doi:10.6041/j.issn.1000-1298.2025.06.026

基于微型光谱仪和 Transformer 模型的便携式土壤全氮含量检测仪研究

剧伟良[1]　杨　玮[1,2]　宋亚美[1]　刘　楠[1]　李民赞[1,2]

(1. 中国农业大学智慧农业系统集成研究教育部重点实验室, 北京 100083;
2. 中国农业大学农业农村部农业信息获取技术重点实验室, 北京 100083)

摘要: 便携式土壤全氮含量近红外光谱检测仪具有快速、非破坏性和高效性等优点,但已开发的仪器多采用滤光片式设计,光谱通道数量有限会导致部分关键信息丢失,且无法采用基于深度学习的预测模型。随着微型光谱仪的商业化,本文开发了基于连续光谱的高精度土壤全氮含量检测仪。检测仪主要由 NIR - R210 型微型光谱仪、树莓派、触控屏、移动电源构成,利用微型光谱仪获取土壤光谱反射率,利用树莓派中嵌入的深度学习模型进行土壤全氮含量预测,然后在显示屏中输出预测结果。在中国农业大学上庄实验站采集了 600 份土壤样本,分别对偏最小二乘法、门控循环单元和 Transformer 3 种模型的预测性能进行了对比分析。结果表明,基于全光谱数据的 Transformer 深度学习模型表现最好,模型决定系数 R^2 为 0.89,均方根误差(RMSE)为 0.19 g/kg,预测偏差(RPD)为 2.96。进一步对检测仪进行田间实时原位测试,田间环境下预测结果 R^2 可达 0.83,精度较高,可为智慧农业中土壤养分实时检测与精准管理提供新的解决方案。

关键词: 土壤全氮; 近红外; Transformer; 深度学习; 树莓派

中图分类号: S126　　**文献标识码**: A　　**文章编号**: 1000-1298(2025)06-0268-09　　**OSID**:

Portable Soil Total Nitrogen Content Detector Based on Miniature Spectrometer and Transformer Model

JU Weiliang[1]　YANG Wei[1,2]　SONG Yamei[1]　LIU Nan[1]　LI Minzan[1,2]

(1. *Key Laboratory of Smart Agriculture System Integration*, *Ministry of Education*, *China Agricultural University*, *Beijing* 100083, *China*
2. *Key Laboratory of Agricultural Information Acquisition Technology*, *Ministry of Agriculture and Rural Affairs*, *China Agricultural University*, *Beijing* 100083, *China*)

Abstract: Portable near-infrared (NIR) spectroscopic detectors for soil total nitrogen content offer the advantages of rapid analysis, non-destructive measurement, and high efficiency. However, most existing instruments adopt filter-based designs with a limited number of spectral channels, which can lead to the loss of critical information and prevent the implementation of deep learning-based prediction models. With the commercialization of miniature spectrometers, a high-precision soil total nitrogen content detector was developed based on continuous spectral data. The detector primarily consisted of an NIR - R210 miniature spectrometer, a Raspberry Pi, a touchscreen display, and a portable power supply. The spectrometer was used to acquire soil spectral reflectance data, which were processed by a deep learning model embedded in the Raspberry Pi to predict soil total nitrogen content. The prediction results were then displayed in real time on the touchscreen. A total of 600 soil samples were collected from the Shangzhuang Experimental Station of China Agricultural University. The predictive performances of three models (partial least squares regression (PLSR), gated recurrent unit (GRU), and Transformer) were compared. Among them, the Transformer model based on full-spectrum data achieved the best performance, with a coefficient of determination (R^2) of 0.89, a root mean square error (RMSE) of 0.19 g/kg, and a residual predictive deviation (RPD) of 2.96. Further real-time in-situ field tests

收稿日期: 2025-03-16　修回日期: 2025-04-18

基金项目: 国家重点研发计划项目(2023YFD1701000)和中国农业大学 2115 人才工程项目

作者简介: 剧伟良(1996—),男,硕士生,主要从事土壤检测设备研究,E-mail: s20223081703@cau.edu.cn

通信作者: 李民赞(1963—),男,教授,博士生导师,主要从事农业信息化技术研究,E-mail: limz@cau.edu.cn

showed that the Transformer model maintained high accuracy under field conditions, with an R^2 of up to 0.83. This portable device provided an effective solution for real-time soil nutrient detection and precision management in smart agriculture.

Key words: soil total nitrogen; near-infrared; Transformer; deep learning; Raspberry Pi

0　引言

土壤全氮含量是土壤肥力的重要指标,确保土壤中充足的氮供应对于促进植物生长和提高作物产量至关重要。然而,过量的氮输入不仅会破坏土壤健康,还会对农业生产造成负面影响。因此,快速且准确地测定土壤全氮含量,确保合理施肥,对实现作物精准管理具有重要意义。随着智慧农业的不断发展,尤其是在田间变量施肥的应用场景中,对土壤检测技术提出了更高要求。在此背景下,能够快速、准确地获取土壤中全氮含量的技术变得尤为重要。近红外光谱技术凭借其快速、精准和非破坏性的优势,逐渐成为土壤全氮含量快速检测中最具潜力的技术之一[1-3]。近年来,研究人员开发了一系列基于近红外技术的便携式土壤全氮含量检测仪器,李民赞等[4]开发了一款基于卤钨灯光源的一分六石英光纤的便携式土壤全氮含量检测仪,根据检测仪吸光度和土壤全氮含量建立了不同预测模型,并选用 6 个敏感波长用于实现土壤全氮含量预测,实现了田间土壤全氮含量实时检测的目的。LIU 等[5]开发了一款基于相位锁定放大技术的便携式土壤全氮检测仪,通过优化遗传算法筛选了 8 个特征波长,并使用数字合成和运算放大器调制光信号,使用光电二极管和锁相放大器对光信号进行解调,有效过滤了环境光噪声,设备集成了 STM32 单片机和树莓派,并嵌入了 BP 神经网络,经过不同光照条件测试,设备预测决定系数 R^2 高于 0.8,均方根误差为 0.30 g/kg,该检测仪为现代农业中土壤养分快速检测和精准管理提供了可靠的技术支持。

上述设备由于滤波片成本和空间限制,只保留了少数光谱通道,仅能捕捉有限的光谱波段,这可能导致部分关键信息丢失,并使得设备只能搭载输入参数量较少的基准机器学习模型进行预测任务。并且上述设备工作原理是使用光电传感器将透过滤光片的反射光进行捕捉,转化为微弱电流,进而转化为电压传输给处理端,受环境以及电磁干扰较大,复杂的信息传递流程也会对测量精准度造成进一步的影响。与此同时,上述设备还需引入功率较高的卤素光源,并需搭配光电传感器与 AD 转换电路等复杂电路模块,这对设备内部空间排列提出了较高要求,开发难度较高。这些因素共同限制了设备在实际应用中的可靠性和效果。

近年来,低成本、符合人体工程学和小型化的微型光谱仪已逐渐实现商业化,使得便携式光谱设备可更方便地部署在农田进行原位检测分析[6-8],商用微型光谱仪的应用也对基于近红外光谱的土壤检测设备提出了新的要求,如何使用较为先进的微型光谱仪,实现便携式检测设备的田间开发与部署,成为亟待解决的重要问题。

随着深度学习的快速发展,结合特征筛选算法的深度学习模型逐渐成为土壤光谱分析领域的主流选择[9-10],然而,特征筛选方法在光谱分析中的应用也存在明显局限性,尽管特征筛选能够减少数据维度并提高计算效率,但其可能导致光谱中其他潜在重要信息的遗漏,限制了模型对全光谱数据的全面利用,从而限制了模型预测能力。此外,特征筛选往往需要根据具体应用场景和数据特性进行计算,以确保选出的特征在对应环境条件下的适用性,在面对不同场景时需要进行重新计算与筛选,这延长了系统开发周期,增大了工程实现与移植难度。

近年来,Transformer 作为一种基于自注意力机制的深度学习模型,凭借其强大的全局特征建模能力,广泛应用于序列分析[11-14]。Transformer 能够同时处理全序列数据,捕捉全局依赖关系,特别适合高通量光谱数据分析。通过自注意力机制,Transformer 有效捕捉长程依赖和复杂模式,能够高效处理大规模、高维数据,无需传统的特征筛选步骤,避免信息丢失,全面保留数据的内在结构和复杂关系,从而实现更精准的预测。树莓派等微处理器的使用,使得在便携式仪器嵌入处理大规模数据的复杂深度学习模型成为可能。因此,本研究提出一种基于 Transformer 模型的土壤全氮含量光谱预测方法,结合光谱仪和树莓派,开发一款便携式实时田间原位土壤全氮含量检测仪,实现高精度、低功耗和高便携性,为智慧农业中的土壤检测技术发展提供有效解决方案。

1　材料与方法

1.1　研究流程

研究流程如图 1 所示。首先,在田间采集土壤样本,在实验室分析采集土壤样本的光谱数据与全氮含量。第 2 步,筛选土壤全氮的光谱特征波段,并

对比不同模型后选择出最优模型。第 3 步,将模型嵌入树莓派中,并对设备进行集成开发。最后,对设备进行实验室与田间性能测试,确保设备稳定性。

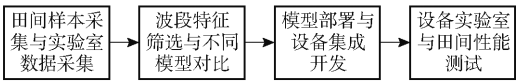

图 1 研究流程图

Fig. 1 Research flowchart

1.2 近红外微型光谱仪模块

所使用的 NIR - R210 型光谱仪,为接触式或微距测量,使用场景为固体样本的定量或定性分析测量,支持 MicroUSB 通信,内置 3 800 mA·h 锂电池与 2.5 W 卤素光源,波长为 900 ~ 1 700 nm,光学分辨率为 10 nm,全光谱波长数为 228(平均采样间隔约为 3.524 nm)。光谱仪整体结构如图 2a 所示,光谱仪光学原理如图 2b 所示,光谱仪模组采用光栅分光设计,光经过狭缝(Slit)并准直(Collimating)照射到光栅(Diffraction)上,由光栅分光后产生不同波长的光,并排射到数字镜像阵列(DLP)反射镜,通过编程控制每个微型反射镜,依次将对应波长的光反射到单点 InGaAs 探测器上,再由 ADC 转换为数字信号并解析成光谱曲线。与传统的分光光度计相比,NIR - R210 型光谱仪在光谱扫描精度上有一定降低,但实现了光源内置处理,避免了光纤接入和外置光源配备,提供了一种成本效益高、空间体积小、便于携带和集成的解决方案,使田间土壤快速实时的原位光谱分析成为可能。

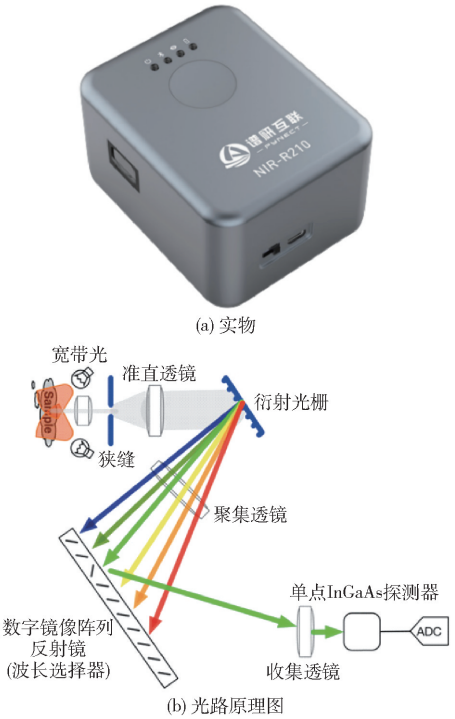

图 2 NIR - R210 型光谱仪实物与原理图

Fig. 2 NIR - R210 spectrometer and optical diagram

1.3 土壤样本与光谱数据采集

土壤样本采集于中国农业大学上庄实验站(40.141 1°N,116.189 4°E)玉米田,共采集 600 份。实验田使用不同施用量尿素对不同区域进行施肥,并在 7 d 后采集土壤样本,采集深度为地表以下 5 ~ 10 cm,样本装袋后带回实验室进行后续处理。去除其中的枯枝与石头等与土壤性质无关的杂质,标号后放入塑料杯中进行烘炉干燥、粉碎与过筛处理,获取干燥、直径小于 0.9 mm 的土壤颗粒以进行后续实验。

使用便携式 NIR - R210 型光谱仪进行光谱数据采集,将干燥过筛的土壤样本放于桌面纸张上,并使用光谱仪探测孔按压土壤样本表面,采集土壤光谱反射率信息。进一步使用杜马斯定氮仪测定样品全氮含量,将扫描光谱后的土壤样品称取 0.5 g,放入小坩埚中,置于定氮仪的样品进样窗口进行燃烧。通过计算机端运行 SKALAR 配套软件进行数据处理,获取样品全氮含量数据。

1.4 特征筛选算法

土壤光谱数据量庞大且含有大量冗余信息和噪声,特征筛选是传统模型提高计算效率和预测性能的重要方法。竞争性自适应重加权采样算法(Competitive adaptive reweighted sampling,CARS)是一种基于回归系数的特征选择方法,通过递归特征消除和加权采样技术,结合偏最小二乘法(PLS),高效选取信息量大的波长特征,剔除冗余波段,避免过拟合,从而提高预测精度[15-17]。

1.5 树莓派模块

选用树莓派 4B 作为运算平台,树莓派体积小巧且功耗低,非常适合便携式设备的开发需求。树莓派 4B 拥有 1.5 GHz 的 64 位四核处理器,并有 4 GB的 RAM 与 40 个 GPIO 引脚,2 个 USB 3.0 端口,2 个 USB 2.0 端口,可连接鼠标键盘等外设,可在田间环境调试设备,此外,树莓派还支持 Python 等多种编程语言,可搭载基于 Python 和 PyTorch 框架的深度学习模型,同时拥有强大的非线性计算能力。

1.6 基于深度学习的回归预测模型

为探索深度学习在近红外光谱分析领域的潜在优势,并充分利用注意力机制的优势,选择 3 种模型进行对比:偏最小二乘回归(PLSR)、门控循环单元(Gated recurrent unit,GRU)和 Transformer 模型。

PLSR 是常规的化学计量学方法,为深度学习模型提供基准参考。GRU 是一种基于更新门和重置门机制的循环神经网络模型,能够高效地处理序列数据[18-20]。GRU 通过动态调整信息的记忆与遗

忘，捕捉时间序列中的依赖关系，适用于非线性和复杂特征的建模任务。Transformer 模型是一种基于自注意力机制的深度学习模型，与 GRU 循环神经网络不同，Transformer 移除了循环结构，通过并行计算实现高效的数据处理能力[21-24]。Transformer 的自注意力机制能够动态计算序列中每个波长与其他波长之间的关系，有效捕捉光谱数据中的全局依赖和复杂特征。同时，多头自注意力机制允许模型在不同子空间中并行学习多样化的特征，从多个角度提取光谱信息。这些特性使 Transformer 模型在光谱数据建模中表现出强大的优势。在土壤全氮含量预测中，光谱数据与全氮含量之间存在复杂的非线性依赖关系，传统方法可能难以全面捕捉这些关系。而 Transformer 模型凭借其自注意力机制，能够动态调整各光谱波段之间的权重，可直接从高维原始数据中捕捉全局和局部的依赖模式，从而更好地表征光谱数据与全氮含量之间的潜在关联。这种机制不仅避免了传统特征筛选可能造成的信息丢失，还能灵活处理波段间复杂的非线性关系，从而提升模型的预测精确性和稳定性。

本文提出一种基于 Transformer 架构的回归预测模型，用于从样本的高维光谱数据中预测对应的土壤全氮含量，并最终嵌入全氮含量检测系统。模型整体结构示意图如图 3 所示，首先，利用线性层对输入的高维光谱数据进行降维，将其映射到低维特征空间，以满足 Transformer 输入的维度要求，同时降低计算复杂度并保留关键信息。之后采用 Transformer 编码器（Encoder）层，通过多头自注意力机制有效捕捉特征之间的全局关系，将输入特征映射到查询（Query）、键（Key）和值（Value）向量，通过点积计算实现信息交互。随后编码器通过前馈神经网络进一步提升特征的非线性表达能力，在此过程中，编码器引入残差连接和层归一化，确保模型训练的稳定性和梯度传播的高效性。通过堆叠多个编码层，模型能够逐层提炼特征，增强对复杂依赖关系的建模能力。最后，通过线性层将编码器输出的低维特征映射为标量值，实现对土壤全氮含量的回归预测。

注意力机制核心公式为

$$\mathrm{Attention}(\boldsymbol{Q},\boldsymbol{K},\boldsymbol{V}) = \mathrm{softmax}\left(\frac{\boldsymbol{Q}\boldsymbol{K}^{\mathrm{T}}}{\sqrt{d_k}}\right)\boldsymbol{V} \quad (1)$$

式中　$\boldsymbol{Q}$——查询矩阵，由输入向量 z 通过线性变换生成

$\boldsymbol{K}$——键矩阵，由输入向量 z 通过线性变换生成

$\boldsymbol{V}$——值矩阵，由输入向量 z 通过线性变换生成

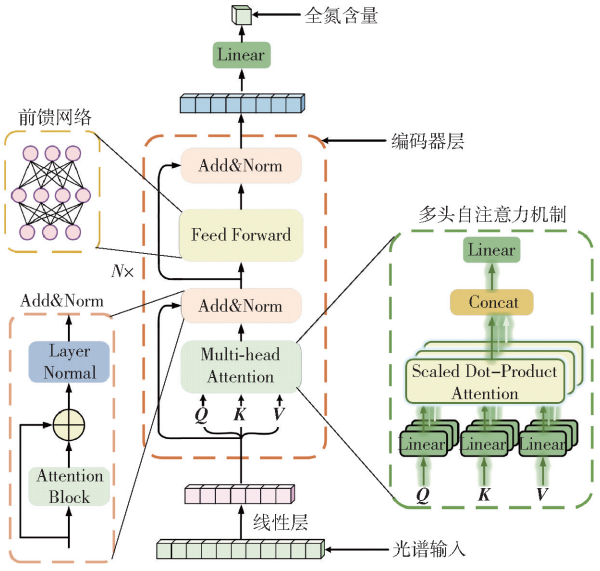

图 3　Transformer 模型与多头注意力机制示意图

Fig. 3　Transformer model and multi-head attention mechanism diagram

$\sqrt{d_k}$——缩放因子，防止点积结果过大

为全面评估模型性能，选取 3 个评价指标：决定系数（R^2）、均方根误差（RMSE）和预测偏差（RPD）。

2　结果与讨论

2.1　系统开发

2.1.1　便携式土壤全氮含量检测仪系统构成

便携式土壤全氮含量检测仪系统构成如图 4a 所示。以 NIR－R210 型光谱仪和树莓派为核心部件，搭配移动电源与 3.5 英寸触控屏，通过整合硬件与软件，实现光谱数据采集、处理及原位实时预测。设备通过内置的模型，能够实时预测土壤全氮含量，并将结果即时显示在屏幕上，为用户提供快速、直观反馈。为满足便携性需求，将关键部件集成在有限空间内，配备屏幕和内置电源，无需外接电源即可运行，支持开机自启动、循环采集及低功耗运行，可在复杂田间环境中长时间稳定工作，满足田间原位实时检测的多样化需求。设备整体结构如图 4b 所示：树莓派作为系统的控制与运算核心，负责设备控制与模型运算；NIR－R210 型光谱仪负责采集样本的光谱信息并传递给树莓派，使用光谱仪探测孔进行土壤光谱采集；为尽可能压缩体积，采用椭圆形移动电源，并将其放置在手柄处以节约设备内部空间；设备还配备了一个可触控 LCD 屏幕，用于显示操作界面和检测结果。

2.1.2　接口开发与模型调用

为实现光谱仪与树莓派间的数据传输，在树莓派端使用 Python 调用光谱仪的动态库文件，通过 USB 接口读取光谱仪传递的数据；树莓派读取

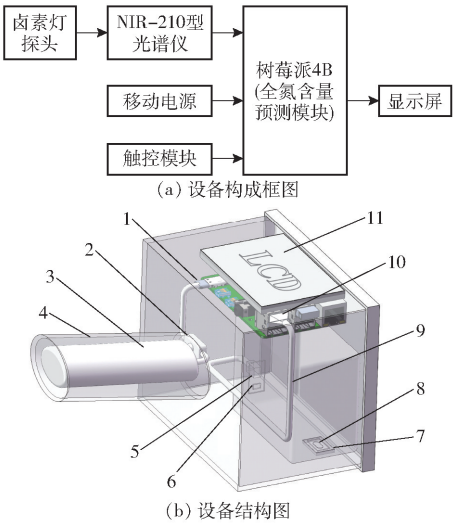

图 4 设备构成框图与结构图

Fig. 4 System block diagram and structural layout of device

1. 树莓派电源接口 2. 移动电源开关 3. 移动电源 4. 检测设备手柄 5. 光谱仪数据接口 6. 光谱仪电源开关 7. 设备扫描孔 8. 光谱仪探测口 9. 光谱仪数据线 10. 树莓派数据接口 11. LCD 屏幕

数据后,将其转换为合适的格式输入至深度学习模型进行预测,并保存采集的光谱数据与全氮含量预测值。

为提升设备的易用性和自动化水平,设计实现程序的封装与自启动功能,将目标程序指向性路径嵌入树莓派的启动配置文件中,实现设备开机后自动加载。系统循环调用函数,可进行连续预测。

2.2 系统性能测试

光谱反射率是土壤全氮含量精准预测的基础。为确保全氮含量检测设备的预测效果,进行反射率白板校正、零点校正与光谱稳定性测试。由于波段 900 ~ 950 nm 及 1 650 ~ 1 700 nm 为光谱仪边缘波段,受设备本身性能影响,信噪比较低,因此硬件性能测试与设备应用仅使用波段 950 ~ 1 650 nm。

2.2.1 白板校正与零点校正实验

为检验光谱仪反射率精度,使用标准白板(反射率为 100%)对光谱仪进行校正测试,使用光谱仪对白板进行 10 次连续扫描,之后进一步对光谱仪进行零点校正,控制卤素灯关闭,模拟无反射光情况,并连续扫描 10 次,扫描结果如图 5 所示。结果表明,在波段 950 ~ 1 650 nm,光谱仪在扫描白板时能够始终保持曲线反射率接近 100%,在关闭卤素灯情况下,接收到的反射率接近 0。表明光谱仪能够准确地获取物质反射率。该过程有效验证了光谱仪的校准精度,确保了仪器能够提供可靠的反射率测量结果,为后续实验或测量提供了基准。

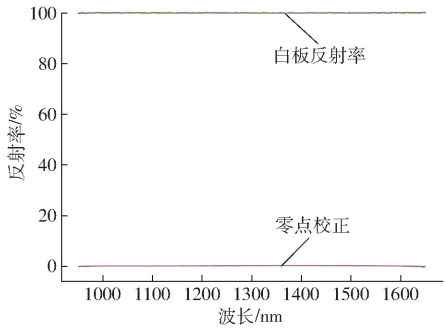

图 5 白板校正与零点校正结果

Fig. 5 Whiteboard calibration and zero point calibration

2.2.2 稳定性测试实验

为评估光谱仪的稳定性,对同一土壤样本进行连续 10 次扫描,结果如图 6 所示。在目标光谱波段范围内(950 ~ 1 650 nm),光谱曲线变化非常小。这表明光谱仪在重复测量中能够保持高度一致的性能,未出现明显的漂移或波动,能够持续、准确地获取土壤的光谱信息。

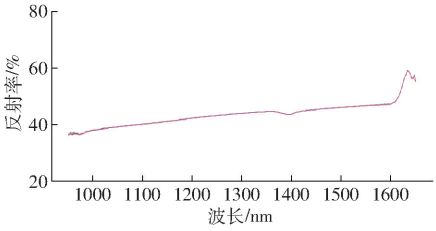

图 6 光谱仪稳定性测试结果

Fig. 6 Spectrometer stability test

2.2.3 树莓派端口光谱采集测试与验证

利用树莓派的 USB - A 接口连接光谱仪的 Micro - USB 数据端口,通过调用光谱仪的动态链接库文件,实现与光谱仪通信。在成功建立连接后,运行程序并采集光强信息。为验证光谱仪在树莓派端口运行下能否稳定进行光谱反射率采集,评估其数据采集和通信的稳定性与精度,采用标准白板作为参照物,将采集的光强信息与标准白板反射光强进行对比,进而转换为反射率形式。使用白板完全遮盖光谱仪的采集端口,对全波段范围内的反射率进行扫描与输出。标准白板在全波段范围内的理论反射率为 100%,通过光谱仪采集光强并转换为反射率,结果发现,除边缘波段存在轻微偏差外,其余波段反射率均接近 100%。结果表明,集成 NIR - R210 型光谱仪、树莓派以及相关软件的土壤全氮含量检测仪性能表现良好,可以稳定、高精度实现光谱采集、数据通信和数据处理。

2.3 深度学习模型建立

2.3.1 土壤数据信息采集

对土壤样本进行处理后,获取 600 份土壤样本,经过干燥过筛等一系列预处理后,对样本进行光

谱与全氮含量采集。通过 SKALAR 软件对杜马斯定氮仪燃烧数据进行处理，可得土壤全氮含量最小值、最大值、平均值和标准差分别为 0.96、2.98、1.84、0.54 g/kg。

分析表明，土壤样本全氮含量范围较广，样本间存在较大差异性，可以保证数据集的多样性和一致性，有利于后续的模型建模和训练。样本原始光谱如图 7 所示。

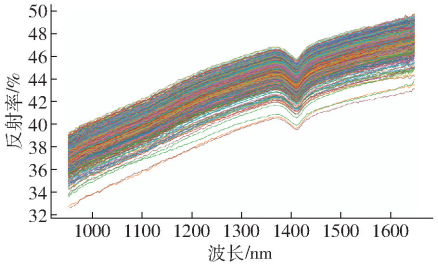

图 7　土壤样本原始光谱曲线

Fig. 7　Spectral curves of soil sample

2.3.2　CARS 特征筛选结果

对光谱数据进行 CARS 特征筛选，结果如图 8 所示。由图 8 可知，在土壤全氮含量预测中，光谱波段 950 ~ 1 650 nm 对预测的贡献程度呈现显著差异。波段 1 360 ~ 1 510 nm 显示出对全氮含量的高度相关性，相比之下，波段 950 ~ 1 300 nm 和 1 510 ~ 1 650 nm 对全氮含量预测的贡献相对较低，这些区域可能包含较多噪声或与目标变量关联不大的特征。

2.3.3　模型预测结果与性能分析

将全谱光谱数据和 CARS 特征光谱数据与对应的全氮含量真值分别输入 PLSR、GRU 和 Transformer 模型中进行训练和预测。预测结果如图 9 所示。

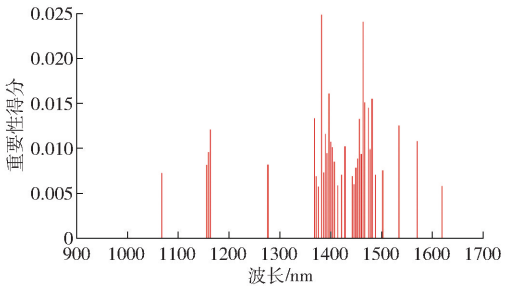

图 8　CARS 特征筛选结果

Fig. 8　CARS feature screening results

从图 9 可见，Transformer 模型在利用土壤光谱数据进行全氮含量预测时展示出卓越的性能，特别是其在全谱条件下的表现优于其他模型组合，其 R^2 达到 0.89，RMSE 为 0.19 g/kg，RPD 为 2.96，显著高于其他模型。在 GRU 与 PLSR 模型中，经过 CARS 特征筛选处理的模型相较于全谱都取得了更好的效果，这表明在 GRU 模型和 PLSR 模型中，特征筛选发挥了较为重要作用。

基于全谱的 Transformer 模型取得了最好的预测效果，这主要归功于它的多头注意力机制，该机制使模型能在全局范围内监测数据波动，并通过精确的注意力计算捕获关键信息，从而能够直接从光谱全谱进行建模。这种能力特别适合于处理数据维度高，波段之间还存在密切相关性以及复杂非线性关系的土壤光谱数据。

多头自注意力的 QKV 机制对于精确预测全氮

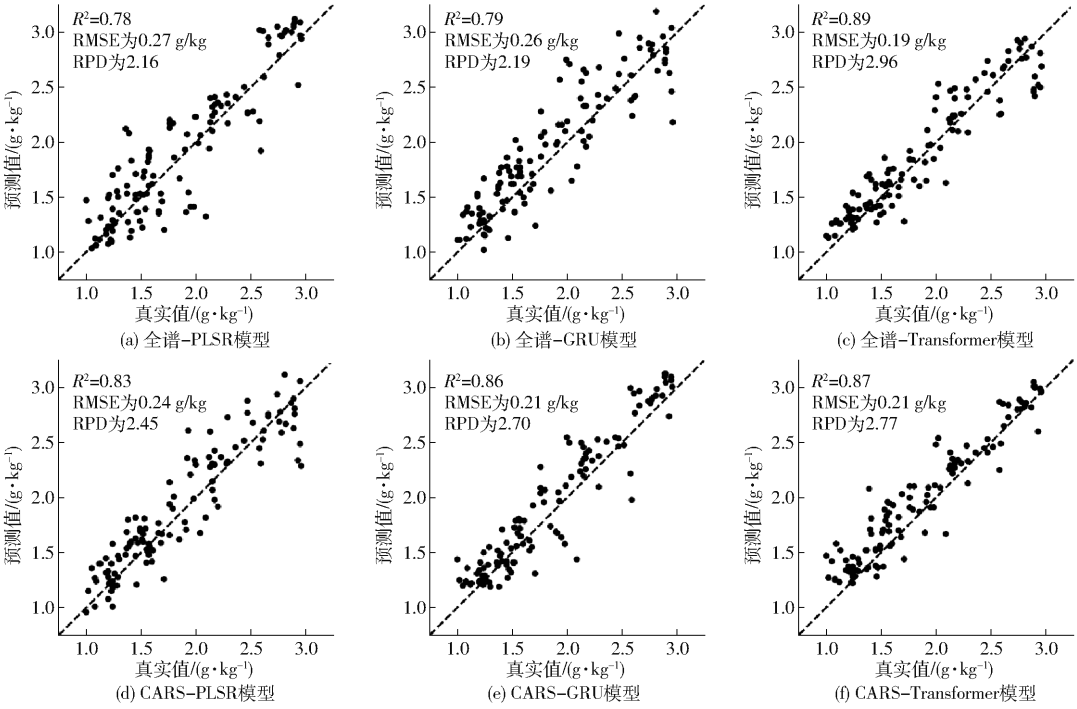

图 9　模型预测结果

Fig. 9　Model prediction results

含量发挥了重要作用。在 QKV 机制中,Query 表示模型需要查询的信息特性,例如某些光谱波段之间的关联性及其对全氮含量预测的重要性。Key 表示可能与 Query 匹配的光谱特性,例如单个波段的特性或波段之间的组合关系,用于判断哪些光谱数据与当前查询相关。Value 则是携带对输入特征进行加权后的具体内容。通过计算 Query 和所有 Key 之间的相似度,模型能够动态调整每个 Value 的权重。如果某个 Key 和 Query 的相似度较高,则表示该波段特性与当前全氮含量预测任务高度相关,模型会赋予其更大的权重,从而使对应的 Value 在预测中发挥更大的作用。反之,如果相似度较低,则该 Value 的影响会被削弱。这一过程使得模型能够自动关注对全氮含量预测最为重要的波段特性,同时忽略无关或冗余数据。这种动态权重调整机制正是 Transformer 模型的核心优势,使其在处理复杂的光谱数据和高维非线性特性时具有强大的建模能力。

CARS 算法通过特征选择在一定程度上优化了输入数据结构,提高了模型的计算效率。然而,这种优化也可能导致部分关键信息丢失。CARS 算法通过迭代选择与响应变量最相关的特征,可能忽略了那些单独作用不显著但与其他特征协同作用时极其重要的信息。这种选择性特征筛选方式可能限制模型对土壤光谱数据中全局性和非线性关系的全面捕捉。

在使用 CARS 特征筛选后,PLSR 和 GRU 模型相较于使用全谱输入的数据取得了更好的预测效果。然而,对于 Transformer 模型而言,特征筛选并未带来相应的性能提升,反而由于简化了输入特征,未能充分发挥 Transformer 在全局特征建模中的优势,导致其预测效果相对全谱输入较差。

相较于 Transformer 模型,特征筛选处理对 GRU 模型预测效果具有显著的正向提升。经过特征筛选处理后,GRU 模型 R^2从 0.79 提升至 0.86,RMSE 减小 0.05 g/kg,RPD 提升 0.51。GRU 作为一种基于序列数据的记忆网络,依靠门控机制动态管理信息的流动和存储,其预测能力与输入数据的序列长度和特征质量密切相关。当输入维度适中且信息量符合模型的记忆容量时,GRU 能够较好地捕捉数据的关键特征,从而表现出较好的预测效果。

然而,当输入包含过长的序列信息时,GRU 的门控机制可能会受到大量冗余特征和噪声的干扰,导致隐藏状态的容量被无关信息占用,削弱模型对关键特征的提取能力。这种情况下,模型不仅效率降低,还可能因信息拥堵而对远距离依赖的建模能力不足。特征筛选方法(如 CARS 算法)能够显著减少输入数据的冗余特征,通过选择与目标变量最相关的特征,使输入数据更加简洁。这种优化有助于减少噪声的干扰,同时提高模型对关键特征的关注度。通过简化数据维度,CARS 特征筛选能够有效缓解 GRU 在处理高维序列时的负担,使模型的门控机制更专注于关键特征之间的关系建模,从而显著提升预测性能。

为减小系统开发工作量并提升预测精度,选用基于全谱的 Transformer 模型作为系统的最终模型。Transformer 模型能够直接处理全光谱数据,无需复杂的特征筛选步骤,这有效简化了数据预处理流程,减小了开发的复杂性。此外,Transformer 模型具有卓越的全局建模能力,可以捕捉光谱数据中的全局依赖关系和非线性特征,为全氮含量预测提供高度精确且稳定的支持。这一模型的选择不仅优化了系统开发过程,还确保了便携式检测仪在性能和精度上的可靠性。

2.4 检测系统开发与性能测试

为实现模型在系统中的嵌入应用,首先需将模型的框架代码集成至树莓派的 Python 程序中,并导入训练好的最优模型参数进行加载。随后,将通过光谱仪实时采集的反射率数据输入该模型,完成土壤参数的计算与预测,并输出最终结果。预测代码体积为 10 KB,模型内存占用量为 1.1 MB,整体模型轻量,具备较快的运行效率。

为提升设备的连续运行能力并实现多次预测功能,系统设计中引入图形用户界面(GUI),通过设置“重复采集”按键,用户可一键触发光谱仪的重复采集与模型预测操作,实现数据的持续获取与实时分析。为防止用户误触其他界面区域,GUI 界面设置为全屏覆盖模式,从而增强操作的稳定性与安全性。

稳定性与可重复性是衡量土壤全氮含量检测系统性能的重要指标。为评估本系统在不同全氮含量条件下的检测稳定性,从实验室现有的土壤样本中选取 5 个全氮含量不同的样本(1.22、1.61、2.13、2.51、2.89 g/kg),进行稳定性检测。对同一样本进行多次独立测试,并分别记录每一次的采样结果,如图 10 所示。从图 10 可以看出,多次检测结果整体较为稳定,数值波动较小,表明本系统在重复测量同一样本时具有较好的稳定性与重复性,能够满足田间快速检测的一致性要求。

2.5 田间验证实验

为评估便携式田间实时检测设备的实际应用精度,进行田间测试,在实验田区域内,东西间隔 3 m,南北间隔 5 m 进行实时预测,并同时对样本进行采

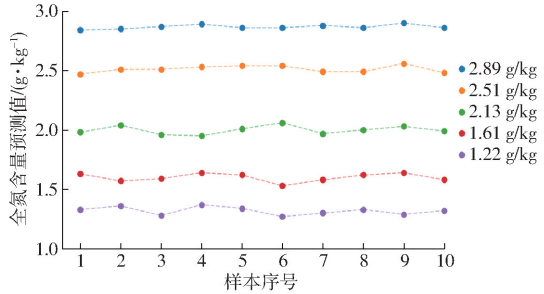

图 10　土壤全氮含量检测稳定性测试

Fig. 10　Stability test of soil total nitrogen detection

样，共采集样本 60 个，并将采集的土壤样本带回实验室进行化学分析，获取全氮含量真实值。田间检测实验如图 11 所示。

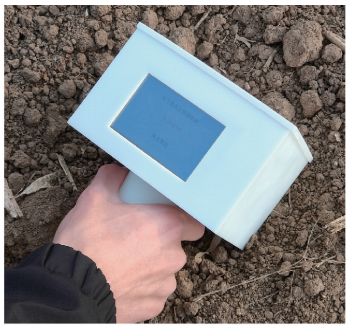

图 11　田间检测实验

Fig. 11　Instrument interface display

田间原位实时预测结果 R^2 为 0.83，RMSE 为 0.23 g/kg，RPD 为 2.43，如图 12 所示。

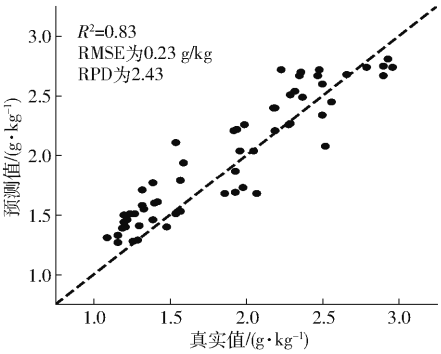

图 12　田间实时预测结果

Fig. 12　Real-time field prediction results

由图 12 可知，相较于模型训练效果，田间实时预测模型决定系数有所降低。分析其原因，首先是土壤水分的影响，田间土壤通常含有较高的水分，水分增加过程中，土壤表面逐渐形成水膜，进而影响土壤光谱反射特性[25-26]。土壤颗粒度差异也可能导致精度降低，在实验室条件下，土壤样本通常被统一进行过筛处理，以减少颗粒结构对光谱信号的影响，然而田间土壤颗粒度较为复杂，颗粒大小不均可能导致光谱信号在采集过程中产生更多散射，使得在采集过程中引入了更多噪声，降低了数据准确性[27]。这些因素使得田间实时预测精度难以达到模型训练时的水平。因此，如何减小水分和粒度对光谱建模的影响，在样本不足情况下通过学习迁移等手段实现更高精度的预测，仍然是实际应用中的重要挑战之一。

3　结论

（1）通过光谱数据预处理及 CARS 特征筛选算法，分别对比了全谱与特征筛选条件下 PLSR、GRU 和 Transformer 3 种模型的性能表现，结果表明，基于全谱的 Transformer 模型表现最优，其 R^2 达到 0.89，RMSE 为 0.19 g/kg，RPD 为 2.96。

（2）基于预测精度与开发需要，选用基于全谱的 Transformer 模型作为系统核心算法，进一步结合微型光谱仪与树莓派，开发了一款便携式土壤全氮含量检测仪。树莓派通过调用光谱仪实现光强数据采集，并将其进一步转化为反射率数据，输入 Transformer 模型进行实时预测并在交互界面进行展示。对 NIR－R210 型微型光谱仪的精准度与稳定性测试结果表明，仪器工作稳定、性能满足要求。

（3）在田间使用开发的检测仪进行土壤全氮含量原位实时检测，R^2 为 0.83，预测精度较好，表明该检测仪可用于田间土壤全氮含量原位实时检测，可为智慧农业中的土壤养分实时检测与精准管理提供技术支持。

参　考　文　献

[1] FU X, LENG G, ZHANG Z, et al. Enhancing soil nitrogen measurement via visible-near infrared spectroscopy: integrating soil particle size distribution with long short-term memory models[J]. Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy, 2025, 327: 125317.

[2] METZGER K, BRAGAZZA L. Prediction of nitrogen, active carbon, and organic carbon-to-clay ratio in agricultural soils by in-situ spectroscopy[J]. European Journal of Soil Science, 2024, 75(3): e13508.

[3] SANTASUP N, THEANJUMPOL P, SANTASUP C, et al. Development of near-infrared spectroscopy (NIRS) for estimating organic matter, total carbon, and total nitrogen in agricultural soil[J]. MethodsX, 2024, 13: 102798.

[4] 李民赞, 姚向前, 杨玮, 等. 基于卤钨灯光源和多路光纤的土壤全氮含量检测仪研究[J]. 农业机械学报, 2019, 50(11): 169－174.
LI Minzan, YAO Xiangqian, YANG Wei, et al. Design of new portable detector for soil total nitrogen content based on high-power tungsten halogen lamp and "one-six" special optical fiber[J]. Transactions of the Chinese Society for Agricultural

Machinery,2019,50(11):169 – 174. (in Chinese)
[5] LIU K, ZHANG Y, GAO T, et al. A handheld rapid detector of soil total nitrogen based on phase-locked amplification technology[J]. Computers and Electronics in Agriculture, 2024, 224: 109233.
[6] BEC K B, GRABSKA J, PFEIFER F, et al. Rapid on-site analysis of soil microplastics using miniaturized NIR spectrometers: key aspect of instrumental variation[J]. Journal of Hazardous Materials, 2024, 480: 135967.
[7] SHI Z, REN Z, YANG Z, et al. Deployment strategy of multiple miniaturized near-infrared spectrometers based on spectral transfer for characterizing soil organic matter and nitrogen[J]. Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy, 2024,320: 124620.
[8] LUCENA P G C, AQUINO R V S, SOUSA J E S, et al. Mineral and particle-size chemometric classification using handheld near-infrared instruments for soil in Northeast Brazil[J]. Geoderma Regional, 2024, 38: e00819.
[9] BAI Z, CHEN S, HONG Y, et al. Estimation of soil inorganic carbon with visible near-infrared spectroscopy coupling of variable selection and deep learning in arid region of China[J]. Geoderma, 2023, 437: 116589.
[10] WANG Y, CHEN S, HONG Y, et al. A comparison of multiple deep learning methods for predicting soil organic carbon in Southern Xinjiang, China[J]. Computers and Electronics in Agriculture, 2023, 212: 108067.
[11] LI C, QIAN G. Stock price prediction using a frequency decomposition based GRU transformer neural network[J]. Applied Sciences, 2022, 13(1): 222.
[12] YANG H, ZHANG Z, LIU X, et al. Monthly-scale hydro-climatic forecasting and climate change impact evaluation based on a novel DCNN-Transformer network[J]. Environmental Research, 2023, 236: 116821.
[13] CHEN Z, ZHOU R, REN P. Spectraformer: deep learning model for grain spectral qualitative analysis based on Transformer structure[J]. RSC Advances, 2024, 14(12): 8053 – 8066.
[14] 廖娟,梁业雄,姜锐,等. 基于 Transformer – FNN 和无人机高光谱遥感技术的棉花黄萎病危害等级分类研究[J]. 农业机械学报,2025,56(2):240 – 251.
LIAO Juan, LIANG Yexiong, JIANG Rui, et al. Classification of cotton *Verticillium wilt* severity levels based on Transformer – FNN and UAV hyperspectral remote sensing technology[J]. Transactions of the Chinese Society for Agricultural Machinery, 2025,56(2):240 – 251. (in Chinese)
[15] QIN G, WANG H, WANG H, et al. Characterization of unplasticized polyvinyl chloride windows by confocal raman microspectroscopy and chemometrics[J]. Analytical Letters, 2024, 57(13): 2157 – 2169.
[16] LU Z, YU H, YIN Y, et al. Determination of the acid and peroxide values of vegetable oils by Raman spectroscopy with competitive adaptive reweighted sampling (CARS) and back propagation neural network (BPNN)[J]. Analytical Letters, 2024, 57(14): 2289 – 2306.
[17] XU Z, LI X, CHENG W, et al. Rapid and accurate determination methods based on data fusion of laser-induced breakdown spectra and near-infrared spectra for main elemental contents in compound fertilizers[J]. Talanta, 2024, 266: 125004.
[18] KEYIMU R, TUERXUN W, FENG Y, et al. Hospital outpatient volume prediction model based on gated recurrent unit optimized by the modified cheetah optimizer[J]. IEEE Access, 2023,11: 139993 – 140006.
[19] FANTINI D G, SILVA R N, SIQUEIRA M B B, et al. Wind speed short-term prediction using recurrent neural network GRU model and stationary wavelet transform GRU hybrid model[J]. Energy Conversion and Management, 2024, 308: 118333.
[20] 郭丰玮,王鹏新,刘峻明,等. 基于遥感多参数和 VMD – GRU 的冬小麦单产估测[J]. 农业机械学报,2024,55(1):164 – 174,185.
GUO Fengwei, WANG Pengxin, LIU Junming, et al. Yield estimation of winter wheat based on multiple remotely sensed parameters and VMD – GRU[J]. Transactions of the Chinese Society for Agricultural Machinery,2024,55(1):164 – 174,185. (in Chinese)
[21] VASWANI A. Attention is all you need[J]. Advances in Neural Information Processing Systems, 2017,30.
[22] XIE W, ZHAO M, LIU Y, et al. Recent advances in Transformer technology for agriculture: a comprehensive survey[J]. Engineering Applications of Artificial Intelligence, 2024, 138: 109412.
[23] ZHANG J, YANG S, HU S, et al. A dairy goat tracking method via lightweight fusion and Kullback Leibler divergence[J]. Computers and Electronics in Agriculture, 2023, 213: 108189.
[24] XIE Y, LU Z. Novel deep-learning method based on LSA-Transformer for fault detection and its implementation in penicillin fermentation process[J]. Measurement, 2024, 235: 114871.
[25] WANG Y, HU B, HONG Y, et al. Minimize of moisture effects from laboratory simulations of in-situ Vis – NIR spectral for the prediction of soil salinity[J]. Infrared Physics & Technology, 2024, 137: 105194.
[26] SEO C, JO H Y, BYUN Y, et al. Montmorillonite content prediction in bentonite using Vis – NIR spectroscopy and PLSR analysis: effects of humidity and mineralogical variability[J]. Geoderma, 2024, 448: 116980.
[27] HE Q, ZHANG H, LI T, et al. NIR spectral inversion of soil physicochemical properties in tea plantations under different particle size states[J]. Sensors, 2023, 23(22): 9107.