International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 07 Issue: 03 | May- 2023 Impact Factor: 7.185 ISSN: 2582-3930

Automated Detection of Grape Leaf Diseases using Machine Learning Techniques

Sarthak Khalkar [1], Sahil Patil[2], Mehul Mulchandani[3] Sakshi Wavdhane

Prof. Y.H Palve sir

[1]*Department of information technology Engineering of Karmveer baburao Thakare College of engineering*

---***---

Abstract – Crop production problems are common in India which severely effect rural farmers, agriculture sector and the country's economy as a whole. In Crops leaf plays an important role as it gives information about the quantity and quality of agriculture yield in advance depending upon the condition of leaf. In this paper we proposed the system which works on preprocessing, feature extraction of leaf images from plant village dataset followed by convolution neural network for classification of disease and recommending Pesticides using Tensor flow technology. The main two processes that we use in our system is android application with Java Web Services and Deep Learning. We have use Convolution Neural Network with different layers five, four & three to train our model and android application as a user interface with JWS for interaction between these systems. Our results show that the highest accuracy achieved for 5-layer model with 95.05% for 15 epochs and highest validation accuracy achieved is for 5- layer model with 89.67% for 20 epochs using tensor flow.

Key Words: ***CNN, Tensor flow, Leaf Disease, ANN.***

1. INTRODUCTION

Technology helps human beings in increasing the production of food. However the production of food can be affected by number of factor such as climatic change, diseases, soil fertility etc. Out of these, disease plays major role to affect the production of food. Agriculture plays an important role in Indian economy. Leaf spot diseases weaken trees and shrubs by interrupting photosynthesis, the process by which plants create energy that sustains growth and defense systems and influences survival[1]. Over 58% smallholder farmer depends on agriculture as their principal means of livelihood. In the developing world, more than 80 percent of the agricultural production is generated by smallholder farmers, and reports of yield loss of more than 50% due to pests and diseases are common[2].The production is decreasing day by day with various factors and one of them is diseases on plants which are not detected early stage. There is various work is done in previous years. Bacterial disease reduces plants growth very fastly so to detect this type of diseases ,Dheeb Al Bashish, Malik Braik, and Sulieman Bani-Ahmad [3] created system which detect the type of disease the plants have using image processing

And color space transformation which creates device independent transformation. Identifying the disease at an early stage and suggesting the solution so that maximum harm can be avoided to increase the crop yield [4] have used ANN and K-means to classify the disease and grade the disease for. There is a need to design the automatic system to detect the leaf disease and recommend the proper pesticide.

Pesticides on rice for controlling the disease damages the rice filed [5] created which will detect diseases at early stage. Which pesticides to use for which type of disease is the important task [6] gives the solution to which type of pesticides use. The paper is organized into following sections. Section 1 gives the introductory part and importance of automatic system design for early detection of leaf disease. Section 2 gives the current work done in this area. Third section include proposed methodology for leaf disease and recommending pesticide using CNN and Tensor flow tech. section 4 discuss the performance analysis and finally section 5th conclude the paper.

LITRARTURE SURVEY:

The system by Xingchun Chen and Ron Roeber [7] focuses on Plant diseases like fungal diseases which reduces crop production. The wetness in leaf, environmental and soil data gathered from different sources. The site is connected to High plain regional climate center (HPRCC) weather server

.The system is built in a zope web server with MYSQL relational databases support. Zope is as an open source web application server, is written in Python and built for content management systems. A model build by J. Duthie [8] and S. Pennypacker Et al. [9] is one of the most famous models for predicting infectiousness of disease. Both papers use a Weibull probability density function (PDF) and consider the effects of temperature and wetness duration. The proposed work is about automatic detection of diseases and diseased part present in the leaf images of plants of grape using SVM. This study contains a unique work that is it will calculate the% of infected area of plants [10]. Demonstration of the technical feasibility of a deep learning approach to enable automatic disease diagnosis through image recognition Classify Crop species & disease status of 38 different classes containing 14 crop species and 26 diseases achieving accuracy of over 99% using Deep Convolution Neural network AlexNet, GoogLeNet, Stochastic Gradian Descent[11]. The study compared the performance of PLSR, v-SVR, and GPR with the PRI and NBNDVI. The experiment was conducted in the greenhouse under controlled conditions to study the different disease symptoms effects on reflectance of the leaves for Wheat leaf rust disease [12]. CNN based on the AlexNet architecture is able to significantly outperform the baseline MLP, showing comparable performance to that of a group experts and outperforming any single expert. They have used Dataset of 2539 images of Apple Tree Species: Maxigala, Fuji Suprema effects on reflectance of the leaves for Wheat leaf rust disease [12]. CNN based on the AlexNet architecture is able to significantly outperform the baseline MLP, showing comparable performance to that of a group experts and outperforming any single expert. They have used Dataset of 2539 images of Apple Tree Species: Maxigala, Fuji Suprema and Pink Lady for Diseases: nutritional imbalances leaves with potassium and magnesium deficiency, damage (apple tree scab and Glomerella's stains), (391,558), herbicide damage (glyphosate) 325,569 healthy leaves)[13]. Robotic

detection system for combined detection of two major threats of greenhouse bell peppers: Powdery mildew (PM) and Tomato spotted wilt virus (TSWV).they have used the tech like Principal Component Analysis for PM, Principal Component Analysis for TSWV,Coefficient of Variation of TSWV Symptom Pattern the result shows For TSWV, PCA- based classification with leaf vein removal, achieved the highest classification accuracy (90%) while the accuracy of the CV methods was also high (85% and 87%). For PM, PCA-based pixel-level classification was high (95.2%) while leaf condition classification accuracy was low (64.3%) since it was determined based on the upper side of the leaf while disease symptoms start on its lower side [14]. In the paper

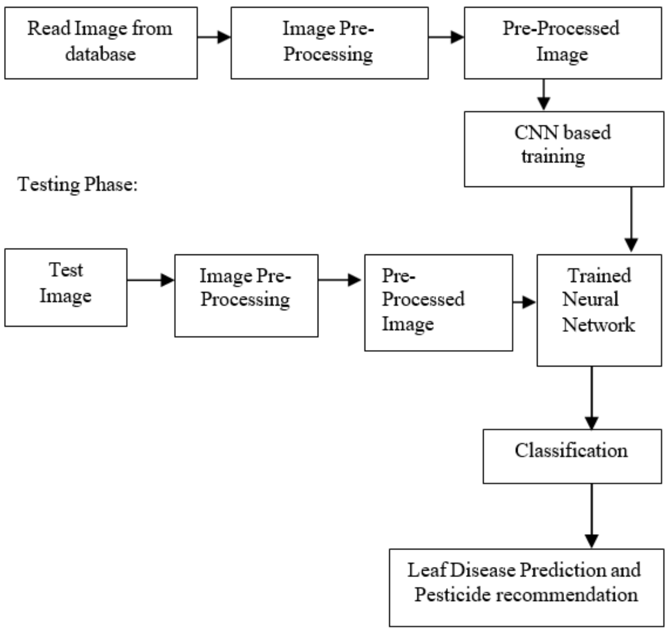
[15] AN FCM clustering and neural network classification based approach is proposed to detect and quantify the severity for late blight disease of potato.[16] A survey on methods that use digital image processing techniques like ANN,SVM,SOM,Fuzzy Logic,Thresholding, color Analysis. To detect, quantify and classify plant diseases from digital images in the visible spectrum. [17] Image analysis and classification techniques for extraction and classification of leaf diseases. Leaf image is captured and then processed to determine the status of each plant using GLCM texture feature and color textures features are extracted for further classification purpose. Finally classification based on SVM. Lucas G. Nachtigall, Ricardo M. Araujo and Gilmar R. Nachtigall [18] consider the apple plants for disease detection on the datasets of 2539 images and 6 known disorders using convolutional neural network. By detecting the different stages of plants disease, it helps to diagnoses and prevent further loss of yield and [19] proposes method which gives different stages of plants disease and for this they used spectral data. The spectral data is collected using ASD spectroradiometer (Analytical Spectral Device, Boulder, CO, USA).By using the wavelength the system is able to find out severity of disease for which they used different math's function. [20] Used a Flash FPGA and a DSP in plant disease remote monitoring and control system. The FPGA is used to acquire and transmit the field plant image or video data for subsequent monitoring and diagnosis, the DSP TMS320DM642 to process and encode video or image data to get high transfer efficiency, the nRF24L01 single chip 2.4GHz radio transceiver is used for wireless data transfer.

PROPOSED METHODOLOGY:

The main aim is to design a system which is efficient and which provide disease name and pesticides name as fast as possible. For that purpose we use two phase: 1st is training phase and 2nd is testing phase. In 1st phase: Image acquisition, Image Pre-processing and CNN based training. In 2nd phase Image acquisition, Image Pre-processing, Classification and disease identification and pesticides identification. For experimentation purpose we have used PlantVillage datasets. The data records contain 54,309 images. The images span 14 crop species: Apple, Blueberry, Cherry, Corn, Grape, Orange, Peach, Bell Pepper, Potato, Raspberry, Soybean, Squash, Strawberry, and Tomato. It contains images of 17 fungal diseases, 4 bacterial diseases, 2 mold (oomycete) diseases, 2 viral disease, and 1 disease caused by a mite. 12 crop species also have images of healthy leaves that are not visibly affected by a disease. The following figures show the overall model of the system and sample images from dataset.

Sample images of leaves from dataset

Training phase:



International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 07 Issue: 03 | May- 2023 Impact Factor: 7.185 ISSN: 2582-3930

System Architecture

A.Image Acquisition:

For training, Image is taken from database. And for testing, you can take image from camera at real time but in this project, we made a particular folder on mobile from that image will be fetched by android application and send through java web services i.e. tomcat server to server side system on which pre-processing is done and later on algorithm test that particular image.

B.Image Pre-Processing:

Image should be processed before sending to the algorithm for testing and training purpose. For that purpose, in this project image is scaled or resize into 150 x 150 dimension. As we used color image so that we don't need any color conversion techniques and that pre-processed image is directly passed to algorithm for training and testing purpose.

E.Convolutional Neural Network:

Once pre-processing is done, then CNN is used for training purpose and after that we get trained model. That CNN method is written with help of tensor flow. By using this model, we classify the image that system is getting after pre-processing of testing image.

Then we get particular disease name or healthy leaf name if there is no disease on that leaf and that disease name is send to android application and with the help of that disease name we get particular pesticide name which help farmer to take respective action in order to decrease percentage of disease.

PERFORMANCE ANALYSIS:

The performance analysis of convolution neural network for classification & prediction of pesticide for leaf disease is performed on PlantVillage datasets. The data records contain 54,309 images. We have used 38 class labels and 16 class labels in which class C00 has 270 images, Class C01 has 280 images etc. for experiment purpose. We have divided the dataset in training data which include 18917 images and testing data which include 3000 images. The results are shown for accuracy for training data and validation accuracy for testing data for different epoch with different layer of CNN i.e. five layer, four layer & three layer CNN and class labels i.e. 38 classes & 16 classes using tensor flow. Loss (cost) function used is cross entropy function and optimizer used is ADAM with learning rate 0.001 in Tensor flow for CNN Model.

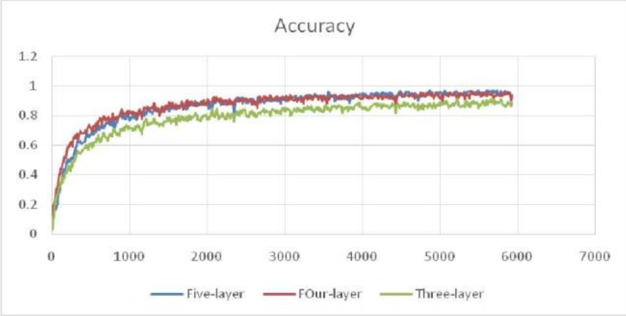


Fig. 3. Comparative graph of 5, 4, 3 CNN layer wrt Epoch

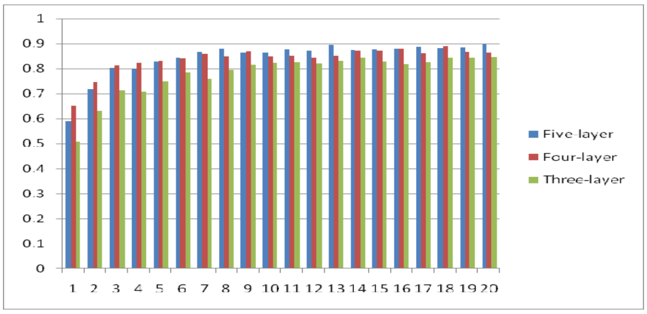


Fig. 4. Validation Accuracy Comparative graph of 5,4, 3 CNN layer wrt Epoch

Analysis of Graph: from the graph we can analyze that the validation accuracy gives best result for five layers with 89.67 as compared to three and four layer CNN.

Following graph shows the Accuracy and validation

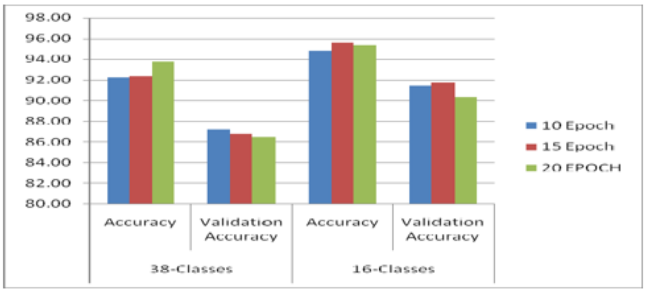


Fig. 5. Comparative graph ofAccuracy & Validation Accuracy for 38&16 Classes wrt Epoch

Analysis of Graph: from the graph we can analyze that the Accuracy for CNN for 38 classes is 93.72 for 20 epoch and 86.43 for validation and for16 classes accuracy is 95.32 & validation accuracy 93.80

TABLE III.OVERALL ACCURACY COMPARISON

	3-layers		4-layers		5-layers	
Total Epoch	*Accur Acy*	*Validat ion*	*Accur acy*	*Validat ion*	*Accur acy*	*Validat ion*
10	88.24%	83.23%	92.19%	87.17%	91.01%	86.93%
15	91.11%	82.97%	92.35%	86.77%	95.05%	86.30%
20	88.62%	84.77%	93.72%	86.43%	94.02%	89.67%

Following graph shows the overall Accuracy for 3 layer

CNN Model,4 Layer CNN model and 5 layer CNN model.

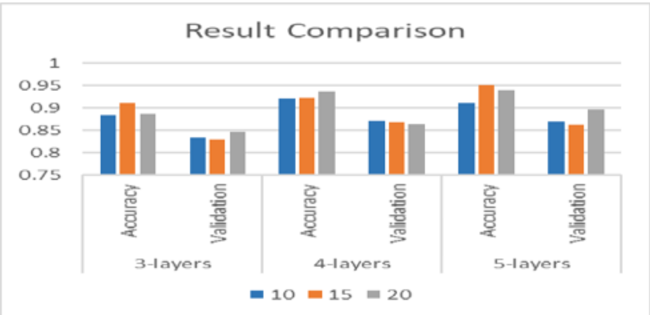


Fig. 6. Comparative analysis chart for overall Accuracy

Analysis of Graph: from the graph we can analyze that the highest Accuracy for CNN Model with 5 layers 95.05%, with 4 layers 93.72% and with 3 layers 81.11%. And highest validation accuracies for CNN model are 84.77%, 87.17% and 89.67% for 3,4 and 5-layer models respectively.

TABLE IV.SAMPLE OF PESTICIDE RECOOMENDATION WITH DISEASE

Disease name	Pesticide name
Apple scab	Spray liquid copper soap
grape leaf blight	Inspire super
orange huanglongbing	Zinkicide
cherry powdery mildew	Lime sulpher

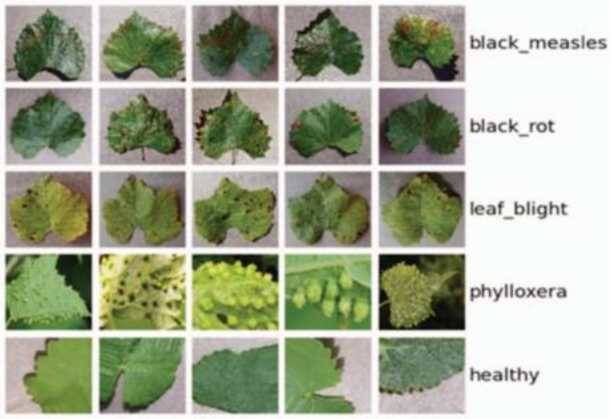


Image Data Examples

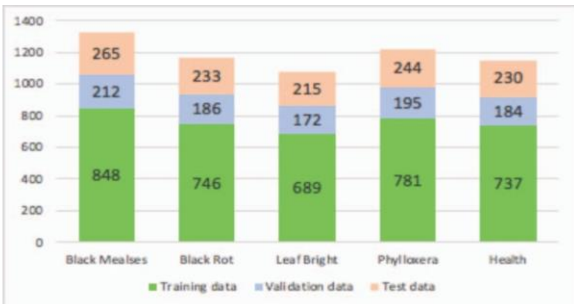
Traning Data Preparation and Validation

To prepare training data for detecting and classifying four grape diseases (leaf blight, black measles, black rot and phylloxera), images in the raw datasets are grouped into five classes, including healthy grape plant images. Figure 2 shows the detailed distributions. There are 5937 images with more than 1000 images in each class. These image data is randomlysplit in two parts: 20% of images are set as test data with number of 1187 images; the 80% of images are set as trainingand validation data. The ratio between training data to validation data is 80% to 20% ratio split.

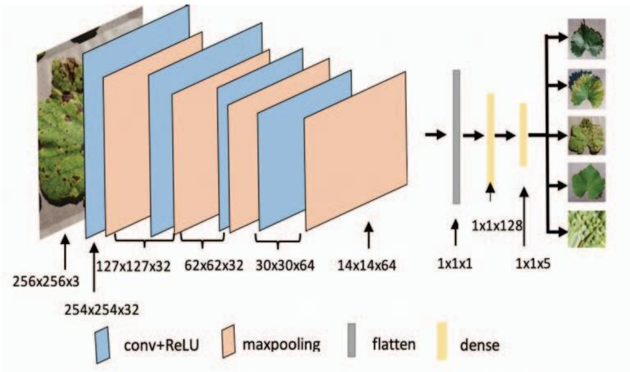
All of these images are placed into three folders. They are:training, validation and test. In each folder, there are five sub− folders named as black_mealses, , black_rot, leave_bright, and phylloxera, and healthy. Using Keras integrated data generation method, we are able to build training, validation and test datasets by reading files from the directory setting.

MACHINE LEARNING MODELS:

For grape disease detection and classification, we have developed four different types of deep learning models to apply our datasets so we can generate a comparative analysis and evaluation results to assess their accuracy andperformance.



Data Distribution



Vanilla CNN Architecture

Vanilla CNN:

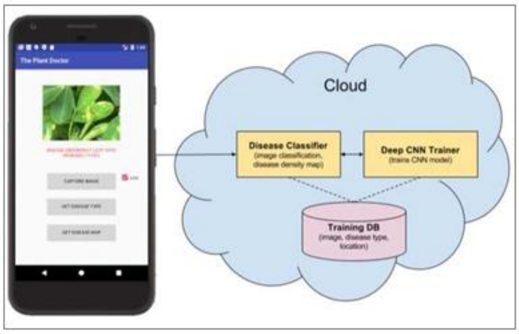
In this research project, we build a simple CNN model from scratch, known as Vanilla CNN, we set this model as basic model which can be used to compare with others.

This modelcontains 4 convolutional and max pooling layers and twodense layers. The model is modified and adjusted to support our 5 categories. The convolution layer's parameters are a set of learnable kernels with size of 3x3. Total parameters of this model are 1,671,973. Figure 3 shows its detailed architecture.

Improved VGG16 model with transfer learning:
VGG16 is a convolutional Network for Classification and Detection proposed by K. Simonyan and A. Zisserman from the University of Oxford. This model is pre−trained on the ImageNet dataset with 14 million images belonging to 1000 classes [14]. VGG16 was used in many deep learning image classification problems. The model showed 92.7% top−5 test accuracy in ImageNet. The structure of this model is very clear and easy to understand. It is easy to implement and can be a good start point for machine learning and classification problem. With the pre−trained model, we could get a very good performance baseline and do transfer learning by leveraging with the existing object feature extraction

The initial model has five convolutional layers. The input tothe first convolutional layer is fixed size 224 x 224 RGB image.The image is passed through a stack of convolutional (conv.) layers, where the filters were used with a very small receptivefield: 3×3 (which is the smallest size to capture the notion ofleft/right, up/down, center). The convolution stride is fixed to1 pixel; the spatial padding of conv. layer input is 1−pixel for 3×3 conv. layers. Spatial pooling is carried out by 5 max− pooling layers, which follow some of the conv. layers (not all the conv. layers are followed by max pooling). Max−pooling is performed over a 2×2 pixel window, with stride 2. Three Fully−Connected (FC) layers follow a stack of convolutional layers: the first two have 4096 channels each, the third performs1000−way ILSVRC classification and thus contains 1000 channels (one for each class). The final layer is the soft−max layer. The configuration of the fully connected layers is the same in all networks. All hidden layers are equipped with the rectification (ReLU) non−linearity. This improved model consists of 138M parameters and the configuration can be varied in depth. We have done the transfer learning and fine−tuning based on VGG16 using our program. The updated model architecture with transfer learning is displayed in Figure 4, in which the red rectangle is the modified part. The detailed transfer learning and fine−tuneare given below,

AN END-TO- END SOLUTION FOR CROP DIAGNOSIS:



System architecture with Cloud and Mobile components

Our proposed solution brings plant disease diagnostics to farmers through a Cloud based scalable collaborative platform.The platform is accessible through a mobile app that enables users to upload images of multiple parts of their plant and get the plant disease automatically diagnosed in real-time. They can also view "disease-density" map for their neighborhood showing geographical spread of diseases. The uploaded image gets classified by our AI engine into the appropriate category of disease for which a previously identified best-known- method solution is provided to the individual. Simultaneously, the geo-location of the image and a time stamp is used to tag the presence of the particular disease in that location. A collective density of diseases stored in a Cloud database is displayed on a map to show its location relative to the user. This allows the user to take preventive measures based on diseases in their neighborhood and serves as an alert for any spreading epidemic. The major components in the end-to-end system architecture of the proposed solution is shown in Fig. 1and the description of the components is provided below.

Mobile App:

- The mobile app contains a simplified frontend for the farmer that is easy to use and hides the complexity of the backend. . It enables the user to take images of the plant (*live* mode) or choose existing

images from the gallery (*offline* mode) and upload themto the Cloud backend for analysis. It allows them to getthe disease type of the uploaded images with a score reflecting the probability or accuracy of classification. It also enables the user to view a disease density map of the local area (if location service is enabled on the phone). Overall, the mobile app has 8 screens (sign-in with mobile number, main page with options, capture new image, load existing image, get disease type, get disease maps, history and expert connect). Android Studio 3.1.3 was used to develop the mobile app in Java with usage of

Google Camera API and Maps API. The mobile app communicates with the Cloud backend running on Amazon Web Services (AWS) over the cellular network using AWS Mobile SDK for Android.

Disease Classifier:

The Classifier is a standalone application running in the Cloud platform that receives the images uploaded via the mobile app and uses a trained deep Convolutional Neural Network (CNN) model to classify the disease type. The CNN model is computed by the Deep CNN Trainer and is used by the Classifier to automatically classify the uploaded images into the correct disease type. The Classifier also performs post-processing such as making a decision on whether the uploaded images should be added to the Training Database based on the classification score or sent to an agricultural expert registered on the platform for further analysis. When the classification score is greater than a preconfigured threshold, the images along with their metadata such as disease type and location ofthe images get added to the Training Database. In case of low classification score, the system forwards the caseand seeks assistance from agricultural expert teams formanual classification which are then sent to the farmer and stored in the Training Database. Low accuracy typically occurs if the user uploads an image with an underlying disease that is so far not known to the trained CNN model, or the image quality is poor Expert intervention in case of low classification score allows addition of new disease types which can bestored for future training runs. After the Training Database has sufficiently large number of images of thenew disease category and a high classification accuracy is achieved, the Classifier can start recognizing the new disease automatically. Over time as more farmers collaborate and contribute images, it enables us to improve the accuracy for automated response to covered diseases, while using the limited expertresources to expand coverage for new diseases.

Deep CNN Trainer:

This Cloud application is responsible for the more intensive work of training the neural network and builds the deep CNN model that is used by the Classifier to classify images into the correct disease types. This application is run asynchronously (without any interference to the Classifier) whenever the number of new images added to the Training Database goes beyond a pre-configured threshold. Eachsubsequent run of this training application works on a larger training dataset, and hence continually improves the deep CNN model used by the Classifier for more accurate disease classification.

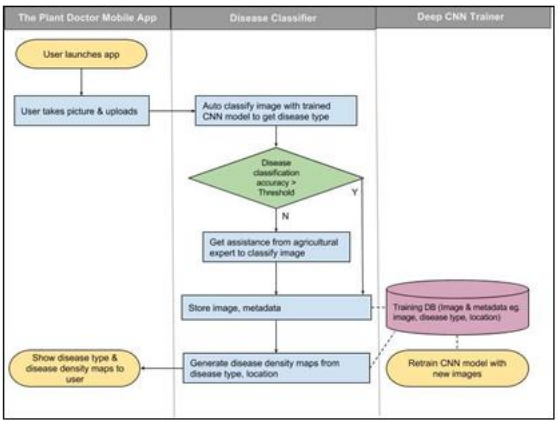
AWS was used to build the entire Cloud platform. The Disease Classifier and the Deep CNN Trainer are applications developed in Python. To make these Python applications accessible over mobile internet, they were developed using a web framework called FLASK and deployed behind an Apache Web Server running on an AWS EC2 machine (Ubuntu 16.04.2 LTS, 2 GiB memory, 8 GiB EBS volume). Disease Classifier and Deep CNN Trainer are built with TensorFlow [4], which is an open source library for Artificial Intelligence by Google.

Training Database:

This is a Cloud based database that stores all images used to train the deep CNN model. In addition to the images, it stores the metadata such as disease type, location of the images and time stamps. This database grows with wider use of themobile app and as farmers upload more images taken from their fields. Growth of the Training Database allows continual retraining of the deep CNN model withlarger datasets. Data in this database is also used to compute disease density relative to the user's location from collective metadata, such as disease types and image geo locations, and the generated disease density maps are rendered in the mobile app. AWS S3 was usedto implement the image database and MySQL running on AWS EC2 was used to store disease metadata such as classification, treatment and location.

Expert Interface:

A web based expert interface has been developed that allows agricultural experts to manually classify images that get low classificationscore. After the expert manually classifies the image, an SMS alert is sent to the user to check the mobile app history to receive the updated classification and remedial suggestions. Another feature of this interface is that it leverages the disease metadata stored by the Cloud platform to allow the experts to render time- based and geographical visualizations of disease data asshown in Fig. 2 for analytics and monitoring purposes. depicts the process flow with the sequence of steps performed by the constituent components of the platform as well as the interactions between them.



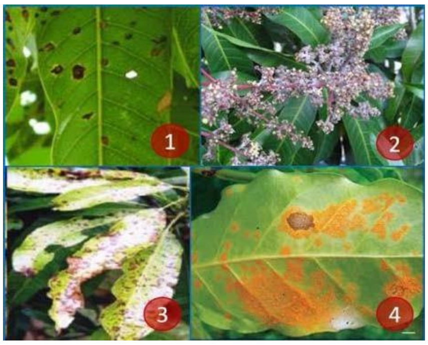
. Process flow of the components

EXPERIMENT, RESULTS AND OBSERVATIOS:

Multiple levels of experimentation were conducted to adequately simulate lab based and field based scenarios for image analysis, which forms the core aspect of this proposal. Experiments done can be broadly categorized into 3 types: Experiment 1 was conducted with training images retrieved from Google search to establish feasibility of the proposal; Experiment 2 was conducted with a large open source public dataset with images taken under controlled conditions to provethat the proposal has a high degree of accuracy even with many disease categories; and Experiments 3 and 4 were conducted with self-collected, high fidelity, high quality images from an agricultural farm to simulate real life with images of common crops such as groundnuts, tomatoes and grapes taken under natural conditions. Seasonality of crops, easy access to them, severity of diseases and their prevalence at the time of experimentation were factors in our choice of crops and our decision to perform aggressive data collection during the season.

A. **Experiment** 1:

As a first level of experimentation, we trained a deep CNN model to build a Disease Classifier using the latest Inception- v3 architecture [10] and Python using the TensorFlow framework [4]. The goal of this experiment was to perform image recognition of different diseases of mangoes as a baseline experiment to demonstrate feasibility of the approach before progressing to a wider data set or to real data collected from fields. Images of mangoes with diseases were bulk downloaded using Google Search. Out of many mango diseases [11], four common diseases were included in this experiment as they had distinct symptoms and their images were more easily available. These diseases were BacterialCanker, Mildew Mango, Phoma Blight and Red Rust



Symptoms of mango diseases (source: Google)

The downloaded images were used as the training data set and the Inception based CNN model was trained (transfer learning) on this data set to teach the network to recognize the four types of mango diseases Our model converts input images to 299X299 RGB for training and classification. Table 1 depicts the details of training and testing of image classification under Experiment 1.

67 images of bacterial canker, 70 images of mildew, 22 images of phoma blight and 37 images of red rust were used asa training data set for the CNN in Experiments1.1 to 1.3. Different number of training steps were used (250, 500, 1000) to evaluate if test accuracy changes with the training steps. Thetrained model was then used to classify test images. The output of classification for each test image is an array of 4 probabilityscores that sum to 1. Each score contains the probability that the test image belongs to one of the four categories.

Overfitting is a potential problem with neural networks where a model may just be memorizing irrelevant details of thetraining images to come up with the right answers, and it may give good results on the images it's seen during training but fails on new images. To avoid the problem of overfitting, someof images were removed from the training set so that the modelcan't memorize them and the process of training and testing were executed and tabulated under Experiments 1.4 and 1.5 in Table 1, keeping the number of training steps constant at 500.Table 1 demonstrates that even with a very low number of low fidelity images used for training, the classification

accuracy is satisfactory. For e.g., Experiment 1.1 shows that abacterial canker test image was classified in the right category with a probability score of 0.749. Keeping the number oftraining images constant, training is repeated by varying the number of training steps in 1.1, 1.2 and 1.3. Results show that as the number of training steps are increased from 250 to 500 to 1000, the score of classification improves (e.g. the scores ofclassification of the same bacterial canker image increases from 0.749 to 0.863 to 0.927 with number of training steps 250, 500 and 1000 respectively).

When the test images were removed from training data, the accuracy of classification of images that the trained model had not seen before was still fairly high as demonstrated by Table 1 Experiment 1.4. Classification of red rust image was incorrect (the model classified the red rust image as bacterial canker with a score of 0.379, marked in red). The experiment was rerun (Experiment 1.5) after swapping the test image of red rust with a better image from the training set (with clearer spots), and the classification score improves, with the model classifying the image as red rust with a score of 0.827. This proves that the accuracy of classification improved with the quality of test data. Training neural networks can be computationally and time intensive (10-60 minutes for our runs), however, the trained CNN models can be used to classify images very quickly (1-3 seconds) which makes the application of neural networks in smartphone apps possible. Results captured in Experiment 1 prove that CNNs can be adopted for image classification in our use case of plant disease diagnosis and motivates further experimentation with wider and high fidelity test data.

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 07 Issue: 03 | May- 2023 Impact Factor: 7.185 ISSN: 2582-3930

B. Experiment 2:

PlantVillage Sample Images

Test2 that failed to classify correctly

In the second level of experimentation, a large public dataset was used that includes images of diseased and healthy plants collected under controlled conditions by agricultural experts. This was to prove the applicability of the solution on a larger scale with more number of disease categories. PlantVillage is an open-source platform [5] for crop health and has released a public dataset of over 50,000 plant images to enable development of computer vision approaches to help solve the problem of loss of crop yields due to infectious diseases. This dataset includes curated images on healthy and infected leaves of crops. It has images of 26 diseases in 14 crops, leading to 38 possible crop-disease pairs (classes/categories labelled as c0 to c37). Fig. 5 is a collage created from samples of images taken from different categories of PlantVillage dataset to give an idea on the type of images used for experimentation.

For our experiment, 8 categories of PlantVillage images were randomly chosen for training and testing. From each category, 5 images were removed from the training set to serve as the test data. Remainder training data set from the 8 categories was used to train our Inception based CNN model, followed by the classification of test images by the trained model. Table 2 shows the statistics for training data set that was used to produce the trained CNN model.

Classification is correct for 37 out of 40 images, hence 92.5% images were classified correctly, proving that the solution will work even with a large dataset with more disease categories. 3 incorrect cases are marked in red in the table.The incorrect classifications are potentially due to reasons such as few categories being visually very similar (e.g. c21 and c24) and also poor quality of the test image as shown in Fig. 6 due to which the classifier fails to identify it correctly.

Experiment 3:

Groundnut was chosen as the main case study for field research and experiment in order to verify end user experience.The goal of this experiment was to simulate real life scenario with images taken in the field by users under natural conditions. Groundnut, also known as peanut, is widely grown and consumed all over the world and has significant economic importance being a rich source of edible oil and protein. We chose groundnut as a case study for field work due to the fact that 80% of the world groundnut crop is produced in developing countries where yields are usually very low and diseases have become a major obstacle to the groundnut output throughout the world [12]. China, India and US are the top three producers of groundnut globally. Although several diseases inflict groundnut crops [12], for purposes of this experiment, two major diseases of groundnut i.e. leaf spot or tikka’ and bud necrosis, were selected for collection and analysis of field data due to their severity, large scale impact onproduction and widespread occurrence in India.

FUTURE WORK AND EXTENSIONS:

Future work involves expanding the model to include more parameters which can improve the correlation to the disease. We can augment the image database with supporting inputs from the farmer on soil, past fertilizer and pesticide treatment along with publicly available environmental factors such as temperature, humidity and rainfall to improve our model accuracy and enable disease forecasting. We also wish to increase the number of crop diseases covered and reduce the need for expert intervention except for new types of diseases. For automatic acceptance of user uploaded images into the Training Database for better classification accuracy and least

Possible human intervention, a simple technique of computing the threshold based on a mean of all classification scores can be used.

Further application of this work could be to support automated time-based monitoring of the disease density maps that can be used to track the progress of a disease and trigger alarms. Predictive analytics can be used to send alerts to the users on the possibility of disease outbreaks near their location.

ACKNOWLEDGMENT:

Our sincere thanks to Dr. Pranav K. Singh, Assistant Professor, GADVASU, Ludhiana and Dr. M. K. Prasanna Kumar, Associate Professor of Plant Pathology, University of Agricultural Sciences, GKVK, Bangalore for consultation and providing access to the farms at both the universities for data collection. Guidance by Mr. Ashutosh Utpal in Android programming is also greatly appreciated.

CONCLUSION

In this work the leaf disease images are classified using convolution neural network with five-layer model, four-layer model and three-layer model and recommend the Pesticides as per the leaf disease. Our result shows that CNN model having 3-convolution layer, 4-convolution layer and 5-convolution layer trained for 10, 15 and 20 epochs (cycles). As we can see from graphs that the highest accuracy achieved for 5-layer model with 95.05% for 15 epochs and highest validation accuracy achieved is for 5-layer model with 89.67% for 20 epochs using tensor flow.

REFERENCES:

[1] Mangement of leaf spot diseasea of trees and shrubs:Trees & Shrubs:Yard & Garden:Garden:University of Minnesota Extension,4/22/2007.

[2] Monzurul Islam, Anh Dinh, Khan Wahid ,Pankaj Bhowmik "Detection of Potato Diseases Using Image Segmentation and Multiclass Support Vector Machine", IEEE 30th Canadian Conference on Electrical and Computer Engineering (CCECE),2017

[3] Dheeb Al Bashish, Malik Braik, and Sulieman Bani-Ahmad,"A framework for detection and classification of plant and stem diseases", International Conference on Signal and Image Processing,2010.

[4] Aakansha Rastogi,Ritika Arora,Shanu Sharma,"Leaf Disease Detection and garding using computer vision Technology & fuzzy logic",IEEE-IC on Signal processing and Integrated network,2015

[5] Santanu Phadik ar and Jaya Sil, "Rice Disease Identification using Pattern Recognition", Proceedings of 11th International Conference on Computer and Information Technology,2008

[6] Neethu K.S1 , P .Vijay ganesh2 , "Leaf Disease Detection and Selection of Pesticides using Artificial Neural Network", International Research Journal of Engineering and Technology (IRJET),2017

[7] Xingchun Chen and Ron, "Monitoring Soybean Disease and Insect Infection Patterns in Nebraska", International Conference on Granular Computing, IEEE 2015

[8] J. A. Duthie , "Models of the Response of Foliar Parasites to the Combined Effects of Temperature and Duration of Wetness", Phytopathology, 87–11, 1088/1095 (1997)

[9] S.P. Pennypacker, H.D. Knoble, C.E. Antle, and L.V. Madden: "A Flexible Model for Studying Plant Disease Progression, Phytopathology", 70–3, 232/235 (1980)

[10] Rajleen Kaur,Dr. Sandeep Singh Kang,"An Enhancement in Classifier Support Vector Machine to Improve Plant Disease Detection",IEEE 3rd International Conference on MOOCs, Innovation and Technology in Education (MITE), 2015

[11] Sharada Prasanna Mohanty,David Hughes and Marcel Salathé,"Using Deep Learning for Image-Based Plant Disease Detection,"April 15, 2016

[12] Davoud Ashourloo, Hossein Aghighi, Ali Akbar Matkan, Mohammad Reza Mobasheri, and Amir Moeini Rad,"An Investigation Into Machine Learning Regression Techniques for the Leaf Rust Disease Detection Using Hyperspectral Measurement,"IEEE JOURNAL OF SELECTED TOPICS IN APPLIED EARTH OBSERVATIONS AND REMOTE SENSING, VOL. 9, NO. 9, SEPTEMBER 2016

[13] Vijai Singh A.K. Misra ,"Detection of plant leaf diseases using image segmentation and soft computing techniques",INFORMATION PROCESSING IN AGRICULTURE 4 (2017) 41–49,Elsevier.

[14] Noa Schor, Avital Bechar,Timea Ignat, Aviv Dombrovsky, Yigal Elad,and Sigal Berman,"Robotic Disease Detection in Greenhouses: Combined Detection of Powdery Mildew and Tomato Spotted Wilt Virus",IEEE ROBOTICS AND AUTOMATION LETTERS, VOL. 1, NO. 1, JANUARY 2016

[15] Sandika Biswas , Bhushan Jagyasi , Bir Pal Singhy and Mehi Lalz,"Severity Identification of Potato Late Blight Disease from Crop Images Captured under Uncontrolled Environment",IEEE Canada International Humanitarian Technology Conference - (IHTC) 2014

[16] Jayme Garcia Arnal Barbedo,"Digital image processing techniques for detecting, quantifying and classifying plant diseases,Sprinjer plus,2013

[17] Ms. Kiran R. Gavhale,Prof. Ujwalla Gawande,Mr. Kamal O. Hajari,"Unhealthy Region of Citrus Leaf Detection Using Image Processing Techniques",International Conference for Convergence of Technology - 2014,IEEE