doi:10.6041/j.issn.1000-1298.2026.03.011

基于无人机遥感图像的水稻稻穗识别与估产方法研究

龙拥兵[1,2]　李　坤[1,2]　刘丹丹[1,2]　龙　腾[1,2]　潘朝阳[1,2]　陈文瀚[1,2]　贺　敏[1,2]
郭炜伦[1,2]　李俊杰[1,2]　陈荣坤[1,2]　康　辉[1,2]　兰玉彬[1,2]

(1.华南农业大学电子工程学院(人工智能学院),广州 510642;
2.国家精准农业航空施药技术国际联合研究中心,广州 510642)

摘要:针对水稻稻穗识别中跨生育期精度低、复杂田间环境易漏检等问题,本文提出一种基于改进 YOLO 11 的旋转目标检测模型 AHF-YOLO 11。该模型通过在 YOLO 11 主干网络中引入卷积视觉模块 AssemFormer 来构建 C3k2_AssemFormer 特征提取模块,显著增强密集遮挡场景下的局部特征表达与跨生育期全局形态学习能力;进一步采用高级筛选功能金字塔网络 HSFPN 替代原模型颈部结构,有效提升模型对不同背景下多尺度稻穗的识别能力。试验表明,改进模型 AHF-YOLO 11 的准确率达到 90.7%,较原 YOLO 11 模型提升 4.3 个百分点,相对于主流模型分别提高 4.4~39 个百分点。跨生育期测试试验结果表明,AHF-YOLO 11 模型在孕穗、抽穗、灌浆、成熟期的识别准确率比原模型分别提高 11.88、7.3、4.78、4.3 个百分点,并在抽穗期的检测精度最高,达到 94.41%。进一步跟踪研究和估产试验表明,抽穗末期-灌浆初期期间为最佳估产时间,估产误差最低仅为 4.66%。本研究为水稻穗数识别和估产提供了重要技术支撑。

关键词:水稻稻穗;目标检测;产量预估;深度学习;无人机
中图分类号:TP391.4　　**文献标识码**:A　　**文章编号**:1000-1298(2026)03-0109-10　　**OSID**:

Rice Panicle Recognition and Yield Estimation Based on UAV Remote Sensing Images

LONG Yongbing[1,2]　LI Kun[1,2]　LIU Dandan[1,2]　LONG Teng[1,2]　PAN Chaoyang[1,2]　CHEN Wenhan[1,2]
HE Min[1,2]　GUO Weilun[1,2]　LI Junjie[1,2]　CHEN Rongkun[1,2]　KANG Hui[1,2]　LAN Yubin[1,2]

(1. *College of Electronic Engineering(College of Artificial Intelligence)*, *South China Agricultural University*, *Guangzhou* 510642, *China*
2. *National Center for International Collaboration Research on Precision Agricultural Aviation Pesticides Spraying Technology*, *Guangzhou* 510642, *China*)

Abstract: In order to address the problems of low detection accuracy across different rice growth stages and easily missed detection under complex field environments, an improved rotated object detection model AHF-YOLO 11 was proposed based on YOLO 11. This model introduced convolutional vision module AssemFormer into YOLO 11 backbone network to construct C3k2_AssemFormer feature extraction module, significantly enhancing the local feature expression and cross reproductive global morphology learning ability in dense occlusion scenes. Furthermore, it adopted the high-level screening feature pyramid network (HSFPN) to replace the original neck structure, effectively improving the ability of the model to recognize multi-scale rice panicles under varying backgrounds. Experimental results demonstrated that the accuracy of the improved AHF-YOLO 11 model reached 90.7%, which represented an improvement of 4.3 percentage points higher than that of the original YOLO 11 model and 4.4~39 percentage points higher than the mainstream models, respectively. The results of cross-growth-stage testing further revealed that the recognition accuracy of AHF-YOLO 11 in the booting, heading, filling, and maturity stages was respectively improved by 11.88, 7.3, 4.78, and 4.3 percentage points when compared with that of the original model. Among these, the highest accuracy of 94.41% was achieved for the model at the heading stage. Further follow-up tracking studies and yield estimation

收稿日期:2025-09-28　修回日期:2025-11-21
基金项目:农业农村部科技项目和高等学校学科创新引智计划项目(D18019)
作者简介:龙拥兵(1979—),男,教授,博士,主要从事作物表型研究,E-mail:yongbinglong@scau.edu.cn
通信作者:兰玉彬(1961—),男,教授,博士,主要从事精准农业航空研究,E-mail:ylan@scau.edu.cn

experiments indicated that the period from the late heading stage to the early filling stage was the optimal period for rice yield estimation, with the lowest yield estimation error of only 4.66%. The research result can provide important technical support for rice panicle number recognition and yield estimation.

Key words: rice panicle; object detection; yield estimation; deep learning; unmanned aerial vehicle

0 引言

水稻作为全球和我国最重要的粮食作物,其产量精准估测对农场种植及政府管理决策具有重要意义[1-3]。稻穗作为水稻的核心生殖器官和表型性状载体,其高精度检测直接关系产量预测的可靠性[4]。传统估产方法依赖人工田间取样,通过统计单位面积穗数与单穗粒数进行推算,存在效率低下、劳动强度大等局限性[5-6]。因此,研发自动化稻穗识别技术,实现高效精准的穗数统计与产量估测,已成为当前农业智能化领域的重点研究方向。

基于深度学习的目标检测技术已成为农业领域小目标检测与计数的研究热点[7],在病虫害[8-9]、玉米雄穗[10-11]和水稻稻穗识别[12-17]等方面取得了重要进展。

在水稻稻穗识别方面,张远琴等[12]基于改进 Faster R-CNN 提出针对小目标优化的稻穗检测方法,mAP 达 80.3%,为视频稻穗自动计数提供了可靠方案。SUN 团队[13]通过改进 YOLO v4 模型,结合 4 个生态点不同类型杂交籼稻穗的无人机图像数据,提出通用弯穗检测方法,实现 90.32% 的检测精度,显著提升了重叠穗与叶片遮挡场景的识别能力。TAN 等[14]借助手机近景成像,提出基于改进 Cascade RCNN 的 RiceRes2Net,在孕穗期、抽穗期和灌浆期平均精度分别为 96.8%、93.7% 和 82.4%,验证了该方法对复杂田间环境和生育期变化的适应性。LAN 团队[15]则采用无人机低空图像,设计 RICE-YOLO 算法,在 YOLO v5 框架中引入 EMA 注意力机制和 SIoU 损失函数平均精度和召回率分别达到 94.8% 和 87.6%。WEI 等[16]提出了基于无人机遥感图像的 YOLO v8s + LSKA + HorNet 模型,结合穗数、千粒质量等关键指标建立回归模型,实现稳定产量预测。LIANG 等[17]使用无人机超低空图像,开发旋转式稻穗检测模型,并进一步验证了田间估产可行性。

这些研究基于小型无人机和手机拍摄图像为稻穗智能检测提供了多样化技术路径,并取得了良好的效果。但在实际应用中,田间稻穗识别仍然面临 3 个核心问题:①稻穗细长形态与随机取向特性导致传统水平检测产生大量冗余像素,显著降低识别精度。②植株结构复杂性与穗部密集分布引发穗-叶-茎相互遮挡,其交叉率与重叠率远超普通目标检测场景,导致误检漏检现象尤为突出。③受品种差异与生育阶段双重影响,穗部形态呈现显著多尺度变化,其尺寸、色泽及几何形状在生长周期中的动态变化进一步增加了特征提取难度[18]。这些因素共同构成了复杂田间环境下稻穗精准检测的技术瓶颈。

针对上述难点和技术瓶颈,本文以华南农业大学增城教学实验基地种植的水稻为研究对象,构建覆盖孕穗-抽穗-灌浆-成熟期的全生育期无人机遥感图像多时序数据集,基于 YOLO 11 旋转检测模型,提出一种跨生育期旋转稻穗精准识别算法。通过在 YOLO 11 中 C3k2 模块内部添加 AssemFormer 卷积视觉模块,构建 C3k2_AssemFormer 模块,有效增强对稻穗跨生育期形态特征学习能力,进一步添加高级筛选功能金字塔网络(HSFPN)实现多尺度特征融合,提高模型识别能力,并同步开展稻穗取样的地面试验,跟踪稻穗数量变化趋势,分析得到最佳测产时间,为稻穗识别估产提供技术参考。

1 材料与方法

1.1 数据采集与数据集建立

1.1.1 无人机遥感图像采集与数据集建立

田间试验在广东省华南农业大学增城宁西基地(23°14′18″N,113°38′8″E)开展,如图 1 所示。田间试验种植了美香占 2 号、象牙香占、南晶香占、华航香银针、美巴香占 5 个水稻品种。同时,设置零氮(0 kg/hm²)、中氮(120 kg/hm²)、标准氮(240 kg/hm²)、高氮(480 kg/hm²) 4 个施肥梯度,以

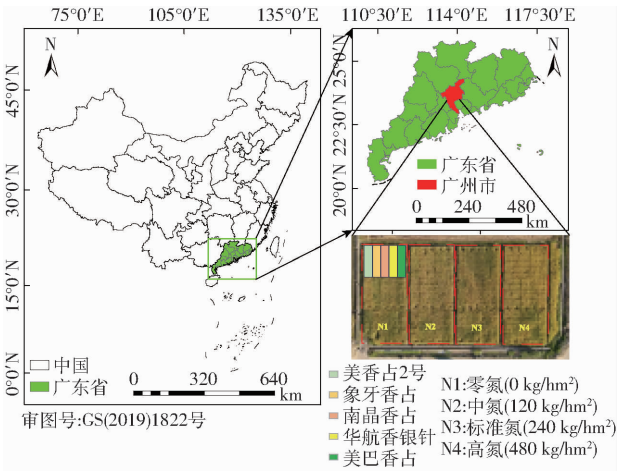


图 1 试验区域示意图

Fig. 1 Schematic of experimental area

增加水稻冠层无人机遥感图像特征的复杂性和多样性[19]。每个品种、每个施肥梯度设计 3 个重复，共有 60 个种植小区，每个小区长和宽分别为 9.7 m 和 2.4 m，种植小区面积为 23.28 m^2。水稻种植为 2024 年 8 月中旬，于 10 月初至 11 月初进行无人机遥感图像采集和水稻稻穗计数试验。

采用大疆 M30T 型无人机在 12:00—14:00 和 15:00—17:00 时段采集图像数据，该无人机配备变焦相机，有效像素为 4 800 万，支持 5 ~ 16 倍光学变焦，最大支持 200 倍数码变焦，通过人工飞行模式控制无人机在 10 m 高空以 10 倍光学变焦拍摄，获取分辨率为 4 000 像素 ×3 000 像素的高清图像，地面采样距离 GSD 为 0.31 cm/像素。为增强数据多样性，研究特别设计了多天气条件和跨生育期的采集方案，增加了样本数据的多样性。

本研究采用 RoLabelimg[20] 旋转标注工具对无人机采集的水稻冠层图像进行标注，构建了适用于 YOLO 框架的水稻穗部检测专用数据集。数据集覆盖抽穗期、灌浆期等关键生育阶段的无人机遥感图像，按比例 8:1:1 划分为训练集、验证集和测试集。数据集中各生育期图像数量与大、中、小目标标签数量如表 1 所示，共包含 45 240 个带注释的稻穗，其中小目标（尺寸小于 32 像素 ×32 像素）、中目标（尺寸在 32 像素 ×32 像素和 96 像素 ×96 像素之间）和大目标（尺寸大于 96 像素 × 96 像素）分别为 40 621、4 044、575 个，所占比例分别为 89.79%、8.94% 和 1.27%。图 2 展示了数据集中对象的大小分布，其中大多数稻穗相对于整个图像的纵横比低于 0.1[21]，图 2g 定量描述目标高宽关系，红色矩形框标示大多数小目标。小尺寸和低纵横比限制了图像所携带的稻穗细节，增加了算法识别难度。

表 1　稻穗数据分布情况

Tab.1　Distribution of rice panicle data

类别	图像数量	小目标	中目标	大目标
孕穗期	175	486	0	0
抽穗期	412	13 801	429	78
灌浆期	148	18 136	2 067	282
成熟期	60	8 198	1 548	215
总数	795	40 621	4 044	575
占比/%		89.79	8.94	1.27

1.1.2　地面稻穗计数

为了验证无人机遥感识别结果，本研究于孕穗末期（10 月初）至成熟期（11 月初）开展地面穗数跟踪计数试验。单生育期内设置 1 次采样节点，共 6 次地面数据采集。每次采样前预留 2 行以消除边际效应，随后在 5 个品种种植区域内分别随机选取 3 穴水稻，通过人工计数记录穗数，用于后续与无人机遥感识别结果进行匹配与识别精度验证。

1.2　水稻稻穗识别模型 AHF－YOLO 11

1.2.1　模型改进

为了提升对复杂环境下不同生育期水稻稻穗的识别精度，本研究提出一种基于 YOLO 11 改进的水稻稻穗识别模型 AHF－YOLO 11。YOLO 11 是 UltralyticsYOLO 实时物体检测器系列中的模型，该模型通过引入 C3k2 模块和 C2PSA 模块实现更高效的特征提取与多尺度感知[22]。在 YOLO 11 模型基础上，进行如下改进：首先在 C3k2 模块中应用了基于组装的卷积视觉 Transformer，称为 AssemFormer[23]，提升模型对跨生育期稻穗表型特征的学习能力；其次，应用 HSFPN 网络结构[24]，替换原有的颈部网络结构，实现多尺度特征融合，提高模型的特征表达能力，增强模型的检测能力。图 3 为改进后的 AHF－YOLO 11 模型结构图。

1.2.2　C3k2_AssemFormer 卷积视觉模块

为了提高模型在复杂背景下捕获稻穗的局部细粒度纹理与跨生育期全局形态，首先在 YOLO 11 中 C3k2 模块内部添加 AssemFormer 卷积视觉模块，构建 C3k2_AssemFormer 模块，该模块在保持 CSP 多分支融合能力的同时，结合 AssemFormer 卷积视觉模块对局部特征和全局信息的感知能力，能够使模型更好地捕获稻穗的局部细粒度纹理与跨生育期全局形态特征，并增强对稻穗关键区域的敏感度，同时抑制叶片遮挡等无关的上下文干扰，提升孕穗期至成熟期稻穗识别的细粒度精度与鲁棒性。

构建 C3k2_AssemFormer 卷积视觉模块分为以下 2 个步骤：

1.2.2.1　构建 AssemFormer 卷积视觉模块

AssemFormer 是一种基于组装的卷积视觉 Transformer 模块[23]，其结构如图 4 所示，该模块利用整合–分离策略，将卷积的局部感知和 Transformer 的全局建模结合起来，进而提高稻穗识别任务的准确性。分为以下几个步骤：

（1）稻穗局部特征提取：3 ×3 和 1 ×1 卷积对稻穗图像特征图 $X \in \mathbf{R}^{B\times H\times W\times C}$ 进行处理，快速捕获稻穗边缘、颜色等低级特征和局部上下文。其中 B 是批次大小，C 是通道数，H 和 W 分别为特征图的高度和宽度。

（2）空间到序列变换：将卷积的输出展开为 N 个不重叠的 patch，即 Transformer 所需的序列格式，降低后续全局自注意力的计算复杂度。

（3）全局关系建模：使用 Transformer 对 P 个 patch 执行自注意力运算，提取全局特征并对全局信

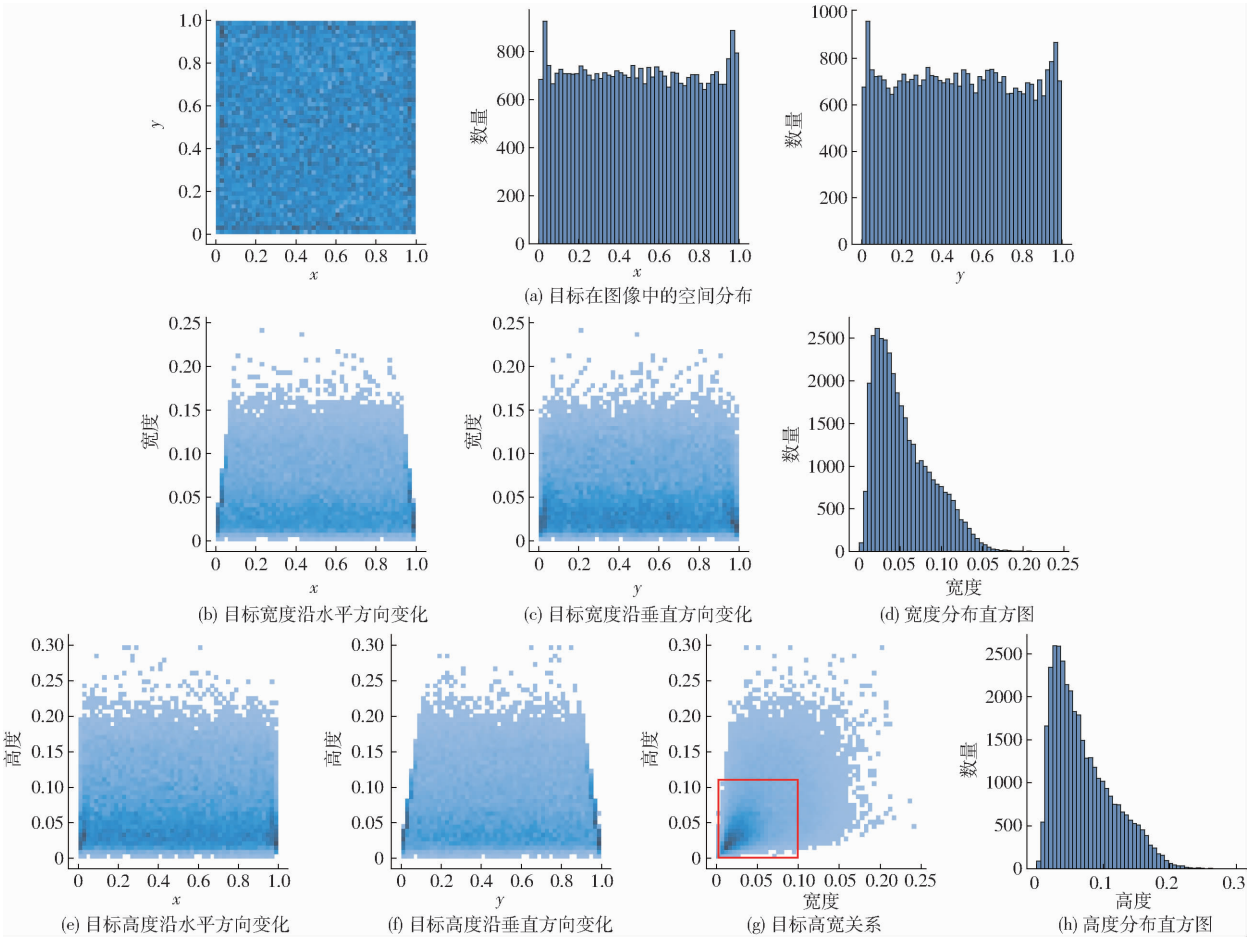


图 2 稻穗在数据集中的分布和尺寸

Fig. 2 Distribution and size of rice panicles in dataset

息进行编码,建立远距离像素间的关系,诸如叶片、茎秆与穗交叉遮挡之间的关系,缓解交叉遮挡导致的漏检。Transformer 算法采用一系列多头自注意力机制(MHSA)和多层感知器(MLP)块,每个块后面都有层归一化自我注意机制,公式[25]为

$$A_{ViT}(\boldsymbol{q},\boldsymbol{k},\boldsymbol{v}) = \text{softmax}\left(\frac{\boldsymbol{q}\boldsymbol{k}^{\mathrm{T}}}{\sqrt{D_h}}\right)\boldsymbol{v}$$

其中 $$D_h = D/m$$

式中 A_{ViT}——采用多头自注意力机制的 Vision Transfomer 模块

$\boldsymbol{q}$、$\boldsymbol{k}$、$\boldsymbol{v}$——输入序列 $z \in \mathbf{R}^{N\times D}$ 的查询、键、值向量

N——块的数量 D——块尺寸

m——自注意操作数量

D_h——$\boldsymbol{q}$ 和 $\boldsymbol{k}$ 的维数

(4)序列到空间变换:将 Transformer 的输出转换为卷积所需的输入格式,即将序列还原为空间特征图,并与原始局部特征拼接,经 1×1 卷积压缩通道,实现局部和全局信息融合。

1.2.2.2 构建 C3k2_AssemFormer 模块

在 C3k2 模块的特征融合末端嵌入 AssemFormer 卷积视觉子模块,在不破坏 CSP 多分支梯度流的前提下,提供更具判别力的特征表示。

1.2.3 高级筛选功能金字塔网络(HSFPN)

为了更好地适应稻穗在复杂田间环境下的多尺度变化并抑制叶片和田间背景等冗余信息,在颈部结构中引入 HSFPN 网络替换原始颈部网络,利用高级特征筛选低级特征重建特征融合部分。HSFPN 主要由特征选择模块 CA 和特征融合模块 SFF 组成[24,26],其结构如图 5a 所示。特征选择模块 CA 首先对不同尺度特征图进行选择,保留关键信息,在此基础上,特征融合模块 SFF 将选择后的高级语义特征与低级细节特征结合,构建富含空间信息且抑制无关信息干扰的融合特征,提升模型对多尺度稻穗的识别能力,降低因尺度变化与特征不足带来的检测误差。HSFPN 构建过程可分为如下 2 个步骤:

(1)特征选择模块 CA

特征选择模块 CA 网络结构如图 5b 所示,CA 模块首先处理输入的特征图$f_{\text{in}} \in \mathbf{R}^{C\times H\times W}$。特征图经并行全局平均池化层和全局最大池化层处理后得到每个通道的平均值和最大值。将二者相加后通过 Sigmoid 激活函数确定每个通道的权重值$f_{\text{CA}} \in$

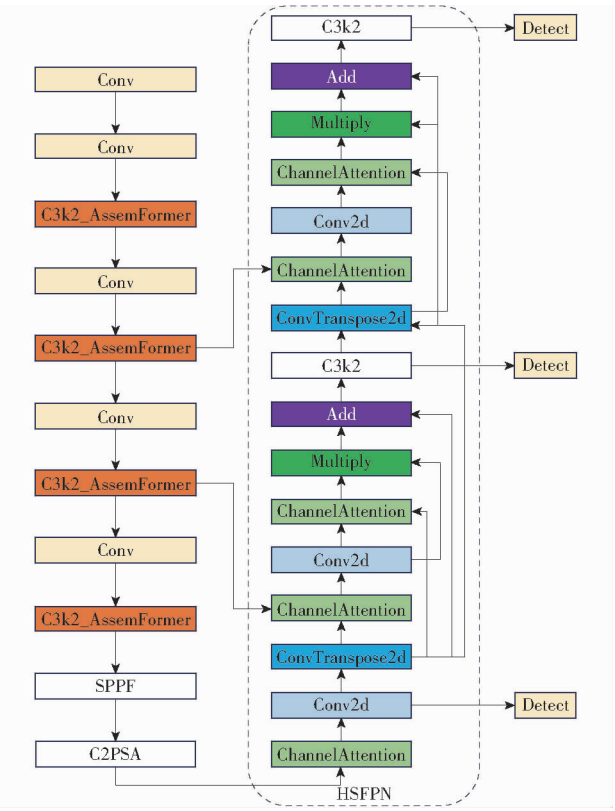


图 3　AHF－YOLO 11 模型结构图

Fig. 3　Architecture of AHF－YOLO 11 model

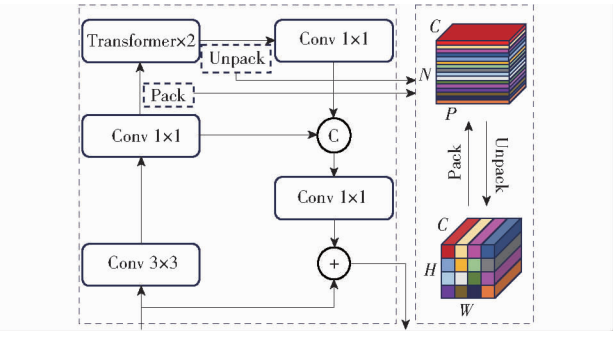


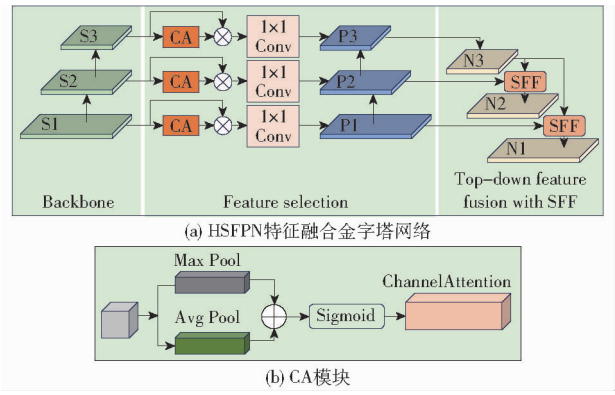
图 4　AssemFormer 结构图

Fig. 4　Structural diagram of AssemFormer module

$\mathbf{R}^{c\times1\times1}$。随后将权重信息与对应比例的特征图相乘，得到滤波后的特征图，并对不同尺度的特征图进行维度对齐，为后续选择特征融合提供尺度统一且富含语义的输入。最大池化层侧重从每个通道中提取最相关的数据，平均池化层侧重从特征图中均匀地获取所有数据，二者组合最大限度地减少过度损失。

（2）特征融合模块 SFF

骨干网络生成的多尺度特征图中，高级特征尽管语义信息丰富，但目标定位精度较低，低级特征定位精确却语义匮乏。为充分利用两者优势，SFF 摒弃了传统的跨尺度特征简单逐像素相加方式，而是以高级特征为权重，对低级特征进行加权融合，在保持低级特征定位精确性的同时强化其语义表达。如



(a) HSFPN特征融合金字塔网络

(b) CA模块

图 5　HSFPN 与特征选择模块结构图

Fig. 5　Architecture of HSFPN and feature selection module

图 6 所示，输入一个高级特征$f_{\text{high}}\in\mathbf{R}^{C\times H\times W}$和一个低级特征$f_{\text{low}}\in\mathbf{R}^{C\times H\times W}$，首先使用转置卷积和双线性插值对两者进行维度统一采样，随后，利用 CA 模块将高层特征转换为注意力权重，以过滤低级特征。最后，将过滤后的低级特征与高层特征融合，增强模型的特征表示能力。双线性插值通过对像素直接操作，实现图像的缩放。转置卷积通过卷积核和可学习参数对扩展后的特征图进行卷积运算，该操作不仅实现特征图的上采样和输入特征的重建，还通过在输出的不同位置对输入的不同区域进行采样，有效处理非均匀采样问题。

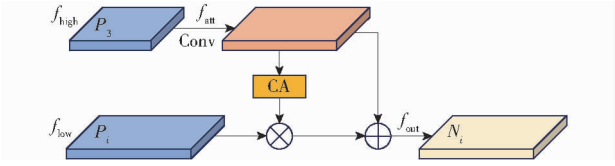


图 6　SFF 模块的框架结构图

Fig. 6　Framework of SFF module

1.3　试验平台与评价指标

（1）试验平台与参数设置

本文试验环境是 13th Gen Intel ® Core™ i9－13900K 中央处理单元（CPU）（3.00 GHz 处理器），64 GB 运行内存，24 GB NVIDIA RTX4090D 图形处理单元（GPU），操作系统采用 Windows，CUDA 版本为 CUDA 12.4，PyTorch 2.4.1 深度学习框架和 Python 3.8 来构建训练、验证和检测模型。

（2）模型评价指标

采用目标识别中常用的准确率（Accuracy，%）、精确率（Precision，P，%）、召回率（Recall，R，%）、F1 分数（F1）、参数量（Params）、计算量（FLOPs）、模型大小作为模型的评价指标[27]。

2　结果与分析

2.1　模型性能与消融试验

为进一步分析每个模块在模型中的作用，本研究开展消融试验，结果如表 2 所示。改进模型

AHF-YOLO 11 的精确率、准确率和召回率分别达到 90.4%、90.7% 和 83.8%，较原模型 YOLO 11n 分别提高 4.7、4.4、4.0 个百分点。另外，AHF-YOLO 11 模型大小仅为 7.3 MB，略高于 YOLO 11n (5.9 MB)，表明改进模型在模型体积与精度之间取得较好平衡。

表 2 消融试验结果

Tab. 2 Ablation experiments

模型	C3k2_AssemFormer	HSFPN	精确率/%	准确率/%	召回率/%	F1 分数/%	参数量	计算量	模型大小/MB
YOLO 11n	×	×	85.7	86.3	79.8	83	2.65×10^{6}	6.60×10^{9}	5.9
YOLO 11n+	√	×	88.1	89.4	83.7	86	2.94×10^{6}	1.03×10^{10}	6.4
YOLO 11n+	×	√	86.8	87.9	82.1	84	1.95×10^{6}	6.60×10^{9}	4.5
AHF-YOLO 11	√	√	90.4	90.7	83.8	86	3.35×10^{6}	1.39×10^{10}	7.3

注：√表示使用该模块，×表示不使用该模块。

从表 2 可知，将 AssemFormer 卷积视觉模块应用到 C3k2 模块中，准确率和召回率相较于 YOLO 11 分别提高了 3.1、3.9 个百分点。这主要是因为水稻稻穗形态随生育期改变而发生明显变化：在孕穗期，稻穗呈纺锤形紧凑结构；在成熟期，稻穗呈弓形松散形态，叶片、茎秆与穗交叉遮挡增加了稻穗形态和所处环境的复杂性。YOLO 11 中的传统卷积池化过程极易丢失稻穗的微弱边缘与纹理特征，造成小目标漏检。而 AssemFormer 以卷积-Transformer 双分支并行提取特征，卷积分支保留稻穗边缘锯齿等高频纹理，避免传统池化造成的细节丢失，Transformer 分支通过多头自注意力捕获跨生育期稻穗形态演变关系，动态抑制叶茎遮挡等无关上下文干扰，C3k2 的并行卷积和通道分割结构进一步增大模型多尺度感受野，提供了更为丰富的特征，从而提高模型精度。

另一方面，水稻田间环境致使稻穗、叶片与茎秆密集交织重叠，导致稻穗形状多尺度形变和特征稀缺问题，原始 YOLO 11 颈部网络难以适应稻穗在田间环境下的尺度变化，且缺乏动态权重抑制机制，冗余信息削弱了对稻穗特征的提取能力。HSFPN 网络替换 YOLO 11 模型原始颈部网络后，模型准确率和召回率相较于 YOLO 11n 分别提高了 1.6、2.3 个百分点。HSFPN 由 CA(特征选择模块)与 SFF(特征融合模块)构成，CA 利用通道注意力对每层特征图加权筛选，突出稻穗关键通道并抑制背景噪声；SFF 依据 CA 生成的动态权重，将筛选后的高级语义与低级细节进行选择性融合，可以保留稻穗纹理、穗轴弯曲等细微特征的同时，生成富含跨生育期多尺度形态信息的融合特征，提升模型在不同背景下对稻穗的检测精度。

2.2 可视化试验

为直观评估改进模型 AHF-YOLO 11 对水稻稻穗的识别能力，本文采用增强版梯度加权激活图[28](Gradient-weighted class activation mapping plus plus，Grad-CAM++)对输入图像生成类激活热力图。相较传统 Grad-CAM[29]，Grad-CAM++采用二阶梯度信息对多通道特征图进行加权，可在目标较小、密集分布的场景下获得更高分辨率的注意力映射，更适合无人机水稻稻穗图像。由图 7 可见，AHF-YOLO 11 的高响应区域主要集中在稻穗目标本身；相比之下，YOLO 11n 的热力图响应存在分散且伴随误识别现象。表明 AHF-YOLO 11 中引入的 AssemFormer 模块通过局部与全局双分支并行特征提取机制，有效增强了对稻穗局部细节与跨生育期形态的特征提取能力；而 HSFPN 借助 CA 与 SFF 策略在特征融合过程中抑制无关上下文干扰，进一步提升了特征融合能力，二者协同作用，使得模型对稻穗的检测精度和鲁棒性均得到显著提升。

2.3 不同深度学习模型在稻穗识别领域性能测试

进一步对比研究了 AHF-YOLO 11 模型与 YOLO v8n[30]、YOLO 11n、YOLO 12n[31]、Oriented R-CNN[32]、Rotated Faster R-CNN[33]、Rotated-Retinanet[34]、ROI-Transformer[35] 主流模型的性能参数，如表 3 所示。在对比试验中，所有模型输入图像尺寸为 4 000 像素×3 000 像素，训练批次设置为

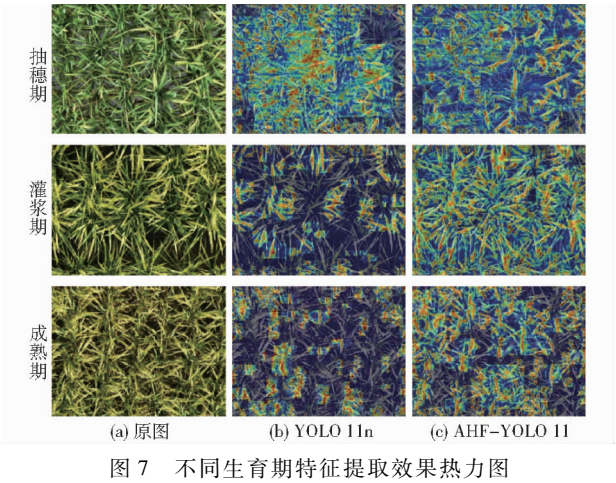


图 7 不同生育期特征提取效果热力图

Fig. 7 Heatmaps of feature extraction effects at different growth stages

8，初始学习率为 0.002。改进网络 AHF－YOLO 11 准确率相较于 YOLO v8n、YOLO 11n、YOLO 12n、Oriented R－CNN、Rotated Faster R－CNN、Rotated－Retinanet、ROI－Transformer 分别高 5.7、4.4、4.6、33.4、38.6、32.3、39 个百分点。本研究在 YOLO 11 的基础上进行改进，改进后的 AHF－YOLO 11 与同系列 YOLO 目标识别检测器相比，引入了 AssemFormer 卷积视觉模块和 HSFPN 网络，增强了模型对稻穗局部特征能力和稻穗跨生育期形态感知能力，提升模型对稻穗的检测精度。Oriented R－CNN、Rotated Faster R－CNN、ROI－Transformer 网络结构复杂，在高分辨率大尺寸图像下计算量过大，且训练过程中容易受到角度回归不稳定的影响，导致精度下降。Rotated－RetinaNet 依赖 Anchor 机制进行角度预测，存在角度不稳定和定位偏差的问题，尤其在稻穗这种密集小目标检测任务中表现不佳。

表 3　不同模型性能对比

Tab.3　Performance comparison of different models

模型	准确率/%	召回率/%	参数量	计算量	模型大小/MB
YOLO 11n	86.3	78.9	2.65×10^{6}	6.6×10^{9}	5.9
YOLO v8n	85.0	78.0	3.08×10^{6}	8.3×10^{9}	6.6
YOLO 12n	86.1	79.0	2.58×10^{6}	6.1×10^{9}	5.8
Oriented R－CNN	57.3	65.8	4.113×10^{7}	$2.273\,9\times10^{11}$	158
Rotated Faster R－CNN	52.1	61.3	4.112×10^{7}	$2.273\,9\times10^{11}$	315
Rotated－Retinanet	58.4	47.7	3.642×10^{7}	$1.621\,8\times10^{11}$	285
ROI－Transformer	51.7	61.0	5.503×10^{7}	$2.273\,9\times10^{11}$	421
AHF－YOLO 11	90.7	83.8	3.35×10^{6}	1.39×10^{10}	7.3

2.4　不同生育期模型检测效果

本文对比研究了 AHF－YOLO 11 和 YOLO 11n 模型对不同生育期图像中稻穗的检测效果，试验使用的图像数量、图像包含的稻穗数量以及模型检测效果如表 4 所示。试验中，针对模型识别后的稻穗图像，本文采用人工统计的方法获得正确检测、错误检测以及重复检测的穗数数量。从表 4 可以得到 2 个结论：

（1）整个生育周期内 AHF－YOLO 11 准确率均优于 YOLO 11n。在孕穗期、抽穗期、灌浆期和成熟期，AHF－YOLO 11 模型的识别准确率为 67.33%、94.41%、83.40% 和 74.43%，与 YOLO 11 相比，识别准确率分别提升 11.88、7.30、4.78、4.3 个百分点。这主要是因为 AHF－YOLO 11 模型中引入的 AssemFormer 视觉卷积模块和 HSFPN 网络，不但增强模型对稻穗局部细微特征的精准捕捉能力，同时还有效提升模型对稻穗跨生育期形态演变过程（从孕穗期、抽穗期的纺锤紧凑形态，到灌浆期和成熟期的弓形松散形态）的动态感知精度，进而生成蕴含稻穗跨生育期形态差异信息的高质量融合特征，最终提升模型在不同生育期条件下对稻穗的检测精度。

表 4　不同生育期条件下模型性能

Tab.4　Comparative performance of model at different rice growth stages

生育期	模型	检测数量/幅	稻穗数量/个	正确检测数/个	误检数/个	重复检测数/个	准确率/%
孕穗期	AHF－YOLO 11	27	101	68	20	0	67.33
	YOLO 11n	27	101	56	18	0	55.45
抽穗期	AHF－YOLO 11	38	1 699	1 622	89	1	94.41
	YOLO 11n	38	1 699	1 480	177	10	87.11
灌浆期	AHF－YOLO 11	33	5 541	4 619	45	20	83.40
	YOLO 11n	33	5 541	4 356	132	38	78.62
成熟期	AHF－YOLO 11	30	5 220	3 885	96	46	74.43
	YOLO 11n	30	5 220	3 661	243	98	70.13

（2）AHF－YOLO 11 在抽穗期的稻穗识别准确率最高，达到 94.41%。主要源于抽穗期稻穗已完全从茎秆中抽出，形态特征比孕穗期更为明显，结构轮廓清晰，易于模型捕捉关键特征。同时，稻穗在该阶段呈现纺锤形紧凑排列，相比灌浆期和成熟期的弓形松散状态，目标形态更加规则，从而识别准确率更高。此外，抽穗期稻穗与背景的对比度较高，干扰因素较少，而灌浆期和成熟期稻穗及茎叶逐渐变黄，背景干扰增加，进一步降低了识别效果。

2.5　水稻稻穗识别模型在估产中的应用

不同生育期稻穗实测数据表明，稻穗数量在齐穗

期以后(即抽穗期末期,80% 的植株完成抽穗)达到最大,而在灌浆期和成熟期,稻穗数量基本保持不变,如图 8(齐穗期即抽穗末期,乳熟期即灌浆初期,蜡熟期即灌浆后期)所示。改进模型 AHF - YOLO 11 在抽穗期识别准确率最高,同时考虑稻穗数量变化趋势,本研究认为在抽穗末期—灌浆初期(乳熟期)期间开展无人机遥感的估产试验,可以取得最好效果。为了验证上述判断,本研究按图 9 所示流程设计了田间试验,以评价和验证 AHF - YOLO 11 在水田稻穗估产中的实用价值。首先,在抽穗末期—灌浆初期

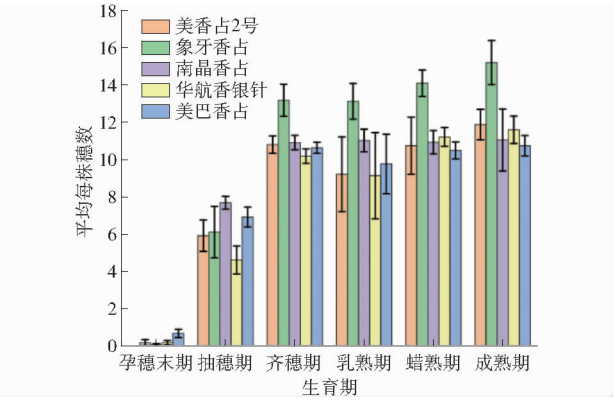


图 8 不同生育期水稻稻穗实测结果

Fig. 8 Measured rice panicle data across different growth stages

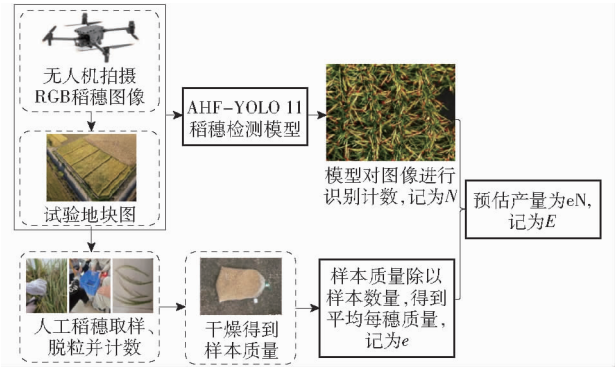


图 9 估产流程图

Fig. 9 Workflow of rice yield estimation

期间,随机选取 5 个不同品种水稻的种植小区,使用无人机对 5 个种植区域进行拍照后,在无人机遥感图像中随机选择面积为 1.2 m × 0.9 m 的区域,利用 AHF - YOLO 11 模型检测每个区域的穗数,检测结果除以区域面积后得到穗数 N(单位为穗/hm^2)。收获前,每个种植小区随机取 3 株水稻,取样计数并干燥至恒质量后,得到平均穗质量 e,按 $E = eN$ 估算该品种产量 E(单位为 kg/hm^2)。为了验证估产结果的准确性,在同一种植小区,人工收割尺寸为 1.2 m × 0.9 m 区域的所有稻穗,经脱粒、干燥至恒质量后称量并计算实际该品种的实际产量 A(单位为 kg/hm^2),最后计算估产误差,相应结果如表 5 所示。

表 5 5 个品种水稻估产结果

Tab. 5 Estimated yield of five rice varieties

区域	面积/(m × m)	水稻品种	识别穗数/(穗·hm^{-2})	平均穗质量/g	真实产量/(kg·hm^{-2})	预估产量/(kg·hm^{-2})	误差/%
1	1.2 × 0.9	美香占 2 号	1.444×10^6	4.35	6 972	6 283	9.94
2	1.2 × 0.9	象牙香占	1.472×10^6	3.54	5 574	5 209	6.56
3	1.2 × 0.9	南晶香占	1.407×10^6	4.68	6 960	6 583	5.43
4	1.2 × 0.9	华航香银针	1.657×10^6	4.54	7 889	7 522	4.66
5	1.2 × 0.9	美巴香占	1.231×10^6	4.99	7 194	6 142	14.64

表 5 的估产结果表明,利用 AHF - YOLO 11 稻穗检测模型进行估产的试验取得了较好的效果。所有估产验证区的误差小于 14.64%,其中华航香银针估产精度最高,误差仅为 4.66%。不同品种稻穗的估产误差存在差异,主要原因在于各品种稻穗自身的形态特征和空间分布特征(表型细长、交叉重叠程度)对模型性能的影响。如图 10(图中白色框区域为漏检区域放大示意图)所示,华航香银针穗型直立,穗长适中且分枝角度小,束状分布,交叉重叠面积小,漏检率低,因而估产精度最高。美香占 2 号、南晶香占和象牙香占,穗长中等、弯曲度小,模型仍能较好地区分相邻稻穗,误差保持在 10% 以内,处于可接受范围。美巴香占穗型细长、分布密度高,稻穗间形成大面积交叉重叠,出现漏穗情况较多,最终估产误差最大,达到 14.64%。

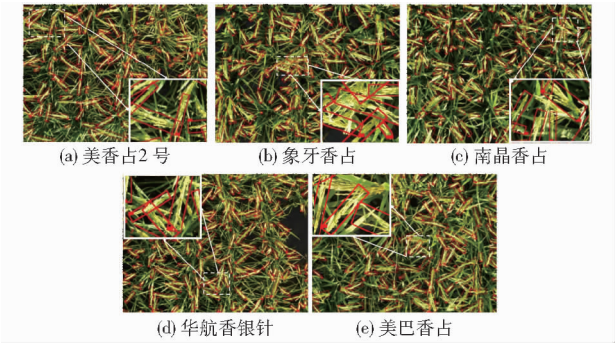


图 10 5 个品种水稻估产与误差分析图

Fig. 10 Yield estimation and error analysis of five rice varieties

3 讨论

尽管改进模型 AHF - YOLO 11 在稻穗检测和估产中取得较高准确率,但叶-稻穗交叉和相互遮挡

等复杂环境仍会导致稻穗漏检，影响检测精度，进而成为产量估算的系统误差源，同时模型训练与验证仅覆盖本研究所提到的 5 个水稻品种，跨品种泛化能力有待进一步验证[36]。在产量预估精度方面，受田间工作量的限制，本研究每个品种仅设置 1 个 1.08 m^2 测产区，样本量不足，同时当前以“穗数与单穗质量乘积”简易模型估产，未纳入千粒质量、结实率等关键产量构成因子，限制了估产模型的严谨性。在未来研究中，将开发具有更高精度和鲁棒性的深度学习模型来识别稻穗，并使这些模型能够在复杂的自然场景中得到实际应用，或者降低跨生育期稻穗形态差异带来的影响。在产量预估方面，可以通过构建稻穗实测总数和无人机遥感识别稻穗数量之间关系，进而建立考虑千粒质量、结实率等关键因素的机器学习回归估产模型。

4　结论

（1）针对目前未充分考虑到稻穗复杂的生长环境以及稻穗跨生育期形态变化对水稻稻穗识别的影响，本文提出了一种基于 YOLO 11 的水稻稻穗识别算法 AHF－YOLO 11。

（2）改进模型的准确率达到 90.7%，相较于原有的 YOLO 11n 模型，提高了 4.3 个百分点，与其他主流旋转目标识别算法相比，分别提高了 4.4 ~ 39 个百分点。

（3）不同生育期测试下，AHF－YOLO 11 抽穗期准确率最高达到 94%，相较于原 YOLO 11 模型在孕穗期、抽穗期、灌浆期和成熟期准确率分别提高 11.88、7.3、4.78、4.3 个百分点，进一步分析表明抽穗期末期–灌浆期初期期间是利用无人机遥感识别稻穗的最佳时间。

（4）针对华航香银针等 5 个品种的估产试验表明，基于 AHF－YOLO 11 模型的估产方法误差在 4.66% ~ 14.64% 范围内，验证了模型在估产中的可行性。本研究为田间水稻穗数的智能计数提供了一种方便、高效、可行的方法，对智慧农业中的水稻检测和估产具有重要意义。

参　考　文　献

[1] QIU F, SHEN X, ZHOU C, et al. Rice ears detection method based on multi-scale image recognition and attention mechanism [J]. IEEE Access, 2024, 12: 68637－68647.

[2] XU C, JIANG H, YUEN P, et al. MHW－PD: a robust rice panicles counting algorithm based on deep learning and multi-scale hybrid window[J]. Computers and Electronics in Agriculture, 2020, 173: 105375.

[3] 翟肇裕，张梓涵，徐焕良，等. YOLO 算法在动植物表型研究中应用综述[J]. 农业机械学报，2024，55(11)：1－20.
ZHAI Zhaoyu, ZHANG Zihan, XU Huanliang, et al. Review of applying YOLO family algorithms to analyze animal and plant phenotype[J]. Transactions of the Chinese Society for Agricultural Machinery, 2024, 55(11): 1－20. (in Chinese)

[4] ZHOU Q, GUO W, CHEN N, et al. Analyzing nitrogen effects on rice panicle development by panicle detection and time-series tracking[J]. Plant Phenomics, 2023, 5: 0048.

[5] WANG X, YANG W, LV Q, et al. Field rice panicle detection and counting based on deep learning[J]. Frontiers in Plant Science, 2022, 13: 966495.

[6] CHEN Y, XIN R, JIANG H, et al. Refined feature fusion for in-field high-density and multi-scale rice panicle counting in UAV images[J]. Computers and Electronics in Agriculture, 2023, 211: 108032.

[7] SANAEIFAR A, GUINDO M L, BAKHSHIPOUR A, et al. Advancing precision agriculture: the potential of deep learning for cereal plant head detection[J]. Computers and Electronics in Agriculture, 2023, 209: 107875.

[8] 王泰华，郭亚州，张家乐，等. 基于改进 YOLO v5s 的水稻害虫识别研究[J]. 农业机械学报，2024，55(11)：39－48.
WANG Taihua, GUO Yazhou, ZHANG Jiale, et al. Rice pest identification based on improved YOLO v5s[J]. Transactions of the Chinese Society for Agricultural Machinery, 2024, 55(11): 39－48. (in Chinese)

[9] 廖娟，刘凯旋，杨玉青，等. 基于 RDN－YOLO 的自然环境下水稻病害识别模型研究[J]. 农业机械学报，2024，55(8)：233－242.
LIAO Juan, LIU Kaixuan, YANG Yuqing, et al. Rice disease recognition in natural environment based on RDN－YOLO[J]. Transactions of the Chinese Society for Agricultural Machinery, 2024, 55(8): 233－242. (in Chinese)

[10] 李金瑞，杜建军，张宏鸣，等. 基于轻量化 MLCE－RTMDet 的人工去雄后玉米雄穗检测算法[J]. 农业机械学报，2024，55(11)：184－192，503.
LI Jinrui, DU Jianjun, ZHANG Hongming, et al. Maize tassel detection algorithm after artificial emasculation based on lightweight MLCE－RTMDet[J]. Transactions of the Chinese Society for Agricultural Machinery, 2024, 55(11): 184－192, 503. (in Chinese)

[11] 马中杰，罗晨，骆巍，等. 基于改进 YOLO v7－tiny 的玉米种质资源雄穗检测方法[J]. 农业机械学报，2024，55(7)：290－297.
MA Zhongjie, LUO Chen, LUO Wei, et al. Tassel detection method of maize germplasm resources based on improved YOLO v7－tiny[J]. Transactions of the Chinese Society for Agricultural Machinery, 2024, 55(7): 290－297. (in Chinese)

[12] 张远琴，肖德琴，陈焕坤，等. 基于改进 Faster R－CNN 的水稻稻穗检测方法[J]. 农业机械学报，2021，52(8)：231－240.
ZHANG Yuanqin, XIAO Deqin, CHEN Huankun, et al. Rice panicle detection method based on improved Faster R－CNN

[J]. Transactions of the Chinese Society for Agricultural Machinery,2021,52(8):231 – 240. (in Chinese)
[13] SUN B, ZHOU W, ZHU S, et al. Universal detection of curved rice panicles in complex environments using aerial images and improved YOLOv4 model[J]. Frontiers in Plant Science, 2022, 13: 1021398.
[14] TAN S, LU H, YU J, et al. In-field rice panicles detection and growth stages recognition based on RiceRes2Net[J]. Computers and Electronics in Agriculture, 2023, 206: 107704.
[15] LAN M, LIU C, ZHENG H, et al. Rice-YOLO: in-field rice spike detection based on improved YOLO v5 and drone images [J]. Agronomy, 2024, 14(4): 836.
[16] WEI J, TIAN X, REN W, et al. A precise plot-level rice yield prediction method based on panicle detection[J]. Agronomy, 2024, 14(8): 1618.
[17] LIANG Y, LI H, WU H, et al. A rotated rice spike detection model and a crop yield estimation application based on UAV images[J]. Computers and Electronics in Agriculture, 2024, 224: 109188.
[18] LU X, SHEN Y, XIE J, et al. Phenotyping of panicle number and shape in rice breeding materials based on unmanned aerial vehicle imagery[J]. Plant Phenomics, 2024, 6: 0265.
[19] 王泽,周钦阳,刘聪,等. 基于无人机和地面图像的田间水稻冠层参数估测与评价[J]. 作物学报,2022,48(5):1248 – 1261.
WANG Ze, ZHOU Qinyang, LIU Cong, et al. Estimation and evaluation of paddy rice canopy characteristics based on images from UAV and ground camera[J]. Acta Agronomica Sinica, 2022, 48(5): 1248 – 1261. (in Chinese)
[20] 张冬冬, 王春平, 付强. 基于 anchor-free 的光学遥感舰船关重部位检测算法[J]. 北京航空航天大学学报, 2022, 50(4): 1365 – 1374.
ZHANG Dongdong, WANG Chunping, FU Qiang. Ship's critical part detection algorithm based on anchor-free in optical remote sensing[J]. Journal of Beijing University of Aeronautics and Astronautics, 2024, 50(4): 1365 – 1374. (in Chinese)
[21] LI C, ZHANG Y, QU Y. Object detection based on deep learning of small samples[C]//2018 Tenth International Conference on Advanced Computational Intelligence (ICACI). IEEE, 2018: 449 – 454.
[22] 李茂,肖洋轶,宗望远. 基于改进 YOLO 11 的多尺度板栗果实识别方法[J]. 农业机械学报,2025,56(5):443 – 454.
LI Mao, XIAO Yangyi, ZONG Wangyuan. Multi-scale chestnut detection method based on improved YOLO 11 model[J]. Transactions of the Chinese Society for Agricultural Machinery,2025,56(5):443 – 454. (in Chinese)
[23] DAI W, LIU R, WU Z, et al. Exploiting scale-variant attention for segmenting small medical objects[J]. arXiv Preprint, arXiv:2407.07720, 2024.
[24] CHEN Y, ZHANG C, CHEN B, et al. Accurate leukocyte detection based on deformable-DETR and multi-level feature fusion for aiding diagnosis of blood diseases[J]. Computers in Biology and Medicine, 2024, 170: 107917.
[25] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need[J]. Advances in Neural Information Processing Systems, 2017, 30:5998 – 6008.
[26] 包强强,唐思源,李擎乾,等. 改进 YOLOv8 模型实现多类型肺结节检测[J]. 计算机科学与探索,2025,19(2):429 – 442.
BAO Qiangqiang, TANG Siyuan, LI Qingqian, et al. Improved YOLOv8 model for multi-type pulmonary nodule detection[J]. Journal of Computer Science and Exploration, 2025, 19(2): 429 – 442. (in Chinese)
[27] YANG M D, BOUBIN J G, TSAI H P, et al. Adaptive autonomous UAV scouting for rice lodging assessment using edge computing with deep learning EDANet[J]. Computers and Electronics in Agriculture, 2020, 179: 105817.
[28] CHATTOPADHAY A, SARKAR A, HOWLADER P, et al. Grad-cam + +: generalized gradient-based visual explanations for deep convolutional networks[C]//2018 IEEE Winter Conference on Applications of Computer Vision (WACV). IEEE, 2018: 839 – 847.
[29] DRAELOS R L, CARIN L. Use HiResCAM instead of Grad-CAM for faithful explanations of convolutional neural networks [J]. arXiv Preprint, arXiv:2011.08891, 2020.
[30] FENG S, HUANG Y, ZHANG N. An improved YOLOv8 obb model for ship detection through stable diffusion data augmentation[J]. Sensors, 2024, 24(17): 5850.
[31] TIAN Y, YE Q, DOERMANN D. YOLO v12: attention-centric real-time object detectors[J]. arXiv Preprint, arXiv:2502.12524, 2025.
[32] XIE X, CHENG G, WANG J, et al. Oriented R – CNN for object detection[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision, 2021: 3520 – 3529.
[33] YANG S, PEI Z, ZHOU F, et al. Rotated Faster R – CNN for oriented object detection in aerial images[C]//Proceedings of the 2020 3rd International Conference on Robot Systems and Applications, 2020: 35 – 39.
[34] 岳冰莹,陈亮,师皓,等. 基于改进 RetinaNet 的 SAR 图像目标检测方法[J]. 信号处理,2022,38(1):128 – 136.
YUE Bingying, CHEN Liang, SHI Hao, et al. Target detection method for SAR images based on improved RetinaNet[J]. Journal of Signal Processing, 2022, 38(1): 128 – 136. (in Chinese)
[35] 刘敏豪,王堃,金睿蛟,等. 基于改进 RoI Transformer 的遥感图像多尺度旋转目标检测[J]. 应用光学,2023,44(5):1010 – 1021.
LIU Minghao, WANG Kun, JIN Ruijiao, et al. Multi-scale rotated object detection in remote-sensing images based on improved RoI Transformer[J]. Journal of Applied Optics, 2023, 44(5): 1010 – 1021. (in Chinese)
[36] YU F, WANG M, XIAO J, et al. Advancements in utilizing image-analysis technology for crop-yield estimation[J]. Remote Sensing, 2024, 16(6): 1003.