RESEARCH ARTICLE OPEN ACCESS

RECOGNITION AND PREVENTION OF LEAF DISEASES USING MACHINE LEARNING

Sahana Vardhan[1], Sanjana M[2], Soumya A V[3], Yashaswini B[4], Mrs Anu H[5]

[1]Student, Electronics & Communication Dept., JSS Academy of Technical Education, Bangalore, India
[2]Student, Electronics & Communication Dept., JSS Academy of Technical Education, Bangalore, India
[3]Student, Electronics & Communication Dept., JSS Academy of Technical Education, Bangalore, India
[4]Student, Electronics & Communication Dept., JSS Academy of Technical Education, Bangalore, India
[5]Asst. Professor, Electronics & Communication Dept., JSS Academy of Technical Education, Bangalore, India

--*********************---

Abstract

India is a land of agriculture and primarily identified for growing variety of crops. Around 60% of Indian population hinge on agriculture. Crop sickness and the growth of weeds along with the crops may affect the income of farmers and they tend to reduce farm and forest productivity. In order to overcome these problems, many fast and accurate automated diagnostic techniques have been technologically advanced for leaf disease detection in the recent years. Hence, our project uses Image Processing and Machine Learning Techniques to detect leaf diseases. The CNN model is trained by a huge dataset to analyze the selected leaf image. If disease is detected, the disease name and its remedy is suggested by the model. The project includes the designing of a robot which sprays the pesticide to the diseased leaf. Also, if weed is detected, the robot plucks it. On a whole, this leads to implementation of technology in agriculture.

Keywords— **Image processing, machine learning, CNN**.

--*********************---

I. INTRODUCTION

Agriculture is a vital source of income for a country. The crop farms get wrecked mainly due to two reasons, i.e. natural calamities and pathogens. 98% of destruction is caused only because of these pathogens. Crop diseases are of three forms namely bacterial, fungal and spots. Visually observable patterns on the plants' leaf, stem or fruits indicate the disease. The chief identification of infected plant is its leaves. The numerous colored spots and patterns present on the leaf are very useful in detecting the disease.

The past scenario for plant disease identification involved direct eye observation, memorizing a specific set of disease as per the climate, season etc. These approaches were time consuming and undeniably inaccurate. Hence, automatic identification of plant diseases is a prime task as it may be remunerative for farmers to supervise large field of plants and to recognize diseases at preliminary stage using machine learning techniques.

II. LITERATURE SURVEY

Pooja V et al in 2017, [1] has proposed a system for "leaf disease detection and classification technique with the help of machine learning and image processing tools". The result achieved by them gave a recognition rate of 92.4%. The technique used for segmentation and classification are K-means clustering and Support Vector Machine respectively. Taking an input of 227 images, 5 different types plant diseases were detected. The drawback of this paper is that the system only detect the disease but doesn't suggest the pesticide.

Usha Kumari et al in 2019, [2] has showcased an algorithm for leaf disease detection using machine learning and image processing techniques. The average accuracy obtained is 92.5%. The algorithm used for feature extraction is K-means clustering and for classification Artificial Neural Networks. The system identifies leaf spots on cotton and tomato plants. It can detect 20 diseases in each plant. The software used here is MATLAB. As ANN works with numerical values, it was difficult to translate the problems into numerals.

Trimi Neha Tete et al 2017, [3] have discussed "different processes for segmentation technique which is applied for different lesion disease detection". Showing the importance of segmentation, they have given a comparative analysis between two algorithms i.e. threshold and k-means clustering and proved that k-means clustering gave relatively effective result compared to threshold.

Appasaheb Gargade et al 2019, [4] have developed a system to "detect leaf diseases and analyse leaf parameter of custard apple using digital image processing and machine learning". The accuracy of leaf parameter measurement in this method relies on number of pixels per unit area. The algorithms used for classification and clustering are SVM and k-means clustering respectively. There is significant scope for improvement and feature addition to leaf parameter analyses using digital image processing. Leaf and fruit weight estimation feature can be added.

In 2017, Monzurul Islam et al [5] developed a system to detect the potato diseases using image segmentation and multiclass SVM where they made use of publically available image database called plant village and carried the processing on 300 potato leaf images as a result two types of diseases i.e. Phytophthora Infestans and Alternaria Soloni were detected. The accuracy obtained was about 93.7%. They used L*a*b color space and used color threshold app from MATLAB instead of Otsu. They used 10 features for the feature extraction and the classification was done using SVM. The advantage of this system is that it required less computational effort and cost. The future scope is to access the system through smart phones.

In 2016, Pranjali B. Padol et al [6] studied two major grape diseases using SVM and classified them. The accuracy obtained was 88.89%. Different steps were encountered in the process where the major methods used were resizing, thresholding, Gaussian filtering, k-means clustering for segmentation. Totally 54 features were used to classify the disease types. This paper has given the comparison of SVM with NN and showed that NN is difficult to implement as it is difficult to know the number of neurons required alone for small database. The future expansion of this project is to automatically suggest and spray the pesticide.

Yusuke Kawasaki et al 2015 [7] explained the detection of disease in cucumber leaf using the training images alone. The implemented system achieved an accuracy of 94.9% in classifying cucumber leaves and the technique used is CNN. The advantage of CNN is that it can automatically achieve the necessary feature for classification and achieve high classification performance. CNN could solve difficult differentiation problems without segmenting the ROI (region of interest). This system can be further improved to suggest the pesticide.

Ms Shruthi U et al 2019, [8] reviewed on "machine learning classification techniques for plant disease detection". This paper concludes that high accuracy can be achieved by Convolutional Neural Network which detects a greater number of diseases of multiple crops. The data is classified into training data and testing data. These data are applied to various algorithms such as SVM, ANN, KNN, FUZZY and Deep Learning. The future work suggested is that, different classification techniques in machine learning like decision trees, Naive Bayes classifiers may be used for disease detection in plants if their accuracy is found to be more.

In 2017 Bharat Mishra et al, [9] gave a review on "various image processing techniques that are used in leaf disease detection". The comparison between various algorithms is shown and the importance of removing the background noise in acquired image is highlighted. This can be achieved using filters. The future scope aims at advancement of hybrid algorithms such as genetic algorithms, cuckoo optimization, particle swarm optimization etc in order to increase the recognition rate of final classification process.

Jyoti Shirahatti et al 2018, [10] surveyed on "plant disease identification using machine learning approach". According to them SVM algorithm should be favoured as it gives improved classification and prediction results which is helpful in identify the plant diseases. Comparison is made between machine learning techniques and their advantages and disadvantages are listed. This helps in detection of disease accurately and increases the productivity of crops.

Priyanka G Shinde et al 2017, [11] implemented a method for detection and classification of leaf disease using IP tools. The information gained about the disease is intimated to the farmer through his mobile using GSM module. To increase the speed and accuracy, they used RASPBERRY PI 3 MODEL B. The advantage of this system is that it gives the pesticide name corresponding to the detected disease. For segmentation and classification, the algorithm used are k-means a KNN. The work can be improved by using of quad copter to capture images in the farm at field level.

Figure 1: Septoria leaf spots

Figure 2: Early Blight

Figure 3: Healthy Tomato Leaf

Prof. Bhavana Patil et al 2017, [12] developed "plant monitoring system using image processing, RASPBERRY PI and IoT". Multiple spots and patterns on the plant leaf are used in detecting the disease accurately. K-means clustering is used for segmentation and SVM for classification. In this paper, the detection as well as the therapy for curing the disease is accomplished. It also uses GSM to send message to every kind of mobile handset.

Adhao, Asmita, Sarangdhar et al 2017, [13] implemented "a system for detection and controlling of diseases on cotton leaf along with soil quality monitoring". The overall accuracy of this system is 83.26%. It uses SVM based system for identification and classification of five cotton leaf diseases. The name of a disease along with its remedies is delivered to farmers using android app. The android app also displays soil parameters and even the farmers can turn on/off motor and sprinkler assembly according to the need. It is only limited to cotton plants.

Kiran R Gavhale et al 2014, [14] developed a model to detect citrus leaf diseases. The accuracy obtained is 96%. K-means clustering is used for segmentation and SVM for classification. The plants can be protected from detrimental chemicals to ensure a healthier environment. Future study on test equipment, Image processing and analysis method can be done to achieve better detection accuracy with various plant species.

III. METHODOLOGY

1. Software

This project proposes an image pattern classification to identify different diseases in leaf with a combination of texture and colour feature extraction.

Steps:

1. The input test image is acquired and preprocessed in the next stage and then it is converted into array form for comparison.
2. The selected database is properly segregated and preprocessed and then renamed into proper folders.
3. The model is properly trained using CNN and then classification takes place.
4. The comparison of the test image and the trained model take place followed by the display of the result.
5. If there is a defect or disease in the plant the software displays the disease along with the remedy.

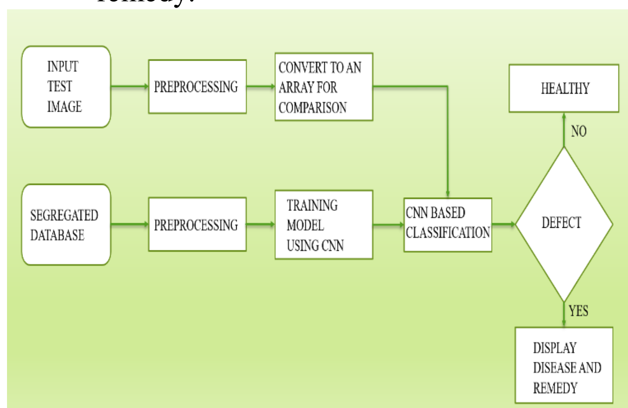

Figure 4: Software methodology

Image test image/Image acquisition: The sample pictures of the infected leaves are gathered and are utilized in training the system. To train and to test the system, unhealthy leaf pictures and few healthy images are taken. The images will be put away in some standard arrangement. The picture foundation ought to give a legitimate difference to the leaf shading. Leaf disease dataset is set up with both high contrast foundation, in light of the near examination dark foundation picture gives better outcomes and thus it is utilized for diseased leaf identification.

Image pre-processing: Picture obtained utilizing the advanced camera is pre-prepared utilizing the clamor expulsion with averaging channels, shading change and histogram equalization. The shading change step changes over the RGB picture to HSI (Hue, Saturation and Intensity) portrayal as this shading space depends on human discernment. Hue alludes to the predominant shading trait similarly as saw by a human onlooker. Saturation alludes to the measure of splendor or white light added to the tone. Intensity alludes to the apparent magnitude of light. After the RGB to HSI transformation, Hue's portion of the picture is considered for the examination as this gives just the necessary data. S and I part are overlooked as it doesn't give any critical data.

Masking green pixels: Since the vast majority of the green hued pixels implies to the healthy leaf and it doesn't increase the value of the illness detection procedures, the green pixels of the leaf are expelled by a specific masking technique, this technique altogether decreases the data processing time. The masking of green pixels is accomplished by processing the intensity estimation of the green pixels, if the intensity is not exactly a predefined limit esteem, RGB segment of that specific pixel is allocated with an estimation of zero. The green pixel masking is a discretionary advance in our disease recognition technique as the infected part of the leaf can be totally segregated in the segmentation procedure.

Disease Identification Using CNN: These days pretty much every tech-innovation firms depend on CNN for progressively productive execution. We should have an essential thought on how the human cerebrum perceives an object disregarding its differing characteristics from each other. Our cerebrum has a mind boggling layer of neurons ,each layer holds some data about the item and all the highlights of the article are extricated by the neurons and put away in our memory, next time when we see a similar object the cerebrum coordinates the stored highlights to perceive the object, yet one can without much of a stretch misstep it as a basic "IF-THEN" function, yes it is somewhat however it has an additional element that gives it an edge over different algorithms that is Self-Learning, in spite of the fact that it can't coordinate a human cerebrum yet at the same time it can give it tough competition. Image is processed utilizing the Basic CNN to recognize the infections in leaves. The information preparing in our CNN model needs to fulfill following requirements:

- There ought to be no missing values in our dataset.
- The dataset should be isolated into training and testing sets, either the training or the testing set shouldn't contain any immaterial information out of our model area. In case of an image dataset all the pictures must be of a same size, one uneven distribution of image size in our dataset can diminish the effectiveness of our neural system.
- The pictures ought to be changed over into black-white format before taking it into the convolution layer since perusing pictures in RGB would include a 3-D NumPy matrix which will decrease the execution time of our model by an impressive sum.
- Any sort of undermined or obscured pictures ought to be cut from the database before taking it into the neural system. Presently we have understood the pre-processing rules, let us make a plunge directly into the working of the convolution neural system.
-

2. Hardware

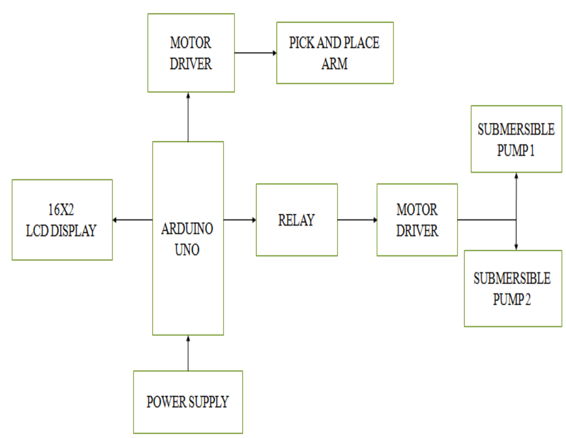

Fig 5: Hardware block diagram

A robot is designed to spray the pesticides and to pluck the weeds.

Arduino UNO is the heart of the robot. Once the disease is identified by the algorithm, a message is communicated to the robot. The algorithm tests for

two diseases of tomato plant. Bacterial Spot and Yellow Curl Virus.
Two tanks are used to store the pesticide. If the communicated message corresponds to the first disease then the first pump connected to UNO is driven by the motor and the pesticide is sprayed.
If the second type is detected then the second pump sprays the pesticide to the affected plant. If the leaf is tested to be healthy, no pesticide is sprayed. Weeds grow along with the crops and are harmful as they reduce the farm and productivity hence, they have to be plucked. If the weeds are detected by the algorithm, a message is communicated to the robot which in turn makes use of a pick and place arm to pluck the weed.

IV. CNN LAYERS

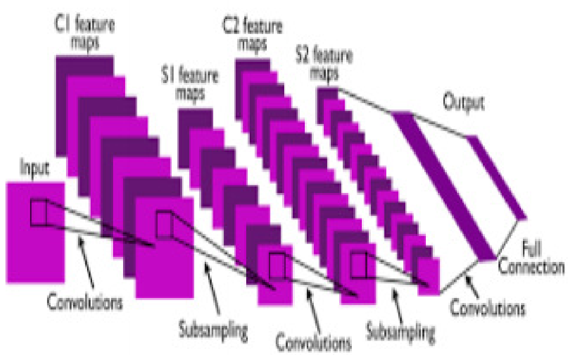

Figure 6: CNN layers

A. Convolution layer

This layer includes filtering the entire image for designs and articulating it as a 3x3 network. This convolved feature matrix of the picture is known as Kernel. Each kernel-unit is called weight vector.

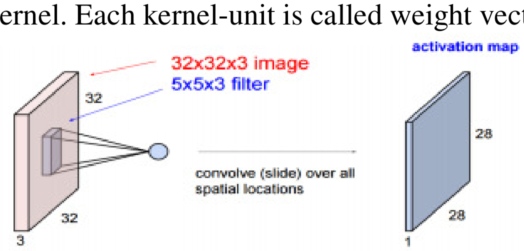

Figure 7: Convolutional layer

B. Pooling layer

After the convolution comes to the pooling here the picture grid is separated into the arrangements of 4 rectangular sections which are non-overlapping. There are two sorts of pooling, Max pooling and average pooling. Max pooling gives the most extreme incentive in the relative lattice area which is taken. Average pooling gives the average value in the relative lattice locale. The key advantage of the pooling layer is that it expands PC execution and diminishes over-fitting possibilities.

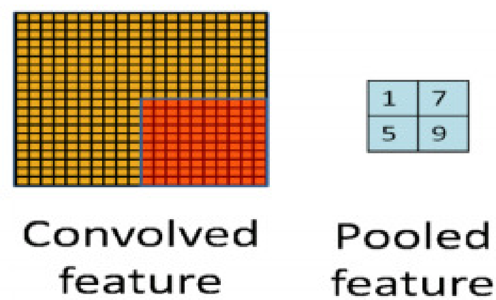

Figure 8: Pooling layer

C. Activation layer

It is the part of the CNN where the units are Normalized that is, they are fitted in a specific range. The utilized convolutional function is ReLU which permits just the positive units and afterward dismisses the negative units. It is the task of low computational expense.

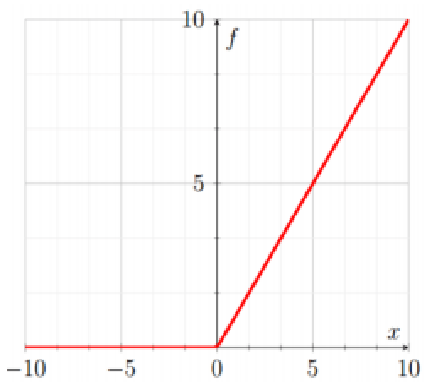

Figure 9: ReLU function

D. Fully connected layer

Here the highlights are compared with the highlights of the test picture and associate comparative highlights with the predetermined label. Generally, labels are encoded as numbers for the computational simplicity, they will be then changed over into their individual strings.

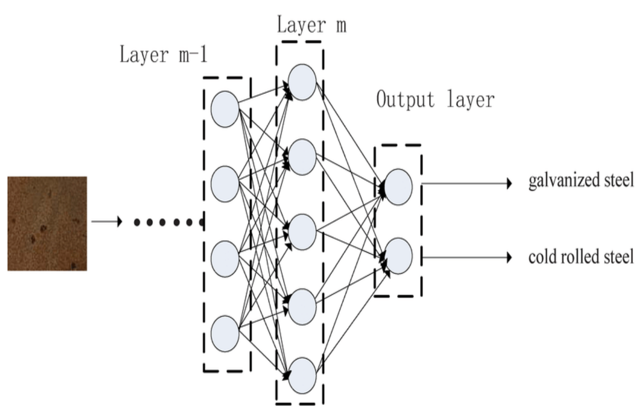

Figure 10: Fully connected layer

V. FLOWCHART AND CIRCUIT DIAGRAM

1. Software flowchart

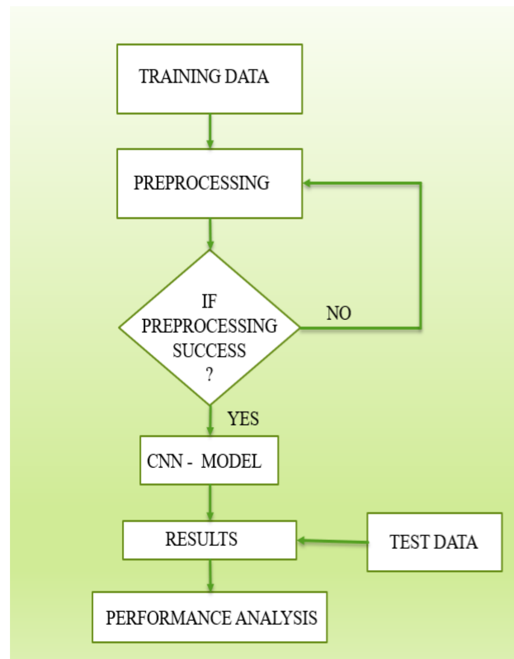

Figure 11: Software flowchart

Database collection: Initial step for any image processing-based project is acquiring proper database which is valid. Most of the time the standard database is preferred but in certain circumstances we do not get proper database. So, in such conditions we can collect the images and can form our own database. The database is accessed from CrowdAI which is plant disease classification challenge. Data available here is not labeled. So, the first task is to clean and label the database. There is a huge database so basically the images with better resolution and angle are selected. After selection of images we should have deep knowledge about the different leaves and the disease they have. Huge research is done from PlantVillage organization repository. Different types of plant images are studied and corresponding. After detail study, labeling in done by segregating the images and with different diseases.

Preprocessing and Training the model (CNN): The database is Preprocessed, for example, Image reshaping, resizing and change to a cluster structure. Comparative processing is additionally done on the test image. A database comprising of around 32000 distinctive plant species is acquired, out of which any picture can be utilized as a test picture for the software. The train database is utilized to train the model (CNN) so it can recognize the test picture and the disease it has. Convolutional Neural Network has various layers that are Dense, Dropout, Activation,Flatten, Convolution2D, MaxPooling2D. After the model is trained effectively, the software can recognize the infection if the plant species is contained in the database. After effective training and preprocessing, examination of the test picture and trained model happens to foresee the illness.

2. Hardware circuit connection

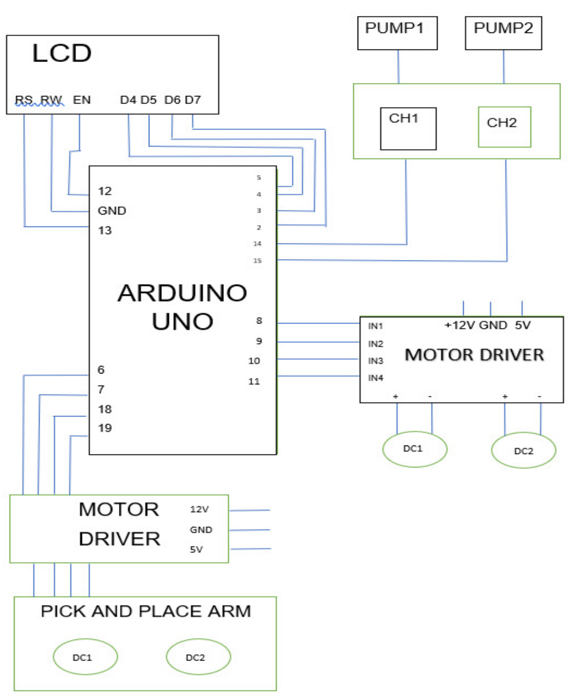

Figure 12: Hardware circuit diagram

1. The robot consists of Arduino UNO which is interfaced with 16X2 LCD display, motor drivers, relay and pick-place arm.
2. It runs on two 45rpm DC motor wheels and two supporting dummy wheels.
3. The submersible pumps are driven by a motor driver.

A relay is used to control the pumps.

1. The pick and place robotic arm is used to pluck weeds.
2. The robot is programmed to move forward, reverse, left, right and to stop.
3. The gears of the pick-place arm help it to move up and down and also support the jaws to open and close.

VI. REMOTE CONTROL OF ROBOT

The designed robot is controlled using an Android Application “TCP/UDP Test Tool”. The connection is established between robot and application by entering the IP address of the Wi-Fi module. Once the connection is established the robot can be controlled by giving the following commands:

$F=Forward
$L=Left
$K=Right
$B= Reverse
$G= Stop
$O= Pump 1
$N= Pump 2
$D= Disease

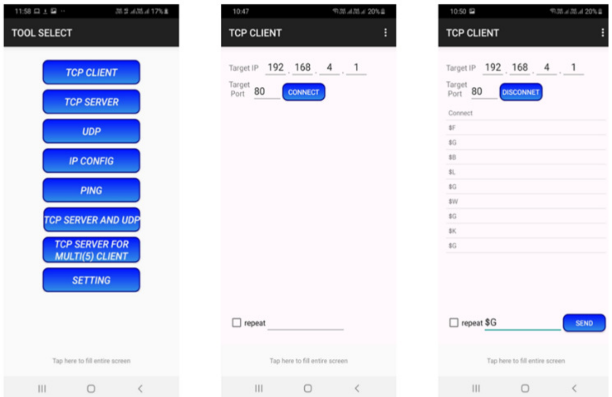

Figure 13: TCP/UDP Application

VII. RESULTS

We use Convolution Neural Network (CNN) which comprises of different layers which are used for prediction. With our code and training model we have achieved an accuracy level of 78%. Our software gives us the name of the plant disease with its confidence level and also the remedy that can be taken as a cure. A robot is designed and controlled using an android application. The robot sprays the suggested pesticide on the diseased leaf. It also plucks the weeds if found.

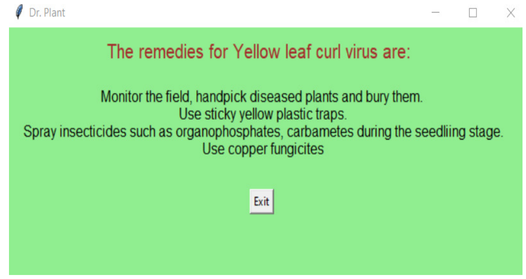

Figure 14: Detected leaf disease and its remedies

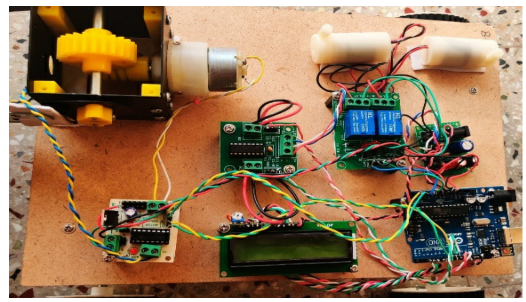

Figure 15: Components of robot

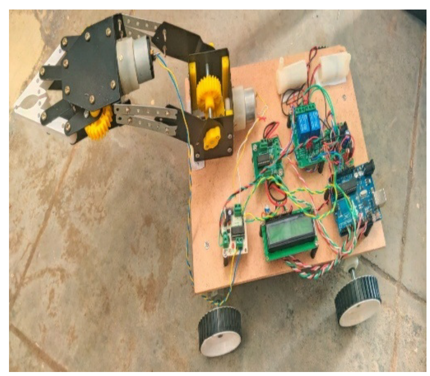

Figure 16: Working model

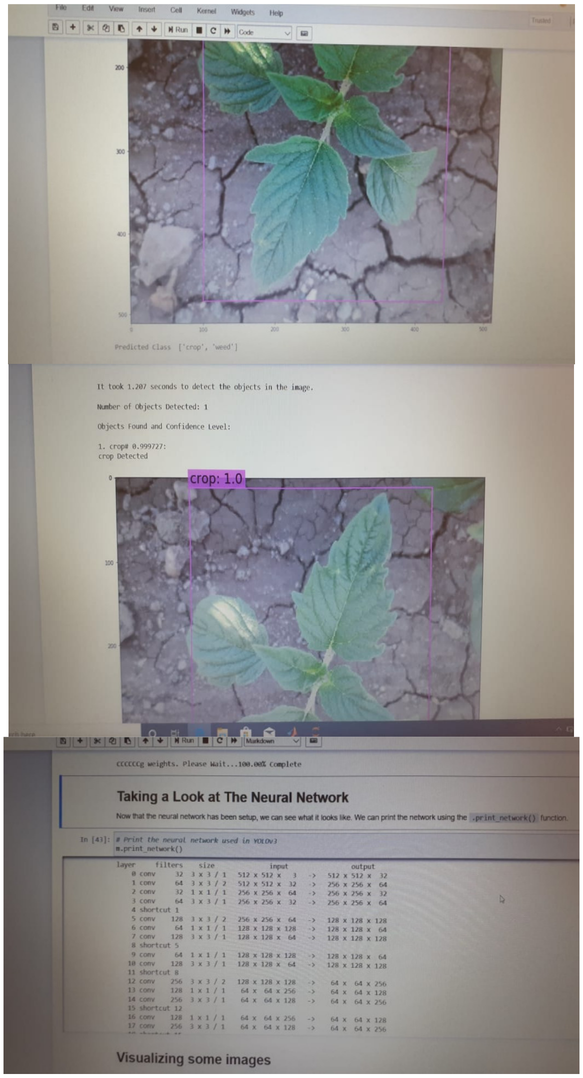

Figure 17: Weed detection result

VIII CONCLUSION AND FUTURE WORK

The developed system helps in recognition of plant disease and give remedies that can be utilized as a resistance system against the malady. The database acquired from the Internet is appropriately segregated and the diverse plant species are distinguished and are renamed to frame a legitimate database and then get test-database which comprises of different plant ailments. At this point, by utilizing training data, we will train our classifier and then yield will be anticipated with ideal precision. Our system gives us the name of the plant species with its certainty level and furthermore the cure that can be taken as a fix. A robot is designed which sprays the suggested pesticide on the leaf. It also plucks the weeds if found.

For real-time detection, a prototype drone model can be designed which can be used for live coverage of large agricultural fields to which a high resolution camera is attached and will capture images of the plants which will act as input for the software, based of which the software will tell us whether the plant is healthy or not.

REFERENCES

[1] “Identification of plant leaf diseases using image processing techniques” by Ms Pooja V, Rahul Das, Kanchana V at IEEE International Conference.

[2] “Leaf Disease Detection: Feature Extraction with K-means clustering and Classification with ANN” by Ch.Usha Kumari, S.Jeevan Prasad, H.Mounika at Proceedings of the Third International Conference on Computer Methodologies and Communication (ICCMC 2019) IEEE Xplore Part Number:CFP19K25-ART;ISBN:978-1-5386-7808-4.

[3] “Dectection of Plant Disease Using Threshold, K-Mean Cluster and ANN Algorithm” by Trimi Neha Tete, Sushma Kamlu at 2017 I2CT.

[4] “A Review: Custard Apple Leaf Parameter Analysis and Leaf Disease Detection using Digital Image Processing” by Appasaheb Gargade, Mr.S.A Khandekar at ICCMC 2019 IEEE Xplore.

[5] “Detection of Potato Diseases Using Image Segmentation and Multiclass Support Vector Machine” by Monzurul Islam, Anh Dinh, Khan Wahid, Pankaj Bhowmik at 2017 CCECE.

[6] “SVM Classifier Based Grape Leaf Disease Detection” by Ms Pranjali B. Padol, Prof. Anjali A. Yadav at 2016 Conference on Advances in Signal Processing (CASP).

[7] “Basic Study of Automated Diagnosis of Viral Plant Diseases Using Convolutional Neural Networks” by Yusuke Kawasaki, Hiroyuki Uga, Satoshi Kagiwada, Hitoshi Iyatomi at c Springer International Publishing Switzerland 2015.

[8] “A Review on Machine Learning Classification Techniques for Plant Disease Detection” by Mrs. Shruthi U, Dr. Nagaveni V, Dr. Raghavendra B K at 2019 ICACCS.

[9] “Recent Technologies of Leaf Disease Detection using Image Processing Approach –

A Review" by Bharat Mishra , Mamta Lambert, Sumit Nema, Swapnil Nema at 2017 ICIIECS.

[10] "A Survey Paper on Plant Disease Identification" by Jyoti Shirahatti, Jyoti Shirahatti, Pooja Akulwar at ICCES 2018 IEEE Xplore

[11] "Plant disease detection using Raspberry Pi by K-means clustering algorithm" by Ms Priyanka G Shinde, Mr Ajay K Shinde, Mr Ajinkya K Shinde, Mr Borate S.P at International Journal Of Electrical, Electronics and Computer Systems (2017).

[12] "Plant Monitoring using Image Processing, Raspberry Pi & IoT" by Ms Bhavana Patil, Hemant Panchal, Mr. Shubham Yadav , Arvind Singh, Dinesh Patil at IRJET.

[13] "Machine Learning Regression Technique for Cotton Leaf Disease Detection and Controlling using IoT" by Adhao Asmita Sarangdhar, Prof. Dr. V. R. Pawar at ICECA 2017.

[14] "Unhealthy region of citrus leaf detection using image processing technique" by Kiran R Gavhale, Prof.Ujwalla Gavwande, Kamal O.Hajari at International Conference for Convergence in Technology(2014).

[15] "Detection of diseases on plant leaf with the help of image processing" by Ms Prajakta S Garud, Ms Rajan Devi at International Journal of Engineering Technology Science and Research (2017).

[16] "Plant disease detection using image processing" by Mr Sachin D.Khirade, Mr A.B.Patil at International Conference on Computing Communication Control and Automation (2015).

[17] "Applying image processing technique to detect plant diseases" by Mr. Anand.H.Kulkarni, Mr Ashwin Patil R.K at International Journal of Modern Engineering Research (2012).

[18] "Crop disease classification using texture analysis" by Loyce Selwyn Pinto et al. at RTEICT 2016.