doi:10.6041/j.issn.1000-1298.2025.06.019

# 水稻收获无人驾驶运粮车粮厢图像轻量化分割模型研究

赵润茂[1,2]　黄嘉涛[1]　满忠贤[1]　罗锡文[1,3]　胡　炼[2,3]　何　杰[1,2]　汪　沛[1,3]　黄培奎[2,3]

(1. 华南农业大学岭南现代农业科学与技术广东省实验室, 广州 510642;
2. 农业装备技术全国重点实验室, 广州 510642;
3. 华南农业大学南方农业机械与装备关键技术教育部重点实验室, 广州 510642)

**摘要**: 针对目前无人驾驶水稻收获机向运粮车转卸稻谷时,依靠收获机和运粮车的北斗定位信息决策卸粮臂位置控制,对靶精度难以保证问题,提出一种粮厢图像视觉分割模型 GTSM,为卸粮臂对靶提供粮厢位置参考信息。在 DeepLabv3 + 结构基础上,使用轻量化主干 ShuffleNetv2 替换 Xception,将 ASPP 模块中空洞卷积替换为深度可分离卷积,然后低秩分解为微因子分解卷积,以减小模型复杂度和提高运行速度;在浅层特征分支引入 SE 通道注意力机制,提高模型对粮厢边缘、纹理等低级特征利用能力。试验结果显示,GTSM 平均交占比和平均像素准确率分别达到 96.06% 和 98.69%,较基准 DeepLabv3 + 分别提升 0.78、0.67 个百分点;同时,模型复杂度明显改善,参数量和内存占用量仅为原来的 1/9,推理速度提高 166%。试验结果表明,提出的 GTSM 兼顾分割精度和推理速度,可为田间运粮车粮厢自动化分割提供参考依据。

**关键词**: 水稻收获; 运粮车; 无人驾驶; 粮厢图像分割; DeepLabv3 +; 多机协同作业

**中图分类号**: S24　　**文献标识码**: A　　**文章编号**: 1000-1298(2025)06-0196-09　　**OSID**:

# Research on Lightweight Image Segmentation Model for Grain Tank of an Unmanned Grain Cart in Rice Harvesting

ZHAO Runmao[1,2]　HUANG Jiatao[1]　MAN Zhongxian[1]　LUO Xiwen[1,3]　HU Lian[2,3]
HE Jie[1,2]　WANG Pei[1,3]　HUANG Peikui[2,3]

(1. *Guangdong Laboratory for Lingnan Modern Agriculture*, *South China Agricultural University*, *Guangzhou* 510642, *China*
2. *State Key Laboratory of Agricultural Equipment Technology*, *Guangzhou* 510642, *China*
3. *Key Laboratory of Key Technology on Agricultural Machine and Equipment*, *Ministry of Education*, *South China Agricultural University*, *Guangzhou* 510642, *China*)

**Abstract**: Aiming to address the issue of low targeting accuracy in controlling the unloading arm position during rice transfer from unmanned rice harvesters to grain transport vehicles, which relies on Beidou positioning information of the harvester and transport vehicle, a GTSM network for visual segmentation of grain compartment images was proposed to provide positional reference information for the unloading arm. Based on the DeepLabv3 + architecture, the lightweight ShuffleNetv2 backbone replaced Xception, and the atrous convolutions in the ASPP module were replaced with depthwise separable convolutions, followed by low-rank decomposition into micro-factorized convolutions to reduce model complexity and improve inference speed. Additionally, an SE channel attention mechanism was introduced in the shallow feature branch to enhance the model's ability to utilize low-level features such as grain compartment edges and textures. Experimental results showed that GTSM achieved a mean intersection over union (mIoU) of 96.06% and a mean pixel accuracy (mPA) of 98.69%, representing improvements of 0.78 and 0.67 percentage points, respectively, over the baseline DeepLabv3 +. Meanwhile, model complexity was significantly reduced, with parameter count and memory usage reduced to 1/9 of the original, and inference speed was increased by 166%. These results demonstrated that the proposed GTSM balanced segmentation accuracy and inference speed, providing a reference for automated grain compartment

收稿日期: 2025-05-03　修回日期: 2025-05-16

**基金项目**: 国家重点研发计划项目(2021YFD2000600)和农业装备技术全国重点实验室(华南农业大学)开放课题项目(SKLAET-202404)

**作者简介**: 赵润茂(1990—),男,副教授,博士,主要从事水田智能农机装备研究,E-mail: rmzhao@scau.edu.cn

**通信作者**: 胡炼(1984—),男,研究员,博士,主要从事智能农机装备和无人农场研究,E-mail: lianhu@scau.edu.cn

segmentation in field grain transport vehicles.

**Key words**: rice harvesting; grain cart; driverless operation; grain tank image segmentation; DeepLabv3 +; multi-vehicle collaboration operation

## 0　引言

在水稻机械化收获过程中,受粮厢容量限制,收获机需多次往返田间与卸粮点[1]。通过合理配置运粮车转运粮食,可有效提升收获效率[2]。人工驾驶模式下,操作人员通过调整收获机与卸粮臂位置和角度,可使其精准对准运粮车粮厢。而在无人驾驶收获机–运粮车协同作业场景中,卸粮行为完全依赖运粮车与收获机的相对位置决策系统[1],若二者位置控制或卸粮臂定位出现偏差,将直接导致卸粮臂与运粮车粮厢对接失效,致使卸粮任务失败。为实现精准对靶与高效卸粮,要准确感知收获机与运粮车的空间相对位置和姿态,图像分割可有效提取粮厢边界与位置信息,为后续的定位与控制决策提供支撑。

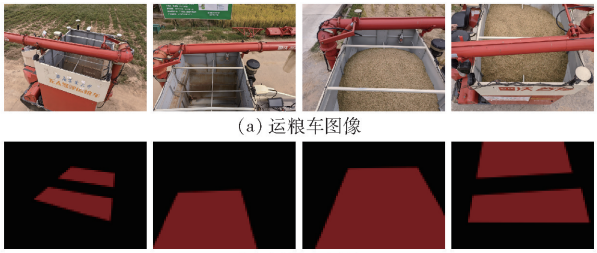
视觉传感器可以感受视野内的语义信息,结合深度学习算法,实现特定目标的分割和提取[3-5]。但是,受推理速度的限制,难以在边缘设备实际应用。为此,部分学者通过网络结构优化[6-11]去除冗余参数,平衡模型分割精度和推理速度,开发适合资源受限等嵌入式设备的部署模型。

基于深度学习的语义分割在田间道路识别、病害分割、作物提取等任务的应用,验证了复杂大田环境下模型分割的有效性及可行性。针对运粮车粮厢颜色与水稻作物颜色相近,卸粮臂出谷过程有大量粉尘影响图像质量,田间光线变化大,粮厢边缘有其它设备导致轮廓不明显的问题,本文基于精度较高的 DeepLabv3 + 模型进行粮厢分割实验。为平衡模型分割准确性与推理速度,满足在嵌入式等轻量化设备的部署,提出一种轻量化运粮车粮厢分割模型(Grain tank segmentation model, GTSM),基于 DeepLabv3 +,选择轻量化的 ShuffleNetv2 替换原网络中的特征提取主干,引入微因子分解卷积替换空间空洞金字塔池化模块(Atrous spatial pyramid pooling, ASPP)中的空洞卷积,以减少模型参数量并提高推理速度;在主干输出浅层特征分支引入通道注意力机制,以提高模型对粮厢边缘特征的提取能力。

## 1　材料与方法

### 1.1　运粮车图像获取和标注

在广东省河源市东源县柳城镇万绿智慧无人农场(24.012 180° N,115.118 903° E)采集运粮车图像共 3 891 幅。采样时间为 2024 年 11 月 1—3 日,采样设备为 Redmi K50 Pro 手机,成像分辨率为 4 000 像素 ×3 000 像素,以 jpg 格式保存,图像示例如图 1a 所示,包含运粮车粮厢有稻谷和无稻谷两种情况。经过人工筛选,选取 678 幅图像作为数据集,按照比例 9∶1 划分为训练集和验证集,得到训练集共 610 幅图像,验证集共 68 幅图像。采用 Labelme 对原始图像的运粮车粮厢和背景进行精确划分,像素值分别被标记为 0 和 1,通过可视化操作赋予对应像素颜色值,标签样例如图 1b 所示。



(a) 运粮车图像

(b) 运粮车标签

图 1　运粮车数据集示例

Fig. 1　Grain transport vehicle dataset example

### 1.2　数据增强

由于较大的图像分辨率影响训练效率,采用等比例缩放并填充图像至 512 像素 ×512 像素。为了提高模型泛化能力,适应复杂田间环境,在训练过程中采用随机样本扩增的方式增加数据量,流程如图 2 所示,包括随机翻转、随机旋转、色域变换和高斯模糊等方法。每种方法触发概率为 50%。其中,随机翻转包括水平和垂直方向,随机旋转有旋转角度 90°、180°或 270°,以解决方向不变性问题;色域

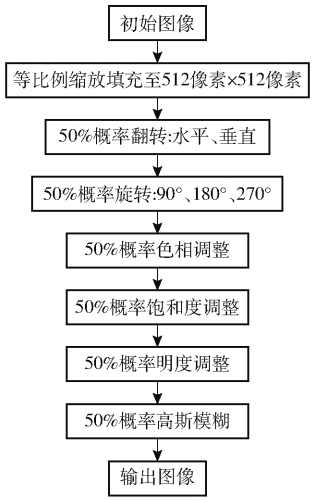



图 2　随机样本扩增流程图

Fig. 2　Random sample augmentation process

变换在 HSV 色彩空间进行,色相调整幅度为 ±10%,饱和度调整幅度为 ±30%,明度调整幅度为 ±25%,可以较好模拟田间复杂光线变化;为模拟农机震动导致成像模糊,设置高斯模糊均值为 0、标准差为 1。数据增强效果如图 3 所示。

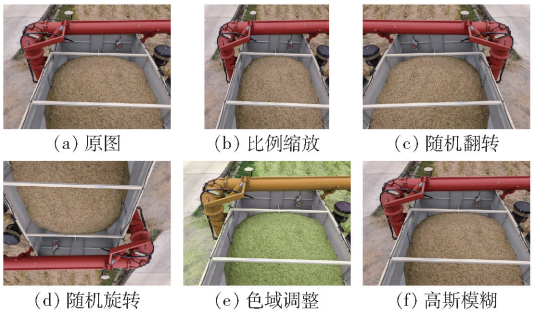



图 3 数据增强效果

Fig. 3 Data augmentation effect

### 1.3 运粮车粮厢快速分割模型

DeepLabv3 +[12] 为经典语义分割网络架构,包含编码器(Encoder)和解码器(Decoder)两部分。编码阶段先通过特征主干网络 Xception 进行下采样,获得浅层特征送入解码器,在空间金字塔池化模块通过不同膨胀率的空洞卷积和池化操作,利用不同感受野进行多尺度信息提取,获得深层特征。在解码阶段,深层特征经过上采样恢复分辨率后与浅层特征融合,最终通过双线性插值(Bilinear interpolation)上采样后恢复图像原始尺寸,输出像素级分割结果。

得益于包含大量参数的 Xception 特征提取主干[13]和复杂度较高的空间空洞金字塔池化模块[14],标准的 DeepLabv3 + 网络分割精度高,但也限制了其在轻量化嵌入式设备的部署。因此,为了减少模型参数量、降低计算复杂度,提高推理速度,采用 ShuffleNetv2 作为轻量化特征提取主干,并引入微因子分解卷积,通过低秩分解降低 ASPP 模块参数量。考虑到运粮车粮厢边缘不规则,存在非目标干扰因素,导致边缘分割精度不高的问题,在主干提取低级特征输出后引入 SE 注意力机制,赋予不同通道关注权重,提高模型对颜色、边缘和纹理等浅层特征的学习能力,GTSM 结构如图 4 所示。

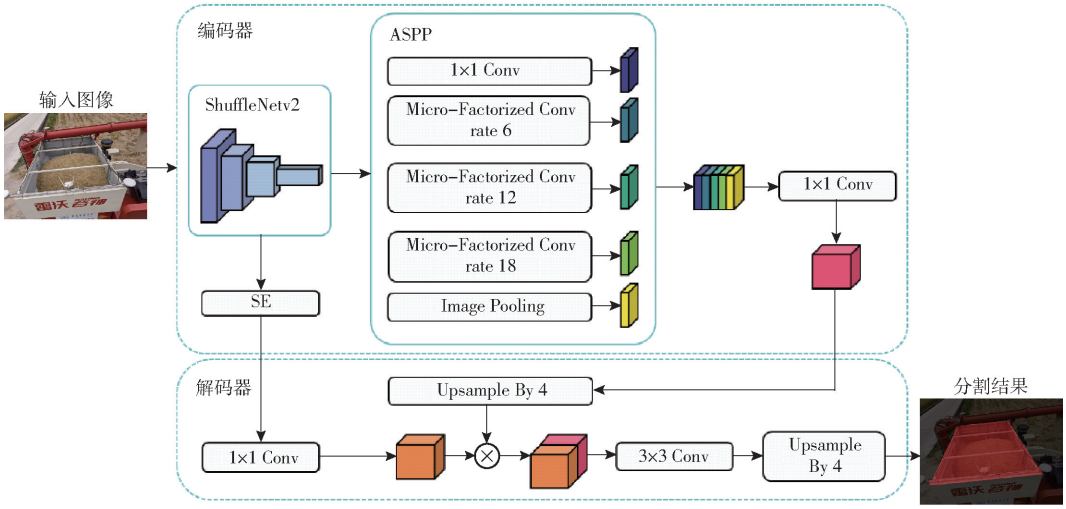



图 4 GTSM 结构

Fig. 4 GTSM structure

#### 1.3.1 ShuffleNetv2 轻量化特征提取主干

ShuffleNetv2[15]是以模型运行速度为优化对象的轻量级网络架构,旨在保证模型性能的同时大幅减小参数量并加快运算速度。输入图像依次经过第 1 阶段的 3 × 3 卷积核和全局最大池化进行两次下采样,然后重复 3 个相同的单元(stage)进行深层特征提取。每个单元都由 stride 为 1 的基本模块和 stride 为 2 的下采样模块组成,两者结构如图 5 所示,每个模块均含 2 个分支,通过拼接(Concat)后进行通道重排(Channel Shuffle)。经过 5 次下采样后,特征图尺寸变为原来的 1/32,通过 1 × 1 卷积调整通道数为类别数,经过全局池化得到全连接层后分类。

本研究提出将 ShuffleNetv2 模块作为 DeepLabv3 + 轻量化特征提取主干,网络结构及详细参数如表 1 所示。在改进标准 ShuffleNetv2 网络过程中,首先,综合考虑分割精度和计算效率,选择宽度因子为 1.5 的 ShuffleNetv2 作为基准特征提取主干;其次,去除用于调整输出通道数的 1 × 1 卷积、GAP 和 FC,使得特征提取主干的输出通道数为 704;随后,替换 32 倍下采样为 8 倍下采样,即将 Stage3 和 Stage4 中下采样模块的深度卷积 stride 调整为 1;最后,为了扩大感受野,将 Stage3 和 Stage4 中的深度卷积替换为深度膨胀卷积,具体地,Stage3

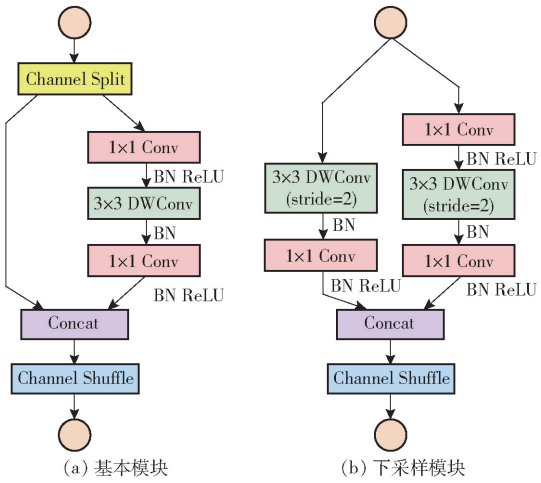



图 5　ShuffleNetv2 模块图

Fig. 5　ShuffleNetv2 module diagram

**表 1　ShuffleNetv2 轻量化特征提取主干结构**

**Tab. 1　Lightweight ShuffleNetv2 feature extraction backbone structure**

| 层 | 输出尺寸/（像素×像素） | 卷积核大小 | 步长 | 空洞率 | 重复次数 | 输出通道数 |
|---|---|---|---|---|---|---|
| Image | 512×512 | | | | | 3 |
| Conv1 | 256×256 | 3×3 | 2 | | 1 | 24 |
| MaxPool | 128×128 | 3×3 | 2 | | | |
| Stage2 | 64×64 | | 2 | | 1 | 176 |
| | 64×64 | | 1 | | 3 | |
| Stage3 | 64×64 | | 1 | 2 | 1 | 352 |
| | 64×64 | | 1 | 2 | 7 | |
| Stage4 | 64×64 | | 1 | 4 | 1 | 704 |
| | 64×64 | | 1 | 4 | 3 | |

中的卷积核膨胀率和填充设置为 2，Stage4 中的膨胀率和填充则设置为 4，以确保特征图输出尺寸保持不变。通过上述改进，ShuffleNetv2 作为 DeepLabv3 + 的轻量化特征提取主干，既降低了模型复杂度，又确保了足够语义信息供 ASPP 模块使用，提升了整体性能和效率。

### 1.3.2　微因子分解卷积和 ASPP 模块优化

微因子分解卷积（Micro-factorized convolution，MFC）由微分解点卷积（Micro-factorized pointwise convolution，MF - PWC）和微分解深度卷积（Micro-factorized depthwise convolution，MF - DWC）两部分组成，通过分解逐点卷积和深度卷积，降低网络连通性，提高计算效率。

微分解深度卷积在空间维度将一个标准卷积核拆分成多个小卷积核，具体表现为分解 $k \times k$ 大小的深度卷积核为 $k \times 1$ 和 $1 \times k$，计算量则由 $O(k^2C)$ 降低到了 $O(kC)$，$C$ 为特征通道数，分解过程如图 6 所示。

微分解点卷积将标准 $1 \times 1$ 点卷积因式分解为

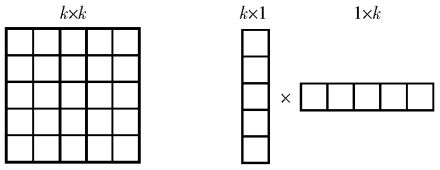



图 6　微分解深度卷积分解

Fig. 6　Micro-factorized depthwise convolution decomposition

2 个自适应点组卷积，中间插入通道重排，进而实现特征通道的压缩与扩展，矩阵分解过程如图 7 所示。计算复杂度由 $O(C^2)$ 变为 $O(2C^2/(RG))$，$G$ 为分组数，$R$ 为压缩因子。为保证计算复杂度降低的同时，使得通道间尽可能连接，应满足关系[16]

$$G = \sqrt{\frac{C}{R}} \tag{1}$$

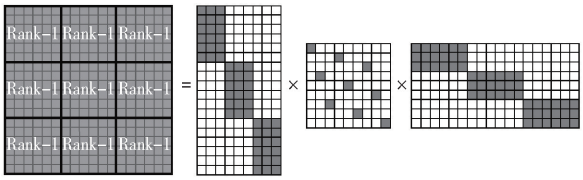



图 7　微分解点卷积矩阵分解

Fig. 7　Micro-factorized pointwise convolution matrix decomposition

ASPP 模块使用不同空洞率的空洞卷积扩展感受野，实现多尺度信息融合，避免了局部特征和全局上下文信息的丢失，然而，不同空洞率的卷积并行排列，增加了模型参数和计算量。为了降低模型复杂度，提高计算效率，在保留空洞卷积空洞率的情况下，将卷积替换为深度可分离卷积，初步减少参数量，再引入微因子分解卷积对深度可分离卷积进行低秩分解，进一步减小模型参数量，降低模型复杂度。

### 1.3.3　注意力机制

注意力机制用于提升模型对重要特征的关注能力，诸如 SE[11-18]、ECA[19-20]、SAM、CBAM[21-23]、CA[24-25]等注意力机制已经大量应用于识别、分类和分割等任务，取得了显著的效果。为准确分割粮厢边缘，使边缘纹理、颜色特征提取更加高效，本研究在浅层特征分支引入 SE 通道注意力机制。

SE 注意力机制[26]通过自适应调整每个通道的权重，支持网络强化更重要的特征通道，通过压缩、激励和重新加权 3 个阶段实现，其结构如图 8 所示。输入特征图 $X \in \mathbf{R}^{H \times W \times C}$，其中，$H$ 为高度，$W$ 为宽度，$C$ 为通道数。

首先，对每个通道进行全局平均池化，将每个通道的空间信息压缩为一个数值，公式为

$$s_c = F_{sq}(x_c) = \frac{1}{HW}\sum_{i=1}^{H}\sum_{j=1}^{W} x_c(i,j) \tag{2}$$

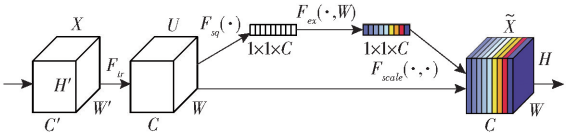



图 8 SE 模块结构

Fig. 8 SE module architecture

得到一个 $C$ 维的向量 $\boldsymbol{s}=(s_1,s_2,\cdots,s_C)$,代表每个通道的全局描述。接下来,通过 2 个全连接层 $W_1$、$W_2$对向量 $\boldsymbol{s}$ 进行非线性变换,将每个通道的评分转换为介于 0 到 1 之间的权重,过程表示为

$$z=F_{ex}(\boldsymbol{s},W)=\sigma(W_2\delta(W_1\boldsymbol{s})) \tag{3}$$

式中 $\sigma$——Sigmoid 激活函数

$\delta$——ReLU 激活函数

最后,将生成的通道权重与原始特征图对应的每个通道进行逐元素相乘,得到特征图 $\widetilde{X}$,表达式为

$$\widetilde{X}=F_{\text{scale}}(z_c,x_c)=z_cx_c \tag{4}$$

#### 1.3.4 Dice - Focal 损失函数

为了应对粮厢分割过程正负样本不平衡,缓解粮厢边缘轮廓不规则导致边缘识别和分割效果欠佳的问题,在原有交叉熵损失的基础上,融合 Dice 损失和 Focal 损失。Focal 损失 $L_{\text{FL}}$、Dice 损失 $L_{\text{Dice}(p,g)}$、总损失 $L_{\text{total}}$的计算式为

$$L_{\text{FL}}(p_t)=-\alpha_t(1-p_t)^{\gamma}\lg p_t \tag{5}$$

$$L_{\text{Dice}(p,g)}=1-\frac{2\sum_{i=1}^{N}p_ig_i}{\sum_{i=1}^{N}p_i+\sum_{i=1}^{N}g_i} \tag{6}$$

$$L_{\text{total}}=\lambda_1L_{\text{CE}}+\lambda_2L_{\text{FL}}+\lambda_3L_{\text{Dice}} \tag{7}$$

式中 $p_t$——模型对真实类别的预测概率

$\alpha_t$——调节正负样本重要性的平衡因子

$\gamma$——聚焦系数,用于调整难易分类样本的权重

$p_i$——第 $i$ 像素预测为目标类别的概率

$g_i$——第 $i$ 像素真实标签

$L_{\text{CE}}$——交叉熵损失函数

$\sum_{i=1}^{N}p_ig_i$——预测与真实标签重叠区域

$\sum_{i=1}^{N}p_i+\sum_{i=1}^{N}g_i$——预测与真实标签的总区域

$\lambda_1$、$\lambda_2$、$\lambda_3$——权重

## 2 结果与讨论

### 2.1 试验环境

本研究的模型试验均在 64 位 Windows 10 操作系统下进行,内存为 64 GB,显卡为 NVIDIA GeForce RTX 3090,搭载 Intel(R) Core(TM) i9 - 10980XE CPU 处理器。计算机采用深度学习框架 Pytorch 1.10.0 版本,安装 CUDA 12.6 版本进行并行计算,Python 版本为 3.8。

### 2.2 评价指标

评价语义分割模型的效果,从性能、推理速度和复杂度等多方面衡量。性能指标采用平均交占比(Mean intersection over union,mIoU)、平均像素准确率(Mean pixel accuracy,mPA);推理速度用帧速(Frames per second,FPS)来衡量;复杂度包括浮点运算量、参数量(Parameter count)和内存占用量(Model size),评价模型在嵌入式设备部署的适用性。

### 2.3 不同注意力机制对比

为了使得模型充分利用粮厢颜色、边缘等低级特征,提升对目标区域的关注度,在浅层特征分支输出添加不同的注意力模块进行对比。表 2 展示了分别将 CA、ECA、SE、SAM、CBAM 注意力机制引入标准 DeepLabv3 + 时粮厢分割性能。

**表 2 不同注意力机制性能对比**

**Tab. 2 Comparison of different attention mechanisms**

%

| 添加注意力模块 | mIoU | mPA |
|---|---|---|
| CA | 95.15 | 97.81 |
| ECA | 95.18 | 98.00 |
| SE | 95.39 | 98.08 |
| SAM | 95.24 | 98.02 |
| CBAM | 95.25 | 97.88 |

由表 2 可得,加入 CA、ECA、SE、SAM、CBAM 5 种注意力机制,mIoU 相较于标准 DeepLabv3 + 模型分别提高 0.05、0.05、0.26、0.11、0.12 个百分点,mPA 分别提高 0.04、0.23、0.31、0.25、0.11 个百分点。相较而言,得益于 SE 注意力模块在捕捉通道信息方面的优势,强化了粮厢边缘、颜色特征表达,从而提高了 DeepLabv3 + 模型分割效果,mIoU 和 mPA 分别达到 95.39% 和 98.08%。

### 2.4 不同模型性能对比

为了评价 GTSM 性能,采用 UNet、DeepLabv3 +、PSPNet 主流语义分割模型,在自建运粮车数据集上采用相同配置参数训练,均添加 Dice - Focal 损失函数。各模型训练过程损失曲线如图 9 所示,损失值随着训练轮次的增加而减小,并趋于收敛,没有发生过拟合现象。在验证集上对比不同模型的 mIoU、mPA、参数量、内存占用量和浮点运算量,结果如表 3 和图 10 所示,图 10 中图像 1 ~ 5 分别对应情况:粮厢有稻谷的标准对靶视角、粮厢无稻谷的标准对靶视角、卸粮臂正在出粮、卸粮臂复杂转角和远距离小目标粮厢。

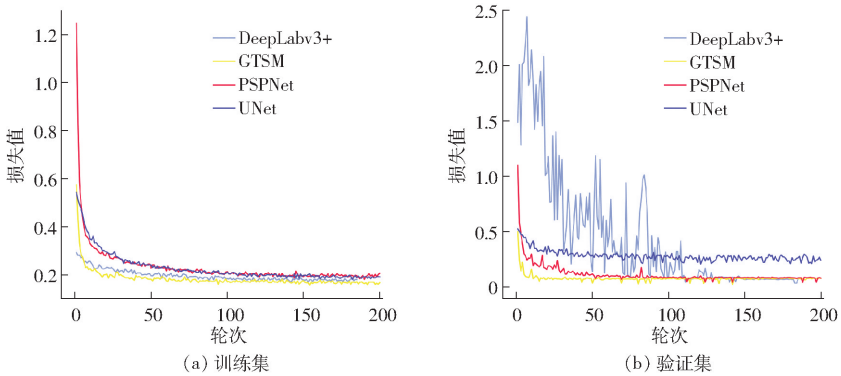



图 9　模型训练损失曲线

Fig. 9　Training loss

**表 3　不同模型性能对比**

**Tab. 3　Comparison of performance of different models**

| 模型 | 主干网络 | 平均交占比/% | 平均准确率/% | 参数量 | 内存占用量/MB | 浮点运算量 | 推理速度/(f·s$^{-1}$) |
|---|---|---|---|---|---|---|---|
| UNet | VGG | 93.39 | 97.51 | $2.489\times10^{7}$ | 94.95 | $2.258\,4\times10^{11}$ | 33.75 |
| DeepLabv3 + | Xception | 95.29 | 98.02 | $5.490\times10^{7}$ | 209.42 | $8.342\,0\times10^{11}$ | 16.14 |
| PSPNet | ResNet | 94.06 | 97.59 | $4.913\times10^{7}$ | 187.40 | $1.944\,0\times10^{11}$ | 33.73 |
| GTSM | ShuffleNetv2 | 96.07 | 98.69 | $5.620\times10^{6}$ | 21.43 | $2.404\,0\times10^{10}$ | 42.93 |

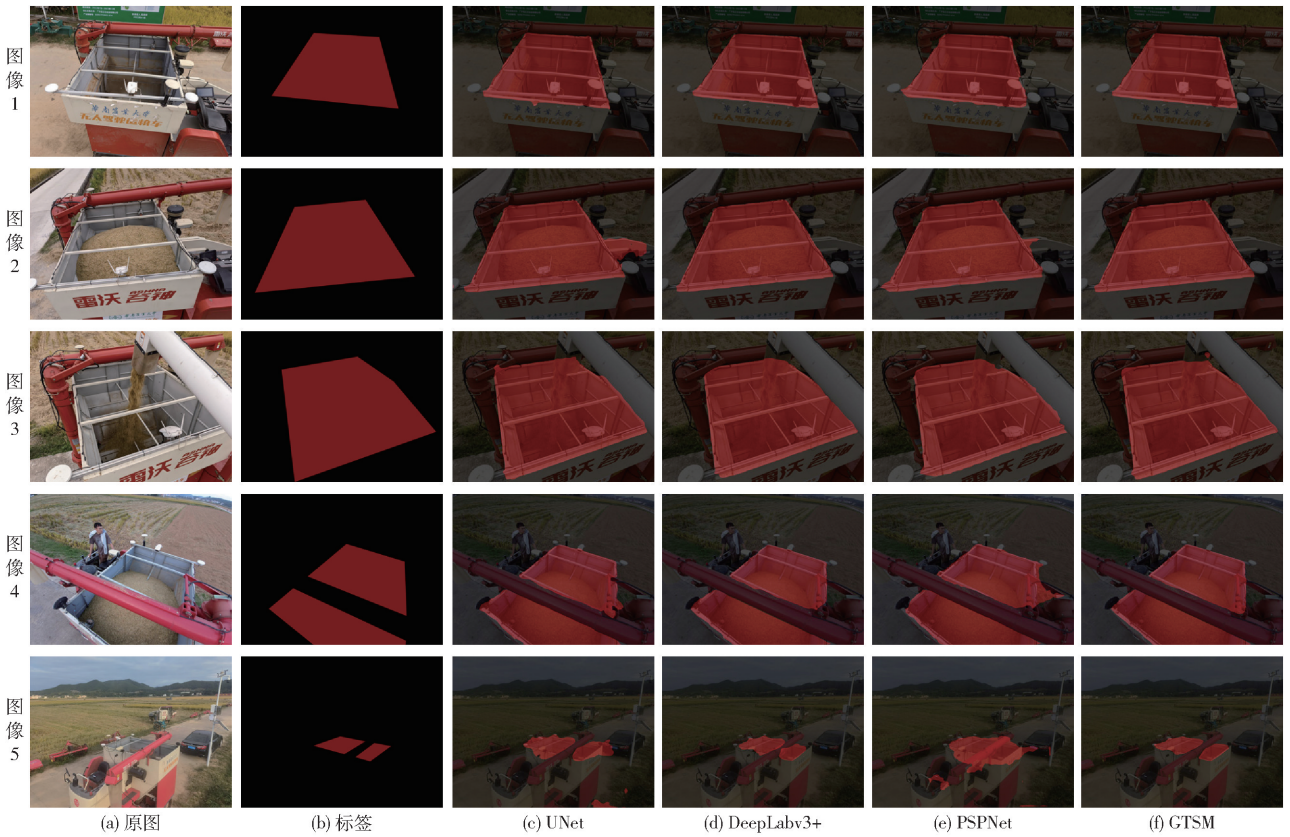



图 10　不同模型分割结果

Fig. 10　Segmentation results of different models

由表 3 可见，GTSM 性能指标表现较好，mIoU 达到 96.07%，mPA 为 98.69%，较基准模型 DeepLabv3 + 分别提高 0.78、0.67 个百分点，较 UNet 模型分别提高 2.68、1.18 个百分点，较 PSPNet 模型分别提高 2.01、1.1 个百分点。此外，GTSM 参数量仅 $5.62\times10^{6}$，内存占用量 21.43 MB，规模远小于其他模型，浮点运算量和推理速度分别为 $2.404\,0\times10^{10}$ 和 42.93 f/s，相比其他模型复杂度更低，推理速度更快，更适合部署到资源受限的移动设备，提供准确、高效的卸粮目标区域实时检测。

观察图像 1,4 种模型分割效果相近,可以很好分割出粮厢区域;由于图像 2 粮厢中有稻谷,PSPNet 和 UNet 对于粮厢内外稻谷区分能力不足,导致错误识别粮厢外水稻为目标区域;图像 3 中除了 PSPNet,其他模型对于粮厢轮廓分割效果较好,但对白色卸粮臂分割效果稍欠;图像 4 中除了 GTSM,其余模型对于卸粮臂复杂转角分割效果不佳,PSPNet 仍错误划分粮厢外稻杆为目标区域;图像 5 的粮厢是远距离小目标,各模型识别能力不足。综合来看,GTSM 分割效果较好,与基准网络 DeepLabv3 + 相近,说明其在模型参数量大幅减少的情况下仍保持较好分割效果,另外,得益于浅层网络分支引入的通道注意力机制,使模型更适合粮厢轮廓分割,得到平滑分割曲线结果,而对于图像 4 复杂转角和图像 5 等小样本视角,后续可以通过扩充数据集的方式提升模型泛化性能。

### 2.5 消融试验

为了验证提出的改进方法对于粮厢分割任务的有效性,在自建数据集进行消融试验,并与基准网络 DeepLabv3 + 对比。消融试验结果如表 4 所示,各模型训练参数一致,均添加 Dice - Focal 损失函数。

**表 4 消融试验结果**

**Tab. 4 Ablation experiments**

| 试验序号 | Backbone | SE | ASPP | Dw - ASPP | Micro - ASPP | mIoU/% | mPA/% | 参数量 | 内存占用量/MB | 浮点运算量 | 推理速度/(f·s$^{-1}$) |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | Xception | - | √ | - | - | 95.29 | 98.02 | $5.49\times10^{7}$ | 209.42 | $8.342\times10^{10}$ | 16.22 |
| 2 | ShufflNetv2 | - | √ | - | - | 95.95 | 98.63 | $1.04\times10^{7}$ | 39.62 | $2.893\times10^{10}$ | 45.79 |
| 3 | ShufflNetv2 | - | - | √ | - | 95.80 | 98.53 | $6.08\times10^{6}$ | 23.20 | $2.452\times10^{10}$ | 44.38 |
| 4 | ShufflNetv2 | - | - | - | √ | 95.96 | 98.61 | $5.62\times10^{6}$ | 21.43 | $2.404\times10^{10}$ | 42.36 |
| 5 | ShufflNetv2 | √ | √ | - | - | 96.10 | 98.64 | $1.04\times10^{7}$ | 39.62 | $2.893\times10^{10}$ | 42.37 |
| 6 | ShufflNetv2 | √ | - | √ | - | 95.94 | 98.59 | $6.08\times10^{6}$ | 23.20 | $2.452\times10^{10}$ | 43.82 |
| 7 | ShufflNetv2 | √ | - | - | √ | 96.07 | 98.69 | $5.62\times10^{6}$ | 21.43 | $2.404\times10^{10}$ | 42.93 |

注:√表示引入此模块;-表示没有引入此模块;Dw - ASPP 表示使用深度可分离卷积改进 ASPP 模块;Micro - ASPP 表示使用微因子分解卷积改进 ASPP 模块。

表 4 显示,DeepLabv3 + 的特征提取主干由 Xception 替换为轻量化模块 ShuffleNetv2 后,虽然参数量减小 81%,但是模型分割精度没有降低,反而有小幅提高,得益于 ShuffleNetv2 以运行速度为直接优化目标,使得模型复杂度大大降低,计算量减少 65%,推理速度提高 29.57 f/s,达到 45.79 f/s。对比试验 2、5,试验 3、6,以及试验 4、7,SE 注意力机制的加入一定程度上提升了模型分割性能,但没有增加较多的参数,而且对计算复杂度和推理速度没有影响,表明 SE 模块能够很好地捕捉和利用低级特征通道间的相关信息,从而改善分割结果,并保持模型的轻量化。对比试验 2、3、4 和试验 5、6、7,采用深度可分离卷积和微因子分解卷积改进 ASPP 模型后,模型参数量、计算量以及内存占用量逐步减小,但性能指标变化不大,表明深度可分离卷积和微因子分解卷积在不影响 ASPP 模块利用深层特征的情况下,均能够有效降低模型复杂度。对比试验 1、7,改进模型不仅便于在轻量化的嵌入式设备部署,还能够很好地完成粮厢实时检测任务。

## 3 结论

(1)围绕无人驾驶收获机向运粮车转卸稻谷时,卸粮臂对运粮车粮厢位置精准定位的需求,提出一种 Grain tank segmentation model(GTSM),对运粮车粮厢图像分割。

(2)以运粮车粮厢快速分割为研究目标,针对现有模型复杂度高、推理速度慢的问题,提出轻量化 GCSM 结构。该模型参数量 $5.62\times10^{6}$,内存占用量 21.43 MB,仅为基准网络的 1/10,浮点运算量为 $2.404\times10^{10}$,平均推理速度达 42.93 f/s,满足轻量化部署和实时性检测需求。

(3)为提高粮厢分割精度,采用 ShuffleNetv2 作为轻量化特征提取主干,在主干浅层特征输出引入 SE 通道注意力,并采用微因子分解卷积对 ASPP 模块低秩分解,进一步降低模型参数量。试验证明,GTSM 性能均高于 UNet、PSPNet 和 DeepLabv3 +,粮厢分割平均交占比达到 96.07%,平均像素精度达到 98.69%,满足田间运粮车粮厢自动化分割需求。

## 参 考 文 献

[1] 张闻宇,张智刚,张帆,等. 水稻收获转运双机协同自主作业策略与试验[J]. 农业工程学报,2022,38(15):1-9.
ZHANG Wenyu, ZHANG Zhigang, ZHANG Fan, et al. Cooperative autonomous operation strategy and experiment of the rice

harvester together with a rice-transporting vehicle[J]. Transactions of the CSAE,2022,38(15):1 - 9. (in Chinese)
[2] 满忠贤,何杰,刘善琪,等. 智能农机多机协同收获作业控制方法与试验[J]. 农业工程学报,2024,40(1):17 - 26.
MAN Zhongxian,HE Jie,LIU Shanqi,et al. Method and test for operating multi-machine cooperative harvesting in intelligent agricultural machinery[J]. Transactions of the CSAE,2024,40(1):17 - 26. (in Chinese)
[3] DING Y J,YANG W T,ZHANG J J,et al. An improved DeepLabV3 + based approach for disease spot segmentation on apple leaves[J]. Computers and Electronics in Agriculture,2025,231:110041.
[4] ZHANG L,MING L,ZHU X H,et al. Navigation path recognition between rows of fruit trees based on semantic segmentation [J]. Computers and Electronics in Agriculture,2024,216:108511.
[5] 李振东,李云飞,杨立伟,等. 基于改进 UNet 的樱桃树枝直径计算方法[J]. 农业机械学报,2024,55(增刊 1):263 - 269.
LI Zhendong,LI Yunfei,YANG Liwei,et al. Calculation method for cherry branch diameter based on improved UNet[J]. Transactions of the Chinese Society for Agricultural Machinery,2024,55(Supp. 1):263 - 269. (in Chinese)
[6] 杨宁,程巍,张钊源,等. 基于 FPGA 加速的 Mask R - CNN 稻瘟病高通量自适应识别模型研究[J]. 农业机械学报,2024,55(7):298 - 304.
YANG Ning,CHENG Wei,ZHANG Zhaoyuan,et al. Mask R - CNN based on-chip recognition algorithm for rice blast image [J]. Transactions of the Chinese Society for Agricultural Machinery,2024,55(7):298 - 304. (in Chinese)
[7] 顾文娟,魏金,阴艳超,等. 基于改进 DeepLabv3 + 的番茄图像多类别分割方法[J]. 农业机械学报,2023,54(12):261 - 271.
GU Wenjuan,WEI Jin,YIN Yanchao,et al. Multi-category segmentation method of tomato image based on improved DeepLabv3 + [J]. Transactions of the Chinese Society for Agricultural Machinery,2023,54(12):261 - 271. (in Chinese)
[8] 侯文慧,周传起,程炎,等. 基于轻量化 U - Net 网络的果园垄间路径识别方法[J]. 农业机械学报,2024,55(2):16 - 27.
HOU Wenhui,ZHOU Chuanqi,CHENG Yan, et al. Path recognition method of orchard ridges based on lightweight U - Net[J]. Transactions of the Chinese Society for Agricultural Machinery,2024,55(2):16 - 27. (in Chinese)
[9] LIU R,WANG D,CHEN Z,et al. Recognition of tea sprouts based on improved DeepLabV3 + model[J/OL]. Multimed Tools Appl. ,2025. https://doi.org/10.1007/s11042 - 024 - 20570 - 8.
[10] SUN J W,ZHOU J,HE Y Q,et al. RL - DeepLabv3 + :a lightweight rice lodging semantic segmentation model for unmanned rice harvester[J]. Computers and Electronics in Agriculture,2023,209:107823.
[11] 曹英丽,赵雨薇,杨璐璐,等. 基于改进 DeepLabv3 + 的水稻田间杂草识别方法[J]. 农业机械学报,2023,54(12):242 - 252.
CAO Yingli,ZHAO Yuwei,YANG Lulu,et al. Weed identification method in rice fielf based on improved DeepLabv3 + [J]. Transactions of the Chinese Society for Agricultural Machinery,2023,54(12):242 - 252. (in Chinese)
[12] CHEN L C,ZHU Y K,PAPANDREOU G,et al. Encoder-decoder with atrous separable convolution for semantic image segmentation[J]. arXiv Preprint,arXiv:1802.02611v3,2018.
[13] CHOLLET F. Xception: deep learning with depthwise separable convolutions[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR),2017:1251 - 1258.
[14] YU F,KOLTUN V. Multi-scale contnxt aggregation by dilated convolutions[J]. arXiv Preprint,arXiv:1511.07122v3,2015.
[15] MA N N,ZHANG X Y,ZHENG H T,et al. ShuffleNet V2: practical guidelines for efficient CNN architecture design[J]. arXiv Preprint,arXiv:1807.11164,2018.
[16] LI Y S,CHEN Y P,DAI X Y,et al. MicroNet: towards image recognition with extremely low FLOP[J]. arXiv Preprint,arXiv: 2011.12289v1,2020.
[17] 赵晋陵,詹媛媛,王娟,等. 基于 SE - UNet 的冬小麦种植区域提取方法[J]. 农业机械学报,2022,53(9):189 - 196.
ZHAO Jinling,ZHAN Yuanyuan,WANG Juan,et al. SE - UNet-based extraction of winter wheat planting areas[J]. Transactions of the Chinese Society for Agricultural Machinery,2022,53(9):189 - 196. (in Chinese)
[18] 武锦龙,吴虹麒,李浩,等. 基于改进 DeepLabV3 + 的荞麦苗期无人机遥感图像分割识别方法研究[J]. 农业机械学报,2024,55(5):186 - 195.
WU Jinlong,WU Hongqi,LI Hao,et al. Segmentation of buckwheat by UAV based on improved lightweight DeepLabv3 + at seedling stage[J]. Transactions of the Chinese Society for Agricultural Machinery,2024,55(5):186 - 195. (in Chinese)
[19] 袁培森,欧阳柳江,翟肇裕,等. 基于 MobileNetV3Small - ECA 的水稻病害轻量级识别研究[J]. 农业机械学报,2024,55(1):253 - 262.
YUAN Peisen,OUYANG Liujiang,ZHAI Zhaoyu,et al. Lightweight identification of rice diseases based on improved ECA and MobileNetV3Small[J]. Transactions of the Chinese Society for Agricultural Machinery,2024,55(1):253 - 262. (in Chinese)
[20] 袁杰,谢霖伟,郭旭,等. 基于改进 YOLO v7 的苹果叶片病害检测方法[J]. 农业机械学报,2024,55(11):68 - 74.
YUAN Jie,XIE Linwei,GUO Xu,et al. Apple leaf disease detection method based on improved YOLO v7[J]. Transactions of the Chinese Society for Agricultural Machinery,2024,55(11):68 - 74. (in Chinese)
[21] 吕继东,翟志强,孟庆建,等. 基于车载成像与深度卷积神经网络的地表残膜识别方法[J]. 农业机械学报,2025,56(5):26 - 37.
LÜ Jidong,ZHAI Zhiqiang,MENG Qingjian,et al. Surface residual film recognition method based on vehicle-mounted imaging and deep convolutional neural networks[J]. Transactions of the Chinese Society for Agricultural Machinery,2025,56(5):26 - 37. (in Chinese)

[22] 王发明,倪昕东,张旗,等. 基于 MobileNetV2 - CBAM 的机收场景下冬小麦成熟期在线分类识别方法[J]. 农业机械学报,2024,55(增刊 1):71 - 80,100.
WANG Faming,NI Xindong,ZHANG Qi,et al. Online classification and identification method of winter wheat maturity under mechanical harvesting scenario based on MobileNetV2 - CBA[J]. Transactions of the Chinese Society for Agricultural Machinery,2024,55(Supp. 1):71 - 80,100. (in Chinese)
[23] 李凯雨,朱昕怡,马浚诚,等. 基于混合扩张卷积和注意力的黄瓜病害严重度估算方法[J]. 农业机械学报,2023,54(2):231 - 239.
LI Kaiyu,ZHU Xinyi,MA Juncheng,et al. Estimation method of leaf disease severity of cucumber based on mixed dilated convolution and attention mechanism[J]. Transactions of the Chinese Society for Agricultural Machinery,2023,54(2):231 - 239. (in Chinese)
[24] 奚小波,丁杰源,翁小祥,等. 基于轻量化 YOLO v5s - MCA 的番茄成熟度检测方法[J]. 农业机械学报,2025,56(3):383 - 391,436.
XI Xiaobo,DING Jieyuan,WENG Xiaoxiang,et al. Tomato maturity detection method based on lightweight YOLO v5s - MCA[J]. Transactions of the Chinese Society for Agricultural Machinery,2025,56(3):383 - 391,436. (in Chinese)
[25] 李尚平,卞俊析,李凯华,等. 基于改进 YOLO v5s 的复杂环境下蔗梢分叉点识别与定位[J]. 农业机械学报,2023,54(11):247 - 258.
LI Shangping,BIAN Junxi,LI Kaihua,et al. Identification and height localization of sugarcane tip bifurcation points in complex environments based on improved YOLO v5s[J]. Transactions of the Chinese Society for Agricultural Machinery,2023,54(11):247 - 258. (in Chinese)
[26] HU J,SHEN L,SUN G,et al. Squeeze-and-excitation network[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR),2018:7132 - 7141.

(上接第 176 页)

[24] LIAN B,PENG J,ZHAN H,et al. Mechanical response of root-reinforced loess with various water contents[J]. Soil and Tillage Research,2019,193: 85 - 94.
[25] SONGÜL G. Soil compaction due to increased machinery intensity in agricultural production: its main causes, effects and management[M]// FIAZ A,MUHAMMAD S. Technology in Agriculture. Rijeka: IntechOpen, 2021.
[26] NIE Y,NI W,LÜ X,et al. Macroscopic mechanical behavior and microstructural evolution of compacted loess in the Chinese Loess Plateau[J]. Soil and Tillage Research,2023,232: 105767.
[27] DA SILVA E E,BAIO F H R,KOLLING D F,et al. Variable-rate in corn sowing for maximizing grain yield[J]. Sci. Rep.,2021,11(1): 10.
[28] 穆璇,吴玉琴. 增湿对斥水土壤压实性的影响研究[J]. 水利水电技术(中英文),2021,52(10): 199 - 208.
MU Xuan,WU Yuqin. Effect of humidification on the compactness of water-repellent soil[J]. Water Resources and Hydropower Engineering,2021,52(10): 199 - 208. (in Chinese)
[29] LAMICHHANE J R,BOIFFIN J,BOIZARD H,et al. Seedbed structure of major field crops as affected by cropping systems and climate: results of a 15-year field trial[J]. Soil and Tillage Research,2021,206: 104845.
[30] 熊乾,高磊,彭新华,等. 黑土区水蚀坡耕地土壤穿透阻力时空变异特征及传递函数[J]. 水土保持学报,2024,38(4): 162 - 169,180.
XIONG Qian,GAO Lei,PENG Xinhua,et al. Spatio-temporal characteristics and pedotransfer of soil penetration resistance in eroded sloping farmland of black soil region[J]. Journal of Soil and Water Conservation,2024,38(4): 162 - 169,180. (in Chinese)
[31] QIU T,LIU C,ZHONG X,et al. Experimental research on the impact of temperature on the adhesion characteristics of soil-structure interface[J]. Geofluids,2020,2020: 6675576.
[32] CHEN X,CRUSE R M,NIU S,et al. Effect of loading time on soil structural failure[J]. Soil and Tillage Research,2019,186: 87 - 93.