2022 IEEE 7th International conference for Convergence in Technology (I2CT)
Pune, India. Apr 07-09, 2022

Plant Diseases Detection Using Image Processing and Suggest Pesticides and Managements

Lathusha Sritharan
Faculty of Computing
Sri Lanka Institute of Information Technology
Malabe, SriLanka
IT18181784@my.sliit.lk

Manikavasagar Anjanan
Faculty of Computing
Sri Lanka Institute of Information Technology
Malabe, SriLanka
IT18186048@my.sliit.lk

Anjalie Gamage
Faculty of Computing
Sri Lanka Institute of Information Technology
Malabe, SriLanka
anjalie.g@sliit.lk

2022 IEEE 7th International conference for Convergence in Technology (I2CT) | 978-1-6654-2168-3/22/$31.00 ©2022 IEEE | DOI: 10.1109/I2CT54291.2022.9825082

***Abstract*- Various plant diseases affect farmers all over the world and there is a very small amount of solutions available online for free in order to assist. In Sri Lanka, in order to address this issue, we have done a study which outputs a mobile application which utilizes image processing and recommend pesticides according to corresponding disease. The disease detection method includes image acquisition, image pre-processing, image segmentation, feature extraction, and classification. This study looked at methods for identifying plant ailments using photos of their leaves. This work also presented unique segmentation and feature extraction techniques for plant disease identification. For feature extraction, the CNN algorithm is utilized. This research paper may be a revolutionary approach to diagnosing plant illnesses by employing a deep convolutional neural network that has been trained and fine-tuned to suit a database of a plant's leaves gathered independently for distinct plant diseases. At the end of the study we achieved an accuracy of 98 percent in detecting the plant diseases and further on implemented mobile system which can suggest pesticide accordingly.**

Keywords—Image processing, features extraction, plant diseases detection, Tensorflow, Keras, Deep learning, CNN algorithm, Pesticide Suggestion

I. INTRODUCTION

Sri Lanka is an agricultural country. Most of the people 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) [20]. 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 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, chilli, paddy 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.

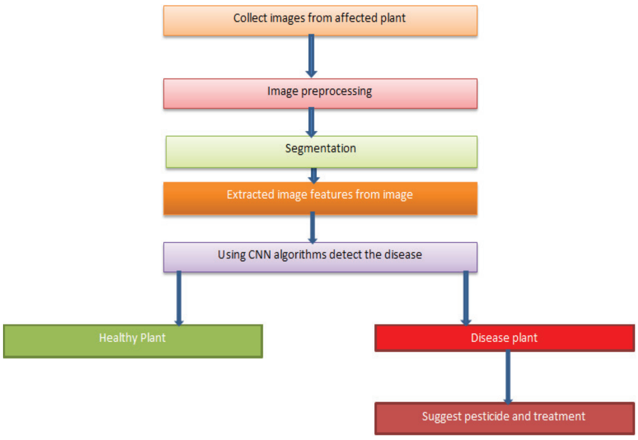


Fig.1.Overview of techniques

II. RELATED WORK

In this publication, a number of researchers provide a brief overview of methods for detecting plant sickness, nutritional deficiency, and insect attack using image processing and machine learning approaches. This section describes the findings of the previous researcher's study papers on the issue.

Lecun et el [1] has done a study which 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. Real-world document recognition systems include field extraction, segmentation recognition, and language modeling. Graph transformer networks (GTN), a revolutionary learning paradigm that allows such multimodule systems to be trained globally using gradient-based approaches in order to minimize an overall performance measure. There are two online handwriting recognition systems reported. Experiments demonstrate the advantages of global training and the flexibility of graph transformer networks. A graph transformer network for reading a bank check is also discussed. By merging convolutional neural network character recognizers with global training approaches, it offers record accuracy on commercial and personal cheques. It is in commercial use and reads millions of checks every day.

Nunik Noviana et el [2] has done a study which is on obtaining paddy characteristics from off-line images. Following image acquisition, automatic thresholding based on local entropy threshold and the Otsu approach is used to convert RGB images to binary images. The noise is removed using a morphological algorithm and the region filling approach. Then, image attributes such as lesion kind, border color, spot color, and broken paddy leaf color are extracted from paddy leaf images. As a result of using the production rule approach, paddy illnesses may be identified with an accuracy rate of 94.7 percent.

Suresh et al. [3] conducted plant disease research that included the evaluation of visually visible patterns on the plant. Plant health monitoring and disease detection are critical for the long-term sustainability of agriculture. It is difficult to monitor plant diseases manually. It takes a lot of work, as well as understanding of plant diseases and long processing periods. As a consequence, image processing is used to diagnose plant ailments by taking photographs of the leaves and comparing them to data sets. Several plants in image format are used in the data collecting. Aside from detection, buyers are directed to an e-commerce website that displays a variety of pesticides, along with their prices and directions for use. This website may be beneficial for comparing the MRPs of different pesticides and purchasing the one required for the illness indicated. The purpose of this article is to properly support and help greenhouse farmers.

Shivani et al. [4] conducted research to provide an overview of how image processing technology may be utilized to detect various plant diseases. Image processing aids in the identification of plant diseases caused by fungus, bacteria, or viruses. Using merely one's sight to detect infections is useless. As pesticides are not cleaned correctly, they create dangerous chronic illnesses in humans. Excessive use also affects plant nutritional quality. Farmers lose a substantial amount of crops as a result of this. As a consequence, image processing techniques might be effective for identifying and categorizing diseases in agricultural applications.

Yallappa D et al[5] shown how to increase the performance of a simple drone vehicle for spraying pesticides on agriculture plants. The major experimental investigation was conducted on crops such as groundnut and paddy. BLDC motor, 2 lithium polymer batteries with capacity 8000mAh, payload 5 kg, 6 cells, 12 V DC motor, front view camera, liquid tank for spraying pesticide with capacity 5 liters, GPS and other components Achieve this Drone from system. The focus is on assessing the performance characteristics of the atomizer attached to the drone and the economics of its operation. This technique is used in places where it is difficult for people to move, such as orchards. The field efficiency is between 62.84 and 60.00 and the running speed is 3.6 km / h1.

Santhosh Kumar S et al highlighted the major important concerns and different obstacles that arise when analyzing plant diseases [6]. We will explain in detail various plant diseases such as rust, coloroga, yellow leaf disease, root rot, and cigar disease, 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. Two types of fungus were detected in cucumber plant leaves using the ANN model, a deep learning method. 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, as well as the Gabor wavelet transform, resulting in the conclusion that image processing should be the primary focus for disease identification in plants.

Jingzhu Li et el [7] has done a study which provides a brief overview of Downey mildew, the most dangerous cucumber mildew, spreads swiftly in cucumber plants. The human eye judgment of disease and the image analysis technique are compared quickly. The leaf scanning technique has been used to calculate the disease index in a developed downy mildew spot extraction algorithm. The mildew picture has an accuracy of 98.3 percent.

Radhika Deshmukh et el [8] has done a study This provides a computationally effective method for identifying paddy leaf disease. The three phases of the proposed approaches are image segmentation, feature extraction, and classification. The image segmentation technique K-means clustering is utilized to find diseased leaf sections. The paddy leaf picture is used to extract features in the feature 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.

Dheeb Al Bashish et el [9], In this research study, we presented and examined a framework for identifying plant leaf/stem illnesses. According to studies, depending solely on professional observation with the naked eye to diagnose such illnesses might be too expensive, particularly in underdeveloped nations. Image-processing-based solutions that are rapid, automated, low-cost, and accurate may be achievable. The proposed framework is based on image processing and consists of the main phases listed below. In the first phase, the available images are segmented according to the KMeans approach, and in the second phase, they are fed by a pre-trained neural network. As a test setting, they employ a set of leaf photographs from Jordan's Argol area. The results of our experiments show that the suggested method can greatly aid in the accurate and automated identification of leaf diseases. The neural network classifier, which was built on the basis of statistical classification, worked well and was able to identify and classify the analyzed diseases with an accuracy of approximately 93%.

Auzi Asfarian et al. [10] conducted research to determine Indonesia's four most frequent paddy illnesses. Diseases examined using fractal descriptors include bacterial leaf blight, brown spots, Leaf blast, and tungro. The pictures of the lesions were manually extracted. The descriptor of the S component of each lesion image was then used in the stochastic neural network classification procedure. 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.

Niket Amoda et el [11] has done a project with a purpose to create, test, and improve an image processing software-based system for detecting and classifying plant leaf disease. However, studies suggest that depending solely on specialists' naked eye observation to diagnose and categorize illnesses may be time-consuming and costly, particularly in rural areas and poor nations. As a result, we offer a method based on image processing that is quick, automated, inexpensive, and accurate. We build a color transformation structure for the RGB leaf picture in the first phase, and then we apply color space transformation to the color transformation structure in the second phase. The pictures are then segmented using the K-means clustering approach in the second step. The texture characteristics for the segmented infected objects are calculated in the third step. Finally, the retrieved features are sent into a pre-trained neural network in the fourth phase.

Nikita Goel et el [12] 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. Image acquisition, picture enhancement, image segmentation, feature extraction, image classification, and image accuracy are included in the first stage of this article, which aids in pre-processing. The k-means method is applied to all pictures in the dataset in the second stage. The feature extraction step, which includes color and form characteristics, is completed in the third stage.

Gina S. Tumang et el, 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 acquisition, image preprocessing, k-means segmentation, gray level co-occurrence matrix or GLCM feature extraction, and gray level co-occurrence matrix or GLCM feature extraction are all used. This method of detecting illness in plants was shown to be 85 percent accurate. [13]

An investigation was conducted by Ananthi et el[14], 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.

Shradha Verma et el [15] found 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. ANN Support Vector Machines (SVM), Extreme Learning Machine (ELM) 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. Here, some technologies are used to data training. 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 image recognition and classification. The ImageDataGenerator class in the Keras deep learning neural network framework allows you to fit models using picture data augmentation. Image data augmentation is used to expand the training dataset in order to enhance model performance and generalization capabilities. NumPy is a free and open-source Python library for numerical computations. NumPy comes with a multi-dimensional array and matrix data structures. It is capable of performing a wide range of mathematical operations on arrays, including trigonometric, statistical, and algebraic procedures.

A. Convolutional Neural Network (CNN)

CNNs have lately gained popularity, and DL is the most common architecture because DL models, like the human brain, may learn substantial properties from input photos at numerous convolutional layers. With a low mistake rate and excellent classification accuracy, DL can tackle complicated problems exceptionally successfully and fast [21]. The DL model is made up of many components (convolutional, pooling layer, and fully connected layers, and activation functions).

It was proposed to prepare the profound convolutional neural organization for extracting an image grouping model from a dataset. The Tensor Flow programming framework is free and open-source that uses information stream diagrams to do mathematical calculations. The diagram's hubs represent numerical jobs, while the diagram's edges represent the multidimensional information displays (tensors) that are shared between them. With a single API, you may send calculations to at least one CPU or GPU in a work area, worker, or, on the other hand, a mobile phone. Tensor Flow was created by scientists and designers working regarding the Google Brain Group, which is part of Google's Machine Intelligence research division and is responsible for AI and deep neural organization research; however the framework is broad enough to be relevant in a number of contexts. A convolutional neural network is a sort of feed-forward fake neural network in artificial intelligence in which the network architecture between its neurons is triggered by the connection of the creature visual brain. Individual cortical neurons respond to changes in a small area of the room known as the responsive field. The receptive fields of different neurons encompass just a portion of the visual field, with the purpose of tiling the visual field. A convolution activity can be used to mathematically simulate the responsiveness of a single neuron to increases in its open field. Natural cycles pushed the development of convolutional networks, which are multilayer perceptron variants designed to need little pre-handling. They're used in image and video recognition, regular language preparation, and recommender systems, among other things. Many layers of receptive fields make up convolutional neural networks

(CNNs). These are small clusters of neurons that pass through sections of the information image. To produce a higher-goal representation of the initial image, These assortments' yields are tiled such that their information districts are covered; This process is repeated for each additional layer. Tiling allows CNNs to survive the interpretation of the information picture. Pooling layers in convolutional networks can be local or global, and they combine the yields of neuron bunches. They, too, are made up of various combinations of convolutional and fully associated layers, with point-savvy nonlinearity applied at the end of or after each layer. A convolution approach on small data areas is known to reduce the number of free borders and enhance speculation. Convolutional networks provide a lot of flexibility, including using shared weight in convolutional layers, which means that every pixel in the layer uses the same channel (loads bank), which reduces memory use and increases performance. The layer's borders are comprised of a collection of learnable segments that contain a small open field but extend throughout the whole depth of the data volume. Amended Linear Units (Re LU) are a nonlinearity soaking option. This exploratory research adapts to rectifier constraints and improves accuracy at a low computing cost. The rectifier is an enactment job defined as: f(x) =max in the uncommon case of fake neural organizations (0, x).

Where x denotes the contribution to a neuron this is also known as inclination work and it is similar to half-wave correction in electrical design. This enactment work was initially introduced to a dynamical organization by Hahn failure et al. in a 2000 study in Nature with substantial organic inspirations and numerical avocations. It has been employed more successfully in convolutional networks than the often used strategic sigmoid (which is triggered by probability hypothesis; see strategic relapse) and its more grounded counterpart, the exaggerated digression. Since 2015, the rectifier has been the most well-known actuation technique for deep neural systems. Profound CNN with ReLU's trains that is a few times faster. This method is used to calculate the yield of each convolutional and totally related layer. Regardless of the yield, information standardization is not required; it is used after ReLU nonlinearity after the first and second convolutional layers because it reduces top-1 and top-5 error rates. Neurons inside a covered-up layer are divided into "include maps" in CNN. The neurons inside an element map have the same weight and inclination. The neurons inside the component map search for a component that is comparable. These neurons are unique in that they are linked to a variety of neurons in the bottom layer. As a result, for the first hidden layer, neurons inside an element guide will be connected with distinct districts of the data image. The veiled layer is divided into highlight maps, in which each neuron in an element map looks for a comparable component in any instance, at various locations across the information picture. The component map is essentially the result of applying convolution on an image. A CNN's convolutional layer is the center structural square [17]. The layer's boundaries are made up of several learnable channels (or pieces) that have a small open field but stretch out through the whole depth of the information volume. During the forward pass, each channel is convolved over the width and height of the input volume, recording the speck item between the channel sections as well as the data and constructing a 2-dimensional actuation guide for that channel. As a result, the organization learns channels that begin when it detects some type of highlight in some geographical circumstance in the data. Stacking the actuation maps for all channels along the profundity measurement constructs the convolution layer's complete yield volume. The depth, stride, and zero-padding hyper parameters govern the size of the convolutional layer's output volume [18].

Depth: The yield volume's depth regulates the number of neurons in the layer that interact with a comparable district of the information volume. These neurons will figure out how to begin for various information highlights. For instance, if the main Convolutional Layer takes the raw image as data, various neurons along the profundity measurement may commence in the presence of various arranged edges, or masses of shading.

Stride: It determines how depth columns are arranged around the spatial dimensions (width and height). When the stride is one, a new depth column of neurons is assigned to spatial places that are only one spatial unit apart. As a result, the receptive fields between the columns are extensively overlapping, resulting in huge output volumes [19]. Conversely, if larger strides are utilized, the receptive fields overlap less and the resulting output volume has lower spatial dimensions.

Zero padding occurs when we include a border of pixels, each with a value of zero, across the boundaries of the input images. This creates a kind of padding of zeros around the edge of the image, hence the term zero padding.

Steps in the CNN Model Conv2D: This is 2D Convolution Layer that generates a convolution kernel that is used with the layer's input to produce a tensor of outputs.

1. keras.layers.Conv2D (filters, kernel_size, strides= (1, 1),
2. Padding='valid', data_format=None, dilation_rate= (1, 1),
3. Activation=None, use_bias=True,
4. kernel_initializer='glorot_uniform',
5. bias_initializer='zeros', kernel_regularizer=None,
6. bias_regularizer=None, activity_regularizer=None,
7. kernel_constraint=None, bias_constraint=None

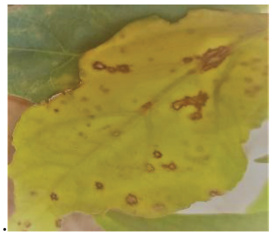
Maxpooling: Maxpooling is a pooling method that selects the best element from the feature map region covered by the filter. As a result, following max-pooling level, the output would be a feature map including the most essential features from the preceding feature map. [20].

Flatten: A 'Flatten' layer exists between the convolutional layer and hence the fully connected layer. Flattening converts a two-dimensional feature matrix into a vector, which is then fed into a completely connected neural network classifier.

Image Data Generator: Image Data Generator swiftly discovered Python generators that would automatically convert image data on disk into batches of preprocessed tensors.

B. Data Collection

Images were collected from farming field and kaggle. The data set has more than twenty thousand images for training and testing. Tomato, paddy, pumpkin, chilli leaves were collected. Eighteen varieties of diseases leaves were collected in above four plants.



Chilli bacterial spot

Tomato early blight

Rice brown spot

Pumpkin Powdery Mildew

Fig. 2. Images from different plants

Figure 2 shows the some images from collected data set. It shows tomato early blight, chilli bacterial spot, rice brown spot and pumpkin powdery mildew. Images were collected from four different types of plants and 18 different classes including diseases and healthy. Table 1 shows the collected plant diseases types.

TABLE I. COLLECTED PLANT DISEASES TYPES

Plant	Disease Type
Tomato	Bacterial Spot
	Target Spot
	Late Blight
	Early Blight
	Leaf Mold
	Healthy
	Two Spotted Spider Mite
	Mosaic Virus
	Septoria Leaf Spot
	Yellow Leaf Curl Virus
Paddy	Brown Spot
	Hispa
	Healthy
	Leaf Blast
Chilli	Bacterial Spot
	Leaf Curl
	Healthy
Pumpkin	Powdery Mildew

We collected data about pesticides, and management from Sri Lanka Agriculture department. Two books from agriculture department helped me to collect data about pest recommendation. These are “Pest Management Recommendations from Ministry of Agriculture, Department of Agriculture”, “and Pest Management Decision Guide from Permanent Crop Clinic Programme Sri Lanka”.

C. Data Preprocessing

Image pre-processing is an important stage in image processing. This technique will remove sounds in the photos, such as hair, clothes, and other artifacts. The major objective of image processing is to remove distracting aspects from damaged plants' images in order to improve image quality

D. Training Process

Effective training starts long before a trainer gives an individual training session and continues long after the training session is over. Training may be described as a process consisting of five interrelated phases or activities: evaluation, motivation, design, implementation, and assessment.

E. Basics steps for the image processing.

Mainly, five steps are used to get accurate output in image processing as shown in Fig.4 below.

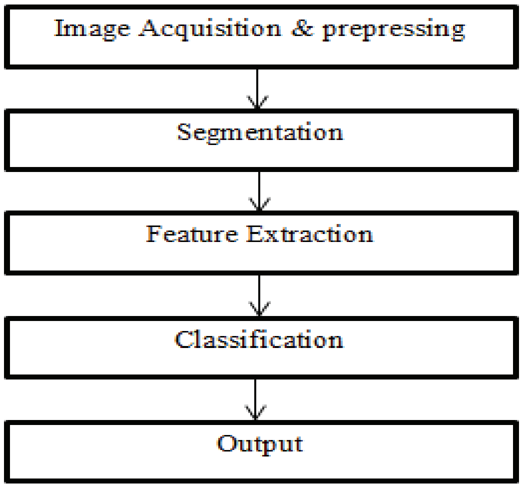


Fig.4. Image processing main steps

1) Image Acquisition

In this phase, digital media such as cameras and mobile phones are used to collect images of plant leaves at the required resolution and size. You may also get images from the internet. 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.

2) Image Segmentation

The goal of this step is to reduce an image's representation so that it gets more significant and simpler to comprehend. This phase is also the essential method to image processing because it is the foundation of feature extraction. Images may be segmented using a variety of approaches, including Otsu's algorithm, k-means clustering, and thresholding, among others. Based on a collection of characteristics, the k-means clustering algorithm divides objects or pixels into K groups. The distances between the items and their respective groups are reduced to reduce the total of squares.

First, impacted leafs will be detected, followed by affected parts and green color pixels. The threshold equation which is Eq.(1) will be used to demonstrate the threshold value of the green color pixels.

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

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.

3) *Image Feature extraction and Classification*

Following segmentation, features from the targeted areas are retrieved using feature extraction techniques. Eq.(2) and Eq.(3) shows that Skewness Asymmetry in the distribution of pixels in the given window around its mean.

$$m = \frac{1}{mn} \sum_{i=1}^{m} \sum_{j=1}^{n} p(i,j) \tag{2}$$

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

Eq. (4) shows that the pixels in the same neighborhood have different intensities.

$$C = \sum_{i,j} (i-j)^2 P(i,j) \tag{4}$$

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. I collected more than 20000 plant diseases images and healthy images from field and internet for data training and testing.

4) *Testing images using classification techniques*

The categorization stage involves assessing if the input image is healthy or not. If a sick image is identified, it has been classified into a number of diseases by various previous works. For classification, a software program called as a classifier must be written in MATLAB. Artificial neural networks (ANN), support vector machines (SVM), k-nearest neighbor (KNN), back propagation neural networks (BPNN), Nave Bayes, and Decision tree classifiers have all been used by researchers in recent years. The most commonly used classifier has been discovered to be SVM. Each classifier has advantages and disadvantages; nonetheless, SVM is a simple and reliable technique.

IV. RESULT AND OUTPUT

The research project's final product is an Android platform-based application. It outputs an APK file that can be installed on Android devices. A single app with many outcomes is available.

- Identify insect attacks
- Recognize disease
- Make pesticide and fertilizer recommendations
- Provide illness management techniques

The keras and tensorflow libraries are used to train the image models. The training of a dataset is depicted in the diagram below.

```python
# fit the model
# Run the cell. It will take some time to execute
r = model.fit_generator(
  training_set,
  validation_data=test_set,
  epochs=20,
  steps_per_epoch=len(training_set),
  validation_steps=len(test_set)
)
```

Fig.5. Python code for model training

The python code for model training is shown in Figure 5.To train the data set, we used 20 epochs. The number of epochs is a hyper parameter that determines how many times the learning algorithm goes through the whole training dataset. Every epoch, each sample in the training dataset has the opportunity to alter the internal model parameters. When the number of epochs is increased, the accuracy improves.

```
[ ]
/usr/local/lib/python3.7/dist-packages/tensorflow/python/keras/engine/training.py:1940: UserWarning: `Model.fit_generator` is deprecated and will be removed in a
  warnings.warn('`Model.fit_generator` is deprecated and '
Epoch 1/20
168/168 [==============================] - 1931s 11s/step - loss: 5.7826 - accuracy: 0.5026 - val_loss: 9.7745 - val_accuracy: 0.4203
Epoch 2/20
168/168 [==============================] - 752s 4s/step - loss: 4.6286 - accuracy: 0.5827 - val_loss: 9.6984 - val_accuracy: 0.3726
Epoch 3/20
168/168 [==============================] - 764s 5s/step - loss: 4.8803 - accuracy: 0.5902 - val_loss: 11.6621 - val_accuracy: 0.3308
Epoch 4/20
168/168 [==============================] - 784s 5s/step - loss: 4.1078 - accuracy: 0.6475 - val_loss: 13.9633 - val_accuracy: 0.3010
Epoch 5/20
168/168 [==============================] - 815s 5s/step - loss: 3.8874 - accuracy: 0.6624 - val_loss: 12.0303 - val_accuracy: 0.4367
Epoch 6/20
168/168 [==============================] - 791s 5s/step - loss: 4.0464 - accuracy: 0.6800 - val_loss: 14.4693 - val_accuracy: 0.3428
Epoch 7/20
168/168 [==============================] - 806s 5s/step - loss: 4.3827 - accuracy: 0.6721 - val_loss: 15.6312 - val_accuracy: 0.3458
Epoch 8/20
168/168 [==============================] - 817s 5s/step - loss: 3.4113 - accuracy: 0.7012 - val_loss: 13.9139 - val_accuracy: 0.3130
Epoch 9/20
168/168 [==============================] - 790s 5s/step - loss: 3.8501 - accuracy: 0.6975 - val_loss: 16.2866 - val_accuracy: 0.3547
Epoch 10/20
168/168 [==============================] - 777s 5s/step - loss: 3.5405 - accuracy: 0.7127 - val_loss: 17.0438 - val_accuracy: 0.2385
Epoch 11/20
168/168 [==============================] - 779s 5s/step - loss: 3.5570 - accuracy: 0.7120 - val_loss: 13.2005 - val_accuracy: 0.3994
Epoch 12/20
168/168 [==============================] - 779s 5s/step - loss: 3.6572 - accuracy: 0.7101 - val_loss: 15.1508 - val_accuracy: 0.3443
Epoch 13/20
```

Fig. 6. Data training

Figure 6 shows the data training of tomato leaves. The entire dataset was trained without any avoidance in this scenario. The dataset accuracy level is displayed in Figure 6 above. When training the same dataset repeatedly, the accuracy of the first epoch will be increased. Fig.6 shows that the training dataset's accuracy in epoch 12 is higher than in prior epochs

After finishing the training process, you must assess the accuracy level of the model you've constructed and determine if it's predicted accurately or not. In our case, we got the most accurate output for the tomato diseases.

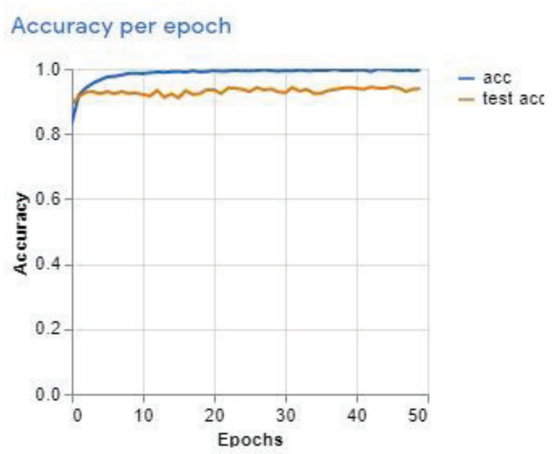


Fig: 7: Tomato Diseases Accuracy per Epoch

The accuracy level increases with epoch, as seen in Figure 7. When we utilize 20 epochs to train the tomato data set, the accuracy level is close to 98 percentages. In 20 to 50 epochs, the accuracy level does not vary. So the data's maximum accuracy is 98 percent, which we attained in 20 epochs.

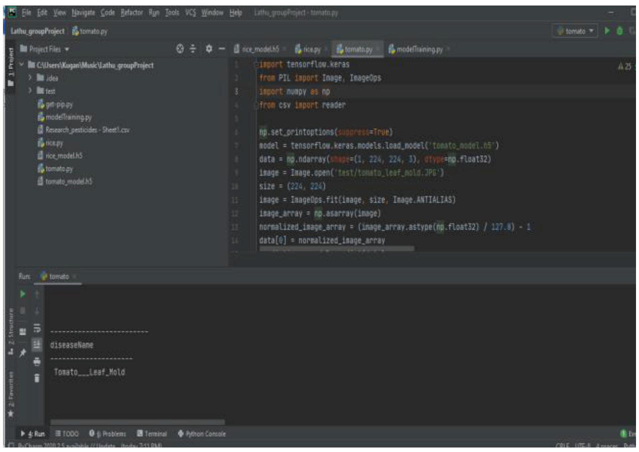


Fig. 8. Tomato Diseases

The final result of detecting tomato illness is shown in Fig 8. I obtained a precise result. Tomato leaf mold disease is the result of the input is tomato leaf mold image.

V. DISCUSSION

The system's main output is to control pests and diseases while also recommending insecticides and fertilizers. To address this issue, we devised a system that would monitor pest and disease activity and then intelligently recommend pesticides and fertilizers. Using image classification, the algorithm primarily recognizes the affected leaf and identifies illness on the leaf. When examining the photograph, the attention was mostly on the leaf's form and color. Deep learning CNN algorithms are used to do this. When recognizing an image, a dataset model must be compared, therefore develop an image model with the tensorflow and keras libraries. Our technology allows users to request assistance from an agriculture officer through email. Because certain pesticides and fertilizers are prohibited in Sri Lanka, and others are not appropriate for our terrain. Because of the differences in climate in different areas, insecticides and fertilizers are ineffective.

VI. 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.

VII. FUTURE WORK

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 farmer's 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] Nunik Noviana Kurniawati, Siti Norul Huda Sheikh Abdullah, Salwani Abdullah, Saad Abdullah, "Investigation on image processing techniques for diagnosing paddy diseases" 2009 International Conference of Soft Computing and Pattern Recognition, IEEE (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] Radhika Deshmukh, 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.

[16] "Agriculture in Sri Lanka - Wikipedia", En.wikipedia.org, 2021. [Online].Available:https://en.wikipedia.org/wiki/Agriculture_in_Sri_Lanka. [Accessed: 13- Nov- 2020].

[17] D.Gupta, "Architecture Of CNN | CNN Image Recognition", Analytics Vidhya, 2017. [Online]. Available: https://www.analyticsvidhya.com/blog/2017/06/architecture-of-convolutional-neural-networks-simplified-demystified/?spm=a2c6h.12873639.0.0.38ef1518HsMIw8. [Accessed: 18- Feb- 2021].

[18] "Convolutional neural network - Wikipedia", En.wikipedia.org, 2021. [Online].Available:https://en.wikipedia.org/wiki/Convolutional_neural_network. [Accessed: 03- May- 2021].

[19] S. Ajantha, G. Mahapriya, J. Mispa, S. Nithya, Measurement of Multimodal Brain Tumor Segmentation using Neural Network in IJERT vol. 5, issue 13

[20] "CNN | Introduction to Pooling Layer - GeeksforGeeks", GeeksforGeeks, 2021. [Online]. Available: https://www.geeksforgeeks.org/cnn-introduction-to-pooling-layer/. [Accessed: 17- Feb- 2021].

[21] Simonyan, K.; Zisserman, A. Very deep convolutional networks for large-scale image recognition.arXiv2014, arXiv:1409.1556.