DOI: 10.19812/j.cnki.jfsq11-5956/ts.20251119005
引用格式: 高宏伟, 李瑞, 孙雯娴, 等. 便携式近红外光谱与集成学习法快速鉴别 8 种食用淀粉[J]. 食品安全质量检测学报, 2026, 17(10): 119–129.
GAO HW, LI R, SUN WX, *et al*. Rapid identification of eight edible starch species by portable near infrared spectroscopy and ensemble learning [J]. Journal of Food Safety & Quality, 2026, 17(10): 119–129. (in Chinese with English abstract).

便携式近红外光谱与集成学习法快速鉴别 8 种食用淀粉

高宏伟*, 李　瑞, 孙雯娴, 李明哲, 张　倩, 陈春光, 周龙龙
(青岛海关技术中心, 青岛　266109)

摘　要: 目的　构建基于便携式近红外光谱(1000~1800 nm)与集成学习算法的智能检测系统快速鉴别 8 种食用淀粉。**方法**　采集了玉米、马铃薯、木薯、小麦、红薯、绿豆、藕、豌豆等 8 类共 239 份市售淀粉样品, 获取 2392 条近红外光谱。采用标准正态变量变换(standard normal variate, SNV)和 Savitzky-Golay 一阶导数(窗口长度 11, 多项式阶数 2)进行预处理, 并直接在全谱范围内建模, 未进行特征波长筛选。**结果**　训练并对比了 K 近邻算法(K-nearest neighbor, KNN)、随机森林算法(random forest, RF)、偏最小二乘判别分析(partial least squares discriminant analysis, PLS-DA)、极端梯度提升库(eXtreme gradient boosting, XGBoost)及集成模型 5 种分类模型。在 59 份独立样品的验证中, 集成模型的准确率达 97.0%, 布里尔分数低至 0.008。**结论**　该系统单次预测耗时仅 0.8 s, 为淀粉掺假监测和物种溯源提供了一种高效、可靠的现场检测技术方案。

关键词: 近红外光谱法; 淀粉物种; 机器学习; 快速检测; 便携式设备

Rapid identification of 8 kinds of edible starch species by portable near infrared spectroscopy and ensemble learning

GAO Hong-Wei*, LI Rui, SUN Wen-Xian, LI Ming-Zhe, ZHANG Qian,
CHEN Chun-Guang, ZHOU Long-Long
(*Technology Center of Qingdao Customs*, *Qingdao* 266109, *China*)

ABSTRACT: Objective　To construct an intelligent detection system based on portable near infrared spectroscopy (1000–1800 nm) and ensemble learning algorithms for rapid identification of 8 kinds of edible starch. **Methods**　A total of 239 commercial starch samples from 8 botanical sources—corn, potato, cassava, wheat, sweet potato, mung bean, lotus root and pea—were collected, yielding 2392 near infrared spectroscopy spectra. The spectra were preprocessed by standard normal variate (SNV) and Savitzky-Golay first derivative (window length=11, polynomial order=2). And it was directly modeled across the full spectral range without conducting characteristic wavelength screening. **Results**　The 5 kinds of classification models—K-nearest neighbors (KNN), random forest (RF), partial

收稿日期: 2025-11-19
基金项目: 青岛海关科研项目(QK202502)
第一作者/*通信作者: 高宏伟(1975–), 女, 研究员, 主要研究方向为种源鉴定、基因检测和食品光谱分析。E-mail: ghw75@126.com

least squares discriminant analysis (PLS-DA), eXtreme gradient boosting (XGBoost), and a weighted soft-voting ensemble model—were trained and compared. In validation using an independent test set of 59 samples, the ensemble model achieved an accuracy of 97.0% with a Brier score as low as 0.008. **Conclusion** The system takes only 0.8 seconds for a single prediction, providing an efficient and reliable *on-site* detection technical solution for starch adulteration monitoring and species traceability.

KEY WORDS: near infrared spectroscopy; starch species; ensemble learning; rapid detection; portable device

0 引 言

淀粉作为自然界中储量最丰富的多糖之一，是食品工业中不可或缺的增稠剂、稳定剂和胶凝剂。其功能特性，如糊化温度、峰值黏度、回生性等，直接取决于其来源物种。例如，马铃薯淀粉具有高黏度和高透明度，适用于酱料和冷冻食品；木薯淀粉则以其良好的冻融稳定性和成膜性著称，常用于糖果和方便食品；而绿豆淀粉因其形成的凝胶强度高、弹性好，是制作粉丝、粉条的首选原料。因此，确保淀粉原料的物种真实性，不仅是企业保证产品质量稳定的关键，更是维护消费者知情权和打击市场掺假行为的核心。

目前针对淀粉物种鉴别的主流方法仍存在明显局限。传统的碘显色反应仅能反映直链/支链比例，无法实现物种区分；显微形态观察虽可识别部分特征颗粒，但主观性强、重复性差，且对加工后样品失效；而基于高效液相色谱法(high performance liquid chromatography, HPLC)、气相色谱-质谱法(gas chromatography-mass spectrometry, GC-MS)或 DNA 条形码等高精度技术虽具有良好准确性，却普遍存在设备昂贵、前处理烦琐、检测周期长等问题，难以满足口岸通关、生产线在线监控、电商仓储等场景快速响应的需求[1]。尤其在我国现行食品安全标准体系中，GB 31637—2021《食用淀粉》虽规定了理化与污染物指标，却未纳入物种真实性检测要求；市场监管部门在 2023 年淀粉专项抽检中仍主要依赖显微镜法或分子生物学手段，缺乏标准化、可推广的现场快检方案。这一监管技术空白，客观上降低了掺假成本，亟需发展一种兼具准确性、便捷性与经济性的新型鉴别技术。近红外光谱(near infrared spectroscopy, NIR)技术因其快速、无损、无需试剂、可原位测量等优势，近年来在食品成分分析与真伪鉴别领域展现出巨大潜力[2−3]。NIR 主要反映分子中 CH、OH、NH 等化学键的倍频与合频振动，这些振动模式与分子的化学环境和空间构型密切相关。不同物种的淀粉，其直链/支链比例、分子链长分布、颗粒表面化学基团等存在细微差异，这些差异在 NIR 上会形成独特的分子指纹图谱。通过结合化学计量学或机器学习算法，可以从这些复杂的光谱信息中提取出与物种相关的特征，从而实现自动化的物种判别。

尽管 NIR 技术潜力巨大，但在实际应用中仍面临诸多挑战[4]。不同淀粉的 NIR 在整体轮廓上高度相似，特征峰重叠严重，尤其是同科属的物种，如玉米与小麦，绿豆与豌豆，其光谱差异微弱，给高精度分类带来困难[5]。样品的颗粒度、填充密度、表面平整度等物理状态会引入显著的散射效应，掩盖了真实的化学信息[6]。多数研究在理想实验室条件下取得了高准确率，但当模型应用于不同批次、不同产地的样品时，性能往往大幅下降[7]。另外，现有研究多集中于台式仪器和实验室环境，缺乏面向实际应用场景的便携式、一体化解决方案[8]。当前国内相关研究多基于台式 NIR 仪器(波段多集中于 900~1700 nm)，样本规模普遍偏小(≤120 份)，且仅聚焦 2~4 种淀粉的二元或三元分类，难以覆盖市场主流品类[9−14]。准确率方面，实验室台式机可达 92%~98%，而便携设备报道的准确率多≤90%[15−18]。

针对上述瓶颈，本研究采用覆盖 1000~1800 nm 波段的便携式近红外光谱仪，结合集成机器学习算法，构建可同时鉴别马铃薯、木薯、玉米、小麦、绿豆、豌豆、红薯和藕粉等 8 种常见食用淀粉的高鲁棒性模型，通过数据增强与交叉验证提升模型对不同批次、产地及物理状态样品的适应性，为实现淀粉的质量监管提供了切实可行的技术路径。

1 材料与方法

1.1 材料与仪器

实验所用淀粉样品共计 239 份，涵盖 8 个物种。其中建立模型用样品 180 份，1802 条光谱(表 1,限于篇幅仅列出部分样品信息)，独立验证样品 59 份。样品购自国内市场、电商平台、实验室留存样品，品牌和产地覆盖广泛，生产日期跨度超过 2 年，详细样品信息见表 1。

SupNIR1500 便携式近红外分析仪[波长范围 1000~1800 nm，分辨率 1.5 nm，聚光科技(杭州)股份有限公司]；ME204E/02 分析天平[精度 0.1 mg，梅特勒-托利多仪器(上海)有限公司]；定制电动转盘(转速可调，淘工厂九九转盘)；聚苯乙烯培养皿(90 mm×15 mm，北京兰杰柯科技有限公司)。

表 1　食用淀粉样品总体信息表

Table 1　General information table of edible starch samples

物种	样品名称	品牌或生产厂家	购买日期	光谱采集日期	采集光谱数
玉米淀粉 (27 个样品, 270 条光谱)	0832A	河北石家庄(无品牌标识)	2025-07	2025-07-29	10
	17A01	烟台中显集团	2025-07	2025-07-29	9
	18A01	中英保健	2025-07	2025-07-29	10
	19A01	石家庄恒禄达食品有限公司	2025-07	2025-07-29	10
	2030A	海宁食品有限公司	2025-07	2025-07-30	10
	3768A	湖北誉承食品有限公司	2025-07	2025-07-30	10
	5204A	海天调味食品股份有限公司	2025-08	2025-08-21	10
	A1020	新良食品有限公司	2025-08	2025-08-21	10
	6779A	焙乖乖(电商品牌)	2025-08	2025-08-21	10
	A0720	金胜星食品有限公司	2025-07	2025-10-10	10
	A0820	赛卡伊(电商品牌)	2025-06	2025-10-10	10
	A0920	福临门食品股份有限公司	2025-08	2025-10-10	10
马铃薯淀粉 (19 个样品, 191 条光谱)	03A01	河南蜜丹儿食品有限公司	2025-07	2025-07-30	10
	3461A	湖北誉承食品有限公司	2025-07	2025-07-30	10
	A0100	浙江田灿科技有限公司	2025-07	2025-07-30	10
	A0300	河南博明食品有限公司	2025-07	2025-07-30	10
	A0400	山西繁峙县懿康食品有限公司	2025-07	2025-07-30	10
	A0500	甘肃定西(无品牌标识)	2025-08	2025-08-22	10
	A0620	姚源香纯食品有限公司	2025-08	2025-08-22	10
	A0720	古香园食品有限公司	2025-08	2025-08-22	10
	A0820	坝上伟伟纯食品有限公司	2025-08	2025-08-22	10
	A0920	海天代代田食品有限公司	2025-08	2025-08-22	10
	A1120	荷兰风车牌生粉	2025-08	2025-08-22	10
	A1220	上海枫未实业有限公司(展艺品牌)	2025-08	2025-08-22	10
小麦淀粉 (23 个样品, 131 条光谱)	02A01	盐城厨鹤粉业有限公司	2025-07	2025-07-30	10
	A0100	上海宝鼎酿造有限公司	2025-07	2025-07-30	10
	A0200	深圳太港食品有限公司	2025-07	2025-07-30	10
	A0300	河南蜜丹儿商贸有限公司	2025-07	2025-07-30	10
	A0400	风满坡食品有限公司	2025-07	2025-07-30	10
	A0500	河南博明食品有限公司	2025-07	2025-07-30	10
	A0600	白鹤牌澄面(广州市白鹤食品有限公司)	2025-08	2025-08-22	10
	A0720	丰联澄粉食品有限公司	2025-08	2025-08-22	10
	A0920	新乡市新良粮油加工有限责任公司	2025-08	2025-08-22	10
	A1020	变地金澄面(变地金食品有限公司)	2025-08	2025-08-22	10
	A1120	上海枫未实业有限公司(展艺品牌)	2025-08	2025-08-22	10
	A1220	广东白鲨食品有限公司	2025-08	2025-08-22	10
藕粉 (28 个样品, 279 条光谱)	藕粉 1	纯无糖(无品牌标识)	2025-07	2025-07-30	10
	藕粉 2	安徽乐饮食品有限公司	2025-07	2025-07-30	9
	藕粉 3	兴化绿帅食品有限公司	2025-07	2025-07-30	10
	藕粉 4	新郑腾发食品有限公司	2025-07	2025-07-30	10
	藕粉 5	杭州醇香园食品有限公司	2025-07	2025-07-30	10
	藕粉 6	智力西湖无糖纯(上海智力食品有限公司)	2025-08	2025-08-22	10
	藕粉 7	庆和原味纯(庆和食品有限公司)	2025-08	2025-08-22	10
	藕粉 8	正宗纯正古法手工(无品牌标识)	2025-08	2025-08-22	10
	藕粉 9	亿显堂食品有限公司	2025-08	2025-08-22	10
	藕粉 10	沁莲馨食品有限公司	2025-08	2025-08-22	10
	藕粉 11	扬州润莲食品有限公司	2025-09	2025-0910	10
	藕粉 12	庆和食品有限公司	2025-09	2025-10-10	10

表 1(续)

物种	样品名称	品牌或生产厂家	购买日期	光谱采集日期	采集光谱数
绿豆淀粉 (17 个样品, 171 条光谱)	A0100	晖雨青岛万家香食品有限公司	2025-07	2025-07-30	10
	A0200	河南蜜丹食品有限公司儿	2025-07	2025-07-30	10
	A0500	新乡良润全谷物食品有限公司	2025-07	2025-07-30	10
	A0720	北京华鑫都糖业	2025-07	2025-07-30	10
	A0820	青岛万家香食品有限公司	2025-07	2025-07-30	10
	A0920	烟台双塔食品股份有限公司	2025-08	2025-07-30	10
	A1020	丝宝宝农家特产	2025-08	2025-08-21	10
	A1220	冀飞食品股份有限公司	2025-08	2025-08-21	10
	A1320	河南蜜丹儿食品有限公司	2025-08	2025-08-21	10
	A1420	新乡新良粮油加工有限责任公司	2025-08	2025-08-21	10
	A1520	烟台双塔食品股份有限公司	2025-08	2025-08-21	10
红薯淀粉 (21 个样品, 210 条光谱)	4037A	京水湾食品有限公司	2025-07	2025-07-30	10
	5375k	河北内黄京统食品有限公司	2025-07	2025-07-30	10
	6928A	济宁泗水利丰食品有限公司	2025-07	2025-07-30	10
	7118A	广东燕保食品有限公司	2025-07	2025-07-30	10
	A0100	农家产品	2025-08	2025-08-22	10
	A0300	新乡新良粮油加工有限责任公司	2025-08	2025-08-22	10
	A0400	烟台双塔食品股份有限公司	2025-08	2025-08-22	10
	A0620	河北古松农副产品有限公司	2025-08	2025-08-22	10
	A1120	恒帆食品有限公司	2025-09	2025-10-10	10
	A1220	财成食品有限公司	2025-09	2025-10-10	10
	A5kua	农家块状红薯粉(无品牌标识)	2025-09	2025-10-10	10
	HENGF	恒丰食品有限公司	2025-09	2025-10-10	10
木薯淀粉 (26 个样品, 259 条光谱)	10356	湖北誉承食品有限公司	2025-07	2025-07-30	10
	10429	河南恒瑞食品有限公司	2025-07	2025-07-30	10
	3030A	鱼台食珍慧盈农业发展有限公司	2025-07	2025-07-30	8
	4568A	泰国进口水妈妈品牌	2025-08	2025-08-22	10
	5387A	百香匠发展有限公司	2025-08	2025-08-22	10
	7276A	河北古松农副产品有限公司	2025-08	2025-08-22	10
	7576A	钰涛食品有限公司	2025-08	2025-08-22	10
	8006A	越南(编号 2511-5387)	2025-04	2025-07-30	10
	A0100	老挝(编号 2511-7276)	2025-04	2025-09-10	10
	A0300	泰国(编号 2511-4159)	2025-03	2025-09-10	10
	A0920	柬埔寨(编号 2511-10429)	2025-06	2025-09-10	10
	A1020	老挝(编号 2511-3381)	2025-02	2025-09-10	10
豌豆淀粉 (19 个样品, 191 条光谱)	01A01	烟台双塔食品股份有限公司	2025-07	2025-07-30	10
	1937A	荆门京山市杰祥食品有限公司	2025-07	2025-07-30	9
	7132A	新乡良润全谷物食品有限公司	2025-07	2025-07-30	10
	A0300	风满坡食品有限公司	2025-08	2025-08-22	10
	A0500	焙乖乖食品有限公司	2025-08	2025-08-22	10
	A0720	盐城厨鹤粉业有限公司	2025-08	2025-08-22	10
	A0820	烟台双塔食品股份有限公司(双塔纯有机系列)	2025-08	2025-08-22	10
	A1020	银河食品有限公司(银河豌豆粉凉粉)	2025-08	2025-08-22	10
	A1120	上海枫未实业有限公司(展艺品牌)	2025-08	2025-08-22	10
	A1220	云山半食品有限公司	2025-08	2025-08-22	10
	A1320	加拿大(编号 2511-7132)	2025-04	2025-09-10	10
	01A07	吉匠食品有限公司	2025-04	2025-09-10	10

1.2　光谱采集

光谱采集时间为 2024 年 7 月、8 月、10 月、11 月, 实验操作由 4 人随机分配任务操作, 实验室温度 25 ℃±5 ℃, 湿度 60%±10%。取适量淀粉样品, 均匀铺于聚苯乙烯培养皿中, 使用不锈钢直尺沿皿口边缘轻轻刮平, 使样品表面平整且厚度一致(图 1)。将培养皿置于电动转盘上, 以 5 r/min 的转速旋转, 确保测量过程中样品受光均匀。使用固定在铁架台上的便携式 NIR 探头进行测量, 探头中心垂直对准培养皿中心, 距离样品表面约 1 cm。在自制暗箱内关闭环境光源, 以消除杂散光干扰。仪器预热 30 min 后, 使用标准白板进行背景校正。对每个样品连续采集 10 次光谱(每次扫描 10 次取平均为 1 条光谱), 以获得高信噪比的平均光谱。



图 1　聚苯乙烯培养皿中刮平表面的淀粉样品
Fig.1　Starch sample scraped from the surface of the polystyrene petri dish

1.3　数据处理与建模

1.3.1　光谱预处理

原始光谱数据常受到物理散射效应和随机噪声的干扰, 直接影响模型性能。本研究采用两步联合预处理策略: 首先采用标准正态变量变换(standard normal variate, SNV)消除因样品颗粒大小、密度差异引起的基线偏移和斜率变化[19–20]; 随后应用 Savitzky-Golay 一阶导数滤波(窗口长度 11, 多项式阶数 2)增强光谱特征分辨率并抑制低频漂移[21–22]。值得注意的是, 多元散射校正(multiplicative scatter correction, MSC)在初步实验中未带来显著性能提升, 故未纳入最终流程。该预处理方案有效提升了光谱信噪比, 为后续建模奠定基础。

1.3.2　特征波长选择

考虑到现代机器学习模型具备较强的抗噪能力和内置特征选择机制, 且全谱信息有助于保留细微判别性特征, 本研究在实验预分析中进行显式的特征波长筛选, 模型准确率不佳, 从而采用全谱建模策略。所有模型均基于完整的原始波长维度(共约 800 个波长点)进行训练, 避免因人为筛选导致的信息损失, 同时简化流程以利于实际部署。

1.3.3　模型构建与评估

所有化学计量学建模遵循 ASTM E1655 标准实践, 确保流程规范性。在全谱特征空间上, 采用五折分层交叉验证评估以下 4 种主流分类器性能: K 近邻(K-nearest neighbor, KNN)、随机森林(random forest, RF)、偏最小二乘判别分析(partial least squares discriminant analysis, PLS-DA)和非分布式梯度提升库(eXtreme gradient boosting, XGBoost)[23–24]。所有模型通过网格搜索进行超参数调优, 评价指标为平衡准确率[25–26]。最终构建集成模型(Ensemble), 根据多次实验预分析结果, 融合 KNN(权重 40%)、RF (30%)和 XGBoost (30%)的预测概率, 进一步提升泛化能力与鲁棒性。

为评估光谱数据集中人工标注标签的可靠性, 设计并实施了一套基于多模型高置信度预测的标签一致性审计方法。该方法的核心假设是, 当多个性能良好的分类模型对某一样本给出高度一致且置信度较高的预测结果, 而该结果与原始标注标签不一致时, 可能存在标签错误[27]。类似基于共识预测的策略在清洗光谱建模的错误标签数据中已被证明有效[28]。针对数据集中存在的类别分布不均衡问题, 本研究对比分析了重采样与代价敏感学习两种主流策略[29]。具体而言, 采用了经典的合成少数类过采样技术[30]以合成新的少数类样本实现训练集均衡化; 同时, 采用类别加权方法[29,31], 通过在损失函数中为少数类别赋予更高的权重来平衡不同类别对模型训练的贡献。

1.4　模型验证

为客观评估模型的泛化能力, 避免交叉验证可能带来的乐观偏差[32], 本研究遵循光谱建模的最佳实践[33], 设置了独立的验证环节, 将训练好的模型应用于该独立数据集, 记录预测结果, 以真实反映系统在实际应用中的性能。

本研究设置了独立的验证环节, 未参与任何预处理参数选择、特征筛选或模型训练。从实验室留存样品中, 另选取 59 份(590 条光谱)未参与任何建模过程的淀粉样品作为独立验证集(表 2)。所有光谱采集均在同 1 d 内完成, 避免日间仪器漂移影响。

1.5　数据处理

模型训练采用 5 折分层交叉验证评估模型性能, 超参数优化采用网格搜索结合平衡准确率作为评价指标。针对类别不平衡问题, 训练集通过合成少数类过采样技术进行 1 次数据增强(augmentation factor=2.0), 具体参数为: 噪声标准差 0.01、基线漂移 0.02、缩放因子 0.03。每个样品平行采集 10 条光谱(每次扫描 10 次取平均为 1 条光谱), 独立验证集采用投票法判定: 仅当某样品大于等于 8 条光谱给出同一物种鉴定结果时, 该样品方被记为相应物种。所有模型训练与验证重复运行 3 次以确保结果稳定性, 报告结果为 3 次运行的平均值±标准偏差。本研究所有数据处理、模型构建、统计分析及可视化均采用 Python 3.9 编程实现。主要依赖包括: Scikit-learn 1.3.0(机器学习模型与交叉验证)、XGBoost 2.0.0(梯度提升算法)、TensorFlow 2.13.0 (PLS-DA 神经网络后端)、Pandas 2.0.3(数据处理)、NumPy 1.24.3(数值计算)、SciPy 1.11.2(信号处理与统计检验)、Matplotlib 3.7.2 与 Seaborn 0.12.2(图表绘制)。光谱预处理

表 2 部分样品独立验证样品信息

Table 2 Information of independent verification some samples

物种	样品名称	品牌/来源	购买/送检日期	光谱采集日期	采集光谱数
玉米淀粉	玉米 1 3678A	河北古松农副产品有限公司(古松牌, 带嘴包装)	2025-08	2025-10-10	10
	玉米 2	青岛万家香食品有限公司	2025-08	2025-10-10	10
	玉米 3	进口(编号 5412)	2024-11	2025-10-10	10
马铃薯淀粉	马铃薯 1	进口(编号 3461)	2025-06	2025-10-10	10
	马铃薯 2	古香园食品有限公司	2025-09	2025-10-11	10
	马铃薯 3	尚品食品有限公司	2025-09	2025-10-11	10
小麦淀粉	小麦 1	新乡市新良粮油加工有限责任公司	2025-08	2025-08-22	10
	小麦 2	上海枫未实业有限公司(展艺品牌)	2025-09	2025-10-11	10
	小麦 3	志龙凯食品有限公司	2025-09	2025-10-11	10
藕粉	藕粉 1	福东盛食品有限公司	2025-09	2025-10-10	10
	藕粉 2	江苏宝应(产地, 无品牌标识)	2025-03	2025-10-10	10
	藕粉 3	杭州三家村藕粉有限公司	2024-11	2025-10-10	10
绿豆淀粉	绿豆 1	美国进口(编号 2512-1937)	2025-02	2025-10-09	10
	绿豆 2	河北古松农副产品有限公司	2025-09	2025-10-11	10
	绿豆 3	晖雨食品有限公司	2025-09	2025-10-11	10
红薯淀粉	红薯 1	菱花集团有限公司(菱花牌)	2025-09	2025-10-10	10
	红薯 2	国产(编号 4037)	2025-04	2025-10-10	10
	红薯 3	国产(编号 5375)	2025-04	2025-10-10	10
木薯淀粉	木薯 1	柬埔寨(编号 2511-8006)	2025-05	2025-10-09	10
	木薯 2	越南(编号 2511-7576)	2025-07	2025-10-09	10
	木薯 3	泰国(编号 2511-4157)	2025-03	2025-10-10	10
豌豆淀粉	豌豆 1	云半山食品有限公司	2025-09	2025-10-11	10
	豌豆 2	烟台双塔食品股份有限公司	2025-09	2025-10-11	10
	豌豆 3	新乡市新良粮油加工有限责任公司	2025-09	2025-10-11	10

(SNV、Savitzky-Golay 一阶导数)及数据增强采用自研 Python 脚本(classification_2.py)实现, Hotelling T^2 统计量计算基于主成分分析(principal component analysis, PCA)得分矩阵通过 SciPy 库实现(post_train_2.py)。模型预测与批量处理采用 prediction_2.py 脚本完成。所有计算在 Windows 10 操作系统(64 位)上完成, 处理器为 Intel Core i7, 内存 16 GB。

2 结果与分析

2.1 NIR 光谱特征分析

8 种食用淀粉的原始近红外光谱显示典型的碳水化合物特征吸收。不同物种淀粉的光谱在整体形状上相似, 但在吸收强度、峰宽和细微峰位上存在差异, 这些差异构成了物种鉴别的光谱基础。如图 2, 在 1190 nm(约 8400 cm^{-1})处观察到 O-H 键第二倍频吸收, 主要与结晶区结合水相关; 1450 nm (约 6900 cm^{-1})处为最敏感的 O-H 键第一倍频吸收, 反映游离水和结合水含量; 1580 nm 和 1760 nm 区域归属为吡喃环 C-H 键的第一倍频及组合频吸收[25,34]。1450 nm 处的吸收峰归属于葡萄糖单元中 C-H 键的弯曲振动, 其强度与淀粉分子中支链结构的数量及分布密切相关。由于支链淀粉含有大量 α-1,6 糖苷键分支, 每个分支点引入额外的亚甲基(-CH_2-)和羟基(-OH), 导致该波段吸收增强, 因此该峰可作为反映支链淀粉相对含量的间接指标。1450 nm 归属于葡萄糖单元中-OH 基团的第一倍频[13]。支链淀粉具有大量 α-1-6 分支, 每个分支点额外产生 1 个非氢键结合-OH 和 2 个处于半无定形区的-CH_2OH; 这些–OH 的氢键网络有序度低、键长分布宽, 导致 1450 nm 吸收峰面积(A_{1450})明显高于直链淀粉[35]。1580 nm 与 1760 nm 归属吡喃环/葡萄糖环 C–H 第一倍频及组合频, 波长 1690 nm 为 O-H 键组合频氢键网络强度吸收峰[36]。不同物种淀粉的直链/支链比例差异导致 1450 nm 处吸收特征不同[35], 这是实现物种鉴别的光谱学基础。

2.2 分类模型性能比较

5 种分类模型在内部测试集上的性能对比如表 3 所示, KNN、RF、XGBoost 和 Ensemble 的综合性能最为优异。KNN 模型的准确率达到 96.7%, 布里尔分数为 0.007, 是所有模型中最低的, 表明其预测概率具有高度可靠性; Ensemble 模型准确率为 97.0%, 布里尔分数为 0.008, 表现稳定且优于多数单一模型。RF 和 XGBoost 模型准确率分

别为 96.4%和 94.8%，布里尔分数较低，展现出较强的泛化能力。PLS-DA 模型准确率仅为 40.0%，布里尔分数高达 0.106，表明该线性判别模型难以有效捕捉本研究中复杂的非线性物种间光谱差异。整体来看，基于树结构的非线性模型(RF、XGBoost)以及集成策略在高维光谱分类任务中表现出显著优势。所有模型均在全谱特征空间上训练，未进行特征降维，确保了信息完整性。经标签审计，KNN、RF、XGBoost、Ensemble 在 8 类别下进行高置信度预测，未发现预测与标签不一致的样本，标签质量良好。

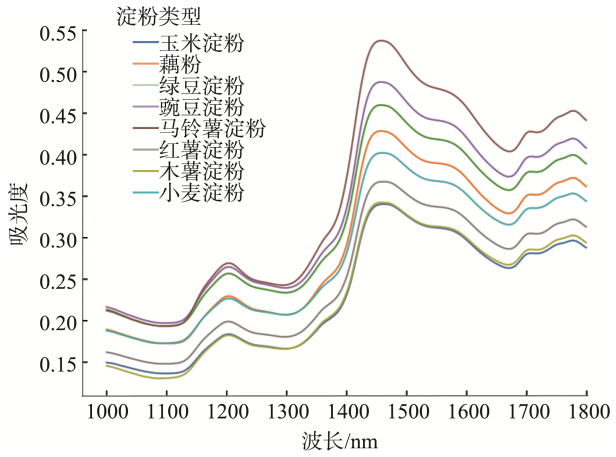


图 2　8 个物种淀粉样品的 NIR 平均曲线

Fig.2　Average curves of NIR for starch samples of 8 species

表 3　5 种算法的内部交叉验证或测试集结果

Table 3　Results of internal cross-validation or test set for the 5 kinds of algorithms

模型名称	准确率/%	布里尔分数	Cohen's κ	MCC
KNN	96.7	0.007	0.962	0.962
RF	96.4	0.015	0.959	0.959
PLS-DA	40.0	0.106	0.297	0.319
XGBoost	94.8	0.012	0.940	0.940
Ensemble	97.0	0.008	0.965	0.966

在内部 5 折交叉验证测试集上，Ensemble 的准确率达到 97.0%，Brier 分数为 0.008(表 3)。此处结果为模型训练过程中的内部验证表现，采用分层抽样确保各类别比例均衡。

进一步分析各类别的识别情况(表 4)，KNN、RF、XGBoost 与 Ensemble 在多数类别上表现出较高的召回率与 F1 分数。其中，玉米、小麦、红薯、木薯、马铃薯和豌豆淀粉的识别性能较为优异，各模型在这些类别的召回率普遍高于 0.92，F1 分数多在 0.94 以上；绿豆和豌豆在 KNN 模型下实现完美识别(召回率=F1=100.0%)，显示出极强的判别能力。藕粉和马铃薯的识别性能相对较低，马铃薯在 XGBoost 模型下的召回率仅为 86.8%，F1 为 91.7%，表明其光谱特征与其他类别存在重叠或变异性较大。集成模型通过加权软投票策略有效融合了基模型的优势，在所有类别上均实现了稳定且均衡的性能，在马铃薯和藕粉等分类性能较差的类别中，集成模型的 F1 分数显著优于各单一模型，表明其在处理类间特征相似性方面更具鲁棒性。

概率校准结果显示，经等渗回归校准后，KNN、RF 和 XGBoost 模型的预测概率可靠性显著提升。Ensemble 在此基础上进一步优化，实现了最佳的概率校准效果。逐类评估显示，豌豆淀粉实现了完美的识别效果(召回率=100.0%，F1=98.7%)；玉米、小麦和木薯淀粉的召回率与 F1 分数均高于 95%；藕粉和红薯虽光谱特征相对复杂，其召回率仍分别达到 94.7%和 97.6%，F1 分数分别为 93.0%和 97.6%，表明模型对各类别均具备良好的判别能力。集成模型通过加权软投票策略[37]，在保持高灵敏度的同时有效提升了分类稳定性，整体性能满足生产线原料验收与市场监管场景现场快速鉴别的精度要求。

2.3　降维可视化分析

PCA 为线性降维方法，其可视化结果表明前两个主成分能够捕捉到数据的主要变异方向，各类别在二维空间中呈现显著重叠，以玉米、小麦和豌豆淀粉为代表，在中心区域高度混杂，反映出线性方法难以有效分离这些光谱

表 4　KNN、RF、XGBoost 与软投票 Ensemble 的 Cohen's κ、MCC 及逐类召回率-F1 明细(%)

Table 4　Cohen's κ, MCC and per-class recall-F1 details of KNN, RF, XGBoost and soft-voting Ensemble (%)

分类	KNN		RF		XGBoost		Ensemble	
	召回率	F1	召回率	F1	召回率	F1	召回率	F1
玉米淀粉	98.3	99.1	100.0	99.1	100.0	96.7	100.0	100.0
藕粉	93.0	95.5	94.7	97.3	94.7	95.6	94.7	96.4
绿豆淀粉	100.0	100.0	90.6	95.1	87.5	93.3	96.9	98.4
豌豆淀粉	100.0	100.0	97.4	94.9	94.7	93.5	100.0	98.7
马铃薯淀粉	92.1	95.9	86.8	93.0	86.8	91.7	89.5	94.4
红薯淀粉	97.6	93.2	100.0	93.3	90.5	89.4	97.6	93.2
木薯淀粉	96.3	94.5	98.1	96.4	100.0	96.4	96.3	95.4
小麦淀粉	97.8	96.8	100.0	100.0	97.8	98.9	100.0	98.9

特征相近的物种。*t*-分布随机邻域嵌入(t-distributed stochastic neighbor embedding, t-SNE)为非线性降维技术，能更好地保留高维数据的局部结构关系。在 *t*-SNE 图中，多数类别形成了相对独立或半独立簇群。木薯淀粉在右下角形成一个紧凑且远离其他类别的簇；马铃薯与红薯淀粉虽存在一定交叉，仍可初步区分；绿豆与豌豆淀粉尽管同属豆科，在空间中展现出清晰的分离趋势，表明所选特征具备较强的判别能力。结合模型训练过程中的预测一致性分析，提示训练集标签质量良好，为后续模型构建提供了可靠的数据基础(图 3)。

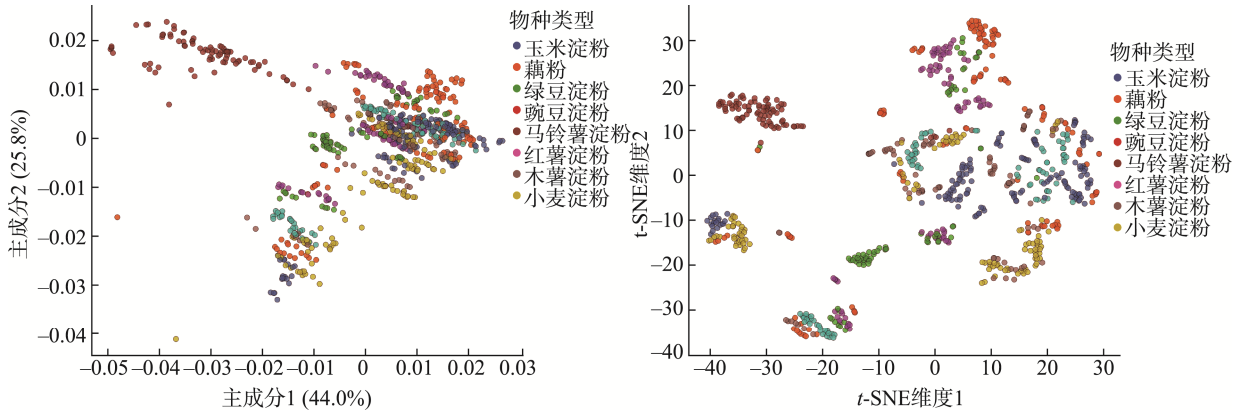


图 3 基于 PCA 和 *t*-SNE 的淀粉来源分类

Fig.3 Starch source classification based on PCA and *t*-SNE

基于 PCA 的特征降维结果显示(图 4)，前两个主成分(PC1 与 PC2)的累计方差贡献率为 69.3%，能够有效捕捉数据的主要变异结构。图 4 中所示的 95% Hotelling T^2 置信椭圆标识了训练集样本在主成分空间中的正常分布范围。从空间分布来看，各类淀粉样本整体呈现一定的聚集趋势，表明其光谱特征具备一定的类间可分性。大多数样本位于椭圆内部，说明模型对主流样本具有良好的拟合能力。然而，玉米淀粉(玉米)与豌豆淀粉(豌豆)在 PC1 轴上高度重叠，难以有效区分；木薯淀粉(木薯)与红薯淀粉(红薯)在二维平面上也存在显著重叠，反映出二者光谱相似性较高。值得注意的是，部分测试样本(灰色点)分布在椭圆外部，提示这些样本可能具有较高的化学或光谱差异性，属于潜在的离群点或边界类别，值得进一步关注。

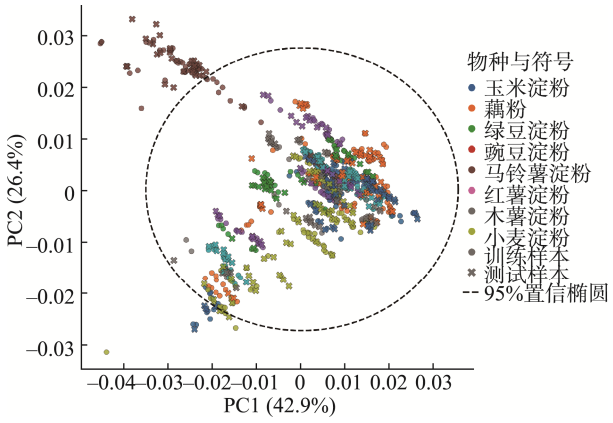


图 4 训练集 PCA 得分与 Hotelling T^2 95%置信椭圆

Fig.4 PCA scores of training set and Hotelling T^2 95% confidence ellipse

软投票 Ensemble 在独立测试集上的归一化混淆矩阵(图 5)显示，模型对玉米、小麦的预测完全正确，召回率达 100.0%；木薯仅有 2 个样本被误判为红薯，召回率为 98.2%；藕粉有 3 个样本被误判为木薯，召回率为 92.9%。绿豆有 1 个样本被误判为豌豆，马铃薯与红薯之间也存在一定混淆。这些误判主要集中在生态特征或光谱特性相近的物种之间，反映此类别间的天然可分性限制。所有样本均未出现跨大类，如豆科与根茎类、谷物与块茎类的混淆，表明集成策略有效抑制过拟合现象，强了模型的泛化性能。

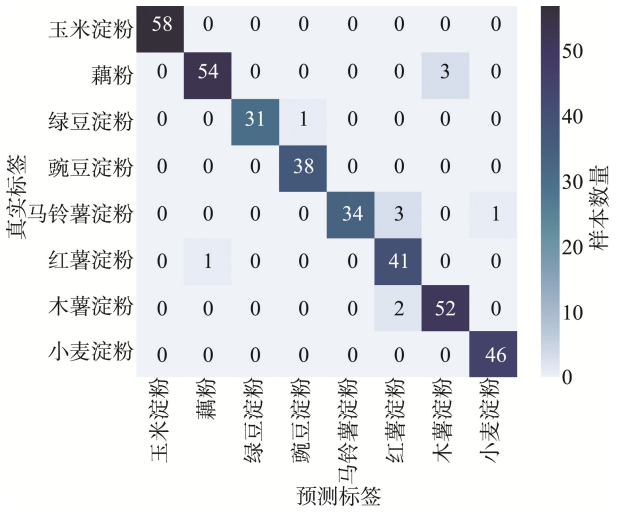


图 5 Ensemble 在测试集上的归一化混淆矩阵

Fig.5 Normalized confusion matrix of Ensemble model on the test set

2.4　完全独立验证

完全独立测试集含 59 个生物学样品, 每类 7~8 个。每个样品采集 10 条平行光谱, 仅当大于等于 8 条光谱给出同一物种鉴定结果时, 该样品方被记为相应物种, 否则标记为不通过, 需进一步确证检测(如 DNA 检测方法)。如表 5 所示, 集成模型在该独立验证集上的总体准确率达到 98.5% (581/590), 较内部验证集(97.0%)略有提升/波动, 表明模型具有良好的泛化能力。从逐类分析可见, 玉米、马铃薯、豌豆和小麦等类别在多数模型中均实现接近或达到 100%的准确率, 表明其光谱特征具有高度可分性; 藕粉与木薯也表现良好, 尤其在 RF 和集成模型下实现完美识别; Ensemble 在保持高精度的同时, 对难分类类别(如红薯)的识别性能仍优于单一模型, 进一步验证了软投票策略在提升模型鲁棒性和抗干扰能力方面的有效性。红薯是唯一存在明显误判的类别, 各模型准确率在 82.9%~90.0%之间。需要注意的是, 光谱采集时发现, 各个品牌的红薯淀粉在颗粒度和颜色外观上呈现较大的差异性, 多个淀粉样品光谱采集的表面无法实现光滑平整, 干扰其光谱的物种来源识别。此外, 不同品牌红薯淀粉中残留的微量蛋白质、脂质等非淀粉成分的差异, 也可能引入额外的光谱变异, 共同构成了该类别识别的难点。

2.5　系统创新与应用价值

本研究在 3 个方面实现了创新: 首次应用便携式近红外光谱仪(1000~1800 nm)对 8 种食用淀粉进行同时鉴别。通过采集来自玉米、马铃薯、木薯、小麦、红薯、绿豆、藕粉、豌豆等 8 种不同来源的市售淀粉样品的 NIR 数据, 实现了快速、无损的物种识别。本研究提出了一种基于 SNV 与 Savitzky-Golay 一阶导数的预处理方案, 并直接在全谱范围内建模。构建一个概率校准软投票集成模型, 在保持高准确率的同时确保了预测概率的高度可靠性。该模型通过结合 KNN、RF、XGBoost 等多种分类算法的优点, 利用加权软投票策略进一步提升了分类性能, 并通过等渗回归优化了概率输出, 使得最终模型不仅在准确性上表现出色, 而且其输出的概率值也具有很高的可信度。

尽管部分已有研究报告了较高的准确率(表 6), 但本研究在 3 个关键维度上实现了突破: 鉴别的淀粉物种数量达到 8 种, 远超多数同类研究; 全部 239 份样品均为真实市售产品, 而非实验室纯品, 更能反映实际应用场景的复杂性; 本研究不仅关注分类准确率, 更通过极低的布里尔分数(0.008)证明了模型预测概率的高度可靠性。另外, 本研究基于两点考虑未选择 SVM 模型。首先在实验数据的预分析中多次建立的 SVM 模型预测准确率均显著低于其他几个模型, 另外 SVM 的概率输出需要通过额外校准, 与构建一个高可靠概率预测系统的目标不符。而 KNN、RF、XGBoost 等模型能直接或更自然地输出校准后概率。其技术框架(便携式设备结合机器学习)亦可推广至其他食品品质指标的现场检测。近期研究表明, 类似技术路线已成功应用于鸡蛋中功能脂肪酸[37]等指标的快速测定, 验证了本方法学的普适性。

本研究所构建的 NIR 快速检测系统不仅在鉴别淀粉物种的数量上有了显著提升, 具有重要的推广应用价值, 但在更广泛的现场部署中仍面临若干挑战, 如环境因素波动、样品物理状态多样性以及长期运行中可能出现的概念漂移等。未来工作将探索集成环境传感器进行动态补偿, 开发在线学习模块支持模型增量更新, 建立可疑样本自动标记、人工复核、模型再训练的闭环反馈机制, 以实现系统的持续进化与自适应能力。

表 5　独立验证结果(%)
Table 5　Results of independent validation (%)

类别	KNN	RF	XGBoost	Ensemble
玉米淀粉	100.0 (80/80)	100.0 (80/80)	100.0 (80/80)	100.0 (80/80)
藕粉	98.8 (79/80)	100.0 (80/80)	100.0 (80/80)	100.0 (80/80)
绿豆淀粉	100.0 (70/70)	97.1 (68/70)	94.3 (66/70)	100.0 (70/70)
豌豆淀粉	100.0 (70/70)	98.6 (69/70)	97.1 (68/70)	100.0 (70/70)
马铃薯淀粉	100.0 (70/70)	100.0 (70/70)	100.0 (70/70)	100.0 (70/70)
红薯淀粉	82.9 (58/70)	90.0 (63/70)	88.6 (62/70)	88.6 (62/70)
木薯淀粉	92.5 (74/80)	100.0 (80/80)	100.0 (80/80)	97.5 (78/80)
小麦淀粉	100.0 (70/70)	100.0 (70/70)	97.1 (68/70)	100.0 (70/70)
总体准确率	97.0 (572/590)	98.5 (581/590)	97.5 (575/590)	98.5 (581/590)

注: 括号内数字表示正确识别的光谱数与该类别总采集光谱数之比, 即正确识别数/总测试数。所有结果基于独立测试集评估, 未参与模型训练。

表 6 国内 NIR 快检食用淀粉研究对比(2021—2024)

Table 6 Comparison of domestic NIR rapid detection studies on edible starch (2021—2024)

波长范围/nm	样本量	物种数	最佳模型	准确率/%	设备类型	文献
900~1700	96	3	PLS-DA	94.2	台式	[5]
1000~2500	120	4	SVM	96.5	台式	[7]
900~1700	80	2	RF	92.0	便携	[8]
1000~1800	110	3	XGBoost	90.1	便携	[10]
900~1700	108	4	CNN	97.3	台式	[11]
1000~2500	90	2	PLS-DA	93.0	台式	[13]
1000~1700	240	6	SVM	98.0	便携	[34]
1000~1800	239	8	Ensemble	97.0	便携	本研究

注: 支持向量机(support vector machine, SVM)。

3 结 论

本研究成功开发了一套基于便携式 NIR (1000~1800 nm)与机器学习的食用淀粉物种快速识别系统。该系统采用 SNV 结合 Savitzky-Golay 一阶导数进行光谱预处理, 并在全谱范围内直接建模, 完整保留原始光谱判别信息。在多种机器学习算法的对比中, Ensemble 表现最优, 准确率为 97.0%, 布里尔分数低至 0.008。在独立验证集(59 份样品, 590 个平行样)上达到 98.5%的准确率, 表明其不仅分类精度高, 且输出概率高度可靠。系统单次预测耗时仅 0.8 s, 显著优于传统方法, 满足现场快检的时效性要求。未来通过持续优化和扩展实验样品, 该系统有望在更广泛的食品真实性鉴别领域发挥重要作用。

参考文献

[1] DAVID CR, IKER G, SILVIA S, *et al*. Detection of nutrients and contaminants in the agri-food industry evaluating the probabilities of false compliance and false non-compliance through PLS models and NIR spectroscopy [J]. Applied Sciences, 2025, 15(9): 4808–4829.

[2] ZHI CL, XIAO HW, JIA JD, *et al*. Deep learning-enhanced spectroscopic technologies for food quality assessment: Convergence and emerging frontiers [J]. Foods, 2025, 14(13): 1–3.

[3] 彭彬倩, 沈福苗. 近红外光谱技术在食品检测中的应用研究进展[J]. 食品安全导刊, 2025(5): 181–184.
PENG BQ, SHEN FM. Research progress on the application of near-infrared spectroscopy technology in food detection [J]. Journal of Food Safety Guide, 2025(5): 181–184.

[4] QI WL, JIANG QQ, ERI HM. Research progress on the application of near-infrared spectroscopy in liquid food quality testing [J]. Beni-Suef University Journal of Basic and Applied Sciences, 2025, 14: 75–90.

[5] PELLACANI S, BORSARI M, COCCHI M, *et al*. Near-infrared and UV-Visible spectroscopy coupled with chemometrics for the characterization of flours from different starch origins [J]. Chemosensors, 2024, 12(1): 1–12.

[6] WU ZS, DU M, SHI XY, *et al*. Robust PLS prediction model for Saikosaponin A in *Bupleurum chinense* DC. coupled with granularity-hybrid calibration set [J]. Journal of Applied Mathematics and Computation, 2015, 271: 218–228.

[7] SILVA CAAC, RIZZO R, SILVA MA, *et al*. Spatio-temporal generalization of VIS-NIR-SWIR spectral models for nitrogen prediction in sugarcane leaves [J]. Remote Sensing, 2024, 16(22): 4250.

[8] ROSSI FAB, TONIAL LMDS. Portable vs benchtop NIR spectrometers for rapid chemical analysis: A study with different samples [M]. Analytical Chemistry: New Developments. Científica Digital, 2025.

[9] TEMIZ HT, ULA SB. A review of recent studies employing hyperspectral imaging for the determination of food adulteration [J]. Photochem, 2021, 1: 125–146.

[10] ALTIERI G, LAVEGLIA S, RASHVAND M, *et al*. Portable NIR spectroscopy combined with machine learning for kiwi ripeness classification: An approach to precision farming [J]. Applied Sciences, 2025, 15(11): 62–79.

[11] WILLIAM V, REBECA SR, GRIMALDO QS, *et al*. Detection of adulterants in powdered foods using near-infrared spectroscopy and chemometrics: Recent advances, challenges, and future perspectives [J]. Foods, 2025, 14(18): 31–47.

[12] TONG PJ, YANG DF. A near-infrared spectroscopy method for detecting corn starch content based on UVE-LightGBM feature selection [J]. Journal of Food Composition and Analysis, 2025, 148: 125–136.

[13] SAMUELE P, MARCO B, MARINA C, *et al*. Near-infrared and UV-visible spectroscopy coupled with chemometrics for the characterization of flours from different starch origins [J]. Chemosensors, 2024, 12(1): 1–12.

[14] CHEN ZH, ZHOU RG, PENG J. Pectraformer: Deep learning model for grain spectral qualitative analysis based on transformer structure [J]. RSC Advances, 2024, 14: 8053–8066.

[15] GUO X, HUANG HY, WANG HY, *et al*. Near-infrared spectroscopy and machine learning for fast quality prediction of bottle gourd [J]. Foods, 2025, 14(14): 2503.

[16] ZOU X, WANG QY, CHEN YJ, *et al*. Fusion of convolutional neural network with XGBoost feature extraction for predicting multi-constituents in corn using near infrared spectroscopy [J]. Food Chemistry, 2025, 15: 463.

[17] ARDIANA T, MAJLINDA S, MARJETA D, *et al*. Applications of

machine learning models in Near-infrared spectroscopy for small-grain quality control [J]. Asian Journal of Dairy and Food Research, 2025, 1: 1–9.

[18] CAO ZHH, WU XH, ZHANG ZX, *et al*. Combining UDT with XGBoost to identify the geographical origin of black beans by near-infrared spectroscopy [J]. Current Research in Food Science, 2025, 11: 101–131.

[19] BARNES RJ, DHANOA MS, LISTER SJ. Standard normal variate transformation and detrending of nearinfrared diffuse reflectance spectra [J]. Applied Spectroscopy, 1989, 43(5): 772–777.

[20] CHANG CJ, LI ZY, LI HY, *et al*. A novel fast method for identifying the origin of Maojian using NIR spectroscopy with deep learning algorithms [J]. Scientific Reports, 2022, 12: 21–31.

[21] SAVITZKY A, GOLAY MJE. Smoothing and differentiation of data by simplified least squares procedures [J]. Analytical Chemistry, 1964, 36(8): 1627–1639.

[22] SAMET O, ALEXANDER B, AHMED R, *et al*. Near-infrared spectroscopy and machine learning for classification of food powders during a continuous process [J]. Journal of Food Engineering, 2023, 341: 111–123.

[23] CHEN T, GUESTRIN C. A scalable tree boosting system: Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining [C]. San Francisco: Association for Computing Machinery, 2016.

[24] NICULESCUMIZIL A, CARUANA R. Predicting good probabilities with supervised learning: Proceedings of the 22nd international conference on machine learning [C]. San Francisco: Association for Computing Machinery, 2005.

[25] ZHAO R, KANG S, XUE Y. A new near-infrared spectral transformation method for quantitative analysis [J]. Optik, 2023: 287:171110

[26] ILEMOBAYO JA, DURODOLA O, ALADE O, *et al*. Hyperparameter tuning in machine learning: A comprehensive review [J]. Journal of Engineering Research and Reports, 2024, 26(6): 388–395.

[27] NORTHCUTT CG, JIANG L, CHUANG IL. Confident learning: estimating uncertainty in dataset labels [J]. Journal of Artificial Intelligence Research, 2021(70): 1373–1411.

[28] CHUNSHENG Y. A review on spectral data preprocessing techniques for machine learning and quantitative analysis [J]. iScience, 2025, 28(7): 112759.

[29] HE H, GARCIA EA. Learning from imbalanced data [J]. IEEE Transactions on Knowledge and Data Engineering, 2009, 21(9): 1263–1284.

[30] CHAWLA NV, BOWYER KW, HALL LO, *et al*. SMOTE: Synthetic minority over-sampling technique [J]. Journal of Artificial Intelligence Research, 2002, 16: 321–357.

[31] ELKAN C. The foundations of cost-sensitive learning: Proceedings of the 17th international joint conference on Artificial intelligence [C]. Seattle: Association for Computing Machinery, 2001.

[32] REFAEILZADEH P, TANG L, LIU H. Cross-validation encyclopedia of database systems [M]. New York: Springer, 2009.

[33] ROGGO Y, CHALUS P, MAURER L, *et al*. A review of near infrared spectroscopy and chemometrics in pharmaceutical technologies [J]. Journal of Pharmaceutical and Biomedical Analysis, 2007, 44(3): 683–700.

[34] WANG Z, WU Z, ZOU M, *et al*. A Voting-Based ensemble deep learning method focused on multi-step prediction of food safety risk levels: Applications in hazard analysis of heavy metals in grain processing products [J]. Foods, 2022, 11: 823.

[35] WANG G, LIU M, XUE H, *et al*. Simultaneous determination of the amylose and amylopectin content of foxtail millet flour by hyperspectral imaging [J]. Frontiers in Remote Sensing, 2025, 6: 1460523.

[36] PAN SHW, ZHANG X, XU WB, *et al*. Rapid *on-site* identification of geographical origin and storage age of tangerine peel by near-infrared spectroscopy [J]. Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy, 2022, 271: 120936.

[37] 张秋芳, 顾先春, 丁昊, 等. 近红外光谱法快速检测鸡蛋中Omega-3多不饱和脂肪酸的含量[J]. 食品安全质量检测学报, 2025, 16(21): 74–82.
ZHANG QF, GU XC, DING H, *et al*. Rapid determination of Omega-3 polyunsaturated fatty acids in eggs by near-infrared spectroscopy [J]. Journal of Food Safety & Quality, 2025, 16(21): 74–82.

(责任编辑: 韩晓红　蔡世佳)