International Journal of Research in Advent Technology, Vol.6, No.7, July 2018
E-ISSN: 2321-9637
Available online at www.ijrat.org

Survey on Leaf Diseases Identification and Recognition using ΔE Color Difference, Multiple Classifier Algorithms and Principle Component Analysis for Dimension Reduction

M. Banu Priya[1], Dr. S. Jayasankari[2]
(Ph. D Research Scholar[1], Associate Professor in Computer Science[2])
Department of Computer Science [1,2], P.K.R Arts College for Women, Gobi[1,2]
Email:banuchandru146@gmail.com[1], jai.nivi.ravi@gmail.com

Abstract- There are different kinds of data such as Transactional level data, Health care data and Large level of sensor data from Internet of Things. Data accumulating from business applications, data means what we listen to, what we watch and so on. There are two different type of data such as structured and unstructured. We make our recommendation from various sites, these are the part of the work we called it as Big Data. We are accumulating amount of data but we don't know how to extract the useful information from the database or data warehouse. The data mining gives a lot of attention. The data mining means the process of analyze, discover and predict insight from large amount of data using quantitative and technical expertise to solve economic, business and social problem. Based on the algorithm or techniques, data mining has been classified into different mining i.e. acting as Image mining, Text mining, and Content based mining. Image mining help as to discover images from different sites or database using alphanumeric keywords and patterns. Image data different from text data because it should be in a form of understandable and more flexible. Image mining plays a vital role in prediction of leaf diseases. Automated system involves predicting Plants leaf disease towards plant pathology. Image processing has different types of segmentation algorithms. K-means clustering is one of the most popular segmentation algorithms in image processing helps to predict leaves diseases. This algorithm is used to separate disease affected part of leaf or abnormal part of leaf from original leaf images. Today agriculture sector took important and indispensable part in India's economy. In 2017-2018 Paramparagat Krishi Vikas Yojana (PKVY) the scheme has been launched to boost the income of farmers. Prediction of leaf diseases support the farmers to improve the estimated production and income.

Index Terms- *Data mining, Segmentation, Image mining, K-means clustering, Image processing*.

1. INTRODUCTION

Today agriculture sector took important and indispensable part in India'seconomy.54.6% of the population is engaged in agriculture. India government took necessary steps to improve its sustainable development. In 2017-2018 Paramparagat Krishi Vikas Yojana (PKVY) the new scheme has been launched to boost the total income of farmers. In nature the variety of plants affected by more than one diseases. Plant pathology lends a helping hand to predict plant diseases. The word pathology derived from two Greek words patho means (suffering, diseases) and logos mean (study, discourse). Pathology is the pure biological term which is "A study plant diseases (leaf, fruit, stem, flower, root) and their causes procedures for controlling and managing them". There are different kinds of leaves such as paddy leaf, sugar cane, tea, coffee, banian leaf and much variety of plants are hampered by leaf diseases. Maximum 82 % Fungus diseases affect the growth of the plant leaf. Remaining percentage of plant affected by bacteria or virus. The contribution of a normal growth of a plant is highly important to secure the life of both human being and environment. The affected plants have not been taken care then it will lead to dies, causes flower, leaf, and fruit drop. These kinds of diseases hampered plant required to take necessary steps for Identification and treatment. Rice, Wheat, Oats, Barley, Millet, Rye are significant agricultural crop and stable food. The growth of the population annually 1.2 percent. We are in the position to improve the production to cope with the country's growing population, but every year above 37% lose in the overall estimated production because of diseases and pests. Disease detection and recognition are a considerable task, generally it is Very difficult to detect disease manually. Due to the unavailability of experts, naked eye observation of farmers and experts may cause error to identify the diseases. The automated system helps us to identify the diseases based on different clustering algorithm, feature selection and classification. After the recognition researchers recommended farmers to use pesticides,

Fertilizers in the earlier stage of diseases to improve the production.

2. LITERATURE SURVEY

In this research the existing system study describe the new method to explain what all the different kind of disease and how to predict. Siddharth singh chouhan et al. pointed clearly the classification, identification of plant leaf diseases leads to costly and time consuming when it has been encompasses of human involvement. In this research work authors proposed a new method of algorithm called "Bacterial Forging Optimization Based Radial Basis Function Neural Network (BRBFNN) for classification and identification. Diagnosis of leaf disease must have required best treatment to control disease affected plant leaves. Here this advanced method makes it easier to separate the affected part of leaf from abnormal leaf images. This process is called segmentation, which is leads to identify the affected part automatically. Bacterial foraging optimization (BFO) used for optimization to improve the speed and accuracy of neural network. The feature extraction process based on different attributes such as color, texture, shape. BFO is multi optimal function used to initializing the weight of RBFNN, it could identify correct region of leaf affected part. Table-1 gives the accuracy result between three different algorithms using five different fungal diseases. The BRBFNN advanced method shows in the table accuracy has been improved.

2.1 Image processing

Image processing facilitates to perform different operation on an image and in order to obtain enhanced image or extract some useful information from it. It should accept input in the form of image and the output may be an image or features/characteristics associated with that image.

2.2 Image processing techniques

1. Image preprocessing
2. Image enhancement
3. Image segmentation
4. Feature extraction
5. Image classification

2.3 Image enhancement

Maximum raw images collected using satellites, because conventional and digital cameras (DSLR- Digital Single Lens Reflex Camera) are lack in brightness and contrast because the image subsystem has some limitation and illumination conditions while capturing images. This type of images may have different noise. Image enhancement does not increase the inherent information content in the original data but it will have emphasized certain specified image characteristics. Enhancement algorithm generally interactive and application dependent.

Few enhancement techniques

1. Contrast stretching
2. Noise filtering
3. Histogram modification

Name of fungal disease	K – means clustering Algorithm	Genetic algorithm	BRBFNN
Common rusts	0.7817	0.8096	0.8213
Late blight	0.8014	0.8205	0.8326
Early blight	0.8211	0.8374	0.8897
Leaf spot	0.8124	0.8318	0.8836
Leaf curl	0.7517	0.7989	0.8879

Table-1 Name of fungal diseases and accuracy with different algorithm

2.4 Image segmentation

During the image segmentation process each pixel in the images has contain label. The pixels with the label share certain characteristics.

There are different methods used for segmentation

- Thresholding methods such as Otsu's method
- Color based segmentation such as K-means clustering
- Transform methods such as watershed segmentation
- Texture methods such as Texture filters

3. PADDY LEAF DISEASES IDENTIFICATION

Rice is the one of the stable food [2]. The production of rice is affected by different kind of paddy leaf diseases. Maximum time the experts are not available in the remote area. Diagnosis of leaf diseases by farmer may also lead to causes problem and naked eye identification not possible because of cost, and time consuming. To minimize these problems through an automated system.

This study done by Farhana Tazmim pinki et al. describe the leaf diseases there are three different kind of diseases affected the paddy leaf like Brown spot, Leaf blight, and Bacterial blight according to the severity of the diseases automated system advised to use fertilizers and/or pesticide. In this research work they used K-means clustering, this is one of the segmentation algorithm in the field of image processing. K-means clustering helps to separate the diseases affected part of the leaf from paddy leaf images this process is called segmentation. Then we need to classify the leaves it will carry out certain visual contents such as color, shape, texture. Feature extraction based on CBIR system i.e. color, texture, shape are commonly used in content based image retrieval (CBIR) system. Support vector machine undergoes the process of recognition. SVM classifier provides sufficient guidance having instantaneous remedies based on the severity of the paddy leaf diseases [2] which is used for classification and identification of diseases.

Name of the diseases	**Identification / Recognition accuracy**	**Over all accuracy**
Bacterial leaf blight	85.71%	92.06%
Brown spot	90.9%	
Leaf blast	94.11%	

Table-2 Automatic system accuracy

Reference	**Name of the classifier**	**Accuracy**
Joshi et al [12]	Minimum Distance classifier	87.02%
T. sumat et al [13]	K-Nearest neighbour classier	89.23%

Table-3 Comparison of performance

Result of this research gives continues improved accuracy. Compare the maximum distance classifier and K-Nearest neighbour classifier, SVM accuracy has been improved the overall accuracy is 92.06%.

4. DETECTION OF CITRUS DISEASES USING ΔE COLOR DIFFERENCE ALGORITHM

This research work done by H. Ali et al. In that they Introduced different algorithm to detect citrus diseases. The major citrus diseases classified using such techniques. Kinnow mandarin being 80% of Pakistan citrus industry was the main focus of study [3]. Due to the little changes in the appearance and symptoms the diagnosis of diseases requires expert ideas. Sometimes incorrect or inappropriate diagnosis also causes economical loss for farmers, the input like fertilizer and/or pesticides.In this research approaches are different the new algorithm ΔE color difference algorithm support to separate the diseases affected part. The algorithm has been used for assessment of color this method is more than a numeric expression. Usually it's an assessment of color difference from a known standard CIE LAB (L*a*b*) & CIE LAB (L*c*h*) are used to compare two objects.

4.1 Feature Extraction

Color descriptors used for image classification, RGB (Red, Green, Blue), HSV (Hug, Saturation, Value), LBP (Local, Binary, Patterns). Based on these color combination, binary value calculated and produce weighted total of its neighbour's pixel value.

4.2 Principle Component Analysis

This is an algorithm which is used to minimize the loss of information. PCA applied for the features set dimension reduction. PCA – (Reducing the original variable into a lower number of orthogonal synthesized variables and change of variable space).

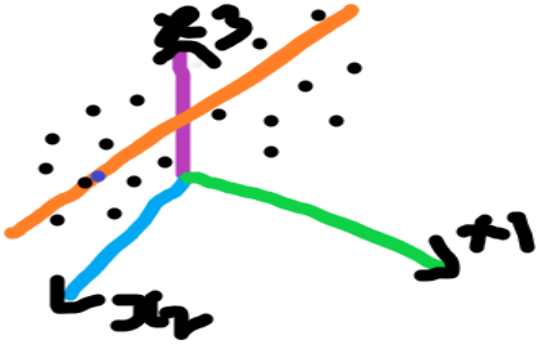

Figure1: Using PCA for dimension reduction.

4.3 Classification

The extracted features are used for training the classifier, collect diseased and non-diseased features are fed into the classifier with the target class [3]. Diseased and non-diseased features of the citrus leaves are taken as input in this stage. In this research K-Cross validation technique used to improve the performance. There are different classifiers are available in the field of data mining which have been used to check the performance. The combination of multiple classifiers used to enhance the accuracy of the overall learning system.

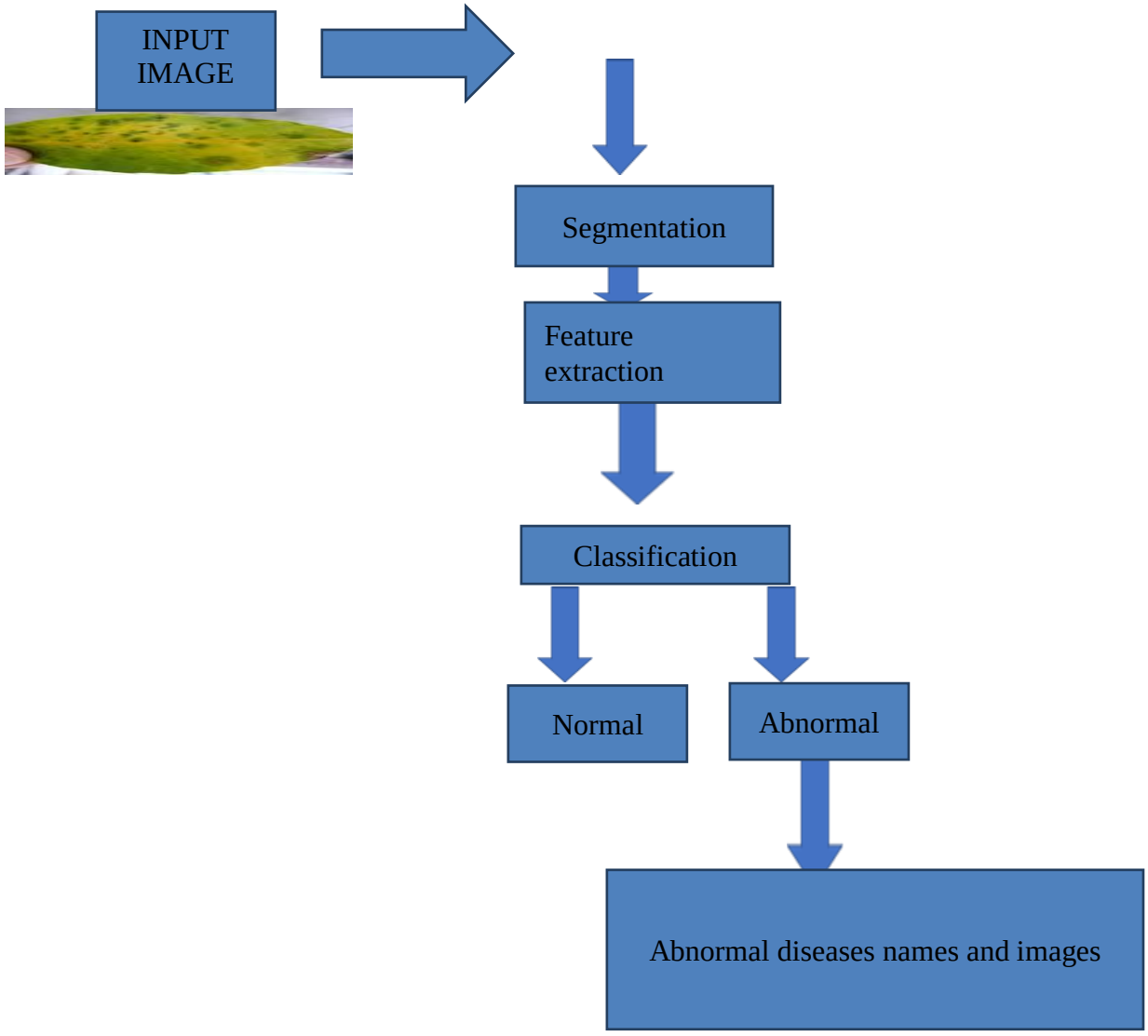

Figure – 1 Leaf diseases prediction (Symptoms in citrus leaves such as Greening, Downy, Anthracnose)

Table-4 Summary of Existing system

Article	Plants	Image prepocessing	Segmentation	Feature Extration	classification	accuracy
Siddarth singh chouhan et al. 2018 January22	Leaf affected by Fungus diseases such as late blight, early blight, leaf blost, brown spot		Bacterial Radial Basis Function	Based on CBIR system	Radial basis Neural Network, Bacterial Optimization (BFO) for optimization and classification which is increase the speed and accuracy.	96%
Farhana Tazmimpinki et al. 2017 December	Paddy leaf diseases (Early Blight, Bacterial blight, Brown spot)	-	K-Means clustering	Based on CBIR	SVM classier	92.06%
H. ali. et al. 2017 April	Kinnow mandarin (Citrus Diseases prediction)	RGB TO LAB	ΔE color difference algorithm. PCA for dimension reduction.	RGB Histogram, HSV Histogram, LBP Histogram	Bagged tree ensemble classifier.	99%
Siddarth singh chouhan et al. 2018 January22	Leaf affected by Fungus diseases such as late blight, early blight, leaf blost, brown spot	-	Bacterial Radial Basis Function	Based on CBIR system	Radial basis Neural Network, Bacterial Optimization (BFO) for optimization and classification which is increase the speed and accuracy.	96%

This system used bagged tree ensemble classifier estimates the generalized error without extra cross validation.

5. RESULT ANALYSIS

This survey paper has given many ideas about image processing and data mining algorithm and view about the performance of algorithm. Finally, we clearly know the meaning of leaf diseases identification and recognition. Citrus diseases prediction researchers explain the what are all symptoms in citrus leaves such as Greening, Downy, and Anthracnose. Automatic detection of citrus diseases which has given 99% accuracy using color descriptor and textual descriptors. We obtain fast and higher accuracy depending on multiple classifiers. This survey paper teaches as to use multiple classifiers to enhance the accuracy of overall learning system.

6. CONCLUSION

In this world plants are highly important and integral part. This automated system helps to the farmers and experts to focus as to create good attention. After prediction we recommended to use fertilizer or pesticides for early stage of diseases control and management. ΔE color difference algorithm and Bagged tree ensemble classifier achieved good performance. Try to save plants that helps us to make our life beautiful. In future it is possible to predict bacteria or virus diseases with multiple classifiers instead of fungus.

Acknowledgment
I have taken effort in this paper however, it not possible without the kind support and help of many individuals and organizations. I would like to express my gratitude and sincere thanks to our college management and I express my sincere respect and gratitude to my college principal and Head of the Department Dr. P. M. Gomathi who has given her valuable support. I am highly indebted to my supervisor Dr. S. Jayasankari for her guidance and constant supervision as well as providing necessary information regarding to my research work.

REFERENCES

[1] Almoda – Ruiz, E., Mattinez – Telllez, M. A, Hernandez -Alamos, mm. Nallejo, S., Prino caisal agent of anthracnose diseases in tropical fruits.193 177-183 Guyer, D.e., Miles, G.E
[2] Chen. Y., Hu X., 2015 High throughput detection of banana plants and aphids using PCR.J, Virol. Methods 193 177-183 Guyer, D.e., Miles, G.E. Gaultnen L.D, Schreiber, M.M Application on machine vision to shape analysis in leaf and plant identification.
[3] L. Cago, Bangladesh population (2017) - World meters, 2017, last accessed Date: 17-July-2017. Rice production in the Asia-pacific region: issues and perspectives M.K. papademetriou*, 2017, last accessed Date: 21-August-2017.
[4] R. Deshmukh, "Detection of paddy leaf diseases," International Conference on Advances in Science and Technology, vol. 2, pp. 8–10, 2015.
[5] R. Islam and M. R. Islam, "An image processing technique to calculate percentage of disease affected area.
[6] S. Phadikar, J. Sil, and A. K. Das, "Classification of rice leaf diseases based on morphological changes," International Journal of Information and Electronics Engineering, vol. 2, no. 3, p. 460, 2012. [6] R. Islam and M. R. Islam, "An image processing technique to calculate percentage of disease affected area.
[7] S. Zhang, X. Wu, Z. You, and L. Zhang, ''Leaf image-based cucumberdisease recognition using sparse representation classification,'' Comput.Electron. Agricult, vol. 134, pp. 135–141, Mar. 2017, doi: 10.1016/j.compag.2017.01.014.