Research on Maize Disease Identification Methods in Complex Environments Based on Cascade Networks and Two-stage Transfer Learning

Hongxin Liu
Qingdao Agricultural University
Haichen Lv
Qingdao Agricultural University
Jiajun Li
Qingdao Agricultural University
Yongshuo Liu
Qingdao Agricultural University
Limiao Deng (denglm68@163.com)
Qingdao Agricultural University

Article

Keywords: Maize leaf disease, Cascade networks, Two-stage transfer learning, Attention mechanism, Complex background

Posted Date: June 22nd, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1732191/v1

Abstract

BACKGROUND: Achieving accurate and reliable maize disease identification in complex environments is a huge challenge. This is because disease images obtained from natural environments are often in complex contexts that may contain elements similar to disease characteristics or symptoms.

RESULTS: Based on cascade network and two-stage transformation learning, the new algorithm is proposed in this paper and applied the improved algorithm to the task of identification and classification of four maize leaf diseases in a complex environment. The proposed algorithm has a cascade structure, which consists of a Faster R-CNN object detector, and a convolutional neural network (CNN) object classifier, named CENet (Complex Environment Network). The Faster R-CNN object detector with an attention mechanism was trained to detect maize leaves from the image, and the CNN object classifier further classified the leaf disease detected in the first stage into four categories: Cercospora leaf spot, Common rust, Northern Leaf Blight, and Healthy, which allowed image features to be extracted more efficiently. The subsequent use of a two-stage transfer learning strategy to train CENet models of disease images in complex contexts allows for faster training of the models while ensuring accuracy. Experimental results showed that the proposed method achieved a test accuracy of 99.34% for four kinds of maize leaf diseases, outperformed some prevalent CNN models, such as VGG16, Inception, and ResNet101, and has a more significant improvement in training speed.

CONCLUSION: The experimental results show that the method proposed in this paper can identify and classify different varieties of maize leaf diseases, proving the feasibility of a convolutional neural network in variety identification and classification. The model proposed in this experiment has a positive significance for exploring other Crop variety identification and classification under complex backgrounds.

1. Introduction

Maize is a major crop in China, with the largest planting area and yield, and also plays an important role in light industry, animal husbandry, and the national economy. Maize diseases not only reduce the maize yield but also affect the development of related industries and economies. At present, the manual method is the main method to identify maize diseases in China. The labor process of using manpower to identify maize diseases is not only inefficient, but also easy to be disturbed by subjective factors such as fatigue and emotion, and can only be identified when the obvious symptoms appear (Chen et al., 2020). Therefore, how recognizing the diseases of maize leaves quickly and accurately and taking appropriate control measures is of great significance to ensure maize production.

The research on crop image disease recognition abroad began in the 1980s. Researchers have extensively used a variety of traditional machine learning methods to study the image recognition technology of agricultural diseases, including the support vector machine classifier method(Semary et al., 2015), PNN method(Shi et al., 2015), K-nearest neighbor classification method(S. W. Zhang et al., 2015), BP network method(Wang et al., 2012), and so on, which has played a positive role in promoting the application of information technology in agricultural disease image recognition research. However, the traditional machine learning method has some shortcomings, such as limited learning and expression ability, manual extraction of features, and unsuitable for processing large amounts of data.

The deep learning method can effectively solve the problem of big data learning and modeling. In recent years, researchers have carried out a lot of research work in agricultural disease image recognition based on deep learning. Hammad Saleem et al. (2020) proposed an image-based deep learning meta-structure model to identify plant diseases. Long et al. (2018) proposed a recognition method based on a convolutional neural network and transfer learning for Camellia oleifera disease image recognition, and the average recognition accuracy reached 96.53%. Based on the characteristics of maize foliar diseases, Zhao et al. (2009) applied the threshold method, area marker method, and Freeman link code method to diagnose five major diseases of maize foliage with an accuracy of more than 80% Y Liu et al. (2018). applied the Triplet loss double convolution neural network structure to study the features of corn images, and then used the SIFT algorithm to extract texture features, and the classification accuracy was above 90%. Zeng and Li (2020)proposed the Self-Attention Convolutional Neural Network(SACNN)to identify crop diseases, and extensive experimental results showed that the recognition accuracy of SACNN on AES-CD9214 and MK-D2 was 95.33% and 98.0%, respectively. Compared with the traditional machine learning methods, a deep learning framework can automatically learn the features contained in the image data. When the data set reaches a certain size, it can achieve better accuracy and robustness in the agricultural disease image recognition task. However, the application of deep learning in agricultural disease image recognition still has some problems, such as large training data set, over-reliance on data annotation, limited generalization ability of the model, and high requirements on hardware computing power.

The deep transfer learning method can use the learned knowledge in the field of big data to assist in the building data model in the field of smaller goals, directly reducing the size of the target domain modeling for data requirements, which includes the research field of agricultural disease image recognition. Researchers have carried out some related research work(Fang et al., 2017; Yuan et al., 2018; K. Zhang et al., 2019), using some existing large image datasets to assist in establishing the image recognition model of target disease with small sample data, and achieved certain results.

Although deep learning models for agricultural disease recognition are becoming more and more mature and some research results have been achieved, however, most of the research is based on disease images collected in the laboratory environment, and few studies focused on disease recognition in the actual farmland environment. When these methods are applied to the actual farmland environment, the detection and recognition results are easily affected by the complex environment and the image shooting environment. The recognition accuracy will be greatly reduced, and the applicability is poor with limitations. Compared with the existing methods, the main contributions of this paper are as follows:

- We proposed an effective maize disease identification method in complex environments based on cascade networks and two-stage transfer learning. The cascade networks were composed of a Faster R-CNN leaf detector (denoted as LS-RCNN) and a CNN disease classifier (denoted as CENet).
- Two-stage transfer learning strategy was proposed to successfully perform transfer learning to train the disease classifier CENet. The transfer of the pre-trained model allowed the model under training to converge faster and to recognize image features with higher accuracy.
- We constructed a maize disease data set containing 7144 images, including 3563 images in the natural environment with a more complex background, and 3581 images in the laboratory environment

- Fifteen augmentation methods were performed on the existing image data (especially the natural environment) for data enhancement to achieve the purpose of increasing data volume, enriching data diversity, and improving the generalization ability of the model. Also, we investigated the effects of different numbers of amplified images and different amplification methods on the recognition performance.

2. Materials And Methods

2.1 Materials

2.1.1 Data collection

In this experiment, corresponding datasets were created for different types of maize leaves. Images in the lab dataset were obtained from Plant Village(Hughes & Salathé, n.d.2015), an open-access repository. Most of the images in the natural environment dataset were acquired through field photography in Qingdao. Due to the limited variety of maize leaves available from field photography, we downloaded some open-source data of the natural environment as a supplement. All experimental protocols complied with all relevant guidelines and regulations.

Table 1
Size of the datase

datasets	Category	Training	Validation	Testing	Total
Laboratory	Cercospora leaf spot	357	104	42	503
	Common rust	777	239	79	1095
	Northern Leaf Blight	668	193	79	940
	Healthy	799	171	73	1043
Natural environment	Cercospora leaf spot	174	47	34	255
	Common rust	1084	293	187	1564
	Northern Leaf Blight	369	103	70	542
	Healthy	809	231	116	1156

The four categories of corn leaf types were Cercospora leaf spot, common rust, Northeast leaf blight, and Healthy. Table 1 shows the number of images collected for each category, the number for training, validation, and testing, and their total number. Figure 1 shows some sample images of the natural environment dataset and the laboratory dataset, and their background variability.

2.1.2 Data augmentation

We performed data enhancement on the existing image data (especially the natural environment) for data enhancement to achieve the purpose of increasing data volume, enriching data diversity, improving the

generalization ability of the model, expanding the sample space, and reducing the influence of unbalanced data.

We used 15 data enhancement methods as shown in Fig. 2. These methods come from the OpenCV-based implementation of the Albumentations library (Buslaev et al., 2020), a fast and flexible open-source library for image enhancement that provides many various image conversion operations. In most image conversion operations, Albumentations enhancement is faster than other commonly used image enhancement tools.

2.2 The proposed method

The proposed disease method had a cascade structure which consisted of a Faster R-CNN maize leaf detector (LS-RCNN) and a CNN leaf disease classifier (CENet), as shown in Fig. 3. First, disease images in the natural environment were input to the LS-RCNN to detect and separate the maize leaf from the complex background. Then the separated maize leaf was input into the trained CENet model to perform disease identification.

2.2.1 Leaf segmentation model based on Faster R-CNN (LS-RCNN)

To reduce the influence of complex background on recognition performance, we constructed the LS-RCNN model based on Faster R-CNN (Ren et al. 2017) to extract the key regions of the maize leaf image from the background before they were fed into CENet model for training and recognition. Figure 4 shows the model structure of LS-RCNN. Structurally, LS-RCNN had integrated feature extraction, proposal extraction, bounding box regression, and classification all into one network, which made its comprehensive performance improved, especially in the detection speed.

First, the LS-RCNN model used a basic set of conv + relu + pooling layers to extract feature maps of maize images, which were shared with the subsequent RPN and fully-connected layers. Then, the RPN network generated region proposals for the maize leaves, which used softmax to determine whether the anchors were positive or negative, and then used the bounding box regression to correct the anchors, eliminated those that were too small and out of bounds, and obtained the exact proposals for the maize leaf region. Next, the Roi Pooling layer collected the input feature maps and proposals and extracted the proposal feature maps after synthesizing the information, which was sent to the subsequent fully connected layer to determine the target class. At last, the category of the proposal was calculated by using the proposal feature maps and the final position of the detection box was obtained by bounding box regression to generate a detection box for the maize leaves. Thus, a new image was generated, which contained the detected maize leaf from each detection box.

2.2.2 CENet model based on two-stage transfer learning

To further solve the disease recognition problem in complex backgrounds, a two-stage transfer learning strategy was proposed to train an effective CNN deep learning model for disease images in complex backgrounds. Figure 5 shows the architecture and the training process of the CENet model for complex environments.

Since Alexnet (Krizhevsky et al., n.d.), the CNN structure has been continuously deepened. VGG(Simonyan & Zisserman, 2015) and GoogLeNet (Szegedy et al., 2015) have 19 and 22 convolution layers respectively. With

the increase of network depth, the existence of gradient disappearance problems makes network training more difficult, and the convergence effect is poor, so ResNet is introduced. ResNet proposed by He et al. (2016) can effectively solve the deep network degradation problem. So, a pre-trained ResNet50 model was used for transfer learning in this paper.

The ResNet50 model was first pre-trained on the ImageNet (Feifei Li et al.) dataset, and then the pre-trained model was trained by parameter transfer on the maize disease dataset obtained in the laboratory, which was the first stage of transfer learning. In the first-stage transfer learning, we replaced the average-pooling-based GlobalPool layer with a max-pooling layer and replaced the fully connected (FC) layer and classification layer with a new FC layer and classification layer. The new classification layer had four output nodes instead of 1000. Then the trained model was further transferred to the domain of natural images, which was the second stage of transfer learning. In the second-stage transfer learning, we replaced the FC layer and classification layer with a new FC layer and classification layer. Specifically, the region of interest was extracted by LS-RCNN to obtain the background simplified natural environment dataset and then was input into the ResNet50 model trained in the previous stage as training samples. In this way, the training process was completed and a well-trained CENet was obtained.

3. Results

3.1 Experimental setup

For training and testing, each image in the dataset is kept at its original size to fit the model. Hardware environment was CPU: AMD Ryzen 7 Mobile 4800H; GPU: NVIDIA GeForce GTX 1660 Ti; Software environment was OS: Windows 10 64 bit; Programming language: Python 3.7.6; Deep learning framework: Pytorch 1.7.0.

After several many trials, we found suitable values for the model parameters. The parameters used in LS-RCNN are shown in Table 2, And the values of the model training parameters are shown in Table 3.

Table 2
Parameters used in LS- RCN

Parameter	Weight decay	Learning rate	Momentum	Gamma	Batch size	Max iters	Step size	Display
values	0.0005	0.001	0.9	0.1	256	15000	9000	10

Table 3
Training parameters of ResNet50 mode

Parameter	Learning rate	Momentum	Gamma	Batch size	Input size	Step size	Num epochs
values	0.001	0.9	0.1	16	224	4	50

3.2 Recognition performance comparison of different convolutional networks

To evaluate the recognition performance of different CNN networks, we compared ReNet50 with some prevalent CNN networks, including VGG16, Inception, AlexNet, ResNet18, Wide_ResNet50, and ResNet101, as illustrated in Fig. 6. After 50 epochs of training on the laboratory (public) dataset, the variation of loss rate of each network model with the number of training rounds is shown in Fig. 6(a), and the variation of recognition accuracy is shown in Fig. 6(b).

Figure 6 shows that all the networks fit quickly in the first 2 epochs and the accuracy rate increases rapidly. Then the loss rate decreases slowly and the accuracy rate increases slowly in about 3 ~ 20 epochs, and then the loss rate tends to be stable and the accuracy rate also tends to be stable after 21 epochs, and the models begin to converge. Among the seven networks, Resnet50, wide_Resnet50_2, and Restnet101 have better recognition, excellent performance, and rapid convergence, with the highest accuracy of 98.52%, 98.66%, and 99.19%, respectively. The following are Resnet18, Alexnet, and Inception with the highest accuracy of 98.25%, 98.25%, and 98.39%, respectively. And the highest accuracy of vgg16 is only 96.37%.

The average training accuracy and consumed time after 50 epochs of training are shown in Fig. 7, in which the accuracy of each model is ranked in ascending order and the consumed time is also shown.

It can be found from Fig. 7 that the models with higher accuracy (e.g., Resnet50, Wide_Resnet50_2, Restnet101) usually take more time. Conversely, models with short time consumption do not have high recognition rates.

Therefore, making a tradeoff between the recognition accuracy and time spent during training, Resnet50 network demonstrated the best performance and was used for further optimization on datasets with complex backgrounds.

3.3 Comparison between two-stage transfer learning and traditional transfer learning

Figure 8 shows the comparison of two-stage transfer learning with traditional transfer learning. Figure 8 (a) is the loss curve, and Fig. 8 (b) is the curve of recognition accuracy.

Figure 8 shows that both methods fit quickly in the first 4 epochs. Then the accuracy increases rapidly, and the loss rate slowly decreases and tends to be smooth in the subsequent epochs. Finally, the accuracy rate slowly increases and tends to be smooth, and the model converges. The accuracy of the two-stage transfer learning technique is higher, with the highest accuracy of 97.22% and the lowest loss rate of 0.1546; the accuracy of traditional transfer learning is relatively lower, with the highest accuracy of 93.06% and the lowest loss rate of 0.2501. The recognition effect of two-stage transfer learning is significantly better than that of traditional transfer learning.

3.4 Recognition effect of different numbers of amplified images

We conducted offline supervised data enhancement on the data set in the natural environment, and the accuracy change with the size of the amplified dataset is shown in Fig. 9.

Experimental results show that on the whole, the accuracy increases with the increase of the size of data sets, which indicates that the relationship between data size and accuracy is proportional, and the larger the data size, the higher the accuracy of the model is. However, when the data is amplified to 1 and 8 times, the accuracy does not increase, which indicates that data augmentation methods do not always have a positive impact on the accuracy. For example, some data augmentation methods such as CoarseDropout and RandomFog will reduce the accuracy of the model.

3.5 Performance evaluation of LS-RCNN model

Comparing the laboratory dataset with the natural dataset, we found that the background of the laboratory data was single, however, the background of the data in the natural environment was more complex and had interference features. Therefore, we used the LS-RCNN model to perform semi-supervised learning on the leaf as the region of interest, so that the natural data can achieve the purpose of separating the leaves from the background and reducing the interference factors of the complex background, as illustrated in Fig. 10.

To evaluate the effect of leaf segmentation model LS-RCNN on the recognition performance, we performed experiments on two datasets: the original dataset with complex background and the dataset with complex background removed by LS-RCNN.100 epochs of training was performed on both datasets using the ResNet50 network, and the training loss curve is shown in Fig. 11(a), and corresponding accuracy curve is shown in Fig. 11(b).

Experimental results show that the two datasets fit quickly in the first 9 epochs and the accuracy increases rapidly; the loss rate decreases slowly and the accuracy increases slowly in about 10 to 26 epochs; after 27 epochs the loss rate leveled off and the accuracy leveled off, and the model converged. The accuracy of the dataset with complex background removed using LS-RCNN is higher, with the highest accuracy of 100% and the lowest loss rate of 0.06297; the accuracy of the original dataset is relatively lower, with the highest accuracy of 94.44% and the lowest loss rate of 0.2285.

3.6 Performance evaluation of our method

To evaluate the performance of our proposed method, we performed some experiments on laboratory datasets and natural datasets and compared our method with some prevalent CNN models and in different conditions. Table 4 shows the comparison of our model with some related CNN models. The training accuracy, validation accuracy, and test accuracy are all the highest accuracy rates among 50 num epochs.

Overall, our proposed model (which uses a cascade network and two-stage transfer learning) has a high accuracy rate of 99.03% among 8 different models; the training time is relatively short among models with the same accuracy, including 72.07% shorter than the time taken for Wide_Resnet50_2. At the model level, the model using two-stage transfer learning takes less time, and the model using LS-RCNN has a good improvement inaccuracy. The experimental results show that the method we proposed has good applicability for the recognition of leaves in complex environments and provides an effective method for object detection in complex environments. Figure 12 shows the confusion matrix of our model recognition test set (all-natural environment images).

Table 4
Comparison of our method with some related CNN model

Model	Training time	Acc_{train}	Acc_{val}	Acc_{test}
VGG16	45m 30s	96.37%	96.86%	98.63%
Inception	70m 1s	98.35%	98.29%	96.72%
Alexnet	20m 60s	97.68%	97.75%	98.29%
Resnet18	28m 45s	98.39%	98.36%	96.95%
Resnet50	61m 11s	99.08%	98.43%	97.49%
Wide_Resnet50_2	103m 42s	99.20%	98.64%	97.95%
Restnet101	89m 16s	99.04%	98.43%	96.72%
Our Model	28m 58s	99.43%	98.67%	99.03%

4. Discussion

4.1 Solutions to low accuracy in complex environments

4.1.1 Two-stage transfer learning

The term transfer was first cited by Lorien Pratt in the field of machine learning. Pratt et al. (1993) proposed a new algorithm called Discriminability-Based Transfer (DBT), where the target network initialized by DBT learns significantly faster than the network initialized randomly. Chuong B Do and Andrew Ng (2006) explored the application of transfer learning in text classification. B Schölkopf et al. (2007) proposed a method for learning a low-dimensional representation that is shared across a set of multiple related tasks. The proposed approach greatly improves the performance compared to learning each task independently. The application of transfer learning to Bayesian networks is discussed by Niculescu-Mizil and Caruana (2012) through transfer learning, the trained network model parameters are saved and reapplied in the new task, which makes the feature parameters of the original network model effectively used and increases the portability.

For the problem of low accuracy in natural scenes that occurs in the experiment, we proposed a two-stage transfer learning method to attempt to solve the problem of recognition accuracy caused by insufficient features of natural data and prevent overfitting problems.

We used the ResNet50 network as the base CNN architecture, set the first sample parameters as trained parameters on the ImageNet dataset, set the second sample parameters as trained parameters on a self-constructed natural environment dataset with a complex background, and used the two-stage transfer learning method to train the maize leaf disease image dataset. Experimental results demonstrated that the accuracy of two-stage transfer learning improved by 4.16% over traditional transfer learning, and had good performance in recognizing images with complex backgrounds in natural environments, which is an effective method to solve the low recognition rate of complex backgrounds.

4.1.2 Image segmentation based on Faster R-CNN

Faster R-CNN was used in the LS-RCNN model to separate maize leave from complex backgrounds for several main reasons: in recent years, Faster R-CNN has been widely used for image target recognition in agriculture(Dyrmann et al., 2016) because of its ability to automatically learn image features, and the Faster R-CNN is one of the most mature target detection algorithms; Faster R-CNN performs well on multiple datasets and is easy to transfer, and changes to the target classes in the dataset can be made to improve the detection speed. Faster R-CNN can integrate feature extraction, candidate region extraction, border regression, and classification into a single network, and use shared convolutional layers to improve detection speed.

For disease recognition in complex background, Li et al. (2020) improved Faster R-CNN for leaf disease detection in bitter melon in the field. Zeng and Li (2020) proposed a Self-Attention Convolutional Neural Network (SACNN), which extracts effective features of crop disease spots to identify crop diseases. Zhou et al. (2021) proposed a vegetable disease recognition model for complex backgrounds based on region proposal and progressive learning (PRP-Net). This model achieves an average recognition accuracy of 98.26%, which is 4.46 percentage points higher than that of the original region proposal network framework. So, we attempted to construct an LS-RCNN model based on Faster R-CNN to detect the regions of interest in natural images. LS-RCNN proved very effective for separating corn leaves from the complex environment and was very helpful to solve the problem of corn leaf disease identification in a complex environment.

4.2 Limited number of images in complex environments

We found that recognition accuracy would be greatly affected by too few images in complex natural environments during two-stage transfer learning. To prevent possible overfitting problems with the limited dataset, we expanded the natural environment dataset in the following two ways: one was to download as many pictures as possible from the Internet, and the other was to use the data augmentation method.

Data enhancement is a common technique to increase the size and diversity of labeled training sets by using input transformations that retain the corresponding output labels. In computer vision, image enhancement has become a common routine technique to combat over-adaptation in deep learning models and is widely used to improve performance. While most deep learning frameworks implemented basic image transformations(Perez & Wang, 2017; Taylor & Nitschke, 2017), which were typically limited to certain variations of flipping, rotating, scaling, and cropping.

In addition, the speed of image processing in existing image enhancement libraries varies. In this paper, we used 15 data enhancement methods and amplified the dataset in complex environments by different orders of magnitude. Experimental results showed that, on the whole, data augmentation improved the recognition performance of the model, and solved the problem of limited data sets to a certain extent, as demonstrated in previous research(Mikołajczyk and Grochowski 2018). However, not all data enhancement methods are effective. Which method is more effective, or how much-amplified data is appropriate remains to be studied in the future.

5. Conclusion

In this paper, we proposed a new method based on cascade networks and two-stage transfer learning to recognize maize leaf disease in the natural environment. Using our proposed method, the overall recognition rate of the proposed model obtained by training on four types of maize leaves reached 99.03%, which was higher than most human experts and traditional neural network models. The proposed method not only eliminated unnecessary feature extraction processes but also improved the accuracy of disease recognition in complex backgrounds. The recognition accuracy of our method in complex backgrounds was 0.99 percentage points higher than that of the original ResNet50 (98.35%) and outperformed the prevalent deep learning methods. The proposed method had provided a new effective method for disease recognition of maize left in complex environments.

In the future, we will conduct research in two directions. First, we will try to integrate multiple region attention to model more complex fine-grained categories. Second, we will try to use a technique that is designed to be used to get more features by removing the complex background rather than focusing on the local area.

Declarations

Acknowledgments

This work was financially supported by the National Natural Science Foundation of China (31872849, 32073029), a key project of the Shandong Provincial Natural Science Foundation (ZR2020KC027), the Qingchuang Talents Induction program of Shandong Higher Education Institution, and Ph.D. Innovation Foundation of Qingdao Agricultural University (1119036).

Data availability

The data that support the plots within this paper and other findings of this study are available from the corresponding author upon reasonable request.

References

1. B Schölkopf, J Platt, & T Hofmann. (n.d.). Multi-Task Feature Learning. Retrieved October 29, 2021, from https://xueshu.baidu.com/usercenter/paper/show?paperid=4efd1b30de34a5aca7d3c6fce3a5cfc5&site=xueshu_se
2. Buslaev, A., Iglovikov, V. I., Khvedchenya, E., Parinov, A., Druzhinin, M., & Kalinin, A. A. (2020). Albumentations: Fast and flexible image augmentations. Information (Switzerland), 11(2). https://doi.org/10.3390/INFO11020125
3. Chen, X., Zhou, G., Chen, A., Yi, J., Zhang, W., & Hu, Y. (2020). Identification of tomato leaf diseases based on combination of ABCK-BWTR and B-ARNet. Computers and Electronics in Agriculture, 178. https://doi.org/10.1016/J.COMPAG.2020.105730
4. Do, C. B., & Ng, A. Y. (n.d.). Transfer learning for text classification.
5. Dyrmann, M., Karstoft, H., & Midtiby, H. S. (2016). Plant species classification using deep convolutional neural network. Biosystems Engineering, 151, 72–80. https://doi.org/10.1016/J.BIOSYSTEMSENG.2016.08.024

6. Fang, S., Yuan, Y., Chen, L., Zhang, J., Li, M., & Song, S. (2017). Crop Disease Image Recognition Based on Transfer Learning. Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 10666 LNCS, 545–554. https://doi.org/10.1007/978-3-319-71607-7_48
7. Hammad Saleem, M., Khanchi, S., Potgieter, J., & Mahmood Arif, K. (2020). Image-based plant disease identification by deep learning meta-architectures. Plants, 9(11), 1–23. https://doi.org/10.3390/PLANTS9111451
8. He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep Residual Learning for Image Recognition (pp. 770–778). http://image-net.org/challenges/LSVRC/2015/
9. Hughes, D. P., & Salathé, M. (n.d.). An open access repository of images on plant health to enable the development of mobile disease diagnostics. http://www.fao.org/fileadmin/templates/wsfs/docs/expert_paper/How_to_Feed_the_World_in_2050.pdf!!!
10. Krizhevsky, A., Sutskever, I., & Hinton, G. E. (n.d.). ImageNet Classification with Deep Convolutional Neural Networks. Retrieved October 29, 2021, from http://code.google.com/p/cuda-convnet/
11. Li, J., Lin, L., Tian, K., & Alaa, A. A. (2020). Detection of leaf diseases of balsam pear in the field based on improved Faster R-CNN. Nongye Gongcheng Xuebao/Transactions of the Chinese Society of Agricultural Engineering, 36(12), 179–185. https://doi.org/10.11975/j.issn.1002-6819.2020.12.022
12. Long, M., Ouyang, C., Liu, H., & Fu, Q. (2018). Image recognition of Camellia oleifera diseases based on convolutional neural network & transfer learning. Nongye Gongcheng Xuebao/Transactions of the Chinese Society of Agricultural Engineering, 34(18), 194–201. https://doi.org/10.11975/J.ISSN.1002-6819.2018.18.024
13. Mikołajczyk, A., & Grochowski, M. (2018). Data augmentation for improving deep learning in image classification problem. 2018 International Interdisciplinary PhD Workshop, IIPhDW 2018, 117–122. https://doi.org/10.1109/IIPHDW.2018.8388338
14. Niculescu-Mizil, A., & Caruana, R. (2012). Inductive Transfer for Bayesian Network Structure Learning. 27, 167–181. https://doi.org/10.5555/3045796
15. Perez, L., & Wang, J. (2017). The Effectiveness of Data Augmentation in Image Classification using Deep Learning. http://arxiv.org/abs/1712.04621
16. Pratt, L. Y., Pratt, L. Y., Pratt, L. Y., Hanson, S. J., Giles, C. L., & Cowan, J. D. (1993). Discriminability-Based Transfer between Neural Networks. ADVANCES IN NEURAL INFORMATION PROCESSING SYSTEMS 5, 204–211. http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.48.6532
17. Semary, N. A., Tharwat, A., Elhariri, E., & Hassanien, A. E. (2015). Fruit-Based Tomato Grading System Using Features Fusion and Support Vector Machine. Advances in Intelligent Systems and Computing, 323, 401–410. https://doi.org/10.1007/978-3-319-11310-4_35
18. Shi, Y., Wang, X. F., Zhang, S. W., & Zhang, C. L. (2015). PNN based crop disease recognition with leaf image features and meteorological data. International Journal of Agricultural and Biological Engineering, 8(4), 60–68. https://doi.org/10.3965/J.IJABE.20150804.1719
19. Simonyan, K., & Zisserman, A. (2015). Very deep convolutional networks for large-scale image recognition. 3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings.

20. Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., & Rabinovich, A. (2015). Going deeper with convolutions. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 07-12-June-2015, 1–9. https://doi.org/10.1109/CVPR.2015.7298594
21. Taylor, L., & Nitschke, G. (2017). Improving Deep Learning using Generic Data Augmentation. http://arxiv.org/abs/1708.06020
22. Wang, H., Li, G., Ma, Z., & Li, X. (2012). Image recognition of plant diseases based on backpropagation networks. 2012 5th International Congress on Image and Signal Processing, CISP 2012, 894–900. https://doi.org/10.1109/CISP.2012.6469998
23. Y Liu, L Bo, C Yan, J Tang, H Liang. (2018). Maize Diseases Identification Based on Deep Convolutional Neural Network. https://xueshu.baidu.com/usercenter/paper/show?paperid=1h4q0v208d260v50nf2r06b0rm161949&site=xueshu_se
24. Yuan, Y., Fang, S., & Chen, L. (2018). Crop Disease Image Classification Based on Transfer Learning with DCNNs. Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 11257 LNCS, 457–468. https://doi.org/10.1007/978-3-030-03335-4_40
25. Zeng, W., & Li, M. (2020a). Crop leaf disease recognition based on Self-Attention convolutional neural network. Computers and Electronics in Agriculture, 172. https://doi.org/10.1016/J.COMPAG.2020.105341
26. Zeng, W., & Li, M. (2020b). Crop leaf disease recognition based on Self-Attention convolutional neural network. Computers and Electronics in Agriculture, 172. https://doi.org/10.1016/J.COMPAG.2020.105341
27. Zhang, K., Zhang, L., & Wu, Q. (2019). Identification of cherry leaf disease infected by podosphaera pannosa via convolutional neural network. International Journal of Agricultural and Environmental Information Systems, 10(2), 98–110. https://doi.org/10.4018/IJAEIS.2019040105
28. Zhang, S. W., Shang, Y. J., & Wang, L. (2015). PLANT DISEASE RECOGNITION BASED ON PLANT LEAF IMAGE. In J. Anim. Plant Sci (Vol. 25, Issue 3).
29. Zhao, Y.-X., Wang, K.-R., Bai, Z.-Y., Li, S.-K., Xie, R.-Z., & Gao, S.-J. (2009). Research of Maize Leaf Disease Identifying Models Based Image Recognition. Crop Modeling and Decision Support, 317–324. https://doi.org/10.1007/978-3-642-01132-0_35
30. Zhou, J., Li, J., Wang, C., Wu, H., Zhao, C., & Wang, Q. (2021). A vegetable disease recognition model for complex background based on region proposal and progressive learning. Computers and Electronics in Agriculture, 184. https://doi.org/10.1016/J.COMPAG.2021.106101
31. Zhou, J., Li, J., Wang, C., Wu, H., Zhao, C., & Wang, Q. (2021). A vegetable disease recognition model for complex background based on region proposal and progressive learning. Computers and Electronics in Agriculture, 184. https://doi.org/10.1016/J.COMPAG.2021.106101□□□, & □□□. (2017). □□□□ □□□ □□□ □□□ □□ □□□ □ □ □□□ □□ □□. □□□□□□□ □□□□, 1291–1294. https://www.dbpia.co.kr/Journal/articleDetail?nodeId=NODE07219480

Figures

Figure 1

Sample images from natural environment datasets and laboratory datasets.

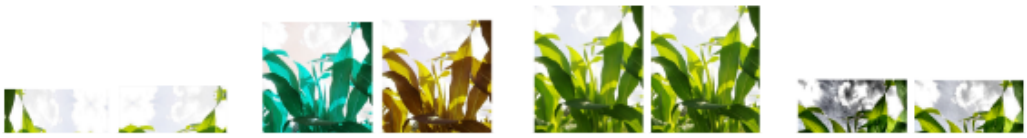


Figure 2

Data enhancement methods.

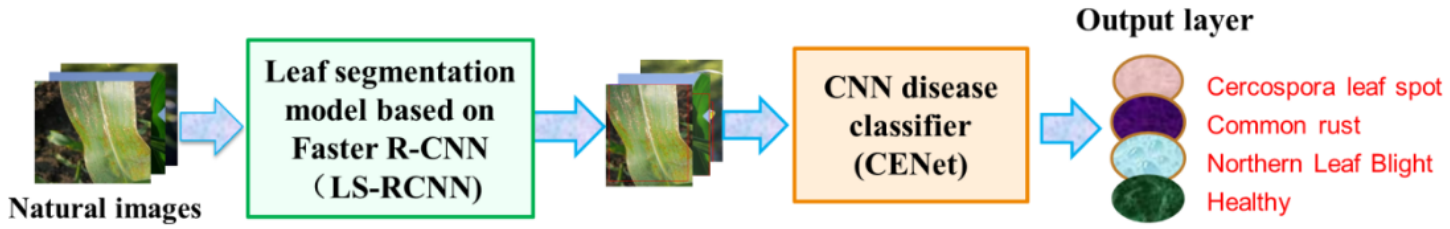


Figure 3

The proposed cascade networks.

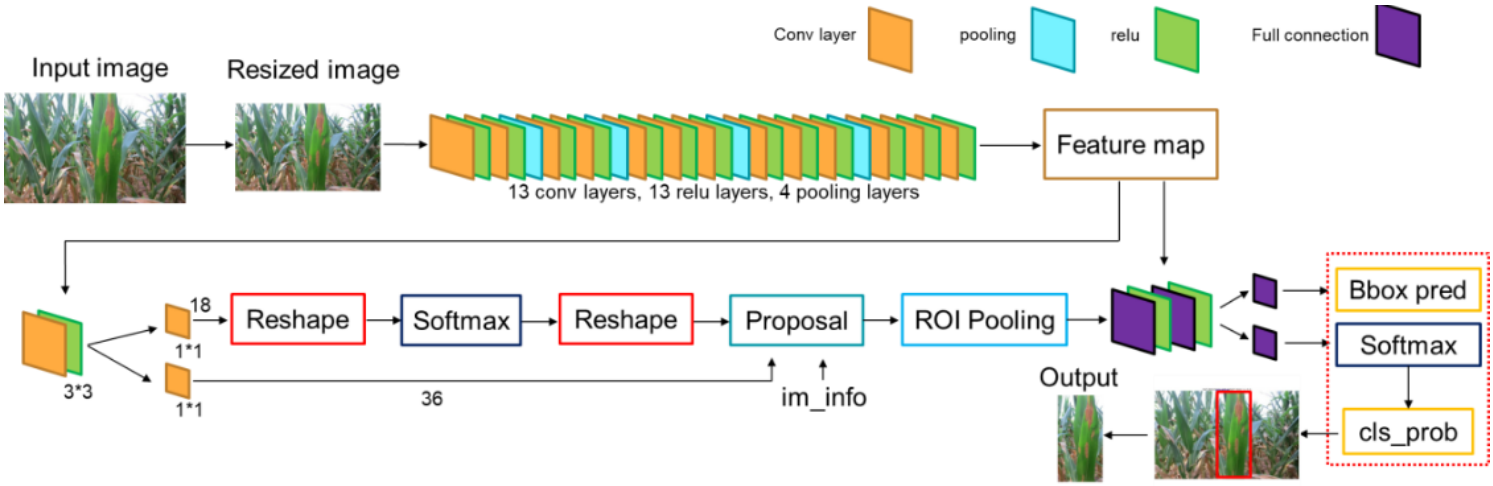


Figure 4

Structure of LS-RCNN model.

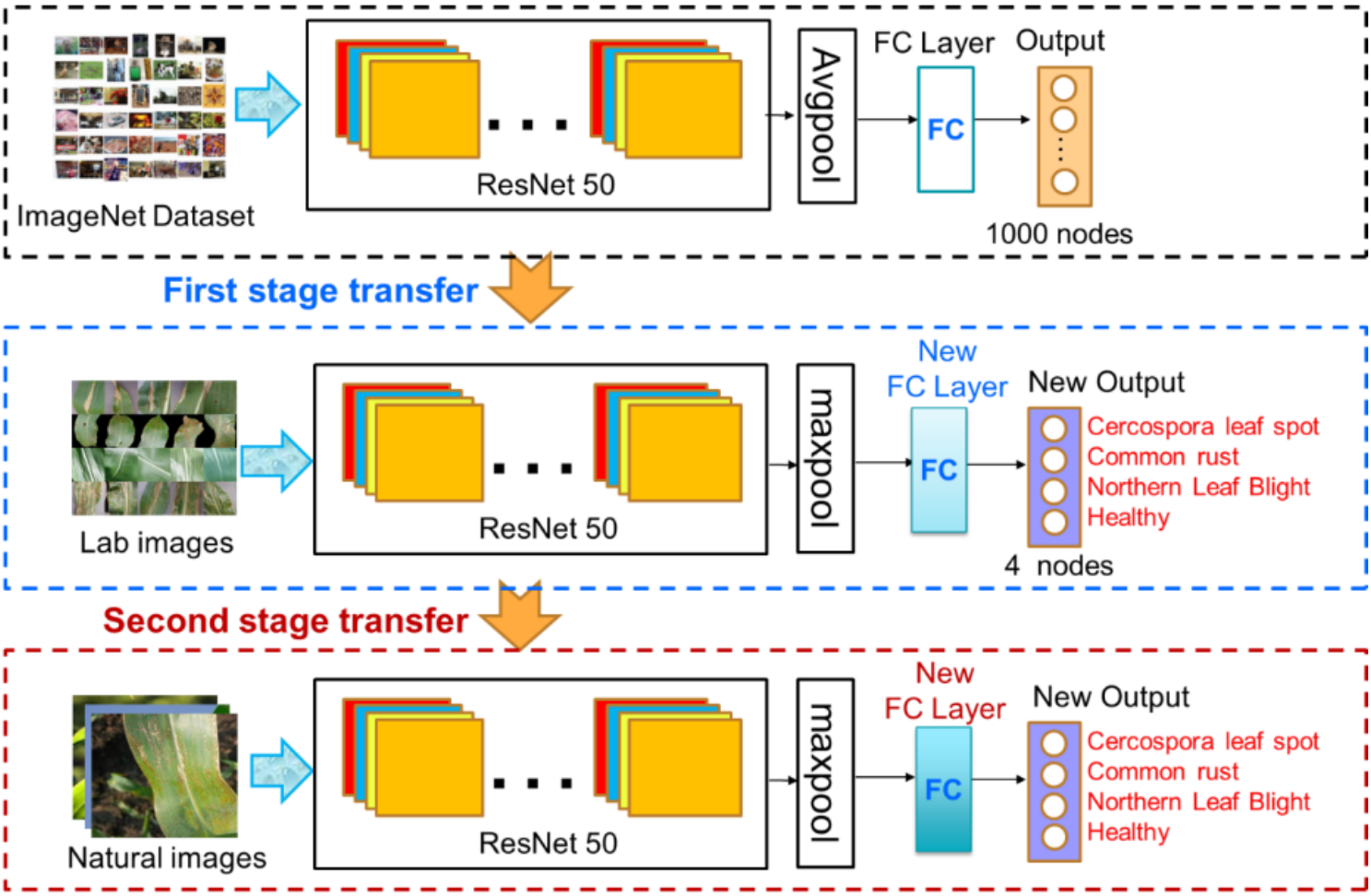


Figure 5

Architecture and training of CENet.

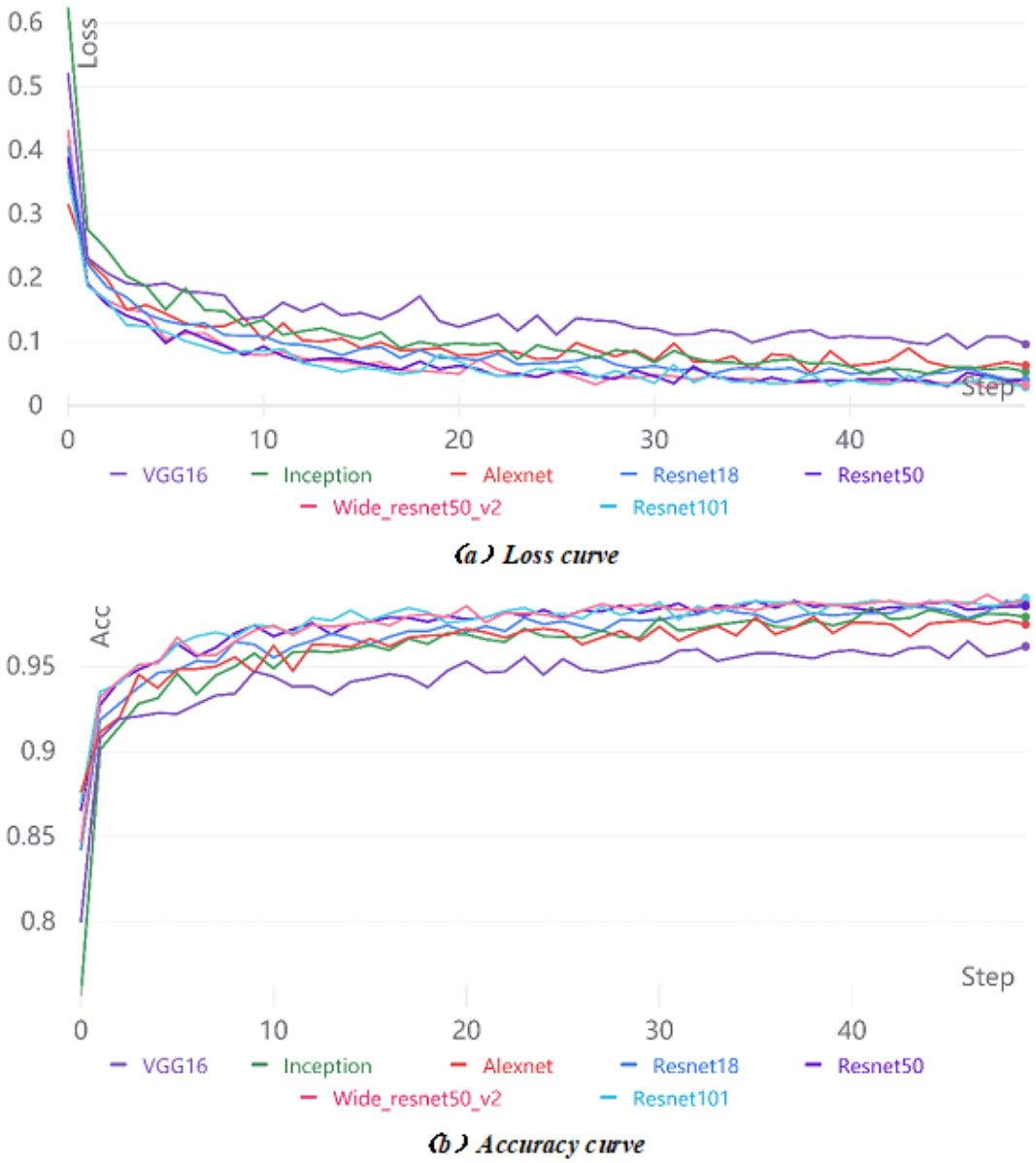
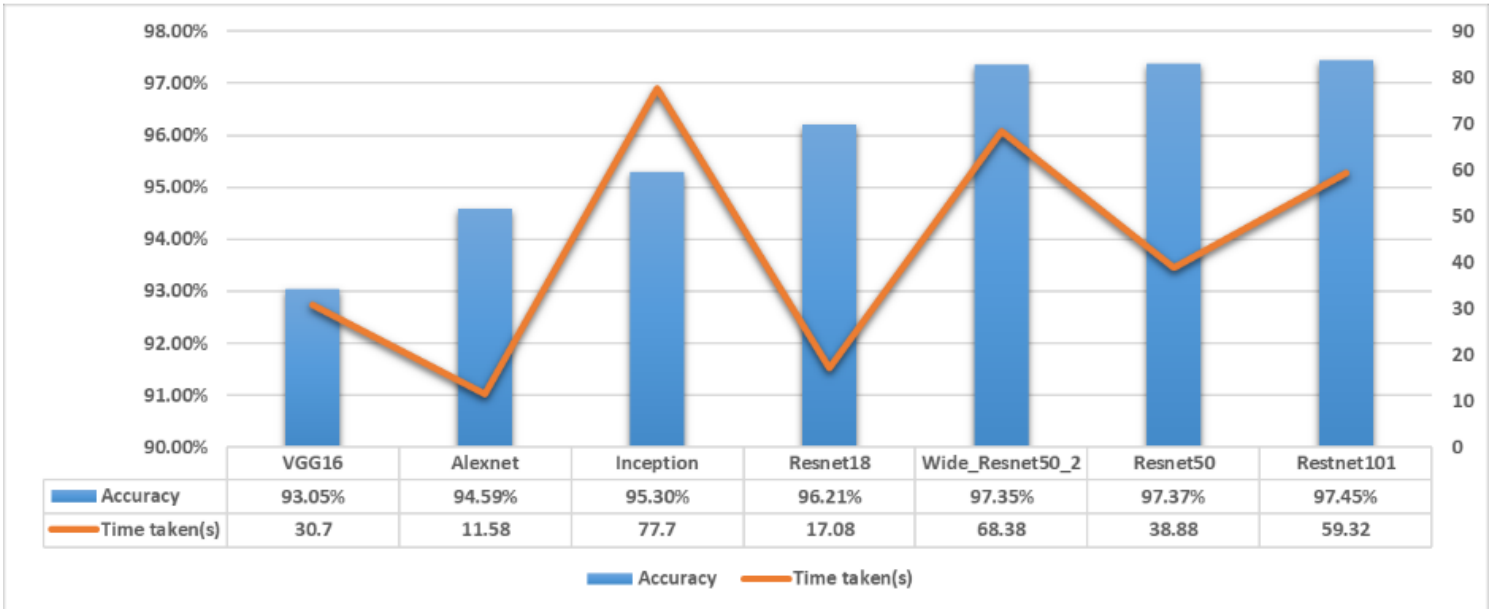


Figure 6

Comparison of recognition results among different convolutional networks.



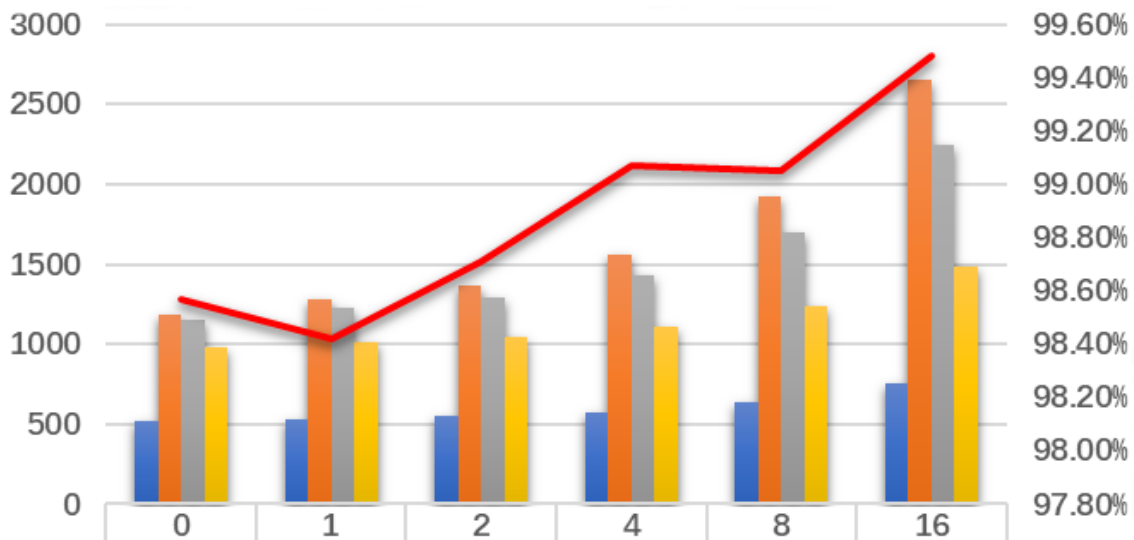
	VGG16	Alexnet	Inception	Resnet18	Wide_Resnet50_2	Resnet50	Restnet101
Accuracy	93.05%	94.59%	95.30%	96.21%	97.35%	97.37%	97.45%
Time taken(s)	30.7	11.58	77.7	17.08	68.38	38.88	59.32

Figure 7

Time of training the model and the accuracy of the model.

Figure 8

Comparison between traditional transfer learning and two-stage transfer learning.



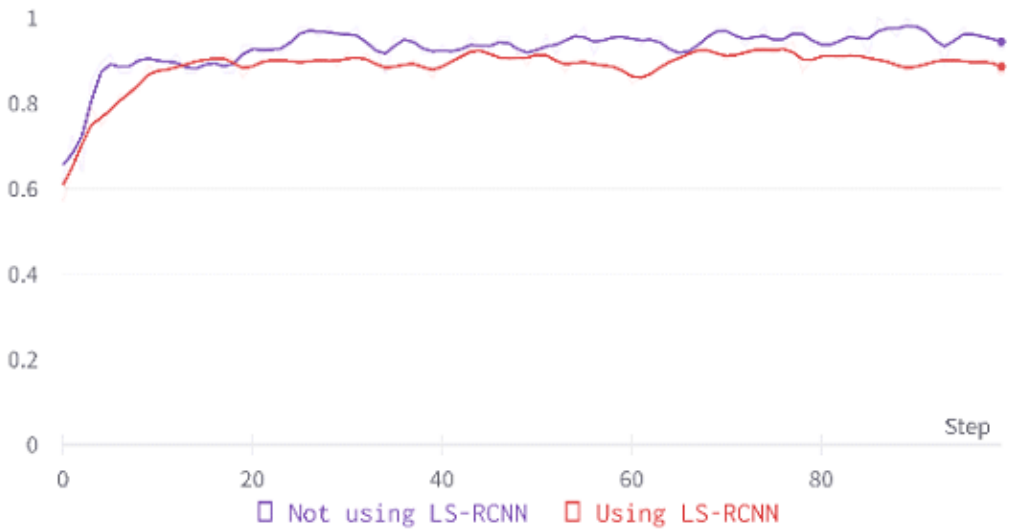
	0	1	2	4	8	16
Cercospora_leaf_spot	518	533	548	578	638	758
Common_rust	1187	1279	1371	1555	1923	2659
healthy	1157	1225	1293	1429	1701	2245
Northern_Leaf_Blight	986	1017	1048	1110	1234	1482
ACC	98.57%	98.42%	98.71%	99.07%	99.05%	99.48%

Figure 9

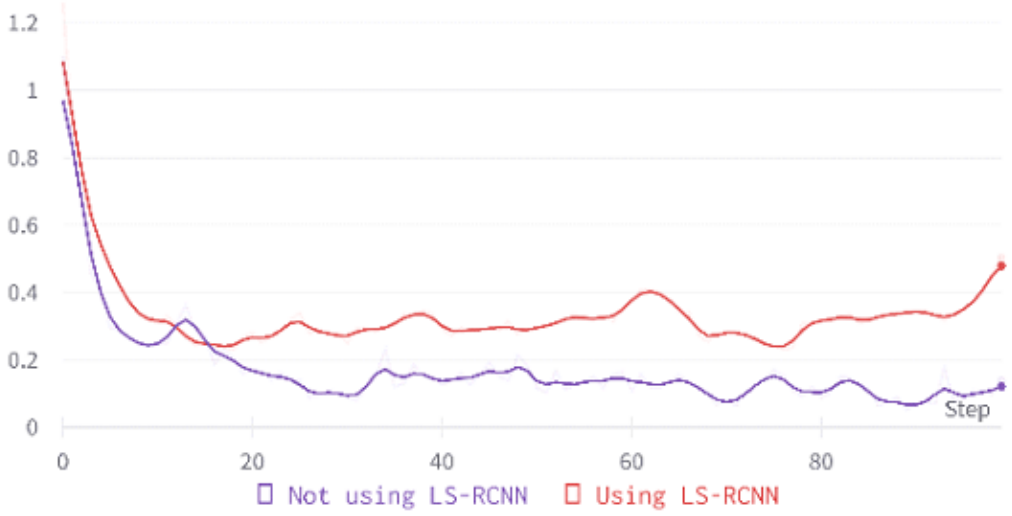
Change of accuracy when natural data sets are expanded exponentially by 2.

Figure 10

The effect of background segmentation using LS-RCNN.



(a) Loss curve



(b) Accuracy curve

Figure 11

Recognition performance of LS-RCNN.

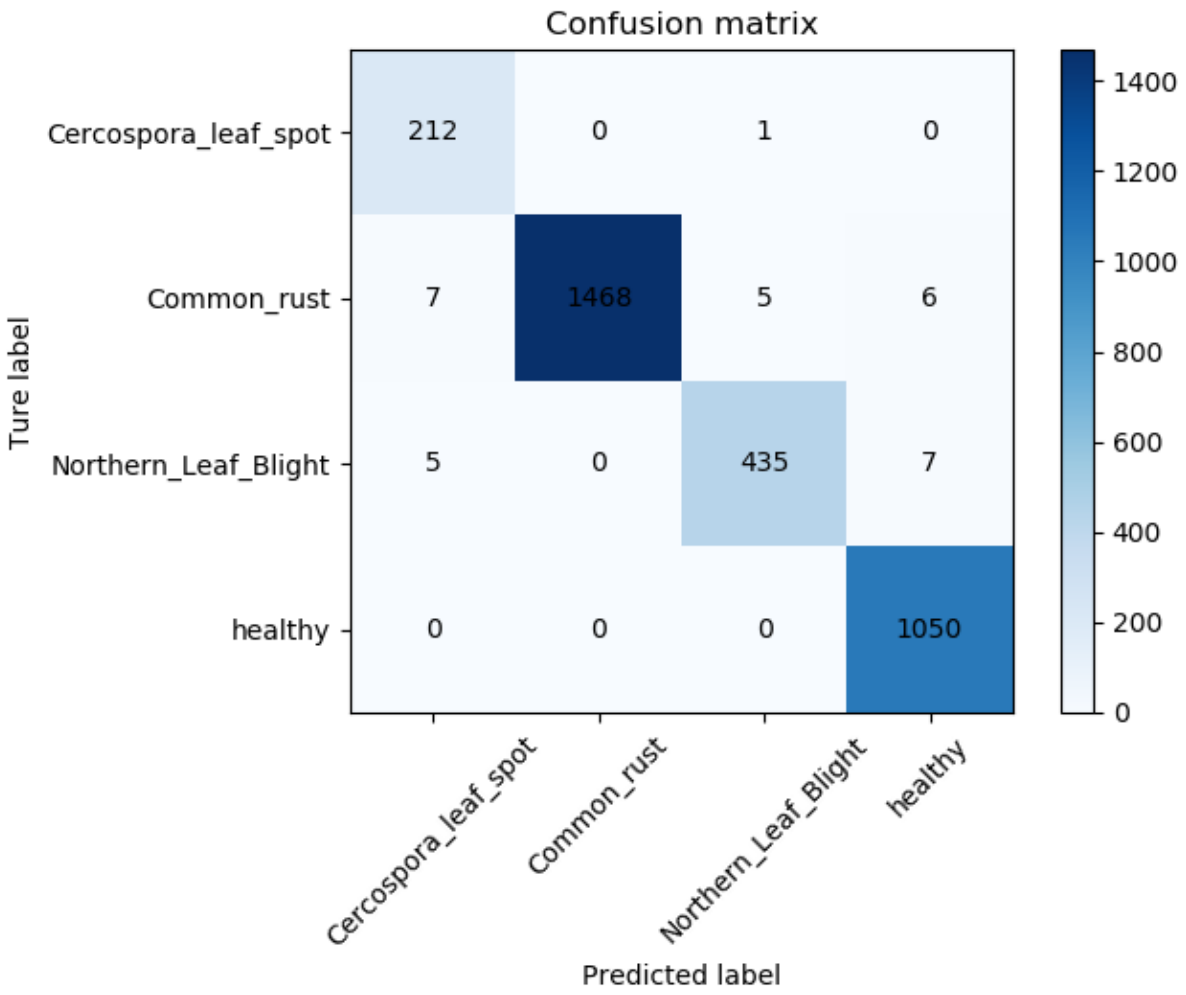


Figure 12

The confusion matrix of our model.