Al-Mustansiriyah Journal of Science
ISSN: 1814-635X (print), ISSN:2521-3520 (online) Volume 32, Issue 5, 2021 DOI: http://doi.org/10.23851/mjs.v32i5.1115

Research Article Open Access

Using Texture Analyses and Statistical Classification for Detection Plant Leaf Diseases

Basim K. M. A. Al-Windi[1,*], Amel H. Abbas[1], Mohammed Shakir Mahmood[2]

[1] Department of Computer Science, College of Science, Mustansiriyah University, Baghdad, IRAQ.
[2] Ministry of Higher Education and Scientific Research, Scholarships and Cultural Relations Directorate, Baghdad, IRAQ.

[*]Correspondent contact: gohn_smith2002@yahoo.com

Article Info

Received
05/07/2021

Accepted
07/08/2021

Published
15/12/2021

ABSTRACT
The proposed method is based on classifying 15 types of plant leaf disease. Hue saturation value was used to delete the background and the healthy areas to show only the affected area in the image. Texture analyses adopted in image features extractions from the R component &G component &B component and creating 3 components which are RG and RB and GB color of the RGB color space images of diseased leaves. Building image classifier using statistical method for classification.

KEYWORDS: GLCM; Texture Analyses; HSV (Hue Saturation Value); Statistical Classification.

INTRODUCTION
Diseases which infect leaves of plants causes color or shape changes. These problems being solved by digital processing of images of plant leaves obtained using unmanned aerial vehicles with an RGB camera and automatic classification tools. Texture analyses are proposed for identifying plant leaf diseases GLCM algorithm providing good result in extraction features for object of study, infected leaf image. Building a classifier that is comfortable with our dataset using Statistical methods.In our method we used 15 plant leaf diseased like 1-Septoria Leaf Spot (Septoria), 2-pirenoforoz (Pyrenophora tritici-repentis),3-Powdery mildew (Erysiphe graminis), 4-brown rust (Puccinia recondita),5, 6 -yellow rust (Puccinia striiformis),7-leaves Septoria Leaf Spot (Septoria tritici), 8 - Snow mold (Fusarium nivale), 9-blotch (*Helminthosporium sativum)*10 - Root rot 11 - Stripe Mosaic (Wheat streake mosaic virus), 12 - Brown (sheet), rust (Fungal diseases (Puccinia triticina)),13-blotch (Pyrenophora tritici-repentis), 14 - Linear (stem) rust diseases (Puccinia triticina)),13-blotch (Pyrenophora tritici-repentis), 14 - Linear (stem) rust (Puccinia graminis), 15 - Head smut (Ustilago tritica)

Literature review
In 2020, S. Bharath, K. Vishal Kumar, R. Pavithran, T. Malathi [1] In their project of identifying plant disease. CNN have been used as feature extractor and classifier, their viewpoint was that using CNN with exact amount of neurons in training and testing could give awesome result because of CNN ability as a texture analyzer and color analyzer. A database of trained image and test where calculated for few plants which are Apple, Grape, and Tomato and others totally 38 kind. They made class for each kind of disease. A module has been done for the classes 38 then the teats show acceptable result.

Healthy Leaves
In 2013, Dong Pixia and Wang Xiangdong,[2]. Suggest method of detections of anthracnose & powdery mildew and downy mildew. They depend on their project greyscale image after conversion from RGB, Median filter & image smoothing used for enhancement. In feature extraction two method are proposed, color extraction is calculated by extracting the value of the pixels for infected area and extracting and average value for the same area

and final feature area. Texture feature used too by using GLCM. Neural network used for classifications.

In 2016, Ms. Poojapawer, Dr. varsha Tukar, prof. parvinpatil [3] proposed method includes image acquisition, image preprocessing, feature extraction with gray level co-occurrence matrix (GLCM) and finally classified with two types: Unsupervised classification and supervised classification.

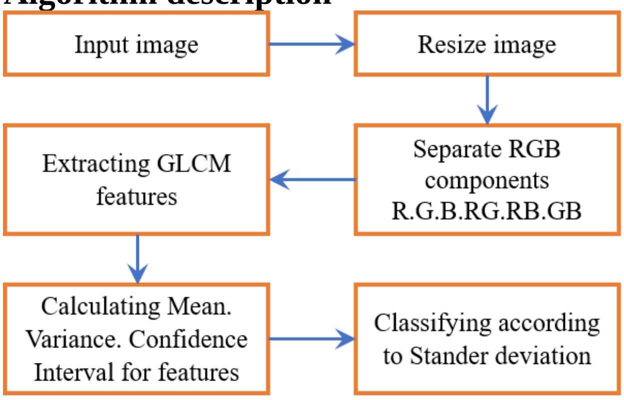
In 2015, Manisha Bhange and H. A. Hingoliwala, [4] Their method of identifying plant disease. In preprocessing step feature extraction done using three ways color extraction second morphology, and CCV. k-means clustering algorithm used, Support vector machine used in classification. In this experiment best result achieved by using morphology. System accuracy reached to 82%

In 2015, Ashwini A, Damini D, Gayatri A, Utkarsha B and Samadhan S, [5] proposed a system using OpenCV library. K-means clustering used in segmentation step. four different methods used in feature extraction (Texture, color, mereology, other). ANN used as classifier. their idea was using all the picture without deleting unnecessary parts like background or healthy region. The system falls to reach dependable detection accuracy.

PROPOSED METHOD

Proposed method consists of special image processing method which is separation of image components R alone, G so as B, then creating 3 additional [6] components by multiplying image components with each other's RG and RB and G with B so from single image or (matrix) we create six matrixes. The aim of this method to get and calculate more features (Hidden) which could not be calculated from image in its original architecture.

Algorithm description



Preprocessing

Is image preparation [6] and rehabilitation to be useful in gathering and extracting information from it, this done by adjusting size, enhancements, noise removing, etc....

Texture analyses

Texture is a feature used to partition images into regions of interest and to classify those regions. And provides information in the spatial arrangement of coolers or intensities in an image. Texture analyses is characterized by the spatial distribution of intensity levels in a neighborhood.

Hue Saturation Value

First step was deleting background using HUE Saturation Value HSV to eliminate any unnecessary information in image.

Grey Level Co-Occurrence Matrix.

Haralick grey level co-occurrence matrix is a feature extractor is doing the arrangements of pairs of voxels for calculation indices of texture. The GLCM is calculated from different directions in 3D with a δδ-voxel distance ("d"−→"d"→)
Relationship between voxels (neighbored voxels). Value of index is the average of the index over the space directions in (X, Y, Z). We used Haralick because it calculates Mean, variance for the features that is needed in classification step.

GLCM Features

contrast:

$$CN = \frac{1}{(G-1)^2}\sum_{u=0}^{G-1}\sum_{v=0}^{G-1}|u-v|^2\, p(u,v);$$

correlation:

$$CR = \frac{1}{2}\sum_{u=0}^{G-1}\sum_{v=0}^{G-1}\frac{(u-\mu_u)(u-\mu_v)}{\sigma_u^2\sigma_v^2}p(u,v)+1;$$

energy:

$$EN = \sum_{u=0}^{G-1}\sum_{v=0}^{G-1}p(u,v)^2\,;$$

homogeneity:

$$HM = \sum_{u=0}^{G-1}\sum_{v=0}^{G-1}\frac{p(u,v)}{1+|u-v|}.$$

Ent:

$$f_5 = -\sum_{i}\sum_{j}p(i,j)\log\big(p(i,j)\big)\,;$$

Energy:

ISSN: 1814-635X (print), ISSN:2521-3520 (online) Volume 32, Issue 5, 2021 DOI: http://doi.org/10.23851/mjs.v32i5.1115

$$f_6 = \sum_i \sum_j \frac{1}{1+(i-j)^2} p(i,j).$$

Figure 1. Various diseases of Leaves images.

Classification is the final step in any module for detecting and recognize the goal of the module, it depends on the training set which has been given before and the testing set which must have the same characteristics as the training set. The most common classifiers are Neural network, support vector machine SVM, Local binary patterns LBP. In our method, we didn't use a traditional classifier. We have built our classifier depending on the source of all classifiers that is statistics. By calculating Mean and variance then standard deviation to calculate the confidence interval for each disease to be the database or the definition for the disease, confidence interval job is same as image features.

Confidence Interval

The confidence interval (CA) is a range of values [7] that's likely to include a population value with a certain degree of confidence. It is often expressed as a % whereby a population means lies between an upper and lower interval. Figure 1 shows R component testing for 15 diseases.

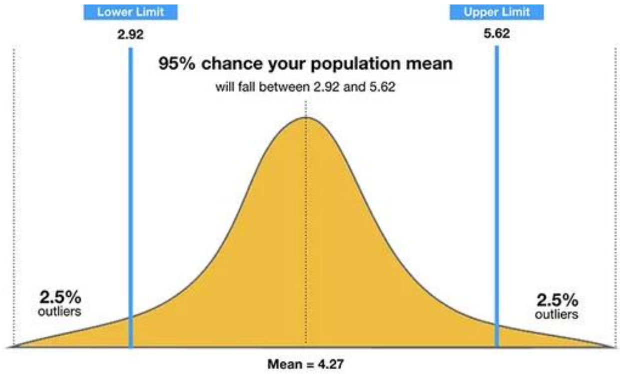


Figure 2. Confidence Interval.

RESULTS AND DISCUSSION

Project program built by python, RGB separation experiment success with detection and classification of the diseases; the minimum amount of images in training phase for every single disease is 30 images that is required to calculate the confidence interval and the standard deviation. Table 1 shows the testing results of one disease according to the reference descriptions (feature map) of the trained diseases strictly using one component (i.g., R) component. Therefore, to get the exact result we must test the entire six component that we generated.

The diagnostic results represent 1, if the value falls within the confidence interval range or, 0 otherwise.

In table 2, the result illustrates exactly the desired class of the trained image of our dataset. In this case, we applied the six component (i.g., RB and GB, and R….B) with confidence level of 0.95. The identification result reaches 96%.

For testing, Table 2 shows that the tested disease belongs to the 4th trained disease. The latter indicates that the values of all six components in fourth disease are 1.

Table 1. Identification result using R component

No of diseases	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Grey level co-occurrences matrix															
Cont.	1	1	0	1	0	0	0	0	0	0	0	0	0	0	0
Corr.	0	1	1	1	0	1	1	1	0	0	1	1	1	0	1

Ener.	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0
Homo.	1	1	0	1	0	1	0	1	0	0	0	1	0	0	0
Entr.	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
Inverse Difference Moment	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
results after voting	4	5	3	6	1	3	2	3	1	1	2	3	2	1	2
Final results	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0

Table 2. Identification result using R, G, B, RG, RB, GB component.

No of diseases	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
GLCM															
R	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0
G	1	0	0	1	0	0	0	1	0	0	0	0	0	0	0
B	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
RG	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
RB	1	1	0	1	0	1	1	1	1	0	0	0	1	0	0
GB	0	1	1	1	0	0	0	1	0	0	0	1	1	0	1
Final result	2	3	1	6	0	1	1	3	1	0	0	1	2	0	1
Binarization results	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0

CONCLUSIONS

Using statistical method to build special classifier proved good accuracy in classification with small amount of dataset. Using one component from RGB components in training and testing falls to provide enough features for building identification system, we must depend on all components. Using 6 components 3 from RGB and RG RB GB provide grate information from picture but it takes long time in training & testing. To calculate confidence interval for the disease, minimum 35 pictures must be used, this applied on test to. In future work we aim to build a classifier using depend on statistical method (confidence interval) can do the test using one image not 35 as our system

ACKNOWLEDGMENTS

Special Acknowledgments for Diyala Department of Agriculture for their help in providing information of infected area, image diseases (Dataset) classification.

REFERENCES

[1] S. Bharath, K.V. Kumar, R. Pavithran, and T. Malathi, "Detection of Plant Leaf Diseases using CNN", International Research Journal of Engineering and Technology (IRJET), 4Computer Science and Engineering, SRM Institute of Science and Technology, Chennai. India, 2020.

[2] D. Pixia, W. Xiangdong, "Recognition of Greenhouse Cucumber Disease Based on Image processing Technology", Open Journal of Applied Sciences, vol. 3, pp.27-3, Mar. 2013.

[3] P. Pawer, V. Tukar, Parvinpatil, "Cucumber Disease detection using artificial neural network", International Conference on Inventive Computation Technologies (ICICT) 2016.

[4] M. Bhange and H. A. Hingoliwala, "Smart Farming: Pomegranate Disease Detection Using Image Processing", proc. of the Second International Symposium on Computer Vision and the Internet (VisionNet‘15), pp. 280-286, 2015.

[5] A. Awate, D. Deshmankar, G. Amrutkar, U. Bagul and S. Sonavane, "Fruit Disease Detection using Color Texture Analysis and ANN", proc. of International Conference on Green Computing and Internet of Things (ICGCIoT), pp. 970-975, 2015.

[6] V. S. Tutygin, B. K. M. Ali Al-Windi, "A system of classification of texture images based on fuzzy logic”, Modern science: topical problems of theory and practice. Series Natural and Technical Sciences. No. 3, 2019.

[7] K.-Man Xu, "Using the Bootstrap Method for a Statistical Significance Test of Differences between Summary Histograms", Monthly Weather Review, 134(5), 1442–1453. doi:10.1175/mwr3133.1 2006