INTERNATIONAL JOURNAL OF ADVANCE RESEARCH, IDEAS AND INNOVATIONS IN TECHNOLOGY

ISSN: 2454-132X
Impact Factor: 6.078
(Volume 7, Issue 3 - V7I3-1452)
Available online at: https://www.ijariit.com

Prediction of plant leaf disease using image pre-processing and filter based optimal feature selection for KNN classifier

Komala T.
komalathippeswamy@gmail.com
Sri Siddhartha Institute of Technology, Tumakuru, Karnataka

Ashwini S. S.
ashwini_249@yahoo.co.in
Sri Siddhartha Institute of Technology, Tumakuru, Karnataka

Dr. M. Z. Kurian
mzkurianvc@yahoo.co.in
Sri Siddhartha Institute of Technology, Tumakuru, Karnataka

ABSTRACT

In plants disease detection is a significant task related to agricultural production of a country. The major economy of a country is linked with the agricultural production which is very important for development of a nation. Any kind of infection on the leaves of the plants leads to loss in crop production and puts down the effort of farmers which in turn hits the economy and livelihood of the country. In this paper we propose an image processing and filter-based feature selection method which distinguishes the disease of the plants leaf and classifies them with using KNN classifier. Using feature selection process, the unwanted redundancy and irrelevant data is filtered which helps the classifier to learn and classify the data more precisely.

Keywords— *Plant Leaf, Disease, image processing, machine Learning, Agriculture.*

1. INTRODUCTION

The backbone of our country is agriculture. Our nation is well-known for agriculture [1-2]. In our country, most of the people have interest to work in agriculture. In the Indian economy, it plays a very significant part. Roughly 70% of rural areas are agriculture dependent. In total, 17 percent of GDP is compensated and 60 percent of the total population is offered job opportunities [3-5]. Therefore, in agriculture the identification of plant diseases plays a key role. Image processing is the suitable methodology utilized in the agriculture sector. Fungus and different bacterial diseases damage plants most [6].

Insecurity of food is caused due to minimized crop yield it is because of pests and diseases. Also, in several under developed and developing countries minimum knowledge regarding diseases and pest control or management leads to minimum food production. Poor disease control capacity, drastic variations in weather, Toxic pathogens are the major factors found in crop production. Several new methodologies have invented to reduce processing of postharvest, to increase the agricultural production and to minimize the food insecurity. Several Laboratory dependent techniques like mass spectrometry, gas chromatography, polymerase chain reaction, and hyper spectral and thermography methods have been emerged to identify the disease. Therefore, these methods consume more time and are not cost effective.

Now a day, mobile dependent and server dependent techniques are invented to identify the diseases. Various attributes of these approaches being maximum operation processing, maximum resolution camera, and also built in equipment are the additional benefits which results auto identification of diseases. To maximize the rate of identification and result accuracy new technologies like machine learning and deep learning algorithm has been employed. Several investigations have been under process in the field of machine learning for identification of plant disease and diagnosis, such traditional machine learning technique being random forest, Support Vector Machine (SVM), artificial neural network, fuzzy logic, Convolution neural networks, K-means method, etc.

The benefit of automatic system helps in case of large crops and also avoids the farmers to continuously visit to their fields. This system identifies the disease by recognizing the disease symptoms. The variations in weather occur due to the increase in population which affects plants [7-8]. Along with the modern approaches diseases in plants can be identified easily. To increase production and post-harvest life [9] leafs requires frequent observations. The image of the affected leaf is considered and processed to identify the diseases called Anthracnose, bacterial blight, Cercospora leaf spot and Alternaria with increased accuracy.

2. LITERATURE REVIEW

V. S. Rajpurohit. And S.S. Sannakki invented a "Classification of Pomegranate Diseases Based on Back Propagation Neural Network" [10], here it majorly operates on segmentation method and it utilizes texture, color and affected area as its features. Here classification uses neural network classifier. The major benefit is 97.30% accurate Categorization is obtained and it Converts to L*a*b to obtain chromaticity layers of the image. It uses for a smaller number of crops this is major drawback of this technique.

R. V. Kshirsagar and P. R. Rothe proposed a" Cotton Leaf Disease Identification using Pattern Recognition Techniques" [11] here method of snake segmentation used, here Hu's moments are considered as distinctive attribute. Vitality within the affected area restricted by using Active contour model, BPNN classifier tackles the various class problems. Here 85.52% average classification is obtained.

Shanu Sharma, Ritika Arora and Aakanksha Rastogi, introduced a" Leaf Disease Detection and Grading using Computer Vision Technology &Fuzzy Logic". [12] For the extraction of features of texture GLCM is utilized. To evaluate the severity of the affected leaf artificial neural network (ANN) as a classifier is utilized. For the disease grading Fuzzy logic is utilized. Segmentation of affected spot is done by using K-means clustering.

Godliver Owomugisha, James Lwasa, Ernest Mwebaze and John A. Quinn, introduced" Automated Vision-Based Diagnosis of Banana Bacterial Wilt Disease and Black Sigatoka Disease "[13] Color histograms are obtained and converted from RGB to L*a*b,RGB to HSV. For classification area under the curve estimation and five shape attributes are utilized. To form max tree Peak elements are utilized. And also utilizes Decision tree, random forest, and SV classifier, nearest neighbors, Naïve bayes and extremely randomized tree. In seven classifiers majorly, randomized trees give flexibility to the application and also give real time details.

Uan Tian, Chunjiang Zhao, XinyuGuo and Shenglian Lu," SVM-based Multiple Classifier System for Recognition of Wheat Leaf Diseases," [14] Color attributes are converted from RGB to HIS, and seven invariant moments are used as shape parameters using GLCM. They also used the MCS SVM classification used for offline wheat plant disease identification. The plant leaves are infected by fungal, bacterial, and viral diseases it involves powdery mildew, brown spot, Downey mildew and bacterial blight etc.
J. R. Yakkundimath and J. D. Pujari, A.S. Byadgi appliedProbablistic neural Network, Support Vector Machines and Artificial Neural Network in the fields of vegetables, commercial crops, and cereals, for disease identification [15-18].

3. PROPOSED METHOD

The figure1 below shows an implementation flow of the proposed method.

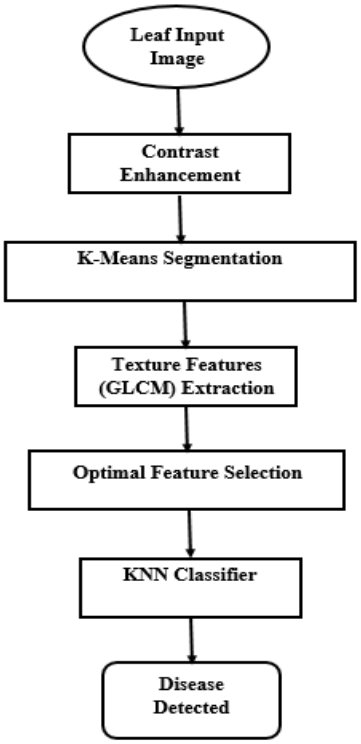

Figure 1: Implementation Flow

In this method we are aimed to predict the illness of the leaves by following the steps given in the implementation flow model for the process using image processing and machine learning with our own method of flow. The flow as follows we take the input color image of leaf which is pre-processed to enhance its contrast and segmented using k-means segmentation algorithm followed by texture(GLCM) features are extracted for the leaf image then the extracted features are optimally selected to remove the redundancy and irrelevant things in the features which we can call it as fine tuning of the features for the better learning and testing process of the machine, Here we use KNN classifier for training and testing the optimal features which are extracted using the filter method which is discussed here.

Feature Selection: Feature selection is done to reduce the redundant and noisy unwanted irrelevant data from the feature set of the leaves. So that, the number of features are more manageable and good for the processes to be performed. It consists of getting the optimal information from the feature dataset which will not only reduce the training cost also helps the learning algorithms to learn accurate and classify accurate which is the final outcome expected in this entire process. In particular the performance of KNN classifier can be increased by doing this feature selection process for the detection of leaf diseases in the plants. In this proposed method the method used for optimal feature selection is filter method.

Filter Method-In this filter method ranking based techniques are used. Ranking is assigned to the variables which have scores assigned based on some threshold values. This kind of approach provides independency to the machine learning algorithms as it is of more general approach. Also, these methods cost less on computation process of the machine during learning and testing phase and avoids over fitting by ignoring the independencies of the features in the feature dataset.

4. RESULTS

The experimental results are discussed as follows. Figure 2 shows the input leaf image, enhanced leaf image and segmented leaf image as part of the implementation process. Figure 3 and table 1 below shows the comparison of accuracy with and without optimal feature selection with the KNN classifier.

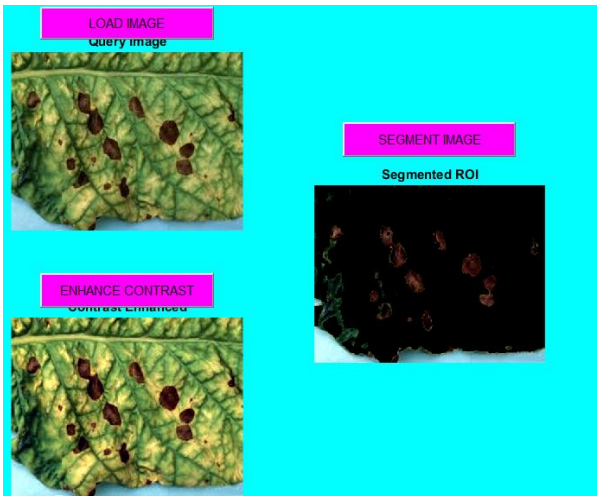

Figure 2: Enhanced and Segmented Images

Table1: Accuracy of proposed

Sl. No.	Disease	Affected Region (%)	Accuracy without feature selection (%)	Accuracy with optimal feature selection (%)
1	Alternaria Alternata	16.321	88.52	95.21
2	Bacterial Blight	15.234	83.54	93.32
3	Anthracnose	15.201	85.25	92.10
4	Cercospora Leaf Spot	15.121	81.81	93.42
5	Healthy leaf	None	90.12	95.28

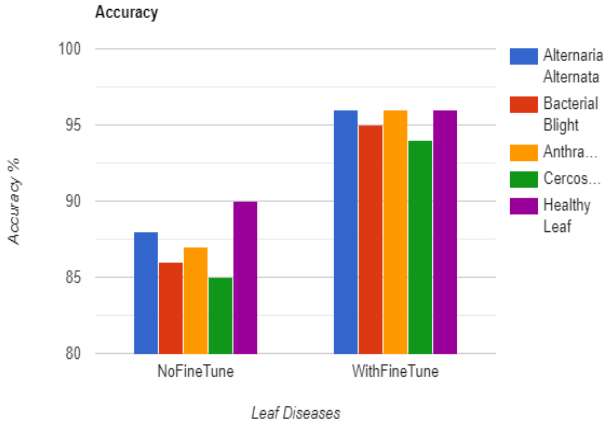

Figure 3: Accuracy Plot

5. CONCLUSION

Recognising abnormalities is the main objective of the proposed method which occurs in the leaves of the plants in their green house environment. Improved accurate results are observed in this model which also contrasts with the other machine learning and image processing algorithms in terms of accuracy. In this method the feature selection process tends to give an improvement in disease detection of the plant leaves. The graph and table in the results are given which gives comparison of results with and without optimal feature selection.

6. REFERENCES

[1] Reena Tijare, PawanKhade, Rashmi Jain, The Survey of DiseaseIdentification of Cotton Leaf, International Journal of InnovativeResearch in Computer and Communication Engineering, 2015.

[2] Sasirekha N, Swetha N, An Identification of Variety of Leaf Diseases Using Various Data Mining Techniques, International Journal ofAdvanced Research in Computer and Communication Engineering, 4(10), 2015.

[3] SujaRadha, Leaf Disease Detection using Image processing, Journal ofChemical and Pharmaceutical Sciences, March 2017.

[4] Prabhjeet Kaur, Sanjay Singla, Sukhdeep Singh, Detection andclassification of leaf diseases using an integrated approach of supportvector machine and particle swarm optimization, International Journal ofAdvanced and Applied Sciences, March 2017.

[5] Saradhambal.G, Dhivya.R, Latha.S, R.Rajesh, Plant disease detection,and its solution using image classification, International Journal of Pureand Applied Mathematics, Volume 119 No. 14 2018, 879-884.

[6] AbiramiDevaraj, KarunyaRathan, SarvepalliJaahnavi, and K Indira,Identification of Plant Disease using Image Processing Technique,International Conference on Communication and Signal Processing,April 4-6, 2019.

[7] Vishal Mani Tiwari&Tarun Gupta "Plant Leaf Disease Analysis usingImage Processing Technique with Modified SVM-CS Classifier"Research Gate2017.

[8] B.V. Ramana Reddy, A. Suresh, M. Radhika Mani, and V.Vijaya 50Kumar, "Classification of Textures Based on Features Extracted fromPre-processing Images on Random Windows", International Journal ofAdvanced Science and Technology, Volume 9, pp 9 – 18, August2009.

[9] H. Al-Hiary, S. Bani-Ahmad, M. Reyalat, M. Braik, and Z. ALRahamneh, "Fast and Accurate Detection and Classification of PlantDiseases", International Journal of Computer Applications (0975 –8887) Volume 17– No.1, March 2011.

[10] S. S. Sannakki and V. S. Rajpurohit," Classification of PomegranateDiseases Based on Back Propagation Neural Network," InternationalResearch Journal of Engineering and Technology (IRJET), Vol2 Issue:02 | May-2015

[11] P. R. Rothe and R. V. Kshirsagar," Cotton Leaf Disease Identificationusing Pattern Recognition Techniques", International Conference onPervasive Computing (ICPC),2015.

[12] AakankshaRastogi, Ritika Arora and Shanu Sharma," Leaf DiseaseDetection and Grading using Computer Vision Technology &FuzzyLogic" 2nd International Conference on Signal Processing andIntegrated Networks (SPIN)2015.

[13] GodliverOwomugisha, John A. Quinn, Ernest Mwebaze and JamesLwasa,"Automated Vision-Based Diagnosis of Banana Bacterial WiltDisease and Black Sigatoka Disease ", Preceding of the 1'st International conference on the use of mobile ICT in Africa, 2014.

[14] uan Tian, Chunjiang Zhao, Shenglian Lu and XinyuGuo," SVM-basedMultiple Classifier System for Recognition of Wheat Leaf Diseases,"Proceedings of 2010 Conference on Dependable Computing (CDC'2010), November 20-22, 2010.

[15] J. D. Pujari, R. Yakkundimath, and A. S. Byadgi, "Identification andclassification of fungal disease affected on agriculture/horticulture cropsusing image processing techniques," *IEEE International Conference onthe Computational Intelligence and Computing Research*, 2014.

[16] BalasubramanianVijayalakshmi and Vasudev Mohan, "Kernel based PSOand FRVM: An automatic plant leaf type detection using texture, shapeand color features," *Computer and Electronics in Agriculture*, vol. 125,pp. 99-112, 2016.

[17] X. Wang, M. Zhang, J. Zhu and S. Geng, "Spectral prediction ofPhytophthorainfestans infection on tomatoes using artificial neuralnetwork (ANN)," *International Journal of Remote Sensing*, pp. 1693–1706, 2008.

[18] Dong Pixia and Wang Xiangdong, "Recognition of Greenhouse CucumberDisease Based on Image Processing Technology," *Open Journal ofApplied Sciences*, vol. 3, pp. 27-3, Mar. 2013.

