

基于嗅觉可视化技术的食用植物油分类识别

杨干, 李大鹏, 文韬, 蒋涵, 龚中良

(中南林业科技大学机电工程学院, 长沙 410004)

摘要:为实现山茶油与3种常见食用植物油(菜籽油、大豆油和玉米油)的区分,制备可视化传感器阵列,采用嗅觉可视化技术对4种不同种类的食用植物油进行分类识别。采用主成分分析(PCA)对4种油样的特征数据进行降维,然后将降维后的数据导入K近邻(KNN)、极限学习机(ELM)、支持向量机(SVM)3种分类模型中进行模型参数优化,对比了3种分类模型的结果。结果表明:建立的SVM分类模型性能最优,当输入主成分向量数为7、 $c=1.7411$ 、 $g=4.5498$ 时,SVM分类模型的测试集分类识别准确率为95.8%,五折交叉验证准确率为89.6%。制得的可视化传感器阵列可以实现4种食用植物油的分类识别,嗅觉可视化技术用于食用植物油检测是可行的。

关键词:嗅觉可视化;食用植物油;分类识别;支持向量机

中图分类号:TS201.7;TS227 文献标识码:A 文章编号:1003-7969(2023)11-0107-05

Classification and recognition of edible vegetable oils based on olfactory visualization technology

YANG Gan, LI Dapeng, WEN Tao, JIANG Han, GONG Zhongliang

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

Abstract: In order to distinguish oil-tea camellia seed oil from three common edible vegetable oils (rapeseed oil, soybean oil and corn oil), visual sensor array was prepared, and four different edible vegetable oils were classified and identified by olfactory visualization technology. Principal component analysis (PCA) was used to reduce the dimension of the characteristic data of the four oil samples. The data after PCA dimensionality reduction was imported into three classification models namely K-Nearest Neighbor (KNN), Extreme Learning Machine (ELM), and Support Vector Machine (SVM), and the model parameters were optimized, and the classification results of the three classification models were compared. The results showed that the established SVM classification model had the best performance. When the number of input principal component vectors was 7, $c=1.7411$, and $g=4.5498$, the classification and recognition accuracy of the test set of the SVM classification model was 95.8%, and the 5-fold validation accuracy was 89.6%. The visual sensor array can achieve the classification and recognition of four edible vegetable oils, and the olfactory visualization technology is feasible for the classification and identification of edible vegetable oils.

Key words: olfactory visualization; edible vegetable oil; classification and recognition; Support Vector Machine

收稿日期:2022-09-13;修回日期:2023-07-27

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(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:杨干(2000),男,在读硕士,研究方向为林产品无损品质检测技术及装备(E-mail)2546274641@qq.com。

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

食用植物油是人们生活中的必需品,不仅可用于各种食材的烹饪,还可用于养生保健^[1]。不同种类食用植物油的组成成分十分相似,但其含量差异较大^[2],营养成分含量不同导致不同种类的食用植

物油价格差异很大,这就驱使不法商贩利用价格差异以次充好来牟取暴利。因此,探究一种能快速检测食用油种类的方法是十分必要的。

目前,用于食用植物油品质检测的方法主要是气相色谱-质谱法(GC-MS),其优点是检测结果准确^[3-4],但同时也有耗时长、仪器昂贵、操作复杂等缺点。近年来,有许多学者采用光谱识别技术^[5-7]与电子鼻技术^[8-9]来对食用植物油进行分类识别,这些技术虽然具有快速、无损等优点,但大都存在仪器昂贵、检测结果受环境影响等问题。嗅觉可视化技术最早由 Rakow 等^[10]于 2000 年提出,2004 年该研究团队首次运用嗅觉可视化技术实现了多种挥发性气体的区分^[11]。相比于其他检测技术,嗅觉可视化技术具有检测快捷、成本低廉等优点。为了有效地分析可视化阵列传感器高维的差分数据,研究人员采用多变量分析方法来分析数据集,主要包括主成分分析(PCA)^[10-15]、K 近邻(KNN)^[12-13]、支持向量机(SVM)^[12-16]、线性判别分析(LDA)^[12, 17]和极限学习机(ELM)^[18]等方法。目前,嗅觉可视化技术已经逐渐应用于酒类识别^[12]、食品贮藏期和发酵程度的判别^[13-14]、肉类新鲜度检测^[15-16]、食醋品质检测^[17-18]等食品检测领域,但有关嗅觉可视化技术对食用植物油的检测还鲜有报道。

山茶油又名油茶籽油,是我国特有的木本油脂,其油酸含量高,且富含角鲨烯、维生素 E、多酚类化合物,营养丰富^[1]。为规范山茶油市场,有必要对其与其他低价植物油进行区分。

本研究制作了用于食用植物油分类识别的可视

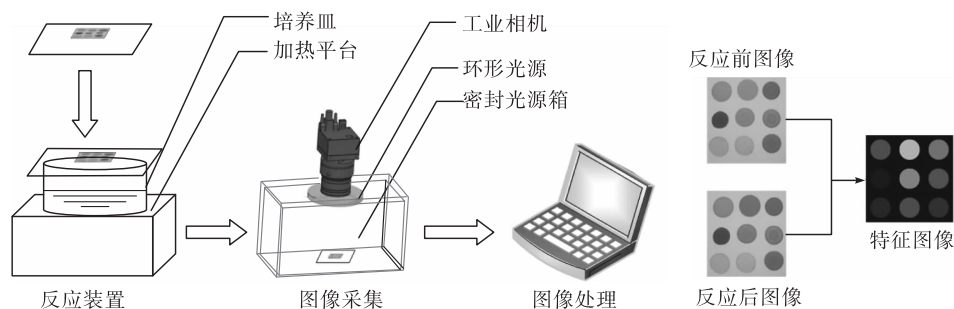


图1 嗅觉可视化试验装置和流程

1.2 试验方法

1.2.1 传感器阵列制备

经过前期研究,筛选出了 9 种色敏材料用于制作 3×3 可视化传感器阵列,选择聚偏氟乙烯(PVDF)膜作为基底材料。将卟啉及其衍生物与酸碱指示剂分别溶于二氯甲烷及无水乙醇中,制成质量浓度为 1 g/L 的溶液。制作可视化传感器阵列时,首先采用微量进样器量取 $1 \mu\text{L}$ 的色敏材料溶液

化传感器阵列,筛选了试验过程中的加热温度,并选用山茶油、大豆油、菜籽油、玉米油 4 种食用植物油进行嗅觉可视化技术识别食用植物油类别的试验。然后将 4 种食用植物油的差分图像数据采用 PCA 降维后导入 KNN、ELM、SVM 3 种分类模型中进行分类识别,以验证嗅觉可视化技术用于食用植物油分类的可行性。

1 材料与方法

1.1 试验材料

山茶油,购自湖南大三湘有限公司;大豆油、菜籽油、玉米油,均购自长沙当地的超市。

2,3,7,8,12,13,17,18-八乙基-21H,23H-卟吩氯化锰、5,10,15,20-四苯基-21H,23H-卟吩氯化铁(III)、5,10,15,20-四(4-甲氧苯基)-21H,23H-卟吩钴、5,10,15,20-四苯基-21H,23H-卟吩钴(II),购于德国默克公司;9 种色敏材料(包含 4 种卟啉及其衍生物和 5 种酸碱指示剂)、甲酚红、甲基红、中性红、溴甲酚绿、溴酚蓝,购于国药集团化学试剂有限公司;二氯甲烷和无水乙醇等均为分析纯。

嗅觉可视化试验装置如图 1 所示,由反应装置、图像采集装置和图像处理部分构成。反应装置包括加热平台和盛有油样的培养皿,加热平台用于保证油样的加热温度恒定;工业相机、环形光源、密封光源箱组成图像采集装置,其中环形光源用于提供均匀的光照条件,密封光源箱用于隔绝外部光照的干扰;图像处理部分运用了 C++ 语言调用 OpenCV 函数库,实现了特征区域分割和图像差值的提取。

滴定在 $3 \text{ cm} \times 3 \text{ cm}$ 的 PVDF 膜上制成 3×3 传感器阵列,然后将传感器阵列放入通风橱中挥发 40 min,待溶液挥发完毕后,立即将传感器阵列用于试验。

1.2.2 嗅觉可视化试验检测流程

嗅觉可视化试验的试验流程如图 1 所示。首先在密闭的光源箱内拍摄传感器阵列的初始图片,然后将传感器阵列粘贴在玻璃盖板上,并将盖板覆于盛有 10 mL 油样的培养皿上。将培养皿放置在加热

平台,在 100 ℃ 下加热 20 min,加热完毕后立即取下传感器阵列,拍摄反应后的图片,最后将反应前后的图片导入计算机中进行图像处理。

1.2.3 图像处理与差分图像数据提取

为了避免色敏材料中心点样点与图像边缘区域的影响,图像处理部分选择了去除色敏材料图像的中心点样点与边缘区域的环形 ROI 区域作为特征提取区域。将该特征区域反应前后的 R 、 G 、 B 平均值分别相减,得到特征差分图像数据。由于一个传感器阵列共有 9 种色敏材料,每种色敏材料有 R 、 G 、 B 3 个特征差值,所以 1 次试验可以得到由 27 个特征差值组成的特征图像数据。

1.2.4 数据处理

本研究选用主成分分析进行差分图像数据的降维处理,建立 KNN、ELM、SVM 3 种分类模型,优化 3 种分类模型的关键参数。采用 Matlab 2018b 软件来处理试验数据和建立模型。

2 结果与分析

2.1 加热温度选择

前期试验发现,在室温下油样的挥发性较小,需要对油样进行加热处理使油样快速挥发。考虑到试验设计的合理性与 PVDF 膜的耐热性,选择 25、50、75、100 ℃ 4 个加热温度梯度,对 4 种油样进行试验,为了得到更为清晰的差分图像,对原始差分图像数据进行了 5 倍放大处理。不同加热温度下 4 种油样的特征差分图像如图 2 所示。

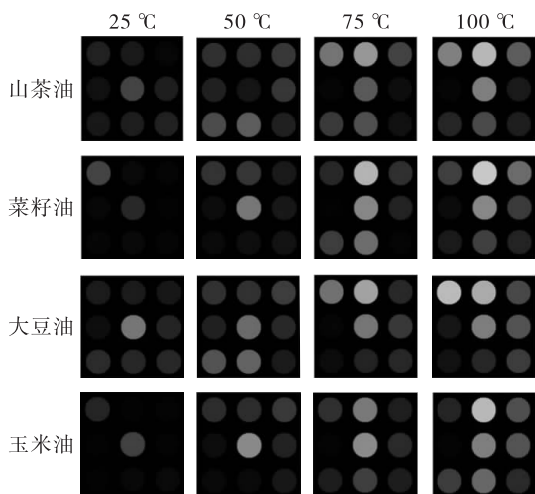


图 2 不同加热温度下 4 种油样的特征差分图像

从图 2 可以看出,加热温度对色敏材料的变色有很大的影响。就山茶油而言,当加热温度为 25 ℃ 时,仅有少部分的色敏材料有轻微变色,当加热温度达到 50 ℃ 时,变色的色敏材料种类略有增加,当加热温度达到 75 ℃ 时,几乎所有色敏材料都发生了变色,加热温度达到 100 ℃ 时,色敏材料的变色程度较

加热 75 ℃ 时进一步加深。总体而言,随着加热温度的升高,变色的色敏材料种类和变色程度都在增加,并且色敏材料在与油样接触前后颜色变化的程度均在加热温度为 100 ℃ 时达到最大。因此,在后续试验中选择 100 ℃ 作为加热温度。

2.2 嗅觉可视化对不同食用油种类的识别

2.2.1 主成分分析结果

本试验的试验样本共 120 份油样,每种油样 30 份,每份油样可得到由 27 个 R 、 G 、 B 差值组成的特征图像数据,所以一共得到了 120 × 27 维度的试验数据。对试验数据进行降维分析,得到的主成分方差贡献率如图 3 所示。

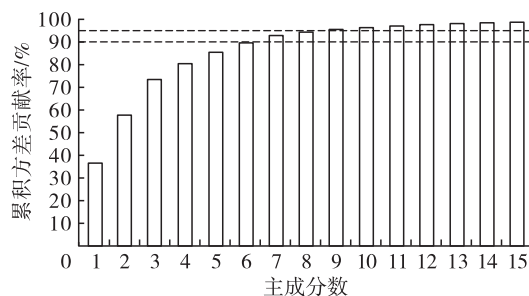


图 3 主成分累积方差贡献率

从图 3 可以看出,可视化传感器阵列需要 7 个主成分才能保证累积方差贡献率达到 90% 以上,需要 9 个主成分才能保证累积方差贡献率达到 95% 以上,这表明制得的嗅觉可视化传感器阵列具有较高的离散性。考虑到输入过多的主成分向量会导致建立的模型过于复杂,本研究选用 1 ~ 10 个主成分向量用于模型的建立,将试验数据按照 4:1 的比例随机划分为训练集与测试集,导入 3 种分类模型中进行模型参数优化,以测试集的准确率作为模型优化结果的判断标准。对优化后的最优模型,采用五折交叉验证的方法来验证模型的稳定性。

2.2.2 KNN 分类模型

建立 KNN 分类模型时,选择欧式距离来度量点与点之间的距离,以 1 ~ 10 个主成分向量作为输入,测试集的准确率作为输出优化 K 值的选取。KNN 分类模型测试集准确率如图 4 所示。

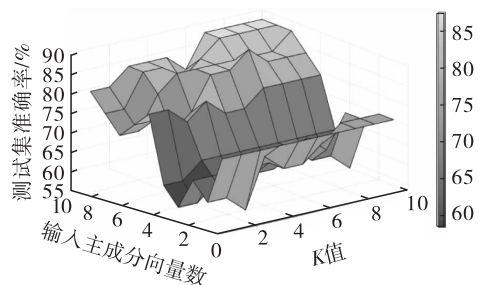


图 4 KNN 分类模型测试集准确率

从图4可以看出,随着输入的主成分向量数和 K 值的增加,测试集的准确率不断上升,直到输入主成分向量数大于7, K 值大于7后,测试集的准确率趋于稳定,不再增加。因此,当输入主成分向量数为7, K 值为7时,建立的KNN分类模型性能最优,测试集的分类准确率为87.5%,五折交叉验证准确率为86.7%。

2.2.3 ELM分类模型

建立ELM分类模型时,设置隐层神经元数为10~50。选择1~10个主成分向量作为输入,以测试集准确率作为输出来优选隐层神经元。ELM分类模型测试集准确率如图5所示。

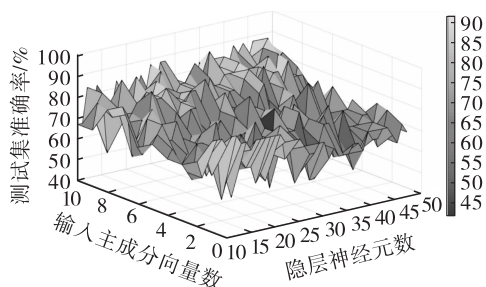


图5 ELM分类模型测试集准确率

从图5可以看出,建立的ELM分类模型没有明显的规律性,测试集的准确率波动较大,原因可能是建立的ELM分类模型训练程度不够,没有很好地提取出数据集的特征。当输入主成分向量数为8时,隐层神经元数为34时,建立的ELM分类模型最优,测试集分类准确率为91.7%,五折交叉验证准确率为82.5%。

2.2.4 SVM分类模型

通过选择不同的核函数,SVM可以实现多种类型数据的分类识别。本研究选用应用最广的RBF核函数来建立SVM分类识别模型,以1~10个主成分向量作为输入,选用网格搜索方法来对关键模型参数 c 、 g 进行寻优。不同输入主成分向量数的最佳 c 、 g 参数下SVM模型的测试集准确率见表1。

表1 不同输入主成分向量数的最佳 c 、 g 参数下SVM模型的测试集准确率

输入主成分向量数	参数优化		测试集准确率/%
	c	g	
1	1.741 1	1.741 1	79.2
2	3.482 2	1.000 0	70.8
3	1.000 0	1.000 0	75.0
4	2.639 0	2.000 0	75.0
5	2.297 4	3.031 4	79.2
6	2.639 0	3.482 2	83.3
7	1.741 1	4.594 8	95.8
8	1.741 1	4.594 8	91.7
9	1.741 1	4.594 8	95.8
10	2.639 0	3.031 4	95.8

由表1可以看出,随着输入主成分向量数的增加,SVM分类模型的准确率总体上不断增加。当输入主成分向量数为7时,测试集准确率达到最高;随着主成分向量数的进一步增加,SVM分类模型的准确率仅有小幅变化,并最终趋于稳定。因此,当输入主成分向量数为7、 $c=1.741 1$ 、 $g=4.594 8$ 时,建立的SVM分类模型性能最优,测试集准确率为95.8%,五折交叉验证准确率为89.6%。这表明建立的SVM分类模型可以很好地实现不同种类食用植物油的区分。

2.3 模型对比分析

3种分类模型的最优分类结果及关键模型参数见表2。

表2 3种分类模型的最优分类结果及关键模型参数

分类模型	输入主成分向量数	关键模型参数	测试集准确率/%	五折交叉验证准确率/%
KNN	7	K 值为7	87.5	86.7
ELM	8	隐层神经元数为34	91.7	82.5
SVM	7	$c=1.741 1$ 、 $g=4.594 8$	95.8	89.6

由表2可以看出,相比于KNN分类模型和ELM分类模型,SVM分类模型具有更高的准确率。分析原因如下:主成分分析的结果显示食用植物油的区分需要较多的主成分向量,在KNN分类模型中,仅通过相邻点之间的多数类别来判断预测点的类别,可能忽略了关键的特征信息,导致分类结果不理想。ELM分类模型属于神经网络模型的一种,神经网络算法需要较多的训练数据,而本研究中用于建立分类模型的数据集为120组数据,数据偏少,模型的训练程度不够,导致建立的ELM分类模型分类效果不理想。而在建立的SVM分类模型中,通过RBF核函数将分类数据集映射到高维空间中来寻找出超平面实现分类,能够保留更多的数据特征信息,分类效果较好。总体而言,建立的SVM分类模型性能最优,其测试集分类准确率为95.8%,五折交叉验证准确率为89.6%。这表明嗅觉可视化技术可以实现食用植物油类别的准确识别。

3 结论

本研究采用嗅觉可视化技术进行4种食用植物油种类的辨识,选定100℃作为试验加热温度进行了油样的分类识别试验。主成分分析的结果表明,制得的可视化传感器阵列具有较高的分散性。选择1~10个主成分向量数作为输入,建立了KNN、ELM、SVM 3种分类模型,结果表明,建立的SVM分类模型性能最优,当输入的主成分向量数为7、

$c = 1.741 1$ 、 $g = 4.594 8$ 时, SVM 分类模型的测试集分类准确率为 95.8%, 五折交叉验证准确率为 89.6%。因此, 制得的可视化传感器阵列可以实现 4 种食用植物油的分类识别, 嗅觉可视化技术用于食用植物油检测是可行的。

参考文献:

- [1] 柯城. 茶油主要营养物质组成和保健功能[J]. 现代食品, 2019(13):105-108.
- [2] 余顺波, 陈长艳, 张品, 等. 11 种食用植物油的脂肪酸组成及主要营养成分含量[J]. 贵州农业科学, 2022, 50(7):113-120.
- [3] JIA X, DENG Q C, YANG Y N, et al. Unraveling of the aroma-active compounds in virgin camellia oil (*Camellia oleifera* Abel) using gas chromatography - mass spectrometry - olfactometry, aroma recombination, and omission studies[J]. J Agric Food Chem, 2021, 69(32): 9043-9055.
- [4] 管方方, 何卓琼, 方燕, 等. GC-MS 分析食用油中甘油三酯的研究进展[J]. 中国油脂, 2014, 39(5):76-80.
- [5] 接昭玮, 李绅, 汪睿璇, 等. 红外光谱快速识别食用植物油种类的研究[J]. 中国油脂, 2023, 48(7):56-61.
- [6] ADERVAL S L, ARNALDO P S, JOAN F, et al. Classification of edible oils and modeling of their physico-chemical properties by chemometric methods using mid-IR spectroscopy [J]. Spectrochim Acta B, 2013, 100: 109-114.
- [7] XUE W H, GUANG L W, XIN W, et al. Convolutional neural network based approach for classification of edible oils using low-field nuclear magnetic resonance[J/OL]. J Food Compos Anal, 2020, 92:103566 [2022-09-13]. <https://doi.org/10.1016/j.jfca.2020.103566>.
- [8] 郭永跃, 马君义, 吕孝飞, 等. 基于电子鼻技术鉴别陇南初榨橄榄油[J]. 中国粮油学报, 2021, 36(8):94-100.
- [9] 张井, 张维一, 徐静, 等. 电子鼻技术在芝麻油品牌识别

及掺假鉴别中的应用[J]. 食品与发酵工业, 2017, 43(6):239-243, 249.

- [10] RAKOW N A, SUSLICK K S. A colorimetric sensor array for odour visualization [J]. Nature, 2000, 406(6797): 710-713.
- [11] KENNETH S S, RAKOW N A, AVIJIT S. Colorimetric sensor arrays for molecular recognition [J]. Tetrahedron, 2004, 60(49):11133-11138.
- [12] YANG M, ZHAI X D, HUANG X W, et al. Rapid discrimination of beer based on quantitative aroma determination using colorimetric sensor array [J/OL]. Food Chem, 2021, 363(30): 130297 [2022-09-13]. <https://doi.org/10.1016/j.foodchem.2021.130297>.
- [13] TONG L, HUI J, CHEN Q S. Qualitative identification of rice actual storage period using olfactory visualization technique combined with chemometrics analysis [J/OL]. Microchem J, 2020, 159:105339 [2022-09-13]. <https://doi.org/10.1016/j.microc.2020.105339>.
- [14] 陈琳, 叶阳, 董春旺, 等. 基于嗅觉可视化技术的工夫红茶发酵程度判定方法[J]. 茶叶科学, 2017, 37(3): 258-265.
- [15] 李闯, 朱瑶迪, 李苗云, 等. 基于天然色素嗅觉可视化技术的牛肉新鲜度的快速检测[J]. 河南农业大学学报, 2019, 53(1):115-119, 127.
- [16] 郑莉晓, 马无锡. 基于可视化嗅觉检测技术的鱼类新鲜度无损检测研究[J]. 浙江工贸职业技术学院学报, 2020, 20(4):72-74, 79.
- [17] 管彬彬, 赵杰文, 林颢. 嗅觉可视化技术鉴别不同原料和不同批次的食醋[J]. 农机化研究, 2013, 35(11): 202-205.
- [18] LIN H, LIN J J, SONG B T, et al. Simulation and non-invasive testing of vinegar storage time by olfaction visualization system and volatile organic compounds analysis [J/OL]. Foods, 2021, 10(3): 532 [2022-09-13]. <http://doi.org/10.3390/foods10030532>.

(上接第 99 页)

- [6] KOVACIC P. Mechanism of drug and toxic actions of gossypol: focus on reactive oxygen species and electron transfer[J]. Curr Med Chem, 2003, 10(24): 2711-2718.
- [7] GIZEJEWSKI Z, SZAFRANSKA B, STEPLEWSKI Z, et al. Cottonseed feeding delivers sufficient quantities of gossypol as a male deer contraceptive [J]. Eur J Wildlife Res, 2008, 54(3): 469-477.
- [8] LORDELO M M, CALHOUN M C, DALE N M, et al.

Relative toxicity of gossypol enantiomers in laying and broiler breeder hens[J]. Poult Sci, 2007, 86(3): 582-590.

- [9] 邵会, 韩文杰, 袁新英, 等. 影响低温棉籽蛋白棉酚含量因素的研究[J]. 中国油脂, 2022, 47(9): 55-58.
- [10] 蒋园园, 吴正奇, 叶文祥, 等. 棉籽蛋白脱酚方法研究[J]. 湖北工业大学学报, 2019, 34(2): 118-120.
- [11] 白玉萍, 王继文, 王文毅, 等. 棉籽膨化加工工艺生产实践[J]. 中国油脂, 2006, 31(11): 83-84.