

基于多特征融合的水稻生长期识别模型

陈明^{1,2} 周世杰^{1,2} 金擎³ 张蕾³ 吴振阳³ 王振华^{1,2}

(1. 上海海洋大学信息学院, 上海 201306; 2. 农业农村部渔业信息重点实验室, 上海 201306;

3. 上海市浦东新区农业农村综合监管事务中心, 上海 201401)

摘要: 我国是全球最大的水稻种植国, 水稻不同生长期的精准识别对于实现稻田智能化管理至关重要。以水稻生长的始穗期、齐穗期、乳熟期和黄熟期为识别目标, 提出了一种基于多特征融合的水稻生长期识别模型。以无人机获取的高分辨率遥感影像作为数据源, 分别采用 Swin Transformer 模型和 ResNet18 模型提取了影像的多尺度特征和深层语义特征; 设计了一种改进的特征融合模块, 包括自适应特征融合 (AFF) 模块和全局上下文建模 (GC) 模块。AFF 模块通过学习特征之间的权重关系, 实现不同特征的自适应融合。GC 模块引入全局上下文信息, 增强模型对全局特征的感知能力, 提高模型对复杂场景的适应性; 采用 Lion 优化器和动态学习率策略, 加速模型的收敛速度。同时, 引入 Dropout 层以防止过拟合, 提高模型的泛化能力。对比实验表明, 本文提出的模型在水稻抽穗期和成熟期识别任务上取得了 97.14% 的准确率, 相较于 MobileNetV2、EfficientNet 和 Swin Transformer 分别提高了 3.29、2.53、1.09 个百分点。在保证高准确率的同时, 该模型的参数量和计算量相对较少, 具有较好的轻量化特性, 适用于实际应用场景。

关键词: 水稻; 生长期; Swin Transformer; ResNet18; 识别模型; 多特征融合

中图分类号: S511; S127

文献标识码: A

文章编号: 1000-1298(2026)11-0325-09

OSID:



Multi-feature Fusion-based Model for Recognizing Different Growth Stages of Rice

CHEN Ming^{1,2} ZHOU Shijie^{1,2} JIN Qing³ ZHANG Lei³ WU Zhenyang³ WANG Zhenhua^{1,2}

(1. College of Information Technology, Shanghai Ocean University, Shanghai 201306, China

2. Key Laboratory of Fisheries Information, Ministry of Agriculture and Rural Affairs, Shanghai 201306, China

3. Shanghai Pudong New Area Agriculture and Rural Comprehensive Supervision Affairs Center, Shanghai 201401, China)

Abstract: China is the largest rice-growing country in the world, and accurate recognition of different growth stages of rice is essential for achieving intelligent rice field management. Targeting the classification of rice growth stages, including initial heading stage, full heading stage, milk ripening stage, and yellow ripening stage, a recognition model for rice growth stages was proposed based on multi-feature fusion. The specific methodology included multi-feature extraction module, utilizing high-resolution remote sensing images obtained from drones as the data source, multi-scale features and deep semantic features of the images were extracted by using the Swin Transformer model and ResNet18 model. An improved feature fusion module was designed, which included an adaptive feature fusion (AFF) module and a global context modeling (GC) module. The AFF module achieved adaptive fusion of different features by learning the weight relationships between features. The GC module incorporated global context information to enhance the model's ability to perceive global features, thereby improving adaptability in complex scenarios. Model optimization, utilizing the Lion optimizer and a dynamic learning rate strategy to accelerate the model's convergence speed. Additionally, a Dropout layer was introduced to prevent overfitting and improve the model's generalization ability. Comparative experiments demonstrated that the proposed model achieved a 97.14% accuracy rate in recognizing rice at the heading and maturity stages, which was 3.29, 2.53, and 1.09 percentage points higher than that of other mainstream models such as MobileNetV2, EfficientNet and Swin Transformer, respectively. Notably,

收稿日期: 2025-01-24 修回日期: 2025-03-14

基金项目: 山东省重点研发计划项目 (2023TZX051)

作者简介: 陈明 (1966—), 男, 教授, 博士, 主要从事计算机技术研究, E-mail: mchen@shou.edu.cn

while maintaining high accuracy, the model had relatively fewer parameters and lower computational costs, showcasing lightweight characteristics suitable for practical applications.

Key words: rice; growth stages; Swin Transformer; ResNet18; recognition model; multi-feature fusion

0 引言

水稻在农业生产和保障粮食安全方面发挥着关键作用^[1-2]。准确识别水稻的不同生长阶段对农业生产具有重要意义^[3]。水稻的生长周期大致分为幼苗期、分蘖期、抽穗期和成熟期4个阶段,每个阶段的生长特点和管理需求各不相同^[4]。精准管理和调控是实现水稻高产优质的关键,而准确识别生长阶段是前提^[5]。

传统的水稻生长阶段监测主要依赖人工进行,但受限于人为因素和环境条件的影响,不同年份的生长阶段起始时间存在较大不确定性。此外,人工识别方法效率低下、主观性强,难以满足现代农业对大规模、精准化管理的需求。

无人机遥感在农业生产中的广泛应用为作物生长阶段的自动化监测提供了重要的数据源^[6-7]。众多学者已将无人机遥感与深度学习技术相结合,开展了丰富的水稻生长阶段识别研究^[8-16]。

深度学习技术能够将无人机获取的大量影像数据转化为可用于精准农业管理的宝贵信息,显著提升了农业生产效率和资源利用率,加速了智慧农业的发展^[17]。现有研究表明,深度学习模型在水稻生长阶段识别方面取得了较好的效果^[18-21]。

然而,当前的水稻生长阶段识别研究仍存在一些挑战,包括:现有方法多基于单一图像数据或单一特征提取模型,无法充分利用多源数据信息;不同类型数据的融合以及特征的有效融合也需要

进一步研究;模型的选择和优化也存在一定难度;在影像获取过程中,振动等因素造成影像模糊。针对上述问题,本文提出一种基于多特征融合的水稻生长阶段识别模型。该模型通过 Swin Transformer 和 ResNet18 模型分别提取影像的全局和局部特征,并通过改进的特征融合模块实现不同特征的有效融合,从而更全面地捕捉水稻生长阶段的特征。此外采用 Lion 优化器和 Dropout 技术,提高模型的泛化能力。

1 实验数据与实验设置

1.1 实验数据

图1所示为实验区域和水稻田示意图。实验区域为上海市临港某水稻田,面积约4 800 hm²;水稻品种分别为花优14、弘优26、松早香1号、南粳46、秀水613、沪软1212、秋优金丰和申优28^[22]。数据采集设备为大疆无人机精灵4(型号 DJI P4M),包含1个可见光成像仪和5个多光谱成像仪(蓝光、绿光、红光、红边和近红外),数据空间分辨率为6 cm;数据采集时间为2022年9—12月。

1.2 数据集构建

在水稻的生长过程中,抽穗期(从始穗期到齐穗期)标志着水稻从营养生长向生殖生长的转变。抽穗期是水稻生长发育的关键时期^[23-24]。精准的灌水、施肥和病虫害防治对确保水稻正常生长至关重要^[25]。成熟期(从乳熟期到黄熟期)是籽粒充实和积累养分的阶段。乳熟期籽粒呈乳白色,逐渐变

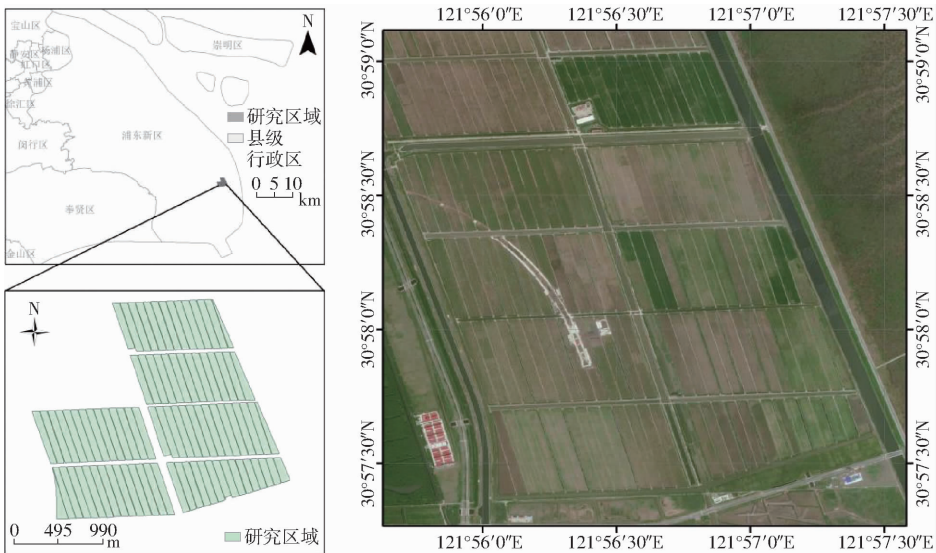


图1 水稻实验田示意图

Fig. 1 Schematic of rice experimental field

硬;黄熟期谷壳变黄,籽粒完全成熟。精准用水、及时断水与收割是水稻丰收的关键^[26-27]。本文选取始穗期、齐穗期、乳熟期和黄熟期 4 个生长期作为识别对象,研究基于无人机遥感监测的水稻不同生长期的识别模型。

以无人机获取的高分辨率遥感影像作为数据源,选取红光(R)、绿光(G)、蓝光(B)组成 RGB 影像作为多通道数据输入,其中红光(R)的中心波长范围为(650±16) nm;绿光(G)的中心波长范围为(560±16) nm;蓝光(B)的中心波长范围为(450±16) nm。选取近红外和红光两个波段进行波段计算,得到归一化植被指数 NDVI (Normalized

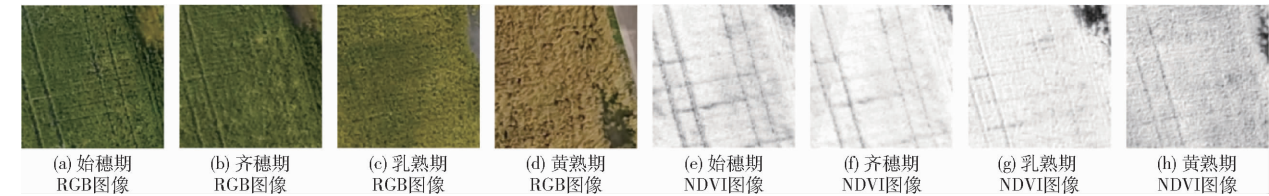


图2 水稻 4 个生长期示意图

Fig. 2 Images of four rice growth stages

表 1 水稻不同生长期的数据集

Tab.1 Data of rice growth stage

水稻不同生长期	数据类型	训练集数据数	验证集数据数
始穗期	RGB 多通道	1 250	300
	NDVI 单通道	1 250	300
齐穗期	RGB 多通道	1 250	300
	NDVI 单通道	1 250	300
乳熟期	RGB 多通道	1 250	300
	NDVI 单通道	1 250	300
黄熟期	RGB 多通道	1 250	300
	NDVI 单通道	1 250	300

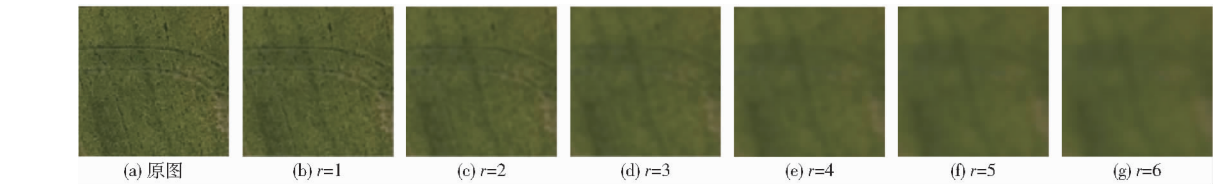


图3 不同模糊半径的高斯模糊效果

Fig. 3 Gaussian blur with different blur radius

与真实采集的模糊水稻影像进行相似度量。哈希算法包括均值哈希、差值哈希和感知哈希。如表 2 所示,将基于 6 种不同模糊半径生成的模糊影像与原始影像进行逐一对比,并通过哈希算法计算相似度。相似度数范围 0~1,其中 0 表示两幅影像差异极大,1 表示两幅影像几乎完全一致。

1.4 评价指标

采用准确率、精确率、召回率和 F1 分数来评价本文模型的识别能力。同时,使用参数量、浮点计算量和 100 幅影像的平均识别时间作为模型复杂度衡

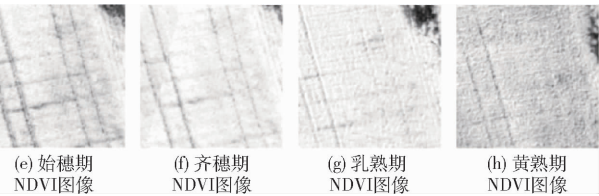
量指标。其中,参数量衡量的是网络中权重和偏置的总数;浮点计算量表示模型在处理一次前向传播

difference vegetation index),将 NDVI 指数作为单通道数据输入,公式为

$$N_{NDVI} = \frac{N_{NIR} - R}{N_{NIR} + R} \tag{1}$$

式中 N_{NIR} ——近红外波长,取(840±26) nm
 R ——红色波长

针对水稻不同生长期的数据,分别进行裁剪,得到 1 250 幅 RGB 影像和对应的 NDVI 影像,图 2 为水稻 4 个生长期示意图。裁剪后的尺寸为 224 像素×224 像素。按照比例 8:2 将数据集随机分为训练集和验证集,确保训练集和验证集之间无重叠。表 1 所示为水稻不同生长期的数据集。



1.3 模糊测试集的构建

为了评估模型在复杂拍摄环境下的识别精度以及对影像模糊的鲁棒性,采用高斯模糊对原始 RGB 水稻影像进行处理,模拟因相机抖动、聚焦不准等因素导致的影像模糊。图 3 为对影像应用了 6 种不同半径 r 的高斯模糊,生成 6 个模糊程度递增的影像数据集。通过在这些数据集上测试模型,评估其不同模糊程度下的性能表现。

为了客观评估模糊影像处理的效果,并避免主观视觉判断带来的偏差,采用哈希算法对模糊影像

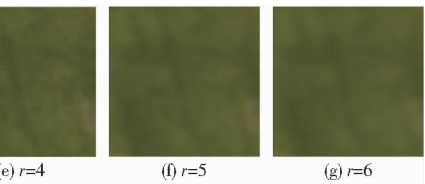


表 2 高斯模糊后影像的相似对比度
Tab.2 Similarity contrast of image after Gaussian blurring

哈希函数	r					
	1	2	3	4	5	6
均值哈希	0.95	0.92	0.93	0.89	0.85	0.84
差值哈希	0.92	0.89	0.84	0.79	0.74	0.72
感知哈希	0.96	0.97	0.92	0.91	0.87	0.82

量指标。其中,参数量衡量的是网络中权重和偏置的总数;浮点计算量表示模型在处理一次前向传播

中所需的总浮点运算量。

采用退化指数、平均退化指数、退化波动指数来评价基于模糊数据集的模型退化性能。

退化指数表示模型准确率下降程度,计算公式为

$$d_i = A - A_i \tag{2}$$

式中 A_i ——第 i 个模糊等级的准确率
 d_i ——第 i 个模糊等级的退化指数

平均退化指数 $\bar{d}$ 表示平均性能的退化程度,计算公式为

$$\bar{d} = \frac{A - A_6}{6} \tag{3}$$

退化波动指数 s 表示整体退化性能的波动程度,计算公式为

$$s = \sqrt{\frac{\sum_{i=1}^6 (d_i - \bar{d})^2}{5}} \tag{4}$$

1.5 实验参数

批量大小设置为 32;使用 Lion 优化器,初始学习率为 0.000 2,权重衰减为 0.01,设置动态学习率减少倍率为 0.1,同时设置在验证性能没有提升的情况下等待的迭代次数为 5;训练迭代次数为 100。模型使用 PyTorch 2.0.0 + cu 118 为框架,通过 Torch 调用 GPU 来对网络进行训练;训练环境为 Windows 11,硬件环境为 NVIDIA RTX 3080 GPU。采用 Lion 优化器^[28]对模型的参数进行优化。为缓解过拟合的发生,在全连接层网络采用随机失活 (Dropout) 参数设置为 0.5。

2 水稻不同生长期的识别模型

图 4 所示为本文提出的多特征融合的水稻不同生长期识别模型架构图。该模型包括:①多特征提取模块:包含两个特征提取部分,分别为 Swin Transformer^[29] 和 ResNet18^[30],其中, Swin Transformer 用来提取多通道 RGB 影像特征; ResNet18 用来提取单通道的 NDVI 影像特征。②特征融合模块:包含改进后的自适应特征融合 (AFF) 模块,使用该模块对提取后的两个特征进行融合。③分类模块:包含 Dropout 层和全连接层,对影像进行分类识别。

2.1 多特征提取模块

Swin Transformer 凭借强大的全局建模能力,可从多通道遥感影像中有效提取全局上下文信息。本文 RGB 影像包含丰富的颜色和纹理信息,适于捕捉作物的整体生长状态。为全面提取特征,本文引入并改造 ResNet18 模型,使其可直接处理单通道

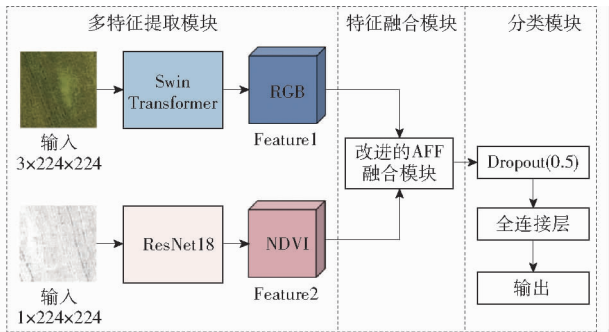


图 4 多特征融合的水稻生长期识别模型架构图
Fig. 4 Architecture diagram of rice growth stages recognition model with multi-feature fusion

NDVI 影像,捕捉水稻生长的局部细节特征。

Swin Transformer 基于 Transformer 架构,在全局上下文信息捕获能力及长距离依赖关系建模方面优势显著。RGB 图像的颜色、纹理等全局特征对水稻生长阶段的识别至关重要,尤其能有效把握水稻从抽穗期过渡至成熟期的颜色演变过程。而 ResNet18 通过残差连接机制,在局部细节特征提取方面表现出色。NDVI 单通道植被指数主要反映植被的生理状况,更多地体现在局部特征上,如叶尖的形态变化或特定区域内的叶绿素分布状况。ResNet18 的深层卷积结构能够高效地捕捉这些细微的局部特征,同时保持较低的计算复杂度。

考虑到 NDVI 单通道数据信息量相对单一,采用 Swin Transformer 处理可能引入冗余计算。因此需结合数据特性与计算效率,实现最优的模型性能。

为充分发挥两种模型的优势,本文设计了一个融合 Swin Transformer 和 ResNet 的特征提取模块。为了减轻模型计算负担,对 Swin Transformer 进行剪枝优化,将各阶段层数调整为 1、1、6、1,在降低计算复杂度的同时,最大程度保留其全局建模能力。

2.2 特征融合模块

本文提出了一种改进的自适应特征融合 (AFF) 模块,用于对提取的两种特征进行有效融合。AFF (Adaptive feature fusion) 模块是一种自适应特征融合模块,通过为不同特征图分配权重,实现灵活且精确的特征融合。传统的 AFF 模块主要关注局部特征的权重分配,在处理全局上下文信息方面存在一定的局限性。为了增强模型对全局信息的感知能力,本文在 AFF 模块基础上引入了一个全局上下文 (Global context, GC) 模块。

GC 模块是一种注意力机制模块,GC 模块通过捕捉全局上下文信息,可以更好地感知特征之间的长距离依赖关系。GC 模块是一种注意力机制模块,通过聚合全局上下文信息,增强对全局信息的感知

能力。改进后的 AFF 模块结构如图 5 所示,GC 模块的结构如图 6 所示。

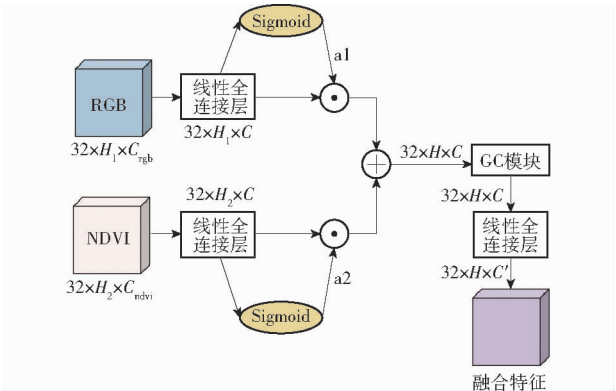


图 5 改进的 AFF 模块结构图

Fig. 5 Structure diagram of improved AFF module

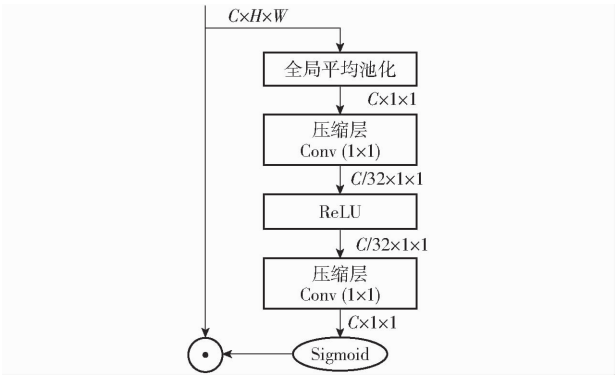


图 6 GC 模块结构图

Fig. 6 Structure diagram of GC (global context) module

2.3 模型优化

为了提升模型的泛化能力并加速训练过程,本文采用了 Lion 优化器。Lion 优化器是一种新型的优化算法,通过符号操作更新模型参数,相比于传统的 Adam 和 SGD 优化器,Lion 优化器具有更快的收敛速度和更强的鲁棒性。

考虑到 Lion 优化器对学习率较为敏感,本文同时引入了动态学习率调整策略,以使模型在训练过程中能够自适应地调整学习率,从而在保证收敛速度的同时提高模型的泛化性能。此外,为了防止过拟合,本文在模型中加入了 Dropout 层,通过随机丢弃神经元来减少模型对训练数据的过拟合。Dropout 层能够有效地提高模型的泛化能力,使其在未见过的样本上表现更好。

3 实验结果与分析

3.1 不同融合方法的对比实验

为验证多源数据融合在提升模型性能方面的优势,进行对比实验,仅采用 RGB 或 NDVI 单模态数据作为模型输入。根据表 3 所示的结果,当模型仅输入 RGB 数据时,其准确率达到 94.23%;而仅

表 3 模型不同输入的对比结果

Tab. 3 Comparison results of different model inputs

模型输入	%		
	准确率	精确率	召回率
仅 RGB 输入	94.23	93.57	94.28
仅 NDVI 输入	91.56	90.25	91.87
本文模型	97.14	97.25	97.34

输入 NDVI 数据时,准确率降至 91.56%。相比之下,当融合 RGB 和 NDVI 两种数据的特征后,模型的准确率显著提升至 96.94%。结果表明,RGB 与 NDVI 特征在模型中具有显著的互补性:RGB 图像为模型提供了丰富的色彩和纹理信息,有助于识别水稻生长阶段的不同视觉特征;而 NDVI 指数则量化了植被的生理状态,反映了水稻的生长活力和健康状况。通过融合这两种数据源,模型能够更全面地捕捉水稻生长过程中的关键信息,从而显著增强其对水稻生长阶段的判别能力。

为探究不同特征融合方法对模型性能的影响,本文分别采用拼接和加和两种传统方法以及改进的 AFF 模块进行对比实验。拼接方法将多个特征图沿着通道维度进行拼接,以获取更丰富的特征表示;加和方法则通过逐元素相加的方式融合特征,操作简单但可能损失部分特征信息。改进的 AFF 模块通过学习不同特征图的权重,自适应地融合特征,并引入全局上下文信息,以更好地捕捉特征之间的长距离依赖关系。如表 4 所示,改进的 AFF 模块在准确率上分别比拼接和加和方法提高了 1.83 个百分点和 2.03 个百分点。虽然改进的 AFF 模块相较于拼接和加和方法引入了更多的参数和计算量(分别增加了 1.6×10^6 和 8×10^7),但其在特征融合方面表现出了更好的效果,证明了引入权重分配和全局上下文信息的有效性。实验结果表明,改进的 AFF 模块能够更好地捕捉特征之间的复杂关系,从而提升模型的性能。

表 4 不同融合方法的对比结果

Tab. 4 Comparison of results by using different

特征融合方法	fusion methods				
	准确 率/%	精确 率/%	召回 率/%	参数量	浮点计 算量
拼接	95.11	96.14	95.45	3.14×10^7	5.18×10^9
加和	94.91	96.25	96.12	3.14×10^7	5.18×10^9
本文模型	96.94	96.75	97.16	3.30×10^7	5.26×10^9
融合模块					

3.2 消融实验

为深入探究改进的 AFF 模块、Lion 优化器以及模型压缩策略对模型性能的影响,本文进行了消融实

验。实验结果如表 5 所示。在引入 AFF 模块后,模型的准确率略有提升,其中 GC 模块的引入对模型性能的提升更为显著,相比基线模型,准确率提高了 1.33 个百分点。虽然 Lion 优化器在准确率上没有带来明显的提升,但其显著加快了模型的收敛速度,如图 7 所示。同时,Lion 优化器在处理稀疏梯度和噪声

梯度方面表现出了更好的鲁棒性。通过模型压缩,本文成功将模型的参数量和计算量分别降低了 7.2×10^6 和 9.5×10^8 ,而准确率的损失相对较小。实验结果表明,改进的 AFF 模块、Lion 优化器以及模型压缩策略在不显著降低模型性能的前提下,有效提升了模型的训练效率和部署性能。

表 5 消融实验结果
Tab.5 Ablation experiments

模块				性能指标			
AFF 模块	GC 模块	Lion 优化器	模型轻量化	准确率/%	参数量	浮点计算量	平均识别时间/s
×	×	√	×	94.81	3.86×10^7	6.13×10^9	3.24
×	×	√	√	94.34	3.14×10^7	5.18×10^9	3.67
√	×	√	√	95.78	3.30×10^7	5.26×10^9	3.14
√	√	×	√	96.94	3.31×10^7	5.26×10^9	3.14
√	√	√	√	97.01	3.31×10^7	5.26×10^9	3.13

注:√表示采用此模块;×表示未采用此模块。

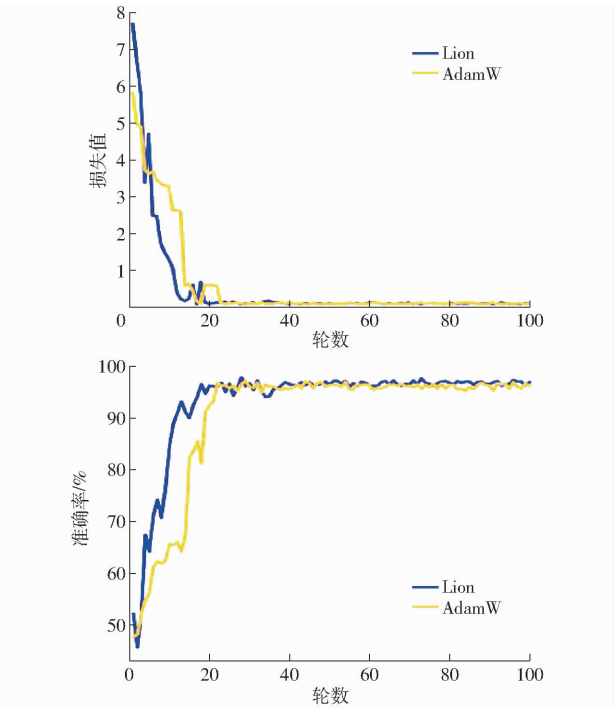


图 7 Lion 优化器与 AdamW 优化器的损失值与准确率曲线

Fig.7 Loss and accuracy curves of Lion optimizer and AdamW optimizer

3.3 不同模型的对比实验

为评估本文所提出模型的性能表现,本研究将其与引言部分所述的相关研究工作中的模型进行了全面的性能对比。对比模型涵盖了机器学习领域的经典算法(即 K 近邻算法(KNN)和支持向量机(SVM))、基于 RAdam 优化器的 ResNet50 深度学习模型,以及采用特征金字塔(FPN)结构的 YOLO v3 目标检测模型。鉴于所有参与对比的生长期识别模型均采用单一模态的输入数据,为确保实验结果的公正性和可比性,本研究统一采用自行构建的 RGB

数据集对各模型进行训练与测试。如表 6 所示,实验结果表明:相较于机器学习模型,本文所提模型在各项评价指标上均表现出更优的性能;在作物生长期的卷积神经网络模型中,本文模型有效地利用了多源数据的融合优势;而在作物生长期识别模型的应用范畴内,本文模型更展现出其在大面积作物生长期识别任务中的适用性和优越性。

表 6 不同作物生长期识别模型的对比实验结果

stage recognition models				%
模型	准确率	精确率	召回率	F1 分数
KNN	87.13	85.46	87.79	86.61
SVM	88.63	86.97	85.58	86.27
基于 RAdam 的 ResNet50 ^[16]	93.07	93.24	92.72	92.98
基于 FPN 的 YOLO v3 ^[10]	90.85	90.90	90.31	90.60
本文模型	97.14	97.25	97.34	97.30

为了全面评估本文所提模型的性能,除了与主流的作物生长期识别模型进行比较之外,将其与目前轻量级网络 MobileNetV2^[31]、高效网络 EfficientNet^[32] 以及近年来备受关注的视觉 Transformer 模型 Swin Transformer 进行了对比实验。考虑到各对比模型的输入均为单一模态,在实验中统一采用本文构建的 RGB 数据集进行训练和测试,以确保实验结果的公平性。

实验结果如表 7 所示。结果表明,本文模型在准确率、精确率、召回率和 F1 分数等多个评价指标上均显著优于对比模型,分别达到了 97.14%、97.25%、97.34% 和 97.30%。这一结果充分说明,相比于仅依赖单一特征提取的模型,本文模型通过融合 SwinTransformer 在全局特征提取方面的优势

表 7 不同模型的对比实验结果

Tab.7 Comparative experiments of different models

模型	%			
	准确率	精确率	召回率	F1 分数
MobileNetV2	93.85	94.55	94.54	94.55
EfficientNet	94.61	94.68	94.48	94.58
Swin Transformer	96.05	95.42	95.82	95.62
本文模型	97.14	97.25	97.34	97.30

和 ResNet18 在局部特征提取方面的优势,实现了对水稻生长阶段的更准确、全面的识别。

为了深入分析模型的识别性能,本文绘制了混淆矩阵(如图 8 所示)。从混淆矩阵可以看出,模型在始穗期和齐穗期这两个相邻生长阶段的识别错误率相对较高。这可能是由于这两个生长阶段的视觉特征差异较为细微,且在时间上较为接近,导致模型难以准确区分。

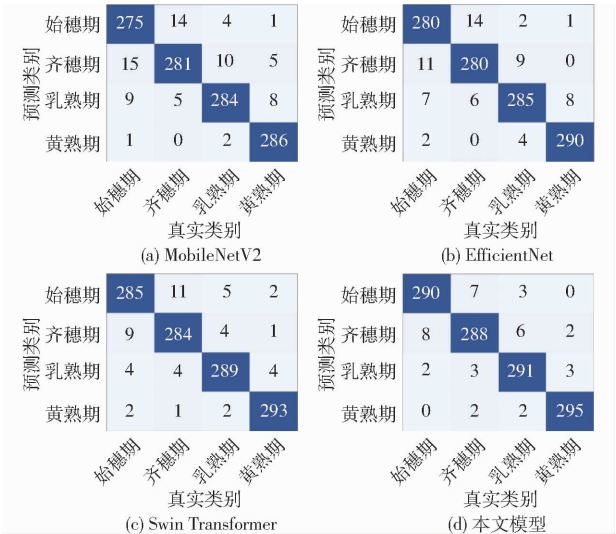


图 8 混淆矩阵

Fig. 8 Confusion matrix

为了深入了解各模型对不同生长期水稻特征的识别能力,本文对模型最后一层卷积层的输出进行了可视化分析,并绘制了相应的热力图。在热力图中,红色区域代表模型对该区域特征的高度关注,黄色区域则表示关注度较低。如图 9 所示,本文模型的热力图显示出更广泛且集中的高激活区域,表明其对不同生长期水稻特征的学习更为全面。相比之下,其他模型的热力图中高激活区域的分布相对分散,表明其对特征的关注不够集中。这种差异说明本文模型能够更有效地提取到影像中与水稻生长阶段密切相关的关键特征。此外,通过对比不同模型的热力图,可以观察到各模型在特征提取方面的侧重点不同,本文模型与其他模型在特征提取上具有互补性,共同促进了模型整体性能的提升。

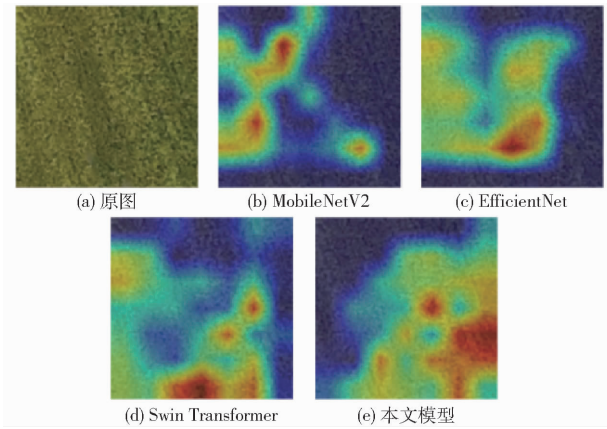


图 9 各个模型的热力图对比

Fig. 9 Comparison of heat maps for various models

3.4 不同水稻品种的识别性能分析

为评估模型针对不同水稻品种的鲁棒性,本文选取了包含花优 14、弘优 26 在内的 8 个水稻品种进行了识别准确率的独立测试。鉴于实验数据集中秋优金丰、申优 28 及南粳 46 的品种占比较高,本研究将分类层级设定为 4 类,以更精准地反映模型性能。如表 8 所示,模型在各品种上的平均识别准确率达到 96.85%,这一结果充分显示了模型对水稻品种间差异的鲁棒性。

表 8 不同水稻品种的对比结果

Tab.8 Comparison results of different rice varieties

水稻品种	%		
	准确率	精确率	召回率
秋优金丰	96.84	96.42	96.48
申优 28	96.73	97.17	96.89
南粳 46	97.11	96.37	96.51
其余 5 个品种	96.73	96.68	96.62

3.5 基于模糊数据集的模型鲁棒性分析

为了评估各模型在实际应用中对影像质量变化的鲁棒性,在水稻生长期识别任务中引入了不同程度的模糊影像。如表 9 所示,随着模糊半径的增加,所有模型的识别准确率均呈现下降趋势。然而,本文提出的模型在不同模糊水平下均表现出优异的性能,尤其在模糊程度最为严重的 $r=6$ 时,仍能保持

表 9 模型对于不同模糊半径影像的准确率

Tab.9 Accuracy of model for images with different

	radius of blurriness					%
模型	r					
	1	2	3	4	5	6
MobileNetV2	91.67	85.84	79.22	54.58	53.63	53.71
EfficientNet	91.11	81.24	78.65	52.15	51.54	52.97
Swin Transformer	92.71	87.33	81.98	64.85	57.47	57.17
本文模型	95.54	94.37	90.73	87.25	83.25	82.75

82.75%的高准确率,显著优于其他对比模型,充分体现了其具有较好的鲁棒性。

如表 10 所示,将本文模型与 MobileNetV2、EfficientNet、Swin Transformer 等模型进行了鲁棒性对比。结果表明,在退化指数 d_6 、平均退化指数以及退化波动指数 s 这 3 个指标上,本文模型均显著优于其他对比模型。具体而言,本文模型的退化指数 d_6 仅为 14.39,平均退化指数为 2.39,退化波动指数为 4.99,远低于其他模型。这表明,当输入影像质量下降时,本文模型的性能衰减幅度明显较小。对比之下, MobileNetV2、EfficientNet 和 Swin Transformer 等模型在面对模糊影像时,性能退化较

为严重。

分析可知,本文模型之所以能够在影像模糊的情况下仍保持较好的性能,主要得益于其独特的多特征提取和特征融合模块。这些模块使得模型能够更有效地提取影像中的关键信息,并对噪声和模糊具有更强的鲁棒性,从而提高了模型的泛化能力。

4 结束语

随着无人机技术的快速发展,利用无人机获取高分辨率遥感影像已成为农业生产中获取农作物生长信息的重要手段。准确识别水稻生长期的关键时期(始穗期、齐穗期、乳熟期和黄熟期)对于指导农业生产、提高粮食产量具有重要意义。本文提出了一种基于深度学习的无人机影像水稻生长阶段识别模型。通过充分利用 Swin Transformer 和 ResNet18 的优势,并结合改进的特征融合模块和优化算法,实现了高效准确的水稻生长阶段识别。该模型为智慧农业的发展提供了新的思路。

表 10 模型的鲁棒性对比

Tab.10 Comparison of model's performance degradation

模型	退化指数 d_6	平均退化指数	退化波动指数 s
MobileNetV2	40.14	6.69	16.05
EfficientNet	41.64	6.94	16.18
Swin Transformer	38.88	6.48	14.32
本文模型	14.39	2.39	4.99

参 考 文 献

[1] 徐春春, 闻军清, 纪龙, 等. 中国水稻种业发展现状、问题与展望[J]. 中国稻米. 2022, 28(5): 74–78.
XU Chunchun, WEN Junqing, JI Long, et al. Current situations, problems and prospects of rice seed industry in China[J]. China Rice, 2022, 28(5): 74–78. (in Chinese)

[2] 王跃星, 魏祥进, 徐春春, 等. 我国水稻种业发展现状与对策浅析[J]. 中国稻米. 2022, 28(5): 62–65.
WANG Yuexing, WEI Xiangjin, XU Chunchun, et al. Current situation and countermeasures of rice breeding and seed industry development in China[J]. China Rice, 2022, 28(5): 62–65. (in Chinese)

[3] 岑海燕, 朱月明, 孙大伟, 等. 深度学习在植物表型研究中的应用现状与展望[J]. 农业工程学报, 2020, 36(9): 1–16.
CEN Haiyan, ZHU Yueming, SUN Dawei, et al. Current status and future perspective of the application of deep learning in plant phenotype research[J]. Transactions of the CSAE, 2020, 36(9): 1–16. (in Chinese)

[4] MA Y, JIANG Q, WU X, et al. Monitoring hybrid rice phenology at initial heading stage based on low-altitude remote sensing data[J]. Remote Sensing, 2021, 13(1): 86.

[5] VITHU P, MOSES J A. Machine vision system for food grain quality evaluation: a review[J]. Trends in Food Science & Technology, 2016, 56: 13–20.

[6] ZHANG C, KOVACS J M. The application of small unmanned aerial systems for precision agriculture: a review[J]. Precision Agriculture, 2012, 13(6): 693–712.

[7] INOUE Y, SAKAIYA E, WANG C. Capability of C-band backscattering coefficients from high-resolution satellite SAR sensors to assess biophysical variables in paddy rice[J]. Remote Sensing of Environment, 2014, 140: 257–266.

[8] ZHENG H, CHENG T, LI D, et al. Evaluation of RGB, color-infrared and multispectral images acquired from unmanned aerial systems for the estimation of nitrogen accumulation in rice[J]. Remote Sensing, 2018, 10(6): 824.

[9] 刘平, 刘立鹏, 王春颖, 等. 基于机器视觉的田间小麦开花期判定方法[J]. 农业机械学报. 2022, 53(3): 251–258.
LIU Ping, LIU Lipeng, WANG Chunying, et al. Determination method of field wheat flowering period baesd on machine vision [J]. Transactions of the Chinese Society for Agricultural Machinery, 2022, 53(3): 251–258. (in Chinese)

[10] 赵春江, 文朝武, 林森, 等. 基于级联卷积神经网络的番茄花期识别检测方法[J]. 农业工程学报, 2020, 36(24): 143–152.
ZHAO Chunjiang, WEN Chaowu, LIN Sen, et al. Tomato florescence recognition and detection method based on cascaded neural network[J]. Transactions of the CSAE, 2020, 36(24): 143–152. (in Chinese)

[11] ZHAO L, GUO W, WANG J, et al. An efficient method for estimating wheat heading dates using UAV images[J]. Remote Sensing, 2021, 13(16): 3067.

[12] GE H X, MA F, LI Z W, et al. Improved accuracy of phenological detection in rice breeding by using ensemble models of machine learning based on UAV–RGB Imagery[J]. Remote Sensing, 2021, 13(14): 2678.

[13] MARSUJITULLA H, ZAINUDDIN Z, MANJANG S, et al. Rice farming age detection use drone based on SVM histogram image classification [C] // Proceedings of the 1st Symposium of Emerging Nuclear Technology and Engineering Novelty (SENTEN)-Discovering Science and Engineering Novelty for Improving Human Life Prosperity. Palembang, 2019.

[14] 张长胜, 李得恺, 杨忠义, 等. 基于深度特征局部重采样融合的多种类水稻种子识别[J]. 农业机械学报, 2025, 56(7):

522 – 531.

ZHANG Changsheng, LI Dekai, YANG Zhongyi, et al. Multi-type rice seed recognition based on local resampling fusion of deep features[J]. Transactions of the Chinese Society for Agricultural Machinery, 2025, 56(7) : 522 – 531. (in Chinese)

[15] 时雷, 雷镜楷, 王健, 等. 基于改进 FasterNet 的轻量化小麦生育期识别模型[J]. 农业机械学报, 2024, 55(5) : 226 – 234.

SHI Lei, LEI Jingkai, WANG Jian, et al. Lightweight wheat growth stage identification model based on improved FasterNet[J]. Transactions of the Chinese Society for Agricultural Machinery, 2024, 55(5) : 226 – 234. (in Chinese)

[16] 罗洋, 何自芬, 张印辉, 等. 主干信息共享与多感受野特征自适应融合的作物叶片等级和病害识别方法[J]. 农业机械学报, 2025, 56(1) : 377 – 387.

LUO Yang, HE Zifen, ZHANG Yinhui, et al. Crop leaf grade and disease recognition method based on backbone information sharing and multi-receptive field feature adaptive fusion[J]. Transactions of the Chinese Society for Agricultural Machinery, 2025, 56(1) : 377 – 387. (in Chinese)

[17] MUHARAM F M, NURULHUDA K, ZULKAFLI Z, et al. UAV- and random-forest-adaBoost (RFA)-based estimation of rice plant traits[J]. Agronomy, 2021, 11(5) : 915.

[18] 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)

[19] IKASARI I H, AYUMI V, FANANY M I, et al. Multiple regularizations deep learning for paddy growth stages classification from LANDSAT – 8[C] // Proceedings of the International Conference on Advanced Computer Science and Information Systems (ICACSIS), Malang, 2016.

[20] MURATA K, ITO A, TAKAHASHI Y, et al. A study on growth stage classification of paddy rice by CNN using NDVI images [C] // 2019 Cybersecurity and Cyberforensics Conference (CCC), 2019 : 85 – 90, 10.

[21] HUANG G, LIU Z, VANDERMAATEN L, et al. Densely connected convolutional networks[C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017: 4700 – 4708.

[22] 董言笑, 周继华, 曹黎明, 等. 上海优质梗稻产业现状及高质量发展对策[J]. 中国稻米, 2022, 28(3) : 88 – 91.

DONG Yanxiao, ZHOU Jihua, CAO Liming, et al. Current situation and development strategies of good-quality Japonica rice in Shanghai[J]. China Rice, 2022, 28(3) : 88 – 91. (in Chinese)

[23] ZHANG Y, LUO L, XU C, et al. Quantitative trait loci for panicle size, heading date and plant height co-segregating in trait-performance derived near-isogenic lines of rice[J]. Theoretical and Applied Genetics, 2006, 113(2) : 361 – 368.

[24] WEI X, LIU L, XU J, et al. Breeding strategies for optimum heading date using genotypic information in rice[J]. Molecular Breeding, 2010, 25(2) : 287 – 298.

[25] GUO W, FUKATSU T, NINOMIYA S. Automated characterization of flowering dynamics in rice using field-acquired time-series RGB images[J]. Plant Methods, 2015, 11 : 1 – 15.

[26] RAMADHANI F, PULLANAGARI R, KERESZTURI G, et al. Mapping a cloud-free rice growth stages using the integration of PROBA – V and Sentinel – 1 and its temporal correlation with sub-district statistics[J]. Remote Sensing, 2021, 13(8) : 1498.

[27] GONG Y, YANG K, LIN Z, et al. Remote estimation of leaf area index (LAI) with unmanned aerial vehicle (UAV) imaging for different rice cultivars throughout the entire growing season[J]. Plant Methods, 2021, 17(1) : 88.

[28] CHEN X N, CHEN L, DA H, et al. Symbolic discovery of optimization algorithms[J/OL]. arXiv preprint arXiv:2302.06675, 2023.

[29] LIU Z, LIN Y, CAO Y, et al. Swin Transformer: hierarchical vision transformer using shifted windows[C] // Proceedings of the 18th IEEE/CVF International Conference on Computer Vision (ICCV), Electr Network, 2021.

[30] 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.

[31] MARK S, ANDREW H, ZHU Menglong, et al. MobileNetV2: inverted residuals and linear bottlenecks[J/OL]. arXiv preprint arXiv:1801.04381, 2018.

[32] TAN M, LE Q. Efficientnet: rethinking model scaling for convolutional neural networks[C] // International Conference on Machine Learning, 2019: 6105 – 6114.