基于机器视觉系统的姜炭炮制程度判别及颜色-成分相关性分析Δ

张一凡[1*]，周苏娟[2]，孟　江[1#]，左　蓉[1]，林华坚[1]，孙　悦[1]，王淑美[1]（1.广东药科大学中药学院/国家中医药管理局中药数字化质量评价技术重点研究室/广东高校（省）中药质量工程技术研究中心，广州　510006；2.广东药科大学信息工程学院，广州　510006）

中图分类号　R917；R283　　**文献标志码**　A　　**文章编号**　1001-0408（2022）22-2712-07
DOI　10.6039/j.issn.1001-0408.2022.22.05

摘　要　目的　基于机器视觉系统探讨姜炭炮制程度判别及颜色-成分含量相关性，为姜炭炮制程度控制和质量评价提供参考。**方法**　采用高效液相色谱法测定干姜及其不同炮制程度姜炭饮片中6-姜酚、8-姜酚、10-姜酚、6-姜烯酚、姜酮5种成分的含量；采用机器视觉技术获得饮片图像并提取饮片RGB、$L^*a^*b^*$、HSV 3个不同颜色空间的颜色特征，采用机器学习，如主成分分析（PCA）、线性判别分析（LDA）、偏最小二乘法-判别分析（PLS-DA）和支持向量机（SVM）等方法对姜炭不同炮制程度姜炭建立定性判别模型，将颜色特征值与测得的5种成分含量进行相关性分析，建立颜色-成分的定量预测模型。**结果**　随着炮制程度的加深，姜酮在炮制后产生且含量先增加后降低，标炭饮片中含量最高；6-姜酚、8-姜酚和10-姜酚的含量逐渐降低；6-姜烯酚含量先增加后降低。基于饮片图像颜色客观量化结合有监督判别模式识别方法的LDA和SVM建立的定性判别模型，其交叉验证训练预测准确率达到100%，外部验证准确率达到95.83%。基于饮片图像颜色客观量化结合SVM建立5种成分含量预测模型，RPD值均大于2，各成分的R^2_P与R^2_C值中姜酮为0.633 9与0.683 3，其他成分值均大于0.75，说明SVM对除姜酮以外的4种成分均有较好的预测能力。**结论**　基于机器视觉系统对姜炭炮制程度的判别和含量预测模型效果良好，可为姜炭饮片的质量控制和炮制程度的判断提供参考。

关键词　姜炭；机器视觉；机器学习；质量评价；炮制程度

Discrimination of processing degree of *Zingiber officinale* charcoal and analysis of the correlation between color and component based on machine vision system

ZHANG Yifan[1], ZHOU Sujuan[2], MENG Jiang[1], ZUO Rong[1], LIN Huajian[1], SUN Yue[1], WANG Shumei[1]（1. School of Traditional Chinese Medicine, Guangdong Pharmaceutical University/Key Laboratory of Digital Quality Evaluation of Chinese Materia Medica, State Administration of Traditional Chinese Medicine/Engineering Technology Research Center for Chinese Materia Medica Quality of Guangdong Universities and Colleges (Province), Guangzhou 510006, China; 2. College of Medical Information Engineering, Guangdong Pharmaceutical University, Guangzhou 510006, China）

ABSTRACT　OBJECTIVE To explore the discrimination of processing degree of *Zingiber officinale* charcoal and the correlation between color and component content based on machine vision system, and provide reference for quality evaluation and processing degree control of *Z. officinale* charcoal. **METHODS** High-performance liquid chromatography method was used to determine the contents of 5 components in *Z. officinale* charcoal and its different processed products, such as 6-gingerol, 8-gingerol, 10-gingerol, 6-shogaol, gingerone. Machine vision system was used to obtain the image of the decoction pieces and extract the color features of the decoction pieces in RGB, $L^*a^*b^*$ and HSV color spaces. Machine learning methods, such as principal component analysis (PCA), linear discriminant analysis (LDA), partial least squares-discriminant analysis (PLS-DA) and support vector machine (SVM), were used to establish qualitative identification model for *Z. officinale* charcoal processed products of different processing degree. The correlation between the color eigenvalues and the contents of measured 5 components were analyzed, and the color-component content prediction model was established. **RESULTS** With the deepening of processing, gingerone was produced after processing and the content firstly increased and then decreased, and the content of gingerone in standard carbon was the highest; the contents of 6-gingerol, 8-gingerol

Δ**基金项目** 国家自然科学基金资助项目（No.81873006）

＊**第一作者** 硕士。研究方向：中药炮制原理与质量标准。E-mail：1416812771@qq.com

＃**通信作者** 教授，硕士生导师，博士。研究方向：中药炮制及饮片质量控制。电话：020-39352175。E-mail：jiangmeng666@126.com

and 10-gingerol decreased gradually; the content of 6-shogaol increased firstly and then decreased. The prediction accuracy of qualitative discriminant model, which was established on the basis of objective quantization of image and color combined with LDA and SVM of supervised discriminant pattern recognition method, reached 100% in cross-validation training and 95.83% in the external validation. Content prediction model of 5 components was established on the basis of objective quantization of image and color combined with SVM, the RPD values were all greater than 2, the R^2_P and R^2_C values of gingerone were 0.633 9 and 0.683 3, and the values of other components were all greater than 0.75, indicating SVM had good prediction ability for the contents of 4 components except for gingerone. **CONCLUSIONS** The machine vision system is excellent for the discrimination of the processing degree of *Z. officinale* charcoal and the content prediction, which can provide a reference for the quality control of *Z. officinale* charcoal decoction pieces and the judgment of the processing degree.

KEYWORDS *Zingiber officinale* charcoal; machine vision; machine learning; quality evaluation; processing degree

姜炭是干姜炒炭后的炮制品，味苦而涩，其性守而不走，长于温经止血，是临床治疗虚寒性出血症的经典温里止血药[1]。目前对姜炭的质量控制，2020年版《中国药典》的标准为“表面焦黑色，内部棕褐色，体轻，质松脆。味微苦，微辣。以含6-姜辣素（$C_{17}H_{26}O_4$）计不得少于0.050%，浸出物不得少于26.0%”[1]。颜色气味等指标存在主观性较强、易受到周围环境的影响和标准不一致等问题[2]。化学成分的含量测定，能在一定程度上反映饮片质量，但其测定耗时耗力，同时化学试剂会对环境造成一定的污染，另外饮片的颜色气味和药效成分的相关性不清楚。

机器视觉也称为计算机视觉，是一门涉及人工智能、神经生物学、计算机科学、图像处理及模式识别等多个领域的科学[3]。机器学习是一门人工智能的科学，研究计算机怎样模拟或实现人类的学习行为，以获取新的知识或技能，重新组织已有的知识结构使之不断改善自身的性能[4]。机器视觉融合机器学习（即机器视觉系统）可以把图像信息数字化，同时利用计算机模拟人的判别准则去理解和识别图像信息，达到分析图像和成分间关系并做出结论的目的，最终用于饮片实际鉴别和质量控制。目前机器视觉系统在中药饮片的定性定量分析中得到广泛应用[5—8]。

姜炭的主要化学成分为姜酚、姜烯酚、姜酮等姜辣素类成分，现代药理研究表明，姜辣素（6-姜酚、8-姜酚、10-姜酚）可以促进血液循环，抑制血栓和血小板聚集，而姜烯酚和姜酮具有良好的抗氧化和抗炎活性[9—12]。因此，本文选择6-姜酚、8-姜酚、10-姜酚、姜酮、6-姜烯酚为姜炭不同炮制程度饮片的指标性成分，采用高效液相色谱（HPLC）法测定这5个成分含量，再通过机器视觉系统对饮片颜色特征进行数字化信息处理，再利用主成分分析（principal component analysis，PCA）、线性判别分析（principal component analysis，LDA）、偏最小二乘法-判别分析（partial least squares-discriminant analysis，PLS-DA）和支持向量机（support vector machine，SVM）等机器学习算法建立基于饮片外观颜色的定性判别模型和颜色-成分的定量模型，对不同炮制程度姜炭进行定性、定量分析，为姜炭炮制程度控制和质量评价提供参考依据。具体研究思路见图1。

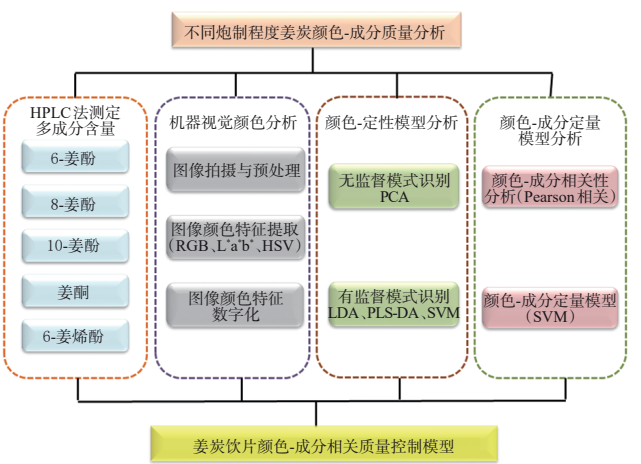


图1 不同炮制程度姜炭饮片颜色-成分质量分析的研究思路图

1 材料

1.1 主要仪器

本研究所用主要仪器有LC-20A型HPLC仪（日本Shimadzu公司）、JP-105A型高速多功能粉碎机（永康市久品工贸有限公司）、BSA124S型万分之一电子天平（德国Sartorius公司）、FUJI X-T2型数码相机（日本Fuji公司）等。

1.2 主要药品与试剂

本实验共收集干姜饮片20批，分别来自四川、云南、安徽等产地，并经广东药科大学中药学院刘基柱教授鉴定为姜科植物姜 *Zingiber officinale* Rosc. 的干燥根茎。课题组采用2020年版《中国药典》（四部）收载的炮制通则项下砂烫法对20批干姜进行炮制，以河砂温度为

310～330 ℃分别炮制7～10 min、13～17 min、18～22 min，得到轻炭饮片（外表焦黑、内黄色）、标炭饮片（外表焦黑、内棕褐色）、重炭饮片（内外均焦黑）[13]，样品来源信息见表1。6-姜酚（批号CHB180306）、8-姜酚（批号CHB180305）、10-姜酚（批号CHB180311）、6-姜烯酚（批号CHB180307）、姜酮（批号CHB19012）均购自成都克洛玛生物科技有限公司，纯度均大于98%。甲醇、乙腈为色谱纯，水为蒸馏水。

表1　干姜及其轻炭、标炭、重炭饮片的样品信息表

干姜	轻炭饮片	标炭饮片	重炭饮片	产地	生产批号	生产日期
SP01	QT01	BT01	ZT01	四川	190201	2019-02-25
SP02	QT02	BT02	ZT02	广东	170901	2017-09-06
SP03	QT03	BT03	ZT03	云南	190201	2019-02-17
SP04	QT04	BT04	ZT04	江苏	HX19K01	2019-04-18
SP05	QT05	BT05	ZT05	四川	181003651	2018-10-24
SP06	QT06	BT06	ZT06	四川	YPA7K0001	2017-11-24
SP07	QT07	BT07	ZT07	四川	YPA8B0001	2018-02-24
SP08	QT08	BT08	ZT08	四川	YPA8H0001	2018-08-07
SP09	QT09	BT09	ZT09	四川	YPA9B0001	2019-02-24
SP10	QT10	BT10	ZT10	四川	190400311	2019-04-02
SP11	QT11	BT11	ZT11	四川	190301	2019-03-15
SP12	QT12	BT12	ZT12	四川	201901116	2019-01-10
SP13	QT13	BT13	ZT13	四川	190400311	2019-04-02
SP14	QT14	BT14	ZT14	安徽	190624	2019-07-01
SP15	QT15	BT15	ZT15	四川	190514FN	2019-05-14
SP16	QT16	BT16	ZT16	四川	180901	2018-09-08
SP17	QT17	BT17	ZT17	四川	190628	2019-07-08
SP18	QT18	BT18	ZT18	四川	190401	2019-04-16
SP19	QT19	BT19	ZT19	四川	190618	2019-06-25
SP20	QT20	BT20	ZT20	云南	190501	2019-05-20

2　方法与结果

2.1　多成分含量测定

2.1.1　混合对照品溶液的制备　分别精密称取6-姜酚对照品5.04 mg、8-姜酚对照品2.40 mg、10-姜酚对照品4.50 mg、6-姜烯酚对照品4.50 mg、姜酮对照品4.80 mg，分别置于1 mL容量瓶中，用甲醇溶解并定容，混匀，即得各单一对照品储备液。分别精密吸取各单一对照品储备液6-姜酚0.5 mL、8-姜酚0.2 mL、10-姜酚0.4 mL、6-姜烯酚0.2 mL、姜酮0.1 mL，置于10 mL容量瓶中，加甲醇至刻度，即得混合对照品溶液。

2.1.2　供试品溶液的制备　取样品粉末（过80目筛）0.5 g，精密称定，置50 mL三角锥形瓶中，精密加入20 mL甲醇，盖好瓶塞，称定质量，超声（功率300 W、频率50 kHz）提取40 min，放冷至室温，用甲醇补足减失质量，摇匀，过0.45 μm微孔滤膜，取续滤液，封口，贴好标签，于4 ℃冰箱中保存备用。

2.1.3　色谱条件　以Ultimate™ XB-C_{18}（4.6 mm×250 mm，5 μm）为色谱柱；甲醇（A）-0.1%磷酸溶液（B）为流动相进行梯度洗脱（0～15 min，45%A→50%A；15～20 min，50%A→70%A；20～60 min，70%A→85%A；60～70 min，85%A→95%A；70～80 min，95%A）；检测波长为280 nm；流速为0.6 mL/min；进样量为10 μL；柱温为30 ℃。

2.1.4　方法学考察　（1）专属性考察：分别取混合对照品溶液、干姜及其不同炮制程度姜炭供试品溶液、空白溶液（甲醇），按“2.1.3”项下色谱条件进样分析，色谱图见图2（空白图略）。结果显示，对照品溶液与供试品溶液中5种成分色谱峰保留时间一致，空白溶液无干扰，表明方法专属性良好。

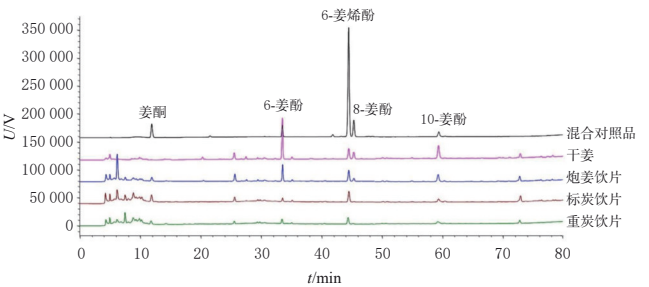


图2　混合对照品溶液、干姜及其不同炮制程度姜炭供试品溶液的HPLC图谱

（2）线性关系考察：精密吸取上述混合对照品溶液，照倍比稀释法，依次制备得到不同质量浓度的混合对照品溶液，按“2.1.3”项下色谱条件进样分析，以对照品质量浓度为横坐标（X）、其峰面积为纵坐标（Y），绘制标准曲线。结果见表2。

表2　6-姜酚等5种成分的线性关系考察结果

成分	回归方程	r	线性范围/（μg/mL）
6-姜酚	Y=10 826 X−13 230.0	0.999 2	0.63～252.00
8-姜酚	Y=10 104 X−2 976.6	0.999 8	0.12～48.00
10-姜酚	Y=8 046.3 X−13 380.0	0.999 4	0.45～180.00
6-姜烯酚	Y=10 333 X−6 118.1	0.999 6	0.23～90.00
姜酮	Y=13 672 X−3 047.2	0.999 7	0.12～48.00

（3）精密度考察：取混合对照品溶液，按“2.1.3”项下色谱条件连续进样6次。结果显示，6-姜酚等5种成分峰面积的RSD均小于2.00%（n=6），表明仪器精密度良好。

（4）重复性考察：取干姜（SP03）粉末6份，精密称定，按“2.1.2”项下方法制备供试品溶液，按“2.1.3”项下色谱条件进样分析，代入线性回归方程计算含量。结果显示，6-姜酚等5种成分含量的RSD均小于3.00%（n=6），表明方法重复性良好。

（5）稳定性考察：将同一供试品溶液（SP03）分别于室温下放置0、2、4、8、12、24 h时按“2.1.3”项下色谱条件进样分析。结果显示，6-姜酚等5种成分峰面积的RSD均小于3.00%（n=6），表明供试品溶液在24 h内基本稳定。

（6）加样回收率考察：精密称定已知含量的干姜（SP03）粉末6份，每份0.25 g，按照约1∶1的比例加入各对照品，按“2.1.2”项下方法制备供试品溶液，按“2.1.3”项下色谱条件进样分析，计算得姜酮、6-姜酚、8-姜酚、10-姜酚、6-姜烯酚的平均加样回收率分别为100.43%、100.47%、100.27%、100.02%、99.84%，RSD分别为1.91%、1.38%、2.18%、2.07%、1.91%（n=6），表明方法准确度良好。

2.1.5　含量测定　取干姜及其不同炮制程度姜炭样品粉末，按“2.1.2”项下方法制备样品，按“2.1.3”项下色谱条件进样分析，每个样品重复测定3次，记录各样品峰面积，代入线性回归方程求出各样品中各成分的含量，结果见图3。随着炮制程度的加深，姜酮在炮制后产生且含量先增加后降低，标炭中含量最高；6-姜酚、8-姜酚和10-姜酚的含量逐渐降低；6-姜烯酚含量先增加后减少。

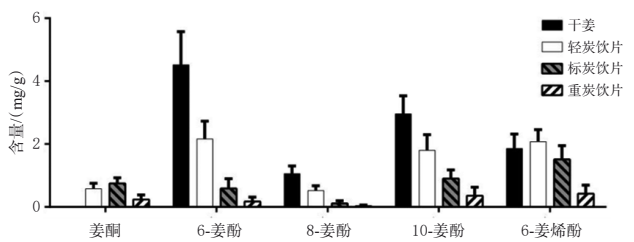


图3　干姜及其不同炮制程度姜炭的多成分含量测定结果（n=3）

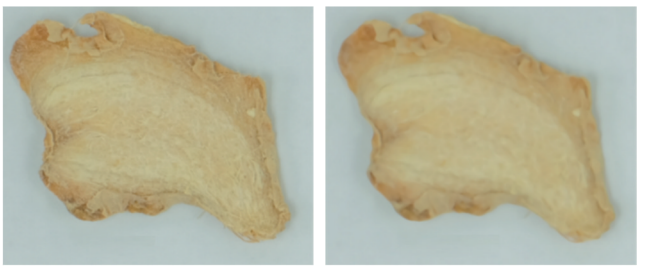
2.2　基于机器视觉技术的干姜及其不同炮制程度姜炭样品的图像采集与处理

2.2.1　拍摄参数设置　拍摄参数为所用型号灯箱推荐参数：手动对焦模式，光圈F/8，快门速度1/60 s，IOS 200，图像色彩为RGB模式，分辨率为4 000×4 000，图像保存格式为jpeg格式。被拍摄样品放置于灯箱底部的白纸正中央上，以减少背景颜色的干扰；采用三脚架固定镜头角度垂直向下，以防止角度偏移或拍摄过程中抖动引起饮片图像失真。

2.2.2　图像预处理　将样品图像进行降噪处理，减少图像颜色失真，提高图像质量；同时对图像中的样本与背景进行分离，提取目标区域，减少背景颜色的干扰。

（1）降噪处理：采用Bitonic滤波降噪器，除去由于光源不稳定、相机抖动等因素，在拍摄饮片的过程中会不可避免地引入噪声或者杂色，结果示例见图4。结果显示，降噪后图像在最大限度保留图像原有信息的情况下对图像噪点进行了一定的抑制或消除，提高了图像的质量。

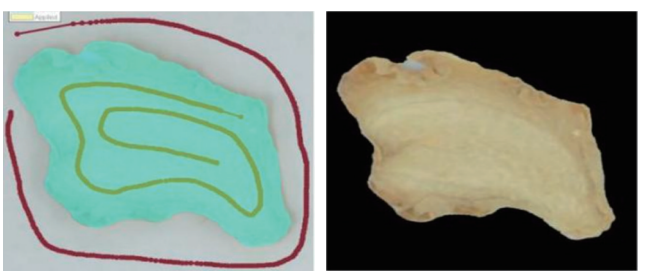
（2）目标区域选择：在RGB颜色空间中，纯黑色（R=0，G=0，B=0）背景能最大限度地减少背景的干扰，因此对饮片图像进行前景与背景的分离。通过使用



A.降噪前　　B.降噪后

图4　Bitonic滤波降噪前后的干姜图片示例

Matlab软件中的Image Segmenter工具箱，对姜炭饮片图像进行目标区域选择。结果如图5所示，蓝色区域为自动识别饮片主体区域，蓝色为勾画前景后计算机学习识别的前景区域，然后通过勾画背景色对背景进行自动识别并将背景色替换为纯黑色。



A.目标区域勾画前　　B.目标区域勾画后

图5　图像目标区域选择结果示例

2.2.3　图片颜色特征值的提取　以RGB颜色空间为基准，通过Matlab软件对样本图像的像素点进行颜色特征提取，得到R、G、B值；同时将RGB颜色空间分别转化为L*a*b*颜色空间、HSV颜色空间，提取得到L^*、a^*、b^*值和H、S、V值。

（1）RGB颜色空间特征值提取：RGB颜色空间，也称三基色模式，是生活中最常用的一个模型。它以R（Red：红）、G（Green：绿）、B（Blue：蓝）3种基本色为基础，进行不同程度的叠加，进而产生广泛且丰富的颜色。R、G、B值的范围均为0～255，通过选取图像上的任一像素点，可获得该像素点的R、G、B值。通过Matlab软件提取图片颜色特征值。

（2）L*a*b*颜色空间特征值提取：L*a*b*颜色空间由L^*、a^*和b^*这3个要素组成，其中L^*表示颜色的亮度（范围0～100，L^*=0表示黑色，L^*=100表示白色），a^*和b^*为2个颜色通道，a^*表示红绿色轴（范围－128～127，+a^*表示红色，－a^*表示绿色），b^*表示黄蓝色轴（范围－128～127，+b^*表示黄色，－b^*表示蓝色）。由于L*a*b*颜色空间所计算处理颜色偏差程度与人们主观感受相似，因此，将图像从RGB转换为L*a*b*[14]。RGB颜色空间无法直接转化为L*a*b*颜色空间，需按照RGB-XYZ-L*a*b*进行转换，转换公式如下：

$$\begin{cases} r=\text{gamma}(\dfrac{R}{255}) \\ g=\text{gamma}(\dfrac{G}{255}) \\ b=\text{gamma}(\dfrac{B}{255}) \end{cases} \quad (1)$$

$$\text{gamma}(x)=\begin{cases} (\dfrac{x+0.055}{1.055})^{2.4}, (x>0.040\ 45) \\ \dfrac{x}{12.92}, \text{otherwise} \end{cases} \quad (2)$$

$$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix}=\begin{pmatrix} 0.412\ 4, 0.357\ 6, 0.180\ 5 \\ 0.212\ 6, 0.715\ 2, 0.072\ 2 \\ 0.019\ 3, 0.119\ 2, 0.950\ 5 \end{pmatrix} \times \begin{pmatrix} r \\ g \\ b \end{pmatrix} \quad (3)$$

$$\begin{cases} L^*=116\text{f}(\dfrac{Y}{Y_n})-16 \\ a^*=500[\text{f}(\dfrac{X}{X_n})-\text{f}(\dfrac{Y}{Y_n})] \\ b^*=200[\text{f}(\dfrac{Y}{Y_n})-\text{f}(\dfrac{Z}{Z_n})] \end{cases} \quad X_n、Y_n、Z_n\text{一般为}95.047、100.0、108.883 \quad (4)$$

$$\text{f}(t)=\begin{cases} t^{1/3}, \text{if } t>(\dfrac{6}{29})^3 \\ \dfrac{1}{3}(\dfrac{29}{6})^2 t+\dfrac{4}{29}, \text{otherwise} \end{cases} \quad (5)$$

（3）HSV颜色空间特征值提取：HSV是一种比较直观的颜色模型，它与人眼对颜色的主观认识相对比较符合[15]。HSV颜色空间由色调（H）、饱和度（S）、明度（V）这3个参数组成。其中H为颜色的基本属性，即为常用的颜色名称，如青色、品红、红色等；S为颜色的鲜艳程度，饱和度越高则颜色越鲜艳，反之颜色越灰；V表示颜色的明亮程度。转换公式如下：

$$\begin{cases} R'=\dfrac{R}{255} \\ G'=\dfrac{G}{255} \\ B'=\dfrac{B}{255} \end{cases} \quad (6)$$

$$\begin{cases} C_{max}=\max(R',G',B') \\ C_{min}=\min(R',G',B') \\ \Delta=C_{max}-C_{min} \end{cases} \quad (7)$$

$$H=\begin{cases} 0°, \Delta=0 \\ 60°\times(\dfrac{G'-B'}{\Delta}+0), C_{max}=R' \\ 60°\times(\dfrac{B'-R'}{\Delta}+2), C_{max}=G' \\ 60°\times(\dfrac{B'-R'}{\Delta}+4), C_{max}=B' \end{cases} \quad (8)$$

$$S=\begin{cases} 0, C_{max}=0 \\ \dfrac{\Delta}{C_{max}}, C_{max}\neq 0 \end{cases} \quad (9)$$

$$V=C_{max} \quad (10)$$

2.2.4　机器视觉技术的精密度和稳定性考察　取干姜（SP15）样品，按“2.2.1”项下拍摄参数在同一时间点内连续拍摄6次，连续6 d在同一时间点拍摄，并按“2.2.2”项下图像预处理和“2.2.3”项下图像颜色特征提取方法提取颜色特征。9个颜色特征值的RSD均小于3%，表明机器视觉技术的精密度、稳定性良好。

2.2.5　干姜及其不同炮制程度姜炭样品颜色测量　为了保证样品的代表性和均匀性，对干姜及其不同炮制程度姜炭进行颜色测量，每批次中随机抽取40个饮片进行拍摄，共采集到3 200张饮片图像。将获得的饮片图像按照“2.2.2”项下图像预处理和“2.2.3”项下提取其颜色特征值，结果见表3。

从表3看出，在RGB颜色空间中，R、G、B 3个特征值总体趋势是随着炮制程度的加深而降低的，这三者的降低造成干姜在炮制过程中随着炮制程度的加深而呈现外观明亮度的降低，即影响L*a*b*颜色空间的L^*与b^*，L^*在炒炭过程中由69.30±1.95降低至23.74±3.82，b^*在炮制过程中由32.29±1.74降低至9.75±0.66，这与干姜在炮制过程中外观颜色由黄白色逐渐炭化变黑相吻合。而a^*有波动，但变化幅度较小，可能因为外观呈黑色以致红绿色变化不明显，因此无法在数值上体现炮制前后的差异性。HSV颜色空间中，H由261.15±6.63降低至86.45±4.65，说明在炮制过程中饮片外观由淡黄色逐渐变为棕褐色；S总体来说呈下降趋势，这与干姜在炮制过程中黄色调逐渐消退从而导致饱和度降低相一致；V值随着炮制过程的加深导致黑色面积逐渐增加而降低。在各颜色值的变化过程中（a^*除外）发现，从生品到轻炭，从轻炭到标炭，从标炭到重炭降低幅度依次减小，其中G、B、L^*、S、V从标炭饮片到重炭饮片反而有略微的增加，这可能和不同炮制程度颜色比较接近，其中标炭饮片和重炭饮片表面均为焦黑色或黑色，颜色更为接近，且黑色对于蓝绿色、饱和度区别不大，而出现数据比较接近的原因。故需要结合多个指标进行综合判断分析。

2.3　基于机器学习的干姜及其不同炮制程度姜炭的颜色定性判别分析

2.3.1　PCA模型判别分析　本实验将干姜及其不同炮制程度姜炭样品颜色特征作为输入变量，建立基于颜色特征的PCA模型判别。该PCA模型中当主成分数

表3　干姜及其不同程度姜炭的颜色特征值测定结果（$\bar{x}\pm s$，$n=40$）

样品	R	G	B	L^*	a^*	b^*	H	S	V
干姜	184.98±4.70	168.23±5.49	110.15±7.58	69.30±1.95	−0.73±0.86	32.29±1.74	261.15±6.63	65.97±2.15	43.19±2.97
轻炭饮片	99.00±11.92	77.87±7.98	43.11±2.98	34.62±3.72	5.12±1.39	23.85±3.58	139.77±6.83	30.54±3.13	16.90±1.17
标炭饮片	68.05±10.63	54.37±8.02	31.10±4.82	23.74±3.82	3.25±1.23	16.89±2.88	96.08±15.01	21.32±3.15	12.20±1.89
重炭饮片	61.23±3.29	60.23±3.10	44.80±2.89	25.13±1.38	−2.13±0.44	9.75±0.66	86.45±4.65	23.62±1.21	17.57±1.13

为2时，累积解释率为99.78%（PC1＝85.47%，PC2＝14.31%），前2个主成分能较好地解释样品99.78%的颜色特征，说明提取的信息具有较好的代表性。该PCA模型的得分如图6所示，由图可知，除了4批轻炭饮片与标炭饮片比较接近外，其余炮制品均能被较好分离，提示通过将颜色客观量化从而区分不同炮制程度姜炭具有一定的可行性。后续实验将采用有监督模式识别方法通过进一步学习提高判别的准确率。

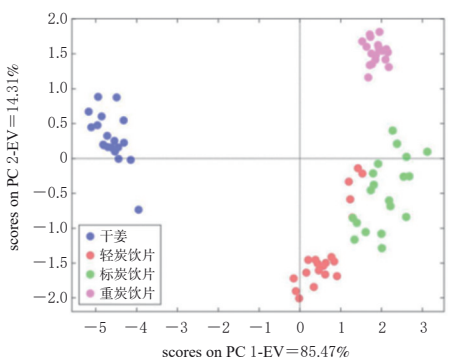


图6　干姜及其不同炮制程度姜炭PCA模型得分结果

2.3.2　有监督模式识别的干姜及其不同炮制程度姜炭的判别　为了提高模型对姜炭不同炮制品的判别效果，基于样品图像提取的颜色特征值，采用有监督模式识别方法，包括LDA、PLS-DA和SVM。以提取的9个颜色特征值（R、G、B、L^*、a^*、b^*、H、S、V）作为自变量，样品类别标签（干姜、轻炭饮片、标炭饮片、重炭饮片）作为因变量，以样品正确判别率为指标，采用10-折交叉验证法和外部验证法来评价模型性能。本次实验共有干姜及其不同炮制程度姜炭样品80批，将其随机分为训练集（56批）与验证集（24批），训练集用来训练模型，从而选出模型的最佳参数；验证集用于测试模型的应用效果。3种模型识别的识别结果如表4所示，PLS-DA对于不同炮制程度姜炭的交叉训练模型效果较差，标炭饮片预测准确度为11/14；LDA和SVM对所有批次样品识别的准确率都较高，两者在交叉验证训练过程中对全部类别的预测准确度均达100%，在外部验证中仅将1个标炭饮片样品误判为轻炭饮片样品。综上所述，LDA、SVM模型识别方法建立的模型对于姜炭炮制程度的预测结果准确快速，其中SVM对于学习和处理样本颜色特征间存在的非共线问题具有较大的优势，故选择SVM用于接下来的不同炮制程度姜炭的定量模型。

2.4　基于机器学习的干姜及其不同炮制程度姜炭的颜色-成分定量预测分析

2.4.1　颜色-成分相关性分析　使用SPSS 23.0软件将干姜及其不同炮制程度姜炭的图像颜色特征值与内在成分含量进行Pearson相关性分析，结果见表5。结果显示，除6-姜酚、8-姜酚、10-姜酚与颜色特征值a^*无差异外，其余各成分和颜色特征值均具有显著差异性。姜酮与颜色特征值R、G、B、L^*、b^*、H、S、V呈显著负相关关系，6-姜酚、8-姜酚和颜色特征值a^*呈负相关关系，其他成分与颜色特征值均呈正相关关系。故进一步对颜色和5种成分进行定量预测模型分析。

表4　不同模型识别方法对干姜及其不同炮制程度姜炭的识别结果

模型识别方法	参数	10-折交叉验证法（预测/真实）					外部验证法（预测/真实）				
		干姜	轻炭饮片	标炭饮片	重炭饮片	正判率/%	干姜	轻炭饮片	标炭饮片	重炭饮片	正判率/%
LDA		14/14	14/14	14/14	14/14	100	6/6	6/6	5/6	6/6	95.83
PLS-DA	components=4	14/14	12/14	11/14	13/14	89.28	6/6	6/6	5/6	6/6	95.83
SVM	RBF核函数 c=0.659 7,g=0.006 8	14/14	14/14	14/14	14/14	100	6/6	6/6	5/6	6/6	95.83

表5　干姜及其不同炮制程度姜炭的图像颜色特征值与内在成分的Pearson相关性分析结果

成分	R	G	B	L^*	a^*	b^*	H	S	V
姜酮	−0.602[a]	−0.689[a]	−0.767[a]	-0.665[a]	0.712[a]	−0.314[a]	−0.602[a]	−0.689[a]	−0.767[a]
6-姜酚	0.910[a]	0.889[a]	0.830[a]	0.898[a]	−0.067	0.877[a]	0.910[a]	0.889[a]	0.830[a]
8-姜酚	0.905[a]	0.885[a]	0.827[a]	0.893[a]	−0.063	0.871[a]	0.905[a]	0.885[a]	0.827[a]
10-姜酚	0.867[a]	0.832[a]	0.759[a]	0.844[a]	0.044	0.875[a]	0.867[a]	0.832[a]	0.759[a]
6-姜烯酚	0.488[a]	0.376[a]	0.230[b]	0.413[a]	0.634[a]	0.732[a]	0.488[a]	0.376[a]	0.230[b]

a：$P<0.01$；b：$P<0.05$

2.4.2　颜色-成分定量分析　将9个颜色特征值（R、G、B、L^*、a^*、b^*、H、S、V）作为自变量，5个成分含量作为因变量，使用SVM建立成分含量回归模型，采用10-折交叉验证法和外部验证法评价模型性能，以决定系数（R^2）、均方根误差（RMSE）和相对分析误差（RPD）作为回归模型的评价指标。R^2越大，RMSE越小，RPD值越大，表明定量模型的性能越好。样品划分方法同“2.3.2”项下，成分含量定量预测模型结果如表6所示。由表6可知，5种成分的RPD值均大于2，说明SVM对5种成分均有着较好的预测能力。各成分的验证集相关系数（R^2_P）与训练集的相关系数（R^2_C）中姜酮为0.633 9与0.683 3，其余4种成分的R^2_P和R^2_C均大于0.75，提示用SVM建立的模型预测5种成分含量时，除姜酮的稳定性较差，其余4种成分稳定性均较好。综上，基于机器视觉颜色数字化融合SVM机器学习方法建立的定量模型可用于预测不同炮制程度姜炭中的6-姜酚、8-姜酚、10-姜酚、6-姜烯酚4种成分的含量。

3　讨论

颜色特征是传统中药质量评价和中药炮制程度判断的重要标准，但由于不可避免的感官差异、主观性等原因，颜色特征评价容易受到干扰。随着技术的发展，

表6　干姜及其不同炮制程度姜炭的颜色-成分含量定量模型结果

成分	参数	R^2_C	RMSEC	R^2_P	RMSEP	RPD
姜酮	c=0.353 6,g=0.007 8	0.683 3	0.106 8	0.633 9	0.136 0	2.403 2
6-姜酚	c=1.414 2,g=9.765 6e^{-04}	0.867 3	0.600 5	0.792 7	0.671 7	2.747 5
8-姜酚	c=45.254 8,g=0.001 4	0.960 9	0.077 5	0.870 2	0.130 9	3.469 7
10-姜酚	c=0.313 7,g=0.002 0	0.897 9	0.310 4	0.817 2	0.361 7	2.925 7
6-姜烯酚	c=0.500 0,g=9.765 6e^{-04}	0.820 9	0.256 7	0.762 7	0.325 3	2.567 4

计算机视觉系统在农业、食品、中药和生物科学等领域得到了广泛的应用。机器学习为人工智能的重要组成部分，通过一定数目样本的学习能够根据颜色特征进行样品的鉴别，同时进行含量的预测。本文采用机器学习算法中应用最多的LDA、PLS-DA和SVM对不同炮制程度姜炭进行定性判别，结果LDA和SVM对所有批次样品识别的准确率达到100%。相对于LDA，SVM是在统计学习理论的基础上发展起来的，是一种专门研究小样本条件下机器学习的理论，它也比PLS-DA具有更强的逼近能力和泛化能力；与LDA相比，SVM对于学习和处理样本存在的非共线问题具有较大的优势，SVM更适合用小样本非共线性问题的建模预测[16]。故本研究将计算机视觉与SVM机器学习相结合，建立干姜及其不同炮制程度姜炭炮制品颜色参数和5种成分含量之间的定量模型，预测炮制过程中外部颜色与内部化学成分含量的变化，结果发现除了姜酮SVM模型的稳定性较差外，6-姜酚、8-姜酚、10-姜酚、6-姜烯酚4种成分的含量预测具有较高的预测精度、较好的稳定性和较小的预估错误率。因此，机器视觉系统可为姜炭等中药饮片的质量鉴别和炮制程度的控制提供一种快速、简单、高效的质量检测方法，可为批量大生产和智能化生产提供科学分析依据。

参考文献

[1] 国家药典委员会.中华人民共和国药典：一部[S]. 2020年版.北京：中国医药科技出版社，2020：15.

[2] LEEMANS V，DESTAIN M F. A real-time grading method of apples based on features extracted from defects [J]. J Food Eng，2004，61(1)：83-89.

[3] 段峰，王耀南，雷晓峰，等. 机器视觉技术及其应用综述[J]. 自动化博览，2002，(3)：59-61.

[4] 田恬. 英国皇家学会发布新的调研项目：机器学习[J]. 科技导报，2016，34(3)：93.

[5] 解达帅，刘玉杰，杨诗龙，等. 基于“内外结合”分析马钱子的炮制火候[J]. 中国实验方剂学杂志，2016，22(8)：1-5.

[6] 张银萍，徐燕，朱双杰，等. 基于机器视觉的金丝皇菊智能分级系统研究[J]. 食品工业科技，2022，43(5)：13-20.

[7] 陶兴宝，卢一，解达帅，等. 基于机器视觉及电子鼻技术的藤椒产地区分研究[J]. 中国调味品，2016，41(4)：122-125.

[8] 郭超，焦淑珍，徐超，等. 基于机器视觉技术的苦豆子颗粒外观质量评价研究[J]. 中国兽药杂志，2021，55(4)：27-38.

[9] 蒋苏贞，宓穗卿，王宁生. 姜酚心血管药理作用研究进展[J]. 时珍国医国药，2007，18(1)：219-221.

[10] 吴英智，傅强，严全能，等. 姜酚在心血管疾病中的药理作用研究进展[J]. 中国临床药理学杂志，2017，33(18)：1824-1827.

[11] 陈琛，耿剑亮，匡海学，等. UHPLC-Oribtrap/HRMS技术分析干姜中非挥发性化学成分在炮制前后的变化[J]. 南京中医药大学学报，2020，36(3)：387-395.

[12] 李田叶，刘卫华，梁娜，等. 评估生姜及其不同炮制品中姜辣素和其抗氧化活性[J]. 食品工业，2016，37(12)：180-183.

[13] 孟江，许舒娅，陈磊，等. 干姜“炒炭存性”质量标准初探[J]. 中国中药杂志，2012，37(4)：453-456.

[14] 牛秋月，李晓蓓，王晓鹏. 基于Lab颜色模型的聚类舌图像分割[J]. 电脑编程技巧与维护，2021，(1)：142-144.

[15] 徐曼菲，吴志生，刘晓娜，等. 从辨色论质谈中药质量评价方法[J]. 中国中药杂志，2016，41(2)：177-181.

[16] FAN W X，XU Q，WANG L，et al. Prediction of chemical component contents of the fruit of *Xanthium strumarium* L. during processing based on a computer vision system combined with a support vector machine[J]. Anal Methods，2019，11(25)：3260-3268.

（收稿日期：2022-05-25　修回日期：2022-09-13）

（编辑：曾海蓉）