

Article

Determination of Coniferous Wood's Compressive Strength by SE-DenseNet Model Combined with Near-Infrared Spectroscopy

Chao Li *, Xun Chen, Lixin Zhang and Saipeng Wang

College of Information and Computer Engineering, Northeast Forestry University, Harbin 150040, China

* Correspondence: lichaonefuzyz@nefu.edu.cn

Abstract: Rapid determination of the mechanical performance of coniferous wood has great importance for wood processing and utilization. Near-infrared spectroscopy (NIRS) is widely used in various production fields because of its high efficiency and non-destructive characteristics, however, the traditional NIR spectroscopy analysis techniques mainly focus on the spectral pretreatment and dimension reduction methods, which are difficult to maximize use of effective spectral information and are time consuming and laborious. Deep learning methods can automatically extract features; data-driven artificial intelligence technology can discover the internal correlation between data and realize many detection tasks in life and production. In this paper, we propose a SE-DenseNet model, which can realize end-to-end prediction without complex spectral dimension reduction compared with traditional modeling methods. The experimental results show that the proposed SE-DenseNet model achieved classification accuracy and F1 values of 88.89% and 0.8831 on the larch's test set, respectively. The proposed SE-DenseNet model achieved correlation coefficients (R) and root mean square errors (RMSE) of 0.9144 and 1.2389 MPa on the larch's test set, respectively. Implementation of this study demonstrates that SE-DenseNet can realize automatic extraction of spectral features and the accurate determination of wood mechanical properties.

Keywords: near-infrared spectroscopy; coniferous wood; compressive strength; deep learning; SE-DenseNet



Citation: Li, C.; Chen, X.; Zhang, L.; Wang, S. Determination of Coniferous Wood's Compressive Strength by SE-DenseNet Model Combined with Near-Infrared Spectroscopy. *Appl. Sci.* **2023**, *13*, 152. <https://doi.org/10.3390/app13010152>

Academic Editor: Giuseppe Lazzara

Received: 7 November 2022

Revised: 14 December 2022

Accepted: 21 December 2022

Published: 22 December 2022



Copyright: © 2022 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/>).

1. Introduction

As an important raw material for production, wood plays a significant role in many aspects of daily life, such as the manufacture of furniture and the construction industry. Currently, there is a shortage of forest resources in China; therefore, it is crucial to use wood resources rationally and efficiently; mechanical strength testing of wood is a crucial aspect of wood resource utilization. The mechanical strength of wood mainly includes compressive strength, tensile strength, bending strength, hardness, etc. Current research focuses primarily on bending strength, while relatively little research has been reported on compressive strength, etc. [1]. The traditional method for determining the mechanical strength of wood requires destructive testing on wood specimens, which is expensive and causes a great waste of resources. Coniferous wood has a regular structure and a soft composition, making it an essential mechanical wood. Therefore, it is essential to develop an effective and nondestructive method for testing the mechanical strength of coniferous wood [2,3]. Commonly used non-destructive testing methods for wood are ultrasonic testing methods [4], stress wave detection method [5], X-ray inspection method [6], near-infrared spectroscopy [7], etc. In recent years, near-infrared spectroscopy (NIRS) methods have received an increasing amount of attention and research due to their ease of operation, safety, and environmental protection [8]. NIRS reveals the structure and composition information of organic matter by using the principle that the NIR absorption wavelength and intensity of different groups or the same group in different chemical environments are significantly different [9]. Researchers have made a number of advances in the non-destructive testing

of the mechanical strength of wood using NIR spectroscopy. Samuel Ayanleye et al. [10] used NIR spectroscopy to predict the density, modulus of elasticity (MOE), and modulus of rupture (MOR) of two coniferous woods. They also investigated the effects of wood surface roughness, infrared spectral range, and machine learning model on the prediction models. The results demonstrated that the prediction accuracy of NIRS data obtained from rough surfaces was higher, and the proposed adaptive neuro-fuzzy inference system (ANFIS) had better prediction accuracy on MOE and MOR compared with multilayer perceptron (MLP) neural network (NN) and partial least squares (PLS) models. M. Mancini et al. [11] collected chestnut wood samples from three different species sources in Europe, preprocessed the raw spectra, constructed a regression model, and applied variable selection techniques to enhance the model performance; finally, the root-mean-squares error of cross-validation (RMSECV) of the MOE regression model was 696.01 MPa, and R^2 was 0.78. Hao Liang et al. [12] extracted the characteristic wavelengths closely associated with the Mongolian oak MOE using a cooperative interval partial least squares and step-by-step projection algorithm and used a back propagation neural network (BPNN) to construct a calibration model with correlation coefficients of prediction (RP) reaching 0.91 and the root mean square error of prediction (RMSEP) reaching 0.76 MPa.

In recent years, with the rapid development of technologies such as artificial intelligence, big data, and cloud computing, new chemometric methods for spectral analysis have become the center of attention and a popular topic among researchers [13]. As one of the preferred research methods, deep learning methods are gradually applied to the field of NIR analysis [14]. Ba Tuan Le [15] applies deep learning-stacked sparse autoencoder (SSAE) methods to extract the advanced features of the NIR spectroscopy, and then builds the prediction model using the affine transformation (AT) and the extreme learning machine (ELM). The results show that the proposed method is superior to other typical NIR analysis methods. Yi Chen et al. [16] proposed convolutional neural networks (CNNs) combined with NIR spectroscopy for tobacco leaf maturity level; experimental results showed that the CNN discriminant models were able to precisely classify the maturity level of tobacco leaves in three categories (upper, middle, and lower position) with accuracies of 96.18%, 95.2%, and 97.31%. Jingru Yang et al. [17] applies the basic convolution block, the residual block, and the inverted residual block into the network architecture designs, proposed TeaNet, TeaResnet, and TeaMobilenet for tea quality control, which reach up to a 100% accuracy rate. Zhe Xu et al. [18] used DenseNet to predict the soil organic matter (SOM) content based on visible and near-infrared spectroscopy, which achieved a 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. Liang Zou et al. [19] presents a one-dimensional squeeze-and-excitation residual network (1D-SE-ResNet) to construct the complex relationship between pork freshness and NIR spectroscopy. Compared with five popular classification models, 1D-SE-ResNet achieves classification accuracy of 93.72%.

This paper proposes a SE-DenseNet method; this model combines the DenseNet and Squeeze-and-Excitation (SE) module, which can realize the automatic extraction of spectral features. The established qualitative and quantitative models based on SE-DenseNet achieve good determination results, which realized the nondestructive determination of mechanical strength of coniferous wood by NIR spectroscopy.

2. Materials and Data

2.1. Specimen Preparation

In this study, three common coniferous woods, larch, hemlock, and mongolica, were chosen as the experimental objects, with wood's smooth grain compressive strength serving as the mechanical index. In accordance with the standard GB/T 1935-2009 [20], standard specimens were created with dimensions of 30 mm × 20 mm × 20 mm, with the length along the grain direction. Finally, 200 specimens of larch, hemlock, and mongolica with flawless smooth-grain compressive strength were obtained.

2.2. Dataset Acquisition

As shown in Figure 1a, the NIR spectrometer utilized in this study is the Ocean Optics (USA) NIRQuest512 spectrometer, which has a detection wavelength range of 900–1700 nm, a spectral resolution of 3 nm, an indium gallium arsenic detector, and a full-signal signal-to-noise ratio of 4000:1. The spectral data should be collected by controlling the stability of the ambient temperature and humidity to eliminate the influence of the environment. Before the collection, the instrument should be calibrated. The environmental spectrum is first measured and stored as dark spectrum. Then, the whiteboard spectrum is measured as a bright spectrum, and the wood spectrum data is measured after subtraction. This can ensure that the collection is not affected by the instrument error. In this experiment, diffuse reflectance spectra were collected at three different locations in the upper and lower sections of the specimen. The average value was taken as the final spectrum of the specimen, and SpectraSuite (Ocean Optics, Dunedin, FL, USA) software was used to control the acquisition process. According to the standard GB/T 1935-2009 [20], the specimen is placed at the center of the support of the WDW-100 universal testing machine (Kexin, Changchun, China), Figure 1b shows the device. The load is applied at a uniform speed, the specimen is destroyed within 1.5 min to 2.0 min, and the damage load is recorded. The compressive strength is calculated as the damage load divides by the area under stress.

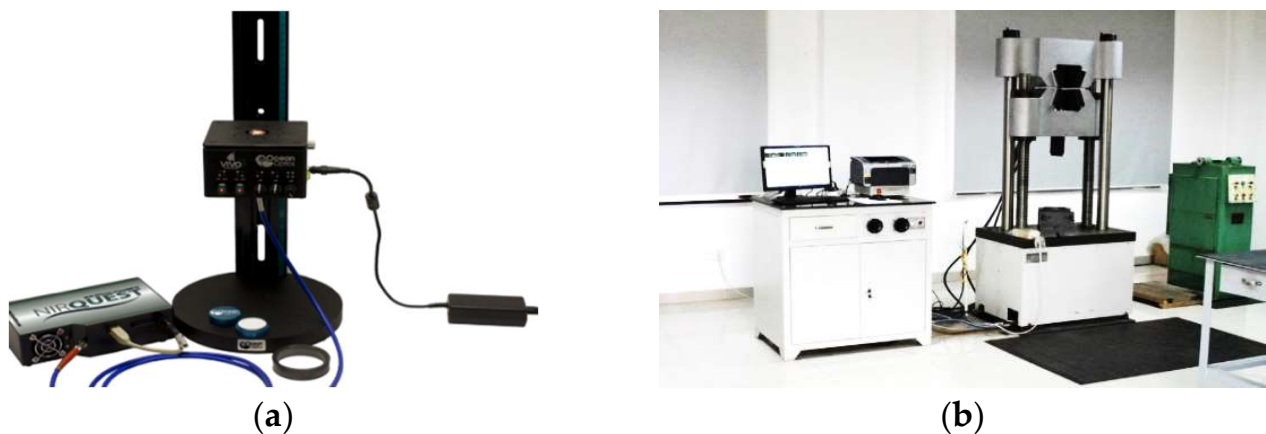


Figure 1. Experimental data acquisition devices: (a) Spectral acquisition device; (b) Universal testing machine.

3. Methods

As shown in Figure 2, after the spectral data and compressive strength data are collected, the outlier samples are rejected first. Then, the spectra are processed, and dimension reduction is performed. Finally, the Partial Least Squares Discriminant Analysis (PLS-DA), Support Vector Machine (SVM), Random Forest (RF), DenseNet and SE-DenseNet methods are used to establish the qualitative classification model. The Partial Least Squares Regression (PLSR), Support Vector Regression (SVR), and Extreme Learning Machine (ELM), DenseNet and SE-DenseNet methods are used to establish the quantitative regression model.

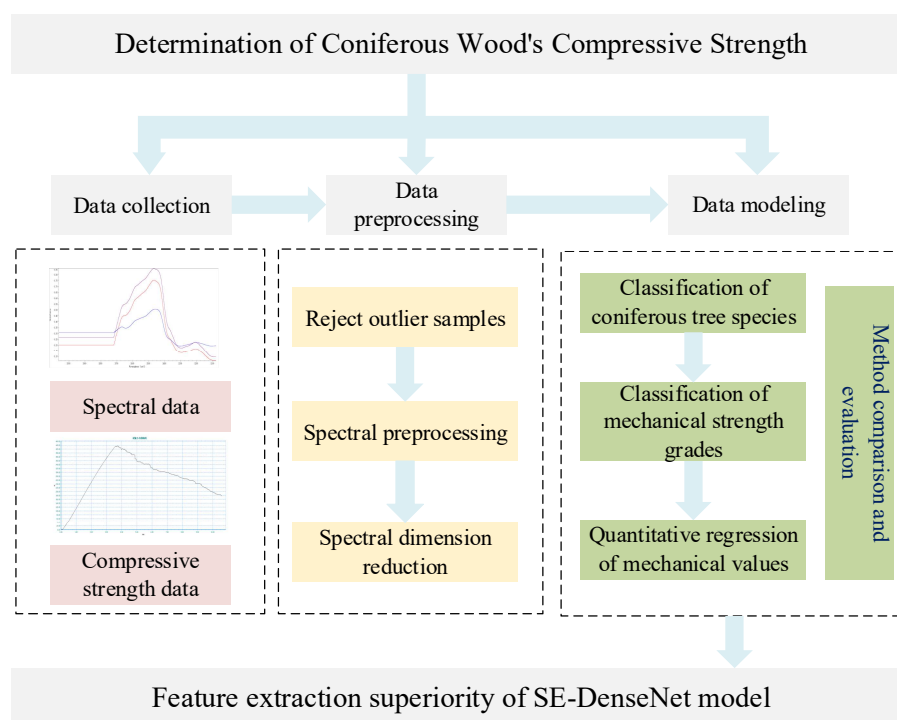


Figure 2. Technology Roadmap.

3.1. Data Preprocessing

3.1.1. Outlier Samples Rejection

Due to improper operation during the acquisition process, value entry error, model misfit, and other reasons, there are usually a small number of outlier samples in the raw data. In accordance with the standard GB/T 37969-2019 [21] and the standard GB/T 29858-2013 [22], outlier samples must be eliminated prior to spectral preprocessing. The outlier samples are divided into two categories: high leverage value samples, whose spectra are significantly different from the average spectrum of the entire model sample, and samples whose predicted values are significantly different from the reference values. In this paper, high leverage value samples and reference value outlier samples are rejected by constructing a full-wavelength PLSR quantitative model in TQ Analyst (Thermo, Waltham, MA, USA).

3.1.2. Spectral Data Preprocessing

In the process of spectral acquisition, due to differences in acquisition time and light, as well as issues with the instrument itself, the original spectrum often contains a great deal of irrelevant information and noise; therefore, the original spectrum must be appropriately preprocessed after acquisition [23]. Common spectral preprocessing methods include baseline calibration methods represented by first-order derivative (D1) and second-order derivative method (D2), smooth denoising methods represented by Savitzky–Golay (SG) convolution smoothing, and Wavelet Transform (WT) to remove high-frequency noise, scattering correction methods represented by Standard Normal Variables (SNV) and Multiplicative Scattering Correction (MSC) to reduce the effect of solid particles. The centering method represented by mean centering, normalization, and Vector Normalization (VN) reduces the error caused by the mathematical operation of matrix inverse. In this study, a large number of combined pretreatment methods were compared, and it was determined that neither the baseline calibration method nor the normalization method played a positive correction role. In this experiment, WT adopts db4 as the wavelet base and uses soft threshold method for noise filtering. Finally, the combination of WT and MSC was selected as the spectral pretreatment method.

3.1.3. Spectral Dimensionality Reduction

Full-spectrum, near-infrared wavelengths contain a great deal of redundant information; however, spectral data dimensionality reduction [24] can preserve the effective information of the spectrum and thus improve the accuracy of the model prediction. Spectral downscaling includes two types of methods, feature selection, and feature extraction. Feature selection is the process of filtering feature wavelengths or feature bands using specific methods, which do not alter the nature of the original feature space but rather select key features to form a new low-dimensional space. In this paper, the Successive Projections Algorithm (SPA) and the Competitive Adaptive Reweighted Sampling (CARS) [25] method are used to select features. The objective of feature extraction is to transform the original high-dimensional spectral data into a low-dimensional space in which each dimension is independent of the others through a mapping relationship. The feature extraction methods used in this paper are Principal Component Analysis (PCA) and locally linear embedding (LLE) [26,27]. In this paper, the combination of CARS for feature selection and LLE for feature extraction can remove redundant information to the greatest extent.

3.2. Traditional Modeling Methods

There are two main types of models for NIR spectral analysis: pattern recognition models and quantitative correction models [28]. In this paper, conventional pattern recognition models are employed. There are PLS-DA, SVM, and RF models, which primarily accomplish the classification of coniferous tree species specimens and the preliminary classification of specimen compressive strength grade and are advantageous to the rapid identification of specimen mechanical properties in wood processing. The quantitative regression models used in this paper are PLSR, SVR, and ELM, which primarily complete the accurate prediction of mechanical values of coniferous tree species specimens.

3.3. SE-DenseNet

3.3.1. DenseNet

DenseNet [29] is a kind of Dense connected neural network, whose network structure is similar to ResNet. First, a large-scale convolution is carried out; then a pooling layer is connected, and then the Dense Block and Transition Layer are moved into several consecutive submodules. Finally, a pooling and fully connected network structure is shown in Figure 3 below. DenseNet accumulates the feature map of each layer with the feature map of all previous layers in the dimension of channel number. Each layer of the network learns the feature value in a small amount, and the feature is reused, which can reduce the amount of computation, reduce redundancy, and solve the problem of gradient disappearance.

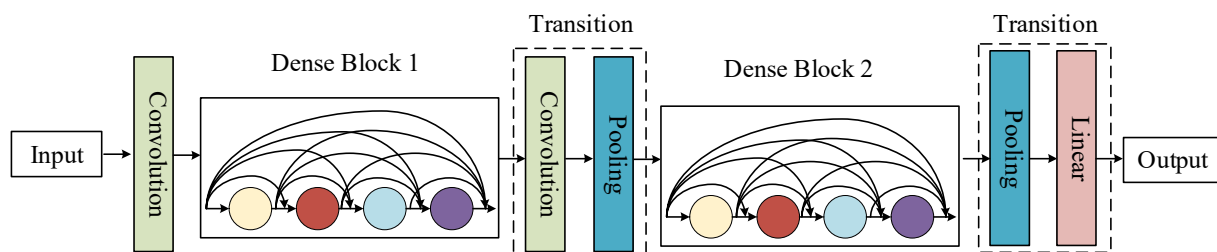


Figure 3. DenseNet model diagram.

3.3.2. SE Module

The Squeeze-and-Excitation (SE) module is a computing unit, which learns the importance of each channel of the input feature map, multiplies the obtained weight with the corresponding channel, and outputs the calibrated feature map by the weight of each channel [30]. It strengthens the useful features and weakens the useless features, so as to improve the discrimination ability of the neural network. As shown in Figure 4, first, the

feature graph X_L of the L -th layer is transformed into a tensor U_1 by convolution operation, that is:

$$U_1 = W^L \otimes f\left(BN\left(W^{L-1} \otimes f(BN([X_0 X_1 \cdots X_{L-1}]))\right)\right) \quad (1)$$

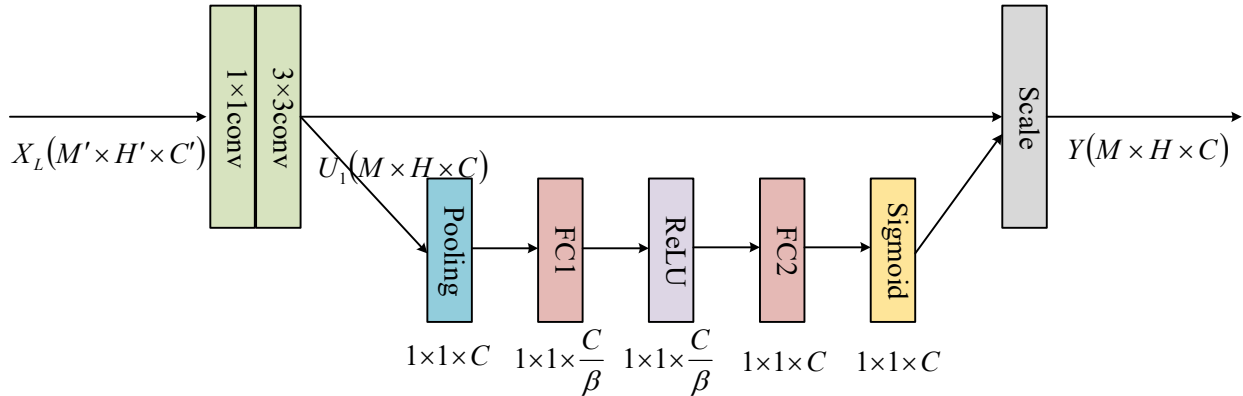


Figure 4. SE module diagram.

In Equation (1), BN represents batch normalization processing; $f(\cdot)$ is the ReLU function; W^{L-1} and W^L are the convolution kernel of size 1 by 1, 3 by 3, respectively.

Then, the Squeeze operation is carried out, which is the global pooling layer. The shape of the feature graph $[M, H, C]$ is compressed into $[1, 1, C]$, and the numerical distribution of c channels in the feature graph of this layer is obtained. The mathematical description is shown in Equation (2):

$$z_c = F_{sq}(u_c) = \frac{1}{M \times H} \sum_{i=1}^M \sum_{j=1}^H u_c(i, j) \quad (2)$$

where, u_c represents the feature diagram of the c channel after the convolution operation. z_c is the feature diagram of the c channel after Squeeze operation. M, H, C represent the three-dimensional information of the eigenmatrix U_1 .

Then the Excitation operation is carried out; its mathematical principle is shown in Equation (3):

$$s_c = F_{ex}(z, W) = \sigma(g(z, W)) = \sigma(W_2 f(W_1 z_c)) \quad (3)$$

where, $W_1 \in R^{(C/\beta) \times C}$; $W_2 \in R^{C \times (C/\beta)}$; σ is the Sigmoid function, β is the dimensional transformation rate.

Finally, the c elements of the obtained scale matrix s_c and the c channels of the feature graph U_1 are corresponding and multiplied one by one to get the output $Y = [y_1 y_2 \cdots y_c]$. The mathematical principle is showed in Formula (4):

$$Y = F_{scale}(u_c, s_c) = s_c \cdot u_c \quad (4)$$

In Formula (4), s_c represents the vector obtained after Squeeze and Excitation, and its dimension is c ; u_c represents the feature map after the convolution operation, and the number of channels is c .

The SE module, first Squeeze and then Excitation on the input, maps the feature graph into a global real number in the channel unit, and finally multiplies the real number corresponding to the input to complete the correlation learning of each channel of the feature graph.

3.3.3. SE-DenseNet

SE-DenseNet proposed in this paper is to add the SE module after the 3×3 convolution layer of each DenseNet structural block, and its model diagram is shown in Figure 5. The dotted box is the process of X_1 to X_2 , and the rest of the processes of X_0 to X_n are the same.

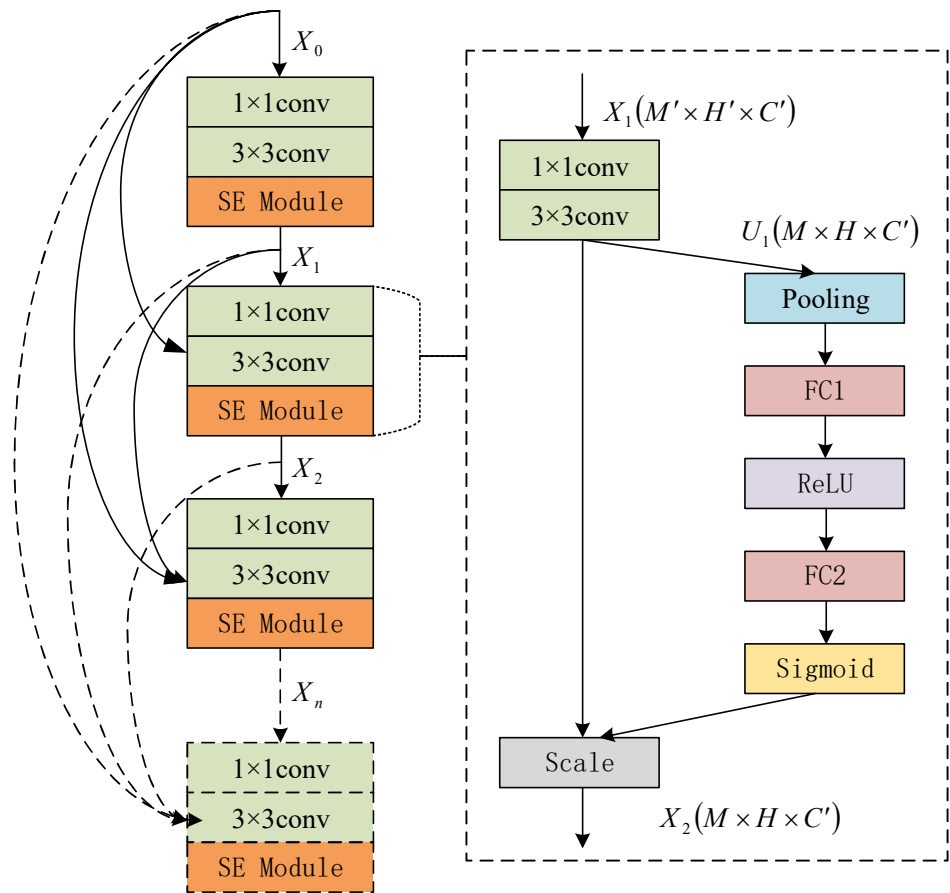


Figure 5. SE-DenseNet model diagram.

Through this fusion mechanism, the network can not only realize the lossless transmission of the original input information, but also automatically learn the global information to obtain the importance of each channel and then enhance the beneficial features and suppress the useless features according to the importance to realize the adaptive calibration of the feature channel. Table 1 lists the configurations of SE-DenseNet and DenseNet respectively. Where c is the number of characteristic channels of the convolution layer, and β is the dimension transformation rate of SE operation.

Table 1. SE-DenseNet network configuration information table.

Network Layer	SE-DenseNet		DenseNet	
	Matrix Dimensions	Structure Configuration	Matrix Dimensions	Structure Configuration
Convolution	$n \times n$	$[3 \times 3, 2c]$	$n \times n$	$[3 \times 3, 2c]$
Pooling	—	—	$n/2 \times n/2$	3×3 Max pooling
Block of structure	$n \times n$	$\begin{bmatrix} 1 \times 1, 4c \\ 3 \times 3, c \\ SE(c/\beta, c) \end{bmatrix} \times 6$	$n/2 \times n/2$	$\begin{bmatrix} 1 \times 1, 4c \\ 3 \times 3, c \end{bmatrix} \times 6$
Transition Layer	$n \times n$	$[1 \times 1, 0.5c]$	$n/2 \times n/2$ $n/4 \times n/4$	$[1 \times 1, 0.5c]$ 2×2 Average pooling
⋮	⋮	⋮	⋮	⋮
Classification Layer	1×1	Global average pool, Fully-connected, softmax		

As shown in Table 1, compared with the traditional DenseNet, SE-DenseNet has the following innovations: (1) After the convolution transformation of each structural block, the SE module is designed to carry out automatic weight calibration for the information of each feature channel; (2) After the first convolution layer, the maximum pooling layer is

removed, which can prevent the loss of low-level features caused by premature pooling operation; (3) The average pooling operation is removed from the transformation layer between structural blocks, only convolution is retained, and the global information is retained under the condition of greatly reducing the computational parameters, so as to enhance the robustness of the whole neural network.

3.4. Model Evaluation Index

The accuracy rate (ACC) and F1 value are the evaluation metrics of pattern recognition models. Tables 2 and 3 and the following formula describe the calculation process of ACC and F1. The F1 score is a statistical measure of the accuracy of a classification model that takes into account both the precision and recall of the class model. In this paper, the F1 score is the average of the F1 scores obtained by building a dichotomy for each class.

$$ACC = \frac{\sum_{g=1}^G n_{gg}}{n}$$

where n is the number of all samples in the training set or validation set.

$$precision = \frac{TP}{TP + FP}$$

$$recall = \frac{TP}{TP + FN}$$

$$F1 = \frac{2 \times precision \times recall}{precision + recall}$$

where TP is true positive, TN is true negative, FP is false positive, FN is false negative.

Table 2. Multiple classification confusion matrix.

		Predicted Class					
		1	2	3	...	G	
Actual class	1	n_{11}	n_{12}	n_{13}	...	n_{1G}	
	2	n_{21}	n_{22}	n_{23}	...	n_{2G}	
	3	n_{31}	n_{32}	n_{33}	...	n_{3G}	
	
	G	n_{G1}	n_{G2}	n_{G3}	...	n_{GG}	

Table 3. Confusion matrix for two types of discriminant analysis.

		Predicted Class	
		Positive	Negative
Actual class	Positive	TP	FN
	Negative	FP	TN

The evaluation indexes of the quantitative regression models are correlation coefficient (R), and root mean square error ($RMSE$). R is used to measure the degree of correlation between variables, and the closer R is to 1 the better the effect. $RMSE$ is used to measure the deviation between the predicted value and the true value, and the smaller its value, the better the effect.

$$R = \sqrt{1 - \frac{\sum_{i=1}^n (y_{i,actual} - y_{i,predicted})^2}{\sum_{i=1}^n (y_{i,actual} - \bar{y}_{actual})^2}}$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_{i,actual} - y_{i,predicted})^2}$$

The software used in this paper is programmed through PyCharm Community Edition 2021.3, except for the anomalous sample rejection part, which is done in TQ software.

4. Results and Discussion

4.1. Data Preprocessing Results

4.1.1. Outlier Sample Rejection Results

The presence of outlier samples can have a significant impact on the effect of modeling; therefore, outlier sample rejection must be performed prior to modeling. The spectra and reference values were imported into the TQ software to establish the full-wavelength PLSR model for larch, hemlock, and mongolica, respectively, and as shown in Figure 6, there are a few samples with abnormally high leverage values and studentized residuals. After removing the outlier samples, 179 samples of larch, 183 samples of hemlock, and 176 samples of mongolica were obtained. Figure 7 shows the spectral wavelengths before and after the removal of the outlier samples, and the spectral distribution was more uniform after the removal. Figure 8 shows the distribution of the reference values of the three conifer samples after the removal of the outlier samples. The compressive strength of larch is the greatest, followed by hemlock and mongolica, which is consistent with the spectral response information.

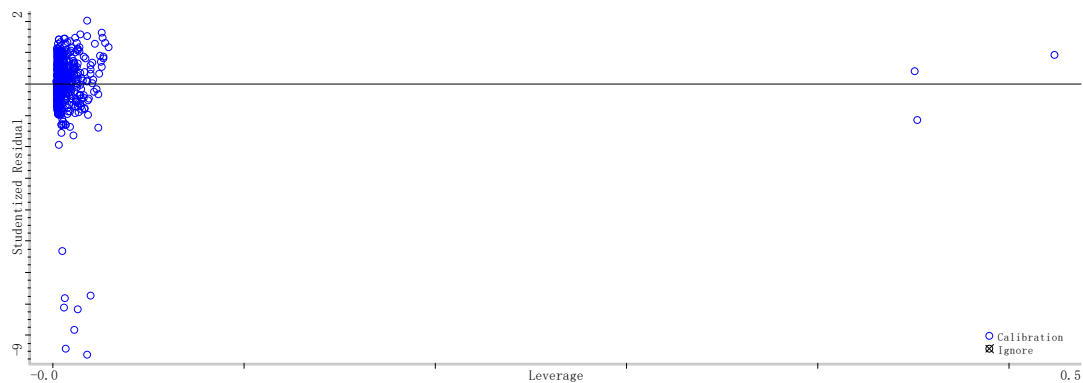


Figure 6. Outlier sample rejection for larch by TQ software.

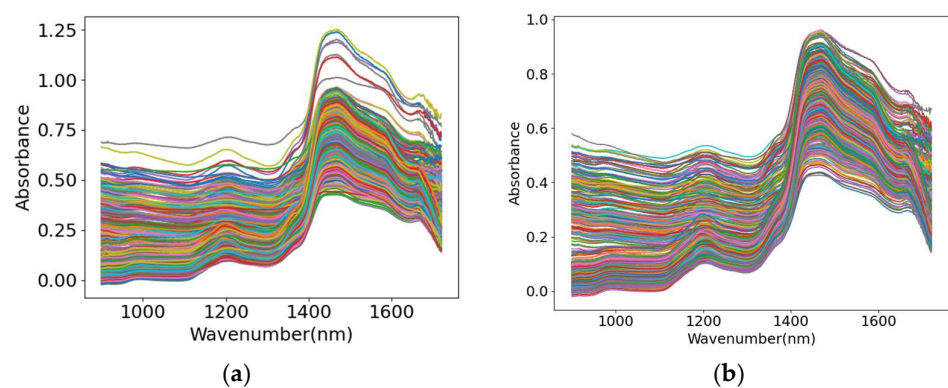


Figure 7. Spectral wavelengths of three coniferous wood: (a) Before the rejection of outlier sample; (b) After the rejection of outlier sample.

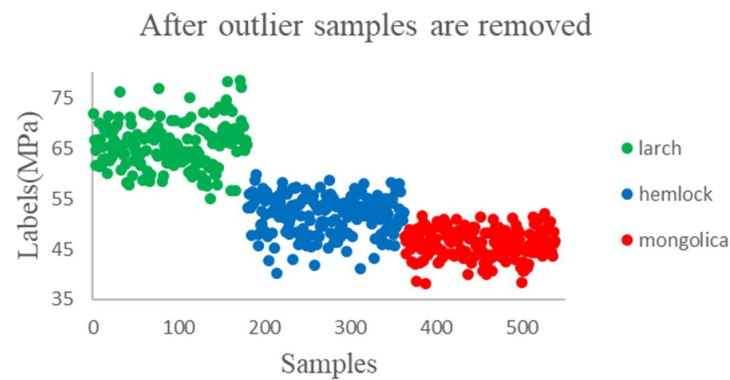


Figure 8. Distribution of reference values of three coniferous wood samples after the rejection of outlier sample.

4.1.2. Spectral Preprocessing Results

After rejecting the outlier samples, the spectra are subjected to preprocessing operations. In general, the processing sequence begins with a derivative for baseline correction, followed by smoothing and denoising, and then scattering correction and normalization. This study establishes a full-wavelength PLSR model, where n components = 3 for PLSR. The commonly used sample selection methods include random selection, Kennard–Stone (K-S), and Sample set Partitioning Based on Joint X-Y Distances (SPXY). The SPXY is generally superior to K-S and random selection [31]. Therefore, the SPXY sample selection method is chosen to divide the training set and test set by 4:1. The remaining one is used to validate the model. In this study, a 5-fold cross-validation is employed, and the training set and validation set are divided in accordance with the 4K principle of the standard GB/T 29858-2013 [22], which states that the number of samples in the validation set must be at least four times the number of principal components. The window length of SG is seven, and the order of polynomial fitting is three. The Daubechies wavelet is chosen as the wavelet transform.

Spectrum of three coniferous woods before and after preprocessing are shown in Figure 9. Using larch as an example, Table 4 compares the prediction results of a number of methods.

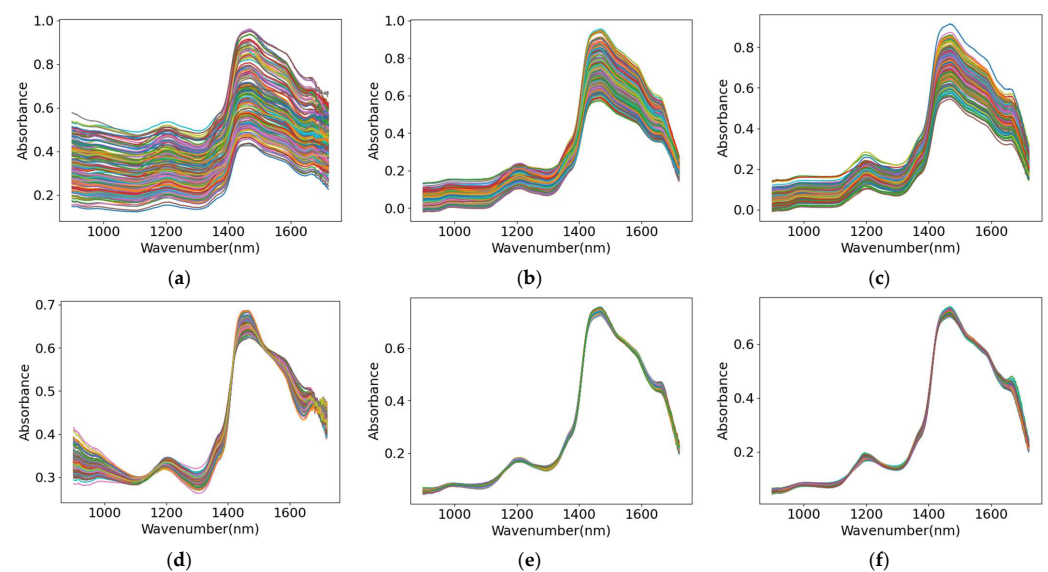


Figure 9. WT + MSC for spectral preprocessing of three coniferous wood: (a) Raw spectrum for larch; (b) Raw spectrum for hemlock; (c) Raw spectrum for mongolica; (d) Preprocessed spectrum for larch; (e) Preprocessed spectrum for hemlock; (f) Preprocessed spectrum for mongolica.

Table 4. Comparison of results of spectral preprocessing methods for larch.

Pretreatment Method	Cross-Validation Set		Test Set	
	RMSECV: MPa	RCV	RMSEP: MPa	RP
Original spectrum	3.6918	0.5594	3.9972	0.5157
D2 + WT + MSC + VN	6.6329	0.0318	6.2443	0.1465
WT + MSC + VN	2.4617	0.7354	2.4609	0.7359
D1 + WT + MSC	5.1621	0.4291	5.6343	0.3716
WT + MSC	2.2640	0.7550	2.3660	0.7453
WT + SNV	2.3657	0.7455	2.4513	0.7373
WT	3.2146	0.6294	3.3429	0.6174
SG	3.2519	0.6289	3.3876	0.6162
SG + MSC	2.5960	0.7278	2.5145	0.7317

As shown in Table 4, the effect of direct modeling of the original spectrum is poor, and the correlation coefficient on the test set is 0.5157. Using smooth denoising methods (e.g., WT, SG, etc.) alone can improve the modeling effect to a certain extent. The smooth denoising method combined with scattering correction can eliminate the spectral background and noise to the greatest extent. For example, the WT + MSC combination method can obtain the most accurate prediction results, and the correlation coefficient on the test set is 0.7453.

However, in this experiment, the derivative methods play a negative correction role and reduce the prediction result. At the same time, the normalization treatment is not helpful to improve the accuracy of the model. This shows that there is a litter baseline drift problem in this experiment. MSC and other scattering correction methods are suitable for diffuse reflection spectrum and can eliminate light scattering caused by uneven sample distribution [32]. WT can eliminate the spectral background and improve the stability of the model [33].

4.1.3. Spectral Dimensionality Reduction Results

After completing the spectral preprocessing, additional spectral dimensionality reduction was performed. In this experiment, the SPA and CARS feature selection methods, and the PCA and LLE feature extraction methods were used [27]. Due to the random nature of the results of both feature selection methods [34], the experiment is repeated five times for feature selection, and the minimum RMSECV is taken as the final result.

Using larch as an example, Figures 10 and 11 display the results of SPA and CARS feature selection. In Figure 10a, RMSE on the cross-validation set showed a trend of decreasing first, then increasing, and then finally stabilizing with the increase of the wavelength variable. The optimal number of variables was finally selected as 10, whose selection position was shown in Figure 10b. CARS algorithm is a method based on Monte Carlo sampling. As shown in Figure 11, RMSE on the cross verification set showed a trend of decreasing first and then increasing with the increase of Monte Carlo iterations. The optimal number of iterations was finally selected as 19, and the number of selected wavelength variables was 66.

The number of principal components retained after feature extraction by PCA, and CARS-LLE method was 10 and 6, respectively. The cumulative variance ratio of the first two principal components retained by PCA and CARS-LLE method was 78.69% and 87.767%, respectively. Figure 12 shows the distribution of principal component scores after PCA and CARS-LLE have extracted features. As shown in Figure 12, PCA, as a linear extraction method, extracts spectral features independently of each other, whereas LLE maintains the local linear features of the sample while downscaling.

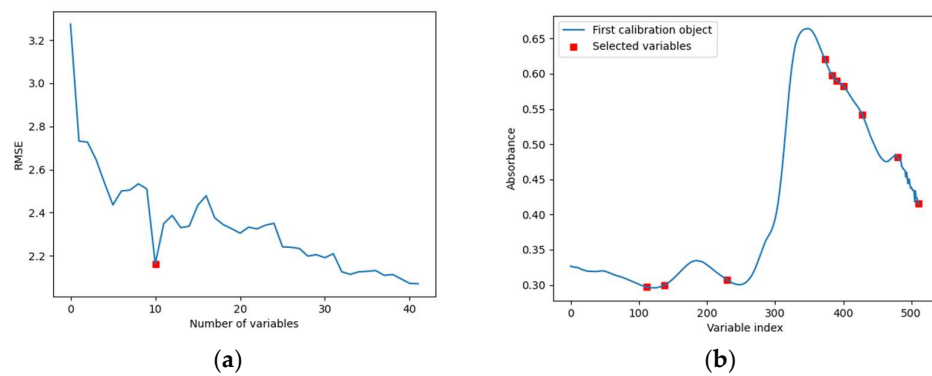


Figure 10. SPA feature selection results: (a) Optimal number of variables; (b) Location of feature.

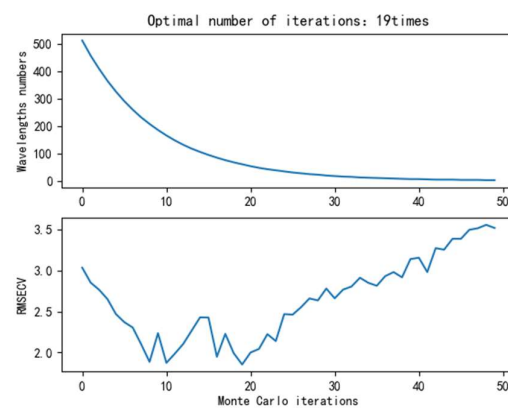


Figure 11. CARS feature selection results.

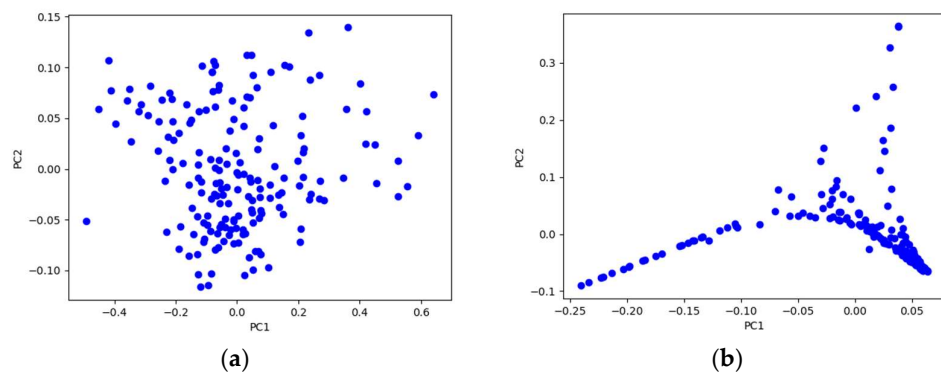


Figure 12. Distribution of principal component scores: (a) PCA (PC1: Variance = 60.613%, PC2: Variance = 18.077%); (b) CARS-LLE (PC1: Variance = 67.685%, PC2: Variance = 20.082%).

Figure 13 depicts the learning curve plots of the training samples before and after spectral dimensionality reduction. Figure 13 shows that the decision coefficient scores of the cross-validation set after dimensionality reduction are higher than those before dimensionality reduction, and that the scores of the training and validation sets tend to remain stable as the number of training samples increases. Figure 14 depicts validation curve plots of PLS principal component number before and after spectral dimensionality reduction. It can be seen that the optimal principal component number of the model is 3, and that increasing the principal component number further will result in overfitting. The comparison of the prediction results of each method is shown in Table 5.

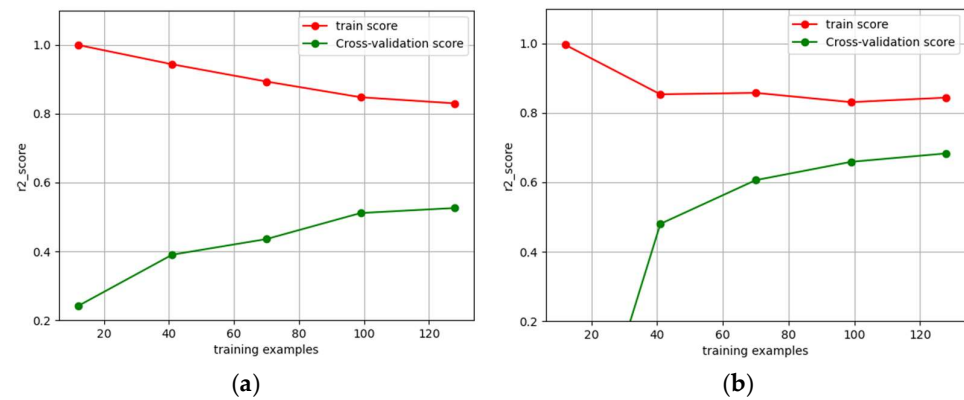


Figure 13. Learning curve of training samples: (a) Before dimension reduction; (b) After dimension reduction with CARS-LLE.

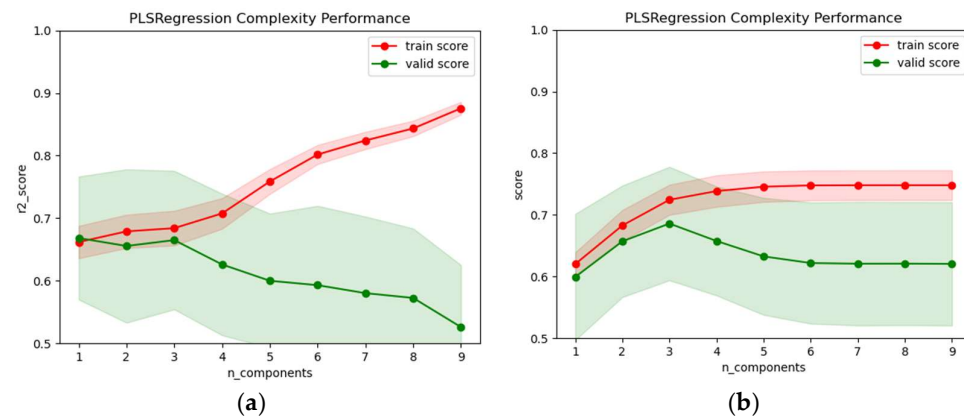


Figure 14. Validation curve of PLS principal component number: (a) Before dimension reduction; (b) After dimension reduction with CARS-LLE.

Table 5. Comparison of results of spectral dimensionality reduction methods.

Methods	Number of Features	Cross-Validation Set		Test Set	
		RMSECv: MPa	RCV	RMSEP: MPa	RP
SPA	10	1.9881	0.7879	2.0378	0.7852
CARS	66	1.9006	0.8106	1.9158	0.8033
PCA	10	2.0214	0.7852	2.0554	0.7827
LLE	7	1.9121	0.8078	1.9189	0.8025
CARS + PCA	13	1.7293	0.8319	1.6970	0.8382
CARS + LLE	6	1.6031	0.8523	1.6658	0.8498

As shown in Table 5, all methods are based on WT + MSC spectrum pretreatment. The modeling effect of the preprocessed spectrum is improved to a certain extent after dimensionality reduction. Wavelength selection methods (SPA, CARS) and feature extraction methods (PCA, LLE) are not very different when used alone. CARS has better prediction results than SPA; LLE is superior to PCA. When the two are combined, the prediction effect of the model is improved to some extent. Among them, CARS-LLE obtained the best prediction results, and the correlation coefficient reached 0.8498 in the larch test set.

Compared with the single variable selection method, the combination method can make use of the complementarity between different algorithms to first select the wavelength variable or wavelength interval, and then select fewer and more effective variables, which can eliminate the redundant information in the high-dimensional spectrum to the greatest extent.

4.2. Modeling Results

4.2.1. Classification of Coniferous Tree Species

After completing the preceding data preprocessing tasks, we began to construct the model using Random Searching and Grid Search to determine the optimal hyperparameters for the employed methods. The method of RF was used for softwood specimen classification. The number of classifiers was 15, and the classification accuracy on the test set was 100%. Figure 15 displays the classification results of three coniferous woods.

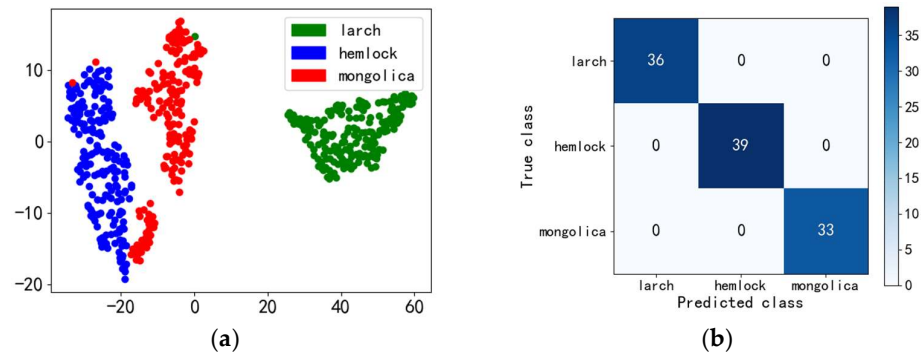


Figure 15. RF classification results of three coniferous woods: (a) Sample distribution of training set; (b) Confusion matrix of the test set.

As shown in Figure 15, the RF model can be used to achieve accurate classification of specimens of three coniferous tree species. Therefore, comparison of other modeling methods is not carried out in this part. However, we can learn that the removal of abnormal samples can make the obtained spectra more representative, and the application of NIRS analysis technology can realize the accurate determination of tree species [8].

4.2.2. The Results of the Classification of the Mechanical Strength Level

The mechanical reference values of the three types of coniferous wood were initially categorized into mechanical strength classes, allowing for a rapid and preliminary classification of the specimen’s mechanical properties during wood processing. The division results are shown in Table 6 below.

Table 6. Mechanical strength level classification.

Range of Values	Grade 1: MPa	Grade 2: MPa
larch	54.9–64.8	64.8–78.6
hemlock	40.2–52.3	52.3–59.8
mongolica	35.2–46.2	46.2–52.4

PLS-DA, SVM, RF, DenseNet, and SE-DenseNet algorithms were used to construct the classification model. The number of PLS factors selected by PLS-DA was 3; SVM penalty coefficient C set at 2.0; sigmoid kernel function selected with kernel function coefficient gamma set at 0.0001; the number of RF classifiers set at 13; the number of layers in the SE-DenseNet model is 58; and dimension transformation rate β is 12.

Table 7 displays a comparison of various methods and modeling results for various tree species. According to Table 7, in the traditional classification model, compared with PLS-DA and SVM, RF has better classification effect, but the results obtained by this method are often unstable. The DenseNet method can achieve end-to-end prediction without complex dimension reduction processing. After the introduction of the SE module, the SE-Densenet model established has obtained the optimal prediction performance. The test set accuracy of larch, hemlock, and mongolica were 0.8889, 0.8108, and 0.8571, respectively. The above results indicate that NIRS combined with deep learning can be used for the preliminary

identification of the mechanical strength grade of the species. Additionally, it has important reference value for rapid identification of mechanical properties in wood processing.

Table 7. Comparison of the results of the classification methods of strength level.

Tree Species	Methods	Cross-Validation Set		Test Set	
		ACCCV	F1CV	ACCP	F1P
larch	WT + MSC, CARS + LLE, PLS-DA	0.7931	0.7752	0.7777	0.7498
	WT + MSC, CARS + LLE, SVM	0.8276	0.8142	0.8333	0.8267
	WT + MSC, CARS + LLE, RF	0.8540	0.8379	0.8333	0.8267
	WT + MSC, DenseNet	0.8342	0.8328	0.8540	0.8454
	WT + MSC, SE-DenseNet	0.8611	0.8601	0.8889	0.8831
hemlock	WT + MSC, CARS + LLE, PLS-DA	0.7688	0.7624	0.7543	0.7499
	WT + MSC, CARS + LLE, SVM	0.7854	0.7745	0.7637	0.7591
	WT + MSC, CARS + LLE, RF	0.8065	0.8012	0.7965	0.7825
	WT + MSC, DenseNet	0.7854	0.7729	0.7965	0.7876
	WT + MSC, SE-DenseNet	0.8276	0.8201	0.8108	0.8016
mongolica	WT + MSC, CARS + LLE, PLS-DA	0.7462	0.7387	0.7354	0.7321
	WT + MSC, CARS + LLE, SVM	0.7688	0.7539	0.7428	0.7456
	WT + MSC, CARS + LLE, RF	0.8067	0.7976	0.8142	0.8078
	WT + MSC, DenseNet	0.7928	0.7863	0.8142	0.8046
	WT + MSC, SE-DenseNet	0.8214	0.8159	0.8571	0.8381

Figure 16 depicts the SE-DenseNet model cross-validation set ROC curve and test set confusion matrix. The ROC curves could evaluate the quality of the models, and the area under the ROC curve (AUC) of the three needle-leaf materials reached 0.8, indicating that these models had good classification effects. As shown in Figure 16, the AUC of the cross-validation set ROC curve of the three coniferous woods were both above 0.8. The results of the confusion matrix of three conifers on the test set also show that SE-DenseNet has a good classification effect on the preliminary identification of mechanical grade.

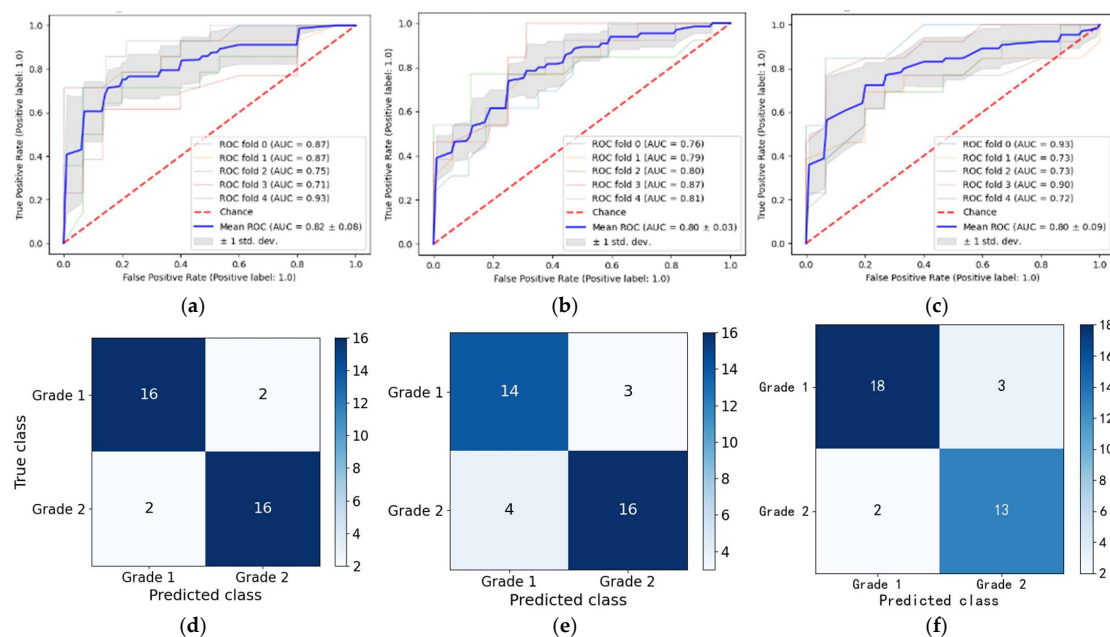


Figure 16. SE-DenseNet for three coniferous wood’s mechanical strength binary classification: (a) Cross-validation set ROC curve for larch; (b) Cross-validation set ROC curve for hemlock; (c) Cross-validation set ROC curve for mongolica; (d) Test set confusion matrix for larch; (e) Test set confusion matrix for hemlock; (f) Test set confusion matrix for mongolica.

4.2.3. Numerical Regression Results of the Mechanical Strength

Finally, a quantitative regression model of mechanical strength values was established, and PLSR, SVR, ELM, DenseNet, and SE-DenseNet algorithms were used for comparison. The number of PLS factors selected by PLSR was 3; SVR penalty factor C was 1.5, and sigmoid kernel function was selected with a kernel function coefficient gamma of 0.001; the number of ELM hidden nodes was 9; the number of layers in the SE-DenseNet model is 42; and dimension transformation rate β is 8. The cross-validation set and test set prediction results of the SE-DenseNet model are shown in Figure 17 below. Table 8 provides a comparison of various modeling techniques and the modeling results for various tree species.

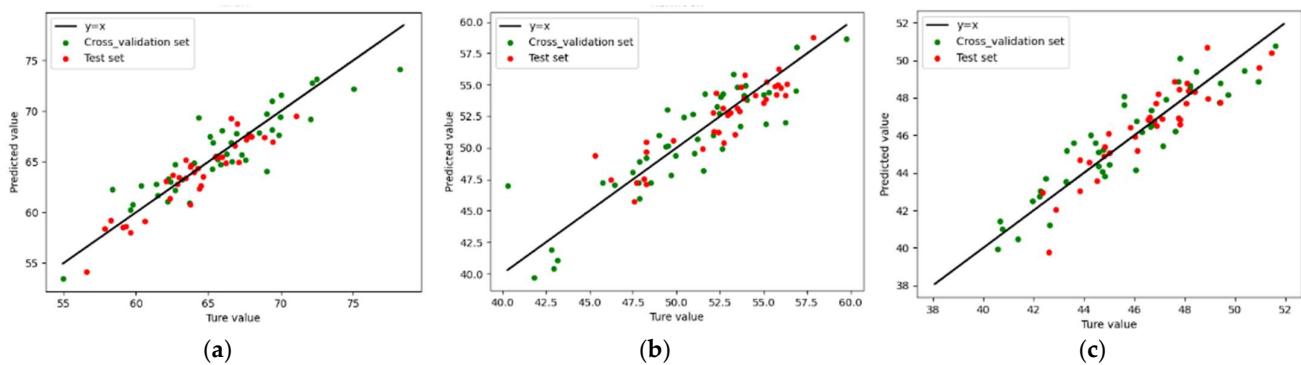


Figure 17. SE-DenseNet for three coniferous wood’s mechanical strength regression: (a) Cross validation and prediction results for larch; (b) Cross validation and prediction results for hemlock; (c) Cross validation and prediction results for mongolica.

Table 8. Comparison of results of numerical regression methods.

Tree Species	Methods	Cross-Validation Set		Test Set	
		RMSECV: MPa	RCV	RMSEP: MPa	RP
larch	WT + MSC, CARS + LLE, PLSR	1.6031	0.8523	1.6658	0.8498
	WT + MSC, CARS + LLE, SVR	1.4019	0.8765	1.4165	0.8733
	WT + MSC, CARS + LLE, ELM	1.4302	0.8722	1.4387	0.8679
	WT + MSC, DenseNet	1.3371	0.8859	1.3582	0.8768
	WT + MSC, SE-DenseNet	1.2636	0.9107	1.2389	0.9144
hemlock	WT + MSC, CARS + LLE, PLSR	1.6215	0.8485	1.6659	0.8368
	WT + MSC, CARS + LLE, SVR	1.4348	0.8542	1.5274	0.8495
	WT + MSC, CARS + LLE, ELM	1.3852	0.8659	1.3518	0.8705
	WT + MSC, DenseNet	1.3055	0.8726	1.3243	0.8717
	WT + MSC, SE-DenseNet	1.1975	0.9117	1.2293	0.8957
mongolica	WT + MSC, CARS + LLE, PLSR	1.4898	0.8577	1.5546	0.8465
	WT + MSC, CARS + LLE, SVR	1.5364	0.8469	1.4966	0.8541
	WT + MSC, CARS + LLE, ELM	1.2895	0.8698	1.2991	0.8684
	WT + MSC, DenseNet	1.2128	0.9015	1.2376	0.8874
	WT + MSC, SE-DenseNet	1.1664	0.9207	1.2244	0.8950

As shown in Figure 17, the SE-DenseNet model has achieved good results on the cross-validation and test set of three coniferous woods. According to Table 8, ELM is slightly superior to PLSR and SVR in traditional methods. The DenseNet model was comparable to the ELM model in larch and hemlock but showed a big improvement in mongolica. After the introduction of the SE module, the SE-Densenet model was established to obtain the optimal prediction performance for the three species. The RP of larch, hemlock, and mongolica were 0.9144, 0.8957, and 0.8950, respectively. Larch had a relatively better predictive performance, which may be due to the spectral pretreatment for larch. The above

results indicate that NIRS combined with SE-DenseNet can be utilized for the accurate determination of coniferous wood's compressive strength, which is of great significance to the rational processing and utilization of wood.

5. Conclusions

In this paper, three kinds of coniferous wood—larch, hemlock, and mongolica— were taken as experimental objects, and the application of NIRS combined with deep learning in the determination of wood mechanical properties was studied. The outlier samples' rejection, spectral preprocessing, spectral dimensionality reduction, model selection, and hyperparameter adjustment all have significant influence on the modeling effect. The SE module improves the sensitivity of the DenseNet model to channel features, and the SE-DenseNet model can automatically and effectively extract low-dimensional features from high-dimensional spectra, which can effectively predict the mechanical strength of coniferous wood. NIRS combined with deep learning methods has a broad application prospect in wood and other fields.

Author Contributions: Conceptualization, C.L.; methodology, C.L. and X.C.; software, X.C.; validation, L.Z. and S.W.; formal analysis, C.L. and X.C.; investigation, L.Z.; resources, S.W.; data curation, L.Z.; writing—original draft preparation, X.C.; writing—review and editing, C.L.; visualization, X.C.; supervision, C.L.; project administration, X.C.; funding acquisition, C.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the National Natural Science Foundation of China (Grant No. 32171777). It was also supported by the Fundamental Research Funds for the Central Universities (Grant No. 2572017PZ04).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: The authors are thankful to Jiajun Wang and Panpan Yang for their contribution to the supervision of this article. We are also grateful to the anonymous reviewers for their constructive remarks that improved this paper.

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

References

1. Jia, R.; Lv, Z.H.; Wang, R.; Zhao, R.; Wang, Y. New advances in wood properties prediction by near-infrared spectroscopy. *For. Ind.* **2021**, *58*, 12–16+23. (In Chinese)
2. Wang, X.P. Recent Advances in Nondestructive Evaluation of Wood: In-Forest Wood Quality Assessments. *Forests* **2021**, *12*, 949. [[CrossRef](#)]
3. Schimleck, L.; Dahlen, J.; Apiolaza, L.A.; Downes, G.; Emms, G.; Evans, R.; Moore, J.; Pâques, L.; van den Bulcke, J.; Wang, X. Non-Destructive Evaluation Techniques and What They Tell Us about Wood Property Variation. *Forests* **2019**, *10*, 728. [[CrossRef](#)]
4. Fang, Y.; Lin, L.; Feng, H.; Lu, Z.; Emms, G.W. Review of the use of air-coupled ultrasonic technologies for nondestructive testing of wood and wood products. *Comput. Electron. Agric.* **2017**, *137*, 79–87. [[CrossRef](#)]
5. Yang, Z.; Jiang, Z.; Hse, C.Y.; Liu, R. Assessing the impact of wood decay fungi on the modulus of elasticity of slash pine (*Pinus elliottii*) by stress wave non-destructive testing. *Int. Biodeterior. Biodegrad.* **2017**, *117*, 123–127. [[CrossRef](#)]
6. Lechner, T.; Sandin, Y.; Klinger, R. Assessment of density in timber using X-ray equipment. *Int. J. Archit. Herit.* **2013**, *7*, 416–433. [[CrossRef](#)]
7. Tsuchikawa, S.; Ma, T.; Inagaki, T. Application of near-infrared spectroscopy to agriculture and forestry. *Anal. Sci.* **2022**, *38*, 635–642. [[CrossRef](#)]
8. Wang, Y.; Xiang, J.; Tang, Y.; Chen, W.; Xu, Y. A review of the application of near-infrared spectroscopy (NIRS) in forestry. *Appl. Spectrosc. Rev.* **2022**, *57*, 300–317. [[CrossRef](#)]
9. Beć, K.B.; Huck, C.W. Breakthrough potential in near-infrared spectroscopy: Spectra simulation. A review of recent developments. *Front. Chem.* **2019**, *7*, 48. [[CrossRef](#)]
10. Ayanleye, S.; Nasir, V.; Avramidis, S.; Cool, J. Effect of wood surface roughness on prediction of structural timber properties by infrared spectroscopy using ANFIS, ANN and PLS regression. *Eur. J. Wood Wood Prod.* **2021**, *79*, 101–115. [[CrossRef](#)]

11. Mancini, M.; Leoni, E.; Nocetti, M.; Urbinati, C.; Duca, D.; Brunetti, M.; Toscano, G. Near infrared spectroscopy for assessing mechanical properties of *Castanea sativa* wood samples. *J. Agric. Eng.* **2019**, *50*, 191–197. [[CrossRef](#)]
12. Liang, H.; Zhang, M.; Gao, C.; Zhao, Y. Non-Destructive Methodology to Determine Modulus of Elasticity in Static Bending of *Quercus mongolica* Using Near- Infrared Spectroscopy. *Sensors* **2018**, *18*, 1963. [[CrossRef](#)]
13. Wang, H.P.; Chen, P.; Dai, J.W.; Liu, D.; Li, J.Y.; Xu, Y.P.; Chu, X.L. Recent advances of chemometric calibration methods in modern spectroscopy: Algorithms, strategy, and related issues. *TrAC Trends Anal. Chem.* **2022**, *153*, 116648. [[CrossRef](#)]
14. Mishra, P.; Passos, D.; Marini, F.; Xu, J.; Amigo, J.M.; Gowen, A.A.; Jansen, J.J.; Biancolillo, A.; Roger, J.M.; Rutledge, D.N.; et al. Deep learning for near-infrared spectral data modelling: Hypes and benefits. *TrAC Trends Anal. Chem.* **2022**, *157*, 116804. [[CrossRef](#)]
15. Le, B.T. Application of deep learning and near infrared spectroscopy in cereal analysis. *Vib. Spectrosc.* **2020**, *106*, 103009. [[CrossRef](#)]
16. Chen, Y.; Bin, J.; Zou, C.; Ding, M. Discrimination of fresh tobacco leaves with different maturity levels by near-infrared (NIR) spectroscopy and deep learning. *J. Anal. Methods Chem.* **2021**, *2021*, 9912589. [[CrossRef](#)]
17. Yang, J.; Wang, J.; Lu, G.; Fei, S.; Yan, T.; Zhang, C.; Lu, X.; Yu, Z.; Li, W.; Tang, X. TeaNet: Deep learning on Near-Infrared Spectroscopy (NIR) data for the assurance of tea quality. *Comput. Electron. Agric.* **2021**, *190*, 106431. [[CrossRef](#)]
18. Xu, Z.; Zhao, X.; Guo, X.; Guo, J. Deep learning application for predicting soil organic matter content by VIS-NIR spectroscopy. *Comput. Intell. Neurosci.* **2019**, *2019*, 3563761. [[CrossRef](#)]
19. Zou, L.; Liu, W.; Lei, M.; Yu, X. An Improved Residual Network for Pork Freshness Detection Using Near-Infrared Spectroscopy. *Entropy* **2021**, *23*, 1293. [[CrossRef](#)]
20. GB/T 1935-2009; Test Method for Longitudinal Compressive Strength of Wood. National Standards of People’s Republic of China: Beijing, China, 2009.
21. GB/T 37969-2019; General Principles for Qualitative Analysis of Near-Infrared Spectroscopy. National Standards of People’s Republic of China: Beijing, China, 2019.
22. GB/T 29858-2013; General Principles for Quantitative Analysis of Molecular Spectral Multivariate Correction. National Standards of People’s Republic of China: Beijing, China, 2013.
23. Bian, X. Spectral Preprocessing Methods. In *Chemometric Methods in Analytical Spectroscopy Technology*; Springer: Singapore, 2022; pp. 111–168.
24. Zebari, R.; AbdulAzeez, A.; Zeebaree, D.; Zebari, D.; Saeed, J. A comprehensive review of dimensionality reduction techniques for feature selection and feature extraction. *J. Appl. Sci. Technol. Trends* **2020**, *1*, 56–70. [[CrossRef](#)]
25. Xing, Z.; Du, C.; Shen, Y.; Ma, F.; Zhou, J. A method combining FTIR-ATR and Raman spectroscopy to determine soil organic matter: Improvement of prediction accuracy using competitive adaptive reweighted sampling (CARS). *Comput. Electron. Agric.* **2021**, *191*, 106549. [[CrossRef](#)]
26. Zhang, D.Y.; Jiang, D.P.; Zhou, B.L.; Cao, J.; Zhao, S.Q. Local linear embedding based on flow learning for near-infrared detection of red pine nut quality. *J. Northeast. For. Univ.* **2019**, *47*, 45–48. (In Chinese)
27. Raghavendra, A.; Guru, D.S.; Rao, M.K. Mango internal defect detection based on optimal wavelength selection method using NIR spectroscopy. *Artif. Intell. Agric.* **2021**, *5*, 43–51. [[CrossRef](#)]
28. Chu, X. *Chemometric Methods in Modern Spectral Analysis*; Chemical Industry Press: Beijing, China, 2022; pp. 176–197.
29. Zhu, Y.; Newsam, S. Densenet for dense flow. In Proceedings of the 2017 IEEE International Conference on Image Processing (ICIP), Beijing, China, 17–20 September 2017; IEEE: Piscataway, NJ, USA, 2017; pp. 790–794.
30. Hu, J.; Shen, L.; Sun, G. Squeeze-and-excitation networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018; pp. 7132–7141.
31. Yun, Y.H. Method of Selecting Calibration Samples. In *Chemometric Methods in Analytical Spectroscopy Technology*; Springer: Singapore, 2022; pp. 297–308.
32. Yang, Y.; Zhao, C.; Huang, W.; Tian, X.; Fan, S.; Wang, Q.; Li, J. Optimization and compensation of models on tomato soluble solids content assessment with online Vis/NIRS diffuse transmission system. *Infrared Phys. Technol.* **2022**, *121*, 104050. [[CrossRef](#)]
33. Huang, L.; Wu, K.; Huang, W.; Dong, Y.; Ma, H.; Liu, Y.; Liu, L. Detection of fusarium head blight in wheat ears using continuous wavelet analysis and PSO-SVM. *Agriculture* **2021**, *11*, 998. [[CrossRef](#)]
34. Wang, L.; Wang, R. Determination of soil pH from Vis-NIR spectroscopy by extreme learning machine and variable selection: A case study in lime concretion black soil. *Spectrochim. Acta Part A Mol. Biomol. Spectrosc.* **2022**, *283*, 121707. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.