40 Zhang S, He G H, Chen H B, Jing N F, Wang Q. Scale adaptive proposal network for object detection in remote sensing images. *IEEE Geoscience and Remote Sensing Letters*, 2019, **16**(6): 864–868
- 41 Zhao W, Ma W P, Jiao L C, Chen P H, Yang S Y, Hou B. Multi-scale image block-level F-CNN for remote sensing images object detection. *IEEE Access*, 2019, **7**: 43607–43621
- 42 Zhang Y K, You Y, Wang R, Liu F, Liu J. Nearshore vessel detection based on scene-mask R-CNN in remote sensing image. In: Proceedings of the 2018 International Conference on Network Infrastructure and Digital Content (IC-NIDC). Guiyang, China: IEEE, 2018. 76–80
- 43 Li Q P, Mou L C, Jiang K Y, Liu Q J, Wang Y H, Zhu X X. Hierarchical region based convolution neural network for multiscale object detection in remote sensing images. In: Proceedings of the 2018 IEEE International Geoscience and Remote Sensing Symposium. Valencia, Spain: IEEE, 2018. 4355–4358
- 44 Xie H N, Wang T, Qiao M N, Zhang M Y, Shan G C, Snoussi H. Robust object detection for tiny and dense targets in VHR aerial images. In: Proceedings of the 2017 Chinese Automation Congress (CAC). Ji'nan, China: IEEE, 2017. 6397–6401
- 45 Wu Z H, Gao Y M, Li L, Fan J L. Research on object detection technique in high resolution remote sensing images based on U-Net. In: Proceedings of the 2018 Chinese Control and Decision Conference (CCDC). Shenyang, China: IEEE, 2018. 2849–2853
- 46 Cao Y S, Niu X, Dou Y. Region-based convolutional neural networks for object detection in very high resolution remote sensing images. In: Proceedings of the 12th International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery (ICNC-FSKD). Changsha, China: IEEE, 2016. 548–554
- 47 Chen G W, Liu L, Hu W L, Pan Z X. Semi-supervised object detection in remote sensing images using generative adversarial networks. In: Proceedings of the 2018 IEEE International

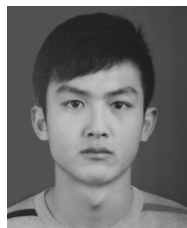
Geoscience and Remote Sensing Symposium. Valencia, Spain: IEEE, 2018. 2503–2506



刘小波 中国地质大学(武汉)自动化学院副教授。2008年获得中国地质大学计算机学院计算机科学硕士学位。2012年获得中国地质大学计算机学院地球科学信息工程博士学位。主要研究方向为机器学习, 演化计算, 高光谱遥感图像处理。本文通信作者。

E-mail: xbliu@cug.edu.cn

(LIU Xiao-Bo Associate professor at the School of Automation, China University of Geosciences, Wuhan. He received his master degree in computer sciences from the School of Computer Science, China University of Geosciences, Wuhan, in 2008. He received his Ph.D. degree in geosciences information engineering from the School of Computer Science, China University of Geosciences, Wuhan in 2012. His research interest covers machine learning, evolutionary computation, and hyperspectral remote sensing image processes. Corresponding author of this paper.)



刘鹏 中国地质大学(武汉)自动化学院硕士研究生。2017年获得湖北理工学院电气与电子信息工程学院学士学位。主要研究方向为遥感图像处理, 机器学习。

E-mail: pengliu0908@cug.edu.cn

(LIU Peng Master student at the School of Automation, China University of Geosciences, Wuhan. He received his bachelor degree from the School of Electrical and Electronics Engineering, Hubei Institute of Technology in 2017. His research interest covers remote sensing image processing and machine learning.)



蔡之华 中国地质大学(武汉)计算机学院教授。1986年获得武汉大学学士学位。1992年获得北京工业大学硕士学位。2003年获得中国地质大学(武汉)博士学位。主要研究方向为数据挖掘, 机器学习, 演化计算。

E-mail: zhcai@cug.edu.cn

(CAI Zhi-Hua Professor at the School of Computer Sciences, China University of Geosciences, Wuhan. He received his bachelor degree from Wuhan University in 1986. He received his master degree from Beijing University of Technology, in 1992. He received his Ph.D. degree from China University of Geosciences, Wuhan in 2003. His research interest covers data mining, machine learning, evolutionary computation and their applications.)



乔禹霖 中国地质大学(武汉)自动化学院硕士研究生. 2018 年获得青岛科技大学自动化学院学士学位. 主要研究方向为深度学习, 图像生成.

E-mail: yljiao@cug.edu.cn

(QIAO Yu-Lin Master student at the School of Automation, China

University of Geosciences, Wuhan. He received his bechelor degree from the School of Automation, Qingdao University of Science and Technology in 2018. His research interest covers deep learning and image generation.)



王凌 清华大学自动化系教授. 1995 年获得清华大学自动化系学士学位. 1999 年获得清华大学自动化系控制理论与控制工程专业博士学位. 主要研究方向为智能优化理论、方法与应用, 复杂生产过程建模、优化与调度.

E-mail: wangling@tsinghua.edu.cn

(WANG Ling Professor in the Department of Auto-

mation, Tsinghua University. He received his bechelor degree from the Department of Automation, Tsinghua University in 1995. He received his Ph.D. degree in control theory and control engineering from the Department of Automation, Tsinghua University in 1999. His research interest covers intelligent optimization theory, method and application, and complex production process modeling, optimization and scheduling.)



汪敏 中国地质大学(武汉)自动化学院硕士研究生. 2017 年获得青岛科技大学自动化学院学士学位. 主要研究方向为显著性检测, 语义分割, 人体解析.

E-mail: wm@cug.edu.cn

(WANG Min Master student at

the School of Automation, China University of Geosciences, Wuhan. She received her bechelor degree from the School of Automation, Qingdao University of Science and Technology in 2017. Her research interest covers saliency detection, semantic segmentation, and human parsing.)