

衰减全反射傅里叶变换红外光谱法结合模式识别 分析橄榄油中掺杂大豆油的研究

黎远鹏¹ 黄富荣¹ 洗瑞仪¹ 夏凯悦¹ 汪勇² 陈星旦^{1,3}

¹暨南大学光电工程系, 广东 广州 510632

²暨南大学食品工程系, 广东 广州 510632

³中国科学院长春光学精密机械与物理研究所, 吉林 长春 130033

摘要 采用衰减全反射傅里叶变换红外光谱(FTIR-ATR)结合模式识别的方法建立橄榄油的快速鉴别模型。实验中测定不同品牌的纯橄榄油样本与纯大豆油样本以及橄榄油中掺入大豆油的混合油样本的 FTIR-ATR 光谱。将 625~4000 cm^{-1} 波段的光谱数据作为鉴别研究对象,应用主成分分析法(PCA)对光谱数据进行降维处理,再结合 Fisher 判别分析和多层感知器神经网络两种模式识别的方法对纯橄榄油、纯大豆油以及三种掺杂有大豆油的橄榄油样本进行鉴别。结果表明:对原始数据所采用的预处理方法中,经小波去噪后的数据的鉴别效果最佳。PCA 结合 Fisher 判别分析建立的判别模型,原始判别的准确率达到 100%,交叉验证的准确率达到 97.1%;多层感知器神经网络鉴别模型训练集分类准确率为 100%,测试集分类准确率为 100%。因此,FTIR-ATR 光谱技术结合模式识别方法可以实现对纯橄榄油、纯大豆油以及三种掺杂有大豆油的橄榄油样本的鉴别,而且具有操作简便、快速、可靠等优点。

关键词 光谱学; FTIR-ATR; Fisher 判别; 神经网络; 鉴别

中图分类号 O433.4

文献标识码 A

doi: 10.3788/LOP52.033001

Discrimination Analysis of Adulterating Olive Oil with Soybean Oil by Attenuated Total Reflectance-Fourier Transform Infrared Spectrometry Combined with Pattern Recognition Methods

Li Yuanpeng¹ Huang Furong¹ Xian Ruiyi¹ Xia Kaiyue¹ Wang Yong² Chen Xingdan^{1,3}

¹Opto-Electronic Engineering Department, Jinan University, Guangzhou, Guangdong 510632, China

²Department of Food Science and Engineering, Jinan University, Guangzhou, Guangdong 510632, China

³Changchun Institute of Optics, Fine Mechanics and Physics, Chinese Academy of Sciences, Changchun, Jilin 130033, China

Abstract Attenuated total reflectance-Fourier transform infrared spectroscopy (FTIR-ATR) combined with pattern recognition is applied to establish a rapid discrimination model of olive oil. Measurements of FTIR-ATR spectra of pure olive oil samples and pure soybean oil samples from different brands and mixed samples of pure olive oil doped with soybean oil are performed. The obtained spectra between 625 cm^{-1} and 4000 cm^{-1} are regarded as the research object of discrimination analysis. The principal component analysis (PCA) is applied to the dimension reduction of these spectra. Identification of the pure olive oil samples, pure soybean oil samples and three samples of olive oil doped with soybean oil is carried out by two pattern recognition methods, the Fisher discrimination and multilayer perception neural network. The experiment results show that the identification effect after the wavelet denoising is the best among all the adopted pretreatment methods. The original discriminant accuracy of discrimination model established by PCA combined with Fisher discriminant analysis reaches 100%. The cross validation accuracy reaches

收稿日期: 2014-08-15; 收到修改稿日期: 2014-09-10; 网络出版日期: 2015-02-10

基金项目: 国家自然科学基金(31371785)、广东省战略新兴产业核心技术攻关项目(2012A032300016)、高等学校优秀青年创新人才培养计划(LYM11026)、中央高校基本科研业务费专项资金(21612436, 21612353)

作者简介: 黎远鹏(1989—),男,硕士研究生,主要从事光谱分析技术与应用等方面的研究。E-mail: 18825062335@139.com

导师简介: 黄富荣(1979—),男,博士,副研究员,主要从事光谱分析技术与应用等方面的研究。

E-mail: furong_huang@163.com(通信联系人)

97.1%. Classification accuracy of the training set and the test set by multilayer perception neural network identification model is 100%. Therefore, FTIR-ATR spectroscopy combined with pattern recognition method can realize the identification of the pure olive oil samples, pure soybean oil samples and three different oil blend samples. This method has some advantages over other methods owing to its easy operation, high speed and reliability.

Key words spectroscopy; FTIR-ATR; Fisher discriminant; neural network; identification

OCIS codes 300.1030; 300.6170; 300.6300; 300.6340

1 引言

橄榄油是指以油橄榄树(*Olea europaea* L.)的果实为原料制取的油脂^[1],含有丰富的不饱和脂肪酸、角鲨烯、多酚、维生素等营养成分,具有抗氧化、调节胆固醇、预防癌症、美容等功效^[2],被称为“液体黄金”。橄榄油中所含油酸、亚油酸和亚麻油酸的比例正好是人体所需的比例,这是其他植物油所不能比拟的,而且其价格比其他食用植物油高很多。一些不法分子常向橄榄油中掺入其他价格低廉的食用油以降低成本,比如大豆油,后者在色泽、脂肪酸组成等方面都与橄榄油非常接近,识别起来十分困难,使得不法商家可以鱼目混珠并从中牟取暴利,不仅损害了消费者的利益,而且严重地影响了橄榄油市场正常发展。因此,橄榄油品质鉴别已经成为迫切需要解决的课题之一。

目前,对于橄榄油的品质分析手段主要有常规理化检测法、气相色谱法^[3]、高效液相色谱法^[4-5]、荧光光度法^[6]、电子鼻技术^[7]、核磁共振^[8]等方法,虽然分析的准确度较高,但存在耗时长、样品预处理过程复杂等缺点。因此研究一种快速、准确鉴别橄榄油的方法,对于保护消费者的利益,提高橄榄油的品质以及维护橄榄油市场的稳定发展有着重要的实际意义。

研究表明橄榄油和豆油主要含有棕榈酸、硬脂酸、油酸、亚油酸、亚麻酸等5种脂肪酸,主要成分为甘油三酯(三酰基甘油酯),是混脂肪酸甘油三酯的混合物^[9]。由于植物油脂在组成上共性非常大,其物理性质和化学性质差异较小。红外光谱作为一种快速无损的检测技术手段应用广泛^[10-13],其中红外波段($400\sim 4000\text{ cm}^{-1}$)体现的是分子基频振动的信息,特征性强,由于油脂中脂肪酸种类的不同,组成比例的不同,在中红外吸收光谱中的信息贡献也会有所差异,这些图谱便是对油脂进行鉴别分析的基础。更重要的是,衰减全反射(ATR)光谱是通过样品表面的反射信号获得样品表层有机成分的结构信息,具有制样简单、无破坏性、对样品的大小和含水量没有特殊要求、可以实现原位测试、实时跟踪等优点,在常规傅里叶变换红外光谱仪(FTIR)上配置ATR附件即可实现测量,仪器价格相对低廉,而且结合模式识别方法可以从数据中最大限度地提取信息,并按照样品的某种性质对样品集进行分类,在农产品^[14]、油脂检测^[15]、生物医学^[16-17]等领域得到了广泛的应用。

本文基于FTIR-ATR光谱结合主成分分析(PCA)与Fisher判别和多层感知器神经网络的模式识别方法建立橄榄油中掺杂大豆油的快速鉴别模型,为实现橄榄油品质的快速无损分析提供技术支撑。

2 实验部分

2.1 实验仪器与测量方法

本实验采用BRUKER公司的VERTEX70型傅里叶变换红外光谱仪和SPECAC公司的衰减全反射样品测量附件,ATR样品槽为ZnSe晶体,45°入射角,3次全反射。光谱采集范围为 $4500\sim 550\text{ cm}^{-1}$,采用氘化三甘氨酸硫酸酯(DTGS)探测器进行检测,分束器为KBr,扫描次数为16次,分辨率为 4 cm^{-1} ,实验温度为 $(25\pm 1)\text{ }^{\circ}\text{C}$,相对湿度为46%。

衰减全反射傅里叶变换红外光谱仪简称FTIR-ATR,其ATR附件根据光内反射原理而设计。红外光从光源发出后经过折射率较大的晶体再投射到折射率较小的试样表面上,当入射角大于临界角时,入射光线就会发生全反射。实际上红外光并不是全部被反射回来,而是穿透到试样表面内一定深度后再返回表面。在这个过程中,试样在入射光频率区域内有选择性地吸收,反射光强度减弱,产生与透射吸收相类似的谱图,从而获得样品表层化学成份的结构信息^[18-19]。

采集光谱时取0.075 mL样品置于样品槽中,用衰减全反射方式获得样品的FTIR吸收光谱,每个样品重复采集3次光谱,计算平均光谱。

2.2 实验样品及其制备

实验材料用特级橄榄油样品 28 个、大豆油样品 12 个以及混合油样品 63 个,共 103 个样品。特级橄榄油和大豆油的纯油样品均由当地质量技术监督局推荐购买,购买于当地超市:百佳、好又多、家乐福、吉之岛。纯油样品详细信息为:大豆油的 3 个品牌分别为福临门(FLM)、萨尔瓦多里(SEW)、金龙鱼(JLY),产地有广西钦州、意大利、广东广州;橄榄油的 7 个品牌分别为红岛(RD)、品利(PL)、金城堡(JCB)、拉维萨斯(LWSS)、迈萨维诺(MS)、福蓓娜(FBN)、亿芭利(YBL),产地有澳大利亚、西班牙、意大利、希腊。

混合油样品按一种品牌的橄榄油与一种品牌的大豆油进行配比,橄榄油体积分数分别为 EVOO(25%)、EVOO(50%)、EVOO(75%)。混合油样品信息如表 1 所示。

表 1 混合油样品
Table 1 Oil blend samples

Number	Blend samples	Number	Blend samples	Number	Blend samples
41	RD (50%)+FLM (50%)	62	RD (25%)+FLM (75%)	83	RD (75%)+FLM (25%)
42	RD (50%)+SEW (50%)	63	RD (25%)+SEW (75%)	84	RD (75%)+SEW (25%)
43	RD (50%)+JLY (50%)	64	RD (25%)+JLY (75%)	85	RD (75%)+JLY (25%)
44	PL (50%)+FLM (50%)	65	PL (25%)+FLM (75%)	86	PL (75%)+FLM (25%)
45	PL (50%)+SEW (50%)	66	PL (25%)+SEW (75%)	87	PL (75%)+SEW (25%)
46	PL (50%)+JLY (50%)	67	PL (25%)+JLY (75%)	88	PL (75%)+JLY (25%)
47	JCB (50%)+FLM (50%)	68	JCB (25%)+FLM (75%)	89	JCB (75%)+FLM (25%)
48	JCB (50%)+SEW (50%)	69	JCB (25%)+SEW (75%)	90	JCB (75%)+SEW (25%)
49	JCB (50%)+JLY (50%)	70	JCB (25%)+JLY (75%)	91	JCB (75%)+JLY (25%)
50	LWSS(50%)+FLM(50%)	71	LWSS(25%)+FLM (75%)	92	LWSS (75%)+FLM (25%)
51	LWSS (50%)+SEW (50%)	72	LWSS (25%)+SEW (75%)	93	LWSS (75%)+SEW (25%)
52	LWSS (50%)+JLY (50%)	73	LWSS (25%)+JLY (75%)	94	LWSS (75%)+JLY (25%)
53	MS (50%)+FLM (50%)	74	MS (25%)+FLM (75%)	95	MS (75%)+FLM (25%)
54	MS (50%)+SEW (50%)	75	MS (25%)+SEW (75%)	96	MS (75%)+SEW (25%)
55	MS (50%)+JLY (50%)	76	MS (25%)+JLY (75%)	97	MS (75%)+JLY (25%)
56	FBN (50%)+FLM (50%)	77	FBN (25%)+FLM (75%)	98	FBN (75%)+FLM (25%)
57	FBN (50%)+SEW (50%)	78	FBN (25%)+SEW (75%)	99	FBN (75%)+SEW (25%)
58	FBN (50%)+JLY (50%)	79	FBN (25%)+JLY (75%)	100	FBN (75%)+JLY (25%)
59	YBL (50%)+FLM (50%)	80	YBL (25%)+FLM (75%)	101	YBL (75%)+FLM (25%)
60	YBL (50%)+SEW (50%)	81	YBL (25%)+SEW (75%)	102	YBL (75%)+SEW (25%)
61	YBL (50%)+JLY (50%)	82	YBL (25%)+JLY (75%)	103	YBL (75%)+JLY (25%)

2.3 模式识别方法原理

2.3.1 Fisher 判别分析

判别分析是根据表明事物特点的变量值和它们所属的类,利用判别函数对未知所属类别的事物进行分类的一种分析方法。Fisher 判别分析的思想是用已知类别的训练样品的观察指标借助方差分析的思想来建立一个或多个判别函数,确定判别函数的系数,原则是使得组间的区别最大而每个组内部的离差最小;有了判别函数以后,将样品测试数据代入判别函数,得出判别得分,从而确定该样品属于哪一类。

应用过程中,通过方差分析构造的判别函数可以把 P 维空间中的已知类别总体以及求知类别归属的样本都变换为一维数据,即要求能够最大限度地缩小每个组内部各个样本点的差异,又能够最大限度地扩大组间各个样本点之间的差异,这样才可能获得较高的判别效率^[20]。

2.3.2 多层感知器神经网络

神经网络是对人脑以及自然神经网络的投影,是一种非线性的数据建模工具集合,由大量神经元组成,神经元之间的连接赋予相关的权重,训练算法在迭代过程中不断调整这些权重,从而使得预测误差最小化并给出预测精度。通过对样本模式的学习模拟信息之间的机制,使用数据对网络进行训练,找到输入与输出之间的关系,实现输入与输出间的高度非线性映射。神经网络广泛应用于模式识别、优化计算、函数逼

近、食品科学等领域^[21]。

本研究采用的多感知器神经网络是一种有监督的前馈神经网络结构,即所有数据均通过输入层、隐藏层最后进入输出层。计算过程中,输入层数值需要通过激活函数计算得到新值进入隐藏层;同样,输出层也有激活函数。这些函数表示输入参数权重的总和,而权重值是由计算的具体函数确定的。隐藏层采用双曲正切作为激活函数,其表达式为

$$\gamma(c) = \tanh(c) = \frac{\exp(c) - \exp(-c)}{\exp(c) + \exp(-c)}, \quad (1)$$

式中 c 为对应某个隐节点激活函数的数据中心。输出层采用的激活函数是 Softmax 函数:

$$\gamma(c) = \frac{1}{1 + \exp(-c)}, \quad (2)$$

协变量的重标度方法为标准化。

本研究采用 Matalab R2010a、SPSS 19.0 软件,对样品的光谱数据进行处理和分析。

3 结果与讨论

3.1 光谱分析

3.1.1 原始光谱分析

橄榄油和大豆油以及混合油样品的 FTIR-ATR 光谱如图 1 所示。由图 1 可知,在 625~4000 cm^{-1} 范围内样品的 FTIR-ATR 光谱非常相似,通过肉眼难以识别。表现在 3008 cm^{-1} 处有不饱和碳链的 C—H 伸缩振动峰,而在 2924 cm^{-1} 和 2854 cm^{-1} 处有饱和碳链 C—H 反对称和对称伸缩振动峰;可以观察到在 1745 cm^{-1} 处有一个最强的吸收峰主要源于甘油三酯 C=O 伸缩振动峰;在 1654 cm^{-1} 附近有一个相对比较弱的吸收峰主要对应于顺式烯烃 C=C 伸缩振动峰;1400~1200 cm^{-1} 波段主要源于脂肪族的亚甲基和甲基 C—H 面内弯曲振动峰;1125~1095 cm^{-1} 波段对应甘油三酯中 C—O 伸缩振动峰和亚甲基 C—H 弯曲振动峰;最后,在低于 1000 cm^{-1} 的波段内,723 cm^{-1} 附近的吸收峰由 $(\text{CH}_2)_n$ 面内摇摆振动和 C—H 面外弯曲振动吸收叠加所引起^[22]。原始光谱分析结果表明:纯橄榄油和纯大豆油以及三种掺杂有大豆油的橄榄油样本红外光谱出现的特征吸收峰、峰位和峰形几乎相同,说明它们的主要组成成分几乎是相同的。

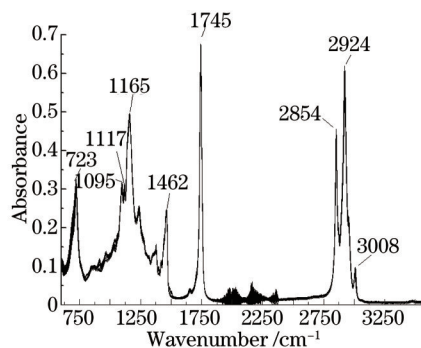


图 1 纯橄榄油、纯大豆油和混合油的 FTIR-ATR 光谱

Fig.1 FTIR-ATR spectra of pure olive oil, pure soybean oil and oil blend samples

3.1.2 预处理后光谱的比较

获取的信息中除了含有待测样品的原始信息外,还包含各种外在的干扰信息,这些噪声信息会导致测得的数值和真实值之间存在一定差异。为了尽可能消除误差,必须运用各种数据处理方法来减弱甚至除去各种干扰因素的影响,为下一步的数据处理奠定基础。因此,对原始光谱进行预处理是非常必要的。本研究采用一阶导数、二阶导数、多元散射校正以及小波去噪的方法对原始光谱进行预处理。

原始光谱进行一阶导数、二阶导数、多元散射校正以及小波去噪预处理后得到的光谱如图 2 所示。橄榄油和大豆油以及混合油样品的光谱特征不明显,难以识别,预处理后光谱比较发现,经一阶导数、二阶导数、多元散射校正处理后的光谱,在 1800~2500 cm^{-1} 波段的噪声仍然很大,而经过小波去噪处理后,此波段变得比较平滑。小波去噪方法噪声减弱的效果非常明显。

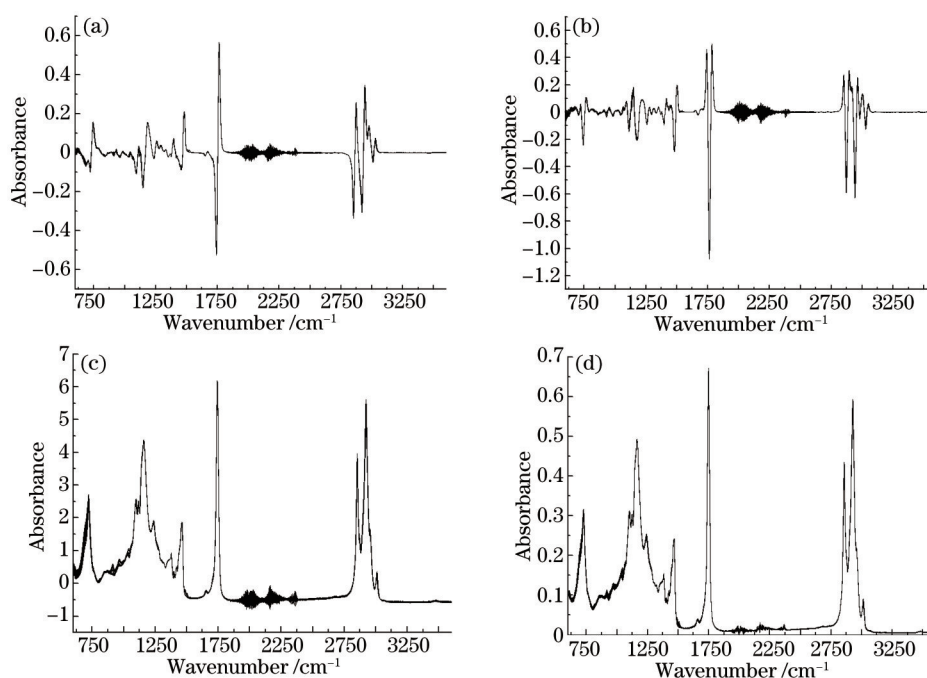


图2 (a) 一阶导数光谱; (b) 二阶导数光谱; (c) 多元散射校正光谱; (d) 小波变换后的光谱

Fig.2 (a) First derivative spectra; (b) second derivative spectra; (c) multiple scattering correction spectra; (d) wavelet domain denoising spectra

结果表明:仅凭ATR原始光谱以及预处理后的光谱无法将5种不同的样品区分开来,因此,本研究进一步结合模式识别的方法对光谱数据进行处理,采用PCA对光谱数据进行降维,再结合Fisher判别分析和多层感知器神经网络的方法对纯橄榄油、纯大豆油以及橄榄油中掺有大豆油的混合油样本进行鉴别,能突出油脂样品之间化学成分的细小差异,从而达到鉴别的目的。

3.2 基于Fisher判别模型的鉴别

3.2.1 不同预处理方法对Fisher判别模型的影响

在判别分析中,数据的选择和数据的预处理是判别分析成功与否的重要因素,为消除样品不均、光散射等因素的影响,对原始光谱数据分别进行一阶微分、二阶微分、多元散射校正以及小波去噪等预处理,处理后的光谱数据分别进行Fisher判别,比较不同的预处理方法对判别效果的影响,如表2所示。与原始数据进行Fisher判别分析比较,预处理后初始识别率都有明显的提高,从表中还可以发现,对预处理后的数据进行Fisher判别分析时,小波去噪的效果最佳,初始识别率为100.00%,交叉验证识别率为97.10%。因此,优选小波去噪的方法对原始光谱数据进行预处理。

表2 不同方法预处理后Fisher判别的识别率

Table 2 Recognition rate of Fisher discriminant under different pretreatment methods

Processing method	Initial recognition rate /%	Cross validation recognition rate /%
Original spectra	94.20	94.20
1-Derivative	97.10	92.20
2-Derivative	97.10	94.20
Multiple scattering correction	96.10	94.20
Wavelet domain denoising	100.00	97.10

3.2.2 建立Fisher判别函数

采用小波去噪后的光谱数据,利用PCA对光谱数据进行降维,取前4个主成分做Fisher判别分析,建立5种不同样品的鉴别模型,得到4个标准化的典型判别式函数为

$$u_1(x) = 0.023x_1 - 0.275x_2 + 1.439x_3 + 1.203x_4, \quad (3)$$

$$u_2(x) = 0.326x_1 + 1.041x_2 + 0.026x_3 + 0.333x_4, \quad (4)$$

$$u_3(x) = 0.299x_1 - 0.154x_2 - 0.250x_3 + 0.773x_4, \quad (5)$$

$$u_4(x) = 1.004x_1 + 0.126x_2 + 0.120x_3 - 0.051x_4, \quad (6)$$

式中 x_1, x_2, x_3, x_4 为各变量: 主成分。

4个函数的累积贡献率如表3所示。

表3 前4个判别函数的贡献率和累计贡献率

Table 3 Variance and cumulative variance of the first four functions

Function	Function 1	Function 2	Function 3	Function 4
Variance /%	79.2	13.9	6.8	0.1
Cumulative variance /%	79.2	93.1	99.9	100.0

针对这4个判别函数进行有效性检验,如表4所示。Wilk's Lambda表示类内离差交叉乘积矩阵行列式与总离差交叉乘积矩阵行列式的比值,间接地检验判别函数的显著性水平,其值越小判别能力越强。1-4表示没有函数被移除,其Wilk's Lambda值为0.003,2-4表示第一个判别函数被移除,其Wilk's Lambda值为0.071,3-4表示前两个判别函数被移除,其Wilk's Lambda值为0.339,4表示前3个判别函数被移除,其Wilk's Lambda值为0.976。因此可以看出前4个函数联合判别能力最强,第4个函数判别能力最弱。1-4、2-4、3-4的显著性水平均为0.000,小于0.050,说明前3个函数的判别能力是显著的。

表4 判别函数的有效性检验

Table 4 Efficacy of discriminant function

Function test	Wilk's Lambda	Chi-square test	Confidence level
1-4	0.003	563.117	0.000
2-4	0.071	258.436	0.000
3-4	0.339	105.376	0.000
4	0.976	2.368	0.124

3.2.3 Fisher判别函数分析结果

橄榄油和大豆油以及掺有不同体积分数大豆油的橄榄油等5种不同样品的初始案例和交叉验证判别分析结果如表5所示。

表5 判别分析结果

Table 5 Discriminant analysis results

		Prediction group					Total	
		Soybean oil	Olive oil	EVOO (25%)	EVOO (50%)	EVOO (75%)		
Training set	Calculation result	Soybean oil	12	0	0	0	0	12
		Olive oil	0	28	0	0	0	28
		EVOO(25%)	0	0	21	0	0	21
		EVOO(50%)	0	0	0	21	0	21
		EVOO(75%)	0	0	0	0	21	21
	Percent	Soybean oil	100.0	0	0	0	0	100.0
		Olive oil	0	100.0	0	0	0	100.0
		EVOO(25%)	0	0	100.0	0	0	100.0
		EVOO(50%)	0	0	0	100.0	0	100.0
		EVOO(75%)	0	0	0	0	100.0	100.0
Cross validation set	Calculation result	Soybean oil	12	0	0	0	0	12
		Olive oil	1	26	1	0	0	28
		EVOO(25%)	0	0	20	0	1	21
		EVOO(50%)	0	0	0	21	0	21
		EVOO(75%)	0	0	0	0	21	21
	Percent	Soybean oil	100.0	0	0	0	0	100.0
		Olive oil	3.6	92.9	3.6	0	0	100.0
		EVOO(25%)	0	0	95.2	0	4.8	100.0
		EVOO(50%)	0	0	0	100.0	0	100.0
		EVOO(75%)	0	0	0	0	100.0	100.0

初始案例分析 103 个样品判别分析结果为:橄榄油和大豆油以及不同体积分数的橄榄油等 5 种不同样品全部被正确分类,总的判别正确率为 100%。

交叉验证案例 103 个样品判别分析结果为:橄榄油样品中有两个被误判,准确率为 92.9%;橄榄油体积分数为 25%的样品中有一个被误判,准确率为 95.2%;其他 3 种样品的准确率都为 100%。总的判别准确率为 97.1%。

分析结果说明建立的判别模型是可靠的。5 种不同样品的判别方程如下:

大豆油:

$$Y_1(x) = -35393421.20 + 70786830.61x_1 + 505478.652x_2 + 277671.423x_3 + 1400837.279x_4, \quad (7)$$

橄榄油:

$$Y_2(x) = -35388395.37 + 70781820.24x_1 + 504731.420x_2 + 276050.699x_3 + 1399401.134x_4, \quad (8)$$

EVOO(50%):

$$Y_3(x) = -35401425.39 + 70794858.22x_1 + 505515.554x_2 + 273601.998x_3 + 1398444.035x_4, \quad (9)$$

EVOO(25%):

$$Y_4(x) = -35394532.17 + 70787957.80x_1 + 505739.849x_2 + 270958.075x_3 + 1393513.477x_4, \quad (10)$$

EVOO(75%):

$$Y_5(x) = -35384718.73 + 70778139.05x_1 + 505290.250x_2 + 270279.909x_3 + 1392423.237x_4. \quad (11)$$

为了直观地观察类内聚类情况和类间距离大小,由表 4 可知前两个函数的累积贡献率已经达到 93.1%。因此,根据前两个典型判别函数得分绘制二维散点图,如图 3 所示。从图 3 两个判别函数特征空间分布可以看出,第一到第四象限分别为 EVOO(25%)、EVOO(50%)和 EVOO(75%)、橄榄油、大豆油。橄榄油和大豆油的距离最远、区分最明显,橄榄油体积分数为 25%的样品与纯橄榄油样品之间的距离次之,随着橄榄油体积分数的增加,其距离也越来越靠近橄榄油,EVOO(25%)和 EVOO(50%)两种样品的距离最近,鉴别效果不明显。

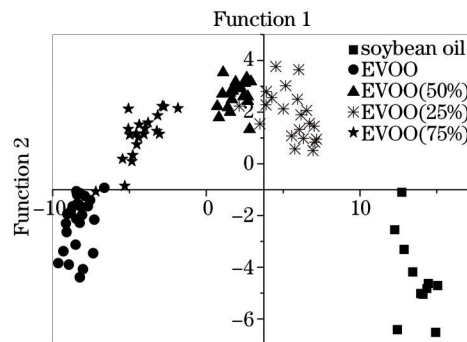


图 3 判别函数得分的二维散点图

Fig.3 Score plots of discriminant functions

3.3 基于多层感知器神经网络的鉴别

3.3.1 多层感知器分类过程

纯橄榄油、纯大豆油以及 3 种掺杂有大豆油的橄榄油样本的多层感知器神经网络鉴别过程为:在原始数据分别进行一阶导数、二阶导数、多元散射校正和小波去噪预处理后的基础上进行主成分分析,对光谱数据进行降维,取前 4 个主成分代表预处理后光谱的特征信息作为多层感知器神经网络的输入;将目标类型["10000"代表大豆油、“01000”代表橄榄油、“00100”代表 EVOO(25%)、“00010”代表 EVOO(50%)、“00001”代表 EVOO(75%)]作为神经网络的输出;将样本按 7:3 的比例进行建模,选择 76 个样品作为训练集来训练网络,选择 27 个样品作为测试集,用训练好的网络对 5 种样本进行鉴别,根据网络的输出值确定每个样本所属的目标类型。

3.3.2 不同预处理方法对模型的影响及分析结果

通过表 6 对比不同预处理方法对模型的影响,结果表明,4 种预处理方法中对原始光谱数据进行小波去噪后,结合 PCA 进行数据降维,对纯大豆油和纯橄榄油以及 3 种掺杂有大豆油的橄榄油这 5 种不同的样品进行多层感知器网络的分析,得到的鉴别结果最佳。

表6 不同的预处理方法对神经网络分类的影响

Table 6 Influence of different pretreatment methods on neural network classification accuracy

Processing method	Training set /%	Test set /%
Original spectra	97.2	90.0
1- Derivative	60.6	53.1
2- Derivative	92.1	77.8
Multiple scattering correction	97.4	100.0
Wavelet domain denoising	100.00	100.0

表7所示为小波去噪后的神经网络分析结果:大豆油、橄榄油、EVOO(25%)、EVOO(50%)、EVOO(75%) 5种样品无论是训练集还是测试集,分类正确率都为100%。

表7 神经网络的分类结果

Table 7 Neural network identification results

Sample	Soybean oil	Olive oil	EVOO(25%)	EVOO(50%)	EVOO(75%)	Accuracy /%
Training set	Soybean oil	8	0	0	0	100
	Olive oil	0	20	0	0	100
	EVOO(25%)	0	0	18	0	100
	EVOO(50%)	0	0	0	15	100
	EVOO(75%)	0	0	0	0	15
Test set	Soybean oil	4	0	0	0	100
	Olive oil	0	8	0	0	100
	EVOO(25%)	0	0	3	0	100
	EVOO(50%)	0	0	0	6	100
	EVOO(75%)	0	0	0	0	6

4 结 论

采用 FTIR-ATR 光谱结合 PCA 与 Fisher 判别和多层感知器网络建立了橄榄油中掺杂大豆油的鉴别模型。结果表明:PCA 结合 Fisher 判别分析建立的判别模型原始判别的准确率达到 100%,交叉验证的准确率达到 97.1%;多层感知器神经网络鉴别模型训练集分类准确率为 100%,测试集分类准确率为 100%。说明运用 ATR 光谱法可以准确、快速地对初榨特级橄榄油的掺杂进行定性鉴别,检测结果都比较理想。本文提出的研究方法快速、简便,为食用油的快速检验提供了一种新的技术手段。

参 考 文 献

- 1 Yu Changqing. Olive oil chemical composition and nutritional value to human body[J]. Food Science and Technology, 2000, (2): 59-60.
于长青. 橄榄油的化学组成及对人体的营养价值[J]. 食品科技, 2000, (2): 59-60.
- 2 G Flores, M L R del Castillo, G P Blanch, *et al.*. Detection of the adulteration of olive oils by solid phase microextraction and multidimensional gas chromatography[J]. Food Chemistry, 2006, 97(2): 336-342.
- 3 L Cercaci, M T Rodriguez-Estrada, G Lercker. Solid-phase extraction-thin-layer chromatography-gas chromatography method for the detection of hazelnut oil in olive oils by determination of esterified sterols[J]. Journal of Chromatography A, 2003, 985: 211-220.
- 4 D Zabarar, M H Gordon. Detection of pressed hazelnut oil in virgin olive oil by analysis of polar components: Improvement and validation of the method[J]. Food Chemistry, 2004, 84(3): 475-483.
- 5 M J Lerma-García, E F Simó-Alfonso, A Méndez, *et al.*. Classification of extra virgin olive oils according to their genetic variety using linear discriminant analysis of sterol profiles established by ultra-performance liquid chromatography with mass spectrometry detection[J]. Food Research International, 2011, 44(1): 103-108.
- 6 K I Poulli, G A Mousdis, C A Georgiou. Rapid synchronous fluorescence method for virgin olive oil adulteration assessment [J]. Food Chemistry, 2007, 105(1): 369-375.

- 7 C C Oliveros, J L P Pavón, C G Pinto, *et al.*. Electronic nose based on metal oxide semiconductor sensors as a fast alternative for the detection of adulteration of virgin olive oils[J]. *Analytica Chimica Acta*, 2002, 495(2): 219–228.
- 8 D Smejkalova, A Piccolo. High-power gradient diffusion NMR spectroscopy for the rapid assessment of extra-virgin olive oil adulteration[J]. *Food Chemistry*, 2010, 118(1):153–158.
- 9 Xu Xuebing. *Chemistry of Fats and Oils*[M]. Beijing: China Commercial Press, 1993: 1–17.
徐学兵. 油脂化学[M]. 北京: 中国商业出版社, 1993: 1–17.
- 10 Zhao Jiewen, Bi Xiakun, Lin Hao, *et al.*. Visible-near-infrared transmission spectra for rapid analysis of the freshness of eggs[J]. *Laser & Optoelectronics Progress*, 2013, 50(5): 053003.
赵杰文, 毕夏坤, 林 颢, 等. 鸡蛋新鲜度的可见-近红外透射光谱快速识别[J]. 激光与光电子学进展, 2013, 50(5): 053003.
- 11 Wang Wei, Liu Wenqing, Zhang Tianshu. Continuous field measurements of stable isotopes in atmospheric water vapor by FTIR spectrometry[J]. *Acta Optica Sinica*, 2014, 34(1): 0130003.
王 薇, 刘文清, 张天舒. 利用傅里叶变换红外光谱技术连续测量环境大气中水汽的稳定同位素[J]. 光学学报, 2014, 34(1): 0130003.
- 12 Ye Weilin, Zheng Chuantao, Wang Yiding. Stability measurement and temperature compensation of mid-infrared methane detection device[J]. *Acta Optica Sinica*, 2014, 34(3): 0323003.
叶玮琳, 郑传涛, 王一丁. 中红外甲烷检测仪稳定性实验及温度补偿研究[J]. 光学学报, 2014, 34(3): 0323003.
- 13 Wu Jiang, Huang Furong, Huang Caihuan, *et al.*. Study on near infrared spectroscopy of transgenic soybean identification based on principal component analysis and neural network[J]. *Spectroscopy and Spectral Analysis*, 2013, 33(6): 1537–1541.
吴 江, 黄富荣, 黄才欢, 等. 近红外光谱结合主成分分析和BP神经网络的转基因大豆无损鉴别研究[J]. 光谱学与光谱分析, 2013, 33(6): 1537–1541.
- 14 Tao Sijia, Li Menghua, Li Jingming, *et al.*. Application of infrared spectroscopy combined with Bayesian information fusion for identification of different original wines[J]. *Chinese Journal of Analytical Chemistry*, 2014, 42(2): 215–220.
陶思嘉, 李梦华, 李景明, 等. 红外光谱及Bayes信息融合技术的葡萄酒鉴别研究[J]. 分析化学, 2014, 42(2): 215–220.
- 15 Yang Benzhi, Pan Qiuyue, Pan Danjie. Rapid determination of soybean oil peroxide value by FTIR/ATR[J]. *Journal of the Chinese Cereals and Oils Association*, 2013, 28(4): 108–112.
杨本志, 潘秋月, 潘丹杰. 利用 FTIR/ATR 快速测定大豆油的过氧化值[J]. 中国粮油学报, 2013, 28(4): 108–112.
- 16 Zhang Lan, Zhang Qingyue, Nan Qun. Application of mid-infrared spectrum in non-invasive blood glucose measurement [J]. *Optics & Optoelectronic Technology*, 2008, 6(2): 81–84.
张 兰, 张清悦, 南 群. 中红外光谱法在无创血糖检测技术中的应用[J]. 光学与光电技术, 2008, 6(2): 81–84.
- 17 Huang Furong, Luo Yunhan, Zheng Shifu, *et al.*. Near-infrared spectroscopic analysis and model optimization on cholesterol and triglyceride in whole blood[J]. *Acta Optica Sinica*, 2011, 31(10): 1030001.
黄富荣, 罗云瀚, 郑仕富, 等. 全血胆固醇、甘油三酯近红外光谱分析与模型优化[J]. 光学学报, 2011, 31(10): 1030001.
- 18 Zhang Peng, Zhang Zhihui. Rapidly changing phase error correction of Fourier transform spectrometer[J]. *Chinese J Lasers*, 2012, 39(s1): s115002.
张 鹏, 张志辉. 傅里叶变换光谱仪非渐变相位误差的校正方法[J]. 中国激光, 2012, 39(s1): s115002.
- 19 Meng Qinghua, Xiang Yang. Infrared horizontal ATR instrument used in Fourier transform spectroscopy[J]. *Optics and Precision Engineering*, 2007, 15(10): 1515–1519.
孟庆华, 向 阳. 傅里叶变换光谱仪用红外水平衰减全反射测试仪[J]. 光学精密工程, 2007, 15(10): 1515–1519.
- 20 Ma Guoqing. *Statistics Management*[M]. Beijing: Science Press, 2007: 346.
马国庆. 管理统计[M]. 北京: 科学出版社, 2007: 346.
- 21 F H F Leung, H K Lam, S H Ling, *et al.*. Tuning of the structure and parameters of a neural network using an improved genetic algorithm[J]. *IEEE Transactions on Neural Networks*, 2003, 14(1): 79–88.
- 22 M J Lerma-García, G Ramis-Ramos, J M Herrero-Martínez, *et al.*. Authentication of extra virgin olive oils by Fourier-transform infrared spectroscopy[J]. *Food Chemistry*, 2010, 118 (1): 78–83.

栏目编辑: 吴秀娟