Proceedings of the New Trends in Information Technology (NTIT-2017).
The University of Jordan, Amman, Jordan. 25-27 April 2017.

An Image Processing Approach for Cucumber Powdery Mildew Infection Detection

Bassam AL-Qarallah, Bashar Al-Shboul, Hazem Hiary, Asmaa Aljawawdeh, Hamad Alsawalqah, Monther Tahat
The University of Jordan,
Amman, 11942, Jordan
bassam702000@yahoo.com, bashar.shboul@gmail.com, hazemh@ju.edu.jo, asma.jawawdeh@gmail.com, h.sawalqah@ju.edu.jo, m.tahat@ju.edu.jo

Abstract

Plant disease detection is one of the most active topics in the modern agriculture literature. This work proposes a new method of leaf disease classification depending on various extracted features from plant leaves utilizing different image processing of image color models/schemes. After validating our proposed work against the human judgement of infectious leaf images, our proposed method show promising results, after comparing various machine learning classification algorithms.

Keywords: Powdery Mildew; Classification; Image Processing; Leaf Disease Detection

1. INTRODUCTION

Plant disease detection is one of the most active topics in the modern agriculture literature. It attracts much attention considering the importance of automating this process in various agricultural crops e.g. cucumber, tomatoes, potatoes.

Many attempts to detect crops' leaf diseases of different types (i.e. powdery mildew) were reported. A general observation is that they follow few major steps: first, an image processing step that clears noise, segment the image. Then, region detection and/or feature extraction. Finally, a machine learning approach is utilized to classify a leaf's image into an infected, or non-infected. In principle, our work in this paper matched the previously major steps; however, details within each of which vary aiming at producing high detection effectiveness. Additionally, we compare few state of the art classification methods against human judgment to evaluate our proposed method.

This paper is organized as follows. In section 2, related works are explained. In section 3, this research methodology is discussed. The experiment setup, and the experimental results are detailed in sections 4 and 5 respectively. Finally, we conclude with section 6.

2. RELATED WORKS

The literature is rich in the field of plant disease detection. In [1], the authors combined information from thermal and stereo visible light images using machine learning techniques for tomato powdery mildew. Their proposed algorithm consists of four parts: image registration; depth estimation; feature extraction and classification. They extracted a novel feature set from the image data using local and global statistics. Then a global approach was used to detect disease from the whole plant. They achieved classification accuracy of 90%.

In [2], the authors proposed an application of image processing and machine learning in identifying three palm oil diseases (anthracnose, hawar leaf, and leaf spot because of Pestalotiopsis Palmarum) based on visual appearances in the early cultivation stage. They classified the image into background, spot, and leaf with iteration for some conflict cases. Image filtering was then applied to remove background and leaf areas. Breadth first search algorithm was used to form all spot pixels. A Neural network was used to learn the pattern and generate a classification model. The average accuracy was 87.75%.

In [3], the authors proposed an automatic approach to detect and classify leaf diseases affecting the fruit crops. They provided a description of leaf disease detection based on color, texture, and shape.

In [4], the authors proposed an automatic crop disease recognition method of cucumber leaves infected with diseases of downy mildew, blight, and anthracnose. They partitioned the image into regions, then statistical features of color, shape, and texture were extracted from the crop disease leaves and meteorological data. Then the probabilistic neural networks (PNNs) classifier was adopted to evaluate the classification accuracy, which was a Bayesian classifier with accuracy rate about 90%.

In [6], the authors proposed an approach to identify the diseases of greenhouse cucumber, namely, downy mildew, powdery mildew, and anthracnose. They applied a range of pre-processing techniques to smooth the images, then they extracted characteristic parameters: energy mean, standard deviation of the entropy, the rectangularity mean, B mean. Based on the shortest distance methods, the experimental results showed that their method achieved recognition accuracy of 96%.

In [7], the researchers presented a study on the image processing techniques used to identify and classify fungal disease symptoms affected on different agriculture crops. They covered the fruit, vegetable, commercial, and cereal crops. In fruit crops, they applied segmentation using K-means clustering, then feature selection and texture selection, after that they used ANN and NN for classification. for the vegetable crops, they used the chan-vese segmentation, then feature selection using local binary patterns and they have used SVM, k-NN for the classification. For commercial crops, they have used a grab-cut for segmentation, a wavelet based on feature selection and Mahalonbies distance and PNN for the classifiers. Finally, for the cereal crops, they have used the K-means clustering, canny edge for segmentation, and for feature selection: color, shapes, texture, color texture and radon transform and they have used SVM, k-NN for classification. Their classification accuracy for the four types ranged between 76.6% and 94%.

In [8], the authors proposed a method for cucumber disease recognition. They applied noise removal using median filtering, followed by extraction of chromaticity moments for texture features in CIE XYZ color space.

In [9], the authors presented an automatic recognition method for cucumber disease, namely, downy mildew, anthracnose, gray mold and powdery mildew. They applied segmentation via thresholding using two-dimensional maximum entropy principle and optimized with differential evolution algorithm. Then they extracted color and texture features, and applied Bayes classifier model.

In [10], the authors proposed a method to recognize diseases in cucumber leaves. In pre-processing stage, images were de-noised using median filtering. Then, segmentation via region growing was effective to extract image features in the regions of interest (leaf area). Classification is then applied using BP neural network, with accuracy rate of 80%.

In [11], the authors detected cucumber downy mildew disease in images with support vector machine (SVM) classification method. They applied image pre-processing, feature parameter extraction, and pattern recognition stages. The accuracy of detection was about 90%.

In [12], the authors proposed an approach to detect cucumber diseases, namely, downy mildew, blight and anthracnose using singular value decomposition (SVD). First, they segmented the images using the watershed algorithm. Second, each image was divided into blocks, and then the combining features of Global-Local singular values were extracted and ordered from each block by SVD. Third, the key-point vectors were constructed and their dimensionalities were adjusted. Finally, an SVM classifier was adopted to recognize the class of the disease leaf image. The experimental results showed that the proposed method was effective and feasible.

3. METHODOLOGY

Images used for this experimental setup were collected from two different resources, namely: cucumber greenhouses at The University of Jordan, and from a dataset of UCI Machine Learning Repository (Leaf Dataset). In this research, images have no special photo shooting conditions similarly to other research (e.g. must be captured in white background, sunlight, etc.). After that, every captured image will be converted into a 3D matrix containing RGB values for each pixel. A round of image pre-processing is performed, trying to remove image noise and isolating colours of interest (i.e. green spectrum, white spectrum) per se. switching between different colour models, and testing different values of colour ranges empirically improved image quality for further steps. The difference between the original photo and the processed one is shown at Fig 1(a) and Fig 1(b).

Afterwards, the resulting image is converted from RGB into grey scale for filtering as shown in Fig 1(c). As used images are from different angles and different illuminations with dark backgrounds, a major challenge is to correct the image illumination; therefore, Flat morphological top-hat structuring filter is applied. The resulting image is shown in Fig 1(d).

In order to detect infected leaf regions, the image will be transformed from grayscale into monochrome in an effort to isolate white spots within leaf black regions; therefore, Otsu's binarization algorithms that involves iterating through all the possible threshold values and calculating a measure of spread for the pixel levels each side of the threshold. The aim is to find the threshold value where the sum of foreground and background spreads is at its minimum, i.e. the intraclass variance of the black and white pixels. The resulting image is shown in Fig 1(e). However, in order to preserve the white regions of the disease, hole filling and noise addition using salt & pepper is applied, and then applying a 2D median filter on it. The resulting image will mask the image in Fig 1(a) as shown in Fig 1(f).

Finally, the resulting image of previous step is converted into the L*a*b model to apply on additional step of noise filtering where image is first transformed into monochrome scale, then filtered similarly to previous step, then used to mask the image in Fig 1(a) producing the image displayed in Fig 1(g).

Previous steps are repeated for another round of filtering to decrease the remaining noise. The final outcome is shown in Fig 1(h).

4. EXPERIMENT SETUP

Images used for this research were collected from two different resources: images were taken from a live environment at The University of Jordan, while another group of images were collected from a dataset at UCI Machine Learning Repository.

A SONY camera model DSC-H55 is used to capture images with a resolution of 2592*1944 pixels. An overall set of 59 images were utilized to apply our method in this experiment.

In this research, MATLAB (R2014b) we used for image processing and programming environment.

Two professors from the school of agriculture, the university of Jordan were asked to annotate the images as infected/non-infected according to their experience, after that results were compared to make sure they match, and then the relevance judgement is made accordingly.

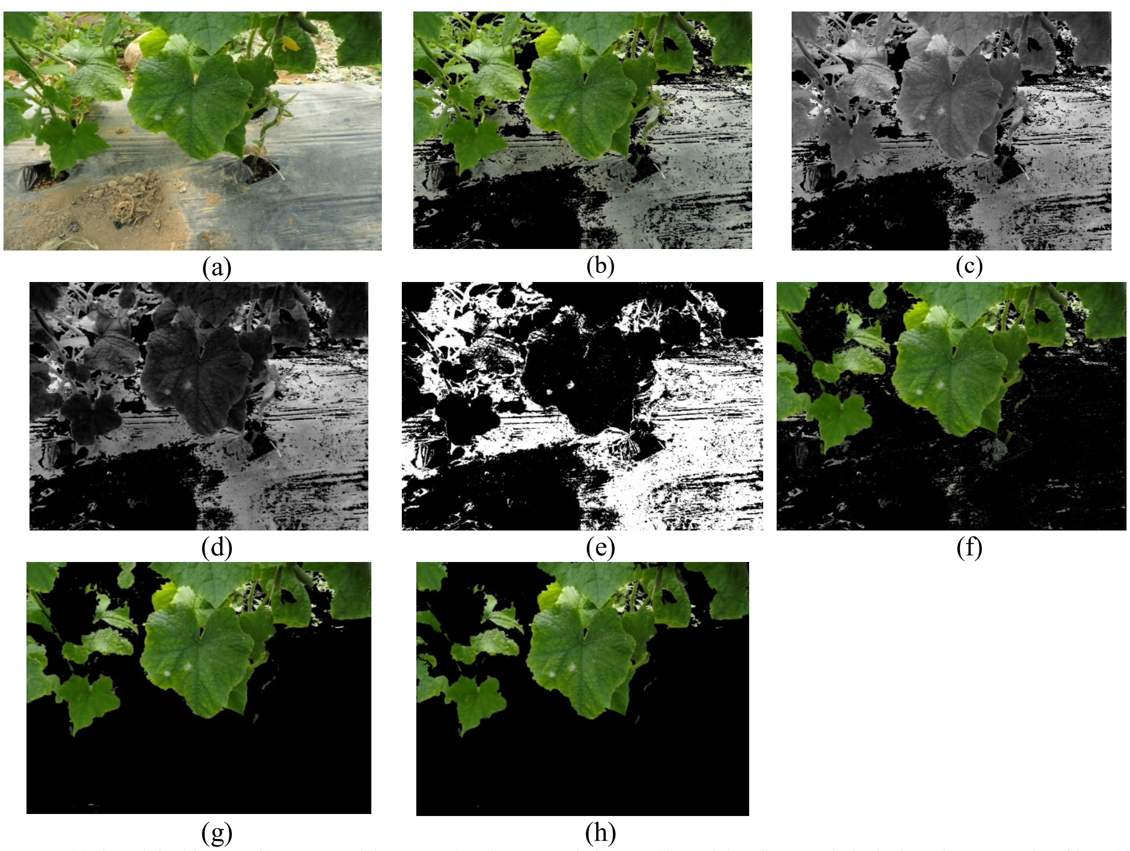

Fig. 1: (a) the original image. (b) HSV Model output. (c) The greyscale image. (d) applying flat morphological top-hat structuring filter. (e) monochrome image (f) after adding noise, filling holes and applying 2-D filter. (g) after using the L*a*b model & adding noise, filling holes and applying 2-D filter. (h) output of the first step.

5. EXPERIMENTAL RESULTS

In order to validate our experimental results, considering that the goal is to automatically identify whether a leaf is infected or not each of the resulting images is used to extract a set of features that will be used later to classify the image, then validated against the human judgment of the infection in the image.

Image features can be summarized as follows: Entropy, mean, standard deviation of the image after the first step of filtering, Entropy, mean, standard deviation of the image after the second step of filtering, 2D correlation coefficient of the resulting image, Mean Squared Error (MSE) of the resulting image, Peak-SNR of the resulting image, and SNR of the resulting image.

Table 1: Confusion matrix

		Actual	
		Inf	Non-Inf
Predicted	Inf	A	B
	Non-Inf	C	D

Table 2: Evaluation methods

Accuracy	(A+D) / (A+B+C+D)
Actual Rate	D / (B+D)
Hit Rate	D / (C+D)

In order to evaluate the image quality for infection classification, we refer to the confusion matrix shown in Table 1 as a primary source for accuracy estimation in classification problems. Based on this confusion matrix, the three-different criterion used are listed in Table 2. Accuracy identifies the percentage of the total number of predictions that were correctly classified where actual rate shows the percentage of predicted infected among actual infected images. Finally, the hit rate shows the percentage of predicted infected in actual infected and actual non-infected.

Several classification algorithms have been tested (i.e. using Weka Machine Learning Toolkit) showing various results. Specifically, Random Forest (RF), Support Vector Machine (SVM), and Instance Based k-nearest neighbor (IBk), and Multilayer Perceptron (MLP) were applied, among others, and reported as shown in Fig 2. All algorithms were run a 10-fold cross validation, and then measures were calculated according the resulting confusion matrix.

The figure shows that at the three evaluation measures IBk and RF performed better than MLP

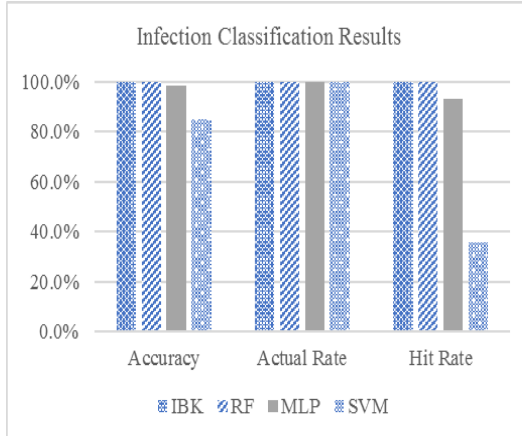

Fig. 2: A classification measures comparison between different classification algorithms

and SVM, which performed poorly in comparison. The results show that reading the previously extracted image features after each processing step of our work helped some classification algorithms to detect infected papers accurately. It is noticeable that SVM detected all non-infected leaves correctly; however, inaccurately identified 9 out of 14 infected leaves as non-infected regardless of the algorithm configurations set in the experiment. With a smaller error, same phenomena appeared with MLP where all non-infected leaves were detected, 13 infected were detected correctly, and 1 non-infected instance was identified as infected.

6. CONCLUSIONS

According to the results obtained, it sounds possible to detect powdery mildew infection in cucumber plants following an image processing approach similarly to what we have achieved. It also sounds interesting that extracting image features after each image processing step helped achieve high infection detection accuracy. Results also show that some classification methods perform better than others considering the dataset class distribution and feature types; nevertheless, few methods performed poorly compared to others. Accordingly, this method can be applied in manufacturing agricultural instruments that can detect the disease on certain leaves, and spray them if necessary.

A possible extension to this work can be identifying the area of the infection, it's location, and the distance from the camera, or an automatic sprayer, can be a challenging, yet valuable, addition to this work.

7. REFERENCES

[1] S. Raza, G. Prince, J. P. Clarkson, and N. M. Rajpoot, "Automatic detection of diseased tomato plants using

thermal and stereo visible light images," PloS one, vol. 10, no. 4, 2015.

[2] A. F., Aji, Q., Munajat, A. P., Pratama, H., Kalamullah, J., Setiyawan, and A. M. Arymurthy, "Detection of palm oil leaf disease with image processing and neural network classification on mobile device," Int'l Jour of Comp Theo and Eng, vol. 5, no. 3, 2013.

[3] P.B. Chopade and K. Bhagyashri "Image processing Based Detection and classification of leaf disease on fruits crops," Int'l Jour of Electri, Electro and Comp Sys, vol. 2, no. 4, 2016.

[4] S. Yun, W. Xianfeng, Z. Shanwen, and Z. Chuanlei, "PNN based crop disease recognition with leaf image features and meteorological data," Int'l Jour of Agr and Bio Eng, vol. 8, no. 4, 2015.

[5] H. H. Kampmann, and O.B. Hansen, "Using colour image analysis for quantitative assessment of powdery mildew on cucumber," Euphytica, vol. 79, no. 1, pp.19-27, 1994.

[6] D. Pixia, and W. Xiangdong, "Recognition of greenhouse cucumber disease based on image processing technology," Open Journal of Applied Sciences, vol. 3, no. 1, 2013.

[7] J. D. Pujari, R. Yakkundimath, and A.S. Byadgi, "Image processing based detection of fungal diseases in plants," Procedia Computer Science, vol. 46, pp.1802-1808, 2015.

[8] Y. Tian, Y. Niu, and T. Li, "A Study on the Method of Image processing and Feature Extraction for Cucumber Diseased," Int'l Conf on Computer and Computing Tech in Agr, Springer, pp.967-972, 2008.

[9] P. Guo, T. Liu, and N. Li, "Design of automatic recognition of cucumber disease image," Info Tech Jour, vol. 13, no. 13, 2014.

[10] Y. Wei, R. Chang, Y. Wang, H. Liu, Y. Du, J. Xu, and L. Yang, "A Study of Image Processing on Identifying Cucumber Disease," Int'l Conf on Computer and Computing Tech in Agr, Springer, pp.201-209, 2011.

[11] B. Zhou, J. Xu, J. Zhao, A. Li, and Q. Xia, "Research on Cucumber Downy Mildew Detection System based on SVM Classification Algorithm," 3rd Int'l Conf on Mat, Mech and Manu Eng, 2015

[12] S. Zhang, and Z. Wang, "Cucumber disease recognition based on Global-Local Singular value decomposition," Neurocomputing, vol. 205, pp.341-348, 2016