

基于智能感官与多源信息融合技术的香附炮制程度快速辨识方法研究

吴鑫雨¹, 邱丽媛^{2#}, 王又迪³, 朱灵昊¹, 应佳璐¹, 梁泽华^{1*}

1. 浙江中医药大学药学院, 浙江 杭州 311400

2. 杭州百诚医药科技股份有限公司, 浙江 杭州 310052

3. 浙江药科职业大学, 浙江 宁波 315010

摘要: **目的** 基于多源信息融合技术, 整合传统的中药性状鉴别, 建立香附 *Cyperi Rhizoma* 炮制程度快速辨识方法, 为香附质量评价标准的制定和炮制过程质量控制的应用研究提供新思路、新方法。**方法** 选取6个产地的生香附饮片, 采用醋炙法炮制, 每隔3 min 取样, 得到72批香附炮制过程样品。然后基于色差仪、电子鼻和近红外光谱 (near infrared spectrum, NIRS) 技术获取上述样品的智能感官信息和 NIRS 数据, 利用主成分分析-判别分析 (principal component analysis-discriminant analysis, PCA-DA)、偏最小二乘-判别分析 (partial least squares-discriminant analysis, PLS-DA)、正交偏最小二乘-判别分析 (orthogonal partial least squares-discriminant analysis, OPLS-DA) 方法、Lasso 回归分析、遗传算法 (genetic algorithm, GA) - 反向传播 (back propagation, BP)、神经网络算法 (GA-BP neural network algorithm, GA-BPNN) 等化学计量学方法, 分别基于单一来源信息和多源信息融合建立香附炮制程度辨识方法。**结果** 基于单源的色差仪、电子鼻和 NIRS 技术建立的香附炮制程度辨识模型均无法准确地判别香附 4 类炮制品, 基于电子鼻和色差仪的二类智能感官与多源信息融合技术建立的炮制程度辨识模型能快速、准确地辨识 4 类香附炮制品, 准确度在 0.93 以上, 模型分类预测效果较好。**结论** 基于二类智能感官与多源信息融合技术建立的香附炮制程度辨识模型可以更加准确地识别香附的炮制程度, 进一步提升预测准确度, 为醋香附及其他中药炮制程度快速辨识提供参考。

关键词: 香附; 智能感官信息; 快速辨识; 多源信息融合技术; 性状鉴别; 质量控制; 醋炙; 近红外光谱; 主成分分析-判别分析; 偏最小二乘-判别分析; 正交偏最小二乘-判别分析; Lasso 回归分析; 遗传算法; 反向传播神经网络算法

中图分类号: R283.6 文献标志码: A 文章编号: 0253-2670(2023)21-7007-10

DOI: 10.7501/j.issn.0253-2670.2023.21.009

Research on rapid identification method of processing degree of *Cyperi Rhizoma* based on intelligent sense and multi-source information fusion technology

WU Xin-yu¹, QIU Li-yuan², WANG You-di³, ZHU Ling-hao¹, YING Jia-lu¹, LIANG Ze-hua¹

1. School of Pharmaceutical Sciences, Zhejiang Chinese Medical University, Hangzhou 311400, China

2. Hangzhou Bio-Sincerity Pharma-Tech Co., Ltd., Hangzhou 310052, China

3. Zhejiang Pharmaceutical University, Ningbo 315010, China

Abstract: Objective Based on multi-source information fusion technology, integrating traditional Chinese medicine character identification, a rapid identification method of processing degree of Xiangfu (*Cyperi Rhizoma*) was established, which provided a new idea and a new method for the formulation of quality evaluation criteria and the application research of processing quality control.

Methods A total of 72 batches of fragrant decoction pieces from six regions were processed with vinegar and sampled at a interval of 3 min. Then, the intelligent sensory information and near-infrared spectral data of the above samples were obtained based on the color difference meter, electronic nose and near-infrared spectrum (NIRS), and the principal component analysis-discriminant analysis (PCA-DA), partial least squares-discriminant analysis (PLS-DA), orthogonal partial least squares-discriminant analysis (OPLS-DA), Lasso regression analysis, genetic algorithm (GA)-back propagation (BP) neural network algorithm (GA-BPNN) and other

收稿日期: 2023-04-05

基金项目: 国家重点研发计划——中药饮片质量识别关键技术研究 (2018YFC1707001)

作者简介: 吴鑫雨 (1998—), 硕士研究生, 研究方向为中药炮制工艺及质量标准。Tel: 18989455650 E-mail: 2916478394@qq.com

*通信作者: 梁泽华 (1976—), 副教授, 从事中药炮制工艺及质量标准研究。Tel: 13858042221 E-mail: 1046698407@qq.com

#共同第一作者: 邱丽媛 (1996—), 硕士研究生, 研究方向为中药炮制工艺及质量标准。Tel: 18806823685 E-mail: 2387841189@qq.com

stoichiometric methods were used, then the processing degree identification method was established based on single source information and multi-source information fusion, respectively. **Results** The processing degree identification model based on single source of color difference meter, electronic nose and NIRS could not accurately identify the four types of processed products of *Cyperi Rhizoma*, while the processing degree identification model based on two kinds of intelligent senses and multi-source information fusion technology could quickly and accurately identify the four types of processed products of *Cyperi Rhizoma*, with an accuracy of more than 0.93, and the model classification and prediction effect were good. **Conclusion** The processing degree identification model of *Cyperi Rhizoma* based on two kinds of intelligent senses and multi-source information fusion technology can identify the processing degree of *Cyperi Rhizoma* more accurately, further improve the prediction accuracy, and provide a reference for the rapid identification of the processing degree of *Cyperi Rhizoma* and other traditional Chinese medicine.

Key words: *Cyperi Rhizoma*; intelligent sensory information; rapid identification; multi-source information fusion technology; character identification; quality control; vinegar-burning method; near infrared spectrum; principal component analysis-discriminant analysis; partial least squares-discriminant analysis; orthogonal partial least squares-discriminant analysis; Lasso regression analysis; genetic algorithm; back propagation neural network algorithm

香附 *Cyperi Rhizoma* 为莎草科莎草属植物莎草 *Cyperus rotundus* L. 的干燥根茎, 始载于《名医别录》^[1], 常用于治疗肝气郁滞所引起的胸胁胀闷疼痛、疝气腹痛、月经不调等妇科疾病^[2]。香附一般醋炙后入药, 醋香附以其抗炎镇痛、理气开郁等药理作用在临床治疗中发挥着重要作用^[3]。传统饮片鉴别方法常对“色、气、味”3 个方面进行评价, 采用眼观、鼻嗅、手摸等方法对饮片进行真伪鉴别或质量评价, 这种传统评价方法存在模糊性、经验性、不确定性、不稳定性等缺点, 只有少数经验十分丰富的老药工和长期从事中药材或中药饮片生产加工的人员才能有较好的判断能力, 这对于中药饮片质量评价的现代化发展和实际应用推广产生了一定的阻碍。然而现行的中药饮片质量评价模式也多以化学指标成分的定性和定量分析为主, 未能体现中医药的整体性^[4], 因此, 亟需一种体现中药质量整体性的香附炮制程度快速辨识方法。

随着人工智能技术的快速发展, 电子鼻、电子舌、电子眼等现代电子智能感官技术相继问世, 它能模拟人的感觉器官, 将传统评价指标“气味、色泽、质地”等特征客观化、数据化、可视化, 更加精确地揭示中药饮片性状特征的变化和主要质控指标或整体性化学成分之间的联系, 可用于中药质量控制^[5-7]。基于“整体观”, 国内学者陆续提出新的中药质量控制研究模式和方法^[8-11]。肖小河等^[12]提出中药整合质量观(中药大质量观)的质控策略, 其核心思想是以生物评价为核心, 并用传统感官评价和化学评价的多元化质量评价控制模式, 凸显了中药整体性的质量属性。在此基础上通过多信息融合技术进行大数据分析处理, 进一步挖掘数据价值, 提升数据分析作用^[13-14]。以附子为例的中药品质整

合评控, 首次利用品质综合指数, 融合性状规格、化学与生物评价结果, 采用归一化与集成化实现综合量化评价药材品质^[15]。因此, 本研究基于色差仪、电子鼻二类智能感官技术, 以及近红外光谱(near infrared spectroscopy, NIRS)技术采集香附炮制过程样品的智能感官信息和 NIRS 信息, 利用主成分分析-判别分析(principal component analysis-discriminant analysis, PCA-DA)等多种化学计量学方法分别建立感官数据单源、NIRS 数据单源、多源数据融合的香附炮制程度辨识模型, 并对二类智能感官与多源信息融合前后所建立模型进行对比分析, 以探讨智能感官与多源信息融合技术应用于香附炮制程度快速辨识的可行性, 同时为其他中药饮片的快速准确辨识提供参考。

1 仪器与材料

1.1 仪器

Thermo Antaris II 型傅里叶变换近红外光谱仪, 美国 Thermos 公司; CM-5 分光测色仪, 日本柯尼卡美能达有限公司; YB-300A 型高速多功能粉碎机, 永康市速锋工贸有限公司; Heracles NEO 超快速气相电子鼻, 法国 Alpha MOS 公司; JA1003B 型千分之一天平, 上海越平科学仪器制造有限公司; DHG-9146A 型电热恒温鼓风干燥箱, 上海精宏实验设备有限公司; XS105DU 型十万分之一电子天平, 瑞士梅特勒-托利多仪器有限公司。

1.2 材料

收集来自广西、广东、四川、湖北、山东、河南 6 个产地的生香附饮片(分别编号为 GX0、GD0、SC0、HB0、SD0、HN0), 其质量均符合《中国药典》2020 年版标准, 经浙江中医药大学黄真教授鉴定为莎草科莎草属植物莎草 *C. rotundus* L. 的干燥根

茎。正构烷烃 C₆~C₁₆ 混合对照品,批号 A10142930, 美国 Restek 有限公司; 乙酸, 分析纯, 批号 A/0406/PB08, 国药集团化学试剂有限公司。

2 方法与结果

2.1 香附炮制过程样品及其粉末的制备

根据《中国药典》2020年版规定,生香附采用醋炙法进行炮制,过程如下:每100 kg 生香附加入米醋 20 kg (总酸≥0.05 g/mL),闷润 2 h。炒药机温度设定 208 °C,加入润好的生香附,炒制 30 min,每隔 3 min 取样 1 次,其中 11 min 为成品取样点,摊开晾凉,即得 72 批香附炮制过程样品。

综合前期检测的挥发油含量和浙江中医药大学中药饮片有限公司老药工师傅杨林加的经验判断,将 72 批香附炮制品区分为生品 6 批 (GX0、GD0、

SC0、HB0、SD0、HN0),炮制不及 18 批 (分别编号为 GX3、GD3、SC3、HB3、SD3、HN3, GX6、GD6、SC6、HB6、SD6、HN6, GX9、GD9、SC9、HB9、SD9、HN9),炮制适度 18 批 (分别编号为 GX11、GD11、SC11、HB11、SD11、HN11, GX12、GD12、SC12、HB12、SD12、HN12, GX15、GD15、SC15、HB15、SD15、HN15),炮制太过 30 批 (分别编号为 GX18、GD18、SC18、HB18、SD18、HN18, GX21、GD21、SC21、HB21、SD21、HN21, GX24、GD24、SC24、HB24、SD24、HN24, GX27、GD27、SC27、HB27、SD27、HN27, GX30、GD30、SC30、HB30、SD30、HN30)。香附炮制过程饮片性状变化图见图 1。将各样品粉碎过四号筛 (65 目),即得 72 批香附样品粉末。

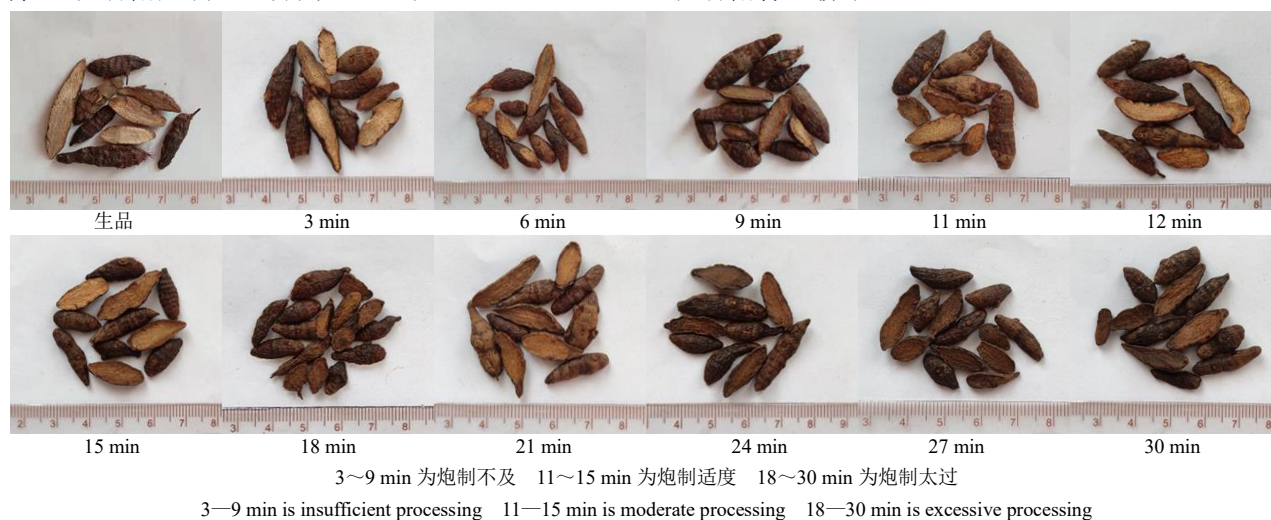


图 1 山东产香附炮制过程饮片性状变化图

Fig. 1 Changes in characteristics of decoction pieces during processing of *Cyperi Rhizoma* in Shandong Province

2.2 智能感官及 NIRS 信息的获取

2.2.1 基于色差仪获取样品色度信息

(1) 测试方法: 色差仪的光源为脉冲氙弧灯,光源观察角度为 8°,测量口径为 30 mm,测量波长为 360~740 nm,测定模式为 SCE 反射光模式。取“2.1”项下香附样品粉末,平铺于测试口测定 L*值 (亮度)、a*值 (红绿色度)、b*值 (黄蓝色度),平行测定 3 次,取平均值。

(2) 色差仪输出信息值: 基于色差仪分别得到 72 批香附样品粉末的 L*值、a*值、b*值,并通过公式 $E_{ab}^* = (L^{*2} + a^{*2} + b^{*2})^{1/2}$ 得到总色度值,共 4 种变量,均可参与建模。

2.2.2 基于电子鼻获取样品气味信息

(1) 测试方法: 取“2.1”项下香附样品粉末,称取 1.000 g,装于电子鼻专用顶空进样瓶中,隔垫

密封,放于自动进样器上进行检测。电子鼻测试条件为孵化温度 50 °C,孵化时间 15 min,捕集阱初始温度 40 °C,捕集阱最终温度 250 °C,捕集阱分流速率 10 mL/min,捕集持续时间 50 s,进样口温度 200 °C,进样口压力 10 kPa;进样量 5000 μL,注射速度 125 μL/s,注射持续时间 45 s,柱温初始炉温 50 °C,柱温的程序升温方式 1 °C/s 至 80 °C,3.0 °C/s 至 250 °C,保持 21 s,采集时间 110 s;FID 增益 12。以正构烷烃 C₆~C₁₆ 标准液作为矫正液。

(2) 电子鼻输出信息值: 计算 Kovats 保留指数和各气味色谱峰双柱 (MXT-5 和 MXT-1701) 相对保留指数,与 Aro Chem Base 数据库匹配,鉴定得到 21 个气味成分信息。基于电子鼻得到 72 批香附样品的 21 个气味成分数据,均可参与建模。

2.2.3 基于 NIRS 技术获取样品 NIRS 信息

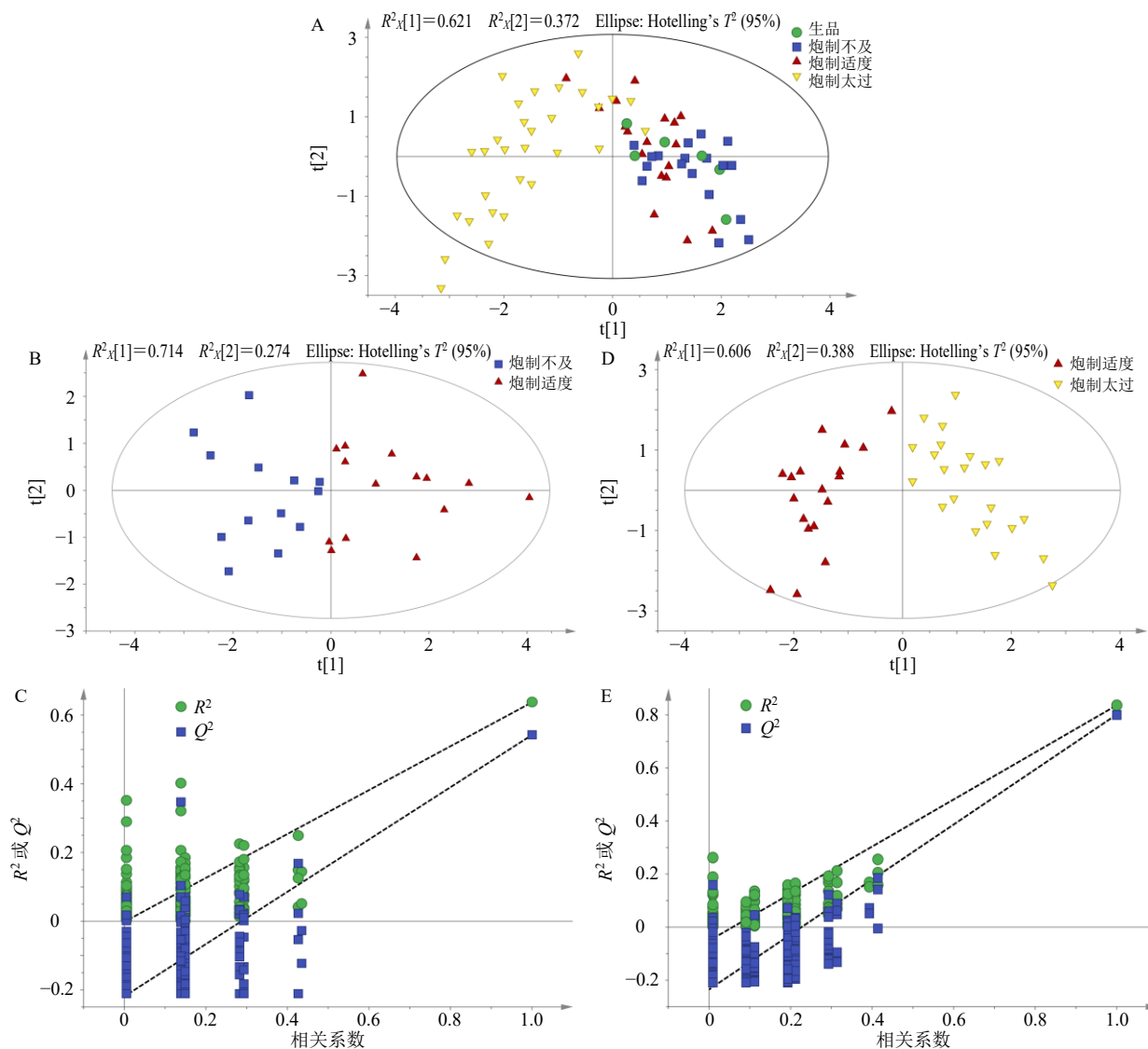
(1) 测试方法: 取“2.1”项下香附样品, 粉碎, 取四号筛和五号筛中间的粉末 (65~80 目), 混匀后平铺在石英杯中, 装样厚度为 0.4 cm, 设置扫描次数为 32 次, 分辨率为 16 cm⁻¹, 采集方式为积分球漫反射, 波数区间为 4000~10 000 cm⁻¹, 温度为 (25±2) °C, 相对湿度为 25%~35%, 以空气为参比, 扣除背景, 采集 NIRS 图。每份样品扫描 3 次, 取平均光谱作为样品的 NIRS。

(2) NIRS 输出信息值: 基于 NIRS 技术得到 72 批香附样品的 NIRS 数据, 均可参与建模。

2.3 香附炮制程度辨识模型的建立及优化

基于色差仪和电子鼻采集的 72 批香附炮制样品的感官信息输入 SIMCA 14.1 软件建立辨识模型; 基于 NIRS 技术采集的 72 批香附炮制样品的 NIRS 信息输入 NIRS TQ Analyst 8.3.126 软件导出光谱数据格式 (csv) 文件, 然后导入 SIMCA 14.1 软件建立辨识模型; 基于多源信息融合技术建立香附炮制程度辨识模型。

2.3.1 基于色差仪的香附炮制程度辨识模型结果 PCA-DA 辨识模型及结果如图 2-A 所示, 生品、炮



A-4 类不同炮制程度粉末色度 PCA 聚类图 B-炮制不及和炮制适度粉末色度 PLS-DA 聚类图 C-炮制不及和炮制适度粉末色度置换检验图 D-炮制太过和炮制适度粉末色度 PLS-DA 聚类图 E-炮制太过和炮制适度粉末色度置换检验图

A-PCA clustering plots of four kinds of powder chromaticity with different processing degrees B-PLS-DA clustering plots of less processed and moderate processed of powder chromaticity C-permutations diagram of less processed and moderate processed of powder chromaticity D-PLS-DA clustering plots of processing too much and processing moderate powder chromaticity E-permutations diagram of processing too much and processing moderate powder chromaticity

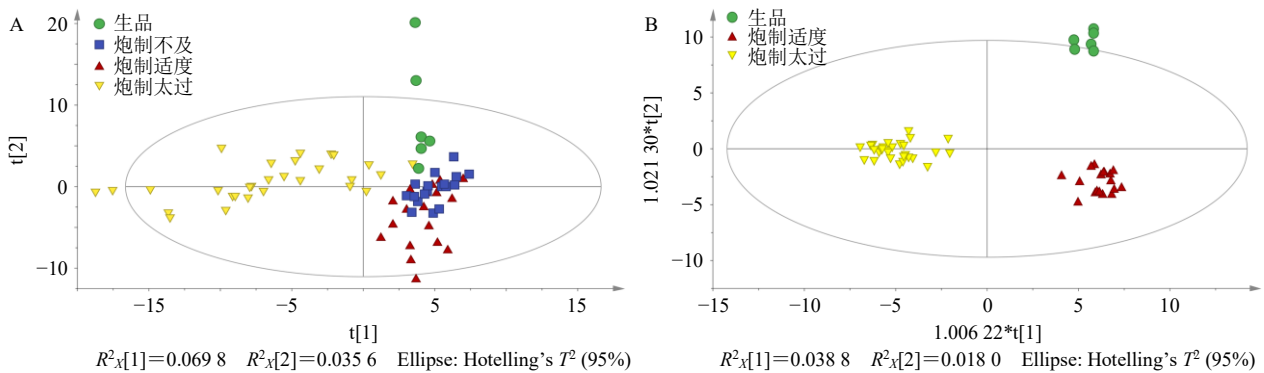
图 2 72 批香附炮制全过程粉末色度分析图

Fig. 2 Powder chromaticity analysis diagrams of whole process of processing 72 batches of *Cyperi Rhizoma*

制不及样品、炮制适度样品、炮制太过样品，这 4 类粉末颜色不能各聚为一类，其中炮制太过和其他 3 类基本能区分，生品、炮制不及、炮制适度 3 类样品有很大程度的重叠。生香附饮片来源于不同产地、不同生产厂家，生品粉末颜色本身差异较大，可能是导致 PCA-DA 辨识模型中样本重叠的原因。PLS-DA 辨识模型及结果如图 2-B 所示，炮制不及和炮制适度样品粉末色度能各聚为一类，经过 200 次置换验证，结果如图 2-C 所示， $R^2 = -0.00171$ ， $Q^2 = -0.219$ ，且左侧随机排列得到的 R^2 和 Q^2 均要小于右侧的原始值，说明模型可靠，未出现过拟合现象。如图 2-D 所示，炮制太过和炮制适度粉末色

度能各聚为一类， $R^2 = -0.0512$ ， $Q^2 = -0.235$ ，且左侧随机排列得到的 R^2 和 Q^2 均要小于右侧的原始值，说明模型可靠，未出现过拟合现象。PLS-DA 辨识模型结果说明凭借色度能区分色度差异较大的香附炮制品，但无法准确地判别香附炮制程度，还需进一步结合其他性状指标的量化来确定。

2.3.2 基于电子鼻的香附炮制程度辨识模型结果 正交偏最小二乘-判别分析 (orthogonal partial least squares-discriminant analysis, OPLS-DA) 辨识模型及结果如图 3 所示，生品、炮制不及样品、炮制适度样品、炮制太过样品 4 类不能各自聚为一类，其中生品、炮制适度、炮制太过 3 组样品气味信息判



A-香附 4 类不同炮制程度样品气味 OPLS-DA 图 B-香附 3 类不同炮制程度样品气味 OPLS-DA 图
A-odor OPLS-DA of four kinds of *Cyperi Rhizoma* with different processing degrees B-odor OPLS-DA of three kinds of *Cyperi Rhizoma* with different processing degrees

图 3 72 批香附炮制全过程气味分析图

Fig. 3 Odor analysis diagrams of whole process of processing 72 batches of *Cyperi Rhizoma*

别效果较好，炮制不及和炮制适度 2 组样品因气味信息比较接近有较大的重叠。以上结果表明单以气味信息也无法准确地判别香附不同炮制程度样品。

2.3.3 基于 NIRS 的香附炮制程度辨识模型结果 OPLS-DA 辨识模型及结果如图 4 所示，生品、炮制不及、炮制适度、炮制太过 4 类样品两两之间区分明显，但无法在一个模型中将 4 组样品进行区分，需进一步对模型进行优化。各模型具体参数见表 1， Q^2 均大于 0.6，表明模型预测效果可靠。

2.3.4 基于多源信息融合的香附炮制程度辨识模型结果

(1) 数据处理：将 72 批样品的色度数据、电子鼻数据、NIRS 数据采用归一化进行无量纲化处理。

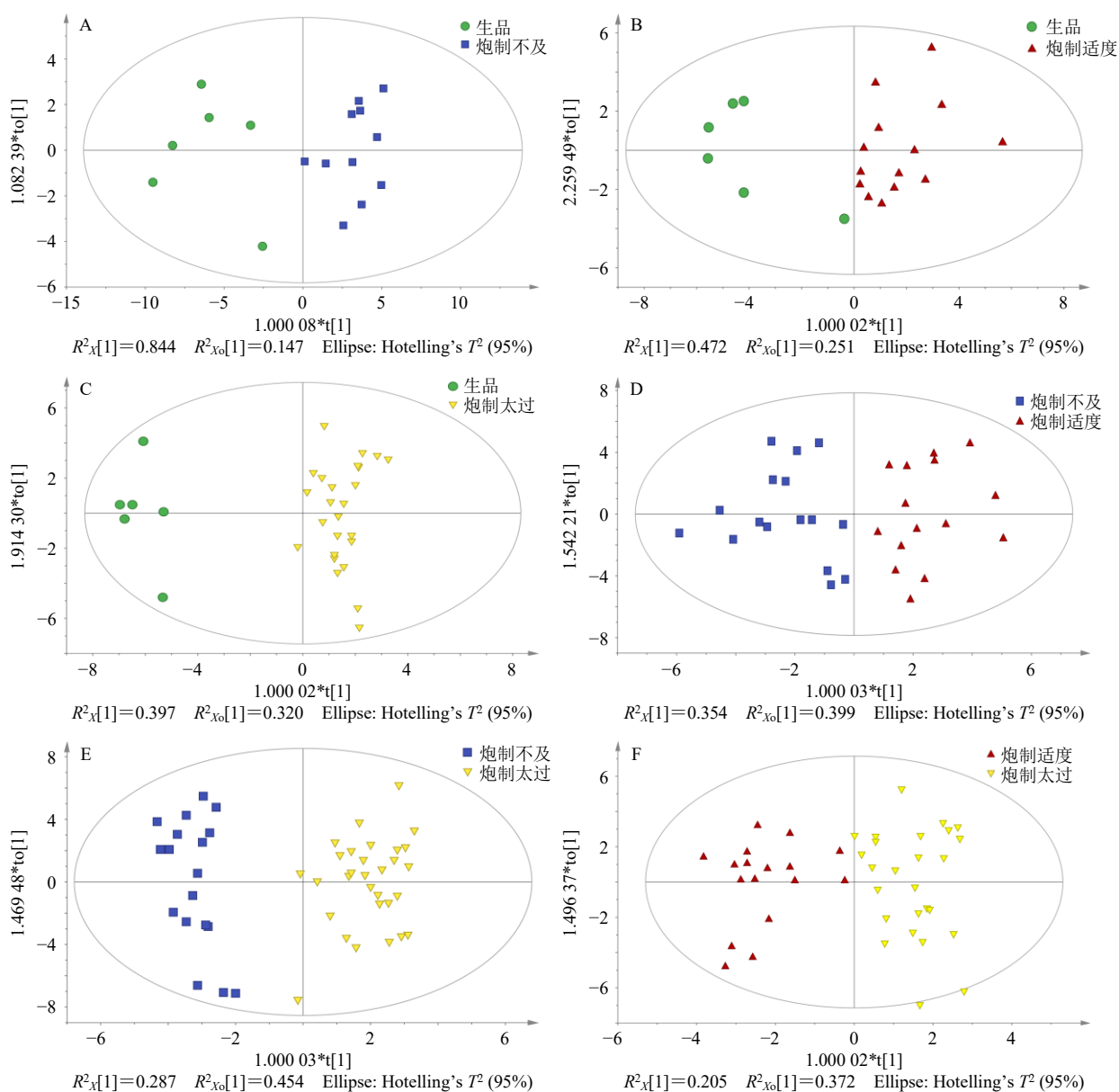
$$x' = [x - \min(x)] / [\max(x) - \min(x)]$$

鉴于 NIRS 和电子鼻峰面积的数据量较大，需先对数据进行降维处理，将原始数据导入 SPSS 25.0 软件，用 PCA 提取前几个累积贡献率较大的主成分，

采用中级数据融合策略，将提取的主成分和 4 个色度值融合。最终采用 PCA 法提取了前 2 个 NIRS 近红外主成分，累积贡献率大于 95%；提取了前 8 个电子鼻主成分，累积贡献率大于 80%，具体结果见表 2，将这 14 个变量用于后续构建定性拟合模型。

(2) 基于遗传算法 (genetic algorithm, GA)-反向传播 (back propagation, BP) 神经网络算法 (GA-BP neural network algorithm, GA-BPNNA) 的香附炮制程度辨识模型：将 72 批香附炮制样品按照类别 1:3:3:5 随机划分训练集 (60 个) 和验证集 (12 个)，根据 Lasso 回归计算各变量的权重系数，选择权重较大的变量作为 BPNNA 的输入层 (X_i) 数据，然后利用 GA 确定最佳权重和偏置，即 ω_i 和 b_i ，激活函数为 Logistic，训练函数为 trainlm，最后输出得到 Z，即预测分类结果，BPNNA 工作流程图见图 5。

基于 GA-BPNNA 构建了 72 批香附炮制程度辨



A-生品 vs 炮制不及模型 B-生品 vs 炮制适度模型 C-生品 vs 炮制太过模型 D-炮制不及 vs 炮制适度模型 E-炮制不及 vs 炮制太过模型 F-炮制适度 vs 炮制太过模型

A-model of the raw vs less processed B-model of the raw vs moderate processed C-model of the raw vs processing too much D-model of the less processed vs moderate processed E-model of the less processed vs processing too much F-model of the moderate processed vs processing too much

图 4 NIRS 定性判别模型分类图

Fig. 4 Classification diagrams of NIRS qualitative discrimination model

表 1 NIRS 的 OPLS-DA 模型的参数

Table 1 Parameters of NIRS OPLS-DA model

模型	$R^2_X(\text{cum})$	$R^2_Y(\text{cum})$	$Q^2(\text{cum})$
生品 vs 炮制不及	0.991	0.848	0.775
生品 vs 炮制适度	0.992	0.746	0.655
生品 vs 炮制太过	0.997	0.940	0.917
炮制不及 vs 炮制适度	0.998	0.750	0.644
炮制适度 vs 炮制太过	0.989	0.823	0.805
炮制不及 vs 炮制太过	0.997	0.901	0.881

识模型, 如表 3 所示, “5 个 X, 6 个神经元” 模型的预测准确度最好, 该模型下的具体权重 (ω) 和偏置 (b) 见表 4。准确度由训练集建模得出, 验证集用来验证模型的准确度。由图 6 可知, 训练次数达 1000 次时模型的准确度良好。

(3) 结果分析: 本研究融合色差仪、电子鼻、NIRS 数据, 采用 GA-BPNN 建立了香附炮制程度辨识模型, 可同时预测 4 类样品 (生品、炮制不及、炮制适度、炮制太过), 相比于单源的色差仪、电子

鼻、NIRS 技术香附炮制程度辨识模型, 该融合模型预测的效率和准确度显著提高, 模型预测准确度 R 大于 0.93, 表明模型建立成功。

$$L = \sum_{i=1}^n \begin{cases} -W_i - X_j - \lg(1 - X_j) & i=j \\ -W_i(1 - X_j)\lg X_i & i \neq j \end{cases}$$

$$i=1, 2, 3, 4; j=1, 2, 3, 4; W_i=[15/28, 5/28, 5/28, 3/28]$$

表 2 14 个变量原始数据
Table 2 Original data of 14 variables

组别	L^*	a^*	b^*	E^*	NIRS				电子鼻							
					PC1	PC2	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8		
GX0	57.120	7.100	18.270	60.390	327.887	-51.130	0.117	-0.074	0.011	0.022	-0.028	0.009	0.028	0.012		
GX3	59.720	6.660	17.660	62.630	299.135	-44.560	0.143	-0.082	0.012	0.022	-0.029	0.007	0.029	0.014		
GX6	56.930	6.400	16.470	59.610	306.332	-40.097	0.164	-0.079	0.012	0.023	-0.030	0.009	0.025	0.007		
GX9	55.900	6.970	17.440	58.970	299.277	-38.214	0.293	-0.098	0.008	0.041	-0.030	-0.001	0.023	-0.002		
GX11	53.780	7.360	17.230	56.950	310.357	-41.322	0.161	-0.074	0.010	0.020	-0.031	0.005	0.023	0.011		
GX12	53.450	7.520	17.240	56.660	319.200	-42.890	0.136	-0.078	0.008	0.028	-0.027	0.004	0.028	0.012		
GX15	53.150	8.370	18.720	56.960	322.052	-47.157	0.123	-0.038	0.012	0.013	-0.015	0.000	0.046	0.005		
GX18	49.920	9.800	20.210	54.740	317.888	-41.813	0.130	-0.071	0.017	0.041	-0.025	0.025	0.000	0.028		
GX21	45.330	10.180	19.850	50.520	319.297	-38.880	0.136	-0.042	0.012	0.007	-0.020	-0.002	0.042	0.008		
GX24	42.130	9.950	18.690	47.150	308.562	-33.218	0.125	-0.078	0.017	0.044	-0.023	0.027	0.010	0.021		
GX27	40.740	9.290	17.480	45.300	322.366	-25.939	0.128	-0.081	0.013	0.037	-0.023	0.017	0.019	0.019		
GX30	37.530	9.100	16.100	41.840	316.673	-25.629	0.132	-0.076	0.016	0.042	-0.020	0.021	0.014	0.023		
GD0	56.540	5.070	16.000	58.980	319.186	-35.055	0.160	-0.077	0.017	0.061	-0.015	0.032	-0.006	0.027		
GD3	58.010	4.980	15.860	60.350	294.439	-33.347	0.145	-0.075	0.018	0.045	-0.022	0.030	0.001	0.026		
GD6	58.080	4.460	15.000	60.150	298.485	-31.713	0.156	-0.032	0.017	0.008	-0.015	0.004	0.053	-0.003		
GD9	55.290	4.740	15.000	57.480	316.990	-33.832	0.147	-0.076	0.021	0.073	-0.020	0.037	-0.004	0.022		
GD11	55.290	5.180	15.420	57.630	318.084	-32.766	0.117	-0.080	0.012	0.054	-0.018	0.025	0.007	0.024		
GD12	52.800	5.310	15.080	55.170	324.607	-30.546	0.152	-0.023	0.017	0.008	-0.013	0.001	0.056	-0.004		
GD15	51.020	6.430	16.040	53.870	324.833	-30.357	0.146	-0.015	0.015	0.012	-0.007	0.002	0.059	-0.010		
GD18	48.300	8.690	18.750	52.540	344.626	-39.077	0.342	-0.006	0.024	-0.078	-0.012	0.004	0.042	0.031		
GD21	44.650	9.270	18.570	49.230	336.263	-31.563	0.270	0.003	0.021	-0.082	-0.008	0.000	0.052	0.039		
GD24	41.330	8.890	17.390	45.720	324.477	-23.631	0.109	-0.077	0.011	0.042	-0.016	0.017	0.027	0.013		
GD27	37.860	8.570	16.200	42.060	331.514	-26.408	0.114	-0.075	0.017	0.063	-0.013	0.033	0.003	0.028		
GD30	35.750	8.210	15.080	39.660	339.553	-12.792	0.127	-0.009	0.016	0.008	-0.006	0.002	0.069	-0.012		
SC0	57.700	6.290	17.920	60.740	339.625	-54.756	0.146	-0.034	0.014	0.001	-0.018	0.002	0.047	0.004		
SC3	58.180	6.400	18.040	61.250	293.864	-39.490	0.123	-0.059	0.011	0.016	-0.025	0.005	0.031	0.010		
SC6	56.820	6.640	18.490	60.120	301.674	-44.662	0.125	-0.078	0.013	0.040	-0.026	0.023	0.008	0.022		
SC9	55.100	7.070	18.390	58.510	337.077	-48.023	0.129	-0.078	0.016	0.033	-0.026	0.020	0.016	0.020		
SC11	53.500	7.150	18.080	56.930	335.718	-47.038	0.135	-0.076	0.019	0.042	-0.023	0.027	0.011	0.021		
SC12	55.900	7.950	20.130	59.940	340.755	-40.908	0.127	-0.046	0.015	0.018	-0.012	0.007	0.053	0.003		
SC15	52.940	9.330	21.640	57.950	331.193	-47.764	0.163	-0.032	0.015	0.003	-0.018	-0.001	0.052	-0.003		
SC18	48.870	10.810	22.750	54.980	327.217	-46.972	0.139	-0.082	0.017	0.045	-0.025	0.029	0.004	0.025		
SC21	44.170	10.760	21.170	50.150	317.251	-37.611	0.168	-0.084	0.014	0.040	-0.025	0.018	0.008	0.024		
SC24	42.410	10.780	20.660	48.390	318.193	-38.189	0.149	-0.021	0.013	0.004	-0.012	-0.001	0.061	-0.008		
SC27	39.760	10.450	19.090	45.330	321.522	-32.846	0.181	-0.026	0.024	0.009	-0.019	0.000	0.055	-0.001		
SC30	38.480	10.490	18.560	44.000	324.338	-34.980	0.164	-0.079	0.020	0.054	-0.024	0.033	-0.003	0.028		

续表 2

组别	L^*	a^*	b^*	E^*	NIRS		电子鼻							
					PC1	PC2	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8
HB0	53.850	7.660	19.090	57.650	348.415	-53.537	0.383	0.014	0.032	-0.087	-0.012	-0.001	0.051	0.032
HB3	54.890	7.090	18.100	58.230	312.014	-42.100	0.197	-0.028	0.011	-0.001	-0.016	-0.010	0.050	0.002
HB6	52.990	7.540	18.490	56.620	320.471	-39.893	0.202	-0.077	0.019	0.051	-0.027	0.030	-0.014	0.031
HB9	51.770	7.530	18.030	55.340	320.664	-41.030	0.113	-0.033	0.011	0.022	-0.005	0.004	0.061	-0.005
HB11	51.760	7.890	18.490	55.530	313.704	-34.388	0.125	-0.024	0.013	0.015	-0.004	0.001	0.065	-0.006
HB12	49.680	9.630	20.250	54.500	324.117	-43.307	0.127	-0.035	0.010	0.016	-0.008	0.000	0.059	-0.004
HB15	51.410	8.680	19.330	55.610	307.814	-34.829	0.192	-0.087	0.012	0.036	-0.031	0.011	0.011	0.020
HB18	46.680	10.630	21.150	52.340	311.669	-40.019	0.140	-0.088	0.012	0.034	-0.029	0.011	0.024	0.019
HB21	42.620	10.430	19.790	48.130	315.326	-31.879	0.176	-0.095	0.011	0.033	-0.031	0.009	0.028	0.015
HB24	40.070	10.080	18.740	45.370	319.271	-29.923	0.200	-0.099	0.012	0.028	-0.036	0.004	0.032	0.016
HB27	37.130	9.440	16.870	41.860	322.547	-23.208	0.192	-0.095	0.015	0.023	-0.038	0.010	0.031	0.015
HB30	35.080	9.210	15.810	39.570	324.043	-21.214	0.126	-0.084	0.018	0.052	-0.018	0.029	0.012	0.027
SD0	50.720	7.760	18.650	54.590	352.276	-49.800	0.143	-0.094	0.021	0.046	-0.029	0.030	0.015	0.025
SD3	50.910	7.280	17.470	54.320	290.316	-33.225	0.133	-0.100	0.017	0.049	-0.025	0.024	0.027	0.021
SD6	52.170	7.470	18.610	55.890	333.451	-45.326	0.134	-0.109	0.019	0.052	-0.025	0.025	0.038	0.017
SD9	51.110	8.070	19.000	55.120	337.120	-47.756	0.138	-0.106	0.021	0.042	-0.032	0.023	0.041	0.013
SD11	50.960	8.620	19.720	55.320	338.354	-48.370	0.161	-0.087	0.013	0.037	-0.026	0.015	0.019	0.020
SD12	50.980	9.310	20.740	55.820	339.615	-51.954	0.254	-0.074	0.027	-0.038	-0.022	0.034	0.029	0.052
SD15	47.730	10.780	21.470	53.440	322.347	-40.928	0.294	-0.089	0.023	-0.017	-0.027	0.025	0.028	0.036
SD18	42.650	12.050	21.330	49.180	334.262	-43.478	0.350	-0.088	0.026	-0.035	-0.024	0.023	0.028	0.050
SD21	37.890	10.960	18.280	43.480	328.366	-35.195	0.285	-0.086	0.026	-0.033	-0.032	0.025	0.030	0.044
SD24	34.900	9.870	15.700	39.520	331.115	-27.928	0.183	-0.087	0.015	0.044	-0.027	0.016	0.011	0.025
SD27	32.590	9.140	13.950	36.610	343.085	-21.868	0.156	-0.093	0.012	0.044	-0.027	0.015	0.023	0.020
SD30	31.360	8.570	12.810	34.940	342.732	-12.079	0.164	-0.097	0.014	0.040	-0.029	0.013	0.030	0.017
HN0	51.790	9.000	19.530	56.080	343.522	-59.191	0.163	-0.100	0.017	0.039	-0.033	0.018	0.033	0.013
HN3	60.330	7.370	18.500	63.530	300.003	-46.239	0.176	-0.105	0.015	0.037	-0.031	0.012	0.038	0.016
HN6	56.830	7.840	18.540	60.290	319.482	-50.548	0.197	-0.087	0.014	0.048	-0.026	0.019	0.004	0.027
HN9	58.160	7.840	18.870	61.640	294.206	-45.694	0.248	-0.092	0.022	0.030	-0.043	0.013	0.013	0.024
HN11	56.070	8.380	19.440	59.930	319.313	-49.001	0.205	-0.103	0.018	0.040	-0.036	0.015	0.023	0.023
HN12	55.650	7.940	18.530	59.190	314.347	-44.530	0.204	-0.104	0.019	0.033	-0.040	0.015	0.030	0.017
HN15	55.410	8.620	19.600	59.400	315.966	-41.253	0.187	-0.103	0.019	0.036	-0.037	0.017	0.033	0.015
HN18	53.320	8.660	19.140	57.310	315.386	-42.846	0.132	-0.015	0.014	0.008	-0.006	0.002	0.068	-0.011
HN21	52.970	9.490	20.370	57.540	314.000	-35.642	0.131	-0.085	0.017	0.065	-0.014	0.034	0.003	0.029
HN24	50.980	9.580	20.710	55.860	327.542	-40.616	0.144	-0.092	0.017	0.073	-0.014	0.040	-0.005	0.034
HN27	48.550	10.160	20.970	53.850	303.037	-35.459	0.137	-0.092	0.018	0.056	-0.021	0.034	0.016	0.020
HN30	43.190	10.190	19.390	48.420	315.205	-36.231	0.153	-0.104	0.017	0.058	-0.023	0.030	0.018	0.025

PC1~8-主成分 1~8

PC1—8-principal component 1—8

以上为 Lasso 具体损失函数, i 为实际类别, j 为预测类别, W_{δ} 为添加的权重值。

3 讨论

本研究融合色差仪技术、电子鼻技术和 NIRS 技

术，可全面反映样品外在性状和内在物质的整体性变化，体现出香附饮片质量评价研究的整体性和全

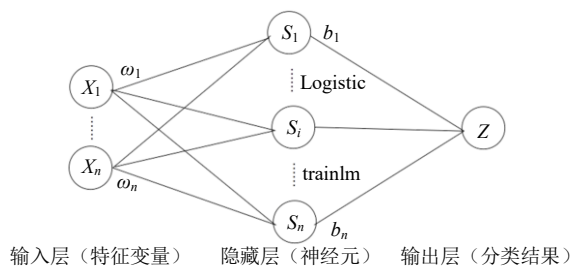


图5 BPNN定性拟合模型工作流程图

Fig. 5 Flow chart of BPNN qualitative fitting model

表3 GA-BPNN定性拟合模型的准确度

Table 3 Accuracies of GA-BPNN qualitative fitting model

序号	模型	总准确度	训练集准确度	验证集准确度
1	4个X, 4个神经元	0.778	0.833	0.767
2	4个X, 5个神经元	0.792	0.833	0.783
3	4个X, 6个神经元	0.806	1.000	0.767
4	5个X, 4个神经元	0.875	0.912	0.867
5	5个X, 5个神经元	0.903	1.000	0.883
6	5个X, 6个神经元	0.931	1.000	0.917

表4 5个X、6个神经元GA-BPNN定性拟合模型具体参数

Table 4 Specific parameters of GA-BPNN qualitative fitting model with five X and six neurons

预测模型	ω_1				b_1		ω_2				b_2	
5个X, 6个神经元	11.039	15.094	-27.264	-6.529	-24.776	11.514	-26.053	-26.386	29.193	-0.662	9.182	
	23.524	29.674	-3.486	-2.647	8.631	-20.321	-18.152	-26.010	30.124	-3.676	-4.711	
	19.561	12.293	-7.705	13.307	8.889	-25.121	-14.162	-13.705	-25.932	26.434	-12.621	
	-25.971	48.673	1.925	-11.033	-4.474	4.033	-13.634	30.857	-15.125	-8.239	-9.006	
	30.548	15.469	-23.749	-17.195	-2.047	-12.577	-1.265	15.819	-29.045	19.685	NAN	
	-16.746	-17.922	-37.944	-0.380	19.288	15.411	23.914	17.690	-19.530	-5.508	NAN	

NAN: 指无数值意义

NAN: no numerical significance

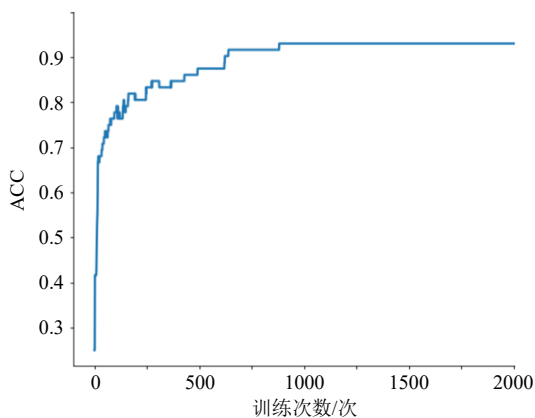


图6 模型训练变化曲线图

Fig. 6 Variation curve of model training

面性。基于色差仪色度信息、电子鼻气味信息、NIRS数据3个指标建立单源和多源的香附炮制程度快速辨识模型，基于单源建立的香附炮制程度辨识模型无法准确辨识香附炮制程度，采用遗传神经算法和数据融合技术建立的香附炮制程度辨识模型能快速准确辨识香附炮制质量，为香附炮制过程质量控制的应用提供参考。GA-BPNN等级预测模型最大的优势为融合了色差仪、电子鼻和NIRS技术，多角度全方面了解中药饮片信息，精准评价中药品质，明显提高炮制程度辨识的准确性和可靠性，同时也

能加快检测速度。GA-BPNN可反映数据之间复杂的非线性关系，相较于常见NIRS定性定量分析方法，如PLS、PCA、last millennium reanalysis (LMR)等具有更强的学习率和自适应率，模型的稳定性和预测准确率更好^[16-18]。

由于中药成分复杂，依据单一的检测分析技术往往无法获得充分的样本信息，信息表征不全面会导致中药质量检测精度较低。数据融合技术是近些年逐渐发展起来的一项分析技术，它能融合不同检测仪器获取的数据，从信息互补的角度增强检测精度，降低检测误差，实现整体质量评价的目的。目前已有研究人员就中药品质评价、产地鉴别、伪劣品鉴定方面进行了相关研究，如基于气味、颜色、化学成分建立的川芎质量等级识别模型；基于紫外可见光谱、NIRS和UHPLC建立的滇重楼来源鉴定模型^[19-21]。

本研究主要收集了山东、湖北、河南、广东、广西、四川共6个产地的生香附，涵盖了市场上绝大部分香附的产地，表明所建立的香附炮制全过程辨识模型具有普适性。而正因为生产产家、生产批次等不同，生品的外观性状不统一，其中颜色是导致基于色差仪建立的香附辨识模型中样本重叠的主

要原因。本研究采用的色差仪、电子鼻技术属于智能感官技术，能对中药香附的颜色、气味进行量化表达，且因其整合了仿生感官和现代仪器分析2类技术，故兼有“感官”的响应和仪器分析“量化”的双重优势，具有分析快速、灵敏度高、重复性好、可靠性强的优点，实现了感官评价的定量化和标准化^[22]。此外，智能感官技术不仅能改善现代仪器检测的耗时长问题，而NIRS技术也同样属于一种快速检测技术，而且具有全息性特点，可以提取药材的化学性质、物理性质，甚至生物学属性^[4]，通过多源信息融合技术即可实现对中药的颜色、气味、化学成分、物理属性等全面的衡量，解决单一检测技术存在的检测精度低、模型预测差等问题，实现对中药香附质量的快速鉴别。

4 结论

本研究中，基于单源的色差仪、电子鼻、NIRS技术建立的香附炮制程度快速辨识模型均无法同时辨识4类香附炮制品，存在重叠样本，归咎于单一信息量少、信息表征不全面导致中药质量检测精度低；而基于多源信息融合的香附炮制程度快速辨识模型具有良好的预测能力，因其来源包含了色差仪、电子鼻、NIRS的相关数据，拥有颜色、气味、密度、组成、化学成分等诸多信息量，更好地代表了样品的综合信息，提高中药质量评价准确性。基于智能感官和多源信息融合的香附炮制程度辨识模型能快速、准确区分不同炮制品，为香附炮制过程质量识别提供了一定的科学依据。

利益冲突 所有作者均声明不存在利益冲突

参考文献

- [1] 梁·陶弘景集. 尚志钧辑校. 名医别录 [M]. 北京: 人民卫生出版社, 1986: 154.
- [2] 中国药典 [S]. 一部. 2020: 270.
- [3] 刘欢, 张孟历, 于猛, 等. 醋制香附挥发油抗抑郁活性及化学成分分析 [J]. 药物评价研究, 2020, 43(3): 436-442.
- [4] 刘晓娜, 车晓青, 李德芳, 等. 基于多源信息融合的中药质量标志物与质量评价研究模式 [J]. 中草药, 2019, 50(19): 4576-4581.
- [5] 李文敏, 吴纯洁, 艾莉, 等. 基于电子鼻、电子舌技术实现中药性状气味客观化表达的展望 [J]. 中成药, 2009, 31(2): 282-284.
- [6] 李学林, 李慧玲, 刘瑞新. 电子舌技术的应用研究述评 [J]. 中医学报, 2013, 28(2): 247-249.

- [7] 任延娜, 冯文豪, 李涵, 等. 基于人工智能感官与多源信息融合技术的五味药性二分类辨识方法探讨 [J]. 中草药, 2023, 54(10): 3080-3092.
- [8] 贾晓斌. 基于整体观中药物质基础认识与创新中药研发 [J]. 世界科学技术—中医药现代化, 2017, 19(9): 1430-1434.
- [9] 姜华, 高原, 杨景明, 等. 源于“整体观”思想的中药质量评价方法研究概述 [J]. 中国中药杂志, 2015, 40(6): 1027-1031.
- [10] 吴茜, 毕志明, 李萍, 等. 基于整体观的中药药效物质基础的生物活性筛选/化学在线分析研究新进展 [J]. 中国药科大学学报, 2007, 38(4): 289-293.
- [11] 孙国祥, 孙万阳, 闫慧, 等. 中药整体质量控制标准体系构建和中药一致性评价步骤 [J]. 中南药学, 2019, 17(3): 321-331.
- [12] 肖小河, 金城, 鄢丹, 等. 中药大质量观及实践 [J]. 中草药, 2010, 41(4): 505-508.
- [13] 陈昭, 马莉, 张百霞, 等. Drupal 在中药信息融合中的应用 [J]. 世界科学技术—中医药现代化, 2014, 16(12): 2563-2567.
- [14] 丁志平, 王家辉, 乔延江. 中药信息学研究浅释 [J]. 中国中医药信息杂志, 2003, 10(4): 92-94.
- [15] 白钢, 侯媛媛, 丁国钰, 等. 基于中药质量标志物构建中药材品质的近红外智能评价体系 [J]. 药学报, 2019, 54(2): 197-203.
- [16] 冯国红, 朱玉杰, 徐华东, 等. 应用遗传算法-主成分分析-反向传播神经网络的近红外光谱识别树种效果 [J]. 东北林业大学学报, 2020, 48(6): 56-60.
- [17] 曲楠, 窦森, 任玉林. 近红外光谱结合遗传算法优化的RBF神经网络对中药萃取物含量的预测 [J]. 东北师大学报: 自然科学版, 2017, 49(1): 98-104.
- [18] 邱丽媛, 梁泽华, 吴鑫雨, 等. 基于模式识别和遗传神经网络算法的醋香附近红外光谱等级评价和含量预测模型研究 [J]. 中草药, 2021, 52(13): 3818-3830.
- [19] 张娇, 王元忠, 杨维泽, 等. ATR-FTIR 和 UV-Vis 结合数据融合策略鉴别滇黄精产地 [J]. 光谱学与光谱分析, 2021, 41(5): 1410-1416.
- [20] 吴思俊, 王龙, 吴红根, 等. 基于多源数据融合技术的盐酸青藤碱制备工艺一致性评价方法研究 [J]. 分析测试学报, 2022, 41(1): 121-127.
- [21] 赵倩, 缪培琪, 李小莉, 等. 数据融合技术在中药分析领域中的应用进展 [J]. 中草药, 2023, 54(11): 3706-3714.
- [22] 侯富国, 桂新景, 王艳丽, 等. 基于智能视觉技术的白及饮片真伪快速辨识方法研究 [J]. 中草药, 2023, 54(2): 509-519.

[责任编辑 郑礼胜]