agronomy

Article

Origin Identification of *Saposhnikovia divaricata* by CNN Embedded with the Hierarchical Residual Connection Block

Dongming Li [1,2], Chenglin Yang [1], Rui Yao [1] and Li Ma [1,*]

[1] Institute of Information Technology, Jilin Agricultural University, Changchun 130118, China
[2] Thing Networking Engineering College, Wuxi University, Wuxi 214063, China
* Correspondence: mali@jlau.edu.cn

Abstract: This paper proposes a method for recognizing the origin of *Saposhnikovia divaricata* using the IResNet model to achieve computer vision-based classification. Firstly, we created a small sample dataset and applied data augmentation techniques to enhance its diversity. After that, we introduced the hierarchical residual connection block in the early stage of the original model to expand the perceptual field of the neural network and enhance the extraction of scale features. Meanwhile, a depth-separable convolution operation was adopted in the later stage of the model to replace the conventional convolution operation and further reduce the time cost of the model. The experimental results demonstrate that the improved network model achieved a 5.03% improvement in accuracy compared to the original model while also significantly reducing the number of parameters required for the model. In our experiments, we compared the accuracy of the proposed model with several classical convolutional neural network models, including ResNet50, Resnest50, Res2net50, RepVggNet_B0, and ConvNext_T. The results showed that our proposed model achieved an accuracy of 93.72%, which outperformed ResNet50 (86.68%), Resnest50 (89.38%), Res2net50 (91.83%), RepVggNet_B0 (88.68%), and ConvNext_T (92.18%). Furthermore, our proposed model achieved the highest accuracy among the compared models, with a transmission frame rate of 158.9 fps and an inference time of only 6.29 ms. The research methodology employed in this study has demonstrated the ability to reduce potential errors caused by manual observation, effectively improving the recognition ability of *Saposhnikovia divaricata* based on existing data. Furthermore, the findings of this study provide valuable reference and support for future efforts to develop lightweight models in this area.

Keywords: computer vision; artificial intelligence and machine learning; *Saposhnikovia divaricata*; convolutional neural networks; IResNet model; origin identification

Citation: Li, D.; Yang, C.; Yao, R.; Ma, L. Origin Identification of *Saposhnikovia divaricata* by CNN Embedded with the Hierarchical Residual Connection Block. *Agronomy* **2023**, *13*, 1199. https://doi.org/10.3390/agronomy13051199

Academic Editors: Baohua Zhang and Yongliang Qiao

Received: 18 February 2023
Revised: 22 April 2023
Accepted: 22 April 2023
Published: 24 April 2023

1. Introduction

The cultivation of Chinese herbal medicine is the cornerstone of the industrialization of traditional Chinese medicine [1], and the authenticity of herbs is a crucial indicator of the level of development of traditional medicine. "Authentic" herbs have been clinically used in Chinese medicine for an extended period and are cultivated in a specific region. These herbs are known for their superior quality and efficacy, and they exhibit consistent quality compared to those produced in other regions. Recently, the environmental impact and conservation and development of authentic herbs have emerged as global concerns [2]. The geographical variation of authentic herbs and the increasing deterioration of the environment have resulted in a significant decline in authentic herbs. As a result, there has been a sharp decrease in the production and quality of authentic herbs each year. These factors pose a significant obstacle to the sustainable development of authentic herbs [3].

Saposhnikovia divaricata is a local herb indigenous to Northeast China [4]. The dried root of the plant, which belongs to the family Umbelliferae, is used to treat various external symptoms, including wind rash and itching, rheumatic paralysis, and pain [5–7]. However, identifying traits in the complex variety of *Saposhnikovia divaricata* is difficult, resulting in many confused products and inaccurate cultivation practices [8–10]. This lack of precision in cultivation leads to the uneven quality of the herbs, confusing both production and medicinal use [11]. Currently, the selection of herbs is still made through manual observation. For minority herbs like *Saposhnikovia divaricata*, the identification process still heavily relies on manual observation compared to more commonly known herbs. This traditional method is time-consuming and subject to the subjective judgment of professionals, resulting in a relatively high margin of error.

Deep learning is a branch of artificial intelligence that utilizes the properties of the human brain to analyze and recognize objects through a multilayer neural network. This approach enables computers to learn and identify various features of objects, including their appearance, sound, and other characteristics. The application of deep learning in the classification and recognition of plants has already yielded many promising results. Krizhevsky et al. [12] developed the AlexNet network for fine-grained image recognition in a large visualization database, known as the ImageNet project, which has been widely used in research related to visual object recognition software Over the last decade, the field of deep learning has seen an influx of talented scholars who have proposed classical convolutional neural networks, including VGGNet [13], GoogLeNet [14], ResNet [15], and DenseNet [16]. These networks have been widely used in image classification tasks, including plant classification [17,18]. For instance, Liu et al. [19] utilized GoogLeNet to classify complex background images of 50 Chinese herbal medicines under natural conditions. This model has significant application value as it can be effectively used for fine-grained recognition of Chinese herbs, leading to improved accuracy and stability of recognition. Gui Yue [20] achieved an impressive accuracy of 95.62% by utilizing ResNet-50 in combination with an attention mechanism and cross-layer bilinear pooling algorithm to classify and recognize 12 plant images. Pei et al. [21] achieved a remarkable accuracy of 99.26% by utilizing ResNet18 as the base model, replacing the fully-connected layer with a convolutional layer, and incorporating a mixed-domain attention mechanism to classify through the Softmax layer. This model was utilized to classify 1360 images into 17 categories of flowers by building a multilayer convolutional neural network and using data enhancement and stochastic gradient descent techniques. Yanjiao Liu [22] achieved an impressive correct classification rate of 94.7% and 86.7% for ten succulent images and nine Saxifrage images, respectively. This was accomplished by leveraging the AlexNet deep model and migration learning technology, utilizing 15,336 images. The trained models could classify succulent and lithophytic images with an accuracy rate of 94.7% and 86.7%, respectively, with an average recognition time of approximately 6 s per image. Yi Zhao [23] and the team were able to classify and recognize 12 different weeds with an impressive accuracy rate of 95.47% by utilizing the VGG16 convolutional neural network model.

Ting Zhou [24] proposed an innovative ResNet-based ethnomedicinal plant image recognition method to classify eight common Tibetan medicines. The method involved an improvement strategy in network structure and loss function, such as replacing the first layer of a convolutional kernel of ResNet34 with a cascaded small convolutional kernel, adding a nonlinear activation layer, utilizing self-adaptive normalization method, Dropout, and optimizing Focal loss function instead of the cross-entropy loss function. Ultimately, this method achieved an accuracy rate of 91.18%.

This paper presents a new convolutional neural network model for the origin identification of authentic herbs by improving and analyzing the capabilities of the IResNet network proposed by C. Duta [25] and others. The new model incorporates the hierarchical residual connection block to enhance the feature extraction and network representation abilities of the IResNet. By leveraging a small number of herbal samples and embedding the hierarchical residual connection block, the model expands and strengthens the neural network's perception field to extract scale features. Additionally, the model incorporates deep separable convolution operations in the later stages of the network to reduce time costs and further enhance feature extraction capabilities. The research not only focuses on the classification of *Saposhnikovia divaricata* but also deepens existing studies' theoretical and practical foundations, providing a reference for future applications in mobile fields.

2. Materials and Methods

2.1. Image Dataset Acquisition and Pre-Processing of *Saposhnikovia divaricata*

According to the 2010 edition of "Pharmacopoeia of the People's Republic of China", we selected *Saposhnikovia divaricata* samples from Jilin, Hebei, Gansu, Northeastern Inner Mongolia, and Heilongjiang provinces. The samples used in this study were obtained from the Central Laboratory of Chinese Herbal Medicine Research Institute of Jilin Agricultural University. They were from the same batch of authentic *Saposhnikovia divaricata* [26–28]. To capture images of the samples based on their origins, we utilized small HD folding studio boxes for photography. The images of *Saposhnikovia divaricata* were captured from various angles and backgrounds with varying image details, with each image having a high resolution of 4000 × 3000 pixels. The dataset consisted of a total of 2001 images, comprising 408 images from Gansu origin, 427 from Hebei origin, 392 from Heilongjiang origin, 406 from Jilin origin, and 368 images from Northeastern Inner Mongolia. Considering the differences in quality, lighting, and categories of the samples, we considered all these factors during data processing. To improve the model's generalization ability and simulate a real-world environment, we utilized the Imgaug Library in Python for data pre-processing. Techniques such as affine transformation, horizontal flip, random crop, and color change were applied to the dataset images to enhance the data samples and create a more diverse dataset (Figure 1) [29–31].

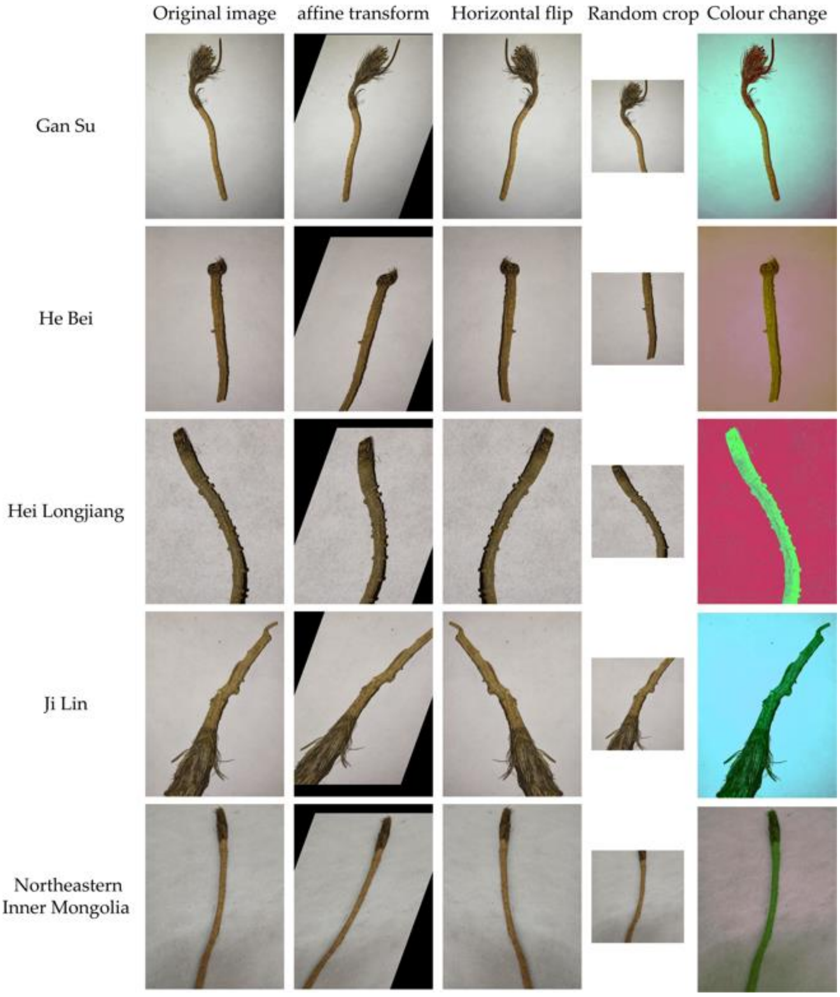


Figure 1. Raw data images of different origins and data-enhanced images.

2.2. Data Set Division of *Saposhnikovia divaricata*

To address the many uncertainties in the experiments, each model was trained using a 5-fold cross-validation method. This involved dividing the entire *Saposhnikovia divaricata* dataset into training and validation sets at an 8:2 ratio using a Python dataset partitioning script, as presented in Table 1. The script divides the dataset into five parts based on the geographic origin of genuine *Saposhnikovia divaricata* and performs random homogenization for each part. This ensures that the ratio of the training set to the validation set is always 8:2 during the network model training. Due to the small number of samples, the validation set was also used as a test set to test the results. Therefore, the final experimental results in this paper were based on the average of the results of five experiments.

Table 1. Data set division of *Saposhnikovia divaricata*.

Place of Origin	Training Set	Validation Set	Totals
Gansu	327	81	408
Hebei	342	85	427
Heilongjiang	314	78	392
Jilin	325	81	406
Northeast Inner Mongolia	295	73	368

3. Building the Model

3.1. IResNet Model

IResNet (Improved Residual Networks) is a backbone model proposed by C. Duta et al. in 2020, improving the traditional ResNet-v2 architecture. Compared to traditional residual networks, the IResNet model has a more rational structure, enabling better information extraction and stronger learning ability. Since the differences in the appearance of *Saposhnikovia divaricata* from different origins are relatively small, we need to extract phenotypic features of *Saposhnikovia divaricata* from multiple angles. The IResNet model can effectively reduce information loss and extract more comprehensive feature information when extracting phenotypic features of *Saposhnikovia divaricata*. Additionally, it can prevent the model from generating gradient dispersion and reduce the need for hyperparameters. For these reasons, we used the IResNet model as the backbone for the origin identification of *Saposhnikovia divaricata*.

3.2. Improved IResNet Model

3.2.1. Hierarchical Residual Connection Block (HRC Block)

The extraction of phenotypic multiscale features of *Saposhnikovia divaricata* is crucial for the task of origin identification presented in this paper. To achieve this, we need to extract phenotypic features at a finer level of granularity using macro receptive fields. S.-H. Gao [32] proposed a simple and effective multi-scale feature processing CNN module called the hierarchical residual connection block (HRC block), as shown in Figure 2.

The HRC block structure replaces a generic single 3 × 3 convolution kernel by constructing similar residual connections with a hierarchy within a single residual block. With each 3 × 3 convolution operation, all feature information to its left can be potentially accepted, and each output can increase the receptive field. As a result, each HRC block can acquire a different number and combination of features with different receptive field sizes, thereby increasing the scale at which the output features can represent information.

$$\mathrm{Y_i} = \begin{cases} \mathrm{x_{\ldots}} & \mathrm{i} = 1; \\ \ldots & \mathrm{i} = 2; \\ \mathrm{K_i}\left(\mathrm{x_i} + \mathrm{y_{i-1}}\right) & 2 < \mathrm{i} \leq \mathrm{s}. \end{cases} \quad (1)$$

At the first 1 × 1 After convolution, the input is divided into s subsets, defined as $\mathrm{x_i}$, i ∈ {1,2,...,s}. Each feature has the same scale size, but the channel is 1/s of the input feature. Except for $\mathrm{x_1}$, other sub-features have corresponding 3 × 3 Convolution kernel, defined as

K_i, whose output is y_i. The sub-feature x_i is added to K_{i-1} and then entered into K_i. The 3×3 convolution of x_1 is omitted to increase s while decreasing the parameters.

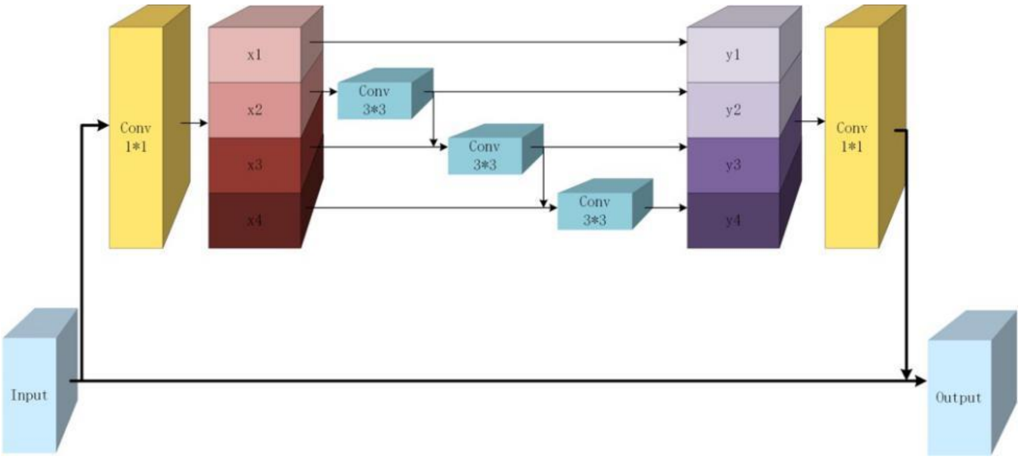


Figure 2. Hierarchical residual connection block.

Unlike most existing approaches to multiscale feature learning through hierarchical representation in neural networks, the HRC block structure improves the multiscale representation of neural networks at the image classification level. It exploits the multiscale potential at a finer granularity level by introducing a new dimension, scale, depth, width, and cardinality dimensions in existing network factors. In the HRC block structure, s is used as a control parameter for the scale dimension. A larger s allows for more features and sensory fields to be learned, which is orthogonal to existing approaches that utilize hierarchical operations.

In summary, we incorporate the HRC block into the *Saposhnikovia divaricata* identification model to increase the multiscale feature extraction capability.

3.2.2. Improved IResStage Structure (IIR Stage)

Although IResNet effectively reduces the problem of gradient dispersion to a certain extent, its network depth is the same as conventional residual networks. Moreover, the conventional convolution operation applied to this network consumes more time and computational resources due to its high computational effort. To reduce the computation and time expenditure of the network model, we replace the conventional convolution with the Depthwise Separable Convolution (DS conv) operation proposed by Chollet [33]. This operation helps to reduce the overall consumption of the network model while ensuring the feature extraction capability.

When the size of the input feature map is $H_f \times W_f \times M$ and the size of the convolution kernel is $K_h \times K_w \times M$ and the number is L, then the total computation of L regular convolutions is:

$$H_f \times W_f \times K_h \times K_w \times M \times L \tag{2}$$

The total calculation for DS conv is:

$$H_f \times W_f \times K_h \times K_w \times M + M \times L \times 1 \times 1 \tag{3}$$

The ratio of the computational effort of DS conv to the normal convolution is:

$$\frac{1}{L} + \frac{1}{K_h \times K_w} \tag{4}$$

The above study shows that using DS conv allows for deeper neural network layers while reducing the time cost of the network and ensuring the same input. Therefore, we replaced the 3 × 3 conventional convolution operation inside the original IResStage structure with a 3 × 3 DS conv to construct a new IIR stage (Figure 3).

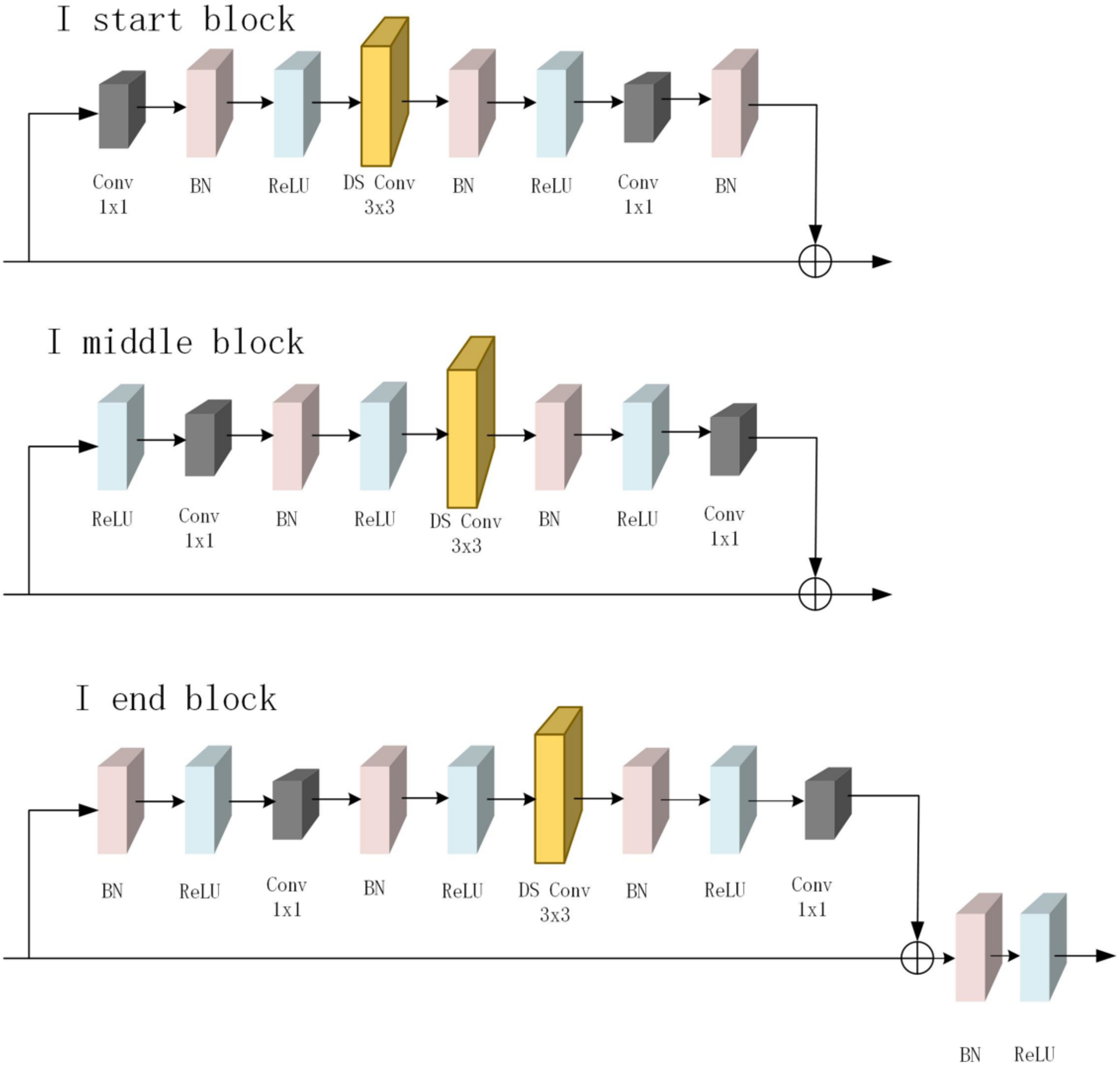


Figure 3. IIRStage structure.

3.3. *Saposhnikovia divaricata* Based Origin Classification Model

To improve the accuracy of the origin recognition model of *Saposhnikovia divaricata*, we introduced the HRC block module in the early network stage. This enhances the multi-scale feature extraction capability of the model, strengthens the correlation between different regions and features within the same residual block, and avoids errors in recognition of different origins of *Saposhnikovia divaricata*. The HRC stage comprises HRC blocks, which differ from the traditional method of extracting scale features by hierarchical scale feature extraction modules. Instead, hierarchical building blocks are used to enhance the scale feature extraction ability and expression of the whole model, increasing the perceptual field of each network layer. This enables the extraction of phenotypic features of anti-fungal at a finer granularity. Furthermore, as the channels are split, the number of parameters does not increase much, which avoids negative problems in model training due to oversized network parameters.

The IIR stage significantly enhances the ability of feature extraction and information propagation. Adjusting the position of the batch normalization layer (BN) and the non-linear activation function (ReLU) effectively avoids the negative impact of information propagation caused by the non-linear activation function. Additionally, the conventional convolution operation is replaced by the DS conv operation, which significantly reduces the amount of computation and the number of parameters that can cause the network model to become too large, leading to high computational costs. The improved network model is shown in Figure 4.

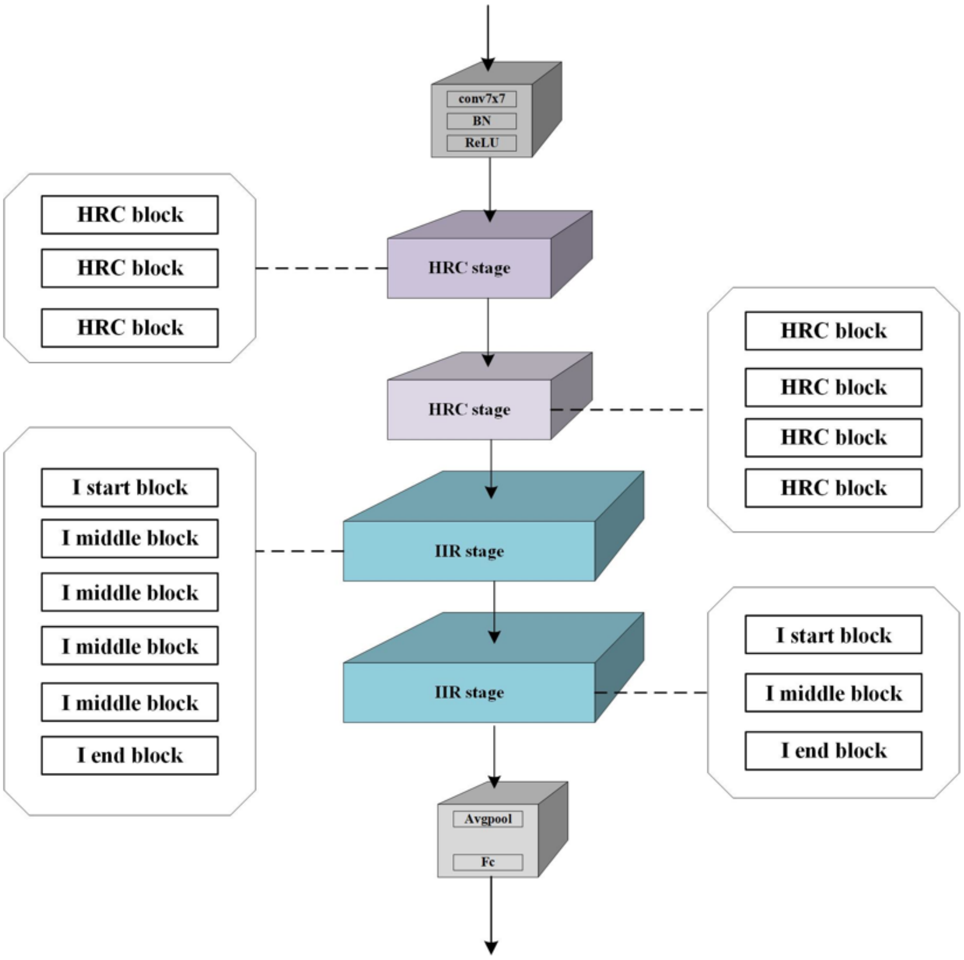


Figure 4. The improved IresNet model.

4. Experimental Environment Configuration and Result Analysis

4.1. Experimental Environment Configuration

The experimental configuration environment was a laptop GPU with a processor of AMD Ryzen9 5900HX with RadeonGraphics 3.30 GHz and a graphics card of Nvidia GeForce RTX 3070 LaptopGPU, with a windows11 operating system, software configuration installed as Anaconda3-2021.11-windows version, and given a Pytorch1. 1 built-in Python3.10.1 programming language. 10 deep learning framework, all comparison algorithms were run in the same environment.

4.2. Experimental Hyperparameter Setting

Table 2 provides information on the specific experimental parameters used in training the new network model proposed in this paper. We set the input size of the 2001 *Saposhnikovia divaricata* dataset to 224 × 224, the number of training rounds to 120, and the initial learning rate to 0.01. The original MultiStep Learning Rate Decay (MultiStepLR) strategy was replaced with an adaptive learning rate decay strategy (Reduce LR On Plateau). The training set loss is used as the adaptive criterion. When the training set loss stops decreasing for six consecutive rounds (i.e., the number of tolerated rounds is six), the learning rate decay is automatically adjusted. This enables the determination of the optimal value of the learning rate at the current stage, according to the network training, thus improving the accuracy and precision of the network model training.

Table 2. Training hyperparameter information.

Set of Parameter	Value or Name
Input Size	224 × 224
Number of samples	2001
Number of epochs	120
Optimizer	SGD
Learning rate	0.01
Learning rate scheduler	Reduce LR On Plateau
Number of patience rounds	6

4.3. Experimental Results and Analysis

4.3.1. Comparison Experiments of Different Models

To verify the effectiveness and advancement of the new network model, we used model accuracy, model loss, model inference time, frames per second transmission, number of model parameters and corresponding arithmetic power as evaluation indexes for model performance. We compared the new network model with four classical convolutional neural networks, namely ResNet50, Resnest50, Res2net50, RepVggNet_B0 and ConvNext_T. The results are presented in Table 3, Figures 5 and 6.

Table 3. Comparison experiments of different models.

Model	Accuracy	Loss	Params	FLOPs	Inference Time	FPS
ResNet50	86.68%	0.3738	23.52 M	4.12 GMac	8.41 ms	100.93 fps
ResNest50	89.38%	0.6024	25.44 M	5.41 GMac	17.6 ms	57.01 fps
Res2Net50	88.68%	0.4019	23.66 M	4.29 GMac	9.14 ms	109.41 fps
RepVggNet_B0	92.18%	0.3212	17.06 M	3.76 Gmac	11.16 ms	89.61 fps
ConvNext_T	91.83%	0.6002	27.82 M	4.47 Gmac	7.12 ms	140.17 fps
Our model	93.72%	0.2192	16.16 M	3.32 Gmac	6.29 ms	158.9 fps

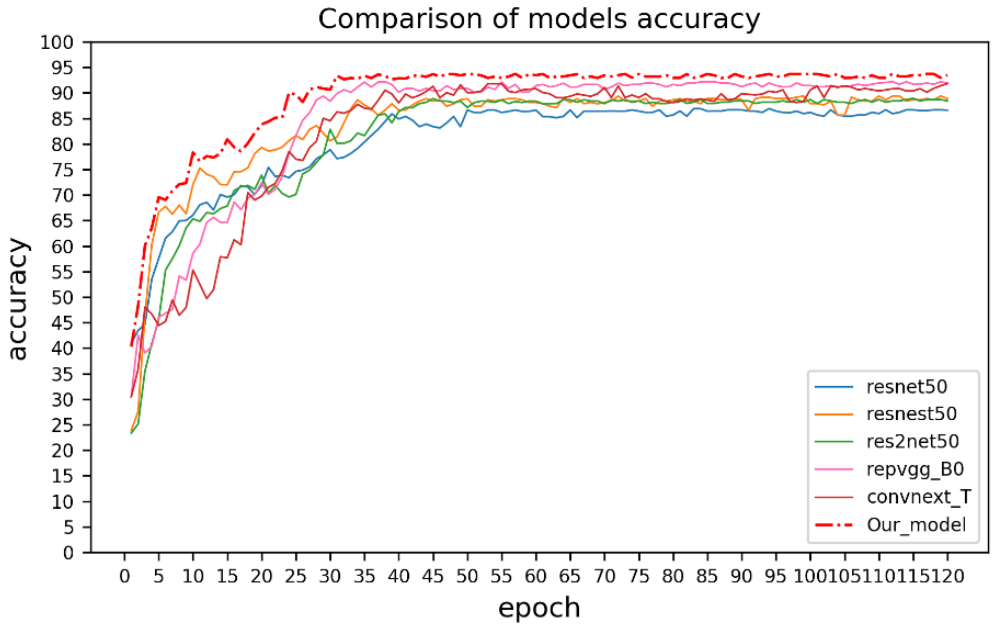


Figure 5. Visualization results of the accuracy of different model comparison experiments.

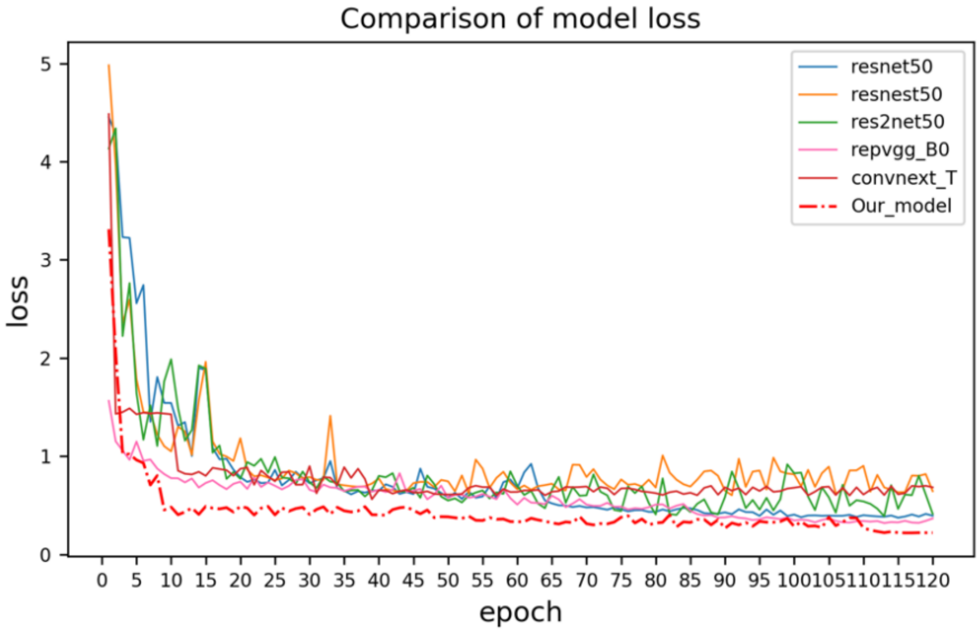


Figure 6. Visualization results of the loss of different model comparison experiments.

The comparison experiments show that the network model proposed in this paper achieves a recognition accuracy of 93.72% on this dataset, with an inference time of only 6.29 milliseconds (ms), a transmission frame rate of 158.9 fps and the lowest loss value compared to other models at 0.2192. The model also has the shortest training time per round, allowing fast convergence to find the optimal value. This proves the superior performance of the model, which converges quickly to find the best value.

In summary, the accuracy of the models proposed in this study is higher than that of other convolutional neural network models. In addition, the models are relatively smaller, reducing computational costs while ensuring improved feature extraction. This also reduces network consumption and the time cost. However, in the accuracy visualization images of the different model comparison experiments, the accuracy curves of the network models proposed in this paper fluctuate, possibly due to the uncertainty of the decay of the adaptive learning rate as the model accuracy increases.

4.3.2. Comparison Experiments for Image Enhancement

To ensure that the augmented *Saposhnikovia divaricata* samples had a practical effect on the model, the original IResNet model was used as the basis for this study and the experimental parameters were kept in the same settings. After using the sample data augmentation strategy, the recognition accuracy of the model was improved by 2.85%. This demonstrates that using the data enhancement strategy can effectively expand the diversity of *Saposhnikovia divaricata* samples and enhance the model's generalization ability. Therefore, the next ablation experiments were conducted using the data-enhanced herbal samples. The recognition accuracy of the model under the data enhancement strategy is shown in Table 4.

Table 4. Model recognition accuracy under data enhancement strategies.

Data Enhancement Strategy	Accuracy of Models
Yes	88.69%
No	85.84%

4.3.3. Ablation Experiments

The Effect of Introducing DS Conv on the Origin Identification Model of the *Saposhnikovia divaricata*

This study replaces the conventional convolution operation by introducing the DS conv operation with a lightweight idea. Compared with the original model, the accuracy improvement is only 0.72%. However, by reconstructing the residual blocks, the number of parameters and the overall arithmetic power of the model is reduced while ensuring a small increase in accuracy, which effectively improves the model's inference time and the number of frames per second transmitted. The overall results of the model after the introduction of DS conv are shown in Table 5.

Table 5. Effect of introducing DS conv into the model.

Model	Accuracy	Inference Time	FPS	Params	FLOPs	F1score
Iresnet	88.69%	8.32 ms	120.1 fps	23.52 M	4.16 GMac	88.3%
Iresnet + IIR stage	89.41%	6.17 ms	162.7 fps	15.51 M	3.04 GMac	89.1%

The Effect of Embedding the Hierarchical Residual Connection Block on the Origin Identification Model of *Saposhnikovia divaricate*

To evaluate the effect of the hierarchical residual connection block on the recognition accuracy of this model, we embedded the HRC stage in the pre-network model. The model effects of embedding the HRC block are shown in Table 6.

Table 6. Effect of the model with embedded HRC block.

Model	Accuracy	Inference Time	FPS	Params	FLOPs	F1score
Iresnet	88.69%	8.32 ms	120.1 fps	23.52 M	4.16 GMac	88.3%
Iresnet + HRC stage	90.25%	6.34 ms	149.4 fps	24.11 M	4.95 GMac	90.5%

Compared with the original model, the model with the HRC stage embedded increased in accuracy by 1.56% and inference time by 1.98 milliseconds (ms), while the transmission frame rate decreased to 149.4 fps. As recognition accuracy rose, the number of parameters and arithmetic power increased accordingly. This is due to the embedded hierarchical residual connection block, which increases the network complexity as the model feature extraction capability increases.

Final Model for Origin Recognition of *Saposhnikovia divaricate*

The accuracy of the improved model for identifying the origin of *Saposhnikovia divaricata* improved by 5.03%, as shown in Table 7. The inference time was reduced by 2.03 milliseconds (ms), and the frames per second (fps) were increased by 38.8 fps. The model with the hierarchical residual connection block achieved a further significant improvement in recognition accuracy at the cost of a small increase in inference time and a number of parameters. A visual comparison of the accuracy before and after model improvement is shown in Figure 7.

Table 7. Comparison results before and after model improvement.

Model	Accuracy	Inference Time	FPS	Params	FLOPs	F1score
Iresnet	88.69%	8.32 ms	120.1 fps	23.52 M	4.16 GMac	88.3%
Iresnet + IIR stage + HRC stage	93.72%	6.29 ms	158.9 fps	16.16 M	3.32 GMac	93.4%

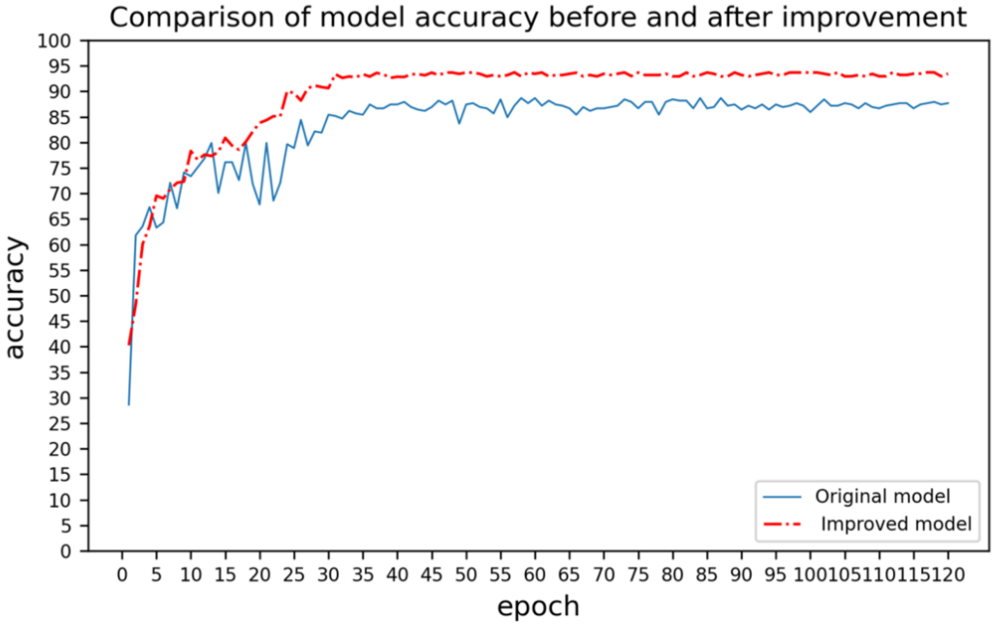


Figure 7. Visual comparison of accuracy before and after model improvement.

The visualization images obtained by the Grad-CAM [34] tool provide a more intuitive view of the regions of interest in the feature maps of each model layer before and after the improvements. Figure 8 shows that the improved model has a significantly enhanced feature extraction capability for *Saposhnikovia divaricata*, effectively analyzing more valuable information. The improved feature extraction capability of the model is particularly important due to the little change in the appearance of *Saposhnikovia divaricata*.

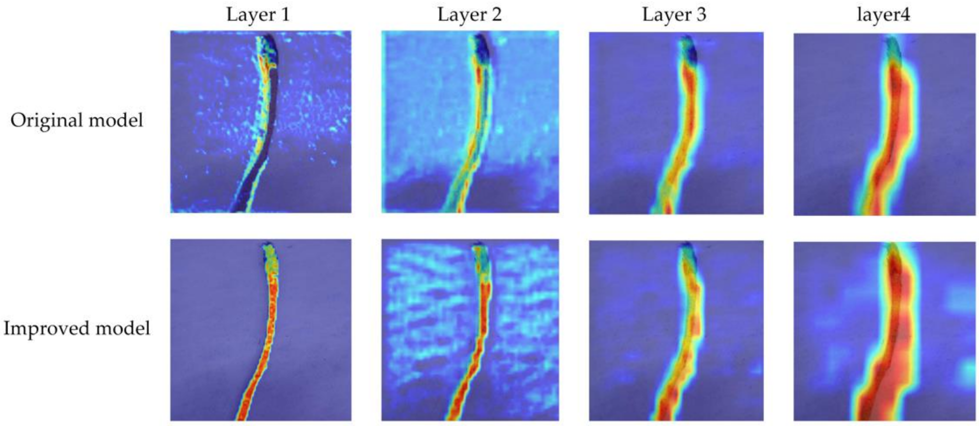


Figure 8. Visualization comparison of the thermal characteristic maps of different layers of the network model before and after improvement.

From the above experimental results, we found that embedding the hierarchical residual connection block enhanced the feature extraction and refinement capabilities of the model for the *Saposhnikovia divaricata* data set. Additionally, the DS conv operation, which introduces the idea of lightness, not only ensures the feature extraction capability of the network model for deeper layers but also effectively reduces the number of network model parameters and inference time.

4.3.4. Model Evaluation

- Analysis of model evaluation indicators

To verify the effectiveness of the enhanced network model in identifying the origin of *Saposhnikovia divaricata* from the five production areas, we utilized Precision, Recall, and F1 score to analyze the model before and after the enhancement.

Precision refers to the proportion of samples predicted to be positive that belong to the positive category.

$$\text{Precision} = \frac{\text{TP}}{\text{TF} + \text{FP}} \tag{5}$$

The recall is the proportion of all positive samples correctly identified as positive categories. That is, it is the proportion of identified predicted positive categories to actual positive categories.

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{6}$$

The F1 score is the average of the sum of precision and recall. It measures the balance between the two and reflects the stability of the model performance, as well as the generalization ability of the prediction model.

$$\text{F1} = 2 \times (\text{Precision} \times \text{Recall}) \div (\text{Precision} + \text{Recall}) \tag{7}$$

The F1 score for the Jilin region increased by 5.5%, while the F1 score for the Northeastern Inner Mongolia region reached a perfect score of 100% with an increase of 0.9%. The improved model for identifying *Saposhnikovia divaricata* production areas exhibited increased accuracy and recall rates. This confirmed that the proportion of correctly classified positive samples was effectively increased, and the improved model demonstrated a more significant improvement in identification accuracy than the previous model. Table 8 shows the evaluation indexes for the model's recognition performance before and after the improvement.

Table 8. Comparison of model recognition performance evaluation metrics.

Place of Origin	**Precision**		**Recall**		**F1 Score**	
	Before	**After**	**Before**	**After**	**Before**	**After**
Gansu	0.8	0.882	0.381	0.714	51.6%	78.9%
Hebei	0.727	0.818	0.711	0.8	71.9%	80.9%
Heilongjiang	0.894	0.962	0.974	0.987	93.2%	97.4%
Jilin	0.867	0.927	0.894	0.944	88%	93.5
Northeast Inner Mongolia	0.991	1.0	0.992	1.0	99.1%	100%

- Visual comparative analysis of model confusion matrices

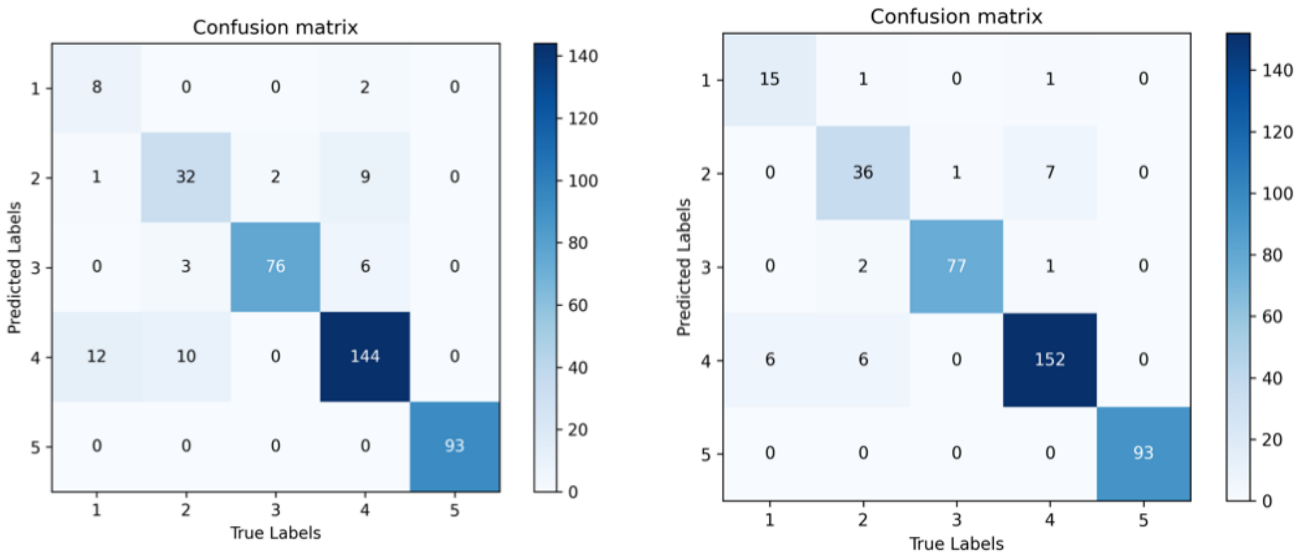
Table 9 displays the corresponding labels of the confusion matrix, while Figure 9 illustrates the visual images of the confusion matrix for the improved network model before and after the enhancement.

To verify the recognition ability of the original identification model proposed in this paper, the confusion matrix visualization results analysis shows that the improved network model incorrectly identifies one and twelve *Saposhnikovia divaricata* from the Gansu production area as the Hebei production area and the Jilin production area. In contrast, the improved network model only incorrectly identifies six *Saposhnikovia divaricata* from the Gansu production area as the Jilin production area and does not incorrectly identify them as the Hebei production area.

In conclusion, the enhanced network model significantly reduced the original error rate, improved the model's accuracy, and proved more effective in identifying the origin of *Saposhnikovia divaricata*. Furthermore, further in-depth analysis using the confusion matrix of the improved network model revealed that the model did not misidentify samples from the northeastern Inner Mongolia production area. However, there were still relatively large errors in identifying samples from the Jilin and Hebei production areas.

Table 9. Corresponding labels and origin.

Label	Place of Origin
0	Gansu
1	Hebei
2	Heilongjiang
3	Jilin
4	Northeast Inner Mongolia



(**a**) (**b**)

Figure 9. Model confusion matrix visualization. (**a**) Original model confusion matrix visualization. (**b**) Our model confusion matrix visualization.

5. Conclusions

This study aims to enhance the original model of the IResNet network by incorporating the concept of lightweight through the use of deep separable convolution operation (DS conv) instead of the traditional convolution operation. This approach reduces the number of parameters and inference time of the model while enhancing the network's refinement extraction ability and overall recognition capacity. Additionally, the study embeds the layered residual connection module (HRC block) to enhance the multi-scale feature extraction ability of the network. This integration further expands the scale feature extraction of the network model and strengthens the network's expression capacity. Combining the features of these two distinct modules and conducting experiments on a dataset from five production areas demonstrate that the proposed method's accuracy is as high as 93.72%. Compared with the original network model, the proposed model in this paper's accuracy rate is improved. In contrast, the arithmetic power, and the number of parameters of the model are significantly reduced, the number of transmitted frames is increased, and the model's performance is greatly improved. By comparing five network models often used for image classification, ResNet50 Resnest50, Res2net50, RepVggNet_B0 and ConvNext_T, the network model proposed in this paper is significantly better than the other network models in terms of performance, and the model complexity is lower, and the requirement for equipment is lower.

To revolutionize the conventional manual collection method for identifying *Saposhnikovia divaricata*, future research aims to introduce a lightweight herb identification model that can be integrated with mobile devices, enabling real-time classification of Chinese herbs using intelligent devices. This will enhance the accuracy and portability of

Chinese medicine collection and identification and contribute to geographical analysis in Chinese medicine.

Author Contributions: Conceptualization, D.L.; data curation, C.Y.; formal analysis, R.Y.; funding acquisition, D.L.; project administration, D.L.; supervision, L.M.; validation, C.Y.; visualization, R.Y.; writing review and editing, C.Y. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Jilin province science and technology development plan project (No. 20210204050YY). Jilin Province Ecological Environment Department Project (2021-07), Jilin Provincial Department of Education Project (No. JJKH20210336KJ).

Data Availability Statement: All relevant data are included in the manuscript. Raw images are available on request from the corresponding author.

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

References

1. Cai, Z.; Wang, G.; Ma, Z.; Zhang, K. Current Situation and development Strategy of Chinese medicinal materials seed industry in Gansu Province. *Gansu Agric. Sci. Technol.* **2022**, *53*, 19–24.
2. Yang, H.; Chen, S. Environmental impact, protection and development of authentic medicinal materials. *Shizhen Tradit. Chin. Med.* **2015**, *26*, 1971–1972.
3. Zhao, L.; Shi, M.; Zhang, Q.; Qin, L. Research progress on quality characteristics and formation mechanism of Sun Yiqi Authentic medicinal materials. *Chin. Herb. Med.* **2022**, *53*, 6931–6947.
4. Wang, Y.; Feng, H.; Qin, R.; Ma, X. Exploration on the cultivation methods of Parsnips in the Greater Khingan Mountains. *Mod. Agric. Res.* **2021**, *27*, 108–109. [CrossRef]
5. Chen, X.; Zeng, X.; Zhou, D. Research status and development direction of Chinese Daodi medicinal materials. *Mod. Distance Educ. Chin. Tradit. Med.* **2021**, *19*, 194–198.
6. Lv, R.; Zhu, T.; Ma, X.; Wang, L.; Jin, L.; Zhang, D. Study on Ecological suitability Zoning of windbreak. *Chin. J. Inf. Tradit. Chin. Med.* **2022**, *29*, 1–4. [CrossRef]
7. Chen, J.; Peng, L.-G. Comparative analysis on the content of effective components of Radix parsnips from different producing areas. *Chin. Natl. Folk. Med.* **2018**, *27*, 34–36.
8. Song, M.; Lv, L.; Li, J.; Fan, C.; Xu, M. Investigation on microscopic characters of Parsnips and related species. *J. Hebei North Univ. (Nat. Sci. Ed.)* **2015**, *31*, 96–99.
9. Su, S. Research on the Recognition Method of Small Sample Grape Leaf Disease Based on Deep Learning. Master's Thesis, Anhui Agri cultural University, Hefei, China, 2022.
10. Xue, Y. Study on Target Detection and Weed Classification of Wheat Ears in Wheat. Master's Thesis, Lanzhou University, Lanzhou, China, 2022.
11. Chen, H.; Li, X.; Liu, X.; Zhang, X.; Yao, Z.; Wang, Q.; Xiong, F. Study on multi-index comprehensive quality evaluation and fingerprint of Parsnips. *World Tradit. Chin. Med.* **2022**, *17*, 1827–1832.
12. Krizhevsky, A.; Sutskever, I.; Hinton, G. ImageNet Classification with Deep Convolutional Neural Networks. *Adv. Neural Inf. Process. Syst.* **2012**, *25*, 84–90. [CrossRef]
13. Simonyan, K.; Zisserman, A. Very Deep Convolutional Networks for Large-Scale Image Recognition. *arXiv* **2015**, arXiv:1409.1556.
14. Szegedy, C.; Liu, W.; Jia, Y.; Sermanet, P.; Reed, S.; Anguelov, D.; Erhan, D.; Vanhoucke, V.; Rabinovich, A. Going Deeper with Convolutions. In Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, MA, USA, 7–12 June 2015; pp. 1–9.
15. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep Residual Learning for Image Recognition. In Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 27–30 June 2016; pp. 770–778.
16. Huang, G.; Liu, Z.; van der Maaten, L.; Weinberger, K.Q. Densely Connected Convolutional Networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 4700–4708.
17. Zhang, R.; Li, Z.; Hao, J.; Sun, L.; Li, H.; Han, P. Convolutional neural network based on transfer learning for peanut pod rank image recognition. *Trans. Chin. Soc. Agric. Eng.* **2020**, *36*, 171–180.
18. Qi, J. Research on Apple target recognition method based on YOLOv5. *Comput. Program. Ski. Maint.* **2022**, *8*, 137–139. [CrossRef]
19. Liu, S. Research on Image Classification of Chinese Herbal Medicine Based on Deep Learning. Master's Thesis, Qilu University of Technology, Jinan, China, 2019.
20. Gui, Y. Classification and Recognition of Crop Seedlings and Weeds Combined with Attention Mechanisms. Master's Thesis, Anhui Agri cultural University, Hefei, China, 2019.
21. Pei, X.; Zhang, Y. Flower image classification algorithm based on improved residual network. *Electron. Device* **2020**, *43*, 698–704.
22. Liu, Y. Research on Image Classification of Succulents Based on Deep Convolutional Network. Master's Thesis, Dalian Jiaotong University, Dalian, China, 2018.

23. Zhao, Y.; Liu, T. Research on field weed recognition method based on Convolutional neural network. *Comput. Simul.* **2019**, *36*, 440–444.
24. Zhou, T. Research on Image Recognition of Ethnomedicine Plants Based on Deep Learning. Master's Thesis, Jiangxi University of Traditional Chinese Medicine, Nanchang, China, 2022.
25. Duta, I.C.; Liu, L.; Zhu, F.; Shao, L. Improved Residual Networks for Image and Video Recognition. In Proceedings of the 2020 25th International Conference on Pattern Recognition (ICPR), Milan, Italy, 10–15 January 2021; pp. 9415–9422.
26. Pan, X.; Yong, L.; Yue, Z.; Xu, X.; Cao, Y. Research on the Algorithm of Small Object Detection of Image Based on Genetic Algorithm. In Proceedings of the International Symposium on Intelligent Information Technology Application Workshops, Shanghai, China, 21–22 December 2008.
27. Li, J.; Li, Y.; Fan, C.; Wang, L. Investigation on characters of Medicinal Parsnips and its related species. *J. Hebei Tradit. Chin. Med.* **2014**, *29*, 38–41. [CrossRef]
28. Zhang, Y.; Feng, F.; Wang, J.; Pan, F.; Tan, M.; Zhang, Z.; Wang, L. Fast classification method of small sample images based on tensor features. *J. Nanjing Univ. (Nat. Sci.)* **2022**, *58*, 1059–1069. [CrossRef]
29. Xia, Q.; He, P.; Luo, J.; Liu, J. Efficient carrier image enhancement algorithm based on universal noise suppression. *Inf. Netw. Secur.* **2022**, *22*, 64–75.
30. Yang, S.; Wang, Y.; Li, Y.; Wang, G. Image Enhancement and Translation for RGB-D Indoor Scene Recognition. In Proceedings of the 2020 4th International Conference on Digital Signal Processing, Association for Computing Machinery, New York, NY, USA, 10 September 2020; pp. 7–11.
31. Hu, X.; Peng, T. Image classification based on Data-enhanced Vision Transformer. *J. Xihua Univ. (Nat. Sci. Ed.)* **2022**, *41*, 9–16.
32. Gao, S.-H.; Cheng, M.-M.; Zhao, K.; Zhang, X.-Y.; Yang, M.-H.; Torr, P. Res2Net: A New Multi-Scale Backbone Architecture. *IEEE Trans. Pattern Anal. Mach. Intell.* **2021**, *43*, 652–662. [CrossRef] [PubMed]
33. Chollet, F. Xception: Deep Learning With Depthwise Separable Convolutions. In Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 21–26 July 2017; pp. 1251–1258.
34. Selvaraju, R.R.; Cogswell, M.; Das, A.; Vedantam, R.; Parikh, D.; Batra, D. Grad-Cam: Visual Explanations from Deep Networks via Gradient-Based Localization. In Proceedings of the IEEE International Conference on Computer Vision, Venice, Italy, 22–29 October 2017; pp. 618–626.