Supervised Learning Based Infection Identification Using Image Processing

Prachi Gadkari[1], Prof. Sufola Das Chagas Silva e Araujo [2], Prof. Megha Vishant Ainapurkar[3]
Master of Information Technology, Padre Conceicao College of Engeneering ,Goa University
Verna-Goa, India

Abstract

Crop production is one of the major occupation of earning in India as more than half of our population relies on agriculture for livelihood but due to the factors like diseases, pest attack and sudden change in the weather condition, the productivity of the crop decreases. Identification of plant and crop infections is the key to prevent losses both in terms of quality and quantity of the agricultural yield. We have traditional methods of checking infection in crop these methods are not relevant and accurate enough in detecting the diseases associated with plants.
So this paper provides a better alternative that can accurately and effectively identify and detect plant infection and helps to enhance the yield quality and quantity of plants.

Keywords - *Put your keywords here, keywords are separated by comma.*

I. INTRODUCTION

Agriculture is the largest economic sector in India and about 75% of the population depends on agriculture which plays a major role in economic development. Infection in the plant leaf may occur due to sudden changes in the climatic condition such as heavy rain fall, drastic changes in temperature or may be due some insects and pesticides resulting in poor quantity in yield. Infection in the plant produces a negative impact on the countries whose economies are primarily dependent on the agriculture. Detecting the infection on the plant in the earlier stages is very important to avoid the loss in terms of productivity and economic progress.

Chilli which is the fruit of capsicum family plant is almost a delicacy in Indian menu. This fruit has a high local demand and can fetch a very handsome economic yield. Producing Chilli is a daunting task as the plant is exposed to the attacks from various organisms, bacterial and viral diseases and pests. The symptoms of the attacks are categorized through the leaves, stems or fruit inspection. It has been observed that Goa faced a major problem in chilly crop cultivation as there where many infections observed at each stages of chilly crop plantation for a variety called G-4 (Guntur-4) open pollinated which lead to poor quality as well as less quantity of chillies that were grown and farmers faced heavy losses. So the identification and detection of infection on the chilli leaf will help to detect infection at the early stage and help to save from the heavy losses.

This project aims to identify the disease accurately from the chilli leaf images, classify them as healthy or unhealthy plant and then on the basis of level of infection, prescribe suitable pesticides to prevent the loss in yield.

Section II shows the study of the literature materials related to the proposed approach. It determines what has already been written on the topics, identifies previous approaches to the topic, identifies central issues in the field, integrates what previous researchers have found and also identifies important issues still unresolved. Section III presents detailed description for the implementation of the application. Section IV describes the methodologies used in the proposed approach. Section V discuss the performance of the system and finally Section VI presents the future scope of this approach.

II. LITERATURE REVIEW

Siddharth Singh Chouhan, Ajay Kauluday[1] proposed method uses Bacterial Foraging Optimization (BFO) to assign optimal weight to Radial Basis Function Neural Network (RBFNN) segmentation technique based on two factors i.e. specificity and sensitivity for the images which is found to be 0.8558 and 0.8705 respectively. Everton Castelão Tetila , Bruno Brandoli Machado[2], this paper presents the Simple Linear Iterative Clustering (SLIC) segmentation algorithm, in order to detect the plant leaves in the images. The method SLIC employs the *k*-means algorithm for the generation of regions, called superpixels. This methodology evaluated the performance of six classifiers and experimental results showed that color and texture attributes lead to higher classification rates, achieving the precision of 98.34%. G. Sandika Biswas, Bhushan Jagyasi, Bir Pal Singhy and Mehi Lalz [3] proposed an FCM clustering and neural network classification based approach to detect and quantify the severity for late blight disease of potato by segmented the affected area using K-Means clustering. The proposed algorithm achieves an accuracy of 93% for 27 images captured in different light condition, from different distances and at different orientations along with

complex background. Pedram Ghamisi, Micael S. Couceiro[5] proposed a novel multilevel thresholding segmentation method called fractional-order Darwinian particle swarm optimization (FODPSO) for grouping the pixels of multispectral and hyperspectral images into different homogenous regions. The proposed method have a overall accuracy of 94.3%.

Siddharth Singh Chouhan, Ajay Kauluday[1] proposed Region Growing Algorithm (RGA) for feature extraction which is a simple approach that starts with the set of seed points and then grows by using these seed points forming a region by appending to each seed the adjoining pixels, having analogous features to the seed such as intensity level, color, or scalar properties for the grayscale images. Sushil R. Kamlapurkar [13] proposed classification based on two principles-discontinuity and similarity. Discontinuity extracts the regions having different properties like intensity, colour, texture etc. Similarity groups the image pixels into groups with some predefined criteria. Based on pixel similarity with the neighbouring pixel, the algorithm used is region based. In leaf disease identification, segmentation is used to identify the diseased area. From this, features of a region are computed. Angel Dacal-Nieto, Esteban Vázquez-Fernández, Arno Formella, Fernando Martin, Soledad Torres-Guijarro, HiginioGonzález-Jorge[15] proposed method taking into account color and projections. Then, features have been extracted from HSV and RGB channels from every segmented potato, using histogram and co-occurrence matrix texture characteristics, creating a learning set. A classifier has been developed by using 1-NN algorithm. This classifier has been optimized with an ad-hoc GA that selects the most discriminant features subset. Muhammad Danish Gondal, Yasir Niaz Khan[17] used texture, shape and color features of color image of disease spot on leaf were extracted, and a classification method of membership function was used to discriminate between the three types of diseases. The analysis of the results showed over 70 percent classification accuracy around 50 sample images. Mr. Hrishikesh P. Kanjalkar, Prof. S.S.Lokhande[18] proposed methodology that uses leaf features to detect disease where feature extraction is done on segmented diseased area. Hue image from HSI gives clear discrimination of diseased spots, and which is more helpful for extracting size, color and centroids.

III.SYSTEM ARCHITECTURE

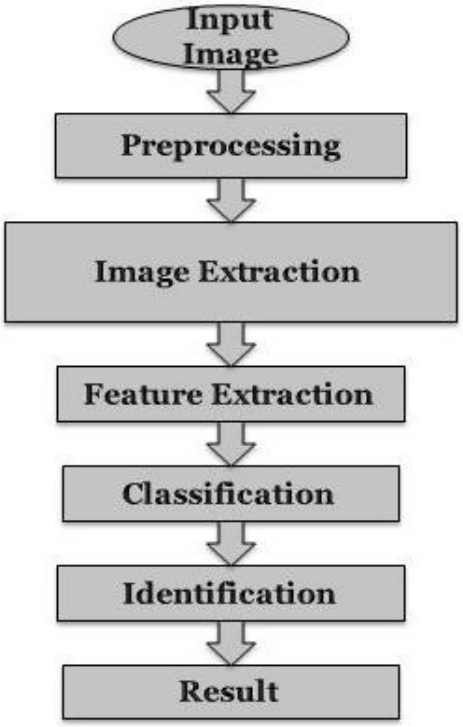

Fig 1: Process Flow Diagram

The overall process flow for the system mentioned is the Fig 1 is described below:

A. Image Acquisition

Chilli leaf images are captured at different stages under laboratory condition using digital camera in a mini photo box of 30cm * 45cm * 35 cm. A total of 180 digital images were taken in the crop year june 2018 – December 2018.

Fig 2: Sample input image

B. Image Pre-processing

The basic idea of image pre-processing is to enhance the image so that it can be well used for further processing. The image is resized to 256×256 pixel. Upon resizing the image is passed through Guassian filter to reduce the noise. The RGB image is converted to HSV in order to make the object discrimination easier.

Fig 3: (a) Resized Image (b)Gaussian Filtered Image (c)HSV Image

C. Image Extraction

Extracting meaningful information from the image is done using segmentation wherein we divide an image into multiple parts.

A soft clustering algorithm – Fuzzy C-Means (FCM) [10] is used to segment the image in which each data object is member of multiple cluster with varying degrees of fuzzy membership between 0 and 1 in FCM.

Fig 5: FCM Segmented Output

A population based stochastic optimization technique – Particle Swarm Optimization (PSO) segments the image by iteratively trying to improve candidate solution.

Fig 6: PSO Segmented Output

D. Feature Extraction

Feature plays important role in image processing to classify and recognize the image. Interesting features of image such as color, shape and texture is extracted from the segmented image and represented as feature vector.

E. Classification

Classification is done using Convolutional Neural Network(CNN). System is trained to classify based on features into different stages varying from healthy stage to partial infection stage to fully infected stage.

F. Identification

Co-relation between the given user test image and the trained data set is found. A GLCM is used to find co-relation. Similarity between the trained and test image is calculated and which stage the test plant leaf image belongs to is identified

G. Result

The proposed approach Specify what stage of infection the plant or crop is in and suggesting the suitable pesticide based on the stage of infection.

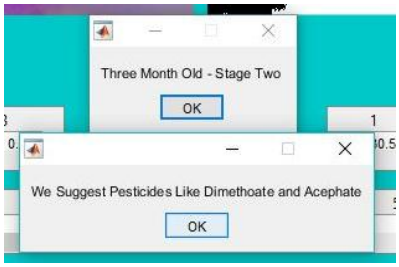

Fig 4: Result of the Proposed System

IV. PROPOSED METHODOLOGY

A. Fuzzy C-Means

Fuzzy algorithms [9] can assign data object partially to multiple clusters and handle overlapping partitions. The degree of membership in the fuzzy clusters depends on the closeness of the data object to the cluster centers. The most popular fuzzy clustering algorithm is FCM which is introduced by Bezdek [10] and now it is widely used.

Fuzzy C-Means Algorithm:

Step 1: select c where, (2 < c < n) , select a value for parameter m, and initialize partition matrix.

Step 2: For each step, compute the centers of each cluster.

Step 3: Update the partition matrix

$$u_{ik} = \left[\sum_{j=i}^{c} \left(\frac{d_{ik}}{d_{jk}} \right)^{2/m-1} \right]^{-1}$$

Step 4. Check for convergence if

$$\left\| u^{(r-1)} - u^{r} \right\| < \epsilon_l,$$

else repeat step 2

B. Particle Swarm Optimization

Particle swarm optimization algorithm (PSO) [10] is an evolutionary computational technique based on the movement and intelligence of swarms. Kennedy and Eberhart proposed the Particle Swarm Optimization technique. PSO is initialized with a group of random particles and then searches for optima by updating generation. Each particle is flown through the search space, having its position adjusted based on its distance from its own personal best position and the distance from best particle of the swarm [11].

Particle Swarm Optimization Algorithm:

Step 1: Initialize cluster centers for each particle randomly.
Step 2: For each particle, assign each pixel to a cluster that has the minimum distance to its cluster center.
Step 3: Calculate the fitness function for each particle and find the global best solution.

$$v' = v + c1.r1.(pBest - x) + c2.r2.(gBest - x)$$
$$x' = x + v'$$

Where,
v is the current velocity, v' the new velocity,
x the current position, x' the new position,
r1 and r2 are even distributed random numbers in the interval [0, 1], and c1 and c2 are acceleration coefficients,
Where c1 is the factor that influences the cognitive behaviour,
and c2 is the factor for social behaviour.
Step 4: Update the cluster centers.
Step 5: Repeat the procedure until the cluster converges.

C. Convolutional Neural Network

A Convolutional Neural Network (CNN) [12] is a Deep Learning algorithm which can take in an input image, assign importance (learnable weights and biases) to various aspects/objects in the image and be able to differentiate one from the other.

CNN are made up of neurons that have learnable weights and biases. Each neuron receives some inputs, performs a dot product and optionally follows it with a non-linearity. The whole network still expresses a single differentiable score function: from the raw image pixels on one end to class scores at the other. And they still have a loss function on the last (fully-connected) layer and all the tips/tricks we developed for learning regular Neural Networks still apply [13].

V. CONCLUSION

Detecting the infection in the plant in the early stage is important as it will help increase the productivity of agricultural product. Several work has been proposed in similar lines. In this approach we use stochastic based Particle Swarm Optimization and Fuzzy C-Means algorithm to improve the accuracy. The system aims to help the farmer to recognize the stage and take effective measures that will help him progress against economic losses. The system is also tested for various performance measures such as accuracy, sensitivity, specificity, precision, recall and F-measure shown in the fig 7.

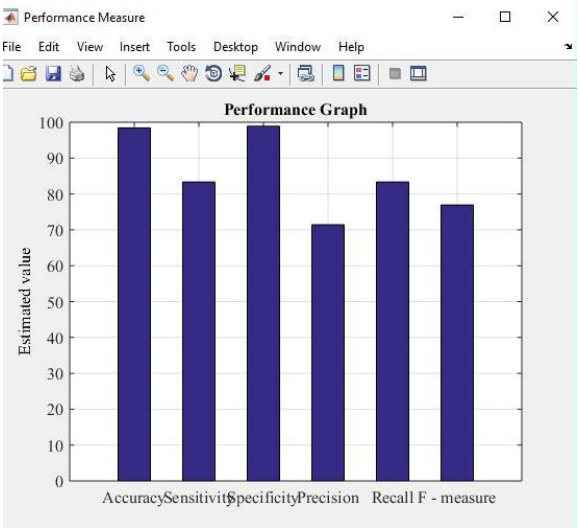

Fig 7: Performance Measure Graph

VI. FUTURE SCOPE

The infections in the leaf, fruit and crop keeps growing due to environmental factors, viruses and pests so there is always a scope to tackle with this issue. Another work to be done is to implement new and better techniques for segmentation and feature extraction that can help identify diseases more efficiently and accurately.

REFERENCES

[1] Siddharth Singh Chouhan, Ajay Kauluday Pratap Singh, and Sanjeev Jain," Bacterial Foraging Optimization Based Radial Basis Function Neural Network (BRBFNN) for Identification and Classification of Plant Leaf Diseases: An Automatic Approach Towards Plant Pathology", IEEE, VOLUME 6, 2018

[2] Everton Castelão Tetila , Bruno Brandoli Machado , Nícolas Alessandro de Souza Belete , David Augusto Guimarães, and Hemerson Pistor," Identification of Soybean Foliar Diseases Using Unmanned Aerial Vehicle Images", IEEE GEOSCIENCE AND REMOTE SENSING LETTERS, VOL. 14, NO. 12, DECEMBER 2017.

[3] Sandika Biswas, Bhushan Jagyasi, Bir Pal Singhy and Mehi Lalz, "Severity Identification of Potato Late Blight Disease from Crop Images Captured under Uncontrolled Environment" 2014 IEEE Canada International Humanitarian Technology Conference - (IHTC)

[4] G. Anthonys, N. Wickramarachchf, "AnImage Recognition System for Crop Disease Identification of Paddy fields in Sri Lanka", IEEE Fourth International Conference on Industrial and Information Systems, ICIIS2009, 28 -3I December2009, SriLanka.

[5] Pedram Ghamisi, Micael S. Couceiro, Fernando M. L. Martins, and Jón Atli Benediktsson, " Multilevel Image Segmentation Based on Fractional-Order Darwinian Particle Swarm Optimization", IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING, VOL. 52, NO. 5, MAY 2014.

[6] R. Revathy, Dr. R. Roselin, "Digital Image Processing Techniques for Bacterial Infection Detection on Tomato and Crape Jasmine Leaves," International Journal of Scientific & Engineering Research, Volume 6, Issue 6, June-2015.

[7] Prashant R. Rothe, Ravindra. V. Kshisagar, "AUTOMATED FEATURE EXTRACTION IN COTTON LEAF IMAGES BY GRAPH CUT APPROACH", RESEARCHWORLD Press, International Conference, Kuala Lumpur Malaysia, 18th October 2014, ISBN: 978-93-84209-57-5.

[8] Jian Zhang and Ling Shen, "An Improved Fuzzy c-Means clustering Algorithm Based on Shadowed set and PSO", Computational Intelligence and Neuroscience Volume 4
[9] Harveen Anand, Vipul Dalal, "Comparative Study of Particle Swarm Optimization and Fuzzy C-means to Data Clustering", International Journal of Computer Application Technology and Research, Volume 3- Issue 1, 2014.
[10] Fahd Mohsen, Mohiy Hadhoud, Kamel Mostafa and Khalid Amin, "A New Image Segmentation based on particle Swarm Optimization", The International Arab Journal of Information Technology, Vol.9, No.5, September 2012.
[11] A Comprehensive Guide to Convolutional Neural Network : https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a53
[12] Convolutional Neural Network for Visual Recognition : http://cs231n.github.io/convolutional-networks/
[13] Sushil R. Kamlapurkar, "Detection of Plant Leaf Disease Using Image Processing Approach," International Journal of Scientific and Research Publications, Volume 6, Issue 2, February 2016.
[14] Jagadeesh D.Pujari, Rajesh Yakkundimath and Abdulmunaf. Syedhusain Byadgi," SVM and ANN Based Classification of Plant Diseases Using Feature Reduction Technique", International Journal of Interactive Multimedia and Artificial Intelligence, Vol. 3, No7.
[15] Angel Dacal-Nieto, Esteban Vázquez-Fernández, Arno Formella, Fernando Martin, Soledad Torres-Guijarro, HiginioGonzález-Jorge, "A genetic algorithm approach for feature selection in potatoes classification by computer vision", 2009 35th Annual Conference of IEEE Industrial Electronics.
[16] B.Ramapriya, D.Rajaselvi, "Disease Segmentation in Vegetables using K-Means Clustering ", International Journal of Pure and Applied Mathematics, Volume 119, Issue No.10, 2018
[17] Muhammad Danish Gondal, Yasir Niaz Khan,"Early Pest Detection from Crop using Image Processing and Computational Intelligence", FAST-NU Research Journal (FRJ), Volume 1, Issue 1, January 2015.
[18] Mr. Hrishikesh P. Kanjalkar, Prof. S.S.Lokhand, "Feature Extraction of Leaf Diseases", International Journal of Advanced Research in Computer Engineering & Technology (IJARCET) Volume 3, Issue 1, January 2014.
[19] GuilingSun, XinglongJia, and TianyuGeng, "Plant Diseases Recognition Based on Image Processing Technology", Journal of Electrical and Computer Engineering,Volume 2018.