Article

Non-Destructive Testing of Moisture and Nitrogen Content in *Pinus Massoniana* Seedling Leaves with NIRS Based on MS-SC-CNN

Zhuo Huang, Tingting Zhu , Zhenye Li and Chao Ni

College of Mechanical and Electronic Engineering, Nanjing Forestry University, Nanjing 210037, China; zhuoh@njfu.edu.cn (Z.H.); zhenye@njfu.edu.cn (Z.L.); chaoni@njfu.edu.cn (C.N.)

* Correspondence: tingtingzhu@njfu.edu.cn

Citation: Huang, Z.; Zhu, T.; Li, Z.; Ni, C. Non-Destructive Testing of Moisture and Nitrogen Content in *Pinus Massoniana* Seedling Leaves with NIRS Based on MS-SC-CNN. *Appl. Sci.* **2021**, *11*, 2754. <https://doi.org/10.3390/app11062754>

Academic Editor: Pietro Picuno

Received: 1 December 2020

Accepted: 17 March 2021

Published: 19 March 2021

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

Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Abstract: *Pinus massoniana* is a pioneer reforestation tree species in China. It is crucial to evaluate the seedling vigor of *Pinus massoniana* for reforestation work, and leaf moisture and nitrogen content are key factors used to achieve it. In this paper, we proposed a non-destructive testing method based on the multi-scale short cut convolutional neural network (MS-SC-CNN) to measure moisture and nitrogen content in leaves of *Pinus massoniana* seedlings. By designing a reasonable short cut structure, the method realized the transmission of loss function gradient across the multi-layer structure in the network and reduced the information loss caused by the multi-layer transmission in the forward propagation. Meanwhile, in the back propagation stage, the loss caused by the multi-layer transmission of gradient was reduced. Thus, the gradient vanishing problem in training was avoided. Since the method realized cross-layer transmission error, the convolutional layer could be increased appropriately to obtain higher measurement accuracy. To verify the performance of the proposed MS-SC-CNN non-destructive measurement method, the near-infrared hyperspectral data of sample leaves of 219 *Pinus massoniana* seedlings were collected from the Huangping Forest Farm in Guizhou Province. The correlation coefficient between the measured and real values of the prediction was as high as 0.977 and the root mean square error was 0.242 for the moisture content of *Pinus massoniana* seedling leaves. For the nitrogen content of *Pinus massoniana* seedling leaves, the correlation coefficient between the measured and real values of the prediction was 0.906 and the root-mean-square error was 0.061. The results showed that the non-destructive testing method based on MS-SC-CNN that we proposed can accurately measure the moisture and nitrogen content in leaves of *Pinus massoniana* seedlings.

Keywords: *Pinus massoniana* seedlings; moisture content; nitrogen content; near-infrared spectroscopy (NIRS); multi-scale short cut convolutional neural network (MS-SC-CNN)

1. Introduction

Pinus massoniana is a pioneer tree species that plays an essential role in reforestation work in South China due to its wide distribution, rapid growth, and strong environmental adaptability [1]. The cultivation of *Pinus massoniana* is not easy, and the quality of its seedlings is directly related to the success of reforestation. The moisture and nitrogen content in leaves of *Pinus massoniana* seedlings are closely related to plants' physiological functions and can reflect the health status of seedlings [2]. Therefore, during the initial stage of reforestation, to evaluate the health status of *Pinus massoniana* seedlings, the most rapid method is to measure the moisture and nitrogen content in *Pinus massoniana* seedling leaves.

Near-infrared spectroscopy (NIRS) is widely used to analyze the components of various plants, such as the moisture content of wheat and soybean [3,4], the nitrogen content of *Pinus massoniana* and rice [5,6], and the chlorophyll content of sunflower [7] by capturing the hyperspectral differences caused by differences of frequency doubling and

frequency vibration of molecular groups in plants. Peng et al. [8] applied multi-spectral technology and the back propagation (BP) neural network to predict the moisture content in corn leaves. However, due to the use of multi-spectral data and an insufficient volume of data, the prediction accuracy was low. Xue et al. [9] used NIRS technology, combined with genetic algorithm screening, to construct a rice moisture content predictive model. However, the method predicts based on raw characteristic wavelength, which is only suitable for the measured objects with prominent characteristics and is challenging to apply to measured objects with insignificant characteristics. Lu et al. [10] used hyperspectral technology to detect moisture content in camellia seeds and applied a radial basis function neural network as the leading predicting mean. This method integrates neural networks and hyperspectral technology, but the network structure is simple, and the prediction result is not sufficiently accurate.

Measurement models commonly used in NIRS include multi-linear regression (MLR) [11], partial least squares regression (PLSR) [12], support vector regression (SVR) [13], and artificial neural network (ANN) [14], among others [15]. MLR and PLSR are linear regression models, which are suitable for describing strong linear relationships. Traditional ANN and SVR introduce nonlinearity through a non-linear activation function and kernel function, respectively, but cannot solve problems with a high degree of abstraction; for the latter case, the deep neural network was devised.

Shen et al. combined the PLSR method based on stack self-coding with hyperspectral imaging technology to realize the rapid non-destructive testing of the solid soluble content in green plums [16]. Although this more in-depth learning network has better nonlinearity, due to the increase of the network's depth, the gradient of the backpropagation loss function may disappear during training, so that the weight of the network cannot be effectively adjusted.

He's team proposed the residual network (Res Net) in 2016, which forms a unique residual block structure by short cut (SC) [17], so that the error of the network can cross the multi-layer network during reverse propagation, overcoming the problem of gradient disappearance, and improve the depth of the network. Res Net application has achieved several research results, effectively solving practical problems. Chen et al. used a deep Res Net for fault diagnosis and defect analysis of a photovoltaic array [18]; Zhong et al. applied Res Net to build a low-dose computed tomography image reconstruction method [19]; Chen et al. built a bearing fault detection method using Res Net, which can locate the defect position accurately by the bearing vibration signal [20].

Convolutional neural network (CNN) is a commonly used deep learning model, mainly utilized to analyze and extract features from input data. A multi-scale convolutional network divides the network into blocks and limits the number of output channels in each layer to reduce the number of parameters and the computational complexity, and to prevent over-fitting. Ni et al. used stacked automatic encoders to extract layered output-related features layer by layer and then applied the SVR model to accurately predict the water content of *Pinus massoniana* leaves [21]. In another study, Wang et al. used variable-weighted CNN to improve the one-dimensional CNN model's generalization ability and predicted the nitrogen content of *Pinus massoniana* leaves.

Here, we proposed a method named multi-scale short cut convolutional neural network (MS-SC-CNN) for measuring moisture and nitrogen content in leaves of *Pinus massoniana* seedlings based on the residual network and CNN's characteristics with NIRS. First, hyperspectral reflectance data of *Pinus massoniana* seedling leaves were collected by NIRS, after which two MS-SC-CNN prediction models were established. The MS-SC-CNN was applied to extract features layer by layer, while the short cut structure was used to make the features generated by the neural network spread across layers accurately, after which the measurement of moisture and nitrogen content was achieved. The main contributions of this study are as follows: (1) We constructed a new MS-CNN structure to estimate the moisture and nitrogen content in *Pinus massoniana* seedling leaves with NIRS; (2) the

short cut was used to transfer data by skipping multiply layers in the proposed MS-CNN network to improve the estimation accuracy and reduce model complexity.

The remainder of this paper is organized as follows: Section 2 describes the collection and processing of the experimental materials and the proposed MS-SC-CNN method to estimate the moisture and nitrogen content in *Pinus massoniana* seedling leaves. Section 3 discusses and analyzes the performance of the MS-SC-CNN model based on several comparative experiments. Finally, Section 4 summarizes the conclusions and future work.

2. Materials and Methods

2.1. Preparation of Experimental Materials

A total of 219 *Pinus massoniana* seedlings were collected from the forest farm of Huangping City, Guizhou Province. We cleaned the leaves of *Pinus massoniana* seedlings to remove impurities (e.g., soil, sand) that could influence the spectrum and chemical content. Then, we gathered infrared spectrum images of the leaves of the seedlings after the leaves had been naturally dried, and measured the moisture and nitrogen content of the leaves. Finally, we randomly select 166 samples to train a model and determine the model parameters, and used the model to generate a prediction using the remaining 53 samples to verify the model's performance.

2.1.1. Near-infrared Spectroscopy Data Acquisition and Processing

An MPA Fourier transforms near-infrared spectrometer equipped (Bruker Inc., Mannheim, Germany) with a PbS detector was used to collect infrared spectrogram of *Pinus massoniana* seedling leaves, and the distance of the camera and leaves maintained the same each time. The spectrum acquisition range of the near-infrared spectrometer was $4000\text{--}12,493\text{ cm}^{-1}$ and the reflection working mode was used; the spectral resolution was 4 cm^{-1} .

All of the leaves from a *Pinus massoniana* seedling were collected and sealed for preservation to form a *Pinus massoniana* seedling leaf sample. All samples were scanned twice at the top, middle, and bottom. The average of the six scans was taken as the final value of the spectral reflectance of the sample for further analysis. Figure 1 shows the raw absorbance spectra of the leaves of 219 *Pinus massoniana* seedlings.

Figure 1. Raw absorbance spectra of *Pinus massoniana* seedling leaves, where each line is a spectrum of a *Pinus massoniana* seedling leaf, including 2203 wavelength variables.

To improve detection accuracy, the raw absorbance spectra of *Pinus massoniana* seedling leaves were pretreated in two steps. First, the quadratic polynomial Savitzky–Golay (SG) smoothing technique [22] was used to reduce noise, and then the first derivative was used to correct the baseline drift, where the window width of SG smoothing was set

to 17. Second, the spectral matrix values were converted to a scale from 0 to 1 through vector normalization to reduce the order of magnitude difference between different dimensional data. Finally, a vector with 2202 numbers between zeros and ones were obtained to represent a sample and were used as input for the proposed model.

2.1.2. Calibration of Moisture and Nitrogen Contents

We used a HB43-S halogen moisture meter to measure moisture content as standard value or target value of the *Pinus massoniana* seedling leaf samples. This instrument works on the thermogravimetric principle, i.e., the moisture content is determined by the weight loss of the sample after heating and drying [23]. The temperature was held 125 °C immediately after the sample was placed in the sampling chamber. The instrument's halogen lamp heated the sample until the sample no longer lost mass. Finally, the instrument measured the moisture content of the sample. The measurement process was typically completed in a few minutes. The measurement results of the moisture content of the experimental *Pinus massoniana* seedlings leaves are shown in Table 1.

Table 1. Results of the moisture content analysis of *Pinus massoniana* seedling leaves.

	Maximum (%)	Minimum (%)	Average (%)	Variance (%)
Calibration	70.740	48.260	60.255	22.753
Prediction	69.250	59.130	64.580	4.731
All	70.740	48.260	61.302	21.796

The nitrogen content of *Pinus massoniana* seedlings leaves was measured by an EA3000 elemental analyzer (EuroVector Instruments & Software, Pavia, Italy). The CHNS-O elemental analyzer of the Euro Vector EA3000 series and its dedicated SW Callus can measure carbon, nitrogen, hydrogen, and sulfur in all existing substances accurately. Before measurement, the test blades were dried to remove internal water and ground. When measuring, we configured the EA3000 in CHNS-O mode with helium as the carrier gas to automatically obtain an accurate nitrogen content reading. The measurement results of the nitrogen content of the experimental *Pinus massoniana* seedlings leaves are shown in Table 2.

Table 2. Results of the nitrogen content analysis of *Pinus massoniana* seedling leaves.

	Maximum (%)	Minimum (%)	Average (%)	Variance (%)
Calibration	1.354	0.072	0.502	0.033
Prediction	1.334	0.126	0.768	0.118
All	1.354	0.072	0.563	0.066

2.2. Non-Destructive Measurement Models

2.2.1. Multi-Scale Convolutional Neural Network Model

CNN is widely applied in the field of machine vision and has powerful functions in classification and recognition. CNN is better-suited to dealing with non-linear and complex systems compared with conventional methods used to process NIRS data. The basic structure of CNN includes an input layer, convolutional layer, downsampling layer, and fully connected layer [24]. However, some information maybe lost during convolutional processing by a basic CNN, while the multi-scale convolutional layer was able to extract more feature information [25]. Therefore, a multi-scale convolutional neural network (MS-CNN) model structure was constructed to estimate the moisture and nitrogen content of *Pinus massoniana* seedlings leaves with NIRS.

The structure of the MS-CNN model is shown in Figure 2, which includes a convolutional layer near the input end of the network, three multi-scale convolutional layers arranged alternately, and two multi-scale downsampling layers. They all contain several branches. The three multi-scale convolutional layers can be used for a multi-scale collection

of the samples' characteristics, and the two multi-scale downsampling layers can reduce the computational complexity of the network and maintain the characteristics contained in the spectral information without losing information.

Figure 2. The structure of the proposed multi-scale convolutional neural network model, where Conv is short for a convolutional layer, MS Conv is multi-scale convolutional layer, MS Sub is multi-scale subsampling or downsampling layer, FC is fully connected layer.

After multi-scale convolution and downsampling, the downsampling layer and a plurality of fully connected layers were applied. Due to the large number of output feature channels in multi-scale convolutional layer 3, the sampling core's size in the downsampling layer was set as 4, which reduced the dimensionality of input data in the fully connected layer. The number of neurons in the two fully connected layers was 200 and 2. The output layer was connected behind it and was composed of one neuron, and the Sigmoid function was used as the activation function to ensure that the output value was constrained to a reasonable range. The input of the loss function for network optimization is the output feature of the rear fully connected layer.

2.2.2. Multi-Scale Short Cut Convolutional Neural Network Model

The gradient may disappear or explode when training the network with many layers. To avoid this, a theory of residual learning [17] was applied to improve the proposed MS-CNN structure as shown in Figure 3. The working principle of residual learning is to insert a short cut into the basic structure of a traditional linear neural network, which is connected with the main path of the network by skipping one or more layers. The introduction of the short cut enables the gradient in the network to be transmitted directly from the bottom layer to the former layer, which avoids the problem of the gradient vanishing.

Figure 3. The structure of the multi-scale short cut convolutional neural network (MS-SC-CNN) model.

In Figure 3, two short cuts were introduced to construct a multi-scale short cut convolutional neural network (MS-SC-CNN) model. The multi-scale convolutional layer and the multi-scale downsampling layer were alternately connected, and after passing through the final downsampling layer, the feature was sent to the fully connected layer, ignoring part of the data to some extent. Therefore, the short cut method was used, and the feature data

of the multi-scale downsampling layer one and the multi-scale downsampling layer two were preserved; a “jump” approach was adopted to cross the multi-layer network into the fully connected layer, which is different from reference [17]. Its structure can be expressed as follows:

$$\begin{cases} S_i = F(x_{n-n_i}, w_{n-n_i}) \\ y_{n-2} = (S_1, S_2, S_3) \\ x_{n-1} = y_{n-2} \end{cases}, \quad (1)$$

$$x_n = (\hat{S}_1, \hat{S}_2, \hat{S}_3) \cdot W_n + b_n. \quad (2)$$

When the structure was trained, the gradient of backpropagation was:

$$\begin{aligned} \frac{\partial loss}{\partial x_{n-n_1}} &= \frac{\partial loss}{\partial x_n} \cdot \frac{\partial x_n}{\partial x_{n-n_1}} \\ &= \frac{\partial loss}{\partial x_n} \cdot \left(\frac{\partial S_1}{\partial x_{n-n_1}} + \frac{\partial (S_2, S_3)}{\partial x_{n-n_1}} \right), \\ &= \frac{\partial loss}{\partial x_n} \cdot \frac{\partial S_1}{\partial x_{n-n_1}} \left(1 + \frac{\partial (S_2, S_3)}{\partial S_1} \right), \end{aligned} \quad (3)$$

where $\frac{\partial loss}{\partial x_n}$ is the partial derivative of the spanning n th layer variable with respect to the loss function. The layer is located at the end of the network structure in the proposed method, so it retains a larger gradient. $\frac{\partial S_1}{\partial x_{n-n_1}}$ only passes through a single layer of the neural network and only reaches the feature extraction layer from the feature, so the gradient cannot die out easily. Although $\frac{\partial (S_2, S_3)}{\partial S_1} + 1$ has undergone multi-layer chain derivation, and it is typical for the gradient to disappear a traditional network structure which only contains $\frac{\partial (S_2, S_3)}{\partial S_1}$, due to the addition of the short cut structure, it is attached with 1, which stops the gradient from disappearing due to the chain rule in the $n - n_1$ layer. Thus, no matter how small $\frac{\partial (S_2, S_3)}{\partial S_1}$ is and how deep the network goes, it can be appropriately trained. This approach not only improves the utilization of data but also limits the number of output channels in each layer and reduces the number of parameters and the risk of overfitting the model.

The model uses the mean square error (MSE) as the loss function, a function of the average error for a batch. The expression is shown as follows:

$$MSE = \frac{\sum_{i=1}^n (y_i - y_i^p)^2}{n}, \quad (4)$$

where y_i is the correct result of the i th data in a batch, and y_i^p is predicted by the neural network model.

2.3. Hardware Platforms and Evaluation Criteria

The proposed algorithm's performance evaluation was carried out with the following hardware platforms: Intel Core i7-8700 CPU@3.2 GHz, with 32 GB memory and a NVIDIA GeForce RTX 2080 Ti (11G video memory) video card, and its software platforms were Ubuntu 18.04, Python 3.6, Tensor Flow-GPU 1.14.0, and Keras 2.0.8.

The correlation coefficient (R^2) and root mean square error (RMSE) were used as evaluation criteria for the calibration model. The correlation coefficient R^2 indicates the squared correlation between the actual output and the predicted output and can be used to reflect a model's reliability. It can be defined as:

$$R^2 = 1 - \frac{\sum_{i=1}^{N_T} (y_i - \hat{y}_i)^2}{\sum_{i=1}^{N_T} (y_i - \bar{y})^2}, \quad (5)$$

where $\bar{y}$ is the average of the actual output for the test samples. The closer R^2 is to 1, the higher the prediction performance of the calibration model is.

RMSE can be defined as:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N_T} (y_i - \hat{y}_i)^2}{N_T}}, \quad (6)$$

where y_i and $\hat{y}_i$ are the actual and predicted target output values for the i th sample, respectively; and N_T is the number of test samples. RMSE represents the model's accuracy, and a smaller RMSE value indicates higher predictive performance.

To sum up, a smaller RMSE value and a larger R^2 indicate better model performance.

3. Results and Discussion

3.1. Structural Performance Evaluation of MS-SC-CNN Model

Firstly, a set of comparative experiments were designed to verify the rationality of the proposed structure of the MS-SC-CNN. In this group of experiments, three kinds of neural networks were set up: CNN, MS-CNN, and MS-SC-CNN. Table 3 shows the results for estimating moisture and nitrogen content in leaves of *Pinus massoniana* seedlings using these three neural network models with different structures, among which the MS-SC-CNN model had the highest estimation accuracy. In contrast, the CNN model had the lowest accuracy. Specifically, with regard to moisture content estimation, the correlation coefficient between the estimated value and the real value of the MS-SC-CNN model was the highest, improved by 0.001 compared with the MS-CNN model, while the RMSE was reduced by 0.03. In nitrogen content estimation, the correlation coefficient of the MS-SC-CNN model was improved by 0.016 compared with the MS-CNN model, and the RMSE of the MS-SC-CNN model was reduced by 0.005, indicating that the short cut structure's generalization ability could be further improved, and the calculation accuracy of the model was improved by adding a short cut structure.

Table 3. Results of moisture and nitrogen content in *Pinus massoniana* seedling leaves using different convolutional neural network models.

Model	Moisture Content		Nitrogen Content	
	R^2	RMSE	R^2	RMSE
CNN	0.965	0.336	0.836	0.084
MS-CNN	0.976	0.272	0.890	0.066
MS-SC-CNN	0.977	0.242	0.906	0.061

Notes: CNN is short for the convolutional neural network; MS-CNN is the multi-scale convolutional neural network; MS-SC-CNN is the multi-scale short cut convolutional neural network.

Figure 4 shows the moisture content prediction effect of the three models on the calibration and prediction. All three methods accurately predicted the moisture content, but the prediction effect was different. Consistent with the results in Table 3, CNN was the method with the least ideal detection effect with the selected data set. In Figure 4a, it can be seen that most of the data points are far from the perfect correlation line, which indicates that the prediction result has a large error and was not highly accurate. The other two models' prediction results were similar, although more data points of the MS-SC-CNN were close to the perfect correlation line, which shows that the prediction error of this method is smaller, and it can better predict the moisture content of *Pinus massoniana* seedlings. Therefore, the best model in terms of its predictive ability was the MS-SC-CNN, followed by the MS-CNN, while the CNN alone had a poor predictive ability. With respect to the moisture content test, the predicted values of the CNN, MS-CNN, and MS-SC-CNN matched the actual values better than the training results.

Figure 4. Measured vs. predicted moisture content for *Pinus massoniana* seedling leaves by the regression models of the convolutional neural network (CNN) (a), the multi-scale convolutional neural network (MS-CNN) (b) and the multi-scale short cut convolutional neural network MS-SC-CNN (c).

Figure 5 shows the model estimation results for nitrogen content. In the nitrogen content test, the three prediction models' predictions were less accurate compared with the moisture content prediction, which indicates that the difficulty of the regression task is greater compared with the moisture content test. However, the estimated performance and moisture content of the three models are consistent. Figure 5a shows that a large number of data points deviate from the line of perfect correlation, and that the CNN model prediction is the worst. In contrast, the other two models had a better predictive ability, and when combined with the results in Table 3. It is clear that the prediction of the MS-SC-CNN model is better than that of the MS-CNN model.

Figure 5. Measured vs. predicted nitrogen content for *Pinus massoniana* seedling leaves by the regression models of CNN (a), MS-CNN (b), MS-SC-CNN (c).

To sum up, the predictive ability of the MS-CNN model, after improving the convolution kernel, is better than that of the CNN model, but these two models cannot be trained because of the deep network depth and the gradient extinction problem. MS-SC-CNN is based on MS-CNN but uses a short cut structure, which makes the backpropagation gradient cross from the end layer to the front network, thus improving the upper limit of the depth of the model and improving the generalization ability of the model. At the same time, the forward prediction provides a more extensive feature space for the subsequent network and improves the prediction accuracy of the model.

3.2. Performance Evaluation of Published Measurement Models

To verify that the performance of the proposed method compares with the existing methods, experiments were set up to compare several published methods for estimating moisture and nitrogen content of *Pinus massoniana* seedlings. In this set of experiments, each method uses the same data to train the model, and determines the model parameters, and then model performance was evaluated using the same test data. The experimental results are shown in Table 4, in which the PLSR, SVR, and ANN models for moisture content measurement are taken from the literature [21], and the PLSR, SVR, and ANN models for nitrogen content measurement are taken from the literature [5].

Table 4. Results of the moisture and nitrogen content in *Pinus massoniana* leaves using different regression models.

Model	Moisture Content		Nitrogen Content	
	R^2	RMSE	R^2	RMSE
PLSR	0.898	0.677	0.771	0.127
SVR	0.902	0.705	0.748	0.132
ANN	0.932	0.556	0.768	0.127
MS-SC-CNN	0.977	0.242	0.906	0.061

Notes: For moisture content measurement, PLSR, SVR, and ANN models are taken from the literature [21], and the PLSR, SVR, and ANN models for nitrogen content measurement are taken from the literature [5].

As shown in Table 4, in the test of moisture and nitrogen content in leaves of *Pinus massoniana* seedlings, using the five-layer neural network leads to a certain degree of model over-fitting, so the correlation coefficient is reduced compared with the traditional PLSR. The SVR method is not ideal because the classification hyperplane does not fully describe the characteristics of non-linear data. Although the ANN method adopts a multi-level linear structure, which makes its result preferable to that of the SVR method, it is easy to for overfitting and gradient extinction to occur because of the multi-level structure and the high number of parameters in the model.

Combined with the model estimation performance results in Table 3, the traditional CNN model's regression effect appears to be better when processing non-linear correlation data. Similarly, in the test of the leaf nitrogen content of *Pinus massoniana* seedlings, PLSR, SVR, and ANN are also not ideal methods for prediction due to various limitations of the models, which are inferior to the CNN-based methods presented in Table 3. The deep model based on CNN had a better prediction accuracy, while MS-CNN and MS-SC-CNN performed better in terms of the R^2 value. Combining the results in Tables 3 and 4 shows that the proposed MS-SC-CNN achieves the best calculation accuracy for both moisture and nitrogen content estimation.

4. Conclusions

The MS-SC-CNN model was proposed to realize the non-destructive measurement of moisture and nitrogen content in leaves of *Pinus massoniana* seedlings. The model replaced the traditional convolutional blocks with one-dimensional multi-scale convolutional blocks, which improved the resistance to over-fitting the model and its accuracy. The short cut structure helped limit the number of output channels, improved the available depth of the model, and meant the target layer changed from crossing a single layer to crossing to the end, ultimately crossing a multi-layer network. It also retained the ability for deep gradient transmission, and prevented the model from over-fitting to the training data. By combining MS-CNN with the short cut method, the moisture and nitrogen content of *Pinus massoniana* leaves was measured quickly, non-destructively, and accurately.

Although the small number of training samples used in this study may limit the predictive ability of some models, the multi-scale short cut neural network can still exhibit good prediction performance when trained with a small number of samples. In this study, the correlation coefficient was 0.998, and the RMSE was 0.087 for the prediction of moisture content, and the correlation coefficient and RMSE was 0.906 and 0.061 for the prediction of nitrogen content, respectively. Compared with the results of several other neural network models, the model we proposed had the best performance in predicting moisture and nitrogen content in leaves of *Pinus massoniana* seedlings, which can provide an accurate data basis for the cultivation and management of *Pinus massoniana* seedlings.

In addition, this study also provided some inspiration for our future work. For example, we can access the performance of the MS-SC-CNN method in another open dataset for other problems or can analyze the generation and complexity of this method and operation time cost.

Author Contributions: All authors designed this work; Z.H., T.Z. and Z.L. contributed equally to this work. Z.H. and Z.L. carried out the experiments and validation of this work; Z.L., and T.Z. wrote original draft preparation; T.Z. and C.N. reviewed and edited the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Natural Science Program of China, grant number 31570714, and the 2019 Jiangsu Province Key Research and Development Plan by the Jiangsu Province Science and Technology Agency, grant number BE2019112.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

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

References

1. You, Y.; Huang, X.; Zhu, H.; Liu, S.; Liang, H.; Wen, Y.; Wang, H.; Cai, D.; Ye, D. Positive interactions between *Pinus massoniana* and *Castanopsis hystrix* species in the uneven-aged mixed plantations can produce more ecosystem carbon in subtropical China. *For. Ecol. Manag.* **2018**, *410*, 193–200. [\[CrossRef\]](#)
2. Cao, L.; Liang, Y.; Wang, Y.; Lu, H. Runoff and soil loss from *Pinus massoniana* forest in southern China after simulated rainfall. *Catena* **2015**, *129*, 1–8. [\[CrossRef\]](#)
3. Zhang, Y.; Wang, S.; Chen, D.; Wang, Y.; Fu, H. Measurement of wheat plants water content based on near-infrared photoelectric sensors. *Trans. Chin. Soc. Agric. Mach.* **2017**, *48*, 118–122, 261.
4. Zhang, Z.; Zhang, K.; Cui, Y.; Qin, C.B.; Li, H. Fast and non-destructive detection of soybean moisture content based on near infrared spectroscopy. *Spectrosc. Spectr. Anal.* **2018**, *38*, 63–64.
5. Ni, C.; Wang, D.; Tao, Y. Variable weighted convolutional neural network for the nitrogen content quantization of Masson pine seedling leaves with near-infrared spectroscopy. *Spectrochim. Acta Part A Mol. Biomol. Spectrosc.* **2019**, *209*, 32–39. [\[CrossRef\]](#)
6. Kusumo, B.H.; Sukartono, S.; Bustan, B.; Purwanto, Y.A. Total nitrogen in rice paddy field independently predicted from soil carbon using Near-Infrared Reflectance Spectroscopy (NIRS). *J. Phys. Conf. Ser. IOP Publ.* **2019**, *1402*, 022096. [\[CrossRef\]](#)
7. Neto, A.J.S.; Lopes, D.C.; Pinto, F.A.C.; Zolnier, S. Vis/NIR spectroscopy and chemometrics for non-destructive estimation of water and chlorophyll status in sunflower leaves. *Biosyst. Eng.* **2017**, *155*, 124–133. [\[CrossRef\]](#)
8. Peng, Y.Q.; Xiao, Y.X.; Fu, Z.T.; Yu-hong, D.; Xin-xing, L.; Hai-jun, Y.; Yong-jun, Z. Water content detection of maize leaves based on multispectral images. *Spectrosc. Spectr. Anal.* **2020**, *40*, 1257–1262.
9. Xue, M.; Ying, M.; Gong, H.; Tao, S.H.; Chen, Y.Z.; Chen, Z.W.; Wang, J.M. Feature wavelength optimization combined with near-infrared technology to detect water content in rice. *Food Sci. Technol.* **2019**, *44*, 335–341.
10. Lu, X.; Luo, Y.; Jiang, P.; Wenwu, H.U. Detection of water content in camellia seeds based on hyperspectrum. *Acta Agric. Zhejiangensis* **2020**, *32*, 1302–1310.
11. Tang, R.; Chen, X.; Li, C. Detection of Nitrogen Content in Rubber Leaves Using Near-Infrared (NIR) Spectroscopy with Correlation-Based Successive Projections Algorithm (SPA). *Appl. Spectrosc.* **2018**, *72*, 740–749. [\[CrossRef\]](#) [\[PubMed\]](#)
12. Jiang, H.; Ru, Y.; Chen, Q.; Wang, J.; Xu, L. Near-infrared hyperspectral imaging for detection and visualization of offal adulteration in ground pork. *Spectrochim. Acta Part A Mol. Biomol. Spectrosc.* **2021**, *249*, 119307. [\[CrossRef\]](#)
13. Liu, J.; Osadchy, M.; Ashton, L.; Foster, M.; Solomon, C.J.; Gibson, S.J. Deep convolutional neural networks for Raman spectrum recognition: A unified solution. *Analyst* **2017**, *142*, 4067–4074. [\[CrossRef\]](#) [\[PubMed\]](#)
14. Yuanyuan, C.; Zhibin, W. Quantitative analysis modeling of infrared spectroscopy based on ensemble convolutional neural networks. *Chemom. Intell. Lab. Syst.* **2018**, *181*, 1–10. [\[CrossRef\]](#)
15. Yang, Y.; Zhou, X.; Liu, Y.; Hu, Z.; Ding, F. Wood Defect Detection Based on Depth Extreme Learning Machine. *Appl. Sci.* **2020**, *10*, 7488. [\[CrossRef\]](#)
16. Shen, L.; Wang, H.; Liu, Y.; Liu, Y.; Zhang, X.; Fei, Y. Prediction of Soluble Solids Content in Green Plum by Using a Sparse Autoencoder. *Appl. Sci.* **2020**, *10*, 3769. [\[CrossRef\]](#)
17. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 770–778.
18. Chen, Z.; Chen, Y.; Wu, L.; Cheng, S.; Lin, P. Deep residual network based fault detection and diagnosis of photovoltaic arrays using current-voltage curves and ambient conditions. *Energy Convers. Manag.* **2019**, *198*, 111793. [\[CrossRef\]](#)
19. Zhong, A.; Li, B.; Luo, N.; Xu, Y.; Zhou, L.; Zhen, X. Image Restoration for Low-Dose CT via Transfer Learning and Residual Network. *IEEE Access* **2020**, *8*, 112078–112091. [\[CrossRef\]](#)
20. Chen, L.; Xu, G.; Tao, T.; Wu, Q. Deep Residual Network for Identifying Bearing Fault Location and Fault Severity Concurrently. *IEEE Access* **2020**, *8*, 168026–168035. [\[CrossRef\]](#)
21. Ni, C.; Zhang, Y.; Wang, D. Moisture Content Quantization of Masson Pine Seedling Leaf Based on Stacked Autoencoder with Near-Infrared Spectroscopy. *J. Electr. Comput. Eng.* **2018**, *2018*, 1–8. [\[CrossRef\]](#)
22. Chen, J. A simple method for reconstructing a high-quality NDVI time-series data set based on the Savitzky–Golay filter. *Remote Sens. Environ.* **2004**, *91*, 332–344. [\[CrossRef\]](#)

-
23. Sun, J.; Cong, S.; Mao, H.; Wu, X.; Zhang, X.; Wang, P. CARS-ABC-SVR model for predicting leaf moisture of leaf-used lettuce based on hyperspectral. *Trans. Chin. Soc. Agric. Eng.* **2017**, *33*, 178–184.
 24. Krizhevsky, A.; Sutskever, I.; Hinton, G.E. Imagenet classification with deep convolutional neural networks. In Proceedings of the Advances in neural information processing systems, Lake Tahoe, NV, USA, 3–6 December 2012; pp. 1097–1105.
 25. Du, X.F.; Qu, X.B.; He, Y.F.; Guo, D. Single Image Super-Resolution Based on Multi-Scale Competitive Convolutional Neural Network. *Sensors* **2018**, *18*, 789. [[CrossRef](#)] [[PubMed](#)]