

中图法分类号: TP392 文献标识码: A 文章编号: 1006-8961(XXXX)XX-0001-15

论文引用格式: Liu Pengdong, Lin Hanran, Wu Maonian, Zheng Bo, Zhu Shaojun, Shi Xingren. A multi-color artificial illumination image dataset for nighttime melon segmentation MelonNightColor[J/OL]. Journal of Image and Graphics, XXXX: 1-15. DOI: 10.11834/jig.260410. (刘鹏栋, 林翰冉, 吴茂念, 郑博, 朱绍军, 施星仁. 多色人工光照夜间甜瓜分割图像数据集 MelonNightColor[J/OL]. 中国图象图形学报, XXXX: 1-15. DOI: 10.11834/jig.260410.) [DOI: 10.11834/jig.260410]

多色人工光照夜间甜瓜分割图像数据集 MelonNight-Color

刘鹏栋^{1†}, 林翰冉^{1†}, 吴茂念^{1,2*}, 郑博¹, 朱绍军¹, 施星仁³

1. 湖州师范大学, 浙江 湖州 313000; 2. 中国-莫桑比克智慧农业“一带一路”联合实验室, 浙江师范大学, 浙江 金华 321004; 3. 湖州吴兴金农生态农业发展有限公司, 浙江 湖州 313000

摘要: 目的 夜间甜瓜识别和采摘严重依赖人工光源。不同人工补光配置下, 甜瓜果实和茎部的图像表现可能不同, 语义分割模型的训练和跨配置测试结果也可能存在差异。为助力夜间农产品识别和采摘, 构建了多颜色人工光照甜瓜分割图像数据集 MelonNightColor。方法 在保持相机位置、拍摄距离和补光灯方位一致的条件下, 切换补光灯红、绿、蓝通道配置, 获取红、蓝、绿、青、紫、黄、白 7 种人工补光配置下的夜间甜瓜图像。当前已发布 5474 张图像及像素级标注(覆盖 782 个拍摄位置, 标注背景、甜瓜果实和茎部 3 个类别)。以拍摄位置为最小划分单元划分训练集、验证集和测试集, 确保同一场景的 7 种颜色图像不跨集合, 并组织为 7 个单颜色子集和 4 个组合颜色子集。采用 9 种代表性语义分割模型进行基线评测, 并以 UNet 构建 7×7 跨颜色泛化矩阵。结果 9 种模型在 7 种光源下的背景、果实和茎部的类别交并比(intersection over union, IoU)均值分别为 99.05%、86.39% 和 38.84%, 茎部因形态细长且易受遮挡影响而分割难度最大。白光和黄光下 9 种模型平均交并比(mean intersection over union, mIoU)均值较高, 分别达 78.17% 和 77.94%; 红光和蓝光下相对较低, 分别为 72.93% 和 73.23%。UNet 同颜色测试平均 mIoU 为 80.38%, 跨颜色测试平均 mIoU 为 44.25%。该结果显示, 在本数据集的设备、品种和采集条件下, 不同补光配置及自动成像过程共同作用下, 训练与测试补光配置不一致时模型性能出现明显下降。结论 MelonNightColor 提供了夜间甜瓜多颜色成组图像和像素级标注。未来, 数据集 MelonNightColor 可为特定夜间人工补光条件下的甜瓜语义分割基线评测和跨配置模型研究提供公开数据基础。数据集已在中国科学院科学数据银行公开, 数字对象标识符(digital object identifier, DOI)为 <https://doi.org/10.57760/sciencedb.j00240.00237>。

关键词: 设施农业; 夜间图像; 甜瓜分割; 人工光源; 多颜色数据集; 语义分割

A multi-color artificial illumination image dataset for nighttime melon segmentation MelonNightColor

Liu Pengdong^{1†}, Lin Hanran^{1†}, Wu Maonian^{1,2*}, Zheng Bo¹, Zhu Shaojun¹, Shi Xingren³

1. Huzhou Normal University, Huzhou 313000, Zhejiang, China; 2. China-Mozambique "Belt and Road" Joint Laboratory on Smart Agriculture, Zhejiang Normal University, Jinhua 321004, China; 3. Huzhou Wuxing Jinnong Ecological Agriculture Development Co., Ltd.,

收稿日期: 2026-07-09; 修回日期: 2026-08-25

* 通信作者: 吴茂念, wmn@huznu.edu.cn

基金项目: 中国-莫桑比克智慧农业“一带一路”联合实验室开放课题项目(项目编号: CMSA-KF202505); 浙江省大学生科技创新活动计划(新苗人才计划)(项目编号: 2025R446C029); 湖州市科技攻关计划项目(项目编号: 2024GN04)

Supported by: the Open Funding of China-Mozambique "Belt and Road" Joint Laboratory on Smart Agriculture (No. CMSA-KF202505); the Zhejiang Provincial College Students' Science and Technology Innovation Program (No. 2025R446C029); the Science and Technology Plan Project of Huzhou (No. 2024GN04)

Huzhou 313000, Zhejiang, China

Abstract: Objective Nighttime melon recognition, localization, and harvesting in protected agriculture rely heavily on artificial illumination. Under different artificial illumination configurations, the apparent color, brightness, contrast, texture, and boundary visibility of melon fruits and stems may change, which can affect semantic segmentation performance. A model trained under one illumination configuration may therefore perform poorly when applied to images captured under another configuration. However, existing agricultural image datasets rarely provide spatially corresponding nighttime images of the same scene under multiple artificial illumination conditions. To address this gap, we constructed and released MelonNightColor, a multi-color artificial illumination image dataset for nighttime melon semantic segmentation. The dataset is designed to support baseline model evaluation, illumination-aware image analysis, and cross-configuration generalization research. **Method** Images were collected from May 16 to June 4, 2025, in a melon cultivation facility in Huzhou, Zhejiang Province, China. The cultivar was the "Xiaobainiang" melon. All images were captured at night, with artificial illumination as the only light source in the acquisition area. A 12 W portable RGB LED lamp was placed beside the camera at an approximately fixed distance from the melon plants. The camera position, shooting distance, lamp position, and illumination direction were kept consistent throughout the acquisition process. Images were captured using a GM1910 camera in professional mode, with a fixed aperture of $f/1.6$ and a focal length of 27 mm (35-mm equivalent). The flash was disabled, while exposure time, ISO sensitivity, white balance, and focus were automatically adjusted. Seven illumination configurations were generated by switching the red, green, and blue LED channels: red, blue, green, cyan, purple, yellow, and white. Red, blue, and green were obtained by activating one channel, whereas cyan, purple, and yellow were generated by simultaneously activating the green-blue, red-blue, and red-green channel combinations, respectively. White was defined as the simultaneous activation of all three channels. At each shooting location, the camera and scene were kept stationary while the seven configurations were applied sequentially. Consequently, each location produced a group of seven spatially corresponding images. The current release includes 5,474 pixel-level annotated images from 782 shooting locations. An additional 1,218 groups (8,526 images) are publicly available without annotations; their pixel-level annotations will be released incrementally in future updates. The original images are 24-bit sRGB JPG files with a resolution of $4,000 \times 3,000$ pixels. For the annotated portion, each image is paired with an 8-bit single-channel PNG mask of the same resolution. The annotation task includes three semantic classes: background, melon fruit, and melon stem. When fruits or stems were partially occluded, only their visible regions were annotated. Annotation quality was controlled through manual labeling, independent cross-checking, automated file-pair and size verification, legal pixel-value inspection, foreground-ratio analysis, scene-level split checking, and visual spot checks. The data were divided into training, validation, and test sets at a ratio of 7:2:1, with the shooting location used as the minimum splitting unit. Thus, the seven images from the same location were assigned to the same subset, preventing scene-level data leakage. MelonNightColor is organized into seven single-color subsets and four combined-color subsets. The latter were formed by merging single-color images at the data-organization level and do not represent additional physical illumination conditions or unique images. Nine representative semantic segmentation models were evaluated, including UNet, PSPNet, UNet 3+, SwinUNet, VM-UNet, DenseNet, DenseUNet, DeepLab, and H-vmunet. All models used the same data split and training protocols. Images and masks were resized to 224×224 pixels, and the models were trained using AdamW with a combined cross-entropy and Dice loss. No color jittering, hue adjustment, channel permutation, or synthetic recoloring was applied. In addition, UNet was used to construct a 7×7 cross-color generalization matrix, with training illumination configurations as rows and testing configurations as columns. **Results** The baseline experiments confirmed that MelonNightColor supports the training and comparison of mainstream semantic segmentation models under nighttime artificial illumination. Averaged across the seven illumination configurations and nine baseline models, the IoUs of the background, melon fruit, and melon stem classes were 99.05%, 86.39%, and 38.84%, respectively. The stem was the most challenging class because it was thin, small, frequently occluded by leaves and vines, and often located near cultivation supports or shadowed regions. The average mIoU varied among illumination configurations. White illumination achieved the highest average mIoU of 78.17%, followed by yellow illumination at 77.94%. Red and blue illumination produced relatively low average mIoUs of 72.93%

and 73.23%, respectively. UNet 3+, UNet, DenseUNet, and VM-UNet achieved comparatively strong results across most configurations. The average mIoUs of the red-blue, red-green, green-blue, and red-green-blue combined-color subsets were 73.30%, 73.90%, 74.52%, and 74.23%, respectively. These results describe the corresponding data organization and evaluation protocols rather than directly demonstrating the superiority of one physical light source over another. The cross-color generalization experiment revealed a substantial performance difference between matched and mismatched illumination conditions. UNet achieved an average mIoU of 80.38% when trained and tested under the same illumination configuration. When the training and testing configurations differed, the average mIoU decreased to 44.25%. In the most challenging example, training under yellow illumination and testing under red illumination produced an mIoU of 17.90%. However, several mismatched combinations retained relatively high performance, such as red-to-yellow, blue-to-white, and purple-to-white testing. These results indicate that the effect of illumination changes varies across color combinations and that MelonNightColor can support research on illumination robustness, color normalization, domain adaptation, and cross-condition semantic segmentation. **Conclusion** MelonNightColor provides a publicly available benchmark for nighttime melon semantic segmentation under multiple artificial illumination configurations. Its main contribution is the grouped acquisition design, in which each shooting location is represented under seven RGB channel combinations. The current annotated portion is accompanied by pixel-level annotations for background, fruit, and stem. The dataset provides pre-defined scene-level data splits, seven single-color subsets, four combined-color subsets, nine-model baseline results, and a UNet-based 7×7 cross-color generalization matrix. These resources enable researchers to study conventional segmentation accuracy as well as the effects of artificial illumination on image appearance, model stability, and cross-configuration transferability. The dataset was collected using a specific cultivar, camera, LED lamp, acquisition site, growth stage, and automatic exposure process. Therefore, the reported results should be interpreted as empirical benchmarks for the stated acquisition conditions rather than causal estimates of the independent effects of spectrum, illumination intensity, or camera parameters. MelonNightColor is publicly available through the Science Data Bank of the Chinese Academy of Sciences under the CC BY-NC-ND 4.0 license. Its DOI is <https://doi.org/10.57760/sciencedb.j00240.00237>.

Key words: facility agriculture; nighttime image; melon segmentation; artificial illumination; multi-color dataset; semantic segmentation

论文引用格式: Liu P D, Lin H R, Wu M N, Zheng B, Zhu S J and Shi X R. 2026. A multi-color artificial illumination image dataset for nighttime melon segmentation MelonNightColor. *Journal of Image and Graphics*, (刘鹏栋, 林翰冉, 吴茂念, 郑博, 朱绍军, 施星仁. 2026. 多色人工光照夜间甜瓜分割图像数据集 MelonNightColor. *中国图象图形学报*, [DOI: 10.11834/jig.260410]).

0 引言

设施农业是实现全年高效作物生产的重要途径,计算机视觉技术正逐步成为温室智能化管理的核心支撑,广泛用于生长监测、产量估计、果实定位与表型分析等环节(Tian等,2020;Ghazal等,2024)。甜瓜是设施栽培中的重要经济作物,对其果实区域进行精确分割,是产量估计和采收决策等下游任务的基础(Kamilaris和Prenafeta-Boldú,2018)。夜间并

非仅是日间作业的时间延伸。在部分设施生产和采后作业场景中,夜间不存在日间直射强光,自然光缺失时作物视觉系统可依靠人工补光获取图像。夜间采集可扩展作物视觉感知、连续监测和采摘的可用时间窗口,并为后续采后处理与物流衔接提供时间上的可能性。因此,建立真实的夜间人工补光图像数据集,有助于研究夜间条件下的作物视觉感知问题。

围绕果实检测与分割,已有大量工作取得了显著进展。在日间自然光条件下,研究者针对苹果(Hani等,2020)、葡萄(Santos等,2020)、草莓(Pérez-Borrero等,2020)、绿色果实(Jia等,2022)、柑橘(梁云等,2024)以及芒果、杏仁等作物(Bargoti和Underwood,2017)开展了检测、分割或采摘定位研究;在温室场景中,甜椒(López-Barrios等,2023)和番茄(Dong等,2025)的果实定位方法也得到了关注。这些工作为农业视觉感知奠定了坚实基础,但它们均在日间自然光或固定补光条件下采集数据,未涉及

夜间多种颜色人工光源的系统性控制。在夜间农业视觉方面,已有学者针对荔枝(Liang等,2020;Cao等,2024)、苹果(Wei等,2018;Zhang等,2024)、柑橘(熊俊涛等,2024)、番石榴(Su等,2025)等作物开展了夜间环境下的检测与识别探索,相关研究还涉及低光照图像增强与语义分割(张航和颜佳,2024)、夜间机器人作业(Waltman等,2024)和果园监测(Kutyrev等,2025)。现有夜间农业视觉研究多聚焦于特定补光条件下的检测或识别任务;在本文调研范围内,同一场景多种人工补光配置下的成组甜瓜分割图像仍较少见。

综合以上分析,同时包含夜间人工补光采集、明确的红、绿、蓝(red, green, and blue, RGB)通道配置以及同一场景多配置成组图像的甜瓜语义分割数据仍较为有限。

上述不足限制了夜间甜瓜分割中补光颜色选择、模型稳定性评估和跨光源泛化分析。在本数据集的夜间人工补光和自动成像条件下,不同补光配置下的图像表现及模型测试结果可能存在差异。在一种补光配置数据上训练的模型,应用于另一种补光配置的测试图像时,性能可能下降。本数据集可为后续开展跨补光配置的模型泛化评测与相关方法研究提供公开数据基础。综述(Kamilaris和Prenafeta-Boldú,2018;Akbar等,2024;Tang等,2023)也指出,农业视觉研究中的环境多样性主要停留在作物种类和生长阶段层面,对光照颜色等变量的系统性覆盖仍然不足。

针对上述空白,本文构建并公开 MelonNightColor 数据集。该数据集聚焦于特定夜间人工补光条件下采集的甜瓜成组图像及像素级标注,为日间数据之外的相关研究提供补充资源。本文公开发布并实验验证其中已完成像素级标注的782组(5474张),另有1218组(8526张)将逐步标注并公开。采集时保持相机位置、拍摄距离和补光灯方位一致,依次切换发光二极管(light-emitting diode, LED)光源的RGB通道配置,在同一拍摄位置获得红、蓝、绿、青、紫、黄、白7种人工补光配置下的成组图像。成组图像保持场景内容和空间位置的对应,可用于比较不同补光配置下的图像表现与模型结果。已标注的每张图像均配备背景、甜瓜果实和甜瓜茎部的像素级标注,数据集按补光配置组织为7个单颜色子集和4个组合颜色子集,共11个子数据集,并以拍摄

位置为最小单元划分为训练集、验证集和测试集(7:2:1),确保同场景7色图像不跨集合。

本文的主要贡献包括四个方面:

1)构建并公开了包含5474张图像、7种人工补光配置和3个语义类别的夜间甜瓜分割数据集 MelonNightColor。

2)在保持拍摄位置、拍摄距离及补光灯方位一致的条件下,构建了7种RGB通道补光配置的成组图像,为比较特定采集条件下不同补光配置的图像表现与模型结果提供数据基础。

3)同时组织7个单颜色子集和4个组合颜色子集,为不同数据组织方式下的模型评测提供数据基础。

4)提供了9种分割模型的基线评测和基于UNet的7×7跨颜色泛化矩阵,为后续研究提供可复现的参照基准。

1 数据集属性

1.1 数据集基本信息

MelonNightColor是一个面向夜间甜瓜语义分割的多颜色人工光照图像数据集,作物品种为“小白娘”甜瓜。当前发布的已标注部分包含5474张已完成像素级标注的图像,原始图像以JPG格式存储,文件扩展名为“.jpg”,位深度为24位,色彩空间为标准红绿蓝色彩空间(standard Red Green Blue, sRGB),分辨率为4000×3000像素,文件大小分布在2-5MB。数据集的核心特色在于:在保持相机位置、拍摄距离和补光灯方位一致的条件下,通过切换LED人工光源的红、绿、蓝通道配置,在同一拍摄位置获得7张对应不同人工补光配置的成组图像。拍摄采用专业模式,组内图像反映补光配置与相机成像过程共同作用下的实际成像差异。数据集按光源颜色组织为7个单颜色子集和4个组合颜色子集,以拍摄位置为最小单元随机划分为训练集、验证集和测试集,确保同组7色图像不跨集合。数据集以知识共享署名-非商业性使用-禁止演绎4.0国际许可(Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International Public License, CC BY-NC-ND 4.0)公开发布。

本版本数据集采集于特定地点、甜瓜品种、生长阶段、拍摄设备和补光设备条件下,其数据分布及基

线结果主要适用于该采集范围。数据集记录了上述条件下不同人工补光配置对应的成组图像和像素级标注,可为后续研究提供公开数据基础。

1.2 标注数据属性

每张原图均对应一张同名的便携式网络图形

(Portable Network Graphics,PNG)格式掩膜文件,掩膜为8位单通道灰度图像,分辨率与原图一致,均为4000×3000像素。每个像素的取值及其语义含义如表1所示。

表1 标注像素值及其含义

Table 1 Pixel values and semantic meanings of annotations

像素值	类别名称	含义
0	背景	叶片、枝蔓、土壤、栽培支架等所有非甜瓜目标区域
1	甜瓜果实	甜瓜果实可见区域,叶片或枝蔓部分遮挡时仅标注可见部分
2	甜瓜茎部	连接甜瓜果实与植株主蔓的茎部可见区域

标注工作采用人工标注与交叉审核相结合的方式。首先由标注人员在原图上逐像素勾画前景区域的轮廓;标注完成后由另一位审核人员逐一核对标注边界是否贴合可见轮廓、是否存在漏标或误标;最后通过自动化脚本对全部标注进行批量质量检查,检查项目包括JPG与PNG文件是否一一对应、原图与掩膜尺寸是否一致、掩膜像素值是否仅含0,1,2三类以及前后景占比是否异常,确保标注数据的一致性和完整性。

1.3 数据集文件的组织结构

MelonNightColor已标注数据集按照子数据集、数据划分和文件类型三级结构组织。第一级目录为11个子数据集,包括7个单颜色子数据集data_red、data_blue、data_green、data_cyan、data_purple、data_yellow、data_white,以及4个组合颜色子数据集data_red_blue、data_red_green、data_green_blue、data_red_green_blue。第二级目录为train、val和test,分别对应训练集、验证集和测试集。第三级目录为images和masks,其中images存放JPG格式原始图像,masks存放与原图同名的PNG格式像素级标注掩膜。

其中,data_red至data_white每个子数据集仅包含对应光源颜色下采集的图像及其标注;data_red_blue、data_red_green、data_green_blue和data_red_green_blue由对应单颜色子集的图像按文件合并而成,用于支持多颜色联合训练的实验需求。所有子数据集采用一致的目录命名和文件配对规则,研究者无需依赖外部元数据文件,直接按子数据集名称、数据划分和文件类型即可读取和使用数据。

1.4 数据集的组成与统计

表2列出了11个子数据集的图像数量及训练、验证、测试集划分。7个单颜色子数据集各包含782张图像,对应782个不同拍摄位置,每个位置在7种光源颜色下各采集1张;4个组合颜色子数据集由对应单颜色子集按图像合并得到,其中两种光源合并子集各为1564张,三种光源合并子集为2346张。已标注数据集共计5474张。

表2 各子数据集的图像数量与划分

Table 2 Image numbers and data splits of each subset

子数据集	光源	train	val	test	合计
data_red	red	547	156	79	782
data_blue	blue	547	156	79	782
data_green	green	547	156	79	782
data_cyan	cyan	547	156	79	782
data_purple	purple	547	156	79	782
data_yellow	yellow	547	156	79	782
data_white	white	547	156	79	782
data_red_blue	red+blue	1094	312	158	1564
data_red_green	red+green	1094	312	158	1564
data_green_blue	green+blue	1094	312	158	1564
data_red_green_blue	red+green+blue	1641	468	237	2346
单颜色合计	不适用	3829	1092	553	5474

表2中cyan、purple、yellow和white在拍摄阶段分别由绿蓝、红蓝、红绿和红绿蓝通道同时开启获得;data_red_blue、data_red_green、data_green_blue和data_red_green_blue则是拍摄完成后将对应单通道

补光图像在数据层面合并得到的组合子数据集,并非拍摄阶段形成的新补光配置。组合子数据集中的样本计数仅用于实验协议说明,不计入5474张唯一已标注图像总量。

2 数据集描述

2.1 数据采集场景与设备

数据的采集工作于2025年5月16日至6月4日在浙江省湖州市一处甜瓜种植基地进行。所有图像均在夜间时段拍摄,采集区域除人工补光外无其他环境光源。采集场景的布局如图1所示,补光灯固定于相机一侧,与甜瓜植株保持约1m的拍摄距离。在完整采集周期内,补光灯的摆放位置、拍摄距离、拍摄设备及拍摄模式均保持恒定,以提高成像条件的可复现性。

由图1可见,采集装置采用近距离固定拍摄方式,补光灯与相机在采集过程中保持相对稳定。该布置使相机、补光灯和植株之间的几何关系在成组采集过程中保持一致,并记录不同作物姿态、遮挡关系和补光配置下的实际图像。



图1 数据采集场景与设备布置

Fig. 1 Data acquisition scene and equipment layout

拍摄相机型号为GM1910,设置为专业模式,光圈固定为 $f/1.6$,35mm等效焦距为27mm,闪光灯强制关闭,曝光时间、感光度、白平衡和对焦均采用自动设置。为说明实际采集参数,本文对7个单颜色子集共5474张JPG图像的可交换图像文件格式(Exchangeable Image File Format, EXIF)信息进行了统计。7种补光配置的曝光时间中位数为1/33~1/20s,ISO中位数为640~1250,具体结果见表3。

由表3可见,不同补光配置下相机采用的曝光时间和ISO组合并不完全相同。因此,本文后续报告的图像表现和模型结果,反映的是相应补光配置与相机自动成像共同形成的实际成像条件。

表3 7种人工补光配置下图像的EXIF参数统计

Table 3 EXIF parameter statistics of images under seven artificial illumination configurations

补光配置	图像数	曝光时间中位数	ISO中位数
red	782	1/20s	1000
green	782	1/33s	640
blue	782	1/20s	1250
cyan	782	1/20s	800
purple	782	1/20s	1000
yellow	782	1/25s	800
white	782	1/20s	800

辅助光源选用速图SUTEFOTO T15 APP RGB便携补光灯,额定功率为12W;CWRGB模式下各通道可独立调节,亮度调节范围为0~100%。本研究采集时仅启用红、绿、蓝三个色彩通道,各启用通道的亮度档位均设为最大值,以保证不同补光配置下甜瓜果实和茎部具有可见性。本文将不同通道组合作为不同人工补光配置进行数据组织与结果分析。

2.2 多颜色人工光源设计

7种人工补光配置通过红、绿、蓝三个通道的开关组合实现。red仅开启红通道,blue仅开启蓝通道,green仅开启绿通道;cyan为绿通道和蓝通道同时开启,purple为红通道和蓝通道同时开启,yellow为红通道和绿通道同时开启;white为红、绿、蓝三个通道同时开启,对应本研究定义的white补光配置。

数据集的核心采集策略为成组拍摄:在每一个拍摄位置,保持相机位置和场景不动,依次改变补光灯的通道配置获得7张图像,构成一组。已标注数据集共覆盖782个不同拍摄位置,每组内的7张图像在空间位置上保持对应。该设计使同组图像在拍摄位置、场景内容、相机与补光灯的几何布局上保持一致,并通过切换补光灯通道配置形成不同采集条件。拍摄时曝光时间和ISO采用自动模式,组内图像记录了不同补光配置下的实际成像结果;场景、视角和拍摄距离在采集过程中保持一致。

cyan、purple、yellow和white四种补光配置是在拍摄阶段通过同时开启相应LED通道获得的图像,而组合颜色子数据集是在拍摄完成后将多个单颜色子集在数据层面合并形成。两类子集的形成方式不同,分别对应拍摄阶段多通道同时开启的实际采集

图像和多个单颜色子集的联合数据组织。

在7种人工补光配置下,同一场景图像呈现出不同的整体色彩表现。red、blue、green、cyan、purple、yellow 和 white 分别对应不同的补光通道组合。

为直观展示同一场景在不同人工补光配置下的图像表现,本文选取同一拍摄位置下的7张成组图像进行对比,如图2所示。



图2 同一场景在7种人工补光配置下的成组图像
Fig. 2 Grouped images of the same scene under seven artificial illumination configurations

由图2可见,同一甜瓜场景在不同补光配置下呈现出不同的图像表现。这些图像记录了当前设备和采集设置下7种人工补光配置对应的实际成像结果。

2.3 单颜色与组合颜色数据集

基于上述7种人工补光配置的采集策略,Melon-NightColor共组织为7个单颜色子集和4个组合颜色子集。7个单颜色子数据集各自仅包含一种补光配置下的图像,是单配置基线评测和跨颜色泛化矩阵中训练与测试的基本单元。4个组合颜色子数据集由多个单颜色子数据集在数据层面合并形成,为呈现训练阶段多颜色图像合并与拍摄阶段多通道同时开启所得图像在各自评测协议下的基线结果提供数据基础。当前两类子数据集的训练样本数量和测试图像构成不同,相关结果分别按各自数据组织与评测协议报告。

具体数据组织关系包括4组:红蓝合并子集与purple子集、红绿合并子集与yellow子集、绿蓝合并子集与cyan子集、红绿蓝合并子集与white子集。这些数据组织用于报告不同训练数据构成及相应评测协议下的基线结果。

2.4 数据划分策略

全数据集统一按7:2:1的比例划分为训练集、验证集和测试集,各子数据集的划分结果如表2所示。划分以拍摄位置为基本单位:同一拍摄位置下采集的7张不同补光配置图像作为一个不可分割的整体,被分配到训练集、验证集或测试集中的同一集合,确保测试集中的所有场景在训练阶段不可见,避免同一场景的不同配置版本分别出现在训练集和测

试集而导致场景级数据泄漏。4个组合颜色子数据集的划分方式继承自其来源的单颜色子数据集:以红蓝合并子集为例,其训练集由红光子集和蓝光子集的训练图像合并构成,验证集和测试集同理。这一设计使组合颜色子数据集与对应的多通道补光子数据集采用一致的场组划分规则,并避免同一拍摄位置的图像跨训练集、验证集和测试集。

2.5 标注规范

数据集的标注目标为三分类语义分割,标注类别如表1所示。标注工作由标注人员在原图上逐像素完成,不同场景遵循统一的标注规则:完整可见的甜瓜果实和茎部标注其全部可见区域;当目标被叶片或枝蔓部分遮挡时,仅标注可见部分,不推断遮挡区域;同一画面内包含多个果实时,所有可见果实和茎部均分别标注对应类别;果实与茎部的连接处按可见区域分别标注,不重复标注;边界模糊或低对比度时按可见轮廓一致性进行标注;高光或反光区域不计入任何前景类别,统一标注为背景。

标注完成后进入交叉审核环节:由另一位审核人员逐一核对每张掩膜的标注边界是否贴合可见轮廓、是否存在漏标或误标。审核未通过的样本返回标注阶段重新处理,直至通过。为展示标注对象和标注边界的实际形态,图3给出了 MelonNightColor 中的典型标注样例。

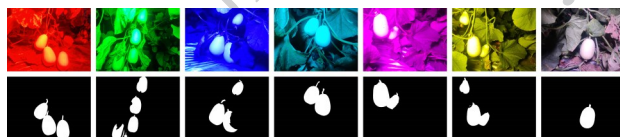


图3 MelonNightColor数据集标注样例
Fig. 3 Annotation examples of the MelonNightColor dataset

由图3可见,甜瓜果实通常具有较完整的椭圆或近圆形可见区域,而茎部区域面积小、形态细长,并常与叶片、枝蔓或支架相邻。因此,数据集既包含相对明确的果实分割目标,也包含边界细碎、易混淆的细小目标类别。

2.6 质量控制方法

为确保数据集在发布前的完整性、一致性和可用性,本研究执行了系统性的质量控制流程,包括文件级检查和内容级检查两个层面,如表4所示。

具体检查步骤为先通过自动化脚本遍历再进行人工抽检。检查中发现异常的样本经人工复核和处

理后,重新纳入数据集或予以剔除,确保最终发布版本的标注数据达到可复现的质量标准。

表4 质量控制检查项
Table 4 Quality control items

检查类型	检查内容	实施方式
文件配对	JPG原图与PNG掩膜是否一一对应	遍历全部11个子数据集的images和masks目录,逐一核对文件名
尺寸一致性	原图与掩膜的像素尺寸是否相同	逐张读取图像尺寸并比对
类别合法性	掩膜中是否仅包含0、1、2三种像素值	统计每张掩膜的唯一像素值集合
空标注检测	掩膜是否全为背景	计算每张掩膜的前景像素占比,标记占比为0的样本
前景占比异常	前景过少或过多	统计每张掩膜的前景像素占比,标记占比明显异常的样本并进行人工复核。
划分泄漏检查	同组7色图像是否跨训练集和测试集	按拍摄位置组号核验所有图像的集合归属
可视化抽检	标注边界是否贴合甜瓜可见轮廓	随机抽取样本,将掩膜叠加至原图进行人工目视复核

3 数据集验证与评估

本实验部分旨在验证MelonNightColor数据集的可用性、挑战性和复用价值,而非提出新的分割模型。围绕这一目标,本文从三个层面开展验证:首先,在11个子数据集上使用多类基础语义分割模型进行基线评测,检验数据集能否支撑主流模型训练和横向比较;其次,比较拍摄阶段多通道同时开启所得子数据集与后期多颜色合并子数据集在当前数据组织与评测协议下的基线表现;最后,基于UNet构建7×7跨颜色泛化矩阵,量化训练光源与测试光源不一致时的性能变化。

3.1 实验设置

本文选取9个常用语义分割模型作为基线模型,包括UNet(Ronneberger等,2015)、PSPNet(Zhao等,2017)、UNet 3+(Huang等,2020)、SwinUNet(Cao等,2022)、VM-UNet(Ruan等,2025)、DenseNet(Huang等,2017)、DenseUNet(Zhou等,2022)、DeepLab(Yurtkulu等,2019)和H-vmunet(Wu等,2025)。这些模型覆盖了经典编码器-解码器结构、金字塔上下文建模、UNet变体、Transformer结构和视觉状态空间模型等不同技术路线,能够从多个角度反映MelonNightColor数据集的分割难度。

所有模型均采用统一的训练设置。输入224×224像素的三通道RGB图像,训练批量大小为4,最大训练轮数为200。优化器采用AdamW,初始学习率为0.001,权重衰减为 5×10^{-4} ;学习率在前20个

epoch进行warm-up,随后在剩余训练阶段采用余弦衰减策略。损失函数由交叉熵项和Dice项组成,Dice损失权重 λ 设置为0.7。随机种子设置为42,每5个epoch在验证集上进行一次评价,并保存验证集mIoU最高的模型用于测试集评价。训练和测试在Linux 5.14系统下完成,实验平台配备Intel Xeon Gold 5418Y处理器和NVIDIA GeForce RTX 4090显卡,图形处理器(graphics processing unit, GPU)显存为24GB,使用Python 3.8.20、PyTorch 2.2.2、计算统一设备架构(Compute Unified Device Architecture, CUDA)11.8和CUDA深度神经网络库(CUDA Deep Neural Network library, cuDNN)8.7.0。

训练阶段对图像和标注同步缩放至224×224像素,并分别以0.5的概率进行水平翻转和垂直翻转;随后转换为张量,并采用各通道均值和标准差均为0.5的方式进行归一化。验证集和测试集仅进行确定性缩放、张量转换和相同的归一化处理,不使用随机几何变换。训练过程中不使用颜色抖动、色相调整、通道置换或合成重着色等颜色增强方法。

所有模型均使用相同的数据划分协议进行训练、验证和测试。对于7个单颜色子数据集,模型仅使用对应颜色的训练集训练,并在同颜色测试集上评估;对于4个组合颜色子数据集,模型使用对应组合子数据集的训练集训练,并在同一组合子数据集的测试集上评估。评价指标包括总体准确率(overall accuracy, OA)、平均交并比(mean intersection over union, mIoU)、平均Dice系数(mean Dice coefficient, mDice)、平均精度均值(mean average pre-

sion, mAP)和每秒帧数(frames per second, FPS)。其中, mIoU 作为语义分割任务的主要评价指标, 用于后续的表格和分析。

由于背景像素在夜间甜瓜图像中占比较高, 单独依赖 OA 容易高估模型性能。本文重点报告 mIoU, 并结合类别 IoU 分析果实和茎部两个前景类别的分割难度。全部基线模型采用统一的数据划分和训练设置; 不同子数据集的结果结合各自的样本构成与评测协议进行解释。

表 5 7 个单颜色子数据集上的基线 mIoU 结果/%

Table 5 Baseline mIoU results on seven single-color subsets/%

数据集	UNet	PSPNet	UNet 3+	SwinUNet	VM-UNet	DenseNet	DenseUNet	DeepLab	H-vmunet
data_red	75.75	67.54	77.33	66.82	76.00	75.21	75.56	74.44	67.75
data_blue	78.44	67.61	78.88	66.94	76.94	76.61	77.94	75.54	60.19
data_green	79.70	69.36	81.37	69.42	79.13	77.72	80.75	76.72	61.07
data_cyan	80.98	69.89	82.87	67.58	79.36	78.57	82.01	77.97	59.66
data_purple	78.13	68.95	79.67	67.03	78.07	75.90	79.83	76.09	59.44
data_yellow	82.53	70.36	83.86	71.98	81.58	79.74	83.27	79.21	68.91
data_white	84.69	70.47	85.32	74.07	82.61	80.53	84.07	79.75	62.02

从不同模型的结果看, UNet 3+、UNet、DenseUNet 和 VM-UNet 在多数补光配置下取得了较高的 mIoU 值。所比较模型在不同补光配置下呈现出相似的数值变化趋势, 但各模型的具体结果仍存在差异。这些结果为 MelonNightColor 上不同模型的基线比较提供参考。

进一步从类别 IoU 统计看, 背景类平均 IoU 达到 99.05%, 甜瓜果实类平均 IoU 为 86.39%, 而甜瓜茎部类平均 IoU 仅为 38.84%。茎部通常呈细长形态, 且容易受到叶片、枝蔓、支架和阴影遮挡影响, 因此是本数据集中最具挑战性的前景类别。这一现象也说明, MelonNightColor 不仅能够支撑果实区域分割, 还可以用于评估模型对细小农业目标的识别能力。

为了更直观地汇总不同人工补光配置下的基线结果, 本文将表 5 中 9 个基线模型的 mIoU 按补光配置求平均, 并绘制平均 mIoU 对比图, 如图 4 所示。

图 4 汇总了 9 个基线模型在 7 个单颜色子集上的平均 mIoU。白光和黄光配置的平均值相对较高, 红光和蓝光配置的平均值相对较低。该图用于展示

3.2 单颜色数据集基线评测

表 5 给出了 9 个基础模型在 7 个单颜色子数据集上的 mIoU 结果。整体来看, 7 种人工补光配置下的测试数值存在差异。在当前数据集的设备、补光与自动成像条件下, 不同补光配置对应的图像表现和模型结果呈现不同的数值表现。以 9 个基础模型的平均 mIoU 计算, 白光和黄光配置下的平均 mIoU 较高, 分别达到 78.17% 和 77.94%; 红光和蓝光配置下相对较低, 分别为 72.93% 和 73.23%。这些数值描述了当前数据集和评测设置下的基线结果。

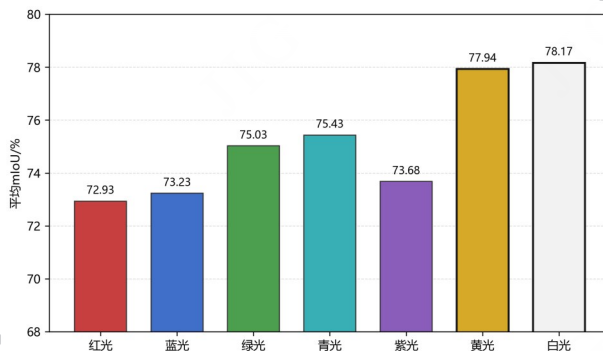


图 4 不同人工补光配置下 9 个基线模型的平均 mIoU 对比

Fig. 4 Average mIoU comparison of nine baseline models under different artificial illumination configurations

当前数据集和评测协议下的数值分布。

3.3 组合颜色数据集评测

表 6 给出了 4 个组合颜色子数据集上的 mIoU 结果。组合颜色子数据集并不对应新的实际补光条件, 而是由多个单颜色子数据集在数据层面合并形成的实验协议, 用于报告相应数据组织方式下的基线表现。

表6 4个组合颜色子数据集上的基线mIoU结果/%

Table 6 Baseline mIoU results on four combined-color subsets/%

数据集	UNet	PSPNet	UNet 3+	SwinUNet	VM-UNet	DenseNet	DenseUNet	DeepLab	H-vmunet
data_red_blue	77.83	68.41	78.70	67.92	76.84	75.49	78.52	74.75	61.21
data_red_green	77.54	69.10	79.56	68.39	77.92	76.54	79.42	75.82	60.83
data_green_blue	78.61	69.03	80.41	69.65	77.90	76.98	80.49	76.50	61.11
data_red_green_blue	78.69	68.54	79.36	69.36	78.31	76.42	80.75	75.52	61.14

绿蓝合并、红绿蓝合并、红绿合并和红蓝合并子集的9个基线模型平均mIoU分别为74.52%、74.23%、73.90%和73.30%。这些数值描述了当前样本构成和评测协议下的模型结果。

为汇总拍摄阶段多通道同时开启所得子集与后期多颜色合并子集在各自评测协议下的结果,本文将红蓝合并与紫色、红绿合并与黄色、绿蓝合并与青色、红绿蓝合并与白色构成4组,并以9个基线模型的平均mIoU进行可视化,如图5所示。

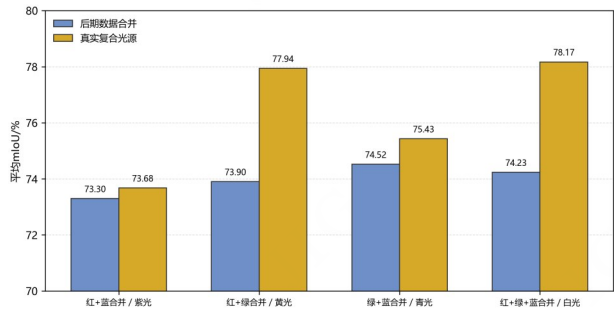


图5 真实复合光源成像与后期多颜色数据合并的基线结果汇总

Fig. 5 Summary of baseline results for real composite illumination imaging and post-hoc multi-color data merging

图5汇总了两类数据组织在各自评测协议下的模型结果。由于两类子数据集的训练样本数量和测试图像构成并不完全一致,图中数值仅用于描述当前实验结果,不构成公平性能比较。

3.4 辅助指标汇总

为避免正文主表过于庞大,表5和表6仅报告语义分割任务中最具代表性的mIoU指标。与此同时,为了更全面地说明不同基线模型在MelonNightColor上的整体表现,本文进一步统计9个模型在11个子数据集上的平均OA、mIoU、mDice、mAP和FPS,结果如表7所示。该表作为辅助指标汇总,用于补充说明模型整体精度、区域重叠质量和推理速度。

表7 9个基础模型在11个子数据集上的辅助指标平均结果

Table 7 Average auxiliary metrics of nine baseline models on eleven subsets

模型	OA/%	mIoU/%	mDice/%	mAP/%	FPS
UNet	99.31	79.35	86.66	88.79	419.87
PSPNet	98.46	69.02	77.97	78.63	161.64
UNet 3+	99.46	80.67	87.51	87.23	209.41
SwinUNet	98.63	69.01	77.58	79.02	101.98
VM-UNet	99.25	78.61	86.13	89.01	58.09
DenseNet	99.20	77.25	84.96	87.24	65.39
DenseUNet	99.45	80.24	87.11	85.34	43.68
DeepLab	99.17	76.57	84.37	86.57	23.60
H-vmunet	98.85	62.12	65.70	74.01	13.25

从表7可以看出,各模型OA均较高,主要原因是背景像素在夜间甜瓜图像中占比较大,因此OA不能充分反映前景类别的分割难度。相比之下,mIoU和mDice对果实与茎部区域的分割质量更加敏感,更适合作为本文主要分析指标。在速度方面,UNet和UNet 3+的FPS较高,而DenseUNet、DeepLab、VM-UNet和H-vmunet的推理速度相对较低。由于本文定位为数据集论文,速度与多指标结果主要作为基线使用,不作为模型优劣排序的核心依据。

3.5 跨颜色泛化矩阵

为量化训练与测试补光配置不一致时的模型表现,本文以UNet为代表模型,构建7个训练颜色与7个测试颜色的跨颜色泛化矩阵。矩阵的行表示训练数据集的补光配置,列表示测试数据集的补光配置;对角线元素表示同配置训练和测试的mIoU,非对角线元素表示配置不一致时的测试mIoU。

记 $mIoU_{ij}$ 为第*i*种补光配置训练、在第*j*种补光配置测试时的mIoU。7个同配置训练与测试组合的

平均 mIoU 由矩阵对角线元素的算术平均得到:

$$\overline{mIoU}_{same} = \frac{1}{7} \sum_{i=1}^7 mIoU_{ii}$$

7×7 矩阵中共有 42 个训练补光配置与测试补光配置不一致的组合, 其跨颜色平均 mIoU 由全部非对角线元素的算术平均得到:

$$\overline{mIoU}_{cross} = \frac{1}{42} \sum_{i \neq j} mIoU_{ij}$$

据此, 同颜色测试平均 mIoU 为 80.38%, 即 7 个对角线元素的算术平均; 跨颜色测试平均 mIoU 为 44.25%, 即 42 个非对角线元素的算术平均。

各训练补光配置与测试补光配置组合对应的 mIoU 结果如表 8 所示。

表 8 UNet 跨颜色泛化 mIoU 矩阵/%
Table 8 Cross-color generalization mIoU matrix of UNet/%

训练颜色/测试颜色	red	blue	green	cyan	purple	yellow	white
red	75.80	33.30	44.70	50.50	31.80	73.30	58.80
blue	52.90	78.40	28.50	71.10	44.60	42.80	73.30
green	43.60	55.90	79.70	59.70	53.00	45.60	62.20
cyan	40.20	58.10	34.70	81.00	39.90	38.80	69.00
purple	51.40	34.00	25.40	33.10	78.10	22.00	72.00
yellow	17.90	33.00	32.80	32.60	36.10	82.50	63.20
white	32.00	31.50	31.50	31.50	37.00	35.50	87.20

由表 8 可知, UNet 在同颜色测试条件下的平均 mIoU 为 80.38%, 而在跨颜色测试条件下的平均 mIoU 为 44.25%。同颜色与跨颜色测试之间存在明显数值差异。该结果显示, 在当前数据集的采集与训练条件下, 训练与测试补光配置不一致时, UNet 的分割性能出现明显下降。其中, yellow 训练、red 测试时 mIoU 为 17.90%, purple 训练、yellow 测试时 mIoU 为 22.00%, 不同配置组合的测试结果存在差异。

同时, 矩阵中也存在若干测试结果相对较好的配置组合。例如 red 训练、yellow 测试的 mIoU 为 73.30%, blue 训练、white 测试的 mIoU 为 73.30%, purple 训练、white 测试的 mIoU 为 72.00%。这些结果可为后续开展跨补光配置模型评测、颜色归一化和鲁棒分割方法研究提供参考。总体而言, 跨颜色泛化矩阵给出了同配置与不同配置测试条件下的 mIoU, 可用于比较模型在不同人工补光配置下的测试表现。

为便于观察跨颜色泛化结果的整体结构, 本文将表 8 中的 mIoU 矩阵进一步绘制为热力图, 如图 6 所示。

由图 6 可见, 对角线区域整体颜色更深, 说明同颜色训练和测试通常能够获得较高性能; 非对角线

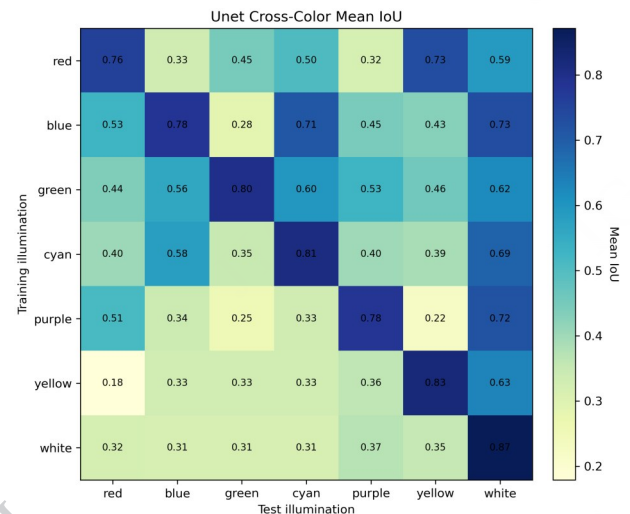


图 6 UNet 跨颜色泛化 mIoU 热力图

Fig. 6 Cross-color generalization mIoU heatmap of UNet

区域普遍较浅, 表明在当前矩阵所覆盖的训练与测试配置中, 补光配置不一致时的 mIoU 通常低于同配置测试结果。热力图还显示, 不同颜色对之间的测试结果并不相同, 部分复合光源组合仍具有较高 mIoU, 可为后续比较不同补光配置下的模型表现提供参考。

3.6 可视化与失败案例分析

基线实验与可视化结果显示, MelonNightColor
© 中国图象图形学报版权所有

中有两方面现象值得关注:一是不同补光配置下的图像表观和模型结果存在差异;二是茎部目标细长、面积小且常被叶片和枝蔓遮挡。在同配置训练和测试时,模型通常能够较完整地分割甜瓜果实主体;在部分补光配置或跨颜色测试条件下,果实边缘容易出现漏分割,茎部区域也更容易被误判为背景。

为了兼顾图像清晰度和论证重点,可视化结果分为同颜色测试和跨颜色测试两部分展示。可视化图均由测试集原始图像、人工标注掩膜和UNet预测掩膜生成。人工标注图中,背景显示为黑色,甜瓜果实显示为洋红色,甜瓜茎部显示为青色;UNet预测图将预测前景以相同颜色半透明叠加到原始图像上,以便同时观察目标位置和分割边界。误差图由人工标注掩膜和预测掩膜逐像素比较得到,其中红色表示漏分割区域,蓝色表示误分割区域,黄色表示前景类别混淆区域,浅灰色表示预测正确区域。该配色避免了黄色果实标记和绿色茎部标记在红光、绿光、黄光图像中与背景混淆的问题。

为展示同颜色训练和测试条件下的典型分割效果,本文选取白光训练并在白光测试、黄光训练并在黄光测试、红光训练并在红光测试3组样本进行可视化,如图7所示。每行依次展示原图、人工标注、预测结果和误差图,用于比较不同补光配置下的同配置测试结果。

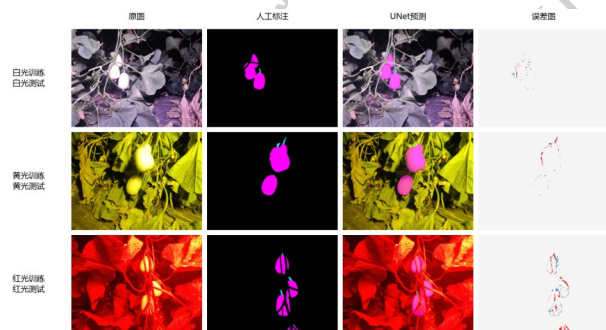


图7 UNet在同颜色测试条件下的可视化分割结果

Fig. 7 Visualization results of UNet under same-color testing conditions

由图7可见,在同颜色测试条件下,UNet对甜瓜果实主体通常能够形成连续且较完整的预测区域,白光和黄光样本的误差区域相对较少;红光样本中果实边缘和茎部附近仍存在一定漏分割或类别混淆。这与单颜色基线结果中白光、黄光表现较优而红光表现相对较差的趋势一致。

为展示跨颜色测试条件下的典型失败情况,本文选取蓝光训练并在绿光测试、黄光训练并在红光测试、紫光训练并在黄光测试3组样本进行可视化,如图8所示。这些组合在跨颜色泛化矩阵中mIoU较低,能够直观展示训练与测试补光配置不一致时的典型分割错误。与图7相同,图8每行依次展示原图、人工标注、预测结果和误差图,重点观察果实边缘漏分割、茎部误判为背景以及前景类别混淆等错误。

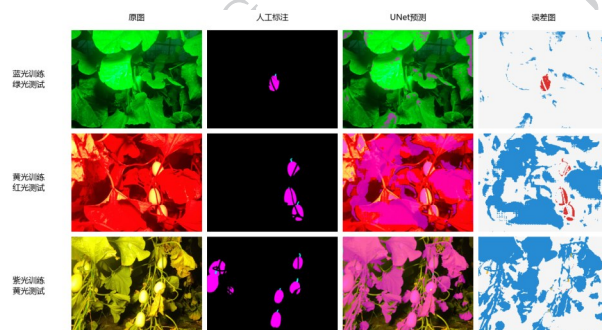


图8 UNet在跨颜色测试条件下的失败案例

Fig. 8 Failure cases of UNet under cross-color testing conditions

由图8可见,在所选跨颜色测试样本中,果实主体边界收缩、局部前景漏检以及茎部区域被预测为背景等错误较为明显。与图7的同颜色结果相比,图8中误差区域增多,说明 MelonNightColor 可用于观察夜间人工补光配置变化下的模型测试表现。

4 数据集使用说明

4.1 数据获取与引用

MelonNightColor 数据集托管于 Science Data Bank 科学数据银行,以 CC BY-NC-ND 4.0 许可公开发布。数据集访问地址为:

<https://cstr.cn/31253.11/sciencedb.j00240.00237>

4.2 使用方法

已标注数据集按照子数据集、训练验证测试集、图像掩膜三级目录结构组织。每个子数据集内部已预划分为 train、val 和 test 三个子目录,研究者可直接按目录路径加载,无需重新划分。以红光子集 (data_red) 为例,其训练集原图位于该子集 train 目录下的 images 子目录,对应标注位于 train 目录下的

masks子目录,原图与标注掩膜通过同名文件一一对应。

不同研究目标可对应选择不同子数据集。若研究目标是评测单一补光配置下的分割结果,可选用相应单颜色子数据集,使用其 train 集训练、val 集验证、test 集测试。若研究目标是开展多配置联合训练,可选用组合颜色子数据集。若研究目标是开展跨配置模型评测,可使用任一单颜色子集的 train 集训练模型,并在其余单颜色子集的 test 集上分别测试。

4.3 应用场景

MelonNightColor 向研究者提供夜间人工补光条件下的甜瓜成组图像和像素级标注,可用于语义分割基线评测、不同补光配置下的图像与模型结果比较以及跨配置模型评测。

夜间甜瓜语义分割基准。数据集提供统一的训练、验证和测试划分,能够用于评测不同语义分割模型在夜间人工光照条件下的表现。由于标注同时覆盖甜瓜果实和甜瓜茎部两个前景类别,研究者不仅可以关注果实主体的分割精度,也可以分析模型对细小、低占比、易遮挡目标的识别能力。

不同补光配置下的成像与模型表现分析。7个单颜色子集在采集时保持拍摄位置、拍摄距离和补光灯方位一致,并在同一拍摄位置切换补光通道配置,可用于比较当前采集条件下不同补光配置对应的图像表现和模型结果。

跨补光配置模型评测。成组采集设计使同一场景具有7种颜色版本,可用于构建训练补光配置与测试补光配置不一致的跨颜色评测协议。模型在不同配置测试中的 mIoU 差异可为颜色归一化、鲁棒分割及相关模型方法研究提供评测数据。

不同数据组织方式下的模型评测。数据集同时包含拍摄阶段多通道同时开启获得的图像子集,以及由多个单颜色子集合并形成的组合子集。两类子集可分别用于报告相应数据组织方式与评测协议下的模型结果。

多配置联合训练。组合颜色子集可用于构建包含多种补光配置图像的训练集,并评价模型在预先定义的测试集上的分割结果。研究者可在统一样本规模和评测协议下进一步研究多配置训练方法。

细小农业目标识别研究。甜瓜茎部面积小、形态细长,且常与叶片、枝蔓和支架相邻,是本数据集中

中最具挑战性的类别之一。因此, MelonNightColor 也可用于评估分割模型对细小目标、低对比度边界和局部遮挡区域的处理能力。

数据质量控制与标注协议复用。数据集提供明确的像素值定义、文件配对规则、按拍摄位置组划分的协议和质量控制流程,可为后续夜间农业图像分割数据集的文件配对、场景组划分、像素值检查和标注审核提供流程参考。

4.4 使用注意事项

使用本数据集时需注意其标注范围、划分方式和采集条件限制。本数据集中的7种补光条件由补光灯通道组合及亮度档位定义。本次采集未对各配置的照度、辐照度或光谱功率进行独立测量,因此相关图像与基线结果用于描述当前人工补光配置和相机成像条件下的表现,不用于区分补光颜色与光照强度的独立作用。

标注类别限制。当前标注仅包括甜瓜果实、甜瓜茎部和背景三类,不支持叶片、枝蔓、病害等其他类别的语义分割。

划分一致性。数据集的训练集、验证集和测试集以拍摄位置组为单位预划分,确保同组7色图像位于同一集合中。研究者不应重新随机打乱划分,否则将引入场景级数据泄漏,导致测试结果虚高。

组合数据集与多通道补光数据集的区别。组合颜色子集由多个单颜色子集在数据层面合并形成,多通道补光子集则来自拍摄阶段相应通道同时开启时采集的图像。使用时应根据研究目的明确训练集和测试集的图像构成,并在比较两类数据组织方式时控制样本数量和测试集条件。

场景泛化限制。本数据集采集自特定地点、特定甜瓜品种和特定生长阶段,使用特定拍摄设备和补光灯。向其他种植环境、其他作物品种或其他生长阶段的泛化能力需在使用前重新评估。

5 结论及讨论

本文构建并公开了面向夜间甜瓜语义分割的 MelonNightColor 数据集。该数据集包含 5474 张已标注图像,覆盖 782 个拍摄位置和 7 种人工补光配置,标注类别包括背景、甜瓜果实和甜瓜茎部。另有 1218 组(8526 张)将逐步标注并公开。成组采集时保持相机位置、拍摄距离和补光灯方位一致,并在同

一拍摄位置依次切换补光通道配置;数据集同时组织为7个单颜色子集和4个组合颜色子集,可为单配置基线评测、多配置联合训练和跨配置模型评测提供数据基础。

综合基线结果表明,在本数据集的特定设备、甜瓜品种、补光设备和自动成像条件下,不同人工补光配置对应的模型分割结果存在差异;白光和黄光配置下的平均mIoU较高,而红光和蓝光配置下的平均mIoU相对较低。跨配置测试中UNet的平均mIoU低于同配置测试,表明当前数据集可用于研究夜间人工补光条件变化下的模型泛化问题。

MelonNightColor采集于特定地点、甜瓜品种、生长阶段、拍摄设备和补光设备条件下,拍摄时曝光时间和ISO采用自动模式。因此,本结果用于描述当前设备和采集条件下不同补光配置对应的图像与模型表现,不作为光谱颜色、光照强度或相机参数独立效应的因果估计。下一轮数据集采集将进一步控制相机参数、补光强度、照射角度和光源距离,并扩大采集条件和标注范围。

参考文献(References)

- Akbar J U M, Kamarulzaman S F, Muzahid A J M, Rahman M A and Uddin M. 2024. A comprehensive review on deep learning assisted computer vision techniques for smart greenhouse agriculture. *IEEE Access*, 12: 4485-4522 [DOI: 10.1109/ACCESS.2024.3349418]
- Bargoti S and Underwood J. 2017. Deep fruit detection in orchards//Proceedings of the 2017 IEEE International Conference on Robotics and Automation (ICRA). Singapore; IEEE: 3626-3633 [DOI: 10.1109/ICRA.2017.7989417]
- Cao H, Wang Y, Chen J, Jiang D, Zhang X, Tian Q, et al. 2022. Swin-UNet: UNet-like pure transformer for medical image segmentation//Karlinsky L, Michaeli T and Nishino K (eds). *Computer Vision - ECCV 2022 Workshops, ECCV 2022. Lecture Notes in Computer Science*, vol.13803. Cham: Springer:205-218 [DOI: 10.1007/978-3-031-25066-8_9]
- Cao H, Zhang G, Zhao A, Wang Q, Zou X and Wang H. 2024. YOLOv8n-CSE: A Model for Detecting Litchi in Nighttime Environments. *Agronomy*, 14 (9) : 1924 [DOI: 10.3390/agronomy14091924]
- Dong L, Zhu L, Zhao B, Wang R, Ni J, Liu S, et al. 2025. Semantic segmentation-based observation pose estimation method for tomato harvesting robots. *Computers and Electronics in Agriculture*, 230: 109895 [DOI: 10.1016/j.compag.2025.109895]
- Ghazal S, Munir A and Qureshi W S. 2024. Computer vision in smart agriculture and precision farming: Techniques and applications. *Artificial Intelligence in Agriculture*, 13: 64-83 [DOI: 10.1016/j.ajaia.2024.06.004]
- Hani N, Roy P and Isler V. 2020. MinneApple: A Benchmark Dataset for Apple Detection and Segmentation. *IEEE Robotics and Automation Letters*, 5(2): 852-858 [DOI: 10.1109/LRA.2020.2965061]
- Huang G, Liu Z, van der Maaten L and Weinberger K Q. 2017. Densely Connected Convolutional Networks//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu: IEEE:4700-4708 [DOI: 10.1109/CVPR.2017.243]
- Huang H, Lin L F, Tong R F, Hu H J, Zhang Q W and Iwamoto Y. 2020. UNet 3+: A Full-Scale Connected UNet for Medical Image Segmentation//Proceedings of the 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). Barcelona: IEEE: 1055-1059 [DOI: 10.1109/ICASSP40776.2020.9053405]
- Jia W K, Wei J M, Zhang Q, Pan N N, Niu Y, Yin X, Ding Y H and Ge X T. 2022. Accurate segmentation of green fruit based on optimized mask RCNN application in complex orchard. *Frontiers in Plant Science*, 13 [DOI: 10.3389/fpls.2022.955256]
- Kamilaris A and Prenafeta-Boldú F X. 2018. Deep learning in agriculture: A survey. *Computers and Electronics in Agriculture*, 147: 70-90 [DOI: 10.1016/j.compag.2018.02.016]
- Kutyrev A, Khort D, Smirnov I and Zubina V. 2025. AI-based orchard monitoring at night: Enhancing sustainable fruit production through real-time apple detection//Proceedings of International Conference on Agritech and Water Management (ICAW). Moscow: EDP Sciences, E3S Web of Conferences, vol.614: 03008 [DOI: 10.1051/e3sconf/202561403008]
- Liang C X, Xiong J T, Zheng Z H, Zhong Z, Li Z H, Chen S M, et al. 2020. A visual detection method for nighttime litchi fruits and fruiting stems. *Computers and Electronics in Agriculture*, 169 [DOI: 10.1016/j.compag.2019.105192]
- Liang Y, Liu Y F, Lin Y S, Jiang W P and Huang Z F. 2024. Real-time citrus picking point localization guided by joint learning network. *Journal of Image and Graphics*, 29(10): 3130-3143 (梁云, 刘云帆, 林毅申, 姜伟鹏, 黄梓帆. 2024. 级联混合模型引导的实时柑橘采摘点定位方法. *中国图象图形学报*, 29(10): 3130-3143) [DOI: 10.11834/jig.230755]
- López-Barrios J D, Escobedo Cabello J A, Gómez-Espinosa A and Montoya-Cavero L E. 2023. Green Sweet Pepper Fruit and Peduncle Detection Using Mask R-CNN in Greenhouses. *Applied Sciences*, 13(10): 6296 [DOI: 10.3390/app13106296]
- Pérez-Borrero I, Marín-Santos D, Gegúndez-Arias M E and Cortés-Ancos E. 2020. A fast and accurate deep learning method for strawberry instance segmentation. *Computers and Electronics in Agriculture*, 178 [DOI: 10.1016/j.compag.2020.105736]
- Ronneberger O, Fischer P and Brox T. 2015. U-Net: Convolutional Networks for Biomedical Image Segmentation//Navab N, Hornegger J,

- Wells W and Frangi A (eds). Medical Image Computing and Computer-Assisted Intervention - MICCAI 2015. Lecture Notes in Computer Science, vol.9351, Cham: Springer:234-241 [DOI: 10.1007/978-3-319-24574-4_28]
- Ruan J C, Li J C and Xiang S C. 2025. VM-UNet: Vision Mamba UNet for Medical Image Segmentation. ACM Transactions on Multimedia Computing, Communications and Applications [DOI: 10.1145/3767748]
- Santos T T, de Souza L L, dos Santos A A and Avila S. 2020. Grape detection, segmentation, and tracking using deep neural networks and three-dimensional association. Computers and Electronics in Agriculture, 170 [DOI: 10.1016/j.compag.2020.105247]
- Su Y M, Xiong J T, Tang K, Huang Q Y, Liao K N, Wu Z X, et al. 2025. Identifying guava and its fruit stem in nighttime environment based on AP-UNet. Journal of Food Composition and Analysis, 140 [DOI: 10.1016/j.jfca.2025.107211]
- Tang Y C, Qiu J J, Zhang Y Q, Wu D X, Cao Y H, Zhao K X, et al. 2023. Optimization strategies of fruit detection to overcome the challenge of unstructured background in field orchard environment: a review. Precision Agriculture, 24: 1183-1219 [DOI: 10.1007/s11119-023-10009-9]
- Tian H K, Wang T H, Liu Y D, Qiao X and Li Y Z. 2020. Computer vision technology in agricultural automation — A review. Information Processing in Agriculture, 7(1): 1-19 [DOI: 10.1016/j.inpa.2019.09.006]
- Waltman J, Buchanan E and Bulanon D M. 2024. Nighttime Harvesting of OrBot (Orchard RoBot). AgriEngineering, 6(2): 1266-1276 [DOI: 10.3390/agriengineering6020072]
- Wei J, Qian Z J, Xu B and Zhao D. 2018. A nighttime image enhancement method based on Retinex and guided filter for object recognition of apple harvesting robot. International Journal of Advanced Robotic Systems, 15: 1-12 [DOI: 10.1177/1729881417753871]
- Wu R K, Liu Y H, Liang P C and Chang Q. 2025. H-vmunet: High-order Vision Mamba UNet for medical image segmentation. Neurocomputing, 624 [DOI: 10.1016/j.neucom.2025.129447]
- Xiong J T, Huo Z W, Huang Q Y, Chen H R, Yang Z G, Huang Y H, et al. 2024. Nighttime citrus detection method combining active light source and improved YOLOv5s model. Journal of South China Agricultural University, 45(1): 97-107 (熊俊涛, 霍钊威, 黄启寅, 陈浩然, 杨振刚, 黄煜华, 苏颖苗. 2024. 结合主动光源和改进 YOLOv5s 模型的夜间柑橘检测方法. 华南农业大学学报, 45(1): 97-107 [DOI: 10.7671/j.issn.1001-411X.202209010])
- Yurtkulu S C, Şahin Y H and Unal G. 2019. Semantic Segmentation with Extended DeepLabv3 Architecture//Proceedings of the 27th Signal Processing and Communications Applications Conference (SIU). Sivas: IEEE:1-4 [DOI: 10.1109/SIU.2019.8806244]
- Zhang G Y, Tian Y, Yin W H and Zheng C. 2024. An Apple Detection and Localization Method for Automated Harvesting under Adverse Light Conditions. Agriculture, 14(3): 485 [DOI: 10.3390/agriculture14030485]
- Zhang H and Yan J. 2024. Low-light image enhancement guided by semantic segmentation and HSV color space. Journal of Image and Graphics, 29(04): 0966-0977 (张航, 颜佳. 2024. 语义分割和 HSV 色彩空间引导的低光照图像增强. 中国图象图形学报, 29(04): 0966-0977 [DOI: 10.11834/jig.230182])
- Zhao H S, Shi J P, Qi X J, Wang X G and Jia J Y. 2017. Pyramid Scene Parsing Network//Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu: IEEE: 6230-6239 [DOI: 10.1109/CVPR.2017.660]
- Zhou Y, Chang H, Lu X and Lu Y. 2022. DenseUNet: Improved image classification method using standard convolution and dense transposed convolution. Knowledge-Based Systems, 254: 109658 [DOI: 10.1016/j.knsys.2022.109658]

作者简介

刘鹏栋,男,湖州师范大学计算机科学与技术专业学生,主要研究方向为人工智能与农业结合。E-mail: pengdong2621@163.com

林翰冉,男,湖州师范大学计算机技术专业学生,主要研究方向为计算机视觉。E-mail: 2023388312@stu.huznu.edu.cn

吴茂念,通讯作者,男,湖州师范大学教授,主要研究方向为人工智能及应用。E-mail: wmn@huznu.edu.cn

郑博,女,湖州师范大学副教授,主要研究方向为人工智能,图像处理,图文识别与推理。E-mail: zhengbo@huznu.edu.cn

朱绍军,男,湖州师范大学副教授,主要研究方向为人工智能在各个领域的应用。E-mail: zhushaojun@huznu.edu.cn

施星仁,男,湖州吴兴金农生态农业发展有限公司负责人。E-mail: jinnongshixingren@gmail.com