

基于高光谱成像技术判别 马铃薯叶片干旱状态

梅轩铭¹, 胡耀华², 张浩天¹, 蔡雨卿¹,
罗凯天¹, 孟玉玲¹, 宋银¹, 单卫星¹

(1. 西北农林科技大学农学院, 陕西 杨凌 712100; 2. 浙江农林大学光电工程学院, 浙江 临安 311300)

摘要:为了准确快速地分级评估马铃薯叶片干旱状态,提出了基于高光谱成像技术的马铃薯叶片干旱状态分类方法。于2022年以青薯9号原原种为材料,通过使用高光谱成像仪获取3种不同干旱状态的马铃薯叶片,提取各类样本光谱反射率信息420个,讨论了4种光谱数据预处理方式对建模的影响。使用基于随机森林的交叉验证递归特征消除算法(RF-RFECV)与竞争性自适应重加权抽样法(CARS)进行特征波长选择,结合极端随机树(extremely randomized trees, EXT)构建了马铃薯叶片干旱状态的分类模型。结果表明,本研究构建的3个干旱状态分类模型,其测试集模型精度均高于85%,其中SNV-RF-RFECV-EXT模型表现最佳,测试集预测准确率达92.14%。同时,为直观地显示马铃薯叶片的干旱状态,选用建立的SNV-RF-RFECV-EXT模型对叶片进行干旱程度可视化,通过不同颜色直观显示叶片干旱状态,为马铃薯叶片干旱状态的判别提供了新方法。

关键词:马铃薯; 干旱状态判别; 高光谱成像; 特征波长; 模型构建

中图分类号:S532; S123 **文献标志码:**A

Evaluation of drought status of potato leaves based on hyperspectral imaging

MEI Xuanming¹, HU Yaohua², ZHANG Haotian¹, CAI Yuqing¹,
LUO Kaitian¹, MENG Yuling¹, SONG Ying¹, SHAN Weixing¹

(1. College of Agronomy, Northwest A&F University, Yangling, Shaanxi 712100, China;

(2. College of Optical, Mechanical and Electrical Engineering, Zhejiang A&F University, Lin'an, Zhejiang 311300, China)

Abstract: To evaluate the drought status of potato leaves accurately and quickly, a hyperspectral imaging method was employed to assess the drought status of potato leaves. In 2022, the original Qingshu No. 9 seed was used as the test material. Three types of drought status were obtained using hyperspectral imaging. A total of 420 spectral reflectance measurements from various samples were extracted and the influence of four spectral data pre-processing methods on modeling was discussed. A classification model for the drought status of potato leaves was constructed using the random forest cross-validation recursive feature elimination algorithm (RF-RFECV) and the feature wavelength selection method of competitive adaptive reweighted sampling (CARS) methods combined with extremely randomized trees (EXT). The results showed that the three drought status classification models constructed in this study had accuracy higher than 85% in the test set, among which the SNV-RF-RFECV-EXT model performed the best with a prediction accuracy of 92.14% in the test set. To visually assess the drought status of potato leaves, the SNV-RF-RFECV-EXT model was applied to depict the degree of drought. Different colors were used to represent varying levels of drought, offering a novel approach for distinguishing the drought conditions of potato leaves.

收稿日期:2024-03-26

修回日期:2024-07-31

基金项目:中国科学院战略性先导科技专项(XDA23070201);杨凌种业创新中心重点研发项目(Ylzy-mls-02)

作者简介:梅轩铭(2000-),男,河南信阳人,硕士研究生,研究方向为智慧农业。E-mail: meixuanming@nwfau.edu.cn

通信作者:单卫星(1967-),男,新疆新源人,教授,主要从事卵菌生物学研究。E-mail: wxshan@nwfau.edu.cn

宋银(1985-),男,陕西汉中,副教授,主要从事作物分子抗病育种研究。E-mail: yin.song@nwfau.edu.cn

Keywords: potato; drought stress identification; hyperspectral imaging; characteristic wavelengths; model construction

马铃薯 (*Solanum tuberosum* L.) 是仅次于水稻、小麦的世界第三大粮食作物^[1]。我国马铃薯种植面积与产量均多年位居世界第一^[2]。在自然生长条件下,马铃薯常受到各种逆境胁迫的影响,其中,干旱胁迫是马铃薯生长发育过程的主要逆境胁迫之一,常导致马铃薯的产量和品质下降^[3]。因此,检测马铃薯叶片干旱状态将有助于了解马铃薯的受旱情况,指导种植者合理灌溉,以保证植株正常的生理活动,降低干旱胁迫对马铃薯产量和品质造成的损失。相较于传统检测,高光谱成像技术只需要进行成像拍摄就能通过可视化结果获取叶片整体干旱状态,具有无损、快速、安全等优点,这些特点成为高光谱技术在作物干旱状态检测领域广泛应用的基础^[4]。

高光谱成像技术 (hyperspectral imaging, HSI) 是一种快速、高通量的植物表型分析方法。它能同时捕获植物的光谱 (λ) 与空间 (x, y) 信息,并将它们合并成一个称为“高光谱数据立方体”的三维数据矩阵;HSI 生成的高光谱图像含有大量的图像与光谱信息,可以用于提取植物叶片的结构和生理信息^[5]。HSI 已被应用于数个不同植物干旱状态的分析之中。Chen 等^[6]利用 HSI 建立了机器学习模型,监测干旱胁迫作用下茶树幼苗的干旱状态。Mishra 等^[7]利用光谱范围为 400~1 000 nm 的便携式高通量高光谱成像装置,采用无监督的 k-means 聚类算法研究拟南芥受干旱胁迫的程度,其结果表明,HSI 能够比传统的归一化植被差异指数 (NDVI) 检测方法更早地检测到拟南芥所遭受的干旱胁迫。Weksler 等^[8]使用高光谱移动相机对莴苣进行成像 (9:00—14:00, 每 1 h 记录 1 次),根据植物的反射率信号计算出 3 种植被指数:红缘叶绿素指数、光化学反射率指数和水分指数,并与植株生理指标测量值进行综合比较,发现在干旱胁迫处理阶段,每小时计算一次的光化学反射率指数值随瞬时蒸腾速率的变化而变化,可以用于监测植物的干旱状态。Zhou 等^[9]通过温室研究,验证了高光谱反射率是一种在干旱胁迫早期快速判别与评价柑橘树叶片光合性能的高通量表型判别方法。孙红等^[10]使用 HSI 实现了马铃薯叶片含水率的检测并且实现了不同干旱状态下马铃薯叶片含水率分布的可视化。当前基于高光谱技术判别叶片干旱状态大多利用

光谱指数或以光谱反射率构建回归模型计算抗旱指标,而直接进行叶片干旱状态分类的报道较少,对叶片干旱状态的可视化分级研究缺乏。

本研究探究运用高光谱成像技术快速判别马铃薯叶片的干旱状态的方法,以期实现马铃薯叶片干旱状态分类的像素级别可视化,为通过高光谱成像技术快速进行作物干旱状态检测提供参考依据。

1 材料与方法

1.1 研究概况与数据收集

研究在 2022 年于陕西省杨凌区西北农林科技大学 (南校区) 玻璃温室内进行,湿度为 50%~60%,温度为 20~25℃。研究用盆为直径 20 cm、高 16 cm 的塑料盆。每盆装有田园土 1.5 kg,营养土 0.5 kg。浇透水后,种植土壤的最大持水量参考土壤烘干法测定^[11]。供试马铃薯材料为青薯 9 号原原种,抗旱性较强,选取大小一致的马铃薯薯块喷施 20 mg · L⁻¹ 赤霉素促进打破休眠^[12],晾干后种入盆中。在马铃薯出苗后的第 20 天进行干旱胁迫处理,共设置 3 个不同的土壤水分梯度:第一组为土壤最大持水量的 70%~80%、第二组为土壤最大持水量的 40%~50%、第三组为土壤最大持水量的 20%~30%,每个处理设置 8 盆,总共 24 盆。结合植株整体生长发育状况以及生理生化状态,将这 3 组水分状况的处理分别定义为无干旱胁迫、中度干旱胁迫和重度干旱胁迫。分别于干旱胁迫处理后第 19 天和第 20 天将各个处理的马铃薯植株主茎上顶枝位置的第 1~3 叶用剪刀剪下,每个处理选取 35 片马铃薯叶片,总共 105 片马铃薯叶片,进行光谱图像数据收集。

本研究的图像采集使用 Resonon 台式高光谱成像系统 (Resonon Pika XC2, Montana, USA),该系统由 Pika 高光谱成像仪 (Pika hyperspectral imaging camera)、线性运动平台 (linear translation stage)、云台 (mounting tower)、照明系统 (lighting assembly) 以及控制软件系统 (software control system) 5 个部分组成 (图 1)。本研究数据收集在室内环境下进行,光源为光谱成像仪自带照明系统,设置仪器曝光时间为 20.4 ms,线性平台平移阶段速率为 1.31 mm · s⁻¹;该成像系统的波长范围为 392~1 009 nm,分辨率为 1.3 nm,总共 462 个波长,由于波长首尾部分噪音较大,故只选择 400~1 000 nm 之间的 449 个波长数据进行后续处理^[13]。



图 1 高光谱成像系统

Fig.1 Hyperspectral imaging system

1.2 高光谱反射率提取、预处理与样本划分

使用 ENVI 5.1 (Exelis Visual Information Solutions, USA) 对得到的马铃薯叶片高光谱图像进行反射率提取, 每个叶片选取 4 个分布在叶面两侧对称的感兴趣的区域 (region of interest, ROI), 每个 ROI 的大小 (长×宽) 为 50×50 像素 (pixels); 总共提取到 420 个原始光谱反射率数据。

在采集光谱原始数据的实际过程中, 由于外界环境以及试验仪器误差等因素的干扰, 造成获取的光谱数据掺杂一些无用信息造成随机噪声, 这些噪声会降低后期光谱数据处理与分析结果的准确性, 因此需要进行光谱预处理。本研究选用标准正态变量变换 (standard normal variate transformation, SNV)、多元散射校正 (multiple scattering correction, MSC)、一阶导数 (first derivative, FD)、二阶导数 (second derivative, SD) 4 种方法对光谱进行预处理并比较选择。将预处理得到的 420 个光谱数据样本使用随机划分 (random state) 算法按照 2:1 的比例分为 280 个训练集样本, 140 个测试集样本。

1.3 特征波长选择方法

为了消除光谱中的冗余信息, 本研究选用竞争性自适应重加权抽样法 (competitive adaptive re-weighted sampling, CARS) 与基于随机森林的交叉验证递归特征消除算法 (random forest-recursive feature elimination with cross-validation, RF-RFECV) 进行特征波长的选择; 去掉那些失去了光谱关键信息的波长, 以提高数据处理速度。

竞争性自适应重加权抽样法 (CARS) 是一种变量选择方法, 该算法试图模仿达尔文的进化理论“适者生存”。每一组光谱波长被视为一个个体, 通过自适应重加权采样技术识别 PLS 模型中具有较大绝对回归系数的波长点, 去除任何低加权波长点, 选择经过交叉验证和连续优化后回归模型中最准确的子集^[14]。

交叉验证递归特征消除算法 (recursive feature elimination with cross-validation, RFECV) 由递归特征消除 (RFE) 和交叉验证 (CV) 2 个部分组成, 递归特征消除原理是利用特征重要性从全集出发进行后向序列选择, 通过反复构建模型, 最终优选出对分类结果最合适的特征子集; 在进行递归特征消除的过程中, 本研究选用随机森林算法 (RF) 来进行选择效能的评估, 因为 RF 运行速度快, 不容易过拟合, 非常适合作为监督学习的估计器。同时为了提高模型选择的泛化能力, 我们利用三折交叉验证的方法来评价模型的性能^[15]。

1.4 叶片干旱状态判别模型的构建与评估

选用极端随机树 (extremely randomized trees, EXT) 来进行模型构建。EXT 是随机森林方法的一个变种, 通过集成多个决策树进行打分, 根据各个决策树预测值的平均值来进行投票^[16]。

分类模型的评估对于模型好坏的评判至关重要, 分类模型的结果会出现以下 4 种情况: 真正例 (true positive, TP), 即预测为正, 实际为正; 假正例 (false positive, FP), 即预测为负, 实际为正; 真反例 (true negative, TN), 即预测为负, 实际为负; 假反例 (false negative, FN), 即预测为负, 实际为正。

选用以下模型评估参数对构建的叶片干旱状态分类模型进行评估:

(1) 准确率 (Accuracy): 预测正确分类的样本个数 (TP+TN) 占总样本个数 (TP+FN+FP+TN) 的百分比:

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN}$$

(2) 召回率 (Recall): 预测正确的正例数据 (TP) 占实际为正例数据 (TP+FN) 的百分比:

$$Recall = \frac{TP}{TP+FN}$$

(3) F1 值 (F1-score): 召回率与精确率 (Precision) 的调和平均值。

$$Precision = \frac{TP}{TP+FP}$$

$$F1-score = 2 \times \frac{Precision \times Recall}{Precision + Recall}$$

1.5 模型可视化

高光谱图像包含样本的空间信息和光谱信息,图像上的每个像素对应 1 个反射率光谱,通过构建好的模型判断图像上每个像素点的干旱状态,使马铃薯叶片干旱状态判别可视化成为可能^[17]。具体步骤如下^[18]:

(1)将校正后的叶片光谱图像数据读入 Python 中,该图像在 Python 中显示为 $X \times Y \times H$ 的三维数组(X, Y 分别代表图像的长与宽; H 代表光谱波长数,本研究为 449 个)。

(2)将该三维数组重塑为 $Z \times H$ 的二维数组($Z = X \times Y$),把得到的二维数组进行 SNV 处理、特征波长选择后导入构建好的模型中即可获得每个像素点预测值的一维数组。然后将一维数组重塑为二维数组 $A(X \times Y)$ 。

(3)使用 ENVI 5.1 对第一步中校正好的叶片高光谱图像进行人工掩膜(MASK)处理,生成二进制掩膜图像 $B(X \times Y)$ 。把 B 读入 Python 中,该图像在 Python 中显示为 $X \times Y$ 的二维数组。

(4)将 A 与 B 进行数组相乘,即可消除背景影响,并进行伪色处理,最后得到马铃薯叶片干旱状态像素分布的二维可视化图。

本研究使用的数据处理方式:使用 ENVI 5.1 进行平均反射率的提取,采用 Python 3.9 进行特征波长的选择、模型构建与数据分析。

2 结果与分析

2.1 高光谱数据预处理结果分析

为了降低环境噪音,校正样品间因散射而引起的光谱误差,本研究对提取到的光谱反射率进行预处理,并通过全光谱波段的模型构建来进行预处理方式的选择。测试集预测结果如表 1 所示,结果显示除了 FD 算法外其他预处理算法都能够提高模型分类性能,尤其是 MSC 与 SNV 算法提升明显,通过 MSC 与 SNV 算法构建的 EXT 模型准确率分别为 85.71% 与 86.43%。整体比较而言,SNV 算法对模型性能的提升最为明显,因此本研究后续选用 SNV 算法对提取到的光谱数据进行预处理。

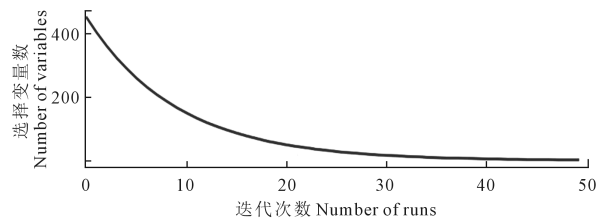
SNV 处理能够消除样本因表面散射、光照变化所产生的噪音。SNV 处理结果如图 2(见 250 页)所示,图 2a 为全部马铃薯叶片样本数据的原始反射率光谱图,图 2b 为 SNV 处理后的光谱图。预处理之后的光谱相比于原始光谱更加集中,增强了光谱数据中有用的光谱信息。

2.2 采用 CARS 与 RF-RFECV 算法筛选的光谱特征波长

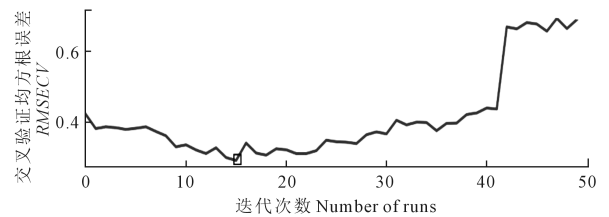
以原始光谱波长作为输入,CARS 算法的具体计算过程如图 3 所示。图 3a 显示,随着 CARS 迭代次数的增加,保留的波长数量逐渐减少,下降速度由快到慢,表明 CARS 算法在筛选特征波长的过程中经历“粗糙选择”和“精细选择”2 个阶段。图 3b 显示随着迭代次数的增加,交叉验证均方根误差 (root mean square error of cross validation, $RMSECV$) 先缓慢下降,说明与干旱状态等级分类的无关冗余波长在 CARS 进行迭代的过程中被剔除;而在迭代 15 次后, $RMSECV$ 开始上升,表明与干旱状态分类关键波长在 CARS 进行迭代的过程中被剔除。图 3b 的“□”代表此时的 $RMSECV$ 最小,依据 $RMSECV$ 最小的选择规则,最佳迭代次数为 15 次,最终选择的波长数目是 86 个,占总波长数的 19.15%。

表 1 不同预处理方式极端随机树模型性能比较/%
Table 1 Comparison of EXT model performance of different pretreatment methods

指标 Index	原始光谱 RAW	多元散 射校正 MSC	标准正态 变量变换 SNV	一阶导数 FD	二阶导数 SD
准确率 Accuracy	76.43	85.71	86.43	71.43	77.14
召回率 Recall	76.47	83.73	86.48	71.42	77.15
F1 值 F1-score	76.39	85.52	86.17	71.53	77.44



(a) CARS 算法在选择过程中波长数量的变化趋势
Change trend of the number of wavelengths in the selection process of CARS algorithm



(b) 随着迭代次数 $RMSECV$ 的变化趋势
Change trend with the number of iterations $RMSECV$

注:图(b)中“□”为对应 $RMSECV$ 最小时迭代次数的标注。

Note: In Figure (b), “□” indicates the minimum number of iterations corresponding to $RMSECV$.

图 3 CARS 算法特征波长选择过程

Fig.3 Characteristic wavelength selection process of CARS algorithm

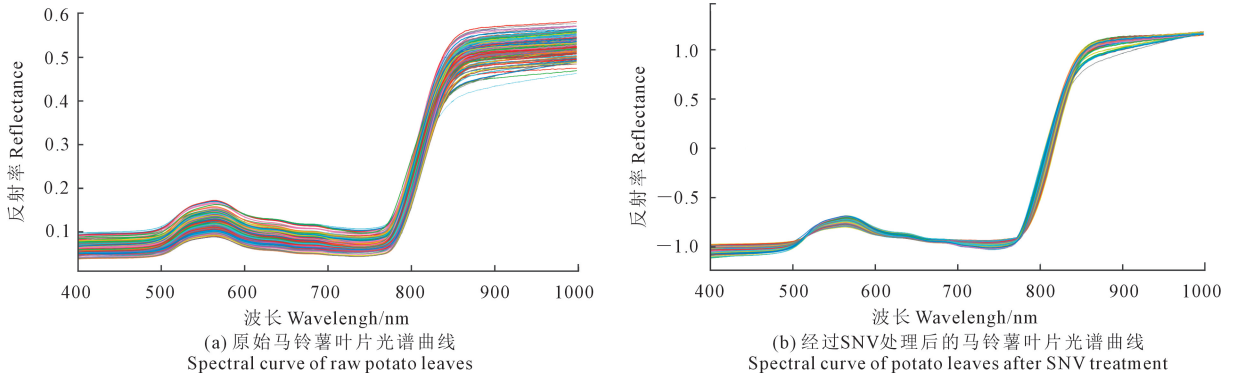
通过 RF-RFECV 算法对整体数据集(测试集+训练集)进行特征波长选择,设置每次递归要删除的特征数量为 1,交叉验证的折数为 3 折。由图 4a 可以看到交叉验证的得分(随机森林模型准确率)。随着选择波长的增加,交叉打折的各个组分得分变化趋势为先增加后在得分 0.85 之间震荡。当选择波长数目为 84 时,交叉验证的各组分平均得分最高达到峰值 0.8738(图 4b 中“□”标注);继续增加选择特征波长的数量并不会明显提高准确率,反而会增加模型负担。因此通过 RF-RFECV 法选择的特征波长数量为 84 个,占总波长数的 18.71%。

RF-RFECV 与 CARS 特征波长选择结果如图 5 所示,两种算法选择的波长数量相似,但是波长的分布范围明显不同。RF-RFECV 算法选择的波长主要在 600~650、680~700、900~950 nm 之间集中。前人研究发现 600~700 nm 之间的波段为幼年番茄

叶片干旱状态划分的主要特征波长,且该波段之间具有较强的相关性^[19],RF-RFECV 算法选择的波长分布与前人研究结果相符。CARS 算法选择的特征波长分布较为全面,特有选择的 550~600 nm,810~850 nm 波长与叶绿素、类胡萝卜素含量相关^[20]。马铃薯幼苗遭受干旱胁迫时,叶片中叶绿素和类胡萝卜素含量降低^[21],从而减少该波长范围内的反射率。反射率的数值大小为马铃薯叶片干旱状态的划分提供了可靠的数学依据,为了对两种特征波长选择方法进行评估,本研究后续通过构建模型的准确率对选择方法进行判断。

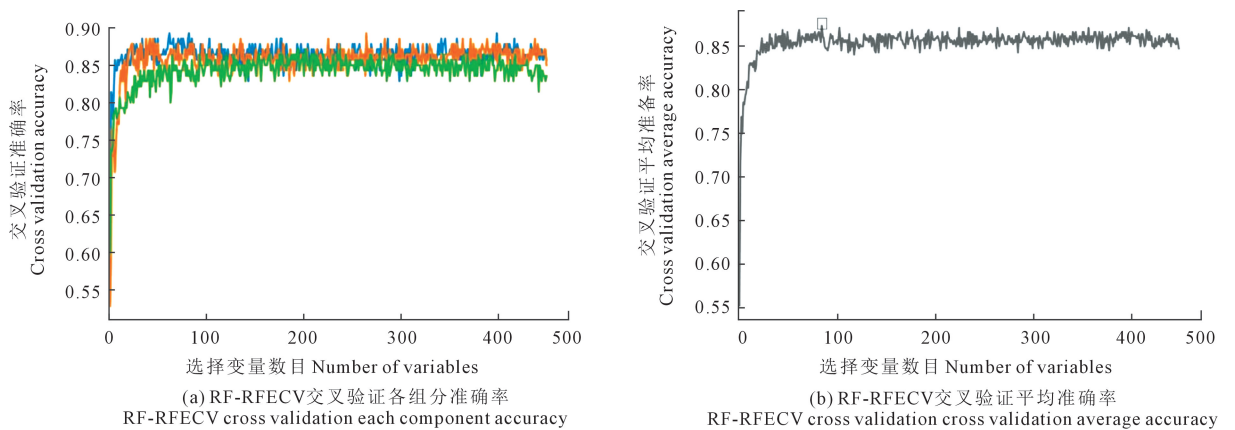
2.3 极端随机树模型建立与比较

以两种算法选择的特征波长与全波长作为叶片干旱状态划分的选择参数,将无干旱胁迫、中度干旱胁迫、重度干旱胁迫 3 种干旱级别分布赋予数值 0、1、2,建立极端随机树 EXT 模型。通过对模型影响



注:图中每个颜色线条代表一个光谱反射率曲线。
Note: Each color line in the figure represents a spectral reflectance curve.

图 2 马铃薯叶片原始与经过 SNV 处理后的光谱曲线
Fig.2 Spectral curves of potato leaves after raw and SNV treatment



注:图(a)中3种颜色线条代表交叉验证中被均分成3份的数据集;图(b)中“□”为最佳选择波长数目的标注。
Note: In Fig.(a), the three colored lines represent the data set that was divided into three parts in the cross-validation. In Fig.(b), “□” is the label for the optimal number of selected wavelengths.

图 4 RF-RFECV 算法特征波长选择过程
Fig.4 Characteristic wavelength selection process of RF-RFECV algorithm

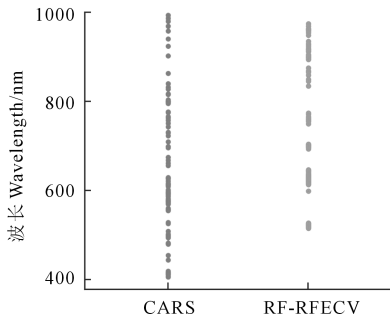


图 5 两种算法的选择波长结果

Fig.5 Selected wavelength results for both algorithms

表 2 极端随机树模型判别结果

Table 2 Discriminant results of extreme random tree model

分类模型 Classification model	决策树数量 Number of estimators	最大深度 Maximum depth	训练集指标 Training set/%			测试集指标 Testing set/%		
			准确率 Accuracy	召回率 Recall	F1 值 F1-score	准确率 Accuracy	召回率 Recall	F1 值 F1-score
RF-RFECV-EXT	113	19	97.50	97.50	97.49	92.14	92.17	92.11
CARS-EXT	22	17	94.79	94.78	94.80	88.57	88.62	88.46
Full-band-EXT	25	10	92.05	92.07	92.19	86.43	86.48	86.17

上升,准确率为 88.57%,*Recall* 为 88.62%,*F1-score* 为 88.46%,但是提升不太明显,可能是进行特征波长选择的时候遗漏了一些有用信息。研究结果说明通过 RF-RFECV 法选择的特征波长在干旱状态判别上更加敏感,划分类别接近的样本更加有效。

3 种模型在测试集上分类结果的混淆矩阵如图 6 所示,图 6 中 3 个混淆矩阵图从左上到右下的主对角线上为模型预测正确的数目(对应为深绿色区域),其余区域则为模型预测错误的数目(对应为浅色区域),图 6a 中基于 RF-RFECV 的 EXT 模型判错样本数为 11 个(图中浅色部分相加之和,下同),该模型能正确识别无干旱胁迫与重度干旱胁迫这 2 类样本,而对中度干旱胁迫样本的交叉错判比较明显,尤其是无干旱胁迫与中度干旱胁迫样本之间(图 6a)。基于 CARS 的 EXT 模型判错数目为 16 个,错误主要是把 7 个中度干旱胁迫样本识别成重度干旱胁迫样本与无干旱胁迫样本(图 6b)。基于全波段的 EXT 模型判错数目为 19 个,产生错误的主要原因是把 9 个中度干旱胁迫样本识别成重度干旱胁迫与无干旱胁迫样本,以及 5 个无干旱胁迫样本被识别成中度干旱胁迫样本(图 6c)。3 种模型错误率产生的主要原因是对中度干旱胁迫识别效果较差。而跨类(无干旱胁迫与重度干旱胁迫)预测错误较少。通过混淆矩阵结果可以看出,特征波长选择方法 RF-RFECV 对模型性能的改善有明显的作用,本研究后续的模型可视化选择表现较好的 SNV-RF-RFECV-EXT 模型进行。

较大的超参数 n estimators 即决策树数量,以及 max depth 即最大深度的测试选择了最佳的模型参数,分类结果如表 3 所示。由表 2 可以看到 3 种模型表现均优异,准确率都在 85% 以上,经过特征波长选择后构建模型的准确率等评价指标均较全波长有所提高;其中基于 RF-RFECV 建立的 EXT 模型的决策树数量为 113,最大深度为 19,测试集准确率为 92.14%,召回率为 92.17%,*F1-score* 为 92.11%,在所有模型中表现最佳。CARS-EXT 模型的决策树数量为 22,最大深度为 17,准确率等参数较全波长有所

2.4 马铃薯叶片干旱状态的可视化分析

利用训练后的 RF-RFECV-EXT 模型对收集到的不同干旱状态的马铃薯叶片高光谱图像进行像素分类,其中不同的颜色(绿色、黄色、红色)代表了模型预测的类别。如图 7 所示,绿色代表模型判断该像素点为无干旱胁迫,黄色代表模型判断该像素点为中度干旱胁迫,红色代表模型判断该像素点为重度干旱胁迫,图 7a 左侧为无干旱胁迫叶片,叶片整体发育正常,右侧为可视化结果图,能够看出颜色主体为代表无干旱胁迫的绿色。图 7b 左侧为中度干旱胁迫叶片,受干旱胁迫影响叶片开始卷曲,右侧可视化结果主体为代表中度干旱胁迫的黄色。图 7c 左侧为重度干旱胁迫叶片,叶片颜色枯黄且发生破裂,右侧可视化结果主体为代表重度干旱胁迫的红色。可视化结果显示预测为对应类别的颜色总是占大部分。上述结果表明通过可视化能够更加直观地显示叶片的干旱状态。

对马铃薯叶片进行干旱状态划分时,把叶片图像导入构建好的模型中,根据每个像素点的分类结果,统计模型预测的叶片上 3 种干旱状态的数量,通过不同干旱状态数量的比重对叶片进行整体评估,如果一个叶片中某个干旱状态像素点的数量明显高于其他干旱状态,就可初步判断该叶片所处的干旱状态。随机选择各个处理 5 片叶子统计 3 种像素分类百分比,与其对应的预测结果如表 3 所示,表 3 中数值为模型输入叶片每个像素点的预测结果占全部像素点的比值,可以看出正确分类的像素(位

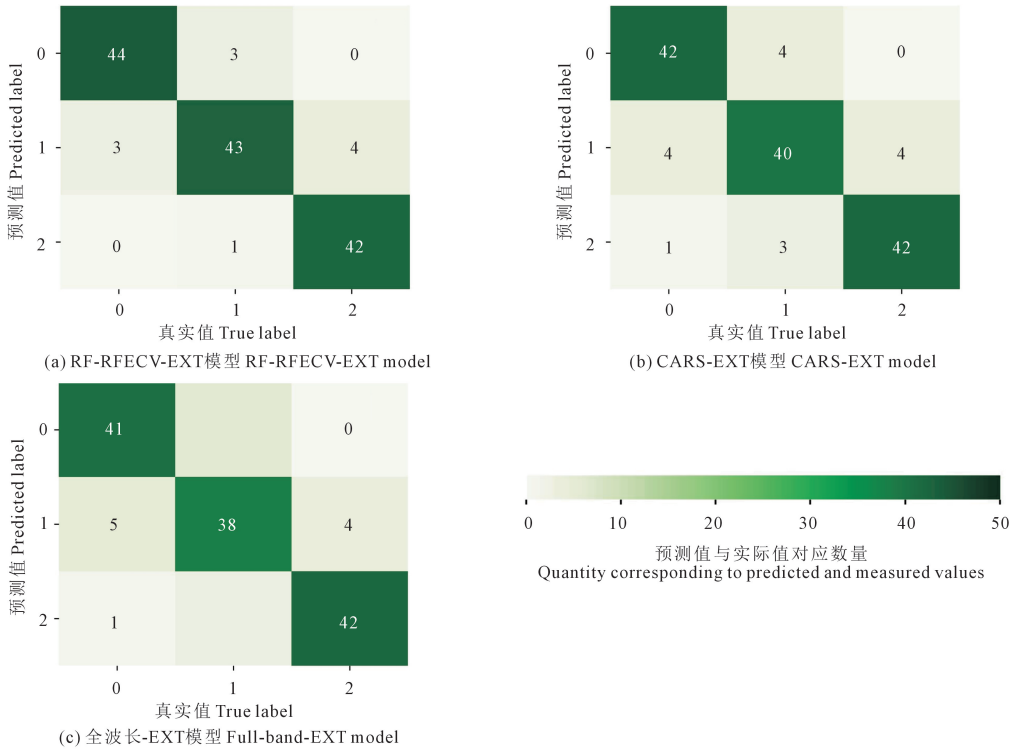
于表 3 的从左上到右下的主对角线上)的百分比总是显著高于那些错误分类的像素百分比,表明该划分方法在马铃薯叶片干旱状态的判别中表现优异。

3 讨论

本研究通过室内盆栽试验获取 3 种干旱状态的马铃薯叶片光谱图像,并获取了 3 种干旱状态马铃薯叶片的光谱反射率数据。运用标准正态变量变换(SNV)、多元散射校正(MSC)、一阶导数(FD)、二阶导数(SD)4 种光谱预处理方法对光谱数据进行预处理,去除数据中的噪声,通过竞争性自适应重加权抽样法(CARS)与基于随机森林的交叉验证递

归特征消除算法(RF-RFECV)筛选出最优波段组合,以此作为自变量,利用极端随机树算法(EXT)建立了马铃薯叶片干旱状态分类模型。最后,通过使用训练后的模型实现了马铃薯叶片干旱状态可视化,为马铃薯干旱胁迫状态的识别与灌溉管理提供了新方法。

相较于原始光谱,经过 SNV 算法处理后的模型性能提升最为明显,这与前人的番茄干旱状态分类研究结果相一致^[22]。以 RF-RFECV 算法得到的特征波长为自变量所构建的极端随机树分类模型效果最好,准确率为 92.14%,召回率为 92.17%,*F1-Score* 为 92.11%,相较于全光谱数据提升明显,表明 SNV-

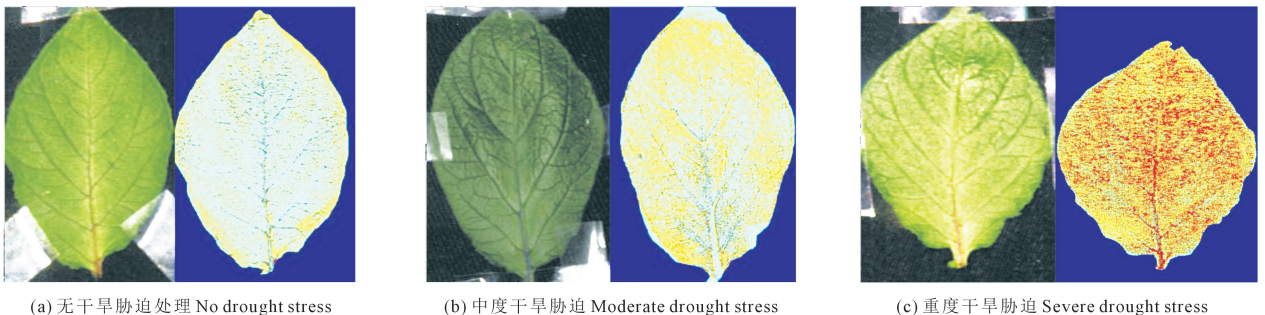


注:横坐标 0 代表无干旱胁迫,1 代表中度干旱胁迫,2 代表重度干旱胁迫。

Note: On the horizontal axis, 0 represents no drought stress, 1 represents moderate drought stress, 2 represents severe drought stress.

图 6 3 种模型在分类结果上的混淆矩阵图

Fig.6 Confusion matrix diagram of three models on the classification results



(a) 无干旱胁迫处理 No drought stress

(b) 中度干旱胁迫 Moderate drought stress

(c) 重度干旱胁迫 Severe drought stress

注:各图中左侧图均为原始叶片,右侧图均为其对应的可视化结果。

Note: The left images in each figure are the original blades, and the right images are their corresponding visualization results.

图 7 不同干旱处理单像素分类的可视化结果

Fig.7 Visual results of single pixel classification of each drought treatment

表 3 不同干旱处理输入模型后输出干旱类型占比/%

Table 3 Proportion of output drought types after input of different drought treatments into the model

模型输入干旱类型 Drought type of model input	模型输出干旱类型 Droughttype of model output		
	无干旱胁迫 No drought	中度干旱胁迫 Moderate drought	重度干旱胁迫 Severe drought
无干旱胁迫 No drought	63.42±11.5a	18.71±8.30b	17.88±7.35b
中度干旱胁迫 Moderate drought	19.19±8.92b	56.23±12.0a	24.59±9.03b
重度干旱胁迫 Severe drought	17.56±5.33b	14.61±7.04b	67.83±9.60a

注:同列不同小写字母表示差异显著($P<0.05$)。

Note: Different lowercase letters in the same column indicate significant differences ($P<0.05$).

RF-RFECV-EXT 模型可以作为马铃薯叶片干旱状态判别的最佳模型。根据建立的分类模型实现了马铃薯叶片干旱状态的可视化,结合图像处理能直观地显示马铃薯叶片的干旱状态,为快速、直观监测马铃薯叶片干旱状态提供了科学参考。

模型的可视化能够为快速、精准的干旱胁迫检测提供依据,如 Navarro 等^[23]对盆栽野生芝麻菜设置 4 种胁迫处理(镰刀菌、丝核菌、干旱胁迫、盐胁迫)并获取光谱图像,通过建模进行胁迫类别分类,并根据可视化结果分析出在胁迫处理 7 d 后就可以鉴别出胁迫类型。Dao 等^[24]通过 HSI 与深度神经网络相结合,使用不同干旱胁迫程度下的无芒雀麦光谱信息进行一阶导数处理与模型构建,实现了干旱胁迫的可视化准确检测。前人构建的模型融合了特征波段与纹理特征,与本研究所建立的马铃薯叶片干旱胁迫分类模型有所差异,而本研究的可视化模型也取得了很好的效果,为目前研究甚少的马铃薯叶片干旱状态可视化分级提供了新的方法和研究思路。

4 结 论

基于 3 种干旱状态的马铃薯叶片样品室内高光谱图像数据、4 种光谱预处理和 2 种特征波长选择算法建立了马铃薯叶片干旱状态分类模型,比较了不同光谱数据预处理和不同特征波长选择方法下模型精度的差异。其中 SNV 算法与 RF-RFECV 特征波长选择算法联合使用能明显提高模型性能,通过两者建立的 SNV-RF-RFECV-EXT 模型是判别马铃薯叶片干旱状态的最优模型,同时通过该模型的可视化能直观地显示马铃薯叶片的干旱状态。

参 考 文 献:

[1] 庞泽,田国奎,王海艳,等.我国马铃薯产业发展现状及展望[J].中国瓜菜,2023,36(7):148-154.
PANG Z, TIAN G K, WANG H Y, et al. Present situation and prospect

of potato industry in China[J]. China Cucurbits and Vegetables, 2023, 36(7): 148-154.

[2] 唐建昭,王仁德,周海涛,等.农牧交错带灌溉和旱作模式下马铃薯耗水差异[J].干旱地区农业研究,2023,41(4):118-125,158.
TANG J Z, WANG R D, ZHOU H T, et al. Difference of water utilization between irrigated and rainfed potato in the agro-pastoral ecotone of northern China[J]. Agricultural Research in the Arid Areas, 2023, 41(4): 118-125, 158.

[3] 胡萌萌,张继宗,张立峰,等.水分胁迫及复水对马铃薯生长发育及产量的影响[J].干旱地区农业研究,2021,39(2):95-101,121.
HU M M, ZHANG J Z, ZHANG L F, et al. Effects of water stress and rewatering on potato growth and yield[J]. Agricultural Research in the Arid Areas, 2021, 39(2): 95-101, 121.

[4] MA J, ZHENG B, HE Y. Applications of a hyperspectral imaging system used to estimate wheat grain protein:a review[J]. Frontiers in Plant Science, 2022, 13: 837200.

[5] SARIC R, NGUYEN V D, BURGE T, et al. Applications of hyperspectral imaging in plant phenotyping[J]. Trends in Plant Science, 2022, 27(3): 301-315.

[6] CHEN S Z, GAO Y, FAN K, et al. Prediction of drought-induced components and evaluation of drought damage of tea plants based on hyperspectral imaging[J]. Frontiers in Plant Science, 2021, 12: 695102.

[7] MISHRA P, FELLER T, SCHMUCK M, et al. Early detection of drought stress in *Arabidopsis thaliana* utilising a portable hyperspectral imaging setup[C]//2019 10th Workshop on hyperspectral imaging and signal processing: evolution in remote sensing (WHISPERS): IEEE, 2019: 1-5.

[8] WEKSLER S, ROZENSTEIN O, BEN DOR E. Continuous seasonal monitoring of nitrogen and water content in lettuce using a dual phenomics system[J]. Journal of Experimental Botany, 2022, 73(15): 5294-5305.

[9] ZHOU J J, ZHANG Y H, HAN Z M, et al. Evaluating the performance of hyperspectral leaf reflectance to detect water stress and estimation of photosynthetic capacities[J]. Remote Sensing, 2021, 13(11): 2160.

[10] 孙红,刘宁,吴莉,等.高光谱成像的马铃薯叶片含水率分布可视化[J].光谱学与光谱分析,2019,39(3):910-916.
SUN H, LIU N, WU L, et al. Visualization of water content distribution in potato leaves based on hyperspectral image[J]. Spectroscopy and Spectral Analysis, 2019, 39(3): 910-916.

[11] 江晓东,李增嘉,侯连涛,等.少免耕对灌溉农田冬小麦/夏玉米作物水、肥利用的影响[J].农业工程学报,2005,21(7):20-24.
JIANG X D, LI Z J, HOU L T, et al. Impacts of minimum tillage and no-tillage systems on soil NO_3^- -N content and water use efficiency of winter wheat/summer corn cultivation[J]. Transactions of the Chinese Society of Agricultural Engineering, 2005, 21(7): 20-24.

[12] 郭华春,沙本才,余杨.赤霉素打破种薯休眠对马铃薯生长及产量的影响[J].中国马铃薯,2003,17(6):336-337.
GUO H C, SHA B C, YU Y. Growth and yield of potato plants as influenced by gibberellin treated seed tubers[J]. Chinese Potato Journal, 2003, 17(6): 336-337.

[13] ZHU H L, WANG M Y, ZHANG J, et al. Prediction of apple hybrid offspring aroma based on hyperspectral[J]. Foods, 2022, 11(23): 3890.

[14] FENG Z H, WANF L Y, YANG Z Q, et al. Hyperspectral monitoring of powdery mildew disease severity in wheat based on machine learning[J]. Frontiers in Plant Science, 2022,13: 828454.

[15] 高宏元,侯蒙京,葛静,等.基于随机森林的高寒草地上生物量高光谱估算[J].草地学报,2021,29(8):1757-1768.
GAO H Y, HOU M J, GE J, et al. Hyperspectral estimation of aboveground biomass of alpine grassland based on random forest algorithm[J]. Acta Agrestia Sinica, 2021, 29(8): 1757-1768.

- [16] 程渊, 李玉霞, 李凡, 等. 基于极端随机树的闪电河流域土壤水分反演[J]. 遥感学报, 2021, 25(4): 941-951.
CHENG Y, LI Y X, LI F, et al. Soil moisture retrieval using extremely randomized trees over the Shandian river basin[J]. National Remote Sensing Bulletin, 2021, 25(4): 941-951.
- [17] YUAN Z R, YE Y, WEI L F, et al. Study on the optimization of hyperspectral characteristic bands combined with monitoring and visualization of pepper leaf SPAD value[J]. Sensors, 2021, 22(1): 183.
- [18] ZHANG L, AN D, WEI Y G, et al. Prediction of oil content in single maize kernel based on hyperspectral imaging and attention convolution neural network[J]. Food Chemistry, 2022, 395: 133563.
- [19] WENG S Z, MA J J, TAO W T, et al. Drought stress identification of tomato plant using multi-features of hyperspectral imaging and subsample fusion[J]. Frontiers in Plant Science, 2023, 14: 1073530.
- [20] 曹晓峰, 任惠如, 李幸芝, 等. 高光谱技术结合特征波长/光谱指数对冬枣成熟度可视化判别[J]. 光谱学与光谱分析, 2018, 38(7): 2175-2182.
CAO X F, REN H R, LI X Z, et al. Discrimination of winter jujube's maturity using hyperspectral technique combined with characteristic wavelength and spectral indices [J]. Spectroscopy and Spectral Analysis, 2018, 38(7): 2175-2182.
- [21] 杨宏伟, 张鑫, 张永彬, 等. 喷施硫酸钾对干旱胁迫下马铃薯生长发育及抗性生理的影响[J]. 华北农学报, 2015, 30(增刊1): 471-477.
YANG H W, ZHANG X, ZHANG Y B, et al. Effects of spraying potassium sulfate on the growth and development and resistance physiology of potato under drought stress[J]. Acta agriculturae boreali-sinica, 2015, 30(S1): 471-477.
- [22] 贺露, 万莉, 高会议. 基于高光谱成像技术识别番茄干旱胁迫[J]. 光谱学与光谱分析, 2023, 43(3): 724-730.
HE L, WAN L, GAO H Y. Recognition of drought stress in tomato based on hyperspectral imaging[J]. Spectroscopy and Spectral Analysis, 2023, 43(3): 724-730.
- [23] NAVARRO A, NICASTRO N, COSTA C, et al. Sorting biotic and abiotic stresses on wild rocket by leaf-image hyperspectral data mining with an artificial intelligence model[J]. Plant Methods, 2022, 18(1): 45.
- [24] DAO P D, HE Y H, PROCTOR C. Plant drought impact detection using ultra-high spatial resolution hyperspectral images and machine learning[J]. International Journal of Applied Earth Observation and Geoinformation, 2021, 102: 102364.

(上接第 234 页)

- [6] 宋恒飞, 辛良杰. 中国耕地利用强度分化特征及影响因素研究[J]. 农业工程学报, 2021, 37(16): 212-222.
SONG H F, XIN L J. Differentiation characteristics and influencing factors of cultivated land use intensity in China[J]. Transactions of the Chinese Society of Agricultural Engineering, 2021, 37(16): 212-222.
- [7] 李建平, 上官周平. 陕西省耕地生产能力的时空分布[J]. 农业工程学报, 2012, 28(10): 239-246.
LI J P, SHANG GUAN Z P. Spatial-temporal distribution of cultivated land production capacity in Shaanxi province [J]. Transactions of the Chinese Society of Agricultural Engineering, 2012, 28(10): 239-246.
- [8] 宋小青, 欧阳竹, 柏林川. 中国耕地资源开发强度及其演化阶段[J]. 地理科学, 2013, 33(2): 135-142.
SONG X Q, OU YANG Z, BAI L C. Evaluation and evolution of exploitative intensity of cultivated land resources in China[J]. Acta Geographica Sinica, 2013, 33(2): 135-142.
- [9] 牛家恒, 江丽, 陈鑫, 等. 耕地利用强度评价方法的县域休耕适用性——以曲周县为例[J]. 中国农业大学学报, 2019, 24(7): 156-166.
NIU J H, JIANG L, CHEN X, et al. Fallow applicability of cropland use intensity assessment methods in county level: A case study of Quzhou County[J]. Journal of China Agricultural University, 2019, 24(7): 156-166.
- [10] 李佳玲, 侯兰金, 朱吉贝, 等. 近 20 年湖南省耕地开发强度时空分异研究[J]. 中国农业资源与区划, 2021, 42(11): 199-207.
LI J L, HOU L J, ZHU J B, et al. Study on the spatial and temporal variation of cultivated land development intensity in Hunan province in the past 20 years[J]. Chinese Journal of Agricultural Resources and Regional Planning, 2021, 42(11): 199-207.
- [11] 王向东, 刘小茜, 裴韬, 等. 基于技术效率测算的京津冀城市土地集约利用潜力评价[J]. 地理学报, 2019, 74(9): 1853-1865.
WANG X D, LIU X Q, PEI T, et al. Potential evaluation of urban land intensive use in Beijing-Tianjin-Hebei region based on measurement of technical efficiency [J]. Acta Geographica Sinica, 2019, 74(9): 1853-1865.
- [12] 周晓林, 吴次芳, 刘婷婷. 基于 DEA 的区域农地生产效率差异研究[J]. 中国土地科学, 2009, 23(3): 60-65.
ZHOU X L, WU C F, LIU T T. Study on productive efficiency difference of regional farmland based on DEA [J]. China Land Science, 2009, 23(3): 60-65.
- [13] 覃成林, 杨威. 中国农村居民收入区域不平衡的动态变化及影响因素——基于人口加权变异系数的分析[J]. 产经评论, 2012, 3(4): 115-124.
QIN C L, YANG W. The dynamic changes of regional income inequality and its determinants in rural China --Based on the analysis of the coefficient of population - weighted variation [J]. Forward Position in Economics, 2012, 3(4): 115-124.
- [14] 王劲峰, 徐成东. 地理探测器: 原理与展望[J]. 地理学报, 2017, 72(1): 116-134.
WANG J F, XU C D. Geodetector: principle and prospective[J]. Acta Geographica Sinica, 2017, 72(1): 116-134.
- [15] 佟庆远, 孙傅, 董欣, 等. 污水处理厂减排效率的统计评价及影响因素分析[J]. 中国人口·资源与环境, 2019, 29(4): 49-57.
TONG Q Y, SUN F, DONG X, et al. Statistical assessment and determinant analysis of the pollutant removal efficiencies of wastewater treatment plants[J]. China Population, Resources and Environment, 2019, 29(4): 49-57.
- [16] 陕西省统计局, 国家统计局陕西调查总队. 陕西统计年鉴 2020[J]. 北京: 中国统计出版社, 2020.
TJJ SHAANXI GOV CN, Survey Office of the National Bureau of Statistics in Shaanxi. Shaanxi statistical yearbook 2020 [J]. Beijing: China Statistics Press, 2020.
- [17] 王小英, 陈占飞, 胡凡, 等. 陕西省农田化肥投入过量与不足的研究[J]. 干旱地区农业研究, 2017, 35(6): 159-165.
WANG X Y, CHEN Z F, HU F, et al. Study on the excessive and insufficient of chemical fertilizer inputs on farmland in Shaanxi Province [J]. Agricultural Research in the Arid Areas, 2017, 35(6): 159-165.