基于改进YOLOv11n的多尺度茶叶病害检测方法

肖瑞宏[1], 谭立新[1,2*], 王日凤[3*], 宋 敏[1,4], 胡程喜[5]

(1. 湖南农业大学 信息与智能科学技术学院,湖南长沙 410125,中国; 2. 湖南信息职业技术学院 电子工程学院,湖南长沙 410200,中国; 3. 广西科技师范学院 人工智能学院,广西来宾 546199,中国; 4. 长沙幼儿师范高等专科学校,湖南长沙 410000,中国; 5. 湖南软件职业技术大学,湖南湘潭 411100,中国)

摘 要: [**目的/意义**] 传统模型使用传统标准化数据集训练后,对实际识别中较远或较近的多尺度茶叶病害目标存在错检或漏检,以及性能不足的情况。针对茶田病害巡检时存在的茶叶病害检测环境中病害形态多样、识别距离不固定且容易受到背景影响而被误判漏判等问题,本研究提出一种集成多尺度特征分解、可切换空洞卷积与自适应空间融合的改进模型YOLO-SADMFA (YOLO Switchable Atrous Dynamic Multi-Scale Frequency-Aware Adaptive)。[**方法**] 增加模型卷积、特征提取、上采样与检测头轮次加强多尺度能力,提出一种多尺度特征分析解算与动态频率调整融合的动态多尺度频率感知上采样模块进行上采样。首先,该模块可以在有效融合多尺度特征的情况下控制上下采样的信息丢失;其次,引入可切换空洞卷积模块代替原有跨阶段部分核心模块,通过结合不同的空洞率结果进一步加强捕捉目标多尺度信息,同时采用权重锁定机制提升了模型性能;最后,在head结构中引入自适应空间特征融合 (Adaptively Spatial Feature Fusion, ASFF),其技术特性形成ASFF检测头,自适应地学习空间融合权重,有效地过滤相冲突的信息。同时建立了1个含有2 880张图像9种茶叶病害类别的茶叶病害数据集。[**结果和讨论**] 该方法在茶叶病害检测的任务中精确度、召回率和平均精度值分别为89.7%、82.6%和86.3%。YOLO-SADMFA较原版YOLOv11n模型精确度、召回率、平均精度值分别提升4.4、8.4、3.7个百分点,尤其在处理病斑面积占比10%~65%的多尺度目标时表现突出。在低光照、复杂背景等田间实际场景下,模型仍保持较高的检测稳定性,能够有效区分形态相似的病害类型,并在边缘计算设备上实现约161帧/s的实时检测速度。[**结论**] 本研究所提出的YOLO-SADMFA有效解决了茶园复杂环境下多尺度病害检测难题,显著提升了检测准确性和鲁棒性,为自动化茶叶病害巡检系统提供了可靠的技术支持,对促进茶产业智能化发展具有重要应用价值。

关键词: YOLO; 茶叶病害; 目标检测; DMF-Upsample; ASFF

中图分类号: S225.99;TP391.41;TP18 **文献标识码:** A **文章编号:** SA202509014

引用格式:肖瑞宏, 谭立新, 王日凤, 宋敏, 胡程喜. 基于改进YOLOv11n的多尺度茶叶病害检测方法[J]. 智慧农业(中英文), 2026, 8(1): 62-71. DOI: 10.12133/j.smartag.SA202509014

XIAO Ruihong, TAN Lixin, WANG Rifeng, SONG Min, HU Chengxi. Multi-Scale Tea Leaf Disease Detection Method Based on Improved YOLOv11n[J]. Smart Agriculture, 2026, 8(1): 62-71. DOI: 10.12133/j.smartag.SA202509014 (in Chinese with English abstract)

0 引 言

茶产业作为中国重要的传统特色农业,市场需求持续增长。截至2024年,全国茶园总面积为349.52万hm^2,干毛茶总产量349.91万t,总产值3 217.84亿元[1]。然而,伴随农村人口结构不断变迁,茶园日常打理所需的人力资源正在快速下降[2-4]。以往的常态化茶田管理,特别是病害的常态化监测与早期防治,造成部分地区业已陷入近乎无人问津的窘境。但茶树病害对茶叶品质与产量构成的威胁却丝毫未减,其发生往往隐蔽性强、蔓延

收稿日期:2025-09-05

基金项目:中国高校产学研创新基金——新一代信息技术创新项目资助课题(2022IT82);广西科技计划项目(桂科AD23026282);湖南省教育科学规划课题(XJK24BZY037)

作者简介:肖瑞宏,硕士研究生,研究方向为计算机视觉及应用。E-mail:909981044@qq.com

*通信作者:1. 谭立新,硕士,教授,研究方向为机器人与智能系统、模式识别与智能信息处理。E-mail:594637823@qq.com;2. 王日凤,博士,教授,研究方向为图像处理、计算机视觉。E-mail:wrfgm@163.com

copyright©2026 by the authors

迅速，一旦爆发极易导致成片茶树受损甚至绝收[5]。加之茶园环境本身复杂多变，即便经验丰富的农技人员亦难在田间复杂背景下做到精准、高效的肉眼识别与诊断[6]。因此，面对人力短缺的刚性约束与病害防控的迫切需求，研发一种能够在茶园场景中稳定运行、对多类别茶树病害实现自动化识别的计算机视觉技术，已成为保障茶产业可持续高质量发展的关键突破口和当务之急。

近年来，随着计算机视觉检测与深度学习技术的进步和训练与部署视觉模型成本的降低，使用视觉检测模型在病害检测领域逐渐崭露头角。不仅可降低漏检率与误判率、提升产品全链条质量管控水平，更是响应制造业智能化转型、践行高质量发展战略的必然技术路径。在已有研究中，学者从多个方向对病害检测模型进行了改进：一部分工作聚焦于复杂背景下的特征增强与多尺度融合。例如，楚家等[7]提出一种融合了可变形卷积的神经网络，引入可变形卷积与卷积注意力机制模块（Convolutional Block Attention Module, CBAM），结合加权交并比（Weighted Intersection over Union, W-IOU）损失函数增强了对野外害虫的鲁棒检测。白凯等[8]设计了多尺度特征融合网络（Multi-scale features inception Neck, MFI Neck），融入可变形注意力跨阶段部分特征融合（Cross-Stage Partial with 2 convolutions Feature fusion - Deformable attention, C2f-DA）结构与降采样（Adaptive Downsampling, ADown）模块，有效识别复杂环境下花生的相似病害。侯文慧等[9]采用超分辨率生成对抗网络（Super-Resolution Generative Adversarial Network, SRGAN）进行图像增强，并引入小目标检测层与三重注意力机制，强化了对番茄病害小目标的检测性能。魏明飞等[10]则提出多尺度注意力自适应空间融合的渐近特征金字塔网络（Efficient Multi-scale Attention-Asymptotic Feature Pyramid Network, EMA-AFPN）与双分支注意力机制，提升了遮挡情况下黄瓜病害的识别能力。前述方法在野外环境多尺度识别方面有一定进展，但算力要求使得边缘计算设备实时部署变得困难，同时对常见的茶园多种类多尺度识别上仍然不够理想仍然存在错检漏检。而在轻量化与结构优化方面，刘博等[11]设计了一种基于图结构的番茄病害识别框架（Tomato leaf Disease Recognition Enhanced via Graph Structures, TDR-EGS），在不增加模型推理阶段复杂度的前提下有效地提升了分类性能。夏顺兴等[12]使用动态卷积、混合标准卷积与深度可分离卷积（Grouped Shuffle Convolution, GSConv）与Slim-neck结构实现轻量化草莓病害检测。高山[13]添加基于视觉注意力模块（Coordinate and Spatial Attention, CSA），在不损失检测性能的情况下降低了计算的开销。胡艳茹和刘德全[14]基于YOLOv8n（You Only Look Once Version 8n）架构，结合高效多尺度注意力模块（Efficient Multi-scale Attention, EMA）与双向特征金字塔网络（Bidirectional Feature Pyramid Network, BiFPN）实现高效多尺度特征融合。这几种轻量化病害检测方法可满足边缘检测设备的部署，但在多尺度目标识别尤其是复杂背景的情况下准确率不尽如人意。上述方法与模型识别在复杂背景下识别多类别病斑时往往只能识别病斑占图片面积10%～30%的较窄范围的尺度，很难胜任茶田巡检时所面临的复杂野外环境下边缘计算设备部署时多尺度、多类别的茶叶病害检测。

针对上述问题，本研究提出基于YOLOv11n的YOLO-SADMFA（YOLO Switchable Atrous Dynamic Multi-Scale Frequency-Aware Adaptive）算法，在原模型的基础上，通过增加卷积、特征提取、上采样轮次，并设计动态多尺度频率感知上采样模块（Dynamic Multi-scale Frequency-aware Upsample, DMF-Upsample）引入可切换空洞卷积模块（Switchable Atrous Convolution, SAConv）和自适应空间特征融合（Adaptively Spatial Feature Fusion, ASFF）检测头，提高模型在田间复杂环境下检测9种病害多尺度病斑目标的性能，以期为茶叶生产中的病害检测和监控提供技术上的支持。

1　数据采集与处理

1.1　数据样本采集

本研究所使用的茶叶病害数据集由课题组于2024年春季与夏季分多次在湖南省长沙市长沙县高桥溪清茶园采集。使用Xiaomi 14和Nikon D3100对样本进行多尺度高分辨率拍摄。病害面积占比为图像的10%～65%，其中，病斑占比大于等于50%的为高尺度目标，小于50%且大于等于20%为中尺度目标，小于20%为低尺度目标。原始图像分辨率像素为3 072×3 072，后经模型自动统一压缩，像素为640×640。高分辨率图像能尽可能提升模型对于特征提取的学习能力[15]。将拍摄好的图片进行筛选分类后共计2 880张图像9种病害类别，涵盖茶枯叶病、茶炭疽病、茶霉粉病、茶藻斑病、茶饼

病、茶赤星病、云纹枯叶病、茶细蛾寄生病和茶红脉病，病害种类图像见图1。



a. 茶枯叶病 b. 茶炭疽病 c. 茶霉粉病 d. 茶藻斑 e. 茶饼病 f. 茶赤星病 g. 云纹枯叶病 h. 茶细蛾寄生病 i. 茶红脉病

图1 9种茶叶病害类别图

Fig. 1 Nine tea leaf disease categories

同时，在数据集中，每种病害还包含在复杂背景下的特写、远景、晴天、阴雨天图像，见图2。



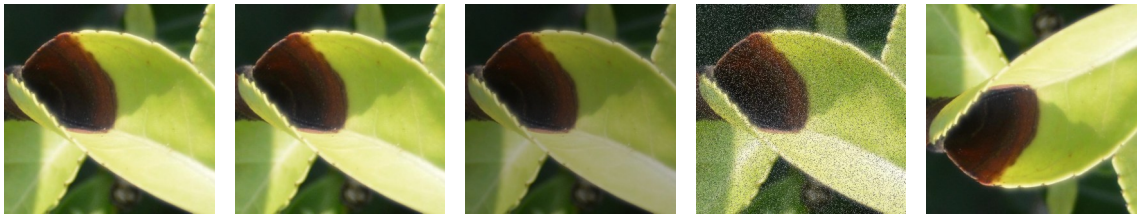
a. 特写 b. 远景 c. 晴天 d. 阴雨天

图2 复杂背景下茶叶病害图像

Fig. 2 Tea leaf diseases images under complex backgrounds

1.2 数据处理

本数据集标签使用LabelImg进行茶叶病害位置与种类的标注。在实际数据集中存在1张图像出现多个或多种病害的情况，标注完成后尽量按照每种标签数量8∶1∶1分成训练集、验证集、测试集。分配完成后对训练集使用随机调整亮度、随机调整对比度、随机加入噪声、随机翻折旋转4种方法进行数据增强，效果如图3所示。实际共得到12 120张图像，训练集、验证集、测试集图像数量分别为11 550、287、283张。



a. 原图 b. 调整亮度图像 c. 调整对比度图像 d. 加入噪声图像 e. 翻折旋转图像

图3 4种增强图像与原图对比图

Fig. 3 Comparison of original images with four data augmentation methods

增强前标签数量总共为3 494个标签，增强后为14 590个标签。其中每种茶叶病害标签数量见表1。

表1 每种茶叶病害标签在数据集中的分布数量

Table 1 Distribution of labels across tea leaf disease categories in the dataset

类别	训练集/个	验证集/个	测试集/个	总数/个
茶枯叶病	2 795	75	69	2 939
茶炭疽病	2 035	51	47	2 133
茶霉粉病	1 060	31	27	1 118
茶藻斑病	1 275	29	32	1 336
茶饼病	835	24	21	880
茶赤星病	1 665	49	42	1 756
云纹枯叶病	2 020	55	51	2 126
茶细蛾寄生病	890	27	22	939
茶红脉病	1 295	36	32	1 363

2 基于YOLOv11n的茶叶病害检测算法

YOLO系列[16-18]算法已成为目标检测领域影响力极大且成功的技术体系[19]。自2016年YOLOv1[20]以来，单阶段（One-stage）检测方法正式登上历史舞台。采用端到端学习方式在图像上并行预测类别与坐标，降低了计算资源依赖的同时减少了检测时间，提高了精准度。YOLOv11是Ultralytics机构最新认证并收录的版本。其核心特点主要体现在轻量化网络结构设计、更高效的多尺度特征融合机制，以及引入针对小目标和遮挡目标的增强策略。YOLOv11包括5个尺寸的版本，分别是YOLOv11n、YOLOv11s、YOLOv11m、YOLOv11l、YOLOv11x，算力需求从低到高各自适合不同的应用场景。

2.1 YOLO-SADMFA构建

本研究选取更适合在巡检装备部署的YOLOv11n为基础改进算法，提出一种YOLO-SADMFA算法。YOLO-SADMFA将骨干网络Backbone中的卷积与特征提取轮次增加，并将C3k2模块替换为SAConv模块。SAConv模块利用空洞卷积的思路，不增加计算量和参数量的前提下在卷积核之间插入空洞来扩大滤波器视野，同时将不同空洞率卷积结果进行结合，加强了对多尺度特征的敏感性。颈部Neck网络增加上采样与特征融合轮次并使用DMF-Upsample代替了原有上采样模块与特征融合模块。利用一种基于小波变换的多尺度特征分解与动态融合来上采样，精准地保留了纹理，强化了语义结构连贯性，加强了多尺度目标的检测性能。最后在检测头Head部分使用4个ASFF检测头模块替换原有检测头模块。通过自适应地学习不同尺度特征的空间特征与融合权重，加强过滤冲突信息，同时增加特征在多尺度下的不变性判断，提升了多尺度检测性能。YOLO-SADMFA模型结构如图4所示。

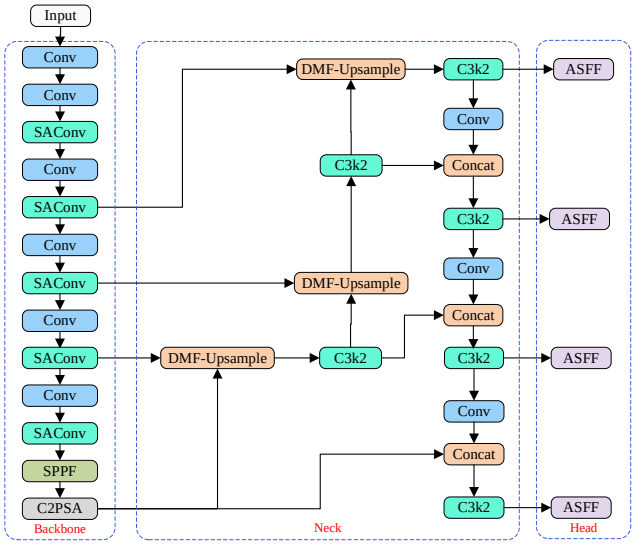


图4　YOLO-SADMFA模型结构图

Fig. 4　Model architecture of YOLO-SADMFA

2.2 DMF-Upsample上采样特征融合模块

在传统编码器-解码器架构中，解码器跨层特征融合常面临频谱混叠问题：直接进行双线性插值或转置卷积等上采样操作在重标定特征尺度时，会破坏叶面纹理等高频分量与轮廓背景等低频分量的能谱分布，导致重建图像出现结构畸变与细节模糊。现有解决方案虽引入跳跃连接传递浅层特征，但未从根本上解决频域干扰传导问题。针对此问题，本研究提出小波特征升级DMF-Upsample上采样特征融合模块，基于离散小波变换（Wavelet Transformation, WT）[21]实现频率感知的特征重组。给定编码器输出的粗尺度特征为F_s与解码器输出的细尺度特征为F_{s+1}，DMF-Upsample模块首先通过小波变换WT将F_s分解为四象限子带，如公式（1）所示。

$$\mathrm{WT}(F_s)=\{A_{LL}^s, H_{LR}^s, V_{RL}^s, D_{RR}^s\} \tag{1}$$

式中：A_{LL}^s为承载低频结构信息；H_{LR}^s，V_{RL}^s，D_{RR}^s分别编码水平、垂直和对角方向的高频细节。子带空间维度自动采样至$\frac{H}{2}\times\frac{W}{2}$与解码器细节尺度$F_{s+1}$几何对齐，无需插值即可实现跨尺度匹配。特征增强阶段，低频通路将A_{LL}^s与F_{s+1}进行Concat拼接，强化全局面部拓扑一致性；同时高频分量经残差块R细化处理，再进行逆小波变换（Inverse Wavelet Transform, IWT）最后获得特征F_s'，如公式（2）所示。

$$F_s'=\mathrm{IWT}\left(\mathrm{Concat}\left(A_{LL}^s, F_{s+1}\right), R\left(H_{LR}^s, V_{RL}^s, D_{RR}^s\right)\right) \tag{2}$$

这种双分支设计在融合过程中显式解耦频率分量，抑制了混叠现象，又完整保留了关键高频特征，通过跨分辨率尺度的无失真特征传递，同时完成了传统上采样与融合模块的工作，有效降低性能开销增加了模型对于跨尺度目标的检测能力。模块结构如图5所示。

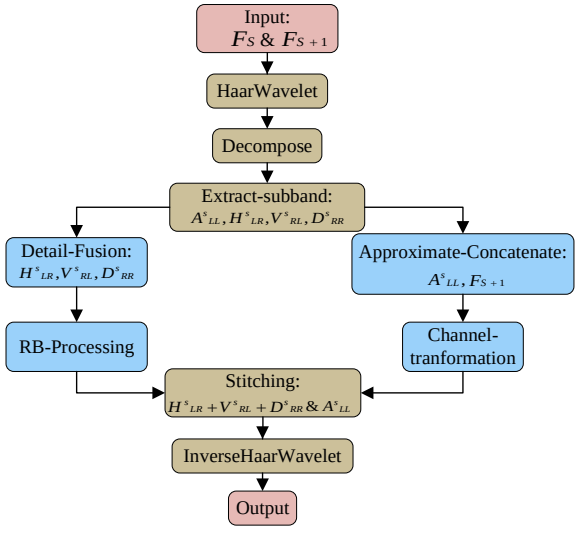


图5　DMF-Upsample上采样特征融合模块结构图

Fig. 5　The architecture of the DMF-Upsample upsampling feature fusion module

2.3 SAConv模块

传统卷积操作受限于固定尺寸的局部感受野，在灰度图像模糊特征提取中无法很好地胜任任务：面对微小目标因感受野过大而丢失细节特征，宏观结构则因局部信息干扰产生定位漂移。空洞卷积SAConv模块[22]，通过结合不同的空洞率的动态多尺度感知与权重锁定机制（一种针对特征的自适配机制）重构特征提取信息。该模块采用3级级联架构：首先，前置全局上下文模块对输入特征进行空间语义建模，建立长程依赖关系以抑制背景噪声；其次，双分支空洞卷积层在权重共享约束下，并行实施差异化采样策略——低空洞率分支聚焦局部纹理精细刻画，高空洞率分支捕获大范围结构关联，同步解耦目标与环境的复杂关系；使用可学习开关函数，其通过门控机制动态评估双路特征贡献度，自适应融合时规避高空洞率导致的网格伪影，同时实现参数效率最大化；最后，后置全局上下文模块对融合特征实施重校准，筛选具有空间判别力的关键信息。在Backbone中使用SAConv代替C3k2模块结合上下文和共享权重可以加强模型对不同尺度目标的泛化能力。SAConv结构如图6所示。SAConv的计算如公式（3）所示。

$$\text{Output} = S(x) \times \text{Conv}(x, y, 1) + (1 - S(x)) \times \text{Conv}(x, y, \Delta y, r) \quad (3)$$

式中：Output为输出信息；x为输入信息；$S(x)$为开关函数；Conv(x，y，Δy，r)是使用权重y进行卷积运算；r为模块超参数；Δy为可训练值。

2.4 基于自适应空间融合ASFF的检测头模块

在复杂场景目标检测任务中，特征金字塔网络（Feature Pyramid Network, FPN）的多尺度融合机制存在着部分问题：浅层特征蕴含高空间分辨率却语义贫乏，深层特征具备丰富语义信息但空间细节丢失，二者直接融合会引发层级间特征冲突。这种空间-语义错位导致梯度反传时出现表征震荡，尤其在对尺度敏感的、微小目标检测中造成定位漂移与漏检。本研究检测头引入自适应空间特征融合ASFF模块[23]，其通过空间动态权重场重构多尺度特征交互逻辑。该模块构建3级处理流：首先将骨干网络提取的Level 1、Level 2、Level 3特征图经可微分插值重采样至统一尺度，消除分辨率差异带来的几何偏差；继而基于轻量级卷积网络生成像素级权重场。使用权重场的空间自适应机制在近距离干扰区域自动降低浅层特征权重、在远距离模糊目标区域提升深层特征贡献度，实现冲突信息的智能过滤。ASFF检测头模块结构如图7所示，计算如公式（4）所示。

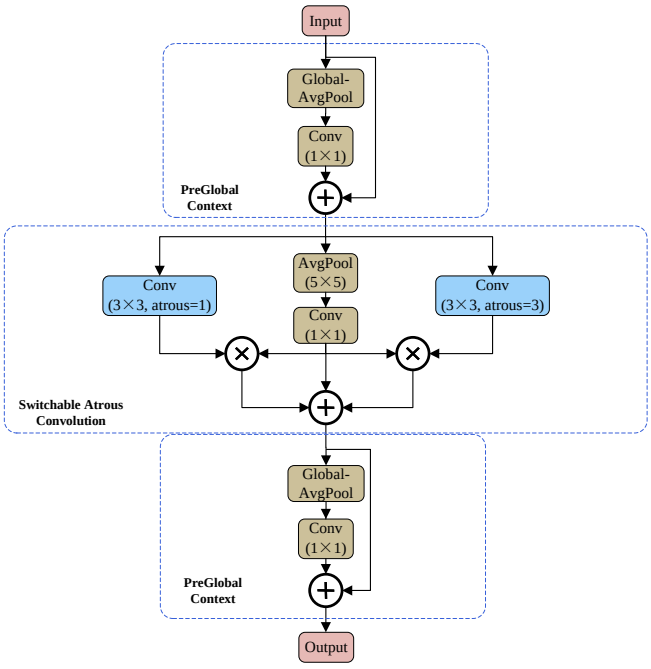


图6 SAConv模块结构图

Fig. 6 SAConv module architecture

$$y_{ij}^{l} = \alpha_{ij}^{l} \times x_{ij}^{1\to l} + \beta_{ij}^{l} \times x_{ij}^{2\to l} + \gamma_{ij}^{l} \times x_{ij}^{3\to l} \quad (4)$$

式中：y_{ij}^{l}为特征输出图结果；α_{ij}^{l}，β_{ij}^{l}，γ_{ij}^{l}为网络自适应学习的空间重要性权重；α_{ij}^{l}为强化浅层特征的边缘响应以锚定叶脉与叶缘等精细结构；β_{ij}^{l}为协调中层特征的部件关联性；γ_{ij}^{l}为增强深层特征的语义一致性以抑制背景虚影；$x_{ij}^{n\to l}$为将第n个尺度的特征图上采样到第1个尺度的特征图。

3 结果与分析

3.1 实验环境

实验中服务器使用Windows 11作为操作系统环境；使用软件为Python 3.11.3；PyTorch 2.6，适配CUDA 12.4、torchvision 0.12；在硬件部分CPU为AMD 7800×3D核处理器，加速频率5 GHz；GPU为RTX4080SUPER；运行内存为64 GB；在模型训练中，图像像素尺寸为640×640，训练轮次为150轮，批次大小为80，初始学习率为0.000 5，最终学习率为0.005，启用随机梯度下降优化器（Sto-

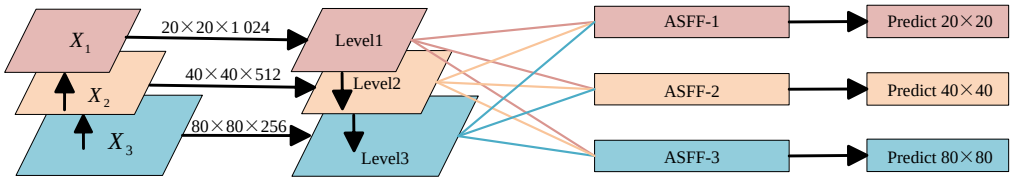


图7　ASFF检测头模块结构图

Fig. 7　ASFF detection head architecture

chastic Gradient Descent, SGD）。

3.2　评价指标

本实验使用3种指标对模型进行评价，分别是精确率（Precision, P）、召回率（Recall, R）、平均精度值（Mean Average Precision, mAP@0.5）。P表示在模型识别出的目标中，检测到正确的目标占全部目标的比值，如公式（5）所示。

$$P = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FP}} \tag{5}$$

式中：TP表示检测到正确目标信息的数量；FP表示模型误判后错误检测的目标数量。

R表示在正确预测为正类别的目标数量与所有实际的真类别目标数量的比值，如公式（6）所示。

$$R = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}} \tag{6}$$

式中：FN表示未能正确预测为正类别的正类别目标数量。

mAP@0.5表示在模型进行目标检测时IoU阈值定为0.5时计算的所有类别目标的平均精度，如公式（7）和公式（8）所示。

$$\mathrm{AP} = \int_0^1 P\mathrm{d}R \tag{7}$$

$$\mathrm{mAP@0.5} = \frac{1}{C}\sum_{k=1}^{C}\mathrm{AP@0.5}_k \tag{8}$$

式中：AP为1个类别的平均精度；C为类别的数量；$\mathrm{AP@0.5}_k$为混淆矩阵IoU阈值取0.5时第k个类别的平均精度。

3.3　消融实验

本研究在YOLOv11n原版模型上逐步引入改进的模块进行消融实验。验证DMF-Upsample上采样特征融合模块、SAConv特征提取模块，以及ASFF多尺度检测头模块对于田间复杂环境下多尺度目标的检测性能。每种具体改进的性能指标见表2。

表2　茶叶病害目标检测研究每种改进方式所对应性能指标

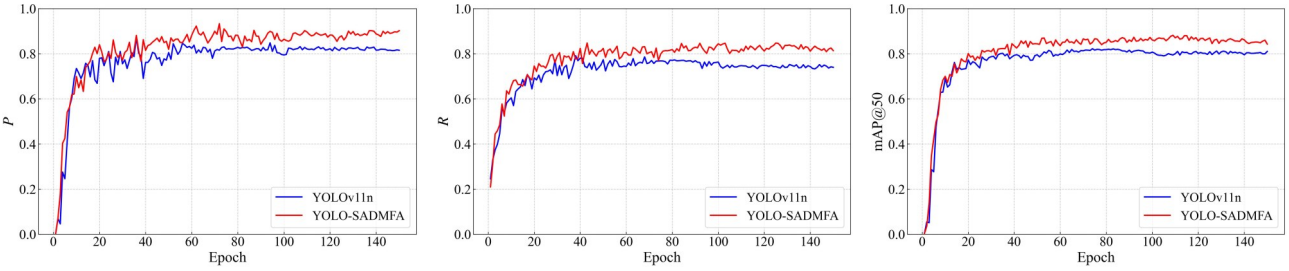
Table 2　Performance metrics of different improvement approaches in tea leaf disease detection

模型	DMF-Upsample	SAConv	ASFF	P/%	R/%	mAP@0.5/%
YOLOv11n	×	×	×	85.3	74.2	82.6
改进1	√	×	×	86.4	78.8	84.2
改进2	×	√	×	87.8	73.4	82.9
改进3	×	×	√	84.7	80.2	84.4
改进4	√	√	×	88.2	81.3	85.5
改进5	√	×	√	87.5	82.1	85.1
改进6	×	√	√	84.4	81.5	81.8
YOLO-SADMFA	√	√	√	89.7	82.6	86.3

注：√表示此改进中使用此模块；×表示此改进中未使用此模块。

由表2可知，引入DMF-Upsample模块后P、R、mAP@0.5分别提升1.1、4.6、1.6个百分点。证明该模块可以有效地在田间复杂环境下对多尺度特征目标分离高频细节和低频结构提升检测准确度减少干扰。将SAConv模块代替C3k2模块后P提升了2.5个百分点，表明该模块可以有效利用空洞卷积获取大范围结构关联提取特征信息提升泛化能力。使用ASFF替换传统检测头之后R和mAP@0.5分别提升了6个百分点与1.8个百分点。表明该模块能有效过滤干扰，减少因为尺度变化带来的尺度信息的误判，能有效地提升模型对于多尺度信息在不同空间对应关系中的敏感性。DMF-Upsample模块分别与SAConv模块、ASFF检测头模块组合显示出了良好的兼容性，3个指标较原模型提升了2.9、7.1、

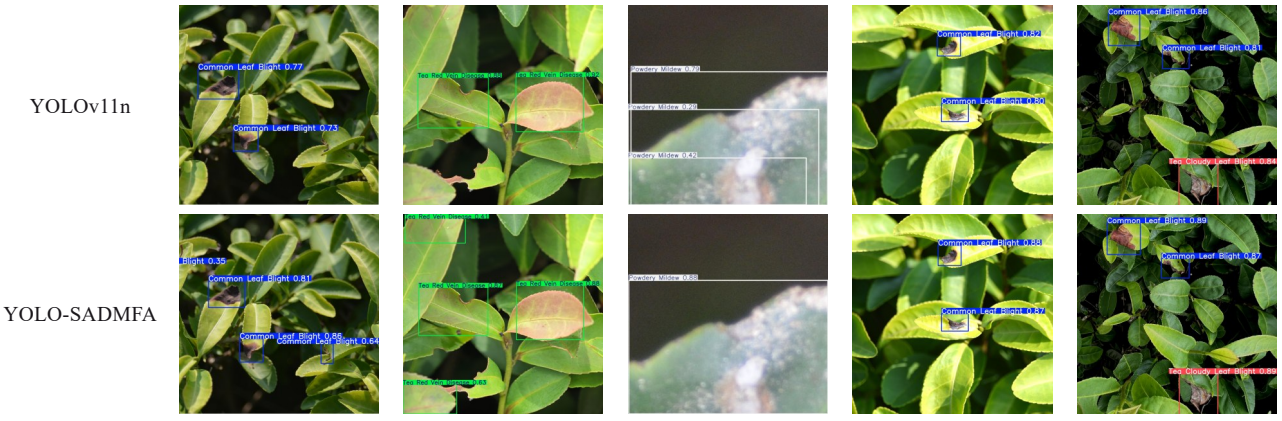
2.9个百分点与2.2、7.9、2.5个百分点。但在缺乏DMF-Upsample模块时，SAConv模块与ASFF模块出现了兼容性问题，*P*和mAP@0.5分别下降了0.9、0.8个百分点。这是因为在缺乏DMF-Upsample上采样特征融合模块分步采样与跨特征融合的情况下，SAConv空洞卷积和全局上下文模块虽然强调了不同尺度的空间信息，但一定程度扰乱了ASFF空间动态权重，进而影响空间自适应机制。实验结果表明，3种模块可有效互补平衡*P*、*R*、mAP@0.5，最终分别达到89.7%、82.6%、86.3%。显著增强了在多尺度复杂场景下的泛化能力和鲁棒性。为了更直观地体现模型改进前后的对比，绘制了YOLOv11n和YOLO-SADMFA训练过程中的收敛曲线，如图8所示。同时展示YOLOv11和YOLO-SADMFA对几组图像的实际检测情况，如图9所示。



a. 两种模型*P*值收敛折线图 b. 两种模型*R*值收敛折线图 c. 两种模型mAP@0.5值收敛折线图

图8 YOLO-SADMFA与YOLOv11n在训练过程中的收敛曲线对比图

Fig. 8 Training convergence comparison between YOLO-SADMFA and YOLOv11n



a. 低尺度图像对比 b. 中尺度图像对比 c. 高尺度图像对比 d. 晴天图像对比 e. 阴雨天图像对比

图9 YOLO-SADMFA与YOLOv11n茶叶病害目标检测效果可视对比

Fig. 9 Visual comparison of YOLO-SADMFA versus YOLOv11n for tea disease detection

从图8可以看出，YOLO-SADMFA比原版模型有明显的性能提升，*P*、*R*、mAP@0.5三个主要指标在20轮之后展示出明显的性能优势，并且在100轮时模型收敛曲线已经趋于平滑，收敛良好处于饱和状态验证了改进后模型的有效性。通过图9可以看出，YOLO-SADMFA较原版模型在复杂背景下低尺度、中尺度与高尺度的漏检错检问题有了明显改善，同时在晴天和阴雨天的*R*有明显提高。改进后的模型在面对复杂环境下的多尺度目标有很强的泛化能力和准确性，既提高了精准程度还降低了漏检率和错检率，具有应用潜力。

3.4 对比实验

由表3可知，通过与其他主流模型的*P*、*R*、mAP@0.5、每秒浮点运算量、帧率5个指标进行对比，本研究提出的YOLO-SADMFA模型在野外复杂环境多尺度目标检测任务下，在*P*、*R*、显著提升的情况下浮点运算量仅为6.9 GFLOPs，帧数保持在每秒161张图像以上。RT-DETR-R18虽在*P*、*R*、mAP@0.5方面表现较好，但57 GFLOPs的浮点运算量并不适合部署在边缘低算力平台。实验证明YOLO-SADMFA在算力需求没有大幅加强的情况下，在处理不同尺度多种类病害，以及模糊背景较

难辨认特征的情况下拥有更好的分辨能力，且准确性、鲁棒性和泛化能力显著高于其他传统模型，适合搭载在实际边缘检测设备。

表3　主流图像检测模型在自建茶叶病害数据集上性能对比

Table 3　Performance comparison of mainstream image detection models on self-built tea disease dataset

模型名称	*P*/%	*R*/%	mAP@0.5/%	FLOPs/G	FPS
Single Shot MultiBox Detector	83.1	67.8	71.9	63.0	41
Faster R-CNN	81.6	65.6	70.1	207.0	24
RT-DETR-R18	90.1	82.2	85.7	57.0	53
YOLOv5n	80.8	71.4	79.3	4.2	192
YOLOv6n	81.5	68.9	79.7	11.5	124
YOLOv7-tiny	80.1	70.6	78.2	13.3	94
YOLOv8n	83.7	73.0	81.8	8.1	170
YOLOv9-tiny	81.3	70.1	80.5	10.7	162
YOLOv10n	85.1	74.4	80.9	8.2	167
YOLOv11n	85.3	74.2	82.6	6.6	179
YOLO-SADMFA	89.7	82.6	86.3	6.9	161

4　结　论

针对复杂茶园环境中茶叶病害检测面积小、形态多样、分布区域随机性大且易受背景干扰导致误判漏判等问题，本研究提出了一种基于YOLOv11n的YOLO-SADMFA检测方法。

实验表明，各模块均能独立带来性能增益，且三者联合使用时表现出良好的互补性，最终YOLO-SADMFA在*P*、*R*和mAP@0.5上分别达到89.7%、82.6%和86.3%，较原始YOLOv11n模型提升4.4、8.4和3.7个百分点。对比实验进一步验证了本方法在多种主流YOLO模型中的优越性，尤其在多类别、小目标和复杂背景场景下表现出更强的识别稳定性与更高的综合检测精度。消融结果表明，仅引入DMF-Upsample后，*P*、*R*和mAP分别提升1.1、4.6和1.6个百分点；缺少DMF-Upsample时，SAConv+ASFF组合导致*P*与mAP下降，凸显DMF-Upsample对跨尺度融合的关键作用。

训练与可视化结果同样与量化指标一致：YOLO-SADMFA在约20轮后即表现出明显优势，至100轮收敛平滑；在复杂背景与多尺度下显著降低漏检与误检，低/中/高尺度目标及晴天与阴雨工况的*R*均得到改善。在效率与部署方面，YOLO-SADMFA仅需6.9 GFLOPs且可达约161帧/s，相较RT-DETR-R18等高算力方案实现更优的精度-算力折中，更适合边缘设备实时检测。

综上所述，YOLO-SADMFA不仅在理论层面创新性地融合了频域分析、动态卷积与自适应特征融合机制。在实际应用层面也为田间自动化茶叶病害巡检装备提供了高效、可靠的视觉检测技术支持。结合上述量化与定性结果，本研究方法在保证实时性的同时显著提升了多尺度、复杂背景条件下的检测准确性与鲁棒性，可作为茶园巡检系统的核心感知方案参考。

利益冲突声明：本研究不存在研究者以及与公开研究成果有关的利益冲突。

参考文献：

[1] 许咏梅, 胡临风, 王慧慧. 中国茶产业数字化赋能对茶叶出口质量影响的实证研究: 基于25个茶叶出口省份的实证分析[J]. 茶叶, 2024, 50(3): 133-144.
XU Y M, HU L F, WANG H H. An empirical study on the influence of digital empowerment of China's tea industry on the quality of tea export: Empirical analysis based on 25 tea export provinces[J]. Journal of Tea, 2024, 50(3): 133-144.

[2] 侯苗. 年龄分层视角下农业人口转移对农地流转的影响研究[D]. 哈尔滨: 东北林业大学, 2024.
HOU M. Study on the influence of agricultural population transfer on agricultural land transfer from the perspective of age stratification[D]. Harbin: Northeast Forestry University, 2024.

[3] 李鹏. 中国农业劳动节约型技术进步对农业人口转移数量的影响[J]. 统计与决策, 2019, 35(20): 99-102.
LI P. Influence of agricultural labor-saving technological progress on the number of agricultural population transfer in China[J]. Statistics & Decision, 2019, 35(20): 99-102.

[4] 李支立. 政策变迁视角下农业转移人口市民化研究: 以湖北省为例[D]. 长春: 吉林大学, 2023.
LI Z L. Research on urbanization of agricultural transfer population from the perspective of policy change: A case study of Hubei province[D]. Changchun: Jilin University, 2023.

[5] 杜英杰, 宗哲英, 王祯, 等. 农作物病害诊断方法现状和展望[J]. 江苏农业科学, 2023, 51(6): 16-23.
DU Y J, ZONG Z Y, WANG Z, et al. Current situation and prospect of diagnostic methods for crop diseases[J]. Jiangsu Agricul-

tural Sciences, 2023, 51(6): 16-23.
[6] 孙艳歌, 吴飞, 姚建峰, 等. 多尺度自注意力特征融合的茶叶病害检测方法[J]. 农业机械学报, 2023, 54(12): 308-315.
SUN Y G, WU F, YAO J F, et al. Tea disease detection method with multi-scale self-attention feature fusion[J]. Transactions of the Chinese Society for Agricultural Machinery, 2023, 54(12): 308-315.
[7] 楚家, 肖敏, 周迅, 等. 基于改进YOLOv8的复杂环境下农田害虫检测算法[J]. 江苏农业科学, 2025, 53(16): 192-204.
CHU J, XIAO M ZHOU X, et al. An algorithm for detecting farmland pests in complex environment based on improved YOLOv8[J]. Jiangsu Agricultural Sciences, 2025, 53(16): 192-204.
[8] 白凯, 张玉杰, 苏邓文, 等. 基于改进YOLOv8n的花生叶片病害检测方法[J]. 农业机械学报, 2025, 56(6): 518-526, 564.
BAI K, ZHANG Y J, SU D W, et al. Peanut leaf disease detection method based on improved YOLOv8n[J]. Transactions of the Chinese Society for Agricultural Machinery, 2025, 56(6): 518-526, 564.
[9] 侯文慧, 龚昌智, 曹文昊, 等. 基于超分辨率增强与改进YOLOv8的番茄叶片病害检测[J]. 农业工程学报, 2025, 41(16): 211-220.
HOU W H, GONG C Z, CAO W H, et al. Tomato leaf disease detection method based on super-resolution enhancement and improved YOLOv8[J]. Transactions of the Chinese Society of Agricultural Engineering, 2025, 41(16): 211-220.
[10] 魏明飞, 郭威, 朱华吉, 等. 复杂环境下改进YOLOX的设施黄瓜病害检测方法[J]. 中国农机化学报, 2025, 46(8): 112-120, 155.
WEI M F, GUO W, ZHU H J, et al. Improving the facility cucumber disease detection method of YOLOX in complex environments[J]. Journal of Chinese Agricultural Mechanization, 2025, 46(8): 112-120, 155.
[11] 刘博, 王斌成, 陶旭, 等. 基于图结构增强的番茄叶部病害识别方法[J]. 中国农机化学报, 2025, 46(5): 125-132.
LIU B, WANG B C, TAO X, et al. Tomato leaf disease recognition method based on enhanced graph structure[J]. Journal of Chinese Agricultural Mechanization, 2025, 46(5): 125-132.
[12] 夏顺兴, 倪铭, 罗友璐, 等. 基于改进YOLOv8n的草莓叶片病害检测方法[J]. 江苏农业学报, 2025, 41(4): 664-675.
XIA S X, NI M, LUO Y L, et al. Strawberry leaf disease detection method based on improved YOLOv8n[J]. Jiangsu Journal of Agricultural Sciences, 2025, 41(4): 664-675.
[13] 高山. 基于注意力机制的茶叶病害识别技术研究[J]. 黄山学院学报, 2025, 27(2): 13-18.
GAO S. Research on tea disease identification technology based on attention mechanism[J]. Journal of Huangshan University, 2025, 27(2): 13-18.
[14] 胡艳茹, 刘德全. 基于GDDL-YOLOv8n的番茄叶病害轻量化检测算法[J]. 电子测量技术, 2025, 48(18): 29-40.
HU Y R, LIU D Q. Lightweight detection algorithm for tomato leaf diseases based on GDDL-YOLOv8n[J]. Electronic Measurement Technology, 2025, 48(18): 29-40.
[15] WU Z Z, WANG X F, ZOU L, et al. Hierarchical object detection for very high-resolution satellite images[J]. Applied Soft Computing, 2021, 113: 107885.
[16] GE Z, LIU S T, WANG F, et al. YOLOX: Exceeding YOLO series in 2021[EB/OL]. arXiv: 2107.08430, 2021.
[17] REDMON J, FARHADI A. YOLO9000: Better, Faster, Stronger [C]// 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, New Jersey, USA: IEEE, 2017.
[18] SELVAM P, SARAVANAN P, MARIMUTHU M, et al. PSDNet: A breakthrough parking space detection network powered by YOLOv8[C]// 2024 International Conference on Advances in Computing, Communication and Applied Informatics (ACCAI). Piscataway, New Jersey, USA: IEEE, 2024: 1-7
[19] 王琳毅, 白静, 李文静, 等. YOLO系列目标检测算法研究进展[J]. 计算机工程与应用, 2023, 59(14): 15-29.
WANG L Y, BAI J, LI W J, et al. Research progress of YOLO series target detection algorithms[J]. Computer Engineering and Applications, 2023, 59(14): 15-29.
[20] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: Unified, real-time object detection[C]// 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, New Jersey, USA: IEEE, 2016: 779-788.
[21] DENG H, ZHANG Y, DUAN X G. Wavelet transformation-based fuzzy reflex control for prosthetic hands to prevent slip[J]. IEEE Transactions on Industrial Electronics, 2017, 64(5): 3718-3726.
[22] QIAO S Y, CHEN L C, YUILLE A. DetectoRS: Detecting objects with recursive feature pyramid and switchable atrous convolution [C]// 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, New Jersey, USA: IEEE, 2021: 10208-10219.
[23] WOO S, PARK J, LEE J Y, et al. CBAM: Convolutional block attention module[C]// Computer Vision-ECCV 2018. Cham, Germany: Springer, 2018: 3-19.

Multi-Scale Tea Leaf Disease Detection Method Based on Improved YOLOv11n

XIAO Ruihong[1], TAN Lixin[1,2*], WANG Rifeng[3*], SONG Min[1,4], HU Chengxi[5]

(1. *College of Information and Intelligence*, *Hunan Agricultural University*, *Changsha* 410125, *China*; 2. *School of Electrical and Electronic Engineering*, *Hunan College of Information*, *Changsha* 410200, *China*; 3. *School of Artificial Intelligence*, *Guangxi Science & Technology Normal University*, *Laibin* 546199, *China*; 4. *Changsha Preschool Education College*, *Changsha* 410000, *China*; 5. *Hunan Software Vocational and Technical University*, *Xiangtan* 411100, *China*)

Abstract: [Objective] Preventing and containing leaf diseases is a critical component of tea production, and accurate identification and localization of symptoms are essential for modern, automated plantation management. Field inspection in tea gardens poses distinctive challenges for vision-based detection algorithms: targets appeared at widely varying scales and morphologies under complex backgrounds and unfixed acquisition distances, which easily misled detectors. Models trained on standardized datasets with uniform distance and background often underperform, leading to false alarms and missed detections. To support method development under realistic constraints, YOLO-SADMFA (You Only Look Once-Switchable Atrous Dynamic Multi-scale Frequency-aware Adaptive), a detector based on the YOLOv11n backbone was proposed. The architecture aims to preserve fine details during repeated re-sampling (down- and up-sampling), strengthen modeling of lesions at varying scales, and refine multi-scale feature fusion. **[Methods]** The pro-

posed architecture incorporated additional convolutional, feature extraction, upsampling, and detection head stages to better handle multi-scale representations, and a DMF-Upsample (Dynamic Multi-scale Frequency-aware Upsample) module that performed upsampling through multi-scale feature analysis and dynamic frequency adjustment fusion was introduced. This module enabled efficient multi-scale feature integration while effectively mitigating information loss during up- and down-sampling. Concretely, the DMF-Upsample analyzed multi-frequency responses from adjacent pyramid levels and fused them with dynamically learned frequency-selective weights, which preserved high-frequency lesion boundaries and textures while retaining low-frequency contextual structure such as leaf contours and global shading. A lightweight gating mechanism estimates per-location and per-channel coefficients to regulate the contribution of different bands, and a residual bypass preserved identity information to further reduce aliasing and oversmoothing introduced by repeated resampling. Furthermore, the baseline C3k2 block was replaced with a switchable atrous convolution (SAConv) module, which enhanced multi-scale feature capture by combining outputs from different dilation rates and incorporates a weight locking mechanism to improve model stability and performance. In practice, the SAConv aggregated parallel atrous branched at multiple dilation factors through learned coefficients under weight locking, which expanded the effective receptive field without sacrificing spatial resolution and suppressed gridding artifacts, while incurring modest parameter overhead. Lastly, an adaptive spatial feature fusion (ASFF) mechanism was integrated into the detection head, forming an ASFF-Head that learned spatially varying fusion weights across different feature scales, effectively filters conflicting information, and strengthens the model's robustness and overall detection accuracy. Together, these components formed a deeper yet efficient multi-scale pathway suited to complex field scenes. **[Results and Discussions]** Compared with the original YOLOv11n model, YOLO-SADMFA improved precision, recall, and mAP by 4.4, 8.4, and 3.7 percentage points, respectively, indicating more reliable identification and localization across diverse field scenes. The detector was particularly effective for multi-scale targets where the lesion area occupied approximately 10%—65% of the image, reflecting the variability introduced by unfixed acquisition distance during tea garden patrols. Under low illumination and in complex backgrounds with occlusions and clutter, it maintained stable performance, reduced both missed detections and false alarms, and effectively distinguished disease categories with similar morphology and color. On edge computing devices, it sustained about 161 f/s, which met real-time requirements for mobile inspection robots and portable systems. These outcomes demonstrated strengthened robustness to background interference and improved sensitivity at extreme scales, which was consistent with practical demands where the acquisition distance was not fixed. From an ablation perspective, DMF-Upsample preserved high-frequency lesion boundaries while retaining low-frequency structural context after resampling, SAConv expanded receptive fields through multi-dilation aggregation under a weight-locking mechanism, and the ASFF-Head mitigated conflicts among feature pyramids. Their combination yielded cumulative gains in stability and accuracy. Qualitative analyses further supported the quantitative results: Boundary localization improved for small, speckled lesions, large blotches were captured with fewer spurious edges, and distractors such as veins, shadows, and soil textures were less frequently misclassified, confirming the benefits of dynamic multi-scale frequency-aware fusion and adaptive spatial weighting in real field conditions. **[Conclusions]** The proposed YOLO-SADMFA effectively addressed the multi-scale disease detection challenge in complex tea garden environments, where acquisition distance was not fixed, lesion morphology and color were diverse, and cluttered backgrounds easily caused misjudgments and omissions. It significantly improved detection accuracy and robustness relative to the original YOLOv11n model across a wide range of target scales, and it maintained stable performance under low illumination and complex backgrounds typical of field inspections. It provided reliable technical support for automated tea leaf disease inspection systems by enabling accurate localization and identification of lesions in real operating conditions and by sustaining real-time inference on edge devices suitable for patrol-style deployment.

Key words: YOLO; tea leaf diseases; object detection; DMF-Upsample; ASFF

Foundation items: Innovation Fund for University-Industry Cooperation in China-Supported Project for New Generation Information-Technology Innovation (2022IT82); Guangxi Science and Technology Program (AD23026282); Hunan Provincial Educational Science Planning Project (XJK24BZY037)

First author: XIAO Ruihong, E-mail: 909981044@qq.com

***Corresponding author:** 1. TAN Lixin, E-mail: 594637823@qq.com; 2. WANG Rifeng, E-mail: wrfgm@163.com