Quality Assessment of Crops using Machine Learning Techniques

Tenzin Chokey[1], Sarika Jain[2]

[1,2]Amity Institute of Information Technology, Amity University Uttar Pradesh
[1]tenzinchokey56@gmail.com, [2]ashusarika@gmail.com

Abstract: Disease in a crop leads to low productivity and in turn leads to huge loss to the farmers. Thus, detection of disease in early stage will be beneficial for farmer so that necessary actions can be taken. This paper discusses supervised machine learning techniques to detect the disease in the plant with the help of the image of the plant. The comparison between classifications techniques are made in order to select model with highest accuracy. The Quadratic SVM results with highest accuracy of 83.3%. This trained model is used for the detection of new disease image.

Keywords: Maize leaf Diseases; Classification; Quality assessment

I. INTRODUCTION

Agriculture is one of the economic source and crucial part of India. There is huge advancement in agriculture sector due to the technology. But farmers are facing loss of crop production due to various factors which leads to even suicidal cases in farmers. Traditionally, quality assessments of crops are based on visual inspection. This results in consumption of time, cost and inaccuracy. Because of this reason researchers had provided solutions through the technology for quality assessments. Previously, image processing was used for the detection of diseases in crops. This process needs image with high resolutions which was difficult to capture and accuracy was a major issue. Image processing along with classification techniques has improved its efficiency and accuracy in detecting the crops diseases. This study aims to develop a system that detects a disease of a leaf at early stage uses machine learning techniques.

In this paper, Maize plant is taken for studying the various diseases. Since maize is third major crop cultivated in India, it has multiple uses in food industry and animal fodder. Usually it is cultivated in Karnataka, Andhra Pradesh, Tamil Nadu, Rajasthan, Maharashtra, Bihar, Uttar Pradesh, Madhya Pradesh and Gujarat. Generally, Maize is sown in the beginning of the monsoon season and harvested in the month of autumn. In case of Tamil Nadu rainfall and the optimum temperature is up to 30°C. In this work, various supervised learning algorithms namely Linear Support Vector Machine (SVM), Medium tree, quadratic SVM and cubic SVM are applied and a comparison is drawn.

state, it is considered as Rabi crop which is sown in winter season. For the proper growth it requires 50-100 cm of Symptoms.

A. TYPES OF DISEASES IN MAIZE AND SYMPTOMS

In Maize, disease are caused by various Pathogens such as bacteria, virus and fungal. These Pathogens mostly attack leaf, stalk and fruit. Some of the main diseases are:

1) Common Rust (Fig. 1)

Pathogen: Fungus (Puccinia sorghi)
Infected area: Leaf

- Presence of brown to red Pustule on both surfaces of leaf.
- The Pustule are spread throughout the leaves and it leads to the ruptured of leaf tissues.

Management and Control:

- Use of Fungicide sprays like copper, oxychloride, Folicur, Amistar and Bravo.
- Use of Hybrid seed

Fig. 1. Common Rust

2) Fusarium Ear Rot (Fig. 2)

Pathogen: Fungus
Infected area: Ear

978-1-5386-9346-9/19/$31.00 ©2019 IEEE

Symptoms:

- Development of powdery white or pink mold.
- Discoloration of Kernel into salmon pink to reddish

Management and control:

- Grow seed co tolerant hybrid
- Use insecticide in order to prevent various insects.

Fig. 2. Fusarium Ear Rot

3) Common Smut or Boil Smut (Fig. 3)
Pathogen: Fungus
Infected area: Kernel
Symptoms:

- Appearance galls in place of kernel.
- Fungus attacks stalk, leaves and ears.
- It affects the plants height.
- Galls release bespores and infect maize plant.

Management and Control:

- Grow hybrid seed
- Careful handling of gall before it ruptures.

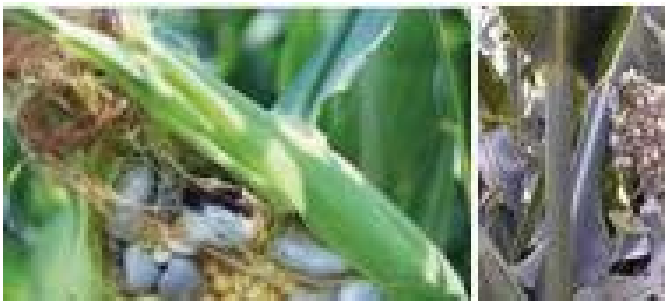

Fig. 3. Common Smut or Boil Smut

II. LITERATURE REVIEW

Ishak*et al.* [1] states classification of Phyllanthus Elegant Wall leaf disease into healthy or unhealthy using ArtificialNeutral network. Images are acquired from herb plant and using image processing algorithms for the colour transformation structure. The classification of image is based on the colour and area of the leaf. A comparison is made between Multi-layer Perceptron (MLP) and Radial Basis Function (RBF). Valenzuela *et al.* [2] considered the features from input image of lettuce as the dataset for training back propagation neural network. The trained network classifies the image with and without defects by outputting 1 for defect lettuce and 0 otherwise with minimum relative error of 0.051. Features of leaf can be obtained from colour, texture and shape of the leaf so that type of disease can be detected using classification algorithms and content based image retrieval system. Padol and Yadav [3] have used Linear Support Vector Machine classification method to classify the type of leaf disease. Firstly image pre processing techniques are applied for input images of grapes and disease regions are detected with k-means clustering from which colour and textures are extracted. SVM performed better with of of 93.33% whereas with powderdly class grape leaf image of accuracy 83.33%. Mohanty*et al.* [4] used deep convolution neural network to detect leaf disease. The author has taken dataset of 54,306 images of diseased and healthy plant leaves from plant village datasets. The three versions of data sets are experimented i.e. color, grayscale, segmented. The model was able to identify 14 crop species and 26 diseases with accuracy of 99.35%. Dandawate and Kokare [5] aimed to detect soy bean leaf disease using image support vector machine. Scale Invariant Feature Transform (SIFT) technique were used to automatically recognise plant based on its shape. The framework provides input to the Decision support System which assists the farmer in minimal efforts over the internet. The proposed method could detect the soybean disease with accuracy of 93.79%. [6]Singh andMisra presented automatic detection of leaf disease usinggenetic algorithms. The proposed method includes input image is pre-processed by image processing techniques and segment using genetic algorithms to classify the leaf disease. The optimal result is obtained with very less computational efforts. The author suggested use of ANN, fuzzy logic and hybrid algorithms for improving the recognition rate in classification process. Patil and Kumar extracted features based on color of diseased tomato leaf. Image is partitioned into green, red and blue components. This extracted feature can be used for the classification of disease. Rushika and Juilee [8] focused on Maharashtra farmers to assist them with suitable crop based on their soil quality. They used machine learning for the prediction of suitable crop. It analyzes the nutrient present in the soil and predicts the crop yield in particular location. Zhihao and Kalbarczyk [9] have proposed a data driven model for developing precision agriculture solution. Soil moisture prediction model is developed to predict the moisture ahead of time based on the environmental attributes and same soil. The prediction output will be accurate over a long period of time if the reliable data are fed into the model. Dahikar and Snehal [10] used artificial neural network for crop prediction based on

parameters like temperature, PH, nitrogen, potassium, phosphate andrainfall. Based on deficiency in nitrogen, phosphate and potassium, it suggests an appropriate fertilizer.

III. PROPOSED METHODOLOGY

Fig. 4 depicts the architecture of proposed methodology. The proposed algorithm has been implemented using classification learner application in MatLab.

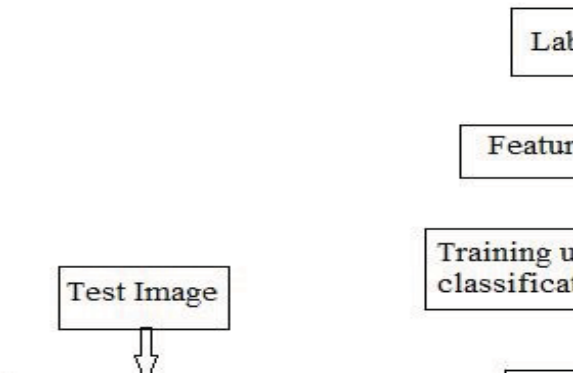

Fig. 4. Flow chart of proposed methodology

Step 1: The flow chart follows supervised machine learning algorithms. Thus, labeled images of three diseases are taken as an input for training.

Step 2: Features extraction are fundamental for any recognition algorithm. From input image, Speeded Up Robust Features (SURF) feature extractor extracts features from all images in all classes and constructs visual vocabulary called bag of features that minimizes the number of features using K-means clustering.

Step 3: The extracted features data are used for training different models with Classification Learner Application. Cross validation helps in evaluating machine learning algorithms for prediction of new datasets. This is achieved by splitting data into training set and testing set. Once the datasets are trained, based on the accuracy of different techniques, select the best model for testing. The export model will be used for testing new input image for the detection of disease.

Step 4: Again input images without a label. Image feature extraction will be followed the same as training. The trained model will predict the corresponding diseases in the plant with the help of images.

To make predictions of unlabelled data, following command is used in command window:

yfit=trainedModel.predictFcn(T), where
T is a data that to be tested.
The output of yfit predicts the disease of an image.

IV. RESULTS AND DISCUSSION

The feature extracted from unlabeled data images will be input to the trained model. The model predicts the disease of an image accordingly.

TABLE 1: COMPARISON BETWEEN CLASSIFICATION TECHNIQUES

Parameter\Model	Linear SVM	Medium tree
Accuracy (%)	79.5	73.1
Prediction Speed	~440	~470

Confusion matrix defines the accuracy of the classification models based on each class that has been taken. It shows the true positive rate and true negative rate of the classification model. From this matrix, we can also analyze which classification model depicts the best for particular diseases.

The x-axis of matrix defines predicted class and y-axis depicts the true class. The diagonal cells correspond to the observation of correctly classified disease. Figures 5-8 show the Confusion matrix of all of the classification models.

Fig. 5. Confusion matrix of Linear SVM

Fig. 6. Confusion matrix of Medium Tree

Fig. 7. Confusion matrix of Quadratic SVM

Fig. 8. Confusion Matrix of Cubic SVM

The figures 9-12 show the Receiver Operating Characteristic (ROC) of linear SVM, Medium Tree, Quadratic SVM, and Cubic SVM. Since, accuracy depends on how well the models are able to distinguish between the various diseases. The ROC graph helps in visualizing the performance whereas area under curve(AUC) summarize the overall performance with a single value.

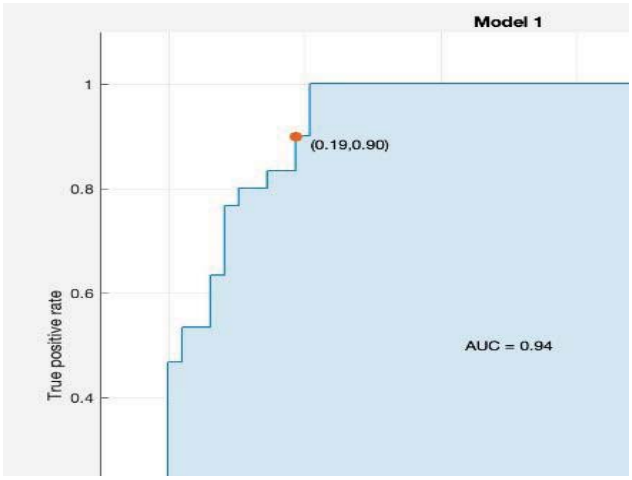

Fig 9. ROC curve of linear SVM

Fig. 10. ROC curve of medium tree

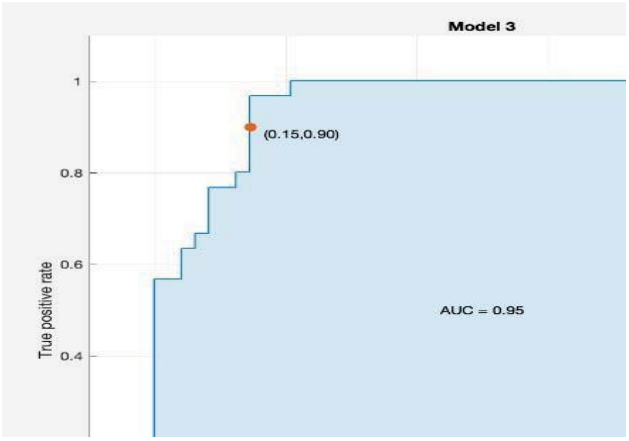

Fig. 11. ROC curve of quadratic SVM

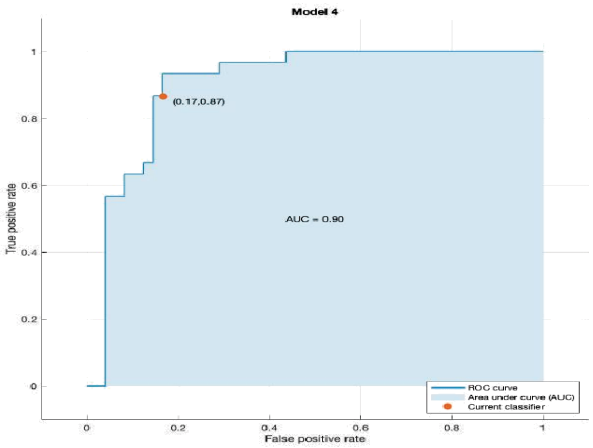

Fig. 12. ROC curve of cubic SVM

Since Quadratic SVM results with highest accuracy during training, this model will be exported for testing or prediction of new image datasets. Unlabeled disease images will be taken as an input and accordingly type of disease is detected as shown in Fig. 13

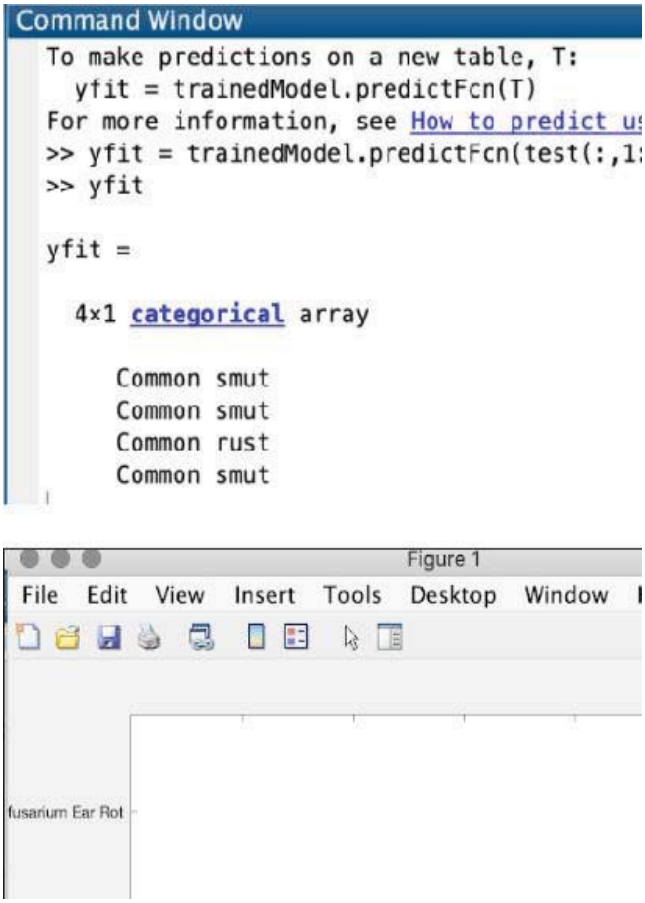

Fig. 13. Predicted Output

V. CONCLUSION

Supervised machine learning techniques namely, linear SVM, medium tree, quadratic SVM and cubic SVM are used to detect the various diseases in Maize leaf. Initially, labeled image data is used to train the classification model and highest accuracy rate is selected for testing the new image data for disease detection. Based on detected disease farmer can take necessary action at the earliest. Quadratic SVM gives the best result in detecting the disease of maize.

REFERENCES

[1] Ishak, S., Rahiman, M. H. F., Aqmariah, S. N., Kanafiahb, M., & Saadc, H, "Leaf Disease Classification Using Artificial Neural Network" *Jurnal Teknologi*, 77(17), pp. 109-114, 2015.

[2] Valenzuela, I. C., Puno, J. C. V., Bandala, A. A., Baldovino, R. G., de Luna, R. G., De Ocampo, A. L., Cuello, J. & Dadios, E. P, "Quality assessment of lettuce using artificial neural network", *IEEE 9th InternationalConference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment and Management (HNICEM),* 2017, pp.1-5.

[3] Padol, P. B., & Yadav, A. A,"SVM classifier based grape leaf disease detection", *IEEE Conference* on*Advances in Signal Processing (CASP),* 2016, pp. 175-179.

[4] Mohanty, Sharada P., David P. Hughes & Marcel Salathé, "Using Deep Learning for Image-Based Plant Disease Detection," *Frontiers in Plant Science* 7, 2016.

[5] Dandawate, Yogesh, and Kokare, Radha "An automated approach for classification of plant diseases towards development of futuristic Decision Support System in Indian perspective," *IEEE International Conference onAdvances in Computing, Communications and Informatics (ICACCI),* ISNB no. 978-1-4799- 8792-4, 2015 , pp.794-799.

[6] Singh, Vijai, and Misra, A. K. "Detection of unhealthy region of plant leaves using image processing and genetic algorithm," *IEEE International Conference on Advances inComputer Engineering and Applications (ICACEA),* 2015,ISBN no 978-1-4673-6911-4, pp. 1028-1032.

[7] Patil, Jayamala K., and Kumar, Raj "Color feature extraction of tomato leaf diseases," *International Journalof Engineering Trends and Technology*, 2(2), pp.72-74,2011.

[8] Ghadge, Rushika, Juilee Kulkarni, Pooja More, Sachee Nene, and Priya RL, "Prediction of Crop Yield using Machine Learning," *International Research Journal ofEngineering and Technology (IRJET)*,Vol.05 ,Feb.2018.

[9] Hong, Z., Kalbarczyk, Z. and Iyer, R.K,"A Data-Driven Approach to Soil Moisture Collection and Prediction,"

[10] *IEEE International Conference on Smart Computing (SMARTCOMP),* 2016, pp. 1-6.

[11] Dahikar, Snehal S., Sandeep V. Rode, and Pramod Deshmukh "An Artificial Neural Network Approach for Agricultural Crop Yield Prediction Based on Various Parameters," *International Journal of Advanced Researchin Electronics and Communication Engineering (IJARECE)* 4.1,2015.