International Journal of Recent Technology and Engineering (IJRTE)
ISSN: 2277-3878, Volume-8 Issue-3, September 2019

Diagnosis of Leaf Disease in Cucurbita Gourd Family using Machine Learning Algorithms

V. Nirmala, B. Gomathy

Abstract: ***PaperThe paper presents the implementation of various machines learning approach for the diagnosis of leaf diseases. For analysis data collection were done towards capturing the images of pumpkin leaf affected by different diseases. The pumpkin leaf samples were taken. The samples correspond to the blight and fungal problems like Alternaria, Powdery Mildew Anthracnose, and Yellow Vine Disease etc. The methods for analysis were implemented and tested which are based on time, frequency and statistical approach. For classification machine learning approaches like neural networks, SVM, KNN etc were analyzed. The implementation issues were presented for future work.***

Keywords : Leaf diseases, blight, yellow vine, neural network, wavelet transform, SVM, classifier, denoising.

I. INTRODUCTION

Botanists study the plants by recognizing the leaves and flowers with various characteristics of comparative tool. This comparative tool can recognize the plant species by using computer vision in leaf image. The features is represented by algorithms like deep learning and hand craft methods . These features are crucial component (Cope et al, 2012) [1]. In this two methods hand craft features is more dependent than deep learning features (Hall et al, 2015) [2]. Some of the characteristics are pre defined by botanists and however they are used as morphological features. In deep learning algorithm, the features can be automatically learning their basic end-to-end advantages of an algorithm. Hence the deep learning method is carried out more often. In figure 1 shows the different size and species and Fig. 1 (a) (b) shows the different growth in same species of plant. The Abies Concolor leaf is the example of leaf detection which is smaller than robur leaf. Some plant leaves has different suitable scale. Juan Zhang and Xin-yuan Huang, (2010) [3] analyzed the single scale features which can learn deeply based on plant leaf recognition method. This is mainly focused for improving the accuracy on existing feature.

Manuscript published on 30 September 2019
* Correspondence Author
V. Nirmala*, Assistant Professor, Department of ECE, M. Kumarasamy College of Engineering, Karur, India.
Email: nirmalavelusamyphd@gmail.com
B. Gomathy, Associate Professor, Department of CSE, Bannari Amman Institute of Technology, Sathyamangalam, Tamilnadu, India.

The leaf recognition method is presented with several deep learning features. Timothy et al (2016) [4] presented the convolution neural network (CNN) for classifying the leaf image to extract some features using support vector machine (SVM). Guillermo et al 2016 [5] presented the UHMT by vein pattern to train the CCN on first segment using original input images. Lee et al 2006 [6] used deep leaning algorithms to recognize the plant disease from the features of CNN model. They used two stream convolution neural network (TwoCNN). The whole and patch images are learned as the two stream structure. This can process as different scales to capture discriminative information. During training progress more samples can be collected and the networks are tested.

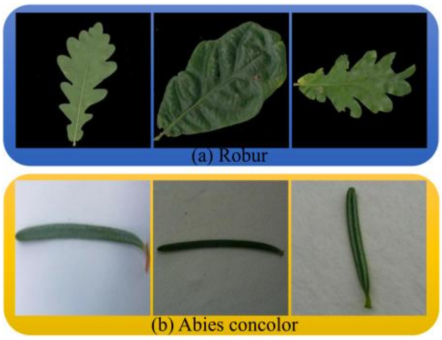

Fig. 1. (a) & (b) Different growth in same species of plants.

Rasti et al 2018 [7] demonstrated the multi-scale convolutional neural networks (MSCNNs) for different scale features for learning branches by different size images. In work space, different patches and different inputs are tested. The MSCNN is learned as the multi scale features which are fused at the layer in step by step process. The segment is done by studying the computer graphics and vision that is mainly focused on 3D images as reconstruction on plants. Photorealistic plant image is the model that appears on leaf shape which has the perfect example of this method. Some of the image has the multi image features that are captured with hand-held camera. This method can produce the photorealistic plant image as the data driven synthesis which is reconstructed as motion structure as 3D point cloud of dense foliage. Li et al, addresses three tasks of leaf SAT which is forward and backward analysis using 3D acquisition with the specific data for model plant growth. Hence the method is validated by the dataset as synthetic all over the specific condition to adapt at large scale experiments in computer vision works. The leaf segment is aligned, tracked and identified without being studied on one are two tasks.

Leaf Segmentation is classified into two main categories such as detached leaf and pixel wise segmentation. The first step is to identify the species and color which is used as pixel based leaf segments with white background. This process is very challenging due to the variation in leafs structure and also overlapping under the background. Tsaftaris et al 2016 [8] investigated a leaf segmentation that is collation of rosette plant. This can be evaluated in three methods. The methods are based on pixels and watershed transformation segmenting methodology. Leaf Alignment is used to find out the structure of leaf that is used for processing the leaf segment. This structure is useful for fitting the leaf shape template in deform polygonal model. This segmentation is tracked automatically. It's measured at some temperature (G. Romay et al, 2014) [9]. The leaves without any overlap on an active contour model are used. This method is achieved by real world application that is presented as the graphical based structure which is exploited in certain angle. Rong et al. (2007) [10] introduced the support vector machine (SVM) using hybrid algorithm for template matching. The SVM uses hyper spectral AISA-Dual data to overcome the disease sensitivity caused by the leaf rust when the data is driven Mewes et al. (2010) [11].

A. Problem statement

Diagnosis of leaf diseases were provided mainly based on visual observations. Growers guide based on soft computing were not available. Some of the control products are only listed. The data and consultancy with experts in certain areas are to be included. An online learning and diagnosis system is not available. Many, many hours were spent to rectify the causes. Sometimes blight can affect the whole farm within hours.

B. Objectives

- Intention to provide a growers guide based on soft computing.
- Control products are listed as a suggestion through a new soft computing based growing methodology with real time pictures.
- Immediate diagnostic steps are to be provided based on the machine learning approaches.

C. Disease of interest in Cucurbita Gourd family

The following diseases are diagnosed in several leaves. But our interest is on the Cucurbita Gourd family. Advanced disease progression, Alternaria, Aphids, Anthracnose, Downy Mildew, Excessive Nitrogen, Gummy Stem Blight, Insect damage, K-def., Mn defficency , Ozone damage. Phytoptora, Powdery mildew, Sunburn, Spray damage, Virus, Yellow Vine Disease Verticilum wilt. Fig. 2 shows some of the major diseases affecting the cucurbita Gourd family.

Fig. 2. Different disease affected leaves of pumpkin plant.

II. BACKGROUND METHODOLOGY

Arti N. Rathod et al (2013) [12] in literature reported the different types of disease. The methods are used to identify the affected area in leaf. The digital camera is used /to take images. The taken image is processed by different types of image-processing techniques. These methods analyze the useful features which are needed to diagnose. Fig. 3 shows generalized block diagram for the leaf disease detection.

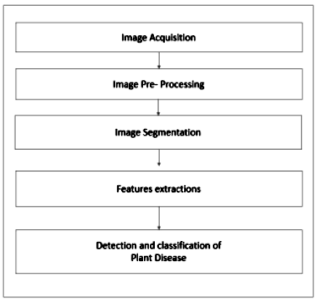

Fig. 3. Generalized Block Diagram for leaf disease detection.

The input image can be acquired by image acquisition technique. The image is taken by digital camera. This will help to improve the quality of image and also to remove the undesired distortion. Isaac Kofi Nti et al, (2017) [13] explains the performance of clipping on leaf image to get the interested image region. Then smoothing is is done by smoothing filter. By this the contrast can be increased. The pixels are computed mostly on green colored pixels that are masked by using a threshold value. Hence the green pixels are masked because it has intensity of green component which is less than the pre-computed threshold value. Then zero value is assigned to the red, green and blue components. The leaf disease is classified using genetic algorithm. Weihua Gui et al, (2013) [14] processed the color co-occurrence for feature extraction. The three major processes are taken to process the detection function which has both the texture and color of an image. The unique features of the leaf is identified. Vijai Singh and A.K.Misra, (2017) [15] represented the RGB image conversion in leaves. The leaf image is processed in the HIS color space to generate the color co-occurrence in matrix. Each of the representation will be processed and compared with the phase of leaf that can be stored on the database.

A. Quality Check

All submitted paper should be cutting edge, result oriented, original paper and under the scope of the journal that should belong to the engineering and technology area. In the paper title, there should not be word 'Overview/brief/ Introduction, Review, Case study/ Study, Survey, Approach, Comparative, Analysis, Comparative Investigation, Investigation'.

III. PUMPKIN LEAF DISEASE DETECTION

A. Image Pre-processing

The pre-processing technique is carried out with different ways to remove the noise in image. Image clipping is one of the preprocessing techniques used to get the interested image region by cropping. Image smoothing is next preprocessing done. Image enhancement is used to increase the contrast. The corresponding RGB image into the grey images using color conversion is given in (1).

Retrieval Number: C5232098319/19©BEIESP
DOI:10.35940/ijrte.C5232.098319
Journal Website: www.ijrte.org

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

$$F(x) = 0.2989*R + 0.5870*G + 0.114.*B \qquad (1)$$

The above equation is used to enhance the plant image and then the histogram equalization is applied to distribute the intensities of the images. P.R. Rothe, and R.V. Kshirsaga, (2012) [16] presented the cumulative distribution function which is used to distribute intensity values and the segmentation can do the partitioning of image into various part of same features like otsu' method, k-means clustering, converting RGB image into HIS model etc.

- ***K-means clustering:*** It is the cumulative distribution that can process the image by partitioning by using classification of object. The classification is based on clustering a set of features into K number of classes. Sonal P Patel and Arun Kumar Dewangan, (2017) [17] uses the classification of object to minimize the sum of the squares of the distance between the object and the corresponding cluster. The algorithm has the pick center of K cluster used randomly on heuristic and also they can assign the pixel of minimized image to find out the distance between the pixel and cluster. In the algorithm, both the process can be averaged with all random clusters by repeating the steps to get the desire convergence.
- ***Otsu Threshold Algorithm:*** This algorithm will process the threshold value that can create binary images from grey-level images by setting all pixels threshold to zero. The Otsu algorithm defines the two clusters by separating their pixels using threshold and it can find the mean by each and every cluster between various square mean and also it can multiply the cluster time. Amruta B. Patil and J.A.shaikh, (2016) [18] presented the masked image of green pixel by calculating the threshold values using otsu algorithm. It can be extracted by RGB component which are infected leaf that is detected.

B. Feature Extraction by GLCM approach

P. Mohanaiah et al (2013) [19] presented Gray level co-occurrence matrix (GLCM) to determine the textural relationship between pixels by performing an operation. This method is based on the feature extraction that is at the second-order statistics in the images. The GLCM will also process the frequency combination of pixels pair that is usually two pixels to brightness the values. The number of rows and columns are with the same gray values in the image but the elements can depend on the frequency of the two specified pixels. Thus the pixel pairs can vary depending on their neighborhood by the second-order statistical probability values. These values are depending on the gray value of the rows and columns.

Autocorrelation, contrast, correlation, cluster prominence, cluster shade, dissimilarity, energy, entropy, homogeneity, maximum probability, sum of squares (variance), sum average, sum variance, sum entropy, difference variance, difference variance, difference entropy, information measure of correlation, inverse difference normalized, inverse difference moment normalized are some of the features computed by the GLCM approach.

C. Classification using Multi-SVM approach

Support Vector machine (SVM) use hyper-plane on multiple way of decision boundary between two of the classes (Gjorgji Madzarov and Dejan Gjorgjevikj, (2009) [20]. The classifier is commonly known as binary classifier. The multi-SVM approach solves some of the problems of pattern recognition that are mapped of nonlinear input data to the linear data. This method provides a classification in high dimensional space with the marginal distance which is used to maximize between different classes. The hyper plane can uses the two classes that are different in kernels. The SVM is uses larger number of support vectors related to training samples that can intended for two-class problems. Hence the mathematical expression separates more than two classes at the same time that can be extended from binary problems to multi classification problems (with k classes where $k > 2$). There are $k\ (k-1)/2$ decision functions for the k-class problem.

IV. RESULTS AND DISCUSSIONS

A. Pre-processing of Pumpkin Leaves

The preprocessing of different pumpkin leaves is done by k-means clustering algorithm, otsu thresholding algorithm. The contrast enhanced pumpkin leaf images are shown in Fig. (4) - (8). The leaves used for the analysis in this work are of five category namely healthy leaf, Downy mildew and Powdery mildew. In preprocessing most of the noise elements are removed.

Fig. 4. Pumpkin Leaf 1 and its contrast enhanced image.

Fig. 5. Pumpkin Leaf 2 and its contrast enhanced image.

Fig. 6. Pumpkin Leaf 3 and its contrast enhanced image.

Fig. 7. Pumpkin Leaf 4 and its contrast enhanced image.

Fig. 8. Pumpkin Leaf 5 and its contrast enhanced image.

B. Features of Pumpkin Leaves

In the feature extraction, several features are extracted like contrast, energy, homogeneity. The features show the characteristics of the various diseases presented in the leaf. The features are used to classify the different categories.

Table- I: Feature Extraction of Pumpkin leaves using GLCM

Features	*Leaf 1*	*Leaf 2*	*Leaf 3*	*Leaf 4*	*Leaf 5*
Contrast	2.5599	2.1231	1.68615	1.415	2.548
Correlation	0.9408	0.94316	0.94787	0.9448	0.956
Energy	0.5276	0.42944	0.60427	0.8238	0.580
Homogeneity	0.9435	0.92691	0.94923	1.0028	0.946
Mean	52.667	60.0387	37.5518	23.22	55.35
Standard Deviation	80.695	76.5521	65.5495	58.868	89.18
Entropy	3.6037	4.33258	3.22383	2.3978	3.474
RMS	9.5749	11.6751	9.31657	8.0565	10.07
Variance	5010	6384.27	4548.35	4276.8	8324
Smoothness	1.0000	1.0000	1.0000	1.0000	1.000
Kurtosis	3.1089	2.42828	4.45938	10.198	3.068 2
Skewness	1.2318	0.84422	1.65472	2.9366	1.275
IDM	255	255	255	255	255

Table. 1 shows the feature extracted information from diseased pumpkin leaves using GLCM method. These features may vary according to the time of capturing the image, the disease concentration on the leaf and age of the plant. So the feature database should contain all the information corresponding to these details. Data collection is made real time. In this work several images were captured using the digital camera with 8mega pixels. The images are analyzed using Matlab. The implementation was done and testing of the images are carried out. From the several observation it is been found that the disease if not diagnosed properly it will damage the whole crop or even the whole field.

C. Classification of Features of Pumpkin Leaves

Table. II. Shows the classification of diseased pumpkin leaves. The accuracy of the methods for different diseased leaves are presented.

Table- II: Classification of Pumpkin leaves by multi SVM technique.

Leaf	*Affected Disease*	*Affected Area (%)*	*Accuracy (%)*
Leaf 1	Healthy Leaf	--	93.7742
Leaf 2	Downy Mildew	35.7875	92.7742
Leaf 3	Healthy Leaf	--	91.3871
Leaf 4	Powdery Mildew	19.7253	92.3871
Leaf 5	Powdery Mildew	17.4835	93.7742

V. CONCLUSION

The paper presents the implementation of various diseases in pumpkin leaves using GLCM methods. The diagnosis of leaf diseases helps the farmers to do early detection of crop damage. For the image analysis, data collection was done and selective diseased pumpkin leaf samples were taken. The leaf categories tested are healthy leaf, downy mildew and powdery mildew.

REFERENCES

1. James S.Cope, DavidCorney, Jonathan Y.Clark, Paolo Remagnino and PaulWilkin, "Plant species identification using digital morphometrics: A review," *Expert Systems with Applications*, vol. 39, no. 08, 2012, pp. 7562-7573.Available: https://www.sciencedirect.com/science/article/pii/S0957417412000863
2. Peter Hall, Hongping Cai, Qi Wu and Tadeo Corradi, "Cross-depiction problem: Recognition and synthesis of photographs and artwork," *Computational Visual Media*, vol. 01, no. 02, 2015, pp. 91-103.
3. Juan Zhang and Xin-yuan Huang, "A Survey of Plant Recognition Method Based on Image Processing;" *International Conference on Electrical and Control Engineering*, 2010, pp. 4974-4977. Available: https://link.springer.com/content/pdf/10.1007/s41095-015-0017-1.pdf
4. Timothy, B.L., Thege, F.I. and Kirby, B.J, "Comparison and optimization of machine learning methods for automated classification of circulating tumor cells," *Cytometry A*, vol. 89, no. 10, 2016, pp. 922-931. https://www.ncbi.nlm.nih.gov/pubmed/27754580
5. L. Grinblat, Guillermo, C. Uzal, Lucas, Larese, Monica and Granitto, Pablo, "Deep learning for plant identification using vein morphological patterns," *Computers and Electronics in Agriculture*, vol. 127, 2016, pp. 418-424.
6. Lee, Sangwook, Chang, Luh-Maan, Skibniewski nad Miroslaw, "Automated recognition of surface defects using digital color image processing," *Automation in Construction*, vol. 15, 2006, pp. 540-549. https://ri.conicet.gov.ar/bitstream/handle/11336/52668/CONICET_Digital_Nro.bfec2c01-4dbd-4407-be27-54e24e26bb62_D.pdf?sequence=5&isAllowed=y
7. Reza Rasti, Hossein Rabbani, Alireza Mehridehnavi and Fedra Hajizadeh , "Macular OCT Classification Using a Multi-Scale Convolutional Neural Network Ensemble," *IEEE Transactions on Medical Imaging*, vol. 37, no. 04, 2018, pp. 1-11. https://www.researchgate.net/publication/321637395_Macular_OCT_Classification_Using_a_Multi-Scale_Convolutional_Neural_Network_Ensemble
8. Tsaftaris, Sotirios, Minervini, Massimo and Scharr Hanno, "Machine Learning for Plant Phenotyping Needs Image Processing," *Trends in Plant Science*, vol. 21, 2016, pp. 989–991. http://tsaftaris.com/preprints/Tsaftaris_TiPS_Letter.pdf

9. G. Romay, H. Lecoq and C. Desbiez, "Cucurbit Crops and their Viral Diseases in Latin America and the Caribbean Islands: A Review," *Journal of Plant Pathology*, vol. 96, no. 02, 2014, pp. 227-242. https://www.researchgate.net/publication/270214758_Cucurbit_crops_and_their_viral_diseases_in_latin_america_and_the_caribbean_islands_A_review
10. Rong Li, Hua-Ning Wang, Han He, Yan-Mei Cui and Zhan-Le Du, "Support Vector Machine combined with K-Nearest Neighbors for Solar Flare Forecasting," *Chinese Journal of Astronomy and Astrophysics*, vol. 07, no. 03, 2007, pp. 441–447. https://iopscience.iop.org/article/10.1088/1009-9271/7/3/15/meta
11. T Mewes, B Waske, J Franke and G Menz, "Derivation of Stress Severities In Wheat Form Hyper spectral Data Using Support Vector Regression," 2nd Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing , pp. 1-4, 2010. https://www.semanticscholar.org/paper/Derivation-of-stress-severities-in-wheat-from-data-Mewes-Waske/6ff0241d9a7665902bddfae01f6351cd25d97092
12. Arti N. Rathod, Bhavesh Tanawal and Vatsal Shah, "Image Processing Techniques for Detection of Leaf Disease," *International Journal of Advanced Research in Computer Science and Software Engineering*, vol. 03, no. 11, 2013, pp. 397-399. https://pdfs.semanticscholar.org/d34a/df4f766d91d7b152d3515c75461f1287aa68.pdf
13. Isaac Kofi Nti, Gyamfi Eric and Yeboah Samuel Jonas, "Detection of Plant Leaf Disease Employing Image Processing and Gaussian Smoothing Approach," *International Journal of Computer Applications*, vol. 162, no. 02, 2017, pp. 20-25. https://pdfs.semanticscholar.org/37f2/b06142e379a89671a7ee577eca22872eda8d.pdf
14. Weihua Gui, JinpingLiu, ChunhuaYang, NingChen and XiLiao, "Color co-occurrence matrix based froth image texture extraction for mineral flotation," *Minerals Engineering*, vol. 46-47, 2013, pp. 60-67. https://www.sciencedirect.com/science/article/abs/pii/S0892687513000988
15. VijaiSingh and A.K.Misra, "Detection of plant leaf diseases using image segmentation and soft computing techniques," *Information Processing in Agriculture*, vol. 04, no. 01, 2017, pp. 41-49. https://www.sciencedirect.com/science/article/pii/S2214317316300154
16. P.R. Rothe, and R.V. Kshirsaga, "A Study on the Method of Image Preprocessing for Recognition of Crop Diseases," *International Conference on Benchmarks in Engineering Science and Technology (ICBEST 2012)*, 2012, pp. 8-10. https://pdfs.semanticscholar.org/7f15/c0c287287f34c2687d26b5eec483a170bfe0.pdf
17. Sonal P Patel and Arun Kumar Dewangan, "Automatic detection of plant leaf disease using K-means clustering and segmentation," International Journal of Scientific Research Engineering & Technology (IJSRET), vol. 06, no. 07, pp. 774-779, 2017. http://www.ijsret.org/pdf/121919.pdf
18. Amruta B. Patil and J.A.shaikh, "OTSU Thresholding Method for Flower Image Segmentation," *International Journal of Computational Engineering Research*, vol. 06, no. 05, 2016, pp. 1-6. http://www.ijceronline.com/papers/Vol6_issue5/A0650106.pdf
19. P. Mohanaiah, P. Sathyanarayana and L. GuruKumar, "Image Texture Feature Extraction Using GLCM Approach," *International Journal of Scientific and Research Publications*, vol. 03, no. 05, 2013, pp. 1-5. http://www.ijsrp.org/research-paper-0513/ijsrp-p1750.pdf
20. Gjorgji Madzarov and Dejan Gjorgjevikj, "Multi-class classification using support vector machines in decision tree architecture," *IEEE EUROCON 2009*, 2009, pp. 288-295. https://www.researchgate.net/publication/224564327_Multi-Class_Classification_Using_Support_Vector_Machines_In_Decision_Tree_Architecture

.