doi:10.6041/j.issn.1000-1298.2026.18.008

基于冠层图像煤污识别的温室白粉虱危害分级方法

杨　月[1,2]　刘云玲[1]　刘亚雄[1]　宋坚利[3]　范　桢[4]

(1.中国农业大学信息与电气工程学院,北京 100083;2.北京外国语大学信息技术中心,北京 100089;
3.中国农业大学理学院,北京 100193;4.北京荷梓科技有限公司,北京 100043)

摘要:针对温室白粉虱成虫体型微小且具有隐蔽性,导致冠层监控难以实现直接监测的问题,利用其诱发的煤污病征进行危害程度识别具有重要意义。然而,由于不同作物背景本底颜色空间的重叠干扰,以及卷积神经网络易捕获背景非因果特征导致跨场景泛化难题,本文提出了一种基于冠层图像煤污识别的温室白粉虱危害分级方法。确立了以煤污面积比与视觉显著度为核心的 4 级危害指数(HI-0~HI-3)评价标准,并构建了覆盖多物候期与异质背景的专家标注数据集,以此提出一种颜色先验引导的双通道融合分级架构(CPG-Fusion)。以 ResNet-18 为骨干网络,通过嵌入颜色先验注意力模块(CPAM),将具备亮度不变性的逐像素相对暗度转换为空间注意力权重,通过残差调制方式强制模型特征向煤污病征区域聚焦,从而抑制背景信号干扰。试验结果显示,在独立温室的跨场景评估中,集成 CPAM 的模型将二次加权 Kappa(Quadratic weighted Kappa, QWK)从 ResNet-18 基线的 0.401 提升至 0.551,且所有误判均严格限制在相邻等级内。横向试验证实,本方法在保持高精度的同时实现了本组对比中最低的折间标准差,可靠性优于主流自学习注意力机制。梯度加权类激活映射(Grad-CAM)可视化证实,模型特征激活区已与真实煤污物理掩膜实现空间对齐。研究结果证明了物理约束在提升定性识别到定量分级中的关键作用,为设施蔬菜精准植保提供了低成本、高鲁棒性的自动监测方法。

关键词:温室白粉虱;煤污识别;冠层图像;序数分级;颜色先验注意力模块

中图分类号:S433.39;TP391.41　**文献标识码**:A　**文章编号**:1000-1298(2026)18-0081-12　**OSID**:

Method for Greenhouse Whitefly Damage Grading Based on Sooty Mold Recognition in Canopy Images

Yang Yue[1,2]　Liu Yunling[1]　Liu Yaxiong[1]　Song Jianli[3]　Fan Zhen[4]

(1. *College of Information and Electrical Engineering*, *China Agricultural University*, *Beijing* 100083, *China*
2. *Information Technology Center*, *Beijing Foreign Studies University*, *Beijing* 100089, *China*
3. *College of Science*, *China Agricultural University*, *Beijing* 100193, *China*
4. *Beijing Hezi Technology Co.*, *Ltd.*, *Beijing* 100043, *China*)

Abstract: Aiming to address the cross-scene generalization deficit of end-to-end CNN classifiers on the task of greenhouse whitefly damage grading in facility vegetable production, where standard models tended to fix on background-correlated cues such as illumination and plant growth stage rather than the physical signature of sooty mold, a color-prior-guided two-channel grading method (CPG-Fusion) was proposed. The core of the method was a color prior attention module (CPAM): the per-image relative darkness, computed as a scene-brightness-independent physical cue, was fed as the attention input signal of a ResNet-18 backbone so that the CNN was consistently guided toward the region that directly corresponded to sooty mold coverage and was steered away from background spurious features. A dataset of 260 images with 4-level hazard index annotations was collected from two solar greenhouses in Changping, Beijing; the CNN was trained under a 5-fold stratified cross-validation protocol, and the deployed prediction was obtained by softmax probability averaging across folds. On 58 cross-scene OOD

收稿日期:2026-06-06　修回日期:2026-06-20

基金项目:国家重点研发计划项目(2023YFD2001200)、国家现代农业产业技术体系项目(CARS-28)和中国高校产学研创新基金项目(2024WA014)

作者简介:杨月(1982—),男,博士生,北京外国语大学高级工程师,主要从事智慧农业与农业图像识别研究,E-mail:yangyuecau@cau.edu.cn

通信作者:刘云玲(1976—),女,副教授,博士生导师,主要从事农业人工智能与智慧农业研究,E-mail:liuyunling@cau.edu.cn

images independently collected in a second greenhouse (GH2), CPAM raised the quadratic weighted Kappa (QWK) from 0.401 of the ResNet-18 baseline to 0.551, improved the recall of the light-damage level by 22 percentage points, and confined all misclassifications to adjacent levels. In the same-protocol horizontal comparison with four representative attention modules (SE-Net, CBAM, ECA and Coordinate Attention), CPAM matched the upper bound of accuracy of the four counterparts on the OOD ensemble while attaining the lowest across-fold standard deviation among the four self-attention modules, indicating better single-model reliability under small-sample weak supervision. Grad-CAM visualization and the near-zero marginal gain of late fusion between CPAM and the pixel prior channel jointly confirmed that the pixel color prior was internalized into the CNN feature map. The proposed method showed that injecting interpretable physical cues into the attention pathway can improve the robustness of deep learning models under complex agricultural environments, and can be directly deployed on existing canopy surveillance cameras without additional hardware or target-scene parameter tuning.

Key words: greenhouse whitefly; sooty mold recognition; canopy image; ordinal grading; color prior attention module

0 引言

设施蔬菜在我国规模化生产中已成为城市菜篮子稳定供给的核心载体。在设施农业数字化战略推动下,温室冠层视频监控已大规模普及,为作物健康连续观测提供了数据基础。在害虫自动识别领域,将海量的视频数据转化为有效的植保决策信息是当前研究的核心。最初对于害虫的监测主要依赖人工田间调查,不仅耗时费力,且难以实现实时动态预警。随着计算机视觉技术的发展,研究者开始尝试利用数字图像处理技术替代人工观测。早期工作多基于人工设计的特征,如利用阈值分割提取虫体颜色、采用边缘检测获取几何形态特征等;随后支持向量机、随机森林为代表的传统机器学习算法被引入害虫识别任务,虽在一定程度上提升了分类精度,但在面对高维冠层图像数据时仍受限于特征工程的复杂性。近年来,以卷积神经网络(Convolutional neural network, CNN)为代表的深度学习技术凭借端到端特征自学习能力,在农业病虫害识别领域展现出主导优势。Wang 等[1]将 CNN 扩展到苹果黑腐病 4 级严重度分级,确立了基于深度学习的病害严重度估计范式。国内农业机械领域亦持续推进,翟肇裕等[2]对农作物病虫害识别关键技术进行了系统综述。赵辉等[3]将改进 YOLO v7 用于复杂农田害虫目标检测,精确率达到 91.6%。何斌等[4]针对温室夜间场景改进 YOLO v5,实现番茄果实快速识别。王志彬等[5]采用动态集成方法,在黄瓜白粉病、霜霉病等 4 类叶片图像上识别准确率达 96.68%。然而,上述以成虫个体或叶部典型病斑为识别对象的研究,在冠层远距监控场景下仍面临严峻挑战。白粉虱成虫体型极小,具有趋光背阴生物特性,在广角视场中像素占比极低且常隐藏于叶背,导致实时监测漏检率高,难以满足大面积、非侵入式的连续预警需求。

作为设施蔬菜中危害最重的害虫之一,白粉虱若虫与成虫吸食叶片汁液并分泌富含糖分的蜜露,蜜露在叶面吸附煤污菌孢子并形成黑色霉膜(统称煤污),不仅遮蔽叶面降低光合作用,更污染果实表面造成价值损失。事实上,温室白粉虱诱发的次生病征为解决上述困局提供了关键切入点。Kamikawa 等[6]通过定量监测证实了白粉虱密度上升导致蜜露分泌增加,进而发现煤污面积扩大的因果链,为煤污推断白粉虱危害提供了农学基础。相比于微小虫体,面积化煤污病征在冠层图像中具有更高的显著性与视觉可辨识度,使得利用现有冠层图像识别煤污并分级危害成为一种更具鲁棒性的方法学路径。

目前围绕煤污的识别研究虽已初见端倪,但多沿用端到端黑盒分类的模式。Apacionado 等[7]在柑橘果园冠层夜间图像上,比较了 YOLO v5m、YOLO v7、CenterNet 3 种检测网络识别煤污的能力,最佳平均精度均值 74.4%,且任务限定在边界框检测而非危害分级,场景为户外柑橘而非温室。Preanto 等[8]采用语义分割方法在叶面上检测煤污像素,仅验证了煤污覆盖与植株健康度之间的负相关,未涉及生产中急需的危害指数输出。事实上,在温室实际生产环境中,由于不同作物背景如樱桃番茄与黄瓜的本底颜色空间存在重叠干扰,且 CNN 在小样本弱监督训练下极易捕获场景光照、植株物候等与危害指数偶然共存的非因果特征,导致模型在换棚、换季等跨场景部署中面临严重的泛化难题。这种对特定场景统计规律的过度依赖,使模型在分布外(Out-of-distribution, OOD)场景下性能显著退化。虽然已有研究尝试利用计算机视觉技术识别煤污,但大多局限于定性识别或目标检测,难以满足生产中对危害指数定量评估的需求[9-10]。

针对以上问题，本文提出由像素先验通道（Pixel-prior channel，PPC）与颜色先验引导 CNN 通道（Color-prior-guided CNN，CPG－CNN）共同组成的双通道融合分级方法（Color-prior-guided fusion，CPG－Fusion）。利用可解释性的像素比例特征建立逻辑基准，通过构建颜色先验注意力模块（Color-prior attention module，CPAM），提取与场景亮度无关的逐像素相对暗度作为物理线索，使得 CNN 锁定真实的煤污病征区域，在特征提取阶段降低模型对环境背景等伪特征的依赖。通过系统对照物理特征与深度特征的判定一致性，验证该方法在规避捷径学习风险、提升跨场景泛化性能方面的优越性，为设施蔬菜精准植保提供更具鲁棒性与可解释性的方法。

1　材料与方法

1.1　CPG－Fusion 架构

融合颜色物理特征与深度特征的双通道分级架构 CPG－Fusion 如图 1 所示，采用物理线索校准深度特征，主要由数据输入层、特征提取双通道以及决策融合层组成。其中，像素先验通道 PPC 针对煤污受害症状的物理本质，通过颜色空间解析提取黑色、黄色及褐色的像素比例特征，并集成决策树序数分类器以提供高透明度的判别基准。颜色先验注意力模块 CPAM 通过提取具备亮度不变性的逐像素相对暗度生成软连续掩膜，作为物理线索注入特征图，强制模型关注符合煤污属性的区域，从而抑制对场景光影等非因果背景特征的依赖。最后，系统采用后期融合策略，通过线性加权平衡像素通道与 CNN 通道的概率输出，利用决策树修正深度学习模型在跨场景部署中判定偏差，确保最终输出 4 级危害指数（HI－0 至 HI－3）在复杂变光及异质背景下具备稳定的分类性能。

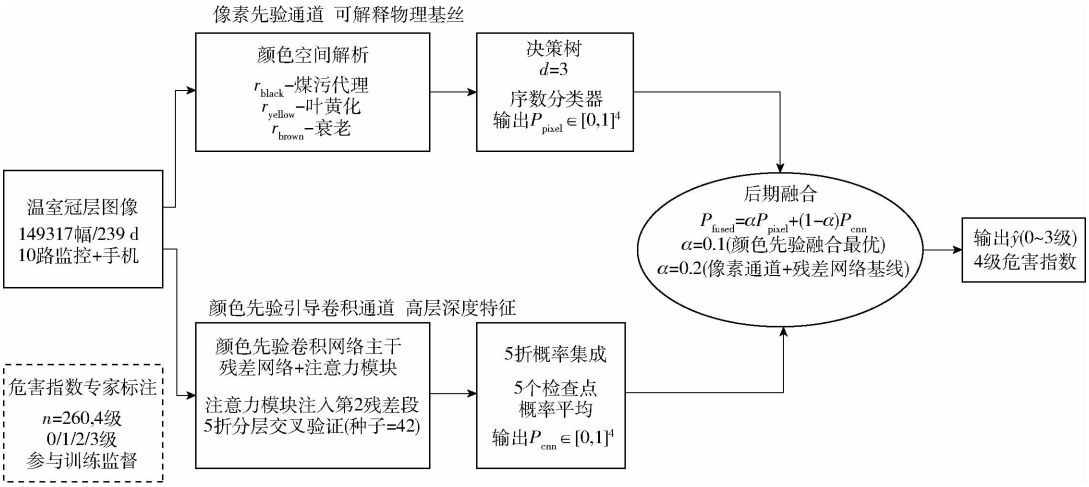

图 1　结合颜色物理特征与卷积特征的双通道危害分级算法框架

Fig. 1　Two-channel damage grading framework combining color physical features and convolutional features

1.2　CPG－CNN 网络

如图 2 所示，CPG－CNN 以 ResNet－18[11] 为骨

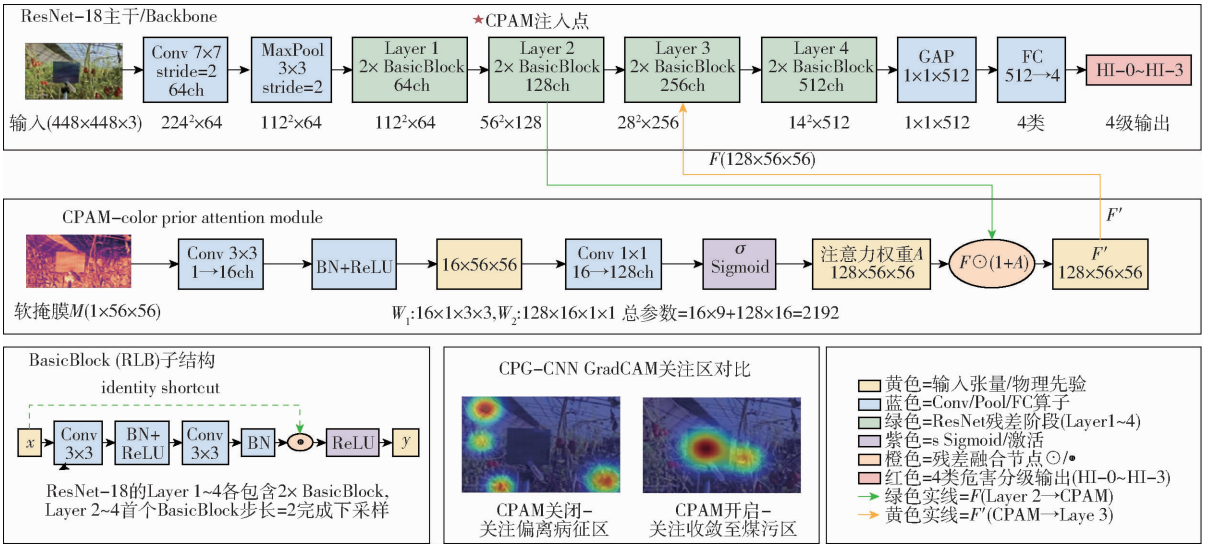

图 2　CPG－CNN 整体网络结构

Fig. 2　Overall architecture of CPG－CNN

干网络,在 Layer 2 输出端插入 CPAM 模块,输出接 4 类全连接分类头。在 ResNet - 18 主干网络中,448 像素 ×448 像素 ×3 输入图像经 Conv1(7 ×7,stride =2,64 通道) + MaxPool 下采样至 112 像素 ×112 像素 ×64,再依次经过 Layer 1 ~ Layer 4 4 个残差段(每段含 2 个 BasicBlock)逐级提取特征,分辨率沿 112 像素 ×112 像素→56 像素 ×56 像素→28 像素 ×28 像素→14 像素 ×14 像素下降,通道数沿 64→128→256→512 上升,最终经全局平均池化与 FC 层得到 4 类危害指数预测。CPAM 注入点位于 Layer 2 输出端(128 通道 ×56 像素 ×56 像素特征图)。在 CPAM 模块中自身的张量流:连续软掩膜 $\hat{M}$ 经 Conv 3 ×3→BN + ReLU→Conv 1 ×1→σ 4 步变换得到逐通道注意力权重 A,再与 Layer 2 输出特征 F 通过残差融合节点 $F\odot(1+A)$ 完成调制,调制后特征 F'注入 Layer 3 继续后续前向传播。

CPAM 插入位置取决于颜色先验掩膜与卷积特征图在空间分辨率和语义抽象度上的匹配。Layer 1 输出(112 像素 ×112 像素)分辨率高而语义弱,过早施加颜色先验近似于对原图作逐像素硬约束,会抑制网络对煤污纹理的自适应表征;Layer 4 输出(14 像素 ×14 像素)空间分辨率过低,煤污斑块的位置信息被压缩,掩膜难精细对齐;Layer 2 输出(56 像素 ×56 像素中层特征)既能将煤污区域在特征图上准确定位、又已具备一定语义,使颜色先验掩膜与特征图的耦合最为充分,故将 CPAM 插入 Layer 2 输出端。

1.3 试验与数据预处理

在北京市昌平区职业学校两栋标准日光温室(GH1、GH2)开展试验。试验期间,每栋温室各部署 5 路监控摄像头,3 路固定顶视广视角机位(GH1:1 - 1、1 - 2、1 - 3;GH2:2 - 1、2 - 2、2 - 3,分辨率 1 920 像素 ×1 080 像素)以及 2 路海康侧视球机(GH1:hk1 - 1、hk1 - 2;GH2:hk2 - 1、hk2 - 2,分辨率 1 920 像素 ×1 080 像素,固定预置位),所有设备通过现场边缘网关定时将图像上传至云端服务器。试验监测周期为 2025 年 8 月—2026 年 4 月共计 239 d。为验证模型在不同作物背景及不同物候期下泛化性能,试验场地设置差异化的种植模式:GH1 全程种植樱桃番茄(圣女果,2025 年 9 月 5 日定植,2026 年 2 月起进入持续采收期);而 GH2 经历黄瓜单种(2025 年 9 月 5 日—12 月 1 日)、番茄与生菜空间分区混种(2025 年 12 月 19 日—2026 年 3 月 10 日)、以及多作物(番茄/黄瓜/生菜/苗盘)多分区混种(2026 年 3 月 26 日—4 月 18 日)3 个茬次。整个试验期共采集有效图像 149 317 幅,覆盖 9 个完整的茬次阶段。为规避夜间红外补光对煤污病征色彩表征的干扰,严格筛选每日 12:00—18:00 时段内 RGB 图像作为研究样本。

1.3.1 危害指数定义

为确保模型学习目标的客观性,确立以物理特征覆盖度为基础、以视觉显著度为核心判定依据的 4 级危害指数(Hazard index,HI)评价标准,如表 1 所示。视觉显著度表示煤污病征在复杂冠层背景下色差辨识度与空间连续性,其中 HI - 0 为健康期,叶面无肉眼可见病征;HI - 1 为轻度危害期,叶面出现零星蜜露点及极少量煤污;HI - 2 为中度危害期,煤污覆盖面积显著扩大且伴随叶片轻微黄化;HI - 3 为重度危害期,煤污形成连续霉膜覆盖叶面并严重阻碍光合作用。

表 1 温室白粉虱诱发煤污病征的危害分级标准

Tab.1 Hazard grading criteria for sooty mold symptoms induced by greenhouse whitefly

危害指数	视觉显著度描述	物理形态特征描述	专家标注逻辑依据
HI - 0(健康)	无显著特征。冠层叶片呈现正常的植被反射光谱,无暗色区域	叶面洁净,无肉眼可见蜜露或菌丝	图像中 R、G、B 通道分布符合典型绿叶特征
HI - 1(轻度)	低显著度。表现为散在的、非连续的暗斑,对比度较低	仅有少量零星蜜露点,煤污面积比小于 5%	需结合原图放大观察,主要为初期侵染,背景干扰最强
HI - 2(中度)	中显著度。暗色区域形成明显连通域,与正常叶片色差增大	煤污呈斑块状分布,部分遮盖叶脉,面积比在 5% ~20% 之间	视觉上能快速锁定病灶区,背景干扰减弱
HI - 3(重度)	高显著度。叶片大面积被黑色菌膜覆盖,呈现极高的视觉反差	煤污连成片状,叶片伴有明显黄化或枯萎,面积比大于 20%	目标特征极强,即使在极暗或极亮光照下仍具有高辨识度

1.3.2 数据集构建

构建 5 个针对性的数据集,分别用于像素特征学习、CNN 训练、CNN 验证、跨设备评估和跨场景 OOD 评估。DS1_expert 数据集由手持 iPhone 手机采集的 202 幅图像组成,经专家校准确立 4 级 HI 标签,用于 PPC 通道决策树的学习以及检验模型在不同成像设备下稳定性。跨场景 OOD DS2_expert 数据集包含手持 iPhone 手机采集于 GH2(黄瓜与番茄混种)的 58 幅图像,作为核心测试集验证模型在异质作物背景、不同物候期及光照偏移下泛化性能。

DS_train_weak 数据集包含 GH1 监控摄像头拍摄的 1 600 幅图像，利用时段伪标签进行弱监督训练。DS_val_internal 数据集包含 400 幅同源图像，用于模型训练过程超参数调整与最优权重筛选。DS_val_48 数据集包含 48 幅由训练摄像头采集的专家标注图像，作为真值，用于建立监控视角下性能基准。数据集见表 2。

DS2 等级分布说明：DS2 采集日（2026 年 4 月 10 日）为温室常规防治介入早期，专家校正后实际样本仅覆盖 HI－0（31 幅）与 HI－1（27 幅）2 级。作为对照，自动启发式标签（hi_auto，基于 DS1 阈值规则在 GH2 上推断）分布为 14/34/8/2，与专家校正存在 22 处差异（38%），差异方向全部为自动标注高于专家标注——即 DS1 校准的颜色阈值在 GH2 黄瓜背景下系统性高估危害指数。本文 4 级序数分级框架在 DS1（含 HI－2/HI－3）上训练并验证，跨场景 OOD 评估口径统一以专家精炼的 hi_final 为准。

表 2　数据集构成

Tab. 2　Dataset composition

数据集	用途	样本量/幅	等级分布（HI－0/1/2/3）（专家校正后）/幅	来源
DS1_expert	像素方法主评估 + CNN 跨设备评估	202	24/56/71/51	GH1 樱桃番茄，2026－01－12 单日手持 iPhone 采集
DS2_expert	跨场景 OOD 评估	58	31/27/0/0	GH2 黄瓜 + 番茄混种，2026－04－10 单日手持 iPhone 采集
DS_train_weak	CNN 训练集	1 600	400/400/400/400	4 路 GH1 顶视摄像头 ×4 类 ×100 幅，时段伪标签
DS_val_internal	CNN 验证集	400	100/100/100/100	同上，25 幅/类/摄像头
DS_val_48	真值（每摄每级 4 幅）	48	12/12/12/12	训练摄像头自身

时段伪标签构建：基于现场观察 + 视觉巡检确认 GH1 樱桃番茄在 2025 年 10 月 1 日至 2026 年 1 月 18 日共 4 个连续时段对应 4 级危害指数，分别为 HI－0 清洁期（10 月 1 日—11 月 15 日）、HI－1 轻度期（11 月 16 日—12 月 27 日）、HI－2 中度期（2025 年12 月 28 日—2026 年 1 月 5 日）、HI－3 重度期（1 月 6 日—1 月 18 日），按时段为该期间的所有图像统一打标签。该方案显著降低了标注成本，但本质属于弱监督，等级标签与单幅图像内容并非一一对应。

DS1 与 DS2 共 260 幅图像在田间现场用 iPhone 12 Pro 单日采集后由作者依据当日实地观察、文献描述与煤污典型视觉特征（黑色斑块覆盖面积比、叶面亮度下降、叶黄化辅助判读）逐幅人工标注 4 级危害指数。

1.3.3　像素特征通道提取

像素通道直接提取冠层图像在 RGB 空间下 3 个颜色比例特征。对每幅输入图像缩放至 200 像素 × 200 像素后，分别计算 3 个像素比例特征，计算式为

$$r_{\text{black}} = \text{mean}[(R<60) \wedge (G<60) \wedge (B<60)] \tag{1}$$

$$r_{\text{yellow}} = \text{mean}[(R>150) \wedge (G>110) \wedge (B<90) \wedge (R>G) \wedge (G>B)] \tag{2}$$

$$r_{\text{brown}} = \text{mean}[(R>80) \wedge (R>1.2G) \wedge (G>B) \wedge (50<L<160)] \tag{3}$$

式中　R、G、B——像素在 RGB 颜色空间下三通道强度（取 0 ~ 255）

L——像素亮度分量

$\text{mean}[\cdot]$——满足括号内条件的像素数占全图像素比例

$\wedge$——逻辑“与”

r_{black}——煤污覆盖面积比（核心特征），取 0 ~ 1

r_{yellow}、r_{brown}——白粉虱吸食致叶片黄化、衰老叶片/背景，取 0 ~ 1

为综合评估多维颜色特征对危害指数贡献度，定义基于先验权重的综合评估得分，计算式为

$$s = 5r_{\text{black}} + 2r_{\text{dark}} + 1.5r_{\text{yellow}} + r_{\text{brown}} - 0.3r_{\text{green}} \tag{4}$$

式中　s——综合评估得分

r_{dark}——深色像素比例（$\text{mean}[L<80]$）

r_{green}——绿色像素比例

各项系数由 GH1、GH2 样本分布确定的先验权重（正系数对应危害相关特征，负系数为健康绿叶抑制项）。

综合得分系数基于 GH1 与 GH2 样本特征分布观测确立，跨场景评估中对比仅用三维原始比例特征（r_{black}、r_{yellow}、r_{brown}）的基准与含综合得分的变体。像素通道序数分类器选用深度为 3 的决策树（scikit-learn），优势在于决策路径完全可解释，仅需 3 层阈值即可还原 4 级判读规则；小数据集上不易过拟合。

1.3.4　颜色先验注意力通道

在小规模时段伪标签训练下，CNN 倾向于将日

期、光照、植株生长阶段等与危害指数共变的场景属性作为分类依据，而非煤污本身的颜色分布，导致跨场景部署性能之间出现显著落差。CPAM 的设计动机在于：图像的逐像素暗度由 RGB 三通道均值无参数计算，其相对空间分布与叶面黑色霉膜的位置直接对应。将此先验显式注入 CNN 特征学习，可迫使网络在训练初期即将注意力锁定在煤污区域，减少对场景伪相关特征的依赖。与 SE - Net[12] 和 CBAM[13] 等从特征图内部自学习通道注意力权重的标准方法不同，CPAM 注意力输入来自与场景无关的颜色物理特征，而非特征图本身，使得注意力权重不依赖于场景特征统计或背景分布，从理论上保证跨场景条件下注意力稳定性；相关研究在农业图像领域已有成功应用[14]。

掩膜生成（软连续）：对输入图像 $x \in \mathbf{R}^{3\times H\times W}$（分辨率 448 像素 ×448 像素），逐像素计算软连续暗度掩膜 $M \in \mathbf{R}^{1\times H\times W}$，计算前先将三通道归一化至 [0,1]，计算式为

$$M(i,j) = (D(i,j) - D_{\min})/(D_{\max} - D_{\min} + \varepsilon) \quad (5)$$

其中

$$D(i,j) = 1 - (R(i,j) + G(i,j) + B(i,j))/3 \quad (6)$$

式中 (i,j)——像素坐标

$R(i,j)$、$G(i,j)$、$B(i,j)$——像素归一化至 [0,1] 的 3 个通道值

$D(i,j)$——相对暗度（值越大越暗）

$D_{\min}$、$D_{\max}$——图像内 D 最小、最大值

ε——防止分母为零的小常数，取 1×10^{-6}

$M(i,j)$——逐图最小-最大归一化后的软连续暗度掩膜，取 0 ~ 1

以逐图最小-最大归一化的相对暗度掩膜替代绝对阈值二值掩膜，消除 OOD 场景绝对亮度漂移导致的掩膜覆盖率系统偏差；以连续值替代硬阈值，保留像素相对暗度的梯度信息，使 CPAM 从渐变的煤污浓度中学习，而非依赖二值边界判读。

掩膜下采样：将 M 以最近邻插值下采样至与 ResNet - 18 Layer 2 输出特征图相同的空间尺寸（448 像素 ×448 像素输入时约为 56 像素 ×56 像素），得到 $\hat{M}$。

注意力权重生成：$\hat{M}$ 经 2 层卷积变换得到逐通道注意力权重，计算式为

$$A = \sigma(W_2\mathrm{ReLU}(\mathrm{BN}(W_1\hat{M}))) \quad (7)$$

式中 W_1、W_2——2 层卷积核参数

BN——批归一化

ReLU——修正线性单元

σ——Sigmoid 函数

残差特征调制：以残差方式将注意力权重叠加到 Layer 2 输出特征 $F \in \mathbf{R}^{C\times56\times56}$ 上，得到调制特征 F'，计算式为

$$F' = F \odot (1 + A) \quad (8)$$

式中 $\odot$——逐元素乘

残差形式 $(1 + A)$ 确保当 A 接近 0 时模块退化为恒等映射，不破坏原始特征的梯度传播；CPAM 仅含 W_1（3 ×3 卷积，1→16 通道）与 W_2（1 ×1 卷积，16→128 通道）2 层卷积核，参数量仅 2 192 个（不含偏置与 BN 参数），计算量可忽略。

1.4 CPG - CNN 训练与 5 折交叉验证

训练时网络输入图像尺寸为 448 像素 ×448 像素，训练轮数为 80，批量大小为 32，AdamW 优化器[15]，初始学习率为 3×10^{-4}，权重衰减为 5×10^{-4}，余弦退火学习率调度。5 折分层交叉验证，将 DS_train_weak 数据集与 DS_val_internal 数据集合并为 2 000 幅图像池（4 类各 500 幅均匀分布），用固定随机种子（42）做按等级分层的 5 折切分，每折 1 600 幅训练/400 幅验证。对每折独立训练 1 个 CPG - CNN（CPAM 软连续掩膜版）和 1 个 ResNet - 18 基线（无 CPAM），共得到 10 个检查点（5 折 ×2 模式）。每个检查点保留 80 轮训练中验证集 Top - 1 最优的权重。OOD 评估时分别报告 2 类结果：各折模型在 DS1/DS2 上独立推理后取 5 折均值 ± 标准差，反映单模型跨折稳定性；5 折模型在每个样本上的概率平均，反映实际部署中可用的稳健预测。

为验证 CPAM 增量贡献，无 CPAM 的 ResNet - 18 基线与 CPG - CNN 仅是否插入 CPAM 模块上存在差异，其余训练数据、超参数、随机种子、5 折划分均完全相同，构成严格消融。

1.5 双通道融合

双通道在推理阶段通过概率的加权平均实现后期融合，最终预测等级计算式为

$$\hat{y} = \mathrm{argmax}\, P_{fused}(c|x) \quad (9)$$

其中

$$P_{fused}(c|x) = \alpha P_{\mathrm{pixel}}(c|x) + (1-\alpha)P_{\mathrm{cnn}}(c|x) \quad (10)$$

式中 $\hat{y}$——最终预测等级

x——输入冠层图像

c——危害等级（HI - 0 ~ HI - 3）

P_{pixel}、P_{cnn}——像素通道与 CNN 通道输出的类别后验概率

α——融合权重，取 0 ~ 1

P_{fused}——融合后概率

在 $\alpha \in \{0, 0.1, \cdots, 1.0\}$ 上扫描，于 GH1→GH2 OOD 评估集上以 4 级序数评估指标二次加权为目标

选取最优 α；最优 α 代表跨场景部署性能上界，实际部署中需在少量目标域有标注样本上重新校准。

1.6　评价指标

评价指标采用 4 级序数评估指标二次加权 Kappa（Quadratic weighted Kappa，QWK）[16-17]。相较于传统的 Top－1 准确率，QWK 能够根据预测等级与真实等级之间的偏移距离施加非线性惩罚，从而实现模型在序数分级任务中对于等级严重程度偏差的评估。此外还采用 4 级 Top－1 准确率、相邻级准确率（即预测值与真实值之差的绝对值不超过 1 的样本比例）、混淆矩阵，从分类精度、离散偏差分布及类别鲁棒性等多个维度综合评价分级性能。所有指标均按 sklearn 标准实现计算[18]。此外，为评估模型在各危害指数等级上的判别性能，进一步采用精确率（Precision，P）、召回率（Recall，R）与 F1 值（F1-score，F1）。

评估场景分为 3 类：①DS1 内部 5 折分层交叉验证（像素通道），评估 PPC＋决策树在同源数据上的拟合能力。②跨场景 OOD（GH1→GH2，CNN 通道），CPG－CNN 与 ResNet－18 基线均以同一 5 折分层交叉验证协议在 4 摄监控摄像头采集的 2 000 幅时段伪标签图像块上训练（每折 1 600 幅训练/400 幅验证），所得 5 折检查点用于在 DS1（GH1_Jan12，$n=202$）与 DS2（GH2_Apr10，$n=58$）上 OOD 推理；DS2 同时覆盖作物变化（樱桃番茄至黄瓜＋番茄混种）、月份变化（1～4 月）与光照变化，是本文方法学贡献的主验证场景，每个数据集上报告 2 类结果：5 折单模型均值 ± 标准差，与 5 折概率平均。③CNN 跨设备评估，训练数据为监控摄像头图像，OOD 数据为手持 iPhone 采集的 DS1/DS2，反映训练–部署设备不一致下鲁棒性（其中 GH1_Jan12 即 2026 年 1 月 12 日采自温室 GH1 的专家标注子集 DS1_expert，含 202 幅；GH2_Apr10 即 2026 年 4 月 10 日采自温室 GH2 黄瓜＋番茄混种场景的跨场景子集 DS2_expert，含 58 幅）。

2　结果与分析

2.1　颜色像素比例特征跨场景鲁棒性

图 3 为煤污代理特征 r_{black} 在 DS1 与 DS2 上按 4 级危害指数分布。图 3a 中，不同危害指数颜色特征分布清晰，随 HI－等级提升，煤污面积 r_{black} 均值 μ 从 0.044 稳步上升至 0.173，各等级界限分明；而图 3b 显示该特征在跨场景（DS2）部署时严重失效，HI－0 与 HI－1 均值分别为 0.112、0.121，数据点高度重叠，算法无法区分健康与轻度危害。失效的物理本质在于，DS2 环境中健康叶片（HI－0）基准值（0.112）高于 DS1 中轻度受害（HI－1）值（0.095），分离度从 1.67 暴降至 0.28，表明黄瓜叶片低反射率及植株阴影使背景颜色系统性变暗，与煤污特征产生逻辑混淆。

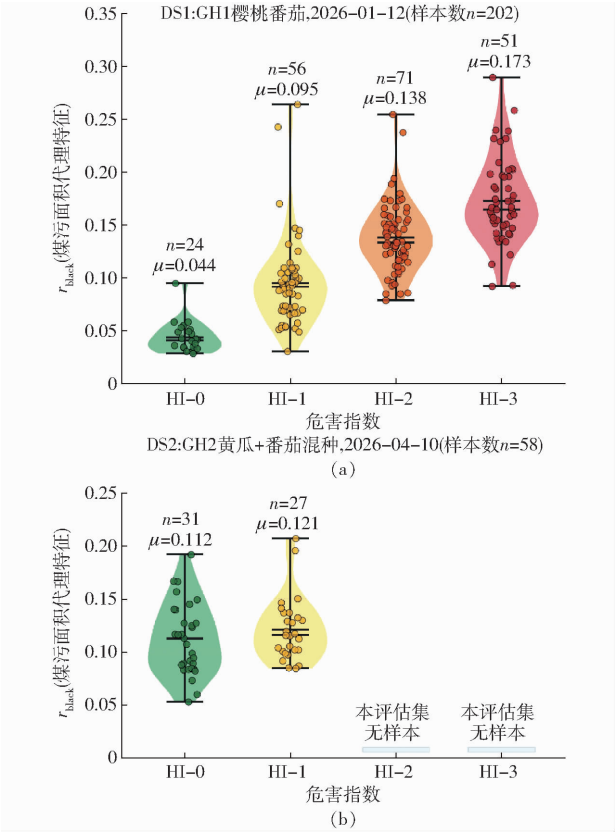

图 3　煤污代理特征 r_{black} 在 DS1 与 DS2 上按 4 级危害指数的分布

Fig. 3　Distribution of sooty-mold proxy r_{black} across 4-level hazard index on DS1 and DS2

通过分析验证了基于 RGB 空间定义的黑色像素比例 r_{black} 作为危害分级指标在不同环境下的适用性。

2.2　CPAM 模块消融试验与跨场景泛化性能

为验证 CPAM 对模型跨场景泛化性能，对 CPG－CNN 三因素在 DS2 上进行消融试验，定量分析注意力机制、掩膜形式及集成策略对分级鲁棒性的影响。5 折分层交叉验证（每折 1 600 幅训练/400 幅/验证，固定种子 42）。编号 1 为无注意力模块的 ResNet－18 基线；编号 3 采用早期绝对阈值（$R<60 \wedge G<60 \wedge B<60$）的二值掩膜版（abs60）；软连续掩膜加入表示本文最终采用的逐图最小–最大归一化暗度连续掩膜。集成未加入时报告 5 折单模型均值 ± 标准差，集成加入时报告 5 折概率平均后的预测。试验结果如表 3 所示，基于 DS1 环境校准的像素阈值模型直接应用于 DS2 环境（即纯像素特征决策树），4 级 QWK 标准差仅为 0.053。这一数值在统计学上接近随机判别水平，其性能退化规律与

图 3 所示的特征空间重叠完全一致。这种跨场景部署的局限性确立了引入物理约束机制的必要性，也是开发 CPAM 模块的直接动因。

表 3 CPG－CNN 三因素消融：DS2（GH2_Apr10）跨场景 OOD 评估

Tab. 3 Three-factor ablation of CPG－CNN: DS2 (GH2_Apr10) cross-scene OOD evaluation

编号	CPAM	软连续掩膜	集成	Top－1/%	QWK	相邻级准确率/%
1	×	—	×	68.3 ±2.3	0.324 ±0.041	96.9 ±2.2
2	×	—	√	70.7	0.401	100
3	√	×	×	62.1 ±8.3	0.268 ±0.106	98.6 ±1.9
4	√	×	√	70.7	0.438	100
5	√	√	×	62.4 ±7.0	0.269 ±0.053	96.2 ±3.1
6	√	√	√	77.6	0.551	100

注：√表示加入该模块，×表示不加入该模块（对软连续掩膜表示采用早期二值掩膜），—表示该配置不适用（无 CPAM 时不存在掩膜）。

由表 3 可知，集成学习是抵消小样本 OOD 环境下注意力权重波动的关键手段。单折评估中 CPAM 的 QWK 略低于基线，这符合注意力机制在有限样本量下对数据切分高度敏感的普遍规律。而软连续掩膜（编号 6）相较于早期二值掩膜（编号 4）的实质性改进在于显著收窄了单折预测的标准差，使其收敛至接近基线的稳健水平。这种稳定性的提升，使得集成架构能够通过概率平均，稳健地从各折模型中提取具有共性的物理主信号。

上述现象可进一步从偏差-方差视角加以阐释：CPAM 以与场景亮度无关的物理暗度作为注意力输入，系统地纠正了基线将 HI－1 误判为 HI－0 的背景偏置（集成 QWK 由 0.401 提升至 0.551），即降低了模型的偏差项；但作为小参数量的附加注意力分支，在 58 幅 OOD 样本与小样本弱监督的折划分下同时抬高了估计的方差。单折读数中方差占主导，故 CPAM 单模型的折间均值 QWK（0.269）略低于基线（0.324）；而软连续暗度掩膜将折间标准差由二值掩膜 0.106 收窄至 0.053（编号 3 与编号 5），使方差向基线的 0.041 收敛。由于偏差降低是系统性的，在 5 折概率平均后依然保留，而近似零均值的随机方差随集成规模按约 $1/k$ 衰减，两者叠加使集成 QWK（0.551）反超基线（0.401）。这正是单折略低、集成更高现象的内在机理。此外，模型在同源验证集上性能与基线基本持平，而跨场景增益显著，证实了 CPG－CNN 优势主要来源于对 OOD 维度的泛化能力提升。通过在特征提取层持续注入逐图像归一化的暗度信号，CPAM 引导模型把颜色物理特征整合进深度特征空间。

图 4 为混淆矩阵分析，性能增益主要源于轻度危害指数（HI－1）召回率的改善，较基线模型提升 22 个百分点。试验证实，基线模型在异质背景下存在明显的判定偏置，易将 HI－1 误判为健康状态（HI－0）。CPAM 模块通过提取物理线索，有效抑制了模型对背景中非因果特征（如光影、日期等）的虚假响应，从而纠正了系统性偏差。从植保决策逻辑分析，由于所有分类误差均被严格限制在相邻等级内，模型在介入防治（HI－1 及以上）与持续观察（HI－0）的临界判定上表现出高度的稳健性，能够有效规避因跨级误判导致的漏防风险，具备支持综合防治（Integrated pest management，IPM）[19] 决策的实际应用价值。

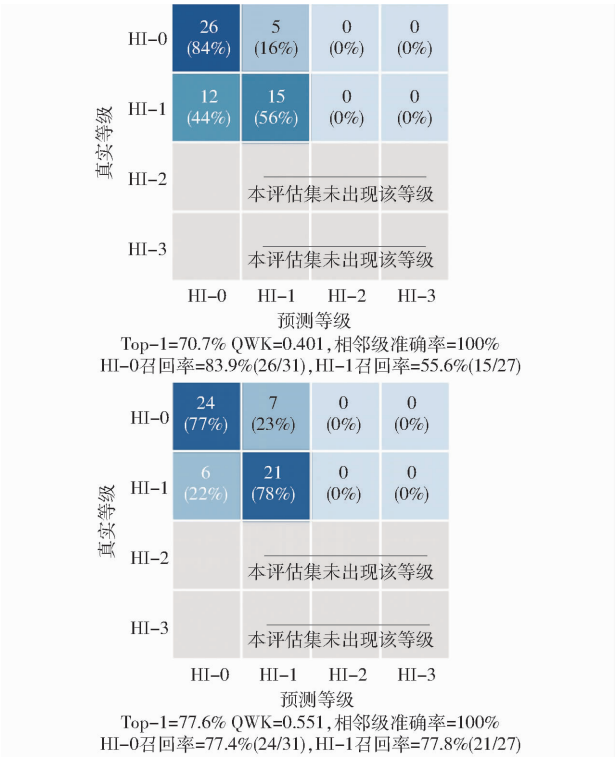

图 4 ResNet－18 基线与 CPG－CNN 在 DS2 OOD 上 5 折集成混淆矩阵结果

Fig. 4 Confusion matrix comparison between 5-fold ensemble of ResNet－18 baseline and CPG－CNN on DS2 OOD

需说明的是，DS2 测试集 HI－2/HI－3 类别样本数为 0（专家标注后全部样本归入 HI－0/HI－1，反映目标温室处于植保介入早期），因此图 4 的 4×4 混淆矩阵在 HI－2/HI－3 行列对应位置为空。为逐级量化 CPAM 对各危害指数等级的判别性能，表 4 为ResNet－18 基线与 CPG－CNN 在 DS2 跨场景 OOD 上逐等级性能（DS2 仅含 HI－0/HI－1 2 级有效样本）。数值基于 5 折集成混淆矩阵（表 3 编号 2 与编号 6）计算；宏平均仅对 DS2 有效的 HI－0/HI－1 2 级取平均。

由表 4 可见，增益集中于 HI－1，召回率由

55.6%提升至 77.8%，与基线相比增加 22.2 个百分点，F1 值由 63.8%提升至 76.4%；HI－0 精确率由 68.4%提升至 80.0%，召回率略降（由 83.9%到 77.4%）；宏平均 F1 值由 69.6%提升至 77.6%，反映 CPAM 纠正基线偏向 HI－0 的系统偏置，2 类判别趋于均衡。

表 4　2 种模型在 DS2 OOD 上各危害指数等级性能

Tab. 4　Per-level performance of two models on DS2 OOD

%

危害指数	模型	精确率	召回率	F1 值
HI－0	ResNet－18 基线	68.4	83.9	75.4
HI－0	CPG－CNN（本文）	80.0	77.4	78.7
HI－1	ResNet－18 基线	75.0	55.6	63.8
HI－1	CPG－CNN（本文）	75.0	77.8	76.4
宏平均	ResNet－18 基线	71.7	69.8	69.6
宏平均	CPG－CNN（本文）	77.5	77.6	77.6

2.3　CPAM 模块与自学习注意力性能试验

表 5 为注意力机制固定 Layer 2 与插入位置（固定 CPAM）在 DS2 OOD 上 5 折交叉验证结果。当各注意力模块均插入 Layer 2 输出端时，其中 SE/CBAM/ECA/Coordinate Attention 的输入为骨干特征图，CPAM 输入为图像颜色物理特征（r_{black}、r_{yellow}、r_{brown}），在准确率方面，CPAM 与 ECA 的集成结果几乎相当，且 5 种注意力机制在 DS2 跨域测试集上准确率彼此接近，说明在本文的数据规模下，不同注意力机制所能达到的准确率上限大体一致；在稳定性方面，CPAM 是 5 种注意力机制中唯一 1 个折间标准差接近基线水平的方案，即在 5 折交叉验证中各折表现最为一致；相比之下，CBAM 在本文数据规模下出现明显退化，单折 QWK 甚至低于基线。CPAM 把与特征图统计独立的图像颜色物理特征（逐图归一化暗度）作为注意力输入，从输入信号上避开了易出错路径，是稳定性优势原因；ECA 在精度上与 CPAM 接近，但缺乏可解释的物理对应，单模型方差远高于 CPAM。综合来看，外部颜色物理特征驱动的注意力，其价值不在于超越自学习注意力的精度上限，而在于保持同等精度的同时，为小样本弱监督任务上单一模型部署提供更可靠的预测。

表 5　注意力机制（固定 Layer 2）与插入位置（固定 CPAM）在 DS2 OOD 上的 5 折交叉验证结果

Tab. 5　Attention mechanism (fixed at Layer 2) and insertion position (fixed to CPAM) under 5-fold CV on DS2 OOD

注意力机制插入位置	模型	参数量	单折 Top－1 准确率（均值±标准差）/%	单折 QWK（均值±标准差）	集成 Top－1 准确率/%	集成 QWK
固定于 Layer 2	ResNet－18 基线	0	68.3±2.3	0.324±0.041	70.7	0.401
	SE－Net[12]	2 048	67.6±2.6	0.372±0.067	74.1	0.477
	CBAM[13]	2 146	57.6±9.2	0.175±0.170	69.0	0.401
	ECA[20]	5	67.2±8.2	0.334±0.167	77.6	0.549
	Coordinate Attention[21]	3 088	69.0±4.0	0.381±0.085	72.4	0.438
	CPAM（本文，Layer 2）	2 192	62.4±7.0	0.269±0.053	77.6	0.551
CPAM 插入位置（注意力机制固定为 CPAM）	Layer 1	112×112	66.6±10.5	0.290±0.227	77.6	0.399
	Layer 2（本文）	56×56	62.4±7.0	0.269±0.053	77.6	0.551
	Layer 3	28×28	61.7±8.5	0.311±0.145	69.0	0.423
	Layer 4	14×14	61.4±10.9	0.230±0.159	74.1	0.508

此外，为判断 CPAM 插入位置对分级性能的影响并确定 CPAM 最优插入层，在保持骨干网络、训练数据、超参数设置与数据划分方式（随机种子 42）完全一致、仅改变注意力模块插入层的前提下，将 CPAM 分别插入 ResNet－18 的 Layer 1～Layer 4 输出端，在 DS2 跨场景 OOD 上测试分级性能。由表 5 可知，当 CPAM 插入 Layer 2 时，DS2 集成 QWK 最高（0.551），优于 Layer 1（0.399）、Layer 3（0.423）与 Layer 4（0.508），与前文的机理分析一致。Layer 2 中层特征既保留了煤污斑块的空间位置、又具备足够语义，最适合与颜色先验耦合，故将颜色先验注入 Layer 2 最为合理。

2.4　CPG－Fusion 双通道融合机制有效性

通过对融合权重 α 进行扫描，验证 CPG－Fusion 双通道融合将像素颜色物理特征整合进 CNN 特征的性能，结果如表 6 所示。在 DS2 跨场景 OOD 任务上，CPG－CNN 集成的 QWK 达到 0.551，高于 YOLO v8m－cls 监督学习基线、EfficientNet－B0 专用分类网络[22]以及单通道 PPC 决策树。后期融合相对早期特征拼接优势在于，无需对骨干网络重训练即可实现既有卷积特征与像素物理特征协同。融合因子 α 作为显式超参数，提升了决策过程可解释性。

图 5 为双通道之间的互补特性，$\alpha=0$ 为纯 CNN

表 6 不同融合方法在 DS2 OOD 上的性能

Tab. 6 Performance of different methods on DS2 OOD

不同融合方法	Top－1 准确率/%	QWK	相邻级准确率/%	备注
经验阈值(YOLO v8m－cls[23])†	53.4	0.150	98.3	DS1 训练
EfficientNet－B0 集成	72.4	0.416	98.3	同量级专用分类网络对照
PPC(3 raw DT,诚实变体)	31.0	0.058	82.8	DS1 4 级训练,主对照
ResNet－18 集成	70.7	0.401	100	表 3 编号 2
PPC＋ResNet－18 集成融合($\alpha=0.2$)‡	70.7	0.445	100	α 扫描最优
CPG－CNN 集成(本文,$\alpha=0$)	77.6	0.551	100	表 3 编号 6
CPG－Fusion(PPC＋CPG－CNN,$\alpha=0.1$)‡	77.6	0.555	100	α 扫描最优

注:†YOLO v8m－cls[23]原始报告基于自动标签 QWK 为 0.443,此处对照专家标签重评。‡后期融合 $P=\alpha P_{\text{pixel}}+(1-\alpha)P_{\text{cnn}}$,$\alpha\in[0,1]$,步长 0.1(详见图 5);同源 DS1 的 PPC 5 折交叉验证 QWK 为 0.728,相邻级准确率 96.5%。

集成;$\alpha=1$ 为纯 PPC。通过对融合权重因子 α 进行步长为 0.1 的参数扫描(即遍历融合权重因子 $\alpha\in[0,1]$ 取最优融合点)。对于缺乏物理约束的 ResNet－18 基线,引入 PPC 像素通道概率可获得显著性能增益,最优融合权重下 QWK 提升至 0.445。然而对于已集成 CPAM 的 CPG－CNN,后期融合的额外性能提升趋于零(最优 α 下 QWK 仅微增至 0.555),说明 CPAM 已在特征提取阶段把像素颜色物理特征整合进 CNN 特征,使模型无需依赖外部参数调优即可在跨场景任务中达到性能上界。

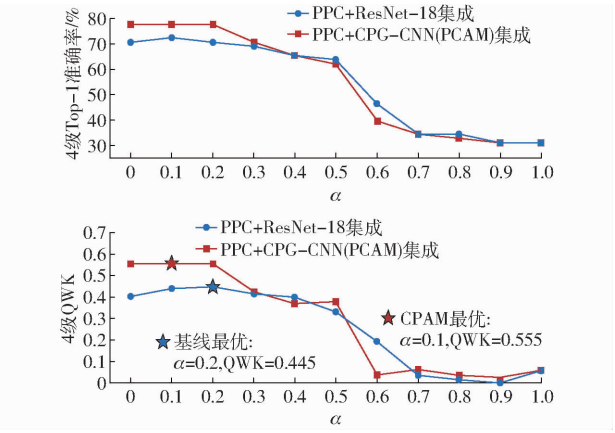

图 5 2 种融合变体在 DS2 OOD 上 α 参数扫描性能曲线

Fig. 5 α-sweep curves of two fusion variants on DS2 OOD

由图 5 可知,若 CPAM 通过软连续掩膜在前向传播中已学到像素颜色信息的判别,则在后期再与 PPC 概率融合应观察到边际收益趋零;反之 ResNet－18 基线缺乏此先验,与 PPC 的互补性应保留可融合空间。2 条曲线的最优 α 均显著偏向 CNN 侧,CPG－Fusion 最优融合点相对纯 CPG－CNN 集成的额外增益接近为零,而 PPC＋ResNet－18 融合相对纯基线集成仍可获得可观增益。这一融合证伪逻辑证实 CPAM 已经在结构上整合了像素颜色物理特征,不需要后期 α 调参也能达到本数据规模下融合可达的近似上限——CPG－CNN 集成单通道由此成为本文推荐的部署形态:无需目标域有标注样本扫描 α,即可在跨场景部署中获得与 CPG－Fusion 等价的预测性能。

2.5 算法决策依据的空间一致性评估

为验证 CPAM 将 CNN 注意力从背景伪相关区域引导至煤污颜色区域,在 DS2 OOD 评估集上对 HI－0 与 HI－1 各取 2 幅代表样本进行试验,结果如图 6 所示,图中 ResNet－18 基线与 CPG－CNN 的 GradCAM 红色覆盖;白虚线圈出经形态学清理后位于叶片附近的深色像素聚集区作为煤污物理参考区,每幅覆盖标注煤污关注比(GradCAM 在参考区内均值/全图均值,大于 1 表示关注集中在煤污区)。

由图 6 可知,通过 Grad－CAM 深度特征可视化分析基准模型与 CPG－CNN 在决策依据上的空间差异,ResNet－18 基线模型的特征激活区(热力图)与真实的煤污区域(物理掩膜)空间对应性较弱,关注点常偏移至叶缘、茎秆连接处或强光高亮区,证实了弱监督训练下 CNN 倾向于将拍摄时段或背景亮度等场景属性作为分级依据,而非真实的病理特征[24]。相比之下,集成 CPAM 模块后,模型的关注区域在相同样本上显著向煤污分布区收敛,背景干扰信号受到有效抑制。这是由于 CPAM 持续将逐像素相对暗度这一颜色物理特征注入注意力路径,强制引导模型回归到煤污本身的颜色表征上,从视觉机理层面解释了前述相邻级误判被消除的原因。定量分析显示,在代表性样本中,HI－1 等级的煤污关注比从基准模型 0.83 提升至 CPG－CNN 的 1.15(增加 0.32),而健康样本(HI－0)因无病征表现,关注比均接近于 1。此外,Grad－CAM 激活区与物理掩膜呈现空间互补而非完全重合的特性(即模型在关注煤污的同时,也兼顾了植株形态特征),进一步证明了颜色物理特征已整合进卷积特征中。

3 结论

(1)提出了颜色先验注意力模块 CPAM。通过把逐图归一化的相对暗度作为与场景亮度无关的物

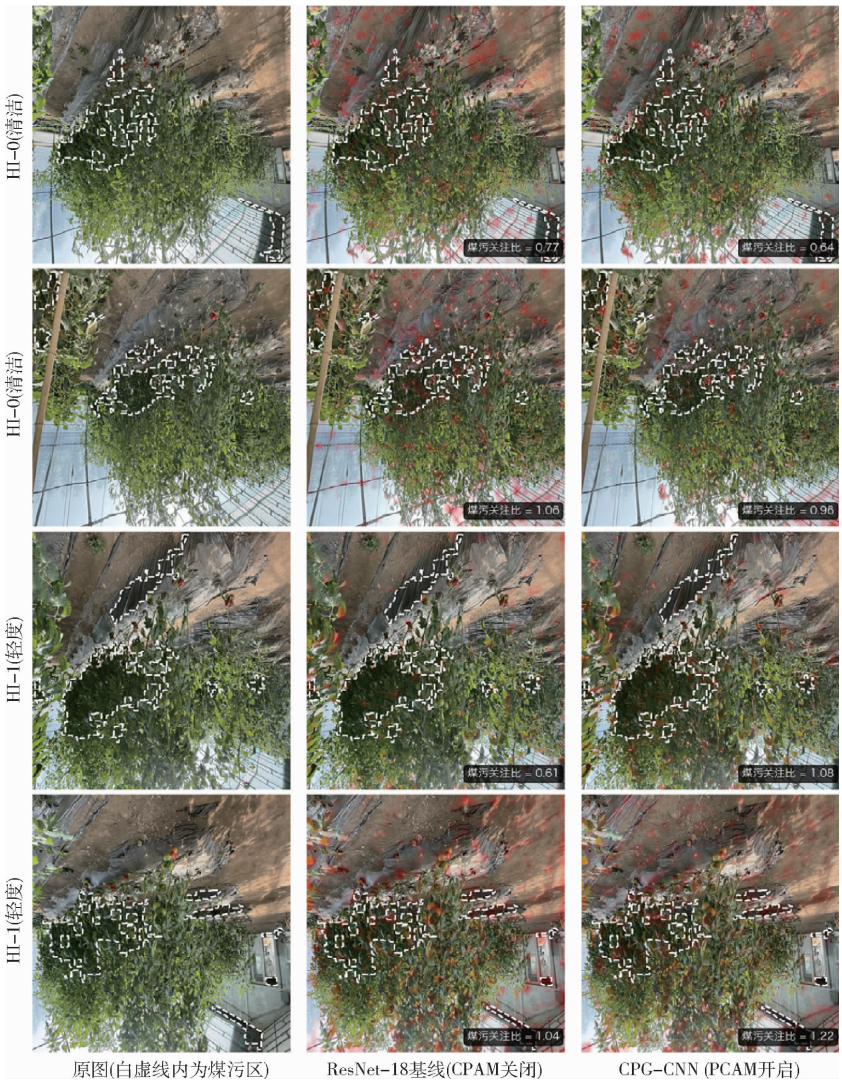

图 6　ResNet－18 基线与 CPG－CNN 的 Grad－CAM 关注区试验结果

Fig. 6　Grad－CAM attention comparison between ResNet－18 baseline (CPAM-off) and CPG－CNN (CPAM-on)

理线索引入 ResNet－18 注意力路径，使模型在小样本弱监督训练中减少对日期、光照等特征的依赖。在 5 折训练、多折概率平均部署下，CPAM 较 ResNet－18 基线在跨场景评估上将 QWK 从 0.401 提升至 0.551，轻度危害召回率与相邻级准确率同步显著改善；与 SE－Net、CBAM、ECA、Coordinate Attention 4 种常用注意力的同协议相比，CPAM 精度与之并列但折间标准差明显更低，在小样本场景下体现了将可解释物理线索引入注意力中稳定性优势。

(2)试验仅覆盖樱桃番茄与黄瓜＋番茄混种两茬作物，DS2 实际仅含低危 2 级，CPAM 单模型折间方差略高于基线、增益主要源于集成平均，且煤污面积与白粉虱虫口密度的定量回归关系仍有待与黄板诱集等直接计数手段同步对照[25]。

参　考　文　献

[1] Wang Guan, Sun Yu, Wang Jianxin. Automatic image-based plant disease severity estimation using deep learning[J]. Computational Intelligence and Neuroscience, 2017, 2017: 2917536.

[2] 翟肇裕, 曹益飞, 徐焕良, 等. 农作物病虫害识别关键技术研究综述[J]. 农业机械学报, 2021, 52(7): 1－18.
Zhai Zhaoyu, Cao Yifei, Xu Huanliang, et al. Review of key techniques for crop disease and pest detection[J]. Transactions of the Chinese Society for Agricultural Machinery, 2021, 52(7): 1－18. (in Chinese)

[3] 赵辉, 黄镖, 王红君, 等. 基于改进 YOLO v7 的农田复杂环境下害虫识别算法研究[J]. 农业机械学报, 2023, 54(10): 246－254.
Zhao Hui, Huang Biao, Wang Hongjun, et al. Pest identification method in complex farmland environment based on improved YOLO v7[J]. Transactions of the Chinese Society for Agricultural Machinery, 2023, 54(10): 246－254. (in Chinese)

[4] 何斌, 张亦博, 龚健林, 等. 基于改进 YOLO v5 的夜间温室番茄果实快速识别[J]. 农业机械学报, 2022, 53(5): 201－208.

He Bin, Zhang Yibo, Gong Jianlin, et al. Fast recognition of tomato fruit in greenhouse at night based on improved YOLO v5 [J]. Transactions of the Chinese Society for Agricultural Machinery, 2022, 53(5): 201 - 208. (in Chinese)

[5] 王志彬, 王开义, 王书锋, 等. 基于动态集成的黄瓜叶部病害识别方法[J]. 农业机械学报, 2017, 48(9): 46 - 52.
Wang Zhibin, Wang Kaiyi, Wang Shufeng, et al. Recognition method of cucumber leaf diseases with dynamic ensemble learning [J]. Transactions of the Chinese Society for Agricultural Machinery, 2017, 48(9): 46 - 52. (in Chinese)

[6] Kamikawa S, Imura T, Sato H. Reduction of sooty mold damage through biocontrol of the greenhouse whitefly Trialeurodes vaporariorum (Hemiptera: Aleyrodidae) using selective insecticides in tomato cultivation greenhouses[J]. Applied Entomology and Zoology, 2018, 53(3): 395 - 402.

[7] Apacionado B V, Ahamed T. Sooty mold detection on citrus tree canopy using deep learning algorithms[J]. Sensors, 2023, 23(20): 8519.

[8] Preanto S A, Ahad M T, Emon Y R. A semantic segmentation approach on sweet orange leaf diseases detection utilizing YOLO [PP/OL]. V1. arXiv(2024 - 09 - 10)[2026 - 05 - 15]. https://doi.org/10.48550/arXiv.2409.06671.

[9] Shoaib M, Shah B, Ei-Sappagh S, et al. An advanced deep learning models-based plant disease detection: a review of recent research[J]. Frontiers in Plant Science, 2023, 14: 1158933.

[10] Selvaraju R R, Cogswell M, Das A, et al. Grad - CAM: visual explanations from deep networks via Gradient-Based localization[C]// Proceedings of the IEEE International Conference on Computer Vision (ICCV), 2017: 618 - 626.

[11] He K M, Zhang X Y, Ren S Q, et al. Deep residual learning for image recognition[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016:770 - 778.

[12] Hu J, Shen L, Sun G. Squeeze-and-excitation networks[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2018:7132 - 7141.

[13] Woo S, Park J, Lee J Y, et al. CBAM: convolutional block attention module[C]// Proceedings of the European Conference on Computer Vision (ECCV), 2018: 3 - 19.

[14] 袁培森, 丁毅飞, 徐焕良. 基于深度主动学习与 CBAM 的细粒度菊花表型识别[J]. 农业机械学报, 2024, 55(2): 258 - 267.
Yuan Peisen, Ding Yifei, Xu Huanliang. Fine-grained chrysanthemum phenotype recognition based on deep active learning and CBAM[J]. Transactions of the Chinese Society for Agricultural Machinery, 2024, 55(2): 258 - 267. (in Chinese)

[15] Loshchilov I, Hutter F. Decoupled weight decay regularization[PP/OL]. V1. arXiv(2019 - 01 - 04)[2026 - 05 - 15]. https://doi.org/10.48550/arXiv.1711.05101.

[16] Cohen J. Weighted Kappa: nominal scale agreement provision for scaled disagreement or partial credit[J]. Psychological Bulletin, 1968, 70(4): 213 - 220.

[17] Landis J R, Koch G G. The measurement of observer agreement for categorical data[J]. Biometrics, 1977, 33(1): 159 - 174.

[18] Pedregosa F, Varoquaux G, Gramfort A, et al. Scikit-learn: machine learning in Python[J]. Journal of Machine Learning Research, 2011, 12: 2825 - 2830.

[19] Barzman M, Bùrberi P, Birch A E, et al. Eight principles of integrated pest management[J]. Agronomy for Sustainable Development, 2015, 35(4): 1199 - 1215.

[20] Wang Q L, Wu B G, Zhu P F, et al. ECA - Net: efficient channel attention for deep convolutional neural networks[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2020: 11531 - 11539.

[21] Hou Q B, Zhou D Q, Feng J S. Coordinate attention for efficient mobile network design[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2021:13713 - 13722.

[22] Tan M X, Le Q V. EfficientNet: rethinking model scaling for convolutional neural networks[C]// Proceedings of the 36th International Conference on Machine Learning (ICML), 2019: 6105 - 6114.

[23] Jocher G, Chaurasia A, Qiu J. Ultralytics YOLO v8[CP/OL]. (2023 - 01 - 10)[2026 - 05 - 08]. https://www.scirp.org/reference/referencepapers? referencecid = 3532980.

[24] Geirhos R, Jacobsen J H, Michaelis C, et al. Shortcut learning in deep neural networks[J]. Nature Machine Intelligence, 2020, 2(11): 665 - 673.

[25] 李善军, 梁千月, 余勇华, 等. 柑橘木虱 YOLO v8 - MC 识别算法与虫情远程监测系统研究[J]. 农业机械学报, 2024, 55(6): 210 - 218.
Li Shanjun, Liang Qianyue, Yu Yonghua, et al. Research on Asian citrus psyllid YOLO v8 - MC recognition algorithm and insect remote monitoring system[J]. Transactions of the Chinese Society for Agricultural Machinery, 2024, 55(6): 210 - 218. (in Chinese)