IJISET - International Journal of Innovative Science, Engineering & Technology, Vol. 8 Issue 2, February 2021
ISSN (Online) 2348 – 7968 | Impact Factor (2020) – 6.72
www.ijiset.com

Plant Leaf Syndrome Revealing and Grouping using ML

Dr.M. Rajesh Khanna [1], Dr. C.Thirumalaiselvan[2], D. Revathi[3] , B.G.Obula reddy[4], S.Anitha[5]

email: rajeshkhanna@veltechmultitech.org,thirumalaiselvan027@outlook.com, bgobulareddy2007@gmail.com, srirangam.anitha@gmail.com

***ABSTRACT*–The project presents leaf characteristics analysis victimization image process techniques for automated vision systems used in the agricultural field. In agriculture analysis of automatic leaf Characteristics detection is important one in watching large fields of crops, and some mechanically detects symptoms of leaf characteristics as shortly as they seem on plant leaves. The planned higher cognitive process system utilizes image content characterization and supervised classifier style of neural network. Image process techniques for this sort of call analysis involves preprocessing, feature extraction and classification stage. At process, associate input images are going to be resized and the region of interest chosen if required. Here, color associated texture options square measure extracted from an input for network coaching and classification .Color options like mean, variance of HSV Color area and texture options like energy, contrast, homogeneity and correlation. The system is going to be accustomed to classify the check pictures automatically to decide leaf characteristics. For this approach, automatic classifier CNN be used for classification supported learning with some coaching samples of that class. This network uses tangent sigmoid operate as kernel operate. Finally, the simulated result shows that the used network classifier provides minimum error throughout coaching and higher accuracy in classification.**

KEYWORD: OpenCV, Python, Numpy, HSV, GLCM

1. INTRODUCTION

Agricultural productivity are some things on that economy extremely depends this can be one in every of the explanations that malady detection in plants plays a crucial role in agriculture field and to extend the productivity, as having malady in plants square measure quite natural. To avoid those reasonably issues in agricultural field early detection of plant disease are going to be useful. We've got taken the project plan supported the agricultural field. In agriculture analysis of automatic plant disease detection is crucial one in observation giant fields of crops, and so mechanically detects symptoms of malady as presently as they seem on plant leaves. To identify the malady within the leaves supported coaching and classification and to send word the farmers so early actions is taken.

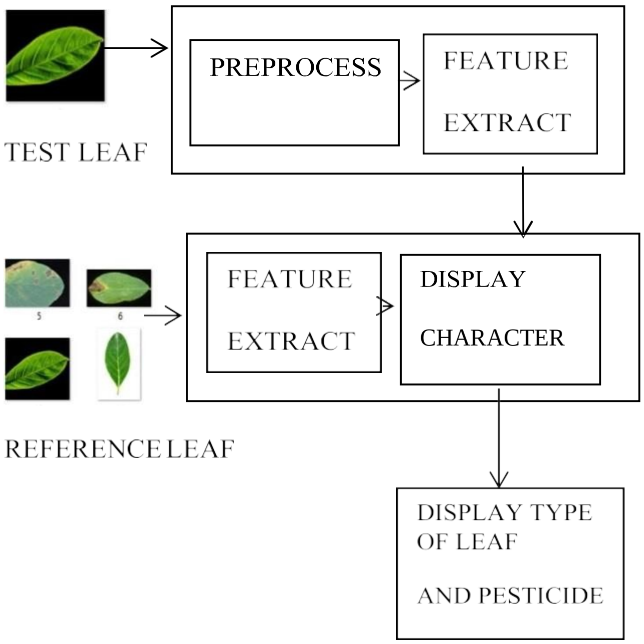
It additionally establish the sort of malady and recommend appropriate pesticides to be additional. Groundnut crop is sometimes cultivated throughout kharif season (May-June to September-October). The plant is vulnerable to attack of the many diseases that squaremeasure for the most part arises throughout summer season. Thanksto low downfall and sudden climate changes. Typically groundnuts square measure cultivated in large-scales by the farmersforthemercantilismandwholesalepurpose.Thusa lot of men and investments square measure needed for the palmy farming of the groundnut. The plant is extremely thought of for its high nutritional values and supermolecule richness of peanuts. The normal approach for observation the sphere includes manual mensuration throughout the huge field. For the higher than reasons, as With the use of image process and machine learning techniques for coming up with a system which will differentiate between traditional and abnormal leaves from the input pictures of the groundnut leaves captured from the sector. The system consists of pre-trained neural network, which collects and learns the essential options from a collection of coaching pictures.

Images square measure loaded for the coaching and classification solely when undergoing a series of image process steps. Each the coaching pictures and testing pictures were processed for the feature extraction and classification. The strategies in our project square measure mathematician Blurring, Bilateral Filter for Edge detection, Gray Conversion, GLCM options Extraction, KMEANS cluster, Thresholding for segmentation, Training and Classification. Supported the on top of strategies the output square measure obtained in an efficientmethod.

2. BACKGROUND

Literature review had been done to understand the concepts and to find the steps in that. In one of the paper they have taken apple leaf. In that disease detection using improved CNN and data augmentation and image annotation technologies. A disease dataset is created and used for training the network. It involves real-time disease detection bycollectingthetestimagefromthefieldandtrainingimages from concerned laboratories. The deep-learning-based approach can automatically extract the features of the diseased apple images and detect the five common types of apple leaf diseases with high accuracy in real time [5]. The next paper considered the wheat leaf. Wheat leaf diseases classification based on convolutional kernel matrix, for fine- grained image classification. This approach yields solid evidence that convolutional kernel matrix is a feasible and useful idea in general, which provides a new path for the identification of crop diseases[11].Likewise they have also considered grape leaf, maize leaf, mango leaf, strawberry leaf, potato leaf and cotton leaf. Likewise they have also taken algorithms like convolutional neural network, k-nearest neighbour, support vector machine etc. There are so many features to be extracted to get the needed area or the area of interest in the particular leaf. Some of the features are homogeneity, contrast etc. And these features have separate formula for calculation purpose. Thus from the above surveys we came to a conclusion that we use convolutional neural network and we have taken features likehomogeneity, contrast, correlation coefficient, energy, entropy in our system.

3. ARCHITECTURE:



3.1 DIAGRAM:

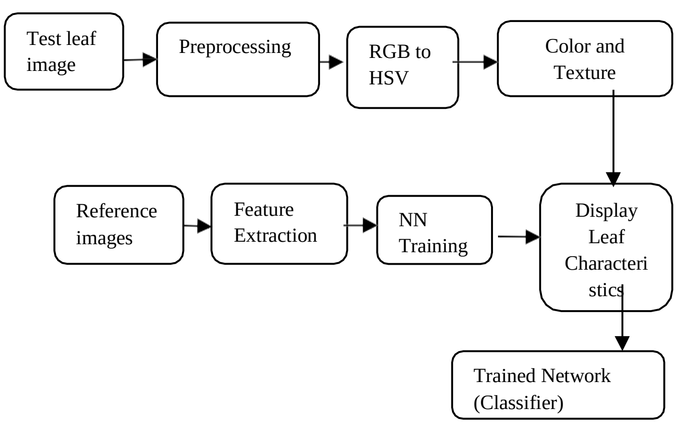


Fig2.1 block diagram

4. METHOD ANDMETHODOLOGY

4.1METHOD:

The method used in our system are Gaussian filtering, Bilateral Filter for Edge detection, Gray Conversion, GLCM Features Extraction, KMEANS clustering, Thresholding for segmentation and Training and Classification. These methods are applied based on the formula for each and every module. The methods will produce a matrix and based on the matrix values the identification and classification of leaf aredone.

4.1.1. GAUSSIAN BLURRING

In this technique, rather than a box filter, a Gaussian kernelisemployed.Weoughttospecifythebreadthandheight of the kernel that should be positive and odd. We conjointly ought to specify the quality deviation within the X and Y directions, sigmaX and sigmaY severally. If solely sigmaX is nominative, sigmaY is taken becausethe same as sigmaX.If each area unit given as zeros, they're calculated from the kernel size. Gaussian blurring is extremely effective in removing Gaussian noise from a picture.

Imageblurringisachievedbyconvolvingtheimagewitha low-pass filter kernel. it's helpful for removing noise. It truly removes high frequency content (eg: noise, edges) from the image. Therefore edges are blurred a touch bit during this operation (there also are blurring techniques that do not blur the edges). OpenCV provides four main varieties of blurring techniques. Among four blurring techniques we tend to getting to use Gaussian blurring technique. In this technique, rather than a box filter, a Gaussian kernel is employed. We ought to specify the dimension and height of the kernel that should be positive and odd. We tend to additionally ought to specify the quality deviation within the X and Y directions, sigmaX and sigmaY severally. If solely sigmaX is given, sigmaY is taken because the same as sigmaX. If each are given as zeros, they're calculated from the kernel size. Gaussian blurring is very effective in removing Gaussian noisefrom apicture.

FORMULA:

$$()= \frac{1}{\sqrt{2^2}}\,{}_{2a^2}\,\frac{x^2}{}$$

Implementation code:

Loading image

os.chdir('E:/project/leaf part/dataset');

img = cv2.imread('17.jpg');

os.chdir('E:/project/leaf part');

cv2.imshow('Input Image',img)

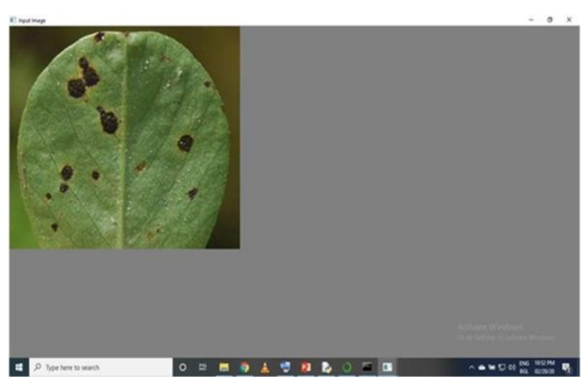


fig 4.1.1. Input image

The above image is given as input to the system and the process of loading is done. The path of the image to be taken are specified in the code. The image are blurred slightly according to the specification.

```
##Gaussian Blurring
kernel = np.ones((7,7),np.float32)/25
img1 = cv2.filter2D(img,-1,kernel)
cv2.imshow('Gaussian Image',img1)
```

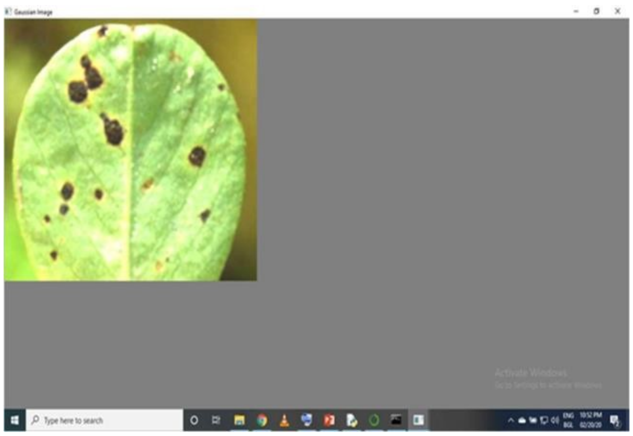


fig 4.1.2. Gaussian output

4.1.2. BILATERALFILTERING

Bilateral Filter is employed for Edge detection. Bilateral Filter is extremely effective in noise removal whereas keeping edges sharp. However the operation is slower compared to alternative filters. Gaussian filter takes the neighborhood round the pixel and finds its Gaussian weighted average. This Gaussian filter may be aperformance of house alone, that is, near pixels are thought-about whereas filtering. It does not contemplate whether or not pixels have virtually constant intensity. It does not contemplate whether or not a pixel is a position pixel or not. Thus it blurs the sides conjointly, that we do not wish to try and do. Bilateral filtering conjointly takes a Gaussian filter in house, however an additional Gaussian filter that may be a performance of pixel distinction. The Gaussian perform of housemakes positive that solely near pixels arethought-aboutforblurring, whereas the Gaussian perform of intensitydistinctionmakes positivethatsolelythosepixelswithsimilarintensitiestothe central pixel are thought-about for blurring. Thus itpreserves the sides since pixels at edges can have massive intensity variation.

FORMULA

$$w(i,j,k,l)=\exp\left(-\frac{(i-k)^2+(j-l)^2}{2\sigma_d^2}-\frac{\|I(i,j)-I(k,l)\|^2}{2\sigma_r^2}\right)$$

Implementation code:

```
## Bilateral Filter for Edge Enhancement
img3 = cv2.bilateralFilter(img1,9,75,75)
cv2.imshow('Bilateral Filtered Image',img3)
```

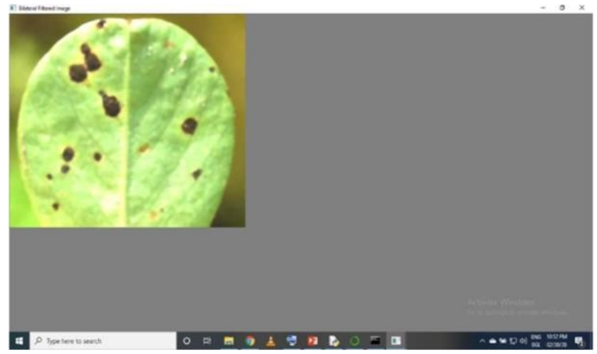


Fig 4.1.2 Bilateral image

The output of the block of code is mentioned above which is a bilateral image.

4.1.3. GRAYCONVERSION

The next method done is Gray conversion. Grayscale is a range of monochromatic shades from black to white. Therefore, a grayscale image contains only shades of gray and no colour. Colour increases the complexity of the model. The inherent complexity of gray level images is lower than that of colourimages,In thatcase,workingdirectlywithagrayimage is going to give you an advantage.

Implementation code:

```
## RGB to Gray conversion
GRAY_Img = cv2.cvtColor(img3,cv2.COLOR_BGR2GRAY)
cv2.imshow('GRAY Image',GRAY_Img)
Data2Ext=GRAY_Img;
cv2.imwrite('ImageRedist.jpg',Data2Ext);
```

FORMULA

$$C_{linear}=\begin{cases}\dfrac{C_{srgb}}{12.92}, & C_{srgb}\le 0.04045\\ \left(\dfrac{C_{srgb}+0.055}{1.055}\right)^{2.4}, & \text{otherwise}\end{cases}$$

Where C_{srgb} represents any of the three gamma-compressed srgbprimaries(C_{srgb}, C_{srgb}, C_{srgb})

$$C_{srgb}=\begin{cases}12.92\,C_{linear}, & C_{linear}\le 0.0031308\\ 1.055\,C_{linear}^{1/2.4}-0.055, & \text{otherwise}\end{cases}$$

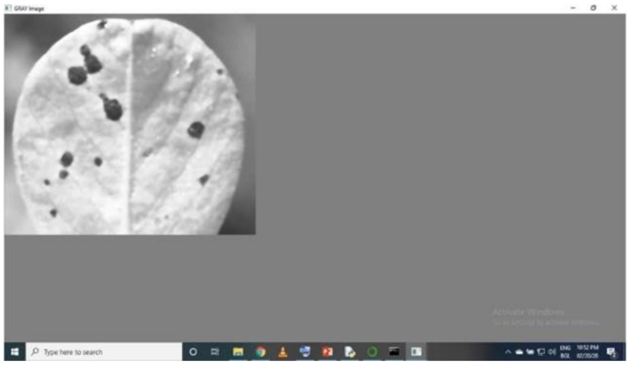


Fig 4.1.3. Gray image

The image which is mentioned above is of only two colors. The colors in the image are black and white. Due to this conversion in the image the disease affected area in the leaf can be easily spotted. The above image shows the disease in the leaf easily.

4.1.4. TEXTUREFEATURE

Energy : It's a livethe uniformness of the image and might be calculated from the normalized COM. it's an acceptable live for detection of disorder in textureimage.

FORMULA

$$=\sum_{i=1}\sum_{j=1}((,))^2$$

Entropy : Entropy offers a live of complexness of the image. advanced textures tend to possess higherentropy

FORMULA

$$= -\Sigma\Sigma(,)((,))$$

$$= =$$

Where,

p(i , j) is the co occurrence matrix

Contrast: Measures the local variations and textureof shadow depth in the gray level co-occurrencematrix.

Correlation Coefficient : Measures the joint probability occurrence of the specified pixel pairs.

sum(sum((x- μx)(y-μy)p(x , y)/σxσy))

Homogeneity : Measures the closeness of the distribution of elements in the GLCM to the GLCM diagonal.

sum(sum(p(x , y)/(1 + [x-y])))

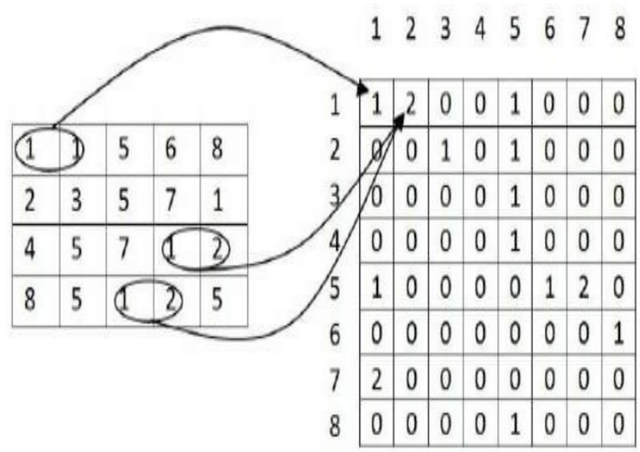
K-Means clustering algorithmic program is an unsupervised algorithmic program and its accustomed section the interest space from the background. It clusters, or partitions the given information into K-clusters or components supported the K-centroids. The algorithmic program is employed once you have unlabeled data (i.e. information while not outlined classes or groups). The goalis to seek out sure teams supported some quite similarity within the information with the amount of teams drawn by K. the target of K-Means agglomeration is to reduce the totalof square distances between all points and therefore the cluster center.

4.1.5. IMAGESEGMENTATION

Image segmentation is the method of partitioning a digital image into multiple distinct regions containing every picture element (sets of pixels, additionally referred to as super pixels) with similar attributes. Image segmentation is that the classification of a picture into completely different teams. Several forms of analysis are worn out the world of image segmentation exploitation bunch. Their area unit completely different strategies and one in every of the foremost well-liked strategies is K-Means bunch algorithmic program

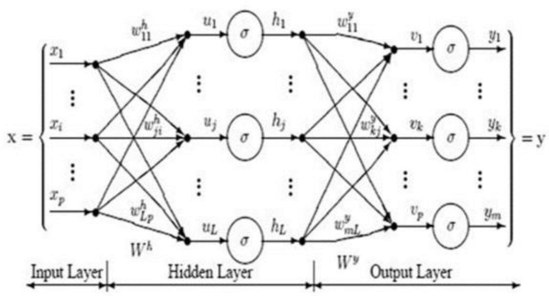
4.1.6. CO-OCCURENCEMATRIX

A Co-occurrence matrix (CCM) by conniving however usually a picture element with the intensity (gray-level) price i happens in an exceedingly specific spatial relationship to a picture element with the worth j. By default, the spatial relationship is outlined because the picture element of interest and also the picture element to its immediate right (horizontally adjacent), however you'll specify alternative spatial relationships between the 2 pixels. Every component (i, j) within the resultant CCM is just the total of the amount of times that the picture element with price i occurred within the such that spatial relationship to a picture element with price j within the input image. The amount of grey levels within the image determines the dimensions of the CCM.

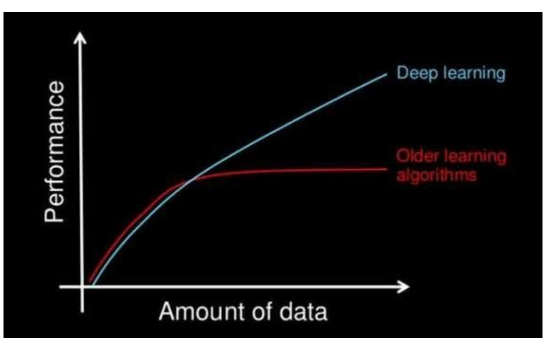


4.1.7. CONVOLUTIONAL NEURALNETWORK

Convolution Layer: The convolution layer contains of a group of freelance filters. Every filter is severally convolved with the image and that we find yourself with half dozen feature maps of form 28*28*1. Pooling is another building block of CNN. Its performance is to increasingly cut back the spatial size of the illustration to cut back the quantityof parameters and computation within the network. Pooling layer operates on every feature map severally. A convolution multiplies a matrix of pixels with a filter matrix or 'kernel' and sums up the multiplication values. Then the convolution slides over to successive constituent and repeats an equivalent method till all the image pixels are lined. The name "convolutional neural network" indicates that the network employs a computation known as convolution. Convolution could be a specialized quite linear operation. Convolutional networks are merely neural networks that use convolution in situ of general matrix operation in a minimum of one amongst their layers. A convolutional neural network consists of associate input associated associate output layer, additionally as multiple hidden layers. The hidden layers of a CNN usually carries with it a series of convolutional layers that twist with a multiplication or alternative scalar product. The activation perform is usually a RELU layer, and is afterward followed by extra convolutions like pooling layers, totally connected layers and standardization layers, stated as hidden layers as a result of their inputs and outputs are disguised by the activation perform and final convolution. Though the layers are conversationally stated as convolutions, this can be often solely by convention. Mathematically, it's technically a slippery scalar product or cross-correlation. This has significance for the indices within the matrix, in this it affects however weight is decided at a selected index purpose.



The above diagram shows the layers in the CNN algorithm.



The above image shows the graph between the parameter of performance and the amount of data. This is done for two algorithms (i.e.) deep learning algorithm and the other learning algorithm. From the graph we are inferring that the deep learning algorithm is the higher performance algorithm.

4.2 DATASET DESCRIPTION

Our dataset consist of both normal and abnormal leaf (i.e)disease affected leaf. We categorize the diseases as bacteria, fungi, bacteria and spotted leaf. And the datasets are categorized based on the disease categories.

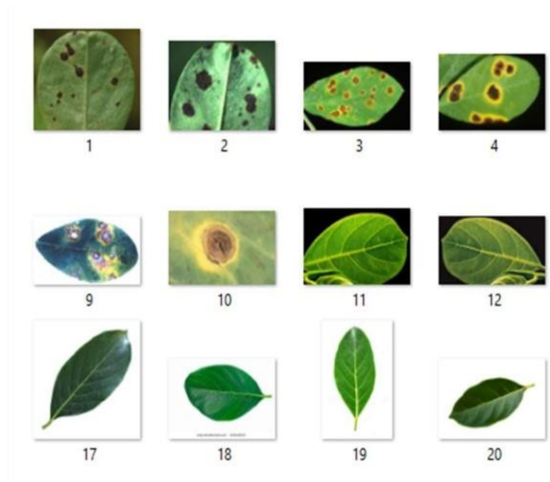


Fig 4.2.1 sample of dataset

The above picture shows the sample of our dataset.Thes datasets are categorized into two as testing dataset and training dataset. The testing dataset are used to test the system

5. EXISTINGSYSTEM

In existing system, the Principal element analysis and texture based mostly segmentation area unit used for image process. KNN is employed for classification. The principal element analysis uses mathematical functions including a vicinity on Statistics that appearance at distribution measurements, or, however the information is unfolded. The opposite section is on algebra and appears at eigenvectors and eigenvalues, vital properties of matrices that area unit elementary to PCA. Texture based mostly segmentation is employed to method the image supported its texture and color contrasts. It involves fixing of ROI (Region of Interest). KNN classifier clusters the processed pictures. It's a non-constant quantity lazy learning algorithm. It implies that it doesn't build any assumptions on the underlying information distribution.

5.1. DRAWBACKS OF EXISTING SYSTEM:

i. High computational load is faced in the existingsystem.

ii. Poor discriminatorypower.

iii. Less accuracy inclassification.

6. PROPOSED SYSTEM:

In proposed system, organization to differentiate among collected pictures of groundnut leaf pictures supported the kind of sickness known. The system conjointly suggests the nameofthechemicaltobeusedonthesphere.atechniquethat build use of image process techniques and segmentation of the processed pictures for each coaching and testing phases. Ittrainsaconvolutionalneuralnetworkwiththeassistanceof assorted datasets and uses the trained network for more classification of take a look at pictures.In this the images can be taken and loaded in the system. The sub-process in the system are done and the classification is done. If the leaf is affected by disease then it will display the name of the disease, pesticide which is needed to be given and the quantity of the pesticide needed. Else if the leaf is normal the system will display “Normal leaf”only.

7. IMPLEMENTATIONRESULT:

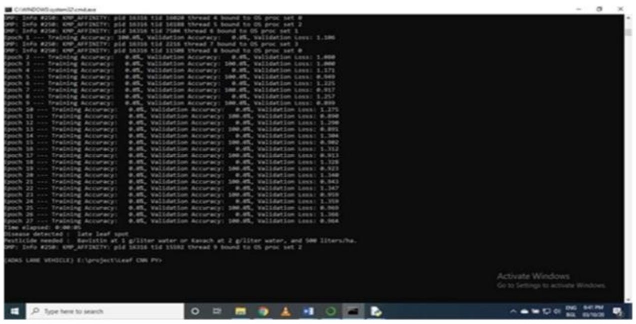


Fig7.1Diseaseleafidentification.

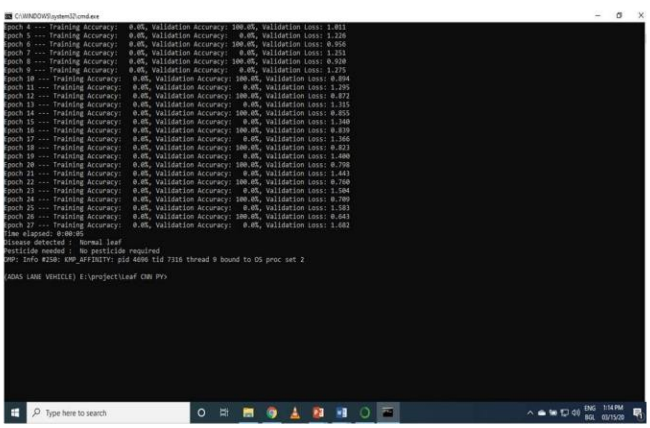


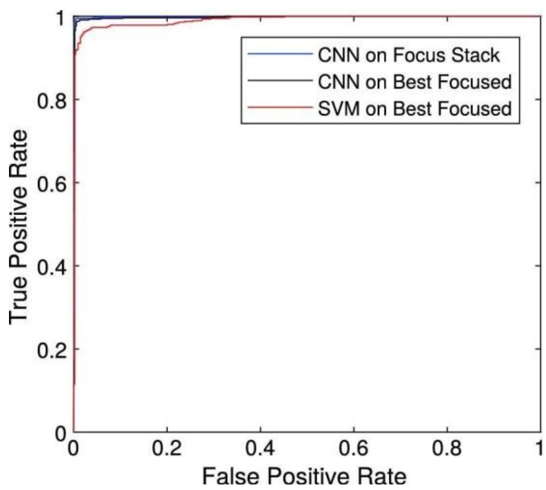
Fig7.2Normalleafidentification

The above shown two figures (fig7.1 and fig 7.2) will be the output of the system. The program is executed by the following steps: First we have to set the directory in the code. The next step is to set the path. Next step will the execution step. Finally the output is displayed in two category, one category is normal leaf and the other category is disease affectedleaf.

8. RESULT ANDDISCUSSSION:

Our proposed system will classify the leaf whether it is normal or diseased leaf. The classification accuracy will be higher and the loss in the training and the loss of data will be less. So our system will be more efficient than the other for the specified leaf.

Trial No.	Plant’s Position	Agriculturist Diagnosis	Output of the System	Assessment
1	1-1	Bacteria	Bacteria	Successful
2	1-2	Fungi	Fungi	Successful
3	1-3	Fungi	Fungi	Successful
4	1-4	Bacteria	Bacteria	Successful
5	2-1	Bacteria	Bacteria	Successful
6	2-2	Fungi	Fungi	Successful
7	2-3	Fungi	Bacteria	Failed
8	2-4	Fungi	Fungi	Successful
9	3-1	virus	virus	Successful
10	3-2	Fungi	Fungi	Successful
11	3-3	Nematodes	Nematodes	Successful
12	3-4	Fungi	Fungi	Successful
13	1-1	Fungi	Fungi	Successful
14	1-2	normal	normal	Successful
15	1-3	Nematodes	Nematodes	Successful
16	1-4	Fungi	Fungi	Successful
17	2-1	Fungi	Fungi	Successful
18	2-2	Nematodes	Nematodes	Successful
19	2-3	Fungi	Fungi	Successful
20	2-4	Fungi	Bacteria	Failed
21	3-1	Bacteria	Bacteria	Successful
22	3-2	Nematodes	Nematodes	Successful
23	3-3	Fungi	Fungi	Successful
24	3-4	normal	normal	Successful
25	1-1	virus	virus	Successful
26	1-2	virus	virus	Successful
27	1-3	Fungi	virus	Failed
28	1-4	virus	virus	Successful
29	2-1	Fungi	Fungi	Successful
30	2-2	Fungi	Fungi	Successful
31	2-3	virus	virus	Successful
32	2-4	Fungi	Fungi	Successful
33	3-1	Fungi	Fungi	Successful
34	3-2	Nematodes	Nematodes	Successful
35	3-3	Fungi	Fungi	Successful
36	3-4	Fungi	Fungi	Successful
Total: (33 out of 36) **91.6%**				



Theaboveimagehastheparameterfalsepositiverateandthe true positive rate. CNN has higher true positive rate and lesser false positive rate. The different colours in the graph shows the different algorithm basedsystem.

9. CONCLUSION

Finally, we have proposed a novel methodology to identify plant leaf disease using the images and classifythose images based on the different types of diseases. The system is able to notify the disease to the farmers and suggest suitable pesticides to be used. Efficient image processing techniques like Hybrid spatial features including color features and texture descriptors has been employed here. It provides low complexity and better feature recognition of leaf images. The method uses Convolutional NeuralNetwork for better precision and accuracy in classification and training.

10. FUTUREENHANCEMENT

The definition of a technique based on such features which by resorting at trained classifiers is able to identify the leaf disease in groundnut plant and also to suggest the suitable pesticides. This module is expected to be deal with more number of datasets in efficient manner. Also it can be implemented as a real-time system that make use of Mobile application facilities in future. The future enhancement can be done as live streaming technique. This live streaming will lead to monitor the large area of fields and to obtain more accuracy in the system. The other enhancement is by adding the feature to be extracted in the image. Due to these additional feature extraction parameter we can improve the accuracy of thesystem.

11. REFERENCES

[1]H. Al-Hiary, S. Bani-Ahmad, M. Reyalat, M. Braik and Z. ALRahamneh, Fast and Accurate Detection and Classification of Plant Diseases‖, IJCA, 2011, 17(1), 31-38, IEEE-2010

[2] Jayamala K. Patil, Raj Kumar, ―Advances In Image Processing For Detection of Plant Diseases‖, JABAR, 2011, 2(2),135-141.

[3] JyotiShirahatti, RutujaPatil, Pooja Akulwar "A Survey Paper on Plant Disease Identification Using Machine Learning Approach" Proceedings of the International Conference on Communication and Electronics Systems (ICCES2018)

[4]N.KRITHIKA, DR.A.GRACE SELVARANI "An Individual Grape Leaf Disease Identification Using Leaf Skeletons and KNN Classification" 2017 International Conference on Innovations in Information, Embedded and Communication Systems (ICIIECS)

[5] PENG JIANG, YUEHAN CHEN, BIN LIU, DONGJIAN HE AND CHUNQUAN LIANG, "Real-Time Detection of Apple Leaf Diseases Using Deep Learning Approach Basedon Improved Convolutional NeuralNetworks,"

[6] Piyush Chaudhary, Anand K. Chaudhari, Dr. A. N. Cheeran and Sharda Godara, Color Transform Based Approach for DiseaseSpotDetectiononPlantLeaf‖,IJCST,2012,3(6),65- 70.

[7] PriyadarshiniPatil, NagaratnaYaligar, Meena S M "Comparision of Performance of Classifiers - SVM, RF and ANN in Potato Blight Disease Detection using Leaf Images" 2017 IEEE International Conference on Computational Intelligence and Computing Research

[8]P.Revathi, M.Hemalatha, ― Classification of Cotton Leaf Spot Diseases Using Image Processing Edge Detection Techniques‖, ISBN, 2012, 169-173, IEEE.

[9]S.M. Jaisakthi, P. Mirunalini, D. Thenmozhi, Vatsala "Grape Leaf Disease Identification using Machine Learning Techniques" Second International Conference on Computational Intelligence in Data Science (ICCIDS-2019)

[10] Yogita K. Dubey*, Milind M. Mushrif, SonamTiple "Superpixel Based Roughness Measure For Cotton Leaf Diseases Detection and Classification" 4th Int'l Conf. on Recent Advances in Information Technology |RAIT-2018|