

许鑫,耿庆,郑凯,等. 基于纹理特征与深度学习的小麦图像中的穗粒分割与计数[J]. 江苏农业学报, 2024, 40(4): 661-674.
doi:10.3969/j.issn.1000-4440.2024.04.010

基于纹理特征与深度学习的小麦图像中的穗粒分割与计数

许鑫¹, 耿庆¹, 郑凯², 石磊², 马新明^{1,3}

(1.河南农业大学信息与管理科学学院,河南 郑州 450002; 2.国家统计局河南调查总队,河南 郑州 450002; 3.河南农业大学农学院,河南 郑州 450002)

摘要: 穗粒数是小麦产量构成的重要因素和估测产量的参数之一,传统的人工计数方法耗时费力,人为因素影响大。为了实现对小麦穗粒数的智能、快速监测,以百农 307、新麦 26、稷麦 336 这 3 个小麦品种为试验材料,利用智能手机于小麦灌浆后期拍摄麦穗图像,随后对麦穗图像进行预处理并归一化为 480×480 像素大小,结合深度学习和迁移学习机制,构建基于冻结-解冻机制的 HRNet 模型的小麦小穗图像分割计数深度学习模型,利用图像处理算法、小麦小穗图像纹理特征确定小穗像素数与穗粒数之间的关系阈值,构建小穗粒数预测模型,实现对小麦穗粒的预测计数。结果表明,对比同样采用冻结-解冻机制的 PSPNet 模型、DeeplabV3+ 分割模型、U-Net 模型及无冻结解冻机制的 HRNet 模型,采用基于冻结-解冻机制的 HRNet 模型对小麦小穗的分割效果更优,且具有更好的鲁棒性,分割精确度为 0.959 4,平均交并比(*mIoU*)为 0.911 9,类别平均像素准确率(*mPA*)为 0.941 9,召回率为 0.941 9;通过 3 个不同品种小麦的麦穗图像对小穗进行计数,所得决定系数(R^2)为 0.92,平均绝对误差为 0.73,平均相对误差为 2.89%;籽粒计数的 R^2 为 0.92,平均绝对误差为 0.43,平均相对误差为 5.51%。由研究结果可知,基于冻结-解冻机制得出的 HRNet 模型的小麦小穗图像分割算法能够有效分割小麦图像中的小穗,并获得更加丰富的语义信息,可用于解决小目标图像分割困难及训练欠拟合问题,通过粒数预测模型可以快速、精确地对小麦的籽粒数进行预测,从而为小麦高效、智能化估产提供算法支撑。

关键词: 麦穗; 籽粒数; 图像处理; HRNet; 深度学习; 图像分割

中图分类号: TP391;S512.1

文献标识码: A

文章编号: 1000-4440(2024)04-0661-14

Segmentation and counting of wheat spikes and grains based on texture features and deep learning

XU Xin¹, GENG Qing¹, ZHENG Kai², SHI Lei², MA Xin-ming^{1,3}

(1.College of Information and Management Science, Henan Agricultural University, Zhengzhou 450002, China; 2.National Bureau of Statistics Henan Survey Corps, Zhengzhou 450002, China; 3.College of Agronomy, Henan Agricultural University, Zhengzhou 450002, China)

收稿日期:2023-02-27

基金项目:国家统计局重大统计专项(2022ZX11);河南省科技创新杰出人才项目(184200510008);河南省现代农业产业技术体系项目(S2010-01-G04)

作者简介:许鑫(1984-),男,河南邓州人,博士,副教授,主要从事智慧农业与大数据技术方面的研究。(E-mail)xuxin468@163.com

通讯作者:马新明,(E-mail)wheatdoctor@163.com

Abstract: The number of grains per ear is an important factor in the composition of wheat yield and one of the parameters for estimating wheat yield. The traditional manual counting method is time-consuming and labor-intensive, and human factors have a great influence. In order to realize the intelligent and rapid monitoring of the number of grains per ear, three varieties of Bainong 307, Xinmai 26 and Jimai 336 were used as test materials, and the

wheat ear images were taken with a smart phone at the late stage of wheat grain filling. Based on the image processing technology, the wheat ear images were preprocessed and normalized to 480×480 pixels. Combining deep learning and transfer learning mechanisms, a HRNet wheat spikelet segmentation and counting deep learning model based on the freeze-thaw mechanism was constructed. Image processing algorithms and wheat spikelet texture features were used to determine the threshold of the relationship between the number of spikelet pixels and the number of grains per spike. The spikelet-grain number prediction model was constructed to realize the prediction and counting of wheat spikes. The results showed that compared with PSPNet, DeeplabV3 + segmentation model, U-Net which also used freeze-thaw mechanism and HRNet without freeze-thaw mechanism, the HRNet model based on the freeze-thaw mechanism had a better segmentation effect on wheat spikelets, and had better robustness. The segmentation accuracy was 0.959 4, the mean intersection over union ($mIoU$) was 0.911 9, the mean pixel accuracy (mPA) was 0.941 9, and the recall rate was 0.941 9. The spikelets were counted by the images of three different wheat varieties. The determination coefficient (R^2) was 0.92, the average absolute error was 0.73, and the average relative error was 2.89%. The R^2 of grain counting was 0.92, the average absolute error was 0.43, and the average relative error was 5.51%. It shows that the HRNet wheat spikelet image segmentation algorithm through the freeze-thaw mechanism can effectively segment wheat spikelets and obtain richer semantic information, which can be used to solve the problems of difficult segmentation of small target images and training underfitting. The model can quickly and accurately predict the number of wheat grains, so as to provide algorithm support for efficient and intelligent yield estimation of wheat.

Key words: wheat ear; grain number; image processing; HRNet; deep learning; image segmentation

小麦是中国重要的粮食作物,2021 年中国小麦总产量达 1.37×10^8 t, 约占粮食总产量的 21%。小麦产量关乎国家粮食安全^[1], 科学准确地预测小麦产量有助于保障粮食供给安全和社会稳定^[2]。小麦灌浆期是决定其产量形成的关键生长期, 小麦的单位面积产量由单位面积穗数、每穗粒数和千粒质量 3 个要素构成^[3], 因此研发快速准确的小麦测产方法是科学估测小麦产量的迫切需求。传统的小麦估产方式主要是在小麦收获前进行人工调查, 估测的产量不仅常常受到人为因素的影响, 而且耗时耗力、效率低, 从而制约了小麦产量估测的时效性、准确度^[1,3]。相反, 图像处理、机器学习和计算机视觉等技术可以快速准确地进行小麦产量估测, 能够为小麦植株表型信息的获取提供技术支持^[4]。

穗粒数是进行禾谷类作物产量估测的 3 个要素之一, 已经成为智能化估产的重要研究方向。赵三琴等^[5]采用稻穗结构图像特征测量方法, 明确了稻穗结构图像特征与籽粒数间的相关关系, 为稻穗穗粒数的估测提供了依据。王栋等^[6]在无需脱粒的条件下设计了一种智能表型解析算法, 通过分析稻穗图像的表型信息, 结合图像处理、深度学习方法考察每穗粒数和多种穗粒的相关性状。通过借鉴水稻穗粒数的估测研究, 杜世伟等^[7]从小麦穗的结构特征入手, 通过小麦小穗面积像素数与穗粒数间的关系间接获取穗粒数, 证明了通过麦穗结构特征研究

穗粒数的可行性。

图像处理和机器学习技术的发展, 为麦穗数、穗粒数的分割识别提供了一种重要的监测手段^[8]。虽然图像处理技术被广泛用于识别麦穗数和穗粒数, 但这些方法侧重于纹理特征、颜色特征和形态特征的提取, 在效率、实际应用方面仍存在问题。基于农学知识, 杜世伟等^[7]采用拟合抛物线的方法对麦穗的小穗进行分割, 并对小穗数、籽粒数进行精确统计, 准确率高达 97%。Fernandez-Gallego 等^[9]采用局部最大峰值法计算大田红-绿-蓝 (RGB) 彩色图像中的麦穗数, 计数成功率高于 90%。在目前的图像处理技术中, 图像特征的提取仍然需要大量人工干预, 这对环境、技术提出了很高要求, 而机器学习被证明在图像分割和识别领域具有显著优势。刘哲等^[10]基于颜色特征、K-means 聚类技术提出了一种全新的小麦麦穗计数算法, 其识别准确率可达 94%。Xu 等^[11]通过 K-means 聚类算法实现了麦穗轮廓特征的自动提取, 进而成功构建了一种高效卷积神经网络 (CNN), 将麦穗识别精度提高到 98.3%。然而, 传统图像处理技术和机器学习方法仍然存在识别分割时间较长且效率较低、复杂图像的识别分割效果差等问题^[5,12]。

深度学习技术在不同领域、应用场景和尺度下展现出了优异的特征提取能力和泛化能力, 被广泛

应用于识别麦穗数、穗粒数。Wei 等^[13]结合迁移学习方法构建了 Faster R-CNN 网络,并优化了小麦籽粒检测与计数模型,使模型的籽粒计数错误率小于 3%。Hu 等^[14]基于注意力机制实现了一种全新的生成对抗网络来计算小麦穗数、小穗数,识别分割小麦穗和小穗的 $F1$ 值为 84.9%。Dandrifosse 等^[15]在灌浆期获取小麦图像,并将深度学习方法与小麦的 RGB 图像相结合,实现了田间小麦麦穗的计数,麦穗检测、分割的平均 $F1$ 值分别为 0.93、0.86,相较于人工观察方法取得了更好的结果。Zhao 等^[16]提出了一种改进的 YOLOv5 方法,用于准确检测无人机图像中的小麦穗,平均检测准确率 (AP) 为 94.1%。虽然深度学习技术在小麦表型信息的提取方面有着不可比拟的优势,能够实现较高精度的图像分割与识别^[17],但仍然存在小麦穗粒图像的获取大多需要专业相机等设备、难以在复杂的生产中应用^[18]、针对密集小目标的图像识别与分割能力不足、目标之间容易出现粘连、精度难以提高等问题^[4,14]。

基于小麦穗粒识别、分割及计数中存在的精度、效率问题,本研究拟以小麦穗粒数为研究对象,选择百农 307、新麦 26 和稷麦 336 这 3 个小麦品种进行大田试验,利用移动终端拍摄麦穗图像,构建麦穗穗粒数深度学习分割模型。利用图像处理技术对小麦小穗的分割结果进行处理,结合小麦小穗纹理特征构建小穗粒数预测模型,以期实现对小麦穗粒的预测计数,获得快速高效的分割计数结果并对小麦产量进行精确估测。

1 材料与方法

1.1 试验设计

试验在河南农业大学原阳科教园区 ($35^{\circ}6'46''$ N, $113^{\circ}56'51''$ E) 进行,选择主栽小麦品种百农 307、新麦 26 和稷麦 336,按照正常施氮量施肥,图 1 为试验地点的无人机图像。小麦播种量采用当地最佳播种量,百农 307 的播种量为 187.5 kg/hm^2 ,新麦 26 的播种量为 187.5 kg/hm^2 ,稷麦 336 的播种量为 142.5 kg/hm^2 ,播种日期为 2021 年 10 月 23 日,行距为 20 cm。试验施用的氮肥均为尿素 (含 46% N),磷肥为过磷酸钙 (含 12% P),钾肥为氯化钾 (含 60% K),氮肥作基肥、追肥的比例为 6:4,追肥在拔节期施入,磷钾肥作基肥一次性施入。田间管理措

施同一般高产田。

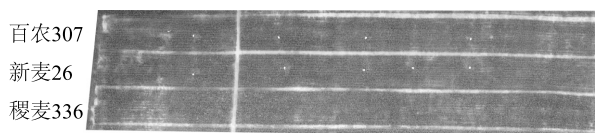


图1 试验地点的无人机图像

Fig.1 An image of the experimental site acquired by the unmanned aerial vehicle

1.2 数据采集

从小麦灌浆期开始进行采样,图像采集设备为 HUAWEI Mate 40 Pro (主摄像头像素 5.000×10^7 , 副摄像头像素 2.000×10^7) 和 Realme Q3 (主摄像头像素 4.800×10^7 , 副摄像头像素 8.000×10^6), 拍摄时间为 9:00–16:00。由于数据集的复杂性和多样性对模型的泛化水平有很大影响,因此试验采用 2 种图像采集方式:一种是离体取样,在实验室环境下对麦穗进行图像采集;另一种是就地取样,在大田环境下对麦穗进行图像采集^[4]。在对每株麦穗进行图像采集后,对麦穗的小穗数、籽粒数进行人工计数。图像采集时,将移动设备平行于麦穗,通过调整与麦穗之间的距离使麦穗完整清晰地出现在移动设备取景框中,并显示清晰的穗部图像。为了降低拍摄距离差异所造成的误差,拍摄设备与麦穗之间的距离保持在 10 cm 左右。根据该方法得到的原始麦穗图像如图 2 所示。

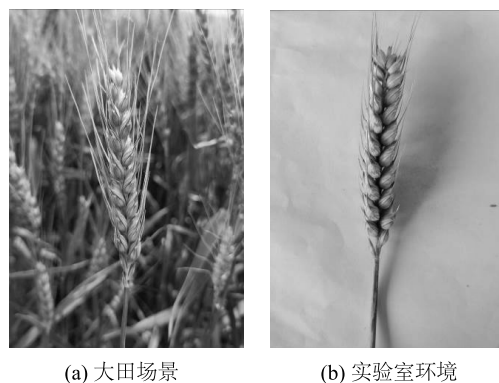


图2 拍摄的原始麦穗示例

Fig.2 Original wheat ear images

分别取不同品种、不同氮肥处理、不同拍摄背景的小麦 30~40 株,获得原始麦穗图像共 660 张,麦穗数据集的情况见表 1。

表 1 麦穗数据集信息
Table 1 Wheat ear dataset information

小麦品种	拍摄背景	拍摄日期	天气	分辨率 (mm)	图像大小 (像素)	拍摄设备	焦距 (mm)	图像数量 (张)
百农 307	大田	2022-05-15	晴	0.26~0.56	3 072×4 096	HUAWEI Mate 40 Pro	7	10
	大田	2022-05-15	晴	0.35~0.54	3 000×4 000	Realme Q3	5	10
	白纸	2022-06-08	多云	0.26~0.56	3 680×5 408	HUAWEI Mate 40 Pro	4	90
	大田	2022-05-15	晴	0.26~0.56	3 072×4 096	HUAWEI Mate 40 Pro	7	10
	大田	2022-05-15	晴	0.35~0.54	3 000×4 000	Realme Q3	5	10
	白纸	2022-06-08	多云	0.26~0.56	3 680×5 408	HUAWEI Mate 40 Pro	4	90
新麦 26	大田	2022-05-15	晴	0.26~0.56	3 072×4 096	HUAWEI Mate 40 Pro	7	10
	大田	2022-05-15	晴	0.35~0.54	3 000×4 000	Realme Q3	5	10
	白纸	2022-06-08	多云	0.26~0.56	3 680×5 408	HUAWEI Mate 40 Pro	4	90
	大田	2022-05-15	晴	0.26~0.56	3 072×4 096	HUAWEI Mate 40 Pro	7	10
	大田	2022-05-15	晴	0.35~0.54	3 000×4 000	Realme Q3	5	10
	白纸	2022-06-08	多云	0.26~0.56	3 680×5 408	HUAWEI Mate 40 Pro	4	90
稷麦 336	大田	2022-05-15	晴	0.26~0.56	3 072×4 096	HUAWEI Mate 40 Pro	7	10
	大田	2022-05-15	晴	0.35~0.54	3 000×4 000	Realme Q3	5	10
	白纸	2022-06-08	多云	0.26~0.56	3 680×5 408	HUAWEI Mate 40 Pro	4	90
	大田	2022-05-15	晴	0.26~0.56	3 072×4 096	HUAWEI Mate 40 Pro	7	10
	大田	2022-05-15	晴	0.35~0.54	3 000×4 000	Realme Q3	5	10
	白纸	2022-06-08	多云	0.26~0.56	3 680×5 408	HUAWEI Mate 40 Pro	4	90

1.3 技术路线

对麦穗图像进行预处理,形成麦穗数据集,并使用深度学习分割网络对小麦小穗进行图像分割,经过训练得到预测模型,并调用预测模型对测试集进

行测试,结合图像处理技术确定小穗面积像素数和穗粒数之间的阈值关系,构建粒数预测模型,从而实现小麦穗粒数的精准预测和计数。图 3 展示了本研究中拟采用的技术路线。

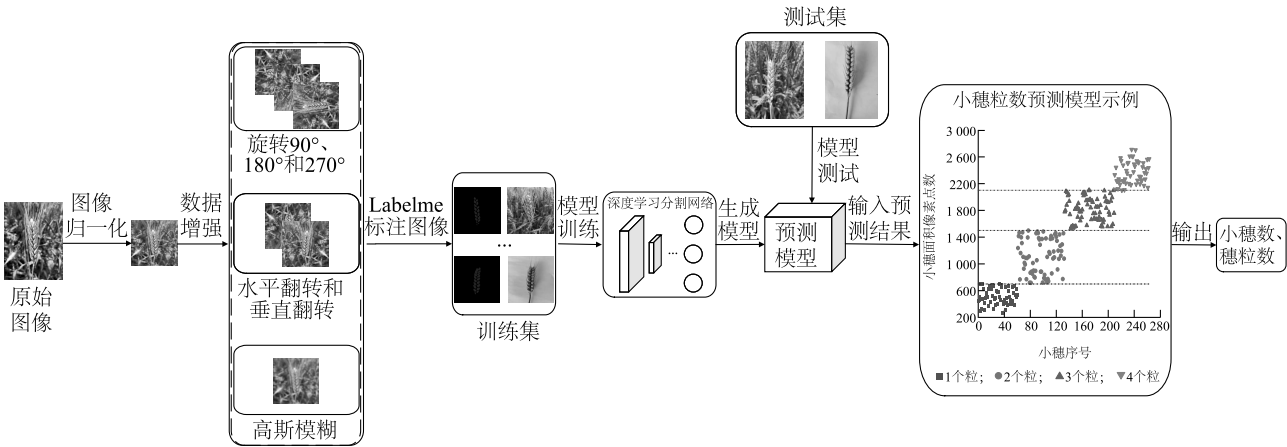


图 3 小麦穗粒预测和计数的技术路线
Fig.3 Technology roadmap for predicting and counting wheat spikes and grains

1.4 数据预处理

1.4.1 数据归一化 数据归一化是将所有图像转换成统一尺寸以方便模型的训练,同时可降低模型运算量、减少过拟合的风险。当原始麦穗图像的尺寸过大时,对设备的要求会较高。为了兼顾设备运算能力、图像数量及质量,需要对原始图像进行归一化处理^[16]。因此,在本研究中的模型进行训练前,先将数据集中的图像尺寸归一化为480×480像素。

1.4.2 数据增强 图像数量会影响模型的训练,导致测试精度下降。数据增强利用原始数据集中的图像进行扩展,能够解决图像数量不足的问题^[14]。随机裁剪可能会去除麦穗图像的特征信息,因此主要通过旋转90°、180°、270°及水平、垂直翻转来增强数据集。图像的细节层次和噪声都能影响模型的训练,这就需要通过高斯模糊对图像进行处理。通过不断调整高斯卷积核大小找出最合适的高斯模糊处理。经比较,将高斯卷积核的大小设置为5×5。

1.4.3 数据集的构建 为了实现麦穗图像的分割任务,本研究采用监督学习模式,即先人工标注数据样本,再训练深度学习网络使其具备一定的泛化能力^[19]。通过 Labelme 图像标注工具手动标记小麦小穗、背景2类分割对象,并通过标注信息将其转为遮罩图像。麦穗图像与其遮罩图像一起构

成了深度学习分割模型所需的数据集。在麦穗数据集中,3个品种、每种处理的图像数量均匀分布,训练集和验证集的分布比例为9:1^[11],其中训练集有594张图像,验证集有66张图像,测试集有66张图像。

1.5 模型的构建

小麦小穗图像的分割是复杂背景下的二分类任务,小穗图像之间像素点相近,图像粘连严重,因此对深度学习模型的解析力和全局信息获取能力要求较高。为了确保分割精度和计算效率,本研究构建 HRNet、PSPNet、DeeplabV3+分割模型及 U-Net 来进行小麦小穗图像的分割研究。

1.5.1 HRNet 模型 HRNet 的特点是多个分辨率分支并行和不同分支信息的交互,能够保持高分辨率,并获得强大的语义信息、精准的位置信息,从而在不断上下采样的过程中避免大量有效信息的丢失,对于位置敏感的语义分割任务而言效果显著。将数据集中的麦穗图像转换为特征图后,经过多个卷积块的并行及多分辨率融合模块得到不同融合输出。针对语义分割任务,HRNet 的表征分支通过上采样来提高低分辨率特征图像的分辨率,并将其堆叠、融合在一起形成小穗预测图^[20]。基于 HRNet 的麦穗小穗数分割网络结构如图4所示。

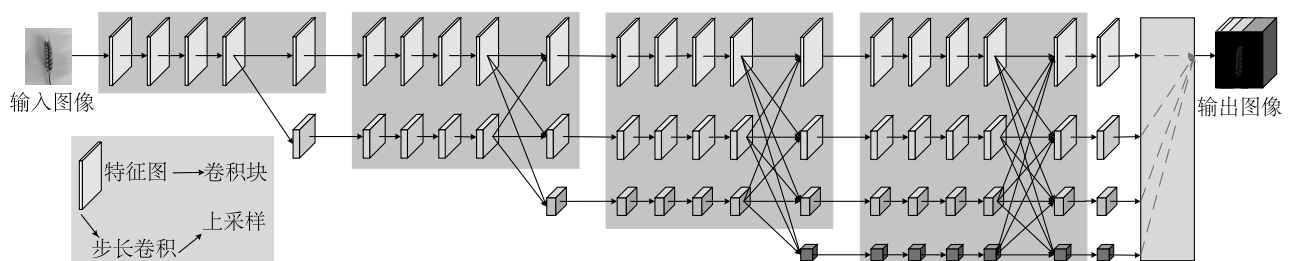


图4 基于 HRNet 的麦穗小穗数分割网络结构

Fig.4 Network structure of wheat spikelets segmentation based on HRNet

1.5.2 PSPNet 模型 PSPNet 在全卷积网络 (FCN) 的基础上进行了改进,以增加了空洞卷积的 ResNet 网络作为特征提取网络,对输入的麦穗图像进行特征提取。将提取的特征输入到金字塔池化模块,便得到不同深度且维度缩减后的金字塔特征。这些金字塔特征经过上采样和特征合并,最终输出小穗特

征图^[21]。基于 PSPNet 的麦穗小穗数分割网络结构如图5所示。

1.5.3 DeeplabV3+分割模型 DeeplabV3+分割模型为编码器-解码器结构^[22]。在编码器部分,数据集中的麦穗图像通过 Backbone 和空洞空间金字塔池化 (ASPP) 处理得到5组特征图并将其融合,最后作

为解码器的输入^[23]。在解码器部分,经过通道降维、插值上采样和卷积块等处理,便能得到分辨率大小与原图相同的小穗预测图^[24]。基于 DeeplabV3+ 分割模型的麦穗小穗数分割网络结构如图 6 所示。

1.5.4 U-Net 模型 U-Net 也是编码器-解码器结构的深度学习网络。数据集中的麦穗图像通过编码器提取小穗特征,其中主要采用卷积模块、池化层获得

输入图像的高层级特征向量,并将其作为解码器的输入^[25]。在解码器中,通过反卷积来增大特征图像分辨率并用解码模块对其进行解码。经过解码器处理之后,能得到特征图中像素属于小穗类别的概率,从而判断像素的类别,最终生成小穗图像分割结果^[26],基于 U-Net 的麦穗小穗数分割网络结构见图 7。

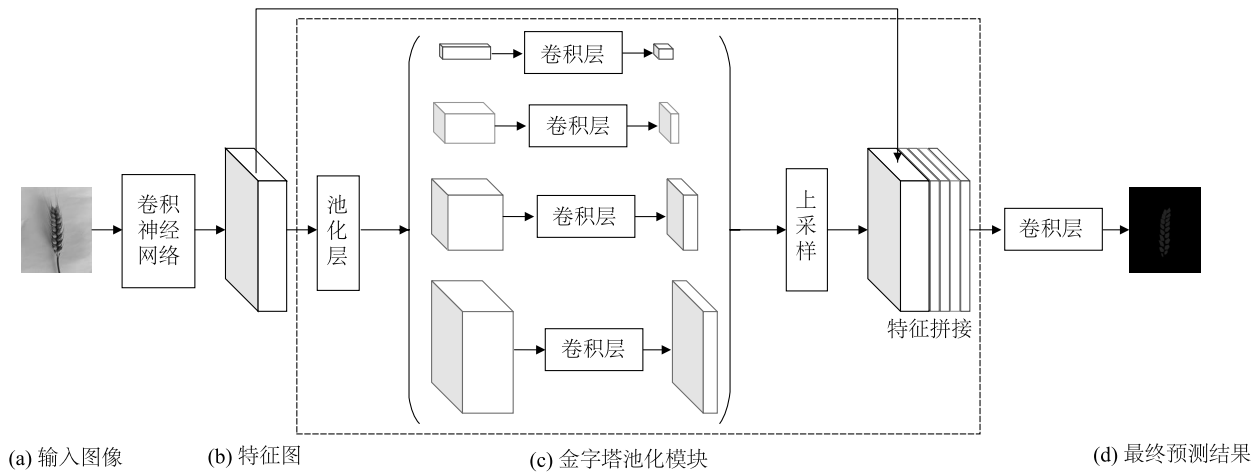
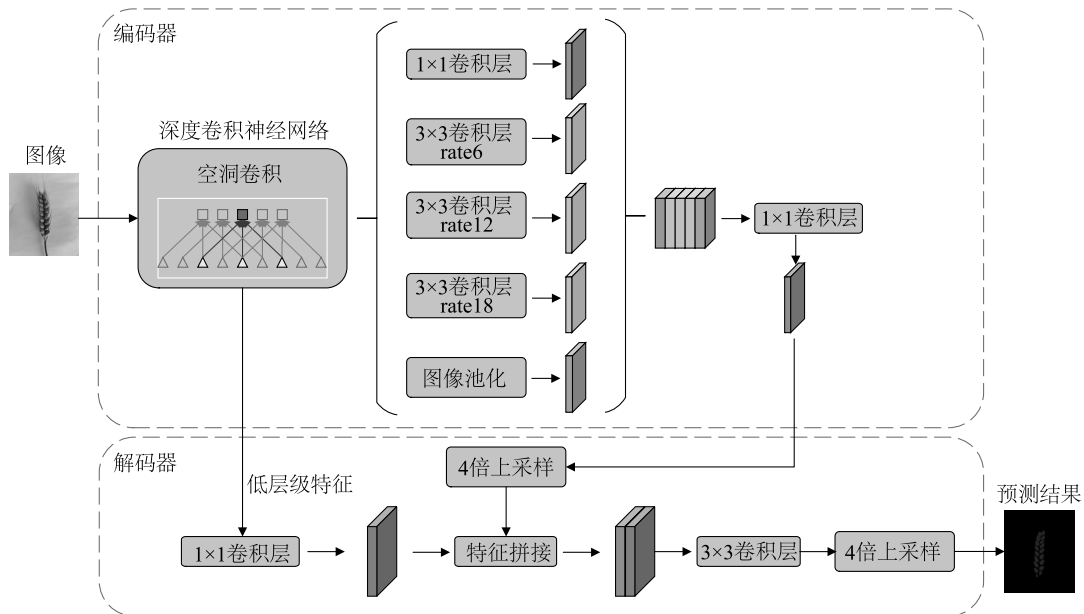


图 5 基于 PSPNet 的麦穗小穗数分割网络结构

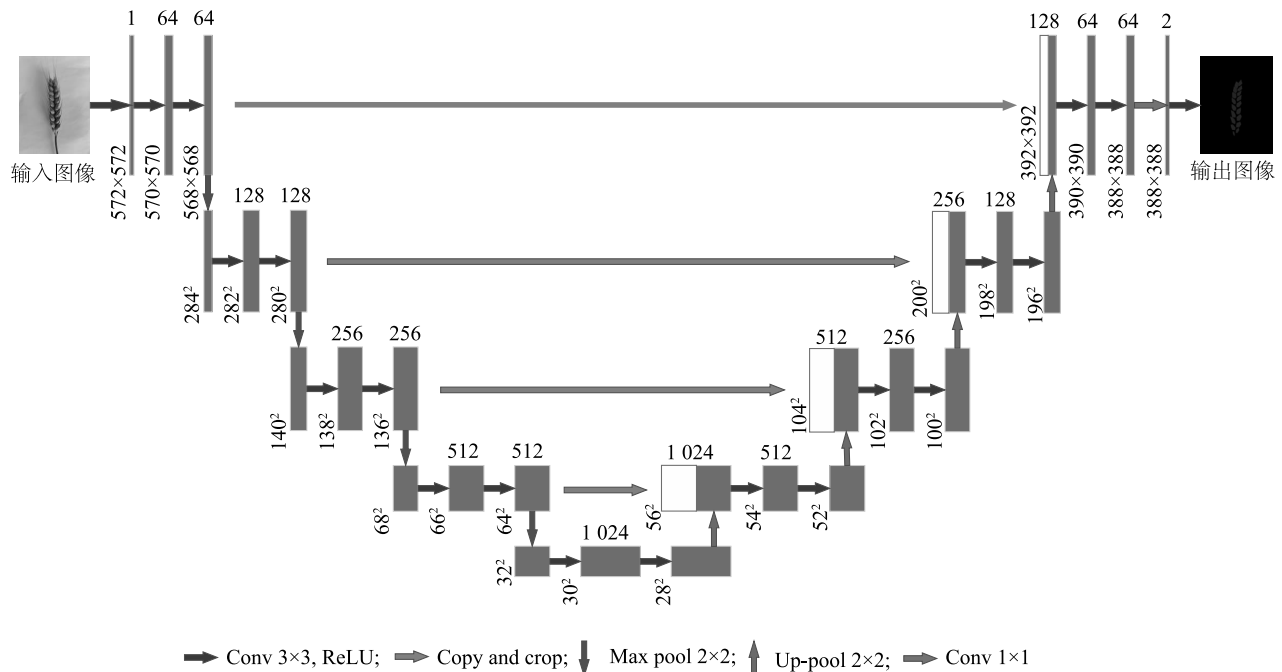
Fig.5 Network structure of wheat spikelets segmentation based on PSPNet



rate6 指 3×3 卷积的膨胀率为 6, rate12, rate18 的含义类推。

图 6 基于 DeeplabV3+ 的麦穗小穗数分割网络结构

Fig.6 Network structure of wheat spikelets segmentation based on DeeplabV3+ model



方块顶部的数据是通道数,左侧的数据是图像的大小(单位:像素)。

图7 基于 U-Net 的麦穗小穗数分割网络结构

Fig.7 Network structure of wheat spikelets segmentation based on U-Net

1.6 基于纹理特征的小穗粒数预测模型

有研究发现,由麦穗结构特征间接获取穗粒数具有可行性,通过观察麦穗的几何和纹理特征,能够明确麦穗结构图像特征与籽粒数间的相关关系^[6-7,12]。小麦小穗通常成对存在于穗轴两侧,但是由于顶端的小穗是幼穗分化后最后出现的小穗,因此并没有与其成对的小穗。在确定小穗数量及其面积像素数(图像面积对应的像素数)之后,通过基于纹理特征的小穗粒数预测模型便能获得各个小穗对应的籽粒数,各小穗对应的籽粒数之和就是小麦总穗粒数。小麦小穗的纹理特征如图8所示。

选取的麦穗样本经过深度学习分割模型预测后,部分小穗之间会有粘连,需要用图像处理方法消除重叠或粘连的部分。通过 OpenCV、NumPy 读入预测结果^[27],并对预测结果进行灰度处理,将色彩空间由 RGB 转换为灰度 (GRAY),之后便可以 120 为阈值进行二值化处理,大于该阈值的像素点的值设置为 255(即白色),小于该阈值的像素点的值设置为 0(即黑色)。二值化图像需要通过腐蚀变换来消除噪点^[28],随后可以发现原本重叠的部分不再重叠,便可根据提供的边界点绘制形状,从而进行轮廓提取与轮廓面积的计算。提取到的轮廓数、轮廓面



图8 小麦小穗的纹理特征

Fig.8 Texture features of wheat spikelets

积即为小麦的小穗数、面积像素数。

不同拍摄距离对小麦小穗像素数和穗粒数之间的关系阈值有一定影响,为了避免这种影响带来的误差,在图像采集过程中使移动设备与麦穗之间的距离基本保持一致,以确保小麦穗部图像的一致性^[5,7]。各个小麦品种均选取 200 个小穗样本,从而确定小穗面积像素数与穗粒数之间的关系阈值,通过该阈值可以获得具有 1~4 个籽粒小穗的面积像

素数范围。

根据分割后各小穗面积的像素数推测其代表的粒数,而粒数不同会导致明显的面积差异。由图 9 可以看出,小穗的像素数(Y)被 3 条阈值线分为 4 个区间,分别是 $Y < 700$ 、 $700 \leq Y < 1\,500$ 、 $1\,500 \leq Y \leq$

$2\,100$ 与 $Y > 2\,100$,小穗面积像素数分别对应 1 个、2 个、3 个、4 个籽粒。该方法可以消除麦穗一侧各小穗之间重叠与粘连的部分,并准确地计算其面积对应的像素数,从而根据划定的阈值预测各小穗代表的籽粒数。

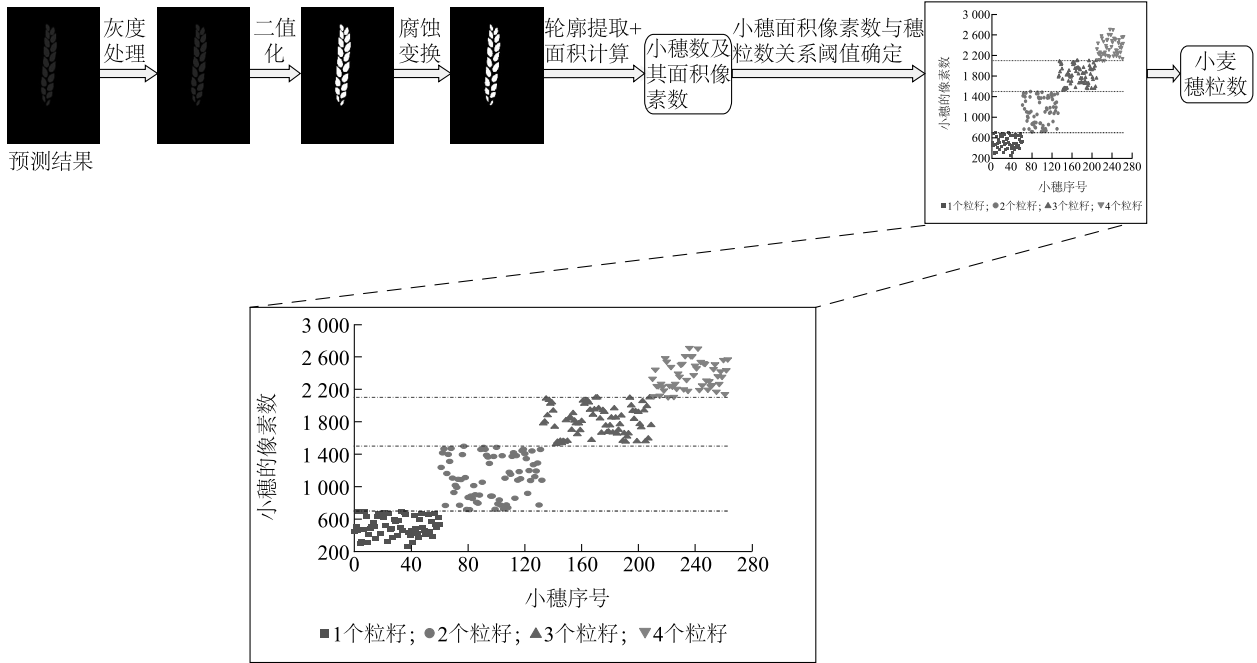


图 9 基于纹理特征的小穗粒数预测模型

Fig.9 Spikelet-grain number prediction model based on textural feature

1.7 精度评价指标

本研究采用精确度 ($Precision$)、召回率 ($Recall$)、类别平均像素准确率 (mPA) 和平均交并比 ($mIoU$) 等评价指标评估分割模型的精度,这些指标都是由混淆矩阵中的参数计算得出的,而混淆矩阵的作用是对每个真实像元的位置和预测像元的位置进行比较,从而对模型进行评估^[29]。

$$Precision = TP / (TP + FP) \quad (1)$$

$$Recall = TP / (TP + FN) \quad (2)$$

$$IoU = TP / (TP + FP + FN) \quad (3)$$

$$mPA = \frac{\sum(P_k)}{k} \quad (4)$$

$$mIoU = \frac{1}{k+1} \sum_{i=0}^k \frac{TP}{TP + FP + FN} \quad (5)$$

式中, TP 为被分割模型预测为真的正样本数, FP 为被分割模型预测为真的负样本数, FN 为被分割模型预测为假的正样本数, IoU 为交并比, k 为总类别数, P_k 为每个类别的像素预测准确率。

为了评估计数模型的精度,本研究采用的指标包括均方根误差 ($RMSE$)、平均绝对误差 (MAE)、平均相对误差 (MRE) 和决定系数 (R^2)^[30]。其中,以人工计数得到的小穗数、穗粒数作为真实值,以根据图像分割算法、粒数预测模型得到的小穗数、穗粒数作为预测值。

$$p_i = \frac{|x_i - y_i|}{x_i} \times 100\% \quad (6)$$

$$R^2 = 1 - \frac{\sum_{i=1}^n (x_i - y_i)^2}{\sum_{i=1}^n (x_i - \bar{x})^2} \quad (7)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - y_i)^2} \quad (8)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |x_i - y_i| \quad (9)$$

$$MRE = \frac{1}{n} \sum_{i=1}^n p_i \times 100\% \quad (10)$$

式中, p_i 为单个样本的相对误差, x_i 为样本麦穗小穗数或籽粒数的真实值, y_i 为样本麦穗小穗数或籽粒数的测量值, \bar{x} 为样本麦穗小穗数或籽粒数真实值的平均值, n 为各个小麦品种的样本数量。

1.8 训练环境

本研究模型的训练环境如下:中央处理器(CPU)为 Intel (R) Xeon (R) Silver 4114 CPU® 2.20 GHz 处理器,内存为 64 GB,图形处理器(GPU)为 NVIDIA Corporation GP104GL (16 GB 显存),操作系统为 Ubuntu,深度学习框架为 Pytorch 1.12,编译环境为 Python 3.9.12,在相同试验环境下对所有算法进行对比。

一般情况下,网络若从 0 开始训练得到的效果并不好,出现这种现象的原因主要是权值太过随机、特征提取效果不明显^[30]。因此根据迁移学习的思想,本研究构建了冻结-解冻机制。由于神经网络主干的特征提取部分提取到的特征具有通用性,因而模型的预训练权重在不同数据集上也有通用性。为了加快训练效率及防止权值被破坏,可以对神经网络的主干部分进行冻结训练^[31]。在冻结训练阶段,主干网络被冻结,使用预训练权重进行训练,特征提取网络不变,只细微地调整网络。在解冻训练阶段,主干网络被解冻,特征提取网络发生改变,所有参数都将被更新。

冻结阶段的训练参数如下:模型开始的训练世代(*Init_Epoch*,取值为 0)、模型冻结训练的迭代次数(*Freeze_Epoch*,取值为 50)及模型冻结训练的批大小(*Freeze_batch_size*,取值为 16)。解冻阶段的训练参数如下:模型训练的总迭代次数(*UnFreeze_Epoch*,取值为 300),模型在解冻后的批大小(*Unfreeze_batch_size*,取值为 8)。

2 结果与分析

2.1 小麦小穗分割模型的训练结果

使用基于冻结-解冻机制的 HRNet 模型、PSPNet 模型、DeeplabV3+分割模型及 U-Net 模型对相同的麦穗训练集进行训练。从图 10 可以看出,在不断迭代的过程中,4 种模型的 *mIoU* 值都处于上升趋势,并且迭代次数的增加使分割模型逐渐稳定,最终达到稳定收敛的状态,基于冻结-解冻机制的 HRNet 模型的 *mIoU* 值基本稳定在 0.910 0 左右,可见该模型对麦穗数据集的分割效果较好。该模型在训练和

验证的过程中表现出良好的稳定性,损失值快速下降并最终收敛到 0.014 0 左右,网络的训练效果、收敛速度较好,不存在损失值突变的情况,且变化幅度较小。损失值在训练集、验证集上的变化趋势基本一致,差距也很小,说明在梯度计算过程中,该模型能够迅速找到合适的梯度方向,具有较好的性能稳定性和学习效果,并且泛化能力较强。

2.2 小麦小穗分割模型的性能比较

为了衡量本研究所提模型的有效性,根据不同分割模型采用不同主干网络、优化器及学习率,并采用多种评价指标对比模型训练的结果和性能。由表 2 可以看出,以 hrnetv2_w32 为主干网络、Adam 为优化器的基于冻结-解冻机制的 HRNet 模型取得了最好的分割精度(*mIoU* = 0.911 9),以 hrnetv2_w32 为主干网络、Adam 为优化器的基于无冻结-解冻机制的 HRNet 模型次之(*mIoU* = 0.901 9),以 MobileNetv2 为主干网络、SGD 为优化器的 PSPNet 模型的分割精度最低(*mIoU* = 0.814 2)。针对 HRNet 模型,通过对比是否使用冻结-解冻机制,证明了基于迁移学习的冻结-解冻机制能够提高模型的精度,加快训练效率,原因在于该机制能够使主干权值不过分随机,从而提高特征提取的效果。4 种模型使用 Adam 优化器的训练结果要优于使用 SGD 优化器的训练结果,原因在于 Adam 通过结合一阶动量和二阶动量来自适应学习率,解决了 SGD 优化器下降速度慢、容易取得局部最优解的问题^[29]。

2.3 小麦小穗分割效果比较

在采用最优主干网络和优化器为 Adam 的情况下,使用麦穗测试集对比分析基于冻结-解冻机制的 HRNet 模型、U-Net 模型、PSPNet 模型及 DeeplabV3+分割模型的分割能力,分割效果如图 11 所示。从分割效果看,对于输入的麦穗图像,PSPNet 模型、DeeplabV3+分割模型的分割效果不是很理想,虽然能够划分小麦小穗和背景,但是小穗之间粘连严重,这是因为小穗之间的灰度特征相似,像素值相近,容易被划分为一个整体,而背景灰度值与麦穗相差较远,容易被划分出来^[27-28]。U-Net 模型的分割效果虽然好一些,但在复杂环境下丢失了一些细节。相比之下,以 hrnetv2_w32 为主干网络、Adam 为优化器的基于冻结-解冻机制的 HRNet 模型对于麦穗图像的分割能力较强,不易受其他噪声影响,能够准确分割出小穗并计算小穗的数量和面积。

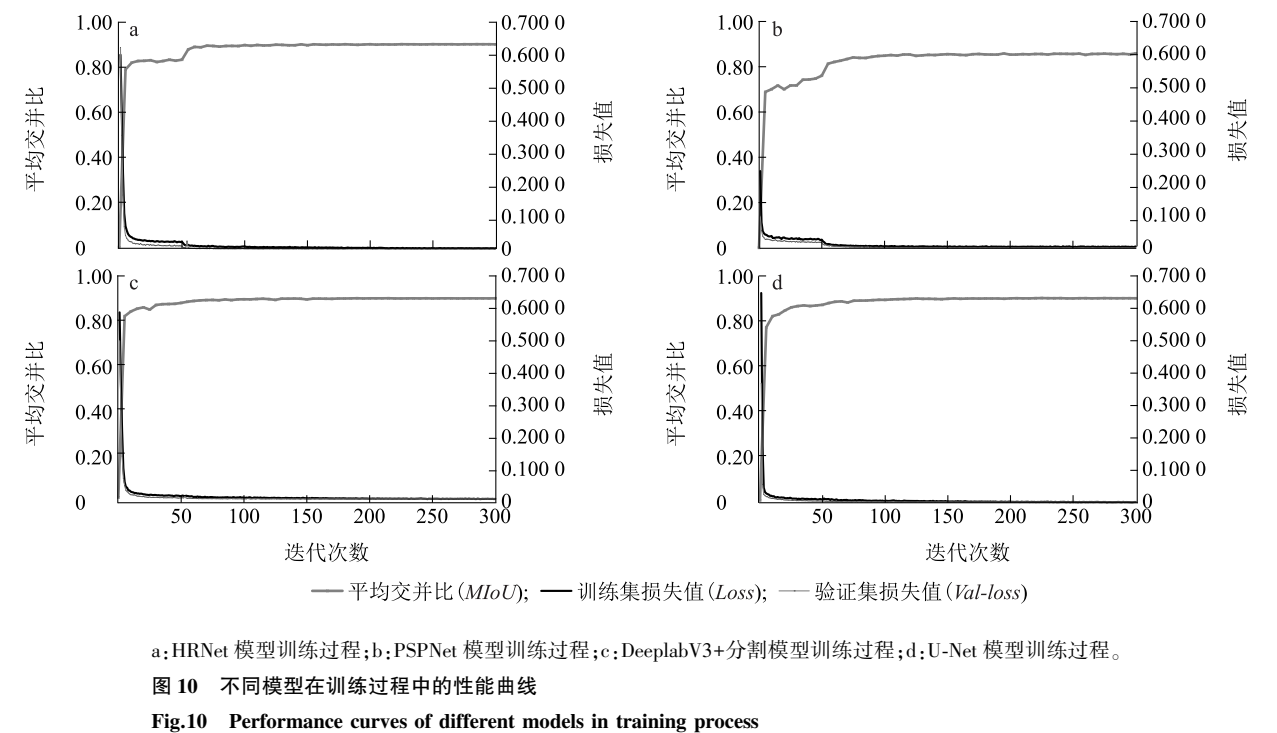


表 2 不同分割模型的精度对比与评价

Table 2 Accuracy comparison and evaluation of different segmentation models

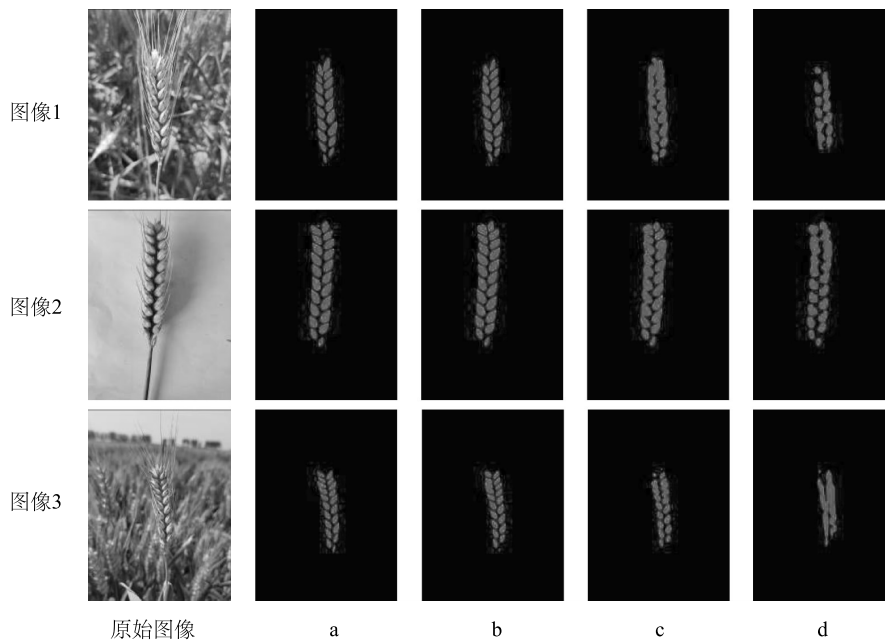
模型	主干网络	优化器	学习率	权值衰减	召回率	精确度	平均交并比	类别平均像素准确率
HRNet	hrnetv2_w32	Adam	0.000 5	0	0.941 9	0.959 4	0.911 9	0.941 9
	hrnetv2_w18	Adam	0.000 5	0	0.940 9	0.949 1	0.900 9	0.940 9
	hrnetv2_w32	SGD	0.004 0	0.000 1	0.930 0	0.946 2	0.888 6	0.930 0
	hrnetv2_w18	SGD	0.004 0	0.000 1	0.927 1	0.942 6	0.883 5	0.927 1
HRNet(无冻结-解冻机制)	hrnetv2_w32	Adam	0.000 5	0	0.941 4	0.950 9	0.901 9	0.941 4
	hrnetv2_w18	Adam	0.000 5	0	0.940 0	0.948 3	0.900 0	0.940 0
	hrnetv2_w32	SGD	0.004 0	0.000 1	0.925 1	0.945 9	0.880 0	0.925 1
	hrnetv2_w18	SGD	0.004 0	0.000 1	0.926 0	0.940 8	0.875 7	0.926 0
PSPNet	MobileNetv2	Adam	0.000 5	0	0.925 3	0.926 1	0.869 2	0.925 3
	ResNet50	Adam	0.000 5	0	0.930 1	0.929 2	0.875 3	0.930 1
	MobileNetv2	SGD	0.010 0	0.000 1	0.865 0	0.914 2	0.814 2	0.865 0
	ResNet50	SGD	0.010 0	0.000 1	0.914 1	0.923 6	0.858 7	0.914 1
DeeplabV3+	MobileNetv2	Adam	0.000 5	0	0.941 5	0.947 2	0.899 0	0.941 5
	Xception	Adam	0.000 5	0	0.938 7	0.946 7	0.896 3	0.938 7
	MobileNetv2	SGD	0.007 0	0.000 1	0.923 8	0.922 6	0.865 3	0.923 8
	Xception	SGD	0.007 0	0.000 1	0.928 3	0.919 9	0.866 8	0.928 3
U-Net	ResNet50	Adam	0.000 1	0	0.940 6	0.949 8	0.900 4	0.940 6
	VGG	Adam	0.000 1	0	0.941 9	0.949 5	0.901 3	0.941 9
	ResNet50	SGD	0.010 0	0.000 1	0.922 8	0.934 0	0.873 4	0.922 8
	VGG	SGD	0.010 0	0.000 1	0.931 0	0.942 7	0.886 7	0.931 0

加粗的数据表示最佳结果。

2.4 麦穗粒数的计数精度分析与评价

每个小麦品种选取 30 张样本图像,采用基于冻结-解冻机制的 HRNet 模型和图像处理算法对小麦样本图像中的小穗进行分割与计数,再根据划定的面积像素数阈值计算各个小穗的籽粒数、麦穗总籽粒数,并与人工计数法所得结果进行对比分析,可得小穗粒数预测模型的计数结果(图 12)及小穗、籽粒的计数精度评价结果。由图 12、表 3 可以看出,3 个小麦品种麦穗图像中小穗计数的均方根误差(RMSE)为 0.63,平均绝对误差为

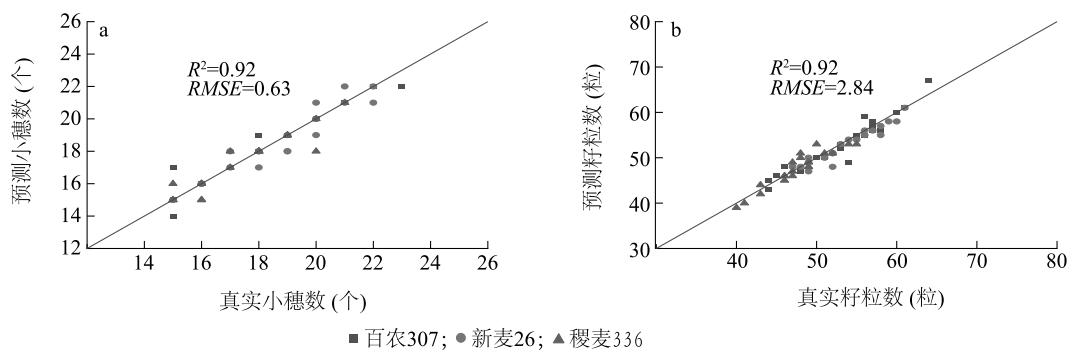
0.73,平均相对误差为 2.89%,决定系数(R^2)为 0.92;籽粒计数的 RMSE 为 2.84,平均绝对误差为 0.43,平均相对误差为 5.51%, R^2 为 0.92。由研究结果可以看出,基于冻结-解冻机制的 HRNet 模型和图像处理算法相较于传统图像处理算法,对小穗、籽粒的计数准确度较高,MAE、MRE 都更低,预测值与真实值的拟合程度较好。因此本方法适用于不同品种的麦穗,能够显著提高小穗数的计数精度,进而实现单株麦穗籽粒数的高精度、自动化计数^[32]。



a: HRNet 模型; b: U-Net 模型; c: DeeplabV3+ 分割模型; d: PSPNet 模型。

图 11 不同模型在测试集上的分割效果

Fig.11 Segmentation effect of different models on the test dataset



a: 小穗计数结果; b: 籽粒计数结果。RMSE: 均方根误差; R^2 : 决定系数。

图 12 小穗粒数预测模型的计数结果分析

Fig.12 Analysis of counting results of grain number prediction model

表 3 基于小穗粒数预测模型的小穗数和籽粒数计数精度评价

Table 3 Evaluation of counting accuracy of spikelets and grains based on spikelet-grain number prediction model

小麦品种	小穗计数精度				籽粒计数精度			
	平均绝对误差	平均相对误差 (%)	RMSE	R^2	平均绝对误差	平均相对误差 (%)	RMSE	R^2
百农 307	0.50	2.71	0.63	0.93	0.38	5.34	1.73	0.91
新麦 26	0.50	2.64	0.63	0.90	0.35	5.52	1.92	0.88
稷麦 336	1.20	3.31	0.67	0.88	0.56	5.68	2.11	0.86

RMSE:均方根误差; R^2 :决定系数。

3 讨论

本研究结果表明,基于冻结-解冻机制的 HRNet 模型和图像处理算法得到的小麦小穗数、穗粒数与人工计数结果基本一致。Alkhubaydi 等^[33]的研究结果表明,在小麦灌浆后期可以获得质量较好的图像,有利于提高深度学习分割模型的精度。由本研究结果可知,分割和计数结果在不同小麦品种间的表现基本一致。麦穗图像的采集时间是 2022 年 5 月 15 日、6 月 8 日,在数据量一般的条件下,基于冻结-解冻机制的 HRNet 模型不仅取得了良好的分割效果,而且在其他日期采集的图像上的表现突出^[5]。研究发现,在晴朗多云的环境下,将移动设备平行于麦穗并相距 10 cm 左右拍摄的图像能够获得更好的效果,因为该距离能够清晰地获取麦穗表型细节,并有助于实际应用。相较于采用互补金属氧化物半导体 (CMOS) 工业相机^[7]、无人机等进行图像采集,使用移动设备采集图像的实用性更好,适用性更强,更方便快捷。赵三琴等^[5]率先确定了稻穗结构图像特征与籽粒数间的相关关系,杜世伟等^[7]将这种方法引入小麦研究中,证明通过小穗像素数间接获取穗粒数的可行性。小穗粒数预测模型通过确定小穗像素数与穗粒数间的关系阈值来预测每个小穗对应的粒数,以深度学习分割模型预测的结果作为输入。相较于以处理后的 RGB 图像作为输入,这种方法能够更快速、准确地计算小穗像素数,从而精确预测穗粒数。

总的来说,通过在相同数据集中对比同样采用冻结-解冻机制的 PSPNet 模型、DeeplabV3+分割模型、U-Net 模型及无冻结-解冻机制的 HRNet 模型,提出了基于冻结-解冻机制的 HRNet 模型,该模型在灌浆期小麦图像中小穗和穗粒的分割方面表现出了较好的性能 ($mIoU=0.9119$)。利用图像处理算

法对预测图像进行各个小穗的分割,计算小穗的数量和像素数,并根据粒数预测模型,由各小穗的像素数得到籽粒数,进而获得整个麦穗的籽粒数。相较于传统图像处理算法^[12],本研究方法极大提高了小穗数的识别精度,能够实现单株麦穗籽粒进行较高精度的自动计数^[34]。

本研究结果有助于开发一种低成本、快速且易于实施的小麦小穗和穗粒计数系统。目前的研究主要通过固定拍摄高度^[35]或设置参考物和地面标准^[36]等措施来降低误差,这在一定程度上影响了该方法的实用性,未来使用无人机低空采集图像来进行小麦小穗、穗粒的计数是研究目标之一。

4 结论

为了对小麦图像中的穗粒进行更加精确的分割与计数,本研究通过获取小麦材料和数据,提出一种基于冻结-解冻机制的 HRNet 模型,结合图像处理算法,实现了小麦穗粒的精确计数,主要结论如下: 1) 针对麦穗图像语义信息比较复杂,图像中的小穗及穗粒之间粘连覆盖现象严重等问题,提出在原始 HRNet 模型基础上,根据迁移学习思想,增加冻结-解冻机制,以提高特征提取效率,防止权值太过随机,从而加快训练效率。基于冻结-解冻机制的 HRNet 模型预测出的效果明显优于对比组的网络模型,证明该模型鲁棒性较好。2) 使用基于冻结-解冻机制的 HRNet 模型对麦穗数据集图像进行训练、验证和测试,用 $mIoU$ 等指标对分割结果的准确性及分割效果进行评估。结果显示,小麦穗粒的分割精确度为 0.959 4, $mIoU$ 值为 0.911 9, mPA 值为 0.941 9,召回率为 0.941 9,与同样采用冻结-解冻机制的 PSPNet 模型、DeeplabV3+分割模型、U-Net 模型及无冻结-解冻机制的 HRNet 模型相比,分割精确度有所提升,表现出较强的泛化能力。3) 本研究

从小麦小穗图像入手,采用图像处理算法确定小穗像素数与穗粒数的关系阈值,根据农学知识找出麦穗总粒数的计算方法,通过计算人工计数结果和预测值的拟合程度和误差,发现相较于传统图像处理算法,本研究方法极大提高了小穗数的计数精度,能够实现对单株麦穗籽粒进行较高精度的自动计数。

4) 本研究方法可用于估算小麦小穗数和穗粒数,提高小麦产量估算效率,同时还可为农业工作者提供快速、自动化的高通量小麦小穗、穗粒计数系统,提高工作效率。本研究方法同样适用于小麦图像中小穗、穗粒的分割和计数,也可以用于其他植物图像中的分割和计数。在未来的工作中,预计目标是降低图像采集成本,提高该方法的计数精度和应用范围。由于本方法是间接获取麦穗籽粒数,后续需要探索新的方法以消除间接误差。

参考文献:

- [1] 张领先,陈运强,李云霞,等. 基于卷积神经网络的冬小麦麦穗检测计数系统[J]. 农业机械学报,2019,50(3):144-150.
- [2] LI Q Y, CAI J, BERGER B, et al. Detecting spikes of wheat plants using neural networks with laws texture energy[J]. Plant Methods,2017,13(1):1-13.
- [3] HU W J, FAN J, DU Y X, et al. MDFC-ResNet: an agricultural IoT system to accurately recognize crop diseases[J]. IEEE Access,2020,8:115287-115298.
- [4] MISRA T, ARORA A, MARWAHA S, et al. SpikeSegNet-a deep learning approach utilizing encoder-decoder network with hourglass for spike segmentation and counting in wheat plant from visual imaging[J]. Plant Methods,2020,16(1):1-20.
- [5] 赵三琴,李毅念,丁为民,等. 稻穗结构图像特征与籽粒数相关关系分析[J]. 农业机械学报,2014,45(12):323-328.
- [6] 王 栋,陈佳玮,沈利言,等. 基于图像的水稻穗粒相关性状智能检测算法研究[J]. 植物生理学报,2022,58(5):957-971.
- [7] 杜世伟,李毅念,姚 敏,等. 基于小麦穗部小穗图像分割的籽粒计数方法[J]. 南京农业大学学报,2018,41(4):742-751.
- [8] QIU R, HE Y, ZHANG M. Automatic detection and counting of wheat spikelet using semi-automatic labeling and deep learning[J]. Frontiers in Plant Science,2022,13:872555.
- [9] FERNANDEZ-GALLEGO J A, KEFAUVER S C, GUTIÉRREZ N A, et al. Wheat ear counting in-field conditions: high throughput and low-cost approach using RGB images[J]. Plant Methods,2018,14(1):1-12.
- [10] 刘 哲,黄文准,王利平. 基于改进 K-means 聚类算法的大田麦穗自动计数[J]. 农业工程学报,2019,35(3):174-181.
- [11] XU X, LI H Y, YIN F, et al. Wheat ear counting using K-means clustering segmentation and convolutional neural network[J]. Plant Methods,2020,16:106.
- [12] 李毅念,杜世伟,姚 敏,等. 基于小麦群体图像的田间麦穗计数及产量预测方法[J]. 农业工程学报,2018,34(21):185-194.
- [13] WEI W, YANG T L, RUI L, et al. Detection and enumeration of wheat grains based on a deep learning method under various scenarios and scales[J]. Journal of Integrative Agriculture,2020,19(8):1998-2008.
- [14] HU G, QIAN L, LIANG D, et al. Self-adversarial training and attention for multi-task wheat phenotyping[J]. Applied Engineering in Agriculture,2019,35(6):1009-1014.
- [15] DANDRIFOSSE S, ENNADIFI E, CARLIER A, et al. Deep learning for wheat ear segmentation and ear density measurement: from heading to maturity[J]. Computers and Electronics in Agriculture,2022,199:107161.
- [16] ZHAO J, ZHANG X, YAN J, et al. A wheat spike detection method in UAV images based on improved YOLOv5[J]. Remote Sensing,2021,13(16):3095.
- [17] MISRA T, ARORA A, MARWAHA S, et al. Web-SpikeSegNet: deep learning framework for recognition and counting of spikes from visual images of wheat plants[J]. IEEE Access,2021,9:76235-76247.
- [18] LI J, LI C, FEI S, et al. Wheat ear recognition based on RetinaNet and transfer learning[J]. Sensors,2021,21(14):4845. DOI:10.3390/s21144845.
- [19] 段凌凤,熊 雄,刘 谦,等. 基于深度全卷积神经网络的大田稻穗分割[J]. 农业工程学报,2018,34(12):202-209.
- [20] WANG J D, SUN K, CHENG T H, et al. Deep high-resolution representation learning for visual recognition[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence,2020,43(10):3349-3364.
- [21] ZHAO H, SHI J, QI X, et al. Pyramid scene parsing network[C]. Honolulu, HI, USA:IEEE,2017:2881-2890.
- [22] CHEN L C, ZHU Y, PAPANDREOU G, et al. Encoder-decoder with atrous separable convolution for semantic image segmentation[C]. Berlin:Springer,2018:801-818.
- [23] 邱云飞,温金燕. 基于 DeepLabV3+与注意力机制相结合的图像语义分割[J]. 激光与光电子学进展,2022,59(4):130-139.
- [24] 马冬梅,李鹏辉,黄欣悦,等. 改进 DeepLabV3+的高效语义分割[J]. 计算机工程与科学,2022,44(4):737-745.
- [25] RONNEBERGER O, FISCHER P, BROX T. U-net:convolutional networks for biomedical image segmentation[C]. Berlin:Springer International Publishing,2015:234-241.
- [26] 梁 波,卢 军,曹 阳. 基于改进 U-Net 卷积神经网络的钢轨表面损伤检测方法[J]. 激光与光电子学进展,2021,58(2):334-340.
- [27] MA J C, LI Y X, DU K M, et al. Segmenting ears of winter wheat at flowering stage using digital images and deep learning[J]. Computers and Electronics in Agriculture,2020,168:105159.
- [28] ALHARBI N, ZHOU J, WANG W. Automatic counting of wheat spikes from wheat growth images[C]. Funchal, Madeira, Portu-

- gal;SciTePress,2018;346-355.
- [29] FERNANDEZ-GALLEGO J A, KEFAUVER S C, GUTIÉRREZ N A, et al. Wheat ear counting in-field conditions: high throughput and low-cost approach using RGB images [J]. *Plant Methods*, 2018,14(1):22.
- [30] KHAKI S, SAFAEI N, PHAM H, et al. Wheatnet: a lightweight convolutional neural network for high-throughput image-based wheat head detection and counting [J]. *Neurocomputing*, 2022, 489:78-89.
- [31] WANG Y, QIN Y, CUI J. Occlusion robust wheat ear counting algorithm based on deep learning [J]. *Frontiers in Plant Science*, 2021,12:645899. DOI:10.3389/fpls.2021.645899.
- [32] MAJI A K, MARWAHA S, KUMAR S, et al. SypNet: spikelet-based yield prediction of wheat using advanced plant phenotyping and computer vision techniques [J]. *Frontiers in Plant Science*, 2022,13:889853.
- [33] ALKHUDAYDI T, ZHOU J, DE LA LGLESIA B. Counting spikelets from infield wheat crop images using fully convolutional networks [J]. *Neural Computing and Applications*, 2022, 34 (20): 17539-17560.
- [34] XU C, JIANG H Y, YUEN P, et al. MHW-PD: a robust rice panicles counting algorithm based on deep learning and multi-scale hybrid window [J]. *Computers and Electronics in Agriculture*, 2020,173:105375.
- [35] HASAN M M, CHOPIN J P, LAGA H, et al. Detection and analysis of wheat spikes using convolutional neural networks [J]. *Plant Methods*, 2018,14(1):1-13.
- [36] SADEGHI-TEHRAN P, VIRLET N, AMPE E M, et al. Deep-Count: in-field automatic quantification of wheat spikes using simple linear iterative clustering and deep convolutional neural networks [J]. *Frontiers in Plant Science*, 2019,10:1176.

(责任编辑:徐 艳)