

基于电子鼻和一维拉普拉斯卷积核的奶粉基粉产地鉴别

张寅升¹, 袁航², 周亚¹, 程永波³, 王海燕^{1*}

(1. 浙江工商大学食药质量安全工程研究院, 浙江杭州 310018) (2. 浙江工商大学管理工程与电子商务学院, 浙江杭州 310018) (3. 南京财经大学管理科学与工程学院, 江苏南京 210023)

摘要: 奶粉基粉是配方奶粉的基础原材料, 其产地影响到终端乳制品的品质。本文提出了一种电子鼻技术与一维拉普拉斯卷积核相结合的基粉奶源地判别方法。通过电子鼻采集样本数据, 经过时域信号对齐, 尝试使用不同阶数的一维拉普拉斯卷积核进行特征提取, 并对比了统计数字特征、快速傅里叶变换、离散余弦变换等其他特征提取方法, 最后使用偏最小二乘及可视化进行可分性分析。实验结果发现快速傅里叶变换、离散余弦变换、二阶的一维拉普拉斯卷积核相对于原始特征均有效提升了可分性, 偏最小二乘的 R^2 效应量从0.61分别提升到0.95、0.96和1.00。一维拉普拉斯卷积核特征提取方法能够准确区分产自国内和国外(澳大利亚)的基粉, 在案例研究中取得了最优判别效果, 说明其能够有效提取到电子鼻各通道序列信号的时间响应特征。该方法能够完成中澳两国奶粉基粉样本的区分工作, 为快速鉴定奶粉基粉来源提供技术支持。

关键词: 电子鼻; 奶粉; 偏最小二乘法; 拉普拉斯卷积核

文章编号: 1673-9078(2024)05-240-246

DOI: 10.13982/j.mfst.1673-9078.2024.5.0299

Identification of Dehydrated Milk Powder Origin Based on Electronic Nose and One-dimensional Laplacian Convolution Kernel

ZHANG Yinsheng¹, YUAN Hang², ZHOU Ya¹, CHENG Yongbo³, WANG Haiyan^{1*}

(1. Zhejiang Food and Drug Quality & Safety Engineering Research Institute, Zhejiang Gongshang University, Hangzhou 310018, China) (2. School of Management and E-Business, Zhejiang Gongshang University, Hangzhou 310018, China) (3. School of Management Science and Engineering, Nanjing University of Finance and Economics, Nanjing 210023, China)

Abstract: Dehydrated milk powder serves as the fundamental raw material for formula milk, and its place of origin affects the quality of terminal dairy products. In this paper, a method for discriminating the source of dehydrated milk powder by employing electronic nose technology and utilizing a one-dimensional Laplacian convolution kernel is proposed. Sample data is collected using an electronic nose, and after temporal signal alignment, various orders of one-dimensional Laplacian convolution kernels are applied to extract features. Additionally, other feature extraction methods such as statistical

引文格式:

张寅升,袁航,周亚,等.基于电子鼻和一维拉普拉斯卷积核的奶粉基粉产地鉴别[J].现代食品科技,2024,40(5):240-246.

ZHANG Yinsheng, YUAN Hang, ZHOU Ya, et al. Identification of dehydrated milk powder origin based on electronic nose and one-dimensional laplacian convolution kernel [J]. Modern Food Science and Technology, 2024, 40(5): 240-246.

收稿日期: 2023-03-13

基金项目: 国家重点研发计划(2023YFD1000400); 国家自然科学基金重大研究计划重点项目(91746202); 国家自然科学基金重点项目(71433006); 国家自然科学基金面上项目(62376249); 国家自然科学基金青年项目(61806177)

作者简介: 张寅升(1986-), 男, 博士, 副教授, 研究方向: 信息系统、机器学习和决策支持, E-mail: oo@zju.edu.cn

通讯作者: 王海燕(1968-), 女, 博士, 教授, 研究方向: 质量可靠性与质量控制理论研究, E-mail: njue2010@163.com

numerical features, fast Fourier transform, and discrete cosine transform were compared. Subsequently, partial least squares and visualization were used for separability analysis. Experimental results revealed that the fast Fourier transform, discrete cosine transform, and second-order one-dimensional Laplacian convolution kernels effectively improved separability relative to the original features. The R^2 effect size of partial least squares was increased from 0.61 to 0.95, 0.96, and 1.00. The one-dimensional Laplacian convolution kernel feature extraction method was able to accurately distinguish between domestically produced and foreign (Australia) base powder. In a case study, it achieved the best discrimination effect, indicating that it can effectively extract time response features of electronic nose channel sequence signals. This method can facilitate the differentiation work of dehydrated milk powder samples in China and Australia, providing technical support for the rapid identification of dehydrated milk powder sources.

Key words: electronic nose; milk powder; partial least squares; Laplacian convolution kernel

电子鼻是一种包括鉴别气味系统和电子化学传感器阵列的仪器。它的功能结构与人类的嗅觉系统类似,专门用来感应气味分子。其软件部分主要由特征提取和模式识别算法组成,能够处理传感器的特征响应,提取和选择有用的信息^[1]。由于电子鼻相较于人工,具有不会疲劳、不易出错、更为精细的优点,因此,自从上世纪九十年代开启相关研究后,电子鼻一直被当作人类嗅觉的延伸^[2]。随着相关技术的不断发展,电子鼻凭借简单易用、无需专业知识即可操作等优势,在环境监测、医学诊断、石油化工、农业等众多领域得到推广使用,尤其是在食品饮料行业,电子鼻技术被广泛用于完成质量控制和分类工作^[3]。在学术研究方面,对于电子鼻技术的应用探索也十分活跃,如鲍薪羽等^[4]使用电子鼻进行牛羊油的差异分析,以及郭淼等^[5]、何阳春等^[6]、Adak等^[7]都用电子鼻进行过研究。

近几年,电子鼻的应用在乳制品研究中成为热点课题^[8]。其中,有关分类及质量检测的研究较多^[9],如徐亚丹等^[10]使用电子鼻研究了伊利牛奶是否掺假,张雪茹等^[11]使用电子鼻检测了羊奶粉中的掺入的牛奶含量,刘立等^[12]使用电子鼻完成了国内外多个品牌奶粉的分类,而对乳制品原材料产地溯源的研究则相对较少。奶粉作为婴幼儿及老年人常食用的产品,同时也是部分乳制品的原材料,应用广泛,对于食品质量与安全的要求较高。而基粉是生产配方奶粉等次级产品的重要原料,它是由动物奶经风干等一系列处理后得到的,在很大程度上影响着终端产品质量,本文旨在探索一种能够区分不同产地基粉的分析方法。

本文受到了数字图像处理领域的二维拉普拉斯边缘检测算子的启发。该算子能够有效地提取像素

梯度和高频信息,在图像边缘检测中得到了广泛应用。考虑到电子鼻的各个通道数据是一种时间序列信号,其时间响应特性同样反映在差分或梯度等数字特征中,因此,本文尝试应用一维的拉普拉斯卷积核,以提取电子鼻各传感器的时间响应特征。在以往的研究中,电子鼻数据主要的特征提取方法有基于原始响应曲线的描述性统计量、基于曲线拟合的特征提取、基于变换域的特征提取^[13]等。本研究将探索有别于传统方法的一维卷积核特征提取方法,并对比评估其特征提取效果,同时为进一步研究提供了思路和技术支撑。

1 材料与方法

1.1 实验材料和样品预处理

取国产(来自国内某乳业公司的黑龙江牧场)奶粉基粉与进口(来自某国际奶企的澳洲墨尔本牧场)奶粉基粉,分别称量5 g奶粉,放入50 mL的离心管中,作为一个电子鼻样品。国产和进口奶粉分别准备四个平行的样本,每一个样本重复测试六次。样本采用专用封口膜密封,静置时间60 min。

实验步骤:

1) 开启电子鼻仪器,传感器阵列清洗20 min,进行预热;

2) 以奶粉样品为检测对象,用电子鼻取样针从离心管中抽取10 mL气体,缓慢打入电子鼻进样口中,进行检测;

3) 每次进样之前,对电子鼻传感器阵列进行300 s的清洗,设置气体流量为0.2 L/min,等待时间10 s,采样时间300 s,每种样品重复6次平行,且为了减小实验积累误差,所有样品以交叉方式进行实验。

1.2 设备介绍

FOX-4000 电子鼻，来自法国 Alpha MOS 公司。电子鼻传感器对应敏感气体如表 1 所示。

表 1 气敏传感器阵列及其属性

传感器序号	敏感气体
S1	氨气、胺类
S2	硫化氢、硫化物
S3	氢气
S4	酒精、有机溶剂
S5	甲苯、丙酮、乙醇、甲醛、氨气和其他有机蒸气
S6	甲烷、沼气、天然气
S7	甲烷、丙烷、异丁烷、天然气、液化气
S8	香烟的烟雾、烹调臭味、挥发性有机化合物 (Volatile Organic Compounds, VOC)、氨气、硫化氢、酒精
S9	丁烷、丙烷、甲烷、液化气、天然气、煤气
S10	液化石油气、可燃气体、丙烷、丁烷
S11	丙烷、烟雾、可燃气体
S12	一氧化碳、乙醇、有机溶剂、其他挥发性气体
S13	烟气、烹调臭味、氨气、一氧化碳、空气污染物
S14	甲烷、天然气

1.3 数据处理

本实验使用 Python 进行数据处理，首先在预处理过程中完成数据对齐和截取，而后使用快速傅里叶变换 (Fast Fourier Transform, FFT)、离散余弦变换 (Discrete Cosine Transform, DCT)、一维拉普拉斯卷积核 (One-dimensional Laplacian Convolution Kernel) 等五种不同的方法完成特征提取，随后使用偏最小二乘法回归 (Partial Least Squares Regression, PLSR) 进行降维，并用可分性得分及可视化等方法完成不同特征提取方法的可分性评估对比，最后得到实验结论。

1.4 实验流程图

实验的整体流程如图 1 所示。

2 数据分析

2.1 预处理

首先，实验对所有的样品进行电子鼻鉴定，获取各个样品的原始电子鼻数据。数据的表现形式为多个通道的传感器阵列所产生的一个时间响应曲

线，如图 2 所示。

由于送样时间不同，每一个样品在检测时所产生的数据在响应时间上也不是同步的，在电子鼻开始识别样本并产生有用信息的波形前，会有一个清洗阶段，这一期间产生的所有数据都是无效的，应当忽略，所以首先需要对这些数据进行时间轴对齐。

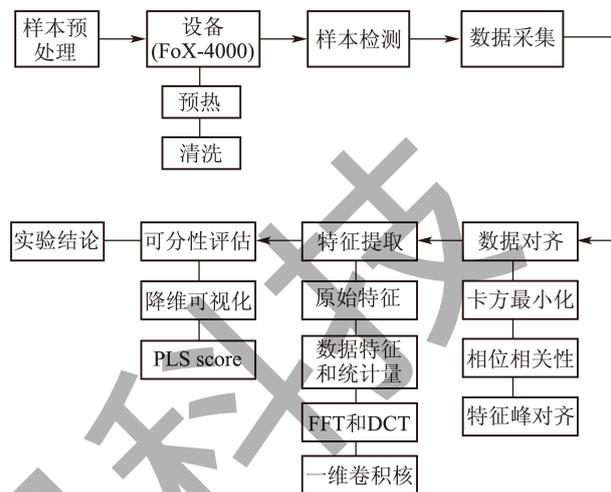


图 1 实验流程图

Fig.1 Experimental flow chart

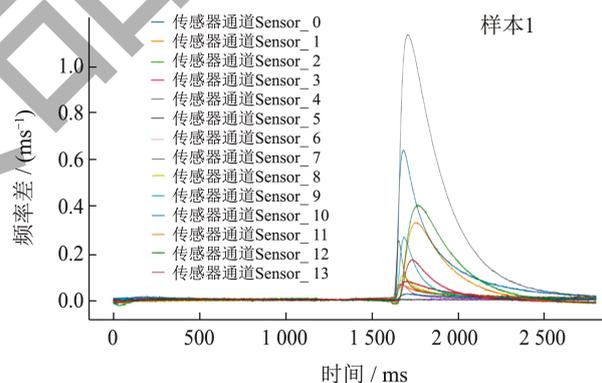


图 2 电子鼻时间响应曲线图

Fig.2 Time response curve of electronic nose

本次实验使用了三种对齐的算法，第一种是基于卡方最小化的算法，该方法使用重采样的方式移动信号，将移位后的信号与“模板”信号进行比较，以卡方最小化来作为标准，进而得出两者之间的最佳偏移。第二种是基于相位相关性的算法，其思想及操作与前一种方法类似，只是更换对齐标准为相位相关性。第三种是特征峰对齐，即使用特征峰的极值进行对齐，然后对比选择最优的对齐算法。因为电子鼻提供的每一个样品都有 14 个通道的数据，为了更好地完成对齐工作，使用每一组样品的平均来对齐。

获取平均值图像后，先确定首个样本作为参考

样本，再依次将其他样本与参考样本对齐，三种方法的对齐效果如图3所示。

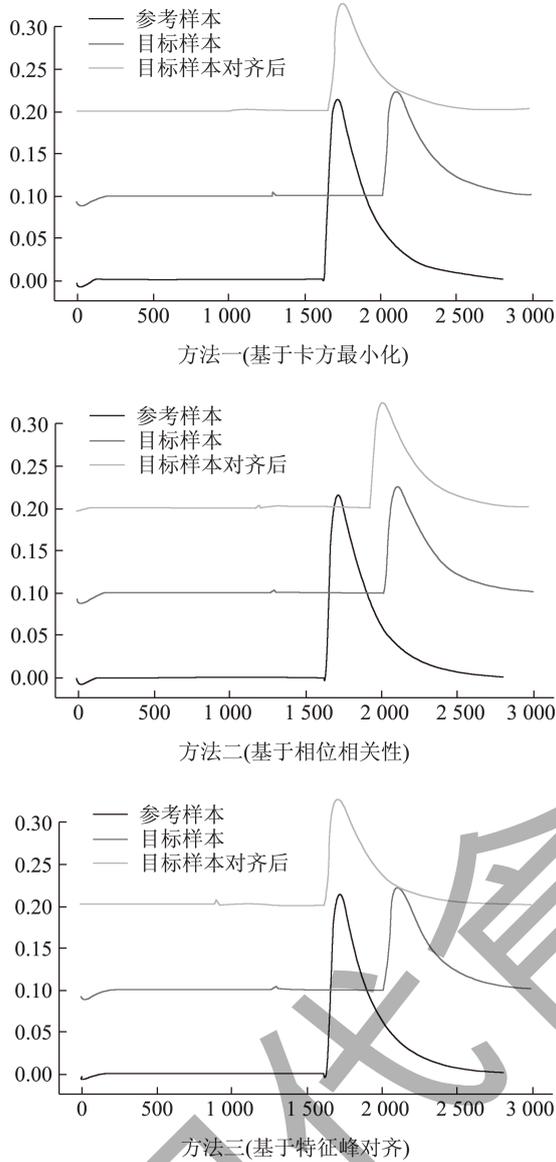


图3 样本对齐示意图

Fig.3 Schematic diagram of sample alignment

通过对下方坐标轴数值的观察，对比三种对齐方法的结果发现，第一种基于卡方最小化的对齐方法整体效果尚可，但在偏移量较大时效果较差。

而第二种基于相位相关性的方法的结果则相较第一种方法更不理想，对齐后仍与目标样本有加大的偏差。

第三种方法使用特征峰对齐的效果最好，与目标样本的偏差最小，同时该方法操作简单，无需额外的超参数调优。

因此，最终选择第三个方案进行对齐，完成后，将每一个样本都截取为相同长度，如图4所示。

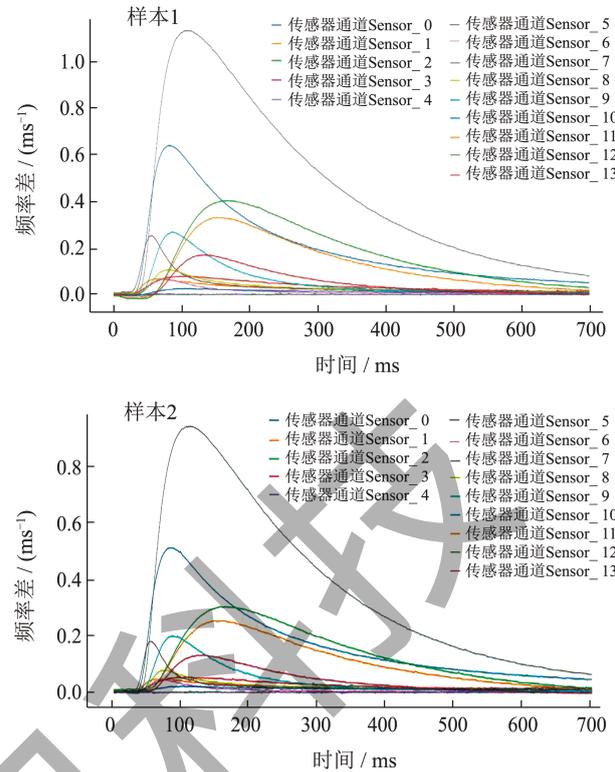


图4 预处理完成后的样本示意图

Fig.4 Samples after preprocessing

2.2 特征提取

由于电子鼻数据反映的是随时间变化的一个响应趋势，因此本实验更加关注的是曲线随时间变化的一些动态特征。根据电子鼻产生曲线的特点，本实验实现了多种特征提取的方法，并观察对比了其对于数据集可分性的影响。

2.2.1 原始特征

第一种方法是直接将每个样本中各个传感器通道的数据拼接成一个长的向量作为特征。这种方法相当于没有做任何的特征提取，而是直接使用原始的特征。该组作为对照组，用来与其他组做对比，以判断各组可分性的区别。

2.2.2 数学特征或统计量

第二种方法，我们对于初始特征进行一些简单处理，即提取图像中原始波形数据的一些数学特征或统计量，曲线最高峰的高度，一阶导数的AUC和最大值、最小值，二阶导数的AUC和最大值、最小值等。

2.2.3 傅里叶变换

第三种方法的基本思路是对各个通道原始的波形进行时域与频域的转化，即离散傅里叶变换（公

式 1)，但离散傅里叶变换计算量太大。为了加快计算，实践中使用快速傅里叶变换算法，即利用离散傅里叶变换的对称性与周期性，奇、偶、虚、实等特性快速实现变换的方法。

$$X(m) = \sum_{n=0}^{N-1} x(n)e^{-j2\pi nm/N} \quad (1)$$

式中：

$x(n)$ ——信号函数的采样值；

$X(m)$ ——信号函数在频域上的变换结果；

N ——信号函数的采样频率；

m ——离散傅里叶变换后的频率系数；

n ——信号函数在时域上每个采样点的标号。

2.2.4 离散余弦变换

第四种方法的思路与第三种方法类似，也是离散傅里叶变换的一种变种，即离散余弦变换算法，其原理是当一个函数为偶函数时，一般其傅里叶变换的虚部为零，因此无需计算，仅需计算余弦项变换，因此计算速度更快。经过 FFT 和 DCT 后的特征示意图如图 5 所示。

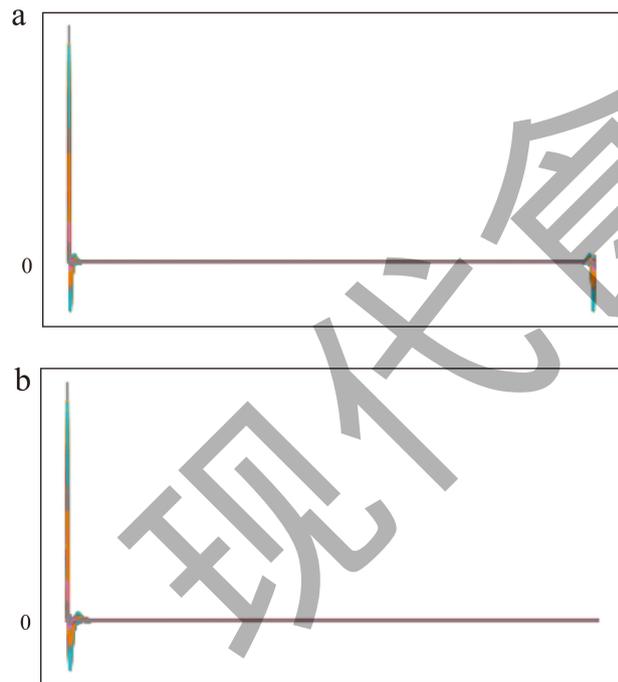


图 5 经过 FFT 及 DCT 的特征示意图

Fig.5 Schematic diagram of the characteristics of FFT(a) and DCT(b)

注：(a) FFT；(b) DCT。

对大部分信号而言，多数有用信息往往存在于低频部分。从图 5 中可以看出，低频之外的大部分频率包含的信息基本为 0，因此本次实验中仅使用

FFT 和 DCT 前端低频的数据，即可获取绝大多数的信息，完成特征提取。本次实验选取 FFT 频域中前 5% 的低频变量和 DCT 频域中前 10% 的低频成分作为提取的特征。

2.2.5 一维拉普拉斯卷积

除以上相关研究采用过的特征提取方法外，本研究还设计了一维卷积核特征提取方法。该方法采用一维拉普拉斯算子与原始信号进行卷积运算，进而得到卷积核提取后的新特征。本次实验采用的拉普拉斯卷积核在图像处理领域应用广泛，对应二阶导数，可检测出不同方向的边缘特征，并具有旋转不变性。其具体公式如下：

$$laplace(f) = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2} + \dots \quad (2)$$

表 2 不同阶数的后向差算子

Table 2 Backward difference operators of different orders

阶数	卷积算子	可视化
0 (原始数据)	[1]	
1	[-1,1]	
2	[1,-2,1]	
3	[-1,3,-3,1]	
4	[1,-4,6,-4,1]	
5	[-1,5,-10,10,-5,1]	
6	[-1,-6,15,-20,15,-6,1]	
7	[-1,7,-21,35,-35,21,-7,-1]	
8	[1,-8,28,-56,70,-56,28,-8,1]	

本研究采用的一维空间中的拉普拉斯核，同时也是一维差分算子的二阶特例。表 2 列出了不同阶数的后向差分算子 (Backward Difference)。

2.3 降维

经过特征提取后，进行数据的降维工作。目前大部分相关研究使用主成分分析法降维，但该方法不能利用目标变量 (因变量) 的信息，而且在本次实验过程中也发现，仅使用主成分分析进行降维效果一般。而偏最小二乘回归法所提取的成分既能很好地概括自变量系统中的信息，又能最好地解释因变量，并排除系统中的噪声干扰^[14]，其最终效果也优于主成分分析法。因此，本实验采用偏最小二乘回归法进行降维。

2.4 可分性计算及评估

最后进行不同特征提取方法的可分性评分与可视化分析，直观地对比各特征提取方法可分性的不同。本次实验的可分性得分标准使用了 PLS 的 R^2 效应量。 R^2 是表示拟合程度的一个指标，主要计算样本的残差。 R^2 的值越大，则拟合程度越高，在本次实验中，代表该特征提取方法的可分性越高，其公式具体为：

$$R^2 = \frac{SSR}{SST} = 1 - \frac{SSE}{SST} = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y})^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (3)$$

式中：

- SSR (Sum of Squared Regression) —— 回归平方和；
- SSE (Sum of Squared Error) —— 总平方和或总方差；
- SST (Sum of Squared Total) —— 残差平方和。

R^2 的最高得分是 1.00 分，说明特征变量的回归效应能够完整刻画分类标签，表示完全的可分性。当提取的特征可分性较差时，则可能会得到比原始特征更低的得分。

3 结果与讨论

3.1 实验结果

实验具体结果如下，经可分性打分后，各特征得分如表 3 所示。

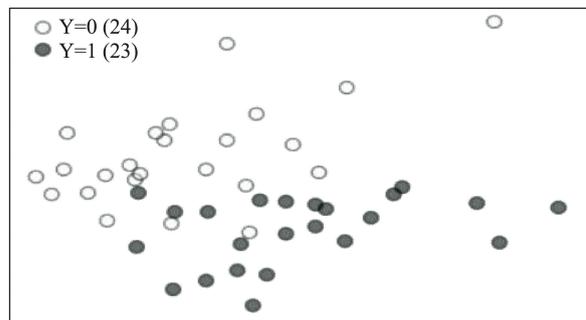
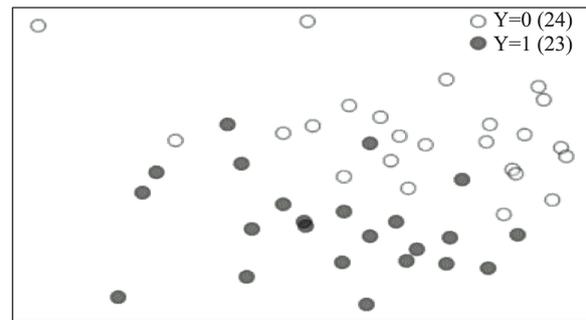
可分性可视化如图 6 所示。

观察实验结果发现，经过数据预处理后，仪器采集到的原始信息不具有可分性。而后应用偏最小二乘回归降维，通过比较可分性得分与可视化散点

图，发现原始特征及其一阶、二阶导数的基本统计量，包括求和、积分、最小值、最大值等，作为特征在可分性上有了改善。利用频域转换的快速傅里叶变换和离散余弦变换作为特征则得到了较强的可分性，而使用一维拉普拉斯卷积核提取到的特征具有最强的可分性。此外，本实验还对比了不同阶数的差分算子，结果表明二阶算子在 PLS 的 R^2 效应量上已经达到了 1.00 的最大值。

表 3 可分性得分表

特征	阶数	可分性得分
原始特征		0.61
数学特征及统计量	N/A	0.66
5% FFT		0.95
10% DCT		0.96
一维拉普拉斯卷积核	1	0.99
	2	1.00
	3	1.00
	4	1.00
	5	1.00
	6	0.93
	7	0.97
	8	1.0



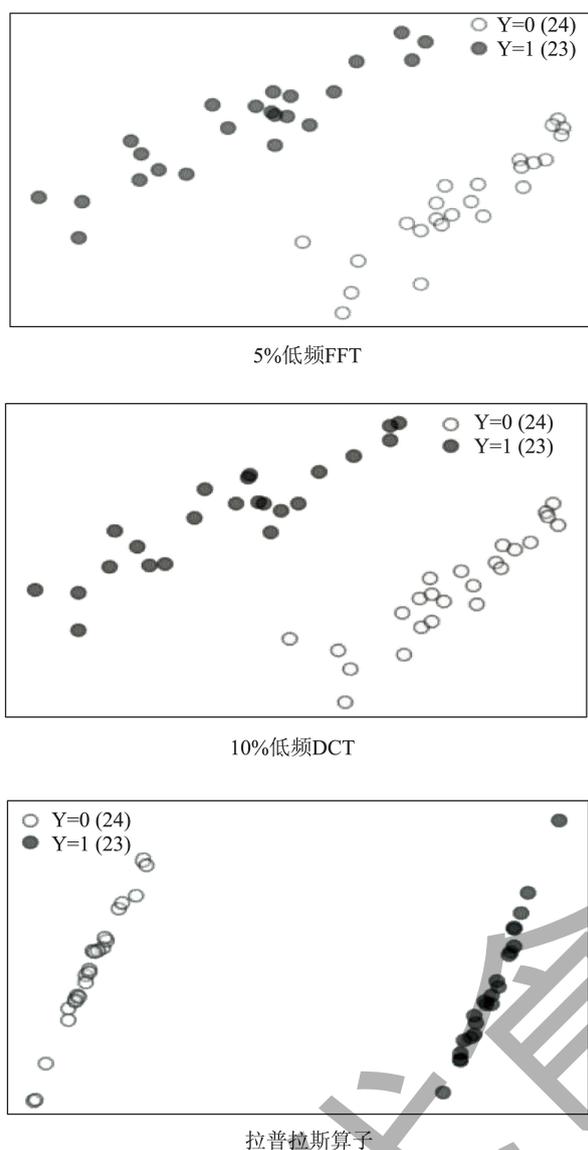


图6 数据集可分性的PLS散点图

Fig.6 PLS scatterplot of dataset separability

3.2 讨论

实验结果显示,快速傅里叶变换、离散余弦变换、一维拉普拉斯卷积算子作为特征的可分性都非常高,在线性分类器上基本上达到了95%、96%、100%的分类准确率,明显优于原始特征,能够完成中澳两国奶粉基粉样本的区分工作,为快速鉴定奶粉基粉来源提供技术支撑。其中一维的拉普拉斯卷积核作为特征,在测试集上达到了100%的分类准确率,说明其能够有效提取到电子鼻各通道序列信号的时间响应特征。

相关可复现代码已发布在Code Ocean (<https://doi.org/10.24433/CO.4661162.v1>)上,以便于其他

研究人员使用。

后续研究可以尝试将三种特征提取方法相结合,并尝试不同阶数的其他一维卷积算子,以适应其他食品及农产品对象的电子鼻检测数据。

参考文献

- [1] 唐辉,钟瑞敏,朱建华,等.电子鼻结合OAV分析臭豆腐卤水的挥发性成分及风味活性物质[J].中国食品学报,2021,21(3):351-359.
- [2] 谢晓铮,夏炎.电子鼻的核心——气体传感器的研究与应用[J].大学化学,2021,36(9):141-148.
- [3] 王建伟,叶升.电子鼻在食品行业的应用进展[J].中国调味品,2022,47(10):198-200,205.
- [4] 鲍薪羽,王丽金,宋焕禄,等.基于电子鼻和GC-O-MS技术分析牛油和羊油中香气成分的差异[J].中国食品学报,2022,22(12):267-281.
- [5] 郭淼,惠国华.基于电子鼻与非线性特征提取建模的猪肉鲜度快速分析方法[J].中国食品学报,2016,16(6):218-224.
- [6] 何阳春,陈丹妮,张千金,等.基于智能电子鼻的食醋识别方法的研究[J].中国食品学报,2012,12(5):168-173.
- [7] ADAK M, YUMUSAK N. Classification of E-nose aroma data of four fruit types by Abc-based neural network [J]. Sensors, 2016, 15(11): 27804-27831.
- [8] TOHIDI M, GHASEMI-VARNAMKHAISTI M, GHAFARRINIA V, et al. Development of a metal oxide semiconductor-based artificial nose as a fast, reliable and non-expensive analytical technique for aroma profiling of milk adulteration [J]. International Dairy Journal Volume, 2018, 77: 38-46.
- [9] 廖美燕,龙鸣,洪晶,等.电子鼻在乳及其制品检测中的应用研究进展[J].中国乳业,2022(10):57-64.
- [10] 徐亚丹,王俊,赵国军.基于电子鼻的对掺假的“伊利”牛奶的检测[C]//中国食品科学技术学会.中国食品科学技术学会第五届年会暨第四届东西方食品业高层论坛论文摘要集.浙江大学生物系统工程与食品科学学院;嘉兴职业技术学院;浙江发展园林有限公司,2007:1.
- [11] 张雪茹,付尚辰,郑卫民,等.商品羊奶粉中DNA的质量评价及牛源性成分掺假检测[J].食品与发酵工业,2022,48(18):249-255.
- [12] 刘立,李代禧,余华星,等.国内外五种著名全脂牛奶感官评价分析及其电子鼻、电子舌甄别初探[J].食品与发酵科技,2014,50(5):90-96.
- [13] YAN J, GUO X Z, DUAN S K, et al. Electronic nose feature extraction methods: a review [J]. Sensors 2015, 15: 27804-27831.
- [14] 王文胜,丁晶,赵玉龙.基于偏最小二乘回归的年用电量预测研究[J].中国电机工程学报,2003,10:17-21.