Simultaneous Prediction of Soil Properties Using Multi_CNN Model

Ruixue Li [1], Bo Yin [1,2,*], Yanping Cong [1] and Zehua Du [1]

[1] College of Information Science and Engineering, Ocean University of China, Qingdao 266000, China; liruixue@stu.ouc.edu.cn (R.L.); congyp@ouc.edu.cn (Y.C.); duzehua@stu.ouc.edu.cn (Z.D.)
[2] Pilot National Laboratory for Marine Science and Technology, Qingdao 266000, China
[*] Correspondence: ybfirst@126.com

Abstract: Soil nutrient prediction based on near-infrared spectroscopy has become the main research direction for rapid acquisition of soil information. The development of deep learning has greatly improved the prediction accuracy of traditional modeling methods. In view of the low efficiency and low accuracy of current soil prediction models, this paper proposes a soil multi-attribute intelligent prediction method based on convolutional neural networks, by constructing a dual-stream convolutional neural network model Multi_CNN that combines one-dimensional convolution and two-dimensional convolution, the intelligent prediction of soil multi-attribute is realized. The model extracts the characteristics of soil attributes from spectral sequences and spectrograms respectively, and multiple attributes can be predicted simultaneously by feature fusion. The model is based on two different-scale soil near-infrared spectroscopy data sets for multi-attribute prediction. The experimental results show that the R_P^2 of the three attributes of Total Carbon, Total Nitrogen, and Alkaline Nitrogen on the small dataset are 0.94, 0.95, 0.87, respectively, and the R_P^2 of the attributes of Organic Carbon, Nitrogen, and Clay on the LUCAS dataset are, respectively, 0.95, 0.91, 0.83, And compared with traditional regression models and new prediction methods commonly used in soil nutrient prediction, the multi-task model proposed in this paper is more accurate.

Keywords: soil; vis-NIR spectroscopy; deep learning; convolutional neural networks; multi-task learning

1. Introduction

Soil is an important natural resource. The rapid acquisition of soil property content and spatial distribution is of great value and significance to agriculture and global change. However, the collection of soil samples consumes a large amount of cost, so the prediction of soil nutrient content has become a hot topic in soil research. Visible light near infrared (Vis-NIR) spectroscopy analysis, with its unique advantages such as rapid detection, non-destructive, non-polluting, and real-time detection, has extensive research and application foundations in soil nutrient content prediction [1–3]. However, the spectral data is susceptible to interference from stray light, noise, baseline drift and other factors, which affect the modeling effect. Therefore, it is necessary to preprocess the spectral data before modeling to improve the predictive ability and robustness of the model. Due to the complex characteristics of spectral data, although traditional mathematical modeling methods can perform a certain degree of analysis and prediction, its more accurate and more universal prediction process faces technical bottlenecks. With the development of machine learning, many new spectral model regression prediction algorithms have been continuously proposed and applied [4–6]. However, compared with traditional mathematical modeling and machine learning methods, neural network models have higher computational efficiency and stronger modeling capabilities, and can independently extract effective feature structures from complex spectral data for learning. The purpose of this paper is

to establish a soil nutrient spectrum prediction model with higher efficiency, higher robustness and accuracy, which is of great significance for accelerating the advancement of my country's agricultural informatization, improving the level of agricultural scientific management and developing my country's agricultural economy.

Early research found that the soil organic matter can be calculated from the reflectance value of the soil reflectance spectrum, and the response of soil properties can be identified from the spectral characteristics. In 2006, Rossel et al. compared the predictions of various soil concentrations using qualitative analysis values of visible light (VIS) (400 nm–700 nm), near infrared (NIR) (700 nm–2500 nm), and medium infrared (MIR) (2500 nm–5000 nm), demonstrating that soil analysis and soil information can be obtained more effectively using VIS, NIR and MIR [7]. Later, due to the complexity of vis-NIR spectroscopy, a variety of methods were applied to the pretreatment of soil spectra, such as Savitzky-Golay smoothing, standardization, and normalization methods [8,9]. In 2016, Lin et al. used a combined method of S-G smoothing and scattering correction to process soil spectral data to minimize irrelevant and useless information of the spectrum and increase the correlation between the spectrum and the measured value [10]. By choosing the best combination of preprocessing methods to process soil vis-NIR data, not only can the interference factors be eliminated to the greatest extent, but also the complementary relationship between each preprocessing method can be used to improve the prediction accuracy of the network model. In the existing literature, researchers mostly focus on the preprocessing of spectral data, and there are few proposals and improvements of correlation spectral regression models. A high-performance spectral data modeling method can simplify the preprocessing requirements of spectral data and is also the key to ensuring the accuracy of spectral prediction [11]. With the development of regression prediction, more and more linear regression methods are applied to soil nutrient prediction, such as the principal component regression (PCR) of Chang [12] and the partial least square regression (PLSR) method of McCarty [13]. After that, random forest, genetic algorithm, least squares-support vector machines (LS-SVM) and the Cubist method in machine learning are also used to improve model prediction ability [14–17]. Because deep neural networks are good at automatically extracting useful feature representations from large amounts of data, they have obvious advantages over shallow models and linear methods in modeling, and have become a hot spot in machine learning research in recent years. In 2015, Veres et al. applied deep learning technology to soil spectroscopy for the first time, proving the feasibility of the CNN model for evaluating certain characteristics of LUCAS soil data [18]. In 2017, Ruder proposed that the use of multi-task models can reduce the risk of overfitting while improving the efficiency of model training [19]. In 2019, Padarian et al. used the convolutional neural network (CNN) model and multi-task CNN network to predict various soil properties based on the LUCAS data set, verifying the effectiveness of multi-task learning in predicting soil properties, but the proposed deep learning method is only suitable for large-scale spectral data set, the prediction result is poor on a small sample [20]. After that, Ndikumana et al. used the spectral data as time series data and input it into the long and short-term memory network (LSTM) for soil prediction, and finally achieved good results. However, before training the model, the article performs PCA linear dimensionality reduction processing on the data, which may cause the loss of non-linear correlation between samples, resulting in the model not being able to fit the data characteristics well [21].

Aiming at the problems of low efficiency and low accuracy of current soil prediction models, this paper proposes a new multi-task model based on near-infrared spectroscopy soil data to simultaneously predict multiple attributes of soil. Since the spectral data presents a non-linear trend with the change of the spectral wavelength, this paper takes the spectral wavelength as the time axis, and the spectral data is a non-stationary time series signal. First, the spectrum signal through the three pre-processing methods of SG smoothing, multi-scattering correction, and centralization to construct a stable spectrum sequence. and the original spectral data is windowed and Fourier transform is used to generate a spectrogram, and multiple input channels are used to construct a dual-stream Multi_CNN network that simultaneously inputs a spectrum sequence and a spectrogram, and realizes multiple

inputs and multiple outputs of the model by fusing one-dimensional convolution and two-dimensional convolution. In addition, the model has an adaptive input selection function, and independently selects single input and multiple input based on the characteristics of two different scale soil spectral data sets. Due to the small number of samples and short wavelength range of the Small dataset, it only uses a single input to use the one-dimensional convolutional network of the Multi_CNN model for attribute prediction, while LUCAS dataset selects multiple inputs for prediction based on the complete Multi_CNN model. The results show that the evaluation results of single-input network predicting small sample data are better than traditional machine learning algorithms. For large sample data sets, the evaluation results of the Multi_CNN model are better than the existing new models.

The structure of the article in this article is as follows. The second part introduces the two soil sample spectral data and preprocessing methods involved in the article, as well as the multi-input multi-output network Multi_CNN built. The third part compares the two data sets with different scales, and discusses and analyzes the results. The fourth part summarizes the article.

2. Materials and Method

2.1. The Soil Dataset

Deep learning methods require a large number of samples to train the network, but soil samples based on large national or even global data sets take a long time to sample, so local soil spectral data sets are generally small. The purpose of this article is to make the most accurate predictions for data sets of different scales. The study uses two soil vis-NIR spectroscopy data sets of different scales for prediction modeling.

The Small dataset selected in this paper is to obtain 180 soil samples from 19 sampling sites in Qingdao's South District, Shibei District, Laoshan District, Huangdao District, and Jiaozhou City. The sampling points are selected to be consistent in color and vegetation coverage. In the region, keep the soil nutrients at each sampling point relatively uniform, and the soil quality is mainly sandy loam or silt loam. After air drying and sifting the soil samples, DH-2000 (Ocean Optics, Dunedin, FL, USA) was used as the light source to conduct soil nutrient spectrum collection by connecting the QE-65000 (Ocean Optics, Dunedin, FL, USA) spectrometer with Y-type optical fiber. The spectral range of the spectrometer is 200 nm–1100 nm, the sampling interval is set to 1 nm, and the integration time is set to 600 ms. In order to eliminate accidental factors, each soil sample in the study was repeated 5 times, and the average value was used for calculation. The obtained reflectance spectrum is greatly interfered in the beginning and ending period, so only the middle reflectance data from 225 nm to 975 nm is retained, and each sample contains 750-dimensional data. The basic information of each sample is shown in Tables A1–A4. After that, the physical and chemical values of soil nutrients are obtained through laboratory methods, and the content of the test is the concentration values of soil total nitrogen (TN), total carbon (TC) and alkali hydrolyzed nitrogen (AN). Among them, the physical and chemical values of TN and TC were directly measured by Perkin-Elmer 2400 (PerkinElmer, Waltham, MA, USA) carbon and nitrogen analyzer, and the soil AN attribute value is measured by alkaline hydrolysis diffusion method. First, sodium hydroxide is used to hydrolyze the sample to make the available nitrogen alkaline hydrolyze into ammonia state, then it is absorbed with boric acid, then titrated with standard acid, and finally the content of alkaline hydrolysis nitrogen is calculated.

The large-scale soil data set LUCAS is composed of 19,036 soil sample data collected from 23 European countries, including cultivated land, grassland, woodland and other land types. The soil samples are quite different [8]. After the soil samples are also air-dried and sieved, they are measured by a FOSS XDS (Foss, Denmark) near-infrared spectrometer. The spectral characteristics are recorded with a spectral resolution of 0.5 nm to obtain 4200-dimensional data with a wavelength range of 400 nm–2500 nm. The LUCAS dataset tested a variety of soil properties, such as coarse debris content, organic carbon, nitrogen, potassium, phosphorus, pH, clay, and so forth, so this article selected three soil properties from the LUCAS data set, including organic carbon (OC), nitrogen (N) and Clay,

the basic information of the two data sets is shown in Table 1, which shows the diversity and difference of LUCAS sample data. Figure 1 shows the original soil vis-NIR spectra of the two data sets, where the horizontal axis is the wavelength range and the vertical axis is the absorbance of the soil sample.

Table 1. Basic information of soil data set.

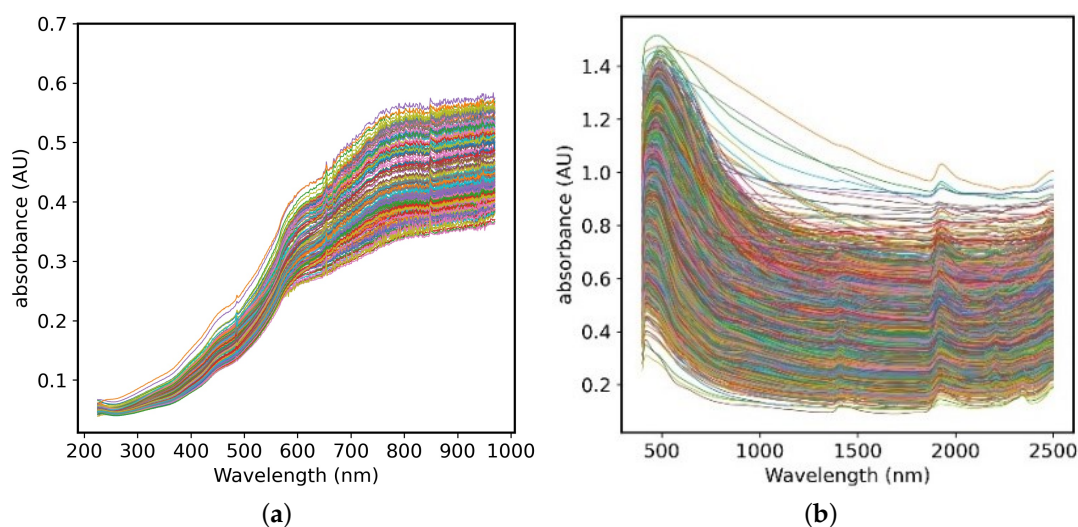|		Minimum	Maximum	Mean	Median	St.Dev.
Small dataset	TC (g·kg^{-1})	1.90	13.40	6.21	4.62	3.42
	TN (g·kg^{-1})	0.21	1.74	0.80	0.56	0.46
	AN (mg·kg^{-1})	17.2	160.00	69.51	62.15	33.18
	OC (g·kg^{-1})	0.00	586.80	50.00	20.80	91.31
LUCAS dataset	N (g·kg^{-1})	0.00	38.60	2.92	1.70	3.76
	Clay (g·kg^{-1})	0.00	79.00	18.88	17.00	13.00

(**a**)　　　　　　　　　　　　　　　(**b**)

Figure 1. Original spectrogram of (**a**) Small dataset (**b**) LUCAS dataset.

2.2. Data Preprocessing

In the collection of soil spectra, due to environmental factors such as temperature and humidity, instrument state, manual operation, and uneven physical state of the soil, the obtained spectra may contain interference factors such as noise, scattering, and baseline drift. Noise concealing spectral characteristics will reduce the accuracy of soil carbon content prediction. Therefore, the use of preprocessing methods to process the original spectrum can reduce the interference that affects the analysis results and establish a more accurate prediction model.

At present, there are many kinds of spectral preprocessing methods. According to the effect of preprocessing, this paper adopts three types of algorithms: smoothing, scattering correction and scale scaling to eliminate high-frequency noise interference in spectral signals. The Savitzky-Golay (S-G) smoothing algorithm is essentially a weighted average method. The smoothing point data is obtained by least square fitting of the data in the smoothing window with the method of polynomial fitting, so as to reduce the loss of spectral information in smoothing by weighting. Multivariate Scattering Correction (MSC) is one of the commonly used algorithms for spectral data preprocessing. It can effectively eliminate the spectral difference caused by different scattering levels and the shift and shift of baseline caused by the influence of scattering between samples, so as to enhance the correlation between spectra and data and improve the signal-to-noise ratio of original absorbance spectrum [22]. In addition, in order to eliminate the influence of the difference between the dimensions and the value range between the standards, it is usually necessary to centralize the data in the regression problem. The absorbance of the average spectrum is removed by calculating the data value of the absorbance

of each spectrum, thereby deducting the value of absolute spectral absorbance. Figure 2a,d are the spectra of the two data sets after S-G smoothing, which can remove the noise peak while retaining the useful spectral information. After the data in Figure 2b,e is processed by MSC, the degree of spectral overlap becomes higher, which reduces the influence of scattering on the original spectrum. Figure 2c,f shows that the data after the centering is scaled based on the origin, eliminating the interference of size difference and different information structure.
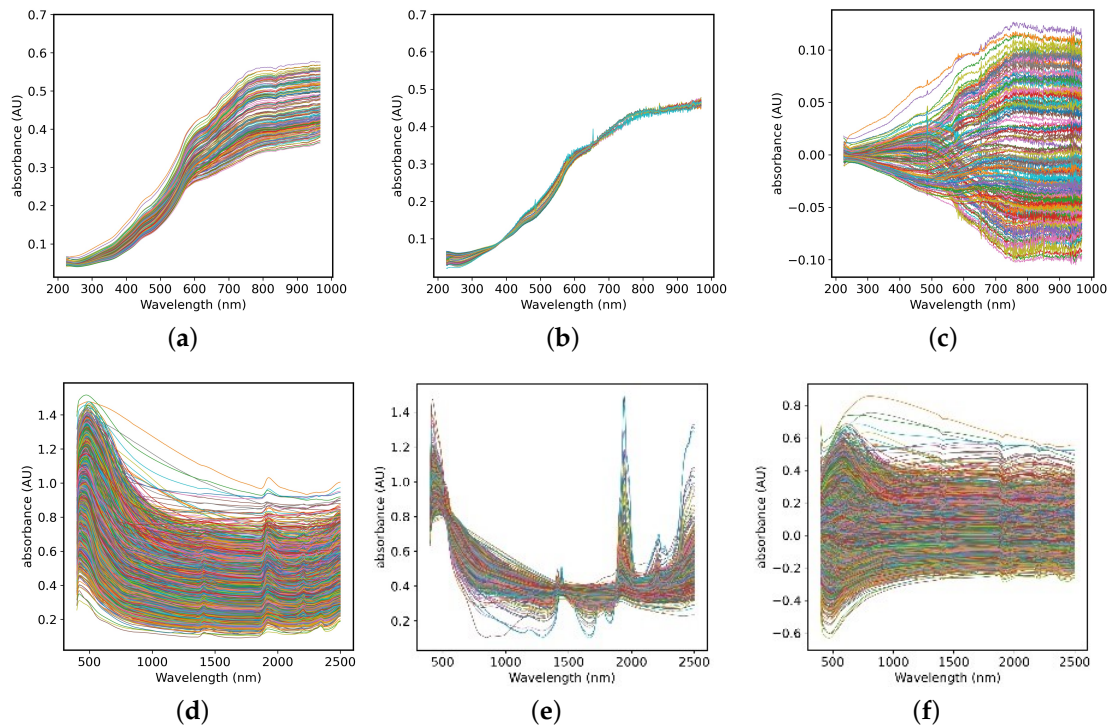

Figure 2. The Spectra of the small dataset (**a**–**c**) and LUCAS dataset (**d**–**f**) preprocessed by Savitzky-Golay (S-G), Multivariate Scattering Correction (MSC) and Centralization methods.

In addition, since this paper uses vis-NIR spectroscopy as time series data, it is possible to decompose signal fragments and do Fourier transform to obtain a spectrum map of soil spectral data. In this experimentr, the Hamming window is used for windowing before the Fourier transformation. The frame length on the Small dataset is determined to be 50, the overlap observation value is 20, and there are 180 pictures in total. The LUCAS dataset contains 19036 pictures, the frame length is determined to be 100, and the overlap observation value is 50. Finally, each sample is represented by a 64×64 spectrogram. Figure 3 shows the spectrum of soil sample data randomly selected from the two data sets. The horizontal axis represents the wavelength range, and the vertical axis represents the frequency.

2.3. The Multiple-Input Multiple-Output Network

Convolutional Neural Networks (CNN) is a nonlinear model, and its unique convolution and pooling structure can extract essential features from complex input information, so it has excellent model characterization capabilities. It is usually composed of multiple convolutional layers, pooling layers and fully connected layers. The convolutional layer is a linear calculation layer that uses a series of convolution kernels to convolve with the input data. It can reduce the number of parameters of the whole model by taking advantage of the localization and positional independence of the features in the input data. The pooling layer is mainly used for feature dimensionality reduction, compressing the number of data and parameters to reduce over-fitting and improve the fault tolerance of the model. The fully connected layer uses neurons to fit the data distribution and improve the model

learning ability. Temporal Convolutional Network (TCN) is an innovative network structure that combines the best practices extracted from convolutional neural networks. It transforms into a model suitable for sequence data by combining one-dimensional full convolution and causal convolution [23]. The residual structure of the TCN model connects causal convolution and dilated convolution, and its structure is shown in Figure 4. It includes two dilated convolutional layers, and WeightNorm and Dropout are added after each layer to achieve regularization. The dilated rate of each dilated convolutional layer increases exponentially with the number of levels, ensuring that the convolution kernel covers all inputs in the effective history information, and also ensuring that the use of deep networks can generate extremely long effective history information.
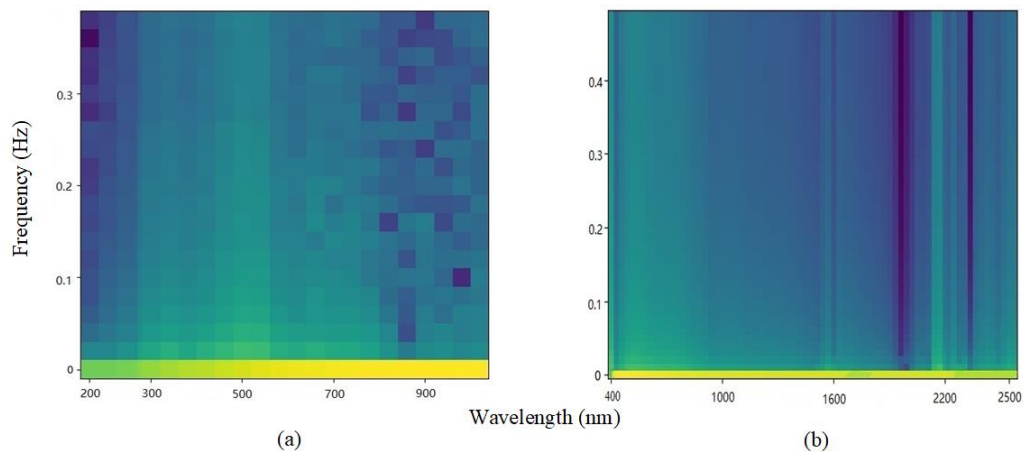
Figure 3. Spectrograms of samples of (**a**) Small dataset (**b**) LUCAS dataset. (The yellower the color represents the greater the spectral frequency density at this wavelength, the bluer the color, the lower the density.)
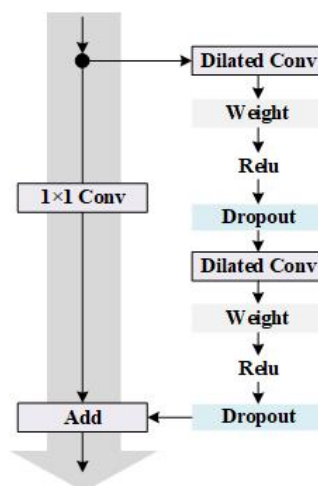
Figure 4. The residual module of Temporal Convolutional Network (TCN).

The multi-task method is a natural choice, and its goal is to obtain predictions for multiple tasks at the same time. The multi-task network mainly includes hidden layers shared between all tasks, as well as maintaining several task-specific output layers, where each output layer is associated with a task. This paper builds a multi-task network structure Multi_CNN, through the fusion of one-dimensional convolution and two-dimensional convolution network, to achieve the model's multiple input and multiple output. The one-dimensional spectral sequence and two-dimensional spectrogram are used as feature data at the same time, which can better fit the spectral features, thereby improving the prediction accuracy of soil properties.

The first input of the model is mainly preprocessed spectral sequence data. By referring to the time convolution network suitable for time series modeling, a one-dimensional convolution network is built to predict multiple soil attributes, and it is named Multi_CNN_1D. The first layer of a one-dimensional convolutional network is a one-dimensional convolutional layer with 64 filters. ReLU is used as the activation function and the convolution kernel weight is normalized to train the network more stably. After that, the BN layer is added, and the maximum pooling layer is used for the down-sampling operation. The third layer is also the convolutional layer with 128 filters. Then the residual module is constructed, which includes two dilated convolutional layers with dilated rate of 2 and 4, and a one-dimensional convolutional layer with exactly the same parameters. Then the fully connected layer connects all the outputs of the previous layer to all the inputs of the next layer and performs information integration. The spectrogram is trained by using a three-layer two-dimensional convolutional layer and adding a pooling layer in the middle, which can reduce the parameter dimension and prevent network overfitting. Then, the feature data type extracted by the dual-stream CNN is converted through the Flatten layer, and finally the prediction results of the three attributes of the soil are output through the three-branch fully connected layer. The network structure is shown in Figure 5, and the specific parameters of the model are shown in Table 2.
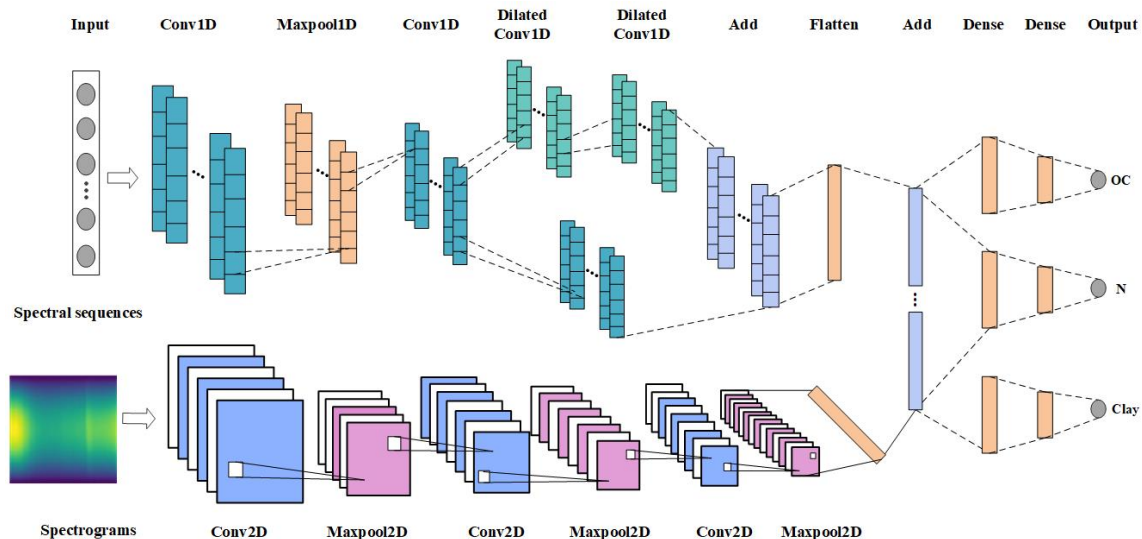
Figure 5. The Multi_CNN network structure.

Table 2. Multi_CNN network specific parameter settings.

Layer	Kernel Size	Filters	Layer	Kernel Size	Filters
Conv1D	3	64	Conv2D	5×5	64
Maxpool1D	5	-	Maxpool2D	2×2	-
Conv1D	3	128	Conv2D	3×3	128
Conv1D	3, dr = 2	64	Maxpool2D	2×2	-
Conv1D	3, dr = 4	64	Conv2D	3×3	256
Conv1D	3	64	Maxpool2D	2×2	-
FC1	-	128			
FC2	-	64			
FC3	-	1			

3. Results and Discussion

The regression fit and accuracy of the network model to the predicted samples are the most important aspects to measure the performance of the model, reflecting the model's ability to predict

unknown samples after training. In order to verify the effectiveness of the multi-task network proposed in this paper, the performance of the model was investigated from the two aspects of regression fitting degree and prediction accuracy. The most unified and objective evaluation criteria were adopted, including determination coefficient (R^2), modeling root mean square error ($RMSEC$), prediction root mean square error ($RMSEP$) and prediction relative analysis error (RPD). R^2 reflects the closeness between the measured value and the predicted value. $RMSEC$ and $RMSEP$ respectively reflect the degree of deviation between the actual measured value and the predicted value in the prediction of the training set and the test set. RPD reflects the predictive power of the model built, The larger the R^2 and RPD obtained at the end, the smaller the $RMSE$, indicating that the performance of the prediction model is better, and vice versa.

3.1. Comparison of Pretreatment Methods

Early experiments proved that a suitable preprocessing method can make the model iterate at a faster convergence rate and improve computational efficiency. In this paper, three single methods of the Savitzky-Golay smoothing, multivariate scattering correction and centralization and four combinations of three methods are used to process the spectrum sequence. Due to the complexity of the spectrum data, the processed spectrum is more messy. In order to see the effect comparison more intuitively, three representative samples are selected in the Small dataset for drawing, as shown in Figure 6. Among them, Figure 6a,b are the spectra of S-G smoothing combined with MSC and centralization, Figure 6c is the spectrum of MSC and centralization, and Figure 6d is the spectrum of S-G smoothing, MSC and centralization.
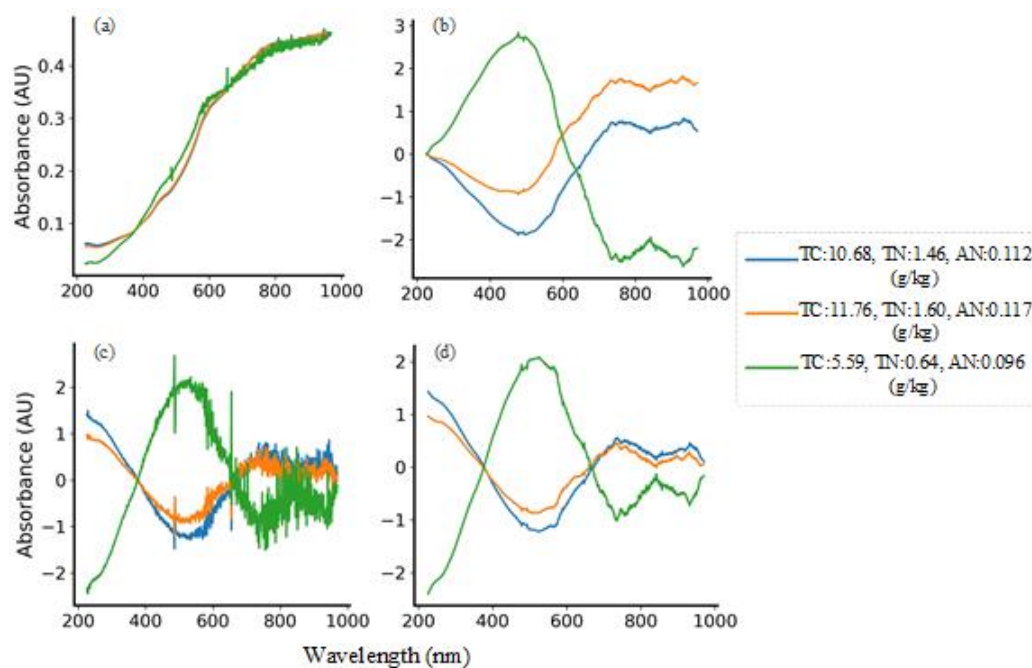
Figure 6. Preprocessed spectra of (**a**) S-G+MSC (**b**) S-G+Centralization (**c**) MSC+Centralization (**d**) S-G+MSC+Centralization.

After that, the Small dataset and LUCAS dataset processed by 3 single preprocessing methods and 4 combinations are input into the Multi_CNN_1D network suitable for time series data modeling for training, and the laboratory measured soil attribute value as true value labels in the network, as the training sample according to the batch sent to the input of network framework. Finally, the model test set evaluation effect diagram in Figure 7 is obtained, where the vertical axis represents R^2. The closer to 1, the higher the accuracy of prediction. It can be seen from Figure 7 that the three soil attributes in the two data sets have the highest R^2 after S-G smoothing, MSC and centralization combined

processing, which means that the gap between the measured value and the predicted value is the smallest, which proves that the S-G smoothing, MSC and centralized combination methods are the best preprocessing methods for processing the spectral sequences of the two data sets.

Figure 7. Comparison diagram of preprocessing effect on test set. (S-G refers to Savitzky-Golay smoothing algorithm, MSC refers to Multivariate Correction, Cent refers to Centralization methods.)

3.2. Experimental Comparison on Small Dataset

Under the condition of small sample of spectral data and selection of the best preprocessing method, the proposed Multi_CNN_1D model is compared with other methods. In this paper, linear regression 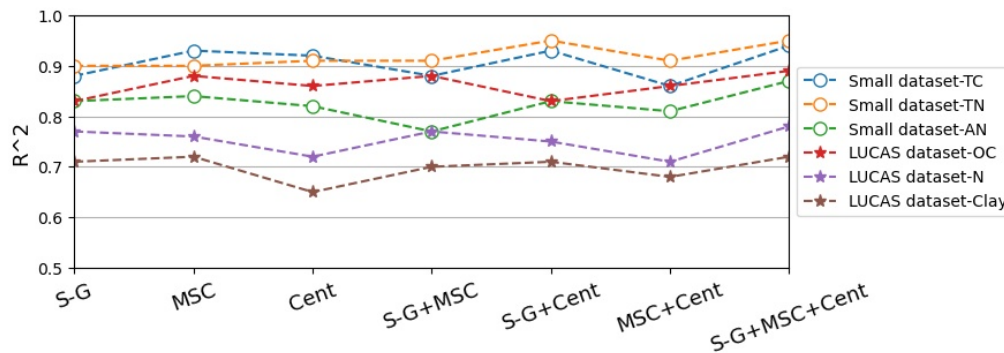method is selected as the model comparison method, including Partial Least Square Regression (PLSR), Random Forest Regression (RFR) and Gradient Boosting Decision Tree (GBR). Considering the problem of too few samples in the small data set, the verification set is not split. Instead, the data set is randomly divided into 125 samples in the training set and 55 samples in the test set according to the ratio of 7:3, and then the Multi_CNN_1D model is input for training. Finally, the prediction sequence corresponding to each soil attribute is obtained. By calculating the deviation between the real value of the training set and the corresponding predicted value of the three soil properties TC, TN and AN, and the real value of the test set and the corresponding predicted value, the training set fitting accuracy of the model and the prediction accuracy of the test set are given quantitatively. Finally, the degree of the advantages and disadvantages of each method is evaluated. The results are shown in Table 3. Where R_C^2 and $RMSEC$ are the results of the training set, R_P^2, $RMSEP$ and RPD are the results of the test set.

Table 3. Comparison of evaluation indexes of each model based on Small dataset.

Attributes		PLSR	RFR	GBR	Multi_CNN_1D
TC	R_C^2	0.96	0.97	0.97	0.99
	$RMSEC$	0.74	0.50	0.57	0.23
	R_P^2	0.89	0.81	0.73	0.94
	$RMSEP$	1.04	1.54	1.79	0.80
	RPD	3.01	2.30	1.95	4.23
TN	R_C^2	0.98	0.96	0.97	0.99
	$RMSEC$	0.07	0.10	0.09	0.04
	R_P^2	0.93	0.82	0.90	0.95
	$RMSEP$	0.11	0.20	0.15	0.09
	RPD	3.80	2.34	3.20	4.71
AN	R_C^2	0.99	0.96	0.89	0.95
	$RMSEC$	3.59	5.98	11.86	7.36
	R_P^2	0.81	0.76	0.83	0.87
	$RMSEP$	14.26	16.84	12.83	11.62
	RPD	2.31	2.04	2.43	2.76

Figure 8 is a box plot of model evaluations obtained by performing multiple experiments on the test set using different regression models, where the solid line in the middle represents the median of the model evaluation value R^2. It can be seen from the Table 3 and Figure 8 that the PLSR method performs best in the linear regression method, and the evaluation results of the Multi_CNN_1D model R^2 and RPD are better than other linear methods. For the same soil attribute TC, the $RMSEP$ error of the Multi_CNN_1D network were 0.24, 0.74 and 0.99 lower than those of traditional PLSR, RFR and GBR networks. Compared with TN, the $RMSEP$ error is reduced by 0.02, 0.11, 0.06. Although R^2_P and RPD of AN were 0.87 and 2.76, which were slightly lower than those of the other two attributes due to the large difference in AN attribute value, the $RMSEP$ error of AN was reduced by 2.64, 5.22 and 1.21 compared with other methods. This indicates that the accuracy of the deep learning method in predicting soil properties based on vis-NIR spectral data is better than the general linear regression method. In addition, when inputting Small dataset samples into the Multi_CNN network for modeling training, the network will have over-fitting problems. This is because the small data set has too few samples and the wavelength range is short, which is more suitable for simple one-dimensional convolutional networks, not multi-input networks. Therefore, for Small dataset, the prediction result of the single-input network is better than that of the dual-stream network.
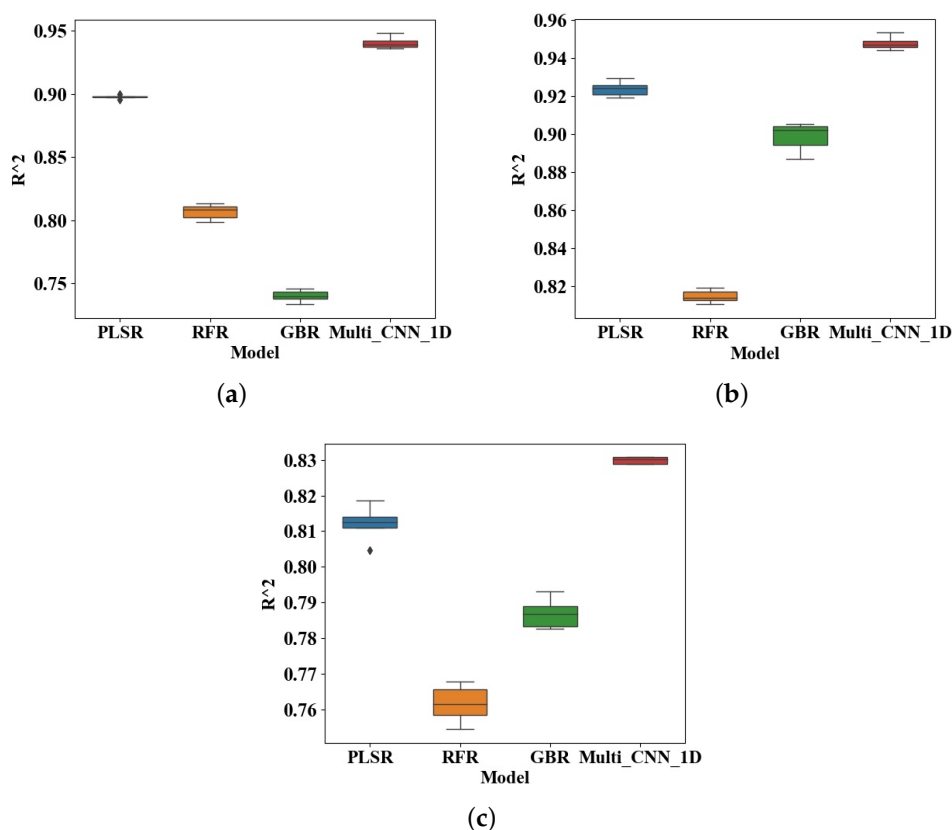
Figure 8. R^2 comparison box diagram of predicted values of (**a**) Total Carbon (**b**) Total Nitrogen (**c**) Alkaline Nitrogen.

Figure 9 shows the comparison of the predicted TC, TN, and AN content values of the test set soil samples with the actual values obtained by laboratory method analysis. It can be clearly seen that the scattered points are closely and evenly distributed on both sides of the regression line, and the predicted values of the three soil properties are positively correlated with the actual values, which proves that the multi-task network proposed in this paper is effective in predicting soil properties with a single input.

(a)

(b)

(c)
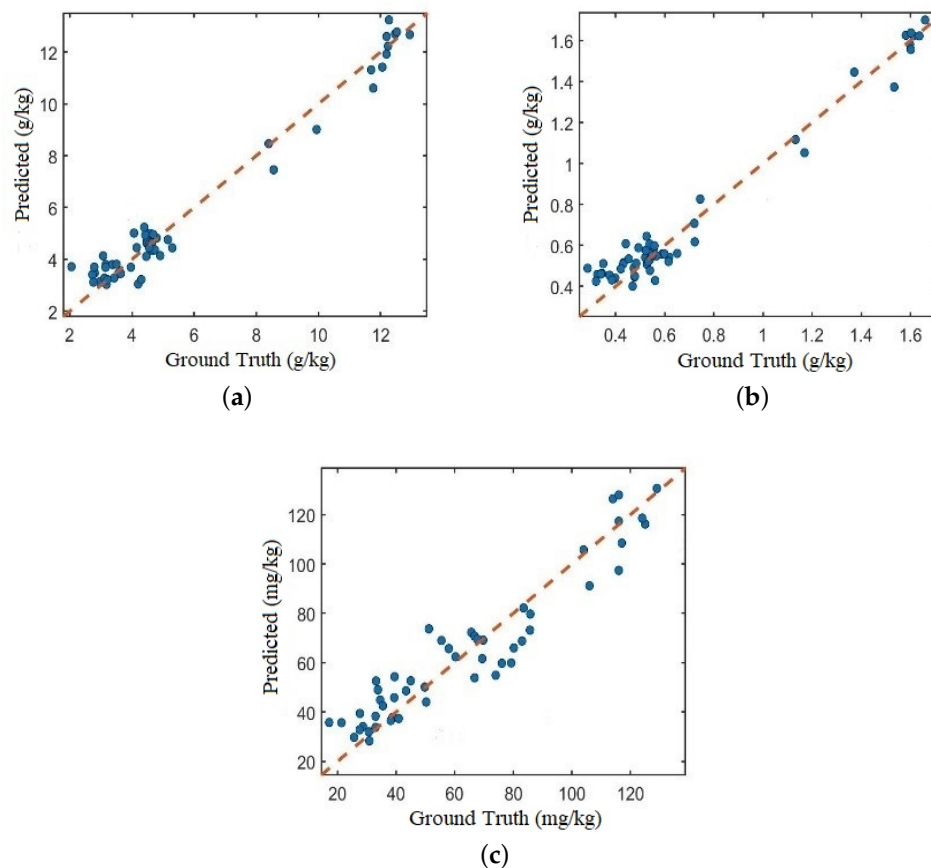
Figure 9. Actual vs Predicted values of Proposed framework (**a**) Total Carbon (**b**) Total Nitrogen (**c**) Alkaline Nitrogen.

3.3. Experimental Comparison on LUCAS Dataset

The sampling of the LUCAS data set spans the European continent, and the soil samples are very diverse. In this paper, the 19,036 samples of the LUCAS data set are shuffled and divided into training set, validation set and test set according to the ratio of 6:2:2. The number of samples in the training set is 11,420, and the number of samples in the validation set and test set is 3808. Figure 10 shows the scatter plot of the predicted and actual values of the soil samples on the test set using the Multi_CNN model. It can be seen that the scatter values of the three attributes of OC, N, and Clay are evenly distributed on both sides of the regression line. The results show that for large samples of soil spectral data, the proposed Multi_CNN model can effectively extract the characteristic information of soil vis-NIR spectral data. It has high regression fitting and regression accuracy for training samples, and has better learning ability, and can achieve maximum training through existing data, and at the same time accurately approximate the actual measured value of the training sample.

In order to more intuitively reflect the effectiveness of the proposed network model, based on the LUCAS data set, this paper compares the two proposed models with the existing advanced models, including the CNN and multi-task CNN models proposed by Padarian [20] and Ndikumana [21] LSTM model and traditional PLSR model. Table 4 shows the comparison of the evaluation results of each model.

Table 4. Comparison of model evaluations on the LUCAS dataset.

Attributes		Multi_CNN_1D	Multi_CNN	CNN [20]	CNN_multi [20]	LSTM [21]	PLSR
OC	R_P^2	0.89	0.95	0.88	0.69	0.94	0.54
	$RMSEP$	29.49	22.69	32.14	31.86	23.25	68.12
N	R_P^2	0.78	0.91	0.83	0.60	0.91	0.55
	$RMSEP$	1.76	1.09	1.54	1.59	1.15	1.73
Clay	R_{P}^2,	0.72	0.83	0.70	0.68	0.80	0.50
	$RMSEP$	7.17	5.53	7.55	7.78	5.95	8.93

The results show that the R_P^2 of the proposed Multi_CNN and LSTM network [21] in predicting soil N attributes is 0.91, but $RMSEP$ is relatively reduced by 0.06, and the results are better than the LSTM network [21] in predicting OC and Clay. Compared with CNN_multi [20], the prediction results R_P^2 and $RMSEP$ of the three attributes of OC, N and Clay proposed in this paper are improved a lot. This is because this paper uses spectral data as time series data to learn the short-term and long-term dependence of sample data, and fuse one-dimensional convolution and two-dimensional convolution makes the feature fit better and the prediction accuracy more accurate. In addition, the prediction effect of the three attributes of the Multi_CNN network is also higher than that of the proposed single-input Multi_CNN_1D network. Therefore, the self-adaptive Multi_CNN network built in this paper can obtain better prediction results for different scale data sets.

(a)

(b)

(c)
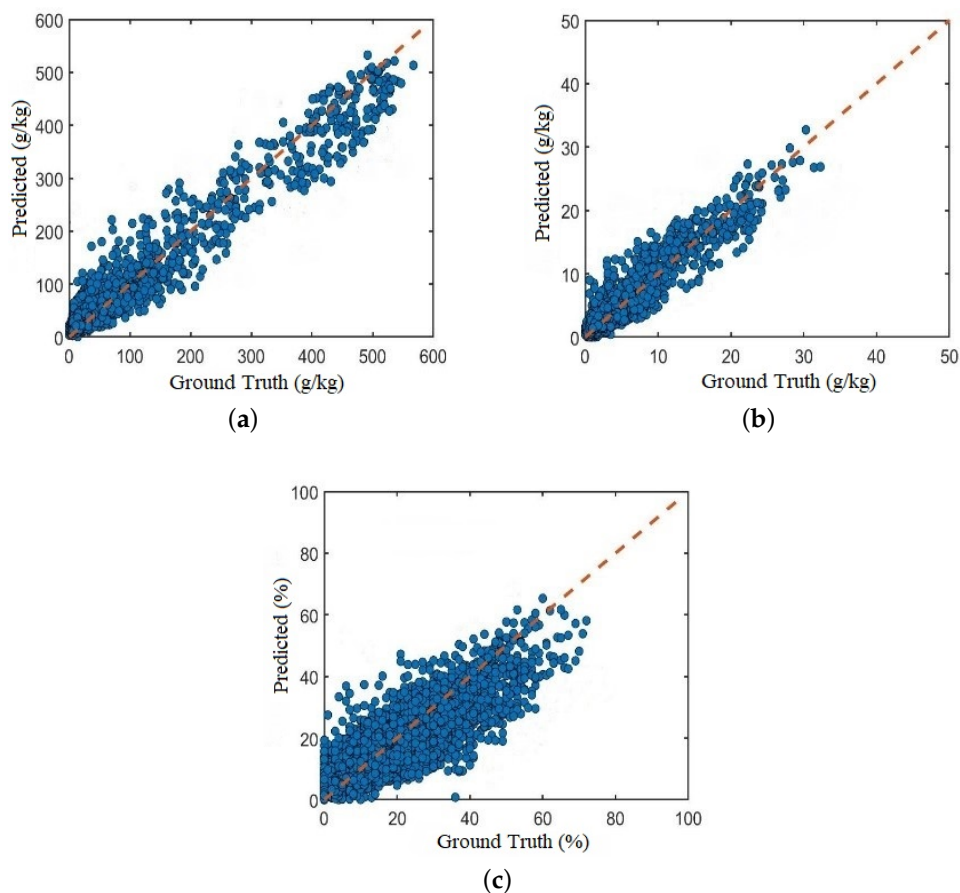
Figure 10. Actual vs Predicted values of Multi_CNN model (**a**) Organic Carbon (**b**) Nitrogen (**c**) Clay.

In order to further verify the effectiveness of the proposed algorithm, this paper selects Qingdao soil spectral data measured in different periods, which contains 500 data samples, and the selected soil attributes are also TC, TN, and AN. The small samples are preprocessed and input into the adaptive network Multi_CNN. Finally, the modulus evaluation parameters R_P^2 of the three attributes are 0.91,

0.98 and 0.95, and the *RMSEP* is 0.17, 0.02, and 2.21. The three soil attributes are shown in the Figure 11. A scatter plot of the predicted value and the actual value of the attribute obtained by laboratory method analysis. It can be seen that the scattered points are evenly distributed on both sides of the regression line, which proves that the Multi_CNN network has high predictive ability and generalization ability.
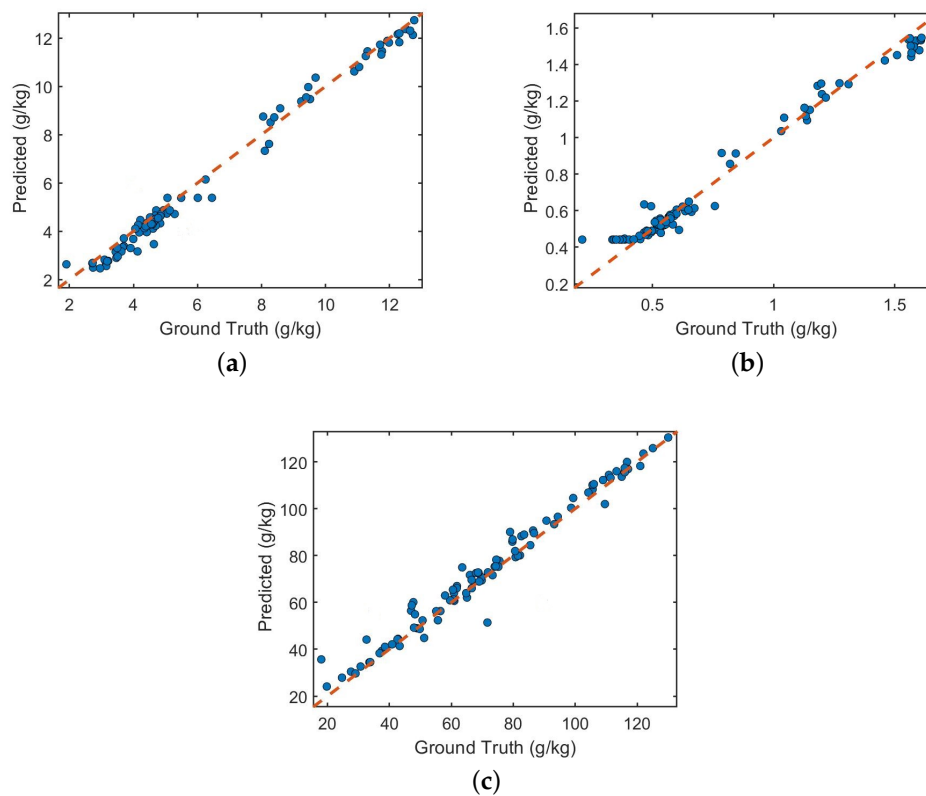
Figure 11. Actual vs. Predicted values of Proposed framework (**a**) Total Carbon (**b**) Total Nitrogen (**c**) Alkaline Nitrogen.

4. Conclusions

This paper proposes a new intelligent network architecture for simultaneous soil multi-attribute prediction in the same task network. The proposed framework is based on soil vis-NIR spectral signals, and a dual-stream convolutional network is built to predict various characteristics of soil. The spectral signal was processed by the combination of pretreatment and the conversion of the original data to the spectral map, which made the soil characteristic information extracted by the network more detailed. In addition, this paper discusses the predictive ability of soil data set networks based on different scales. Due to the small number of samples and short wavelength range of the Small dataset, it is more suitable for one-dimensional convolution input, but not for the complex network structure with multiple inputs and outputs. However, for the large-scale LUCAS dataset, the multi-input and multi-output network significantly improves the prediction accuracy, and the results are better than the existing methods. This paper fully proves the feasibility and accuracy of multi-task network in soil attribute prediction.

Author Contributions: The manuscript was written through contributions of all authors, and all authors contributed equally. Conceptualization, R.L. and B.Y.; methodology, R.L. and B.Y.; validation, Y.C.; visualization, R.L. and Z.D.; supervision, Y.C. and Z.D.; writing—original draft, R.L.; writing—review and editing, R.L. and B.Y. All authors have read and agreed to the published version of the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Table A1. Basic information on samples from S1 to S55.

Soil Sample	TC (g·kg^{-1})	TN (g·kg^{-1})	AN (mg·kg^{-1})	Minimum	Maximum	Mean	Median
S1	2.90	0.33	32	5.20	55.16	32.18	37.02
S2	3.99	0.53	39.5	4.89	52.26	30.43	35.04
S3	2.29	0.27	20.6	5.74	57.38	34.28	39.98
S4	5.04	0.54	64.6	4.68	49.23	28.38	32.13
S5	4.64	0.53	90.4	5.20	52.12	30.65	35.17
S6	3.18	0.38	93.2	5.01	55.68	32.50	37.36
S7	3.07	0.45	33.2	4.93	54.29	31.34	35.94
S8	3.10	0.42	52.7	5.02	56.59	32.84	37.77
S9	3.50	0.50	34.7	4.57	51.58	29.45	33.50
S10	1.90	0.21	18	4.93	56.00	32.76	38.12
S11	2.94	0.32	31.3	4.90	50.68	29.34	33.42
S12	3.15	0.34	43.4	5.38	53.90	31.88	36.57
S13	3.16	0.48	34.6	4.91	53.69	30.95	35.30
S14	3.41	0.40	30.3	4.99	55.59	31.95	36.30
S15	4.37	0.47	38.4	6.31	58.38	35.37	41.19
S16	2.98	0.33	35.3	4.96	54.08	31.00	35.21
S17	2.73	0.35	33.8	5.16	56.83	32.92	37.85
S18	3.16	0.44	37.8	4.70	51.18	29.49	33.71
S19	3.70	0.46	42.6	4.59	52.60	29.64	33.42
S20	3.14	0.40	36	5.01	54.16	31.41	35.85
S21	4.66	0.62	50.3	5.93	54.07	32.44	37.60
S22	2.83	0.43	32.6	5.12	56.00	32.87	38.00
S23	3.61	0.44	38.5	5.07	50.55	30.20	35.48
S24	3.84	0.56	43.6	4.87	51.30	30.05	34.71
S25	2.79	0.43	27.7	4.99	53.59	31.60	36.65
S26	3.37	0.43	35.5	5.13	52.99	31.39	36.52
S27	3.03	0.47	35.3	4.65	49.66	28.88	33.18
S28	3.96	0.52	38.7	5.05	56.36	32.69	37.41
S29	3.00	0.47	32.3	5.06	52.85	31.49	36.68
S30	3.58	0.56	60.6	4.87	53.40	31.19	35.86
S31	4.00	0.53	46.4	4.82	50.56	29.73	34.56
S32	4.01	0.57	47	5.18	55.00	32.57	37.75
S33	2.96	0.34	33	4.98	56.31	33.30	38.50
S34	3.56	0.47	38.8	5.02	55.82	32.81	37.87
S35	2.96	0.36	24.7	4.93	55.55	32.60	37.37
S36	4.07	0.55	35.4	4.82	54.93	32.01	36.74
S37	3.51	0.42	30.7	4.74	55.52	32.28	36.73
S38	2.99	0.41	25.5	4.85	54.91	32.42	37.56
S39	3.12	0.38	29.4	5.27	56.88	33.90	39.23
S40	3.96	0.54	38.2	4.93	53.43	31.38	36.14
S41	2.74	0.33	19.8	4.76	55.77	32.53	37.22
S42	3.43	0.47	25.7	4.59	52.40	30.68	35.32
S43	3.64	0.48	33	4.55	51.33	30.04	34.41
S44	3.52	0.54	32.2	4.62	52.08	30.57	35.03
S45	2.76	0.31	23.2	4.57	53.49	30.98	35.27
S46	4.30	0.48	40.9	4.42	49.98	29.37	33.67
S47	3.11	0.40	28.7	4.67	54.10	32.04	37.17
S48	3.63	0.39	32.1	4.74	50.69	30.17	34.77
S49	2.05	0.29	17.2	4.53	52.20	30.92	35.76
S50	3.59	0.40	34.1	4.45	51.93	30.59	35.36
S51	3.20	0.38	32.2	4.39	50.65	29.52	33.69
S52	4.07	0.47	47	5.71	57.35	35.39	41.25
S53	2.80	0.33	30.9	4.54	53.88	31.33	35.79
S54	3.21	0.39	27.6	4.29	51.97	30.28	34.58
S55	3.39	0.47	34.6	4.39	47.43	28.10	32.49

Table A2. Basic information on samples from S56 to S105.

Soil Sample	TC (g·kg^{-1})	TN (g·kg^{-1})	AN (mg·kg^{-1})	Minimum	Maximum	Mean	Median
S56	3.20	0.40	36.8	4.25	51.35	29.92	34.34
S57	4.59	0.59	45.3	4.33	52.67	31.06	35.93
S58	2.79	0.43	30.4	4.49	56.21	32.82	37.72
S59	3.18	0.36	30.2	4.65	56.45	33.27	38.14
S60	2.76	0.32	21.4	4.18	51.39	29.72	33.83
S61	5.49	0.57	80.6	4.65	41.94	25.24	29.81
S62	5.57	0.61	64.8	5.24	45.97	27.99	33.04
S63	4.27	0.52	47.4	4.65	36.73	22.33	26.22
S64	5.15	0.62	80.2	4.50	40.33	24.34	28.79
S65	4.54	0.57	74	5.06	41.88	25.56	30.03
S66	5.08	0.59	91	4.83	42.41	26.03	31.08
S67	4.56	0.56	73.1	4.66	41.23	25.12	29.94
S68	4.58	0.57	74.6	4.50	40.36	24.50	29.05
S69	4.68	0.52	67.7	4.88	43.72	26.51	31.30
S70	4.84	0.59	68.6	5.04	43.56	26.77	31.79
S71	4.25	0.48	61.9	4.95	40.68	24.75	29.39
S72	4.31	0.51	55.7	4.72	39.26	23.69	27.91
S73	4.73	0.60	83	4.80	41.69	25.33	30.10
S74	4.42	0.51	59.6	4.98	41.96	25.51	30.18
S75	4.38	0.52	60.9	4.94	43.08	26.14	30.83
S76	5.43	0.69	92.9	4.65	42.18	25.52	30.10
S77	5.30	0.65	85.7	5.01	41.97	25.96	30.84
S78	4.15	0.53	58	4.77	41.15	24.81	29.42
S79	4.74	0.61	67.4	5.22	42.88	26.26	31.20
S80	4.48	0.52	60.3	4.91	42.52	25.83	30.47
S81	4.31	0.54	51.4	4.93	39.94	24.17	28.33
S82	4.14	0.51	69	5.06	44.39	27.02	32.22
S83	4.29	0.53	79.3	4.95	42.10	25.98	30.92
S84	4.94	0.55	74.5	4.69	41.62	25.01	29.52
S85	5.13	0.65	81.5	4.41	41.22	24.69	29.26
S86	4.48	0.59	69.7	4.92	42.57	25.91	30.68
S87	4.91	0.54	65.7	5.11	43.86	27.02	32.20
S88	4.49	0.53	62.4	4.91	43.47	26.38	31.34
S89	4.46	0.56	76.1	4.82	40.40	24.67	29.24
S90	4.92	0.54	61.4	4.47	40.76	24.38	28.62
S91	4.65	0.53	66.8	4.74	38.86	23.64	27.82
S92	5.66	0.59	60.6	4.53	40.86	24.49	28.84
S93	4.43	0.44	39.4	4.50	38.49	23.04	27.07
S94	4.22	0.48	45.1	4.28	37.14	22.22	26.10
S95	4.32	0.50	61.6	4.78	41.85	25.32	29.81
S96	5.21	0.58	60.6	4.77	42.98	25.87	30.53
S97	4.99	0.55	49.7	4.35	36.58	22.01	25.87
S98	4.39	0.49	45	4.46	38.88	23.18	27.19
S99	4.46	0.52	53.3	4.64	39.50	23.98	28.44
S100	4.79	0.54	49.8	4.53	38.32	22.85	26.76
S101	4.96	0.54	54.5	4.43	39.78	23.67	27.81
S102	4.48	0.53	56.6	4.45	40.27	24.03	28.36
S103	5.03	0.55	48.8	4.48	38.69	23.26	27.34
S104	4.80	0.58	47.7	4.40	38.79	23.11	27.11
S105	4.38	0.67	71.8	4.66	41.52	24.97	29.43

Table A3. Basic information on samples from S106 to S150.

Soil Sample	TC (g·kg^{-1})	TN (g·kg^{-1})	AN (mg·kg^{-1})	Minimum	Maximum	Mean	Median
S106	4.06	0.72	66.8	4.76	43.43	26.13	30.94
S107	4.69	0.58	64	4.76	43.24	25.96	30.51
S108	4.22	0.51	59.1	4.56	41.57	24.91	29.33
S109	4.43	0.52	63.3	4.31	40.09	24.02	28.22
S110	4.67	0.60	71	4.57	41.64	25.23	29.78
S111	4.18	0.46	56.6	4.59	39.27	23.89	28.09
S112	4.55	0.55	69.4	4.57	42.13	25.22	29.76
S113	4.15	0.53	61.7	4.42	40.93	24.82	29.41
S114	4.71	0.52	49.8	4.69	40.64	24.79	29.36
S115	4.61	0.55	69.8	4.41	39.32	23.77	27.82
S116	4.65	0.49	56.9	4.12	38.17	22.49	26.25
S117	4.58	0.56	55.5	3.89	38.82	22.66	26.40
S118	5.14	0.60	67.9	4.55	46.39	27.65	32.65
S119	4.84	0.53	50.7	3.90	38.10	22.23	25.98
S120	5.60	0.64	95.7	4.80	44.19	27.46	32.81
S121	10.12	1.37	108	4.27	44.00	25.32	29.45
S122	11.27	1.54	120	4.34	44.77	25.91	30.04
S123	11.70	1.61	130	4.58	47.10	27.56	32.23
S124	12.57	1.55	131	4.18	41.94	23.84	27.22
S125	7.31	1.02	80.7	4.13	45.00	25.68	29.85
S126	7.26	1.00	96.2	4.30	45.74	26.28	30.81
S127	12.27	1.58	125	4.31	40.96	23.80	27.58
S128	9.52	1.22	90.7	4.10	43.54	24.96	28.90
S129	12.64	1.61	130	4.08	41.27	23.82	27.47
S130	12.52	1.53	136	4.74	44.82	26.32	30.68
S131	10.80	1.25	102	4.41	42.96	24.80	28.52
S132	13.40	1.55	115	4.48	41.44	24.10	27.82
S133	12.30	1.57	115	4.68	49.12	28.70	33.51
S134	12.47	1.63	122	4.30	42.54	24.52	28.47
S135	7.21	1.01	69.8	4.10	45.77	26.08	30.46
S136	12.69	1.63	118	4.73	48.43	28.34	33.07
S137	6.01	0.82	61.8	4.05	48.30	27.21	31.74
S138	12.31	1.56	106	4.45	44.62	26.21	30.81
S139	4.20	0.72	54.5	4.24	51.14	29.16	34.03
S140	12.06	1.37	104	4.51	44.22	25.76	29.88
S141	8.40	1.13	83.5	4.17	45.20	25.70	29.72
S142	10.99	1.49	109	4.37	43.85	25.49	29.86
S142	9.85	1.29	97.3	4.40	42.44	24.64	28.55
S144	11.24	1.50	103	4.32	47.09	26.68	30.71
S145	10.68	1.49	107	4.75	49.05	28.48	33.19
S146	11.98	1.57	121	4.69	46.77	27.60	32.46
S147	9.94	1.32	106	4.53	49.55	28.66	33.51
S148	12.47	1.64	116	4.60	46.26	26.93	31.34
S149	9.28	1.27	96.7	4.02	46.39	26.26	30.65
S150	11.78	1.45	160	4.20	43.71	25.00	28.77

Table A4. Basic information on samples from S151 to S180.

Soil Sample	TC (g·kg^{-1})	TN (g·kg^{-1})	AN (mg·kg^{-1})	Minimum	Maximum	Mean	Median
S151	8.28	1.03	79.7	4.18	45.43	25.95	30.16
S152	12.77	1.59	122	4.51	44.94	25.98	30.10
S153	11.61	1.61	118	4.40	42.26	24.43	28.31
S154	11.93	1.61	116	4.22	42.20	24.07	27.85
S155	12.94	1.60	114	4.28	43.31	24.95	28.83
S156	11.10	1.47	133	4.33	46.08	26.15	30.20
S157	12.94	1.74	136	4.56	47.11	27.32	31.94
S158	12.51	1.65	125	4.42	43.79	25.11	28.89
S159	5.18	0.78	61.1	4.36	48.26	27.47	31.84
S160	10.61	1.50	109	4.03	45.55	25.92	30.24
S161	12.27	1.65	125	4.37	42.85	24.86	28.93
S162	11.99	1.51	111	4.28	42.76	24.63	28.64
S163	12.24	1.60	124	4.34	42.79	24.67	28.52
S164	12.51	1.72	122	4.37	45.00	26.04	30.34
S165	9.95	1.35	97.7	4.37	44.02	25.22	29.39
S166	10.68	1.46	112	4.52	47.42	27.20	31.54
S167	11.12	1.51	113	4.17	42.86	24.45	28.27
S168	9.59	1.32	99.4	4.31	45.04	25.90	30.39
S169	10.03	1.41	100	4.11	41.29	23.71	27.46
S170	8.05	1.13	79.8	4.18	46.09	26.39	30.67
S171	3.79	0.74	51.2	3.97	49.46	27.71	32.25
S172	11.08	1.52	116	4.51	45.33	26.04	30.09
S173	9.22	1.18	94.1	4.02	42.33	24.23	28.25
S174	7.84	1.03	79.3	4.37	49.09	27.95	32.50
S175	12.19	1.61	116	4.41	43.37	24.93	28.97
S176	11.76	1.60	117	4.62	48.21	28.05	32.87
S177	11.70	1.52	116	4.25	39.88	23.12	26.70
S178	8.56	1.17	85.8	4.24	45.89	26.50	31.09
S179	12.19	1.66	129	4.64	46.80	27.23	31.75
S180	11.26	1.61	116	4.41	44.57	25.54	29.50

References

1. Yan, L.; Escobar, M.S.; Kaneko, H.; Funatsu, K. Detection of Nonlinearity in Soil Property Prediction Models Based on Near-infrared Spectroscopy. *Chemom. Intell. Lab. Syst.* **2017**, *167*, 139–151. [CrossRef]
2. Schimann, H.; Joffre, R.; Roggy, J.C.; Lensi, R.; Domenach, A.M. Evaluation of the recovery of microbial functions during soil restoration using near-infrared spectroscopy. *Appl. Soil Ecol.* **2007**, *37*, 223–232. [CrossRef]
3. Zhang, Y.; Li, M.; Zheng, L.; Zhao, Y.; Pei, X. Soil nitrogen content forecasting based on real-time NIR spectroscopy. *Comput. Electron. Agric.* **2016**, *124*, 29–36. [CrossRef]
4. Nawar, S.; Buddenbaum, H.; Hill, J.; Kozak, J.; Mouazen, A.M. Estimating the soil clay content and organic matter by means of different calibration methods of vis-NIR diffuse reflectance spectroscopy. *Soil Tillage Res.* **2016**, *155*, 510–522. [CrossRef]
5. Nawar, S.; Mouazen, A.M. Comparison between Random Forests, Artificial Neural Networks and Gradient Boosted Machines Methods of On-Line Vis-NIR Spectroscopy Measurements of Soil Total Nitrogen and Total Carbon. *Sensors* **2017**, *17*, 2428. [CrossRef]
6. Vohland, M.; Besold, J.; Hill, J.; Fründ, H.C. Comparing different multivariate calibration methods for the determination of soil organic carbon pools with visible to near infrared spectroscopy. *Geoderma* **2011**, *166*, 198–205. [CrossRef]
7. Rossel, R.A.V.; Walvoort, D.J.J.; McBratney, A.B.; Janik, L.J.; Skjemstad, J.O. Visible, near infrared, mid infrared or combined diffuse reflectance spectroscopy for simultaneous assessment of various soil properties. *Geoderma* **2006**, *131*, 59–75. [CrossRef]
8. Stevens, A.; Nocita, M.; Tóth, G.; Montanarella, L.; Van, W.B. Prediction of Soil Organic Carbon at the European Scale by Visible and Near InfraRed Reflectance Spectroscopy. *PLoS ONE* **2013**, *8*, e66409. [CrossRef] [PubMed]

9.   Vasques, G.M.; Grunwald, S.; Sickman, J.O. Modeling of Soil Organic Carbon Fractions Using Visible–Near-Infrared Spectroscopy. *Soil Sci. Soc. Am. J.* **2009**, *73*, 176–184. [CrossRef]

10.   Lin, Z.D.; Wang, Y.B.; Wang, R.J.; Wang, L.S.; Lu, C.P.; Zhang, Z.Y.; Song, L.T.; Liu, Y. Improvements of Vis-NIRS Model in The Prediction of Soil Organic Matter Content Using Wavelength Optimization. *J. Appl. Spectrosc.* **2017**, *84*, 529–534. [CrossRef]

11.   Liu, S.; Shen, H.; Chen, S.; Zhao, X.; Biswas, A.; Jia, X.; Shi, Z.; Fang, J. Estimating forest soil organic carbon content using vis-NIR spectroscopy: Implications for large-scale soil carbon spectroscopic assessment. *Geoderma* **2019**, *348*, 37–44. [CrossRef]

12.   Chang, C.W.; Laird, D.; Mausbach, M.; Hurburgh, C. Near-Infrared Reflectance Spectroscopy–Principal Components Regression Analyses of Soil Properties. *Soil Sci. Soc. Am. J.* **2001**, *65*, 480–490. [CrossRef]

13.   Mccarty, G.W.; Reeves, J.B.; Reeves, V.B.; Follett, R.F.; Kimble, J.M. Mid-Infrared and Near-Infrared Diffuse Reflectance Spectroscopy for Soil Carbon Measurement. *Soil Sci. Soc. Am. J.* **2002**, *66*, 640–646. [CrossRef]

14.   Rossel, R.A.V.; Behrens, T. Using data mining to model and interpret soil diffuse reflectance spectra. *Geoderma* **2010**, *158*, 46–54. [CrossRef]

15.   Hosseini, M.; Agereh, S.R.; Khaledian, Y.; Zoghalchali, H.J.; Brevik, E.C.; Naeini, S.A.R.M. Comparison of multiple statistical techniques to predict soil phosphorus. *Appl. Soil Ecol.* **2017**, *114*, 123–131. [CrossRef]

16.   Morellos, A.; Pantazi, X.E.; Moshou, D.; Alexandridis, T.; Whetton, R.; Tziotzios, G.; Wiebensohn, J.; Bill, R.; Mouazen, A.M. Machine learning based prediction of soil total nitrogen, organic carbon and moisture content by using VIS-NIR spectroscopy. *Biosyst. Eng.* **2016**, *152*, 104–116. [CrossRef]

17.   Yang, M.; Xu, D.; Chen, S.; Li, H.; Shi, Z. Evaluation of Machine Learning Approaches to Predict Soil Organic Matter and pH Using vis-NIR Spectra. *Sensors* **2019**, *19*, 263. [CrossRef]

18.   Veres, M.; Lacey, G.; Taylor, G.W. Deep Learning Architectures for Soil Property Prediction. In Proceedings of the 2015 12th Conference on Computer and Robot Vision, Halifax, NS, Canada, 3–5 June 2015; pp. 8–15. [CrossRef]

19.   Ruder, S. An Overview of Multi-Task Learning in Deep Neural Networks. *arXiv* **2017**, arXiv:1706.05098.

20.   Padarian, J.; Minasny, B.; McBratney, A.B. Using deep learning to predict soil properties from regional spectral data. *Geoderma Reg.* **2019**, *16*, e00198. [CrossRef]

21.   Ndikumana, E.; Minh, D.H.T.; Baghdadi, N.; Courault, D.; Hossard, L. Deep Recurrent Neural Network for Agricultural Classification using multitemporal SAR Sentinel-1 for Camargue, France. *Remote Sens.* **2018**, *10*, 1217. [CrossRef]

22.   Sum, S.T. Spectral Signal Correction for Multivariate Calibration. Ph.D. Thesis, University of Delaware, Newark, DE, USA, 1998.

23.   Bai, S.; Kolter, J.Z.; Koltun, V. An Empirical Evaluation of Generic Convolutional and Recurrent Networks for Sequence Modeling. *arXiv* **2018**, arXiv:1803.01271.

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.