

基于LASSO-LSSVM算法赋能拉曼光谱技术 鉴别葡萄籽油的掺假行为

龚泽华^{1*}, 田保华², 薛艳茹³, 周尉⁴

(1. 山西开放大学乡村振兴学院, 山西太原 030027) (2. 太原理工大学生态学院, 山西太原 030024)
(3. 太原理工大学生物医学工程学院, 山西太原 030024) (4. 山西大学计算机与信息技术学院, 山西太原 030006)

摘要: 针对葡萄籽油掺假的问题, 研究了葵花籽油掺伪葡萄籽油掺假的定量分析的技术方法。采用便携式近红外拉曼光谱技术结合最小绝对收缩选择算子和最小二乘支持向量机 (Least Absolute Shrinkage and Selection Operator-Least Squares Support Vector Machine, LASSO-LSSVM) 算法开展了不同掺伪体积分数的模拟实验。实验中共制备了11种不同掺伪体积分数的混合油样, 通过便携式785拉曼光谱仪分别采集所有油样的拉曼光谱, 采用小波算法对原始光谱数据进行基线校正、降噪和归一化处理, 引入机器学习算法分别提取拉曼光谱的特征向量并建立量化分析模型, LASSO算法实现光谱数据降维和特征提取, LSSVM算法则构建掺伪量化的分析模型, 模型测试集的决定系数 R^2 为0.9846, RMSE为0.0249。基于LASSO-LSSVM算法实现了赋能拉曼光谱技术鉴别葡萄籽油的掺假量, 该方法对促进国内葡萄籽油消费市场具有重要的应用价值和商业潜力。

关键词: 拉曼光谱; 最小绝对收缩选择算子; 最小二乘支持向量机; 葡萄籽油; 葵花籽油; 定量模型

文章编号: 1673-9078(2025)07-12-19

DOI: 10.13982/j.mfst.1673-9078.2025.7.0587

Detection of Adulteration of Grape Seed Oil Using Raman Spectroscopy Coupled with LASSO-LS-SVM Algorithm

GONG Zehua^{1*}, TIAN Baohua², XUE Yanru³, ZHOU Wei⁴

(1. Rural Revitalization Institute, Shanxi Open University, Taiyuan 030027, China)

(2. School of Ecology, Taiyuan University of Technology, Taiyuan 030024, China)

(3. College of Biomedical Engineering, Taiyuan University of Technology, Taiyuan 030024, China)

(4. School of Computer & Information Technology, Shanxi University, Taiyuan 030006, China)

Abstract: To address the issue of adulteration in grape seed oil, a quantitative analysis method for detecting adulteration of grape seed oil with sunflower seed oil was studied. Simulation experiments were conducted by applying portable near-infrared Raman spectroscopy combined with a least absolute shrinkage and selection operator-least squares support vector machine (LASSO-LS-SVM) algorithm to samples with varying adulteration levels. A total of 11 mixed oil samples with varying adulteration concentrations were prepared in this research. A portable 785 Raman spectrometer was

引文格式:

龚泽华, 田保华, 薛艳茹, 等. 基于LASSO-LSSVM算法赋能拉曼光谱技术鉴别葡萄籽油的掺假行为[J]. 现代食品科技, 2025, 41(7): 12-19.

GONG Zehua, TIAN Baohua, XUE Yanru, et al. Detection of adulteration of grape seed oil using raman spectroscopy coupled with LASSO-LS-SVM algorithm [J]. Modern Food Science and Technology, 2025, 41(7): 12-19.

收稿日期: 2024-04-29

基金项目: 国家自然科学基金项目 (62006146)

作者简介: 龚泽华 (1986-), 女, 博士, 讲师, 研究方向: 作物学, E-mail: gongzehua1986@126.com

utilized to acquire the Raman spectra of all mixed oil samples. A wavelet algorithm was adopted for the baseline correction, noise reduction, and normalization of the raw spectral data. A machine learning algorithm was used to extract the eigenvectors of the Raman spectra, and a quantitative analysis model was established. The LASSO algorithm reduced the dimensionality and extracted features of the spectral data, and the LSSVM algorithm enabled successful construction of a quantitative analysis model for adulteration. The coefficient of determination (R^2) of the model for the test set is 0.984 6, and the RMSE is 0.024 9. The integration of the LASSO-LS-SVM algorithm enables the application of Raman spectroscopy for quantifying adulteration levels in grape seed oil. The proposed method demonstrates substantial application value and promising commercial potential for promoting the domestic grape seed oil market.

Key words: Raman spectroscopy; least absolute shrinkage and selection operator; least squares support vector machine; grape seed oil; sunflower seed oil; quantitative model

葡萄籽油是从葡萄籽中提取的一种植物油，其含有丰富的抗氧化剂，可保护细胞免受自由基的损害。为了防止植物油市场出现大规模的掺假造假行为，对葡萄籽油的掺伪定量分析具有重要的价值和意义，而实现快速、无损伤、准确地检测要求是目前发展的重要方向。

传统的植物油检测技术包括色谱法^[1]、酸碱滴定法^[2]及紫外吸收法^[3]，其中大部分方法已形成国家标准，但这些手段很难实现快速、无损的检测需求，且上述方法通常需要复杂的样品前处理，对操作人员的专业具有一定的要求，部分设备维护和管理费也是制约其大面积推广的关键因素。激光光谱技术不仅具备快速、无损的检测优势，而且对设备维护和样品前处理要求较低，对操作人员的专业能力要求也较低，因此激光光谱技术是当前进行物质成分定性定量分析的重要研究方向和热点领域^[4,5]。激光光谱技术包含众多分支，其中激光诱导击穿光谱等属于原子光谱^[6]，而拉曼光谱、红外吸收光谱、布里渊散射光谱及荧光光谱等属于分子光谱^[7-10]。这众多激光光谱技术中拉曼光谱技术是目前小型化和商业化最具代表性的技术之一，而近红外拉曼光谱技术在小型化和仪器性能表现中都表现出突出优势。

葡萄籽油掺假的拉曼光谱研究存在许多空白，尤其是针对廉价植物油掺假葡萄籽油的量化分析。万恒兴等^[11]将廉价植物油掺入葡萄籽油中并开展了基于偏最小二乘算法定量分析，实现了葵花籽油掺假葡萄籽油测试集的浓度预测，模型决定系数为 0.975 5。唐云峰等^[12]研究设计一套低成本、高准确度的可见 / 近红外光谱仪检测系统来实现葡萄籽油品质掺假鉴别，检测系统操作简单、快速，不仅对已知掺假类别能够 100% 识别，而且对新掺假类别能够实现定性判别。张莉等^[13]采用傅里叶红外光谱

技术建立了葡萄籽油红外光谱的标准方法，发现葡萄籽油、橄榄油和杏仁油在 $1\ 120\ \text{cm}^{-1}$ 波数附近的吸收峰相对强弱存在明显差异。拉曼光谱量化分析的准确性受仪器的稳定性和定量分析模型的复杂度等因素的影响。尤其是便携式仪器的稳定性不及台式系统的性能表现，这时对定量分析模型的要求更高。此外激励源的波长同样会影响拉曼光谱信号，因为存在荧光干扰，选着合适的激励源波长是提高定量分析精度的关键。传统的定量分析模型包括多元线性回归模型^[14]、偏最小二乘回归模型^[15]及支持向量回归模型^[7]等。李明亮等^[16]比较了机器学习回归模型在开展金属合金材料中铜元素的定量分析，发现偏最小二乘回归模型的预测效果最优。除此之外，Lv 等^[17]提出采用线性回归与稀疏欠完全自动编码器相结合的特征提取方法，实现了对陶瓷高维光谱数据的非线性特征提取和降维，在减少输入数据维数和冗余信息的同时，提高了定量分析性能。

总之，将便携式拉曼光谱技术应用于定量分析葡萄籽油掺假研究对于维护消费者合法权益具有重要的价值。本文采用近红外便携式拉曼光谱技术分析葡萄籽油与葵花籽油的光谱差异，并将最小绝对收缩和选择算子（Least Absolute Shrinkage and Selection Operator, LASSO）与最小二乘支持向量机（Least Squares Support Vector Machine, LSSVM）模型相结合，建立了葵花籽油掺假葡萄籽油的定量回归模型，通过定量模型评价参数评价模型的稳定性和准确性。

1 材料与方法

1.1 原料及设备

1.1.1 实验原料

实验原料为纯植物油样本，包含了葡萄籽油和

葵花籽油两类,按照表1葡萄籽油和葵花籽油体积掺伪体积分数配料表配置混合油样,其中每种混合油样各独立配置50个样本。混合油样经过人工充分搅拌后,静置放置在阴凉通风处,存放时间大于15 d。

表1 葡萄籽油和葵花籽油体积掺伪体积分数配料表
Table 1 Proportioning table of volume adulteration concentration of grape seed oil and sunflower seed oil

编号	葡萄籽油体积/mL	葵花籽油体积/mL	数量
1	1.0	0.0	50
2	0.9	0.1	50
3	0.8	0.2	50
4	0.7	0.3	50
5	0.6	0.4	50
6	0.5	0.5	50
7	0.4	0.6	50
8	0.3	0.7	50
9	0.2	0.8	50
10	0.1	0.9	50
11	0.0	1.0	50

1.1.2 仪器设备

实验设备为长春新产业生产的近红外便携式拉曼光谱仪(RamanSys-785),仪器激励源波长为785 nm,线宽为3.2 cm⁻¹,能量最大为500 mW(可调),标配拉曼探头的工作距离为7.5 mm。光谱采集范围为300~3 500 cm⁻¹,信噪比为1 000:1,积分时间为18 ms~10 s(可调)。实验测量参数为拉曼探头出射的激光能量为69.5 mW,激光焦点位置侵入油样2 mm,光谱积分时间为800 ms,每次采集的光谱作为特征信息保存一条,每个混合油样各随机采集了2条光谱数据,因此每种混合油样获取了100条光谱数据,共计1 100条光谱数据。所有的光谱采集实验在恒温恒湿的实验室暗环境下开展,激光初次照射样品时,原始光谱会存在一定的抖动,因此,仪器实际采集光谱数据时需带光谱稳定后开展数据保存的操作。

1.2 算法及评价参数

本文主要使用的方法学主要涉及到光谱预处理算法和定量分析算法。光谱预处理算法为小波算法,定量分析算法为最小绝对收缩和选择算子与最小二乘支持向量机的组合算法。

1.2.1 小波算法

在拉曼光谱数据处理中去基线和降噪处理是最基本的预处理过程。去基线是从原始信号中去除基线或趋势成分,在拉曼光谱中荧光和背景光等宽谱信息常被视为基线信息。降噪是指从信号中去除噪声或干扰,以提取出信号的有效成分,在拉曼光谱中降噪处理可以提高拉曼光谱特征谱峰的信噪比和对比度,利于光谱的定性定量分析。小波算法作为一种有效的信号分析和处理工具,在去基线和降噪中发挥着重要的作用^[18]。小波算法去基线和降噪的原理是通过小波变换将原始信号分解为不同频率的小波系数,然后去除包含基线或趋势成分的低频小波系数,其次对小波系数进行阈值处理,去除包含噪声或干扰的高频小波系数,最后进行逆小波变换重构去基线和降噪后的信号。因此该过程包括了小波变换、系数选择和去除、阈值处理和逆小波变换等过程,选择合适的小波基函数和阈值处理方法对去基线的效果具有重要影响。不同的小波基函数适用于不同类型的信号,而阈值处理方法可以控制去除低频小波系数的程度,从而影响去基线后的信号质量。

1.2.2 LASSO-LSSVM算法

最小绝对收缩和选择算子(LASSO)是一种压缩估计方法,其使用L1正则化,通过对系数进行惩罚实现稀疏性,而无关冗余信息的权重变为0,LASSO的损失函数表达式为:

$$\min_{\omega} \left\{ \frac{1}{2} \|X_{\omega} - y\|_2^2 + \alpha \|\omega\|_1 \right\} \quad (1)$$

式中:

α ——正则化系数;

ω ——回归系数。

最小二乘支持向量机(LSSVM)是在支持向量机算法的基础上进一步改进的机器学习算法,因其改进后的算法在处理非线性问题上具有了一定的优势,LSSVM算法在处理数据分析时选用的核函数为径向基函数,其表达式为:

$$K(x_i, x_j) = \exp \left(-\frac{\|x - x_k\|^2}{2\sigma^2} \right) \quad (2)$$

式中:

σ ——核参数。

总之,LSSVM回归模型^[19]可表达为:

$$f(x) = \sum_{i=1}^n a_i k(x_i, x_j) + b \quad (3)$$

LASSO-LSSVM回归模型非常适用于处理复

杂、不确定性的高维光谱数据,可有效减少输入数据的维数和冗余信息,因此,LASSO-LSSVM 该模型减少了 LSSVM 的过拟合问题^[19]。

1.2.3 模型预测效果评价参数

通常采用决定系数 (Coefficient of Determination, R^2) 和均方根误差 (Root Mean Square Error, RMSE) 作为模型预测效果评价参数,其具体表达为:

$$R^2 = 1 - \frac{\sum_{i=1}^n (f_i - \hat{f}_i)^2}{\sum_{i=1}^n (f_i - \bar{f}_i)^2} \quad (4)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (\hat{f}_i - f_i)^2}{n}} \quad (5)$$

式中:

f_i ——真实值;

$\hat{f}_i$ ——模型预测值;

$\bar{f}_i$ ——平均值;

n ——总数量。

1.2.4 数据分析

该研究采用统计学方法—皮尔逊相关系数研究光谱特征谱峰峰强与掺伪浓度的统计学规律,皮尔逊相关系数是一种衡量两个变量之间线性关系强度和方向的统计量。它的取值范围在 -1 到 1 之间,其中 -1 表示完全负相关,0 表示无相关,1 表示完全正相关。皮尔逊相关系数可以用来衡量两个变量之间的线性关系强度。当两个变量之间存在线性关系时,相关系数的绝对值趋近于 1;当两个变量之间不存在线性关系时,相关系数的绝对值趋近于 0。皮尔逊相关系数可以用来判断两个变量之间的关系是否显著。一般来说,当相关系数的绝对值大于 0.5 时,可认为两个变量之间存在显著的线性关系。

2 结果与讨论

2.1 拉曼光谱分析

2.1.1 拉曼光谱预处理

图 1 为经过小波算法处理的葡萄籽油和葵花籽油的拉曼光谱预处理过程,小波算法实现了对两种纯植物油拉曼光谱的基线去除和降噪,图 1a 为葡萄籽油拉曼光谱的基线校正和光谱降噪过程,图 1b 为葵花籽油拉曼光谱的基线校正和光谱降噪过程,图 1c 为小波算法预处理后强度累加和归一化后葡萄籽油、葵花籽油的拉曼光谱信号,其中红色曲线为葡萄籽油的拉曼光谱,黑色曲线为葵花籽油的拉曼

光谱。图 1c 表明预处理后的光谱只包含了不同分子振动模式产生的拉曼光谱信号,且两种纯植物油的拉曼光谱特征谱峰信息存在很大的差异,这种光谱差异以拉曼光谱特征峰强的差异性为主,或解释为不同拉曼光谱特征峰的比值存在显著的差异性。这种差异性通常可以反映掺伪体积分数梯度变化。

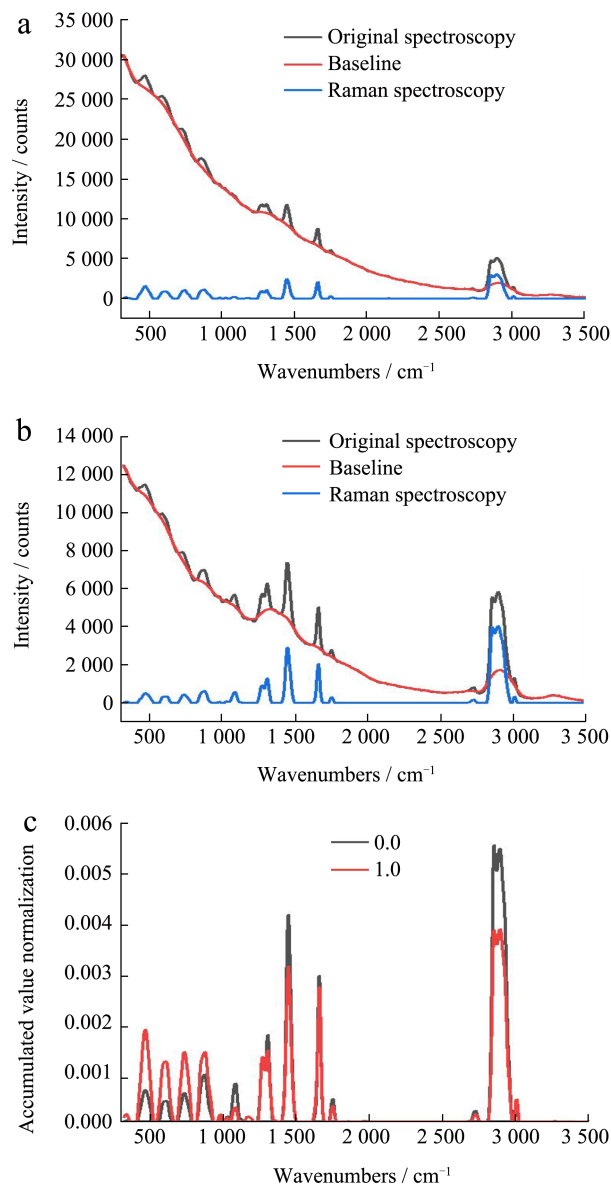


图 1 拉曼光谱的预处理过程

Fig.1 Pretreatment of Raman spectra

注: (a) 葡萄籽油; (b) 葵花籽油; (c) 归一化结果。

2.1.2 葡萄籽油不同掺伪体积分数的拉曼光谱

为了进一步验证并对比不同掺伪体积分数引起的拉曼光谱特征峰强的差异,本文绘制了 11 种混合油样的拉曼光谱,如图 2 所示。实验样本制备环节每种混合油样各单独制备了 50 个,而在光谱采集环节每个混合油样各随机采集了 2 条光谱,因此

每种掺伪体积分数的混合油样共有 100 条拉曼光谱数据, 取这 100 条拉曼光谱的平均光谱作为每种掺伪体积分数混合油样的代表性拉曼光谱, 可得到图 2 的 11 条代表性拉曼光谱曲线。图 2 表明掺伪体积分数变化引起的拉曼光谱差异主要体现在两种纯植物油典型拉曼光谱的特征峰强差异。

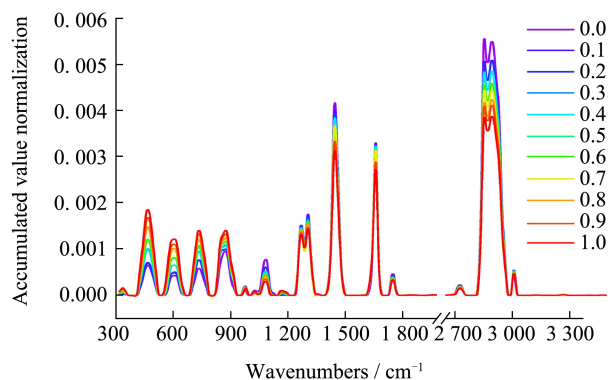


图 2 11 种混合油样的代表性拉曼光谱

Fig.2 Representative Raman spectra of 11 mixed oils

2.1.3 植物油拉曼光谱分析

拉曼光谱是反映物质成分分子层面的化学信息, 因此拉曼光谱特征峰代表了某种或某类化学物质的官能团信息, 基于拉曼光谱技术结合机器学习算法实现葡萄籽油掺伪的定量分析有必要开展植物油拉曼光谱特征峰的光谱信息解译。拉曼光谱分析主要为特征峰的归属判定, 图 3 为 11 种混合油样的代表性拉曼光谱, 其包含了 12 个离散的拉曼光谱特征峰和 1 个叠加的拉曼光谱特征带(众多拉曼光谱特征峰的叠加区域)。

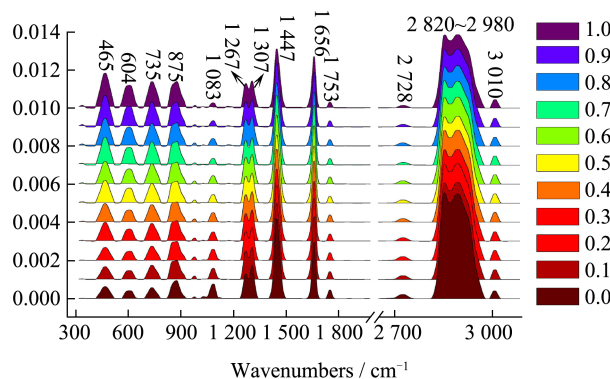


图 3 11 种混合油样的代表性拉曼光谱的特征峰位

Fig.3 Characteristic peak positions of representative Raman spectra of 11 mixed oils

低波数的拉曼光谱特征峰位脂肪酸碳长链产生的振动光谱, 其官能团为 $-(CH_2)_n-$, 这些官能团产生了 465、604、735、875 以及 1083 cm^{-1} 附近的拉曼光谱特征峰^[20-22]。1267 和 1656 cm^{-1} 附近的拉

曼光谱特征峰由不饱和碳双键的振动模式产生^[20-22]。1307、1447、3010 以及 2820~2980 cm^{-1} 的拉曼光谱特征峰主要由脂肪酸分子的甲基或亚甲基的 C-H 分子键产生^[20-22]。根据对拉曼光谱特征峰的分析可以发现葡萄籽油和葵花籽油的拉曼光谱反映了两种植物油的脂肪酸的信息, 这表明两种纯植物油所含有的脂肪酸的种类及其含量上存在一定的不同, 且这种区别完美映射到拉曼光谱的特征谱峰信息上, 这也是机器学习算法实现定量分析的根本原因。

2.2 定量分析

2.2.1 统计学规律

本文在前一个小节中分析了植物油的拉曼光谱, 且拉曼光谱的特征谱峰反映了不同掺伪体积分数的信息, 但拉曼光谱特征峰的差异在反映掺伪体积分数上是否具有统计学规律需要进一步分析和讨论。因此, 本小节重点研究了 12 个离散的拉曼光谱特征峰和 1 个叠加的拉曼光谱特征带与掺伪体积分数的统计学规律, 其中 12 个离散的拉曼光谱特征峰为数据预处理后对应的峰强与体积分数的变化规律, 而拉曼光谱特征带为该区域内所有强度信息的平均值与体积分数的变化规律, 如图 4 所示。

图 4a 为拉曼光谱特征峰(带)对应的峰强与掺伪体积分数的变化规律, 可以发现不同拉曼光谱特征峰强变化与掺伪体积分数变化存在正相关或负相关的特性, 每个拉曼光谱特征峰的变化反映的是某类脂肪酸分子在整体脂肪酸含量中对应官能团分子键的增加或减少。图 4b 为拉曼光谱特征峰(带)对应的峰强与掺伪体积分数的皮尔逊相关系数, 可以发现皮尔逊相关系数不同反映了不同拉曼光谱特征峰强对应掺伪体积分数的相关性, 而且给出了具体的数值。

当皮尔逊相关系数的绝对值越接近于 1 时, 说明两变量的正(负)相关性越强, 当系数趋近于 0 时, 说明两变量不存在相关性, 结合图 4b 的柱状图可以发现位于 465、604、735 及 875 cm^{-1} 附近的拉曼光谱特征峰峰强与掺伪体积分数变化存在正相关性, 且相关系数大于 0.98; 位于 1083、1307、1447、1753、2728 以及 2820~2980 cm^{-1} 附近的拉曼光谱特征峰峰强与掺伪体积分数变化存在负相关性, 且相关系数小于 -0.95; 位于 1267、1656 以及 3010 cm^{-1} 附近的拉曼光谱特征峰峰强与掺伪体积分数变化存在负相关性, 且相关系数分别

为-0.47、-0.73和0.18。通过对拉曼光谱特征峰强信息与掺伪体积分数的统计学规律的研究表明拉曼光谱具有反映葡萄籽油掺伪体积分数信息的能力,且间接证明了本文制备的不同掺伪体积分数实验样品的可靠性(统计学规律受人为制备过程中引入的浓度误差的影响)。

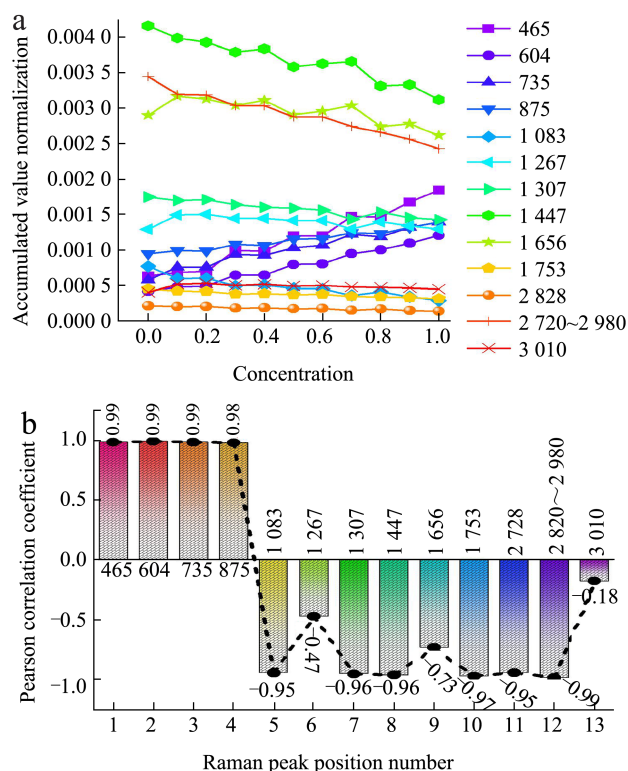


图4 拉曼光谱特征峰/带与掺伪体积分数的统计学规律

Fig.4 Statistical characteristics of characteristic peaks/bands and adulteration concentration of Raman spectra

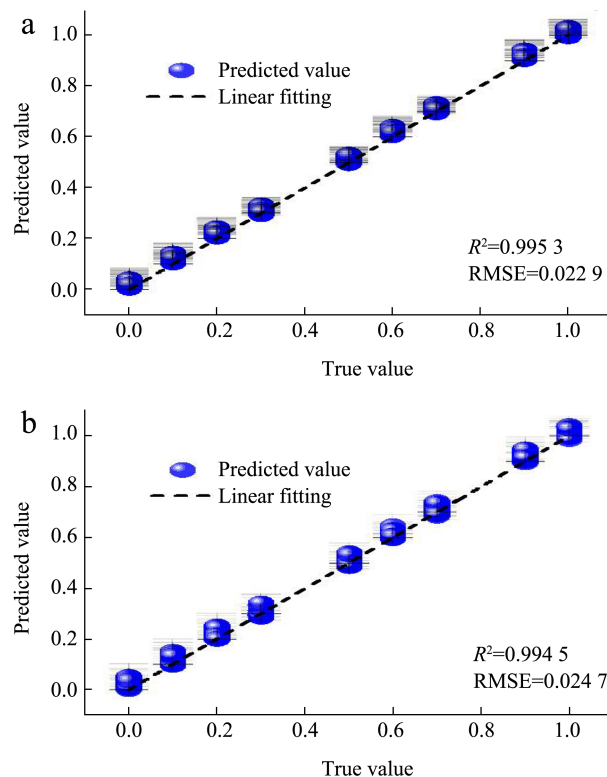
注:(a) 拉曼光谱特征峰/带与掺伪体积分数的点线图;
(b) 拉曼光谱特征峰/带与掺伪体积分数皮尔逊相关系数的柱状图。

2.2.2 LASSO-LSSVM算法模型

原始拉曼光谱在经过预处理后得到了去除基线、降噪和归一化的光谱数据,利用 LASSO 算法对光谱数据进行特征向量选择,选择均方根误差最小的 Alpha 值作为 LASSO 算法的最优参数,将所有的拉曼光谱数据输入 LASSO 模型中,选择具有非零 ω 系数的特征谱峰的信息,基于 LASSO 算法搜索的降维且有效的预测因子子集,因此,定量预测模型首先采用 LASSO 算法选择合理的特征向量信息。

将 1 100 条 LASSO 算法处理过的数据划分训练集、验证集和测试集,其中 11 种掺伪体积分数的

混合油样按照 9:2 的比例原则划分训练-验证集和测试集,掺伪体积分数为 0.4 和 0.8 的混合油样的 200 条数据作为模型的测试集,其余的混合油样的 900 条数据作为模型的训练-验证集。在划分训练集和验证集的过程中,按照 4:1 的原则在 900 条数据中随机选择 720 条数据作为训练集,剩余的 180 条数据用于验证集。总之,将 LASSO 算法选择的特征数据用于建立 LSSVM 算法模型,其中用于建立 LSSVM 模型的数据有 720 条,用于验证 LSSVM 模型的数据有 180 条,而最终验证 LSSVM 模型的数据为 200 条。需要说明一点的是 LSSVM 模型的训练集和验证集来自于相同的掺伪体积分数的混合油样,而用于测试 LSSVM 模型的测试集则未参与定量模型的训练,因此,测试集的预测效果直接可以反映 LSSVM 模型的泛化能力。图 5 为 LASSO-LSSVM 算法模型的预测效果,其中图 5a 为建立定量模型训练集的预测效果,图 5b 为验证定量模型验证集的预测效果,图 5c 为测试定量模型测试集的预测效果。图 5 表明 LASSO-LSSVM 算法模型在预测葵花籽油掺伪葡萄籽油的定量结果中表现出色,其中定量模型训练集的 R^2 为 0.995 3, RMSE 为 0.022 9;定量模型验证集的 R^2 为 0.994 5, RMSE 为 0.024 7;定量模型训练集的 R^2 为 0.984 6, RMSE 为 0.024 9。



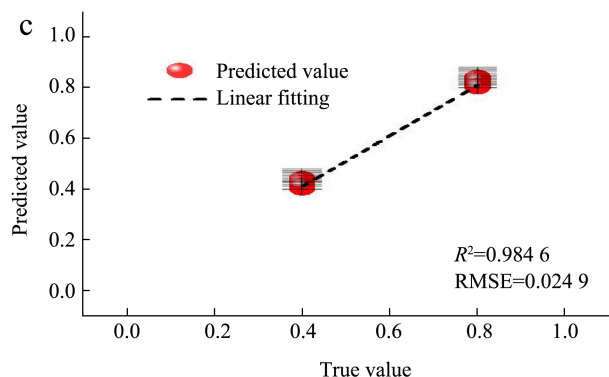


图 5 LASSO-LSSVM 算法模型的定量曲线

Fig.5 Quantitative curve of lasso-lssvm algorithm model

注: (a) 训练集; (b) 验证集; (c) 测试集。

总之, 根据图 5 表现出的 LASSO-LSSVM 算法模型的定量曲线可以确定基于 785 拉曼光谱和 LASSO-LSSVM 算法相结合的技术手段可以实现对葵花籽油掺伪葡萄籽油的掺伪体积分数预测, 且定量模型的鲁棒性和泛化能力均表现较为理想。

2.2.3 检测限分析

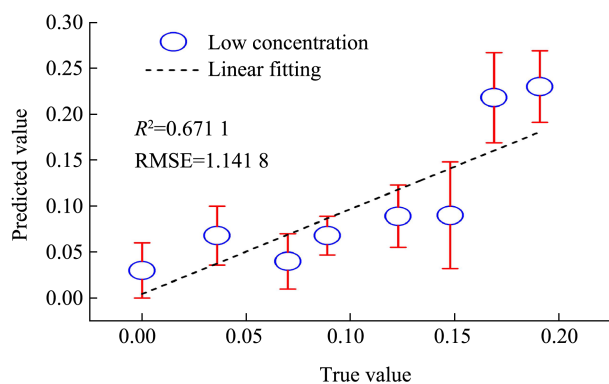


图 6 低浓度掺伪的预测能力

Fig.6 The predictive ability of low concentration adulteration

检测限问题是将拉曼光谱技术结合 LASSO-LSSVM 算法推广至商业化的重要考量, 为此本小节评估并讨论了低浓度掺伪的模拟实验, 单独制备了 8 种含有低葡萄籽油浓度的测试样本 (0.00、0.036、0.070、0.089、0.123、0.148、0.169 及 0.191), 且每种混合油样各随机采集 10 条光谱, 将 10 条光谱取平均并送入建立的 LASSO-LSSVM 模型中进行定量预测, 预测结果如图 6 所示。图 6 表明 LASSO-LSSVM 算法模型在低葡萄籽油浓度的预测效果较差, 这可能由两方面因素导致的结果, 一是建立定量模型的混合油样的掺伪体积分数梯度为 0.1, 二是配制低掺伪体积分数的混合油样时引入的人为误差较大。总之, 以上因素单独或同时存在导致的 LASSO-LSSVM 模型在低浓度掺伪预测效果差

的结果, 即模型的检测限不可能高于本小节测试样本的浓度梯度。

3 结论

为实现葵花籽油掺假葡萄籽油的定量预测, 本文提出一种基于 785 便携式拉曼光谱结合 LASSO-LSSVM 算法实现葡萄籽油掺假的定量预测, 其中定量模型测试集的评价参数 R^2 为 0.984 6, RMSE 为 0.024 9。近红外拉曼光谱技术具有表征葡萄籽油和葵花籽油脂肪酸成分化学信息的能力, 且拉曼光谱特征峰强度变化反映了掺伪体积分数变化, 部分特征峰强与掺伪体积分数两个变量存在极强的相关性。LASSO 算法剔除了拉曼光谱中冗余的无用信息, LSSVM 算法则基于 LASSO 算法提取的特征向量建立了葵花籽油掺伪葡萄籽油的量化分析模型。总之, 本文提出的技术方法具有具有较好的性能表现和潜在的商业潜力, 此外本文的研究结果对规范国内食用油消费市场, 促进葡萄籽油消费的健康发展具有一定的研究价值。

参考文献

- [1] 胡昆,张成龙,杨瑞琴.基于气相色谱法结合化学计量学识别微量植物油的比较研究[J].中国油脂,2025,50(2): 123-127.
- [2] 马葱,刘钟栋.酸碱滴定法在植物油质量检测中应用的改进[J].中国食品添加剂,2018,2:188-195.
- [3] 龙伶俐,薛雅琳,张蕊,等.利用紫外吸收光谱研究植物油精炼程度[J].农业机械,2012,24:61-63.
- [4] 蒋雪松,莫欣欣,孙通,等.食用植物油中反式脂肪酸含量的激光拉曼光谱检测[J].光谱学与光谱分析,2019, 39(12):3821-3825.
- [5] 张凤娟,黄敏.基于激光拉曼光谱的橄榄油掺伪定量检测[J].应用激光,2023,43(10):164-169.
- [6] 吴宜青,刘津,莫欣欣,等.共轴双脉冲激光诱导击穿光谱结合双谱线内标法定量分析植物油中的铬[J].物理学报, 2017,66(5):127-135.
- [7] 彭楠,方俊,毛潭.基于PCA-SVM结合共聚焦拉曼光谱的特级初榨橄榄油掺伪压榨菜籽油定量分析[J].中国油脂,2024,49(2):70-74.
- [8] 杨丽,王珂,徐新忠,等.基于近红外光谱法的哈萨克斯坦亚麻籽油品质指标定量预测模型的建立[J].现代食品科技,2023,39(11):302-309.
- [9] 曹晨鹏,郝仕国,罗宁宁,等.基于受激布里渊散射及紫外-可见吸收光谱的橄榄油等级鉴别与分析[J].中国激光, 2018,45(9):165-170.
- [10] 王九玲,罗文,李文凯.反向传播神经网络算法结合拉曼荧光光谱法定量检测特级初榨橄榄油掺假[J].食品安全

- 质量检测学报,2023,14(22):126-133.
- [11] 万恒兴,冯丽雄,余展旺.基于拉曼光谱的葡萄籽油掺伪定量检测研究[J].山东化工,2021,50(20):68-74.
- [12] 唐云峰,柴琴琴,林双杰,等.可见/近红外光谱的葡萄籽油掺伪检测系统[J].光谱学与光谱分析,2020,40(1):202-208.
- [13] 张莉,魏冬梅,王华.FTIR法建立葡萄籽油的红外光谱特征[J].中国酿造,2013,32(7):121-122.
- [14] WANG Z, AFGAN MUHAMMAD SHER, GU W, et al. Recent advances in laser-induced breakdown spectroscopy quantification: from fundamental understanding to data processing [J]. TrAC Trends in Analytical Chemistry, 2021, 143: 116385.
- [15] 赵雅惠,廖鹏飞,陈璐,等.GC-MS/GC-O-MS结合PLSR分析新型青稞白酒的关键香气物质[J].食品工业科技, 2024,45(19):228-239.
- [16] 李明亮,戴宇佳,秦爽,等.LIBS分析模型对铝合金定量分析精度的影响[J].光谱学与光谱分析,2022,42(2):587-591.
- [17] LV Z J, YU H X, SUN L X, et al. Composition analysis of ceramic raw materials using laser-induced breakdown spectroscopy and autoencoder neural network [J]. Analytical Methods, 2022, 14: 1320-1328.
- [18] GALLOWAY C M, RU E L, et al. An iterative algorithm for background removal in spectroscopy by wavelet transforms [J]. Applied Spectroscopy, 2009, 63(12): 1370-1376.
- [19] 戴宇佳,高勋,刘子源.LASSO-LSSVM与激光诱导击穿光谱技术结合提高铝合金中Mn成分检测精度研究[J].光谱学与光谱分析,2024,44(4):977-982.
- [20] FAN D S, HUANG W, LIU T C, et al. Quantitative analysis of blended oils by confocal Raman spectroscopy and chemometrics in situ [J]. Food Control, 2022, 142: 109244.
- [21] WU X J, ZHANG X, DU Z R, et al. Raman spectroscopy combined with multiple one-dimensional deep learning models for simultaneous quantification of multiple components in blended olive oil [J]. Food Chemistry, 2024, 431: 137109.
- [22] ROCIO R R, JOSE Á S D, CLEMENTE O R, et al. A comparative study of fluorescence and Raman spectroscopy for discrimination of virgin olive oil categories: Chemometric approaches and evaluation against other techniques [J]. Food Control, 2024, 158: 110250.