

基于反射光谱的油茶籽油掺伪量快速测定 及特征波长特性研究

刘 强, 龚中良, 李大鹏, 文 韬, 汪志强, 管金伟, 郑文峰

(中南林业科技大学 机电工程学院, 长沙 410004)

摘要:为了探索紫外-可见-近红外反射光谱测定油茶籽油掺伪量的方法,按照不同掺伪比例制备了244个油茶籽油掺伪大豆油、菜籽油、花生油、玉米油的样本,以自主搭建的实验平台采集所制备样本在200~1100 nm范围内的反射光谱。将原始光谱进行Savitzky-Golay(SG)-连续小波变换(CWT)预处理后,利用Kennard-Stone(K-S)算法以2:1的比例将样本划分成校正集和预测集。采用竞争性自适应重加权算法(CARS)、连续投影算法(SPA)、自主软收缩算法(BOSS)、迭代变量子集优化算法(IVSO)进行特征波长选择,分别建立基于支持向量机(SVM)、极限学习机(ELM)、随机森林(RF)的油茶籽油掺伪量快速预测模型,同时对特征波长的特性进行了研究。结果表明:原始光谱经过SG-CWT(L5)预处理和BOSS特征波长筛选后,建立的基于SVM的油茶籽油掺伪量快速预测模型能够鉴别掺伪量为1%及以上的油茶籽油,该模型在十折交叉验证和网格搜索法下得到最佳惩罚因子(c)和核函数(γ)分别为5.278 0和0.108 8,其预测决定系数(R_p^2)、预测均方根误差($RMSE_p$)、预测平均绝对误差(MAE_p)分别为0.998 5、0.013 4、0.010 2。特征波长聚集程度和陡度对模型预测结果存在一定影响。综上,建立的基于反射光谱的油茶籽油掺伪量快速预测模型预测误差小,预测效果较好。

关键词:油茶籽油;紫外-可见-近红外反射光谱;反射率;BOSS-SVM;特征波长特性

中图分类号:TS225.1;TS227 文献标识码:A 文章编号:1003-7969(2024)03-0132-08

Rapid prediction of oil-tea camellia seed oil adulteration amount based on reflection spectroscopy and characteristic wavelength characteristics

LIU Qiang, GONG Zhongliang, LI Dapeng, WEN Tao,
WANG Zhiqiang, GUAN Jinwei, ZHENG Wenfeng

(College of Electrical and Mechanical Engineering, Central South University of Forestry Technology, Changsha 410004, China)

Abstract: In order to explore the method of UV-Vis-NIR reflection spectroscopy to identify blended oil-tea camellia seed oil (CAO), 244 samples of CAO adulterated with soybean oil, rapeseed oil, peanut oil and corn oil were prepared according to different adulteration amounts, and the reflectance spectra of the prepared samples in the range of 200-1100 nm were collected by an experimental platform built independently. After pretreating the raw spectra with SG-continuous wavelet transform (CWT), the

samples were divided into correction and prediction sets using the Kennard-Stone (K-S) algorithm in a ratio of 2:1. Competitive adaptive reweighting sampling (CARS) algorithm, successive projections algorithm (SPA), bootstrapping soft shrinkage (BOSS) algorithm, and iteratively variable subset optimization (IVSO) algorithm were used for characteristic

收稿日期:2022-11-26;修回日期:2023-11-02

基金项目:湖南省科技计划重点研发项目(2022NK2048);湖南省教育厅科学项目(18B192, 20A515);湖南省自然科学基金(2020JJ4142);湖南省林业杰青培养科研项目(XLK202108-7)

作者简介:刘 强(1995),男,硕士研究生,研究方向为农业装备自动化(E-mail)1340320841@qq.com。

通信作者:李大鹏,讲师,博士(E-mail)dapengli@csuft.edu.cn。

wavelength selection, and rapid identification models based on support vector machine (SVM), extreme learning machine (ELM), and random forest (RF) were established for CAO adulteration amount, respectively, and the characteristics of characteristic wavelength were studied. The results showed that the SVM model established after the SG - CWT (L5) pretreating and BOSS characteristic wavelength screening could discriminate the amount of adulteration 1% and above, and the model obtained the best penalty factor c (5.278 0) and kernel function γ (0.108 8) under the ten - fold cross - validation and grid search method, with R_p^2 , $RMSE_p$ and MAE_p of 0.998 5, 0.013 4 and 0.010 2, respectively. At the same time, the degree of aggregation and steepness of the characteristic wavelength had some influence on the model prediction results. In conclusion, the established rapid prediction model for the adulteration amount of oil - tea camellia seed oil based on reflection spectroscopy has low error and good prediction effect.

Key words: oil - tea camellia seed oil; UV - Vis - NIR reflection spectroscopy; reflectance; BOSS - SVM; characteristic wavelength characteristics

油茶籽油具有优越的理化性能,且营养丰富,长期食用有助于人体健康^[1-2]。油茶籽油价格为普通植物油的5~10倍^[3-4],面对悬殊的价格差异,有不良商贩对油茶籽油掺伪,或虚假标注调和油中油茶籽油的配比^[1],严重破坏了市场的公平性。因此,研究快速、有效的油茶籽油掺伪量测定方法至关重要。

食用油掺伪鉴别的方法主要有色谱法^[5]、核磁共振法^[6]、电子鼻法^[7]、光谱法^[8]等。色谱法往往需要化学药剂的前处理,造成样本污染;核磁共振法所使用设备价格昂贵且操作复杂;电子鼻法的稳定性有待研究^[9];而光谱法具有直接、无污染的优点^[10]。目前基于光谱法的食用油掺伪鉴别研究多以近红外透射光谱(1 000~2 500 nm)获取吸光度进行建模预测^[4,11]。相比传统油脂透射光谱采集计算吸光度的方法,反射光谱具有波长范围宽、扫描时间短、采集更加便捷等优点^[12]。如:翁士状等^[12]测定了菜籽油、大豆油、葵花籽油、玉米油、橄榄油、芝麻油及花生油在350~2 500 nm范围内的反射光谱,实现了对食用油中饱和脂肪酸含量的快速检测;Su等^[13]通过自主搭建的反射光谱采集平台在350~2 500 nm范围内采集110个掺伪大豆油的芝麻油、菜籽油样本的反射光谱并计算反射率,得到偏最小二乘回归-多元散射校正-竞争性自适应重加权算法(PLSR - MSC - CARS)为芝麻油掺伪大豆油和菜籽油掺伪大豆油的最优预测模型,但未对掺伪量10%以下做出研究。另外,现有的油茶籽油鉴别多以价格昂贵的石英比色皿完成,成本较高。因此,本研究以低成本玻璃杯代替石英比色皿,并采用自主搭建的实验平台采集紫外-可见-近红外反射光谱(200~1 100 nm),首先对反射光谱进行了

Saritzky - Golay (SG) - 连续小波变换 (CWT) 预处理,然后以CARS、连续投影算法 (SPA)、自主软收缩算法 (BOSS)、迭代变量子集优化算法 (IVSO) 4种方法从全光谱 (FS) 中提取特征波长,最后分别建立了基于支持向量机 (SVM)、极限学习机 (ELM)、随机森林 (RF) 的油茶籽油掺伪量快速预测模型,以期开发基于反射光谱的油茶籽油掺伪检测装置奠定基础。

1 材料与方法

1.1 实验材料

1.1.1 实验原料

油茶籽油(大三湘,各项指标符合GB/T 11765—2018《油茶籽油》标准要求)、大豆油(金龙鱼)、菜籽油(金龙鱼)、玉米油(金龙鱼)、花生油(鲁花),均在保质期内,从长沙大型超市购买。

1.1.2 实验装置

自主搭建的反射光谱采集平台由暗箱、玻璃杯固定槽、可转动探头固定支架、Y型光纤、Maya2000 Pro光谱仪(美国OceanView公司)、HL1000卤钨灯光源(上海闻奕光电科技有限公司)、Ocean View光谱采集软件(美国OceanView公司)、玻璃杯组成,如图1所示。首先Y型光纤分叉端的两个接口分别连接HL1000卤钨灯光源和Maya2000 Pro光谱仪,Y型光纤的采集端固定在暗箱中(距离玻璃杯液面约5 mm)。然后通过光纤将光源装置发射的光源输送至采集端位置处,并将光源射入玻璃杯的油液内。所发射光源在油液中发生反射现象,一部分光源会被玻璃杯、空气、油液中的杂质和气泡等吸收从而改变传播方向,另一部分光源会继续传播。最后有效光源在经过反射后重新被Y型光纤固定采

集端采集,并通过光纤传送至光谱仪接收端,从而完成油茶籽油掺伪样本反射光谱的采集。

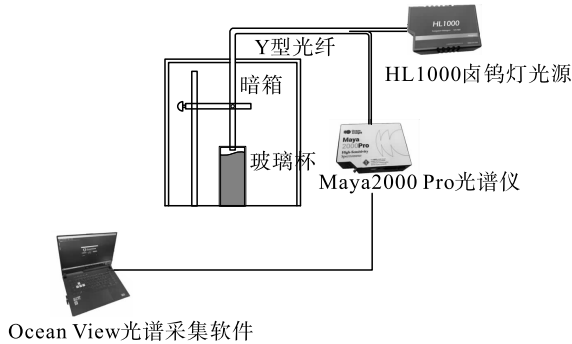


图1 反射光谱采集平台

1.2 实验方法

1.2.1 掺伪样品的制备

按照掺伪比例 0%、1%、3%、5%、7%、10%、15%、20%、30%、40%、50%、60%、70%、80%、90%、100% 将大豆油、菜籽油、玉米油、花生油分别混入油茶籽油中(0%指纯油茶籽油样品,100%指纯掺伪油样品)。为确保所制备样品的独立性,以4个组别分批制备样本,每组包括14个掺伪大豆油样品、掺伪菜籽油样品、掺伪玉米油样品、掺伪花生油样品和5个纯油(油茶籽油、大豆油、菜籽油、玉米油、花生油)样品,每组各61个油样,共计244个样品。制备时,将油样放入磁力搅拌机中,在35℃下搅拌1h,随后静置24h。

1.2.2 光谱采集

以自主搭建的反射光谱采集平台(图1)获取原始光谱数据,光谱扫描范围为200~1100nm,设置光谱积分时间为32ms,扫描次数为100次,滑动平均宽度为5,非线性矫正开启。实验前将光谱仪与光源设备预热40min。以10mL液体容量为标准,对244个掺伪样品进行光谱采集,同一样品采集3次后计算其平均值,并通过式(1)计算样品反射率。

$$T_{\lambda} = \frac{S_{\lambda} - D_{\lambda}}{R_{\lambda} - D_{\lambda}} \quad (1)$$

式中: T_{λ} 为反射率; S_{λ} 为采集样本光谱强度; D_{λ} 为暗光谱背景强度; R_{λ} 为空玻璃杯参考光谱强度。

1.2.3 光谱预处理与样本划分

光谱采集过程中容易产生杂散光、噪声等无用信息。为了提高模型精度,首先采用SG^[14]对原始反射光谱进行平滑处理,然后利用CWT^[15]增强光谱特征间的差异性,其中SG平滑算法设置窗口大小为15,CWT算法设置10个变换尺度($2^1 \sim 2^{10}$)。采用基于欧氏距离的Kennard-Stone(K-S)样本划分法^[16]将预处理后的数据分为校正集和预测集。

1.2.4 特征波长选择

冗余的波长量影响模型预测的精度和速度,为了简化模型,分别以CARS、SPA、BOSS、IVSO对预处理后的数据进行特征波长选择。CARS通过十折交叉验证循环筛选出交叉验证均方根误差(RMSECV)最小的波长子集作为特征波长集^[17],设置蒙特卡罗运行次数为100次,每次抽取80%样品作为校正集。SPA利用向量的投影分析将向量最大的波长作为待选,然后以校正模型选出冗余度低的特征波长组合^[18],将均方根误差(RMSE)最小的波长子集作为特征波长组合,设置SPA降维后的波长数量范围为1~40。BOSS以各子模型回归系数绝对值来分配各波长的权重,并通过加权自主抽样(WBS)令波长空间优化收缩,将RMSECV最小的波长子集作为特征波长组合^[19],设置bootstrap采样次数为1000次。IVSO通过结合加权二进制矩阵采样(WBMS)和顺序加法来模拟竞争过程,进而逐步消除无用及干扰信息波长^[20],将RMSECV最小的波长子集作为特征波长组合,设置WBMS采样次数为1000次。

1.2.5 掺伪量预测模型的建立

建立基于SVM、ELM、RF的油茶籽油掺伪预测模型。SVM通过选取不同的核函数使特征映射到高维空间构建最优超平面求解最优解^[21]。选取径向基函数作为核函数,并对惩罚因子(c)和核函数(γ)进行循环优选,建立最高预测率下的SVM鉴别模型。ELM以神经元的数量来调节模型VC维的大小^[22],选取Sigmoidal函数作为激活函数,同时通过优选神经元数量建立预测率最高的ELM鉴别模型。RF利用每棵决策树的判断做出投票,以投票最多的类别作为最终结果,RF决策树的数量是影响模型精度的重要因素之一^[23],选取预测准确率最高时的决策树数量。

1.2.6 模型评价

以决定系数(R^2)、RMSE、平均绝对误差(MAE)作为油茶籽油定量预测模型的评价指标。 R^2 分为校正决定系数(R_c^2)和预测决定系数(R_p^2),其代表油茶籽油预测模型拟合数据的准确程度, R^2 越接近于1表示模型效果越好。RMSE分为校正均方根误差(RMSE_c)和预测均方根误差(RMSE_p),其代表预测值和实际值之间的差距。MAE分为校正平均绝对误差(MAE_c)和预测平均绝对误差(MAE_p),相较于RMSE其能够减少异常值对模型精度的评价干扰。RMSE和MAE越小表示模型效果越好。

2 结果与讨论

2.1 光谱分析

样品原始光谱见图2。

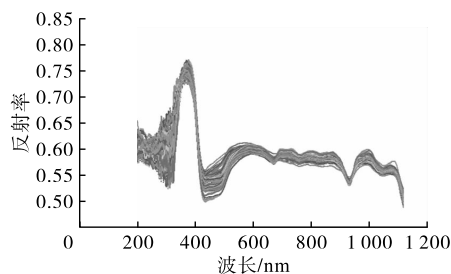


图2 样本原始光谱

如图2所示,430 nm左右处为索雷特特征峰,

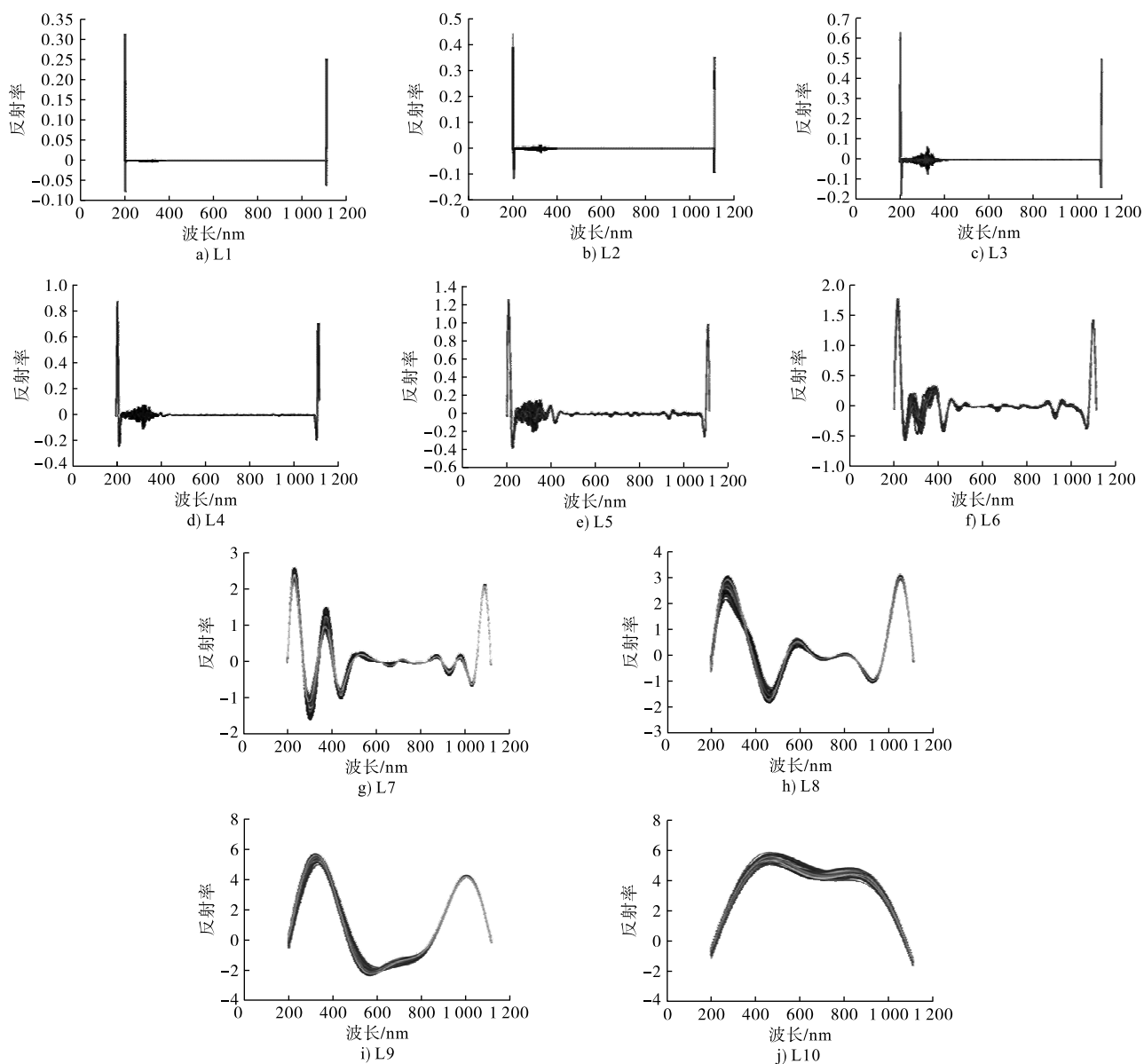


图3 CWT在不同分解尺度(L_n)下的预处理结果

由图3可知,随着分解尺度的增大,SG-CWT预处理将光谱曲线中不显著的特征峰谷先逐步放大,后续光谱曲线变得更加平滑,从而使一些不明显的特征峰谷被逐步去除,增大了光谱特征信息的捕

捉难度,证明CWT在低分解尺度下可以有效平滑数据,降低噪声。将SG-CWT预处理后的光谱矩阵带入SVM预测模型中,各光谱矩阵的 R_c^2 和 $RMSE_c$ 见表1。

2.2 光谱预处理和样本划分

对反射光谱数据进行了SG-CWT预处理。CWT对SG平滑处理后的光谱进行分解变换时,依据其分解尺度(L_n)按照 2^n 应小于全光谱(FS)波长数量(2068)的原则,将其分为10个尺度,CWT在不同分解尺度下的预处理结果见图3。

表1 SG-CWT在不同分解尺度下的 R_c^2 和 $RMSE_c$

预处理方法	R_c^2	$RMSE_c$
FS	0.999 7	0.005 4
SG-CWT(L1)	0.999 9	0.004 8
SG-CWT(L2)	0.999 9	0.004 6
SG-CWT(L3)	0.999 9	0.004 6
SG-CWT(L4)	0.999 9	0.004 5
SG-CWT(L5)	0.999 9	0.004 2
SG-CWT(L6)	0.999 8	0.004 5
SG-CWT(L7)	0.999 8	0.004 5
SG-CWT(L8)	0.999 7	0.006 2
SG-CWT(L9)	0.999 4	0.008 1
SG-CWT(L10)	0.980 0	0.047 0

由表1可知,SG-CWT(L5)为最优预处理方法,其 R_c^2 为0.999 9, $RMSE_c$ 为0.004 2。因此,后续只对SG-CWT(L5)预处理后的数据进行分析。

采用K-S样本划分法按照2:1的比例将SG-CWT(L5)预处理后的反射样本划分为校正集和预测集,得到校正集163个、预测集81个,如表2所示。

表2 样本划分结果

掺伪比例/%	校正集	预测集	掺伪比例/%	校正集	预测集
0	3	1	40	11	5
1	9	7	50	11	5
3	9	7	60	12	4
5	9	7	70	12	4
7	11	5	80	11	5
10	10	6	90	12	4
15	11	5	100	10	6
20	11	5	合计	163	81
30	11	5			

2.3 特征波长选择

CARS在第60次筛选时其 $RMSECV$ 为最小值(0.080 68),SPA在选择33个波长时 $RMSE$ 基本趋于稳定状态(0.048 483),BOSS在选择30个波长时其 $RMSECV$ 为最小值(0.065 418),IVSO在选择47个波长时其 $RMSECV$ 为最小值(0.097 77),此时CARS、SPA、BOSS、IVSO筛选出的特征波长数量分别为33、33、30、47,分别将波长数量降至全光谱(2 068个波长)的1.60%、1.60%、1.45%、2.27%,波长量的缩减可以有效提高模型的预测速度。对4种方法所选择波长在紫外、可见、近红外3个波段中的分布数量进行分析,结果见表3。

表3 CARS、SPA、BOSS、IVSO所提取波长在紫外、可见和近红外波段的分布情况

特征波长选择方法	波长分布数量		
	紫外 (200~360 nm)	可见光 (360~780 nm)	近红外 (780~1 100 nm)
CARS	0	25	8
SPA	12	9	12
BOSS	0	23	7
IVSO	0	41	6

由表3可知,CARS、BOSS、IVSO提取波长集中分布在可见光波段和近红外波段,而SPA所提取波长在紫外和近红外波段的波长数量较多。CARS、SPA、BOSS、IVSO特征波长选择结果见图4。

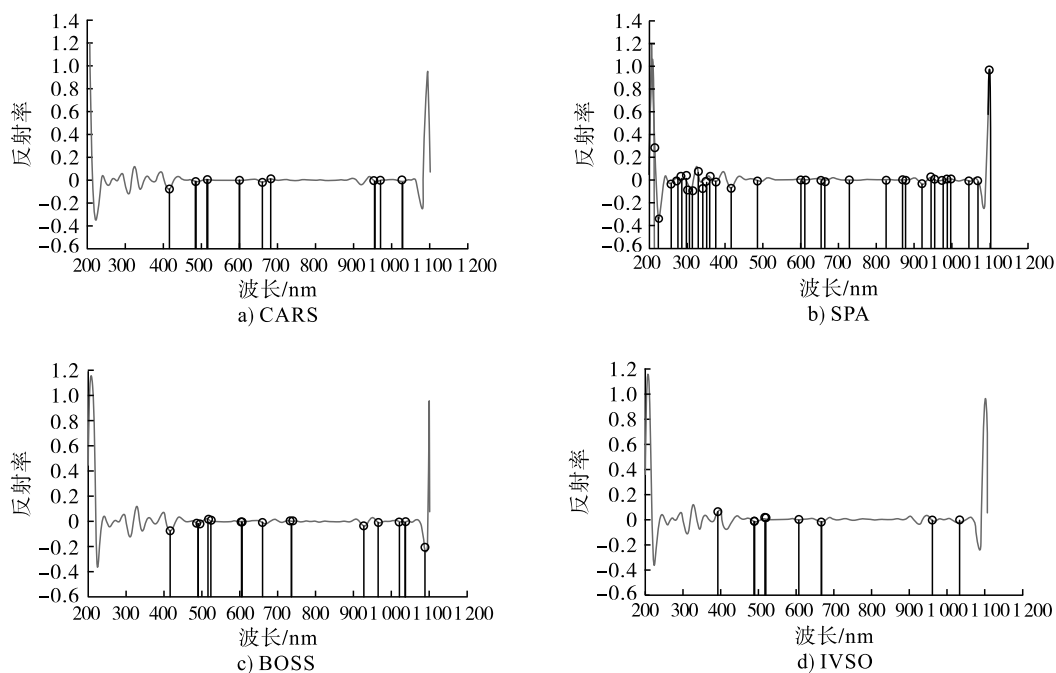


图4 CARS、SPA、BOSS、IVSO特征波长选择结果

由于4种光谱特征提取算法的提取原理具有各自的特点^[13],因此造成所提取波长存在一定差异。由图4可知,相较于CARS、BOSS、IVSO,SPA筛选波长的分布位置多集中于光谱两端处,而两端处存在较多的微弱震荡给预测模型增加了干扰,同时中间部分的波长分布较为分散且多处于特征峰谷不显著的平稳波段(700~900 nm)。相比于BOSS、CARS,IVSO在近红外波段提取的波长数量最少,并且近红外波段波长在IVSO所提整个特征波长中的比例较小(可见光与近红外比约7:1),单一波段的一小片连续波长难以代表整个光谱最具特征的信息^[12]。CARS和IVSO所提取波长均忽略了930 nm左右处特征峰谷的有效信息,而BOSS所提取的波长比CARS更加倾向分布于波形陡度大的位置处。

2.4 建模结果及建模方法比较

建立油茶籽油掺伪量快速预测模型时,通过主要参数的寻优进一步提升各模型的预测效果。SVM

模型通过十折交叉验证和网格搜索法确定最优的惩罚因子(c)与核函数(γ)。RF模型以十折交叉验证和粒子群优化算法确定最优决策树数量(n)。ELM模型设置For循环在1~100内以1为步长寻优隐含层神经元数量(n)。表4为RF、SVM、ELM模型寻优后的最佳参数。在此基础上,采用模型最优参数建立的SVM、ELM、RF模型评价指标如表5~表7所示。

表4 RF、SVM、ELM最佳模型参数选择

特征波长选择方法	RF	SVM		ELM
	n	c	γ	n
FS	98	1	0.003 9	79
CARS	33	3.031 4	0.189 5	98
SPA	29	16	0.020 6	88
BOSS	15	5.278 0	0.108 8	70
IVSO	25	9.189 6	0.108 8	90

表5 SVM油茶籽油掺伪量预测模型的评价指标

特征波长选择方法	波长数量	校正集			预测集		
		R_c^2	RMSE _c	MAE _c	R_p^2	RMSE _p	MAE _p
FS	2 068	0.998 9	0.011 4	0.005 7	0.995 1	0.023 5	0.014 5
CARS	33	0.999 5	0.007 3	0.005 6	0.997 2	0.017 7	0.011 7
SPA	33	0.999 2	0.009 2	0.006 5	0.995 9	0.021 8	0.015 4
BOSS	30	0.999 6	0.006 9	0.005 5	0.998 5	0.013 4	0.010 2
IVSO	47	0.999 2	0.009 6	0.006 7	0.997 0	0.018 2	0.013 9

表6 ELM油茶籽油掺伪量预测模型的评价指标

特征波长选择方法	波长数量	校正集			预测集		
		R_c^2	RMSE _c	MAE _c	R_p^2	RMSE _p	MAE _p
FS	2 068	0.998 2	0.013 9	0.010 3	0.988 2	0.037 1	0.028 6
CARS	33	0.999 6	0.006 5	0.005 0	0.997 2	0.017 6	0.012 9
SPA	33	0.999 1	0.009 8	0.007 4	0.996 1	0.021 0	0.016 7
BOSS	30	0.999 1	0.009 6	0.007 3	0.996 4	0.020 6	0.014 7
IVSO	47	0.999 3	0.008 7	0.006 7	0.996 3	0.020 9	0.014 4

表7 RF油茶籽油掺伪量预测模型的评价指标

特征波长选择方法	波长数量	校正集			预测集		
		R_c^2	RMSE _c	MAE _c	R_p^2	RMSE _p	MAE _p
FS	2 068	0.993 2	0.030 8	0.022 3	0.982 7	0.053 5	0.040 4
CARS	33	0.996 4	0.020 5	0.013 5	0.995 6	0.023 0	0.017 7
SPA	33	0.996 3	0.021 0	0.013 9	0.978 7	0.051 1	0.030 6
BOSS	30	0.995 8	0.022 0	0.015 4	0.990 5	0.033 2	0.020 9
IVSO	47	0.996 5	0.020 9	0.013 7	0.988 9	0.026 3	0.037 1

由表5~表7可知,在SVM预测模型中得到的最佳模型为BOSS-SVM,其 R_p^2 、RMSE_p、MAE_p分别为0.998 5、0.013 4、0.010 2。在RF、ELM模型中所得到的最佳模型都是基于CARS所建立,CARS-

ELM模型的 R_p^2 、RMSE_p、MAE_p分别为0.997 2、0.017 6、0.012 9,CARS-RF模型的 R_p^2 、RMSE_p、MAE_p分别为0.995 6、0.023 0、0.017 7。与CARS-ELM模型相比,BOSS-SVM模型的 R_p^2 升高了0.13

百分点, $RMSE_p$ 和 MAE_p 分别下降了 23.9% 和 20.9%。与 CARS - RF 模型相比, BOSS - SVM 模型的 R_p^2 升高了 0.29 百分点, $RMSE_p$ 和 MAE_p 分别下降了 41.7% 和 42.3%。综上, 最优预测油茶籽油掺伪量的模型为 BOSS - SVM。另外, 也可看出, 仅在 RF 模型中 SPA - RF 模型的 R_p^2 (0.978 7) 低于全光谱模型的 (0.982 7), RF 模型的预测精度降低。原因在于相比于 CARS、BOSS、IVSO, SPA 提取的波长更加倾向于紫外波段且波长分布陡度较低, 从而造成了模型精度的下降。结合波长特性分析, 发现所提取波长位置的聚集程度和陡度 (图 4) 对模型的预测效果具有一定影响。

BOSS - SVM 模型在校正集和预测集的预测性能如图 5 所示。

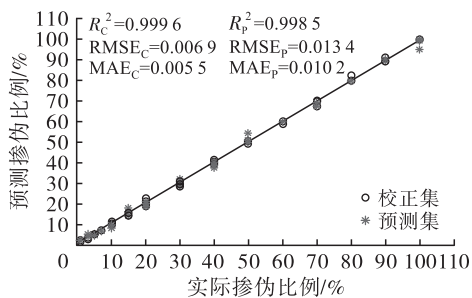


图 5 BOSS - SVM 模型预测性能

由图 5 可知, 各点拟合程度较高、预测误差较小, 证明 BOSS - SVM 模型的预测效果较好。

表 8 为 BOSS - SVM 模型预测集不同掺伪比例下的 $RMSE_p$ 。

表 8 BOSS - SVM 模型预测集不同掺伪比例下的 $RMSE_p$

掺伪比例/%	$RMSE_p$ /%
0	0.539 0
1	1.080 2
3	2.056 1
5	0.694 8
7	0.664 2
10	1.153 5
15	1.608 4
20	1.152 2
30	1.408 9
40	1.437 4
50	1.806 5
60	0.447 0
70	1.506 9
80	0.769 2
90	0.731 7
100	1.999 9

由表 8 可知, 不同掺伪比例下预测集的 $RMSE_p$

均较小, 说明该模型能够有效实现掺伪比例 1% 及以上的快速鉴别。

3 结论

在紫外 - 可见 - 近红外波段内利用反射率对油茶籽油中的大豆油、菜籽油、花生油、玉米油掺伪量进行鉴别。首先, 对比了 SG - CWT 在不同分解尺度下 (L1 ~ L10) 的全光谱的预处理结果, 得到在 SG - CWT (L5) 的预处理效果为最佳。分别以 CARS、SPA、BOSS、IVSO 算法对预处理后的全光谱进行特征波长筛选, 并建立 SVM、ELM、RF 的油茶籽油掺伪量预测模型。结果表明, 反射光谱所得到的最优模型为 BOSS - SVM, 其预测模型的 R_p^2 、 $RMSE_p$ 、 MAE_p 分别为 0.998 5、0.013 4、0.010 2。同时通过对特征波长特性进行分析得到, 波长位置的聚集程度和陡度对模型的预测效果存在一定影响。建立的基于反射光谱的油茶籽油掺伪量快速预测模型预测误差小, 可实现掺伪比例 1% 及以上的快速鉴别, 预测效果较好。

参考文献:

- [1] 岳超, 喻宁华, 李红爱, 等. 湖南茶油中脂肪酸及活性成分研究[J]. 食品安全质量检测学报, 2021, 12(5): 1972 - 1977.
- [2] 许俊道. 茶油的保健功能与开发前景[J]. 中国果菜, 2018, 38(10): 41 - 43.
- [3] 李丽, 吴雪辉, 寇巧花. 茶油的研究现状及应用前景[J]. 中国油脂, 2010, 35(3): 10 - 14.
- [4] 郭文川, 朱德宽, 张乾, 等. 基于近红外光谱的掺伪油茶籽油检测[J]. 农业机械学报, 2020, 51(9): 350 - 357.
- [5] AGHILI N S, RASEKH M, KARAMI H, et al. Detection of fraud in sesame oil with the help of artificial intelligence combined with chemometrics methods and chemical compounds characterization by gas chromatography - mass spectrometry [J/OL]. LWT - Food Sci Technol, 2022, 167:113863 [2022 - 11 - 26]. <https://doi.org/10.1016/j.lwt.2022.113863>.
- [6] RAY C L, GAWENIS J A, GREENLIEF C M. A new method for olive oil screening using multivariate analysis of proton NMR spectra [J/OL]. Molecules, 2021, 27(1): 213 [2022 - 11 - 26]. <https://doi.org/10.3390/molecules27010213>.
- [7] MARTÍNEZ GILA D M, SANMARTIN C, NAVARRO SOTO J, et al. Classification of olive fruits and oils based on their fatty acid ethyl esters content using electronic nose technology [J]. J Food Meas Charact, 2021, 15(6): 5427 - 5438.
- [8] MOUSA M A A, WANG Y, ANTORA S A, et al. An

- overview of recent advances and applications of FT - IR spectroscopy for quality, authenticity, and adulteration detection in edible oils[J]. *Crit Rev Food Sci Nutr*, 2022, 62(29): 8009 - 8027.
- [9] 沈乐丞, 曾秀英, 温志刚, 等. 基于近红外光谱技术的赣南茶油掺假快速鉴别[J]. *中国油脂*, 2022, 47(6): 62 - 67.
- [10] 林涛, 于海燕, 应义斌. 可见/近红外光谱技术在液态食品检测中的应用研究进展[J]. *光谱学与光谱分析*, 2008, 28(2): 285 - 290.
- [11] 姚婉清, 彭梦侠, 陈梓云, 等. 山茶油多元掺假近红外模型的建立与研究[J]. *食品安全质量检测学报*, 2020, 11(2): 493 - 499.
- [12] 翁士状, 储昭结, 王满琴, 等. 反射光谱结合光谱基二维卷积回归网络快速检测食用油中饱和脂肪酸[J]. *光谱学与光谱分析*, 2022, 42(5): 1490 - 1496.
- [13] SU N, WENG S, WANG L, et al. Reflectance spectroscopy with multivariate methods for non - destructive discrimination of edible oil adulteration [J/OL]. *Biosensors*, 2021, 11(12): 492[2022 - 11 - 26]. <https://doi.org/10.3390/bios11120492>.
- [14] CHEN H, PAN T, CHEN J, et al. Waveband selection for NIR spectroscopy analysis of soil organic matter based on SG smoothing and MWPLS methods [J]. *Chemom Intell Lab Syst*, 2011, 107(1): 139 - 146.
- [15] LI L, GENG S, LIN D, et al. Accurate modeling of vertical leaf nitrogen distribution in summer maize using in situ leaf spectroscopy via CWT and PLS - based approaches [J/OL]. *Eur J Agron*, 2022, 140(4): 126607[2022 - 11 - 26]. <https://doi.org/10.1016/j.eja.2022.126607>.
- [16] HARTIGAN J, WONG M A. Algorithm AS 136: A K - means clustering algorithm[J]. *J R Stati Soc C - Appl*, 1979, 28(1):100 - 108.
- [17] LI H, LIANG Y, XU Q, et al. Key wavelengths screening using competitive adaptive reweighted sampling method for multivariate calibration[J]. *Anal Chim Acta*, 2009, 648(1): 77 - 84.
- [18] ARAÚJO M C U, SALDANHA T C B, GALVÃO R K H, et al. The successive projections algorithm for variable selection in spectroscopic multicomponent analysis [J]. *Chemom Intell Lab Syst*, 2001, 57(2): 65 - 73.
- [19] DENG B C, YUN Y H, CAO D S, et al. A bootstrapping soft shrinkage approach for variable selection in chemical modeling[J]. *Anal Chim Acta*, 2016, 908: 63 - 74.
- [20] WANG W, YUN Y, DENG B, et al. Iteratively variable subset optimization for multivariate calibration[J]. *RSC Adv*, 2015, 5(116):95771 - 95780.
- [21] SUN T, MENG T, LIU Y. CAMELYON 17 challenge: A comparison of traditional machine learning (SVM) with the deep learning method [J/OL]. 2022:9910471[2022 - 11 - 26]. <https://doi.org/10.1155/2022/9910471>.
- [22] WANG H, CHU X, CHEN P, et al. Partial least squares regression residual extreme learning machine (PLSRR - ELM) calibration algorithm applied in fast determination of gasoline octane number with near - infrared spectroscopy [J/OL]. *Fuel*, 2022, 309(3): 122224[2022 - 11 - 26]. <https://doi.org/10.1016/j.fuel.2021.122224>.
- [23] VIGNEAU E, COURCOUX P, SYMONEAUX R, et al. Random forests: A machine learning methodology to highlight the volatile organic compounds involved in olfactory perception [J]. *Food Qual Prefer*, 2018, 68: 135 - 145.
- [24] 丁俭, 齐宝坤, 王立敏, 等. 5种不同植物油脂氧化程度与脂肪酸比例变化的相关性研究[J]. *中国粮油学报*, 2017, 32(8): 84 - 91.
- [25] STRYER L. *Biochemistry*[M]. 4th ed. New York: W. H. Freeman, 1995.
- [26] 沃克曼, 文依. 近红外光谱解析实用指南[M]. 褚小立, 许育鹏, 田高友, 译. 北京: 化学工业出版社, 2009.

· 公益广告 ·



节能减排，提质增效！

《中国油脂》宣