International Journal of Engineering & Technology, 7 (4.11) (2018) 49-54

SPC

International Journal of Engineering & Technology

Website: www.sciencepubco.com/index.php/IJET

Research paper



A Practical Plant Diagnosis System for Field Leaf Images and Feature Visualization

E. Fujita[1]*, H. Uga[2], S. Kagiwada[3], H. Iyatomi[1]

[1]*Applied Informatics, Graduate School of Science and Engineering, Hosei University, Japan*
[2]*Saitama Agricultual Technology Research Center, Japan*
[3]*Clinical Plant Science, Faculty of Bioscience and Apllied Chemistry, Hosei University, Japan*
**Corresponding author E-mail: iyatomi@hosei.ac.jp*

Abstract

An accurate, fast and low-cost automated plant diagnosis system has been called for. While several studies utilizing machine learning techniques have been conducted, significant issues remain in most cases where the dataset is not composed of field images and often includes a substantial number of inappropriate labels. In this paper, we propose a practical automated plant diagnosis system. We first build a highly reliable dataset by cultivating plants in a strictly controlled setting. We then develop a robust classifier capable of analyzing a wide variety of field images. We use a total of 9,000 original cucumber field leaf images to identify seven typical viral diseases, Downy mildew and healthy plants including initial symptoms. We also visualize the key regions of diagnostic evidence. Our system attains 93.6% average accuracy, and we confirm that our system captures important features for the diagnosis of Downy mildew.

Keywords: *convolutional neural networks; feature visualization; image processing; plant diagnosis.*

1. Introduction

Plant diseases affect agricultural production all over the world [1-3]. To minimize the damage and avoid secondary infection, we have to identify the infected plants and apply an appropriate treatment as soon as possible (e.g., removal of infected plants or pesticide application). Plant diagnosis is generally conducted through visual examination by experts with subsequent genetic testing applied as necessary, thus it is usually expensive and time-consuming.

In such circumstances, methodologies for automated plant diagnosis characterized by accuracy, speed and low costs have been requested by the agricultural industry. Several studies have been carried out in response to such requests [4-23]. In [4] used support vector machines (SVM) to classify rice plant diseases and attained 92.7% accuracy. In [5] analysed leaf and stem images of plants with an artificial neural network classifier. Their classifier achieved around 93% accuracy in classifying them into six classes (five diseases and a healthy state). In [7] also used an artificial neural network classifier and showed 87.8% in fungal disease diagnosis. In [12] discriminated cassava diseases in five categories (four diseases and a healthy state) and estimated their severity in five grades from healthy (= 1) to terminal (= 5). They used a combination of their original feature descriptors and classifiers such as linear SVM. They claimed 99.98% and nearly 99% accuracy in disease severity estimation and classification, respectively. In [18] investigated six kinds of Cercospora leaf spots of sugar cane with an evaluation of common statistical and handmade image features. Their method attained 82% accuracy. These methods successfully established preferable performance for their own target task. However, since they are designed based on conventional pattern recognition techniques, i.e. a sequential process of (1) preprocessing including segmentation, detection of the regions of interests (ROI), etc., (2) development of hand-crafted features specially designed for a specific task and (3) classification. Thus, they usually have constraints on their usage.

In recent years, a new machine learning schema called deep learning has demonstrated many promising achievements in a wide range of industries. Convolutional neural networks (CNNs) are a principal aspect of deep learning techniques specialised for machine learning including computer vision. CNNs automatically capture efficient image features for classification from the training images as a part of their learning process. Due to that, they not only significantly reduced the need for the complicated hand-made processes mentioned previously but also achieved high classification performance. Recently, several applications for automated plant diagnosis relying on deep learning have also been proposed [11, 15, 17, 20-23]. In [15] used a total of 54,306 plant leaf images consisting of 14 crop species and 26 diseases for a total of 38 classes of crop-disease pairs from PlantVillage [24] and built CNNs classifiers. Their best score reached an overall accuracy of 99.35%. However, all the leaves used in their study were physically cropped and each leaf was separately placed in front of a uniform colored background and photographed. The conditions are quite different to what we observe in the field, thus we see a noticeable difference in performance in practical situations. In fact, they also noted in their manuscript that the accuracy dropped to around 31% in a different setting from the training images. In addition, we found a significant number of inappropriate label assignments in the PlantVillage dataset. This is a serious problem that open dataset inherently has. Note that the PlantVillage dataset is not currently available to the public. In [22] analysed apple leaves for classifying four kinds of diseases with CNNs. They attained an excellent average accuracy of 97.62%. However, their study also used cropped leaf images, as well as the PlantVillage dataset and therefore these systems cannot be directly applied to practical situations.

Copyright © 2018 Authors. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Meanwhile, robust and practical plant diagnosis systems have been investigated [11, 17, 20-21]. In [20] analysed their wheat disease database 2017 (WDD2017 Dataset) (9,230 images consisting of seven types of diseases) with a full convolutional VGG-based [25] neural network. Their system identifies the location of lesion areas and performs their diagnosis simultaneously, and they reported the best score for mean recognition accuracy of 97.95%. In [21] investigated a total of 5,000 tomato field leaf images. Their pre-trained CNN models are combined with three types of object detection methods (e.g. VGG-net and Faster region-based CNN [26]). Their system also detects and diagnoses infected regions simultaneously and achieved 0.836 mean average precision (mAP) at maximum. We agree that these methodologies work well in some practical situations.

On the other hand, there is still room for further investigations. The first thing that comes to mind is that it is desirable to be able to provide the evidence of the outcome of the system in an easy-to-understand manner. This is an important topic from the viewpoint of system reliability. The following is the limitation of simultaneous identification processes of localization and diagnosis. Since these systems are composed of general object detection and recognition methods, their final output is in the form of boundary boxes and associated recognition results. Accordingly, the target symptoms need to be locally present, i.e. they should have clear boundaries and thus be relatively easy to identify. In general, plant symptoms are highly diverse, especially when the disease is in early stages or when plants are infected with a virus. Symptoms of these cases may appear as a very tiny spot or be spread throughout the leaf and difficult to discern. In either of these cases, urgent treatment is required. To address these issues, we believe heat-map representation is more suitable for visualization purposes in general. Furthermore, this representation is useful as it can be regarded as the evidence of automatic diagnosis.

In the meantime, we have been aiming at developing a practical plant diagnosis system that can reduce the burden on farmers and experts [11, 17, 19]. First, we avoided using open datasets due to their low reliability. We built a strictly controlled field environment, cultivated the target plant and made a highly reliable dataset of which each datum had a confirmed diagnosis.

In our former study [17], we formed our cucumber leaf dataset of which the images were photographed under various conditions, i.e. various backgrounds and photographic conditions as well as intractable initial symptoms. The dataset was composed of a total 7,320 cucumber leaf images (targeting seven typical viral diseases and a healthy diagnosis). We investigated this dataset with CNNs and achieved 83.2% in the four-fold cross validation strategy. We confirmed from this achievement that CNNs have a capability of diagnosing cucumber viral diseases under practical conditions. While this system provided classification results, there was no evidence of the outcome, and thus there is a need for performance improvement.

In this paper, we propose a practical and accurate plant diagnosis system based on our extended highly reliable dataset. Our system accepts field images and provides evidence of the outcome in heat-map representation. More concretely, we used a pre-trained VGG-net model [15] and performed transfer learning. For visualization, we applied Grad-CAM [27] to identify important regions of the image to be the basis of the diagnosis that is the evidence of diagnosis.

2. Methodology

2.1. Dataset and Preprocess

For building an accurate diagnosis system, appropriate training data collection is indispensable. To address this issue, a strictly controlled field environment was built for each disease. The target disease was transmitted in an isolated environment to avoid contamination. The dataset was created from cucumber leaves including seven typical viral diseases (MYSV: melon yellow spot virus, ZYMV: zucchini yellow mosaic virus, CCYV: cucurbit chlorotic yellows virus, CMV: cucumber mosaic virus, PRSV: papaya ring spot virus, WMV: watermelon mosaic virus, KGMMV: kyuri (= cucumber) green mottle mosaic virus), leaves infected with Downy mildew, and healthy leaves. The dataset consisted of a total of 9,000 leaf images (1,000 images per class). Each image was photographed under only one loose requirement; a leaf must be near the middle of the picture. Fig. 1 shows examples of images in the dataset. The dataset includes images of different sizes and aspect ratios. Therefore, the image was cropped to a square around its center. As for data augmentation, the same method proposed in the literature was used [17]. Each training image was rotated clockwise with the incremental step size of 20 degrees and their mirror image was then generated with respect to the horizontal axis. In summary, this process artificially increased the training data by 36 times. Finally, we resized the images to 224×224 pixels.

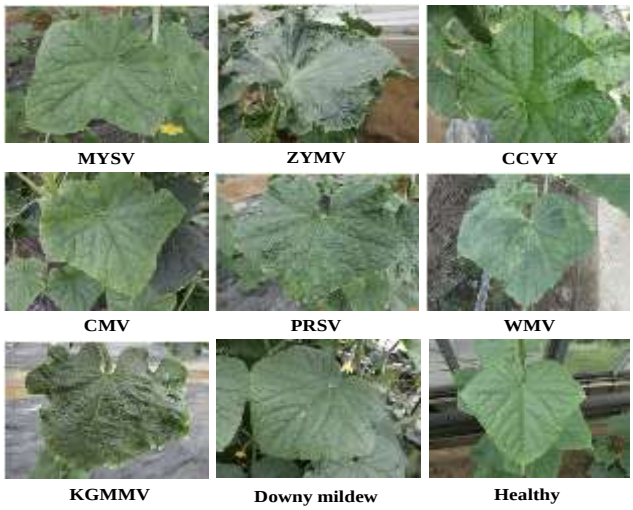

Fig. 1: Example images in our dataset.

2.2. Architecture

Several studies reported that performing transfer learning on the pre-trained large-scale networks such as VGG-net [25], and Res-net [28] with the transfer learning tend to give better performance than models developed from scratch [29]. In plant applications, some studies have used this strategy and showed good results [15, 20, 23]. Therefore, a pre-trained VGG-model was used and fine-tuned using our dataset. Our system was trained with "Caffe" deep learning framework [30]. Model setting and initial learning rate were the same as VGG-net excluding the node size of the output-layer (i.e. from original 1000 to 9 as number of result category).

2.3. Visualization

As a highly reliable system, it is necessary that the system is highly capable on the target task and furthermore it is also necessary to be able to present grounds for the diagnosis in a readable format. Traditionally, this problem was not easy with most machine learning approaches. With recent advancements in deep learning techniques, several promising visualization techniques mainly for CNNs have been proposed to address this issue.

In this study, Grad-CAM [27] was applied to see which leaf regions in the input image make large impacts on the diagnosis that is to say it enable us to see the evidence for diagnosis visually. The Grad-CAM firstly calculates a gradient of the target class on each feature map and averages them to estimate the degree of importance of each map. The visualization is obtained by calculation of a weighted sum of each feature map activation with the importance associated with the input image. The Grad-CAM algo-

rithm is reasonable, and since it does not need additional specially designed components, it does not sacrifice execution speed.

2.4. Experimental

In this study, the abovementioned CNN-based classifier for plant disease diagnosis was built and its validity was investigated by visualising key regions as the basis. For classification performance, our system was evaluated under the 10-fold cross validation strategy. For evaluation of visualization of "evidence", we applied both qualitative (visual) and quantitative assessments. Here, the focus was more on visualising features for Downy mildew than others in this experiment. This is in accordance with suggestions from experts that the typical symptom of viral diseases appears in the whole area of a leaf, but is often not clearly discernible. Even experts cannot determine the exact location of involved areas (see Fig. 4 mentioned later), so visual evaluation of our results for viral diseases was not thought to be very reliable. In order to provide a quantitative evaluation of detected regions presumed to be like the grounds of Downy mildew, those regions were replaced with the mean pixel value of the training data. The transition of the classification result was investigated, while changing the replacement rate of the region of interest for entire Downy mildew images.

3. Results and Analysis

We trained and evaluated the performance of our diagnosis system under 10-fold cross validation. The result is summarized in Table 1 where the right column in bold indicates the result of our proposed model (VGG-based transfer learning) and the left column is our former CNN-model (shallow CNN model trained from scratch) [17]. Using VGG-based transfer learning, our model attains the average 93.6% of accuracy, which is much higher than the scratch-CNN with 86.6% accuracy. The sensitivity of each disease and the specificity of our proposed model also outperform than the previous one. For more detail of our classification performance, the confusion matrix can be found in Fig. 2. Even our model achieves high accuracy, there is still not a small number of misclassification due to the similarity of disease symptoms. The detail of this problem will be discussed in the next section.
Fig. 3 and 4 are examples of estimated regions that were used as the basis of the diagnosis for Downy mildew and viral diseases, respectively. We can see examples from Fig. 3 that our method appropriately determines lesions on leaves infected with Downy mildew by visual evaluation. On the other hand, it is difficult to find the validity on viral diseases from Fig. 4 as mentioned. In general, symptoms of viral diseases often spread out to larger parts of the leaf. Experts diagnose them with the appearance of the whole leaf and they usually do not identify the exact location of lesions. Results observed in Fig. 4 meet the experience of experts and are difficult to evaluate with this quantitative method. For that reason, we evaluate the validity of the heat-map quantitatively only for Downy mildew in this study. Fig. 5 shows an (a) examples of leaves infected by Downy mildew, (b) heat-maps obtained from layer conv2_1 (please refer original manuscript [25] in detail) and (c) images with the region of high interest replaced with the average leaf color.

Table 1: Our diagnosis performance under 10-fold cross validation.

Index	Scratch-CNN* [%]	Proposed [%]
Accuracy	86.6	93.6
Sensitivity of MYSV	86.3	94.3
Sensitivity of ZYMV	86.4	92.2
Sensitivity of CCYV	93.0	96.2
Sensitivity of CMV	86.0	93.5
Sensitivity of PRSV	82.3	91.7
Sensitivity of WMV	79.0	90.8
Sensitivity of KGMMV	81.7	90.1
Sensitivity of Downy mildew	98.0	99.6
Specificity	86.7	94.5

*Scratch-CNN model is the same architecture used in [17]



Fig. 2: Example of misclassified images due to similar features. Left image: actual class = KGMMV (5.4%), prediction class = PRSV (94.5%) Right image: actual class = CMV (17.5%), prediction class = healthy (80.9%)

Table 2: Confusion matrix of our proposed system

		Predicted Class								
		MYSV	ZYMV	CCYV	CMV	PRSV	WMV	KGMMV	Downy Mildew	Healthy
Actual Class	MYSV	943	4	4	14	12	7	4	1	11
	ZYMV	12	922	2	4	12	11	31	0	6
	CCYV	14	0	962	9	2	0	0	1	12
	CMV	16	0	3	935	3	9	0	6	28
	PRSV	20	19	0	3	917	6	34	0	1
	WMV	8	15	3	17	7	908	22	0	2
	KGMMV	15	35	1	5	29	12	901	0	2
	Downy mildew	0	0	0	0	2	0	0	996	2
	Healthy	2	3	6	34	0	7	2	1	945

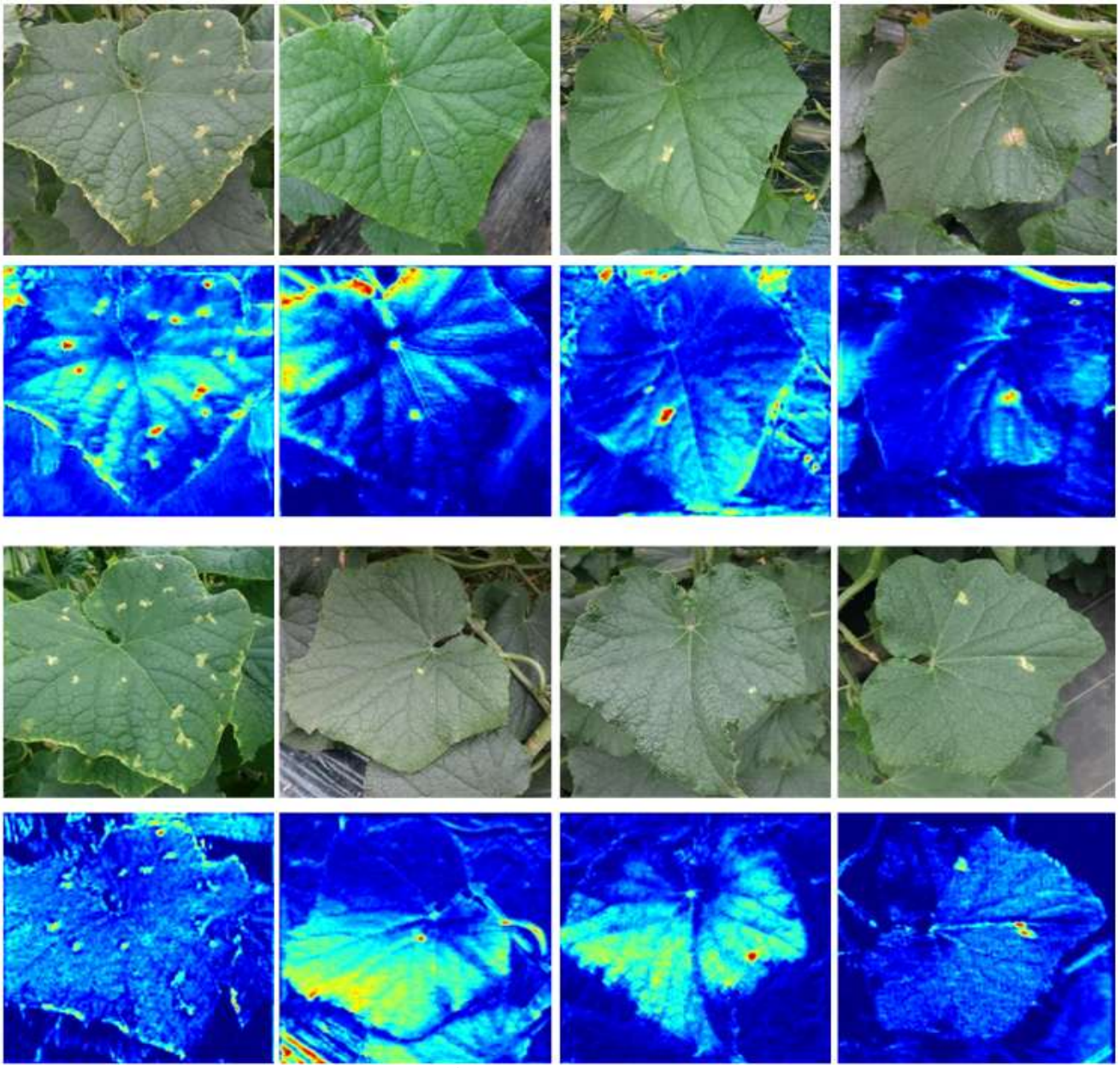

Fig. 3: Example of Downy mildew and corresponding heatmaps.

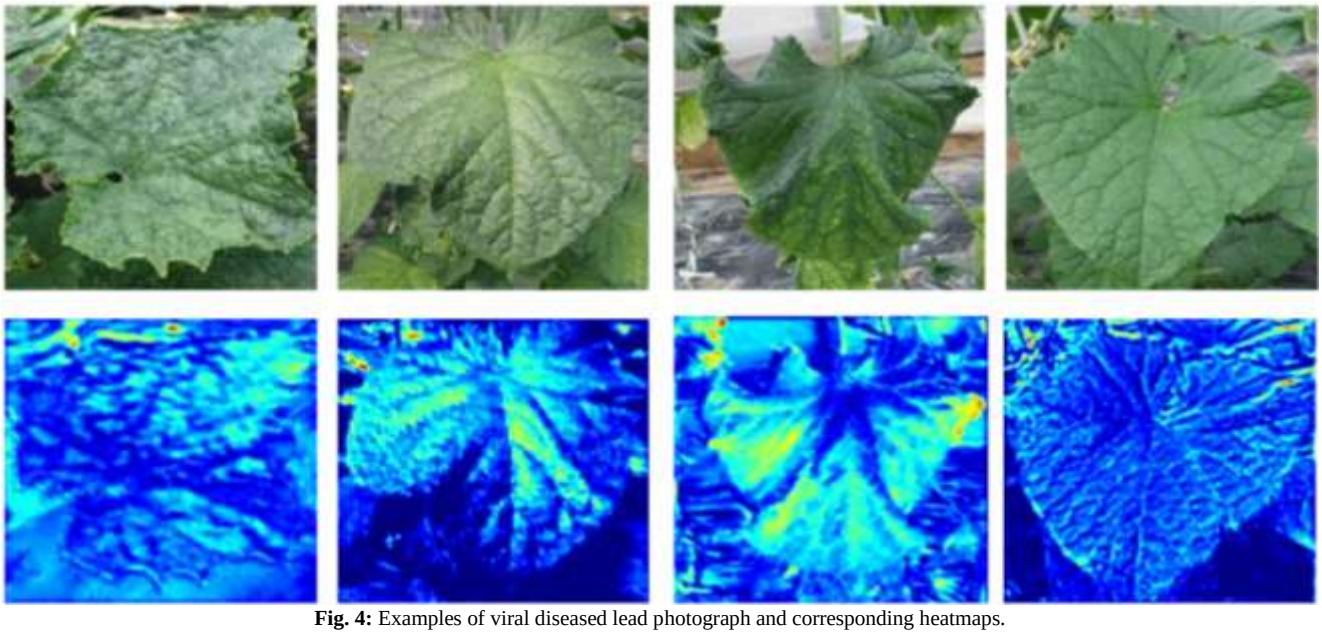

Fig. 4: Examples of viral diseased lead photograph and corresponding heatmaps.

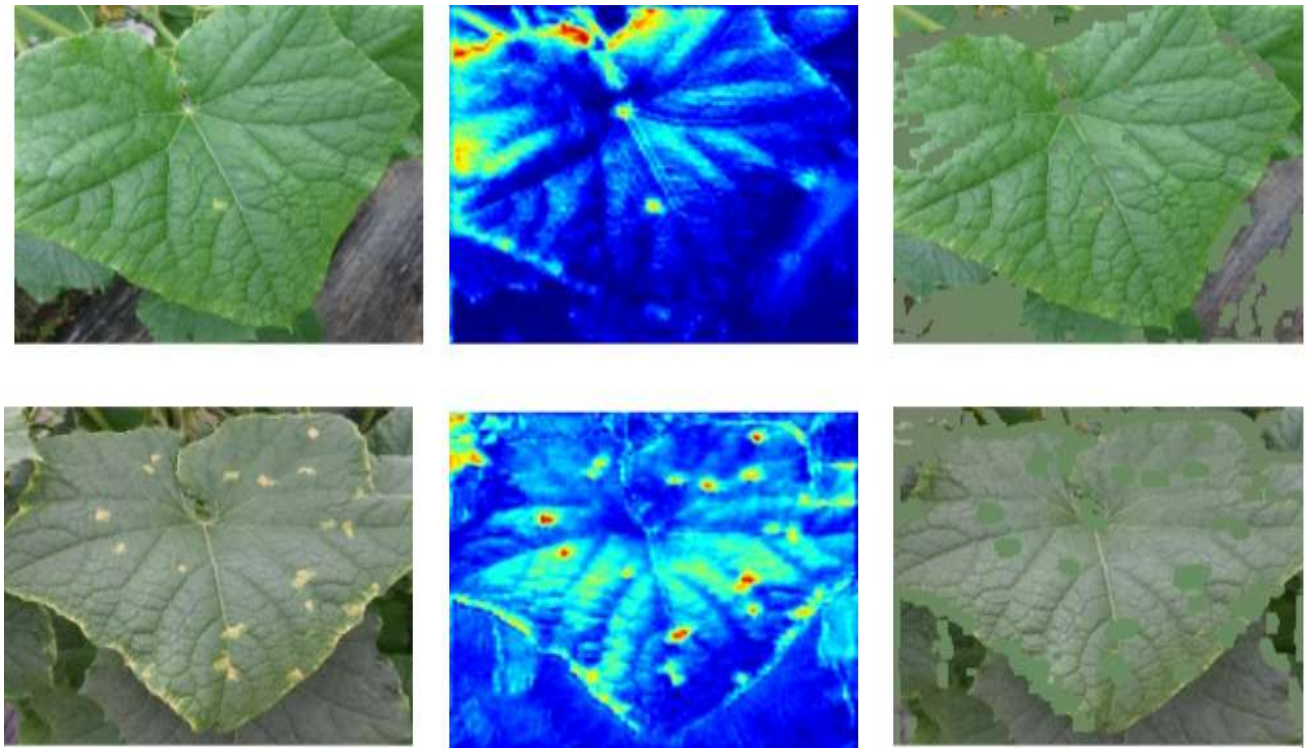

Fig. 5: Examples of (a) Downy mildew leaves, (b) corresponding heat-maps and (c) with high interest regions replaced with the average pixel value of the training images.

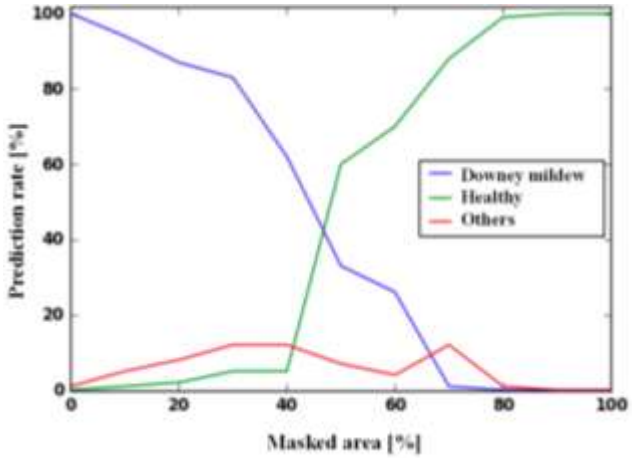

Fig. 6: The transition of the classification result for Downy mildew images seen while changing the replacement rate of high interest heat-map regions.

4. Discussion

4.1. Evaluation of Classification Performance

The performance of our proposed system achieved over 90% accuracy for all diseases (Fig. 1). This is superior to our previous CNN model trained from the scratch. Our results support the hypothesis that utilizing transfer learning is also effective for automated plant diagnosis. For Downy mildew, a typical mold disease, our model attained quite high accuracy (99.6%). We think this is because their symptoms are clearer than those of other diseases. Next, we would like to examine the degree of misclassification by getting feedback from experts. From Table 2, we can see a relatively large number of misclassifications were found between: (i) ZYMV←→KGMMV, (ii) CMV←→healthy, (iii) WMV→KGMMV and WMV→healthy, (iv) PRSV←→ KGMMV. Here, the source and destination of arrows indicate the true and incorrectly estimated label respectively. A double headed arrow indicates misclassification occurring in both directions. Experts commented on our experimental difficulties as follows: "severe symptoms of ZYMV, KGMMV, PRSV, MYSV sometimes show similar appearances to each other." "The initial symptoms of KGMMV sometimes resemble that of WMV." "For recent breed improved cucumbers, CMV sometimes delays their growth and withers leaves, but their symptoms are usually very mild." "Most of initial stage of WMV yields no clear symptoms, so it is hard to discriminate them from healthy ones." "In any case, viral diseases cause mild mosaic symptoms in young leaves and they often look similar each other." According to their comments, we find most of the misclassifications observed in our experiments meet expectations. That is those classifications are inherently difficult not only for automated systems, but also for experts. Despite these difficult situations, our system shows very small number of misclassification. Therefore, we think that the error is in an acceptable range.

As for the reason of high classification performance on Downy mildew, the symptoms basically appear locally and most of them have a clear boundary except in the severe terminal stage. This tendency is much different from that with viral infection. Diagnosis of Downy mildew by humans is relatively easy compared with that of viral diseases. From these results, we confirmed that the trend of classification accuracy is similar in human experts and in our automated system according to the discrimination difficulty level.

4.2. Evaluation of Visualization

We confirmed from Fig. 3 that heat map intensity changes (from yellow to red) appear in the area considered to be mold symptom in many cases. We can also evaluate the validity of the displayed heat map area quantitatively as illustrated in Fig. 6. This figure shows the transition of the classification result for whole (1,000) Downy mildew images under 10-fold cross validation while changing the replacement rate of the ROIs. We can see as the replaced area increases, the number of Downy mildew diagnoses is decreased while the number of healthy diagnoses is increased. This trend indicates that the heat map represents the important local pattern allowing the system to identify diseases in the target

image. Thus, we conclude that our system captured characteristics of Downy mildew appropriately.

However, on the other hand, some heat map responses are observed from outside of the leaf region. These imply that our system felt overfitting for training data. In order to avoid this problem, we primarily need to increase the number of training images with a much wider variety and it then might be necessary to introduce some mask process to eliminate or reduce the effect of out of ROI regions as a second option.

In this study, we obtained heat-map representation also for viral diseases (Fig. 4). We believe that those results are reasonable because the heat-map is generated by the GradCAM algorithm, and this algorithm only depends on the combination of activation from each feature map of the trained classification CNNs. Our CNNs attained more than 90% classification accuracy for all the viral diseases. Still, we did not evaluate the validity of them in a quantitative manner. This is because, as we mentioned earlier, the typical symptoms appear in the whole area of the leaf and are often not discernible even by experts. Thus, we do not have good evaluation criteria that we (humans) can understand. Our experience is that human-oriented features are useful for our understanding, but utilizing only those features seriously deteriorates the performance of learning machines. We need to find good balance between them in future.

5. Conclusion

In this paper, we proposed an accurate and practical cucumber diagnosis system based on a highly reliable dataset and CNNs with transfer learning. Our system achieved a 93.6% classification accuracy on average for nine classes. We also visualized the evidence of diagnosis for Downy mildew, a typical mold disease and confirmed the validity through both visual and quantitative assessment.

Acknowledgement

This research was supported in part by the Ministry of Education, Science, Sports and Culture, Grant-in-Aid for Fundamental Research (C), 17K08033, 2017–2020.

References

[1] Majumdat D, Kole D K, Chakraborty A, Majumder D D. Review: Detection and diagnosis of plant leaf disease using integrated image processing approach. International Journal of Computer Engineering and Applications, 2014, 6(3), 10-16.

[2] Themen D. 2014. Reduction of food losses and waste in Europe and Central Asia for improved food security and agrifood chain efficiency. http://www.fao.org/publications/card/en/c/3cc68649-a7c4-401d-ad3b-14bb238f14a6.

[3] Little E L. 2014. Georgia plant disease loss estimates. http://extension.uga.edu/publications/detail.cfm?number=AP102-7.

[4] Yao Q, Guan Z, Zhou Y, Tang J, Hu Y, Yang B. Application of support vector machine for detecting rice diseases using shape and color texture features. Proceedings of the Engineering and Computation, 2009, pp. 79-83.

[5] Bashish D A, Braik M, Bani-Ahmad S. A framework for detection and classification of plant leaf and stem diseases. Proceedings of the International Conference on Signal and Image Processing, 2010, pp. 113-118.

[6] Sannakkiet S S, Rajpurohit V S, Nargund V B, Kulkarni P. Diagnosis and classification of grape leaf diseases using neural networks. Proceedings of the International Conference on Computing, Communications and Networking Technologies, 2013, pp. 1-5.

[7] Pujari J D, Yakkundimath R, Byadgi A S. Recognition and classiffcation of produce affected by identically, looking powdery Mildew disease. Acta Technologica Agriculturae, 2014, 17(2), 29-34.

[8] Barbedo G A. An automatic method to detect and measure leaf disease symptoms using digital image processing. Plant Disease. 2014, 98, (12) 1709-1716.

[9] Patil S P, Zambre R S. Classification of cotton leaf spot disease using support vector machine. Engineering Research and Applications, 2014, 4(5), 92-97.

[10] Li B, Hulin M T, Brain P, Mansfield J W, Jackson R W, Harrison R J. Rapid, Automated detection of stem canker symptoms in woody perennials using artificial neural network analysis. Plant Methods, 2015, 111(), 1-9.

[11] Kawasaki Y, Uga H, Kagiwada S, Iyatomi H. Basic study of automated diagnosis of viral plant diseases using convolutional neural networks. Lectures Notes in Computer Science, 2015, 9475, 638-645.

[12] Mwebaze E, Owomugisha O G. Machine learning for plant disease incidence and severity measurements from leaf images. Proceedings of the International Conference on Machine Learning and Application, 2016, pp. 158-163.

[13] Paware P, Tukar V, Patil P. Cucumber disease detection using artificial neural network. Proceedings of the International Conference on Inventive Computation Technologies, 2016, pp. 1-5.

[14] Abdullah S N H S, Abdullah S, Petrou M, Razalan S A. A portable rice disease diagnosis tool basedon bi-level colour image thredholding. Applied Engineering in Agriculture, 2016, 32(4), 295-310.

[15] Mohanty S P, Hughes D P, Salathe M. Using deep learning for image-based plant disease detection. Frontiers in Plant Science, 2016, 7(1419), 1-10.

[16] Qin F, Liu D, Sun B, Ruan L, Ma Z, Wang H. Identification of alfalfa leaf diseases using image recognition technology. Plos One, 2016, 11(12), 1-26.

[17] Fujita E, Uga H, Kagiwada S, Iyatomi H. Basic investigation on a robust and practical plant diagnostic system. Proceedings of the. International Conference on Machine Learning and Application, 2016, pp. 989-992.

[18] Hallau L, Neumannbc M, Klattc B, Kleinhenzd B, Kleind T, Kuhnd C, Roehrige M,Bauckhageb C, Kerstingf K, Mahleina A K, Steinera U, Oerke E C. Automated identification of sugar beet diseases using smartphones. Plant Pathology, 2017, 67(2), 399-410.

[19] Cap H Q, Fujita E, Kagiwada S, Uga H, Iyatomi H. A deep learning approach for on-site plant leaf detection. Proceedings of the Signal Processing and its Applications, 2018, pp. 120-124.

[20] Lu J, Hu J, Zhao G, Mei F, Zhang C. An field automatic wheat disease diagnosis system. Computers and Electronics in Agriculture, 2017, 142, 369-379.

[21] Fuentes A, Yoon S, Kim S C, Park D S. A robust deep-learning-based detector for real-time tomato plant diseases and pests recognition. Sensors, 2017, 17(9), 1-21.

[22] Liu B, Zhang Y, He D J, Li Y. Identification of apple leaf diseases based on deep convolutional neural networks. Symmetry, 2018, 10(1), 1-16.

[23] Atabay H A. Deep residual learning for tomato plant leaf disease identification. Journal of Theoretical and Applied Information Technology, 2018, 95(24), 6800-6808.

[24] Hughes D P, Salathe M. An open access repository of images on plant health to enable the development of mobile disease diagnostics. 2015, https://arxiv.org/ftp/arxiv/papers/1511/1511.08060.pdf.

[25] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. 2014, https://arxiv.org/pdf/1409.1556.pdf%20http://arxiv.org/abs/1409.1556.pdf.

[26] Ren S, He K, Kaiming G, Girshick R, Sun J. Faster R-CNN: Towards real-time object detection with region proposal networks. Proceedings of the Advances in Neural Information Processing Systems, 2015, pp. 91-99.

[27] Selvaraju R R, Cogswell M, Das A, Vedantam R, Parikh D, Batra D. Grad-CAM: Visual explanations from deep networks via gradient-based localization. Proceedings of the International Conference Computer Vision, 2017, pp. 618-626.

[28] He K, Zhang X, Ren S, Sun J. Deep residual learning for ImageNet recognition. Proceedings of the International Conference on Computer Vision and Pattern Recognition, 2016, pp. 770-778.

[29] Tajbakhsh N, Shin J Y, Gurudu S T, Hurst R T, Kendall C B, Gotway M B, Liang J. Convolutional neural networks for medical image analysis: Full training or fine tuning? IEEE Transactions on Medical Imaging, 2016, 35(5), 1299-1312.

[30] Jia Y, Shelhamer E, Donahue J, Karayev S, Long J, Girshick R, Guadarrama S, Darrell T. Caffe: Convolutional architecture for fast feature embedding. Proceedings of the 22nd ACM International Conference on Multimedia, 2014, pp. 675-678.