

网络出版时间:2025-07-09

网络出版地址:https://link.cnki.net/urlid/61.1359.S.20250709.1153.006

基于无人机视频流的小麦穗数精准监测研究

韩桐鹤^{1,2}, 张博涵^{1,2}, 费帅鹏², 李雷², 孙海艳², 王多霞², 孟亚雄¹, 肖永贵²

(1. 甘肃农业大学农学院, 甘肃兰州 730070; 2. 中国农业科学院作物科学研究所, 北京 100081)

摘要:为了解利用深度学习技术实现小麦穗数高效精准测定的可行性,选取黄淮冬麦区10个小麦主栽品种(泛麦8号、周麦36、中麦895、马兰1号、新麦26、豫麦49、济麦22、中麦578、郑麦1860和中麦255),设置120万、240万和360万株·hm⁻²三个种植密度,采用YOLO v5、YOLO v6、YOLO v8和YOLO v10四种深度学习算法,构建通过视频流实时识别计数小麦穗数的模型,并以小麦品种中麦578的生产田群体(300万株·hm⁻²)进行预测精度验证。结果表明,不同模型损失函数起点和下降幅度均存在差异,各模型对麦穗的检测性能均随迭代次数的增加而提升。在训练时间和检测效率方面,YOLO v6和YOLO v10模型速度较快,但检测性能和准确率均低于YOLO v5和YOLO v8。YOLO v5和YOLO v8模型耗时相对较长,但YOLO v8模型的召回率、F1分数、平均精度及准确率表现最佳,分别为90.90%、93.00%、97.20%和88.00%。YOLO v8模型识别穗数与真实计数的相关性随种植密度增加而降低,在三个种植密度下 r^2 分别为0.92、0.81和0.79。经生产田群体验证,综合考虑 r^2 、RMSE和MAE值,YOLO v8模型相对其他模型在小麦灌浆各阶段的穗数识别中表现最稳定,识别穗数精度最高;在视频流中YOLO v8模型检测穗数的 r^2 最高,达到0.90。这说明YOLO v8模型在小麦群体穗数识别方面具有较高的可靠性,受种植密度、品种和穗数发育阶段影响较小,在复杂田间环境下同样表现出较好的鲁棒性和高效性,适用于产量预测、育种研究、栽培管理等场景。

关键词: 麦穗计数; 视频流; 深度学习; 无人机; YOLO

中图分类号: S512.1; S314

文献标识码: A

文章编号: 1009-1041(2026)02-0264-12

Research on Accurate Monitoring of Wheat Ear Count Based on UAV Video Streams

HAN Tonghe^{1,2}, ZHANG Bohan^{1,2}, FEI Shuaipeng², LI Lei², SUN Haiyan², WANG Duoxia²,
MENG Yaxiong¹, XIAO Yonggui²

(1. College of Agronomy, Gansu Agricultural University, Lanzhou, Gansu 730070, China;

2. Institute of Crop Sciences, Chinese Academy of Agricultural Sciences, Beijing 100081, China)

Abstract: To evaluate the feasibility of using deep learning for efficient and accurate wheat spike counting, ten major winter wheat cultivars from the Huang-Huai wheat region (Fanmai 8, Zhoumai 36, Zhongmai 895, Malan 1, Xinmai 26, Yumai 49, Jimai 22, Zhongmai 578, Zhengmai 1860, and Zhongmai 255) were selected as materials. Three planting densities (1.2 million, 2.4 million, and 3.6 million plants · hm⁻²) were tested, and four deep learning algorithms (YOLO v5, YOLO v6, YOLO v8, and YOLO v10) were employed to construct real-time video-based wheat spike detection and counting models. The models were validated using field-grown Zhongmai 578 populations (3 million plants · hm⁻²). The results showed that the initial loss functions and convergence rates varied among models, with all models improving in spike detection performance as iterations increased. In terms of

收稿日期: 2025-03-28

修回日期: 2025-05-06

基金项目: 科技创新 2030—“新一代人工智能”重大项目(2022ZD0115703); 国家自然科学基金项目(32372196); 科技创新工程—小麦稳产保供应急技术支撑项目(S2024XC08)

第一作者 E-mail: hantonghe2025@163.com (韩桐鹤)

通讯作者 E-mail: yxmeng1@163.com (孟亚雄); xiaoyonggui@caas.cn (肖永贵)

training speed and inference efficiency, YOLO v6 and YOLO v10 were faster but exhibited lower detection accuracy compared to YOLO v5 and YOLO v8. Although YOLO v5 and YOLO v8 required longer processing time, YOLO v8 achieved the best performance in recall(90.90%), F1-score(93.00%), mean average precision(97.20%), and overall accuracy(88.00%). The correlation(r^2) between YOLO v8's spike counts and manual counts decreased with increasing planting density, yielding values of 0.92, 0.81, and 0.79 for the three densities, respectively. Field validation demonstrated that YOLO v8 outperformed other models in stability and precision across different grain-filling stages, with the highest r^2 (0.90) for real-time video-based spike counting. The model also showed robustness against variations in planting density, cultivar, and spike growth stage, maintaining high performance in complex field environments. These findings suggest that YOLO v8 is a reliable algorithm for wheat spike counting, suitable for yield prediction, breeding, and cultivation management applications.

Keywords: Wheat spike counting; Video streams; Deep learning; Unmanned Aerial Vehicle(UAV); YOLO

小麦是中国第三大粮食作物^[1]。由于人口增长和气候变化的影响,市场对粮食的需求呈持续刚性增长态势,因而高产稳产依然是小麦育种的最主要目标。穗数是小麦产量构成因素之一^[2-3],与籽粒产量呈显著正相关,在小麦遗传育种与产量估测研究中尤为重要。传统小麦估产主要依靠人工田间计数麦穗,需要大量人力、物力和时间,尤其是在大区域尺度下传统估产方法最易受人为等因素限制,进而影响结果的准确性和工作效率^[4-6]。因此,开发一种能够快速、准确识别小麦穗数的方法对于小麦育种和产量评估工作具有重要意义^[7]。

虽然结合麦穗颜色、纹理和形态特征的计算机视觉和机器学习算法能够对麦穗进行计数,但该方法的泛化性能较弱^[8-9]。在不同种植环境和不同密度条件下,对田间麦穗进行实时检测时,算法需经过特定调整以适应复杂现场条件。通常做法是优化算法参数、改进模型结构以及增强对环境适应性。随着计算机性能的提高,基于深度学习的麦穗图像识别算法得到了广泛应用^[10]。麦穗的检测和计数可通过图像分割^[11-12]和目标检测^[13-15]实现。目标检测算法又可分为单阶段目标检测算法和两阶段目标检测算法。其中,单阶段目标检测算法主要通过目标检测算法直接从图像中探测目标的位置和类别,无需生成候选框,从而显著提高检测速度,如 SSD(single shot multibox detector)^[16]和 YOLO(you only look once)系列算法^[17-18],因其快速、高效的检测性能和易于部署的特点,在农业领域得到广泛应用;两阶段目标检

测算法,如 Fast R-CNN^[19-20]和 SPP-net^[21-22]模型,其通过引入区域候选网络(region proposal network,RPN)生成候选区域,虽然在检测精度上有所提高,但速度相对较慢。为了解决这一问题,专家提出了多种改进策略,如动态颜色变换网络通过调整图像的颜色通道,减少了假阴性,提高了检测准确性^[23-24]。

前人研究发现,在计算麦穗数量时可将 CBAM(convolutional block attention module)与 YOLO v5 融合为一种复合网络结构,该结构在麦穗检测任务上可以达到 90.40%的 F1 分数和 90.20%的平均精度^[25]。通过在 YOLO v7x 框架中引入 SPD(space to depth)模块,并结合 NWD(normalized gaussian wasserstein distance)损失函数,构建出改进的检测模型 YOLO v7xSPD,可显著提升视频流中麦穗计数的效率和准确性^[26]。视频流检测小麦穗数相较于图片检测,能够提供连续实时数据,有效捕捉动态变化,减少因静态图像导致的漏检,对于大面积农田检测效率更高,可为农业管理提供强有力的技术支持^[27-28]。

为了解单阶段目标检测算法 YOLO v5、YOLO v6、YOLO v8 和 YOLO v10 在小麦不同种植密度图像数据集下的麦穗检测性能表现,本研究建立了能够准确识别麦穗的深度学习模型,确保即使在环境光线变化、种植密度不一的情况下也能保持高准确度,重点分析各模型间监测精度差异、模型泛化能力的提升策略,探索视频流技术在小麦穗数识别计数中的实用性,为小麦田间生长发育检测、估产、品种鉴定等提供技术支持。

1 材料与方法

1.1 试验区概况

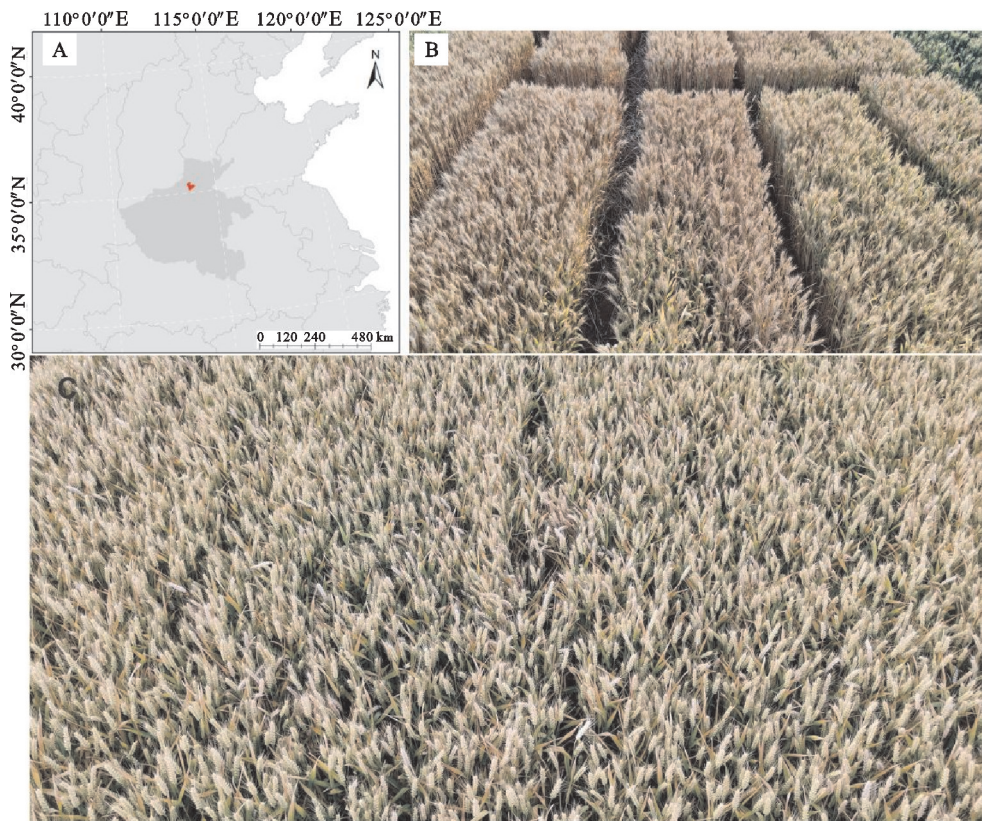
试验采用品种和种植密度双因素随机区组设计。供试小麦材料为泛麦 8 号、周麦 36、中麦 895、马兰 1 号、新麦 26、豫麦 49、济麦 22、中麦 578、郑麦 1860 和中麦 255 等 10 个当前黄淮海区主栽品种,包含不同穗型,具有多样性和代表性。种植密度设置 120 万、240 万和 360 万株· hm^{-2} 3 个水平。2 个重复,小区长 6 m,宽 1.20 m,面积 7.20 m^2 。小麦种植管理同当地生产田。验证群体来自小麦品种中麦 578 的生产田,种植密度为 300 万株· hm^{-2} 。试验材料均于 2023—2024 年生长季种植在河南省新乡市新乡县七里营镇中国农业科学院作物科学研究所新乡综合试验基地(113°45′40″E, 35°8′11″N)(图 1)。

1.2 技术路线

按图 2 所示流程采集不同种植密度小区和生产田无人机视频流数据,并结合四种前沿的深度学习算法,对小麦穗数进行预测研究。

1.3 视频流数据采集

小区图像数据采集时间为 2024 年 5 月 22 日,小麦正值灌浆后期;生产田图像数据采集时间为 2024 年 5 月 12 日、5 月 17 日、5 月 22 日,小麦正值灌浆前期、中期与后期。数据采集选择晴朗的早晨(约 7 点到 9 点之间)进行,以避免因光照强度过高而造成干扰。利用消费级无人机大疆 DJI Air 3 搭载 4 800 万像素的摄像头来捕获视频流,其分辨率为 3 840×2 160,帧率为 60 FPS。为获得小麦冠层的详细视频,镜头倾斜 37°,手动控制飞机在麦穗冠层上方 3 m,以速度 0.50~1 $\text{m} \cdot \text{s}^{-1}$ 飞行。此外,为进一步验证模型在复杂环境下的可靠性,采集了生产田麦穗视频用于评估不同模型在视频流识别方面的效果,并在测试集中加入了灌浆前期(2024 年 5 月 12 日)与灌浆中期(2024 年 5 月 17 日)数据。这些数据能够更全面地评估模型在不同生育时期的性能,确保其在复杂环境下的准确性和稳定性。按照当前国际通用方式,本研究小麦穗数识别采用视频与图片中的麦穗由两人单独计数后取平均值作为真实值^[20,26,29]。

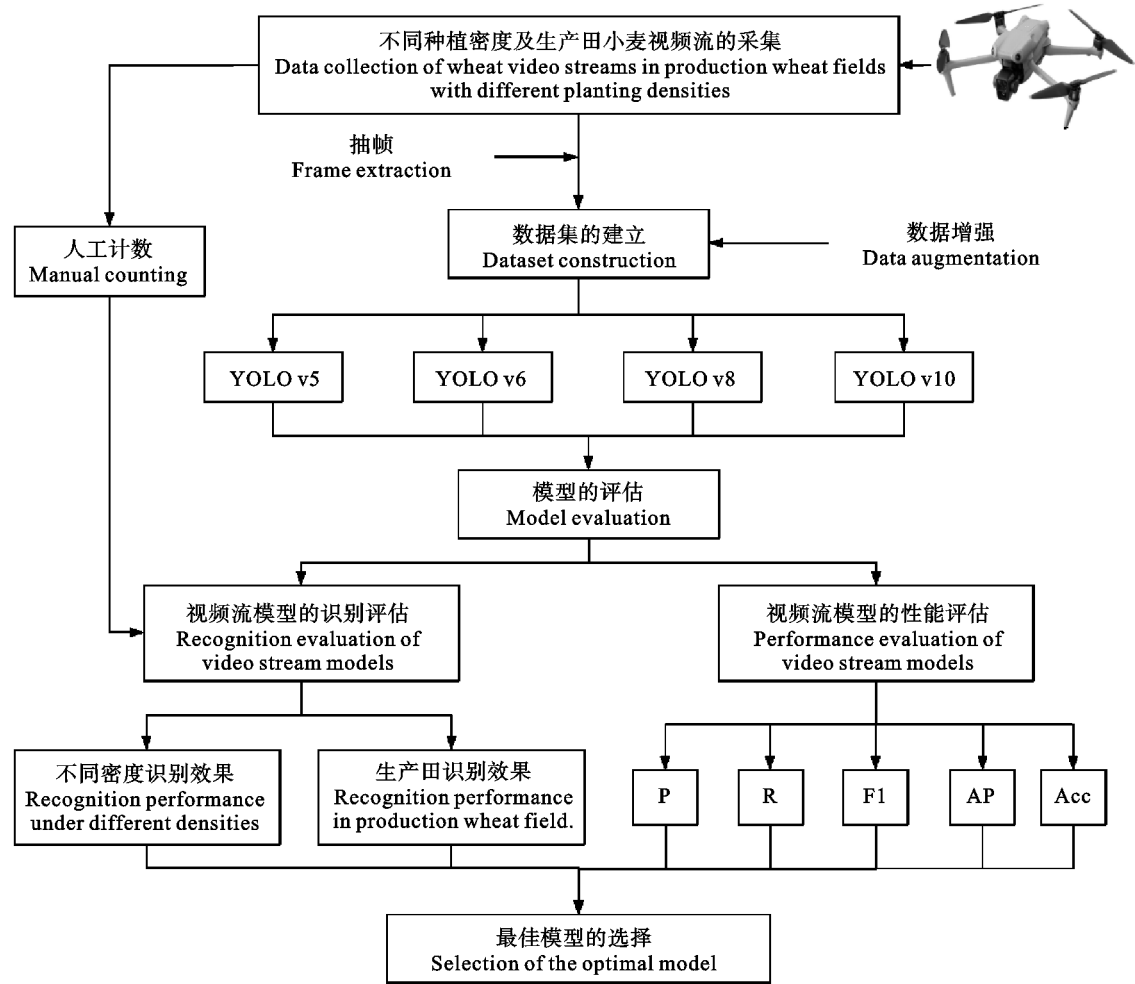


A: 研究区位置; B: 不同种植密度小区; C: 生产田。

A: Location of the study area; B: Experimental plots with different planting densities; C: Wheat field.

图 1 小麦种植区地理位置与 UAV 视频采集

Fig. 1 Geographical location of wheat planting area and UAV video acquisition



P: 精确率;R: 召回率;F1: F1 分数;AP: 平均精度;Acc: 准确率。
P: Precision; R: Recall; F1: F1 Score; AP: Average precision; Acc: Accuracy.

图 2 技术路线图
Fig. 2 Technical roadmap

1.4 数据集创建

将灌浆后期不同种植密度的图像数据用于模型训练。由于麦穗在视野范围内太密集,考虑到数据存储空间和处理效率,本研究将原始视频(分辨率为 3 840×2 160)裁剪为分辨率为 450×450 的小视频。为了降低抽帧时图像的重叠率,每 45 帧提取 1 帧,共得到 528 幅图像并用于模型训练。模型检结果及人工计数结果均以单幅图像中的麦穗数为单位。每幅子图像(450×450 像素)覆盖的实际田间面积取决于原始视频中的位置,范围为 0.12~0.24 m²。使用 labellmg 对数据集进行标注,图 3 为原始图像与标注图像对比图。此外,为了提高模型的泛化性能,减少小麦冠层受到光照条件、天气变化和风速等因素影响,采用水平翻转、垂直翻转、水平垂直翻转、高斯模糊、增减对比度 6 种方案对图像进行数据增强,共获得 3 168 张

图像数据。图像数据增强效果如图 4 所示。使用 Python 语言编写的脚本对增强后的图像进行注释和转换,以获得其相应的注释文件。按照 8:1:1 的比例将数据集分为训练集、测试集和验证集。

1.5 麦穗检测算法原理

YOLO v5 是一个基于 PyTorch 的深度学习算法,主要通过在输入端实施 Mosaic 数据增强策略,以提升模型对多变场景的泛化能力。此外,YOLO v5 引入了自适应锚框计算策略,优化了对不同数据集中目标尺寸和形状的适应性,进而提高检测精度。为了增强多目标的实时检测与跟踪性能,研究者在 YOLO v5 中集成了 ReID(re-identification)特征识别模块,加快了跟踪速度,在实现输出目标位置的同时,也能提供目标特征信息^[11,25]。

YOLO v6 由美团视觉智能部研发,专为工业应用设计,具有高精度和快速的推理效率。相较

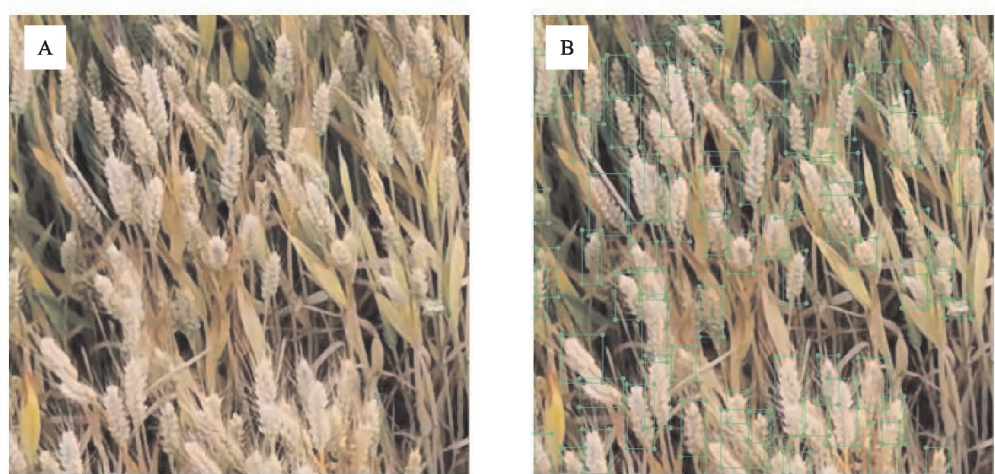
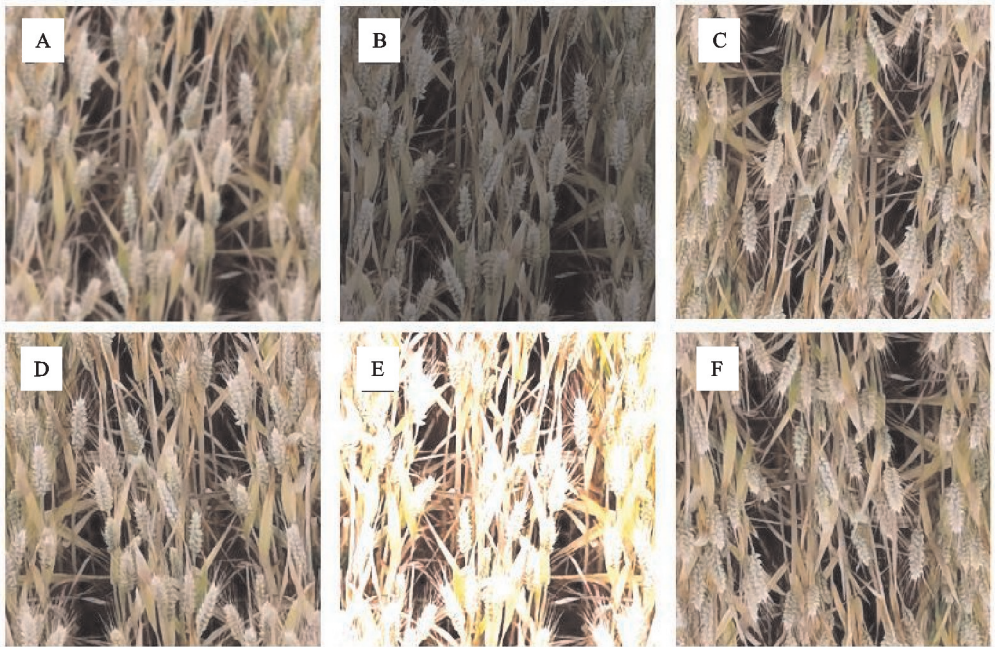


图 3 原始图像(A)与标注图像(B)的对比

Fig. 3 Comparison between the original image(A) and the annotated image(B)



A: 高斯模糊; B: 减少对比度; C: 水平垂直翻转; D: 水平翻转; E: 增加对比度; F: 垂直翻转。
A: Gaussian blur; B: Decrease contrast; C: Horizontal and vertical flip; D: Horizontal flip; E: Increase contrast; F: Vertical flip.

图 4 数据增强效果图

Fig. 4 Data augmentation effect diagram

于 YOLO v5, 采用了 Hardware-friendly 的骨干网络设计, 使得模型更加适合在不同硬件平台上运行, 其更简洁高效的检测有助于提升模型性能, 并引入双向串联模块(bidirectional series module), 可以增强模型的特征提取能力^[30]。

YOLOv8 支持多种视觉任务, 如图像分类、物体检测和实例分割, 具有多功能性和灵活性, 在保持高精度的同时提高了处理速度, 实现了速度与精度的平衡。其提供了从 nano、small、medi-

um、large 至 x-large 多种规模模型, 以适应不同需求。在骨干网络架构上, YOLO v8 采用了 C2f 结构, 取代了 YOLO v5 中的 C3 结构, 并为不同模型定制了通道数, 显著提升了性能^[31-33]。

YOLO v10 在 YOLO v8 的基础上进行了优化, 是 YOLO 系列的最新迭代, 由清华大学研发团队开发, 旨在实现端到端的高性能目标检测。特别是在非极大值抑制(non-maximum suppression)训练方面取得了突破, 减少推理延迟并提高

了部署效率。YOLO v10 还引入了空间一通道解耦下采样技术,提升了特征提取效率和准确性。此外,YOLO v10 提供了多种算法,以适应从资源受限环境到追求高精度的不同应用需求^[34]。

1.6 模型训练

模型采用 Pytorch v2.40 深度学习框架,将训练集和验证集分为 16 个批次,遍历所有批次后完成 1 次迭代。优化器选择 SGD(stochastic gradient descent),设置初始学习率为 0.01,共进行 150 次迭代。随着迭代次数的提升,降低学习率至 0.000 1。

1.7 评价指标

1.7.1 不同模型性能的评估指标

为更加全面地衡量模型的识别性能,采用精准率(precision,Pr)、召回率(recall,Re)、F1 分数(F1 score,F1)、平均精度(average precision,AP)和准确率(accuracy,Acc)对模型进行评价,指标在 0~1 范围内取值,其值越大,评估效果越好。其计算公式如下,TP 代表真阳性,FP 代表假阳性,FN 代表假阴性,TN 代表真阴性,AP 是 PR 曲线中曲线和坐标轴包围的区域的面积。

$$\begin{aligned} Pr &= \frac{TP}{TP+FP}; Re = \frac{TP}{TP+FN}; \\ F1 &= \frac{2(Pr \times Re)}{Pr+Re}; AP = \int_0^1 Pr(Re) dRe \\ Acc &= \frac{TP+TN+FP+FN}{TP+TN} \end{aligned}$$

1.7.2 不同模型识别效果的评估指标

模型计数与田间真实值之间的相关性决定系数(coefficient of determination, r^2)来评估,均方根误差(root mean square error, RMSE)用于评估模型计数和田间真值数之间的差异程度,平均绝对误差(mean absolute error, MAE)为衡量预测值与真实值之间绝对误差的平均值。

$$\begin{aligned} r^2 &= 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y}_i)^2} \\ RMSE &= \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}} \\ MAE &= \frac{\sum_{i=1}^n |y_i - \hat{y}_i|}{n} \end{aligned}$$

各定义式如下,其中, n 为图像个数, y_i 表示人工统计到的地面真实值, $\hat{y}_i$ 表示模型计数结果, $\bar{y}_i$ 表示真实值的平均值, $|y_i - \hat{y}_i|$ 为地面真实值与模型计数结果之间的绝对误差。

2 结果与分析

2.1 不同模型训练损失函数及性能提升分析

在目标检测模型的训练过程中,损失函数能够量化模型预测结果与实际结果之间的差异,从而指导模型通过优化过程提高预测准确性。从 YOLO v5、YOLO v6、YOLO v8 和 YOLO v10 四种模型的训练过程(图 5)看,随着迭代次数的增加,各模型的边界框损失(Train/box_loss)、分类损失(Train/cls_loss)和分布焦点损失(Train/dfl_loss)均逐渐降低并趋于平稳,清晰地表明模型在训练过程中在不断优化并最终收敛。尽管各模型损失函数的起始点和下降幅度存在差异,但所有模型对麦穗的检测性能经过迭代均有所提升。

2.2 不同麦穗检测模型的综合性能评价

为了确保模型可靠性,所有模型使用相同数据集进行训练。从训练时间来看,YOLOv6 训练时间最短,仅需 6.37 h,而 YOLOv8 训练时间最长,为 24.37 h,是 YOLO v6 的 3.83 倍(表 1)。在检测效率方面,YOLO v6 的每秒传输帧数最高,为 13.04 帧·s⁻¹,推理时间为 130.60 ms·张⁻¹,预处理时间和输出时间分别为 6.80 和 17.90 ms·张⁻¹,检测效率高于其他模型;YOLO v10 检测效率次之;YOLO v5 检测效率与 YOLO v8 接近,均较低。从训练结果(表 2)看,虽然 YOLO v10 的精确率最高,为 96.60%,但其召回率(72.70%)、F1 分数(82.00%)、平均精度(88.60%)以及准确率(69.80%)均低于 YOLO v8,YOLO v6 各项与之相近。YOLO v8 在召回率(90.90%)、F1 分数(93.00%)、平均精度(97.20%)和准确率(88.00%)上均表现最好,说明其综合检测性能最佳。YOLO v5 的精确率、召回率、F1 分数、平均精度和准确率分别为 92.10%、90.10%、92.00%、96.60%和 85.70%,均略低于 YOLO v8。尽管 YOLO v6 和 YOLO v10 在时间消耗方面具有优势,但检测性能和准确率均低于 YOLO v5 和 YOLO v8。综合各方面表现,YOLO v8 虽相对耗时,但检测性能均衡,整体准确性较高,可作为最佳模型。

2.3 不同种植密度下模型识别麦穗精度验证分析

在不同种植密度条件下,四种模型对小区麦穗数量识别的情况各有不同。随着种植密度的增加,整体上各模型的识别穗数与真实穗数的相关性(r^2)均有所下降,误差(RMSE 和 MAE)则相应增加(图 6)。在低种植密度(120 万株·hm⁻²)条件下,各模型的预测精度相对接近。其中,YOLO v8

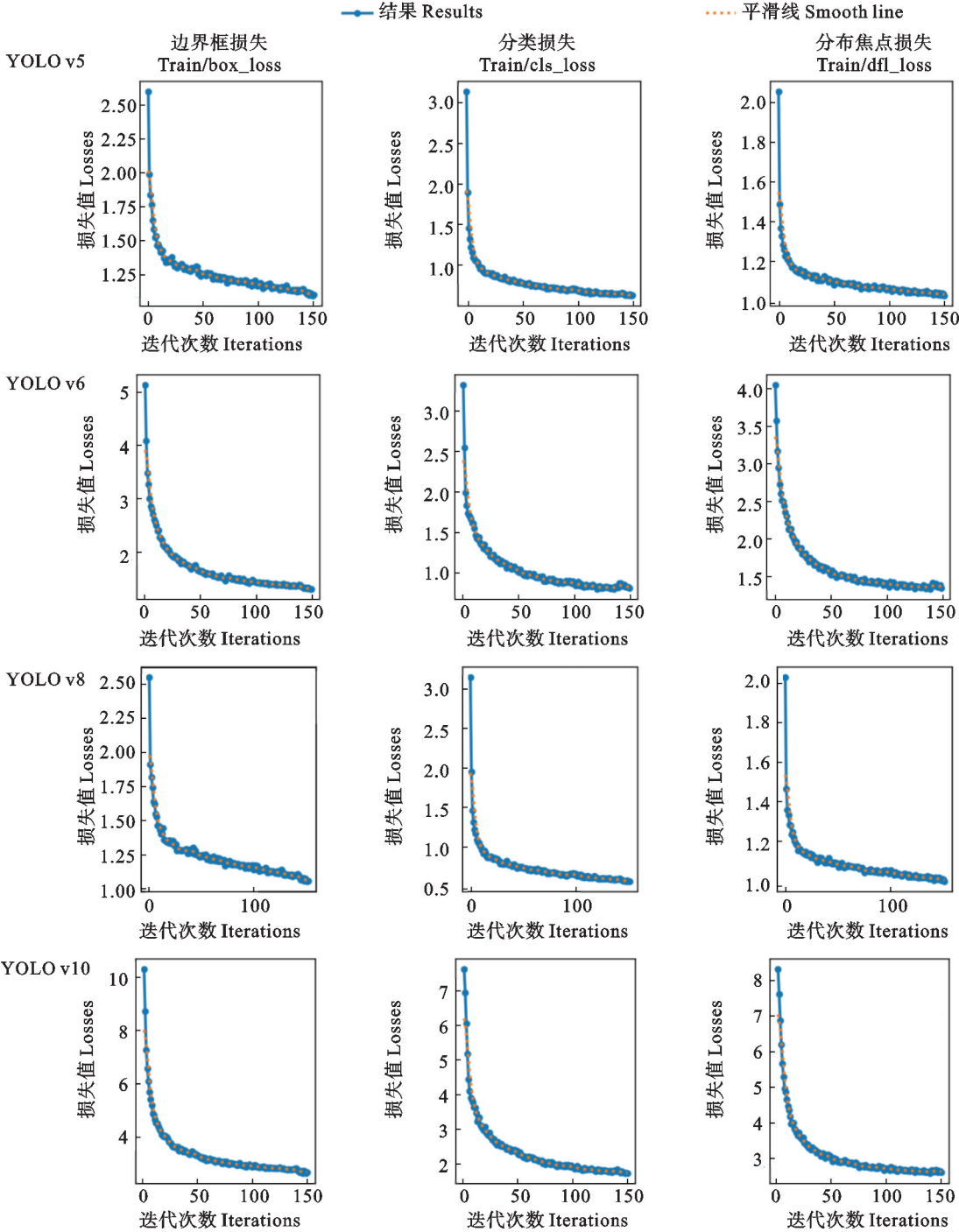


图 5 不同模型训练损失变化趋势

Fig. 5 Training loss variation trends of different models

表 1 不同模型的训练时间与检测效率对比

Table 1 Comparison of training time and detection efficiency among different models

模型名称 Model name	训练时间 Training time/h	预处理时间 Preprocessing time/ (ms · image ⁻¹)	推理时间 Inference time/ (ms · image ⁻¹)	输出时间 Output time/ (ms · image ⁻¹)	每秒传输帧数 Frames per second	每秒浮点运算次数 Floating-point operations per second
YOLO v5	11. 15	5. 80	146. 40	19. 50	12. 80	95. 97G
YOLO v6	6. 37	6. 80	130. 60	17. 90	13. 04	97. 83G
YOLO v8	24. 37	6. 20	144. 80	22. 20	11. 95	89. 59G
YOLO v10	8. 57	5. 70	136. 90	0. 60	10. 07	75. 53G

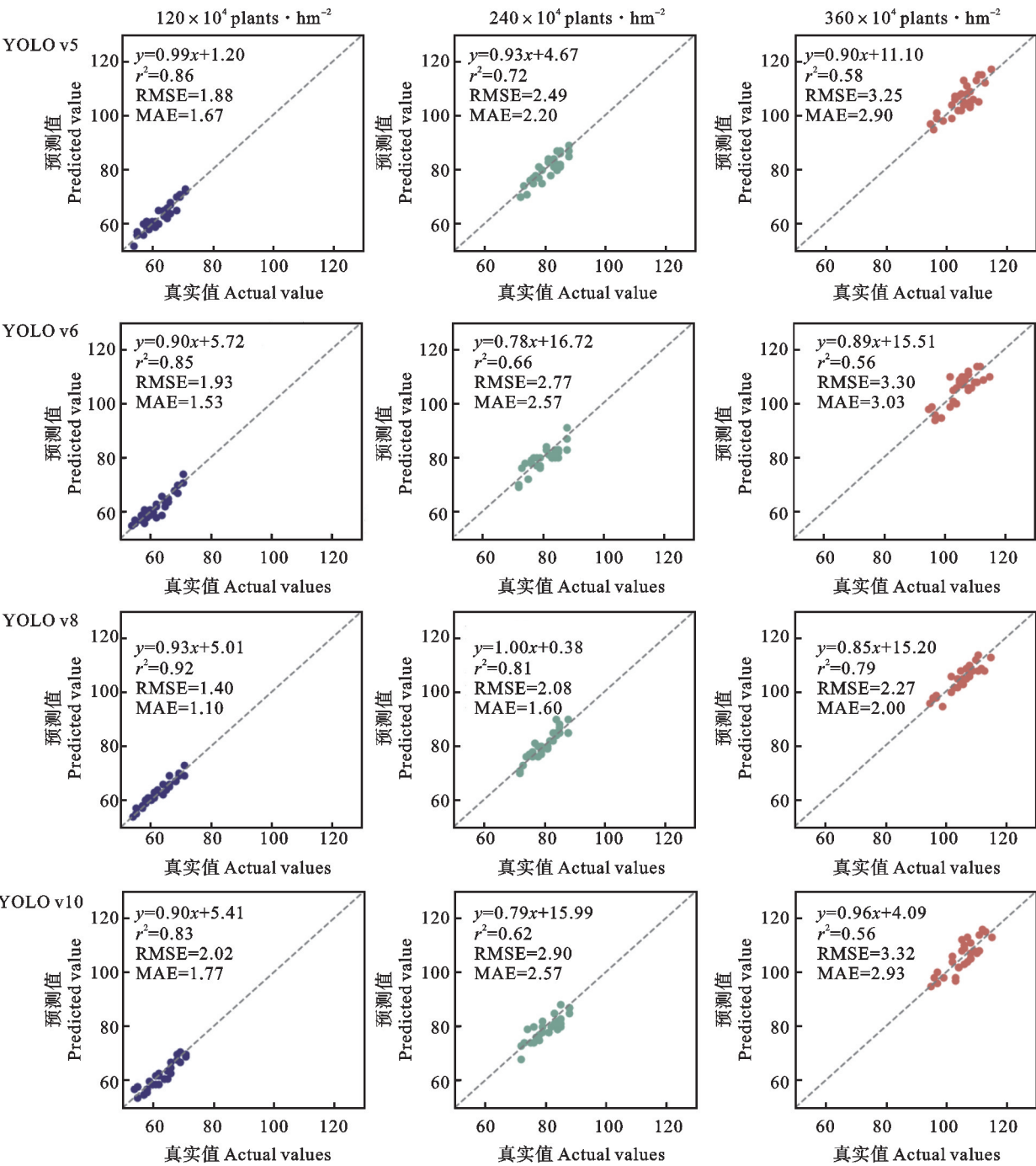
每秒浮点运算次数下的“G”表示十亿次(10⁹)浮点运算。

“G” in floating-point operations per second(FLOPS) denotes one billion(10⁹) floating-point operations.

表 2 不同模型训练结果

Table 2 Training results of different models

模型名称 Modelname	精确率 Precision/%	召回率 Recall/%	F1 分数 F1 score/%	平均精度 Average precision/%	准确率 Accuracy/%
YOLO v5	92.10	90.10	92.00	96.60	85.70
YOLO v6	96.10	80.30	89.00	93.80	80.90
YOLO v8	93.10	90.90	93.00	97.20	88.00
YOLO v10	96.60	72.70	82.00	88.60	69.80



预测值,真实值, RMSE 和 MAE 的单位均为穗 · 幅⁻¹。图 7 和图 8 同。

The units of the predicted values, actual values, RMSE and MAE are spikes · image⁻¹. The same in figure 7 and figure 8.

图 6 不同种植密度下小麦穗数预测值与真实值相关性

Fig. 6 Correlation between the predicted and actual values of wheat spikes under different planting densities

具有最高的拟合优度($r^2=0.92$),MAE 和 RMSE 分别为 1.10 和 1.40 穗,其预测精确度高于 YOLO v5、YOLO v6 和 YOLO v10。当种植密度提升至 240 万株 $\cdot\text{hm}^{-2}$ 时,YOLO v8 的拟合优度仍较高($r^2=0.81$),而其他模型的 r^2 值均有所降低。当种植密度达到 360 万株 $\cdot\text{hm}^{-2}$ 时,YOLO v8 的 r^2 值为 0.79,依然优于其他模型,其他模型的 r^2 值进一步降低。

通过统计分析,在种植密度为 120 万株 $\cdot\text{hm}^{-2}$ 时,YOLO v8 的预测性能与 YOLO v5 和 YOLO v10 差异分别达到 0.05 和 0.01 显著水平;在种植密度为 240 万株 $\cdot\text{hm}^{-2}$ 时,YOLO v8 的预测性能与 YOLO v6 和 YOLO v10 差异均达到 0.01 显著水平;在种植密度为 360 万株 $\cdot\text{hm}^{-2}$ 时,YOLO v8 的预测性能与 YOLO v5、YOLO v6 和 YOLO v10 差异均达到 0.01 显著水平(表 3)。这进一步印证了 YOLO v8 在高种植密度条件下对麦穗表现出较高的识别能力。

表 3 不同种植密度下模型的预测性能比较

Table 3 Comparison of model prediction performance under different planting densities

种植密度 Planting density/ ($\times 10^4 \cdot \text{hm}^{-2}$)	模型组合 Model combination	<i>t</i>	<i>P</i>
120	YOLO v5-YOLO v6	0.49	0.63
	YOLO v5-YOLO v8	2.48	0.02
	YOLO v5-YOLO v10	-0.41	0.68
	YOLO v6-YOLO v8	1.60	0.12
	YOLO v6-YOLO v10	-0.82	0.42
	YOLO v8-YOLOv10	-2.73	0.01
240	YOLO v5-YOLO v6	-1.26	0.21
	YOLO v5-YOLO v8	1.83	0.07
	YOLO v5-YOLO v10	-1.10	0.28
	YOLO v6-YOLO v8	3.06	<0.01
	YOLO v6-YOLO v10	0.00	1.00
	YOLO v8-YOLO v10	-2.74	0.01
360	YOLO v5-YOLO v6	-0.37	0.72
	YOLO v5-YOLO v8	2.67	0.01
	YOLO v5-YOLO v10	-0.08	0.93
	YOLO v6-YOLO v8	3.31	<0.01
	YOLO v6-YOLO v10	0.27	0.79
	YOLO v8-YOLO v10	-2.68	0.01

2.4 模型识别麦穗精度的大田验证结果

利用生产田群体进行验证,结果(图 7)表明,

在灌浆前期、中期和后期,各模型的识别精度随着时间的变化而逐渐增加。在灌浆前期,YOLO v8 的 r^2 为 0.78,RMSE 和 MAE 值分别为 3.62 和 3.30 穗 $\cdot\text{幅}^{-1}$;而 YOLO v5、YOLO v6 和 YOLO v10 的 r^2 为 0.60~0.76,误差指标相对较高。在灌浆中期,YOLO v8 的 r^2 提升至 0.88,而 YOLO v5、YOLO v6 和 YOLO v10 的 r^2 分别为 0.83、0.74 和 0.75。到灌浆后期,YOLO v8 的 r^2 达到 0.91,MAE 和 RMSE 值分别为 3.51 和 3.10 穗 $\cdot\text{幅}^{-1}$,各项指标均优于其他模型。综合来看,YOLO v8 在小麦灌浆不同阶段的麦穗识别计数表现最为稳定和可靠,适合用于麦穗数量监测与评估。

由表 4 可知,YOLO v5、YOLO v6、YOLO v8、和 YOLO v10 四个模型在麦穗检测中都展现了各自的优势和特点。YOLO v6 和 YOLO v10 的整体结果相近,但 YOLO v10 的误检数量最低,YOLO v6 在漏检数量、准确识别数量和识别数量上优于 YOLO v10。YOLO v5 和 YOLO v8 的整体结果相近,且均优于 YOLO v6 和 YOLO v10,但 YOLO v8 的识别数量、漏检数量、误检数量和准确识别数量均优于 YOLO v5。综上所述,YOLO v8 无论是在检测覆盖率、准确性还是整体识别效率上展现出了较好性能。

2.5 生产田视频流下模型识别麦穗精度分析

生产田视频流下对比模型预测的穗数与生产田真实穗数,结果(图 8)表明,YOLO v5、YOLO v6、YOLO v8 和 YOLO v10 的 r^2 值分别为 0.84、0.74、0.90 和 0.72,RMSE 值分别为 13.05、14.97、10.00 和 15.88 穗 $\cdot\text{幅}^{-1}$,MAE 值分别为 9.70、12.95、7.80 和 12.70 穗 $\cdot\text{幅}^{-1}$,表明 YOLO v8 是所有模型中最佳的,具有最高的 r^2 值(0.90)与最低的 RMSE 值(10.00 穗 $\cdot\text{幅}^{-1}$)和 MAE 值(7.80 穗 $\cdot\text{幅}^{-1}$),说明它在小麦穗部计数任务中最为准确和稳定,YOLO v5 是个可替代的选择,YOLO v6 和 YOLO v10 预测误差较大。

3 讨论

基于无人机视频流的小麦穗数精准计数方法是一种能够准确、实时地在田间进行小麦穗计数的智能技术。本研究表明,YOLO v8 在动态视频流处理上具有明显优势,适用于多尺度和多宽高比的目标检测,尤其在密集目标排列的情况下,锚框机制进一步提升了检测性能^[31-32]。与 YOLO v5 相比,YOLO v8 在速度和准确性方面的优化使

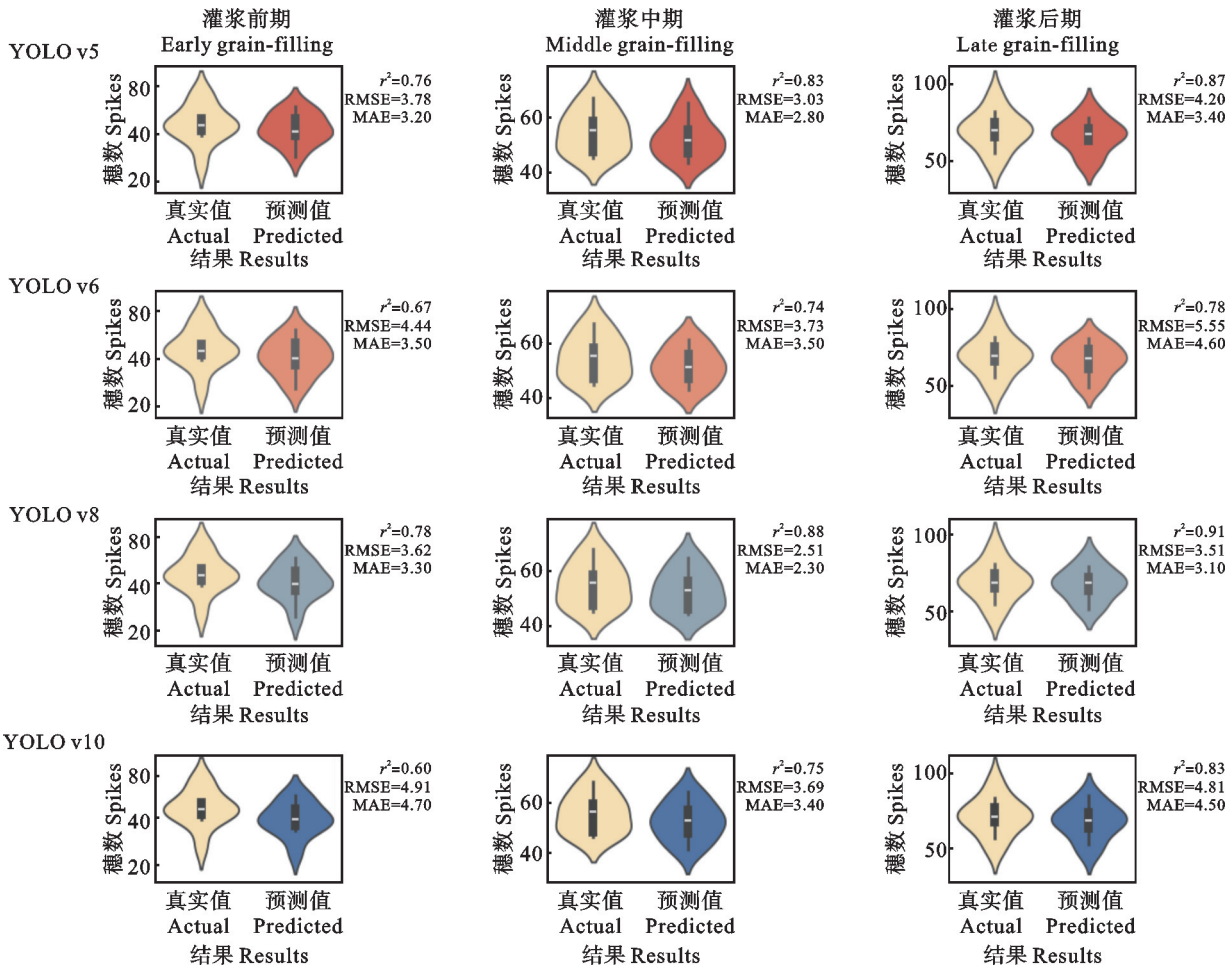


图 7 不同生育期下预测值与真实值对比

Fig. 7 Comparison of the predicted and actual values at different wheat growth stages

表 4 不同模型识别麦穗结果对比

模型 Model	穗数 Spike number/(spikes · image ⁻¹)				
	真实数量 True number	识别数量 Detected number	误检数量 False positive number	漏检数量 False negative number	准确识别数量 True positive number
YOLO v5	2 759	2 708	191	242	2 517
YOLO v6	2 759	2 386	122	495	2 264
YOLO v8	2 759	2 713	169	215	2 544
YOLO v10	2 759	2 114	76	721	2 038

其在农业实时监测领域展现出巨大的应用潜力^[33]。差异性显著分析的结果也证实,在不同种植密度条件下,YOLO v8 与其他模型的预测性能存在显著差异,其能够更好地在实际应用中发挥作用,展现出较强的稳定性和适应性。此外,尽管有研究表明,YOLO v8 的精确率略低于 YOLO v7xSPD(P 为 95.80%)^[26],但其通用性和易用性使其在更广泛的田间条件下具有更高的推广价值。此外,本研究进一步针对不同小麦穗型品种和种植密度进行试验,验证了模型在不同田间环境下的适应性和鲁棒性,证明在不同密度的种植

条件以及摄影角度方面,穗数识别效果存在显著差异^[35]。

随着无人机技术和深度学习算法的不断进步,基于深度学习的无人机图像处理技术在小麦穗计数中展现出显著优势,能够有效解决复杂田间环境下的检测问题^[5]。例如,轻量级密度估计模型^[36](FIDMT-GhostNet)和 OSWSDet(oriented and small wheat spike detection)方法^[5],通过结合方向敏感的特征提取模块,显著提升了小目标和密集排列场景下的检测精度,为算法优化提供了重要参考。同时,数据增强技术在提升深

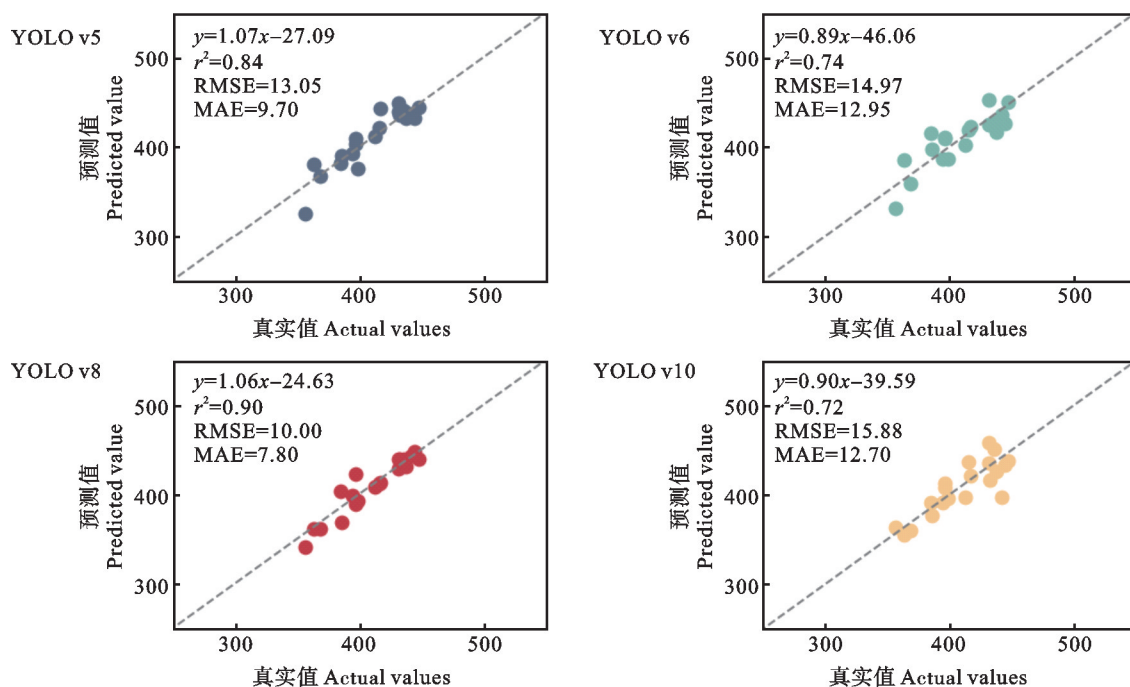


图8 生产田视频中小麦穗数预测值与真实值相关性

Fig. 8 Correlation between the predicted and actual values of spikes in wheat field videos

度学习模型的泛化能力方面发挥着关键作用,使其在不同光照和天气条件下保持较高的检测精度。这与 Misra 等^[37]的研究结果一致,证实了数据增强在提高模型鲁棒性中的重要性。

随着视频流技术的发展,实时监测与实际应用在农业管理中的重要性日益凸显。农业领域的深度学习应用正逐步从静态图像转向动态视频流^[26],这一转变虽然增加了数据处理的复杂性,但也为提高检测的实时性和准确性提供了新的可能性。例如,通过无人机视频流技术,农民可以实时监测作物生长状况,及时发现病虫害等问题,从而提高农业生产效率。此外,探索更高效的数据如旋转、缩放、颜色变换等增强技术,以适应复杂多变的田间环境,也是未来的研究重点,可以实现更高效、精准的作物管理。

4 结论

本研究通过无人机视频流结合 YOLO v5、YOLO v6、YOLO v8 和 YOLO v10 四种深度学习算法构建的麦穗检测模型,比较了不同模型在麦穗识别计数上的性能与识别效果,发现 YOLO v5 在小样本数据集上表现出较高的精度,但处理速度较慢,可能限制其在实时、大规模农业场景中的应用,YOLO v6 和 YOLO v10 虽在时间消耗

方面具有优势,但检测性能和准确率较低,漏检数量较高。YOLO v8 在召回率、F1 分数、平均精度及准确率上均表现最佳,尽管耗时相对较长,但其综合性能和整体准确性最佳,且与生产田种植材料进行验证分析,YOLO v8 性能在小麦灌浆各阶段的穗数识别中表现最稳定可靠,其准确识别穗数最多,漏检穗数最少,在视频流中检测穗数的 r^2 最高(0.90),适合用于田间麦穗数量监测与评估。

参考文献:

- [1]何中虎,庄巧生,程顺和,等.中国小麦产业发展与科技进步[J].农学学报,2018,8(1):99.
HE Z H, ZHUANG Q S, CHENG S H, et al. Wheat production and technology improvement in China [J]. *Journal of Agriculture*, 2018, 8(1): 99.
- [2]张胜全,王拯,杨卫兵,等.群体调控对杂交小麦源库系统的影响[J].华北农学报,2022,37(S1):74.
ZHANG S Q, WANG Z, YANG W B, et al. Effect of population regulation on source-sink system of hybrid wheat [J]. *Acta Agricultrae Boreali-Sinica*, 2022, 37(S1): 74.
- [3]李向东,吕凤荣,张德奇,等.小麦田间测产和实际产量转换系数实证研究[J].麦类作物学报,2016,36(1):69.
LI X D, LÜ F R, ZHANG D Q, et al. Empirical study on conversion factors between yield monitoring and actual yield of winter wheat [J]. *Journal of Triticeae Crops*, 2016, 36(1): 69.
- [4]JIN X L, YANG W N, DOONAN J H, et al. Crop phenotyping studies with application to crop monitoring [J]. *The Crop Journal*, 2022, 10(5): 1221.
- [5]ZHAO J Q, YAN J W, XUE T J, 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.
- [6]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.
- [7] 张小虎,李向东,赵振东,等. 基于深度学习的无人机影像麦穗识别方法[J]. 农业工程学报, 2021,37(6):123.
ZHANG X H, LI X D, ZHAO Z D, *et al.* Wheat ear recognition method based on deep learning from UAV images [J]. *Transactions of the Chinese Society of Agricultural Engineering*, 2021,37(6):123.
- [8] LI Q Y, CAI J H, BERGER B, *et al.* Detecting spikes of wheat plants using neural networks with Laws texture energy [J]. *Plant Methods*, 2017,13:83.
- [9] FERNANDEZ-GALLEGO J A, KEFAUVER S C, GUTIÉRREZ N A, *et al.* Wheat ear counting in-field conditions: High throughput and low-cost approach using RGB images [J]. *Plant Methods*, 2018,14:22.
- [10] 薛媛. 小麦麦穗的目标检测与杂草分类识别研究[D]. 兰州: 兰州大学, 2022.
XUE Y. Research on target detection and weed classification and identification of wheat ears [D]. Lanzhou: Lanzhou University, 2022.
- [11] SHEN X J, ZHANG C, LIU K, *et al.* A lightweight network for improving wheat ears detection and counting based on YOLOv5s [J]. *Frontiers in Plant Science*, 2023, 14: 1289726.
- [12] 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.
- [13] HASAN M M, CHOPIN J P, LAGA H, *et al.* Correction to: Detection and analysis of wheat spikes using Convolutional Neural Networks [J]. *Plant Methods*, 2019,15:27.
- [14] SADEGHI-TEHRAN P, VIRLET N, AMPE E M, *et al.* DeepCount: In-field automatic quantification of wheat spikes using simple linear iterative clustering and deep convolutional neural networks [J]. *Frontiers in Plant Science*, 2019, 10:1176.
- [15] XIONG H P, CAO Z G, LU H, *et al.* TasselNetv2: In-field counting of wheat spikes with context-augmented local regression networks [J]. *Plant Methods*, 2019,15:150.
- [16] LIU W, ANGUELOV D, ERHAN D, *et al.* SSD: Single shot MultiBox detector [M]//Computer Vision - ECCV 2016. Cham: Springer International Publishing, 2016:21.
- [17] YAO Z S, LIU T, YANG T L, *et al.* Rapid detection of wheat ears in orthophotos from unmanned aerial vehicles in fields based on YOLOX [J]. *Frontiers in Plant Science*, 2022,13:851245.
- [18] 赵小丹,胡林. 基于深度学习的农业科技政策知识抽取方法研究[J]. 数据与计算发展前沿(中英文), 2024,6(4):106.
ZHAO X D, HU L. Research on knowledge extraction method for agricultural science and technology policies based on deep learning [J]. *Frontiers of Data & Computing*, 2024,6(4):106.
- [19] REN S Q, HE K M, GIRSHICK R, *et al.* Faster R-CNN: Towards real-time object detection with region proposal networks [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017,39(6):1137.
- [20] LI L, HASSAN M A, YANG S R, *et al.* Development of image-based wheat spike counter through a Faster R-CNN algorithm and application for genetic studies [J]. *The Crop Journal*, 2022,10(5):1303.
- [21] 陈庆超,韩松,毛钧毅. 采用多层次特征融合 SPP-net 的暂态稳定多任务预测[J]. 控制与决策, 2022,37(5):1279.
CHEN Q C, HAN S, MAO J Y. Multi-task prediction for transient stability using, multi-level feature fusion based SPP-net [J]. *Control and Decision*, 2022,37(5):1279.
- [22] HE K M, ZHANG X Y, REN S Q, *et al.* Spatial pyramid pooling in deep convolutional networks for visual recognition [C]//Computer Vision - ECCV 2014. Cham: Springer, 2014:346.
- [23] 王姗姗,余山山,张文毅,等. 基于特征点邻域 Hough 变换的水稻秧苗行检测[J]. 农业机械学报, 2020,51(10):18.
WANG S S, YU S S, ZHANG W Y, *et al.* Detection of rice seedling rows based on Hough transform of feature point neighborhood [J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2020,51(10):18.
- [24] LIU C X, WANG K W, LU H, *et al.* Dynamic color transform networks for wheat head detection [J]. *Plant Phenomics*, 2022,2022:9818452.
- [25] 黄硕,周亚男,王起帆,等. 改进 YOLOv5 测量田间小麦单位面积穗数[J]. 农业工程学报, 2022,38(16):235.
HUANG S, ZHOU Y N, WANG Q F, *et al.* Measuring the number of wheat spikes per unit area in fields using an improved YOLOv5 [J]. *Transactions of the Chinese Society of Agricultural Engineering*, 2022,38(16):235.
- [26] XU X M, ZHOU L, YU H L, *et al.* Winter wheat ear counting based on improved YOLOv7x and Kalman filter tracking algorithm with video streaming [J]. *Frontiers in Plant Science*, 2024,15:1346182.
- [27] WU T L, ZHONG S Y, CHEN H, *et al.* Research on the method of counting wheat ears via video based on improved YOLOv7 and DeepSort [J]. *Sensors*, 2023,23(10):4880.
- [28] 黄小毛,张维,邱天,等. 基于无人机视频影像的油菜苗检测与计数[J]. 农业工程学报, 2024,40(10):147.
HUANG X M, ZHANG W, QIU T, *et al.* Rapeseed seedling detection and counting based on UAV videos [J]. *Transactions of the Chinese Society of Agricultural Engineering*, 2024,40(10):147.
- [29] LI Y, MA R, ZHANG R T, *et al.* A tea buds counting method based on YOLOv5 and Kalman filter tracking algorithm [J]. *Plant Phenomics*, 2023,5:30.
- [30] GUPTA C, GILL N S, GULIA P, *et al.* A novel finetuned YOLOv6 transfer learning model for real-time object detection [J]. *Journal of Real-Time Image Processing*, 2023,20(3):42.
- [31] 赵艳芹,崔翊超. 基于改进 YOLOv8 算法的草莓采摘目标检测算法[J]. 高师理科学刊, 2024,44(10):46.
ZHAO Y Q, CUI Y C. Strawberry picking target detection algorithm based on improved YOLOv8 algorithm [J]. *Journal of Science of Teachers' College and University*, 2024, 44(10):46.
- [32] 姜庆玲. 基于 YOLOv8 的鲁棒西红柿目标检测方法[J]. 辽宁师专学报(自然科学版), 2024,26(3):64.
JIANG Q L. Robust tomato target detection method based on YOLOv8 [J]. *Journal of Liaoning Normal Colleges (Natural Science Edition)*, 2024,26(3):64.
- [33] TANG L, LI T T, XU C W. Stratigraphic division method based on the improved YOLOv8 [J]. *Applied Sciences*, 2024,14(20):9485.
- [34] GUAN S T, LIN Y M, LIN G Y, *et al.* Real-time detection and counting of wheat spikes based on improved YOLOv10 [J]. *Agronomy*, 2024,14(9):1936.
- [35] HONG Q Q, LIU W, ZHU Y, *et al.* CTHNet: A network for wheat ear counting with local-global features fusion based on hybrid architecture [J]. *Frontiers in Plant Science*, 2024, 15:1425131.
- [36] KHAN M A, MENOUAR H, HAMILA R. LCDnet: A lightweight crowd density estimation model for real-time video surveillance [J]. *Journal of Real-Time Image Processing*, 2023,20(2):29.
- [37] MISRA T, ARORA A, MARWAHA S, *et al.* SpikeSegNet-a deep learning approach utilizing encoder-decoder network with hourglass for spike segmentation and counting in wheat plant from visual imaging [J]. *Plant Methods*, 2020,16(1):40.