doi:10.6041/j.issn.1000-1298.2025.11.048

# 基于 YOLO v8 - TRP 模型的小麦麦穗识别方法

苑迎春[1,2]　耿　俊[1]　许　楠[1,2]　何振学[1,3]　王克俭[1,3]

(1. 河北农业大学信息科学与技术学院, 保定 071001; 2. 农业遥感应用河北省工程研究中心, 保定 071001;
3. 河北省农业大数据重点实验室, 保定 071001)

**摘要**: 针对麦穗个体在图像中显著性不足导致的目标特征提取不充分及识别准确率低的问题,提出一种基于 YOLO v8 - TRP 的小麦麦穗识别方法。该方法在骨干网络前设计三注意力融合(Triple attention fusion,TAF)模块,通过结合多头自注意力(MHSA)、像素注意力(PA)和通道注意力(CA),在保留麦穗细节信息不被破坏的前提下,充分提高了麦穗目标与其他复杂背景的对比度。并在骨干网络中引入感受野注意力卷积(RFAConv)和 CBAM 注意力机制改进卷积模块,使模型在提取特征时,能在感受野范围内为麦穗目标加权,有效提升了模型对麦穗目标的特征提取能力和查全率。同时在模型颈部和头部增加 P2 小目标层,提高模型的小目标检测能力和查准率。实验结果表明,在 Global Wheat Head Detection 2021(GWHD2021)公开数据集上,改进模型的平均精度均值达到 92.5%,相较于原模型提高 2.9 个百分点。改进模型在公共数据集 WEDD、Spike、RGWHD、ACID 和自建数据集 HBAUW(Hebei Agricultural University Wheat)上的平均精度均值较原模型平均提升 4.6 个百分点,充分说明该模型能够有效提升复杂背景下小麦麦穗的检测精度。

**关键词**: 麦穗识别; 公开数据集; YOLO v8n; RFAConv; CBAM

**中图分类号**: TP391　　**文献标识码**: A　　**文章编号**: 1000-1298(2025)11-0499-10　　**OSID**:

# Wheat Ear Recognition Method Using YOLO v8 - TRP Model

YUAN Yingchun[1,2]　GENG Jun[1]　XU Nan[1,2]　HE Zhenxue[1,3]　WANG Kejian[1,3]

(1. *College of Information Science and Technology*, *Hebei Agricultural University*, *Baoding* 071001, *China*
2. *Hebei Engineering Research Center for Agricultural Remote Sensing Application*, *Baoding* 071001, *China*
3. *Hebei Key Laboratory of Agricultural Big Data*, *Baoding* 071001, *China*)

**Abstract**: Aiming to address the issue of insufficient target feature extraction and low recognition accuracy due to the low saliency of individual wheat ears in images, an improved YOLO v8 - TRP wheat ear recognition method based on YOLO v8n was proposed. A triple attention fusion (TAF) module before the backbone network was designed, which combined multi-head self-attention (MHSA), pixel attention (PA), and channel attention (CA) to fully enhance the contrast between the wheat ear target and the complex background without destroying the detailed information of the wheat ear. Additionally, the receptive field attention convolution (RFAConv) and CBAM attention mechanism were introduced to improve the convolution module in the backbone network, enabling the model to weight the wheat ear target within the receptive field when extracting features, effectively enhancing the model's feature extraction ability and recall rate for the wheat ear target. Moreover, a P2 small target layer was added to the neck and head of the model to improve the model's ability to detect small targets and the precision rate. Experimental results showed that on the Global Wheat Head Detection 2021 (GWHD2021) public dataset, the mean average precision of the improved model reached 92.5%, an increase of 2.9 percentage points compared with that of the original model. The mean average precision of the improved model on the public datasets WEDD, Spike, RGWHD, ACID, and the self-built dataset HBAUW (Hebei Agricultural University Wheat) was on average 4.6 percentage points higher than that of the original model, fully demonstrating that this model can effectively improve the detection accuracy of wheat ears in complex backgrounds.

**Key words**: wheat ear recognition; public datasets; YOLO v8n; RFAConv; CBAM

收稿日期: 2025 - 05 - 12　修回日期: 2025 - 06 - 19

**基金项目**: 国家自然科学基金项目(62102130)和河北省重点研发项目(21307405D)

**作者简介**: 苑迎春(1970—),女,教授,博士生导师,主要从事智能信息处理与大数据技术研究,E-mail: nd_hd_yyc@163.com

## 0 引言

小麦是我国的主要粮食作物之一,在国家粮食安全体系构建中发挥着重要作用[1],及时可靠的产量预测是保障粮食市场稳定、优化农业资源配置的核心环节,而精准高效的小麦麦穗检测可以显著提升产量预估的效率。因此,如何准确检测出自然场景下小麦麦穗成为预测小麦产量的重中之重。目前,基于机器学习[2-3]的小麦麦穗识别方法依赖于麦穗颜色、纹理、光谱等特征,受光照影响较大,难以满足计数需求。

随着计算机视觉技术的发展,基于深度学习[4-5]的麦穗检测模型逐渐成为研究热点,许多学者在自建的数据集上开展基于深度学习的小麦麦穗计数方法研究[6-7],获得了较高的麦穗检测准确率,但实验性能仅在自建数据集上进行了验证,具有一定局限性。

深度学习算法的性能与训练数据样本的丰富性和多样性具有很强的相关性。为验证算法的有效性,近年来,在小麦麦穗识别领域构建了GWHD2020[8]、GWHD2021[9]等代表性公共数据集,开展了小麦麦穗识别算法研究。目前学者对GWHD2020 数据集的研究相对成熟,准确率也达到了一定的高度[10-13]。

GWHD2021 数据集在 GWHD2020 数据集的基础上,从数据集规模、麦穗多样性、标签可靠性、图像质量性等多个方面进行扩展,因此 GWHD2021 数据集的检测难度大幅提升。现有研究[14-15]对GWHD2021 数据集的研究虽取得了一定进展,但均未从根本上解决由麦穗目标小、多、散和麦穗目标与背景对比度低导致的目标主体显著性不足的问题,进而引起的大量分散小目标的特征提取能力与小目标检测能力矛盾问题。

因此,本文以 GWHD2021 数据集为研究对象,针对上述问题,提出一种改进 YOLO v8n 的小麦麦穗识别方法。通过设计三注意力融合(TAF)模块,在小麦图像输入到骨干网络前,提高麦穗目标个体在图像中的显著性。并在骨干网络中引入基于 CBAM 的感受野注意力卷积(RFCBAMConv),以提升大量分散麦穗小目标的特征提取能力,保证查全率,解决漏检问题。同时在颈部和头部增加 P2 小目标检测层,补充深层网络丢失的细节信息,以提高小目标检测能力,保证查准率,解决误检问题。进而从根本上解决漏检和误检问题,同时提升查全率和查准率。

## 1 实验数据

### 1.1 2021 全球小麦穗数据集(GWHD2021)

小麦麦穗的图像数据是 2021 年全球小麦挑战赛计算机视觉国际会议提供的公共数据集,包含几十种不同的种类以及 3 个不同时期,其中成熟期图像 1 869 幅,灌浆期图像 3 724 幅,开花期图像 922 幅,如表 1 所示。

**表 1 数据集的分布**

**Tab. 1 Distribution of dataset**

| 生长阶段 | 国家 | 图像数量/幅 | 麦穗数量 | 来源 |
|---|---|---|---|---|
| 灌浆期 | 瑞士 | 747 | 49 603 | 苏黎世联邦理工学院 |
| 灌浆期-成熟期 | 英国 | 432 | 19 210 | 洛桑研究院 |
| 成熟期 | 比利时 | 30 | 1 847 | 列日大学/让布鲁克斯农业大学 |
| 灌浆期-成熟期 | 挪威 | 180 | 12 556 | 挪威生命科学大学 |
| 开花期-灌浆期-成熟期 | 法国 | 2 268 | 80 474 | 法国农业科学研究院、法国国家农业食品与环境研究院 |
| 灌浆期-成熟期 | 加拿大 | 200 | 5 985 | 萨斯喀彻温省大学 |
| 开花期-灌浆期-成熟期 | 美国 | 605 | 24 873 | 堪萨斯州立大学、TERRA - REF 项目 |
| 开花期 | 墨西哥 | 206 | 7 175 | 国际玉米小麦改良中心 |
| 开花期-成熟期 | 日本 | 1 174 | 32 950 | 东京大学、京都大学 |
| 开花期-灌浆期 | 中国 | 220 | 10 754 | 南京农业大学 |
| 开花期-灌浆期-成熟期 | 澳大利亚 | 423 | 28 872 | 昆士兰大学 |
| 灌浆期 | 苏丹 | 30 | 888 | 苏丹农业研究委员会 |
| 总计 | | 6 515 | 275 187 | |

根据文献[13]的经验,将数据集划分为训练、验证和测试集。其中训练集包含 3 657 幅图像,测试集包含 1 382 幅图像,验证集包含 1 476 幅图像,图像分辨率为 1 024 像素 ×1 024 像素,涵盖小麦开花期、灌浆期、成熟期等多个发育阶段。如图 1 所示。

### 1.2 数据采集

HBAUW 图像数据于河北省保定市清苑区石桥乡黄陀村河北农业大学示范田基地(38°46′24.90″N,115°32′33.23″E)采集,采集时间为 2024 年 5 月 22 日和 2024 年 5 月 29 日的 09:00—11:00,天气晴朗,温

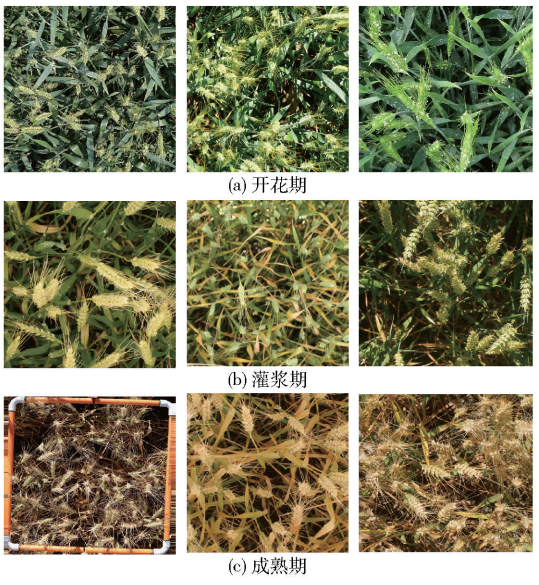



图 1　小麦各生长周期图像示例

Fig. 1　Example images of each growth cycle of wheat

度在 20～30℃，小麦分别在这 2 个时间段处于灌浆期和成熟期，麦穗分明，利于采集。拍摄设备为小米 14 型智能手机和尼康 D7500 型单反数码相机，分辨率分别为3 072 像素×3 072 像素和 2 784 像素×1 856 像素，图像格式为 JPG，拍摄角度垂直向下以及侧上方，拍摄高度在麦穗上方 0.2～0.5 m，采集小麦灌浆期和成熟期图像共 713 幅，其中灌浆期图像 435 幅，成熟期图像 278 幅。部分数据集样本如图 2 所示。

### 1.3　数据集制作

本文使用 LabelImg 工具对采集的田间小麦麦穗进行标注，在标注过程中，确保每个小麦麦穗都完

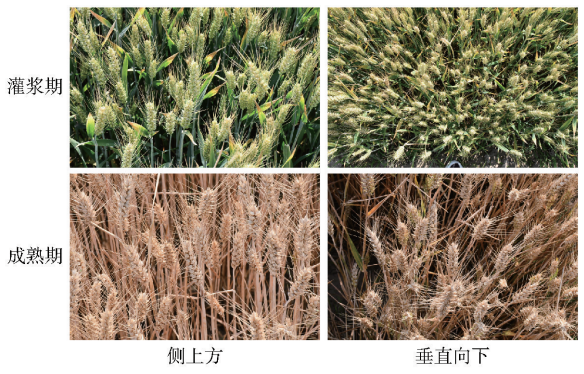



图 2　采集的小麦麦穗样本图

Fig. 2　Map of wheat ears sampled from collection

全封闭在外接矩形边界框内，选择适当的外接矩形边界框，尽量减少框内多余的背景信息，并将图像中的小麦麦穗定义为 0，在标注完成后，就会生成一个 TXT 文件，其中包含标签名称、图像位置等信息。

## 2　模型改进

由于小麦图像中的麦穗目标小、多、散以及麦穗目标与背景对比度低，导致麦穗目标主体显著性不足，因此会出现大量分散小目标的特征提取不充分以及小目标检测准确率低的问题。针对该问题本文提出一种改进 YOLO v8n 的小麦麦穗识别方法，该方法以 YOLO v8n 模型为基础，首先在骨干网络前设计三注意力融合 TAF 模块，其次引入感受野注意力卷积 RFCBAMConv 替换骨干网络中的传统普通卷积，最后在模型颈部和头部添加 P2 小目标检测层，改进后的网络结构如图 3 所示。

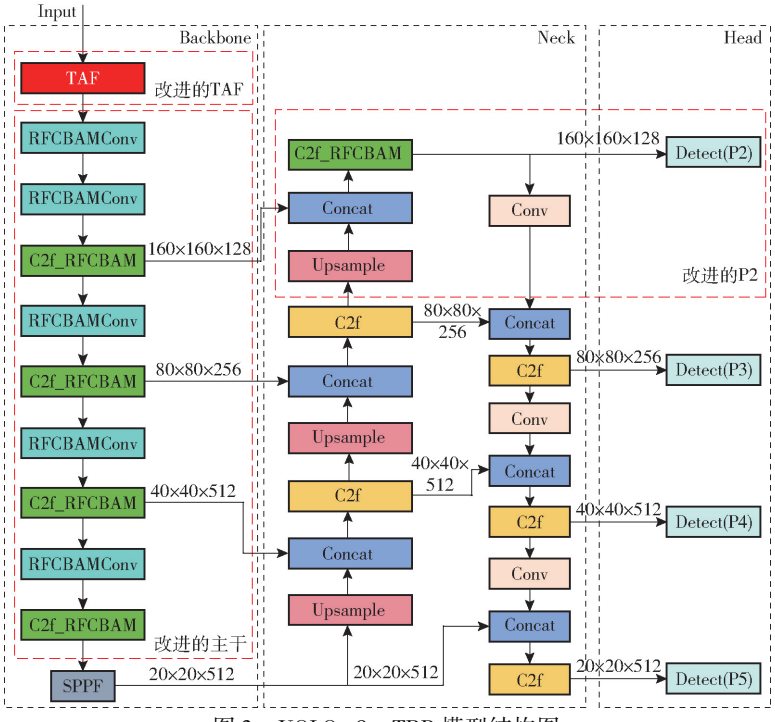



图 3　YOLO v8－TRP 模型结构图

Fig. 3　YOLO v8－TRP model structure diagram

### 2.1 TAF 模块

由于小麦公共数据集中的图像普遍存在背景复杂的问题,再加之小麦麦穗和小麦秸秆在各生长周期的颜色也几乎相同,导致麦穗目标与图像中其他干扰背景差异小、对比度低,目标主体不明显。因此,本文设计了一种基于多头自注意力 MHSA、像素注意力 PA 和通道注意力 CA 的三重注意力融合 TAF 模块。该模块由预处理卷积层、组结构(G1 ~ G4)层、双注意力层(Dual - Attention,DA)和后处理卷积层共 4 部分构成,其结构如图 4 所示。

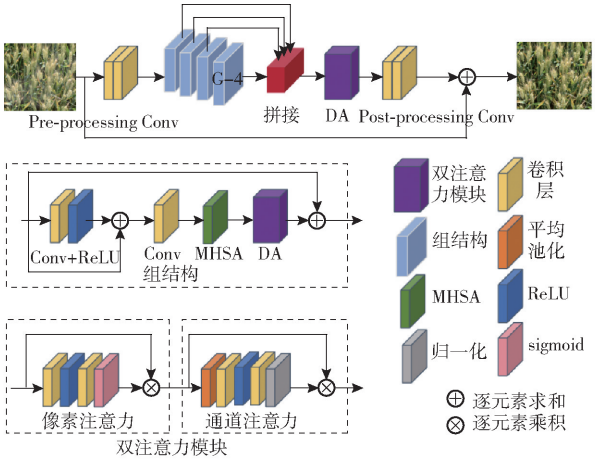



图 4 TAF - Net 网络架构

Fig. 4 TAF - Net network architecture

首先经过 2 个 3 × 3 的预处理卷积层进行浅层特征提取,其次将提取到的特征输入到 4 个组结构中给麦穗目标加权,随后将 4 个组结构的输出特征图进行拼接后经过双注意力融合模块(DA)进行降维下采样融合与二次加权,最后通过 2 个 3 × 3 的后处理卷积层对融合后的特征图进行重构。

(1)组结构

由于 TAF 的位置处于网络最前端,主要目的是对图像中的麦穗目标特征进行加强,因此需保证麦穗目标的特征信息不被破坏和丢失,以供后续网络进行提取识别。因此设计组结构层,组结构层由 4 个组结构模块 G1 ~ G4 组成,每个组结构模块由 3 × 3 卷积层、MHSA 模块和 DA 模块构成,其结构如图 4 所示。其中,卷积层主要用于特征提取。多头自注意力层通过捕获特征间的长距离依赖关系,尽可能保留小麦目标的浅层关键细节特征信息。MHSA 模块结构如图 5 所示。

然后通过 DA 双注意力对 MHSA 输出的特征图在像素和通道维度上分别进行权重分配,进而聚焦小麦目标信息。并通过 shortcut 捷径连接集成局部残差学习,允许通过多个局部残差连接绕过不太重要的干扰背景信息,如颜色有明显区别的根茎部或

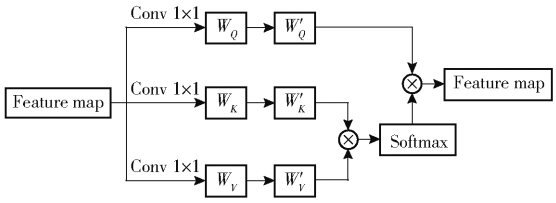



图 5 多头自注意力机制

Fig. 5 Multi-head self-attention mechanism

深黑色的背景边缘信息,增强目标与干扰背景信息的区分度。

(2)双注意力 DA

双注意力 DA 模块由通道注意力模块 CA 和像素注意力模块 PA 串联组成,其通过结合通道注意力与像素注意力,可以灵活区分出不同质量和场景图像中的麦穗目标与背景,并对图像中的不同区域进行加权处理,提高小麦目标区域权重,降低其他干扰背景区域权重,其结构如图 4 所示。其表达式为

$$DA = \{CA(PA(F))\} \tag{1}$$

式中 DA——双注意力模块

CA——通道注意力模块

PA——像素注意力模块

$F$——初始特征图

像素注意力机制负责为每个像素特征分配权重,表达式为

$$PA = \delta(Conv(ReLU(Conv(F)))) \tag{2}$$

式中 $\delta$——sigmoid 激活函数

ReLU——ReLU 激活函数

Conv——卷积操作

通道注意力机制负责为每个通道特征分配权重,表达式为

$$CA = \delta(Conv(ReLU(Conv(g)))) \tag{3}$$

式中 $g$——具有像素注意力的特征图

在 TAF 中,组结构中的 DA 为 4 个不同层次尺度的特征图分配权重,组结构后的 DA 对 4 个组结构拼接融合后输出的特征图进行二次加权,并通过 CA 中 1 × 1 的卷积层进行降维下采样,使特征图尺寸恢复一致。最后再通过后处理卷积层进行重构后引入全局残差学习的思想,使保留的浅层细节信息能够传递到深层网络,实现在保证麦穗目标纹理等细节特征不被破坏的前提下充分提高麦穗目标与其他复杂背景的区分度。

### 2.2 RFCBAMConv 模块

由于麦穗目标多而分散,在图像中分布不集中,目标主体不明显,导致模型在提取特征时无法强调感受野中每个特征的重要性,限制了特征提取能力和表达能力。为进一步从经过 TAF 模块处理后得

到的特征图中提取目标特征，本文引入了一种由感受野注意力卷积（Receptive-field attention convolutional operation，RFAConv）[16] 和 CBAM[17]（Concentration-based attention module）注意力机制改进的卷积块 RFCBAMConv，RFCBAMConv 与 CBAM 结构对比如图 6 所示。

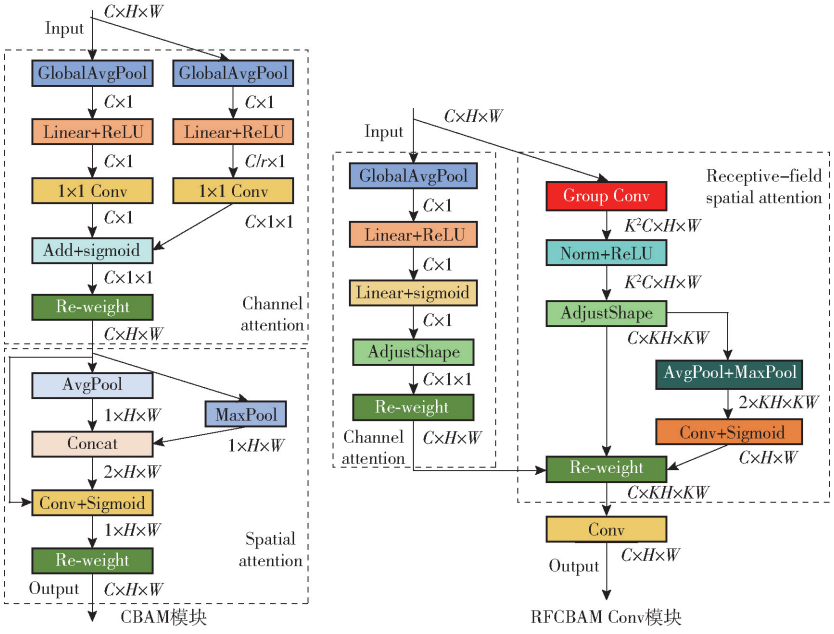



图 6　RFCBAMConv 与 CBAM 结构图

Fig. 6　RFCBAMConv and CBAM structural diagram

RFCBAM 模块主要由通道注意力模块和感受野空间注意力模块组成，通道注意力模块使用 SE 注意力机制代替 CBAM 中的通道注意力，并去除全局最大池化分支，减少计算开销与特征重叠。改用 RFAConv 中的分组卷积结合 CBAM 中的空间注意力提取感受野空间特征作为感受野空间注意力模块。通道注意力和空间注意力不再串联执行，而是并联执行，同时对输入特征图进行加权。

在模型主干网络中，使用 RFCBAMConv 替换传统普通卷积，保持 C2f 中 Bottleneck 的双卷积层基础架构，将第 2 个卷积层替换为 RFCBAMConv 作为 RFCBAM_neck 替换 Bottleneck 形成 C2f_RFCBAM 模块，如图 7 所示，图中 $W$、$H$、$C$ 分别表示特征图的宽度、高度、通道数。其中，RFCBAMConv 已内建有跨层残差信息融合机制，因此 RFCBAM_neck 在结构上省略了传统残差连接操作，同时降低了网络复杂度。

RFCBAMConv 卷积层通过感受野特征的自适应权重分配，重构了传统卷积操作，不仅实现了真正意义上非共享参数的大核卷积操作，还通过全局池化考虑了远程信息，进而解决了使用传统卷积进行特征提取操作时因卷积核参数共享导致的感受野中每个特征的重要性不明显问题，即特征位置不敏感问题。使小麦图像在提取特征时从空间和通道两个维度为每个卷积滑动窗口的感受野范围内的麦穗目标提供权重加持，进而整幅图像的麦穗目标均能得到注意，有效提升了模型的麦穗目标检测效果。

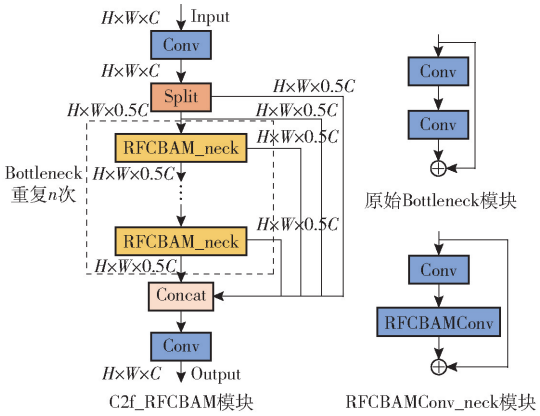



图 7　C2f_RFCBAM 模块与子模块示意图

Fig. 7　Structure of C2f_RFCBAM and sub modules

## 2.3　P2 小目标检测层

由于小麦图像中的麦穗目标普遍较小，且经过骨干网络进行多层特征提取和颈部大倍数的下采样后，较深层次的特征图会丢失部分麦穗小目标的细节信息，导致小目标麦穗识别精确率不高。因此，为了解决深层网络在下采样过程中小目标麦穗细节信息丢失的问题，在 YOLO v8n 模型的颈部和头部引入小目标检测层 P2，使模型能够结合更浅层次的高分辨率细节特征信息，提高对小目标麦穗的检测识别能力。

P2 小目标检测层在保留原模型 C3～C5 层的基

础上,额外提取 C2 阶段输出作为高分辨率特征源,然后将 P3 层上采样 2 倍后与 C2 层进行通道对齐拼接,生成初始 P2 特征图。并通过引入 C2f_RFCBAMConv 优化 P2 层特征,抑制背景噪声的同时保留麦穗小目标高频纹理信息。

### 2.4 评价指标

为评价小麦麦穗检测模型的性能,选取精确率(Precision,P)、召回率(Recall,R)、F1 值(F1 score)、平均精度均值(Mean average precision,mAP)和参数量(Parameters,M)作为模型性能的评价指标。最后为更好地验证模型的有效性和鲁棒性,选用平均精度均值 mAP50 作为评价指标。

## 3 实验结果与分析

本实验的硬件环境为:Gen Intel(R) Core(TM) i7 - 12700F 2.10GHz 中央处理器、32GB 内存、NVIDIA GeForce RTX 3060 12 GB 显存显卡。软件环境为 Pycharm 编辑器、64 位 Windows 10 操作系统,编程语言为 Python 3.9.19,深度学习框架为 Pytorch 2.3.0 + cu118。训练迭代次数为 100 次,初始学习率为 0.01,线程数(Num workers)为 0,同时由于硬件设备资源有限,故将本模型的 batch size 设置为 10。

### 3.1 可视化分析

为了更直观地展示 TAF 模块在各类复杂情境下区分麦穗目标与其他复杂背景方面的性能,本文对麦穗图像使用 Grad CAM 热力图进行可视化展示,红色区域越深,表示该区域对模型最终预测的贡献越大,以观察 TAF 模块对麦穗目标的特征响应能力。如图 8 所示,从全球麦穗数据集中手动筛选出密集和重叠、模糊、强光、弱光等 4 组具有代表性的场景,分别使用原图、YOLO v8n、YOLO v8 + TAF 还原小麦麦穗的检测效果。

由图 8 可知,与 YOLO v8n 相比,YOLO v8 + TAF 模型不仅能精准聚焦于麦穗目标的纹理细节(如芒刺等细微结构),还能有效抑制秸秆、土壤等复杂背景的干扰响应。在原始图像中,由于麦穗与秸秆具有高度相似的颜色特征,YOLO v8n 的热力图显示其对茎秆部分存在显著的误激活现象;而引入 TAF 模块后,MHSA 通过长距离依赖建模有效强化了麦穗的形态特征表达,同时 DA 模块的双重注意力机制在通道维度上突出麦穗的颜色通道特征,在空间维度上抑制干扰区域的像素响应,从而实现了更精准的目标-背景特征解耦。

### 3.2 不同模型识别结果分析

为验证不同模型对小麦麦穗的检测效果以及本文所提出的 YOLO v8 - TRP 模型的性能,将改进模型分别与 YOLO v9 - T[18]、YOLO v7 - Tiny[19]、YOLO v6n[20]、YOLO v5n、Faster R - CNN[21]、SSD[22]等主流的目标检测模型进行比较,使用相同的数据集,对比结果如表 2 所示。

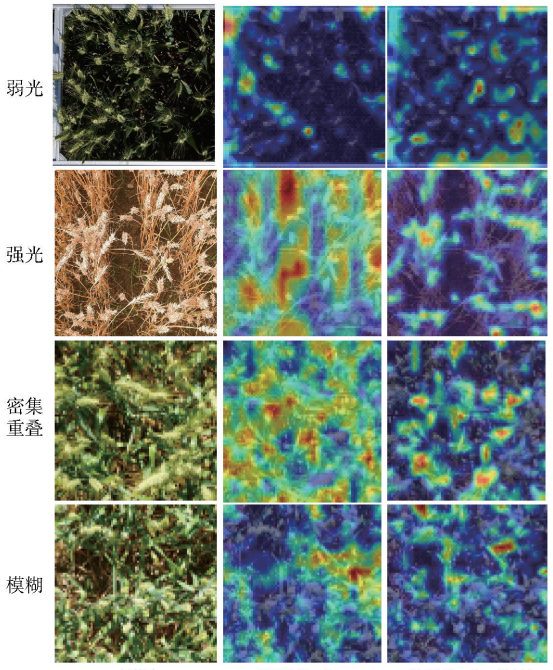



图 8 热力图可视化效果对比

Fig. 8 Comparison of heatmap visualization

由表 2 可知,与 SSD、Faster R - CNN、Centernet、DETR 等传统模型相比,YOLO v8n 具备更高的精确率、召回率、F1 值、平均精度均值以及更少的参数量。与 YOLO v6n、YOLO v9 - T、YOLOX 等 YOLO 系列模型相比,YOLO v8n 具备更高的精确率、召回率、F1 值、平均精度均值以及更少的参数量,与 YOLO v7 - Tiny、YOLO v8s 模型相比,平均精度均值仅低 0.7、0.7 个百分点,但 YOLO v7 - Tiny 模型的参数量近乎是 YOLO v8n 模型的 2 倍,YOLO v8s 模型的参数量近乎是 YOLO v8n 模型的 4 倍。与 YOLO v5n 相比,YOLO v8n 在保持相同精确率和召回率的情况下,实现了更高的 mAP50。尽管 YOLO v8n 的参数量略有增加,但其 mAP50 的提升表明该模型具有更强的泛化能力和鲁棒性,能够更好地适应复杂多变的检测场景。与当前主流模型相比,比 RT - DETR 模型的 mAP50 仅低 0.1 个百分点,比 RT - WEDT 模型的 mAP50 仅低 0.6 个百分点,但 RT - WEDT 模型的参数量近乎是 YOLO v8n 模型的 4 倍,RT - DETR 模型的参数量近乎是 YOLO v8n 模型的 10 倍,因此选用 YOLO v8n 模型作为基础网络更适用于不同种类麦穗的检测任务。从检测效果来看,改进模型 YOLO v8 - TRP 的精确率、召回率、

表 2　不同模型识别结果

Tab. 2　Recognition results of different models

| 类别 | 模型 | 精确率/% | 召回率/% | F1 值/% | 平均精度均值/% | 参数量 |
|---|---|---|---|---|---|---|
| 传统模型 | SSD | 74.9 | 65.7 | 69.9 | 67.3 | $2.25\times10^7$ |
| | Faster R－CNN | 63.8 | 55.7 | 59.5 | 59.3 | $2.59\times10^7$ |
| | Centernet | 88.4 | 79.3 | 83.6 | 85.8 | $3.33\times10^7$ |
| | DETR | 81.1 | 76.7 | 78.8 | 80.1 | $3.68\times10^7$ |
| YOLO 系列 | YOLO v5n | 90.6 | 81.6 | 85.9 | 89.4 | $2.5\times10^6$ |
| | YOLO v6n | 89.8 | 80.7 | 85.0 | 86.1 | $4.7\times10^6$ |
| | YOLO v7－Tiny | 90.1 | 85.1 | 87.5 | 90.3 | $6\times10^6$ |
| | YOLO v8n | 90.6 | 81.6 | 85.9 | 89.6 | $3\times10^6$ |
| | YOLO v8s | 91.2 | 83.0 | 86.9 | 90.3 | $1.11\times10^7$ |
| | YOLO v9－T | 89.8 | 80.7 | 85.0 | 88.9 | $2.6\times10^6$ |
| | YOLOX | 88.2 | 78.7 | 83.1 | 86.3 | $8.9\times10^6$ |
| 当前主流模型 | RT－DETR[15] | | | | 89.7 | $3.2\times10^7$ |
| | RT－WEDT[15] | | | | 90.2 | $1.2\times10^7$ |
| 本文方法 | YOLO v8－TRP | 92.3 | 84.5 | 88.2 | 92.5 | $3.9\times10^6$ |

F1 值和平均精度均值分别为 92.3%、84.5%、88.2%、92.5%，mAP50 相较于传统模型平均提高 19.3 个百分点，相较于 YOLO 系列模型平均提高 3.8 个百分点，相较于当前主流模型平均提高 2.5 个百分点。综上，通过对不同模型的综合分析比较，改进后的 YOLO v8－TRP 模型在检测精度上拥有较大优势。

### 3.3　消融实验

为验证本文所提出改进算法的有效性，以 YOLO v8n 为基准模型，精确率、召回率和平均精度均值作为评估指标，对 3 种不同的改进方法进行系统的消融实验，实验过程中，逐步引入各个模块，以观察它们对模型性能的影响。实验结果见表 3，揭示了各种方法对提升模型性能的显著效果。

表 3　模型消融实验结果

Tab. 3　Results of model ablation experiments　%

| TAF | RFCBAM | P2 | P | R | mAP50 |
|---|---|---|---|---|---|
| × | × | × | 90.6 | 81.6 | 89.6 |
| √ | × | × | 91.4 | 82.7 | 90.9 |
| × | √ | × | 91.6 | 82.9 | 90.6 |
| × | × | √ | 91.4 | 81.8 | 90.4 |
| √ | √ | × | 92.1 | 83.8 | 91.7 |
| √ | × | √ | 91.9 | 83.6 | 91.3 |
| × | √ | √ | 91.8 | 83.4 | 91.6 |
| √ | √ | √ | 92.3 | 84.5 | 92.5 |

注：√表示使用该模块，×表示不使用该模块。

由表 3 可知，单独使用 TAF 模块、RFCBAM 模块、P2 模块对小麦麦穗识别的精确率、召回率、平均精度均值均有明显提升。当仅使用 RFCBAM 模块后，最高召回率为 82.9%，能够有效提升模型的查全率，因此实现了召回率的最大提升。

若使用 TAF、RFCBAM、P2 模块中的任意 2 个，模型的精确率、召回率和平均精度均值也均有提升，其中平均精度均值分别为 91.7%、91.3%、91.6%，相较于原模型分别提高 2.1、1.7、2 个百分点。

若同时使用 3 个模块，模型的精确率、召回率和平均精度均值均为最高，其中，平均精度均值为 92.5%，相较于原模型提升 2.9 个百分点，充分表明 3 个模块能够有效提升模型对小麦麦穗的检测性能。

### 3.4　不同生长时期对比结果分析

为全面验证 YOLO v8－TRP 在小麦麦穗不同生长时期下的检测效果，在 GWHD2021 数据集的测试集中选取灌浆期图像 100 幅，开花期图像 100 幅，成熟期图像 100 幅，分别使用训练好的 YOLO v8n 和 YOLO v8－TRP 模型验证不同时期的图像，具体实验结果如表 4 所示。

由表 4 可知，YOLO v8－TRP 模型在不同生长时期的小麦麦穗检测中平均达到 85.5% 的精确率、67.3% 的召回率、75.9% 的 mAP50。YOLO v8－TRP 模型在小麦麦穗成熟期、灌浆期、开花期的平均精度均值相较于原模型分别提升 8.8、8.6、9.0 个百分点，平均提升 8.8 个百分点。召回率相较于原模型分别提升 8.1、8.0、9.8 个百分点，平均提升 8.6 个百分点。精确率相较于原模型分别提升 5.2、4.5、8.1 个百分点，平均提升 5.9 个百分点。其中，开花期的各项评价指标分别提升 8.1、9.8、9.0 个百分点，在不同生长时期下的提升率均为最高。灌浆期的各项评价指标分别提升 4.5、8.0、8.6 个百分点，在不同生长时期下的提升率均为最低。

同时，由于开花期小麦麦穗较小，且颜色与麦秆相近，形状也不够完整，容易产生误检和漏检，因此

表 4 不同生长时期对比实验结果

Tab. 4 Comparative experimental results at different growth stages

| 生长时期 | 精确率 | | | 召回率 | | | F1 值 | mAP50 | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | YOLO v8n/% | YOLO v8-TRP/% | 提升率/百分点 | YOLO v8n/% | YOLO v8-TRP/% | 提升率/百分点 | YOLO v8-TRP/% | YOLO v8n/% | YOLO v8-TRP/% | 提升率/百分点 |
| 成熟期 | 78.3 | 83.5 | 5.2 | 61.4 | 69.5 | 8.1 | 75.9 | 69.1 | 77.9 | 8.8 |
| 灌浆期 | 83.9 | 88.4 | 4.5 | 65.9 | 73.9 | 8.0 | 80.5 | 74.0 | 82.6 | 8.6 |
| 开花期 | 76.6 | 84.7 | 8.1 | 48.8 | 58.6 | 9.8 | 69.3 | 58.1 | 67.1 | 9.0 |
| 平均值 | 79.6 | 85.5 | 5.9 | 58.7 | 67.3 | 8.6 | 75.2 | 67.1 | 75.9 | 8.8 |

YOLO v8-TRP 模型验证后得到的平均精度均值最低,提升效果最好。而灌浆期小麦麦穗长度达到最大值,麦粒饱满,麦穗颜色特征逐渐由绿转黄,容易识别,因此 YOLO v8-TRP 模型验证后得到的各项评价指标均为最高,提升效果最少。成熟期小麦麦穗分布密集,容易遮挡,与麦秆等背景的区分度不如灌浆期,在同等密集、遮挡的干扰情况下难以实现高精度识别,因此 YOLO v8-TRP 模型验证后得到的各项评价指标均低于灌浆期。综上所述,YOLO v8-TRP 模型能够有效应对小麦麦穗在不同生长时期的特点,在不同场景下均展现出优异的检测性能。

最后,对原图、YOLO v8n、YOLO v8-TRP 在小麦麦穗不同时期下的检测效果进行可视化,并将 conf 设置为 0.5,如图 9 所示。

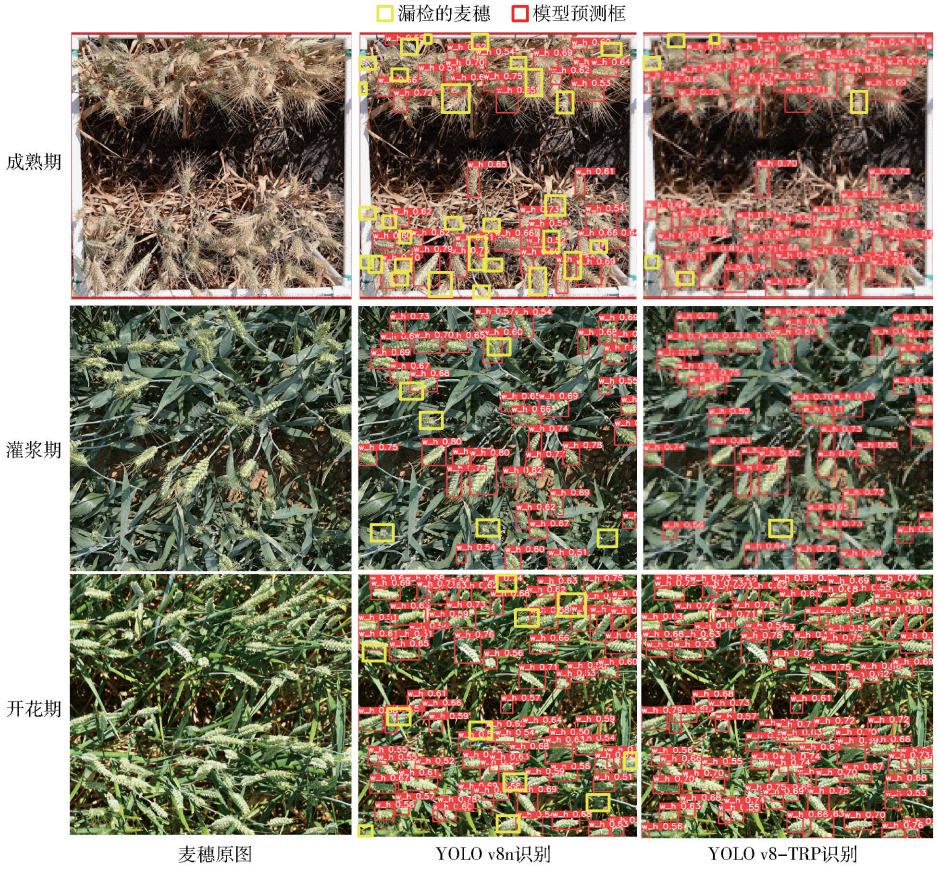



图 9 YOLO v8-TRP 模型可视化示例图

Fig. 9 YOLO v8-TRP model visualization example

由图 9 可知,原始 YOLO v8n 模型能够识别出大部分小麦麦穗,但在一些情况下,检测框的位置不够精确,有些小麦植株的顶部或底部未被完全包含在框内,同时,也存在一些漏检的情况,尤其是在麦穗目标较小且密集的区域。通过对 2 个模型的对比,明显观察到 YOLO v8-TRP 模型生成的红色矩形框更紧密地匹配小麦植株的实际轮廓和尺寸,边界框的定位更加准确。此外,YOLO v8-TRP 模型对小麦麦穗的预测概率普遍高于原模型,少有漏检、误检情况,体现出模型对自身预测结果的高度自信。综上所述,通过可视化结果,可以得出结论,YOLO v8-TRP 模型在小麦植株检测任务中表现出更高的准确性和鲁棒性,以及对复杂环境的适应性。

### 3.5　不同数据集对比结果分析

为验证改进模型的提升能力，本文除了在原有的全球小麦 2021 数据集验证其评价指标外，还通过使用训练好的 YOLO v8n 和改进的 YOLO v8－TRP 模型在公共数据集 Spike[23]、WEDD[24]、ACID[25]、RGWHD[26]和自建数据集 HBAUW 上进行验证实验，其中 Spike 数据集包含开花期、灌浆期、成熟期的 320 幅小麦麦穗图像，WEDD 数据集包含开花期、灌浆期的 236 幅小麦麦穗图像，ACID 数据集包含灌浆期的 534 幅小麦麦穗图像，RGWHD 数据集包含开花期、灌浆期、成熟期的 700 幅小麦麦穗图像，使用相同的实验环境和参数，用于确保模型在不同场景下的可移植性和鲁棒性，具体实验结果如表 5 所示。

表 5　不同数据集的对比实验结果

Tab. 5　Comparative experimental results of different datasets

| 数据集 | 精确率 | | | 召回率 | | | mAP50 | | | mAP50:90 |
|---|---|---|---|---|---|---|---|---|---|---|
| | YOLO v8n/% | YOLO v8－TRP/% | 提升率/百分点 | YOLO v8n/% | YOLO v8－TRP/% | 提升率/百分点 | YOLO v8n/% | YOLO v8－TRP/% | 提升率/百分点 | YOLO v8－TRP/% |
| WEDD | 84.6 | 87.9 | 3.3 | 71.8 | 81.6 | 9.8 | 80.9 | 88.7 | 7.8 | 48.9 |
| Spike | 54.6 | 59.7 | 5.1 | 50.8 | 58.2 | 7.4 | 49.3 | 57.4 | 8.1 | 21.7 |
| RGWHD | 90.9 | 91.3 | 0.4 | 86.4 | 88.1 | 1.7 | 92.3 | 93.0 | 0.7 | 41.8 |
| ACID | 54.7 | 57.9 | 3.2 | 48.7 | 53.1 | 4.4 | 44.7 | 50.2 | 5.5 | 17.6 |
| HBAUW | 71.6 | 73.2 | 1.6 | 72.8 | 73.9 | 1.1 | 70.9 | 71.8 | 0.9 | 24.8 |
| 平均值 | 71.3 | 74.0 | 2.7 | 66.1 | 71.0 | 4.9 | 67.6 | 72.2 | 4.6 | 31.0 |

由表 5 可知，YOLO v8－TRP 模型在不同数据集上的小麦麦穗检测中平均精确率达到 74%、召回率达到 71%、mAP50 达到 72.2%。YOLO v8－TRP 模型在 WEDD、Spike、RGWHD、ACID、HBAUW 数据集上相较于原模型平均精度均值分别提升 7.8、8.1、0.7、5.5、0.9 个百分点，平均提升 4.6 个百分点；召回率分别提升 9.8、7.4、1.7、4.4、1.1 个百分点，平均提升 4.9 个百分点；精确率分别提升 3.3、5.1、0.4、3.2、1.6 个百分点，平均提升 2.7 个百分点。其中，Spike 和 WEDD 数据集的各项评价指标提升率较好，HBAUW 和 RGWHD 数据集的各项评价指标提升率较差。YOLO v8－TRP 模型在不同数据集上的精确率、召回率和平均精度均值均优于原始的 YOLO v8n 模型，充分说明本文所提出的改进 YOLO v8n 算法在小麦麦穗的检测任务中具有更高的准确性和较强的提升能力。

用训练好的模型验证 Spike 数据集和 ACID 数据集得到的平均精度均值较低，这是由于前者取自复杂的大田环境下，多为大视角图像，图像模糊，目标较小且密集，遮挡严重；后者取自温室，背景统一为黑色，麦穗遮挡严重，且品种少见，外观与大众小麦品种差异巨大，各数据集如图 10 所示。

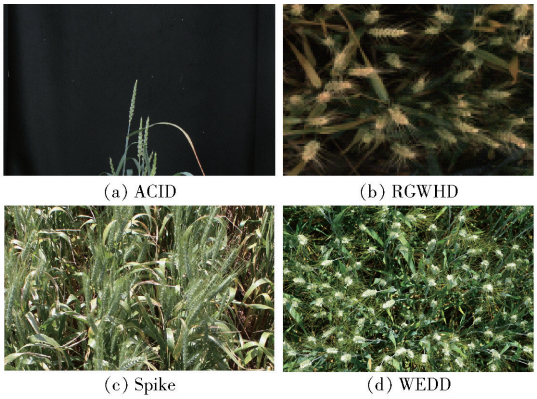



图 10　各数据集麦穗示例

Fig. 10　Examples of wheat ears for each dataset

## 4　结论

(1)针对麦穗个体在图像中显著性不足导致的目标特征提取不充分及识别准确率低的问题，提出一种基于改进 YOLO v8n 的小麦麦穗检测方法。结果显示，改进模型 YOLO v8－TRP 的平均精度均值相较于原模型 YOLO v8n 提升 2.9 个百分点，充分说明本文模型参照官方推荐的数据集划分方式，能够有效识别小麦麦穗，可为基于深度学习的麦穗识别研究提供参考。

(2)为验证本文模型的有效性和鲁棒性，使用改进后的 YOLO v8－TRP 模型对不同数据集的小麦进行验证实验。结果显示，改进的模型在公共数据集 WEDD、Spike、RGWHD、ACID、HBAUW 上的识别精确率，较原始模型也能平均提升 2.7 个百分点，都展现出优越的性能，验证了本文模型的有效性。

## 参　考　文　献

[1]　时雷，杨程凯，雷镜楷，等. 基于改进 YOLO v8s 的小麦小穗赤霉病检测研究[J]. 农业机械学报，2024，55(7)：280－289.

SHI Lei, YANG Chengkai, LEI Jingkai, et al. Wheat spikelet detection of fusarium head blight based on improved YOLO v8s [J]. Transactions of the Chinese Society for Agricultural Machinery, 2024, 55(7): 280 - 289. (in Chinese)
[2] 刘哲,黄文准,王利平. 基于改进 K - means 聚类算法的大田麦穗自动计数[J]. 农业工程学报, 2019, 35(3): 174 - 181.
LIU Zhe, HUANG Wenzhun, WANG Liping. Field wheat ear counting automatically based on improved K - means clustering algorithm[J]. Transactions of the CSAE, 2019, 35(3): 174 - 181. (in Chinese)
[3] 范梦扬,马钦,刘峻明,等. 基于机器视觉的大田环境小麦麦穗计数方法[J]. 农业机械学报, 2015, 46(增刊): 234 - 239.
FAN Mengyang, MA Qin, LIU Junming, et al. Counting method of wheatear in field based on machine vision technology[J]. Transactions of the Chinese Society for Agricultural Machinery, 2015, 46(Supp.): 234 - 239. (in Chinese)
[4] 鲍文霞,张鑫,胡根生,等. 基于深度卷积神经网络的田间麦穗密度估计及计数[J]. 农业工程学报, 2020, 36(21): 186 - 193,323.
BAO Wenxia, ZHANG Xin, HU Gensheng, et al. Estimation and counting of wheat ears density in field based on deep convolutional neural network[J]. Transactions of the CSAE, 2020, 36(21): 186 - 193, 323. (in Chinese)
[5] 陶兆胜,宫保国,李庆萍,等. 基于残差网络和特征融合的小麦图像修复模型[J]. 农业机械学报, 2023, 54(3): 318 - 327.
TAO Zhaosheng, GONG Baoguo, LI Qingping, et al. Wheat image inpainting based on residual networks and feature fusion[J]. Transactions of the Chinese Society for Agricultural Machinery, 2023, 54(3): 318 - 327. (in Chinese)
[6] 张震明,黄子琦,王东,等. 基于深度学习的田间麦穗检测[J]. 山东农业大学学报(自然科学版), 2022, 53(5):790 - 795.
ZHANG Zhenming, HUANG Ziqi, WANG Dong, et al. Field wheat ear detection based on deep learning[J]. Journal of Shandong Agricultural University Natural Science Edition, 2022, 53(5): 790 - 795. (in Chinese)
[7] ZHAO J, YAN J, XUE T, et al. A deep learning method for oriented and small wheat spike detection (OSWSDet) in UAV images[J]. Computers and Electronics in Agriculture, 2022, 198: 107087.
[8] DAVID E, MADEC S, SADEGHI-TEHRAN P, et al. Global wheat head detection (GWHD) dataset: a large and diverse dataset of high-resolution RGB-labelled images to develop and benchmark wheat head detection methods[J]. Plant Phenomics, 2020, 2(1):12.
[9] DAVID E, SEROUART M, SMITH D, et al. Global wheat head detection 2021: an improved dataset for benchmarking wheat head detection methods[J]. Plant Phenomics, 2021,2021:9846158.
[10] YANG X, FANG C. Lightweight YOLOv8 for wheat head detection[J]. IEEE Access, 2024,2024:12.
[11] LI R, WU Y. Improved YOLO v5 wheat ear detection algorithm based on attention mechanism[J]. Electronics, 2022, 11(11): 1673.
[12] WEN C, WU J, CHEN H, et al. Wheat spike detection and counting in the field based on SpikeRetinaNet[J]. Frontiers in Plant Science, 2022, 13: 821717.
[13] 董燕,刘运东,李卫杰,等. 基于三目注意力机制的高效轻量化麦穗检测算法[J]. 计算机工程与设计, 2023, 44(7): 2062 - 2069.
DONG Yan, LIU Yundong, LI Weijie, et al. Efficient and lightweight wheat ear detection algorithm with trinocular attention mechanism[J]. Computer Engineering and Design, 2023, 44(7): 2062 - 2069. (in Chinese)
[14] QING S, QIU Z, WANG W, et al. Improved YOLO - FastestV2 wheat spike detection model based on a multi-stage attention mechanism with a LightFPN detection head[J]. Frontiers in Plant Science, 2024, 15: 1411510.
[15] 李婕,杨子豪,郑权,等. 基于 RT - WEDT 的麦穗检测与计数方法[J]. 农业工程学报, 2024, 40(21): 146 - 156.
LI Jie, YANG Zihao, ZHENG Quan, et al. Method for detecting and counting wheat ears using RT - WEDT[J]. Transactions of the CSAE, 2024, 40(21): 146 - 156. (in Chinese)
[16] ZHANG X, LIU C, YANG D, et al. RFAConv: innovating spatial attention and standard convolutional operation[J]. arXiv Preprint arXiv:2304.03198, 2023.
[17] WOO S, PARK J, LEE J Y, et al. CBAM: convolutional block attention module[J/OL]. Springer, Cham, 2018. DOI:10.1007/978 - 3 - 030 - 01234 - 2_1.
[18] WANG C Y, YEH I H, LIAO H Y M. YOLOv9: learning what you want to learn using programmable gradient information [J]. arXiv Preprint, arXiv:2402.13616, 2024.
[19] WANG C Y, BOCHKOVSKIY A, LIAO H Y M. YOLOv7: trainable bag-of-freebies sets new state-of-the-art for real-time object detectors[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023: 7464 - 7475.
[20] LI C, LI L, JIANG H, et al. YOLOv6: a single-stage object detection framework for industrial applications[J]. arXiv Preprint, arXiv:2209.02976, 2022.
[21] REN S, HE K, GIRSHICK R, et al. Faster R - CNN: towards real-time object detection with region proposal networks[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2017, 39(6):1137 - 1149.
[22] WEI L, DRAGOMIR A, DUMITRU E, et al. SSD: single shot multibox detector[J/OL]. Springer, Cham, 2016. DOI:10.1007/978 - 3 - 319 - 46448 - 0_2.
[23] BATIN M A, ISLAM M, HASAN M M, et al. WheatSpikeNet: an improved wheat spike segmentation model for accurate estimation from field imaging[J]. Frontiers in Plant Science, 2023, 14: 1226190.
[24] MADEC S, JIN X, LU H, et al. Ear density estimation from high resolution RGB imagery using deep learning technique[J]. Agricultural and Forest Meteorology, 2019, 264: 225 - 234.
[25] POUND M P, ATKINSON J A, WELLS D M, et al. Deep learning for multi-task plant phenoty[C]//Proceedings of the IEEE International Conference on Computer Vision Workshops, 2017: 2055 - 2063.
[26] YU J, CHEN W, LIU N, et al. Oriented feature pyramid network for small and dense wheat heads detection and counting[J]. Scientific Reports, 2024, 14(1): 8106.