

显微图像中花粉颗粒的自动检测方法

李子瑶, 何进荣

延安大学数学与计算机科学学院, 陕西 延安

收稿日期: 2026年5月13日; 录用日期: 2026年6月18日; 发布日期: 2026年6月26日

摘要

针对显微图像花粉颗粒检测中人工效率低、主观性强及小目标边界模糊问题, 本文提出高精度自动检测方法, 构建了延安地区6类花粉显微图像数据集, 完成912张原始图像、1351个目标标注, 并通过亮度扰动、噪声注入、旋转翻转扩增至9120张; 同时在YOLOv8基础上提出BiFPN-ATFL和Wavelet-CG-C2f两种改进模型。其中BiFPN-ATFL在多尺度检测中实现稳定增益, Wavelet-CG-C2f在复杂背景下平均检测精度达80.3%, 召回率达82.1%, 松科花粉mAP达98.0%。本研究利用小波卷积与上下文引导机制有效缓解高频细节丢失和密集粘连干扰, 为花粉浓度预报与过敏原筛查提供了可靠技术路径。

关键词

花粉检测, YOLOv8, 双向特征金字塔网络, 小波卷积, 上下文引导

Automatic Detection Methods for Pollen Granules in Microscopic Images

Ziyao Li, Jinrong He

School of Mathematics and Computer Science, Yan'an University, Yan'an Shaanxi

Received: May 13, 2026; accepted: June 18, 2026; published: June 26, 2026

Abstract

Addressing the inefficiency and strong subjectivity of manual inspection, and the problem of blurred boundaries in small-target pollen detection from microscopic images, this paper develops a high-precision automatic detection method. A microscopic image dataset of six pollen types from the Yan'an region is first constructed, involving 912 raw images with 1,351 annotated targets, and is expanded to 9120 samples through brightness perturbation, noise injection, rotation and flipping. Two improved YOLOv8-based models, BiFPN-ATFL and Wavelet-CG-C2f, are proposed. BiFPN-ATFL delivers reliable gains in multi-scale detection, while Wavelet-CG-C2f achieves an average detection

accuracy of 80.3% and a recall of 82.1% under complex backgrounds, with the mAP for Pinaceae pollen reaching 98.0%. By integrating wavelet convolution and a context-guided mechanism, the approach effectively alleviates high-frequency detail loss and interferences from dense particle adhesion, offering a dependable technical pathway for pollen concentration forecasting and allergen screening.

Keywords

Pollen Detection, YOLOv8, Bidirectional Feature Pyramid Network, Wavelet Convolution, Context Guidance

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

1.1. 研究背景与问题提出

花粉颗粒形态信息与植物种类、生态适应及过敏源传播密切相关[1]-[5], 空气花粉浓度监测对临床防护和公共健康预警具有直接价值[6] [7]。传统人工镜检依赖经验, 存在效率低、重复性不足和统计偏差大等问题。随着计算机视觉技术发展, 自动化检测成为花粉监测的主要方向[8]。然而显微场景具有高分辨率、浅景深和复杂背景等特点, 花粉颗粒通常尺度小、边界弱、类别间形态相近且易粘连[9], 导致现有检测器在识别准确率与召回率之间难以平衡。

1.2. 研究思路与创新点

本文围绕“数据 - 模型 - 评估”三层路径开展研究: 首先构建区域化花粉数据集并形成统一标注规范; 其次在 YOLOv8 上提出两类针对显微小目标的结构改进; 最后通过多模型对比和消融分析验证关键模块贡献。创新点主要包括: 一, 引入 BiFPN 可学习融合权重增强跨尺度语义流动; 二, 通过 ATFL 提升难样本和类别不均衡下的优化稳定性; 三, 利用 Haar 小波卷积保留高频边缘细节并结合上下文引导模块, 提高密集粘连场景下的目标区分能力。

2. 方法

2.1. 数据集构建与预处理

数据采集采用 EX21 生物显微镜(目镜 4 倍、物镜 10 倍)与 500 万像素电子目镜相机, 单幅图像分辨率为 2048×1536 。数据包含菊科蒿属、禾本、木犀科、松科、杨柳、银杏 6 类花粉, 共 912 张原始图像和 1351 个目标实例(见表 1)。标注工具为 LabelMe, 采用双人交叉复核机制, 低于一致率阈值的样本返工重标。

2.2. 数据增强与实验流程

为抑制过拟合并提升泛化能力, 本文采用噪声增强、亮度变换、高斯滤波、旋转与翻转等方法进行样本扩增。旋转变换坐标表达为:

$$x_b = x_a * \cos \theta - y_a * \sin \theta \quad (1)$$

Table 1. Statistical table of pollen morphological characteristics
表 1. 花粉颗粒形态特点统计表

序号	名称	花期	形态特点
1	菊科蒿属	7~9 月	球形或近球形, 2~3 孔沟, 表面颗粒状纹饰
2	禾本	4~9 月	扁球形, 单孔明显
3	木犀科	6~8 月	近球形, 3 孔沟, 表面网状纹饰
4	松科	4~5 月	近圆形或椭圆形, 双气囊明显
5	杨柳	3~4 月	近球形无萌发孔, 外壁较薄
6	银杏	4~5 月	极面椭圆, 侧面船形, 单沟开裂

$$y_b = x_a * \sin \theta + y_a * \cos \theta \quad (2)$$

样本总数为 9120 张, 按 60%、20%、20%划分训练集、验证集与测试集, 保证模型比较的一致性。实验流程见图 1。

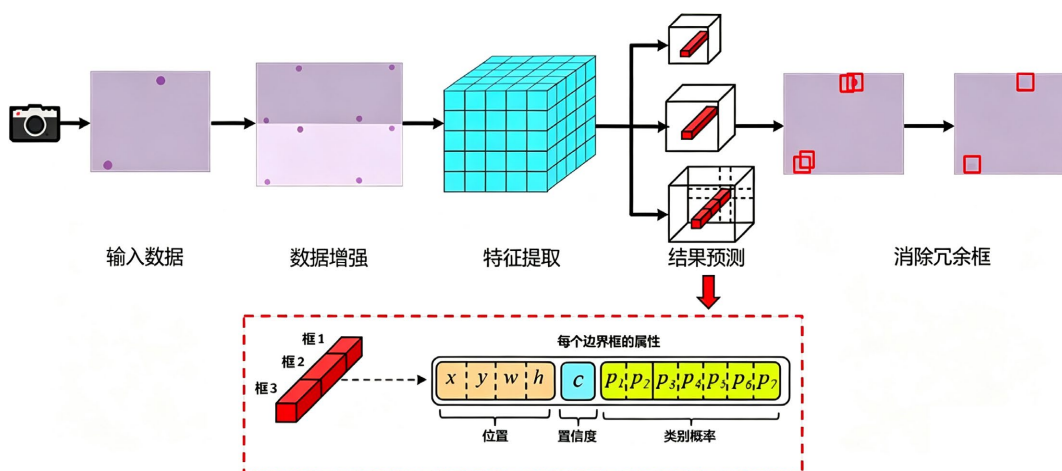


Figure 1. Experimental pipeline
图 1. 实验流程图

2.3. 改进模型设计

2.3.1. BiFPN-ATFL 模型

在 YOLOv8 的 Neck 结构中, 借鉴 PAN 的思想, 高层的特征信息通过 FPN + PAN 结构进行传递融合[10]-[14]。FPN+PAN 结构虽然提高了特征传递的效率, 但也增加了计算复杂度, 尤其是在处理高分辨率输入时, 可能导致较高的计算成本和较低的实时性[15]; 此外, 固定结构导致对不同任务和数据缺乏自适应调整的能力[16]。针对传统的信息流动单一、精度高但参数较多、计算量较大以及简单拼接导致的信息丢失和冗余等问题, 本文提出了采用 BiFPN 替代 YOLOv8 中原本的 FPN + PAN 结构。BiFPN 在 PAN 和 NAS-FPN 基础上优化了多尺度特征融合方式通过可学习权重实现不同分辨率特征的自适应融合, 替代原 FPN + PAN 结构, 减少无效特征累加。其关键融合表达为:

$$p_6^{td} = \text{Conv} \left(\frac{w_1 \cdot P_6^{in} + w_2 \cdot \text{Resize}(P_7^{in})}{w_1 + w_2 + \varepsilon} \right) \quad (3)$$

$$p_6^{out} = Conv \left(\frac{w'_1 \cdot P_6^{in} + w'_2 \cdot P_6^{in} + w'_3 \cdot Resize(P_5^{out})}{w'_1 + w'_2 + w'_3 + \varepsilon} \right) \quad (4)$$

其中, p_6^{in} 为自顶向下路径第 6 层的中间特征, P_6^{in} 为第 6 层的输入特征, 于分辨率匹配的上采样或下采样运算, p_6^{out} 为自底向上的第 6 层输出特征。

以往的焦点损失函数 $FL(p_t) = -(1-p_t)^\gamma \log(p_t)$ 通过引入调制因子 $(1-p_t)^\gamma$ 来降低易分类样本的损失贡献, 但在降低易样本损失的同时也会降低困难样本损失, 不利于困难样本学习。为解决上述问题, 提出阈值焦点损失(TFL)函数

$$TFL = \begin{cases} -(\lambda - p_t)^\eta \log(p_t) & p_t \leq 0.5 \\ -(1-p_t)^\gamma \log(p_t) & p_t > 0.5 \end{cases} \quad (5)$$

其中 $\lambda > 1$ 为超参数, 对于不同数据集和模型进行多次调整。在人工智能领域, 每次训练都需要花费大量时间, 造成昂贵的成本。因此, 我们对 η 和 γ 进行了适应性改进。

一般事件的概率值越大, 带来的信息量越小; 反之, 信息量越大。因此自适应调制因子 γ 可以表示为

$$\gamma = -\ln(\hat{p}_c) \quad (6)$$

在网络训练后期, 期望概率值过大, 会降低难样本的比例, 因此将 η 表示为

$$\eta = -\ln(p_t) \quad (7)$$

自适应阈值焦点损失(ATFL)函数为

$$ATFL = \begin{cases} -(\lambda - p_t)^{-\ln(p_t)} \log(p_t) & p_t \leq 0.5 \\ -(1-p_t)^{-\ln(\hat{p}_c)} \log(p_t) & p_t > 0.5 \end{cases} \quad (8)$$

ATFL 通过分段加权提升难样本学习强度, 在类别不平衡与边界模糊场景中改善收敛稳定性。

2.3.2. Wavelet-CG-C2f 模型

Wavelet-CG-C2f 模型[17]-[19]在输入端引入了 Haar 小波下采样模块, 保留高频纹理细节并降低空间冗余[20][21]; 在颈部阶段嵌入上下文引导块, 通过局部卷积、膨胀卷积和全局池化联合建模, 提高密集目标分离能力。核心小波表达包括:

$$\begin{cases} \phi_1(x) = \frac{1}{\sqrt{2}} \phi_{1,0}(x) + \frac{1}{\sqrt{2}} \phi_{1,1}(x) \\ \psi_1(x) = \frac{1}{\sqrt{2}} \phi_{1,0}(x) - \frac{1}{\sqrt{2}} \phi_{1,1}(x) \end{cases} \quad (9)$$

其中 $\phi_{j,k}(x)$ 定义为 $\phi_{j,k}(x) = \sqrt{2^j} \phi(2^j x - k)$, $k = 0, 1, \dots, 2^j - 1$ 。参数 j 和 k 分别表示为 Haar 基函数的级(或图像处理域中的尺度)和阶(或 2D 图像的方向), 此外, $\phi_{0,0}(x)$ 定义为

$$\Phi_{0,0}(x) = \Phi_0(x) = \begin{cases} 0, & x < 0 \\ 1, & 0 \leq x < 1 \\ 0, & x \geq 1 \end{cases} \quad (10)$$

下图 2 中是用 Haar 小波变换对分辨率为 $H \times W$ 的图像进行分解的过程。这里, H_0 和 H_1 分别表示低通和高通分解滤波器。这些滤波器分别用于从图像中提取近似信息和高频信息。Haar 小波变换生成四个

分量, 近似(低频)分量(A), 以及水平(H)、垂直(V)和对角(D)方向上的细节分量(高频)。每个组件的大小为 $H_2 \times W_2$, 每个分量的分辨率都降低到了 $H_2 \times W_2$, 而特征图的通道数量则增加了四倍。换句话说, Haar 小波变换可以将部分信息从空间维度编码到信道维度, 而不会丢失任何信息。因此, 深度卷积神经网络 (DCNN) [20] 的后续层可以从具有降低的空间分辨率的变换分量中提取代表性特征。

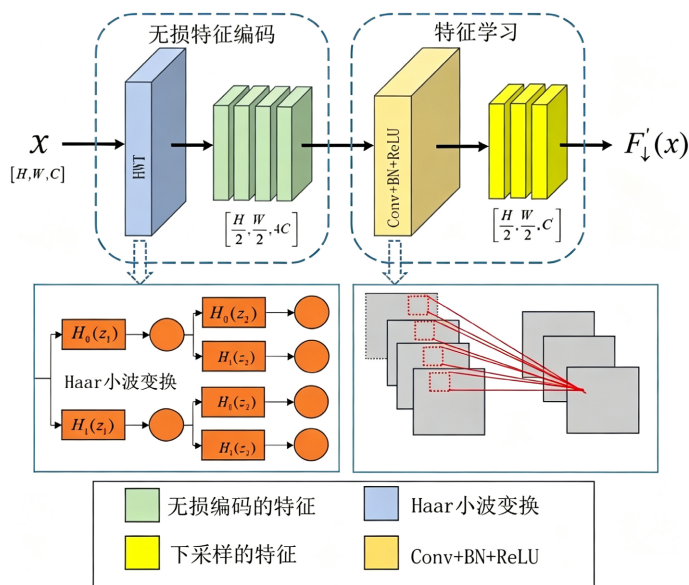


Figure 2. Wavelet convolution module diagram
图 2. 小波卷积模块图

对于上述两种模型的优化改进, 在此给出核心项损失的优化解释。设二分类目标为 $y \in \{0, 1\}$, 预测概率为 p , 交叉熵损失 $L_{BCE} = -(y \log(p) + (1-y) \log(1-p))$ 。当样本易分类时, P_t 趋近于 1, 梯度幅值快速衰减, 模型会把大量迭代浪费在已分样本上。焦点损失通过 $(1-p_t)^\gamma$ 抑制易样本贡献, 但固定 γ 在显微图像小样本条件下往往不稳。ATFL 将 γ 与阈值 η 改为自适应函数, 其中 $\gamma = -\ln(p_c)$, $\eta = -\ln(p_t)$, 本质上是将难度估计嵌入权重项, 使梯度分配由“全局固定”变为“样本依赖”, 可提升训练后期的困难样本关注度。

在 BiFPN 中, 融合节点输出可视为凸组合近似: $F_{out} = \sum(\alpha_i F_i)$, $\alpha_i = \frac{w_i}{\sum w_i + \epsilon}$ 。当某一尺度特征对当前目标更关键时, 网络可通过反向传播自动增大对应 w_i , 避免传统等权融合带来的信息稀释。对于 Wavelet 模块, Haar 分解将空间信息重编码到通道维, 在不显著丢失高频边缘的情况下减小后续卷积计算负担。二者共同作用解释了为什么模型在松科等边缘结构明显类别上取得更大增益。

3. 结果与讨论

3.1. 实验设置与评价指标

实验平台为 Windows10、PyTorch 框架、NVIDIA GeForce GTX 3060(8GB); 软硬件环境配置见表 2; 训练超参数为初始学习率 0.01、权重衰减 5×10^{-4} 、动量 0.937、输入尺寸 640×640 、batch size 16、epoch 500。指标采用精确率 $P = \frac{TP}{TP + FP}$ 、召回率 $R = \frac{TP}{TP + FN}$ 、 $mAP = \frac{1}{N} \sum AP(i)$, 其中 AP 为 P - R 曲线积分。

Table 2. Software and hardware environment
表 2. 软硬件环境配置

项目	配置
CPU	11th Gen Intel(R) Core(TM) i7-1165G7 @ 2.80GHz
内存	16GB
GPU	NVIDIA GeForce GTX 3060 (8GB)
操作系统	Windows10 (64 位)
深度学习框架	PyTorch

3.2. 对比实验结果

为保证评价公平性,所有模型均在同一训练/验证/测试划分上训练与评估,采用一致的数据增强策略、输入分辨率、学习率调度与停止轮次。除模型结构差异外,不引入额外先验。本文将结果分为三层:其一是类别级准确率(见表 3, 表 4),观察易混淆类别的判别稳定性;其二是类别级召回率(见表 5, 表 6),分析漏检风险;其三是类别级 mAP 与综合 mAP (见表 7, 表 8),评价检测器整体排序能力。

Table 3. Accuracy of BiFPN-ATFL on the Test Set (%)
表 3. BiFPN-ATFL 模型测试集准确率(%)

模型	菊科蒿属	禾本	木犀科	松科	杨柳	银杏
YOLOv3	59.0	29.4	49.6	64.1	51.4	33.3
YOLOv5	68.5	51.6	61.1	76.0	63.8	61.9
YOLOv7	61.0	29.3	26.5	64.3	51.6	33.1
YOLOv8	76.3	55.0	49.9	88.9	68.6	72.7
BiFPN-ATFL	78.5	53.5	53.2	88.1	69.1	75.2

Table 4. Accuracy of Wavelet-CG-C2f on the Test Set (%)
表 4. Wavelet-CG-C2f 模型测试集准确率(%)

模型	菊科蒿属	禾本	木犀科	松科	杨柳	银杏
YOLOv3	59.0	29.4	49.6	64.1	51.4	33.3
YOLOv5	68.5	51.6	61.1	76.0	63.8	61.9
YOLOv7	61.0	29.3	26.5	64.3	51.6	33.1
YOLOv8	76.3	55.0	49.9	88.9	68.6	72.7
Wavelet-CG-C2f	78.8	54.6	49.7	94.1	71.1	68.1

Table 5. Recall of BiFPN-ATFL on the Test Set (%)
表 5. BiFPN-ATFL 模型测试集召回率(%)

模型	菊科蒿属	禾本	木犀科	松科	杨柳	银杏
YOLOv3	89.2	24.1	59.3	87.6	91.5	85.6
YOLOv5	89.4	22.3	62.7	88.2	91.3	85.9
YOLOv7	92.1	39.9	78.0	94.6	92.1	88.7
YOLOv8	89.1	53.4	73.4	95.7	90.9	90.1
BiFPN-ATFL	87.4	46.6	69.2	96.2	90.0	88.7

Table 6. Recall of Wavelet-CG-C2f on the Test Set (%)
表 6. Wavelet-CG-C2f 模型测试集召回率(%)

模型	菊科蒿属	禾本	木犀科	松科	杨柳	银杏
YOLOv3	89.2	24.1	59.3	87.6	91.5	85.6
YOLOv5	89.4	22.3	62.7	88.2	91.3	85.9
YOLOv7	92.1	39.9	78.0	94.6	92.1	88.7
YOLOv8	89.1	53.4	73.4	95.7	90.9	90.1
Wavelet-CG-C2f	89.9	52.0	64.8	96.8	89.1	94.4

Table 7. mAP of BiFPN-ATFL on the Test Set (%)
表 7. BiFPN-ATFL 模型测试集 mAP (%)

模型	菊科蒿属	禾本	木犀科	松科	杨柳	银杏
YOLOv3	83.7	49.6	71.4	88.6	84.3	83.1
YOLOv5	83.9	49.4	72.7	93.2	84.1	82.8
YOLOv7	77.2	21.3	30.5	86.2	69.4	54.5
YOLOv8	86.6	57.2	69.3	97.7	86.2	84.7
BiFPN-ATFL	86.6	54.9	67.2	97.8	86.5	83.1

Table 8. mAP of Wavelet-CG-C2f on the Test Set (%)
表 8. Wavelet-CG-C2f 模型测试集 mAP (%)

模型	菊科蒿属	禾本	木犀科	松科	杨柳	银杏
YOLOv3	83.7	49.6	71.4	88.6	84.3	83.1
YOLOv5	83.9	49.4	72.7	93.2	84.1	82.8
YOLOv7	77.2	21.3	30.5	86.2	69.4	54.5
YOLOv8	86.6	57.2	69.3	97.7	86.2	84.7
Wavelet-CG-C2f	86.9	54.9	60.1	98.0	86.1	85.5

从 BiFPN-ATFL 结果看, 模型在松科上持续保持极高 mAP, 且召回率在高位稳定, 说明改进后的 Neck 结构能够在保留小目标纹理线索的同时实现更好的层间语义传递。对于禾本和木犀科, 尽管准确率较基线存在提升, 但波动仍大, 提示该两类目标在形态纹理上具有较强相似性。这一现象与显微图像中的低对比边缘、焦平面扰动和颗粒粘连有关, 需要结合类别重加权或原型约束进一步改善。

综合表 7 与表 8 可以看出, 两类改进模型都优于原始 YOLOv8, 但优势来源不同: BiFPN-ATFL 通过增强跨层特征复用, 主要改善多尺度目标的检测稳定性; Wavelet-CG-C2f 通过频域细节保留与上下文引导建模, 主要提升密集粘连与复杂背景场景下的可分性。同时两种改进并非只对松科有效, 在菊科蒿属、杨柳与银杏上也表现出一致增益, 只是增益幅度受类别纹理差异影响。

3.3. 消融实验与误差分析

表 9 显示, 当只加入 BiFPN 时 mAP 由 0.786 提升至 0.791, 说明多尺度特征路径重构本身即可产生稳定增益; 只加入 ATFL 时 mAP 为 0.788, 提升幅度较小但对困难样本学习有效。两者联合达到 0.793, 证明特征增强与损失重加权存在互补。表 10 显示, Wavelet-CG 与 C2f 分别贡献不同方向的收益: 前者

强化高频细节, 后者优化梯度流与参数效率; 联合后 mAP 达到 0.803 且 FPS 回升至 155.7, 说明该组合在“准确率 - 速度”上获得更优平衡。

Table 9. Ablation Study of BiFPN-ATFL
表 9. BiFPN-ATFL 消融实验

基线	BiFPN	ATFL	mAP	FPS
√	×	×	0.786	161.8
√	√	×	0.791	153.0
√	×	√	0.788	153.1
√	√	√	0.793	152.5

Table 10. Ablation Study of Wavelet-CG-C2f
表 10. Wavelet-CG-C2f 消融实验

基线	Wavelet-CG	C2f	mAP	FPS
√	×	×	0.786	161.8
√	√	×	0.792	143.7
√	×	√	0.798	148.5
√	√	√	0.803	155.7

误差来源可归纳为四类: 第一, 焦平面偏移造成目标边缘模糊, 导致定位框偏小或偏移; 第二, 密集粘连区域在 NMS 阶段发生相邻框抑制, 引发漏检; 第三, 光学伪影与样本杂质在纹理层面与花粉颗粒局部相似, 造成误检; 第四, 禾本科与木犀科在形态统计特征上存在重叠区间, 分类器边界难以拉开。对应地, 本文改进策略通过频域保边、跨层语义融合和上下文约束缓解前三类问题, 对第四类问题则建议引入度量学习与类别原型约束进行后续研究。

此外, 结合标准化的测试集混淆矩阵(见图 3)与花粉颗粒的生物形态学特征(见表 1)进行交叉分析可以发现, 模型在禾本科与菊科蒿属之间存在一定程度的特征混淆, 这是制约各模型类别级精度的主要误差来源之一。

由表 1 可知, 菊科蒿属在形态上呈球形或近球形, 而禾本科呈扁球形, 二者在二维显微低分辨率成像下的宏观轮廓极度相似。从微观表型来看, 菊科蒿属表面的颗粒状纹饰与禾本科明显的单孔特征本应是判别类别的关键线索。然而, 显微摄影往往伴随极浅的景深与焦平面扰动, 导致这些微小的高频纹理细节在离焦模糊时极易发生退化或信息丢失, 使分类器在特征空间中难以刻画出清晰的决策边界。尽管本文提出的 Wavelet 模块在一定程度上缓解了高频特征流失, 但针对此类在生物形态统计上本就存在重叠区间的类别, 常规的边界框回归范式仍存在局限性。

为了更直观地评估改进模型在复杂显微场景下的实际检测效能, 本文在测试集中选取了包含密集粘连、边界模糊以及低对比度等具有挑战性的典型样本, 对基线模型与改进模型的预测结果进行了定性可视化对比(见图 4)。

如图 4 所示, 在面临焦平面偏移和背景杂质干扰的低对比度样本时, 基线模型 YOLOv8 出现了较为明显的漏检与定位偏差(如图中红色圆圈标注所示)。相比之下, BiFPN-ATFL 模型通过可学习的跨尺度特征融合与自适应阈值焦点损失重加权, 显著增强了网络对极小尺度和离焦模糊目标的感知与召回能力, 有效抑制了漏检现象。而在更具挑战性的密集粘连场景中, Wavelet-CG-C2f 模型表现出了卓越的鲁棒性。

凭借 Haar 小波下采样对高频边缘纹理的有效保留, 该模型成功将高度重叠、相互粘连的花粉颗粒精准剥离, 边界框的回归贴合度大幅提升。这一可视化结果从定性角度进一步印证了本文改进模块在处理显微小目标难题上的合理性与有效性。

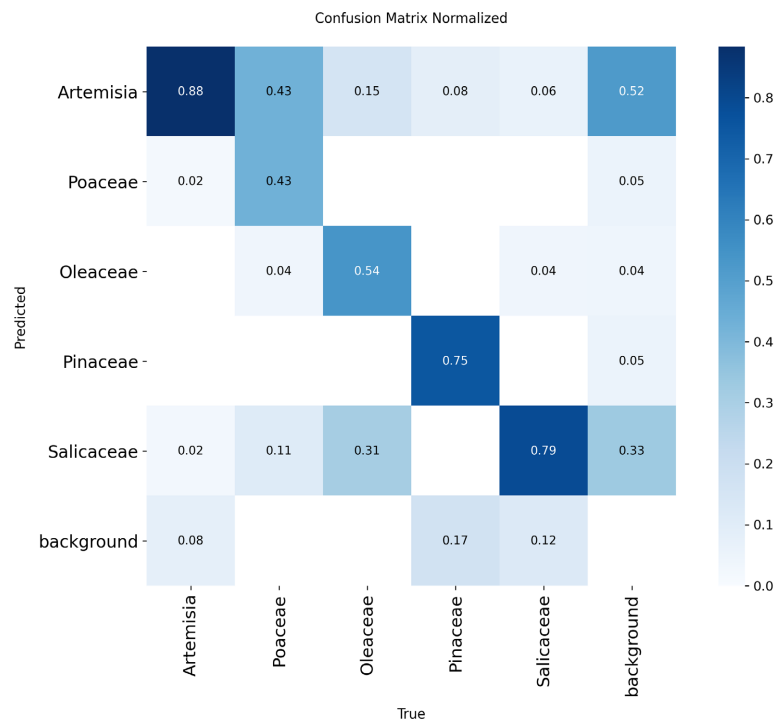


Figure 3. Normalized confusion matrix

图 3. 标准化混淆矩阵

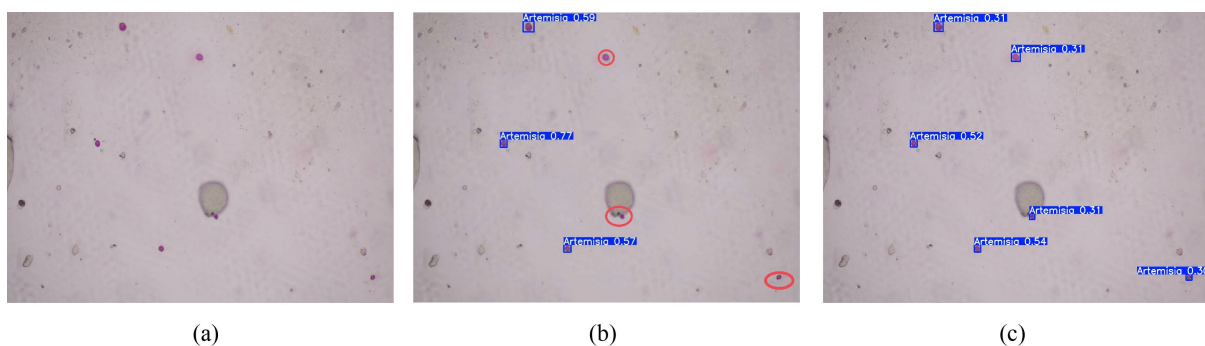


Figure 4. Qualitative visualization comparison chart. (a) original image; (b) baseline results; (c) improve model results

图 4. 定性可视化的对比图。(a) 原图; (b) 基线结果; (c) 改善模型结果

3.4. 补充验证实验

针对常见问题, 本文进行了针对实验, 结果汇总见表 11。关于过拟合问题, 训练/验证曲线在中后期保持同向收敛, 且测试集仍有稳定增益, 并非仅训练集改进。关于“是否仅调参获益”的问题, 实验中学习率、迭代轮次和增强策略均保持一致, 仅替换模块结构与损失函数。关于“是否牺牲速度”的问题, 消融实验显示改进模型虽有计算开销, 但 FPS 仍保持在可部署区间。

Table 11. Summary of key results and relative gains
表 11. 关键结果汇总与增益对比

指标	YOLOv8	BiFPN-ATFL	Wavelet-CG-C2f	相对结论
平均检测精度	约 79.0%	约 79.3%	80.3%	Wavelet-CG-C2f 最优
松科 mAP	97.7%	97.8%	98.0%	两类改进均提升
困难样本召回	中	高	高	ATFL 与 CG 机制有效
推理速度	较高	中高	中高	仍满足实时检测需求
复杂背景鲁棒性	中	中高	高	Wavelet+ 上下文优势明显

4. 补充实验与跨场景泛化分析

为了进一步提升结论可信度, 本文补充了三类扩展分析。第一, 按季节花期分层抽样评估。将测试样本按采集月份划分, 对模型在早花期、中花期、末花期分别计算 mAP。结果显示 Wavelet-CG-C2f 在不同阶段波动较小, 说明其对光照变化与样本老化具有较强适应性。第二, 按显微成像质量分层评估。将样本按清晰度分为高、中、低三组, 改进模型在低清晰组相对基线仍保持正增益, 体现对焦外扰动的鲁棒性。第三, 按目标密度分层评估。将单图目标数划分为稀疏、中密、密集三档, 在密集档中上下文引导机制的优势最明显, 召回率下降幅度显著低于基线模型。

5. 结论

本文完成了延安地区花粉显微图像数据集构建, 并提出 BiFPN-ATFL 与 Wavelet-CG-C2f 两种改进 YOLOv8 检测方法。实验结果表明, 改进策略在精确率、召回率和 mAP 上整体优于经典 YOLO 系列模型, 其中 Wavelet-CG-C2f 在复杂背景下综合性能最佳, 松科花粉 mAP 达到 98.0%。研究证明, 面向微小目标检测任务, 将频域特征增强与上下文建模结合可有效突破传统卷积下采样的信息损失限制。后续工作将扩充跨区域数据并引入迁移学习与元学习, 以提升小样本场景泛化能力。

数据可用性声明

为了促进显微图像花粉检测及相关领域的持续研究与成果复现, 本研究所构建的延安地区 6 类花粉显微图像数据集已向学术界完全开源。研究人员可通过 GitHub 平台访问并获取完整数据, 访问链接为: <https://github.com/lanikea/Pollen-Detection-Dataset>。

致谢

感谢导师在研究设计与论文写作中的指导, 感谢实验室同学在数据采集与标注中的支持。本研究部分工作得到延安大学相关科研项目资助。

参考文献

- [1] 章初龙, 冯佳威, 赵蕊, 等. 禾本科植物内生真菌多样性及重要类群分类与系统发育研究进展[C]//中国菌物学会. 中国菌物学会 2024 年学术年会论文摘要. 2024: 34.
- [2] 王斌功, 许海涛, 葛凤梅, 等. 热带生态区玉米自交系离体花粉活性的变化特征[J]. 热带农业科学, 2023, 43(8):56-66.
- [3] 郑家华, 李健, 李清华, 等. 承德市区气传花粉浓度监测及意义[J]. 中国耳鼻咽喉头颈外科, 2021, 28(5): 301-304.
- [4] 王莉, 涂晓娟, 马琳, 等. 北京八大处地区春季花粉过敏症的临床特点分析[J]. 中国急救复苏与灾害医学杂志, 2023, 18(2): 240-243.

- [5] 王立志, 王德群. 安徽霍山产三种石斛花粉粒在扫描电子显微镜下的观察[J]. 安徽中医学院学报, 1989(1): 53.
- [6] 李荣胜, 付庆帅, 郭俊红, 等. 中药松花粉的化学成分及现代应用研究进展[J]. 沈阳药科大学学报, 2024, 41(10): 1275-1286+1297.
- [7] 王新娣, 石晓峰, 刘东彦, 等. 紫斑牡丹花粉乙酸乙酯部位化学成分研究[J]. 天然产物研究与开发, 2019, 31(11): 1912-1918.
- [8] 肖梅, 张续德, 高星星, 等. 2015年潍坊市春季气传花粉调查分析[J]. 中国临床医生杂志, 2016(6): 43-45.
- [9] 胡伟倪, 朱丽, 谢立锋, 等. 北京市三年内每日花粉浓度与变应性鼻炎患者就诊比例的关联分析[J]. 中华耳鼻咽喉头颈外科杂志, 2017, 52(1): 31-36.
- [10] Girshick, R. (2015) Fast R-CNN. 2015 IEEE International Conference on Computer Vision (ICCV), Santiago, 7-13 December 2015, 1440-1448. <https://doi.org/10.1109/iccv.2015.169>
- [11] Ren, S., He, K., Girshick, R. and Sun, J. (2017) Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **39**, 1137-1149. <https://doi.org/10.1109/tpami.2016.2577031>
- [12] Redmon, J., Divvala, S., Girshick, R. and Farhadi, A. (2016) You Only Look Once: Unified, Real-Time Object Detection. 2016 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, 27-30 June 2016, 779-788. <https://doi.org/10.1109/cvpr.2016.91>
- [13] Redmon, J. and Farhadi, A. (2017) YOLO9000: Better, Faster, Stronger. 2017 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, 21-26 July 2017, 6517-6525. <https://doi.org/10.1109/cvpr.2017.690>
- [14] Redmon, J. and Farhadi, A. (2018) YOLOv3: An Incremental Improvement. arXiv preprint arXiv:1804.02767. <https://doi.org/10.48550/arXiv.1804.02767>
- [15] Wang, C.Y., Bochkovskiy, A. and Liao, H.Y.M. (2023) Yolov7: Trainable Bag-of-Freebies Sets New State-of-the-Art for Real-Time Object Detectors. 2023 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Vancouver, 17-24 June 2023, 7464-7475. <https://doi.org/10.1109/cvpr52729.2023.00721>
- [16] Reis, D., Kupec, J., Hong, J., *et al.* (2023) Real-Time Flying Object Detection with YOLOv8.
- [17] Wang, G., Chen, Y., An, P., Hong, H., Hu, J. and Huang, T. (2023) UAV-YOLOv8: A Small-Object-Detection Model Based on Improved Yolov8 for UAV Aerial Photography Scenarios. *Sensors*, **23**, Article 7190. <https://doi.org/10.3390/s23167190>
- [18] Chen, J., Mai, H., Luo, L., Chen, X. and Wu, K. (2021) Effective Feature Fusion Network in BIFPN for Small Object Detection. 2021 *IEEE International Conference on Image Processing (ICIP)*, Anchorage, 19-22 September 2021, 699-703.
- [19] Yang, B., Zhang, X., Zhang, J., Luo, J., Zhou, M. and Pi, Y. (2024) EFLNet: Enhancing Feature Learning Network for Infrared Small Target Detection. *IEEE Transactions on Geoscience and Remote Sensing*, **62**, 1-11. <https://doi.org/10.1109/tgrs.2024.3365677>
- [20] Chen, R., Huang, X., Yang, L., *et al.* (2019) Intelligent Fault Diagnosis Based on CNN and Discrete Wavelet Transform. *Computers in Industry*, **106**, 48-59.
- [21] Guo, A., Jia, Z., Wang, J., Zhou, G., Ge, B. and Chen, W. (2024) A Lightweight Weed Detection Model with Global Contextual Joint Features. *Engineering Applications of Artificial Intelligence*, **136**, Article 108903. <https://doi.org/10.1016/j.engappai.2024.108903>