

• 药剂与工艺 •

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

任延娜¹, 冯文豪², 李涵¹, 范雪花¹, 李海洋¹, 姚静^{3,4,5,6}, 王艳丽^{3,4,5}, 刘瑞新^{3,4,5,6*}, 李学林^{3,4,5,6*}

1. 河南中医药大学, 河南 郑州 450046
2. 河南省科学院同位素研究所有限责任公司, 河南 郑州 450000
3. 河南中医药大学第一附属医院, 河南 郑州 450000
4. 河南省中药临床应用、评价与转化工程研究中心, 河南 郑州 450000
5. 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
6. 河南省中药临床药理学中医药重点实验室, 河南 郑州 450000

摘要: **目的** 探究人工智能感官与多源信息融合技术用于中药五味药性二分类辨识方法的可行性, 为中药药性评价提供新的方法借鉴。**方法** 选取 122 种仅含单一味不含兼味的 5 类代表性中药饮片(源自《中国药典》2020 年版)及 14 种常用的食品类样本, 使用 PEN3 型电子鼻及 ASTREE、SA402B 型电子舌采集 136 种样本的智能感官信息, 以得到的信息矩阵作为自变量(X), 药典项下的性味描述等作为标杆信息(Y), 利用主成分分析-判别分析(principal component analysis-discriminant analysis, PCA-DA)、最小二乘-支持向量机(least squares-support vector machine, LS-SVM) 2 种化学计量学方法, 分别基于单一型号智能感官设备(单源)和多智能感官信息融合(多源)建立五味二分类(酸/非酸、咸/非咸、辛/非辛、甘/非甘、苦/非苦)辨识模型 $Y=F(X)$, 以交互验证的正判率作为模型优选指标。**结果** 经留一法交互验证, 基于单源信息的五味二分类模型中最大正判率分别是 98.53% (ASTREE/PCA-DA、LS-SVM)、97.06% (ASTREE/PCA-DA)、84.56% (ASTREE/LS-SVM)、89.71% (ASTREE/LS-SVM)、84.56% (ASTREE/LS-SVM), 基于多源信息的五味二分类辨识模型的最大正判率分别是 99.26% [(ASTREE+SA402B)/PCA-DA]、99.26% [(ASTREE+SA402B、PEN3+ASTREE+SA402B)/PCA-DA]、88.97% [(PEN3+ASTREE+SA402B)/LS-SVM]、91.91% [(PEN3+ASTREE+SA402B)/PCA-DA]、86.76% [(ASTREE+SA402B)/PCA-DA], 多源信息融合后模型的正判率有所提高, 平均提高 2.35% ($P<0.01$)。**结论** ASTREE 型电子舌对于中药五味二分类辨识方面表现良好; 多源信息融合后模型的正判率较单独使用任一电子鼻或电子舌均有所提高, 为中药五味药性评价提供一定参考。**关键词:** 五味; 药性; 二分类; 智能感官; 多源信息融合

中图分类号: R283.6 文献标志码: A 文章编号: 0253-2670(2023)10-3080-13

DOI: 10.7501/j.issn.0253-2670.2023.10.005

Discussion on binary classification identification method of five flavors of drug properties based on artificial intelligence sensory and multi-source information fusion technology

REN Yan-na¹, FENG Wen-hao², LI Han¹, FAN Xue-hua¹, LI Hai-yang¹, YAO Jing^{3,4,5,6}, WANG Yan-li^{3,4,5}, LIU Rui-xin^{3,4,5,6}, LI Xue-lin^{3,4,5,6}

1. Henan University of Chinese Medicine, Zhengzhou 450046, China

收稿日期: 2022-12-22

基金项目: 国家自然科学基金资助项目(81774452); 国家自然科学基金资助项目(81001646); 河南省中医药拔尖人才培养项目资助(2019ZYBJ07); 国家重点研发计划中医药现代化重点专项(2017YFC1703402); 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); 国家中医临床研究基地科研专项(2021JDZY106); 河南省中医科学研究基地专项(2021JDZY104); 河南中医药大学 2022 年度研究生科研创新能力提升计划项目(2022KYCX027); 河南省科技攻关项目(222102310377)

作者简介: 任延娜(1996—), 女, 硕士, 研究方向为中药饮片临床应用现代化关键技术研究。E-mail: ryn2021@163.com

***通信作者:** 刘瑞新(1980—), 硕士生导师, 主任药师, 主要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(1960—), 男, 教授, 博士生导师, 主任药师, 主要从事中药应用形式研究。Tel: (0371)66245142 E-mail: xuelinli450000@163.com

2. Institute of Isotope Research, Henan Academy of Sciences, Zhengzhou 450000, China
3. The First Affiliated Hospital of Henan University of Chinese Medicine, Zhengzhou 450000, China
4. Henan Province Engineering Research Center for Clinical Application, Evaluation and Transformation of Traditional Chinese Medicine, Zhengzhou 450000, China
5. Co-construction Collaborative Innovation Center for Chinese Medicine Prevention and Treatment of Respiratory Diseases by Henan & Education Ministry, Henan University of Chinese Medicine, Zhengzhou 450046, China
6. Henan Key Laboratory for Clinical Pharmacy of Traditional Chinese Medicine, Zhengzhou 450000, China

Abstract: Objective To explore the feasibility of artificial intelligence sensory and multi-source information fusion technology for the binary classification identification method of five flavors of traditional Chinese medicine (TCM), and provide a new method for the evaluation of drug properties of TCM. **Methods** A total of 122 representative TCM decoction pieces with only one flavor (derived from the 2020 edition of *Chinese Pharmacopoeia*) and 14 commonly used food samples were selected. The intelligent sensory information of 136 samples was collected by PEN3 electronic nose, ASTREE electronic tongue and SA402B electronic tongue. The obtained information matrix was used as the independent variable X , and the description of the nature and flavor under the items in pharmacopoeia was used as the benchmark information Y . Two chemometric methods, principal component analysis-discriminant analysis (PCA-DA) and least squares support vector machine (LS-SVM), were used. Based on a single type of intelligent sensory equipment (single source) and multi-intelligent sensory information fusion (multi-source), the five flavors two classification (acid/non-acid, salty/non-salty, spicy/non-spicy, sweet/non-sweet, bitter/non-bitter) identification model $Y = F(X)$ was established, and the positive rate of cross-validation was used as the model optimization index. **Results** Through the leave-one-out cross-validation method, the maximum correct judgment rates in the five-flavor two-classification model based on single-source information were 98.53% (ASTREE/PCA-DA, LS-SVM), 97.06% (ASTREE/PCA-DA), 84.56% (ASTREE/LS-SVM), 89.71% (ASTREE/LS-SVM), 84.56% (ASTREE/LS-SVM). The maximum correct rate of the two-classification identification model based on multi-source information was 99.26% [(ASTREE + SA402B)/PCA-DA], 99.26% [(ASTREE + SA402B, PEN3 + ASTREE + SA402B)/PCA-DA], 88.97% [(PEN3 + ASTREE + SA402B)/LS-SVM], 91.91% [(PEN3 + ASTREE + SA402B)/PCA-DA], 86.76% [(ASTREE + SA402B)/PCA-DA], respectively. The correct rate of the model after multi-source information fusion was improved, with an average increase of 2.35% ($P < 0.01$). **Conclusion** ASTREE electronic tongue performs well in the identification of five flavors and two classifications of TCM. The correct judgment rate of the model after multi-source information fusion is higher than that of any single electronic nose or electronic tongue, which provides a reference for the evaluation of five flavors of TCM.

Key words: five flavors; medicinal properties; binary classification; intelligent senses; multi-source information fusion

中药药性理论包括四气、五味、升降浮沉、归经、毒性等，其中五味理论是中药药性理论的核心内容之一^[1]。《神农本草经》指出：“药有酸、咸、甘、苦、辛五味”，首次把药味的概念引入本草著作，产生了药物五味，并将其作为药物五味理论的基本思想^[2]。“五味”最初的定义是源于人们对中药滋味、气味的实际感受，故有“非口不能味也”。中药五味药性理论虽然来源于口尝味并和口尝味有着密不可分的关系，且其大部分与口尝味相一致（吻合率为74%^[3]），但二者不能完全等同，其含义更多的是中药的功能味、物质基础以及发挥的中药药理作用等。随着广大学者将目光聚焦在传统中医药文化上，越来越多的科学技术被运用在对传统中药饮片五味药性的辨识中。

在生物层面，众多学者采用味觉受体的方式选择五味的受体，探索不同的研究方法，包括对辛^[4]、

苦^[5-6]、酸^[7]、甘^[8]、咸^[9]传统中药五味药性理论的辨识及其化学生物学实质的系统分析，基于五味药性的定义和功效内涵，从滋味气味、生物效应、体内过程及动力学规律等不同方面，提出中药五味药性的表征方法；在分子层面，韩彦琪等^[10]基于五味药性的基本定义以及药物分子-味觉、嗅觉受体的相互作用关系，提出了基于分子对接技术的中药五味药性表征的技术方法和研究路径，尝试从分子学角度区分五味等。以上2方面仅提出了五味药性表征的方法，如何客观地对五味药性进行辨识并量化，人工智能感官（artificial intelligence sense, AIS）技术为其提供了可能。AIS技术是模拟人体感官（眼、舌、鼻、耳和皮肤等）的一类仿生学技术，是判定滋味、气味等的客观方法，最早应用于食品领域^[11]，主要用于质量评价^[12-13]、风味等级划分^[14]等。随后在医学保健、药品等多个领域均有所涉及^[15]，其在

中药饮片研究中的应用也越来越受到重视^[16]。在中药五味药性评价的过程中, AIS 技术能够提取中药的药性特征并依据此进行模式识别^[7], 这为中药五味药性评价带来新的曙光^[17]。中药五味是基于药物的成分, 电子舌、电子鼻的响应也是基于药物的成分与物质基础, 因此, 二者具有一定的联系。

国外学者 Kataoka 等^[18]尝试将电子舌用于有苦味或涩味的药用植物和中药复方的味觉评价和质量控制; 国内也有学者对 22 种常用中药的水煎液进行电子舌检测, 得到酸、苦、甜、咸的规律聚类分布图和特征雷达图谱, 以及酸、苦、甜味判别分析模型, 通过电子舌技术实现了中药不同“滋味”的分类^[19]; 本课题组前期也已采用经典人群口尝法 (traditional human taste panel method, THTPM) 与电子舌法相结合的方式, 以 47 种天然药物及食品类成分为标准物质, 开展了中药 4 类味觉 (酸、甜、苦、咸) 的定量与定性研究, 达到了 80% 以上的准确率^[20]。以上是对苦味、涩味或是对酸、苦、甜、咸四类“滋味”的研究, 并没有对中药的功能味进行系统的研究, 也没有对中药功能味做出准确的客观描述, 且以上均采用单一的智能感官技术, 虽然使用单一智能感官能够通过该仪器的多个传感器获得样本的综合信息, 一定程度上具有整体性, 但表达的依然是某一感官的单方面信息。那么, 将不同的智能感官联合使用, 融合多源信息, 能否得到更好的五味功能味的分类效果, 从而提高中药五味药性评价的准确度, 值得深入研究。因此, 本研究拟基于智能感官与多源信息融合技术, 在前期研究的基础上, 扩大样本数量, 基于《中国药典》2020 年版筛选 122 种单一味道不含兼味的中药饮片及 14 种常用食品类成分作为载体, 采集 136 种样本的感官数据, 利用主成分分析-判别分析 (principal component analysis-discriminant analysis, PCA-DA)、最小二乘-支持向量机 (least squares-support vector machine, LS-SVM) 2 种化学计量学方法分别建立感官数据单源、多源的五味二分类 (酸/非酸、咸/非咸、辛/非辛、甘/非甘、苦/非苦) 辨识模型, 以模型交互验证正判率为指标, 探索 AIS 与多源信息融合技术用于中药五味判别分析的可行性, 以期能够实现新的药性定量评价方法的探索。

1 仪器与材料

1.1 仪器

BSA224S-CW 型万分之一电子天平, 赛多利斯

科学仪器有限公司; LCD-A200 型电子天平, 福州华志科学仪器有限公司; 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; 药典筛, 四号筛, 0.25 μm 孔径, 浙江上虞市道墟五四仪器厂; 30MF5 3L 型多功能养生壶, 深圳市正云科技有限公司; HK250 型科导台式超声清洗器, 上海科导超声仪器有限公司; HHS 型电热恒温水浴锅, 常州普天仪器制造有限公司; TD5M 型高速离心机, 上海卢湘仪离心机仪器有限公司; PEN3 型电子鼻 (传感器信息见表 1), 德国 Airsense 公司; ASTREE 型电子舌 (传感器信息见表 2), 法国 Alpha M. O. S. 公司; SA402B 型电子舌 (传感器信息见表 3), 日本 Insent 公司。

1.2 材料

所选 122 种中药饮片均收载于《中国药典》2020 年版 (酸味 12 种, 辛味 31 种, 咸味 7 种, 甘味 36 种, 苦味 36 种), 因所记载酸味无兼味与咸味无兼味中药较少, 同时考虑甘味饮片水煎液味淡, 对于后续模型分类贡献较小, 增补乳酸、无水柠檬酸、苹果酸、黑醋、白醋为酸味样本; 增补食盐、NaCl 为咸味样本; 增补红糖、蔗糖、果糖、饴糖、麦芽糖、葡萄糖、核糖为甘味样本; 共计 136 种样本, 见表 4。

表 1 PEN3 型电子鼻传感器敏感物质

Table 1 Sensitive substances of PEN3 electronic nose sensor

传感器编号	传感器代码	敏感物质
1	W1C	芳烃化合物
2	W5S	氮氧化合物
3	W3C	氨, 芳香分子
4	W6S	氢化物
5	W5C	烯烃、芳族, 极性分子
6	W1S	烷类
7	W1W	硫化合物
8	W2S	醇, 部分芳香族化合物
9	W2W	芳烃化合物, 硫的有机化合物
10	W3S	烷类和脂肪族

表 2 ASTREE 型电子舌传感器与味觉信息

Table 2 ASTREE electronic tongue sensor and taste information

序号	传感器名称	味觉信息	序号	传感器名称	味觉信息
1	SRS	酸味	5	SPS	复合味觉
2	GPS	复合味觉	6	SWS	甜味
3	STS	咸味	7	BRS	苦味
4	UMS	鲜味			

表3 SA402B型电子舌8根传感器与11种味觉值的匹配信息

Table 3 Matching information between eight sensors of SA402B electronic tongue and 11 taste values

传感器	对应的味觉	味觉信息		传感器	对应的味觉	味觉信息	
		先味	回味			先味	回味
C00	酸性苦味	苦味 (bitterness)	酸性苦味回味 (aftertaste-B, AB)	AE1	涩味	涩味 (astringency)	涩味回味 (aftertaste-A, AA)
AN0	碱性苦味	×	碱性苦味回味 (B-bitterness2, BB)	GL1	甜味	甜味 (sweetness)	×
BT0	碱基盐类 苦味	×	碱基盐类苦味回味 (H-bitterness, HB)	CA0	酸味	酸味 (sourness)	×
				CT0	咸味	咸味 (saltness)	×
				AAE	鲜味	鲜味 (umami)	丰富度 (richness)

“×”表示无味觉信息值

“×” indicates no taste information value

表4 136种样本来源信息

Table 4 Sources information of 136 kinds of samples

编号	类别	名称	产地	批号	编号	类别	名称	产地	批号	编号	类别	名称	产地	批号
1	酸味	山茱萸	河南	2007033	30	辛味	薄荷	安徽	2005212	59	甘味	淡竹叶	四川	1909212
2		沙棘	-	-	31		浮萍	安徽	1908240062	60		密蒙花	四川	1912192
3		山楂叶	-	-	32		半边莲	江苏	1911092	61		金银花	河南	191119
4		乌梅	四川	191103	33		鱼腥草	四川	1912052	62		忍冬藤	山东	2003170212
5		石榴皮	安徽	2007071	34		地锦草	浙江	1805152	63		地骨皮	山西	190801
6		布渣叶	-	-	35		闹羊花	-	-	64		地黄	河南	2009081
7		白矾	浙江	1905131	36		徐长卿	山东	2001112	65		车前草	江西	2008130072
8		花蕊石	-	-	37		广藿香	广西	2003162	66		炙甘草	内蒙古	200714
9		五倍子	-	-	38		佩兰	安徽	2007182	67		丝瓜络	湖北	1910072
10		杠板归	-	-	39		豆蔻	广西	20062805	68		茯苓	安徽	200902
11		马齿苋	河北	1912102	40		干姜	云南	200701	69		猪苓	山西	2007041
12		木瓜	湖北	2008052	41		花椒	-	-	70		泽泻	四川	2008043
13		乳酸	河南	-	42		丁香	广西	2007121	71		薏苡仁	贵州	200901
14		无水柠檬酸	河南	2019121603A	43		红花	新疆	101102	72		通草	四川	191002
15		L-苹果酸	-	004-00345	44		卷柏	-	-	73		冬瓜皮	河北	2002222
16		黑醋	山西	-	45		高良姜	广西	2001311	74		炒鸡内金	山东	2008163
17		白醋	山西	-	46		小茴香	-	-	75		炒麦芽	安徽	2008193
18	辛味	生姜	-	-	47		八角茴香	-	-	76		白茅根	河北	2005103
19		香薷	湖北	1909292	48		荜茇	广西	2007071	77		竹茹	湖北	2006101
20		荆芥	河北	2002182	49	咸味	煅瓦楞子	山东	2006151	78		胖大海	广西	200501
21		大腹皮	广西	2008112	50		石决明	-	-	79		桑白皮	安徽	2001122
22		千金子	-	-	51		昆布	山东	2002182	80		柏子仁	河南	20090101
23		白芷	安徽	2007151	52		牡蛎	山东	2009051	81		合欢皮	湖北	2008293
24		藁本	辽宁	1911281	53		珍珠母	浙江	2004131	82		钩藤	江西	20072805
25		大蒜	-	-	54		大青盐	青海	191004	83		百合	湖南	200101
26		川芎	四川	200701	55		地龙	广西	191101	84		枸杞子	宁夏	191105
27		麸煨肉豆蔻	广西	2008103	56		食盐	云南	6923229821758	85		玉竹	湖南	1911292
28		辛夷	河南	2007171	57		NaCl	天津	160904	86		龙眼肉	广西	200401
29		细辛	辽宁	2002152	58	甘味	芦根	河北	1912192	87		大枣	河南	200801

续表 4

编号	类别	名称	产地	批号	编号	类别	名称	产地	批号	编号	类别	名称	产地	批号
88	甘味	山药	河南	2007031	105	苦味	黄连	重庆	2007231	122	苦味	丹参	山西	2007223
89		月季花	-	-	106		黄柏	四川	2007153	123		马鞭草	-	-
90		黄芪	甘肃	191103	107		龙胆	辽宁	2001022	124		浙贝母	浙江	2008101
91		石斛	四川	200401	108		苦参	吉林	2001080110	125		莲子心	湖南	191102
92		酒黄精	湖北	191001	109		赤芍	内蒙古	200901	126		绵草薹	江西	2006223
93		蜂蜜	吉林	1909307	110		连翘	陕西	191105	127		虎杖	安徽	2002181
94		红糖	山东	6926636302689	111		大青叶	甘肃	1912032	128		射干	河北	191119
95		蔗糖	山东	692663630267	112		板蓝根	安徽	191102	129		白苡	湖南	1901080062
96		果糖	河南	-	113		山豆根	广西	19110209	130		漏芦	内蒙古	1904092
97		饴糖	河南	190421	114		大黄	甘肃	2007093	131		大血藤	湖北	1908101
98		麦芽糖	河南	20170519	115		防己	江西	1912311	132		千里光	-	-
99		葡萄糖	山西	20170321	116		路路通	湖北	2001292	133		川木通	四川	2001101
100		D-核糖	河南	-	117		荷叶	湖南	2006122	134		北刘寄奴	湖北	1912012
101	苦味	栀子	江西	200901	118		篇蓄	安徽	1908240102	135		北豆根	-	-
102		王不留行	河南	200601	119		瞿麦	河北	1909160362	136		预知子	-	-
103		黄芩	山西	2009101	120		槐花	山东	1812292					
104		胡黄连	西藏	1904091	121		茜草	山西	1912172					

“-”表示从市场零售购买,无批号记载

“-” indicates purchase from retail in the market, without batch number record

2 方法

聘请河南中医药大学第一附属医院在中药鉴定方向工作多年有丰富经验的施钧瀚副主任药师对所购买的中药饮片进行初步筛选,根据结果将存在争议的饮片筛选出,按照《中国药典》2020年版项下性状鉴别、显微鉴别、薄层鉴别等方法对其进行检验与验证,若不符合要求,仍需购买新的样本作为补充,最终确定所有样本均为正品且符合《中国药典》2020年版规定。

2.1 智能感官信息(X)值的获取

2.1.1 基于PEN3型电子鼻获取X值

(1) 样本制备:称取所购买的样本适量,置于万能粉碎机打碎,过4号筛得到粉末15g,备用。

(2) 信息采集:取样本粉末2g,置于25mL的烧杯中,双层保鲜膜封口,于室温24℃左右、相对湿度82%左右的条件下放置30min后测试。采用直接顶空吸气法直接将进样针头插入含样本的密封样本杯中,电子鼻进行测定。

测定条件:采样时间为每组1s;传感器自清洗时间为80s;传感器归零时间为5s;样本准备时间为5s;进样流量为400mL/min;分析采样时间为60s。每个样品采集3次,取平均值,得到10根传感器嗅觉信息矩阵X(136×10)。

2.1.2 基于ASTREE型电子舌获取X值

(1) 饮片水煎液的制备^[21-22]:参考《中国药典》2020年版推荐用量的均值制备122种中药饮片水煎液。取药典用药量范围平均值(如山茱萸用量为6~12g,则取9g)的10倍量90g制备样本,将其置于煎药壶中,一煎加水以液面淹没过药材约2~5cm,浸泡30min,武火煮沸后用文火煎30min;二煎加水以液面没过药材约1~2cm,武火煮沸后再用文火煎煮20min,合并滤液,混匀,冷至室温,4000r/min离心15min,取上清液,定容至1000mL,取100mL灌装,压盖,高温蒸汽灭菌,备用(煎药所用水均为纯水)。

(2) 添加剂样本的配制:所增补食品或食品添加剂类样本的配制:将麦芽糖、饴糖、葡萄糖、红糖、蔗糖、果糖、D-核糖、乳酸、L-苹果酸、无水柠檬酸、食盐、NaCl配制为质量浓度为4%的溶液;将黑醋、白醋配制为体积浓度为5%的溶液。

(3) 信息采集:参考文献报道方法^[23-24],每次采样时间120s,后接清洗时间10s×2次,每个样本重复测量5次,取后3次测试结果的平均值,实验结束最后的清洗时间为120s×3次。输出数据发现传感器UMS(鲜味)对于所有样本检测值均为-9997.21,该变量对于模型构建或分类无贡献,故

删去该传感器,留下6根传感器(如表2所示),得到6根传感器味觉信息矩阵 $X(136 \times 6)$ 。

2.1.3 基于 SA402B 型电子舌获取 X 值

(1) 样本制备:同“2.1.2”项下样本制备方法。

(2) 信息采集:参考文献报道方法^[25],得到8根传感器输出的11种味觉值的信息矩阵 $X(136 \times 11)$ 。

2.2 标杆信息的获取

122种中药饮片以药典【性味与归经】项下味的描述作为标杆信息(Y),14种食品类样本以参考文献及古籍等分类为标准,若无参考信息,暂将其“滋味”定义为“药味”。其中,甜味添加剂红糖味甘;麦芽糖性味甘温^[26];饴糖、葡萄糖功效与果糖类似^[27],具有补体虚寒冷、健壮脾胃、润肺止咳、补中益气,治虚劳腹疼、吐血健脾的作用,认为其为甘味;酸味样本补充为黑醋、白醋,记载于《本草别录》,味酸。样本共分为酸和非酸样本2类,标杆信息“1”代表酸,“2”代表非酸(咸、辛、甘、苦分类同)。

2.3 辨识模型的构建

2.3.1 基于单源智能感官信息的辨识模型构建 分别以PEN3型电子鼻、ASTREE型电子舌、SA402B型电子舌3种感官信息矩阵为自变量 $X(136 \times n)$, n 表示传感器/变量个数),以中国药典【性味与归经】项下味的描述等为标杆信息 Y ,利用PCA-DA、LS-SVM2种化学计量学方法建立感官数据的五味二分类(酸/非酸、咸/非咸、辛/非辛、甘/非甘、苦/非苦)单源辨识模型 $Y=F(X)$,模型验证方法为留一法交互验证。具体建模方法参考文献报道^[20]。

2.3.2 基于多源信息融合的辨识模型构建 辨识模型构建方法同“2.3.1”项,其中自变量 $X(136 \times n)$, n 表示传感器/变量个数)分别是以PEN3型电子鼻、ASTREE型电子舌和SA402B型电子舌的感官信息两两组合(二源数据融合)与PEN3型电子鼻、ASTREE型电子舌和SA402B型电子舌的感官信息共同组合(三源数据融合)组成的信息矩阵。

2.4 模型原始变量(传感器)对分类贡献度的追踪

以PEN3型电子鼻的10根传感器、ASTREE型电子舌6根传感器与SA402B型电子舌11个味觉变量信息为原始变量,结合模型原始变量载荷图,通过各传感器所携带的变异信息值(Wilk's Lambda值),追踪各传感器对模型分类的贡献度大小。

2.5 模型的建立与评价

由于传统数据分析和提取方法不适合于多种感

官信息交叉的矩阵,本研究在确定自变量与因变量的基础上,建立数学模型实现对感官信息的预测和辨识,且由于中药的特殊属性以及采集的数据量较大,在此基础上还要节约时间成本,因此,选择PCA-DA、LS-SVM2种化学计量学方法建立感官数据的五味二分类辨识模型,以交互验证的正判率作为模型优选指标,并综合 F_1 值对模型判别结果进行评价与探析。PCA-DA算法,即将PCA与DA结合,是一种基于PCA的分析方法,通过最大化类间方差的比率和最小化类内方差的比率来进一步压缩高维数据,从而探索能够解释数据集主要趋势的变量的组合^[28-29]。LS-SVM是SVM在二次损失函数形式下的一种扩展,通过求解一组线性方程代替支持向量机中复杂的二次优化问题的非线性建模方法,其保留了SVM处理小样本、高维、非线性问题的能力^[30]。其中 F_1 值为精确率(precision)和召回率(recall)的调和平均,其取值在0~1,数值越大表示模型效果越好。

$$F_1 = (\text{精确率} \times \text{召回率} \times 2) / (\text{精确率} + \text{召回率})$$

$$\text{精确率} = TP / (TP + FP)$$

$$\text{召回率} = TP / (TP + FN)$$

TP代表真阳性,FP代表假阳性,FN代表假阴性

3 结果

3.1 酸和非酸二分类辨识结果

3.1.1 基于PEN3型电子鼻的酸和非酸二分类辨识结果

(1) PCA-DA辨识结果:变量选择10个,数据缩放选择自动缩放,识别模式选择线性,验证方法选择百叶窗交互验证。选取4个主成分时判别错误率最低,4个主成分得分之和达99%以上,可解释样本大部分变异信息。模型交互验证正判率为87.81%,且由模型主成分得分图(图1)可以看出,2类样本在二维空间中有一部分重叠,但根据模型正判率结果来看,都实现了良好的分类,且不存在未分类样本。

如表5所示,根据PEN3型电子鼻10根传感器所携带的变异信息值,变异信息值越小表明变量所携带变异信息越多。二分类辨识模型中7、10、8、6号传感器所携带变异信息较多,对酸和非酸模型分类的贡献度较大,1号传感器所携带的变异信息较少。结合图2模型变量载荷图,7、10、8、6号传感器均为负相关变量。

(2) LS-SVM辨识结果:变量选择10个,通过

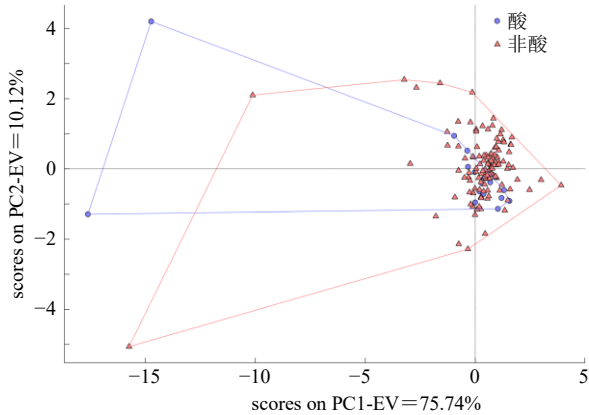


图1 PCA-DA 模型主成分得分图 (PEN3 型电子鼻)

Fig. 1 PCA-DA model principal component score chart (PEN3 electronic nose)

表5 PEN3 型电子鼻的变异信息值

Table 5 Wilk' Lambda value of PEN3 electronic nose

序号	变量	变异信息值	序号	变量	变异信息值
1	7	0.932 74	6	5	0.972 24
2	10	0.942 85	7	4	0.973 59
3	8	0.945 16	8	3	0.979 95
4	6	0.949 22	9	9	0.984 11
5	2	0.951 44	10	1	0.992 11

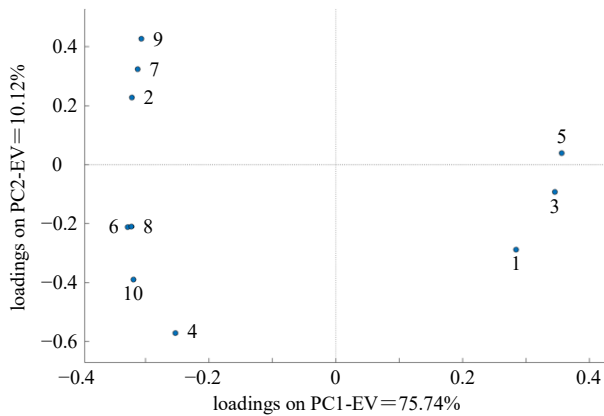


图2 PCA-DA 模型变量载荷图 (PEN3 型电子鼻)

Fig. 2 PCA-DA model variable load diagram (PEN3 electronic nose)

优化 LS-SVM 的核函数类型及 gam 值(LA-SVM 中的正则化参数, 决定了适应误差的最小化和平滑程度) 等参数来优化模型, 选取最优模型正判率为 86.03%。

3.1.2 基于 ASTREE 型电子舌的酸和非酸二分类辨识结果

(1) PCA-DA 辨识结果: 变量选择 6 个, 数据缩放选择自动缩放, 识别模式选择线性, 验证方法选择百叶窗交互验证。选取 4 个主成分时判别错误

率最低, 4 个主成分得分之和达 99%以上, 可解释样本大部分变异信息。模型交互验证正判率为 98.53%, 且由模型主成分得分图(图 3)可以看出, 2 类样本在二维空间中都实现了良好的分类, 且不存在未分类样本。

如表 6 所示, 根据 ASTREE 型电子舌 6 根传感器所携带的变异信息值, 二分类辨识模型中 1、6、2、3 号传感器所携带变异信息较多, 对酸和非酸二分类模型的贡献度较大。由图 4 可知, 1、6、2 号传感器为正相关变量, 3 号传感器为负相关变量。

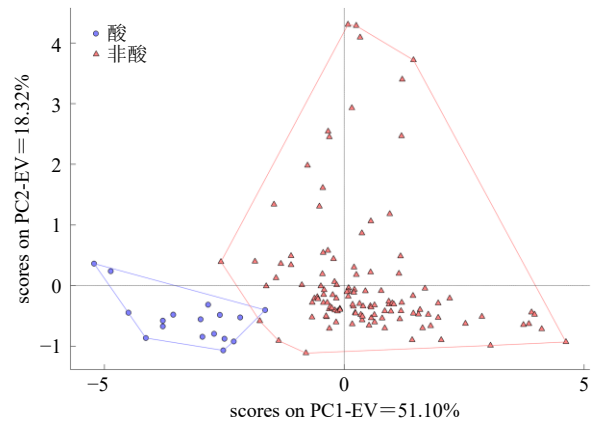


图3 PCA-DA 模型主成分得分图 (ASTREE 型电子舌)

Fig. 3 PCA-DA model principal component score chart (ASTREE electronic tongue)

表6 ASTREE 型电子舌的变异信息值

Table 6 Wilk's Lambda value of ASTREE electronic tongue

序号	变量	变异信息值	序号	变量	变异信息值
1	1	0.480 57	4	3	0.686 45
2	6	0.587 14	5	5	0.969 99
3	2	0.656 44	6	4	0.987 29

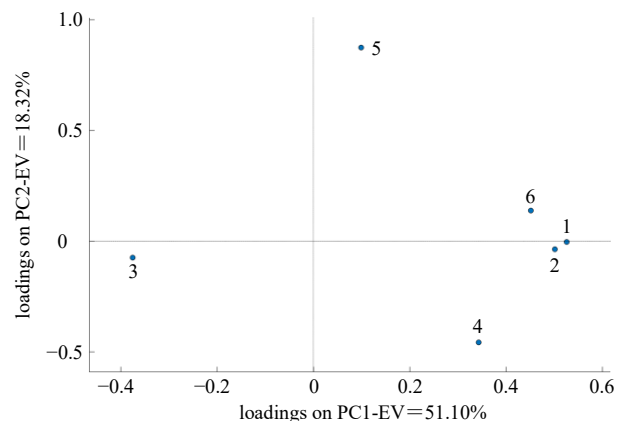


图4 PCA-DA 模型变量载荷图 (ASTREE 型电子舌)

Fig. 4 PCA-DA model variable load diagram (ASTREE electronic tongue)

(2) LS-SVM 辨识结果: 变量选择 6 个, 通过优化 LS-SVM 的核函数类型及 gam 值等参数来优化模型, 选取最优模型正判率为 98.53%。

3.1.3 基于 SA402B 型电子舌的酸和非酸二分类辨识结果

(1) PCA-DA 辨识结果: 变量选择 11 个, 数据缩放选择自动缩放, 识别模式选择线性, 验证方法选择百叶窗交互验证。选取 5 个主成分时判别错误率最低, 5 个主成分得分之和达 99% 以上, 可解释样本大部分变异信息。模型交互验证正判率为 94.85%, 且由模型主成分得分图 (图 5) 可以看出, 两类样本在二维空间中有一部分重叠, 但根据模型正判率结果来看, 都实现了良好的分类, 且不存在未分类样本。

如表 7 所示, 根据 SA402B 型电子舌 11 个味觉变量信息所携带的变异信息, 二分类辨识模型中 4、5、9、8 号传感器所携带变异信息较多, 对酸和非酸模型分类的贡献度较大, 6 号传感器所携带的变异信息较少。结合图 6 模型变量载荷图可知, 5 号传感器为正相关变量, 4、9、8 号传感器为负相关变量。其中 2 号传感器距离原点较近, 说明该特征

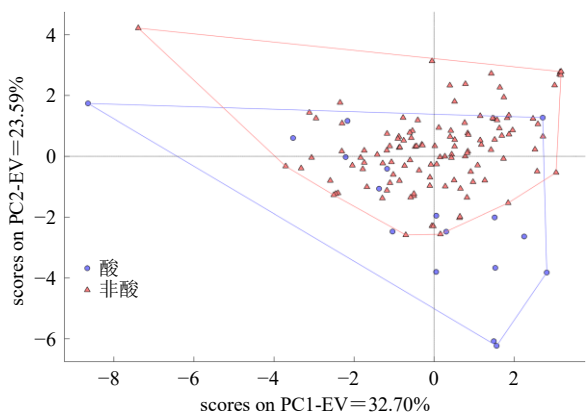


图 5 PCA-DA 模型主成分得分图 (SA402B 型电子舌)
Fig. 5 PCA-DA model principal component score chart (SA402B electronic tongue)

表 7 SA402B 型电子舌的变异信息值

Table 7 Wilk's Lambda value of SA402B electronic tongue					
序号	变量	变异信息值	序号	变量	变异信息值
1	4	0.597 29	7	10	0.991 41
2	5	0.850 46	8	2	0.994 72
3	9	0.858 74	9	11	0.996 52
4	8	0.943 77	10	7	0.998 34
5	1	0.972 89	11	6	0.999 01
6	3	0.975 36			

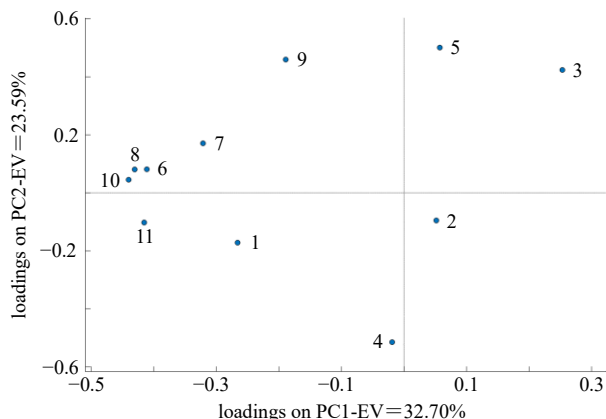


图 6 PCA-DA 模型变量载荷图 (SA402B 型电子舌)
Fig. 6 PCA-DA model variable load diagram (SA402B electronic tongue)

属性的波动对样本之间的区别贡献不大。

(2) LS-SVM 辨识结果: 变量选择 11 个, 通过优化 LS-SVM 的核函数类型及 gam 值等参数来优化模型, 选取最优模型正判率为 91.91%。

3.1.4 基于 PEN3 电子鼻+ASTREE 型电子舌+SA402B 型电子舌的酸和非酸二分类辨识结果

(1) PCA-DA 辨识结果: 变量选择 27 个, 数据缩放选择自动缩放, 识别模式选择线性, 验证方法选择百叶窗交互验证。选取 17 个主成分时判别错误率最低, 17 个主成分得分之和达 99% 以上, 可解释样本大部分变异信息。模型交互验证正判率为 97.79%, 且由模型主成分得分图 (图 7) 可以看出, 2 类样本在二维空间中有一部分重叠, 但根据模型正判率结果来看, 都实现了良好的分类, 且不存在未分类样本。

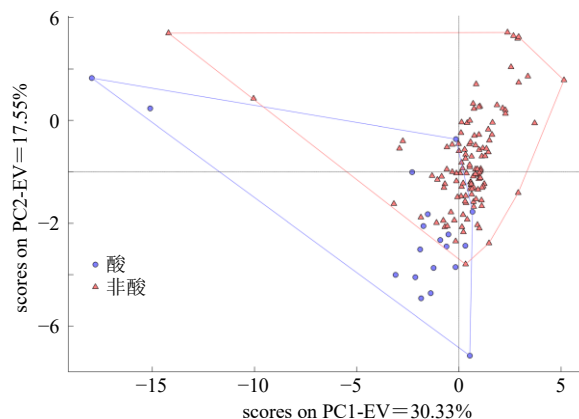


图 7 PCA-DA 模型主成分得分图 (PEN3 电子鼻+ASTREE 型电子舌+SA402B 型电子舌)
Fig. 7 PCA-DA model principal component score chart (PEN3 electronic nose + ASTREE electronic tongue + SA402B electronic tongue)

如表 8 所示, 根据 27 个变量所携带的变异信息值, 二分类辨识模型中 11 号 (SRS 传感器/ASTREE 型电子舌)、16 号 (BRS 传感器/ASTREE 型电子舌)、20 号 (sourness/SA402B 型电子舌)、12 号 (GPS 传感器/ASTREE 型电子舌)、13 号 (STS 传感器/ASTREE 型电子舌) 所携带变异信息较多, 对酸和非酸模型分类的贡献度较大, 22 号 (astringency/SA402B 型电子舌)、23 号 (AB/SA402B 型电子舌) 所携带变异信息较少。且结合图 8 所示, 11、12、16 号为正相关变量, 20、13 号为负相关变量。18 号 (HB/SA402B 型电子舌) 距离原点较近, 说明其特征属性的波动对样本分类影响较小。

表 8 3 类传感器融合的变异信息值

Table 8 Wilk's Lambda values of three types of sensor fusion

序号	变量	变异信息值	序号	变量	变异信息值
1	11	0.501 38	15	9	0.972 26
2	16	0.585 77	16	5	0.973 71
3	20	0.605 96	17	17	0.977 94
4	12	0.662 43	18	19	0.979 83
5	13	0.665 24	19	4	0.980 01
6	21	0.854 45	20	3	0.981 11
7	25	0.858 52	21	14	0.989 30
8	7	0.925 30	22	26	0.992 05
9	24	0.947 75	23	1	0.992 71
10	8	0.953 42	24	18	0.993 56
11	2	0.953 51	25	27	0.997 62
12	10	0.954 42	26	23	0.998 18
13	6	0.956 90	27	22	0.998 30
14	15	0.966 56			

表 9 单源多源数据的酸和非酸二分类模型及参数对比

Table 9 Acid and non-acid binary classification model and parameter comparison of single source and multi-source data

分类		PCA-DA		LS-SVM	
		正判率/%	F ₁ /%	正判率/%	F ₁ /%
单源	PEN3 型电子鼻 (n=10)	87.81	19.05	86.03	9.52
	ASTREE 型电子舌 (n=6)	98.53*	94.44	98.53*	94.12
	SA402B 型电子舌 (n=11)	94.85	75.86	91.91	56.00
多源	PEN3 型电子鼻+ASTREE 型电子舌 (n=16)	98.53	94.44	86.76	NA
	PEN3 型电子鼻+SA402B 型电子舌 (n=21)	94.85	75.86	86.76	NA
	ASTREE 型电子舌+SA402B 型电子舌 (n=17)	99.26*	97.14	98.53	94.12
	PEN3 型电子鼻+ASTREE 型电子舌+SA402B 型电子舌 (n=27)	97.79	91.89	97.06	86.67

“*”表示最高正判率, “n”表示传感器/变量个数, “NA”表示无法得到有效值, 下表同

“*” indicates the highest positive judgment rate, “n” indicates the number of sensors/variables, “NA” indicates not available, the table below is the same

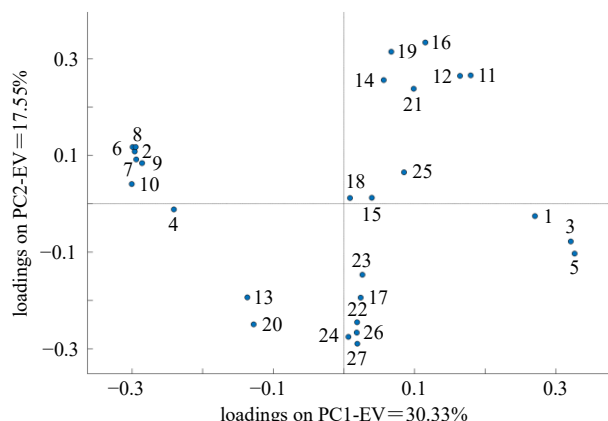


图 8 PCA-DA 模型变量载荷图 (PEN3 电子鼻+ASTREE 型电子舌+SA402B 型电子舌)

Fig. 8 PCA-DA model variable load diagram (PEN3 electronic nose + ASTREE electronic tongue + SA402B electronic tongue)

(2) LS-SVM 辨识结果: 变量选择 27 个, 通过优化 LS-SVM 的核函数类型及 gam 值等参数来优化模型, 选取最优模型正判率为 97.06%。

因基于 PEN3 型电子鼻+SA402B 型电子舌、PEN3 型电子鼻+ASTREE 型电子舌与 SA402B 型电子舌+ASTREE 型电子舌的二源感官数据融合下的模型二分类辨识方法与“3.1.4 (1)”项均类似, 仅自变量有所不同, 因此, 本项仅保留模型判别结果 (下文同), 如表 9 所示。

将 136 种样本分为酸和非酸 2 类, 二分类辨识结果基本达到了预期。其中, 单源智能感官辨识结果中, 除电子鼻数据分类结果相对较差, 其余都达到了 90%以上, 且 ASTREE 型电子舌正判率较高于 PEN3 型电子鼻和 SA402B 型电子舌; 多源信息融合的结果显示, 模型最优正判率高于单源智能感官

模型最优正判率。由 F_1 值结果可知, PCA-DA 模型的 F_1 值整体高于 LS-SVM, LS-SVM 模型中的部分 F_1 值不可用, 且基于电子鼻的 F_1 值较小, 表明 PCA-DA 模型准确度较好。

3.2 咸和非咸二分类辨识结果

本项下模型二分类辨识方法与“3.1”项均类似, 自变量与因变量有所不同, 因此, 仅保留模型判别结果, 如表 10 所示。

基于 PEN3 型电子鼻、ASTREE 型电子舌和

SA402B 型电子舌的咸与非咸二分类辨识模型可以得到, 单源智能感官辨识结果中, ASTREE 型电子舌正判率表现相对较好于 PEN3 型电子鼻和 SA402B 型电子舌。多源信息融合结果显示, 模型正判率相对较高于单源智能感官模型正判率。由 F_1 值可知, 除电子鼻外, 其他单源或多源的 PCA-DA 模型的 F_1 值整体高于 LS-SVM, LS-SVM 模型中的部分 F_1 值不可用。基于单源、多源数据融合模型正判率结果见表 10。

表 10 单源多源数据的咸和非咸二分类模型及参数对比

Table 10 Salty and non-salty binary classification model and parameter comparison of single source and multi-source data

分类	PCA-DA		LS-SVM		
	正判率/%	F_1 /%	正判率/%	F_1 /%	
单源	PEN3 型电子鼻 ($n=10$)	94.12	55.56	94.85	58.82
	ASTREE 型电子舌 ($n=6$)	97.06*	71.43	94.85	36.36
	SA402B 型电子舌 ($n=11$)	93.38	18.18	93.38	NA
多源	PEN3 型电子鼻+ASTREE 型电子舌 ($n=16$)	97.79	82.35	95.59	62.50
	PEN3 型电子鼻+SA402B 型电子舌 ($n=21$)	97.79	82.35	94.85	46.15
	ASTREE 型电子舌+SA402B 型电子舌 ($n=17$)	99.26*	94.12	97.06	71.43
	PEN3 型电子鼻+ASTREE 型电子舌+SA402B 型电子舌 ($n=27$)	99.26*	94.12	97.79	82.35

3.3 辛和非辛二分类辨识结果

本项基于 PEN3 型电子鼻、SA402B 型电子舌和 ASTREE 型电子舌的辛与非辛二分类辨识模型可以得到, 单源智能感官辨识结果中, 基于 ASTREE 型的 PCA-DA 算法的模型正判率较低于 PEN3 型电子鼻的, 而基于 ASTREE 型电子舌的 LS-SVM 算法的正判率较高于 PEN3 型电子鼻和 SA402B 型电子舌; 多源信息融合结果显示, 模型最优正判率高于单源智能感官模型最优正判率。由 F_1 值可知, 基于 PEN3 型电子鼻+ASTREE 型电子舌的 PCA-DA 模型的 F_1 值大于 LS-SVM, 其他单源或多源的 PCA-

DA 模型的 F_1 值小于 LS-SVM, PCA-DA 模型中的部分 F_1 值不可用。基于单源、多源数据融合模型正判率结果见表 11。

3.4 甘和非甘二分类辨识结果

基于 PEN3 型电子鼻、SA402B 型电子舌和 ASTREE 型电子舌的甘与非甘二分类辨识模型可以得到, 单源智能感官辨识结果中, ASTREE 型电子舌正判率结果较高于 PEN3 型电子鼻和 SA402B 型电子舌。多源信息融合结果显示, 模型最优正判率高于单源智能感官模型最优正判率。由 F_1 值可知, 基于 PEN3 型电子鼻+ASTREE 型电子舌以及 PEN3

表 11 单源多源数据的辛和非辛二分类模型及参数对比

Table 11 Spicy and non-spicy binary classification model and parameter comparison of single source and multi-source data

分类	PCA-DA		LS-SVM		
	正判率/%	F_1 /%	正判率/%	F_1 /%	
单源	PEN3 型电子鼻 ($n=10$)	83.09	43.90	82.35	45.45
	ASTREE 型电子舌 ($n=6$)	77.21	NA	84.56*	63.16
	SA402B 型电子舌 ($n=11$)	75.00	5.56	83.82	67.65
多源	PEN3 型电子鼻+ASTREE 型电子舌 ($n=16$)	83.09	48.89	82.35	45.45
	PEN3 型电子鼻+SA402B 型电子舌 ($n=21$)	82.35	47.83	83.09	48.89
	ASTREE 型电子舌+SA402B 型电子舌 ($n=17$)	77.94	40.00	88.24	71.43
	PEN3 型电子鼻+ASTREE 型电子舌+SA402B 型电子舌 ($n=27$)	80.88	50.00	88.97*	73.68

型电子鼻+ASTREE 型电子舌+SA402B 型电子舌数据融合的 PCA-DA 模型的 F_1 值大于 LS-SVM, 其他单源或多源的 PCA-DA 模型的 F_1 值小于 LS-SVM。基于单源、多源数据融合的甘与非甘二分类辨识模型正判率见表 12。

3.5 苦和非苦二分类辨识结果

基于 PEN3 型电子鼻、SA402B 型电子舌和 ASTREE 型电子舌的苦与非苦二分类辨识模型可以得到, 单源智能感官辨识结果中, ASTREE 型电子

舌正判率结果较高于 PEN3 型电子鼻和 SA402B 型电子舌。多源信息融合结果显示, 模型最优正判率高于单源智能感官模型最优正判率。由 F_1 值可知, 基于 PEN3 型电子鼻+ASTREE 型电子舌+SA402B 型电子舌数据融合的 PCA-DA 模型的 F_1 值大于 LS-SVM, 其他单源或多源的 PCA-DA 模型的 F_1 值小于 LS-SVM, PCA-DA 模型中的部分 F_1 值不可用。基于单源、多源数据融合的苦与非苦二分类辨识模型正判率见表 13。

表 12 单源多源数据的甘和非甘二分类模型及参数对比

Table 12 Sweet and non-sweet binary classification model and parameter comparison of single source and multi-source data

分类		PCA-DA		LS-SVM	
		正判率/%	F_1 /%	正判率/%	F_1 /%
单源	PEN3 型电子鼻 ($n=10$)	72.06	34.48	81.62	64.79
	ASTREE 型电子舌 ($n=6$)	83.09	70.13	89.71*	83.72
	SA402B 型电子舌 ($n=11$)	72.06	45.71	69.85	46.75
多源	PEN3 型电子鼻+ASTREE 型电子舌 ($n=16$)	88.97	82.76	82.35	66.67
	PEN3 型电子鼻+SA402B 型电子舌 ($n=21$)	75.00	55.26	81.62	64.79
	ASTREE 型电子舌+SA402B 型电子舌 ($n=17$)	88.24	80.95	91.18	85.37
	PEN3 型电子鼻+ASTREE 型电子舌+SA402B 型电子舌 ($n=27$)	91.91*	87.06	81.62	64.79

表 13 单源多源数据的苦和非苦二分类模型及参数对比

Table 13 Bitter and non-bitter binary classification model and parameter comparison of single source and multi-source data

分类		PCA-DA		LS-SVM	
		正判率/%	F_1 /%	正判率/%	F_1 /%
单源	PEN3 型电子鼻 ($n=10$)	73.53	NA	83.09	63.49
	ASTREE 型电子舌 ($n=6$)	83.09	54.90	84.56*	66.67
	SA402B 型电子舌 ($n=11$)	78.68	49.12	81.62	59.02
多源	PEN3 型电子鼻+ASTREE 型电子舌 ($n=16$)	79.41	54.84	85.29	69.70
	PEN3 型电子鼻+SA402B 型电子舌 ($n=21$)	79.41	53.33	83.09	62.30
	ASTREE 型电子舌+SA402B 型电子舌 ($n=17$)	86.76*	68.97	86.03	69.84
	PEN3 型电子鼻+ASTREE 型电子舌+SA402B 型电子舌 ($n=27$)	86.03	67.80	83.09	63.49

4 讨论

4.1 酸味、咸味样本的分析

由本研究结果可知, 基于单源、多源信息融合构建的模型对酸味、咸味二分类能力较好, 甘味、辛味次之, 苦味最小。且根据 F_1 值显示, 对于酸味、咸味样本, 即使 PCA-DA、LS-SVM 2 种算法模型的正判率较高, 但其对应的 F_1 值部分差值悬殊, 且部分 F_1 值不可用。推测其原因是酸味、咸味样本与其他味样本之间差异大, 特征明显, 甘味与辛味样本基数较大, 所能筛选的共有特点较多, 对于模式识别结果存在影响。而苦味正判率相对较小, 与中

药水煎液的共性——苦有关, 因此苦味样本与非苦味样本的差异不够明显, 导致模型不易区分, 这点从电子舌苦味输出均较高可以看出。

以 PEN3 型电子鼻响应分析为例, 对于酸味样本, 由表 5 可知, 对酸和非酸二分类结果贡献度较大的原始变量是 7、10、8、6 号传感器, 根据 PEN3 型电子鼻 10 根传感器对应的敏感物质, 这些传感器响应的物质主要为硫化物、烷类、脂肪族、醇和部分芳香族化合物; 对于咸味样本, 对咸和非咸二分类结果贡献度较大的原始变量是 1、3、5、9 号传感器, 这些传感器响应的物质主要为芳烃化合物、

氨、芳香分子、烯烃、芳族、极性分子、硫的有机化合物。结合文献研究,如山茱萸、乌梅等酸味样本,其中山茱萸包括环烯醚萜、鞣质、黄酮、三萜、芳香酚酸、挥发油类等化学成分^[31],乌梅中主要化学成分类型为有机酸类、黄酮类、萜类、醇类、芳香类化合物等^[32],说明电子鼻能够实现酸味样本辨识是具有相应物质基础的。

4.2 辛味样本的分类

PCA-DA 算法模型中,单源信息建模时,电子鼻对于辛+非辛的二分类判别显示出明显的优势,主要原因是所选择的辛味饮片大多气味强烈,如生姜、荆芥、大蒜、薄荷、花椒、丁香、小茴香、八角茴香等,具有明显可嗅的辛辣味道,所含有的挥发性成分多,因此,电子鼻信息值建模效果较好。本研究中,虽然所选择的电子舌不含有辛味传感器,但依然将辛味药物纳括在内,原因在于,一方面,查阅文献可知,已经有相关学者通过电子舌数据分析了辛味药物的物质基础并实现了其分类^[33];另一方面,辛味药物常含有较多的挥发性成分,且具有明显的刺激性气味,本研究添加了电子鼻技术作为相应支撑,利用多传感器融合技术将多种感官信息进行组合,提取特征分析,作为分类判别的依据。

4.3 未分类与误分类样本的分析

经统计发现,2种判别模型均存在误分类的样本,不存在未分类的样本,其中,误分类样本在PCA-DA或LS-SVM判别模型中分布较集中,以在本研究中分类结果较好的酸味、咸味为例,如酸味,单源及多源数据融合下,误分类的样本主要集中在山楂叶、布渣叶、花蕊石、黑醋,花蕊石描述为“气微、味淡”,多外用,水煎液味道淡,略苦,不具有明显酸味,其功效化痰止血,与酸“能收、能涩”相对应,但其味淡可能对于电子舌数据融合表征造成影响以至于分类错误;而黑醋出现多次误分类,考虑其主要原因是所购买的为市售黑醋样本,其中存在的添加剂等成分影响其特征。咸味,单源及多源数据融合下,误分类的样本主要集中在瓦楞子、石决明、珍珠母、地龙等。地龙的药效成分主要为蛋白及多肽、酶类、核苷酸、氨基酸、微量元素等,与普通咸味药不同之处在于其含有的矿物质较少^[34];且有学者曾从地龙水煎液中检测出32个挥发性成分,有12个成分能够嗅出气味,地龙的腥臭味大多源于此^[35]。这可能是其与其他咸味样本相比较独特之处。而其他样本未正确分类的原因尚需进

一步研究。

4.4 多源信息融合分析

多源信息融合即把多类传感器数据融合并做预处理,在一定准则下加以自动分析、综合的信息处理过程。从本研究结果来看,多传感器信息融合建模的正判率显然更为理想,其中酸味提高的较少为0.73%,辛味提高的较多为4.41%,总体来讲,多传感器信息融合建模后模型正判率较单一的AIS技术平均提高2.35%,多种信息的互补初步实现了中药五味的二分类。本研究是基于多传感器间的数据级融合,未来也可探讨特征级融合和决策级融合是否可以进一步改善其分类性能。特征级融合又称中级融合,即首先提取各传感器原始信息特征矢量,然后基于融合的特征矢量对样本识别和评估。决策级融合又称高级融合,首先分别提取各传感器原始数据的特征信息,生成特征矢量来识别和预测目标,然后融合各传感器的识别结果,得到被测目标的一致性解释说明。且在本研究中,将电子舌与电子鼻的信息值融合在一起,其输出值一定条件下属于两种不同系统的感官信息,在模式识别之前,如何将不同系统的智能感官信息进行前处理,为不同的模式识别提供多样的预处理方法,可能是优化模式识别结果的突破点之一。

5 结论

本研究中,以3种智能感官对五味二分类的辨识结果可见,相比于PEN3型电子鼻与SA402B型电子舌,ASTREE型电子舌输出的信息值整体正判率较高。经多传感器信息融合后,新组成的数据集辨识模型的正判率较单独使用任一电子鼻或电子舌均有所提高。综上,本研究基于AIS与多源信息融合技术建立的中药五味二分类模型,达到了分类需求,能够为中药五味辨识提供一定的参考。

利益冲突 所有作者均声明不存在利益冲突

参考文献

- [1] 鲁耀邦,赵权. 中药学 [M]. 北京: 中国农业大学出版社, 2015: 17-25.
- [2] 张铁军,刘昌孝. 中药五味药性理论辨识及其化学生物学实质表征路径 [J]. 中草药, 2015, 46(1): 1-6.
- [3] 周祯祥,汪琼,李晶晶,等. 基于《中国药典》(2010年版)中的五味研究 [J]. 湖北中医药大学学报, 2015, 17(2): 63-64.
- [4] 孙玉平,张铁军,曹煌,等. 中药辛味药性表达及在临证配伍中的应用 [J]. 中草药, 2015, 46(6): 785-790.
- [5] 张静雅,曹煌,许浚,等. 中药苦味药性表达及在临证

- 配伍中的应用 [J]. 中草药, 2016, 47(2): 187-193.
- [6] Matsunami H, Montmayeur J P, Buck L B. A family of candidate taste receptors in human and mouse [J]. *Nature*, 2000, 404(6778): 601-604.
- [7] 曹煌, 张静雅, 龚苏晓, 等. 中药酸味的药性表达及在临证配伍中的应用 [J]. 中草药, 2015, 46(24): 3617-3622.
- [8] 张静雅, 曹煌, 龚苏晓, 等. 中药甘味的药性表达及在临证配伍中的应用 [J]. 中草药, 2016, 47(4): 533-539.
- [9] 张静雅, 曹煌, 龚苏晓, 等. 中药咸味药性表达及在临证配伍中的应用 [J]. 中草药, 2016, 47(16): 2797-2802.
- [10] 韩彦琪, 许浚, 龚苏晓, 等. 基于味觉、嗅觉受体分子对接技术的中药性味物质基础研究的路径和方法 [J]. 中草药, 2018, 49(1): 14-19.
- [11] 刘淼. 智能人工味觉分析方法在几种食品质量检验中的应用研究 [D]. 浙江大学, 2012.
- [12] Fei C H, Ren C C, Wang Y L, *et al.* Identification of the raw and processed *Crataegi Fructus* based on the electronic nose coupled with chemometric methods [J]. *Sci Rep*, 2021, 11(1): 1849.
- [13] Li Y, Fei C H, Mao C Q, *et al.* Physicochemical parameters combined flash GC E-nose and artificial neural network for quality and volatile characterization of vinegar with different brewing techniques [J]. *Food Chem*, 2022, 374: 131658.
- [14] 韦真博. 伏安型电子舌的研发及其在食品检测中的应用 [D]. 浙江大学, 2011.
- [15] Nissen F, Quint J K, Morales D R, *et al.* How to validate a diagnosis recorded in electronic health records [J]. *Breathe*, 2019, 15(1): 64-68.
- [16] 白杰, 高利利, 张志勤, 等. 电子舌技术的原理及在中药领域的应用 [J]. 中南药学, 2021, 19(1): 78-84.
- [17] 贾灿潮, 李俊妮, 李荣, 等. 中药快速分析技术的发展及其前景 [J]. 药物分析杂志, 2018, 38(9): 1476-1483.
- [18] Kataoka M, Tokuyama E, Miyanaga Y, *et al.* The taste sensory evaluation of medicinal plants and Chinese medicines [J]. *Int J Pharmaceut*, 2008, 351(1/2): 36-44.
- [19] 杜瑞超, 王优杰, 吴飞, 等. 电子舌对中药滋味的区分辨识 [J]. 中国中药杂志, 2013, 38(2): 154-160.
- [20] 王艳丽, 陈鹏举, 桂新景, 等. 基于电子舌的天然药物的四类味觉分类辨识研究 [J]. 中华中医药杂志, 2021, 36(1): 423-433.
- [21] 蒋志, 蒋丽霞, 李智韬, 等. 中药汤剂煎煮方法及影响因素研究进展 [J]. 广州中医药大学学报, 2022, 39(2): 458-462.
- [22] 李艳, 白明, 宋亚刚, 等. 中药标准汤剂的研究与思考 [J]. 中草药, 2018, 49(17): 3977-3980.
- [23] Lin Z Z, Zhang Q, Liu R X, *et al.* Evaluation of the bitterness of traditional Chinese medicines using an E-tongue coupled with a robust partial least squares regression method [J]. *Sensors*, 2016, 16(2): 151.
- [24] 刘瑞新, 李慧玲, 李学林, 等. 基于电子舌的穿心莲水煎液的掩味效果评价研究 [J]. 中草药, 2013, 44(16): 2240-2245.
- [25] 谢亚妮, 沈思涵, 陈光宇, 等. 基于电子舌智能感官评定与人工评价结合的五汁饮矫味技术研究 [J]. 中国现代应用药学, 2022, 39(6): 772-776.
- [26] 卢传坚. 孩子肝脾不和, 可以吃点麦芽糖 [J]. 家庭医药. 快乐养生, 2019(1): 44.
- [27] 刘宗利, 李克文, 王京博, 等. 低聚果糖的理化特性、生理功效及其应用 [J]. 中国食品添加剂, 2016(10): 211-215.
- [28] 王小鹏, 张璐, 陈鹏举, 等. 近红外光谱技术应用于中药四类味觉分类辨识的可行性分析 [J]. 中草药, 2023, 54(4): 1076-1086.
- [29] Jiang H Z, Hu Y L, Jiang X S, *et al.* Maturity stage discrimination of *Camellia oleifera* fruit using visible and near-infrared hyperspectral imaging [J]. *Molecules*, 2022, 27(19): 6318.
- [30] 丁海泉, 温江北, 高洪智, 等. 最小二乘支持向量机用于人参皂苷近红外光谱分析研究 [J]. 现代农业装备, 2021, 42(5): 17-22.
- [31] 范倩, 陈雪冰, 荣莉, 等. 山茱萸化学成分、生物活性、复方应用及质量控制研究进展 [J]. 天然产物研究与开发, 2020, 32(7): 1244-1258.
- [32] 杨亚焱, 王瑞, 钱程程, 等. 乌梅化学成分、药理作用研究进展及质量标志物预测 [J/OL]. 中成药. [2023-04-24]. <http://kns.cnki.net/kcms/detail/31.1368.R.20220328.1319.002.html>.
- [33] 曹煌, 张铁军, 张静雅, 等. 基于电子鼻和电子舌技术的辛味中药气-味的表征研究 [J]. 中草药, 2016, 47(11): 1962-1967.
- [34] 李思维, 郝二伟, 杜正彩, 等. 广地龙化学成分和药理作用的研究进展及其质量标志物 (Q-Marker) 的预测分析 [J]. 中草药, 2022, 53(8): 2560-2571.
- [35] 张萌, 徐硕, 高志, 等. 地龙酒炙前后水煎液气味的吹扫捕集-GC-MS-嗅闻联用研究 [J]. 中国医药导报, 2012, 9(35): 129-131.

[责任编辑 郑礼胜]