DOI: 10.19812/j.cnki.jfsq11-5956/ts.20260203002
引用格式: 张译匀, 姚博伟, 朱秋楠, 等. 基于高光谱成像的玉米渣中伏马菌素快速无损检测模型研究[J]. 食品安全质量检测学报, 2026, 17(7): 110–119.
ZHANG YY, YAO BW, ZHU QN, *et al*. Study on rapid and nondestructive detection model of fumonisin in corn residue based on hyperspectral imaging [J]. Journal of Food Safety & Quality, 2026, 17(7): 110–119. (in Chinese with English abstract).

基于高光谱成像的玉米渣中伏马菌素快速无损检测模型研究

张译匀*, 姚博伟*, 朱秋楠, 张文轩, 段丽娟
(宁夏计量质量检验检测研究院, 银川 750200)

摘 要: 目的 探索采用高光谱成像技术(hyperspectral imaging, HSI)(光谱范围 900~1700 nm)结合深度学习方法构建玉米渣中伏马菌素(fumonisin, FUMs)快速无损检测模型的可行性。**方法** 通过人工制备含不同浓度梯度 FUMs (0~6 mg/kg)的玉米渣样品(共 175 个), 经多元散射校正(multiplicative scatter correction, MSC)预处理消除干扰后, 采用无信息变量消除法(uniformative variable elimination, UVE)与变量组合群体分析法(variable combination population analysis, VCPA)筛选出 26 个关键特征波长。**结果** 对比偏最小二乘回归(partial least squares regression, PLSR)、支持向量机(support vector machine, SVM)、卷积神经网络(convolutional neural network, CNN)、长短期记忆网络(long short-term memory, LSTM)、门控循环单元(gated recurrent unit, GRU)和注意力机制(Attention)等模型后, 发现 CNN-GRU-Attention 融合模型表现最优: FUMs 检测的预测决定系数达 0.9328、预测均方根误差为 0.3931、预测偏差比为 3.8247。**结论** 本研究建立的检测方法实现了 FUMs 的快速、精准、无损同步检测, 为谷物真菌毒素检测提供了新的技术路径, 也为深度学习在光谱分析领域的应用提供了参考。

关键词: 玉米渣; 伏马菌素; 高光谱成像技术; 深度学习

Study on rapid and nondestructive detection model of fumonisin in corn residue based on hyperspectral imaging

ZHANG Yi-Yun*, YAO Bo-Wei*, ZHU Qiu-Nan, ZHANG Wen-Xuan, DUAN Li-Juan
(*Ningxia Academy of Metrology & Quality Inspection*, *Yinchuan* 750200, *China*)

ABSTRACT: Objective To explore the feasibility of constructing a rapid and nondestructive detection model for fumonisin (FUMs) in corn residue using hyperspectral imaging (HSI, spectral range 900–1700 nm) combined with deep learning methods. **Methods** Corn bran samples containing a concentration gradient of FUMs (0–6 mg/kg) were artificially prepared (total of 175 samples). After removing interference through multiplicative scatter correction

收稿日期: 2026-02-03
基金项目: 2023 年度宁夏回族自治区青年科技托举人才培养项目(宁科协发组字〔2024〕6 号); 宁夏科技惠民项目(2024CMG03049)
第一作者/*通信作者: 张译匀(1996—), 女, 硕士, 助理工程师, 主要研究方向为食品质量安全。E-mail: 793787105@qq.com
***通信作者:** 姚博伟(1993—), 女, 硕士, 工程师, 主要研究方向为应用微生物及食品质量安全。E-mail: ybw2016@sina.com

(MSC) preprocessing, the univariate variable elimination (UVE) and variable combination population analysis (VCPA) methods were used to identify key feature wavelengths for FUMs (26 wavelengths), respectively. **Results** After comparing models including partial least squares regression (PLSR), support vector machine (SVM), convolutional neural network (CNN), long short-term memory (LSTM), gated recurrent unit (GRU) and Attention, the CNN-GRU-Attention fusion model demonstrated optimal performance: For FUMs detection, the prediction coefficient of determination reached 0.9328, the root mean square prediction error was 0.3931, and the prediction deviation ratio was 3.8247. **Conclusion** The established detection method enables rapid, precise and nondestructive simultaneous detection of FUMs, providing a novel technical pathway for grain mycotoxin detection and offering reference for the application of interpretable deep learning in spectral analysis.

KEY WORDS: corn residue; fumonisins; hyperspectral imaging; interpretable deep learning

0　引　言

玉米是世界范围内三大重要粮食作物之一, 其加工产品玉米渣因质地细腻、营养价值较高, 在食品加工、动物饲料以及工业应用中具有广泛用途。然而, 玉米在种植、生长、采收和储藏等环节易引发真菌毒素污染。其中, 伏马菌素(fumonisins, FUMs)是玉米制品中最典型的污染物之一, 具有明显的神经毒性、免疫抑制效应以及潜在致癌风险。在所有 FUMs 类型中, B 族 FUMs 在环境分布范围和毒性危害方面最为突出, 其中伏马菌素 B_1 (fumonisins B_1, FB_1)和伏马菌素 B_2 (fumonisins B_2, FB_2)毒性最强, 也是粮食与饲料中最主要的存在形式。研究表明, 在玉米类谷物 FUMs 污染中, FB_1 的检出量通常占总量的 70%~80%, FB_2 约占 15%~25%[1]。

FUMs 具有较强的热稳定性, 即使在 100 ℃条件下处理 30 min, 其结构仍难以破坏, 因此在食品或饲料加工过程中很难彻底清除。目前国际上针对食品中 FUMs 的限量标准尚未统一, 我国相关研究起步较晚, 尚未制定食品领域的明确限量要求。相比之下, 欧盟已对未加工玉米、直接食用玉米制品及婴幼儿谷物食品设定限量值, 分别为 4.0、1.0 和 0.2 mg/kg; 美国规定食用玉米中 FUMs 最大允许值为 2 mg/kg。而在饲料及原料方面, 我国 GB 13078—2017《饲料卫生标准》规定 FB_1 与 FB_2 总量限值为 60 mg/kg[2–4]。

近年我国食用玉米制品及饲料中 FUMs 的阳性检出率持续保持高水平, 且污染趋势呈上升态势[1]。不同区域调查结果显示, 中部地区 FUMs (FB_1+FB_2)检出率最高(约 50.0%), 华北地区次之(约 45.0%); 而在检测浓度方面, 西北、中部及东北地区均出现超过 2 mg/kg 的样品[5]。此外, 我国主要玉米产区中 FUMs (FB_1+FB_2)的检出率范围为 28.00%~100.00%[6]。XING 等[7]对河北、山东和河南 3 省 44 份玉米样品检测结果显示 FUMs 检出率达到 100%; 谭莉等[8]在吉林省 240 份玉米样品中发现 FB_1 与 FB_2 检出率分别高达 96.7%与 98.3%。因此, 构建一种快速、精准且无损的玉米渣中 FUMs (FB_1+FB_2)检测技术, 对于提升食品安全监管水平和维护市场秩序具有重要意义。

目前真菌毒素检测主要依赖高效液相色谱法(high performance liquid chromatography, HPLC)以及超高效液相色谱-串联质谱法(ultra performance liquid chromatography-tandem mass spectrometry, UPLC-MS/MS), 这类方法灵敏度与特异性较高, 但通常需要复杂的前处理, 检测耗时长、试剂成本高, 对操作人员专业能力要求较强, 难以满足食品生产现场对实时快速筛查和大批量检测的需求。在此背景下, 无损快速检测技术逐渐成为研究重点。高光谱成像技术由于能够同时获取样品的光谱信息与空间分布特征, 在不破坏样品的条件下, 通过捕捉毒素与玉米渣基质在特定波段下的光谱响应差异(如官能团振动吸收与散射特征), 实现毒素含量的快速预测[9]。

传统高光谱定量分析中常采用偏最小二乘回归(partial least squares regression, PLSR)、支持向量机(support vector machine, SVM)等机器学习方法, 这些模型结构简单、训练效率较高, 能够在一定程度上挖掘光谱数据与毒素浓度之间的线性或低阶非线性关系, 并具有较好的应用基础。然而, 面对复杂的高维光谱信息以及粒径、水分等基质干扰因素时, 这类方法对高阶特征的提取能力有限, 模型鲁棒性仍有待提升。

近年来, 深度学习模型如卷积神经网络(convolutional neural network, CNN)、长短期记忆网络(long short-term memory, LSTM)、门控循环单元(gated recurrent unit, GRU)以及融合注意力机制(Attention)的网络结构显示出更强的优势。这类模型能够以端到端方式自动学习光谱数据中的深层非线性特征, 减少对人工特征工程的依赖; 其中 CNN 擅长提取局部光谱细节, LSTM/GRU 能够刻画序列依赖特征, 而 Attention 可进一步强化关键波段信息, 使模型在处理高维复杂光谱数据时预测精度显著优于传统方法[10–12]。

基于上述背景, 本研究提出一种结合高光谱成像与深度学习的玉米渣 FUMs 检测方案, 利用 CNN-GRU-Attention 等模型的深层特征提取能力, 挖掘与

毒素含量相关的微弱光谱差异，从而提升复杂基质条件下的检测准确性。本研究将系统介绍实验设计、模型构建，为玉米渣中真菌毒素的快速无损检测提供新的技术路径，并为深度学习在光谱分析领域的进一步应用提供参考依据。

1 材料与方法

1.1 材料与试剂

均匀适中颗粒(0.5~1.0 mm)光谱信号稳定性最佳，能更好地反映样品内部化学组分(如 FUMs)的特征光谱，有助于提升高光谱成像技术检测模型的准确性与稳定性，契合高光谱成像技术在谷物加工制品检测中的应用需求，而过细颗粒(如小于 0.25 mm)虽光谱重复性较好，但可能增加样品团聚风险，导致光谱信号出现偏差，可能因颗粒团聚影响检测准确性，过粗颗粒(如大于 1.5 mm)表面粗糙度较高，光散射不均匀，会降低光谱数据的一致性，导致检测模型的分类误差增大，影响 FUMs 等污染物的检测精度[3]。故本研究选用了产自宁夏山逗子杂粮绿色食品科技开发有限公司的玉米渣预包装产品，其玉米颗粒度(中位粒径)为 0.5~1.0 mm(符合 GB/T 5507—2008《粮油检验 粉类粗细度测定》标准)。

FB_1标准溶液、FB_2标准溶液(色谱纯，质量浓度 1 g/mL，上海安谱实验科技股份有限公司)。

1.2 仪器与设备

HyperSIS-NIR 高光谱成像系统(光谱范围 900~1700 nm，北京卓立汉光仪器有限公司)，该系统可连续采集 256 个光谱波段，光谱分辨率为 5 nm，系统硬件由高光谱成像光谱仪、CCD 相机、4 盏 35 W 卤钨灯、电子位移平台及计算机组成。

1.3 样品制备

先对购买来的玉米渣中 FUMs (FB_1+FB_2)进行原始检测，确认其为无毒素基质。将 FB_1 和 FB_2 标准溶液配制为 100 mg/mL 的 FUMs (FB_1+FB_2)混合标准溶液，将标准品添加至无毒素玉米渣基质中并充分混匀，制备 FUMs (FB_1+FB_2)原始污染样(10 mg/kg)。参考欧盟规定玉米中 FUMs (FB_1+FB_2) 的 MRLs 为 4 mg/kg，将上述原始污染样与无毒素基质定量混合，构建毒素的污染浓度梯度：FUMs (FB_1+FB_2)梯度设为 0、0.5、1.5、3.0、6.0 mg/kg。每个浓度梯度制备 35 个平行样品，每个样品取 15 g 平铺于直径 5 cm 的培养皿中，总计 175 个样品用于后续高光谱数据采集。

1.4 高光谱成像系统及数据采集

由于高光谱成像过程易受外界光散射、噪声等干扰，为保证数据稳定性，系统整体置于密闭金属箱内，以完全消除环境光的影响。样品采集前需将系统开机预热 30 min，确保设备处于最佳工作状态。将样品置于电子位移平台后，设定核心采集参数：平台移动速度 20 mm/s，光源与样品间距 40 mm，曝光时间 3 ms[2]。

采集完成的高光谱图像需进行黑白校正，以消除系统暗电流、光源不均匀性等带来的误差，校正计算如公式(1)所示：

$$R=\frac{I-D}{W-D}\times 100\% \tag{1}$$

式中：I 为样本初始光谱图像；W 为白色参考图像；D 为暗电流图像；R 为校正后光谱图像。

其中，暗电流图像 D 通过盖上相机镜头盖(反射率≈0%)采集，白色参考图像 W 通过白色聚四氟乙烯平板(反射率大于 99%)采集，以分别模拟“无反射”与“全反射”的极端场景，确保校正效果。

1.5 玉米渣样品光谱提取

在本研究的光谱检测阶段，实验样本的光谱扫描图谱易受到实验平台稳定性及光源照射条件(如光强波动、波长稳定性差异)的干扰，进而可能引发图谱基线漂移或特征峰偏移[13–14]。为消除上述干扰对数据质量的影响，研究通过在各样本校正图像中选取感兴趣区域(region of interest, ROI)，并对所提取 ROI 的光谱反射率值进行平均计算，最终获得用于后续数据分析的标准化光谱数据。

1.6 数据处理方法

1.6.1 剔除异常值

受限于样品采集环节的复杂变量(包括人为操作偏差、环境光线干扰、仪器运行误差)，采集数据中不可避免会存在异常值[15]。这些异常值会对模型的学习过程产生干扰，导致其预测能力下降，故需在模型构建前排除异常值的影响。本研究采用蒙特卡洛方法进行异常值剔除，该方法借助随机模拟生成多组数据集，通过统计分析手段精准定位并筛选离群值，在实际应用中展现出良好的异常值剔除可靠性[16–17]。

1.6.2 数据分集

本研究在剔除数据集中的异常值后，获得 160 个有效样本。基于随机抽样(random sampling, RS)算法，将两类指标的有效样本均按 3:1 的比例划分为校正集与预测集[18]。校正集被用于模型的构建与参数校正，预测集则作为独立数据集，用于检验所建模型的泛化能力与评估精度[19]。

1.6.3 数据预处理

预处理的核心目的在于进一步消除光谱数据中的噪声、散射及基线漂移，以此获取更有效的建模光谱信息并提升模型的预测精度[20]。为验证不同预处理方法的实际效能，本研究选取 5 种预处理算法开展迭代性测试，具体涵盖滤波(savitzky-golay, SG)、归一化(Normalize)、基线校正

(Baseline)、标准正态变量变换(standard normal variate transformation, SNV)及多元散射校正(multiplicative scatter correction, MSC)。

1.6.4　特征波长的选择

高光谱数据普遍存在维度高、波段间相关性强的特点, 若直接采用全谱段构建分析模型, 易引入大量冗余信息与噪声干扰, 不仅会增加模型计算的复杂性, 还可能降低其分析精度[21−22]。为此, 本研究综合运用多种变量选择与降维方法对光谱数据进行处理, 以筛选出能表征全波段核心特征的重要变量, 具体包括竞争性自适应重加权采样法(competitive adaptive reweighted sampling, CARS)、无信息变量消除法(uninformative variable elimination, UVE)、变量组合群体分析法(variable combination population analysis, VCPA), 旨在通过有效降维缩短计算时长, 提升模型的准确性与稳定性[23]。

1.7　建模方法

1.7.1　机器学习

选取单一算法类模型中的 PLSR 与 SVM 进行建模分析。为保证模型训练与验证的客观性, 所有样本数据均通过随机抽样方式按 3:1 的比例划分为校正集和预测集; 同时, 为减少数据划分随机性可能带来的偏差并避免模型过拟合, 建模过程中采用交叉验证策略以优化模型参数、提升最终模型的稳定性与可靠性[24]。

1.7.2　深度学习

选用深度学习领域中的两种融合架构模型, 分别为 CNN-LSTM-Attention 与 CNN-GRU-Attention 进行建模分析。其中, CNN 擅长捕获光谱数据中的局部空间特征与细节模式, LSTM 和 GRU 则分别通过特殊的门控机制有效处理时序依赖关系(LSTM 更侧重缓解长序列梯度消失问题, GRU 则以更简洁的结构实现高效时序建模), 而 Attention 的引入可进一步增强模型对关键特征的权重分配能力, 三者融合能充分发挥各自优势以提升复杂光谱数据的建模效能, 最终提升模型的稳定性与预测可靠性[25−28]。

1.8　模型评价标准

本研究通过多项性能指标对校准模型的准确性开展系统评估, 具体包括校准决定系数(R_c^2)、校准均方根误差(root mean square error of calibration, RMSEC)、预测决定系数(R_P^2)、预测均方根误差(root mean square error of prediction, RMSEP)及预测偏差比(residual prediction deviation, RPD)[29]。模型稳健性的判别标准明确为: 较高的决定系数(R_c^2、R_P^2)、较大的 RPD 值与较低的均方根误差(RMSEC、RMSEP)。各指标的计算均参照公式(2)~(4)执行:

$$R^2 = 1 - \frac{\Sigma_{i=1}^{n}(y_i - \hat{y})^2}{\Sigma_{i=1}^{n}(y_i - \bar{y})^2} \tag{2}$$

$$\text{均方根误差} = \sqrt{\frac{\Sigma_{i=1}^{n}(\hat{y}_i - y_i)^2}{n}} \tag{3}$$

$$\text{RPD} = \frac{\sigma_y}{\text{RMSEP}} \tag{4}$$

式中: y_i、$\hat{y}_i$、$\bar{y}$ 和 σ_y 分别代表响应变量的测量值、预测值、均值和标准差值; n 代表样本数量。

1.9　数据处理

本研究中光谱数据预处理方法包括: SG、Normalize、Baseline、SNV、MSC 等 5 种方法。定量方法包括 PLSR、SVM、CNN-LSTM-Attention 及 CNN-GRU-Attention 等 4 种方法。模型效果评价主要由 RMSEC、RMSEP 和 RPD 等参数决定。上述数据处理均采用 MATLAB R2024b 软件。数据预处理、质量评价模型构建与效果评价均在 The Unscrambler X 10.4 软件中进行, 数据生成图片处理由 Origin Pro 2025 软件完成。

2　结果与分析

2.1　光谱分析结果

由图 1 可见, 玉米样本集的平均光谱反射率曲线整体趋势高度一致, 这一实验现象暗示各样本的化学组成未发生根本性差异, 仅存在微量成分波动。在 900~1700 nm 的目标光谱区间内, 玉米样本的反射率数值稳定在 0~0.8 范围, 且所有样本均在 970、1200、1450 nm 附近出现特征吸收峰, 其形成机制与玉米体内官能团的振动特性密切相关: 970 nm 附近的吸收峰源于 O-H 伸缩振动的二级倍频, 这一振动模式是淀粉分子中羟基(-OH)的典型光谱响应[30], 淀粉分子通过-OH 形成分子间氢键, 构成稳定的结晶结构。当玉米渣受到 FUMs 污染时, 产生 FUMs 的镰刀菌会分泌淀粉酶, 分解淀粉分子中的糖苷键, 导致淀粉结晶结构破坏, 游离-OH 数量增加; 同时, FUMs 分子自身含有的多个-OH 会与淀粉游离-OH 形成氢键结合, 进一步增强该波段的吸收强度。根据朗伯-比尔定律, 吸收强度与物质浓度正相关, 因此随着 FUMs 污染浓度升高, 970 nm 波段的吸收峰强度基本增强, 对应的光谱反射率呈现整体下降趋势; 1200 nm 附近的吸收峰由 C-H 伸缩振动的一级或二级倍频引发, 该峰既能反映玉米油脂成分的光谱特征, 也可通过几丁质的 C-H 振动间接关联真菌污染信息[31], FUMs 污染过程中, 镰刀菌在玉米渣中生长繁殖会消耗油脂等营养物质, 导致油脂含量下降, 进而减弱 C-H 振动的吸收强度; 但同时, FUMs 分子结构中的亚甲基(-CH_2-)数量较多(分子结构中含多个-CH_2-链), 其 C-H 伸缩振动会在该波段产生特征吸收, 且吸收强度随 FUMs 浓度升高而增强。二者协同作用下, 1200 nm 波段的吸收峰强度与 FUMs 浓度呈现一定相关性, 且该波段的光谱信号能同时反映真菌生长状态与 FUMs 积累量, 成为区分污染样本与未污染样本的关键

波段；1450 nm 附近的吸收峰则是水分(O-H)与蛋白质(N-H)共同作用的结果，可综合指示玉米的水分含量与蛋白分布[32]，FUMs 污染会显著改变玉米渣的水分与蛋白质含量：一方面，真菌生长繁殖过程中会消耗蛋白质作为营养来源，导致蛋白质分子中的 N-H 键数量减少，减弱该波段的 N-H 振动吸收；另一方面，真菌代谢会产生大量代谢水，使玉米渣水分含量升高，增强 O-H 振动吸收。此外，FUMs 分子中的-OH 会与水分中的 O-H 形成氢键，进一步调节该波段的吸收强度。最终，1450 nm 波段的吸收峰强度随 FUMs 浓度变化呈现规律性波动，且与毒素浓度呈负相关，可作为特征波段，提升 FUMs 检测的准确性。

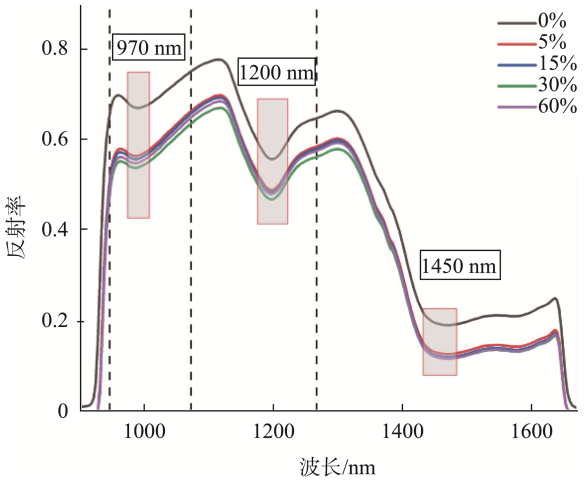


图 1 FUMs 的平均光谱曲线

Fig.1 Average spectral curves of FUMs

综上，FUMs 污染引发玉米渣光谱反射率下降的内在逻辑是：FUMs 自身分子官能团(-OH、-CH_2-)的特征振动与玉米渣基质成分(淀粉、蛋白质、水分)的官能团振动叠加，同时真菌生长导致基质成分发生结构性变化，共同改变了 970、1200、1450 nm 等关键波段的吸收强度，进而表现为光谱反射率随毒素浓度升高而规律性下降。这一构效关系明确了上述特定波段作为 FUMs 检测特征波段的科学依据，也解释了本研究中 UVE 方法筛选的特征波长集中于这些波段的合理性，为后续模型简化与检测精度提升提供了坚实的理论支撑。

进一步观察发现，FUMs 的污染水平会显著影响样本的光谱反射率：随着毒素污染浓度的增加，样本的光谱反射率呈现规律性下降。利用不同样本原始光谱在吸收峰强度、反射率绝对值上的差异特征，可构建有效的区分依据，实现对各类污染样本的准确鉴别。

2.2 光谱数据预处理结果

光谱数据易受噪声、杂散光等无关信息干扰，会降低后续定量建模的准确性，因此在构建模型前需对光谱数据进行预处理，以提升模型的预测精度。针对 FUMs 的研究中，均采用 SG、Normalize、SNV、Baseline 及 MSC 5 种方法，对 ROI 内的光谱数据进行去噪预处理，不同预处理方法处理后的光谱平均反射率曲线如图 2 所示。

为筛选最优预处理方法，将原始光谱数据及经上述 5 种方法预处理后的光谱数据分别输入 PLSR、SVM、CNN-LSTM-Attention 及 CNN-GRU-Attention 4 种模型中进行对比分析，结果如表 1 所示。研究发现毒素在相同模型下不同预处理方法的建模效果存在显著差异。其中，MSC 预处理方法在 PLSR、CNN-LSTM-Attention 及 CNN-GRU-Attention 模型中均表现最优。具体表现为：FUMs 经 CNN-GRU-Attention-MSC 建模的 R_P^2 达 0.9467，RMSEP 为 0.3568；上述指标均优于其他预处理方法(表 1)。因此，确定 MSC 为后续研究的最优光谱预处理方法。

2.3 特征波长的选择结果

高光谱数据具有高维度、相邻变量强相关性及显著共线性特征，其蕴含的部分低相关性成分会加剧模型训练的复杂度与耗时，同时降低建模效果与预测精度[33]。因此，特征波长提取为光谱数据分析的核心环节，对简化模型结构、提升计算效率与预测精度至关重要，数据预处理后需通过特征波长选择算法完成[4]。图 3 呈现毒素不同特征提取方法的最优波长分布。结果显示，对于 256 个光谱波长点，FUMs 通过 CARS、UVE、VCPA 方法分别提取 55、26、11 个关键波长。这些光谱区主要对应色素、蛋白质、水分及脂肪等物质。基于此，将上述关键波段信息引入模型，构建适用于深度分析的简化模型，以实现对掺假水平更快速、精准的预测。

CARS 算法联合蒙特卡洛采样(设定采样次数为 50)与 PLSR，借助自然选择原理筛选对模型贡献最优的波长变量。针对 256 组光谱数据，该算法通过运行 PLSR 模型，以交叉验证 RMSECV 为迭代优化目标，将 RMSECV 达到最小值时对应的波长确定为最终的最优变量组合。筛选得到的特征波长多集中于 970、1200、1450 nm 附近，与平均光谱的吸收峰高度契合，且多数特征波长与既往研究已确定的相关波长存在关联性，能够有效提升模型的预测准确率。

UVE 特征波长提取采用十折交叉验证策略，对数据进行中心化处理，基于交叉验证 RMSECV 筛选最优潜变量。该方法通过随机噪声矩阵评估光谱变量信息量，结合 PLSR 筛选有效波长，提取的特征波长在全波段分布更均匀且多集中于波峰、波谷位置，表明其特征提取更注重波长间的综合作用。

VCPA 的特征波长提取基于相关性筛选强相关且非冗余的变量组合，借助主成分分析或因子分析实现数据降维，通过聚类算法划分同质群体，并结合交叉验证评估群体区分度选择少量特征波长。该算法运行时采用五折交叉验证

与数据中心化, 设定最大主成分得分 20、最佳子集采样数 1000、有效自由度迭代 50 次及最佳与最差子集比例 a=0.1; 其选取的关键波长多位于波峰或波谷位置, 表明 VCPA 能有效保留有效信息, 提升模型准确度。

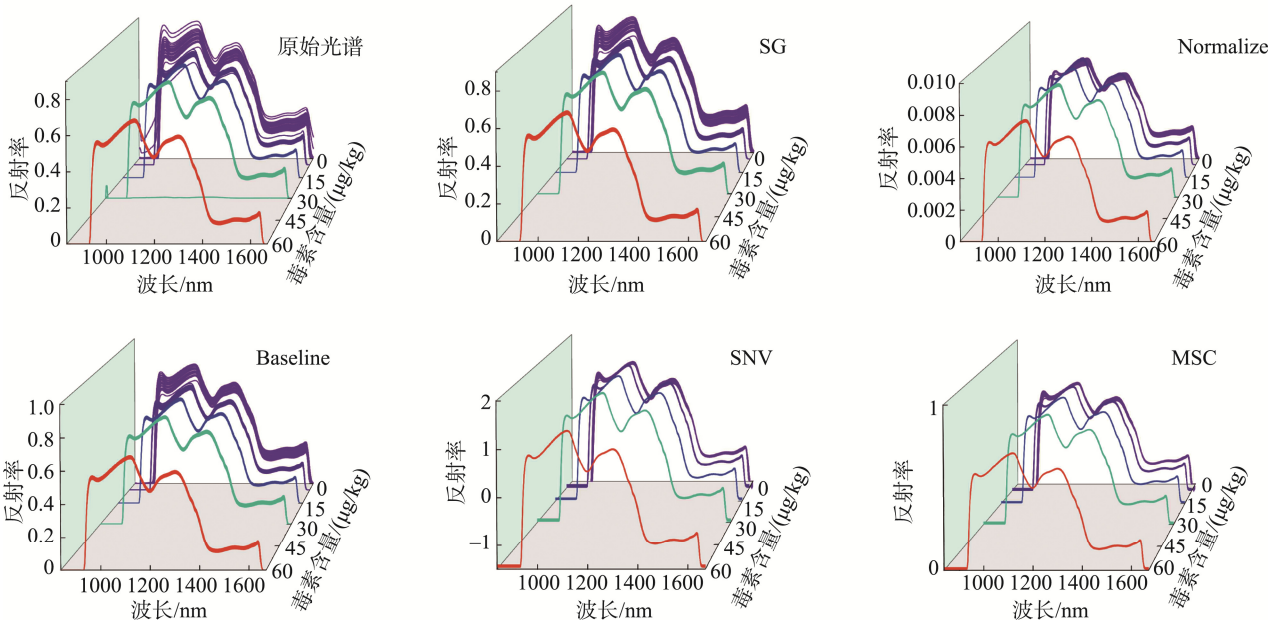


图 2　FUMs 的预处理光谱曲线

Fig.2　Preprocessed spectral curve of FUMs

表 1　样品光谱预处理数据建模结果

Table 1　Modeling results based on spectral preprocessing of samples

建模方法	预处理方法	校正集		预测集	
		R_c^2	RMSEC	R_P^2	RMSEP
PLSR	SG	0.8516	0.5261	0.8428	0.9562
	Normalize	0.8704	0.4736	0.8531	0.5534
	Baseline	0.8449	0.5027	0.8465	0.5851
	SNV	0.9071	0.4158	0.8412	0.6023
	MSC	0.9152	0.3919	0.8736	0.5311
SVM	SG	0.9072	0.2685	0.9125	0.2987
	Normalize	0.9651	0.1310	0.8954	0.3643
	Baseline	0.9122	0.2601	0.9136	0.2503
	SNV	0.9875	0.0790	0.8763	0.3979
	MSC	0.9832	0.0747	0.8765	0.3998
CNN-LSTM-Attention	SG	0.8904	0.4529	0.8777	0.5247
	Normalize	0.8496	0.5306	0.8729	0.5383
	Baseline	0.7991	0.6123	0.8121	0.6591
	SNV	0.9072	0.4154	0.8949	0.4877
	MSC	0.9168	0.3949	0.9083	0.4538
CNN-GRU-Attention	SG	0.9180	0.3857	0.8907	0.4821
	Normalize	0.9112	0.4126	0.9034	0.4614
	Baseline	0.9231	0.3789	0.8925	0.4825
	SNV	0.9316	0.3503	0.9200	0.4332
	MSC	0.9628	0.2665	0.9467	0.3568

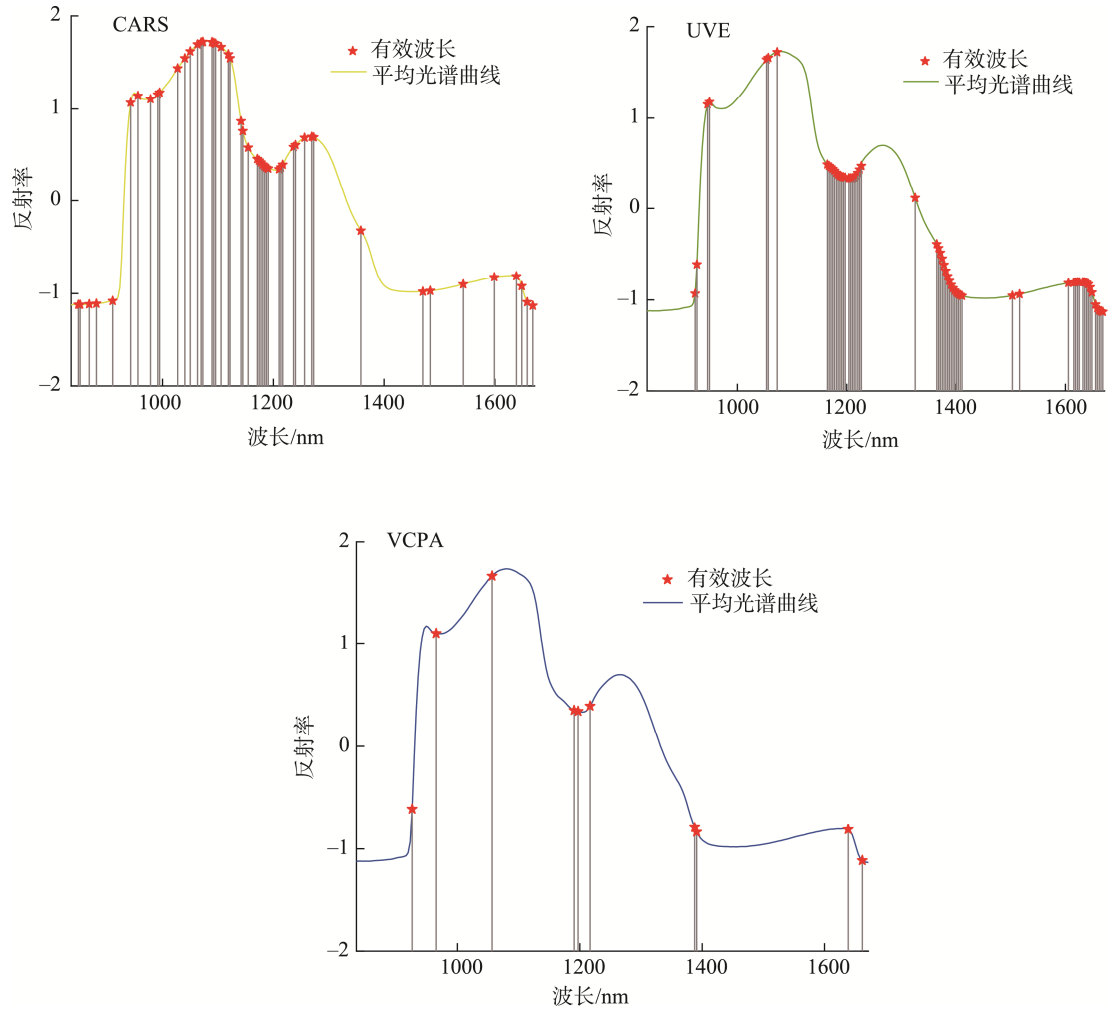


图 3 FUMs 的特征波长提取结果

Fig.3 Results of characteristic wavelength extraction for FUMs

2.4 模型的建立结果

有效波长筛选为精准回归模型的构建提供关键支撑。挑选最优效能特征波长，既能够达成数据维度压缩，又能最大程度削弱噪声冗余；此外，该操作还能增强模型可解释性，赋能潜在关联及内在规律的深度挖掘，最终实现预测精度优化与过拟合风险削减的作用[34−35]。分别使用全波段和特征变量构建 PLSR、SVM、CNN-LSTM-Attention 和 CNN-GRU-Attention 4 种模型，对玉米渣中 FUMs 进行定量分析，通过 R^2、RMSEC、RMSEP 以及 RPD 对模型性能进行评价，各模型分析结果如表 2 所示。

对于 FUMs 分析的机器学习模型，PLSR 的最优定量分析模型为 CARS，它的 R_P^2、RMSEP、RPD 分别为 0.9119、0.4714、3.0872，SVM 的最优定量分析模型为 UVE，它的 R_P^2、RMSEP、RPD 分别为 0.9165、0.4347、3.5011，对于深度学习模型，CNN-LSTM-Attention 的最优定量分析模型为 UVE，它的 R_P^2、RMSEP、RPD 分别为 0.9108、0.4580、3.3624，CNN-GRU-Attention 的最优定量分析模型为 UVE，它的 R_P^2、RMSEP、RPD 分别为 0.9328、0.3931、3.8247。结果表明，在 4 种定量分析模型中 UVE 的特征提取方法具有明显的优势，该算法提取 26 个特征变量，表现出良好的模型拟合性，在机器学习模型中，经过特征提取的预测精度大体都有一定程度的提升，但提升的程度不高，而对于深度学习模型，只有 UVE 算法对全波段的预测精度有了大幅度的提升，CARS 和 VCPA 的预测精度并未得到改善。在 4 种建模方法中 CNN-GRU-Attention-MSC-UVE 成为最佳回归模型，相较于 PLSR、SVM、CNN-LSTM-Attention 模型准确率分别提高了 0.0209%、0.0163%、0.0220%，RMSEP 分别降低了 0.0783、0.0416、0.0649，RPD 分别提高了 0.7375、0.3236、0.4623。

综合分析表明，CNN-GRU-Attention 模型在准确性、特异性、灵敏度及精密度等评估指标上均取得最优性能，其动态学习机制可实现对 FUMs 的精准检测与有效评估。

表 2　基于有效波长预测毒素含量的回归模型性能

Table 2　Performance of regression models based on effective wavelengths for toxin prediction

建模方法	特征提取方法	校正集		预测集		
		R_c^2	RMSEC	R_P^2	RMSEP	RPD
PLSR	RAW	0.9153	0.3958	0.8732	0.5337	2.8449
	CARS	0.9618	0.2656	0.9119	0.4714	3.0872
	UVE	0.9025	0.4667	0.9027	0.4662	3.1273
	VCPA	0.8736	0.4842	0.8959	0.4896	3.0001
SVM	RAW	0.9883	0.0711	0.8712	0.3985	3.1531
	CARS	0.9486	0.3085	0.8953	0.4833	3.1575
	UVE	0.8963	0.4403	0.9165	0.4347	3.5011
	VCPA	0.8651	0.5027	0.8468	0.5898	2.5792
CNN-LSTM-Attention	RAW	0.9160	0.3946	0.9075	0.4525	3.3265
	CARS	0.8773	0.4769	0.8686	0.5436	2.8146
	UVE	0.9001	0.4328	0.9108	0.4580	3.3624
	VCPA	0.8709	0.4902	0.8821	0.5147	2.9228
CNN-GRU-Attention	RAW	0.9343	0.3487	0.9201	0.4239	3.5491
	CARS	0.9441	0.3214	0.8713	0.5388	2.8238
	UVE	0.9498	0.3043	0.9328	0.3931	3.8247
	VCPA	0.9156	0.3977	0.8759	0.5293	2.8900

3　结论与讨论

本研究针对玉米渣中 FUMs (B_1+B_2)快速无损检测需求, 构建了高光谱成像技术结合深度学习的同步检测技术体系, 有效解决了传统检测方法周期长、破坏性检测等关键问题, 为食品安全领域真菌毒素检测提供了新的理论支撑与技术方案。

研究通过系统的实验设计与数据分析, 明确了各环节最优技术参数: 光谱预处理阶段, MSC 方法表现最优, 可有效消除噪声与基质干扰, 使 FUMs 经 CNN-GRU-Attention-MSC 建模的 R_P^2 达 0.9467, RMSEP 降至 0.3568; 特征波长选择中, UVE 方法对 FUMs 通过深度学习建模检测的特征提取效能突出, 筛选的特征波长均集中于 970、1200、1450 nm 等与淀粉、水分、蛋白质相关的特征波段, 为模型简化与精度提升奠定基础。建模性能对比显示, CNN-GRU-Attention 融合模型展现出最优的定量预测能力: 针对 FUMs, CNN-GRU-Attention-MSC-UVE 模型的 R_P^2 达 0.9328, RMSEP 为 0.3931, RPD 为 3.8247; 其融合架构可充分挖掘光谱数据的局部特征与时序依赖关系, 优于 PLSR、SVM 及 CNN-LSTM-Attention 模型。

在国内研究方面, 卢萍萍等[6]针对市售爆米花中真菌毒素污染检测, 采用近红外光谱结合机器学习方法, 构建的 FUMs 检测模型 R_P^2 为 0.897, RMSEP 为 0.512, 模型精度低于本研究(R_P^2 0.9328、RMSEP 0.3931), 可能由于本研究采用的 CNN-GRU-Attention 融合模型, 相比传统机器学习算法, 能更充分挖掘光谱数据的深层特征, 同时通过 UVE 特征波长选择进一步降低了基质干扰。此外, ZHANG 等[4]针对玉米中黄曲霉毒素 B_1 的高光谱成像技术检测, 构建的多尺度特征注意力网络模型 R_P^2 达 0.921, RPD 为 3.56, 与本研究模型性能接近, 但该研究检测对象为完整玉米籽粒。

国外 XING 等[7]针对华北平原玉米在自然干燥前后的真菌及霉菌毒素分布开展研究, 采用高光谱成像技术结合传统建模方法, FUMs 检测的 R_P^2 为 0.883, RMSEP 为 0.557, 其研究重点为毒素分布规律, 且未采用深度学习算法, 难以充分挖掘光谱数据的时序与局部关联特征。KANDPAL 等[30]采用短波红外高光谱成像技术检测玉米籽粒中黄曲霉毒素 B_1, 模型 R_P^2 为 0.905, RMSEP 为 0.428, 但检测对象为完整籽粒, 本研究针对玉米渣中 FUMs 检测, 基质干扰更强, 均匀性更差, 更能体现模型的抗干扰能力与实际应用价值, 模型精度更优(RMSEP 0.3931)。TAO 等[31]采用近红外高光谱成像技术评估玉米籽粒中黄曲霉毒素污染, 构建的模型 R_P^2 为 0.912, RPD 为 3.27, 低于本研究的 RPD 值(3.8247), 表明本研究模型的预测稳定性与泛化能力更具优势, 这得益于 CNN-GRU-Attention 模型对光谱特征的深度挖掘, 以及 UVE 与 MSC 的协同优化作用。此外, VENTURINI 等[23]采用荧光光谱结合机器学习开展多霉菌毒素检测, 虽能同时检测多种毒素, 但检测速度与无损性均不及本研究采用的高光谱成像技术, 且模型精度低于本研究模型。

综合来看，本研究的创新点主要体现在 3 个方面：(1)聚焦玉米渣这类易被忽视的谷物加工制品，相比同类研究中针对完整谷物籽粒的检测，更贴合实际生产加工中的质量管控需求；(2)构建的 CNN-GRU-Attention 融合模型，结合 UVE 特征波长选择与 MSC 预处理，在检测精度(R_P^2 0.9328)、稳定性(RPD 3.8247)上优于同类研究采用的传统机器学习模型及单一深度学习模型；(3)明确了 970、1200、1450 nm 等特征波段与玉米渣中淀粉、水分等基质组分及 FUMs 的关联，为后续模型简化与便携式检测设备开发提供了理论依据。同时，本研究也发现，与部分针对单一、纯净基质的研究相比，本研究模型在复杂基质下的精度仍有提升空间，这也是未来需要重点优化的方向。

参考文献

[1] 刘晓萌, 王寿南, 姜德铭, 等. 伏马毒素的污染现状及其生物脱毒研究进展[J]. 中国畜牧杂志, 2024, 60(2): 143–149.
LIU XM, WANG SN, JIANG DM, *et al*. Current status of contaminationand advances in bio-degradation of fumonisins [J]. Chinese Journal of Animal Science, 2024, 60(2): 143–149.

[2] 熊振杰. 基于高光谱成像技术的鸡肉品质快速无损检测[D]. 广州: 华南理工大学, 2015.
XIONG ZJ. Rapid and non-destructive detection of chicken meat quality based on hyperspectral imaging technique [D]. Guangzhou: South China University of Technology, 2015.

[3] LIU YM, ZHANG JC, YUAN HL, *et al*. Non-destructive quality-detection techniques for cereal grains: A systematic review [J]. Agronomy, 2022, 12(12): 3187.

[4] ZHANG Y, HUAN K, LIU X, *et al*. Multi-scale feature attention network for rapid and non-destructive quantification of aflatoxin B_1 in maize using hyperspectral imaging [J]. Foods, 2025, 14(21): 3769.

[5] 程玲云, 傅骏青, 杨振东, 等. 玉米制品中伏马菌素污染情况调查研究[J]. 食品安全导刊, 2024(26): 93–95, 99.
CHENG LY, FU JQ, YANG ZD, *et al*. Investigation on the contamination of fumonisin in maize products [J]. Journal of Food Safety Guide, 2024(26): 93–95, 99.

[6] 卢萍萍, 曾祥辉, 何颖, 等. 2024 年湖南省市售爆米花真菌毒素污染状况及人群暴露风险评估[J]. 食品安全质量检测学报, 2025, 16(3): 65–73.
LU PP, ZENG XH, HE Y, *et al*. Assessment of mycotoxin contamination in commercial popcorn in Hunan Province in 2024 and associated human exposure risk [J]. Journal of Food Safety & Quality, 2025, 16(3): 65–73.

[7] XING FG, LIU XA, WANG LM, *et al.* Distribution and variationof fungi and major mycotoxins in pre- and post-nature dryingmaize in North China Plain [J]. Food Control, 2017, 80: 244–251.

[8] 谭莉, 孟繁磊, 范宏, 等. 超高效液相色谱-串联质谱法同时测定玉米中 9 种真菌毒素[J]. 食品工业科技, 2021, 42(7): 240–245.
TAN L, MENG FL, FAN H, *et al.* Simultaneous determination of nine mycotoxins in corn by ultra-high-performance liquid chromatography-tandem mass spectrometry [J]. Science and Technology of Food Industry, 2021, 42(7): 240–245.

[9] LIU T, ZHANG WJ, XU HL, *et al*. Non-destructive quantitative detection of lymphatic meat adulterated in beef by hyperspectral imaging combined with interpretable deep learning [J]. Food Control, 2026, 181: 111742.

[10] AQEEL M, SOHAIB A, IQBAL M, *et al.* Hyperspectral identification of oil adulteration using machine learning techniques [J]. Current Research in Food Science, 2024, 8: 100773.

[11] DHAKAL K, SIVARAMAKRISHNAN U, ZHANG X, *et al.* Machine learning analysis of hyperspectral images of damaged wheat kernels [J]. Sensors, 2023, 23(7): 3523.

[12] HUANG S, LIU Q, WU Y, *et al.* Edible mushroom greenhouse environment prediction model based on attention CNN-LSTM [J]. Agronomy, 2024, 14(3): 473.

[13] DONG F, HAO J, LUO R, *et al*. Identification of the proximate geographical origin of wolfberries by two-dimensional correlation spectroscopy combined with deep learning [J]. Computers and Electronics in Agriculture, 2022, 198: 107027.

[14] FENG Y, LV Y, DONG F, *et al.* Combining Vis-NIR and NIR hyperspectral imaging techniques with a data fusion strategy for prediction of norfloxacin residues in mutton [J]. Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy, 2024, 322: 124844.

[15] DONG F, BI Y, HAO J, *et al.* A Combination of near-infrared hyperspectral imaging with two-dimensional correlation analysis for monitoring the content of alanine in beef [J]. Biosensors, 2022, 12(11): 1043.

[16] WANG G, YAN X, FENG Y, *et al*. Deep learning-driven hyperspectral imaging for real-time monitoring and growth modeling of psychrophilic spoilage bacteria in chilled beef [J]. International Journal of Food Microbiology, 2025, 439: 111254.

[17] YU Y, CHEN W, ZHAO D, *et al*. Meat species authentication using portable hyperspectral imaging [J]. Frontiers in Nutrition, 2025, 12: 1577642.

[18] LAI Y. A comparison of traditional machine learning and deep learning in image recognition [J]. Journal of Physics: Conference Series, 2019, 1314(1): 12148.

[19] FAN Y, YAO X, WANG Z, *et al.* Combining dual-wavelength laser-induced fluorescence hyperspectral imaging with mutual information decomposition and redundancy elimination method to detect aflatoxin B_1 of individual maize kernels [J]. Computers and Electronics in Agriculture, 2025, 239: 110938.

[20] ZHANG C, TANG Z, TAN X, *et al.* Detection of rice prolamin and glutelin content using hyperspectral imaging combined with feature selection algorithms and multivariate regression models [J]. Foods, 2025, 14(19): 3304.

[21] LONG Y, WANG Q, TANG X, *et al.* Detection of starch content in maize kernel based on Raman hyperspectral imaging technique [J]. Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy,

2025, 336: 126025.

[22] SHI R, ZHANG H, WANG C, *et al*. Data fusion-driven hyperspectral imaging for non-destructive detection of single maize seed vigor [J]. Measurement, 2025, 253: 117416.

[23] VENTURINI F, MICHELUCCI U, MAGNUS I, *et al.* Multi-mycotoxin detection using fluorescence spectroscopy and machine learning [J]. Food Control, 2026, 181: 111728.

[24] ADESOKAN M, OTEGBAYO B, ALAMU EO, *et al*. Evaluating the dry matter content of raw yams using hyperspectral imaging spectroscopy and machine learning [J]. Journal of Food Composition and Analysis, 2024, 135: 106692.

[25] HU X, ZENG J, DAI M, *et al*. Hyperspectral-driven PSO-SVM model and optimized CNN-LSTM-Attention fusion network for qualitative and quantitative non-destructive detection of adulteration in strong-aroma Baijiu [J]. Food Chemistry, 2025, 490: 145197.

[26] HUANG P, YANG P, XU L, *et al*. Moisture content detection of Tibetan tea based on hyperspectral technology, machine vision and machine learning [J]. Journal of Food Measurement and Characterization, 2025, 19(2): 1167–1185.

[27] LONG Z, WANG T, ZHANG Z, *et al.* Quality and maturity detection of korla fragrant pears via integrating hyperspectral imaging with multiscale CNN-LSTM [J]. Foods, 2025, 14(20): 3561.

[28] WANG XY, LIAO WX, AN D, *et al*. Maize haploid identification via LSTM-CNN and hyperspectral imaging technology [J]. ArXiv, 2018. https://doi.org/10.48550/arXiv.1805.09105

[29] ZOU J, QU H, ZHANG P, *et al*. Conventional to deep learning methods for hyperspectral unmixing: A review [J]. Remote Sensing, 2025, 17(17): 2968.

[30] KANDPAL LM, LEE S, KIM MS, *et al*. Short Wave Infrared (SWIR) hyperspectral imaging technique for examination of aflatoxin B_1 (AFB_1) on corn kernels [J]. Food Control, 2015, 51: 171–176.

[31] TAO F, YAO H, HRUSKA Z, *et al*. Near-infrared hyperspectral imaging for evaluation of aflatoxin contamination in corn kernels [J]. Biosystems Engineering, 2022, 221: 181–194.

[32] ZHANG H, LI W, WANG G, *et al*. Predicting stomatal conductance of chili peppers using TPE-optimized LightGBM and SHAP feature analysis based on UAVs' hyperspectral, thermal infrared imagery, and meteorological data [J]. Computers and Electronics in Agriculture, 2025, 231: 110036.

[33] EKRAMIRAD N, KHALED AY, DOYLE LE, *et al.* Nondestructive detection of codling moth infestation in apples using pixel-based NIR hyperspectral imaging with machine learning and feature selection [J]. Foods, 2021, 11(1): 8.

[34] ZHANG J, DAI L, ZHUANG R, *et al*. Corn seed freezing damage identification of different sides based on hyperspectral imaging and SPA-2DCOS fusion algorithm [J]. Molecules, 2025, 30(10): 2178.

[35] ZHOU Y, XIAO Z, BAO T, *et al*. Estimation network for multiple chemical parameters of astragalus leaves based on attention mechanism and multivariate hyperspectral features [J]. Agriculture, 2024, 14(11): 1859.

(责任编辑: 安香玉　蔡世佳)