doi:10.6041/j.issn.1000-1298.2024.S1.008

基于MobileNetV2-CBAM的机收场景下冬小麦成熟期在线分类识别方法

王发明[1]　倪昕东[1]　张　旗[1]　陶　伟[2]　陈　度[1,3]　毛　旭[1,3]

（1. 中国农业大学工学院，北京 100083；2. 博创联动科技股份有限公司，北京 100083；3. 智能农业动力装备全国重点实验室，北京 100083）

摘要：小麦成熟期在线精准分类识别将为实现联合收获机的智能化调控提供有效支撑。本文提出一种基于车载相机和深度学习结合的冬小麦成熟期在线分类方法。以车载相机拍摄的实时图像为主，无人机拍摄的图像为辅，构建小麦乳熟-蜡熟初期、蜡熟后期-完熟初期、完熟后期-枯熟期和已收割区数据集（4 400幅）。针对机收环境复杂、小麦图像模糊等问题，以MobileNetV2为基础网络结构，在特征提取后添加卷积注意力模块（Convolutional block attention module, CBAM）提升对图像特征的自适应提取能力。为了评估模型可信度，采用可视化技术观察模型对图像的关注区域。以不同分类模型为对比，对建立的MobileNetV2-CBAM模型性能进行评价。试验结果表明，MobileNetV2-CBAM模型在测试集中的分类识别准确率达到99.5%，相比于MobileNetV2高0.7个百分点；与ResNet和Swin Transformer模型相比，在分类精度未发生明显差异的前提下，MobileNetV2-CBAM模型内存占用量（8.73 MB）仅为其1/8和1/11。为了验证模型实际应用效果，田间试验结果表明，在车速4~6 km/h条件下，每隔1 s识别1幅图像，成熟期分类识别精度为96.8%，满足机收场景下的小麦成熟期在线分类准确性和实时性要求。

关键词：小麦；成熟期；MobileNetV2-CBAM；深度学习；车载相机

中图分类号：TP391　　**文献标识码：**A　　**文章编号：**1000-1298(2024)S1-0071-10

Online Classification and Identification Method of Winter Wheat Maturity under Mechanical Harvesting Scenario Based on MobileNetV2-CBAM

WANG Faming[1]　NI Xindong[1]　ZHANG Qi[1]　TAO Wei[2]　CHEN Du[1,3]　MAO Xu[1,3]

(1. *College of Engineering, China Agricultural University, Beijing* 100083, *China*

2. *Universal Mobile Linking Technology Co.,Ltd., Beijing* 100083, *China*

3. *State Key Laboratory of Intelligent Agricultural Power Equipment, Beijing* 100083, *China*)

Abstract: The precise online classification and identification of wheat maturity stages will offer valuable support for the intelligent control of combine harvesters. An online classification method was proposed for wheat maturity stages that combined vehicle-mounted cameras with deep learning techniques. By using real-time images captured by vehicle-mounted cameras, along with additional images from drones, a dataset of 4 400 images was constructed, which included various wheat maturity stages, including milk ripening-early wax ripening stage, late wax ripening-early full ripening stage, late full ripening-dry ripening stage and harvested area. To address challenges such as complex harvesting environments and blurry wheat images, the MobileNetV2 was employed as the foundational network structure. Additionally, a convolutional block attention module (CBAM) was incorporated after feature extraction to enhance the adaptive capability of image feature extraction. To assess the credibility of the model, visualization techniques were employed to examine the areas of interest identified by the model in the images. The performance of the MobileNetV2-CBAM model was compared with other classification models. Results indicated that the MobileNetV2-CBAM model achieved a classification accuracy of 99.5% on the test set, which was 0.7 percentage points higher than that of MobileNetV2. When compared with ResNet and Swin Transformer models, the MobileNetV2-CBAM model demonstrated similar

收稿日期：2024-07-18　修回日期：2024-09-20

基金项目：中央高校基本科研业务费专项资金项目(2024TC183)和中国农业大学横向课题项目(202405410710092)

作者简介：王发明(1996—)，男，博士生，主要从事智能农机装备、计算机测控技术和机器视觉研究，E-mail: wfm@cau.edu.cn

通信作者：陈度(1981—)，男，教授级高级工程师，博士生导师，主要从事农业自动化及测控技术研究，E-mail: tchendu@cau.edu.cn

classification accuracy but with a significantly smaller model memory usage (8.73 MB)—only 1/8 and 1/11 of the memory usage of ResNet and Swin Transformer, respectively. Field experiments further validated the model's practical application: at vehicle speeds of 4 km/h to 6 km/h, the system recognized an image every second with a maturity classification accuracy of 96.8%, meeting the accuracy and real-time requirements for online wheat maturity classification in mechanical harvesting scenarios.

Key words: wheat; maturity; MobileNetV2-CBAM; deep learning; vehicle-mounted camera

0 引言

我国小麦收获环节已基本实现机械化,机械化作业质量将影响机收小麦的质量和粮食储藏品质。机收场景下,小麦秸秆和籽粒含水率是影响联合收获机作业效率和作业质量的重要指标之一[1-2],对联合收获机的切割喂入、输送、脱粒分离及清选过程的作业性能均会产生较大影响[3-5]。小麦植株含水率和不同成熟期阶段密切相关,通过实现小麦成熟期在线准确分类,预测联合收获机前方小麦植株的含水率,进而将为联合收获机的智能化调控提供重要支撑。现阶段,小麦成熟期判别主要包括烘干[6-9]、遥感[10-12]、气象预测[13-15]等方法,对指导宏观生产具有重要意义。但对于微观层面,存在空间和时间分辨率不足、样本处理时间长等问题。

近年来,机器视觉、深度学习等新兴技术在农业领域得到广泛应用,目标检测和图像分割方法已经成功应用于作物计数和产量预测[16-17],国内外学者将深度学习模型部署于农机装备车载终端,开展了倒伏面积在线提取[18]、稻麦高度在线测量[19]以及玉米籽粒杂质在线检测[20]等研究工作。但目标检测和图像分割关注的是图像中像素的归属,需要理解检测物的详细信息,由于小麦种植密度高以及受大田复杂环境影响,在收获过程中利用目标检测和图像分割算法对小麦成熟期进行在线分类存在一定难度。

相较于目标检测和图像分割方法,图像分类算法关注小麦整体生长状态的变化,不需要详细的麦穗和秸秆轮廓信息,因此能够更好适应机收场景下的小麦成熟期在线分类识别。国内外研究学者将图像分类算法应用于大田作物的生育期分类[21-23],已取得一定进展,但对于作物收获期成熟状态特征的细分研究尚不充分,并且基于图像分类算法识别机收场景下小麦成熟期的研究未见报道。

针对上述需求,本文提出基于改进MobileNetV2网络的机收场景下小麦成熟期在线分类识别方法,构建小麦不同成熟期数据集,研究改进MobileNetV2的轻量级图像分类模型,以期实现小麦成熟期在线精准分类。

1 材料与方法

1.1 成熟期类别

冬小麦生育期可以分为12个阶段,分别为播种期、出苗期、分蘖期、越冬期、返青期、起身期、拔节期、挑旗期、抽穗期、开花期、灌浆期和成熟期[24],根据作物含水率、籽粒饱满程度不同,成熟期又可分为乳熟期、蜡熟期、完熟期以及枯熟期4个阶段[25],冬小麦部分生育期及特征如图1所示。

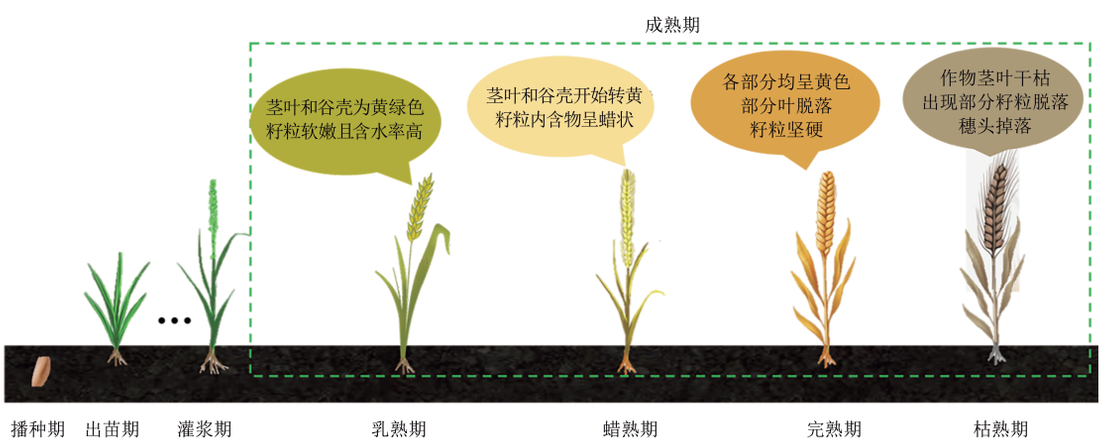


图1 冬小麦部分生育期及特征示意图

Fig. 1 Schematic of partial growth stages and characteristics of winter wheat

小麦最佳机收期集中在蜡熟后期-完熟初期,这一阶段的小麦植株含水率相近,品质最好,因此本文将蜡熟后期-完熟初期定义为一类。小麦乳熟-蜡熟初期,籽粒含水率较高,未达到收割条件,因此将乳熟-蜡熟初期定义为一类。完熟后期-枯熟期为非最佳机收期,故定义为一类。同时,小麦成熟期在线识别将为实现联合收获机智

能化调控提供决策依据，作业过程中收获机满仓后需要经过已收割区域前往地头卸粮，在已收割区域行驶时需要断开割台-过桥-滚筒主离合并提高行驶速度，因此，已收割区可作为小麦成熟期的泛化类别，丰富数据集多样性。建立4个分类的小麦成熟期数据集，如表1所示。

1.2　田间图像采集

为了增加小麦成熟期数据集的多样性，选择多个小麦产区典型品种作为数据集构建对象。由于乳熟-蜡熟初期小麦含水率过高，无法采集车载相机（型号为ZED2i）拍摄的数据集，同时为了扩充其他成熟期数据集，借助大疆御MAVIC AIR 2型无人机模拟收获机视角进行拍摄，构建以车载相机拍摄图像为主，无人机拍摄图像为辅的小麦成熟期数据集，详见表2；为了提高模型对不同视角、高度和环境条件下图像适应性，无人机镜头距地面高度分别设置为1、3、5 m，镜头与水平地面夹角设置为30°，飞行速度设置为1.5 m/s和3 m/s。车载相机通过云台安装于联合收获机驾驶室顶部（图2）。

表1　小麦成熟期分类

Tab. 1　Wheat maturity classification

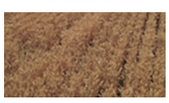
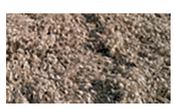
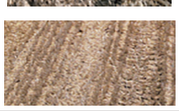
分类示意图	分类标签	特征
	乳熟-蜡熟初期	颜色为绿/黄绿色，含水率较高
	蜡熟后期-完熟初期	茎叶和谷壳为黄色，籽粒饱满
	完熟后期-枯熟期	茎叶和谷壳为枯黄色，麦穗干枯
	已收割区	无麦穗，剩麦茬

表2　图像数据采集详情

Tab. 2　Image data acquisition details

采集地点	采集时间	采集工具	分辨率/（像素×像素）	小麦品种	成熟期
北京市平谷区	2023-05-29	无人机	4 000×2 250	京麦183	乳熟-蜡熟初期
	2023-06-16	车载相机	22 08×1 242;1 920×1 080	京麦183	蜡熟后期-完熟初期
		无人机	4 000×2 250		已收割区
江苏省新曹农场	2023-06-06—06-10	车载相机	2 208×1 242;1 920×1 080	镇麦12号	完熟后期-枯熟期
		无人机	4 000×2 250		已收割区
江苏省东辛农场	2023-06-11—06-12	车载相机	2 208×1 242;1 920×1 080	淮麦33	蜡熟后期-完熟初期
		无人机	4 000×2 250		已收割区

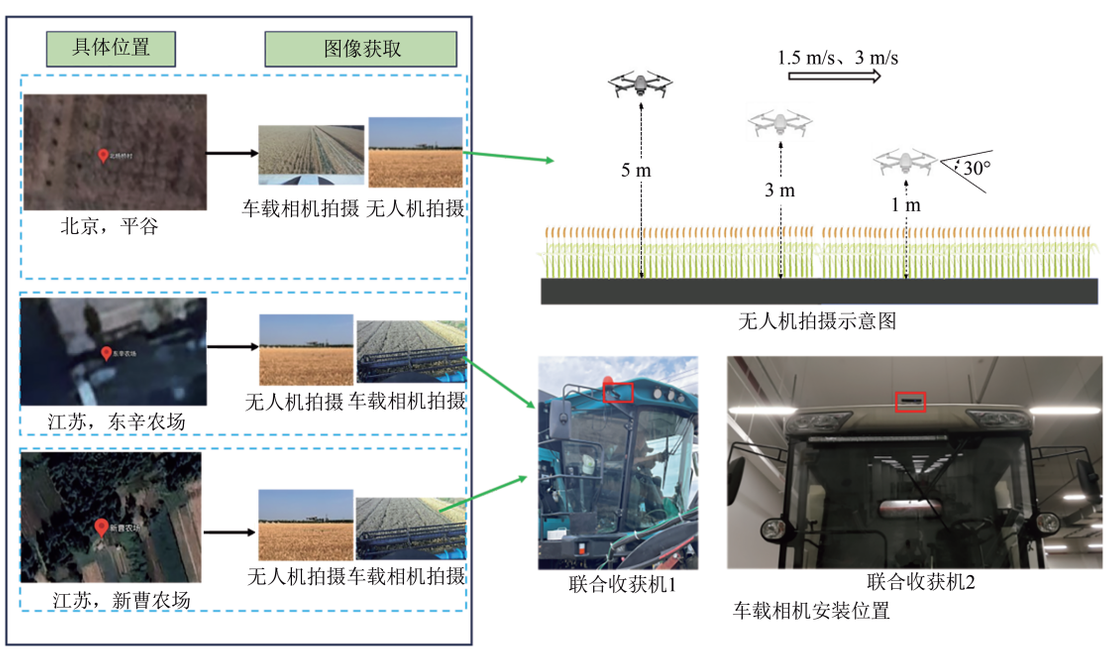


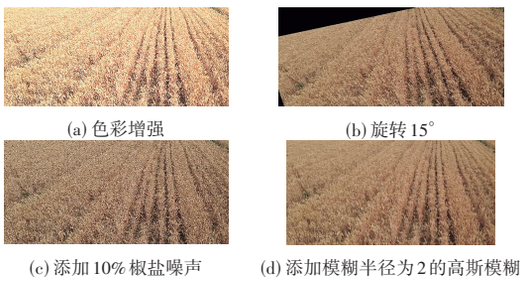
图2　数据采集方式

Fig. 2　Data collection method

1.3　数据预处理

联合收获机在作业过程中产生振动会使相机出现一定抖动，造成图像采集角度发生改变，图像讯号受到干扰产生噪声。同时，机收环境会产生较大尘土，可能会造成图像模糊不清，降低模型泛化能力。丰富图像数据集有助于减少网络训练中的过拟合，提高模型泛化能力[26]。采用数据裁剪方法，将来自不同采集源的图像尺寸统一调为224×224×3，然后对裁剪后图像分别进行色

彩增强、角度旋转、椒盐噪声、高斯模糊数据增强等处理，数据增强后图像如图3所示。



(a) 色彩增强 (b) 旋转15°

(c) 添加10%椒盐噪声 (d) 添加模糊半径为2的高斯模糊

图3 数据增强样本

Fig. 3 Data augmentation samples

本文制作的小麦成熟期数据集共包含4 400幅图像。首先，选取2 000幅原始图像，其中每个小麦成熟期类别各500幅，并从中随机选取100幅作为测试集，共计400幅。剩余1 600幅图像经过数据增强处理，用于构建训练集和验证集，训练集-验证集制作规则如表3所示。每类标签下原始图像为400幅，经过4种数据增强方法后，每类标签下图像为2 000幅，其中包括：原始图像400幅、色彩增强400幅、角度旋转400幅、椒盐噪声400幅、高斯模糊400幅，然后从每个数据集中随机选取200幅，因此每类标签下图像共选取1 000幅，4类标签共选取4 000幅，将其按比例8∶2划分为训练集和验证集。

表3 训练集-验证集制作规则

Tab. 3 Rules for making training set and validation set 幅

项目	乳熟-蜡熟初期	蜡熟后期-完熟初期	完熟后期-枯熟期	已收割区	总计
原始图像	400 (随机选取200)	400 (随机选取200)	400 (随机选取200)	400 (随机选取200)	1 600 (随机选取800)
色彩增强	400 (随机选取200)	400 (随机选取200)	400 (随机选取200)	400 (随机选取200)	1 600 (随机选取800)
角度旋转	400 (随机选取200)	400 (随机选取200)	400 (随机选取200)	400 (随机选取200)	1 600 (随机选取800)
椒盐噪声	400 (随机选取200)	400 (随机选取200)	400 (随机选取200)	400 (随机选取200)	1 600 (随机选取800)
高斯模糊	400 (随机选取200)	400 (随机选取200)	400 (随机选取200)	400 (随机选取200)	1 600 (随机选取800)
总计	2 000 (随机选取1 000)	2 000 (随机选取1 000)	2 000 (随机选取1 000)	2 000 (随机选取1 000)	8 000 (随机选取4 000)

1.4 网络结构设计

1.4.1 MobileNetV2-CBAM总体架构

小麦不同成熟期的图像特征(颜色、麦穗纹理等)存在明显差异，利用图像分类算法进行在线分类识别。MobileNetV2是一种在移动设备和嵌入式系统上高效运行的轻量级神经网络架构，采用深度可分离卷积、倒残差连接和线性瓶颈等关键技术[27]，可有效减少网络参数量和计算复杂度，因此采用MobileNetV2网络模型作为骨干网络。

虽然MobileNetV2在目标识别及图像分类领域表现出了较好效果，但其对图像特征的自适应提取能力相对较弱，特别是在大田机收场景中，视频设备捕捉的图像质量会受到机体振动、光照以及灰尘影响，导致图像包含噪声、清晰度降低。同时，小麦成熟期识别属于较大面积识别，在识别过程中极易出现误判现象。因此本文在MobileNetV2网络中引入卷积注意力机制模块(Convolutional block attention module，CBAM)[28]来解决这一问题，在关注图像中关键信息的同时，过滤无关信息。

对MobileNetV2网络模型整体框架进行改进(图4)，即在特征提取层后加入卷积注意力机制模块，扩大网络感知野，并将改进MobileNetV2模型命名为MobileNetV2-CBAM，其具体结构参数如表4所示。模型在识别过程中，首先输入224×224×3成熟期图，经过一个由3×3卷积层、批归一化(Batch normalization，BN)层和ReLU6组成的下采样层，图像尺寸变为原尺寸一半，特征通道数变为32；随后经过17个瓶颈(Bottleneck)层，包括4个步长为2下采样层，图像尺寸变为7×7，特征通道数变为320。将7×7×320图像输入1×1卷积层，通道数变为1 280，随后经过CBAM层，输出7×7×1 280，最后经过一个平均池化层和全连接层输出预测成熟期种类。CBAM具体操作流程如下所述。

1.4.2 卷积注意力机制模块

CBAM包括2个独立子模块：通道注意力机制(Channel attention module，CAM)和空间注意力机制(Spatial attention module，SAM)，在通道和空间维度上同时引入注意力机制，使网络更加聚焦于图像中重要特征，从而提高深度卷积神经网络性能和泛化能力。相较于其他注意力模块，CBAM设计相对轻量化，易集成到不同卷积神经网络架构中，适用于嵌入式设备和资源受限的环境，并且已经成功应用于杂草以及病虫害识别等大田场景[29-30]。

(1)通道注意力

图5为通道注意力模块示意图，输入的小麦成熟期

特征图分别经过全局最大池化层和全局平均池化层，得到2个包含一个通道的特征图；然后将它们分别送入一个共享多层感知机（Multilayer perceptron，MLP），将MLP输出的特征进行加和操作，利用Sigmoid函数进行激活操作，生成最终的成熟期图像通道注意力特征图；最后，将特征图和初始输入特征图进行乘法操作，生成空间注意力模块需要的成熟期图像输入特征，即

$$\begin{cases} \boldsymbol{M}_c(F) = \sigma(MLP(\mathrm{GAvgPool}(F)) + MLP(\mathrm{GMaxPool}(F))) \\ F' = F \otimes \boldsymbol{M}_c(F) \end{cases} \tag{1}$$

式中　$\boldsymbol{M}_c(F)$——通道注意力的特征权重矩阵

σ——Sigmoid激活函数

MLP——多层感知机

F'——融合后输出

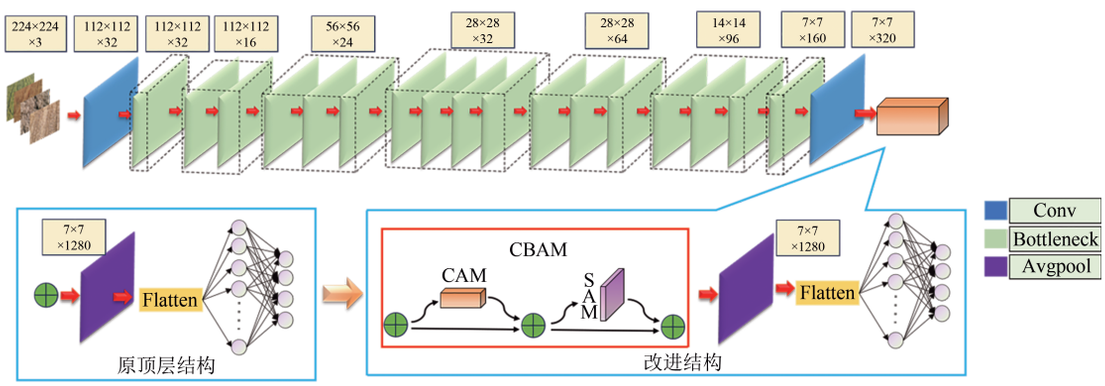


图4　MobileNetV2-CBAM模型总体框架图

Fig. 4　MobileNetV2-CBAM model overall framework diagram

表4　MobileNetV2-CBAM网络结构

Tab. 4　MobileNetV2-CBAM network structure

输入	操作	扩展因子	输出特征矩阵深度	重复次数	步距
$224^2\times3$	Conv2d	–	32	1	2
$112^2\times32$	Bottleneck	1	16	1	1
$112^2\times16$	Bottleneck	6	24	2	2
$56^2\times24$	Bottleneck	6	32	3	2
$28^2\times32$	Bottleneck	6	64	4	2
$28^2\times64$	Bottleneck	6	96	3	1
$14^2\times96$	Bottleneck	6	160	3	2
$7^2\times160$	Bottleneck	6	320	1	1
$7^2\times320$	Conv2d 1×1	–	1 280	1	1
$7^2\times1\ 280$	CBAM	–	1 280	1	1
$7^2\times1\ 280$	Avgpool 7×7	–	–	1	–
1×1×1 280	Conv2d 1×1	–	k	–	

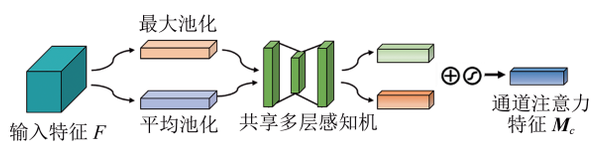


图5　通道注意力模块

Fig. 5　Convolutional block attention module

（2）空间注意力

图6为空间注意力模块示意图，式(2)给出空间注意力模块过程表达式：将通道注意力模块输出的特征图作为本模块的输入特征图。首先经过全局最大池化层和全局平均池化层，得到2个一通道的特征图；然后将这2个特征图做通道拼接操作(f_c)，经过一个7×7$K^{7\times7}$标准卷积操作，降维为1个通道，再经过Sigmoid函数生成最终的成熟期图像空间注意力特征图$M_s(F)$；最后将该特征图和该模块的原始输入特征做乘法，得到最终的成熟期图像输出CBAM特征F''，即

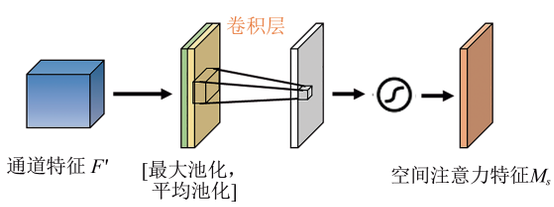


图6　空间注意力模块

Fig. 6　Spatial attention module

$$\begin{cases} M_s(F) = \sigma(K^{7\times7}(f_c[\mathrm{GAvgPool}(F);\ \mathrm{GMaxPool}(F)])) \\ F'' = F' \otimes M_s(F) \end{cases} \tag{2}$$

1.4.3　评价指标

在实验中，利用混淆矩阵作为主要评价标准来衡量模型性能。从混淆矩阵中计算准确率（Accuracy，A）、精确率（Precision，P）、召回率（Recall，R）和特异度（Specificity，S）4个指标，以评估模型训练效果。

2　结果与分析

2.1　训练与推理结果

使用的MobileNetV2-CBAM模型基于PyTorch 2.0.1深度学习模型框架构建，深度学习模型训练和推理的系统环境为Windows10，python版本为3.9，GPU加速库使用CUDA11.8；硬件环境为CPU i9-13900K、NVIDIA GeForce RTX 4080-16 GB、主板Z790 D DDR4。训练时采用随机梯度下降法，学习率设置为0.000 1，批处理大小为32，模型迭代次数为100，防止过度训练导致过拟合现象。

2.1.1　预训练模型训练结果

为验证所设计分类模型的识别性能，将提出的

MobileNetV2-CBAM 模型分别与经典 CNN 模型(VGG16[31]、ResNet[32])、轻量化 CNN 模型(MobileNetV2、ShufflenetV2[33]、MobileViT[34])、Transformer 模型(Swin Transformer[35])进行对比。在训练过程中引入迁移学习加速模型收敛,加载基于 ImageNet 数据集的预训练权重至各个模型骨干网络,再将权重值与训练样本特征元素进行相乘得到主干特征。

图 7a 为不同模型在训练集上的损失值,VGG16、ResNet 以及 Swin Transformer 在训练集上损失值从训练开始均保持在 0.3 以下,完成第 8 次迭代后,损失值均保持在 0.2 以下。与此不同,MobileViT 和 ShufflenetV2 收敛后损失值分别为 0.35 和 0.21,均高于 0.2,且轻量级网络 ShufflenetV2 在接近迭代 100 次后才出现收敛趋势。图 7b 为模型在验证集上的准确率,由图可知,MobileNetV2、Swin Transformer、MobileViT 以及 ResNet 模型在训练初期的准确率达到 95% 以上,这是由于引入迁移学习,加速了模型收敛。但在训练过程中,VGG16 模型准确率波动较大,这是由于经典模型结构较为复杂,输入的批处理大小较小造成曲线上下震荡[23]。

改进 MobileNetV2-CBAM 在迭代 14 次前损失值呈线性急速下降,准确率基本呈线性上升,完成 14 次迭代后,训练集损失值收敛至 0.01,验证集准确率达到 99.9%,并趋于稳定,这是由于顶层结构中添加了卷积注意力机制模块,改变了原有的分类器结构,训练初期需结合成熟期数据进行自适应训练,当训练至一定程度,损失值和准确率会呈现较好的结果。

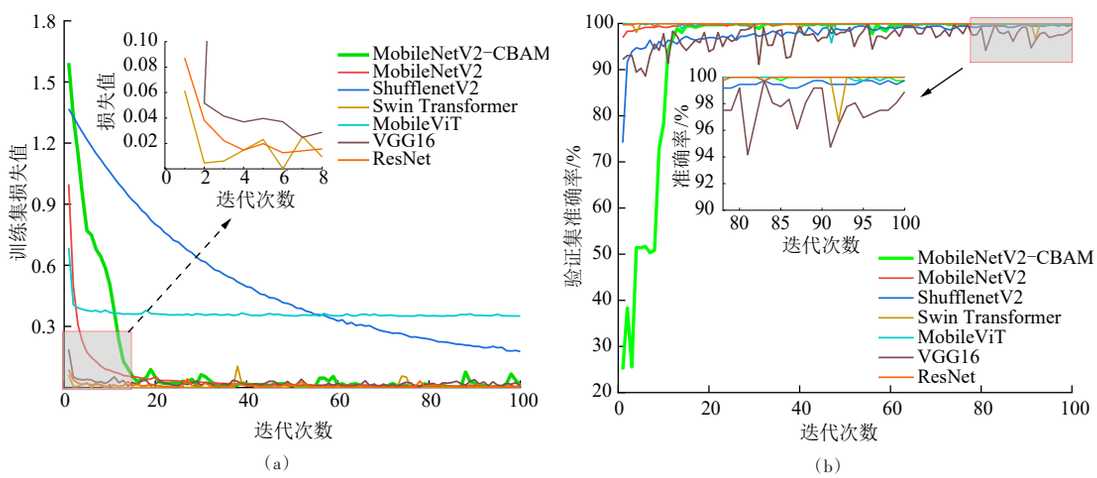


图 7 不同模型的准确率和损失值对比

Fig. 7 Comparison of accuracy and loss values of different models

2.1.2 评价结果

表 5 为 7 个分类模型在测试集上的准确率和模型内存占用量,分析模型内存占用量,只有 MobileViT、ShufflenetV2、MobileNetV2 和 MobileNetV2-CBAM 小于 10 MB,能够更好地实现边缘部署,同时说明添加 CBAM 后,模型内存占用量变化较小。测试集准确率有 3 种模型达到 99% 以上,其中 Swin Transformer 分类效果最好,达到 99.8%;MobileNetV2-CBAM 模型识别效果次之,准确率为 99.5%,相较于 MobileNetV2 模型提高 0.7 个百分点,这是由于引入 CBAM,提高了模型识别性能;经典网络中 ResNet 识别准确率为 99.2%。在 ResNet 和 Swin Transformer 的模型内存占用量分别为 MobileNetV2-CBAM 的 8 倍和 11 倍的情况下,三者预测准确率基本相同,因此综合考虑识别准确率和模型内存占用量,本文模型优于其他分类模型。

表 5 不同模型的分类结果

Tab. 5 Classification results of different models

模型	训练集损失值(收敛后)	验证集准确率(收敛后)/%	模型内存占用量/MB	测试集准确率/%
VGG16	0.02	97.6	512.00	97.2
ResNet	0.005	99.9	81.31	99.2
Swin Transformer	0.000 5	99.8	105.00	99.8
MobileViT	0.35	99.9	3.76	98.0
ShufflenetV2	0.21	99.5	4.96	94.2
MobileNetV2	0.007	99.9	8.73	98.8
MobileNetV2-CBAM(本文)	0.01	99.9	9.52	99.5

为了进一步评价模型分类性能,列出准确率前 4 位模型在测试集上的混淆矩阵,如图 8 所示。其中,L0 为乳熟-蜡熟初期,L1 为完熟后期-枯熟期,L2 为蜡熟后期-完熟初期,L3 为已收割区。行对应预测类别,列对

应真实类别，对角线绿色单元格表示正确分类，非对角线表示错误分类，每个单元格中显示分类数量与对应的占比。从图中可以计算出正确预测和错误预测数量和对应百分比。ResNet、Swin Transformer、MobileNetV2和MobileNetV2-CBAM误判图像数分别为3、1、5、2幅，其中将完熟后期-枯熟期和已收割区互相误判的图像数分别为3、1、3、2幅；此外，MobileNetV2模型将乳熟-蜡熟初期和已收割区互相误判的图像数为2幅。综上分析，MobileNetV2-CBAM相比于MobileNetV2误判率降低。



图8　不同模型的混淆矩阵

Fig. 8　Confusion matrices for different models

通过混淆矩阵计算轻量化模型对小麦成熟期分类任务的精确率、召回率和特异度，如表6所示。在精确率方面，MobileViT、ShuffnetV2和MobileNetV2对L1的精确率表现稍差，ShuffnetV2在L0、L1、L2的精确率均低于95%，并且L1精确率低于90%，这表明3种模型对完熟后期-枯熟期的误判较多。MobileNetV2只在L2精确率达到100%，本文提出的MobileNetV2-CBAM在L0、L1、L2的精确率均达到100%。在召回率方面，MobileViT和MobileNetV2只有在L2召回率为100%，并且MobileViT在L0召回率只有96%，ShufflenetV2在4个类别召回率均不高于98%，MobileNet-CBAM在L0、L2、L3召回率均达到100%。在特异度方面，MobileNetV2-CBAM在L0、L1和L2特异度均达到100%。综上，MobileNetV2-CBAM模型在精确率、召回率和特异度3个方面均表现最好，适用于机收场景下小麦成熟期在线分类任务。

从上述模型误判类别及各评价指标分析发现，误判范围主要集中在完熟后期-枯熟期和已收割区之间。这是因为在联合收获机驾驶室顶部视角下，完熟后期-枯熟期和已收割区图像特征相似度较大，如图9所示，由于光照和收割时产生的尘土影响，少量完熟后期-枯熟期和已收割区图像的颜色和背景相似，小麦麦穗与麦茬特征相近，造成颜色区别减弱，混淆了相似的分类特征。

表6 不同模型的评价指标

Tab. 6 Evaluation indicators for different models %

模型	精确率				召回率				特异度			
	L0	L1	L2	L3	L0	L1	L2	L3	L0	L1	L2	L3
MobileViT	100	94.2	98.0	100	96.0	98.0	100	98.0	100	98.0	99.3	100
ShufflenetV2	94.2	89.6	94.2	100	97.0	95.0	97.0	88.0	98.0	96.3	98.0	100
MobileNetV2	99.0	98.0	100	98.0	99.0	99.0	100	97.0	99.7	99.3	100	99.3
MobileNetV2-CBAM	100	100	100	98.0	100	98.0	100	100	100	100	100	99.3

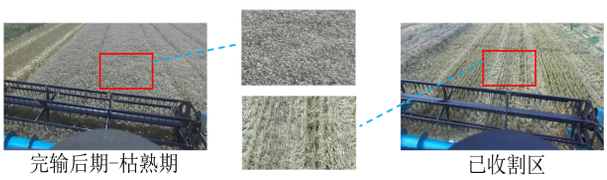


图9 误判图像类别

Fig. 9 Misclassification of image categories

2.1.3 可视化分析

为探究不同分类模型对图像特征的提取范围，使用Grad-CAM（Gradient-weighted class activation mapping）[36]可视化技术对小麦成熟期图像进行热力图分析。该方法通过显示高激活区域分析模型对图像的关注范围，有助于提高模型可解释性，并可作为提高网络性能的基础。

由于田间小麦成熟期识别属于较大面积识别，需要对图像中所有小麦信息特征进行提取，模型关注范围应足够广，同时联合收获机收割时，需要关注的小麦区域在割台正前方。

图10左侧4幅图展示了轻量化模型ShuffnetV2、MobileViT、MobileNetV2以及MobileNetV2-CBAM在小麦乳熟-蜡熟初期图像上的可视化效果。ShuffnetV2和MobileViT的高激活区域（红色部分）较为分散，MobileNetV2的次激活区域（绿色部分）覆盖的范围较大，红色区域主要集中在图像右侧，出现注意力偏移现象。图10右侧4幅图为MobileNetV2-CBAM分别在乳熟-蜡熟初期、蜡熟后期-完熟初期、完熟后期-枯熟期和已收割区的可视化效果，可以得到模型关注区域几乎覆盖整个图像，并且激活区域主要集中在图像正前方部分，因此本文提出的模型能够很好地实现上述要求。

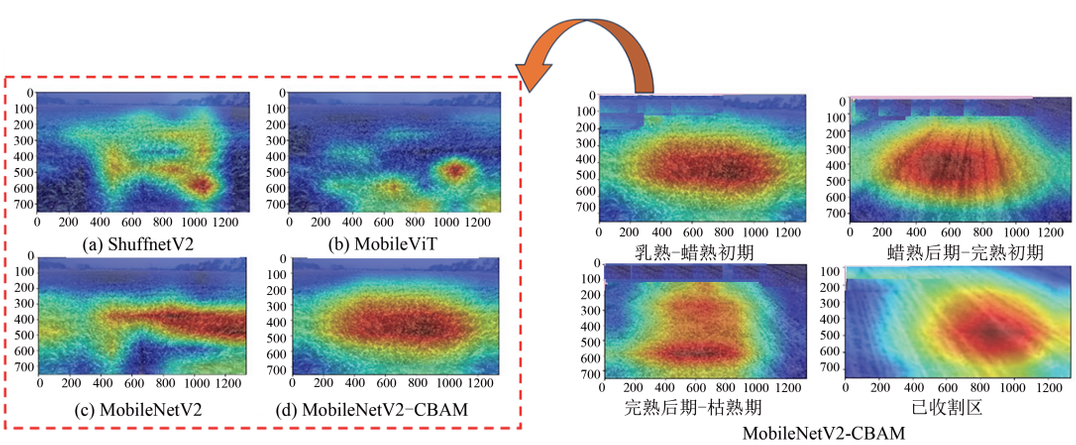


图10 不同模型热力图对比

Fig. 10 Comparisons of heat maps of different models

2.2 田间验证试验

2.2.1 试验条件

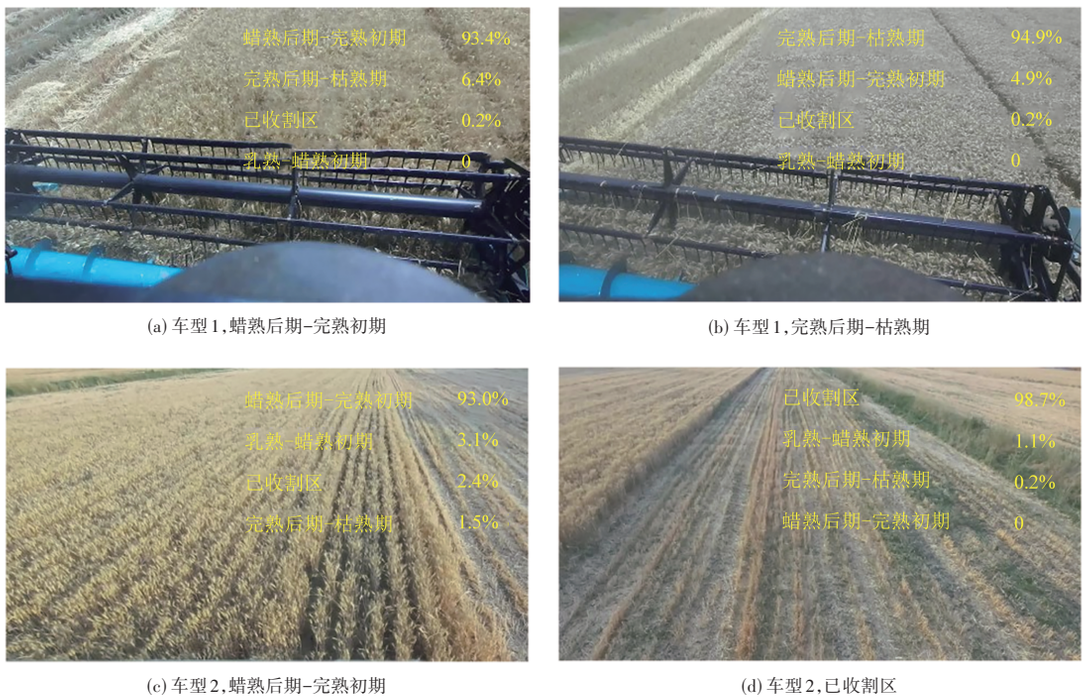
为验证本文模型在机收环境下的成熟期在线识别效果，在江苏省盐城市、连云港市以及北京市平谷区开展田间收获试验，试验选用国内两款谷物联合收获机，作业速度为4~6 km/h，分别开展小麦蜡熟后期-完熟初期、完熟后期-枯熟期和已收割区的在线分类识别试验，其中在车型1上验证了蜡熟后期-完熟初期和完熟后期-枯熟期的在线识别效果，在车型2上验证了蜡熟后期-完熟初期和已收割区的在线识别效果。

2.2.2 田间试验结果与分析

随机从田间试验识别结果中选取1幅图像进行分析，如图11所示，4种工况下成熟期均正确分类，以图11a为例进行分析，将图像识别为蜡熟后期-完熟初期概率为93.4%，识别为完熟后期-枯熟期概率为6.4%，识别为已收割区概率为0.2%，识别为乳熟-蜡熟初期概率为0，识别分类效果良好。

在总体识别准确率方面，车型1收割蜡熟后期-完熟初期小麦时，测试时长为52 s，每隔1 s处理1幅图像，其中4幅误判为完熟后期-枯熟期，识别准确率为92.3%；车型2在已收割区行驶时，1幅误判为乳熟-蜡熟初期，识别准确率为94.7%，其余两类成熟期识别准确率均为100%，平均准确率达到96.8%，具体试验结果如表7所示。试验结果表明，本文方法能够满足大田机收场景下小麦成熟期在线分类识别准确性和实时性要求。

车载视频分类误判原因主要有两点：①收割过程中出现坑洼强烈颠簸，出现图像丢失现象；在强光照射下，



(a) 车型1，蜡熟后期-完熟初期　(b) 车型1，完熟后期-枯熟期

(c) 车型2，蜡熟后期-完熟初期　(d) 车型2，已收割区

图11　部分图像分类结果

Fig. 11　Partial image classification results

表7　田间试验在线识别结果

Tab. 7　Field trial online identification results

参数	盐城(车型1)	连云港(车型1)	平谷(车型2)	
成熟期	完熟后期-枯熟期	蜡熟后期-完熟初期	蜡熟后期-完熟初期	已收割区
连续测试时长/s	17	52	26	19
准确率/%	100	92.3	100	94.7
总体识别准确率/%		96.8		

图像出现过度曝光呈现全白色现象，导致出现误判现象。②在已收割区内和小麦密度低的区域工作时，模型将地面生长的杂草与乳熟-蜡熟初期的特征混淆，进而影响模型对图像的判别效果。

3　结论

(1)建立了包含4 400幅图像的小麦成熟期数据集，以车载相机拍摄图像为主，无人机图像为辅，丰富数据集多样性。同时结合色彩增强、角度旋转、添加噪声和模糊度的数据增强方法，扩充图像数据集，提高模型对复杂农田环境适应性。

(2)提出一种基于车载相机和深度学习结合的机收场景下的小麦成熟期在线分类识别方法，采用卷积注意力机制模块对轻量化网络MobileNetV2顶层结构进行优化，并引入迁移学习进行预训练。在小麦成熟期测试集准确率高达99.5%，比传统MobileNetV2模型高0.7个百分点；与最优CNN分类模型和Transformer模型相比，分类精度未发生明显差异，模型内存占用量仅为其1/8和1/11；分析不同模型可视化结果，MobileNetV2-CBAM模型关注区域范围更聚焦，能够适应机收环境下较大面积成熟期识别任务。

(3)选用国内两款小麦联合收获机开展田间试验验证，试验结果表明，在车速4~6 km/h工况下，每隔1 s识别1幅图像，小麦成熟期总体识别分类准确率为96.8%，能够满足在线识别准确性和实时性要求。

(4)该研究成果不仅能够为收获作业过程中联合收获机智能化调控提供决策依据，而且可以为其它谷类作物的成熟期在线识别提供技术支撑。

参　考　文　献

[1]　张亚伟，王书茂，陈度，等. 基于近红外的小麦植株含水率检测方法[J]. 农业机械学报，2017，48(增刊)：118－122.
ZHANG Yawei，WANG Shumao，CHEN Du，et al. Measurement of wheat plants water content based on near-infrared photoelectric sensors [J]. Transactions of the Chinese Society for Agricultural Machinery，2017，48(Supp.)：118－122. (in Chinese)

[2]　陈度，王书茂，康峰，等. 联合收获机喂入量与收获过程损失模型[J]. 农业工程学报，2011，27(9)：18－21.
CHEN Du，WANG Shumao，KANG Feng，et al. Mathematical model of feeding rate and processing loss for combine harvester [J].

Transactions of the CSAE,2011,27(9):18 - 21. (in Chinese)

[3] YIN Y X ,QIN W C ,ZHANG Y W , et al. Compensation control strategy for the cutting frequency of the cutterbar of a combine harvester [J]. Biosystems Engineering, 2021,204:235 - 246.

[4] 杨立权,王万章,张红梅,等. 切流-横轴流玉米脱粒系统改进设计及台架试验[J]. 农业工程学报, 2018,34(1):35 - 43.
YANG Liquan, WANG Wanzhang, ZHANG Hongmei, et al. Improved design and bench test based on tangential flow-transverse axial flow maize threshing system[J]. Transactions of the CSAE, 2018,34(1):35 - 43. (in Chinese)

[5] DING B C, LIANG Z W, QI Y Q, et al. Improving cleaning performance of rice combine harvesters by DEM - CFD coupling technology[J]. Agriculture,2022,12(9):1457.

[6] PEPLER S, GOODING M J, ELLIS R H. Modelling simultaneously water content and dry matter dynamics of wheat grains[J]. Field Crops Research, 2006,95(1):49 - 63.

[7] ALVAREZ PRADO S, GALLARDO J M, SERRAGO R A, et al. Comparative behavior of wheat and barley associated with field release and grain weight determination[J]. Field Crops Research, 2013,144:28 - 33.

[8] PARVEJ M R, HURBURGH C R, HANNA H M, et al. Dynamics of corn dry matter content and grain quality after physiological maturity[J]. Agronomy Journal, 2020,112(2):998 - 1011.

[9] MENENDEZ Y C, BOTTO J F, GOMEZ N V, et al. Physiological maturity as a function of seed and pod water concentration in spring rapeseed (*Brassica napus* L.)[J]. Field Crops Research, 2019,231:1 - 9.

[10] TEODORO P E, TEODORO L P R, BAIO F H R, et al. Predicting days to maturity, plant height, and grain yield in soybean: a machine and deep learning approach using multispectral data[J]. Remote Sensing, 2021,13(22):4632.

[11] 蒙继华, 吴炳方, 杜鑫,等. 基于HJ-1A/1B数据的冬小麦成熟期遥感预测[J]. 农业工程学报, 2011, 27 (3):225 - 230.
MENG Jihua, WU Bingfang, DU Xin, et al. Predicting mature date of winter wheat with HJ-1A/1B data[J]. Transactions of the CSAE, 2011,27(3):225 - 230. (in Chinese)

[12] 杜颖. 基于多源遥感数据的冬小麦灌浆进程及成熟度监测研究[D]. 扬州:扬州大学, 2019.
DU Ying. Monitoring of grain filling and maturity of winter wheat based on multi-source remote sensing imagery[D]. Yangzhou: Yangzhou University,2019. (in Chinese)

[13] YUE Y, LI J H, FAN L F, et al. Prediction of maize growth stages based on deep learning[J]. Computers and Electronics in Agriculture, 2020,172:105351.

[14] KUMUDINI S, ANDRADE F H, BOOTE K J, et al. Predicting maize phenology: intercomparison of functions for developmental response to temperature[J]. Agronomy Journal, 2014,106(6):2087 - 2097.

[15] ZHAO J F, ZHAN X, JIANG Y Q, et al. Variations in climatic suitability and planting regionalization for potato in northern China under climate change[J]. PLOS ONE, 2018,13(9):e203538.

[16] MA J C, LI Y X, LIU H J, et al. Towards improved accuracy of UAV-based wheat ears counting: a transfer learning method of the ground-based fully convolutional network[J]. Expert Systems with Applications, 2022,191:116226.

[17] 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.

[18] HANG D Y, DING Y, CHEN P F, et al. Automatic extraction of wheat lodging area based on transfer learning method and Deeplabv3+ network[J]. Computers and Electronics in Agriculture, 2020,179:105845.

[19] 秦成帅. 智能收获机作物高度及割幅宽度视觉感知方法[D]. 南京:东南大学, 2021.
QIN Chengshuai. Crop height and cutting width visual perception method of intelligent harvester[D]. Nanjing: Southeast University,2021. (in Chinese)

[20] LIU L, DU Y F, CHEN D, et al. Impurity monitoring study for corn kernel harvesting based on machine vision and CPU-Net[J]. Computers and Electronics in Agriculture, 2022,202:107436.

[21] 付景枝,马悦,宏观,等. 基于I_CBAM-DenseNet模型的小麦发育期识别研究[J/OL]. 南京信息工程大学学报(自然科学版),1-16[2023-11-28]http://kns.cnki.net/kcms/detail/32.1801. N.20230918.1728.002.html.
FU Jingzhi, MA Yue, HONG Guan, et al. Research on wheat development stage recognition based on I_CBAM-DenseNet model[J/OL]. Journal of Nanjing University of Information Science & Technology (Natural Science), 1-16[2023-11-28] http://kns.cnki.net/kcms/detail/32.1801.N.20230918.1728.002.html. (in Chinese)

[22] ZHANG S X, FENG H, HAN S Y, et al. Monitoring of soybean maturity using UAV remote sensing and deep learning[J]. Agriculture, 2023,13(1):110.

[23] 扶兰兰, 黄昊, 王恒,等. 基于Swin Transformer模型的玉米生长期分类[J]. 农业工程学报, 2022,38(14):191 - 200.
FU Lanlan, HUANG Hao, WANG Heng, et al. Classification of maize growth stages using the Swin Transformer model[J]. Transactions of the CSAE, 2022,38(14):191 - 200. (in Chinese)

(下转第100页)

[16] JAFARI A, ASADI V. Precise measurement of residue cover by means of image processing techniques[J]. International Journal of Natural and Engineering Sciences, 2011, 5(3):53 - 56.
[17] 苏艳波,张东远,李洪文,等.基于自动取阈分割算法的秸秆覆盖率检测系统[J].农机化研究,2012,34(8):138 - 142.
SU Yanbo, ZHANG Dongyuan, LI Hongwen, et al.Measuring system for residue cover rate based on automation threshold[J]. Journal of Agricultural Mechanization Research, 2012, 34(8): 138 - 142.(in Chinese)
[18] 王丽丽,徐岚俊,魏舒,等.基于Sauvola与Otsu算法的秸秆覆盖率图像检测方法[J].农业工程,2017,7(4):29 - 35.
WANG Lili, XU Lanjun, WEI Shu, et al. Straw coverage detection method based on Sauvola and Otsu segmentation algorithm[J]. Agricultural Engineering, 2017, 7(4): 29 - 35.(in Chinese)
[19] 李洪文,李慧,何进,等.基于人工神经网络的田间秸秆覆盖率检测系统[J].农业机械学报,2009,40(6):58 - 62.
LI Hongwen, LI Hui, HE Jin, et al. Measuring system for residue cover rate in field based on BP neural network[J].Transactions of the Chinese Society for Agricultural Machinery, 2009, 40(6): 58 - 62.(in Chinese)
[20] 于国庆,郝若帆,马洪涛,等.基于图像处理和支持向量粉碎性秸秆覆盖率的图像识别方法研究[J].河南农业科学,2018,47(11):155 - 160.
YU Guoqing, HAO Ruofan, MA Hongtao, et al.Research on image recognition method based on SVM algorithm and ESN algorithm for crushed straw mulching rate[J]. Journal of Henan Agricultural Sciences, 2018, 47(11):155 - 160.(in Chinese)
[21] 李佳,吕程序,苑严伟,等.快速傅里叶变换结合SVM算法识别地表玉米秸秆覆盖率[J].农业工程学报,2019,35(20):194 - 201.
LI Jia, LÜ Chengxu, YUAN Yanwei, et al. Automatic recognition of corn straw coverage based on fast Fourier transform and SVM[J]. Transactions of the CSAE, 2019, 35(20): 194 - 201.(in Chinese)
[22] 冯为栋.双目视觉在田间秸秆覆盖率监测系统中的应用[J].南方农机,2019,50(20):17.
[23] 刘媛媛,周小康,王跃勇,等.改进U-Net模型的保护性耕作田间秸秆覆盖检测[J].光学精密工程,2022,30(9):1101 - 1112.
LIU Yuanyuan, ZHOU Xiaokang, WANG Yueyong, et al.Straw coverage detection of conservation tillage farmland based on improved U-Net model[J]. Optics and Precision Engineering, 2022, 30(9): 1101 - 1112.(in Chinese)
[24] 马钦,万传峰,卫建,等.基于U-Net和特征金字塔网络的秸秆覆盖率计算方法[J].农业机械学报,2023,54(1):224 - 234.
MA Qin, WAN Chuanfeng, WEI Jian, et al. Calculation method of straw coverage based on U-Net network and feature pyramid network[J].Transactions of the Chinese Society for Agricultural Machinery, 2023, 54(1): 224 - 234.(in Chinese)
[25] 安晓飞,代均益,路振田,等.基于彩色空间距离优化的秸秆覆盖率检测算法研究[J].农业机械学报,2023,54(增刊2):229 - 234.
AN Xiaofei, DAI Junyi, LU Zhentian, et al. Development of rapid detection algorithm of straw coverage rate based on color spatial distance gray and K-means method[J].Transactions of the Chinese Society for Agricultural Machinery, 2023, 54(Supp. 2): 229 - 234.(in Chinese)

(上接第80页)

[24] 农敢网.冬小麦生长周期[EB/OL].(2022-10-28)[2023-1-16].https://www.nonggan.com/a/34887.html.
[25] GB/T 5262—2008 农业机械试验条件测定方法的一般规定[S].北京:中华人民共和国国家质量监督检验检疫总局、中国国家标准化管理委员会,2008.
[26] 孔建磊,金学波,陶治,等.基于多流高斯概率融合网络的病虫害细粒度识别[J].农业工程学报,2020,36(13):148 - 157.
KONG Jianlei, JIN Xuebo, TAO Zhi, et al. Fine-grained recognition of diseases and pests based on multi-stream Gaussian probability fusion network[J]. Transactions of the CSAE, 2020,36(13):148 - 157.(in Chinese)
[27] SANDLER M, HOWARD A, ZHU M L, et al. MobileNetV2: inverted residuals and linear bottlenecks[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. IEEE, 2018:4510 - 4520.
[28] WOO S, PARK J, LEE J, et al. CBAM: convolutional block attention module[C]//Proceedings of the European Conference on Computer Vision (ECCV),2018:3 - 19.
[29] BAO W X, YANG X H, LIANG D, et al. Lightweight convolutional neural network model for field wheat ear disease identification[J]. Computers and Electronics in Agriculture, 2021,189:106367.
[30] CAI Y L, ZENG F G, XIAO J Y, et al. Attention-aided semantic segmentation network for weed identification in pineapple field[J]. Computers and Electronics in Agriculture, 2023,210:107881.
[31] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition[J]. arXiv Preprint, arXiv: 1409.1556, 2014.
[32] HE K M, ZHANG X Y, REN S Q, et al. Deep residual learning for image recognition[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016:770 - 778.
[33] MA N N, ZHANG X Y, ZHENG H T, et al. ShuffleNet V2: practical guidelines for efficient CNN architecture design[C]//Proceedings of the European Conference on Computer Vision (ECCV), 2018:116 - 131.
[34] MEHTA S, RASTEGARI M. MobileVit: light-weight, generalpurpose, and mobile-friendly vision transformer[J]. arXiv Preprint, arXiv: 2110.02178,2021.
[35] LIU Z, LIN Y T, CAO Y, et al. Swin transformer: hierarchical vision transformer using shifted windows[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. IEEE, 2021:10012 - 10022.
[36] SELVARAJU R R, COGSWELL M, DAS A, et al. Grad-CAM: visual explanations from deep networks via Gradient-based localization[C]//IEEE International Conference on Computer Vision, 2017:618 - 626.