

基于不同卷积神经网络模型的红壤有机质高光谱估算

钟 亮, 郭 熙*, 国佳欣, 徐 喆, 朱 青, 丁 萌

(1. 江西农业大学国土资源与环境学院, 南昌 330045; 2. 江西省鄱阳湖流域农业资源与生态重点实验室, 南昌 330045)

摘 要: 以卷积神经网络 (Convolutional Neural Network, CNN) 为代表的深度学习方法因具有强大的特征学习能力已被广泛应用于计算机视觉、自然语言处理等领域, 但在土壤高光谱遥感领域研究较少。为探究其在小样本数据集下, 通过高光谱数据估算土壤有机质 (Soil Organic Matter, SOM) 的可行性, 以江西省奉新县北部为研究区, 248 个红壤样本为研究对象。对比分析深度学习方法 CNN、多层感知器 (Multilayer Perceptron, MLP)、常用的机器学习方法随机森林 (Random Forest, RF) 和支持向量机 (Support Vector Machine, SVM) 在不同光谱预处理下的建模效果, 在此基础上分别建立 5 种各具特点的 CNN 结构模型, 以探讨不同网络结构的建模效果, 包括最早提出的 LeNet-5、具有大卷积核的 AlexNet-8、采用小卷积核的 VGGNet-7、含有 Inception 结构的 GoogLeNet-7 以及使用残差学习的 ResNet-13。此外, 还探讨了 VGGNet 模型在 5 种不同网络深度下的模型效果。结果表明: 在使用原始光谱的情况下, CNN 模型依然能够取得较好的建模效果 (相对分析误差 > 2.5); 浅层 CNN 结构优于深层建模效果, 超参数较小的卷积核、步长和池化范围有助于提取更多的特征数量, 提高建模精度; VGGNet-7 网络结构在所有模型中表现最为突出, 在训练集上决定系数为 0.895, 均方根误差为 4.145 g/kg, 相对分析误差为 3.447, 在验证集上决定系数为 0.901, 均方根误差为 4.647 g/kg, 相对分析误差为 3.291, 具有极好的模型估测能力; 680、1 360、1 390、1 920、2 310 nm 及其附近是 VGGNet-7 建模过程中所提取的 SOM 重要特征波长。因此, CNN 能够简化光谱预处理过程, 在土壤高光谱遥感小样本建模中具备可行性, 具有非常广阔的应用前景, VGGNet-7 可以应用于红壤地区通过高光谱数据快速、准确的估算 SOM 含量。

关键词: 土壤; 模型; 卷积神经网络; 有机质; 高光谱

doi: 10.11975/j.issn.1002-6819.2021.01.025

中图分类号: S127; S153.6⁺21

文献标志码: A

文章编号: 1002-6819(2021)-01-0203-10

钟亮, 郭熙, 国佳欣, 等. 基于不同卷积神经网络模型的红壤有机质高光谱估算[J]. 农业工程学报, 2021, 37(1): 203-212.

doi: 10.11975/j.issn.1002-6819.2021.01.025 <http://www.tcsae.org>

Zhong Liang, Guo Xi, Guo Jiaxin, et al. Hyperspectral estimation of organic matter in red soil using different convolutional neural network models[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2021, 37(1): 203-212. (in Chinese with English abstract) doi: 10.11975/j.issn.1002-6819.2021.01.025 <http://www.tcsae.org>

0 引 言

土壤有机质 (Soil Organic Matter, SOM) 是土壤的重要组成部分, 是植物营养的主要来源之一, 在提高土壤肥力、改善土壤物理性质和促进作物生长等方面发挥着重要的作用^[1]。传统通过化学分析方法检测 SOM 的过程复杂、耗时耗力^[2]。近年来, 高光谱技术以其光谱分辨率高和波段信息丰富的优点迅速发展起来, 在快速估测土壤属性信息上有其强大的优势^[3-4], 能够为精准农业的实施提供监测依据, 有着广泛的研究和应用前景^[5-6]。目前, 已有众多学者进行了高光谱技术估算 SOM 的研究, Shi 等^[7]通过空间约束局部偏最小二乘法 (Partial Least Squares Regression, PLSR) 和可见光近红外光谱估算 SOM, 其建模精度比局部 PLSR 和单独 PLSR 高; 章涛

等^[8]使用小波能量特征降低高光谱的噪声, 提升了 SOM 含量的估测精度; 张东辉等^[9]利用神经网络方法通过航空高光谱数据估算黑土 SOM 含量; 章海亮等^[10]应用遗传算法与连续投影算法相结合的方法, 提取光谱特征波段, 检测 SOM 含量; 谢文等^[11]和国佳欣等^[12]分别采用径向基函数 (Radial Basis Function, RBF) 神经网络组合模型和 PLSR-BP (PLSR and Back Propagation, PLSR-BP) 组合模型进行红壤 SOM 的高光谱估测, 结果都比单一模型效果好。

现有研究利用光谱分析估算土壤有机质时, 较多通过构建多种机器学习方法的组合模型以提高建模精度, 但由于高光谱数据具有成百上千个波段, 各波段间共线性较强且关系复杂, 传统机器学习方法处理能力较为有限。在研究过程中需要进行大量光谱数据的预处理, 如光谱增强、光谱曲线去噪以及光谱的数学变换等^[13], 并且预先提取特征波段, 然后做建模对比试验, 寻求适用于研究区域或者研究对象的光谱预处理与建模方法, 过程较为繁琐。探索一种既能保证建模精度又能简化数据预处理过程的方法显得十分重要。

随着计算机技术的发展, 以卷积神经网络 (Convolutional Neural Network, CNN) 为代表的深度学

收稿日期: 2020-07-23 修订日期: 2020-10-10

基金项目: 国家重点研发计划项目 (2017YFD0301603); 国家自然科学基金项目 (41361049); 研究生创新专项基金 (NDYC2020-S008)

作者简介: 钟亮, 研究方向为农业遥感与信息技术。

Email: zhongliang1007@163.com

*通信作者: 郭熙, 博士, 教授, 研究方向为农业遥感与信息技术及土地资源利用。Email: xig435@163.com

习逐渐发展起来,其通过卷积和池化逐层提取特征,同时具有权值共享和局部连接的特性,减少了需要训练参数的数量,使得模型更容易优化^[14]。虽然 CNN 在自然语言处理、图像和语音识别等领域发展迅速,但 CNN 在土壤高光谱遥感方面的研究较少。Xu 等^[15]通过高光谱数据应用多种深度学习模型估算 SOM 含量,结果表明 CNN 能够从复杂的光谱数据中有效提取特征进行学习,相较于 BP 和多层感知器(Multilayer Perceptron, MLP)具有更强的模型表达能力;王璨等^[16]使用 CNN 通过近红外光谱估算土壤含水率,结果表明随着训练样本数量的增加, CNN 建模效果显著优于传统模型。CNN 发展至今已衍生出多种网络结构,为探讨不同 CNN 模型在土壤高光谱遥感领域的建模效果,本研究从光谱预处理、网络结构以及深度 3 个方面探讨采用 CNN 方法估算 SOM 的可靠性,并挖掘 CNN 模型所提取的特征波段,以期对红壤地区基于高光谱数据估算 SOM 含量提供参考。

1 材料与方法

1.1 研究区概况

研究区位于江西省奉新县北部(115°03'~115°23'E, 28°40'~28°47'N),总面积约 20 000 hm²,属中亚热带湿润气候,年平均降雨量 1 612 mm,年平均气温 17.3℃,海拔介于 31~133 m 之间,地势较平坦,处于低丘平原地带。研究区土地利用类型主要包括耕地、园地和林地,土壤类型主要为红壤。

1.2 土壤样本采集与光谱数据处理

样本的采集时间为 2018 年 7 月 23 日—8 月 11 日,将研究区划分 1 km × 1 km 的网格,综合考虑网格内的地理环境后共采集了 248 个样本,如图 1 所示。样本在实验室风干、研磨和过 2 mm 筛后,均匀分为两部分,一部分通过重铬酸钾容量法测定土壤有机质含量,另一部分采用美国 ASD 公司的 FieldSpec4 地物光谱仪获取光谱数据,波谱范围为 350~2 500 nm,在 350~1 000 nm 范围内采样间隔为 1.4 nm,在 >1 000~2 500 nm 范围内采用间隔为 2 nm,重采样间隔为 1 nm。将信噪比低的边缘波段 350~399 和 2 451~2 500 nm 去除。

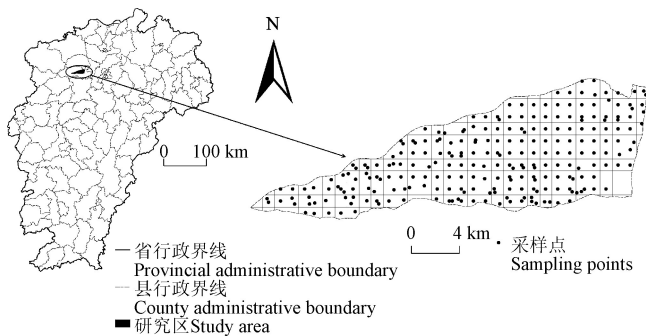


图 1 研究区采样点分布

Fig.1 Sampling points distribution in the study area

为了降低数据维数和减少冗余度,对光谱数据每 10 nm 间隔取平均值进行重采样,每个样本得到 205 个光

谱反射率波段,共计 50 840 个数据的原始光谱反射率(R_f)用于建模。同时,在此基础上进行一定的光谱预处理作为对比:首先采用 Daubechies6 小波进行三层分解,通过软阈值去除光谱中的高频噪声^[17];然后对光谱反射率进行微分变换^[18],包括 0.5 阶微分(0.5 Order Derivative Reflectance, 0.5DR)、1 阶微分(1DR)、1.5 阶微分(1.5DR)、2 阶微分(2DR)、对数的微分(Log-Derivative Reflectance, LDR)。

1.3 基于深度学习方法的 SOM 模型

1.3.1 CNN 模型

CNN 建模过程如下:首先对 SOM 值进行归一化处理,输入层可以看成是行、列为(1, 205)的二维光谱信息矩阵,以适应 CNN 模型的学习结构;卷积层通过一定大小的多个卷积核按照一定的步长对输入数据进行特征提取;池化层也称下采样层,以一定大小采样范围的最大值或者平均值替代原范围的值,以达到减少数据处理量并且保留重要特征信息的目的;全连接层位于卷积层和池化层之后,其作用则是对提取的特征进行非线性组合以得到输出结果,其超参数主要为神经元的数量;输出层为 0~1 范围内的值,最后通过反归一化得到 SOM 估测值。数据集每迭代训练和验证完 1 次为 1 个迭代周期(Epoch),整个学习过程通过不断的 Epoch 循环更新权重参数值以降低损失函数值,从而自主提取光谱特征。在模型构建过程中,激活函数一般位于卷积层和全连接层后面,其作用是使用非线性的激活函数提高模型的表达能力;优化器作用则是计算和更新模型参数,使其逼近或达到最优值,从而最小化损失函数;随机失活法和早停法目的都是为了防止模型过拟合的现象,使用随机失活方法在每个训练批次中按一定比例屏蔽掉一些神经元,早停法则是在训练过程中损失函数在一定忍耐度内没有得到改善时模型提前停止训练。

由于本研究主要探讨不同 CNN 结构和深度情况下的建模效果,在保证网络结构和大部分超参数与对应的 CNN 模型相同的情况下,为了有可比性,所有模型固定某些模型函数和超参数的值,具体设置如下:输出层激活函数为 Sigmoid,其他层的激活函数为 Tanh,优化器为 Nadam, Loss 为均方根误差(Root Mean Squared Error, RMSE),学习率为 0.000 1,忍耐度为 300 Epoch,在最后的 3 个全连接层中神经元数量分别为 200、100 和 1,并且在全连接层中引入 Dropout 随机失活 30%神经元。

本研究构建了 5 种不同的 CNN 模型,如表 1 所示。为了增加各 CNN 模型的可比性,对部分模型的原始结构和参数进行了调整以相对统一,其中 LeNet-5 与文献[19]中提到的 LeNet-5 结构不同在于第 2 个卷积层中卷积核数量由 16 变为 12,全连接层数量由 120、84、10 变为 200、100、1,并且加入了随机失活率为 30%的 Dropout 方法。AlexNet-8^[20]特点是深度比 LeNet-5 更深、使用更大的卷积核和步长并且池化范围也增大。VGGNet^[21]首先采用 7 层的 VGGNet-7 网络结构,由 4 个卷积层、2 个池化层和 3 个全连接层构成;之后通过增减卷积层和池化层分别进

行了 VGGNet 5、7、10、13 和 16 层网络结构的建模，以探讨不同深度下的建模效果。GoogLeNet^[22]的 Inception 结构将上层的特征通过不同大小的卷积核，以 4 种方式提取特征，并且采用 1×1 卷积核提高计算效率，

GoogLeNet-7 包含 1 个 Inception 结构，其卷积核数量为 6。ResNet^[23]通过跨层连接构建残差结构，缓解了深度 CNN 中容易出现梯度消失的问题，ResNet-13 包含 3 个残差结构，每个残差结构由 3 个卷积层构成。

表 1 不同卷积神经网络模型结构
Table 1 Different Convolutional Neural Network (CNN) model structures

CNN 模型结构 CNN model structure	SOM 值归一化 Normalized SOM	输入层 Input layer	卷积层 Convolutional layer	池化层 Pooling layer	全连接层 Fully-connected layer	输出层 Output layer	反归一化 Inverse normalization
LeNet-5	-	-	Conv1(1,5)-1-6 Conv2(1,5)-1-12	Maxpool1(1,2)-2 Maxpool2(1,2)-2	FC3-200 (Dropout-0.3) FC4-100 (Dropout-0.3)	FC5-1	-
AlexNet-8	-	-	Conv1(1,11)-4-6 Conv2(1,5)-1-12 Conv3(1,3)-1-24 Conv4(1,3)-1-24 Conv5(1,3)-1-24	Maxpool1(1,3)-2 Maxpool2(1,3)-2 Maxpool5(1,3)-2	FC6-200 (Dropout-0.3) FC7-100 (Dropout-0.3)	FC8-1	-
VGGNet-7	Normalization (SOM)	(1,205)	Conv1(1,3)-1-6 Conv2(1,3)-1-6 Conv3(1,3)-1-12 Conv4(1,3)-1-12	Maxpool2(1,2)-2 Maxpool4(1,2)-2	FC5-200 (Dropout-0.3) FC6-100 (Dropout-0.3)	FC7-1	SOM
GoogLeNet-7	-	-	Conv1(1,7)-2-6 Conv2(1,3)-1-12 Inception-6	Maxpool1(1,3)-2 Maxpool2(1,3)-2	FC5-200 (Dropout-0.3) FC6-100 (Dropout-0.3)	FC7-1	-
ResNet-13	-	-	Conv1(1,7)-2-6 Res1-Conv2(1,1)-1-6 Res1-Conv3(1,3)-1-6 Res1-Conv4(1,1)-1-24 Res2-Conv5(1,1)-1-6 Res2-Conv6(1,3)-1-6 Res2-Conv7(1,1)-1-24 Res3-Conv8(1,1)-1-6 Res3-Conv9(1,3)-1-6 Res3-Conv10(1,1)-1-24	Maxpool1(1,3)-2 Maxpool10(1,3)-2	FC11-200 (Dropout-0.3) FC12-100 (Dropout-0.3)	FC13-1	-

注：Normalization(SOM)为归一化有机质值；(1,205)表示 1 行 205 列的光谱矩阵；Res1-Conv2(1,1)-1-6 表示第 1 个残差结构，卷积层位于模型第 2 层，卷积核大小为 (1,1)，步长为 1，卷积核数量为 6；Inception-6 表示 Inception 结构中卷积核数量为 6；Maxpool1(1,2)-2 表示最大池化层位于模型第 1 层，池化范围为 (1,2)，步长为 2；FC3-200 (Dropout-0.3) 表示全连接位于模型第 3 层，神经元数量为 200，30% 的神经元随机失活；FC5-1 表示全连接位于模型第 5 层，神经元数量为 1。其他依次类推。

Note: Normalization(SOM) is the normalized Soil Organic Matter (SOM) value; (1,205) denotes the spectral matrix with 1 row and 205 columns; Res1-Conv2(1,1)-1-6 denotes the 1st residual structure, the convolution layer is located in the 2nd layer of the model, the convolution kernel size is (1,1), the step size is 1, and the number of convolution kernels is 6; Inception-6 denotes inception structure with the number of convolutional kernels of 6; Maxpool1(1,2)-2 indicates that the maximum pooling layer is located at layer 1 of the model with a pooling range of (1,2) and a step size of 2; FC3-200 (Dropout-0.3) indicates that the full connectivity is located at layer 3 of the model with the number of neurons of 200 and 30% of the neurons randomly deactivated; FC5-1 indicates that the full connection is located at layer 5 of the model and the number of neurons is 1. And so on for the others.

1.3.2 其他模型

MLP 是一种构建多隐含层的深度学习模型，通过学习构建的深层非线性网络结构，从数据集中挖掘潜在的特征规律，其网络结构中的隐含层由全连接层构成^[24]。本研究的 MLP-5 通过 BP 反向传播算法进行训练，包含 5 个全连接层构成，神经元数量分别为 200、200、100、100、1，并且每个全连接层中间都加入了 30%Dropout 机制以防止过拟合，其他超参数与 LeNet-5 模型一致。

随机森林 (Random Forest, RF) 是一种集成学习方法，通过构建多个树模型，对单个模型结果通过平均或者多数表决原则来决定最终结果^[25]。经过多次调试，本研究的 RF 使用了 10 个基评估器。

支持向量机 (Support Vector Machine, SVM) 基于统计学习理论，通过核函数将数据映射到高维特征空间，以找出一个超平面作为决策边界，使估测误差尽可能小^[26]。本研究的 SVM 模型，通过多次调试，核函数为 linear 时建模效果较其他核函数更优。

1.4 高光谱特征提取

可解释学习法 (SHapley Additive exPlanations, SHAP)^[27]是由 Lundberg 在 2017 年所提出，其将博弈论

中的经典 Shapley 值^[28]与局部解释联系起来，构建一个解释模型预测的统一方法，可用于解释机器学习模型的输出，是一种特征重要性的度量。SHAP 值能反映出每个样本中的特征对模型输出的贡献，而且还表现出贡献的正负性。因此，通过计算 CNN 建模过程中各波长的 SHAP 值，可以得到各波长的贡献程度，从而提取高光谱特征。假设第 i 个样本为 x_i ，第 i 个样本的第 j 个特征为 x_{ij} ，模型对该样本的预测值为 y_i ， ϕ_0 为整个模型的基值（通常是所有样本的目标变量的均值），那么第 i 个样本第 j 个特征的 SHAP 值 $f(x_{ij})$ 服从如下等式：

$$y_i = \phi_0 + f(x_{i1}) + f(x_{i2}) + \dots + f(x_{ij}) \quad (1)$$

1.5 精度评价

模型的精度通过决定系数 (R^2)、RMSE 以及相对分析误差 (Relative Percent Deviation, RPD) 3 个指标进行评价， R^2 和 RPD 越大以及 RMSE 越小说明模型的估测效果和稳定性越好，其中，当 RPD 为 2.5~3 时，表明模型估测能力很好；当 RPD>3 时，表明模型估测能力极好^[29]。

本研究在 OriginPro 9.1、ArcGIS 10.2 和 Microsoft

Excel 2010 软件中完成数据统计分析, CNN 和 MLP 模型在 Spyder 软件中通过深度学习 keras 库使用 Python3.7 语言编写脚本搭建, RF 和 SVM 模型调用 Sklearn 接口中相应的机器学习模块实现, SHAP 值通过调用 shap 接口中的 DeepExplainer 模块得到。

2 结果与分析

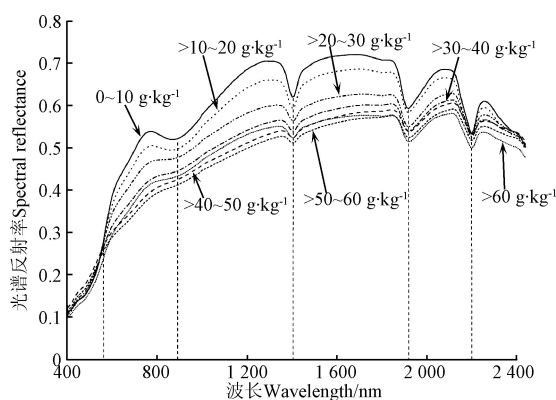
2.1 红壤有机质统计特征和光谱特性

将 248 个土壤样本依据 K-S (Kennard-Stone) 算法^[30]得到样本间的欧式距离, 按照 3:1 的比例分为 2 个部分, 共得到 186 个训练样本集, 62 个验证样本集, 如表 2 所示, 土壤样本的 SOM 含量介于 5.27~64.00 g/kg 之间, 3 个样本类型的平均 SOM 含量分别为 33.97、35.04、30.78 g/kg, 变异系数均为中等程度变异。

表 2 土壤有机质含量统计特征
Table 2 Statistical characteristics of SOM content

样本类型 Sample type	样本量 Sample size	最小值 Min/ (g·kg ⁻¹)	最大值 Max/ (g·kg ⁻¹)	平均值 Mean/ (g·kg ⁻¹)	标准差 Standard deviation/ (g·kg ⁻¹)	变异系数 Coefficient of variation/%
总体样本 Whole set	248	5.27	64.00	33.97	14.70	43.27
训练样本 Training set	186	5.56	64.00	35.04	14.39	41.07
验证样本 Validation set	62	5.27	61.60	30.78	15.28	49.64

以 10 g/kg 为梯度, 将土壤样本按照 SOM 含量进行分组, 测定各组土壤光谱曲线, 取各组平均值, 获得不同 SOM 含量的土壤光谱反射率特征, 如图 2 所示。可以看出各光谱曲线形状和走势相似, 总体上, 光谱反射率随 SOM 含量增大而降低, 但在 550 nm 附近存在交叉现象。SOM 含量>60 g/kg 土样的反射率, 在 600~1 400 nm 波段内高于 SOM 含量为>40~60 g/kg 的土样, 在 1 400~1 900 nm 波段内高于 SOM 含量为>50~60 g/kg 的土样。在 900 nm 左右受氧化铁的影响存在吸收谷, 在 1 400、1 900 和 2 200 nm 附近存在水分吸收谷。



注: 图中数值如 0~10 g·kg⁻¹ 等为有机质含量。

Note: Data in figure such as 0-10 g·kg⁻¹ is SOM content.

图 2 不同 SOM 含量土壤样本的光谱曲线

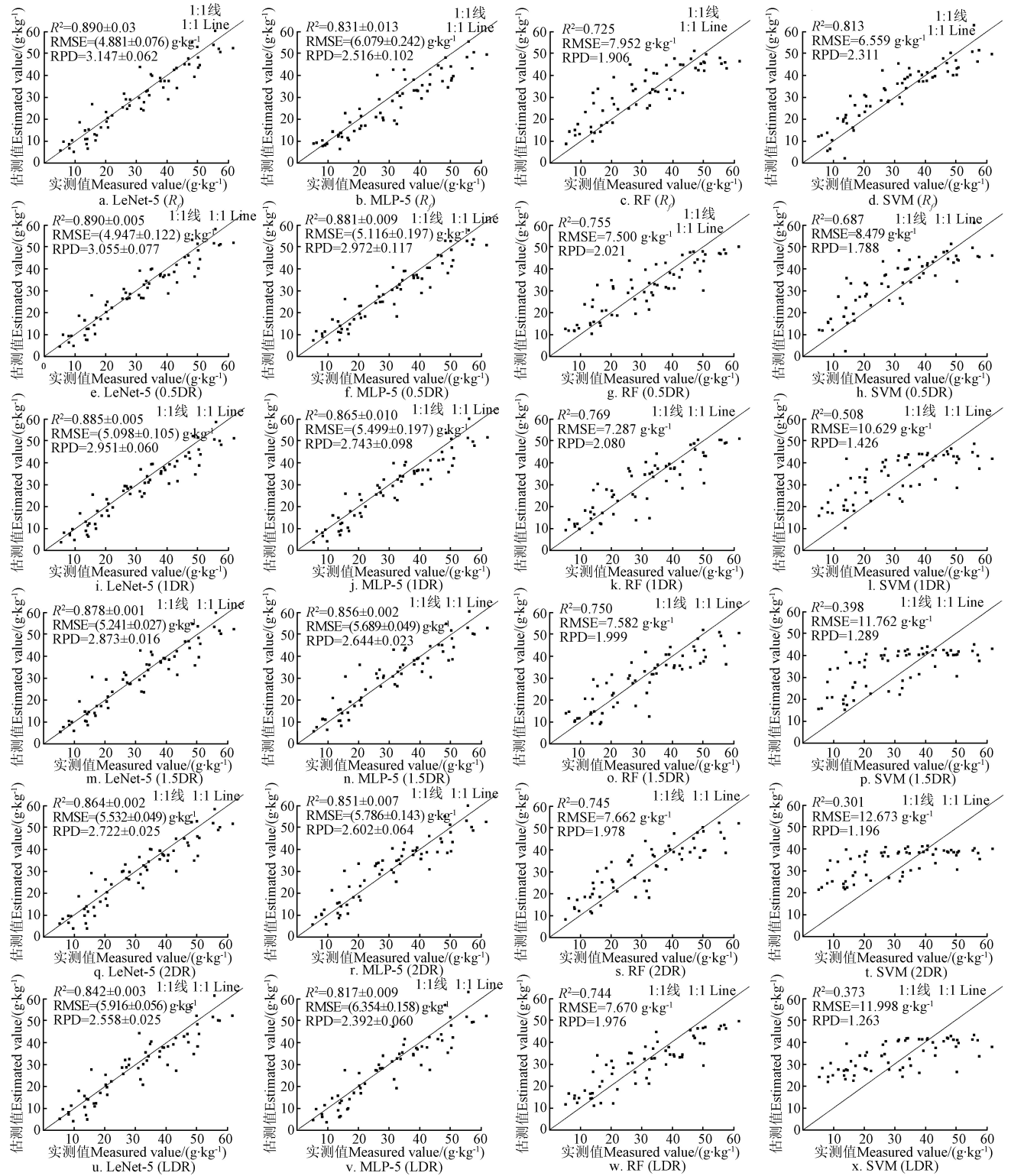
Fig.2 Spectral curves of soil samples with different SOM contents

2.2 光谱预处理对不同 SOM 模型估算精度的影响

以验证集为例, 比较 6 种光谱预处理对不同模型 (LeNet-5、MLP-5、RF、SVM) 估算土壤 SOM 含量的影响, 结果如图 3 所示。图中值为最后 1 个 Epoch 的估测值, 从不同光谱预处理来看, 光谱微分变换仅在 MLP-5 和 RF 模型中精度有所提高, 从 R^2 、RMSE 和 RPD 可以看出, 对于原始光谱、0.5DR 和 1DR 附近是较好的预处理选择。从不同模型来看, LeNet-5 在不同预处理下的表现都要优于其他模型, 且在原始光谱作为模型输入时建模效果最好; LeNet-5 和 MLP-5 估测与实测值较为靠近 1:1 线, 但在 SOM 含量>40 g/kg 时, 存在一定的估测值小于实测值; LeNet-5 在所有预处理中, 相较于 MLP-5, 将网络结构中的全连接层部分替换为卷积和池化层不仅能够提高模型精度, 同时精度波动也较小。RF 与 SVM 的估测与实测值的散点图较为离散, 在 SOM<40 g/kg 时, 估测值普遍大于实测值, SOM>40 g/kg 时, 估测值普遍小于实测值。

2.3 CNN 结构对 SOM 模型估算的影响

为探索不同 CNN 结构在原始光谱反射率作为模型输入时的建模效果, 得到各 CNN 模型训练和验证精度曲线和精度对比分别如图 4 和表 3 所示。从整体来看, 5 种模型的精度曲线非常相似, 学习过程大致可分为迅速学习、缓慢学习、趋于稳定 3 个阶段; 在精度对比中 5 种模型在训练集和验证集的 $R^2>0.85$ 以上, 训练集的 RMSE<5 g/kg, RPD>3, 验证集 RMSE<6 g/kg, RPD>2.5, 表明各模型具有很好或者极好的估测能力。从各模型来看, LeNet-5 在训练集和验证集上表现都较为稳定, R^2 稳定在 0.89 左右, RMSE<5 g/kg, RPD>3, 具有极好的估测能力。AlexNet-8 和 ResNet-13 都是在训练集上的建模效果好, R^2 超过 0.9, RMSE 低于 4 g/kg, 但在验证集上的精度偏低, 从精度曲线中也可以看到各评价指标在迭代后期训练集曲线和验证集曲线之间拉开较大并且在验证集上的震荡较大, 产生了过拟合现象并且模型不稳定, 分析其原因是这 2 种模型在第 1 个卷积层的卷积核大小和步长太大, 并且池化层的池化范围相较 LeNet-5 和 VGGNet-7 更大, 虽然增大了提取特征的范围, 但是减少了提取到的特征数量, 导致模型的泛化能力减弱^[21]。另外, ResNet-13 仅迭代了 2 000 Epoch 就训练完成了, 而其他模型需要训练 4 000 或者 5 000 Epoch 左右, 原因是 ResNet-13 的残差学习结构, 解决了梯度消失的问题, 加快了学习的过程^[23]。VGGNet-7 在 5 种模型中训练集和验证集的 R^2 指标差距以及 RMSE 指标差距最小, 并且 RPD 最高, 达到 3.291 ± 0.053 , 建模效果最好, 其原因是采用了较小卷积核、步长和池化范围, 能够提取到更多的特征数量, 提高学习能力^[21]。相比其他模型, GoogLeNet-7 虽然在训练集上的建模效果最差, 但在验证集上的精度却比 AlexNet-8 和 ResNet-13 高, RPD 也趋近于 3, 可能由于 Inception 结构中多种大小不同的卷积核, 提取到不同范围的特征, 增强了模型的泛化能力^[22]。



注：MLP-5 为 5 层感知器模型；RF 为随机森林模型；SVM 为支持向量机模型； R_f 为原始光谱反射率；0.5DR~2DR 分别为 0.5 阶、1 阶、1.5 阶、2 阶微分；LDR 为对数的微分；RPD 为相对分析误差；LeNet-5 和 MLP-5 模型中 R^2 、RMSE 和 RPD 值均用最后 300 个迭代周期的均值和标准差表示。下同。
 Note: MLP-5 is a five layers perceptron model; RF is a Random Forest Model; SVM is Support Vector Machine Model; R_f is original spectral reflectance; 0.5DR is 0.5 order derivative reflectance; 1DR is first order derivative reflectance; 1.5DR is 1.5 order derivative reflectance; 2DR is second order derivative reflectance; LDR is log-derivative reflectance; RPD is Relative Percent Deviation; The R^2 , RMSE and RPD for LeNet-5 and MLP-5 models are represented by the mean and standard deviation of the last 300 epochs. Same below.

图 3 验证集不同预处理下各模型 SOM 含量实测值与估测值比较

Fig.3 Comparisons of measured and estimated SOM contents for validation set based on each model under different pretreatments

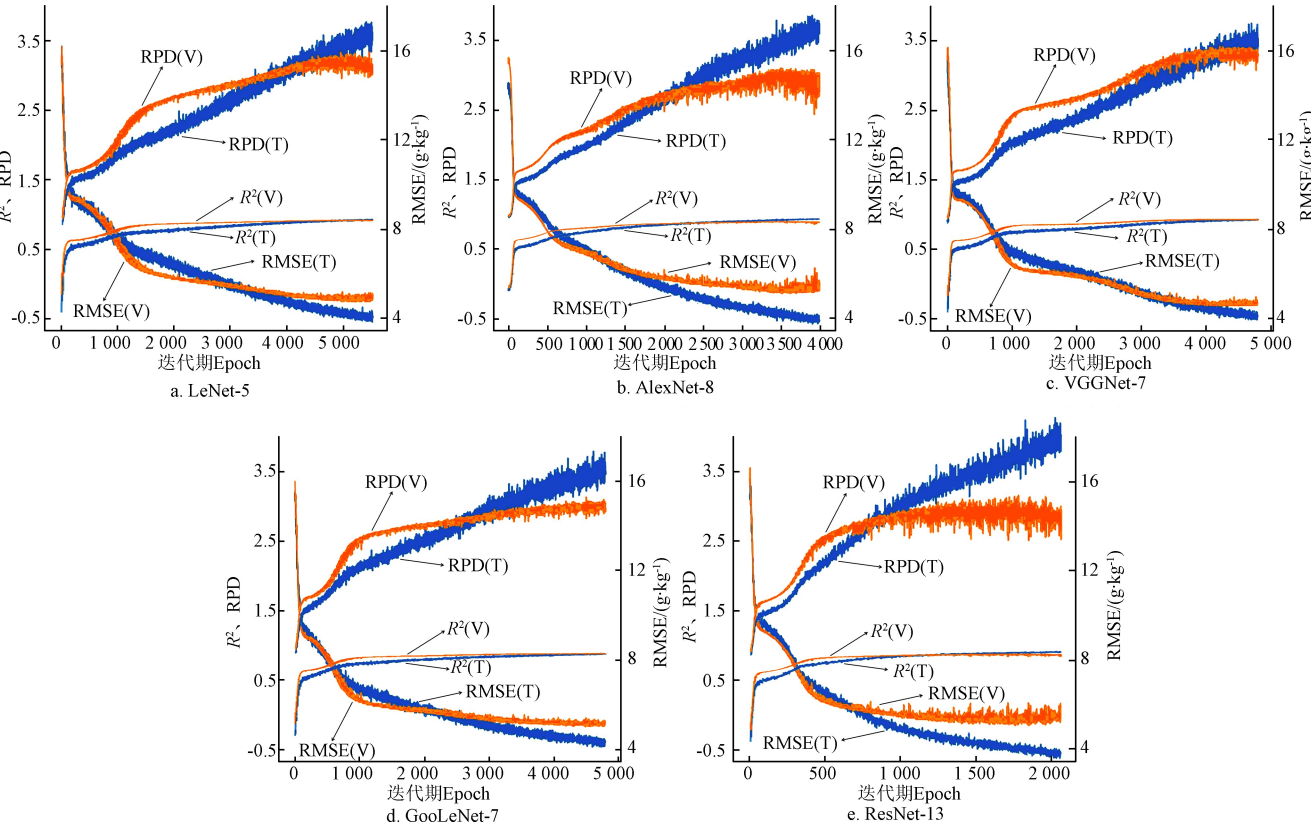


图4 训练集 (T) 和验证集 (V) 基于不同 CNN 模型的 SOM 估算精度

Fig.4 Accuracy of SOM estimation by different CNN models for training set (T) and validation set (V)

表 3 基于不同 CNN 模型的 SOM 估算精度对比

Table 3 Comparison of SOM estimation accuracy based on different CNN models

模型 Models	训练集 Training set			验证集 Validation set		
	R^2	RMSE/(g·kg ⁻¹)	RPD	R^2	RMSE/(g·kg ⁻¹)	RPD
LeNet-5	0.896±0.005	4.060±0.092	3.544±0.091	0.890±0.003	4.881±0.076	3.147±0.062
AlexNet-8	0.906±0.005	3.981±0.102	3.605±0.101	0.868±0.006	5.299±0.145	2.939±0.100
VGGNet-7	0.895±0.005	4.145±0.097	3.447±0.094	0.901±0.002	4.647±0.057	3.291±0.053
GoogLeNet-7	0.877±0.006	4.290±0.097	3.461±0.098	0.882±0.002	5.090±0.061	2.989±0.045
ResNet-13	0.902±0.006	3.857±0.119	3.836±0.141	0.869±0.008	5.337±0.183	2.876±0.121

2.4 CNN 建模结果比较

为进一步探索在原始光谱反射率作为模型输入时不同深度 VGGNet 模型建模效果,其训练和验证精度对比如表 4 所示,从中可以发现,模型在训练集上由 VGGNet-13 取得 R^2 最高值(0.914)和 RMSE 最低值(3.719 g/kg),但在验证集上却是 VGGNet-7 模型取得最高 R^2 值(0.901)、最高 RPD 值(3.291)和最低 RMSE 值(4.647 g/kg)。随着网络深度的进一步(7~16)增加,

模型在训练集和验证集的 R^2 指标差距以及 RMSE 指标差距逐渐变大,说明模型随着深度增加容易出现过拟合现象;从 R^2 、RMSE 和 RPD 指标在训练集的标准差逐渐增大,以及在验证集也是深层模型标准差更大,可以看出深度的增加使得模型的稳定性也逐渐降低。综合 3 个指标来看,VGGNet-7 建模效果最好,其在验证集上 SOM 含量实测值与估测值比较如图 5 所示,从图中可以看出大部分样点的估测值都接近 1:1 线。

表 4 基于不同 VGGNet 模型的 SOM 估算精度对比

Table 4 Comparison of SOM estimation accuracy based on different VGGNet models

模型 Models	训练集 Training set			验证集 Validation set		
	R^2	RMSE/(g·kg ⁻¹)	RPD	R^2	RMSE/(g·kg ⁻¹)	RPD
VGGNet-5	0.859±0.008	4.738±0.109	3.038±0.079	0.887±0.002	4.895±0.044	3.199±0.046
VGGNet-7	0.895±0.005	4.145±0.097	3.447±0.094	0.901±0.002	4.647±0.057	3.291±0.053
VGGNet-10	0.901±0.006	3.895±0.105	3.734±0.114	0.889±0.009	4.965±0.207	3.044±0.131
VGGNet-13	0.914±0.008	3.719±0.149	3.945±0.165	0.881±0.009	5.181±0.192	2.909±0.101
VGGNet-16	0.913±0.014	3.734±0.250	3.982±0.276	0.875±0.012	5.240±0.255	2.915±0.145

注: VGGNet-5 为 5 层网络结构的 VGGNet 模型,其他模型以此类推。
Note: VGGNet-5 is a VGGNet model with five layers of net structure, and so on.

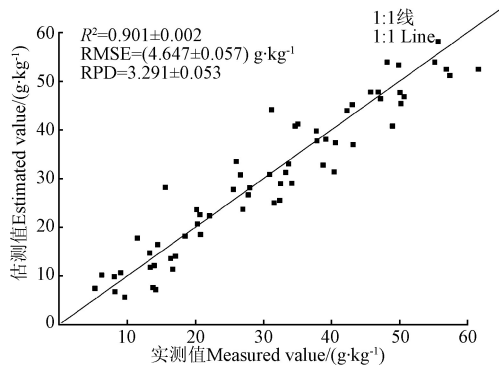
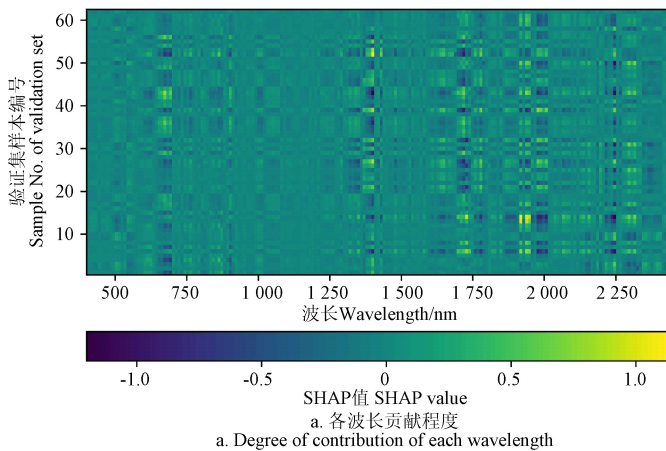
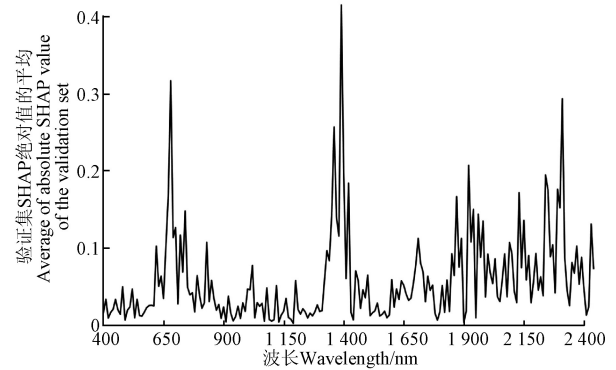


图 5 基于 VGGNet-7 模型的验证集 SOM 含量实测值与估测值比较

Fig.5 Comparisons of measured and estimated SOM contents based on VGGNet-7 model for validation set



a. Degree of contribution of each wavelength



b. Average degree of contribution of each wavelength

注：SHAP 为样本的各波长对模型输出的贡献。

Note: SHAP is the contribution of each wavelength of the sample to model output.

图 6 VGGNet-7 模型验证集各波长贡献程度分布

Fig.6 Distribution of contribution degree of each wavelength for validation set for VGGNet-7 model

3 讨论

本研究首先比较了 CNN、MLP、RF 和 SVM 在不同光谱预处理下的建模效果，发现光谱微分变换能够在部分模型中提高建模精度，且 0.5DR 和 1DR 附近是较好的光谱微分变换选择，这与李雪萍等^[31]得到的光谱一阶微分处理更加突出 SOM 隐藏的信息，光谱二阶微分处理反而起到抑制作用结论相似。但 LeNet-5 在不同预处理下的表现都要优于其他模型，且在原始光谱作为模型输入时建模效果最好，说明 CNN 能够充分挖掘出原始光谱数据中所包含的丰富且复杂的特征信息。在此基础上进一步探讨了 CNN 模型在何种网络结构以及何种网络深度下的建模效果更好，结果表明，各 CNN 模型 RPD 都大于 2.5，都具有很好或者极好的预测能力，这得益于 CNN 强大的特征学习能力，使其在处理高维度以及大数据集时具有特殊的优势^[32]。本研究的样本数量虽然不多，但每个样本的光谱波段较多，数据量相对较大，在这种背景下，将 CNN 应用于土壤高光谱建模是具备可行性的，试验结果数据证明结构简单并且深度较浅的 VGGNet-7 效果最好，未出现过拟合现象，体现了奥卡姆剃刀“简单有效”的原理，这在推广到实际应用的过程中是有益的。但 CNN

2.5 CNN 模型特征波段提取

VGGNet-7 模型验证集各波长重要性分布如图 6 所示，其中 SHAP 值大于 0 说明该波长提升了 SOM 的估测值，有正向作用，SHAP 值小于 0 说明该波长使得估测值降低，有负向作用，SHAP 的绝对值越大说明该波长对 SOM 估测的贡献程度越大。从图 6a 中可以发现，不同的波长对于样本的贡献程度不同，同一波长对于不同的样本既可以起正向作用，也可能有负向作用；图 6b 反映了各波长的平均贡献程度，其中排名前五的波长为 1 390、680、2 310、1 360、1 920 nm，可以发现这些波峰附近的波长对 SOM 估测贡献程度较高，是 VGGNet-7 模型在利用高光谱数据估测 SOM 时所提取的重要特征波长。

模型也存在解释性差和一定的不稳定性等缺点，相信在样本数量增加时，模型的稳定性会进一步提高。而 CNN 模型的“黑箱”探索是目前学者们正在研究的热点，本研究通过获取解释模型预测的 SHAP 值，得到 VGGNet-7 模型各波长对于 SOM 估测的贡献程度，发现 1 390、680、2 310、1 360、1 920 nm 及其附近是 VGGNet-7 所提取的重要特征波长。680 nm 及其附近与 Galvao 等^[33]的 550~700 nm 和纪文君等^[34]的 600~800 nm 是 SOM 的敏感波段结果一致，其他几个特征波长可能是由于高岭石、脂肪族化合物、酚类物质、甲基、碳水化合物等因素 O-H、C-H、Al-OH、C=O 基团的基频振动以及合频和倍频振动吸收的影响^[35]。

在图 3 估测与实测值的散点图中发现，当 SOM 含量大于 40 g/kg 附近时，各模型估测值都出现了偏小的现象，在 SOM 含量小于 40 g/kg 附近时，RF 和 SVM 估测值普遍偏大。分析其原因，图 2 中虽然光谱反射率随 SOM 含量增大而降低，但由于交叉现象以及 SOM 含量大于 60 g/kg 的样本在部分波段范围反射率高于 SOM 含量 40~50 g/kg 和 50~60 g/kg 的样本，使得模型在训练过程中当 SOM 含量在 40 g/kg 附近时特征学习较为混乱，容易降低高实测值的估算，提高低实测值的估算。但从图 5 中

VGGNet-7 的估测与实测值的散点图可以看出, 其散点分布相对较为均匀, 没有明显的估测偏低或者偏高的现象, 表明 VGGNet-7 具有更好的特征学习和区分能力。

土壤光谱是土壤内在理化特性的综合反映, 环境因素会造成成土条件的差异, 从而导致光谱响应波段的不同^[36]。本研究在未考虑环境协变量的情况下, 通过构建一个综合环境因素的模型, 仅利用光谱数据估算 SOM 含量, 以挖掘 CNN 模型的建模能力。结果证明模型精度较为理想, 这在实际应用过程中可以在不考虑环境条件下, 直接通过光谱数据进行土壤属性的估算, 从而减少工作量。但由于本文的研究区较小, 环境变量相对较为统一, 以上结论的适用性还有待考证, 今后可以对不同区域、不同土壤类型、不同土壤属性进行高光谱数据的 CNN 建模验证。

4 结 论

本研究以 248 个红壤样本的可见近红外光谱数据作为输入数据, 土壤有机质 (Soil Organic Matter, SOM) 含量为输出数据, 在小样本数据集下, 探讨了 6 种光谱预处理、5 种不同卷积神经网络 (Convolutional Neural Network, CNN) 结构和 5 种不同网络深度情况下 CNN 的建模效果, 得到结论如下:

1) 得益于 CNN 模型强大的特征学习能力, 在仅使用原始光谱数据的情况下, 各 CNN 模型在验证集上的相对分析误差大于 2.5, 具有很好的模型估测能力, 是一种较好的利用高光谱数据估算 SOM 含量的方法。

2) 在不同网络结构的对比中, GoogLeNet-7 和 ResNet-13 虽然都加入了特殊的结构, 但拥有超参数较小的卷积核、步长和池化范围的 LeNet-5 和 VGGNet-7 网络结构建模效果更好, 因此, CNN 模型中一些超参数的设置可能比网络结构本身更关键; 在不同深度的对比中, 随着网络深度的增加, 模型容易产生过拟合现象且不稳定, 浅层的 CNN 结构比深层建模效果更好。

3) VGGNet-7 网络结构在所有模型中建模效果最好, 在训练集上决定系数为 0.895, 均方根误差为 4.145 g/kg, 相对分析误差为 3.447, 在验证集上决定系数为 0.901, 均方根误差为 4.647 g/kg, 相对分析误差为 3.291, 具有极好的模型估测能力。

4) 680、1 360、1 390、1 920、2 310 nm 及其附近是 VGGNet-7 所提取的 SOM 重要特征波长, 为便携式土壤有机质监测设备的研发和应用提供参考。

参 考 文 献

- [1] Gong W, Yan X, Wang J, et al. Long-term manure and fertilizer effects on soil organic matter fractions and microbes under a wheat-maize cropping system in northern China[J]. *Geoderma*, 2009, 149(3): 318-324.
- [2] Tian Y, Zhang J, Yao X, et al. Laboratory assessment of three quantitative methods for estimating the organic matter content of soils in China based on visible/near-infrared reflectance spectra[J]. *Geoderma*, 2013, 202: 161-170.
- [3] 童庆禧, 张兵, 张立福. 中国高光谱遥感的前沿进展[J]. 遥

感学报, 2016, 20(5): 689-707.

Tong Qingxi, Zhang Bing, Zhang Lifu. Current progress of hyperspectral remote sensing in China[J]. *Journal of Remote Sensing*, 2016, 20(5): 689-707. (in Chinese with English abstract)

- [4] 于雷, 洪永胜, 周勇, 等. 高光谱估算土壤有机质含量的波长变量筛选方法[J]. *农业工程学报*, 2016, 32(13): 95-102.
Yu Lei, Hong Yongsheng, Zhou Yong, et al. Wavelength variable selection methods for estimation of soil organic matter content using hyperspectral technique[J]. *Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE)*, 2016, 32(13): 95-102. (in Chinese with English abstract)
- [5] 陈仲新, 任建强, 唐华俊, 等. 农业遥感研究应用进展与展望[J]. *遥感学报*, 2016, 20(5): 748-767.
Chen Zhongxin, Ren Jianqiang, Tang Huajun, et al. Progress and perspectives on agricultural remote sensing research and applications in China[J]. *Journal of Remote Sensing*, 2016, 20(5): 748-767. (in Chinese with English abstract)
- [6] 史舟, 梁宗正, 杨媛媛, 等. 农业遥感研究现状与展望[J]. *农业机械学报*, 2015, 46(2): 247-260.
Shi Zhou, Liang Zongzheng, Yang Yuanyuan, et al. status and prospect of agricultural remote sensing[J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2015, 46(2): 247-260. (in Chinese with English abstract)
- [7] Shi Z, Ji W, Rossel R A V, et al. Prediction of soil organic matter using a spatially constrained local partial least squares regression and the Chinese vis-NIR spectral library[J]. *European Journal of Soil Science*, 2015, 66(4): 679-687.
- [8] 章涛, 于雷, 易军, 等. 高光谱小波能量特征估测土壤有机质含量[J]. *光谱学与光谱分析*, 2019, 39(10): 3217-3222.
Zhang Tao, Yu Lei, Yi Jun, et al. Determination of soil organic matter content based on hyperspectral wavelet energy features[J]. *Spectroscopy and Spectral Analysis*, 2019, 39(10): 3217-3222. (in Chinese with English abstract)
- [9] 张东辉, 赵英俊, 赵宁博, 等. 航空高光谱预测黑土地 SOM 含量的神经网络方法[J]. *光谱学与光谱分析*, 2018, 38(增刊 1): 328-329.
Zhang Donghui, Zhao Yingjun, Zhao Ningbo, et al. Neural network method for prediction SOM content in black soil using aerial hyperspectral data[J]. *Spectroscopy and Spectral Analysis*, 2018, 38(Supp.1): 328-329. (in Chinese with English abstract)
- [10] 章海亮, 罗微, 刘雪梅, 等. 应用遗传算法结合连续投影算法近红外光谱检测土壤有机质研究[J]. *光谱学与光谱分析*, 2017, 37(2): 584-587.
Zhang Hailiang, Luo Wei, Liu Xuemei, et al. Measurement of soil organic matter with near infrared spectroscopy combined with genetic algorithm and successive projection algorithm[J]. *Spectroscopy and Spectral Analysis*, 2017, 37(2): 584-587. (in Chinese with English abstract)
- [11] 谢文, 赵小敏, 郭熙, 等. 基于 RBF 组合模型的山地红壤有机质含量光谱估测[J]. *林业科学*, 2018, 54(6): 16-23.
Xie Wen, Zhao Xiaomin, Guo Xi, et al. Spectrum based estimation of the content of soil organic matters in mountain

- red soil using RBF combination model[J]. *Scientia Silvae Sinicae*, 2018, 54(6): 16-23. (in Chinese with English abstract)
- [12] 国佳欣, 赵小敏, 郭熙, 等. 基于 PLSR-BP 复合模型的红壤有机质含量反演研究[J]. *土壤学报*, 2020, 57(3): 636-645.
- Guo Jiaxin, Zhao Xiaomin, Guo Xi, et al. Inversion of organic matter content in red soil based on PLSR-BP composite model[J]. *Acta Pedologica Sinica*, 2020, 57(3): 636-645. (in Chinese with English abstract)
- [13] 史舟. 土壤地面高光谱遥感原理与方法[M]. 北京: 科学出版社, 2014.
- [14] Lecun Y, Bengio Y, Hinton G. Deep learning[J]. *Nature*, 2015, 521(7553): 436-444.
- [15] Xu Z, Zhao X, Guo X, et al. Deep learning application for predicting soil organic matter content by VIS-NIR spectroscopy[J]. *Computational Intelligence and Neuroscience*, 2019, 2019(6): 1-11.
- [16] 王璨, 武新慧, 李恋卿, 等. 卷积神经网络用于近红外光谱预测土壤含水率[J]. *光谱学与光谱分析*, 2018, 38(1): 42-47.
- Wang Can, Wu Xinhui, Li Lianqing, et al. Convolutional neural network for application in prediction of soil moisture content[J]. *Spectroscopy and Spectral Analysis*, 2018, 38(1): 42-47. (in Chinese with English abstract)
- [17] Hu Y, Jiang T, Shen A, et al. A background elimination method based on wavelet transform for Raman spectra[J]. *Chemometrics and Intelligent Laboratory Systems*, 2007, 85(1): 94-101.
- [18] 王敬哲, 塔西甫拉提·特依拜, 丁建丽, 等. 基于分数阶微分预处理高光谱数据的荒漠土壤有机碳含量估算[J]. *农业工程学报*, 2016, 32(21): 161-169.
- Wang Jingzhe, Tashpolat-Tiyip, Ding Jianli, et al. Estimation of desert soil organic carbon content based on hyperspectral data preprocessing with fractional differential[J]. *Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE)*, 2016, 32(21): 161-169. (in Chinese with English abstract)
- [19] Lecun Y, Bottou L, Bengio Y, et al. Gradient-based learning applied to document recognition[J]. *Proceedings of the IEEE*, 1998, 86(11): 2278-2324.
- [20] Krizhevsky A, Sutskever I, Hinton G E. ImageNet classification with deep convolutional neural networks[J]. *Advances in Neural Information Processing Systems*, 2012, 25(2): 1-9.
- [21] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition[EB/OL]. 2014-09-04 [2015-04-10]. <https://arxiv.org/pdf/1409.1556.pdf>
- [22] Szegedy C, Liu W, Jia Y, et al. Going deeper with convolutions[C]. Boston, MA, USA: 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2015: 1-9.
- [23] He K, Zhang X, Ren S, et al. Deep residual learning for image recognition[C]. Las Vegas, NV, USA: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016: 770-778.
- [24] 孙志军, 薛磊, 许阳明, 等. 深度学习研究综述[J]. *计算机应用研究*, 2012, 29(8): 2806-2810.
- Sun Zhijun, Xue Lei, Xu Yangming, et al. Overview of deep learning[J]. *Application Research of Computers*, 2012, 29(8): 2806-2810. (in Chinese with English abstract)
- [25] Breiman L. Random forests[J]. *Machine Learning*, 2001, 45(1): 5-32.
- [26] 张学工. 关于统计学习理论与支持向量机[J]. *自动化学报*, 2000, 26(1): 36-46.
- Zhang Xuegong. Introduction to statistical learning theory and support vector machines[J]. *Acta Automatica Sinica*, 2000, 26(1): 36-46. (in Chinese with English abstract)
- [27] Lundberg S, Lee S I. A unified approach to interpreting model predictions[C]. Long Beach, CA, USA: NIPS'17 Proceedings of the 31st International Conference on Neural Information Processing Systems, 2017: 4768-4777.
- [28] Lipovetsky S, Conklin M. Analysis of regression in game theory approach[J]. *Applied Stochastic Models in Business & Industry*, 2010, 17(4): 319-330.
- [29] 史舟, 王乾龙, 彭杰, 等. 中国主要土壤高光谱反射特性分类与有机质光谱预测模型[J]. *中国科学: 地球科学*, 2014, 44(5): 978-988.
- Shi Zhou, Wang Qianlong, Peng Jie, et al. Classification of hyperspectral reflectance characteristics of main soils and spectral prediction model of organic matter in China[J]. *Scientia Sinica: Terrae*, 2014, 44(5): 978-988. (in Chinese with English abstract)
- [30] Bilgili A V, Es H M V, Akbas F, et al. Visible-near infrared reflectance spectroscopy for assessment of soil properties in a semi-arid area of Turkey[J]. *Journal of Arid Environments*, 2010, 74(2): 229-238.
- [31] 李雪萍, 张飞, 王小平. 微分算法的艾比湖湿地自然保护区土壤有机质多光谱建模[J]. *光谱学与光谱分析*, 2019, 39(2): 535-542.
- Li Xueping, Zhang Fei, Wang Xiaoping. Study on differential-based multispectral modeling of soil organic matter in Ebinur lake wetland[J]. *Spectroscopy and Spectral Analysis*, 2019, 39(2): 535-542. (in Chinese with English abstract)
- [32] Yuan Q, Shen H, Li T, et al. Deep learning in environmental remote sensing: Achievements and challenges[J]. *Remote Sensing of Environment*, 2020, 241: 111716.
- [33] Galvao L S, Vitorello I. Role of organic matter in obliterating the effects of iron on spectral reflectance and color of Brazilian tropical soils[J]. *International Journal of Remote Sensing*, 1998, 19(10): 1969-1979.
- [34] 纪文君, 史舟, 周清, 等. 几种不同类型土壤的 VIS-NIR 光谱特性及有机质响应波段[J]. *红外与毫米波学报*, 2012, 31(3): 277-282.
- Ji Wenjun, Shi Zhou, Zhou Qing, et al. VIS-NIR reflectance spectroscopy of the organic matter in several types of soils[J]. *Journal of Infrared and Millimeter Waves*, 2012, 31(3): 277-282. (in Chinese with English abstract)
- [35] Viscarra R R A, Behrens T. Using data mining to model and interpret soil diffuse reflectance spectra[J]. *Geoderma*, 2010, 158(1/2): 46-54.
- [36] 周清, 周斌, 张杨珠, 等. 成土母质对水稻土高光谱特性

及其有机质含量光谱参数模型影响的初步研究[J]. 土壤学报, 2014, 41(6): 905-911.
Zhou Qing, Zhou Bin, Zhang Yangzhu, et al. Influence of

parent materials on paddy soil hyperspectral characteristics and SOM spectral-parameter-models[J]. Acta Pedologica Sinica, 2014, 41(6): 905-911. (in Chinese with English abstract)

Hyperspectral estimation of organic matter in red soil using different convolutional neural network models

Zhong Liang, Guo Xi^{*}, Guo Jiaxin, Xu Zhe, Zhu Qing, Ding Meng

(1. College of Land Resources and Environment, Jiangxi Agricultural University, Nanchang 330045, China;

2. Key Laboratory of Poyang Lake Watershed Agricultural Resources and Ecology of Jiangxi Province, Nanchang 330045, China)

Abstract: Deep learning represented by Convolutional Neural Networks (CNN) has been increasing rapidly in recent years, due to its powerful feature learning for computer vision and natural language processing. But there are few studies in the field of hyperspectral remote sensing in soil. Therefore, this study aims to estimate Soil Organic Matter (SOM) using hyperspectral images in small sample dataset, thereby to investigate the modeling effects of different network structures. A total of 248 red soil samples were collected from the northern Fengxin county, Jiangxi province, China. A geospectrometer was used to capture the spectral data. The original spectral data was resampled at 10 nm intervals, after removing the edge bands of 350-399 nm and 2 451-2 500 nm with a low signal-to-noise ratio. A total of 205 original spectral bands and their derivative transformation were obtained as input data, while the SOM content as output data of the model. Firstly, the modeling effects of CNN were compared, such as Multilayer Perceptron (MLP), Random Forest (RF) and Support Vector Machine (SVM) under different spectral pretreatments. Five CNN structures were established, including the earliest LeNet-5, AlexNet-8 with large convolutional core, VGGNet-7 with small convolutional core, GoogLeNet-7 with inception structure, and ResNet-13 with residual learning, particularly on the modeling effects of VGGNet model at five depths. Secondly, all models were evaluated using random deactivation (Dropout) and early stopping to prevent overfitting of the model by three indicators: decision coefficient (R^2), Root Mean Square Error (RMSE) and Relative Analytical Error (RPD). Finally, the black box of CNN model was explained. The results showed that: 1) Due to the strong capability of feature learning in CNN models, the RPD of each CNN model in the validation set was greater than 2.5 in the case of the original spectral data, indicating excellent prediction capability and a better way to predict SOM content using hyperspectral images. 2) In the comparison of different network structures, an optimal model was determined in the network structures of LeNet-5 and VGGNet-7 with small convolutional nuclei, step length, and pooling range of hyper parameters, although the later GoogLeNet-7 and ResNet-13 both incorporated special structures. Therefore, the setting of some hyper parameters in the CNN model can be more critical than the network structure. In different depths, the model was prone to overfitting and unstable, as the network depth increased, where the shallow CNN structure was better than the deep one. 3) An optimal model was achieved in the VGGNet-7 network structure with the excellent model estimation power: R^2 was 0.895 and RMSE was 4.145 g/kg on the training set, while R^2 was 0.901, RMSE was 4.647 g/kg and RPD was 3.291 on the verification set. 4) The wavelengths of 680, 1 360, 1 390, 1 920, 2 310 nm and its vicinities were the important for SOM and they were extracted from the process of VGGNet-7 model establishment. The CNN can be expected for very broad application prospects, due to its simple spectral pre-processing, and feasibility in small samples of soil hyperspectral remote sensing. Therefore, the VGGNet-7 can be applied to the red soil area for rapid and accurate estimation of SOM content using hyperspectral data.

Keywords: soils; models; convolutional neural network; organic matter; hyperspectrum