

Research Article

Deep Learning Application for Predicting Soil Organic Matter Content by VIS-NIR Spectroscopy

Zhe Xu ^{1,2}, Xiaomin Zhao ², Xi Guo,² and Jiaxin Guo²

¹College of Forestry, Jiangxi Agricultural University, Nanchang 330045, China

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

Correspondence should be addressed to Xiaomin Zhao; zhaoxm889@126.com

Received 5 June 2019; Revised 17 August 2019; Accepted 30 August 2019; Published 29 November 2019

Academic Editor: Luis Javier Herrera

Copyright © 2019 Zhe Xu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Deep learning is characterized by its strong ability of data feature extraction. This method can provide unique advantages when applying it to visible and near-infrared spectroscopy for predicting soil organic matter (SOM) content in those cases where the SOM content is negatively correlated with the spectral reflectance of soil. This study relied on the SOM content data of 248 red soil samples and their spectral reflectance data of 400–2450 nm in Fengxin County, Jiangxi Province (China) to meet three objectives. First, a multilayer perceptron and two convolutional neural networks (LeNet5 and DenseNet10) were used to predict the SOM content based on spectral variation and variable selection, and the outcomes were compared with that from the traditional back-propagation neural network (BPN). Second, the four methods were applied to full-spectrum modeling to test the difference to selected feature variables. Finally, the potential of direct modeling was evaluated using spectral reflectance data without any spectral variation. The results of prediction accuracy showed that deep learning performed better at predicting the SOM content than did the traditional BPN. Based on full-spectrum data, deep learning was able to obtain more feature information, thus achieving better and more stable results (i.e., similar average accuracy and far lower standard deviation) than those obtained through variable selection. DenseNet achieved the best prediction result, with a coefficient of determination (R^2) = 0.892 ± 0.004 and a ratio of performance to deviation (RPD) = 3.053 ± 0.056 in validation. Based on DenseNet, the application of spectral reflectance data (without spectral variation) produced robust results for application-level purposes (validation R^2 = 0.853 ± 0.007 and validation RPD = 2.639 ± 0.056). In conclusion, deep learning provides an effective approach to predict the SOM content by visible and near-infrared spectroscopy and DenseNet is a promising method for reducing the amount of data preprocessing.

1. Introduction

Soil organic matter (SOM) content, a key indicator of soil fertility, substantially impacts the physicochemical properties and quality of soil; thus, the SOM content must be considered in scientific fertilization. Visible (VIS, 400–780 nm) and near-infrared (NIR, 780–2526 nm) spectroscopy is a convenient and efficient technique for quickly and inexpensively monitoring SOM [1], since spectral reflectance of soil is negatively correlated with the SOM content and the SOM content could be obtained from measured soil reflectance spectrum [2, 3].

Many studies have proposed and tested various spectral data modeling techniques, including linear regression (LR), partial least squares regression (PLSR), back-propagation (BP) neural network (BPN), and support vector machine (SVM). Xie et al. [4] predicted the SOM content in mountain red soil using PLSR, BPN, SVM, and a combination model based on the radial basis function (RBF) neural network applied on a full spectrum (450–2450 nm); they found that the RBF-based combination model gave the best results with a ratio of the performance to deviation (RPD = 2.06), followed by the SVM (RPD = 1.67). Ye et al. [5] found that compared with LR, BPN yielded better results for prediction

of the SOM content based on hyperspectral data. In other work, Shi et al. [6] were able to use PLSR to predict the SOM content from the Chinese VIS-NIR spectral library.

After data preprocessing by principal components analysis, Zeng et al. [7] found that the SVM gave the best prediction results of the SOM content (RPD = 2.28). Additionally, Ji et al. [8] used 441 soil samples (400–2450 nm) to predict the SOM content, for which the SVM had a RPD = 2.16, but the best accuracy came from PLSR-BP with a RPD = 2.36. The inversion accuracy of models such as PLSR, BPN, and SVM is generally higher than that of LR, the most commonly used modeling technique [5]. Chen et al. [9] used the BPN to update the importance variables generated from a random forest to provide a variable selection strategy. According to prior studies, the current approach to predict the SOM content by VIS-NIR spectroscopy finds the feature spectrum and then establishes a prediction model [9–14]. Most of these studies have focused on the preprocessing of soil spectral data and the screening of useful feature spectra. Yet we still need a high-performance modeling technique to simplify the preprocessing requirements of spectral data, which is also crucial for ensuring accurate predictions.

Deep learning has developed rapidly in recent years. This method allows for a computational model consisting of multiple processing layers to learn data representations with multiple levels of abstraction [15]. By learning the deep nonlinear network structure, a complex function approximation is realized using the BP algorithm. The obtained results indicate how the deep learning machine should change its internal parameters to discover the complex structure of larger data sets, demonstrating the powerful ability to learn the essential features of a data set from a smaller sample set [16].

Recently, Chen et al. [17] proposed a deep learning method using a multilayer perceptron (MLP) structure to predict the soil organic carbon content. In addition, the convolutional neural network (CNN) has been applied to studies of image recognition. The CNN uses the convolution and pooling operations to extract the abstract feature maps of the data, layer by layer, thereby learning the structural features and their essential relationships within the spectral data [18]. In this way, the spectral curve can be regarded as a wavelength \times 1 gray scale image, and the same-padding skill may be used for convolution operations in a deeper network. Therefore, it should be feasible to perform SOM inversion directly using raw or transformed spectral data. As computing power improves and deep learning rapidly develops, exploring how deep learning may be applied for predicting the SOM content from VIS-NIR wavelengths is increasingly necessary.

To this end, this study applied three deep learning models to estimate the SOM content and compared their accuracy to that of the traditional BPN model. Since deep learning can learn the essential features of a data set, we also compared outcomes based on the full spectrum versus selected characteristic spectrum. Furthermore, the best-performing model was used to fit the spectral reflectance data and test whether some data preprocessing steps could be removed.

2. Materials and Methods

2.1. Study Area and Sampling and Data Collection. The study area was located in the central part of Fengxin County, Jiangxi Province, in China, which has typical red soil. The soil samples were collected from gardens, woodlands, and paddy fields throughout this area. Specifically, 1 km \times 1 km grid was used to select the sampling point, taking into account its topography, vegetation cover, and land use type. From each grid one sample was collected and used; for areas with complex geographies, more sampling points were used per grid to ensure adequate data representation. Figure 1 shows the spatial distribution of the 248 soil samples obtained in total. Each composite sample was obtained by a four-point mixing method: the sampling depth was 0–20 cm for paddy fields or 0–30 cm for gardens and woodlands. All samples were air-dried in the laboratory.

After removing debris, the soil samples were grounded and passed through a 2 mm sieve. Each sample was then divided into two parts, for soil spectroscopy and SOM analysis, respectively. The SOM was determined using a potassium dichromate solution [19], while a FieldSpec4 spectrometer (ASD Inc., Cambridge, United Kingdom) was used to measure the spectral reflectance of a sample. The spectral acquisition range of the FieldSpec4 spectrometer is 350–2500 nm, and its spectral sampling interval is 1.4 nm (350–1000 nm) and 2 nm (1001–2500 nm), with a resampling interval of 1 nm. A total of 2151 wavelength variables were generated for SOM content prediction. A soil sample was placed in a black sample dish 6 cm in diameter and 2 cm deep, filled to the brim and its surface flattened with a ruler. The built-in light source included with the MugLite device was used for measurements; it was positioned above the sample dish in the slot atop the instrument. Both dark current and standard whiteboard calibrations were performed on the instrument before each sample's data acquisition. Five spectral data were collected per sample, for which their arithmetic mean was taken as the spectral curve for that sample to reduce measurement error.

2.2. Data Preprocessing. Because the influence from the immediate environment and the instrument itself together generated substantial noise in the edge band of the measured spectrum, the wavelengths spanning 350–399 nm and 2451–2500 nm were removed. Wavelet transform was used to reduce the noise generated during the measurement process: three-layer decomposition was performed by the Daubechies6 wavelet and soft thresholding was used to detail high-frequency coefficients [2, 20, 21]. To reduce data dimensions and data redundancy, resampling by 10 nm interval was carried out to accelerate the training process, whose result was similar to the original data. Figure 2 shows the spectral curves of red soils after preprocessing, in which the SOM content was divided into six groups of <15, 15–25, 25–35, 35–45, 45–55, and >55 g/kg and their spectra average taken. Evidently, a distinct iron oxide absorption valley was present in the samples around 900 nm, accompanied

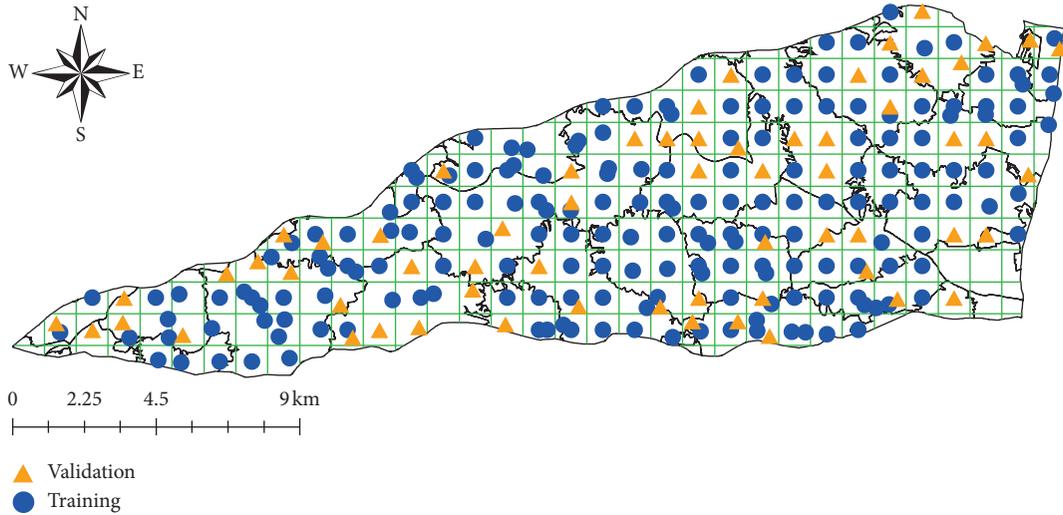


FIGURE 1: Spatial distribution of soil sampling points in Fengxin County (Jiangxi Province, China).

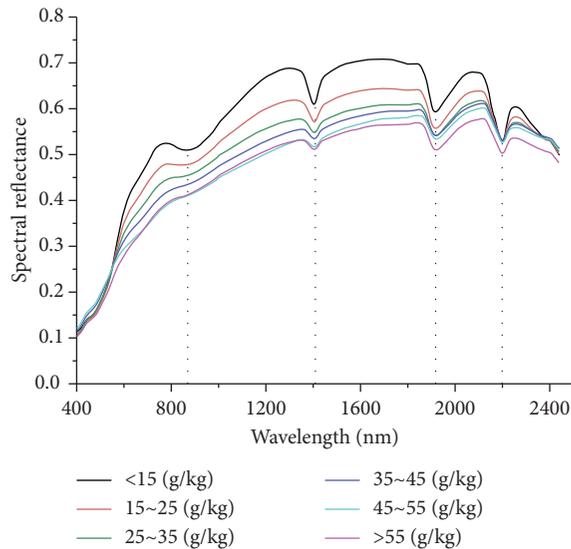


FIGURE 2: Spectral curves of soils differing in their organic matter content.

by apparent water absorption valleys at 1400, 1900, and 2200 nm [22].

Following the work of Hong et al. [12], Zhang et al. [23], and Xu et al. [24], the fractional-order derivative (FOD) algorithm was used as a mathematical method to analyze the obtained reflection spectra. This allows interpolation between integer derivatives and thereby extracts more exceptional details from the spectral signals. The 1.5 order derivative with the Grünwald–Letnikov method was applied here to transform the spectral data, which generated 203 wavelength band variables. This Grünwald–Letnikov process is shown in formula (1), where ν is the order, $\Gamma(x)$ is the Gamma function, and n is the difference between the upper and lower limits of the derivative:

$$\begin{aligned} \frac{d^\nu f(x)}{dx^\nu} \approx & f(x) + (-\nu)f(x-1) + \frac{(-\nu)(-\nu+1)}{2}f(x-2) + \dots \\ & + \frac{\Gamma(-\nu+1)}{n!\Gamma(-\nu+n+1)}f(x-n). \end{aligned} \quad (1)$$

The Pearson correlation coefficients between transformed the spectral data and SOM content are shown in Figure 3(a). In all, 67 variables (at 620–650, 670, 730–840, 970, 980, 1220, 1270–1390, 1420, 1430, 1530, 1580–1620, 1720–1770, 1850–1940, 1990–2030, 2230, and 2290–2310 nm) having r^2 values >0.4 and P values <0.01 (Figure 3(b)) were selected for further analysis, and the wavelength around 900 nm affected by iron oxide was removed.

Based on the studies of Xie et al. [4] and Ji et al. [8], 186 training samples and 62 validation samples were then generated using a 3:1 ratio, with the geostatistics module in ArcGIS 10.5 (ESRI Inc., Redlands, USA). The spatial distribution of these training and validation samples is shown in Figure 1.

2.3. Model Architecture. The MLP is a forward-structured artificial neural network, consisting of multiple layers of neurons and their connections. In addition to the input nodes, each node functions as a neuron (or processing unit) with a nonlinear activation function [25]. Here, the BP algorithm was used to train the MLP. Figure 4 depicts the eight-layer deep MLP architecture used, which had seven hidden layers.

The typical architecture for the CNN is LeNet5 [26], in which the convolutional layer and the pooling layer are alternated with the fully connected layer. It was Harley [27] who realized the 2D and 3D visualization of LeNet5’s architecture, demonstrating the scale and complexity of typical CNN architecture (<http://www.cs.cmu.edu/~aharley/vis/>).

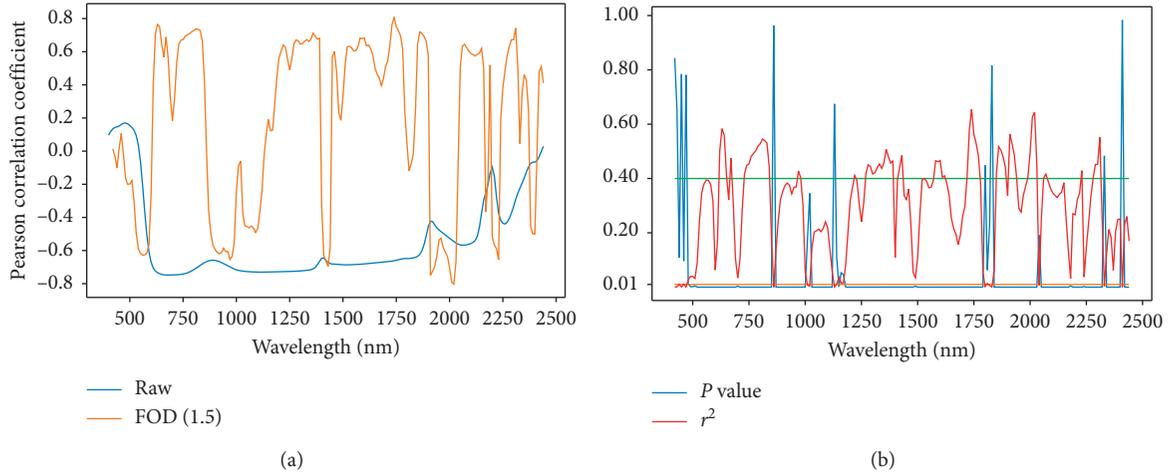


FIGURE 3: Correlation coefficients of raw spectral reflectance/fractional-order derivative (1.5) with the soil organic matter content (a) and 67 variables selected from the fractional-order derivative (1.5) with $r^2 > 0.4$ and $P < 0.01$ (b).

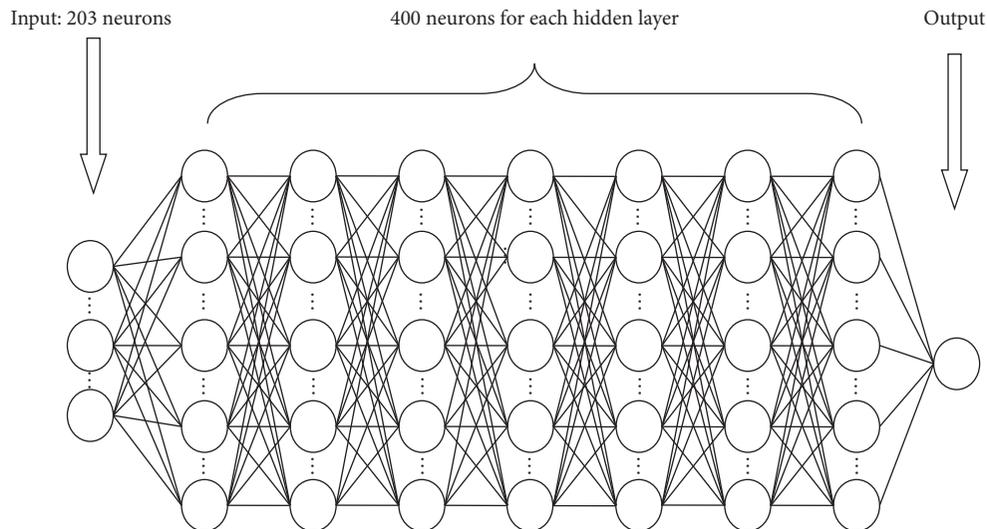


FIGURE 4: The architecture of multilayer perceptron for deep learning.

DenseNet [28] is a type of CNN having dense connections and drawing on the shortcut idea of ResNet [29]. In such a network, there is a direct connection between any two layers, for which the input of each layer of the network is the union of outputs from all prior layers, and the feature map learned by a given layer is also directly transmitted to all layers behind it and used as input. Figure 5 shows a diagram for the DenseNet architecture with three dense blocks, consisting mainly of two components, a dense block and a transition layer. Dense connections have many obvious advantages, since each module uses the available information from all layers in front of the module and each layer has a dense connection to the preceding layer. Such connections could strengthen the transfer of gradients, enhance feature reuse, and reduce overfitting of small-sized sample data sets [30].

3. Deep Learning Proposal

3.1. Activation Function. Admittedly, some problems can occur in the BP process of deep neural networks, such as the disappearance of the gradient and slow training. Nevertheless, the process can be optimized, by adjusting both the activation function and the optimizer. Each neuron node in the neural network first accepts the output value of the upper neuron as the input value of that neuron and transmits this input value to the next neuron. The input neuron node will directly transfer the input attribute value to the next neuron. In multilayer neural networks, there is a functional relationship between the output of the upper nodes and the input of lower nodes, which is called an activation function. In this respect, traditional MLP uses the sigmoid function (formula (2), where $f(x)$ is a nonlinear function), such that

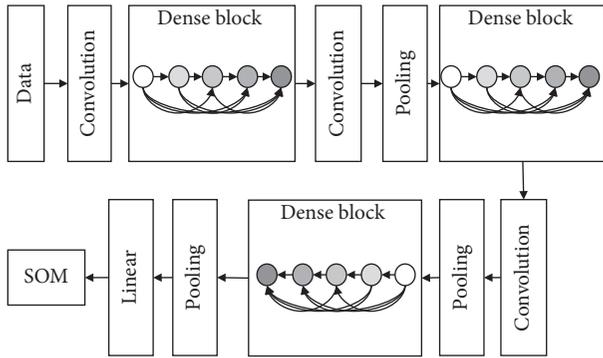


FIGURE 5: The architecture of DenseNet for deep learning [28].

the BP process of MLP multiplies the partial derivatives of the function, layer by layer, and the derivative interval of sigmoid ranges from 0 to 0.25. Therefore, when MLP's layers are profound its gradient will disappear; this leaves the problem that the power operation will increase the training time in a large-scale deep network [31]:

$$f(x) = \frac{1}{1 + e^{-x}}. \quad (2)$$

In LeNet5 convolution, the tanh activation function (formula (3), where $f(x)$ is a nonlinear function) is used, which is beset by same problem as the sigmoid function. However, the difference between them is that tanh's derivative range is 0 to 1; thus, the tanh function is better equipped than the sigmoid function for practical applications [32]:

$$f(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}. \quad (3)$$

The rectified linear unit (ReLU) has the advantage of easy optimization. According to formula (4), the output of half of the definition domain is zero, whereas the second derivative of the ReLU is almost zero everywhere and the first derivative of the modified linear element is 1 when it is in the active state. Thus, when the parameters of affine transformation are initialized, b can be set to a small positive value, such as 0.1; this makes it possible for linear rectifier units to activate most of the inputs in the training set at the beginning, allowing the derivatives to pass [31]. Currently, the ReLU is the most widely used activation function in deep learning applications [33]:

$$f(x) = \max(0, w^T x + b). \quad (4)$$

3.2. Avoiding Model Overfitting. Generalization ability is a model's adaptability to validation samples and it is an important index to evaluate the overall performance of a given model [34]. Taking the BPN as an example, it often suffers from the phenomenon of overfitting, in which the model performs well with the training data but performs poorly with the validation data. Deep learning has a strong ability to fit data; thus, some robust methods are also needed to prevent overfitting and to build models with excellent generalization ability.

In this respect, the most commonly used method to avoid overfitting is the L2 regularization. It adds the sum of squares of weighted parameters directly from the original loss function, as represented by formula (5), where L is loss, E_{in} is the training sample error without the regularization term, and λ is an adjustable regularization parameter:

$$L = E_{in} + \lambda \sum_j \omega_j^2. \quad (5)$$

Including a hyperparameter dropout can ignore half of the feature detectors in each training batch (when the dropout is set to 0.5), that is, let half of the hidden layer nodes have a value of 0 to reduce the mutual correlation between the feature detectors (hidden layer nodes). When the network is propagated in its forward direction, the activation value of an individual neuron can be stopped with a certain probability. However, because the whole network does not rely too much on some local features, this can significantly reduce model overfitting [35].

Early stopping is also a technique used to prevent overfitting. In a deep neural network, overfitting problems are more of a risk. Therefore, while generating models in the training iteration process the model is simultaneously evaluated with a verification set. Each training iteration output performance is saved. If there is no better result within a certain number of iterations, the training is terminated, and the better weighted parameters are used as output [36].

3.3. Model Experimentation. This was carried out using a desktop computer equipped with an Intel Core i9 7920X CPU and 64 GB of memory. Its operating system was Windows 10, and two GeForce 1080TI GPUs with 11 GB of memory each provided acceleration for model training and validation. The Keras framework with Tensorflow backend supported the implementation of all neural network models. Keras is a simple and easy-to-use neural network library that provides most of the building blocks needed to build a relatively complex model [37].

In all experiments, the Nadam optimizer [38] was used to accelerate the training process, and the batch size was set to 32. The BPN was built with an input layer, a hidden layer, and an output layer, wherein the input layer had 203 neurons and the single hidden layer had 400 neurons, the activation function was sigmoid, and the learning rate was 0.001. The MLP was based on the architecture depicted in Figure 3, using a dropout value set to 0.3, and its learning rate was 0.00001. The LeNet5 architecture changed the activation function of the output neuron to the sigmoid function for prediction value, and the output neuron performed an L2 regularization to reduce overfitting; its dropout rate was set to 0.3, with a learning rate of 0.001. DenseNet also changed the activation function of output neurons to the sigmoid function for regression but it removed all the batch-normalization operations in the architecture, adjusting the kernel size of the pooling layer to adapt the input data; its dropout rate was set to 0.5, with a learning rate of 0.001.

SOM content data were normalized to speed up the training process. Root mean square error (RMSE) was used

as the loss function in the training process with all neural network models tested. The coefficient of determinant (R^2), RMSE, and RPD were used to evaluate the optimization of the model. An RPD value between 1.5 and 2 indicates the model can achieve rough estimation, 2.0 to 2.5 indicates the model has moderate predictive ability, 2.5 to 3 indicates the model has good predictive ability, and any value higher than 3 indicates that the model has excellent predictive ability [39]. To prevent overfitting, all the neural network models were engaged in the abovementioned “early stopping” method. The indicator display form used the average index plus standard deviation of the last P (P stands for the patience of iterations according to the early stopping setting) models before an exit.

4. Results and Discussion

4.1. Approach Process of Different Models. Figure 6(a) shows the approach process of the BPN, MLP, LeNet5, and DenseNet10 models applied to the 203 variables. Their corresponding evaluation curves demonstrated that the training and the verification loss functions of each model were small, suggesting all were characterized by good generalization ability. The entire training process revealed that use of dropout and L2 regularization can effectively suppress the overfitting phenomenon; the prediction accuracy of a given model was stable in the later stage of training, and its prediction results could be accurately calculated using the early exit technique. Additionally, the BPN used the sigmoid activation function and LeNet5 used the tanh activation function. Comparing the MLP with DenseNet10 using the ReLU activation function, the training process using the ReLU activation function model was smoother and more efficient at fitting the training data.

Figure 6(b) shows the training and approach process for BPN, MLP, LeNet5, and DenseNet10 applied to the 67 selected variables. For MLP, LeNet5, and DenseNet10, the amplitude of their validation curves was more significant than for the 203 variables. For MLP, as the input decreased, it would easily undergo overfitting had the dropout method not been used, but the dropout randomly chosen would ignore the hidden layer nodes. For every batch training process, since each hidden layer node was randomly ignored, the network of each epoch was somewhat different. The dropout effect was the same for the CNN. Further, less information reduced receptive fields (i.e., the region in the input space corresponding to a particular feature of the CNN); hence, fewer neurons might lose some features. As shown by our results, a smaller amount of data increased the difference in accuracy between each epoch, resulting in a greater gap between each epoch of the validation set and a higher standard deviation. These results collectively indicated that full spectrum is more suitable for deep learning.

4.2. Prediction Accuracy of Different Models. Table 1 summarizes the prediction accuracy results of the MLP, CNN, and BPN models. The accuracy of the validation set samples is the most important indicator to measure the performance

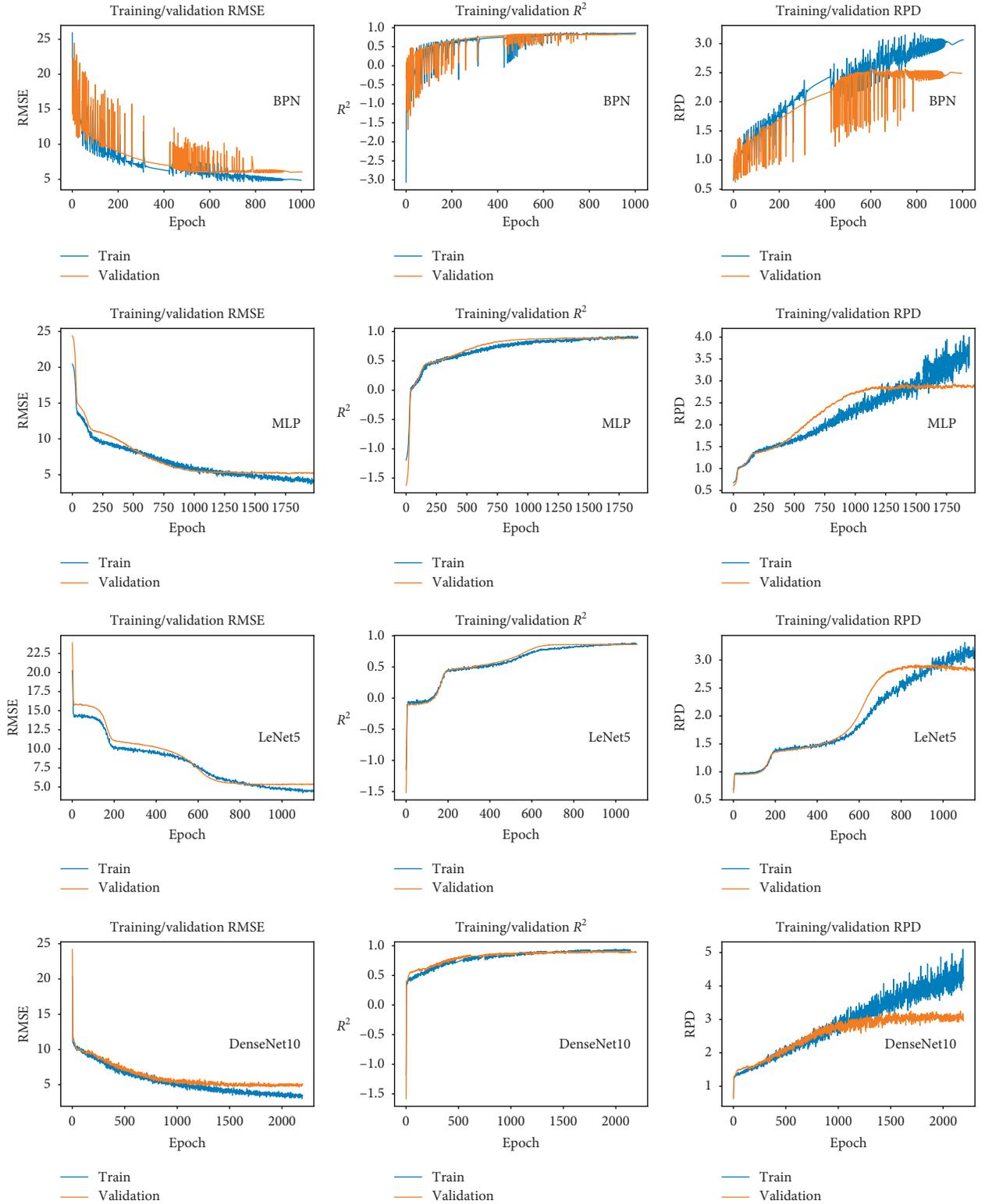
of a given model. With 203 variables, DenseNet10 had the largest coefficient of determination ($R^2 = 0.892 \pm 0.004$) and the smallest root mean square error ($\text{RMSE} = 4.933 \pm 0.091$), while its performance deviation ratio was the highest ($\text{RPD} = 3.053 \pm 0.056$). The prediction accuracy of MLP did not differ from that of DenseNet10. LeNet5 had an RPD around 2.8, whereas the BPN had an RPD below 2.5. DenseNet10 increased its R^2 value by more than 0.06, decreased its RMSE by ~ 1.18 , and increased its RPD by ~ 0.59 compared with the BPN. The traditional modeling method, BPN, had a particular gap in verification accuracy when compared with the deep learning model, with full-spectrum data.

The MLP entails a structural evolution of the BPN. The results obtained with 203 bands showed that the MLP has a stronger ability to fit data fitting ability because its artificial neural network with multiple hidden layers has an excellent feature learning ability and the learned features are essential for data characterization. The drawback of shallow structure algorithms is that their ability to represent complex functions is limited in the case of finite samples and computational units, hindering the generalization ability for complex problems.

The number of parameters for DenseNet10 (43,273) was far lower than that of LeNet5 (108,941). The advanced architecture of the CNN gave an absolute improvement in the prediction accuracy with fewer parameters. After testing both DenseNet40 and DenseNet121 architectures of the deeper network, the results did not improve, which meant that Occam’s razor law should be invoked. The eight-layer MLP had a slight gap vis-à-vis DenseNet10 in its prediction accuracy. The MLP used full connections, so its total number of parameters, at 1,044,401, was much higher than that of DenseNet10, demonstrating that DenseNet has clear advantages for parsimonious modeling.

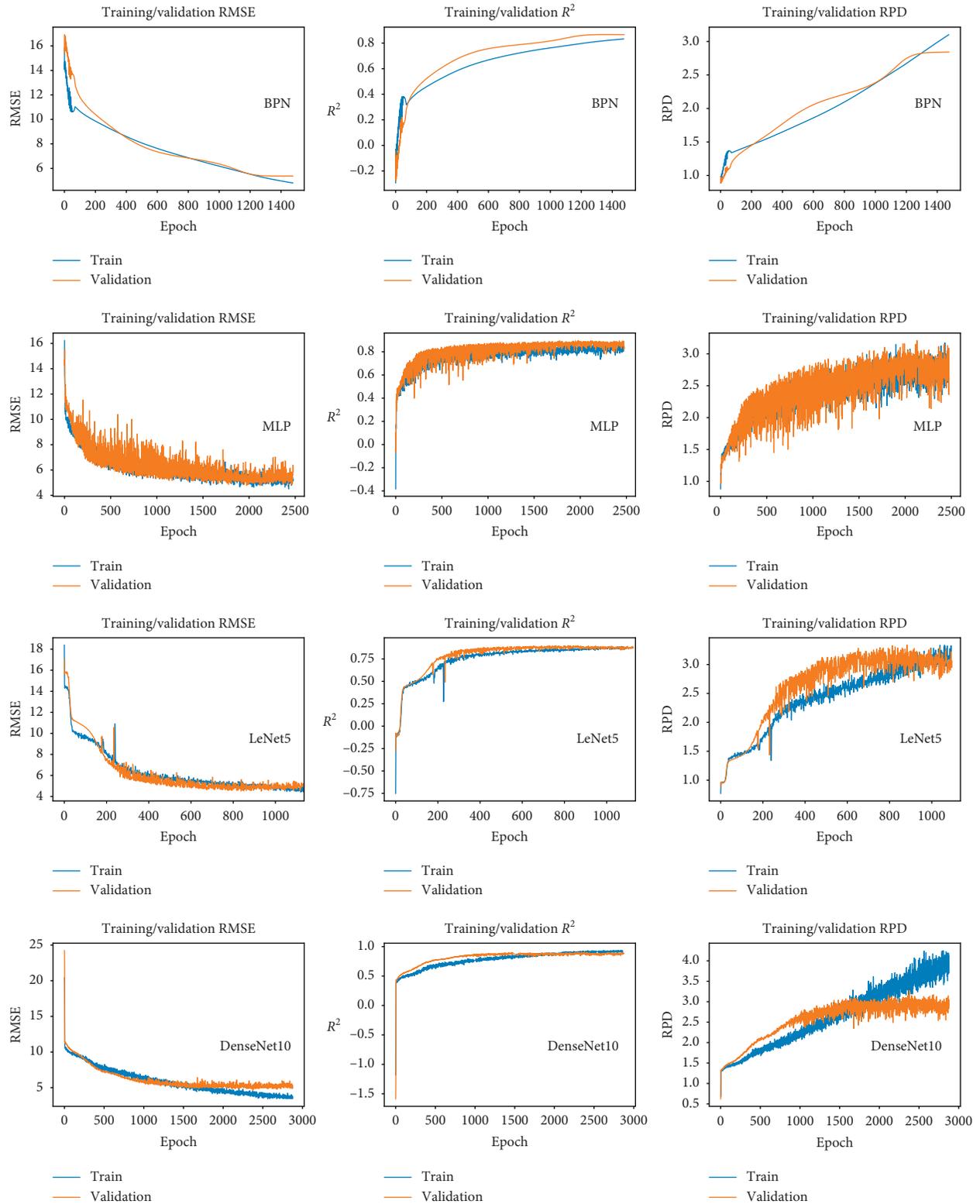
Table 1 also provides the predicted accuracy results of different models from selected feature variables. Among MLP, LeNet5, and DenseNet10, their overall accuracy results were not substantially different between the 67 and 203 variables used. However, the BPN was better adjusted to selected feature variables because of its shallow structure. For MLP, fewer bands would quickly lead to a gradient disappearance problem, resulting in lower model accuracy. For both BPN and MLP, with fewer variables, the R^2 values in the training set were lower than those in the validation set. By contrast, the CNN had better generalization ability with fewer bands. In conclusion, deep learning can achieve the same prediction accuracy without screening sensitive variables and is more suited to fit SOM data obtained from the full spectrum with using feature variables.

4.3. Prediction without Any Spectral Variation. During data preprocessing, FOD can effectively improve the correlation between soil spectral reflectance and SOM content, making the full-spectrum data more useful for analytical modeling. However, FOD is a time-costly algorithm, one not conducive to real-time monitoring. Table 2 summarizes the prediction accuracy results of using different depths of DenseNet to fit



(a)

FIGURE 6: Continued.



(b)

FIGURE 6: Training and validation curves for model evaluation using the root mean square error (RMSE), coefficient of determination (R^2), and ratio of performance to deviation (RPD): (a) approach of different models using the full spectrum with 203 variables and (b) approach of different models applied to a subset of 67 selected variables.

TABLE 1: Prediction accuracy results of different model with the fractional-order derivative (1.5) transformation of soil spectral reflectance data.

Model	Variables	Training		Validation		
		R^2	RMSE	R^2	RMSE	RPD
BPN	67	0.831 ± 0.003	4.900 ± 0.062	0.867 ± 0.000	5.382 ± 0.001	2.843 ± 0.001
BPN	203	0.835 ± 0.018	5.107 ± 0.290	0.834 ± 0.018	6.112 ± 0.290	2.465 ± 0.092
MLP	67	0.806 ± 0.021	5.657 ± 0.288	0.861 ± 0.019	5.578 ± 0.364	2.709 ± 0.164
MLP	203	0.890 ± 0.011	4.249 ± 0.199	0.886 ± 0.002	5.044 ± 0.037	2.987 ± 0.022
LeNet5	67	0.872 ± 0.015	4.542 ± 0.224	0.870 ± 0.020	5.304 ± 0.401	2.901 ± 0.220
LeNet5	203	0.902 ± 0.006	3.909 ± 0.114	0.863 ± 0.001	5.471 ± 0.033	2.790 ± 0.022
DenseNet10	67	0.927 ± 0.009	3.376 ± 0.191	0.888 ± 0.012	4.925 ± 0.274	3.133 ± 0.177
DenseNet10	203	0.907 ± 0.009	3.367 ± 0.177	0.892 ± 0.004	4.933 ± 0.091	3.053 ± 0.056

Data are the mean \pm standard deviation. R^2 = coefficient of determination, RMSE = root mean square error, and RPD = the ratio of performance to deviation.

TABLE 2: Prediction accuracy results of using different depths of DenseNet based on the full-spectrum data without the fractional-order derivative transformation.

Model	Learning rate	Training		Validation		
		R^2	RMSE	R^2	RMSE	RPD
DenseNet10	0.001	0.832 ± 0.012	5.334 ± 0.177	0.806 ± 0.007	6.519 ± 0.128	2.345 ± 0.050
DenseNet10	0.0006	0.837 ± 0.013	5.188 ± 0.203	0.836 ± 0.010	6.037 ± 0.199	2.508 ± 0.085
DenseNet13	0.001	0.878 ± 0.010	4.376 ± 0.164	0.833 ± 0.005	6.117 ± 0.095	2.462 ± 0.038
DenseNet13	0.0006	0.861 ± 0.012	4.697 ± 0.199	0.828 ± 0.006	6.207 ± 0.108	2.431 ± 0.044
DenseNet16	0.001	0.856 ± 0.012	4.782 ± 0.197	0.827 ± 0.008	6.172 ± 0.135	2.461 ± 0.051
DenseNet16	0.0006	0.858 ± 0.011	4.792 ± 0.177	0.848 ± 0.007	5.847 ± 0.131	2.576 ± 0.058
DenseNet19	0.001	0.870 ± 0.010	4.515 ± 0.179	0.853 ± 0.007	5.722 ± 0.124	2.639 ± 0.056
DenseNet19	0.0006	0.916 ± 0.009	3.628 ± 0.177	0.853 ± 0.011	5.745 ± 0.222	2.622 ± 0.099

Data are the mean \pm standard deviation. R^2 = coefficient of determination, RMSE = root mean square error, and RPD = the ratio of performance to deviation.

the model based on raw reflectance data. Across different depths, DenseNet19 gave the best result with the validation set, with $R^2 = 0.853 \pm 0.007$, $RMSE = 5.722 \pm 0.124$, and $RPD = 2.639 \pm 0.056$; however, this was worse than the results obtained from transformed spectral data. Thus, although deep learning has robust data mining ability, its prediction results are still based on data learning. For good prediction accuracy, it is therefore necessary to improve the correlation between the SOM content and spectral data through transformation of spectral data. However, the result of DenseNet is deemed acceptable for practical purposes.

Our result for DenseNet19 has an improved RPD that is 0.97 greater than that of the SVM, and 0.58 more than that from the RBF combination model reported by Xie et al. [4] who predicted the SOM content in mountain red soil based on spectral reflectance data. Additionally, the RPD for DenseNet19 exceeds that for both SVM ($RPD = 2.16$) and PLSR-BP ($RPD = 2.36$) obtained by Ji et al. [8] who applied training models on a full spectrum (450–2450 nm). These findings therefore suggest that DenseNet is a powerful tool for data feature extraction.

Due to the overlapping absorption characteristics of spectral activity, the VIS-NIR spectra of soils are multilinear, broad, and nonspecific, which may weaken the model performance of SOM estimation. A deep learning algorithm embodies the powerful ability of data feature extraction, excludes outlier data, and finds hidden patterns in the data set, which can especially solve nonlinear problems with high

model accuracy. However, in the process of model building, local optimal problems occur frequently in training raw reflectance in DenseNet, the gradient tends to disappear in deep MLP, while the setting of hyperparameters and the optimization of model structure are time costly. Although this modeling process can be complicated and time-consuming, the accuracy of its prediction results is generally high after the model is built. The prediction based on spectral reflectance data enables robust prediction accuracy, which could effectively reduce the amount and time spent on data preprocessing, thereby improving the efficiency of real-time monitoring.

5. Conclusions

In this study we investigated deep learning framework algorithms for predicting the SOM content by VIS-NIR spectroscopy. Based on FOD (1.5) spectral variation, we compared BPN, MLP, and CNN (including LeNet5 and DenseNet10) with full-spectrum data (203 variables) and a subset of 67 variables highly correlated with the SOM content (r^2 values >0.4). Our results indicate that deep learning methods including the MLP and CNN can be used to predict the SOM content from VIS-NIR soil spectra, each displaying state-of-the-art performance. Hence, these methods are better suited to fit the full-spectrum data where more information leads to stable results, as their averaged accuracy is similar to that obtained with selected variables, but standard deviations are much lower.

The multilayer artificial neural network model has a strong feature learning ability, and the feature data obtained by the deep learning model could capture a more essential representation of the original soil data. As a high-performance deep learning model, the CNN can extract effective feature structures from complex spectral data for learning, displaying stronger model expression ability than traditional shallow learning models. Moreover, the CNN reduces the number of parameters needed for SOM prediction and improves the generalization ability of the model via its network structure of local connection and weight sharing.

Overall, the DenseNet architecture gives the best prediction accuracy with fewer calculation parameters. It also achieves high accuracy without FOD (1.5) transformation of soil spectra data. As DenseNet reduces the data preprocessing of variable selection and spectral variation, it is suitable for real-time monitoring. Hence, we suggest DenseNet is a promising solution for predicting the SOM content by VIS-NIR spectroscopy. This method could also be widely used in other similar spectral applications.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This research was supported by the National Natural Science Foundation of China (41361049).

References

- [1] C. R. Lobsey, R. A. Viscarra Rossel, P. Roudier, and C. B. Hedley, "rs-localdata-mines information from spectral libraries to improve local calibrations," *European Journal of Soil Science*, vol. 68, no. 6, pp. 840–852, 2017.
- [2] Z. Shi, *Principle and Method of Soil Surface Hyperspectral Remote Sensing*, Science Press, Beijing, China, 2014.
- [3] W.-J. Ji, Z. Shi, Q. Zhou, and L.-Q. Zhou, "VIS-NIR reflectance spectroscopy of the organic matter in several types of soils," *Journal of Infrared and Millimeter Waves*, vol. 31, no. 3, pp. 277–282, 2012.
- [4] W. Xie, X. M. Zhao, X. Guo, Y. C. Ye, X. X. Sun, and L. H. Kuang, "Spectrum-based estimation of the content of soil organic matters in mountain red soil using RBF combination model," *Scientia Silvae Sinicae*, vol. 54, no. 6, pp. 16–23, 2018.
- [5] Q. Ye, X. Q. Jiang, X. C. Li, and Y. Lin, "Comparison on inversion model of soil organic matter content based on hyperspectral data," *Transactions of the Chinese Society for Agricultural Machinery*, vol. 48, no. 3, pp. 169–177, 2017.
- [6] Z. Shi, W. Ji, R. A. Viscarra Rossel, S. Chen, and Y. Zhou, "Prediction of soil organic matter using a spatially constrained local partial least squares regression and the Chinese VIS-NIR spectral library," *European Journal of Soil Science*, vol. 66, no. 4, pp. 679–687, 2015.
- [7] Y. Zeng, Y. Z. Lu, C. W. Du, and J. M. Zhou, "Applying infrared photoacoustic spectroscopy and support vector machine model to quantify soil organic matter content," *Acta Pedologica Sinica*, vol. 51, no. 6, pp. 1262–1269, 2014.
- [8] W. J. Ji, X. Li, C. X. Li, Y. Zhou, and Z. Shi, "Using different data mining algorithms to predict soil organic matter based on visible-near infrared spectroscopy," *Spectroscopy and Spectral Analysis*, vol. 32, no. 9, pp. 2393–2398, 2012.
- [9] H. Chen, X. Liu, Z. Jia, Z. Liu, K. Shi, and K. Cai, "A combination strategy of random forest and back propagation network for variable selection in spectral calibration," *Chemometrics and Intelligent Laboratory Systems*, vol. 182, pp. 101–108, 2018.
- [10] Y. Hong, Y. Chen, L. Yu et al., "Combining fractional order derivative and spectral variable selection for organic matter estimation of homogeneous soil samples by VIS-NIR spectroscopy," *Remote Sensing*, vol. 10, no. 3, p. 479, 2018.
- [11] L. Yu, Y. S. Hong, Y. Zhou et al., "Wavelength-variable selection methods for estimation of soil organic matter content using hyperspectral technique," *Transactions of the Chinese Society of Agricultural Engineering*, vol. 32, no. 13, pp. 95–102, 2016.
- [12] Y. S. Hong, Y. X. Zhu, X. P. Su, Q. Zhu, Y. Zhou, and L. Yu, "Estimation of soil organic matter content using hyperspectral techniques combined with normalized difference spectral index," *Spectroscopy and Spectral Analysis*, vol. 37, no. 11, pp. 3537–3542, 2017.
- [13] S. Nawar, H. Buddenbaum, J. Hill, J. Kozak, and A. M. Mouazen, "Estimating the soil clay content and organic matter by means of different calibration methods of vis-nir diffuse reflectance spectroscopy," *Soil and Tillage Research*, vol. 155, pp. 510–522, 2016.
- [14] D. H. Zhang, Y. J. Zhao, K. Qin, N. B. Zhao, and Y. C. Yang, "Influence of spectral transformation methods on nutrient content inversion accuracy by hyperspectral remote sensing in black soil," *Transactions of the Chinese Society of Agricultural Engineering*, vol. 34, no. 20, pp. 141–147, 2018.
- [15] G. E. Hinton and R. Salakhutdinov, "Reducing the dimensionality of data with neural networks," *Science*, vol. 313, no. 5786, pp. 504–507, 2006.
- [16] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015.
- [17] H. Z. Chen, Z. Y. Liu, J. Gu, W. Ai, J. B. Wen, and K. Cai, "Quantitative analysis of soil nutrition based on FT-NIR spectroscopy integrated with BP neural deep learning," *Analytical Methods*, vol. 10, no. 41, pp. 5004–5013, 2018.
- [18] C. Wang, X. H. Wu, L. Q. Li, Y. S. Wang, and Z. W. Li, "Convolutional neural network application in prediction of soil moisture content," *Spectroscopy and Spectral Analysis*, vol. 38, no. 1, pp. 24–47, 2018.
- [19] V. M. Marta, E. B. Anne, M. Felipe, and K. Heike, "Biochemically altered charcoal residues as an important source of soil organic matter in subsoils of fire-affected subtropical regions," *Geoderma*, vol. 262, pp. 62–70, 2016.
- [20] C. H. Ma and L. Y. Liu, "Spectral data processing based on wavelet analysis," *Metallurgical Analysis*, vol. 32, no. 1, pp. 34–37, 2012.
- [21] Y. Hu, T. Jiang, A. Shen, W. Li, X. Wang, and J. Hu, "A background elimination method based on wavelet transform for Raman spectra," *Chemometrics and Intelligent Laboratory Systems*, vol. 85, no. 1, pp. 94–101, 2007.
- [22] X. M. Zhao and M. H. Yang, "Hyper-spectral characteristics of major types of soils in red soil region of Jiangxi province," *Acta Pedologica Sinica*, vol. 55, no. 1, pp. 31–42, 2018.

- [23] D. Zhang, T. Tashpolat, F. Zhang, K. Ardak, and N. Xia, "Effect of fractional differential algorithm on hyperspectral data of saline soil," *Acta Optica Sinica*, vol. 36, no. 3, 2016.
- [24] J. G. Xu, X. L. Feng, L. Guan, S. Wang, and Q. L. Hu, "Fractional differential application in reprocessing infrared spectral data," *Control and Instruments in Chemical Industry*, vol. 39, no. 3, pp. 347–351, 2012.
- [25] D. E. Rumelhart, G. E. Hinton, R. J. Williams, and R. J. Williams, "Learning representations by back-propagating errors," *Nature*, vol. 323, no. 6088, pp. 533–536, 1986.
- [26] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278–2324, 1998.
- [27] A. W. Harley, "An interactive node-link visualization of convolutional neural networks," *Advances in Visual Computing*, vol. 9474, pp. 867–877, 2015.
- [28] G. Huang, Z. Liu, V. Laurens, V. D. Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proceedings of the 2017 The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 4700–4708, Honolulu, July 2017.
- [29] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the 2016 The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 770–778, Las Vegas, NV, USA, June–July 2016.
- [30] S. Jegou, M. Drozdal, D. Vazquez, A. Romero, and Y. Bengio, "The one hundred layers tiramisu: fully convolutional DenseNets for semantic segmentation," in *Proceedings of the 2017 The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 11–19, Honolulu, HI, USA, July 2017.
- [31] G. Ian, Y. Bengio, and C. Aaron, *Deep Learning*, MIT Press, Cambridge, MA, USA, 2017.
- [32] X. Li, "A comparison between information transfer function sigmoid and tanh on neural," *Journal of Wuhan University of Technology*, vol. 28, no. 2, pp. 313–314, 2004.
- [33] K. Hara, D. Saito, and H. Shouno, "Analysis of function of rectified linear unit used in deep learning," in *Proceedings of the 2015 International Joint Conference on Neural Networks (IJCNN)*, pp. 1–8, Killarney, Ireland, July 2015.
- [34] Z. H. Zhou, *Machine Learning*, Tsinghua University Press, Beijing, China, 2016.
- [35] N. Srivastava, G. E Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "A simple way to prevent neural networks from overfitting," *Journal of Machine Learning Research*, vol. 15, no. 1, pp. 1929–1958, 2014.
- [36] P. Lutz, "Automatic early stopping using cross validation: quantifying the criteria," *Neural Networks*, vol. 11, no. 4, pp. 761–767, 1998.
- [37] K. Choi, D. Joo, and J. Kim, "Kapre: on-gpu audio pre-processing layers for a quick implementation of deep neural network models with keras," in *Proceedings of the Machine Learning for Music Discovery Workshop at 34th International Conference on Machine Learning*, Sydney, Australia, August 2017.
- [38] T. Dozat, "Incorporating nesterov momentum into Adam," in *Proceedings of the 2016 International Conference on Learning Representations (ICLR)*, San Juan, Puerto Rico, May 2016.
- [39] Z. Shi, Q. Wang, J. Peng et al., "Development of a national VNIR soil-spectral library for soil classification and prediction of organic matter concentrations," *Science China Earth Sciences*, vol. 57, no. 7, pp. 1671–1680, 2014.