

基于智能视觉技术的白及饮片真伪快速辨识方法研究

侯富国^{1,4}, 桂新景^{1,2,3,4}, 王艳丽^{2,3}, 李海洋¹, 范雪花¹, 李 涵¹, 施钧瀚^{2,3,4}, 刘瑞新^{2,3,4*}, 李学林^{2,3,4*}

1. 河南中医药大学药学院, 河南 郑州 450046
2. 河南中医药大学第一附属医院 药学部, 河南 郑州 450000
3. 河南省中药饮片临床应用现代化工程研究中心, 河南 郑州 450000
4. 河南中医药大学呼吸疾病中医药防治省部共建协同创新中心, 河南 郑州 450046

摘 要: 目的 利用电子眼和色差仪分别对白及 *Bletilla striata* 及其伪品进行辨识, 旨在建立白及真伪的快速辨识方法。方法 首先基于《中国药典》、地方标准和 HPLC 指纹图谱对多采集来源的 134 批白及及其伪品样品进行综合鉴别, 以确定标杆辨识信息 (Y)。然后基于电子眼和色差仪分别获取上述样品的智能视觉信息 (X)。最后利用 Matlab 建立 $Y=F(X)$ 模型, 即分别建立白及真伪二分类 (白及、非白及) 和四分类 (白及、天麻 *Gastrodia elata*、玉竹 *Polygonatum odoratum*、黄花白及 *Bletilla ochracea*) 的主成分分析-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)、最小二乘-支持向量机 (least squares-support vector machine, LS-SVM) 和 K 最近邻 (K nearest neighbor, KNN) 辨识模型并验证。结果 经留一法交互验证, 基于电子眼建立的二分类最优辨识模型为 KNN 模型, 正判率为 99.25%; 四分类最优辨识模型为 LS-SVM 模型, 正判率为 97.01%。基于色差仪建立的二分类最优辨识模型为 KNN 模型, 正判率为 99.25%; 四分类最优辨识模型为 PLS-DA 模型, 正判率为 97.67%。基于 2 类智能视觉与多源信息融合技术建立二分类最优辨识模型为 PCA-DA 和 KNN 模型, 正判率均为 98.51%, 相较于融合前有所降低; 四分类最优辨识模型为 LS-SVM 模型, 正判率为 97.76%, 相较于融合前有所升高。结论 2 类智能视觉技术均可用于白及真伪的快速辨识。可为中药品质的快速准确辨识提供参考。

关键词: 白及; 智能视觉技术; 真伪快速辨识; 多源信息融合; 电子眼; 色差仪; 天麻; 玉竹; 黄花白及; 主成分分析-判别分析; 偏最小二乘-判别分析; 最小二乘-支持向量机; K 最近邻

中图分类号: R283.6 文献标志码: A 文章编号: 0253-2670(2023)02-0509-11

DOI: 10.7501/j.issn.0253-2670.2023.02.018

Authenticity rapid identification method of *Bletilla striata* decoction pieces based on intelligent vision technology

HOU Fu-guo^{1,4}, GUI Xin-jing^{1,2,3,4}, WANG Yan-li^{2,3}, LI Hai-yang¹, FAN Xue-hua¹, LI Han¹, SHI Jun-han^{2,3,4}, LIU Rui-xin^{2,3,4}, LI Xue-lin^{2,3,4}

1. School of Pharmacy, Henan University of Chinese Medicine, Zhengzhou 450046, China
2. Department of Pharmacy, The First Affiliated Hospital of Henan University of Chinese Medicine, Zhengzhou 450000, China
3. Henan Modernization Engineering Research Center for Clinical Application of Traditional Chinese Medicine Decoction Pieces, Zhengzhou 450000, China
4. Provincial Co-construction Collaborative Innovation Center for Prevention and Treatment of Respiratory Diseases with Traditional Chinese Medicine, Henan University of Chinese Medicine, Zhengzhou 450046, China

收稿日期: 2022-07-20

基金项目: 国家重点研发计划中医药现代化重点专项课题 (2017YFC1703400); 国家重点研发计划中医药现代化重点专项课题 (2017YFC1703401); 河南省科技攻关项目 (222102310377); 河南省中医科学研究基地专项 (2021JDZY104); 河南省中医药拔尖人才培养项目资助 (2019ZYBJ07); 河南省高层次人才特殊支持“中原千人计划”——“中原青年拔尖人才”项目 (ZYQR201912158); 河南省卫生健康中青年学科带头人专项 (HNSWJW-2020014); 2022 年协同创新中心研究生科研创新基金项目 (协同中心[2022]002 号)

作者简介: 侯富国, 硕士研究生, 从事中药品质评价及应用形式研究。Tel: (0371)66245142 E-mail: 1344864266@qq.com

***通信作者:** 刘瑞新, 硕士生导师, 主任药师, 从事中药饮片临床应用现代化关键技术研究。Tel: (0371)66233562 E-mail: liuruixin7@163.com
李学林, 博士生导师, 主任药师, 从事中药应用形式研究。Tel: (0371)66245142 E-mail: xuelinli450000@163.com

Abstract: Objective To identify Baiji (*Bletilla striata*) and its approximate decoction pieces by electronic eye and colorimeter, so as to establish a rapid identification method for authenticity of *B. striata* decoction pieces. **Methods** Firstly, 134 batches of *B. striata* and its counterfeit samples from multiple sources were comprehensively identified based on *Chinese Pharmacopoeia*, local standards and HPLC fingerprint to determine the benchmarking information (Y). Then intelligent visual information (X) of the above samples was obtained based on electronic eye and colorimeter. Finally, the $Y = F(X)$ model was established by Matlab, that is, the principal component analysis-discriminant analysis (PCA-DA), partial least squares-discriminant analysis (PLS-DA), least squares support vector machine (LS-SVM) and K nearest neighbor (KNN) of 134 batches of samples were established and verified for the Authenticity binary classification (*B. striata*, non-*Bletilla striata*) and the four classification [*B. striata*, Tianma (*Gastrodia elata*), Yuzhu (*Polygonatum odoratum*) and Huanghuabaiji (*Bletilla ochracea*)]. **Results** After the leave-one-out interactive verification, the optimal identification model based on electronic eye is KNN model, and the positive rate is 99.25%; The four-class optimal identification model is the LS-SVM model, and the positive rate is 97.01%. The optimal identification model based on colorimeter is KNN model, and the positive rate is 99.25%; the optimal identification model of four classifications is PLS-DA model, and the positive rate is 97.67%. Based on two types of intelligent vision and multi-source information fusion technology, the optimal identification model of two classifications is PCA-DA and KNN model, and the positive rate is 98.51%, which is lower than that before fusion; The optimal identification model of four classifications is LS-SVM model, and the positive rate is 97.76%, which is higher than that before fusion. **Conclusion** Two kinds of intelligent vision technology can be used for rapid identification of authenticity of *B. striata* decoction pieces. The establishment of this method can provide reference for the rapid and accurate identification of the quality of traditional Chinese medicine.

Key words: *Bletilla striata* (Thunb.) Reichb. f.; intelligent vision technology; authenticity rapid identification; multi-source information fusion; electronic eye; colorimeter; *Gastrodia elata* Bl.; *Polygonatum odoratum* (Mill.); *Bletilla ochracea* Schltr.; PCA-DA; PLS-DA; LS-SVM; KNN

白及为兰科白及属植物白及 *Bletilla striata* (Thunb.) Reichb. f. 的干燥块茎, 又名白芨, “及”和“芨”通用, 具有收敛止血、消肿生肌的功能, 是研究较多、应用较广的一味中药^[1]。随着人们对白及功能的不断挖掘, 其应用范围不断拓展, 市场需求逐步扩大^[2]。白及对生态环境要求严格, 繁殖、生长缓慢, 再加上人为长期掠夺性挖采, 导致野生资源遭到严重破坏^[3], 已被《国家重点保护野生植物名录》(第 2 批) 列为二级保护植物^[4]。近年来, 由于其资源紧缺而用量又大等特点, 市场上常出现其同属植物黄花白及 *B. ochracea* Schltr.、水白及(市场上白及一类混伪品的统称)等冒充或掺伪的现象, 同时也有出现玉竹 *Polygonatum odoratum* (Mill.) Druce 碎片、质量次等的天麻 *Gastrodia elata* Bl. 等白及近似品种掺伪的情况^[5]。另外, 白及伪品种类复杂, 性状、显微、薄层、高效液相等单一鉴别方法只能区分部分伪品, 很难区分所有伪品^[6]。虽然通过 DNA 条形码技术^[7]及鉴别方法的联用, 可成功区分白及真伪品, 但存在操作繁琐、耗时较长及技术要求高等问题。因此, 亟需一种能够快速、精准的白及真伪鉴别方法。

随着计算机和传感器技术的快速发展, 智能视觉、智能嗅觉、智能味觉等智能仿生技术相继出现。

其中应用最多的是智能视觉技术, 其将各种图像处理技术与神经网络、支持向量机等机器学习技术相结合, 将客观物体的形状、颜色、纹理等外观特征数字化, 抽取物体的外观特征参数并实现对目标对象的客观化鉴定或评价^[8]。其具有对观测物与被观测者都不会产生损伤、分辨精度高、速度快、操作简单、成本低等特点^[9], 且目前已成功应用于川贝母^[10]、地黄^[11]等中药饮片的鉴别。因此, 本研究利用《中国药典》、地方标准和 HPLC 指纹图谱的综合鉴别以获取标杆辨识信息 (Y), 然后基于常用的 2 类智能视觉技术电子眼和色差仪获取智能视觉信息 (X), 最后结合化学计量学方法对白及、天麻、玉竹及黄花白及进行分类辨识, 并对 2 类智能视觉数据融合前后所建立模型进行对比分析, 以探讨智能视觉技术应用于白及其伪品快速准确辨识的可行性, 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

UltiMate-3000 型高效液相色谱仪、紫外检测器, 赛默飞世尔科技有限公司; IRIS VA400 型电子眼, 法国 Alpha-MOS 公司; MiniScan EZ 型手持式色差仪, 美国 Hunter Lab 公司; BSA2245-CW 型万分之一电子天平、CP225D 型十万分之一电子分析

天平, 德国 Sartorius 公司; FW-100 型高速万能粉碎机, 60~180 目, 北京科伟永兴仪器有限公司; 药典四号筛, 0.25 μm 孔径, 浙江上虞市道墟五四仪器厂; LEICA DM1000 型电子显微镜, 上海 LAS V4.0 系统; TLC ViSuALi ZER2 型薄层成像仪, 瑞士卡玛公司; GZX-9146MBE 型电热鼓风干燥箱, 上海博迅实业有限公司医疗设备厂。

1.2 材料

共收集 134 批白及及其伪品样品, 其中 1~45 为白及饮片样品, 46~75 为天麻样品, 76~105 为玉竹样品, 106~134 为黄花白及样品。1~29 号样品来源于饮片公司 (具体信息见表 1), 剩余样品均来源于成都荷花池中药材专业市场, 且从药材市场

购买的样品需自行进行净制和切制, 经河南中医药大学第一附属医院施钧瀚副主任药师按照《中国药典》2020 年版鉴定, 所有样品均为正品, 其中白及基原为兰科白及属植物白及 *B. striata* (Thunb.) Reichb. f. 的干燥块茎, 天麻基原为兰科天麻属植物天麻 *G. elata* Bl. 的干燥块茎, 玉竹基原为百合科黄精属植物玉竹 *P. odoratum* (Mill.) Druce 的干燥根茎, 黄花白及基原为兰科白及属植物黄花白及 *B. ochracea* Schltr. 的干燥块茎。白及对照药材, 批号 121261-201706, 中国食品药品检定研究院; 对照品 1,4-二[4-(葡萄糖氧)苄基]-2-异丁基苹果酸酯 (militarine, 质量分数 $\geq 98\%$), 批号 K18O9B72711, 上海源叶生物科技有限公司; 对照品天麻素, 批号

表 1 样品信息

Table 1 Information of samples

编号	品名	批号	产地	采集地点/生产厂家	编号	品名	批号	产地	采集地点/生产厂家
S1	白及	180311	贵州	安徽人民中药饮片有限公司	S16	白及	180903	贵州	安徽人民中药饮片有限公司
S2	白及	181001	贵州	安徽人民中药饮片有限公司	S17	白及	190213	贵州	安徽人民中药饮片有限公司
S3	白及	181202	贵州	安徽人民中药饮片有限公司	S18	白及	181125	贵州	安徽人民中药饮片有限公司
S4	白及	20181201	贵州凯里	安徽德昌药业股份有限公司	S19	白及	181216	贵州	安徽人民中药饮片有限公司
S5	白及	170801	贵州	亳州市先锋中药饮片有限公司	S20	白及	190113	贵州	安徽人民中药饮片有限公司
S6	白及	1901191	云南	安徽普仁中药饮片有限公司	S21	白及	180520	贵州	安徽人民中药饮片有限公司
S7	白及	1808061	贵州	安徽普仁中药饮片有限公司	S22	白及	190120	贵州	安徽人民中药饮片有限公司
S8	白及	1812131	云南	安徽普仁中药饮片有限公司	S23	白及	181110	贵州	安徽人民中药饮片有限公司
S9	白及	1812230102	湖北	亳州市沪淮药业有限公司	S24	白及	181005	贵州	安徽人民中药饮片有限公司
S10	白及	181219	贵州	安徽人民中药饮片有限公司	S25	白及	190214	贵州	安徽人民中药饮片有限公司
S11	白及	18051001	云南普洱	亳州市永刚饮片厂有限公司	S26	白及	180707	贵州	安徽人民中药饮片有限公司
S12	白及	18072501	云南普洱	亳州市永刚饮片厂有限公司	S27	白及	190101	贵州	安徽人民中药饮片有限公司
S13	白及	20181001	贵州凯里	安徽德昌药业股份有限公司	S28	白及	18030102	贵州	郑州瑞龙制药股份有限公司
S14	白及	D1812074	云南	四川新荷花中药饮片股份有限公司	S29	白及	18030101	贵州	郑州瑞龙制药股份有限公司
S15	白及	180715	贵州	安徽人民中药饮片有限公司					

110807-201809, 质量分数 96.7%, 中国食品药品检定研究院; 稀甘油, 批号 20180708, 南昌白云药业有限公司; 超纯水; 甲醇、乙腈、磷酸均为色谱纯。

2 方法与结果

2.1 标杆辨识信息 (Y) 的获取

2.1.1 《中国药典》及地方标准鉴别 参照《中国药典》2020 年版一部白及项下白及的外观性状描述及检查方法依次对 134 批样品进行外观性状鉴别、显微鉴别、薄层鉴别; 参照《四川省中药饮片炮制规范》2015 年版^[12]所刊载黄花白及、《中药炮制规范》2009 年版^[13]和《甘肃省中药材标准》2009 年版^[14]中小白及 (实为黄花白及) 项下的性状鉴别方

法。其中性状鉴别内容包括饮片的形状、外表皮颜色、切面颜色、质感、气和味等的描述。结果表明, 除 1~45 号样品及 106~120 号样品外, 其余样品外观性状均不符合《中国药典》2020 年版白及项下标准, 其中 121~134 号样品相较于白及标准其木化严重。1~45 号样品不符合地方标准中规定的黄花白及标准, 而 106~134 号符合地方标准中规定的黄花白及标准, 部分黄花白及样品不易与白及区分; 显微和薄层结果显示, 1~45 号样品均符合药典白及项下描述, 106~134 号样品与其相似, 而其余样品与其差异明显; 46~75 号样品、76~105 号样品经河南中医药大学第一附属医院施钧瀚副主任药师鉴

定均为正品天麻和玉竹。

2.1.2 HPLC 指纹图谱鉴别

(1) 对照品溶液的制备：分别精密称取 *militarine* 对照品 9.06 mg、天麻素对照品 10.66 mg，置于 10 mL 棕色量瓶中，加乙腈-水 (3:97) 溶解并定容，摇匀，即得对照品溶液。

(2) 供试品溶液的制备：取样品适量，粉碎并过 4 号筛，精密称定 2.0 g，置于 50 mL 具塞瓶中，加稀乙醇 40 mL，称定质量，超声 (功率 200 W、频率 40 kHz) 提取 30 min，取出，放冷，用稀乙醇补足减失的质量，滤过，滤液浓缩至无醇味，残渣加乙腈-水 (3:97) 混合溶液溶解，转移至 25 mL 量瓶中，用乙腈-水 (3:97) 混合溶液稀释至刻度，摇匀，离心，经 0.45 μ m 微孔滤膜滤过，取滤液，即得供试品溶液。

(3) 色谱条件：Shim-Pack GIST C₁₈-AQ 色谱柱 (250 mm×4.6 mm, 5 μ m)；流动相为 0.1% 磷酸水溶液-乙腈，梯度洗脱：0~5 min, 95%~80% 乙腈；5~10 min, 80%~76% 乙腈；10~20 min, 76%~68.5% 乙腈；20~25 min, 68.5%~65% 乙腈；25~30 min, 65%~58% 乙腈；30~45 min, 58%~40% 乙腈；体积流量 1.0 mL/min；检测波长 280 nm；柱温 30 $^{\circ}$ C；进样量 10 μ L。

(4) 精密度实验：按照上述方法处理 14 号供试品，吸取 10 μ L，按上述色谱条件连续进样 6 次，以 2 号峰 *militarine* 为参照峰，计算各共有峰的平均相对保留时间和平均相对峰面积 RSD 值。精密度实验结果显示 8 个共有峰的平均相对保留时间和平均相对峰面积 RSD 分别为 0.04%、0.13%，说明仪器

精密度良好。

(5) 重复性实验：精密称定 14 号样品 6 份，按照上述方法处理 14 号样品，制备供试品溶液，吸取 10 μ L，按上述色谱条件进样，以 2 号峰 *militarine* 为参照峰计算各共有峰的平均相对保留时间和平均相对峰面积 RSD 值。重复性实验结果显示，8 个共有峰的平均相对保留时间和平均相对峰面积 RSD 分别为 0.05%、0.27%，说明该方法的重复性良好。

(6) 稳定性实验：取上述制备好的 2 号供试品溶液，吸取 10 μ L，并按上述色谱条件分别在 0、2、4、8、12、24 h 进样，以 2 号峰 *militarine* 为参照峰计算各共有峰的平均相对保留时间和平均相对峰面积的 RSD 值。结果显示，8 个共有峰的平均相对保留时间和平均相对峰面积的 RSD 分别为 0.07%、0.33%，说明该方法制备的供试品溶液在恒温恒湿的室内条件下能够在 24 h 内保持稳定。

(7) 相似度分析：部分代表性样品的指纹图谱见图 1、2。以 45 批白及和白及对照药材为代表性样品获取白及共有模式。采用共有峰加权方法分别对 134 批样品的色谱峰与白及共有模式进行相似度分析。考虑到 *militarine* 为白及主要有效成分且白及与黄花白及不易区分等情况，对共有成分进行赋权，经过多次尝试最终赋予 *militarine* 权重为 2.5，天麻素权重为 1，其他组分权重为 1。45 批白及的指纹图谱相似度为 0.88~0.98 (均值为 0.931)，30 批天麻指纹图谱相似度为 0.20~0.34 (均值为 0.263, $P<0.01$)，30 批玉竹的指纹图谱相似度为 0.10~0.20 (均值为 0.134, $P<0.01$)，黄花白及的指纹图谱相似度为 0.70~0.88 (均值为 0.822, $P<0.05$)。

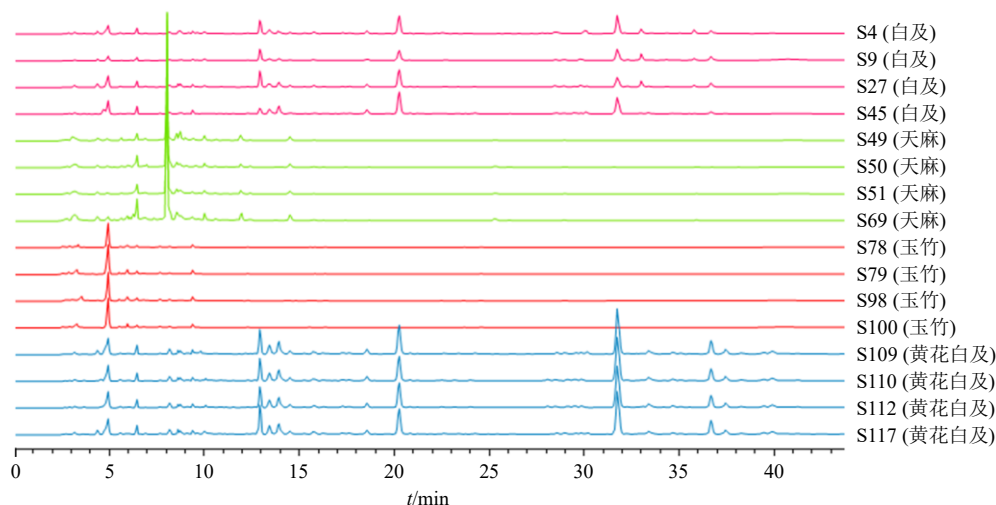


图 1 部分代表性样品指纹图谱

Fig. 1 Fingerprint of some representative samples

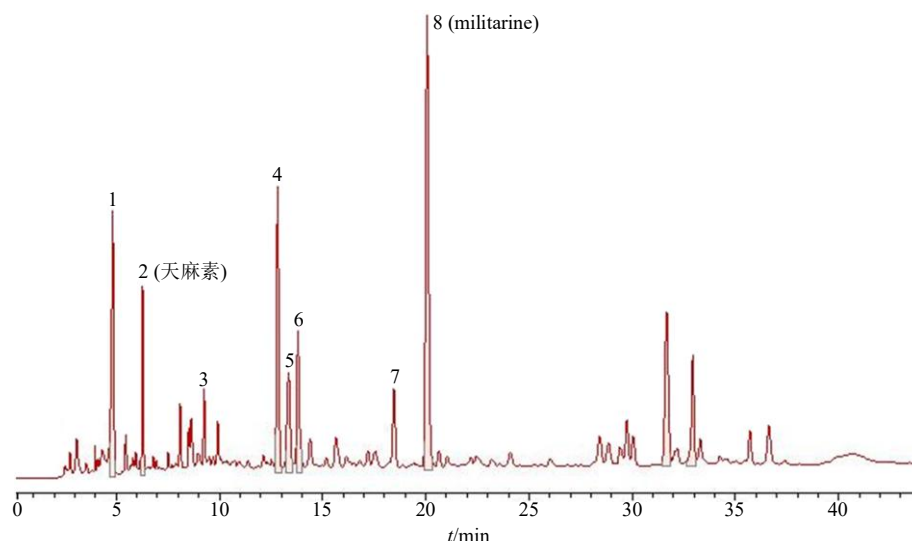


图 2 白及样品共有峰图谱

Fig. 2 Common peak spectrum of *B. striata* samples

2.1.3 综合鉴别结果 综上, 药典、地方标准、HPLC 指纹图谱法的鉴别结果均与原始采购信息一致, 即所购样品 1~45 为白及、46~75 为天麻、76~105 为玉竹、106~134 为黄花白及。

2.2 智能视觉信息(X)的获取

2.2.1 基于电子眼获取 X

(1) 样品制备: 模拟人工的颜色鉴别方法, 不破坏样品的完整性, 随机抽取样品适量放置在 A4 纸上, 待测。

(2) 测试程序: 电子眼开机后稳定几分钟, 待光照稳定后(仪器指示绿灯后)开始测试。首先筛选照明条件, 根据预实验结果选择底部照明, 使用 24 色色彩校正板进行校正, 使用 5 mm 光圈同时打开上下背光灯, 电子眼经过校正后对样品进行逐个采集图像, 每个样品变换位置采集 3 次图像。

(3) 电子眼输出信息值: 该研究中 134 批样品的电子眼颜色光谱数据为 85 种色号, 经筛选后 85 种色号均为有效信息, 均可参与建模, 数据矩阵为 134×85 。

2.2.2 基于色差仪获取 X

(1) 样品制备: 称取 5 g 样品于比色皿中, 使样品可以均匀平铺于整个比色皿。平行测定 3 次。

(2) 测试方法: 色差仪模式类型设定为 RSIN-包括镜面反射, 观察孔面积为 0.711 cm^2 , 分别使用 Hunter Lab Color Quest XE Light Trap、Instrument Standard 和 Diagnostic Tile 进行校准, 分别测定 134 批样品的红绿色差(a^*)、黄蓝色差(b^*)和光质明亮度(L^*)值。

(3) 色差仪输出信息值: 基于色差仪分别得到 134 批样品的 a^* 、 b^* 和 L^* 共 3 种变量, 均可参与建模, 数据矩阵为 134×3 。

2.3 辨识模型的建立及优化

以电子眼和色差仪采集的 134 批样品的视觉信息作为自变量 X, 基于药典、地方标准、HPLC 指纹图谱法的综合鉴别结果作为因变量 Y, 利用 Matlab 矩阵实验室建立 $Y=F(X)$ 模型, 即分别建立 2 类智能视觉数据融合前后的白及真伪二分类(白及、非白及)和四分类(白及、天麻、玉竹、黄花白及)辨识模型, 辨识结果以留一法交互验证的结果为准。

2.3.1 基于电子眼的白及真伪二分类辨识模型结果

(1) PCA-DA 辨识模型及交互验证结果: 如表 2 所示, 当主成分个数为 9 时, 解释变异信息之和达 90%以上, 可解释原始变量大部分信息。判别结果见表 3, 经留一法交互验证所得 PCA-DA 判别模型的正判率为 98.51%, 2 类样本在二维空间中部分重叠, 该模型区分度较高, 且不存在未分类样品。其中, 85 个变量分布较为分散, 均携带有显著的变异信息, 其中对潜变量 1 影响较大的是 34 和 67 号变量, 对潜变量 2 影响最大的是 28 和 83 号变量, 见图 3。

(2) PLS-DA 辨识模型及交互验证结果: 如表 4 所示, 潜变量个数为 4 时模型性能达到最优。判别结果见表 5, 经留一法交互验证所得 PLS-DA 判别模型的正判率为 97.01%, 2 类样本在二维空间中部分重叠, 该模型区分度较高, 且不存在未分类样

表 2 PCA-DA 主成分优选结果及变异信息

Table 2 Selection results and variation information of PCA-DA main component

主成分 数 (N)	误差率 (CV)/%	解释变量 (X)/%	主成分 数 (N)	误差率 (CV)/%	解释变量 (X)/%
1	0.50	32	11	0.02	94
2	0.23	56	12	0.02	95
3	0.27	72	13	0.03	96
4	0.04	80	14	0.02	96
5	0.03	84	15	0.02	97
6	0.02	88	16	0.02	97
7	0.03	90	17	0.02	97
8	0.03	91	18	0.02	98
9	0.02	93	19	0.03	98
10	0.02	94	20	0.02	98

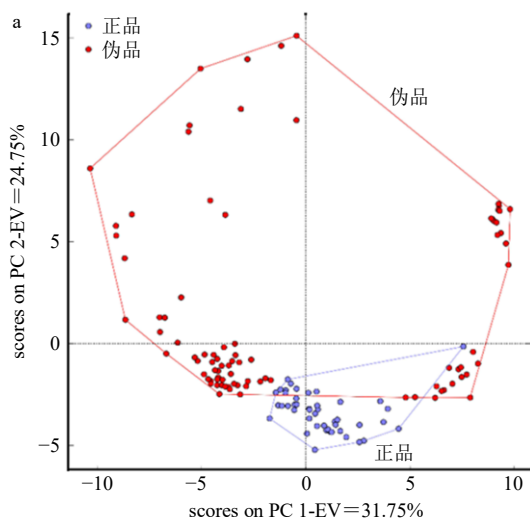


表 3 PCA-DA 留一法交互验证模型判别结果

Table 3 Discriminant results of PCA-DA leave-one-out interactive verification model

预测值	实际值		未分类	正判率/%
	白及	非白及		
白及	44	1	0	98.51
非白及	1	88	0	

品。其中对潜变量 1 影响较大的是 50、51 和 72 号变量,对潜变量 2 影响较大的是 26、33、35 号变量,见图 4。

(3) LS-SVM 辨识模型及交互验证结果:经留一法交叉验证对 LS-SVM 的核函数类型等参数进行优化。在本次真伪定性判别中,对线性核函数、多项式核函数 (polynomial kernel function)、多层感知

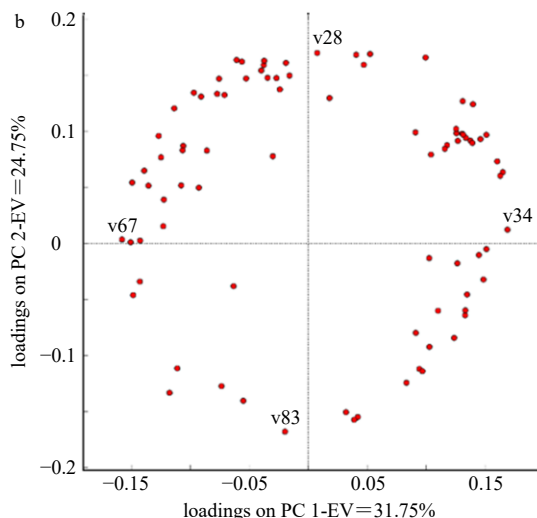


图 3 PCA-DA 得分 (a) 及载荷图 (b)

Fig. 3 PCA-DA core (a) and load graph (b)

表 4 PLS-DA 主成分优选结果及变异信息

Table 4 Selection results and variation information of PLS-DA main component

N	CV/%	X/%	N	CV/%	X/%	N	CV/%	X/%
1	0.09	24	8	0.04	90	15	0.05	95
2	0.03	36	9	0.04	91	16	0.07	96
3	0.05	61	10	0.04	92	17	0.06	96
4	0.03	72	11	0.03	93	18	0.04	96
5	0.04	83	12	0.04	93	19	0.06	97
6	0.03	86	13	0.04	94	20	0.06	97
7	0.03	88	14	0.04	95			

表 5 PLS-DA 留一法交互验证模型判别结果

Table 5 Discriminant results of PLS-DA leave-one-out interactive verification model

预测值	实际值		未分类	正判率/%
	白及	非白及		
白及	44	1	0	97.01
非白及	3	86	0	

器核函数 (multi layer perceptron kernel function)、径向基核函数 [(radial basis function (RBF) kernel function)] 等核函数类型进行筛选,最终选择判别结果最优的径向基核函数 (核函数参数 1 值,核函数参数 2 值);对单纯形 (simplex)、网格搜索 (gridsearch) 等优化函数 (optimization function) 进行筛选,最终选择单纯形优化函数。判别结果见表 6,经留一法交互验证所得 LS-SVM 判别模型的正

器核函数 (multi layer perceptron kernel function)、径向基核函数 [(radial basis function (RBF) kernel

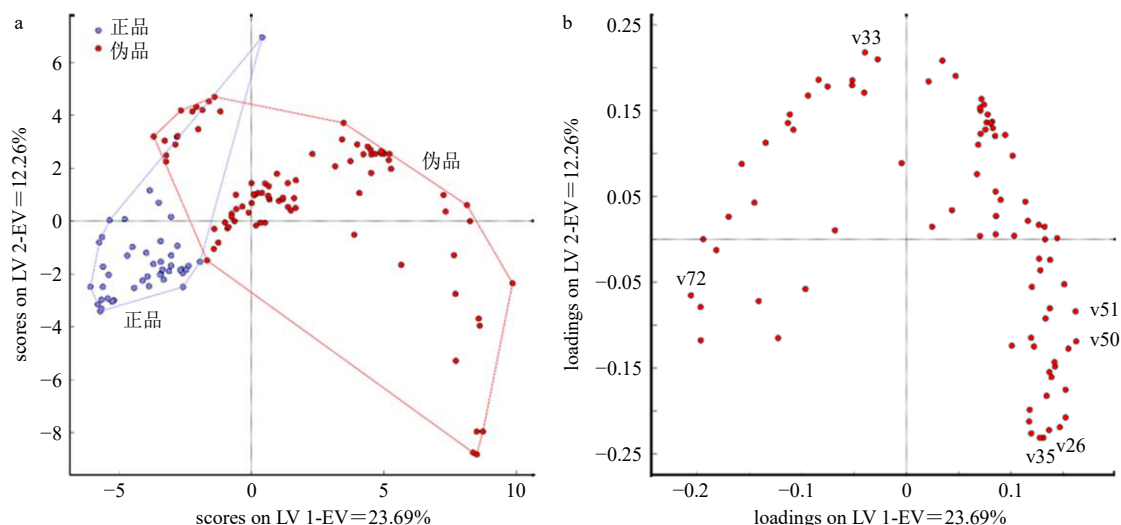


图 4 PLS-DA 得分 (a) 及载荷图 (b)

Fig. 4 PLS-DA core (a) and load graph (b)

表 6 LS-SVM 留一法交互验证模型判别结果

Table 6 Discriminant results of LS-SVM leave-one-out interactive verification model

预测值	实际值		未分类	正判率/%
	白及	非白及		
白及	31	14	0	89.55
非白及	0	89	0	

判率为 89.55%。

(4) KNN 辨识模型及交互验证结果：如表 7 所示，当 K 值为 3 时，模型错误率最低。判别结果见表 8，经留一法交互验证所得 KNN 判别模型的正判率为 99.25%。

(5) 各模型正判率结果对比：基于电子眼建立

表 7 K 参数优选结果

Table 7 Optimization results of K parameter

K 值	CV/%	K 值	CV/%	K 值	CV/%
1	0.03	5	0.02	9	0.02
2	0.03	6	0.02	10	0.02
3	0.01	7	0.02		
4	0.03	8	0.02		

表 8 KNN 留一法交互验证模型判别结果

Table 8 Discriminant results of KNN leave-one-out interactive verification model

预测值	实际值		未分类	正判率/%
	白及	非白及		
白及	44	1	0	99.25
非白及	0	89	0	

的白及真伪二分类辨识模型均未出现样本未分类情况，模型正判率结果比较为 $KNN > PCA-DA > PLS-DA > LS-SVM$ ，以 KNN 模型为最优。

2.3.2 基于色差仪的白及真伪二分类辨识模型结果

(1) PCA-DA 辨识模型及交互验证结果：如表 9 所示，当主成分个数为 2 时，解释变异信息之和达 90% 以上，可解释原始变量大部分信息。判别结果见表 10，经留一法交互验证所得 PCA-DA 判别模型的正判率为 67.16%，2 类样本在二维空间中部分重叠，该模型区分度较低，但不存在未分类样品。其中，3 个变量分布较为分散，均携带有显著的变异信息，其中对潜变量 1 影响较大的是 2 号变量，对潜变量 2 影响最大的是 3 号变量，见图 5。

表 9 PCA-DA 主成分优选结果及变异信息

Table 9 Selection results and variation information of PCA-DA main component

主成分数	CV/%	X /%
1	0.56	67
2	0.42	97
3	0.42	100

表 10 PCA-DA 留一法交互验证模型判别结果

Table 10 Discriminant results of PCA-DA leave-one-out interactive verification model

预测值	实际值		未分类	正判率/%
	白及	非白及		
白及	13	32	0	67.16
非白及	12	77	0	

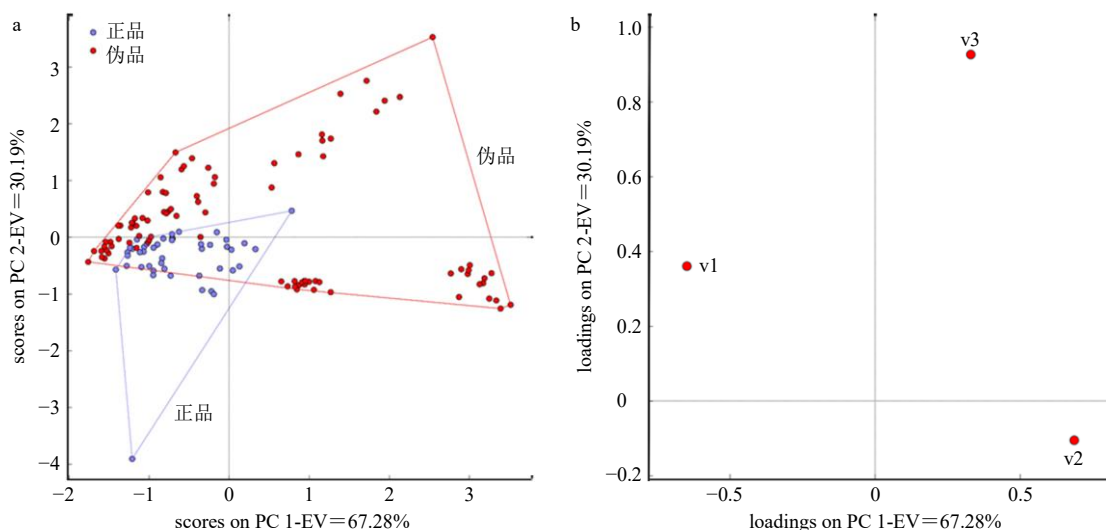


图 5 PCA-DA 得分 (a) 及载荷图 (b)

Fig. 5 PCA-DA core (a) and load graph (b)

(2) PLS-DA 辨识模型及交互验证结果: 如表 11 所示, 潜变量个数为 3 时模型性能达到最优, 所选 2 个潜变量可解释 90% 以上的自变量差异信息和 10% 左右的因变量差异信息。判别结果见表 12, 经留一法交互验证所得 PLS-DA 判别模型的正判率为

表 11 PLS-DA 主成分优选结果及变异信息

Table 11 Selection results and variation information of PLS-DA main component

主成分数	CV/%	X/%
1	0.25	62
2	0.23	95
3	0.12	100

表 12 PLS-DA 留一法交互验证模型判别结果

Table 12 Discriminant results of PLS-DA leave-one-out interactive verification model

预测值	实际值		未分类	正判率/%
	白及	非白及		
白及	41	4	0	87.31
非白及	13	76	0	

87.31%, 2 类样本在二维空间中部分重叠, 该模型区分度较低, 但不存在未分类样品, 且比 PCA-DA 判别模型结果较好。其中对潜变量 1 影响较大的是 2 号变量, 对潜变量 2 影响较大的是 3 号变量, 见图 6。

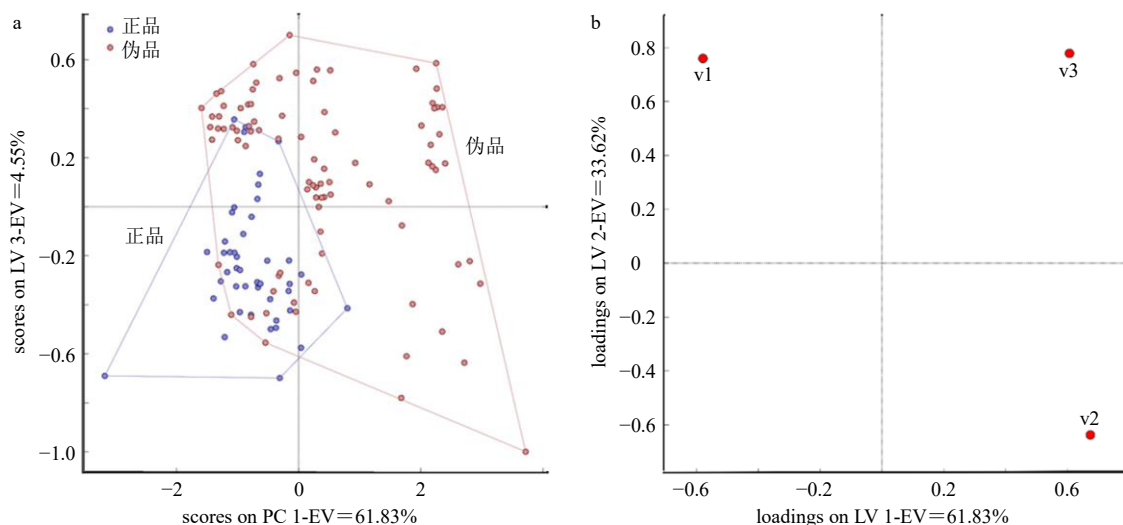


图 6 PLS-DA 得分 (a) 及载荷图 (b)

Fig. 6 PLS-DA core (a) and load graph (b)

(3) LS-SVM 辨识模型及交互验证结果: 同“2.3.1 (3)”, 判别结果如表 13 所示, 经留一法交互验证所得 LS-SVM 判别模型的正判率为 98.51%。

(4) KNN 辨识模型及交互验证结果: 如表 14 所示, 当 K 值为 3 时, 模型误差率达到最低。判别结果如表 15, 经留一法交互验证所得 KNN 判别模型的正判率为 99.25%。

表 13 LS-SVM 留一法交互验证模型判别结果

Table 13 Discriminant results of LS-SVM leave-one-out interactive verification model

预测值	实际值		未分类	正判率/%
	白及	非白及		
白及	43	2	0	98.51
非白及	0	89	0	

表 14 K 参数优选结果

Table 14 Optimization results of K parameter

K 值	CV/%	K 值	CV/%	K 值	CV/%
1	0.016 7	5	0.011 4	9	0.016 9
2	0.016 7	6	0.011 4	10	0.016 9
3	0.005 5	7	0.011 4		
4	0.005 5	8	0.011 4		

表 15 KNN 留一法交互验证模型判别结果

Table 15 Discriminant results of KNN leave-one-out interactive verification model

预测值	实际值		未分类	正判率/%
	白及	非白及		
白及	45	0	0	99.25
非白及	1	88	0	

(5) 各模型正判率结果对比: 基于色差仪建立的白及真伪二分类辨识模型均未出现样本未分类情况, 模型正判率结果比较为 $KNN > LS-SVM > PLS-DA > PCA-DA$, 以 KNN 模型为最优。

2.3.3 基于 2 类智能视觉技术的白及真伪四分类辨识模型结果 经留一法交互验证后, 基于电子眼建立的白及真伪四分类辨识模型的正判率结果比较为 $LS-SVM > PLS-DA > KNN = PCA-DA$, 以 LS-SVM 模型为最优; 基于色差仪建立的白及真伪四分类辨识模型的正判率结果比较为 $PLS-DA > LS-SVM > KNN > PCA-DA$, 以 PLS-DA 模型为最优。正判率如表 16 所示。其中, 建立的白及四分类 PLS-DA 模型均存在未分类情况。

表 16 基于 2 类智能视觉技术的白及真伪四分类辨识模型正判率结果

Table 16 Positive rate results of *B. striata* authenticity four-classification identification models of two types of intelligent vision technologies

视觉技术类型	正判率/%			
	PCA-DA	PLS-DA	LS-SVM	KNN
电子眼	96.27	96.67	97.01	96.27
色差仪	76.12	97.67	90.23	89.55

2.3.4 基于 2 类智能视觉与多源信息融合技术的白及真伪辨识模型结果 结果如表 17 所示, 将 2 类智能视觉技术所得的标杆信息 X 进行数据级融合后, 所得白及真伪二分类辨识模型的正判率结果比较为 $PCA-DA = KNN > PLS-DA > LS-SVM$, 以 PCA-DA 和 KNN 模型为最优, 最优模型正判率相较于融合前有所降低; 建立的白及真伪四分类辨识模型的正判率结果比较为 $LS-SVM > PLS-DA > PCA-DA > KNN$, 以 LS-SVM 模型为最优, 最优模型正判率相较于融合前有所升高。相较于融合前所建白及四分类 PLS-DA 模型, 未分类样本明显减少。

表 17 基于 2 类智能视觉与多源信息融合技术的白及真伪辨识模型正判率结果

Table 17 Positive rate results of *B. striata* authenticity identification model based on two types of intelligent vision and multi-source information fusion technology

模型类别	正判率/%			
	PCA-DA	PLS-DA	LS-SVM	KNN
二分类	98.51	97.01	87.31	98.51
四分类	97.01	97.56	97.76	96.27

3 讨论

3.1 标杆辨识信息的获取方法分析

《中国药典》是国家药品标准的核心组成部分, 也是药品研制、生产、经营、使用和监督管理等必须遵循的法典。药典检测方法通过定性和定量方法对中药质量信息进行表征, 能够很好的反映所购中药的品质, 其最大特点是相对准确、重复性较好。该研究性状鉴别结果显示部分黄花白及与白及不易区分, 为使鉴定结果更加准确和具有说服力, 增加了 HPLC 指纹图谱鉴别法。另外, 在指纹图谱建立过程中, 本研究比较了白及液相色谱在 220、230、270、280 nm 波长下的紫外吸收情况, 发现其在 280 nm 波长下的紫外吸收情况更好。同时考察了不同

流动相、体积流量、柱温及进样量对色谱峰的影响并选取最优条件用于指纹图谱的建立。

3.2 2 种智能视觉技术特点分析

在该研究建立的模型中,除二分类 LS-SVM 和四分类 PLS-DA 模型外,基于电子眼建立的白及真伪二分类和四分类模型具有更强的预测能力,其正判率均大于或等于色差仪建立的同类模型的正判率。其中,该研究所用手持式色差仪易于携带和便于使用,可以在生产现场进行使用以节约大量来往于实验室与生产现场的时间,但精度相对较低,可应用高光谱成像技术提高其预测精度。其含有 a^* 、 b^* 和 L^* 共 3 个变量,分别代表被测物的红绿色差、黄蓝色差和光质明亮度值;而该研究所用电子眼在控制照明和成像条件下使用高分辨率成像技术,对颜色差异有高度的敏感性。该研究通过电子眼获取了白及其伪品的 85 种色号信息,其变量数目相较于色差仪更多,所测颜色信息更精细化,推测其能涵盖和表达的质量信息也更多。

3.3 人工智能感官技术在中药鉴别领域的应用分析

“辨状论质”是历代医药学家长期用药经验的智慧结晶,即通过观察药材的外观性状,如形状、颜色、味道、质地等,对观察的结果进行分析与总结,来判断药材的内在品质^[15],在中药品质评价中始终承担着重要作用^[16]。但传统性状鉴别描述比较模糊,缺乏客观量化的标准,在实际应用中常受到局限^[17]。随着人工智能感官 (artificial intelligence sense, AIS) 技术^[18]的发展,研究者开始不断尝试采用电子眼、电子鼻、电子舌等^[19]仿生技术对中药的性状特征进行客观化表达。因其整合了仿生感官和现代仪器分析两类技术,故兼有“感官”响应的“快速”和仪器分析“量化”的双重优势,具有分析快速、灵敏度高、可靠性强、重复性好、整体性强等特点,可使感官评价定量化和标准化^[20]。在弥补中药传统性状鉴别不足的同时, AIS 还能改善现代仪器鉴别方法比较耗时等问题。其在中药鉴别方面的应用已初步显现^[21-23],且拥有良好的应用前景和发展潜力。另外,多源信息融合技术^[24]可通过对获取的不同感官数据进行融合,即可对中药气味、滋味、色泽等方面的信息进行全面衡量^[25],以解决单一感官可能存在鉴别结果差、模型传递性差等问题,最终实现 AIS 技术对中药的准确快速鉴别。

4 结论

基于电子眼和色差仪建立的白及真伪快速辨识

模型均具有良好的预测能力,但在模型相同的情况下,相较于色差仪,基于电子眼所建模型性能整体上更优。2 类智能视觉数据融合后,二分类最优辨识模型正判率有所降低,四分类最优辨识模型正判率有所升高。总体而言,该技术能够实现对白及的快速准确辨识,为中药品质的快速准确辨识奠定了基础。

利益冲突 所有作者均声明不存在利益冲突

参考文献

- [1] 孙乐乐,杨永红,刘军凯,等.白及的本草考证[J].中药材,2010,33(12):1965-1968.
- [2] 钱丽华,尹舒雅,陆娜,等.白芨资源综合利用现状及展望[J].江苏农业科学,2021,49(19):64-71.
- [3] 曹琦,王学平.药用白芨的生物学特性及其保护[J].安徽农业科学,2015,43(18):175-176.
- [4] 张曼,韩亭亭,胡春芳,等.白及产业现状及可持续发展策略[J].中草药,2019,50(20):5103-5108.
- [5] 李媛媛,王艳丽,姚静,等.基于电子舌的白及及其近似饮片的快速辨识研究[J].世界科学技术—中医药现代化,2021,23(5):1532-1539.
- [6] 李明华,程显隆,李宁新,等.白及的质量问题及真伪鉴别方法[J].中国药事,2018,32(11):1490-1499.
- [7] 陈美君,李峰庆,吕蒙,等.白及与其混伪品 ITS2 序列二级结构比较与鉴别[J].中国实验方剂学杂志,2017,23(15):46-52.
- [8] 谢树莹,周金海.机器视觉在中药领域的应用现状及展望[A]//第四届中国中医药信息大会论文集[C].北京:中国中医药信息学会,2017:333-336.
- [9] 黎江华,吴纯洁,孙灵根,等.基于机器视觉技术实现中药性状“形色”客观化表达的展望[J].中成药,2011,33(10):1781-1784.
- [10] 刘瑞新,郝小佳,张慧杰,等.基于电子眼技术的中药川贝母真伪及规格的快速辨识研究[J].中国中药杂志,2020,45(14):3441-3451.
- [11] 甄臻,李慧芬,刘静,等.基于粉末和显微特征颜色数字化的生地黄与熟地黄判别[J].中草药,2021,52(24):7438-7446.
- [12] 四川省食品药品监督管理局.四川省中药饮片炮制规范(2015年版)[M].成都:四川科技出版社,2015:5.
- [13] 甘肃省卫生局.中药炮制规范[M].兰州:甘肃人民出版社,2009:12-13.
- [14] 甘肃省食品药品监督管理局.甘肃省中药材标准:2009版[M].兰州:甘肃文化出版社,2009:13-14.
- [15] 荆文光,程显隆,刘安,等.基于“辨状论质”综合评价指数的厚朴饮片等级划分和优质优效研究[J].中草药,2021,52(8):2285-2293.
- [16] 钟可,周汉华,丁宁,等.《中药鉴定学》性状鉴别经

- 验术语的传承与创新 [J]. 中国中医基础医学杂志, 2016, 22(8): 1105-1106.
- [17] 李捷. 基于化学轮廓分析和“辨状论质”对肉苁蓉产地加工与炮制工艺的研究 [D]. 天津: 天津中医药大学, 2021.
- [18] Galvan D, Aquino A, Efting L, *et al.* E-sensing and nanoscale-sensing devices associated with data processing algorithms applied to food quality control: A systematic review [J]. *Crit Rev Food Sci Nutr*, 2022, 62(24): 6605-6645.
- [19] 刘瑞新, 陈鹏举, 李学林, 等. 人工智能感官: 药学领域的新技术 [J]. 药物分析杂志, 2017, 37(4): 559-567.
- [20] Xie D S, Peng W, Chen J C, *et al.* A novel method for the discrimination of Hawthorn and its processed products using an intelligent sensory system and artificial neural networks [J]. *Food Sci Biotechnol*, 2016, 25(6): 1545-1550.
- [21] 秦宇雯, 费程浩, 苏联麟, 等. 基于 Heracles NEO 超快速气相色谱电子鼻对温郁金醋制前后气味差异标志物的快速识别研究 [J]. 中草药, 2022, 53(5): 1313-1319.
- [22] 郑玉, 谢兴亮, 刘俊, 等. 基于 pH 调控多酚-唾液蛋白结合作用的三果汤含片涩味掩蔽方法研究 [J]. 中草药, 2022, 53(13): 3953-3961.
- [23] 杨诗龙. 基于智能感官分析技术的贝母及黄连饮片鉴别研究 [D]. 成都: 成都中医药大学, 2015.
- [24] 朱建锡, 马晓钟, 张津阳, 等. 信息融合及其与电子鼻和电子舌相关的应用 [J]. 现代食品, 2020(11): 118-122.
- [25] 田晓静. 基于电子鼻和电子舌的羊肉品质检测 [D]. 杭州: 浙江大学, 2014.

[责任编辑 郑礼胜]