Detection of Plant Leaf Diseases Using Random Forest Classifier

N. Srinivasa Gupta[1], P, Venkata ramana[2], M.Anusha Triveni[3], V.Sai Harika[4], P.Bhanu Prasad[5]
[1]*Asst Prof, Electronics and communications Department, Raghu Institute of Technology, Dakamarri, Visakhapatnam, Andhra Pradesh*
[2]*UG Students, Electronics and communications Department, Raghu Institute of Technology, Dakamarri, Visakhapatnam, Andhra Pradesh*

***Abstract*—Detection of plant disease through some automatic technique is beneficial as it reduces a large work of monitoring in big farms of crops, and at very early stage itself it detects the symptoms of diseases i.e., when they appear on plant leaves. This project presents an algorithm which is used for automatic detection and classification of plant leaf diseases. It also covers survey on different diseases classification techniques that can be used for plant leaf disease detection. Feature extraction which is an important aspect for disease detection in plant leaf disease and GLCM features are extracted and clustering is done. SVM and Random Forest classifier is used to identify the and the results are compared. The experimental results are evaluated using MATLAB tool.**

***Index Terms*— ImageProcessing, random Forest Classifier, MATLAB.**

1. INTRODUCTION

Farming is one of the main occupations in India. Agriculture is important to the economy of all countries. Advancement in the field of agriculture is mainly intended to meet the growing demand of population. The field of agriculture needs an upgradation to survive the present condition. Crops get affected by both the bacterial and fungal diseases. This causes huge damage to the productivity of farmers. To get an optimal yield, crop should be healthy. Detection of diseases through our naked eye will always remain as a complicated process. To do so, constant monitoring of the farm is necessary. This is a tedious process. This is also very costly, when the size of the farm is huge. Due to this difficulty, even agricultural experts are not able to diagnose the diseases easily and find a solution to the problem. Anautomated system that could identify the diseases in plants will be of great help to the farmers. This system may act as a tool to inform the farmers at the right time and take necessary precaution. Various diseases that affect the plant can cause harm to plant parts like leaf, fruit, seed etc. These diseases are unique to different body parts of plant. Leavesare the most important part of plant. If the leaf of the plant is affected by disease, it directly disturbs the plant life cycle.The diseases that are commonly affecting the leaves arebacterial disease, fungal disease etc. Hence, the earlydetection of plant disease is crucial.

2.METHODOLOGY

Leaves in the plant are prone to several disorders. This can be due to humidity, environmental conditions. The usual diseases include virus disease, bacterial disease, and fungal disease. This may lead to change in color shapes. These variations are difficult to recognize because of similar patterns. Therefore, early detection of these diseases can avoid loss. A machine learning approach for classifying plant disease is proposed in this paper.

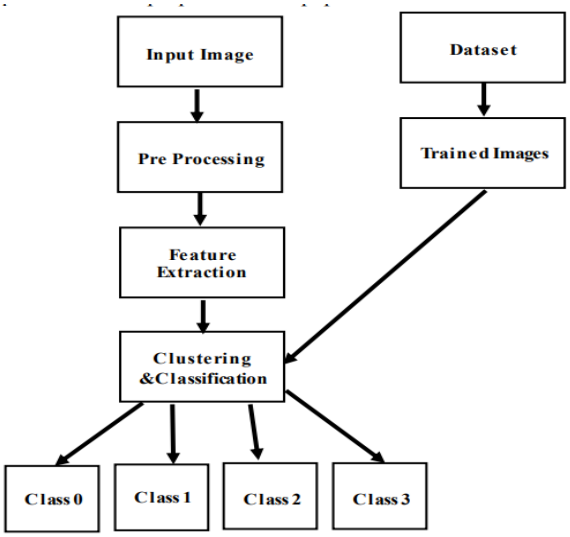


Fig. 1: Flow chart of the proposed method

Input image:
The image from the leaf of the plant is captured. The image is in the RGB form.

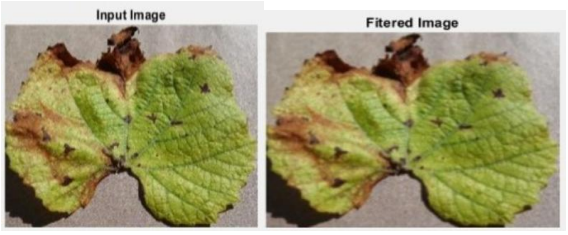
Preprocessing:
Preprocessing technique is applied to remove any kind of noise enhance image features. The images are preprocessed using contrast enhancement. It enhances image features where the contrast of image is increased by mapping input intensity to new value.

Feature extraction:
This step is crucial for classification of image. Instead of selecting the entire image, we extract features only from the infected region. Gray co-occurrence matrix (GLCM) is made used to extract features. It is used to find the spatial difference of fray level in an image. Some of the features that can be extracted from the image using GLCM are Contrast, Correlation, Energy, and Homogeneity. The other features such as Mean, Standard Deviation, Entropy, RMS, Variance, Smoothness, Kurtosis, and Skewness are extracted using MATLAB commands.
Dataset:
The data set used for the training is taken from internet that includes images of plant infected with various types of diseases and healthy leaves.

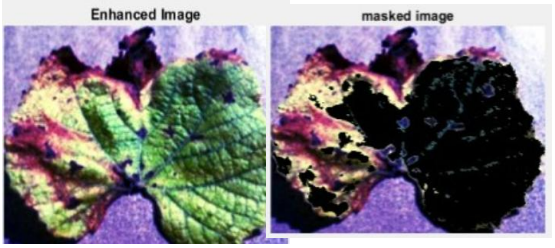
Training:
Here the sample input leaf images containing infected image and healthy images are used for training and fed to the classifier.

Classification:
Random Forest:
Random forest is a decision tree based supervised learningmethod. Unlike decision tree, Random Forest can be used forboth Classification and Regression.In decision tree, Entropy and Gini Index is used to identifythe best attribute to split the tree. The leaf node in the treeconsists of labelled data. The decision tree often leads tooverfitting. Unlike decision trees, Random forests overcomethe disadvantage of over fitting of their training data set andit handles both numeric and categorical data. In Random Forest, of the M available features, N featuresare selected in random. from the available features/ attributes. Random Forest creates more than one tree(N) by randomlyselecting the features. The input data is applied to all N treesin the forest. Each tree predicts the outcome independent ofthe rest of the trees. The output of prediction of the RandomForest is the maximum vote received for each of the typepredicted in the forest.

3.EXPERIMENTAL RESULTS



1.Input Image 2.Filtered Image



3.Enhanced Image 4.Masked Image

5.Disease type

Accuracy	97.3333
Precision	97.0000
Sensitivity	98.9795
Specificity	94.2307

Performance Metric of RF Classifier

4.CONCLUSION

With very less computational efforts the optimum results were obtained, which also shows the efficiency of proposed algorithm in recognition and classification of the leaf diseases. Another advantage of using this method is that the plant diseases can be identified at early stage or the initial stage. The result of the preliminary test shows the better result of disease identification is done using SVM and RFC techniques, in which the RFC achieved good rate of accuracy.

REFERNCES

[1] NonikNovianakurniawati & Salwani Abdullah, "Texture Analysis for Diagnosing Paddy Disease (2009)."
[2] Arnalbarbedo& Jayme Garcia, Digital Image Processing Technique for Detecting, Quantifying and Classifying Plant Diseases (2008).
[3] Shenweizheng&wcyachun, "Grading method of Leaf Spot Disease based on Image Processing (2008)."
[4] Manoj Mukherjee & Titan Pal, "Damaged Paddy Leaf Detection using Image Processing (2012)".
[5] P.Tamijeseivy&kowsalya.N, "Paddy Crop Disease Finder using Probabilistic Approach.(2016)".
[6] Alhouse, Mark L.G. & Chein-I Chang. 2010. Image Segmentation by Local Entropy Methods. Proc. International Conference on Image Processing 3:61-64
[7] R. Pydipati, T.F. Burks, W.S. Lee, "Identification of citrus disease using color texture features and discriminate analysis", Computers and Electronics in Agriculture 52
[8] J.B. Cunha. 2003. Application of Image Processing Techniques in the Characterization of Plant Leaf's. Proc. IEEE Intl' Symposium on Industrial Electronics.
[9] K.J. Runtz. 1991. Electronic Recognition of Plant Species for Machine Vision Sprayer Control Systems. Proc. IEEE Western Canada Conference on Computer, Power, and Communication System in a Rural Environment (WESCANEX'91)
[10] S.Dave and K. Runtz. 1995. Image Processing Methods for Identifying Species of Plants. Proc. IEEE Western Canada Conference on Computer, Power, and Communication System in a Rural Environment (WESCANEX'95),
[11] Rick Cartwright and Fleet Lee, 10 – Management of Rice Diseases, pp.-87
[12] Kobayashi, T., Kanda, E., Kitada, K., Ishiguro, K., Torigoe, Y., 2001. Detection of rice panicle Blast with multispectral radiometer and the Potential of using airborne multispectral Scanners. Phytopathology, 91(3):pp. 316-323.
[13] A.SenthilRajan , Image Processing Techniques for Diagnosing Paddy Disease Proceedings of the World congress on Engineering 2012 Vol II WCE 2012, July 4-6, 2012, London, U.K.