Academy journal for Basic and Applied Sciences (AJBAS) Volume 5 # 4 August 2023

Intelligent System for Potato Diseases Using Image Processing and Support Vector Machine

Maged. S. Almuzoghi[1], Mohammed. G. Alnaas[2]., Najeh. I. Allafi[3]

[1]Higher institute of Science and Technology -Azizia, Libya
mageed.d.learing@gmail.com
[2]Libyan Academy of graduate Studies- Tripoli, Libya
info.cs@academy.edu.ly
[3]Higher institute of Science and Technology -Azizia, Libya
annajeh73@gmail.com

Abstract

Modern plant disease detection technologies are a promising step toward food security and long-term agriculture. Imaging and computer vision, in particular, aid in the study of quantitative plant physiology. Manual interpretation, on the other hand, necessitates a great deal of effort, knowledge of plant diseases, and a long processing time. This paper proposes a method for identifying illnesses using leaf photos that combines image processing and artificial intelligence. From a plant picture collection, this automated technique classifies illnesses on potato plants. Use of segmentation method and a support vector machine result in a 95 percent accuracy in disease classification. As a result, the proposed method paves the way for large-scale automated plant disease diagnostics

.

Keywords: *- Machine learning, Plant disease, Brown, Rot infection, Pink rot.*

I. Introduction

One of the most important food crops is the potato, diseases that cause significant yield losses in potatoes. Early discovery of these diseases allows for preventative actions to be taken, as well as the reduction of economic and production losses. The most widely used method for detecting and identifying plant diseases in recent decades has been professional naked eye monitoring [1].

However, in many circumstances, this strategy is impractical due to long processing times and a scarcity of experts on farms in remote places. As a result, the use of image analysis technologies has proven to be an excellent way for maintaining continuous plant health monitoring and early diagnosis of plant diseases.

Image analysis of visual patterns on leaves can be used to diagnose diseases, as illnesses leave visible signals on plants, particularly leaves. As a consequence, merging imaging technology and machine learning to solve the problem of agricultural productivity while simultaneously assuring food security is a win-win situation. The purpose of this research is to use imaging and machine learning to produce an effective and error-free disease diagnostic system for plants.

II. Literature Review

RGB imaging, X-ray, ultrasound, and multispectral and hyper spectral technologies have all been developed to monitor and diagnose crop illnesses [2]. Macedo-Cruz et al. presented a method in [3] to assess the damage caused by frost in oat crops, the RGB color space is converted to the L*a*b* color space. It used Otsu's approach, Isodata algorithm, and fuzzy thresholding as their thresholding methodologies.

Yao et al. presented a technique in [4] that attempted to discover and classify some categories of rice-related disorders. Otsu's approach is used to segment the picture, following which the sick parts are separated. Color, form, and texture are retrieved, with the latter coming from the HSV color space

The approach presented by Phadikar and Sil in [5] identifies and distinguishes two illnesses that impact rice harvests by transforming the picture to the HSI color space and segmenting the image using an entropy-based thresholding. After applying an edge detector to the segmented picture, the intensity of the green components used to detect spots.

Wang et al. are a group of researchers that have worked on a variety suggested a strategy for distinguishing between disease pairings in wheat and grapevines. A K-means algorithm is used to segment the photos, and then 50 color, shape and texture attributes which are retrieved [6].

The goal of [7] was to identify and distinguish between four different forms of mineral deficiency (nitrogen, phosphorus, potassium and magnesium). The photos are transformed to the HSI and L*a*b* color spaces before color analysis. Euclidean distances measured in both color spaces are used to quantify the disparities.

Kurniawati et al. [8] provided a system for identifying and labelling three different types of paddy crop diseases. Thresholding is used to segment healthy and unhealthy zones, as is the case with many other approaches. Sharada P. M et al. [9] recently employed a deep convolutional neural network to recognize 14 crop species and 26 illnesses (or lack thereof).

III. Proposed System

This study utilizes the salient features of RGB images and machine learning for disease detection in potato. The dataset was separated into two independent data sets, with 80% and 20% used for training and testing respectively as illustrated in Figure 1.

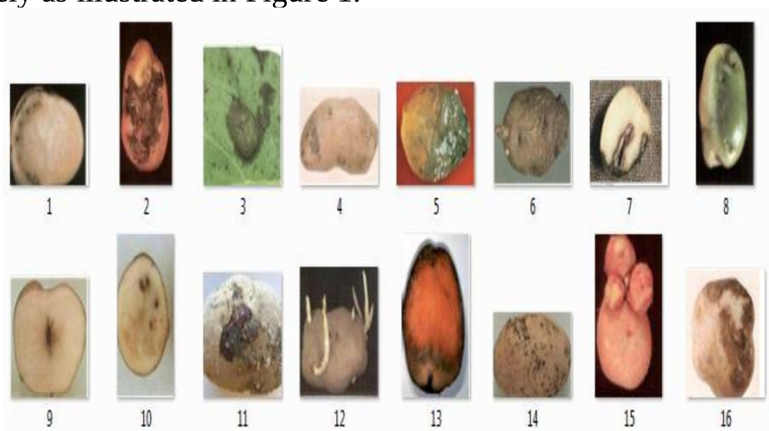


Fig.1: Some of Testing Image

IV. Feature Extraction

The basic information gathered from photos to identify them is known as color, texture and all features that have an influence on the picture of a potato leaf, three feature descriptors are used:

Academy journal for Basic and Applied Sciences (AJBAS) Volume 5 # 4 August 2023

A -Hu Moments

Hu Moments measure the form of an item in a picture. It usually refers to the object's contours. When a color image is transformed to a greyscale image, the image's moments are computed. Then it returns shape feature vectors.

A. Haralick Texture

The Haralick Texture feature descriptor extracts texture features, because the Haralick feature descriptor demands pictures in greyscale format, conversion of a color image to a greyscale image is required to extract features from it. The Grey Level Co-occurrence Matrix is the foundation for computing the Haralick texture feature (GLCM).

B. Color Histogram

It shows the amount of pixels in each color range and represents the color distribution in a picture. The color intensity of a picture is computed using the color histogram descriptor.

V. Classification using Vector Machine (SVM)

The kind of illness must be determined when the feature extraction procedure is done. The approach utilized for this challenge is the SVM classifier. SVM is a machine learning method that is already being used to discover public domains. The supervised learning model support vector machine is explicitly disproved by separating hyperplane. Figure 2 illustrated plant disease proposed system.

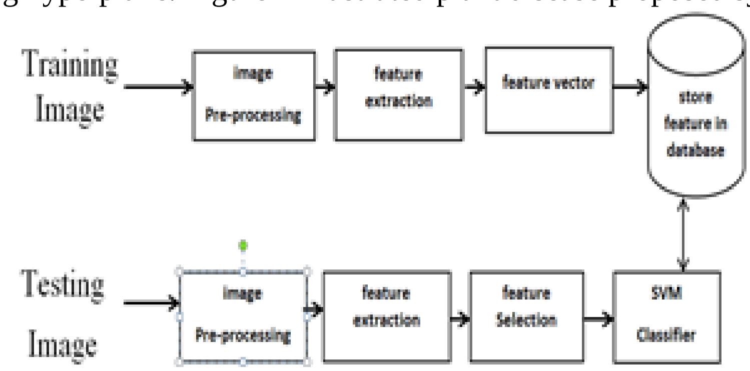


Fig. 2: Plant Disease Proposed System

SVM is a supervised learning approach that uses a pre-defined training data set to develop a decision-making mechanism to a better describe the data. The approach aims to identify decision boundaries that differentiate data points from different groups using a hyperplane, which is frequently used in pattern recognition and classification. Figure 3 illustrated a line separating between two categories using SVM.

The equation of the hyperplane is defined by [10]:

$$f_1 x1 + vc_1 \qquad (1)$$

Where:

f_1 is weight vector and b is bias.

Provided the training labeled data set {x i,y i} (i=1^N1) with x (i∈) R^d being the input vector and $\{x_i, y_i\}_{i=1}^{N}$ with $x_{i\in}\mathbb{R}^d$ being the input vector and $y_i \epsilon \{-1, +1\}$. As x_i is input vector and y_i is its corresponding label.

Academy journal for Basic and Applied Sciences (AJBAS) Volume 5 # 4 August 2023

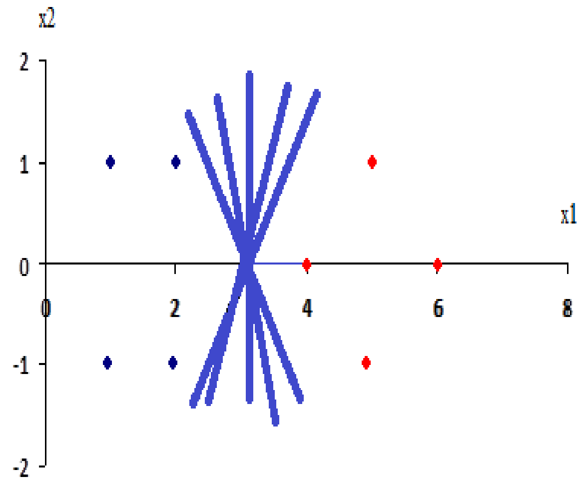


Fig. 3: Line Separating between two Categories

SVMs map the d-dimensional input vector x from the input space to the d_h -dimensional feature space by non-linear function $\varphi(\cdot)$: $\mathbb{R}^d \rightarrow \mathbb{R}_h^d$. Hence hyper plane defined as:

$$w1^T \varphi(x1) + vc1 = 0 \qquad (2)$$

With $vc1 \epsilon \mathbb{R}$ and w1 an unknown vector with the same dimension as $\varphi(x1)$. The resulting problem with SVM optimization is defined as:

$$\underset{w1,\xi,b}{min\ j1}(w1,\xi) = \frac{1}{2} w1^T w1 + c \sum_{i=1}^{n} \xi i \qquad (3)$$

SVM problem is usually described in dual space, by using Lagrange multipliers to impose restrictions on the minimizing function. The dual formula is:

$$\underset{\propto}{max1} \sum_{i1=1}^{m1} \propto_{i1} - \frac{1}{2} \sum_{i1,j1=1}^{N} \propto_{i1} \propto_{j1} y_{i1} y_j 1(x_{i1,} x1_{j1}) \quad (4)$$

Due to the $\propto_{i1} \geq 0$ for all $i1 = 1, \dots \dots m1$ and $1 \sum_{i1=1}^{m1} \propto_{i1} y_{i1} = 0$, the hyperplane can therefore be defined as a dual optimization problem

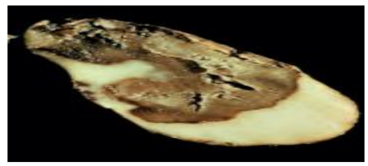
$$f_1(x_1) = sgn\left[\sum_{i1=1}^{m1} y_{1i} \propto_{i1} \left(x_{i1,} x_{1,}\right) + b_1\right] \qquad (5)$$

VI. Implementation

Mat lab used to carry out the execution of this study, photos are pre-processed by rescaling, converting to HSV color space and texture features are extracted. The features extracted for classification using SVM, eventually the classification results detect the name of disease.

On a database of 18 classes of photos for potato disease, the suggested method is tested. Multiclass SVM has been used to classify the data. Performance criteria like accuracy, sensitivity, recall and testing accuracy of the classification are used to evaluate the model's performance.
Table 1 shows some of the database photos, which present the user interface. The image that the user wants to anticipate is provided by the user.

Table 1: Database Images

Disease Name	**Disease Image**
Brown-Rot infection	
Watery Wound Rot(Leak)	
Early blight lesions on a potato leaf	
Pink Rot	
Normal Sprouts	

The algorithm processing the information and analyses it to determine the disease's name. If a leaf is expected to be sick, it shows the unhealthy leaf along with the name of the illness that affects it (be it bacterial, fungal or viral) as described in Figure 4.

Figure 5 illustrated recognition rate comparisons for the three image feature extraction methods and Table 2 Feature vector for some plant images.

The suggested system's accuracy is measured using the evaluation and comparison of the suggested approach in terms of time and percentage of effective outcomes according to the equation:

Academy journal for Basic and Applied Sciences (AJBAS) Volume 5 # 4 August 2023

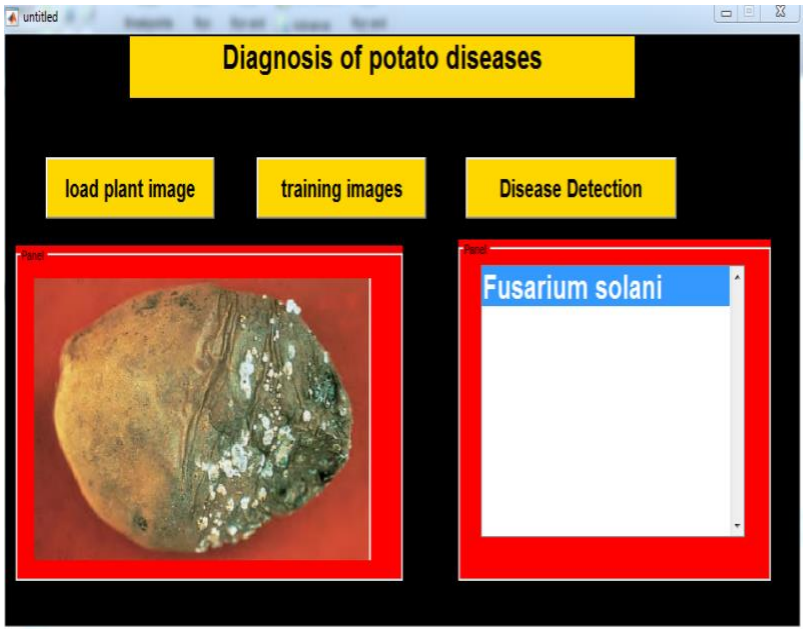


Figure 4: Graphic Proposed Matlab System

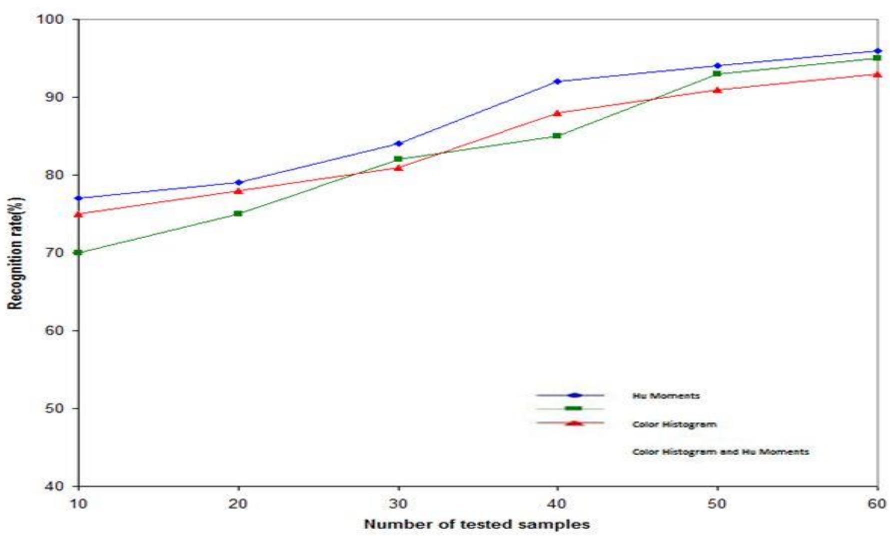


Figure 5: Recognition Rate Comparisons

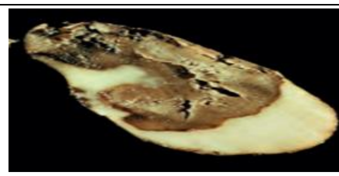
$$\text{Recognition rate} = \frac{\text{the number of success identification}}{\text{Total number of identification trials}}$$

- Success test refers to the number of correct results.
- failure test refers to the number of error results.

VII. Conclusion

Image segmentation with multiclass SVM used to create an automated and user-friendly system, with a little computational effort, the two most important potato diseases, late blight and early blight, may be recognized. This study provides farmers with a practical, effective and time-saving method of disease detection. It intends to add a larger number of illnesses from diverse plant species to the system. In the future, study will focus on calculating the severity of a sickness that has been recognized automatically

Table 2 Feature Vecto

Image	Feature Vector
	$\begin{bmatrix} -0.7658 \\ -0.0725 \\ -35.0485 \\ 15.8810 \\ 3.9851 \\ -35.0485 \\ 10.9506 \end{bmatrix}$
	$\begin{bmatrix} -1.0088 \\ -0.0590 \\ -43.0064 \\ 25.4109 \\ 5.0409 \\ -43.0064 \\ 6.9178 \end{bmatrix}$
	$\begin{bmatrix} -0.4008 \\ -0.0133 \\ -46.3286 \\ 24.8511 \\ 4.9851 \\ -46.3286 \\ 13.7102 \end{bmatrix}$
All images	$\begin{bmatrix} -0.6138 \\ -0.0465 \\ -41.5197 \\ 22.5970 \\ 4.7536 \\ -41.5197 \\ 12.8159 \end{bmatrix}$

References

[1] Vibhute A, Bodhe SK., Applications of image processing in agriculture: a survey, International Journal of Computer Applications. 52(2) 2012, pp. 34-40.

[2] Mirwaes Wahabzada, Anne-Katrin Mahlein, Christian Bauckhage, lrike Steiner, Erich Christian Oerke, Kristian Kersting, 'Plant Phenotyping using Probabilistic Topic Models: Uncovering the Hyperspectral Language of Plants, Scientific Reports 6, Nature, Article number: 22482, 2016.

[3] Macedo-Cruz A, Pajares G, Santos M, Villegas-Romero I., ―Digital image sensor-based assessment of the status of oat (Avena sativa L.) crops after frost damage, Sensors 11(6), 2011, pp. 6015–6036.

[4] Yao Q, Guan Z, Zhou Y, Tang J, Hu Y, Yang B, ―Application of support vector machine for detecting rice diseases using shape and color texture features, 2009 International Conference on Engineering Computation, IEEE, Hong Kong, 2009, pp. 79–83.

[5] Phadikar S, Sil J, ―'Rice disease identification using pattern recognition techniques, IEEE, Khulna, 2008, pp. 420_423.

[6] Wang H, Li G, Ma Z, Li X, ―Application of neural networks to image recognition of plant diseases, Proceedings of the 2012 International Conference on Systems and Informatics (ICSAI). IEEE, Yantai, 2012, pp. 2159–2164.

[7] Wiwart M, Fordonski G, Zuk-Golaszewska K, Suchowilska, Early diagnostics of macronutrient deficiencies in three legume species by color image analysis, ‖ Comput Electron Agric 65, 2009, pp. 125–132.

[8] Kurniawati NN, Abdullah SNHS, Abdullah S, Abdullah S, ―Investigation on image processing techniques for diagnosing paddy diseases, 2009 International Conference of Soft Computing and Pattern Recognition, IEEE, Malacca, 2009, pp. 272–277.

[9] Sharada Prasanna Mohanty, David Hughes, Marcel Salathe, Using Deep Learning for Image-Based Plant Disease Detection, Computer Vision and Pattern Recognition, to be published.

[10] S. Hossain, R. M. Mou, M. M. Hasan, S. Chakraborty and M. A. Razzak, ―Recognition and detection of tea leaf's diseases using support vector machine, 2018 IEEE 14th International Colloquium on Signal Processing & Its Applications (CSPA), Batu Feringghi, 2018, pp. 150–154.