

《农业信息高质量图谱数据专题》卷首语

饶元

安徽农业大学

智慧农业是农业信息化发展的高级阶段，是未来农业发展的必然趋势。中国农业也正向以数据、信息为生产要素，以互联网、物联网、大数据、云计算、区块链、人工智能和智能装备应用为特征的智慧农业迈进。光学传感、深度学习等新一代信息技术正成为驱动农业新质生产力快速发展的重要推动力。

现代农情信息采集过程中，RGB 图像、多光谱、高光谱、激光雷达、实景三维等多源和多维度的成像感知手段提供了丰富的图谱数据源，充分发挥深度学习技术的特征提取与知识挖掘优势，准确进行动植物生命信息、农产品等目标特征的检测与识别，为农业生产管理高质量决策提供支持，是当前农业信息化技术研究领域的热点问题。

深度学习模型的训练往往对数据的质和量有着较高的要求。由于农业生产管理的特殊性与复杂性，农业信息图谱数据，尤其是高质量标注数据尚较为匮乏，深度学习模型在农业领域落地测试缺少足量数据支撑，不同的方法也难以在同一基准下进行比较。

为了促进农业信息领域数据的开放共享，提升数据的可发现性和可复用性，保护数据拥有者的知识产权，本专题在《中国科学数据》（2025 年第 1 期）集中发表了 10 篇农业信息领域高质量图谱数据论文。这些数据论文主要提供了以下几个方面的基础数据：

1) 牛羊动物养殖数据。杜林煊等^[1]构建了基于可见光与红外热成像的奶牛头部多模态数据集，涵盖 2636 组经过配准的可见光图像（VIS）和红外热成像图像（IRT），并对 IRT 图像中的奶牛头部区域进行了精确标注，且通过模糊检测、数据增强等技术优化了质量，为奶牛健康监测、跨模态学习及精准养殖提供基础数据。侯现坤等^[2]采集了多个角度、光照条件和群体规模下的 2538 张奶牛脸部图像，标注了奶牛脸部、左右眼睛中心、鼻子中间、左右嘴角等 5 个关键点，为牛脸识别、姿态估计及行为分析提供数据支撑，助力智能畜牧业发展。黄小平等^[3]整理了涵盖不同光照、奶牛姿态与背景条件下奶牛体况评分目标检测数据集，包含 53566 张奶牛尾部图像，完成了体况评分区域（3.25 - 4.25 分）标注，为自动化体况评分模型研发提供了关键数据。黄小平等^[4]构建了自然状态下纯色山羊和自然环境中绵羊的羊脸图像数据集并完成了高质量标注，覆盖不同遮挡/数量/角度、舍内/外等多样化场景，为羊脸检测模型研究与应用提供了宝贵数据资源。

2) 植物病虫害类数据。池美香等^[5]整理了柑橘黄龙病及相关病害田间症状数据集，涵盖自然生态、白色背景板和黑色背景板等多样化采集环境，并提供了 yolo 格式记录照片中柑橘黄龙病症状种类和发病叶片、果实位置信息的标注文件以及基于 YOLOv5l 训练获得的模型文件，为黄龙病智能诊断、症状分类及复杂环境病害检测提供多场景数据资源和基准模型。杨满仙等^[6]采集了新疆棉田中苜蓿盲蝽、牧草盲蝽、棉铃虫、瓢虫等 13 种（类）常见的棉田昆虫图像，并对每张图像的棉田昆虫进行了人工标注，构建了复杂背景下的小目标棉田昆虫分类数据集和目标检测数据集，为基于深度学习的棉花害虫和益虫精准检测提供基础数据，助力棉花虫害精准防治与生态平衡研究。

3) 农业收获物类数据。张雨等^[7]构建了基于可见光、深度、近红外的不同成熟度番茄

果实多模态图像数据集,涵盖自然光、人工光、微弱光、钠黄光等4种光照条件下的番茄样本图像,包含未熟、半熟、成熟三类成熟度的目标检测和语义分割标注,为番茄智能管理、采收设备的视觉智能模型训练与系统研发提供基础数据支撑。陈文骏等^[8]采集了在不同光照和遮挡条件下青色、紫色两类6个品种的葡萄串高质量多模态视频流数据,并从中提取3954张图像样本进行语义分割和目标检测标注,为多模态视觉数据融合、葡萄串语义分割和目标检测等领域提供宝贵的基础数据资源,对促进葡萄采摘机器人视觉系统优化具有重要价值。常骞等^[9]通过苹果园现场图像拍摄采集与增强处理构建了自然环境下富士苹果花药开裂散粉状态表型数据图集,图像清晰、表征显著,完整展示盛花期苹果花及花药散粉生理变化,弥补了常见花类数据集中反映苹果花散粉生理表征形态数据的不足,适用于深度学习中的图像分类与识别任务,为苹果散粉花的生物生理、精准授粉时间预测模型研发提供可靠基础数据。张哲等^[10]整理了离体豆荚与植株图像集,其中离体豆荚包括无重叠离体豆荚和重叠离体豆荚,植株包括单分枝型、双分枝型和复杂分枝型;标注文件数据包含了大豆植株主茎的实例分割标注、主茎节的检测框标注、在体豆荚的检测框标注、离体豆荚的实例分割和检测框标注,且在体豆荚和离体豆荚的检测框标注采用了正框和旋转框两种标注方式,标注主茎节、豆荚检测框及实例分割。该数据集为大豆植株的离体豆荚检测、在体豆荚检测与植株形态分析的目标检测和实例分割模型方法研究提供了宝贵基础图像数据资源,对促进大豆智能化考种方法研究亦具有重要价值。

动物类聚焦奶牛、羊的生理监测与识别,提供多模态图像与关键点标注数据;病虫害类侧重病害症状与害虫检测,为复杂环境下的农业病虫害智能诊断提供数据基础;农业收获物类覆盖番茄、大豆等作物成熟度与表型分析,为智能作业装备视觉智能系统研发提供基础数据。这些数据论文不仅提供了数据集细节描述,还提供了数据收集和加工处理方法、数据质量评估和验证方法,以及其他便于理解和使用数据的相关信息。

我们希望通过《农业信息高质量图谱数据专刊》推出的有关农业信息图谱系列数据论文的发表,促进中国数字农业科学数据的开放共享,助力农业数字化应用场景研发、提高智慧农业应用落地水平,推进中国现代农业高质量发展。

参考文献

- [1] 杜林煊, 曹姗姗, 刘婷婷, 等. 奶牛个体头部多模态图像深度学习训练数据集[J/OL]. 中国科学数据, 2025, 10(1). (2025-02-12). DOI: 10.11922/11-6035.nasdc.2025.0001.zh. [DU L X, CAO S S, LIU T T, et al. A dataset of deep learning training for multimodal head images of individual cows[J/OL]. China Scientific Data, 2025, 10(1). (2025-02-12). DOI: 10.11922/11-6035.nasdc.2025.0001.zh.]
- [2] 侯现坤, 黄小平, 黄飞, 等. 奶牛脸部及关键点检测数据集[J/OL]. 中国科学数据, 2025, 10(1). (2024-09-29). DOI:10.11922/11-6035.csd.2024.0129.zh. [HOU X K, HUANG X P, HUANG F, et al. A dataset of cow face and keypoint detection[J/OL]. China Scientific Data, 2025, 10(1). (2024-09-29). DOI:10.11922/11-6035.csd.2024.0129.zh.]
- [3] 黄小平, 豆子豪, 黄飞, 等. 奶牛体况评分目标检测数据集[J/OL]. 中国科学数据, 2025, 10(1). (2025-02-24). DOI: 10.11922/11-6035.csd.2024.0126.zh. [HUANG X P, DOU Z H,

- HUANG F, et al. A dataset of target detection for dairy cow body condition scores[J/OL]. China Scientific Data, 2025, 10(1). (2025-02-24). DOI: 10.11922/11-6035.csd.2024.0126.zh.]
- [4] 黄小平, 豆子豪, 黄飞, 等. 山羊和绵羊脸部检测数据集[J/OL]. 中国科学数据, 2025, 10(1). (2024-10-12). DOI:10.11922/11-6035.csd.2024.0085.zh. [HUANG X P, DOU Z H, HUANG F, et al. A dataset of goat and sheep facial images[J/OL]. China Scientific Data, 2025, 10(1). (2024-10-12). DOI:10.11922/11-6035.csd.2024.0085.zh.]
- [5] 池美香, 陈韶萍, 黄婷, 等. 柑橘黄龙病田间症状识别图像数据集[J/OL]. 中国科学数据, 2025, 10(1). (2025-02-17). DOI: 10.11922/11-6035.csd.2023.0161.zh. [CHI M X, CHEN S P, HUANG T, et al. A dataset of field symptoms for identifying Citrus Huanglongbing[J/OL]. China Scientific Data, 2025, 10(1). (2025-02-17). DOI: 10.11922/11-6035.csd.2023.0161.zh.]
- [6] 杨满仙, 陈燕红, 李雨晴, 等. 新疆棉田主要昆虫图像数据集 CottonInsect [J/OL]. 中国科学数据, 2025, 10(1). (2024-04-03). DOI: 10.11922/11-6035.csd.2023.0130.zh. [YANG M X, CHENG Y H, LI Y Q, et al. A dataset of common insects in Xinjiang cotton fields: CottonInsect[J/OL]. China Scientific Data, 2025, 10(1). (2024-04-03). DOI: 10.11922/11-6035.csd.2023.0130.zh.]
- [7] 张雨, 饶元, 陈文骏, 等. 不同成熟度番茄果实多模态图像数据集[J/OL]. 中国科学数据, 2025, 10(1). (2025-02-17). DOI: 10.11922/11-6035.csd.2023.0154.zh. [ZHANG Y, RAO Y, CHEN W J, et al. A dataset of multimodal images of tomato fruits at different stages of maturity[J/OL]. China Scientific Data, 2025, 10(1). (2025-02-17). DOI: 10.11922/11-6035.csd.2023.0154.zh.]
- [8] 陈文骏, 饶元, 王丰仪, 等. 葡萄多模态目标检测和语义分割数据集[J/OL]. 中国科学数据, 2025, 10(1). (2023-10-07). DOI: 10.11922/11-6035.csd.2023.0132.zh. [CHEN W J, RAO Y, WANG F Y, et al. A dataset of grape multimodal object detection and semantic segmentation[J/OL]. China Scientific Data, 2025, 10(1). (2023-10-07). DOI: 10.11922/11-6035.csd.2023.0132.zh.]
- [9] 常骥, 尹龙, 李壮, 等. 辽宁绥中富士苹果花药开裂散粉状态表型数据图集 ABP51[J/OL]. 中国科学数据, 2025, 10(1). (2025-02-20). DOI:10.11922/11-6035.csd.2024.0156.zh. [CHANG J, YIN L, LI Z, et al. A dataset of anther dehiscence and pollen dispersal phenotypic atlas for Fujii apples in Suizhong, Liaoning: ABP51[J/OL]. China Scientific Data, 2025, 10(1). (2025-02-20). DOI:10.11922/11-6035.csd.2024.0156.zh.]
- [10] 张哲, 饶元, 束雅莉, 等. 大豆室内考种图像数据集[J/OL]. 中国科学数据, 2025, 10(1). (2024-08-21). DOI: 10.11922/11-6035.csd.2023.0104.zh. [ZHANG Z, RAO Y, SHU Y L, et al. A dataset of images for indoor soybean survey[J/OL]. China Scientific Data, 2025, 10(1). (2024-08-21). DOI: 10.11922/11-6035.csd.2023.0104.zh.]

论文引用格式

饶元. 《农业信息高质量图谱数据专题》卷首语[J/OL]. 中国科学数据, 2025, 10(1). (2025-03-13). CSTR: 32001.14.11-6035.csd.2025.0027.zh. DOI: 10.11922/11-6035.csd.2025.0027.zh.