

留兰香高光谱特征波长稳健提取及产地鉴别模型构建

李晓龙, 殷 勇*, 于慧春, 袁云霞

(河南科技大学食品与生物工程学院, 河南 洛阳 471023)

摘 要: 本研究通过高光谱技术解决留兰香的产地鉴别问题。首先, 选取了5个产地共375个留兰香样本进行高光谱数据采集, 并采用多元散射校正方法进行预处理, 对预处理后的数据进行主成分分析(principal component analysis, PCA), 将得到的主成分变量构建Wilks λ 统计量, 且从小到大排序。然后, 分别绘制前3小Wilks λ 值所对应的主成分在各个波长下的权重系数曲线, 系数曲线上的峰谷所对应的波长为特征波长, 共得到37个特征波长。随后用Fisher判别分析构造鉴别模型的输入变量。最后, 分别构建支持向量机(support vector machine, SVM)和反向传播神经网络(back propagation neural network, BPNN)产地鉴别模型。结果表明: SVM模型的鉴别效果优于BPNN模型, 其训练集和测试集鉴别正确率分别为99.67%和98.67%。因此, 在PCA联合Wilks λ 统计量提取特征波长基础上构建的留兰香产地SVM鉴别模型可有效实现产地的鉴别, 并且所提取的特征波长不受到理化指标数量的影响, 使该鉴别模型具有较强的鲁棒性。

关键词: 留兰香; 高光谱技术; 主成分分析; Wilks λ 统计量; Fisher判别分析; 产地鉴别

Robust Extraction of Hyperspectral Feature Wavelengths and Development of a Model for Geographical Origin Identification of Spearmint

LI Xiaolong, YIN Yong*, YU Huichun, YUAN Yunxia

(College of Food and Bioengineering, Henan University of Science and Technology, Luoyang 471023, China)

Abstract: In this study, hyperspectral technology was used to solve the problem of identifying the geographical origin of spearmint. First, 375 spearmint leaf samples from 5 geographical origins were selected for hyperspectral data collection, and the hyperspectral data were preprocessed by multiplicative scatter correction (MSC) and analyzed by principal component analysis (PCA); the principal component variables were used to construct Wilks λ statistics, which were then ranked in an increasing order. Then, the weight coefficient curves of the principal components corresponding to the first three smallest Wilks λ values were drawn at each wavelength; a total of 37 feature wavelengths were obtained, namely the peak and valley wavelengths in the coefficient curves. Next, Fisher discriminant analysis (FDA) was used to construct the input variables for a model for graphical origin identification of spearmint. At last, a support vector machine (SVM) model and a back propagation neural network (BPNN) model for identifying the geographical origin of spearmint were constructed. The results indicated that the SVM model outperformed the BPNN model, with discrimination accuracy of 99.67% and 98.67% in the training and test sets, respectively. Therefore, the SVM model, constructed using PCA combined with Wilks λ statistics, can effectively identify the geographical origin of spearmint. In this model, the extracted feature wavelengths are not affected by the number of physicochemical indexes, making the model robustness.

Keywords: spearmint; hyperspectral technology; principal component analysis; Wilks λ statistics; Fisher discriminant analysis; geographical origin identification

DOI:10.7506/spkx1002-6630-20240513-093

中图分类号: TS205.9

文献标志码: A

文章编号: 1002-6630 (2024) 22-0262-07

收稿日期: 2024-05-13

基金项目: 国家重点研发计划项目 (2017YFC1600802)

第一作者简介: 李晓龙 (2000—) (ORCID: 0009-0000-5458-5532), 男, 硕士研究生, 研究方向为食品检测与智能分析。

E-mail: 3054886640@qq.com

*通信作者简介: 殷勇 (1966—) (ORCID: 0000-0002-4023-3656), 男, 教授, 博士, 研究方向为食品检测与智能分析。

E-mail: yinyong@haust.edu.cn

引文格式:

李晓龙, 殷勇, 于慧春, 等. 留兰香高光谱特征波长稳健提取及产地鉴别模型构建[J]. 食品科学, 2024, 45(22): 262-268.

DOI:10.7506/spkx1002-6630-20240513-093. <http://www.spkx.net.cn>

LI Xiaolong, YIN Yong, YU Huichun, et al. Robust extraction of hyperspectral feature wavelengths and development of a model for geographical origin identification of spearmint[J]. Food Science, 2024, 45(22): 262-268. (in Chinese with English abstract) DOI:10.7506/spkx1002-6630-20240513-093. <http://www.spkx.net.cn>

留兰香作为药食同源的作物^[1], 其富含的挥发油可作为食品添加剂被用于香料中, 在饮料、海鲜制品、奶茶、面包等食品加工领域被广泛使用^[2]。目前, 市场上留兰香的需求不断增加, 尽管不同产地的留兰香在外观上没有明显差别, 但其香味和营养价值存在较大差异, 导致存在品质低劣产区的留兰香冒充优质产区的留兰香以次充好的现象, 严重扰乱留兰香的供销市场。因此, 对不同产地的留兰香进行鉴别非常必要。

对于留兰香产地的判别, 传统的感官检测不仅工作量大且主观性强, 结果可靠性低; 而化学检测方法虽分类准确率高, 但属于有损检测, 且所需时间长。因此, 亟需建立一种高效、无损且准确率高的留兰香产地鉴别方法。近红外光谱技术因为具有快速、无损检测的特点, 已被广泛使用于食品质量领域^[3-4]。然而, 近红外光谱也有缺点, 作为一种点测量技术, 它的分辨率相对较低, 这可能限制其在分析异质结构样品时的应用^[5]。高光谱成像技术不仅检测精准、快速、易操作, 并且属于无损检测, 可以准确获得样品的光谱信息和空间信息, 是一种将光谱学与图像学结合的新兴无损检测技术^[6-7], 已经在地质领域^[8]、文物保护^[9]、医学诊疗^[10]、农产品品质评价^[11]及矿物信息提取^[12]等方面得到广泛使用。近年来, 高光谱技术在农产品品种和产地鉴别中的应用也有不少报道, 刘子健等^[13]采用连续投影算法(successive projections algorithm, SPA)对高光谱数据进行特征波长筛选, 通过多变量数据分析发现, SPA-支持向量机(support vector machine, SVM)模型对山楂产地的判别效果最佳, 训练集和预测集准确率分别为95.2%和93%; 杨洋等^[14]利用高光谱数据结合SPA降维, 凭借均方根误差(root mean squared error, RMSE)值选择了25个特征波长, 同时利用粒子群算法(particle swarm optimization, PSO)优化反向传播神经网络(back propagation neural network, BPNN)的初始权重和阈值, 构建PSO-BP模型作为分类器进行检测, 实现了花生种类的快速分类。以上研究表明了高光谱技术可以实现对不同品种、不同产地的农产品进行鉴别。

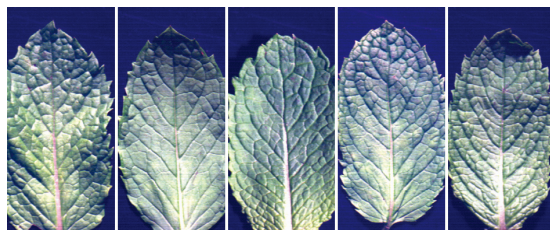
因此, 本研究拟利用高光谱成像技术对留兰香样本进行数据采集, 光谱数据在经过预处理后, 运用不受到理化指标限制的特征波长提取方法获得特征波长, 再借助于Fisher判别分析(Fisher discriminant analysis, FDA)构造具有可分性的鉴别模型输入变量, 进而构建稳健的

留兰香产地鉴别模型, 旨在为快速、无损、准确地鉴别留兰香产地提供一种新方法。

1 材料与方法

1.1 材料

新鲜留兰香叶分别购自安徽宣城、北京、广东广州、河南周口和上海5个产地。外形如图1所示, 每个产地随机选择75个生长状况良好且大小均匀一致、外观无明显损伤的新鲜留兰香叶片, 共375个样本。从每个产地的样本中以4:1的比例随机选取训练集和测试集, 得到的300个样本组成训练集, 75个样本组成测试集。



从左至右分别是安徽宣城、北京、广东广州、河南周口和上海5个产地的新鲜留兰香叶。

图1 5个不同产地的留兰香叶样本

Fig. 1 Spearmint leaves from five different regions

1.2 仪器与设备

IST50-3810高光谱成像仪 德国Inno-Spec公司; RK90000420108可调节光纤卤素灯(500 W) 德国Esylux公司; UV-2600紫外-可见分光光度计 日本岛津公司; TSQ9000气相色谱-质谱(gas chromatography-mass spectrometry, GC-MS)联用仪 美国赛默飞世尔公司。

1.3 方法

1.3.1 高光谱信息采集

实验室所用的高光谱图像采集系统详见文献[15]所述。通过预实验, 确定设置232 mm的物距、1.30 mm/s的物料滑台传送速度。采集样本的高光谱信息时, 将挑选的留兰香叶样本正面平铺在传送带上采集样本的高光谱信息。采集前进行黑白校正以减少暗电流及外界光源对数据的影响。其校正值按式(1)计算:

$$R = \frac{R_0 - R_b}{R_w - R_b} \quad (1)$$

式中： R 为校正后图像的光谱平均值； R_0 为原始图像的光谱平均值； R_b 为全黑标定图像的光谱平均值； R_w 为全白标定图像的光谱均值。

1.3.2 理化指标的测定

留兰香的成分比较复杂，不仅富含VC和总黄酮，而且留兰香中挥发油也是极其重要的成分，在挥发油中左旋香芹酮和柠檬烯的含量占有较大比例，并且这些理化成分对其品质优劣的影响比较大^[16]，因此选择左旋香芹酮、柠檬烯、VC和总黄酮作为测量的理化指标。

1.3.2.1 左旋香芹酮和柠檬烯含量的测定

左旋香芹酮和柠檬烯的测定均采用GC-MS仪，GC-MS工作条件参考文献[17]，GC条件：色谱柱HP-5MS弹性石英毛细管柱（30 m×0.25 mm，0.25 μm），进样量1 μL，载气为氦气，载气流量1 mL/min，分流比50:1，进样口温度230 °C。升温程序为起始温度50 °C，保持1 min；接着升温至150 °C，升温速率为10 °C/min，保持1 min；再升温至230 °C，升温速率为15 °C/min，保持5 min。MS条件：离子源温度230 °C，接口温度280 °C；电离能70 eV；扫描范围35~500 u。根据文献[18]，采用外标法定量，先进行标准溶液的配制，绘制出浓度和峰面积的线性关系曲线，依据关系曲线测定留兰香样本左旋香芹酮和柠檬烯的含量。

1.3.2.2 VC含量的测定

采用GB/T 5009.86—2016《食品中抗坏血酸的测定》中的2,6-二氯靛酚滴定法对不同产地留兰香中VC的含量进行测定。

1.3.2.3 总黄酮含量的测定

参考GB/T 20574—2006《蜂胶中总黄酮含量的测定方法 分光光度比色法》对不同产地留兰香中总黄酮的含量进行测定。

1.3.3 高光谱图像预处理

采集的样本图像中，利用ENVI 5.6软件在样品主脉中心点上方附近选取一个矩形区域为感兴趣区域，该区域的大小为2 499个像素点，将该区域的平均光谱反射值作为样本的高光谱值。在采集样本高光谱信息的过程中，两端信噪比较低。因此，去除波段范围为371.05~435.82 nm和961.60~1 023.82 nm的光谱数据，选择波段范围为435.82~961.60 nm的光谱数据，最终获得的光谱数据中包含1 039个波长。并且提取的高光谱值还受到随机噪声、光散射等因素的影响，因此，采用多元散射校正（multiplicative scatter correction, MSC）^[19]、标准正态变量变换（standard normal variate, SNV）^[20]、卷积平滑（Savitzky-Golary smoothing, S-G）^[21]分别进行平滑降噪，减少干扰，并对比预处理结果选出最合适的方法。

1.3.4 鲁棒特征波长提取

高光谱数据规模大、处理复杂度高，为了减少冗余

信息、简化计算、提高模型的可靠性，对预处理之后的光谱数据需提取特征波长。

SPA是最常用的特征波长提取方法^[22]，该方法提取特征波长时需要考虑理化指标的影响，但测量全部理化指标不切实际，而且不同理化指标是不同属性的体现，理化指标所体现的产地差异性还有待进一步研究。因此，仅靠部分理化指标进行特征波长的提取也仅仅是表达样本间的部分属性，而且这些属性能否表现产地差异性也很难判断。由此，提出了不依赖理化指标的特征波长提取方法，即主成分分析（principle component analysis, PCA）^[23]联合Wilks λ 统计量^[24]的特征波长稳健提取方法。

PCA联合Wilks λ 统计量是用PCA对预处理数据进行分析，然后通过式（2）得到各主成分变量的Wilks λ 统计量，将 λ 值从小到大排列。 λ 值越小，表示不同产地样本的差异性越显著。考虑到只用最小 λ 值对应的主成分提取出的特征波长光谱信息可能不足以充分表征留兰香产地的差异性，因此分别绘制最小 λ_1 、次小 λ_2 、再次小 λ_3 所对应的主成分在各个波长下的权重系数曲线，曲线上峰谷对应的波长为特征波长（第4小 λ 值对应主成分的权重系数曲线已不存在明显的波峰、波谷，所以只绘制至第4小 λ 值）。接着将 λ_1 、 λ_2 、 λ_3 对应的主成分所提取出的特征波长进行组合，作为最终的特征波长。该方法提取特征波长仅借助光谱信息，不需要理化指标，因此，所提取特征波长不受理化指标局限性的影响，使其具有鲁棒性。

$$\lambda = \frac{|D|}{|T|} \quad (2)$$

式中： D 为预处理数据经PCA后得到的主成分矩阵中某一列的组内离差； T 为该列的总离差。

另外，为了验证上述特征波长提取方法的有效性，采用SPA提取了4种主要理化指标融合下的特征波长，以尽可能多地表征留兰香的品质属性。这是因为对应于每个理化指标所提取的特征波长仅是留兰香局部属性的体现，若是用4种理化指标分别提取特征波长，也仅是它们表征属性的分别体现，不能反映指标间的内在关联性。所以，需要对它们进行信息融合，用融合后的变量信息提取特征波长。4种指标的选择见1.3.2节；信息融合采用PCA方法，选取累计贡献率达到95%以上的主成分与SPA结合进行特征波长提取，然后将提取到的特征波长进行组合，作为由4个理化指标所提取特征波长的最终特征波长。

1.3.5 鉴别模型输入变量构造

原始变量经FDA线性组合后生成具有可分离性的FD新变量，且新变量间相互正交，所以认为FDA实现了数据的可分性融合^[25]，并且不同产地样本的差异性信息也包含在新融合变量中。因此，可将新生成的FD变量作为

鉴别模型的输入变量。依照1.3.4节的方法将提取出的特征波长的光谱值进行FDA, 构造具有可分性的FD新变量, 并选取累计判别能力达到99%以上的作为鉴别模型的输入变量。

1.3.6 鉴别模型构建

选取SVM和BPNN构建产地鉴别模型, 并比较两种模型的鉴别效果。

1.3.6.1 SVM模型构建

SVM是一种用于解决线性可分和不可分问题的机器学习算法, 具有很强的处理非线性问题能力。它就是先将低维的输入向量转换成高维特征空间, 接着在高维特征空间中寻找最优的分离超平面对数据点进行分类, 使得分类误差在这个空间中最小化。

SVM在构建过程中, 首先将留兰香样本高光谱数据集以4:1的比例随机划分为训练集和测试集, 接着寻找最优参数。SVM模型中核函数的选择对模型性能影响显著, 本研究选择的核函数是性能较稳定的径向基函数(radial basis function, RBF), 为了使模型训练效果更好, 利用网格搜索法和K折交叉验证方法寻找最优核参数 c 和 σ , 此时 c 和 σ 的变化范围为 $-10 \sim 10$, 步长为0.2^[26]。

1.3.6.2 BPNN模型构建

BPNN是一种应用广泛的多层前馈神经网络。BPNN的原理是按照误差逆向传播算法进行训练, 不仅能够映射多维函数, 还具有优秀的模式分类能力^[27]。

在构建BPNN模型过程中, 先将留兰香样本高光谱数据集以4:1的比例随机划分为训练集和测试集, 接着选用了非线性映射能力较强的3层BPNN模型, 包括输入层、隐含层和输出层。tansig函数为隐含层传递函数, purelin函数为输出层传递函数, trainlm函数为训练函数^[28]。此时的目标误差为0.000 1, 学习率为0.001, 训练迭代次数为1 000次。

1.4 数据处理

所有理化指标的测定均重复3次, 从紫外-可见分光光度计和GC-MS联用仪中导出数据, 同时也从GC-MS联用仪中导出样本色谱图, 利用Excel软件进行理化指标数据的处理并绘制标准曲线, 使用SPSS软件进行差异显著性分析。

2 结果与分析

2.1 原始光谱数据预处理

通过3种光谱数据预处理结果的对比分析, 发现MSC是最合适的光谱数据预处理方法。图2、3分别给出5个产地所有留兰香原始光谱图及MSC预处理后光谱图。对比分析发现MSC预处理后的光谱曲线中毛刺明显减少, 有效消除了光谱噪声。从图3还发现在435.82~961.60 nm波段范围内, 平均光谱反射值波动幅度较大, 整体呈现“两峰一谷”的明显变化趋势,

这表明该波段包含着丰富的留兰香信息。因此, 筛选435.82~961.60 nm波段范围以进行后续处理。

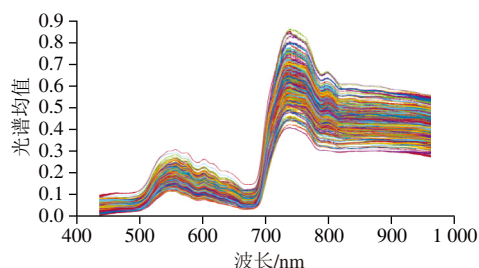


图2 留兰香原始光谱图

Fig. 2 Original spectra of spearmint

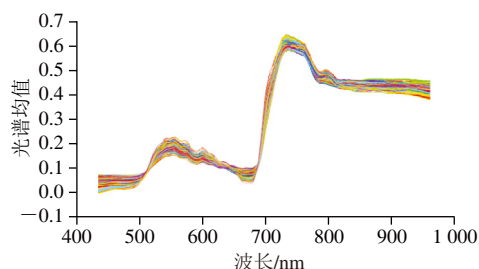


图3 MSC预处理后留兰香光谱图

Fig. 3 MSC preprocessed spectra of spearmint

2.2 理化指标分析

由图4、5及表1可以看出, 虽然不同产地的理化指标含量均有差异, 但其中上海产地留兰香的左旋香芹酮和柠檬烯含量明显少于其他产地, 这可能是在运输、实验等过程中操作不当造成的, 并且5个产地留兰香之间的左旋香芹酮、柠檬烯和总黄酮含量普遍存在显著差异($P < 0.05$)。

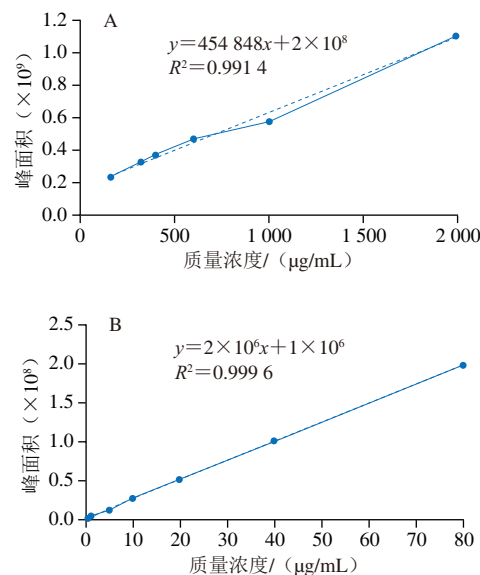
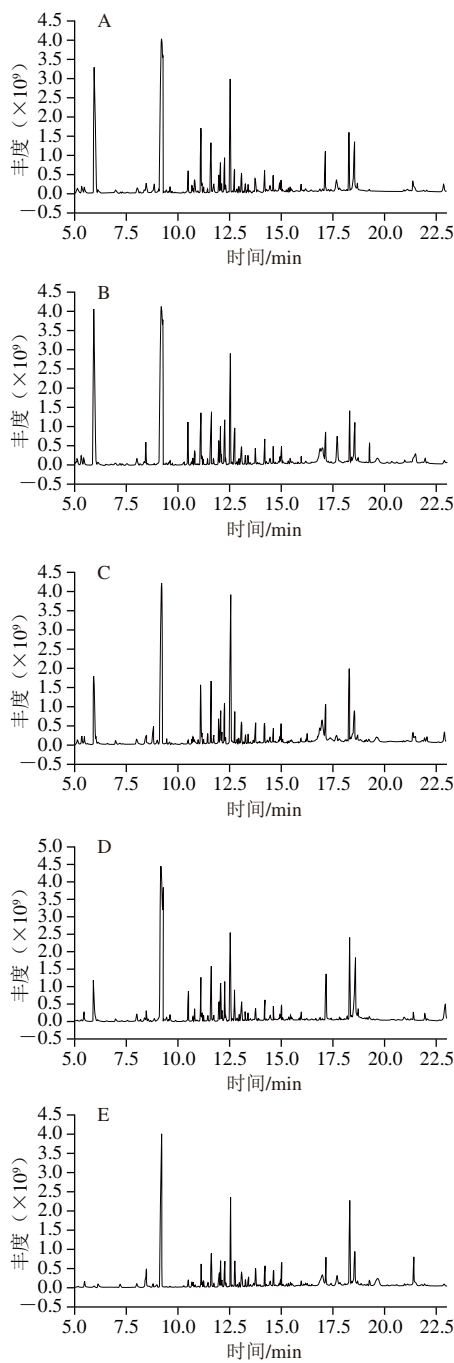


图4 左旋香芹酮(A)和柠檬烯(B)标准曲线

Fig. 4 Standard curves of *L*-carvone (A) and limonene (B)



A~E分别为安徽宣城、北京、广东广州、河南周口和上海5个产地留兰香。

图5 不同产地留兰香的样本色谱图

Fig. 5 Chromatograms of spearmint from different regions

表1 不同产地留兰香的理化指标分析

Table 1 Analysis of physicochemical indexes of spearmint from different regions

产地	左旋香芹酮含量/ (mg/g)	柠檬烯含量/ (mg/g)	VC含量/ (mg/g)	总黄酮含量/ (mg/g)
安徽宣城	10.05±0.01 ^d	1.10±0.00 ^b	0.34±0.03 ^{bc}	23.75±0.04 ^b
北京	15.50±0.03 ^a	1.60±0.03 ^a	0.29±0.01 ^{bc}	20.00±0.02 ^d
广东广州	12.98±0.01 ^b	1.07±0.04 ^b	0.17±0.02 ^c	10.00±0.09 ^c
河南周口	11.45±0.09 ^c	0.40±0.10 ^c	0.60±0.26 ^a	26.25±0.05 ^a
上海	2.10±0.13 ^e	0.01±0.00 ^d	0.41±0.01 ^{ab}	21.25±0.03 ^c

注：同列字母不同表示差异显著（ $P<0.05$ ）。

2.3 特征波长提取

2.3.1 PCA联合Wilks λ 统计量法提取特征波长

最小、次小、再次小的 λ 值对应的主成分变量分别为第2主成分、第3主成分以及第5主成分。因为权重系数的绝对值越大，该波长对主成分的贡献度就越大，也越能够体现不同产地样本间的差异性，故选取权重系数曲线上峰谷对应的波长为特征波长。如表2所示，由于有3个波长为重复波长，因此共筛选出37个特征波长，

表2 前3小 λ 值对应的主成分提取的特征波长个数

Table 2 Number of characteristic wavelengths extracted from top three principal components with the smallest λ values

主成分	第2主成分	第3主成分	第5主成分
特征波长个数	9	19	12

采用PCA对高光谱数据进行特征波长的提取，选取累计贡献率达99%的前9个主成分作为输入变量输入到SVM和BPNN模型中，其训练集鉴别正确率分别为96.67%和92.33%，测试集鉴别正确率分别为86.67%和85.33%，鉴别效果不理想。所以，单独用PC提取的特征波长构建的鉴别模型不是最优模型。

2.3.2 SPA提取特征波长验证

在SPA中将最大波长数目设置为50，用PCA对4种理化指标进行信息融合，融合后的主成分1与主成分2的累计方差贡献率已达95%以上。因此，由SPA分别与第1主成分和第2主成分结合对预处理后的光谱数据提取特征波长，结果如表3所示，SPA与第1主成分结合提取24个特征波长，与第2主成分结合也提取24个特征波长。将这些特征波长进行组合，筛选出35个特征波长，其中有13个波长为重复波长。

表3 SPA结合理化指标主成分提取的特征波长数

Table 3 Number of characteristic wavelengths extracted by SPA combined with either the first or second principal component

主成分	第1主成分	第2主成分
特征波长数	24	24

2.4 鉴别模型输入变量构造

将上述得到的特征波长按照1.3.5节的方法进行FDA，结果见表4，将生成的FD变量输入到SVM和BPNN模型中，将两种方法提取到的特征波长光谱值经FDA融合后，生成的前4个FD变量累计判别能力均已达到99%，所以将两种方法通过FDA得到的前4个FD变量都分别输入到SVM和BPNN的模型中，以构建产地鉴别模型。

表4 两种特征波长提取方法经FDA生成的新FD变量的判别能力

Table 4 Discriminant capacity of new FD variables generated by FDA using two feature wavelength extraction methods

提取方法	FD1	FD2	FD3	FD4	累计判别能力
PCA+Wilks λ 统计量	0.953 21	0.030 97	0.010 37	0.005 44	0.999 99
PCA+SPA	0.929 01	0.046 06	0.013 28	0.011 64	0.999 99

2.5 鉴别模型构建

SVM和BPNN的产地鉴别模型结果如表5所示, 4个理化指标经PCA融合后, 借助于SPA提取的特征波长也可较好地表征产地的差异性, 这充分说明了运用多个主要理化指标融合提取特征波长是适宜的, 也是必需的。然而, 高光谱信息借助PCA联合Wilks λ 统计量提取特征波长的效果更好, 优于SPA, 且不受测量的理化指标数量影响, 进而使产地鉴别模型具有较强的鲁棒性。同时, Wilks λ 统计量 λ 值本身就能体现出产地差异性, λ 值越小, 样本之间产地差异性就越大, 而SPA却无法判断。

表5 两种特征波长提取方法对应的产地鉴别模型的鉴别结果
Table 5 Identification results of SVM and BPNN models with two feature wavelength extraction methods

特征波长提取方法	FD变量 个数	SVM鉴别模型正确率/%		BPNN鉴别模型正确率/%	
		训练集	测试集	训练集	测试集
PCA+Wilks λ 统计量	4	99.67	98.67	98.67	97.33
PCA+SPA	4	99.33	97.33	98.67	94.67

由表5还可知, 对于PCA联合Wilks λ 统计量提取的特征波长, SVM模型对训练集鉴别的正确率为99.67%, 对测试集的鉴别正确率为98.67%, 被错误鉴别的样本个数均为1; 而BPNN模型的训练集样本鉴别中有4个样本被错误鉴别, 正确率为98.67%, 测试集样本鉴别中有2个样本被错误鉴别, 正确率为97.33%。

对比两种模型结果可知, SVM模型的训练集鉴别正确率比BPNN模型的高1%, 测试集鉴别正确率高1.34%, SVM的鉴别结果优于BPNN。

3 结论

为了实现留兰香产地鉴别, 本研究采集了安徽宣城、北京、广东广州、河南周口和上海5个产地留兰香的高光谱数据, 通过对特征波长提取方法、可分性变量构造方法的研究, 分别构建了SVM和BPNN产地鉴别模型。主要结论如下:

1) 提出了一种基于PCA联合Wilks λ 统计量的稳健特征波长提取方法。将预处理数据经PCA处理后, 计算各主成分变量的 λ 值, 并从小到大排序。选出前3小 λ 值对应的3个主成分变量, 绘出这3个主成分在各个波长下的权重系数曲线, 提取系数曲线上峰谷对应的波长为特征波长, 并将分别提取到特征波长进行组合, 作为最终的特征波长。该方法不仅有效减少了数据中的冗余信息, 而且提取的特征波长仅借助光谱信息, 不受因测量的理化指标较少而产生的不利影响, 使其具有鲁棒性。

2) 利用FDA对特征波长的光谱值进行可分性融合, 构建了产地鉴别模型的输入变量。对比两种鉴别模型的结果, SVM对训练集与测试集的鉴别结果均要优于BPNN, 其正确率分别是99.67%和98.67%。

综上, 对于留兰香产地的鉴别, 在PCA联合Wilks λ 统计量稳健提取特征波长的基础上所构建的SVM模型可有效实现留兰香产地的快速准确鉴别, 并且所提取的特征波长不受到理化指标的影响, 进而使该鉴别模型具有较强的鲁棒性。

参考文献:

- [1] 郝瑞芬, 贾会玲, 钱骅, 等. 不同生长期留兰香精油含量与品质评价[J]. 中国食品添加剂, 2017, 28(9): 55-60. DOI:10.3969/j.issn.1006-2513.2017.09.002.
- [2] 王莹莹, 陈鑫沛, 董颖, 等. 留兰香调味油的研制[J]. 中国调味品, 2022, 47(4): 117-121. DOI:10.3969/j.issn.1000-9973.2022.04.022.
- [3] YUAN W D, JIANG H Z, SUN M M, et al. Geographical origin identification of Chinese tomatoes using long-wave Fourier-transform near-infrared spectroscopy combined with deep learning methods[J]. Food Analytical Methods, 2023, 16(4): 664-676. DOI:10.1007/s12161-023-02444-1.
- [4] BACA-BOCANEGRA B, NOGALES-BUENO J, HERNÁNDEZ-HIERRO J M, et al. Evaluation of extractable polyphenols released to wine from cooperage byproduct by near infrared hyperspectral imaging[J]. Food Chemistry, 2018, 244: 206-212. DOI:10.1016/j.foodchem.2017.10.027.
- [5] 袁伟东, 姜洪喆, 杨诗雨, 等. 基于高光谱成像技术的宁夏枸杞产地溯源鉴别[J]. 食品科学, 2024, 45(6): 254-260. DOI:10.7506/spkx1002-6630-20230620-159.
- [6] NIRERE A, SUN J, KAMA R, et al. Nondestructive detection of adulterated wolfberry (*Lycium Chinense*) fruits based on hyperspectral imaging technology[J]. Journal of Food Process Engineering, 2023, 46(4): e14293. DOI:10.1111/jfpe.14293.
- [7] ROMÁN S M S, FERNÁNDEZ-NOVALES J, CEBRIÁN-TARANCÓN C, et al. Monitorization of varietal aroma composition dynamics during ripening in intact *Vitis vinifera* L. tempranillo blanco berries by hyperspectral imaging[J]. Journal of Agricultural and Food Chemistry, 2023, 71(5): 2616-2627. DOI:10.1021/acs.jafc.2c07425.
- [8] 黄净晴, 王浩, 韩菁, 等. 光谱遥感技术在地质领域的应用研究——评《高光谱遥感地质作用建模及应用》[J]. 有色金属工程, 2022, 12(1): 2. DOI:10.3969/j.issn.2095-1744.2022.01.022.
- [9] 张立福, 王飒, 张燕, 等. 文博领域的高光谱遥感技术应用进展[J]. 测绘学报, 2023, 52(7): 1126-1138. DOI:10.11947/j.AGCS.2023.20220655.
- [10] 黎珊珊, 谢志茹, 黄燕秋, 等. 高光谱成像技术用于医学诊疗的研究进展[J]. 广东医学, 2018, 39(13): 1921-1925. DOI:10.13820/j.cnki.gdyx.20180717.001.
- [11] 张伏, 王新月, 崔夏华, 等. 高光谱结合极限学习机的玉米品种鉴别[J]. 光谱学与光谱分析, 2023, 43(9): 2928-2934. DOI:10.3964/j.issn.1000-0593(2023)09-2928-07.
- [12] 白杨林, 吕凤军, 苏鸿博, 等. 高光谱遥感蚀变矿物信息提取研究综述[J]. 遥感信息, 2023, 38(1): 1-10. DOI:10.20091/j.cnki.1000-3177.2023.01.001.
- [13] 刘子健, 顾佳盛, 周聪, 等. 基于高光谱成像技术的山楂产地判别研究[J]. 食品工业科技, 2024, 45(10): 282-291. DOI:10.13386/j.issn1002-0306.2023090074.
- [14] 杨洋, 徐熙平, 薛航, 等. 基于SPA-PSO-BP的花生高光谱图像分类方法研究[J/OL]. 激光技术, 1-14[2024-03-15].
- [15] 贾梦梦, 殷勇, 于慧春, 等. 高光谱成像融合特征波长筛选监测番茄贮藏中品质变化的方法[J]. 光谱学与光谱分析, 2023, 43(3): 969-975. DOI:10.3964/j.issn.1000-0593(2023)03-0969-07.

- [16] 荆晓艳, 张思文, 刘利锋, 等. 留兰香提取物挥发性成分分析[J]. 中国调味品, 2013, 38(12): 68-70; 74. DOI:10.3969/j.issn.1000-9973.2013.12.017.
- [17] 薛子瞻, 张玄兵, 王健, 等. 不同罗勒品种香气成分GC-MS分析[J]. 南方农业学报, 2016, 47(3): 449-453. DOI:10.3969/j.issn.2095-1191.2016.03.449.
- [18] 刘梦文, 沈静, 阿依努尔·白克热, 等. 一测多评法同时测定桑叶中9种绿原酸类和黄酮类成分[J]. 食品工业科技, 2024, 45(15): 295-303. DOI:10.13386/j.issn1002-0306.2023080195.
- [19] GUO W, LI X X, XIE T H. Method and system for nondestructive detection of freshness in *Penaeus vannamei* based on hyperspectral technology[J]. Aquaculture, 2021, 538: 736512. DOI:10.1016/j.aquaculture.2021.736512.
- [20] ZHANG J, TIAN H Q, WANG D, et al. A novel spectral index for estimation of relative chlorophyll content of sugar beet[J]. Computers and Electronics in Agriculture, 2021, 184: 106088. DOI:10.1016/j.compag.2021.106088.
- [21] MA L, ZHANG Y, ZHANG Y Y, et al. Rapid nondestructive detection of chlorophyll content in muskmelon leaves under different light quality treatments[J]. Agronomy, 2022, 12(12): 3223. DOI:10.3390/agronomy12123223.
- [22] FENG S, CAO Y L, XU T Y, et al. Rice leaf blast classification method based on fused features and one-dimensional deep convolutional neural network[J]. Remote Sensing, 2021, 13(16): 3207. DOI:10.3390/rs13163207.
- [23] 王琪, 常庆瑞, 李铠, 等. 基于主成分分析和随机森林回归的冬小麦冠层叶绿素含量估算[J]. 麦类作物学报, 2024, 44(4): 532-542. DOI:10.7606/j.issn.1009-1041.2024.04.14.
- [24] 薛书凝, 殷勇, 于慧春, 等. 香蕉贮藏中腐败基准确定与高光谱信息表征及腐败预警模型构建[J]. 光谱学与光谱分析, 2020, 40(12): 3871-3877. DOI:10.3964/j.issn.1000-0593(2020)12-3871-07.
- [25] 郭林鸽, 殷勇, 于慧春, 等. 基于Fisher判别分析可分性信息融合的马铃薯VC含量高光谱检测方法[J]. 食品科学, 2024, 45(7): 164-171. DOI:10.7506/spkx1002-6630-20230808-034.
- [26] 任永杰, 殷勇, 于慧春, 等. 三维荧光光谱融合小波包分解融合Fisher判别分析及支持向量机识别紫苏[J]. 食品科学, 2024, 45(1): 198-203. DOI:10.7506/spkx1002-6630-20230518-174.
- [27] 窦嘉铭, 马鸿雁. 基于多神经网络组合的建筑能耗预测研究[J]. 计算机仿真, 2022, 39(5): 438-443. DOI:10.3969/j.issn.1006-9348.2022.05.088.
- [28] 朱焯炜, 阙立志, 陈国庆, 等. 三维荧光光谱结合平行因子及神经网络对清香型白酒的年份鉴别[J]. 光谱学与光谱分析, 2015, 35(9): 2573-2577. DOI:10.3964/j.issn.1000-0593(2015)09-2573-05.