doi:10.6041/j.issn.1000-1298.2026.09.029

基于 GB－YOLO v5su 的轻量化小麦麦穗检测方法

沈淑娟[1]　杨　婷[2]　张杜娟[3]　曹建春[4]

（1.郑州大学计算机与人工智能学院，郑州 450001；2.郑州航空工业管理学院航空发动机学院，郑州 450015；
3.国家超级计算郑州中心，郑州 450001；4.黄河水利职业技术大学信息工程学院，开封 475001）

摘要：针对复杂农田环境中因植株遮挡及不同生长阶段形态差异导致的小麦麦穗检测精度不足问题，本文提出一种融合轻量化设计与多尺度特征增强的检测算法 GB－YOLO v5su（Ghost and BiFPN-enhanced YOLO v5su）。该算法以 YOLO v5su（YOLO v5 small anchor-free）为基线模型，采用 Ghost 模块重构骨干网络，以构建轻量化 GhostNet 结构，在保持模型表征能力的同时显著降低计算复杂度与参数量；采用加权双向特征金字塔网络（BiFPN）替换原 PANet 结构，通过跨尺度连接与可学习的自适应权重机制，增强多尺度特征融合能力；同时新增 P2 检测层构建四尺度特征金字塔，充分利用高分辨率浅层特征提升对小尺度麦穗的检测灵敏度；引入 MPDIoU 损失函数替代 CIoU，通过最小化预测框与真实框角点距离提升定位精度。在全球小麦检测数据集 GWHD2021 上的实验结果表明，本文算法在参数量仅为 5.38×10^6 时，平均精度 AP50 达到 92.8%，实现了精度与效率的良好平衡。系统的消融实验与跨数据集测试进一步验证了各模块的有效性及算法的强鲁棒性。本研究为实现资源受限田间边缘计算场景下的实时精准麦穗检测，提供了一种高效、实用的轻量化部署方案。

关键词：麦穗检测；GB－YOLO v5su；轻量化目标检测；多尺度特征融合；边缘计算

中图分类号：TP391　　**文献标识码**：A　　**文章编号**：1000-1298(2026)09-0310-09　　**OSID**：

Lightweight Wheat Head Detection Method Using GB－YOLO v5su

SHEN Shujuan[1]　YANG Ting[2]　ZHANG Dujuan[3]　CAO Jianchun[4]

(1. *School of Computer Science and Artificial Intelligence, Zhengzhou University, Zhengzhou* 450001, *China*
2. *School of Aero Engine, Zhengzhou University of Aeronautics, Zhengzhou* 450015, *China*
3. *National Supercomputing Center in Zhengzhou, Zhengzhou* 450001, *China*
4. *School of Information Engineering, Yellow River Conservancy Technical University, Kaifeng* 475001, *China*)

Abstract: Aiming to address the issue of insufficient detection accuracy for wheat heads in complex field environments caused by plant occlusion and growth stage variations, a detection algorithm named GB－YOLO v5su, was proposed which integrated lightweight design with multi-scale feature enhancement. Building upon the YOLO v5su model, the algorithm reconstructed the backbone network by incorporating Ghost modules to form a lightweight GhostNet structure, substantially reducing computational complexity and model parameters while maintaining representational capacity. The original PANet was replaced with a weighted bidirectional feature pyramid network (BiFPN), which enhanced multi-scale feature representation through efficient cross-scale connections and a learnable adaptive weighting mechanism. Subsequently, a P2 detection layer was introduced to construct a four-scale feature pyramid, effectively leveraging high-resolution shallow features to improve detection sensitivity for small-scale wheat heads. Furthermore, the MPDIoU loss function was adopted to replace CIoU, which optimized bounding box regression by directly minimizing the corner point distance, thereby improving localization accuracy. Experimental results on the public GWHD2021 dataset demonstrated that the proposed algorithm reduced the model parameter count to 5.38×10^6 while achieving an average precision (AP50) of 92.8%, striking an effective balance between accuracy and efficiency. Extensive ablation studies and cross-dataset tests further validated the contribution of each module and confirmed the strong robustness of the algorithm. The

收稿日期：2025－11－18　修回日期：2025－12－02

基金项目：国家自然科学基金项目(42501427)

作者简介：沈淑娟(1972—)，女，讲师，主要从事计算机视觉研究，E-mail：iesjshen@zzu.edu.cn

通信作者：曹建春(1970—)，男，教授，主要从事人工智能与智慧农业应用研究，E-mail：cjcssj@163.com

research result can provide an efficient, practical, and lightweight deployment solution for achieving real-time and accurate wheat head detection in resource-constrained field edge computing scenarios.

Key words: wheat head detection; GB－YOLO v5su; lightweight object detection; multi-scale feature fusion; edge computing

0　引言

小麦作为世界主要粮食作物，其产量的准确预测对于保障粮食安全与农业决策至关重要。单位面积麦穗数是构成产量的关键要素，因此实现麦穗高效、精准自动检测，已成为推动小麦估产数字化与智能化亟待解决的核心技术问题[1]。

早期基于颜色与纹理特征的麦穗计数方法[2-3]，尽管取得了一定成效，但其对成像条件变化敏感，环境适应性有限，泛化能力较弱。近年来，深度学习技术的引入显著提升了麦穗检测的精度与鲁棒性。与传统方法依赖手工设计浅层特征不同，深度学习可通过多层神经网络自动学习从像素到麦穗概念的层次化特征表达[4-6]。当前基于深度学习的目标检测框架主要分为 2 类：一类是锚框（Anchor-based）框架，如 R－CNN[7]、早期 YOLO[8] 系列。针对麦穗检测中小目标特征弱、密集遮挡并且目标与背景相似度高等问题，通过引入注意力机制[9-11]、优化多尺度特征融合结构[11-12]提升检测性能。但这类方法依赖固定尺寸与比例的预设锚框，在麦穗不同生长阶段尺度差异大、种植密度多样的田间场景中易出现锚框与真实目标不匹配，导致漏检与误检。另一类是无锚框（Anchor-free）检测范式，如基于 Transformer 的 DETR[13] 及后续改进方法[14]，虽能自适应麦穗形态与尺度变化，但存在计算复杂度高、训练收敛慢等问题，难以满足田间设备的资源约束。相比之下，无锚框 YOLO 模型因其兼顾检测精度与推理效率，在农业实时检测研究中更受青睐。

然而，无锚框 YOLO 模型直接应用于田间麦穗检测时面临三重关键挑战：① 特征融合适配性不足，传统特征金字塔网络（Feature pyramid network，FPN）[15]、路径聚合网络（Path aggregation network，PANet）[16] 在融合麦穗细粒度纹理浅层特征与植株密集区域语义信息深层特征时存在信息损耗，导致灌浆期密集麦穗易漏检。② 小目标表征能力有限，仅通过 P3－P4－P5 三尺度检测头，难以捕捉像素尺寸小于 32 像素 ×32 像素的小尺度麦穗细节特征，下采样过程中易丢失麦穗芒刺等关键区分信息。③ 定位精度与部署效率双重受限，现有 CIoU 损失函数在麦穗遮挡场景下，如成熟期麦穗重叠度超 30%，难以精准优化边界框，定位精度不足；并且模型参数量往往超出田间边缘设备资源约束，制约了模型在实际应用场景中的部署与应用。

针对上述问题，本文提出 GB－YOLO v5su 轻量化检测算法。该算法从特征提取、多尺度融合与边界框回归 3 个层面进行改进，在提升定位精度的同时显著降低模型参数量：提出一种融合 GhostNet[17] 与双向加权特征金字塔网络（Bidirectional feature pyramid network，BiFPN）[18] 的轻量化增强型网络结构；构建包含 P2 层的四尺度检测体系，强化小目标感知能力；采用 MPDIoU 损失函数，提升遮挡场景下的定位精度。以期为资源受限的田间边缘计算场景提供一种实时、精准的麦穗检测解决方案，实现模型精度与效率的良好平衡。

1　实验数据

2020 年，DAVID 等[19] 发布了全球小麦穗检测数据集（Global Wheat Head Detection，GWHD），为该领域算法研究提供了重要的基准数据支撑，有力推动了麦穗检测研究的发展。该团队在次年进一步扩展了数据集，新版本在规模、多样性、标注质量及研究实用性方面均显著优于 2020 年版，已成为该领域关键的研究资源，本文以 GWHD2021[20] 为实验数据。该数据集图像来自 12 个国家采集的不同地点的实验数据。图像使用多种地面表型平台和相机进行采集，高度范围为 1.8 ~ 3 m，焦距从 10 mm 到 50 mm 不等。数据集涵盖了不同的种植密度、生长阶段和环境条件，图像示例如图 1 所示。图像在光照条件、麦穗方向、视图方向、麦穗外观等方面呈现多样性。GWHD2021 共有图像 6 515 幅，图像分辨率均为 1 024 像素 ×1 024 像素，其中训练集有 3 657 幅图像，验证集和测试集分别为 1 476 幅和1 382 幅图像。

2　模型设计

YOLO（You only look once）系列算法由于其速度与精度的平衡、生态成熟度高等优势，在目标检测领域中被广泛应用[21]。在 Ultralytics 团队开发的系列模型中，YOLO v5 因其具有高度的工程成熟度与优良的稳定性被广泛采用。后续发布的 YOLO v8 等更新模型扩展了实例分割、姿态估计等多任务能力，而 YOLO v5 更专注于目标检测任务，并针对实

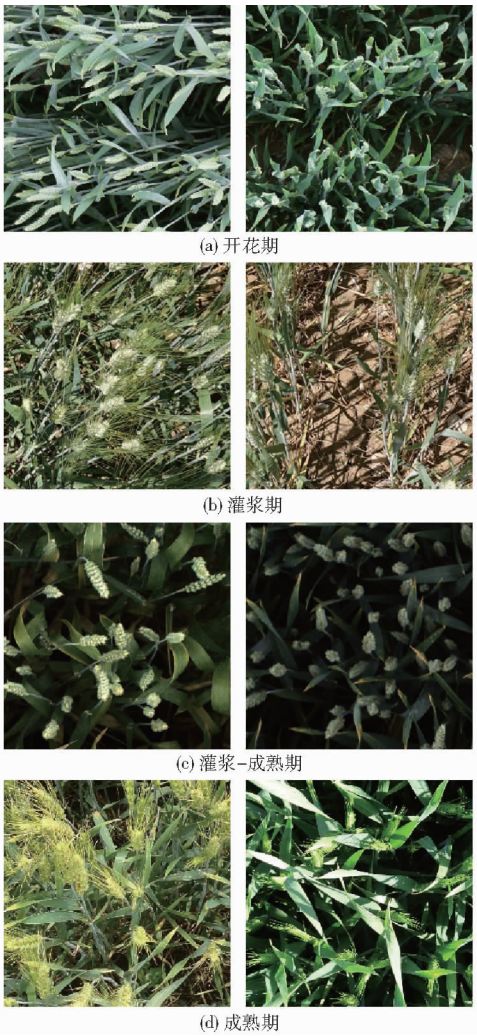


图 1 GWHD2021 样本集示例图像

Fig. 1 Examples from GHWD2021 dataset

时应用进行了深度优化,在资源受限,如无 GPU 支持的部署环境中表现出显著优势。YOLOBench 基准测试表明,在 Intel CPU 和 ARM CPU 等嵌入式平台上,YOLO v5 系列模型与 YOLO v6、YOLO v7 等后续模型相比,在精度与速度权衡上保持了强大的竞争力[22]。值得注意的是,YOLO v5 仍在不断进步中,YOLO v5u 整合了无锚框等机制,优化了检测头的设计,这些优化进一步提升了模型性能。

基于上述原因,本文选用 YOLO v5su 作为基线模型,提出一种融合轻量化设计与多尺度特征的小麦麦穗检测算法。该算法在特征提取阶段采用 GhostNet 结构[17]对卷积过程进行简化;使用双向加权特征金字塔网络(BiFPN)进行多尺度特征融合;新增 P2 检测头构建四尺度检测体系;使用 MPDIoU 替代 CIoU 精准优化边界框定位。改进后的网络结构如图 2 所示。

2.1 无锚框检测机制

早期的 YOLO 系列普遍采用基于锚框的方法,需要预先定义几组不同尺度、不同长宽比的先验锚框,通过预测其与真实框的偏移和尺寸缩放来进行目标定位。这种方法不仅引入了对锚框尺寸、数量和比例敏感的超参数,还增加了计算复杂度。

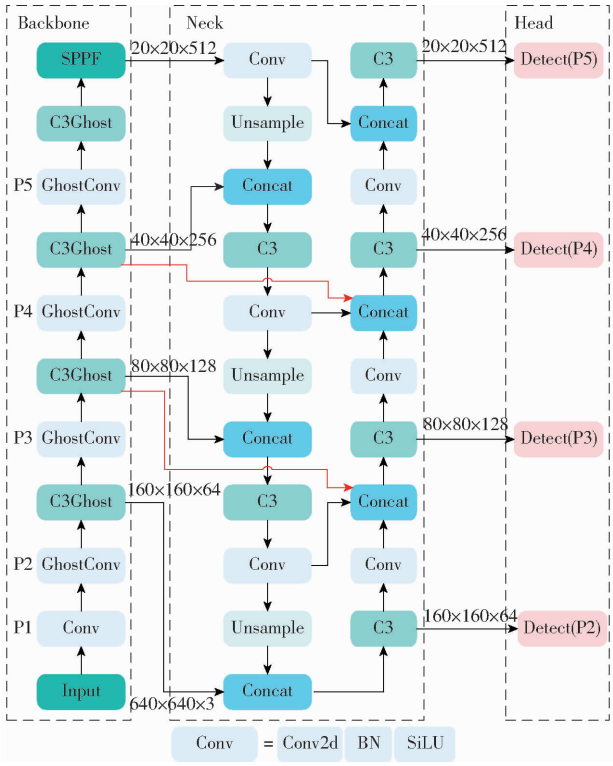


图 2 GB-YOLO v5su 网络结构

Fig. 2 Network structure of GB-YOLO v5su

本文采用的 YOLO v5su 无锚框机制则简化了这一流程。其核心是直接预测目标边界框的宽高和中心点相对于所在网格左上角的归一化偏移,无需预定义锚框。具体而言,检测头为每个输出网格预测目标置信度、类别概率,以及边界框的中心偏移量和尺寸。该机制省去了锚框设计与匹配过程,使检测流程更简洁,降低了计算开销。无锚框机制通过减少超参数依赖,增强了模型对尺度多变的田间麦穗的泛化能力与定位灵活性。

2.2 YOLO v5su 骨干网络的轻量化

已有研究表明,常规卷积操作生成的特征图存在冗余,如图 3 所示,其中 3 对相似的特征图用相同颜色框进行了标注。部分特征图可通过低复杂度的线性变换从其他特征图有效生成,为此,本文引入 GhostNet 结构对卷积过程进行简化,从而实现参数量的有效压缩与模型复杂度的降低。

与常规卷积不同,GhostNet 中的 Ghost 卷积模块将传统卷积拆分为“生成固有特征图”和“生成幽灵特征图”两步,避免冗余特征图的重复高成本计算[17]。具体而言,通过标准卷积生成 m 个固有特征图,m 小于目标通道数 n,其结构如图 4b 所示。随后对每个固有特征图应用廉价线性变换,叠加恒

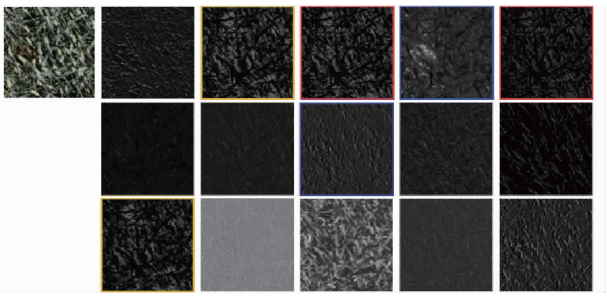


图 3　相似特征图对示例

Fig. 3　Examples of similar feature map pairs

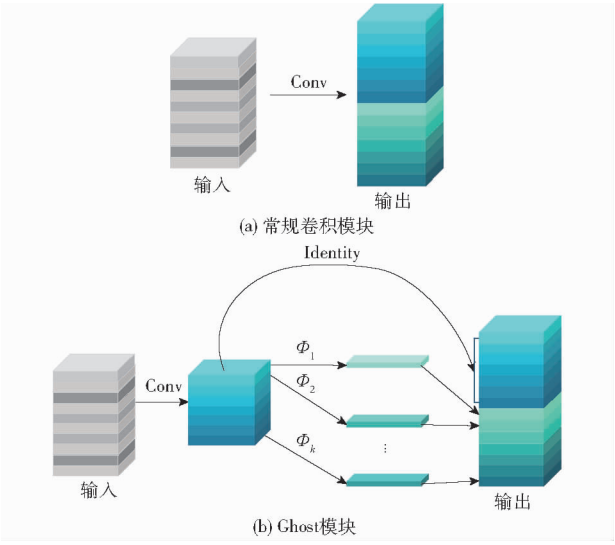


图 4　卷积层与 Ghost 模块结构对比

Fig. 4　Convolutional layer vs. Ghost module

等映射，最终生成 n 个特征图。

在本算法的实现中，针对 YOLO v5 模型的多尺度特征提取，对其骨干网络进行了结构重构：采用 GhostConv 模块替代常规卷积模块 Conv，C3Ghost 模块替换 C3 模块。该改进在保持模型特征表征能力的同时，显著降低了计算复杂度和参数量。

为应对不同下采样需求，基于 GhostConv 模块构建了 2 种 GhostBottleneck 结构。如图 5 所示，当 stride 为 1 时，模块由 2 个 GhostConv 构成主路径，结合恒等映射的捷径路径，实现特征增强而不改变分辨率；当 stride 为 2 时，在主路径中插入深度卷积实现下采样，同时捷径路径通过深度卷积和 1×1 卷积完成分辨率和通道调整，确保残差连接的有效性。在 YOLO v5 原有的 C3 模块基础上，其内部的 Bottleneck 单元替换为 GhostBottleneck，形成 C3Ghost 模块。该模块继承了 C3 的多分支特征融合优势，同时通过轻量化设计显著降低计算量。

经过上述操作后，取出第 1 个 Ghost 卷积后所得的前 15 幅特征图，如图 6 所示。通过对比图 3 和图 6 的结果，可以观察到，采用 Ghost 卷积模块后，模型生成的特征图冗余度明显降低，相似特征图数

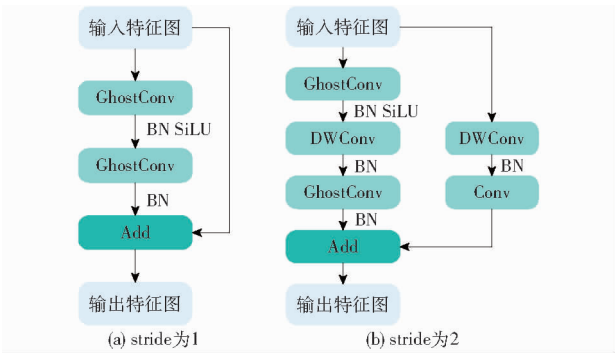


图 5　两种 GhostBottleneck 结构

Fig. 5　Two GhostBottleneck architectures

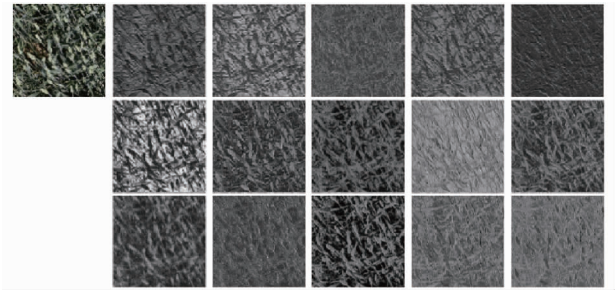


图 6　卷积输出结果可视化

Fig. 6　Visualization of Ghost convolution outputs

量减少。这表明，Ghost 模块在减少计算复杂度的同时，仍能提取出较为完整的特征信息，从而验证了其在轻量化网络设计中的实用价值。

2.3　多尺度特征融合与增强

2.3.1　基于 BiFPN 的特征融合

多尺度特征融合的核心目标在于聚合来自不同分辨率的特征信息[23]。传统的特征金字塔网络 FPN 只有自上而下的单向信息流，路径聚合网络 PANet 在 FPN 的基础上增加了一条额外的自底向上的路径进行信息增强，有效保留更多浅层特征，结构如图 7a 所示。考虑到在不同的特征层中，输入特征具有不同的分辨率，对输出特征的贡献通常并不相等，谷歌公司团队提出了双向加权特征金字塔网络 BiFPN，通过双向跨尺度连接与可学习权重机制，实现对不同尺度特征的动态融合，从而强化有效特征的传递，削弱冗余、噪声特征的干扰，进一步提升多尺度特征融合效果，其结构如图 7b 所示。

YOLO v5 原网络生成 P1 至 P5 5 个尺度的特征层，采用 PANet 结构实现上下层特征融合，并且仅选取 P3、P4 和 P5 3 层参与融合，如图 7a 所示。本文借鉴 BiFPN 思想，采用图 7c 所示结构替换原网络中的 PANet 模块，并增加 P2 特征层参与特征融合。P2 层蕴含丰富的高分辨率细节信息，对于提升小尺度麦穗的检测性能具有关键作用。在融合路径上，本文引入跨尺度连接，将特征提取网络中提取的特征直接与自底向上路径中对应尺度的

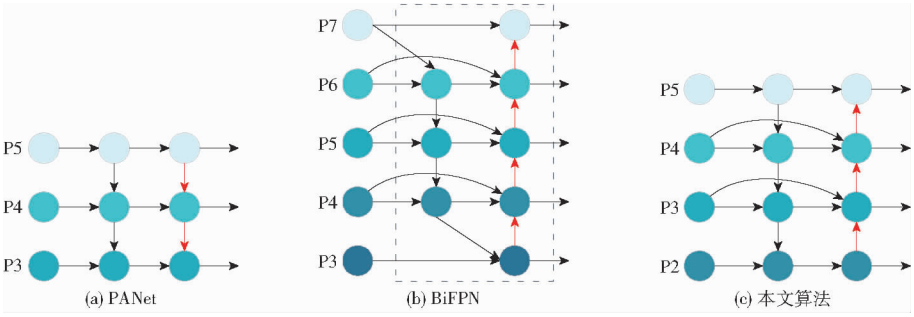


图 7 特征融合结构

Fig. 7 Feature fusion architecture

特征进行融合,从而在保留更多浅层语义信息的同时,尽可能不丢失深层语义信息。具体而言,第 24 层特征新融入第 4 层的浅层特征,第 27 层新融入第 6 层的浅层特征,提高网络对于深浅层特征的融合度。此外,对于参与融合的多个尺度特征,为每个输入特征分配一个可学习权重,模型在训练过程中根据检测损失自动调整这些权重,最终让模型自主判断不同尺度特征的重要性,实现自适应的加权特征融合。尽管此举会引入少量的计算开销,但在目标分布密集、特征相对简单的农业场景数据集上,该结构能够更有效地融合多尺度信息,从而取得更优异的检测性能。

2.3.2 P2 小目标检测层

为进一步强化小尺寸麦穗的感知能力,本文在颈部网络末端构建了四尺度“P2 - P3 - P4 - P5”检测体系。原 YOLO v5 模型采用“P3 - P4 - P5”三尺度检测头,对于图像中占比较小的麦穗目标特征响应较弱。新增的 P2 检测层基于网络早期下采样特征图,分辨率为 160 像素 ×160 像素,能够有效保留麦穗的细粒度纹理特征和边缘信息,直接匹配小目标的特征尺度需求。具体而言,将颈部网络得到的 P3 特征图通过卷积和上采样至 160 像素 ×160 像素,与相同尺寸的第 2 层特征图进行通道对齐拼接,形成增强的 P2 特征层。该层作为检测头的输入端之一,与 P3、P4、P5 层共同构成四尺度检测结构,既保留了浅层特征的高分辨率,又融合了深层特征的丰富语义信息,显著提升了模型对小尺寸麦穗的检测能力。

2.4 MPDIoU 损失函数

精确的边界框回归(Bounding box regression, BBR)是目标检测的关键。现有的损失函数主要分为基于 L_n-norm 范数和基于交并比(Intersection over Union, IoU)2 大类。然而,包括 CIoU 在内的主流 IoU 系列损失函数存在一个固有局限:当预测框与真实框中心点重合且宽高比相同,但宽高绝对尺寸不同时,其损失值相同,无法为模型提供有效的梯度来区分并优化框的尺寸[24]。这限制了模型在需要对边界框尺寸进行精细回归时的收敛速度与精度。

为了解决上述问题,本文采用基于最小点距离的 MPDIoU(Minimum point distance IoU)构建损失函数[24]。该损失函数的核心思想是直接最小化预测框与真实框左上角和右下角 2 个关键点的欧氏距离平方,从而能够同时且直接地驱动边界框的位置与尺寸回归。

给定一个预测框 $\boldsymbol{B}_{prd}=[x_1^{prd} \quad y_1^{prd} \quad x_2^{prd} \quad y_2^{prd}]^{\mathrm{T}}$,其对应的真实框 $\boldsymbol{B}_{gt}=[x_1^{gt} \quad y_1^{gt} \quad x_2^{gt} \quad y_2^{gt}]^{\mathrm{T}}$,MPDIoU 的计算公式为

$$MPDIoU = IoU - \frac{d_1^2}{w^2+h^2} - \frac{d_2^2}{w^2+h^2} \tag{1}$$

其中

$$d_1^2 = (x_1^{prd} - x_1^{gt})^2 + (y_1^{prd} - y_1^{gt})^2 \tag{2}$$

$$d_2^2 = (x_2^{prd} - x_2^{gt})^2 + (y_2^{prd} - y_2^{gt})^2 \tag{3}$$

$$IoU = \frac{\boldsymbol{B}_{gt} \cap \boldsymbol{B}_{prd}}{\boldsymbol{B}_{gt} \cup \boldsymbol{B}_{prd}} \tag{4}$$

式中 w、h——输入图像的宽度、高度

IoU——预测框与真实框的交并比

相应的损失函数定义为

$$L_{MPDIoU} = 1 - MPDIoU \tag{5}$$

相较于 CIoU,其在预测框与真实框中心点重合且宽高比相同但尺寸不同时,依然能提供有效的优化梯度,有助于获得更快的收敛速度和更高的定位精度。此外,其计算仅依赖边界框的角点坐标,无需复杂的高宽比计算或寻找最小外接矩形,形式简洁。它所涵盖的重叠/非重叠面积、中心点距离及宽高比偏差等几何因素,均可直接由两个边界框的角点坐标推导得出。在本文的 GB - YOLO v5su 模型中,使用此 MPDIoU 损失函数替代原 CIoU 损失函数,旨在更精准地优化模型对小麦麦穗目标的定位能力。

3 实验结果与分析

3.1 评价指标

为评价小麦麦穗检测模型的性能,使用两类评价指标。一类是精度指标,包括:精确率(Precision,

P）、召回率（Recall，R）和平均精度（Average precision，AP），AP 为 P－R 曲线与坐标轴围成的面积，用于衡量模型对一类目标的检测性能。本文采用 IoU 为 0.5 时的平均精度 AP50 作为主要评价指标。另一类是效率指标，包括：参数量（Parameters）和浮点计算量（Floating point operations，FLOPs），用来衡量模型的复杂度和计算量。

3.2　网络训练与超参数设置

本实验运行环境配置：硬件平台采用 Intel Core i7－14700KF CPU 与 NVIDIA GeForce RTX 4070 Ti SUPER GPU，显存 16 GB。软件环境为 Windows 11 操作系统，使用 Python 3.13 编程语言，采用 PyTorch 2.7.1 深度学习框架与 CUDA 12.6 进行模型训练。训练超参数设置如下：输入图像尺寸为 640 像素×640 像素。训练共进行 200 个轮次，批量大小为 16。优化器使用 Adam，其参数设置：初始学习率为 0.01，权重衰减系数为 5×10^{-4}，采用余弦退火算法作为学习率调整策略。

3.3　不同模型性能对比

表 1 列出了不同模型在 GWHD2021 数据集上的检测结果。从表 1 可以看出，传统目标检测模型 SSD、Faster R－CNN 的综合表现明显落后于 YOLO 模型。其中两阶段目标检测模型 Faster R－CNN 的精确率、召回率和 AP50 分别为 64.7%、57.2% 和 63.8%。与之相比，单阶段目标检测模型 SSD 表现稍好，精确率、召回率和 AP50 分别为 75.6%、66.8% 和 74.3%。而作为单阶段目标检测模型中突破传统单阶段性能瓶颈的典范，YOLO 系列模型均展现出更优的性能，各项核心指标均大幅超越传统模型，参数量也更具轻量化优势。

作为本文改造后的模型，GB－YOLO v5su 的性能优势尤为突出，在核心指标和轻量化程度上都表现最优。精确率达 92.7%，召回率达 86.1%，AP50 达 92.8%，优于基线模型 YOLO v5su，其精确率为 92.4%，召回率为 85.3%，AP50 为 92.1%，也优于 YOLO v8s、YOLO v10s 和 YOLO 11s，展现出更强的目标检测精确率和召回率。

表 1　不同模型检测结果

Tab.1　Detection results of different models

模型	P/%	R/%	AP50/%	参数量
Faster R－CNN	64.7	57.2	63.8	2.56×10^{7}
SSD	75.6	66.8	74.3	2.31×10^{7}
YOLO v5n	91.6	83.9	91.1	1.76×10^{6}
YOLO v5su	92.4	85.3	92.1	7.01×10^{6}
YOLO v8s	92.5	85.0	92.3	1.11×10^{7}
YOLO v10s	91.6	85.3	92.0	8.07×10^{6}
YOLO 11s	92.0	85.7	92.6	9.42×10^{6}
GB－YOLO v5su	92.7	86.1	92.8	5.38×10^{6}

GB－YOLO v5su 的参数量仅为 5.38×10^{6}，是基线模型 YOLO v5su 参数量（7.01×10^{6}）的 76.7%，更低于 YOLO v8s 参数量（1.11×10^{7}）、YOLO v10s 参数量（8.07×10^{6}）和 YOLO 11s 参数量（9.42×10^{6}），仅高于极致轻量化 YOLO v5n 参数量（1.76×10^{6}），但 YOLO v5n 的精确率为 91.6%、召回率为 83.9%，AP50 为 91.1%，与 GB－YOLO v5su 存在明显差距。

3.4　模型消融实验

为系统评估各改进模块的贡献，在基线模型 YOLO v5su 上设计了消融实验。实验固定数据集、超参数及硬件配置，依次引入 BiFPN、Ghost 和 P2 模块，共构成 8 种策略组合进行对比。

实验评估指标涵盖检测性能与模型效率两方面。检测性能指标除精确率（P）、召回率（R）和平均精度（AP50）外，还包括多区间平均精度 AP50－75。AP50－75 指标通过计算交并比（IoU）阈值从 0.5 到 0.75（步长 0.05）区间内各阈值下 AP 值的均值，能更综合地衡量模型在不同定位精度下的鲁棒性。效率指标则包括参数量（Parameters）与浮点计算量（FLOPs）。实验结果如表 2 所示。

表 2　模型消融实验结果

Tab.2　Model ablation study results

BiFPN	Ghost	P2	P/%	R/%	AP50/%	AP50－75/%	AP_S/%	AP_M/%	AP_L/%	参数量	浮点计算量
×	×	×	92.4	85.3	92.1	76.3	16.3	70.8	80.6	7.01×10^{6}	1.58×10^{10}
✓	×	×	92.9	86.9	92.9	78.5	17.3	71.2	80.6	7.08×10^{6}	1.60×10^{10}
×	✓	×	91.7	84.3	91.3	75.6	13.8	70.1	77.6	5.08×10^{6}	1.05×10^{10}
×	×	✓	93.0	86.9	93.0	78.6	19.9	72.5	81.7	7.16×10^{6}	1.85×10^{10}
✓	✓	×	92.6	85.6	92.3	76.5	14.8	71.1	79.6	5.14×10^{6}	1.08×10^{10}
×	✓	✓	92.8	85.2	92.2	76.6	16.5	70.5	80.3	5.24×10^{6}	1.33×10^{10}
✓	×	✓	93.2	87.1	93.4	79.0	21.7	72.3	82.0	8.48×10^{6}	2.26×10^{10}
✓	✓	✓	92.7	86.1	92.8	78.1	18.9	70.6	81.4	5.38×10^{6}	1.40×10^{10}

注：“✓”表示有此模块，“×”表示无此模块。AP_S、AP_M、AP_L 表示在小目标、中目标和大目标范围内的 AP50－75 值。

以基线模型的性能作为基准,其精确率为92.4%,召回率为85.3%,平均精度AP50为92.1%,多区间平均精度AP50-75为76.3%,参数量为7.01×10^6。

单一模块效果:BiFPN模块仅以7×10^4的参数量增加显著提升了检测性能,AP50提升至92.9%,AP50-75提升2.2个百分点至78.5%,验证了多尺度特征融合的有效性。Ghost模块实现了明显的轻量化效果,参数量降低27.5%至5.08×10^6,但以精度损失为代价,AP50降至91.3%。P2检测层对小目标改善最为明显,AP_S从16.3%大幅提升至19.9%,但参数量增至7.16×10^6。

模块组合分析:BiFPN与Ghost组合在保持参数量5.14×10^6轻量化优势的同时,AP50回升至92.3%,表明特征增强可以补偿轻量化带来的部分精度损失。Ghost与P2组合与前一组合结果相似,AP50与基线模型基本持平。BiFPN与P2组合达到最高检测精度,AP50达93.4%,但参数量也最高至8.48×10^6。GB-YOLO v5su模型集成了3个模块,在参数量降低至5.38×10^6的前提下,仍保持优良的检测性能,AP50达92.8%,并且小目标检测精度AP_S较基线模型提升了16.0%,实现了精度与效率的最佳平衡。

实验表明,各改进模块均有效并且功能互补:BiFPN增强多尺度特征融合,P2提升小目标检测性能,Ghost实现模型轻量化,三者协同工作为边缘计算场景提供了理想的解决方案。

3.5 模型鲁棒性验证

为验证改进模型的鲁棒性,本文在另外3个公开数据集上进行实验。其中ACID(Annotated crop image dataset)[4]是专注于小麦穗和穗粒的高分辨率图像数据集,在温室条件下拍摄,背景为黑色以确保一致性,共520幅图像,原始分辨率为1 956像素×1 530像素,未经过裁剪或增强处理。CFWED(Contrasted-fertilization wheat ear dataset)[25]基于2020年比利时两处田间实验的小麦RGB图像,覆盖从抽穗期到成熟期的全生长阶段。SPIKE数据集由HASAN等[26]于2018年首次发布,用于不同品种和施肥处理的小麦穗的检测与分析。

表3给出了GB-YOLO v5su算法与基线模型YOLO v5su在3个公开小麦数据集上的性能对比。实验结果表明,GB-YOLO v5su在ACID、CFWED与SPIKE数据集上的性能都显著提升,展现出优良的检测能力与跨数据集适应性。具体而言,在ACID数据集上,GB-YOLO v5su的精确率、召回率和AP50分别达到93.2%、87.6%、92.3%,较YOLO v5su分别提升1.7、10.3、2.5个百分点;在CFWED数据集上,其精确率、召回率和AP50分别达90.9%、92.7%、92.1%,较YOLO v5su提升2.3、7.6、1.5个百分点;在SPIKE数据集上,精确率、召回率和AP50达89.5%、91.5%、91.6%,较YOLO v5su提升6.8、13.6、7.0个百分点。本文方法召回率提升最为显著,相较于YOLO v5su,在3个数据集上分别提高10.3、7.6、13.6个百分点。

表3 不同数据集的检测结果对比

Tab.3 Detection results comparison on multiple datasets

数据集	YOLO v5su			GB-YOLO v5su			性能提升值		
	P/%	R/%	AP50/%	P/%	R/%	AP50/%	P/百分点	R/百分点	AP50/百分点
ACID	91.5	77.3	89.8	93.2	87.6	92.3	1.7	10.3	2.5
CFWED	88.6	85.1	90.6	90.9	92.7	92.1	2.3	7.6	1.5
SPIKE	82.7	77.9	84.6	89.5	91.5	91.6	6.8	13.6	7.0

为进一步从视觉方面分析模型改进效果,图8展示了在3个数据集上的热力图与检测结果对比(图中,N_{TP}表示正确判定为正样本的数量,N_{FP}表示实际负样本错误判定为正样本的数量,N_{FN}表示实际正样本被错误判定为负样本的数量)。从热力图可以看出,在每组热力图中,右边GB-YOLO v5su对应的热力图,其高响应区与麦穗目标的轮廓吻合得更精确,尤其在植株密集和遮挡严重的区域,模型依然能够产生强烈的激活响应,表明模型对细微特征与重叠目标具有更强的辨识能力。这解释了召回率提升的原因:BiFPN结构促进了多尺度特征的深度融合,而增设的P2检测头增强了对浅层细节的感知,从而有效缓解了因目标遮挡或重叠造成的漏检问题。此外,MPDIoU损失函数的引入,通过优化边界框回归,进一步提升了模型的定位精度。

综合检测数据与可视化分析可知,GB-YOLO v5su能够适应不同生长阶段与成像环境的麦穗检测任务,在保持高精度的同时显著提升了对复杂场景中目标的召回能力,具有良好的鲁棒性和实用性。

4 结论

(1)本文针对田间复杂环境下小麦麦穗检测任

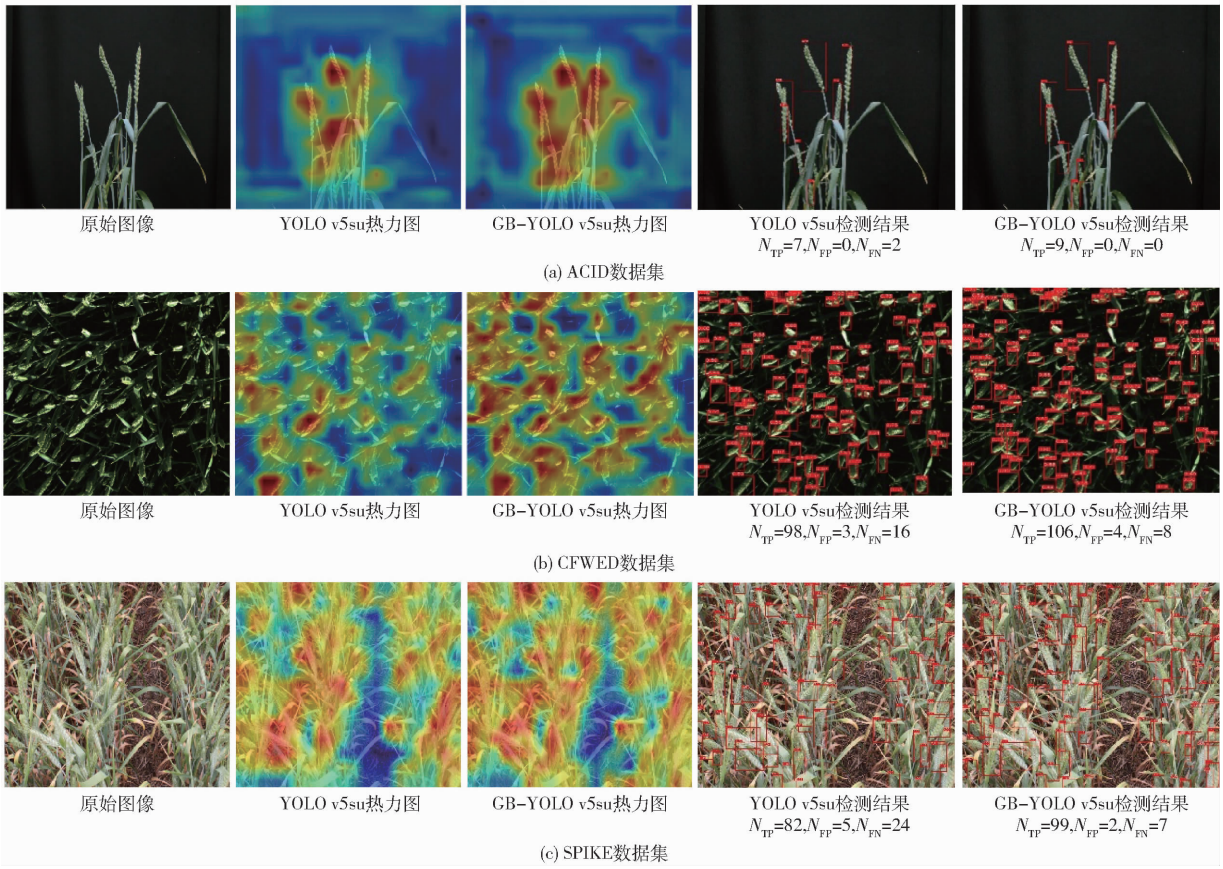


图 8　不同数据集上的热力图可视化与检测性能对比

Fig. 8　Heatmap visualization and detection performance comparison across different datasets

务中存在的目标尺度多变、遮挡严重等难题，提出了轻量化改进检测算法 GB－YOLO v5su。通过集成 GhostNet 轻量化主干、BiFPN 多尺度融合机制、P2 高分辨率检测层与 MPDIoU 损失函数，在显著降低模型参数量的同时，保持了高检测精度。

(2)在 GWHD2021 数据集上的实验结果表明，GB－YOLO v5su 在参数量仅为 5.38×10^6 的情况下，AP50 达到 92.8%，优于多数主流轻量模型。消融实验进一步验证了各改进模块的有效性与协同作用。此外，在 ACID、CFWED、SPIKE 等多个公开数据集上的鲁棒性测试表明，本文算法具有良好的泛化能力与场景适应性。

参　考　文　献

[1] 鲍文霞，苏彪彪，胡根生，等. 基于 FE－P2Pnet 的无人机小麦图像麦穗计数方法[J]. 农业机械学报，2024，55(4)：155－164，289.
BAO Wenxia，SU Biaobiao，HU Gensheng，et al. Method for counting wheat ears in UAV images based on FE－P2Pnet[J]. Transactions of the Chinese Society for Agricultural Machinery，2024，55(4)：155－164，289.（in Chinese）

[2] 刘涛，孙成明，王力坚，等. 基于图像处理技术的大田麦穗计数[J]. 农业机械学报，2014，45(2)：282－290.
LIU Tao，SUN Chengming，WANG Lijian，et al. In-field wheatear counting based on image processing technology[J]. Transactions of the Chinese Society for Agricultural Machinery，2014，45(2)：282－290.（in Chinese）

[3] 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）

[4] POUND M P，ATKINSON J A，WELLS D M，et al. Deep learning for multi-task plant phenotyping[C]// IEEE International Conference on Computer Vision Workshops（ICCVW），2017：2055－2063.

[5] ZHU J P，YANG G F，FENG X P，et al. Detecting wheat heads from UAV low-altitude remote sensing images using deep learning based on transformer[J]. Remote Sensing，2022，14(20)：5141.

[6] 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.

[7] WANG Donglin，SHI Longfei，YIN Huiqing，et al. A detection approach for wheat spike recognition and counting based on UAV images and improved Faster R－CNN[J]. Plants，2025，14(16)：2475.

[8] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: unified, real-time object detection[C]// IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016: 779 - 788.

[9] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need[C]// Advances in Neural Information Processing Systems, 2017: 5999 - 6009.

[10] 苑迎春,耿俊,许楠,等. 基于 YOLO v8 - TRP 模型的小麦麦穗识别方法研究[J]. 农业机械学报, 2025, 56(11): 499 - 508.
YUAN Yingchun, GENG Jun, XU Nan, et al. Wheat ear recognition method using YOLO v8 - TRP model[J]. Transactions of the Chinese Society for Agricultural Machinery, 2025, 56(11): 499 - 508. (in Chinese)

[11] 陈森,徐伟峰,王洪涛,等. 基于改进 YOLOv7 的麦穗检测算法[J]. 吉林大学学报(理学版), 2024, 62(4): 886 - 894.
CHEN Sen, XU Weifeng, WANG Hongtao, et al. Wheat ear detection algorithm based on improved YOLOv7[J]. Journal of Jilin University (Science Edition), 2024, 62(4): 886 - 894. (in Chinese)

[12] MENG X, LI C, LI J. YOLOv7 - MA: improved YOLOv7 based wheat head detection and counting[J]. Remote Sensing, 2023, 15(15): 3770.

[13] CARION N, MASSA F, SYNNAEVE G, et al. End-to-end object detection with transformers[C]// European Conference on Computer Vision (ECCV), 2020: 213 - 229.

[14] 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)

[15] LIN T Y, DOLLAR P, GIRSHICK R, et al. Feature pyramid networks for object detection[C]// Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR), 2017: 936 - 944.

[16] LIU S, QI L, QIN H, et al. Path aggregation network for instance segmentation[C]// Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR), 2018: 8759 - 8768.

[17] HAN K, WANG Y, TIAN Q, et al. GhostNet: more features from cheap operations[C]// Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR), 2020: 1577 - 1586.

[18] TAN M, PANG R, LE Q V. EfficientDet: scalable and efficient object detection[J]. arXiv Preprint, arXiv: 1911.09070, 2020.

[19] 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, 2020: 3521852.

[20] 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.

[21] VIJAYAKUMAR A, VAIRAVASUNDARAM S. YOLO-based object detection models: a review and its applications[J]. Multimedia Tools and Applications, 2024, 83: 83535 - 83573.

[22] LAZAREVICH I, GRIMALDI M, KUMAR R, et al. YOLOBench: benchmarking efficient object detectors on embedded systems[C]// IEEE/CVF International Conference on Computer Vision Workshops (ICCVW), 2023: 1161 - 1170.

[23] DOHERTY J, GARDINER B, KERR E, et al. BiFPN - YOLO: one-stage object detection integrating bi-directional feature pyramid networks[J]. Pattern Recognition, 2025, 160: 111209.

[24] MA S, XU Y. MPDIoU: a loss for efficient and accurate bounding box regression[J]. arXiv Preprint, arXiv: 2307.07662, 2023.

[25] DANDRIFOSSE S, ENNADIFI E, CARLIER A, et al. Deep learning for wheat ear segmentation and ear density measurement: from heading to maturity[J]. Computers and Electronics in Agriculture, 2022, 199: 107161.

[26] HASAN M M, CHOPIN J P, LAGA H, et al. Detection and analysis of wheat spikes using convolutional neural networks[J]. Plant Methods, 2018, 14: 1 - 13.