

曹英丽, 陈晓安, 蔺雨桐, 等. 基于多尺度信息融合的杂草语义分割方法[J]. 沈阳农业大学学报, 2024, 55(6): 743-751.

CAO Yingli, CHEN Xiaolan, LIN Yutong, et al. Semantic segmentation of weeds based on multiscale information fusion[J]. Journal of Shenyang Agricultural University, 2024, 55(6): 743-751.

## 基于多尺度信息融合的杂草语义分割方法

曹英丽<sup>1,2</sup>, 陈晓安<sup>1</sup>, 蔺雨桐<sup>1</sup>, 李 严<sup>1</sup>, 郭忠辉<sup>1</sup>

[1. 沈阳农业大学 信息与电气工程学院, 沈阳 110161; 2. 国家数字农业区域创新分中心(东北), 沈阳 110161]

**摘 要:** 玉米田间杂草有效分割是无人机精准变量施药的必要前提, 针对传统 CNN 语义分割模型难以克服因玉米与杂草相互遮挡、小目标杂草漏检等问题, 以玉米 3-5 叶龄期田间杂草无人机正射数码影像为对象, 提出 Transformer 与 CNN 多尺度信息融合的杂草语义分割模型 TFPSP-CA。首先, 将基础模型 PSPNet 特征金字塔替换为 BiFPN, 加强特征融合, 增强模型对图像细节的学习能力和上下文信息获取能力; 使用 MobileNet 系列网络替换原有的 ResNet 网络, 加快模型预测速度、减小模型规模, 得到改进的杂草分割模型 FPSPNet。结果表明: 改进的分割模型 FPSPNet 在小目标、中目标、大目标的 mIoU、PA 分别为 84.48%、89.36%、82.05%、89.34% 和 83.32%、89.26%, 尤其对小目标杂草分割精度提升明显, 其 mIoU、PA 比基础分割模型分别提高 4.60% 和 2.42%。其次, 进一步在 FPSPNet 金字塔模块的输出端引入 Transformer 特征输出模块并行连接从而获得更多尺度信息, 并加入结合坐标与通道信息的 CA 注意力机制, 得到 TFPSP-CA 杂草分割模型, 该模型能够更好地捕捉全局特征和长距离依赖关系, 从而提高对复杂环境下杂草分割问题的处理能力。结果表明: 改进后的 TFPSP-CA 模型在无遮挡、轻度遮挡、严重遮挡情况下杂草分割 mIoU 和 PA 分别为 90.21%、91.98%、89.44%、89.11% 和 87.59%、87.53%, 改进模型在严重遮挡情况下精度提升明显, 对比原始模型 PSPNet 和 FPSPNet, mIoU、PA 分别提升 6.34%、2.27% 和 10.96%、5.22%。

**关键词:** 杂草分割; 玉米苗期; 精准农业; Transformer; 卷积神经网络

中图分类号: S513; TP391.41

文章编号: 1000-1700(2024)06-0743-09

文献标识码: A

开放科学(资源服务)标识码(OSID):



## Semantic Segmentation of Weeds Based on Multiscale Information Fusion

CAO Yingli<sup>1,2</sup>, CHEN Xiaolan<sup>1</sup>, LIN Yutong<sup>1</sup>, LI Yan<sup>1</sup>, GUO Zhonghui<sup>1</sup>

[1. College of Information and Electrical Engineering, Shenyang Agricultural University, Shenyang 110161, China; 2. National Regional Innovation Sub-center for Digital Agriculture (Northeast), Shenyang 110161 China]

**Abstract:** Effective segmentation of weeds in maize fields is a prerequisite for accurate variable herbicide application by UAVs. Aiming at the problems of missing detection of the traditional CNN semantic segmentation model for the mutual occlusion of maize and weeds, and the small target weeds, etc., taking UAV ortho-digital image of the field weed at three-five-leaf-age of maize as the object, we propose the weed semantic segmentation model TFPSP-CA based on multi-scale information fusion of Transformer and CNN. Firstly, the base model PSPNet feature pyramid is replaced with BiFPN to strengthen feature fusion and enhance the model's ability to learn image details and acquire contextual information; the original ResNet network is replaced with the MobileNet series network to speed up the model prediction speed and reduce the model size, and the improved weed segmentation model FPSPNet is obtained. The results show that the mIoU and PA of

收稿日期: 2024-10-10

基金项目: 辽宁省教育厅重点攻关项目(JYTZD2023123)

第一作者: 曹英丽(1977-), 女, 博士, 教授, 博士生导师, 从事精准农业航空领域研究, E-mail: caoyingli@syau.edu.cn

the improved segmentation model FPSPNet are 84.48%, 89.36%, 82.05%, 89.34%, and 83.32%, 89.26% for small, medium, and large targets, respectively. The accuracy of weed segmentation is especially improved significantly for small targets, and its mIoU and PA are increased by 4.60% and 2.42% respectively compared to the base segmentation model. Secondly, the Transformer feature output module is further introduced at the output end of the FPSPNet pyramid module to connect in parallel to obtain more scale information. The CA attention mechanism combining coordinate and channel information is added to obtain the TFPSP-CA weed segmentation model. The model can better capture the global features and long-distance dependence, thus improving the processing ability of the weed segmentation problem in complex environments. The results show that the improved TFPSP-CA model has 90.21%, 91.98%, 89.44%, 89.11%, and 87.59%, 87.53% of weed segmentation mIoU and PA in the case of no occlusion, mild occlusion, and severe occlusion, respectively, and the accuracy of the improved model in the case of severe occlusion improves significantly, compared with the original models PSPNet and FPSPNet, the mIoU and PA are improved by 6.34%, 2.27% and 10.96% and 5.22%, respectively.

**Key words:** weed segmentation; maize seedling stage; precision agriculture; Transformer; convolutional neural network

玉米田间杂草抢夺水肥光等资源滋生细菌加剧土壤环境恶化、引发病虫害,造成玉米生产质量下降<sup>[1-2]</sup>。据联合国粮农组织(FAO)研究表明<sup>[3-4]</sup>,世界粮食由于受到病、虫、草害等威胁平均造成30%~35%的产量损失,因杂草干扰造成的损失达10%。因此田间杂草有效查杀和管理对于提升玉米产量、保障粮食安全具有重要意义。我国农业农村部粮食作物田杂草科学防控技术方案(2023年)建议,北方一年一熟玉米种植区,在播种季节土壤墒情较好的地块,杂草防控采用“一封一杀”策略<sup>[5]</sup>。“一封”是指玉米播种后至出苗前,选择合适的除草剂对土壤进行封闭处理;“一杀”是指玉米出苗后3~5叶期,通过喷施除草剂的方式,以达到杀灭杂草的目的。目前,玉米田间杂草“一杀”管理中仍主要采用化学防治手段,不同地块采用无差别喷洒除草剂,大大增加了化学农药喷施量,对环境造成巨大影响。为实现除草剂精准喷施、减少农药不良影响,亟须探索快速高效的农田杂草分布检测、施药专题图制作、植保无人机变量喷施作业等技术方法,为更好支撑变量作业处方图制定,田间杂草语义分割是首要前提<sup>[6]</sup>。

基于无人机平台的田间杂草语义分割模型构建研究中,一些经典的语义分割网络如SegNet、U-Net和Deeplab系列等陆续应用于水稻、棉花和玉米等杂草识别之中,皆取得了一定的识别和分割效果。ESPEJO-GARCIA等<sup>[7]</sup>基于迁移学习和对抗神经网络相结合的方式,使用Xception网络结合GANs来对杂草进行识别分割研究,测试后性能达到99.07%,并且经过试验在其他数据集上也有很好的效果准确率达到93.23%。王璨等<sup>[8]</sup>提出了一种基于双注意力机制语义分割网络的苗期玉米与杂草分割模型,在FCN网络的基础上引入双注意力机制,并改进损失函数和改进迁移学习策略,相比于原模型提高1.47%,识别速度达到 $15.9 \text{ f} \cdot \text{s}^{-1}$ 。上述研究针对Transformer结构进行了杂草分割识别研究,可更好地应对复杂环境下分割问题,但是模型预测速度往往未能达到无人机施药作业中对实时检测的要求。

PSPNet是一种深度学习语义分割模型,它最主要的特点是采用金字塔池化模块PPM(Pyramid Pooling Module)实现不同尺度信息的有效融合。相比上述语义分割模型,PSPNet能够更加有效地捕获全局信息<sup>[9]</sup>,增强语义分割的效果。张艺等<sup>[10]</sup>提出一种基于改进PSPNet的森林火烧迹地检测模型。该模型以带空洞卷积的ResNet34作为主干网络,并在主干网络内部融合RFB(Receptive Field Block)模块和ULSAM模块来增强其特征提取能力。改进PSPNet模型平均交并比和总体准确率分别提高1.52%和0.67%。陈奕伏等<sup>[11]</sup>提出一种改进的生物回波提取网络:MFF-PSPnet。MFF-PSPnet是一种轻量级网络,其参数量相比一般生物回波提取网络减少了10.2%,提取生物回波的准确率达到98.1%,IoU达到94.8%。尽管PSPNet在语义分割任务中表现良好,但其对小目标的检测及遮挡物体的识别仍存在改善的空间。

为解决玉米与杂草相互遮挡与无人机尺度下小目标杂草漏检等问题,提出一种基于改进基础模型PSPNet的玉米苗期杂草识别模型FPSPNet模型,通过在特征金字塔模块中引入BiFPN加强特征融合模块提升模型对杂草分割精度。为增强模型复杂环境处理能力,提出Transformer结合CNN多尺度信息融合的语义分割方法,通过并行的方式将FPSPNet的特征金字塔模块的输出与Transformer模块连接,从而获得更多尺度的信息,并在模型中引入CA注意力机制,得到TFPSP-CA分割模型。通过使

用不同遮挡、不同高度的杂草图像对模型进行测试。

## 1 玉米杂草图像采集与处理

### 1.1 玉米杂草图像采集

田间试验地点位于辽宁省沈阳市沈阳农业大学试验基地,经度  $123.5^{\circ}$ ,纬度  $41.8^{\circ}$ 。试验使用大疆公司的精灵4RTK无人机配备2000万像素数码相机用于捕获RGB图像,最大上升速度为  $6\text{ m}\cdot\text{s}^{-1}$ ,最大下降速度为  $3\text{ m}\cdot\text{s}^{-1}$ 。保障地块拍摄的全面性,相机配置为垂直俯视角,并设定20%的航向和旁向重叠度,分辨率为  $5472\times 3648\text{ px}$ 。玉米田间杂草无人机影像数据采集时期为玉米3~5叶期,于播种后15 d开始采集3次。图1a为玉米幼苗破土时期图像,图像中玉米幼苗矮小,并且由于玉米播种前的除草工作,图像杂草样本较少;图1b为玉米生长至3叶期的玉米图像,此阶段为除草的关键时期,图像中玉米与杂草像素比例均衡,玉米幼苗与杂草遮挡问题较轻;图1c时期玉米与杂草的遮挡问题较为严重。无人机拍摄高度为2,3,5 m,不同高度、不同时间段共采集玉米杂草图像763幅。

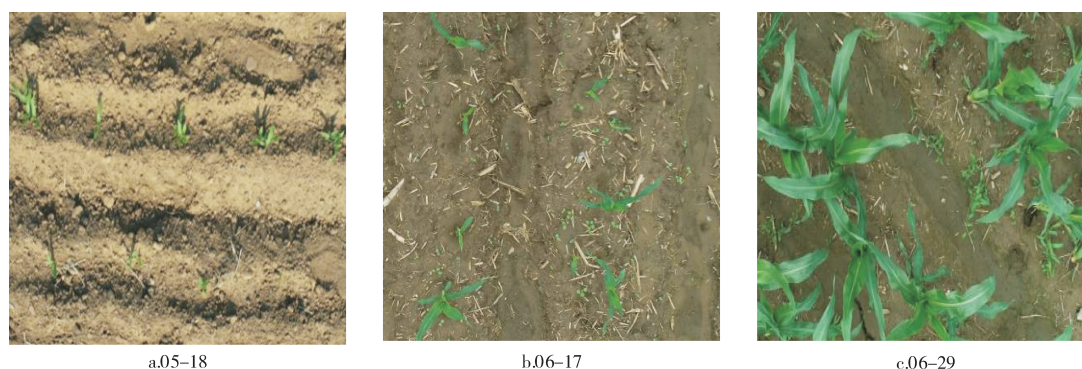


图1 玉米苗多时期数据采集

Figure 1 Data collection of maize seedling at multi-period

### 1.2 玉米杂草图像预处理与数据标注

对无人机获取的763幅数码影像进行预处理,原始图像分辨率为  $5472\text{ px}\times 3648\text{ px}$ ,模型训练所需时间和计算量较大,所以首先在原图片进行了随机裁剪,得到  $512\text{ px}\times 512\text{ px}$  大小图片共1526幅。筛选包含杂草的图像1300幅,并将图像分为禾本杂草、阔叶杂草两类进行标注,使用图像标注工具LabelMe生成杂草的标注掩码,使用多边形选择工具来标注玉米和杂草。标注结果以JSON文件格式保存。

为保证试验数据集多样性和丰富性,使用数据增强技术扩展数据规模,减少模型对特定图像的依赖。本研究使用几何翻转、增强对比度与饱和度等方式对玉米杂草原图像和从JSON文件中提取的标签图像进行数据增强。获得7360张玉米杂草图像,按照7:2:1随机分为5152张训练集、1472张验证集和736张测试集。为后续模型测试,筛选出无遮挡、轻微遮挡、严重遮挡影像数据集(图2);按照coco数据集大中小目标定义标准:小目标(目标边界框面积  $<32^2\text{ px}$ )、中目标(目标边界框面积在  $32^2\sim 96^2\text{ px}$  之间)、大目标(目标边界框面积  $>96^2\text{ px}$ ),筛选出大、中、小杂草数据集。



图2 玉米苗与杂草不同遮挡影像

Figure 2 Images of maize seedlings and weeds under different covered situations

## 2 玉米杂草语义分割模型TFPSP-CA构建

### 2.1 玉米杂草分割基线模型PSPNet

PSPNet(Pyramid Scene Parsing Network)是一种基于CNN的语义分割模型,适用于复杂场景及包含多尺度信息的图像语义分割任务。由于对计算资源要求相对较低,其在实时语义分割中也具有广泛应用。综合考虑分割效果及推理速度<sup>[12]</sup>,本研究选择PSPNet作为玉米杂草分割的基线模型,其主要由特征提取网络与金字塔池化网络两个部分组成图3。

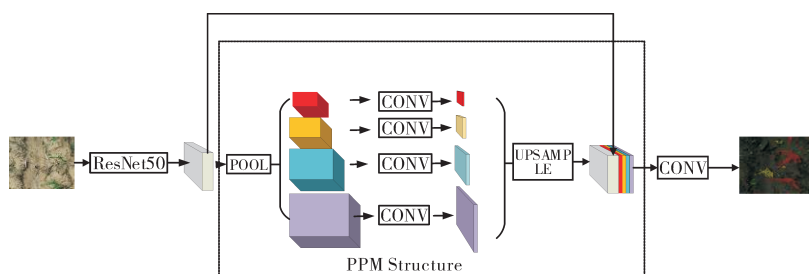


图3 PSPNet结构图

Figure 3 PSPNet structure diagram

### 2.2 FSPSPNet网络模型构建

为提高PSPNet网络模型对杂草的分割性能,本研究对其在两方面进行了改进,构建FSPSPNet网络模型。首先,在原始PSPNet的金字塔池化网络中引入BiFPN模块,增强网络的特征提取能力;其次,将PSPNet的主干网络替换为MobileNet系列网络,实现轻量化处理,提升运算速度。

2.2.1 特征金字塔模块改进 玉米田间环境复杂、杂草目标较小,传统金字塔结构难以提取特征。PSPNet模型在特征融合时会丢失信息。因此,改进PSPNet的金字塔模块,引入BiFPN加强特征融合,提升模型对图像细节和上下文信息的学习能力,减少信息丢失<sup>[13-14]</sup>。本研究BiFPN利用快速归一化融合方法,计算如式(1),通过结合不同层特征,以提升计算速度和准确性。模型结构如图4,主干网络输出特征层到金字塔模块,经平均池化得到四维特征图,输入BiFPN中,通过横向连接和双向融合增强联系,用1×1卷积调整通道数减运算量,输出四种融合特征图。各尺度特征图融合不同上下文信息,网络获取信息能力更强。BiFPN融合特征计算方式见式(2)。

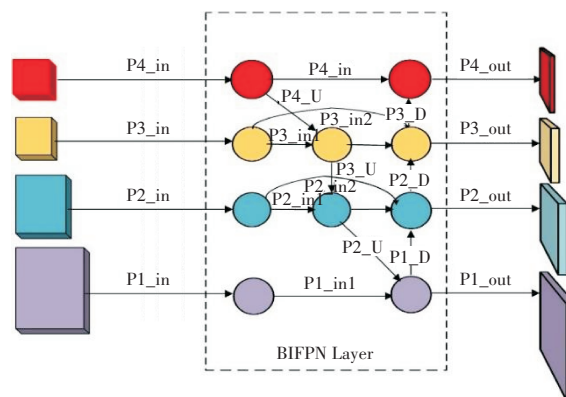


图4 BiFPN网络结构示意图

Figure 4 BiFPN structure diagram

$$0 = \sum_i \frac{w_i}{\epsilon + \sum_j w_j} \cdot I_i \quad (1)$$

$$P_3^{out} = Conv \left[ \frac{w_1 \cdot P_3^{in} + w_2 \cdot P_3^{fd} + w_3 \cdot Resize(P_2^{out})}{w_1 + w_2 + w_3 + \epsilon} \right] \quad (2)$$

式中: $w_i$ 为可学习的权重; $Resize$ 为用于分辨匹配的上采样或下采样操作; $Conv$ 为用于特征处理的卷积操作。

2.2.2 主干网络替换MobileNet 系列网络分为V1、V2和V3共3个版本<sup>[15]</sup>,其具有模型参数量少,识别精度较高,计算复杂度低等优点。PSPNet网络在基础模型中采用ResNet网络作为主干特征提取网络,具有较多的网络层参数,网络尺寸较大,推理速度较慢<sup>[16]</sup>。综合考虑计算资源、模型性能与训练时间成本等因素,本研究将使用MobileNet系列网络替换原始PSPNet的主干网络,以构建高效、轻量的杂草

分割模型。改进后的FPSPNet网络结构如图5。

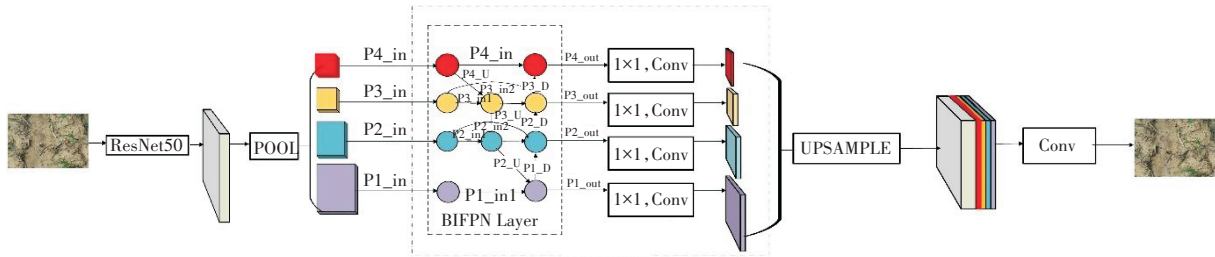


图5 FPSPNet网络结构示意图  
Figure 5 FPSPNet structure diagram

### 2.3 Transformer与卷积融合注意力的多尺度杂草识别模型

为进一步提升FPSPNet在处理玉米和杂草间存在遮挡时的分割效果,本研究在FPSPNet的金字塔网络输出端引入Transformer特征输出模块,以获取更多尺度信息,并加入CA注意力机制,得到TF-PSP-CA杂草分割模型。

2.3.1 引入Transformer特征输出模块 为解决模型应对遮挡环境与高飞行高度拍摄的杂草图像分割预测能力差的问题,本研究结合ViT(Vision Transformer)结构<sup>[17]</sup>,使用8层Transformer编码层搭建基于Transformer的特征输出模块如图6。旨在克服传统卷积神经网络在处理小目标和遮挡情况下的局限性,通过引入Transformer结构,使模型能够更好地捕捉全局特征和长距离依赖关系,从而提高对复杂环境下杂草分割问题的处理能力<sup>[18]</sup>。经过多头注意力层和多层感知机操作后<sup>[19]</sup>,生成了编码的序列数据,为了将这些编码的序列数据恢复到原始的空间顺序,进行逆操作将序列数据重新重塑,以确保编码的序列数据能够恢复到与特征金字塔输出的特征图大小相同,本研究搭建的Transformer结构特征输出模块结合ViT架构,使用8层Transformer编码层对图像的全局特征进行捕捉,每个Transformer编码层包含多头注意力层和全连接前馈网络,特征输出模块如图7。

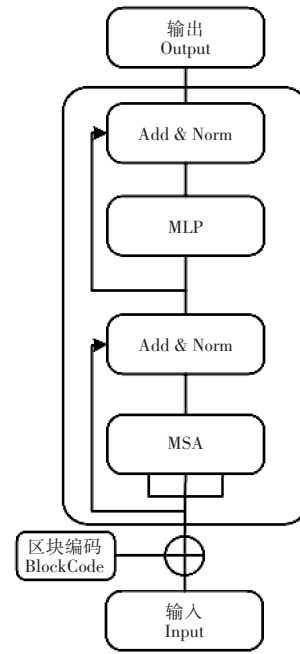


图6 Transformer编码层结构示意图  
Figure 6 Transformer structure diagram of output module

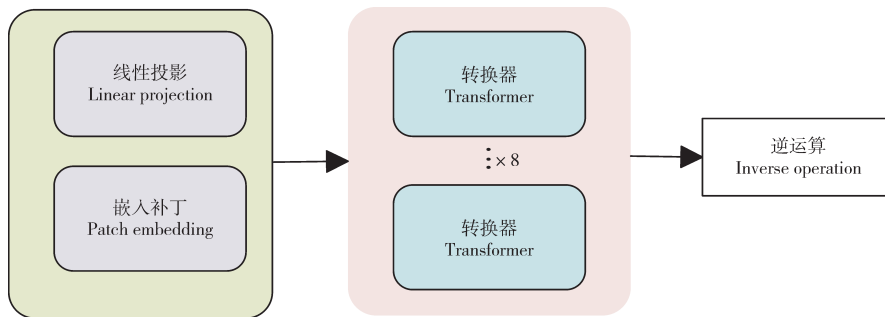


图7 Transformer结构特征输出模块结构示意图  
Figure 7 Transformer structural feature output module

2.3.2 CA注意力机制 CA(Coordinate Attention)注意力机制是一种结合空间坐标和通道信息的注意力机制<sup>[20]</sup>。其通过计算特征图上每个位置的关注度,获得更加全面的特征提取。图8为CA注意力机制结构图。CA注意力机制主要有两个部分构成:坐标信息嵌入(coordinate information embedding)和坐标注意力生成(coordinate attention generation)<sup>[21]</sup>,具体为对输入的特征图使用尺寸为(1,W)和(H,1)的池化核沿着特征图的X、Y方向进行池化操作得到两方向的特征向量 $z_c^h(h)$ 、 $Z_c^w(w)$ ,将得到的特征向量通过特征融合后使用1×1卷积进行降维后经使用Sigmoid激活,随后沿空间维度分离融合后的特征向量得到 $g^h$ 、 $g^w$ ,最终输入出结果。

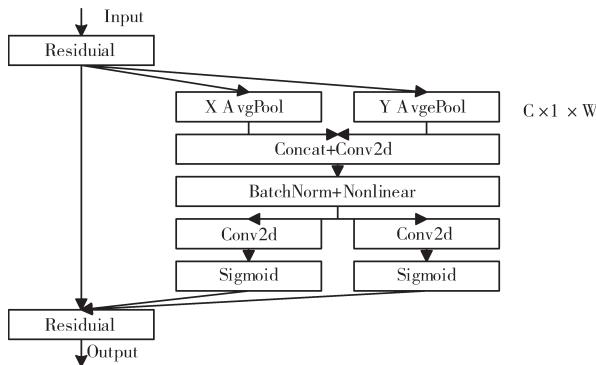


图8 CA注意力机制示意图

Figure 8 CA attention mechanism diagram

2.3.3 TFPSP-CA杂草分割模型 为应对杂草遮挡带来的分割精度下降问题,本研究提出Transformer与卷积相结合的多尺度杂草识别模型。在FPSPNet基础上,通过并行方式添加Transformer的特征输出模块,根据ViT架构构建8层Transformer的特征输出模块。并在Transformer分支与主分支输出位置添加CA注意力机制,构建本章的TFPSP模型,结构如图9。

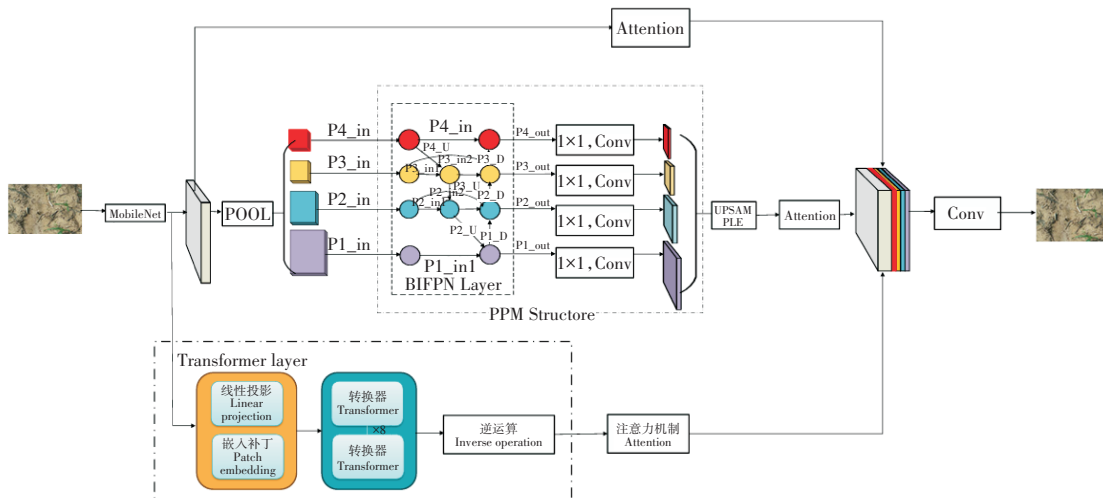


图9 TFPSP-CA结构示意图

Figure 9 TFPSP-CA structure diagram

### 3 试验结果与分析

#### 3.1 试验环境配置及参数设置

本试验平台选择在 Windows 10 操作系统下使用 PyCharm 平台结合 Pytorch 框架搭建玉米苗期杂草语义分割模型,使用 NVIDIA RTX3090 显卡搭配 CUDA 架构进行训练,模型训练参数使用随机梯度下降算法作为优化器,其内部使用的动量参数 momentum 设置为 0.9,最小学习率设置为 1e-4,权值衰减系数设置为 1e-2,初始学习率设置为 0.005 最小值设置为 1e-4,学习率权重衰减系数设置为 0.000 1,训练 Batchsize 设置为 64,迭代次数设置为 300 次。

#### 3.2 评价指标

对于玉米苗期杂草语义分割任务来说,模型需要满足的要求有 3 点:(1)在图像中尽可能准确地标注出像素的类别;(2)在满足玉米幼苗与杂草分割精度的前提下,模型尽可能满足实时检测的任务;(3)模型在玉米幼苗与杂草相互遮挡的条件下有较好的适应能力。所以本研究评价指标为像素准确率(Pixel Accuracy, PA)、平均交并比(Mean Intersection over Union, mIoU)、参数量(Params)、FPS(Frames

Per Second)等<sup>[22]</sup>。

### 3.3 杂草分割模型FPSPNet改进效果测试

由表1可知,改进后模型的mIoU和PA均有不同程度提高,其中小目标mIoU提高4.6%,PA提高2.42%,改善效果较为明显。

表1 基本模型改进前后的分割试验结果

Table 1 Segmentation test results before and after the improved model

模型 Model	mIoU/%			PA/%		
	小目标	中目标	大目标	小目标	中目标	大目标
	SmallTarget	MidTarget	MajorTarget	SmallTarget	MidTarget	MajorTarget
PSPNet	79.88	78.94	80.26	86.94	87.08	87.32
FPSPNet	84.48	82.05	83.32	89.36	89.34	89.26

由表2可知,模型在精度方面有不同程度的损失,分别降低2.21%、1.10%和2.04%,但在计算复杂度与FPS上均有显著提升,分别达到50.1,48.3,46.5  $f \cdot s^{-1}$ ,满足实时检测中FPS高于30  $f \cdot s^{-1}$ 的要求,因此选择精度下降最少的MobileNetV2作为模型主干网络<sup>[23]</sup>。

表2 替换PSPNet主干网络的对比试验结果

Table 2 Comparative test results of replacing the PSPNet backbone network

模型 Model	PA/%	mIoU/%	FPS /( $f \cdot s^{-1}$ )	Params/M
PSPNet	87.73	80.26	20.6	22.891
PSPNet-MobileNetV1	85.52	77.62	50.1	2.376
PSPNet-MobileNetV2	86.63	78.53	48.3	2.404
PSPNet-MobileNetV3	85.69	78.35	46.5	2.532

### 3.4 杂草分割模型TFPSP-CA改进效果测试

3.4.1 引入Transformer模块与不同注意力机制效果测试 由表3可知,CBAM注意力机制有着较高PA但是mIoU提升较低。添加CBAM注意力机制对大目标分割能力提升较大,而对小目标杂草分割提升较小。而使用CA注意力机制后PA于mIoU都有不同幅度提升,并且对比CBMA注意力机制对模型FPS影响较小,所以后续试验使用CA注意力机制为本章模型的注意力机制得到本章结合Transformer结构的TFPSP-CA杂草分割模型。

表3 TFPSP使用不同注意力机制对比试验结果

Table 3 Comparison of experimental results of TFPSP using different attention mechanisms

模型 Model	PA/%	mIoU/%	FPS /( $f \cdot s^{-1}$ )
TFPSP	91.32	88.92	29.3
TFPSP-SE	91.26	88.71	29.2
TFPSP-CBAM	92.05	89.73	28.4
TFPSP-CA	91.98	90.21	28.7

3.4.2 不同遮挡程度玉米杂草分割结果 由表4可知,改进后模型在不同遮挡条件下都取得了较好的效果,其中在无遮挡下,mIoU、PA分别达到90.21%和91.98%,比FPSPNet模型分别提高1.60%和1.29%,比PSPNet基础分割模型分别提高9.95%和4.25%。在轻度遮挡下,mIoU、PA分别达到89.44%和89.11%,比FPSPNet模型分别提升4.12%和0.51%,比PSPNet基础分割模型分别提高11.1%和3.17%。在严重遮挡下mIoU、PA分别达到87.59%和87.53%,比FPSPNet模型分别提升6.34%和2.27%,比PSPNet基础模型分别提高10.96%和5.22%。结合Transformer模块后,模型应对杂草分割问题上分割效果有一定提升。图10为不同遮挡条件下PSPNet模型、FPSPNet模型与TFPSP-CA模型分割对比图,其中方框为后续标注,从标注区域可以看出改进后模型在不同遮挡条件下都有较大的提升,并且在遮挡程度严重时模型分割的平均交并比下降较少,所以使用Transformer结构与CNN结合的杂草分割方式应对复杂环境下遮挡问题的能力较强,所以本节模型更适用于不同遮挡条件下无人机变量施药。

表4 不同遮挡条件下模型性能测试结果

Table 4 Test results of model performance under different covered conditions

模型 Model	mIoU/%			PA/%		
	无遮挡	轻度遮挡	严重遮挡	无遮挡	轻度遮挡	严重遮挡
	No covered	Mild covered	Heavy covered	No covered	Mild covered	Heavy covered
PSPNet	80.26	78.34	76.63	87.73	85.94	82.31
FPSPNet	88.61	85.32	81.25	90.69	88.60	85.26
TFPSP-CA	90.21	89.44	87.59	91.98	89.11	87.53

## 4 讨论与结论

针对无人机尺度下玉米田间杂草分割的精度下降问题,构建了Transformer与CNN多尺度信息融合的杂草语义分割模型TFPSP-CA。首先,通过无人机拍摄并预处理图像数据,构建包含7360幅图像的玉米杂草数据集,为模型训练提供了丰富的样本。其次,改进了PSPNet模型,构建了FPSPNet模型,通过引入BiFPN加强特征融合模块并替换主干网络为MobileNetV2,有效解决了小目标杂草漏检问题。最后,针对田间玉米与杂草相互遮挡的问题,在FPSPNet模型基础上引入了Transformer特征输出模块和CA注意力机制,构建了TFPSP-CA模型。对比原始模型与其他改进方法,TFPSP-CA模型在不同遮挡条件下的分割精度均有显著提升,实现了玉米田间杂草的有效分割。

综上所述,本研究提出的改进策略可为后续玉米杂草的有效分割提供参考。

本研究在玉米3~5叶龄杂草补杀关键期,采用无人机搭载数码相机,拍摄了不同飞行高度、多时序玉米田间杂草正射影像;并对图像数据进行随机裁剪、筛选等预处理,使用LabelMe进行图像标注;使用几何变换、颜色变换、Mosaic数据增强等方法丰富样本图像,最终获得玉米杂草图像数据集共7360幅图像,按照7:2:1随机分为训练集、验证集和测试集。并根据杂草分割模型测试需求,将测试集分别按照杂草种类、目标大小、田间遮挡程度等进行了分类。基于杂草分割基本模型PSPNet改进的分割模型FPSPNet解决了无人机尺度下小目标杂草漏检问题,在特征金字塔模块中引入BiFPN加强特征融合模块,将主干网络替换为MobileNetV2,通过不同层级间进行信息交换帮助网络整合不同尺度的特征信息,减少了后续在特征融合中信息丢失的现象,测试结果表明改进后模型对大、中、小目标检测精度均有提升,尤其对小目标杂草分割mIoU、PA比基本分割模型分别提升4.60%和2.42%,有效解决了小目标杂草漏检问题。针对田间玉米与杂草相互遮挡而引起的杂草分割精度降低问题,在FPSPNet模型基础上并行引入Transformer特征输出模块,并添加了结合坐标与通道信息的CA注意力机制,构建了杂草分割模型TFPSP-CA。田间不同遮挡数据集测试结果表明,相比于基础杂草分割模型,TFPSP-CA模型在严重遮挡、轻度遮挡、无遮挡条件下mIoU、PA分别提升10.96%、5.22%、11.1%、3.17%和9.95%、4.25%。最后应用模型TFPSP-CA实现玉米田间杂草有效分割。

### 参考文献:

- [1] CALLO A, MANSOURI M. Food security in global food distribution networks: A systems thinking approach[C]//2024 IEEE International Systems Conference (SysCon). Montreal, QC, Canada. IEEE, 2024: 1-6.
- [2] JAHAN N, AKILAN T, PHALKE A R. Machine learning for global food security: A concise overview[C]//2022 IEEE International Humanitarian Technology Conference (IHTC). Ottawa, ON, Canada. IEEE, 2022: 63-68.
- [3] KOVALCHUK O, BEREZKA K, DANYLYUK I, et al. Modeling Russian-Ukrainian war impact on global food safety. Preliminary evaluations[C]//2023 13th International Conference on Advanced Computer Information Technologies (ACIT). Wrocław, Poland. IEEE, 2023: 382-387.

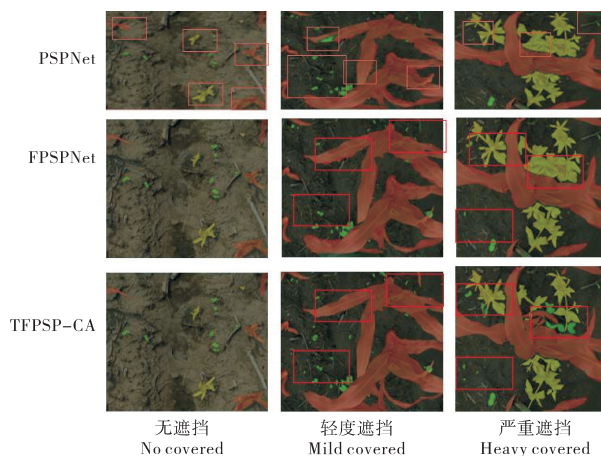


图10 不同遮挡条件对比图

Figure 10 Comparison of different covered conditions

- [4] PRADEEP G, RAYEN T D V, PUSHPALATHA A, et al. Effective crop yield prediction using gradient boosting to improve agricultural outcomes[C]//2023 International Conference on Networking and Communications (ICNWC). Chennai, India. IEEE, 2023:1-6.
- [5] 中华人民共和国农业农村部. 2023年粮食作物田杂草科学防控技术方案[EB/OL]. [2023-02-16]. [https://www.moa.gov.cn/ztl/2023cg/jszd\\_29356/202302/t20230220\\_6420966.htm](https://www.moa.gov.cn/ztl/2023cg/jszd_29356/202302/t20230220_6420966.htm).
- [6] SHAHRIN F, ZAHIN L, RAHMAN R, et al. Agricultural analysis and crop yield prediction of habiganj using multi-spectral bands of satellite imagery with machine learning[C]//2020 11th International Conference on Electrical and Computer Engineering (ICECE). Dhaka, Bangladesh. IEEE, 2020:21-24.
- [7] ESPEJO-GARCIA B, MYLONAS N, ATHANASAKOS L, et al. Combining generative adversarial networks and agricultural transfer learning for weeds identification[J]. Biosystems Engineering, 2021, 204:79-89.
- [8] 王 璨, 武新慧, 张燕青等. 基于双注意力语义分割网络的田间苗期玉米识别与分割[J]. 农业工程学报, 2021, 37(9): 211-221.
- [9] 王海舰, 刘丽丽, 赵雪梅, 等. 基于改进 PSPnet-MobileNetV2 的煤岩界面快速精准识别[J]. 振动. 测试与诊断, 2024, 44(4):793-800, 832-833.
- [10] 张 艺, 马永军, 王广来, 等. 基于改进 PSPNet 的森林火烧迹地检测[J]. 现代信息科技, 2024, 8(17):43-48, 55.
- [11] 陈奕伏, 邓志良, 吴东丽, 等. 基于深度学习特征融合的天气雷达生物回波提取研究[J/OL]. 南京信息工程大学学报, 1-15[2024-11-07]. <https://doi.org/10.13878/j.cnki.jnuist.20240418002>.
- [12] 申瑞超. 基于 SSD 和 PSPNet 算法的轨道扣件形位偏移检测方法研究[D]. 上海: 上海工程技术大学, 2021.
- [13] 孟 勃, 史伟大. 改进 YOLOv7 的交通标志识别模型[J]. 中国图象图形学报, 2024, 29(9):2737-2752.
- [14] 陆 彬, 毛义萱, 王露. 基于改进 YOLOv8 的侧扫声呐图像目标检测方法研究[J/OL]. 水利水电快报, 1-8[2024-11-07]. <http://kns.cnki.net/kcms/detail/42.1142.TV.20241022.1921.006.html>.
- [15] ZHANG Y J, CHEN X. Lightweight semantic segmentation algorithm based on MobileNetV3 network[C]//2020 International Conference on Intelligent Computing, Automation and System (ICICAS). Chongqing, China. IEEE, 2020:429-433.
- [16] 苏 琳, 魏国峰, 汤 鹏, 等. 基于数据融合 RESNET 网络的外辐射源定位方法[J]. 哈尔滨工程大学学报, 2025(1):1-8.
- [17] 王丽华. 基于 Vision Transformer 的图像语义分割方法研究[D]. 曲阜: 曲阜师范大学, 2024.
- [18] 陈 鹏, 白 勇, 陈 旭, 等. 融合点云 Transformer 的多尺度抓取检测模型[J/OL]. 计算机工程与应用, 1-11[2024-11-08]. <http://kns.cnki.net/kcms/detail/11.2127.TP.20241030.1252.004.html>.
- [19] JHA R R, BHARDWAJ A, GARG D, et al. MHATC: Autism spectrum disorder identification utilizing multi-head attention encoder long with temporal consolidation modules[J]. Annual International Conference of the IEEE Engineering in Medicine and Biology Society, IEEE Engineering in Medicine and Biology Society Annual International Conference, 2022, 2022:337-341.
- [20] 芦碧波, 梁 迪, 杨 洁, 等. 基于改进 ENet 的复杂背景下山药叶片图像分割方法[J/OL]. 智慧农业(中英文), 1-12[2024-11-07]. <http://kns.cnki.net/kcms/detail/10.1681.s.20241107.1151.002.html>.
- [21] WEI Z H, GAO H, WANG G. Water level detection based on ghost convolution and CA attention mechanism[C]//2023 5th International Conference on Frontiers Technology of Information and Computer (ICFTIC). Qiangdao, China. IEEE, 2023, 278-282.
- [22] CHEN W X, ZHANG T, ZHAO X. Semantic segmentation using generative adversarial network[C]//2021 40th Chinese Control Conference (CCC). Shanghai, China. IEEE, 2021:8492-8495.
- [23] 庄前伟, 王志明, 吴龙貽, 等. 基于改进 SOLOv2 的穴盘幼苗图像分割方法[J]. 南京农业大学学报, 2023, 46(1):200-209.

[责任编辑 亓 国]