葡萄牙科英布拉大学联合里斯本大学-俯瞰林木：农林业遥感真实与合成数据集综述 MDPI Remote Sensing

作者：writer 来源：科学网

本文原地址：https://www.iikx.com/news/progress/42189.html

本文仅供学习交流之用，版权归原作者所有，请勿用于商业用途！

葡萄牙科英布拉大学联合里斯本大学-俯瞰林木：农林业遥感真实与合成数据集综述 MDPI Remote Sensing。论文标题：Seeing the Trees from Above: A Survey on Real and Synthetic Agroforestry Datasets for Remote Sensing Applications

论文链接：https://www.mdpi.com/2072-4292/17/19/3346

期刊名： Remote Sensing

期刊主页：https://www.mdpi.com/journal/remotesensing

一、引言

随着无人机（UAV）技术的快速发展和深度学习在遥感领域的广泛应用，高分辨率航空林业数据集的需求日益增长。然而，公开可用的数据集在质量、标注规范和可获取性方面参差不齐，严重制约了相关算法的发展与验证。来自葡萄牙科英布拉大学的学者Babak Chehreh及其合作者在Remote Sensing 期刊发表了综述文章，系统梳理了面向农业林业遥感应用的真实与合成数据集现状，并基于FAIR原则对数据集的可用性进行了量化评估，为研究人员选择合适的数据资源提供了重要参考。

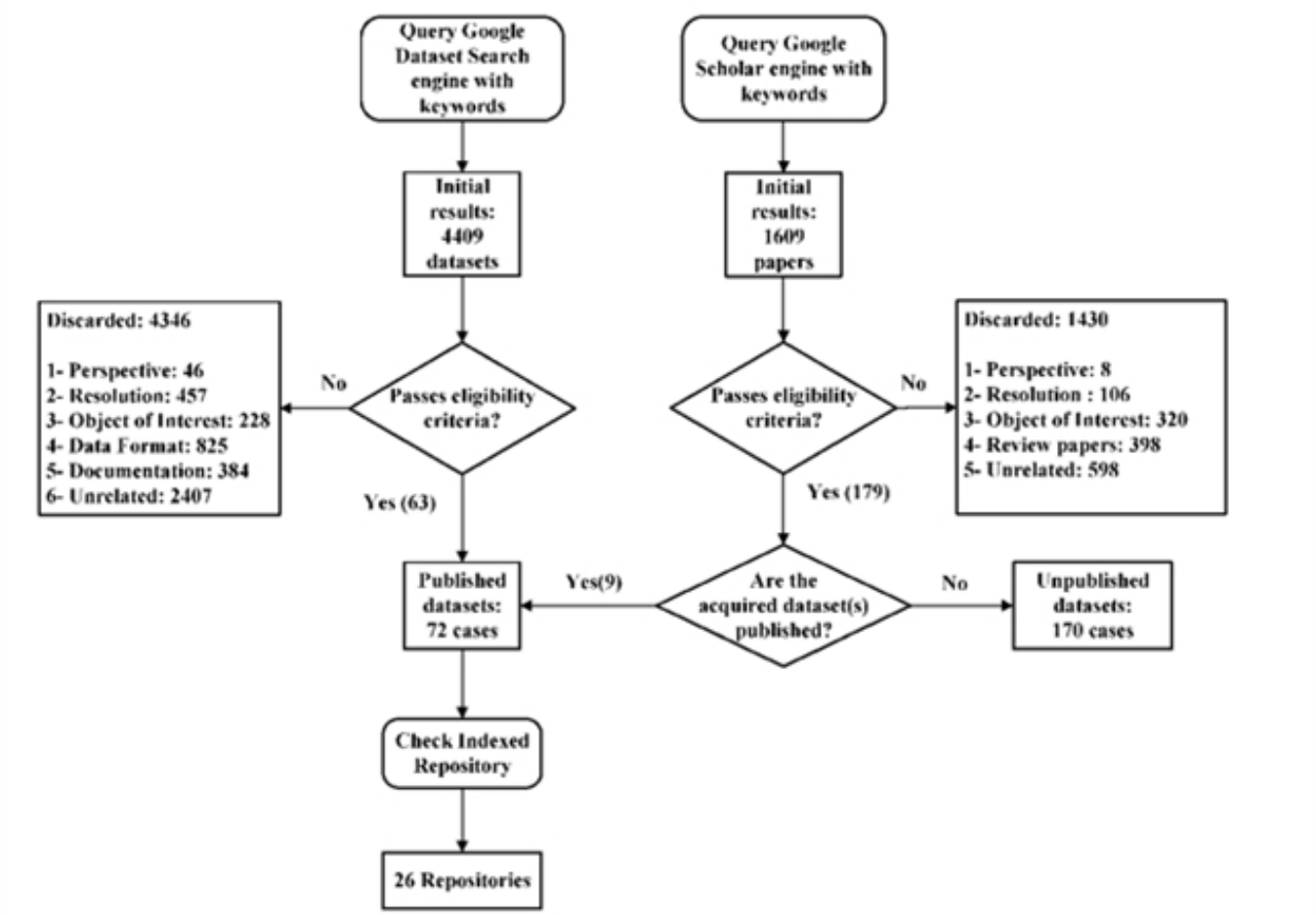


图1 数据集搜索与筛选流程

二、研究过程与结果

作者在文中提出了一套系统的数据集检索与评估框架。研究团队通过Google Dataset Search和Google Scholar双引擎并行检索，最终从4409个数据集候选和1609篇论文中，识别出72个已发布数据集、170个未发布数据集以及26个数据存储库。

在分辨率标准方面，作者创新性地以"单株树木可见性"作为核心判据，将数据分为超高分辨率（UHR，1-5cm GSD）和高分辨率（HR，5-30cm GSD）两类。这一分类方法突破了传统仅以数值界定分辨率的局限，更贴合林业遥感实际应用需求。

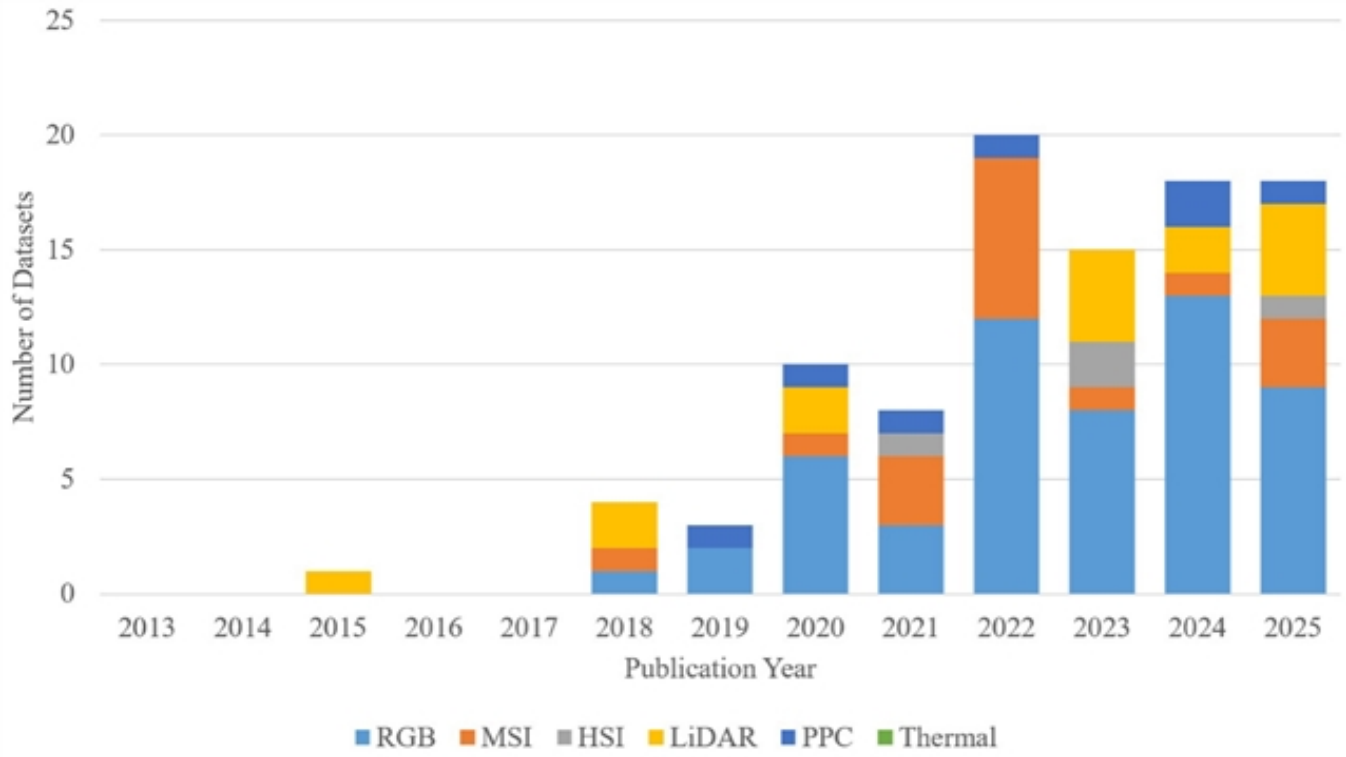


图2 数据源类型年度分布（RGB、多光谱、高光谱、LiDAR、摄影测量点云、热红外）

研究最重要的贡献之一是基于FAIR原则（可查找、可访问、可互操作、可重用）构建了量化评分体系。作者针对F（Findable）、A（Accessible）、I（Interoperable）、R（Reusable）四个维度共13项子指标设计了0-1分的评分标准，并采用加权平均法计算综合FAIR得分（0-4分）。评估结果显示，在72个已发布数据集中，仅3个数据集获得满分4.0分，多数数据集在元数据丰富度（F2）、许可协议明确性（R1.1）和来源追溯（R1.2）方面存在明显不足。得分最低的数据集仅为2.04分，反映出当前林业遥感数据在标准化共享方面仍有较大提升空间。

在应用分类方面，作者将数据集按用途划分为10个类别：单株树冠勾绘（ITCD）、单株树木检测（ITD）、语义分割（SS）、病害检测（DD）、物种识别（SI）、碳储量估算（CSE）、健康状态分析（HSA）、几何测量（GM）、物候监测（PM）及其他。统计发现，ITCD和ITD是最主流的应用方向，合计占比超过50%；而病害检测和碳储量估算等具有重要生态价值但标注难度较高的应用，数据集数量相对匮乏。

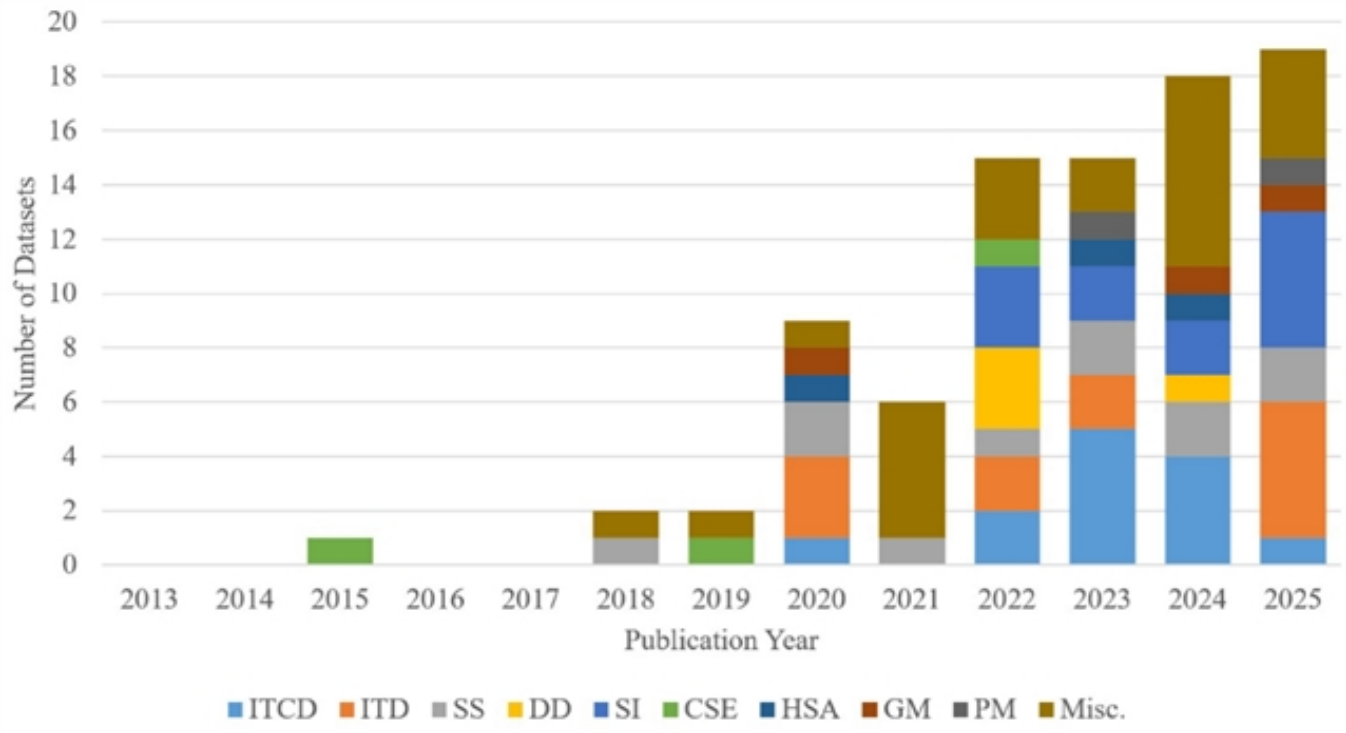


图3 数据集应用类型年度分布

平台分析揭示了一个显著趋势：无人机（UAV）已成为高分辨率林业数据获取的主导平台，在已发布数据集中占比超过70%，远超有人驾驶飞机和卫星平台。这一趋势与无人机成本效益高、机动灵活、可搭载多类型传感器的优势密切相关。在传感器方面，RGB相机仍是最主流配置，但多光谱、LiDAR和高光谱数据的占比近年来呈明显上升态势，多模态数据融合成为重要发展方向。

值得关注的是，作者特别关注了合成数据（Synthetic Data）这一新兴领域。研究识别出3个具有代表性的合成数据集，包括面向低空无人机飞行的MidAir多模态数据集、基于真实无人机图像合成的西伯利亚落叶松树冠数据集SiDroForest，以及利用扩散模型从单张RGB图像重建3D树模型的Tree-D Fusion数据集。作者指出，合成数据在控制变量、降低标注成本、模拟极端场景等方面具有独特优势，有望缓解真实林业数据获取成本高、标注困难、场景覆盖不全等瓶颈问题，但目前仍处于早期发展阶段，距离大规模实际应用尚需进一步成熟。

三、研究总结

本文系统综述了面向农林业遥感的高分辨率航空数据集现状，通过双引擎检索策略和严格的筛选标准，构建了目前最为全面的林业遥感数据集清单。研究基于FAIR原则提出的量化评估方法，为数据集质量评价提供了可操作的框架，有助于推动数据共享标准化。分析表明，无人机平台主导了高分辨率林业数据获取，但多模态协同观测数据仍显稀缺；已发布数据集数量远低于未发布数据，数据开放共享文化亟待加强；合成数据展现出巨大潜力，有望成为未来重要的数据补充来源。

作者强调，当前林业遥感数据集建设面临三大核心挑战：一是未发布数据占比过高（170 vs. 72），大量有价值的数据因各种原因未能公开，制约了领域整体发展；二是多模态对齐数据稀缺，R

GB、多光谱、高光谱、LiDAR等数据的协同采集与配准仍面临技术瓶颈，而这对深度学习中的注意力机制等先进方法至关重要；三是合成数据的真实性与多样性仍需提升，以更好地支撑算法在真实场景中的泛化能力。该综述为农业林业遥感研究人员提供了系统性的数据资源导航，对促进数据共享、推动算法基准测试标准化、加速无人机林业遥感技术的实用化进程具有重要的参考价值。

Chehreh, B.; Moutinho, A.; Viegas, C. Seeing the Trees from Above: A Survey on Real and Synthetic Agroforestry Datasets for Remote Sensing Applications. Remote Sens. 2025, 17, 3346.

期刊介绍

主编：Prasad S. Thenkabail, USGS Western Geographic Science Center (WGSC), USA;

Dongdong Wang, Peking University, China

期刊范围涵盖遥感科学所有领域，从传感器的设计、验证和校准，到遥感在地球科学、环境生态、城市建筑等各方面的广泛应用。

2025 Impact Factor：4.3

2025 CiteScore：9.4

Time to First Decision：22 Days

Acceptance to Publication：2.9 Days

来源：Remote Sensing

更多 科学进展 请访问 https://www.iikx.com/news/progress/

本文版权归原作者所有，请勿用于商业用途，爱科学iikx.com转发