

基于多种算法建立牡蛎及石决明的拉曼光谱鉴别模型

李君翔¹, 倪琳^{2*}, 王亚飞¹, 马寒露³, 宋平顺², 杨平荣², 王红球³ (1.兰州大学药学院, 兰州 730000; 2.甘肃省药品检验研究院/国家药监局中药材及饮片质量控制重点实验室/甘肃省中藏药检验检测工程技术实验室, 兰州 730000; 3.北京鉴知技术有限公司, 北京 100081)

摘要: 目的 利用拉曼光谱建立多种算法的定性模型, 用于贝壳类动物药牡蛎和石决明的鉴别分析。方法 采集 32 批牡蛎和 29 批石决明共计 366 个样品的拉曼光谱, 光谱在 Matlab 软件中作 Savitzky-Golay 平滑滤波、Scale Normalization for Image Pyramids 基线校正、归一化预处理, 并对全部及特征波段进行主成分分析降维。使用 K 最近邻法、决策树算法、判别分析法、集成学习法、支持向量机、人工神经网络 6 种分类算法建立鉴别模型, 通过贝叶斯优化对前 5 种模型性能进行提升。结果 预处理后拉曼信号峰明显, 降维后牡蛎和石决明有聚类趋势。建立以 200~310, 670~740, 1 050~1 100 cm⁻¹ 为特征波段的 5 种算法鉴别模型, 相比于全波段模型, 除集成学习法、支持向量机外, 其他算法准确率均有一定的提高。贝叶斯优化后, 有 3 种模型训练集准确率达到 98%, 测试集准确率达到 100%。采用人工神经网络建模分类, 训练集、验证集、测试集的正确率均达到 100%, 能够准确地区分牡蛎和石决明。结论 基于多种算法或 IP 建立了牡蛎及石决明拉曼光谱鉴别模型, 预测效果均较为理想, 其中人工神经网络模型可实现牡蛎和石决明的百分百准确、快速鉴别。

关键词: 拉曼光谱; 分类算法; 鉴别模型; 牡蛎; 石决明

中图分类号: R917 文献标志码: B 文章编号: 1007-7693(2023)04-0477-06

DOI: 10.13748/j.cnki.issn1007-7693.2023.04.007

引用本文: 李君翔, 倪琳, 王亚飞, 等. 基于多种算法建立牡蛎及石决明的拉曼光谱鉴别模型[J]. 中国现代应用药学, 2023, 40(4): 477-482.

Identification Models of Raman Spectrum for Ostreae Concha and Haliotidis Concha Based on Multiple Algorithms

LI Junxiang¹, NI Lin^{2*}, WANG Yafei¹, MA Hanlu³, SONG Pingshun², YANG Pingrong², WANG Hongqiu³ (1.School of Pharmacy, Lanzhou University, Lanzhou 730000, China; 2.Gansu Provincial Institute of Drug Control/Key Laboratory of Quality Control of Chinese Medicinal Materials and Decoction Pieces of NMPA/Gansu Province Chinese and Tibetan Medicine Testing Engineering Technology Laboratory, Lanzhou 730000, China; 3.JINSP Co., Ltd., Beijing 100081, China)

ABSTRACT: OBJECTIVE To establish qualitative models of Raman spectroscopy for identification of shellfish animal drugs Ostreae Concha and Haliotidis Concha based on multiple algorithms. **METHODS** The Raman spectrograms of 366 samples from 32 batches of Ostreae Concha and 29 batches of Haliotidis Concha were collected respectively. The spectrograms were processed by Savitzky-Golay smoothing filter, Scale Normalization for Image Pyramids, and normalized preprocessing in Matlab. Principal component analysis was used to reduce dimension of full band and specific bands. Six classification algorithms including K-nearest neighbor, decision tree, discriminant analysis, ensemble learning, support vector machine and artificial neural network were used to establish the identification models, and the performance of the first five models was improved by Bayesian optimization. **RESULTS** The Raman signal peaks were obvious after preprocessing, and there was obvious clustering trend of Ostreae Concha and Haliotidis Concha after dimension reduction. Five algorithmic identification models with 200–310, 670–740 and 1050–1100 cm⁻¹ as characteristic bands were established. Compared with the full band model, the accuracy of all classifiers were improved except ensemble learning and support vector machines. After Bayesian optimization, the accuracy of three models achieved 98% in training set and 100% in test set. Using artificial neural network modeling for classification, the accuracy of training set, validation set and test set reached 100%, which could distinguish Ostreae Concha and Haliotidis Concha well. **CONCLUSION** Qualitative models of multiple algorithms by Raman spectroscopy for the identification of Ostreae Concha and Haliotidis Concha are established. The predictions of each model are satisfactory. Among them, the artificial neural network model can achieve 100% accurate and fast identification of Ostreae Concha and Haliotidis Concha.

KEYWORDS: Raman spectroscopy; classification algorithm; identification model; Ostreae Concha; Haliotidis Concha

基金项目: 中药材及饮片质量控制重点实验室项目(2021GSMPLA-KL01, 2021GSMPLA-KL02)

作者简介: 李君翔, 男, 硕士生 E-mail: 980830104@qq.com *通信作者: 倪琳, 女, 主任中药师 E-mail: 1724010367@qq.com

动物药作为传统中药的重要组成部分，具有广泛的临床应用。牡蛎和石决明均为贝壳类动物药，主产于中国沿海地区，牡蛎具有重镇安神、平肝潜阳的功效，石决明除平肝潜阳的功效外，还能够清肝明目^[1]。牡蛎和石决明外观大都呈卵圆或椭圆形，表面多不平整，珍珠层不规则碎块，棱柱层少见，断面观呈棱柱状，因此普通性状鉴别容易将两者混淆^[2]。贝壳是软体动物的外骨骼，其主要成分为碳酸钙，通过物理化学方法检测专属性低。由此可见，需要建立一种快速、简便、专属性强的方法用于区分牡蛎和石决明。目前有关贝壳类药材鉴别的研究较少，杨文哲等^[3]采用红外光谱法对牡蛎、石决明、珍珠母等样品进行鉴别，发现可以将部分贝壳类药材进行区分。林丽君^[4]用同步热分析法对贝壳类药材及其伪品进行分析，发现真伪品间的热性质存在差异可以作为区分条件，但热分析受试样、仪器、实验条件等因素限制，具有一定的局限性。

拉曼光谱是基于拉曼效应通过测定频率发生变化的散射光来提供分子振动的相关信息，进而分析物质的结构^[5]。拉曼光谱不仅可以提供分子结构与组成的信息，还可以用于晶相、应力改变、构象变化等方面的研究。拉曼光谱对样品种类、样品量、样品形态要求不高，具有无损检测、高灵敏度、快捷简便等优点，目前在食品、医药、环境、军事等领域有广泛应用^[6-7]。近年来手持式拉曼系统的出现加强了拉曼光谱用于现场快检的优势。因此，本研究基于多种算法结合化学计量学方法尝试建立牡蛎以及石决明 2 种贝壳类药材的拉曼光谱鉴别模型，以期传统中药的快速鉴别提供新方法。

1 仪器与试剂

RS2000 便携式拉曼光谱仪(北京鉴知技术有限公司，激光功率为 500 mw，波长范围 200~3 200 cm^{-1} ，分辨率为 6 cm^{-1})；乙腈校正溶液(分析纯，国药集团化学试剂有限公司)。

样品牡蛎共 32 批，石决明共 29 批，部分为平凉市青松中药饮片有限公司送检，其余来自全国各药材市场。各批药材均经甘肃省药品检验研究院宋平顺主任药师鉴别为正品，信息见表 1。每批样品取 6 个样本用于采集拉曼光谱，其中牡蛎样本为 192 个，石决明样本为 174 个，共 366 个。

表 1 样品信息

Tab. 1 Information of samples

批次	名称	来源	批次	名称	来源	批次	名称	来源
1	牡蛎	送检	22	牡蛎	浙江	43	石决明	送检
2	牡蛎	送检	23	牡蛎	浙江	44	石决明	送检
3	牡蛎	送检	24	牡蛎	广东	45	石决明	送检
4	牡蛎	送检	25	牡蛎	广东	46	石决明	送检
5	牡蛎	送检	26	牡蛎	广东	47	石决明	送检
6	牡蛎	送检	27	牡蛎	广东	48	石决明	安徽
7	牡蛎	送检	28	牡蛎	河北	49	石决明	安徽
8	牡蛎	送检	29	牡蛎	河北	50	石决明	安徽
9	牡蛎	送检	30	牡蛎	山东	51	石决明	安徽
10	牡蛎	送检	31	牡蛎	山东	52	石决明	浙江
11	牡蛎	送检	32	牡蛎	山东	53	石决明	浙江
12	牡蛎	送检	33	石决明	送检	54	石决明	浙江
13	牡蛎	送检	34	石决明	送检	55	石决明	浙江
14	牡蛎	送检	35	石决明	送检	56	石决明	浙江
15	牡蛎	送检	36	石决明	送检	57	石决明	广东
16	牡蛎	安徽	37	石决明	送检	58	石决明	广东
17	牡蛎	安徽	38	石决明	送检	59	石决明	广东
18	牡蛎	安徽	39	石决明	送检	60	石决明	山东
19	牡蛎	浙江	40	石决明	送检	61	石决明	山东
20	牡蛎	浙江	41	石决明	送检			
21	牡蛎	浙江	42	石决明	送检			

注：来源中“送检”表示平凉市青松中药饮片有限公司送检，其他为购买地。

Note: Unit submitted for “inspection” was Pingliang Qingsong TCM Decoction Pieces Co., Ltd., and others were the place of purchase.

本研究所有数据分析均使用 Matlab2018b 软件。

2 方法与结果

2.1 拉曼光谱采集

分别取上述样品适量，在暗箱中用拉曼光谱仪的探头轻抵样品表面，通过调节样品平套找到激光最佳聚焦距离后，顺序采集样品的拉曼光谱。激光功率 80%，积分时间 3 000 ms。各样品分别采集 3 张光谱，取平均光谱用于分析。牡蛎和石决明的原始拉曼光谱图见图 1。牡蛎在波数 280，713，1 085 cm^{-1} 处存在特征峰，其中 1 085 cm^{-1} 为碳酸钙 V1 振动特征峰，V4 特征峰 713 cm^{-1} 及晶格振动特征峰 235 cm^{-1} 为三方晶系方解石的特征峰。石决明在 235，703，1 085 cm^{-1} 处存在特征峰。703 cm^{-1} 为斜方晶系文石 V4 振动峰的特征峰。由此可见，牡蛎和石决明虽然都是以碳酸钙为主要的组成成分，但是它们的晶型并不相同，牡蛎以方解石结构为主，石决明以文石结构为主。石决明的拉曼光谱有较强的荧光，这可能是由于石决明中含有一定的胆素等其他成分造成。

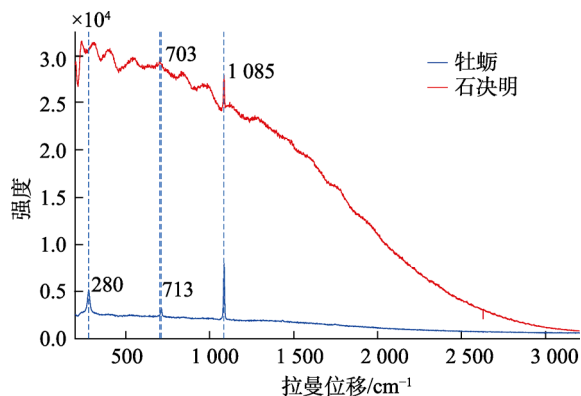


图1 牡蛎和石决明拉曼原始图谱

Fig. 1 Original raman spectra of *Ostrea Concha* and *Haliotis Concha*

2.2 多种算法的拉曼光谱鉴别模型建立

2.2.1 预处理方法 由于拉曼信号受样品表面的平整性以及颜色、仪器噪声、光线等环境因素的影响,因此需要选择合适的图谱预处理方法减少干扰并提高模型性能。从图1可以看到石决明有较强的荧光背景,虽然可以通过荧光对牡蛎和石决明大致进行鉴别分类,但不同批次的样品荧光强度波动较大,荧光也会掩盖一些强度较弱的拉曼峰,因此为提高模型精度和抗干扰性能需要去除荧光干扰。Savitzky-Golay平滑滤波是光谱预处理中被广泛应用的除噪方法,它通过对数据点进行 k 阶多项式拟合,进而得到拟合后的结果。这种方法在滤除噪声的同时可以确保信号的形状和宽度不变^[8]。本研究中的预处理方式采用 Savitzky-Golay 卷积平滑去噪、Scale Normalization for Image Pyramids(SNIP)荧光扣除基线校正等方法组合减少原始图谱中噪声以及荧光的影响。随后进行归一化处理,将样品的拉曼信号强度映射到 $[-1, 1]$ 内,进而提高模型的训练效率^[9]。牡蛎和石决明样品谱图经过前处理后得到的拉曼光谱见图2。经过预处理的谱图能够清晰地看到尖锐的拉曼特征峰,为后续的鉴别分析提供了良好的基础。

2.2.2 波段的选择和光谱数据降维 原始谱图中每个光谱有1704个数据点,数据量庞大,冗余信息较多,可能涉及无关信号影响模型预测效果,因此需结合拉曼图谱对波段进行选择 and 降维后再进行建模分析。牡蛎和石决明的特征峰主要集中在 $200\sim 310$ 、 $670\sim 740$ 、 $1050\sim 1100\text{ cm}^{-1}$ 处,采用主成分分析(principal component analysis, PCA)对这些波段进行降维。降维后前10个主成分的累计

贡献率为93.70%,保留了谱图的大部分信息,见表2。在后续的分类中,采用保留>99%谱图信息的主成分进行,降低了其他无关信号的干扰,节约了计算时间和成本。以特征波段PCA降维后所有样本前3个主成分为坐标建立了PCA得分图,见图3。结果表明,牡蛎和石决明的分布存在明显的聚类趋势。

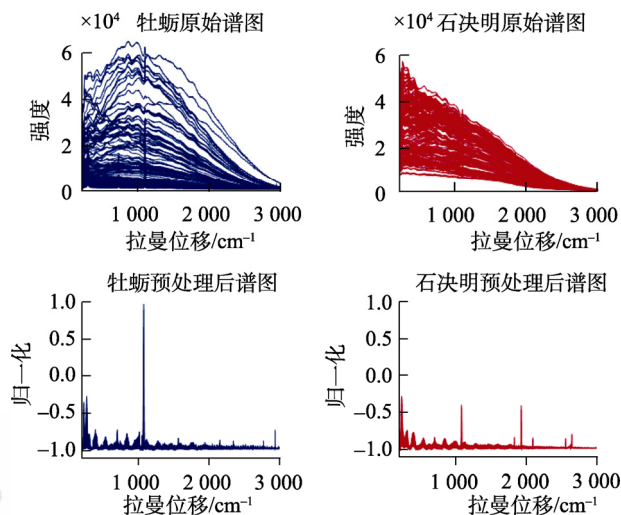


图2 牡蛎和石决明预处理前后图谱

Fig. 2 Atlas of *Ostrea Concha* and *Haliotis Concha* before and after pretreatment

表2 主成分贡献率

Tab. 2 Principal component contribution rate

主成分	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10
贡献率/%	53.18	27.53	4.75	3.22	2.31	1.65	0.42	0.27	0.23	0.14

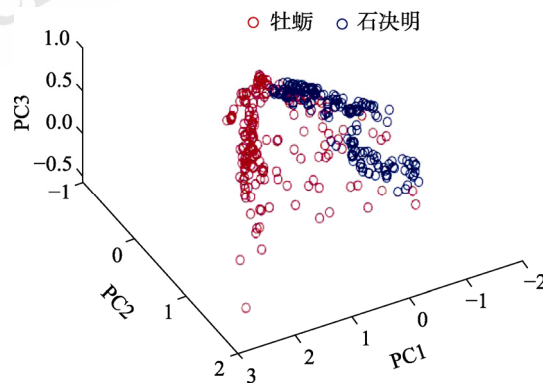


图3 特征波段降维后样本前3个主成分为坐标建立的PCA得分图

Fig. 3 PCA score plot by the first 3 principal components of the sample after feature dimension reduction as the coordinates

K最近邻(K Nearest Neighbor, KNN)法是一种基本分类和回归方法,通过判断与样品最邻近的 K 个样本类别来定义目标样品的类别^[10]。决策树算法采用树形结构,使用层层推理实现最终分类。

决策树易于实现,可解释性强,在多个领域都有广泛的应用^[11]。判别分析法是根据明确分类的样品建立判别函数,进而将新样品进行分类的监督学习算法。集成学习法是通过将多个监督模型集成在一起,以不同模型间相互纠正实现高准确率分类目的的方法^[12]。支持向量机(support vector machines, SVM)是基于统计学习理论的机器学习方法,通过寻找最佳分割超平面将两类数据区分开,同时借助内积的回旋避免了维灾难问题,从而实现在低样本量的情况下获得良好统计规律的目的^[13]。本研究将累计方差贡献率达 99%的主成分采用 KNN 法、决策树、判别分析法、集成学习法、SVM 对所有样品进行建模,所有参数均使用默认参数,五折交叉验证后,比较前处理后未降维的谱图分类的正确率,结果见表 3。从表中可以看出,所用的这些分类算法在使用全波段降维前对牡蛎和石决明分类鉴别的准确率最高可达 97.5%,其中 KNN 分类器的准确率为 80.3%,模型分类鉴别的能力有待提升。在全波段谱图 PCA 降维后,各分类器鉴别准确率均有所提升,其中 KNN 分类器的准确率由 80.3%提升至 97.8%,有较明显的提升。集成学习法的准确率从 95.9%提升至 99.2%,能够很好地区分鉴别牡蛎和石决明这 2 种动物药。可见 PCA 降维不仅能够保留样本信息,同时还能够降低一些干扰,提高模型鉴别的准确率。在选择特征波段降维后,除集成学习法准确率降低, SVM 准确率保持不变外,其他分类器的准确率均得到进一步提高。

表 3 经预处理后不同分类算法准确率

Tab. 3 Accuracy of different classification algorithms after pretreatment %

分类算法	全波段降维前	全波段降维后	特征波段降维后
决策树算法	89.1	94.4	95.4
KNN 法	80.3	97.8	98.6
判别分析法	—	97.4	97.5
集成学习法	95.9	99.2	96.4
SVM	97.5	98.9	98.9

2.2.3 拉曼光谱鉴别多种算法模型 为了进一步提高各分类算法的鉴别准确率并防止模型的过拟合,随机选择 70%(256 个)的拉曼光谱作为训练集,另外 30%(110 个)作为测试集,在前处理的基础上采用多种算法建立牡蛎和石决明的鉴别模型,并使用贝叶斯优化对各算法进行优化,五折交叉验证后,预测测试集分类结果,分类正确率见表 4。

表 4 经贝叶斯优化后不同分类算法准确率

Tab. 4 Accuracy of different classification algorithms after Bayesian optimization %

分类算法	训练集	测试集	优化后训练集	优化后测试集
决策树算法	94.9	95.4	96.9	96.7
KNN 法	94.5	95.5	97.7	96.4
判别分析法	97.3	97.3	98.8	100.0
集成学习法	98.0	99.1	98.4	100.0
SVM	98.4	98.2	98.4	100.0

贝叶斯优化是一种求解表达式位置的函数极值问题的黑盒优化算法。算法根据一组采样点处的函数值通过高斯回归预测出任意点处函数值的概率分布。根据高斯过程回归的结果构造采集函数,用于衡量每一个点值得探索的程度,求解采集函数的极值从而确定下一个采样点。最后返回这组采样点的极值作为函数的极值。SVM 分类器贝叶斯优化的优化过程见图 4, SVM 贝叶斯优化后训练集和测试集准确率分别为 98.4%, 100%, 其中优化后 SVM 分类器测试集较优化前测试集的准确率明显提高。此外,从表 4 可以看到,这 5 种分类算法经贝叶斯优化后训练集和测试集的准确率均>95%。除 SVM 分类器以外,判别分析、集成学习法 2 种分类算法优化后训练集准确率也>98%,优化后测试集准确率达到 100%,可以较好地对牡蛎、石决明进行分类鉴别。

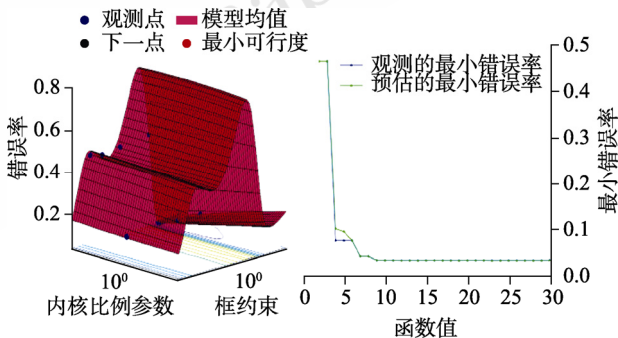


图 4 SVM 分类器贝叶斯优化结果

Fig. 4 Bayesian optimization results of SVM classifier

随着近年来深度学习的高度发展,神经网络已经能够出色地完成聚类、拟合、模式识别(分类)、时序预测等各种任务。神经网络是模仿大脑神经突触的链接和信号传导模式的一种信息处理的数学模型。神经网络的构成主要包括神经元、网络拓扑和算法 3 个主要部分,可以分为输入层、隐藏层和输出层^[14],神经网络结构见图 5。将经过前处理后的拉曼谱图作为输入层,采用 2 层前向式网络,隐藏层为 10 层,训练函数为反向共轭梯度法,激活函数为双曲正切函数。采用留出法

通过程序随机分层选取 70%(256)谱图作为训练集, 15%(55)作为验证集, 15%(55)作为测试集。

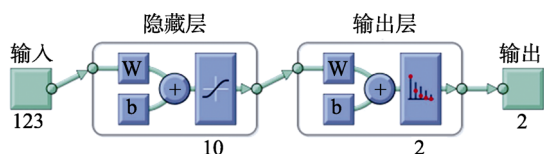


图 5 模式识别神经网络结构

Fig. 5 Pattern recognition of neural network structure

当神经网络迭代次数为 73 次时, 模型分类效果最优, 此时测试集的损失函数最低为 0.000 17。随着迭代次数的增加, 训练集的损失系数降低, 但是验证集的损失系数先降低后波动, 在第 73 次迭代时达到最低, 见图 6。为防止模型的过拟合, 最佳迭代次数为 73 次。此时分类结果显示, 训练集、验证集和测试集的分类准确率均为 100%, 可以准确地区分牡蛎和石决明, 见图 7。

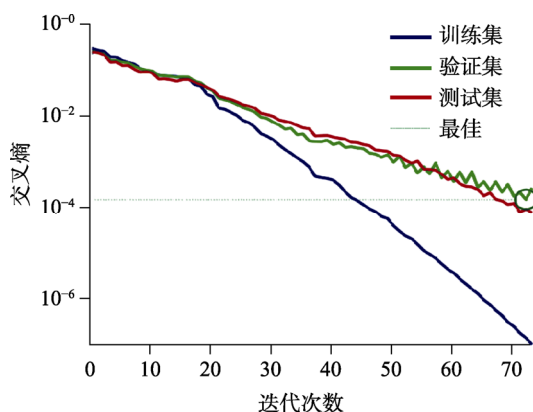


图 6 训练集、验证集及测试集的损失函数

Fig. 6 Loss functions of training set, validation set and test set

训练集混合矩阵			验证集混合矩阵		
输出类	1	139 54.3%	0 0.0%	26 47.3%	0 0.0%
	2	0 0.0%	117 45.7%	0 0.0%	29 52.7%
	目标类	100% 0.0%	100% 0.0%	100% 0.0%	100% 0.0%
测试集混合矩阵			全部混合矩阵		
输出类	1	24 43.6%	0 0.0%	189 51.6%	0 0.0%
	2	0 0.0%	31 56.4%	0 0.0%	117 48.4%
	目标类	100% 0.0%	100% 0.0%	100% 0.0%	100% 0.0%

图 7 牡蛎、石决明分类混淆矩阵

Fig. 7 Confusion matrix of *Ostreae Concha* and *Haliotidis Concha* classification

3 讨论

牡蛎和石决明作为 2 种常见的贝壳类药物, 来源虽然不同, 但外形具有相似性。目前鉴别 2 种药材的方法十分有限, 需要建立有关贝壳类药材完整的定性模型。前人对部分贝壳类药材的近红外光谱进行 PCA 发现能够大致进行区分, 却没有建立程序性的模型并进行验证。Ming 等^[6]曾基于 X 射线衍射解析物相组成, 尝试用拉曼光谱对含碳酸钙的矿石粉末建立聚类分析和偏最小二乘分析鉴别模型, 模型准确率为 95%和 90%, 存在误判状况。且矿石粉末的粉碎需要破碎和磨粉过程, 成本高, 更无法实现快速鉴定。

本研究采用拉曼光谱对牡蛎和石决明药材进行鉴别分析, 发现该方法可以很好地鉴别牡蛎和石决明。来自全国各地的 61 批共计 366 个样品具有一定代表意义。同时大批量样品能够避免建模中因样本量不足导致的整体任务欠拟合等问题, 也能满足部分算法所需的参数量。原始谱图经过去噪、基线拉平、归一化、PCA 降维后, 牡蛎和石决明显示了明显的聚类趋势, 但仍有一部分重叠。使用 KNN 法、决策树算法、判别分析法、集成学习法、SVM 5 种分类器建立定性分析模型, 五折交叉验证后, 使用贝叶斯超函数优化。相较于其他算法, 判别分析法分类效果最好, 训练集准确率为 98.8%, 测试集准确率为 100%。其次为集成学习法和 SVM。KNN 法、决策树算法的准确率相对较低, 可能是由于分类器自身分类能力存在差异。特征波段降维后的主成分结果中少量牡蛎样本作为离群点落在石决明样本中, KNN 法、决策树算法可能对这些样本的判断能力较弱。这些离群点的产生原因可能有以下几个方面: ①牡蛎或石决明个体差异; ②炮制过程中有一定磕碰损伤或检测面不平导致检测到贝壳不同的结构层引起; ③牡蛎和石决明生长时寄生在一起, 未能分开。虽然离群点会对分类结果产生一定影响, 但为了获得更具有实际意义的定性模型, 在建模过程中也应将这些点纳入。总体来看, 这 5 种分类算法训练集和测试集虽然不能达到百分百准确识别, 但是准确率均>95%, 说明预测效果较为理想。由于人工神经网络分类准确度高、分布处理能力强、对噪声神经具有较强的容错能力。为了实现更高的准确率, 采用人工神经网络对预处理后数据进行建模分类。所建立的人工神经网络模型训练集

和验证集和测试集准确率均能达到 100%，可以准确地区分牡蛎、石决明这 2 种易混淆药材。

相较于此前研究对中药粉末进行拉曼或红外光谱分析，本研究在应用拉曼光谱进行检测的过程中，被检样本均无需进行粉碎等预处理，且检测时间非常短，只需要几秒钟，操作简便，对检测环境和检测人员无特殊要求，有利于建立大样本量的数据库。本研究基于拉曼光谱结合多种算法所建立的鉴别模型为贝壳类动物药的补充检验方法提供了新思路，同时也表明拉曼光谱作为一种无损、高灵敏、响应快、指纹图谱性的技术可以结合化学计量学以及机器学习方法应用在更多类型药材的分析中。

REFERENCES

- [1] HAO R, SHAO M S, ZHANG B X. Exploration of the clinical application and dosage of abalone shell[J]. Jilin J Chin Med(吉林中医药), 2021, 41(4): 536-539.
- [2] 中国药典. 一部[S]. 2020: 93, 180.
- [3] YANG W Z, GONG H L, QIN Y H, et al. Identification of marine-derived shell TCM by near infrared spectroscopy[J]. China J Chin Mater Med(中国中药杂志), 2014, 39(17): 3291-3294.
- [4] LIN L J. Application of synchronous thermal analysis in identification and evaluation of pearl powder[J]. Guangzhou Chem Indust(广州化工), 2020, 48(6): 101-103.
- [5] CAO L, ZHU J S, GUAN Y Y, et al. Research progress of Raman spectroscopy in pharmaceutical analysis[J]. J Light Scatt(光散射学报), 2019, 31(2): 101-111.
- [6] MING J, CHEN L, CHEN K L, et al. Identification of 4 mineral medicine powders containing CaCO_3 with XRD and Raman spectroscopy[J]. China J Tradit Chin Med Pharm(中华中医药杂志), 2017, 32(3): 1340-1343.
- [7] LÜ Q H, WANG X H, SHEN A G, et al. Application of Raman spectroscopy techniques in on-site fast detection[J]. J Instrum Anal(分析测试学报), 2019, 38(5): 612-617.
- [8] ZHANG Y Z, XU T F, LIU Z W, et al. Correction method for stripe nonuniformity in infrared images based on Savitzky-Golay weighted fitting algorithm[J]. Chin Opt(中国光学), 2015, 8(1): 51-59.
- [9] BAI H X, YANG F, LI D Y, et al. Multi-component substance classification and recognition based on surface-enhanced Raman spectroscopy[J]. Acta Optica Sinica(光学学报), 2021, 41(20): 164-173.
- [10] GENG L J, LI X Y. Improvements of KNN algorithm for big data classification[J]. Appl Res Comput(计算机应用研究), 2014, 31(5): 1342-1344, 1373.
- [11] MA L H, XU C Y, YANG X M, et al. Application status of decision tree in traditional Chinese medicine[J]. World Chin Med(世界中医药), 2021, 16(17): 2648-2651, 2656.
- [12] ZHANG C X, ZHANG J S. A survey of selective ensemble learning algorithms[J]. Chin J Comput(计算机学报), 2011, 34(8): 1399-1410.
- [13] YAO Y K. Research of key technologies for support vector machine and their applications on human activity recognition [D]. Lanzhou: Lanzhou University, 2015.
- [14] ZHANG C, GUO Y, LI M, et al. Review of development and application of artificial neural network models[J]. Comput Engin Appl(计算机工程与应用), 2021, 57(11): 57-69.

收稿日期: 2022-03-21

(本文责编: 李艳芳)