doi:10.6041/j.issn.1000-1298.2025.09.049

基于改进 YOLO v11n 的寒地旱直播水稻植株检测方法研究

赵　斌[1,2]　李明洋[1]　王　淞[1]　王　迪[1]　孙　博[1]　肖延伟[1]　陈修成[1]

（1.黑龙江八一农垦大学工程学院，大庆 163319；2.黑龙江省农机智能装备重点实验室，大庆 163319）

摘要：针对三至四叶期的旱地直播水稻秧苗种植密度高，秧苗检测难度大的问题，为实现东北地区复杂旱地环境中对直播秧苗植株精准计数，本文提出基于改进 YOLO v11n 模型的旱直播幼苗检测方法。首先，采用全维度动态卷积（Omni-dimensional dynamic convolution，ODConv）优化 C3k2 卷积块，增强复杂旱地环境秧苗特征捕获能力，提升模型对秧苗浅层特征的提取；其次，将快速空间金字塔池化（Spatial pyramid pooling fast，SPPF）替换为焦点调制机制（FocalModulation，FM），通过多尺度特征交互与注意力融合，保留秧苗关键特征与提升多尺度目标适应性；最后，在主干网络 C2PSA 模块中引入三重注意力机制（Triple Attention，TA），构建多维度特征增强策略，强化深层特征表征。构建旱田直播秧苗图像的数据集，利用该数据集对改进后的模型进行训练和验证。测试结果表明：改进的 YOLO v11n－TFO 旱直播秧苗检测算法的检测精度（P）、召回率（R）、平均精度均值（mAP50）和检测速度分别为 88.2%、79.6%、89.4%、217.9 f/s，其中 mAP50 较改进前提升 6.5 个百分点，模型浮点计算量 6.0×10^9，低于主流目标检测模型，构建的检测平台中旱直播秧苗计数精度达到 87% 以上。本文提出的基于 YOLO v11n－TFO 的寒地旱直播水稻检测模型有效平衡检测精度与算力的关系，且模型具有较好的可解释性，可满足旱直播水稻计数需求。

关键词：旱直播水稻；目标检测；YOLO v11n；多维度特征提取

中图分类号：TP391.41；S511.05　　**文献标识码**：A　　**文章编号**：1000-1298(2025)09-0585-11　　**OSID**：

Cold-region Dry Direct-seeded Rice Plant Detection Method Based on Improved YOLO v11n

ZHAO Bin[1,2]　LI Mingyang[1]　WANG Song[1]　WANG Di[1]　SUN Bo[1]　XIAO Yanwei[1]　CHEN Xiucheng[1]

(1. *College of Engineering*, *Heilongjiang Bayi Agricultural University*, *Daqing* 163319, *China*
2. *Heilongjiang Provincial Key Laboratory of Intelligent Agricultural Machinery Equipment*, *Daqing* 163319, *China*)

Abstract: Aiming to address the challenges of high planting density and difficult seedling detection in dryland direct-seeded rice at 3－4 leaf stages, an improved YOLO v11n－TFO model was proposed for precise seedling counting in complex northeastern dryland environments. Firstly, the C3k2 convolution block was enhanced by omni-dimensional dynamic convolution (ODConv) to strengthen shallow feature extraction and improve environmental adaptability. Secondly, spatial pyramid pooling fast (SPPF) was replaced with FocalModulation (FM) mechanism, achieving multi-scale feature interaction and attention fusion to preserve critical seedling characteristics. Thirdly, Triple Attention (TA) mechanism was integrated into the C2PSA module of backbone network, establishing multi-dimensional feature enhancement strategy for deeper feature representation. A dedicated dataset of dryland direct-seeded rice seedling images was constructed for model training and validation. The test results showed that the detection precision (P), recall (R), mean average precision (mAP50) and detection speed of the improved YOLO v11n－TFO dry live seedling detection algorithm were 88.2%, 79.6%, 89.4%, and 217.9 f/s, respectively, of which the mAP50 was improved by 6.5 percentage points compared with the pre-improvement period, and the model floating-point computation was 6.0×10^9, which was lower than the mainstream target detection model, and more than 87% accuracy in counting dry-drilled seedlings in

收稿日期：2025－03－25　修回日期：2025－04－23

基金项目：国家自然科学基金青年基金项目（32201655）、智慧农机电源与电网融合互动关键技术与商业模式研究项目（SGJNO000SCJS2310075）和黑龙江八一农垦大学研究生创新科研项目（YJSCX2024－Z02）

作者简介：赵斌（1970—），男，教授，主要从事农业信息获取与图像处理研究，E-mail：616283364@qq.com

the constructed testing platform. The proposed YOLO v11n – TFO-based detection model effectively balanced detection accuracy and computational efficiency while maintaining good interpretability, meeting the requirements for dry direct-seeded rice plant counting in cold regions.

Key words: dry direct-seeded rice; object detection; YOLO v11n; multi-dimensional feature extraction

0 引言

水稻旱直播是一种不经过育秧、插秧,利用直播机在旱地开沟播种的水稻种植模式[1]。在大田种植生产中具有节约用水、通风良好、减少人力等优点[2-4]。黑龙江省作为我国优质水稻的重要产区,在劳动力转移与耐旱水稻改良的大背景下,逐步推广发展寒地旱直播水稻种植技术,由于旱直播水稻种植易受气候的影响,因而水稻种植苗情不易掌握[5-6]。通过获取幼苗期秧苗数量信息可量化出苗率等苗情信息,在促进肥精准调控提供数据支撑的同时,为检验播种效果和水稻育种成效提供科学依据[7]。

传统水稻出苗信息统计通过种植户人工判断,耗时长、成本高、实施难度大[8]。目标检测技术通过分析目标数量与空间位置关系,在农业信息获取中发挥重要作用[9-10]。这类传统算法的特征提取依赖人工设计,具有较强解释性,适用于等株距种植,对于非均匀分布的旱直播秧苗存在训练构建周期长、环境适应性差的局限。随着深度学习发展,以 Faster R – CNN 为代表的两阶段算法和以 YOLO、SSD 为首的单阶段算法显著提升了检测效率与精度[11-15]。但两阶段算法存在算力需求大、效率低的问题。相较之下,单阶段算法通过端到端设计有效平衡精度与效率关系。部分学者[16-17]对 YOLO v5 优化实现油菜苗植株计数,均提高了一定检测精度;李兆东等[18]改进 YOLO v7 提升油菜幼苗识别性能;LIN 等[19]结合 MobileNet 与 CenterNet 实现棉花植株检测,但易受自然光线干扰。研究表明,单阶段算法凭借高效特征提取与灵活性,逐渐成为农业复杂环境下的主流选择,未来需进一步优化模型轻量化与抗干扰能力以适配多样化农田场景。

此外部分学者采用图像分割法对植株计数[20-22]。研究表明,图像分割方法对连续大目标(如成熟植株)具有较高精度和解释性优势,但对密集小目标(如旱直播秧苗)存在漏检率较高的问题,易受目标尺寸和图像分辨率限制。

综上所述,尽管基于深度学习的目标检测算法在农业应用中展现出较好的适应性,但针对形态特征丰富、密集非均匀分布的旱直播秧苗,存在检测精度低、算力需求大的局限性。因此本文根据旱条播种植环境,提出一种基于改进 YOLO v11n 的旱直播秧苗植株精准检测的方法,优化主干网络的特征提取环节,增强模型对多尺度、密集秧苗深层特征的发掘,以满足本地 PC 设备对旱直播秧苗的可靠检测。

1 数据集构建

1.1 图像采集

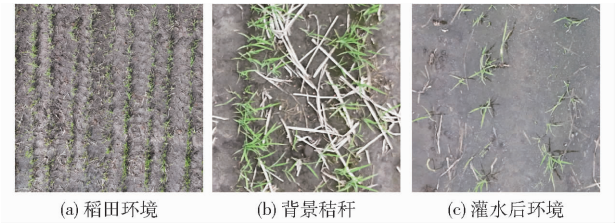
实验图像数据采集于黑龙江省佳木斯市建三江七星农场(47°20′56.234 8″N,132°46′27.978 2″E),图像采集时间为 2024 年 6 月 21 日 07:00—10:00,天气晴朗。如图 1 所示,图像采集设备选择大疆 Mavic3T 型无人机,在兼顾设备性能与图像质量的前提下,以高度 3 m(±0.1 m)、飞行速度 1 m/s 正射角度采集图像,获取原始 4 000 像素×3 000 像素图像 768 幅,筛选去除包含地块边界与部分过曝的图像后,对每幅图像不同区域随机提取 640 像素×640 像素的图像共 1 509 幅。



图 1 图像采集设备与采集区域航线图

Fig. 1 Map of equipment and area for image acquisition

水稻种植品种为“魔稻 119”,按照 300 mm、200 mm 间距的宽窄行播种,实行“旱播水管”的种植模式,并于四叶后期进行灌水作业。如图 2 所示,灌水后形成 2~4 cm 浅水层,覆盖土壤表面至秧苗下部叶片,局部秸秆、化肥浮沫漂浮水面,干扰模型对植株特征的判断,故选择在灌水前期采集图像。



(a) 稻田环境 (b) 背景秸秆 (c) 灌水后环境

图 2 旱直播水稻种植环境

Fig. 2 Live streaming of dry rice paddy environment

1.2　图像预处理与数据集建立

为增强模型对于旱直播水稻表型特征的捕捉、弱化背景的干扰，并提高模型的泛化能力，避免图像重复出现，随机从提取图像中筛选 750 幅图像，借助 Adobe Lightroom Classic 进行预处理。通过调整图像对比度、曝光度、饱和度，适当提高秸秆与秧苗的区分度；通过锐化与降噪处理突出秧苗植株细节、降低背景噪点。如图 3 所示，预处理后的图像背景色更深，秧苗植株形态特征更为突出，有助于训练中算法对植株形态特征的有效提取。

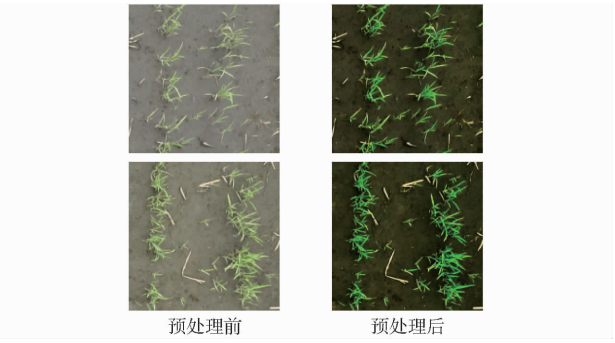


图 3　预处理前后对比

Fig. 3　Comparison chart before and after preprocessing

将预处理图像与未处理图像混合，在 Python 3.9 的环境下利用 Labelimg 对混合图像中的秧苗植株标注，并将标注命名为“rice”，标注数量共计 93 469 个。最后将构建好的数据集，按照比例 8∶1∶1 的将图像依次划分为训练集、验证集、测试集，各数据集间相互独立，在满足数据多样性的前提下，确保模型训练结果与测试结果可靠与准确。

2　模型构建与优化

2.1　模型选择

构建的旱直播秧苗数据集中，整体秧苗的植株个体形态异质、叶片呈现典型的线性狭窄，部分湿润土壤表层覆盖墨绿色青苔，浅绿色叶片与墨绿色背景的对比度较低，经典目标检测算法在检测中易受光照与背景干扰，无法有效应对。此外，旱地秧苗受播种深度、土壤墒情差异影响，秧苗种群分布呈条状局部聚集与随机稀疏并存的非均匀分布态，单株形态受遮挡等因素影响较大，要求检测模型具备秧苗形态本质特征提取和复杂空间分布的强适应能力。

YOLO v11 在模型设计中强化了对低对比度、小尺度目标的敏感性，其自适应特征融合机制能够在不依赖人工增强干预的情况下，通过多层级语义信息协同，能够捕捉秧苗与背景的微弱边界差异。

2.2　模型构建

YOLO v11n 模型在检测旱直播秧苗时存在以下问题：①传统卷积的固定核参数限制了对秧苗浅层表型多尺度特征信息捕捉能力和深层特征融合表达能力[23]。②SPPF 采用固定池化层数和池化窗口导致不能有效聚焦秧苗表型信息、秧苗个体结构信息感知能力不足的问题，产生较多冗余信息[24]。③对于部分小目标特征信息不能较好提取，检测中出现明显的漏检错检问题[25]。

为提高模型对旱直播秧苗检测的综合性能，拟采用 ODConv 优化 C3k2 卷积块，增强浅层网络的特征复用与提取能力[26]；利用 FM 机制替换 SPPF，改善模型对于多尺度秧苗特征的融合能力[27]；最后对主干网络 C2PSA 模块引入 TA 机制，提高小目标感知能力[28]。优化后模型结构如图 4 所示。

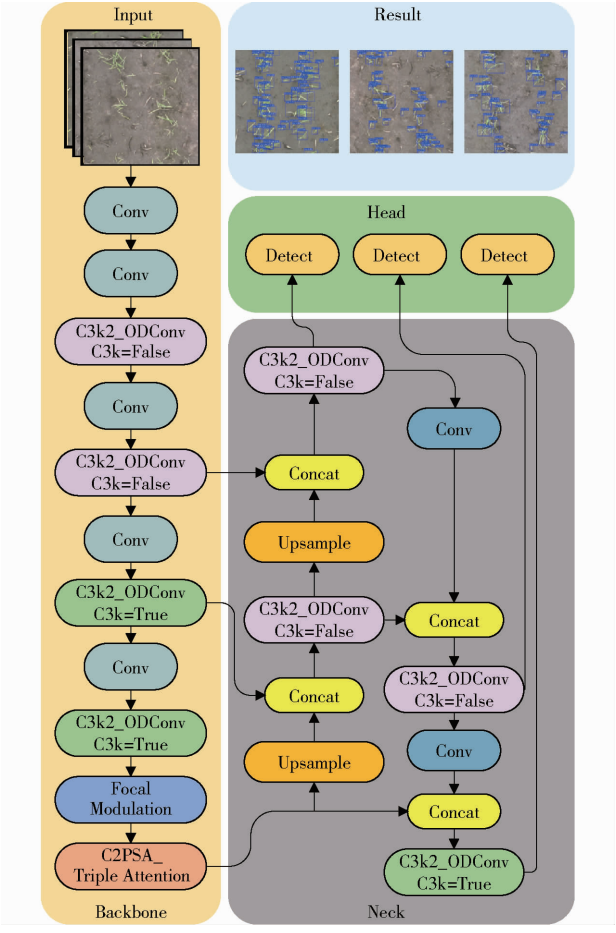


图 4　YOLO v11n－TFO 模型结构图

Fig. 4　YOLO v11n－TFO model structure diagram

2.2.1　ODConv 全维动态卷积

基于深度学习的目标检测模型，常采用固定的卷积头尺寸，限制了对大目标整体特征的捕捉和对小目标信息的获取。旱直播水稻秧苗形态异质，分布呈非均匀态，对于密集秧苗相互遮挡，为避免密集目标错判，影响后续特征增强与特征融合，本文将颈部和主干网络的 C3k2 卷积块替换为 ODConv 卷积，旨在提高浅层网络中特征的提取能力，同时增强深层网络中语义表达并降低跨尺度信息融合难度。

ODConv 是一种采用多维的注意力机制构成的动态卷积，相较于经典动态卷积通过多个卷积核进行加权组合以适应不同的输入特征实现动态性，在相关对比研究中，单个卷积核 ODConv 精度可达到甚至超越传统多核动态卷积精度，图 5 为经典动态卷积与 ODConv 结构对比[29]。

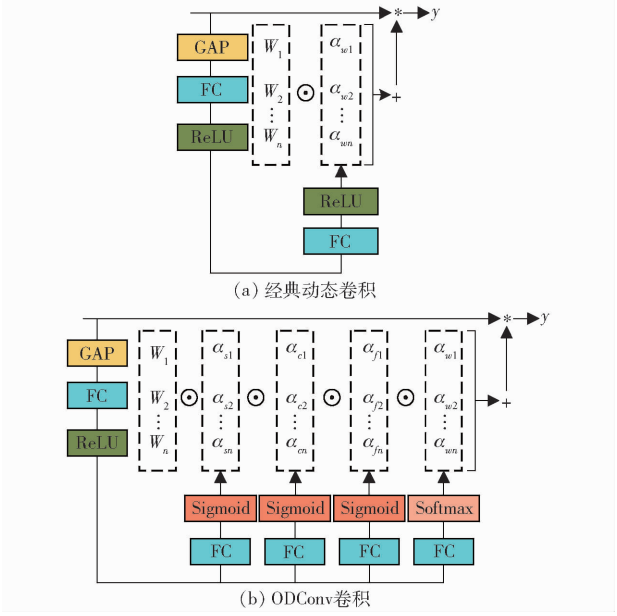


图 5 经典动态卷积与 ODConv 卷积结构对比

Fig. 5 Structural differences between classical dynamic convolution and ODConv convolution

ODConv 通过并行策略来学习卷积核在卷积核数量、核空间大小、输入通道数和输出通道数 4 个维度上的互补注意力[30]。图 6 为 ODConv 中 4 种类型的注意力与卷积作用示意图。

其中引入 C3k2 模块的 ODConv 卷积通过空间注意力权重 a_{si} 计算公式为

$$a_{si}=\sigma(\mathrm{Conv}_{1\times1}(X)) \tag{1}$$

式中 σ——Sigmoid 激活函数

X——输入特征图

输入通道注意力权重 a_{ci} 计算公式为

$$a_{ci}=\sigma(\mathrm{MLP}(\mathrm{GAP}(X))) \tag{2}$$

式中 MLP——含降维的全连接网络

GAP——全局平均池化

输出通道注意力权重 a_{fi} 计算公式为

$$a_{fi}=\sigma(\mathrm{MLP}(\mathrm{GAP}(X))) \tag{3}$$

卷积核数量注意力权重 a_{wi} 计算公式为

$$a_{wi}=\mathrm{Softmax}(\mathrm{MLP}(\mathrm{GAP}(X))) \tag{4}$$

式中 Softmax——归一化指数函数

分别求取 a_{si}、a_{ci}、a_{fi}、a_{wi}，并将所求权重与基础卷积核 W_i 逐元素相乘形成动态卷积核权重 W_{ODConv}，计算公式为

$$W_{\mathrm{ODConv}}=W_i a_{si} a_{ci} a_{fi} a_{wi} \tag{5}$$

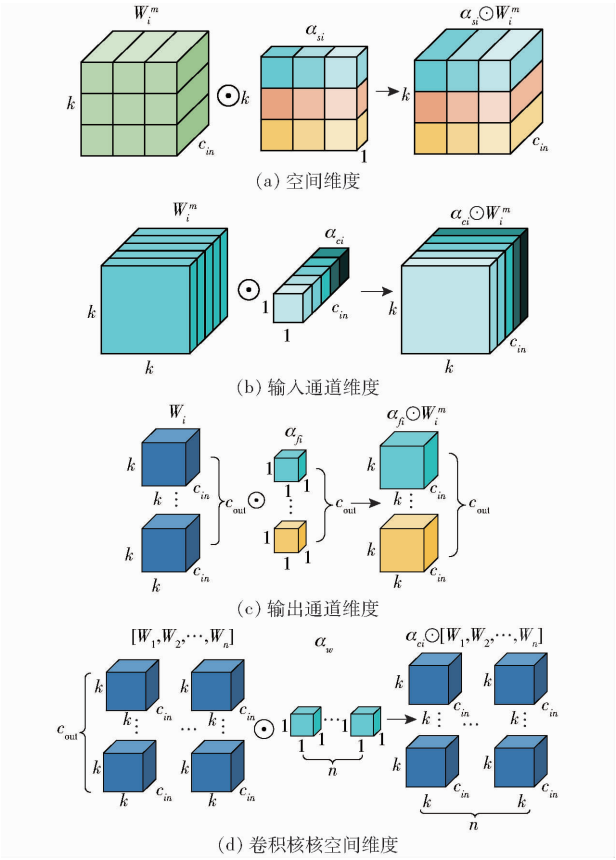


图 6 ODConv 中 4 种类型的注意力与卷积作用示意图

Fig. 6 Schematic of four types of attention and convolution effects in ODConv

ODConv 可以更精细地适应输入数据的特征，其中并行策略确保各维度之间的互补性和协同作用，在大幅减少网络参数量，同时在处理每个维度的特征时更加高效[31]。

根据 ODConv 的特性将其引入颈部和主干网络中的 C3k2 卷积块，改进的 C3k2_ODConv 由 3 个 ODConv 层构成。第 1 个 ODConv 对旱直播秧苗特征图降维，压缩特征并去除冗余，降维后的中间特征 Y_1 计算公式为

$$Y_1=\mathrm{ODConv}_{1\times1}(X;W^{(1)}_{\mathrm{ODConv}}) \tag{6}$$

第 2 个 ODConv 提取秧苗空间特征，空间特征增强后的输出 Y_2 计算公式为

$$Y_2=\mathrm{ODConv}_{k_2\times k_2}(Y_1;W^{(2)}_{\mathrm{ODConv}}) \tag{7}$$

式中 k_2——卷积核尺寸

第 3 个 ODConv 通过升维将低维特征映射到高维特征中，在恢复特征表达能力的同时，保证残差连接的输入与输出维度一致性，升维后的最终特征 Y_3 计算公式为

$$Y_3=\mathrm{ODConv}_{1\times1}(Y_2;W^{(3)}_{\mathrm{ODConv}}) \tag{8}$$

最后通过残差连接，稳定梯度流动，其中 X 直接参与最终输出，保留原始特征信息，缓解梯度消失问题。残差连接 Y_{Output} 计算公式为

$$Y_{\text{Output}} = X + Y_3 \tag{9}$$

2.2.2 FocalModulation 特征增强

SPPF 池化层采用无参数设计的固定池化核的非学习机制，避免浅层网络过拟合，计算具有高效性，但该机制对于密集秧苗的复杂空间分布，上下文交互不足，难以动态适配，易导致高密度秧苗特征错位。为了能有效聚焦 ODConv 提取的秧苗浅层表型特征，减小秧苗的异质性影响，过滤冗余信息，对秧苗多尺度特征信息融合，优化模型对不同尺度秧苗的适应能力，故采用 FocalModulation（FM）机制替换原有的 SPPF 机制。

FM 机制通过焦点调制机制替换自注意力机制，捕捉图像中长距离依赖和上下文信息。该功能实现可大致包括焦点上下文化、门控聚合、逐元素仿射变换三大关键环节[32]。

FM 结构图如图 7 所示，图中 Context Aggregate 表示焦点调制层级化的上下文聚合和门控聚合过程。焦点上下文化使用深度卷积层堆叠来编码不同范围的视觉上下文信息，从而允许网络在不同层次上理解图像内容，增强对全局结构的认识；门控环节通过计算上下文信息筛选重要秧苗特征信息，根据门控操作的结果选择性地聚合上下文信息到一个调制器中，它通过选择性地集中处理重要的上下文信息来提升网络的效率和性能；最后逐元素仿射变换，通过门控聚合得到的调制器注入到每个查询令牌中，使得模型能够利用上下文信息来有效地调整查询令牌，增强了模型对关键秧苗视觉特征的长距离依赖关系捕捉和表达能力，提升特征表达的丰富性与多样性。

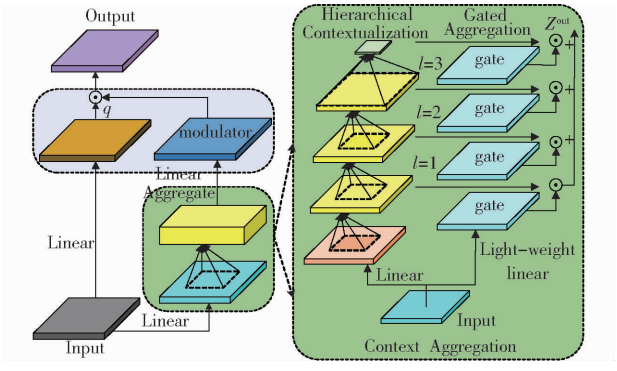


图 7　FocalModulation 结构图

Fig. 7　FocalModulation structure diagram

焦点调制机制通过层级化上下文聚合和门控聚合，实现对自注意力机制的替代，提升了模型的灵活性和可解释性，简化了交互和聚合操作，降低了计算复杂度。

2.2.3 注意力机制引入

FM 机制在建立全局特征信息关联性的过程中，易忽略部分小尺度秧苗特征信息，引起个别较小秧苗漏检。为提高对小目标秧苗检测准确性，优化密集小目标秧苗特征提取，避免算法模型的漏检，在主干网络的 C2PSA 模块引入注意力机制。

目前常用的目标检测注意力机制包括通道注意力（Squeeze excitation，SE）机制[33]、通道-空间注意力（Convolutional block attention，CBAM）机制[34]、全局上下文注意力（Global context attention，GC）机制[35]。

3 种注意力机制对比如图 8 所示。SE 机制的作用机理是通过获取通道的全局信息，再学习通道间的注意力权重 A_c，对特征图进行通道维度重标定。CBAM 机制则是作为一种混合注意力机制，融合了通道特性与空间特性，通过通道维度的池化生成空间权重图，在输入特征图 F 后与原始特征图通道注意力权重 A_c 逐乘运算，增强重要通道的响应，并生成新的特征图 F'，再将其与空间注意力权重 A_s 逐乘，调整空间维度特征权重，聚焦任务关键区域，生成通道-空间注意力调整后的特征图 F''。GC 机制则是一种结合全局上下文建模与轻量化设计的注意力机制，通过全区特征提取、特征变换和特征增强，实现增强模型对长距离信息的感知能力。

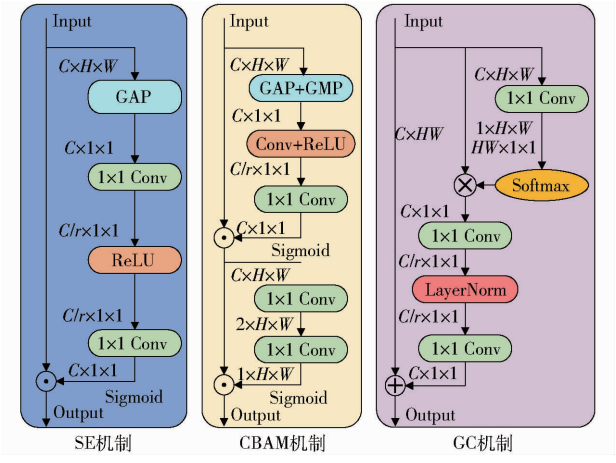


图 8　3 种注意力机制结构对比

Fig. 8　Comparison of different attention mechanism structures

三者通过动态权重分配和多维度特征增强，在一定程度上提高模型对关键信息的捕捉能力。在应用中从不同角度获取旱直播秧苗特征信息，为了充分利用三者特点，本研究提出在 YOLO v11n 模型的基础上引入融合 3 种注意力策略的 TA 注意力机制。不同注意力机制结构对比如图 9 所示。TA 注意力机制[36]由 3 个分支构成，每个分支负责捕获空间维度 H 或 W 与通道维度 C 之间的交互特征，利用每个分支中的输入张量进行排列变换，再通过 Z 池将 3 维通道维度“压缩”为一个大小为 $k \times k$ 的卷积

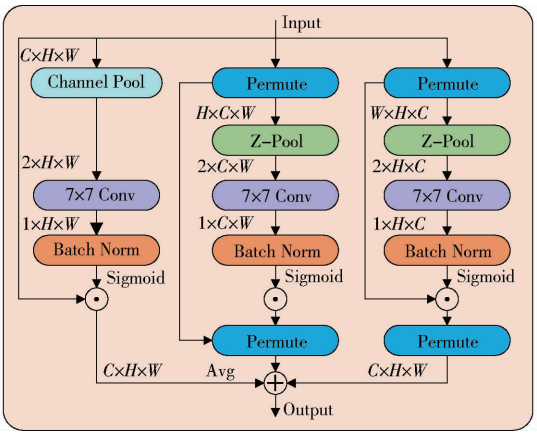


图 9 TA 机制结构图

Fig. 9 Structure diagram of TripleAttention mechanism

层,最后经 Sigmoid 激活函数生成注意力权重。生成的注意力权重经排列变换后应用于输入张量,再变换回原来的形状。

其中通道注意力分支计算公式为

$$A_c = \sigma(\mathrm{MLP}(\mathrm{GAP}(F))) \tag{10}$$

空间注意力分支计算公式为

$$A_s = \sigma(f^{k\times k}([\mathrm{GAP}(F);\mathrm{GMP}(F)])) \tag{11}$$

式中 $f^{k\times k}$——含降维的全连接网络

GMP——全局最大池化

跨维度注意力分支计算公式为

$$A_d = \mathrm{Softmax}\left(\frac{\boldsymbol{QK}^{\mathrm{T}}}{\sqrt{d}}\right)\boldsymbol{V} \tag{12}$$

式中 A_d——跨维度注意力权重

$\boldsymbol{Q}$——跨维度分支的查询

$\boldsymbol{K}$——跨维度分支的键

$\boldsymbol{V}$——跨维度分支的值

d——查询、键、值的特征投影维度

最后对 3 个注意力分支融合计算公式为

$$F_{\mathrm{out}} = \alpha(F\otimes A_c) + \beta(F\otimes A_s) + \gamma A_d \tag{13}$$

式中 α、β、γ——可学习的融合权重

F_{out}——融合生成的最终输出特征

图 10 为三重注意力机制原理图,其中每个子图表示三重注意力机制的一个分支[37]。分支 1 对高度(H)和通道(C)的维度旋转输入张量,经残差变换提取特征;分支 2 则对宽度(W)和通道(C)的维度旋转输入张量,经残差变换提取特征;分支 3 没有旋转操作,直接对输入张量处理,经残差变换提取特征。通过这样的设计,三重注意力模型能够有效地捕获输入张量中空间和通道维度之间的交互关系。该方法使模型能够构建通道与空间位置之间的相互依赖性,从而高效综合不同维度上的信息,更好地捕获秧苗目标的内在特征,提高模型对图像中复杂信息的理解能力。

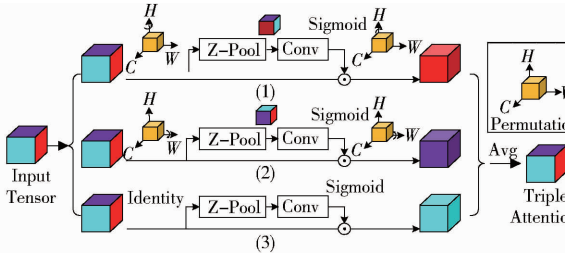


图 10 TA 机制原理图

Fig. 10 TA mechanism schematic

2.3 实验环境与评价指标

2.3.1 实验环境

本实验硬件设备平台包括:Intel Core i7-13700KF CPU、32GB RAM、Nvidia GeForce RTX 4070TI Super 16G GPU,软件平台包括:64 位 Windows 10 操作系统、Cuda 11.8、Cudnn 11.x、Python 3.11、Pytorch 2.0.1 深度学框架、PyCharm 2023.3.5 编辑器。相关超参数设置:迭代次数 400,批处理大小 16,优化器选择 SGD,为提高收敛速率设置初始学习率为 0.02。

2.3.2 评价指标

为验证改进模型对旱直播秧苗检测的性能,选取关键性能指标进行多维评价:在检测精度方面,采用检测精度、召回率、平均精度均值作为核心指标;针对模型复杂度,选取浮点计算量进行量化分析;同时以检测速度评估实时性表现。通过与主流目标检测算法进行对比实验,综合验证所提方法在精度-效率平衡性上的提升效果。

3 结果与分析

选择构建的旱直播秧苗的训练集,对改进后的模型进行训练,训练结果表明,改进后的 YOLO v11n-TFO 模型对于旱直播秧苗的检测精度、召回率、平均精度均值分别为 88.2%、79.6%、89.4%,平均精度均值较改进前提升 6.5 个百分点,检测速度达到 217.9 f/s。

图 11 为损失函数曲线,前 50 次迭代验证集损失值与训练集损失值同步快速下降,在 150 次迭代后 2 个损失函数曲线平行收敛,模型达到帕累托最优状态。其中最后 10 次训练损失中由于 Mosaic 增

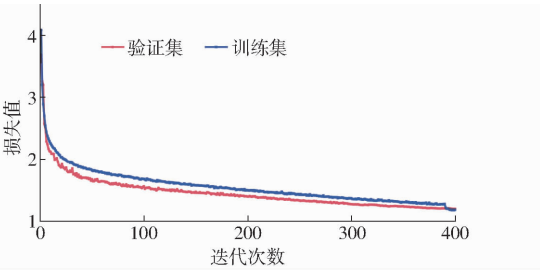


图 11 损失函数曲线

Fig. 11 Loss function curves

强关闭出现断崖下降，并在 1.18 附近波动。训练过程中二者趋势相同，损失验证没有出现较大起伏，二者最终收敛比均小于 2%，针对旱直播秧苗构建的 YOLO v11n－TFO 模型没有出现过拟合与欠拟合，具备良好的记忆能力和泛化能力。

3.1　实验结果

3.1.1　消融实验

为进一步验证本文提出模型各改进部分对旱直播秧苗检测的有效性，以 YOLO v11n 为基准模型对改进部分进行消融实验。实验结果如表 1 所示，ODConv 通过提高特征提取效果增强网络的灵活性和适应性，同时有效控制计算成本；FM 机制通过对秧苗表型特征信息融合与增强提升特征表达质量；TA 机制则强化模型对秧苗特征提取理解能力。结果表明 TA 机制、FM 机制和 ODConv 的独立使用均能促进模型对旱直播秧苗检测能力，其中 ODConv 在提高模型检测性能同时降低了模型浮点计算量。

表 1　消融实验结果

Tab. 1　Ablation test results

YOLO v11n	TA	FM	OD	检测精度/%	召回率/%	平均精度均值/%	浮点计算量
√	×	×	×	80.6	74.5	82.9	6.3×10^9
√	√	×	×	83.0	75.3	85.4	6.3×10^9
√	×	√	×	82.0	75.4	85.0	6.4×10^9
√	×	×	√	83.6	75.8	85.6	5.9×10^9
√	√	√	×	85.3	78.4	87.5	6.4×10^9
√	×	√	√	85.5	77.8	87.3	6.0×10^9
√	√	×	√	84.4	74.9	86.7	5.9×10^9
√	√	√	√	88.2	79.6	89.4	6.0×10^9

注：×表示未加入此模块；√表示加入此模块。

在各机制组合实验中，各改进方案展现出良好的兼容性，其中在 TA 机制和 FM 机制分别与 ODConv 的组合中仍能保持良好的检测精度，并显著降低模型计算量，进一步证实 ODConv 有助于模型在旱直播秧苗检测中降低浮点计算量。

本文提出的 YOLO v11n－TFO 模型的改进机制通过协同作用形成 3 级优化网络，模型平均精度均值达到 89.4%，较基准模型 YOLO v11n 提升 6.5 个百分点，整体浮点计算量较改进前减少约 4.8 %，对旱直播秧苗的检测较好地平衡了检测精度与检测效率的关系，为旱直播秧苗检测任务提供了一个较好的解决方案。

3.1.2　模型性能对比实验

对比实验可以较好地反映改进后的模型在旱直播秧苗检测任务中与同类模型的性能差异，故选择将 YOLO v11n－FTO 与 Faster R－CNN、SSD、YOLO 同类主流模型对比，对比结果如表 2 所示，通过归一化后的雷达图直观对比性能差异，如图 12 所示。在模型对比实验结果中，检测精度、召回率、平均精度均值有较为突出的表现，浮点计算量也相对较低，模型表现出较好的检测性能。在测试平台中模型的检测速度略低于其他模型，但可以满足图像秧苗植株检测速度要求。

表 2　不同模型对比实验结果

Tab. 2　Comparative experimental results of different models

模型	检测精度/%	召回率/%	平均精度均值/%	浮点计算量	检测速度/(f·s^{-1})
SSD	59.7	32.6	53.12	3.4×10^{10}	107.1
Faster R－CNN	63.9	73.1	69.5	2.0×10^{11}	18.7
YOLO v5n	80.3	69.0	79.1	4.2×10^9	461.8
YOLO v8n	78.6	72.2	80.9	8.1×10^9	408.4
YOLO v9t	76.5	70.4	79.2	1.1×10^{11}	122.4
YOLO v10n	77.7	72.5	80.8	8.2×10^9	255.9
YOLO v11n	80.6	74.5	82.9	6.3×10^9	303.6
YOLO v12n	80.1	75.1	83.0	6.3×10^9	206.6
YOLO v11n－TFO	88.2	79.6	89.4	6.0×10^9	217.9

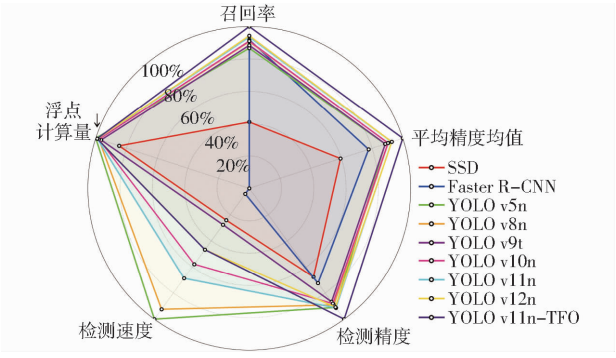


图 12　对比模型雷达图

Fig. 12　Comparison model radar image

图 13 为不同模型平均精度均值随训练轮次的变化曲线。随着训练轮次的增加，各模型的平均精度均值整体呈增长趋势，所有模型约在 300 次迭代后实现基本收敛，仅存在小幅度波动。其中改进前 YOLO v11n 模型在 200 次迭代后曲线增长变缓，较同类模型更快实现基本收敛，在旱直播秧苗检测中

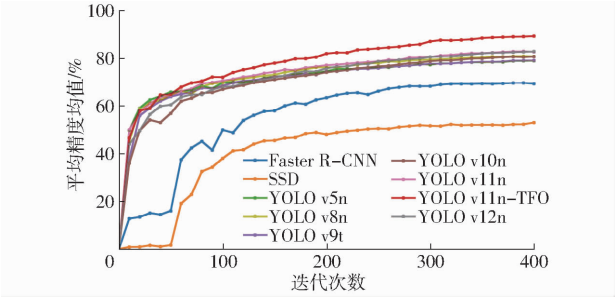


图 13　平均精度均值变化曲线

Fig. 13　mAP50 curves for each comparison model

表现出更好的学习能力，改进后的 YOLO v11n－TFO 在 100 次迭代到 200 次迭代中较改进前表现出更强的学习能力，在 400 次迭代后平均精度均值没有明显增长，实现基本收敛。对比结果表明，针对骨干网络特征提取部分的改进较同类目标检测网络，对旱直播秧苗检测任务具有一定先进性。

3.2 模型可视化

模型的可视化分析可突破深度网络中“黑箱”模型的局限性，其中热力图定位分析可以通过量化模型对目标区域的关注强度，验证改进后的模型是否抑制噪声、强化秧苗植株目标区域，同时检查改进后的 YOLO v11n－TFO 模型对秧苗植株检测的可解释性。

实验采用梯度权重类激活映射（Grad－CAM）与特征通道可视化结合的方法，对改进后的 YOLO v11n－TFO 的主干网络进行多层次性能评估。对卷积层可视化分析，验证改进的 ODConv 在浅层网络中对秧苗叶片轮廓等低级特征获取能力；对 FM 机制层与原模型 SPPF 层可视化分析，检查秧苗目标的语义信息保留与丢失情况，并验证秧苗有效特征信息关注程度；最后对引入 TA 机制的 C2PSA 模块与改进前的 C2PSA 模块可视化分析，测试改进后的主干网络对秧苗深度特征提取能力。

可视化分析结果如图 14 所示，为 YOLO v11n 同 YOLO v11n－TFO 的主干网络不同层对旱直播秧苗的热力图。对比图 14b、14e，改进后的动态卷积层对旱直播秧苗的线形特征有更好地捕获能力，可以有效区分背景与植株特征，但对秸秆同为线性干扰物特征进行捕获；对比图 14c、14f，替换 SPPF 后的 FM 机制能更好地聚焦于秧苗植株语义信息的保留，对动态卷积层捕获的信息更有针对性地保留，从全局对秧苗特征融合，但同时对部分分散目标的信息存在部分丢失；对比图 14d、14g，在引入注意力机制 TA 后的 C2PSA 模块能强化对 FM 层秧苗语义信息的理解，并对丢失的小目标信息进行有效提取，最大程度保留秧苗形态特征信息并减少对秸秆干扰物特征的关注度。

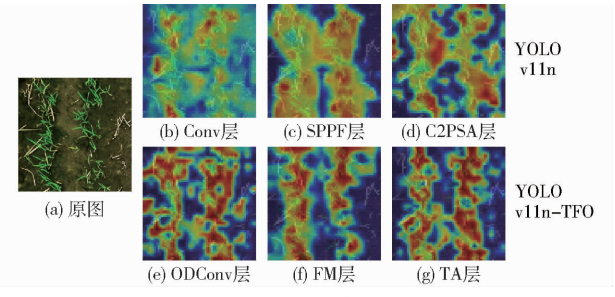


图 14 主干网络热力图

Fig. 14 Backbone network heat maps

通过模型可视化分析验证，改进后模型的主干网络在旱直播秧苗特征提取过程中表现出更集中的注意力分布，其关注区域与秧苗形态学特征匹配度显著提升，模型的可解释性增强，验证了算法改进方案的有效性。

3.3 秧苗计数平台

为了能便捷获取秧苗植株数量信息，在 YOLO v11n－TFO 旱直播秧苗检测算法基础上，开发了旱直播秧苗检测系统，采用模块化设计实现高效检测与数据管理功能，操作界面如图 15 所示。

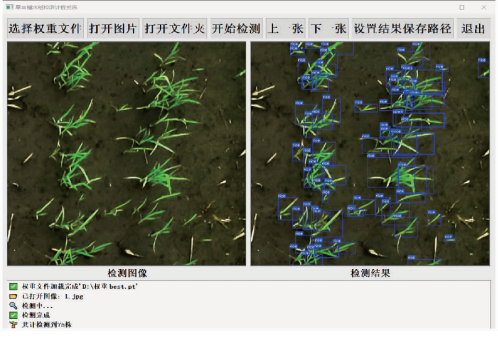


图 15 旱直播秧苗检测系统操作界面

Fig. 15 Drought live seedling detection system

利用 PyQt5 框架设计多功能的可视化界面，核心功能模块包括：①图像加载模块，用户通过“打开图片”按钮导入待检图像，并在界面左侧预览区实现图像可视化加载。②智能检测模块，点击“开始检测”后系统根据选择的预权重解析输入图像，利用改进后的 YOLO v11n－TFO 网络进行特征提取与目标识别。③结果输出模块，检测结果实时显示于界面下侧，以边界框标注形式直观呈现秧苗空间分布特征，同时在下方的操作日志窗口同步更新检测统计信息。系统自动生成包含秧苗坐标信息的分布及统计数量的文本文件，利用 UTF－8 编码格式保存至用户指定路径，并采用“时间戳＋检测参数”的格式规范文件命名，确保数据溯源性，为后续的农艺分析、种植密度计算等应用场景提供可靠数据支持。

利用测试集中的 150 幅图像对旱直播秧苗检测系统进行测试，将平台的预测值与人工计数的真实值对比，对比结果如图 16 所示，检测平台在旱直播秧苗密植合理和较为稀疏的图像检测中有较高的检测精度，对于出苗较密、相互遮挡较多的秧苗簇，预测量略高，检测精度偏低。检测平台对旱直播秧苗计数精度达到 87% 以上，与训练结果偏差较小，证明模型在旱直播秧苗检测中具有一定泛化能力与可行性。

基于预处理后的无人机图像数据，聚焦于 YOLO v11n－TFO 旱直播秧苗计数模型的优化与应

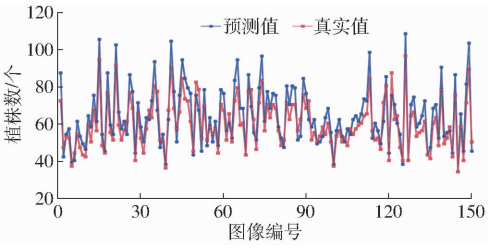


图 16　预测值和真实值对比曲线

Fig. 16　Comparison curve between predicted and actual values

用，在保证计算效率的前提下提升检测精度与可靠性。针对幼苗期田间管理需求，提出后续优化策略：采用剪枝-蒸馏联合压缩技术，通过结构化剪枝保留关键权重，降低模型复杂度，结合知识蒸馏实现大模型知识迁移，满足实时流媒体检测需求。构建元宇宙农业信息平台，融合三维点云重建与数字孪生技术，将检测结果与稻田环境数据结合，形成可视化种植管理方案，提升旱直播稻田复杂背景下的决策支持能力。

4　结论

（1）针对旱直播秧苗检测中因植株形态异质性、空间分布非均匀性导致的漏检与误判问题，提出一种 YOLO v11n－TFO 模型，实现了检测精度与计算效率的协同优化。

（2）对 YOLO v11n 模型中颈部与骨干网络 C3k2 卷积块改进 C3k2_ODConv，改善模型对于旱直播秧苗浅层表型特征捕获能力并丰富颈部网络对多尺度特征融合，优化对遮挡植株的检测能力；利用 FM 机制替代原有 SPPF 层，增强对秧苗多尺度特征信息的筛选能力；引入 TA 机制从而优化 C2PSA 模块，提高模型对旱直播秧苗的关注度和表型特征信息提取能力。

（3）改进后的 YOLO v11n 模型，平均精度均值达到 89.4%，较改进前提升 6.5 个百分点，检测精度为 88.2%，召回率为 79.6%，尽管参数量较改进前有所增加，但检测速度仍达到 217.9 f/s，且浮点计算量仅为 6.0×10^9；与 YOLO v5n、Faster R－CNN、SSD 等主流基于深度学习的目标检测模型相比，改进后的 YOLO v11n－TFO 对寒地旱直播秧苗植株的检测具有较高的检测精度和检测速度。

（4）面向 PC 设备设计基于 YOLO v11n－TFO 的旱直播秧苗检测系统，秧苗计数精度达到 87% 以上。在本地 PC 设备部署系统并集成 PYQT5 开发的可视化交互界面，通过上传检测秧苗图像，界面直观显示检测结果，并在后台保存分布信息，为稻田出苗评估与稻田农艺管理提供数据支持。

参　考　文　献

[1] 魏永霞，冀俊超，刘慧，等. 水分管理对旱直播稻温室气体排放与土壤无机氮的影响[J]. 农业机械学报，2021，52(11)：305－314.
WEI Yongxia，JI Junchao，LIU Hui，et al. Effects of water management on greenhouse gas emission and soil in organic nitrogen of dry direct seeding rice[J]. Transactions of the Chinese Society for Agricultural Machinery，2021，52(11)：305－314.（in Chinese）

[2] 魏永霞，侯景翔，吴昱，等. 不同水分管理旱直播水稻生长生理与节水效应[J]. 农业机械学报，2018，49(8)：253－264.
WEI Yongxia，HOU Jingxiang，WU Yu，et al. Effects of different water management on growth physiology and water-saving of dry direct seeding rice[J]. Transactions of the Chinese Society for Agricultural Machinery，2018，49(8)：253－264.（in Chinese）

[3] 廖萍，翁文安，高辉，等. 我国直播稻栽培技术应用现状与发展建议[J]. 中国农业科学，2024，57(24)：4854－4870.
LIAO Ping，WENG Wenan，GAO Hui，et al. Application status and development suggestion of direct-seeding rice cultivation in China[J]. Scientia Agricultura Sinica，2024，57(24)：4854－4870.（in Chinese）

[4] 薛里，张忠学，齐智娟，等. 节水灌溉下秸秆还田形式对黑土区稻田 N_2O 排放与产量的影响[J]. 农业机械学报，2024，55(4)：280－289.
XUE Li，ZHANG Zhongxue，QI Zhijuan，et al. Effects of straw returning on N_2O emission and yield under water-saving irrigation in black soil paddy field[J]. Transactions of the Chinese Society for Agricultural Machinery，2024，55(4)：280－289.（in Chinese）

[5] 李茉，薛敏，沙炎，等. 基于 Meta 分析的黑龙江省水稻水土肥资源协同优化调配[J]. 农业机械学报，2024，55(5)：302－311.
LI Mo，XUE Min，SHA Yan，et al. Synergistic and optimal allocation of water land and fertilizer resources of rice in heilongjiang province based on Meta-analysis[J]. Transactions of the Chinese Society for Agricultural Machinery，2024，55(5)：302－311.（in Chinese）

[6] 聂堂哲，张忠学，齐智娟，等. 1960—2015 年黑龙江省水稻需水量时空分布特征[J]. 农业机械学报，2019，50(5)：279－290.
NIE Tangzhe，ZHANG Zhongxue，QI Zhijuan，et al. Spatial and temporal distribution characteristics of rice water requirement in Heilongjiang Province during 1960—2015[J]. Transactions of the Chinese Society for Agricultural Machinery，2019，50(5)：279－290.（in Chinese）

[7] 刘双喜，张玮平，胡宪亮，等. 基于改进 YOLO v8s 的水稻种植机械作业质量检测[J]. 农业机械学报，2024，55(增刊

1):61 - 70.
LIU Shuangxi, ZHANG Weiping, HU Xianliang, et al. Rice planting machinery operation quality detection based on improved YOLO v8s[J]. Transactions of the Chinese Society for Agricultural Machinery, 2024, 55(Supp. 1): 61 - 70. (in Chinese)
[8] 朱文静,戴世元,冯展康,等. 基于混合像元分解的分蘖期水稻基本苗数量估测方法研究[J]. 农业机械学报, 2024, 55(6): 202 - 209.
ZHU Wenjing, DAI Shiyuan, FENG Zhankang, et al. Estimation of rice basic seedling number based on mixed pixel decomposition[J]. Transactions of the Chinese Society for Agricultural Machinery, 2024, 55(6): 202 - 209. (in Chinese)
[9] 汤颖辉. 小麦出苗质量快速监测技术及系统构建[D]. 南京:南京农业大学,2022.
TANG Yinghui, Research and development of rapid monitoring technology and system for wheat emergance quality[D]. Nanjing: Nanjing Agricultural University, 2022. (in Chinese)
[10] BAI Y, NIE C, WANG H, et al. A fast and robust method for plant count in sunflower and maize at different seedling stages using high-resolution UAV RGB imagery[J]. Precision Agriculture, 2022, 23(5): 1720 - 1742.
[11] 张立杰,周舒骅,李娜,等. 基于改进 SSD 卷积神经网络的苹果定位与分级方法[J]. 农业机械学报, 2023, 54(6): 223 - 232.
ZHANG Lijie, ZHOU Shuhua, LI Na, et al. Apple location and classification based on improved SSD convolutional neural network[J]. Transactions of the Chinese Society for Agricultural Machinery, 2023, 54(6): 223 - 232. (in Chinese)
[12] 宋怀波,尚钰莹,何东健. 果实目标深度学习识别技术研究进展[J]. 农业机械学报, 2023, 54(1): 1 - 19.
SONG Huaibo, SHANG Yuying, HE Dongjian. Review on deep learning technology for fruit target recognition[J]. Transactions of the Chinese Society for Agricultural Machinery, 2023, 54(1): 1 - 19. (in Chinese)
[13] 张志远,罗铭毅,郭树欣,等. 基于改进 YOLO v5 的自然环境下樱桃果实识别方法[J]. 农业机械学报, 2022, 53(增刊 1): 232 - 240.
ZHANG Zhiyuan, LUO Mingyi, GUO Shuxin, et al. Cherry fruit detection method in natural scene based on improved YOLO v5 [J]. Transactions of the Chinese Society for Agricultural Machinery, 2022, 53(Supp. 1): 232 - 240. (in Chinese)
[14] 朱红春,李旭,孟炀,等. 基于 Faster R - CNN 网络的茶叶嫩芽检测[J]. 农业机械学报, 2022, 53(5): 217 - 224.
ZHU Hongchun, LI Xu, MENG Yang, et al. Tea bud detection based on Faster R - CNN network[J]. Transactions of the Chinese Society for Agricultural Machinery, 2022, 53(5): 217 - 224. (in Chinese)
[15] VELUMANI K, LOPEZ-LOZANO R, MADEC S, et al. Estimates of maize plant density from UAV RGB images using Faster - RCNN detection model: impact of the spatial resolution[J]. Plant Phenomics, 2021, 2021: 9824843.
[16] 张昌松,李伟. 基于 YOLO v5 的直播油菜激光间苗系统设计与试验[J]. 农业机械学报, 2024, 55(4): 40 - 52.
ZHANG Changsong, LI Wei. Research and experiment of direct seeding rape laser seedling system based on YOLO v5[J]. Transactions of the Chinese Society for Agricultural Machinery, 2024, 55(4): 40 - 52. (in Chinese)
[17] 李春鹏. 基于改进卷积神经网络的苗期油菜植株计数算法研究[D]. 合肥:安徽农业大学,2023.
LI Chunpeng. Research on seedling rape plant counting algorithm based on improved neural network[D]. Hefei: Anhui Agricultural University, 2023. (in Chinese)
[18] 李兆东,章艳芳,汪蕴红,等. 基于 YC - YOLO v7 模型的油菜幼苗株数识别方法[J]. 农业机械学报, 2024, 55(12): 322 - 332.
LI Zhaodong, ZHANG Yanfang, WANG Yunhong, et al. Identification of rapeseed seedling number based on YC - YOLO v7 model[J]. Transactions of the Chinese Society for Agricultural Machinery, 2024, 55(12): 322 - 332. (in Chinese)
[19] LIN Z, GUO W. Cotton stand counting from unmanned aerial system imagery using mobilenet and centernet deep learning models[J]. Remote Sensing, 2021, 13(14): 2822.
[20] BARRETO A, LOTTES P, YAMATI F R I, et al. Automatic UAV-based counting of seedlings in sugar-beet field and extension to maize and strawberry[J]. Computers and Electronics in Agriculture, 2021, 191: 106493.
[21] LU H, LIU L, LI Y N, et al. TasselNetV3: explainable plant counting with guided upsampling and background suppression [J]. IEEE Transactions on Geoscience and Remote Sensing, 2021, 60: 1 - 15.
[22] 朱启兵,张梦,刘振方,等. 基于点云配准的盆栽金桔果实识别与计数方法[J]. 农业机械学报, 2022, 53(5): 209 - 216.
ZHU Qibing, ZHANG Meng, LIU Zhenfang, et al. Identification and counting method of potted kumquat fruits based on point cloud registration[J]. Transactions of the Chinese Society for Agricultural Machinery, 2022, 53(5): 209 - 216. (in Chinese)
[23] 杜小强,李卓林,马锃宏,等. 基于空间注意力和可变形卷积的无人机田间障碍物检测[J]. 农业机械学报, 2023, 54(2): 275 - 283.
DU Xiaoqiang, LI Zhuolin, MA Zenghong, et al. UAV field obstacle detection based on spatial attention and deformable convolution[J]. Transactions of the Chinese Society for Agricultural Machinery, 2023, 54(2): 275 - 283. (in Chinese)
[24] 林珊玲,张雪,陈燕,等. 融合全局信息与双域注意力机制的光学遥感飞机目标检测[J]. 光学精密工程, 2024, 32(20): 3085 - 3098.
LIN Shanling, ZHANG Xue, CHEN Yan, et al. Integrating global information and dual-domain attention mechanism for optical remote[J]. Optics and Precision Engineering, 2024, 32(20): 3085 - 3098. (in Chinese)
[25] 李彬,李生林. 改进 YOLOv11n 的无人机小目标检测算法[J]. 计算机工程与应用, 2025, 61(7): 96 - 104.
LI Bin, LI Shenglin. Improved YOLOv11n small object detection algorithm in UAV view[J]. Computer Engineering and

Applications, 2025, 61(7): 96 - 104. (in Chinese)
[26] 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)
[27] TAN S, KUANG Z, JIN B. AppleYOLO: apple yield estimation method using improved YOLO v8 based on deep OC - SORT [J]. Expert Systems with Applications, 2025, 272: 126764.
[28] TAO H. A triple interference removal network based on temporal and spatial attention interaction for forest smoke recognition in videos[J]. Computers and Electronics in Agriculture, 2024, 218: 108756.
[29] LI C, ZHOU A, YAO A. Omni-dimensional dynamic convolution[J]. arXiv Preprint, arXiv: 2209.07947, 2022.
[30] ZHAO L, FU L, JIA X, et al. YOLO - BOS: An emerging approach for vehicle detection with a novel BRSA mechanism[J]. Sensors, 2024, 24(24): 8126.
[31] 杨断利,齐俊林,陈辉,等. 基于改进 YOLO v8n 模型的散养蛋鸡个体行为识别方法与差异分析[J]. 农业机械学报, 2024, 55(11): 112 - 123.
YANG Duanli, QI Junlin, CHEN Hui, et al. Individual behavioral identification and differential analysis of free-range laying hens based on improved YOLO v8n model[J]. Transactions of the Chinese Society for Agricultural Machinery, 2024, 55(11): 112 - 123. (in Chinese)
[32] YANG J, LI C, DAI X, et al. Focal modulation networks[J]. Advances in Neural Information Processing Systems, 2022, 35: 4203 - 4217.
[33] LIU Y, HU Y, CAO G, et al. SENet: super-resolution enhancement network for crowd counting[J]. Pattern Recognition, 2025, 162: 111420.
[34] FU H, WANG C, CHEN Z. Adaptive selection of spectral-spatial features for hyperspectral image classification using a modified-CBAM-based network[J]. Neurocomputing, 2025, 615: 128877.
[35] WU J, SUN M, JIANG C, et al. Context-based local-global fusion network for 3D point cloud classification and segmentation [J]. Expert Systems With Applications, 2024, 251: 124023.
[36] LI G, ZHAO T. Efficient image analysis with triple attention vision transformer[J]. Pattern Recognition, 2024, 150: 110357.
[37] 王芯蕊,王慧琴,王可,等. 基于改进 YOLO v8 的无人机视角下高速公路异常目标检测方法[J]. 计算机工程与应用, 2025, 61(11): 105 - 118.
WANG Xinrui, WANG Huiqin, WANG Ke, et al. Anomalous target detection method for highway from UAV viewpoint based on improved YOLO v8[J]. Computer Engineering and Applications, 2025, 61(11): 105 - 118. (in Chinese)

(上接第 575 页)

[29] YANG H, MING B, NIE C, et al. Maize canopy and leaf chlorophyll content assessment from leaf spectral reflectance: estimation and uncertainty analysis across growth stages and vertical distribution[J]. Remote Sensing, 2022, 14(9): 2115.
[30] ANJUM S, XIE X, WANG L, et al. Morphological, physiological and biochemical responses of plants to drought stress[J]. African Journal of Agricultural Research, 2011, 6(9): 2026 - 2032.
[31] 孙杰,杨静,丁书杰,等. 基于多注意力机制与编译图神经网络的高光谱图像分类[J]. 农业机械学报, 2024, 55(3): 183 - 192, 212.
SUN Jie, YANG Jing, DING Shujie, et al. Hyperspectral image classification based on multi-attention mechanism and compiled graph neural networks[J]. Transactions of the Chinese Society for Agricultural Machinery, 2024, 55(3): 183 - 192, 212. (in Chinese)
[32] 马文勇,王训明. 基于高光谱分析的草地叶绿素含量估算研究进展[J]. 地理科学进展, 2016, 35(1): 25 - 34.
MA Wenyong, WANG Xunming. Progress on grassland chlorophyll content estimation by hyperspectral analysis[J]. Progress in Geography, 2016, 35(1): 25 - 34. (in Chinese)
[33] HAN L, YANG G, YANG X, et al. An explainable XGBoost model improved by SMOTE - ENN technique for maize lodging detection based on multi-source unmanned aerial vehicle images[J]. Computers and Electronics in Agriculture, 2022, 194: 106804.
[34] 王丽萍,张新明,李继伟,等. BP 网络激活函数选择及在径流预报模型中的应用[J]. 水力发电学报, 2014, 33(1): 29 - 36.
WANG Liping, ZHANG Xinming, LI Jiwei, et al. BP neural network activation function's selection and application to runoff forecasting model[J]. Journal of Hydroelectric Engineering, 2014, 33(1): 29 - 36. (in Chinese)
[35] 徐睿,梁循,齐金山,等. 极限学习机前沿进展与趋势[J]. 计算机学报, 2019, 42(7): 1640 - 1670.
XU Rui, LIANG Xun, QI Jinshan, et al. Advances and trends in extreme learning machine[J]. Chinese Journal of Computers, 2019, 42(7): 1640 - 1670. (in Chinese)