Article

A Novel Tran_NAS Method for the Identification of Fe- and Mg-Deficient Pear Leaves from N- and P-Deficient Pear Leaf Data

Xiu Jin,[#] Wenjing Ba,[#] Lianglong Wang, Tong Zhang, Xiaodan Zhang, Shaowen Li, Yuan Rao,* and Li Liu*

Read Online

ACCESS | Metrics & More | Article Recommendations

ABSTRACT: Trace element deficiency diagnosis plays a critical role in pear cultivation. However, high-quality diagnostic models are challenging to investigate, making it difficult to collect samples. Therefore, this manuscript developed a novel transfer learning method, named Tran_NAS, with a fine-tuning neural network that uses a neural architecture search (NAS) to transfer learning from nitrogen (N) and phosphorus (P) to iron (Fe) and magnesium (Mg) to diagnose pear leaf element deficiencies. The best accuracy of the transferred NAS model is 89.12%, which is 11% more than that of the model without the transfer of trace element-deficient samples. Meanwhile, Tran_NAS also has better performance on source datasets after comparing with different proportions of training sets. Finally, this manuscript summarizes the transfer model coincident characteristics, including the methods of batch normalization (BN) and dropout layers, which make the model more generalizable. This manuscript applies a symmetric homogeneous feature-based transfer learning method on NAS that is designed explicitly for near-infrared (NIR) data collected from nutrient-deficient pear leaves. The novel transfer learning method would be more effective for the micro-NIR spectrum of the nondestructive diagnosis.
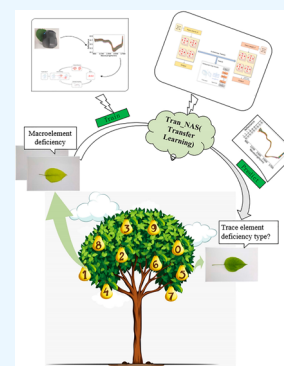
1. INTRODUCTION

Pears are one of the main fruits in the world, and the production of nutritious pears is very important. The nutrients in a pear tree determine the quality of the pears, and the various nutrient elements of a pear tree will affect the normal growth of the fruit.[1] Therefore, a reduction in nutrients can reduce the yield and quality of fruit.[2,3] Among them, nitrogen (N) deficiency reduces the photosynthetic rate of pear trees, and excess N weakens the disease resistance of pear trees.[2] Phosphorus (P) deficiency can lead to slow plant growth and poor root development, affecting the fruit quality and yield.[4] An iron (Fe) deficiency impairs the photosynthetic leaf rate as well as the transpiration rate.[5] Magnesium (Mg) deficiency affects carbohydrate production and translocation, thereby affecting the fruit ripening process.[6] For steady growth and increased fruit production, it is necessary to know the nutritional status of a pear tree.

Traditional pear leaf deficiency identification methods include morphological diagnoses and physical and chemical analyses. Morphological diagnoses are the determination of whether the leaves are deficient in N, P, Fe, and Mg using external morphological characteristics, such as the color and shape of the leaves.[7] This technique has low recognition accuracy and requires skilled fruit farmers. Physicochemical experiments can be used to diagnose deficiencies by detecting the chemical compositions of pear leaves; however, it is costly and has a slow result turnaround.[8] Nutrient deficiencies in pear leaves can result in different color characteristics, so RGB
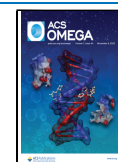
cameras can be used to identify pear leaves for deficiency diagnosis. However, the leaf surface features were almost the same when trace element deficiency was less, and the RGB camera could not detect pear leaf deficiency in time. Therefore, it is of great significance to explore an accurate, rapid, cost-efficient detection method for improving the yield and quality of pears.
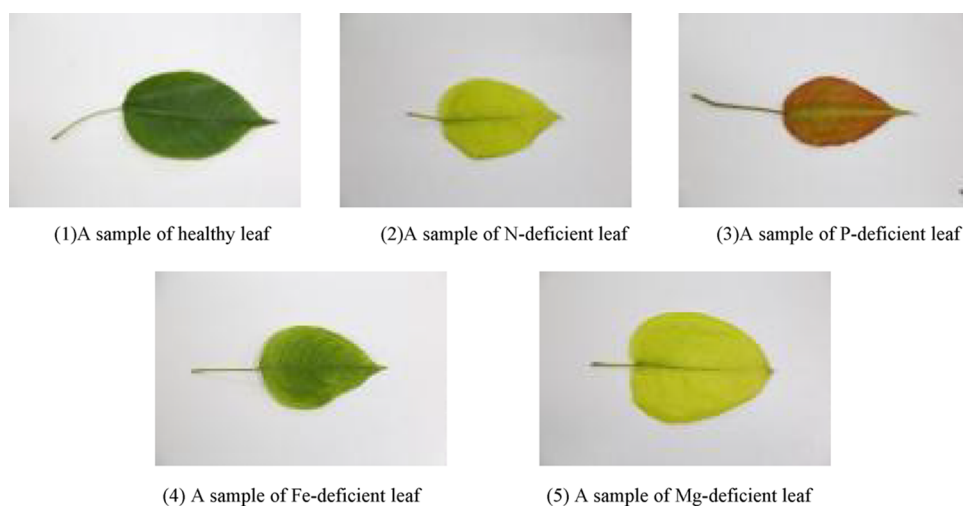
The identification of pear leaf deficiencies based on near-infrared (NIR) spectroscopy is a nondestructive and low-cost method and is therefore gaining increasing attention. The NIR spectra range from 780 to 2526 nm, and the absorption intervals can be roughly attributed to the first, second, third, or higher octaves of molecular vibrations of the hydrogen-containing groups, that is, C−H, N−H, O−H, S−H radicals, and their combined frequencies.[9] Because the macronutrients and trace elements do not have these vibration modes, they cannot be detected directly with NIR spectra. However, most of the nutrient minerals of fruit trees are present in organic matter and combined with one of the abovementioned moieties, so quantitative and qualitative analyses of nutrients can be achieved indirectly by NIR spectroscopy.[10] For

(1)A sample of healthy leaf (2)A sample of N-deficient leaf (3)A sample of P-deficient leaf

(4) A sample of Fe-deficient leaf (5) A sample of Mg-deficient leaf

Figure 1. Representative samples of pear leaves.

example, Whittier et al.[11] used NIR spectroscopy to predict the foliar nutrient levels of hydroponically grown teak seedlings. Bedin et al.[12] detected C, N, P, and K in poultry litter by NIR spectroscopy via support vector machine (SVM) and achieved good results. Jamil Maia et al.[13] used NIR spectroscopy to predict potentially toxic elements in soil and sediments from a semiarid and coastal humid tropical transitional river basin through the random forest (RF), and satisfactory results were obtained. Huang et al.[14] conducted nondestructive detection of pork with NIR by Adaptive Boosting (AdaBoost) and obtained ideal results. However, the generalization of the identify model of NIR spectra has inferior performance by machine learning (ML) approaches that have failing implementation on the target dataset because of varied feature spaces and distributions.

It is desired to develop a robust transfer method to reduce the model uncertainty and make the model more generalizable. In recent years, some advanced ML methods, for example, deep learning, have been quickly developed. Sathyavani et al.[15] used a variant of convolutional neural networks (denseNet-BC) to identify nutrient deficiencies in rice crops. However, a better trained deep model requires a large sample size and high computational cost, which makes dealing with complex problems in real scenarios difficult. With deep learning, the transfer learning method learns the knowledge in the target domain using the knowledge in the source domain to solve the learning problem when the target domain only has few samples that cannot be accurately classified.[16] In general, if one neural network can extract features in a specific dataset, transfer learning can strengthen its generalization to similar tasks without requiring too much data in the other datasets. In recent years, many researchers have studied transfer learning.[17,18] Wan et al.[18] developed a new transfer learning method by coupling transfer component analysis with support vector regression (TCA-SVR) to transfer LNC (leaf nitrogen concentration) assessment models across different plant species. Guerrero et al.[17] used image recognition of banana leaves using convolutional neural networks trained by transfer learning and fine-tuning to determine the absence of nutrients in banana leaves. The cost of obtaining a large number of different pear leaf deficiency samples is too high, so transfer learning is very important for modeling of pear leaf deficiency by NIR spectroscopy. Feature transfer learning is the major

researched part of transfer learning and can be further classified into symmetrical feature transfer learning and asymmetrical feature transfer learning.[19]

The symmetric feature-based transfer learning approach discovers potentially meaningful structures between domains to find a common latent feature space that has predictive qualities while reducing the marginal distribution between the domains.[20] Zhang et al. proposed a method to diagnose susceptibility to alcoholism by extracting features using deep learning algorithms combined with transfer learning.[21] Despite it being a popular topic, to the best of the authors' knowledge, there is no research on applying transfer learning to NIR based on nutrient-deficient pear leaf samples. The application of symmetric homogeneous feature-based transfer learning of the nutrient-deficient pear leaf domain is novel.

In this manuscript, the cost is reduced because of the development of a novel transfer learning Tran_NAS method in which the neural architecture search (NAS) model of macroelement-deficient identification is transferred to trace element-deficient pear leaves. The current manuscript aims to (1) establish a determination model using spectral data; (2) propose the Tran_NAS method to explore the feasibility of trace element transfer; and (3) explore the NAS model with datasets of different proportions and neural architectures. This manuscript proposes a method to identify the deficient elements in pear leaves based on NIR spectroscopy; the method will enable pear growers to quickly determine the physiological condition of pear leaves, and when a problem occurs, they can respond to it accordingly.

2. MATERIALS AND METHODS

2.1. Materials and Experiments. *2.1.1. Pear Leaf Deficiency Samples.* In this manuscript, an experiment on elemental deficiency stress of pear trees was conducted in the Dayangdian experimental field of Anhui Agricultural University, Hefei, Anhui Province, China (31°55′59.3″N, 117°11′52.7″E). Nitrogen-deficient, phosphorus-deficient, iron-deficient, and magnesium-deficient pear trees were cultivated separately, and healthy pear trees were set as the control group. Pear leaf samples were collected on April 13, April 23, and May 10, 2021. It was the growing season of the pear trees, and they had luxuriant foliage. Under the guidance of garden experts, a total of 1078 pear leaf samples were
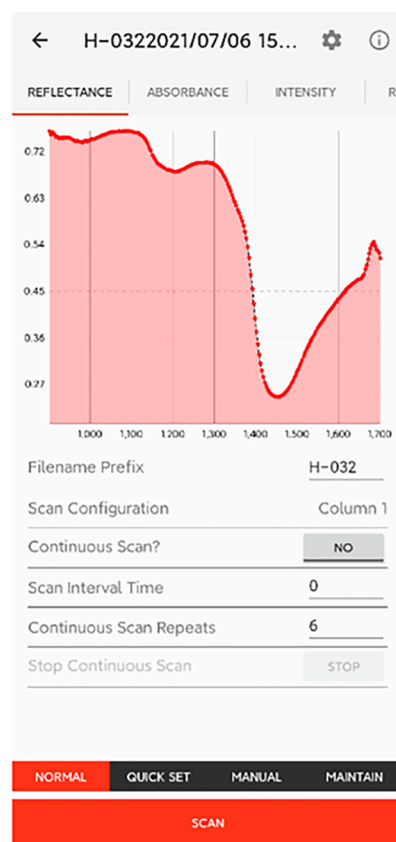
(a) Handheld miniature NIR spectrometer (NIR-S-G1)
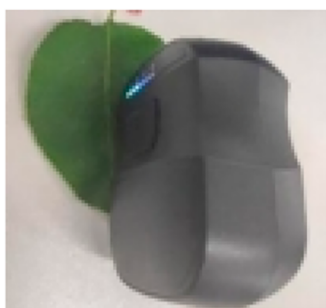
(b) Experimental method

(c) Operation interface

Figure 2. Instruments and experimental method: (a) Handheld miniature NIR spectrometer (NIR-S-G1). (b) Experimental method. (c) Operation interface.

collected, including 223 healthy samples, 268 N-deficient samples, 325 P-deficient samples, 178 Fe-deficient samples, and 84 Mg-deficient samples. On the same day, the spectral reflectance of the pear leaf samples was collected and then sent to professional institutions for the determination of nutrient elements. Figure 1 shows representative samples of different types of pear leaves. Nitrogen-deficient leaves had a yellowish-green color, phosphorus-deficient leaves had a purplish-red color, iron-deficient leaves had a yellowish color in the vein part, and magnesium-deficient leaves had a yellowish color in the flesh part.

2.1.2. Micro-NIR Experiment. In this manuscript, the NIR reflectance data were collected using a handheld miniature NIR spectrometer (product model: NIR-S-G1, created by Shenzhen Puyan Internet Technology Co., LTD) with a spectral range of 900−1700 nm, a spectral resolution of 3.89 nm, and a signal-to-noise ratio (SNR) of 5000:1, as shown in Figure 2a. The instrument was closely fitted with the leaf sample, as shown in Figure 2b. The light source emitted by the instrument is reflected into the spectrometer through the leaf surface, and the reflected light from the leaf surface is converted into brightness values (BV) inside the spectrometer. The NIR spectrometer was calibrated with a polytetrafluoro-ethylene (PTFE) compound, which gives the highest reflectance and lowest absorbance so that the device can provide the best result with the smallest calibration error.

Before the measurement of each leaf sample, the spectrometer was first corrected by standard white reference and dark reference, and the BV of the white reference and dark reference were saved. The reflection data of the pear leaf spectrum can be obtained by calculating eq 1. *I* represents the BV of the sample, B represents the BV of the dark reference, and W represents the BV of the white reference. Then, we selected five points on each leaf, that is, one central point and four random points, for measurement. Each point was measured three times, and the average spectrum was taken as the sample spectrum for modeling. The results are shown in Figure 2b.

$$\text{Reflectance} = \frac{I - B}{W - B} \times 100\% \tag{1}$$

2.1.3. Chemical Diagnosis of Pear Leaf Deficiency. The deficient leaves and normal leaves of the pear trees were collected, and the petioles were removed for professional treatment. After the treatment, a quantitative sample was taken and processed in an infrared digestion oven using the H_2SO_4−H_2O_2 decoration method. Nitrogen exists in the form of inorganic ammonium salts, phosphorus in the form of inorganic phosphate, and potassium in the form of ions. Nitrogen (N) was measured using the Kjeldahl method,[22] phosphorus (P) using the molybdenum−antimony anticolori-metric method, and magnesium (Mg) and iron (Fe) using the atomic absorption method.[23] The normal range of nutrients in

pear leaves[24] is shown (Table 1). An *N* value that is between 2.5 and 2.7 is considered normal, and a value that is less than 2.5 is judged as deficient.

Table 1. Reference Values of the Adequate Nutrient Status of Pear Leaves

nutrient	adequate values	deficiency condition
N	2.5−2.7%	<2.5
P	0.14−0.20%	<0.14
Mg	0.25−0.80%	<0.25
Fe	100 mg/kg	21~30

2.2. Preprocessing Methods. To improve the spectral characteristics, correct baseline drift and spectral scattering, and eliminate random noise to improve the SNR, Savitzky−Golay (SG) smoothing,[25] multiple scattering correction (MSC),[26] standard normal variate (SNV), first derivative (FD), second derivative (SD), LG transform (LG: log(1/Reflectance)), and two hybrid SG smoothing with multiple scattering correction (SG + MSC), SG smoothing with standard normal variate (SG + SNV) models were used to preprocess the pear leaf spectral data. The derivatives (FD and SD) can eliminate spectral shifts as well as baseline effects.[26] SNV and MSC can reduce the scattering error caused by the tightness of the object surface. SG can effectively eliminate high-frequency random noise. To improve the robustness of the model, the best preprocessing method is selected by comprehensively evaluating the performance of each model with different evaluation indexes.

2.3. Modeling Methods. In this manuscript, the random forest (RF), extreme gradient boosting (XGBoost), adaptive boosting (AdaBoost), support vector machine (SVM), multi-layer perceptron (MLP), and neural network architecture search (NAS) methods were used to develop models. RF is an algorithm that combines Breiman's idea of bagging[27] and Ho's random subspace method.[28] AdaBoost and XGBoost are two boosting algorithms.[29] SVM is a widely used supervised classification learning algorithm[30] and predictively classifies data by finding the maximum boundary in the feature space. SVM is based on empirical risk minimization for the prediction of nonlinear data,[31] and overfitting is prevented by setting the penalty parameter C as well as the kernel function parameter gamma. The MLP model[32] is a neural network classifier that maps input datasets to output datasets using a feedforward neural network. It is designed to learn the feature space of

nonlinear spectra without considering their statistical properties.[33] Because the MLP model requires the manual setting of the hyperparameters, number of layers, and network structure, it takes a large amount of time to reset a network in case of failure.

NAS is an algorithm for searching the best neural architecture. In recent years, many studies have been conducted on this topic. Among them, Weng et al.[34] used NAS for medical image segmentation and achieved good results, and Koh et al.[35] investigated the application of NAS in image-based plant phenotype analysis by using a drone to photograph wheat collapse. NAS mainly includes a search strategy and a performance evaluation strategy.[36] The search space lists all available network architectures, and the search strategy describes how to explore the search space. The final evaluation strategy identifies those network structures that will produce the best results with the new data. In this manuscript, a novel framework using Bayesian optimization guides network morphism for efficient neural architecture search.[37] NAS automatically sets up the network architecture and its hyperparameters, while the newly generated subnetworks can have the advantages and weights of the previous network for relevant training, making the whole process more efficient. Then, NAS designs an editable neural network kernel to solve the problem that the NAS space is not Euclidean space, and the specific kernel function formula is shown in eq 2:

$$k(f_a - f_b) = e^{-p^2(d(f_a - f_b))} \qquad (2)$$

where the function $(f_a - f_b)$ represents the editable distance between the two neural networks and p is a mapping function that maps the distance in the original metric space to the corresponding distance in the new space.

The workflow of NAS is shown in the following diagram. NAS is based on an initial network, and the search space is constructed using subnetworks, which can be obtained by shrinking the layers of the initial network. NAS is learned by training a super network to jointly train all the subnetworks with shared weights. The super network has the same architecture as the original architecture and contains these shared weights. The super network contains different layer widths and blocks. An optimizer is used to find the best performing dense architecture after training the super network. The optimizer interactively samples the search space, generates a set of samples, and determines the next set of samples based on the current sample performance. This process continues
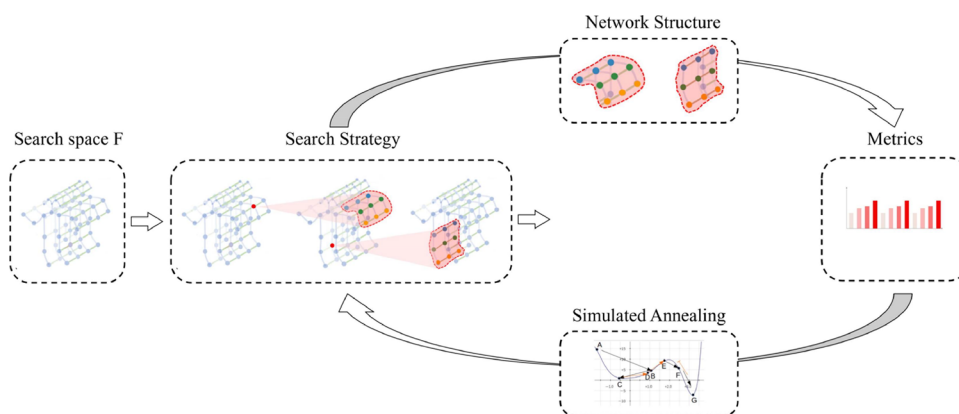
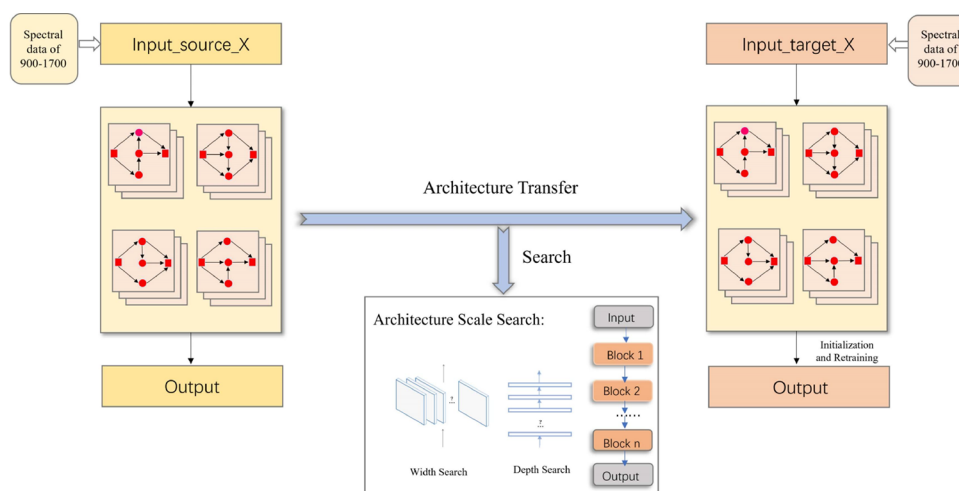
Figure 3. NAS algorithm process.

Figure 4. Tran_NAS algorithm process.

until a given stopping criterion is satisfied, and then, the discovered dense architecture is trained until it converges. As a result of the trained super network, the accuracy of the samples can be evaluated directly using the shared weights without further training (Figure 3). A simulated annealing algorithm is used to extract the optimal architecture from the trained network state projections. The simulated annealing equation is shown in eq 3:

$$f_{\mathrm{fit}}(x) = \exp\left(\frac{f_{\mathrm{obj}}^{\mathrm{best}}(x) - f_{\mathrm{obj}}(x)}{t} \right) \tag{3}$$

where $f_{\mathrm{obj}}(x)$ is the value of the objective function for each chromosome, $f_{\mathrm{obj}}^{\mathrm{best}}$ is the best value of the objective function, and t is the annealing temperature.

2.4. Transfer Learning Methods. In this manuscript, a novel Tran_NAS method based on the NAS algorithm is used for transfer learning from source domain datasets to target domain datasets. Transfer learning is an important research problem in ML and is the adaptation of prior learning experiences to new learning to improve the learning efficiency of new data.[38] The workflow of this Tran_NAS method is shown in the following:

First, using source domain data to develop a model, different blocks (preprocessing, dropout, softmax, etc.) are selected from the network state projections to form a network that has the best performance evaluation, and processing is repeated with a prespecified number of iterations. The final selected neural network structure is used as the transfer model. Second, the feature layer of the NAS model is frozen and transferred. Third, the last set of fully connected layers of the model is cut off and replaced with a new set of fully connected layers with random initialization. The weights of the frozen layer are not updated.

Next, we train the neural network in the target domain with a very small learning rate. Then, the rest of the network is thawed sequentially, and the optimal results are saved. Finally, the optimal classification results are obtained by comparison. The process is shown in Figure 4.

2.5. Transfer Experiments with Nutrient-Deficient Pear Leaves. To prove the feasibility of transfer learning, four experiments were designed, in which the data of healthy, N-deficient, and P-deficient pear leaves were set as the source

domain, and the data of healthy, Fe-deficient, and Mg-deficient pear leaves were set as the target domain. The source domain datasets are used to train the transferable model, and the target domain datasets are used to evaluate the model performance. SNV was selected for preprocessing, and NAS was selected as the modeling method. Experiments (1)−(4) were designed based on this, and the four experiments are as follows:

(1) The NAS model was trained on the target domain datasets and was tested on the target domain datasets. It was used to determine whether a good prediction model can be built under the condition of a small sample size.

(2) The NAS model was trained on the source domain datasets and was tested on the target domain datasets. It was used to investigate whether the model can be used directly in the target domain.

(3) The NAS model was trained on mixed data from the target domain and source domain datasets and was tested on the target domain datasets. It was used to determine whether the prediction accuracy of the model can be improved.

(4) The NAS model was trained on the source domain datasets and was transferred by Tran_NAS to fine-tune part of the target domain data, and the model was tested on the target domain datasets. It was used to determine whether the transfer learning model can significantly improve the prediction accuracy.

The accuracy of the transfer model is affected by many factors; therefore, this manuscript discusses it from the following two directions. First, the size of a sample from the mixed source and target domain dataset was investigated. Second, the structural difference was investigated. Different neural network architectures are obtained, and then, the transfer effects of different network structures are studied.

For an analysis of the performance of different models, evaluation metrics, including accuracy, confidence, precision (P) (eq 4), recall (R) (eq 5), and F1 value (F1) (eq 6), it is assumed that the sample has N labels, where $x_{i,j}(i, j = 1,2, ..., N)$ indicates the number of samples, in which label is i and the prediction is j. For neural networks, a reliable and practical measure to predict the confidence lever is essential. The most commonly used evaluation metric in classification modeling is accuracy, but for models with uneven data distributions, the F1 value is better than accuracy for evaluating the goodness of fit

of the model. However, in general, when evaluating good and bad models, both accuracy and the *F1* values are considered to ensure accuracy.

$$P = \frac{1}{N} \sum_{j=1}^{N} \frac{x_{i,j}}{\sum_{i=1}^{N} x_{i,j}} \tag{4}$$

$$R = \frac{1}{N} \sum_{i=1}^{N} \frac{x_{i,j}}{\sum_{j=1}^{N} x_{i,j}} \tag{5}$$

$$F1 = \frac{2 \times P \times R}{P + R} \tag{6}$$

In this chapter, a brief description of the methods used is provided. In Section 3, the results and discussion of preprocessing, modeling, and transferring are mainly described.

3. RESULTS AND DISCUSSION

3.1. Analysis of Pear Leaf Samples. A total of 1078 fresh leaf samples were collected in this experiment; 223 samples were healthy, 268 samples were mainly N-deficient, 325 samples were mainly P-deficient, 178 samples were mainly Fe-deficient, and 84 samples were mainly Mg-deficient, and they accounted for 20, 24, 30, 17, and 9% of the total number of samples, respectively. For researching the transfer learning experiment, the training set and test set are set to 3:7 and shown below (Figure 5).
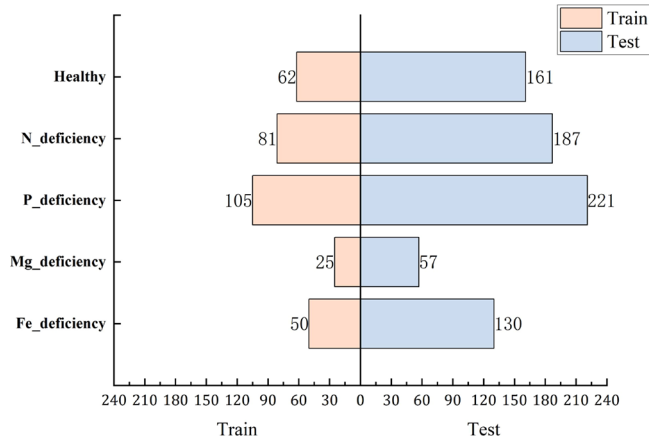

Figure 5. Sample statistics.

Then, the spectral data were preprocessed by the first-order derivatives (FD), second-order derivatives (SD), Savitzky−Golay (SG) smoothing, multiple scattering correction (MSC), standard normal variate (SNV), LG transform (LG), SG smoothing with multiple scattering correction (SG + MSC), and SG smoothing with standard normal variate (SG + SNV) because they are susceptible to interference from background, noise, and external environmental factors during measurement. To prevent visualizing the differences in the sample data after correction using different preprocessing methods, the transformed spectra were compared, and the results are shown in the following figures.

The spectral range of the leaves measured in this manuscript was 900−1700 nm (228 bands in total), where the average spectral reflectance of the samples in different categories is represented in Figure 6a, and its overall trend showed a consistent state. Figure 6b shows the raw spectra of 1078

samples, and the appearance of absorption bands in the NIR spectra is related to functional groups associated with water, proteins, and carbohydrates. There are four distinct peaks and valleys in the original spectrum, and the absorption at these wavelengths corresponds to O−H, C−H, N−H stretch first and second overtones, and combination bands that could be attributed to water absorption and protein changes.[39] N, P, and K are important components of organic and inorganic matter in plants. They are important element in the synthesis of proteins, amino acids, and carbohydrates. N and P are both components of organic matter and are important components of chlorophyll and protein, which can be detected by NIR due to the C−H bond in chlorophyll and the C−H and N−H bonds in proteins. The NIR regions had considerable influence on the spectra due to the strong relationship between trace minerals and other constituents, mainly with O−H tones (water) and with C−H combination tones (organic functional groups).[40] Figure 6c shows the log-transformed spectrum, which shows an opposite trend to the original spectrum due to the conversion of the reflectance to the absorbance. Figure 6d,e shows the spectra after FD and SD transformations, respectively, compared with the original spectra, and the difference in the spectra after the derivative transformations is more obvious. The reason is that derivatives can effectively eliminate baseline drift. Figure 6f shows that the spectra processed by SG are smoother than the original spectrum because SG effectively reduces the noise in the spectrum. Figure 6g,h shows that the spectra processed by multiple scattering correction (MSC) and standard normal variate (SNV) are more compact compared with the original spectrum. The effect of the two pretreatment methods on eliminating the uneven particle distribution is caused by the rough surface of pear leaves. The data after the hybrid SG smoothing with multiple scattering correction (SG + MSC), and SG smoothing with standard normal variate (SG + SNV) processing methods are shown in Figure 6i,j. Both figures contain the preprocessing results of SG, which are compared to the original spectrum; SG removes part of the noise and makes the results more compact as a whole.

Prior to modeling, the dataset was divided into a training set and a test set according to the Kennard−Stone algorithm[41] to show the classification ability of the model. To assess the relative robustness of the various preprocessing methods, modeling is performed next, and the modeling results are presented and analyzed in Section 3.2.

3.2. Evaluating Nutrient-Deficient Pear Leaves with Different Models. To further analyze the type of pear leaf deficiency, six modeling methods—the RF, XGBoost, AdaBoost, SVM, MLP, and NAS—were selected to build 54 prediction models with five-fold cross-validation training and different preprocessing methods. The evaluation analysis of each model is shown in Figure 7.

From the above figures, each model has poor classification ability on the data after the spectra were processed by SD, while other preprocessing methods improved the accuracy on the original data. This indicates that the spectra processed by SD still contain noise. The optimal preprocessing method is SNV, and the optimal value is 0.74 in the NAS model. The classification accuracy of SNV-processed spectral data with the RF, XGBoost, SVM, MLP, and NAS is better than the accuracy with other preprocessing methods (Table 2). The accuracy of most of the data processed using other preprocessing methods after AdaBoost modeling was less than 50%, but the
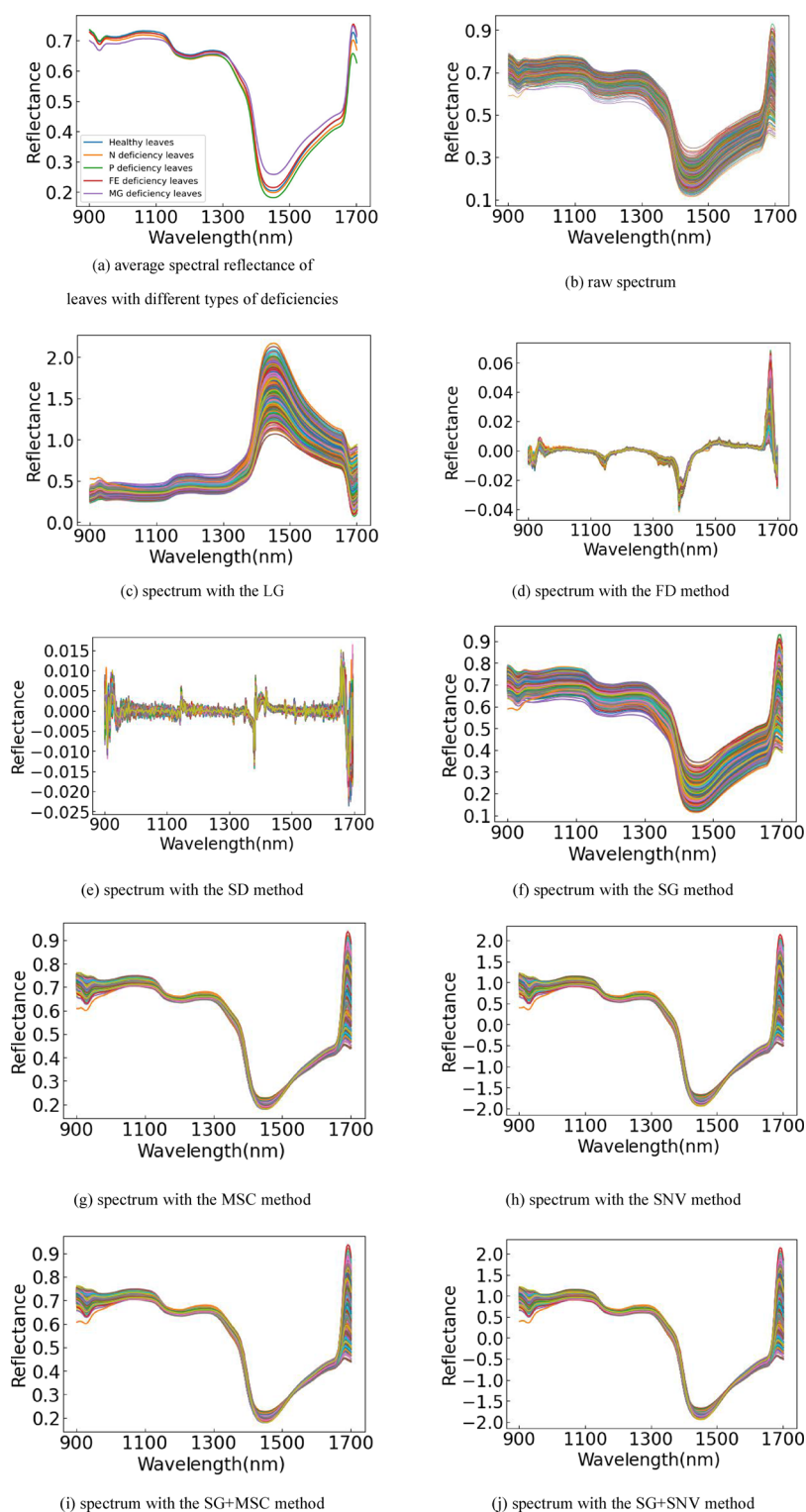
Figure 6. Preprocessed spectral curves. (a) average spectral reflectance; (b) raw spectrum; (c) spectrum with the LG; (d) spectrum with the FD method; (e) spectrum with the SD method; (f) spectrum with the SG method; (g) spectrum with the MSC method; (h) spectrum with the SNV method; (i) spectrum with the SGM + MSC method; (j) spectrum with the SG + SNV method.

classification accuracy of the data modeled after SNV processing reached 50%. This shows that preprocessing removes the main noise and background interference and improves the prediction precision and stability of the model. In summary, SNV was determined to be the best preprocessing method for this experiment. The classification effects of different modeling methods are shown in the following table.

The original data set was divided into the training set and the test set in a ratio of 3:7. The training set was then trained in a cross-validation under each model, and the test set was predicted. In order to ensure the stability of the model, this experiment chose the method of 5-fold cross-validation. In 5-fold cross-validation, the training set was divided into five subsets of equal size. Sequentially, one subset was tested using
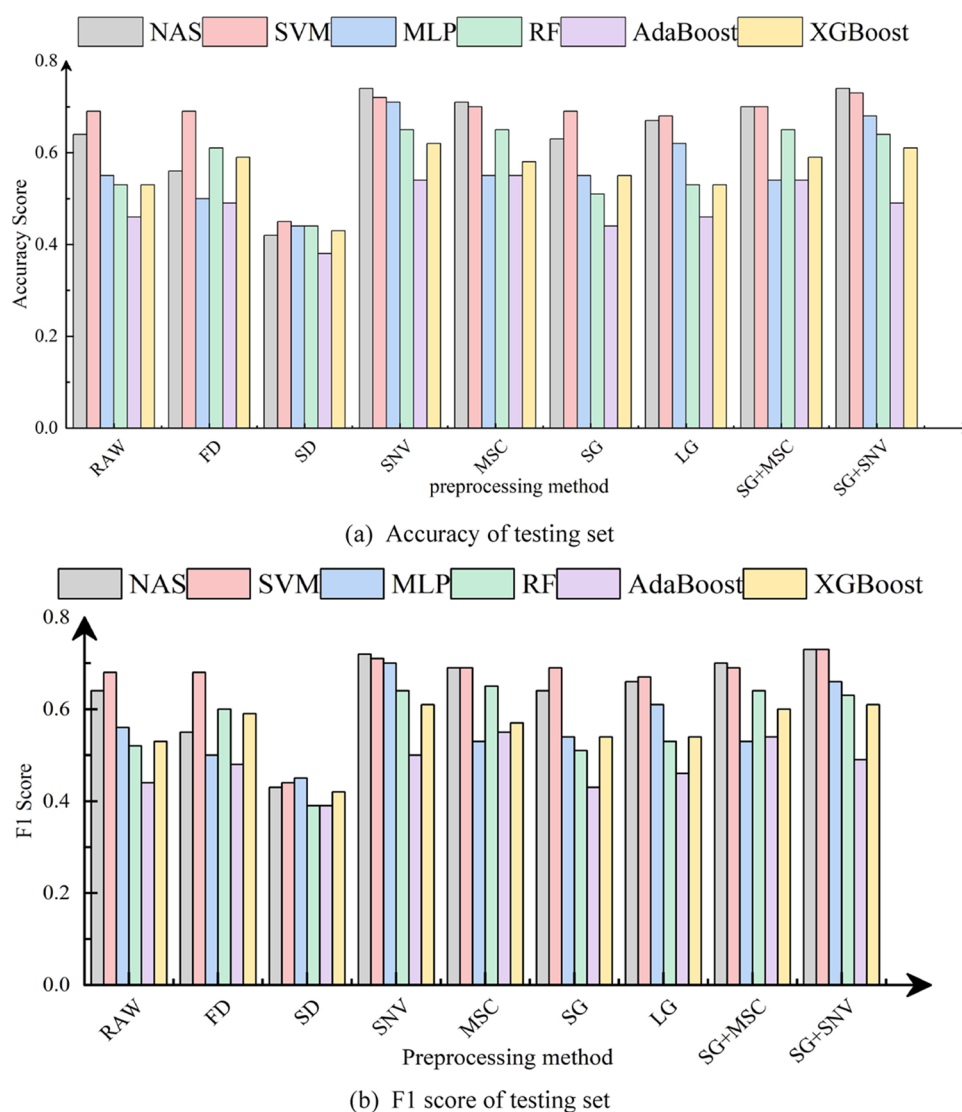
(a) Accuracy of testing set

(b) F1 score of testing set

Figure 7. Performance of different models with different preprocessing methods.

Table 2. SVM, RF, MLP, XGBoost, AdaBoost, and NAS Classification Performance of the Training Set and Test Set in the Full-Band NIR Spectrum[a]

modeling method	parameter	training set				test set			
		P^b (%)	R^b (%)	$F1^b$ (%)	accuracy^b (%)	P (%)	R (%)	$F1$ (%)	accuracy^b (%)
SNV-NAS	mls = 50, epochs = 300	96.94	95.72	96.27	96.59	73.01	72.11	72.49	74.34
SG + SNV-SVM	C = 610, gamma = 0.1	86.23	85.33	85.38	84.52	73.19	72.41	72.68	73.15
SNV-MLP	hidden_layer_sizes = [50,50]	76.44	77.23	76.48	76.47	70.86	69.93	70.13	70.50
SNV-RF	ndt = 390, d = 15	86.00	83.13	84.29	85.14	64.84	64.59	64.46	65.21
SNV-XGBoost	nwl = 220, learning rate = 0.03	98.69	97.41	98.00	98.14	61.50	60.96	61.06	61.77
SG + MSC-AdaBoost	nwl = 290, learning rate = 0.5	71.54	69.78	70.31	70.28	57.84	52.57	54.44	54.37

[a]SNV−NAS: The data are preprocessed by SNV and then modeled by NAS. All the methods in the above table include the modeling method on the left and the preprocessing method on the right. [b]The results of the five-fold cross validation.

the classifier trained on the remaining four subsets. This process was repeated for each one of the subsets. Then, the average of the five cross-validation results is taken as the final result. After cross-validation, SNV combined with NAS modeling achieves the best result of 74.34%. Therefore, the modeling is all processed using cross-validation. The following analysis is specific to each model.

The kernel of SVM modeling is polynomial(poly), and C and gamma are hyperparameters. The C parameter adds a

penalty for each error, gamma affects the kernel function, and this manuscript uses a grid search to optimize the hyperparameters. The best classification accuracy of SVM is 74.34%. RF is based on a decision tree-based learner with the addition of random attribute selection. The depth (d) and the number of decision trees (ndt) are hyperparameters. The larger d is, the better the fit. The greater ndt is, the lower the generalization error. The best classification accuracy of RF is 65.21%, and the prediction is relatively good. The results show that the

precision, recall, accuracy and $F1$ scores of RF are all smaller than those of SVM. In addition, the modeling accuracy of SVM is close to the prediction accuracy, indicating that the SVM model has strong generalization ability and stability.

XGBoost and AdaBoost are two kinds of integrated algorithms; the learning rate and the number of weak learners (nwl) are hyperparameters, and the learning rate and nwl determine the fit of the algorithm at the same time. The integrated algorithm performed poorly because of the unbalanced distribution of the pear leaf deficiency datasets. Among them, the classification effect of AdaBoost is 54.37%, and the classification accuracy of XGBoost is 61.77%. The classification effect of XGBoost is relatively better compared with that of AdaBoost because XGBoost adds regularization on top of it to improve the generalization ability of the model, and finally, its results are relatively better.

MLP and NAS are both neural network algorithms. MLP needs to set the number of neurons in the hidden layer and the number of layers (hidden_layer_sizes), both of which affect the complexity of the model and are crucial to the neural network. The MLP effect is better, and the accuracy can reach up to 70.50%. NAS is an automatic search, and the initial number of models (mls) and the number of iterations (epochs) are set in this experiment. The classification accuracy of NAS reaches 74.34%. Because MLP is a traditional neural network, it requires the manual setting of parameters, such as the number of layers and number of neurons, which requires much empirical knowledge, while NAS uses automatic exploration to find optimal parameters, which can eventually lead to better classification results.

The above figure also shows that the best modeling method is NAS, and the highest classification accuracy is achieved by using SNV preprocessed data to classify the samples obtained after NAS modeling. The NAS architectures employed are very suited for modeling NIR spectra. The current structure of the NAS architecture fosters the modeling of the spectra as continuous data series. These models are particularly suitable for NIR spectra, which are typically smooth functions with broad features. NAS selects the parameters by itself and remembers the advantages of the trained model during modeling, so it can inherit these advantages in the next training phase, and the final classification results obtained by NAS are the best (Figure 8).

However, its accuracy only reached 74.34%. To explore the reason, the classification results of various categories in NAS and the results are shown in Table 3. The confidence level of identifying healthy individuals in the best model is 0.15, the confidence level of identifying an N deficiency is 0.15, the confidence level of identifying a P deficiency is 0.28, the confidence level of identifying Fe deficiency is 0.10, and the confidence level of identifying an Mg deficiency is 0.05.

The confidence levels for the Fe deficiency and Mg deficiency were the lowest. It is possible that the number of samples is too low, resulting in a lower identification accuracy for both. The model developed by NAS predicts the NIR spectra, which is feasible for improving the accuracy of Fe- and Mg-deficient identification single classes and building a model with trace element data.

3.3. Transfer Experiment from N- and P- to Fe- and Mg-Deficient Pear Leaves. In the previous subsection, it was found that the confidence levels of Fe and Mg were low in the model trained on the five types of data. Then, it is significant for exploring to improve the confidence of Fe and
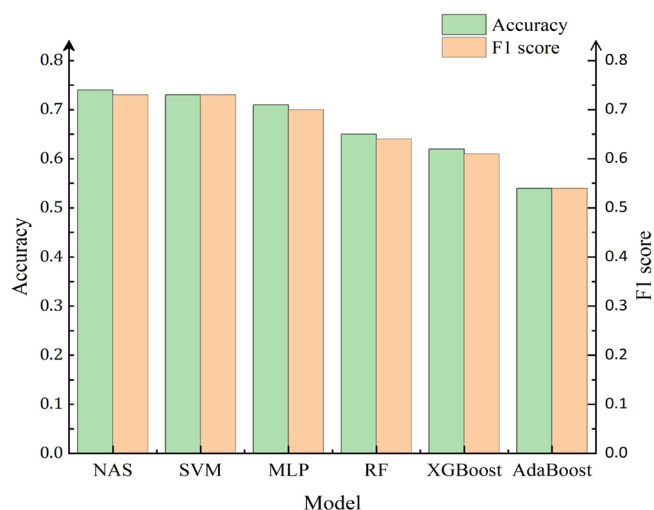
Figure 8. SVM, RF, MLP, XGBoost, AdaBoost, and NAS classification performance.

Table 3. SNV-NAS Classification Performance for Each Deficiency Type

modeling method	type	test set			
		P	R	$F1$	confidence
SNV-NAS	healthy	76.00%	70.81%	73.31%	15.07%
	N deficiency	64.86%	64.17%	64.52%	15.87%
	P deficiency	88.70%	95.93%	92.17%	28.04%
	Fe deficiency	58.14%	57.69%	57.92%	9.92%
	Mg deficiency	77.36%	71.93%	74.55%	5.42%

Mg deficiency by building model with trace element deficiency. Therefore, the trace element data were processed by SNV, and the model was built by NAS. This experiment analysis sets the trace element training set and the test set ratio to 3:7. The detailed results are shown in the following table (Table 4); hopefully, the single class confidence level will improve with a small amount of data.

Table 4. Analyzing the Performance of the SNV-NAS Model on the Target Domain Dataset

modeling method	deficiency type	test set			
		P	R	$F1$	confidence
3/10tar → Direct 7/10tar SNV-NAS	healthy	81.08%	77.42%	79.21%	35.29%
	Fe deficiency	78.91%	82.11%	80.48%	29.10%
	Mg deficiency	71.88%	74.19%	73.02%	13.53%

From the above table, the confidence level of Fe and Mg deficiencies improved by using the NAS method. However, the overall accuracy only reached 78.53%. Overall, $F1$ only reached 77.57%. To improve the accuracy and confidence of trace elements, transfer learning training is carried out in this manuscript.

Although the confidence of a single class was improved after the modeling of trace elements, the overall classification accuracy showed a decreasing trend. The data were partitioned into source domain data (src), which included N deficiency

data, P deficiency data, healthy data, and target domain data (tar), which included Fe deficiency data, Mg deficiency data, and healthy data. In the procedure, four experiments are designed to verify the feasibility of transfer learning. The following table shows the experimental results of the five experiments. From Table 6, the accuracy of experiment (1) is only 78.53%, indicating that the modeling accuracy when there are few data is worse. The accuracy of experiment (2) is only 60%, indicating that the source domain model cannot be directly applied to the target domain. The classification result of experiment (3) is 76.34% and experiment (4) is 63.29%, while the classification accuracy of experiment (5) is 89.12%, and it is the best (Table 5).

Table 5. Comparison of the Performance of Transfer Learning Models

| experiment | number of samples | | evaluation indicator | |
	training set	test set	F1 (%)	accuracy (%)^a
① 3/10tar →direct 7/10tar	146	339	77.57	78.53
② 3/10src →direct 7/10tar	247	339	45.20	60.00
③ 3/10src + 3/10tar →direct 7/10tar	392	339	71.94	76.34
④ 3/10src + 3/10tar →TCA_NAS 7/10tar	392	339	63.18	63.29
⑤ 3/10src + 3/10tar →Tran_NAS 7/10tar	392	339	87.35	89.12

^{a}The results of the five-fold cross validation

Experiment (2) shows that model failure will occur. Although NAS learns the same spectral characteristics, the model cannot learn the relevant characteristics of trace elements, resulting in the model being trained with source domain data that cannot be applied directly to the target domain. Experiment (3) shows a decreasing trend in accuracy. This is because the training data that include the source domain are too large, which leads to model bias and makes the final classification effect not very good. The accuracy of experiment (5) is significantly improved compared with those of the previous experiments. The prediction accuracy of experiment (1) was also significantly lower than that of experiment (5). The model developed with macroelements based on Tran_NAS can be used to predict trace elements. The spectral bands at 1100−1250 and 1650−1700 nm to identify the lack of N, P, Fe, and Mg are coincident in NIR spectroscopy (Figure 6.). Sharing layers with the source task reduces overfitting and promotes better reuse of features extracted from the source task. Then, the rest of the full model is fine-tuned by using the best hyperparameters. Therefore, the classification performance of experiment (5) was significantly improved. As a transfer learning algorithm for feature transformation, TCA was performed on the NIR feature matrixes of the source domain and target domain. Then, NAS algorithm was used to construct transfer learning models using the transformed features of the source domain. The accuracy of TCA_NAS models was 63.29%, and the results are much lower than Tran_NAS. The reason is that fine-tuning can better overcome the variability between different datasets.[42] In summary, the model developed by Tran_NAS is feasible for pear leaf trace element deficiency. The confidence results for the single class are shown in Table 6.

Table 6. Performance of Trace Element Deficiency Types by Tran_NAS

| modeling method | deficiency type | test set | | | |
		P	R	F1	confidence
3/10src + 3/10tar → Tran_NAS 7/10tar SNV-NAS	healthy	92.26%	92.26%	92.26%	42.06%
	Fe deficiency	90.16%	89.43%	89.80%	32.35%
	Mg deficiency	79.37%	80.65%	80.00%	14.71%

In Figure 9, the confidence level of identifying healthy individuals is 0.42, the confidence level of identifying an Fe
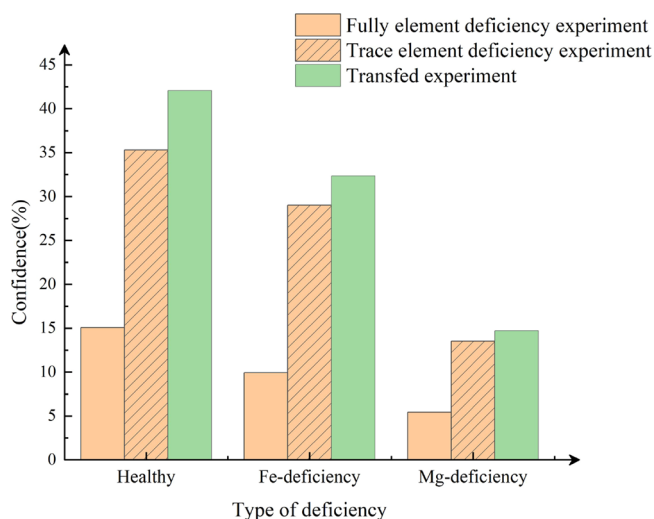

Figure 9. Confidence levels of trace element deficiency under three experiments.

deficiency is 0.32, and the confidence level of identifying an Mg deficiency is 0.14. The level for Fe improved by 3.25%, the level for Mg improved by 1.18%, and the overall accuracy improved by 10.59% compared with those values before the transfer. The shared features of this section are selected by NAS to develop the model. Quick learning of relevant features is performed by transfer learning. The specific trend is shown in the following chart.

In Figure 9, fully element deficiency experiment stands for the model developed by NAS with five types of element deficiencies, trace element deficiency experiment stands for the model developed by NAS only with trace element deficiencies, and transfed experiment stands for the model developed by Tran_NAS with trace element deficiencies.

The results show that the confidence of the model developed by Tran_NAS with trace element deficiencies is improved compared with the model developed by NAS with five types of element deficiencies and trace element deficiencies. The experiment shows that transfer learning approaches perform better than ML models for target domain datasets. Compared with training from scratch, fine-tuning a pretrained neural network on a target dataset can significantly improve performance while reducing the target labeled data requirements.

In this manuscript, the source domain training set and the test set are divided in a ratio of 3:7, and the initial hyperparameters (max_trials = 50 and epochs = 500) are set. The top ten neural network structures in terms of accuracy

a. The loss value of the training set

b. The loss value of the test set
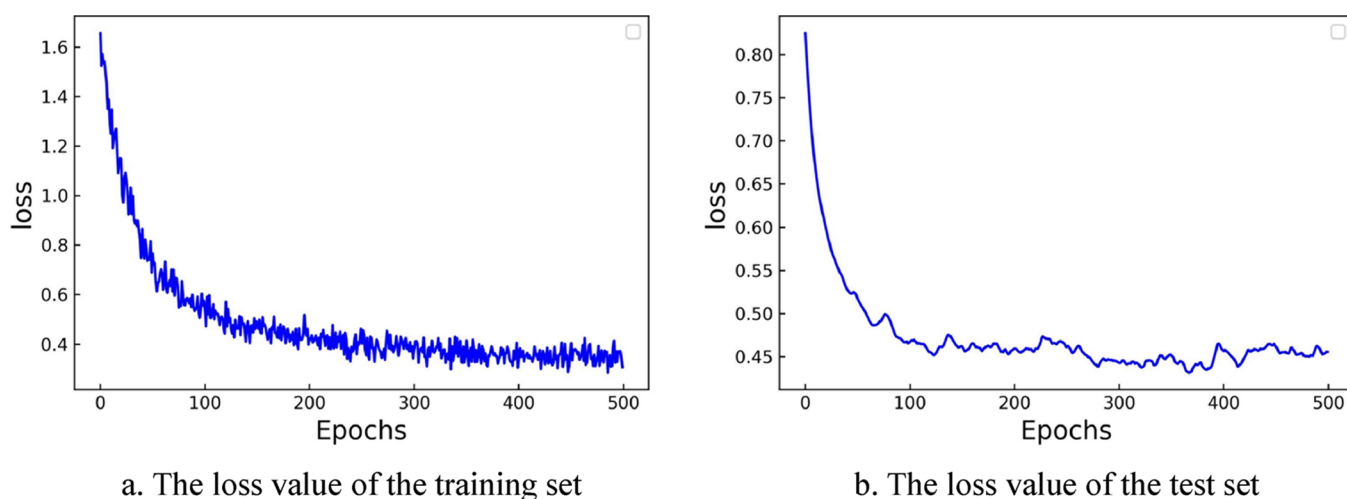
Figure 10. Plot showing loss in the training and test datasets.

ranking are identified as the subnet architecture. The target domain is divided into a training set and a test set at a ratio of 3:7. The pretrained model is retrained on the target domain dataset, and the weight values of the model are updated according to a new task. A fine-tuning process occurs on the network using backpropagation with labels. Finally, the nutrient deficiency type of pear leaves was determined. The fine-tuning operation is made by removing the last two layers and replacing them with new layers. These layers are fully connected softmax layers. The number of outputs of a fully connected layer should be equal to the number of classes counts in the training dataset. The softmax layer is the same as the previous layers. The NAS model with new layers is trained and updated according to the target domain dataset. The next layer is unfrozen when the neural network is trained to convergence until an optimal result is found, as shown in the figure below.

Figure 10 shows that when the number of iterations reaches 300, the neural network starts to converge. The changes in the loss values of the training set and test set tend to be consistent, and the model has relatively better generalization ability. In this section, the model developed with macroelements based on Tran_NAS can be used to predict trace element deficiencies. The architecture of the neural network also plays a role in the recognition accuracy, which is further discussed in Section 3.4.

3.4. Further Discussion of the Tran_NAS Method. In this manuscript, we further discuss the proposed method from the following two directions. First, the size of the sample, which is a mixture of the source and target domain datasets, is discussed. Second, the structural difference is discussed. The source domain is divided into training and test sets using a ratio of 3:7, and the target domain was divided into training and test sets using ratios of 5:5, 4:6, 3:7, 2:8, and 1:9 for the experimental comparison (Table 7).

The result shows that all accuracy exceeds 70%. The results are 71.40 and 81.70% when the dataset is divided with 1:9 and 2:8 ratios, respectively. However, the prediction accuracy of the model developed with the target domain data is 78.53%, and the accuracy of the model developed with data divided with 1:9 and 2:8 ratios is not meaningful. The accuracy of the model developed with data divided with a 3:7 ratio is 89.12%. The results are 90.72 and 92.59% when the dataset is divided

Table 7. Dataset Division and Performance of the Target Domain[a]

experiment	sample number			evaluation indicator	
	training set	test set	proportion $(a{:}b)$	F1 (%)	accuracy (%)
3/10src + a/10tar → Tran_NAS b/10tar	296	436	1:9	66.69	71.40
	344	388	2:8	79.85	81.70
	392	339	3:7	87.35	89.12
	194	291	4:6	89.40	90.72
	242	243	5:5	90.90	92.59

[a]In the table, a represents the size of the training set in the target domain, and b represents the size of the test set in the target domain.

with 4:6 and 5:5 ratios, respectively. The results show that the larger the data are, the better the modeling effect. However, it is very difficult to obtain a large amount of trace element deficiency samples and its cost is high. Therefore, the accuracy of the model developed with target domain datasets divided into a ratio of 3:7 is better. Using a small amount of data to build the model will result in low costs.

In this manuscript, based on the model developed with the target domain dataset, which is divided into a 3:7 ratio, different kinds of neural architectures are explored to find a more suitable transfer of pear leaf deficiency data.

The structure of the neural network has a few changes every search iteration, and the purpose of this manuscript is to explore the model best suited for transfer learning. The difference in the optimizer, preprocessing, and other modules in the neural network leads to differences in accuracy. The following table shows the specific components of the neural architecture of each model (Table 8):

By comparing the accuracy and F1 values of the above 10 models, the NAS_Model_1 model has good performance with an accuracy of 89.12% and an F1 value of 87.35%. The network structure consists of three layers: the first layer is the input layer, which contains data input and data preprocessing layers; the second layer is the dense layer, which includes batch normalization (BN), rule, and dropout layers; and the last layer is the output layer, which is composed of dense and softmax layers.
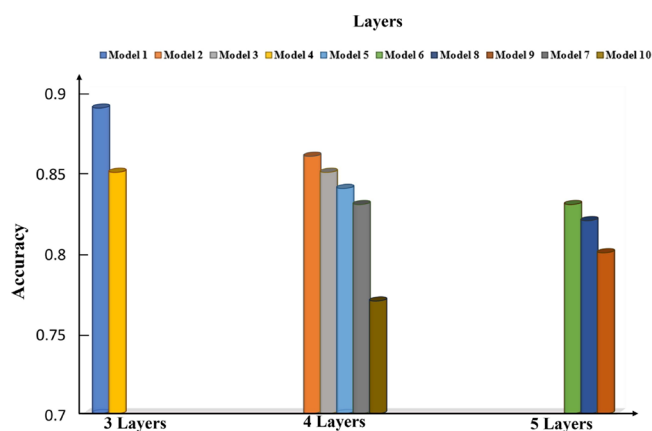
In this manuscript, the modules owned by each of the ten models in the tree search space are shown (Figure 11).

Table 8. Performance of Different Neural Network Architectures

neural networks	layer	dropout	batch normalization	evaluation indicators	
				F1	accuracy
NAS_Model_1 (model 1)	3	1 layer	1 layer	87.35%	89.12%
NAS_Model_2 (model 2)	4	none	2 layers	84.54%	86.47%
NAS_Model_3 (model 3)	4	1 layer	2 layers	85.00%	85.59%
NAS_Model_4 (model 4)	3	none	1 layer	82.54%	85.00%
NAS_Model_5 (model 5)	4	2 layers	2 layers	81.64%	84.41%
NAS_Model_6 (model 6)	5	3 layers	3 layers	80.25%	83.53%
NAS_Model_7 (model 7)	4	1 layer	None	81.49%	83.52%
NAS_Model_8 (model 8)	5	4 layers	3 layers	80.15%	82.35%
NAS_Model_9 (model 9)	5	3 layers	2 layers	72.64%	80.58%
NAS_Model_10 (model 10)	4	2 layers	none	70.01%	77.65%

**Figure 12.** NAS model performance with different layers.

The result shows that the search space of the tree includes three-layer, four-layer, and five-layer structures, and most of the better models use the four-level structure and a few use the three-level structure. Most NAS methods are based on the a priori knowledge of human experts.[43] Considering the model of the architecture as a fixed element, the depth and width of the architecture usually affect the performance of the architecture. To obtain an architecture with better transmission performance, the size of the architecture is searched by manipulating the width factor. By comparing the different preprocessing, optimizer, and number of network layer options, the result shows which kind of neural network architecture can obtain the most stable transfer results.

The result shows that the optimal classification model has a three-layer structure, and the others have four-layer structures (Figure 12). Although the number of models in the three-layer

category is less, its classification effect on transfer learning is relatively strong. The four-layer structure contains poorly performing models. Among them, NAS_Model_1 is the same as NAS_Model_3 except for the different number of layers, but the NAS_Model_1 classification results are higher than those of NAS_Model_3. For neural networks, the bottom layer of the network learns some general features, the middle and top layers of the network learn abstract high-level semantic features, and the transfer ability of features decreases as the number of layers increases.[42] In summary, a three-layer architecture is the most suitable neural network architecture for the Tran_NAS method with the spectra dataset of pear leaves.

The results of the analysis of BN are shown in Figure 13. The better models of NAS_Model_1 to NAS_Model_6 have BN, while the models of NAS_Model_7 to NAS_Model_10 with relatively poor prediction results do not have BN. BN ensures the expressiveness of the input data to some extent, and BN results in a more stable distribution of the input data at each layer and has some regularization effect.[44] In summary, batch processing is necessary in neural networks.
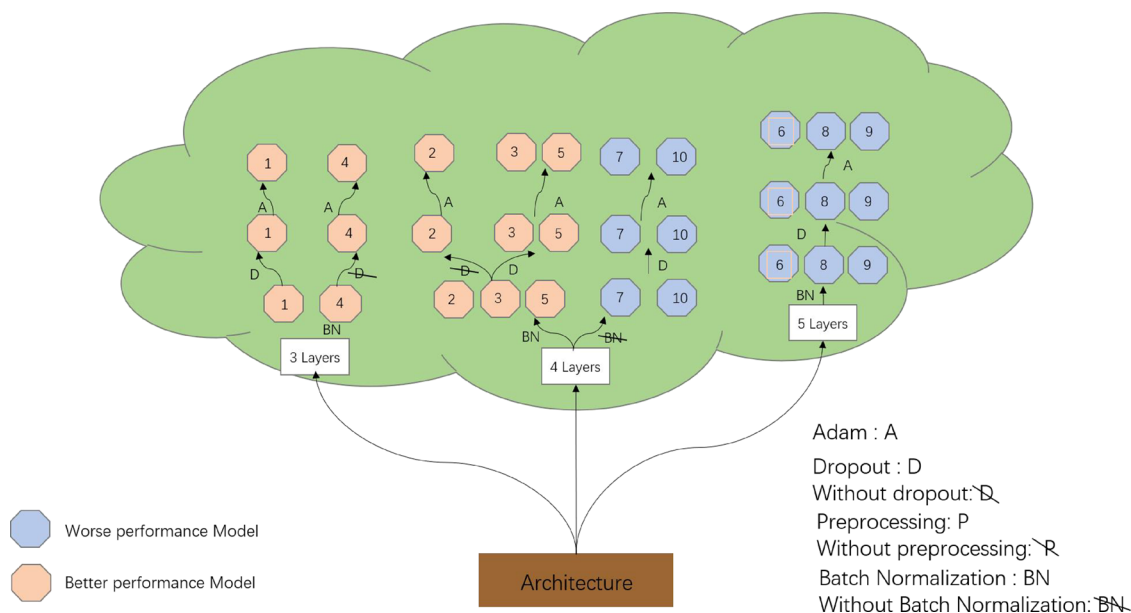
Figure 11. Tree search space of the NAS model. Orange graphs 1−5 represent better models and blue graphs 6−10 represent less accurate models.
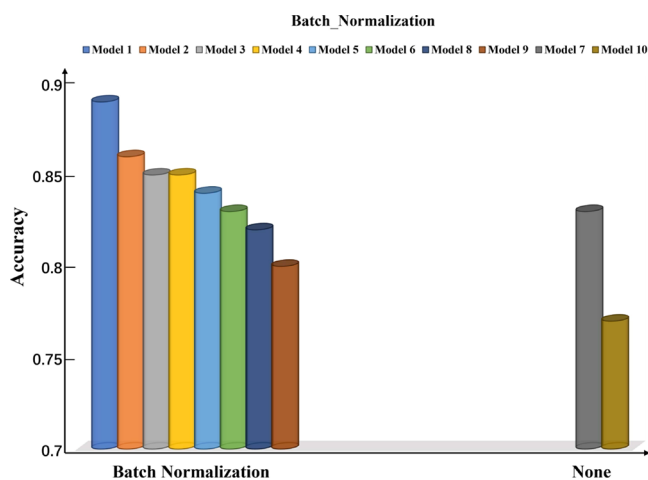
Figure 13. NAS model performance of the batch normalization layer.

Finally, the dropout layers in the neural network are compared (Figure 14). NAS_Model_2 and NAS_Model_4 do
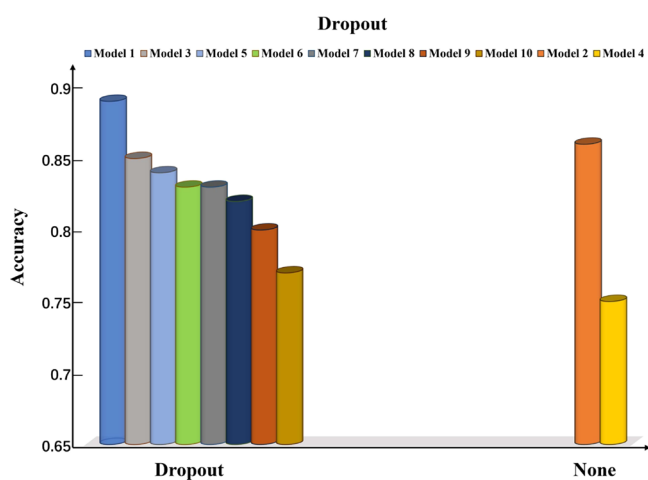

Figure 14. NAS model performance of the dropout layer.

not have dropout layers; the rest of the models have dropout layers. The above table shows that NAS_Model_2 and NAS_Model_4 have relatively good prediction results, which means that the dropout layer has little effect on the structure. However, NAS_Model_8 has a dropout layer, and the results are still poor because too much regularization has a negative impact on the training dynamics, especially in the early stage.[45] In summary, data preprocessing is crucial in neural networks,[46] and the correct preprocessing approach usually contributes to the results.

In summary, this manuscript analyses the number of layers, normalization approach, and dropout layers. The neural structure is related to the number of parameters, and the number of parameters varies with the neurons.[47] However, too many neurons can lead to an overfitting situation; in contrast, too few neurons can limit the ability of the neural network to model the mechanism, leading to over generalization.[48] Finally, the most suitable modeling features for transfer learning are derived by comparison (Figure 15).

From what has been discussed above, the Tran_NAS model is feasible for identifying pear leaves with Fe and Mg deficiency. However, using Tran_NAS to identify other trace
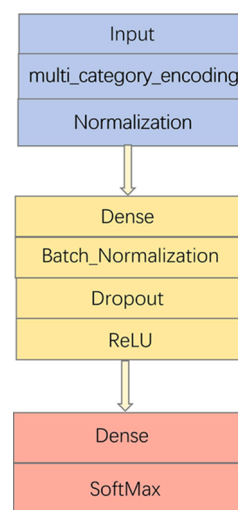

Figure 15. Optimal neural network structure.

element deficiencies requires the following three considerations. First, it can be seen from Table 7 that the sample size is larger, the transfer effect is better. Therefore, for other trace elements, such as Cu, Zn, and Mn, at least the same sample size in this manuscript should be achieved when the Tran_NAS method is used to fit the normal effect. Second, in this manuscript, the micro NIR spectrometer is small and portable, but its resolution is low. The higher content of nutrient elements in leaves can be better identified. For trace elements with low content, the spectral recognition effect is poor, so the transfer effect of NAS will also be reduced. Higher-resolution spectroscopic instruments can be used when identifying nutrients at lower levels, and their transfer identification may be more satisfactory. Finally, the results in Table 8 show that the depth and width of the architecture usually affect the performance of the architecture. For the transfer of other trace elements, the time of model training can be extended so that NAS can fully search for the most suitable framework for transfer. This will also help obtain the desired recognition results. In the future, we will explore further the transfer of other elements.

4. CONCLUSIONS

This manuscript investigates that diagnosing the N, P, Fe, and Mg deficiencies of pear tree leaves based on NIR reflectance spectra is feasible. The relevant results show that the confidence of trace element deficiencies is relatively poor because of the lack of samples. In this research, a feature-based homogenous transfer learning approach named Tran_NAS was explored for the classification domain to improve the identification accuracy of trace element deficiencies by the advantage of NAS and fine-tuning methods. The relevant results show that the novel learned Tran_NAS model (89.12%) outperformed the traditional NAS model (78.53%) by 10.59%. To evaluate the influence of different searching neural networks, the characteristics of efficient architecture searches are explored for ten different neural network models. The results show that a three-layer neural network with a preprocessing layer, BN layer, and dropout layer is optimal. In the future, the Tran_NAS method may be used for the fast transfer of structured data with a small spectral dataset, and it can effectively improve the classification performance. However, it is also noted that transfer learning still depends

on people's empirical knowledge, and we leave this research thread for future work.

■ AUTHOR INFORMATION

Corresponding Authors

Yuan Rao − *Anhui Province Key Laboratory of Smart Agricultural Technology and Equipment and College of Information and Computer Science, Anhui Agricultural University, Hefei 230001, China*; Email: raoyuan@ahau.edu.cn

Li Liu − *Anhui Province Key Laboratory of Smart Agricultural Technology and Equipment and College of Horticulture, Anhui Agricultural University, Hefei 230001, China*; Email: liulireal@ahau.edu.cn

Authors

Xiu Jin − *Anhui Province Key Laboratory of Smart Agricultural Technology and Equipment and College of Information and Computer Science, Anhui Agricultural University, Hefei 230001, China*; ⊙ orcid.org/0000-0003-3017-7328

Wenjing Ba − *Anhui Province Key Laboratory of Smart Agricultural Technology and Equipment and College of Information and Computer Science, Anhui Agricultural University, Hefei 230001, China*

Lianglong Wang − *Anhui Province Key Laboratory of Smart Agricultural Technology and Equipment and College of Information and Computer Science, Anhui Agricultural University, Hefei 230001, China*

Tong Zhang − *Anhui Province Key Laboratory of Smart Agricultural Technology and Equipment and College of Information and Computer Science, Anhui Agricultural University, Hefei 230001, China*

Xiaodan Zhang − *Anhui Province Key Laboratory of Smart Agricultural Technology and Equipment and College of Information and Computer Science, Anhui Agricultural University, Hefei 230001, China*

Shaowen Li − *Anhui Province Key Laboratory of Smart Agricultural Technology and Equipment and College of Information and Computer Science, Anhui Agricultural University, Hefei 230001, China*

Complete contact information is available at:
https://pubs.acs.org/10.1021/acsomega.2c03596

Author Contributions

[#]X.J. and W.B. contributed equally to this work

Notes

The authors declare no competing financial interest.

■ REFERENCES

(1) Fu, H.; Ma, Q.; Ma, Z.; Hu, Y.; Liu, F.; Chen, K.; Pan, W.; Tang, S.; Zhang, X.; Wu, L. Quantifying Key Internal and External Yield-limiting Factors for Chinese Pear in Smallholder Dominant Areas. *Hortscience* **2021**, *56*, 1395−1401.

(2) Sete, P. B.; Comin, J. J.; Nara Ciotta, M.; Almeida Salume, J.; Thewes, F.; Brackmann, A.; Toselli, M.; Nava, G.; Rozane, D. E.; Loss, A.; et al. Nitrogen fertilization affects yield and fruit quality in pear. *Sci. Hortic.* **2019**, *258*, No. 108782.

(3) Aruani, M. C.; Reeb, P. D.; Barnes, N. E. Influence of soil properties on yield and fruit maturity at harvest of 'Williams' pear. *Chil. J. Agric. Res.* **2014**, *74*, 460−467.

(4) Mu, X.; Chen, Y. The physiological response of photosynthesis to nitrogen deficiency. *Plant Physiol. Biochem.* **2021**, *158*, 76−82.

(5) Sorrenti, G.; Toselli, M.; Marangoni, B. Use of compost to manage Fe nutrition of pear trees grown in calcareous soil. *Sci. Hortic.* **2012**, *136*, 87−94.

(6) Zhao, Z.; Wang, P.; Jiao, H.; Tang, C.; Liu, X.; Jing, Y.; Zhang, S.; Wu, J. Phylogenetic and expression analysis of the magnesium transporter family in pear, and functional verification of PbrMGT7 in pear pollen. *J. Hortic. Sci. Biotechnol.* **2018**, *93*, 51−63.

(7) Park, B. B.; Park, G. E.; Bae, K. Diagnosis of plant nutrient and growth responses on fertilization with vector analysis and morphological index. *For. Ind. Sci. Technol.* **2015**, *11*, 1−10.

(8) Osco, L. P.; Ramos, A. P. M.; Moriya, É. A. S.; de Souza, M.; Junior, J. M.; Matsubara, E. T.; Imai, N. N.; Creste, J. E. Improvement of leaf nitrogen content inference in Valencia-orange trees applying spectral analysis algorithms in UAV mounted-sensor images. *Int. J. Appl. Earth Obs. Geoinf.* **2019**, *83*, No. 101907.

(9) Xiaobo, Z.; Jiewen, Z.; Povey, M. J.; Holmes, M.; Hanpin, M. Variables selection methods in near-infrared spectroscopy. *Anal. Chim. Acta* **2010**, *667*, 14−32.

(10) Lin, H.; Ying, Y. Theory and application of near infrared spectroscopy in assessment of fruit quality: a review. *Sens. Instrum. Food Qual. Saf.* **2009**, *3*, 130−141.

(11) Whittier, W. A.; Hodge, G. R.; Lopez, J.; Saravitz, C.; Acosta, J. J. The use of near infrared spectroscopy to predict foliar nutrient levels of hydroponically grown teak seedlings. *J. Near Infrared Spectrosc.* **2021**, *29*, 301−310.

(12) Bedin, F. C. B.; Faust, M. V.; Guarneri, G. A.; Assmann, T. S.; Lafay, C. B. B.; Soares, L. F.; de Oliveira, P. A. V.; dos Santos-Tonial, L. M. NIR associated to PLS and SVM for fast and non-destructive determination of C, N, P, and K contents in poultry litter. *Spectrochim. Acta, Part A* **2021**, *245*, No. 118834.

(13) Jamil Maia, A.; Cabral Nascimento, R.; da Silva, Y. J. A. B.; do Nascimento, C. W. A.; de Sousa Mendes, W.; Veras Neto, J. G.; de Araújo Filho, J. C.; Tiecher, T.; da Silva, Y. J. A. B. Near-infrared spectroscopy for prediction of potentially toxic elements in soil and sediments from a semiarid and coastal humid tropical transitional river basin. *Microchem. J.* **2022**, *179*, No. 107544.

(14) Huang, Q.; Chen, Q.; Li, H.; Huang, G.; Ouyang, Q.; Zhao, J. Non-destructively sensing pork's freshness indicator using near infrared multispectral imaging technique. *J. Food Eng.* **2015**, *154*, 69−75.

(15) Sathyavani, R.; JaganMohan, K.; Kalaavathi, B. Classification of nutrient deficiencies in rice crop using denseNet-BC. *Mater. Today: Proc.* **2022**, *56*, 1783−1789.

(16) Pan, S. J.; Yang, Q. A survey on transfer learning. *IEEE Trans. Knowl. Data Eng.* **2010**, *22*, 1345−1359.

(17) Guerrero, R.; Renteros, B.; Castañeda, R.; Villanueva, A.; Belupú, I.Detection of nutrient deficiencies in banana plants using deep learning. In *2021 IEEE International Conference on Automation/XXIV Congress of the Chilean Association of Automatic Control (ICA-ACCA)*; IEEE, 2021; pp 1−7.

(18) Wan, L.; Zhou, W.; He, Y.; Wanger, T. C.; Cen, H. Combining transfer learning and hyperspectral reflectance analysis to assess leaf nitrogen concentration across different plant species datasets. *Rem. Sens. Environ.* **2022**, *269*, No. 112826.

(19) Weiss, K.; Khoshgoftaar, T. M.; Wang, D. A survey of transfer learning. *J. Big Data* **2016**, *3*, 1−40.

(20) Farahani, A.; Pourshojae, B.; Rasheed, K.; Arabnia, H. R.A Concise Review of Transfer Learning. In *2020 International Conference on Computational Science and Computational Intelligence (CSCI)*; IEEE, 2020; pp 344−351.

(21) Zhang, H.; Silva, F. H.; Ohata, E. F.; Medeiros, A. G.; Rebouças Filho, P. P. Bi-dimensional approach based on transfer learning for alcoholism pre-disposition classification via EEG signals. *Front. Hum. Neurosci.* **2020**, *14*, No. 00365.

(22) Wang, H.; Pampati, N.; McCormick, W. M.; Bhattacharyya, L. Protein nitrogen determination by kjeldahl digestion and ion chromatography. *J. Pharm. Sci.* **2016**, *105*, 1851−1857.

(23) Ma, B.-L.; Zheng, Z. Nutrient uptake of iron, zinc, magnesium, and copper in transgenic maize (Zea mays) as affected by rotation

systems and N application rates. *Nutr. Cycl. Agroecosyst.* **2018**, *112*, 27−43.

(24) Reuter, D.; Robinson, J. B.Plant analysis: an interpretation manual. In *Tree Physiology*, 2nd ed.; Reuter, D. J., Ed.; Academic Press: CSIRO publishing, 1997; pp 11−18.

(25) Savitzky, A.; Golay, M. J. Smoothing and differentiation of data by simplified least squares procedures. *Anal. Chem.* **1964**, *36*, 1627−1639.

(26) Zhao, N.; Wu, Z.; Cheng, Y.; Shi, X.; Qiao, Y. MDL and RMSEP assessment of spectral pretreatments by adding different noises in calibration/validation datasets. *Spectrochim. Acta, Part A* **2016**, *163*, 20−27.

(27) Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5−32.

(28) Ho, T. K. The random subspace method for constructing decision forests. *IEEE Trans. Pattern Anal. Mach. Intell.* **1998**, *20*, 832−844.

(29) Schapire, R. E.The boosting approach to machine learning: An overview. In Nonlinear estimation and classification; Springer: New York, NY, 2003; 149−171.

(30) Hearst, M. A.; Dumais, S. T.; Osuna, E.; Platt, J.; Scholkopf, B. Support vector machines. *IEEE Intell Syst. App.* **1998**, *13*, 18−28.

(31) Li, J.; Sun, L.; Li, R. Nondestructive detection of frying times for soybean oil by NIR-spectroscopy technology with Adaboost-SVM (RBF). *Optik* **2020**, *206*, No. 164248.

(32) Castillo, P. A.; Carpio, J.; Merelo, J. J.; Prieto, A.; Rivas, V.; Romero, G. Evolving Multilayer Perceptrons. *Neural Process. Lett.* **2000**, *12*, 115−128.

(33) Zhang, C.; Pan, X.; Li, H.; Gardiner, A.; Sargent, I.; Hare, J.; Atkinson, P. M. A hybrid MLP-CNN classifier for very fine resolution remotely sensed image classification. *ISPRS J. Photogramm. Remote Sens.* **2018**, *140*, 133−144.

(34) Weng, Y.; Zhou, T.; Li, Y.; Qiu, X. Nas-unet: Neural architecture search for medical image segmentation. *IEEE Access* **2019**, *7*, 44247−44257.

(35) Koh, J. C.; Spangenberg, G.; Kant, S. Automated machine learning for high-throughput image-based plant phenotyping. *Remote Sens.* **2021**, *13*, 858.

(36) Elsken, T.; Metzen, J. H.; Hutter, F. Neural architecture search: A survey. *J. Mach. Learn. Res.* **2019**, *20*, 1997−2017.

(37) Jin, H.; Song, Q.; Hu, X.Auto-keras: Efficient neural architecture search with network morphism, 2018, *5*, *arXiv preprint arXiv:1806.10282.*

(38) Lu, J.; Behbood, V.; Hao, P.; Zuo, H.; Xue, S.; Zhang, G. Transfer learning using computational intelligence: A survey. *Knowl. Based Syst.* **2015**, *80*, 14−23.

(39) Barbin, D. F.; Maciel, L. F.; Bazoni, C. H. V.; Ribeiro, M. D. S.; Carvalho, R. D. S.; Bispo, E. D. S.; Miranda, M. D. P. S.; Hirooka, E. Y. Classification and compositional characterization of different varieties of cocoa beans by near infrared spectroscopy and multivariate statistical analyses. *J. Food Sci. Technol.* **2018**, *55*, 2457−2466.

(40) González-Martín, I.; Hernández-Hierro, J. M.; González-Cabrera, J. M. Use of NIRS technology with a remote reflectance fibre-optic probe for predicting mineral composition (Ca, K, P, Fe, Mn, Na, Zn), protein and moisture in alfalfa. *Anal. Bioanal. Chem.* **2007**, *387*, 2199−2205.

(41) Ferreira, R. D. A.; Teixeira, G.; Peternelli, L. A. Kennard-Stone method outperforms the Random Sampling in the selection of calibration samples in SNPs and NIR data. *Cienc. Rural* **2022**, *52*, 5.

(42) Long, M. S.; Cao, Y.; Cao, Z. J.; Wang, J. M.; Jordan, M. Transferable Representation Learning with Deep Adaptation Networks. *EEE Trans. Pattern Anal. Mach. Intell.* **2019**, *41*, 3071−3085.

(43) Fang, J.; Chen, Y.; Zhang, X.; Zhang, Q.; Huang, C.; Meng, G.; Liu, W.; Wang, X. EAT-NAS: Elastic architecture transfer for accelerating large-scale neural architecture search. *Sci. China Inf. Sci.* **2021**, *64*, 1−13.

(44) Takagi, S.; Yoshida, Y.; Okada, M. Impact of Layer Normalization on Single-Layer Perceptron - Statistical Mechanical Analysis. *J. Phys. Soc. Jpn.* **2019**, *88*, No. 074003.

(45) Yu, K.; Ranftl, R.; Salzmann, M.Landmark regularization: Ranking guided super-net training in neural architecture search. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2021; pp 13723−13732.

(46) Deng, L.; Li, G.; Han, S.; Shi, L.; Xie, Y. Model Compression and Hardware Acceleration for Neural Networks: A Comprehensive Survey. *Proc. IEEE* **2020**, *108*, 485−532.

(47) Han, S.; Mao, H.; Dally, W. J.Deep compression: Compressing deep neural networks with pruning, trained quantization and huffman coding, 2015, *arXiv preprint arXiv:1510.00149.*

(48) Cruz, I. A.; Chuenchart, W.; Long, F.; Surendra, K.; Andrade, L. R. S.; Bilal, M.; Liu, H.; Figueiredo, R. T.; Khanal, S. K.; Ferreira, L. F. R. Application of machine learning in anaerobic digestion: Perspectives and challenges. *Bioresour. Technol.* **2022**, *345*, No. 126433.