

基于U-Net++的部分可见标签辅助监督作物 - 杂草分割

贾劲武, 易云康

内蒙古师范大学地理科学学院, 内蒙古 呼和浩特

收稿日期: 2026年7月20日; 录用日期: 2026年8月13日; 发布日期: 2026年8月26日

摘 要

作物 - 杂草像素级分割是实现变量除草和田间精细管理的重要基础。针对PhenoBench数据集中部分可见作物(partial-crop)和部分可见杂草(partial-weed)在常规三分类合并后可见性信息被弱化的问题, 本文以U-Net++为统一骨干, 比较三分类直接监督(B0)、五分类直接监督(B1)、部分可见像元渐进式重加权(B2)和训练期五分类辅助监督(B3) 4种标签利用策略。四种方法均采用3个随机种子在官方训练集上训练, 并在完整官方验证集上评价。结果表明, B3的平均交并比(mean intersection over union, mIoU)、杂草召回率和部分可见杂草召回率分别为 0.7943 ± 0.0131 、 0.6170 ± 0.0117 和 0.5413 ± 0.0045 。相较B0, B3的杂草召回率提高2.30个百分点, 但杂草精确率下降2.88个百分点, mIoU仅提高0.24个百分点。B1在普通五分类交叉熵(cross-entropy, CE)与Dice联合损失下, 部分可见杂草召回率较低且种子间波动较大。本文结果仅反映验证集上的描述性差异, 总体而言, 训练期五分类辅助监督能够在不改变三分类推理输出的前提下改善部分可见杂草的检出, 但其精确率 - 召回率权衡及跨场景泛化仍需进一步验证。

关键词

PhenoBench, U-Net++, 作物 - 杂草分割, 部分可见标签, 辅助监督

U-Net++-Based Crop-Weed Segmentation with Partial-Visibility Label-Assisted Supervision

Jinwu Jia, Yunkang Yi

College of Geographical Science, Inner Mongolia Normal University, Hohhot Inner Mongolia

Received: July 20, 2026; accepted: August 13, 2026; published: August 26, 2026

Abstract

Pixel-level crop-weed segmentation is fundamental to variable-rate weeding and fine-scale field management. To address the weakened visibility information caused by merging partially visible crop (partial-crop) and partially visible weed (partial-weed) labels into a conventional three-class task in the PhenoBench dataset, this study uses U-Net++ as a unified backbone and compares four label-utilization strategies: direct three-class supervision (B0), direct five-class supervision (B1), progressive reweighting of partially visible pixels (B2), and training-only five-class auxiliary supervision (B3). All four methods are trained on the official training set using three random seeds and evaluated on the complete official validation set. The results show that B3 achieves a mean intersection over union (mIoU), weed recall, and partial-weed recall of 0.7943 ± 0.0131 , 0.6170 ± 0.0117 , and 0.5413 ± 0.0045 , respectively. Compared with B0, B3 increases weed recall by 2.30 percentage points, while weed precision decreases by 2.88 percentage points and mIoU increases by only 0.24 percentage points. Under the five-class cross-entropy (CE) and Dice joint-loss configuration without class-balancing strategies, B1 exhibits lower partial-weed recall and greater variation across random seeds. The reported results represent descriptive differences on the validation set only. Overall, training-only five-class auxiliary supervision can improve the detection of partially visible weeds without changing the three-class inference output, although its precision-recall trade-off and cross-scenario generalization require further validation.

Keywords

PhenoBench, U-Net++, Crop-Weed Segmentation, Partial-Visibility Labels, Auxiliary Supervision

Copyright © 2026 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

作物与杂草的空间分布是精准农业作业决策、杂草防控和变量施药的重要依据。传统人工调查能够获得较可靠的地面信息,但受作业效率和空间连续性限制,难以满足田间快速、连续和精细化制图的需求。无人机遥感平台能够灵活搭载传感器并获取高空间分辨率影像,为精准农业中的田间信息采集提供了重要数据来源[1]。在杂草识别任务中,目标检测可用于定位独立植株或候选区域[2];语义分割则通过逐像素预测表达土壤、作物和杂草的空间范围[3],为杂草面积统计、作业区域划定和变量除草控制提供基础信息。

田间作物-杂草分割面临显著的类别不平衡、尺度差异和类间相似性。根据本文对 PhenoBench 标签的统计,土壤像元占据影像主体,杂草及部分可见杂草所占比例较低;与此同时,作物幼苗与杂草在外部形态和局部特征上具有一定相似性[4][5],进一步增加了两类目标的区分难度。在此情况下,仅依据总体准确率评价模型,容易掩盖少数类别的性能变化。因此,本文除报告整体分割指标外,还重点分析杂草类别的精确率、召回率和 F1 分数,并统计部分可见杂草的主要误分方向,以区分减少漏检与控制误检两类不同的性能变化。

围绕田间杂草识别,已有研究从轻量化分类[4]、半监督语义分割[6]、域内与跨域半监督检测[7]以及基于作物行信息的伪标签构建[8]等方面展开。U-Net 通过编码-解码结构和跳跃连接实现定位信息与语义信息的融合[9],U-Net++进一步采用嵌套跳跃连接聚合不同语义尺度的特征[10]。然而,PhenoBench 数

据集同时提供完整目标与部分可见目标的细粒度标签。若将部分可见作物和部分可见杂草直接合并至普通作物和杂草类别, 可简化最终输出空间, 但会弱化可见性信息; 若将其直接设置为独立类别, 又可能受到极端类别不平衡的影响。因此, 有必要在统一网络和训练条件下比较不同的标签利用方式。

基于上述问题, 本文以 U-Net++ 为统一骨干, 比较 4 种部分可见标签利用策略: 三分类直接监督、五分类直接监督、渐进式损失重加权和训练期五分类辅助监督。研究主要包括三方面内容: 一是在网络主体和基础损失保持一致的条件下开展控制变量比较; 二是依据原始五类标签分别统计完整杂草与部分可见杂草的识别结果; 三是结合混淆矩阵和真实样本预测结果, 分析不同策略对少数类别漏检与误检的影响。

2. 数据集与实验设计

2.1. 数据集与标签设置

本文采用公开的 PhenoBench 数据集[11], 实验使用 v1.1.0 版本。该数据集由无人机在甜菜田自然光照条件下采集, 单幅 RGB 影像大小为 1024×1024 像素。本文使用官方训练集 1407 幅和官方验证集 772 幅; 所有模型均在官方训练集上训练。PhenoBench 原始语义标注包括土壤(soil)、作物(crop)、杂草(weed)、部分可见作物(partial-crop)和部分可见杂草(partial-weed)5 类, 其中部分可见类别指目标可见像元比例低于 50% 的区域[11]。本文将 crop 与 partial-crop 合并为作物, 将 weed 与 partial-weed 合并为杂草, 以构建土壤、作物和杂草三分类任务; B1 在训练和预测阶段保留原始五类输出, 评价时再按同一规则合并为三类。训练集和验证集中, 土壤像元分别占 87.66% 和 89.92%, 杂草像元分别占 0.48% 和 0.51%, partial-weed 像元进一步降至 0.017% 和 0.023%, 表明该任务具有明显的类别不平衡特征。

2.2. 对比实验设置

为比较部分可见标签的不同利用方式, 本文设置 B0、B1、B2 和 B3 四组实验。B0 将原始五类标签合并为三类后直接监督; B1 保留原始五类标签进行训练, 并在评价阶段合并为三类; B2 保持三分类输出, 对部分可见像元实施渐进式损失重加权; B3 设置三分类主分支和五分类辅助分支, 推理阶段仅保留三分类主输出。四组实验均采用交叉熵(cross-entropy, CE)损失与 Dice 损失(Dice loss)的联合形式, 除标签利用方式外, 网络主体、数据增强、优化策略和最大训练轮数保持一致。每组采用随机种子 42、2026 和 3407 重复训练, 结果以算术平均值和样本标准差表示。四组对比实验的具体设置见表 1。

Table 1. Settings of the four comparative experiments

表 1. 四组对比实验设置

实验组	标签处理	监督方式	损失函数	训练输出	推理输出/评价映射
B0	五类合并为三类	三分类直接监督	CE + Dice	三类	三类分类得分
B1	保留原始五类	五分类直接监督	五分类 CE + Dice	五类	1/3→作物; 2/4→杂草, 按三类评价
B2	三类标签; 标记部分可见像元	渐进式重加权	加权 CE + 加权 Dice	三类	三类分类得分
B3	三类主标签 + 原始五类标签	三分类主分支 + 五分类辅助分支	$L_{main} + 0.2L_{aux}$	训练期三类主分支和五类辅助分支	仅三分类主分支; 辅助分支仅训练使用

2.3. 评价指标

本文采用交并比(intersection over union, IoU)、精确率(precision)、召回率(recall)、F1 分数(F1-score)、

平均交并比(mean intersection over union, mIoU)、平均 F1 分数(mean F1-score, mF1)和总体准确率(overall accuracy, OA)评价模型性能。其中, 真正例(true positive, TP)表示真实类别与预测类别均为目标类别的像元数; 假正例(false positive, FP)表示被误判为目标类别的像元数; 假负例(false negative, FN)表示真实属于目标类别但未被正确识别的像元数。mIoU 和 mF1 分别为土壤、作物和杂草三类 IoU 与 F1 分数的算术平均值, OA 表示正确分类像元占全部像元的比例。评价时先在完整验证集上累计像元混淆矩阵, 再统一计算各项指标; 混淆矩阵以真实类别为行、预测类别为列, 类别顺序依次为土壤、作物和杂草。

2.4. 实验环境与实现细节

输入影像保持 1024×1024 像素的原始分辨率, 并将像元值归一化至[0, 1], 不进行均值 - 标准差归一化。训练阶段采用随机水平翻转、垂直翻转、 90° 旋转及亮度 - 对比度扰动进行数据增强, 其中三种几何变换的执行概率均为 0.5, 光度扰动概率为 0.3。几何变换同步作用于影像及对应标签, 光度扰动仅作用于影像。

U-Net++的输入通道数为 3, 基础通道数为 16, 五级特征通道依次为 16、32、64、128 和 256。每个卷积块由两组 3×3 卷积、批量归一化(batch normalization)和修正线性单元(rectified linear unit, ReLU)组成; 下采样采用 2×2 最大池化, 上采样采用双线性插值。模型未启用深监督, 最高分辨率解码特征节点 $x_{0.4}$ 连接 1×1 分类头。B0 和 B2 采用 3 通道输出头, B1 采用 5 通道输出头; B3 在共享特征上同时设置 3 通道主头和 5 通道辅助头。模型采用 AdamW 优化器, 初始学习率和权重衰减均为 1×10^{-4} ; 以验证集 mIoU 为依据自适应降低学习率, 衰减系数为 0.5, 连续 8 轮未改善时触发, 最低学习率为 1×10^{-7} 。批量大小为 1, 梯度累积步数为 4, 最大训练轮数为 80; 早停耐心值为 15, 梯度裁剪阈值为 1.0, 并采用混合精度训练。为降低随机性影响, 各方法采用统一的随机种子和确定性设置, 并在 NVIDIA GeForce RTX 4060 Ti 图形处理器(graphics processing unit, GPU)上完成训练。

3. 部分可见标签的利用策略

3.1. 三分类直接监督 B0

设输入 RGB 影像张量为 $X \in \mathbb{R}^{3 \times H \times W}$, 原始五类标签 y 经映射函数 $M(\cdot)$ 转换为三类标签, 具体映射关系见式(1):

$$M(y) = \begin{cases} 0, & y = 0, \\ 1, & y \in \{1, 3\}, \\ 2, & y \in \{2, 4\}. \end{cases} \quad (1)$$

式(1)中, 标签 0 表示土壤, 标签 1 和 3 分别表示作物与部分可见作物, 统一映射为作物类别; 标签 2 和 4 分别表示杂草与部分可见杂草, 统一映射为杂草类别。B0 输出土壤、作物和杂草三类分类得分, 损失由 CE 损失和多分类 Dice 损失等权相加。CE 损失采用像元均值约简; Dice 损失首先对分类得分执行 Softmax 函数, 并将真实标签转换为独热编码(one-hot encoding), 随后跨批次维度和空间维度逐类统计。平滑项设为 1.0, 土壤在内的全部类别均参与算术平均, 验证批次中暂未出现的类别不予剔除。U-Net++ 通过嵌套跳跃连接聚合不同语义尺度的特征[10], 有利于兼顾区域语义与目标边界。

B0 不区分完整目标与部分可见目标, 其输出空间与最终三分类任务一致, 结构和训练流程较为直接; 但在标签合并后, 部分可见目标包含的可见性信息不再单独参与监督。B1、B2 和 B3 均以 B0 为参照, 仅改变原始部分可见标签的利用方式。

3.2. 五分类直接监督 B1

B1 直接使用土壤、作物、杂草、部分可见作物和部分可见杂草 5 类标签训练 U-Net++, 输出通道数

由 3 扩展为 5, 损失由式(2)给出。评价阶段先对五类预测结果取最大概率类别, 再按照式(1)合并为土壤、作物和杂草三类; 学习率调度、早停和模型选择均依据合并后的验证集 mIoU。除三分类主指标外, 本文还报告 B1 的五分类 mIoU 和 partial-weed IoU, 用于判断模型能否稳定区分完整目标与部分可见目标。

$$L_{B1} = L_{CE}^{(5)} + L_{Dice}^{(5)} \quad (2)$$

B1 旨在保持 U-Net++主体结构及 CE 与 Dice 联合损失不变, 单独考察将 partial-crop 和 partial-weed 设置为独立输出类别的影响。因此, B1 属于控制变量条件下的普通五分类对照, 未进一步引入类别加权或重采样策略。

3.3. 渐进式重加权 B2

B2 保持三分类输出, 并依据原始五类标签中的部分可见像元位置构建权重图。设第 e 轮训练进度 $p = \min(1, e/30)$, partial-crop 像元权重由 0.5 线性增加至 1.0, partial-weed 像元权重由 0.5 线性增加至 1.5, 其余像元权重恒为 1, 两类权重均在第 30 轮达到设定上限。加权 CE 损失先计算逐像元损失, 再按权重求和并除以权重总和; 加权 Dice 损失中, 权重同时作用于交集项以及预测项和真实项。B2 的推理结构与 B0 一致, 仍输出土壤、作物和杂草三类。

3.4. 五分类辅助监督 B3

B3 在 U-Net++最高分辨率解码特征节点 $x_{0.4}$ 上设置三分类主头和五分类辅助头, 两者均为 1×1 卷积, 输出通道分别为 3 和 5。主分支预测土壤、作物和杂草, 辅助分支使用原始五类语义标签监督; 训练阶段同时优化三分类主任务损失 L_{main} 和五分类辅助损失 L_{aux} , 总损失见式(3)。

$$L = L_{main} + 0.2L_{aux} \quad (3)$$

其中, L_{main} 和 L_{aux} 均采用 CE 损失与 Dice 损失等权相加, 辅助损失权重 λ 固定为 0.2。辅助头仅增加 85 个参数, 两项损失共同更新共享特征; 验证和推理阶段只保留三分类主头。本文未开展 λ 的系统敏感性分析, 因此 0.2 仅代表本实验配置, 不能视为最优取值。

B3 并不要求主分支显式输出 partial-weed, 而是利用训练期辅助目标约束共享特征, 使三分类主任务能够保留更多部分可见目标信息。与利用作物行结构构建伪标签的 RoWeeder [8]以及利用未标注数据开展半监督学习的方法[6] [7]不同, B3 不引入额外数据, 而是重新利用 PhenoBench 已有的细粒度语义标注。

4. 实验结果与讨论

4.1. 总体与杂草类别性能

四种方法在官方验证集上的三分类指标见表 2。B3 的 mIoU、mF1 和 OA 分别为 0.7943 ± 0.0131 、 0.8632 ± 0.0114 和 0.9915 ± 0.0005 , 平均值均为四组实验中最高。B0 的 mIoU 为 0.7919 ± 0.0199 , 与 B3 相差 0.24 个百分点; 该差值小于两种方法各自的随机种子标准差, 因此仅能解释为平均值上的小幅变化, 不作统计显著性判断。B2 和 B1 的 mIoU 分别为 0.7836 ± 0.0150 和 0.7726 ± 0.0181 , 说明在当前训练配置下, 渐进式重加权和普通五分类直接监督均未超过三分类基线。

杂草类别的差异更为明显。B3 的杂草召回率为 0.6170 ± 0.0117 , 高于 B0、B1 和 B2 的 0.5940 ± 0.0742 、 0.5448 ± 0.0839 和 0.5585 ± 0.0639 ; 杂草 F1 分数为 0.6294 ± 0.0311 。与此同时, B3 的杂草精确率为 0.6434 ± 0.0546 , 低于 B0 的 0.6722 ± 0.0884 。由此可见, B3 的主要作用表现为减少杂草漏检, 而非同时提高精确率和召回率。B1 的杂草 IoU 和 F1 分数最低, 表明普通五分类 CE 与 Dice 联合损失在合并后的三分类

任务上表现较弱。

Table 2. Overall and weed-class segmentation metrics of different experimental groups

表 2. 不同实验组总体与杂草类别分割指标

指标类别	指标	B0	B1	B2	B3
总体	mIoU	0.7919 ± 0.0199	0.7726 ± 0.0181	0.7836 ± 0.0150	0.7943 ± 0.0131
总体	mF1	0.8613 ± 0.0182	0.8434 ± 0.0173	0.8536 ± 0.0139	0.8632 ± 0.0114
总体	OA	0.9911 ± 0.0005	0.9905 ± 0.0006	0.9911 ± 0.0006	0.9915 ± 0.0005
杂草类	IoU	0.4574 ± 0.0583	0.4048 ± 0.0479	0.4327 ± 0.0419	0.4597 ± 0.0335
杂草类	精确率	0.6722 ± 0.0884	0.6152 ± 0.0219	0.6631 ± 0.0646	0.6434 ± 0.0546
杂草类	召回率	0.5940 ± 0.0742	0.5448 ± 0.0839	0.5585 ± 0.0639	0.6170 ± 0.0117
杂草类	F1 分数	0.6263 ± 0.0538	0.5752 ± 0.0485	0.6032 ± 0.0402	0.6294 ± 0.0311

4.2. 部分可见杂草区域诊断

部分可见杂草区域诊断结果见表 3。该表分别统计完整杂草和部分可见杂草区域的召回率, 并给出 partial-weed 误分为作物和土壤的比例。B3 的部分可见杂草召回率为 0.5413 ± 0.0045 , 较 B0 提高 5.56 个百分点, 较 B1 提高 19.22 个百分点; 其 partial-weed 误分为作物和土壤的比例分别为 0.2872 ± 0.0174 和 0.1716 ± 0.0174 , 均低于 B0 和 B1。B1 的部分可见杂草召回率为 0.3491 ± 0.1340 , 且随机种子间波动较大。

Table 3. Diagnostic results for partially visible weed regions

表 3. 部分可见杂草区域诊断结果

指标	B0	B1	B2	B3
完整杂草召回率	0.5990 ± 0.0781	0.5538 ± 0.0825	0.5632 ± 0.0650	0.6204 ± 0.0122
部分可见杂草召回率	0.4857 ± 0.0657	0.3491 ± 0.1340	0.4579 ± 0.0411	0.5413 ± 0.0045
部分可见杂草→作物	0.3194 ± 0.0485	0.4361 ± 0.0722	0.3425 ± 0.0371	0.2872 ± 0.0174
部分可见杂草→土壤	0.1948 ± 0.0181	0.2148 ± 0.0623	0.1996 ± 0.0082	0.1716 ± 0.0174

4.3. B1 五分类学习表现

B1 的五分类诊断结果为: 五分类 mIoU = 0.5252 ± 0.0339 、五分类宏平均 F1 分数(macro-averaged F1-score) 0.6191 ± 0.0459 、五分类 OA = 0.9791 ± 0.0011 、partial-crop IoU = 0.3827 ± 0.0367 、partial-weed IoU = 0.0491 ± 0.0851 。该结果仅用于解释 B1 对原始五类标签的学习情况, 不与其他方法的三分类主指标直接比较。partial-weed IoU 在随机种子 2026 和 3407 下均为 0, 说明在当前普通五分类 CE 与 Dice 联合损失配置中, 极少量 partial-weed 类别的学习稳定性不足。类别极端不平衡可能是原因之一, 但现有实验不能确定唯一成因, 也不能外推至采用类别平衡损失或采样策略的其他五分类方案。

4.4. 混淆矩阵分析

四种方法在官方验证集上的行归一化三分类混淆矩阵如图 1 所示。该图由每种方法 3 个随机种子的最佳模型在完整验证集上重新评价所得的混淆矩阵按像元计数求和, 并进行行归一化。B0、B1、B2 和 B3 对真实杂草的正确识别比例分别为 0.594、0.545、0.559 和 0.617。B1 将部分可见标签设置为独立类别后,

合并三分类的杂草召回率仍低于 B0 与 B3; B3 的杂草误分为作物的比例为 0.182, 低于 B0 的 0.193 和 B1 的 0.247。由于四种方法的土壤正确识别比例均接近 0.997, 仅依据 OA 难以反映少数类别的性能差异。

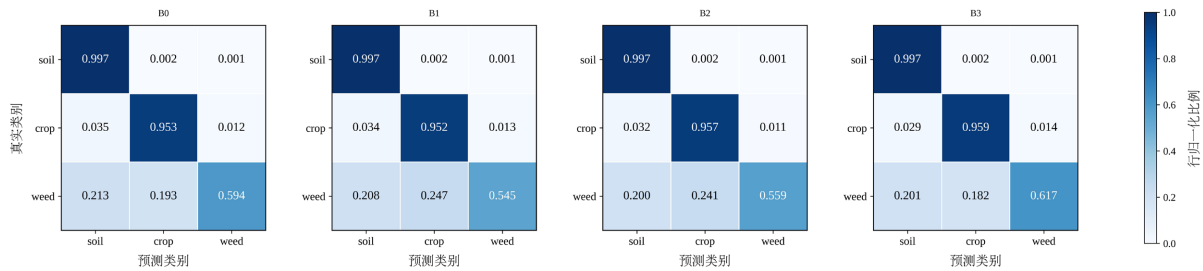


Figure 1. Row-normalized three-class confusion matrices of B0, B1, B2, and B3 on the PhenoBench official validation split
图 1. B0、B1、B2 和 B3 在 PhenoBench 官方验证集上的行归一化三分类混淆矩阵

4.5. 定性结果分析

不同部分可见场景下的定性分割结果如图 2 所示。图中选取 3 组真实样本, 自上而下分别对应连续部分可见杂草、作物邻近或遮挡场景以及分散小目标。B0、B1 和 B2 在不同程度上出现杂草漏分或向作物类别误分, B3 的预测区域整体更接近真实标签。第一组样本中, B3 对连续部分可见杂草区域的恢复更完整; 第二组样本表明, 在作物与杂草空间邻近或局部遮挡时, 辅助监督有助于减轻类别混淆; 第三组样本中, B3 对分散小目标的响应更接近真实标签。

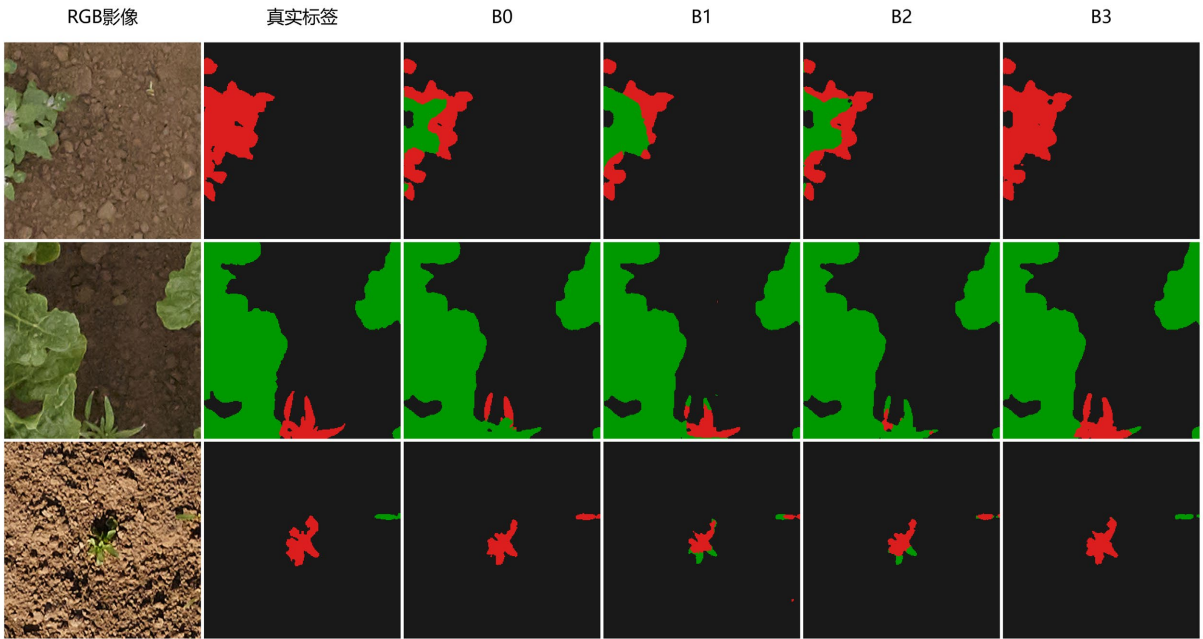


Figure 2. Qualitative comparison of real B0~B3 segmentation results under different partial-visibility scenarios
图 2. 不同部分可见场景下 B0~B3 的真实分割结果比较

4.6. 讨论与局限

四组对比实验表明, 不同部分可见标签利用策略对少数类别的识别效果存在差异。B1 直接扩展为五分类输出, 在未采用类别平衡策略的 CE 与 Dice 联合损失配置下, 其三分类 mIoU 和部分可见杂草召回

率均低于 B0, 且两个随机种子下部分可见杂草类别的 IoU 均为 0, 说明仅扩展输出类别难以充分发挥细粒度标签的监督作用。B2 虽然也利用了部分可见标签, 但整体性能未呈现与 B3 相同的改善趋势。相比之下, B3 保留五类标签作为辅助监督, 在不改变三分类推理结构的条件下取得了最高的部分可见杂草召回率, 表明辅助监督能够更充分地利用部分可见标签信息。

B3 与 B0 的差异主要体现在共享特征学习过程。B0 将部分可见作物和部分可见杂草分别合并至普通类别, 网络仅学习土壤、作物和杂草三类目标之间的差异。B3 则通过三分类主任务与五分类辅助任务共同约束共享特征提取, 使网络在完成三类目标判别的同时, 进一步学习完整目标与部分可见目标之间的差异。由此, 目标边缘、局部遮挡和弱纹理等细粒度信息能够持续参与模型优化, 减少标签合并造成的细粒度差异信息丢失, 为作物与杂草的判别提供更多依据。

实验结果显示, B3 的杂草召回率较 B0 提高 2.30 个百分点, 精确率下降 2.88 个百分点。混淆矩阵结果表明, 真实杂草的正确识别比例由 0.594 提高至 0.617, 杂草误分为作物的比例由 0.193 降至 0.182, 说明辅助监督在一定程度上减少了杂草漏检。然而, 部分可见杂草与作物边缘、阴影及弱纹理植被之间具有一定的特征相似性。当模型对这些局部特征更加敏感时, 类别边界附近的非杂草像素也更容易被识别为杂草, 从而增加假阳性。综合上述结果推测, 辅助监督可通过强化弱杂草相关特征的共享表达, 提高模型对少数类别的检出能力, 但同时也导致增强模型对部分模糊区域的响应, 进而造成召回率提高而精确率下降。

B3 相较 B0 的 mIoU 仅提高 0.24 个百分点, 小于随机种子间的波动范围, 说明在类别不均衡条件下, mIoU 对少数类别性能变化的敏感性有限。因此, 本文结合精确率、召回率、F1 分数、部分可见杂草召回率及混淆矩阵, 对不同策略进行综合评价。由于辅助监督仅作用于训练阶段, 不改变最终推理结构, 因此不会增加模型部署阶段的输出复杂度, 在漏检代价较高的田间杂草识别场景中具有一定应用价值。但在误检较为敏感的精准除草任务中, 仍需结合类别平衡、阈值调整或后处理方法控制假阳性。该策略在其他网络结构和数据集上的适用性仍需进一步验证。

本文实验仅基于 PhenoBench 数据集和 U-Net++ 网络开展, 每组实验进行了 3 次随机重复训练, 尚未进行统计显著性检验, 也未系统分析类别平衡策略、辅助损失权重和决策阈值的影响, 因此后续研究将结合不同数据集和网络结构开展跨场景验证, 并通过特征响应可视化和类别特征分布分析, 进一步揭示辅助监督对共享特征学习及分类边界的影响。

5. 结论

本文在统一 U-Net++ 框架下比较了 4 种部分可见标签利用方式。B3 在总体 mIoU、杂草召回率和部分可见杂草召回率上取得最高平均值, B0 则具有更高的杂草精确率, 表明训练期辅助监督带来了精确率 - 召回率之间的权衡。在未采用类别平衡策略的五分类 CE 与 Dice 联合损失配置下, B1 对极少量部分可见杂草类别的学习稳定性不足, B2 的渐进式重加权也未超过三分类基线。结果提示, 在当前数据分布和训练配置下, 训练期五分类辅助监督有助于在保持三分类推理输出的同时利用原始五类语义信息, 但其相对 B0 的总体增益有限。上述结论仅适用于当前 PhenoBench 官方验证协议和 U-Net++ 配置, 仍需通过独立测试、跨场景验证、类别平衡策略比较及辅助损失权重敏感性分析进一步检验。

参考文献

- [1] Zhang, Z. and Zhu, L. (2023) A Review on Unmanned Aerial Vehicle Remote Sensing: Platforms, Sensors, Data Processing Methods, and Applications. *Drones*, 7, Article 398. <https://doi.org/10.3390/drones7060398>
- [2] Rahman, A., Lu, Y. and Wang, H. (2023) Performance Evaluation of Deep Learning Object Detectors for Weed Detection for Cotton. *Smart Agricultural Technology*, 3, Article ID: 100126. <https://doi.org/10.1016/j.atech.2022.100126>

-
- [3] Long, J., Shelhamer, E. and Darrell, T. (2015) Fully Convolutional Networks for Semantic Segmentation. 2015 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Boston, 7-12 June 2015, 3431-3440. <https://doi.org/10.1109/cvpr.2015.7298965>
 - [4] 李亚, 陈晓东, 王海瑞, 等. 基于优化MobileViT模型的轻量化田间杂草识别[J]. 华中农业大学学报, 2025, 44(4): 192-203.
 - [5] Ilyas, T., Arsa, D.M.S., Ahmad, K., Lee, J., Won, O., Lee, H., *et al.* (2025) CWD30: A New Benchmark Dataset for Crop Weed Recognition in Precision Agriculture. *Computers and Electronics in Agriculture*, **229**, Article ID: 109737. <https://doi.org/10.1016/j.compag.2024.109737>
 - [6] 蔡雨霖, 肖佳仪, 余超然, 等. 基于 UANP-MT 的半监督菜心杂草分割方法[J]. 农业工程学报, 2023, 39(11): 183-191.
 - [7] Deng, B., Lu, Y. and Brainard, D. (2025) Semi-Supervised Weed Detection in Vegetable Fields: In-Domain and Cross-Domain Experiments. *IFAC-PapersOnLine*, **59**, 96-101. <https://doi.org/10.1016/j.ifacol.2025.11.770>
 - [8] De Marinis, P., Vessio, G. and Castellano, G. (2025) RoWeeder: Unsupervised Weed Mapping through Crop-Row Detection. In: Del Bue, A., Canton, C., Pont-Tuset, J. and Tommasi, T., Eds., *Computer Vision—ECCV 2024 Workshops*, Springer, 132-145. https://doi.org/10.1007/978-3-031-91835-3_9
 - [9] Ronneberger, O., Fischer, P. and Brox, T. (2015) U-Net: Convolutional Networks for Biomedical Image Segmentation. In: Navab, N., Hornegger, J., Wells, W. and Frangi, A., Eds., *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2015*, Springer, 234-241. https://doi.org/10.1007/978-3-319-24574-4_28
 - [10] Zhou, Z., Siddiquee, M.M.R., Tajbakhsh, N. and Liang, J. (2020) UNet++: Redesigning Skip Connections to Exploit Multiscale Features in Image Segmentation. *IEEE Transactions on Medical Imaging*, **39**, 1856-1867. <https://doi.org/10.1109/tmi.2019.2959609>
 - [11] Weyler, J., Magistri, F., Marks, E., Chong, Y.L., Sodano, M., Roggiolani, G., *et al.* (2024) PhenoBench: A Large Dataset and Benchmarks for Semantic Image Interpretation in the Agricultural Domain. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **46**, 9583-9594. <https://doi.org/10.1109/tpami.2024.3419548>