doi:10.6041/j.issn.1000-1298.2026.22.029

基于改进 YOLO v12 与主方向压缩聚类的水稻苗带线提取

王姗姗[1]　李东方[2]　张天慈[1]　游祥龙[1]　陶倩文[1]　陈学深[3]
(1. 南京林业大学汽车与交通工程学院, 南京 210037; 2. 南京农业大学工学院, 南京 211800;
3. 华南农业大学工程学院, 广州 510642)

摘要: 水稻秧苗行精准识别对提升水田除草机械的自主导航与对行作业精度至关重要。针对水田中普遍存在的浮萍覆盖、杂草干扰、秧苗缺失及光照不均等问题,本文提出了一种融合目标检测与主方向压缩聚类的秧苗行识别算法。首先构建了集成多尺度门控注意力和轻量化共享检测头的水稻秧苗检测模型(RSD),有效增强了对秧苗上下文特征提取能力及对复杂背景的抑制能力。进而,构建基于方向性压缩迭代聚类策略,通过自适应地弥合同行秧苗点间纵向间隙,解决传统方法在缺苗场景下苗行断裂问题,实现了从离散检测点到连续苗带线稳健拟合。试验结果表明,秧苗检测中,RSD 模型 $AP_{0.5}$ 达 92.5%,检测速度为 104.2 f/s,性能优于主流对比模型;在苗带线提取中,该算法在不同缺苗情况下行断裂率与误合并率均低于 1%,平均识别角度偏差仅为 0.203°,试验结果充分验证了算法性能优势。

关键词: 水稻秧苗行识别; YOLO v12; 主方向压缩; 迭代聚类; 秧苗缺失
中图分类号: TP391.4　　**文献标识码**: A　　**文章编号**: 1000-1298(2026)22-0339-10　　**OSID**:

Seedling Band Line Extraction Method Based on Improved YOLO v12 and Clustering with Principal Direction Compression

Wang Shanshan[1]　Li Dongfang[2]　Zhang Tianci[1]　You Xianglong[1]　Tao Qianwen[1]　Chen Xueshen[3]
(1. *School of Automotive and Transportation Engineering*, *Nanjing Forestry University*, *Nanjing* 210037, *China*
2. *College of Engineering*, *Nanjing Agricultural University*, *Nanjing* 211800, *China*
3. *College of Engineering*, *South China Agricultural University*, *Guangzhou* 510642, *China*)

Abstract: Accurate recognition of rice seedling rows is crucial for improving the autonomous navigation and row-following precision of weeding machinery in paddy fields. To solve the problems in paddy fields, such as duckweed coverage, weed interference, missing seedlings, and uneven illumination, a seedling row recognition algorithm that integrated object detection with principal direction compression clustering was proposed. The algorithm firstly constructed a rice seedling detection model, which combined a multi-scale gated attention mechanism with a lightweight shared detection head, effectively enhancing the extraction of seedling contextual features and the suppression of complex backgrounds. Furthermore, an iterative clustering strategy based on directional compression was developed, which adaptively bridged the longitudinal gaps between seedling points within the same row, thereby solving the problem of row breakage faced by traditional methods in seedling gap scenarios and achieving robust fitting from discrete detection points to continuous seedling band lines. Experimental results demonstrated that the proposed RSD model achieved an $AP_{0.5}$ of 92.5% and a detection speed of 104.2 f/s in the seedling detection task, outperforming mainstream comparison models. The seedling band line extraction method achieved an average angular deviation of 0.203°, with both the row breakage and mis-merging rates below 1% across various seedling gap conditions. Comparative experiments fully validated the performance advantages of this method.

Key words: recognition of rice seedling row; YOLO v12; directional compression; iterative clustering; seedling gaps

收稿日期: 2026-05-11　修回日期: 2026-05-30
基金项目: 江苏省科技计划项目(BE2021304)
作者简介: 王姗姗(1992—),女,讲师,主要从事农业机器人作业感知研究,E-mail: wss92me@163.com
通信作者: 陈学深(1977—),男,教授,主要从事智能农机装备研究,E-mail: chenxs@scau.edu.cn

0 引言

人工除草和喷洒除草剂是水稻田常用的除草方法[1]。长期大量喷洒除草剂会导致水稻农药残留超标、生态链破坏和环境污染等诸多问题[2-5],并因杂草抗药性日益增强而进一步削弱防效[6-8]。人工除草存在劳动强度大、效率低、成本高等问题[9-10]。相比之下,机械除草作为一种环境友好型除草方法,在有效清除杂草的同时,可疏松土壤、增强通透性、提高肥料利用率并促进水稻根系生长,具有良好的生态与农艺效益[11]。作物行精准识别是提高机械除草对行作业精度的关键前提,同时也为精准施肥、变量施药等现代农业精准作业体系提供了必要的信息感知与决策依据。

早期研究主要依赖传统图像处理技术,通过颜色、几何或光谱特征分离作物与背景,再利用聚类算法将空间分布上属于同一行点聚合,最后进行苗带线拟合。孟笑天等[12]以苗期玉米为研究对象,采用改进超绿法和 Otsu 法进行图像预处理,利用改进后的 K 均值聚类算法对玉米苗特征点聚类,最后用最小二乘法进行直线拟合,识别率达 91%,但聚类数目受杂草影响较大。王爱臣等[13]提出基于区域生长与均值漂移的聚类方法,5 种作物平均行识别率为 98.18%,平均误差角度为 1.21°。傅灯斌等[14]利用顺序聚类结合 RANSAC 算法,实现了水稻苗带中心线的鲁棒检测,平均误差角度为 2.34°。

随着人工智能发展,基于深度学习的方法迅速成为研究热点,并衍生出目标检测与语义分割 2 条主要技术路线[15]。目标检测路线(如 YOLO 系列模型)通过定位单株作物拟合作物行,在效率与实用性上优势显著[16-19]。Huang 等[20]提出了一种两阶段深度定位网络,用于稻田中单株水稻幼苗检测,定位精度达 93.22%。Wei 等[21]利用基于行锚选择的轻量级 GhostNet 网络,在生长早期作物行识别中准确率达 97.90%。Wang 等[22]通过改进 CS_YOLO v5 网络定位秧苗,并设计一种基于初始聚类和子区域生长的直线-弯曲秧苗行识别算法,平均识别角度偏差小于 1°。此外,Tang 等[23]在 YOLO v5 中引入自适应锚框与全局注意力机制,在飞行高度 15、30 m 下,玉米幼苗平均检测精度为 94.5% 和 88.2%,每幅图像平均检测时间为 0.025 s。基于语义分割方法将图像逐像素分类为作物区域与背景,再通过后处理提取苗带线[24-25]。Li 等[26]采用基于 Transformer 的语义分割模型识别生长阶段早期的水稻秧苗根茎的形态学掩模,提出了一种基于动态搜索方向的聚类算法,将稀疏分布的锚点分组成行,同时完成秧苗行拟合,检测精度达 92.93%。Liao 等[27]提出一种条带语义分割网络 SC-Net,构建了并行多级卷积块(PMCB)和条带多级卷积块(SMCB),用于设计分割网络编码器和解码器,平均识别精度达 87.48%。为克服语义分割依赖后处理、难以区分相邻作物行的局限,Chen 等[28]提出了双路径实例分割网络(TP-ISN),通过并行语义分割分支和像素嵌入分支,直接为每个独立的苗带生成实例掩膜。然而,这类像素级密集预测方法模型复杂,其训练依赖大量精细到像素级别的数据标注。

上述研究多集中于背景相对简单、作物连续分布均匀的旱田或早期生长环境。当应用于水田环境时,现有检测模型易受水面反光、秧苗缺失、浮萍及杂草影响,导致秧苗定位精度下降。此外,由缺苗、遮挡造成的秧苗点分布离散问题,使得基于传统聚类方法的行线拟合方法在断裂处极易失效,难以生成完整、连续的苗带线。本文在 YOLO v12 网络中集成一种多尺度门控注意力模块和轻量化共享检测头,以增强模型对秧苗细长结构的判别能力,实现精度与轻量化有效平衡。提出一种基于主方向压缩的聚类优化策略,通过自适应缩减同苗带特征点沿主方向间距,克服缺苗造成的苗带断裂,以显增强缺苗场景下苗带线提取完整性与鲁棒性。

1 材料与方法

1.1 图像采集

在江苏省太仓市水稻试验区采集图像,试验田总面积约 2 hm^2。选择机插秧移栽稻作为研究对象,移栽稻田主要有 2 个杂草萌发高峰期:栽后 8~12 d 和栽后 25~30 d。在机插秧后 15 d 和 25 d 左右进行水稻秧苗行图像采集,在晴天和阴天情况下,在每块试验田不同位置(随机杂草分布)进行图像采集。第 1 个杂草萌发期秧苗高度在 20 cm 左右,第 2 个杂草萌发期秧苗高度在 30 cm 左右,秧苗行间距为 30 cm,秧苗株距为 10~13 cm。

采用 Intel Realsense D455 型相机作为图像采集设备,摄像头安装位置距离地面约 120 cm,摄像头光轴与水平线夹角约为 50°,在机器人以速度 0.5 m/s 行进过程中,对秧苗进行实时图像采集(图 1),图像分辨率为 640 像素×480 像素。共采集水稻秧苗图像 7 055 幅,整个数据集划分为训练集(4 405 幅)、验证集(1 101 幅)与测试集(1 549 幅)。数据集覆盖浮萍蓝藻、水面反光、水层土壤及不同杂草密度等典型田间场景。其中,浮萍与秧苗颜色相似易造成视觉混淆,水面反光则导致图像局部过曝与细节模糊,二者均严重影响秧苗目标检测。同时,

为评估杂草干扰影响，根据杂草与秧苗数量比例，将图像划分为高杂草密度（杂草数量大于等于秧苗数量 5%）与低杂草密度（杂草数量小于秧苗数量 5%）。采集的水稻秧苗图像示例如图 2 所示，其中图 2a 为大量浮萍场景，秧苗行有缺失情况，作物行不连续；图 2b 为水田反光场景，秧苗图像模糊；图 2c 为高杂草密度场景，杂草较多增加识别难度；图 2d 为低杂草密度场景。



图 1　水稻秧苗图像动态采集

Fig. 1　Dynamic acquisition of rice seedling images

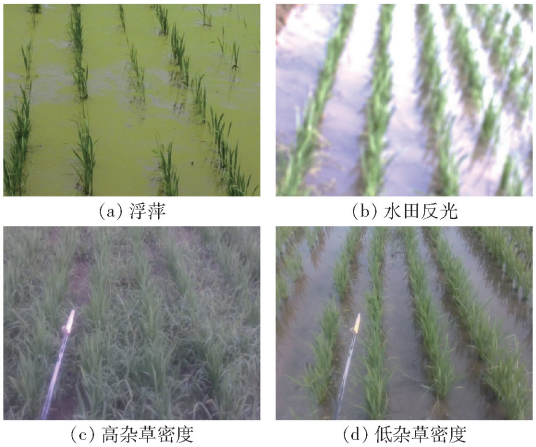


图 2　水稻秧苗图像示例

Fig. 2　Representative images of rice seedling rows

1.2　水稻秧苗检测模型

1.2.1　网络整体架构

以 YOLO v12 为基线模型，提出多尺度门控注意力模块（Multi-scale gated attention module，MSGA）和轻量化共享检测头（Lightweight shared detection head，LSDH），构建水稻秧苗检测模型 RSD。

该网络采用经典的端到端架构，由 Backbone、Neck 和 Head 3 部分组成，整体结构如图 3 所示。Backbone 通过 Conv、C3k2 和 A2C2f 等模块逐步提取多尺度特征。其中，C3k2 模块通过跨阶段部分连接（CSP）增强特征传递效率，结合 Bottleneck 残差结构缓解深层网络梯度消失问题，确保底层秧苗特征能够有效传递到高层；A2C2f 模块集成区域注意力机制，在大感受野下捕捉全局上下文信息。Neck 作为特征融合的关键环节，其性能直接影响最终检测精度。将 MSGA 模块嵌入 Neck 的 P3 层 A2C2f 模块中，通过多尺度卷积和双路径门控精炼增强特征的空间上下文信息。网络检测头 Head 接收来自 Neck 的 3 个尺度特征图，分别对应输入图像 8 倍、16 倍和 32 倍下采样。为降低计算开销，设计了轻量化共享检测头 LSDH，通过参数共享在 3 个尺度上统一检测结构，显著减少模型参数量与推理延迟。

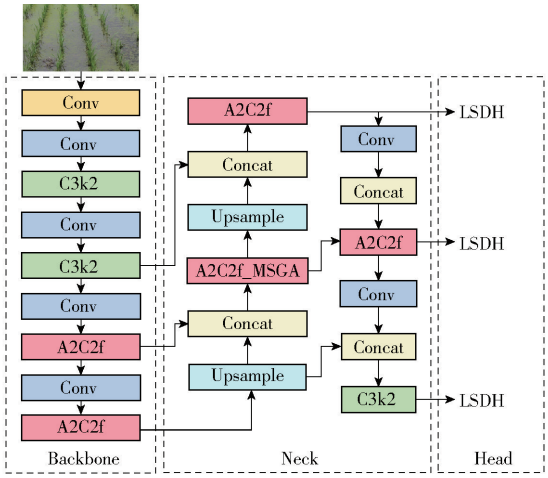


图 3　整体网络结构

Fig. 3　Overall network structure

1.2.2　多尺度门控注意力模块 MSGA

针对传统目标检测网络在复杂背景下对全局上下文与局部细节的联合建模能力不足的问题，提出多尺度门控注意力模块 MSGA，通过多尺度特征提取、通道注意力校准与双路径门控精炼的协同作用，增强模型对秧苗特征的判别性能。模块结构如图 4 所示。

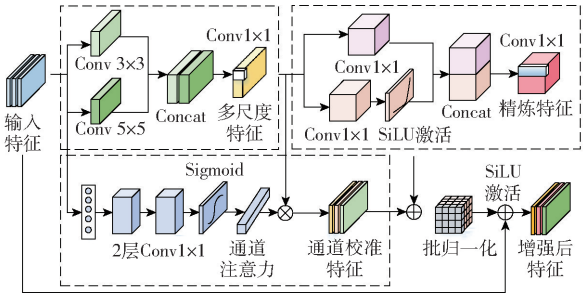


图 4　MSGA 模块结构

Fig. 4　Structure diagram of MSGA module

为捕捉不同粒度秧苗特征，该分支并行使用 3×3 和 5×5 2 尺寸深度可分离卷积核。3×3 卷积关注局部细节，5×5 卷积感知更大范围的上下文信息，两者互补形成对秧苗目标的多层次空间表达。2 个尺度特征经拼接后通过 1×1 卷积进行融合，得到多尺度空间特征。

通道注意力校准对原始输入特征进行全局上下文建模，通过全局平均池化聚合空间信息，经 2 层 1×1 卷积与 Sigmoid 激活函数生成通道注意力权重。该权重与多尺度空间特征逐通道相乘，实现全局上下文对局部多尺度特征动态校准。

双路径门控精炼包含 2 条路径,直接映射路径通过 1×1 卷积保留线性特征传递,门控残差路径在 1×1 卷积后引入 SiLU 激活函数,增强非线性表达能力。2 条路径输出特征经拼接后通过 1×1 卷积进行融合,实现对原始多尺度空间特征门控式特征筛选与重组。

通道校准后多尺度特征与双路径精炼特征通过逐元素相加进行融合,经批归一化和 SiLU 激活后,与模块原始输入进行残差连接,在保持特征图尺寸不变的前提下实现空间上下文特征显著增强。

1.2.3 轻量化共享检测头 LSDH

轻量共享检测头 LSDH 结构如图 5 所示。为降低检测头参数量,LSDH 采用跨尺度权重共享策略,同时引入尺度独立的组归一化层,以适配不同层级特征的分布差异。

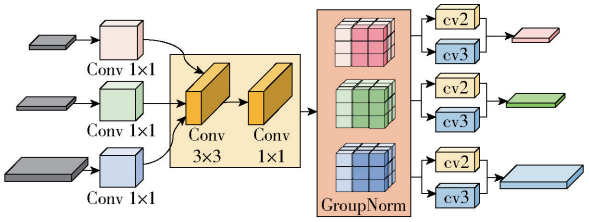


图 5 LSDH 模块结构

Fig. 5 Structure diagram of LSDH module

LSDH 首先对颈部网络输出 3 尺度特征进行统一通道处理,即各尺度特征经独立的 1×1 卷积映射至 256 维统一通道空间,以消除尺度间通道维度差异。统一后的特征依次通过共享的 3×3 卷积和 1×1 卷积处理,前者用于提取局部空间信息,后者用于融合跨通道特征,所有卷积参数在 3 个尺度间共用。考虑到不同尺度特征图具有不同的统计特性,共享卷积的输出由各尺度独立的 GroupNorm 层进行归一化,以保持各层级特征特异性。最后,归一化特征分别送入回归分支和分类分支,生成边界框参数与类别概率。

1.3 苗带线提取算法

1.3.1 整体思想

通过水稻秧苗检测模型获得预测秧苗位置后,需要对秧苗检测框包含秧苗特征点进行聚类分组,以获取各秧苗行点集用于中心线拟合。提出了一种基于方向性压缩与迭代聚类的苗带线提取方法,通过对特征点集沿主方向进行空间压缩,主动缩小同行秧苗点之间的欧氏距离,从而促使断裂的秧苗行在聚类过程中合并为完整苗带。

图 6 为苗带聚类原理。由图 6a 可知,压缩前原始秧苗特征点因田间缺苗形成断裂带,导致欧氏聚类将每行误分为多个子类,如绿色行被分为簇 A 和簇 B;由图 6b 可知,通过对特征点集沿主特征向量方向进行空间压缩,显著缩短了同行断裂点间的纵向距离,而保持行间横向距离不变;由图 6c 可知,在相同阈值下,压缩后同行点因距离满足阈值而被成功合并为完整苗带聚类,为后续苗带中心线拟合提供完整而准确的点集。

1.3.2 中心线拟合

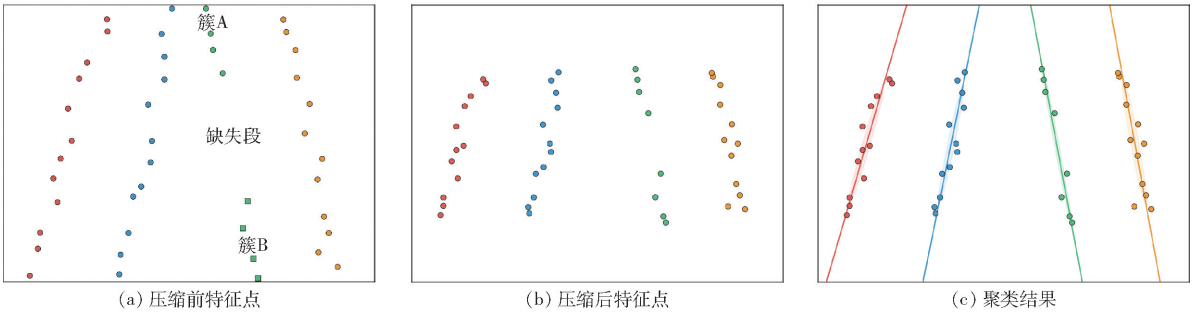
该方法是一个迭代过程,中心线拟合主要包括特征点预处理、初始欧氏聚类、秧苗行主方向提取、全局主方向判定、特征点方向性压缩与重聚类、以及迭代滤除 6 个核心步骤,最终输出完整的秧苗行聚类用于中心线拟合。

(1)特征点预处理

创建大小为 $B \times B$ 的网格滤波对象,B 为网格边长像素数。经网格滤波处理后,网格内所有点均由该网格重心点表示。对原始秧苗特征点进行下采样,使初始特征点分布变得稀疏,减少算法处理时间,提高运行速度。

(2)初始欧氏聚类

特征点初始聚类过程具体为:选取图像下半部分的一个特征点 K 作为初始点,基于最近邻搜索算法获取距离该点 K 最近的 m 个特征点;若这 m 个最近点中任意一点与点 K 欧氏距离小于设定阈值,则将该点归入聚类 Q 中(此处距离阈值设定为 d_T);在聚类 Q 中,选取除点 K 外某一个点作为新初始点 K,重复新的聚类过程;当聚类 Q 中点数不再变化时,初始聚类过程结束。

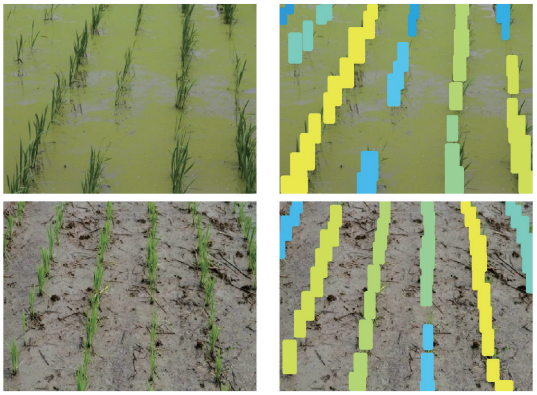


(a) 压缩前特征点 (b) 压缩后特征点 (c) 聚类结果

图 6 苗带聚类合并示意图

Fig. 6 Schematics of seedling band clustering based on directional compression and iterative clustering

基于欧氏距离的初始聚类结果如图 7 所示。其中，不同类别的秧苗点以不同颜色区分。由图 7 可以看出，在秧苗存在缺失的区域，原本应属于同一秧苗行的点群可能被划分为多个类别。因此后续需对同一行秧苗分散聚类进行合并处理，以得到完整秧苗行。



(a) 原始图像　　(b) 初始聚类图像

图 7　水稻秧苗图像初始聚类结果

Fig. 7　Initial clustering results of rice seedling images

(3)秧苗行主方向提取

在初始聚类形成的各个聚类中，首先计算每个类坐标均值(x_{im}, y_{im})，然后将各类中每个特征点坐标减去均值即可获得去中心化秧苗坐标(x_{ij}, y_{ij})；在计算各类协方差矩阵后，通过奇异值分解(SVD)可得到协方差矩阵特征值与特征向量，其中各类最大与最小特征值分别记为 λ_{i1} 和 λ_{i2}，由于秧苗点坐标为二维数据，与最大特征值 λ_{i1} 对应的特征向量表示平行于秧苗行方向向量，与最小特征值 λ_{i2} 对应的特征向量则表示垂直于秧苗行方向向量，这 2 个特征向量共同构成 1 个 2×2 特征向量矩阵 $\boldsymbol{M}_i$。

(4)全局主方向判定

首先计算每个聚类秧苗行置信度 r_i，计算式为

$$r_i = \frac{\lambda_{i1}}{\lambda_{i2}} \tag{1}$$

比值越大，表明该聚类的点分布越接近直线，其指示方向越可靠。

选取置信度最高的聚类，将其特征向量 $\boldsymbol{M}_i$ 定义为本轮迭代的全局主方向 $\boldsymbol{M}$，最大置信度 r_i 对应的秧苗聚类记为 i_m，并记录该聚类中特征点原始索引。

(5)特征点方向性压缩与重聚类

假设秧苗特征点的坐标为(x, y)，将秧苗特征点投影到以主特征向量作为基的坐标系 V，投影坐标为(a, b)，两者之间投影关系为

$$(x, y)^{\mathrm{T}} = \boldsymbol{M}(a, b)^{\mathrm{T}} \tag{2}$$

由式(2)可知，秧苗特征点在坐标系 V 下投影(a, b)为 $\boldsymbol{M}^{-1}(x, y)^{\mathrm{T}}$；在坐标系 V 下对投影进行压缩，设压缩对角阵为 $\boldsymbol{A} = \begin{bmatrix} \mu & 0 \\ 0 & 1 \end{bmatrix}$，其中 μ 表示沿主特征向量方向压缩比，压缩后秧苗点可表示为 $\boldsymbol{AM}^{-1}(x, y)^{\mathrm{T}}$；最后将压缩后秧苗点投影回原始坐标系，压缩后秧苗特征点投影坐标为(x_c, y_c)，计算式为

$$\begin{bmatrix} x_c \\ y_c \end{bmatrix} = \boldsymbol{MAM}^{-1} \begin{bmatrix} x \\ y \end{bmatrix} \tag{3}$$

采用和初始聚类相同的距离阈值，对压缩后特征点集进行重聚类。通过沿主特征向量方向的压缩变换，使得原本因间隔过大而离散的、属于同一苗带的特征点在压缩空间中满足聚类条件。因此，此次重聚类能够将离散子类有效合并为完整苗带。最后，记录重聚类结果中各点在原始点集中的索引，以建立与原始数据映射关系。

(6)迭代滤除

通过遍历压缩后的秧苗聚类中每类特征点索引，当压缩前和压缩后秧苗聚类中具有相同索引的秧苗点数量超过 T 后，即认为压缩前聚类 i_m 对应压缩后聚类 i_{mc}，将其判定为一个完整秧苗行，并根据索引从原始特征点集中滤除这些点。将滤除后剩余特征点作为新的输入，重复执行步骤(2)～(6)。当剩余特征点集数量为 0 时，循环终止。所有被滤除的独立点集即为最终秧苗行聚类结果，如图 8 所示。由图 8 可知，对于缺失秧苗行，沿主特征向量方向压缩后，原本因缺失而分散于不同类别的同列秧苗点，最终被有效合并至同一聚类中，并以统一颜色呈现，形成完整的秧苗行结构。

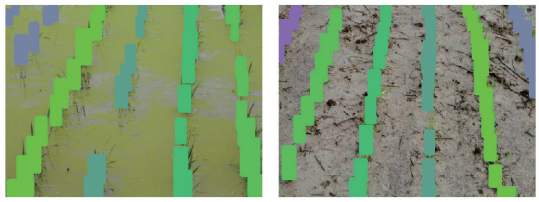


图 8　秧苗行最终聚类结果

Fig. 8　Final clustering results of rice seedling images

基于方向性压缩与迭代聚类的苗带线提取方法的整体流程如图 9 所示。

2　试验结果与分析

2.1　试验设置

试验硬件环境为 AMD Ryzen 7 4800HS 处理器(8 核 16 线程，2.90 GHz)、NVIDIA GeForce RTX 2060 显卡(6 GB GDDR6)和 16 GB DDR4 内存，软件环境在 Ubuntu 20.04 系统上采用 CUDA 11.8、Python 3.8 和 Pytorch 2.0 搭建深度学习框架。

为了充分利用硬件资源，在不超过 GPU 内存限制的前提下，设置图像批量大小为 16，网络初始学

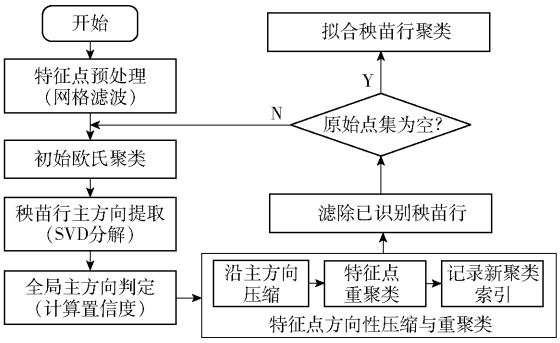


图 9 基于方向性压缩与迭代聚类的苗带线提取算法流程

Fig. 9 Flowchart of seedling band line extraction based on Euclidean clustering and principal direction compression

习率设置为 0.001,权重衰减系数设置为 0.000 5,动量系数为 100。

2.2 秧苗检测试验

设计逐层递进的 3 组试验。首先,通过消融试验验证各模块对秧苗检测性能的贡献;其次,开展不同检测模型对比试验,在公开基准上证明 RSD 模型相较于主流 YOLO 系列模型的综合性能优势;最后,进行场景适应性试验,在复杂田间环境(如浮萍、反光、杂草干扰)专用测试集上,进一步验证模型鲁棒性。

2.2.1 消融试验

为验证提出的 MSGA 模块和 LSDH 检测头有效性,基于 YOLO v12n 基线模型在自建秧苗数据集上开展消融试验,试验结果如表 1 所示。

表 1 不同模块消融试验结果

Tab. 1 Ablation experiment results of different modules

MSGA	LSDH	准确率/%	召回率/%	$AP_{0.5}$/%
–	–	86.6	83.1	91.5
√	–	87.3	83.9	91.9
–	√	87.0	83.8	91.8
√	√	87.7	84.2	92.5

注:√表示采用此模块,–表示未采用此模块。

由表 1 可知,基线模型准确率、召回率和 $AP_{0.5}$ 分别为 86.6%、83.1% 和 91.5%。单独引入 MSGA 模块后,准确率、召回率和 $AP_{0.5}$ 分别提升 0.7、0.8、0.4 个百分点,达到 87.3%、83.9% 和 91.9%;单独引入 LSDH 后,分别提升 0.4、0.7、0.3 个百分点,达到 87.0%、83.8% 和 91.8%,表明二者均能有效提升检测性能。当 MSGA 与 LSDH 结合使用时,准确率、召回率和 $AP_{0.5}$ 分别达 87.7%、84.2% 和 92.5%,均为最优,说明 MSGA 与 LSDH 具有互补作用。

2.2.2 不同检测模型性能试验

将提出的 RSD 模型和 YOLO v7[29]、YOLO v8n、YOLO v9t、YOLO v10n、YOLO 11n 和 YOLO v12n 进行对比试验,试验采用准确率、召回率、平均精度和速度作为评价指标,试验结果如表 2 所示。

表 2 对比试验结果

Tab. 2 Comparative experimental results

模型	准确率/%	召回率/%	$AP_{0.5}$/%	速度/(f·s^{-1})
YOLO v12n	86.6	83.1	91.5	99.0
YOLO 11n	82.2	78.2	87.7	102.0
YOLO v10n	75.8	73.4	83.4	153.8
YOLO v9t	75.9	66.6	77.9	114.9
YOLO v8n	81.1	75.7	86.1	138.9
YOLO v7	84.2	86.1	87.9	82.6
RSD	87.7	84.2	92.5	104.2

从表 2 可以看出,提出的 RSD 模型在准确率和 $AP_{0.5}$ 指标上均超越了所有对比模型,其准确率与 $AP_{0.5}$ 分别达到 87.7% 与 92.5%,较 YOLO v12n 分别提升 1.1、1.0 个百分点,较 YOLO 11n 分别提升 5.5、4.8 个百分点。在召回率方面,RSD 达到 84.2%,较 YOLO v12n 提升 1.1 个百分点。虽然 YOLO v7 的召回率略高(86.1%),但 RSD 准确率与 $AP_{0.5}$ 均高出 3.5、4.6 个百分点,表明 RSD 在保持较高召回率的同时,具备更强的正样本判别能力。在推理速度方面,RSD 达到 104.2 f/s,与 YOLO v12n(99 f/s)性能相近。尽管 RSD 的推理速度逊于 YOLO v10n(153.8 f/s)等模型,但仍远高于农业机器人实时作业所需的 10 ~ 30 f/s,展现出良好的工程部署潜力。

2.2.3 场景适应性试验

构建了 4 个典型场景的测试集:浮萍、水面反光、低杂草密度及高杂草密度。其中浮萍场景有 484 幅,水面反光场景有 426 幅,低杂草密度和高杂草密度场景分别为 1 057、492 幅。在各场景下,对基线模型 YOLO v12n 及提出的 RSD 进行了对比测试,结果如表 3 所示。

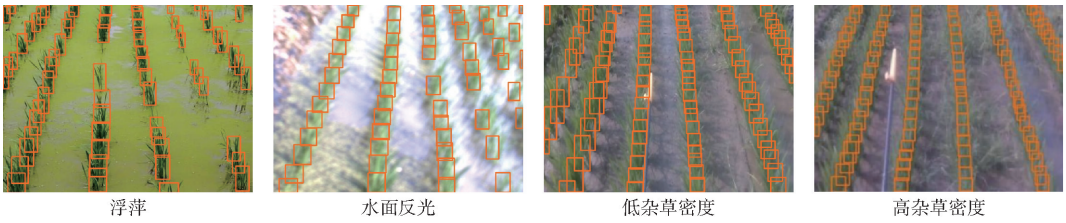
表 3 不同场景性能试验结果

Tab. 3 Performance experimental results across different scenarios %

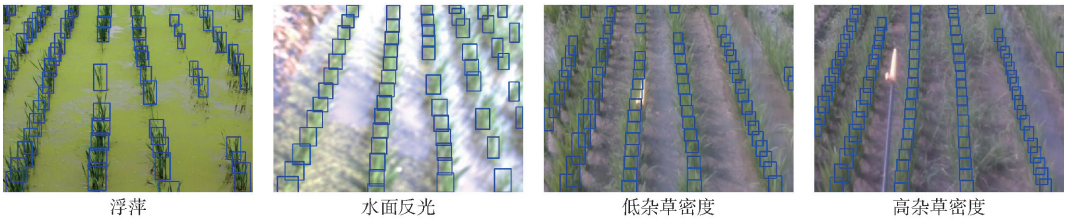
场景	模型	准确率	召回率	$AP_{0.5}$
浮萍	RSD	94.3	93.1	95.6
	YOLO v12n	91.0	90.2	95.0
水面反光	RSD	96.4	95.1	98.5
	YOLO v12n	95.6	93.8	98.3
低杂草密度	RSD	92.5	90.2	96.0
	YOLO v12n	90.5	88.5	94.7
高杂草密度	RSD	82.6	76.9	84.9
	YOLO v12n	80.2	76.4	84.3

在 4 种典型水田环境场景下,RSD 模型展现出卓越的检测性能,其中在水面反光场景下检效果最

佳，在高杂草密度场景下检测精度有所下降。总体而言，RSD 在所有场景下准确率、召回率和 $AP_{0.5}$ 3 项指标上均优于 YOLO v12n。具体而言：在浮萍场景中，RSD 准确率、召回率和 $AP_{0.5}$ 分别达到 94.3%、93.1% 和 95.6%，相对 YOLO v12n 分别提升 3.3、2.9、0.6 个百分点，提升幅度最为显著；在水面反光场景中，RSD 较 YOLO v12n 分别提升 0.8、1.3、0.2 个百分点；在低杂草密度场景中，RSD 准确率、召回率和 $AP_{0.5}$ 分别为 92.5%、90.2% 和 96.0%，相比 YOLO v12n 分别提升 2.0、1.7、1.3 个百分点；在高杂草密度场景下，RSD 仍保持优势，准确率、召回率和 $AP_{0.5}$ 分别为 82.6%、76.9% 和 84.9%，相比 YOLO v12n 分别提升 2.4、0.5、0.6 个百分点。不同场景识别结果如图 10 所示。



(a) RSD



(b) YOLO v12n

图 10　不同场景秧苗检测结果

Fig. 10　Seedling detection results in different scenarios

2.3　苗带线提取试验

为评估算法在不同缺苗场景下性能与鲁棒性，依据单行连续缺失株数，将缺苗程度划分为 3 个等级：轻度（缺失 1 株）、中度（连续缺失 2 ~ 3 株）与重度（连续缺失大于等于 4 株）。考虑到田间缺苗常呈非均匀分布，同一秧苗行内可能包含多种缺失模式，采用基于最严重缺苗段的等级判定原则：若一行中存在不同长度的连续缺苗段，则以其中最长的连续缺苗段所含缺失株数为依据，将该行整体划入对应等级。

基于上述标准，从总体测试集中随机抽取 100 幅图像，对其中每 1 条秧苗行进行独立的缺苗等级标注与统计，最终构建 1 个包含 62 行轻度缺苗、41 行中度缺苗与 21 行重度缺苗的专用测试样本集。基于该数据集，设计 3 个逐层递进试验：参数敏感性试验（探究参数对算法的影响规律）、缺苗鲁棒性试验（评估算法对不同缺苗等级的适应性）以及算法性能对比试验（与主流方法进行性能比较），以从参数优化、场景鲁棒性与方法优越性 3 个维度全面验证所提方法有效性。

2.3.1　参数敏感性试验

苗带线提取算法核心参数包括：网格滤波边长 B、欧氏聚类搜索距离阈值 d_T 与压缩比例 μ。为探究各参数对算法性能的影响规律并确定最优配置，设计了系统参数敏感性试验。其中网格滤波边长 B 主要控制特征点数量：B 越大，特征点越稀疏，算法运行速度越快；搜索距离阈值 d_T 影响聚类合并效果：阈值过大会导致不同苗行特征点误合并，过小则难以整合缺苗造成的离散点，且阈值增大会增加计算负担；压缩比例 μ 表示压缩程度，μ 过小难以克服严重缺苗导致的聚类分离问题，μ 过大则会造成行间特征点混淆。

在初始聚类阶段评估算法性能，暂不考虑压缩比例的影响，专门分析滤波边长和搜索距离阈值 2 个参数对聚类效果的影响。试验设计多组参数组合：B 为 4、5、6 像素，搜索距离阈值为 12、15、18 像素，共构成 9 种参数组合。针对每种组合，计算识别错误率与平均运行时间 2 个关键性能指标，测试结果如图 11 所示。

由图 11 可知，网格滤波边长 B 和搜索距离阈值 d_T 对识别错误率与运行时间具有显著影响。随着 B 增大，运行时间明显下降，但错误率变化呈非线性波动；同时，增大阈值可在轻微增加耗时的情况下有效降低错误率，有助于连接缺苗区域的特征点。在各种参数组合中，$B=5$ 像素、$d_T=18$ 像素组合取得了最低错误率（57.3%），但运行时间（0.035 9 s）相对较长。相比之下，$B=6$ 像素、$d_T=18$ 像素组合将运行时间显著缩短至 0.031 3 s（减少约 13%），而错误率仅小幅上升至 59.7%。该组合以较小的精度代价实现显著的计算效率提升，在效率与准确性之

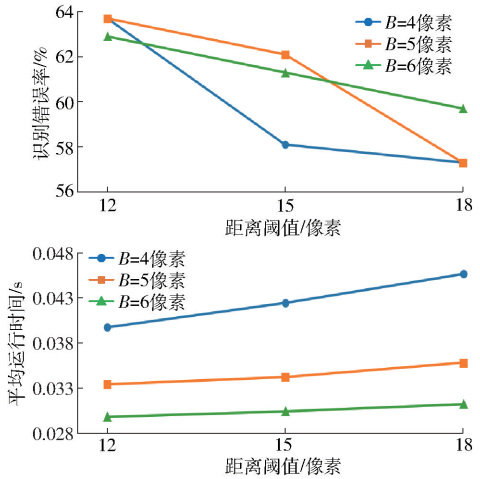


图 11 不同参数组合秧苗行初始聚类结果

Fig. 11 Initial clustering results of seedling rows with various parameter settings

间达到了更优的工程平衡。

为评估压缩比例对秧苗行识别结果的影响，设置 3 个压缩比例 μ 为 0.1、0.2、0.3，分别统计行断裂率与误合并率，结果如表 4 所示。其中，行断裂指同一秧苗行因中间严重缺苗，导致其被错误识别为多条不连续短行段的现象；误合并则指因相邻行间距过小或特征点交错，导致不同秧苗行特征点被错误聚类为同一行的现象。

表 4 不同压缩比例秧苗行识别结果

Tab. 4 Recognition results of seedling rows at different compression ratios

压缩比例	行断裂率/%	误合并率/%
0.1	0.81	0.81
0.2	8.06	0.81
0.3	23.39	0

由表 4 可知，压缩比例 μ 对行断裂率和误合并率存在显著的权衡关系。随着 μ 从 0.1 增大至 0.3，行断裂率从 0.81% 单调上升至 23.39%，而误合并率在 $\mu=0.1$ 和 $\mu=0.2$ 时保持在 0.81%，在 $\mu=0.3$ 时降至零。变化趋势表明，较高的 μ 有助于避免误合并，但会显著增加行断裂的风险；较低的 μ 虽能减少行断裂，却会导致更多误合并。

在秧苗行识别场景中，综合考虑作业连续性和导航精度需求，选择 μ 为 0.1 作为压缩阈值更为适宜。该设定下行断裂率仅为 0.81%，能够最大程度保持秧苗行结构连续性，这对于农业机械的稳定导航与路径跟踪至关重要；尽管存在 0.81% 误合并，但其在视觉上影响有限，并可通过形态学后处理予以修正。相比之下，μ 为 0.2 在误合并率不变的情况下，行断裂率上升至 8.06%，显著提高了行断裂风险，不利于实际作业的连续性保障。

2.3.2 缺苗鲁棒性试验

在构建的缺苗测试集上开展缺苗鲁棒性试验，试验选取 3 项核心评价指标：秧苗行识别正确率、行断裂率与误合并率。不同缺苗等级秧苗行识别结果如表 5 所示。

表 5 不同缺苗等级秧苗行识别结果

Tab. 5 Recognition results of seedling rows with different gap levels

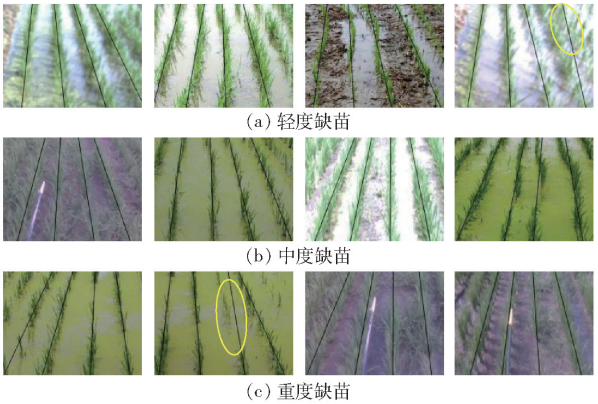
缺苗等级	测试行数	正确行数	识别正确率/%	行断裂率/%	误合并率/%
轻度	62	61	98.39	0	1.61
中度	41	41	100	0	0
重度	21	20	95.24	4.76	0
整体	124	122	98.38	0.81	0.81

由表 5 可知，提出的苗带线提取算法在不同缺苗等级下均展现出稳定的性能，且性能变化呈明确规律性。在轻度缺苗条件下，识别准确率达 98.39%，其中误合并率为 1.61%，误差主要源于轻度缺苗区域与田块边缘区域在图像空间中的邻接与视觉特征的相似性，导致算法在苗行边界判断上出现局部混淆。中度缺苗时识别正确率为 100%，说明该条件下缺苗特征已足够显著且苗行整体结构保持完整，算法处于最佳识别状态；而在重度缺苗条件下，算法识别正确率为 95.24%，行断裂率上升到 4.76%，反映出连续严重缺苗可能导致行结构断裂，从而引发过分割错误。总体而言，算法在整体测试集上平均识别正确率为 98.38%，且行断裂与误合并的总体发生率均低于 1%，验证了算法对于田间不同程度的缺苗情况具有良好的适应性与鲁棒性。

图 12 为轻度、中度及重度缺苗情况下秧苗行识别效果，识别错误区域均已用圆圈标注。由图 12a 可知，在轻度缺苗条件下，若缺苗区域位于第 1 或 4 行且与边缘秧苗行间距过近，易因空间邻近与视觉特征相似导致边缘行被误合并。由图 12c 可知，重度缺苗时，若同一行存在 2 处及以上缺苗区域，该行内零星秧苗形成的聚类结果可能被误判为独立秧苗行，引发过分割错误。

2.3.3 模型性能对比试验

将本文方法与 Wang 等[22]提出的基于初始聚类和子区域生长的方法以及传统的 Hough 变换算法进行性能比较试验。评估指标包括苗带识别平均角度偏差和平均运行时间，以综合评价算法精度与效率。表 6 为 3 种方法性能试验结果，本文算法在苗带识别任务中平均角度偏差 0.203°，显著优于文献[22]的 0.585°与 Hough 变换算法的 1.573°；在平



(a) 轻度缺苗

(b) 中度缺苗

(c) 重度缺苗

图 12　不同缺苗程度秧苗行识别效果

Fig. 12　Recognition effects of seedling rows with different levels of seedling gaps

均耗时方面，本文算法运行时间为 0.086 s，虽高于 Hough 变换算法的 0.036 s，仍完全满足田间实时作业时效要求。

图 13 为 3 种方法识别效果，其中黑色线条表示人工标注苗带基准线。由图 13 可知，在秧苗线拟合时，Hough 变换算法主要依据图像中满足共线约束的秧苗点确定直线参数，然而，该方法倾向于选取共线点数量最多的局部区域进行拟合，当秧苗行弯曲曲率较大时，拟合的直线往往仅覆盖部分苗带区域，无法有效跟踪整体走向，导致识别精度显著下降。文献[22]方法依赖于图像底部信息定位初始苗带线，在秧苗分布连续、底部结构清晰的场景中具有较好的识别性能。然而，当图像底部出现缺苗或严重遮挡时，该方法因无法准确获取初始苗带信息而导致识别失败，限制了其在复杂田间场景下的适用性。本文算法采用沿苗带主方向进行图像压缩与密度聚类策略，有效降低了对局部共线性和底部初始信息的依赖。试验结果验证了该方法在应对缺苗、曲率变化等复杂场景时具有更高的鲁棒性与环境适应性。

表 6　不同算法性能试验结果

Tab. 6　Performance test results of different algorithms

方法	平均角度偏差/(°)	单帧平均耗时/s
本文算法	0.203	0.086
文献[22]	0.585	0.042
Hough 变换算法	1.573	0.036

将本文算法部署至 Jetson TX2 边缘移动设备进行实际工况测试。结果表明，算法平均运行速度为 5.1 f/s。当除草装备以速度 0.5 m/s 作业时，车辆在单帧处理时间内前进距离为 9.8 cm。根据田间秧苗种植农艺参数(株距 10 ~ 13 cm)，该帧间位移小于秧苗最小株距(10 cm)，确保每株秧苗至少被 1 帧图像捕获，消除了整株秧苗落入检测盲区的风险。

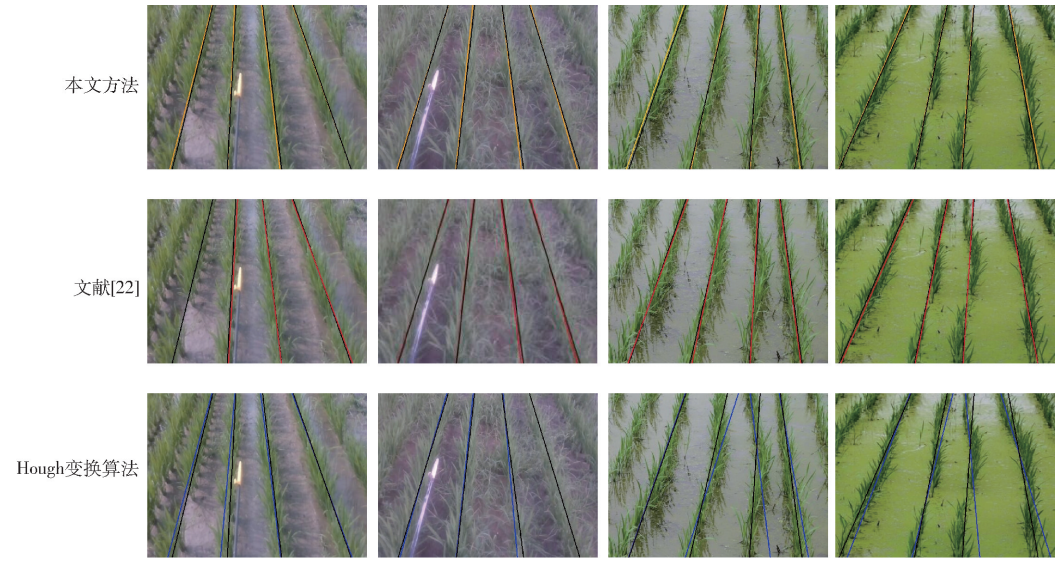


图 13　3 种方法识别效果

Fig. 13　Recognition performance of three methods

3　结论

(1)构建了水稻秧苗检测模型，采用多尺度门控注意力机制增强了小目标秧苗特征的判别能力，并设计了轻量化共享检测头以降低计算冗余。提出一种基于方向性压缩的迭代聚类方法，自适应地弥合了因缺苗造成的同行点纵向间隙，解决了传统聚类在缺苗场景下面临的苗行断裂问题。

(2)模型平均检测精度达 92.5%，在多项指标上均显著超越主流对比模型，尤其是在高杂草密度、水面反光等挑战性场景下，其性能优势与稳定性最为突出。在苗带线提取方面，在不同缺苗等级鲁棒

性测试中,行断裂率和误合并率均低于1%,所提取苗带线平均识别角度偏差仅为0.203°,单帧处理耗时0.086 s,行识别精度优于Hough变换算法及文献[22]方法。

参 考 文 献

[1] 唐伟, 徐红星, 董卉, 等. 我国水稻田除草剂同步用药现状与发展趋势[J]. 杂草学报, 2022, 40(2): 1-5.
Tang Wei, Xu Hongxin, Dong Hui, et al. Current status and development prospects of rice planting synchronized with herbicides application in China[J]. Weed Science, 2022, 40(2): 1-5. (in Chinese)

[2] 方会敏, 牛萌萌, 薛新宇, 等. 玉米田间机械-化学协同除草的杂草防除效果[J]. 农业工程学报, 2022, 38(6): 44-51.
Fang Huimin, Niu Mengmeng, Xue Xinyu, et al. Effects of mechanical-chemical synergistic weeding on weed control in maize field[J]. Transactions of the CSAE, 2022, 38(6): 44-51. (in Chinese)

[3] Wen X, Fu Y H, Xiang L L, et al. Impact of long-term use of multiple herbicides on soil microbial communities and nutrient cycling across different agricultural land-use types[J]. Agriculture, Ecosystems & Environment, 2025, 394: 109879.

[4] Ndukuryay O, Gong X Y, Yang W C, et al. Advances in understanding the toxicity of 4-hydroxyphenylpyruvate dioxygenase-inhibiting herbicides[J]. Journal of Agricultural and Food Chemistry, 2024, 72(32): 17762-17770.

[5] Benaragama D I, Willenborg C I, Shirtliffe S J, et al. Revisiting cropping systems research: an ecological framework towards long-term weed management[J]. Agricultural Systems, 2024, 213: 103811.

[6] Man J, Lei C, Xing W, et al. Development of herbicide resistance genes and their application in rice[J]. The Crop Journal, 2022, 10(1): 26-35.

[7] Rehman M U, Eesaar H, Abbas Z, et al. Advanced drone-based weed detection using feature-enriched deep learning approach[J]. Knowledge-Based Systems, 2024, 305: 112655.

[8] Zhu J, Wang J, Ditommaso A, et al. Weed research status, challenges, and opportunities in China[J]. Crop Protection, 2020, 134: 104449.

[9] Li Y, Guo Z, Shuang F, et al. Key technologies of machine vision for weeding robots: a review and benchmark[J]. Computers and Electronics in Agriculture, 2022, 196: 106880.

[10] 张馨悦, 王庆杰, 王超, 等. 旱田智能化机械除草技术与装备研究进展[J]. 农业机械学报, 2025, 56(4): 22-41,71.
Zhang Xinyue, Wang Qingjie, Wang Chao, et al. Research progress on intelligent mechanical weeding technology and equipment in dry farmland[J]. Transactions of the Chinese Society for Agricultural Machinery, 2025, 56(4): 22-41,71. (in Chinese)

[11] 朱惠斌, 李仕, 白丽珍, 等. 基于改进YOLO v8的玉米大豆间套复种作物行导航线提取方法[J]. 农业机械学报, 2025, 56(6): 205-217.
Zhu Huibin, Li Shi, Bai Lizhen, et al. Extraction method of navigation lines for maize soybean intercropping based on improved YOLO v8[J]. Transactions of the Chinese Society for Agricultural Machinery, 2025, 56(6): 205-217. (in Chinese)

[12] 孟笑天, 徐艳蕾, 王新东, 等. 基于改进K均值特征点聚类算法的作物行检测[J]. 农机化研究, 2020, 42(8): 26-30.
Meng Xiaotian, Xu Yanlei, Wang Xindong, et al. Crop line detection based on improved K-means feature point clustering algorithm[J]. Journal of Agricultural Mechanization Research, 2020, 42(8): 26-30. (in Chinese)

[13] 王爱臣, 张敏, 刘青山, 等. 基于区域生长均值漂移聚类的苗期作物行提取方法[J]. 农业工程学报, 2021, 37(19): 202-210.
Wang Aichen, Zhang Min, Liu Qingshan, et al. Seedling crop row extraction method based on regional growth and mean shift clustering[J]. Transactions of the CSAE, 2021, 37(19): 202-210. (in Chinese)

[14] 傅灯斌, 江茜, 齐龙, 等. 基于区域生长顺序聚类-RANSAC的水稻苗带中心线检测[J]. 农业工程学报, 2023, 39(7): 47-57.
Fu Dengbin, Jiang Qian, Qi Long, et al. Detection of the centerline of rice seedling belts based on region growth sequential clustering-RANSAC[J]. Transactions of the CSAE, 2023, 39(7): 47-57. (in Chinese)

[15] Xu F, Yao X C, Zhang K X, et al. Deep learning in cropland field identification: a review[J]. Computers and Electronics in Agriculture, 2024, 222: 109042.

[16] 朱立成, 韩振浩, 赵博, 等. 基于VINS-MONO和改进YOLO v4-Tiny的果园自主寻筐方法[J]. 农业机械学报, 2023, 54(8): 97-109.
Zhu Licheng, Han Zhenhao, Zhao Bo, et al. Autonomous basket searching method for orchards transporter based on VINS-MONO and improved YOLO v4-Tiny[J]. Transactions of the Chinese Society for Agricultural Machinery, 2023, 54(8): 97-109. (in Chinese)

[17] Liu T H, Zheng Y, Lai J S, et al. Extracting visual navigation line between pineapple field rows based on an enhanced YOLOv5[J]. Computers and Electronics in Agriculture, 2024, 217: 108574.

[18] Kong X, Guo Y, Liang Z, et al. A method for recognizing inter-row navigation lines of rice heading stage based on improved enet network[J]. Measurement, 2025, 241: 115677.

(下转第 426 页)

Signal Processing, 2019,133(1):106289.
[23] Pintelon R, Guillaume P, Schoukens J. Uncertainty calculation in (operational) modal analysis[J]. Mechanical Systems and Signal Processing, 2007,21(6):2359 - 2373.
[24] Pintelon R, Schoukens J. System identification: a frequency domain approach[M]. New York: IEEE Press,2012.
[25] Pi Y L, Mickleborough N C. Modal identification of vibrating structures using ARMA model[J]. Journal of Engineering Mechanics,1989,115(10):2232 - 2250.
[26] Pintelon R, Schoukens J. Frequency domain maximum likelihood estimation of linear dynamic errors-in-variables models[J]. Automatica, 2007,43(4):621 - 630.
[27] Dennis J E, Schnabel R B. Numerical methods for unconstrained optimization and nonlinear equations[M]. Philadelphia: SIAM,1996.
[28] Song Shenglin, Zhang Erliang. Identification of ARMAX models with noisy input: a parametric frequency domain solution[J]. IEEE Transactions on Signal Processing, 2025,73:292 - 304.
[29] Söderström T, Mahata K, Soverini U. Identification of dynamic errors-in-variables models: approaches based on two-dimensional ARMA modeling of the data[J]. Automatica, 2003,39(5):929 - 935.
[30] Pi Y L, Mickleborough N C. Modal identification of a vibrating structure in the time domain[J]. Computers & Structures, 1989,32(5):1105 - 1115.
[31] 张青松,郑中原,廖庆喜,等. 农业工程领域耦合仿真技术应用现状与展望[J]. 农业机械学报,2025,56(7):20 - 37.
Zhang Qingsong, Zheng Zhongyuan, Liao Qingxi, et al. Application status and prospect of coupled simulation in agricultural engineering research[J]. Transactions of the Chinese Society for Agricultural Machinery, 2025,56(7):20 - 37. (in Chinese)
[32] 胡国友,夏品奇,杨劲松. 渐变刚度钢板弹簧刚度特性计算的曲率-载荷混合法[J]. 南京航空航天大学学报,2008,4(1):46 - 50.
Hu Guoyou, Xia Pinqi, Yang Jinsong. Curvature-force hybrid method for calculating properties of leaf springs with variable stiffness[J]. Journal of Najing University of Aeronautics & Astronauties, 2008,4(1):46 - 50. (in Chinese)
[33] Kwon Y W, Bang H. The finite element method using MATLAB[M]. Boca Raton: CRC Press,2000.
[34] 邢铭宗,赵飞,姜歌东,等. 一种确定自回归滑动平均模型最小阶次的新方法[J]. 西安交通大学学报, 2011,45(12):99 - 103.
Xing Mingzong, Zhao Fei, Jiang Gedong, et al. Novel minimum order fixed method for autoregressive moving average models [J]. Journal of Xi'an Jiaotong University, 2011,45(12):99 - 103. (in Chinese)

(上接第 348 页)

[19] Diao Z, Guo P, Zhang B, et al. Navigation line extraction algorithm for corn spraying robot based on improved YOLOv8s network[J]. Computers and Electronics in Agriculture, 2023, 212: 108049.
[20] Huang S P, Wu S H, Sun C, et al. Deep localization model for intra-row crop detection in paddy field[J]. Computers and Electronics in Agriculture, 2020, 169: 105203.
[21] Wei C Y, Li H L, Shi J Y, et al. Row anchor selection classification method for early-stage crop row-following[J]. Computers and Electronics in Agriculture, 2022, 192: 106577.
[22] Wang S S, Yu S S, Zhang W Y, et al. The identification of straight-curved rice seedling rows for automatic row avoidance and weeding system[J]. Biosystems Engineering, 2023, 233: 47 - 62.
[23] Tang B Y, Zhou J P, Pan Y C, et al. Recognition of maize seedling under weed disturbance using improved YOLOv5 algorithm [J]. Measurement, 2025, 242: 115938.
[24] Zhang S, Liu Y, Xiong K, et al. A review of vision-based crop row detection method: focusing on field ground autonomous navigation operations[J]. Computers and Electronics in Agriculture, 2024, 222: 109086.
[25] 韩振浩, 李佳, 苑严伟, 等. 基于 UNet 网络的果园视觉导航路径识别方法[J]. 农业机械学报, 2021, 52(1): 30 - 39.
Han Zhenhao, Li Jia, Yuan Yanwei, et al. Path recognition of orchard visual navigation based on UNet[J]. Transactions of the Chinese Society for Agricultural Machinery, 2021, 52(1): 30 - 39. (in Chinese)
[26] Li D F, Li B L, Long S F, et al. Rice seedling row detection based on morphological anchor points of rice stems[J]. Biosystems Engineering, 2023, 226: 71 - 85.
[27] Liao J, Chen M, Zhang K, et al. Sc - net: a new strip convolutional network model for rice seedling and weed segmentation in paddy field[J]. Computers and Electronics in Agriculture, 2024, 220: 108862.
[28] Chen Z Y, Cai Y H, Liu Y X, et al. Towards end-to-end rice row detection in paddy fields exploiting two-pathway instance segmentation[J]. Computers and Electronics in Agriculture, 2025, 231: 109963.
[29] Wang C, Bochkovskiy A, Liao H. YOLOv7: trainable bag-of-freebies sets new state-of-the-art for real-time object detector[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. IEEE, 2023: 7464 - 7475.