Farmer Advisory: A Crop Disease Detection System

Anuradha Badage [1], Aishwarya C [2], Ashwini N [3], Navitha K Singh [4], Neha Vijayananda [5]

[1]*Assistant Professor, Department of CSE, Sapthagiri College of Engineering, Karnataka, India*
[2]*Student, Department of CSE, Sapthagiri College of Engineering, Karnataka, India*
[3]*Student, Department of CSE, Sapthagiri College of Engineering, Karnataka, India*
[4]*Student, Department of CSE, Sapthagiri College of Engineering, Karnataka, India*
[5]*Student, Department of CSE, Sapthagiri College of Engineering, Karnataka, India*

***

Abstract - *Agriculturists are facing loss due to various crop diseases. It becomes difficult for the cultivators to monitor the crops on regular basis when the cultivated area is huge in terms of acres. If proper care is not taken in this area then it causes serious effects on plants and due to which respective product quality, quantity and productivity is affected. Smart farming is need of the hour of the Indian economy. There is a need of an automatic, accurate and less expensive system for detection of diseases from the image and to suggest a proper pesticide as a solution. The most significant part of our research is early detection the disease as soon as it starts spreading on the top layer of the leaves using remote sensing images. This approach has two phases: first phase deals with training the model for healthy and as well as diseased images , second phase deals with monitoring of crops and identification of particular disease using KNN algorithm and also intimate the agriculturists with an early alert message immediately.*

Key Words: **Crop monitoring, Disease detection, Remote Sensing image, Canny Edge algorithm, KNN algorithm.**

1. INTRODUCTION

The agricultural land mass is more than just being a feeding sourcing in today's world . Indian economy is agriculture based and it is the main source of rural livelihood. Indian economy is highly dependent of agricultural productivity. Therefore in field of agriculture, detection of disease in plants plays an important role. Every living being depends on agriculture for food. But for better yield, the crops should be healthy therefore some highly technical method is needed for periodic monitoring. Plant disease is one of the important factor where it can cause significant reduction of quality and quantity of agriculture products. Due to the exponential inclination of population, the climatic conditions also cause the plant disease. The plants suffer from diseases that can drastically affect the quantity and quality of the yield. Usually the detection and identification of leaf diseases is performed by farmers by naked eye observation . It leads to incorrect diagnosis as the farmer's judge the symptoms by their experience. This will also cause needless and excess use of costly pesticides. Therefore the automatic detection of disease is important which will help in early and accurate diagnosis of leaf diseases The major challenges of sustainable development is to reduce the usage of pesticides, cost to save the environment ad to increase the quality. Precise, accurate and early diagnosis may reduce the usage of pesticides. Consequently, effective monitoring of the incidence and severity of crop diseases is of great importance to guide the spray of pesticides. The existing method for plant disease detection is simply naked eye observation by experts through which identification and detection of plant diseases is done. For doing so, a large team of experts as well as continuous monitoring of plant is required, which costs very high when we do with large farms. At the same time, in some countries, farmers do not have proper facilities or even idea that they can contact to experts. Due to which consulting experts even cost high as well as time consuming too. In such conditions, the suggested technique proves to be beneficial in monitoring large fields of crops. Automatic detection of the diseases by just seeing the symptoms on the plant leaves makes it easier as well as cheaper. Machine Learning provides a possible way to detect the incidence and severity of the disease rapidly. This approach starts with training of images for both the samples such as healthy and disease leaf images.

2. LITERATURE SURVEY

There are different types of algorithms used for different types of crops. Canny edge detection algorithm is used for rice and wheat crops and the images of crops must be converted to gray scale. This algorithm obtains the edges accurately since many diseases causes leaf deformities. SVM (Support Vector Machine) classifier is used for sugarcane which classifies normal and diseased images. All features from these images are extracted and classified into normal and diseased crops and then accuracy is calculated. Adaptive neuro fizzy algorithm is used for cotton leaf and helps in classification. It combines the principles of neural network and fuzzy logic therefore it provides advantages of both in a single structure. Active Contour model is used for image segmentation and Hu's moments are extracted as features for cotton leaf diseases. Back propagation neutral networks are used for solving multiple class problems. K-means clustering algorithm is used for pomegranate plant diseases which is used to divide the image into its constituent regions and objects. Images are partitioned into four clusters in which one cluster contains the majority of the diseased part of the image. From these images,the texture features are extracted using GLCM method.

Santanu Phadikar and Jaya Sil [3] proposed Rice Disease Identification using Pattern Recognition Techniques .The proposed system is software prototype system for rice disease detection based on the infected images of various rice plants .The proposed system consists of following steps :first step is the feature extraction method that include region segmentation , boundary detection and spot detection . In feature extraction the images have been grabbed using digital cameras .Then in segmentation method brightness and contrast of images are increased and then transformed into hue intensity model . In boundary detection method boundary detection algorithm is applied and based on that spot detection is done. In second step zooming algorithms are used which is used to interpolate points in the detected spots. Final step is classification in which som neural network algorithm is used. Advantage of this system is that the zooming algorithm extracts features of the images using simple computational efficient technique, which results in satisfactory classification of test images. Disadvantage is that transformation of image in the frequency domain does not yield in better classification. Aging is not sensed. Historical analysis is not performed.

Dheeb Al Bashish, Malik Braik and Sulieman Bani-Ahmad[4] proposed and evaluated a framework for detection of plant leaf/stem diseases. The proposed framework is image processing based solution for the automatic leaf disease detection and classification and is composed of the following main steps: first step the images are segmented using k-means clustering to partition the leaf images into four cluster in which one or more cluster contain the disease k-means uses squared Euclidean distances. The method is followed for extracting the feature set is called the Color Co-occurrence method, in this both color and texture of an image are taken into account to arrive at unique features. Second step the segmented images are passed through a pre-trained neural network. A set of leaf images are used for experimenting. The system is tested on five diseases which effect the plants; they are Early scroch, cotton mold, ashen mold, late scroch, tiny whiteness. Advantage of this system is that it detects plant disease based on both leaf and stem giving us 93% accuracy ,it detects at early stages because of sensor used in it,in this images are segmented using k-means cluster for proper extraction of feature and classified using Neural network based on back propagation algorithm. Disadvantage is that it can detect only five disease of plants,and the feature extraction may require lot of mathematical calculation which makes it complicated and for detection few sensors used in this system which may increase in cost.

YuanTian, ChunjiangZhao, ShenglianLu and XinyuGuo[5] proposed an SVM-based Multiple Classifier System (MCS) for wheat leaf diseases. It uses a stacked generalization structure to join the classification decisions obtained from three kinds of support vector machines (SVM) based classifiers. The features like color, texture and shape features are used as training sets for classifiers. Firstly, features are classified using a classifier in low-level of MCS to corresponding mid-level categories, which can partially detect the symptom of crop diseases according to the knowledge of plant pathology. Then the mid-level features are generated from these mid-categories generated from low-level classifiers. Finally, high-level SVM has been trained and correct errors made by the color, texture and shape SVM to improve the performance of detection. Compared with other classifiers, it can provide better success rate of detection. The classifiers like SVM Artificial Neural Network classifier, k-nearest neighbour (kNN) classifier's, the MCS can obtain better recognition accuracy than others classifiers. Color, texture and shape SVMs to improve the performance of detection. Compared with other classifiers, it can provide better success rate of detection. Segmentation is difficult here.

Suhaili Beeran Kutty et al in 2013, [6] have considered an artificial neural network based system to classify the watermelon leaf diseases of Downney Mildew and Anthracnose. The classification is based on color feature extraction from RGB color model where the RGB pixel color indices have been extracted from the identified Regions of Interest (ROI). The classification model involves the process of disease classification using Statistical Package for the Social Sciences (SPSS) and Neural Network Pattern Recognition Toolbox. This paper describes the analysis of watermelon leaf diseases by using image processing technique with respect to its mean value of RGB color component. Each group consists of images captured under specific requirement. The cropped data from each leaves images then were analyzed using SPSS software and then tested their performance by using Error Bar plot and Neural Network Pattern Recognition Toolbox. Determinations in this work have shown that the type of leaf diseases achieved 75.9% of accuracy based on its RGB mean color component. In order to improve the effectiveness of this classification system for the watermelon leaf diseases, it is recommended to use a high pixel of digital camera to get the best images. Also recommended to increase the number of data for the training and testing to get the best result. In addition, the lighting setup must be in proper position because it also can affect the image captured.

Ratih Kartika Dewi and R. V. Hari Ginardi [7] proposed Feature Extraction for Identification of Sugarcane Rust Disease. data collection which contain normal and diseased images, image pre-processing, feature extraction and classification. Real data of sugarcane images are collected from sugarcane fields survey . Observation is conducted by capturing photos of sugarcane leaf on a paper. There are 200 data of normal sugarcane leaves and 200 data of sugarcane leaves with rust disease. After collecting the image data, these images are pre-processed. Firstly, leaf area that represents rust disease are selected. To extract shape feature, the color image is processed to be a binary using Otsu thresholding Combination of color, shape and texture feature of an image are used in feature extraction. To extract color feature, we transform image from RGB to LAB color We use shape, texture and color feature extraction to identify rust disease in sugarcane. SVM is used to classify normal and diseased images. Analysis is conducted by evaluate combination of several features to find distinctive features for identification of rust disease. When single feature is used, shape has the lowest accuracy of 51% and texture has the highest accuracy of 96.5%. The combination of texture and color feature extraction results a highest classification accuracy of 97.5%. To improve the quality of proposed system data sets can be increased to get best results. Texture feature extraction is quite complex process.

P.R. Rothe and R. V. Kshirsagar in 2014 [8] proposed Adaptive Neuro-fuzzy Inference System for Recognition of Cotton Leaf Diseases develop and to evaluate adaptive neuro-fuzzy inference system . Initially the digital images are acquired using a digital camera. Then image preprocessing techniques are applied to these images to smooth the images. The object of interest from the image is separated by using the segmentation method and the features extracted are analyzed to determine the discriminating features. The neural network used for classification is trained with the help of features of the training data set. The images of the testing data set are used to determine accuracy of the system. The extraction process of color descriptor consists of four stages: Image partitioning, Representative color selection, DCT transformation and Zigzag scanning. Adaptive neuro fuzzy inference system is used for classification. It combines the principles of neural network and fuzzy logic therefore it provides advantages of both in a single structure .The premise parameters which defines the membership function are updated by using gradient decent method and consequent parameters are identified by using least square method. Digital image analysis combined with neural network has proven to be a feasible approach to the identification of cotton leaf disease. The results obtained have shown that the identification of Bacterial Blight, Myrothecium and Alternaria is feasible using color layout descriptors as features for training the anfis model. In order to improve the effectiveness of this classification system for the cotton leaf diseases, it is recommended to use a high pixel of digital camera to get the best images. Also Color descriptor process is quite complex and long process. This system can further extended on identification of diseases on other plants.

P. R. Rothe and R. V. Kshirsagar [9] proposed that a pattern recognition system for identification and classification can be done. An Active Contour model (Snake segmentation) technique is used for segmenting the diseased region from the cotton leaf. Hu's moments are used as the features for the classification. For training it uses a set of seven moments and for classification it uses Back Propagation Neural network. Advantage is that Back propagation neural networks are highly efficient for solving Multiple Class problems. Active contour model is used to minimize the energy inside the disease spot, BPNN solves the multiple class problems and average classification is found to be 85.52%. Snake segmentation is a very slow process which is a disadvantage here. It processes the images captured in natural conditions from varying distances. After that image segmentation techniques are used to isolate the diseased spots from background. Features are then extracted and utilized to train the network which carries out the classification. Noise in the images are removed by low pass filter. The proposed methods can be applied to other crops like orange, citrus, wheat, corn and maize etc.

Aakanksha Rastogi, Ritika Arora and Shanu Sharma [10] suggested a Fuzzy system for leaf disease detection and grading. K-means clustering technique has been used for segmentation which groups similar pixels of an image. RGB color space is converted to L*a*b space, where L is the luminosity and a*b are the color space. The reason for this conversion is that luminosity factor is not important for the color image. GLCM matrix including contrast, correlation, energy and homogeneity has been measured for disease grading. Artificial Neural Networks as been used for training the data. Fuzzy logic is used for grading the disease. Severity of the disease is checked. It is fast and highly efficient.. Disadvantage of this system is that it has low-level segmentation.

Mrunmayee Dhakate and Ingole A.B [11] suggested a Back-propagation Neural Network based classifier (BPNN) for detecting the disease in Pomegranate leaf. The work proposes an image processing and neural network method to deal with th e main issues of phytopathology i.e disease detection and classification. The experiment is performed n pomegranate fruit .the diseases are like bacterial blight ,fruit spot, fruit rot and leaf spot. It uses some images for training and testing purpose. It even uses k-means algorithm and feature extraction using GLCM method ,and given to the artificial neural network. The overall accuracy is 90%. Features have been selected as color and texture. BPNN detects and classifies the diseases. RGB image is converted to L*a*b to extract chromaticity layers of image, classification is found to be 97.30%.It is only applicable for limited crops.

Huu Quan Cap,Satoshi Kagiwada,Hiroyuki Uga and Hitosi Iyatomi[12] proposed leaf localization method from on-site wide angle images with deep learning approach. This method achieves a detection performance of 78.0% in F1-measure at 2.0 fps. The objective of this system is to localize the "fully leaf" part from the input image. "fully leaf" indicates the region contains almost whole leaf. "Not fully leaf" and "None leaf" are the regions that contain part of a leaf and none leaf object respectively. Leaf detection consists of three steps. Firstly, given a wide-angle image, the system extracts a numerous candidate boxes that may contain fully leaf regions. Secondly, specially trained CNN classifier analyzes those boxes to find location of fully leaf. Finally, the non-maximum suppression is used to remove overlapping bounding boxes . The detected fully leaf regions will be diagnosed by an external diagnosis system. The images used here are wide angle images and takes 0.5 seconds to detect all the leaves per image. Full leaf part is considered for disease detection and other leaf parts are discarded. For study they have considered healthy leaves. This approach is affective only for private and small scale facilities.

3. PROPOSED APPROACH

The proposed approach starts with the training of images for both the samples such as healthy and diseased leaf images. Once the database is acquired of healthy and infected images of samples the feature is extracted. The canny edge algorithm is used for feature extraction i.e. to detect whether the particular image is healthy or infected. It is then compared with the images obtained by remote sensing that are taken periodically. If the particular plant is not infected message is sent to the farmer that the crop is healthy. If the crop is infected then KNN algorithm is used for identifying particular plant disease and intimate with an alert message to farmer using android application. The proposed method has two phase. The first phase deals with training the model for healthy and as well as diseased images and second phase deals with monitoring of crops and identification of particular disease using KNN algorithm and also intimate the agriculturists with an early alert message immediately.

3.1 Training model

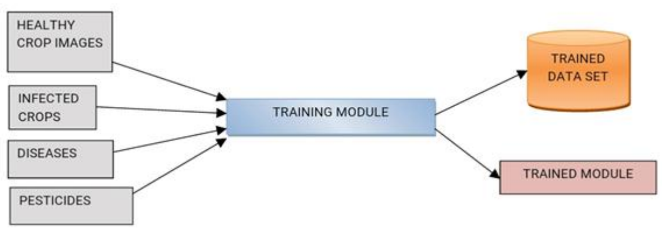

Fig-1: **Trained Model**

This module is the basic building block of the system. This model will used in the disease detection model for taking the proper decision about the diseases. The different crop types along with the diseases and pesticides are given as the input to training model. For each crop, plenty of Healthy and defected crop images are considered. The module will be trained in a such a way that it will be easy to take the proper decision for all types of crops. Here the canny edge algorithm is used for training the model. Fig.1 shows that Healthy crop images, Infected crop images, different diseases of different Crops and Pesticides for that diseases are all considered to get trained model. The module is trained to identify whether the particular image is healthy or infected and also to give all the possible diseases and pesticides for that particular crop if it is infected .

3.2 Disease Detection

Fig-2: **Decision Support System**

Reference image will be given as input to Decision module. This module makes use of trained module to take the further decision. Figure-2 shows the reference image as input and also output from trained module. Trained module uses the dataset to provide a proper suggestion. If there is no disease is detected then the message is sent to the farmer that the crop is healthy. If not, KNN algorithm is used for disease detection. Once the disease is identified , the name of the disease and the amount of pesticide to treat that disease is intimated to the farmer using an android application.

3.3 System architecture

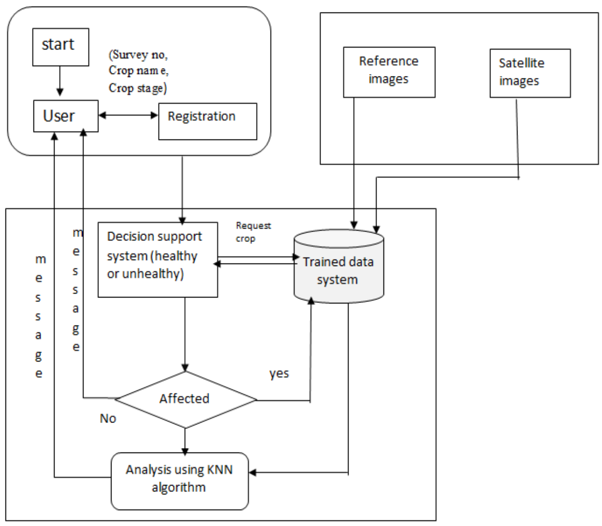

Fig-3: **System Architecture.**

Figure-3 shows the system architecture. Some of the terms used here are:

- Satellite images: These are the images of the plants that are taken periodically from the farmers land using satellite.
- Reference images: These are all the possible healthy and diseased images of the plant that are used for training the model.
- Trained data system: This is the trained model referred in fig 1 which is used to identify whether the plant is healthy or diseased.
- Decision support system: This is the model referred in fig 2 which is used to detect the disease of the plant.

User registers using land number , crop name and crop stage. If the user has already registered then he can login to the application using login credentials. The trained data system is used to identify whether the plant is healthy or diseased. The canny edge algorithm is used to identify whether the plant is healthy or diseased. If the crop is healthy the message is displayed to the farmer that the crop is healthy. If the crop is unhealthy then in decision support system the KNN algorithm to used to identify the crop disease.

Once the disease is detected then a message is sent to farmer along with the disease predicted and the amount of pesticide to be used.

3.4 Algorithm

3.4.1 Canny Edge Detection ALGORITHM

The Canny edge detection algorithm is composed of 5 steps:

1. Noise reduction: To get rid of the noise on the image, is by applying Gaussian blur to smooth it.

2. Gradient calculation: The Gradient calculation step detects the edge intensity and direction by calculating the gradient of the image using edge detection operators.

3. Non-maximum suppression: non-maximum suppression is performed to thin out the edges.

4. Double threshold: It aims at identifying 3 kinds of pixels: strong, weak, and non-relevant.

5. Edge Tracking by Hysteresis: Based on the threshold results, the hysteresis consists of transforming weak pixels into strong ones, if and only if at least one of one.

3.4.2 K-Nearest Neighbor ALGORITHM

The K-nearest neighbor (KNN) classifier is a non parametric classifier because it makes no underlying assumptions about the statistical structure of data. The KNN algorithm measures the distance between a test query features and a set of training data features that store in the database. The distance between these two feature vector is estimated using a distance function d(x,y), where x, y are feature vector can be represents as

X = {x1, x2, x3, … .}
Y = {y1, y2, y3, … .}

The feature vectors must be normalized before classification algorithm run. The overall KNN algorithm is running in the following steps:

1. KNN algorithm uses training set that consist features and labeled classes store in database.
2. Calculate distance between test features and all training feature.
3. Sort the distance and determine k nearest neighbor.
4. Use simple majority of the category of nearest neighbor assign to the test.

The KNN classification algorithm is performed by using a training set which contains both the input feature and the labeled classes and then by comparing test feature with training feature a set of distance of the unknown K nearest neighbors determines. Finally test class assignment is done by either averaging the class numbers of the K nearest reference points or by obtaining a majority vote for them.

4. CONCLUSION

The proposed system implements an innovative idea to identify the affected crops and provide remedy measures to the agriculture industry. Recognizing the disease is the main purpose of our proposed system. The images are fed to the application for identification of diseases. Feature extraction technique helps to extract the infected leaf and to classify plant diseases. Machine learning techniques are used to train the model which helps to take a proper decision regarding the diseases. The pesticide as a remedy is suggested to the farmer for infected diseases to control it. Thus this system would be useful for saving farmers from huge loss.

REFERENCES

[1] Leninisha Shanmugam, A. L Agasta Adline , N Aishwarya, "Disease detection in crops using remote sensing images", 2017 IEEE Technological Innovations in ICT for Agriculture and Rural Development (TIAR), 2017

[2] International Research Journal of Engineering and Technology, "Crop disease detection using Machine Learning: Indian Agriculture", Anuradha Badage, no.100, 2018.

[3] Santanu Phadikar and Jaya Sil,"Rice Disease Identification using Pattern Recognition Techniques",2008 International Conference on Computer and Information Technology.

[4] Dheeb Al Bashish, Malik Braik, and Sulieman Bani-Ahmad, "A Framework for Detection and Classification of Plant Leaf and Stem Diseases", IEEE 2010.

[5] Yuan Tian,Chunjiang Zhao, Shenglian Lu and Xinyu Guo," SVM-based Multiple Classifier System for Recognition of Wheat Leaf Diseases," Proceedings of 2010 Conference on Dependable Computing (CDC'2010),November 20-22, 2010.

[6] Suhaili Beeran Kutty, Noor Ezan Abdullah, Dr. Hadzli Hashim, A'zraa Afhzan Ab Rahim, Aida Sulinda Kusim,Tuan Norjihan Tuan Yaakub, Puteri Nor Ashikin Megat Yunus, Mohd Fauzi Abd Rahman "Classification of Watermelon Leaf Diseases Using Neural Network Analysis" 2013 IEEE Business Engineering and Industrial Applications[BEIAC].

[7] Ratih Kartika Dewi and R. V. Hari Ginardi," Feature Extraction for Identification of Sugarcane Rust Disease",

International Conference on Information, Communication Technology and System, 2014.

[8] P.R. Rothe and R. V. Kshirsagar, "Adaptive Neuro-fuzzy Inference System for Recognition of Cotton Leaf Diseases" International Conference on Innovative Applications of Computational Intelligence on Power, Energy and Controls with their Impact on Humanity (CIPECH14), (2014) .

[9] P. R. Rothe and R. V. Kshirsagar," Cotton Leaf Disease Identification using Pattern Recognition Techniques", International Conference on Pervasive Computing (ICPC),2015.

[10] Aakanksha Rastogi, Ritika Arora and Shanu Sharma," Leaf Disease Detection and Grading using Computer Vision Technology &Fuzzy Logic" 2nd International Conference on Signal Processing and Integrated Networks (SPIN)2015.

[11] Mrunmayee Dhakate and Ingole A.B "Diagnosis of Pomegranate Diseases Based on Back Propagation Neural Network," International Research Journal of Engineering and Technology (IRJET), Vol2 Issue: 02 | May-2015.

[12] Huu Quan Cap,Satoshi Kagiwada,Hiroyuki Uga,Hitosi Iyatomi,"A Deep Learning Approach for on-site Plant Leaf Detection" 2018