doi:10.6041/j.issn.1000-1298.2026.18.012

基于高光谱成像与深度学习的乌菜叶片亚硝酸盐含量无损检测

张国祥[1]　吴佳闻[1]　权龙哲[1]　李鑫星[2]　王翔宇[3]

（1. 安徽农业大学机械与车辆工程学院，合肥 230036；2. 中国农业大学信息与电气工程学院，北京 100083；
3. 长治学院物理系，长治 046011）

摘要：为实现乌菜亚硝酸盐含量快速无损检测，本研究提出一种基于高光谱成像与深度学习的高精度定量检测方法。在培养箱中设置 4 组不同光强环境，每组放置 50 株乌菜，获取不同光强环境下乌菜叶片样本，每组 200 片作为试验数据集。利用高光谱成像设备采集乌菜叶片样本高光谱图像信息，进行亚硝酸盐含量预测。通过感兴趣区域（ROI）和阈值分割处理图像获取平均光谱信息。利用不同预处理方法对光谱图像进行预测模型构建并分析性能，确立小波变换去噪（WTD）算法作为最佳预处理方法，并对使用 WTD 预处理的光谱数据进一步分析。使用 4 种传统降维方法对预处理后最佳光谱数据进行降维，分别通过支持向量机回归（SVR）建立模型与引入注意力块（SE）的一维卷积神经网络（1D－CNN）深度特征提取后建立的 SVR 模型进行分析。结果表明，基于 1D－CNN－SE 提取深度特征的 SVR 模型对不同光强环境中乌菜叶片亚硝酸盐含量预测效果较好，测试集决定系数（R_p^2）、均方根误差（RMSEP）和相对分析误差（RPD）分别为 0.884 9、0.073 81 μmol/g 和 2.763 1，表明在光强环境下建立的 WTD－1D－CNN－SE－SVR 模型性能较佳，深度学习方法结合高光谱成像无损检测技术能够有效实现乌菜叶片亚硝酸盐毒害的无损检测，为叶菜无机物毒害无损检测提供技术支持。

关键词：乌菜叶片；亚硝酸盐含量；无损检测；高光谱图像；卷积神经网络；深度学习

中图分类号：TS255.7；O657.3　　**文献标识码**：A　　**文章编号**：1000-1298(2026)18-0133-11　　**OSID**：

Non-destructive Detection of Nitrite Content in Tatsoi Leaves Based on Hyperspectral Imaging and Deep Learning

Zhang Guoxiang[1]　Wu Jiawen[1]　Quan Longzhe[1]　Li Xinxing[2]　Wang Xiangyu[3]

(1. *College of Mechanical and Vehicle Engineering*, *Anhui Agricultural University*, *Hefei* 230036, *China*
2. *College of Information and Electrical Engineering*, *China Agricultural University*, *Beijing* 100083, *China*
3. *Department of Physics*, *Changzhi University*, *Changzhi* 046011, *China*)

Abstract: Aiming to achieve rapid and non-destructive detection of nitrite content in Tatsoi (*Brassica rapa* subsp. *narinosa*), a high-precision quantitative detection method was proposed based on hyperspectral imaging and deep learning. Four groups of different light intensities were set up in a growth chamber, with 50 Tatsoi plants placed in each group. Leaf samples of Tatsoi under different light intensity environments were collected, and 200 leaf slices per group were used as the experimental dataset. Hyperspectral images of the leaf samples were acquired by using a hyperspectral imaging device to predict their nitrite content. The average spectral information was extracted by processing the images through region of interest (ROI) selection and threshold segmentation. Various preprocessing methods were employed to construct prediction models on the spectral data, and their performances were compared. Wavelet transform denoising (WTD) was identified as the optimal preprocessing method. The preprocessed spectral data using WTD were further analyzed. Four traditional dimensionality reduction methods were applied to reduce the dimensionality of the optimal preprocessed spectral data. Support vector regression (SVR) models were established on the reduced data, and further analysis was conducted by using SVR models incorporating deep features extracted by a one-dimensional convolutional

收稿日期：2026－04－22　修回日期：2026－05－18

基金项目：安徽省高校自然科学研究项目（2024AH050450）和山西省基础研究计划项目（202203021212170）

作者简介：张国祥（1994—），男，副教授，博士，主要从事智能设施农业装备研究，E-mail：zgxaau@ahau.edu.cn

通信作者：王翔宇（1989—），男，副教授，博士，主要从事图像信息处理研究，E-mail：wangxiangyu@czc.edu.cn

neural network (1D - CNN) integrated with a squeeze-and-excitation (SE) attention block. The results demonstrated that the SVR model based on deep features extracted by 1D - CNN - SE achieved superior predictive performance for nitrite content in Tatsoi leaves under different light intensity environments. On the test set, the coefficient of determination (R_p^2), root mean square error of prediction (RMSEP), and residual predictive deviation (RPD) were 0.884 9, 0.073 81 μmol/g, and 2.763 1, respectively. These findings indicated that the WTD - 1D - CNN - SE - SVR model established under light intensity conditions exhibited favorable performance. The combination of deep learning methods with hyperspectral imaging non-destructive detection technology can effectively achieve non-destructive detection of nitrite toxicity in Tatsoi leaves, providing technical support for non-destructive detection of inorganic toxin stress in leafy vegetables.

Key words: Tatsoi leaves; nitrite content; non-destructive detection; hyperspectral image; convolutional neural network; deep learning

0 引言

近年来,蔬菜亚硝酸盐毒害问题日益凸显,成为亟待重视的食品安全隐患。当人体中亚硝酸盐含量过高时,可导致血液缺氧产生中毒反应;若长期摄入含有过量亚硝酸盐蔬菜,可能增加消化系统癌变风险[1]。在植物中过量 NO_2^- 会抑制 PSⅡ修复[2]和蛋白质广泛 S-亚硝基化[3],引发光抑制,降低光合效率,干扰植物正常代谢。亚硝酸盐过量积累可能导致氮素同化失衡并加剧亚硝酸盐毒害、影响细胞膜透性出现叶柄黑点症[4]和根系无氧呼吸加剧,根系会逐渐腐烂直至叶片和植株完全枯死;严重影响植株生长发育和产量。因此,亚硝酸盐检测已成为目前叶菜生产病虫害防治、智能预警的关键生理性目标。

研究表明,无论是在大田及简易棚室等自然光波动环境下,还是在玻璃温室、植物工厂等高可控恒定光环境中,光照条件对叶菜亚硝酸盐含量均有显著影响。光照过强或过弱,均会抑制叶菜对亚硝酸盐的分解能力,从而加速其累积[5]。当前,亚硝酸盐的主流测定方法主要涵盖光谱比色法、色谱法、离子选择电极法和亚硝酸盐试纸条-反射仪速测法等[6]。光谱比色法凭借较广适用性,在实践中应用最为普遍。但以上方法多存在一定局限性,不仅需对样品进行复杂前处理,还通常具有破坏性,实现样品的无损检测非常困难。因此,亚硝酸盐无损检测技术成为保障叶菜食用品质的迫切需求。研究一种无需破坏性取样、操作简便的新型蔬菜亚硝酸盐检测技术,已成为植物生理指标测定的迫切研究方向。

可见光-近红外高光谱成像技术(Visible near-infrared hyperspectral imaging, Vis - NIR HIS)作为前沿的植物信息检测手段[7],融合了光谱与图像双重信息,具备对农作物进行无损检测能力,在农业信息化和精准栽培领域展现出广阔的应用前景。目前,该技术在植物信息分析方面已取得一定进展,岳学军等[8]采用高光谱检测技术对柑橘叶片叶绿素含量进行研究,使用小波去噪-偏最小二乘回归(Wavelet denoising - partial least square regression, WD - PLSR)构建无损检测模型,验证集模型决定系数(R_p^2)达到 0.867 0,可以有效识别叶绿素含量。Chu 等[9]提出了一种多尺度光谱分解和差异化特征选择对多生育期条件下大蒜叶面积指数(LAI)反演的高光谱建模方法,使用轻量化 CAM - CNN - Transformer 模型验证了多尺度光谱分解结合差分选择方法在提高光谱建模性能方面具有显著优势。Faqeerzada 等[10]使用高光谱图像融合技术挖掘在高通量检测干旱胁迫草莓植株中的应用潜力,PLS - DA 模型预测集精度达到 99%。周鑫等[11]利用感兴趣区域(Region of interest, ROI)对高光谱图像进行掩膜处理以提取有效区域,通过深度迁移学习方法结合高光谱成像无损检测技术构建的 SNV -(T - SAE)- SVR 模型能够有效实现油菜叶片锌含量检测。Hao 等[12]提出了基于 CNN - LSTM 算法的 H2D - COS 分析与迁移学习相结合方法研究番茄叶片超氧化物歧化酶(SOD)活性,对高光谱图像中敏感波长的微观模型和宏观模型进行了迁移学习,能够实时精准反映 SOD 活性动态变化。从众多学者对不同作物营养元素检测、逆境胁迫响应和生长指标反演的研究中发现,可见光-近红外高光谱成像技术结合先进的信号处理和深度学习方法,在植物生理检测方面具有独特优势。

本文依托高光谱图像无损检测技术,融合深度学习算法和经典分析化学方法,对不同光环境下叶菜植株叶片亚硝酸盐含量开展定量检测研究。技术路线分为高光谱数据采集、预处理、深度特征提取和模型构建 4 部分。验证通过高光谱实现乌菜叶片亚硝酸盐含量快速测定和深度提取模型的技术路线高效性与优越性,以期为叶菜品质无损快速检测提供

理论支撑与技术参考。

1　材料与方法

1.1　试验材料

以当前叶菜主产区黄淮海地区常见叶菜品类乌菜为试验对象，共选取 200 株形态相近、生长状态良好的乌菜植株作为试验试材，质量为 100 ~ 200 g，采用营养土育苗、移栽定植，生长周期为 60 d。样本植株随机均分成 4 份（每组 50 株）并施加相同质量浓度（1 000 mg/L）叶面氮肥（光合生物，高氮型水溶肥），分别放入 4 个独立植物培养箱中，相对湿度统一设定 70%，恒湿，温度维持在最适生长温度（19℃），光周期为 11 h，确保除光照强度外其他环境变量均保持不变。

试验采用 HWS - 50 型培养箱，控温范围 0 ~ 80℃，温度波动 ± 0.3℃，相对湿度 20% ~ 99.9%，相对湿度波动范围 ± 3%。叶菜植株亚硝酸盐含量采用 752N 型分光光度计（上海仪电科学仪器股份有限公司）法检测。采用自主搭建的叶菜高光谱图像采集平台，包括：x20P - LIR 一体式激光雷达红外高光谱成像仪，光谱范围为 350 ~ 1 002 nm，光谱波段数为 161，波段间隔为 4 nm；采用 2 个功率为 36 W 的全光谱植物补光灯，1 个体积为 1 m^3 铝型材拍摄平台，1 台装有 Cuvistouch 图像采集软件和 CuvisExport 图像输出软件的计算机，图像采集平台结构如图 1 所示。

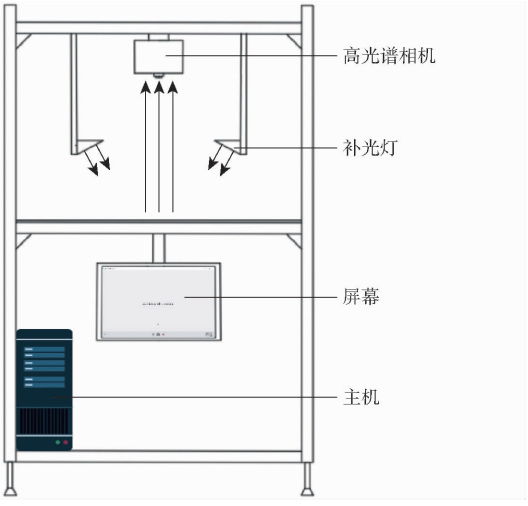

图 1　高光谱图像采集平台结构示意图

Fig. 1　Structure diagram of hyperspectral image acquisition platform

1.2　试验光强环境设定

模拟乌菜种植自然光波动环境和高可控设施下恒定光强条件，其中 A 组自然光波动采用光照强度与时间拟合的正弦函数控制。为更好地模拟冬季日照规律，设定 08:00—17:00 为控制光照变化时间段，自然光照函数为

$$Y(t_m)=\begin{cases}25\,000\sin^2\left(\dfrac{\pi}{600}(t_m-420)\right) & (420\ \text{min}\leqslant t_m\leqslant 1\,020\ \text{min})\\ 0 & (t_m<420\ \text{min 或 } t_m>1\,020\ \text{min})\end{cases} \tag{1}$$

式中　$Y(t_m)$——在时间t_m的光照强度，lx

在此基础上，设计 3 种不同梯度恒定光照强度环境：B 组（弱光组，10 000 lx）、C 组（对照组，20 000 lx）、D 组（强光组，30 000 lx）。

1.3　高光谱图像信息采集

采集乌菜叶片样本时，1 株乌菜取 2 ~ 3 片形态大小相近的成熟叶片进行图像采集。首先需预热高光谱相机 10 min，使系统完全启动，然后对高光谱相机进行黑白校正，以减少暗电流效应对图像产生的影响[13]：打开框架两侧补光灯，使用标定白板和全黑镜头盖分别获得纯白标定图像和纯黑标定图像，再将放置有乌菜样品的托盘放置在拍摄台架中央，拍摄原始高光谱图像，最后通过公式进行校正，得到校正后高光谱图像。校正公式为

$$I=\frac{I_0-I_w}{I_b-I_w} \tag{2}$$

式中　I——校正后乌菜叶片样本图像

I_0——原始图像

I_w——纯白标定图像

I_b——纯黑标定图像

使用 ENVI 5.6 软件对高光谱图像进行处理，通过分析样本与背景区域光谱差异，识别出 2 个特征区分明显的波段（546、806 nm），基于二者比值生成比值图像，对比值图像进行阈值分割处理（最小阈值设为 1.2），将得到的二值化掩模图像应用于经归一化处理的高光谱图像，通过 ROI 提取出目标区域的光谱信息。在此基础上，计算 ROI 范围内全部像元的平均光谱，并将其作为预处理阶段的输入数据。

1.4　亚硝酸盐含量测定

高光谱图像信息采集完成后，使用蒸馏水洗净乌菜叶片并使用吸水纸吸干残留水分。将叶片样本剪碎后使用电子天平称量，每个样品称取 0.1 g。随后按照亚硝酸盐试剂盒说明书（南京建成生物科技）对样本进行处理。处理完成后取测定管上层清液 0.8 mL，用于亚硝酸盐含量测定。

紫外分光光度计光路模式选择高光通量的单光束[14]，波长为 550 nm，采用 1 mL 比色皿，使用前需将仪器预热 30 min。NO_2^- 含量（μmol/g）计算式为

$$H_{NO_2^-}=\frac{\Delta A_1}{\Delta A_0}c\,\frac{V}{W} \tag{3}$$

式中 ΔA_0——标准管与空白管在波长 550 nm 吸光值之差

ΔA_1——测定管与空白管在波长 550 nm 吸光值之差

c——亚硝酸盐标准溶液浓度,取 100 μmol/L

W——样本质量,g

V——样本提取时加入提取液体积,mL

1.5 高光谱数据处理方法

1.5.1 数据预处理

为使光谱数据能够平滑去噪和增强特征,需对光谱进行预处理,采用一阶导数(First derivative, FD)[15]、二阶导数(Second derivative, SD)[15]、标准正态变换(Standard normalized variate, SNV)[15]、SG 平滑(Savitzky - Golay, SG)[16]、多元散射校正(Multiplicative scatter correction, MSC)[17]、小波变换去噪(Wavelet transform denoising, WTD)[18]和移动平均平滑(Moving average, MA)[19]对 ROI 提取后的光谱进行预处理。

1.5.2 数据降维

为消除波段间的多重共线性和大量无关噪声,避免维度过高,采用主成分分析法(Principal component analysis, PCA)[20]、竞争性自适应重加权采样(Competitive adaptive reweighted sampling, CARS)、无信息变量消除法(Uninformative variables elimination, UVE)[21]和连续投影算法(Successive projections algorithm, SPA)[22]对最佳预处理后光谱数据进行数据降维并与使用深度特征提取进行对比。

1.5.3 深度特征提取

由于研究所涉及的高光谱数据具有连续一维特性,每个样本是 1 条连续的一维信号,因此采用一维卷积神经网络(1 - dimensional convolutional neural network, 1D - CNN)[23]结合注意力块(Squeeze-and-excitation, SE)[24]实现对光谱数据的深度特征提取。

(1)1D - CNN 网络结构设计

1D - CNN 主要由滤波级组成,承担深度特征提取功能。首先通过“卷积-批归一化-激活-池化”堆叠操作,实现从输入样本中完成多尺度局部特征提取,选用修正线性单元(Rectified linear unit, ReLU)[25]作为滤波级的激活函数。使用最大池化(Max pooling, MP)策略进行池化,池化窗口尺寸设为 3 ×1,步幅为 2,窗口在输入特征序列上无重叠滑动,每步保留局部区域内最大响应值。在多层卷积后使用 SE 模块对特征通道进行自适应重标定,增强与亚硝酸盐相关的光谱特征,抑制无关信息,提供质量更高的特征表示。针对本研究对象和目标,还需结合 1 ×1 卷积对通道数进行压缩降维,并使用全局平均池化层固定输出维度,便于建模处理。搭建的 1D - CNN - SE 深度特征提取网络模型结构如图 2 所示。

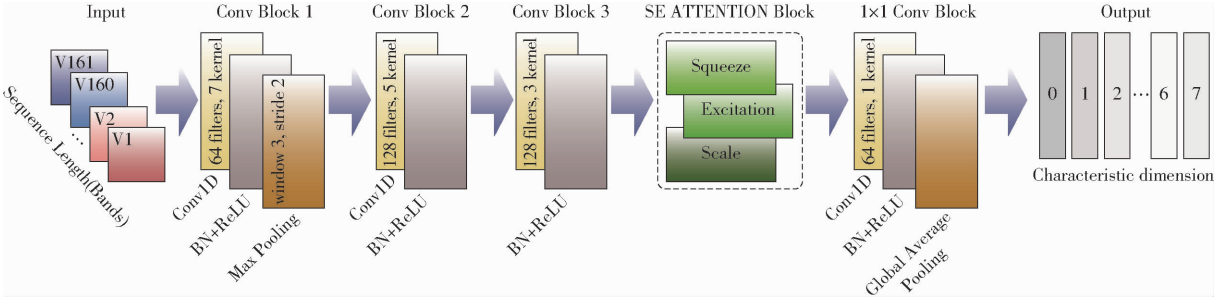

图 2 深度特征提取网络结构

Fig. 2 Network structure of deep feature extraction

在滤波级中,卷积承担了核心的特征提取功能。设第 l 层为卷积层,则该层适用于乌菜叶片高光谱数据处理的一维卷积运算公式为

$$\boldsymbol{y}_j^{(l)}(x,y)=\sum_{i=1}^{M}(\boldsymbol{x}_i^{(l)}(x,y)*\boldsymbol{\mu}_{ij}^{(l)})+b_j^{(l)} \quad (4)$$

式中 $\boldsymbol{y}_j^{(l)}(x,y)$——第 l 层在空间位置(x,y)处,第 j 个输出通道的叶片光谱特征向量

$\boldsymbol{x}_i^{(l)}(x,y)$——第 l 层在空间位置(x,y)处,第 i 个输入通道的叶片光谱特征向量

$\boldsymbol{\mu}_{ij}^{(l)}$——第 l 层连接第 i 个输入与第 j 个输出的一维卷积核

$b_j^{(l)}$——第 j 个输出通道偏置

M——输入光谱通道数

为适配回归模型输入要求,需对特征维度进行降维。传统的降维方法如全连接层参数量较大,多级卷积逐步降维计算复杂度高。为此,引入 1 ×1 卷积层对输入特征通道维度进行特征降维和线性重组,用最小参数量实现通道维度的有效压缩。对于输入批次大小为 B、特征维度为 M 的特征向量 $\boldsymbol{X}\in\mathbf{R}^{B\times M}$的 1 ×1 卷积运算过程可表示为[26]

$$Y_{b,c} = \sum_{k=1}^{M_1} W_{c,k} X_{b,k} + b_c \tag{5}$$

式中　$Y_{b,c}$——输出特征中第 b 个样本和第 c 个通道值

$W_{c,k}$——第 c 个输出通道对应第 k 个输入通道权重

$X_{b,k}$——输入特征中第 b 个样本和第 k 个通道值

b_c——第 c 个输出通道偏置

M_1——原始高光谱数据波段数

在网络末端，采用全局平均池化层将最后一个卷积层输出的特征图聚合为固定长度的特征向量。全局平均池化通过对每个特征通道的所有空间位置求均值，将变长序列转换为固定长度向量，增强了模型的平移不变性，数学表达式为[27]

$$F_c = \frac{1}{L} \sum_{i=1}^{L} a_c^{(\text{last})}(i) \tag{6}$$

式中　F_c——第 c 个通道全局特征值

L——特征图序列长度

$a_c^{(\text{last})}(i)$——最后 1 个卷积层输出的第 c 个特征通道在位置 i 处激活值

1D－CNN 网络的设计充分考虑了高光谱数据的一维序列特性，通过多尺度卷积捕获局部光谱模式，使用堆叠结构构建层次化的特征表示。最终，输出一个维度为 C 的全局特征向量 $\boldsymbol{F} \in \mathbf{R}^C$，该向量即为从原始高光谱数据中提取的深度光谱特征，作为 SVR 回归模型输入。

(2)SE 注意力块设计

在卷积神经网络中，针对高光谱数据中不同波段组合对亚硝酸盐检测贡献度不同的特点，在一维卷积神经网络中引入注意力块。SE 通过显式建模通道间的相互依赖关系，自适应地重标定各通道的特征响应[24]，使网络聚焦在与亚硝酸盐含量有关的光谱特征。

SE 注意力块是一种轻量级的通道注意力机制，核心操作包括压缩(Squeeze)、激励(Excitation)和重标定(Reweighting)[24]。在 SE 注意力块内引入双路径混合池化技术(Dual-path hybrid pooling)，将全局平均池化与全局最大池化结果联合，形成统一的压缩描述子。对于输入特征图 $\boldsymbol{U} = (u_1, u_2, \cdots, u_C) \in \mathbf{R}^{B \times L \times C_\text{out}}$(其中 C_out 为通道数)，双路径混合池化压缩操作表达式为

$$z_c = \frac{1}{L} \sum_{j=1}^{L} u_c(j) \parallel \max_{j \in [1,L]} u_c(j) \tag{7}$$

式中　$u_c(j)$——在第 c 个通道的特征图中，在第 j 个序列位置上的值

z_c——第 c 个通道的通道描述子

在 1D－CNN 网络中，SE 注意力机制使得网络在完成多尺度局部特征提取后，能够对特征通道进行自适应重标定，从而增强与亚硝酸盐相关的光谱特征，同时抑制无关信息，为后续处理提供更高质量的特征表示。SE 模块设计如图 3 所示。

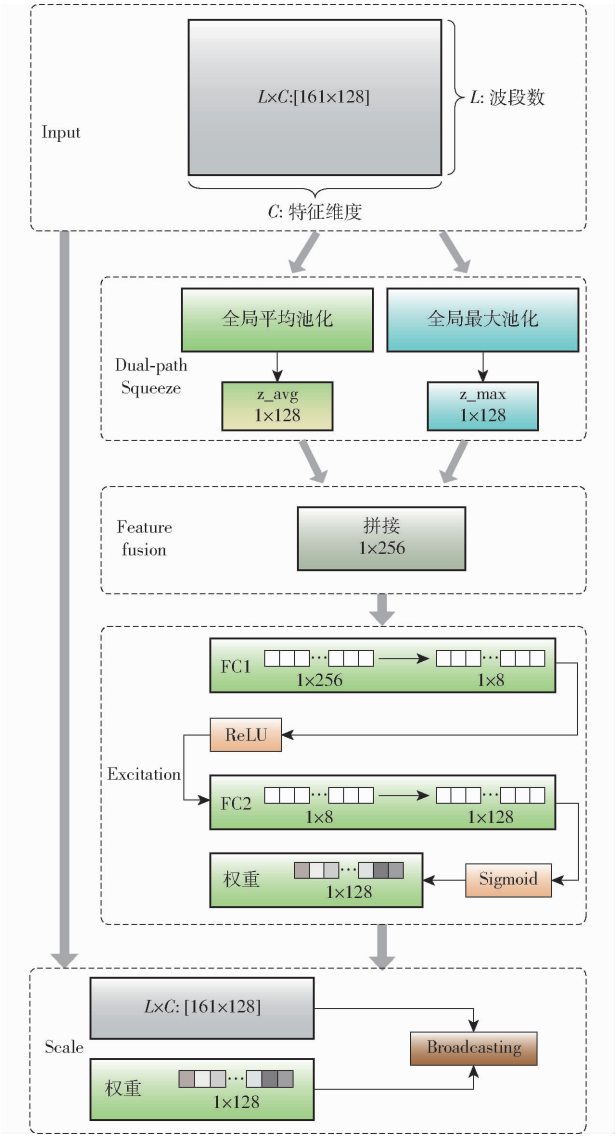

图 3　SE 注意力块结构

Fig. 3　Structure of SE attention block

1.6　SVR 算法

支持向量回归(SVR)是一种基于统计学习理论的回归方法，通过寻找一个能够容忍一定偏差 ε 的最优超平面，同时保持模型复杂度最低，因此 SVR 在小样本和高维数据上具有优异的泛化能力[28]，适用于本研究样本有限的场景。在建模前需对数据进行划分，将 A、B、C、D 4 组光环境处理下采集的叶片样本光谱数据划分为 NL、LL、ML、HL 4 个数据集，分别按比例 6∶2∶2将数据集划分为训练集、验证集和测试集，如表 1 所示。

表 1 SVR 数据集划分

Tab. 1 SVR dataset partitioning

数据集	训练集	验证集	测试集
NL	120	40	40
LL	120	40	40
ML	120	40	40
HL	120	40	40

在乌菜叶片高光谱数据中,亚硝酸盐含量与光谱反射率之间存在非线性关系。径向基(RBF)核函数可以将数据映射到无限维特征空间,有效捕捉非线性模式。由于 RBF 核参数少,仅需优化惩罚系数 C、核参数 γ 和不敏感带宽度 ε,适合小样本情况下网格搜索优化[29]。因此选用 RBF 核函数作为 SVR 核函数。

1.7 评价指标

为全面评估模型性能,使用的回归评价指标包括决定系数(R^2)、均方根误差(RMSE)和相对分析误差(RPD)。R^2 衡量模型对目标变量方差的解释度,RMSE 衡量预测值和真实值之间的偏差,RPD 用于评估模型预测能力[30]。

2 试验结果与分析

2.1 不同光环境下叶片亚硝酸盐含量

4 组不同光环境数据集中乌菜叶片平均亚硝酸盐含量测定结果如图 4 所示。从图 4 可看出,A 组平均亚硝酸盐含量为 5.031 μmol/g,显著高于 B、C、D 组($p<0.05$)。在恒定光照组中,B 组显著高于 C 组和 D 组($p<0.05$),C 组与 D 组间差异未达到显著水平($p>0.05$)。这一变化与光照强度胁迫下薤菜亚硝酸盐含量检测研究结果一致[31]。

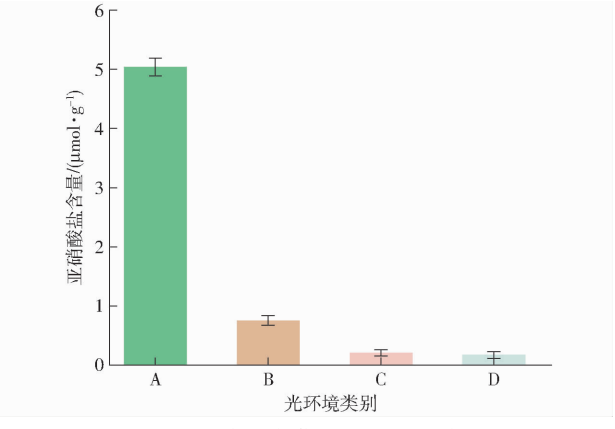

图 4 不同光环境下乌菜叶片平均亚硝酸盐含量

Fig. 4 Average nitrite concentration in leaves of Tatsoi under different light environments

不同光环境下叶片样本亚硝酸盐含量平均光谱曲线如图 5 所示。从图 5 可看出,4 组样本光谱曲线整体趋势一致,但在不同光照强度下,乌菜叶片光谱反射率存在明显差异,利用高光谱技术实现乌菜叶片中亚硝酸盐含量测定可行且有效。

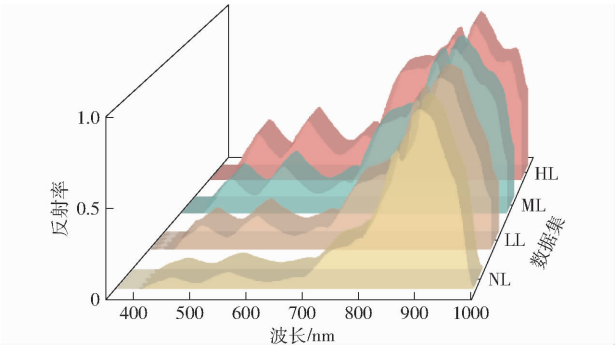

图 5 不同光环境下乌菜叶片亚硝酸盐含量平均光谱曲线

Fig. 5 Average spectral curves of nitrite content in leaves of Tatsoi under different light environments

2.2 高光谱数据预处理

使用 1.5 节预处理方法对原始 ROI 平均光谱数据进行处理,其中一阶导数和二阶导数处理窗口长度为 30,多项式阶数为 3。SG 平滑处理窗口尺寸为 11,多项式阶数为 3。小波变换去噪的小波基函数选用 Daubechies 4[17],分解层数为 3 层,阈值模式为软阈值。移动平均平滑处理窗口尺寸为 5。预处理后光谱数据如图 6 所示,从图 6 可发现,不同预处理后光谱曲线存在一定差异性。对使用不同预处理方法的 ROI 光谱数据构建的 SVR 模型评估结果如表 2 所示。从表 2 可得出,使用不同预处理方法的高光谱数据所建立的 SVR 亚硝酸盐含量预测模型性能存在差异,经对比发现,采用 WTD 预处理后光谱数据集经过深度特征提取后所构建的 SVR 模型性能较佳,在不同光环境下测试集平均最优 R_p^2、RMSEP 和 RPD 分别为 0.884 9、0.073 81 μmol/g 和 2.763 1。因此,将使用 WTD 对高光谱数据进行预处理。

2.3 高光谱数据降维

采用 PCA、CARS、UVE 和 SPA 对经过 WTD 预处理后的数据进行降维。PCA 算法使用累计方差解释率法选择主成分数[32],PCA 对 NL、LL、ML 和 HL 4 个不同数据集进行降维后主成分数贡献率如图 7 所示。从图 7 可知,PCA 选择的主成分数分别为 7、7、10 和 11,累积贡献率分别为 99.09%、99.11%、99.20% 和 99.08%。

CARS 算法通过迭代剔除权重较小的波长,保留对模型贡献大的波长。由于 CARS 需要的样本数较多,因此将 NL、LL、ML 和 HL 4 个数据集合并后再使用 CARS 对预处理后的高光谱数据进行特征波段筛选,结果如图 8 所示。从图 8 可看出,随着迭代次数增加,交叉验证均方根误差呈先下降后上升趋势,在第 33 次迭代时达到最小值 0.069,保留特征

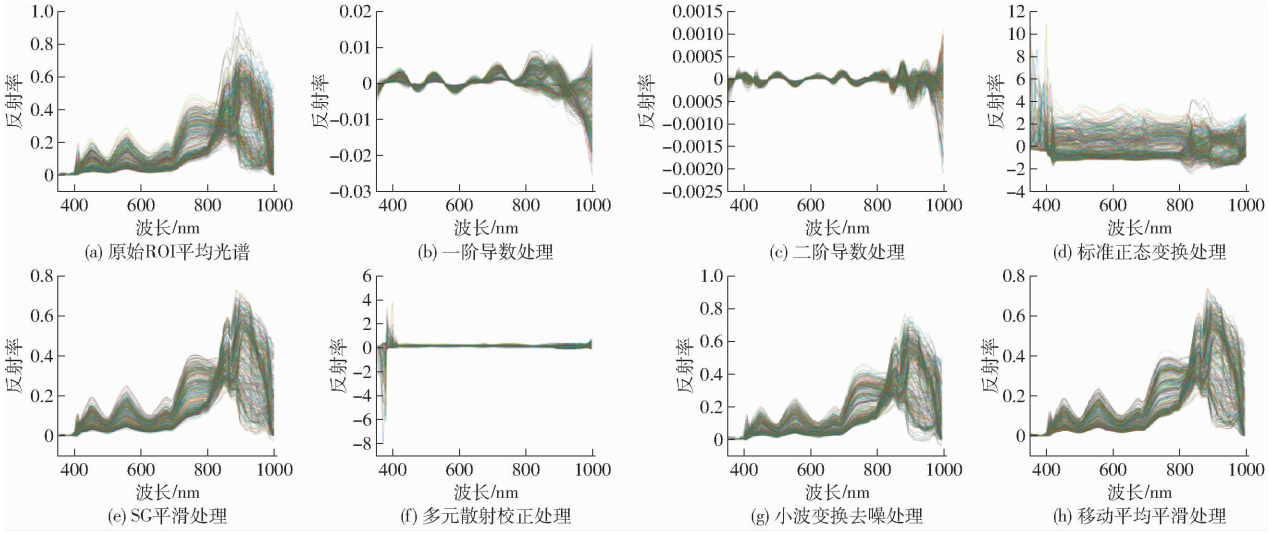

图 6　预处理后光谱数据曲线

Fig. 6　Spectral data curves after pretreatment

表 2　不同预处理后高光谱数据集的 SVR 模型性能

Tab. 2　Performance of SVR model for hyperspectral dataset after different preprocessing

数据集	数量	预处理	训练集		验证集		测试集		
			R_c^2	RMSEC/(μmol·g^{-1})	R_{cv}^2	RMSECV/(μmol·g^{-1})	R_p^2	RMSEP/(μmol·g^{-1})	RPD
NL、LL、ML、HL	400	原始	0.812 7	0.098 41	0.694 3	0.157 31	0.683 1	0.151 26	1.543 7
		FD	0.823 7	0.096 59	0.714 7	0.141 46	0.723 1	0.129 23	1.729 6
		SD	0.847 6	0.073 72	0.722 6	0.138 19	0.714 3	0.136 15	1.713 1
		SNV	0.784 3	0.113 74	0.716 1	0.146 83	0.695 4	0.148 37	1.643 2
		SG	0.807 6	0.090 63	0.701 2	0.149 67	0.721 4	0.138 11	1.732 9
		MSC	0.814 9	0.089 14	0.725 7	0.131 39	0.692 1	0.149 60	1.627 4
		WTD	0.964 3	0.056 49	0.893 2	0.069 17	0.884 9	0.073 81	2.763 1
		MA	0.837 9	0.074 67	0.802 4	0.097 21	0.782 5	0.125 61	1.954 0

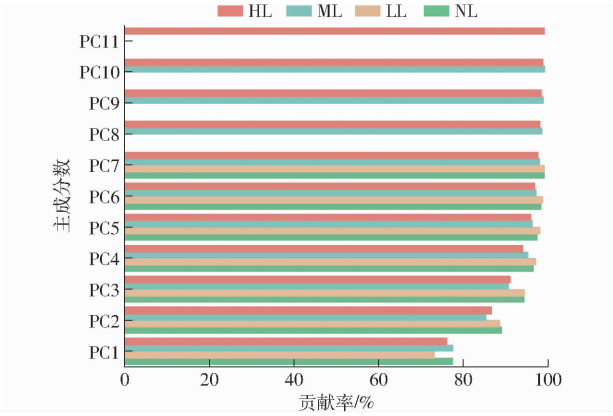

图 7　PCA 降维处理后累积贡献率

Fig. 7　Cumulative contribution rate after PCA dimensionality reduction

波段数为 7 个，压缩比为 4.29%，特征波段主要集中在 790 nm 和 874 nm 附近。

UAE 算法通过添加随机变量矩阵进行 PLS 建模[21]。NL、LL、ML 和 HL 4 个数据集中 UAE 算法阈值为 1.1，随机变量为 100。SPA 算法利用向量的投影分析，通过多次迭代选出最优波长组合[22]。使用 UVE 和 SPA 对预处理后数据集进行降维提取的特征波段分布如图 9 所示。从图 9 可知，在不同光环境下 NL、LL、ML 和 HL 数据集中，UAE 算法提取的相关特征波段最终为 147、126、141 和 146；SPA 算法提取的相关特征波段为 4、6、32 和 5。2 种降维提取算法处理后光谱特征波段主要分布在 430、602、758、834、930 nm 附近。

2.4　1D－CNN－SE 深度特征提取

为更好地挖掘高光谱图像中光谱特征，设计并训练了一种 1D－CNN－SE 深度提取网络。为更好地提高泛化能力，将 WTD 预处理后的 4 个高光谱数据集进行合并且按比例 6∶2∶2 随机划分为训练集、验证集和测试集。在模型构建中采用多重优化策略，通过卷积设计保持特征图空间维度的一致性；SE 注意力模块采用瓶颈结构，能够大幅减少参数量。采用 1×1 卷积实现跨通道信息融合与特征降维，将高维特征压缩至 8 维。采用 AdamW 优化器的解耦权重衰减机制[33]，学习率为 1×10^{-3}，衰减权重系数为 1×10^{-4}，通过 L2 正则化约束模型参数范

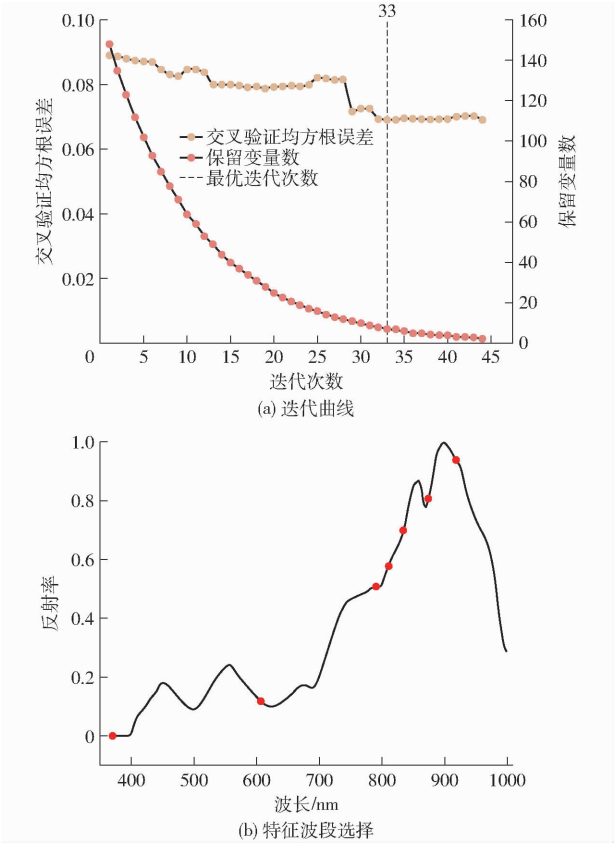

图 8 CARS 迭代次数和波段选择结果

Fig. 8 CARS iteration times and band selection

数。使用 Smooth L1 损失函数作为优化目标[34]。为防止模型过拟合,模型轮次设置为 500,批次大小为 32,采用早停机制(Early stopping)进行训练控制,当连续 100 轮验证集损失未出现下降时,提前终止训练并恢复最佳模型权重。训练集和验证集损失函数和决定系数如图 10 所示。从图 10 可知,训练损失和验证损失在 130 轮之前迅速下降,在 335 轮之后逐渐收敛,最终稳定在 0.05 ~ 0.1 之间,R^2 分别为 0.937 7 和 0.948 5。

2.5 消融试验

为验证 1D - CNN - SE 网络中各组件有效性,设计 6 种不同的特征提取网络(Net1 ~ Net6)。提取网络均输出 8 维深度特征,并使用超参数相同的 SVR 模型进行建模和预测,不同特征提取网络设计如表 3 所示。

使用 WTD 对数据进行预处理后,使用不同深度学习网络对数据集进行深度特征提取后构建的 SVR 模型测试集性能评估结果如表 4 所示。从表 4 可看出,在 SVR 模型完全相同的情况下,不同特征提取网络导致显著的性能差异。1D - CNN - SE 3 层卷积神经网络性能效果较好,R_p^2、RMSEP 和 RPD 分别为 0.884 9、0.073 81 μmol/g 和 2.763 1,证明了多尺度卷积、SE 注意力和 1 ×1 卷积模块的有效性。

2.6 SVR 模型构建

SVR 模型采用 RBF 核函数,通过网格搜索方法确定 C、γ 和 ε 分别为 10、0.1 和 0.05。使用 WTD 对数据进行预处理后,由传统降维方法和 1D - CNN - SE 所构建的 SVR 模型性能如表 5 所示。从表 5 可看出,不同降维和特征提取方法对 SVR 模型性能产生显著影响。WTD - SVD - SVR 和 WTD - PCA - SVR 通过线性变换进行降维,WTD - SVD - SVR 模型 R_p^2、RMSEP 和 RPD 分别为 0.712 6、0.143 72 μmol/g和 1.632 1,WTD - PCA - SVR 模型为 0.730 4、0.124 25 μmol/g 和 1.894 3。WTD - UVE - SVR 和 WTD - SPA - SVR 通过特征选择策略保留原始波长。WTD - UVE - SVR 模型R_p^2、RMSEP 和 RPD 分别为 0.781 4、0.103 71 μmol/g 和 2.234 5。WTD - SPA - SVR 模型为 0.721 3、0.156 53 μmol/g

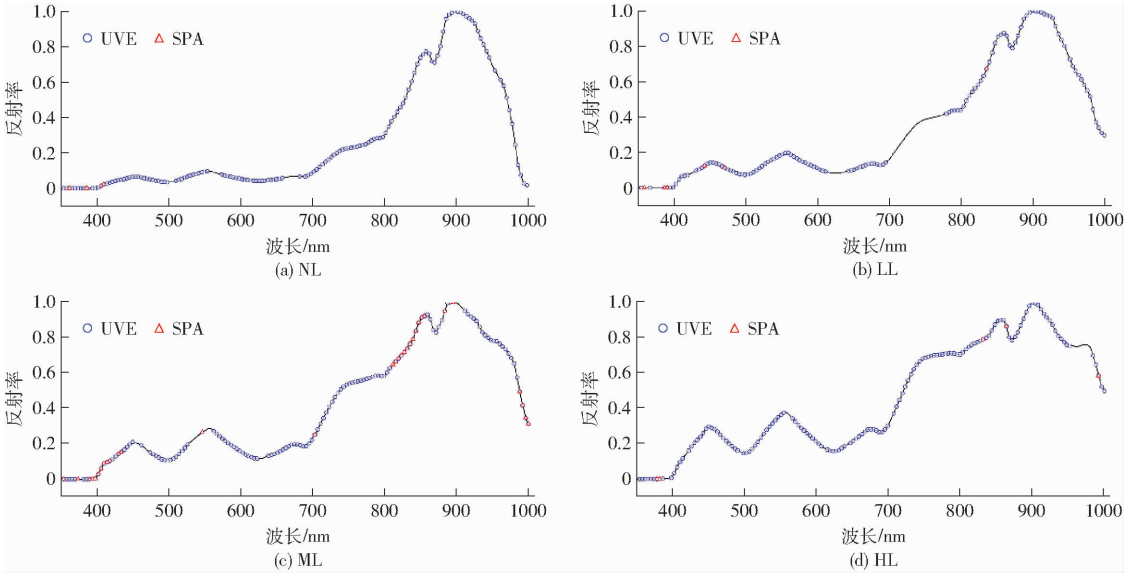

图 9 UVE 和 SPA 降维处理后特征波段

Fig. 9 Characteristic band after dimension reduction processing of UVE and SPA

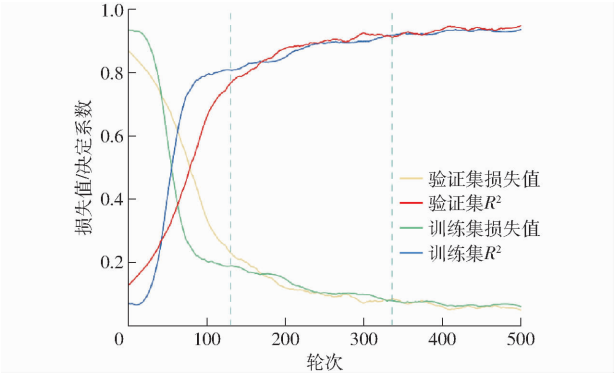

图 10　1D－CNN－SE 损失函数和决定系数曲线

Fig. 10　1D－CNN－SE loss function and R^2 curves

表 3　不同特征提取网络设计

Tab. 3　Design of different feature extraction networks

特征提取网络	卷积层数	SE 模块	1×1 降维
Net1	1	●	●
Net2	2	●	●
Net3	3	○	○
Net4	3	○	●
Net5	3	●	○
Net6	3	●	●

注：●表示包含该模块，○表示未包含该模块。

表 4　不同特征提取网络的 SVR 模型性能

Tab. 4　Performance of SVR models with different feature extraction networks

特征提取网络	R_p^2	RMSEP/(μmol·g^{-1})	RPD
Net1	0. 742 7	0. 138 31	1. 824 1
Net2	0. 794 3	0. 113 63	2. 128 1
Net3	0. 769 1	0. 144 72	1. 984 2
Net4	0. 806 2	0. 105 94	2. 251 7
Net5	0. 811 3	0. 093 57	2. 345 6
Net6	0. 884 9	0. 073 81	2. 763 1

和 1. 775 1。WTD－1D－CNN－SE－SVR 模型R_p^2、RMSEP 和 RPD 分别为 0. 884 9、0. 073 81 μmol/g 和 2. 763 1。WTD－1D－CNN－SE－SVR 模型的所有测试集样本真实值与预测值残差分布如图 11 所示。从图 11 可观察到，所有残差点随机分布在零线两侧，无明显的趋势或模式，残差点之间无明显关联，相邻点不存在同向偏离趋势，模型表现出较好的无偏性。4 个不同光环境处理数据集预测值和真实值如表 6 所示，由表 6 可知，NL、LL、ML 和 HL 数据集R_p^2分别为 0. 897 1、0. 891 6、0. 880 4 和 0. 870 5，RMSEP 为 0. 063 06 ~ 0. 092 36 μmol/g，

表 5　不同特征提取后高光谱数据 SVR 模型结果

Tab. 5　Results of SVR model after different feature extractions on hyperspectral data

数据集	模型	训练集		验证集		测试集		
		R_c^2	RMSEC/(μmol·g^{-1})	R_{cv}^2	RMSECV/(μmol·g^{-1})	R_p^2	RMSEP/(μmol·g^{-1})	RPD
NL、LL、ML、HL	WTD－SVD－SVR	0. 804 6	0. 093 25	0. 732 8	0. 137 36	0. 712 6	0. 143 72	1. 632 1
	WTD－PCA－SVR	0. 824 1	0. 090 59	0. 751 6	0. 115 17	0. 730 4	0. 124 25	1. 894 3
	WTD－UVE－SVR	0. 840 9	0. 087 24	0. 790 6	0. 098 19	0. 781 4	0. 103 71	2. 234 5
	WTD－SPA－SVR	0. 801 2	0. 094 51	0. 734 7	0. 146 83	0. 721 3	0. 156 53	1. 775 1
	WTD－1D－CNN－SE－SVR	0. 964 3	0. 056 49	0. 893 2	0. 069 17	0. 884 9	0. 073 81	2. 763 1

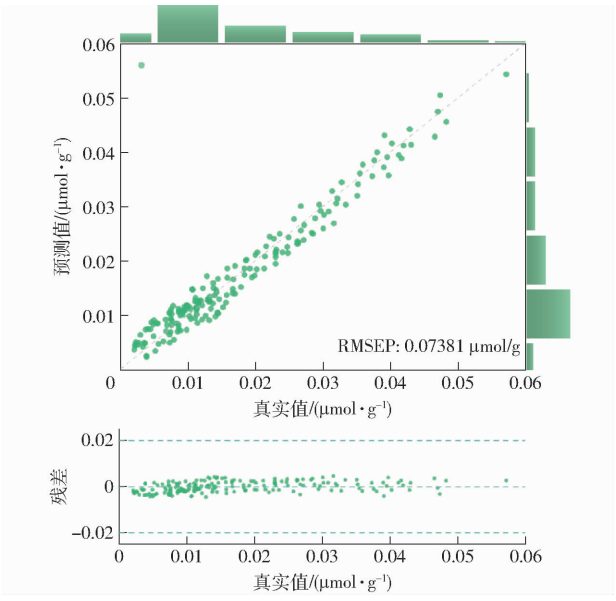

图 11　WTD－1D－CNN－SE－SVR 模型残差分析

Fig. 11　Residual analysis of WTD－1D－CNN－SE－SVR model

RPD 为 2. 433 0 ~ 3. 118 3。4 组间R_p^2极差仅为 0. 026 6，RMSEP 波动小于 0. 03 μmol/g，平均偏差均接近于零且无系统性趋势，但不同光照处理仍然对预测结果产生了一些影响。以上结果验证了基于高光谱成像与深度学习的乌菜叶片亚硝酸盐毒害无损检测方法的可行性与有效性。经分析发现，基于 1D－CNN－SE 的 SVR 模型在不同光环境数据集下整体表现最佳，较优于其他传统建模方法，说明 WTD－1D－CNN－SE－SVR 模型在不同光照强度条件下均能保持稳定的预测性能，对光谱测量条件的变化具有良好鲁棒性，能够实现对乌菜叶片亚硝酸盐含量的有效预测。

2.7　讨论

构建的 WTD－1D－CNN－SE－SVR 组合模型，实现了基于高光谱成像技术和深度学习的乌菜叶片亚硝酸盐毒害无损检测，为叶菜无损监测亚硝酸盐毒害提供了新的解决方案。研究结果表明，深度学

表 6 不同光环境下 SVR 模型测试集预测结果

Tab. 6 Prediction results of SVR model on test set under different lighting conditions

数据集	光照强度/lx	实测均值 ± 标准差/(μmol·g^{-1})	预测均值 ± 标准差/(μmol·g^{-1})	平均偏差/(μmol·g^{-1})	R_p^2	RMSEP/(μmol·g^{-1})	RPD
NL	自然光照	5.124 ± 0.288	5.080 ± 0.312	−0.044	0.897 1	0.092 36	3.118 3
LL	10 000	0.753 ± 0.195	0.779 ± 0.221	0.026	0.891 6	0.068 73	2.837 2
ML	20 000	0.181 ± 0.167	0.204 ± 0.189	0.023	0.880 4	0.063 06	2.648 3
HL	30 000	0.198 ± 0.173	0.219 ± 0.195	0.021	0.870 5	0.071 10	2.433 0

习方法结合高光谱成像技术能够有效提取与亚硝酸盐含量相关的深层光谱特征。本研究验证了方法的可行性,但仍存在试验样本均来源于人工培养箱控制环境,未涵盖不同采收期、不同含水量及不同生育期等田间实际因素对叶片物质构成的潜在影响的局限性。此外,仅针对乌菜单一品种展开验证,模型在不同叶菜种类间的迁移适用性尚待检验。后续研究应进一步纳入多生育期、多水分梯度及多品种样本,以提升模型在实际生产条件下鲁棒性与通用性。可以进一步扩大样本规模,探索迁移学习在不同作物和环境中的应用。推动该技术在实际农业生产中的推广应用。

3 结论

(1)不同光照强度对乌菜植株叶片亚硝酸盐含量有着明显影响。其中,恒定光照组中叶片亚硝酸盐含量随着光照强度增加表现出逐级下降趋于平缓趋势,自然光波动组(A 组)亚硝酸盐含量最高。

(2)在对光谱数据进行预处理过程,试验发现小波变换去噪(WTD)表现最佳。通过与其他预处理方法的 SVR 模型性能对比,WTD 处理后光谱数据优于原始光谱和其他预处理方法。

(3)使用 SVD、PCA、UVE 和 SPA 4 种传统降维方法与 1D − CNN − SE 深度特征提取方法进行性能对比。试验结果表明,基于 1D − CNN − SE 深度特征提取的 SVR 模型性能最优,R_p^2、RMSEP 和 RPD 分别为 0.884 9、0.073 81 μmol/g 和 2.763 1,明显优于其他 4 种传统方法。对特征提取网络的消融试验进一步说明,多尺度卷积、SE 注意力机制和 1 × 1 卷积模块的协同作用对提升模型性能有效。

参 考 文 献

[1] 朱卫芳,占绣萍,宋佳,等. 上海某郊区蔬菜硝酸盐和亚硝酸盐含量状况及膳食摄入风险评估[J]. 农产品质量与安全,2023(2):89 − 93.

[2] Kasten D, Mthöfer A, Georgii E, et al. Nitrite is the driver, phytohormones are modulators while NO and H_2O_2 act as promoters of NO_2-induced cell death[J]. Journal of Experimental Botany, 2016, 67(22): 6337 − 6349.

[3] Jedelská T, Luhová L, Petřivalský M. Decoding nitric oxide signals: the S-denitrosation machinery in plants[J]. Plant Science, 2026, 362: 112801.

[4] 雷娟利,李必元,岳智臣,等. 氮素水平对大白菜叶柄黑点症发生的影响[J]. 浙江农业科学,2017,58(11):2010 − 2012.

[5] 占绣萍,李建勇,方朝阳,等. 蔬菜中硝酸盐、亚硝酸盐残留状况及调控对策[J]. 上海农业学报,2024,40(2):89 − 93.
Zhan Xiuping, Li Jianyong, Fang Chaoyang, et al. Residual status and regulation strategies of nitrate and nitrite in vegetables[J]. Shanghai Journal of Agricultural Sciences, 2024, 40(2): 89 − 93. (in Chinese)

[6] 薛利红,杨林章. 基于可见近红外高光谱的菠菜硝酸盐快速无损测定研究[J]. 光谱学与光谱分析,2009,29(4):926 − 930.
Xue Lihong, Yang Linzhang. Rapid and nondestructive determination of nitrate in spinach based on visible near infrared hyperspectral spectroscopy[J]. Spectroscopy and Spectral Analysis, 2009, 29(4): 926 − 930. (in Chinese)

[7] Wang C, Liu B, Liu L, et al. A review of deep learning used in the hyperspectral image analysis for agriculture[J]. Artificial Intelligence Review, 2021, 54(7): 5205 − 5253.

[8] 岳学军,全东平,洪添胜,等. 柑橘叶片叶绿素含量高光谱无损检测模型[J]. 农业工程学报,2015,31(1):294 − 302.
Yue Xuejun, Quan Dongping, Hong Tiansheng, et al. Non-destructive hyperspectral measurement model of chlorophyll content for citrus leaves[J]. Transactions of the CSAE, 2015, 31(1): 294 − 302. (in Chinese)

[9] Chu Zhiqiang, Wu Qiulan, Sun Wenqi, et al. Hyperspectral inversion method of garlic LAI based on multi-scale spectral decomposition and differentiated feature selection[J]. International Journal of Remote Sensing, 2026, 47(6): 2692 − 2722.

[10] Faqeerzada M A, Kim H, Kim M S, et al. Hyperspectral imaging VIS − NIR and SWIR fusion for improved drought-stress identification of strawberry plants[J]. Computers and Electronics in Agriculture, 2025, 237: 110702.

[11] 周鑫,王坚,赵春江,等. 基于高光谱深度特征的油菜叶片锌含量检测[J]. 农业工程学报,2024,40(19):262 − 271.
Zhou Xin, Wang Jian, Zhao Chunjiang, et al. Detection of zinc content in rape leaves based on hyperspectral depth characteristics[J]. Transactions of the CSAE, 2024, 40(19): 262 − 271. (in Chinese)

[12] Hao J, Yan Y, Zhang Y, et al. A cross-scale transfer learning framework: prediction of SOD activity from leaf microstructure to macroscopic hyperspectral imaging[J]. Plant Biotechnology Journal, 2025, 23(4): 1091 − 1100.

[13] Blaksley C, Casolino M, Cambié G. Design and performance of a hyperspectral camera for full-face in vivo imaging[J]. Review of Scientific Instruments, 2021, 92(5): 055108.

[14] 侯耀斌, 冯巍巍, 蔡宗岐,等. 基于神经网络模型的海水硝酸盐测量方法研究[J]. 光谱学与光谱分析,2020,40(10): 3211 - 3216.
Hou Yaobin, Feng Weiwei, Cai Zongqi, et al. Research on measurement method of seawater nitrate based on neural network model[J]. Spectroscopy and Spectral Analysis, 2020, 40(10): 3211 - 3216. (in Chinese)

[15] 霍迎秋,赵士超,凌晨东,等. 基于高光谱的猕猴桃叶片氮磷钾元素快速检测方法研究[J]. 中国农机化学报,2026, 47(4):195 - 204.
Huo Yingqiu, Zhao Shichao, Ling Chendong, et al. Rapid detection method of nitrogen, phosphorus and potassium elements in kiwi leaves based on hyperspectral technology[J]. Journal of Chinese Agricultural Mechanization, 2026, 47(4): 195 - 204. (in Chinese)

[16] 梁晓颖,张紫婷,杨硕,等. 基于高光谱成像的小麦赤霉病严重度轻量化检测方法[J]. 农业机械学报,2025,56(6):218 - 227.
Liang Xiaoying, Zhang Ziting, Yang Shuo, et al. Hyperspectral imaging-based lightweight detection method for rapid detection of fusarium head blight severity in wheat[J]. Transactions of the Chinese Society for Agricultural Machinery, 2025, 56(6): 218 - 227. (in Chinese)

[17] Ryu J, Hong S, Park S, et al. Nondestructive freshness evaluation of mackerel fish using Vis/NIR hyperspectral imaging and multivariate analysis[J]. Journal of Food Engineering, 2024, 377: 112086.

[18] Ibrahim I, AlRowaily M, Arof H, et al. Performance comparison of selected filters in fast denoising of oil palm hyperspectral data[J]. Applied Sciences, 2024, 14(19): 8895.

[19] Zang Yuanpu, Liu Jiangping. Nondestructive detection of milk nutritional components based on hyperspectral imaging[J]. Journal of Food Science, 2025, 90(1): e17621.

[20] 孙俊,胡双齐,周鑫,等. 便携式蔬菜叶片重金属镉含量无损检测仪设计与试验[J]. 农业机械学报,2022,53(2):195 - 202,220.
Sun Jun, Hu Shuangqi, Zhou Xin, et al. Design and experiment of portable nondestructive detector for cadmium content in vegetable leaves[J]. Transactions of the Chinese Society for Agricultural Machinery, 2022, 53(2): 195 - 202, 220. (in Chinese)

[21] 封润泽,韩鑫,兰玉彬,等. 基于高光谱和 CNN - LSTM 的白菜叶片铜胁迫分析与分类模型研究[J]. 农业机械学报, 2025,56(6):477 - 486.
Feng Runze, Han Xin, Lan Yubin, et al. Analysis and non-destructive monitoring of Chinese cabbage leaf copper stress based on hyperspectral and CNN - LSTM[J]. Transactions of the Chinese Society for Agricultural Machinery, 2025, 56(6): 477 - 486. (in Chinese)

[22] Ouyang H, Tang L, Ma J, et al. Application of hyperspectral technology with machine learning for Brix detection of pastry pears[J]. Plants, 2024, 13(8): 1163.

[23] Huang Jiadong, He Hongyuan, Lv Rulin, et al. Non-destructive detection and classification of textile fibres based on hyperspectral imaging and 1D - CNN[J]. Analytica Chimica Acta, 2022, 1224: 340238.

[24] Jin Xin, Xie Yanping, Wei Xiushen, et al. Delving deep into spatial pooling for squeeze-and-excitation networks[J]. Pattern Recognition, 2022, 121: 108159.

[25] Bai Yuhan. ReLU function and derived function review[C]// SHS Web of Conferences. EDP Sciences, 2022,144:02006.

[26] Grattarola D, Zambon D, Bianchi FM, et al. Understanding pooling in graph neural networks[J]. IEEE Transactions on Neural Networks and Learning Systems, 2022, 35(2): 2708 - 2718.

[27] Li Jianjun, Han Yu, Zhang Ming, et al. Multi-scale residual network model combined with Global Average Pooling for action recognition[J]. Multimedia Tools and Applications, 2022, 81(1): 1375 - 1393.

[28] Montesinos L O A, Montesinos L A, Crossa J. Support vector machines and support vector regression[M]// Multivariate statistical machine learning methods for genomic prediction. Cham: Springer International Publishing, 2022: 337 - 378.

[29] Campos D, Wayo D D K, De Santis R B, et al. Evolutionary automated radial basis function neural network for multiphase flowing bottom-hole pressure prediction[J]. Fuel, 2024, 377: 132666.

[30] 张慧春,周子阳,边黎明,等. 基于 1DCNN 融合多源表型数据的杨树干旱胁迫评估方法[J]. 农业机械学报,2024, 55(9):286 - 296.
Zhang Huichun, Zhou Ziyang, Bian Liming, et al. Assessment of poplar drought stress level based on 1DCNN fusion of multi-source phenotypic data[J]. Transactions of the Chinese Society for Agricultural Machinery, 2024, 55(9): 286 - 296. (in Chinese)

[31] 王忠全,吴志鹏,周子平. 氮源配比及光强对水培蔬菜生长和品质的影响[J]. 上海蔬菜,2005(1):70 - 71.

[32] Zheng Anyuan, Peng Yang, Xu Shanfa, et al. Application of SVD combined with PCA in delineation and evaluation of ore-prospecting targets in the Gejiu tin polymetallic cluster region, SW China[J]. Ore Geology Reviews, 2023, 160: 105571.

[33] Zhou Pan, Xie Xingyu, Lin Zhouchen, et al. Towards understanding convergence and generalization of AdamW[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2024, 46(9): 6486 - 6493.

[34] Xu Peng, Xiang Zhiyu, Qiao Chengyu, et al. Adaptive multi-modal cross-entropy loss for stereo matching[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2024.