Advanced and Sustainable Technologies
ASET
E-ISSN 2976-2294
Volume 3, No 1, June 2024 [22-28]

Plant Disease Classification Using Image Processing Technique

Shahrul Fazly Man[1], Shafie Omar[1], Muhammad Imran Ahmad[1,2*], Wan Mohd Faizal Wan Nik[1], Tan Shie Chow[1,2], Mohd Nazri Abu Bakar[1], Fadhilnor Abdullah[1], Asbhir Yuusuf Omar[2]

[1]Institute of Sustainable Agrotechnology, Universiti Malaysia Perlis, Sg Chuchuh, 02100 Padang Besar, Perlis, Malaysia.
[2]Faculty of Electronic Engineering Technology, Universiti Malaysia Perlis, Kampus Pauh Putra, 02600 Arau, Perlis, Malaysia.

Received 20 August 2023, Revised 21 September 2023, Accepted 28 September 2023

ABSTRACT

Agriculture remains pivotal to our economy, with farming playing a central role in revenue generation. Challenges such as pests, plant diseases, and evolving climate patterns pose threats to crop yield and production. Addressing these challenges, timely and accurate detection of plant diseases emerges as imperative. Manual detection, however, remains resource-intensive and often lags. Addressing this gap, this project proposes an innovative image processing-based system for rapidly detecting plant diseases. The system proficiently identifies specific diseases by analyzing images of plant leaves against a curated dataset. The emphasis of this study was on three major diseases: Bacterial Blight (with an accuracy of 98.6%), Alternaria Alternata (98.5714%), and Cercospora Leaf Spot (97.5%). The compelling results underline the system's capacity to swiftly and effectively categorize diseases, offering monoculture farmers an indispensable tool for obtaining prompt, disease-specific insights.

Keywords: Agriculture, Plant Disease, SVM classifier.

1. INTRODUCTION

Agriculture, instrumental in the evolution of human civilization, faces challenges from plant diseases leading to devastating economic losses, as exemplified by the $653.06 million loss experienced by Georgia (USA) in 2007. The potential severity in regions like India is also speculated to be profound, yet remains undocumented [1]. Traditional detection methods, based on direct observation, often fall short in early disease detection due to their time-consuming nature and reliance on expert knowledge. These challenges are exacerbated in monoculture farming, where the effect of a single disease can be catastrophic. Moreover, erroneous disease detection often leads to excessive and indiscriminate pesticide use, further impacting the environment and sustainability of agricultural practices.

In light of these challenges, this study introduces a digital solution that harnesses image processing and computer vision techniques. Utilizing MATLAB, we propose an automated system to diagnose plant diseases from 2D leaf images, offering farmers a timely and accurate alternative to manual methods. This system classifies the diseases that were faced by the plantation. Furthermore, our approach employs a Support Vector Machine (SVM) model trained with diverse leaf images to ensure robustness in detection. By aiming for precise disease classification, this research strives to foster more sustainable agricultural practices, reducing unnecessary chemical use and its environmental repercussions.

*Corresponding author: m.imran@unimap.edu.my

2. LITERATURE REVIEW

Arivazhagen et al. developed a plant disease detection system targeting diseases on brinjal leaves [2]. Their approach synergizes the K-Means Clustering method with the Artificial Neural Network (ANN). Leaf images undergo histogram equalization and are then segmented using K-means clustering. Features from these clusters are extracted using gray-level co-occurrence matrix parameters, with the ANN determining the presence of disease. Gavhale introduced a system for detecting unhealthy regions on citrus leaves [3]. This approach, too, employed K-means clustering for image segmentation and utilized the Support Vector Machine (SVM) as the classifier, achieving an impressive 96% Genuine Acceptance Rate.

Selvarag introduced a mobile application to identify banana diseases and pests in banana plantations [4]. This system harnesses the power of Deep Convolutional Neural Networks for detection. It was trained on an extensive dataset of 18,000 banana field images across three distinct convolutional networks, generating six models. These models demonstrated 90% accuracy in detecting both banana diseases and pests. Another notable mobile-based plant disease detection method utilized a dataset of 7,386 images.[5] Here, features related to leaf color and shape were extracted. A Linear Support Vector classifier was subsequently trained on these features. This method excelled in gauging the severity of plant infections, with the training data representing five distinct stages of disease severity.

Lastly, in light of these reviews, our proposed method seeks to amalgamate the strengths of previous research. We present a user-friendly GUI that enables users to effortlessly upload and process images for disease detection. Importantly, our approach is efficient and doesn't necessitate vast amounts of data, unlike some deep learning methods, making it more accessible and cost-effective.

3. METHODOLOGY

Our proposed methods can be separated into two stages: the training and testing stages. Both stages have similar processes, starting with image acquisition, pre-processing, segmentation, and feature extraction. The difference is that in the testing stage, the feature extracted and stored during the training stage was used to compare with the new input image. The classification result will be computed by comparing the newly extracted and trained features. Figure 1 shows the workflow of the system.

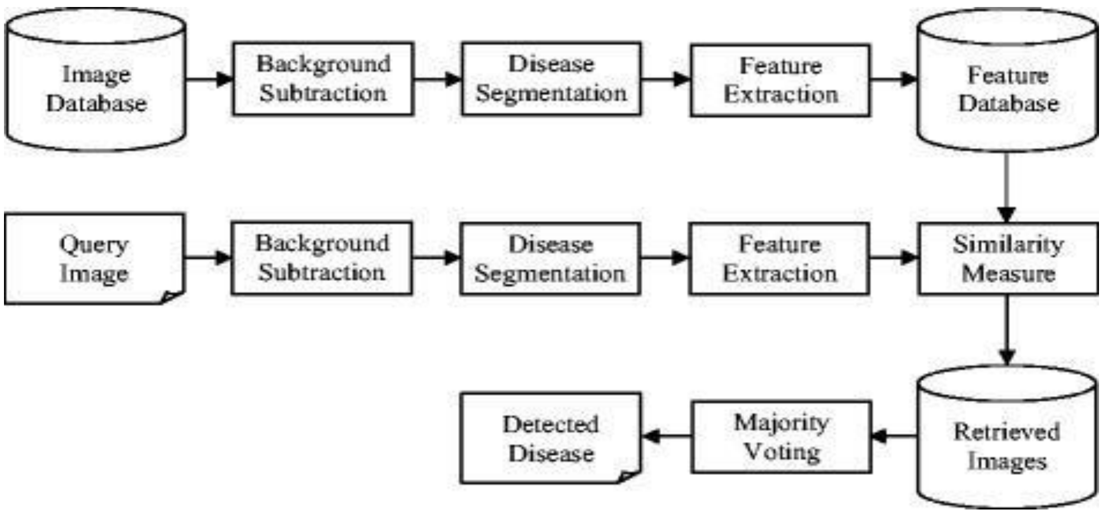

Figure 1: Overall process of the system.

3.1 Image Acquisition

Image acquisition serves as the foundational phase in our image-processing pipeline. In this crucial step, digital cameras or smartphones capture high-resolution photographs of the plant's leaves, ensuring that the leaf's texture, color, and any potential anomalies are clearly discernible. Once captured, these images are transferred and stored securely on a computer. This repository of images then becomes the primary input for the subsequent stages of the system, ensuring accurate disease detection and classification. Figure 2 shows the sample of the leaf database.

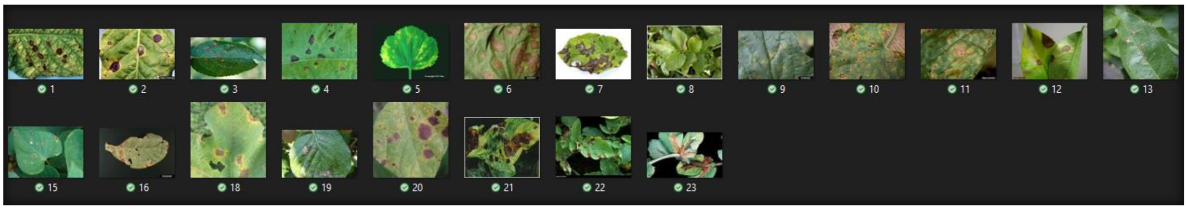

Figure 2: Alternaria Alternate leaf database

3.2 Image Pre-processing

In pre-processing, images are meticulously refined to ensure optimal disease detection. This entails removing extraneous noise such as spores and dust and subsequently transforming the color profile from the conventional RGB to the Lab* color space, which better aligns with human vision by delineating colors based on luminance (L*) and chromaticity (a* and b*). Such modifications accentuate disease-specific patterns and standardize the images, facilitating more accurate analyses in the succeeding stages. Figure 3 shows an example of a leaf that has already been pre-processed.

Figure 3: Enhanced Leaf Image.

3.3 Image Segmentation

Image segmentation is vital for subdividing the input image into coherent segments or sub-images, emphasizing distinct regions. Various methods such as region-based segmentation, edge detection, and -means clustering [6] can achieve this segmentation. The K-means method is particularly prominent due to its efficacy. At its core, K-means aims to partition n observations into k clusters, where each observation belongs to the cluster with the nearest mean. Mathematically, this can be thought of as minimizing the within-cluster sum of squares:

$$\sum_{i=1}^{k} \sum_{x \in S_i} \left\| x - \mu_i \right\|^2 \quad (1)$$

Where:

- S_i is the i^{th} cluser
- μ_i is the means of S_i

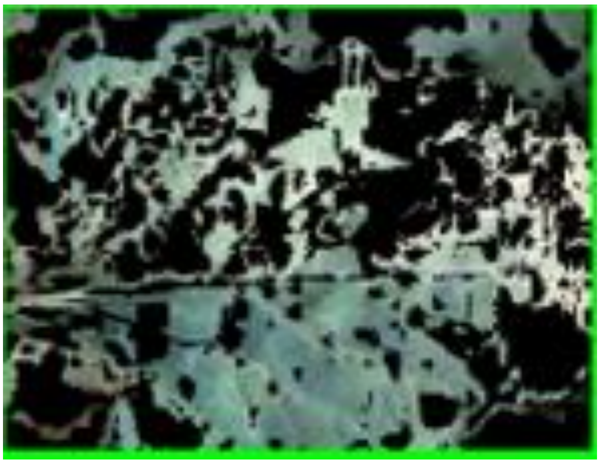

Figure 4: Segmented Leaf Image.

3.4 Feature Extraction

At this point, we must glean the relevant data from the image. The size of the selected area has been reduced compared to the whole image. Many techniques are available for feature extraction, but the Grey Level Co-occurrence Matrix is the most frequent. The GLCM is well-praised for its prowess in extracting texture features that may be used in subsequent analysis. The approach uses the image's intensity data to determine which pixel has the given value. Since the selected image is currently in RGB format, it will be converted to grayscale. Below is a table detailing the thirteen features and their respective expressions; these features are stored in an array.

GLCM is a statistical method renowned for capturing texture information [7], which is particularly crucial in discerning subtle changes in plant leaf health. The matrix quantifies how often specific combinations of pixel brightness values (gray levels) occur in an image. This co-occurrence captures the spatial relationship of pixels, revealing patterns and textures often indicative of disease.

To employ GLCM, it's imperative first to convert the segmented image from RGB to grayscale, as this simplifies the matrix by focusing solely on intensity variations rather than color information. The resulting GLCM then provides a variety of texture measures. While a comprehensive list of thirteen features derived from GLCM is available, some commonly used ones include contrast, correlation, energy, and homogeneity. Each feature offers insights into the textural patterns within the leaf, and they are stored in an array for subsequent processing and analysis. Table 1 below shows the features derived from GLCM.

Table 1: Measurements and conversions for refractive index of water.

Features	Expression
Mean =M	$\sum_{i=0}^{N-1} g(i)P(g(i))$
Standard Deviation -S	$\sqrt{\sum_{i=0}^{N-1} (g(i)-M)^2 P(g(i))}$
Entropy	$\sum_{i=0}^{N-1} P(g(i))\log_2(P(g(i)))$
RMS	$\sqrt{\frac{1}{NxN}\sum_{i=0}^{N-1}\sum_{j=0}^{N-1} (g(i,j)-I)^2}$
Variance	$\sum_{i=0}^{N-1} (i-\mu)^2 p(i)$
Smoothness	$\sum_i \sum_j \frac{1}{1+(i-j)^2} g_{i,j}$
Kurtosis	$\frac{1}{S^k}\sum_{i=0}^{N-1}(g(i)-M)^3 P(g(i))$
Skewness	$\frac{1}{S^3}\sum_{i=0}^{N-1}(g(i)-M)^3 P(g(i))$
Inverse Difference	$\sum_{i=0}^{N-1}\sum_{j=0}^{N-1} \frac{p_{i,j}}{1+(i-j)^2}$
Contrast	$\sum_i \sum_j (i-j)^2 g_{i,j}$
Correlation	$\frac{\sum_i \sum_y (ij) g_{i,j} - \mu_x \mu_y}{\sigma_x \sigma_y}$
Energy	$\sum_i \sum_j g_{ij}$
Homogeneity	$\sum_i \sum_j \frac{1}{1+(i-j)^2} g_{i,j}$

3.5 Image Classification

Plant diseases are primarily induced by fungi, bacteria, viruses, and insect damage, typically manifesting as spots on the leaves. For the precise classification of these diseases, the Support Vector Machine (SVM) is employed.

SVM is a supervised machine learning algorithm that identifies the hyperplane that best divides a dataset into classes. In the context of plant disease identification, features derived from the Gray

Level Co-occurrence Matrix (GLCM) become vital. These features, encapsulating the texture characteristics of the leaf images, serve as the input vectors for the SVM.

During training, the SVM is supplied with images of various diseases and their corresponding GLCM-derived features. This helps the algorithm learn the correlation between texture patterns and their associated diseases. The SVM tries to maximize the margin while minimizing the classification error for these training images, thus identifying an optimal hyperplane for classification. Once trained, this classifier can then discern and categorize unseen leaf images based on the texture features extracted from them.

The mentioned dataset, illustrated in the subsequent figure, comprises these training images. It acts as a foundational reference for the SVM classifier, ensuring that the SVM has a robust set of vector characteristics for comparison during the classification stage.

4. RESULTS AND DISCUSSION

The developed plant disease classification system is accessible via a Graphic User Interface (GUI) meticulously crafted by our team. Users can easily upload a new leaf image into the system. Once uploaded, the system undertakes several processing steps. It begins by showcasing different stages of the image as it undergoes processing. By the end of the procedure, the detected disease, alongside its affected region, is distinctly highlighted within the GUI. An illustrative example of the interface in action can be seen in Figure 5.

Diving deeper into the system's methodology, a series of image processing techniques, such as contrast enhancement and segmentation using K-means clustering, are employed after the image acquisition. Feature extraction follows suit, utilizing the Gray Level Co-occurrence Matrix (GLCM) method to discern vital texture features. These extracted attributes then feed into the SVM classifier, which discerns the type of plant disease present.

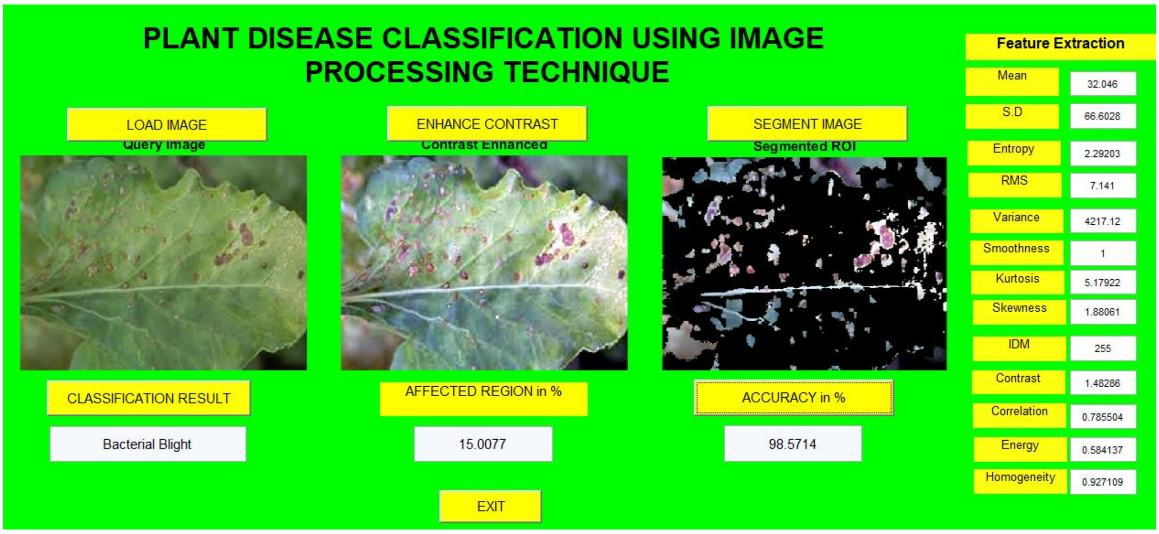

Figure 5: Plant disease classification system.

Using K-means clustering for image segmentation and SVM for classification, our system has been optimized to identify three prominent infectious diseases: Anthracnose, Cercospora Leaf Spot, and Bacterial Blight. The system performs a texture analysis on the segmented image derived from the Gray Level Co-occurrence Matrix (GLCM) features to discern the disease. As shown in Table 1 below, the percentages of infected and healthy leaves from various samples have been cataloged. Impressively, our system achieved an accuracy rate of 98.57% in detecting and classifying the infections correctly.

Table 1: Measurements and conversions for refractive index of water.

Sample NO:	Disease Classified	Affected Area (Percentage)	Accuracy (Percentage)
1	Bacterial Blight	15.0077	98.6
2	Alternata alternaria	15.0433	98.5714
3	Cercospora Leaf Spot	21.89	97.5

5. CONCLUSION

Leveraging image processing for plant disease detection has proven pivotal for consistent crop yield. Our system achieved an impressive accuracy of 98.57% for Alternata Alternaria disease, underscoring its efficacy in disease diagnosis. This technological advancement ensures precise identification and paves the way for farmers to optimize their crop production.

In the future, integrating MATLAB with smartphone applications and utilizing drones can further enhance the system. The aim is to make this technology accessible to farmers, ensuring timely interventions and driving agricultural practices towards unparalleled precision.

ACKNOWLEDGEMENTS

This project is supported by the UniMAP Research Grant.

REFERENCES

[1] Arivazhagan, S., Shebiah, R. N., Ananthi, S., & Varthini, S. V. Detection of unhealthy region of plant leaves and classification of plant leaf diseases using texture features. Agricultural Engineering International: CIGR Journal, vol 15, issue 1 (2013) pp. 211-217.

[2] Anand, R., Veni, S., & Aravinth, J. An application of image processing techniques for detection of diseases on brinjal leaves using k-means clustering method. In 2016 international conference on recent trends in information technology (ICRTIT) (2016) pp. 1-6.

[3] Gavhale, K. R., Gawande, U., & Hajari, K. O. Unhealthy region of citrus leaf detection using image processing techniques. In International Conference for Convergence for Technology-2014 (2014) pp. 1-6.

[4] Selvaraj, M. G., Vergara, A., Ruiz, H., Safari, N., Elayabalan, S., Ocimati, W., & Blomme, G. AI-powered banana diseases and pest detection. Plant methods, vol 15, (2019) pp. 1-11.

[5] Owomugisha, G., & Mwebaze, E. Machine learning for plant disease incidence and severity measurements from leaf images. In 2016 15th IEEE international conference on machine learning and applications (ICMLA) (2016) pp. 158-163.

[6] He, C., Mo, H., Pan, L., & Zhao, Y. (Eds.). Bio-inspired Computing: Theories and Applications: 12th International Conference, BIC-TA 2017, Harbin, China, December 1–3, 2017, Proceedings, vol 791, (2017).

[7] Umaselvi, M., Suresh Kumar, S., and Athithya, M., Chennai and Vivekanandha College of Technology for women, VCTW, (2012).