

基于 RCH-UNet 的新疆密植棉花图像快速分割及产量预测

刘祥, 田敏*, 梁金艳

(石河子大学机械电气工程学院, 石河子 832003)

摘要: 针对复杂棉田环境下传统图像分割技术存在分割精度低、实时性弱和鲁棒性差等问题, 该研究以脱叶期新疆密植棉花为对象, 结合低空无人机遥感平台, 提出一种 RCH-UNet (resnet coordinate hardswish UNet) 棉花产量快速预测模型。将 UNet 中传统的 CBR (convolution batch normalization ReLU) 下采样模块替换为 ResNet50, 同时将 CA (coordinate attention) 注意力机制和 hardswish 激活函数引入 UNet, 以提高图像特征的提取能力, 增强图像分割效果。基于无人机采集的棉花图像数据集评估 RCH-UNet 模型性能。试验结果表明, 在该研究构建的棉花图像数据集上, RCH-UNet 模型的棉花分割交并比达到 92.79%, 像素准确率达到 96.22%, 精确率为 96.30%, 与原始 U-Net、PSPNet 和 DeepLabv3 相比, 像素准确率分别提高了 9.85、17.67、6.31 个百分点。通过 RCH-UNet 提取棉花像素比和灰度共生矩阵提取纹理特征, 结合岭回归分析构建多因素棉花产量预测模型, 模型的 R^2 为 0.92, 预测产量与实际产量平均绝对百分比误差为 9.254%。研究结果可为新疆密植棉花产量预测提供技术支持。

关键词: 深度学习; 图像处理; 语义分割; UNet; 自注意力机制; 密植棉花; 产量预测

doi: 10.11975/j.issn.1002-6819.202310082

中图分类号: S126; TP79

文献标志码: A

文章编号: 1002-6819(2024)-07-0230-10

刘祥, 田敏, 梁金艳. 基于 RCH-UNet 的新疆密植棉花图像快速分割及产量预测[J]. 农业工程学报, 2024, 40(7): 230-239. doi: 10.11975/j.issn.1002-6819.202310082 <http://www.tcsae.org>

LIU Xiang, TIAN Min, LIANG Jinyan. Prediction of cotton yield densely planted in Xinjiang of China using RCH-UNet model[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2024, 40(7): 230-239. (in Chinese with English abstract) doi: 10.11975/j.issn.1002-6819.202310082 <http://www.tcsae.org>

0 引言

棉花具有重要经济价值^[1-2]。新疆是中国重要的棉花种植基地之一, 在全球棉花产业中具有举足轻重的地位^[3-4]。据 2022 年国家统计局数据显示, 中国的棉花总产量为 598.02 万吨, 其中 90% 的棉花产量来自新疆地区^[5]。因此, 快速而准确地在新疆棉田中进行棉花产量预测对于新疆棉花种植管理、棉花贸易和政策制定都具有重要意义。

目前, 为了增加棉花产量, 新疆棉花普遍采用窄行和宽行交替分布的“矮、早、密”的密集种植模式^[6-7]。该模式在防止自然灾害和提高产量方面具有明显优势, 但是窄行的棉花植株相互交错并且生长茂密, 给棉花图像分割和棉花产量预测工作带来了诸多挑战^[8-10]。“智慧农业”的兴起推动了深度学习在农业领域的广泛应用和无人机遥感平台的发展。深度学习凭借其强大的特征学习能力能准确地给出生产决策从而被广泛应用于作物病虫害识别和产量预测。SAEED 等^[11]提出使用 3D 分割方法进行棉铃分割, 该方法可有效地解决部分棉铃遮挡问

题, 然而该方法数据处理较为繁琐且训练周期长, 无法实现棉花产量快速预测。LI 等^[12]提出一种基于区域的语义分割方法, 使用无监督区域生成、特征提取和随机森林进行棉花吐絮完成度检测。该方法仅针对绿色棉铃进行了检测, 未能扩展到成熟棉花的检测与识别; 同时区域的生成参数选择较为简单, 不同环境下模型的鲁棒性仍然有待提高。SINGH 等^[13]提出 3 个轻量级全卷积神经网络模型, 用于棉铃语义分割, 以便部署到棉花收割机器人上进行实时识别。但其使用的数据集仅包含 100 张图像, 样本量较小, 图像背景也较为简单。考虑到新疆棉田植株密度大、光照条件复杂等因素, 模型在新疆密植棉花图像上的泛化能力还有待验证。LV 等^[14]将 PSPNet 模型结构进行优化, 在模型中加入上下文编码模块, 补充更多的全局先验知识, 并在模型解码部分整合了棉花图像浅层特征。但模型仍存在误判和分割结果不完整的问题, 预测精度仍需进一步提升。无人机凭借其灵活性和低空飞行能力被广泛应用于作物产量预测。戴建国等^[15]提出利用无人机遥感平台收集多时相遥感数据预测棉花产量, 从而得到棉花产量预测结果。然而, 该方法需要收集多时段的棉花遥感数据, 数据获取繁琐且模型训练周期较长。FENG 等^[16]从配备有可见光相机无人机获取的棉田图像中提取了 8 种图像特征, 探究了不同图像特征对于产量预测模型的影响, 由于数据集背景简单且棉花种植模式相对稀疏, 因此并不适用于新疆密植棉花产量预测工作。MA 等^[17]提出从无人机遥感影像

收稿日期: 2023-10-13 修订日期: 2023-11-27

基金项目: 科技创新 2030—“新一代人工智能”重大项目 (2022ZD0115803); 国家自然科学基金项目 (61962053)

作者简介: 刘祥, 研究方向为计算机视觉技术农业领域应用。

Email: lx@stu.shzu.edu.cn

*通信作者: 田敏, 教授, 博士生导师, 研究方向为农业信息化。

Email: tm_inf@shzu.edu.cn

中提取植被指数和纹理特征进行棉花产量预测工作，但是在农田中棉花往往与其他植被混杂，同时遥感图像中可能存在地物遮挡，导致提取植被指数存在一定的局限性。DUBE 等^[18]提出使用检测棉铃数量进行棉花产量预测，该方法在美国棉花产量预测中具有较好的效果，然而在密植的新疆棉花中表现仍然欠佳。

为实现新疆密植棉花产量快速预测，本文提出了一种新的新疆密植棉花产量预测方法。该方法基于无人机遥感影像，使用基于深度学习的语义分割技术将棉花从 RGB (red green blue) 图像中分割同时使用灰度共生矩阵 (gray-level co-occurrence matrix, GLCM) 提取图像纹理特征，最终借助岭回归分析模型构建棉花产量预测模型。本文的主要内容如下：1) 借助无人机遥感平台构建一种具有复杂背景的航拍棉花图像数据集。2) 对 UNet 模型进行优化，将 ResNet50 作为 UNet 的主干网络，同时在模型中加入 CA (coordinate attention) 注意力机制和 hardswish 激活函数构建 RCH-UNet 棉花图像分割模型。3) 对改进后的 UNet 模型进行测试和验证，并将改进后的 UNet 模型与其他经典的语义分割算法进行性能比较。4) 通过岭回归分析寻找棉花图像像素比、棉花图像纹理特征和棉花产量之间的关系，构建棉花产量预测模型。

1 材料与方法

本文主要技术流程如下，首先在棉花收获前 7 d 使用无人机进行棉花图像数据采集，并人工采集样方中的棉花，回到实验室通过称量记录样方产量，在收获当天记录机采棉收获产量以及人工复采的产量。然后通过改进的 RCH-UNet 模型进行棉花图像分割得到相应图片的像素比，使用 GLCM 提取棉花图像纹理特征。最后通过像素比和纹理特征与记录的样方产量进行多次岭回归分析从而得到棉花产量预测模型，之后采用 5 个大田棉花地块进行产量预测模型评估。

1.1 数据采集

1.1.1 无人机遥感影像获取与预处理

本文以脱叶期新疆密植棉花为研究对象，由于目前没有优质的公开新疆密植棉花图像和相关产量数据，因此在中国新疆沙湾县朱家庄农场 (44°51'N, 85°99'E) 采集棉花图像，研究区域位置如图 1 所示。数据采集时间为 2022 年 10 月 11 日—10 月 18 日，无人机飞行高度为 10 m。在棉田中通过五点采样方法随机选取 5 个样方点进行数据采集，每个样方大小为 1 m×1 m。数据采集设备为 DJI M200 工业级无人机，相机为 ZENMUSE X4S，有效分辨率为 2 000 万像素。

1.1.2 数据集制作

本研究借助无人机遥感平台获取棉花脱叶期遥感影像数据样本，为了便于模型训练同时充分利用高分辨率图片所包含的信息，本文将分辨率为 5 472×3 078 像素的图片连续分割为多个大小为 480×480 像素的子图片，每个子图片拥有 10% 的重叠区域。

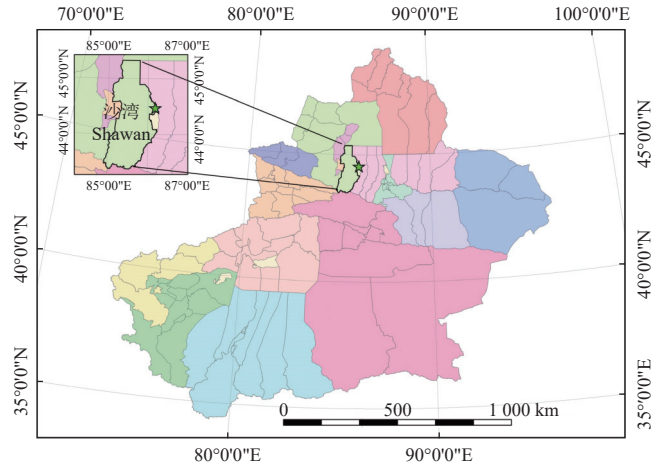


图 1 研究区域示意图

Fig.1 Schematic of study area

为增强模型的学习能力，对原始图像分别进行亮度增强、噪声模糊、翻转和旋转；同时将 RGB 颜色空间转换为 CMYK (cyan magenta yellow black) 颜色空间。通过上述数据增强方法，共得到 2 000 张样本数据，其样本增强和颜色空间转换效果如图 2 所示。最终将样本数据集按照 7:2:1 划分为训练集、验证集和测试集。

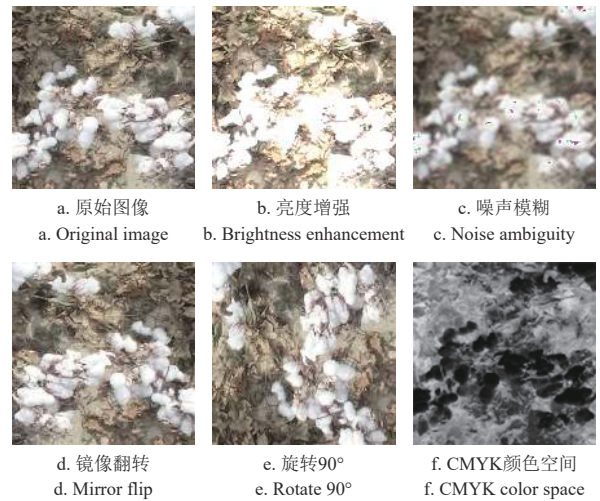


图 2 数据样本扩充

Fig.2 Expansion of data sample

1.2 RCH-UNet 模型构建

1.2.1 UNet 分割模型

UNet^[19]在农业领域被广泛应用，例如，孙俊等^[20]人在复杂环境中葡萄快速检测研究中使用 UNet 进行田间葡萄分割。朱立学等^[21]使用一种轻量化 UNet 模型进行香蕉图像分割。SAHIN 等^[22]使用 UNet 与植被指数相结合分割杂草。DIAO 等^[23]使用 UNet 解决针对玉米作物在杂草、断行、叶片粘附等复杂农田环境中难以识别中心线的问题。上述模型将 UNet 模型应用于各类农业领域中，然而场景大多是空间较为宽阔，目标较大且遮挡较少的环境。上述研究人员使用 UNet 在相关研究领域中取得了一定的成果，但是在复杂的新疆密植棉花场景中存在很多小目标棉花，UNet 模型很难精准分割。同

时密植棉田多变性大, UNet 对环境变化鲁棒性较差, 泛化性能不佳, 不同环境下性能落差明显。因此将 UNet 模型通过网络结构优化、增加感受野等方式提升模型鲁棒性和泛化性, 以适应新疆密植棉花图像分割以及后续的棉花产量预测工作。

原始 UNet 模型主要由主干特征提取网络和加强特征提取网络组成。通过主干特征提取部分获得 5 个初步有效特征层。加强特征提取部分利用该 5 个初步有效特征层进行上采样, 并且进行特征融合, 最终获得一个融合所有特征的有效特征层。UNet 结构如图 3 所示。

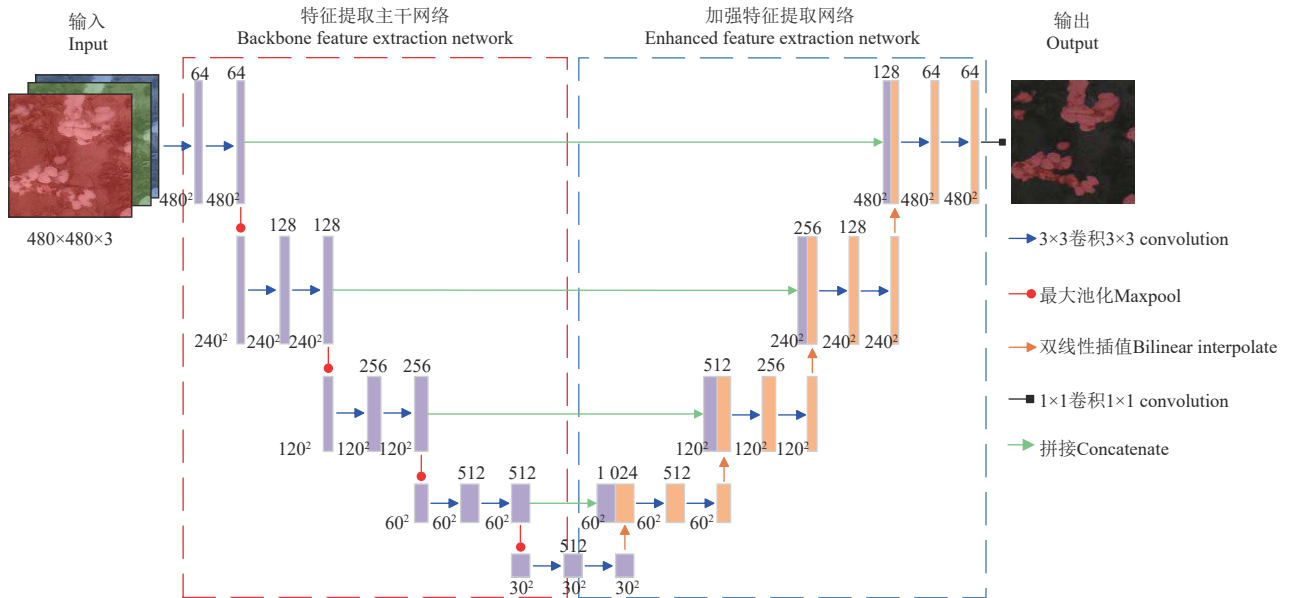


图 3 UNet 结构
Fig.3 UNet structure

1.2.2 ResNet50 网络结构

ResNet^[24] 因提出的残差结构具有强大的特征提取能力而被广泛应用于各类深度学习网络中。深度残差网络有效地克服了由网络深度加深而产生的学习效率变低与准确率无法有效提升的问题。ResNet 主要由卷积残差块 (conv block) 和跳跃残差块 (identity block) 构成, 其中卷积残差块用于改变网络维度, 结构如图 4a 所示; 跳跃残差块用于加深网络, 结构如图 4b 所示。

1.2.3 CA 注意力机制

SE^[25] 和 CBAM^[26] 注意力机制的提出证实了通道注意力机制具有提高深度学习网络性能和有效筛选特征的潜力。SE 和 CBAM 在求取通道注意力机制时, 通道的处理一般是采用全局最大池化或者是平均池化。然而, 池化操作会丢失物体的空间信息。因此, 本文采用通道注意力机制和空间注意力机制相结合的 CA 注意力机制^[27]。

CA 注意力机制作为一种轻量级注意力机制, 能够将横向和纵向的位置信息编码到通道注意力中, 在不增加计算量的基础上使得网络能够关注大范围的位置信息。CA 注意力机制结构如图 5 所示。

CA 注意力机制主要分为以下 3 个阶段:

- 1) 将输入特征图分别在为宽度和高度 2 个方向分别进行全局平均池化, 分别获得 2 个方向的特征图。
- 2) 将 2 个并行阶段合并, 将宽和高转置到同一个维度, 然后进行堆叠, 将宽高特征合并。
- 3) 再次分为 2 个并行阶段, 通过非线性转换将宽度和高度分为: $[C, 1, H]$ 和 $[C, 1, W]$, 之后进行转置, 获得 2 个特征层 $[C, H, 1]$ 和 $[C, 1, W]$ 。然后利用 1×1 卷积调

整通道数后使用 sigmoid 函数获得宽高维度上的注意力, 其中 C 、 H 和 W 分别表示通道数、图像高度和图像宽度。

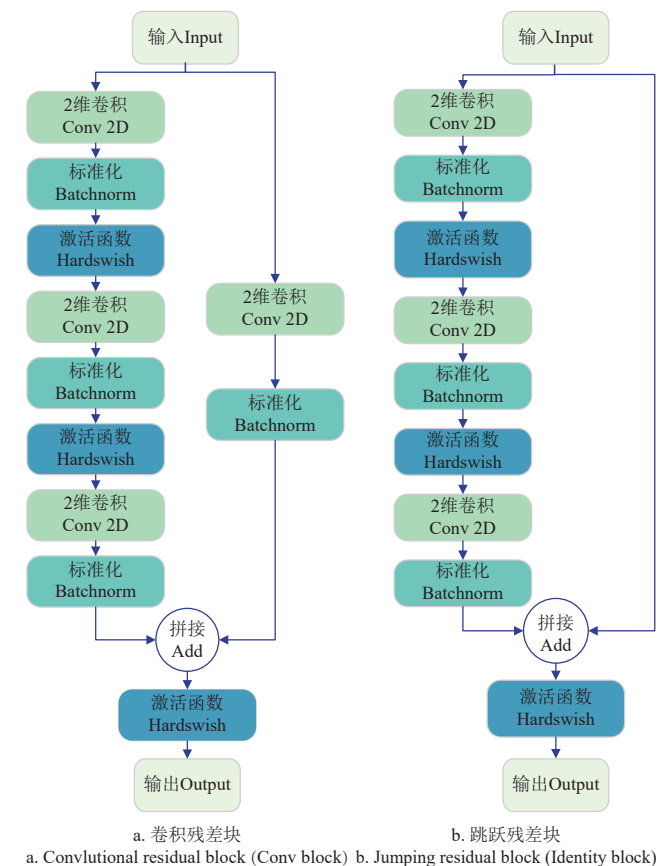


图 4 ResNet 结构
Fig.4 ResNet structure

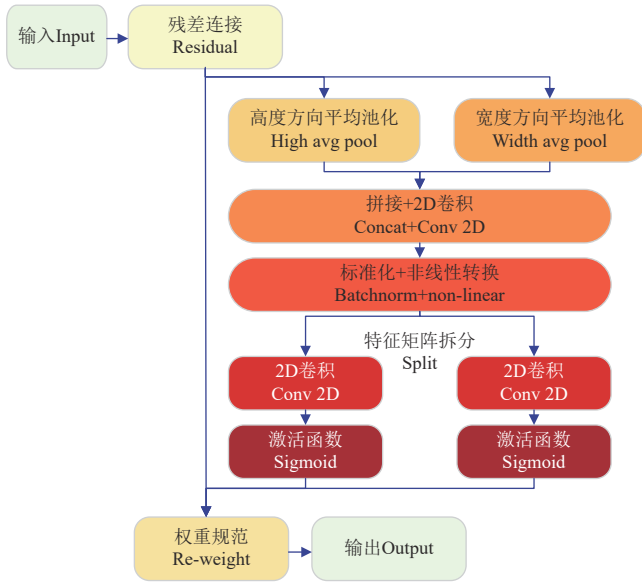


图 5 CA 注意力机制

Fig.5 Coordinate attention mechanism

1.2.4 hardswish 激活函数

深度网络中激活函数的选择对训练动态和任务性能有显著影响^[28]。ReLU 和 swish 激活函数被广泛应用于各类语义分割网络之中^[29]。当 ReLU 函数的输入值为负的时候，该函数输出始终为 0，其一阶导数也始终为 0 导致神经元不能更新参数，无法充分学习。swish 函数相较于 ReLU 函数显著地提高了网络的准确性，但是 swish 函数中的 sigmoid 函数是由指数构成的，会额外增加计算成本。因此，本研究将 hardswish 函数作为激活函数，如式 (1) 所示，式中 ReLU6 表示 ReLU 激活函数。

$$\text{hardswish}(x) = x \frac{\text{ReLU6}(x+3)}{6} \quad (1)$$

hardswish 函数相较于普通 swish 函数降低了由 sigmoid 产生的计算成本，同时引入 ReLU6 函数作为门限函数将输入值限制在 [-3, 3] 之间，通过截断和线性缩

放的操作使得 hardswish 函数更具有鲁棒性，对于模型的训练和推理过程中的梯度传播更有利。

1.2.5 RCH-UNet 分割模型

在大田环境下进行棉花图像分割存在诸多因素影响分割精度。例如，棉花种植密度较高，棉株之间、棉株与杂草之间均会产生遮挡。为了解决上述问题，本文提出一种改进的 UNet 分割模型，即 RCH-UNet。

RCH-UNet 结构与原始 UNet 结构相似，仍然采取编码-解码结构。RCH-UNet 将 UNet 模型作为基准模型，将图 3 中的主干特征提取网络替换为 ResNet50，优化后的特征提取主干网络如图 6 所示。本文在主干特征提取网络提取的 5 个特征层之后引入 CA 注意力机制增加模型的感受野，以提升模型分割精度。此外，RCH-UNet 模型中所有的激活函数被替换为 hardswish 函数，优化训练过程中模型的梯度传播。

1.3 分割模型评价指标与试验环境

1.3.1 评价指标

为了测试和评估模型的各项性能，基于混淆矩阵本文采用语义分割模型常用的交并比 (I_u)、像素准确率 (P_a) 和精确率 (P_e) 作为评价指标，计算方法如式 (2) ~ (4) 所示。

$$I_u = \frac{T_p}{T_p + F_p + F_n} + \frac{T_n}{T_n + F_n + F_p} \times 100\% \quad (2)$$

$$P_a = \frac{T_p + T_n}{T_p + T_n + F_p + F_n} \times 100\% \quad (3)$$

$$P_e = \frac{T_p}{T_p + F_p} \times 100\% \quad (4)$$

式中 T_p 为样本真实类别是棉花，同时模型正确识别； F_n 为样本真实类别是棉花，但是模型错误识别； F_p 为样本真实类别是背景，模型将其识别为棉花； T_n 为样本真实类别是背景，模型将其识别为棉花。

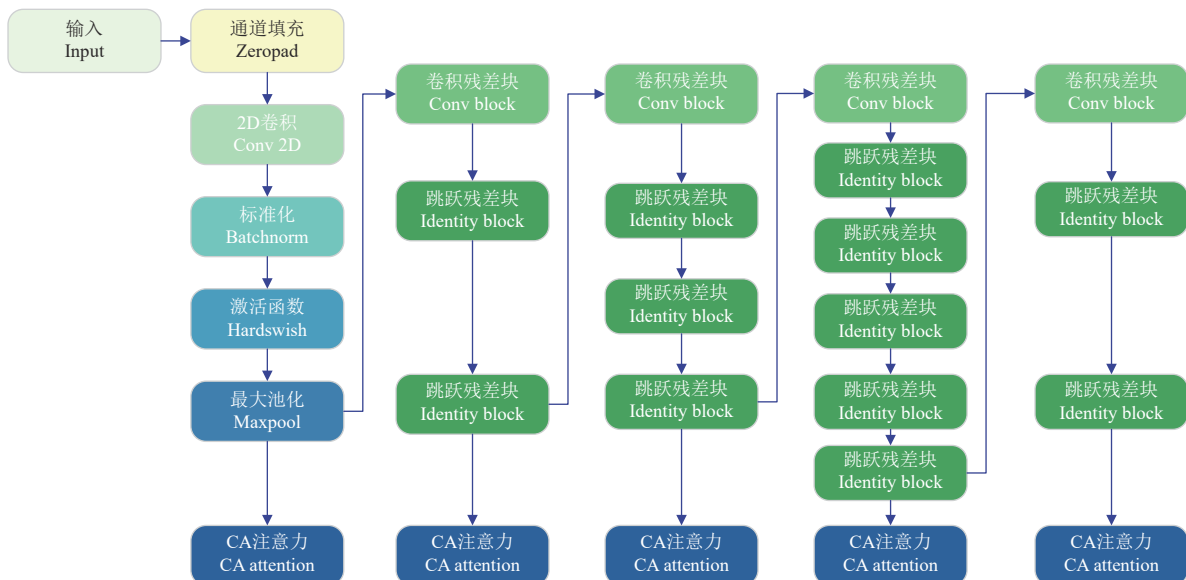


图 6 RCH-UNet 特征提取主干

Fig.6 Feature extraction backbone of RCH-UNet

1.3.2 试验环境

本研究模型训练和验证均在同一台设备上完成,设备采用 Linux 操作系统,中央处理器为 AMD EPYC 7T83 64-Core,显卡型号为 NVIDIA GeForce RTX 4090;编程语言版本为 Python 3.9.16;模型训练架构为 Torch 1.10.0,加速平台为 CUDA 11.3, CUDNN 8.4.1。模型训练中采用 adam 优化器防止权重衰减导致模型过拟合;初始学习率设置为 10^{-4} ,最小学习率设置为 10^{-6} ,学习率下降方式为余弦下降。训练轮次(epoch)设置为 100,其中冻结训练 epoch 为 50;模型冻结时批次大小(batch size)为 64,解冻训练时 batch size 为 32。

1.4 棉花产量预测模型与评价指标

1.4.1 棉花产量预测模型

本研究使用 RCH-UNet 模型计算棉花像素比,并将其作为产量预测工作中重要的特征值之一。棉花像素比计算方法如式(5)所示。

$$P_r = \frac{P_c}{P_t} \quad (5)$$

式中 P_r 为棉花像素比, P_c 为图像中棉花像素数量, P_t 为图像总像素数量。

此外,由于 GLCM 可以提取图像的对比度、能量、相关性、均匀性等特征,并且 GLCM 作为一种灰度不变的特征提取方法,它不受图像的亮度、色彩、光照等因素的影响,因此可以适应不同的拍摄条件和设备,可以有效地提高棉花产量预测模型的鲁棒性和准确性^[30-32]。对比度 (C_t)、能量 (E_y)、相关性 (C_o) 和均匀性 (A_m) 计算方法如式(6)~(9)所示。

$$C_t = \sum_i \sum_j (i, j)^2 P(i, j) \quad (6)$$

$$E_y = \sum_{i, j=0}^{N-1} P(i, j)^2 \quad (7)$$

$$C_o = \frac{\sum_i \sum_j (i, j) P(i, j) - \mu_x \mu_y}{\sigma_x \sigma_y} \quad (8)$$

$$A_m = \sum_i \sum_j P(i, j)^2 \quad (9)$$

式中 i 和 j 表示灰度级; $P(i, j)$ 表示灰度级 i 和 j 同时出现的概率; μ_x 和 μ_y 分别是灰度级 i 和 j 的均值; σ_x 和 σ_y 分别是灰度级 i 和 j 的标准差。

本研究在使用 RCH-UNet 提取棉花像素的同时还使用 GLCM 提取棉花图像中对比度、能量、相关性、均匀性 5 个纹理特征作为棉花产量预测模型的产量预测因子。最终通过岭回归分析寻找产量预测因子与产量之间的关系。

1.4.2 产量预测模型评价指标

在本研究中使用决定系数 (R^2)、相对误差 (R_e) 和平均绝对百分比误差 (M_e) 评估模型拟合度和模型预测精度;计算方法如式(10)~(12)所示。

$$R^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})^2}{n \sum_{i=1}^n (x_i - \bar{x})^2 n \sum_{i=1}^n (y_i - \bar{y})^2} \quad (10)$$

$$R_e = \frac{|y_i - x_i|}{x_i} \times 100\% \quad (11)$$

$$M_e = \frac{1}{n} \sum_{i=1}^n \left| \frac{y_i - x_i}{x_i} \right| \times 100\% \quad (12)$$

式中 x_i 表示第 i 个样方的棉花产量的测量值; y_i 表示第 i 个样方的棉花产量预测值; \bar{x} 表示实测棉花产量的平均值; \bar{y} 表示模型估算的棉花产量平均值; n 表示地块数量。

2 结果与分析

2.1 改进模块性能比较

2.1.1 不同注意力机制的比较试验

由于棉田环境较为复杂,UNet 会受环境影响从而导致分割精度下降。为了降低复杂背景对模型分割性能的影响,本文在主干中加入了不同的注意力机制进行比较。表 1 展示了分别将 SE、CBAM、ECA 和 CA 四种注意力机制引入 UNet 之后模型针对棉花类别的分割性能。

表 1 不同注意力机制对模型的影响

模型 Model	像素准确率 Pixel accuracy P_a /%	精确率 Accuracy P_p /%	交并比 Intersection and union ratio I_u /%
UNet	86.37	87.62	78.45
UNet+SE	87.65	90.80	78.90
UNet+CBAM	87.95	91.46	80.54
UNet+ECA	87.08	91.73	79.93
UNet+CA	94.54	95.15	83.47

结果表明,加入 SE、CBAM、ECA 和 CA 四种注意力机制模块, P_a 相较于原始 UNet 模型分别提高了 1.28、1.58、0.71、8.17 个百分点。其中 CA 模块对模型的分割性能提升最大,CA 模块还将模型的 P_e 和 I_u 分别提高了 7.53、5.02 个百分点。由上述结果可知,CA 注意力机制同时关注通道注意力机制和空间注意力机制,明显地提升了模型的分割性能。

本文通过 Grad-CAM (gradient-weighted class activation map)^[33] 绘制热力图展示模型在复杂田间场景下对棉花的关注程度与分割效果。Grad-CAM 将与分割结果相关的区域用明亮的色彩标注,色彩越明亮表示与棉花相关度越高,为模型的检测结果提供可视化依据,如图 7 所示。由图 7b 和 7c 可以看出添加 CA 注意力机制模块之后,模型对图像中的棉花像素关注更加集中和丰富,进一步证明了 CA 注意力机制能有效地提升模型的分割精度。

2.1.2 不同激活函数的比较试验

为了增强棉花图像分割模型的表达能力,本文将 UNet 分别与 Leaky ReLU、SiLU 和 hardswish 函数相结合,分别测试其分割性能。表 2 展示了 3 种激活函数在 UNet 模型的性能表现。

从表 2 可知,将 Leaky ReLU 函数作为原始 UNet 模型主干部分的激活函数时,模型的各项指标均是最低的。与 Leaky ReLU 函数相比, SiLU 激活函数虽然略微提升了模型的性能,但是并不突出。当 hardswish 作为激活函数时,由于 hardswish 特有的截断和线性缩放的操作增强了训练过程中模型反向传播的鲁棒性,同时 hardswish

对边界值的影响更有利于分割田间小目标和复杂背景下的遮挡棉花目标。因此，模型获得了最佳的分割精度和分割速度，其中 hardswish 函数的 P_e 比 Leaky ReLU 和 SiLU 分别高出 0.83、0.62 个百分点。

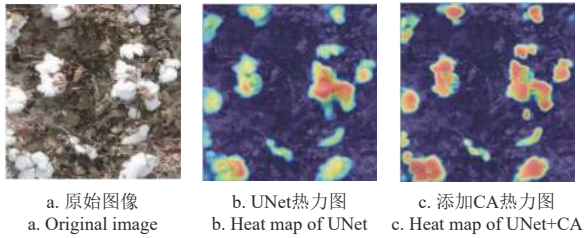


图 7 不同模型的热力图
Fig.7 Heat map for different models

表 2 不同激活函数对 UNet 模型的影响

Table 2 Effect of different activation functions on the UNet model

激活函数 Activation function	P_a /%	P_e /%	I_u /%
Leaky ReLU	86.57	87.42	79.25
SiLU	87.11	87.63	80.15
hardswish	87.16	88.25	84.15

2.1.3 不同主干网络的比较试验

在棉花图像分割任务中，由于不同高度的棉花存在着大小差异，模型难以分割出小目标棉花像素。这种尺寸差异容易导致模型的分割精度较低，是分割工作面临的重大挑战之一。为了增强模型对小尺寸的棉花像素的分割性能，使用 ResNet50 网络作为 UNet 模型的主干特征提取部分，充分结合浅层位置信息和深层语义信息，进一步增强模型的多尺度像素分割能力。表 3 展示了 UNet 模型不同特征提取网络分割的结果，使用 VGG16、MobileNetv3 和 ResNet50 网络进行试验研究。由表 3 可知，相比原始 UNet 模型，优化主干网络之后的模型检测精度均有提升，其中 ResNet50 表现最佳， P_a 、 P_e 和 I_u 分别提升了 2.50、3.16、2.29 个百分点。

表 3 不同主干网络分割的结果

Table 3 Segmentation results of different backbone

模型 Model	P_a /%	P_e /%	I_u /%
UNet	86.37	87.62	78.45
UNet+VGG16	85.14	87.98	78.72
UNet+MobileNetv3	85.08	87.86	78.68
UNet+ResNet50	88.87	90.78	80.74

2.2 RCH-UNet 模型消融试验

为了进一步分析本文所提出的一系列改进方法在执行棉花图像分割任务中的优势，本文基于 UNet 模型，

结合 2.1 节方法设计消融试验，试验结果如表 4 所示。

表 4 消融试验结果
Table 4 Results of ablation experiments

UNet	R	C	H	P_a /%	P_e /%	I_u /%
√				86.37	87.62	78.45
√	√			88.87	90.78	80.74
√		√		94.54	95.15	83.47
√			√	87.16	88.25	84.15
√	√	√		95.45	95.81	88.02
√	√		√	93.16	94.25	88.15
√		√	√	94.74	95.85	87.40
√	√	√	√	96.22	96.30	92.79

注：UNet 表示 UNet 分割模型，R 表示 ResNet50，C 表示 CA 注意力机制，H 表示 hardswish 激活函数。

Note: UNet represents UNet segmentation model, R represents ResNet50, C represents coordinate attention mechanism, and H represents hardswish activation function.

由表 4 可知，本文提出的 RCH-UNet 模型相比原始 UNet 模型，各项评价指标均存在明显提升。根据表 4 可以看出，在原始 UNet 模型基础上单独添加 R (ResNet50)、C (CA) 和 H (hardswish) 模块均能提升模型的分割效果。其中 R 和 C 对于分割模型影响最大，分别使 P_e 提升了 3.16、7.53 个百分点。ResNet50 和 CA 注意力机制的结合强化了特征融合和语义信息，从而增强分割模型的特征提取能力；同时，由于 hardswish 函数是非单调光滑的。非单调性有助于保持负值较小，从而稳定网络的梯度。结合 hardswish 函数对模型分割结果具有良好的泛化能力和有效的优化能力。因此 RCH-UNet 模型拥有最佳分割精度。

2.3 不同分割模型的对比试验

本文使用了经典的分割模型 UNet、PSPNet 和 DeepLabv3，在同一数据集上进行训练和测试，并将分割结果与本文提出的 RCH-UNet 模型进行比较。分割结果如图 8 所示，模型各项指标如表 5 所示。

根据表 5 中的数据可以看出，RCH-UNet 相较于原始 UNet 的 I_u 、 P_a 、 P_e 分别提高了 14.34、9.85、8.68 个百分点，相较于 PSPNet 的 I_u 、 P_a 、 P_e 分别提高了 23.04、17.67、10.14 个百分点，相较于 DeepLabv3 的 I_u 、 P_a 、 P_e 分别提高了 9.15、6.31 和 3.99 个百分点。PSPNet 模型的棉花图像分割效果最差， P_a 仅为 78.55%， I_u 仅为 69.75%。相比之下，DeepLabv3 模型表现稍好， P_a 为 89.91%， I_u 为 83.64%。尽管 DeepLabv3 模型的性能已经相当可观，但对于高精度的棉花图像分割任务仍然需要进一步改进。

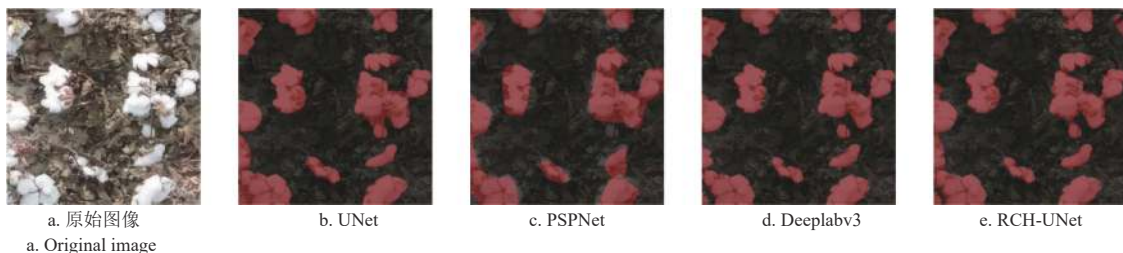


图 8 不同模型分割结果
Fig.8 Segmentation results with different models

表5 不同模型的对比试验

Table 5 Experimental comparison of different models

模型 Model	P_e /%	P_u /%	I_u /%
UNet	86.37	87.62	78.45
PSPNet	78.55	86.16	69.75
DeepLabv3	89.91	92.31	83.64
RCH-UNet	96.22	96.30	92.79

与前两种模型相比,本文提出的 RCH-UNet 模型展现出最佳的性能,各项指标都得到了大幅度提升,这意味着 RCH-UNet 在棉花图像分割方面具有更高的准确度和更好的边界定义能力。通过结合 ResNet50、CA 和 hardswish 激活函数, RCH-UNet 能够更好地捕捉棉花图像中的细节和纹理信息,从而得到更准确的分割结果。该研究表明,对于棉花图像分割任务,选择合适的特征提取网络非常关键。尽管传统的 PSPNet 和 DeepLabv3 模型在其他场景下可能具有良好的表现,但它们在棉花图像分割方面的效果受到限制。相比之下, RCH-UNet 模型通过特定的特征提取网络和激活函数选择,能够更好地适应棉花图像的特征,提高了分割性能。

此外,本研究提出的 RCH-UNet 模型和经典的 DeeplabV3、PSPNet、U-Net 模型在棉花图像数据集训练完成后,将 4 种分割模型部署到边缘设备 Jetson Xavier NX。采用同一张棉花图片在 Jetson Xavier NX 上进行模型测试,将模型分割单张棉花图片的时间作为计算效率评估指标。各模型单张图片分割耗时结果如表 6 所示。

表6 单张图片分割耗时对比

Table 6 Comparison of time consumption for single image segmentation

模型 Model	时间 Time/s
UNet	0.227
PSPNet	0.174
DeepLabv3	0.205
RCH-UNet	0.247

由表 6 可知, RCH-UNet 模型平均耗时为 0.247 s。由于在棉花图像分割任务中,分割精度要求比分割时间更关键,因此牺牲小部分分割速度选取最高分割精度的 RCH-UNet 模型。

2.4 棉花产量预测模型结果分析

本文基于 RCH-UNet 所提取的棉花像素比构建单因素产量预测模型,单因素产量预测模型如式(13)所示。

$$y = 38.23x_1 + 31.35 \quad (13)$$

式中 x_1 为棉花像素比, y 为棉花产量。单因素产量预测模型的 R^2 为 0.84。

为了进一步验证模型精度,本文基于 RCH-UNet 所提取的棉花像素比和 GLCM 所提取的棉花图像对比度、能量、相关性、均匀性 5 个特征值借助岭回归分析构建多因素产量预测模型,多因素产量预测模型如式(14)所示。

$$y = 28.9x_1 + 0.7x_2 + 4.2x_3 + 4.5x_4 + 1.5x_5 + 31.4 \quad (14)$$

式中 x_1 为棉花像素比、 x_2 为对比度、 x_3 为能量、 x_4 为相关性、 x_5 为均匀度、 y 为棉花产量。多因素棉花产量预测模型的 R^2 为 0.92。

在数据采集期间,本研究选择了 5 块棉田,并记录了机采棉产量和人工复采之后的棉花实际产量数据,同时对相应棉田应用本文预测模型进行产量预测,结果如表 7~表 8 所示。预测结果显示,单因素棉花产量预测模型平均绝对百分比误差为 9.98%;多因素棉花产量预测的平均绝对百分比误差为 9.254%。

比较两种模型的各项评价指标以及最终产量预测结果与误差可知,利用图像的更多隐藏信息,能够建立更准确和泛化能力更强的产量预测模型。

表7 单因素产量预测模型结果及误差

Table 7 Results and errors of one-factor yield prediction model

地块编号 Plot No.	实际产量 Actual yield/ (kg·hm ⁻²)	预测产量 Prediction yield/ (kg·hm ⁻²)	相对误差 Relative error R_e /%
1	7 139	6 232	12.7
2	6 427	5 823	9.4
3	5 430	4 947	8.9
4	5 743	5 157	10.2
5	6 273	5 727	8.7

表8 多因素产量预测模型结果及误差

Table 8 Results and errors of multifactor yield prediction model

地块编号 Plot No.	实际产量 Actual yield/ (kg·hm ⁻²)	预测产量 Prediction yield/ (kg·hm ⁻²)	R_e /%
1	7 139	6 375	10.7
2	6 427	5 755	10.5
3	5 430	5 058	6.9
4	5 743	5 262	8.4
5	6 273	5 653	9.9

3 讨论

将 RCH-UNet 与原始 UNet、PSPNet 和 DeepLabv3 进行比较分析, RCH-UNet 模型的性能指标明显优于三种对比分割模型。此外,与 SINGH 等^[13]提出的轻量化卷积神经网络模型相比, RCH-UNet 模型的 P_e 高出了 0.54 个百分点, I_u 高出 1.76 个百分点,并且 RCH-UNet 模型推理时间快出 0.391s。与 SAEED 等^[11]提出的三位分割棉花植株方法相比, RCH-UNet 模型的 P_e 高出 0.11 个百分点, I_u 高出 3.67 个百分点,推理时间快 0.633 s。与 LV 等^[14]提出的改进 PSPNet 分割模型相比, RCH-UNet 模型的 P_e 高出 16.76 个百分点, I_u 高出 10.77 个百分点。上述对比结果表明 ResNet50、CA 注意力机制和 hardswish 函数使 RCH-UNet 模型学习到了更具代表性和区分性的特征,从而大大提高了模型的整体性能,因此相较于其他分割模型获得了明显的分割精度提升。

将研究区域棉花实际产量与本文模型预测结果进行比较,接近程度表明本文所提出的方法在无人机图像的正射影像中进行棉花产量预测工作是有效的;然而,在垂直视图中下部棉花可能会被上部棉花、棉枝和棉叶所遮住,会对无人机图像的分析造成一定程度的影响。产量预测的准确性受到上述遮挡情况的影响,特别是在种植密度较高的区域,更多的棉花被遮蔽,导致预测结果产生更高的误差。此外,与 FENG 等^[16]提出基于无人机可见光图像提取植被指数预测棉花产量相比, RCH-UNet

模型实现了图像隐藏特征的自动提取; 同时与该方法相比本文所提出的棉花产量预测方法平均绝对百分比误差低 5.046 个百分点; 与王汇涵等^[2]提出基于 Sentinel-2A 卫星预测棉花产量相比, 本文产量预测方法的平均绝对百分比误差低 6.832 个百分点, 同时从时间成本角度考虑, 本文提出的棉花产量预测模型远远优于 Sentinel-2A 卫星预测方法; 与戴建国等^[15]提出基于无人机遥感平台与 CNN-BiLSTM 方法预测棉花产量研究结果相比, 本文所提出的产量预测方法平均绝对百分比误差高出 2.543 个百分点, 棉花产量预测模型在精度上仍有一定的提升空间, 分析原因主要包括以下两个方面: 1) 本研究的棉花图像数据集仅仅考虑了脱叶期一个时期, 图像中的时序特征变化较少, 忽视了棉花生长过程中生长动态信息对于棉花最终产量产生的影响; 2) 本研究在提取图像特征时没有充分利用植被指数、株高等相关特征参数。针对以上存在的问题, 在后续试验中可以进行如下两方面的改进: 1) 采集不同生育期的棉花图像, 将时序特征进行融合以提高模型的精度; 2) 提取棉花生长过程中的植被指数、株高和土壤肥力等特征参数构建多元棉花产量预测模型从而提高棉花产量预测精度。

4 结 论

本研究基于 UNet 分割模型构建 RCH-UNet 分割模型, 使用 RCH-UNet 模型从棉花脱叶期的无人机遥感影像中提取棉花像素比和图像纹理特征。最终通过棉花像素比和纹理特征与实际产量进行岭回归分析从而得到棉花产量预测模型。本文的主要结论如下:

1) 将 UNet 中传统的 CBR (convolution batch normalization ReLU) 下采样模块替换为 ResNet50 网络, 可以增强模型提取图像特征的能力, 为后续的棉花图像分割与产量预测奠定基础。

2) 借助 CA 注意力机制的通道和空间双重关注策略, 模型可以更加聚焦于目标相关的特征, 减少非目标特征的干扰, 提高模型的鲁棒性。有效地提高了模型对复杂背景中棉花图片的分割能力。

3) 在复杂的大田环境中, 棉花具有各种不规则的形状、细小的纹理等, ResNet50 和 hardswish 有效地帮助模型更好地学习和表示复杂的特征, 提高了图像中棉花的分割精度。

4) 通过 RCH-UNet 提取的棉花像素比成功实现了新疆密植棉花产量快速预测。结果显示, 通过棉花像素比和纹理特征构建的多因素产量预测模型的平均绝对百分比误差为 9.254%。研究结果表明, 从棉花图像中提取棉花像素比和提取纹理特征进行产量预测具有可行性。

基于目前的研究结果, 未来的研究将集中在提取棉花生长过程中存在的多个特征参数、解决棉花之间的互相遮挡问题、模型轻量化和将模型直接部署到无人机设备上, 实现在实际应用场景中的棉花产量快速预测。

[参 考 文 献]

- [1] 王振华, 王菲, 吕德生, 等. 不同土质下灌水盐度对滴灌棉花生理及产量品质的影响[J]. *农业工程学报*, 2023, 39(24): 69-78.
WANG Zhenhua, WANG Fei, LYU Desheng, et al. Effects of brackish water irrigation on the physiological characteristics, yield and quality of mulched drip irrigation cotton under different soil textures[J]. *Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE)*, 2023, 39(24): 69-78. (in Chinese with English abstract)
- [2] 王汇涵, 张泽, 康孝岩, 等. 基于 Sentinel-2A 的棉花种植面积提取及产量预测[J]. *农业工程学报*, 2022, 38(9): 205-214.
WANG Huihan, ZHANG Ze, KANG Xiaoyan, et al. Cotton planting area extraction and yield prediction based on Sentinel-2A[J]. *Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE)*, 2022, 38(9): 205-214. (in Chinese with English abstract)
- [3] 胡灿, 王兴旺, 王旭峰, 等. 基于 Android 手机的田间棉花产量预测系统设计[J]. *农业机械学报*, 2023, 54(s2): 252-259, 277.
HU Can, WANG Xingwang, WANG Xufeng, et al. Field Cotton Yield Prediction System Based on Android Mobile Phone[J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2023, 54(s2): 252-259,277. (in Chinese with English abstract)
- [4] 李兴. 新疆棉花生产的历史、现状与前景研究[D]. 陕西: 陕西师范大学, 2022.
LI Xing. Research on the History, Current Situation and Prospect of Cotton Production in Xinjiang[D]. Shaanxi: Shaanxi Normal University, 2022. (in Chinese with English abstract)
- [5] 国家统计局关于 2022 年棉花产量的公告[N]. *中国信息报*, 2022-12-27.
- [6] 刘传迹, 金晓斌, 徐伟义, 等. 2000-2020 年南疆地区棉花种植空间格局及其变化特征分析[J]. *农业工程学报*, 2021, 37(16): 223-232.
LIU Chuanji, JIN Xiaobin, XU Weiyi, et al. Analysis of the spatial distribution and variation characteristics of cotton planting in southern Xinjiang from 2000 to 2020[J]. *Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE)*, 2021, 37(16): 223-232. (in Chinese with English abstract)
- [7] 胡宇凯, 陈国栋, 王佳乐, 等. 密度对间作棉花新陆中 82 冠层结构及产量的影响[J]. *农业与技术*, 2021, 41(21): 64-66.
- [8] 李侃, 谢章书, 何玉玺, 等. 播期和种植密度对不同熟性棉花生理特性及产量的影响[J]. *湖南农业大学学报 (自然科学版)*, 2023, 49(3): 260-267.
LI Kan, XIE Zhangshu, HE Yuxi, et al. Effects of sowing date and planting density on physiological characteristics and yield of cotton with different maturity properties[J]. *Journal of Hunan Agricultural University (Natural Sciences)*, 2023, 49(3):

- 260-267. (in Chinese with English abstract)
- [9] 孙巨龙, 刘帅, 胡启星, 等. 不同种植密度对棉花空间成铃分布的影响[J]. 棉花科学, 2021, 43(1): 31-36.
SUN Julong, LIU Shuai, HU Qixing, et al. The influence of different planting density on the spatial distribution of cotton boll[J]. Cotton Sciences, 2021, 43(1): 31-36. (in Chinese with English abstract)
- [10] 陈兵, 张国蕾, 王静, 等. 机采棉等行距密植模式下棉花脱叶催熟效果研究[J]. 西北农业学报, 2019, 28(10): 1594-1601.
CHEN Bing, ZHANG Guolei, WANG Jing, et al. Effect of different defoliant on machine-picked cotton under equal spacing and dense planting model[J]. Acta Agriculturae Boreali-occidentalis Sinica, 2019, 28(10): 1594-1601. (in Chinese with English abstract)
- [11] SAEED F, SUN S, RODRIGUEZ-SANCHEZ J, et al. Cotton plant part 3D segmentation and architectural trait extraction using point voxel convolutional neural networks[J]. *Plant Methods*, 2023, 19(1): 1-23.
- [12] LI Y, CAO Z, LU H, et al. In-field cotton detection via region-based semantic image segmentation[J]. *Computers and Electronics in Agriculture*, 2016, 127: 475-486.
- [13] SINGH N, TEWARI V K, BISWAS P K, et al. Lightweight convolutional neural network models for semantic segmentation of in-field cotton bolls[J]. *Artificial Intelligence in Agriculture*, 2023, 8: 1-19.
- [14] LV Q, WANG H. Cotton boll growth status recognition method under complex background based on semantic segmentation[C]//2021 4th International Conference on Robotics, Control and Automation Engineering (RCAE). Wuhan, China: IEEE, 2021: 50-54.
- [15] 戴建国, 蒋楠, 薛金利, 等. 基于 CNN-BiLSTM 的棉花产量预测方法[J]. 农业工程学报, 2021, 37(17): 152-159.
DAI Jianguo, JIANG Nan, XUE Jinli, et al. Method for predicting cotton yield based on CNN-BiLSTM[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2021, 37(17): 152-159. (in Chinese with English abstract)
- [16] FENG A, ZHOU J, VORIES E D, et al. Yield estimation in cotton using UAV-based multi-sensor imagery[J]. *Biosystems Engineering*, 2020, 193: 101-114.
- [17] MA Y, MA L, ZHANG Q, et al. Cotton yield estimation based on vegetation indices and texture features derived from RGB image[J]. *Frontiers in Plant Science*, 2022, 13: 925986.
- [18] DUBE N, BRYANT B, SARI-SARRAF H, et al. Cotton boll distribution and yield estimation using three - dimensional point cloud data[J]. *Agronomy Journal*, 2020, 112(6): 4976-4989.
- [19] RONNEBERGER O, FISCHER P, BROX T. U-Net: Convolutional networks for biomedical image segmentation[J]. Springer International Publishing, 2015, 9351: 234-241
- [20] 孙俊, 宫东见, 姚坤杉, 等. 基于通道特征金字塔的田间葡萄实时语义分割方法[J]. 农业工程学报, 2022, 38(17): 150-157.
SUN Jun, GONG Dongjian, YAO Kunshan, et al. Real-time semantic segmentation method for field grapes based on channel feature pyramid[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2022, 38(17): 150-157. (in Chinese with English abstract)
- [21] 朱立学, 伍荣达, 付根平, 等. 基于多尺度串联空洞卷积的轻量化 UNet 香蕉图像分割[J]. 农业工程学报, 2022, 38(13): 194-201.
ZHU Lixue, WU Rongda, FU Genping, et al. Segmenting banana images using the lightweight UNet of multi-scale serial dilated convolution[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2022, 38(13): 194-201. (in Chinese with English abstract)
- [22] SAHIN H M, MIFTAHUSHUDUR T, GRIEVE B, et al. Segmentation of weeds and crops using multispectral imaging and CRF-enhanced U-Net[J]. *Computers and Electronics in Agriculture*, 2023, 211: 107956.
- [23] DIAO Z, GUO P, ZHANG B, et al. Maize crop row recognition algorithm based on improved UNet network[J]. *Computers and Electronics in Agriculture*, 2023, 210: 107940.
- [24] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas, NV, USA: IEEE, 2016: 770-778.
- [25] HU J, SHEN L, ALBANIE S, et al. Squeeze-and-Excitation Networks[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2020, 42(8): 2011-2023.
- [26] WOO S, PARK J, LEE JY, et al. Cbam: Convolutional block attention module[C]//15th European Conference on Computer Vision (ECCV). Munich, Germany: Springer, 2018: 3-19.
- [27] HOU Q, ZHOU D, FENG J. Coordinate attention for efficient mobile network design[C]//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Nashville, TN, USA: IEEE, 2021: 13713-13722.
- [28] LI Y, GENG T, STEIN S, et al. GAAF: Searching activation functions for binary neural networks through genetic algorithm[J]. *Tsinghua Science and Technology*, 2022, 28(1): 207-220.
- [29] ZHANG Y, MA B, HU Y, et al. Accurate cotton diseases and pests detection in complex background based on an improved YOLOX model[J]. *Computers and Electronics in Agriculture*, 2022, 203: 107484.
- [30] ALABI T R, ABEBE A T, CHIGEZA G, et al. Estimation of soybean grain yield from multispectral high-resolution UAV data with machine learning models in West Africa[J]. *Remote Sensing Applications: Society and Environment*, 2022, 27: 100782.
- [31] LIU Y, FENG H, YUE J, et al. Estimation of potato above-ground biomass based on unmanned aerial vehicle red-green-

- blue images with different texture features and crop height[J]. *Frontiers in Plant Science*, 2022, 13: 938216.
- [32] ROMANO E, BERGONZOLI S, PECORELLA I, et al. Methodology for the definition of durum wheat yield homogeneous zones by using satellite spectral indices[J]. *Remote Sensing*, 2021, 13(11): 2036.
- [33] SELVARAJU R R, COGSWELL M, DAS A, et al. Grad-cam: Visual explanations from deep networks via gradient-based localization[C]//2017 IEEE International Conference on Computer Vision (ICCV). Venice, Italy: IEEE, 2017: 618-626.

Prediction of cotton yield densely planted in Xinjiang of China using RCH-UNet model

LIU Xiang , TIAN Min^{*} , LIANG Jinyan

(College of Mechanical and Electrical Engineering, Shihezi University, Shihezi 832003, China)

Abstract: Cotton is one of the most important economic crops in China. It is of great significance to predict the yield of cotton. In this study, the Unmanned Aerial Vehicle (UAV) remote sensing platform was first used to collect some data on the densely planted cotton in Xinjiang, China. Five-point sampling was also utilized in the period of defoliation. A total of five cotton fields were then selected as well. Secondly, each image was divided into multiple sub-images and then enhanced by color space conversion, brightness, noise blurring, flipping and rotating operation. A high-quality dataset was constructed to estimate the cotton yield. Better training was achieved, compared with the original. Eventually, the cotton dataset with the construction site was used to extract the element rate from the cotton images using RCH-UNet (resnet coordinate hardswish UNet). Among them, the UNet was used as a baseline model to construct an improved U-type convolutional neural network. While the IoU (intersection over union), PA (pixel accuracy), and precision were taken as the evaluation metrics. The overall RCH-UNet improved the three metrics by 14.34, 9.85, and 8.68 percentage points, respectively, compared with the original UNet. Specifically, ResNet50 backbone feature extraction network was selected to replace the traditional CBR (convolution batch normalization ReLU) downsampling structure in UNet; The CA (coordinate attention) mechanism was combined with the UNet; The ReLU activation function in UNet was replaced with the hardswish activation function. The results showed that: 1) The richer semantic information in the ResNet50 was learned through the residual structure, indicating the better feature extraction and expression of the improved model. 2) The CA mechanism was significantly enhanced to learn the detailed features. At the same time, there was an effective reduction in the interference of irrelevant features to the model. 3) The hardswish activation function presented the stronger expression and feature fusion of UNet when performing up-sampling and jump connection. Subsequently, the performance of RCH-UNet was tested with the PSPNet and DeepLabv3 models under the same experimental conditions. The IoU, PA, and precision of RCH-UNet had improved by 9.15, 6.31, and 3.99 percentage points, respectively, compared with the DeepLabv3 model. A prediction model of cotton yield was constructed with the ridge regression from the cotton pixels extracted by the RCH-UNet model. The image texture features were also extracted by GLCM (gray-level co-occurrence matrix). The R^2 value of the improved model was 0.92, and the average relative error between the predicted and actual yield was 9.254%. The RCH-UNet model was accurately and effectively extracted from the cotton images; Meanwhile, the prediction model of cotton yield was effectively verified using deep learning and image processing. The UAV low-altitude imaging can also be expected to predict the yield of densely planted cotton in Xinjiang, China.

Keywords: deep learning; image processing; semantic segmentation; UNet; self-attention mechanism; dense cotton; yield prediction