High-Order Residual Convolutional Neural Network for Robust Crop Disease Recognition

Weihui Zeng*
Institute of Intelligent Machines,
Chinese Academy of Sciences,
University of Science and Technology of China
P.O. Box 230026, China
whzeng@iim.ac.cn

Miao Li
Institute of Intelligent Machines,
Chinese Academy of Sciences
P.O. Box 230031
China
mli@iim.ac.cn

Jian Zhang
Institute of Intelligent Machines,
Chinese Academy of Sciences
P.O. Box 230031
China
jzhang@iim.ac.cn

Lei Chen
Institute of Intelligent Machines,
Chinese Academy of Sciences
P.O. Box 230026
China
chenlei@iim.ac.cn

Sisi Fang
Institute of Intelligent Machines,
Chinese Academy of Sciences
P.O. Box 230026
China
fss135721@163.com

Jingxian Wang
Institute of Intelligent Machines,
Chinese Academy of Sciences
P.O. Box 230026
China
wjx2016@mail.ustc.edu.cn

ABSTRACT

Fast[1] and robust recognition of crop diseases is the basis for crop disease prevention and control. It is also an important guarantee for crop yield and quality. Most crop disease recognition methods focus on improving the recognition accuracy on public datasets, but ignoring the anti-interference ability of the methods, which result in poor recognition accuracy when the real scene is applied. In this paper, we propose a high-order residual convolutional neural network (HOResNet) for accurate and robust recognizing crop diseases. Our HOResNet is capable of exploiting low-level features with object details and high-level features with abstract representation simultaneously in order to improve the anti-interference ability. Furthermore, in order to better verify the anti-interference ability of our approach, we introduce a new dataset, which contains 9,214 images of six diseases of Rice and Cucumber. This dataset is collected in the natural environment. The images in the dataset have different sizes, shooting angles, poses, backgrounds and illuminations. Extensive experimental results demonstrate that our approach achieves the highest accuracy on the datasets tested. In addition, when the input images are added to different levels of noise interference, our approach still obtains higher recognition accuracy than other methods.

CSAE '18, October 22–24, 2018, Hohhot, China

ACM ISBN 978-1-4503-6512-3/18/10...$15.00
https://doi.org/10.1145/3207677.3277952

CCS CONCEPTS

• **Computing methodologies → Artificial intelligence**; *Computer vision*; Computer vision representations; Image representations

KEYWORDS

High-order residual convolutional neural network, crop disease recognition, robustness

1 INTRODUCTION

Fast and accurate recognition of crop diseases is the basis for crop disease prevention and control, and is also an important guarantee for crop yield and quality. The diagnosis of crop diseases in traditional agriculture mainly depends on naked eyes. However, because of the lack of professional knowledge of farming and the lack of opportunity to stay with farmers with agricultural expertise, the best time to prevent and control crop disease is easy to be missed. In recent years, with the rapid development of image processing, pattern recognition, and computer vision in various fields, using computer technology to automatically recognize and diagnose crop diseases provide a feasible solution to the problems mentioned above [1-12].

When crops are infected with diseases, the leaves typically appear in a certain form of disease. These disease forms are diverse and complex in size, color, texture, and venation, as shown in Fig. 1. Therefore, crop diseases can be automatically recognized and diagnosed by processing leaf disease images. Current methods devote huge efforts on improving recognition accuracy on the public datasets, while ignoring the important anti-interference ability of the methods [1-4]. When such approaches are applied to real scenes, they may fail to deal with those disease leaves with various noise disturbances. However, improving the anti- interference ability of recognition algorithm is the key to the practical application.

In this paper, we propose a high-order residual convolutional neural network (HOResNet) for accurate and robust recognizing crop diseases. In summary, the main contributions of this work are:

- We propose a HOResNet deep network and demonstrate its effectiveness on the anti-interference task.
- To better verify the anti-interference ability of our HOResNet, we introduce a new dataset, which is collected in the natural environment and contains a variety of complex scene changes.
- We provide extensive experiments to evaluate the effectiveness of our approach. Our approach achieves the highest accuracy on the datasets tested, even after adding noises to the input images.

Figure 1: Visual examples of our introduced new dataset: *AES-CD9214*, which contains various natural scenes, such as different object sizes, shooting angles, poses, and backgrounds.

2 RELATED WORKS

In this section, we review previous work related to this work. Numerous work of identifying crop diseases have been proposed in the last decade [4-15]. According to the basic ideas, we can classify them into two categories: approach based on handcrafted representations and approach based on deep representation.

2.1 Handcrafted Representation based Approaches

Approach based on handcrafted representations typically extract handcrafted representations from crop disease images, and then select a proper classifier to identify classes [16-20]. Omrani *et al.* [1] propose a method to address the recognition of three leaf diseases of apple. They compare the recognition results of Support Vector Machine (SVM) and Artificial Neural Network (ANN) based classifier. Their results show that the SVM classifier obtains higher recognition accuracy than the ANN classifier when employing handcrafted features. Wang *et al.* [2] exploit color features to recognize the leaf disease of Cucumber. Ma *et al.* [3] use color, texture, and shape features to identify the Cucumber Downy Mildew in Greenhouse. Because of the limitations of handcrafted features, the accuracy of these approaches is not satisfactory, especially when dealing with complex scenes with multiple categories.

2.2 Deep Representation based Approaches

Benefit from the successful application of deep neural network in the field of image classification, numerous classic classification models have been proposed, such as LeNet [4], AlexNet [5], GoogleNet [6], VGG [7], ResNet [8], and SENet [9]. For crop disease recognition, Mohanty *et al.* attempt to use GoogLeNet and AlexNet to classify 26 disease types on PlantVillage dataset [11]. Their results show that their approach based on deep representation outperforms approach based on handcrafted representations. Besides, the classification accuracy of GoogLeNet [6] is higher than AlexNet [5]. Durmu *et al.* propose to identify tomato disease based on deep learning. The AlexNet [5] and SqueezeNet [10] models are used in their approach to classify the images of tomato disease in PlantVillage dataset [11]. It shows that the recognition accuracy of the AlexNet model is slightly higher than that of the SqueezeNet model, but the model size and the computation time is also doubled. Their approach uses entire face images at all stages instead of only using local patches. These work focus on improving the recognition accuracy, while ignoring the important anti-interference ability of deep network [12-15]. Albeahdili et al. [21] propose a multi-stage convolutional neural network to extract multi-scale features for improving the robustness of image recognition. Their method is evaluated on MNIST, CIFAR-10, and CIFAR-100 datasets and achieves comparable results. To tackle face images with large variation in head pose, Kowalski et al. [22] propose a deep alignment network to align face robustly. These two work indicate that the robustness of network has attracted much attention from researchers.

3 THE PROPOSED APPROACH

In this section, we introduce our high-order residual convolutional neural network (HOResNet) in details. Firstly, we detail the architecture of our HOResNet, especially the residual block. Then, we introduce the network parameter in each layer and the loss function employed, which will clearly show the simplicity of our HOResNet. Finally, implementation details are introduced.

3.1 Architecture of the Proposed HOResNet

The residual network [8] has achieved great success on Large Scale Visual Recognition Challenge (ILSVRC). In fact, the effect of high-order residual network is not discussed in details in previous work; especially its anti-interference ability in identifying crop diseases is ignored. In this work, we demonstrate that the high- order residual network is able to improve the anti-interference ability of deep network.

We define a residual block, as shown in Fig. 2, in which the sum layer sends the sum of two inputs to the next layer. There are three convolution layers in a residual block. Note that the middle

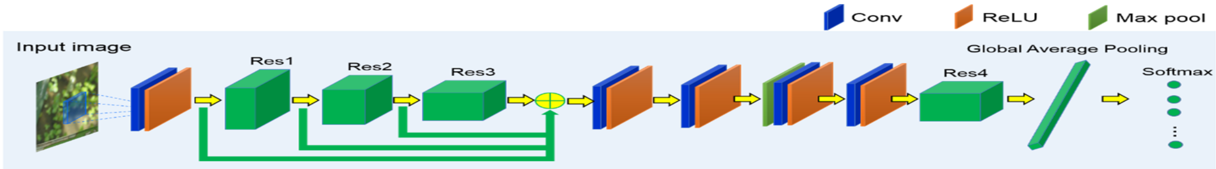

Figure 3: Overall architecture of our high-order residual convolutional neural network (HOResNet) for accurate and robust recognizing crop diseases.

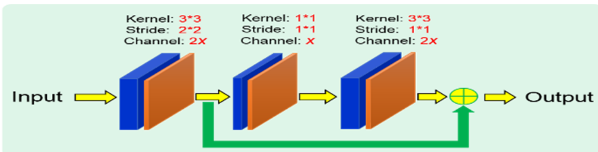

Figure 2: Demonstration of our residual block containing three convolution layers. This residual block is the basic unit of our HOResNet network and is denoted as a green block in Fig. 3.

layer uses 1×1 kernel to change the channels of previous layer. Fig. 3 shows the overall architecture of our HOResNet. In Fig. 3, a green block represents a residual block defined in Fig. 2. Given a crop disease image, our network uses five convolutional layers, four residual blocks, a global average pooling layer, and a softmax layer to directly output the class probability of input image. In Fig. 3, we concatenate outputs of first three residual blocks in order to effectively exploit different layer features, because there is a key observation that this concatenation operation can provide richer features (containing low-level and high level features) for the later network layer and improve the robustness of the network. We call this concatenation operation a high-order residual block (HORB). In the experiment, we will demonstrate the effectiveness of HORB.

3.2 Network Parameter and Loss Function

Our HOResNet is a simple and fast convolutional neural network. Table 1 lists the parameter details of the proposed HOResNet. Note that the parameter details of residual block are shown in Fig. 2. We only use 1 × 1 and 3 × 3 two types of kernel size in order to reduce the number of network parameters, which is helpful for avoiding overfitting. Besides, we use a deeper rather than a wider strategy to design each layer of the channel number. Specifically, we use small channels on each layer. Besides, we use softmax as the objective function, which is formulated as follows: where x_n denotes n^{th} training sample and $y_n \in 1,2,3, \ldots, C$ is the corresponding label. M and C denote the numbers of training samples and classes. $p(.)$ is an indicator function.

$$J(\theta) = -\frac{1}{M}\left[\sum_{j=1}^{M} \sum_{k=1}^{C} p\left(y_j == k\right) \log \frac{e^{\theta_k^T x_i}}{\sum_{p=1}^{C} e^{\theta_p^T x_i}}\right] \quad (1)$$

Table 1: Parameter Details of the Proposed HOResNet

Layer	Output size (h/w/c)	Filter size (h × w)/stride	Pooling size (h × w)/stride
Conv1	256/256/8	3 × 3/1	-
ResBlock1	128/128/16	-	-
ResBlock2	64/64/32	-	-
ResBlock3	32/32/64	-	-
Conv2	32/32/64	3 × 3/1	-
Conv3	32/32/64	3 × 3/1	-
Max pool	16/16/64	-	2 × 2/2
Conv4	16/16/128	3 × 3/1	-
Conv5	16/16/64	1 × 1/1	-
ResBlock4	8/8/192	-	-
GAP	1 × 192	-	-
Softmax	1 × 6	-	-

3.3 Implementation Details

We train our HOResNet network on the train sets from two datasets: PlantVillage [11] and AES-CD9214. It takes around 20 minutes for 100 epochs on a machine with an Nvidia GPU 1080i. Typically, after training 50 epochs, our HOResNet is capable of producing satisfactory accuracy. The loss function of Eq. (1) is optimized using Adam algorithm with an initial learning rate of 2×10^{-3}. The batch size is set to 120. During the testing stage, to demonstrate the anti-interference ability of the proposed HOResNet network, we add different degrees of Gaussian and Salt & Pepper noises to the test image to evaluate the recognition accuracy of our network.

4 EXPERIMENTAL RESULT

4.1 Experimental Setup

Dataset: Our approach is tested on the PlantVillage dataset [11] and our introduced AES-CD9214 dataset. The PlantVillage [11] is a large collection to solve the problem of plant disease diagnosis, as shown in Table 2. It is open to all researchers. This dataset contains both the disease and healthy leaf images of multiple plants. The AES-CD9214 is a challenging dataset, which is collected in the natural environment. The images in the dataset have different sizes, shooting angles, poses, backgrounds and illumination. Table 3 shows each class and number of corresponding images. The number of examples in AES-CD9214

Table 2: The disease name and number of corresponding images in PlantVillage dataset [11]

Disease name	Tomato bacterial spot	Tomato healthy	Tomato late blight	Tomato septoria leaf spot	Tomato spider mites two spotted spider mite	Tomato target spot
Number of Images	2127	1591	1909	1771	1676	1404

Table 3: The disease name and number of corresponding images in our introduced AES-CD9214 dataset

Disease name	Rice sheath blight	Rice blast	Rice flax spot	Cucumber powdery mildew	Cucumber downy mildew	Cucumber target spot
Number of Images	3559	2741	795	763	780	576

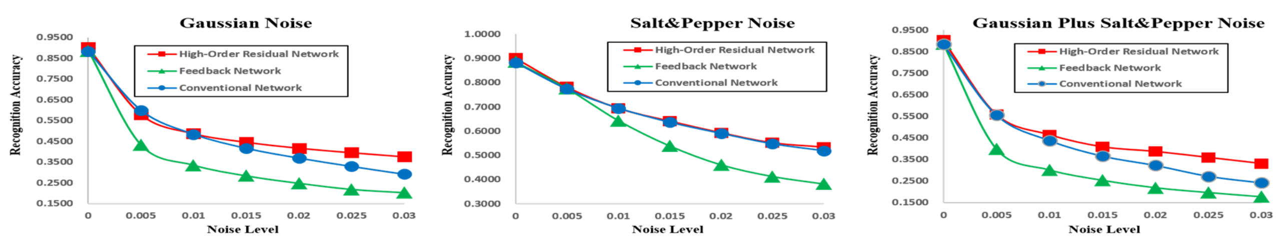

Figure 4: Comparison results of noise level vs recognition accuracy on challenging AES-CD9214 dataset. Our high-order residual network obviously outperforms other methods by a large margin.

Table 4: Comparison of Recognition Accuracy on PlantVillage Dataset

Approaches	Recognition Accuracy
Conventional Network	0.8859
Feedback Network	0.9093
Our High-Order Residual Network	**0.9179**

Table 5: Comparison of Recognition Accuracy on AES-CD9214 Dataset

Approaches	Recognition Accuracy
Conventional Network	0.8829
Feedback Network	0.8856
Our High-Order Residual Network	**0.9014**

dataset is unbalance, which is challenging to the recognition methods. The 20% images of the PlantVillage [11] and AES-CD9214 datasets are used as test sets, 80% as training sets.

Compared methods: We compare with the conventional CNN without high-order residual block and the feedback CNN. The feedback CNN is similar to recurrent neural networks and has the ability to resist interference. Therefore, there are two methods to compare with our approach.

4.2 Comparison of Recognition Accuracy

Table 4 and Table 5 report the comparison of recognition accuracy on PlantVillage [11] and AES-CD9214 datasets, respectively. The results show that our high-order residual network outperforms other methods on both two datasets, which well demonstrates the effectiveness of high-order residual block for obtaining high recognition accuracy. Note that the recognition accuracy is not very high in Table 4 and Table 5, because we adopt a relatively simple network (see Fig. 3 and Table 1). We believe that by increasing the high-order residual blocks and the training data, the recognition accuracy can be further improved.

4.3 Comparison of Recognition Robustness

To evaluate the anti-interference ability of the proposed HOResNet network, we add different degrees of Gaussian and Salt & Pepper noises to the test images to evaluate the recognition accuracy of our network.

Results on challenging AES-CD9214 dataset: Fig. 4 demonstrates the experimental results. The horizontal axis and vertical axis represent the noise intensity added and the recognition accuracy, respectively. Obviously, with the increase of noise intensity, the accuracy of all recognition methods decreases. Besides, the accuracy of our high-order residual network is higher than that of other methods in terms of different noise types and noise levels. The performance advantage of our approach is especially obvious when the noise intensity is large. These results well demonstrate the anti-interference ability of our approach.

Results on PlantVillage dataset [11]: Table 6-Table 8 show the experimental results of noise level vs. recognition accuracy on PlantVillage dataset. When adding Gaussian noise and hybrid noise (both Gaussian and Salt& Pepper noises), our approach achieves better performance than other methods. When adding Salt& Pepper noise, the performance of Feedback Network method is better, but our approach achieves the similar accuracy. Overall, our approach outperforms other methods in most cases.

Table 6: Comparison of Recognition Accuracy on PlantVillage Dataset when Adding Gaussian Noise

Approaches/Noise Level	0.005	0.01	0.015	0.02
Conventional Network	0.5661	0.3709	0.2936	0.2778
Feedback Network	**0.6520**	0.3981	0.3079	0.2797
Our High-order Residual Network	0.6401	**0.5103**	**0.3613**	**0.2821**

Table 7: Comparison of Recognition Accuracy on PlantVillage Dataset when Adding Salt& Pepper Noise

Approaches/Noise Level	0.005	0.01	0.015	0.02
Conventional Network	0.7566	0.6010	0.5241	0.4697
Feedback Network	**0.8907**	**0.8301**	**0.7017**	0.5809
Our High-order Residual Network	0.8826	0.7642	0.6477	**0.6038**

Table 8: Comparison of Recognition Accuracy on PlantVillage Dataset when Adding Gaussian and Salt& Pepper Noise

Approaches/Noise Level	0.005	0.01	0.015	0.02
Conventional Network	0.3208	0.2897	0.2811	**0.3208**
Feedback Network	0.5480	0.3399	**0.2835**	0.2530
Our High-order Residual Network	**0.5947**	**0.4010**	0.2797	0.2363

5 CONCLUSIONS

We have introduced a high-order residual convolutional neural network (HOResNet) for accurate and robust recognizing crop diseases. Besides, to better verify the anti-interference ability of our approach, we introduce a new dataset: *AES-CD9214,* which can contribute to the study of whole community. Extensive experiments are conducted to verify the effectiveness of the proposed approach. Our approach outperforms other methods on the datasets tested. In addition, when the input image is added to different levels of noise interference, our approach still obtains higher recognition accuracy than other methods. In the future, we consider combining the high-order residual block and feedback block to obtain more accurate and robust recognition model.

ACKNOWLEDGMENTS

This work was partially supported by the 13th Five-year Informatization Plan of Chinese Academy of Sciences, Grant No. XXH13505-03-104.

REFERENCES

[1] Omrani E, Khoshnevisan B, and Shamshirband S, et al. 2014. Potential of radial basis function-based support vector regression for apple disease detection. *Measurement*, 55(9), 512-519.

[2] Wang ZhiBin, and Wang Kaiyi,et al. 2017. Recognition Method of Cucumber Leaf Diseases with Dynamic Ensemble Learning. *Transactions of The Chinese Society of Agricultural Machinery*, 48(9), 46-52.

[3] David A. Anisi. 2003. *Optimal Motion Control of a Ground Vehicle.* Master's Thesis. Royal Institute of Technology.

[4] Yann LeCun, L´eon Bottou, and Yoshua Bengio, et al. 1998. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11), 2278–2324.

[5] Krizhevsky A, Sutskever I, and Hinton G E, et al. 2012. ImageNet Classification with Deep Convolutional Neural Networks. *Neural Information Processing Systems*, 1097-1105.

[6] C Szegedy,Wliu, and Y Jia, et al. & Rabinovich, A. 2015. Going deeper with convolutions. *Proceedings of the IEEE conference on computer vision and pattern recognition*, 1–9.

[7] Karen Simonyan, and Andrew Zisserman. 2014. Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556, 2014.

[8] Kaiming He, Xiangyu Zhang, and Shaoqing Ren, et al. 2016. Deep residual learning for image recognition. *Proceedings of the IEEE conference on computer vision and pattern recognition*, 770–778.

[9] Jie Hu, Li Shen, and Gang Sun. 2017. Squeeze-and-excitation networks. arXiv preprint arXiv: 1709.01507.

[10] Iandola F N, Han S, and Moskewicz M W, et al. 2016. SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and <0.5MB model size. arXiv: Computer Vision and Pattern Recognition.

[11] Hughes, D. P., and Salathe, M. 2015. An open access repository of images on plant health to enable the development of mobile disease diagnostics through machine learning and crowdsourcing. *Computer Science.*

[12] Liu Z, Zhu L, and Zhang X, et al. 2015. Hybrid Deep Learning for Plant Leaves Classification. *International Conference on Intelligent Computing*, 115-123.

[13] Lee S H, Chan C S, and Wilkin P, et al. 2015. Deep-Plant: Plant Identification with convolutional neural networks. *International Conference on Image Processing*, 452-456.

[14] Lee S H, Chan C S, and Mayo S J, et al. 2017. How deep learning extracts and learns leaf features for plant classification. Pattern Recognition, 1-13.

[15] Grinblat G L, Uzal L C, and Larese M G, et al. 2016. Deep learning for plant identification using vein morphological patterns. *Computers & Electronics in Agriculture*, 127, 418-424.

[16] Kalyoncu C, and Önsen Toygar. 2015. Geometric leaf classification. *Computer Vision & Image Understanding*, 133, 102-109.

[17] Neto J C, Meyer G E, and Jones D D, et al. 2006. Plant species identification using Elliptic Fourier leaf shape analysis. *Computers & Electronics in Agriculture*, 50(2), 121-134.

[18] Naresh Y G, and Nagendraswamy H S. 2016. Classification of medicinal plants: An approach using modified LBP with symbolic representation. *Neurocomputing*, 173, 1789-1797.

[19] Charters J, Wang Z, and Chi Z, et al. 2014. EAGLE: A novel descriptor for identifying plant species using leaf lamina vascular features. *International Conference on Multimedia And Expo*, 1-6.

[20] Kumar N, Belhumeur P N, and Biswas A, et al. 2012. Leafsnap: a computer vision system for automatic plant species identification. *European Conference on Computer Vision*, 502-516.

[21] Albeahdili H M, Alwzwazy H A, and Islam N E, et al. 2015. Robust Convolutional Neural Networks for Image Recognition. *International Journal of Advanced Computer Science and Applications*, 6(11).

[22] Kowalski M, Naruniec J, and Trzcinski T, et al. 2017. Deep Alignment Network: A Convolutional Neural Network for Robust Face Alignment. Computer Vision and Pattern Recognition, 2034-2043.