doi:10.6041/j.issn.1000-1298.2024.05.021

# 基于改进 FasterNet 的轻量化小麦生育期识别模型

时 雷<sup>1,2</sup> 雷镜楷<sup>1</sup> 王 健<sup>1</sup> 杨程凯<sup>1</sup> 刘志浩<sup>1</sup> 席 磊<sup>1,2</sup> 熊蜀峰<sup>1</sup> (1.河南农业大学信息与管理科学学院,郑州 450046; 2.河南粮食作物协同创新中心,郑州 450046)

**摘要:**针对现阶段小麦生育期信息获取需依靠人工观测,效率低、主观性强等问题,本文构建包含冬小麦越冬期、返 青期、拔节期和抽穗期4个生育期共计4599幅小麦图像数据集,并提出一种基于 FasterNet 的轻量化网络模型 FSST(Fast shuffle swin transformer),开展4个关键生育期的智能识别。在 FasterNet 部分卷积的基础上引入 Channel Shuffle 机制,以提升模型计算速度。引入 Swin Transformer 模块来实现特征融合和自注意力机制,用来提升小麦关 键生育期识别准确率。调整整个模型结构,进一步降低网络复杂度,并在训练中引入 Lion 优化器,加快网络模型收 敛速度。在自建的数据集上进行模型验证,结果表明,FSST 模型参数量仅为1.22×10<sup>7</sup>,平均识别准确率、F1 值和 浮点运算量分别为 97.22%、78.54%和3.9×10<sup>8</sup>,与 FasterNet、GhostNet、ShuffleNetV2和 MobileNetV34种模型相 比,FSST 模型识别精度更高,运算速度更快,并且识别时间分别减少84.04%、73.74%、72.22%和77.01%。提出 的 FSST 模型能够较好地进行小麦关键生育期识别,并且具有识别快速精准和轻量化的特点,可以为大田作物生长 实时监测提供信息技术支持。

关键词:小麦;生育期识别;FasterNet;轻量化;Lion优化器 中图分类号:TP391.41 文献标识码:A 文章编号:1000-1298(2024)05-0226-09



# Lightweight Wheat Growth Stage Identification Model Based on Improved FasterNet

SHI Lei<sup>1,2</sup> LEI Jingkai<sup>1</sup> WANG Jian<sup>1</sup> YANG Chengkai<sup>1</sup> LIU Zhihao<sup>1</sup> XI Lei<sup>1,2</sup> XIONG Shufeng<sup>1</sup>
 (1. College of Information and Management Science, Henan Agricultural University, Zhengzhou 450046, China
 2. Collaborative Innovation Center of Henan Grain Crops, Zhengzhou 450046, China)

Abstract: In response to the problems of low efficiency and strong subjectivity in obtaining information about the current stage of wheat development that relies on manual observation, a wheat image dataset consisting of four key growth stages of winter wheat: winterovering stage, green-turning stage, jointing stage, and heading stage, totaling 4 599 images were constructed. A lightweight model FSST (fast shuffle swin transformer) based on FasterNet was proposed to carry out intelligent recognition of these four key growth stages. Firstly, based on the partial convolution of FasterNet, the Channel Shuffle mechanism was introduced to improve the computational speed of the model. Secondly, the Swin Transformer module was introduced to achieve feature fusion and self attention mechanism, it can improve the accuracy of identifying key growth stages of wheat. Then the structure of the whole model was adjusted to further reduce the network complexity, and the Lion optimizer was introduced into the training to accelerate the training speed of the model. Finally, model validation on the self-built wheat dataset with four key growth stages was performed. The results showed that the parameter quantity of the FSST model was only  $1.22 \times$  $10^7$ , the average recognition accuracy was 97.22%, the F1 score was 78.54%, and the FLOPs was  $3.9 \times 10^8$ . Compared with that of the FasterNet, GhostNet, ShuffleNetV2 and MobileNetV3 models, the recognition accuracy of the FSST model was higher, the operation speed was faster, and the recognition time was reduced by 84.04%, 73.74%, 72.22% and 77.01%, respectively. The FSST model

作者简介:时雷(1979—),女,教授,博士,主要从事智慧农业和数据挖掘研究,E-mail: shilei@ henau. edu. cn

收稿日期: 2023-12-12 修回日期: 2024-03-01

**基金项目:** 国家自然科学基金项目(31501225)、河南省科技研发计划联合基金(优势学科培育类)项目(222301420113)和河南省自然科 学基金项目(222300420463、232300420186)

通信作者:熊蜀峰(1980—),男,副教授,博士,主要从事机器学习和大数据处理研究, E-mail: xsf@ whu. edu. cn

proposed can effectively identify the key growth stage of wheat, and had the characteristics of fast, accurate, and lightweight recognition. It can provide a reference for optimizing the application of deep learning models in smart agriculture and offerring information technology support for real-time monitoring of field crop growth on resource-constrained mobile devices.

Key words: wheat; growth stage identification; FasterNet; lightweight; Lion optimizer

# 0 引言

小麦是我国主要的粮食作物之一,2022 年全国 小麦总产量1.37×10°t<sup>[1]</sup>,小麦生长发育与产量息 息相关,对其进行生长监测对我国粮食安全具有重 要意义。小麦生育期识别是小麦生长发育监测的重 要组成部分,而现阶段小麦生育期信息获取主要依 靠人工观测,效率低且耗费大量人力物力,无法满足 实时、快速的监测需求。

遥感具有大面积和快速无损的优势,为农作物 物候期识别提供了有效的技术手段。杨蜀秦等<sup>[2]</sup> 使用无人机可见光遥感影像对小麦种植行检测划 分。周琦等<sup>[3]</sup>使用无人机多光谱遥感数据对不同 生育期小麦冠层 SPAD 值进行估算,其中拔节期决 定系数 *R*<sup>2</sup>达到 0.827。兰仕浩等<sup>[4]</sup>利用卫星光谱尺 度反射率对冬小麦生物量估算,*R*<sup>2</sup>达到 0.70。杨欣 等<sup>[5]</sup>利用高光谱对冬小麦进行氮素估算,*R*<sup>2</sup>达到 0.76。但遥感在地块级别技术支持与实时精准服务 上略受限,无法在田间地块小尺度上进行实时物候 监测。

近年来,深度学习算法的出现极大地提升了图像分类的准确率,其中卷积神经网络<sup>[6]</sup>在农业图像 识别领域备受青睐。随着智能化技术在场景识别、 物体分类等方面的研究不断深入,深度学习在农业 领域的应用前景日益广阔。利用深度学习的高效率 和准确性,逐渐成为了农作物物候特性识别的理想 选择。徐建鹏等<sup>[7]</sup>使用卷积神经网络模型,对水稻 生育期进行分类,分类精度达到 97.33%。申华磊 等<sup>[8]</sup>使用基于 U2 – Net 的网络模型完成了倒伏面积 提取,准确率可达 97.25%。李云霞等<sup>[9]</sup>基于 Faster R – CNN 和 NMS 构建了小麦计数模型,其  $R^2$ 可达 0.89。孙少杰等<sup>[10]</sup>使用卷积神经网络与 BP 神经网 络结合对小麦产量进行估算, $R^2$ 达到 0.96。传统深 度学习参数量较大,通常难以在移动端设备部署,不 能较好地满足实际应用需求。

轻量化网络模型在推理速度和参数量方面具 有很大优势,越来越多的研究者开始使用 EfficientNet<sup>[11]</sup>、ShuffleNet<sup>[12]</sup>、MobileNet<sup>[13-15]</sup>、 GhostNet<sup>[16]</sup>和FasterNet<sup>[17]</sup>等轻量化网络模型。杨 森森等<sup>[18]</sup>对MobileViT 轻量化网络进行了改进, 对田间杂草进行识别,准确率达到 99.61%。苗荣 慧等<sup>[19]</sup>利用 MobileNetV3 改进 YOLO V7 并对番茄 成熟度进行检测,精度达到 98.6%,同时将内存占 用量 控 制 在 66.5 MB。范 晓 飞 等<sup>[20]</sup>利用 MobileNetV1 替换 YOLO V4 主干网络,单幅图像检 测耗时达到 1.85 s。孙 道宗等<sup>[21]</sup>使用改进的 SqueezeNet 对茶树叶片分类,参数量减少 49.1%, 准确率提高 7.7%。

真实环境下大田图像背景较为复杂,当前大多 数模型识别准确率和效率都不能满足实时识别小麦 生育期的需求。本文对 FasterNet 模型进行改进。 在 FasterNet 模型中引入 Channel Shuffle,减少网络 参数量;在 FasterNet 模型中嵌入 Swin Transformer 模 块,加强特征融合并实现自注意力机制,以提高模型 精度;对 FasterNet 的网络框架进行重新编排,提出 轻量化模型 FSST(Fast shuffle swin transformer),同 时在训练过程中引入 Lion 优化器<sup>[22]</sup>,使模型快速 收敛。最终将模型应用于小麦生育期数字图像识 别,以期代替部分人工观测任务,为作物物候特性监 控提供有效支持。

# 1 小麦图像数据集

#### 1.1 数据来源

以河南省冬小麦为研究对象,其生长周期在 220~270 d, 一般划分为播种期、出苗期、分蘖期、越 冬期、返青期、起身期、拔节期、孕穗期、抽穗期、开花 期、灌浆期和成熟期12个生育时期。选取小麦越冬 期、返青期、拔节期、抽穗期4个生育时期,其冠层图 像采集于河南省许昌市河南农业大学许昌校区试验 田(34°12′06″N,113°58′26″E),采集时间为 2021 年 1-4月。试验田属暖温带大陆性季风气候,四季分 明,年平均气温 14.6℃,年平均无霜期 216 d,年平 均降水量 728.9 mm,适合小麦等粮食作物生长。使 用智能手机(小米8,后置1.2×10<sup>7</sup>像素的双摄像 头,手机拍摄模式设定为 AI 模式下的光学变焦,最 高分辨率4344 像素 ×4344 像素) 在距离地面定高 1 m 进行拍摄,所有图像均在 10:00-11:00、 14:00—15:00 时间段内且在自然光照、自然环境条 件下进行采集。所处生育期、拍摄日期与图像数量 如表1所示。小麦生育期部分原始图像如图1所示。

表1 不同生育期图像数量

Tab.1 Number of images at different growth stages 幅

生育期	拍摄日期	图像数量	分割后图像数量
越冬期	2021年1月15日	102	918
	2021年3月5日	80	720
返青期	2021年3月19日	80	720
	2021年3月30日	80	720
拔节期	2021年4月9日	86	774
抽穗期	2021年4月15日	83	747



#### 1.2 数据集划分

本研究将扩充和增强后的数据集按照比例 7:3 划分成训练集和验证集。原始图像像素尺寸大,数 量有限(共计 511 幅图像),具有高精度小样本数据 集特点。为扩充小麦生育期图像数据集,将原有 511 幅小麦生育期图像进行等比分割,分割为9份, 然后将分割后图像尺寸重新调整至 448 像素 × 448 像素。至此图像数据集扩充至 4 599 幅,图像分 割方法如图 2 所示。



Fig. 2 Method of image splite

对于输入图像,尽管进行了分割扩充操作,但仍 建立在原始图像之上,为了进一步提高模型鲁棒性, 对图像进行增强操作,包括随机裁剪、随机旋转、随 机翻转、随机缩放和随机噪声等数据增强技术来增加数据集多样性和提高模型泛化能力,以促进在训练过程中更好地模拟实际场景中的变化和不确定性。每次训练迭代时增强函数会再对输入图像进行增强处理,可以增加数据多样性,从而提升模型泛化能力。

# 2 小麦生育期识别模型

#### 2.1 FasterNet

FasterNet 是一种高效的神经网络,通常用于完成目标检测任务,并在速度和精度方面进行了优化, 优于 MobileVit<sup>[23]</sup>等网络。FasterNet 的核心思想是 在保持轻量级和高速度的基础上,提高特征表达能 力和感受野的覆盖范围。FasterNet 网络结构由4个 阶段组成,如图3所示。图中每个阶段都是对特征 进行提取,仅在卷积核大小上有区别,Embedding 模 块由一个步长为4的正则卷积组成,Merging 层由一 个步长为2的卷积组成,用于空间下采样和通道数 扩展。



Faster Block 模块是 FasterNet 的核心模块,该模 块设计思想来源于 GhostNet,在一定程度上解决了 特征卷积通道上具有一定冗余的问题,但与 GhostNet 相比,并没有采用 DWConv(深度可分离卷 积),而是提出了一种新的被称为 PConv(部分卷积) 的算子,PConv 通过减少冗余计算和内存访问来更 有效地提取空间特征。Faster Block 模块结构如图 4 所示。首先使用 3 × 3 的 PConv 提取特征,提取后使 用 1 × 1 的 Conv(标准卷积)进行特征降维,经过 BN (Batch normalization)归一化和 GELU 激活函数后使 用 1 × 1 的卷积进行升维,最后升维特征与输入特征 进行相加得到最终输出。



### 2.2 改进 FasterNet 的轻量化小麦生育期识别模型

为解决大田环境下小麦背景图像复杂,以及图 像尺寸偏大而导致的识别准确率较低、模型尺寸大、 处理效率不高的问题,在轻量化、特征融合和网络结 构调整3方面对FasterNet进行了改进。

# 2.2.1 模型轻量化改进

Channel Shuffle 操作能够有效地增强特征的跨 通道交流,从而提高网络模型性能,同时在一定程度 上减少模型计算复杂性。它通过将特征图分组并使 用不同的卷积核进行卷积,降低计算复杂度的同时 保证模型性能。此外,Channel Shuffle 操作还打破 了输入特征图不同通道之间的依赖关系,增加了模 型泛化能力。同时,它也提高了不同分组特征之间 的信息流动,有助于挖掘潜在的特征信息。在 FasterNet中,PConv操作虽大幅度提高了计算效率, 但模块末尾的1×1卷积仍然存在较大的资源消耗, 为进一步提升模型性能,将末尾的1×1卷积块替换 为GConv(分组卷积)模块,同时加入 Channel Shuffle 机制,进而降低参数量和计算量。

原有的 Faster Block 将最初的特征先进行 3×3的部分卷积,然后经过1×1的标准卷积,通过 BN 归一化和 GELU 激活函数后再进行1×1的卷 积,最终与初始特征进行 Add 操作。而对 Faster Block 的改进点主要是在 Channel Shuffle 的基础上 先通过全局平均池化层再与最初的特征进行相加。 首先将 Faster Block 中最后一层 Conv 操作替换为 GConv,降低标准卷积的计算量。同时引入 Channel Shuffle 机制,然后将 3×3 的部分卷积模块接入 1× 1 的卷积块,改进的模块结构如图 5 所示。



# 2.2.2 特征融合与注意力机制

特征融合能够综合利用各种信息,使得模型能 够更好地理解数据的内在结构和规律。注意力机制 是指通过有选择地关注输入信息中的一部分,而忽 略其他不太相关的信息来处理任务。Swin Transformer 是一种特征融合与注意力机制的综合实 现,其结构如图 6 所示,该模块能够有效地利用 Transformer 的自注意力机制,捕捉图像中的高级特 征,从而提高分类准确率。此外,Swin Transformer 通过将图像分解为多个小块并分别进行自注意力计 算,降低模型参数量和计算复杂度。



Swin Transformer 由 3 个核心模块组成,分别是 多头 自 注 意 力 机 制 (Windows multi-head selfattention, W-MSA)模块、多层感知机模块 (Multilayer perceptron, MLP)和偏移的多头自注意力机制 (Shifted windows multi-head self-attention, SW-MSA) 模 块。每 个 模 块 前 都 会 使 用 LN (Layer normalization)归一化,在模块后又会使用 Add 操作 将特 征 相 加, 通 过 组 合 3 个 模 块 得 到 了 Swin Transformer 结构。将 Swin Transformer 模块接入到 第 4 层卷积后,使模型融合多尺度特征并实现自注 意力机制,可提高模型精度。

# 2.2.3 网络结构改进

轻量化 Swin Transformer 模块具有的特征融合 和自注意力机制能够显著提升模型综合性能,但 Transformer 仍然是一个较为复杂的模块,不可避免 地增加模型复杂度。FasterNet 最初网络结构设计的 目的是在千万数量级别的 ImageNet 数据集<sup>[24]</sup>进行 分类.图像类别多达1000类。使用的小麦生育期 数据集无论是图像数量还是图像类别均远低于 ImageNet 公共数据集。为此,本文重新设计了 FasterNet 模型的整体架构以减少模型复杂度,提出 一种基于 FasterNet 的轻量化网络模型 FSST (Fast shuffle swin transformer)。FSST 模型将模型整体层 数减至原来一半,且只包含4个改进的 Faster Block 模块,并将第1个 FasterNet Block 中 PConv 的输出 通道数从 50 降至 32,最后一个 FasterNet Block 中 PConv 的输出通道数从 1 280 降至 512, 其整体架构 如图7所示。



# 2.3 Lion 优化器

Lion(EvoLved sign momentum)是由谷歌在 2023 年通过数千 TPU小时的算力搜索并结合人工干预, 得到的一个具有更高的速度且占用内存更少的优化 器,其在图像分类、图文匹配、扩散模型、语言模型预 (2)

训练和微调等多个任务上进行了充分的实验,多数 任务都显示 Lion 比目前主流的 Adam<sup>[25]</sup>、Radam<sup>[26]</sup> 和 AdamW<sup>[27]</sup>优化器有着更好的效果。

Lion 优化器更新过程为

$$\begin{cases} \mu_{t} = \operatorname{sign}(\beta_{1}m_{t-1} + (1 - \beta_{1})g_{t}) + \lambda_{t}\theta_{t+1} \\ \theta_{t} = \theta_{t-1} - \eta_{t}\mu_{t} \\ m_{t} = \theta_{t}m_{t} + (1 - \theta_{t})g_{t} \end{cases}$$
(1)

其中 
$$g_t = \nabla_{\theta} L(\theta_{t-1})$$

其中

$$\operatorname{sign}(\theta) = \begin{cases} 1 & (\theta \ge 0) \\ -1 & (\theta < 0) \end{cases}$$
(3)

式中 g,——损失函数梯度

AdamW 优化器更新过程为  $(m_1 = \beta_1 m_{1-1} + (1 - \beta_1)g_1)$ 

$$\begin{cases} v_{i} = \beta_{2} v_{i-1} + (1 - \beta_{2}) g_{i} \\ \hat{m}_{i} = m_{i} / (1 - \beta_{1}^{i}) \\ \hat{v}_{i} = v_{i} / (1 - \beta_{2}^{i}) \\ \hat{\mu}_{i} = \hat{m}_{i} / (\sqrt{v_{i}} - \varepsilon) + \lambda_{i} \theta_{i-1} \\ \theta_{i} = \theta_{i-1} - \eta_{i} \mu_{i} \end{cases}$$
(4)

使用 Lion 替换 AdamW 优化器, Lion 相比 AdamW 参数更少,同时少缓存一组参数,所以更节 约显存,并且去掉 AdamW 更新过程中计算量最大 的除法和开根号运算,使得计算量减少,提升了运算 速度。

#### 2.4 实验设置

本实验硬件环境为内存:32 GB,CPU:Intel(R) Core i7 11800F, GPU: NVIDIA V100 32 GB, 操作系统 为 Windows 11, 选用的开源深度学习框架为 Torch, 通过 Torch 调用 GPU 实现不同神经网络的训练。 实验时,将图像尺寸固定为448 像素×448 像素.优 化器分别采用 AdamW 和 Lion, 批次大小设置为 16, 学习率初始值为 0.001, 迭代次数为 120。通过与 FasterNet、GhostNet、ShuffleNetV2 和 MobileNetV3 的 比较来评估本研究方法的性能。

#### 2.5 评价指标

使用模型在自建的小麦生育期数据集测试集上 的准确率(Accuracy)、精确率(Precision)、召回率 (Recall)和F1 值(F1 score)衡量模型性能,并且使 用模型参数量(Parameters)、模型内存占用量、浮点 计算量(FLOPs)和100幅生育期图像上平均识别时 间作为模型复杂度衡量指标。

#### 结果与分析 3

#### **3.1 Channel Shuffle** 对模型的影响

为探究 Channel Shuffle 模块的位置与网络性能 的关系,分别将 Channel Shuffle 接入模型每个 FasterNet Block 内部(A1),每个 FasterNet Block 外 部(A2), 第1个 FasterNet Block 和最后1个 FasterNet Block 之间(A3),并进行对比,如图 8 所示。



小麦生育期数据集测试集上结果如表2所示。 从表2可以看出,在A1区域添加Shuffle Channel模 块,准确率明显升高。其原因是将输入特征与输出 特征连接后进行通道混洗,从而达到不同通道间信 息交换的目的,增加了模型非线性表示能力<sup>[28-29]</sup>。 在 A2 区域添加 Channel Shuffle,虽然准确率相较于 在 A1 位置提升 0.12 个百分点,由于添加的 Channel Shuffle 模块过多,参数量增加 5.34 × 10<sup>6</sup>,同时准确 率提升不大,总体不利于轻量化性能提升。而在 A3 区域添加 Channel Shuffle 模块后,参数量基本不变 且准确率下降,效果最差。

表 2 Channel Shuffle 在不同位置的效果

### Tab. 2 Effect of Channel Shuffle at different positions

模块位置	准确率/	精确率/	召回率/	F1 值/	会粉导
	%	%	%	%	沙奴里
A1	98.15	92.13	92.31	89.47	9. $26 \times 10^{6}$
A2	98.27	92.24	97.32	96.58	$1.46 \times 10^7$
A3	96.48	90. 58	92.10	76.27	9. $26 \times 10^{6}$

#### **3.2** Swin Transformer 对模型的影响

为探究 Swin Transformer 模块对网络性能的影 响,分别将 Swin Transformer 加入到网络中第3个特 征提取模块(B1)和第4个特征提取模块(B2)后进 行验证,如图9所示。



实验结果如表3所示。从表3可以看出,在B1 位置添加 Swin Transformer 模块后在验证集上最高 准确率为 99.87%, 相较于不添加 Swin Transformer 模块的基准模型 FasterNet 准确率 96.19%, 提高 3.68 个百分点,其主要原因可能是抽穗期高分

表 3 Swin Transformer 在不同位置的效果 Tab. 3 Effect of Swin Transformer at different positions

模块位置	准确率/ %	精确率/ %	召回率/ %	F1 值/ %	参数量
B1	99.87	99.69	99.31	99.50	$1.81 \times 10^{7}$
B2	95.75	96.78	96.32	96.55	$1.81 \times 10^7$

辦率图像上具有大量的麦穗,而 Swin Transformer 模块能够更加关注密集的小目标<sup>[30]</sup>,同时能减 少背景干扰带来的影响。而在 B2 处添加 Swin Transformer 模块后准确率相较于基准模型下降 0.44 个百分点,同样遇到了在最后一层卷积上添 加模块导致准确率降低的问题,其原因可能是由 于特征在 B2 之前就已经固定。由于增加的模块 仅位置不同而数量相同,并没有造成参数量 改变。

### 3.3 消融实验

为 了 探 究 使 用 Channel Shuffle、Swin Transformer、Lion 优化器和架构调整的改进方式对 FasterNet 模型性能的影响,在小麦生育期数据集上 进行了消融实验,结果如表4所示。

表 4 消融实验结果 Tab. 4 Results of ablation experiment

模型	因素				评价指标			
	Channel Shuffle	Swin Transformer	Lion 优化器	架构调整	准确率/%	参数量	浮点运算量	模型内存占用量/MB
	$\sim$	×	×	×	98.27	9. $26 \times 10^{6}$	$1.17 \times 10^{9}$	182.14
	×	$\checkmark$	×	×	99.87	$1.71 \times 10^{7}$	$1.64 \times 10^{9}$	304. 61
FasterNet	×	×	$\checkmark$	×	93.25	$1.01 \times 10^{7}$	$1.25 \times 10^{9}$	200.71
	$\checkmark$	$\checkmark$	×	×	99.29	$1.63 \times 10^{7}$	$1.52 \times 10^{9}$	273.93
	$\checkmark$	$\checkmark$	$\checkmark$	×	99.17	$1.63 \times 10^{7}$	$1.52 \times 10^9$	271.62
FSST	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	97.22	$1.22 \times 10^{7}$	$3.90 \times 10^{8}$	167.30

注:√表示使用该模块,×表示不使用该模块。

由表4可知,加入Swin Transformer 模块带来的 精度提升最明显,比原始FasterNet 模型提高3.68 个百分点。加入Channel Shuffle 后准确率提升虽然 较少,但参数量有所降低,Channel Shuffle 模块能够 同时提升模型准确率并降低参数量和内存占用量。 调整网络结构后,准确率下降1.95个百分点,但参 数量直接从1.52×10°优化至3.90×10<sup>8</sup>,内存占用 量也从271.62 MB下降至167.30 MB,与此相比调 整网络结构带来准确率降低可忽略不计。

为了提高训练速度,使得模型快速收敛,使用 Lion 优化器替换 AdamW 优化器,两种优化器在训 练过程中损失值与准确率如图 10 所示。优化器本 身的改变并不能直接影响模型的尺寸、参数量、浮点 运算量与内存占用量等相关指标。在小麦生育期验 证集中 Lion 优化器带来的精度提升是模型收敛速 度变快的结果,随着轮数持续增加,Lion 优化器与 AdamW 优化器最终准确率接近。

# 3.4 FSST 与其他模型性能对比

采用 FSST、FasterNet、GhostNet、MobileNetV3 和 ShuffleNetV2 共5 种深度卷积网络模型获得的小麦 生育期图像分类结果如表5 所示。

从表 5 可以看出,FSST 模型对小麦生育期图像 识别的准确率最高,且识别速度最快。在参数量上, 由于 FSST 模型中含有 Swin Transformer 模块,因此 相比于 FasterNet 参数量增加 2.10 × 10<sup>6</sup>。在内存占



图 10 Lion 与 AdamW 训练中损失值与准确率对比 Fig. 10 Comparison of training loss and accuracy between Lion and AdamW

用量、准确率和识别时间3个指标上,相比于 FasterNet、GhostNet、ShuffleNetV2和MobileNetV3,均 取得了最优的结果。

#### 3.5 小麦生育期识别结果

为了验证在实际场景下模型对小麦各生育期的

2024年

识别率,将上述模型在不同生育期下的测试集上进行识别,结果见表6,FSST 网络模型4个生育期平均 正确识别率达到97.22%,训练集识别错误样本仅 有128个,验证集识别错误样本有31个。各个生育 期正确识别率也均优于其他对比模型。其中FSST 模型在4个生育期上的混淆矩阵如图11所示。

表 5 模型性能对比 Tab.5 Comparison of model performance

		模型内	准确	率/%	识别	河占运管
模型	参数量	存占用	训练	验证	时间/	日点均异
		量/MB	集	集	s	里
FSST	$1.22 \times 10^7$	167.30	96.37	97.22	12.54	$3.90 \times 10^{8}$
FasterNet	$1.01\times 10^7$	200.71	95.44	96.19	78.56	$1.\ 25\times 10^9$
GhostNet	$1.99 \times 10^7$	462.69	94.31	95.65	47.76	5. $40 \times 10^8$
ShuffleNetV2	8.75 $\times 10^{6}$	202.78	97.32	96.58	45.14	5. 50 $\times 10^8$
MobileNetV3	2. $10 \times 10^7$	444. 50	92.10	94.76	54. 54	8. $30 \times 10^8$

注:其中识别时间是在 Intel i3 550 CPU 环境下,累计识别 100 幅 像素尺寸为 448 像素 × 448 像素图像的总消耗时间。

表6 不同模型在小麦生育期验证集下识别准确率

Tab. 6 Recognition accuracy of different models on

whea	t growth	stage	validation	set	%
whica	i growin	stage	vanuation	SUL	70

模型	越冬期	返青期	拔节期	抽穗期	平均值
FSST	98.36	96.53	98.05	96.94	97.22
FasterNet	97.71	95.83	96.11	95.42	96.19
GhostNet	96.18	95.00	97.36	95.14	95.65
ShuffleNetV2	97.06	96.11	97.22	96.67	96.58
MobileNetV3	96.18	94.02	95.83	94.03	94.76

由图 11 可知,大部分识别错误的图像集中在拔 节期和抽穗期。而越冬期图像与其他 3 个生育期图 像差异较为明显,识别准确率最高。返青期、拔节 期、抽穗期可能由于图像经过一系列增强处理,导致 部分关键特征被抹除,如果能够拥有更多不同的训 练集图像,或未来采用更好的图像增强策略,可能会 提高模型在这 3 个生育期的识别准确率。



图 11 FSST 模型在 4 个生育期上的混淆矩阵

Fig. 11 Confusion matrix of FSST in four growth stages

# 4 结论

(1)将 Channel Shuffle 引入到 FasterNet 中,参数量减少 8.4×10<sup>5</sup>。将 Swin Transformer 加入 FasterNet 后,模型对特征的提取能力明显增强,在测试集上准确率提高 3.68 个百分点。

(2)调整网络结构后,调整 FasterNet 的输出通 道数和基本单元的堆叠次数可以大幅度降低模型复 杂度,提出的 FSST 在小麦生育期图像测试集识别 准确率为 97.22%,并且在 Intel i3 550 CPU 环境下 平均 100 幅为 448 像素 × 448 像素生育期图像推理 时间为 12.54 s。在训练中使用 Lion 优化器替换 AdamW 优化器,能够加速网络收敛,并提升网络推 理精度。

(3) FSST 与 FasterNet、GhostNet、ShuffleNetV2、
MobileNetV3 4 种卷积神经网络模型相比,FSST 模型性能更佳,综合表现更好;基于 FasterNet 改进的
FSST 模型能够兼顾性能和精度。

(4)提出的方法能够快速精准和轻量化地识别 小麦关键生育期,为作物物候特性实时智能监控提 供了技术支持。

#### 参考文献

- [1] 中国人民共和国国家统计局. 中国统计年鉴 2022 [M]. 北京:中国统计出版社,2022.
- [2] 杨蜀秦,林丰山,徐鹏辉,等. 基于无人机遥感影像的多生育期冬小麦种植行检测方法[J]. 农业机械学报,2023,54(2): 181-188.

YANG Shuqin, LIN Fengshan, XU Penghui, et al. Planting pow detection of multi-growth winter wheat field based on UAV remote sensing image [J]. Transactions of the Chinese Society for Agricultural Machinery, 2023, 54(2):181 - 188. (in Chinese)

[3] 周琦,王建军,霍中洋,等. 不同生育期小麦冠层 SPAD 值无人机多光谱遥感估算[J]. 光谱学与光谱分析,2023,43(6): 1912-1920.

ZHOU Qi, WANG Jianjun, HUO Zhongyang, et al. UAV multi-spectral remote sensing estimation of wheat canopy SPAD value in different growth periods [J]. Spectroscopy and Spectral Analysis, 2022, 42(10):3269-3274. (in Chinese)

[4] 兰仕浩,李映雪,吴芳,等. 基于卫星光谱尺度反射率的冬小麦生物量估算[J]. 农业工程学报,2022,38(24):118-128.
 LAN Shihao, LI Yingxue, WU Fang, et al. Winter wheat biomass estimation based on satellite spectral-scale reflectance[J].
 Transactions of the CSAE, 2022, 38(24):118-128. (in Chinese)

YANG Xin, YUAN Ziran, YE Yin, et al. Winter wheat total nitrogen content estimation based on UAV hyperspectral remote sensing[J]. Spectroscopy and Spectral Analysis, 2022,42(10):3269-3274. (in Chinese)

- [6] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition [C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016:770 - 778.
- [7] 徐建鹏, 王杰, 徐祥, 等. 基于 RAdam 卷积神经网络的水稻生育期图像识别[J]. 农业工程学报, 2021, 37(8):143-150.
   XU Jianpeng, WANG Jie, XU Xiang, et al. Image recognition for different developmental stages of rice by RAdam deep convolutional neural networks[J]. Transactions of the CSAE, 2021, 37(8):143-150. (in Chinese)
- [8] 申华磊,苏歆琪,赵巧丽,等. 基于深度学习的无人机遥感小麦倒伏面积提取方法[J]. 农业机械学报,2022,53(9): 252-260,341.

SHEN Hualei, SU Xinqi, ZHAO Qiaoli, et al. Extraction of lodging area of wheat barieties by unmanned aerial vehicle remote sensing based on deep learning[J]. Transactions of the Chinese Society for Agricultural Machinery, 2022, 53(9): 252 - 260, 341. (in Chinese)

[9] 李云霞,马浚诚,刘红杰,等. 基于 RGB 图像与深度学习的冬小麦田间长势参数估算系统[J]. 农业工程学报,2021, 37(24):189-198.

LI Yunxia, MA Juncheng, LIU Hongjie, et al. Field growth parameter estimation system of winter wheat using RGB digital images and deep learning[J]. Transactions of the CSAE, 2021, 37(24):189-198. (in Chinese)

[10] 孙少杰,吴门新,庄立伟,等. 基于 CNN 卷积神经网络和 BP 神经网络的冬小麦县级产量预测[J]. 农业工程学报, 2022,38(11):151-160.

SUN Shaojie, WU Menxin, ZHUANG Liwei, et al. Forecasting winter wheat yield at county level using CNN and BP neural networks[J]. Transactions of the CSAE, 2022,38(11):151-160. (in Chinese)

- [11] TAN M, LE Q. Efficientnet: rethinking model scaling for convolutional neural networks [C] // International Conference on Machine Learning, 2019:6105-6114.
- [12] MA N, ZHANG X, ZHEN H T, et al. Shufflenet v2: practical guidelines for efficient CNN architecture design [C] // Proceedings of the European Conference on Computer Vision, 2018:116-131.
- [13] ANDREW H, ZHU M L, CHEN B, et al. MobileNets: efficient convolutional neural networks for mobile vision applications [J]. arXiv preprint arXiv: 1704.04861,2017,4.
- [14] MARK S, ANDREW H, ZHU Menglong, et al. MobileNetV2: inverted residuals and linear bottlenecks[J]. arXiv preprint arXiv: 1801.04381,2018,1.
- [15] ANDREW H, MARK S, CHU G, et al. Searching for MobileNetV3[J]. arXiv preprint arXiv: 1905.02244,2019,5.
- [16] HAN K, WANG Y H, TIAN Q, et al. GhostNet: more features from cheap operations [J]. arXiv preprint arXiv: 1911. 11907,2019,11.
- [17] CHEN J R, KAO S H, HE H, et al. Run, don't walk: chasing higher FLOPS for faster neural networks [J]. arXiv preprint arXiv: 2303.03667, 2023, 3.
- [18] 杨森森,张昊,兴陆,等. 改进 MobileViT 网络识别轻量化田间杂草[J]. 农业工程学报,2023,39(9):152-160.
   YANG Sensen, ZHANG Hao, XING Lu, et al. Light weight recognition of weeds in the field based on improved MobileViT network[J]. Transactions of the CSAE, 2023,39(9):152-160. (in Chinese)
- [19] 苗荣慧,李志伟,武锦龙. 基于改进 YOLO v7 的轻量化樱桃番茄成熟度检测方法[J]. 农业机械学报,2023,54(10):225-233.
   MIAO Ronghui, LI Zhiwei, WU Jinglong. Lightweight maturity detection of cherry tomato based on improved YOLO v7[J].
   Transactions of the Chinese Society for Agricultural Machinery, 2023, 54(10):225-233. (in Chinese)
- [20] 范晓飞,王林柏,刘景艳,等. 基于改进 YOLO v4 的玉米种子外观品质检测方法[J]. 农业机械学报,2022,53(7):226-233.
   FAN Xiaofei, WANG Linbai, LIU Jingyan, et al. Corn seed appearance quality estimation based on improved YOLO v4[J].
   Transactions of the Chinese Society for Agricultural Machinery, 2022,53(7):226-233. (in Chinese)
- [21] 孙道宗,丁郑,刘锦源,等. 基于改进 SqueezeNet 模型的多品种茶树叶片分类方法[J]. 农业机械学报,2023,54(2):223-230,248.

SONG Daozong, DING Zheng, LIU Jinyuan, et al. Classification method of multi-variety tea leaves based on improved SqueezeNet model[J]. Transactions of the Chinese Society for Agricultural Machinery, 2023,54(2):223 - 230,248. (in Chinese)

- [22] CHEN X N, CHEN L, DA H, et al. Symbolic discovery of optimization algorithms [J]. arXiv preprint arXiv: 2302.06675, 2023,2.
- [23] SACHIN M, MOHAMMAD R. MobileViT: light-weight, general-purpose, and mobile-friendly vision transformer[J]. arXiv preprint arXiv: 2110.02178, 2021, 5.
- [24] DENG J, DONG W, SOCHER R, et al. ImageNet: a large-scale hierarchical image database [C] // IEEE Conference on Computer Vision and Pattern Recognition. IEEE Computer Society, 2009:248 - 255.
- [25] KINGMA D, BA J. Adam: a method for stochastic optimization [J]. arXiv preprint arXiv: 1412.6980, 2014, 12.

- [26] LIU L Y, JIANG H M, HE P C, et al. On the variance of the adaptive learning rate and beyond [J]. arXiv preprint arXiv: 1908.03265, 2019, 8.
- [27] ILYA L, FRANK H. Decoupled weight decay regularization [J]. arXiv preprint arXiv: 1711.05101, 2017, 11.
- [28] 宋怀波,华志新,马宝玲,等. 基于 SimCC ShuffleNetV2 的轻量化奶牛关键点检测方法[J]. 农业机械学报, 2023, 54(10): 275-281,363.

SONG Huaibo, HUA Zhixin, MA Baoling, et al. Lightweight keypoint detection method of dairy cow based on SimCC – ShuffleNetV2[J]. Transactions of the Chinese Society for Agricultural Machinery, 2023, 54(10): 275 – 281, 363. (in Chinese)

[29] 陈丽,王世勇,高思莉,等. Sentinel-2卫星的多光谱轻量级船舶目标检测算法[J]. 光谱学与光谱分析,2022,42(9): 2862-2869.

CHEN Li, WANG Shiyong, GAO Sili, et al. Multispectral lightweight ship target detection algorithm for Sentinel - 2 satellite [J]. Spectroscopy and Spectral Analysis, 2022, 42(9): 2862 - 2869. (in Chinese)

[30] 赵辉,黄镖,王红君,等. 基于改进 YOLO v7 的农田复杂环境下害虫识别算法研究[J]. 农业机械学报,2023,54(10): 246-254.

ZHAO Hui, HUANG Biao, WANG Hongjun, et al. Pest identification method in complex farmland environment based on improved YOLO v7 [J]. Transactions of the Chinese Society for Agricultural Machinery, 2023, 54 (10): 246 - 254. (in Chinese)

#### (上接第 206 页)

- [24] WU J, LI B, NI W, et al. An adaptively weighted multi-feature method for object-based change detection in high spatial resolution remote sensing images[J]. Remote Sensing Letters, 2020, 11(4): 333 - 342.
- [25] NIU X, GONG M, ZHAN T, et al. A conditional adversarial network for change detection in heterogeneous images [J]. IEEE Geoscience and Remote Sensing Letters, 2019, 16(1): 45 - 49.
- [26] PATRA S K, SHEKHER M, SOLANKI S S, et al. A technique for generating natural colour images from false colour composite images[J]. International Journal of Remote Sensing, 2006, 27(14): 2977 - 2989.
- [27] 张祖勋,姜慧伟,庞世燕,等. 多时相遥感影像的变化检测研究现状与展望[J]. 测绘学报, 2022, 51(7): 1091-1107. ZHANG Zuxun, JIANG Huiwei, PANG Shiyan, et al. Review and prospect in change detection of multi-temporal remote sensing images[J]. Acta Geodaetica et Cartographica Sinica, 2022, 51(7): 1091-1107. (in Chinese)
- [28] ZHU Q Q, GUO X, LI Z Q, et al. A review of multi-class change detection for satellite remote sensing imagery [J]. Geospatial Information Science, 2022, 27(1): 1-15.
- [29] 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), 2016: 770-778.
- [30] 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.
- [31] ZHAO H, SHI J, QI X, et al. Pyramid scene parsing network [C] // 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017: 6230 - 6239.
- [32] CHEN L, PAPANDREOU G, KOKKINOS I, et al. DeepLab: semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2018, 40(4): 834-848.
- [33] HUANG Z, WANG X, WEI Y, et al. CCNet: criss-cross attention for semantic segmentation [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2023, 45(6): 6896-6908.