doi:10.6041/j.issn.1000-1298.2026.12.027

# 动态尺度特征挖掘与双向注意力交互的黄瓜病害检测

王心霖　邹治涵　金海波
(辽宁工程技术大学软件学院, 葫芦岛 125105)

**摘要**: 针对黄瓜病害检测中各类病斑难以区分和识别效率低等问题,提出动态尺度特征挖掘与双向注意力交互的黄瓜病害检测算法。在特征提取阶段引入感受野与通道注意力卷积 RFCAConv,强化病斑区域的特征表征,提升模型对形态变异较大病害目标的识别能力;采用自适应特征增强模块 AFEM,有效抑制背景噪声,增强目标与背景的区分度;设计 C3k2 - DSC 模块,有效减少因目标形状变化导致的特征信息丢失。在构建的黄瓜病害数据集上进行消融试验,结果表明平均精度均值达到 85.9%,相较于基线模型提升 3.8 个百分点,召回率提升 5.1 个百分点,各项改进均对模型识别性能有提升效果,其中,AFEM 模块对模型提升效果最为明显;模型在包含 3 070 幅图像的公开数据集上试验结果表明,精确率和召回率分别提升 2.3 个百分点和 4.4 个百分点。将模型部署在 Android 手机端,自然环境下测试结果表明,平均识别准确率为 84.8%,平均检测耗时为 0.352 s。所建立的黄瓜病害检测系统能实现对黄瓜病害的精准快速识别,为黄瓜病害的及时防治提供有效方法。

**关键词**: 黄瓜; 病害检测; 动态尺度特征挖掘; 双向注意力交互; 黄瓜病害检测系统
**中图分类号**: TP391.41; S432.1　**文献标识码**: A　**文章编号**: 1000-1298(2026)12-0285-11　**OSID**:

# Cucumber Disease Detection with Dynamic Scale Feature Mining and Bidirectional Attention Interaction

WANG Xinlin　ZOU Zhihan　JIN Haibo
(*College of Software*, *Liaoning University of Engineering Technology*, *Huludao* 125105, *China*)

**Abstract**: Aiming to address the challenges of difficult-to-distinguish and inefficiently recognized disease spots in cucumber disease detection, a cucumber disease detection algorithm combining dynamic scale feature mining and bidirectional attention interaction was proposed. In the feature extraction stage, receptive field and channel attention convolution (RFCAConv) was introduced to enhance the feature representation of disease spot areas and improve the model's ability to recognize disease targets with significant morphological variations. An adaptive feature enhancement module (AFEM) was employed to effectively suppress background noise and enhance the distinction between targets and backgrounds. The C3k2 - DSC module was designed to effectively reduce the loss of feature information caused by target shape variations. Ablation experiments conducted on a constructed cucumber disease dataset showed that the mean average precision reached 85.9%, an improvement of 3.8 percentage points compared with that of the baseline model, and the recall rate was increased by 5.1 percentage points. All improvements enhanced the model's recognition performance, with the AFEM module showing the most significant improvement. The model was tested on a public dataset containing 3 070 images, and the experimental results showed that the precision and recall rates were increased by 2.3 percentage points and 4.4 percentage points, respectively. When deployed on an Android mobile phone, the test results showed an average recognition accuracy of 84.8% and an average detection time of 0.352 s. The established cucumber disease detection system achieved accurate and rapid identification of cucumber diseases, providing an effective method for timely prevention and control of cucumber diseases.

**Key words**: cucumber; disease detection; dynamic scale feature mining; bidirectional attention interaction; cucumber disease detection system

收稿日期: 2025-03-28　修回日期: 2025-04-20
**基金项目**: 国家自然科学基金面上项目(62173171)
**作者简介**: 王心霖(1979—),男,讲师,博士,主要从事人工智能和计算机视觉等研究,E-mail: wangxinlin1107@foxmail.com

## 0 引言

黄瓜是我国重要的经济型农作物之一,种植区域遍布全国,是我国保护地种植面积最大的蔬菜作物之一,在中国拥有超过 2000 年栽培历史[1]。作为全球最大的黄瓜生产国,我国自上世纪 70 年代起持续扩大种植面积,其产量贡献率始终保持世界第一[2]。截止 2021 年底,我国黄瓜种植面积达 $1.29\times10^6$ hm$^2$,年产量突破 $7.6\times10^7$ t,占据全球黄瓜生产总量的 4/5;单位面积产量达 $5.85\times10^4$ kg/hm$^2$,较同期全球平均单产($4.31\times10^4$ kg/hm$^2$)提升 35.8%[3]。借助温室大棚技术,黄瓜能够实现四季栽培,具有成本较低、产量较高的显著优势。黄瓜的广泛种植显著改善菜农收益结构,助力农村脱贫工作,但同时对病害防治也提出了更高要求[4]。在中国黄瓜炭疽病[5]、霜霉病[6]、蔓枯病[7]等病害在各地普遍发生。若在发病初期未能及时采取防治措施,病害可在短时间内迅速传播,进而导致黄瓜大面积减产[8]。

随着深度学习技术的迅猛发展,在过去 10 年间,神经网络技术在农作物病害识别领域得到了广泛应用。目前,神经网络技术在农作物病害识别领域的应用分为图像分类[9]和目标检测[10]。在图像分类算法中,MOHANTY 等[11]利用谷歌神经网络 GoogLeNet 模型对标注图像进行训练,通过 1×1 卷积核降低通道维度,减少了模型参数量和计算成本。RANGARAJAN 等[12]使用改进的亚历克斯神经网络 AlexNet 和视觉几何组网络 VGG16 模型对番茄病害进行训练,结合 AlexNet 快速低级特征提取与 VGG16 深层细节捕捉。ZHANG 等[13]提出了一种多通道自动定向递归注意力网络,以丰富番茄叶片病害的特征,递归注意力网络能动态聚焦于病害区域,抑制无关背景噪声,提高了神经网络表征能力。图像分类算法由于在检测图像时无法确定目标大小与位置和多个目标,具体应用不如目标检测算法。目标检测算法能够自动从图像中提取有效特征,有效地解决单幅图像中多目标分类与定位问题[14]。樊湘鹏等[15]提出了一种改进的卷积神经网络模型,用于复杂背景的玉米病害识别,通过优化网络结构和数据预处理策略,显著提升了田间环境中病害识别准确率和泛化能力,设计了包含 5 层卷积、4 层池化和 2 个全连接层的深度网络结构,并结合 L2 正则化与 Dropout 技术抑制过拟合,使模型在 9 种玉米病害识别任务中大幅提高了平均精度均值;樊湘鹏等[16]提出了一种基于迁移学习和改进 VGG16 的葡萄叶部病害检测模型,通过优化网络结构与训练策略,实现了复杂田间环境下高效病害识别,通过批归一化(BN)加速训练收敛,利用 GAP 替代全连接层显著降低计算复杂度,并通过中心损失函数增强类间特征分离能力,开发的 Android 移动端检测系统,在自然环境中平均准确率为 95.67%,实现了田间病害的快速诊断,为葡萄病害精准防控提供了智能化解决方案。张惠莉等[17]提出了一种基于 YOLO v8 - GSGF 的葡萄病害识别模型,通过多模块改进提升识别精度与速度,引入 GhostNetV2 主干网络,采用分解全连接层和 DFC 注意力机制增强特征提取能力,同时嵌入 SPPFCSPC 金字塔池化保持感受野并加速推理,通过 GAM 注意力机制聚焦病斑区域减少背景干扰,结合 Focal EIoU 损失函数优化边界框回归性能。黄丽明等[18]提出了一种改进的 YOLO v4 模型,用于无人机遥感影像中松材线虫病变色木检测模型,使用倒残差结构以增强浅层与深层特征的交互,为了降低模型参数量使用了深度可分离卷积,加强了对松材线虫病变色木的细节捕捉能力。刘诗怡等[19]使用改进的 YOLO v7 模型对黄瓜叶片病虫害进行检测与识别,可同时识别黄瓜叶片的多种病害,支持复杂场景下综合诊断。尽管上述方法在农作物病害检测领域取得了一定进展,但在实际应用中仍面临很多挑战。比如,病害早期检测精度较低,对复杂背景的病害识别能力不足,以及对不规则目标病害特征的提取不够精确,这些问题限制了智能农业技术的推广应用。

综上所述,通过深度学习目标检测技术检测植物病害具有可行性,但仍存在一些问题亟需解决。当叶片面积和病斑大小存在较大差异或遇到细长或弯曲的病斑时,难以精准捕捉复杂的病斑特征,造成漏检情况。复杂背景也常干扰模型,导致模型识别精度大幅降低。因此,为了优化模型结构,提高病害检测精确率和对复杂背景的病害识别能力,本研究采用 YOLO 11 模型作为基线模型,提出动态尺度特征挖掘与双向注意力交互(Dynamic scale feature mining and bidirectional attention interaction, DSFM - BAI)目标检测算法。

## 1 材料与数据集

### 1.1 黄瓜病害图像采集

选用黄瓜常见病害(炭疽病、细菌性枯萎病、腹腐病、霜霉病、蔓枯病、腐霉果腐病)作为研究对象,并加入新鲜无病害黄瓜和黄瓜叶片作为对照。在辽宁省绥中县黄瓜种植基地采集黄瓜和黄瓜叶片的高质量图像,使用佳能 90D 型相机于 2024 年 10 月 30 日—11 月 6 日 08:00—15:00 期间采集,共获得图

像 960 幅，所有黄瓜病害图像统一裁剪分辨率为 640 像素×640 像素。并借助网络爬虫技术，从网络公开资源中获取 509 幅患病黄瓜和黄瓜叶片的高质量图像，共得到 1 469 幅黄瓜和黄瓜叶片的高质量图像。数据集中图像背景和光强条件差异显著，包括 6 类患有病害及 2 类无病害的图像。筛选删除不适用图像以确保数据集独立分布，最后使用 1 280 幅图像制作数据集，包含 160 幅炭疽病、160 幅细菌性枯萎病、160 幅腹腐病、160 幅霜霉病、160 幅蔓枯病、160 幅腐霉果腐病以及 320 幅新鲜无病害黄瓜叶片和黄瓜图像，黄瓜病害图像如图 1 所示。




图 1　黄瓜病害

Fig. 1　Diseases of cucumber

### 1.2　数据集构建

对所有图像进行旋转、亮度调节和加噪等方式将数据集扩充至 2 560 幅。原始图像没有标注，所以使用目标检测标注工具 LabelImg 进行人工标注，标签分别为 An（炭疽病）、BW（细菌性枯萎病）、BR（腹腐病）、DM（霜霉病）、FC（新鲜黄瓜）、FL（新鲜叶片）、GS（蔓枯病）和 PF（腐霉果腐病）。

为了再次扩充黄瓜病害图像，采用 MixUp[20] 方法通过代码再次进行数据增强。该方法核心思想是通过线性融合 2 个不同样本生成新的样本，如图 2 所示。MixUp 方法先对 2 幅图像分别进行几何变换（随机翻转、尺度缩放）与像素操作（高斯模糊处理），随后将变换后的图像进行线性融合，针对融合过程产生的边缘空白区域采用灰色像素填充。数据集图像包含原始图像和 2 次数据增强图像，从 1 280 幅增扩至 3 840 幅，为保证模型训练的科学性，将数据集按比例 8∶1∶1 随机划分训练集、验证集和测试集。

## 2　检测模型构建

### 2.1　YOLO 11 网络模型

YOLO 是一种主流的目标检测算法，主要通过




图 2　MixUp 数据增强结果

Fig. 2　MixUp data enhancement results

卷积神经网络完成目标检测任务[21]。YOLO 11 是 Ultralytics 团队于 2024 年开发的新一代目标检测模型，支持目标检测、图像分类等任务，由特征提取主干、特征融合颈部和多任务预测头部构成。YOLO 11 的主干网络部分采用 C3k2 模块替换 YOLO v8 中 C2f 模块，C3k2 模块优化了特征提取能力，同时减少了计算量。对于颈部网络，在 SPPF 模块后新增 C2PSA 注意力机制，由 2 个卷积层和 1 个多头自注意力模块构成，增强了特征提取能力。YOLO 11 头部网络参考 YOLO v10 设计，采用深度可分离方法减少了计算冗余并提高推理速度。YOLO 11 在实时检测的同时保持了高精度检测，在多种复杂场景中依旧表现出色。

YOLO 11 模型在黄瓜病害识别任务中仍然有一些不足之处：①采用普通卷积进行特征提取，导致网络对不同位置的差分信息敏感度低，其固定的感受野难以有效捕捉细小病斑特征。②无法有效表征跨空间的像素关系，难以有效抑制无用背景信息，进而严重降低病害检测效率。③传统卷积操作无法自适应调整卷积核的形状与尺寸，难以精准提取细长或弯曲的病斑特征，易造成漏检情况。

### 2.2　DSFM－BAI 算法

针对基线模型 YOLO 11 对黄瓜病害检测各类病斑难以区分、识别效率低等问题，本文提出 DSFM－BAI 黄瓜病害识别模型，网络结构如图 3 所示，基于基线模型主要改进 4 个部分：①在主干网络中引入感受野与通道注意力卷积[22]（Receptive field and content-aware convolution，RFCAConv）代替主干网络中 Conv 模块，使模型能更有效地处理从小面积初期病斑到大面积严重感染病斑等不同尺度的病害特征。②在 Neck 中，提出并引入自适应特征增强模块（Adaptive feature enhancement module，AFEM），有效抑制无用背景信息。③设计并引入 C3k2－DSC 模块，使用动态蛇形卷积 DySnakeConv[23] 替换 Bottleneck 结构中普通卷积 Conv，增加模型对细长或弯曲病害特征的有效识别，减少因目标形状变化导致的特征信息丢失。④使用损失函数 WIoU[24] 替换 YOLO 11 原本的损失函数 CIoU，使用动态非单调

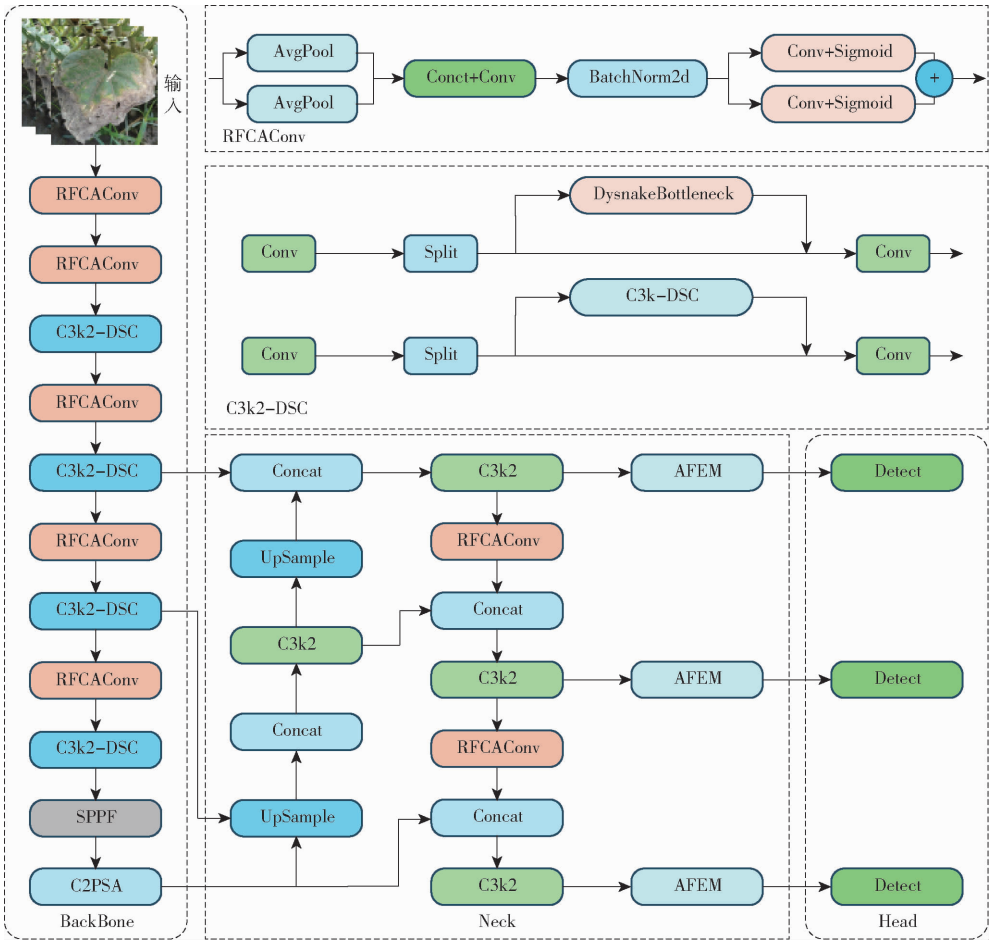


图 3 DSFM－BAI 模型结构

Fig. 3 DSFM－BAI model structure

聚焦机制进行损失计算，通过预测框与真实框的 IoU 统计特性动态调整梯度增益。

### 2.2.1 RFCAConv 模块

由于黄瓜病害形态多样性，叶片面积和病斑大小存在较大差异，基线模型特征提取时采用普通卷积，普通卷积核进行特征提取时采用共享参数，导致网络对不同位置的差分信息敏感度低，在处理形状、尺寸各异的病斑时，其固定的感受野难以有效捕捉细小病斑特征，进而影响模型对变形叶片病斑的检测精度。普通卷积在特征提取时缺乏对不同区域信息的重要性评估，使得网络难以充分关注病斑区域的关键特征。为缓解这一问题，在主干网络将普通卷积 Conv 替换为 RFCAConv，网络结构如图 4 所示。

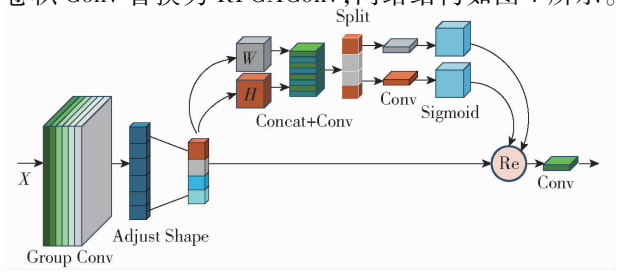


图 4 RFCAConv 网络结构

Fig. 4 RFCAConv network structure

图 4 中，$X$ 表示上一层网络输出，RFCAConv 通过自适应感受野调整和图像感知机制，通过优先考虑感受野区域内空间特征分布以突出不同特征的重要性，缓解卷积层中参数共享带来的局限性，使卷积核能够动态调整感受野范围，同时关注不同区域的重要性，使采样点更贴近叶片病斑的形状和尺寸，使模型能更有效地处理从小面积初期病斑到大面积严重感染病斑等不同尺寸的病害特征，从而增强模型对不同尺寸病斑的提取能力。

RFCAConv 在特征提取时不仅动态权衡感受野内多尺度特征的语义重要性，更通过空间结构建模捕捉感受野内局部几何分布特性，获取更丰富的特征信息；RFCAConv 在感受野调节的基础上引入特征权重自适应机制，使模型在检测叶片病害时能更有效地提取关键病斑信息，提高检测精度。

为了更精准地提取与病害检测相关的关键特征，RFCAConv 充分考虑每个感受野内特征的重要性差异，同时忽略无关信息，如叶脉或光照干扰等。RFCAConv 通过空间通道协同注意力机制解决了传统卷积因参数共享导致的特征表征能力受限问题。在主干网络中替换标准卷积后，增强了模型表征能

力和感受野有效捕捉病斑特征的能力，带来的计算增量也比较少。

### 2.2.2　自适应特征增强模块 AFEM

在黄瓜病害检测任务中，复杂的背景常干扰模型，难以有效抑制无用背景信息，进而严重降低病害检测效率。特征增强模块作为提升检测精度的关键组件，能有效表征跨空间的像素关系，抑制背景噪声，增强目标与背景的区分度。

受 SCP[25] 启发，提出自适应特征增强模块 AFEM，AFEM 由 3 个紧密协作的分支构成，显著优化了特征提取流程，降低了复杂背景对模型的干扰。SCP 结构如图 5（H 表示广播哈达玛积（Broadcast Hadamard product））所示。

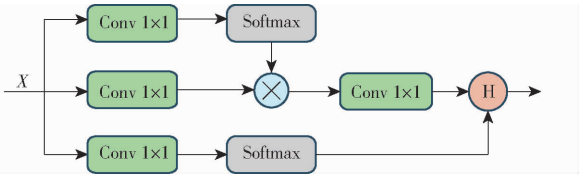


图 5　SCP 网络结构

Fig. 5　SCP network structure

AFEM 结构如图 6（M 表示 GMP 全局最大池化，A 表示 GAP 全局平均池化，LW 表示可学习权重模块）所示，第 1 个分支经过并行的全局平均池化 GAP 与全局最大池化 GMP 操作，整合输入特征图的全局信息。GAP 能够捕获特征图的整体平均特征，反映全局特征分布；GMP 则突出显著特征以强调关键信息。区别于 SCP 固定池化，紧接着经过学习权重模块（Learnable weights，LW），以提升模型对不同病害特征分布的自适应性。LW 通过动态调整 GAP 与 GMP 输出的权重，以动态平衡全局与局部特征，使模型能够精准聚焦病害区域，大幅提高对病害特征的关注度。第 2 个分支主要负责增强模型对大范围扩散病害的感知能力，如霜霉病、炭疽病。首先采用 1×1 卷积对输入的特征图进行初步处理，在不改变空间维度的前提下，通过调整通道维度以整合特征信息。随后，通过额外的卷积层对特征图进行深度线性变换，生成具有丰富语义信息的 Value 特征图，为特征融合提供关键信息；第 3 个分支采用 1×1 卷积对特征图进行处理，生成用于注意力计算的 Query 和 Key。这种简化运算不仅减少了计算量，还保留了特征中关键的空间与语义信息，确保在注意力机制中能够有效捕捉特征间的长距离依赖关系。

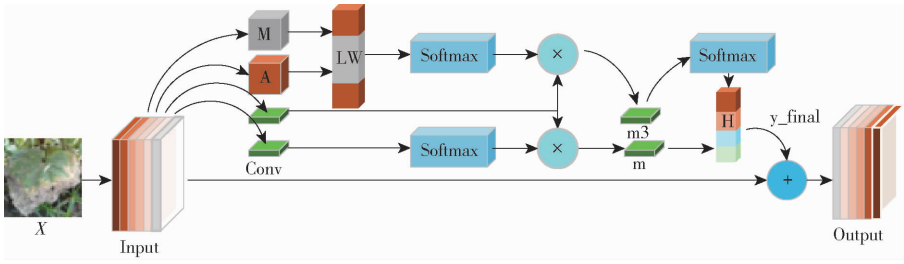


图 6　AFEM 网络结构

Fig. 6　AFEM network structure

3 个分支分别处理后，第 2 个分支生成的 Value 特征图与第 3 个分支生成的特征图，通过广播哈达玛积操作进行融合。哈达玛积进行逐元素相乘，使不同分支提取的特征信息在对应位置上相互增强，得到融合后的特征图 y_final。最后，将输入特征图 $X$ 与融合后的特征图 y_final 进行相加操作，把原始输入特征与经过多分支处理、融合增强后的特征相结合，充分保留原始信息的同时强化了关键特征。

在黄瓜病害检测任务中，将 AFEM 引入到模型中，优化了模型特征提取过程，可大幅提升模型的性能。AFEM 能够有效抑制背景噪声，突出病害特征，有助于提升目标检测算法在农作物病害检测的精确率。在 AFEM 的层级结构中，像素空间上下文可以表示为

$$Q_i^j = P_i^j + a_i^j \sum_{j=1}^{N_i} \frac{\exp(\boldsymbol{\omega}_{qk} P_i^j)}{\sum_{n=1}^{N_i} \exp(\boldsymbol{\omega}_{qk} P_i^n)} \boldsymbol{\omega}_v P_i^j \tag{1}$$

其中

$$a_i^j = \frac{\exp([\mathrm{avg}(P_i);\max(P_i)]P_i^j)}{\sum_{n=1}^{N_i} \exp([\mathrm{avg}(P_i);\max(P_i)]P_i^n)} \boldsymbol{\omega}_v \tag{2}$$

式中　$P_i^j$——第 $i$ 层特征图中第 $j$ 个像素的输入特征

$P_i$——第 $i$ 层特征图输入特征集合

$Q_i^j$——增强后像素特征

$a_i^j$——第 $i$ 层特征图中第 $j$ 个像素对应的注意力权重

$\boldsymbol{\omega}_{qk}$——用于计算 Query－Key 相似度的可学习权重矩阵

$N_i$——特征图总像素数

$\boldsymbol{\omega}_v$——Value 特征映射的可学习权重矩阵

通过 avg(·)和 max(·)分别执行全局平均池化 GAP 和全局最大池化 GMP。2 种池化操作协同引导特征图进行通道维度的重要信息筛选，使 AFEM

能有效了解通道上下文信息。

### 2.2.3 C3k2 - DSC 模块

在黄瓜病害检测过程,病斑形态多样性是影响准确率的主要因素。在黄瓜各类病害中,病斑大小和形状千差万别,特别是细长或弯曲病斑特征,由于传统卷积操作无法自适应调整卷积核的形状与尺寸,因此难以精准提取,极易造成漏检,进而很大程度上影响检测精度。为有效解决这一问题,本研究对基线模型中 Bottleneck 模块、C3k 模块和 C3k2 模块进行了重新设计[26],如图 7 所示。

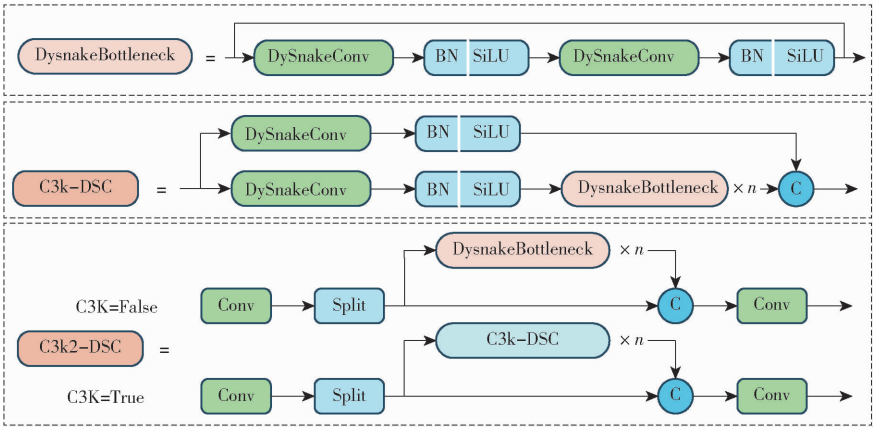


图 7 C3k2 - DSC 相关模块结构

Fig. 7 Overall frame of C3k2 - DSC related modules

基于动态蛇形卷积 DySnakeConv[23] 重新设计 DysnakeBottleneck 模块,由 2 个 DySnakeConv 层构成,并且采用跳跃连接方式进行数据传递。跳跃连接的引入能加强对特征详细学习,使模型在训练过程更容易学习到不同层次的特征,跳跃连接的引入还避免了梯度消失等问题。基于 DysnakeBottleneck 和 DySnakeConv,重构 C3k 模块,得到 C3k - DSC 模块;在 C3k2 基础上,使用 DysnakeBottleneck 模块和 C3k - DSC 模块替换 Bottleneck 模块和 C3k 模块,从而通过重新设计得到 C3k2 - DSC 模块。这种替换增强了对复杂形状病斑的特征提取能力,使模型能更好地识别黄瓜病害中不规则形态病斑的特征。并使用 C3k2 - DSC 模块替代了骨干网络中 C3k2 模块。

传统卷积操作在处理规则和网格状数据时具有良好表现,但当遇到目标形状变化较大情况时,其局限性便表现出来。动态蛇形卷积核心优势在于其动态可变形的卷积核设计,通过自适应调整卷积核形状与尺寸,进而对目标特定形态进行特征捕获,在保持计算效率的同时实现特征表达与目标形状的精准适配。动态蛇形卷积通过分析目标形态特征自适应调节卷积核参数,有效提升对不规则目标的识别能力。与固定尺寸及方向的传统卷积核相比,DySnakeConv 通过动态调节参数增强了对不规则目标的适应性,显著优化了困难检测样本的特征表征效果。在 DysnakeBottleneck 模块中加入该卷积后,系统能自主适应病灶区域的线状特征与干扰背景,尤其在线性不规则病灶检测中展现出优越性能。该方法核心机制聚焦于管状组织的局部特征提取,对于标准二维卷积坐标集合 $P$ 的中心点 $P_i=(x_i,y_i)$,以 1 个 3×3 卷积核为例,可用 $P$ 表示为

$$P=\{(x-1,y-1),(x-1,y),\cdots,(x+1,y+1)\} \tag{3}$$

为使卷积核能更灵活地聚焦于目标的复杂几何特征,DySnakeConv 引入变形偏移概念,变形偏移是指在卷积核原有位置基础上,通过学习额外的偏移量,使卷积核能根据目标的形状和几何特性调整自身位置和方向。然而,如果任由模型自由学习变形偏移,在处理细长管状结构时,卷积核的感知场很容易偏离目标区域。为克服这一问题,DySnakeConv 采用如图 8 所示的迭代策略。

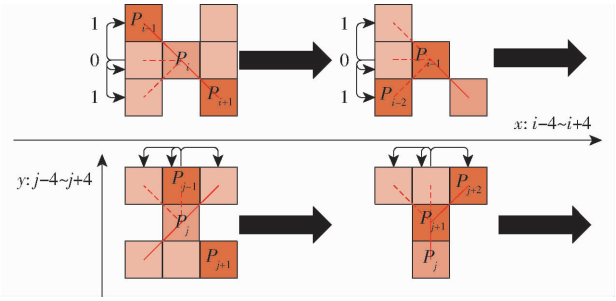


图 8 DySnakeConv 的坐标计算图解

Fig. 8 Illustration of coordinate calculation for DySnakeConv

该策略按照一定顺序依次选择每个待处理目标的下一个位置进行观察,这样可以保证卷积核关注的连续性,避免因变形偏移过大使感知范围过度扩散,计算式为

$$P_{i+c} = (x_{i+c}, y_{i+c}) = \left(x_{i+c}, y_i + \sum_{k=1}^{c} \Delta y_{i+k}\right) \quad (4)$$

$$P_{j+c} = (x_{j+c}, y_{j+c}) = \left(x_j + \sum_{k=1}^{c} \Delta x_{j+k}, y_{j+c}\right) \quad (5)$$

式中　$c$——中心位置采样步长

$\Delta x_{j+c}$——第 $j+c$ 个采样位置在 $x$ 方向学习得到的偏移增量

$\Delta y_{i+c}$——第 $i+c$ 个采样位置在 $y$ 方向学习得到的偏移增量

动态蛇形卷积通过将标准卷积核沿坐标轴方向线性重构实现形态适应。以 9×9 卷积核为例，在 $x$ 轴方向上各网格位置可定义为 $P_{i\pm c} = (x_{i\pm c}, y_{i\pm c})$，其中参数 $c \in \{0,1,2,3,4\}$ 表示相对于中心网格的水平偏移量。该结构的构建采用多阶段动态生成机制：首先通过 3×3 基础卷积生成初始偏移，再经 tanh 激活函数将偏移量约束至 $[-1,1]$ 区间，形成归一化偏移量 $\delta$。值得注意的是，卷积核的线性形态通过偏移量的逐级累积实现，从中心位置 $P_i$ 出发，后续网格位置 $P_{i+1}$ 由前序位置 $P_i$ 叠加增量偏移量 $\Delta(\Delta \in [-1,1])$ 计算得出，最终通过累积运算 $\Sigma$ 确保卷积核沿特定方向形成连续线性结构。这种递进式偏移调节机制，使卷积核能动态贴合目标的几何特征。

#### 2.2.4　损失函数

YOLO 11 模型使用 CIoU Loss 作为损失函数，在考量边界框坐标与类别误差方面具有良好表现，能综合考虑重叠面积、中心点距离和宽高比等因素，精准判断预测框与真实框的相似程度。然而，CIoU Loss 在进行损失计算时未考虑不同目标的检测难度与重要性差异。病害检测中，目标因大小、形状、遮挡等情况各不相同，检测难度和重要性有区别，CIoU Loss 不做区分的处理方式会影响困难样本的检测效果，为解决这一问题，采用 WIoUv1 代替 CIoU Loss。

WIoUv1 使用动态非单调聚焦机制进行损失计算，通过预测框与真实框的 IoU 值统计特性动态调整梯度增益。对低质量样本降低梯度惩罚强度，如标注噪声或严重遮挡的极低 IoU 样本，避免模型被低质量样本影响。同时为中等质量样本分配更高关注度，引导模型优先优化对检测精度提升最有帮助的样本。

通过损失权重与 IoU 值的动态关联，实现对样本质量的自适应区分，抑制低质量样本干扰并显著提高模型对有效目标检测精度。

WIoUv1 计算式为

$$L_{\mathrm{WIoUv1}} = R_{\mathrm{WIoU}} L_{\mathrm{IoU}} \quad (6)$$

其中　$$R_{\mathrm{WIoU}} = \exp\left(\frac{(b_{c_x}^{g^t} - b_{c_x})^2 + (b_{c_y}^{g^t} - b_{c_y})^2}{c_w^2 + c_h^2}\right) \quad (7)$$

$$L_{\mathrm{IoU}} = 1 - IoU \quad (8)$$

式中　$L_{\mathrm{WIoUv1}}$——加权交并比损失函数

$R_{\mathrm{WIoU}}$——加权交并比惩罚项

$L_{\mathrm{IoU}}$——交并比损失函数

$b_{c_x}^{g^t}$、$b_{c_y}^{g^t}$——真实边界框中心 $x$、$y$ 坐标

$b_{c_x}$、$b_{c_y}$——预测边界框中心 $x$、$y$ 坐标

$c_w$、$c_h$——同时包围预测框与真实框的最小外接矩形宽度与高度

## 3　试验与结果

### 3.1　试验环境及参数设置

本试验环境基于 14v CPU Intel(R) Xeon(R) Platinum 8362 CPU@2.80 GHz 的 1 TB 固态硬盘进行配置，内存为 45 GB。显卡采用 Nvidia GeForce RTX 3090 24 GB，软件环境部署在 Windows 系统平台，CUDA 版本选择 12.1，Python 版本选择 3.12，深度学习框架为 Pytorch 2.3.0。

训练参数设置：输入图像分辨率固定为 640 像素×640 像素，批次大小为 16，初始学习率为 0.01，训练迭代次数为 200。

### 3.2　评价指标

使用 4 个指标评估 DSFM-BAI 模型的检测性能，包括精确率、召回率、平均精度均值和推理速度。

### 3.3　损失函数试验

WIoU 损失函数有 3 个版本，在相同的数据集和参数配置下对 3 个版本依次进行训练与测试，试验结果如表 1 所示。

**表 1　WIoU 改进模型测试结果**

**Tab.1　Improved model test results**

| 模型 | 精确率/% | 召回率/% | 平均精度均值/% | 推理速度/(f·s$^{-1}$) |
|---|---|---|---|---|
| YOLO 11n | 79.6 | 78.1 | 82.1 | 62.24 |
| WIoUv1 | 80.3 | 79.3 | 83.2 | 61.56 |
| WIoUv2 | 79.0 | 78.3 | 82.5 | 61.83 |
| WIoUv3 | 80.1 | 78.8 | 82.6 | 61.22 |

试验结果表明，WIoUv1 改进模型效果最好，与基线模型 CIoU 损失函数相比，精确率提高 0.7 个百分点，召回率提高 1.2 个百分点，平均精度均值提高 1.1 个百分点，虽然推理速度下降 0.68 f/s，且比 WIoUv2 低 0.27 f/s，但综合来看 WIoUv1 在检测精度与计算效率间拥有更好的平衡效果。WIoUv1 通过动态非单调聚焦机制，显著降低了低质量样本对

损失值的干扰,同时优先优化中等质量样本,使模型能够更高效地学习对检测性能提升贡献最大的样本。基于上述优势,选用 WIoUv1 损失函数代替 CIoU。

### 3.4 消融试验

为验证 DSFM - BAI 模型改进的各个模块以及最终组合的有效性,进行消融试验,试验结果如表 2 所示。试验在相同的数据集和参数配置下进行训练与测试,通过不同模块组合策略分别进行 6 组试验,对比不同改进方案的性能表现,重点考察 DSFM - BAI 模型在精确率、平均精度均值 mAP、召回率和推理速度上对比基线模型提升效果。

**表 2 不同消融方法试验结果**

**Tab. 2 Results of different experimental ablation methods**

| 序号 | RFCAConv | C3k2 - DSC | AFEM | WIoU | 精确率/% | 召回率/% | 平均精度均值/% | 推理速度/(f·s$^{-1}$) |
|---|---|---|---|---|---|---|---|---|
| 1 | - | - | - | - | 79.6 | 78.1 | 82.1 | 62.24 |
| 2 | √ | - | - | - | 80.2 | 81.3 | 83.7 | 60.53 |
| 3 | - | √ | - | - | 80.5 | 82.6 | 83.4 | 65.51 |
| 4 | - | - | √ | - | 81.8 | 82.0 | 85.1 | 56.29 |
| 5 | - | - | - | √ | 80.3 | 79.3 | 83.2 | 61.56 |
| 6 | √ | √ | √ | √ | 81.0 | 83.2 | 85.9 | 58.88 |

注:试验以基线模型为基准,√表示引入模块,-表示未引入模块。

由表 2 可知,试验 1 为基线模型未采用任何改进模块,取得精确率 79.6%、平均精度均值 82.1% 和召回率 78.1%,推理速度达到 62.24 f/s;单独引入 RFCAConv 模块,通过增强特征捕获能力,在精确率微增 0.6 个百分点至 80.2% 的同时,平均精度均值提升 1.6 个百分点至 83.7%,召回率提高 3.2 个百分点,推理速度降至 60.53 f/s;单独引入 C3k2 - DSC 模块,在保持较高运行效率的前提下,精确率提升 0.9 个百分点至 80.5%,平均精度均值提升 1.3 个百分点达到 83.4%,召回率显著提高 4.5 个百分点;单独引入 AFEM 模块,通过抑制背景噪声,突出病害特征,带来 3.0 个百分点最大单模块增长,平均精度均值达到 85.1%,精确率提升 2.2 个百分点至 81.8%,召回率提升 3.9 个百分点至 82.0%,但推理速度下降明显至 56.29 f/s;WIoU 损失函数通过动态非单调聚焦机制改进损失计算,精确率提升 0.7 个百分点至 80.3%,召回率提升 1.2 个百分点至 79.3%,平均精度均值提升 1.1 个百分点至 83.2%,推理速度维持在 61.56 f/s 与基线模型接近;4 个改进模块同时引入,模型展现出显著的协同效应:精确率达到 81.0%,平均精度均值达到 85.9%,较基线模型提升 3.8 个百分点;召回率同步提升至 83.2%,较基线模型提升 5.1 个百分点,尽管推理速度降至 58.88 f/s,但实现了精度与速度之间的平衡优化,仍满足实时检测能力。试验结果表明,改进算法 DSFM - BAI 在复杂场景下黄瓜病害检测中较基线模型取得了显著效果。

### 3.5 主流模型性能试验

选择 YOLO v5n 模型、YOLO v8n 模型、YOLO v10n 模型、YOLO 11n 模型与 DSFM - BAI 进行试验,验证 DSFM - BAI 在目标检测任务中的有效性。试验在相同数据集和参数配置下进行训练与测试,测试结果如表 3 所示。

**表 3 主流模型训练结果**

**Tab. 3 Mainstream model training results**

| 模型 | 精确率/% | 召回率/% | 平均精度均值/% | 推理速度/(f·s$^{-1}$) |
|---|---|---|---|---|
| YOLO v5n | 75.2 | 78.8 | 80.3 | 56.32 |
| YOLO v8n | 77.2 | 80.4 | 83.0 | 56.96 |
| YOLO v10n | 72.4 | 76.4 | 80.8 | 57.44 |
| YOLO 11n | 79.6 | 78.1 | 82.1 | 62.24 |
| DSFM - BAI | 81.0 | 83.2 | 85.9 | 58.88 |

由表 3 可知,与其他模型对比,DSFM - BAI 模型在检测精确率、召回率和平均精度均值都表现出明显优势。DSFM - BAI 与 YOLO v5n、YOLO v8n、YOLO v10n、YOLO 11n 模型相比,精确率分别提高 5.8、3.8、8.6、1.4 个百分点,召回率分别提高 4.4、2.8、6.8、5.1 个百分点,平均精度均值分别提高 5.6、2.3、5.1、3.8 个百分点。而 DSFM - BAI 模型推理速度比 YOLO v5n、YOLO v8n、YOLO v10n 分别提高 2.56、1.92、1.44 f/s,但比 YOLO 11n 降低 3.36 f/s。DSFM - BAI 模型比其他 4 种模型具有更好的在检测精确率和推理速度之间平衡能力,尤其在召回率和平均精度均值方面,DSFM - BAI 都展现了明显的优势,说明 DSFM - BAI 能够胜任具有复杂背景的黄瓜叶片病害检测任务。

使用本文模型与基线模型在相同的数据集和参数配置下进行训练与测试得到平均精度均值、召回率曲线如图 9 所示。从图 9 可以看出,DSFM - BAI 对比基线模型,平均精度均值存在明显优势,一直高

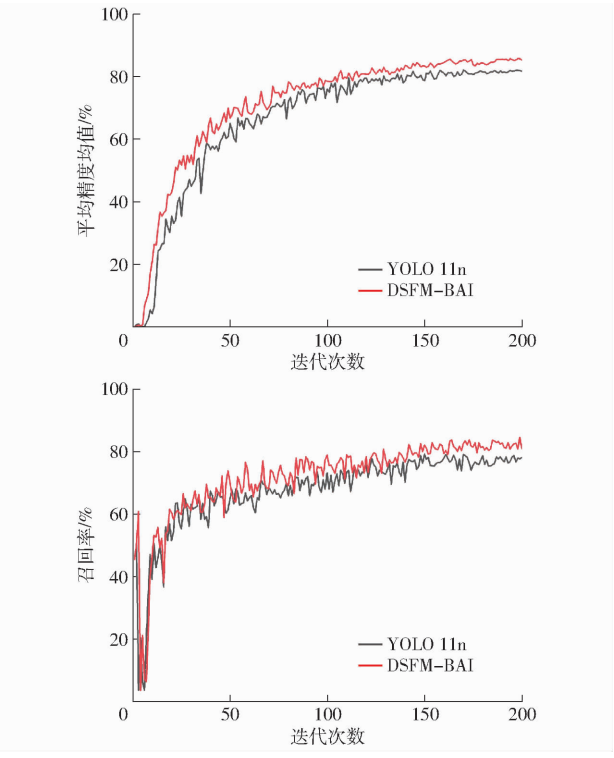


图 9　DSFM－BAI 与 YOLO 11n 平均精度均值和召回率结果

Fig. 9　mAP and recall results of DSFM－BAI and YOLO 11n

于基线模型；在试验前期 2 个模型召回率难分伯仲，随着试验的进行 DSFM－BAI 召回率在后期展现出明显优势。

### 3.6　在公开数据集上泛化试验

为了全面验证 DSFM－BAI 模型的场景适应能力，在公开的 leaf disease detection computer vision project 数据集上进行模型验证试验，将改进模型与基线模型进行对比，以进一步评估改进模型的检测能力。该数据集集成跨光照条件、多目标类别及多视角采集的图像，同时融合旋转、缩放、色彩扰动等数据增强策略生成的合成样本，包括交替霜霉病、白粉病、孢霉病、尾孢霉病和病毒性疾病等 5 个不同类别，数据集共包括 3 070 幅图像，数据集背景的复杂性和病灶特征，与本研究背景和问题相符。

试验结果如表 4 所示。由表 4 可知，DSFM－BAI 模型精确率、召回率和平均精度均值分别达到 66.2%、52.3% 和 55.1%，与基线模型相比，分别提升 2.3、4.4、2.7 个百分点。通过引入可变卷积模块，增强了模型对不规则病灶特征捕捉能力，在复杂背景下漏检率显著降低。尽管模型因为复杂度增加导致推理速度降至 58.7 f/s，但仍能满足农作物病害检测需求。

**表 4　YOLO 11n 和改进模型在公开数据集训练结果**

**Tab. 4　Training results of YOLO v11n and improved models on public datasets**

| 模型 | 精确率/% | 召回率/% | 平均精度均值/% | 推理速度/(f·s$^{-1}$) |
|---|---|---|---|---|
| YOLO 11n | 63.9 | 47.9 | 52.4 | 63.36 |
| DSFM－BAI | 66.2 | 52.3 | 55.1 | 58.70 |

试验结果验证了本文模型在复杂农业场景中检测有效性，实现了在检测精度与计算效率之间取得更优的平衡。

### 3.7　不同模型黄瓜病害检测可视化

使用训练的 DSFM－BAI、基线模型分别对自然场景下黄瓜叶片病害进行检测，得到可视化图像如图 10 所示，第 1 行为原始图像，第 2、3 行为基线模型和 DSFM－BAI 模型检测结果。DSFM－BAI 模型比基线模型对病害识别的精确率更高且检出了基线模型漏检的病害。在第 1 列新鲜无病害叶子、第 3

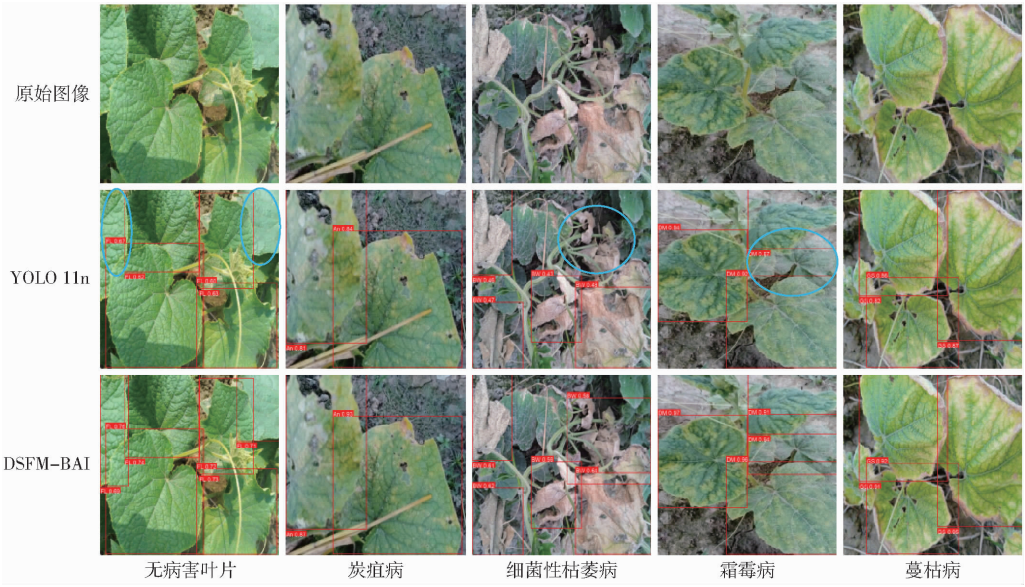


图 10　不同模型可视化结果

Fig. 10　Visualization results of different models

列细菌性枯萎病和第 4 列霜霉病的检测与识别中，在自然环境下，基线模型出现了漏检现象，DSFM－BAI 模型能有效识别被遮挡和形状不规则从而导致漏检的黄瓜叶片，检测精度较高。由第 2 列和第 5 列检测结果可看出，DSFM－BAI 模型有效识别了炭疽病与蔓枯病病斑的细节差异，检测精确率比基线模型均有大幅提高，解决了现有模型在实际应用中对复杂背景、形状不规则病害的检测精度不高，对相似病害的细节差异的识别精度不够的问题。

### 3.8 模型部署试验

为验证 DSFM－BAI 算法在实际应用中的效果，利用开发的 Android 端黄瓜病害识别系统开展现场试验，检测结果如图 11 所示。通过手机摄像头拍摄黄瓜图像或者选择图库中图像，图像进入系统后自动裁剪为 640 像素×640 像素，随后点击检测按钮，图像通过 DSFM－BAI 算法进行识别，识别结束后输出检测结果。




图 11 模型部署检测结果

Fig. 11 Model deployment detection result

部署试验在辽宁省绥中县黄瓜种植基地进行，共拍摄 500 幅图像进行试验测试，其中正确判断为 424 幅，错误判断为 76 幅，平均准确率为 84.8%，平均检测耗时为 0.352 s。判断错误的主要原因是拍摄过程抖动导致部分图像模糊，因此平均准确率低于实验室环境的准确率。所开发的黄瓜病害识别系统集成了训练的 DSFM－BAI 算法，具有较高的准确率和较短的检测耗时，在后续开发中，应加入去模糊算法以降低低质图像对检测结果的影响，进一步提高 DSFM－BAI 算法在自然环境中实用性。

## 4 结论

（1）提出了一种动态尺度特征挖掘与双向注意力交互的黄瓜病害检测算法，以 YOLO 11 为基础，在特征提取阶段引入 RFCAConv，通过感受野与通道协同注意力机制强化病斑区域的特征表征，使网络能够更精准地捕捉病斑特征。提出了自适应特征增强模块 AFEM，有效抑制背景噪声，增强目标与背景的区分度，突出病害特征。设计了 C3k2－DSC 模块，通过卷积核形状与尺度的动态适配机制，有效减少因目标形状变化导致的特征信息丢失。

（2）分别引入 RFCAConv、AFEM、C3k2－DSC 模块及 WIoU 损失函数，试验结果表明各改进模块均能提升检测效果，四者协同作用使模型的精确率、召回率、平均精度均值达较基线模型均有所提升，推理速度为 58.88 f/s，实现精度与实时性平衡。与主流检测模型相比，DSFM－BAI 综合性能更优。

（3）在 Android 手机端部署 DSFM－BAI 模型，自然环境下测试结果显示，平均识别精度为 84.8%，平均检测耗时为 0.352 s。引入的 RFCAConv 和 AFEM 虽提升了检测精度，但增加了计算复杂度，导致检测速度降低，在低算力移动设备上的实时性仍需优化，同时需要提高对强模糊或低分辨率图像的识别能力。

## 参 考 文 献

[1] 夏天，谢纯，李琳一，等. 基于改进 YOLO v8 的复杂温室环境黄瓜果实分割方法[J]. 农业机械学报，2025，56(5)：433－442.
XIA Tian, XIE Chun, LI Linyi, et al. Improved YOLO v8 method for cucumber fruit segmentation in complex greenhouse environments[J]. Transactions of the Chinese Society for Agricultural Machinery, 2025, 56(5): 433－442. (in Chinese)

[2] 李凯雨，朱昕怡，马浚诚，等. 基于混合扩张卷积和注意力的黄瓜病害严重度估算方法[J]. 农业机械学报，2023，54(2)：231－239.
LI Kaiyu, ZHU Xinyi, MA Juncheng, et al. Estimation method of leaf disease severity of cucumber based on mixed dilated convolution and attention mechanism[J]. Transactions of the Chinese Society for Agricultural Machinery, 2023, 54(2): 231－239. (in Chinese)

[3] 惠成章，刘术均，刘爱群. 新时期我国黄瓜产业高质量发展研究[J]. 农业科技与装备，2023(6)：85－87.
HUI Chengzhang, LIU Shujun, LIU Aiqun. Research on high quality development of cucumber industry in China in the new era [J]. Agricultural Science and Equipment, 2023(6): 85－87. (in Chinese)

[4] 张放. 2018 年我国主要水果生产统计简析[J]. 中国果业信息，2020，37(7)：32－43.
ZHANG Fang. Brief analysis of the production statistics of major fruits in China in 2018[J]. China Fruit News, 2020, 37(7): 32－43. (in Chinese)

[5] 王艳华,张锦东.黄瓜炭疽病的发生、传播与防治[J].农业工程技术,2021,41(8):41 - 42.
WANG Yanhua, ZHANG Jindong. Occurrence, transmission and prevention of cucumber anthracnose[J]. Agricultural Engineering Technology, 2021, 41(8): 41 - 42. (in Chinese)
[6] 秦立峰,李博棣,林敬轩,等.基于 SGD - YOLO 模型的黄瓜霜霉病检测[J].农业机械学报,2026,57(2):203 - 214.
QIN Lifeng, LI Bolai, LIN Jingxuan, et al. Cucumber downy mildew detection based on SGD - YOLO model[J]. Transactions of the Chinese Society for Agricultural Machinery, 2026,57(2):203 - 214. (in Chinese)
[7] 梁兆合.黄瓜蔓枯病的发生与综合防治技术[J].上海蔬菜,2024(3):68 - 69.
LIANG Zhaohe. Occurrence and comprehensive control technology of cucumber gummy stem blight[J]. Shanghai Vegetables, 2024(3): 68 - 69. (in Chinese)
[8] 王翔宇,李鑫星,张领先,等.农业主要病害检测与预警技术研究进展分析[J].农业机械学报,2016,47(9):266 - 277.
WANG Xiangyu, LI Xinxing, ZHANG Lingxian, et al. Research progress analysis of main agricultural diseases detection and early warning technologies[J]. Transactions of the Chinese Society for Agricultural Machinery, 2016, 47(9): 266 - 277. (in Chinese)
[9] 孙杰,杨静,丁书杰,等.基于多注意力机制与编译图神经网络的高光谱图像分类[J].农业机械学报,2024,55(3):183 - 192,212.
SUN Jie, YANG Jing, DING Shujie, et al. Hyperspectral image classification based on multi-attention mechanism and compiled graph neural networks[J]. Transactions of the Chinese Society for Agricultural Machinery, 2024,55(3):183 - 192,212. (in Chinese)
[10] 刘艺豪,周名扬,陈度,等.基于深度学习和视觉显著性算法的自然环境中烟叶成熟度判别[J].农业机械学报,2025,56(7):502 - 512.
LIU Yihao, ZHOU Mingyang, CHEN Du, et al. Maturity discrimination of tobacco leaves in natural environments based on deep learning and visual saliency algorithm[J]. Transactions of the Chinese Society for Agricultural Machinery, 2025,56(7): 502 - 512. (in Chinese)
[11] MOHANTY S P, HUGHE D P, SALATHÉ M. Using deep learning for image-based plant disease detection[J]. Frontiers in Plant Science, 2016,7: 1419 - 1429.
[12] RANGARAJAN A K, PURUSHOTHAMAN R, RAMESH A. Tomato crop disease classification using pre-trained deep learning algorithm[J]. Procedia Computer Science, 2018, 133: 1040 - 1047.
[13] ZHANG Y, HUANG S, ZHOU G, et al. Identification of tomato leaf diseases based on multi-channel automatic orientation recurrent attention network[J]. Computers and Electronics in Agriculture, 2023, 205: 107605.
[14] LECUN Y, BENGIO Y, HINTON G. Deep learning[J]. Nature, 2015, 521(7553): 436 - 444.
[15] 樊湘鹏,周建平,许燕,等.基于改进卷积神经网络的复杂背景下玉米病害识别[J].农业机械学报,2021,52(3):210 - 217.
FAN Xiangpeng, ZHOU Jianping, XU Yan, et al. Corn disease recognition under complicated background based on improved convolutional neural network[J]. Transactions of the Chinese Society for Agricultural Machinery, 2021, 52(3): 210 - 217. (in Chinese)
[16] 樊湘鹏,许燕,周建平,等.基于迁移学习和改进 CNN 的葡萄叶部病害检测系统[J].农业工程学报,2021,37(6):151 - 159.
FAN Xiangpeng, XU Yan, ZHOU Jianping, et al. Detection system for grape leaf diseases based on transfer learning and improved CNN[J]. Transactions of the CSAE, 2021, 37(6): 151 - 159. (in Chinese)
[17] 张惠莉,代晨龙,任景龙,等.基于 YOLO v8 - GSGF 模型的葡萄病害识别方法研究[J].农业机械学报,2024,55(11):75 - 83.
ZHANG Huili, DAI Chenlong, REN Jinglong, et al. Grape disease identification method based on YOLO v8 - GSGF model [J]. Transactions of the Chinese Society for Agricultural Machinery, 2024, 55(11): 75 - 83. (in Chinese)
[18] 黄丽明,王懿祥,徐琪,等. 采用 YOLO 算法和无人机影像的松材线虫病异常变色木识别[J]. 农业工程学报,2021,37(14):197 - 203.
HUANG Liming, WANG Yixiang, XU Qi, et al. Recognition of abnormal discolored trees caused by pine wilt disease using YOLO algorithm and UAV images[J]. Transactions of the CSAE, 2021, 37(14): 197 - 203. (in Chinese)
[19] 刘诗怡,胡滨,赵春.基于改进 YOLOv7 的黄瓜叶片病虫害检测与识别[J].农业工程学报,2023,39(15):163 - 171.
LIU Shiyi, HU Bin, ZHAO Chun. Detection and identification of cucumber leaf diseases and pests based on improved YOLOv7 [J]. Transactions of the CSAE, 2023, 39(15): 163 - 171. (in Chinese)
[20] ZHANG H, CISSE M, DAUPHIN Y N, et al. Mixup: beyond empirical risk minimization[J]. arXiv Preprint arXiv: 1710.09412,2017.
[21] 罗臻,蒋瑜,李维诚.基于改进 YOLO11n 的道路缺陷检测模型[J].微电子学与计算机, 2025,42(11):1 - 13.
LUO Zhen, JIANG Yu, LI Weicheng. Road defect detection model based on improved YOLO11n[J]. Microelectronics and Computer, 2025, 42(11): 1 - 13. (in Chinese)
[22] ZHANG X, LIU C, YANG D, et al. Rfaconv: innovating spatial attention and standard convolutional operation[J]. arXiv Preprint arXiv:2304.03198,2024.
[23] QI Y L, HE Y T, QI X M, et al. Dynamic snake convolution based on topological geometric constraints for tubular structure segmentation[C]// IEEE/CVF International Conference on Computer Vision,2023: 6070 - 6079.
[24] ALE L H, SHETA A, LI L Z, et al. Deep learning based plant disease detection for smart agriculture[C]// 2019 IEEE Globecom Workshops(Gc Wkshps). Waikoloa Hi: IEEE,2019:1 - 6.
[25] LIU Y, LI H, HU C. Learning to aggregate multi-scale context for instance segmentation in remote sensing images[J]. arXiv Preprint arXiv:2111.11057,2021.
[26] 李耀龙,陈晓林,林浩,等. DySnake - YOLO:改进的 YOLOv9c 电路板表面缺陷检测方法[J].计算机工程与应用,2025,61(3):242 - 252.
LI Yaolong, CHEN Xiaolin, LIN Hao, et al. DySnake - YOLO: improved surface defect detection method for YOLOv9c circuit board[J]. Computer Engineering and Applications, 2025, 61(3): 242 - 252. (in Chinese)