



刘涛,刘望,杨奉源,张寰,殷冬梅,焦有宙,张梅凤,张全国. 基于无人机高光谱影像和机器学习算法的花生生物量估算方法研究[J]. 中国农业大学学报,2025,30(03):206-217.

LIU Tao, LIU Wang, YANG Fengyuan, ZHANG Huan, YIN Dongmei, JIAO Youzhou, ZHANG Meifeng, ZHANG Quanguo. Research on estimating peanut biomass using UAV hyperspectral imaging and machine learning algorithm[J]. *Journal of China Agricultural University*, 2025, 30(03): 206-217.

DOI: 10.11841/j.issn.1007-4333.2025.03.19

基于无人机高光谱影像和机器学习算法的花生生物量估算方法研究

刘涛^{1,2} 刘望¹ 杨奉源¹ 张寰^{2*} 殷冬梅³ 焦有宙² 张梅凤⁴ 张全国²

(1. 河南财经政法大学 城乡规划学院, 郑州 450046;

2. 河南农业大学 机电工程学院, 郑州 450002;

3. 河南农业大学 农学院, 郑州 450002;

4. 郑州轻工业大学 工程训练中心, 郑州 450103)

摘要 为评估无人机高光谱遥感技术在作物生物量估算应用中的潜力,以荥阳市花生种植试验田作为研究对象,采用无人机搭载高光谱相机收集多个品种花生在成熟期的高光谱影像数据,结合多种机器学习算法,构建花生生物量估算模型,并进行模型的精度评价与对比分析。首先,通过使用 Savitzky-Golay 滤波器对高光谱影像的反射率进行平滑预处理,并应用 Gaussian4 小波基函数进行连续小波变换,筛选了 53 个植被指数作为特征输入;然后,通过皮尔逊相关系数法进行敏感植被指数筛选,并利用筛选出的植被指数分别构建支持向量回归(Support Vector Regression, SVR)、随机森林(Random Forest, RF)、卷积神经网络(Convolutional Neural Networks, CNN)、粒子群优化(Particle Swarm Optimization)的 SVR、粒子群优化的 RF、粒子群优化的 CNN 等花生生物量估算模型,后进行模型精度评价。结果表明:深度学习模型 CNN 相比于传统的机器学习模型如 RF 和 SVR 等,在花生生物量的预测精度上表现更优;CNN 模型在测试集上的决定系数(R^2)为 0.710, RMSE 为 0.371 kg/m², MSE 和 MAE 分别为 0.138 和 0.329 kg/m²;通过粒子群算法 PSO 进行参数优化后,RF、SVR、CNN 模型的预测精度都有提升,其中 CNN 的提升较为明显,决定系数(R^2)提升约为 8.2%。因此,在花生的收获期使用 PSO 对 CNN 参数优化后的模型对于花生整体生物量的估算最为准确。本研究可为精确预测花生生物量提供科学方法,为智慧乡村建设提供有力支撑。

关键词 高光谱影像;连续小波变换;花生生物量;机器学习;智慧乡村

中图分类号 S107

文章编号 1007-4333(2025)03-0206-12

文献标志码 A

Research on estimating peanut biomass using UAV hyperspectral imaging and machine learning algorithm

LIU Tao^{1,2}, LIU Wang¹, YANG Fengyuan¹, ZHANG Huan^{2*}, YIN Dongmei³, JIAO Youzhou²,
ZHANG Meifeng⁴, ZHANG Quanguo²

(1. School of Urban and Rural Planning, Henan University of Economics and Law, Zhengzhou 450046, China;

2. School of Mechanical and Electrical Engineering, Henan Agricultural University, Zhengzhou 450002, China;

3. College of Agronomy, Henan Agricultural University, Zhengzhou 450002, China;

4. Engineering Training Center, Zhengzhou University of Light Industry, Zhengzhou 450103, China)

收稿日期: 2024-05-21

基金项目: 河南省科技研发计划联合基金青年科学家项目(225200810089);河南省高校科技创新人才支持计划(23HASTIT028);河南省青年人才托举工程项目(2024HYTP018);河南省博士后科研項目(202101041)

第一作者: 刘涛(ORCID:0009-0004-3859-0090),副教授,主要从事农业遥感监测、位置服务等领域研究,E-mail:liutao@huel.edu.cn

通讯作者: 张寰(ORCID:0000-0002-9450-0625),教授,主要从事生物质高值化利用研究,E-mail:zhanghuan5754@163.com

Abstract Accurate and efficient crop biomass estimation is of remarkable importance to identify superior varieties, regional production management and food security assessment. In order to evaluate the potential of UAV hyperspectral remote sensing technology in crop biomass estimation, a peanut planting test field in Xingyang City was taken as the research object. The hyperspectral image data of multiple varieties of peanuts at maturity were collected by UAV equipped with hyperspectral cameras. A peanut biomass estimation model was constructed by combining various machine learning algorithms. The accuracy of the model was evaluated and compared. Firstly, the reflectance of the hyperspectral images was smoothed and preprocessed using the S-G filter, and continuous wavelet transform was applied using the Gaussian4 wavelet basis function to screen 53 vegetation indices as feature inputs. Secondly, the sensitive vegetation indexes were selected through the Pearson correlation coefficient method. Lastly, the identified vegetation index was used to construct Support Vector Regression (SVR), Random Forest (RF), Convolutional Neural Network (CNN), Particle Swarm Optimization Support Vector Machine (PSO-SVR), Particle Swarm Optimization Random Forest (PSO-RF), and Particle Swarm Optimization Convolutional Neural Network (PSO-CNN) models for estimating peanut biomass were constructed and their accuracy was evaluated. The results indicated that: The deep learning model CNN demonstrated a superior performance in the prediction accuracy of peanut biomass compared to traditional machine learning models such as RF and SVM. The CNN model had a determination coefficient (R^2) of 0.710, RMSE of 0.371 kg/m², MSE of 0.138 kg/m², and MAE of 0.329 kg/m² on the test set. After parameter optimization using the Particle Swarm Optimization (PSO) algorithm, the prediction accuracies of RF, SVR, and CNN models were improved, with the CNN model showing the most significant improvement, with an increase of approximately 8.2% in the determination coefficient. Therefore, using the PSO-optimized CNN model during the peanut harvest period provides the most accurate estimation of overall peanut biomass. This study provides a scientific method for accurate peanut biomass prediction and a strong support for the development of smart rural areas.

Keywords hyperspectral imaging; continuous wavelet transformation; peanut biomass; machine learning; smart rural areas

花生作为我国主要的油料作物和经济作物之一,因其具有较高的产量和出油率,在油料作物中占据着独特地位。与其他油料作物如大豆相比,花生果实不仅具有较高的油脂含量,较强的耐旱性和适应性,而且具有极高经济价值^[1]。准确地获取花生种植区域的生物量信息对于农田生产管理和保障农户经济利益至关重要。因此,及时、准确且无损地获取花生生物量信息对于监控植株的生长健康状况、优化农田管理措施,实现花生精细化种植管理具有重要的实际意义^[2-4]。

传统的作物生物量估算方法,如统计调查和气象估产等,虽然被广泛使用,但往往耗时费力,且估算结果存在时间和空间精度上的误差^[5],难以满足精细化种植管理的需求。近年来,无人机遥感技术因其操作简单、机动性强和高时空分辨率的特点,在农业领域得到了广泛应用。特别是在作物产量、生物量监测方面,使用无人机遥感技术已经成为现代农业研究和实践中的重要趋势^[6-7]。其中,高光谱

能够探究农作物光谱特征上的微小差异,常用于建立作物的各种反演模型,能够提供快速、准确的生物量数据,极大地增强了对农业生产的监控和管理^[8]。陶惠林等^[9]利用无人机搭载的高光谱相机获取冬小麦各个时期的遥感影像,选取了9种植被指数和5种红边参数,构建了不同生育期的小麦产量估算模型;Ashapure等^[10]成功地应用高分辨率无人机影像和机器学习方法构建了棉花产量估算模型,该模型综合了时序、非时序和定性特征,在多次测试的平均情况下,模型展示了出色的预测能力;向友珍等^[11]以西北地区的大豆为研究对象,从无人机多光谱影像中筛选出最具相关性的植被指数,基于支持向量机(Support Vector Machine, SVM)、随机森林(Random Forest, RF)和反向神经网络(Back Propagation Neural Network, BPNN)分别构建了大豆LAI、地上部生物量和产量的估计模型。为了提高模型的实用性并增强其稳定性,一些研究在原始光谱数据的基础上采用了连续小波变换和光谱微

分等技术,能够充分挖掘更深层次的和潜在的有效光谱信息,从而有效提高对产量和生物量等参数的估算精度。比如:刘杨等^[12]结合使用多个植被指数和连续小波转换技术,从中提取小波系数并进行相关性分析,筛选出最具相关性的前10个波段,并使用偏最小二乘回归(Partial Least Squares Regression, PLSR)方法构建了马铃薯地上生物量的估算模型;李长春等^[13]将冬小麦在4个生育期的原始光谱与小波能量系数进行相关性分析,后进行筛选,并基于支持向量回归、随机森林和高斯过程回归3种算法构建了冬小麦生物量估算模型。

已有研究主要集中在玉米、小麦、大豆等作物的地上生物量估算,而利用无人机高光谱技术探究花生整体生物量估算的研究仍较为缺乏^[14]。随着近年来作物秸秆在发酵制氢^[15]和制肥^[16]等农业废弃物高值化利用技术中的应用日益受到关注,本研究综合考虑了花生地上部分和地下部分的生物量,全面评估了花生的整体生物量,涵盖其副产品的应用价值。因此,本研究拟以河南荥阳花生种植试验

田为研究对象,通过对高光谱影像使用小波变换得到小波能量系数,计算植被指数,结合SVR、RF、CNN及使用粒子群优化这些模型PSO-RF、PSO-SVR以及PSO-CNN的方法对花生生物量进行建模和评估,分析不同模型对于生物量的预测效果,以期寻找构建花生生物量反演的最佳模型,为农田地块尺度的花生生物量监测提供一种有效的参考。

1 材料与方法

1.1 研究区概况

研究区域荥阳市位于河南省中部,地处黄河的中下游平原地区,研究区的地理位置以及试验田样本点示意图见图1。该地区地势相对平坦,土壤肥沃,土地资源丰富,适宜农业种植,特别是粮食作物的生产。荥阳市的气候类型属于温带大陆性季风气候,四季分明,历年平均气温14.8℃,降水集中在6—9月,历年平均降水量608.8 mm,该地区以农作物种植为主,常见的作物包括花生、小麦、玉米等。

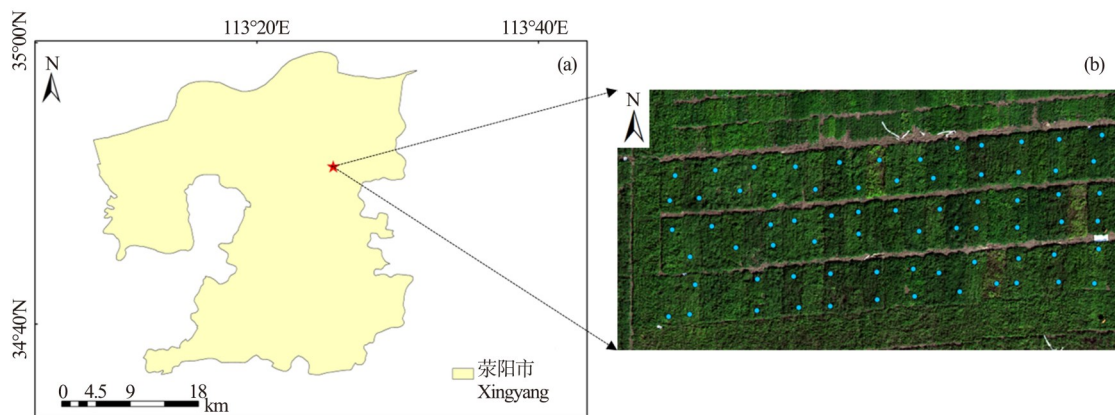


图1 研究区的地理位置(a)以及试验田样本点示意图(b)

Fig. 1 Geographical location of the study area (a) and the schematic diagram of the sample plots (b)

1.2 数据获取及处理

1.2.1 光谱数据获取和处理

本研究于2022年10月10日花生收获当天,利用大疆M600 Pro六旋翼无人机搭载四川双利合谱有限公司的GaiaSky-mini2-VN高光谱成像仪,获取了研究区域内花生田块的高光谱影像数据。飞行高度设置为50 m,旁向和航向重叠率均为80%,空间分辨率约为1.8 cm/像素。该高光谱成像仪具有176个光谱反射率通道,覆盖的光谱范围为

400~1 000 nm,光谱分辨率为3.5 nm。采集的影像数据通过高光谱系统处理软件SpecView进行了一系列预处理,包括黑白帧校准、反射率校准、大气校正以及镜头变换等。在光谱数据的采集过程中,高光谱传感器可能受到外部环境因素或传感器自身的影响,通常表现为光谱曲线上的不平滑或毛刺。为解决上述问题,本研究采用Savitzky-Golay(以下简称S-G)滤波器进行卷积平滑处理,以提高光谱的信噪比并减少外部干扰的

影响;具体而言,使用了窗口大小为13、阶数为5的S-G滤波参数,从而显著改善了影像的平滑效

果。平滑处理前后的影像反射率效果如图2所示。

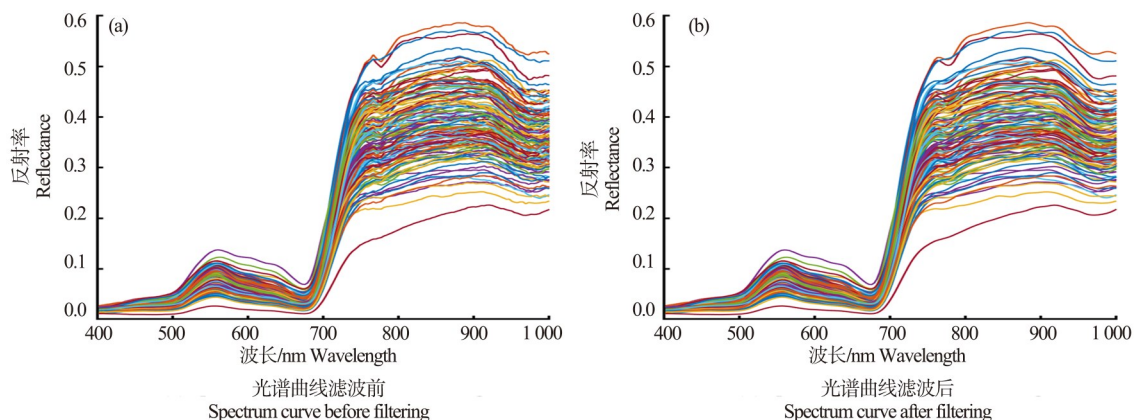


图2 SG平滑滤波前(a)后(b)光谱曲线对比

Fig. 2 S-G smooth filtering before (a) and after (b) comparison

1.2.2 花生生物量获取

本研究选取多品种花生试验田进行数据采集,该试验田种植:豫科花、龙乡红、驻花、周科花、星花、濮花、宛花、周花、远杂、豫花等在内11个大类花生品种,种植试验田共66个小区,每个小区的长为6.7 m,宽为2 m,种5行空一行,每个品种重复3次。在花生饱果成熟期,对采样点0.5 m×0.5 m范围内的所有花生植株进行采摘,

剔除根部泥土后装袋密封,将采集的花生植株带回实验室进行清洗;放置在通风良好处,自然风干后,将采样点植株放入烤箱中65~75℃烘至恒重,进行称重作为各采样点的生物量干重实测值。为了进行模型训练和验证,从70个采样点中随机选取80%样本点作为训练集,剩余的20%样本点作为验证集。样本点生物量的统计结果如表1所示。

表1 样本点花生生物量统计结果

Table 1 Statistical results of peanut biomass at sampling points

样本类型 Sample type	样本数 Sample size	最大值/(kg/m ²) Maximum	最小值/(kg/m ²) Minimum	均值/(kg/m ²) Mean	标准差/(kg/m ²) Standard deviation	变异系数/% Coefficient of variation
训练集 Training set	56	5.52	0.88	1.95	0.76	38.75
测试集 Test set	14	4.16	1.24	2.00	0.71	35.67
全集 Complete set	70	5.52	0.88	1.98	0.76	38.34

1.3 数据处理及模型构建

1.3.1 小波变换

小波变换是一种常见的数据处理方法^[17-18]。该方法基于傅里叶变换增加了时间尺度,使得信号可以在时间域和频域进行分解,从而得到原始信号中的深层次隐含的信息。在高光谱数据处理中,基于小波母函数的平移和缩放对高光谱数据进行卷积运算,在不改变原始波段范围和位置的情况下将高

光谱信息分解成不同尺度下和不同波段长度的小波能量系数。其计算公式如下:

$$W_f(a, b) = \int_{-\infty}^{+\infty} f(\lambda) \phi_{a,b}(\lambda) d\lambda \quad (1)$$

$$\Psi_{a,b}(\lambda) = \frac{1}{\sqrt{a}} \Psi\left(\frac{\lambda - b}{a}\right) \quad (2)$$

式中: $W_f(a, b)$ 为小波能量系数; $f(\lambda)$ 为地上花生叶高光谱反射率; λ 为光谱波段(400~1 000 nm);

$\psi_{a,b}$ 为选取 Gaussian4 作为小波基函数; a 为尺度因子; b 为平移因子。本研究将经过 SG 滤波平滑处理的高光谱数据进一步通过使用 Matlab 2023a 软件进行连续小波变换分析。在 1~10 尺度 (2^1 、 2^2 、 2^3 、 2^4 、 2^5 、 2^6 、 2^7 、 2^8 、 2^9 、 2^{10}) 进行分解, 使用 Gaussian4 作为小波分析的母小波函数, 以提取高光谱数据中的隐含有效信息。

1.3.2 植被指数构建方法

利用小波变换处理后的光谱反射率数据来构建植被指数。植被指数通过线性或非线性计算获得, 不仅减少了干扰信息, 还突出了植被的关键信息, 从而能够反映出植被的深层次特征。基于已有研究成果^[19-21], 选用了 53 个与作物产量或生物量相关的经验植被指数, 分别与实测花生生物量进行皮

尔逊相关性进行分析。考虑到模型的简洁性以及特征冗余易造成的多重共线性的问题, 本研究仅筛选出相关性最高的前 1%, 即前 5 个植被指数作为特征向量。

所筛选的植被指数其计算公式如表 2 所示。其中, PSRI 对于监测植被冠层胁迫性变化十分敏感, 常用于作物生产力的预测模型; DCNII 则被视为预测植物氮浓度的最佳光谱指标, 对作物生物量预测具有极大的潜力; MMTCI 作为经过改性的陆地叶绿素指数 (MTCI), 展现了比原始 MTCI 更高的提取精度和抗干扰能力, 常用于估算作物的叶绿素含量; SAVI 是一个广泛使用于作物估产的植被指数; VARI 因其在预测作物水分胁迫方面的性能而被广泛应用于小麦和苕麻等作物的估产模型。

表 2 植被指数及其计算公式

Table 2 Vegetation index and its calculation formula

植被指数 Vegetation index	计算公式 Calculation formula	参考文献 Reference
植被衰老反射率指数 PSRI	$(R_{680} - R_{500}) / R_{750}$	[19]
双峰冠层氮素指数 I DCNII	$((R_{750} - R_{670} + 0.09)(R_{750} - R_{700})) / (R_{700} - R_{670})$	[20]
修正地面叶绿素指数 MMTCI	$((R_{750} - R_{680} + 0.03)(R_{750} - R_{710})) / (R_{710} - R_{680})$	[20]
土壤调节植被指数 SAVI	$1.5(R_{800} - R_{670}) / (R_{800} + R_{670} + 0.5)$	[21]
可见抗大气指数 VARI	$(R_{555} - R_{680}) / (R_{555} + R_{680} - R_{480})$	[21]

注: R 表示反射率, 其下角标代表具体波长, nm。

Note: R represents the reflectance, with the subscript representing its specific wavelength (nm).

1.3.3 模型构建方法

1) 粒子群优化算法。粒子群优化算法 (PSO) 是一种启发式优化算法, 每个粒子被初始化速度和位置之后在解空间中迭代寻找全局最优解, 并通过适应性评价来记录和比较解的品质, 具有收敛速度快且易于实现等优点, PSO 算法被广泛应用于模型优化等问题^[22-23]。为了提升 RF、SVR 和 CNN 模型的拟合能力, 并减少参数设置的误差影响, 本研究采用了 PSO 来优化这些模型的参数, 具体算法流程如图 3 所示。具体而言, 设置学习因子 $c1$ 和 $c2$ 均为 1.5, 惯性权重设置为 0.73。在初始化种群和模型超参数后, 并将其应用于各种模型中, 对每次迭代的适应度进行记录, 并判断是否达到终止条件; 如果未达到终止条件, 则继续更新粒子的位置和速度

参数, 直至满足终止条件。

2) 随机森林。随机森林 (RF) 是一种应用于分类和回归问题的集成学习方法, 基于多个决策树构建。此方法通过构建并组合多个决策树的预测结果来提高整体预测性能, 每棵决策树都是独立地从原始数据集中随机选择样本和特征进行训练。这种随机选择增加了模型的多样性, 使得 RF 在回归预测中显示出优异的鲁棒性^[24]。在本研究中, 利用 PSO 算法对 RF 模型的关键参数进行优化, 具体包括决策树的数量和叶节点的最小样本数。参数的初始范围详见表 3。

3) 支持向量回归。支持向量回归 (SVR) 是从支持向量机算法衍生而来, 可用于解决回归拟合问题。SVR 通过使用基核函数, 如高斯核和多项式

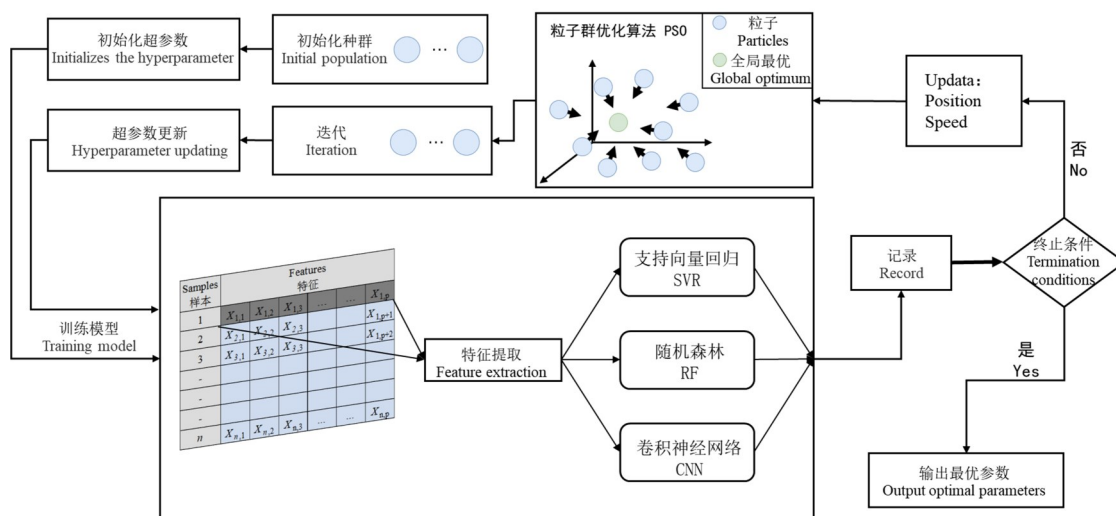


图3 粒子群算法流程图

Fig. 3 Flow chart of particle swarm optimization

核,以及通过梯度下降算法优化权重系数,来构建最优的决策边界;这一过程包括构建一个超平面,将输入样本映射到高维空间中,从而在该空间中最小化数据点的距离,使得数据更加线性可分。这样构建出来的回归模型不仅预测性能出色,而且具有很高的鲁棒性^[25]。在本研究中,进行优化的参数包括惩罚因子和径向基函数参数。参数的初始范围详见表3。

4)卷积神经网络。卷积神经网络(CNN)是一种深度学习架构,主要由输入层、卷积层、池化层和全连接层组成,本研究使用的CNN网络结构如图4所示。

其中,卷积层负责提取输入数据的特征,池化层用于减小数据的维度,而全连接层则利用Softmax激活函数整合这些特征以产生最终输出^[26]。CNN的性能和准确性在很大程度上受到其超参数的影响,如初始学习率、批处理大小和L2正则化系数等^[27]。因此,优化这些超参数是提高模型训练效率和预测精度的关键。本研究采用PSO算法对CNN的超参数进行优化,以加速收敛过程并提高模型的总体性能。超参数的初始优化范围详见表3。这种方法通过系统地调整关键参数,旨在实现更优的网络训练结果和提升模型的泛化能力。

表3 超参数初始值范围设置

Table 3 Hyperparameter initial value range setting

模型 Model	粒子 Particle	超参数 Hyperparameter	初始化范围 Initialize the scope
RF	ntree	决策树数目 Number of decision trees	50~500
	leaf	最小叶子数 Minimum leaf number	1~10
SVR	c	惩罚因子 Penalty factor	1~20
	g	径向基函数参数 Radial basis function parameter	1~5
CNN	X1	批处理规模 Batch scale	32~512
	X2	初始学习率 Initial learning rate	0.001~0.1
	X3	L2正则化参数 L2 regularization parameter	0.001~0.1

1.3.4 模型精度评价

为了评估花生生物量估算模型的精度,将随机选取的20%样本点的样本数据带入已训练好的模

型中进行测试,采用决定系数(R^2)、均方根误差(RMSE)、均方误差(MSE)、平均绝对误差(MAE)对模型的性能进行评价。具体计算公式如下:

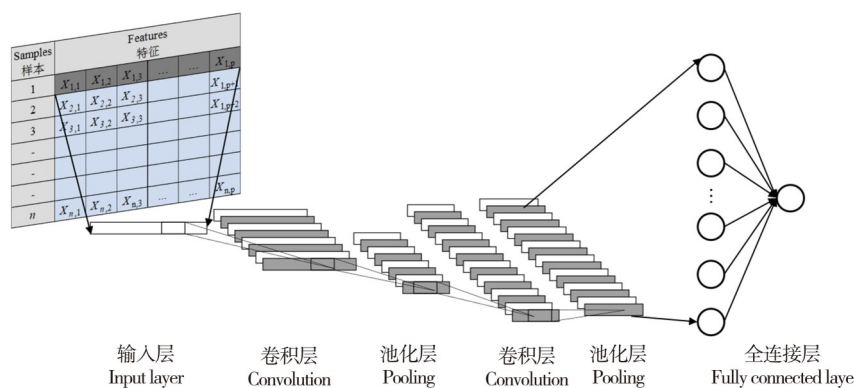


图4 卷积神经网络结构图

Fig. 4 Convolutional neural network structure

$$R^2 = 1 - \frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (3)$$

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}} \quad (4)$$

$$\text{MSE} = \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n} \quad (5)$$

$$\text{MAE} = \frac{\sum_{i=1}^n |y_i - \hat{y}_i|}{n} \quad (6)$$

式中: i 为样本集的个数, $i = 1, 2, \dots, n$; y 为花生生物量实际测量值; $\hat{y}_i$ 为花生生物量的预测值; $\bar{y}$ 为花生生物量的平均值。其中, 当 R^2 接近于 1, RMSE、MSE 和 MAE 接近于 0, 说明模型的拟合效果越好。

2 结果与分析

2.1 小波分析结合植被指数特征构建

图 5 展示了基于高光谱原始反射率直接计算的植被指数与花生生物量之间的相关性。包括相关系数以及通过小波变换分解为 10 个尺度后计算的相关系数。可见: 不同植被指数与花生生物量的相关系数范围为 $-0.49 \sim 0.45$, 部分植被指数与生物量呈正相关; 另一些则呈负相关。这种广泛的相关性范围反映了不同植被指数在捕捉植被生长状况及其生理反应方面的差异, 尤其是在不同小波尺度下, 揭示了不同尺度的小波分解在捕捉与生物量相关的重要光谱信息方面的有效性。

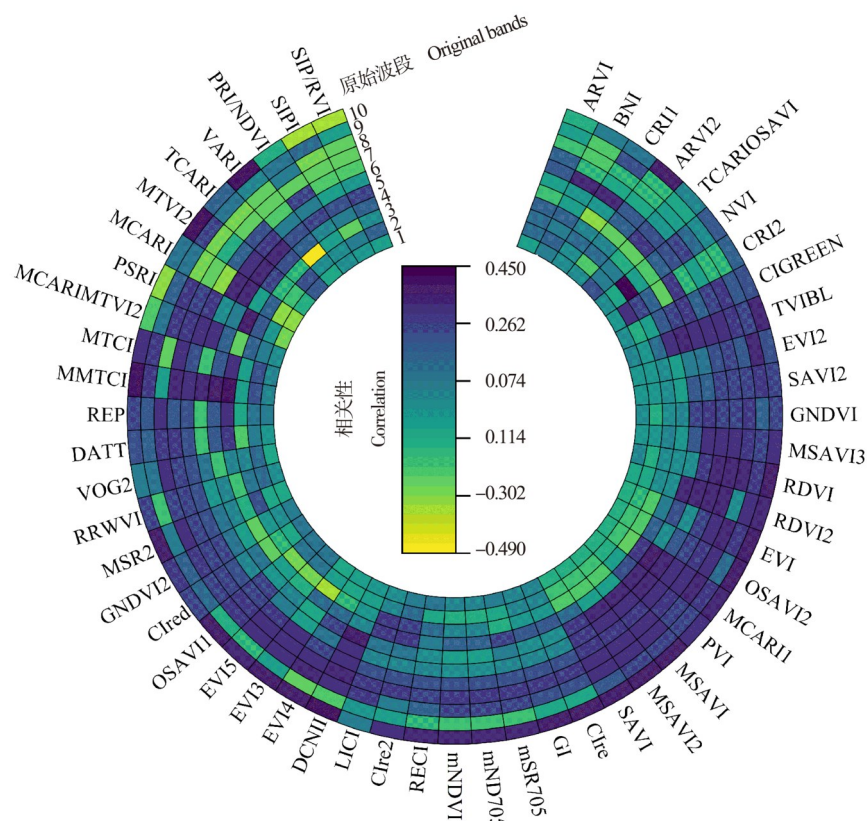
由图 5 可见, 相关性较高的植被指数主要集中在

在第 3、4、5、7、8 和 9 尺度, 表明这些尺度的小波变换能够更有效地捕捉与花生生物量密切相关的重要光谱特征。具体而言, 本研究使用第 3 尺度下的土壤调节植被指数 (SAVI)、可见光大气修正植被指数 (VARI)、第 4 尺度下的植被衰老反射率指数 (PSRI)、修正地面叶绿素指数 (MMTCI) 和第 5 尺度下的双峰冠层氮素指数 I(DCNII) 作为模型输入, 其相关性值分别为 0.448、 -0.486 、0.377、0.389 和 0.370。可以看出, 不同尺度的植被指数在捕捉与花生生物量相关的光谱特征方面具有显著差异, 且某些特定尺度下的指数与生物量之间存在较强的相关性。因此, 小波变换有助于解析植被指数与生物量之间的复杂关系, 经过小波变换分解后的植被指数能够有效保留植被指数在捕捉生物量信息方面的敏感性, 为多尺度分析方法提供了有力依据。

2.2 不同模型构建及精度评价

利用筛选出的植被指数作为特征变量, 以构建花生生物量反演模型。具体而言: 使用 Matlab 2023a 软件与多种机器学习模型建模, 包括 SVR、RF、CNN 及其通过粒子群优化算法 (PSO) 增强的版本, PSO-SVR、PSO-RF 和 PSO-CNN, 共计获得 6 种模型。表 4 中展示了每个模型的 4 种性能评价指标的检测结果, 包括: 决定系数 (R^2)、均方根误差 (RMSE)、均方误差 (MSE) 和平均绝对误差 (MAE)。由表 4 可见, PSO 优化后的模型普遍表现更优, 尤其是在 R^2 和 RMSE 方面, 显示出更高的预测准确性和较低的误差, 这表明使用 PSO 优化后的超参数有助于提高模型拟合效果。

6 种模型在训练集和测试集上的检验结果如图 6 所示。可见在这 6 个模型中, 大部分数据点紧邻



1~10表示小波变换分解的10个尺度。

1 to 10 represent the ten scales of wavelet transform decomposition.

图 5 各植被指数与花生生物量相关性热力图

Fig. 5 Heat map of the correlation between all vegetation index and peanut biomass

表 4 不同建模方法下花生生物量估算的精度检验结果

Table 4 Comparison of precision test results of peanut biomass estimation model

建模方法 Modeling method	R^2		RMSE /(kg/m ²)		MSE /(kg/m ²)		MAE /(kg/m ²)	
	训练集 Training set	测试集 Test set	训练集 Training set	测试集 Test set	训练集 Training set	测试集 Test set	训练集 Training set	测试集 Test set
RF	0.673	0.674	0.429	0.393	0.184	0.155	0.305	0.28
SVR	0.674	0.651	0.428	0.407	0.183	0.165	0.272	0.346
CNN	0.778	0.710	0.353	0.371	0.125	0.138	0.264	0.329
PSO-RF	0.671	0.689	0.430	0.384	0.185	0.147	0.301	0.274
PSO-SVR	0.667	0.674	0.433	0.393	0.187	0.154	0.279	0.334
PSO-CNN	0.805	0.768	0.331	0.331	0.110	0.110	0.244	0.283

1:1虚线,表明所建模型能够有效地预测训练集和测试集的花生生物量,反映了模型的预测准确性和泛化能力。由图6(a)、(c)和(e)可见:在机器学习模型中,RF在测试集上的决定系数(R^2)为0.674,均方根误差(RMSE)为0.429 kg/m²,均方误差(MSE)和平均绝对误差(MAE)分别为0.155和

0.28 kg/m²;SVR测试集的 R^2 为0.651, RMSE为0.407 kg/m², MSE和MAE分别为0.165和0.346 kg/m²。在深度学习模型中,使用默认参数的CNN测试集的 R^2 为0.710, RMSE为0.371 kg/m², MSE和MAE分别为0.138和0.329 kg/m²,表明深度学习CNN模型的预测性能相比机器学习模型(RF和

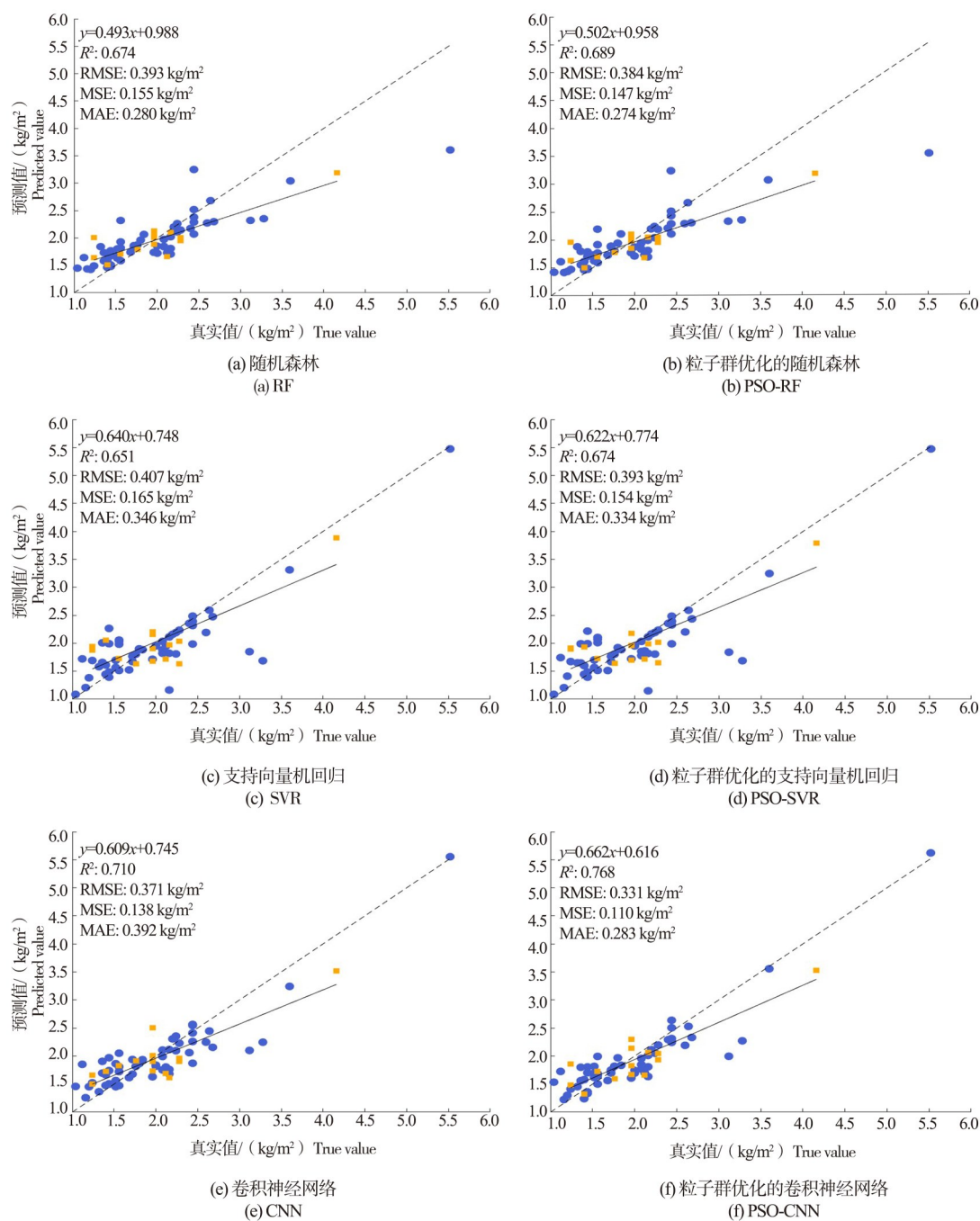


图6 6种模型对花生生物量估算的拟合效果

Fig. 6 Fitting effect of 6 models on peanut biomass estimation

SVR)有一定提升,即深度学习算法在对花生生物量的建模中的拟合效果要优于机器学习模型。

由图6(b)、(d)和(f)可见: PSO-RF和PSO-SVR在测试集中的 R^2 、RMSE、MSE和MAE分别为0.689、0.384、0.147、0.274和0.674、0.393、0.154、0.334 kg/m²。PSO-CNN在测试集上表现最佳, R^2 为0.768,RMSE为0.331 kg/m²,MSE和MAE分别为0.11和0.283 kg/m²。与PSO-RF和PSO-SVM模型相比,PSO-CNN模型的拟合效果达到最优,说明经过PSO优化后的深度学习模型优于PSO优化后的机器学习模型。将模型预测精度从高到低排序,未经PSO优化的模型为CNN、RF、SVR;而经过PSO优化的模型排序为PSO-CNN、PSO-RF、PSO-SVR。因此经过PSO优化的深度学习模型(PSO-CNN)在预测花生生物量方面表现最佳,显著优于其他未经优化和经过优化的机器学习模型,进一步验证了PSO方法在提升模型预测精度方面的有效性。

3 讨论

相比以往研究利用原始光谱反射率构建植被指数估算作物生物量,本研究利用小波变换对相关较强的植被指数进行分解,并利用分解后的特征构建生物量估算模型。结果发现在未分解的53种植被指数中,双峰冠层氮素指数I(DCNI I)与实测生物量的相关性最高,相关系数为0.403,这与裴浩杰等^[20]研究结论相似,但仍低于本研究中使用小波变换后的植被指数SAVI(0.448),本研究使用相同的方法,挑选相关性最强的5个未经小波变换的植被指数,并构建6种花生生物估算模型(SVR、RF、CNN、PSO-SVR、PSO-RF、PSO-CNN),结果表明PSO-CNN模型精度最高,决定系数 R^2 为0.65。

利用PSO优化的模型精度普遍高于未优化之前,PSO-CNN模型在使用高光谱影像预测花生生物量方面表现出最高的精度,试验结论与王晗等^[22]和刘兰军等^[23]研究结果一致。其中,PSO对于RF和SVR的提升效果有限,对于CNN提升效果较为明显,训练集的决定系数从0.710提升到了0.768,提升了约为8.2%。这一结果也突显了PSO在优化深度学习模型参数中的有效性,为精确估算花生生物量提供了一个强有力的工具。

本研究展示了无人机高光谱遥感技术在精准

农业中的应用潜力,特别是在花生生物量估算方面,所提出的高效、无损的花生生物量估测方法,不仅增强了现代化作物管理的科学性,为农业监测、农业保险等应用提供了可靠的依据,有助于在农业估产领域实现更高水平的推广应用,为发展低空经济和数字化农业提供重要参考。

4 结论

为提高田块尺度花生生物量估算精度、实现精准农业管理,本研究基于无人机高光谱成像技术,探索了结合小波变换和植被指数的特征构建方法,从53种经验植被指数中筛选出相关性较强的特征,并通过机器学习模型、深度学习模型以及粒子群优化算法(PSO)构建花生生物量估算模型。主要结论如下:

1)在小波变换第3尺度下的SAVI和VARI、第4尺度下的PSRI和MMTCI和第5尺度下的DCNII与实测花生生物量具有较强的相关性。相较于直接利用原始波段构建的植被指数,本研究构建的最优模型在决定系数(R^2)上提升了约16.8%;

2)深度学习模型CNN在花生生物量预测精度上优于传统的机器学习模型,如RF和SVR,在花生生物量预测模型精度上表现出更优的效果;

3)通过粒子群算法PSO进行参数优化后,RF、SVR、CNN模型的预测精度都有提升,其中CNN的提升较为明显,模型精度提升约为8.2%。使用PSO-CNN模型对于花生生物量的估算能够获得最优结果,测试集 R^2 为0.768。

参考文献 References

- [1] 全伟,马紫涛,石方刚,高晨,吴明亮. 2BHG-160型花生播种机关键部件设计与试验[J]. 中国农业大学学报,2024,29(3):164-174
Quan W, Ma Z T, Shi F G, Gao C, Wu M L. Design and experiment of the key components of 2BHG-160 peanut planter [J]. *Journal of China Agricultural University*, 2024, 29(3): 164-174 (in Chinese)
- [2] Qiao D, Yang J T, Bai B, Li G W, Wang J G, Li Z H, Liu J C, Liu J Y. Non-destructive monitoring of peanut leaf area index by combing UAV spectral and textural characteristics[J]. *Remote Sensing*, 2024, 16(12): 2182
- [3] 王靖,彭琦,刘小娟,莫佳才,梁婷. 基于无人机多光谱数据的水稻LAI反演与应用[J]. 中国农业大学学报,2021,26(12):145-156
Wang J, Peng Y, Liu X J, Mo J C, Liang T. Inversion and application of rice LAI based on UAV multispectral data [J]. *Journal of China Agricultural University*, 2021, 26(12): 145-156 (in Chinese)
- [4] 杨北萍,陈圣波,于海洋,安秦. 基于随机森林回归方法的水稻产量遥感

- 估算[J]. 中国农业大学学报, 2020, 25(6): 26-34
- Yang B P, Chen S B, Yu H Y, An Q. Remote sensing estimation of rice yield based on random forest regression method [J]. *Journal of China Agricultural University*, 2020, 25(6): 26-34 (in Chinese)
- [5] Cheng E H, Zhang B, Peng D L, Zhong L H, Yu L, Liu Y, Xiao C C, Li C J, Li X Y, Chen Y, Ye H C, Wang H Y, Yu R Y, Hu J K, Yang S L. Wheat yield estimation using remote sensing data based on machine learning approaches[J]. *Frontiers in Plant Science*, 2022, 13: 1090970
- [6] 刘涛, 张寰, 王志业, 贺超, 张全国, 焦有宙. 利用无人机多光谱估算小麦叶面积指数和叶绿素含量[J]. 农业工程学报, 2021, 37(19): 65-72
- Liu T, Zhang H, Wang Z Y, He C, Zhang Q G, Jiao Y Z. Estimation of the leaf area index and chlorophyll content of wheat using UAV multi-spectrum images[J]. *Transactions of the Chinese Society of Agricultural Engineering*, 2021, 37(19): 65-72 (in Chinese)
- [7] 韦春宇, 杜娅丹, 程智楷, 周智辉, 谷晓博. 基于无人机遥感植被指数优选的覆膜冬小麦估产研究[J]. 农业机械学报, 2024, 55(4): 146-154, 175
- Wei C Y, Du Y D, Cheng Z K, Zhou Z H, Gu X B. Yield estimation of mulched winter wheat based on UAV remote sensing optimized by vegetation index[J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2024, 55(4): 146-154, 175 (in Chinese)
- [8] 赵龙才, 李粉玲, 常庆瑞. 农作物遥感识别与单产估算研究综述[J]. 农业机械学报, 2023, 54(2): 1-19
- Zhao L C, Li F L, Chang Q R. Review on crop type identification and yield forecasting using remote sensing [J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2023, 54(2): 1-19 (in Chinese)
- [9] 陶惠林, 徐良骥, 冯海宽, 杨贵军, 杨小冬, 牛亚超. 基于无人机高光谱遥感数据的冬小麦产量估算[J]. 农业机械学报, 2020, 51(7): 146-155
- Tao H L, Xu L J, Feng H K, Yang G J, Yang X D, Niu Y C. Winter wheat yield estimation based on UAV hyperspectral remote sensing data [J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2020, 51(7): 146-155 (in Chinese)
- [10] Ashapure A, Jung J, Chang A J, Oh S, Yeom J, Maeda M, Maeda A, Dube N, Landivar J, Hague S, Smith W. Developing a machine learning based cotton yield estimation framework using multi-temporal UAS data [J]. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2020, 169: 180-194
- [11] 向友珍, 安嘉琪, 赵笑, 金琳, 李志军, 张富仓. 基于无人机多光谱遥感的大豆生长参数和产量估算[J]. 农业机械学报, 2023, 54(08): 230-239
- Xiang Y Z, An J Q, Zhao X, Jin L, Li Z J, Zhang F C. Soybean growth parameters and yield estimation based on UAV multispectral remote sensing [J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2023, 54(8): 230-239 (in Chinese)
- [12] 刘杨, 孙乾, 冯海宽, 杨福芹. 基于小波分析的马铃薯地上生物量估算[J]. 光谱学与光谱分析, 2021, 41(4): 1205-1212
- Liu Y, Sun Q, Feng H K, Yang F Q. Estimation of above-ground biomass of potato based on wavelet analysis [J]. *Spectroscopy and Spectral Analysis*, 2021, 41(4): 1205-1212 (in Chinese)
- [13] 李长春, 李亚聪, 王艺琳, 马春艳, 陈伟男, 丁凡. 基于小波能量系数和叶面积指数的冬小麦生物量估算[J]. 农业机械学报, 2021, 52(12): 191-200
- Li C C, Li Y C, Wang Y L, Ma C Y, Chen W N, Ding F. Winter wheat biomass estimation based on wavelet energy coefficient and leaf area index [J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2021, 52(12): 191-200 (in Chinese)
- [14] Oliveira M F, Carneiro F M, Ortiz B V, Thurmond M, Oliveira L P, Bao Y, Sanz-Saez A, Tedesco D. Predicting below and above-ground peanut biomass and maturity using multi-target regression[J]. *Computers and Electronics in Agriculture*, 2024, 218: 108647
- [15] Liu T, Li Y M, Zhang H, Zhang N Y, Tahir N, Zhang Q G. Estimating the potential of biohydrogen production and carbon neutralization contribution from crop straw [J]. *Bioresource Technology*, 2023, 373: 128718
- [16] Guo T T, Zhang T P, Lim E, López-Benítez M, Ma F, Yu L M. A review of wavelet analysis and its applications: Challenges and opportunities[J]. *IEEE Access*, 2022, 10: 58869-58903
- [17] 曾斌, 唐敏, 唐伟, 于桂阳, 曾健青. 花生粕和花生秸秆在动物饲用中的研究进展[J]. 中国饲料, 2023(5): 152-162
- Zeng B, Tang M, Tang W, Yu G Y, Zeng J Q. Research progress of peanut meal and peanut straw in animal feeding[J]. *China Feed*, 2023 (5): 152-162 (in Chinese)
- [18] 邹佳琪, 王仲林, 谭先明, 陈燎原, 杨文钰, 杨峰. 基于连续小波变换估测干旱胁迫下玉米籽粒产量[J]. 作物学报, 2024, 50(4): 1030-1042
- Zou J Q, Wang Z L, Tan X M, Chen L Y, Yang W Y, Yang F. Estimation of maize grain yield under drought stress based on continuous wavelet transform [J]. *Acta Agronomica Sinica*, 2024, 50(4): 1030-1042 (in Chinese)
- [19] 刘杨, 冯海宽, 黄珏, 孙乾, 杨福芹, 杨贵军. 基于无人机高光谱影像的马铃薯株高和地上生物量估算[J]. 农业机械学报, 2021, 52(2): 188-198
- Liu Y, Feng H K, Huang J, Sun Q, Yang F Q, Yang G J. Estimation of potato plant height and above-ground biomass based on UAV hyperspectral images [J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2021, 52(2): 188-198 (in Chinese)
- [20] 裴浩杰, 冯海宽, 李长春, 李振海, 杨贵军, 王衍安, 郭建华. 基于多元线性回归和随机森林的苹果叶绿素含量高光谱估测方法比较[J]. 江苏农业科学, 2018, 46(17): 224-230
- Pei H J, Feng H K, Li C C, Li Z H, Yang G J, Wang Y A, Guo J H. Comparison of hyperspectral estimation methods for chlorophyll contents of apple based on multiple linear regression and random forest [J]. *Jiangsu Agricultural Sciences*, 2018, 46(17): 224-230 (in Chinese)
- [21] 郑恒彪, 吉文翰, 郭彩丽, 邱小雷, 姚露, 江冲亚, 朱艳, 曹卫星, 程涛. 无人机遥感作物估产研究进展[J/OL]. 南京农业大学学报, 1-12
- Zheng H B, Ji W H, Guo C L, Qiu X L, Yao X, Jiang C Y, Zhu Y, Cao W X, Cheng T. Uav remote sensing for crop yield estimation research progress [J/OL]. *Journal of Nanjing Agricultural University*, 1-12 (in Chinese)
- [22] 王晗, 向友珍, 李汪洋, 史鸿棹, 王辛, 赵笑. 基于无人机多光谱遥感的冬油菜地上部生物量估算[J]. 农业机械学报, 2023, 54(8): 218-229
- Wang H, Xiang Y Z, Li W Y, Shi H Z, Wang X, Zhao X. Estimation of winter rapeseed above-ground biomass based on uav multi-spectral remote sensing[J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2023, 54(8): 218-229 (in Chinese)
- [23] 刘兰军, 翟永庆, 郑俊俊, 范萍萍, 邓莉. 基于 PSO-CNN 的土壤氮含量可见/近红外光谱建模[J]. 光学技术, 2021, 47(4): 438-445
- Liu L J, Zhai Y Q, Zheng J J, Fan P P, Deng L. Visible/near infrared spectroscopy modeling of soil nitrogen content based on PSO-CNN[J]. *Optical Technique*, 2021, 47(4): 438-445 (in Chinese)
- [24] 李冠稳, 高小红, 肖能文, 肖云飞. 基于 sCARS-RF 算法的高光谱估算土壤有机质含量[J]. 发光学报, 2019, 40(8): 1030-1039
- Li G W, Gao X H, Xiao N W, Xiao Y F. Estimation soil organic matter contents with hyperspectra based on sCARS and RF algorithms [J]. *Chinese Journal of Luminescence*, 2019, 40(8): 1030-1039 (in Chinese)
- [25] 黎锐, 李存军, 徐新刚, 王纪华, 杨小冬, 黄文江, 潘瑜春. 基于支持向量

- 回归(SVR)和多时相遥感数据的冬小麦估产[J]. 农业工程学报, 2009, 25(7): 114-117
- Li R, Li C J, Xu X G, Wang J H, Yang X D, Huang W J, Pan Y C. Winter wheat yield estimation based on support vector machine regression and multi-temporal remote sensing data[J]. *Transactions of the Chinese Society of Agricultural Engineering*, 2009, 25(7): 114-117 (in Chinese)
- [26] Kiranyaz S, Avcı O, Abdeljaber O, Ince T, Gabbouj M, Inman D J. 1D convolutional neural networks and applications: A survey [J]. *Mechanical Systems And Signal Processing*, 2021, 151: 107398
- [27] Yan H, Zhong C Q, Wu Y H, Zhang L Y, Lu W. A hybrid-model optimization algorithm based on the Gaussian process and particle swarm optimization for mixed-variable CNN hyperparameter automatic search [J]. *Frontiers of Information Technology & Electronic Engineering*, 2023, 24(11): 1557-15
- 责任编辑: 杨爱东



第一作者简介: 刘涛, 博士, 副教授。河南省农业工程学会常务理事、农业信息工程专委会副主任, 主要从事农业遥感监测、定位导航与位置服务等领域的研究工作, 主持1项国家自然科学基金项目、1项中国博士后基金项目、多项省级项目, 获省级科技进步奖1项、出版专著1部。



通讯作者简介: 张寰, 博士, 教授, 正高级实验师。主要从事生物质高值化利用、农业遥感监测等领域的研究工作, 主持1项国家自然科学基金项目、多项省级项目, 获省级科技进步奖2项、出版专著3部。