PlantNet: Transfer learning Based Fine-grained Network for High-throughput Plants Recognition

Ziying Yang[1#], Wenyan He[1#], Xijian Fan[1]*, Tardi Tjahjadi[2]

1. *College of Information Science and Technology, Nanjing Forestry University, Nanjing, China*
2. *School of Engineering, University of Warwick, CV4 7AL, Coventry, United Kingdom*

Abstract: In high-throughput phenotyping, recognizing individual plant category is an important support process for plant breeding. However, different plant categories have different fine-grained characteristics, i.e., intra-class variation and inter-class similarity, which make the process challenging. Existing deep learning-based recognition methods fail to effectively address this recognition task under such challenging requirements, leading to technical difficulties such as low accuracy and lack of generalization robustness. To address these requirements, this paper proposes PlantNet, a fine-grained network for plant recognition based on transfer learning and a bilinear convolutional neural network, which achieves high recognition accuracy in high-throughput phenotyping requirements. The network operates as follows. First, two deep feature extractors are constructed using transfer learning. The outer product of the different spatial locations corresponding to the two features are then calculated, and the bilinear convergence is computed for the different spatial locations. Finally, the fused bilinear vectors are normalized via maximum expectation to generate the network output. Experiments on a publicly available Arabidopsis dataset show that the proposed bilinear model performed better than related state-of-the-art methods. The interclass recognition accuracy of the four different species of Arabidopsis Sf-2, Cvi, Landsberg and Columbia are found to be 98.48%, 96.53%, 96.79% and 97.33%, respectively, with an average accuracy of 97.25%. Thus, the network has good generalization ability and robust performance, satisfying the needs of fine-grained plant recognition in agricultural production.

Keywords: fine-grained recognition; convolutional neural network; Bilinear-CNN; transfer learning

1.Introduction

Plant phenotype is a set of observable and measurable characteristics and traits of a plant, which is significantly affected by the interaction between plant gene expression and environmental influences [1]. The monitoring of phenotype is used to provide guidance for plant cultivation, a prerequisite for intelligent production and planting, and information/data management. The identification of plant species is an important application of plant phenotype detection and plays an important role in ecological monitoring to effectively detect biological growth and protect biodiversity [2, 3]. Accurate and efficient identification of plant species to obtain their physiological information are essential for effective monitoring of the distribution of biological species and the impact of ecological changes on the distribution of species in a geographical area. It also enables the realization of information technology, data gathering and automation in agriculture [4].

Traditional plant identification methods mainly rely on manual observation and measurement to analyze the appearance of plants in terms of shape, texture [5, 6], colour [7, 8]，and other characteristic morphological phenotypes. These methods are not efficient and have low recognition accuracy. With the recent development of deep learning in computer vision [9, 10], the application of deep convolutional neural network (CNN) or deep learning to process two-dimensional natural images has become one of the popular research topics. Deep learning techniques have been employed in agriculture and forestry for plant phenotyping, opening up an era of intelligent plant phenotyping [11]. Due to significant advances in image acquisition system, high throughput plant image collection is becoming widespread. In particular, deep learning has demonstrated reliability and efficiency in processing large amount of data under high throughput plant analysis requirements [11]. However, little research has been reported in the area of high-throughput plant recognition task.

In this paper, we propose a plan recognition network (PlantNet) model for plant phenotyping task. The novel contributions of this paper are as follows. PlantNet is based on transfer learning and takes into account the unique characteristics of plant image so as to address the fine-grained problem, i.e., inter-class similarity and intra-class variation in high-throughput plant phenotyping task. PlantNet utilizes weakly supervised learning for fine-grained image classification and includes three main modules. First, two deep feature extractors are constructed using transfer learning. Then, the outer product of the different spatial locations corresponding to the two features is calculated, and the bilinear convergence is computed for the different spatial locations to uncover the subtle local differences in plant phenotype. Finally, the bilinear fused vectors are normalized to obtain the output of the recognition model via maximum expectation. We also introduce a data augmentation process to improve the generalization ability of the PlanNet model. Experiments on a dataset of four different species of Arabidopsis achieved a high recognition accuracy of 97.25%, outperforming related state-of-the-art methods, and demonstrating the effectiveness of PlantNet for plant recognition.

The rest of this paper is organized as follows. Section 2 presents the related work. Section 3 presents the details of the PlantNet. Section 4 presents the data sources and pre-proposing. Section 5 presents the experiments and evaluation results. Finally, Section 6 concludes the paper.

2.Related Work

Deep learning techniques have been exploited for plant phenotyping analysis [12, 13], e.g., plant image recognition, leaf counting, plant segmentation, etc. Grinblat et al. [14] developed a method for extracting leaf vein texture features to recognize three legumes of different species with a recognition rate of 96.9% and a tolerance of ± 0.2%. Liu et al. [15] improved the accuracy of leaf classification of 220 different plants to 93.9% using multi-feature fusion and an improved deep confidence network. Zhu et al. [16] achieved 100% classification based on multi-channel sparse coding feature extraction method with Scale-Invariant Feature Transform (SIFT) and multi-channel sparse coding on images. Taghavi et al. [17] identified and classified a dataset consisting of top-view images of different species of Arabidopsis thaliana through a coupled algorithm of CNNs, and CNN combined with long short-term memory network (CNN-LSTM). Their networks successfully improved the identification and classification of Arabidopsis thaliana to 93%, an improvement of 5.4% over the average recognition rate of 87.6% with CNNs and CNN combined with conditional random fields (CNN-CRFs). Nguyen et al. [18] used transfer learning to build a pervasive crop recognition system that adapts to the uneven distribution of plants in different regions through flexible data collection. They evaluated the effectiveness of crop recognition using different backbone network architectures, e.g., AlexNet [19] and VGG [20]. The results show that the method extracts richer and more reliable features, gaining better recognition performance.

The above-mentioned approaches directly apply deep learning models without any specific analysis of the plant images, thus, ignoring the unique characteristics that exist in the plant images. However, the following need to be considered to directly leverage image classification in plant images: 1) external factors, e.g., camera shooting angle, lighting, complex backgrounds during data acquisition, result in a wide variation of samples belonging to the same type of image, i.e., **intra-class variation**; and 2) there is a detailed division of plants into subclasses or subgroups, and each subclass is bio-morphologically similar to the other, leading to the problem of fine-grained identification, i.e., **inter-class similarity.** Therefore, a recognition method that can effectively characterize small local differences in different plant categories and capture common information of the same plant category is needed to data mine the knowledge structure associated with the plant images. There are two common approaches on fine-grained image classification: strongly supervised learning [21] and weakly supervised learning [22]. The former requires a large amount of manual annotation of local positions in order to achieve more fine-grained localization. However, this is time-consuming and labor-intensive, leading to significant limitations in practical applications [23]. The latter tends to automatically extract local features and focus on the intrinsic connection between the local and the entire image area, which has gradually gain attention. These methods include Navigator-Teacher-Scrutinizer Network (NTS-Net) [24], Faster Training of Multi-global Covariance Pooling Networks (FAST-MPN) [25] and Discriminative Filter Bank Learning (DFL) Networks [26]. These have emerged in public fine-grained datasets and challenge competitions, continuing to improve the recognition accuracy. However, the above-mentioned methods are not practical for high-throughput plant recognition task.

3.Proposed PlantNet

Due to the differences in capability of different deep learning networks in extracting features for identifying plant phenotypes, we only exploit the ideas presented in [27] and apply bilinear CNN model to enhance their discriminative ability for feature extraction from a plant image. The structure of the proposed PlanNet is illustrated in Figure 1. The use of bilinear CNN, CNN stream A and CNN stream B, enable the recognition model to effectively extract subtle differences among high-throughput plant species, where the bilinear local feature descriptors are extracted using bilinear pooling for better robustness in recognizing plant phenotypes. The feature descriptors are fused together and then normalized via maximum expectation.

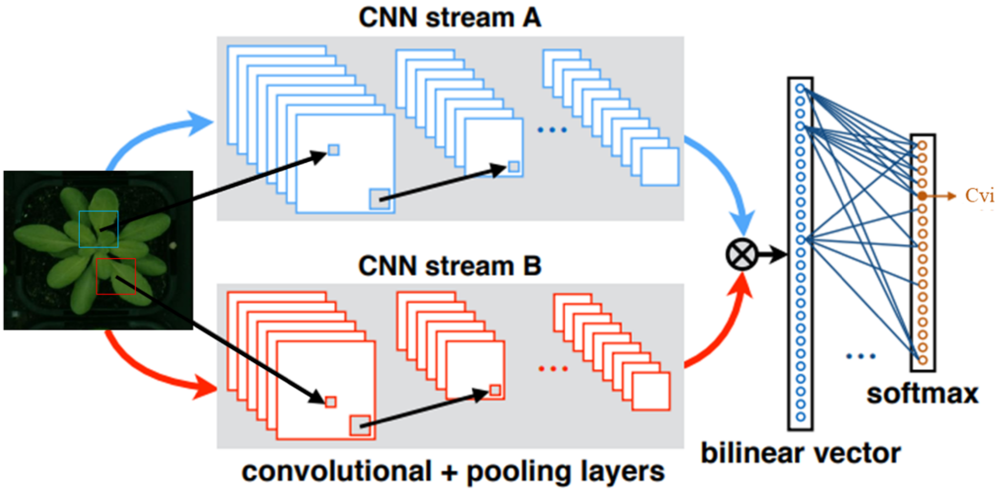


Figure 1 The proposed bilinear CNN model for image classification.

3.1. Deep CNN

Deep CNN is one of the most widely used machine learning models in general image classification task which has shown promising performance. The sharing of weights in deep CNN architecture enables the discovery of discriminative and robust features for image classification. The convolutional layers, pooling layers and fully connected layers are three main modules of CNN based models [28], that are invariant to a certain extent to translation, warping, and deflation of a two-dimensional graph [29]. The convolutional layer is modelled after the biological visual perception mechanism and consists of a fixed-size convolutional kernel which acts as a filter to extract features from plant leaf images. The pooling layers follow the convolutional layers and perform down-sampling and retain the most important features in the plant images [30]. PlanNet uses CNN to construct classification models.

3.2. Bilinear Model and Transfer Learning

Transfer learning [31] is a machine learning technique which solves new problems by transferring relevant network model parameters learned in an existing problem domain to a task in the target domain to learn new feature information. When the original domain dataset is much larger than the target domain, transfer learning is a powerful method for model learning and training.

The bilinear recognition model consists of two CNN feature extractors [27]. It obtains the bilinear descriptor features of an image by multiplying the external products of different spatial locations and averaging them. Its architecture interactively models pairwise correlations among feature channels in a translation-invariant manner, providing a stronger feature representation for detecting local regions without cumbersome image annotation.

The proposed PlantNet for feature extraction of plant images using bilinear model consists of two weakly supervised networks. Both of the two base networks use VGG16 [20] as backbone networks to perform coarse-grained feature extraction. The networks are pre-trained on ImageNet [19] for initialization, and then transferred to the Arabidopsis dataset used for fine-grained feature extraction and model parameter updating. We slightly modify VGG16 by removing the final pooling layer, the two fully connected and softmax layers and globally using a 3-by-3 convolutional layer and a 2-by-2 pooling layer. The ReLU activation function i.e.,

$$ReLU(z) = \begin{cases} x & x > 0 \\ 0 & x <= 0 \end{cases} \tag{1}$$

is applied for more efficient gradient calculation and back propagation while simplifying the computation to avoid gradient disappearance and gradient explosion. The feature extraction function $f(\cdot)$ is constructed to derive the coarse-grained features X_1, i.e.,

$$X_1 = H_{vgg}(x, y, \{W_1, b_1, \delta_{\mathrm{relu}}\}) \tag{2}$$

where H_{vgg} denotes the conv layer of VGG16, (x, y) is the input feature parameter of the image, W_1 denotes the weight parameter of the network model to be iteratively updated, b_1 is the bias, and δ denotes the Relu activation function.

The overall representation of the proposed bilinear model is

$$B = (f_A, f_B, P, C) \tag{3}$$

where P denotes the pooling layer, which serves to down-sample the feature map, and C denotes the classification function. The common classifiers include logistic regression, support vector machine and naive bayes classifier. There are two feature extraction functions A and B in the model, which serve to map the image I with location l into a $C \times D$ dimensions feature. The two features $f_A(l, \mathrm{I}) \in {}^{\sim c\times \mathrm{M}}$ and $f_B(l, \mathrm{I}) \in {}^{\sim c\times N}$ are bilinearly fused and multiplied at the same location to obtain the $M \times N$ dimensions of matrix.

$$b(l, \mathrm{I}, f_A, f_B) = f_A^T(l, \mathrm{I}) f_B(l, \mathrm{I}). \tag{4}$$

To streamline the computational complexity, this algorithm uses summation pooling to accumulate the matrices b at all positions to obtain the matrix.

$$\xi(\mathrm{I}) = \sum_l b(l, \mathrm{I}, f_A, f_B) \tag{5}$$

A multidimensional vector expansion is performed on the matrix ξ to obtain the feature vector.

$$x = vec(\xi(\mathrm{I})), x \in {}^{\sim MN\times 1} \tag{6}$$

Finally, the obtained feature vectors are normalized, and a combination of moment normalization operation and L2 normalization operation is used to obtain the final features for fine-grained classification. The normalized feature description is denoted as

$$z : z = y / \|y\|_2 \ , \ y = sign(x)\sqrt{|x|}. \tag{7}$$

In the practical application of the model, let f_A and f_B output the tensor to get $M \times N \times \mathrm{P}$ dimensions at position L, giving M×N-dimensional position points. Each position point is a P×P-dimensional vector extracted after bilinear transformation and a $P \times 1$-dimensional feature vector after cumulative pooling. Finally, the end-to-end training is completed by back-propagating the derivative chain rule. Assuming that the bilinear feature after cumulative pooling is

$$x = A^T B, \tag{8}$$

then the gradient expression of the loss function on the feature vector is

$$\frac{dl}{dA} = B(\frac{dl}{dx})^T, \frac{dl}{dB} = A(\frac{dl}{dx})^T. \tag{9}$$

There are currently two types of bilinear CNN implementations [27]. One is multimodal bilinear pooling, where the two features extracted from a uniform sample are derived from two different feature extraction functions. The second is homologous bilinear pooling or second-order pooling, where two features are extracted by the same feature extractor. To combine features such as plant leaf texture and shape of plant phenotypic species, PlanNet uses a second-order pooling representation which is more suitable for our application requirements. The overall structure of the PlantNet model using transfer learning is shown in Figure 2.

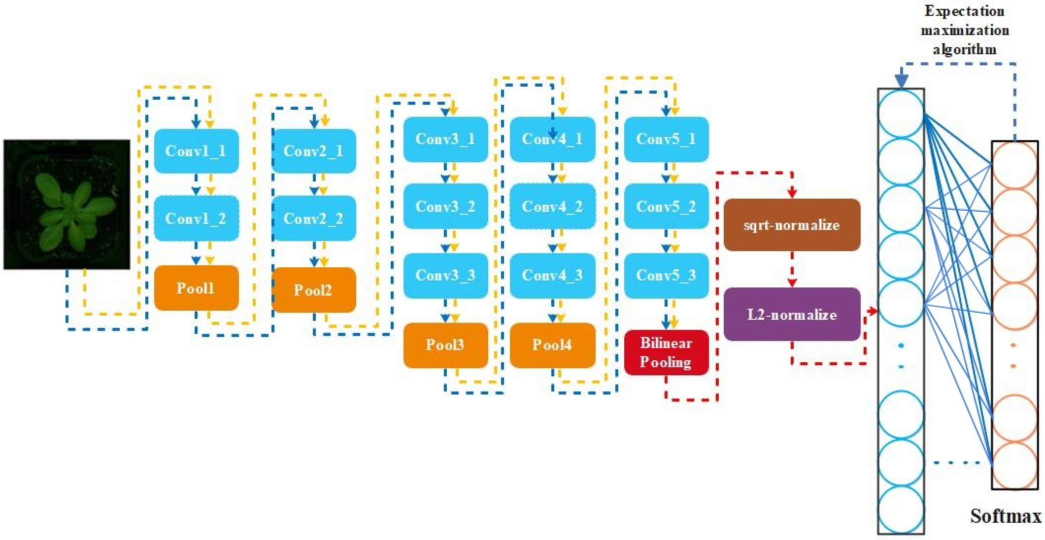


Figure 2 Overall structure of the PlantNet model using transfer learning.

4. Data Sources

4.1. Data

Arabidopsis is selected for our research work because it has a good genome sequence that can be used for plant phenotyping studies. We use a published phenotypic dataset [17] for this experiment. The dataset consists of consecutive top-view images of four different species of Arabidopsis, i.e., Sf-2, Cvi, Landsberg (Ler-1) and Columbia (Col-0). Figure 3 shows some data samples. In this dataset different species of Arabidopsis are grown in substrates with strictly controlled environmental conditions such as soil and light. Stationary cameras are mounted above the plants and 22 top-view images were taken at a fixed rate to construct a data sequence for each plant recorded at 12:00 pm each day.

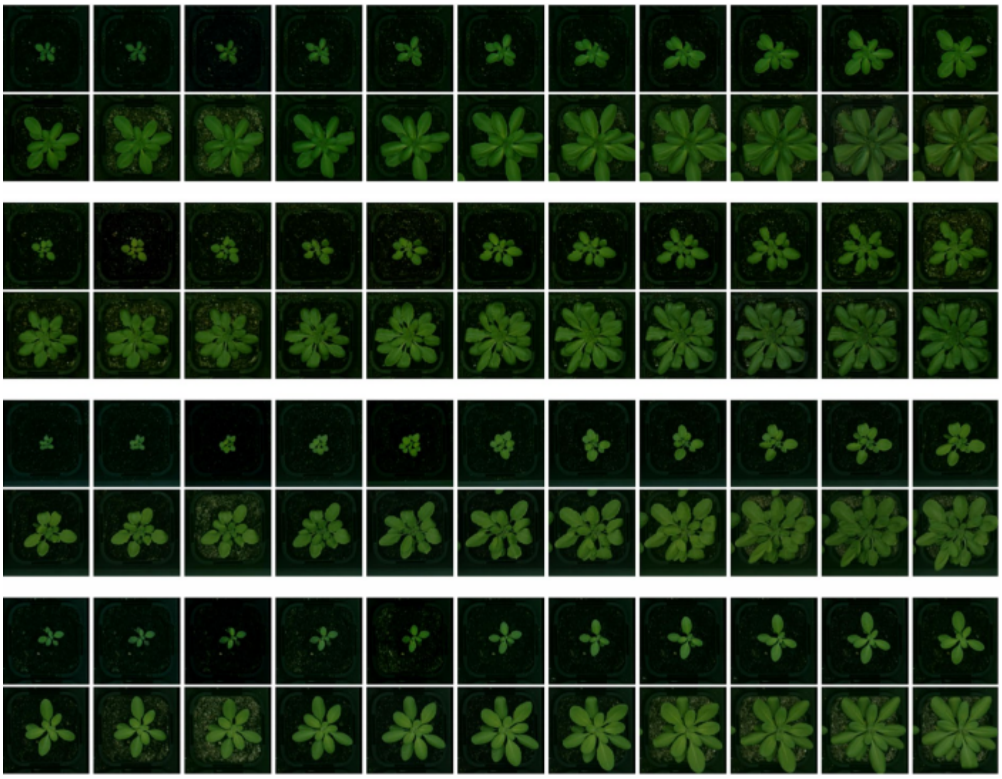


Figure 3 Data samples from Arabidopsis.

4.2 Data Enhancement

In order to improve the generalization ability of the PlanNet model and to avoid over-fitting, we increase the amount of data using data augmentation methods [32] before training. The data augmentation methods used in this paper are as follows: 1) Random crops: Randomly crop different areas of a larger image to extract local information, then resizing the pixels in the cropped area to 448×448; 2) Horizontal flip and random rotation: Flip the image horizontally and rotate it randomly with probability p (p=0.5) to simulate the difference in angle of real plant growth to improve deformation adaptability; 3) Colour enhancement: Change colour brightness, saturation, etc. to suit lighting and other camera shooting conditions; 4) Normalization: Normalization of images with mean (0.485, 0.456, 0.406) and standard deviation (0.229, 0.224, 0.225). The pre-treatment results of each data augmentation method are shown in Figure 4.

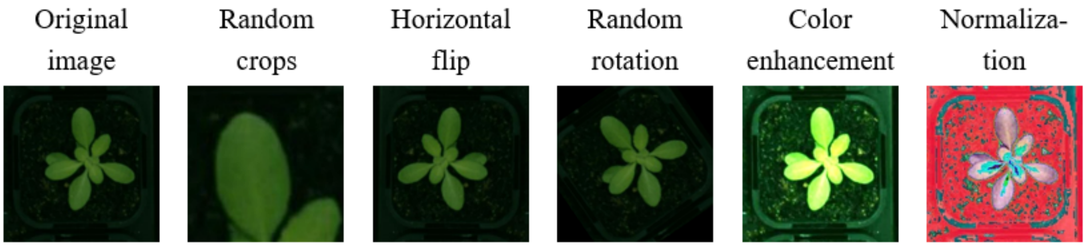


Figure 4 Illustrations of data augmentation.

5. Experiments

5.1. Experimental Environment Setup

To improve the reliability of training, 8/11 of each class of the Arabidopsis dataset (1552 sheets in total) are randomly used as the training set and the remaining 582 sheets are used as the test set. Based on this division, we built a cloud server platform with Ubuntu 16.04LTS as the operating system, which is equipped with dual-core Intel Core i7-8600@3.6 GHz x8 processor, 256 G of RAM and 4×4T solid state drives, NVIDIA Tesla p40 GPU as the graphics card, computational cache of 96 G, and the deep learning framework Pytorch.

5.2. Comparative Results

To verify the performance advantages of our bilinear CNN model for fine-grained plant recognition, three related state-of-the-art coarse-grained deep networks, i.e., VGG16 [20], ResNet18 [33], and DenseNet161 [34] were trained. The parameters used in the training process are batch_size, activation function, optimization function and number of iterations, and as shown in Table 1. The classification of the four different species of Arabidopsis, i.e., Columbia (Col-0), Landsberg (Ler-1), Sf-2 and Cvi using the four above-mentioned network models are shown in Table 2.

Table 2 shows that the recognition accuracy using VGG16 and ResNet18 are low because the coarse-grained networks focus on the obvious differences in the feature maps but ignores the subtle differences between classes. The average accuracy using DenseNet161 is 94.85%. Although DenseNet161 has a deeper network structure and is able to extract deep plant phenotypic features, it does not take into account subtle differences between classes. The model proposed in this paper, which uses a bilinear convolutional neural network for fine-grained plant recognition, achieves an accuracy of 97.25%, i.e., higher than the general coarse-grained networks.

Table 1 Parameters setting.

Parameter Category	Value
Batch_size	32
Activation function	ReLU
Optimizer	SGD
Number of iterations	55

Table 2 Classification results.

	Col-0	Ler-1	Sf-2	Cvi	Avg
Visual Geometry Group Networks-16 layers (VGG16) [20]	87.33	83.33	93.94	83.33	86.77
Residual network-18 layers (ResNet18) [33]	89.33	86.54	92.42	85.42	88.32
Dense Networks-161 layers (DenseNet161) [34]	97.33	94.87	95.45	91.67	94.85
Bilinear-CNN (proposed)	97.33	96.79	98.48	96.53	97.25

5.3. Analysis of experimental results

Gradient-weighted class activation mapping (Gram-CAM) [35] is a method that presents the activation level of a neural network for each pixel of an image in the form of weight. It uses the output of the model to derive the output features of the convolutional layer to obtain the feature activation map. In order to more visually represent the degree of response of the model to a region, the bilinearly fused features are visualized by Gram-CAM as shown in Figure 5. In the figure, the blue region indicates where the model network is not highly activated, and the red region indicates where the model is more sensitive. The visualization shows the fused feature is capable of uncovering the essential local parts of plant images.

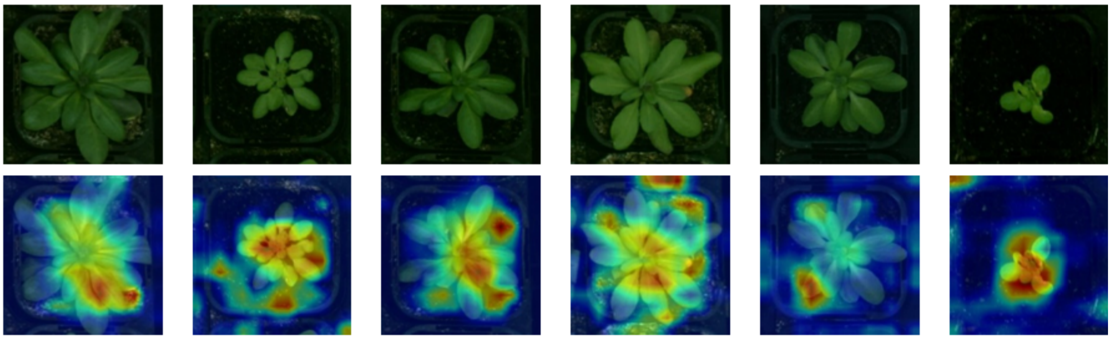
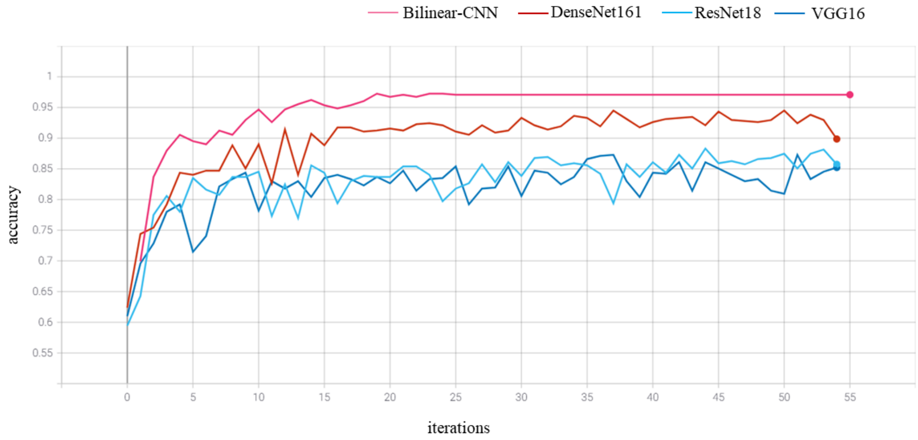


Figure 5. Grad-CAM visualization of the bilinearly fused features.

The average accuracy and loss function curves for plant recognition of VGG16, ResNet18 and DenseNet161 and bilinear models are plotted to provide an accurate portrayal of the recognition performance of each type of network, as shown in Figure 6 and Figure 7. The figures show that the bilinear model has better average accuracy and loss function than the other three types of classical CNN networks, and converges faster to reach stability. The results show that the bilinear CNN model effectively enhances the model robustness and better adapts to the fine-grained plant identification than the coarse-grained CNNs.



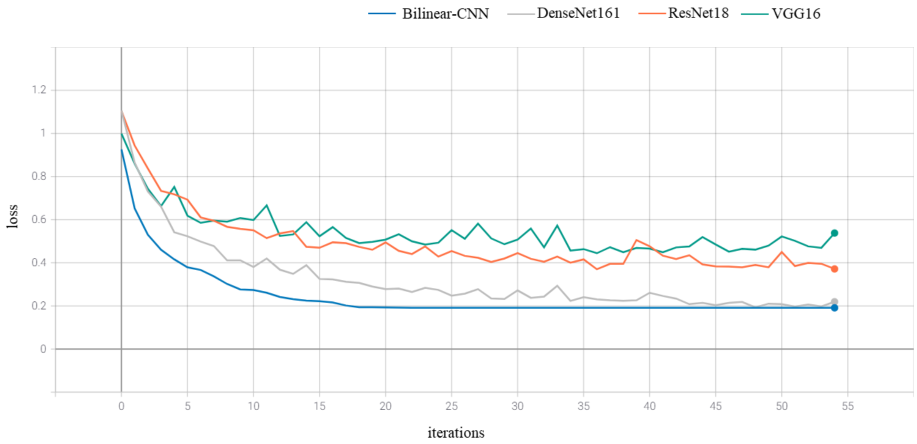


Figure 7 Curve of loss comparison.

6. Conclusion

This paper presents PlanNet to address the challenge of fine-grained plant recognition in complex agricultural requirements. PlanNET is a bilinear CNN-based recognition model based on transfer learning and fine-tuned CNN after data augmentation. Our experiments on a publicly available Arabidopsis dataset show that PlanNET has high recognition performance, with recognition accuracy up to 97.25%. Its generalization ability and robustness are better than other comparative deep network models. These satisfy the high demand of fine-grained plant recognition in agricultural production practice, providing accurate and reliable plant recognition for the majority of agricultural producers with high efficiency and low cost. The proposed network model provides a theoretical and technical basis for expanding the application of fine-grained plant recognition in agricultural production.

Acknowledgements

The authors acknowledge the National Natural Science Foundation of CHINA (Grant, 61902187), Jiangsu Province Innovative and Entrepreneurial Talent Project, Nanjing Forestry University High-level Start-up Foundation for Research, and National College Students Innovation and Entrepreneurship Training Program (Grant, 202010298123Y).

References

[1] Siebner, H.R., Callicott, J.H., Sommer, T. and Mattay, V.S., (2009). From the genome to the phenome and back: linking genes with human brain function and structure using genetically informed neuroimaging, 1-6.

[2] Farnsworth, E.J., Chu, M., Kress, W.J., et al., (2013). Next-generation field guides. BioScience, 63(11), 891-899.

[3] Elphick, C.S., (2008). How you count counts: the importance of methods research in applied ecology. Journal of Applied Ecology, 45(5), 1313-1320.

[4] Gao, W., Li, Z., Yu, L. and Wang, J., (2010). Speed up development of agricultural informatization and improve construction of agricultural modernization. Research of Agricultural Modernization, 31(3),

257-261.

[5] Cointault, F., Journaux, L., Miteran, J. and Destain, M.F., (2008). Improvements of image processing for wheat ear counting. Agricultural and Biosystems Engineering for a Sustainable World.

[6] Cointault, F., Journaux, L., Rabatel, G., Germain, C., Ooms, D., Destain, M.F., Gorretta, N., Grenier, G., Lavialle, O. and Marin, A., (2012). Texture, color and frequential proxy-detection image processing for crop characterization in a context of precision agriculture. pp. 49-70, InTech.

[7] Duan, L., Huang, C., Chen, G., Xiong, L., Liu, Q. and Yang, W., (2015). Determination of rice panicle numbers during heading by multi-angle imaging. The Crop Journal, 3(3), 211-219.

[8] Cointault, F., Guerin, D., Guillemin, J.P. and Chopinet, B., (2008). In-field Triticum aestivum ear counting using colour-texture image analysis. New Zealand Journal of Crop and Horticultural Science, 36(2), 117-130.

[9] Qiaolin Y., Zechao L., and Liyong F., et al. (2019) Nonpeaked Discriminant Analysis. IEEE Transactions on Neural Networks and Learning Systems, 30(12), 3818-3832.

[10] Qiaolin Y., Henghao Z., and Zechao L., et al. (2018) L1-norm Distance Minimization Based Fast Robust Twin Support Vector k-plane clustering, IEEE Transactions on Neural Networks and Learning Systems, 29(9), 4494-4503.

[11] Granier, C. and Vile, D., (2014). Phenotyping and beyond: modelling the relationships between traits. Current opinion in plant biology, 18, 96-102.

[12] Albawi, S., Mohammed, T.A. and Al-Zawi, S., (2017), August. Understanding of a convolutional neural network. In 2017 International Conference on Engineering and Technology (ICET). 1-6, IEEE.

[13] Lee, S.H., Chan, C.S., Wilkin, P. and Remagnino, P., (2015), September. Deep-plant: Plant identification with convolutional neural networks. In 2015 IEEE international conference on image processing (ICIP). 452-456, IEEE.

[14] Grinblat, G.L., Uzal, L.C., Larese, M.G. and Granitto, P.M., (2016). Deep learning for plant identification using vein morphological patterns. Computers and Electronics in Agriculture, 127, 418-424.

[15] Liu, N. and Kan, J.M., (2016). Improved deep belief networks and multi-feature fusion for leaf identification. Neurocomputing, 216, 460-467.

[16] Zhu, H., Huang, X., Zhang, S. and Yuen, P.C., (2017). Plant identification via multipath sparse coding. Multimedia Tools and Applications, 76(3), 4599-4615.

[17] Namin, S.T., Esmaeilzadeh, M., Najafi, M., Brown, T.B. and Borevitz, J.O., (2018). Deep phenotyping: deep learning for temporal phenotype/genotype classification. Plant methods, 14(1), 1-14.

[18] Nguyen, T.T.N., Le, T.L., Vu, H. and Hoang, V.S., (2019). Towards an Automatic Plant Identification System without Dedicated Dataset. International Journal of Machine Learning and Computing, 9(1), 26-34.

[19] Krizhevsky, A., Sutskever, I. and Hinton, G.E., (2012). Imagenet classification with deep convolutional neural networks. Advances in neural information processing systems, 25, 1097-1105.

[20] Simonyan, K. and Zisserman, A., (2014). Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556.

[21] Xu, Z., Huang, S., Zhang, Y. and Tao, D., (2015). Augmenting strong supervision using web data for fine-grained categorization. In Proceedings of the IEEE international conference on computer vision, 2524-2532.

[22] Xu, Z., Tao, D., Huang, S. and Zhang, Y., (2016). Friend or foe: Fine-grained categorization with weak supervision. IEEE Transactions on Image Processing, 26(1), 135-146.

[23] Song, K., Yang, H. and Yin, Z., (2018). Multi-scale attention deep neural network for fast accurate object detection. IEEE Transactions on Circuits and Systems for Video Technology, 29(10), 2972-2985.

[24] Yang, Z., Luo, T., Wang, D., Hu, Z., Gao, J. and Wang, L., 2018. Learning to navigate for fine-grained classification. In Proceedings of the European Conference on Computer Vision (ECCV) (pp. 420-435).

[25] Li, P., Xie, J., Wang, Q. and Gao, Z., (2018). Towards faster training of global covariance pooling networks by iterative matrix square root normalization. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 947-955.

[26] Wang, Y., Morariu, V.I. and Davis, L.S., (2018). Learning a discriminative filter bank within a cnn for fine-grained recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition, 4148-4157.

[27] Lin, T.Y., RoyChowdhury, A. and Maji, S., 2015. Bilinear cnn models for fine-grained visual recognition. In Proceedings of the IEEE international conference on computer vision, 1449-1457.

[28] Yu, F., Liu, L., Xiao, L., Li, K. and Cai, S., (2019). A robust and fixed-time zeroing neural dynamics for computing time-variant nonlinear equation using a novel nonlinear activation function. Neurocomputing, 350, 108-116.

[29] O'Shea, K. and Nash, R., 2015. An introduction to convolutional neural networks. arXiv preprint arXiv:1511.08458.

[30] Chen, W.R., Midtgaard, J. and Shepherd, G.M., (1997). Forward and backward propagation of dendritic impulses and their synaptic control in mitral cells. Science, 278(5337), 463-467.

[31] Donahue, J., Jia, Y., Vinyals, O., Hoffman, J., Zhang, N., Tzeng, E. and Darrell, T., A deep convolutional activation feature for generic visual recognition. UC Berkeley & ICSI, Berkeley, CA, USA.

[32] Van Dyk, D.A. and Meng, X.L., 2001. The art of data augmentation. Journal of Computational and Graphical Statistics, 10(1), pp.1-50.

[33] He K, Zhang X, Ren S, et al., (2016) Deep residual learning for image recognition, Proceedings of the IEEE conference on computer vision and pattern recognition. 770-778.

[34] Huang G., Liu Z., Van Der Maaten L., et al. Densely connected convolutional networks. Proceedings of the IEEE conference on computer vision and pattern recognition. 4700-4708.

[35] Selvaraju R., Cogswell M., Das A., et al. Grad-cam: Visual explanations from deep networks via gradient-based localization. Proceedings of the IEEE international conference on computer vision. 618-626.