JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

Advancements in Disease Recognition and Characterization in Horticulture: A Study Utilizing K-Means and Multi-Class SVM

Jitendra Kumar Katariya[1] , Kunal Kumar[2], Sneha Jain[3], Rajnish Kumar[4], Aman Kumar Singh[5], Abhinav Kumar[6]

Scholar[2,3,4,5,6] , Assistant Professor[1]

Department of Computer Science and Application, Vivekananda Global University , Jaipur

Abstract : This research delves into enhancing the modified structure for disease recognition and characterization in horticulture. Over the past few decades, various societies have investigated different aspects of plant pathology. This study particularly focuses on analyzing leaf-related images. Utilizing the K-Means algorithm, the concentration of infections on leaves is determined. The system not only identifies infected and healthy leaves but also distinguishes between them. Disease determination employs a Multi-Class SVM-based classifier, utilizing an information database derived from extensive leaf analysis. The system successfully identifies diseases such as Alternaria Alternata, Anthracnose, Bacterial Blight, and Cercospora Leaf Spot, achieving an impressive accuracy level of 96 to 97 percent, surpassing previous approaches in performance.

IndexTerms **Disease recognition, Characterization, Horticulture, K-Means, Multi-Class SVM, Leaf analysis, Alternaria Alternata, Anthracnose, Bacterial Blight, Cercospora Leaf Spot, Accuracy.**

I. INTRODUCTION

Disease recognition and characterization in horticulture [1] are critical components of modern agricultural practices, aiming to mitigate the detrimental impacts of plant diseases on crop yield and quality. With the global population on the rise, ensuring food security has become increasingly paramount, making the effective management of plant diseases a top priority for agricultural researchers and practitioners [1]. Horticultural crops, including fruits, vegetables, and ornamental plants, are susceptible to a wide array of diseases caused by various pathogens such as fungi, bacteria, and viruses. Therefore, developing accurate and efficient methods for early detection and classification of plant diseases is essential for implementing timely interventions and minimizing yield losses [1].

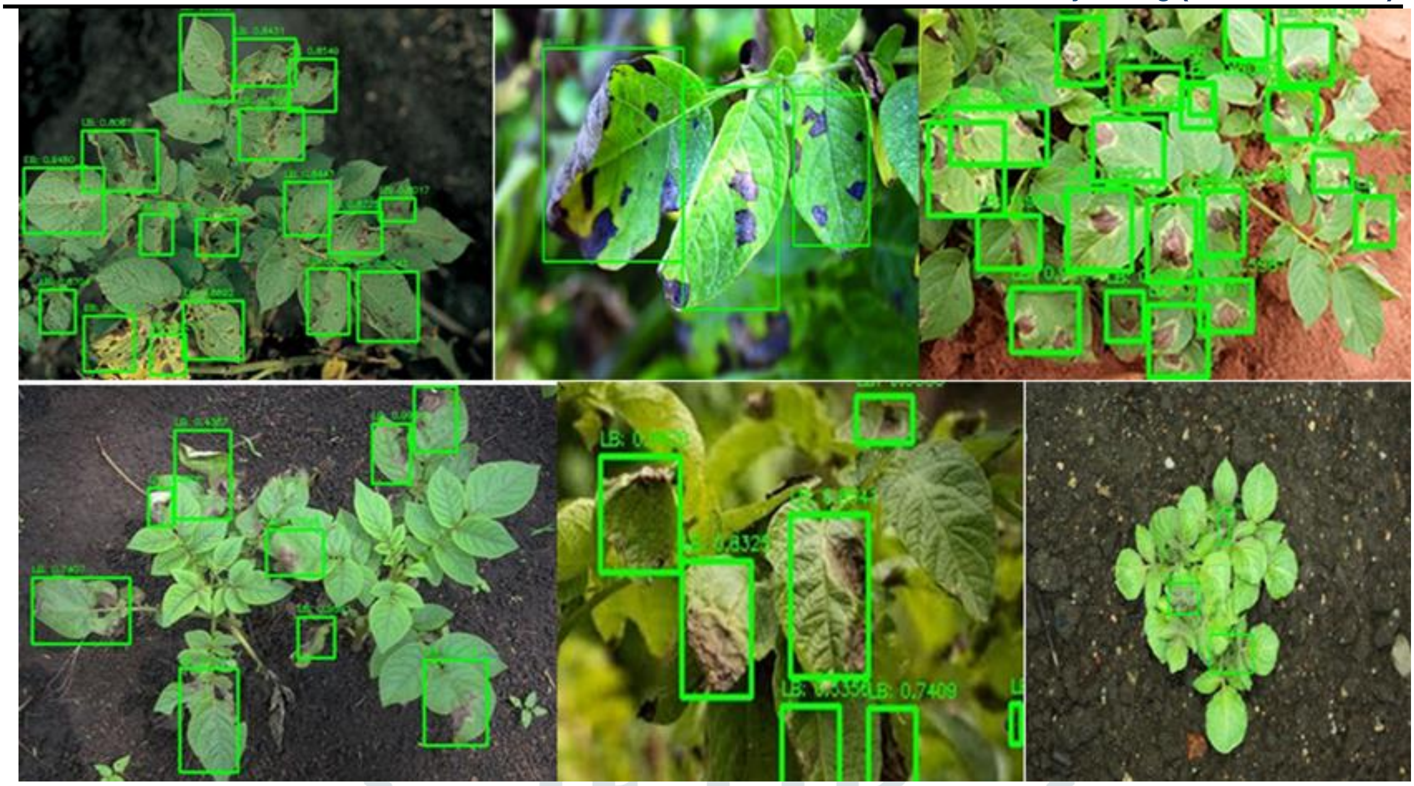

Fig 1.1. Disease Detection

Over the years, numerous approaches have been proposed and implemented for disease recognition and characterization in horticulture [2]. Traditional methods often relied on manual inspection by trained experts, which can be time-consuming, labor-intensive, and subjective. Additionally, visual symptoms of diseases may not always be distinct or easily identifiable, leading to misdiagnosis and ineffective management strategies [2]. While advancements in technology, particularly in imaging and machine learning, have enabled the development of automated disease recognition systems, many existing approaches still face certain limitations [2].

Previous studies have primarily focused on utilizing image processing techniques and machine learning algorithms to analyze digital images of plant leaves for disease diagnosis. While these approaches have shown promise in accurately identifying specific diseases based on visual symptoms, they often lack scalability and robustness across different environmental conditions and plant species. Furthermore, the reliance on handcrafted features and limited training data may hinder the generalization ability of these models, leading to suboptimal performance in real-world scenarios [3].

The primary objective of this study is to improve upon existing methods for disease recognition and characterization in horticulture by proposing a modified structure that integrates advanced machine learning techniques. Specifically, the study aims to address the limitations of previous approaches by leveraging the power of unsupervised learning with the K-Means algorithm to identify and quantify the concentration of infected areas on plant leaves [4]. Additionally, the study seeks to enhance disease classification accuracy by employing a Multi-Class Support Vector Machine (SVM) [4]based classifier, trained on a comprehensive information database derived from extensive leaf analysis. By achieving higher accuracy levels and improved robustness, the proposed approach aims to facilitate more effective disease management strategies and contribute to the overall sustainability and resilience of horticultural systems [4].

II. Literature Review

Chouhan et al. [5] emphasize the criticality of plant disease management for both human sustenance and environmental health, discussing the impact of diseases on various parts of plants and the necessity for accurate identification and treatment. They propose a method called BRBFNN for automated identification and classification of plant leaf diseases, achieving higher accuracy compared to existing methods.

Guo et al. [6] focus on image segmentation as a crucial step in plant disease diagnosis. They propose a technique called SFLA-PCNN, which enhances the performance of PCNN through parameter optimization, aiding in effective extraction of diseased regions from images.

Sardogan et al.[7] stress the importance of early detection of diseases for agricultural productivity, particularly for tomato plants affected by various diseases. They advocate for customized disease classification systems to facilitate prompt action upon detecting symptoms.

Sarkar and Pramanik [8] discuss the significance of image segmentation in plant disease recognition and introduce a modified region growing algorithm for this purpose, improving upon existing grayscale-based methods.

Singh et al. [9] underscore the efficiency of automated methods in detecting plant diseases, presenting a genetic algorithm-based approach for image segmentation in disease identification and classification.

Ganesan et al. [10] highlight the pivotal role of plant disease identification and control in agricultural economies, proposing a fuzzy-based approach for early detection and segmentation of disease-affected plant parts.

Yang et al. [11] explore the application of Near-field Acoustical Holography in analyzing plant physiology under disease stress, developing mathematical models for disease prediction and optimal pesticide spraying to enhance crop protection.

Khitthuk et al. [12] introduce a framework for plant leaf disease analysis using neural networks, focusing on feature extraction and disease classification from color imagery. Their system demonstrates high accuracy in classifying various grape leaf diseases and has potential for broader application in plant disease diagnosis.

III. RESEARCH METHODOLGY

The research methodology described in the proposed algorithm outlines the steps followed to analyze and identify plant leaf diseases from images. Here's a breakdown:

1. **Selection of Leaf Image:** The process begins by selecting a leaf image for examination, typically obtained from field observations or digital imaging.
2. **Setting the Image as Query Image:** The selected leaf image is designated as the query image, which will undergo analysis to identify any diseases present.
3. **Contrast Enhancement and Image Resizing:** To improve clarity and facilitate analysis, the contrast of the image is enhanced, and it is resized as needed.
4. **Color Image Segmentation:** The image is segmented based on color, separating different regions or components within the image.
5. **K Means Clustering:** K Means clustering, a machine learning algorithm, is applied to the segmented image to further partition it into distinct clusters based on color similarity.
6. *Conversion to Lab Color Space:** The RGB color space of the image is converted to the L*a*b* color space, which is more perceptually uniform and often used in color-based image processing tasks.
7. **Multi-class SVM for Training Set Creation:** A multi-class Support Vector Machine (SVM) algorithm is employed to create a training set for identifying various disease classes present in the images.
8. **Loading Training Data:** The training data set, typically stored in a file like Training_Data.mat, is loaded into the system. This dataset likely contains labeled examples of diseased and healthy leaf images.
9. **Disease Classification:** The SVM classifier is utilized to classify the diseases present in the leaf image. Common diseases such as Alternaria Alternata, Anthracnose, Bacterial Blight, Cercospora Leaf Spot, and Healthy Leaf are typically considered.
10. **Stop:** Once the classification process is completed, the algorithm terminates.

Overall, this methodology combines image processing techniques, machine learning algorithms, and color space conversions to automate the identification and classification of plant leaf diseases from digital images.

IV. Simulations and Result Analysis

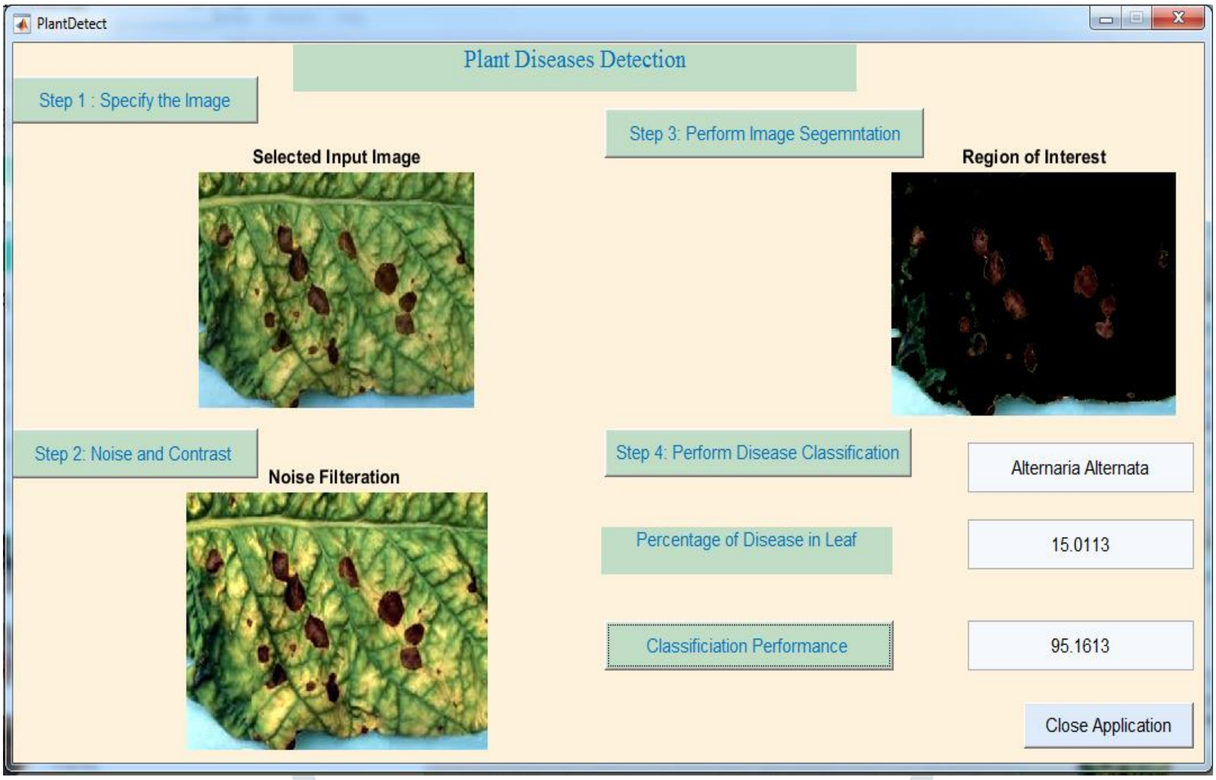

Fig 4.1 Implementation

Table 4.1 "Accuracy Table for *Alternaria Alternata*"

"Sample Number"	"Affected Region Identified %"	"Accuracy Achieved %"
1	15.019	95.1613
2	16.214	98.3871
3	15.314	96.7742
4	15.85	96.7742
5	14.1935	96.7742
	Average Accuracy	96.7742

The base paper which is taken is ,"**M. Bhagat, et.al 2020, Plant Leaf Disease Classification Using Grid Search Based SVM"** the accuracy which is achieved in plant disease identification is 84%.

Table 4.2 "Accuracy Table for AlternariaAlternata Comparison with Base"

"Classification Result"	"Base Percentage %"	"Proposed Accuracy Achieved %"
AlternariaAlternata	84%	96.7742

V. Conclusion

The conclusion highlights the significance of advancements in automated disease detection systems for agriculture, noting researchers' efforts in exploring various plant aspects. Using the K-Means algorithm, the system identifies infection concentrations on leaves and distinguishes between healthy and infected ones. Employing a Multi-Class SVM classifier, diseases like Alternaria

Alternata, Anthracnose, Bacterial Blight, and Cercospora Leaf Spot are accurately identified with an impressive 96 to 97 percent accuracy rate, surpassing previous methods.

In future work, the focus will be on further enhancing plant disease detection and improving classification accuracy. This will involve exploring additional techniques such as Artificial Neural Networks, Bayes classifiers, Fuzzy Logic, and hybrid algorithms to elevate recognition rates during classification.