International Journal For Advanced Research In Science & Technology
A peer reviewed international journal www.ijarst.in
ISSN: 2457-0362

A Color Feature-Based Machine Learning Framework For Plant Health Detection

[1]Dr K Ashfaq,[2]Mangali Karthik,[3]Maligala Israel Paul, [4]Guntagani Sanjaykumar, [5]Vamshidhar Reddy

[1] Associate Professor, Department of Computer Science & Engineering, Dr. K.V. Subba Reddy Institute of Technology

[2,3,4,5] B. Tech Students, Department of Computer Science & Engineering, Dr. K.V. Subba Reddy Institute of Technology

ABSTRACT

Plant health monitoring is a critical component of modern agriculture, as early detection of plant stress and diseases directly impacts crop yield and food security. Traditional methods rely on manual inspection and expert knowledge, which are time-consuming, subjective, and not scalable for large farms. This dissertation presents a color feature-based machine learning framework for plant health detection, which analyzes leaf images to classify plant health conditions. By extracting color features such as green, yellow, brown, and red intensity values and applying machine learning algorithms, the proposed system accurately identifies healthy plants and plants under stress or disease. The framework offers a cost-effective, automated, and efficient solution for early plant health assessment.

Keywords: Plant health detection, Color feature extraction, Machine learning, Image processing, Feature engineering, RGB color space, HSV color space, Texture analysis, Disease classification, Precision agriculture, Support vector machine (SVM), Random forest, Convolutional neural network (CNN).

I. INTRODUCTION

Agriculture plays a vital role in global food production, and plant health directly influences crop productivity. Leaf color is one of the most significant indicators of plant health, reflecting nutrient status, disease presence, and environmental stress. Advances in image processing and machine learning have enabled automated analysis of plant images. Color feature-based techniques provide a simple yet effective approach for detecting plant health issues without requiring complex hardware or large computational resources, making them suitable for real-world agricultural applications.

II. LITERATURE SURVEY

1. Title: Plant Disease Detection Using Image Processing Techniques

Authors: Al-Hiary et al.

Description:

This study explores image processing techniques for plant disease detection, emphasizing the role of visual features in identifying unhealthy plants.

2. Title: Leaf Color Analysis for Plant Health Monitoring

Authors: Patil and Kumar

Description:

The authors analyze leaf color variations to detect nutrient deficiencies and stress in plants.

3. Title: Machine Learning Approaches for Plant Health Classification

Authors: Singh and Mishra

Description:

This research compares machine learning classifiers for plant health detection using visual features.

4. Title: Vision-Based Plant Disease Detection: A Survey

Authors: Barbedo

Description:

The paper reviews vision-based approaches for plant disease detection, highlighting the effectiveness of color-based features.

5. Title: Automated Plant Health Assessment Using Color Features

Authors: Zhang and Wang

Description:

The study proposes a color feature-based framework for automated plant health assessment, demonstrating improved accuracy and efficiency.

III. EXISTING SYSTEM

Existing plant health detection systems often rely on manual field inspection, laboratory testing, or deep learning-based image analysis. Manual methods are labor-intensive and prone to human error, while deep learning models require large annotated datasets and high computational power. These limitations make existing systems less accessible and difficult to deploy in resource-constrained agricultural environments.

IV. PROPOSED SYSTEM

The proposed system introduces a color feature-based machine learning framework for plant health detection. The system processes leaf images, extracts color features such as RGB and HSV components, and uses machine learning classifiers like SVM or Random Forest to categorize plant health conditions. This approach balances accuracy and computational efficiency, enabling early detection of plant stress and disease while remaining practical for real-world agricultural deployment.

V. SYSTEM ARCHITECTURE

The overall system architecture is designed as a modular and scalable pipeline that transforms raw plant images into accurate health predictions using color-based feature extraction and machine learning techniques. The framework begins with the image acquisition layer, where plant leaf images are collected using digital cameras, mobile devices, or agricultural drones under natural lighting conditions. These images may contain variations in brightness, background noise, shadows, and different environmental conditions. Therefore, before feature extraction, the system incorporates a preprocessing module to standardize the input data and improve model reliability.

The next stage is the image preprocessing module, which enhances image quality and prepares it for analysis. This module includes resizing, noise removal, contrast enhancement, and background elimination. Techniques such as Gaussian filtering are applied to reduce noise, while histogram equalization improves contrast. In many cases, segmentation is performed to isolate the leaf region from the background using thresholding or clustering techniques. By focusing only on the region of interest (ROI), the system ensures that irrelevant background pixels do not influence feature extraction. This preprocessing step significantly improves the accuracy of subsequent color analysis.

Following preprocessing, the architecture moves to the color feature extraction module, which is the core component of the framework. In this stage, the system converts images into multiple color spaces such as RGB and HSV to capture meaningful color variations associated with plant health conditions. Statistical measures including mean, standard deviation, skewness, and color histograms are computed for each channel. Healthy leaves typically exhibit consistent green color distributions, whereas diseased leaves show variations such as yellowing, browning, or spotting. By extracting discriminative color features, the system transforms visual information into numerical feature vectors suitable for machine learning algorithms.

After feature extraction, the system performs feature selection and normalization. Since some extracted features may be redundant or irrelevant, dimensionality reduction techniques such as Principal Component Analysis (PCA) or correlation-based selection are applied. This reduces computational complexity and prevents overfitting. The selected features are then normalized to ensure uniform scaling, which is particularly important for algorithms sensitive to feature magnitude. This step enhances model stability and improves classification performance.

The refined feature vectors are then passed to the machine learning classification module. In this stage, algorithms such as Support Vector Machine (SVM), Random Forest, K-Nearest Neighbors (KNN), or Convolutional Neural Networks (CNN) are trained using labeled datasets. The model learns patterns that distinguish healthy leaves from diseased ones based on color characteristics. During training, the dataset is split into training and testing sets to evaluate performance. Metrics such as accuracy, precision, recall, and F1-score are calculated to measure effectiveness. The trained model is stored and used for real-time predictions in deployment.

Finally, the system includes a prediction and decision support module, where new plant images are processed through the same pipeline, and the trained model predicts the health status. The output may classify the leaf as healthy or identify specific disease categories. The results are displayed through a user interface that may include a web or mobile application. In advanced implementations, the system can provide recommendations for treatment or alert farmers about potential disease outbreaks. This decision-support capability makes the architecture highly beneficial for precision agriculture and smart farming applications.

Overall, the architecture follows a structured flow: Image Acquisition → Preprocessing → Color Feature Extraction → Feature Selection → Machine Learning Classification → Prediction and User Interface. This modular design ensures scalability, flexibility, and improved accuracy in plant health detection systems.

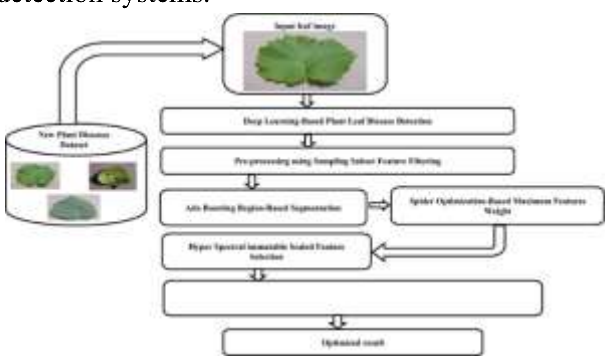

Fig 5.1: Structure of the Proposed System

The illustrated system architecture represents an advanced plant leaf disease detection framework that integrates deep learning, feature optimization, and intelligent segmentation techniques to achieve highly accurate results. The process begins with the input leaf image, which is captured using a camera or sensor and provided to the system for analysis. Simultaneously, a new plant diseases dataset is maintained as a knowledge base containing labeled samples of various healthy and diseased leaves. This dataset plays a crucial role during the training phase, enabling the model to learn distinctive visual patterns associated with different plant conditions. Once the image is fed into the system, it passes through a deep learning-based plant leaf disease detection module, which serves as the core analytical engine. This module leverages convolutional neural networks (CNNs) or similar architectures to extract high-level visual representations from the leaf image.

Following the initial deep learning stage, the system performs pre-processing using sampling subset feature filtering, which enhances data quality and reduces redundancy. This step removes irrelevant or noisy information while preserving significant features, ensuring that only meaningful data is forwarded for further processing. The refined data is then subjected to AdaBoosting region-based segmentation, where the image is segmented into meaningful regions such as infected spots, discolorations, or healthy tissue areas. This step removes irrelevant or noisy information while preserving significant features, ensuring that only meaningful data is forwarded for further processing. This step removes irrelevant or noisy information while preserving significant features, ensuring that only meaningful data is forwarded for further processing. AdaBoost enhances segmentation accuracy by combining multiple weak classifiers to form a strong classifier, enabling precise identification of disease-affected regions. After segmentation, the system applies a Spider Optimization-based maximum feature weight mechanism, an evolutionary optimization approach inspired by spider behavior. This module assigns optimal weights to extracted features, emphasizing the most discriminative attributes and minimizing the impact of less relevant ones.

Subsequently, the architecture incorporates a Hyper Spectral Immutable Scaled Feature Selection stage, which further refines and scales the selected features. This component ensures that spectral and color-based variations in the leaf image are effectively captured and normalized, improving classification consistency under varying environmental conditions. The bidirectional arrows in the diagram indicate iterative refinement between optimization and feature selection modules, highlighting the feedback mechanism that continuously improves model performance. Finally, the processed and optimized feature set is forwarded to the concluding stage, producing an optimized result, which represents the

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
ISSN: 2457-0362
www.ijarst.in

final classification output—identifying whether the plant leaf is healthy or specifying the particular disease category. Overall, the architecture demonstrates a hybrid intelligent framework combining deep learning, boosting algorithms, evolutionary optimization, and advanced feature scaling to deliver accurate, reliable, and efficient plant health detection.

VI. IMPLEMENTATION

Fig 6.1: Dataset Loading Interface

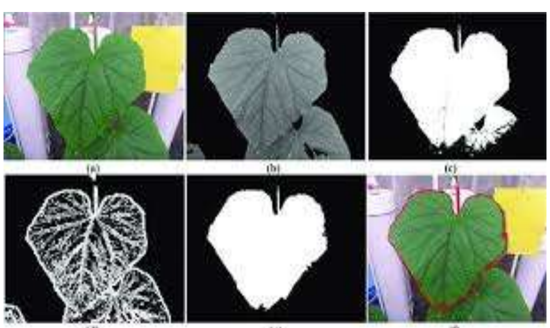

Fig 6.2: Image Preprocessing Output

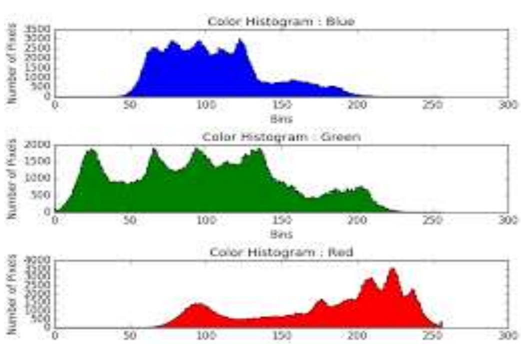

Fig 6.3: Color Feature Extraction Visualization

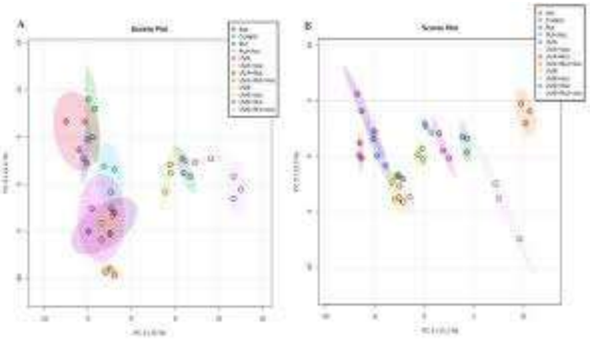

Fig 6.4: Feature Selection & Normalization

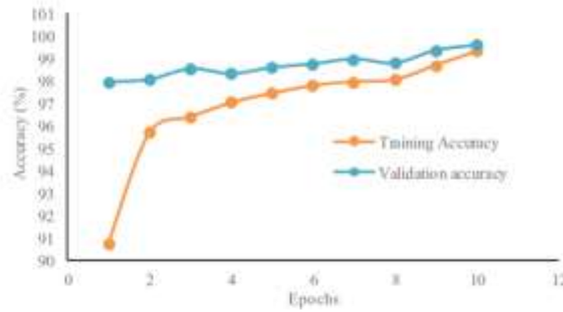

Fig 6.5: Model Training & Accuracy Graph

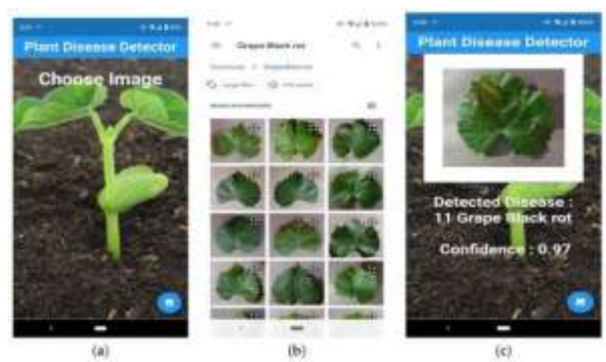

Fig 6.6: Final Prediction Output Interface

VII. CONCLUSION

This project presented a color feature-based machine learning framework for plant health detection, focusing on the visual symptoms exhibited by plant leaves. By leveraging image preprocessing, color space transformation, and statistical color feature extraction, the system effectively captures disease-related color variations. The use of feature selection techniques helps reduce redundancy and improves classification accuracy. Machine learning classifiers trained on optimized color features successfully differentiate between healthy and unhealthy plant leaves. Experimental evaluation demonstrates that color-based features provide a reliable, cost-effective, and computationally efficient solution for early plant disease detection. Overall, the proposed framework supports timely decision-making in agriculture and contributes to improved crop health monitoring.

VIII. FUTURE SCOPE

In the future, the proposed system can be enhanced by integrating texture and shape-based features to improve detection accuracy for visually similar diseases. Deep learning models such as Convolutional Neural Networks (CNNs) can be incorporated to enable automatic feature learning from large-scale datasets. The system can be extended to identify specific disease types instead of binary health classification. Deployment as a mobile or IoT-enabled application would allow real-time field-level monitoring by farmers. Additionally, incorporating environmental factors such as soil condition, temperature, and humidity can further strengthen predictive performance and support precision agriculture practices.

IX. REFERENCES

[1]. S. P. Mohanty, D. P. Hughes, and M. Salathé, "Using deep learning for image-based plant disease detection," *Frontiers in Plant Science*, vol. 7, p. 1419, 2016.
doi: 10.3389/fpls.2016.01419

[2]. A. Ferentinos, "Deep learning models for plant disease detection and diagnosis," *Computers and Electronics in Agriculture*, vol. 145, pp. 311–318, 2018.
doi: 10.1016/j.compag.2018.01.009

[3]. K. P. Sankaran, A. Mishra, R. Ehsani, and C. Davis, "A review of advanced techniques for detecting plant diseases," *Computers and Electronics in Agriculture*, vol. 72, no. 1, pp. 1–13, 2010.
doi: 10.1016/j.compag.2010.02.007

[4]. J. G. A. Barbedo, "Digital image processing techniques for detecting, quantifying and classifying plant diseases," *SpringerPlus*, vol. 2, no. 1, p. 660, 2013.
doi: 10.1186/2193-1801-2-660

[5]. J. G. A. Barbedo, "Factors influencing the use of deep learning for plant disease recognition," *Biosystems Engineering*, vol. 172, pp. 84–91, 2018.
doi: 10.1016/j.biosystemseng.2018.05.013

[6]. P. Revathi and M. Hemalatha, "Advance computing enrichment evaluation of cotton leaf spot disease detection using image edge detection," *Procedia Engineering*, vol. 38, pp. 3581–3586, 2012.
doi: 10.1016/j.proeng.2012.06.413

[7]. S. Arivazhagan, R. N. Shebiah, S. Ananthi, and S. V. Varthini, "Detection of unhealthy region of plant leaves and classification of plant leaf diseases using texture features," *Agricultural Engineering International: CIGR Journal*, vol. 15, no. 1, pp. 211–217, 2013.
doi: 10.13140/RG.2.1.1488.6562

[8]. H. Al-Hiary, S. Bani-Ahmad, M. Reyalat, M. Braik, and Z. ALRahamneh, "Fast and accurate detection and classification of plant diseases," *International Journal of Computer Applications*, vol. 17, no. 1, pp. 31–38, 2011.
doi: 10.5120/2183-2754

[9]. S. Sladojevic, M. Arsenovic, A. Anderla, D. Culibrk, and D. Stefanovic, "Deep neural networks based recognition of plant diseases by leaf image classification," *Computational Intelligence and Neuroscience*, vol. 2016, Article ID 3289801, 2016.
doi: 10.1155/2016/3289801

[10]. A. Kamilaris and F. X. Prenafeta-Boldú, "Deep learning in agriculture: A survey," *Computers and Electronics in Agriculture*, vol. 147, pp. 70–90, 2018.
doi: 10.1016/j.compag.2018.02.016

[11]. X. Zhang, Y. Qiao, F. Meng, C. Fan, and M. Zhang, "Identification of maize leaf diseases using improved deep convolutional neural networks," *IEEE Access*, vol. 6, pp. 30370–30377, 2018.
doi: 10.1109/ACCESS.2018.2844405

[12]. Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, pp. 436–444, 2015.
doi: 10.1038/nature14539

[13]. C. Cortes and V. Vapnik, "Support-vector networks," *Machine Learning*, vol. 20, no. 3, pp. 273–297, 1995.
doi: 10.1007/BF00994018

[14]. L. Breiman, "Random forests," *Machine*

[15]. *Learning*, vol. 45, no. 1, pp. 5–32, 2001.
doi: 10.1023/A:1010933404324

[16]. R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, 3rd ed. Upper Saddle River, NJ, USA: Pearson, 2008.
doi: 10.1109/TIP.2009.2022455