doi:10.6041/j.issn.1000-1298.2025.11.042

基于无人机遥感图像的茶叶枯病病斑半监督检测方法

江永成[1]　吕蕴娴[1]　胡根生[2]

(1.安徽大学电气工程与自动化学院,合肥 230601;
2.安徽大学农业生态大数据分析与应用技术国家地方联合工程研究中心,合肥 230601)

摘要:从茶树低分辨率图像中准确检测茶叶枯病是一项重要且具有挑战性的任务。由于无人机遥感图像中茶叶枯病病斑密集、病斑面积较小且边缘模糊,现有方法检测精度较低,且通常依赖全监督学习,需要大量的专业标注,耗时耗力。因此,提出了一种基于无人机遥感图像的茶叶枯病病斑模糊密集小目标半监督检测方法(SSTLBdet)。该方法通过迭代标注,充分利用有限的标签信息,筛选出更多具有高不确定性、视觉信息丰富和大小分布多样化的目标进行检测,且引入加权联合置信度估计和自适应样本区分,有效增强了对密集、模糊和小病斑的处理能力。结果表明,SSTLBdet 在不同标注数据比例下的 mAP 分别达到 51.28%(1%标注数据)、59.85%(5%标注数据)、75.09%(10%标注数据)、77.99%(20%标注数据)和 78.12%(30%标注数据)。与 FCOS(Fully convolutional one-stage object detection)模型相比,SSTLBdet 在测试数据上的 mAP 提高 32.70 个百分点。其 mAP@0.5 和 F1 分数达到 78.12%和 71.24%,也显著高于其他检测方法。所提方法显著提高了对茶叶枯病病斑的检测精度,同时大幅减少了数据标注工作量。

关键词:茶叶枯病;深度学习;半监督;目标检测;无人机遥感

中图分类号:S127　　**文献标识码**:A　　**文章编号**:1000-1298(2025)11-0441-12　　**OSID**:

Semi-supervised Object Detection Method for Spots of Tea Leaf Blight Based on UAV Remote Sensing Images

JIANG Yongcheng[1]　LÜ Yunxian[1]　HU Gensheng[2]

(1. *School of Electrical Engineering and Automation*, *Anhui University*, *Hefei* 230601, *China*
2. *National Engineering Research Center for Agro-Ecological Big Data Analysis and Application*, *Anhui University*, *Hefei* 230601, *China*)

Abstract: Accurate detection of tea leaf blight (TLB) from low-resolution images of tea trees is an important and challenging task. Due to the dense blight spots, small areas, and blurred edges of TLB in unmanned aerial vehicle (UAV) remote sensing images, existing methods exhibited low detection accuracy. Moreover, current TLB detection methods typically relied on fully supervised learning, which required a large amount of expert annotation, making it time-consuming and labor-intensive. To address this, a semi-supervised target detection method of semi-supervised object detection method for spots of tea leaf blight in UAV remote sensing images (SSTLBdet) was proposed for detecting small, dense, and blurred TLB spots in UAV remote sensing images. This method utilized limited labeled information through iterative annotation to screen out more targets with high uncertainty, rich visual information, and diverse size distributions for detection, additionally, it introduced weighted joint confidence estimation (WJCE) and adaptive sample selection (ASS), effectively enhancing the ability to handle dense, ambiguous, and small lesions. Experimental results showed that SSTLBdet achieved mAP values of 51.28% (1% labeled data), 59.85% (5% labeled data), 75.09% (10% labeled data), 77.99% (20% labeled data), and 78.12% (30% labeled data). Compared with the FCOS model, SSTLBdet improved the mAP value on the test data by 32.70 percentage points. In terms of mAP@0.5 and F1 score, SSTLBdet achieved 78.12% and 71.24%, respectively, significantly outperforming other

收稿日期:2024-07-17　修回日期:2024-08-19
基金项目:国家自然科学基金项目(32372632)和安徽省高等学校自然科学研究重大项目(KJ2020ZD03)
作者简介:江永成(1971—),男,副教授,博士,主要从事自动控制与目标检测算法研究,E-mail:jiang_hugo@ahu.edu.cn
通信作者:胡根生(1971—),男,教授,博士,主要从事机器学习、图像处理与计算机视觉研究,E-mail:hugs2906@sina.com

detection methods. The proposed method significantly enhanced the detection accuracy of tea leaf blight spots while significantly reducing the workload of data annotation.

Key words: tea leaf blight; deep learning; semi-supervised; object detection; unmanned aerial vehicle remote

0 引言

茶树在生长过程中可能会感染一些病害,严重影响茶叶品质[1]。茶叶枯病是茶树中最常见的病害之一,广泛分布于全国各茶区[2]。当茶树感染茶叶枯病,起初会表现为叶片黄褐色和水浸病斑,病程持续发展,叶子会形成不均匀的浅棕色和灰色斑块,最终导致叶片受损,甚至整棵茶树死亡[3]。

目前,大多数茶叶枯病检测由经验丰富的茶农或植保专家现场进行,但由于很多茶园地处偏远,分布面积较大,加之传统的人工检测方式工作量大、时间周期长且易遗漏,难以全面地掌握大面积茶园中病害茶树的空间分布情况,以致错过采取防治措施的最佳时机,造成病情快速扩散[4]。

遥感工具和方法已被广泛应用于作物病虫害检测中。将深度学习与无人机遥感技术相结合,为作物病虫害检测开辟新的研究方向。KERKECH 等[5]利用卷积神经网络构建葡萄病害检测模型,该模型提取 HSV、LAB、YUV 颜色和纹理特征。检测结果表明,YUV 色彩空间与过量绿减过量红指数植被指数相结合的检测效果最好。YOLO 系列[6-7]、RetinaNet[8]、SSD[9]等也被广泛应用于作物病虫害检测中。文献[10-14]采用全部标注的训练数据训练深度学习模型,但由于标注数据集较少,造成模型检测精度低。在此情况下,如何充分利用大量的未标注数据和少量的有标注数据进行基于深度学习的植物病害检测,是一个亟待解决的问题。

近年来出现的半监督学习方法使用少量标注数据和大量未标注数据来训练网络模型,提高模型性能[15]。XU 等[16]利用基于半监督学习框架 Noisy Student 的 SOLOv2 模型来准确有效地检测和对玉米幼苗叶片计数。NONG 等[17]引入 Hard Example Maining(OHEM)对改进标注数据训练,解决作物和杂草之间的相似性和重叠问题。NAJAFIAN 等[18]提出一种半监督的小麦抽穗检测方法,该方法基于少量已标注图像和短视频来模拟标注数据集,缓解分布偏移的问题。这些半监督学习方法显著减少与标注相关的耗费,提高农作物病害检测的效率。现有的半监督学习方法用于无人机图像中的茶叶枯病病斑检测,存在如下问题:①无人机遥感图像中茶叶分布密集(茶叶数量多,位置近,茶叶枯病病斑形状相似),难以满足分类和定位的一致性,神经网络对这种密集问题易产生漏检和误检。②无人机遥感图像覆盖的空间面积往往较大,图像分辨率不足,图像中病斑边界模糊,不易区别前景和背景,造成正负样本分类和定位不准确。③茶叶枯病初期病斑较小,现有神经网络模型一般难以检测到小目标,因此会出现漏检问题。

为解决上述问题,本文提出一个半监督目标检测方法 SSTLBdet:对标注方法进行迭代优化改进;对置信度评估进行加权联合改进;对样本划分进行自适应阈值改进。以期提升检测精度与泛化能力,实现对密集模糊小目标的精准检测。

1 材料与方法

1.1 图像采集

无人机遥感图像和地面图像采集自安徽省芜湖市天井山茶园(东经 117°34′~117°39′、北纬 31°15′~31°17′,海拔 179 m)。数据采集区域如图 1 所示。为了尽量减少风、阴影、强光、弱光和反射对图像质量的影响,采集时间为 2021 年 2 月 19 日 14:00—17:00 和 2024 年 3 月 22 日 14:00—17:00。所用无人机型号为 DJI 147 Mavic Air2,相机型号为 DJI FC3170,该相机可拍摄 4 800 万像素的可见光图像,光圈值为 f/2.8,曝光时间为 1/60 s,ISO 感光度为 ISO 100,焦距为 4 mm,最大续航时间为 34 min,最长续航里程为 18.5 km,最大飞行速度为 19 m/s,抗风等级 5 级。拍摄图像为 JPEG 格式,分辨率为 8 000 像素×6 000 像素。飞行过程中,无人机飞行速度和方向由人工手动控制,飞行高度为距离茶叶冠层 5 m,且摄像机角度保持垂直于地面。无人机具有 GPS 定位

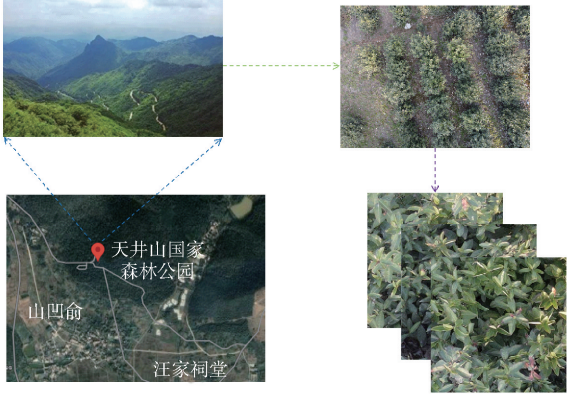


图 1 数据获取区域

Fig. 1 Data acquisition area

能力，能够精确记录每幅图像位置和高度。

1.2　图像预处理

采集到的无人机图像覆盖空间较广，直接将其用于训练和测试会导致计算量过大、模型预测效率低，因此本文先对图像进行裁剪。随机选取其中 40 幅无人机遥感图像，将其分别放大至原始图像尺寸 2 倍和缩小至原始尺寸一半，然后围绕图像中心点随机选取 4 个方向上距离中心点 200 ~ 400 像素的点，以这 4 个点形成的最大边界矩形作为图像的裁剪区域。所选图像中部分存在阴影或过曝情况，这些因素会在数据中引入噪声和不一致性，使模型难以从有限的有标签数据中学习到可靠特征，并影响对无标签数据的有效利用。因此，扩展数据集后，经过图像归一化处理，剔除一些阴影和过曝严重图像，最终获得分辨率为 640 像素 ×640 像素的高清图像 800 幅。所有图像由国家农业生态大数据分析与应用工程研究中心的专家使用图像标注工具 LabelImg 软件进行标注。对于筛选后图像中仍存在的部分阴影及过曝问题，选择 Retinex 算法[19]来增强对比度，有效减轻了图像中阴影和过曝现象的不利影响。

茶叶枯病图像以比例 7∶1随机分为训练集和测试集，其中训练集 700 幅，测试集 100 幅。为了增强训练样本的丰富度和多样性，需要对样本进行数据增广。考虑到使用的数据集较小以及茶园环境特殊性[20]，采用平移、旋转 30°、翻转、缩放、添加噪声和 mosaic 增强等数据增广方法。

1.3　半监督目标检测模型 SSTLBdet

提出的 SSTLBdet 模型基于基本的教师-学生半监督检测网络，其核心构造主要包括迭代标注、加权联合置信度估计、自适应样本区分模块。SSTLBdet 遵从伪标签框架，可以与 FCOS[21] 单阶段检测器兼容。

受 STAC(自训练和一致正则化驱动的增强策略)[22]框架启发，方法包括 2 个阶段：burn-in 阶段和 self-training 阶段。如图 2 所示，在 burn-in 阶段，将有标注的 TLB 图像输入 FCOS 模型，用于初始化教师模型和学生模型。burn-in 阶段完成后，进入如图 3 所示的核心迭代阶段，即 self-training 阶段。在每次迭代中，通过 metric score 来动态选择最重要的未标记茶叶样本进行标注，逐步提高病斑伪标签的质量和数量。该方法优先利用模型早期训练阶段生成的伪标签，对含有密集、模糊和较小病斑的茶叶图像进行标注，从而最大化利用有限标签信息。标注后的 TLB 图像在输入教师模型后获得预测得分，WJCE 模块结合每个病斑预测的分类得分和定位得分，通过加权方式生成伪标签的加权置信度，此模块解决了对 TLB 预测时分类定位不一致的问题，有效减少了在病斑密集时伪标签的误差，并通过合理加权提高了对病斑检测的精度。经 WJCE 筛选后的伪标签，在 ASS 模块中自适应地划分为正样本、模糊区域和负样本阈值。其中，正样本表示需要预测的病斑；负样本表示地面、杂草或正常茶叶等背景；模糊区域可能含有潜在的正样本(即病斑)，需对其进行进一步分类和定位挖掘。通过上述步骤完成对图像中 TLB 的检测。模型通过计算损失不断迭代优化，最终达到稳定状态，输出对整个 TLB 数据集的检测结果。

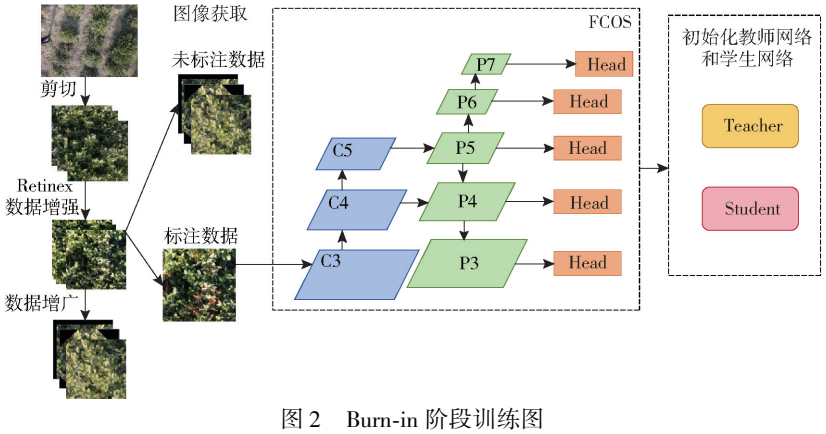


图 2　Burn-in 阶段训练图

Fig. 2　Burn-in phase training diagram

1.3.1　基本教师-学生半监督检测网络

与前人工作[22-23]类似，本文半监督模型同样采用经典的教师-学生检测网络架构。使用 FCOS 作为基线检测网络。教师网络和学生网络通常具有相同的体系结构，训练流程如下：首先，教师网络在有标签数据上进行监督学习，以学习任务的基本特征和知识。该阶段的学习结果形成初始模型，并将其权重同步到教师模型和学生模型中；然后将无标签数据输入教师网络中进行预测，预测结果进行加权联合置信度筛选后，作为伪标签；最后使用有标签数

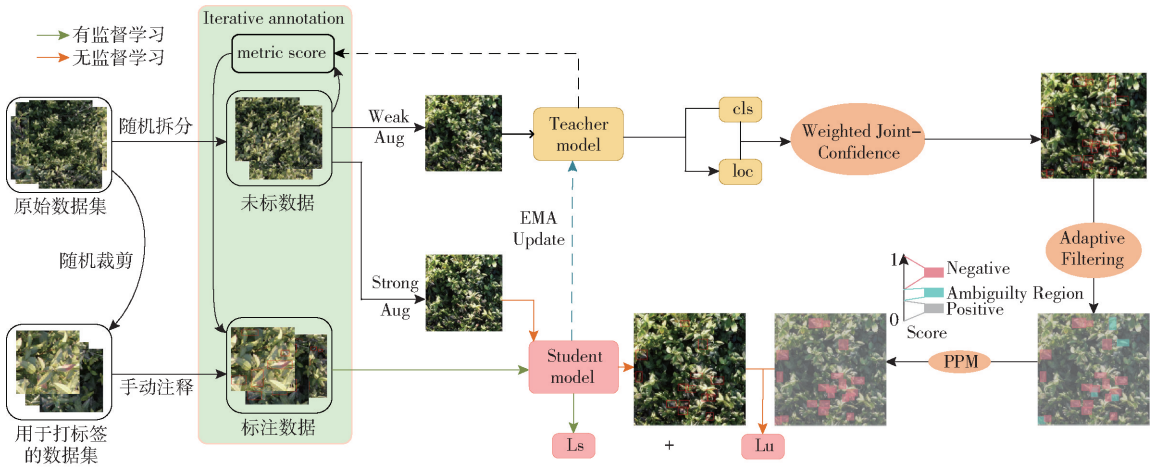


图 3 SSTLBdet 结构

Fig. 3 Structure of SSTLBdet

据和上述生成的伪标签数据共同训练学生网络。学生网络优化目标为最小化其损失函数 L,即

$$L = L_{sup} + \lambda L_{unsup} \tag{1}$$

式中 L_{sup}——有监督损失

L_{unsup}——无监督损失

λ——控制未标注数据贡献的超参数

输入有标签数据集 $D_L = \{X_L, Y_L\}$ 和无标签数据集 $D_U = \{X_U\}$,其中 X 为原图,Y 为打好的标签数据。有标签数据部分损失 L_{sup} 为分类损失 L_{cls} 和回归损失 L_{loc},即

$$L_{sup} = \frac{1}{N_l}\sum_{i=1}^{N_l}\left(L_{cls}(x_l^i, y_{cls}^i) + L_{loc}(x_l^i, y_{cls}^i)\right) \tag{2}$$

其中

$$L_{cls}(x_l^i, y_{cls}^i) = L_{cls}^{rpn}(x_l^i, y_{cls}^i) + L_{cls}^{roi}(x_l^i, y_{cls}^i) \tag{3}$$

$$L_{loc}(x_l^i, y_{loc}^i) = \sum_{c\in(x,y,h,w)} Smooth_{L1}(t_c^i - y_c^i) \tag{4}$$

式中 x_l——振动幅度

y_{cls}、y_{loc}——图像标签信息

t_c^i——图像 X_i 第 c 个坐标

y_c^i——图像 Y_i 第 c 个坐标

边界盒回归损失采用 $L-1$ 损失,即

$$Smooth_{L1}(x) = \begin{cases} 0.5x^2 & (|x|<1) \\ |x| - 0.5 & (|x|\geqslant 1) \end{cases} \tag{5}$$

式中 x——模型预测边界盒参数(坐标或偏移量)

而无监督损失不仅包括定位损失和回归损失,还包括 GIoU 损失,即

$$L_{unsup} = \frac{1}{N_{cls}}\sum_{i=1}^{N_{cls}} L_{cls}(x_u, \hat{y}_{cls}) + \frac{1}{N_{loc}}\sum_{i=1}^{N_{cls}} L_{loc}(x_u, \hat{y}_{loc}) + \frac{\alpha}{N_{loc}}\sum_{i=1}^{N_{loc}} L_{iou}(x_u, \hat{x}_u) \tag{6}$$

式中 N_{cls}——分类样本数量

N_{loc}——定位样本数量

$\hat{y}_{cls}$、$\hat{y}_{loc}$——教师网络生成伪标签

L_{iou}——GIoU 损失值

L_{cls} 和 L_{loc} 的定义与有监督相同,加权项 α 设为 0.5。

为了平滑数据、减少噪声干扰,并提升模型的稳定性和性能,采用半监督训练[22,24]中通过学生网络更新教师网络参数(Exponential moving average, EMA)[25]的策略,即

$$\theta_t^i = \beta\theta_t^{i-1} + (1-\beta)\theta_s^i \tag{7}$$

式中 θ_t——教师网络参数

θ_s——学生网络参数

i——训练步数索引

β——控制参数更新速度的超参数

1.3.2 教师学生网络初始化

目标检测模型可分为两大类:两阶段模型和单阶段模型。两阶段模型中的典型是 Faster R-CNN[26],虽精度较高,但检测速度较慢,难以满足实时性需求。基于回归方法的单阶段目标检测模型无需生成候选区域,可直接输出图像分类与边界框回归结果。这类模型只需对输入图像进行一次前向传播,具有更快的处理速度。在病害检测任务中,除需要保证检测精度外,实际应用场景对实时性也有严格要求,因此本文采用 FCOS[21]为基本目标检测器。FCOS 网络结构由 4 部分组成:输入端、主干网络、特征融合结构和输出端,其详细架构如图 4 所示。

在 burn-in 阶段(图 2),本文使用部分标注数据训练初始 FCOS 检测器,并将训练得到的网络参数同时初始化给教师网络和学生网络(二者初始配置完全相同)。该阶段主要目标是通过有监督训练使教师网络具备生成高质量伪标签的能力。

1.3.3 迭代标注

与传统的教师-学生模型使用静态伪标签集不同,本文提出的迭代标注通过动态迭代方式持续优

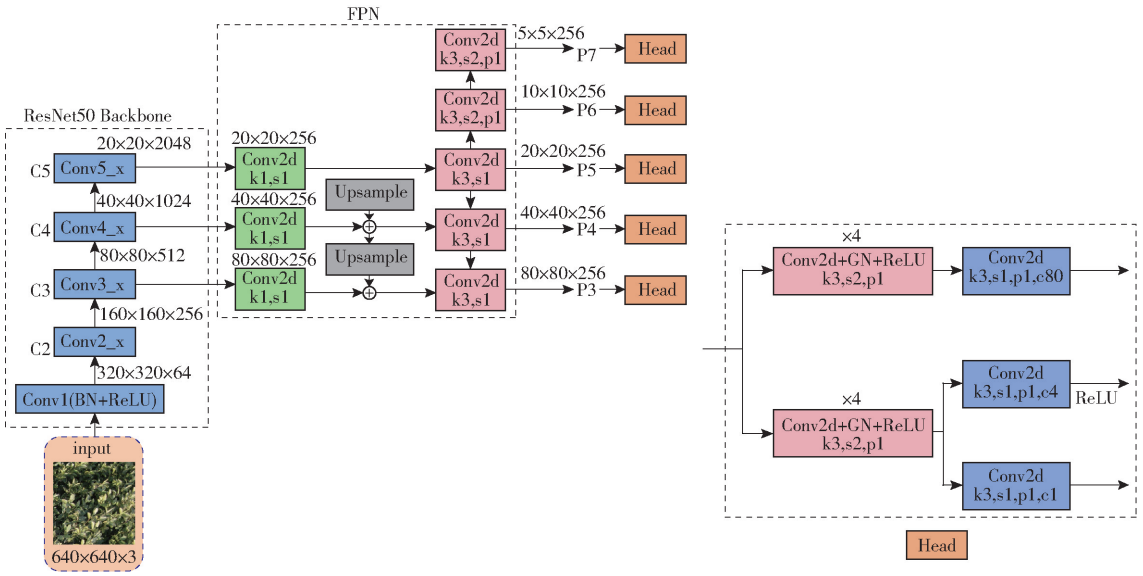


图 4　FCOS 结构图

Fig. 4　Structure diagram of FCOS

化和扩充标签集。在每次训练迭代后，教师模型基于度量分数（Metric score）动态筛选关键样本进行标注，显著提升伪标签质量和训练效率。

受主动教师模型[27]启发，从 3 个维度评估图像重要性，并针对茶叶枯病特性调整指标权重，最终综合判定样本保留价值。

困难度（Difficulty）：得分高的目标具有高不确定性，表现为预测概率分布的高熵值。通过引入目标尺寸权重系数，确保模型对模糊目标和小目标赋予更高分数。在茶叶枯病检测中，含模糊病斑或小病斑图像将获得更高困难度评分。计算公式为

$$s_i^{diff} = -\frac{1}{n_b^i}\sum_{j=1}^{n_b^i}\sum_{k=1}^{N_c}\omega_{size}(b_j)\cdot p(c_k;b_j,\theta_t)\lg p(c_k;b_j,\theta_t) \tag{8}$$

式中　$\omega_{size}(b_j)$——与目标相关的权重，较小目标会赋予较大权重

n_b^i——第 i 幅图像经过 NMS 和 WJCE 筛选后预测的边界框数量

$p(c_k;b_j,\theta_t)$——教师网络在参数 θ_t 下对边界框 b_j 属于类别 c_k 的概率

信息度（Information）：得分高的目标包含丰富的视觉信息，通常表现为存在多个高置信度边界框。这类目标能提供更多有效检测模式，有助于模型学习更全面的特征表征。在密集目标场景中，图像往往包含大量高置信度边界框。通过引入目标尺寸权重系数，确保具有丰富视觉信息的密集目标获得更高信息度分数，从而增强模型对此类区域的学习能力。在茶叶枯病检测中，含有密集病斑和小病斑的图像将获得更高评分。计算公式为

$$s_i^{inf} = \omega_{size}(b_j)\sum_{j=1}^{n_b^i}\text{confidence}(b_j,\theta) \tag{9}$$

式中　$\text{confidence}(b_j,\theta)$——教师网络对第 j 个边界框的预测置信度

复杂度（Diversity）：得分高的图像通常呈目标尺寸多样化分布，这有助于模型学习更通用的检测模式（尤其是小目标检测）。针对茶叶枯病检测任务，虽然目标类别单一，但通过计算目标尺寸的变异系数来衡量多样性，仍可有效提升模型对小病斑的检测性能。在计算过程中，对较小目标赋予较大的 $\omega_{size}(b_j)$，以确保其对多样性分数的贡献度。计算公式为

$$s_i^{dive} = C_V(\{A(b_j)\}_{j=1}^{n_b^i})\sum_{j=1}^{n_b^i}\omega_{size}(b_j) \tag{10}$$

其中

$$C_V = \frac{\sqrt{\frac{1}{n_b^i}\sum_{j=1}^{n_b^i}(A(b_j)-\mu_A)^2}}{\mu_A} \tag{11}$$

式中　$A(b_j)$——第 j 个边界框面积

μ_A——目标大小均值

C_V——变异系数

综合得分计算：通过整合调整后的各项得分，可确保每个未标注样本的得分准确反映其检测难度、信息量和多样性特征，从而在半监督学习过程中实现更有效的样本选择和标注。由于各指标取值范围存在显著差异，首先对单个指标进行归一化处理，即

$$\hat{s}_i^m = \frac{s_i^m}{s_{m,\max}} \tag{12}$$

式中　$\hat{s}_i^m$——归一化后第 i 个样本第 m 个指标分数

s_m^i——原始第 i 个样本第 m 个指标分数

$s_{m,\max}$——第 m 个指标最大值

归一化后的 3 个指标分数可表示为三维空间

中的一个点，每个未标注样本对应一个三维特征向量，表征其在困难度、信息度和多样性 3 个维度的特征

$$\boldsymbol{s}_i = (\hat{s}_i^{diff}, \hat{s}_i^{inf}, \hat{s}_i^{dive}) \tag{13}$$

为将这些三维向量转换为单一标量得分，采用 L－p 范数计算

$$s_{Lp} = L_p(\boldsymbol{s}) = \left(\sum_{i=1}^{3} s_i^p \right)^{1/p} \tag{14}$$

s_{Lp}作为样本筛选的直接依据，其值越大，样本被选中的概率越高。

传统方法通常在训练过程中随机或固定比例选择未标注样本，而本文提出的迭代标注方法动态更新标注集：基于当前模型状态和未标注数据得分，逐步选择最具价值的样本进行标注。相较于传统半监督目标检测方法中常见的伪标签质量不佳问题，本文方法通过主动采样机制和高品质伪标签生成策略，结合真实标签与伪标签协同训练，显著提升了模型性能，同时有效抑制了伪标签噪声的负面影响。

1.3.4 加权联合置信度估计

在茶叶枯病病斑检测中，病斑的密集分布使得分类与定位任务尤为复杂。高分辨率图像中，分类分数 S_{cls}和定位分数 S_{loc}的显著差异会严重影响检测结果。这种不一致性主要表现为：高分类置信度但低定位置信度（或反之）、密集目标导致的边界框重叠、复杂背景对分类和定位的干扰、高分辨率图像细节带来的处理复杂度。

加权联合置信度估计（WJCE）的核心目标是通过建立分类与定位的加权联合置信度来优化伪标签选择。如图 5 所示，其中 C 表示通道数，H、W 表示图像的高、宽。WJCE 采用双分支并行架构：分类分支 S_{cls}专注于目标类别识别，输出分类置信度分数；定位分支 S_{loc}精确计算目标位置，生成定位置信度分数。两个分支在同步监督训练后，其输出通过加权相乘的方式融合，最终生成联合置信度分数作为检测结果的综合评估指标 S，即

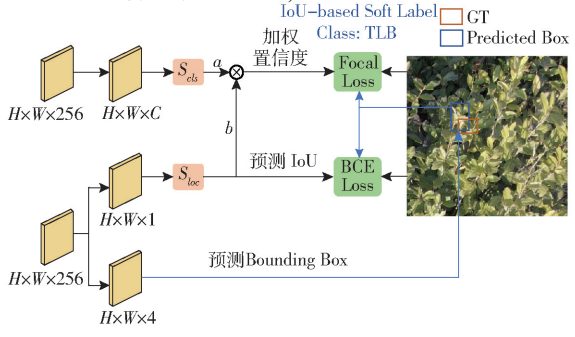


图 5 加权联合置信度估计原理图

Fig. 5 Weighted joint confidence estimation

$$S = S_{cls}^a S_{loc}^b \tag{15}$$

式中 a——分类分数权重

b——定位分数权重

加权联合置信度估计中的权重 a 和 b 通常设定在[0.5, 1.5]范围内。经过实验验证，由于分类分数在目标检测任务中具有相对更高的重要性，当 $a=0.8$、$b=1.2$ 时，模型性能达到最优。S 与原图 S_0 进行比较即可获得对应损失，对于有标注数据，将预测结果与人工标注真值进行对比计算损失；对于无标注数据，则使用预测结果与最高置信度的伪标签进行对比求取损失

$$L_{cls} = FL(S, S_0) \tag{16}$$

为使网络更加专注于 IoU 计算，额外增加 IoU 损失 L_{iou}为

$$L_{iou} = BCE(S_{loc}, IoU) \tag{17}$$

WJCE 通过为分类分数和定位分数分配不同权重，有效综合了两类信息，减少了单一分数导致的误差，增强了置信度估计的鲁棒性，尤其适用于茶叶枯病病斑检测中高分辨率和复杂背景图像的处理。此外，权重 a 和 b 可根据具体任务需求调整，提供了应用灵活性。

1.3.5 自适应样本区分

传统锚框方法在处理小目标（如小病斑）时存在局限性，难以准确捕捉细微特征。相比之下，FCOS 模型通过像素级监督减少了对预定义锚点的依赖，在处理小目标时展现出更好的灵活性和准确性。然而，FCOS 模型像素级监督对标签质量较为敏感，使用伪标签时可能引入噪声，影响模型训练效果和检测精度。

为解决这一问题，采用 DSL[28]（密集学习）中的自适应滤波（AF）策略来精细化处理密集学习中的伪标签，并通过分类挖掘和定位挖掘进一步处理模糊区域（图 6）。传统单一阈值方法易导致大量实例被错误标注，引入训练噪声。AF 策略引入 2 个阈值 τ_{neg}和 τ_{pos}，其计算式为

$$\tau_{neg} = 0.1 \tag{18}$$

$$\tau_{pos} = \left(\frac{\sum_{i=1} 1\{\hat{S}_i == 1\} p_i}{N_{pos}} \right)^{\gamma} \tau \tag{19}$$

式中 γ——控制对尾类的倾斜程度，取 0.7

τ_{pos}——正阈值

$1\{\cdot\}$——指示器函数，如果满足条件$\{\cdot\}$，则输出 1，否则输出 0

$\hat{S}_i$——第 i 个样本在经加权联合置信度估计得到的 IoU 分数

τ——用作参考阈值，取 0.35

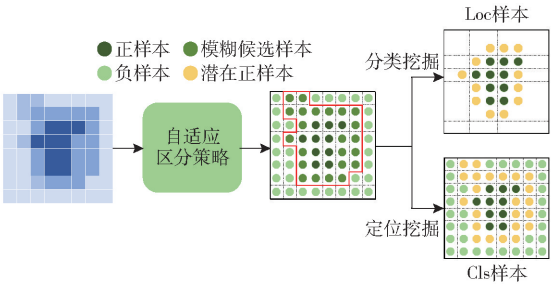


图 6　自适应样本区分示意图

Fig. 6　Sketch map of adaptive sample segregation

N_{pos}——一个小批中的正像素数量

将预测像素分为 3 类：正样本、模糊区域和负样本。这种多阈值分区方法能更准确地识别模糊区域中的潜在正样本，特别适用于小病斑等细小目标。正样本阈值 τ_{pos} 则采用类自适应的策略动态调整，以提升小目标检测精度

$$\hat{S}_i=\begin{cases}Negative & (\mathrm{Max}(S_i)<\tau_{neg})\\ Ambiguilty\ region & (\tau_{neg}\leqslant\mathrm{Max}(S_i)\leqslant\tau_{pos})\\ Positive & (\mathrm{Max}(S_i)>\tau_{pos})\end{cases} \tag{20}$$

当 $\gamma=0.7$ 时模型取得最优 mAP。该参数通过调节对模糊区域和尾部类别的关注程度，实现了阈值设置的动态优化。这种机制使模型能够自适应不同的数据分布特性，满足多样化的任务需求，最终在像素级标注中实现了更精细、多层次的信息分配。

由于对阈值周围样本的判定相对模糊，因此在这些模糊样本里可能存在潜在正样本。在通过 AF 策略得到这 3 种样本后，对模糊候选样本进行进一步分类和定位挖掘：

(1)分类挖掘。根据置信度将样本划分为负样本、正样本和模糊候选样本。但模糊候选样本可能包含值得学习的前景信息，所以将模糊候选样本也纳入分类任务的一致性学习中，以提高分类模型鲁棒性。

(2)定位挖掘。参考文献[29]的处理方式，从模糊候选样本中提取潜在的正样本，以改进定位任务中的学习过程，提取潜在正样本需满足以下相似性度量：①分类相似性：候选样本应与正样本具有相同的预测类别。②定位相似性：候选样本边界框与正样本边界框之间的 IoU 应大于某个阈值(默认 0.6)。③位置相似性：候选样本位置应在正样本边界框内。

一旦选择潜在正样本，其定位目标根据匹配正样本加权平均值来计算，公式为

$$B=\frac{\sum_{i=1}^{N}\mathrm{Max}(\hat{S}_i)\hat{B}_i}{\sum_{i=1}^{N}\mathrm{Max}(\hat{S}_i)} \tag{21}$$

式中　N——匹配正样本数量

$\hat{B}_i$——第 i 个正样本边界框

综上所述，采用 ASS 策略进行小病斑检测的优势在于其能够有效处理模糊区域和小目标检测中常见的伪标签问题，通过精细化的像素级监督和自适应的阈值设定，显著提高了深度学习模型对小病斑等细微目标识别精度和准确性。

1.4　病害检测步骤和流程

针对无人机拍摄的高分辨率图像，首先将其裁剪并调整为 640×640×3 的标准尺寸以适应模型输入要求。为降低自然光照不均对模型性能的影响，采用 Retinex 算法对图像进行预处理，并进行数据复杂性分析。预处理后的图像按比例 7∶1∶2 划分为训练集、验证集和测试集。

基于增广后的数据训练 SSTLBdet 模型，训练过程如图 2 所示。模型优化过程中，通过监控验证集损失值，在其达到最小值时保存最优模型参数。最终，使用训练完成的模型对测试集图像进行预测，输出茶叶枯病的类别标签及对应置信度，实现茶叶枯病检测。检测方法流程如图 7 所示。

2　结果和分析

2.1　预处理结果

图 8 展示了经 Retinex 算法处理前后的图像对比效果，其中对包含病斑的关键区域进行了局部放大显示，直观呈现了该算法在图像增强方面的作用。数据增强阶段共新增 1 000 幅茶叶枯病训练图像，图 9 为部分代表性增广样本，反映数据增广后的图像特征变化。

2.2　实验设置和评价指标

实验硬件配置详见表 1，采用随机梯度下降(SGD)作为优化算法。使用的数据集包含 1 800 幅经过数据增广的茶叶枯病图像，并采用 2 种标注协议进行评估：①部分标注协议：随机选取训练数据的 1%、5%、10%、20% 和 30% 作为标注数据[22]，其余作为未标注数据。②完全标注协议：全部训练数据作为标注样本，额外未标注数据作为补充。采用 F1 分数、平均精度(AP)与平均精度均值(mAP)作为主要评价指标，其中 mAP 针对所有类别进行计算，采用 mAP@0.5 和 mAP@0.5∶0.95；AP 针对单个类别进行计算，采用 AP@0.5 和 AP@0.5∶0.95 来评估模型性能。

采用 FCOS[21] 作为基线模型，其主干网络为 ResNet-50[30]，该网络使用 ImageNet[31] 数据集进行预训练。重量衰减和动量分别设置为 0.000 1 和 0.9。基本学习率设置为 0.002 5，无监督损失

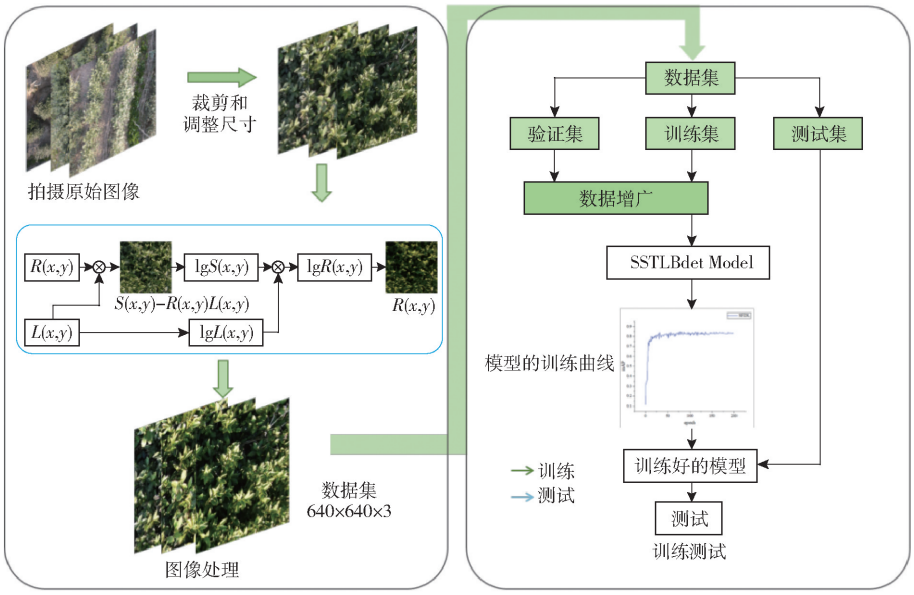


图 7 茶叶枯病检测方法流程图

Fig. 7 Schematic of flow of proposed TLB detection method

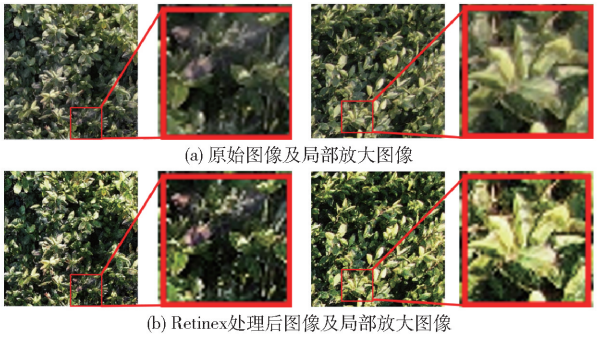


图 8 Retinex 处理前后的图像

Fig. 8 Images before and after Retinex processing

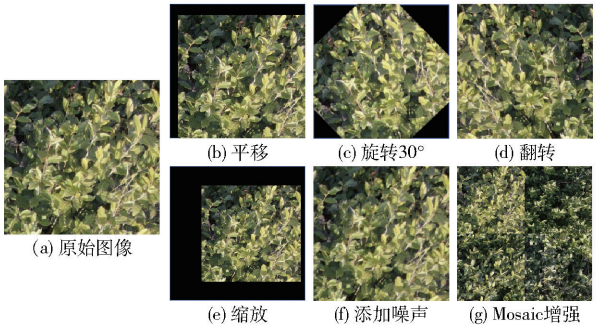


图 9 图像增广样本

Fig. 9 Image augmentation sample

权重设为 2,教师模型通过 EMA 更新动量为 0.999 5。经 200 训练轮次、4 500 次迭代训练后,损失函数变化曲线如图 10 所示。实验结果表明,SSTLBdet 网络最终获得了更优的参数组合,其 mAP 显著优于基线模型。虽然 ARSL 方法表现出更低的损失值,但这反映了 SSTLBdet 模型结构复杂度带来的更强拟合能力。当训练达到 4 500 次时,AP@0.5 和损失曲线均趋于稳定,表明模型已充分收敛。

表 1 实验配置

Tab. 1 Experimental configuration

配置	参数
CPU	Intel 8375C CPU@2.9 GHz 32 cores and 64 threads
GPU	NVIDIA GeForce RTX 3080Ti GPU
操作系统	Ubuntu 20.04 LTS
环境	Pytorch 1.8.1;CUDA 11.1; CUDNN 8.1.1

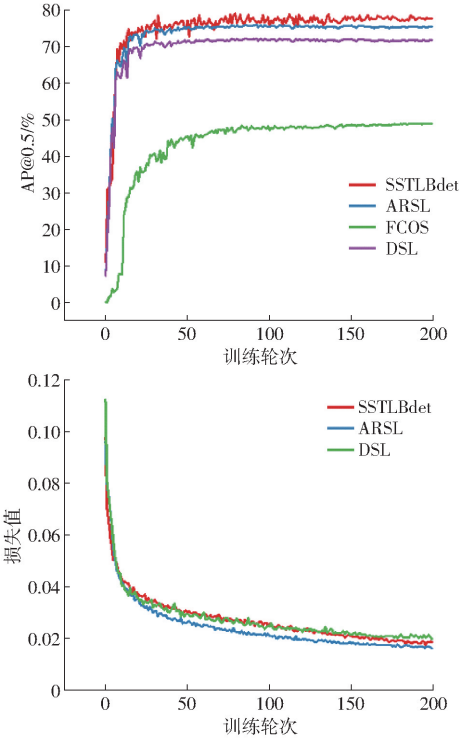


图 10 模型训练曲线

Fig. 10 Training curves of model

2.3 不同半监督方法对比

如表 2 所示,通过对比不同方法及不同无标注比例[22]在无人机茶叶枯病数据集上的检测结果

表 2　不同半监督模型对比实验结果(mAP)

Tab. 2　Comparative experimental results with different semi supervised models (mAP)　%

模型	标注比例/%				
	1	5	10	20	30
Faster R－CNN[26]	21.19	31.81	44.05	50.41	52.04
STAC[22]	29.56	38.60	51.58	54.58	56.36
unbiased teacher[23]	43.90	51.41	59.82	60.04	61.98
unbiased teacher V2[32]	46.21	55.31	57.25	63.86	65.92
humble teacher[33]	35.87	45.95	58.61	59.72	61.65
soft teacher[24]	43.28	58.05	62.05	64.88	66.98
DSL[28]	46.60	53.29	65.32	69.05	71.27
FCOS[21]	17.15	28.47	39.78	43.99	45.42
FCOS[21](全监督)	81.43	81.43	81.43	81.43	81.43
ARSL[29]	48.31	57.25	72.89	73.39	77.76
SSTLBdet	51.28	59.85	75.09	77.99	78.12

(mAP),可以客观评价 SSTLBdet 的性能表现。与现有 SOTA 方法(包括基于 Faster R－CNN 的两阶段检测器和基于 FCOS 的单阶段检测器)相比,SSTLBdet 在所有半监督比例下均优于监督基线和其他半监督检测方法。值得注意的是,FCOS 基线达到与 Faster R－CNN 基线相当的性能水平。

图 11 展示了 SSTLBdet 与 STAC、unbaised teacher、soft teacher、ARSL 模型检测效果对比。其中黄色标注框表示漏检或误检实例。图 11a 中 STAC 方法存在显著漏检;图 11b、11c 表明 unbiased teacher 和 soft teacher 在遮挡和光照不足条件下出现误检和漏检;图 11d 显示 ARSL 虽在模糊密集场景表现良好,但对小病斑仍存在少量漏检;图 11e 证明 SSTLBdet 能准确检测各类茶叶枯病病斑。

综合结果表明,SSTLBdet 在模糊密集小目标检测方面取得显著进展,特别是对小病斑的检测性能提升明显,验证了该方法在茶叶枯病检测任务中的实用价值。

2.4　不同茶叶枯病检测方法对比

为全面评估方法性能,本研究在遥感数据集上与其他茶叶病害检测方法进行了对比实验。本文方法采用 30% 标注数据的半监督 SSTLBdet,而对比方法分别在标注数据比例 30% 和 100% 条件下进行测试,结果如表 3 所示。

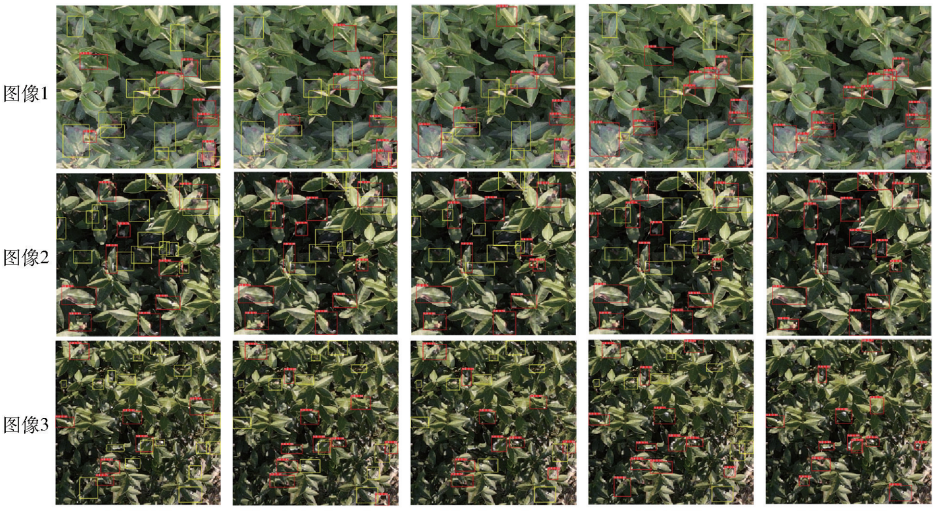


图 11　不同半监督模型检测结果对比

Fig. 11　Comparison of detection results with different semi-supervised models

表 3　不同茶叶枯病检测方法结果对比

Tab. 3　Comparison of quantitative results of different tea leaf blight detection methods　%

参数	LC3Net[34]		DDMA YOLO[11]		LWDNet[13]		LiTLBNet[14]		SSTLBdet
	30%	100%	30%	100%	30%	100%	30%	100%	30%
F1 分数	41.94	76.39	41.58	73.20	42.42	76.58	43.15	73.79	71.24
mAP@0.5	46.32	79.21	45.92	78.52	46.85	80.12	47.66	81.50	78.12
mAP@0.5:0.95	37.50	71.12	26.61	70.50	27.76	71.41	28.01	72.19	69.95

由表 3 可知,通过对比 30% 和 100% 标注数据条件下的实验结果,可以发现在关键性能指标方面,SSTLBdet 方法展现出显著优势。当仅使用 30% 标注数据时,该方法在 F1 分数达到 71.24%,较其他方法在相同条件下的最佳表现(43.15%)提升约 65%;mAP@0.5 达到 78.12%,较对比方法的最佳值(47.66%)提升约 64%。这一优异表现已接近全监督方法使用 100% 标注数据时的性能。

LiTLBNet mAP@0.5:0.95 最优(72.19%),SSTLBdet 略次之(69.95%),与最优结果仅相差 3.1 个百分点。值得注意的是,虽然全监督方法 LC3Net 和 LWDNet 的 F1 分数(分别达到 76.39% 和 76.58%)和 mAP@0.5(分别达到 79.21% 和 80.12%)表现较好,但 mAP@0.5:0.95 指标与 SSTLBdet 差距不超过 3.5 个百分点。

其他全监督方法在使用 100% 标注数据时的综合表现,与 SSTLBdet 仅使用 30% 标注数据时的性能基本相当,各关键指标的差异均保持在 5 个百分点以内。实验结果充分验证了 SSTLBdet 方法在半监督学习框架下的有效性和优越性,特别是在标注数据有限的条件下仍能保持出色的检测性能。

2.5 消融实验

为评估本文方法在无人机遥感图像茶叶枯病检测中的有效性,对迭代标注、加权联合置信度估计、自适应样本区分模块进行了消融实验。选择仅使用 30% 有标签数据进行训练的 FCOS 网络作为基线(即不使用任何半监督方法)。表 4 展示了不同方法在无人机茶叶枯病数据集上的结果。

表 4 消融实验结果

Tab. 4 Results of ablation experiment %

半监督	迭代标注	WJCE	ASS	AP	AP@0.5	AP@0.5:0.95
×	×	×	×	42.36	58.98	45.59
√	×	×	×	48.99	70.23	53.21
√	√	×	×	51.78	74.13	57.64
√	√	√	×	52.58	77.29	62.77
√	√	√	√	56.55	78.12	63.32

注:√表示采用此改进方法,×表示不采用此改进方法。下同。

由表 4 可知,基线网络 FCOS 仅依赖 30% 有标签数据时,检测性能相对较低。在引入剩余的 70% 未标注图像(即加入半监督方法)后,性能显著提升:AP 提升至 48.99%,AP@0.5 提升至 70.23%,AP@0.5:0.95 提升至 53.21%。这说明半监督目标检测方法有效利用了未标注数据,从而提高了检测性能。加入迭代标注模块后,性能进一步提升:AP 达到 51.78%,AP@0.5 达到 74.13%,AP@0.5:0.95 达到 57.64%。这表明迭代标注有助于逐步提升模型的检测能力,尤其是在处理小目标和背景复杂模糊的密集病斑时。引入 WJCE 模块后,AP 提升至 52.58%,AP@0.5 提升至 77.29%,AP@0.5:0.95 提升至 62.77%。这说明 WJCE 模块有效缓解了定位与分类不一致的问题,进一步提高了检测准确性,尤其对小目标检测精度提升明显。最终集成 ASS 模块后,模型 AP、AP@0.5 和 AP@0.5:0.95 均达到最高值。这验证了 ASS 模块通过其精细化的像素级监督和自适应的阈值设定,显著提升了模型对小病斑等细微目标的识别精度和准确性。

上述实验结果为今后进一步优化密集物体检测模型提供了有力参考,同时也展示了权重调整在提高检测准确性和鲁棒性方面的潜力。

在不同比例的标注数据下,对迭代标注模块进行了消融实验(表 5)。其基线为移除了整个迭代标注模块的 SSTLBdet 方法(即使用传统半监督目标检测框架)。

表 5 迭代标注中不同采样策略的消融实验结果

Tab. 5 Experimental results of ablation using different sampling strategies in iterative annotation %

困难度	信息度	丰富度	标注比例/%		
			5	10	20
×	×	×	47.84	64.39	67.26
√	×	×	56.03	61.13	64.59
×	√	×	55.92	60.98	62.87
×	×	√	57.40	62.26	69.05
√	√	√	59.85	75.09	77.99

由表 5 可知,困难度、信息度和丰富度 3 个评分指标对于半监督目标检测是必要的。此外,在训练过程中,AutoNorm 评分优于其他指标,最终实现了最佳的整体性能,这充分证实了迭代标注模块的有效性。

3 结论

(1)针对自然场景下茶叶枯病检测的困难,提出了一种基于无人机遥感图像的茶叶枯病半监督检测方法。利用最少的标注数据,在茶叶枯病病害模糊密集图像的小目标检测中实现与完全监督模型相当的检测性能。

(2)SSTLBdet 对茶叶枯病的检测性能优于 ARSL、DSL 等半监督模型。当标注数据比例为 30% 时,学生模型获得最佳检测性能,mAP 达到 78.12%。SSTLBdet 对小病斑的预测优于其他半监督模型,显著减少了漏检误检。

(3)在不同标注数据比例下,SSTLBdet 均有效提升了基线 FCOS 的检测精度。当标注数据比例为 30% 时,学生模型茶叶枯病检测性能与完全监督模型相当。基于半监督学习方法在茶叶枯病检测方面进行了初步探索,实验结果表明,该方法具有一定的有效性,能够在密集模糊的图像中准确检测出小病斑,同时减少了数据标注工作量。

参 考 文 献

[1] HU G, WU H, ZHANG Y, et al. A low shot learning method for tea leaf's disease identification[J]. Computers and Electronics in Agriculture, 2019, 163: 104852.

[2] 李应祥, 陈跃华, 杨青, 等. 茶云纹叶枯病病原鉴定[J]. 贵州农业科学, 2015, 43(1): 65-67.
LI Yingxiang, CHEN Yuehua, YANG Qing, et al. Identification of pathogenic fungus causing tea brown blight[J]. Guizhou Agricultural Sciences, 2015, 43(1): 65-67. (in Chinese)

[3] HU G, WANG H, ZHANG Y, et al. Detection and severity analysis of tea leaf blight based on deep learning [J]. Computers & Electrical Engineering, 2021, 90: 107023.

[4] 华黔星. 茶叶病虫害绿色防控技术的应用探讨[J]. 农业开发与装备, 2023(12): 233-235.

[5] KERKECH M, HAFIANE A, CANALS R. Deep leaning approach with colorimetric spaces and vegetation indices for vine diseases detection in UAV images [J]. Computers and Electronics in Agriculture, 2018, 155: 237-243.

[6] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: unified, real-time object detection[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016: 779-788.

[7] REDMON J, FARHADI A. Yolov3: an incremental improvement[J]. arXiv Preprint arXiv:1804.02767, 2018.

[8] LIN T Y, GOYAL P, GIRSHICK R, et al. Focal loss for dense object detection[C]// Proceedings of the IEEE International Conference on Computer Vision, 2017: 2980-2988.

[9] LIU W, ANGUELOV D, ERHAN D. SSD: single shot multibox detector[C]// European Conference on Computer Vision. Cham: Springer International Publishing, 2016: 21-37.

[10] CHEN C J, HUANG Y Y, LI Y S, et al. An AIoT based smart agricultural system for pests detection[J]. IEEE Access, 2020, 8: 180750-180761.

[11] BAO W, ZHU Z, HU G, et al. UAV remote sensing detection of tea leaf blight based on DDMA-YOLO[J]. Computers and Electronics in Agriculture, 2023, 205: 107637.

[12] 黄亮, 吴春燕, 李小祥, 等. 基于弱监督语义分割的灯盏花无人机遥感种植信息提取[J]. 农业机械学报, 2022, 53(4): 157-163.
HUANG Liang, WU Chunyan, LI Xiaoxiang, et al. Extraction of erigeron breviscapus planting information by unmanned aerial vehicle remote sensing based on weakly supervised semantic segmentation[J]. Transactions of the Chinese Society for Agricultural Machinery, 2022, 53(4): 157-163. (in Chinese)

[13] HU G, YE R, WAN M, et al. Detection of tea leaf blight in low-resolution UAV remote sensing images[J]. IEEE Transactions on Geoscience and Remote Sensing, 2023, 62: 1-18.

[14] 胡根生, 谢一帆, 鲍文霞, 等. 基于轻量型网络的无人机遥感图像中茶叶枯病检测方法[J]. 农业机械学报, 2024, 55(4): 165-175.
HU Gensheng, XIE Yifan, BAO Wenxia, et al. Lightweight network for tea leaf blight detection in UAV remote sensing images[J]. Transactions of the Chinese Society for Agricultural Machinery, 2024, 55(4): 165-175. (in Chinese)

[15] ZHOU P, CHENG G, YAO X, et al. Machine learning paradigms in high-resolution remote sensing image interpretation [J]. National Remote Sensing Bulletin, 2021, 25(1): 182-97.

[16] XU X, WANG L, LIANG X, et al. Maize seedling leave counting based on semi-supervised learning and UAV RGB images [J]. Sustainability, 2023, 15(12): 9583.

[17] NONG C, FAN X, WANG J. Semi-supervised learning for weed and crop segmentation using UAV imagery [J]. Frontiers in Plant Science, 2022, 13: 927368.

[18] NAJAFIAN K, GHANBARI A, STAVNESS I, et al. A semi-self-supervised learning approach for wheat head detection using extremely small number of labeled samples[C]// Proceedings of the IEEE/CVF International Conference on Computer Vision, 2021: 1342-1351.

[19] COOPER T J, BAQAI F A. Analysis and extensions of the Frankle-McCann Retinex algorithm [J]. Journal of Electronic Imaging, 2004, 13(1): 85-92.

[20] 林成创, 单纯, 赵淦森, 等. 机器视觉应用中的图像数据增广综述[J]. 计算机科学与探索, 2021, 15(4): 583-611.
LIN Chengchuang, SHAN Chun, ZHAO Gansen, et al. Review of image data augmentation in computer vision[J]. Journal of Frontiers of Computer Science & Technology, 2021, 15(4): 583-611. (in Chinese)

[21] TIAN Z, SHEN C, CHEN H, et al. FCOS: fully convolutional one-stage object detection. [J]. arXiv Preprint, arXiv: 190401355, 2019.

[22] SOHN K, ZHANG Z, LI C L, et al. A simple semi-supervised learning framework for object detection [J]. arXiv Preprint, arXiv: 200504757, 2020.

[23] LIU Y C, MA C Y, HE Z, et al. Unbiased teacher for semi-supervised object detection [J]. arXiv Preprint, arXiv: 210209480, 2021.

[24] XU M, ZHANG Z, HU H, et al. End-to-end semi-supervised object detection with soft teacher[C]// Proceedings of the

IEEE/CVF International Conference on Computer Vision, 2021: 3060 - 3069.

[25] TARVAINEN A, VALPOLA H. Mean teachers are better role models: weight-averaged consistency targets improve semi-supervised deep learning results[J/OL]. Advances in Neural Information Processing Systems, 2017. DOI:10.48550/arXiv.1703.01780.

[26] REN S, HE K, GIRSHICK R, et al. Faster R-CNN: towards real-time object detection with region proposal networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2016, 39(6): 1137 - 1149

[27] MI P, LIN J, ZHOU Y, et al. Active teacher for semi-supervised object detection[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2022: 14482 - 14491.

[28] CHEN B, LI P, CHEN X, et al. Dense learning based semi-supervised object detection[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2022: 4815 - 4824.

[29] LIU C, ZHANG W, LIN X, et al. Ambiguity-resistant semi-supervised learning for dense object detection[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023: 15579 - 15588.

[30] RUSSAKOVSKY O, DENG J, SU H, et al. Imagenet large scale visual recognition challenge[J]. International Journal of Computer Vision, 2015, 115(3): 211 - 252.

[31] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016: 770 - 778.

[32] LIU Y C, MA C Y, KIRA Z. Unbiased teacher v2: semi-supervised object detection for anchor-free and anchor-based detectors[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2022: 9819 - 9828.

[33] TANG Y, CHEN W, LUO Y, et al. Humble teachers teach better students for semi-supervised object detection[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021: 3132 - 3141.

[34] JIANG Y, LU L, WAN M, et al. Detection method for tea leaf blight in natural scene images based on lightweight and efficient LC3Net model [J]. Journal of Plant Diseases and Protection, 2024, 131(1): 209 - 225.