

基于YOLOv11的杂粮成分识别

彭雪峰, 瞿士骐, 王雨晗, 李涛, 曾曙光

三峡大学数理学院、核能科学与工程学院, 湖北 宜昌

收稿日期: 2026年4月26日; 录用日期: 2026年5月19日; 发布日期: 2026年5月29日

摘要

杂粮由于其粮食种类丰富、营养全面等优点存在着极大的市场空间。为便于消费者了解杂粮的品质、工商部门检测杂粮的质量以及农民设备的局限性, 呼应市场的需求, 本文设计了一种基于YOLO模型的谷物分类计数器, 且仅用手机相机即可实现数据收集。针对杂粮识别场景中颗粒目标细小、密集易粘连、类别外观高度相似、复杂背景干扰强等问题, 本文通过针对性数据增强、优化分类与定位损失权重、引入测试时增强TTA及差异化后处理策略。实验结果表明, 改进模型在10类杂粮识别任务中, 有8类杂粮都具备较高分类精度与计数稳定性, 模型对大多数谷物的识别准确率达到85%以上。改进模型的平均绝对误差由1.130降至0.880, 平均相对误差由3.5%降至2.9%。本文也通过混淆矩阵、性能曲线、对比实验、重复性检验与误差分析验证了鲁棒性与实用性, 可有效实现杂粮成分快速判别与精准计数。

关键词

谷物分类, 谷物计数, 图像增强, YOLO模型

Based on YOLOv11 for Multi-Grain Component Recognition

Xuefeng Peng, Shiqi Qu, Yuhan Wang, Tao Li, Shuguang Zeng

College of Mathematics and Physics, College of Nuclear Science and Engineering, China Three Gorges University, Yichang Hubei

Received: April 26, 2026; accepted: May 19, 2026; published: May 29, 2026

Abstract

Coarse cereals hold significant market potential due to their rich variety and comprehensive nutritional value. To help consumers assess product quality, support market regulation, and address the limitations of equipment in agricultural settings, this paper presents a grain classification and counting system based on a YOLO model, which only requires a mobile phone camera for data

collection. To address challenges in coarse cereal recognition, such as small and densely packed particles, high inter-class visual similarity, and complex background interference, this study employs targeted data augmentation, optimizes the weights of classification and localization losses, introduces Test Time Augmentation (TTA), and applies differentiated post-processing strategies. Experimental results show that the improved model achieves high classification accuracy and counting stability for 8 out of 10 coarse cereal types, with recognition accuracy exceeding 85% for most grains. The mean absolute error of the enhanced model is reduced from 1.130 to 0.880, and the mean relative error decreases from 3.5% to 2.9%. The robustness and practicality of the system are further validated through confusion matrices, performance curves, comparative experiments, repeatability tests, and error analysis, demonstrating its effectiveness for rapid composition identification and accurate counting of mixed grains.

Keywords

Grain Classification, Grain Counting, Image Enhancement, Yolo Model

Copyright © 2026 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

杂粮由于其粮食种类丰富、营养全面等优点存在着极大的市场空间。在各个超市或批发厂商，均可买到不同粮食混合而成的杂粮。在各个谷物自动化生产工厂，对不同谷物的分类也存在极大的需求。因此，为便于消费者了解杂粮的品质以及工商部门检测杂粮的质量，谷物分类的研究具有极大的价值。

由于近年来基于深度学习的目标检测方法在各领域都得到十分充分、成功的应用，因此目前针对杂粮种类识别与计数已有若干有层次、有特色的方法：韩雪晨、程征等人[1]从不同种类杂粮 HSV 颜色空间及尺寸差异出发来区分杂粮中各谷物，孙明、王一鸣等人[2]提出了十分巧妙的基于色调的黄粒米检测算法，卢亚玲[3]提出了图像分割最佳阈值算法，再据此可靠地进行粮虫对象计数。但毋庸讳言，现有传统杂粮检测方法大多仍依靠人工设计的特征提取规则，在复杂场景下显然难以为实际应用提供理想性能。

许晓琳、牛思琪等人[4]提出一种基于改进 YOLOv8 的算法 MEGS_YOLOv8 对玉米籽粒品种进行鉴别，但只对单一玉米进行鉴别；李云、邱述金等人[5]提出一种基于 YOLOv5 的轻量化检测方法对谷穗进行实时检测；上述研究都基于 YOLO 算法对单一谷物进行研究，而目前尚未有研究将 YOLOv11 应用于杂粮检测。

YOLOv11 是目前应用最广泛、最成熟的检测算法，在检测速度、识别精度、工程部署诸方面都表现出极好的性能。而杂粮颗粒检测任务中不可避免地会遇到颗粒堆叠密集、互相遮挡、小目标居多诸种困难，传统方法对此都难有理想解决办法。因此，YOLOv11 有十分自然且有力的优势：它能用多类别联合训练的方式自动、可靠地提取目标细节特征，又不需要人工设计复杂的分割规则，对光照、背景变化都具有极好的鲁棒性。更重要的是，其改进的特征融合网络配合动态锚框设计，能非常合理地融合不同尺度的特征信息，故能对小颗粒做到精准定位，也由此直接提高密集杂粮检测的计数精度。

2. YOLOv11 算法

2.1. 算法原理

YOLOv11 在模型结构、损失函数、训练策略等方面进行了多项创新。其算法原理如下：

YOLOv11 在架构上对 YOLOv8 进行了针对性的改进，核心在于引入了两个关键模块：C3K2 和 C2PSA

[6], 显著提升了模型性能与效率。

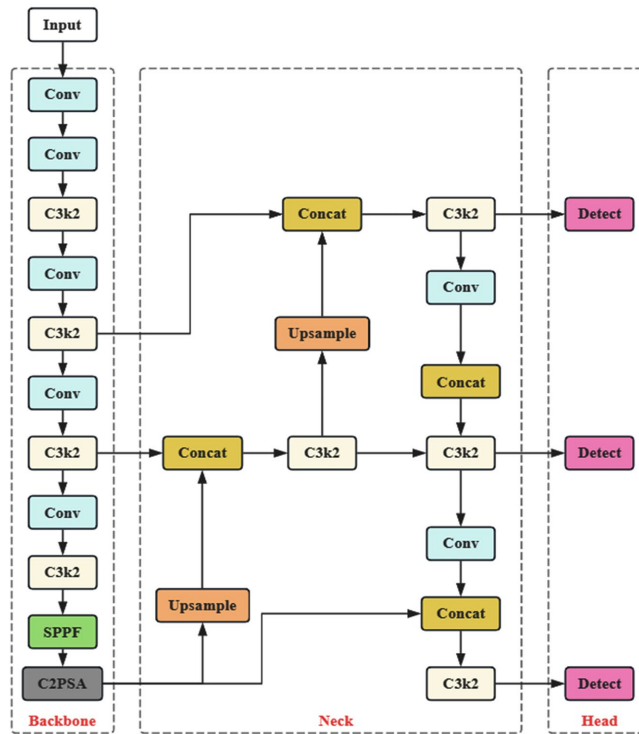


Figure 1. YOLOv11 network architecture diagram
图 1. YOLOv11 网络结构图

C3K2 和 C2PSA 两者都以 CSP 结构为基础, 结构如图 1 所示, 先把输入特征合理地拆分成两条支路予以单独处理, 之后自然、巧妙地做特征融合, 因此都兼顾了效率又有利于特征提取。更难得的是, C3K2 模块使用了可变卷积核结构, 由此自然地扩大了网络的感受范围, 在处理大目标、复杂背景时有极好的表现。与此形成极好补充的是 C2PSA 模块在另一支路中所引入的 PSA 注意力机制, 它将金字塔卷积多尺度特征提取的优势与通道注意力机制二者完美结合, 能根据任务需要主动、可靠地强化所关心的特征通道, 故对遮挡目标的检测能力有极大提升。

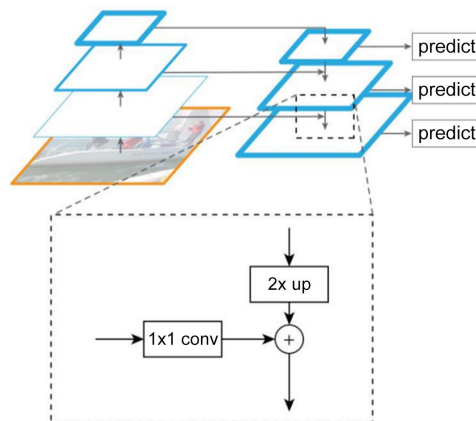


Figure 2. FPN structure diagram
图 2. FPN 结构图

YOLOv11 保留了 FPN 结构进行多尺度特征融合，并进行了轻量化与自适应改进，FPN 结构中包括自下而上，自上而下和横向连接三个部分，如图 2 所示。这种结构可以将各个层级的特征进行融合，使其同时具有强语义信息和强空间信息。

YOLOv11 的总体损失函数 L [7] 由定位损失 λ_{coord} 、置信度损失 L_{conf} 和分类损失 L_{class} 加权组成，其一般形式为：

$$L = \lambda_{coord} L_{loc} + L_{conf} + L_{class} \quad (1)$$

其中， λ_{coord} 为协调权重。在定位损失上，本文采用了能更好优化边界框重合度、中心点距离及长宽比的 CIoU 损失(公式见下文 2.2 节)，以提升回归精度。

2.2. 算法公式

YOLOv11 算法学习涉及的公式主要如下：

2.2.1. 预测输出公式

YOLOv11 的预测输出为三维张量

$$S \times S \times B \times (5 + C) \quad (2)$$

对于每个边界框，预测值可表示为

$$p = (x, y, w, h, confidence, c_1, c_2, \dots, c_C) \quad (3)$$

其中， S 是网格大小； B 是每个网格单元预测的边界框数量；5 表示边界框的 x , y , w , h 四个坐标参数及置信度； C 是类别数； x , y 是边界框中心相对于网络单元的偏移； w , h 是边界框的宽度和高度相对于整幅图像的比例； $confidence = P(object) \cdot IOU_{pred}^{truth}$ ，即对象存在的概率乘以预测框与真实框的交并比； c_i 是第 i 类别的条件概率。

2.2.2. 特征融合公式

改进的特征金字塔网络采用双向特征金字塔结构，公式为

$$P_{out} = \sum_i \omega_i \cdot Resize(P_{in}^i) \quad (4)$$

其中， $\sum \omega_i = 1$ ， ω_i 是可学习的权重参数，确保不同层级特征的贡献度能够自适应调整。

2.2.3. CIoU 损失公式

$$L_{CIoU} = 1 - IoU + \frac{\rho^2(b, b^{gt})}{c^2} + \alpha v \quad (5)$$

其中， ρ 表示中心点欧氏距离， c 是最小外接矩阵对角线长度， v 衡量长宽比一致性。

3. 数据预处理

3.1. 数据采集

由于要让所训练的模型能真正适应真实复杂环境中谷物的识别任务，因此本文自主建立了包含大米、糙米、黑豆、黑米、红豆、红米、花生、绿豆、糯米、玉米共 10 类常见谷物的数据集，有 600 张训练图像及 150 张验证图像，所有图像都由人工用 labeling 工具严格标注，得到标准 txt 格式的标注文件。图像尺寸统一为 1440×1440 像素，训练时设置 250 epoch 的轮次，每批 12 个样本，且对数据作了统一分辨率、统一标签格式的预处理，因而既加快了训练速度，又有利于缓解训练中的数据过拟合问题，也为此

后的图像增强做好了充分准备。训练集部分示例如图 3。



Figure 3. Training set samples
图 3. 训练集示例

3.2. 数据增强

为更合理地模拟实际场景，所用图像均为手机拍摄，故先解决其成像质量偏低的问题，采用两阶段图像增强策略。第一阶段对所有图像做统一的轻度 CLAHE 对比度增强及边缘锐化处理[8]，之后针对形状十分相似的白米、糙米和糯米，设计了专门的增强流程：先将图像转为 LAB 颜色空间，在 L 通道做自适应直方图均衡，再适当增强 a、b 通道的对比度以放大颜色差异，同时结合锐化及高斯差分算法突出颗粒表面纹理及光泽差异。最后对此类图像做合理复制，切实提高模型对相似谷物的区分能力。第一次增强数据示例如图 4。

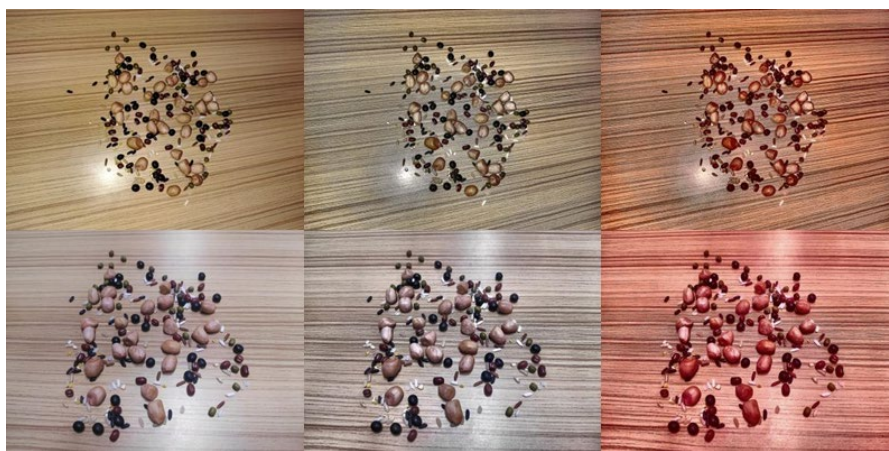


Figure 4. Enhanced data samples (First Round)
图 4. 第一次增强数据示例

第二阶段在模型训练时同时使用 YOLO 自带的图像增强方法，故以 Mosaic [9]和 Mixup 两种增强方式作为核心。具体而言，Mosaic 用四张训练图像做随机拼接，因此很合理地模拟了杂粮堆放时密集拥挤、彼此遮挡的真实情况，而 Mixup 用图像及标签都做线性加权融合，有极好的正则化效果，因而所训练出

的模型对光照偏移、颜色变化等干扰都更加鲁棒。HSV 颜色空间扰动、随机旋转、随机缩放诸种操作又进一步丰富了数据集的多样性。变换后部分示例如图 5。

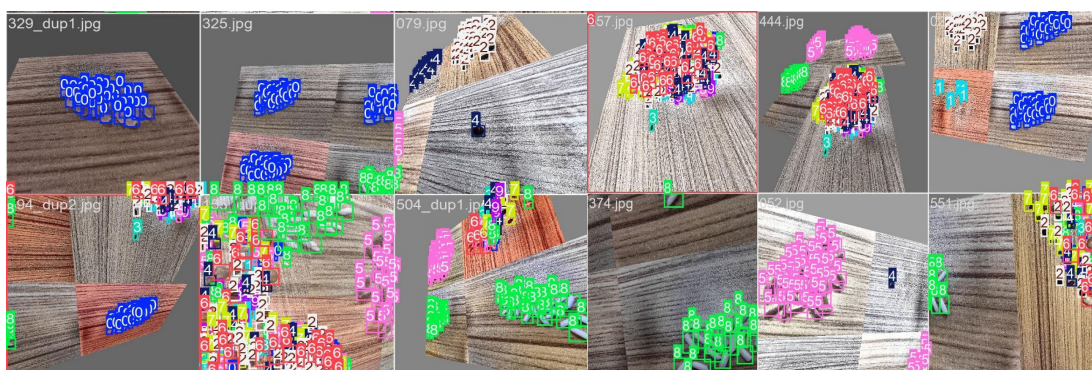


Figure 5. YOLO internal data transformations
图 5. YOLO 内部数据变换

4. 实验过程

4.1. 损失函数

由于小目标检测存在困难，本文对损失函数权重做了合理调整：杂粮识别任务中不同品类谷物认错后果不同，把糯米错判为白米的影响远大于检测框位置稍有偏差，故本实验把分类损失权重调高到 4.0，定位损失权重调低到 0.3，训练时所选初始学习率取 0.0008，并配合余弦退火学习率衰减方式，保证模型训练时平稳、可靠地收敛。模型训练与验证曲线如图 6。

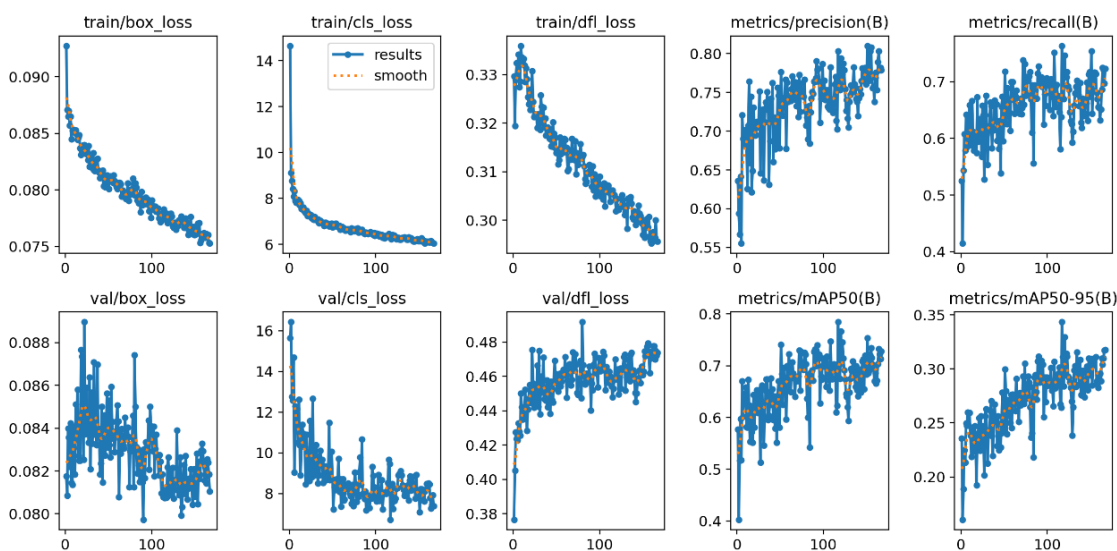


Figure 6. Model training and validation curves
图 6. 模型训练与验证曲线

模型训练与验证曲线可以很自然、合理地检验所提出的重分类、轻定位的损失函数策略及保守学习率的有效性，故本文给出了模型训练及验证曲线。

由于模型训练时各损失量都平稳下降并自然趋于稳定，故可十分清楚、可靠地判断：分类损失快速

降到低位，因此模型已经很好地学得了谷物各类别的细小特征，而定位损失的收敛过程也十分平滑，与实际应用中边框位置不必极端精确的要求十分契合。更重要的是，所得到的模型在验证数据集上表现稳定，精确率、召回率都超过 0.7，mAP50 约为 0.75。故能很好地处理复杂背景下的多类杂粮识别任务。

4.2. 优化方法

由于实验中对训练时所用的 HSV 亮度扰动幅度作了合理调小的处理，因此很好地避免了亮度变化幅度过大而引起的目标边缘模糊的问题，又自然、妥帖地加大了饱和度调整的强度以抑制光照强弱变化的干扰，更难得的是网络中引入了 0.15 的 Dropout 比率，由此十分自然地防止了对训练集中某种光照条件下背景过拟合。

4.3. 后处理阶段

对模型做测试时自然、合理地采用了 Test-Time Augmentation (TTA) 技术：先对单张测试图像作原图、水平翻转、轻微亮度提高、对比度调整诸种方式的独立变换推理，再用非极大值抑制算法对各次结果予以可靠融合。由于本实验中所用图片全为手机拍摄，故成像质量参差不齐，对焦模糊、画面抖动等情况十分常见，因此引入 TTA 能让不同版本的图像互为补充、彼此校验，很自然地降低错误检测的概率。更难得的是，根据验证集的结果，作者设置了类别特异的置信度阈值：白米和糯米易混淆，故阈值定为 0.35，糙米属于中等难度，故取 0.3，其余各类取 0.25。由此顺理成章地减少了相似谷物之间的错分问题。

5. 实验分析

5.1. 归一化混淆矩阵的模型性能分析

为了深入评估模型对各谷物类别及背景的区别能力，本文绘制了归一化后的混淆矩阵，如图 7 所示。

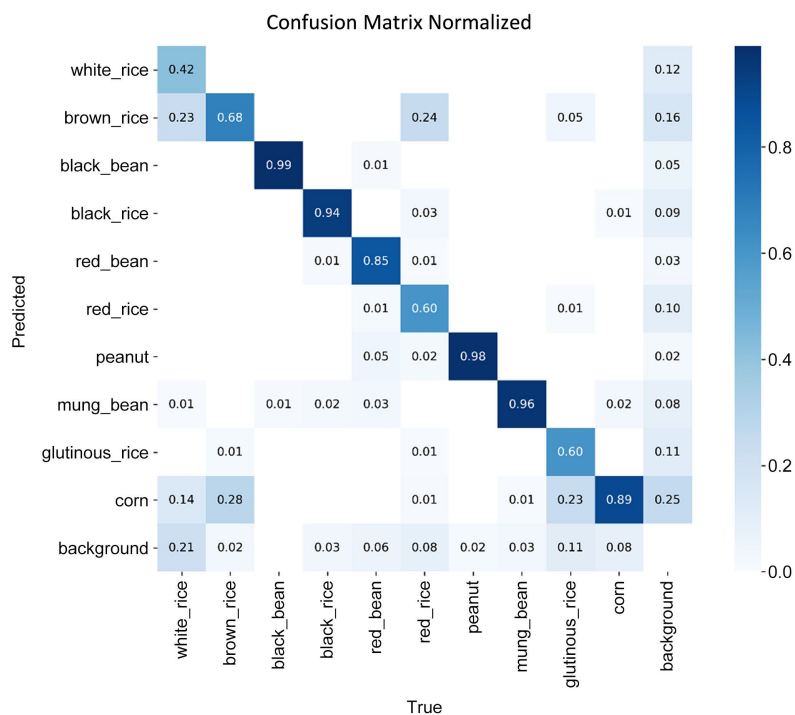


Figure 7. Normalized confusion matrix
图 7. 归一化后的混淆矩阵

从矩阵中可以清晰看出，模型对大多数谷物的识别准确率达到 85% 以上，其中黑豆识别准确率高达 0.99，表明模型能极其精准地捕捉黑豆的颜色与外形。黑米、花生、绿豆识别准确率分别达到 0.94、0.98、0.96。由于谷物本身颜色和形状相近，白米、糙米、红米、糯米识别准确率均在 0.42~0.68 区间，整体来看，模型把背景区域误判成谷物目标的比例很低，证明网络可以很好区分前景目标与背景杂物，抗干扰能力较强。

整体而言，所改进的模型对多品类杂粮识别有极好的区分能力，检测结果十分稳定。为后续构建高精度的谷物检测系统奠定了基础。

5.2. 性能曲线分析

图中展示了模型在 10 类谷物上的整体性能变化曲线图，蓝色粗线代表所有类别的平均水平，彩色细线为单一谷物的指标变化，四张子图分别为召回率 - 置信度曲线、精确率 - 召回率曲线、F1 分数 - 置信度曲线以及精确率 - 置信度曲线。

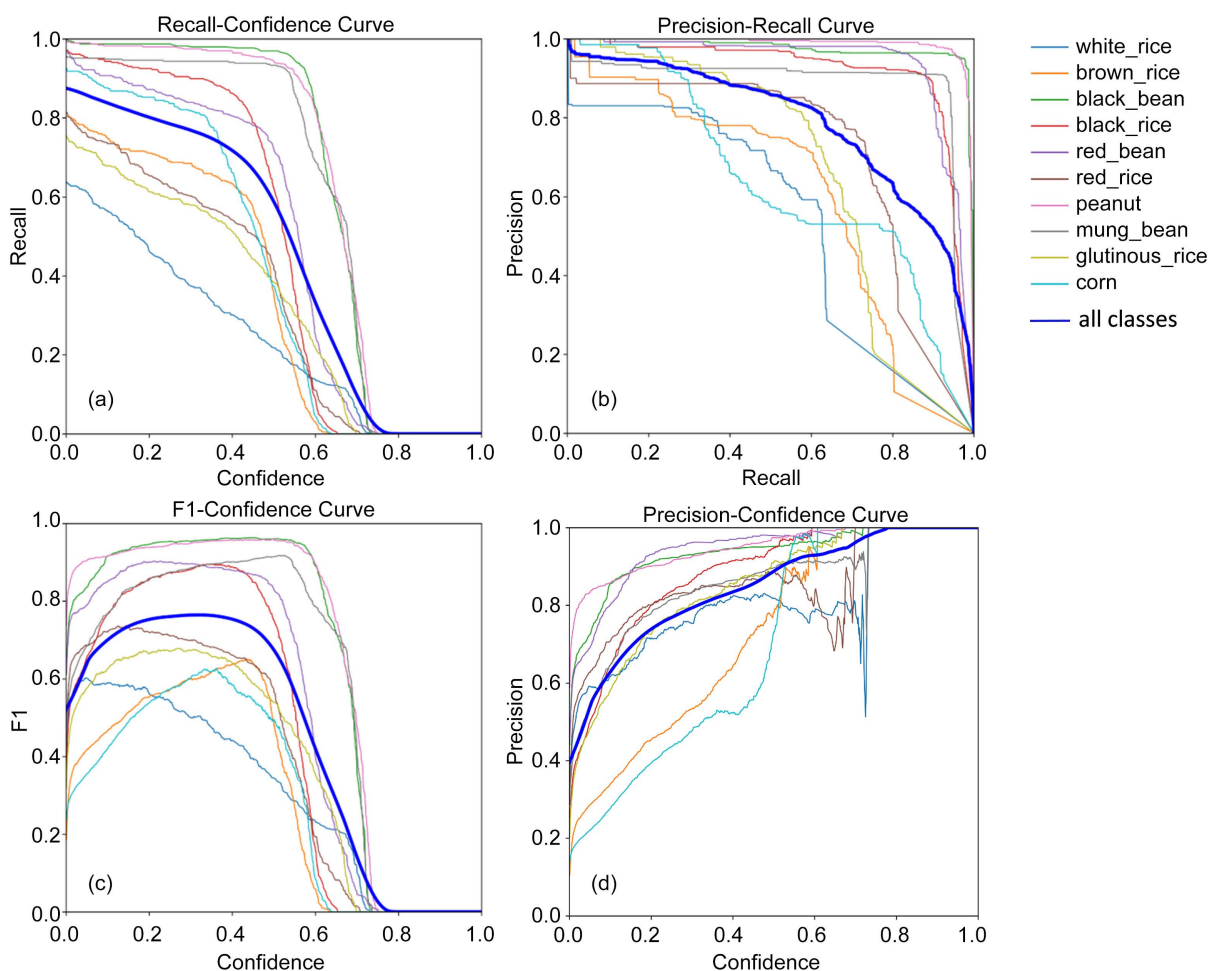


Figure 8. Comprehensive performance evaluation curves for each grain class

图 8. 类谷物上的综合性能评估曲线

图 8 中可以看到所选各谷物类别的模型性能的综合评价曲线，首先，召回率 - 置信度曲线(a)说明低置信度时召回率仍然很高，因此漏检率极低。其次，精确率 - 变率曲线(b)说明总体检测性能很好，各类

谷物的检测结果都稳定可靠。第三, F1-置信度曲线(c)有明确的单峰形状, 故极易确定最优置信度取值点。第四, 精确率-置信度曲线(d)表明随置信度增加, 精确率迅速上升并在较高置信度时趋于饱和, 误检率很小。综合来看, 整体各项指标分布均匀、表现均衡, 完全可以满足日常杂粮自动计数的实际使用要求。

5.3. 结果分析

实验结果表明, 模型对黑豆、花生、黑米等特征显著的类别识别效果优良, 但对白米、糙米的识别效果相对较差。考虑到在杂粮成分检测的实际场景中, 黑豆、花生等谷物的检出具备更大需求, 而白米和糙米本身属于同类主食, 营养构成和市场售价差距并不大, 实际应用中对二者精细区分的需求并不高。因此, 尽管模型对两者的识别效果不佳, 但是仍可应用在大多数场景。

接下来, 本文将剔除白米和糙米这两种谷物后, 对模型的实际性能进行检测。

对比实验

为验证测试时增强(TTA)对谷物计数精度的提升效果, 本文开展了对比实验, 结果如表 1 所示(方法一不使用 TTA; 方法二使用 TTA)。

Table 1. Comparative experimental results

表 1. 对比实验结果

	玉米	糯米	红米	黑米	花生	红豆	绿豆	黑豆	平均误差
人工计数	20	30	50	50	20	30	30	20	/
方法一计算个数	18	30	54	49	21	29	30	20	/
方法一绝对误差	2	0	4	1	1	1	0	0	1.130
方法一相对误差	0.10	0.00	0.08	0.02	0.05	0.03	0.00	0.00	0.035
方法二计算个数	18	30	54	50	21	30	30	20	/
方法二绝对误差	2	0	4	0	1	0	0	0	0.880
方法二相对误差	0.10	0.00	0.08	0.00	0.05	0.00	0.00	0.00	0.029

实验表明, 方法二相较于方法一平均绝对误差由 1.130 降至 0.880, 平均相对误差由 3.5% 降至 2.9%, 整体提升效果明显。其中, TTA 对黑米、红豆这些易混淆类别的计数误差实现了完全消除, 体现了 TTA 在提升模型鲁棒性、优化谷物计数准确性方面的有效性。

为验证两种谷物计数方法的重复性与可靠性, 本文进行了重复性检验与多梯度误差分析。重复性检验结果显示, 两种方法在连续三次测试中对八类谷物的计数结果完全一致, 变异系数均为 0, 表明二者均具备优异的操作稳定性和结果可重复性。表明二者均具备优异的操作稳定性和结果可重复性。

进一步的误差分析选取了玉米、红米和花生三类代表性谷物, 在 10 至 50 粒的数量梯度下进行。结果表明, 两种方法在多数情况下均能保持较高精度, 但其误差特征随样本量增加而呈现差异。具体而言, 玉米计数在 35 粒与 45 粒处分别出现了方法特异性误检; 红米在样本量较高(40~50 粒)时, 方法一出现颗粒粘连引起的严重漏检, 而方法二对此作了极其有效的改进, 稳定性显著提升; 花生由于形态特征十分明显, 故两种方法在全梯度测试中只有极少数点位存在 1 粒偏差。综上所述, 方法二在处理高样本量及较难检测类别时更稳健。

6. 结论

本文针对杂粮识别场景中颗粒目标细小、密集易粘连、类别外观高度相似、复杂背景干扰强等问题,

构建并改进基于 YOLOv11 的杂粮成分识别与计数模型。通过针对性数据增强、优化分类与定位损失权重、引入测试时增强 TTA 及差异化后处理策略,强化模型对细微纹理与颜色特征的提取能力,改善小目标与粘连目标的检测效果。实验结果表明,改进模型在 10 类杂粮识别任务中,有 8 类杂粮都具备较高分类精度与计数稳定性。

本文也通过混淆矩阵、性能曲线、对比实验、重复性检验与误差分析验证了鲁棒性与实用性,可有效实现杂粮成分快速判别与精准计数。但受限于小目标密集分布与类别高度相似性,模型在高数量、多相似谷物场景仍存在问题,对此未来将进一步优化轻量化骨干网络与特征融合结构,进一步提升检测精度和推理速度。也会考虑构建更丰富更复杂的数据集,强化模型泛化能力,降低漏检率与误检率,使其应用在越来越多的场景。

参考文献

- [1] 韩雪晨,程征,曾曙光,等. 基于图像处理的杂粮成分智能识别[J]. 光学与光电技术, 2014, 12(2): 58-62.
- [2] 孙明,王一鸣,凌云,等. 基于色调的黄粒米检测方法[J]. 农业机械学报, 2005(8): 78-81.
- [3] 卢亚玲. 基于图像处理的粮虫自动计数系统的研究[J]. 微计算机信息, 2007(24): 311-312.
- [4] 许晓琳,牛思琪,梁敖,等. 基于改进 YOLOv8 的玉米籽粒品种鉴别算法研究[J]. 中国农机化学报, 2025, 46(9): 137-144.
- [5] 李云,邱述金,赵华民,等. 基于轻量化 YoloV5 的谷穗实时检测方法[J]. 江苏农业科学, 2023, 51(6): 168-177.
- [6] 孟庆宽,蔡成宇,张漫,等. 基于改进 YOLO11n 的自然环境下玉米幼苗与杂草轻量化检测方法[J/OL]. 农业工程学报, 1-10. <https://link.cnki.net/urlid/11.2047.S.20260420.1621.006>, 2026-04-22.
- [7] 张毅杰,宋泽祺,张志涛,等. 基于 YOLO11n 改进的自然环境下云南澳洲坚果花识别[J/OL]. 农业工程学报, 1-10. <https://link.cnki.net/urlid/11.2047.S.20260420.1623.016>, 2026-04-22.
- [8] 戴一奇,张子豪,奚美娟,等. 基于 YOLO 神经网络的中药材自动识别模型及其手机应用的开发[J]. 北京生物医学工程, 2026, 45(2): 153-162.
- [9] 许佳琪. 基于 CLAHE-Mosaic 算法的铝模天花图像增强方法研究[D]: [硕士学位论文]. 上海: 东华理工大学, 2024.

附录

附录 1. 数据增强具体参数表

技术环节	具体参数/策略	参数值/设置
训练期在线增强	输入分辨率	1440
	色调扰动	0.02
	饱和度扰动	0.6
	旋转角度	10
	Mosaic 概率	0.85
	Mixup 概率	0.2
	Copy-Paste 概率	0.15
预处理离线增强	易混淆样本复制	3
	LAB 空间 a 通道增强	1.3
	LAB 空间 b 通道增强	1.2
	纹理增强	GaussianBlur sigma = 1.0/2.0

附录 2. 损失函数权重参数表

具体参数/策略	参数值/设置
分类损失权重	4
定位损失权重	均为 0.3

附录 3. 差异化后处理阈值参数表

具体参数/策略	参数值/设置
白米、糯米置信度阈值	0.35
糙米置信度阈值	0.3
其余 7 类置信度阈值	0.25