

张芳, 邓照龙, 田有文, 等. 基于高光谱成像技术的南果梨酸度无损检测方法[J]. 沈阳农业大学学报, 2024, 55(2): 231-239.

ZHANG Fang, DENG Zhaolong, TIAN Youwen, et al. Non-destructive testing method for acidity of nanguo pear based on hyperspectral imaging technology[J]. Journal of Shenyang Agricultural University, 2024, 55(2): 231-239.

基于高光谱成像技术的南果梨酸度无损检测方法

张芳¹, 邓照龙¹, 田有文¹, 高鑫¹, 王开田¹, 徐正玉²

(1. 沈阳农业大学信息与电气工程学院, 沈阳 110161; 2. 合肥市财政局, 合肥 230031)

摘要: 南果梨是一种重要的水果品种, 其酸度是评估果品质量的重要指标之一。然而, 传统的南果梨酸度检测方法通常需要破坏性采样和化学分析, 不仅耗时费力, 而且容易导致样品污染和浪费。因此, 旨在探索一种基于高光谱成像技术的无损检测方法, 以实现对南果梨酸度的快速、准确、无损检测。首先, 采集室温 20 °C 下不同贮藏天数南果梨的高光谱数据, 其光谱波长范围为 400~1 000 nm, 并且通过理化实验测量南果梨样本的可滴定酸; 其次, 采用多元散射校正 (multiplicative scatter correction, MSC)、标准正态变换 (standard normal variate, SNV)、Savitzky-Golay 平滑滤波等多种方法对光谱数据进行预处理, 建立偏最小二乘回归模型 (partial least squares regression, PLSR), 选择出建模效果最佳的预处理方法, 结果显示 MSC 方法效果最优; 然后结合连续投影算法 (successive projection algorithm, SPA) 提取特征波段, 在 700~900 nm 范围内确定 9 个特征光谱变量; 最后, 以提取出的 9 个特征光谱变量作为输入矢量, 分别建立 PLSR 模型、极限学习机 (extreme learning machine, ELM) 模型以及遗传算法 (genetic algorithm, GA) 和粒子群算法 (particle swarm optimization, PSO) 优化的 BP 神经网络模型。研究结果表明, 基于 MSC 预处理和 SPA 算法特征提取的 PSO-BP 模型预测精度最高, 效果最好, 预测集决定系数 $R_p^2=0.911$, RMSEP=0.032。可见, 基于高光谱成像技术的 SPA-PSO-BP 模型可用于南果梨酸度的检测, 为南果梨的品质评价提供参考。

关键词: 高光谱成像技术; 南果梨; 酸度; BP 神经网络; PSO-BP 模型

中图分类号: TS255.3

文章编号: 1000-1700(2024)02-0231-09

文献标识码: A

开放科学 (资源服务) 标识码 (OSID):



Non-destructive Testing Method for Acidity of Nanguo Pear Based on Hyperspectral Imaging Technology

ZHANG Fang¹, DENG Zhaolong¹, TIAN Youwen¹, GAO Xin¹, WANG Kaitian¹, XU Zhengyu²

(1. College of Information and Electrical Engineering, Shenyang Agricultural University, Shenyang 110161, China;

2. Hefei Treasury, Hefei 230031, China)

Abstract: Nanguo pear is an important fruit variety, and its acidity is one of the important indicators for evaluating fruit quality. However, traditional methods for detecting acidity in Nanguo pear often require destructive sampling and chemical analysis, which is not only time-consuming and laborious, but also prone to sample contamination and waste. Therefore, a non-destructive testing method based on hyperspectral imaging technology was explored to achieve rapid, accurate, and non-destructive detection of acidity in Nanguo pear. Firstly, the hyperspectral data of Nanguo pear stored for different days at room temperature of 20 °C was collected, the wavelength range is 400-1 000 nm, and the titratable acid of Nanguo pear samples was measured through physical and chemical experiments; secondly, multiple methods such as multiple scatter

收稿日期: 2024-01-10

基金项目: 辽宁省教育厅基础研究项目 (JYTMS20231285)

第一作者: 张芳 (1976-), 女, 博士, 教授, 从事农产品品质检测研究, E-mail: lsjzflbh@163.com

correction (MSC), standard normal variation (SNV), Savitzky Golay smoothing filtering were used to preprocess spectral data. A partial least squares regression (PLSR) model was established, and the best preprocessing method was selected. The results showed that the MSC method had the best performance; then, combined with the continuous projection algorithm (SPA), feature bands are extracted, and 9 feature spectral variables are determined in the range of 700–900 nm; finally, using the extracted 9 feature spectral variables as input vectors, a PLSR model, an extreme learning machine (ELM) model, and a BP neural network model optimized by genetic algorithm (GA) and particle swarm optimization (PSO) were established respectively. The research results indicate that the PSO–BP model based on MSC preprocessing and SPA algorithm feature extraction has the highest prediction accuracy and the best performance, with a prediction set determination coefficient $R_p^2=0.911$ and RMSEP=0.032. It can be seen that the SPA–PSO–BP model based on hyperspectral imaging technology can be used for the detection of acidity in Nanguo pear, providing reference for the quality evaluation of Nanguo pear.

Key words: hyperspectral imaging technology; Nanguo pear; acidity; BP neural network; PSO–BP model

南果梨是辽宁省鞍山市的特产水果,以其果肉细腻、酸甜可口、风味香浓而深受赞誉,虽然当地种植面积较有规模,但采摘后产业化水平和商品化处理技术仍不成熟。目前对其糖度、酸度等内部品质主要采用传统的人工检测,不仅耗时耗力、操作繁琐,而且损坏检测样本造成了极大浪费。因此,在实际生产中,寻找到能够实现南果梨品质快速、无损检测的方法具有重要意义。

高光谱成像是目前在水果研究中最常用的无损检测技术之一,它是通过结合图像和光谱技术来实现检测的方法。图像技术主要用于检测样品的外部特性,而光谱技术则主要用于检测样品的内部特性。然而,高光谱成像能够真正实现对水果内外部质量特性的同时检测。目前,已经有很多学者将高光谱成像用于水果内部品质的研究。例如,孟庆龙等^[1]利用高光谱系统采集了贮藏期内的“贵长”猕猴桃样本的图像和光谱数据,通过标准正态变换技术,有效消除了光谱数据中的噪声干扰;随后采用主成分分析法,从这些特征变量中筛选出关键信息,最终成功构建出一个用于预测猕猴桃酸度的回归模型,预测集相关系数 r_p 为0.90,具有非常好的预测性能。卢娜等^[2]对草莓硬度进行了预测研究,利用高光谱系统获取草莓光谱数据,采用多种预处理方法分别建立偏最小二乘(partial least squares, PLS)定量预测模型,结果表明经过标准正态变换预处理方法的PLS模型效果最好,进而证明高光谱成像技术对草莓硬度的检测方法可行。TAN等^[3]提出了一种基于多光谱融合的方法检测樱桃番茄的可溶性固形物(soluble solids content, SSC)和可滴定酸(titratable acidity, TA)值,通过对比单个光谱信息和融合光谱信息在检测樱桃番茄时对其SSC和TA值的影响,结合偏最小二乘回归、支持向量回归、反向传播神经网络等多种机器学习方法进行建模,结果表明与单一光谱相比,多光谱融合方法对检测樱桃番茄内部品质有更好的预测效果。KALOPESA等^[4]提出了一种多输入输出卷积神经网络来检测葡萄糖的糖度和酸度值,通过与偏最小二乘回归、随机森林回归、支持向量机等算法对比,结果表明,其在估计糖度和酸度方面具有很高的准确性。蒋蕻等^[5]以油茶籽为研究对象,找到一种油茶籽脂肪酸成分含量的最佳预测模型,首先利用高光谱获取反射光谱图像,对图像进行平滑与多元散射校正,通过相关性分析和逐步回归分析得到最佳优化波段,然后对其采用偏最小二乘回归、主成分回归、径向基神经网络3种建模方法建模,经对比得出径向基神经网络建立的预测模型最好。ÇETIN等^[6]对苹果硬度和SSC进行预测,利用高光谱技术采集苹果三个收获阶段的光谱数据,对比人工神经网络、决策树、偏最小二乘回归和多元线性回归的预测性能,研究表明,人工神经网络和决策树方法对硬度更有效,而决策树和多元线性回归对SSC更有效,这些方法在实际应用中被证明是非常可行的。WANG等^[7]利用杨梅敏感波段光谱和光谱指标,采用简单统计回归模型建立了杨梅各理化参数的光谱模型,并对模型进行验证;最后,建立了杨梅糖酸度模型,该模型具有较好的准确性和稳定性;同时在此模型的基础上,研制了一种低成本的便携式糖酸检测器。常洪娟等^[8]利用芒果的近红外高光谱及SSC数据建立反向传播神经网络模型与偏最小二乘回归模型来对芒果的可溶性固形物进行预测,预测集决定系数达到0.864 1。尽管高光谱成像技术在水果品质检测方面已有诸多应用,但关于南果梨酸度的研究还未见报道。

本研究以南果梨为研究对象,使用高光谱成像技术获取其酸度数据,并对特征波段进行筛选,提取其品质指标的相关特征波段,并进一步建立南果梨酸度的预测模型,获得了良好的预测效果,为南果梨品质无损检测提供了重要参考。

1 材料与方法

1.1 材料

试验用品选用来自辽宁鞍山海城王石镇果园大小、形状基本一致的南果梨120个,将它们分别编号后置于实验室内(室温20℃)备用。为了样品的酸度值有明显变化,将样品按照10个一组分为12组,在贮藏0(置于实验室当天),1,2,3,4,5,6,7,8,9,10,11 d时,分别在室温情况下进行高光谱数据采集和酸度数据测定。



图1 南果梨样品

Figure 1 Sample of Nanguo pear

1.2 数据获取

1.2.1 高光谱图像采集 高光谱反射成像系统的硬件包括光谱仪、CCD相机、位移台、光源、计算机和软件控制系统等。在本研究中使用的高光谱成像系统由ImSpector V10E光谱仪、IGV-B1410M CCD相机(像素为392×1040)、IRCP0076-1 COM位移控制平台、2个150 W的光纤卤素灯(3900 Illuminator, Illumination Technologies, Inc. USA)、1个暗箱(尺寸为120 cm×50 cm×140 cm)以及1台戴尔Vostro 5560D-1528数据处理机组成。该成像系统通过PC机进行控制,并利用HSI Analyzer高光谱图像采集软件来采集信息。成像光谱仪光谱范围为400~1 000 nm,光谱率为1.27 nm。开机后首先进行预热30 min^[9],预热后进行焦距调整,同时调试平台的移动速度,以确保图像的清晰度。经多次调试,本试验设置光源强度为2 cd,平台移动速度1.40 mm·s⁻¹。

为了避免由于光照不均、周围环境以及仪器的暗电流对光谱数据的噪声影响,需要在采集样本高光谱图像前,分别利用标准白板和镜头盖采集白板信息(I_w)和黑板信息(I_b)后按照式(1)对采集的样本原始高光谱图像(I_0)计算得到校正后的图像信息(I)^[10]。

$$I = \frac{I_0 - I_b}{I_w - I_b} \quad (1)$$

式中: I 是校正图像信息; I_0 是原始的高光谱图像; I_w 是白板信息; I_b 是黑板信息。

1.2.2 可滴定酸测定 首先使用电子天平称取20 g(精确至10 mg)南果梨样品,在其中加入少量无CO₂蒸馏水,研钵研成糊状并用无CO₂蒸馏水清洗,装入移液管放入250 mL容量瓶中,用蒸馏水定容且充分摇匀,然后使用干燥滤纸和漏斗进行过滤备用。然后100 mL三角瓶中配制0.1 mol·L⁻¹氢氧化钠标准溶液50 mL备用,在250 mL烧杯中配制酚酞指示剂100 mL。吸取滤液 $V_s=50$ mL,加入酚酞指示剂2~3滴,用氢氧化钠标准溶液0.1 mol·L⁻¹和碱式滴定管滴定,滴定至微红色,1 min内不褪色即为终点,最后记录消耗的标准碱液体积(V_1),同一被测样品测定3次,最后通过式(2)计算南果梨中可滴定酸含量。

$$\text{可滴定酸含量} = \frac{V \times c \times (V_1 - V_n) \times f}{V_s \times m} \times 100\% \quad (2)$$

南果梨中主要的有机酸为苹果酸,折算系数 f 为0.067^[11]。本试验采用可滴定酸(titratable acidity, TA)含量表示南果梨的酸度。

1.3 光谱数据预处理与特征波段筛选

首先利用ENVI5.3软件对南果梨高光谱影像中的感兴趣区域(region of interesting, ROI)进行分割,选取30×30大小的像素点,获取各像素点的平均光谱信息,并利用Pycharm2021.1.1对其进行二次分析。在此基础上,利用多元散射修正(MSC)、标准正态变换(SNV)、二阶差分、小波变换(Wave)等多种方法,对所采集到的原始光谱进行预处理。全光谱波段数据含有大量无关信息和冗余信息,与目标参数相关性不高,需要进一步筛选特征波段以提高信噪比,减少计算量^[12]。

采用提取特征波长的方法进行建模,此方法可以有效解决在建模过程中波长数目不足,建模效率

低、模型繁琐复杂等问题,同时提高模型的稳定性。此外,在特征波长建模方法中连续投影算法最适合本模型的样本空间,故选用该方法进行后续研究。该算法是一种前向循环选择算法,主要用来消除光谱变量之间共线性的影响,SPA算法是在所有光谱中循环投影,是最少冗余信息变量的组合,以此来达到变量之间共线性最低,优化模型的精确度和效率^[13]。其中SPA算法的主要作用就是可以避免信息冗余的同时提取,降低了模型的误差^[14]。

1.4 定量建模方法

定量建模是将检测样本中得到的信息进行量化并建立数学模型的过程,该建模方法可分为线性方法(MLR模型和PLSR模型等)和非线性方法(BP神经网络模型和ELM模型等)。本研究中主要使用PLSR模型、极限学习机(extreme learning machines, ELM)模型、BP神经网络模型,以及利用遗传算法、粒子群算法来优化BP神经网络模型。

PLSR(partial least squares regression, PLSR)算法是一种基于偏最小二乘回归的统计建模方法。它是在偏最小二乘算法(partial least squares, PLS)基础上发展而来,主要用于处理回归问题。PLSR算法能够解决多重共线性和高维数据等挑战,同时能够保持对目标变量的最大相关性^[15]。极限学习机是一种基于神经网络的快速训练算法,用于解决分类和回归问题。与传统的神经网络相比,ELM具有更快的训练速度和较好的泛化能力^[16]。BP神经网络(backpropagation neural network, BNN)是一种常用的人工神经网络模型,常用来解决分类、回归和模式识别等问题^[17]。

BP神经网络虽然表达能力强,适应性好,但是存在着易陷入局部最优,训练速度慢等问题。本研究提出采用遗传算法和粒子群优化算法来优化BP神经网络模型的输入权重和偏置^[18],目的是提升BP神经网络模型在预测精度和性能方面的表现。

粒子群优化算法(particle swarm optimization, PSO)是一种群体智能优化算法,基于模拟鸟群或鱼群等群体行为而提出的,它通过模拟群体内个体之间的协作和竞争,来寻找问题的全局最优解^[19]。遗传算法(genetic algorithm, GA)是一种模拟自然进化过程的优化算法,它通过模拟生物进化中的遗传、交叉和变异等操作,以逐步搜索和优化问题的解空间^[20]。

粒子群优化算法优化BP神经网络模型流程为:(1)初始化粒子群:随机生成一组粒子,每个粒子表示一个解(BP神经网络的参数),包括权重和偏置。(2)初始化速度和最佳位置:为每个粒子随机初始化速度,并将当前位置作为最佳位置。(3)确定适应度函数:将BP神经网络的训练误差作为适应度函数,即粒子群的目标函数。(4)更新粒子速度和位置:根据一定的更新规则,更新每个粒子的速度和位置。速度的更新受到个体历史最佳位置和群体历史最佳位置的影响。(5)更新最佳位置:根据适应度函数的值,更新每个粒子的历史最佳位置。(6)终止条件确定:基于预先确定的终止条件(如果达到了最大迭代次数,或者达到了某种误差要求)确定算法是否被终止。(7)重复步骤4至步骤6,直到达到终止条件。

在PSO算法中,粒子通过搜索空间来寻找最优解,其速度和位置的更新受到个体历史最佳位置和群体历史最佳位置的引导^[21]。通过不断迭代更新,粒子群逐渐趋向于全局最优解,即BP神经网络模型的最优参数。

预测模型建成后,通过计算训练集的决定系数(decision coefficient of calibration set, R_c^2)和均方根误差(root mean square error of calibration set, RMSEC)、测试集的决定系数(decision coefficient of prediction set, R_p^2)和均方根误差(root mean square error of prediction set, RMSEP)4个指标来评估模型的预测性能。具体来说, R_c^2 和 R_p^2 越接近1,表明模型在训练集和测试集上的拟合度越好;RMSEC和RMSEP越小,表明模型的预测误差越低^[22]。

2 结果与分析

2.1 南果梨可滴定酸分析

TA的变化与南果梨的成熟度有一定关系,图2为南果梨在贮存过程中TA的变化情况。由图2可知,由于采后的呼吸作用等生理过程,南果梨果内的有机酸会继续降解并转化为可溶性固形物^[23],因此,南果梨会随储存时间的增加而逐步变得越来越成熟,而TA的含量却下降。

2.2 光谱处理

在原始高光谱中,除了样品的品质信息外,还可能包含有仪器产生的电噪声、样本背景以及环境中的杂散光等无关信息。为了确保数据的准确性和可靠性,本实验需要对这些无关信息进行预处理,以消除或减少其对结果的影响^[24]。首先对高光谱首端的大噪声区域进行人工剔除,并保留 500~1 000 nm 的区域光谱为研究光谱,并分别使用二阶差分、多元散射校正(MSC)和小波变换(Wave)等不同预处理方法,分别建立了 PLSR 模型,通过对比和分析发现 MSC 对样品的预处理方式是最优的,因此以 MSC 对样品进行处理后得到的光谱为研究光谱,其中样品的目标区初始光谱和 MSC 预处理光谱如图 3。

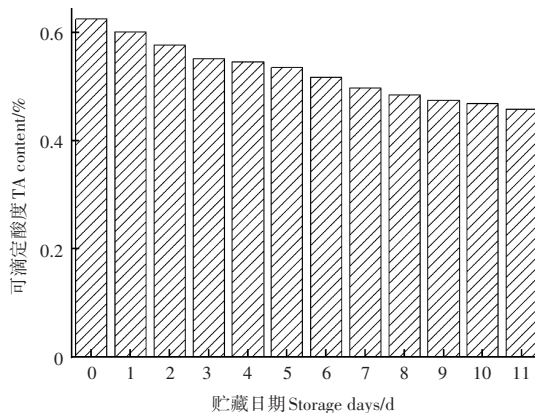


图2 南果梨贮藏过程中可滴定酸的变化

Figure 2 Changes in titratable acids during storage of Nanguo pear

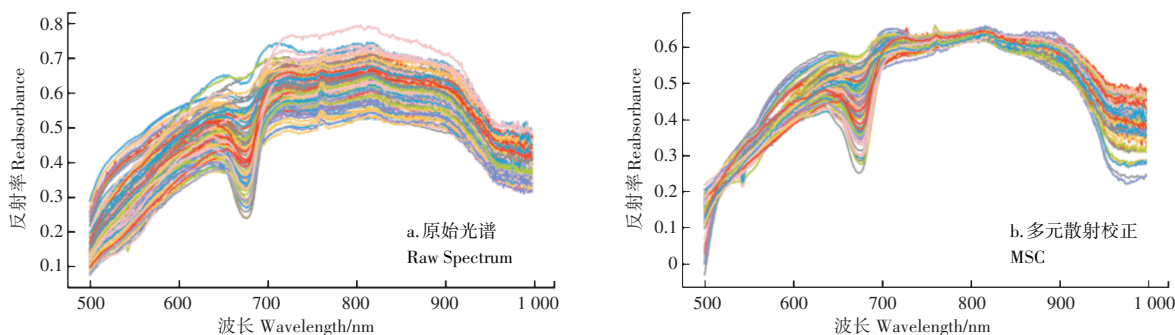


图3 ROI区域光谱反射率分布图

Figure 3 Distribution of spectral reflectivity in ROI region

2.3 数据集划分

利用 SPXY 方法将南果梨每组各 10 个样品(共 120 个)分成训练集(96 个)和预测集(24 个),得到的分割数据见表 1。由表 1 可知,训练集中包含了测试集中酸度的最大值和最小值,并且其分布范围广泛,这表明 SPXY 划分的数据集是合理的,有助于后续的定量预测建模效果。这种合理的数据集划分能够确保模型的预测性能和准确性^[25]。

表1 南果梨数据集划分

Table 1 Division of Nanguo pear dataset

样本集 Set	样本数 Number	最大值 Maximum	最小值 Minimum	平均值 Average	标准差 Standard deviation	变异系数 Coefficient of variation
训练集 Training set	96	0.625	0.399	0.534	0.045	8.43
预测集 Prediction set	24	0.564	0.399	0.475	0.055	11.58
总样本 Total samples	120	0.625	0.399	0.505	0.050	10.00

2.4 光谱预处理

利用不同预处理方式得到的南果梨高光谱数据和南果梨可滴定酸数据进行 PLSR 建模,得到的数据如表 2。由表 2 可知, MSC 对南果梨高光谱数据的预处理方法具有较高的 R^2 和较低的 RMSE 值,说明 MSC 具有较好的降噪能力,提高了模型的预报精度^[26]。

2.5 特征波段提取

首先对高光谱数据进行了 MSC 的预处理,然后再使用 SPA 算法来提取特征波长,并设定特征光

谱变量数在5~50之间变动。由图4可知,在RMSE约为0.049的情况下,提取了9个特征光谱变量,这些变量占据了全光谱范围的2%。由图5可知,SPA算法筛选的特征波段主要集中在排序在200~350之间的波段点,对应的是光谱波长700~900 nm范围。

表2 不同预处理方法建模对比

Table 2 Comparison of modeling with different preprocessing methods

方法 Methods	训练集 Training set		预测集 Prediction set	
	R_c^2	RMSEC	R_p^2	RMSEP
None	0.753	0.053	0.702	0.058
SNV	0.783	0.050	0.732	0.055
MSC	0.830	0.044	0.791	0.049
SG	0.805	0.047	0.776	0.050
Move_avg	0.782	0.050	0.740	0.054
D2	0.802	0.047	0.767	0.051
Wave	0.780	0.051	0.729	0.055
Mean_centralization	0.757	0.052	0.723	0.056

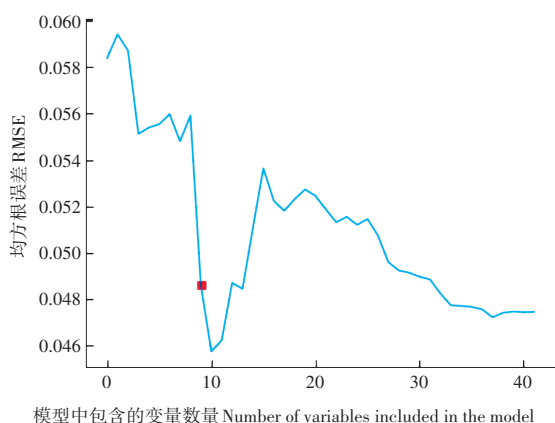


图4 SPA算法选择的RMSE分布图

Figure 4 RMSE distribution map for SPA algorithm selection

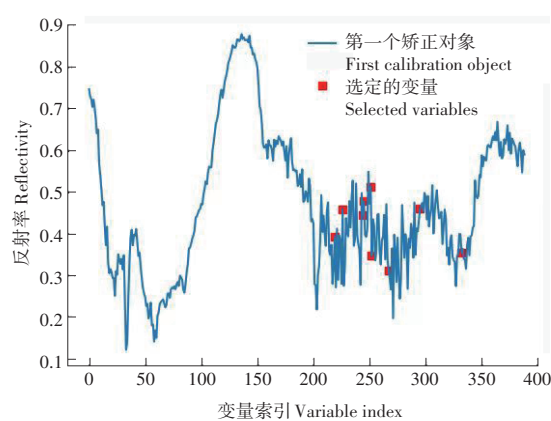


图5 SPA算法选择的最优特征波长

Figure 5 The optimal feature wavelength selected by the SPA algorithm

2.6 模型建立与比较

将SPA算法提取出的9个特征波长作为输入,分别建立不同模型进行训练。在这个过程中模型会学习数据的内在规律和模式^[27]。然后通过训练集和测试集的决定系数和均方根误差来对模型进行评估,如果发现模型效果不理想,就需要采取调整模型参数和更换模型等措施^[28]。

2.6.1 ELM模型的建立 以SPA算法中提取出的特征波长为输入,构建基于ELM模型的TA预测模型。ELM模型激活函数使用ReLU,隐藏层神经元个数为16,预测结果如图6。预测结果显示,ELM模型预测效果比PLSR模型有所降低,因此继续采用BP神经网络模型对南果梨酸度进行预测。

2.6.2 BP神经网络模型的建立 利用SPA算法提取经过MSC预处理后的光谱数据的特征波段,然后设置BP神经网络模型参数:激活函数使用ReLU,输入层节点9,学习率0.000 1,迭代次数10 000次。建模结果如图7。结果表明:BP神经网络模型在训练集上达到比较理想的预测效果, R_c^2 与RMSEC达到0.805与0.047,但是预测集的预测效果却有了显著降低, R_p^2 与RMSEP分别为0.729与0.055,说明利用SPA算法对特征波长进行提取所建立起来的BP神经网络模型有一定程度的过拟合^[29]。因此,本实验使用GA算法和PSO算法对BP模型进行优化,提高对新数据的泛化能力。

2.6.3 GA、PSO优化BP神经网络模型的建立 利用SPA算法提取的特征波长作为输入,构建基于GA-BP、PSO-BP的TA预测模型。设置GA-BP参数:激活函数使用ReLU,输入层节点9,种群大小为150,交叉概率0.7,学习率0.000 1,迭代次数10 000次;PSO-BP的参数:粒子群规模60,学习系数 $c_1=1.8$, $c_2=$

1.6,群体中的速率和空间区间为[-1,1],[−5,5]。建模结果如图8和图9,由图8和图9可知,采用不同方法对BP神经网络模型进行了优化,得到了良好的预测效果,其预测的数据集中在拟合曲线上,只有少数点偏离了直线,表明该方法的预测效果与实际情况相差不大,具有良好的预测效果。

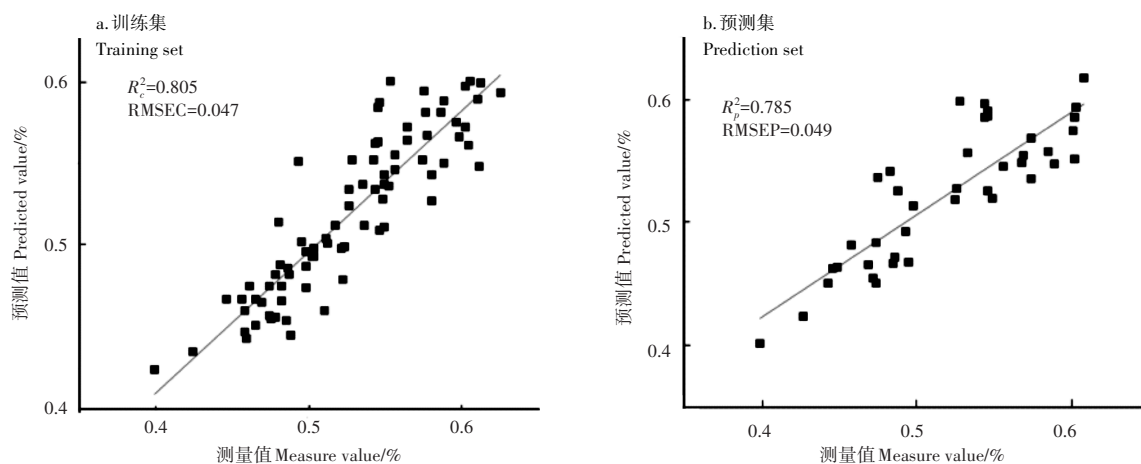


图6 ELM模型预测效果

Figure 6 ELM model prediction performance

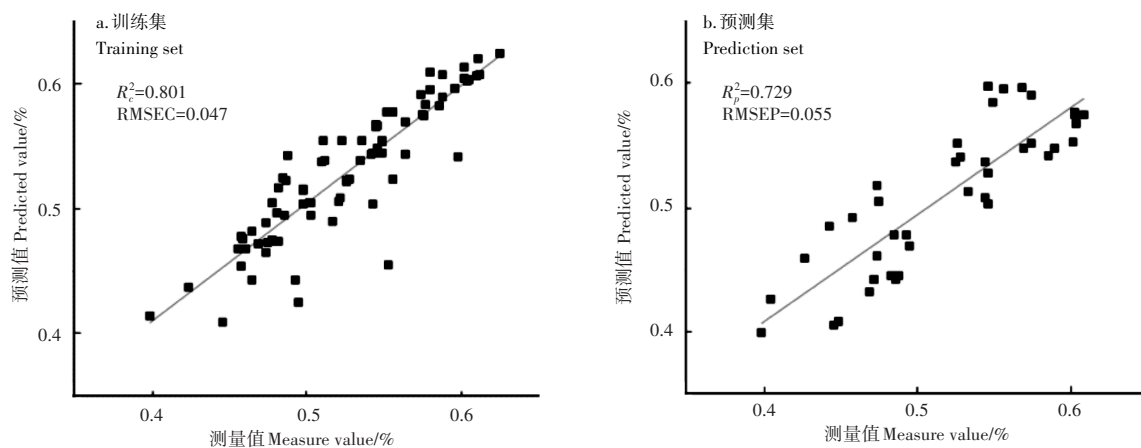


图7 BP模型预测效果

Figure 7 BP model prediction performance

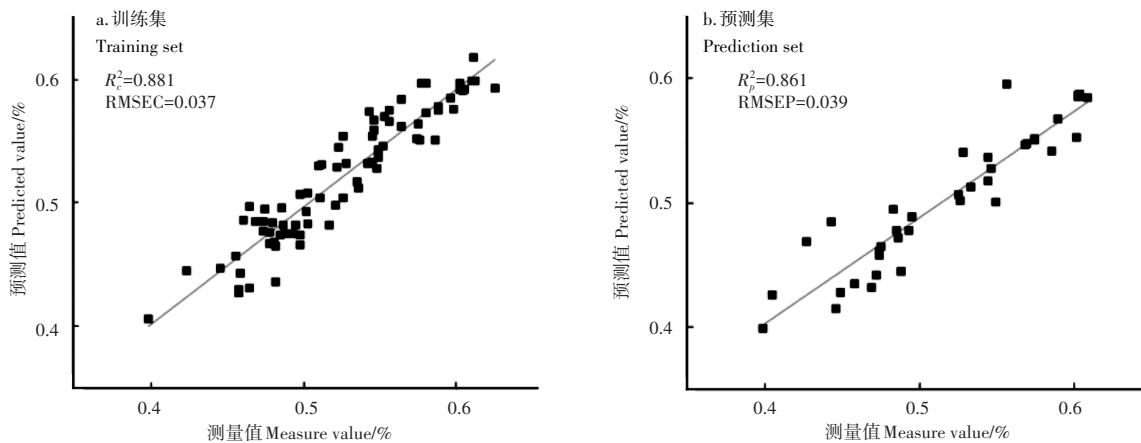


图8 GA-BP模型预测效果

Figure 8 GA-BP model prediction performance

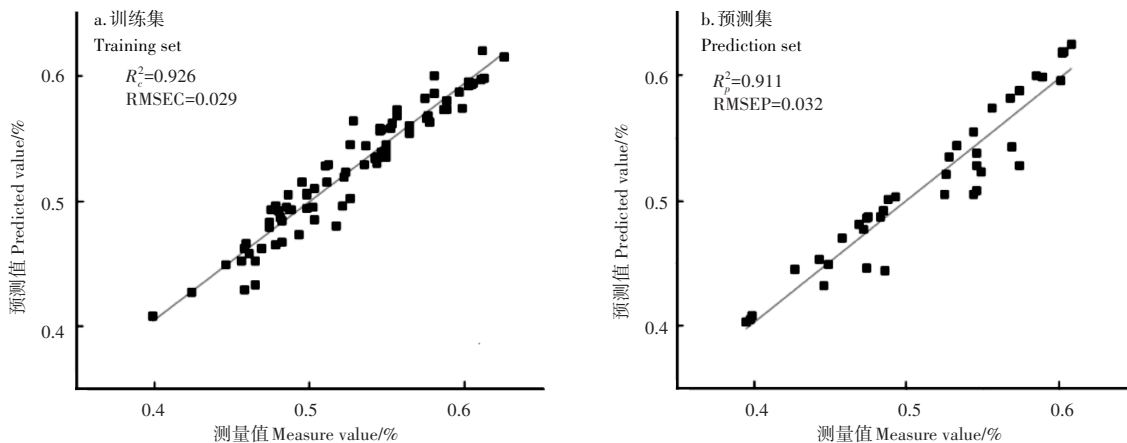


图9 PSO-BP模型预测效果

Figure 9 PSO-BP model prediction performance

由此进一步说明,GA算法和PSO算法对BP神经网络模型的输入权重和偏置起到了良好的优化作用,提高了模型的泛化能力和预测精度。

2.6.4 SPA-PLSR、SPA-GA-BP和SPA-PSO-BP模型结果比较 将SPA-PLSR、SPA-GA-BP、SPA-PSO-BP的3种模型进行分析和比较,结果如表3。SPA-PSO-BP与其他两个模型相比具有最好的预测准确性。与SPA-PLSR、SPA-GA-BP相比,训练集的预测效果 R_c^2 比SPA-PLSR和SPA-GA-BP的预测效果好,提高11.6%和5.1%,RMSEC值也有所降低,对比前者降低34%,对比后者降低21.6%;预测集 R_p^2 预测结果较SPA-PLSR模型和SPA-GA-BP模型分别提高15.2%和5.9%,RMSEP分别下降34.7%和17.9%。可见,无论是训练集模型还是预测集模型,SPA-PSO-BP模型对南果梨酸度的预测都有较好的表现。说明PSO算法具有更好的全局优化能力,更快速准确地帮助BP神经网络找到最优解。

表3 不同模型对比

Table 3 Comparison of different models

模型 Model	训练集 Training set		预测集 Prediction set	
	R_c^2	RMSEC	R_p^2	RMSEP
SPA-PLSR	0.830	0.044	0.791	0.049
SPA-GA-BP	0.881	0.037	0.861	0.039
SPA-PSO-BP	0.926	0.029	0.911	0.032

3 讨论与结论

采用高光谱成像技术对南果梨内部酸度进行无损检测研究,结果表明:(1)基于MSC、SNV、SG平滑、滑动平均滤波、二阶差分、小波变换、均值中心化预处理原始光谱数据,并结合理化实验测出的南果梨TA值建立了PLSR模型。数据显示,MSC预处理方法表现最优。(2)将MSC预处理光谱利用SPA进行特征波长提取,得到9个特征波长变量。(3)输入SPA方法提取出的特征波长,构造ELM、PLSR、BP、GA-BP及PSO-BP南果梨TA预测模型。对比分析,SPA-PSO-BP模型对南果梨TA的预测最佳, R_c^2 和 R_p^2 分别为0.926和0.911,RMSEC和RMSEP分别为0.029和0.032。

本研究使用高光谱成像技术对南果梨酸度数据的特征波段进行筛选,提取其品质指标的相关特征波段,并进一步建立了南果梨酸度的预测模型,获得了良好的预测效果,为南果梨品质无损检测提供了参考依据。在今后的研究中,可以增加南果梨的种类、选用更多特征波长提取方法进行对比,以及优化特征波长的提取算法等,以便进一步提高模型预测的精度。

参考文献:

[1] 孟庆龙,黄人帅,张 艳,等.贮藏期内猕猴桃酸度的快速无损检测[J].农产品加工,2022(11):66-68.
 [2] 卢 娜,韩 平,王纪华.基于高光谱成像技术的草莓硬度预测[J].软件导刊,2018,17(3):180-182.

- [3] TAN F,MO X M,RUAN S W,et al.Combining vis-NIR and NIR spectral imaging techniques with data fusion for rapid and nondestructive multi-quality detection of cherry tomatoes[J].Foods,2023,12(19):3621.
- [4] KALOPESA E,GKRIMPIZIST,SAMARINAS N,et al.Rapid determination of wine grape maturity level from pH,titratable acidity,and sugar content using non-destructive in situ infrared spectroscopy and multi-head attention convolutional neural networks[J].Sensors,2023,23(23):9536.
- [5] 蒋 蕙,罗亚辉,胡文武,等.基于高光谱的油茶籽内部品质检测最优预测模型研究[J].农机化研究,2015,37(7):56-60.
- [6] ÇETIN N,KARAMAN K,KAVUNCUOĞLUE,et al.Using hyperspectral imaging technology and machine learning algorithms for assessing internal quality parameters of apple fruits[J].Chemometrics and Intelligent Laboratory Systems,2022,230:104650.
- [7] WANG J R,WUW Z,TIAN S P,et al.Non-destructive determination of bayberry sugar and acidity by hyperspectral remote sensing of Si-sensor and low-cost portable instrument development[J].Sensors,2023,23(24):9822.
- [8] 常洪娟,蒙庆华,吴哲锋,等.基于反向传播神经网络和高光谱成像的芒果可溶性固形物含量检测[J].食品安全质量检测学报,2024,15(2):141-148.
- [9] 庄浩轩,魏明生,王 波,等.基于高光谱成像技术的糯玉米种子分类研究[J].现代农业研究,2024,30(1):51-57.
- [10] 马文强,张 漫,李 源,等.基于高光谱成像的核桃仁品质检测与分类方法[J].分析化学,2020,48(12):1737-1746.
- [11] AKBAR M,ALI N,IMRAN M,et al.Spherical Fe₂O₃ nanoparticles inhibit the production of aflatoxins (B1 and B2) and regulate total soluble solids and titratable acidity of peach fruit[J].International Journal of Food Microbiology,2024,410:110508.
- [12] SHANTHINI K S,GEORGE S N,GEORGE S,et al.Hyperspectral fruit image restoration using non-convex optimization[J].Journal of Food Engineering,2023,358:111662.
- [13] 刘宝花.基于高光谱成像技术的鲜食葡萄内部品质无损检测研究[D].泰安:山东农业大学,2020.
- [14] 叶伟欣.基于高光谱成像技术和机器学习的干枣品质检测及平台开发[D].石河子:石河子大学,2023.
- [15] XU M,SUN J,CHENG J H,et al.Non-destructive prediction of total soluble solids and titratable acidity in Kyoho grape using hyperspectral imaging and deep learning algorithm[J].International Journal of Food Science & Technology,2023,58(1):9-21.
- [16] LE B T,HA T T L.Hyperspectral image classification based on average spectral-spatial features and improved hierarchical-ELM[J].Infrared Physics & Technology,2019,102:103013.
- [17] LIN Y X,XU C Q,LYU S Q.Disease regions recognition on mural hyperspectral images combined by MNF and BP neural network[J].Journal of Physics:Conference Series,2019,1325(1):012095.
- [18] 许童羽,方健羽,郭忠辉,等.东北地区稻田土壤氮含量无人机高光谱反演建模研究[J].沈阳农业大学学报,2023,54(6):759-768.
- [19] CHAKRABORTY R,SUSHIL R,GARG M L.Hyper-spectral image segmentation using an improved PSO aided with multilevel fuzzy entropy[J].Multimedia Tools and Applications,2019,78(23):34027-34063.
- [20] 崔小涛,常庆瑞,屈春燕,等.基于高光谱和MLSR-GA-BP神经网络模型油菜叶片SPAD值遥感估算[J].东北农业大学学报,2020,51(8):74-84.
- [21] 孙 通,李翰林,孔令飞,等.光源光斑直径对苹果霉心病近红外检测的影响[J].农业工程学报,2023,39(24):298-305.
- [22] WANG Q Y,SONG S,LI L,et al.An extreme learning machine optimized by differential evolution and artificial bee colony for predicting the concentration of whole blood with Fourier Transform Raman spectroscopy[J].Spectrochimica Acta Part A,Molecular and Biomolecular Spectroscopy,2023,292:122423.
- [23] 陈敬鑫,张亚丽,高晓雅,等.减压贮藏对南果梨后熟进程及抗氧化能力的影响[J].中国食品学报,2023,23(11):202-211.
- [24] 汪晓慧.基于高光谱荧光成像技术的水蜜桃品质参数无损检测研究[D].雅安:四川农业大学,2023.
- [25] 姜凤利,杨 磊,田有文,等.基于高光谱成像的软枣猕猴桃SSC检测研究[J].沈阳农业大学学报,2023,54(3):318-326.
- [26] 李 斌,万 霞,刘爱伦,等.基于高光谱成像技术的涌泉蜜桔糖度最优检测位置[J].中国光学(中英文),2024,17(1):128-139.
- [27] 徐厚生,王继春.基于高光谱技术的一种TPC-TELM模型应用于土壤重金属含量反演研究[J].沈阳农业大学学报,2023,54(2):224-230.
- [28] 宋 科,潘磊庆,杨崇龙,等.受精种蛋理化指标与光谱信息的响应关系[J].南京农业大学学报,2023,46(6):1187-1195.
- [29] 刘小花,周彬静,彭 菁,等.基于电子鼻和高光谱成像技术的冷鲜牛肉微生物的生长模型构建[J].南京农业大学学报,2023,46(3):595-605.