

DOI: 10.19812/j.cnki.jfsq11-5956/ts.20250205002

引用格式: 郭欣悦, 陈国良, 朱良宽, 等. 基于高光谱成像技术的蓝莓糖度无损检测模型优化研究[J]. 食品安全质量检测学报, 2025, 16(11): 207-214.

GUO XY, CHEN GL, ZHU LK, *et al.* Optimization research on non-destructive detection model of *Vaccinium* spp. sugar content based on hyperspectral imaging technology [J]. Journal of Food Safety & Quality, 2025, 16(11): 207-214. (in Chinese with English abstract).

基于高光谱成像技术的蓝莓糖度无损检测 模型优化研究

郭欣悦¹, 陈国良², 朱良宽^{1*}, 刘大洋², 孙枭雄¹

(1. 东北林业大学机电工程学院, 哈尔滨 150040; 2. 东北林业大学计算机与控制工程学院, 哈尔滨 150040)

摘要: **目的** 基于高光谱成像技术优化蓝莓糖度的无损检测模型。**方法** 以产自丹东的 L25 品种蓝莓为实验对象, 采用高光谱成像技术获取 900~1700 nm 波长范围内的蓝莓高光谱图像, 并计算感兴趣区域的平均光谱作为原始数据。通过多元散射校正(multiple scatter correction, MSC)、标准正态变量变换(standard normal variable, SNV)和 Savitzky-Golay (SG) 3 种预处理方法改善光谱数据质量。基于预处理后的全波长数据, 分别建立偏最小二乘回归法(partial least squares, PLS)、反向传播神经网络(back propagation neural network, BPNN)、支持向量机回归法(support vector regression, SVR)等糖度预测模型。**结果** 在 MSC 和 SNV 预处理下, PLSR 模型在糖度预测过程中表现出较好的预测性能, 均方根误差(root mean squared error of prediction, RMSEP)分别为 0.3586、0.3599。**结论** 本研究优化了基于高光谱成像技术的蓝莓糖度无损检测模型, 为蓝莓糖度的快速、准确预测提供了有效的技术支持, 具有较强的实际应用潜力。

关键词: 蓝莓糖度; 无损检测; 高光谱成像技术; 机器学习

Optimization research on non-destructive detection model of *Vaccinium* spp. sugar content based on hyperspectral imaging technology

GUO Xin-Yue¹, CHEN Guo-Liang², ZHU Liang-Kuan^{1*}, LIU Da-Yang², SUN Xiao-Xiong¹

(1. College of Mechanical and Electrical Engineering, Northeast Forestry University, Harbin 150040, China; 2. College of Computer and Control Engineering, Northeast Forestry University, Harbin 150040, China)

ABSTRACT: Objective To optimize a non-destructive detection model for predicting *Vaccinium* spp. sugar content using hyperspectral imaging technology. **Methods** The L25 variety of blueberries from Dandong was selected as the subject, and hyperspectral imaging technology was acquired in the wavelength range of 900–1700 nm. The average spectrum of the region of interest was calculated as the raw data. The 3 kinds of preprocessing methods, including multiple scatter correction (MSC), standard normal variate (SNV) and Savitzky-Golay (SG), were applied to improve the spectral

收稿日期: 2025-02-05

基金项目: 中央高校基本科研业务经费创新团队项目(2572023CT15); 国家自然科学基金项目(32202147)

第一作者: 郭欣悦(1997—), 女, 硕士研究生, 主要研究方向为高光谱林业经济作物无损检测。E-mail: guoxy_1028@163.com

*通信作者: 朱良宽(1978—), 男, 博士, 教授, 主要研究方向为农林产品无损检测技术与应用。E-mail: zhulk@126.com

data quality. Non-destructive sugar content prediction models were established using partial least squares regression (PLSR), back propagation neural network (BPNN), and support vector regression (SVR) based on the full-wavelength data after preprocessing. **Results** The experimental results demonstrated that the PLSR model, with MSC and SNV preprocessing, exhibited the best performance, achieving root mean square error of prediction (RMSEP) values of 0.3586 and 0.3599, respectively. **Conclusion** This study provides an optimized non-destructive detection model for *Vaccinium* spp. sugar content, offering effective technical support for rapid and accurate sugar content prediction with significant practical potential.

KEY WORDS: *Vaccinium* spp. sugar content; non-destructive detection; hyperspectral imaging technology; machine learning

0 引言

蓝莓作为一种营养价值丰富且具有多种保健功能的浆果,近年来受到越来越多消费者的青睐,其市场需求持续增长^[1-2]。蓝莓不仅可以直接鲜食,还可以加工成果干、果汁等多样化的营养产品,具有广泛的市场应用前景。在蓝莓的众多品质指标中,糖度是决定其风味和品质的关键因素之一。糖度不仅影响蓝莓的口感,还是合成维生素和类胡萝卜素等营养成分的基础原料^[3-4]。此外,糖度也是确定蓝莓采摘和储存时间的重要指标,在农业生产和食品加工领域,准确测定蓝莓的糖度对于生产者和消费者都具有至关重要的意义^[5-6]。因此,有必要采用一种快捷可靠且无破坏性的方法对蓝莓糖度进行测定^[7-8]。

高光谱成像技术是一种融合光谱分析与图像采集的先进技术^[9-10],能够同时获取样本的空间和光谱信息,具有非接触、无损、高精度等特点。在水果品质检测领域,该技术已被广泛应用,如猕猴桃、苹果等水果糖度的无损检测。例如,冯树南等^[11]利用高光谱成像技术对枇杷糖度进行无损检测,通过连续投影法(successive projections algorithm, SPA)筛选特征光谱并建立多元线性回归(multiple linear regression, MLR)模型,成功实现了枇杷糖度的可视化预测。许丽佳等^[12]开展了基于高光谱成像技术的猕猴桃糖度无损检测研究,采用直接正交信号校正(orthogonal signal correction, DOSA)预处理和多种特征波长提取算法,结合极限学习机(extreme learning machine, ELM)建立了预测模型,取得了良好的预测效果。班兆军等^[13]将高光谱成像技术与深度学习相结合,对苹果品质进行无损检测,通过多种预处理方法和特征波长提取算法,利用支持向量机(support vector machine, SVM)、卷积神经网络(convolutional neural network, CNN)及基于深度学习的定量光谱数据分析模型(DeepSpectra)对苹果的糖酸比进行预测,其中基于深度学习的定量光谱数据分析(DeepSpectra)模型预测准确率高达 93.70%。这些研究表明,高光谱成像技术在水果内部品质检测中具有广阔的应用前景,并且在品质检测方面取得了显著成果。

蓝莓作为小型水果,其果实相较于苹果、猕猴桃等水果更小且汁水较少,这使得其糖度测量在操作上需要额外

的处理和更精细的技术支持。近年来,研究者们开始探索将高光谱成像技术应用于蓝莓等小型水果的品质检测。例如,在《基于高光谱图像的蓝莓糖度和硬度无损测量》一文中,研究人员采用高光谱图像技术对蓝莓的糖度与硬度进行预测。然而,糖度预测的结果存在一定偏差,且准确度明显低于苹果、柑橘等大果实的糖度测量效果。这一现象主要归因于图像分辨率较低和光源系统设计不完善,导致采集到的高光谱图像质量不高,从而影响了预测模型的性能。因此,未来的研究需要采用更高分辨率的图像采集系统,并优化光源照明设计,以提高蓝莓高光谱图像数据的准确性,从而实现更加可靠的糖度预测。

综上,在多数水果(例如猕猴桃、苹果、梨等)品质检测方面,高光谱成像技术已被广泛应用于糖度的无损检测,然而针对蓝莓等小型水果的糖度检测研究还存在预测精度不足、成像质量有待提升等问题,研究还处于探索阶段。因此,本研究以高光谱成像系统作为实验平台,获取蓝莓高光谱成像信息,并对原始光谱进行预处理以改善光谱数据质量,进一步基于统计学习与机器学习理论建立蓝莓糖度预测模型,并对所建立的模型进行优化及对比分析,旨在为蓝莓糖度的无损检测研究提供一种可行的思路。

1 材料与方法

1.1 实验样本

本研究选用丹东地区的蓝莓品种 L25。该品种具有早熟、果实硬脆、极甜、伴有浓郁花蜜香气的特点。为了确保实验数据的准确性和一致性,剔除了形状异常及表面有损伤的样本,最终选取了 300 颗大小均匀、表面光滑、外观完整的蓝莓果实作为实验材料。总体实验步骤包括。

(1)实验前,将筛选后的蓝莓果实密封保存于冰箱中以维持其新鲜度。

(2)实验开始前,提前将果实从冰箱中取出,待果实达到室温后进行实验操作。

(3)实验中首先对蓝莓果实表面用蒸馏水进行除尘处理,以去除灰尘和其他可能影响光谱采集的表面干扰物。随后,将蓝莓果实按照顺序编号,并将 300 颗蓝莓分成 6 组,每组 50 颗,将其按照每行 5 颗每列 10 颗的排列方式进行布局。每组蓝莓依次进行高光谱图像的采集与分析。

1.2 光谱采集系统

本研究采用由芬兰 SPECIM 公司开发的高光谱成像系统作为实验平台。该系统主要由光谱相机(Specim FX17, Specim, Finland)、CCD 相机、6 个 20 W 的卤素灯、电动扫描平台及计算机组成。光谱相机的波长范围为 900~1700 nm, 采样分辨率为 8 nm, 像素分辨率为 1198×512。光谱数据通过 ENVI 5.3 软件采集, 使用的数据采集软件为 SpectralSENS (Spectral Imaging Lid, Finland), 如图 1 所示。

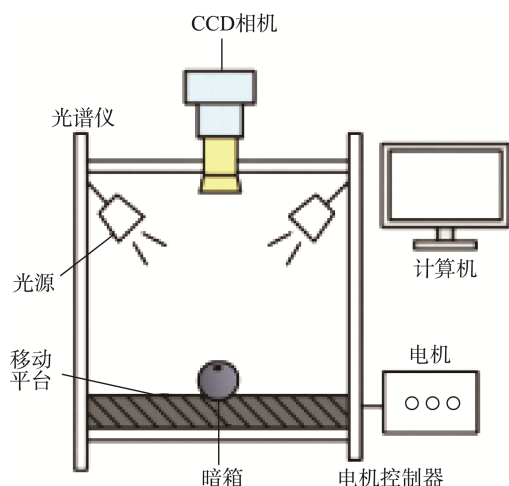


图1 高光谱成像系统

Fig.1 Hyperspectral imaging systems

在进行高光谱成像数据采集前, 需对地面进行校正以确保数据的准确性。此外, 为了保证图像的清晰度和避免失真, 采集前进行了对焦调整。实验过程中, 测试平台以 2 cm/s 的速度移动, 曝光时间设置为 20 ms。在数据采集前, 打开卤素灯并稳定光源。将蓝莓样品按照编号分组, 每组 50 颗, 按果柄端朝上放置, 如图 2 所示, 蓝莓样品整齐地排列在平台上。在数据采集过程中, 平台以设定的速度移动, 蓝莓样品在卤素灯下依次经过。当样品被光束照射时, 一部分光被样品吸收, 其余部分被反射, 而光谱信息则反映了蓝莓的内部特性。通过采集样品的反射光谱信息, 实验能够获得蓝莓的高光谱数据。



图2 蓝莓高光信息采集过程

Fig.2 Acquisition process of *Vaccinium* spp. hyperspectral information

1.3 光谱数据提取

为消除仪器和光谱数据的影响, 本研究对原始光谱数据进行了黑白校正^[14], 校正公式如式(1):

$$R = \frac{R_i - R_d}{R_w - R} \quad (1)$$

式中: R 为校正后的样本光谱数据; R_i 为原始光谱数据; R_d 为全黑标定图像; R_w 为标准白板图像。

在完成黑白校正后, 利用 ENVI 软件从高光谱图像中选取感兴趣区域(region of interest, ROI), 并提取该区域的平均光谱。提取的平均光谱被作为蓝莓样本的原始光谱曲线^[15-16]。经过分析, 原始光谱曲线表现优异, 因此选取波长范围为 900~1700 nm 的有效光谱段进行研究, 共包含 224 个波段。

1.4 糖度值测定

在样本光谱采集完成后, 立即对蓝莓糖度值进行测定。糖度测定采用全自动便携式数字折射仪(PAL-3, ATAGO CO, LTD, Tokyo, Japan)。实验过程中, 将蓝莓沿赤道方向切开, 用无纺布包裹果柄端后挤压榨汁, 滴加在折射仪的测试平台上进行测量。为了确保测量结果的准确性, 每个样品重复测定 3 次, 并取 3 次测量值的平均值作为最终糖度值记录^[17-18], 每次测量完成后, 使用蒸馏水彻底清洗折射仪的镜面, 以避免交叉污染, 确保下一次测量的准确性。

1.5 模型评价

为了评价模型的预测性能, 本研究采用以下 5 个指标^[19-21]。训练集相关系数(R_c), 衡量模型在训练集上的拟合程度, 值越接近 1, 模型在训练集上的表现越好; 训练集均方根误差(root mean square error of validation, RMSEV), 表示训练集上的预测误差, 值越小, 模型的拟合效果越好; 预测集相关系数(R_p), 衡量模型在预测集上的拟合能力, 值越接近 1, 模型的预测能力越强; 预测集均方根误差(root mean square error of prediction, RMSEP), 表示预测集上的预测误差, 值越小, 模型的预测精度越高; 相对分析误差(residual predictive deviation, RPD), 定义为预测集标准偏差与 RMSEP 的比值, 用于衡量模型的定量预测能力。RPD 值越大, 模型预测性能越优。RPD<1.5: 模型预测能力很差, 无法用于定量分析; 1.5≤RPD<2.0: 模型具有初步的区分能力, 可用于粗略分类; 2.0≤RPD<2.5: 模型适合用于筛选工作; 2.5≤RPD<3.0: 模型具有良好的定量预测能力; RPD≥3.0: 模型预测性能优异, 可用于高精度定量分析。

2 结果与分析

2.1 样本划分

本研究的实验样本划分采用 KS 法^[22], 并结合 10 折交叉验证方法进行模型评估。如表 1 所示, 训练集和预测集

的糖度预测值在训练集范围为 7.9%至 17.05%，数据划分合理且具有代表性。10 折交叉验证方法可以更全面地评估模型的泛化能力，并有效减少模型的偏倚。通过交叉验证，确保了模型在不同数据子集上的表现一致性，从而进一步增强了实验结果的可靠性与稳健性。

表 1 样本划分结果
Table 1 Results of sample division

样本集	样本数量	最小值/%	最大值/%	平均值/%	标准差/%
训练集	225	7.9	17.05	11.78±1.76	1.76
预测集	75	8.3	16.55	11.39±1.72	1.72
总样本	300	7.9	17.05	11.68±1.75	1.75

2.2 光谱预处理

原始高光谱数据常受到仪器误差、背景光干扰以及随机噪声等多种因素的影响，从而掩盖与样品自身性质相关的有效信息。为了降低噪声干扰，增强光谱特征的表达能力，有必要对原始光谱进行预处理。预处理的目的在于去除与目标变量无关的物理或化学干扰，改善光谱数据质量，降低模型复杂性，并提高后续建模的准确性与稳定性。

本研究针对原始光谱中可能存在的散射效应与基线漂移，选用了 3 种有效的光谱预处理方法，分别是多元散射校正(multiplicative scatter correction, MSC)、标准正态变量变换(standard normal variate, SNV)、Savitzky-Golay (SG)^[23-26]。图 3~6 展示了蓝莓样本在应用不同预处理方法后的光谱变化情况。通过比较不同预处理策略对光谱特征的影响，为后续建模选择最优的数据输入形式提供了依据。

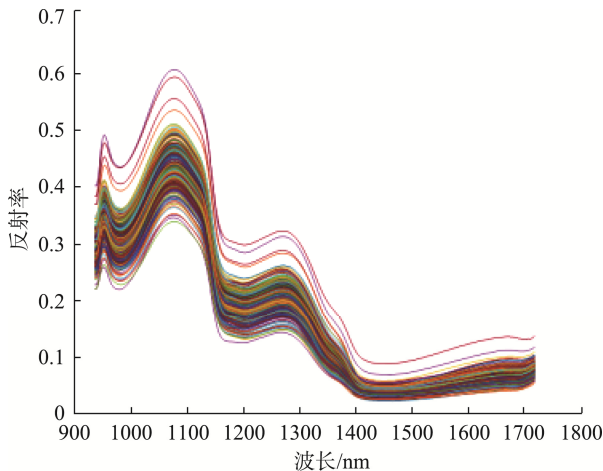


图3 蓝莓样本的原始光谱曲线

Fig.3 Raw spectral curves of *Vaccinium* spp. samples

2.3 回归模型的建立与分析

在 900~1700 nm 的全波段范围内，采用预处理后的高光谱反射率数据进行建模分析。为确保模型构建的科学性与泛化能力，采用 Kennard-Stone (KS)算法对样本进行划分，将数据集合理分为训练集和预测集。随后，分别建立

蓝莓糖度的预测模型，选取 3 种回归方法进行建模，偏最小二乘回归(partial least squares regression, PLSR)、反向传播神经网络(back propagation neural network, BPNN)与支持向量回归(support vector regression, SVR)^[27-30]。

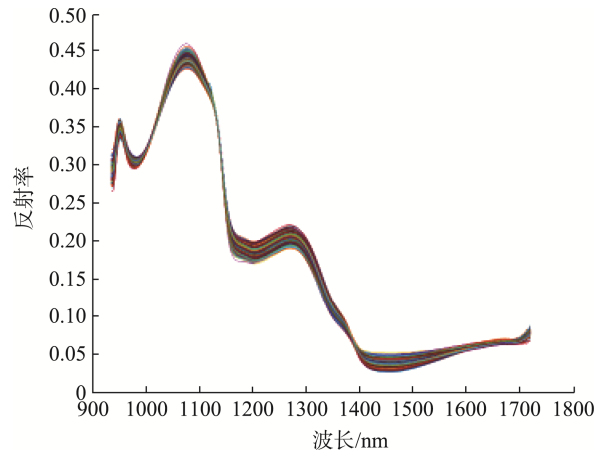


图4 MSC处理后的光谱曲线

Fig.4 Spectral curves after MSC treatment

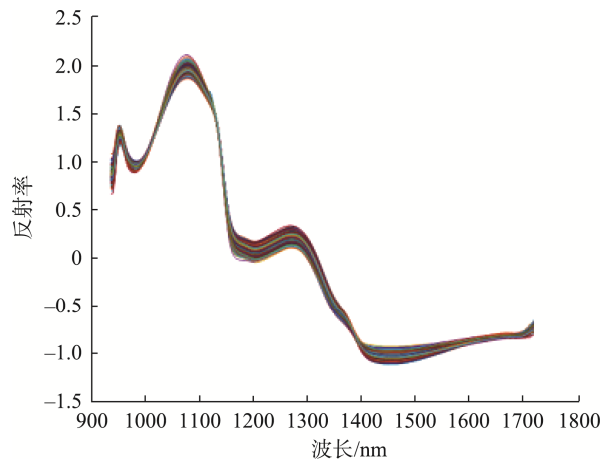


图5 SNV处理后的光谱曲线

Fig.5 Spectral curves after SNV treatment

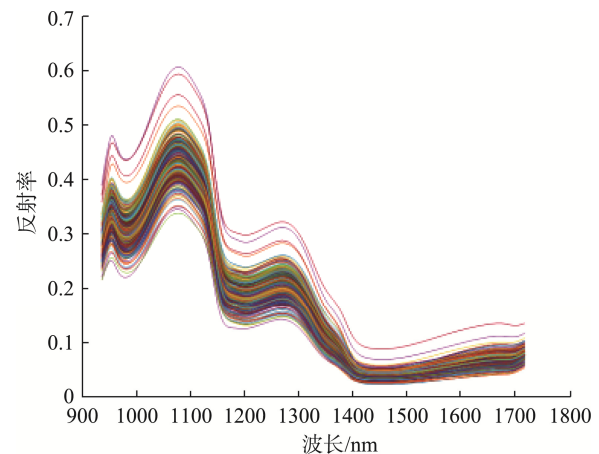


图6 SG处理后的光谱曲线

Fig.6 Spectral curves after SG treatment

PLSR^[27]是一种适用于处理高维、多重共线性问题的数据建模方法,它结合了主成分分析(principal component analysis, PCA)与 MLR 的优势,在化学计量学、光谱分析等高维数据建模领域应用广泛。PLSR 通过将原始自变量矩阵(X)与因变量矩阵(Y)同时投影至一个新的低维潜变量空间中,提取出一组既能最大程度解释 X 方差,又与 Y 高度相关的潜变量(latent variables, LVs)。本研究采用五折交叉验证方法确定 PLSR 模型的最优潜变量数,以避免模型出现过拟合或欠拟合现象。最终将预测集的光谱数据输入至训练完成的模型中,进行蓝莓糖度的预测。

BPNN^[29]是一种经典的前馈神经网络,通过反向传播算法不断调正网络权重和偏置,从而优化模型性能。由于其具备良好的非线性关系建模能力,广泛应用于分类、回归、模式识别等领域。在本研究中,构建的 BPNN 模型采用三层网络结构,包括输入层、隐藏层和输出层网络结构的输入层特征维度为 224,输入层的节点数为提取的特征维数,隐藏层的节点数为 20,输入层的节点数为 224,输出层为单节点,用于输出蓝莓糖度值。激活函数为 Sigmoid 函数,以增强模型的非线性拟合能力。模型训练过程中,学习率为 0.01,训练次数为 100 次,以确保网络收敛并获得稳定的预测性能。

SVR^[30]是一种基于 SVM 理论的回归分析方法,能够在高维特征空间中实现复杂的非线性函数拟合。SVR 通过引入核函数将原始输入映射到高维空间,在该空间中构建一个对误差具有容忍度的回归函数,从而在保证模型复杂度的同时提高预测精度。与传统回归方法相比,SVR 在处理高维、小样本、非线性和多重共线性问题方面具有显著优势。在本研究中,SVR 模型采用线性核函数进行建模,并通过五折交叉验证对其关键参数进行优化,包括惩罚因子(C)和核函数参数(如 γ)。通过参数优化过程,提升模型的拟合效果和泛化能力,从而实现蓝莓糖度的准确预测。

采用 MSC、SNV、SG 3 种预处理方法,分别建立 MSC-PLSR、MSC-BPNN、MSC-SVR、SNV-PLSR、SNV-BPNN、SNV-SVR、SG-PLSR、SG-BPNN、SG-SVR 9 种回归模型。

2.3.1 PLSR 建模与分析

对原始光谱分别利用多种预处理方法进行预处理,采用 KS 法进行样本划分,基于 PLSR 建立蓝莓糖度预测模型,基于不同预处理方法下 PLSR 模型对蓝莓糖度的预测结果如表 2 所示。

表 2 基于不同预处理方法下 PLSR 模型对蓝莓糖度的预测结果
Table 2 Prediction results of *Vaccinium* spp. sugar content based on PLSR model under different pretreatment methods

建模方法	预处理	训练集		预测集		
		R_c	RMSEV	R_p	RMSEP	RPD
PLSR	MSC	0.992	0.2094	0.978	0.3586	4.91
	SNV	0.992	0.2087	0.978	0.3599	4.88
	SG	0.973	0.4005	0.972	0.4173	4.11

实验结果表明,在不同预处理方法下构建的 PLSR 模型预测性能存在明显差异。其中 MSC 预处理下的 PLSR 模型表现最佳, $R_c=0.992$, $R_p=0.978$, $RPD=4.91$,是实现蓝莓糖度高精度预测的最优方案。SNV 次之,其模型性能接近 MSC, $RPD=4.88$,略低于 MSC,但仍表现优秀,可作为 MSC 的备选方法。而 SG 方法下的模型性能相对较弱, RPD 值为 4.11,虽然仍具有良好的预测能力,但整体精度略逊一筹,误差相对较大。综上,通过 MSC 预处理与 PLSR 建模,可以为蓝莓糖度的无损检测提供可靠的理论依据和技术支持。

2.3.2 BPNN 建模与分析

对原始光谱分别利用多种预处理方法进行预处理,采用 KS 法进行样本划分,基于 BPNN 建立蓝莓糖度预测模型,基于不同预处理方法下 BPNN 模型对蓝莓糖度的预测结果如表 3 所示。

表 3 基于不同预处理方法下 BPNN 模型对蓝莓糖度的预测结果
Table 3 Prediction results of *Vaccinium* spp. sugar content based on BPNN model under different pretreatment methods

建模方法	预处理	训练集		预测集		
		R_c	RMSEV	R_p	RMSEP	RPD
BPNN	MSC	0.975	0.4290	0.945	0.5871	2.99
	SNV	0.971	0.4197	0.963	0.4829	3.64
	SG	0.964	0.4618	0.922	0.6611	2.59

实验结果表明,SNV 预处理下构建的 BPNN 模型蓝莓糖度预测效果最佳, $R_c=0.971$, $R_p=0.963$, $RPD=3.64$,具有较高的预测精度和优秀的定量分析能力。MSC 预处理结合 BPNN 模型的性能略逊一筹, $RPD=2.99$ 值接近优秀预测标准,但与 SNV 相比,RMSEP 略高,整体模型精度稍低。SG 预处理的效果较差, RPD 值仅为 2.59,模型预测能力一般,RMSEP 最高为 0.6611,表明模型误差较大,仅适用于对精度要求较低的场景。

2.3.3 SVR 建模与分析

对原始光谱分别利用多种预处理方法进行预处理,采用 KS 法进行样本划分,基于 SVR 建立蓝莓糖度预测模型,基于不同预处理方法下 SVR 模型对蓝莓糖度的预测结果如表 4 所示。

表 4 基于不同预处理方法下 SVR 模型对蓝莓糖度的预测结果
Table 4 Prediction results of *Vaccinium* spp. sugar content based on SVR model under different pretreatment methods

建模方法	预处理	训练集		预测集		
		R_c	RMSEV	R_p	RMSEP	RPD
SVR	MSC	0.937	0.6324	0.952	0.5998	2.93
	SNV	0.937	0.6294	0.953	0.5892	2.99
	SG	0.819	1.0357	0.740	1.2112	1.41

实验结果表明,SNV 预处理结合 SVR 模型在蓝莓糖度预测中表现最佳, $R_c=0.937$, $R_p=0.953$, $RPD=2.99$,具有

较高的预测精度和良好的定量分析能力。MSC 预处理表现次优, R_c 、 R_p 等指标接近 SNV, 但 $RPD=2.93$ 略低, 预测能力良好, 可作为 SNV 的备选方案。而 SG 预处理的效果较差, RPD 值仅为 1.41, 表明模型预测能力较差, 不适用于高精度定量分析。

2.4 不同方法的比较

2.4.1 不同预处理方法的比较

MSC 和 SNV 预处理效果相近, 在两者预处理下, PLSR 模型表现最佳, 均达到优秀水平($RPD>4.8$)。SG 预处理效果稍弱, 在 SG 预处理下, 虽然 PLSR 表现尚可($RPD=4.11$), 但整体预测能力不如 MSC 和 SNV。

2.4.2 不同建模方法的比较

PLSR 优势较为明显, 在所有预处理方法下, PLSR 的 R_c 、 R_p 、RMSEP 和 RPD 均优于 BPNN 和 SVR, 表现出更高的预测精度和更小的误差。其中效果最佳的是采用 MSC 预处理建立的 PLSR 模型(图 7 为 MSC-PLSR 模型预测结果散点图), 为了进一步验证模型的稳定性和可靠性, 本研究采用了 10 折交叉验证方法, 并计算了交叉验证后的平均 RPD 值为 4.08, RPD 的标准差为 0.78。这些结果表明, 模型在不同数据集上的表现较为稳定, 且未出现过拟合现象。因此, 交叉验证进一步验证了 MSC-PLSR 模型的可靠性和泛化能力。BPNN 次优, 尽管 BPNN 的表现不及 PLSR, 但在 SNV 预处理下的 RPD 达到 3.64, 具有良好的预测能力。SVR 表现最弱, SVR 在所有预处理方法下的 R_c 和 R_p 较低, RMSEP 较大, RPD 未达到良好的预测能力($RPD<3$)。为了提高模型的泛化能力, 购买了山东临沂地区的莱克西品种蓝莓, 并对其进行了补充实验。实验中, 使用了 Kennard-Stone (KS) 算法进行样本划分, 采用了 MSC 进行数据预处理, 并使用 PLSR 进行建模。根据实验结果, 莱克西蓝莓的模型预测效果为: R_p 值为 0.965, RPD 为 3.65, 图 8 为莱克西品种的模型预测结果散点图。

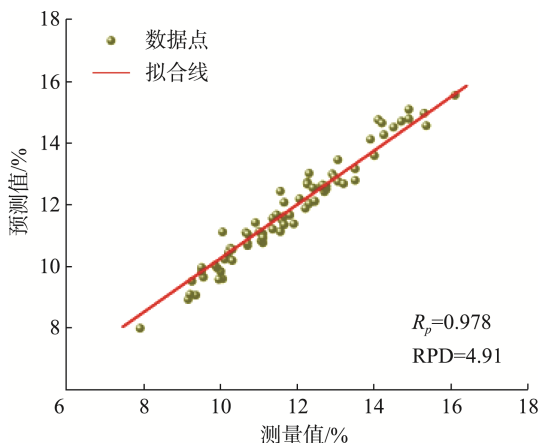


图7 MSC-PLSR模型预测集预测结果散点图

Fig.7 Scatter plot of the prediction results for the MSC-PLSR model prediction set

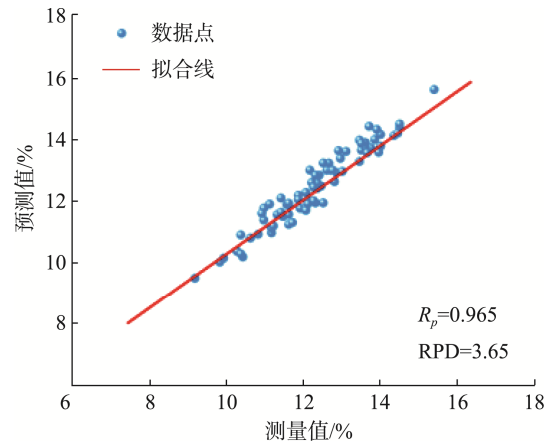


图8 莱克西品种的模型预测结果散点图

Fig.8 Scatter plot of the prediction results for the Lexi variety

通过对比使用 KS-MSC-PLSR 模型预测的 L25 蓝莓糖度值与实际测量值, 计算得到了预测集的平均回收率为 100.12%。这个回收率略大于 100%, 说明本研究的模型在某些情况下可能略微高估了糖度值。尽管如此, 这并不影响方法的可靠性, 且这一现象与实验设计和模型特性相关。在某些情况下, 模型可能会对糖度的预测略有偏差, 但这一点不会显著影响模型的整体性能, 且仍然能提供准确的糖度预测结果。

3 结论

本研究提出了一种基于高光谱图像技术的蓝莓糖度无损检测方法, 通过使用 MSC、SNV 和 SG 3 种方法对原始光谱进行预处理以改善光谱数据质量, 并基于预处理后的全波长数据分别构建了 PLSR、BPNN、SVR 蓝莓糖度 3 种预测模型, 在此基础上对所建立的回归模型进行对比分析。实验表明: 在 MSC 和 SNV 预处理下, PLSR 模型的表现出较好的预测性能, 但其中模型效果最佳的是采用 MSC 预处理结合 PLSR 建模方法(MSC-PLSR), 其 $R_c=0.992$, $R_p=0.978$, RMSEV=0.2094, RMSEP=0.3586, RPD=4.91, 说明 PLSR 能够充分挖掘光谱数据中的信息, 且具有良好的泛化能力。本研究为实现蓝莓糖度的无损检测提供了一种可行的思路和技术参考。

参考文献

- [1] SILVA S, COSTA EM, VEIGA M, *et al.* Health promoting properties of blueberries: A review [J]. *Critical Reviews in Food Science and Nutrition*, 2020, 60(2): 181-200.
- [2] 李亚东, 刘成, 魏鑫, 等. 2024 年中国蓝莓产业发展报告[J]. *吉林农业大学学报*, 2025, 47(1): 1-14.
- LI YD, LIU C, WEI X, *et al.* Development report of 2024 China *Vaccinium* spp. industry [J]. *Journal of Jilin Agricultural University*, 2025, 47(1): 1-14.
- [3] 赵慧芳, 吴文龙, 黄正金, 等. 34 个蓝莓品种果实品质评价[J]. *植物资源*

- 源与环境学报, 2023, 32(4): 44–53, 72.
- ZHAO HF, WU WL, HUANG ZJ, *et al.* Evaluation on fruit quality of 34 *Vaccinium* spp. cultivars [J]. Journal of Plant Resources and Environment, 2023, 32(4): 44–53, 72.
- [4] TOTAD MG, SHARMA RR, SETHI S, *et al.* Genotypic variability in nutritional and functional attributes of *Vaccinium* spp. varieties grown in northern-western Himalayas [J]. Journal of Food Science and Technology, 2020, 57: 2251–2258.
- [5] CHIABRANDOV, GIACALONEG, ROLLE L. Mechanical behaviour and quality traits of high bush *Vaccinium* spp. during postharvest storage [J]. Journal of the Science of Food and Agriculture, 2009, 89(6): 989–992.
- [6] 纪淑娟, 马超, 周倩, 等. 蓝莓果实贮藏期间软化及相关指标的变化 [J]. 食品科学, 2013, 34(12): 341–345.
- JI SJ, MA C, ZHOU Q, *et al.* Softening and softening indices change of *Vaccinium* spp. during storage [J]. Food Science, 2013, 34(12): 341–345.
- [7] ZHAO J, CHEN Q, SARITPORN V, *et al.* Determination of apple firmness using hyperspectral imaging technique and multivariate calibrations [J]. Transactions of the Chinese Society of Agricultural Engineering, 2009, 25(11): 226–231.
- [8] ZHANG BH, LI JB, FAN SX, *et al.* Principles and applications of hyperspectral imaging technique in quality and safety inspection of fruits and vegetables [J]. Spectroscopy and Spectral Analysis, 2014, 34(10): 2743–2751.
- [9] MITCHELL JBO. Machine learning methods in chemoinformatics [J]. Wiley Interdisciplinary Reviews: Computational Molecular Science, 2014, 4(5): 468–481.
- [10] GAO S, XU J. Hyperspectral image information fusion-based detection of soluble solids content in red globe grapes [J]. Computers and Electronics in Agriculture, 2022, 196: 106822.
- [11] 冯树南, 谭涛, 尚静, 等. 枇杷糖度无损检测及可视化研究 [J]. 中国农业化学报, 2024, 45(2): 157–163.
- FENG SN, TAN T, SHANG J, *et al.* Non-destruction detection and visualization of sugar content in loquat [J]. Journal of Chinese Agricultural Mechanization, 2024, 45(2): 157–163.
- [12] 许丽佳, 陈铭, 王玉超, 等. 高光谱成像的猕猴桃糖度无损检测方法 [J]. 光谱学与光谱分析, 2021, 41(7): 2188–2195.
- XU LJ, CHEN M, WANG YC, *et al.* Study on Non-destructive detection method of kiwifruit sugar content based on hyperspectral imaging technology [J]. Spectroscopy and Spectral Analysis, 2021, 41(7): 2188–2195.
- [13] 班兆军, 高喧翔, 马肆恒, 等. 基于高光谱和深度学习的苹果品质无损检测方法 [J]. 江苏农业学报, 2024, 40(8): 1446–1454.
- BAN ZJ, GAO XX, MA YH, *et al.* Non-destructive detection method of apple based on hyperspectral and deep learning [J]. Jiangsu Journal of Agricultural Sciences, 2024, 40(8): 1446–1454.
- [14] HU R, ZHANG L, YU Z, *et al.* Optimization of soluble solids content prediction models in ‘Hami’ melons by means of Vis-NIR spectroscopy and chemometric tools [J]. Infrared Physics & Technology, 2019, 102: 102999.
- [15] 王子轩, 杨良, 黄凌霞, 等. 高光谱结合深度学习的桑椹采后 TSS 含量无损检测 [J]. 光谱学与光谱分析, 2024, 44(6): 1724–1730.
- WANG ZX, YANG L, HUANG LX, *et al.* Nondestructive determination of TSS content in postharvest mulberry fruits using hyperspectral imaging and deep learning [J]. Spectroscopy and Spectral Analysis, 2024, 44(6): 1724–1730.
- [16] 王广来, 王恩凤, 王聪聪, 等. 基于高光谱图像技术与迁移学习的水晶梨早期损伤检测 [J]. 光谱学与光谱分析, 2022, 42(11): 3626–3630.
- WANG GL, WANG ENF, WANG CC, *et al.* Early bruise detection of crystal pear based on hyperspectral imaging technology and transfer learning [J]. Spectroscopy and Spectral Analysis, 2022, 42(11): 3626–3630.
- [17] 黄文倩, 李江波, 陈立平, 等. 以高光谱数据有效预测苹果可溶性固形物含量 [J]. 光谱学与光谱分析, 2013, 33(10): 2843–2846.
- HUANG WQ, LI JB, CHEN LP, *et al.* Effectively predicting soluble solids content in apple based on hyperspectral imaging [J]. Spectroscopy and Spectral Analysis, 2013, 33(10): 2843–2846.
- [18] 邵园园, 王永贤, 玄冠涛, 等. 高光谱成像快速检测壳聚糖涂膜草莓可溶性固形物 [J]. 农业工程学报, 2019, 35(18): 245–254.
- SHAO YY, WANG YX, XUAN GT, *et al.* Rapid detection of soluble solids content in strawberry coated with chitosan based on hyperspectral imaging [J]. Transactions of the Chinese Society of Agricultural Engineering, 2019, 35(18): 245–254.
- [19] 马本学, 肖文东, 祁想想, 等. 基于漫反射高光谱成像技术的哈密瓜糖度无损检测研究 [J]. 光谱学与光谱分析, 2012, 32(11): 3093–3097.
- MA BX, XIAO WD, QI XX, *et al.* Nondestructive measurement of sugar content of Hami melon based on diffuse reflection hyperspectral imaging technology [J]. Spectroscopy and Spectral Analysis, 2012, 32(11): 3093–3097.
- [20] 郭志明, 王郡艺, 宋焯, 等. 手持式可见近红外苹果品质无损检测系统设计与试验 [J]. 农业工程学报, 2021, 37(22): 271–277.
- GUO ZM, WANG JY, SONG Y, *et al.* Design and experiment of the handheld visible-near infrared nondestructive detecting system for apple quality [J]. Transactions of the Chinese Society of Agricultural Engineering, 2021, 37(22): 271–277.
- [21] 薛建新, 张淑娟, 张晶晶. 高光谱成像技术对鲜枣内外部品质检测的研究 [J]. 光谱学与光谱分析, 2015, 35(8): 2297–2302.
- XUE JX, ZHANG SJ, ZHANG JJ. Simultaneous detection of external and internal quality parameters of Huping jujube fruits using hyperspectral imaging technology [J]. Spectroscopy and Spectral Analysis, 2015, 35(8): 2297–2302.

- [22] XU Y, GOODACRE R. On splitting training and validation set: A comparative study of cross-validation, bootstrap and systematic sampling for estimating the generalization performance of supervised learning [J]. *Journal of Analysis and Testing*, 2018, 2(3): 249–262.
- [23] 霍迎秋, 张晨, 李宇豪, 等. 高光谱图像结合机器学习方法无损检测猕猴桃[J]. *中国农机化学报*, 2019, 40(4): 71–77.
HUO YQ, ZHANG C, LI YH, *et al.* Nondestructive detection for kiwifruit based on the hyperspectral technology and machine learning [J]. *Chinese Journal of Agricultural Mechanization*, 2019, 40(4): 71–77.
- [24] 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.
LI B, WAN X, LIU AIL, *et al.* Optimal position for suger content detection of Yongquan honey oranges based on hyperspectral imaging technology [J]. *China Optics*, 2024, 17(1): 128–139.
- [25] 尚静, 黄人帅, 张艳, 等. 高光谱成像结合模式识别无损检测猕猴桃成熟度[J]. *中国农机化学报*, 2022, 43(8): 90–95.
SHANG J, HUANG RS, ZHANG Y, *et al.* Nondestructive detection for maturity of kiwifruit based on hyperspectral imaging and pattern recognition [J]. *Journal of Chinese Agricultural Mechanization*, 2022, 43(8): 90–95.
- [26] 董朋欣, 董安国, 李楚婷, 等. 基于全卷积网络和自编码的高光谱图像分类[J]. *计算机工程与应用*, 2022, 58(5): 256–263.
DONG PX, DONG ANG, LI CT, *et al.* Hyperspectral image classification based on fully convolutional network and auto-encoder [J]. *Computer Engineering and Applications*, 2022, 58(5): 256–263.
- [27] 杜馨, 孙晓荣, 刘翠玲, 等. 拉曼光谱结合偏最小二乘法对食用油品质快速检测研究[J]. *中国酿造*, 2019, 38(12): 171–174.
- DU X, SUN XR, LIU CL, *et al.* Rapid detection of edible oil quality by Raman spectroscopy combined with partial least squares [J]. *China Brewing*, 2019, 38(12): 171–174.
- [28] 刘燕德, 周延睿, 潘圆媛. 基于最小二乘支持向量机的辣椒可溶性固形物和维生素 C 含量近红外光谱检测[J]. *光学精密工程*, 2014, 22(2): 281–288.
LIU YD, ZHOU YR, PAN YY. Determination of soluble solid and contents and vitamin C of fresh peppers based on NIR spectrometry and least square support vector machine [J]. *Optics and Precision Engineering*, 2014, 22(2): 281–288.
- [29] 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.
CHANG HJ, MENG QH, WU ZF, *et al.* Detection of soluble solids content in mango based on backpropagation algorithm neural network and hyperspectral imaging [J]. *Journal of Food Safety & Quality*, 2024, 15(2): 141–148.
- [30] 王立国, 赵亮, 刘丹凤. SVM 在高光谱图像处理中的应用综述[J]. *哈尔滨工程大学学报*, 2018, 39(6): 973–983.
WANG LG, ZHAO L, LIU DF. A review on the application of SVM in hyperspectral image processing [J]. *Journal of Harbin Engineering University*, 2018, 39(6): 973–983.

(责任编辑: 于梦娇 安香玉)