基于深度学习的光学遥感图像目标检测研究进展

廖育荣[1]，王海宁[2]，林存宝[1]，李阳[2]，方宇强[1]，倪淑燕[1]
（1. 航天工程大学电子与光学工程系，北京 101416；2. 航天工程大学研究生院，北京 101416）

摘　要：目标检测是光学遥感图像解译中的核心问题，在情报侦察、目标监视、灾害救援等领域均具有广泛应用。首先，结合深度学习光学遥感图像目标检测算法研究进展，对基于候选区域和回归分析的两类遥感目标检测算法进行了综述。其次，针对旋转目标、小目标、多尺度、密集目标四类常见特定任务场景目标检测算法改进进行了总结。再次，结合常用遥感图像数据集对不同算法性能进行了对比分析。最后，对未来遥感图像目标检测研究值得关注的问题进行了展望，为后续相关研究提供思路。

关键词：光学遥感图像；目标检测；深度学习；卷积神经网络
中图分类号：TP391.41
文献标志码：A
DOI: 10.11959/j.issn.1000−436x.2022071

Research progress of deep learning-based object detection of optical remote sensing image

LIAO Yurong[1], WANG Haining[2], LIN Cunbao[1], LI Yang[2], FANG Yuqiang[1], NI Shuyan[1]
1. Department of Electronic and Optical Engineering, Space Engineering University, Beijing 101416, China
2. Department of Graduate Management, Space Engineering University, Beijing 101416, China

Abstract: Object detection is the core issue in the interpretation of optical remote sensing images, and it is widely used in fields such as intelligence reconnaissance, target monitoring, and disaster rescue. Firstly, combined with the research progress of deep learning optical remote sensing image object detection algorithms, the two types of algorithms based on candidate regions and regression analysis were reviewed. Secondly, the improvement of object detection algorithms for four types of common task-specific scenes were summarized, including rotating objects, small objects, multi-scales, and dense objects. Then, combined with commonly used remote sensing image data sets, the performance of different algorithms was compared and analyzed. Finally, the issues worthy of attention in remote sensing image object detection in the future were prospected, and ideas for follow-up related research were provided.

Keywords: optical remote sensing image, object detection, deep learning, convolutional neural network

0 引言

光学遥感图像目标检测的主要任务是在光学遥感图像中对感兴趣目标进行定位和分类[1]，在情报侦察、目标监视、灾害救援、工业应用及日常生活等领域都发挥着重要作用，同时也是后续目标跟踪、场景分类、图像理解等工作的基础。遥感图像目标检测关注的对象通常包括港口、桥梁、飞机、舰船、车辆等多类型目标，由于遥感目标自身存在任意方向旋转、小目标、多尺度、密集分布等情形，叠加不同成像条件及环境因素等对遥感图像的影响，导致针对多场景遥感图像目标的检测具有很大挑战。

传统的遥感图像目标检测算法一般可以分为

收稿日期：2021−12−23；**修回日期**：2022−03−23
通信作者：倪淑燕，daninini@163.com
基金项目：国家自然科学基金资助项目（No.61805283,No.61805284,No.61906213）
Foundation Item: The National Natural Science Foundation of China (No.61805283, No.61805284, No.61906213)

基于模板匹配的算法和基于人工特征建模的算法。具体地，基于模板匹配的算法计算输入图像中特定区域的特征向量与模板特征向量的匹配度，如结构相似性测量[2]与 Hausdorff 距离匹配[3]，根据匹配度大小来检测目标。基于人工特征建模的算法基于专家先验规则以选取候选区域，进而建立目标的特征表示，如视觉显著性检测[4]、尺度不变特征变换[5]、定向梯度直方图[6]和稀疏多模态学习[7]等，再通过分类器实现对目标的检测。传统遥感目标检测算法虽然理论完备，但其难以对多种复杂任务场景及目标进行充分特征表示，且滑动窗口式等操作限制了算法效率，导致其目标检测的精度和速度难以满足实际需求，针对不同类型的检测对象泛化能力仍需提高。

随着人工智能技术的发展以及数据样本不断扩充，研究人员开始致力于使用深度学习技术来解决遥感图像目标检测问题[8-9]，目前仍呈增加之势。当前基于深度学习的遥感图像目标检测算法可以分为两大类：基于候选区域[10-12]和基于回归分析[13-14]的算法，二者有时又被称为双阶段与单阶段算法。虽然基于深度学习的目标检测算法可解释性仍需进一步研究，但其在多个公开数据集上的目标检测效果已远超传统算法，特别是经过近 10 年的迅速发展已形成一批优秀成果。文献[15-17]针对遥感图像目标检测相关问题进行综述，重点围绕代表性方法、数据集、遥感图像所存在的特殊问题、特定目标类型等进行归纳总结。正是因为遥感图像智能目标检测算法的飞速发展，珍贵的遥感图像数据才得以被充分解译，这也促进了遥感技术服务于当今社会发展的潜力。因此，总结该领域中的深度学习目标检测算法研究进展具有重要的意义。

本文在充分参考相关研究及综述工作的基础上，首先，对遥感图像目标检测领域基于候选区域与回归分析的两大类算法进行了归纳综述；其次，针对最常见的遥感图像目标检测特定任务场景算法改进进行了总结；再次，介绍了遥感图像目标检测常用数据集，并对典型算法的测试结果进行了对比分析；最后，对未来遥感图像目标检测研究值得关注的问题进行了展望。

1　基于不同检测策略的遥感目标检测算法

1.1　基于候选区域的深度学习遥感目标检测算法

光学遥感图像在成像视角、空间分辨率、下视背景、目标属性等方面相对常规图像具有明显差异，因此对其目标检测需结合任务特点进行算法设计。基于候选区域的目标检测算法将问题分为 2 个阶段：第一阶段用于生成一系列可能包含目标的候选区域；第二阶段则对这些候选区域进行目标或背景分类及边界框回归。基于候选区域的光学遥感目标检测算法流程如图 1 所示，其中最具代表性的算法为区域卷积神经网络（R-CNN, region convolutional neural network）及其改进系列算法。

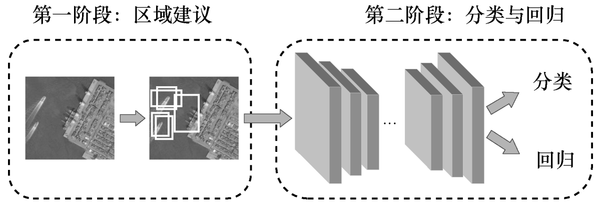


图 1　基于候选区域的光学遥感目标检测算法流程

文献[18]于 2014 年提出 R-CNN 算法，其首先通过选择性搜索算法生成 2 000 个候选区域，然后对每个候选区域进行尺度归一化，并送入 CNN 进行特征提取，最后基于所提取特征进行支持向量机分类和边界框回归。2015 年，文献[19]提出 Fast R-CNN 算法，使用感兴趣区域池化（ROI Pooling, region of interest pooling）层进行区域特征选择。2017 年，Ren 等[20]提出了 Faster R-CNN 算法，将候选区域生成融入整体网络中，实现了端到端的深度学习目标检测框架，极大地提高了算法检测速度和精度。基于 Faster R-CNN 的优异性能，研究人员将其广泛拓展到遥感图像检测领域，以期实现对遥感目标的有效检测。现有工作主要聚焦于提高目标特征表示的有效性、感兴趣区域定位的准确性，以及预测位置回归的准确性这 3 个方面。

1.1.1　增强目标特征表示的改进研究

特征表示是将目标信息映射到特征向量的过程，其直接影响后续分类的精度。Faster R-CNN 算法的浅层网络中虽含有较多目标特征信息，但语义信息不够丰富，容易造成目标定位与分类误差。其深层网络中语义信息相对丰富，但计算量较大，难以满足遥感图像目标检测速度要求。为了提高 Faster R-CNN 算法对遥感图像目标的特征表示能力，通常采用特征融合策略进行特征增强[21-25]，同时，也可以通过改进卷积层结构或处理策略以增强目标特征表示，如使用层次选择滤波层[26]、非局部特征增强卷积层[27]等。此外，Cheng 等[28]在浅层卷积特征提取时使用旋转不变正则化器和 Fisher 判别

正则化器，消除了目标特征旋转与类间相似性对目标特征提取的影响。Han 等[29]通过将生成候选区域与检测网络这 2 个阶段的特征进行共享集成，增强了检测网络的目标特征表示能力。

1.1.2 优化感兴趣区域生成与处理的改进研究

候选区域生成网络（RPN, region proposal network）与 RoI Pooling 层的引入是 Faster R-CNN 相比原有 R-CNN 系列算法的重大改进，对提高目标检测性能至关重要。针对遥感目标多方向、多尺度、多成像条件等特点，研究人员主要聚焦从候选 RPN 与 ROI Pooling 层 2 个方面进行改进，以提高算法对遥感目标的检测效果。一方面，RPN 用于为 ROI 提供有效的候选框集合，其准确性将直接影响后续目标定位与分类精度，因此对 RPN 的改进是优化 ROI 最直接有效的方式[30-33]。另一方面，ROI Pooling 层在 RPN 的基础上筛选训练样本，并且由不同大小的候选区域生成相同尺度的特征图，使遥感图像的 ROI 更好地保留或融合空间位置信息[34-40]。

1.1.3 提高目标定位精度的改进研究

遥感图像中普遍存在目标方向各异、密集排列的情况，导致准确确定目标位置具有较大难度，并且目标与目标之间可能出现区域重叠，干扰预测位置回归的准确性。一方面，交并比（IoU, intersection over union）值作为衡量锚点框定位精度的优化目标，其无法区分 2 个重叠对象在不同方向上有相同交叉的情况。因此对 IoU 的改进重点在于增强其自适应能力，如文献[38]设计的自适应 IoU 计算，文献[41]提出的旋转框 IoU 计算，文献[42]提出的联合交集 IoU 计算等方法，有效缓解了因阈值设置带来的检测结果敏感性问题。另一方面，由于目标定位精度依赖于非极大值抑制（NMS, non-maximum suppression）算法筛除冗余预选框，因而针对传统 NMS 的改进也是提高遥感目标定位精度的一个重要途径[43-44]。基于候选区域的遥感目标检测算法相关信息如表 1 所示。

1.2 基于回归分析的深度学习遥感目标检测算法

推扫型和凝视型成像遥感卫星生成的可见光遥感图像通常具有较高的像元数量，而对于遥感卫星典型应用，通常具有较高的目标检测时效要求。基于回归分析的目标检测算法不需要单独候选区域生成，而是从输入图像的多个位置直接回归分析出目标的边界框与类别，具有比基于候选区域算法更快的处理速度，因此在遥感图像目标检测领域获得了广泛关注。基于回归分析的光学遥感图像目标

表 1 基于候选区域的遥感目标检测算法相关信息

文献	算法亮点	适用范围
文献[21-22]	引入特征融合模块，在准确提取深层网络特征的同时简化计算空间复杂度	针对一般遥感目标特征表示阶段改进
文献[23-25]	采用区域上下文信息增强目标特征与对应场景之间的联系	
文献[26]	调整特征图尺度变换的方式，增强遥感目标检测的特征表征能力	
文献[27]	固定角度旋转进行数据增强，调整特征图尺度变换的方式	
文献[28]	浅层卷积特征提取时使用正则化器	
文献[29]	引入特征共享集成机制	
文献[30]	提出位置敏感平衡策略	针对一般遥感目标感兴趣区域生成与处理阶段改进
文献[31]	提出选择性搜索匹配策略	
文献[32]	提出旋转不敏感 RPN，有效提高了对遥感图像中多方向目标的检测效果	
文献[33]	提出一种改进的 NMS 模块，以更好地进行目标区域筛选	
文献[34]	提出一种兴趣区域特征金字塔池化模块	
文献[35]	提出一种 ROI 形变器	
文献[36]	提出一种 ROI 排序方法	
文献[37]	提出一种 ROI 旋转策略	
文献[38]	将 ROI Pooling 层与不同交并比加权损失值相结合以提高目标区域定位精度	
文献[40]	提出了一种结合掩模预测及语义分割算法，使 ROI 中保留更多目标空间特征向量	
文献[41]	计算旋转 IoU 损失	针对一般遥感目标提升定位精度改进
文献[42]	计算联合交集 IoU 损失	
文献[43]	设计了比例约束 NMS 算法，能够筛除遥感目标预选框中的错误结果	
文献[44]	提出了衰减 NMS 算法，能够解决多种遥感目标不同尺度重叠的问题	

检测主要包含基于 YOLO（you only look once）[45]和 SSD（single shot multibox detector）[46]框架的两类算法，其基本流程如图 2 所示。

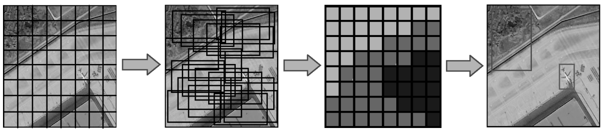


图 2 基于回归分析的光学遥感目标检测基本流程

1.2.1 YOLO 系列遥感目标检测算法

遥感目标检测的实时性是性能评价的关键因素，因此更快速的 YOLO 系列目标检测算法得到了广泛关注。Redmon 等[45]提出了 YOLOv1 算法，通过端到端的单阶段网络同时实现对目标的分类与定位，在兼顾检测精度的同时大幅提高了检测速度。YOLO 算法在发展过程中不断被改进[47-49]，逐步向高效率检测方向发展。Jocher 等[50]针对实时目标检测提出了 YOLOv5 算法，通过采用模型剪裁等技术进一步压缩了模型大小，提高了检测速度。Ge 等[51]提出了 YOLOX 算法，通过数据增强、无锚点检测以及标签分类等巧妙的集成方案，实现了算法精度与速度的进一步提升。

YOLO 系列目标检测算法因其便于轻量化部署的优势，在高像素、高实时，特别是资源受限的在轨处理等遥感领域具有较大潜力。研究人员基于改进的 YOLO 系列算法，已成功探索其在地震房屋受损检测[52]、深海石油探测[53]、特殊农作物种植检测[54]等遥感领域应用。当前提升 YOLO 算法对遥感目标的检测精度一方面可以从特征提取角度展开，能够更加有效地使网络进行特征迭代[55-58]；另一方面，使用注意力机制可以突出遥感目标的资源分配权重[59-60]，或采用密集连接网络以增强层级之间的信息传输[61]，从而达到基于 YOLO 网络进行遥感图像目标检测任务更深入、更准确的训练。此外，由于不同的遥感场景中目标的形态特点各式各样、方向排序千差万别，因此 YOLO 检测算法衍生出多种专项组合网络，分别用于小目标[62-63]、多尺度[64-65]、多模态变化[66]等具有挑战性的遥感检测场景，并取得了比原网络更优的检测效果。

1.2.2 SSD 框架在遥感检测中的应用

Faster R-CNN 的锚点框机制和 YOLO 算法的回归思想在遥感目标检测中有各自的优势。基于两者的优势，2016 年 Liu 等[46]提出了 SSD 算法，将 Faster R-CNN 中锚点框改为先验框，并且直接使用回归分析过程，有效提升了遥感目标检测的速度。此外，SSD 算法从多尺度特征图中生成预测，克服了光学遥感图像常出现的目标尺寸不同的问题。

基于 SSD 算法在目标检测速度和精度上的均衡优势，研究人员分别从网络结构[67-69]、回归策略[70]，以及特定场景适应性[71-75]等角度出发，进行算法改进研究，以此抑制光学遥感图像中噪声对目标检测的影响，提高检测精度。表 2 总结了基于回归分析的遥感目标检测算法相关信息。尽管当前针对上述问题设计出多种改进方案，但遥感目标检测仍是一项开放的、具有挑战性的创新型工作。

2 面向特定任务场景的遥感目标检测算法

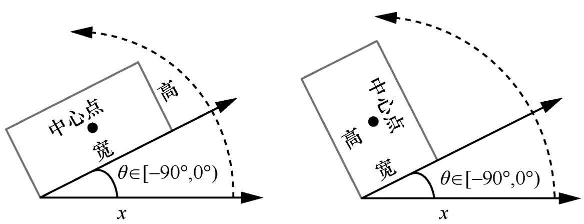
2.1 针对遥感旋转目标的改进研究

遥感图像中非规则排列的目标因呈现方向多样性特点，被称作旋转目标。常规深度学习目标检测网络由于并未针对旋转目标进行针对性设计，因此通常对该类目标特征提取能力相对较弱，进而影响对该类目标的检测效果。为了解决这一问题，一种解决思路是通过旋转变换来扩充训练样本[27,76]，该方法依旧使用水平框标记检测到的目标，虽然能够在一定程度上增强模型对旋转目标的稳健性，但也容易造成水平框与旋转目标实际尺寸不符的结果。另一种解决思路是采用回归旋转目标的角度信息，在目标周围生成合适的倾斜框，从而提高对旋转目标的特征表示能力。目前常用旋转框参数表示方法包括以 90° 或 180° 为周期的五参数表示法和有序四边形表示法[77-78]，如图 3 所示。其中，五参数表示法使用目标中心点坐标、框宽高、x 轴与框边夹角来表示旋转框的位置；有序四边形表示法采用 4 个顶点坐标对水平框与衍生旋转框做逆时针排列。

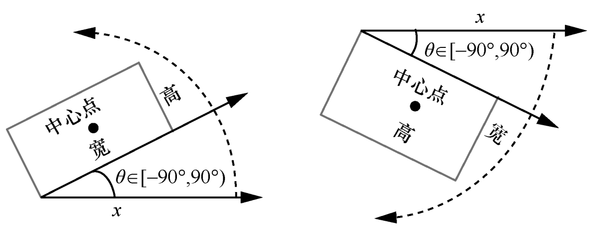
现阶段大多数研究工作使用上述表示法进行旋转角度回归，生成准确的预测边界框。文献[79-80]将旋转锚定采样与多层特征相结合，构造了针对旋转目标的特征提取与融合结构，然而在角度预测过程中会产生较大的损失值，形成难以消除的边界误差。文献[81-83]较好地平衡了旋转角度分类所带来的理论误差问题。除了角度回归方法以外，研究人员还提出了自适应感受野[84]、中线预测[85-86]、掩模分支预测[87]等技术，文献[88]则提出在损失函数中添加正则化项来约束旋转前后的特征映射关系，从而实现对旋转目标的有效检测。

表 2 基于回归分析的遥感目标检测算法相关信息

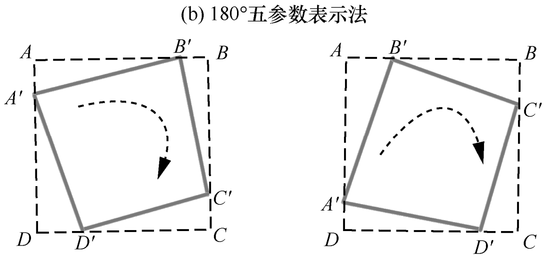
文献	算法亮点	适用范围
文献[52-54]	在特殊应用背景下的成功探索	对检测速度有更高要求的快速遥感目标检测场景
文献[55]	设计特征增强模块，能够提升目标的特征表达能力	
文献[56]	设计特征精细化模块，能够实现目标特征的细致区分	
文献[57]	结合空洞卷积，有效提升 YOLO 算法针对大目标的精度提升	
文献[58]	引入多感受域算法，提高了目标特征提取能力	
文献[59-60]	引入注意力机制实现精度与速度间的最佳权衡	
文献[61]	采取密集连接网络以增强目标特征提取能力	
文献[62-66]	分别针对小目标、多尺度、多模态等专项遥感检测场景定制化改进	
文献[67]	引入空间和通道注意力模块以筛选融合特征信息	对检测速度与精度间有均衡要求的一般遥感目标检测场景
文献[68]	结合空洞卷积与特征融合策略，能够增强浅层特征的语义信息	
文献[69]	设计新的编码–解码模块，结合空间和通道注意力级联结构	
文献[70]	采取连续两次回归的策略，使检测结果更加精确	
文献[71]	通过定向边界框生成角度修正参数	
文献[72]	通过旋转边界框学习遥感目标的正确角度方向	
文献[73]	利用密集连接网络提取特征	
文献[74]	结合激活函数 CReLU 和特征金字塔网络，能够提升网络的浅层特征的利用率	
文献[75]	利用浅层信息并增大小目标训练权重以提升特征传递效率	



(a) 90°五参数表示法



(b) 180°五参数表示法



(c) 有序四边形表示法

图 3 常用旋转框参数表示方法

2.2 针对遥感小目标检测的改进研究

遥感图像中通常会包含一些仅有几十甚至十几个像素的小目标。小目标给遥感目标检测带来了一定挑战：一方面小目标特征相对较少，传统 CNN 的降采样处理会减小特征图的尺度，因此小目标容易在降采样的特征图中消失，导致算法难以有效进行特征提取；另一方面当小目标与背景颜色接近时，易出现目标与背景难以区分，导致对小目标定位不准的问题。当前改进算法主要通过增大小目标特征图尺度和设计有效的特征融合提取模块的方式来提高检测能力。

关于增大特征图尺度，常用的方法是增强浅层网络的特征信息，或者在深层特征中进行上采样操作[35,64,75,89-90]。文献[44]提出了一种小目标检测增强架构，通过上采样操作增大深层特征图尺度。值得注意的是，在保证深层特征中始终存在小目标时上采样才有意义，否则需要重新优化浅层网络。关于设计小目标特征融合提取模块，其目的是丰富特征图中的小目标信息，并在后续传递过程中减少信息损失。文献[80]提出了级联特征的融合策略，通过整合浅层位置信息和深层语义信息，达到融合各层特征并增强级联效果的目的。文献[68]利用扩展卷积设计了一种高效特征融合网络，增强了深层特征的有效感受野。由于真实遥感图像中的小目标尺度可能不同，特征融合模块难以在深层特征中拼接，

因此仅使用特征融合模块的方式可能会成为检测的瓶颈，而现有工作已验证注意力机制与特征融合模块相结合在小目标检测的有效性[67]。此外，特征嵌入[26]、特征迁移[33]等技术也被应用于遥感小目标检测特征融合方法中。在实际检测中，根据任务需要将特征图尺度调节与特征融合策略相结合，使 2 种思路实现优势互补，能够缓解目标尺度过小对检测带来的影响。

2.3 针对遥感多尺度目标检测的改进研究

在卷积神经网络中，目标的语义结构信息可能出现在任何卷积层，各层特征的分布与目标的尺度大小有关。在遥感场景中，经常会出现不同类别尺度的目标或者同一类别不同大小的目标在同一视野的情况。这些目标的尺度差距较大，使不同目标的特征难以同步传递至深层网络，导致以特征图表征作为预测基础的深度卷积网络不能有效获取多尺度目标的特征，进而影响多尺度目标的检测精度。因此尺度变化范围大的遥感图像目标检测仍是一项具有挑战性的问题。

当前遥感多尺度目标检测主要围绕多尺度信息融合模块进行改进[22,91-93]。一种较具代表性的算法是利用特征金字塔结构来充分获取多尺度目标的特征信息，使不同尺度的特征得以保留[94-96]。多尺度目标中通常同时包含小目标，利用金字塔结构等特征提取融合模块能够有效兼顾浅层与深层特征，在对多尺度目标进行有效检测的同时也兼顾了对小目标的高精度检测[97]。此外，Transformer 模型[98]所引入的注意力机制被证明在多尺度目标检测中具有良好效果，文献[99]利用 Transformer 模型改进 YOLOv5 的预测网络，结合自注意力机制实现了对多尺度目标的有效检测。

2.4 针对遥感密集目标检测的改进研究

光学遥感图像密集目标检测的难点在于各个目标位置相距较近导致易出现位置预测不准确和边界框重叠的情况。同时，当采用水平边界框进行目标位置预测时，由于水平边界框包围面积一般大于目标实际面积，使在对密集目标检测时会加剧边界框重叠问题，特别是当目标密集斜向排列时。为提高对遥感密集目标检测效果，当前主要从特征增强和精确定位 2 个方向对算法进行改进。

在特征增强方面，文献[79]通过采用监督式空间注意力和通道注意力组合机制，在增强密集目标特征的同时削弱了背景特征，使密集目标之间的界限更清晰。文献[56]设计了特征精细化模块来进行特征重建，使特征与目标对齐，同时采用了微调旋转锚点框以适应密集目标检测任务场景。文献[100]通过采用多尺度检测算法来增强对密集目标的特征提取能力，结合圆形平滑标签（CSL, circular smooth label）技术来改进损失函数，以此降低对密集目标重复检测的概率。在精确定位方面，文献[31]提出了基于无监督的边界框回归算法，并利用非极大值抑制方式优化密集目标的边界框，从而实现对密集目标的准确定位。此外，密集标签编码[82]、特征选择与动态优化[84]等方式从特征匹配角度出发，也为密集遥感目标检测提供了新的研究思路。以上算法的相关信息如表 3 所示。

综上，遥感图像目标检测中的复杂任务场景激发了多类算法改进。需要注意的是，虽然上述总结分析是对四类特定任务场景进行阐述，但实际中一幅遥感图像可能同时包含其中的多类情形，针对某一任务场景的改进有时也会有利于其他任务场景。因此，在光学遥感图像目标检测领域，面向各类问题的多任务优化目标可以协同解决。

3 遥感图像数据集及算法性能评价

3.1 遥感图像目标检测公开数据集

光学遥感图像数据集在遥感目标检测任务中发挥着至关重要的作用，它能够为模型训练提供珍贵标准的遥感数据，同时也为不同网络及算法对比提供客观统一的基准。近年来随着卫星遥感技术的发展，一些高质量的光学遥感图像目标检测数据集流行起来，本文选取具有代表性的 15 个数据集进行介绍，各数据集样例的统计信息如表 4 所示，包括发布者及内容描述、数据集所含目标类别数与图像数。正是这些公开的光学遥感图像数据集，使基于深度学习的遥感目标检测技术快速发展。

3.2 算法性能评价

当前评价光学遥感图像目标检测算法常用的性能指标是精确度（Precision）、召回率（Recall）、平均精度均值（mAP, mean average precision）以及帧率（FPS, frame per second）。精确度反映了检测结果中的真实正样本占比。召回率反映了在所有待检测正样本中被正确检测到的正样本占比，精确度和召回率存在权衡组合的关系。其中，将精确度作为纵坐标，召回率作为横坐标作图，可得到精度–

表 3 面向特定任务场景的遥感目标检测算法相关信息

文献	算法亮点	适用范围
文献[76]	随机角度旋转以进行数据增强	针对遥感旋转目标的改进
文献[77-78]	有序四边形表示法学习矩形各边的偏移	
文献[79]	结合多层特征与锚定采样的特征融合网络	
文献[80]	监督空间注意力机制与多步检测子网络级联相结合	
文献[81-82]	将旋转角度回归转化为分类问题并限制分类结果的范围	
文献[83]	2 个任意方向高斯分布的 Wasserstein 距离作为旋转 IoU 损失	
文献[84]	提出自适应调整的神经元感受野模块	
文献[85]	提出目标中线以进行旋转角度回归	
文献[86]	提出学生 T-分布将分类任务与旋转角度回归相关联	
文献[87]	提出掩模分支预测下获取目标旋转边界框	
文献[88]	添加正则化项约束训练样本在前后的特征表示	
文献[89]	减小网络降采样的倍数，并将浅层与深层特征融合拼接	针对遥感小目标的改进
文献[90]	在浅层特征图中聚合上下文信息，并增加小目标训练权重	
文献[91]	构建包含不同尺度信息的特征融合架构	针对遥感多尺度目标的改进
文献[92]	寻找合适的 ROI 尺寸进行多尺度特征映射与融合	
文献[93]	将视觉注意力网络引入多尺度信息融合模块	
文献[94]	结合全局注意力模块将多尺度特征图输入显著性金字塔	
文献[95]	提出定向引导的锚点框生成网络以生成合适的锚点框	
文献[96]	利用特征金字塔进行不同尺度目标信息的提取	
文献[97]	使用超尺度块作为核心结构构建超尺度目标检测框架	
文献[99]	利用 Transformer 模型的预测主体提升多尺度预测能力	
文献[100]	基于圆形平滑标签设计损失函数	针对遥感密集目标的改进

表 4 常用的光学遥感图像目标检测数据集概述

数据集	发布者及内容描述	目标类别数/个	图像数/幅
TAS [101]	由 Stanford 大学发布的车辆目标数据集	1	30
OIRDS [102]	由 Raytheon 公司发布的车辆目标数据集	5	900
SZTAKI [103]	由 Mta Sztaki 发布的旋转建筑物目标数据集	1	9
UCAS-AOD [104]	由中国科学院发布的车辆和飞机目标数据集，以及背景负样本	2	976
NWPU VHR-10 [105]	由西北工业大学发布的飞机、舰船、油罐、棒球场、网球场、篮球场等目标数据集	10	1 510
VEDAI [106]	由 Caen 大学发布的车辆目标数据集	9	1 210
HRSC2016 [107]	由西北工业大学发布的舰船目标数据集	1	1 061
DLR3k [108]	由 German Aerospace Center 发布的车辆目标数据集	7	20
RSOD [31]	由武汉大学发布的飞机、油箱、运动场和立交桥目标数据集	4	976
TGRS-HRRSD [109]	由中国科学院发布的舰船、桥梁、田径场、油罐、篮球场、网球场等目标数据集	13	21 761
LEVIR [110]	由北京航空航天大学发布的飞机、舰船、油罐目标数据集	3	22 000
ITCVD [111]	由 Twente 大学发布的车辆目标数据集	1	135
DIOR [15]	由西北工业大学发布的飞机、机场、篮球场、桥梁、烟囱、水坝等目标数据集	20	23 463
DOTA [112]	由武汉大学发布的舰船、游泳池、田径场、港口、直升机、足球场等目标数据集	16	2 806
FAIR1M [113]	由中国科学院等发布的飞机、舰船、车辆、球场、道路等 5 种大类别、37 个细粒度类别的数据集，是目前全球规模最大的光学遥感图像细粒度目标检测识别数据集	37	15 000

召回率曲线（PR, precision-recall curve）；曲线下的面积表示某类别目标的平均精确度（AP, average precision）；多个类别的 AP 均值即平均精度均值 mAP，表示算法在数据集上的整体性能。帧率是每秒可以处理的图片数量，用于评估算法的检测速度。

各算法性能评价大多采用 NWPU VHR-10 与 DOTA 数据集，典型光学遥感图像目标检测算法性能对比如表 5 所示。结合前述对各类型算法原理特点的阐述分析，以及表 5 中典型算法在相同数据集上的性能对比，可以得出以下结论。

表 5　典型光学遥感图像目标检测算法性能对比

数据集	算法	主干网络	改进功能				mAP
			目标旋转	微小目标	多尺度	目标密集	
DOTA 数据集	GWD[83]	ResNet152	√			√	80.23
	HSP[21]	ResNet101	√		√		78.01
	NLFE[27]	ResNet50	√				77.46
	DCL[82]	ResNet152	√	√			77.37
	R^{3}Det[56]	ResNet152	√			√	76.47
	CLS[81]	ResNet152	√			√	76.24
	APERO[42]	ResNet101	√	√			75.75
	FFA[22]	ResNet101	√		√		75.70
	SCFPN[23]	ResNet101			√		75.22
	Gliding vertex[78]	ResNet101	√				75.02
	RSDet[77]	ResNet152	√			√	74.10
	DR-Net[84]	CenterNet	√				73.23
	GLS-Net[94]	ResNet101	√		√		72.96
	IoU-Adaptive[38]	ResNet101		√	√		72.72
	FMSSD[75]	VGG16		√	√		72.43
	CAD-Net[25]	ResNet101	√				69.60
	RoITransformer[41]	ResNet101	√		√		69.56
	TEANS[44]	ResNet101		√			66.01
	AFF-SSD[67]	ResNet50		√			52.60
NWPU VHR-10 数据集	HyNet[97]	ResNet50		√	√		98.52
	SCRDet[79]	ResNet101	√	√		√	91.75
	MS-OPN AODN[38]	GoogleNet		√	√	√	91.52
	CAD-Net[25]	ResNet101	√				91.50
	FMSSD[75]	VGG16		√	√		90.40
	AF-SSD[69]	ShuffleNetV2		√			88.70
	RICA[32]	ZFNet	√				87.12
	Deformable Faster RCNN[35]	ResNet50					84.40
	Sig-NMS[33]	VGG16		√			82.90
	PSB[30]	ResNet101					81.20
	HSF-Net[26]	VGG16		√	√		81.15
	Deformable R-FCN[43]	ResNet101			√		79.10
	R-P-Faster R-CNN[29]	VGG16			√		76.50
	DFSSD[68]	VGG16		√			65.35

1) 由于不同遥感图像数据集在空间分辨率、背景复杂性、目标多样性、成像质量上存在差异，同一算法在场景复杂与简单的数据集的 mAP 值差异明显，如 CAD-Net[25]算法。这表明复杂任务场景下的目标检测依然是当前一个具有挑战性的问题。

2) 由于算法结构本身的约束使选取不同的主干网络会得出不同的结果，因此要在实验中匹配合适的主干网络层数，使模型的性能整体上得到提升。从表 5 中可以看出，高性能网络模型逐渐成为主干网络的首选。此外，从不同层数 ResNet 主干网络对应算法性能来看，主干网络层数也并非越深越好，需要与算法框架进行匹配设计。主干网络层的非线性拟合能力与算法的匹配结合问题，也是当前进行深度学习解释性研究工作的难点。

3) 算法针对确切场景问题从而做出的改进是有效的。例如 GWD[83]专门用以解决目标方向旋转的问题，HyNet[97]专门用以解决目标多尺度变化的问题。此外，由 FMSSD[75]和 HSF-Net[26]等的性能对比可以看到，针对小目标检测所改进的增强网络特征提取能力，在一定程度上也会提升对密集目标的检测效果。因此，针对某一任务场景的改进有时也会有利于其他任务场景，面向各类问题的多任务优化目标可以协同解决。

4 深度学习遥感目标检测研究展望

近年来，基于深度学习的遥感图像目标检测研究迅速发展。从遥感图像目标检测技术未来能够广泛应用于民用领域来看，提高各类复杂任务场景的检测精度和速度，进而提高算法整体性能依然是遥感图像目标检测的主要发展方向。与此同时，伴随着航天、载荷等技术的发展，遥感图像目标检测未来在以下 4 个方面值得进一步研究。

1) 遥感小样本无监督学习目标检测。尽管近年来搭载有光学载荷的成像卫星频繁发射，公开遥感图像数据集也接连发布，但受限于卫星总体数量、特定遥感数据的敏感性，以及对目标类别型号检测的更高要求，依然缺乏稀有目标或偶发事件的足够训练数据，通常需要在小样本或仿真数据下进行模型训练和验证，文献[114-115]已开展相关研究。由于部分遥感数据存在无标注或标注不明确情况，因此还需要采用无监督学习来解决问题。

2) 遥感视频影像动态目标检测。高分辨率视频影像在提供更加丰富信息的同时，也对遥感影像目标检测及后续处理提出了新的需求。遥感视频影像具有目标微小、干扰复杂的特点，因此基于深度学习的遥感视频目标检测仍面临巨大的挑战[116]。当前遥感目标检测主要针对单帧影像，探索基于视频影像的高效目标检测策略和算法，包括构建遥感视频影像目标检测数据集及评价准则，将会是未来计算机与遥感通信交叉领域的一项重要研究内容。

3) 多源数据融合目标检测。当前遥感目标检测主要基于可见光图像，限制了其在情报侦察、灾害救援等涉及夜间、微光、云雾、伪装等条件下的应用。伴随着多光谱、红外、合成孔径雷达（SAR, synthetic aperture radar）载荷先后发射升空，未来天基遥感数据将包含多源异构数据。遥感图像目标检测算法也将在主要处理可见光图像的基础上，进一步拓展到对多光谱、红外、SAR 图像的处理，以及对多源异构遥感数据的融合目标检测[117]。

4) 遥感在轨实时处理检测。在航天情报侦察与目标跟踪等应用领域，往往对情报信息的准确性和时效性有着严格要求。当前，一方面受限于算法性能，另一方面受限于卫星硬件算力及资源消耗，遥感目标检测算法主要针对卫星下传数据进行离线处理和人机结合应用。从航天技术未来发展来看，在轨实时目标检测将成为遥感领域的必然发展方向之一，也必将促进适用于卫星在轨处理的目标检测硬件加速技术的发展[118]。

5 结束语

随着卫星技术的迅速发展与遥感数据集规模的日益扩增，基于深度学习的光学遥感图像目标检测技术已经证明其在特征表征、训练表达等方面的优越性。本文首先对当前基于候选区域与回归分析的两大类光学遥感图像目标检测算法进行了归纳，在此基础上针对四类常见特定任务场景目标检测改进算法进行了综述；然后结合常用遥感目标数据集对不同算法的性能进行了对比分析；最后对未来遥感图像目标检测值得关注的问题进行了展望，进一步说明了该领域面临的挑战与巨大应用前景。

参考文献：

[1] CHENG G, HAN J W. A survey on object detection in optical remote sensing images[J]. ISPRS Journal of Photogrammetry and Remote

Sensing, 2016, 117: 11-28.

[2] WANG Z, BOVIK A C, SHEIKH H R, et al. Image quality assessment: from error visibility to structural similarity[J]. IEEE Transactions on Image Processing, 2004, 13(4): 1-14.

[3] WU W. Quantized Gromov-Hausdorff distance[J]. Journal of Functional Analysis, 2006, 238: 58-98.

[4] HAREL J, KOCH C, PERONA P. Graph-based visual saliency[J]. Advances in Neural Information Processing Systems, 2006, 19: 545-552.

[5] SEDAGHAT A, EBADI H. Remote sensing image matching based on adaptive binning SIFT descriptor[J]. IEEE Transactions on Geoscience and Remote Sensing, 2015, 53(10): 5283-5293.

[6] DALAL N, TRIGGS B. Histograms of oriented gradients for human detection[C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2005: 886-893.

[7] GÓMEZ C L, TUIA D, MOSER G, et al. Multimodal classification of remote sensing images: a review and future directions[J]. Proceedings of the IEEE, 2015, 103(9): 1560-1584.

[8] ZHU X X, TUIA D, MOU L C, et al. Deep learning in remote sensing: a comprehensive review and list of resources[J]. IEEE geoscience and remote sensing magazine, 2017, 5(4): 8-36.

[9] 罗会兰, 陈鸿坤. 基于深度学习的目标检测研究综述[J]. 电子学报, 2020, 48(6): 1230-1239.
LUO H L, CHEN H K. Survey of object detection based on deep learning[J]. ACTA ELECTRONICA SINICA, 2021, 48(6): 1230-1239.

[10] LI Y S, ZHANG Y J, HUANG X, et al. Deep networks under scene-level supervision for multi-class geospatial object detection from remote sensing images[J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2018, 146: 182-196.

[11] XUE B, TONG N N. DIOD: fast and efficient weakly semi-supervised deep complex ISAR object detection[J]. IEEE Transactions on Cybernetics, 2019, 49(11): 3991-4003.

[12] YU D W, JI S P. A new spatial-oriented object detection framework for remote sensing images[J]. IEEE Transactions on Geoscience and Remote Sensing, 2022, 60: 1-16.

[13] 周勇, 陈思霖, 赵佳, 等. 基于弱语义注意力的遥感图像可解释目标检测[J]. 电子学报, 2021, 49(4): 679-689.
ZHOU Y, CHEN S L, ZHAO J, et al. Weakly semantic based attention network for interpretable object detection in remote sensing imagery[J]. ACTA ELECTRONICA SINICA, 2021, 49(4): 679-689.

[14] ZHANG Y J, SHENG W G, JIA J F, et al. Priority branches for ship detection in optical remote sensing[J]. Remote Sensing, 2020, 12(7): 1196.

[15] LI K, WAN G, CHENG G, et al. Object detection in optical remote sensing images: a survey and a new benchmark[J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2020, 159: 296-307.

[16] 聂光涛, 黄华. 光学遥感图像目标检测算法综述[J]. 自动化学报, 2021, 47(8): 1749-1768.
NIE G T, HUANG H. A survey of object detection in optical remote sensing images[J]. ACTA AUTOMATICA SINICA, 2021, 47(8): 1749-1768.

[17] 宋志娜, 眭海刚, 李永成. 高分辨率可见光遥感图像舰船目标检测综述[J]. 武汉大学学报(信息科学版), 2021: doi.org/10.13203/j.whugis20200481.
SONG Z N, SUI H G, LI Y C. A survey on ship detection technology in high-resolution optical remote sensing images[J]. Geomatics and Information Science of Wuhan University, 2021: doi. org/10.13203/j.whugis20200481.

[18] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2014: 580-587.

[19] GIRSHICK R. Fast R-CNN[C]//Proceedings of IEEE Conference on Computer Vision. Piscataway: IEEE Press, 2015: 1440-1448.

[20] REN S Q, 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, 2017, 39(6): 1137-1149.

[21] XU C Y, LI C Z, CUI Z, et al. Hierarchical semantic propagation for object detection in remote sensing imagery[J]. IEEE Transactions on Geoscience and Remote Sensing, 2020, 58(6): 4353-4364.

[22] FU K, CHANG Z H, ZHANG Y, et al. Rotation-aware and multi-scale convolutional neural network for object detection in remote sensing images[J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2020, 161: 294-308.

[23] CHEN C Y, GONG W G, CHEN Y L, et al. Object detection in remote sensing images based on a scene-contextual feature pyramid network[J]. Remote Sensing, 2019, 11(3): 339.

[24] YAN D C, LI G Q, LI X Q, et al. An improved faster R-CNN method to detect tailings ponds from high-resolution remote sensing images[J]. Remote Sensing, 2021, 13(11): 2052.

[25] ZHANG G J, LU S J, ZHANG W. CAD-Net: a context-aware detection network for objects in remote sensing imagery[J]. IEEE Transactions on Geoscience and Remote Sensing, 2019, 57(12): 10015-10024.

[26] LI Q P, MOU L C, LIU Q J, et al. HSF-Net: multiscale deep feature embedding for ship detection in optical remote sensing imagery[J]. IEEE Transactions on Geoscience and Remote Sensing, 2018, 56(12): 7147-7161.

[27] 赵彤洲, 杨成万, 刘威. 非局部特征增强的遥感图像目标检测方法[J]. 华中科技大学学报(自然科学版), 2021: doi.org/10.13245/j.hust.210909.
ZHAO T Z, YANG C W, LIU W. Remote sensing image object detection method based on non-local feature enhancement[J]. Journal of Huazhong University of Science & Technology. (Natural Science Edition), 2021: doi.org/10.13245/j.hust.210909.

[28] CHENG G, HAN J W, ZHOU P C, et al. Learning rotation-invariant and fisher discriminative convolutional neural networks for object detection[J]. IEEE Transactions on Image Processing, 2019, 28(1):

265-278.

[29] HAN X B, ZHONG Y F, ZHANG L P. An efficient and robust integrated geospatial object detection framework for high spatial resolution remote sensing imagery[J]. Remote Sensing, 2017, 9(7): 666.

[30] ZHONG Y F, HAN X B, ZHANG L P. Multi-class geospatial object detection based on a position-sensitive balancing framework for high spatial resolution remote sensing imagery[J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2018, 138: 281-294.

[31] LONG Y, GONG Y P, XIAO Z F, et al. Accurate object localization in remote sensing images based on convolutional neural networks[J]. IEEE Transactions on Geoscience and Remote Sensing, 2017, 55(5): 2486-2498.

[32] LI K, CHENG G, BU S H, et al. Rotation-insensitive and context-augmented object detection in remote sensing images[J]. IEEE Transactions on Geoscience and Remote Sensing, 2018, 56(4): 2337-2348.

[33] DONG R C, XU D Z, ZHAO J, et al. Sig-NMS-based faster R-CNN combining transfer learning for small target detection in VHR optical remote sensing imagery[J]. IEEE Transactions on Geoscience and Remote Sensing, 2019, 57(11): 8534-8545.

[34] 仲伟峰, 郭峰, 向世明, 等. 旋转矩形区域的遥感图像舰船目标检测模型[J]. 计算机辅助设计与图形学学报, 2019, 31(11): 1935-1945.

ZHONG W F, GUO F, XIANG S M, et al. Ship detection in remote sensing based with rotated rectangular region[J]. Journal of Computer-Aided Design & Computer Graphics, 2019, 31(11): 1935-1945.

[35] REN Y, ZHU C R, XIAO S P. Deformable Faster R-CNN with aggregating multi-layer features for partially occluded object detection in optical remote sensing images[J]. Remote Sensing, 2018, 10(9): 1470.

[36] DING P, ZHANG Y, DENG W J, et al. A light and faster regional convolutional neural network for object detection in optical remote sensing images[J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2018, 141: 208-218.

[37] ZHANG S M, WU R Z, XU K Y, et al. R-CNN-based ship detection from high resolution remote sensing imagery[J]. Remote Sensing, 2019, 11(6): 631.

[38] YAN J Q, WANG H Q, YAN M L, et al. IoU-adaptive deformable R-CNN: make full use of IoU for multi-class object detection in remote sensing imagery[J]. Remote Sensing, 2019, 11(3): 286.

[39] HE K M, GKIOXARI G, DOLLÁR P, et al. Mask R-CNN[C]//Proceedings of 2017 IEEE International Conference on Computer Vision. Piscataway: IEEE Press, 2017: 2980-2988.

[40] WU Q F, FENG D Q, CAO C Q, et al. Improved Mask R-CNN for aircraft detection in remote sensing images[J]. Sensors, 2021, 21(8): 2618.

[41] DING J, XUE N, LONG Y, et al. Learning RoI transformer for oriented object detection in aerial images[C]//Proceedings of 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE Press, 2019: 2844-2853.

[42] ZHU Y X, DU J, WU X Q. Adaptive period embedding for representing oriented objects in aerial images[J]. IEEE Transactions on Geoscience and Remote Sensing, 2020, 58(10): 7247-7257.

[43] XU Z Z, XU X, WANG L, et al. Deformable ConvNet with aspect ratio constrained NMS for object detection in remote sensing imagery[J]. Remote Sensing, 2017, 9(12): 1312.

[44] CHEN H B, JIANG S, HE G H, et al. TEANS: a target enhancement and attenuated nonmaximum suppression object detector for remote sensing images[J]. IEEE Geoscience and Remote Sensing Letters, 2021, 18(4): 632-636.

[45] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: unified, real-time object detection[C]//Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2016: 779-788.

[46] LIU W, ANGUELOV D, ERHAN D, et al. SSD: single shot multibox detector[C]//Proceedings of European Conference on Computer Vision. Berlin: Springer, 2016: 21-37.

[47] REDMON J, FARHADI A. YOLO9000: better, faster, stronger[C]//Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2017: 6517-6525.

[48] REDMON J, FARHADI A. YOLOv3: an incremental improvement[J]. arXiv Preprint, arXiv: 1804. 02767, 2018.

[49] BOCHKOVSKIY A, WANG C Y, LIAO H Y. YOLOv4: optimal speed and accuracy of object detection[J]. arXiv Preprint, arXiv: 2004. 10934, 2020.

[50] JOCHER G, STOKEN A, BOROVEC J, et al. Ultralytics/yolov5: v4.0[EB]. 2021.

[51] GE Z, LIU S T, WANG F, et al. YOLOX: exceeding YOLO series in 2021[J]. arXiv Preprint, arXiv: 2107. 08430, 2021.

[52] MA H J, LIU Y L, REN Y H, et al. Detection of collapsed buildings in post-earthquake remote sensing images based on the improved YOLOv3[J]. Remote Sensing, 2019, 12(1): 44.

[53] CHANG Y L, ANAGAW A, CHANG L N, et al. Ship detection based on YOLOv2 for SAR imagery[J]. Remote Sensing, 2019, 11(7): 786.

[54] WANG C S, WANG Q, WU H R, et al. Low-altitude remote sensing opium poppy image detection based on modified YOLOv3[J]. Remote Sensing, 2021, 13(11): 2130.

[55] XU D Q, WU Y Q. FE-YOLO: a feature enhancement network for remote sensing target detection[J]. Remote Sensing, 2021, 13(7): 1311.

[56] YANG X, YAN J C, FENG Z M, et al. R^3Det: refined single-stage detector with feature refinement for rotating object[C]//Proceedings of the 35th AAAI Conference on Artificial Intelligence. Palo Alto: AAAI Press, 2021: 3163-3171.

[57] LIU W C, MA L, WANG J, et al. Detection of multiclass objects in optical remote sensing images[J]. IEEE Geoscience and Remote Sensing Letters, 2019, 16(5): 791-795.

[58] XU D Q, WU Y Q. MRFF-YOLO: a multi-receptive fields fusion network for remote sensing target detection[J]. Remote Sensing, 2020, 12(19): 3118.

[59] HU J M, ZHI X Y, SHI T J, et al. PAG-YOLO: a portable attention-guided YOLO network for small ship detection[J]. Remote Sensing, 2021, 13(16): 3059.

[60] CHEN L Q, SHI W X, DENG D X. Improved YOLOv3 based on attention mechanism for fast and accurate ship detection in optical remote sensing images[J]. Remote Sensing, 2021, 13(4): 660.

[61] XU D Q, WU Y Q. Improved YOLO-V3 with DenseNet for multi-scale remote sensing target detection[J]. Sensors, 2020, 20(15): 4276.

[62] WU K J, BAI C S, WANG D C, et al. Improved object detection algorithm of YOLOv3 remote sensing image[J]. IEEE Access, 2021, 9: 113889-113900.

[63] PHAM M T, COURTRAI L, FRIGUET C, et al. YOLO-fine: one-stage detector of small objects under various backgrounds in remote sensing images[J]. Remote Sensing, 2020, 12(15): 2501.

[64] VAN ETTEN A. You only look twice: rapid multi-scale object detection in satellite imagery[J]. arXiv Preprint, arXiv: 1805. 09512, 2018.

[65] HONG Z H, YANG T, TONG X H, et al. Multi-scale ship detection from SAR and optical imagery via a more accurate YOLOv3[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2021, 14: 6083-6101.

[66] SHARMA M, DHANARAJ M, KARNAM S, et al. YOLOrs: object detection in multimodal remote sensing imagery[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2020, 14: 1497-1508.

[67] LU X C, JI J, XING Z Q, et al. Attention and feature fusion SSD for remote sensing object detection[J]. IEEE Transactions on Instrumentation and Measurement, 2021, 70: 1-9.

[68] QU J S, SU C, ZHANG Z W, et al. Dilated convolution and feature fusion SSD network for small object detection in remote sensing images[J]. IEEE Access, 2020, 8: 82832-82843.

[69] YIN R H, ZHAO W, FAN X D, et al. AF-SSD: an accurate and fast single shot detector for high spatial remote sensing imagery[J]. Sensors, 2020, 20(22): 6530.

[70] BAO S Z, ZHONG X, ZHU R F, et al. Single shot anchor refinement network for oriented object detection in optical remote sensing imagery[J]. IEEE Access, 2019, 7: 87150-87161.

[71] TANG T Y, ZHOU S L, DENG Z P, et al. Arbitrary-oriented vehicle detection in aerial imagery with single convolutional neural networks[J]. Remote Sensing, 2017, 9(11): 1170.

[72] LIU L, PAN Z X, LEI B. Learning a rotation invariant detector with rotatable bounding box[J]. arXiv Preprint, arXiv: 1711. 09405, 2017.

[73] 郭智, 宋萍, 张义, 等. 基于深度卷积神经网络的遥感图像飞机目标检测方法[J]. 电子与信息学报, 2018, 40(11): 2684-2690.
GUO Z, SONG P, ZHANG Y, et al. Aircraft detection method based on deep convolutional neural network for remote sensing images[J]. Journal of Electronics & Information Technology, 2018, 40(11): 2684-2690.

[74] 李晖晖, 周康鹏, 韩太初. 基于 CReLU 和 FPN 改进的 SSD 舰船目标检测[J]. 仪器仪表学报, 2020, 41(4): 183-190.
LI H H, ZHOU K P, HAN T C. Ship object detection based on SSD improved with CReLU and FPN[J]. Chinese Journal of Scientific Instrument, 2020, 41(4): 183-190.

[75] WANG P J, SUN X, DIAO W H, et al. FMSSD: feature-merged single-shot detection for multiscale objects in large-scale remote sensing imagery[J]. IEEE Transactions on Geoscience and Remote Sensing, 2020, 58(5): 3377-3390.

[76] SCHILLING H, BULATOV D, NIESSNER R, et al. Detection of vehicles in multisensor data via multibranch convolutional neural networks[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2018, 11(11): 4299-4316.

[77] QIAN W, YANG X, PENG S L, et al. Learning modulated loss for rotated object detection[C]//Proceedings of the AAAI Conference on Artificial Intelligence. Palo Alto: AAAI Press, 2021: 2458-2466.

[78] XU Y C, FU M T, WANG Q M, et al. Gliding vertex on the horizontal bounding box for multi-oriented object detection[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2021, 43(4): 1452-1459.

[79] YANG X, YANG J R, YAN J C, et al. SCRDet: towards more robust detection for small, cluttered and rotated objects[C]//Proceedings of 2019 IEEE/CVF International Conference on Computer Vision (ICCV). Piscataway: IEEE Press, 2019: 8231-8240.

[80] HOU L P, LU K, XUE J, et al. Cascade detector with feature fusion for arbitrary-oriented objects in remote sensing images[C]//Proceedings of 2020 IEEE International Conference on Multimedia and Expo. Piscataway: IEEE Press, 2020: 1-6.

[81] YANG X, YAN J C, HE T. Arbitrary-oriented object detection with circular smooth label[J]. arXiv Preprint, arXiv: 2003. 05597, 2020.

[82] YANG X, HOU L P, ZHOU Y, et al. Dense label encoding for boundary discontinuity free rotation detection[C]//Proceedings of 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE Press, 2021: 15814-15824.

[83] YANG X, YAN J C, MING Q, et al. Rethinking rotated object detection with Gaussian Wasserstein distance loss[C]//Proceedings of International Conference on Machine Learning. New York: ACM Press, 2021: 11830-11841.

[84] PAN X J, REN Y Q, SHENG K K, et al. Dynamic refinement network for oriented and densely packed object detection[C]//Proceedings of 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE Press, 2020: 11204-11213.

[85] WEI H R, ZHANG Y, CHANG Z H, et al. Oriented objects as pairs of middle lines[J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2020, 169: 268-279.

[86] FENG P M, LIN Y T, GUAN J, et al. TOSO: student's-T distribution aided one-stage orientation target detection in remote sensing images[C]//Proceedings of 2020 IEEE International Conference on Acoustics, Speech and Signal Processing. Piscataway: IEEE Press, 2020: 4057-4061.

[87] LI Y Y, HUANG Q, PEI X, et al. RADet: refine feature pyramid network and multi-layer attention network for arbitrary-oriented object detection of remote sensing images[J]. Remote Sensing, 2020, 12(3): 389.

[88] CHENG G, ZHOU P C, HAN J W. Learning rotation-invariant convolutional neural networks for object detection in VHR optical remote sensing images[J]. IEEE Transactions on Geoscience and Remote Sensing, 2016, 54(12): 7405-7415.

[89] LI C Y, CONG R M, GUO C L, et al. A parallel down-up fusion network for salient object detection in optical remote sensing images[J]. Neurocomputing, 2020, 415: 411-420.

[90] ZHANG Q J, CONG R M, LI C Y, et al. Dense attention fluid network for salient object detection in optical remote sensing images[J]. IEEE Transactions on Image Processing, 2021, 30: 1305-1317.

[91] DENG Z P, SUN H, ZHOU S L, et al. Multi-scale object detection in remote sensing imagery with convolutional neural networks[J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2018, 145: 3-22.

[92] DONG Z P, WANG M, WANG Y L, et al. Object detection in high resolution remote sensing imagery based on convolutional neural networks with suitable object scale features[J]. IEEE Transactions on Geoscience and Remote Sensing, 2020, 58(3): 2104-2114.

[93] WANG C, BAI X, WANG S, et al. Multiscale visual attention networks for object detection in VHR remote sensing images[J]. IEEE Geoscience and Remote Sensing Letters, 2019, 16(2): 310-314.

[94] LI C Y, LUO B, HONG H L, et al. Object detection based on global-local saliency constraint in aerial images[J]. Remote Sensing, 2020, 12(9): 1435.

[95] YU Y T, GUAN H Y, LI D L, et al. Orientation guided anchoring for geospatial object detection from remote sensing imagery[J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2020, 160: 67-82.

[96] HUA X, WANG X Q, RUI T, et al. A fast self-attention cascaded network for object detection in large scene remote sensing images[J]. Applied Soft Computing, 2020, 94: 106495.

[97] ZHENG Z, ZHONG Y F, MA A L, et al. HyNet: hyper-scale object detection network framework for multiple spatial resolution remote sensing imagery[J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2020, 166: 1-14.

[98] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need[C]//Annual Conference on Neural Information Processing Systems. Cambridge: MIT Press, 2017: 5998-6008.

[99] ZHU X K, LYU S C, WANG X, et al. TPH-YOLOv5: improved YOLOv5 based on transformer prediction head for object detection on drone-captured scenarios[C]//Proceedings of 2021 IEEE/CVF International Conference on Computer Vision Workshops (ICCVW). Piscataway: IEEE Press, 2021: 2778-2788.

[100] QING Y H, LIU W Y, FENG L Y, et al. Improved YOLO network for free-angle remote sensing target detection[J]. Remote Sensing, 2021, 13(11): 2171.

[101] HEITZ G, KOLLER D. Learning spatial context: using stuff to find things[C]//Lecture Notes in Computer Science. Berlin: Springer, 2008: 30-43.

[102] FRANKLIN T, BRIAN C, CRAIG P, et al. Overhead imagery research data set - an annotated data library & tools to aid in the development of computer vision algorithms[C]//Proceedings of 2009 IEEE Applied Imagery Pattern Recognition Workshop (AIPR 2009). Piscataway: IEEE Press, 2009: 1-8.

[103] MAGGIORI E, TARABALKA Y, CHARPIAT G, et al. Can semantic labeling methods generalize to any city? the inria aerial image labeling benchmark[C]//Proceedings of 2017 IEEE International Geoscience and Remote Sensing Symposium. Piscataway: IEEE Press, 2017: 3226-3229.

[104] ZHU H G, CHEN X G, DAI W Q, et al. Orientation robust object detection in aerial images using deep convolutional neural network[C]//Proceedings of 2015 IEEE International Conference on Image Processing. Piscataway: IEEE Press, 2015: 3735-3739.

[105] 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: 119-132.

[106] RAZAKARIVONY S, JURIE F. Vehicle detection in aerial imagery: a small target detection benchmark[J]. Journal of Visual Communication and Image Representation, 2016, 34: 187-203.

[107] LIU Z K, YUAN L, WENG L B, et al. A high resolution optical satellite image dataset for ship recognition and some new baselines[C]//Proceedings of the 6th International Conference on Pattern Recognition Applications and Methods. Francisco: SciTePress, 2017: 324-331.

[108] LIU K, MATTYUS G. Fast multiclass vehicle detection on aerial images[J]. IEEE Geoscience and Remote Sensing Letters, 2015, 12(9): 1938-1942.

[109] ZHANG Y L, YUAN Y, FENG Y C, et al. Hierarchical and robust convolutional neural network for very high-resolution remote sensing object detection[J]. IEEE Transactions on Geoscience and Remote Sensing, 2019, 57(8): 5535-5548.

[110] ZOU Z X, SHI Z W. Random access memories: a new paradigm for target detection in high resolution aerial remote sensing images[J]. IEEE Transactions on Image Processing, 2018, 27(3): 1100-1111.

[111] YANG Y M. ITCVD dataset[EB]. 2018.

[112] XIA G S, BAI X, DING J, et al. DOTA: a large-scale dataset for object detection in aerial images[C]//Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2018: 3974-3983.

[113] SUN X, WANG P J, YAN Z Y, et al. FAIR1M: a benchmark dataset for fine-grained object recognition in high-resolution remote sensing imagery[J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2022, 184: 116-130.

[114] CHENG G, YAN B W, SHI P Z, et al. Prototype-CNN for few-shot object detection in remote sensing images[C]//Proceedings of IEEE

Transactions on Geoscience and Remote Sensing. Piscataway: IEEE Press, 2022: 1-10.

[115] LI X, DENG J Y, FANG Y. Few shot object detection in remote sensing images[C]//Proceedings of IEEE Transactions on Geoscience and Remote Sensing. Piscataway: IEEE Press, 2021: 76-81.

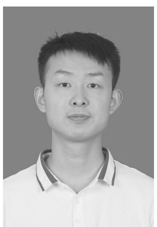
[116] CUI Y Y, HOU B, WU Q, et al. Remote sensing object tracking with deep reinforcement learning under occlusion[J]. IEEE Transactions on Geoscience and Remote Sensing, 2022, 60: 1-13.

[117] MERANER A, EBEL P, ZHU X X, et al. Cloud removal in Sentinel-2 imagery using a deep residual neural network and SAR-optical data fusion[J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2020, 166: 333-346.

[118] DU Q, TANG B, XIE W Y, et al. Parallel and distributed computing for anomaly detection from hyperspectral remote sensing imagery[J]. Proceedings of the IEEE, 2021, 109(8): 1306-1319.

[作者简介]

廖育荣（1972– ），男，四川德阳人，博士，航天工程大学研究员，主要研究方向为空间智能信息处理、航天测控通信、卫星信号处理技术等。



王海宁（1998– ），男，河南南阳人，航天工程大学硕士生，主要研究方向为遥感图像处理、深度学习、目标检测。

林存宝（1987– ），男，吉林延边人，博士，航天工程大学助理研究员，主要研究方向为空间光学载荷、空间智能信息处理。

李阳（1993– ），女，辽宁大连人，航天工程大学博士生，主要研究方向为遥感图像处理、目标检测。

方宇强（1984– ），男，四川乐山人，博士，航天工程大学副教授，主要研究方向为计算机视觉、机器学习、深度学习。

倪淑燕（1981– ），女，河北清河人，博士，航天工程大学副教授，主要研究方向为数字图像处理、空间智能信息处理等。