

doi:10.6041/j.issn.1000-1298.2025.06.021

基于高光谱成像的小麦赤霉病严重度轻量化检测方法

梁晓颖^{1,2} 张紫婷^{1,2} 杨 硕^{1,2} 陈 煦^{1,2} 姚志凤^{1,2} 宋怀波^{1,2}

(1. 西北农林科技大学机械与电子工程学院, 陕西杨凌 712100;

2. 农业农村部农业物联网重点实验室, 陕西杨凌 712100)

摘要: 为实现对小麦赤霉病(*Fusarium head blight*, FHB)严重度等级快速无损检测,采用高光谱成像技术结合机器学习模型进行建模分析。通过对小麦麦穗中部籽粒进行镰刀菌真菌接种,共获取1 660个不同程度的患病麦穗样本。利用高光谱成像设备采集麦穗样本高光谱信息,将整个麦穗作为感兴趣区域获取其平均光谱信息。通过对比归一化(Normalization)、标准正态变量变换(Standard normal variate, SNV)、多元散射校正(Multiplicative scatter correction, MSC)和平滑导数(Savitzky-Golay, SG)4种不同预处理后光谱对FHB严重度等级识别的准确率,选用SNV算法作为最佳预处理方法,并对其处理后的光谱数据进行进一步分析。随后对预处理后的光谱数据采用连续投影算法(Successive projections algorithm, SPA)、竞争性自适应重加权采样(Competitive adaptive reweighted sampling, CARS)、统一流形逼近与投影(Uniform manifold approximation and projection, UMAP)和线性判别分析(Linear discriminant analysis, LDA)算法进行降维,通过比较最终选择能降到3维,且保持分类准确率以及较低时间复杂度的LDA算法。揭示了LDA判别FHB严重度等级的特征波段处于540 nm叶绿素反射峰至650 nm红光吸收谷波段区间,是由于随着病情不断加重,叶绿素含量急速减少和叶片结构损伤的协同效应。最终构建结合SNV和LDA的轻量级支持向量机(Support vector machine, SVM)FHB严重度等级识别的最优模型。结果表明,该研究样本数据在构建的算法模型下测试集和训练集的准确率分别为96.05%和94.71%,且LDA将256维数据降到3维空间的时间复杂度仅为0.09 s,能够快速高效地进行FHB严重度判别且具有优秀的泛化能力,为将来田间大面积实时快速的识别FHB奠定了基础。

关键词: 小麦赤霉病; 高光谱成像; 无损检测; 特征提取; 机器学习

中图分类号: S435.121.4⁺5; TP391.4 **文献标识码:** A **文章编号:** 1000-1298(2025)06-0218-10 **OSID:** 

Hyperspectral Imaging-based Lightweight Detection Method for Rapid Detection of *Fusarium* Head Blight Severity in WheatLIANG Xiaoying^{1,2} ZHANG Ziting^{1,2} YANG Shuo^{1,2} CHEN Xu^{1,2} YAO Zhifeng^{1,2} SONG Huaibo^{1,2}

(1. College of Mechanical and Electronic Engineering, Northwest A&F University, Yangling, Shaanxi 712100, China

2. Key Laboratory of Agricultural Internet of Things, Ministry of Agriculture and Rural Affairs, Yangling, Shaanxi 712100, China)

Abstract: Aiming to achieve rapid and non-destructive detection of *Fusarium* head blight (FHB) severity levels, the hyperspectral imaging technology combined with machine learning models for analysis was employed. A total of 1 660 wheat ear samples with varying degrees of infection were obtained by inoculating *Fusarium* fungi into the middle grains of wheat ears. Hyperspectral information of the samples was collected by using a hyperspectral imaging system, with the entire wheat ear designated as the region of interest (ROI) to extract average spectral data. By comparing the classification accuracy of four preprocessing methods—normalization, standard normal variate (SNV), multiplicative scatter correction (MSC), and Savitzky-Golay (SG) smoothing derivatives—the SNV algorithm was selected as the optimal preprocessing method. Subsequent analyses were conducted on the SNV-processed spectral data. Feature wavelength selection was performed by using the successive projections algorithm (SPA) and competitive adaptive reweighted sampling (CARS), while dimensionality reduction was implemented via uniform

收稿日期: 2025-03-20 修回日期: 2025-04-24

基金项目: 国家自然科学基金项目(32201662)和陕西省重点研发计划项目(2025NC-YBXM-215)

作者简介: 梁晓颖(2000—),女,硕士生,主要从事农业病害无损检测研究,E-mail: liangxiaoying507@163.com

通信作者: 陈煦(1989—),女,副教授,主要从事农业信息快速感知研究,E-mail: chenxu029@126.com

manifold approximation and projection (UMAP) and linear discriminant analysis (LDA). After comparing these algorithms, LDA was ultimately chosen for its ability to reduce data to three dimensions while maintaining classification accuracy (96.05% for the test set and 94.71% for the training set) and low computational complexity (0.09 s processing time). It was revealed that the critical spectral range for LDA-based FHB severity discrimination lay between 540 nm (chlorophyll reflection peak) and 650 nm (red light absorption valley), attributed to the synergistic effects of rapid chlorophyll degradation and structural tissue damage as infection progresses. A lightweight support vector machine (SVM) model integrating SNV and LDA was developed as the optimal framework for FHB severity classification. The results demonstrated that the proposed algorithm achieved high accuracy with excellent generalization capability, enabling efficient FHB severity assessment. The research result can lay a foundation for future large-scale, real-time field detection of FHB.

Key words: wheat *Fusarium* head blight (FHB); hyperspectral imaging; non-destructive detection; feature extraction; machine learning

0 引言

小麦赤霉病(*Fusarium* head blight, FHB)是一种由禾谷镰刀菌(*Fusarium graminearum*)引起的全球性真菌病害,其流行性爆发具有显著的气候依赖性^[1]。近年来,受气候变暖、秸秆粗放还田导致的田间菌源积累以及镰刀菌抗药性增强等多重因素的影响^[2],该病害的高发地区已由长江中下游延伸至黄淮南部麦区,导致小麦产量损失率高达 15% ~ 40%,同时,脱氧雪腐镰刀菌烯醇(Deoxynivalenol, DON)等真菌毒素污染,对食品安全与人类健康构成双重威胁^[3]。传统 FHB 防控技术主要依赖化学药剂喷施和抗病品种种植,通过及时阻断 FHB 的传播,制约 DON 毒素的产生提升防控效果^[4]。然而,盲目施药会造成双重困境,过量喷施会增加环境负担,加剧病原体抗药性的演变,若施药不足则无法阻隔病害扩散,影响作物生长和产量^[5]。

现有田间调查仍以人工目视评估为主,其主观性强、效率偏低,难以满足对病害严重程度的大面积动态监测需求^[6]。据农业农村部发布的《2025 年全国小麦重大病虫害发生趋势》,全国预计发生小麦赤霉病风险面积约 1.3×10^7 hm²,防控形势严峻。可见,开发智能化检测技术与装备,实现 FHB 严重程度的高效精准分级,提升智慧测报水平已成为构建小麦安全生产及绿色防控体系的关键^[7]。近年来,融合机器学习与深度学习的无损检测技术迅速兴起并广泛应用^[8-9]。然而,RGB 技术对早期潜伏期病害的生化特征敏感性不足,极大地制约了检测能力,导致对轻症感染样本的误判率较高。

高光谱成像技术(Hyperspectral imaging, HSI)作为一种能够同时获取被测对象图像和光谱信息的新兴技术,凭借非破坏性、灵敏度高、信息量大等优点在农业领域中展现出巨大潜力^[10]。林芬芳等^[11]利用堆栈稀疏自编码器深度学习算法,通过增加隐

含层数及隐含神经元数,成功从小麦赤霉病高光谱遥感数据中提取到全方面的深层光谱特征,其赤霉病严重度预测模型决定系数达到 0.927,但模型复杂度较高。为满足田间实时检测需求,兼具精确度和较低复杂度的轻量化模型是目前重要的研究方向之一。DIAO 等^[12]利用改进的 BS-Net-FC 筛选出 15 个特征波段,结合轻量级 CNN 模型实现了玉米幼苗与杂草的快速识别。结果表明,基于特征波段的分类精度相较于全波段下降 3% 左右,但可训练参数减少近 10 倍,显著降低了模型复杂度。沈广辉等^[13]通过在 PLS-DA 模型上对比特征波段与全波段识别小麦赤霉病瘪粒的正确率,发现 CARS 筛选 8 个波段就能达到 100% 的识别准确率,成功降低了模型复杂度,但样本仅限于患病与未患病,对轻微感染的样本存在误判现象。目前,现有的轻量化模型研究多聚焦于植物类型、病害类型和病害有无的判别,对病害严重程度分级缺乏系统性研究,且与病害生理指标的关联机制尚未明确,导致模型泛化能力不足,限制了其在实际田间的广泛应用。

针对上述问题,为平衡分类精度和计算效率,本研究首先通过人工接种获取不同发病程度的病穗样本,采用高光谱成像系统获取样本数据。对比不同降维算法,优选适用于田间部署的低维特征空间,通过关键波段反演与生理指标的关联关系,阐明光谱响应机制。本研究提出的轻量化模型架构,可为田间大面积快速识别 FHB 发生情况和精准施药提供实时、低成本的决策支持,同时为赤霉病智能监测装备的开发提供技术支撑。

1 实验与方法

1.1 实验材料与样本

小麦选择西北农林科技大学植物保护学院提供的 Fielder 品种,将种子种入塑料花盆中,在温室中保持生长环境温度为 25℃、相对湿度为 70%、光照

周期为 12 h 光照和 12 h 黑暗交替进行。待到小麦扬花期时,在 2024 年 8 月 5—8 日陆续接种,根据接种规范^[14]具体操作为:制备禾谷镰刀菌孢子悬浮液($1 \times 10^5 \sim 5 \times 10^5$ 个/mL);选择发育程度一致的穗轴中部,花药完全伸出,颖壳开放角度大于 45°的小麦麦穗;接种前,充分摇晃震荡孢子悬浮液,以确保孢子分散均匀。采用单小花滴注法接种,注入孢子悬浮液 10 μ L,在每穗选择的小穗处接种并对接种点进行标记;对接种小穗处套上保鲜袋保湿 2 d。

从接种 16 d 开始,为选取不同患病程度的小麦样本,在 2024 年 8 月 22、26 日,分别依据 GB/T 3543.5—2022《小麦赤霉病田间调查规范》选取感染赤霉病且发病程度不同的小麦穗作为实验样本,剔除机械损伤或外界干扰个体,最终构建四级病害样本库共 255 穗。

1.2 高光谱数据采集

分别对实验选取的麦穗样本采集其高光谱和 RGB 图像数据,在西北农林科技大学农业农村部农业物联网重点实验室进行数据采集。

高光谱图像采用美国 Surface Optics 公司生产的 SOC-710VP 型成像仪采集,该成像仪光谱范围为 400 ~ 1 000 nm,涵盖可见光到近红外波段,光谱分辨率为 2.1 nm。系统配备卤素灯阵列(100 W $\times$ 3)提供均匀照明,光源入射角 45°以避免镜面反射干扰。镜头类型为 C-Mount,固定焦距为 50 mm,镜头光圈调整至 F2.1,调整镜头与样本之间距离为 55 cm。数据采集通过 HyperScanner 软件控制,默认为低增益模式,设置积分时间(24 ms),扫描模式(Scanner settings)选择快速(Quick)模式,扫描时间 5 ~ 10 s,空间像素为 696 像素 $\times$ 520 像素,光谱通道数为 256。同步利用 SRAna710e 软件将原始 DN 值转换为标准反射率数据^[15]。

为减少环境干扰,实验在温湿度恒定的密闭暗室中进行,并在采集图像前提前打开高光谱系统预热 30 min。高光谱图像采集方式如图 1 所示。在采集过程中,首先将样本放置于电动升降台上,麦穗以穗轴垂直于扫描方向固定,每组 3 穗间隔 2 cm 平铺于黑色天鹅绒背景板(反射率小于 0.02)上,确保单次扫描覆盖完整穗部。每穗从 0°、90°、180°、270° 4 个正交角度重复采集,共获取 225 穗的 4 个角度共 900 组高光谱立方体数据,以提高数据的代表性和可靠性。

RGB 图像采集与高光谱扫描同步进行,用以对麦穗赤霉病感染严重度进行辅助分级,图像保存为 RAW 格式,后期通过 Adobe Lightroom 统一白平衡与曝光参数,确保视觉分级结果与光谱数据时空

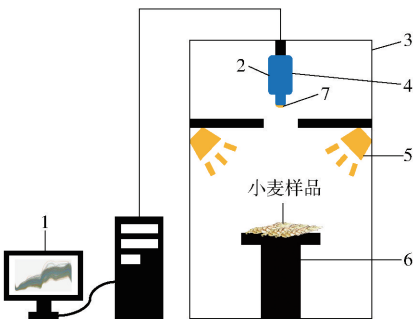


图 1 高光谱检测系统示意图

Fig. 1 Hyperspectral system schematic

1. 计算机 2. 相机 3. 暗箱 4. 光谱仪 5. 卤素灯 6. 升降台
7. 镜头

对齐。

根据 GB/T 15796—2011《小麦赤霉病测报技术规范》,以病小穗占总小穗的比例作为分级依据,将赤霉病严重程度分为 4 个等级^[16],如表 1 所示。为了对麦穗样本的赤霉病感染严重度级别进行准确判定,本研究采用多源数据融合方法实现病害的自动化分级。具体为:对 4 个角度的病穗面积占比取算术平均值作为该麦穗的最终病害程度指标,并根据计算得到的病害程度指标参照国家标准中的分级阈值将样本划分为 4 个严重度等级,结果与农艺专家目视评估结果一致。为避免样本不均导致的模型训练出现偏差,采用几何变换(旋转 180°)、图像亮度调节(0.7 ~ 1.4)和添加高斯噪声等方式^[17]对样本进行扩充,扩充前后的样本分布如表 1 所示。

表 1 单穗小麦赤霉病严重度分级标准与扩充前后的样本严重度分级结果

Tab. 1 FHB severity grading standards and augmented sample results

赤霉病等级	症状	扩充前样本数	扩充后样本数
1 级	病小穗数占全部小穗的 1/4 以下	12	372
2 级	病小穗数占全部小穗的 1/4 ~ 1/2	196	392
3 级	病小穗数占全部小穗的 1/2 ~ 3/4	448	448
4 级	病小穗数占全部小穗的 3/4 以上	224	448

1.3 高光谱数据校正

为消除传感器本身的系统误差和暗电流噪声,首先对原始 DN 值进行辐射校正^[18],其校正公式为

$$L_{\lambda} = G_{\lambda} DN + B_{\lambda}$$

(1)

式中 L_{λ} ——校正后的辐射亮度

G_{λ} ——增益系数

B_{λ} ——偏移量

反射率计算采用标准参考板法^[19]。标准参考板在相同光照条件下具有稳定的反射特性,其反射率 $R_{ref,\lambda}$ 已知。通过对比样本与参考板的辐射亮度

计算样本的反射率,即

$$R_{\lambda} = \frac{L_{sample,\lambda}}{L_{ref,\lambda}}R_{ref,\lambda}$$

(2)

式中 R_{λ} ——样本在波长 λ 处反射率

$L_{sample,\lambda}$ ——样本辐射亮度

$L_{ref,\lambda}$ ——参考板辐射亮度

通过上述辐射校正和标准化处理,有效消除了仪器系统误差和环境干扰因素,能够准确反映样本的光谱特征。

最后使用 ENVI 5.6 软件对高光谱图像进行感兴趣区域 (Region of interest, ROI) 提取,图 2 为小麦麦穗健康、患病以及整体的 ROI 高光谱反射率示意图。具体 ROI 提取步骤如图 3 所示,将整个麦穗区域定义为 ROI,并将其裁剪为高度420像素、宽度

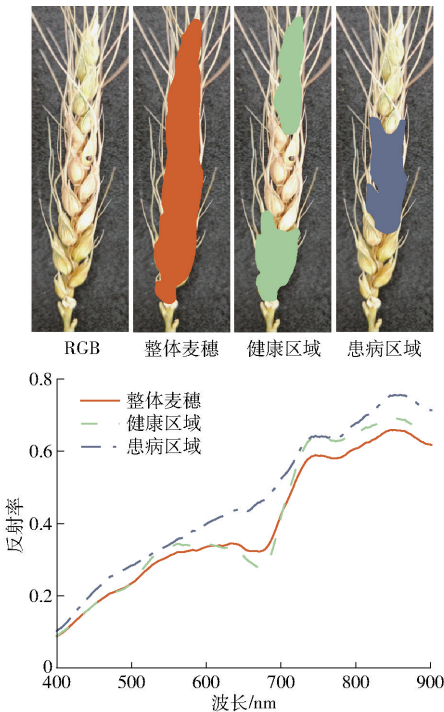


图2 小麦赤霉病健康、患病和整体平均光谱反射率变化曲线

Fig.2 Spectral reflectance comparison: healthy vs. diseased vs. overall FHB

150 像素、光谱通道数 256 的高光谱图像,随后对裁剪后的图像进行掩膜处理,通过二值化掩膜技术去除无效背景区域,最终仅保留 ROI 内的麦穗目标区域,以确保后续分析仅针对有效区域进行。

1.4 小麦赤霉病高光谱数据预处理

为消除仪器噪声、光照不均、麦芒及麦穗表面曲率引起的散射干扰,提高数据的质量和后续分析的准确性,本研究对比了归一化 (Normalization)、标准正态变量变换 (Standard normal variate, SNV)、多元散射校正 (Multiplicative scatter correction, MSC) 和平滑导数 (Savitzky-golay, SG) 4 种预处理方法^[20-23],处理效果通过建模后评价指标综合评估。

1.5 小麦赤霉病高光谱数据降维

由于高光谱数据每个像素点都包含连续的光谱信息,但相邻波段间信息相关性高^[24],为去除冗余信息,探究波段间强相关性以及样本标签信息的可用性,本文对比统一流形逼近与投影 (Uniform manifold approximation and projection, UMAP)、线性判别分析 (Linear discriminant analysis, LDA)、连续投影算法 (Successive projections algorithm, SPA) 和竞争性自适应重加权采样 (Competitive adaptive reweighted sampling, CARS) 的降维效果。其中,UMAP 是一种基于流形学习的无监督降维方法,通过同时保持数据局部邻域结构与全局拓扑特性,实现将高维数据到低维空间的非线性映射^[25-26];LDA 则作为有监督降维方法,通过投影将高维数据映射到低维空间,同时最大化类间方差并最小化类内方差达到降维目的^[27];SPA 是一种通过向量的投影分析逐步选择特征的前向特征选择方法^[28];CARS 结合了蒙特卡洛采样 (Monte Carlo sampling) 和偏最小二乘回归 (Partial least squares regression, PLS) 模型的回归系数,通过迭代过程不断调整样本权重,最终选择出对模型预测最有贡献的特征变量^[29]。

为客观评估降维效果,分别对 UMAP、LDA、SPA 及 CARS 的降维结果进行 t 分布随机邻域嵌入

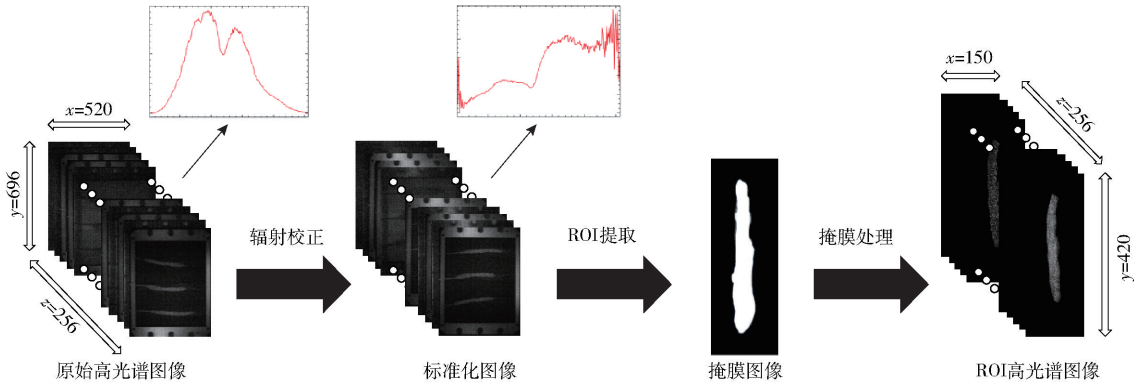


图3 FHB 高光谱 ROI 数据采集

Fig.3 FHB hyperspectral ROI data acquisition

(t-distributed stochastic neighbor embedding, t-SNE)^[30] 三维可视化,通过计算基于马氏距离的类间距离与类内距离比值进行对比分析,结合观测不同方法在低维空间的类簇可分性以及在不同模型中的表现程度,利用 SVM 模型综合评判降维方法的可行性以及深入探寻影响等级判别的特征波段。

1.6 分类模型建立

为综合考虑模型复杂度和计算效率,本研究采用传统机器学习支持向量机 (Support vector machine, SVM)^[31] 进行建模分析,旨在保持可解释性的同时,更高效地挖掘光谱的内在特征。为提高模型性能,首先通过网格搜索 (Grid search) 系统遍历超参数空间,结合 5 折分层交叉验证 (Stratified 5-Fold CV) 确保参数选择的最优性。继而根据准确率 (Accuracy)、召回率 (Recall) 和 F1 分数 (F1 score) 等评价指标最终构建兼顾效率与精度的光谱分类模型。

2 结果与讨论

2.1 光谱反射率特征

小麦赤霉病样本的反射率曲线以及 4 种严重程度级别平均光谱反射率如图 4 所示。

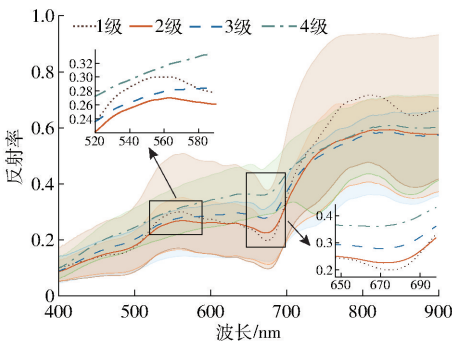


图 4 小麦赤霉病平均光谱反射率和不同严重程度级别平均光谱反射率

Fig. 4 Spectral reflectance and severity level of FHB

由图 4 可看出,不同病害等级样本的反射率曲线在特定波段呈规律性响应,小麦麦穗在 550 nm 附近存在明显反射峰,这是由于叶绿素对绿光的吸收较弱,未被吸收的光被反射形成峰值^[32],随着病害级别逐渐加重,该波段处的反射峰值呈降低趋势,表明叶绿素的持续降解,造成光合作用能力衰退。4 级样本的绿峰反射率反弹至 0.36,高于 1 级样本,可能是当病害程度进一步增强时,组织坏死导致的细胞结构崩解,增加了光散射效应^[33]。与此同时,叶绿素吸收谷随病害发展,在 650 ~ 680 nm 红光区域逐渐填平,红光吸收谷的减弱主要归因于叶绿素含量减少导致的色素浓度下降^[34]。图 4 中具有显著变化的波段位于 680 ~ 750 nm,红边位置发生了

偏移,红边蓝移 (Blue shift)^[35] 表明了叶绿素降解加快和麦穗衰老加剧。随着病害级别加重,该区域反射率显著提升且红边斜率逐渐减小,这是由于麦穗细胞结构被破坏,组织开始松散化^[36];在近红外 750 ~ 1 000 nm 波段,由于患病麦穗细胞壁破裂、海绵组织塌陷,从而导致光散射减少,反射能力减弱。但 1 级样本麦穗的反射率最高,可能由于病害初期会触发植物应激反应以及防御反应,导致接种部位麦穗通过加厚细胞壁、增加表皮蜡质或产生抗病化合物来抵御病原体增强了散射,而其他组织仍保持正常^[37-38]。970 nm 左右为水敏感波段,小麦麦穗随着病害的加重,反射率的上升需结合水分与细胞结构变化的综合效应,严重脱水更可能降低近红外反射率^[39]。

2.2 光谱预处理

不同预处理后小麦赤霉病光谱反射率数据去除前后噪声保留 400 ~ 900 nm 波段进行后续处理。基于 SVM 的交叉验证结果如表 2 所示,表明不同预处理方法对模型性能产生不同影响。首先,不同方法处理后的 F1 分数与准确率差异均小于 0.1 个百分点,说明数据增强后的样本数据类别达到均衡。其中,SG 平滑因对高频细节的抑制导致训练集与测试集的准确率 (94.79%、91.18%) 均表现不佳;原始光谱和归一化处理虽在训练集呈现高准确率,但测试集性能明显下降,其训练-测试落差分别达到 6.89% 和 5.38%,表明这两种方法均存在过拟合风险。

表 2 基于不同处理后小麦赤霉病光谱反射率数据的 SVM 建模结果

Tab.2 SVM modeling results using preprocessed FHB spectral data					%
预处理	训练集		测试集		训练-测试落差
	准确率	F1 分数	准确率	F1 分数	
原始光谱	97.98	97.98	91.18	91.09	6.89
归一化	97.14	97.13	91.76	91.63	5.38
SNV	97.98	97.98	94.51	94.47	3.47
MSC	95.54	95.53	94.31	94.27	1.23
SG	94.79	94.75	91.18	91.12	3.61

MSC 通过散射校正有效缓解麦穗麦芒造成的光谱散射效应,其训练集与测试集准确率落差较小,仅为 1.23%,体现出较好的泛化稳定性,但准确率受限于样本光谱参考的病理特异性。SNV 预处理通过全局标准化,有效消除不同麦穗表面曲率的差异,在训练集和测试集上均达到最高准确率 (97.98% 和 94.51%),分别较 MSC 提升 2.44、0.2 个百分点,且训练-测试落差 (3.47%) 显著低于原

始光谱和归一化方法。

综上所述, 权衡精度与鲁棒性, 本研究选择 SNV 作为核心预处理方法, 为后续数据降维奠定低偏置数据基础。

2.3 数据降维

2.3.1 维度选择

为了在保持分类性能的同时确定最小特征维度, 本研究对比了有监督和无监督算法对经 SNV 预处理的数据进行光谱降维, 包括 SPA、CARS、UMAP 和 LDA。通过比较这些降维方法在 SVM 模型测试集上的分类准确率, 筛选出最优的特征数量。

图 5 表示不同特征选择算法在维度优化过程中的性能差异情况。在波长特征选择算法中, SPA 在选取 26 个特征波长时达到最优性能, 测试集准确率为 94.71%; CARS 在 21 个特征波长时测试集准确率为 92.94%, 性能次于 SPA 算法。在降维算法中, 无监督算法 UMAP 在降至 11 维时达到性能峰值 (90.39%), 相比原始光谱测试集准确率下降 4.32 个百分点; 有监督算法 LDA 则通过最大化类间可分性, 在 3 维投影空间 (类数 - 1 准则) 即可实现 94.71% 测试准确率, 与原始 256 维数据精度 (94.51%) 相比提升 0.2 个百分点, 且特征维度压缩率达 98.8%。对比较低性能的 UMAP 说明无监督方法在光谱数据降维时存在一定的信息损失风险, 有监督降维方法在维度与性能方面均具有优势。分析监督降维的维度效率优势, 主要表现在 LDA 的判别准则, 即通过构建类间散布矩阵与类内散布矩

阵的广义特征分解, 前 3 个特征向量已包含 98.6% 的类别判别信息。相较之下, UMAP 虽能保持数据拓扑结构, 但其无监督特性导致与病害等级相关的生物特征响应在降维过程中出现了明显的权重衰减。

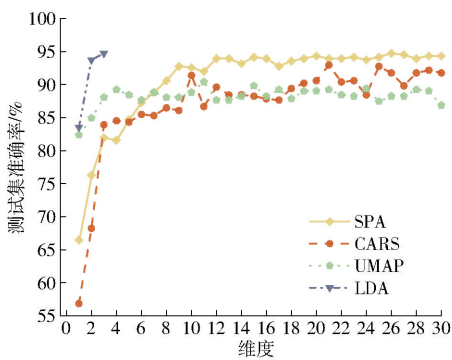


图 5 不同维度下 SVM 测试集准确率变化曲线
Fig. 5 Variation trend of SVM test set accuracy across different feature dimensions

2.3.2 重要性选择

为解析和揭示 FHB 等级判别的关键光谱响应机制, 本研究对 4 种降维算法过程中选择的特征进行重要性评估, 其重要性分布特性如图 6 所示。

SPA 基于正交投影残差最大化的前向选择策略, 首先选取光谱能量最大的初始波长, 随后迭代选择与前序特征正交投影残差最大的波长, 最终通过正交投影残差值量化各波段重要性, 如图 6a 所示。由于正交约束下对光谱全频段信息的系统性覆盖, 筛选出的 26 个特征波长近似均匀分布, 表明 FHB

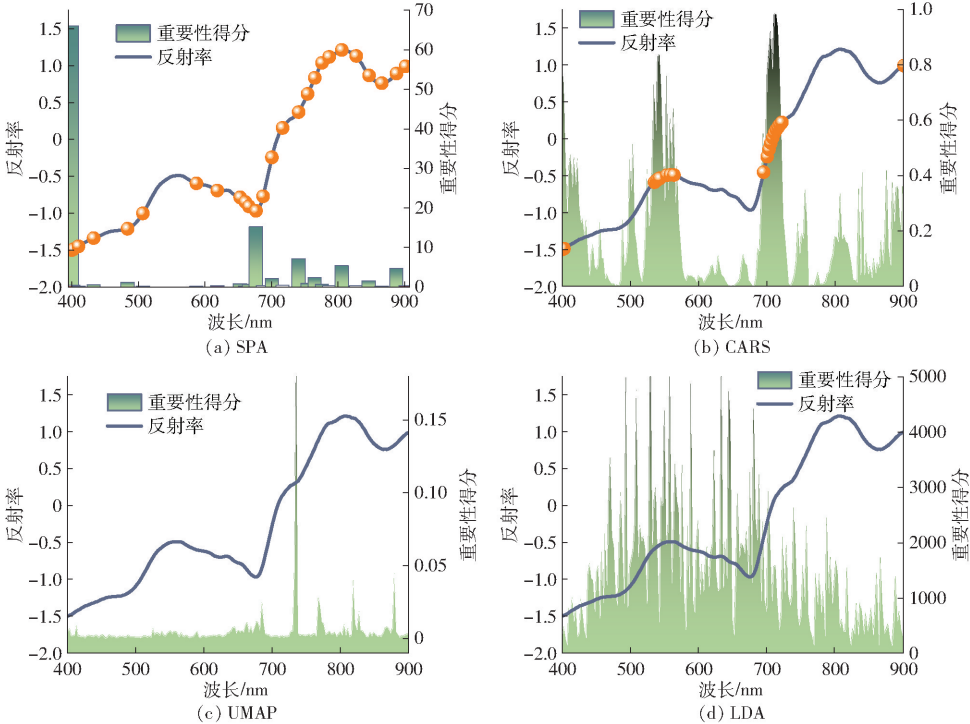


图 6 特征波段的选择及不同算法下波长重要性分布

Fig. 6 Feature band selection and wavelength importance by algorithm

引发的光谱扰动具有广谱性,涉及叶绿体降解(550 nm)、细胞壁多糖分解(780 nm)等多病理进程协同作用。

CARS通过迭代随机抽取80%样本构建PLS模型,以回归系数的绝对值作为特征重要性指标,经50次迭代累加后获得各波长重要性评分。如图6b所示,CARS最终保留的21个特征波长集中在叶绿素反射波段520~580 nm和红边蓝移区域680~740 nm,与叶绿素a/b下降及红边位置偏移量显著相关。

UMAP将256维光谱数据映射至11维空间后,各波长的重要性通过构建随机森林(Random forest, RF)模型计算各波长在决策树节点分裂过程中产生的基尼不纯度下降值,量化其在非线性特征空间中的判别贡献。如图6c所示,贡献度最高的波长集中在740 nm近红外波段附近,对应海绵组织塌陷引发的散射特性发生变化,但其无监督特性导致可见光区生物标志响应被弱化。

LDA通过最大化类间散布与类内散布的比值计算投影方向,将其权重系数直接作为各波长对类别线性可分的贡献度。如图6d所示,450~700 nm区间内存在连续的重要性分布,其中在540 nm叶绿素反射波段和650 nm类胡萝卜素吸收波段附近出现局部极大值。说明在判别FHB严重度等级时,由于患病过程中叶绿素与类胡萝卜素的变化,主要光

谱特征位于绿光和红光区域。

上述4种降维算法在900~1 000 nm无选择,说明在水分胁迫波段仍存在感知盲区,分析原因可能为数据增强过程中人工湿度控制削弱了自然状态下的水分胁迫梯度。

为进一步验证上述发现,结合表3基于马氏距离的类间距离与类内距离比值和图7 t-SNE 3D可视化后进行对比分析。

表 3 不同降维方法处理后类间距离与类内距离比值

Tab.3 Inter/intra-class distance ratios after different dimensionality reduction methods

降维算法	SPA	CARS	UMAP	LDA
类间距离/类内距离	1.318 9	1.504 3	0.537 9	2.479 1

由表3可以看出,LDA降维后4类样本在投影空间形成明显区分,其类间距离与类内距离比值2.479 1显著优于UMAP(0.537 9);而SPA与CARS虽能保持较高分类精度,但特征波长冗余导致投影面存在维度交叉干扰。该结果证实,LDA降维方法可为田间便携式传感器波段优化提供精准靶点。

通过LDA投影权重谱(图8)及总重要性分布(图9)联合分析光谱响应机制,其中每个LDA成分对应一个判别方向,纵坐标表示该波长在投影方向

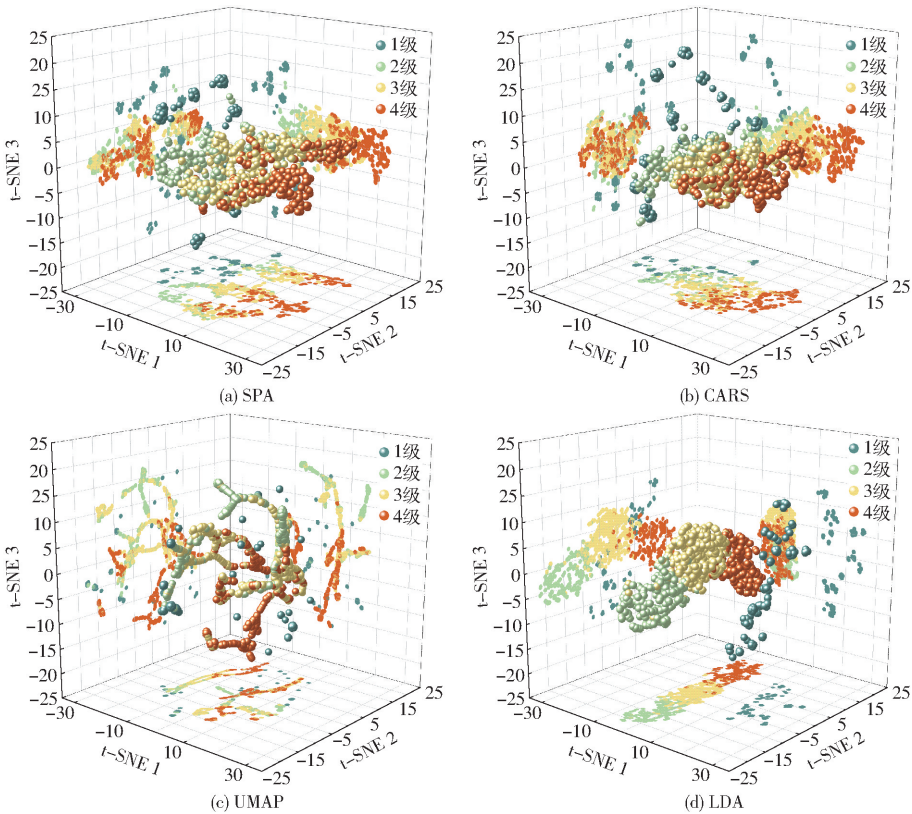


图 7 t-SNE 3D 可视化
Fig.7 t-SNE 3D visualization

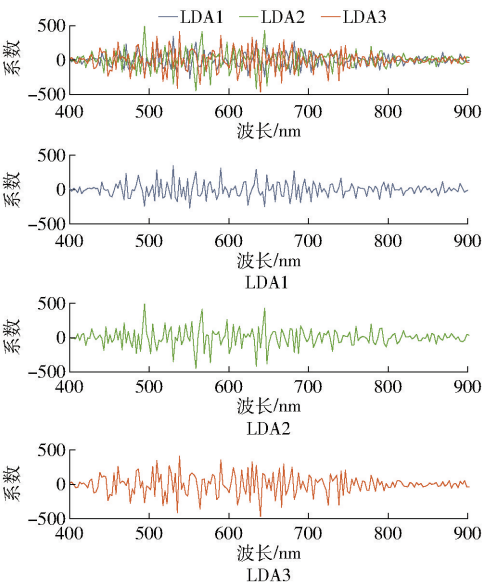


图 8 LDA 投影方向可视化

Fig. 8 LDA projection directions visualization

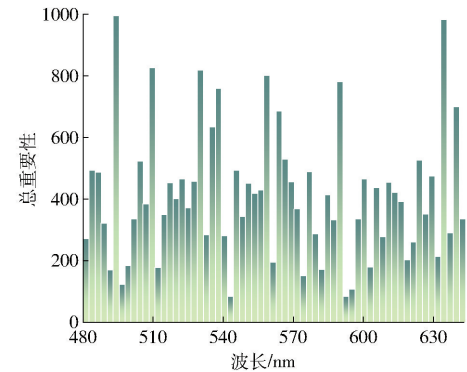


图 9 480 ~ 640 nm 波长总重要性 (LDA)

Fig. 9 480 ~ 640 nm total importance (LDA)

的标准化权重,系数值越大代表该波段对类别区分的判别能力越强。

对比图 8 可以看出,LDA1 在 509.5 nm (蓝光吸收区)和 634.5 nm (红光吸收谷边缘)处具有最高权重,分别为 286.19 和 290.09,主要集中在可见光与近红外波段,其判别贡献源于叶绿素含量随病害发展的特异性衰减与叶片结构损伤;LDA2 在 494.1、566.5、645.1 nm 达到极大值,在 558.7、576.8、634.5、647.7 nm 达到极小值,500 nm 左右对应病原侵染初期类胡萝卜素吸收特性,600 nm 反映叶肉细胞排列紊乱或光系统 II 损伤,与前文分析的在该波段下 4 级较 1 级峰值幅度增加相吻合;LDA3 在 509.5 nm (蓝光)和 634.5 nm (红光)处权重最高,分别为 347.00 和 275.87,主要集中在 500 ~ 550 nm 和 630 ~ 680 nm 区间,与叶绿素吸收和光系统损伤 II 相关。400 ~ 460 nm 和 680 ~ 900 nm 区间内系数变化较为平稳,在降维过程中权重较低,表明该区域对当前病害级别的区分贡献较小。

图 9 为总重要性峰值在 480 ~ 640 nm 区间发生

剧烈震荡,将每个波长下的 LDA1、LDA2 和 LDA3 的绝对值相加计算的总重要性分布情况。在 494.1 nm 处总重要性达到最高 (994.35),位于近红外波段,该波段在 LDA2 和 LDA3 中权重较高,可能捕捉到了细胞结构损伤或病害引起的代谢产物变化;在 634.5 nm 总重要性为次高 (983.67),处于红光波段,LDA1 和 LDA3 有较强贡献度,由于病害发生过程中叶绿素会发生不同程度的降解,红光波段能够敏感感知叶绿素含量变化;509.5、530.2、558.7、589.9 nm 总重要性均位于 800 附近,能够敏感感知在绿光波段范围中病害所引起的颜色变化,随着病害程度加重,叶绿素含量逐渐减少和类胡萝卜素含量逐渐升高会导致反射率发生显著变化。值得注意的是,近红外波段 (680 ~ 900 nm) 权重普遍较低,可能与实验环境湿度恒定弱化了自然条件下水分胁迫的光谱响应有关。

综合来看,绿光和红光波段的权重较高,且高权重波长如 634.5 nm 与叶绿素吸收峰高度重叠,说明基于 LDA 的降维方法能够同步捕获叶绿素代谢、细胞壁修饰及光能转化效率下降等多尺度病理进程,以此为后续病害发生程度的精准判别提供理论依据,同时为开发面向早期诊断的便携式检测传感器研发提供技术支撑。

2.4 模型建立

表 4 为不同降维方法下 SNV 预处理后数据的 SVM 建模结果。SPA 在选择 26 个特征波长的情况下训练集和测试集的准确率分别为 97.81% 和 94.71%,均高于 CARS 选择 21 个特征波长之后的准确率 (96.30% 和 92.94%)。UMAP 降至 11 维时测试集准确率高达 98.15%,但测试集准确率仅有 90.39%,具有一定的过拟合风险,泛用性不高。LDA 的训练集准确率虽非最高仅有 96.05%,但其测试集的准确率与 SPA 持平,具有良好的泛用性。其次 LDA 低至 3 维,远低于其他方法,能有效地减少模型复杂度同时保持训练和预测速度及准确率。

表 4 基于不同降维方法下 SNV 预处理后数据的 SVM 建模结果

Tab. 4 SVM model results of SNV-preprocessed data under different dimensionality reduction methods

降维方法	维度	训练集			测试集			时间复杂度/s
		准确 率/%	召回 率/%	F1 分 数/%	准确 率/%	召回 率/%	F1 分 数/%	
SPA	26	97.81	97.81	97.81	94.71	94.71	94.65	2.94
CARS	21	96.30	96.30	96.31	92.94	92.94	92.90	0.18
UMAP	11	98.15	98.15	98.14	90.39	90.39	90.37	12.68
LDA	3	96.05	96.05	96.05	94.71	94.71	94.68	0.09

在 1 660 × 256 数据集,对应的最佳维度为选择

特征数,进行重复 5 次实验计算平均运行时间作为时间复杂度。UMAP 在降维过程中不仅准确率较低,且时间复杂度最高达到 12.68 s;CARS 的运行时间较短(0.18 s)但分类准确率小于 SPA 与 LDA;在 SPA 和 LDA 中,时间复杂度为 2.94 s 和 0.09 s,LDA 远小于其他 3 种算法。上述结果表明,在有限样本与噪声干扰并存的田间场景下,LDA 在测试集性能、泛化能力和降维效率 3 方面达到最佳平衡,能够满足田间赤霉病在线检测系统的毫秒级响应需求。

3 结论

(1)融合高光谱成像技术与数据降维算法,构建了小麦赤霉病严重程度的轻量化判别模型,实现了 FHB 的高精度快速无损分级。

(2)通过对比多种预处理方法,明确 SNV 能够有效消除噪声干扰、提升光谱数据质量。在降维方法选择中,对比各算法最佳维度下的病害程度识别准确率和模型复杂度,LDA 在准确率、泛化能力以及复杂度方面均表现较优。进一步揭示了该降维算法能在 540 nm(叶绿素反射峰)至 650 nm(红光吸收谷)波段区间对 FHB 严重度级别进行显著区分,综合反映了叶绿素降解、光系统 II 损伤及叶片结构变化的协同效应。

(3)在 SNV 预处理和 LDA 降维算法之后,构建了满足田间实时检测需求的 FHB 严重度分类模型,训练集和测试集准确率最高可达 96.05% 和 94.71%,时间复杂度仅为 0.09 s,实现在显著降低模型时间复杂度的同时保留了关键病害信息,大幅降低计算资源需求。

参 考 文 献

- [1] SUN X, YANG R, TANG H, et al. Diversity and pathogenicity of *Fusarium* species associated with *Fusarium* head blight in wheat and maize cropping systems in Sichuan Province[J]. Scientific Reports, 2025, 15(1): 5984.
- [2] 李兵, 梁晋刚, 朱育攀, 等. 我国小麦赤霉病成灾原因分析及防控策略探讨[J]. 生物技术进展, 2021, 11(5): 647–652.
LI Bing, LIANG Jin'gang, ZHU Yupan, et al. Epidemiological analysis and management strategies of *Fusarium* head blight of wheat[J]. Current Biotechnology, 2021, 11(5): 647–652. (in Chinese)
- [3] 吴海霞, 谷莉莉, 吴佳文, 等. 小麦赤霉病化学农药减量控害的技术途径[J]. 江苏农业科学, 2021, 49(11): 76–79.
- [4] MORELLOS A, PANTAZI X E, ALMOUJAHED M B, et al. Forecasting of *Fusarium* head blight spatial distribution in winter wheat using machine learning[J]. Computers and Electronics in Agriculture, 2025, 231: 10996.
- [5] LI L, DONG Y, XIAO Y, et al. Combining disease mechanism and machine learning to predict wheat *Fusarium* head blight[J]. Remote Sensing, 2022, 14(12): 2732.
- [6] 徐玮, 周佳良, 钱啸, 等. 基于 AR 眼镜和改进 YOLOv8m-seg 的田间小麦赤霉病严重度识别方法[J]. 农业大数据学报, 2024, 6(4): 497–508.
- [7] 全国农技中心组织召开 2025 年全国农作物重大病虫害发生趋势会商会[J]. 中国植保导刊, 2024, 44(12): 107.
- [8] 李志伟, 霍静琦, 蒿晟昆. 机器视觉技术在农业智能装备中应用的研究进展[J]. 当代农机, 2021(7): 5–7.
- [9] 韩文霆, 崔家伟, 崔欣, 等. 基于特征优选与机器学习的农田土壤含盐量估算研究[J]. 农业机械学报, 2023, 54(3): 328–337.
HAN Wenting, CUI Jiawei, CUI Xin, et al. Estimation of farmland soil salinity content based on feature optimization and machine learning algorithms[J]. Transactions of the Chinese Society for Agricultural Machinery, 2023, 54(3): 328–337. (in Chinese)
- [10] 雷雨, 韩德俊, 曾庆东, 等. 基于高光谱成像技术的小麦条锈病病害程度分级方法[J]. 农业机械学报, 2018, 49(5): 226–232.
LEI Yu, HAN Dejun, ZENG Qingdong, et al. Grading method of disease severity of wheat stripe rust based on hyperspectral imaging technology[J]. Transactions of the Chinese Society for Agricultural Machinery, 2018, 49(5): 226–232. (in Chinese)
- [11] 林芬芳, 陈星宇, 周维勋, 等. 基于堆栈稀疏自编码器的小麦赤霉病高光谱遥感检测[J]. 作物学报, 2023, 49(8): 2275–2287.
LIN Fenfang, CHEN Xingyu, ZHOU Weixun, et al. Hyperspectral remote sensing detection of *Fusarium head blight* in wheat based on the stacked sparse auto-encoder algorithm[J]. Acta Agronomica Sinica, 2023, 49(8): 2275–2287. (in Chinese)
- [12] DIAO Z, YAN J, HE Z, et al. Corn seedling recognition algorithm based on hyperspectral image and lightweight-3D-CNN[J]. Computers and Electronics in Agriculture, 2022, 201: 107343.
- [13] 沈广辉, 曹瑶瑶, 刘馨, 等. 近红外高光谱成像结合特征波长筛选识别小麦赤霉病瘪粒[J]. 江苏农业学报, 2021, 37(2): 509–516.
SHEN Guanghui, CAO Yaoyao, LIU Xin, et al. Identification of *Fusarium* damaged kernels using near infrared hyperspectral imaging and characteristic bands selection[J]. Jiangsu Journal of Agricultural Sciences, 2021, 37(2): 509–516. (in Chinese)
- [14] 史文琦, 刘美玲, 郑磊, 等. 55 个小麦主栽及后备品种抗赤霉病评价与籽粒 DON 毒素含量分析[J]. 植物保护, 2024, 50(6): 262–269, 320.
SHI Wenqi, LIU Meiling, ZHENG Lei, et al. Evaluation of resistance to *Fusarium* head blight and analysis of DON toxin content in kernel among 55 wheat cultivars and reserve varieties[J]. Plant Protection, 2024, 50(6): 262–269, 320. (in Chinese)
- [15] PLAZA A, BENEDIKTSSON J A, BOARDMANET J W, et al. Recent advances in techniques for hyperspectral image processing[J]. Remote Sensing of Environment, 2007, 113(Supp. 1): S110–S122.
- [16] 中华人民共和国农业部. 小麦赤霉病测报技术规范: GB/T 15796—2011[S]. 北京: 中国标准出版社, 2011.
- [17] 丰阳, 罗天, 戴元杰, 等. 基于正样本重构对比学习的小样本高光谱图像分类研究[J]. 成都工业学院学报, 2024, 27(4): 34–38.

FENG Yang, LUO Tian, DAI Yuanjie, et al. Hyperspectral image classification with limited samples based on the positive samples reconstruction and contrast learning[J]. Journal of Chengdu Technological University, 2024, 27(4): 34 – 38. (in Chinese)

[18] HOLTZ U S T. Introductory digital image processing: a remote sensing perspective, third edition[J]. Environmental & Engineering Geo Science, 2007, 13(1): 89 – 90.

[19] 居雷,于洁,吴炎森,等. 高光谱预处理方法与多模型在分类判别中的对比研究[J]. 光谱学与光谱分析, 2025, 45(1): 125 – 132.

[20] GAI Z, AUN L, BAI H, et al. Convolutional neural network for apple bruise detection based on hyperspectral[J]. Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy, 2022, 279: 121432.

[21] 赵昕,郑树亮,牛晓颖,等. 基于高光谱成像技术和近红外光谱技术的金冠苹果货架期判别及其品质分析[J]. 食品工业科技, 2025, 46(11): 1 – 11.

ZHAO Xin, ZHENG Shuliang, NIU Xiaoying, et al. Shelf life identification and quality analysis of golden delicious apples based on hyperspectral imaging and near infrared spectroscopy[J]. Science and Technology of Food Industry, 2025, 46(11): 1 – 11. (in Chinese)

[22] 刘静,汪泓,张磊,等. 基于高光谱的辣椒叶片氮素含量反演[J]. 中国农业科学, 2025, 58(2): 252 – 265.

LIU Jing, WANG Hong, ZHANG Lei, et al. Inversion of nitrogen content in chili pepper leaves based on hyperspectral analysis[J]. Scientia Agricultura Sinica, 2025, 58(2): 252 – 265. (in Chinese)

[23] 张春琪,刘梦姝,晁金阳,等. 基于特征波长的接触式作物叶绿素检测系统[J]. 农业机械学报, 2024, 55(增刊2): 255 – 262.

ZHANG Chunqi, LIU Mengshu, CHAO Jinyang, et al. Contact-based crop chlorophyll detection system based on feature wavelengths[J]. Transactions of the Chinese Society for Agricultural Machinery, 2024, 55(Supp. 2): 255 – 262. (in Chinese)

[24] ZHANG L, ZHANG S, WEI Y, et al. Maize seed variety identification using hyperspectral imaging and self-supervised learning: a two-stage training approach without spectral preprocessing[J]. Expert Systems With Applications, 2024, 238(PF): 122113.

[25] BECHT E, MCLNNES L, HEALY J, et al. Dimensionality reduction for visualizing single-cell data using UMAP[J]. Nature Biotechnology, 2018, 37(1): 38.

[26] MCINNES L, HEALY J, MELVILLE J. UMAP: uniform manifold approximation and projection for dimension reduction[J]. arXiv Preprint, arXiv:1802.03426, 2018.

[27] LI X, MA Z, PENG P, et al. Supervised latent Dirichlet allocation with a mixture of sparse softmax[J]. Neurocomputing (Amsterdam), 2018. 312: 324 – 335.

[28] 杨林婧,杨莎,张圣杨,等. 农田土壤有机碳高光谱特征及定量监测研究[J]. 激光生物学报, 2024, 33(4): 316 – 325.

YANG Linjing, YANG Sha, ZHANG Shengyang, et al. Hyperspectral characteristics and quantitative monitoring of soil organic carbon in farmland[J]. Acta Laser Biolgy Sinica, 2024, 33(4): 316 – 325. (in Chinese)

[29] 刘玮佳,张晓彤,杨睿,等. 基于特征波段选择的芦苇 LAI 高光谱遥感估测[J]. 华北理工大学学报(自然科学版), 2025, 47(1): 79 – 87.

LIU Weijia, ZHANG Xiaotong, YANG Rui, et al. Hyperspectral remote sensing estimation of LAI reed based on characteristic band selection[J]. Journal of North China University of Science and Technology (Natural Science Edition), 2025, 47(1): 79 – 87. (in Chinese)

[30] 殷秀丽,谢丽蓉,杨欢,等. 特征选择与 t-SNE 结合的滚动轴承故障诊断[J]. 机械科学与技术, 2023, 42(11): 1784 – 1793.

YIN Xiuli, XIE Lirong, YANG Huan, et al. Rolling bearing fault diagnosis combined feature selection with t-distributed stochastic neighbor embedding[J]. Mechanical Science and Technology for Aerospace Engineering, 2023, 42(11): 1784 – 1793. (in Chinese)

[31] 韩青妍,崔洪博,刘燕妮,等. 基于高光谱反射率估测分蘖洋葱叶片净光合速率的研究[J]. 东北农业科学, 2024, 49(6): 83 – 88.

HAN Qingyan, CUI Hongbo, LIU Yanni, et al. Research on estimating the net photosynthetic rate of tillering onion leaves based on hyperspectral reflectance[J]. Journal of Northeast Agricultural Sciences, 2024, 49(6): 83 – 88. (in Chinese)

[32] MAHLEIN A, OERKE E, STEINER U, et al. Recent advances in sensing plant diseases for precision crop protection[J]. European Journal of Plant Pathology, 2012, 133(1): 197 – 209.

[33] CARTER G A, KNAPP A K. Leaf optical properties in higher plants: linking spectral characteristics to stress and chlorophyll concentration[J]. American Journal of Botany, 2001, 88(4): 677 – 84.

[34] BLACKBURN G A. Hyperspectral remote sensing of plant pigments[J]. Journal of Experimental Botany, 2006, 58(4): 855 – 867.

[35] ZHU L, CHEN Z, WANG J, et al. Monitoring plant response to phenanthrene using the red edge of canopy hyperspectral reflectance[J]. Marine Pollution Bulletin, 2014, 86(1 – 2): 332 – 341.

[36] ZHANG D, WANG Q, LIN F, et al. Development and evaluation of a new spectral disease index to detect wheat Fusarium head blight using hyperspectral imaging[J]. Sensors, 2020, 20(8): 2260 – 2260.

[37] WEBER V S, ARAUS J L, CAIRNS J E, et al. Prediction of grain yield using reflectance spectra of canopy and leaves in maize plants grown under different water regimes[J]. Field Crops Research, 2012, 128: 82 – 90.

[38] SANKARAN S, MISHRA A, EHSANI R, et al. A review of advanced techniques for detecting plant diseases[J]. Computers and Electronics in Agriculture, 2010, 72(1): 1 – 13.

[39] DAMIEN V, DAMIEN E, GUILLAUME J, et al. Near infrared hyperspectral imaging method to assess Fusarium head blight infection on winter wheat ears[J]. Microchemical Journal, 2023, 191: 108812.