doi:10.6041/j.issn.1000-1298.2026.09.027

基于先验嵌入与多尺度特征融合的耕后水稻单根秸秆语义分割方法

王　昱　奚小波　丁杰源　韩连杰　邹赟涵　沈　辉　张瑞宏
（扬州大学机械工程学院，扬州 225127）

摘要：耕后土壤表面秸秆覆盖率是评估秸秆还田质量的重要参数，现有方法难以有效识别水稻单根秸秆，为此提出一种融合先验嵌入与多尺度特征融合的语义分割方法，以实现水稻单根秸秆的准确识别。采用基于颜色距离的固定阈值分割法对原始图像进行预处理并生成先验图，在抑制背景干扰情况下为后续识别任务提供先验信息输入；设计了一种改进 U－Net 的 MRCF－DA－CDPE 模型，采用并行多尺度卷积捕获从细碎秸秆到大型秸秆聚集区等不同尺度的信息特征，利用通道与空间注意力筛选关键特征并抑制土壤亮斑等干扰，将预处理中的连续距离信息作为额外输入通道嵌入网络，为模型分割提供物理引导。图像预处理策略表明，该方法使各基础模型的平均精度提升了 2～4 个百分点，其中 U－Net 的识别精度最高。改进模型验证试验表明，MRCF－DA－CDPE 模型的平均交并比、平均精度和 Kappa 系数分别达到了 86.93%、94.89% 和 0.850 2，相比基础 U－Net 模型分别提升了 2.72、2.98 个百分点和 0.056 7。本方法实现了对单根秸秆的精细识别，可为秸秆还田质量检测、耕整效果评估等提供技术支撑。

关键词：水稻秸秆识别；语义分割；图像分割；深度学习；U－Net
中图分类号：TP391.41；S24　　**文献标识码**：A　　**文章编号**：1000-1298(2026)09-0289-10　　**OSID**：

Semantic Segmentation of Individual Straw Stalks after Plowing Based on Prior Embedding and Multi-scale Feature Fusion

WANG Yu　XI Xiaobo　DING Jieyuan　HAN Lianjie　ZOU Yunhan　SHEN Hui　ZHANG Ruihong
(*School of Mechanical Engineering*, *Yangzhou University*, *Yangzhou* 225127, *China*)

Abstract: Straw coverage on the soil surface after tillage is a key parameter for evaluating the quality of straw return to the field. Existing methods struggle to effectively identify individual rice straw stalks. To address this, a semantic segmentation method that integrated prior embedding with multi-scale feature fusion was proposed to achieve accurate recognition of individual rice straw stalks. A fixed-threshold segmentation method based on color distance was employed to preprocess the original image and generate a prior map, providing prior information input for subsequent recognition tasks while suppressing background interference. An enhanced U－Net model, MRCF－DA－CDPE, was designed. It employed parallel multi-scale convolutions to capture information features across scales—from fragmented straw fragments to large straw clusters—while utilizing channel and spatial attention to select key features and suppress disturbances like soil bright spots. Continuous distance information from preprocessing was embedded into the network as an additional input channel, providing physical guidance for segmentation. Image preprocessing strategies demonstrated that this approach enhanced the average accuracy of each base model by 2 ~ 4 percentage points, with U－Net achieving the highest recognition accuracy. Validation tests of the improved model demonstrated that the MRCF－DA－CDPE model achieved 86.93% average intersection-over-union ratio (IOU), 94.89% average precision, and 0.850 2 Kappa coefficient, representing improvements of 2.72, 2.98 percentage points and 0.056 7 respectively over the

收稿日期：2025－11－21　修回日期：2025－12－09
基金项目：国家重点研发计划项目(2022YFD1500404)、江苏省重点研发计划项目(BE2022338)、江苏省农业科技自主创新资金项目(CX(24)1026)、江苏省现代农机装备与技术示范推广项目(NJ2025－02)和扬州大学“高端人才支持计划”项目
作者简介：王昱(1997—)，男，博士生，主要从事智能农业装备研究，E-mail：MX120200512@stu.yzu.edu.cn
通信作者：奚小波(1989—)，男，副教授，博士生导师，主要从事智能农业装备研究，E-mail：xxbctg@126.com

baseline U-Net model. This method achieved precise recognition of individual straw stalks, providing technical support for straw return quality inspection and tillage effectiveness evaluation.

Key words: rice straw recognition; semantic segmentation; image segmentation; deep learning; U-Net

0 引言

近年来,由于不科学耕作与过度施肥,我国耕地耕作层土壤有机质含量显著下降[1-2]。秸秆还田是实现农业可持续发展的重要措施,可有效减少风蚀、增加土壤有机质、降低化肥和农药依赖[3-4]。我国水稻秸秆年产量约 2 亿吨,秸秆还田量占比达到 60%[5-6],因此确保还田后秸秆覆盖质量是实现资源高效利用的关键。由于水稻收获与下茬作物播种间隔短,还田质量检测具有时间紧、任务重的特点。传统检测方法主要包括视觉估计法[7]、线点样带法[8]、米尺测量法[9]、钉轮法[10]和摄影技术法[11]等,普遍存在人力消耗大、效率低、精度差且无法实时检测的缺点,不利于大面积推广应用。基于深度卷积神经网络的机器视觉已在作物成熟度检测、病虫害识别和农产品自动分拣等领域广泛应用[12-14]。目前虽有光谱传感器可精确识别土壤表面秸秆,但其价格昂贵、硬件脆弱,难以适配耕作机械使用。为解决秸秆还田质量实时检测难题,仍需研究更加精准、高效的检测方法,为还田作业质量评估和耕整效果调整提供技术支撑。

当前,秸秆覆盖率检测方法主要包括传统图像处理[15-17]、多光谱与高光谱遥感[18-19]以及深度学习[20-22]3 类。然而现有方法多针对成片或块状秸秆,对耕后细小、细长、断裂零散且极易与土壤亮斑、干燥土块混淆的单根秸秆仍存在严重漏检、断裂和边界模糊等问题。常规深度学习模型直接在原始 RGB 图像上端到端训练,易受复杂背景干扰;标准卷积感受野单一,难以同时捕获细碎片段与大面积聚集区的多尺度特征;对传统预处理生成的连续颜色距离先验信息也缺乏有效利用。以 U-Net 为例,其在细长结构捕获、残余干扰抑制以及先验信息挖掘方面均存在明显不足。

为此,本文提出一种融合先验嵌入与多尺度特征融合的单根秸秆语义分割方法。首先采用颜色距离固定阈值生成连续先验图抑制背景干扰,再以 U-Net为骨干构建 MRCF-DA-CDPE 模型,通过多分辨率上下文融合(MRCF)、动态注意力(DA)和颜色距离先验嵌入(CDPE)实现复杂田间环境下单根秸秆的精细鲁棒识别,为秸秆还田质量实时检测提供技术支撑。

1 秸秆图像数据集构建与预处理方法对比

1.1 数据采集与预处理

数据采集于 2024 年 11 月 12—16 日,地点为江苏省扬州市江都区七里镇。试验随机选取 10 块稻麦轮作田块,土壤类型为重粘土,耕层平均坚实度 172.37 kPa,水稻收获后秸秆铺设厚度约 9.7 cm,基本铺满田面。秸秆还田作业采用东方红 2204 型拖拉机配套双轴旋耕机,作业速度 5 km/h。图像采用小米 13 Pro 型智能手机采集,分辨率 4 096 像素 × 3 072 像素。拍摄时手机固定于 1.2 m 高三脚架,镜头轴线与地面夹角 30°,每隔 2 m 拍摄 1 幅,该高度与角度经多次验证,可良好模拟后续工业相机安装于拖拉机旋耕机挂件后的实际视角,同时保证近场地面分辨率约为 0.41 mm/像素。所有图像均使用无闪光手动模式拍摄,避免手机自动优化对颜色分布造成干扰。为充分覆盖田间光照变化,分别在 08:00、10:00、12:00、14:00 和 16:00 5 个时段拍摄,每时段图像数量相等,一半迎光拍摄、一半背光拍摄。人工剔除模糊、过曝、欠曝图像后,共获得有效图像 4 946 幅。

原始图像依次进行镜头畸变校正、透视变换、正射标准化和统一裁剪。畸变校正后边缘失真小于 2 像素;透视变换时,每幅图像手动选取田块四角附近特征点作为控制点,由同一人完成以降低主观误差,平均重投影误差小于 3.8 像素;变换后图像统一缩放至前进方向 1.26 m、横向 0.61 m 的固定物理尺寸,并裁剪中心 1 500 像素 × 3 072 像素区域,确保所有图像视野范围、物理尺度和像素分辨率完全一致,尺寸偏差小于 0.3%。最终得到 4 712 幅几何标准化图像用于模型训练与测试。秸秆真实掩膜采用 Labelme 逐像素手动标注,完全破碎秸秆不予标注。

1.2 图像分割方法比较

为筛选最适合作为深度学习先验输入的预处理方法,本研究选用来自不同耕深田块、覆盖多种光照时段的图像,对比了基于颜色距离的固定阈值分割、LAB-OTSU 自适应阈值分割、HSV 多通道阈值分割以及 K-means 聚类分割 4 种传统算法,如图 1 所示。结果表明,颜色距离固定阈值法计算复杂度最低且易于农机实时部署,在多数场景下分割边界最清晰、秸秆主体结构和细长连续性保持最完整,仅在强日照土壤亮斑处出现 5 ~ 10 像素局部误分割;

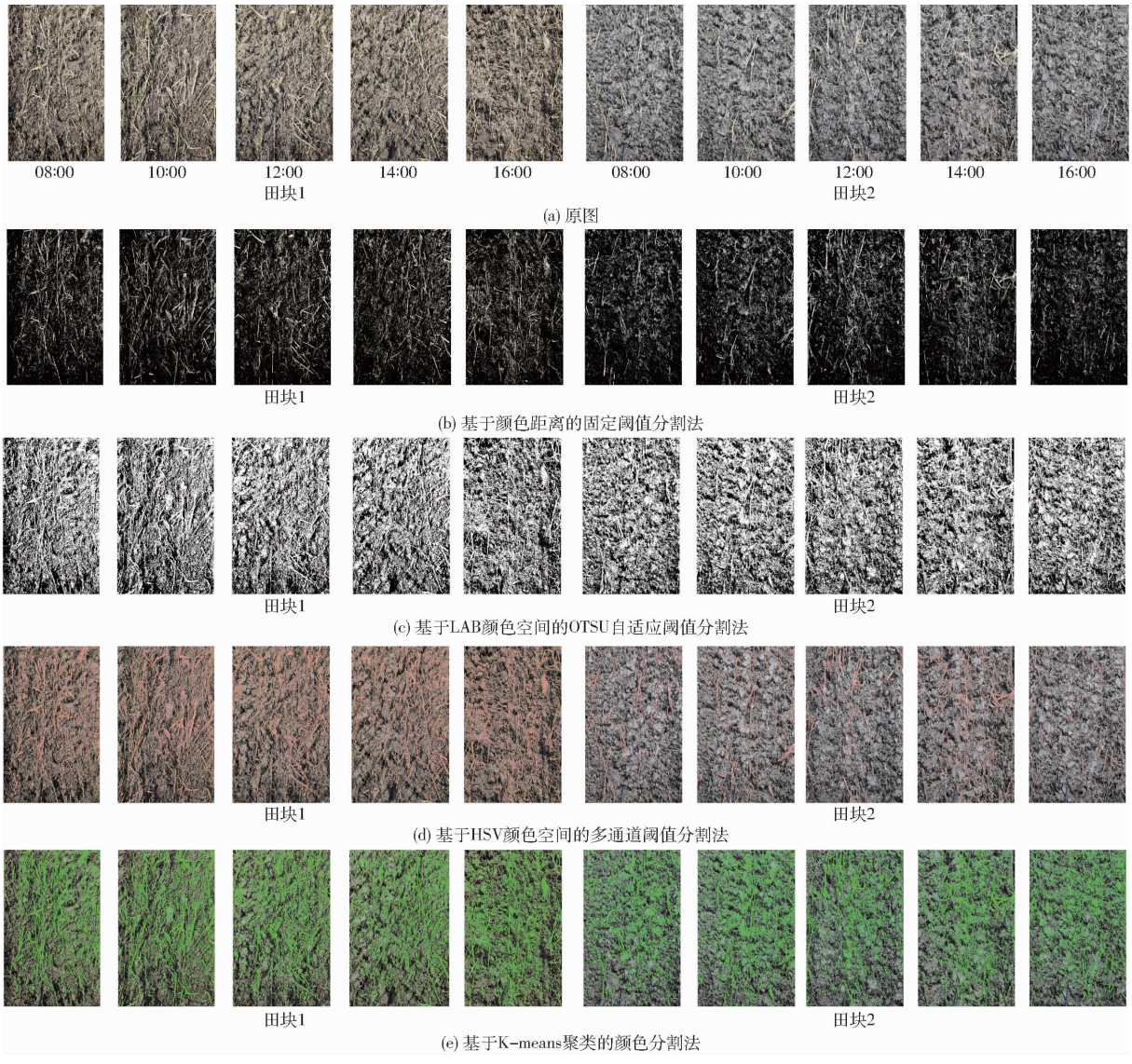

图 1　图像分割结果

Fig. 1　Image segmentation results

LAB－OTSU 法在亮斑干扰下易产生大面积前景背景粘连；HSV 多通道法在背光区域细小秸秆易丢失；K－means 聚类法存在结果不稳定、计算量大以及对噪声敏感等问题。

综合对比显示，颜色距离固定阈值法在分割清晰度、秸秆结构保持率、前景纯度和计算效率等方面均表现最佳。尽管局部强光区仍存在少量误分割，其轻量化特性和对秸秆整体结构的优良保留能力仍使其最适合生成连续颜色距离先验图并为后续深度学习模型提供可靠物理引导。因此，本研究最终选用该方法作为预处理策略。

2　MRCF－DA－CDPE 识别模型构建

基于训练策略与模型架构对比试验，U－Net 在多种语义分割模型中展现了最佳的识别性能。据此，本文采用 U－Net 作为基础架构，并对其进行改进。

针对基础 U－Net 在细长秸秆感受野不足、残余干扰抑制能力弱以及先验信息利用率低等问题，提出多分辨率上下文融合、动态注意力优化策略和颜色距离先验嵌入方法来提高单根秸秆的识别率，整体框架如图 2 所示。多分辨率上下文融合（MRCF）模块专门解决多尺度特征提取的技术难题，通过并行多尺度卷积架构捕获多层次的空间特征；动态注意力（DA）模块重点解决复杂环境下的干扰抑制和重要特征增强问题，通过三重注意力机制进行特征筛选；颜色距离先验嵌入（CDPE）模块负责利用预处理阶段获得的先验知识，将传统图像处理与深度学习进行融合。三者形成互补关系：MRCF 提供跨尺度结构特征，DA 强化语义表达并抑制伪特征，CDPE 则通过先验引导缩小搜索空间，从而实现对单根秸秆的稳定识别。

整个 MRCF－DA－CDPE 优化策略的数学目标可以表述为复合损失函数优化问题

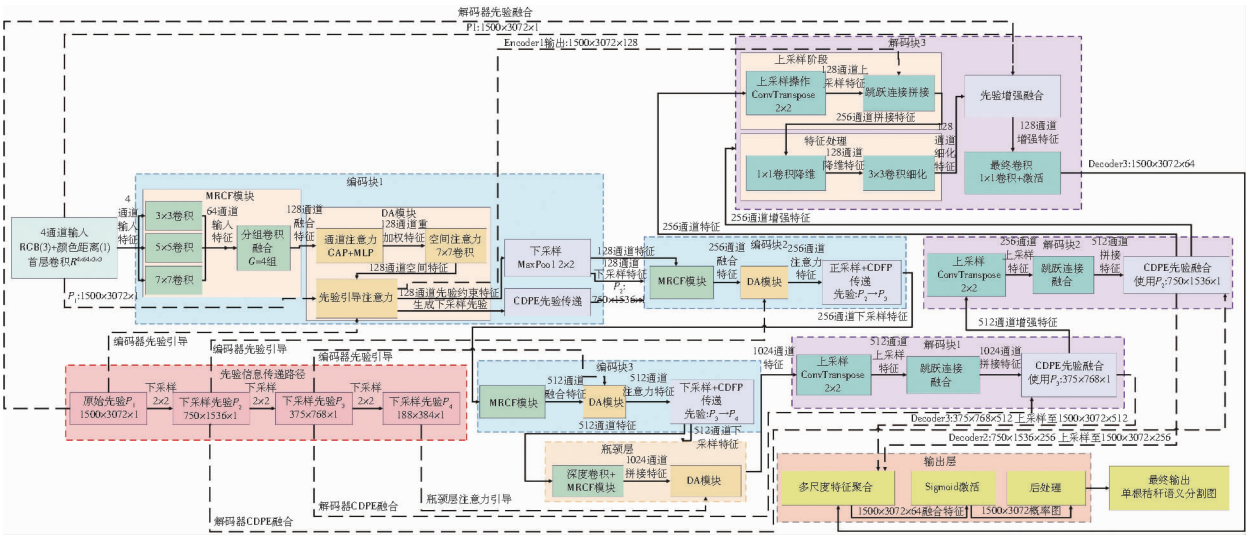

图 2 MRCF－DA－CDPE 优化策略架构框架

Fig. 2 Framework of network architecture for MRCF－DA－CDPE

$$L_{total} = L_{seg} + \lambda_1 L_{mrcf} + \lambda_2 L_{da} + \lambda_3 L_{cdpe} \quad (1)$$

式中 L_{seg}——标准的像素级分割损失

L_{mrcf}——多分辨率特征一致性损失

L_{total}——MRCF－DA－CDPF 模型的总损失

L_{da}——动态注意力正则化损失

L_{cdpe}——颜色距离先验嵌入损失

λ_1——多分辨率特征一致性损失的权重系数

λ_2——动态注意力正则化损失的权重系数

λ_3——颜色距离先验嵌入损失的权重系数

参数经过试验优化确定:$\lambda_1 = 0.1$,$\lambda_2 = 0.05$,$\lambda_3 = 0.15$。

2.1 多分辨率上下文融合(MRCF)

基于对随机抽取的 300 幅样本图像进行人工测量统计,细碎秸秆段长度主要集中在 2 ~ 8 像素范围,占总秸秆实例数量的约 30%。这类秸秆主要由收获过程中产生的小段秸秆和自然破碎的片段构成,其边缘在下采样过程中易被卷积核平滑或直接消失,导致小目标特征断裂;同时其纹理与土壤颗粒高度相似,容易被模型误判为背景。中等长度秸秆段长度为 10 ~ 25 像素,是最常见的秸秆形态,具有明显的条状或弯曲形状,虽具有可辨识的条状或弯曲结构,但常受到局部遮挡和形变影响,需要模型具备一定的形状保持能力。大面积秸秆聚集区通常面积超过 50 像素 × 50 像素,多出现在收获机转弯处或秸秆自然堆积的区域,内部重叠与堆积结构复杂,需依赖长距离上下文信息。上述差异导致单一尺度的标准卷积难以同时兼顾微观纹理、中等形状与宏观布局。

针对上述问题,MRCF 模块采用 3 × 3、5 × 5、7 × 7这 3 个并行卷积分支设计,核尺寸选择直接对应前述 3 类主要秸秆尺度,既保证覆盖完整,又保持较低计算成本。3 × 3 分支负责捕获微观纹理与精细边界,5 × 5 分支聚焦中等长度秸秆的完整形状与空间连接关系,7 × 7 分支则提取大范围分布模式与聚集区整体布局。3 分支输出在通道维度拼接后经 1 × 1卷积融合。为控制并行结构带来的计算量增长,每个分支均采用分组卷积方式,将输入通道均匀分组后独立卷积,最后合并输出,从而在几乎不增加计算开销的前提下实现高效多尺度特征提取。

综上,不同尺度秸秆在像素空间中具有显著的尺寸跨度与结构复杂性,需要模型同时具备对局部纹理、细长结构和大范围聚集模式的多尺度建模能力。MRCF 模块通过并行 3 × 3、5 × 5、7 × 7 卷积核构建分层感受野,并在特征融合阶段形成跨尺度上下文一致性表达,从而有效弥补基础 U－Net 在细长目标捕获、聚集区域整体识别和尺度连续性建模方面的不足。

2.2 动态注意力(DA)

预处理后的图像中非秸秆区域虽占比相对较低,但其干扰模式具有高度结构化的特征,主要表现为光照诱导的高反射干扰、颜色与纹理相似性干扰以及结构类异物干扰。在强日照条件下,土壤表面高亮反射区域在 RGB 空间中与干燥秸秆呈现极为相似的亮度特征,导致浅层卷积难以在响应强度上有效区分两者,进而引发显著的伪秸秆激活。干燥土块在色相与亮度上与秸秆高度重叠,尤其在土壤水分梯度较大或地表起伏明显区域更为突出,这类区域往往缺乏清晰边缘与纹理差异,颜色距离预处理难以实现可靠分离。石块与植物残留等异物则呈片状或块状分布,其局部形态结构与大面积秸秆聚集区高度相似,在缺乏足够上下文约束时易被卷积核错误识别为秸秆。

传统卷积仅依赖固定权重，对光照产生的高反射区域与真实秸秆纹理表现出相似响应，导致模型在复杂背景下缺乏稳定的区分能力。针对这一问题本文设计了一种动态注意力（DA）模块。该模块集成了通道注意力、空间注意力以及先验引导注意力 3 种机制，并以级联方式构建，实现从通道到空间再到先验的逐级特征筛选，从而赋予模型在复杂土壤背景下的精细区域响应能力。通道注意力首先对特征图执行全局平均池化以压缩空间维度获得每个通道的全局描述子，随后通过 2 层小型全连接网络学习通道间非线性依赖关系并生成归一化权重，最后将权重逐通道乘回原始特征图，从而在保持空间结构完整的前提下完成对关键通道的有效重校准。空间注意力并行执行通道平均池化与通道最大池化分别捕获平滑背景与显著边缘信息，将两者拼接后送入 7×7 卷积并经 Sigmoid 激活生成空间权重图，该权重图能覆盖更大上下文窗口，对捕获秸秆长条状结构及局部聚集模式尤为关键，最终通过逐元素乘法施加到前一级输出特征。先验引导注意力则将预处理生成的颜色距离场作为软先验，首先将其调整至当前特征图分辨率，再与前 2 级输出在通道维度拼接，经 1×1 卷积与 Sigmoid 激活生成引导权重，最后通过逐元素乘积得到最终增强特征，以显式方式将传统图像处理知识融入深度网络。

综上，通道注意力强化了对秸秆判别关键的纹理与边缘通道，空间注意力提升了对长条状结构与局部聚集模式的感知能力，先验引导注意力则将颜色距离先验作为软约束进一步抑制歧义区域。三者协同作用显著提升了模型在复杂田间背景下的特征辨识与区域响应能力。

2.3　颜色距离先验嵌入（CDPE）

颜色距离图可视为一种连续的软先验概率场，其数值与像素属于秸秆的后验置信度呈严格负相关。距离越小，该像素颜色分布越接近参考秸秆分布，前景概率越高；中等距离区域通常对应秸秆边缘或过渡带，其类别判定需依赖局部上下文；接近阈值的区域则处于前景与背景分布高度重叠的歧义区，最易产生误判。因此，颜色距离图不仅提供了前景强度的连续量化表达，同时隐含空间连续性与不确定性分布，对深度网络特征选择具有重要的物理引导意义。CDPE 模块通过保留完整颜色距离数值并将其作为第 4 输入通道嵌入网络，实现传统图像处理结果向深度学习的直接信息传递。颜色距离由像素与标准参考颜色在 RGB 空间的欧氏距离计算获得，并经归一化处理至区间（0，1），以确保与 RGB 通道在量纲与统计特性上的一致性。

在 U－Net 中引入该通道时，首层卷积核由 3 通道扩展为 4 通道，其中新增权重专用于学习颜色距离特征，从而在网络最底层即可建立外观信息与先验概率的联合表征。这种显式多模态融合策略使模型对高反射、颜色相似等易混区域的判别稳定性显著提升，同时保证先验信息在后续深层传播中得到有效保留。

为进一步克服传统方法仅在输入层单点利用先验的局限性，本文设计了多尺度先验传递路径。该路径基于如下观察：浅层特征主要关注纹理与边缘等低级视觉信息，与颜色距离局部相似性高度相关；深层特征则转向高级语义表达，仅需弱化全局先验作为辅助约束。为此，在编码器各层保留对应分辨率的颜色距离图，并通过专用路径逐级注入解码器对应层。注入权重采用指数衰减模式，浅层赋予较高权重以强化局部引导，深层权重逐步减小以避免抑制网络自主学习。最终增强特征通过加权求和方式获得，即

$$F_l^{\text{enhanced}} = (1-\alpha_l)F_l + \alpha_l(F_l \odot P_l) \tag{2}$$

式中　F_l^{enhanced}——增强后的特征输出

α_l——第 l 层的先验融合权重系数

F_l——当前层特征图

P_l——对应尺度的先验图

该渐进式融合策略确保先验信息在网络不同抽象层次均发挥尺度匹配的约束作用，既保留原始语义特征，又缓解 U－Net 中 skip-connection 常见的语义鸿沟问题。

综上，CDPE 模块通过颜色距离场构建连续的概率先验，并在编码器各尺度逐级注入先验信息，形成跨尺度一致的软约束，从而实现对秸秆区域的显式引导与对高反射干扰区域的抑制。该机制有效弥合了传统图像处理与深度语义特征之间的表征差异，是提升单根秸秆识别精度的重要组成部分。

3　试验设计与模型训练

试验在配备英特尔酷睿 6 代处理器、英伟达 GeForce RTX 4070 显卡、64 GB 内存和 12 GB 显存的工作站上进行。训练环境在 Windows 系统上使用 Anaconda 搭建，可视化编辑器采用 PyCharm，训练环境基于 Windows 10，采用 Python 3.8 与 PyTorch 1.12。

所有模型采用完全相同的训练设置以保证公平对比：训练 10 000 个轮次，使用交叉熵损失，学习率采用余弦退火策略，批大小为 16。每训练 10 个 epoch 记录一次损失，每 500 个轮次保存一次当前最优模型。训练结束后，以曲线图形式展示各模型的

损失收敛情况、验证集指标以及总训练时长。

3.1 评价指标

模型性能采用平均交并比(mIoU)[23]、平均精度(mAP@0.5,COCO 标准)[24]和 Kappa 系数[25]3 个指标综合评估。其中,mIoU 衡量预测分割区域与真实标注的重叠程度;mAP 反映不同置信度阈值下的分类精度,适合类别不均衡场景;Kappa 系数评估预测结果与随机分类的一致性。3 个指标从边界精度、分类准确性以及整体一致性 3 个维度量化模型在耕后单根秸秆分割任务中的表现。

3.2 训练策略与模型架构对比

为评估预处理策略的有效性并选择最优的基础模型,设计了 2 层递进式的对比试验。通过训练策略对比验证颜色距离预处理相对于直接端到端训练的优势。

表 1 展示了端到端训练与预处理后训练每个模型在 10 000 次迭代后的最终的 mAP、mIoU 和损失值,所有模型均表现为使用预处理图像训练的模型精准性更高,其中 U-Net 和 SegMenter 的最终 mAP 和 mIoU 最高。此外,使用基于颜色距离的固定阈值分割法处理后的图像训练模型,迭代过程中趋向的损失值更低,其中无论是否使用预处理图像,U-Net的最终损失值均低于其他模型。

表 1 不同训练策略对比

Tab.1 Comparison of different training strategies

模型	端到端训练			预处理后训练		
	损失值	mIoU/%	mAP/%	损失值	mIoU/%	mAP/%
U-Net	0.019 1	86.23	87.21	0.008 1	91.05	92.29
U^2-Net	0.152 5	84.17	76.01	0.102 5	84.34	80.43
MobileSeg	0.193 2	72.45	68.36	0.042 9	76.72	72.33
SegNet	0.081 1	78.76	74.40	0.053 2	79.13	74.85
SegNeXt	0.122 1	83.91	79.85	0.037 2	85.67	82.44
Topformer	0.165 9	77.34	70.74	0.114 2	77.89	81.35
DeepLab	0.152 2	84.78	87.15	0.096 7	88.95	87.35
SegMenter	0.186 3	86.47	85.55	0.118 3	90.82	87.11

为全面评估预处理策略的有效性,本文选取了 8 种代表性深度学习模型进行性能对比,该对比测试旨在不同架构间建立性能参照,为后续的模型改进提供依据,结果如表 2 所示。

图 3 展示了每个模型的损失值,表 2 则列出了在 2 000、6 000、10 000 次迭代时的训练损失值。由图 3 可知,在秸秆数据集优化中,SegMenter 损失最高、波动最大、收敛最慢;其他模型损失随迭代增加而相对稳定,10 000 次迭代后 U-Net 达最低值(0.008 1)。图 4 和图 5 展示了每个模型的精度趋势。经过 10 000 次迭代后,所有模型都达到了稳定状态,其中

表 2 每个对比模型在不同迭代轮次下的损失值

Tab.2 Loss value statistics of each comparison model across different iterations

模型	迭代轮次		
	2 000	6 000	10 000
U-Net	0.045 4	0.011 6	0.008 1
U^2-Net	0.208 7	0.127 9	0.102 5
SegNet	0.764 9	0.042 6	0.053 2
SegNeXt	0.062 4	0.044 8	0.037 2
SegMenter	0.131 5	0.129 1	0.118 3
DeepLab	0.176 1	0.118 4	0.096 7
TopFormer	0.136 1	0.079 7	0.144 2
MobileSeg	0.060 9	0.049 2	0.042 9

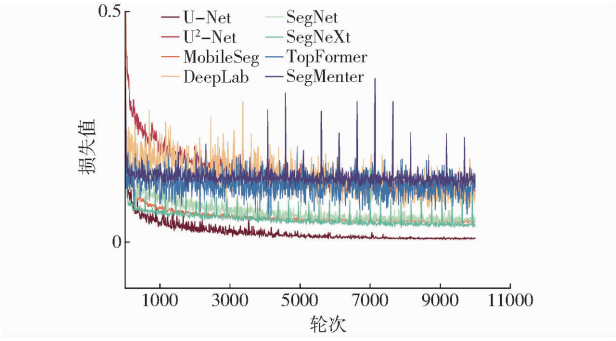

图 3 每个对比模型在验证集上的损失值

Fig.3 Loss value of each comparison model on validation set

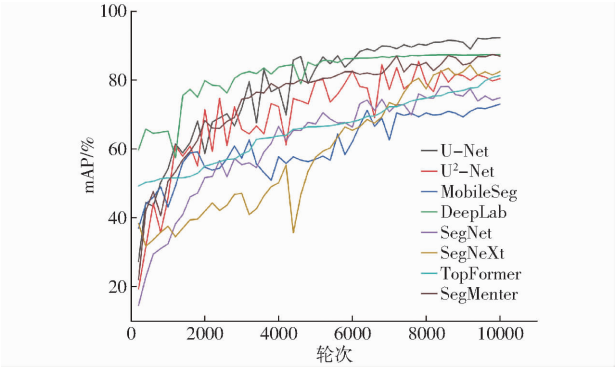

图 4 各对比模型在验证集上的 mAP 曲线

Fig.4 mAP curves of each comparison model on validation set

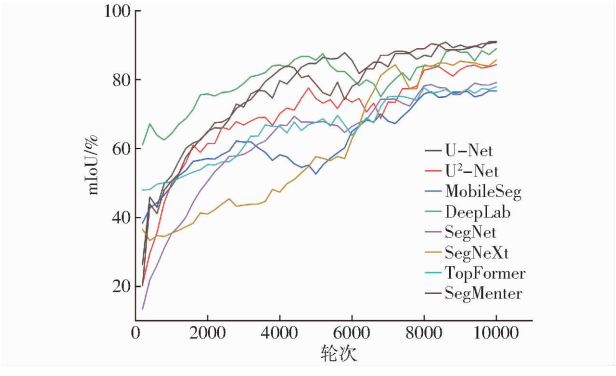

图 5 各对比模型在验证集上的 mIoU 曲线

Fig.5 mIoU curves of each comparison model on validation set

SegNeXt 波动最大，而 U－Net 与 DeepLab 表现出较低方差与更快收敛，迅速逼近全局最优。在精度方面，所有模型在训练集上的 mAP 都超过了 72%，其中 U－Net 的准确率最高达到 93%，所有模型在训练集上的 mIoU 都超过了 75%，其中 U－Net 的准确率最高达到 91.05%。

综上所述，基于颜色距离的预处理策略通过显式剥离背景干扰，在 8 种主流分割模型上均表现出良好的普适性，使平均识别精度提升 2～4 个百分点，验证了该先验构建方法的有效性。

3.3　改进模型与基础模型对比

为量化改进模型所带来的性能增益，本文在验证集上对 MRCF－DA－CDPE 与基础 U－Net 模型进行了对比试验。图 6 展示了 2 种模型的 mIoU 收敛曲线。在 10 000 次迭代结束后，二者均趋于稳定，但在前 6 000 次迭代范围内，基础 U－Net 波动明显大于 MRCF－DA－CDPE。随着迭代推进，两模型 mIoU 均持续提升，在迭代 6 000～7 000 轮次区间内波动渐趋平缓，迭代至 7 500 轮次后完全稳定。最终，MRCF－DA－CDPE 的 mIoU 达到 87.29%，较基础 U－Net 提升 3.2 个百分点，表明其在秸秆覆盖条件下的结构完整性与边界连续性优于基准模型。

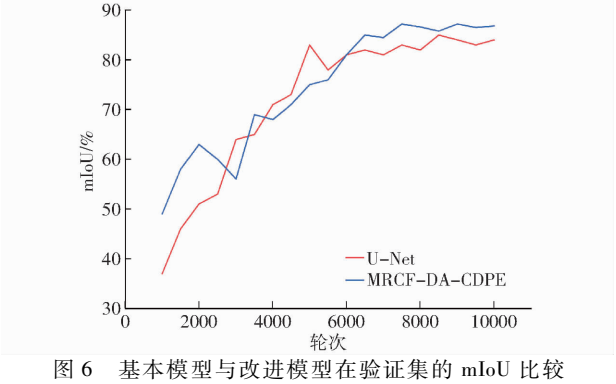

图 6　基本模型与改进模型在验证集的 mIoU 比较

Fig. 6　Comparison of mIoU between baseline model and improved model on validation set

图 7 给出了 2 种模型的 mAP 收敛曲线。mIoU 的提升表明模型在秸秆结构完整性、边界连续性方面有所改善；mAP 的提升表明模型在区分秸秆与干扰区域方面具备更高的判别能力。随着迭代增加，两模型 mAP 均呈上升趋势，最显著提升出现在迭代 0～4 000 轮次阶段。在迭代 4 000～6 500 轮次区间内，二者进入优化期并出现小幅波动，随后迅速稳定。MRCF－DA－CDPE 最终 mAP 达 95.41%，较基础 U－Net 提升 3.84 个百分点。mIoU 与 mAP 的双重提升反映了改进模型在秸秆结构保持、边界精确性以及干扰区域判别方面改善显著。

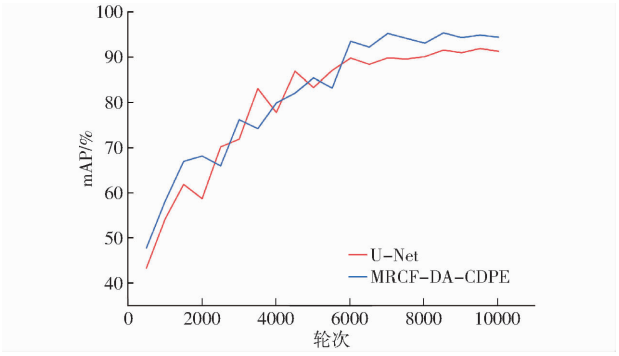

图 7　基本模型与改进模型在验证集上的 mAP 比较

Fig. 7　Comparison of mAP between baseline model and improved model on validation set

3.4　MRCF－DA－CDPE 模型消融试验

为评估各改进模块的有效性，本文设计通过消融试验分析多分辨率上下文融合、动态注意力机制及颜色距离先验嵌入策略的独立与协同贡献。以基础 U－Net 为基准，逐步引入单个模块、双模块组合直至全模块配置。MRCF 模块通过扩展感受野与强化多尺度结构捕获，单独引入后 mIoU 提升 1.05 个百分点、mAP 提升 1.13 个百分点、Kappa 提升 0.025 1，显著改善细碎秸秆纹理连续性与大聚集区边界完整性。DA 模块经通道、空间与先验三重注意力增强特征筛选与干扰抑制，单独引入后 mIoU 下降 2.11 个百分点，但 mAP 提升 1.71 个百分点、Kappa 提升 0.044 1，有效降低土壤亮斑伪激活，尽管对结构保持略有影响。CDPE 模块嵌入颜色距离先验作为软约束，单独引入后 mIoU 下降 3.22 个百分点、mAP 提升 0.33 个百分点、Kappa 提升 0.049 6，提升幅度有限，主要因依赖网络自主学习，但为后续模块提供物理引导基础。双模块组合进一步凸显互补效应。MRCF＋DA 使 mIoU 提升 1.29 个百分点、mAP 提升 1.34 个百分点、Kappa 提升 0.039 6，动态注意力强化了多尺度输出的语义一致性；MRCF＋CDPE 使 mIoU 提升 0.99 个百分点、mAP 提升 1.09 个百分点、Kappa 提升 0.036 6，先验嵌入辅助了多尺度融合稳定性；DA＋CDPE 使 mIoU 提升 0.64 个百分点、mAP 提升 0.89 个百分点、Kappa 提升 0.033 0，注意力机制充分利用先验权重实现更精准干扰抑制。

消融结果汇总于表 3。完整 MRCF－DA－CDPE 模型在测试集上 mIoU 达 86.93%、mAP 达 94.89%、Kappa 达 0.850 2，相较任意子集，3 项指标同步改善，表明模型在结构完整性、语义判别与预测稳定性方面均获显著优化。

图 8 展示了不同背景下消融模型的测试结果。其中，示例图 1 代表了秸秆离散分布、背景相对简单

表 3 在测试集上对消融模型的评估指标

Tab. 3 Test results of ablation model

模型	mIoU/%	mAP/%	Kappa 系数
U－Net	84.21	91.91	0.793 5
MRCF	85.26	93.04	0.818 6
DA	82.10	93.62	0.837 6
CDPE	80.99	92.24	0.843 1
MRCF－DA	85.50	93.25	0.833 1
MRCF－CDPE	85.20	93.00	0.830 1
CDPE－DA	84.85	92.80	0.826 5
MRCF－DA－CDPE	86.93	94.89	0.850 2

的典型场景，示例图 2 则包含了秸秆聚集且存在土壤亮斑等干扰的复杂场景。基础 U－Net 在细碎秸秆和局部聚集区域表现出结构断裂与背景混淆，主要源于其有限感受野导致的尺度不一致性。MRCF 模块基于多尺度感受野构建原则，通过并行 3×3、5×5 和 7×7 卷积捕获不同粒度空间结构，显著改善长秸秆整体连贯性与细碎秸秆局部纹理表达，但由于依赖卷积局部建模，对于多层叠加或结构模糊区域，跨区域语义依赖仍不足，部分聚集区未完全恢复。DA 模块利用通道、空间和先验三重注意力动

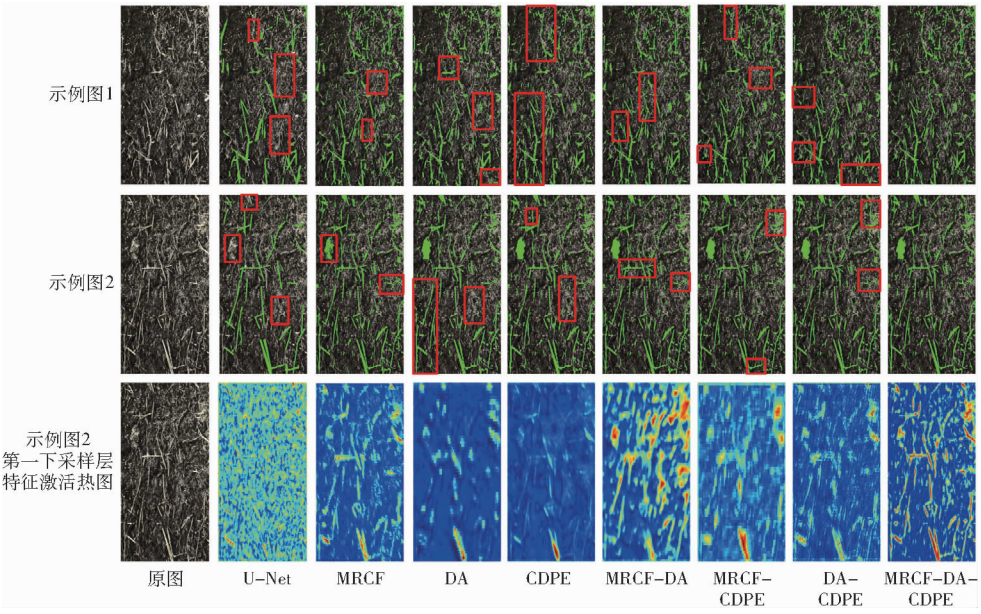

图 8 不同背景下消融模型的测试结果

Fig. 8 Results of ablation model testing under varying background

态重加权特征，强化干扰抑制如亮斑和干燥土块，但特征选择依赖网络学习，在纹理相似或边界模糊区域易受干扰，表现为局部漏检。CDPE 模块引入颜色距离场作为软先验，提供歧义区域概率约束，提升边界判别能力，但其主要基于颜色分布，对结构信息建模有限，在多秸秆重叠或颜色剧变背景下，形状细节恢复仍受限；双模块组合进一步凸显互补：MRCF＋DA 通过结构建模与动态筛选优化尺度一致性；MRCF＋CDPE 在多尺度先验融合中辅助边界稳定性；DA＋CDPE 利用注意力权重强化干扰抑制。MRCF－DA－CDPE 模型整合所有组件，在结构建模、干扰抑制与先验约束上形成互补，实现最佳效果。

为剖析模型内部的注意力分配机制，图 8 中可视化了示例图 2 在第一下采样层的特征激活热图，红色编码高激活关注度，直观揭示特征学习演化。基础 U－Net 热图红色区域稀少，黄色与黄绿色中激活区为主，蓝色低激活少，红色虽指向秸秆但混杂非目标。MRCF 后红色仍少，但黄色与黄绿色高亮集中于秸秆轮廓，初步突出结构表达。DA 后高亮减少且模糊，秸秆轮廓响应减弱，体现干扰抑制伴随细节丢失。CDPE 后高亮进一步稀少，仅显现秸秆微微轮廓，提供初步边界引导。MRCF＋DA 中红色增多，黄色与黄绿色围绕分布，高亮覆盖秸秆但仍混杂非目标，轮廓略模糊。MRCF＋CDPE 红色减少，黄色与黄绿色高亮增多但模糊，协同增强多尺度稳定性。DA＋CDPE 红色与高亮更少，秸秆轮廓相对清晰，模糊度降低。全模型 MRCF－DA－CDPE 红色区域丰富且集中于秸秆，轮廓清晰明显，黄色与黄绿色高亮限于细碎秸秆处，整体关注度最优。这些可视化结果证实模块逐步优化机制：MRCF 提升轮廓表达，DA 强化抑制，CDPE 辅助平滑，双模块协同放大清晰度，为单根秸秆精细识别提供支撑。

综上，MRCF－DA－CDPE 通过联合多尺度结构建模、动态特征重加权与先验约束，提升复杂田间背景下结构预测稳定性和降低误分类率，为水稻单根秸秆精准检测提供可靠方案。

3.5 边缘设备部署分析

在农机装备实际作业场景中，视觉检测系统需

部署于低功耗、小体积、抗恶劣环境的边缘计算设备，如 NVIDIA Jetson 系列。因此，模型在边缘硬件上的可部署性与实时性是评估其工程应用价值的关键指标。MRCF－DA－CDPE 模型在设计中强调结构轻量化与高效特征处理路径，通过并行多尺度卷积结合分组策略、级联注意力机制及特征图尺寸控制，在维持高表达能力的同时显著降低计算量与显存占用量，参数规模适中、计算密度较低，完全适配嵌入式 GPU 加速推理。

NVIDIA Jetson 系列作为嵌入式平台的主流选择，具有优异能效比、完善的深度学习推理框架支持及强环境适应性。其内置 CUDA 加速单元与 TensorRT 引擎可自动融合卷积、高维特征处理及跳跃连接等操作，大幅提升推理效率。MRCF－DA－CDPE 的规则化卷积结构、适中模块数量与清晰计算路径高度契合 Jetson 特性，推理显存占用量稳定，不依赖复杂自注意力或过深堆叠，从而在低功耗平台上展现良好可部署性与运行稳定性。

4　结论

（1）提出一种基于颜色距离的固定阈值预处理方法，通过构建连续的颜色距离先验图显式剥离背景干扰，降低输入分布复杂度并提升前景可分性，8 种主流语义分割模型识别精度均提升了 2～4 个百分点，验证了该先验构建策略的普适性。

（2）构建了基于 U－Net 的 MRCF－DA－CDPE 模型，MRCF 模块提供多尺度结构建模能力，DA 通过动态特征重加权增强有效区域响应，CDPE 模块将颜色距离先验作为软约束引入深度网络，实现传统图像处理与深度学习的融合表达。

（3）MRCF－DA－CDPE 模型 mIoU、mAP 和 Kappa 系数分别达到 86.93%、94.89% 和 0.850 2，相比基础 U－Net 提升 2.72、2.98 个百分点和 0.056 7，改进模型在细碎、离散以及弱边缘秸秆识别方面表现尤为突出，反映出其对复杂田间场景具有更高的鲁棒性。

（4）全消融试验表明，MRCF 提升多尺度结构表达能力，DA 增强干扰抑制与特征选择，CDPE 提供物理先验约束，三模块在结构建模、注意力强化和先验融合之间形成互补，形成从局部纹理到高层语义的一致性增强路径，实现显著的协同增益。

参　考　文　献

[1] ZHANG Y, SUI B, SHEN H, et al. Estimating temporal changes in soil pH in the black soil region of Northeast China using remote sensing[J]. Computers and Electronics in Agriculture, 2018, 154: 204－212.

[2] 李伟秋. 垦造水田规划设计中土壤有机质的提升及 pH 调整作用分析[D]. 广州：华南农业大学, 2019.
LI Weiqiu. Analysis of soil organic matter improvement and pH adjustment in reclamation paddy field planning and design[D]. Guangzhou: South China Agricultural University, 2019. (in Chinese)

[3] 王庆杰, 曹鑫鹏, 王超, 等. 东北黑土地玉米免少耕播种技术与机具研究进展[J]. 农业机械学报, 2021, 52(10): 1－15.
WANG Qingjie, CAO Xinpeng, WANG Chao, et al. Research progress of no/minimum tillage corn seeding technology and machine in northeast black land of China[J]. Transactions of the Chinese Society for Agricultural Machinery, 2021, 52(10): 1－15. (in Chinese)

[4] 汪可欣, 付强, 张中昊, 等. 秸秆覆盖与表土耕作对东北黑土根区土壤环境的影响[J]. 农业机械学报, 2016, 47(3): 131－137.
WANG Kexin, FU Qiang, ZHANG Zhonghao, et al. Effects of straw mulching mode and tillage methods on soil environment of root zone in northeast black soil[J]. Transactions of the Chinese Society for Agricultural Machinery, 2016, 47(3): 131－137. (in Chinese)

[5] 成慧婷, 赵楠, 曾尚鹏, 等. 基于 BMP 法微生物驱动水稻秸秆选择性降解木质素能力研究[J]. 太阳能学报, 2024, 45(12): 485－493.
CHENG Huiting, ZHAO Nan, ZENG Shangpeng, et al. Study on microbial-driven selective lignin degradation capacity of rice straw based on BMP method[J]. Acta Energiae Solaris Sinica, 2024, 45(12): 485－493. (in Chinese)

[6] 杨励华. 秸秆及有机粪肥对覆膜春玉米生长发育和氮素吸收与利用的影响[D]. 杨凌：西北农林科技大学, 2024.
YANG Lihua. Effects of straw and organic manure on growth nitrogen uptake and utilization of mulched spring maize[D]. Yangling: Northwest A&F University, 2024. (in Chinese)

[7] MCNAIRN H, PROTZ R. Mapping corn residue cover on agricultural fields in Oxford County, Ontario, using Thematic Mapper[J]. Canadian Journal of Remote Sensing, 1993, 19(2): 152－159.

[8] HARTWIG R, LAFLEN J M. A meterstick method for measuring crop residue cover[J]. Journal of Soil and Water Conservation, 1978, 33: 90－91.

[9] LAFLEN J M, AMEMIYA M, HINTZ E A. Measuring crop residue cover[J]. Journal of Soil and Water Conservation, 1981, 36(6): 341－343.

[10] MORRISON J E, HUANG C H, LIGHTLE D T, et al. Residue measurement techniques[J]. Journal of Soil and Water

Conservation, 1993, 48(6): 478 – 483.

[11] WILLIAMS D J, WILKINS E D, DOUGLAS L C, et al. Mow-plow crop residue management influence on soil erosion in north-central Oregon[J]. Soil & Tillage Research, 2000, 55(1): 71 – 78.

[12] 丁杰源, 奚小波, 张宝峰, 等. 基于 YOLOv5s – CBS 的番茄成熟度识别方法[J]. 扬州大学学报(自然科学版), 2025, 28(2): 29 – 37.
DING Jieyuan, XI Xiaobo, ZHANG Baofeng, et al. Tomato maturity recognition method based on YOLOv5s – CBS[J]. Journal of Yangzhou University(Natural Science Edition), 2025, 28(2): 29 – 37. (in Chinese)

[13] 黄思琪,张正华,郭丽瑞,等. 基于 Cutout – ResNet50 的野外环境水稻病害识别系统[J]. 扬州大学学报(自然科学版), 2024, 27(3): 37 – 45.
HUANG Siqi, ZHANG Zhenghua, GUO Lirui, et al. Rice disease identification system in the field environment based on Cutout – ResNet50[J]. Journal of Yangzhou University(Natural Science Edition), 2024, 27(3): 37 – 45. (in Chinese)

[14] XU Yan, YU Hongpeng, WU Liyan, et al. Contingency planning of visual contamination for wheeled mobile robots with chameleon-inspired visual system[J]. Electronics, 2023, 12(11): 2365.

[15] 王丽丽, 徐岚俊, 魏舒, 等. 基于 Sauvola 与 Otsu 算法的秸秆覆盖率图像检测方法[J]. 农业工程,2017,7(4):29 – 35.
WANG Lili, XU Lanjun, WEI Shu, et al. Straw coverage detection method based on Sauvola and Otsu segmentation algorithm[J]. Agricultural Engineering, 2017, 7(4): 29 – 35. (in Chinese)

[16] 安晓飞, 代均益, 路振田, 等. 基于彩色空间距离优化的秸秆覆盖率检测算法研究[J]. 农业机械学报, 2023, 54(增刊 2): 229 – 234.
AN Xiaofei, DAI Junyi, LU Zhentian, et al. Development of rapid detection algorithm of straw coverage rate based on color spatial distance gray and K-means method[J]. Transactions of the Chinese Society for Agricultural Machinery, 2023, 54(Supp. 2): 229 – 234. (in Chinese)

[17] MA J T, WU F, XIE H X, et al. Uniformity detection for straws based on overlapping region analysis[J]. Agriculture, 2022, 12(1): 80.

[18] 岳继博, 李婷, 宋洁, 等. 基于高光谱和深度学习的水稻秸秆覆盖度遥感估算[J]. 河南农业大学学报, 2024, 58(5): 838 – 851.
YUE Jibo, LI Ting, SONG Jie, et al. Estimation of rice residue cover by remote sensing based on hyperspectral and deep learning[J]. Journal of Henan Agricultural University, 2024, 58(5): 838 – 851. (in Chinese)

[19] CAI W, ZHAO S, WANG Y, et al. Estimation of winter wheat residue coverage using optical and SAR remote sensing images[J]. Remote Sensing, 2019, 11(10): 1163.

[20] 马钦, 万传峰, 卫建, 等. 基于 U – Net 和特征金字塔网络的秸秆覆盖率计算方法[J]. 农业机械学报, 2023, 54(1): 224 – 234.
MA Qin, WAN Chuanfeng, WEI Jian, et al. Calculation method of straw coverage based on U – Net network and feature pyramid network[J]. Transactions of the Chinese Society for Agricultural Machinery, 2023, 54(1): 224 – 234. (in Chinese)

[21] 刘媛媛, 周小康, 王跃勇, 等. 改进 U – Net 模型的保护性耕作田间秸秆覆盖检测[J]. 光学精密工程, 2022, 30(9): 1101 – 1112.
LIU Yuanyuan, ZHOU Xiaokang, WANG Yueyong, et al. Straw coverage detection of conservation tillage farmland based on improved U – Net model[J]. Optics and Precision Engineering, 2022, 30(9): 1101 – 1112. (in Chinese)

[22] ZHOU D, LI M, LI Y, et al. Detection of ground straw coverage under conservation tillage based on deep learning[J]. Computers and Electronics in Agriculture, 2020, 172: 105369.

[23] EVERINGHAM M, GOOL L, WILLIAMS I K C, et al. The pascal visual object classes(VOC) challenge[J]. International Journal of Computer Vision, 2010, 88(2): 303 – 338.

[24] LIN T Y, MAIRE M, BELONGIE S, et al. Microsoft COCO: common objects in context[C]// Computer Vision-ECCV2014: 13th European Conference. Zurich: Springer, 2014: 740 – 755.

[25] FENG Y, GHADA Z, SANDEEP A, et al. Assessing Inter-annotator agreement for medical image segmentation[J]. IEEE Access: Practical Innovations, Open Solutions, 2023, 11:21300 – 21312.