International Journal on Future Revolution in Computer Science & Communication Engineering (IJFRCSCE)
ISSN: 2454-4248 Volume: 9 Issue: 1
DOI: https://doi.org/10.17762/ijfrcsce.v9i1.2123
***Article Received**: 18 November 2022 **Revised**: 24 December 2022 **Accepted**: 25 January 2023*

Numerical Modeling and Design of Machine Learning Based Paddy Leaf Disease Detection System for Agricultural Applications

Sitaram Meena[1], Abhigya Saxena[2]
[1]M. Tech Scholar, Department of Computer Science, YIT, Jaipur
[2]Assistant Professor, Department of Computer Science, YIT, Jaipur
Email: Sitarammeena79@gmail.com

Abstract

In order to satisfy the insatiable need for ever more bountiful harvests on the global market, the majority of countries deploy cutting-edge technologies to increase agricultural output. Only the most cutting-edge technologies can ensure an appropriate pace of food production. Abiotic stress factors that can affect plants at any stage of development include insects, diseases, drought, nutrient deficiencies, and weeds. On the amount and quality of agricultural production, this has a minimal effect. Identification of plant diseases is therefore essential but challenging and complicated. Paddy leaves must thus be closely watched in order to assess their health and look for disease symptoms. The productivity and production of the post-harvest period are significantly impacted by these illnesses. To gauge the severity of plant disease in the past, only visual examination (bare eye observation) methods have been employed. The skill of the analyst doing this analysis is essential to the caliber of the outcomes. Due to the large growing area and need for ongoing human monitoring, visual crop inspection takes a long time. Therefore, a system is required to replace human inspection. In order to identify the kind and severity of plant disease, image processing techniques are used in agriculture. This dissertation goes into great length regarding the many ailments that may be detected in rice fields using image processing. Identification and classification of the four rice plant diseases bacterial blight, sheath rot, blast, and brown spot are important to enhance yield. The other communicable diseases, such as stem rot, leaf scald, red stripe, and false smut, are not discussed in this paper. Despite the increased accuracy they offer, the categorization and optimization strategies utilized in this work lead it to take longer than typical to finish. It was evident that employing SVM techniques enabled superior performance results, but at a cost of substantial effort. K-means clustering is used in this paper segmentation process, which makes figuring out the cluster size, or K-value, more challenging. This clustering method operates best when used with images that are comparable in size and brightness. However, when the images have complicated sizes and intensity values, clustering is not particularly effective.

Keywords: SVM, K-Means, Paddy Leaf Disease, Machine Learning.

I. INTRODUCTION

The plant disease damages the plant physiological function and generates significant destruction to the field. Further, the plant diseases may transmit to other plants through different mechanisms. The existence of each disease in the plant is detected by its symptoms, which may present in many areas of the crops, such as roots, fruits, leaves, flowers and stem. Disease in plants can generate needless changes in appearance, size of fruits, leaves, flowers and stem.

Various Diseases in Paddy Leaf

Every year, diseases, pathogen invasions, and weather conditions have a major impact on agricultural production. While the spread of illness and infections can't be stopped, bad weather is something that can't be ignored. The following are some of the most common diseases that attack paddy leaf.

Deterioration of the sheath: Sheath rot reduces yields by preventing or delaying panicle development, which results in empty panicles and undeveloped seeds. As a result of this illness, the quality of the crop declines, and the panicles that form on the grains decay and change color. This illness spreads more quickly when the weather is moist than when it is dry. The spread of this disease is accelerated by a high planting density. Insects like stem borer produce wounds and damage in plants that allow this disease to enter. To prevent leaf rot and damage caused by the fungal disease sheath rot, fungicides should be used at the time of emergence, when the disease is at its most active. Climate factors such as warm temperatures (20–28 degrees Celsius), frequent applications of nitrogen fertilizer, and high relative humidity all contribute to the spread of this disease. Fig. 1.1 depicts a paddy leaf with sheath rot.

International Journal on Future Revolution in Computer Science & Communication Engineering (IJFRCSCE)
ISSN: 2454-4248 Volume: 9 Issue: 1
DOI: https://doi.org/10.17762/ijfrcsce.v9i1.2123
***Article Received**: 18 November 2022 **Revised**: 24 December 2022 **Accepted**: 25 January 2023*

Fig. 1.1 Sheath rot affected paddy leaf

Image Processing

The term "image processing" refers to a set of procedures that may be applied to a picture in order to either improve it or get actionable intelligence from it. The picture serves as input for this form of signal processing, and the output might be the same image or some of the image's attributes or characteristics. Image processing is one of the technologies that is expanding quickly today. It's a major focus of study in computer science and engineering.

Pre-processing of a picture might involve either the removal of unnecessary details or the improvement of existing ones. As a kind of signal processing, it is defined as having an image as both the output and the input. Various industries are now making use of image processing techniques, therefore this field is expanding quickly. Image processing techniques improve the quality of many technological and scientific fields. This includes fingerprint detection, optical character recognition, industries, augmented reality, computer vision, remote sensing, forecasting, biometric, data transmission, face detection, medical image processing, etc.

Application of image processing in Agriculture

Image processing can be used in agricultural application for following purposes:

- Pest management could be done in an efficient way using image processing tech- niques, that are employed to determine the insects, diseases and other organisms.
- Weeds are known as unwanted crops, in crop cultivation, because it utilize most of space, nutrients and water, therefore, image processing could be applied for the detection of Weeds.
- In order to classify the horticultural products such as potatoes, pears, nuts, grapes, banana, apples, tomatoes, pomegranates, mangoes, peaches and oranges, the im- age processing techniques are employed. Further, image processing is applied to perform quality inspection of food products and classification of grain.
- Image processing can be used during the harvest phase of both vegetable and fruits.
- The various diseases like Alternaria leaf spot, foliar leaf spot, and fungus on cotton leaves can be identified by image processing methods.
- Multi spectral image sensors are used in fields to find out the nitrogen content.
- Both texture and colour segmentation are used to perform Real time object track- ing.

Machine Learning Techniques Used In Agriculture

Different types of contemporary technology have emerged to improve agricultural productivity, sustainability, and efficiency in post-harvest processing. Many different methods, including thermography, gas chromatography, polymerase chain reaction, hyper spectral techniques, and mass spectrometry, are used in the detection of plant diseases. While these methods are inexpensive, they take longer to diagnose illnesses. The diagnosis of plant diseases has recently expanded to include mobile and server-based methods. These methods are paired with others, such as an extensive set of in-built accessories, a high-resolution camera, and powerful processing, to enable autonomous detection of plant diseases. Deep learning or machine learning algorithms are used to improve accuracy and recognition rate in such methods. Machine learning techniques such as convolutional neural networks, fuzzy logic, artificial neural networks, the k-means technique, the random forest classifier, and the support vector machine (SVM) are used to identify plant diseases. During training, a random forest is built up of several decision trees to perform tasks like regression, classification, and more.

a. Support Vector Machines (SVM)

When it comes to classification, SVM is one of the most important algorithms available. SVM may be used for classification of nonlinear and linear data. In the first step, SVM employs kernel functions to perform a nonlinear mapping of the data into a high-dimensional space. High-dimensional space is used to locate the ideal linear hyper plane, which partitions the data with a large margin.

SVM's benefits include its scalability to large datasets, improved accuracy, and resilience in the face of distorted training samples.

International Journal on Future Revolution in Computer Science & Communication Engineering (IJFRCSCE)
ISSN: 2454-4248 Volume: 9 Issue: 1
DOI: https://doi.org/10.17762/ijfrcsce.v9i1.2123
***Article Received**: 18 November 2022 **Revised**: 24 December 2022 **Accepted**: 25 January 2023*

Data mapping to generate higher dimensional data is a time-consuming and laborious process due to the importance of careful consideration of the kernel parameters and the kernel function.

b. Artificial Neural Network (ANN)

One way to think of artificial neural networks is as an amalgamation of several types of classification strategies. Based on the same principles as the human nervous system, ANN excels in interpolating massive volumes of uncertain data with great precision. When it comes to methods for diagnosing plant diseases, artificial neural networks are among the most prominent tools of the trade. In conjunction with other pre-processing methods, it helps extract relevant data from images.

To sum up, ANN has the benefit of being accurate even with complicated and noisy data, and it is also quite resilient. Although it may take more time to practice for high accuracy, it is possible to do so. It also has problems with scaling.

c. K - Nearest Neighbour Classifier (KNN)

KNN uses a distance metric to evaluate the closeness (similarity) of samples. Closest k neighbors data is classified based on its proximity to those neighbors. The K-Nearest Neighbor (KNN) method is used as a classifier in disease identification systems to sort leaves into groups based on their area, perimeter, and roundness.

The benefits of KNN include its easy and reliable implementation, its small number of tuning parameters (distance metric and k), and its low sensitivity to irrelevant and noisy input.

d. Random Forests

Commonly used for both regression and classification, random forest (RF) is a popular machine learning technique. It can effectively categorize a big portion of the dataset. It seems the tree is functioning properly based on the random values sampled from the forest. The data is sent into the system from the top of the tree down. The information is sampled at random, but only certain subsets are being used. The distribution of the sample is decided by a random number of trees in a random forest.

Motivation and Objectives

In the recent years, rice production has been mostly impacted by rice plant leaf diseases due to a lack of knowledge on effective management measures. Consistent damage to the paddy leaves from pathogens such sheath rot, leaf blast, brown spot, and bacterial blight is a major source of financial loss for farmers. The farmer uses naked-eye inspection to evaluate the paddy leaves, but this method is time-consuming and subject to human error. Natural vision observation is notoriously challenging and prone to human mistake. Getting around these issues calls for a reliable and quick recognition mechanism. Determining illnesses in paddy leaf necessitates hence the application of suitable methods. When it comes to providing farmers with an efficient, affordable, and trustworthy solution, the image processing methodology is often regarded as a non-invasive method. As a result, the purpose of this paper is to develop the optimized SVM algorithm in order to give the quick identification system to identify leaf illnesses in paddy crops.

Problem Definition

The detection of leaf diseases has always relied on human visual methods. However, it has a longer learning curve, lower output quality, and fewer usable features. The expert hired or the person's own eyesight determines the degree of precision and accuracy in a human vision method. Methods based on machine learning may correctly identify illnesses, choose effective therapies, and evaluate outcomes. When compared to human specialists, machine learning consistently outperforms the competition. In order to resolve the issues that have plagued traditional classification techniques, a new machine learning based approach has become necessary. When it comes to recognizing diseases in plant leaves, existing models have demonstrated relatively low recognition accuracy, notably in paddy leaf.

II. LITERATURE REVIEW

Sun et al. (2019) This research proposes a new approach that combines SLIC (Simple Linear Iterative Cluster) with SVM (Support Vector Machine) for better disease saliency map extraction from tea plant leaves against complicated backdrops. Before anything else, a super-pixel block is created using the SLIC technique, a significant point is identified using the Harris algorithm, and the contour of the fuzzy salient region is recovered using the convex hull approach. Second, the salient and background super-pixel blocks are classified using a support vector machine (SVM) classifier to extract their four-dimensional texture properties; finally, the classification map is generated. Finally, a system for fixing categorized super-pixel blocks using morphological and algebraic operations is built. This process yields a single, very precise saliency map of a diseased tea plant leaf. Testing on a dataset of 261 photos with illness showed an accuracy of 98.5%, precision of 98.6%, recall of

International Journal on Future Revolution in Computer Science & Communication Engineering (IJFRCSCE)
ISSN: 2454-4248 Volume: 9 Issue: 1
DOI: https://doi.org/10.17762/ijfrcsce.v9i1.2123
***Article Received**: 18 November 2022 **Revised**: 24 December 2022 **Accepted**: 25 January 2023*

97.7%, and F-value of 98.7% for the quality evaluation measure. These results show that the suggested technique outperforms the other three SLIC-based algorithms in terms of visual effects and quality assessment index. The suggested approach successfully extracts the disease saliency map from the complicated backdrop of tea plant leaves. This study should thus serve as a solid groundwork for future investigations into tea plant leaf disease diagnosis. The suggested technology, which extracts a saliency map of crop or plant disease, has promising future applications.

Zhang *et al.* (2019) This year, K-means clustering was presented as a method for diagnosing diseases in cucumbers. This strategy for identifying cucumber diseases involves three sequential steps: segmenting sick leaf pictures with K-means clustering; extracting shape and color characteristics from lesion information; and classification of diseased leaf images using Sparse Representation (SR). An important advantage of this method is its ability to categorize the SR space, which can reduce computational complexity and boost recognition performance. Additionally, a leaf image dataset related to cucumber illnesses is used to evaluate four different feature extraction based strategies. The suggested method is superior to the other methods in its ability to accurately identify seven key illnesses affecting cucumbers.

Maize leaf disease prediction was the focus of research conducted by Lin et al. (2018), who investigated the use of Homomorphic Filtering, Region of Interest (ROI) segmentation, and Multichannel Convolutional Neural Network (MCNN). MCNN has eight layers: input, five convolutions, three subsampling, three completely connected, and output. With this method of video saliency detection, the first completely linked layer is connected directly to the first and second subsampling layers, simulating the way humans' eyes work. Rectified Linear Units (ReLU), which combine features of pooling and normalizing, are also included. We also demonstrate the learning procedure inherent to the network architecture. In order to train the MCNN, researchers collected 10,820 RGB photos of five different diseases plaguing maize fields in China's Shandong Province. Real photos were useless for detection studies because of background noise and irrelevant details. In order to create a uniform database, they were first denoised, then segregated by homomorphic filtering and ROI segmentation. Experiments using 8 GB Graphics Processing Units (GPUs) demonstrated that the MCNN might aid in achieving improved accuracy in recognizing the maize leaf diseases. For the purpose of performance enhancement, multichannel design and the ensemble of numerous innovations have been shown to be effective methods.

The k-means clustering method was developed by Wahab et al. (2019) to identify illnesses in chili crops. In order to detect illnesses affecting the chilli plant, a new image processing system based on artificial intelligence is presented. The k-means clustering technique is emphasized as a key component of the proposed strategy for segmenting images. Following this, categorization is accomplished by a Support Vector Machine (SVM). Features extracted from the computed pictures are then utilized to assign categories to the resulting images. A wide range of SVM classification methods may be computed by adjusting the parameters and using different kernel functions. Results are broken down into three sections: baseline, healthy, and ill (Cucumber Mosaic), and can accurately discriminate between the two.

For paddy leaf disease diagnosis and prediction, Pinki et al. (2017) uses median filtering, K-means clustering, and Support Vector Machine (SVM). Brown spot, Leaf blast, and Bacterial blight are the three most common diseases affecting paddy leaves, and an automated system diagnoses them and provides recommendations for pesticides and/or fertilizers according on the severity of the illnesses. An application of K-means clustering is used to extract the contaminated region from the paddy leaf picture. Diseases are categorized based on visible qualities such color, texture, and form. SVM classifier is utilized to distinguish between different paddy leaf diseases. Upon completion of the identification procedure, a single predictive therapy is advised that can aid agricultural peasants and organizations in taking effective measures against these illnesses.

K-means and the Support Vector Machine (SVM) were proposed by Oo and Htun (2018) to identify and categorize diseases in plant leaves. Picture capture, image preprocessing, segmentation, feature extraction, and classification are the five main components of a plant leaf disease detection system. A collection of photographs of diseased leaves is made and kept for future use in scientific research. The photographs are improved with pre-processing. To generate sters, K-Means clustering methods are applied to the obtained leaf pictures for segmentation. Following K-means, we employ Grey Level Co-occurrence Matrices (GLCMs) and Local Binary Patterns (LBPs) as features. To classify and detect leaf diseases such bacterial blight, cercospora leaf spot, powdery mildew, and rust, researchers have turned to the Support Vector Machine (SVM).

International Journal on Future Revolution in Computer Science & Communication Engineering (IJFRCSCE)
ISSN: 2454-4248 Volume: 9 Issue: 1
DOI: https://doi.org/10.17762/ijfrcsce.v9i1.2123
***Article Received**: 18 November 2022 **Revised**: 24 December 2022 **Accepted**: 25 January 2023*

In 2018, Sardogan et al. proposed a method for disease identification and classification in tomato leaves using a Convolutional Neural Network (CNN) model and the Learning Vector Quantization (LVQ) algorithm. A total of 500 photos of diseased tomato leaves are included in the collection. You may use a convolutional neural network to automatically extract features and classify them. Leaf color data is often used in the study of plant diseases. The model employs the filters in a three-channel setup that is based on the RGB color space. After the convolution step, the feature vector output is given to the LVQ for training the network. According to experimental data, the suggested method is useful for accurately diagnosing four distinct types of tomato leaf diseases.

III. PROPOSED METHODOLOGY

In our research work, the role of Support Vector Machine is to classify the paddy pests, disease and weeds. Before the classification process established the images should be preprocessed and it will be helpful for the classification accuracy. Collection of data is from the dataset (ImageNet), pre-processing is used to removing the irrelevant and noise information from the image which helps to improve the overall efficiency of the paddy pest, disease and weed detection process. Then the general noise removal processing steps such as image acquisition, image preprocessing, image segmentation, feature extraction and classification. Once the preprocessing is done, the extracted features are given as an input to the Support Vector Machine (SVM) classifier for classification.

Classification using Support Vector Machine

The class H1 does not separate the two classes; H2 separates but with a very tinny margin between the classes and H3 separates the two classes with much better margin than H2. Thus the hyper plane creates the border between two different classes that is namely known as maximum-margin hyper plane or linear classifier or maximum margin classifier. Further, the Support Vector Machine method has been worked in both linear and non-linear way to analyze the features from the dimensional space.

(i) Linear Model

The Support Vector Machine method separates the data by determining the hyper plane [70]. Let as consider „W" is the hyper plane and „b" is the displacement of the origin data. Then the input feature is determined by using

D(x)=W.x-b --------------------------------3.1

where,

$$x \epsilon \begin{cases} A & if\ D(x) > 0 \\ B & if\ D(x) < 0 \end{cases}$$
-----------------3.2

Based on the input feature estimation, the hyper plane is determined,

(D(x))/‖W‖ ---------------------------------- 3.3

In the eqn (3.4) and (3.5), is the opposite signs which is belongs to the different sets when both inputs are belongs to the opposite sides of the hyper plane separation which is represented as follows,

w.x-b=1 ------------------------------------- 3.4

w.x-b=-1 ------------------------------------ 3.5

Based on the hyper plane in the linear model, the output of that is relevant group is estimated as,

$$\boldsymbol{y_i = \begin{cases} +1 & if\ x_i \in A \\ -1\ if\ x_i \in B \end{cases}}$$
---------------3.6

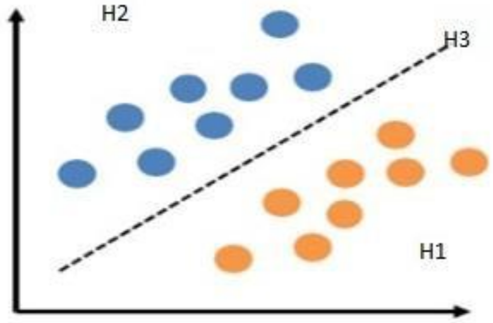

Figure 3.1: Separation hyper planes in SVM Linear

According to the output estimation process, the training data"s present in two groups is linearly separable with effective manner. The distance between the hyper plane. The linear separable based Support Vector Machine classification only fit for the defined dimensional space which is difficult to process while using the dynamic dimensional space for this issue.

Demerits of Linear

(i) Only linear associations between dependent and independent variables may be modeled using linear regression. It makes the

International Journal on Future Revolution in Computer Science & Communication Engineering (IJFRCSCE)
ISSN: 2454-4248 Volume: 9 Issue: 1
DOI: https://doi.org/10.17762/ijfrcsce.v9i1.2123
***Article Received**: 18 November 2022 **Revised**: 24 December 2022 **Accepted**: 25 January 2023*

sometimes-incorrect assumption that they are related in a straight line. To put it another way, linear regression is hyper vigilant for outliers in the data (or outliers).

(ii) Consider that the majority of your information has values between 0 and 10. For example, the regression coefficients would be drastically different if only one of the data points (out of a total of twenty) was outside the acceptable range of -15 to 15.

(iii) Another drawback is that the model begins modeling the noise rather than the link between the variables if there are more parameters than available samples.

The non-linear separable based SVM classification is used which is explained as follows.

(ii) Non-linear Model

There are some difficulties present in the linear separable Support Vector Machine model is resolved by using the non-linear based classification which was introduced by the Isabelle Guyon, Bernhard Boser and Vapin in 1992. They introduced a new method which works better in high dimensional feature space with the easiest way by modifying the quadratic ogramming to the feature space transform function which provides the better result. Then the support vector decision function is defined as follows,

$$D(x) = \sum_{i=1}^{p} \propto_i \, y_i K(x_i, x) - b \qquad \text{----------3.7}$$

So, in this work, the extracted paddy features are fed into the non-linear support vector machine which separates the healthy and affected image with the help of the hyper plane and radial basis kernel function. The kernel function is calculated as, and it‟s denoted as the Squared Euclidean distance between feature space.

$$K(x,x') = exp\left(-\frac{\|x-x'\|^2}{2\sigma^2}\right) \qquad \text{---------------3.8}$$

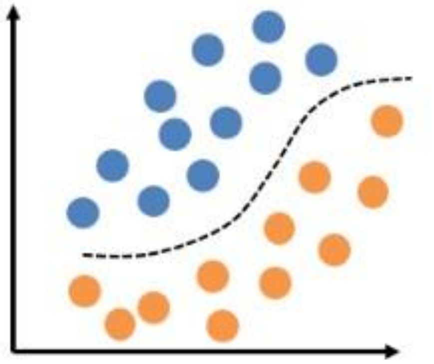

Figure 3.2: Separation hyper planes in SVM Non-Linear

The SVM separation hyper planes are shown in Figure 3.2. The non-linear support vector machine (SVM) is a relatively recent development in machine learning. Among the various applications of support vector machines (SVM), paddy categorization is a common one. Due to the way it is programmed, SVM can only distinguish between two groups of data. The hyper plane edge is maximized to achieve this. Support vectors are the samples that were chosen to be near to the edge in order to find the hyper plane. A number of one-versus-all and one-class support vector machines (SVMs) provide the basis of multiclass classification's appropriateness and fundamental construction.

Performance Analysis

Performance analysis examines the excellence of the preprocessing, segmentation, and feature extraction and classification process. The effectiveness is tested by running the algorithm in MATLAB tool and the efficiency is determined by using the following performance metrics.

Performance Metrics

The performance of the proposed system is analyzed with the help of the objective base quality measures such as (i)Mean Square Error (MSE), (ii)Peak Signal to Noise Ratio (PSNR), (iii)Sensitivity, (iv)Specificity and (v) Accuracy .

The proposed approach was evaluated using sensitivity and specificity method. Confusion matrix indicates the evaluation process of Support Vector Machine (SVM) classifier.

These images were classified as TP (True Positive), TN (True Negative), FP (False Positive), and FN (False Negative) based on results generated by the system.

True Positive: The image was infected and its pest, disease and weed was predicted truly.

True Negative: The image was not infected by a certain

International Journal on Future Revolution in Computer Science & Communication Engineering (IJFRCSCE)
ISSN: 2454-4248 Volume: 9 Issue: 1
DOI: https://doi.org/10.17762/ijfrcsce.v9i1.2123
***Article Received**: 18 November 2022 **Revised**: 24 December 2022 **Accepted**: 25 January 2023*

pest, disease and weed the system predicted it as negative.

False Positive: The image lacked a certain pest, disease and weed but the system predicted it positive.

False Negative: The image had a certain pest, disease and weed but the system predicted absence of it.

The summary of the results is shown in Table 3.1, which depicted 225 TN, 600 TP, 95 FN and 80 FP cases.

Accuracy: The proportion of times that the system makes correct predictions. Measures the system's ability to correctly forecast outcomes and expresses results as a percentage. Specifically, it measures how well the system can spot true positives.

Mean Square Error (MSE)

It measures the average of the square error [72] which means calculating the error amount by the pixel value of the original image differs from the image during the noise estimation process.

$$\text{Mean Square Error} = \frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N}\left(f(i,j)-f'(i,j)\right)^2 \quad \text{----------3.9}$$

The eqn (3.9) $f(i,j)$ represented as the original image and $f'(i,j)$ denoted as the noise estimated image. M is the height of the image and N is the width of the image.

Peak Signal to Noise Ratio (PSNR)

Peak Signal to Noise Ratio is the performance metric measures which are used to identify the quality of the original image which justifies the similarity between the original and noise image.

$$\text{Peak Signal Noise Ratio} = 20\log 10\left(\frac{255}{\sqrt{MSE}}\right) \quad \text{---------------3.10}$$

Mean Squared Error which computes the difference between the actual value and the estimated value and then the advantage of the PSNR metric is the easiest computation.

Accuracy

Accuracy is a statistical measure which is used to analyze how well the binary classifier recognizes the pest, disease and weed with optimized way. In addition, the accuracy is the proportion of the true results that include both true positives and true negatives among the total number of cases examined, then accuracy value is calculated 82.50 % by the eqn (3.11).

$$Accuracy = \frac{True\ Positive + True\ Negative}{True\ Positive + False\ Positive + False\ Negative + True\ Negative} \quad \text{-------------3.11}$$

The excellence of the proposed noise removal, segmentation and feature extraction and classification process is evaluated using the following implementation results.

Sample Calculation:

True Positive=600, True Negative=225, False Positive=80 and False Negative=95

Accuracy=TP+TN/ TP +FP+FN+TN

$$\frac{600+225}{600+80+95+225} = \frac{825}{1000} = 82.50$$

Sensitivity

Sensitivity is a measure is used to how the proposed system correctly classifies the pest, disease and weed in paddy with efficient manner. The sensitivity is measured the value is 86.33% by the eqn (3.24)

$$Sensitivity = \frac{True\ Positive}{(True\ positive + False\ Negative)} \quad \text{---------------3.12}$$

The efficiency of the system is examined with the help of the experimental results and discussions. Even though the Support Vector Machine method successfully recognizes the affected paddy pests, diseases and weeds feature, the testing phase consumes more time which reduces the efficiency of the system.

IV. RESULT ANALYSIS

Implementation of Proposed Methodology

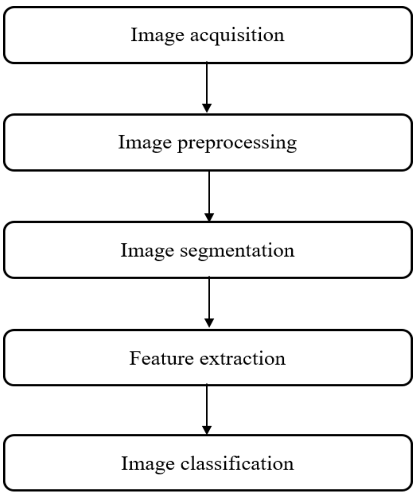

Figure 4.1 Flow Chart of Application of Disease Detection Framework

International Journal on Future Revolution in Computer Science & Communication Engineering (IJFRCSCE)
ISSN: 2454-4248 Volume: 9 Issue: 1
DOI: https://doi.org/10.17762/ijfrcsce.v9i1.2123
***Article Received**: 18 November 2022 **Revised**: 24 December 2022 **Accepted**: 25 January 2023*

The aforementioned approach has been built using the MATLAB software suite on an Intel Quad Core CPU. Figures 4.1 and 4.2 detail the proposed methodology's execution flow chart in more detail.

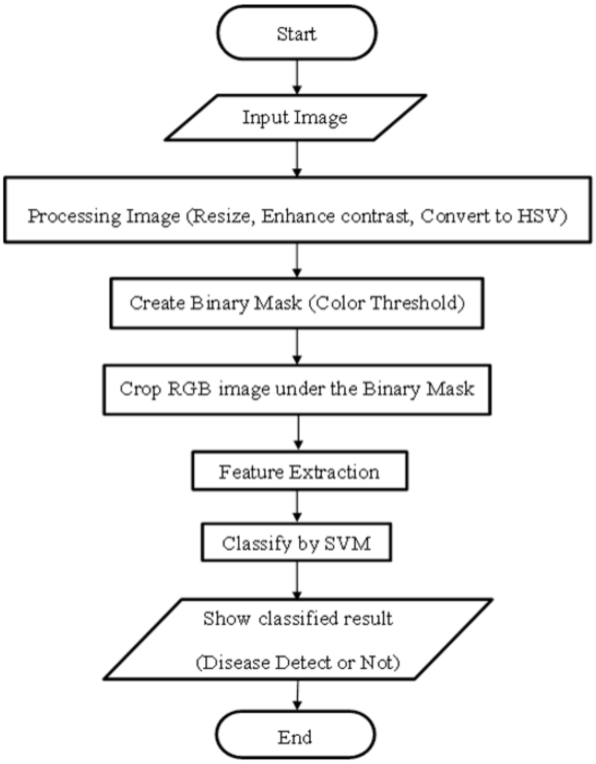

Figure 4.2 Detailed Flow Chart of Pre Processing and Classification Process

(a) Convert RGB to HSV (Healthy Leaf)

(b) Convert RGB to HSV (Unhealthy Leaf)

Figure 4.3 Detailed Flow Chart of Pre Processing of Unhealthy and Healthy Image

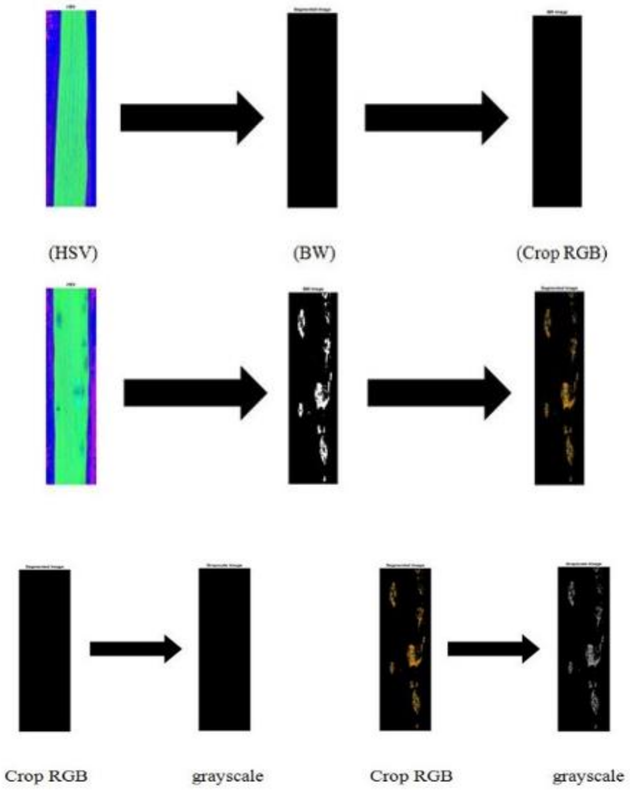

Figure 4.5 Step by Step Process of Segmentation of Healthy and Unhealthy Leaf

Table 4.1
Analysis of Accuracy for the Proposed System

	Healthy	Disease
Healthy	10	0
Disease	1	9

Images of three illnesses, including blast impacted disease and others like bacterial blight, sheath rot, and brown spot, were taken into account to increase the quantity of study done. In this study, we used enhanced SVM to categorize the various paddy leaf diseases separately. On top of that, the suggested classifier's accuracy is compared to that of the current classifier. The sections that follow classify the results obtained by various classifiers and the accompanying image disorders. Table 4.1 summarizes the results of many different classifiers with regards to accuracy. When it comes to identifying paddy crop diseases, the SVM classifier supports maximum accuracy output with a 90% accuracy rate. Classifiers of different sorts are evaluated for their role in a comprehensive disease detection system for agricultural crops. When it comes to diagnosing plant diseases, SVM classifier yields more reliable results. SVM with enhanced

International Journal on Future Revolution in Computer Science & Communication Engineering (IJFRCSCE)
ISSN: 2454-4248 Volume: 9 Issue: 1
DOI: https://doi.org/10.17762/ijfrcsce.v9i1.2123
Article Received: 18 November 2022 Revised: 24 December 2022 Accepted: 25 January 2023

classifier is tailored to detect the remaining illnesses, such as bacterial blight, sheath rot, and brown spot, with higher precision. In the end, evidence accumulated to show that paddy crops grown in greenhouses were more productive than those grown in a traditional farm setting.

V. CONCLUSION AND FUTURE SCOPE

Conclusion

This paper goes into great detail about how image processing may be used to spot various illnesses in rice fields. In order to maximize output, it is necessary to identify and classify four rice plant diseases: bacterial blight, sheath rot, blast, and brown spot. This paper does not address the other obtainable diseases, such as stem rot, leaf scald, red stripe, or false smut. This job takes longer than average to complete because of the categorization and optimization techniques used, despite the improved accuracy they provide. It was clear that using SVM methods facilitated better performance outcomes, albeit at the expense of considerable effort. Throughout this paper segmentation procedure, k-means clustering plays a role, increasing the difficulty of determining the cluster size, or K-value. When working with photographs of similar dimensions and brightness, this clustering technique performs most well. Clustering, however, is not very useful when the pictures are of complex sizes and intensity values.

Future Scope

The study focused on developing efficient machine learning algorithms to use with image processing tools for the diagnosis of four distinct rice plant diseases. Clustering is performed using k-means clustering during the segmentation process. Extraction of features is based on a combination of two aspects, often texture and color. Use of machine learning methods, and specifically the SVM optimization algorithm, aids in the classification process, allowing for the best possible results to be obtained.

REFERENCES

[1.] Sun, Y, Jiang, Z, Zhang, L, Dong, W & Rao, Y 2019, 'SLIC_SVM based leaf diseases saliency map extraction of tea plant', Computers and Electronics in Agriculture, vol. 157, pp. 102-109.

[2.] Tian, K, Li, J, Zeng, J, Evans, A & Zhang, L 2019, 'Segmentation of tomato leaf images based on adaptive clustering number of K-means algorithm', Computers and Electronics in Agriculture, vol. 165, pp. 1-7.

[3.] Hernández-Rabadán, DL, Ramos-Quintana, F & Guerrero Juk, J 2014, ' Integrating SOMs and a Bayesian classifier for segmenting diseased plants in uncontrolled environments', The Scientific World Journal, pp.1-13.

[4.] Shinde, RC, Mathew, CJ & Patil, CY 2015, Segmentation technique for soybean leaves disease detection', International Journal of Advanced Research (IJAR), vol. 3, no. 5, pp.522-528.

[5.] Gui, J, Hao, L, Zhang, Q & Bao, X 2015, 'A new method for soybean leaf disease detection based on modified salient regions', International Journal of Multimedia and Ubiquitous Engineering, vol. 10, no. 6, pp. 45-52.

[6.] Xu, L, Xu, X, Hu, M, Wang, R, Xie, C & Chen, H 2015, 'Corn leaf disease identification based on multiple classifiers fusion', Transactions of the Chinese Society of Agricultural Engineering, vol. 31, no. 14, pp. 194-201.

[7.] Lin, Z, Mu, S, Shi, A, Pang, C & Sun, X 2018, 'A novel method of maize leaf disease image identification based on a multichannel convolutional neural network', Transactions of the ASABE, vol. 61, no. 5, pp. 1461-1474.

[8.] Wahab, AHBA, Zahari, R & Lim, TH 2019, 'Detecting diseases in Chilli Plants Using K-Means Segmented Support Vector Machine', 3rd International Conference on Imaging, Signal Processing and Communication (ICISPC), pp. 57-61.

[9.] Pinki, FT, Khatun, N & Islam, SM 2017, 'Content based paddy leaf disease recognition and remedy prediction using support vector machine', 20th International Conference of Computer and Information Technology (ICCIT), pp. 1-5.

[10.] Oo, YM & Htun, NC 2018, 'Plant leaf disease detection and classification using image processing', International Journal of Research and Engineering, vol. 5, no. 9, pp. 516-523.

[11.] Karmokar, BC, Ullah, MS, Siddiquee, MK & Alam, KMR 2015, 'Tea leaf diseases recognition using neural network ensemble', International Journal of Computer Applications, vol. 114, no. 17, pp. 27-30.

[12.] Sardogan, M, Tuncer, A & Ozen, Y 2018, 'Plant leaf disease detection and classification based on CNN with LVQ algorithm', 3rd International Conference on Computer Science and Engineering (UBMK), pp. 382-385.

[13.] Amara, J, Bouaziz, B & Algergawy, A 2017, 'A deep learning-based approach for banana leaf diseases classification', Datenbanksysteme für Business, Technologie und Web (BTW 2017)-Workshopband, pp. 1-10.

[14.] Liu, B, Zhang, Y, He, D & Li, Y 2018, 'Identification of apple leaf diseases based on deep convolutional neural networks', Symmetry, vol. 10, no. 1, pp. 1-16.

[15.] Mia, MR, Roy, S, Das, SK & Rahman, MA 2020, 'Mango leaf disease recognition using neural network and support vector machine', Iran Journal of Computer Science, pp. 1-9.

[16.] Trang, K, TonThat, L & Thao, NGM 2020, 'Plant Leaf Disease Identification by Deep Convolutional Autoencoder as a Feature Extraction Approach', 17th

International Journal on Future Revolution in Computer Science & Communication Engineering (IJFRCSCE)
ISSN: 2454-4248 Volume: 9 Issue: 1
DOI: https://doi.org/10.17762/ijfrcsce.v9i1.2123
***Article Received**: 18 November 2022 **Revised**: 24 December 2022 **Accepted**: 25 January 2023*

International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON), pp. 522-526.

[17.] Chen, J, Liu, Q & Gao, L 2019, 'Visual tea leaf disease recognition using a convolutional neural network model', Symmetry, vol. 11, no. 3, pp. 1-13.

[18.] Xie X, Ma Y, Liu B, He J, Li S and Wang H (2020) A Deep-Learning-Based Real-Time Detector for Grape Leaf Diseases Using Improved Convolutional Neural Networks. *Front. Plant Sci.* 11:751. doi: 10.3389/fpls.2020.00751

[19.] Liu B, Ding Z, Tian L, He D, Li S and Wang H (2020) Grape Leaf Disease Identification Using Improved Deep Convolutional Neural Networks. *Front. Plant Sci.* 11:1082. doi: 10.3389/fpls.2020.01082

[20.] Zhang, X, Qiao, Y, Meng, F, Fan, C & Zhang, M 2018, 'Identification of maize leaf diseases using improved deep convolutional neural networks', IEEE Access, vol. 6, pp. 30370-30377.

[21.] Omrani, E, Khoshnevisan, B, Shamshirband, S, Saboohi, H, Anuar, NB & Nasir, MHNM 2014, 'Potential of radial basis function-based support vector regression for apple disease detection', Measurement: Journal of the International Measurement Confederation, vol. 55, pp. 512-519.

[22.] Zhu, J, Wu, A, Wang, X & Zhang, H 2020, 'Identification of grape diseases using image analysis and BP neural networks', Multimedia tools and applications, vol. 79, no. 21, pp. 14539-14551.

[23.] Jiang, P, Chen, Y, Liu, B, He, D & Liang, C 2019, 'Real-time detection of apple leaf diseases using deep learning approach based on improved convolutional neural networks', IEEE Access, vol. 7, pp. 59069-59080.

[24.] Zhang, S, Zhang, S, Zhang, C, Wang, X & Shi, Y 2019, 'Cucumber leaf disease identification with global pooling dilated convolutional neural network', Computers and Electronics in Agriculture, vol. 162, pp. 422-430.

[25.] Adeel, A, Khan, MA, Akram, T, Sharif, A, Yasmin, M, Saba, T & Javed, K 2020, 'Entropy-controlled deep features selection framework for grape leaf diseases recognition', Expert Systems, pp. 1-17.

[26.] Barbedo, JGA 2019, 'Plant disease identification from individual lesions and spots using deep learning', Biosystems Engineering, vol. 180, pp.96-107.

[27.] Damayanti, F, Muntasa, A, Herawati, S, Yusuf, M & Rachmad, A 2020, 'Identification of Madura Tobacco Leaf Disease Using Gray- Level Co-Occurrence Matrix, Color Moments and Naïve Bayes', Journal of Physics: Conference Series, pp. 1-8.

[28.] Halder, M, Sarkar, A & Bahar, H 2019, 'Plant Disease Detection by Image Pro-cessing: A Literature Review', Image, vol. 3, no. 6, pp. 534-538.

[29.] Islam, M, Dinh, A, Wahid, K & Bhowmik, P 2017, 'Detection of potato diseases using image segmentation and multiclass support vector machine', IEEE 30th Canadian conference on electrical and computer engineering (CCECE), pp. 1-4.

[30.] Manik, FY, Saputra, SK & Ginting, DSB 2020, 'Plant Classification Based on Extraction Feature Gray Level Co-Occurrence Matrix Using k-nearest Neighbour', Journal of Physics: Conference Series, pp. 1-9.

[31.] Wahyuni, S, Zarlis, M & Widia Sembiring, R 2020, 'Digital Image Processing For Detecting Yellow Lines (Patch Yellow) In Palm Oil Plant Using Sobel Algorithm', Journal of Physics: Conference Series, vol. 1471, no. 1, pp. 1-7.

[32.] Zhang, S, Huang, W & Zhang, C 2019, 'Three-channel convolutional neural networks for vegetable leaf disease recognition', Cognitive Systems Research, vol. 53, pp. 31-41.

[33.] Zhang, S, You, Z & Wu, X 2019, 'Plant disease leaf image segmentation based on superpixel clustering and EM algorithm', Neural Computing and Applications, vol. 31, no. 2, pp. 1225-1232.

[34.] Vyas, S., Mukhija, M.K., Alaria, S.K. (2023). An Efficient Approach for Plant Leaf Species Identification Based on SVM and SMO and Performance Improvement. Intelligent Systems and Applications. Lecture Notes in Electrical Engineering, vol 959. Springer, Singapore. https://doi.org/10.1007/978-981-19-6581-4_1

[35.] Koul, Deepali, and Satish Kumar Alaria. "A new palm print recognition approach by using PCA & Gabor filter."." *International Journal on Future Revolution in Computer Science & Communication Engineering* 4 (2018): 38-45.

[36.] Dogiwal, Sanwta Ram, Y. S. Shishodia, Abhay Upadhyaya, Hanuman Ram, and Satish Kumar Alaria. "Image Preprocessing Methods in Image Recognition." *International Journal of Computers and Distributed Systems* 1, no. 3 (2012): 96-99.

[37.] Alaria, S. K. "A.. Raj, V. Sharma, and V. Kumar."Simulation and Analysis of Hand Gesture Recognition for Indian Sign Language Using CNN"." *International Journal on Recent and Innovation Trends in Computing and Communication* 10, no. 4 (2022): 10-14.