

Automatic Recognition of Vegetable Crops Diseases based on Neural Network Classifier

Ismail El Massi
IRF-SIC Laboratory,
Ibn Zohr University
Agadir, Morocco

Youssef Es-saady
IRF-SIC Laboratory,
Ibn Zohr University
Agadir, Morocco

Mostafa El Yassa
IRF-SIC Laboratory,
Ibn Zohr University
Agadir, Morocco

Driss Mammass
IRF-SIC Laboratory,
Ibn Zohr University
Agadir, Morocco

Abdeslam Benazoun
Hassan 2 Institute of
Agronomy and Veterinary,
Complex of Agadir
Agadir, Morocco

ABSTRACT

This paper presents a pattern recognition system for the identification of the vegetable crops diseases from images. The proposed system is based on three main phases: Segmentation, feature extraction and classification. The segmentation of the images is carried out using k-means clustering method. Then, three type of features are extracted from the segmented images including color, texture and shape. These features are used for training and classification using neural networks. The tests of this study are carried out on 300 images of three vegetable crops diseases (Early blight, Late blight and Powdery mildew). The results, with a recognition rate of 95,3 %, show that the proposed system would be interesting to use as means of diagnosis of vegetable crops diseases.

Keywords

Pattern recognition, Classification, Neural network, Vegetable crops, Diseases, Image.

1. INTRODUCTION

Vegetable crops (Tomato, potato...) are among the most important crops that are grown in Morocco. These crops suffer from many parasitic diseases, which can cause the reduction of crop yields. Therefore, significant economic losses can be result to farmers and producers [1].

Nowadays, the methods used for the diagnosis of plant diseases are the traditional methods such as naked eye observation of the experts. However, these methods have many disadvantages; on the one hand, they are expensive, since the need of continuous monitoring of the experts, and on the other hand they make time consuming, because farmers may have to go long distance to contact experts [2].

For this purpose, other fast, accurate and less expensive methods as image processing and pattern recognition methods can be used to design machine vision systems for the diagnosis and detection of plant diseases ([3], [4] and [5]). In recent years, several approaches are proposed for automatic recognition of plant diseases [6]. However, in many of the previous approaches; the authors adopt a small local database of images and also their systems are used for detection of the diseases that attack only one crop.

This study presents a machine vision system for automatic recognition of diseases that attack different vegetable crops as tomato, potato and other. The proposed system is based on the previous approaches with the use of some specific features that are relevant and distinct for the classification performance.

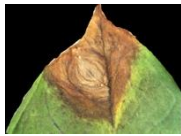





This work focuses on three type of vegetable crops diseases including Early blight *Alternaria solani*, Late blight *Phytophthora infestans* and Powdery mildew *Leveillula taurica*, which are among the major enemies of vegetable crops in Souss Massa region (located in the south of Morocco).

The remainder of this paper is organized as follows: Section 2 describes the proposed system architecture. Section 3 presents the experimental results. Conclusion and future work are given in section 4.

2. PROPOSED SYSTEM DESIGN

The proposed approach in this study aims at designing an automatic recognition system of plant leaves diseases. It is based on three main phases: Preprocessing and segmentation, feature extraction and classification. Figure 1 illustrates the architecture of the proposed system.

Table 1. Some images of the database of the three adopted classes Early blight, Late blight and Powdery mildew

Early blight	Late blight	Powdery mildew
		
		

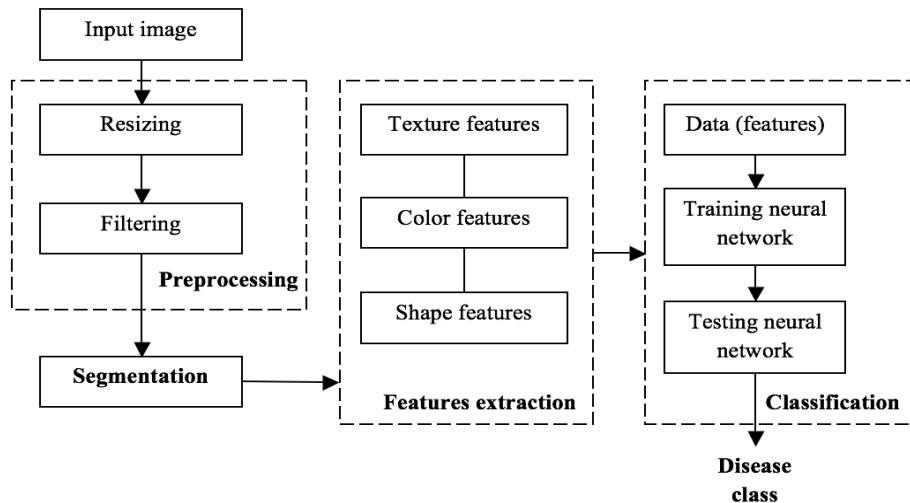


Fig 1: Architecture of proposed system

2.1 Image acquisition

The first step of our system is the collection of the images of infected leaves in order to build the database. The images were captured using digital camera, Canon 600D, in several farms in Morocco with the help of an agricultural expert. Other images are downloaded from Internet to have different environments and to increase our database size. Table 1 shows some images of the database. We demonstrate in this table two examples of each class of the three used classes Early blight, Late blight and Powdery mildew.

2.2 Preprocessing

The input image should be preprocessed in order to improve its quality and to facilitate the next steps (segmentation and analysis) [7].

The adopted preprocessing methods in this work include resizing and filtering. Firstly, the image is resized with a standard size. Then, a filter median is applied as filtering method to reduce the noise in the image. The noise is generally due to acquisition process.

2.3 Segmentation

The segmentation is an important step in automatic recognition of plant leaves diseases [8]. It allows portioning the input image in order to extract the infected area from leaf area.

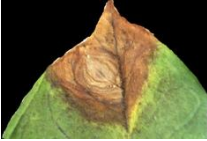
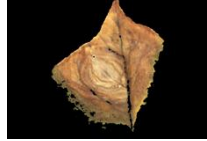


K-means clustering [9] method is used in this context. This method is the most used in the previous works, since it gives good results in segmentation of colored images. In this work, k-means clustering segments the input image into three clusters ($k=3$) in which one cluster contains the lesion area.

In order to do that, the k-means clustering algorithm classifies the objects (pixels in our case) into k number of clusters based on a set of features ($L*a*b$ color space).

The classification of pixels is carried out by minimizing the sum of squares of distances between the data objects and the corresponding cluster. The squared Euclidean distance is used in this context.

Table 3 demonstrates the segmentation results of two images of two infected leaves using k-means clustering method. The segmented image represents the lesion area.

Table 2. Segmentation results using k-means clustering method

Original image	Segmented image
	
	

2.4 Feature extraction

After segmentation, the feature extraction is the next necessary step to carry out. It consists of representing the segmented image on a vector of fixed features. They should be distinct and relevant for the classifier performance. The adopted features, in this study, include color, texture and shape features.

Color moments [10] method is used in this work for color features extraction. This method is adopted because of its ease of use and it provides important results. Color moment method is defined by three moments: Mean, Standard deviation and Skewness. These moments are extracted in this study for each R/G/B component of RGB color space and for each H/S/V component of HSV color space. In total, 18 color features are calculated.

For texture, we adopted Grey Level Co-occurrence Matrix (GLCM) [11] method. This method is the most adopted for texture features in the majority of previous works, since it gives extra information for discrimination between diseases. GLCM is a statistical analysis tool of an image in gray levels, which measures the distribution of gray levels in the image based on the spatial relations of pixels. Haralick [12] introduces 14 attributes of texture based on GLCM. Five attributes were only used in this study including Contrast, Energy, Entropy, Homogeneity and Correlation. These five attributes are the most used in previous works, because they are relevant and distinct. In this work, these attributes are calculated for each component R/G/B of RGB model space

and for each component H/S/V of HSV model space, in which 30 texture features are extracted in total.

Twelve shape features are adopted in this study including Area, Perimeter, Circularity, Complexity, Solidity, Extent, Major axis length, Minor axis length, Eccentricity, Centroid and Diameter. We associate in this work the most used shape attributes in previous works, and then the results are significant.

2.5 Classification

The classification phase consists of assigning a disease form in the input image to one class of the three output classes Early blight, Late blight and Powdery mildew. In this study, the back propagation neural network classifier [13] is adopted for learning and classification. It consists on tree layers (Figure 2): an input layer (Inputs), a hidden layer and an output layer (Outputs). Neurons of each layer are connected completely and directly to the nodes of its adjacent layer. The layers of the neural network classifier in this work are defined as:

- **Input layer:** Consists of sixty (60) calculated features including 18 color, 30 texture and 12 shape.
- **Hidden layer:** Consists of 10 nodes. This number gives best results.
- **Output layer:** Consists of the three output classes: Early blight, Late blight and Powdery mildew.

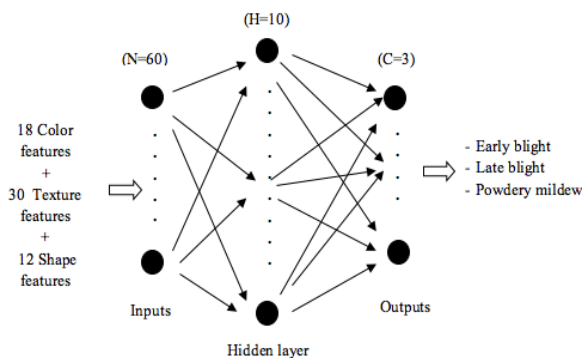


Fig 2: Architecture of neural network classifier

3. EXPERIMENTAL RESULTS

In order to test and evaluate the proposed system, we have built a database of 300 images including 100 images in each class of the three adopted classes: Early blight, late blight and powdery mildew.

In this experiment, the dataset is divided in two subsets: a set of 210 images (70%) used for training and a set of 82 images (30%) for test.

The tool nntool of Matlab 2015a is used in this work to train and test the neural network classifier. Figure 3 demonstrates the confusion matrix of the proposed System. The global recognition rate is 95,3 %. From 300 images, 286 are well classified. The powdery mildew class is the most classified class with the higher recognition rate (99 %). The Late blight class has the less recognition rate (92 %).

Output Class \ Target Class	1	2	3	Total
1	95 31.7%	4 1.3%	0 0.0%	99 33.0%
2	5 1.7%	92 30.7%	1 0.3%	98 32.7%
3	0 0.0%	4 1.3%	99 33.0%	103 34.3%
Total	95.0% 5.0%	92.0% 8.0%	99.0% 1.0%	95.3% 4.7%

Fig 3: Confusion matrix of the proposed system

The results of the proposed system are significant and encouraging. The examination of Figure 3 indicates that the Late blight class, with a rate of 92%, cause the diminution of the global recognition rate of the proposed system. This rate is due to the complexity of this class, in which it is difficult to discriminate between forms of Late blight class. The symptoms of Late blight are characterized by different characteristics (color, texture and shape) in different stages of development. Figure 4 shows two forms of Late blight symptoms in two different stages of development. It appears, through the images, the difference, in the characteristics, between the two forms of Late blight.



Fig 4: Difference, in characteristics, between forms of Late blight class. (a) and (b): Two forms of Late blight symptoms

Also, the classification errors are due to the the similarity in the color between Early blight and Late blight. Some forms of these two classes have a brown color which can cause problems in classification. Figure 5 demonstrates examples of images that represent the symptoms of two classes Early blight and Late blight. We see through this figure the similarity in the color between the two classes.



Fig 5: Similarity in the color between symptoms of the two classes Early blight and Late blight. (a) Early blight symptoms. (b) Late blight symptoms

The implementation of the system is illustrated in Figure 6. This figure demonstrates an example of Powdery mildew symptoms that are well recognized by the proposed system.

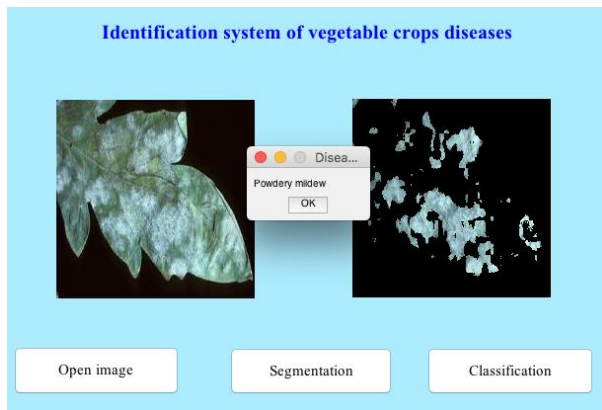


Fig 6: Powdery mildew symptoms are well recognized by the proposed system

4. CONCLUSION AND FUTURE WORK

Automatic recognition of vegetable crops diseases on plant leaves is the aim of this study. The proposed system is based on neural network classifier and the combination of three types of features including color, texture and shape. Thus, the approach is tested on three important vegetable crops diseases which are Early blight, Late blight and Powdery mildew. The results of this study indicate that the proposed system would be of importance to use as a means of diagnosis and phytosanitary problem recognition from images.

In future work, we plan to improve the proposed system in order to reduce the classification errors, especially the errors that is caused by the Late blight class, and therefore achieving a more efficient pattern recognition system of vegetable crops diseases.

5. ACKNOWLEDGMENTS

Our thanks to the experts in agricultural field Abdellah Mofrir and Abdelaziz Hsissou, who have contributed in the elaboration of this work

6. REFERENCES

- [1] Sankaran, S., Mishra, A., Ehsani, R., and Davis, C., A review of advanced techniques for detecting plant diseases. *Computers and Electronics in Agriculture*, vol. 72, 2010, pp. 1-13.
- [2] Al Bashish, D., Braik, M. and Bani-Ahmad, S., Detection and classification of leaf diseases using k-mans clustering based segmentation and neural networks based classification. *Information Technology Journal*, 2011, pp. 267 -275.
- [3] Sannakki, S.S., Rajpurohit, V.S., Nargund, V.B. and Kulkarni, P., Diagnosis and Classification of Grape Leaf Diseases using Neural Network. In 4th IEEE International Conference on Computing,

Communications and Networking Technologies (ICCCNT), 2013, pp. 1-5

- [4] El Massi, I., Es-Saady, Y., Elyassa, M., Benazoun, A., and D. Mammass, Automatic recognition of the damages and symptoms on plant leaves using parallel combination of two classifiers. 13th IEEE International Conference Computer Graphics, Imaging and Visualization (CGIV), 2016, pp. 131-136.
- [5] Es-Saady, Y., EL Massi, I., El Yassa, M., Benazoun, A., and D. Mammass, Automatic recognition of the damages and symptoms on plant leaves using parallel combination of two classifiers. 2nd IEEE Electrical and Information Technologies ICEIT'2016, Tangier, Morocco.
- [6] Barbedo, J. 2013. Digital image processing techniques for detecting, quantifying and classifying plant diseases. In Springer plus 2, (2013) 1–12.
- [7] Geng, Y., Miao, L., Yuan, Y., and Zelin, H., A Study on the Method of Image Pre-Processing for Recognition of Crop Diseases. In International Conference on Advanced Computer Control, 2008, pp. 202-206.
- [8] Camargo, A., and Smith, J., An image-processing based algorithm to automatically identify plant disease visual symptoms. *Biosystems Engineering*, Vol. 102, 2008, pp. 9–21.
- [9] MacQueenn, J. B., Some methods for classification and analysis of multivariate observations. *Proceedings of 5th Berkeley Symposium on Mathematical Statistics and Probability*, 1967, pp. 281-297.
- [10] Patil, J. K., and Kumar, R., Color feature extraction of tomato leaf diseases. *International Journal of Engineering Trends and Technology*, Vol. 2, 2011, pp.72-74.
- [11] Al Bashish, D., Braik, M., and Bani-Ahmad, S., A Framework for Detection and Classification of Plant Leaf and Stem Diseases. In IEEE International Conference on Signal and Image Processing, 2010, pp. 113 -118.
- [12] Haralick, R.M., Shanmugan, K., and Dinstein, I., Textural Features for Image Classification. In IEEE Transactions on Systems, Man, and Cybernetics, Vol. 3, 1973, pp. 610-621.
- [13] Basheer, I.A., and M. Hajmeer, Artificial neural networks: fundamentals, computing, design, and application. *Journal of Microbiological Methods*, 2000, pp. 1-31.