

多波长激光诱导荧光麦卢卡蜂蜜掺杂分类识别

陈思颖¹, 贾亦文¹, 蒋玉蓉^{1*}, 陈和¹, 杨文慧²,
罗宇鹏¹, 李中石¹, 张寅超¹, 郭磐¹

1. 北京理工大学光电学院, 北京 100081

2. 中国人民解放军军事科学院军事医学研究院, 北京 100071

摘要 麦卢卡蜂蜜产自新西兰, 具有很强的抗菌及抗氧化作用, 其售价较高, 近年来掺假事件时有发生, 利用激光诱导荧光技术对掺杂糖浆的麦卢卡蜂蜜进行分类识别研究。选用 266, 355, 405 和 450 nm 四种常用激光作为激发源, 选择三种品牌的新西兰进口麦卢卡蜂蜜(编号 A, B, C)中掺杂烘焙糖浆作为实验样品, 掺杂比例为 0%~90%, 间隔 10%; 每个激发波长下每种样本溶液重复测试 60 次, 共 7 200 组数据。光谱数据首先进行荧光波段截取、平滑及归一化等预处理; 然后随机选取 80% 的数据做训练集, 20% 的数据做测试集; 对训练集数据使用主成分分析(PCA)结合线性判别分析(LDA)做数据降维; 最后对降维后的数据分别建立 K 最近邻(KNN)和支持向量机(SVM)分类模型, 对测试集数据进行分类识别, 重复进行 50 次随机分组及分类识别后对得到的分类识别率求平均值及标准差。实验分析结果表明, 激发光波长对最终识别结果影响较大, 266 nm 激发的荧光光谱分类识别正确率最高, 三种麦卢卡蜂蜜掺杂溶液的分类识别率均能达到 98.5% 以上, 最高能达到 100%; 355 和 405 nm 激发的分类识别效果次之, 所有样品的分类识别率均大于 92%; 而 450 nm 激发的荧光光谱分类识别率最低, 不同样品的分类识别率均低于 66%。因此, 后续分类算法的比较仅使用 266, 355 和 405 nm 激发的荧光光谱数据, 分析结果表明, KNN 算法的分类效果要优于 SVM 算法, 不同激发波长下三种蜂蜜掺杂溶液的分类识别率均是 KNN 算法更高, 且对 266 nm 激发的三种蜂蜜掺杂样品, KNN 算法的分类识别率比 SVM 算法要高 1% 以上。实验结果表明, 使用激光诱导荧光技术对掺假麦卢卡蜂蜜进行分类识别是可行的, 对于掺杂糖浆的麦卢卡蜂蜜, 在使用的所有组合中, 266 nm 激发结合 PCA-LDA 降维和 KNN 分类算法的分类识别率最高, 分类效果最好, 可用于掺假麦卢卡蜂蜜的快速准确鉴别。

关键词 激光诱导荧光; 多波长; 麦卢卡蜂蜜; 掺假检测; 分类识别

中图分类号: O657.3 **文献标识码:** A **DOI:** 10.3964/j.issn.1000-0593(2022)09-2807-06

引言

麦卢卡(Manuka)是新西兰的一种天然红茶树, 蜜蜂采集这种茶树花酿制而成的蜜即为麦卢卡蜂蜜, 麦卢卡蜂蜜是一种医疗级蜂蜜, 具有很强的抗菌作用及较高的药用价值, 近年来在我国的进口蜂蜜中占比越来越大^[1-2]。与普通蜂蜜相比, 麦卢卡蜂蜜售价更高, 且不同等级的麦卢卡蜂蜜间价格差异较大。在 2014 年 1 月国家质检总局发布的进口食品化妆品不合格名单中, 来自法国和新西兰的 7 个批次共计 1 614 千克的麦卢卡蜂蜜因被检出掺假而遭到了销毁或退

货, 因此, 需要对掺假麦卢卡蜂蜜进行快速准确地分类识别。

激光诱导荧光技术检测速度快、灵敏度高; 由于蜂蜜本身含有维生素、氨基酸、酚类等可以产生荧光的物质, 近年来国内外学者越来越多地开始利用激光诱导荧光技术对蜂蜜的产地、植物来源以及真伪进行鉴别。葛雪峰等^[3]利用荧光光谱法得到了蜂蜜浓度和荧光峰值强度间的线性关系, 其决定系数为 0.97。Karoui 等^[4]利用正面荧光光谱法对瑞士产的 7 种蜂蜜进行了植物来源分类, 其校准样品和验证样品的正确分类率分别为 100% 和 90%, 证明了基于荧光光谱对蜂蜜植物来源进行分类的可行性。Chen 等^[5]将 KS-GMM 算法

收稿日期: 2021-08-09, 修订日期: 2021-11-16

基金项目: 国家自然科学基金项目(61505009)资助

作者简介: 陈思颖, 1975 年生, 北京理工大学光电学院教授 e-mail: csy@bit.edu.cn

* 通讯作者 e-mail: yrkitty@bit.edu.cn

与激光诱导荧光技术相结合,对不同种类的纯蜂蜜进行识别,识别准确率最大为 96.52%。周杰文等^[6]基于三维荧光光谱技术对掺杂大米糖浆的蜂蜜进行无损鉴别研究,最终 LDA 模型识别率为 94.44%,BP-ANN 模型识别率为 100%。与普通蜂蜜相比,麦卢卡蜂蜜中含有 leptosperin、lepterdine 等独特的荧光标志物,其中 leptosperin 是鉴定麦卢卡蜂蜜的权威化学标志物;新西兰麦卢卡蜂蜜协会规定麦卢卡蜂蜜中的 leptosperin 含量需 $>100 \text{ mg} \cdot \text{kg}^{-1}$ ^[1,7-8]。这些独特的荧光标志物的存在使得麦卢卡蜂蜜的荧光光谱与普通蜂蜜或糖浆的荧光光谱有一定差异,这为利用激光诱导荧光技术对掺假麦卢卡蜂蜜进行分类识别提供了理论基础,但目前这方面的国内外研究还很少。

本工作基于激光诱导荧光技术使用多个激发波长对三种品牌的新西兰进口麦卢卡蜂蜜进行了掺杂研究,激发光波长选用 266, 355, 405 和 450 nm, 掺杂物选用烘焙糖浆,以 0%~90%(间隔 10%)的比例与麦卢卡蜂蜜进行混合,对不同激发波长下三种蜂蜜掺杂溶液的荧光光谱分别进行处理分析。预处理后使用主成分分析(PCA)结合线性判别分析(LDA)对数据降维,降维后的数据分别使用 K 最近邻算法(KNN)和支持向量机算法(SVM)进行训练得到分类模型,最终对测试集数据进行分类识别。不同激发波长下各样品的分类识别率相差较大,其中 266 nm 激发的荧光光谱分类识别正确率最高,均在 98.5%以上,最高能达 100%,且标准差均小于 0.01。450 nm 激发的荧光光谱分类识别正确率最低,均低于 66%,且标准差均高于 0.02。与激发波长相比,不同分类算法对分类识别的影响稍小,但比较而言,KNN 算法的分类识别效果要优于 SVM 算法。

1 实验部分

1.1 样本

选择三种品牌的新西兰进口麦卢卡蜂蜜(编号 A, B, C)和国产烘焙糖浆(主要成分为麦芽糖和果葡糖浆)作为样品。由于蜂蜜和糖浆粘度过大,实验前先将其分别配置成 5 g/100 mL 的水溶液,然后以 10%为间隔,在 0%~90%的掺杂范围内将不同比例的糖浆溶液掺进三种蜂蜜溶液,即每种蜂蜜有十种样品溶液,共三十种样品溶液。使用移液枪将其分别移至干燥洁净的 4 mL 比色皿中准备实验。

1.2 仪器及参数

实验系统如图 1 所示,激光器发射 266/355/405/450 nm 激光至盛有待测样品的比色皿上,激发出的荧光信号首先经过透镜收集,高通滤光片滤除弹性散射光,然后经透镜会聚至光纤($NA=0.22$,芯径 600 μm),最后由光谱仪分光探测后传输至计算机进行显示存储。其中 266 nm 激光器(Beamtech Dawa-200)脉冲能量 20 mJ,频率 20 Hz,355 nm 激光器(Beamtech Dawa-200)脉冲能量 40 mJ,频率 20 Hz,405 nm 激光器(创联科技)和 450 nm 激光器(新产业 LSR450NL-1W)功率均为 500 mW,滤光片为 Semrock 公司的 272/355/405/458 nm 长通滤光片,需随激发光波长的改变进行更换,光谱仪(BWTEK-Exemplar Plus)光谱探测范围

190~1100 nm,分辨率设置为 1 nm。每个激发波长下每种样本溶液重复测试 60 次,共 7200 组数据。

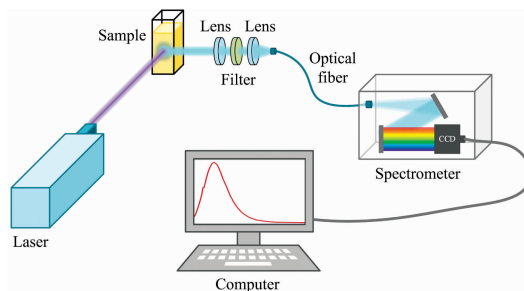


图 1 激光诱导荧光实验系统示意图

Fig. 1 Schematic diagram of laser-induced fluorescence experiment system

2 结果与讨论

2.1 数据处理流程及算法介绍

对于荧光光谱,通常需要先做荧光波段截取、平滑和归一化等预处理;然后对光谱数据进行降维,提取主要特征信息,降维不仅可以缩减数据的运算量,还能够去除部分噪声干扰信息;最后通过分类算法训练降维后的数据得到分类模型,进行荧光光谱的分类识别。

采用 PCA-LDA 算法对掺假麦卢卡蜂蜜的荧光光谱做降维处理,其中 PCA 是一种常用的降维方法,通过正交变换得到少数几个主成分,使其尽可能多地保留原始数据的信息,而 LDA 在选择投影直线时使同类样本的投影点尽可能接近,异类样本的投影点尽可能远离,即在降维的同时还能进行分类,但 LDA 在对“小样本”降维时可能会出现过拟合现象^[9],而如果在使用 LDA 前先使用 PCA 进行特征提取,就可以在一定程度上缓解过拟合问题,因此选用 PCA 结合 LDA 的方法进行光谱数据降维,先使用 PCA 将数据降至低维,然后使用 LDA 继续降至二维。基于 PCA-LDA 的降维结果,使用 KNN 和 SVM 两种分类算法对不同激发波长下的掺假蜂蜜进行分类。简单来说,KNN 算法采用距离度量,计算待测样本与每一类训练样本集的距离,距离越小,则属于该类别的可能性越大。而 SVM 算法的目的是寻找一个最优平面,在尽可能对样本进行准确分类的同时最大化该平面与最近样本点的距离^[10]。

图 2 所示为数据处理流程。光谱数据首先进行预处理,对不同激发波长下每种样品溶液的 60 组数据,随机选择 80%的数据做训练集,其余 20%的数据做测试集,训练集数据先经过 PCA-LDA 降维,然后对降维后的数据分别训练得到 KNN 和 SVM 分类模型,最终使用测试集数据对得到的分

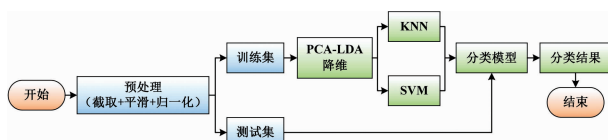


图 2 数据处理流程

Fig. 2 Data processing flowchart

类模型进行测试。随机分组及分类识别重复进行 50 次后对分类识别率取平均值，最终分析比较不同激发波长下不同样品的平均分类识别率及对应的标准差。

2.2 不同激发波长比较

图 3 为四个激发波长下三种蜂蜜掺杂溶液光谱数据预处理后的平均归一化光谱图，其中(a), (b)和(c)为 266 nm 激发的荧光光谱，截取 300~600 nm 荧光波段，(d), (e)和(f)为 355 nm 激发的荧光光谱，截取 360~650 nm 荧光波段，(g), (h)和(i)为 405 nm 激发的荧光光谱，截取 415~750 nm 荧光波段，(j), (k)和(l)为 450 nm 激发的荧光光谱，截取 458~750 nm 荧光波段，且(a), (d), (g)和(j)为编号 A 麦卢卡蜂蜜掺杂溶液，(b), (e), (h)和(k)为编号 B 麦卢卡蜂蜜掺杂溶液，(c), (f), (i)和(l)为编号 C 麦卢卡蜂蜜掺杂

溶液。从图中可以看出，当激发波长为 266 nm 时，不同掺杂比例的荧光光谱有明显区别，谱峰位置随掺杂比例的升高出现蓝移现象，且 315 和 520 nm 处的小谱峰逐渐凸显，光谱的这些差异有助于后续进行分类识别。当激发波长为 355 和 405 nm 时，不同掺杂比例下的荧光光谱谱型接近，但从局部放大图可以看出，谱峰位置有一定差异，其随糖浆掺杂比例的升高出现轻微蓝移现象。当激发波长为 450 nm 时，如图 3 (j), (k)和(l)所示，不同掺杂比例下的荧光光谱谱型和谱峰位置均差别不大，对后续的分类识别造成一定困难。

光谱数据预处理后，随机选择 80% 的数据作训练集，20% 的数据作测试集，使用 PCA-LDA 算法对训练集数据做降维，得到的二维散点图如图 4 所示，其中(a), (b)和(c)为 266 nm 激发，(d), (e)和(f)为 355 nm 激发，(g), (h)和(i)

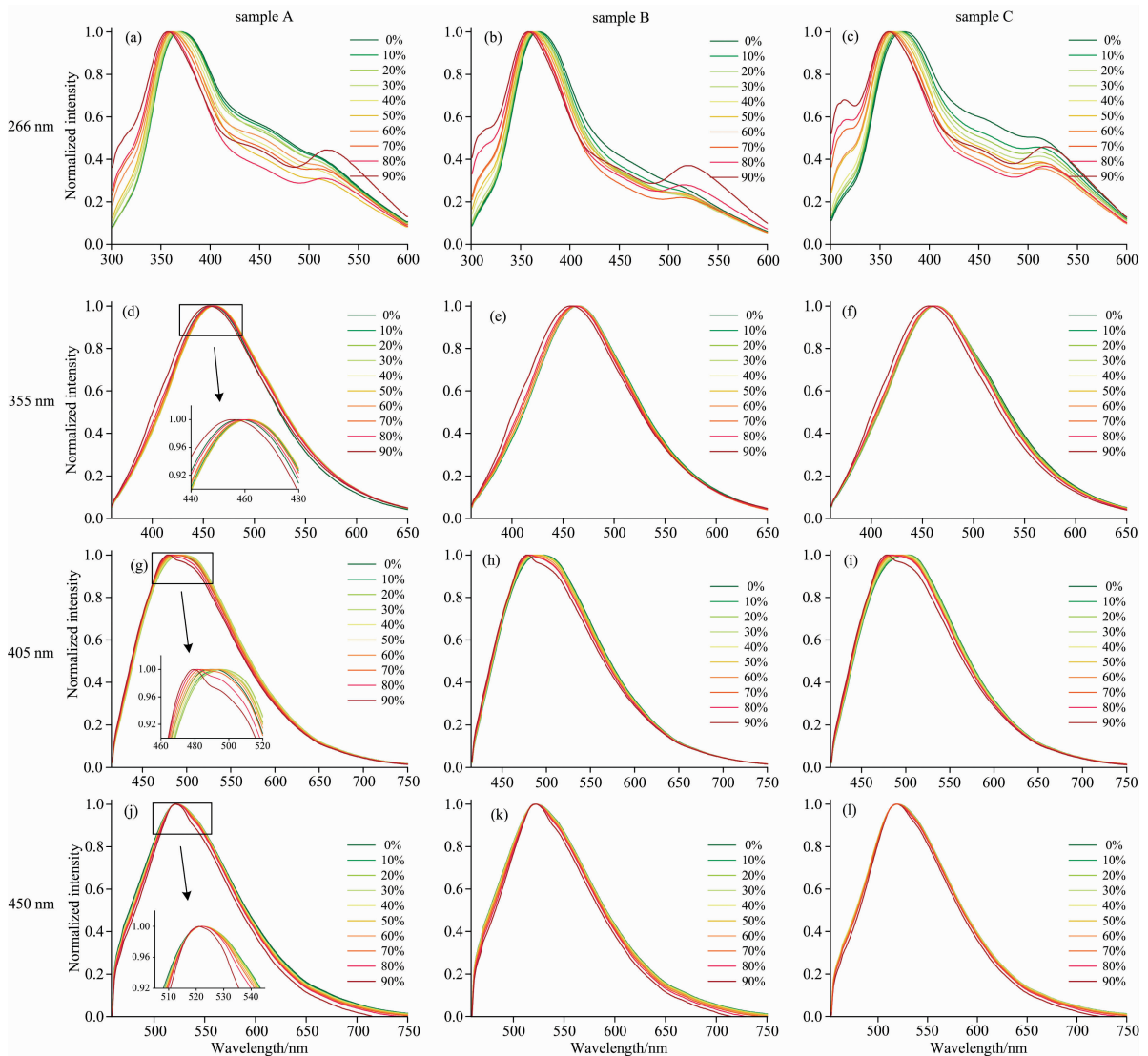


图 3 预处理后荧光光谱图

(a), (b), (c): 266 nm 激发; (d), (e), (f): 355 nm 激发; (g), (h), (i): 405 nm 激发; (j), (k), (l): 450 nm 激发;
(a), (d), (g), (j): 蜂蜜 A; (b), (e), (h), (k): 蜂蜜 B; (c), (f), (i), (l): 蜂蜜 C(下同)

Fig. 3 Fluorescence spectra after pretreatment

(a), (b), (c): 266 nm excitation; (d), (e), (f): 355 nm excitation; (g), (h), (i): 405 nm excitation; (j), (k), (l): 450 nm excitation;
(a), (d), (g), (j): Sample A; (b), (e), (h), (k): Sample B; (c), (f), (i), (l): Sample C (the same below)

为 405 nm 激发, (j), (k) 和 (l) 为 450 nm 激发, 且 (a), (d), (g) 和 (j) 为编号 A 麦卢卡蜂蜜掺杂溶液, (b), (e), (h) 和 (k) 为编号 B 麦卢卡蜂蜜掺杂溶液, (c), (f), (i) 和 (l) 为编号 C 麦卢卡蜂蜜掺杂溶液。可以看出, 随着激发光波长的增加, 同类散点逐渐分散, 不同类散点出现重叠, 降维效果变差, 与归一化光谱图的结果吻合。其中 266 nm 激发得到的降维图效果最好, 同掺杂比例的散点聚合得较好, 不同掺杂

比例的散点间很少出现重叠, 有利于后续分类识别。与 266 nm 激发相比, 355 和 405 nm 激发时相同掺杂比例的散点更加分散, 当掺杂比例大于等于 60% 时, 不同类的散点间距离较大, 而掺杂比例小于 60% 的不同类散点会有一定程度的重叠, 说明高掺杂比例更容易进行区分。显然, 450 nm 激发得到的降维图效果最差, 同掺杂比例的散点分散, 不同掺杂比例的散点重叠, 大大增加了后续分类识别的难度。

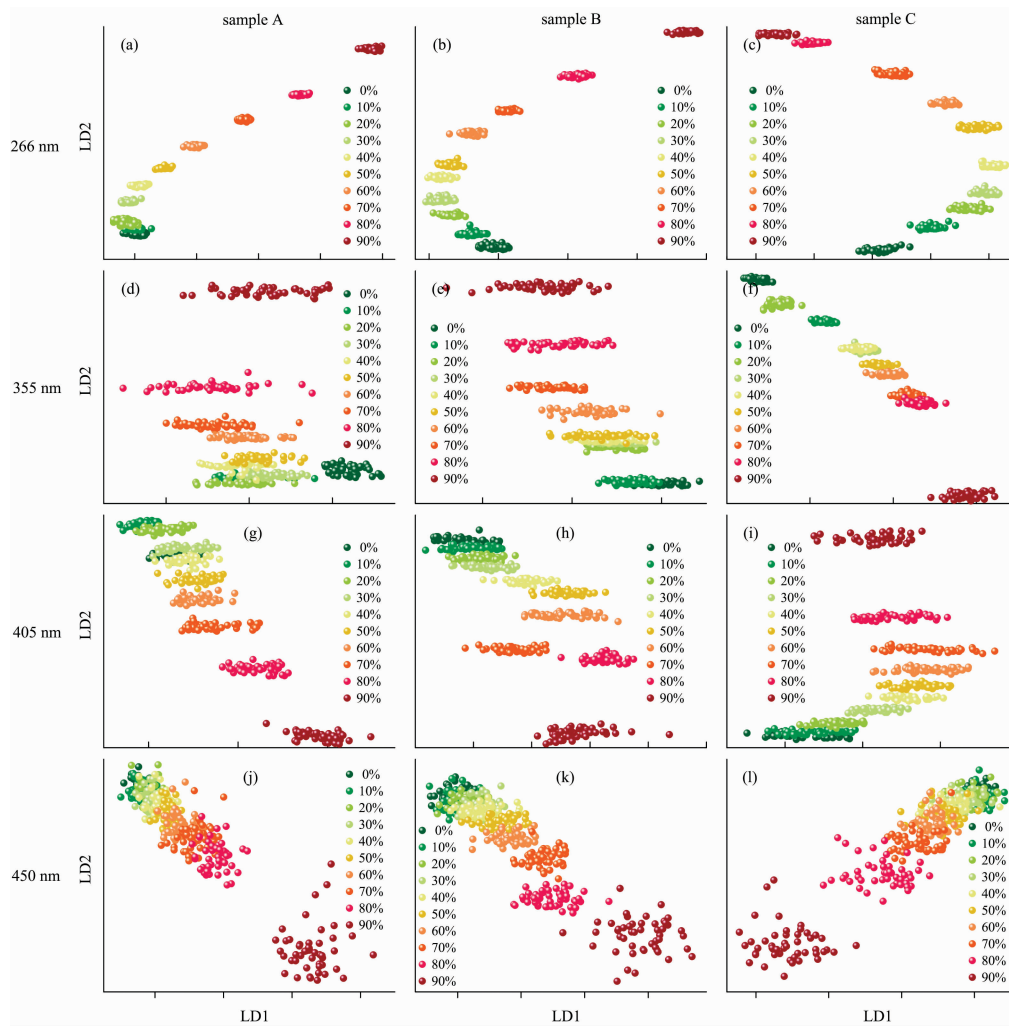


图 4 训练集数据降维图

Fig. 4 Dimensionality reduction diagram of training set data

对降维后的训练集数据建立 KNN 模型, 然后对测试集数据进行分类识别, 重复进行 50 次随机分组及分类识别后对得到的分类识别率求平均值及标准差。表 1 为四个激发波长下三种蜂蜜掺杂溶液的平均分类识别率及对应标准差, 可以看出, 三种蜂蜜掺杂溶液都是 266 nm 激发时得到的分类识别率最高, 均在 98.5% 以上, 且标准差最小, 三个标准差均值为 0.003 2, 其中 B 掺杂溶液的分类识别率最高, 达到 100%, 标准差为 0, 说明 50 次计算得到的分类识别率每次都为 100%, 结果稳定。355 和 405 nm 激发得到的分类识别率接近, 三种掺杂溶液的分类识别率均在 92% 以上, 最高达

98.88%, 不同样品的标准差均值分别为 0.019 8 和 0.016 4, 且比较而言, 对于 B 掺杂溶液, 405 nm 激发的分类效果更好, 对于 C 掺杂溶液, 355 nm 激发得到的分类识别率更高。450 nm 激发得到的分类识别率最低, 均低于 66%, 且标准差最大, 标准差均值为 0.040 3。整体来看, 不同激发波长下三种麦卢卡蜂蜜掺杂溶液的平均分类识别率结果与图 2 归一化光谱图和图 3 降维图的情况相符, 即 266 nm 激发分类效果最好, 355 和 405 nm 激发次之, 450 nm 激发分类效果最差。

表1 不同激发波长下三种蜂蜜的分类识别率及标准差

Table 1 Recognition rates and standard deviations of three kinds of honey under different excitation wavelengths

| 激发波长/nm | 分类算法 | A | | B | | C | |
|---------|------|-------|---------|-------|---------|-------|---------|
| | | 识别率/% | 标准差 | 识别率/% | 标准差 | 识别率/% | 标准差 |
| 266 | KNN | 98.75 | 0.008 1 | 100 | 0 | 99.97 | 0.001 6 |
| 355 | KNN | 92.77 | 0.027 7 | 93.62 | 0.018 0 | 98.20 | 0.013 8 |
| 405 | KNN | 92.85 | 0.025 1 | 98.88 | 0.009 5 | 96.58 | 0.014 6 |
| 450 | KNN | 57.52 | 0.044 0 | 65.48 | 0.029 2 | 57.42 | 0.047 8 |

2.3 不同分类算法比较

由于450 nm激发分类效果较差,后续分类算法的比较仅使用266, 355和405 nm三种激发波长。其他处理方式不变,仅改变分类算法,对训练集数据建立SVM分类模型,然后对测试集数据进行测试,随机分组及分类识别重复进行50次,对分类识别率求均值。不同样品溶液在两种分类算法下的分类识别率、识别率差值及对应标准差如表2所示。从表中可以看出,相比于SVM算法,KNN算法的分类效果更好,对于三种激发波长和三种麦卢卡蜂蜜的掺杂溶液样品的

任意组合,KNN算法的分类识别率比SVM算法均有提升,且除405 nm激发A掺杂溶液,其他情况下KNN算法50次分类识别率的标准差均小于SVM算法。具体来看,266 nm激发A掺杂溶液两种算法的分类识别率相差1.57%,提升最为明显,405 nm激发A掺杂溶液两种算法的分类效果仅相差0.10%,最为接近。整体来看,266 nm激发下KNN算法的分类识别率比SVM算法的分类识别率均提高1%以上,效果最好,而对于C样品,KNN算法的识别率对三种激发波长都能提升大约1%,效果最为明显。

表2 不同激发波长及分类算法下的分类识别率及标准差

Table 2 Recognition rates and standard deviations for different excitation wavelengths and classification algorithms

| 激发波长/nm | 分类算法 | A | | | B | | | C | | |
|---------|------|-------|------|---------|-------|------|---------|-------|------|---------|
| | | 识别率/% | 差值/% | 标准差 | 识别率/% | 差值/% | 标准差 | 识别率/% | 差值/% | 标准差 |
| 266 | KNN | 98.75 | 1.57 | 0.008 1 | 100 | 1.13 | 0 | 99.97 | 1.09 | 0.001 6 |
| | SVM | 97.18 | | 0.014 4 | 98.87 | | 0.0105 | 98.88 | | 0.009 8 |
| 355 | KNN | 92.77 | 0.67 | 0.027 7 | 93.62 | 0.14 | 0.018 0 | 98.20 | 1.40 | 0.013 8 |
| | SVM | 92.10 | | 0.034 0 | 93.48 | | 0.018 3 | 96.80 | | 0.016 4 |
| 405 | KNN | 92.85 | 0.10 | 0.025 1 | 98.88 | 1.10 | 0.009 5 | 96.58 | 0.95 | 0.014 6 |
| | SVM | 92.75 | | 0.023 3 | 97.78 | | 0.018 6 | 95.63 | | 0.017 7 |

3 结论

基于激光诱导荧光技术对掺杂糖浆的麦卢卡蜂蜜进行了分类识别研究,选用266, 355, 405和450 nm四个激发波长,预处理后使用PCA-LDA算法做数据降维,然后基于降维后的训练集数据建立KNN和SVM分类模型,最后对测试集数据进行分类识别。通过对实验数据处理分析发现,激

发光波长对分类结果影响较大,450 nm激发的分类识别效果最差,而266, 355和405 nm激发下不同样品的平均分类识别率均可达到92%以上,其中266 nm激发配合PCA-LDA降维及KNN分类得到的分类识别率最高,均在98.5%以上,且B品牌麦卢卡蜂蜜掺杂溶液的分类识别率能够稳定达到100%,证明了激光诱导荧光技术用于麦卢卡蜂蜜掺杂分类识别的可行性,可用于掺假麦卢卡蜂蜜的快速准确鉴别。

References

- [1] Hegazi N M, Elghani G, Farag M A. Journal of Food Science and Technology, 2022, 59: 2527.
- [2] KANG Hai-ning, CHEN Bo, LAN Fang, et al(康海宁, 陈波, 蓝芳, 等). Modern Food Technology(现代食品科技), 2019, 35(12): 328.
- [3] GE Xue-feng, WU Yan-wei, ZHAN Zhi-min(葛学峰, 吴彦玮, 赵志敏). Analysis Laboratory(分析实验室), 2017, 36(6): 660.
- [4] Karoui R, Dufour E, Bosset J O, et al. Food Chemistry, 2007, 101(1): 314.
- [5] Chen H, Xu Q, Jia Y, et al. Applied Optics, 2021, 60(21): 6140.
- [6] ZHAO Jie-wen, HAN Xiao-yan, CHEN Quan-sheng, et al(赵杰文, 韩小燕, 陈全胜, 等). Spectroscopy and Spectral Analysis(光谱学与光谱分析), 2013, 33(6): 188.
- [7] Bong J, Prijic G, Braggins T J, et al. Food Chemistry, 2017, 214: 102.
- [8] Bong J, Loomes K M, Lin B, et al. Food Chemistry, 2017, 267(30): 355.
- [9] Szymańska E, Gerretzen J, Engel J, et al. TrAC Trends in Analytical Chemistry, 2015, 69: 34.

[10] ZHOU Zhi-hua(周志华). Machine Learning(机器学习). Beijing: Tsinghua University Press(北京: 清华大学出版社), 2016.

Classification and Recognition of Adulterated Manuka Honey by Multi-Wavelength Laser-Induced Fluorescence

CHEN Si-ying¹, JIA Yi-wen¹, JIANG Yu-rong^{1*}, CHEN He¹, YANG Wen-hui², LUO Yu-peng¹, LI Zhong-shi¹, ZHANG Yin-chao¹, GUO Pan¹

1. School of Optics and Photonics, Beijing Institute of Technology, Beijing 100081, China

2. Academy of Military Medical Sciences, Academy of Military Sciences, Beijing 100071, China

Abstract Manuka honey is produced in New Zealand and has strong antibacterial and antioxidant effects. The price is relatively high, and adulteration incidents have occurred frequently in recent years. This paper uses laser-induced fluorescence (LIF) technology to classify and identify Manuka honey adulterated with syrup. Four commonly used lasers of 266, 355, 405 and 450 nm are selected as excitation sources, and three brands of New Zealand Manuka honey (No. A, B and C) adulterated with baking syrup are used as experimental samples. The adulteration ratio ranged from 0% to 90%, with an interval of 10%. Each sample solution has been tested 60 times under different excitation wavelengths, with a total of 7 200 sets of data. For the spectral data, firstly, pretreat with fluorescence band interception, smoothing and normalization; Then, randomly select 80% of the data as training and 20% as test sets. Principal component analysis (PCA) combined with linear discriminant analysis (LDA) is used to reduce the dimension of the training set data; Finally, K-nearest neighbor (KNN) and support vector machine (SVM) classification models are established for the dimensionality reduction data respectively, and the test set data are classified and identified by the models. After 50 times of random grouping and classification calculation, the recognition rate's average value and standard deviation are obtained. The experimental results show that the excitation wavelength greatly influences the final recognition results. The recognition rate of 266 nm excitation is the highest. The recognition rates of the three Manuka adulterated solutions are more than 98.5%, and the highest can reach 100%; 355 and 405 nm excitation are the second, and the recognition rates of all samples are greater than 92%; The classification effect of 450 nm excitation is the worst, with the recognition rates less than 66%. Therefore, the comparison of classification algorithms only uses the spectral data excited by 266, 355 and 405 nm. The analysis results show that the classification effect of the KNN algorithm is better than the SVM algorithm. For the three honey adulterated solutions excited by 266 nm, the recognition rates of the KNN algorithm are more than 1% higher than that of the SVM algorithm. According to the experimental results, using LIF to classify and identify adulterated Manuka honey is feasible. For Manuka honey adulterated with syrup, among all combinations used in this paper, 266nm excitation combined with PCA-LDA and KNN algorithms has the highest recognition rate and the best classification effect, which provides an effective method for rapid and accurate identification of adulterated Manuka honey.

Keywords Laser-induced fluorescence; Multi-wavelength; Manuka honey; Adulteration detection; Classification and identification

(Received Aug. 9, 2021; accepted Nov. 16, 2021)

* Corresponding author