

光学遥感图像目标检测数据集综述

袁一钦, 李浪, 姚西文, 李玲君, 冯晓绪, 程臻, 韩军伟

西北工业大学 自动化学院, 西安 710021

摘要:近年来,随着深度学习等人工智能技术在光学遥感目标检测领域中的快速发展,大量相关研究算法不断涌现,逐渐形成了一种基于数据驱动的光学遥感图像目标检测新范式。高质量的遥感数据成为了此类范式算法研究的前置条件和必要储备,遥感数据的重要性日益凸显。迄今为止,国内外各大研究机构已相继发布了数量众多且规模不一的光学遥感图像目标检测数据集,为基于深度学习的光学遥感图像目标检测算法的发展奠定了研究基础。然而,当前尚未有相关学者对已发布的光学遥感图像检测数据集进行全面的归纳整理与分析,针对此问题,本文全面调研领域文献,对2008年—2023年期间已发布的公开光学遥感图像检测数据集进行整合分析,并依据不同的数据标注方式进行划分,对其中的11个典型数据集进行了全面阐述,以表格的形式对所有的数据集信息进行归纳总结,同时采用3种分析方式去描述数据集的发展情况,即:元数据分析,从数量分布、地域分布、来源分布、规模分布着手;分辨率分析,从空间分辨率与光谱分辨率着手;基本信息分析,从类别数量、图像数量、实例数量及图像宽度信息着手,有效论证了光学遥感图像目标检测数据集必然朝着高质量、大规模、多类别的方向发展。此外,针对已发布的数据集,从水平框目标检测、旋转框目标检测以及细分检测方向(小目标检测和细粒度检测)等多个角度对相关算法的应用和发展进行了概述,证实了遥感数据对目标检测算法的研究具有积极的推动作用。综上,本文将为基于深度学习的目标检测算法在遥感领域的应用提供参考。

关键词:深度学习,光学遥感图像,数据源,目标检测,数据集发展

中图分类号: TP701/P2

引用格式:袁一钦,李浪,姚西文,李玲君,冯晓绪,程臻,韩军伟.2023.光学遥感图像目标检测数据集综述.遥感学报,27(12):2671-2687

Yuan Y Q, Li L, Yao X W, Li L J, Feng X X, Cheng G and Han J W. 2023. A comprehensive review of optical remote-sensing image object detection datasets. National Remote Sensing Bulletin, 27(12): 2671-2687 [DOI: 10.11834/jrs.20233457]

1 引言

光学遥感(聂光涛和黄华,2021)主要通过可以发射可见光、紫外光以及红外光的观测器进行对地拍摄,由此获取的影像称为光学遥感图像,其一般具有时效强、范围广的特点,通常包含大量高分辨率的空间地理信息。因此,遥感影像解析(姚艳清等,2021)在城乡规划、救援救灾、机场和港口监测等领域都有着很大的应用前景。

目标检测是遥感影像解析(周培诚等,2021)多种任务的基础,即从光学遥感图像中检测出感兴趣的目标或区域,如汽车(Mundhenk等,2016)、桥梁(Nogueira等,2019)、船舶(Liu等,2017)

等。近年来,光学遥感图像目标检测发展迅速,以下从数据和算法两个方面进行阐述。

早期一般通过卫星等提取某一地理空间的整体属性,比如,辨别某一区域的地形属于高原还是山地。由于空间分辨率较低,很难通过相关影像解译技术实现某种特定类别目标的检测与识别。近年来,光学遥感技术逐渐成熟,遥感影像空间分辨率也在不断提升,光学遥感图像中常见类型的目标均已达到肉眼可分辨的程度。因而,近年出现了大量大规模遥感数据集,为相关技术的发展提供了必要条件。

另一方面,近年来深度学习的快速崛起,极大地促进了遥感影像解析技术的发展。传统的遥

收稿日期:2023-11-01;预印本:2023-11-14

基金项目:国家自然科学基金(编号:62071388,62136007);陕西省重点研发计划(编号:2023-YBGY-224)

第一作者简介:袁一钦,研究方向为深度学习,遥感图像处理。E-mail:NPUyuan@163.com

通信作者简介:姚西文,研究方向为资源受限下的遥感智能解译。E-mail:yaowen@nwpu.edu.cn

感影像解译算法, 由于缺少有效的图像特征表示方法, 需要手动设计特征表示算子。因此, 特征提取网络的设计决定了目标检测算法的最终性能。深度学习算法则不需要设计精巧的特征提取方法, 而是直接通过深度神经网络进行自动地特征提取与表征, 因而深度学习算法的目标检测性能相比传统算法显著提升。

深度学习算法通过提取目标的有效特征、表征大量数据之间的分布特性来达到学习的目的。因此, 数据集的规模与质量决定了算法模型的性能。近年来, 出现了许多与遥感目标检测相关的数据集。根据其标注方式大致可以分为两种类型: 水平框 (HBB) 标注和旋转框 (OBB) 标注 (聂光涛和黄华, 2021)。水平框常见的标注方式有3种: (x, y, w, h) , 其中 (x, y) 表示目标边界框的中心点坐标, w 和 h 分别表示目标边界框的宽和高; (x_1, y_1, x_2, y_2) , 其中 (x_1, y_1) 和 (x_2, y_2) 分别表示目标边界框的左上角点和右下角点的坐标; $(x_{\min}, y_{\min}, w, h)$, 其中 (x_{\min}, y_{\min}) 表示目标边界框的左上角点坐标, w 和 h 分别表示目标边界框的宽和高。常见的旋转框标注方式有两种: (x, y, w, h, θ) , 其中 (x, y) 、 w 、 h 分别表示目标边界框的中心点坐标以及宽和高, 而 θ 表示目标边界框相对于水平方向的旋转角度; $(x_1, y_1, x_2, y_2, x_3, y_3, x_4, y_4)$, 其中 (x_i, y_i) 、 $(i = 1, 2, 3, 4)$ 分别表示目标边界框的4个顶点坐标。

本文收集了15年来与遥感领域目标检测相关的数据集, 对数据集进行了全面的介绍, 并从多个角度进行了简明扼要的分析, 同时, 论述了相关深度学习算法的发展情况, 旨在对光学遥感图像目标检测数据集提供一个综合全面的概述。

2 光学遥感图像目标检测数据集

光学遥感图像目标检测任务的关键是对遥感图像中存在的目标进行准确定位并将其归属至对应类别 (Cheng 和 Han, 2016)。深度学习技术的发展不断对遥感图像数据提出新的要求 (Li 等, 2020a), 截至目前, 国内外已有多家单位累计发布了数十种光学遥感图像目标检测数据集。具体而言, 主流的标注方式可将数据集划分为两类: 旋转框标注与水平框标注, 数据集相关标注样例如图1所示, 典型的遥感图像水平框标注数据集包

括: xView (Lam 等, 2018)、VisDrone (Zhu 等, 2022)、DIOR (Li 等, 2020a)、AI-TOD (Wang 等, 2021b)、AI-TOD-v2 (Xu 等, 2022a)、RarePlanes (Shermeyer 等, 2021) 等; 典型的遥感图像旋转框标注数据集包括: VEDAI (Razakarivony 和 Jurie, 2016)、FAIR1M (Sun 等, 2022)、DIOR-R (Cheng 等, 2022)、SODA-A (Cheng 等, 2023) 等。此外, HRSC2016 (Liu 等, 2017) 同时具备水平框、旋转框与关键部位3种标注方式, DOTA (Xia 等, 2018)、DOTA-v2.0 (Ding 等, 2022)、FGSD (Chen 等, 2020) 等同时具备水平框与旋转框两种标注方式, COWC (Mundhenk 等, 2016) 则采用的是中心点标注。

2.1 光学遥感图像数据集

2.1.1 xView

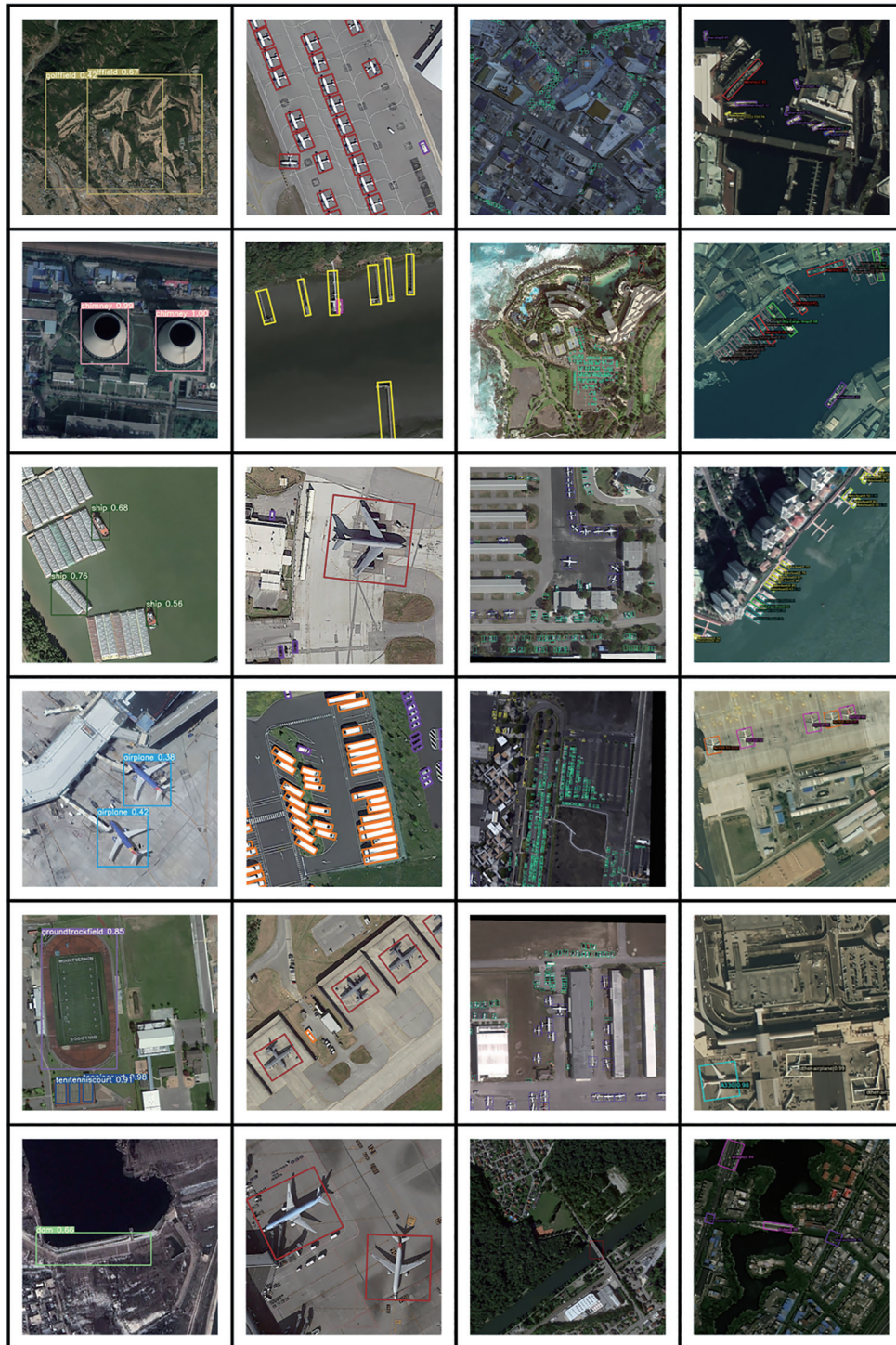
xView (Lam 等, 2018) 数据集是 Darius Lam 等于2018年发布的面向光学遥感图像目标检测任务的大型水平框数据集。数据集影像来源于 WorldView-3 卫星, 空间分辨率为 0.3 m, 高分辨率为目标检测任务提供了极佳的图景信息。整个数据集包含图像 1129 张, 覆盖了超过 1400 km² 土地上的 60 类目标, 实例数量约 100 余万。为了适应于细粒度目标检测任务的需求, 60 个类别里有 44 个类别从属于其中的 7 个父类 (固定翼飞机、客运车辆、卡车、铁路车辆、工程车、海船、建筑)。该数据集的提出有效解决了已有数据集存在的类别数量少、地理多样性差、训练实例少以及类别范围过窄等问题, 仿真了现实世界目标分布不均的情形。

2.1.2 VisDrone

VisDrone (Zhu 等, 2022) 数据集是由天津大学 Zhu 等于2018年从无人机航拍视角获取图像, 面向光学遥感图像目标检测所发布的水平框数据集, 并于2019年在原先发布的数据集基础之上进行了补充, 该数据集除了用于图像目标检测, 数据集的拓展部分 (视频标注) 还可用于视频目标检测、目标追踪、密集计数等任务。该数据集的图像源于各类无人机平台, 所选图像涵盖了中国 14 个城市的差异化场景, 包含了 10209 张图像, 涉及 10 个类别 (行人、小轿车、厢式货车、公共汽车、卡车、人、摩托车、自行车、遮阳篷三轮

车和三轮车) 超 54 万个对象实例, 图像的最大尺寸达到 2000 像素×1500 像素。该数据集所选图像关注了大尺度目标、视点变化以及相关目标存在

严重遮挡的问题, 极大增强了数据集对真实场景的再现, 推动了领域相关算法的发展。



(a) 水平框 (b) 旋转框 (c) 小目标 (d) 细粒度
(a) Horizontal bounding box (b) Oriented bounding box (c) Small object (d) Fine grain

图 1 光学遥感图像目标检测数据集样例
Fig. 1 Dataset samples for object detection in optical remote sensing images

2.1.3 DIOR

DIOR (Li 等, 2020a) 数据集是西北工业大学程臻等于2019年发布的面向光学遥感图像通用目标检测的大规模水平框数据集。数据集图像来源于 Google Earth 等, 空间分辨率普遍位于 0.5—30 m。该数据集由 23463 张图像组成, 覆盖了 20 个常见目标类别的 192472 个实例, 图像大小均为 800 像素×800 像素。在该数据集中, 飞机、船等类别的实例数量较多, 而机场、港口则较少, 与现实世界的实例分布相一致; 数据集中同一类别的实例可能源于不同成像条件、不同天气、不同季节、不同图像质量的生成, 导致呈现出完全不同的效果, 极大增强了数据集的鲁棒性; 数据集中实例的类间相似性较低, 而类内多样性较高, 且实例尺寸变化范围大, 进一步增强了数据集的可应用度。DIOR 数据集的提出极大弥补了已有数据集图像数量和对象类别规模较小, 图像多样性和变化性不足的问题。西北工业大学的程臻等于 2021 年在原始的 DIOR 数据集基础之上重新进行了有向框的标注, 提出了 DIOR-R (Cheng 等, 2022) 旋转框数据集, 该数据集包含 23463 张图像, 共 192512 个实例。

2.1.4 AI-TOD

AI-TOD (Wang 等, 2021b) 数据集是由武汉大学 Wang 等于 2020 年面向光学遥感图像小目标检测任务所发布的大规模水平框数据集。该数据集的图像来源于 xView (Lam 等, 2018)、VisDrone (Zhu 等, 2022)、DIOR (Li 等, 2020a)、DOTA (Xia 等, 2018) 以及 Kaggle 挑战赛中的船舶检测数据集, 共 28036 张图像, 涵盖 8 个目标类别 (飞机, 桥梁, 储罐, 船舶, 游泳池, 车辆, 人, 风车), 700621 个实例。该数据集中的图像大小均为 800 像素×800 像素, 对象实例的平均尺寸大小约为 12.8 像素, 最大对象实例的大小不超过 64 像素。该数据的提出填补了光学遥感图像小目标检测领域数据集的空白, 为后续光学遥感图像小目标检测任务的发展奠定了基础。武汉大学 Xu 等 (2022a) 于 2022 年对原有数据集中存在的注释缺失和标注错误问题进行了修补, 对数据集进行了更精细地标注, 提出了 AI-TOD-v2, 实例的数量经修补扩充后为 752745 个。

2.1.5 RarePlanes

RarePlanes (Shermeyer 等, 2021) 数据集是由 Shermeyer 等于 2020 年发布的面向光学遥感图像目标检测的水平框数据集。该数据集的影像源于两部分, 一部分为真实拍摄所得的遥感图像, 包含 253 个 Maxar WorldView-3 卫星场景, 覆盖了 22 个国家 112 处地点, 达 2142 km², 涵盖了 14707 个飞机实例; 另一部分为人工合成的图像, 总计 50000 张, 模拟总面积为 9331.2 km², 生成了约 63 万个飞机实例。该数据集的真实飞机实例和合成飞机实例均具有 10 种精细的纹理属性 (飞机长度、翼展、机翼形状、机翼位置、翼展等级、推进类型、引擎数量、垂直稳定器数量、鸭翼以及飞机类型) 和 33 种子属性, 基于上述属性可以构建从 1—110 个飞机类别。

2.1.6 VEDAI

VEDAI (Razakarivony 和 Jurie, 2016) 数据集是由 Razakarivony 等于 2015 年发布的面向光学遥感图像目标检测的旋转框数据集。该数据集主要聚焦于遥感图像中小型交通工具的检测, 其中图像的背景涵盖森林、城市、道路、停车场、建筑工地及田野等复杂情形。数据集总共包含 1210 张图像, 覆盖了 9 个常见的交通工具类别 (船舶、小轿车、野营车、飞机、皮卡、拖拉机、卡车、厢式货车和其他类别), 共 3640 个实例, 图像大小分为 1024 像素×1024 像素与 512 像素×512 像素两种类型。数据集图像的空间分辨率为 0.125 m, 每幅图像包含 4 个通道 (RGB 与 1 个近红外)。数据集中的目标除了体积小、方向不定之外, 还面临光照变化、遮挡等情形, 使得数据集具备了较强的鲁棒性和抗干扰性。该数据集的提出促进了高分遥感图像目标检测相关算法的发展, 使得评估图像分辨率或颜色波段对检测结果的影响成为可能。

2.1.7 HRSC2016

HRSC2016 (Liu 等, 2017) 数据集是由 Liu 等于 2016 年发布的面向光学遥感图像细粒度船舶检测的数据集。该数据集影像均来源于 6 个标志性的港口 (Google Earth 平台提供), 只包含海上船舶与近岸船舶两种场景, 图像空间分辨率位于 0.4—2 m, 图像大小不一, 从 300 像素×300 像素到 1500 像素×900 像素均有分布, 但多数图像尺寸均大于 1000 像

素×600像素。具体而言,数据集共有1061张图像,2976个对象实例,采用了3种标注方式:水平框标注、旋转框标注与关键部位标注,并对船舶进行了三级分类,以适应于细粒度船舶检测任务。HRSC2016是第一个从公开高分辨率遥感图像中建构的聚焦于细粒度船舶检测的公开数据集,有力促进了光学遥感图像相关算法的发展,具有重要意义。

2.1.8 DOTA

DOTA (Xia等, 2018)数据集是由武汉大学的Xia等于2018年发布的面向光学遥感图像目标检测的大规模旋转框数据集。该数据集的图像来源于中国资源卫星数据与应用中心提供的Google Earth、GF-2卫星、JL-1卫星及CycloMedia B.V.航拍影像,具备了多种分辨率,有效保证了数据的多样性。具体而言,数据集总共有2806张图像,图像大小普遍位于800像素×800像素到4000像素×4000像素之间,涵盖了常见的15个目标类别(飞机、船舶、储罐、棒球场、网球场、游泳池、地面田径场、港口、桥梁、大型车辆、小型车辆、直升机、环形交叉路口、足球场和篮球场),共计188282个对象实例,且以任意四边形的方式对所有实例进行了标注。与在此之前提出的数据集相比,如TAS (Heitz和Koller, 2008)、VEDAI (Razakarivony和Jurie, 2016)、UCAS-AOD (Zhu等, 2015)、HRSC2016 (Liu等, 2017),DOTA数据集实例数量极大且涵盖了各种常见的场景,对相关算法的研究推进更具帮助。武汉大学于2019年发布了DOTA-v1.5数据集,在原有数据集的基础上新增了一个类别(集装箱起重机),并对原有图像进行了更为细致的标注,使得实例数量达到402089个。武汉大学的Ding等于2021年发布了DOTA-v2.0 (Ding等, 2022)数据集,相较于DOTA-v1.5数据集,新增了机场和直升机停机坪两个类别,达到18类,包含图像共计11268张,涉及1793658个对象实例。

2.1.9 FGSD

FGSD (Chen等, 2020)数据集是由北京邮电大学的Chen等于2020年发布的面向高分辨率光学遥感图像细粒度舰船检测的数据集。数据集的图像来源于Google Earth平台,覆盖4个国家(中国、

日本、美国、西班牙)的17个大型港口,图像的空间分辨率处于0.12—1.93 m。数据集共包含图像2612张,对象实例5634个,图像大小均为930像素×930像素,所有图像都被采用了两种标注方式:水平框标注和旋转框标注。为适应细粒度检测任务需求,对所有存在的舰船目标采用了三级分类的方式,共1个一级类,4个二级类以及43个三级类,此外还添加了一个码头类。与HRSC2016 (Liu等, 2017)数据集相比,FGSD数据集类别数和对象实例数更多,为相关细粒度船舶检测算法的应用提供了更为复杂的检测背景与特殊的目标视角,促进了相关强鲁棒算法的提出,对海事管理、贸易估算等领域具有极大的实际应用价值。

2.1.10 FAIR1M

FAIR1M (Sun等, 2022)数据集是由Sun等于2021年发布的面向高分辨率光学遥感图像细粒度目标检测的大规模旋转框数据集。该数据集的图像来源于高分卫星和Google Earth平台,空间分辨率位于0.3—0.8 m,图像大小普遍位于1000像素×1000像素到10000像素×10000像素之间,图像中的场景横跨多个大洲,覆盖了全世界100多个民用机场、港口和城市。具体而言,该数据集包含高分遥感图像15266张,超过100万个实例,划分为5个大类(飞机、船舶、车辆、球场、道路),进一步可细分为37个小类。与已有细粒度检测数据集相比,FAIR1M数据集实例数量和图像数量均大得多,且它为遥感图像中的目标提供了更丰富的细粒度分类信息以及更好的图像质量,对细粒度目标检测的发展起到了极大推动作用。

2.1.11 SODA-A

SODA-A (Cheng等, 2023)数据集是由西北工业大学的程堃等于2022年发布的面向光学遥感图像小目标检测的大规模旋转框数据集。该数据集的图像来源于Google Earth平台,大部分图像的尺寸超过4700像素×2700像素。具体而言,该数据集共计2513张图像,872069个对象实例,划分为9个类别(飞机、直升机、小型车辆、大型车辆、船舶、集装箱、储罐、游泳池、风车)。与已有的数据集相比,SODA-A中的图像背景十分多样,图像中的目标密度变化剧烈且经常以任意方向出现,不同类别间的对象实例数量相对均衡(直升机类

别除外), 高分辨率的图像为小目标的识别提供了更精细的细节信息, 故该数据集的提出对遥感图像小目标检测具备十分重要的意义。

除了上述较为典型、应用较为广泛的光学遥感图像目标检测数据集之外, 还有一些提出时间较早、规模相对较小的相关数据集, 水平框标注的包括: TAS (Heitz 和 Koller, 2008)、UCAS-AOD (Zhu 等, 2015)、NWPU VHR-10 (Cheng

等, 2014)、LEVIR (Zou 和 Shi, 2018)、HRRSD (Zhang 等, 2019)、RSD-GOD (Zhuang 等, 2019) 等, 旋转框标注的包括: OIRDS (Tanner 等, 2009)、SZTAKI-INRIA (Benedek 等, 2012)、DLR 3K Vehicle (Liu 和 Mattyus, 2015) 等。所有常用的光学遥感图像目标检测数据集的相关信息汇总如表 1 所示。

表 1 光学遥感图像目标检测数据集

Table 1 Object detection datasets in optical remote sensing images

数据集名称	数据集内容	标注方式	数据集大小
TAS (Heitz 和 Koller, 2008)	包含 30 张大小为 792×636 的遥感影像, 1319 辆汽车	HBB	921.32 MB
SZTAKI-INRIA (Benedek 等, 2012)	包含 9 张遥感图像和 665 个建筑物目标, 图像宽约 800 像素	OBB	4.49 MB
OIRDS (Tanner 等, 2009)	900 张遥感影像、1800 辆汽车。包括 4 个目标类型, 分别为小汽车、卡车、皮卡和未知车型	OBB	154 MB
NWPU VHR-10 (Cheng 等, 2014)	650 张有目标的图像和 150 张背景图像, 共计 800 张。其中 715 张来自谷歌地球, 空间分辨率为 0.5—2 m, 另外 85 张为来自 Vaihingen 数据的全锐彩色红外图像, 空间分辨率为 0.08 m。共包含 10 个目标类别	HBB	73 MB
UCAS-AOD (Zhu 等, 2015)	两类目标, 汽车和飞机, 还有一部分是背景样本。汽车类别包括 310 张图像和 2819 个实例; 飞机类别包括 600 张图像和 3210 个实例	HBB	3.24 GB
VEDAI (Razakarivony 和 Jurie, 2016)	由 3640 辆交通工具组成, 包括 9 个类别: 船、轿车、野营车、飞机、皮卡、拖拉机、卡车、厢式货车和其他类别。包含 1210 张航拍图像, 空间分辨率为 12.5 cm	OBB	2.76 GB
DLR 3K Vehicle (Liu 和 Mattyus, 2015)	包含 20 张图像、14235 个目标。图像尺寸为 5616×3744, 空间分辨率 0.13 m。包括两个类别: 汽车和卡车	OBB	162.43 MB
HRSC2016 (Liu 等, 2017)	共 1061 幅图像, 2976 个目标。其中海洋图像 70 幅, 样本 90 个; 海陆图像 991 幅, 样本 2886 个。其中, 训练集包含 436 张图像 (1207 个样本)、验证集 181 张图像 (541 个样本)、测试集 444 张图像 (1228 个样本)	HBB/OBB	3.95 GB
COWC (Mundhenk 等, 2016)	包含 32716 辆汽车和 58247 个负样例。目标采用中心点注释; 包含与汽车相似的负样例; 目标标注包含一定的上下文信息	Point	54 GB
CARPK (Hsieh 等, 2017)	包括 1448 张图像和 89777 辆汽车。图像是在大约 40 m 高度的无人机视角下收集的, 单个场景中最多含有 188 辆汽车	HBB	2.0 GB
RSOD (Xiao 等, 2015)	976 张遥感图像, 6950 个实例。包括飞机、操场、立交桥和油罐 4 个场景	HBB	309.5 MB
VisDrone (Zhu 等, 2022)	10209 张无人机航拍图像, 其中训练集 6471 张, 验证集 548 张, 测试挑战集 1580 张, 测试开发集 1610 张。包括 10 个类别, 每个类别约 54.2 K 个实例	HBB	2.07 GB
xView (Lam 等, 2018)	1129 张图像, 约 1M 个实例。包括 60 个父类与子类, 其中 7 个父类为固定翼飞行器、客运车辆、卡车、铁路车辆、工程车辆、海上船舶、建筑物。另外, 并不是所有的子类都有父类	HBB	19.72 GB
LEVIR (Zou 和 Shi, 2018)	21952 张图像、11028 个实例。图像大小为 800×600、分辨率 0.2—1 m。包括 4724 架飞机、3025 艘船和 3279 个油罐	HBB	1.6 GB
ITCVD (Yang 等, 2018)	原始数据包含 228 张图像, 分辨率为 5616×3744。筛选其中的 173 张图像作为最终的数据集、包含 29088 辆汽车。训练集: 135 张图像、23543 辆汽车; 测试集: 38 张图像、5545 辆汽车	HBB	1.27 GB
DOTA (Xia 等, 2018)	2806 张遥感影像与 188282 个实例。包含 15 个目标类别。为了确保训练数据和测试数据分布近似匹配, 随机选择一半的原始图像作为训练集, 1/6 作为验证集, 1/3 作为测试集	HBB/OBB	18.8 GB
DOTA-v1.5	DOTA-v1.5 使用与 DOTA-v1.0 相同的图像, 但注释了极小的实例 (小于 10 像素)。此外, 增加了一个新的类别“集装箱起重机”, 总共包含 402089 个实例	HBB/OBB	18.8 GB

续表

数据集名称	数据集内容	标注方式	数据集大小
DOTA-v2.0 (Ding等,2022)	包括11268张图片和1793658个实例。与DOTAv1.5相比,它进一步增加了“机场”和“直升机停机坪”两个类别。为了避免过拟合问题,训练集和验证集的比例小于测试集的比例	HBB/OBB	137.5 GB
RSD-GOD (Zhuang等,2019)	包含18187张图像、40990个实例,包括5个类别:机场、飞机、直升机、军舰、油罐	HBB	1.25 GB
HRRSD (Zhang等,2019)	包括21761幅空间分辨率为0.15—1.2 m的Google Earth彩色影像和4961幅空间分辨率为0.6—1.2 m的百度地图彩色影像。包含13个类别、55740个实例	HBB	约4 GB
Bridges Dataset (Nogueira等,2019)	包含来自全球各地的500张不同桥梁的图像,由于图像和桥梁属性的不同,如建筑的方向、景观背景(河流或山)、区域(城市或乡村)等,增加了其多样性和代表性	HBB	1.49 GB
AI-TOD (Wang等,2021b)	共28036张航拍图像,包含8个类别、700621个实例。AI-TOD中目标的平均大小约为12.8像素,比其他数据集小得多。训练集、验证集和测试集的比例分别为2/5、1/10、1/2	HBB	23.5 GB
AI-TOD-v2 (Xu等,2022a)	针对AI-TOD数据集标签噪声问题,重新标注了AI-TOD数据集。AI-TOD-v2和AI-TOD图像相同,但注释不同。包含752745个实例	HBB	23.5 GB
SIMD (Haroon等,2020)	包括5000张大小为1024×768的遥感影像、45096个实例。总共包含15个目标类型	HBB	1.1 GB
FGSD (Chen等,2020)	包含2612幅遥感图像,其中标注船舶样本5634个。训练集共包含1917张图像、3964个船舶样本,验证集268张图像、590个样本,测试集427张图像、1080个样本	OBB	—
FAIR1M (Sun等,2022)	15266幅高分辨率遥感影像和超过100万个实例。包含5个类别和37个子类别,5个类别分别为:飞机、船舶、车辆、球场和道路。训练集、验证集和测试集的比例分别为1/2、1/6、1/3	OBB	45.4 GB
DIOR (Li等,2020a)	23463张遥感图像和192472个实例,覆盖20个常见的遥感目标类型。图像大小为800×800像素,空间分辨率为0.5—30 m	HBB	6.9 GB
DIOR-R (Cheng等,2022)	在原始的DIOR数据集上重新进行了旋转框标注,即图像数量不变,重新标注192512个旋转框实例	OBB	6.9 GB
RarePlanes (Shermeyer等,2021)	包括253个Maxar WorldView-3卫星场景,最大地面采样距离(GSD)为0.31—0.39 m,这些场景覆盖22个国家112个位置,总面积为2142 km ² 。共标注了14700架飞机。另外,在50000张合成卫星图像上标注了约630000架飞机	HBB	488.96 MB
MAR20 (禹文奇等,2022)	包括来自全球60个军用机场的3842张高分辨率遥感图像,20种飞机型号以及22341个目标示例。MAR20提供了水平边界框和有向边界框两种标注方式,分别对应于水平框与有向框目标识别任务	HBB/OBB	1.12 G
DOSR (Han等,2022b)	包含1066幅光学遥感影像和6127个船舶实例。图像大小为600至1300像素,分辨率为0.5—2.5 m。图像包括832个近岸场景(5212个实例)和234个离岸场景(915个实例)	OBB	841.7 M
HRPlanes (Bakirman和Sertel,2022)	包括3092张4800×2703大小的影像、18477架飞机。数据集以7:2:1的比例划分为训练集、验证集与测试集	HBB	—
CORS-ADD (Shi等,2023)	包含7337张图像,涵盖了典型的机场和各种罕见的场景。数据集由32285架民用和军用飞机实例组成,包括轰炸机,战斗机和预警机等。目标尺寸的范围分布在4×4到240×240像素	HBB/OBB	6.73 G
SODA-A (Cheng等,2023)	2513张图像,872069个实例。包括9个类别:飞机、直升机、小型车辆、大型车辆、船舶、集装箱、油罐、游泳池、风车。训练集、验证集和测试集的占比分别为0.4、0.25、0.35	OBB	10.2 GB
TinyPerson (Yu等,2020)	用于行人检测与小目标检测,包含1610个标记图像和759个未标记图像(两者主要来自同一视频集),共标注72651个极小的实例	HBB	1.47 GB

2.2 相关数据集分析

(1) 元数据分析。元数据是描述基本数据及其信息资源的描述性概念。针对光学遥感图像目标检测数据集而言,元数据描述了遥感数据集的数据分布、数据来源等原始信息。本节从数据集数量分布、数据集地域分布、数据来源分布、数

据集规模分布4个角度对15年来遥感领域目标检测数据集的发展情况做简要分析。

图2(a)描述了近年来数据集随时间发展的发布情况。可以看到,数据集的发布数量大致呈现出增长的趋势。同时,2016年之前,数据集的发布数量整体偏少,2016年之后,数据集发布明

显增多。这种现象与深度学习的发展有着密不可分的关系。近年来,以数据为驱动的深度学习方法在图像分类、分割、检测与识别等领域得到了

广泛地应用。因而,对于数据的需求远远超过早期的传统算法。

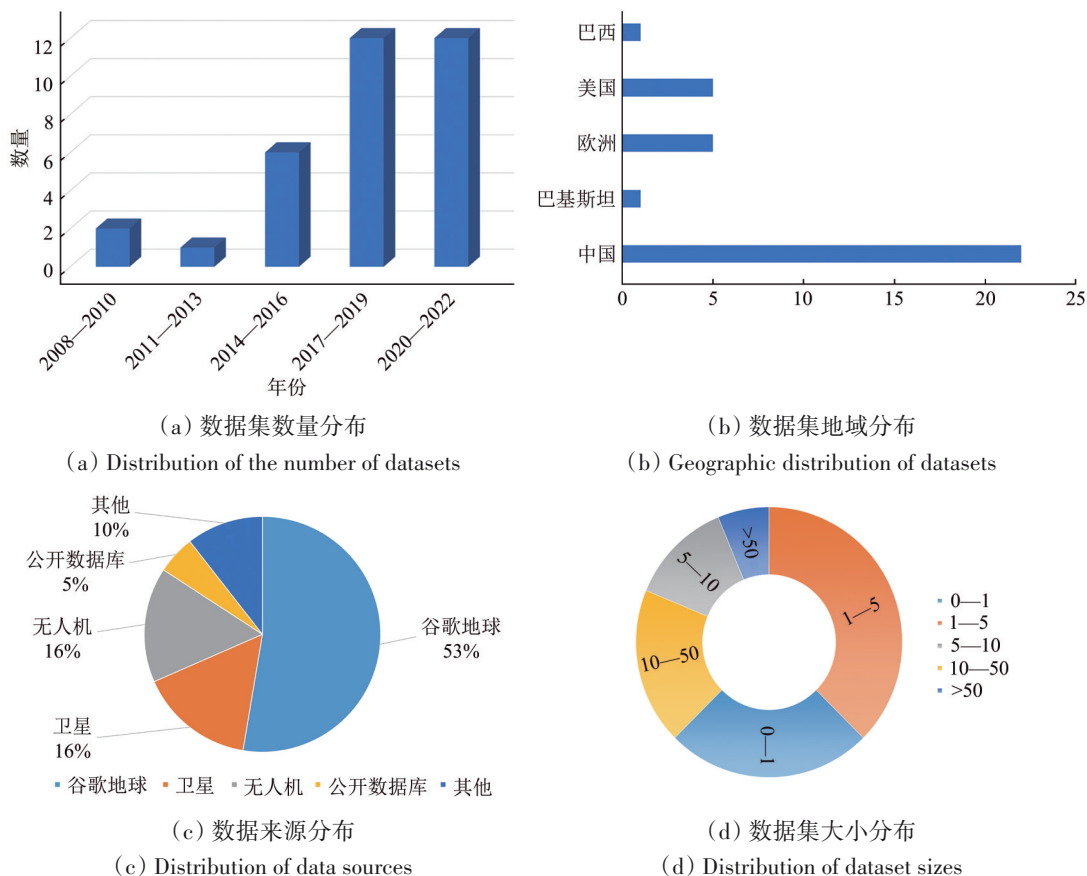


图2 数据集元数据统计

Fig. 2 Metadata statistics for datasets

数据集的地域分布情况如图2(b)所示,主要集中在中国、美国和欧洲,其中,中国数据集分布数量远远超过其他国家和地区。近年来,中国已成为推动遥感领域快速发展的主要力量。得益于得天独厚的人才与数据优势,以及相关政策的支持,以高校和科研院所为核心的学术机构纷纷发布公开数据集。例如,西北工业大学、武汉大学、中国科学院等相关实验室陆续发布了多个遥感目标检测数据集,满足了日益增长的数据需求,促进了相关科研成果的产出。

图2(c)展示了数据集的数据来源。其中,有53%的数据来源于谷歌地球,16%的数据来源于卫星,16%的数据由无人机拍摄得到。另外,还有少数数据来自公开的数据库。值得一提的是,随着国产卫星技术的快速迭代,遥感影像解译的不断成熟,国产系列自主卫星已成为一股不可忽

视的力量。近年来,国产卫星(如GF-2卫星, JL-1卫星等)提供了大量对地观测遥感数据,极大地推动了中国乃至世界遥感技术的发展与应用。

数据集规模的信息统计如图2(d),其中,1—5 GB大小的数据集发布数量最多。近年来,相关算法模型复杂度不断提升,对于数据量的需求也日益增长,一些超过50 GB的大型数据集逐渐出现。随着视觉大模型的诞生,数据量的需求也将更上一个台阶。

(2) 分辨率分析。不同的数据集,由于分辨率的不同,影像特点和差异显著。本节从空间分辨率和光谱分辨率两个角度出发,对相关数据集进行分析。

早期发布的数据集,其空间分辨率往往在0.5 m以上,如NWPU VHR-10(Cheng等,2014)的空间分辨率为0.5—2 m。近年,光学遥感影像解

析技术逐渐成熟, 其空间分辨率也在不断攀升。例如 xView (Lam 等, 2018) 数据集的空间分辨率为 0.3 m, LEVIR (Zou 和 Shi, 2018) 数据集的最小分辨率为 0.2 m。同时, 由于空间分辨率的提升, 目标检测的实例大小也呈现逐渐减小的趋势, 光学遥感图像小目标检测目前已成为遥感领域的一个研究热点与难点。近年发布的 AI-TOD (Wang 等, 2021b)、SODA-A (Cheng 等, 2023) 等数据集均聚焦于光学遥感小目标检测。

目前大多数的数据集图像格式采用 jpg 或 png 格式, 仅包含红绿蓝 3 个波段的光谱信息。但是, 随着算法性能及算力水平的提升, 近几年大幅面多光谱数据集成为一个新兴的趋势。多光谱数据除了常规的红绿蓝波段, 还会增加红外波段等其他波段信息, 如 RarePlanes (Shermeyer 等, 2021)、xView (Lam 等, 2018) 等均包含一部分多光谱图像。

(3) 基本信息分析。本节对数据集的基本信息做简要分析, 数据集的类别数量、图像数量、实例数量及图像宽度信息如表 2 所示。从类别数量角度来说, 大多数数据集只适用于特定领域, 甚至出现了一些单一类别的数据集, 如船舶检测数据集 HRSC2016 (Liu 等, 2017)、汽车检测数据集 COWC (Mundhenk 等, 2016) 等。同时, 也有部分数据集囊括了遥感领域常见的类别, 适合通用遥感图像目标检测, 如 DIOR (Li 等, 2020a)、DOTA (Xia 等, 2018) 等。不同数据集的图像数量与图像大小差异显著。早期受限于算法与算力的限制, 数据集中的图像数量与大小均较小, 而近几年则出现了一些规模较大的数据集。数据集的实例数量也呈现出不断增加的趋势, 例如 DOTA-v2 (Ding 等, 2022)、FAIR1M (Sun 等, 2022) 的实例数量均达到百万级。

表 2 数据集基本信息

Table 2 Basic information of datasets

数据集名称	类别数量	图像数量	实例数量	图像宽度
TAS (Heitz 和 Koller, 2008)	1	30	1319	792
SZTAKI-INRIA (Benedek 等, 2012)	1	9	665	~800
OIRDS (Tanner 等, 2009)	4	900	1800	—
NWPU VHR-10 (Cheng 等, 2014)	10	800	3775	~1000
UCAS-AOD (Zhu 等, 2015)	2	910	6029	1280
VEDAI (Razakarivony 和 Jurie, 2016)	9	1210	3640	512, 1024
DLR 3K Vehicle (Liu 和 Mattyus, 2015)	2	20	14235	5616
HRSC2016 (Liu 等, 2017)	1	1061	2976	300—1500
COWC (Mundhenk 等, 2016)	1	53	32716	2000
CARPK (Hsieh 等, 2017)	1	1448	89777	1280
RSOD (Xiao 等, 2015)	4	976	6950	~1000
VisDrone (Zhu 等, 2022)	10	10209	~542 K	~1490
xView (Lam 等, 2018)	60	1129	~1 M	2000—4000
LEVIR (Zou 和 Shi, 2018)	3	21952	11028	800
ITCVD (Yang 等, 2018)	1	173	29088	5616
DOTA (Xia 等, 2018)	15	2806	188282	800—13000
DOTA-v1.5	16	2806	402089	800—13000
DOTA-v2.0 (Ding 等, 2022)	18	11268	1793658	800—20000
RSD-GOD (Zhuang 等, 2019)	5	18187	40990	300—600
HRRSD (Zhang 等, 2019)	13	26722	55740	152—10569
Bridges Dataset (Nogueira 等, 2019)	1	500	—	4800
AI-TOD (Wang 等, 2021b)	8	28036	700621	800
AI-TOD-v2 (Xu 等, 2022a)	8	28036	752745	800

续表

数据集名称	类别数量	图像数量	实例数量	图像宽度
SIMD (Haroon 等, 2020)	15	5000	45096	1024
FGSD (Chen 等, 2020)	1/4/43	2612	5634	930
FAIR1M (Sun 等, 2022)	5/37	15266	1.02M	1000—10000
DIOR (Li 等, 2020a)	20	23463	192472	800
DIOR-R (Cheng 等, 2022)	20	23463	192512	800
RarePlanes (Shermeyer 等, 2021)	1—110	253	14707	512
MAR20 (禹文奇 等, 2022)	20	3824	22341	~800
DOSR (Han 等, 2022b)	20	1066	6172	600—1300
HRPlanes (Bakirman 和 Sertel, 2022)	1	3092	18477	4800
CORS-ADD (Shi 等, 2023)	1	7737	32285	640
SODA-A (Cheng 等, 2023)	9	2513	872069	~4761
TinyPerson (Yu 等, 2020)	1	2369	72651	—

综上所述,近年来,随着深度学习等技术在遥感领域的不断推广,对数据集的需求规模与数据质量要求不断提高。另一方面,新型遥感数据集的不断发布也为光学遥感图像检测任务注入新的活力并促进了相关算法的不断优化。从长远来看,光学遥感图像目标检测数据集必然朝着高质量、大规模、多类别的方向发展;同时,光学遥感图像小目标检测、细粒度检测等研究痛点也必将催生出更多具备倾向性的数据集。

3 数据集相关算法分析

上述公开的数据集已经被广泛应用于以深度学习为基础的光学遥感图像目标检测领域,并为相关算法的发展奠定了基础。本节将对依托数据集推动的光学遥感图像目标检测算法进行归纳和分析。

3.1 水平框目标检测

本节介绍基于深度学习的水平框目标检测算法在遥感领域的应用。深度神经网络最早在自然图像领域得到应用,因其性能远超当时的传统算法,很快被迁移到遥感影像解析等其他领域。伴随着深度学习在遥感领域的应用,UCAS-AOD (Zhu 等, 2015)、LEVIR (Zou 和 Shi, 2018)、DIOR (Li 等, 2020a) 等数据集相继公开,进一步促进了光学遥感目标检测算法的发展。以下,将根据目标检测器是否依赖锚框,从两个方面介绍算法的发展情况。

R-CNN (Girshick 等, 2014) 算法最早在遥感目标检测领域得到应用。例如, Cheng 等 (2016)

针对遥感目标的旋转不变特性,在R-CNN的基础上提出旋转不变模型(RICNN),用于多类别地理空间目标检测。RICNN在现有CNN架构的基础上引入旋转不变层,同时施加正则化约束来优化目标函数,使得不同旋转角度的样本具有相似的特征,从而实现旋转不变性。Long等(2017)指出非极大值抑制算法(NMS)保留的边界框难以进行精准定位,进而提出一种基于得分的无监督边界框回归方法(USB-BBR),对NMS保留后的边界框进一步回归优化,解决了遥感图像复杂背景信息下的精准定位问题。然而,以上算法均采用选择性搜索算法(Uijlings等,2013)进行提议框的生成,计算速度较慢,制约了相关算法的应用。为进一步提高检测性能,随后一些工作将Faster R-CNN (Ren等,2015)用于遥感目标检测。如Zhong等(2018)提出了一种基于位置敏感平衡(PSB)的遥感图像多类别地理空间目标检测框架。PSB框架采用位置敏感的池化层,以缓解分类阶段的不变性和目标检测阶段的不变性之间的矛盾。然而,在预测结果与真实结果位置偏差较大时,位置敏感策略很难起到作用。Shivappriya等(2021)提出AAF激活函数,利用傅里叶变换和线性组合来更新损失,使网络更好地收敛。另外,为了满足实时检测的需求,一些研究聚焦于结构更简单的一阶段网络。Li等(2022b)将一种基于元学习的方法引入YOLOv3 (Redmon和Farhadi,2018),用于遥感图像小样本目标检测。该方法通过元特征提取器从输入图像提取元特征图,同时

通过特征重加权模块学习重加权向量, 然后利用重加权向量对元特征图进行校准, 很好地改善了少量样本下模型的泛化能力。Ye等(2022)提出一种基于自适应注意力融合机制(AAFM)的一阶段网络。AAFM机制通过将最佳融合因子应用于通道注意力与空间注意力, 来获取最优融合信息表征。然而, 该网络需平衡多尺度目标的比例, 增加了图像预处理的复杂度。同时, AAFM机制会引入额外的可学习参数, 进一步增加了模型的计算成本。

锚框的设计不仅使检测流程变得复杂, 而且还需要预设合适的超参数。为此, 研究人员陆续提出一些适用于遥感场景的无锚框算法。Li等(2020b)通过逐像素密集预测的方式进行目标检测, 从而避免了锚框的设计。同时, 考虑到目标的多尺度差异, 提出密集路径聚合特征金字塔网络(DPAFPN), 以更好地利用多尺度信息, 并在一定程度上较少了特征传递过程中的信息丢失问题。针对场景信息缺失等问题, Liu等(2022)提出一种场景相关的无锚网络(SRAF-Net)。SRAF-Net首先通过设计场景增强特征金字塔网络(SE-FPN)获取目标的场景上下文特征, 然后通过场景辅助检测头(SADH)进行更准确的检测。但是, 该方法仅考虑了背景信息对目标检测的辅助作用, 而未考虑不同前景类别之间的上下文信息。Fu等(2023)提出了Xnet, 以解决正负样本不均衡的问题。该模型设计了一种可变形卷积, 用于动态分析感受野的边界和偏移, 使模型更适应目标的尺度变化。

3.2 旋转框目标检测

由于特殊成像视角, 遥感图像中的目标通常是以任意方向出现, 单纯移植水平框检测算法无法对检测目标的方向进行描述, 且生成的检测框中可能存在冗余的背景信息。此外, 非极大值抑制方法的应用导致极易丢失有效的水平框(Wang等, 2023)。针对上述问题, 相关旋转框目标检测算法被提出。

DOTA(Xia等, 2018)、DOTA-v2.0(Ding等, 2022)、DIOR-R(Cheng等, 2022)等是光学遥感图像旋转框目标检测算法中最为常用的数据集, 为众多算法的发展奠定了基准。同样, 旋转框检测算法依据是否基于锚框也可分为两大类: 基于

锚框的算法和无锚框算法(Wang等, 2023)。

基于锚框的两阶段检测算法最早可追溯到Faster RCNN OBB算法(Xia等, 2018), 它在Faster RCNN(Ren等, 2015)算法的基础上, 将RPN网络的标签修正为旋转框的最大外接矩形框, 并通过在检测头部分增加一个额外的输出通道来回归每个检测框的方向。该算法的缺点是RPN部分生成的水平候选框无法紧密包围目标导致映射的层级特征携带了过多的背景信息限制了最终的检测性能。RRPN(Ma等, 2018)对RPN进行修正, 预设了多个旋转框, 以适应不同方向的目标, 进而生成旋转区域候选框, 并设计旋转RoI池化层来提取特征, 有效解决了特征不对齐问题。该算法的缺点是锚框预设方向有限且带来的计算和存储消耗均很大。RoI Transformer(Ding等, 2019)设计了一个RoI学习器来学习生成旋转框, 并使用旋转位置敏感的RoI Align进一步提取旋转不变特征, 一定程度上缓解了遥感图像对比度低、信息丢失严重的情况, 有效提升了效率和精度, 但其对于大长宽比目标的检测效果不佳。Redet(Han等, 2021)通过设计旋转等变网络和旋转不变RoI Align来提取目标的旋转不变特征, 大幅降低了对方位信息进行编码的冗余参数量, 并在网络设计中切实建模了方位变换, 由此大大降低模型尺寸, 但受限于其架构设计, 检测速度仍待提升。Oriented RCNN(Xie等, 2021)提出了一种轻量化的架构Oriented RPN来生成高质量的旋转候选框区域, 实现了精度和速度的高效平衡。基于锚框的一阶段检测算法较早提出的是R3Det(Yang等, 2021a), 它在RetinaNet(Lin等, 2017)的基础上使用了一个特征细化模块(FRM), 根据前一阶段的预测边界框信息对响应的特征点进行重新编码, 以对齐和重构特征图, 此外还提出了SkewIoU损失使训练更加稳定、定位更加准确, 但其检测速度慢且精度提升也较为有限。S2A-Net(Han等, 2022a)通过设计一个特征对齐模块(FAM)和旋转检测模块(ODM), 实现特征对齐的同时, 缓解了回归和分类的不一致性。GWD(Yang等, 2021b)提出了一种基于高斯Wasserstein距离的损失来建模两个旋转框之间的偏差, 有效弥补了边界不连续性及其与最终检测度量的不一致性问题, 但由于涉及人为指定的分布距离度量, 需要额外超参数调整。KFIOU(Yang等, 2023)提出了一

种称之为 KFIoU 的损失, 通过采用高斯分布的乘积来模拟 SkewIoU 损失的计算机制, 由于其完全的可微性和处理非重叠情况的能力, 故对基于梯度的训练十分友好且效果更佳。

基于锚框的检测方法需要预设锚框, 不仅过程繁琐且对一些特殊尺寸的目标很难有效匹配, 为此, 无锚框检测算法被提出来。无锚框算法又可分基于中心域的算法与基于关键点的算法 (Wang 等, 2023)。基于中心域的检测方法较为典型的是 IENet (Lin 等, 2021), 它在 FCOS (Tian 等, 2019) 基础上添加了一个角度回归分支, 并在其上设计了分支交互模块, 将来自分类分支和边框回归分支的特征利用自注意力模块进行融合, 组合到角度回归分支中, 生成最终特征以进行角度回归, 在降低计算复杂度的同时提升了存储效率, 但仍需要过滤大量冗余框且其检测精度依旧有较大改善空间。CenterMap-Net (Wang 等, 2021c) 通过预测中心概率图来消除目标和背景间的歧义, 并以此生成旋转框, 很大程度上消除了目标定义和背景像素上的模糊性, 弱化了背景像素对检测性能的影响, 更好地达成了从杂乱背景中区分出感兴趣的对象。基于关键点的算法先检测目标的关键点, 再对目标的空间范围进行限制。O²-Dnet (Wei 等, 2020) 通过从中心点回归偏移量来定位旋转框四边的中点, 以确定检测目标的尺寸及方向, 既不需要预设锚框也不需要过滤冗余框, 大大降低了模型的计算复杂性。CenterRot (Wang 等, 2021a) 在 CenterNet 的基础上, 提出了一种可变形特征金字塔网络并在检测头部分增加角度分支, 简化了模型的设计, 同时也减少了与锚相关的计算, 使得针对尺度变化剧烈的对象保留了鲁棒的特征提取能力, 并准确预测了旋转对象的方向, 提升了检测精度。Oriented RepPoints (Li 等, 2022a) 提出了一种自适应点集评估和分配策略来预测旋转目标, 并提出了 3 种转换方式将代表点转换为旋转框, 不仅能够从相邻对象及背景噪声中剥离相关特征, 实现了具有精确方位的精确检测, 而且捕获了任意方位实例的底层几何结构, 以便于准确定向的分类和定位。

3.3 其他目标检测方向

HRSC2016 (Liu 等, 2017)、xView (Lam 等, 2018)、FGSD (Chen 等, 2020)、FAIR1M (Sun

等, 2022) 等是光学遥感图像细粒度目标检测算法中最为常用的数据集, 为众多细粒度检测算法的发展提供了数据基础。Han 等 (2022a) 提出一种高效信息复用网络, 利用密集融合子网络最大限度地提取多层信息, 利用双层掩码注意力模块抑制复杂背景噪声的同时突出目标信息。Devaki 等 (2023) 等设计了一系列用于特征增强的模块, 包括分组增强模块、峰值抑制模块和随机抑制模块, 使得模型可以学习到更具有判别性的特征, 从而提高细粒度目标检测能力。Song 等 (2023) 提出精细平衡金字塔网络来解决特征金字塔间的语义信息稀疏问题, 并设计中心尺度感知标签分配策略来充分利用旋转目标的统计特性。

VisDrone (Zhu 等, 2022)、AI-TOD (Wang 等, 2021b)、AI-TOD-v2 (Xu 等, 2022a)、SODA-A (Cheng 等, 2023) 等是光学遥感图像小目标检测算法中最为常用的数据集, 有力推动了小目标检测任务的进展。Xu 等 (2021) 采用提出的 DotD 度量方式来取代常用的 IoU, 以更好地实现标签分配。RFLA (Xu 等, 2022b) 通过在标签分配中度量每个特征点的高斯接受域与 ground truth 之间的相似度, 以提高对微小目标的检测性能。DCFL (Xu 等, 2023) 一方面通过一种动态的方式对先验、标签分配和目标表征进行建模, 另一方面利用粗糙的先验匹配和更为精细的后验约束来动态分配标签, 实现了小目标检测性能的大幅提升。

3.4 遥感大模型

随着通用大模型的不断深入研究, 人工智能领域取得了突破性进展, 同时, 也极大地影响了遥感领域的走向。近来的一些研究表明, 自监督、多模态逐渐成为遥感视觉基础模型的发展趋势。例如, Cha 等 (2023) 指出遥感领域基础模型的研究主要集中在预训练方法、数据集规模与模型参数的数量 3 个方面, 并就模型参数数量对基础模型在旋转目标检测、语义分割等下游任务的性能表现进行了深入的研究。通过不同参数数量的 MAE (He 等, 2022) 和大规模遥感数据集 MillionAID (Long 等, 2021) 证明基础模型的性能随着模型参数的数量增加而提高, 进而, 提出了遥感领域第一个十亿级参数规模的基础模型。Sun 等 (2023) 基于生成式自监督学习, 构建出遥感基础模型框架 RingMo。该框架构建了一个拥有两百万张遥感

影像的大规模数据集,同时,设计了一种遥感基础模型训练方法,旨在解决复杂遥感场景下的密集小目标检测问题。通过在8个数据集、4个下游任务(场景分类、目标检测、语义分割、变化检测)上的实验,证明了RingMo框架的有效性。(Liu等,2023)针对遥感基础模型缺乏语言理解的问题,提出了第一个用于遥感的视觉语言基础模型RemoteCLIP。该模型可以同时学习丰富的语义信息和对齐的文本映射,实现了鲁棒的视觉特征表示。

4 结 论

高质量的数据是基于深度学习的光学遥感图像目标检测任务发展的基础,也是相关算法研究的前置条件和必要储备。随着深度学习等人工智能技术不断在遥感领域应用和推广,遥感数据的重要性日益凸显,新型光学遥感图像目标检测数据集不断发布,本文旨在对已发布的数据集进行梳理归纳与总结,并对算法的应用进行概述,为相关领域科研人员的后续研究提供借鉴。首先,本文将2008年—2023年已发布的光学遥感图像目标检测数据集依据标注方式进行划分,对其中的11个典型数据集进行了全面阐述,以表格的形式对所有的数据集信息进行了全面归纳总结,并从元数据、分辨率、基本信息3个角度对数据集进行了全面对比分析,论证了光学遥感图像目标检测数据集必然朝着高质量、大规模、多类别的方向发展。此外,针对已发布的数据集,从水平框目标检测、旋转框目标检测以及细分检测方向(小目标检测和细粒度检测)等多个角度对相关算法发展应用进行了概述,证实了遥感数据对算法的推动作用。综上,本文将为后续基于深度学习的光学遥感图像目标检测算法的发展提供数据与研究思路的参考。

参考文献(References)

Bakirman T and Sertel E. 2022. HRPlanes: high resolution airplane dataset for deep learning. arXiv preprint arXiv:2204.10959

Benedek C, Descombes X and Zerubia J. 2012. Building development monitoring in multitemporal remotely sensed image pairs with stochastic birth-death dynamics. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 34(1): 33-50 [DOI: 10.1109/TPAMI.2011.94]

Cha K, Seo J and Lee T. 2023. A billion-scale foundation model for remote sensing images. arXiv:2304.05215

Chen K Y, Wu M, Liu J M and Zhang C. 2020. FGSD: a dataset for fine-grained ship detection in high resolution satellite images. arXiv:2003.06832

Cheng G and Han J W. 2016a. A survey on object detection in optical remote sensing images. *ISPRS Journal of Photogrammetry and Remote Sensing*, 117: 11-28 [DOI: 10.1016/j.isprsjprs.2016.03.014]

Cheng G, Han J W, Zhou P C and Guo L. 2014. Multi-class geospatial object detection and geographic image classification based on collection of part detectors. *ISPRS Journal of Photogrammetry and Remote Sensing*, 98: 119-132 [DOI: 10.1016/j.isprsjprs.2014.10.002]

Cheng G, Wang J B, Li k, Xie X X, Lang C B, Yao Y Q and Han J W. 2022. Anchor-free oriented proposal generator for object detection. *IEEE Transactions on Geoscience and Remote Sensing*, 60: 5625411 [DOI: 10.1109/TGRS.2022.3183022]

Cheng G, Yuan X, Yao X W, Yan K B, Zeng Q H, Xie X X and Han J W. 2023. Towards large-scale small object detection: survey and benchmarks. arXiv preprint arXiv:2207.14096

Cheng G, Zhou P C and Han J W. 2016b. Learning rotation-invariant convolutional neural networks for object detection in VHR optical remote sensing images. *IEEE Transactions on Geoscience and Remote Sensing*, 54(12): 7405-7415 [DOI: 10.1109/TGRS.2016.2601622]

Devaki P, Vineetha P N, Reddy C H, Bharathi P, Karimulla S and Kumar S U. 2023. Fine-grained feature enhancement for object detection in remote sensing images. *International Research Journal of Modernization in Engineering Technology and Science*, 5(3): 2112-2118 [DOI: 10.56726/IRJMETS34594]

Ding J, Xue N, Long Y, Xia G S and Lu Q K. 2019. Learning roi transformer for oriented object detection in aerial images//*Proceedings of the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Long Beach: IEEE: 2844-2853 [DOI: 10.1109/cvpr.2019.00296]

Ding J, Xue N, Xia G S, Bai X, Yang W, Yang M Y, Belongie S, Luo J B, Datcu M, Pelillo M and Zhang L P. 2022. Object detection in aerial images: a large-scale benchmark and challenges. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 44(11): 7778-7796 [DOI: 10.1109/TPAMI.2021.3117983]

Fu S S, He Y F, Du X F and Zhu Y. 2023. Anchor-free object detection in remote sensing images using a variable receptive field network. *EURASIP Journal on Advances in Signal Processing*, 2023(1): 53 [DOI: 10.1186/s13634-023-01013-2]

Girshick R, Donahue J, Darrell T and Malik J. 2014. Rich feature hierarchies for accurate object detection and semantic segmentation//*Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition*. Columbus: IEEE: 580-587 [DOI: 10.1109/CVPR.2014.81]

Han J M, Ding J, Li J and Xia G S. 2022a. Align deep features for oriented object detection. *IEEE Transactions on Geoscience and Re-*

- remote Sensing, 60: 5602511 [DOI: 10.1109/TGRS.2021.3062048]
- Han J M, Ding J, Xue N and Xia G S. 2021. ReDet: a rotation-equivariant detector for aerial object detection//Proceedings of the 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Nashville: IEEE: 2785-2794 [DOI: 10.1109/cvpr46437.2021.00281]
- Han Y Q, Yang X Y, Pu T and Peng Z M. 2022b. Fine-grained recognition for oriented ship against complex scenes in optical remote sensing images. *IEEE Transactions on Geoscience and Remote Sensing*, 60: 5612318 [DOI: 10.1109/TGRS.2021.3123666]
- Haroon M, Shahzad M and Fraz M M. 2020. Multisized object detection using spaceborne optical imagery. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 13: 3032-3046 [DOI: 10.1109/jstars.2020.3000317]
- He K M, Chen X L, Xie S N, Li Y H, Dollár P and Girshick R. 2022. Masked autoencoders are scalable vision learners//Proceedings of the 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition. New Orleans: IEEE: 15979-15988 [DOI: 10.1109/cvpr52688.2022.01553]
- Heitz G and Koller D. 2008. Learning spatial context: using stuff to find things//10th European Conference on Computer Vision. Marseille, France: Springer: 30-43 [DOI: 10.1007/978-3-540-88682-2_4]
- Hsieh M R, Lin Y L and Hsu W H. 2017. Drone-based object counting by spatially regularized regional proposal network//Proceedings of the 2017 IEEE International Conference on Computer Vision. Venice: IEEE: 4165-4173 [DOI: 10.1109/iccv.2017.446]
- Lam D, Kuzma R, McGee K, Dooley S, Laielli M, Klaric M, Bulatov Y and McCord B. 2018. xView: objects in context in overhead imagery. arXiv preprint arXiv:1802.07856
- Li K, Wan G, Cheng G, Meng L Q and Han J W. 2020a. Object detection in optical remote sensing images: a survey and a new benchmark. *ISPRS Journal of Photogrammetry and Remote Sensing*, 159: 296-307 [DOI: 10.1016/j.isprsjprs.2019.11.023]
- Li W T, Chen Y J, Hu K X and Zhu J K. 2022a. Oriented RepPoints for aerial object detection//Proceedings of the 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition. New Orleans: IEEE: 1819-1828 [DOI: 10.1109/cvpr52688.2022.00187]
- Li X, Deng J Y and Fang Y. 2022b. Few-shot object detection on remote sensing images. *IEEE Transactions on Geoscience and Remote Sensing*, 60: 5601614 [DOI: 10.1109/TGRS.2021.3051383]
- Li Y Y, Pei X, Huang Q, Jiao L C, Shang R H and Marturi N. 2020b. Anchor-free single stage detector in remote sensing images based on multiscale dense path aggregation feature pyramid network. *IEEE Access*, 8: 63121-63133 [DOI: 10.1109/ACCESS.2020.2984310]
- Lin T Y, Goyal P, Girshick R, He K M and Dollár P. 2017. Focal loss for dense object detection//Proceedings of the 2017 IEEE International Conference on Computer Vision. Venice: IEEE: 2999-3007 [DOI: 10.1109/iccv.2017.324]
- Lin Y T, Feng P M, Guan J, Wang W W and Chambers J. 2021. IENet: interacting embranchment one stage anchor free detector for orientation aerial object detection. arXiv:1912.00969
- Liu F, Chen D L, Guan Z Q Y, Zhou X C, Zhu J L and Zhou J. 2023. RemoteCLIP: a vision language foundation model for remote sensing. arXiv:2306.11029
- Liu J M, Li S J, Zhou C S, Cao X Y, Gao Y and Wang B. 2022. SRAF-Net: a scene-relevant anchor-free object detection network in remote sensing images. *IEEE Transactions on Geoscience and Remote Sensing*, 60: 5405914 [DOI: 10.1109/TGRS.2021.3124959]
- Liu K and Mattyus G. 2015. Fast multiclass vehicle detection on aerial images. *IEEE Geoscience and Remote Sensing Letters*, 12(9): 1938-1942 [DOI: 10.1109/LGRS.2015.2439517]
- Liu Z K, Yuan L, Weng L B and Yang Y P. 2017. A high resolution optical satellite image dataset for ship recognition and some new baselines//Proceedings of the 6th International Conference on Pattern Recognition Applications and Methods. Porto: SciTePress: 324-331 [DOI: 10.5220/0006120603240331]
- Long Y, Gong Y P, Xiao Z F and Liu Q. 2017. Accurate object localization in remote sensing images based on convolutional neural networks. *IEEE Transactions on Geoscience and Remote Sensing*, 55(5): 2486-2498 [DOI: 10.1109/TGRS.2016.2645610]
- Long Y, Xia G S, Li S Y, Yang W, Yang M Y, Zhu X X, Zhang L P and Li D R. 2021. On creating benchmark dataset for aerial image interpretation: reviews, guidances, and million-aid. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 14: 4205-4230 [DOI: 10.1109/JSTARS.2021.3070368]
- Ma J Q, Shao W Y, Ye H, Wang L, Wang H, Zheng Y B and Xue X Y. 2018. Arbitrary-oriented scene text detection via rotation proposals. *IEEE Transactions on Multimedia*, 20(11): 3111-3122 [DOI: 10.1109/TMM.2018.2818020]
- Mundhenk T N, Konjevod G, Sakla W A and Boakye K. 2016. A large contextual dataset for classification, detection and counting of cars with deep learning//14th European Conference on Computer Vision. Amsterdam, The Netherlands: Springer: 785-800 [DOI: 10.1007/978-3-319-46487-9_48]
- Nie G T and Huang H. 2021. A survey of object detection in optical remote sensing images. *Acta Automatica Sinica*, 47(8): 1749-1768 (聂光涛, 黄华. 2021. 光学遥感图像目标检测算法综述. *自动化学报*, 47(8): 1749-1768) [DOI: 10.16383/j.aas.c200596]
- Nogueira K, Cesar C, Gama P H T, Machado G L S and dos Santos J A. 2019. A tool for bridge detection in major infrastructure works using satellite images//2019 XV Workshop de Visão Computacional (WVC). São Bernardo do Campo: IEEE: 72-77 [DOI: 10.1109/WVC.2019.8876942]
- Razakarivony S and Jurie F. 2016. Vehicle detection in aerial imagery: a small target detection benchmark. *Journal of Visual Communication and Image Representation*, 34: 187-203 [DOI: 10.1016/j.jvcir.2015.11.002]
- Redmon J and Farhadi A. 2018. YOLOv3: an incremental improvement. arXiv:1804.02767

- Ren S Q, He K M, Girshick R and Sun J. 2015. Faster R-CNN: towards real-time object detection with region proposal networks// Proceedings of the 28th International Conference on Neural Information Processing Systems. Montreal: MIT Press: 91-99
- Shermeyer J, Hossler T, Van Etten A, Hogan D, Lewis R and Kim D. 2021. RarePlanes: synthetic data takes flight// Proceedings of the 2021 IEEE Winter Conference on Applications of Computer Vision. Waikoloa: IEEE: 207-217 [DOI: 10.1109/wacv48630.2021.00025]
- Shi T J, Gong J N, Jiang S K, Zhi X Y, Bao G Z, Sun Y and Zhang W. 2023. Complex optical remote-sensing aircraft detection dataset and benchmark. IEEE Transactions on Geoscience and Remote Sensing, 61: 5612309 [DOI: 10.1109/TGRS.2023.3283137]
- Shivappriya S N, Priyadarsini M J P, Stateczny A, Puttamadappa C and Parameshachari B D. 2021. Cascade object detection and remote sensing object detection method based on trainable activation function. Remote Sensing, 13(2): 200 [DOI: 10.3390/rs13020200]
- Song J J, Miao L J, Ming Q, Zhou Z Q and Dong Y P. 2023. Fine-grained object detection in remote sensing images via adaptive label assignment and refined-balanced feature pyramid network. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 16: 71-82 [DOI: 10.1109/JSTARS.2022.3224558]
- Sun X, Wang P J, Lu W X, Zhu Z C, Lu X N, He Q B, Li J X, Rong X E, Yang Z J, Chang H, He Q L, Yang G, Wang R P, Lu J W and Fu K. 2023. RingMo: a remote sensing foundation model with masked image modeling. IEEE Transactions on Geoscience and Remote Sensing, 61: 5612822 [DOI: 10.1109/TGRS.2022.3194732]
- Sun X, Wang P J, Yan Z Y, Xu F, Wang R P, Diao W H, Chen J, Li J H, Feng Y C, Xu T, Weinmann M, Hinz S, Wang C and Fu K. 2022. FAIR1M: a benchmark dataset for fine-grained object recognition in high-resolution remote sensing imagery. ISPRS Journal of Photogrammetry and Remote Sensing, 184: 116-130 [DOI: 10.1016/j.isprs.2021.12.004]
- Tanner F, Colder B, Pullen C, Heagy D, Eppolito M, Carlan V, Oertel C and Sallee P. 2009. Overhead imagery research data set—An annotated data library and tools to aid in the development of computer vision algorithms// 2009 IEEE Applied Imagery Pattern Recognition Workshop (AIPR 2009). Washington: IEEE: 1-8 [DOI: 10.1109/AIPR.2009.5466304]
- Tian Z, Shen C H, Chen H and He T. 2019. FCOS: fully convolutional one-stage object detection// Proceedings of the 2019 IEEE/CVF International Conference on Computer Vision. Seoul: IEEE: 9626-9635 [DOI: 10.1109/iccv.2019.00972]
- Uijlings J R R, Van De Sande K E A, Gevers T and Smeulders A W M. 2013. Selective search for object recognition. International Journal of Computer Vision, 104(2): 154-171 [DOI: 10.1007/s11263-013-0620-5]
- Wang J, Yang L and Li F. 2021a. Predicting arbitrary-oriented objects as points in remote sensing images. Remote Sensing, 13(18): 3731 [DOI: 10.3390/rs13183731]
- Wang J W, Yang W, Guo H W, Zhang R X and Xia G S. 2021b. Tiny object detection in aerial images// 2020 25th International Conference on Pattern Recognition (ICPR). Milan: IEEE: 3791-3798 [DOI: 10.1109/ICPR48806.2021.9413340]
- Wang J W, Yang W, Li H C, Zhang H J and Xia G S. 2021c. Learning center probability map for detecting objects in aerial images. IEEE Transactions on Geoscience and Remote Sensing, 59(5): 4307-4323 [DOI: 10.1109/TGRS.2020.3010051]
- Wang K, Wang Z, Li Z, Su A, Teng X C, Liu M H and Yu Q F. 2023. Oriented object detection in optical remote sensing images using deep learning: a survey. arXiv preprint arXiv:2302.10473
- Wei H R, Zhang Y, Chang Z H, Li H, Wang H Q and Sun X. 2020. Oriented objects as pairs of middle lines. ISPRS Journal of Photogrammetry and Remote Sensing, 169: 268-279 [DOI: 10.1016/j.isprs.2020.09.022]
- Xia G S, Bai X, Ding J, Zhu Z, Belongie S, Luo J B, Datcu M, Pelillo M and Zhang L P. 2018. DOTA: a large-scale dataset for object detection in aerial images// Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City: IEEE: 3974-3983 [DOI: 10.1109/cvpr.2018.00418]
- Xiao Z F, Liu Q, Tang G F and Zhai X F. 2015. Elliptic Fourier transformation-based histograms of oriented gradients for rotationally invariant object detection in remote-sensing images. International Journal of Remote Sensing, 36(2): 618-644 [DOI: 10.1080/01431161.2014.999881]
- Xie X X, Cheng G, Wang J B, Yao X W and Han J W. 2021. Oriented R-CNN for object detection// Proceedings of the 2021 IEEE/CVF International Conference on Computer Vision. Montreal: IEEE: 3500-3509 [DOI: 10.1109/iccv48922.2021.00350]
- Xu C, Ding J, Wang J W, Yang W, Yu H, Yu L and Xia G S. 2023. Dynamic coarse-to-fine learning for oriented tiny object detection// Proceedings of the 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Vancouver: IEEE: 7318-7328 [DOI: 10.1109/CVPR52729.2023.00707]
- Xu C, Wang J W, Yang W, Yu H, Yu L and Xia G S. 2022a. Detecting tiny objects in aerial images: a normalized Wasserstein distance and a new benchmark. ISPRS Journal of Photogrammetry and Remote Sensing, 190: 79-93 [DOI: 10.1016/j.isprs.2022.06.002]
- Xu C, Wang J W, Yang W, Yu H, Yu L and Xia G S. 2022b. RFLA: gaussian receptive field based label assignment for tiny object detection// 17th European Conference on Computer Vision. Tel Aviv: Springer: 526-543 [DOI: 10.1007/978-3-031-20077-9_31]
- Xu C, Wang J W, Yang W and Yu L. 2021. Dot distance for tiny object detection in aerial images// Proceedings of the 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops. Nashville: IEEE: 1192-1201 [DOI: 10.1109/cvprw53098.2021.00130]
- Yang M Y, Liao W T, Li X B and Rosenhahn B. 2018. Deep learning for vehicle detection in aerial images// 2018 25th IEEE International Conference on Image Processing (ICIP). Athens: IEEE: 3079-3083 [DOI: 10.1109/ICIP.2018.8451454]

- Yang X, Yan J C, Feng Z M and He T. 2021a. R3Det: refined single-stage detector with feature refinement for rotating object//Proceedings of the 35th AAAI Conference on Artificial Intelligence. Palo Alto: AAAI Press: 3163-3171 [DOI: 10.1609/aaai.v35i4.16426]
- Yang X, Yan J C, Ming Q, Wang W T, Zhang X P and Tian Q. 2021b. Rethinking rotated object detection with gaussian wasserstein distance loss//Proceedings of the 38th International Conference on Machine Learning. Virtual Event: [s.n.]: 11830-11841
- Yang X, Zhou Y, Zhang G F, Yang J R, Wang W T, Yan J C, Zhang X P and Tian Q. 2023. The KFIOU loss for rotated object detection. arXiv:2201.12558
- Yao Y Q, Cheng G, Xie X X and Han J W. 2021. Optical remote sensing image object detection based on multi-resolution feature fusion. National Remote Sensing Bulletin, 25(5): 1124-1137 (姚艳清, 程焱, 谢星星, 韩军伟. 2021. 多分辨率特征融合的光学遥感图像目标检测. 遥感学报, 25(5): 1124-1137) [DOI: 10.11834/jrs.20210505]
- Ye Y X, Ren X Y, Zhu B, Tang T F, Tan X, Gui Y and Yao Q. 2022. An adaptive attention fusion mechanism convolutional network for object detection in remote sensing images. Remote Sensing, 14(3): 516 [DOI: 10.3390/rs14030516]
- Yu W Q, Cheng G, Wang M J, Yao Y Q, Xie X X, Yao X W and Han J W. 2022. MAR20: a benchmark for military aircraft recognition in remote sensing images. National Remote Sensing Bulletin: 1-11 (禹文奇, 程焱, 王美君, 姚艳清, 谢星星, 姚西文, 韩军伟. 2022. MAR20: 遥感图像军用飞机目标识别数据集. 遥感学报: 1-11) [DOI: 10.11834/jrs.20222139]
- Yu X H, Gong Y Q, Jiang N, Ye Q X and Han Z J. 2020. Scale match for tiny person detection//Proceedings of the 2020 IEEE Winter Conference on Applications of Computer Vision. Snowmass: IEEE: 1246-1254 [DOI: 10.1109/WACV45572.2020.9093394]
- Zhang Y L, Yuan Y, Feng Y C and Lu X Q. 2019. Hierarchical and robust convolutional neural network for very high-resolution remote sensing object detection. IEEE Transactions on Geoscience and Remote Sensing, 57(8): 5535-5548 [DOI: 10.1109/tgrs.2019.2900302]
- Zhong Y F, Han X B and Zhang L P. 2018. Multi-class geospatial object detection based on a position-sensitive balancing framework for high spatial resolution remote sensing imagery. ISPRS Journal of Photogrammetry and Remote Sensing, 138: 281-294 [DOI: 10.1016/j.isprsjprs.2018.02.014]
- Zhou P C, Cheng G, Yao X W and Han J W. 2021. Machine learning paradigms in high-resolution remote sensing image interpretation. National Remote Sensing Bulletin, 25(1): 182-197 (周培诚, 程焱, 姚西文, 韩军伟. 2021. 高分辨率遥感影像解译中的机器学习范式. 遥感学报, 25(1): 182-197) [DOI: 10.11834/jrs.20210164]
- Zhu H G, Chen X G, Dai W Q, Fu K, Ye Q X and Jiao J B. 2015. Orientation robust object detection in aerial images using deep convolutional neural network//2015 IEEE International Conference on Image Processing (ICIP). Quebec City: IEEE: 3735-3739 [DOI: 10.1109/ICIP.2015.7351502]
- Zhu P F, Wen L Y, Du D W, Bian X, Fan H, Hu Q H and Ling H B. 2022. Detection and tracking meet drones challenge. IEEE Transactions on Pattern Analysis and Machine Intelligence, 44(11): 7380-7399 [DOI: 10.1109/TPAMI.2021.3119563]
- Zhuang S, Wang P, Jiang B R, Wang G and Wang C. 2019. A single shot framework with multi-scale feature fusion for geospatial object detection. Remote Sensing, 11(5): 594 [DOI: 10.3390/rs11050594]
- Zou Z X and Shi Z W. 2018. Random access memories: a new paradigm for target detection in high resolution aerial remote sensing images. IEEE Transactions on Image Processing, 27(3): 1100-1111 [DOI: 10.1109/TIP.2017.2773199]

A comprehensive review of optical remote-sensing image object detection datasets

YUAN Yiqin, LI Lang, YAO Xiwen, LI Lingjun, FENG Xiaoxu, CHENG Gong, HAN Junwei

School of Automation, Northwestern Polytechnic University, Xi'an 710021, China

Abstract: With the introduction of artificial-intelligence technologies such as deep learning into the field of optical remote-sensing detection, various algorithms have emerged. The use of these algorithms has gradually formed a new paradigm of data-driven optical remote-sensing image object detection. Consequently, high-quality remote-sensing data has become a prerequisite and a necessary resource for researching these paradigm algorithms, highlighting the increasing importance of remote-sensing data. To date, numerous optical remote-sensing image object detection datasets have been published by major research institutions domestically and internationally. These datasets have laid the foundation for the development of deep learning-based remote-sensing image detection tasks. However, no comprehensive summarization and analysis of the published optical remote-sensing image detection datasets have been conducted by scholars. Therefore, this paper aimed to provide a comprehensive review of the published datasets and an overview of algorithm applications. We also aimed to

provide a reference for subsequent research in related fields.

This paper presents an overview and synthesis of the optical remote-sensing image object detection datasets published between 2008 and 2023. The synthesis is based on an extensive and comprehensive survey of literature in the field. By reviewing and analyzing these datasets, we enable a comprehensive understanding of the progress and trends in optical remote-sensing image object detection dataset research.

This paper categorizes the optical remote-sensing image object detection datasets published from 2008 to 2023 based on the annotation method. A comprehensive description of 11 representative datasets is provided, and all dataset information are summarized in tabular form. The analysis considers the information in the datasets themselves and also the spatial and spectral resolution of the images in the datasets. Other basic information including the number of categories, number of images, number of instances, and image-width information are also considered. This analysis effectively demonstrates the trend toward high quality, large scale, and multi-category development of object-detection datasets for optical remote-sensing images. Additionally, we provide an overview of the development and application of algorithms related to published datasets from different perspectives (e.g., horizontal bounding box object detection and rotated bounding box object detection), as well as a subdivision of detection directions (e.g., small object detection and fine-grained detection). Our findings confirm the influential role of remote-sensing data in driving algorithmic advances.

In summary, we offer a comprehensive review of optical remote-sensing image object detection datasets from various perspectives. To our best knowledge, this comprehensive review is the first one on such datasets in the field. The work serves as a valuable reference for subsequent research on deep learning-based optical remote-sensing image object detection, providing insights into data availability and research directions. This study is expected to contribute to the advancement of this field by offering a solid foundation for further investigation and innovation.

Key words: deep learning, optical remote sensing imagery, data source, object detection, development of datasets

Supported by National Natural Science Foundation of China (No. 62071388, 62136007); Key Research and Development Program of Shaanxi Province (No. 2023-YBGY-224)