Detection of Cotton Leaf Diseases Using Image Processing and Machine Learning Approach

K. Praveen Chowdhary¹,², Sainath³, Yaratapalli Sukumar Reddy³, M. Yashas Kumar⁴, B. G. Girish⁵  
¹,²,³,⁴Student, Department of Computer Science and Engineering, S. J. C. Institute of Technology, Chikkaballapur, India  
⁵Assistant Professor, Department of Computer Science and Engineering, S. J. C. Institute of Technology, Chikkaballapur, India  
*Corresponding author: pchowdharyk@gmail.com

Abstract: Cotton crop is an important crop all over the world to farmers. The problem faced by our farmers is the diseases that affects the Cotton crops. So the farmers need to know the diseases attacked on the crops and take necessary measures to avoid poor yield. The proposed work presents an image processing and machine learning approach for identification and classification of cotton leaf diseases such as Bacterial Blight, Leaf Crumple and Alternaria. The project is designed to detect different types of diseases which can be further used to find other diseases that may affect various crops. Disease is identified and the farmer can precautionary measures can be taken to save the cotton yield. Through this system the detection of the diseases by analyzing the different symptoms of the disease is more cost efficient.

Keywords: Cotton leaf diseases, Machine Learning, Image Processing, K-Means, ANN, KNN, SVM.

1. Introduction

Cotton is an important cash crop in India and affects India’s economy in many ways. As India is a land of irregular climates the farmers will have a number of different crops for example food, cash, plantation, Horticulture and many more. Many new technologies can help in yielding the crops efficiently although with the advancement of the technology it requires the constant monitoring by the producers. Huge number of people depends on cotton crop by its cultivation or processing.

Cotton crops can be attacked by various types of diseases like Alternaria alternata, Bacterial blight, crown gall, leaf crumple, cotton rust and worm attack which decreases the yield. The existing methods of prediction of the diseases are not accurate as the experts may get wrong with prediction. For doing so, a large team of experts is required, and experts use their previous knowledge and provide solutions which are less accurate which costs very high when we do with large farms. So the proposed system methods such as image processing and machine learning techniques are used to detect and classify cotton leaf diseases. The system includes the image processing techniques such as the K-Nearest Neighbor(KNN), Support Vector Machine(SVM) and Artificial Neural Network(ANN) are implemented to get the accurate results.

2. Literature Survey

P. Revathi, M. Hemalatha, in "Classification of Cotton Leaf Spot Disease Using Image Processing Edge Detection Technique" presents mobile captured symptoms of cotton leaf spot images and categorize the diseases using HPCCDD Proposed Algorithm. The classifier is trained to achieve intelligent farming, including early Identification of diseases in the groves, selective fungicide application, etc. [1]

Nikhil Shah and Sarika Jain in “Detection of Disease in Cotton Leaf using Artificial Neural Network” presents a machine learning based approach is proposed which can test the image of the leaf of plant and detect the disease and the quality of the cotton plants using machine learning approach. The main method of the research is to detect the different diseases of cotton by applying artificial neural network tool, which apply image pre-processing method to process the image and based on colours changes on the image the main portion of the affected leaf is highlighted and detect the type of disease based on data [2].

P. R. Rothe and R. V. Kshirsagar in “Cotton Leaf Disease Identification using Pattern Recognition Techniques” presents classification using Back propagation neural network where the training is performed by extracting seven invariant moments from three kinds of textured leaves images. The average accuracy of classification is found to be 85.52%. The snake segmentation algorithm provides efficient technique to isolate the diseased spot but is a very slow process. This results in longer training and testing phase for the system. Various other promising features can be added in feature extraction process for making the system more robust. This helps in increasing the performance of the system [3].

Ratih Kartika Dewi and R. V. Hari Ginardi.in “Feature Extraction for Identification of Sugarcane Rust Disease” presents an efficient method for image pattern classification in sugarcane rust disease. Firstly, normal and diseased leaf images are collected and pre-processed. Then, features of color and
texture are extracted from these images. Shape feature extraction used in this paper are solidity, extent, minor axis length and eccentricity. Texture feature extraction used in this paper are contrast, correlation, energy and homogeneity. Skewness and kurtosis are used to represent color as features. Images are captured in RGB format. RGB cannot represent color in terms of human perception. We transform image color from RGB to LAB. After that, these images are classified using support vector machine [4].

Ajay A. Gurjar and Viraj A. Gulhane in “Disease Detection On Cotton Leaves by Eigen feature Regularization and Extraction Technique” presents an approach that regularizes and extracts eigenfeature from cotton leaf image. Here scatters matrix is developed which is within class type, now this matrix is decomposed into various subspaces, related to various diseases i.e. fungal disease, leaf crumple and by providing the thousand no of sample images, i.e. here considering the various variation of pixel value. Eigen features are regularized differently in these subspaces based on that eigen spectrum is modeled this enables the discriminant evaluation performed in the whole space feature extraction or dimensionality reduction occurs at the final stage, after comparison of this feature results to disease identification [5].

Prof. Sanjay B. Dhaygude and Mr. Nitin P. Kumbhar in “Agricultural plant Leaf Disease Detection Using Image Processing” presents application of texture statistics for detecting the plant leaf disease has been explained. Firstly by color transformation structure RGB is converted to HSV space because HSV is a good color descriptor. Masking and removing of green pixels with pre-computed threshold level. Then in the next step segmentation is performed using 32X32 patch size and obtained useful segments. These segments are used for texture analysis by color co-occurrence matrix. Finally, if texture parameters are compared to texture parameters of normal leaf. The extension of this work will focus on developing algorithms and Neural Networks in order to increase the recognition rate of classification process [6].

Namrata R. Bhimte and V. R. in “Diseases Detection of Cotton Leaf Spot using Image Processing and SVM Classifier” presents image processing techniques along with segmentation and classification techniques are used for identification of cotton leaf diseases such as Bacterial blight and Magnesium Deficiency done using SVM. Colour segmentation where RGB image is changed to l * a * b color space by using k-means clustering to extract diseased part (area of interest) from leaf images with efficient features. Features such as color, shape and texture are useful for pattern recognition, classification, accurate and error free classification are calculated. Future work will consist of developing a more efficient, robust machine vision system for early automatic detection of various type of diseases in plants [7].

3. Proposed Approach
The proposed System uses image processing and machine learning techniques to detect and classify cotton leaf diseases which is efficient when we want to work with large farms.

A. Modules
The individual modules that compromise the building blocks of the system are distinguished and are discuss below:

1) Image Acquisition
Creating a dataset of Cotton leaf disease images, such that at least 3 different type of cotton leaf disease, more than 40 images of each disease are collected and stored.

2) Image Pre-Processing
The Background subtraction of the image is done using K-Means Segmentation in order to avoid the background noise for feature extraction. Colored image of leaf is converted into grey scale image. And then adaptive histogram equalization is applied.

3) Feature Extraction
This module deals with extracting Features from the Cotton leaf image. Feature vector is obtained. This is fed as input to classifier.

4) Classification
Based on the extracted features, KNN SVM or ANN classifier returns matched Cotton leaf disease type. Cotton leaf disease name will be displayed to user

![Fig. 1. Flowchart of proposed approach](image)

B. Implementation
The Implementation is done by running the Cotton leaf disease detection system (GUI) module as follows:

1) Platform: MATLAB
MATLAB is an interactive programming environment for scientific computing. MATLAB is heavily used in many technical fields for data analysis, problem solving, and for experimentation and algorithm development

2) Steps of Implementation
- The image is to be loaded from the system into the Cotton leaf disease detection system
- Then the read image is stored in a matrix for image
processing.
- The system checks whether the input image is cotton leaf or not.
- Background subtraction through K-means is done on the cotton image to isolate the diseased part and displayed in the GUI.
- Histogram Equalization is applied to the Image to improve contrast.
- Then extract the Color, Shape and Texture features and store it as Dataset which can be further used to classify the image.
- For the test image steps above steps are repeated. Machine Learning techniques such as Multi class support vector machine, K Nearest Neighbor Algorithm and Artificial Neural Network are used to classify the cotton leaf. Based on the extracted features in the dataset are compared with Test image features. The Disease type is detected and possible remedies to be taken are provided.

4. Results
Classification is done using the classifiers such as Support Machine, K Nearest Neighbor or Artificial Neural Network. Disease is identified and the farmer can take precautionary measures to save the cotton yield.

Table 1
<table>
<thead>
<tr>
<th>Algorithms</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>ANN</td>
<td>90%</td>
</tr>
<tr>
<td>KNN</td>
<td>85%</td>
</tr>
<tr>
<td>SVM</td>
<td>70%</td>
</tr>
</tbody>
</table>

5. Conclusion
In this paper, cotton leaf diseases are classified using image processing and machine learning techniques are used to detect and classify cotton leaf diseases. The system includes the image processing techniques such as, the K-Means Segmentation technique for Background subtraction and Adaptive Histogram Equalization for feature extraction. Classification algorithms such as the KNN (K-Nearest Neighbor), SVM (Support Vector Machine) and ANN (Artificial Neural Network) are implemented to get the accurate results. For Support Vector Machine 70% accuracy is achieved and using K-Nearest Neighbor Classifier 85% accuracy is achieved and for Artificial Neural Network 90% accuracy is achieved. The work presents the techniques that can be used for cotton leaf species identification and disease classification. This provides a mechanism to identify the disease of the species through just an image of the leaf.

References