Plant Diseases Detection Using Image Processing and Suggest Pesticides and Managements

***Abstract*-- Agriculture is one of Sri Lanka's most significant industries as they contribute to the economy of the country. But farmers are facing a huge economic loss due to the diseases which spread among their crops. Not only experienced farmers but inexperienced youths are also attempting to participate in farming. But these people lack knowledge about farming and insects which can cause harm to the plants. So they need a solution for that problem. There is no public internet forum for farmers in Sri Lanka to help them out. We intended to create a mobile application for farmers that would allow them to readily access information and help farmers gather knowledge easily to tackle their problems. There are several studies on illness detection. But there is no implemented system currently present in Sri Lanka in order to help farmers identify diseases. This solution uses the Convolutional Neural Network (CNN) and TensorFlow framework, as well as the Keras library, to classify and recognize images. The VGG16 model was tweaked so that it could be used for feature extraction. In the implementation machine learning as well as deep learning are employed. Application's main function is to identify pests and diseases. Here, we take a photo of a leaf that has been affected by an insect or illness, then use an application to evaluate the image and forecast the outcome. The projected outcome directs the user to the most appropriate insecticides. This solution will have a significant influence on the agricultural area, and when it reaches the Sri Lankan agriculture department, there will be another major target for this application. The suggested application's significance is to improve plant product rate and generate large income at no expense.**

***Index Terms*—image processing, plant disease detection, pesticide suggestion**

I. INTRODUCTION

Sri Lanka is an agriculture country. A relatively huge amount comparing to the total population in Sri Lanka are farmers. 27.1 percent of Sri Lanka's population is involved in agricultural pursuits. In 2020, agriculture contributed for 7.4% of GDP (gross domestic product). Sri Lanka is mostly an agricultural country, with a concentration on vegetable farming. Each leaf should be free of pests and illnesses in order to produce healthy plants. Farmers may have trouble cultivating plants if they do not have adequate understanding of the corresponding pesticides. The majority of Sri Lankan farmers have little knowledge about plant management. Plant diseases are one of the major problems which farmers are facing today. Many experience farmers know about diseases and how to manage them. but today many new youngers are involved in farming. Also well educated people are interested in farming and home gardening. In this situation new comers to farming don't know about plant diseases and how to use pesticide management to particular diseases. So they need someone who has knowledge about plant diseases and pesticides management to treat the affected plants. To fix the situation, they must meet and seek guidance from an agriculture department officer or an agricultural specialist. Farmers must demonstrate sample-affected leaf to the experts, which takes time. Farmers must thus apply proper pesticides; otherwise, they may encounter challenges such as decreased agricultural production quality and quantity. it is time consuming. To control above problems, pesticides functions will be implement. Recognize the indications of insect attack and nutritional shortages in the leaves and find a remedy early on will assist to reduce the impact of faults in the plant. When symptoms develop, samples of the afflicted plant parts are gathered and analyzed using the conventional method, which is time consuming, labor involved, and costly. Therefore, a system has been implemented to detecting plant diseases and suggests pesticides in this research project.

The system's goal is to provide precise and scalable visual cues to identify Vegetables plant diseases, particularly pumpkin, tomato, chili plants diseases, and to provide a solution for farmers by providing exact information about diseases, pests, nutrient imbalance, and how to apply fertilizer and pesticides effectively.

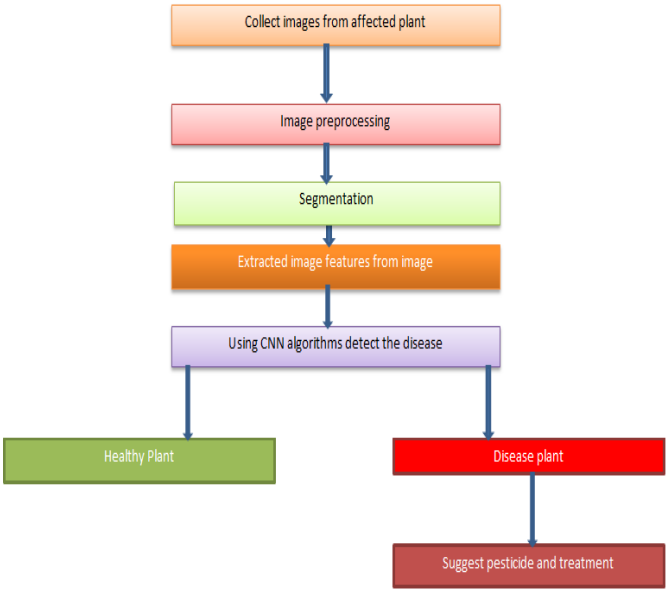

Figure 1 overview of techniques

Detection of disease is achieved by image processing. Here, some technologies are used to data training. Machine Learning and Deep Learning which are the branches of Artificial Intelligence. Deep Learning is also type of Machine Learning. Deep learning is modeled using architecture modeled after the human brain. Where as in machine learning the human involvement is minimized most of the time. Deep Learning is used for data training and testing. And, in this research TensorFlow library, Keras library and NumPy library are used in python language. TensorFlow combines a variety of techniques and models to allow users to build deep neural networks for applications such as classifying as well as identification of image. With the use of augmentation of picture data, the models which are provided by the ImageDataGenetor in Keras Framework. Image data augmentation is used to increase the size of the training dataset in order to improve the model's performance and generalization capabilities. In order to do numeral computation and store structure such as matrix and multi dimensional array a python library called Num Py is used. This can perform various form of statistical and algebraic computations.

Convolutional neural network (CNN) algorithm is used in this research. The over all theory of the functions of CNN is that it is able to take a picture as an input, then assign weights and biases to depending on the various aspects of images and identify them. Only a subarea of the preceding layer's input is received by neurons in a convolutional layer. Each neuron in a completely linked layer receives information from all elements of the preceding layer. A CNN extracts characteristics from photos and uses them to make decisions. Feature detection is taught to CNNs over tens or hundreds of hidden layers.

II. RELATED WORK

A number of researchers give a brief review of ways to detecting plant illness, nutrient insufficiency, and insect assault using image processing with machine learning in this paper. This section discusses the conclusions made from the earlier researcher's study publications on the subject.

In 1998[1], this study examines and analyzes several techniques for handwritten character recognition using a conventional handwritten digit recognition problem. Convolutional neural networks have been proven to surpass all other approaches when it comes to dealing with the variety of 2D forms. Modeling of language , segmentation of recognition, field extraction, are all components of real-world document recognition systems. Graph transformer networks (GTN), a novel learning paradigm, in order to reduce the overall metric of performance the gradient based techniques which enables multi component modules to be taught all over the world.. Descriptions of 2 online handwriting recognition system are provided. Global training and graph transformers presents benefits according to the experiments done. Also discussion on reading a bank check using graph transformer also have been done. With the combination global training methods and CNN character recognizer the checks which are personal as well as business achieves accuracy on records. It is in commercial use and reads millions of checks every day.

In 2009 [2], The focus of this study is on obtaining paddy characteristics from off-line images. Image capture is followed by automated using a local entropy threshold as a threshold and the Otsu technique to transform RGB pictures to binary images. The noise is removed using a morphological algorithm and the region filling approach. Then, from paddy leaf pictures, image features such as lesion kind, the colors of border, spot and diseased paddy leaf are retrieved. As a result of using the production rule approach, paddy illnesses may be identified with an accuracy rate of 94.7 percent.

In 2020 [3], the study focuses on identifying the patterns in plants which can be useful in finding the diseases using visual elements that are common in the diseased plants.. To have a prolonged healthy agricultural activity, constant care of health and detecting diseases in plants is crucial. But constantly watching over the plant in real life could be tedious task in real life, especially if there is a huge number of plants to be scanned with our bare eyes and if it happens in action it might take a lot of the valuable time. Also lack of knowledge about plant diseases and how they visually affect a plant make the task of detecting a diseases almost an impossible chore So as a solution to this complication, image processing used to identify the patterns which are similar to the dataset which has pictures of the diseased leaves and the ones which are fed for identifying the disease. The data set contains a varied range of plants in form of a picture. And further the users in this study are led into commercial websites where pesticides are up for sale with their user manual is present. This activity is considered to be beneficial when considering the evaluation of MRPs of pesticides and to cure the disease using it. Also this study gives a hands to those who grow their plants in an green house environment.

In 2016 [4], the objectives of this article is to provide an insight of how image processing technology can be applied to represent various plant diseases. With the use of image processing, the diseases cause by microorganisms such as virus, bacteria and fungus are easily identified. The use of only one's eyes to identify illnesses is ineffective. As pesticides are not cleaned correctly, they create dangerous chronic illnesses in humans. Also when there is an over the limit usage, it can lead to more harms such as loss of quality and also loss in the amount of production. So, image processing methodologies might be useful in agricultural applications for detecting and classifying illnesses.

In 2020 [5], Yallappa D et al presented a basic drone vehicle and improved its performance for spraying pesticides on crop plants. The major experimental investigation was conducted on crops such as groundnut and paddy. BLDC motors, six cells in two lithium polymer batteries with 8000mAh capacity, 5kg payload capacity, a 12v dc motor, a camera for front view, a fluid tank to spray pesticide with a 5 liter capacity, GPS, and other components make up this drone system. The major focus will be on evaluating the performance characteristics of a drone-mounted sprayer as well as the economics of its operation. This technique is utilized in areas where humans cannot readily travel, such as orchard crops. The field efficiency is between 62.84 and 60.00, and the forward speed is at 3.6 km/h1.

In 2019 [6], in his study effort, Santhosh Kumar S et al highlighted the major important concerns and different obstacles that arise when analyzing plant diseases. Various plant diseases, such as rust, yellow leaf, rotting of leaf, and leaf curling , are described in depth, along with their symptoms and appearance. Using the Gabor wavelet transform and the hybrid clustering approach, mobile-based client-server architecture is utilized to identify leaf disease. The ANN model, a deep learning approach, was utilized to find fungus of different kinds in cucumber plant leaves. It aids in illness categorization and detection. Automation, multi-feature, and genetic algorithms are used to identify plant diseases. The literature review detailed the BP neural network, the method for segmenting plant leaf disease, and the Gabor wavelet transform, concluding that image processing should be the major perspective for disease detection in plants.

In 2017 [7], This study paper provides a brief overview of the most deadly cucumber mildew, Downey mildew, which spreads rapidly in cucumber plants. There is a quick comparison between the human eye assessment of illness and the image analysis approach. A developed downy mildew spot extraction algorithm has been suggested, which employs the leaf scanning approach to determine the illness index.

The mildew picture has an accuracy of 98.3 percent.

In 2015 [8], this study offers a computationally efficient approach for detecting paddy leaf disease. Segmenting of images, feature extraction, and classification are the three steps of the suggested methods. The image segmentation technique K-means cluster algorithm is utilized to find diseased leaf sections. The paddy leaf picture is used to extract features in thefeature extraction step. These characteristics are utilized as input to the classifier in order to classify the data. The classifier is utilized as an artificial neural network in this experiment. For many years, numerous researchers have been researching on real-time plant leaf diseases. This project will be used to identify leaf disease in real time in the future. Farmers would benefit greatly from this initiative since it will allow them to detect paddy illnesses at an early stage.

In 2010 [9], the study focuses on a framework which enables the detection of plant leaf diseases. According to studies, depending solely on professional observation with the naked eye to diagnose such illnesses might be too expensive, particularly in underdeveloped nations. Providing quick, automated, low-cost, and accurate image-processing-based solutions for that task can be quite feasible. The suggested framework is image-processing-based and consists of the key phases listed below. The pictures at hand are segmented using the K-Means approach in the first phase, and then fed through a pre-trained neural network in the second. They use a collection of leaf pictures from Jordan's Al-Ghor region as a testbed. The results of our experiments show that the suggested method can greatly aid in the accurate and automated detection of diseased leaves. The statistical based Neural network classifier was able to detect and also categorize the disease with the accuracy of 93%.

In 2013 [10], this study is structured in away to analyze and categorize, brown pot , leaf blast, bacterial blight and tungro which are the most common Indonesian paddy diseases with the help of fractal descriptors. The pictures of the lesions were manually extracted. Afterwards, the S component descriptors out of each lesion image were used in a stochastic neural network classification technique. When it came to detecting illnesses, this approach had an accuracy of at least 83.00 percent. If coupled with additional characteristics, this approach has the potential to be utilized as one of the features, particularly when two illnesses of similar hue are involved.

In 2018 [11], the main aim of their study is to classify and detect diseases of plants using image of the leaves with the help of software system which consist image processing, by testing and improving the image processing algorithm. This study also suggest that identifying the illness can be consumes a lot of time and expensive as it might require a lot of labor specifically in the rural areas of the countries as it should be done with just the observation of the normal eyes scanning carefully in order to detect a disease. As a result, we offer a method based on image processing that is quick, automated, inexpensive, and accurate. The first process consists of creation of color transformation structure for the data which is the image of the leaf. Second, this structure is changed by the application of color space transformation.
Then the pictures are segmented using the K-means clustering .On the third step, texture characteristics of segmentation of the diseased object is estimated.. On the final stage the received feature a fed into the already trained neural network model.

In this paper [12], Nikita Goel et al. presented a model that uses a trained collection of pomegranate leaf pictures to inform about the illness of the leaf. This study introduces a new computer-assisted segmentation and classification approach. Obtaining image, picture enhancement, segmentation of image, extracting features, classifying image, and image accuracy are included in the first stage of this article, which aids in pre-processing. The k-means method is employed on all pictures in dataset in the second stage. The feature extraction step, which includes color and form characteristics, is completed in the third stage.

Gina S. Tumang spoke on the most prevalent illnesses that affect mango trees. Anthracnose, fruit borer, and sooty mold are the culprits. The illnesses are detected utilizing image processing techniques such as multi-SVM and GLCM. For image classification, a support vector machine classifier is employed. Image capture, image pre-processing, segmenting using k-means algorithm, feature extraction using gray level co-occurrence matrix, are all utilized. This method of detecting illness in plants was shown to be 85 percent accurate. [13]

In this paper [14], an investigation was conducted by author ananthi, s.vishnovarthini into the categorization of plant leaves and the identification of illnesses that damage them. They focused on controlling pests that cause illnesses that reduce crop output, and they mostly summarized machine learning and image processing approaches such as BPNN, SVM, and KNN.

In 2018 [15], Tomatoes are one of India's most important agricultural crops. In that situation, early illness identification in the plant is critical. The tomato plant diagnosis would be aided by image processing and a prediction model based on IoT sensors., Extreme Learning Machine ,Support Vector Machines, ANN classifier, Deep learning, and other modeling approaches have been created in response to a variety of symptoms and illnesses identified in tomato plants. The majority of the times, pictures are utilized as system input.

III. METHODOLOGY

Image processing to identify plant diseases is the most often addressed topic at various stages. To train and test the data set, a combination of DL and ML approaches are utilized. CNN algorithms are used to extract and categorize information from pictures. Mainly, five steps are used to get accurate output.

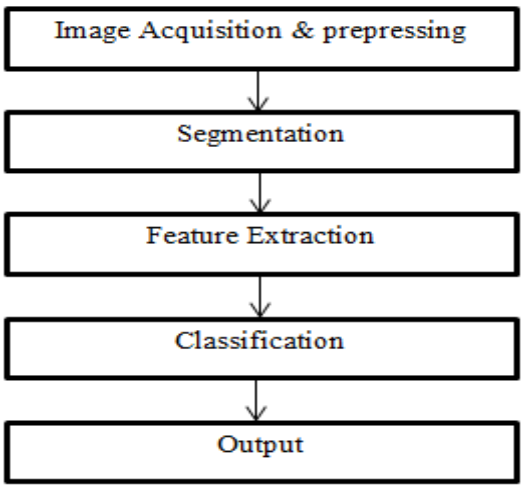

Figure 2 Image Processing Steps

a) Obtaining images

Images of plant leaves are photographed with professional camera or mobile smart phones in order to achieve the optimal resolution and size of image required for the study. Also the images could acquired from the internet if they satisfy the requirements. The recorded image was preprocessed to improve image quality, such as removing the non-uniform background and resizing the image with the dimensions as RGB training data. To eliminate the noise from the picture, nonlinear spatial filtering is used.

b) Image Segmentation

This steps requires to optimize image in a way to make them viable for analyzation. So accordingly the representation of the image is reduced. This phase is also the essential method to image processing because it is the foundation of feature extraction. Otsu' algorithm and k-means clustering are among some of the methods in which image may get segmented. Based on a collection of characteristics, the k-means clustering algorithm divides objects or pixels into K groups. The reduction of total of the squares of distance between items and their groups help in categorizing .

First, impacted leaves will be detected, followed by affected parts and green color pixels. The threshold equation will be used to demonstrate the threshold value of the green color pixels.

$$g(x,y) = \begin{cases} 1, & if\ f(x,y) \geq T \\ 0, & if\ f(x,y) < T \end{cases}$$

Enhance the image utilizing K means methods, which result in a high-quality image for detecting nutritional insufficiency. Edge detection technology will be utilized to determine the insect's carrying pattern, which will be used to detect the pest.

c) Image Feature extraction and Classification

Following segmentation, features from the targeted areas are retrieved using feature extraction techniques. Skewness Asymmetry in the distribution of pixels in the given window around its mean

$$m = \frac{1}{mn}\sum_{i=1}^{m}\sum_{j=1}^{n} p(i,j)$$

$$m = \frac{1}{mn}\sum_{i=1}^{m}\sum_{j=1}^{n}\left(\frac{p(i,j)-m}{\sigma}\right)^3$$

Pixels in the same neighborhood have different intensities.

$$C = \sum_{i,j}(i-j)^2 P(i,j)$$

The component extracts the image model using a deep learning method and approach after extracting the features from the damaged leaf picture. As a result, the picture model is reliant on an external library. Tensor flow is a toolkit for dealing with machine learning techniques like the convolutional neural network (CNN).

It's used to classify images, such as determining whether or not a leaf is healthy. The neural networks (CNN) method learns a complicated image in small layers, and each image will have many layers. In addition, keras is most likely tensorflow when it comes to deep neural networks. For this study, we collected more than 4000 plant diseases images and healthy images from field and internet for data training and testing.

d) Testing images using classification techniques

Determining whether a leave is healthy or not using the image provided is done in this classifying phase. Based upon the categorization arrived using the prior training the unhealthy ones are identified. A classifier is created in MATLAB environment in this study.

Among the huge varieties of classifiers such as Back propagation network, Naive Bayes, Decision Trees, Support Vector Machine (SVM),and K-nearest neighbor, SVM is chosen even though each one of them has its pros and cons, This is because of the reliable approach that is found in the SVM.

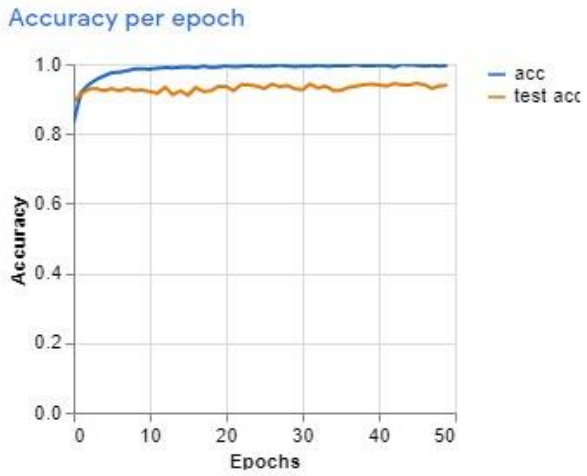

Figure 3 Accuracy per Epoch for Tomato Diseases

The figure shows accuracy of tomato leaf diseases detection in both of training and testing data.

e) Output

Using the training data file, created python file to get accurate result.

Accuracy per class

CLASS	ACCURACY	# SAMPLES
Bacterial_spot	0.97	256
Early_blight	0.90	290
healthy	0.97	289
Late_blight	0.96	278
Leaf_Mold	0.93	283
Septoria_leaf_spot	0.90	262
Spider	0.91	262
Target_Spot	0.90	275
mosaic_virus	0.97	269
Yellow_Leaf_Curl_V...	0.98	295

Figure 4 Image Processing Steps

Figure 4 depicts the tomato disease samples and accuracy. When utilizing 50 epochs in data training, the most accurate result is close to 98 percent.

IV. CONCLUSION

The system's main function is to detect insect assaults and nutritional deficiencies and recommend pesticides for farmers to purchase through our "E-commerce for Farmers system." Methods concentrate on image processing techniques used to damaged or diseased leaf images, as well as the application of machine learning. CNN and K means clustering approach detection methods may be efficient and accurate in classifying the ailment and suggesting solutions. For automated detection, this approach is efficient and accurate.

V. FUTURE WORKS

Three varieties of vegetables are utilized to collect data for identifying illnesses on their leaves in this study. This study focuses on tomato, chili, and pumpkin plant diseases. English is also utilized at the research level. In the future, we will concentrate on all plant diseases. This approach will assist farmers in the future in purchasing pesticides for plant diseases. In addition, Tamil and Sinhala are the native languages of the people of Sri Lanka. Thus, Tamil and Sinhala will be used in research in the future. Accuracy and speed will be prioritized in the future.

REFERENCES

[1] Y. Lecun, L. Bottou, Y. Bengio, P. Haffner, "Gradient-based learning applied to document recognition," Proceedings of the IEEE, 1998.

[2] Kurniawati, N. N. et al, "Investigation on image processing techniques for diagnosing paddy diseases" (2009).

[3] Mr.V Suresh, D Gopinath, M Hemavarthini., K Jayanthan, Mohana Krishnan, "Plant Disease Detection using Image Processing., " International Journal of Engineering Research & Technology (IJERT), 2020.

[4] Shivani K. Tichkule, Dhanashri. H. Gawali, "Plant diseases detection using image processing techniques" Online International Conference on Green Engineering and Technologies (IC-GET), 2016.

[5] Yallappa D, "development and evaluation of drone mounted sprayer for pesticide applications to crops" IEEE, 2017.

[6] Santhosh Kumar.S, B.K.Raghavendra, "Diseases Detection of Various Plant Leaf Using Image Processing Techniques" A Review, IEEE, 2019.

[7] Jingzhu Li, Peng Wang, changing geng, "The Disease Assessment of Cucumber Downy Mildew Based on Image Processing" IEEE, 2017.

[8] RadhikaDeshmukh, ManjushaDeshmukh, "Detection of Paddy Leaf Diseases" ICAST, 2015.

[9] Dheeb Al Bashish, Malik Braik, Sulieman Bani-Ahmad, "A framework for detection and classification of plant leaf and stem diseases" IEEE, 2010.

[10] Auzi Asfarian, Yeni Herdiyeni, Aunu Rauf, Kikin Hamzah Mutaqin, "Paddy diseases identification with texture analysis using fractal descriptors based on Fourier spectrum," The International Conference on Computer, Control, Informatics and its Applications (IC3INA), IEEE, 2013.

[11] Niket Amoda, Bharat Jadhav, Pradnya Kurle, Sharon Kunder, Smeeta Naikwadi, "DETECTION AND CLASSIFICATION OF PLANT DISEASES BY IMAGE PROCESSING" IJERT, 2018

[12] Nikita Goel, Dhruv Jain, Adwitiya Sinha, “Prediction Model for Automated Leaf Disease Detection & Analysis” IEEE, 2019.

[13] Gina S. Tumang, “Pests and Diseases Identification in Mango using MATLAB”, IEEE, 2019.

[14] Ananthi, S.Vishnu Varthini, “Detection & Classification of Plant Leaf Diseases”, International Journal of Research in Engineering & Applied Sciences, Volume 2, Issue 2, February 2012.

[15] Shradha Verma, Anuradha Chug, Amit Prakash Singh, “Prediction Models for Identification and Diagnosis of Tomato Plant Diseases”, IEEE, 2018.