Postprint: Desert Plant Recognition Method in Natural Backgrounds Integrating Transfer Learning and Ensemble Learning

Authors: Wang Yapeng, Cao Shanshan, Li Quansheng, Sun Wei

Date: 2023-08-14T00:00:00+00:00

Abstract

[Purpose/Significance] Accurate identification of desert plants constitutes an indispensable task in their understanding and conservation, forming the foundation of desert ecological research and protection. Automatic machine vision classification and recognition of wild desert plant images under natural conditions can effectively enhance the efficiency of plant resource surveys, mitigate the influence of human subjective factors, and holds significant importance for the precise classification, diversity conservation, and resource utilization of desert plants. [Method] Taking whole-plant desert plant images in natural environments as the research object, we constructed a desert plant image dataset for the arid region of Xinjiang. Based on EfficientNet B0-B4 networks as the backbone, we proposed a desert plant image recognition algorithm that integrates transfer learning and ensemble learning, and conducted comparative validation on the public dataset Oxford Flowers102. [Results and Discussion] For single sub-models based on the EfficientNet B0 network, the Top-1 accuracy reached a maximum of 93.35% and a minimum of 92.26%. The accuracies of the soft-voting Ensemble-Soft model, hard-voting Ensemble-Hard model, and Ensemble-Weight model integrated via weighted voting were 93.63%, 93.55%, and 93.67%, respectively, with F1 Scores comparable to the accuracies. For single sub-models based on EfficientNet B0-B4 networks, the Top-1 accuracy reached a maximum of 96.65% with an F1 Score of 96.71%, while the accuracies of the Ensemble-Soft, Ensemble-Hard, and Ensemble-Weight models were 99.07%, 98.91%, and 99.23%, respectively. Compared with single sub-models, the accuracy was further improved, with F1 Scores essentially matching the accuracies, demonstrating significant model performance enhancement. Comparative experiments on the public dataset Oxford Flowers102 showed that compared with the five sub-models, the three ensemble models achieved maximum improvements in accuracy and F1 Score of 4.56% and 5.05%, respectively, and minimum improvements of 1.94% and 2.29%, respectively, proving that the transfer and ensemble learning strategies proposed in this study can effectively enhance model performance. [Conclusion] This method can improve the recognition accuracy of desert plants, enabling accurate identification after cloud transmission to the server, and provides solutions for problems such as low accuracy of plant image recognition, weak model robustness, and poor generalization in real wild environments. It serves scenarios including field surveys, educational outreach, and scientific experiments.

Full Text

Desert Plant Recognition Method Under Natural Background Incorporating Transfer Learning and Ensemble Learning

WANG Yapeng[1,2], **CAO Shanshan**[2,3], **LI Quansheng**[1], **SUN Wei**[2,3*]

[1] Computer and Information Engineering College, Xinjiang Agricultural University, Urumqi 830052, China
[2] Agricultural Information Institute, Chinese Academy of Agricultural Sciences, Beijing 100081, China
[3] National Agriculture Science Data Center, Beijing 100081, China

Abstract

[**Objective/Significance**] Accurate identification of desert plants is an indispensable task in their understanding and protection, forming the foundation of desert ecological research and conservation. Machine vision-based automatic classification and recognition of field desert plant images under natural conditions can effectively enhance the efficiency of plant resource surveys while reducing subjective human factors. This holds significant importance for the precise classification, diversity conservation, and resource utilization of desert plants. [**Methods**] This study focused on whole-plant desert plant images in natural environments, constructing a desert plant image dataset for the arid regions of Xinjiang. Using the EfficientNet B0 network as the backbone, we proposed a desert plant image recognition algorithm that integrates transfer learning and ensemble learning, with comparative validation conducted on the public Oxford Flowers102 dataset. [**Results and Discussion**] The highest accuracy achieved by single sub-models based on the EfficientNet B0 network reached 93.35%, with the weighted voting ensemble achieving comparable performance. After integrating models through ensemble learning and voting strategies, the accuracy of the Ensemble-Soft, Ensemble-Hard, and Ensemble-Weight models reached 93.63%, 93.55%, and 93.67% respectively, with F1 Scores essentially matching accuracy rates. Compared to single sub-models, the ensemble models showed further improved precision, with significant model performance. When validated on the public Oxford Flowers102 dataset, the three ensemble models achieved maximum improvements of 4.56% and 5.05% in accuracy and F1 Score respectively over the five sub-models, while minimum improvements were 1.94%

and 2.29%. These results demonstrate that the proposed transfer and ensemble learning strategies effectively enhance model performance. [**Conclusion**] This method improves desert plant recognition accuracy, enabling accurate identification after cloud transmission to servers. It provides solutions to challenges such as low recognition accuracy, weak model robustness, and poor generalization for plant images in real field environments, serving applications in field surveys, educational outreach, and scientific experiments.

Keywords: desert plant recognition; natural background; ensemble learning; transfer learning; voting method; dataset
CLC number: TP183; J522.3
Document code: A
Article ID: SA202305001

Citation format: WANG Yapeng, CAO Shanshan, LI Quansheng, SUN Wei. Desert plant recognition method under natural background incorporating transfer learning and ensemble learning[J]. Smart Agriculture (Chinese and English), 2023, 5(2): 93-103.

Introduction

Desert vegetation maintains the material and energy cycles of desert ecosystems, playing crucial roles in wind control, desertification prevention, and microclimate improvement, while providing medicinal materials, forage, and timber as by-products for human use. As the core of desert ecosystems, it holds significant ecological and economic importance [1]. Currently, due to factors such as overgrazing, excessive development, and climate change, many desert plants have not been thoroughly understood or protected, suffering destruction or even extinction [2,3]. This is particularly true for nationally protected wild plants such as *Ammopiptanthus mongolicus* and *Betula halophila*, which are in endangered states [4]. Various resource surveys constitute long-term, fundamental work for desert plant research, protection, and utilization. During field operations, accurate and rapid determination of plant scientific names, families, genera, and traits is essential. However, the accuracy and consistency of plant classification are often constrained by surveyors' technical levels and subjective experience, making the process time-consuming and labor-intensive.

In recent years, rapid development in computer hardware technology and the proliferation of multi-source data acquisition methods such as smartphones, digital cameras, and UAVs have enabled breakthrough progress in deep learning algorithms, particularly Convolutional Neural Networks (CNNs) [5,6], for plant image classification and feature recognition [7]. Current plant classification research predominantly focuses on plant leaves or local organs in laboratory environments [8-10], achieving high recognition accuracy but proving difficult to apply to whole-plant image recognition in natural environments. Complex growth environments caused by lighting, soil, shadows, and other vegetation increase recognition difficulty [11], resulting in poor model generalization and compromised accuracy. Ensuring plant recognition accuracy in real-world environments has become a research hotspot and challenge. Feng et al. [12] proposed a tree species recognition method based on whole-tree images, achieving 99.15% accuracy—a significant improvement over leaf-based approaches. Song et al. [13] focused on plant leaves and flowers, proposing an effective region screening method for plant image recognition that improved accuracy and provided new insights for natural background plant image recognition. Zhou et al. [14] developed a vegetable disease recognition method based on region proposal and progressive learning for six types of natural complex background disease images, guiding the model to focus on key regions while avoiding costly manual labeling, achieving 98.26% accuracy. Li et al. [15] compared a series of high-performance CNN models on 24 desert plant species, concluding that MobileNet V2 achieved the optimal balance between accuracy, parameters, and Floating Point Operations (FLOPs). Currently, the key challenges in natural background plant image recognition lie in improving accuracy for plants within the same family or genus [16] and overcoming various natural noise factors affecting precision [17].

This study addresses whole-plant desert plant image recognition in natural environments by proposing a desert plant recognition method based on transfer learning and ensemble learning. To maximize utilization of desert plant images, K-fold cross-validation was employed for dataset partitioning, followed by transfer learning of K sub-models on the source ImageNet dataset. Voting-based ensemble methods were then used to achieve accurate desert plant recognition. Intelligent recognition algorithms can assist traditional manual identification of desert plants, providing technical support for desert plant understanding, diversity conservation, sand-fixing plant development, forage and medicinal plant utilization, and excellent fuelwood development, while reducing dependence on expert identification and saving time and labor costs.

2.1 Data Collection and Preprocessing

2.1.1 Dataset Construction

Desert plant image data were collected during two periods: late September 2021 and July-August 2022, in Changji and Tacheng regions of Xinjiang. Images were captured using smartphones and digital cameras with pixel dimensions of 3968$\times 2232 and 4800 \times$3200, respectively. Multi-angle and multi-lighting photography was employed, ensuring that desert plants occupied the central portion of images while incorporating noise factors such as shadows, sand and gravel, lighting variations, and occlusion by other vegetation to enhance dataset robustness. Expert identification results were recorded for each desert plant species, constructing a desert plant image dataset comprising 12,507 images across 50 plant species from 13 families and 43 genera, with each species represented by

183-339 images. Example desert plant images are shown in Figure 1 [Figure 1: see original paper].

2.1.2 Data Preprocessing

Due to large original image sizes, all images were uniformly resized to 600$×$600 pixels. The dataset was partitioned into test (2,481 images) and training sets (10,026 images) at a 2:8 ratio. The test set underwent no data augmentation to ensure data independence and non-repetition. The training set was then divided into four training subsets and one validation subset through five-fold cross-validation. Given the limited number of training subset images and the difficulty covering diverse real-world field conditions, models risked overfitting—exhibiting high validation accuracy but poor test performance, resulting in weak generalization and robustness. Therefore, data augmentation was applied during training subset training.

Through analysis of desert plant characteristics, five augmentation methods were employed: image translation, flipping, rotation, random color adjustment, and random black block occlusion (with long sides in [100,600] and short sides in [10,60]). These methods were randomly selected and combined with different trigger probabilities during augmentation, expanding the training subset fivefold. Using *Ephedra sinica* as an example, data augmentation examples are shown in Figure 2 [Figure 2: see original paper], where Figure 2(a) shows the original image, Figures 2(b)-2(f) demonstrate single-method effects, and Figures 2(g)-2(i) show combined effects from random selection of the five augmentation methods.

2.2 EfficientNet Network

The EfficientNet B0 network architecture consists of nine stages. Stage 1 performs downsampling through 3$×3convolution, stages2-8extractimagefeaturesbystackingsevenMBConvBlo convolution, Global Average Pooling, and a fully connected layer. The EfficientNet B0 network architecture is detailed in Table 1 .

The input image size for EfficientNet B0 is the commonly used 224$×$224 in engineering applications. Network depth and width are increased by stacking MBConv Blocks, while B1-B7 networks amplify depth, width, and image resolution based on B0. Dropout operations within MBConv Blocks randomly discard entire MBConv Blocks before shortcut connections merge with main branches, leaving only shortcut connections. This introduces stochastic depth during training, further amplifying differences among EfficientNet series networks and making them suitable as sub-models for ensemble learning.

2.3 Transfer Learning

Broadly defined, transfer learning involves leveraging existing knowledge, models, or structures to achieve learning objectives on target data [21]. Pre-training and fine-tuning have long been considered the most important manifestation of transfer learning, referring to training a network on a source domain and directly applying it to target domain data for fine-tuning. This study conducted transfer learning based on EfficientNet B0-B4 networks, freezing weights from stages 1-8 and training only the final 1$×$1 convolutional layer and fully connected layer, enabling rapid training of ideal models even with limited image quantities.

2.4 Ensemble Learning

2.4.1 Model Architecture

While classical machine learning methods are less effective than deep learning with abundant data, ensemble learning continues to play a significant role in deep learning [22]. Currently, introducing ensemble learning on top of deep learning remains an important means for many researchers to improve performance. The core difference between ensemble learning and other deep learning methods lies in its focus on the bias-variance tradeoff, making it valuable for all machine learning methods including deep learning. In statistical learning, model quality is primarily measured through bias and variance, with low bias and low variance considered ideal outcomes. The bias and variance diagram for ensemble learning is shown in Figure 4 [Figure 4: see original paper].

As illustrated in Figure 4, strategies are needed to reduce bias and variance when models exhibit high bias, high variance, or high bias-high variance scenarios. High bias can be reduced through increased model complexity, Boosting [23], and Stacking [24], while high variance can be addressed through reduced model complexity, Bagging [25], and Stacking. The core idea of ensemble learning involves training multiple models and combining them through specific methods to reduce bias and variance, thereby improving model performance.

This study employed an ensemble learning strategy combining Bagging and Stacking based on EfficientNet networks. The first layer adopted Stacking methods, introducing K-fold cross-validation for dataset partitioning and training K sub-models. Considering identical output features across models for this classification problem, the second layer used Bagging to integrate first-layer models through voting methods, comparing identical base learners with K base learners to select optimal base learners for constructing the ensemble model, reducing model bias and variance while enhancing recognition performance. The integrated model architecture for desert plant image recognition is shown in Figure 5 [Figure 5: see original paper].

2.4.2 Model Integration Strategy

Voting is an ensemble learning strategy following the majority rule that reduces variance by integrating multiple models to achieve high accuracy. Ideally, voting predictions should outperform any single model. For this desert plant image classification problem, voting predictions represent the most frequent prediction result across all models. Classification voting methods include soft voting and hard voting, with differences detailed in Table 2 .

As shown in Table 2, soft voting (Ensemble-Soft) and hard voting (Ensemble-Hard) can yield different conclusions for the same sample. For a simple binary classification, a sample's result is either A or B, with three sub-models producing three recognition results (and prediction probabilities). Hard voting follows the majority rule, while soft voting considers prediction probability as additional information, averaging probabilities and classifying as A if exceeding 50%, otherwise B. Relative to hard voting, soft voting incorporates prediction probability information, yielding more accurate predictions. Considering varying desert plant image feature extraction capabilities and recognition accuracy differences among sub-models during ensemble integration, a more reasonable weighted voting method (Ensemble-Weight) should be adopted. The prediction classification result for the i-th sub-model is represented by matrix J as shown in equation (1), where i = 1, 2, …, 5 and n = 1, 2, …, 50. Each row represents sub-model prediction probabilities for desert plant images. The integrated model obtains prediction result E by weighted averaging across sub-models, as shown in equation (2), where w_i = m_i / Σ_{i=1}^5 m_i, and m_i values are assigned based on recognition accuracy ranking on the test set, with the lowest accuracy sub-model receiving m_i = 1 and the highest receiving m_i = 5.

2.5 Evaluation Metrics

2.5.1 Top-k Accuracy

Image classification problems use test set top(k) accuracy to evaluate model performance, calculated as shown in equation (3), where k = 1, 2, …, 5; a represents the total number of test images, and r represents the number of correctly predicted images among the top k results. Commonly used Top-1 accuracy refers to the accuracy where the top-ranked category matches the actual result.

2.5.2 Classification Metrics

The confusion matrix is an indicator for evaluating model results, commonly used to assess classifier performance [26]. Based on the confusion matrix, precision, recall, and F1 Score are derived. For predictive classification models, higher accuracy is ideal, corresponding to larger TP (True Positive) and TN

(True Negative) quantities and smaller FP (False Positive) and FN (False Negative) quantities in the confusion matrix. The F1 Score physically represents the weighted average of precision and recall, with formula (6) assuming equal weights. Its value ranges from 0 to 1, where 1 represents optimal model output and 0 represents the worst. Classification indicators from the confusion matrix are shown in Table 3 .

2.6 Experimental Environment and Parameter Settings

To ensure consistent model operating environments and eliminate environmental impacts on results, all model training and testing processes in this experiment used identical experimental environments, detailed in Table 4 . During model training, batch size was set to 32, epochs to 100, loss function to cross-entropy loss, optimizer to stochastic gradient descent, and learning rate scheduling to cosine annealing with initial learning rate 0.01 decreasing to 0.0001 after 100 epochs.

3.1 Desert Plant Recognition Based on EfficientNet B0 Network

The study examined 50 desert plant species using EfficientNet B0 as the base network, constructing desert plant image recognition models through transfer learning and ensemble learning. Based on image quantities for the 50 classes, 20% of data was partitioned as the test set, with the remaining data divided through five-fold cross-validation, represented as DPi (i = 1, 2, …, 5) for each training/validation set. Recognition results on the test set based on EfficientNet B0 network are shown in Table 5 .

As shown in Table 5, five EfficientNet B0 models trained on different training sets achieved varying Top-1 accuracies on the same test set, ranging from 92.26% to 93.35% (a 1.09% difference), indicating dataset heterogeneity after five-fold cross-validation. F1 Scores also varied from 93.01% to 93.56% (a 0.55% difference), with Top-1 accuracy and F1 Score values being comparable, demonstrating high recognition accuracy and strong performance. After ensemble integration through voting strategies, the Ensemble-Soft model showed Top-1 accuracy and F1 Score improvements of 1.37% and 1.3% respectively over the EfficientNet B0-DP3 model. Soft voting marginally outperformed hard voting in accuracy and F1 Score, validating that soft voting's consideration of prediction probabilities yields more accurate results. The Ensemble-Weight model integrated through weighted voting showed minimal improvement over Ensemble-Soft and Ensemble-Hard models, yet demonstrated superior performance, confirming the effectiveness of the proposed weighted voting method.

However, the three ensemble models showed limited improvement over the five sub-models, suggesting that high homogeneity among the five EfficientNet B0 models constrained voting strategy effectiveness. The recognition performance significantly depended on the EfficientNet B0-DP5 model, indicating that networks with greater differentiation should be considered as sub-networks. Loss value variation curves for the five EfficientNet B0 networks on the validation set are shown in Figure 6 [Figure 6: see original paper].

3.2 Recognition Results Based on Differentiated Networks

EfficientNet B1-B7 networks amplify network depth, width, and image resolution based on EfficientNet B0, with Dropout operations in MBConv Blocks randomly discarding entire blocks before shortcut connections, leaving only shortcut connections. This introduces stochastic depth during training, further amplifying differences among EfficientNet series networks and making them suitable as sub-models for ensemble learning.

Considering model recognition efficiency, EfficientNet B0-B4 networks were selected as sub-models under identical training conditions. The B0 network remained unchanged, while DPi (i = 2, 3, 4, 5) datasets served as training/validation sets for EfficientNet B1-B4 networks. Desert plant image recognition models were constructed through transfer learning and ensemble learning, with results shown in Table 6 .

As shown in Table 6, Top-1 accuracy reached 96.65% and F1 Score 96.71% on the same test set, demonstrating that the EfficientNet B4 single network already achieved high precision with significant model performance. Since EfficientNet B1-B4 networks amplify all three scales based on B0, both Top-1 accuracy and F1 Score improved, proving that continuous expansion yields higher recognition accuracy. However, F1 Score improvements slowed, and EfficientNet B4-DP5 precision slightly exceeded EfficientNet B3-DP4, suggesting that further expansion might cause model degradation, increasing training costs and reducing recognition efficiency.

After ensemble integration through voting strategies, the Ensemble-Soft model improved Top-1 accuracy and F1 Score by 6.08% and 5.79% respectively over EfficientNet B0-DP1, with significant enhancement. Compared to EfficientNet B4-DP5, Ensemble-Soft showed 2.42% and 2.36% improvements, demonstrating that voting strategies are more suitable for sub-networks with low homogeneity in ensemble learning. Soft voting again marginally outperformed hard voting, while the Ensemble-Weight model achieved the highest accuracy and F1 Score of 99.23%, confirming superior performance of weighted voting.

Additionally, Top-1 accuracy for all 50 plant species tested using the Ensemble-Weight model reached 97% or higher. Loss value variation curves for EfficientNet B0-B4 networks on the validation set are shown in Figure 7 [Figure 7: see original paper].

3.3 Validation of Ensemble Learning Effectiveness

To validate the effectiveness of the proposed ensemble learning method, comparative experiments were conducted on the public Oxford Flowers102 dataset. This dataset contains 102 common UK flower species with 40-258 images per species (8,189 total images), featuring substantial variations in pose and lighting. Some categories show significant differences within the category and across several very similar categories. Given its public availability and morphological feature distribution similarity to our desert plant dataset—both having numerous categories with large quantity distribution differences—Oxford Flowers102 serves as an appropriate validation dataset. Example images are shown in Figure 8 [Figure 8: see original paper].

Based on image quantities for the 102 flower species, 20% of data was partitioned as the test set, with the remaining data divided through five-fold cross-validation, represented as OFi (i = 1, 2, …, 5). Model training employed identical strategies to the desert plant recognition models. Recognition results on the test set based on differentiated networks are shown in Table 7 .

As shown in Table 7, five sub-models demonstrated similar performance on Oxford Flowers102 as on the desert plant dataset, with comparable Top-1 accuracy and F1 Score at high levels. The three ensemble models improved accuracy and F1 Score over the five sub-models by maximums of 4.56% and 5.05%, and minimums of 1.94% and 2.29%, validating that the proposed ensemble learning strategy enhances model recognition accuracy and performance. Notably, F1 Score improvements for the five sub-models slowed, suggesting that further model volume increases might cause degradation, raising training costs and reducing recognition efficiency.

4.1 Discussion

Current plant image recognition research is largely limited to single-background or laboratory environments, focusing primarily on plant leaves or local organs. However, most plant images collected in practical applications feature natural backgrounds containing complex noise such as lighting variations, soil, weeds, other plants, and clustered growth. Consequently, whole-plant desert plants in natural backgrounds cannot be recognized using traditional methods. This study addresses desert plant image recognition in natural backgrounds, overcoming background interference and natural noise to improve accuracy while enhancing model generalization.

At present, desert plant classification and identification domestically and internationally rely mainly on traditional manual recognition and expert experience, consuming significant time and labor. Intelligent recognition algorithms can effectively save time and labor costs while providing support for field investigators and reducing dependence on expert identification. This study constructed a desert plant image dataset for 50 species from Changji and Tacheng regions in Xinjiang under natural backgrounds, with 183-339 images per species (12,507 total images). References [27-29] focused on whole-plant appearance features but studied grassland plants, field weeds, and indoor potted plants. Currently, no domestic research exists on desert plant image dataset construction and recognition, and limited international research addresses desert plant recognition in natural backgrounds. Reference [15] constructed a desert plant dataset containing 24 species and 2,331 images, achieving optimal balance between accuracy and volume after comparing multiple models. Our dataset exceeds this in both species count and quantity, with recognition accuracy far surpassing previous work. All desert plant images in this study were captured during field surveys without combining public datasets or web crawling, more authentically reflecting growth characteristics over time. The employed transfer and ensemble learning strategies effectively improved recognition accuracy, achieving 99.23% accuracy on the 50-species desert plant dataset with minimal classification errors across species. However, several issues require improvement:

1. The desert plant image dataset requires further supplementation and improvement, considering both species diversity and collection stages.
2. Other ensemble learning strategies or lightweight networks should be considered as sub-networks to address training difficulties and slow inference speeds.

Conflict of Interest Statement: This study has no conflicts of interest among researchers or with publicly disclosed research findings.

4.2 Conclusion

Based on the PyTorch deep learning framework, this study proposed a whole-plant desert plant image recognition algorithm integrating transfer learning and ensemble learning, using desert plant images collected under natural conditions in Changji and Tacheng, Xinjiang. The method addresses problems of low recognition accuracy, weak model robustness, and poor generalization for desert plant images in real field environments, applicable to field surveys, educational outreach, and scientific experiments, while filling domestic research gaps in desert plant image dataset construction and machine vision classification.

This study employed EfficientNet series networks, pre-training on ImageNet through transfer learning, then adopting an ensemble learning strategy combining Bagging and Stacking. The first layer used Stacking with K-fold cross-validation for dataset partitioning and model training, while the second layer used Bagging to integrate first-layer model output features through voting methods, yielding the complete ensemble learning model.

Single sub-models based on EfficientNet B0 achieved Top-1 accuracy ranging from 92.26% to 93.35%, already reaching relatively high precision. After ensemble integration, soft voting, hard voting, and weighted voting achieved accuracies of 93.63%, 93.55%, and 93.67% respectively, with F1 Scores comparable to accuracy rates. Improvements over single sub-models were limited due to high homogeneity among sub-models constraining voting strategy effectiveness. Using more differentiated networks as sub-models (EfficientNet B0-B4), single sub-models achieved Top-1 accuracy up to 96.65% with F1 Score of 96.71%, demonstrating excellent performance. Soft voting, hard voting, and weighted voting achieved accuracies of 99.07%, 98.91%, and 99.23% respectively, with further improved accuracy and F1 Scores essentially matching accuracy rates, showing significant model performance. Validation on the public Oxford Flowers102 dataset confirmed that the proposed ensemble learning strategy effectively improves model recognition accuracy and performance.

References

[1] SONG Z F. Response of *Seriphidium transiliense* vegetation characteristics to grazing disturbance in desert grasslands[D]. Urumqi: Xinjiang Agricultural University, 2018.

[2] TENG Y F. Studies on diversity of the plants in Shahu nature reserve, Ningxia, China[D]. Yinchuan: Ningxia University, 2013.

[3] YAN H. The response of two representative desert shrubs to salt stress in northwest arid region[D]. Yangling: Northwest A & F University, 2012.

[4] HE H B. Studies on communities and rhizobium of *Ammopiptanthus mongolicus* (maxim.)[D]. Beijing: Beijing Forestry University, 2008.

[5] LECUN Y, BENGIO Y, HINTON G. Deep learning[J]. Nature, 2015, 521(7553): 436-444.

[6] GOODFELLOW I, BENGIO Y, COURVILLE A. Deep learning[M]. Cambridge, Massachusetts: The MIT Press, 2016.

[7] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. ImageNet classification with deep convolutional neural networks[J]. Communications of the ACM, 2017, 60(6): 84-90.

[8] JEON W S, RHEE S Y. Plant leaf recognition using a convolution neural network[J]. The international journal of fuzzy logic and intelligent systems, 2017, 17(1): 26-34.

[9] LEE S H, CHAN C S, WILKIN P, et al. Deep-plant: Plant identification with convolutional neural networks[C]//2015 IEEE International Conference on Image Processing (ICIP). Piscataway, NJ, USA: IEEE, 2015: 452-456.

[10] HAN B, ZENG S W. Plant leaf image recognition based on multi-feature integration and convolutional neural network[J]. Computer science, 2021, 48(S1): 113-117.

[11] JIN L T. Research on plant image recognition with complex background based on convolution neural network[D]. Lanzhou: Lanzhou Jiaotong University, 2020.

[12] FENG H L, HU M Y, YANG Y H, et al. Tree species recognition based on overall tree image and ensemble of transfer learning[J]. Transactions of the Chinese society for agricultural machinery, 2019, 50(8): 235-242, 279.

[13] SONG X Y, JIN L T, ZHAO Y, et al. Plant image recognition with complex background based on effective region screening[J]. Laser & optoelectronics progress, 2020, 57(4): 181-191.

[14] ZHOU J, LI J X, WANG C S, et al. A vegetable disease recognition model for complex background based on region proposal and progressive learning[J]. Computers and electronics in agriculture, 2021, 184: ID 106101.

[15] LI J C, SUN S D, JIANG H R, et al. Image recognition and empirical application of desert plant species based on convolutional neural network[J]. Journal of arid land, 2022, 14(12): 1440-1455.

[16] CAO X Y, SUN W M, ZHU Y X, et al. Plant image recognition based on family priority strategy[J]. Journal of computer applications, 2018, 38(11): 3241-3245.

[17] GUO X L. Research and implementation on plant image segmentation algorithm based on neural network[D]. Hohhot: Inner Mongolia University, 2021.

[18] RAGHU M, POOLE B, KLEINBERG J, et al. On the expressive power of deep neural networks[C]//Proceedings of the 34th International Conference on Machine Learning - Volume 70. New York, USA: ACM, 2017: 2847-2854.

[19] ZAGORUYKO S, KOMODAKIS N. Wide residual networks[EB/OL]. arXiv: 1605.07146, 2016.

[20] TAN M, LE Q. EfficientNet: Rethinking model scaling for convolutional neural networks[EB/OL]. International conference on machine learning. arXiv:1905.11946, 2019.

[21] PAN S J, YANG Q. A survey on transfer learning[J]. IEEE transactions on knowledge and data engineering, 2010, 22(10): 1345-1359.

[22] DONG X B, YU Z W, CAO W M, et al. A survey on ensemble learning[J]. Frontiers of computer science, 2020, 14(2): 241-258.

[23] WANG B, PINEAU J. Online bagging and boosting for imbalanced data streams[J]. IEEE transactions on knowledge and data engineering, 2016, 28(12): 3353-3366.

[24] HUI Y, MEI X S, JIANG G D, et al. Milling tool wear state recognition by vibration signal using a stacked generalization ensemble model[J]. Shock and vibration, 2019, 2019: 1-16.

[25] ANDIOJAYA A, DEMIRHAN H. A bagging algorithm for the imputation of missing values in time series[J]. Expert systems with applications, 2019, 129: 10-26.

[26] FIELDING A H, BELL J F. A review of methods for the assessment of prediction errors in conservation presence/absence models[J]. Environmental conservation, 1997, 24(1): 38-49.

[27] GAO H Y, GAO X H, FENG Q S, et al. Approach to plant species identification in natural grasslands based on deep learning[J]. Pratacultural science, 2020, 37(9): 1931-1939.

[28] PENG W, LAN Y B, YUE X J, et al. Research on paddy weed recognition based on deep convolutional neural network[J]. Journal of South China agricultural university, 2020, 41(6): 75-81.

[29] CHEN S J, ZHOU Y X, FANG Y J. The plant species recognition based on the whole appearance features[J]. Computer applications and software, 2017, 34(9): 222-227.