

基于无人机多源数据的花生表型估算模型

何宁^{1,2} 王剑³ 卢宪菊^{1,2} 陈博^{1,2} 白波^{4,5} 樊江川^{1,2}

(1.北京市农林科学院信息技术研究中心,北京 100097; 2.数字植物北京市重点实验室,北京 100097;

3.北京市数字农业农村促进中心,北京 101117; 4.山东省农业科学院农作物种质资源研究所,济南 250100;

5.山东省作物遗传改良与生态生理重点实验室,济南 250100)

摘要:花生作为重要的油料作物,对粮油产量安全起到至关重要的作用,准确、无损、实时的表型监测对花生生产管理具有重要意义。本研究利用无人机平台获取关键生育期多光谱及图像数据,提取冠层光谱(Multispectral, MS)、结构(Canopy height model, CHM)和纹理参数(Textural, TEX)信息,采用偏最小二乘回归(Partial least squares regression, PLSR)、支持向量机(Support vector machine, SVM)、人工神经网络(Artificial neural network, ANN)和随机森林回归(Random forest regression, RFR)4种算法构建花生株高、叶片叶绿素相对含量(SPAD值)、地上部生物量估算模型。研究结果表明:花生地上生物量和株高与近红外波段有强相关性(皮尔森相关系数分别为0.77和0.69),融合纹理、结构和光谱特征后的随机森林模型取得了对生物量最优的模型反演效果(决定系数 R^2 为0.96),融合纹理和光谱特征后的偏最小二乘回归模型对株高的反演效果最优(R^2 为0.94);融合纹理和结构特征后的偏最小二乘回归对SPAD值反演效果相对较好(R^2 为0.39,均方根误差(RMSE)为3.06,归一化均方根误差(nRMSE)为0.062,百分偏差比率(RPD)为1.30)。本研究明确了不同机器学习方法对花生不同表型指标估算所需的特征指标,构建的基于无人机多源数据的表型估算模型可以实现对花生株高和生物量的准确、无损、高效估算,为花生长势监测和生产管理提供了一种有效技术手段。

关键词:花生;表型性状估算模型;多源数据融合;机器学习;无人机遥感

中图分类号: S31; TP79

文献标识码: A

文章编号: 1000-1298(2026)01-0114-11

OSID:



Peanut Phenotype Estimation Model Based on Multi-source Data from Unmanned Aerial Vehicles

HE Ning^{1,2} WANG Jian³ LU Xianju^{1,2} CHEN Bo^{1,2} BAI Bo^{4,5} FAN Jiangchuan^{1,2}

(1. Information Technology Research Center, Beijing Academy of Agriculture and Forestry Sciences, Beijing 100097, China

2. Beijing Key Laboratory of Digital Plant, Beijing 100097, China

3. Beijing Digital Agriculture Rural Promotion Center, Beijing 101117, China

4. Institute of Crop Germplasm Resources, Shandong Academy of Agricultural Sciences, Jinan 250100, China

5. Shandong Provincial Key Laboratory of Crop Genetic Improvement and Ecological Physiology, Jinan 250100, China)

Abstract: Peanut (*Arachis hypogaea* L.), a critical oilseed crop, plays a crucial role in ensuring food and oil production security. Accurate, nondestructive, and real-time phenotypic monitoring is essential for optimizing peanut production management. Multispectral data acquired by an unmanned aerial vehicle (UAV) platform during key growth stages were leveraged to extract canopy multispectral (MS), structural (CHM), and textural (TEX) parameters. Four machine learning algorithms, partial least squares regression (PLSR), support vector machine (SVM), artificial neural network (ANN), and random forest regression (RFR), were employed to construct estimation models for plant height, SPAD values, and aboveground biomass. Results demonstrated strong correlations between peanut aboveground biomass/plant height and the near-infrared band (Pearson correlation coefficients were 0.77 and 0.69, respectively). The random forest model, integrating textural, structural, and spectral features, achieved optimal biomass estimation accuracy ($R^2 = 0.96$). For plant height inversion, the PLSR model

收稿日期: 2025-03-20 修回日期: 2025-04-21

基金项目: 山东省重点研发计划项目(2021LZGC026)和北京市农林科学院作物表型组学协同创新中心项目(KJCX20240406)

作者简介: 何宁(1997—),男,博士生,主要从事农业信息技术研究,E-mail: a807299517@163.com

通信作者: 王剑(1975—),男,高级工程师,主要从事农业信息技术研究,E-mail: wang_aojie@126.com

combining textural and spectral features performed best ($R^2 = 0.94$). SPAD estimation using PLSR with fused textural and structural features yielded moderate accuracy ($R^2 = 0.39$, RMSE = 3.06, nRMSE = 0.062, RPD = 1.30). The research identified feature-specific requirements for machine learning-based estimation of distinct peanut phenotypic traits and established a UAV multi-source data fusion framework capable of accurate, nondestructive, and efficient assessment of plant height and biomass. These findings can provide a robust technical approach for growth monitoring and precision management in peanut cultivation systems.

Key words: peanut; phenotypic trait estimation model; multi-source data fusion; machine learning; UAV remote sensing

0 引言

花生是大宗经济作物和粮油产量的主要来源，其作为重要的油料作物在中国国民经济和社会发展以及农民收入中占有重要地位。株高、叶片叶绿素相对含量 (SPAD 值) 以及地上部干物质量对花生的生长发育起着至关重要的作用^[1]，其中株高通过影响作物的冠层结构对花生的光合生产等生理过程产生影响，叶片 SPAD 值直接影响光合生产的效率，而地上部干物质量是光合同化物积累分配后的结果^[2]。目前对花生株高和叶绿素相对含量的测定，可使用测量工具对植株进行直接测定，对花生植株破坏较小，但地上部干物质量多采用破坏性取样，其中 SPAD 值也可通过紫外分光光度法、荧光分析法和活体叶绿素分析法测量，也需要进行破坏性取样，这些分析方法不仅效率较低而且对试验区其他植株有一定影响^[3]。

随着无人机在农业上的快速应用，大量研究已证明其在作物不同生育时期中表型研究的可行性^[4-5]。基于无人机平台搭载的传感器镜头在作物表型参数估计中精度较高，不同传感器镜头获取的数据类型在表型反演中存在差异^[6]。例如，多光谱图像数据可通过光谱、纹理以及结构等信息对作物的表型指标、成熟度和作物病害进行检测分析^[2,7]，热红外传感器主要通过测量作物冠层温度和作物对水分胁迫的响应来监测作物生长^[8]，高光谱镜头通过比多光谱相机更广的波段对作物表型有更加精确的反演^[9]。虽然无人机具有高效性和快捷性，同时可搭载多种传感器，其获得的单独数据源应用效果较好，但单数据源反演多种表型指标的精度均有待提高。

传感器获取的多种数据源进行多源数据融合相对于单数据源的处理精度高。冠层结构和光谱信息融合可改善植物性状估计^[5,10]，如作物生物量、叶面积指数 (LAI) 和籽粒产量。通过结合不同传感器系统中包含的冠层光谱、结构、热、纹理信息，在多种农业应用中使用多种传感器的特征来提高植物性状估计的潜力^[11-12]，同时，有研究报道将冠层热信息与

光谱及结构特征相结合，提高了作物在不同天气条件和发育阶段下产量预测的鲁棒性^[13]。纹理信息以图像中像素强度的空间变化为特征，可突出植物冠层结构和几何特征，并有助于抑制不适合景观高空间异质性模型饱和^[14]。尽管纹理信息在前人的许多研究中已被广泛用于图像分类和分割^[15]，但直接评估图像纹理在作物表型性状预测^[16]中的潜力鲜有报道；纹理参数 (TEX) 中固有的空间信息在花生表型性状估算中的应用还未见报道。因此，纹理信息有应用于花生表型预测的潜力。

许多基于统计和机器学习的回归方法^[13,17-18]，如多元线性回归、偏最小二乘回归、人工神经网络、随机森林回归和支持向量回归等，已被用于实现多种作物的精确表型指标预测。结合多模态数据融合配合机器学习方法在多种作物的表型性状和产量估算上得到了一定应用^[4]，并表现出较高的预测精度和鲁棒性^[19]，但在花生的表型性状估算中未见报道。

综上所述，基于无人机搭载多类型传感器，通过多源数据融合结合机器学习的回归模型已被用于多种作物性状估计，但目前在花生方面研究还较少。因此，本研究通过无人机搭载多光谱传感器获取花生冠层多模态数据，通过多源数据融合结合机器学习算法构建株高、生物量和 SPAD 值估算模型，以实现花生冠层表型的高通量精准解析。

1 材料与方法

1.1 试验地概述

本试验于山东省泰安市宁阳县山东省农业科学院试验基地 (35°50'57"N, 116°39'50"E) 开展，试验地土壤为壤土。于 2023 年 5 月 10 日播种，采用等行距起垄种植，垄长 4 m，垄宽 85 cm，垄面宽 60 cm，一垄双行，行距 25 cm，株距 10 cm，各小区面积均为 2.4 m²，种植密度为 3.3 × 10⁵ 株/hm²，选取山东地区主推品种共 43 个品种作为试验材料。

1.2 测定项目与方法

分别于花生关键生育期苗期 (6 月 7 日)、开花下针期 (7 月 5 日)、结荚期 (8 月 1 日)、成熟期 (9

月 5 日)进行无人机数据获取,同步开展地面实测工作。

1.2.1 株高、SPAD 值和地上部生物量测定

地上部干物质量:在 4 个关键生育时期对 43 个小区选取 3 株长势一致且具有代表性的花生植株,取植株地上部分将其放入干燥箱中 105℃ 杀青 30 min,将温度调至 75℃ 干燥至恒质量并称量记录器官干质量。

株高和 SPAD 值:在 4 个关键生育期对 43 个小区选取 5 株长势一致且具有代表性的花生植株进行原位株高测量,同时采用 SPAD 仪(SPAD-502 型)测定顶部成熟叶片 SPAD 值。

1.2.2 无人机多光谱、图像数据获取

无人机平台使用 DJI M300 RTK 型多旋翼无人机(深圳大疆科技有限公司),配备 Micasense Altum 多光谱相机(AgEagle Sensor Systems Inc.,美国),该相机可同时收集 5 个波段数据,包括可见光(红光、绿光和蓝光)、近红外光和红边波段,能实现多传感器数据同步获取。为获得高质量数据,选择在晴朗、无风或微风、无云的天气状况下在 10:00—14:00 采

集,使用反射标定板和光照传感器对数据进行校准。无人机航线规划采用航点飞行模式,航线覆盖度为 85%,无人机飞行高度为 30 m,飞行速度设置为 1 m/s,航线转弯半径为 5 m。

1.3 数据处理

1.3.1 预处理

将所获取的关键生育期光谱影像使用 Pix4D 软件进行拼接。将不同生育期获取的无人机原始影像和各波段校准板数据导入 Pix4D 软件,进行辐射定标、图像对齐、优化相机以及构建正射影像。将合成好的各波段影像导入 ArcMap 软件进行波段组合并导出。将正射影像导入 ENVI 并进行地理配准、感兴趣区域绘制、掩膜消除土壤背景以及提取植被指数等操作。

1.3.2 冠层光谱参数

在总结前人工作的基础上,选择用于估算作物生物量、株高和 SPAD 值的冠层光谱参数(MS,表 1)。这些参数不仅包括传统植被指数,基于 RGB 图像,还考虑了近红外波段和红边波段信息,以更全面地反映作物的生长状况。

表 1 多光谱图像提取的冠层光谱参数以及其计算公式

Tab.1 Canopy spectral parameters extracted from multispectral images

植被指数	计算公式	文献序号
比值植被指数(RVI)	$RVI = R_{NIR} / R_R$	[20]
绿度植被指数(GCI)	$GCI = R_{NIR} / R_G - 1$	[21]
归一化差值植被指数(NDVI)	$NDVI = (R_{NIR} - R_R) / (R_{NIR} + R_R)$	[22]
绿色归一化差值植被指数(GNDVI)	$GNDVI = (R_{NIR} - R_G) / (R_{NIR} + R_G)$	[23]
绿红植被指数(GRVI)	$GRVI = (R_G - R_R) / (R_G + R_R)$	[24]
增强型植被指数(EVI)	$EVI = 2.5(R_{NIR} - R_R) / (R_{NIR} + 6R_R - 7.5R_B + 1)$	[25]
双波段增强型植被指数(EVI2)	$EVI2 = 2.5(R_{NIR} - R_R) / (R_{NIR} + 2.4R_R + 1)$	[25]
转换型叶绿素吸收反射率指数(TCARI)	$TCARI = [R_{NIR} - R_R - 0.2(R_{NIR} - R_G)R_{NIR} / R_R]3$	[25]
优化型土壤调节植被指数(OSAVI)	$OSAVI = (R_{NIR} - R_R) / (R_{NIR} + R_R + 0.16)$	[22]
修正型叶绿素吸收比指数(MCARI)	$MCARI = [R_{RE} - R_R - 0.2(R_{RE} - R_G)]R_{RE} / R_R$	[26]
组合植被指数(MCARI/OSAVI)	$MCARI / OSAVI$	[26]
宽动态范围植被指数(WDRVI)	$WDRVI = (0.12R_{NIR} - R_R) / (0.12R_{NIR} + R_R)$	[27]
重归一化差值植被指数(RDVI)	$RDVI = (R_{NIR} - R_R) / \sqrt{R_{NIR} + R_R}$	[28]
修正型三角植被指数 2(MTVI2)	$MTVI2 = 1.5[1.2(R_{NIR} - R_G) - 2.5(R_R - R_G)] / \sqrt{2(2R_{NIR} + 1) - (6R_{NIR} - 5\sqrt{R_R})} - 0.5$	[26]
归一化差值红边指数(NDRE)	$NDRE = (R_{NIR} - R_{RE}) / (R_{NIR} + R_{RE})$	[29]
简化冠层叶绿素含量指数(SCCCI)	$SCCCI = NDRE / NDVI$	[30]
三角植被指数(TVI)	$TVI = 0.5[120(R_{NIR} - R_G) - 200(R_R - R_G)]$	[31]
差值植被指数(DVI)	$DVI = R_{NIR} - R_R$	[32]
修正型土壤调节植被指数(MSAVI)	$MSAVI = 0.5[2R_{NIR} + 1 - \sqrt{((2R_{NIR} + 1)^2 - 8(R_{NIR} - R_R))}]$	[33]
归一化色素叶绿素指数(NPCI)	$NPCI = (R_R - R_B) / (R_R + R_B)$	[34]
红光波段(R)		
绿光波段(G)		
蓝光波段(B)		
红边波段(RE)		
近红外波段(NIR)		

注: R_B 为蓝光波段反射率, R_G 为绿光波段反射率, R_R 为红光波段反射率, R_{RE} 为红边波段反射率, R_{NIR} 为近红外波段反射率。

1.3.3 冠层结构参数

从无人机 RGB 图像提取冠层高度,并将其作为冠层结构特征用于表型指标的反演和预测。在各关键生育时期利用无人机获取高分辨率的 RGB 图像,使用拼接好的 RGB 图像构建的三维点云创建裸露地面数字高程模型 (Digital elevation model, DEM) 和地面地形基础数据的数字地面模型 (Digital terrain model, DTM),用 DEM 图像和 DTM 图像做差值运算,得到冠层结构参数 (CHM) 图像,再提取对应小区的 CHM 来作为模型的特征值。

1.3.4 冠层纹理参数

将无人机多光谱 5 个波段的光谱图像进行二阶灰度共生矩阵计算并选择均值 (Mean)、方差 (Variance)、同质性 (Homogeneity)、对比度 (Contrast)、相异性 (Dissimilarity)、熵 (Entropy)、二阶矩 (Second Moment) 和相关性 (Correlation), 8 种纹理信息参数作为冠层纹理参数 (TEX) 特征输入。使用 ENVI 5.4.1 软件计算相关参数。

1.4 相关性分析和模型精度评价

1.4.1 相关性分析

使用皮尔逊相关系数进行相关性分析,使用皮

尔逊相关系数绝对值来衡量各个指标与花生多维表型指标之间的相关性^[35]。相关系数绝对值 0.8 ~ 1 为极强相关,0.6 ~ 0.8 为强相关,0.4 ~ 0.6 为中等相关,0.2 ~ 0.4 为弱相关,0 ~ 0.2 为不相关。

1.4.2 模型精度评价

选取偏最小二乘回归 (PLSR)、支持向量机 (SVM)、随机森林回归 (RFR) 和人工神经网络 (ANN) 4 种常见且性能良好的机器学习算法进行建模分析,基于 1.2 节中 4 个关键生育期 (每个时期均对 43 个品种进行取样) 所采集的 172 组样本数据构建模型,选择的训练和验证样本数量分别为 114 和 58 (2:1)。采用决定系数 (R^2)、均方根误差 (RMSE)、归一化均方根误差 (nRMSE) 和百分偏差比率 (RPD) 评估模型精度^[36]。其中 RPD 大于 3 为模型预测精度较好,2 ~ 3 为模型预测精度良好,1.5 ~ 2 为模型预测精度一般,小于 1.5 为较差。

试验流程如图 1 所示。通过对无人机光谱图像提取不同特征源,同时将单源特征和多源特征融合后在不同模型中预测效果进行比较分析,进一步确定最优模型。

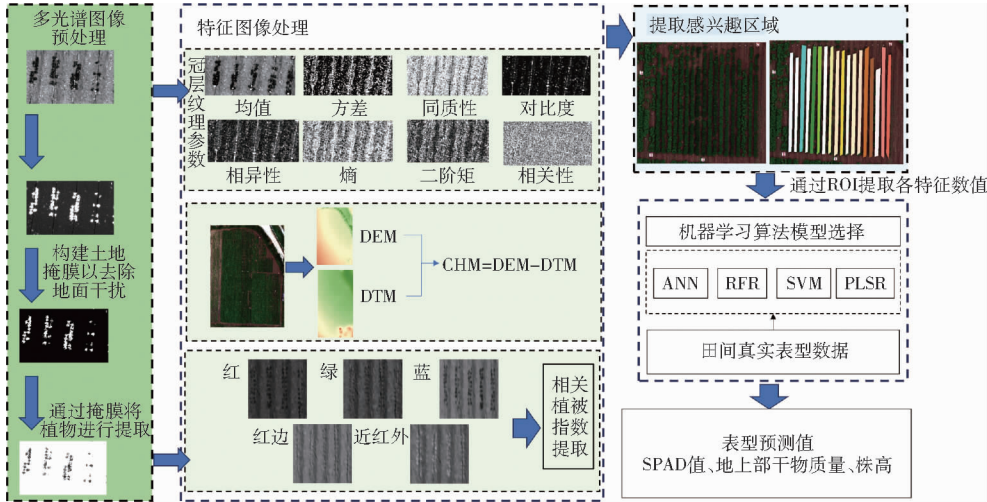


图 1 试验流程图

Fig.1 Flow chart of test

2 结果与分析

2.1 不同表型指标与特征输入相关性分析

3 种表型指标 (地上部干物质量 (AGB)、株高 (PH) 和 SPAD 值) 与多光谱单波段、植被指数以及 CHM 相关性分析如图 2 所示。由图 2 可以看出,光谱特征与株高和地上部生物量之间存在较为显著的相关性,与 SPAD 值的相关性较低。具体而言,株高与可见光波段相关性较弱,与红边波段呈中等相关性 (相关系数为 0.57),近红外波段显示为强相关性 (相关系数为 0.77),地上部生物量表现出同样的规

律,而 SPAD 值则对可见光波段表现出中等相关性 (蓝、绿、红波段相关系数分别为 0.54、0.41 和 0.52),对红边和近红外波段则相关性较弱。

大多数植被指数与株高和地上部生物量有较高的相关性,与 SPAD 值相关性较低。具体来看,EVI、EVI2、TCARI、OSAVI、RDVI、MTVI2、TVI、DVI、MSAVI 与株高和地上部干物质量均表现出强相关,SPAD 值则只有 RVI、GCI、NDVI、GNDVI、WDRVI、NDRE 显示出了中等相关性。

CHM 与株高、SPAD 值、地上部生物量的相关系数分别为 0.67、0.25 和 0.70。可见结构特征对株

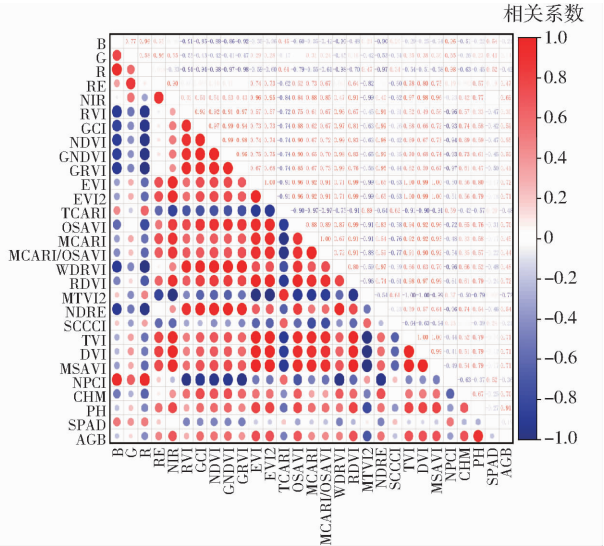


图2 表型指标与冠层光谱和结构参数相关性分析

Fig.2 Correlation analysis of phenotypic indicators with canopy spectral and structural parameters

高和地上部生物量的解释性较好。

对可见光波段(蓝绿红)提取的8种冠层纹理参数与3种表型指标进行相关性分析,如图3所示。由图3可以看出,可见光波段的冠层纹理参数与3种表型指标相关性较弱,其中,SPAD值与蓝波段冠层纹理参数呈中等相关性,绿波段冠层纹理参数与株高和地上部干物质量均呈中等相关性,地上部干物质量与红波段冠层纹理参数呈中等相关性。具体而言,株高与绿波段冠层纹理参数中的同质性(相关系数为-0.58)和相异性(相关系数为0.56)相关性较高,SPAD值与蓝波段冠层纹理参数中的均值(相关系数为0.46)相关性较高,地上部干物质量与红波段冠层纹理参数中的方差(相关系数为0.64)相关性较高。

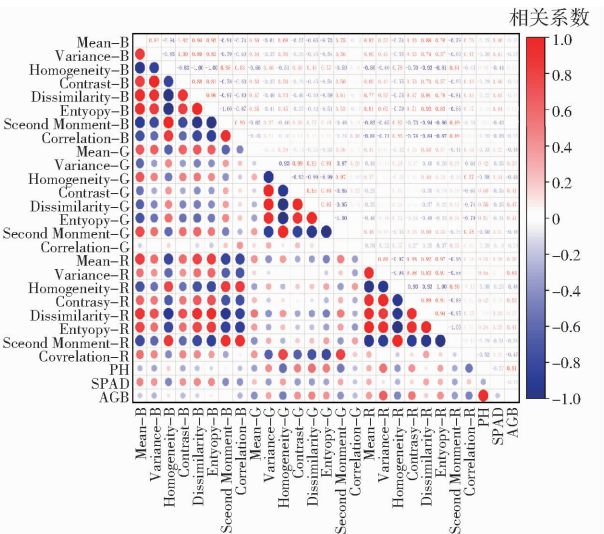


图3 表型指标与可见光波段冠层纹理参数相关性分析

Fig.3 Correlation analysis of phenotypic indicators with canopy texture parameters in visible band

红边和近红外波段提取的8种冠层纹理参数与3种表型指标相关性分析如图4所示。由图4可以看出,近红外波段的TEX对株高和地上部干物质量的相关性远高于红边波段,其中株高与近红外波段冠层纹理参数中的均值(相关系数为0.79)、方差(相关系数为0.89)、对比度(相关系数为0.89)以及相异性(相关系数为0.86)表现出强相关性,地上部干物质量与近红外波段冠层纹理参数中的均值(相关系数为0.80)、方差(相关系数为0.91)、对比度(相关系数为0.92)以及相异性(相关系数为0.86)表现出极强相关性。SPAD值与红边波段的冠层纹理参数呈中等相关性,其中均值的相关性最高,为负相关关系,相关系数为-0.53。

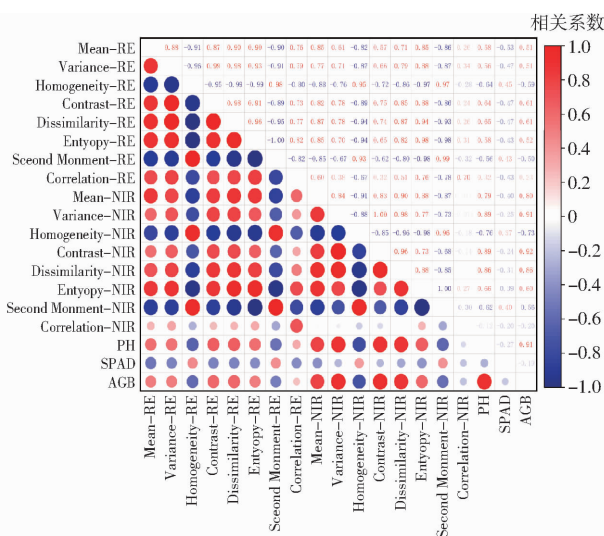


图4 表型指标与红边和近红外波段冠层纹理参数相关性分析

Fig.4 Correlation analysis of phenotypic metrics with canopy texture parameters in red edge and near-infrared bands

2.2 不同特征融合下花生表型指标模型估算性能评价

利用支持向量机、随机森林、偏最小二乘以及人工神经网络4种机器学习模型对融合不同特征数据源进行花生3种表型指标估算分析。

2.2.1 多源特征融合对花生株高的估算模型性能评价

各机器学习训练集反演株高的评价指标如图5(图中横坐标的A~G分别表示不同数据融合类型,其中A为CHM、B为MS、C为TEX、D为MS+CHM、E为CHM+TEX、F为MS+TEX、G为MS+CHM+TEX,下同)所示。由图5中可以看出,融合多源数据在4个模型中训练精度提高,融合多源数据后随机森林较其他模型反演精度(R^2 为0.95~0.96)、误差(RMSE为1.48~1.63 cm, nRMSE为0.091~0.096)以及精准性(RPD为4.6~5.2)均最稳定。具体而言,SVM对冠层纹理参数特征值较为

敏感,冠层纹理参数以及融合纹理和结构两者的特征源的株高训练精度最好 (R^2 分别为 0.99 和 0.98),误差和精准度同样也较好;RF、PLSR 以及 ANN 在融合多源数据后,各模型对株高的训练效果显著提升,同时可以发现将 TEX 作为特征值输入时,这 3 种模型均取得了最优的训练效果,精度分别为 0.96、0.85 和 0.86。

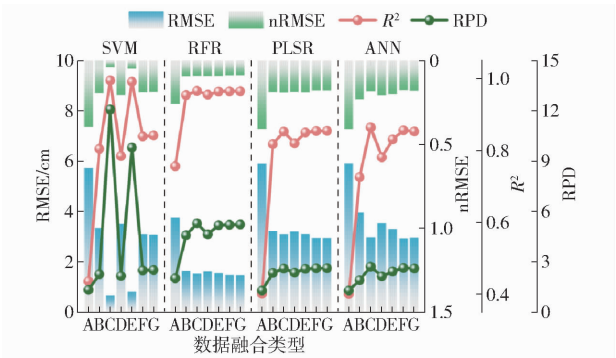


图 5 不同模型在不同输入特征类型和数量下训练集株高预测性能

Fig.5 Strain height prediction performance of different models with different input feature types and numbers in training set

从图 6 可以看出,RF 对多源数据融合的响应最小,其他 3 个模型均在 TEX 的基础上融合其他特征源后模型性能显著提高。融合多源数据在 4 个模型中测试精度均提高,融合光谱特征和冠层纹理参数后 PLSR 取得了最优训练反演精度 (R^2 为 0.94)、误差 (RMSE 为 1.62 cm,nRMSE 为 0.158) 以及精准性 (RPD 为 4.34)。除 PLSR 外,融合光谱、结构以及冠层纹理参数在 SVM、PLSR 和 ANN 模型中表现出最优的株高反演结果。

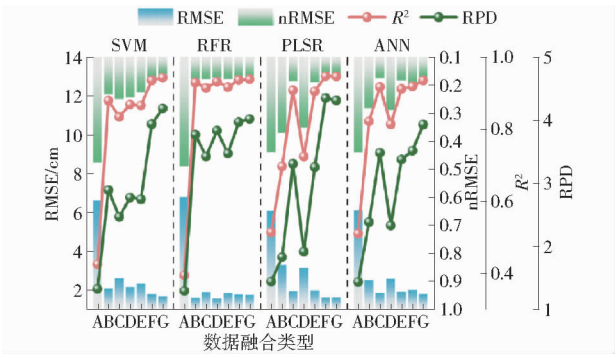


图 6 不同模型在不同输入特征类型和数量下测试集株高预测性能

Fig.6 Strain height prediction performance of different models for different input feature types and numbers in test set

2.2.2 多源特征融合对花生 SPAD 值的模型估算性能评价

从各机器学习训练集反演 SPAD 值的评价指标

(图 7) 中可以看出,融合光谱特征和结构特征在 SVM 模型中表现出较强的训练效果,融合多源数据后随机森林较其他模型训练反演精度 (R^2 为 0.82 ~ 0.84)、误差 (RMSE 为 1.87 ~ 1.92,nRMSE 为 0.039 ~ 0.040) 以及精准性 (RPD 为 2.36 ~ 2.57) 均最稳定。

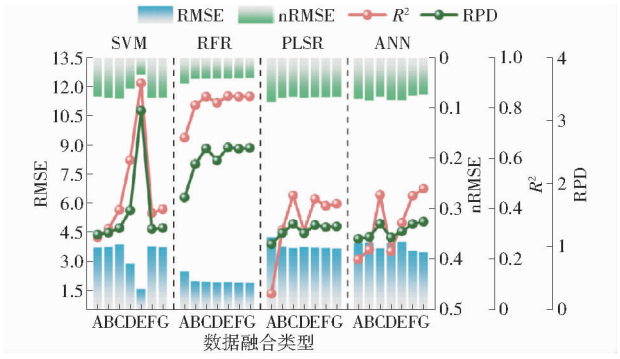


图 7 不同模型在不同输入特征类型和数量下训练集 SPAD 值预测性能

Fig.7 SPAD prediction performance of different models with different input features types and numbers in training set

从图 8 可以看出,除 SVM 外,RF、PLSR 以及 ANN 模型在融合结构和冠层纹理参数反演结果达到最优,其中反演精度分别为 36.6%、39.8% 和 37.7%。单数据源中各机器学习模型在将冠层纹理参数作为输入时均表现为最优的反演效果,反演精度 R^2 分别为 0.34 (SVM)、0.34 (RF)、0.39 (PLSR) 和 0.33 (ANN)。

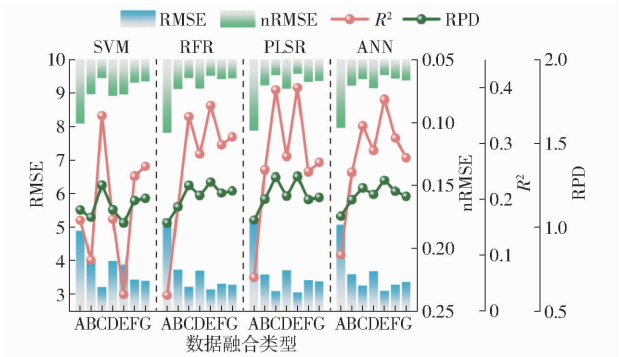


图 8 不同模型在不同输入特征类型和数量下测试集 SPAD 值预测性能

Fig.8 SPAD prediction performance of different models for different input feature types and numbers in test set

2.2.3 多源特征融合对花生地上部干物质质量的模型估算性能评价

从图 9 可以看出,融合多源数据在 4 个模型中训练精度均提高,融合多源数据后随机森林较其他模型训练精度 (R^2 为 0.95 ~ 0.96)、误差 (RMSE 为 2.76 ~ 3.07 g/株,nRMSE 为 0.132 ~ 0.140) 以及精准性 (RPD 为 4.9 ~ 5.6) 均最稳定。具体而言,融合结构特征和光谱特征对各模型的训练效果均有一定程度的降低。

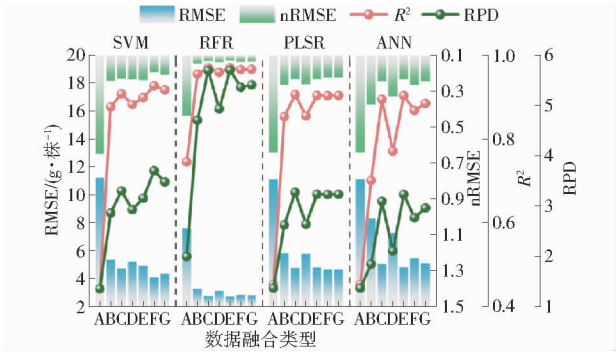


图 9 不同模型在不同输入特征类型和数量下训练集地上干物质质量预测性能

Fig. 9 Above-ground dry matter quality prediction performance of different models for different input feature types and numbers in training set

从图 10 可以看出,融合结构、光谱和冠层纹理参数后 RF 模型反演效果最优(R^2 为 0.96, RMSE 为 2.52 g/株, nRMSE 为 0.267, RPD 为 5.19),在其他模型中融合结构和冠层纹理参数表现出较好的反演效果,单数据源中冠层纹理参数在各模型中效果最优。

2.3 多源数据融合下花生表型指标最优估计模型

株高实测值与最优模型预测值交叉验证散点图如图 11 所示,其中 SVM 特征值输入为 MS + CHM + TEX, RFR 特征值输入为 MS + CHM + TEX, PLSR 特征值输入为 MS + TEX, ANN 特征值输入为 MS + CHM + TEX。总体来看,PLSR 反演生育期的效果最

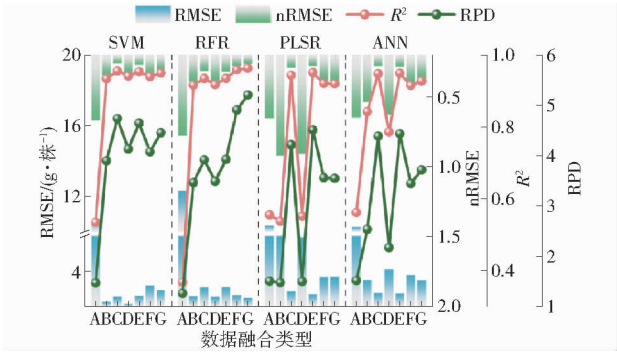


图 10 不同模型在不同输入特征类型和数量下测试集地上干物质质量预测性能

Fig. 10 Above-ground dry matter quality prediction performance of different models for different input feature types and numbers in test set

优(R^2 为 0.94)。由图 11 可知,模型对在 20 cm 以下花生植株株高反演结果较为精准,其中 SVM、RF 以及 ANN 融合多源数据后预测株高结果均达到最佳。

SPAD 实测值与最优模型预测值交叉验证散点图如图 12 所示,其中 SVM 特征值输入为 TEX, RFR 特征值输入为 CHM + TEX, PLSR 特征值输入为 CHM + TEX, ANN 特征值输入为 CHM + TEX。总体来看,模型对 SPAD 值的反演效果均不理想,其中 PLSR 反演生育期的结果最优(R^2 为 0.39),RF 在训练集中预测结果较为集中,但在测试集中预测结果一般。

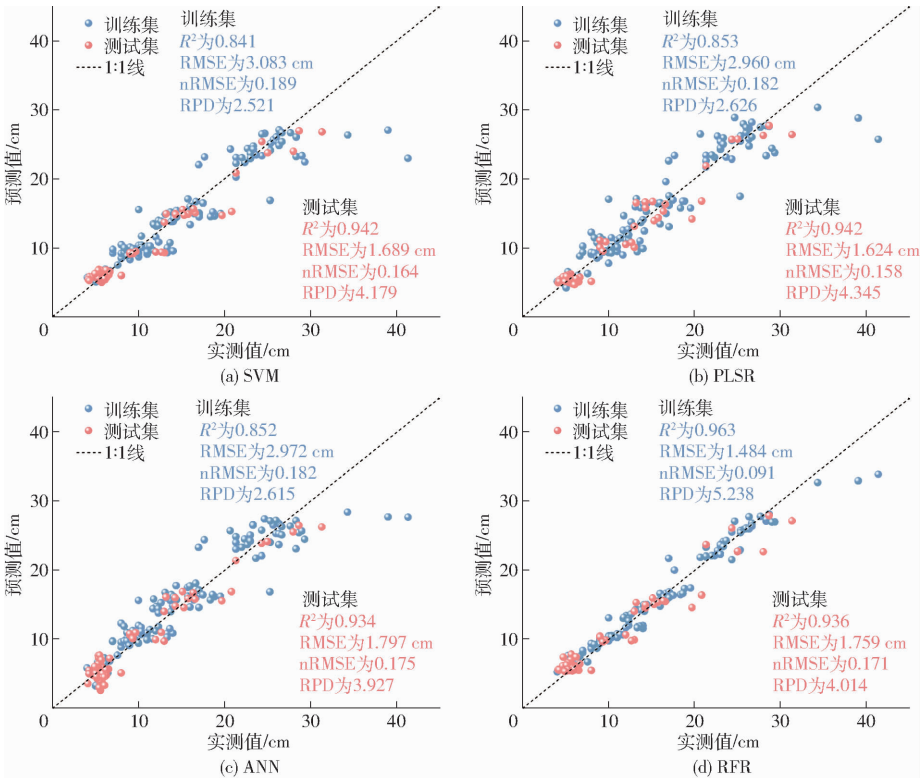


图 11 株高实测值与预测值的交叉验证散点图

Fig. 11 Cross-validated scatterplots of measured and predicted values of plant height

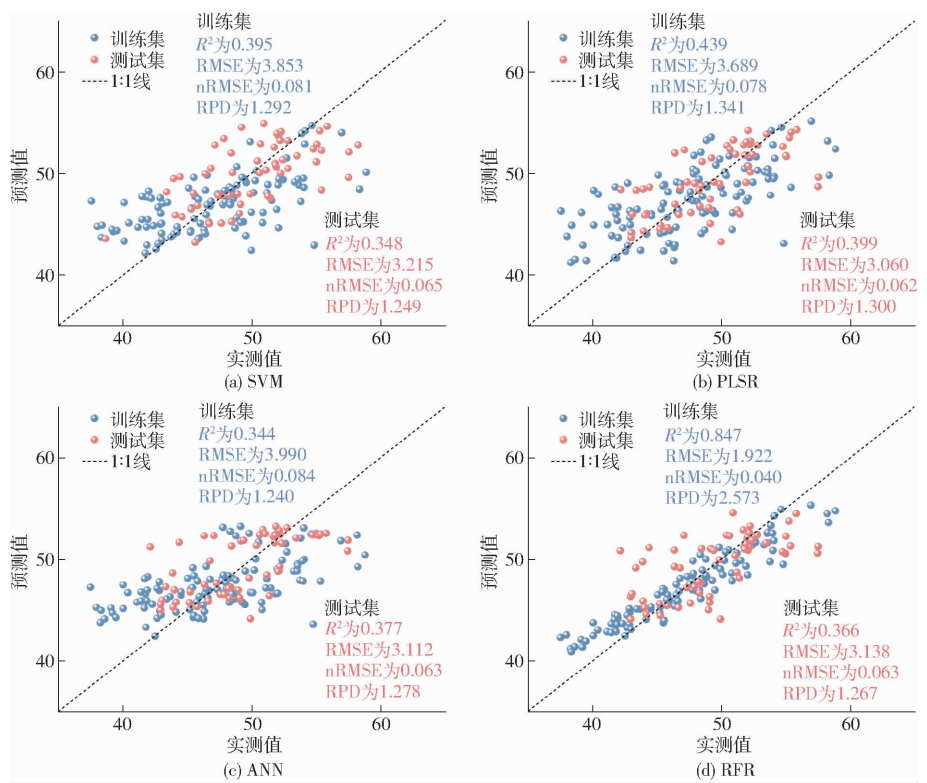


图 12 SPAD 值实测值与预测值的交叉验证散点图

Fig. 12 Cross-validated scatterplots of SPAD measured and predicted values

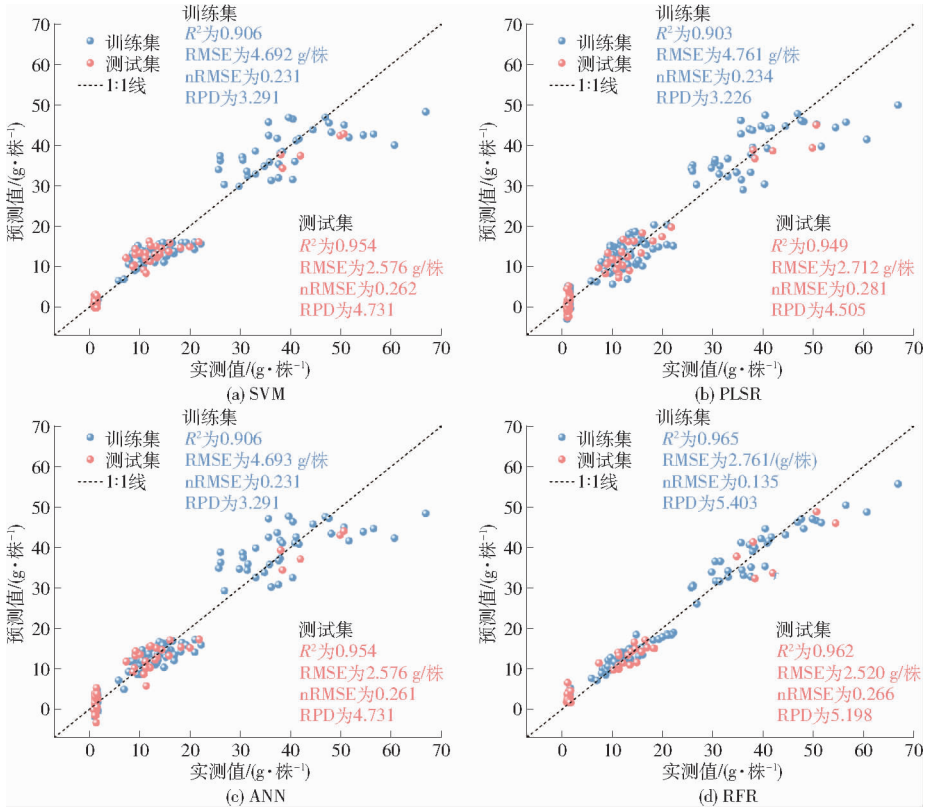


图 13 地上部干物质质量实测值与预测值的交叉验证散点图

Fig. 13 Cross-validated scatterplots of measured and predicted dry matter mass in above-ground sector

通过对地上部干物质质量实测值与最优模型的预测值交叉验证散点图如图 13 所示,其中 SVM 特征值输入为 TEX,RFR 特征值输入为 MS + CHM + TEX,PLSR 特征值输入为 CHM + TEX,ANN 特征值

输入为 CHM + TEX。总体来看,各模型对于干物质积累量在 30 g/株以下花生植株的反演结果较为精准,其中 RF 同时输入 3 种冠层特征参数进行预测全生育期地上部干物质质量结果达到最优 (R^2 为 0.96),

SVM 在仅使用一个特征参数作为输入源的情况下取得较好模拟结果(R^2 为 0.95),可见其对冠层纹理参数表现出较好的特征响应。

3 讨论

作物表型是农学研究中的重要课题之一,随着无人机等高通量获取手段不断更新,使高效快速获取表型数据成为可能^[37]。花生是我国重要的粮油作物,其生长发育情况直接影响花生产量进而对粮油产量产生影响。目前花生高通量表型获取和研究并未有详细报道,大多侧重于对花生单维度表型指标的研究。

比较分析 3 种表型指标与不同数据源的相关性,可以看出光谱特征中蓝、绿、红波段对 SPAD 值的影响高于红边和近红外波段,而株高和地上部干物质量则与红边和近红外有较好的相关性,由于 SPAD 值主要反映颜色层面信息^[38],故在可见光波段内表现出一定相关性,而地上部干物质量和株高则更多地表示长势和结构信息,在可见光波段难以直观地进行反映,有研究也表明在近红外波段能更好地解析结构信息^[39]。本研究中根据 5 个波段所计算的植被指数中,只有 RVI、GCI、NDVI、GNDVI、WDRVI、NDRE 与 SPAD 值呈中等相关,这也导致在模型估算 SPAD 值出现了不理想的效果(R^2 小于 0.50),而株高和地上部干物质量则与 EVI、EVI2、TCARI、OSAVI、RDVI、MTVI2、TVI、DVI、MSAVI 呈强相关,这导致株高和地上部干物质量在只将光谱作为输入时估算准确率达到 60% 以上,融合其他特征源后 R^2 能达到 0.90 以上。本研究的纹理指数将无人机获取多光谱图像的 5 个波段均进行了二阶概率统计滤波^[40] 计算,得到的 8 个冠层纹理参数对于不同表型指标相关性表现出差异,其中蓝光波段和红边波段的 8 种纹理指数对 SPAD 值影响较大,而近红外波段下冠层纹理参数对株高和地上部干物质量均表现出强相关,与绿光波段和红边波段的部分冠层纹理参数也呈中等相关,这也解释了在单数据源模型估算中纹理指数在 3 种表型指标中均取得最优效果。

根据最优模型(图 11~13)结果可明显看出,机器学习模型对株高 20 cm 以下、地上部干物质量小于 30 g/株花生植株的反演效果较好,由于花生植株

比较矮小,生物量和株高较低,反演容易出现拟合聚集的情况,可以考虑分开时期进行模型构建。通过对特征值与多维表型指标的相关性分析可以明确,近红外波段的相关特征值对株高和地上部干物质量的反演预测起到重要作用,而可见光波段对 SPAD 值的预测和反演有一定促进作用,根据图像所提取出的冠层纹理参数也有同样规律。但可以明显看出,本试验根据文献[20-34]所提出的光谱植被指数对模型并未产生重要影响,只有 RVI、GCI、NDVI、GNDVI、WDRVI、NDRE 对模型产生了一定影响。

总体来看,融合多源数据后机器学习模型均会提升反演精度,但并不是将结构、光谱和冠层纹理参数都进行融合作为特征值输入就一定会使模型精度有所提高。具体来说,RFR 对于融合多源数据后训练集的反演效果较为稳定,同时对多维表型指标的反演也取得了很好效果。SVM、PLSR 和 ANN 融合冠层纹理参数后的特征输入反演效果更好,SPAD 值各模型反演效果中,融合 TEX 的特征输入比融合 TEX + CHM + MS 的效果好。

4 结论

(1)花生叶片 SPAD 值与可见光波段以及关联特征相关性较高,株高和地上部生物量与红波段、红边波段、近红外波段相关性较高。

(2)多源数据融合对模型预测精度和模型稳定性有一定提高作用,但并不一定融合全部特征的估算结果最优。株高方面,RF 反演结果最稳定(R^2 为 0.91~0.93,RPD 为 3.4~4.0),同时融合 TEX 后各机器学习模型对株高的估算精度均有显著提高;SPAD 值方面,融合 TEX 后的 PLSR 模型反演结果最优(R^2 为 0.39, RMSE 为 3.06, nRMSE 为 0.062, RPD 为 1.30);地上部生物量方面,融合全数据特征后 RF 模型取得最优的反演结果(R^2 为 0.96, RMSE 为 2.52 g/株, nRMSE 为 0.267, RPD 为 5.19)。

(3)根据最优模型,该反演模型对株高 20 cm 以下以及地上部生物量小于 30 g/株的花生植株有较好的预测模拟效果,说明该模型对花生生长前期的估算效果优于生长后期。

(4)本研究为花生株高、SPAD 值和生物量的无损、高通量估算提供了一种新的有效工具,可满足低成本、大范围的试验需求。

参 考 文 献

[1] ZHANG J L, GENG Y, GUO F, et al. Research progress on the mechanism of improving peanut yield by single-seed precision sowing[J]. Journal of Integrative Agriculture, 2020, 19(8): 1919-1927.
[2] 陈龙跃, 段丹丹. 利用无人机多光谱监测大田花生长势[J]. 农业与技术, 2025, 45(1): 21-25.
CHEN Longyue, DUAN Dandan. Multi-spectral monitoring of peanut growth in field by UAV was used[J]. Agriculture and

- Technology, 2025, 45(1): 21–25. (in Chinese)
- [3] MONTEIRO S T, MINEKAWA Y, KOSUGI Y, et al. Prediction of sweetness and amino acid content in soybean crops from hyperspectral imagery[J]. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2007, 62(1): 2–12.
- [4] ZHAI W, LI C, FEI S, et al. CatBoost algorithm for estimating maize above-ground biomass using unmanned aerial vehicle-based multi-source sensor data and SPAD values[J]. *Computers and Electronics in Agriculture*, 2023, 214: 108306.
- [5] 承达瑜, 何伟德, 付春晓, 等. 融合无人机光谱信息与纹理特征的冬小麦综合长势监测[J]. *农业机械学报*, 2024, 55(9): 249–261.
- CHENG Dayu, HE Weide, FU Chunxiao, et al. Comprehensive growth monitoring of winter wheat by integrating UAV spectral information and texture features[J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2024, 55(9): 249–261. (in Chinese)
- [6] GALÁN R J, BERNAL-VASQUEZ A M, JEBSEN C, et al. Integration of genotypic, hyperspectral, and phenotypic data to improve biomass yield prediction in hybrid rye[J]. *Theoretical and Applied Genetics*, 2020, 133(11): 3001–3015.
- [7] 王猛, 隋学艳, 梁守真, 等. 利用无人机遥感技术提取农作物植被覆盖度方法研究[J]. *作物杂志*, 2020(3): 177–183.
- WANG Meng, SUI Xueyan, LIANG Shouzhen, et al. Research on the method of extracting crop vegetation coverage using UAV remote sensing technology[J]. *Crops*, 2020(3): 177–183. (in Chinese)
- [8] MAES W H, STEPPE K. Perspectives for remote sensing with unmanned aerial vehicles in precision agriculture[J]. *Trends in Plant Science*, 2019, 24(2): 152–164.
- [9] 刘涛, 刘望, 杨奉源, 等. 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(3): 206–217.
- LIU Tao, LIU Wang, YANG Fengyuan, et al. Research on estimating peanut biomass using UAV hyperspectral imaging and machine learning algorithm[J]. *Journal of China Agricultural University*, 2025, 30(3): 206–217. (in Chinese)
- [10] 张庆, 李金阳, 石文强, 等. 基于无人机光谱图像多数据融合的大豆地上生物量估测研究[J]. *中国农机化学报*, 2025, 46(1): 151–156, 163.
- ZHANG Qing, LI Jinyang, SHI Wenqiang, et al. Study on estimation of soybean aboveground biomass based on multi data fusion of unmanned aerial vehicle spectral images[J]. *Journal of Chinese Agricultural Mechanization*, 2025, 46(1): 151–156, 163. (in Chinese)
- [11] MAIMAITIJANG M, GHULAM A, SIDIKE P, et al. Unmanned aerial system (UAS)-based phenotyping of soybean using multi-sensor data fusion and extreme learning machine[J]. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2017, 134: 43–58.
- [12] SHI Y, THOMASSON J A, MURRAY S C, et al. Unmanned aerial vehicles for high-throughput phenotyping and agronomic research[J]. *PloS One*, 2016, 11(7): e0159781.
- [13] RISCHBECK P, ELSAYED S, MISTELE B, et al. Data fusion of spectral, thermal and canopy height parameters for improved yield prediction of drought stressed spring barley[J]. *European Journal of Agronomy*, 2016, 78: 44–59.
- [14] SIBANDA M, MUTANGA O, ROUGET M, et al. Estimating biomass of native grass grown under complex management treatments using worldview-3 spectral derivatives[J]. *Remote Sensing*, 2017, 9(1): 55.
- [15] MONGUS D, ŽALIK B. Segmentation schema for enhancing land cover identification: a case study using Sentinel 2 data[J]. *International Journal of Applied Earth Observation and Geoinformation*, 2018, 66: 56–68.
- [16] SHIU Y, CHUANG Y. Yield estimation of paddy rice based on satellite imagery: comparison of global and local regression models[J]. *Remote Sensing*, 2019, 11(2): 111.
- [17] KUWATA K, SHIBASAKI R. Estimating corn yield in the united states with MODIS EVI and machine learning methods[J]. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 2016, 3: 131–136.
- [18] JIN Z, AZZARI G, LOBELL D B. Improving the accuracy of satellite-based high-resolution yield estimation: a test of multiple scalable approaches[J]. *Agricultural and Forest Meteorology*, 2017, 247: 207–220.
- [19] ZHAI W, LI C, CHENG Q, et al. Exploring multisource feature fusion and stacking ensemble learning for accurate estimation of maize chlorophyll content using unmanned aerial vehicle remote sensing[J]. *Remote Sensing*, 2023, 15(13): 3454.
- [20] TUCKER C J. Red and photographic infrared linear combinations for monitoring vegetation[J]. *Remote Sensing of Environment*, 1979, 8(2): 127–150.
- [21] GITELSON A A, GRITZ Y, MERZLYAK M N. Relationships between leaf chlorophyll content and spectral reflectance and algorithms for non-destructive chlorophyll assessment in higher plant leaves[J]. *Journal of Plant Physiology*, 2003, 160(3): 271–282.
- [22] 周科, 陈鹏飞. 耦合多源无人机遥感数据和机器学习方法的玉米 SPAD 估测[J]. *作物学报*, 2025, 51(5): 1389–1399.
- ZHOU Ke, CHEN Pengfei. Maize SPAD estimation by combining multi-source unmanned aerial vehicle remote sensing data and machine learning methods[J]. *Acta Agronomica Sinica*, 2025, 51(5): 1389–1399. (in Chinese)
- [23] GITELSON A A, KAUFMAN Y J, MERZLYAK M N. Use of a green channel in remote sensing of global vegetation from

- EOS – MODIS[J]. Remote Sensing of Environment, 1996, 58(3): 289 – 298.
- [24] SRIPADA R P, HEINIGER R W, WHITE J G, et al. Aerial color infrared photography for determining early in-season nitrogen requirements in corn[J]. Agronomy Journal, 2006, 98(4): 968 – 977.
- [25] HABOUDANE D, MILLER J R, TREMBLAY N, et al. Integrated narrow-band vegetation indices for prediction of crop chlorophyll content for application to precision agriculture[J]. Remote Sensing of Environment, 2002, 81(2 – 3): 416 – 426.
- [26] HABOUDANE D, MILLER J R, PATTEY E, et al. Hyperspectral vegetation indices and novel algorithms for predicting green LAI of crop canopies: modeling and validation in the context of precision agriculture[J]. Remote Sensing of Environment, 2004, 90(3): 337 – 352.
- [27] GITELSON A A, ZUR Y, CHIVKUNOVA O B, et al. Assessing carotenoid content in plant leaves with reflectance spectroscopy[J]. Photochemistry and Photobiology, 2002, 75(3): 272 – 281.
- [28] ROUJEAN J L, BREON F M. Estimating PAR absorbed by vegetation from bidirectional reflectance measurements[J]. Remote Sensing of Environment, 1995, 51(3): 375 – 384.
- [29] DAVIDSON C, JAGANATHAN V, SIVAKUMAR A N, et al. NDVI/NDRE prediction from standard RGB aerial imagery using deep learning[J]. Computers and Electronics in Agriculture, 2022, 203: 107396.
- [30] RAPER T B, VARCO J J. Canopy-scale wavelength and vegetative index sensitivities to cotton growth parameters and nitrogen status[J]. Precision Agriculture, 2015, 16(1): 62 – 76.
- [31] BROGE N H, LEBLANC E. Comparing prediction power and stability of broadband and hyperspectral vegetation indices for estimation of green leaf area index and canopy chlorophyll density[J]. Remote Sensing of Environment, 2001, 76(2): 156 – 172.
- [32] LIU H Q, HUETE A. A feedback based modification of the NDVI to minimize canopy background and atmospheric noise[J]. IEEE Transactions on Geoscience and Remote Sensing, 1995, 33(2): 457 – 465.
- [33] QI J, CHEHBOUNI A, HUETE A R, et al. A modified soil adjusted vegetation index[J]. Remote Sensing of Environment, 1994, 48(2): 119 – 126.
- [34] RIEDELL W E, BLACKMER T M. Leaf reflectance spectra of cereal aphid-damaged wheat[J]. Crop Science, 1999, 39(6): 1835 – 1840.
- [35] CHEN B, HUANG G, LU X, et al. Prediction of vertical distribution of SPAD values within maize canopy based on unmanned aerial vehicles multispectral imagery[J]. Frontiers in Plant Science, 2023, 14: 1253536.
- [36] WU T, CHEN H, LIN Z, et al. Identification and quantitation of melamine in milk by near-infrared spectroscopy and chemometrics[J]. Journal of Spectroscopy, 2016, 2016(1): 6184987.
- [37] 刘仕元, 梁晋, 王帅斌, 等. 基于无人机遥感的花生叶片叶绿素含量监测研究[J]. 花生学报, 2020, 49(2): 21 – 27, 35.
- LIU Shiyuan, LIANG Jin, WANG Shuaibin, et al. Research on monitoring of peanut leaf chlorophyll content based on UAV remote sensing[J]. Journal of Peanut Science, 2020, 49(2): 21 – 27, 35. (in Chinese)
- [38] SHU M, LI Q, GHAFOR A, et al. Using the plant height and canopy coverage to estimation maize aboveground biomass with UAV digital images[J]. European Journal of Agronomy, 2023, 151: 126957.
- [39] MAIMAITIJIANG M, SAGAN V, SIDIKE P, et al. Soybean yield prediction from UAV using multimodal data fusion and deep learning[J]. Remote Sensing of Environment, 2020, 237: 111599.
- [40] NICHOL J E, SARKER M L R. Improved biomass estimation using the texture parameters of two high-resolution optical sensors [J]. IEEE Transactions on Geoscience and Remote Sensing, 2010, 49(3): 930 – 948.