IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Crop Disease Detection using CNN

Vineet Kumar[1], Shubham Mishra[2], ShivanshTrivedi[3], Sonali Choudhary[4]

[1]Associate Professor, [2,3,4]B.tech Students

[1,2,3,4]Department of Computer Science and Engineering

[1,2,3,4]Noida Institute of Engineering and Technology, Gr. Noida, India

Abstract:

In developing countries, farming land can be huge and farmers are unable to observe each and every plant, every day. Farmers are unaware of non-native diseases. Consultation of experts for this is time consuming & costly as well. So there is a need to develop an automated system to detect plant disease and classify them using image processing.

Therefore deep learning and convolution neural network are used to detect the disease in agricultural harvest. Convolution neural networks provide a better accuracy and result by classifying the images to healthy and non-healthy categories as they are specially designed to process pixel data. This approach has two phases: first phase deals with training the model for healthy as well as unhealthy images, second phase involves the monitoring of crops and identification of particular disease in the plant which ultimately results in early detection of the disease.

Keywords: **Deep Learning, Convolution Neural Network, Image Processing, Leaf Diseases and Detection.**

I. INTRODUCTION

Agriculture is amongst an important source of livelihood/income in India. The country's population is directly or indirectly related to the agricultural sector. Therefore, the production of high quality agricultural products is necessary to support the country's economic development. Therefore, in the agricultural sector, the discovery of plant diseases plays an important role. All living things depend on agriculture for their livelihood. But to get the best yields, the plants need to be healthy so another technical approach is needed to look after them from time to time. Plant diseases are one of the most important factors that can lead to a significant decline in the quality and quantity of agricultural products. The growing social trend, weather conditions also cause plant disease. Crops suffer from diseases that can significantly affect the number and quality of the crop. Often the diagnosis of leaf diseases is done by farmers by eye contact. It leads to a misdiagnosis as a farmer predicts the symptoms by their experience. This will also create unnecessary and excessive use of expensive pesticides. The advancement in technology, specifically the use of image processing in combination with the deep learning approach would help the farmers in terms of discovering the plant disease in the initial stages. Accurate and early detection of Disease would help reducing the use of more pesticide and increase the quality of the crop. Detection through naked eyes would involve a large team of experts to diagnose it which is too costly. We know that presence of disease on the plant is mainly reflected by symptoms on leaves. So an automatic, accurate and less expensive Machine Vision System for detection of diseases from the image suggests a better solution to the problem.

II. PLANT DISEASES

A. Bacterial Blight/Spot :

It is a bacterial infection mainly caused by germs "Xanthomonas Campestris pv Malvacearum "[1].Effects of bacterial blight damage start to occur as dark grass, water spots of 1 to 5 mm also occur sideways in red-brown leaf border. Initially, these angular areas appear as Water logged areas that later change from dark brown in black [2].

Fig 1: Bacterial Blight

B. Leaf Mold:

The disease leaf mold is caused mainly by the fungus Passalora fulva (previously called Fulvia fulva). It is not known to be pathogenic on any plant.Mainly targets tomato where the infected blossoms turn black and fall off.

Fig 2: Leaf Mold

C. Spider Mite:

Spider mite are insects that generally feed on the lower surface of the leaf as a result the infected leaf turns yellow finally leading to defoliation they spread to all parts of the plant as their population increases[3].

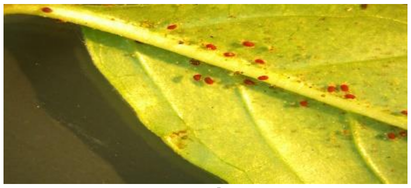


Fig 3: Spider Mite

D. Leaf Curl:

Leaf curl is a crop disease in which distortion and coloration of the leaves occur and is caused by the fungus Begomovirus(Geminiviridae). It may even be caused by abiotic disorders and herbicides.Mostly occurs in tropical and subtropical regions it is one of the most destructive disease of tomato.

Fig 4: Leaf Curl

E. Mosaic Virus:

It causes the plant to have a mottled appearance and foliage.It is very contagious when infected leaves rub against healthy leaves it gets transfered.It attacks more than 150 types of plants including vegetables and fruits.

Fig 5: Mosaic Virus

III LITERATURE SURVEY

Visual Analysis, Image Processing and Optical Sensor are few implementations for the disease detection method. By using these methods a system can be developed to detect diseases earlier and be used to overcome this problem. On comparing these methods we find that detection by using visual analysis is not accurate and optical sensor are not easy to implement and costly. Therefore image processing is the only way which can
be used to build a simple robust and accurate disease detection system [4]. On working with image processing, database collection is the most challenging task. So you must have knowledge about the diseases you are going to predict.

There are different types of algorithms used for detection of crop diseases to attain better results and accuracy. Few of the models that we went through involved the use of PNN (probabilistic neural network), Canny Edge Detection and SVM (support vector machine) to detect the disease. Canny edge detection is used for rice, wheat crops and the images of crops are converted to gray scale. SVM algorithm obtains the edges accurately/precisely as many disease cause leaf deformities. SVM classifier is used for crops which classifies normal and diseased images. Features from these images are extracted and classified into healthy and diseased crops and then accuracy is calculated[5].

Probabilistic Neural Network (PNN) is adopted for it has fast speed on training and simple structure features are extracted and treated by Principal Component Analysis (PCA) to form the input vector of PNN. Experimental result indicates that the algorithm is workable with accuracy greater than 90%[6].

BP (back propagation) distribution networks have been used as dividers to differentiate grape and wheat diseases, respectively. Results show such diagnostics can be successfully achieved using BP networks. While the size of element details were not downgraded using (PCA).

Out of all these approaches we find that CNN comes out to provide us with the best and accurate results.

As it is specially designed to perform calculations over pixel data. Convolution network is better than feed forward network because CNN has features for sharing parameters and reducing size. Due to the parameter sharing on CNN, the number of computations is therefore reduced. It extract features from images by dividing it, pooling and stacking small areas of the image.

A Convolutional Neural Network (ConvNet/CNN) is a Deep Learning Algorithm which can take in an input image, assign importance (learnable weights and biases) to various aspects/objects in the image and be able to differentiate one from the other thus providing better accuracy than traditional approaches.

On validating the datasets we came out with an accuracy of 95% with 70 iterations (epochs-no of passes) of training and expect the accuracy to increase up to 98-99% with a rise in number of epochs.

Table I: Review of Papers

Title	**Date**	**Author**	**Methodology**
Leaf Disease Detection and Classification using Neural Networks[7]	November 2016	*V. Ramya , M. Anthuvan Lydia*	PNN,BP
Crop Disease Detection using Machine Learning : Indian Agriculture[8]	September 2018	*Anuradha Badage*	Canny Edge Detection
Farmer Advisory: A Crop Disease Detection System [9]	May 2019	*Anuradha Badage , Aishwarya C , Ashwini N , Navitha K Singh , Neha Vijayananda*	SVM

IV PROPOSED METHODOLOGY

The proposed approach starts with firstly collecting the dataset of the crops for which the diseases would be predicted after acquiring the dataset we need to train the images for both healthy as well as unhealthy leaf images. Once we are done with training a full connected neural network is formed for the classification of the images.

Let`s drive through how the images would be trained and classified and finally a result would be obtained using CNN.

A. Working of Convolution Neural Networks

Convolution Neural Networks architecture looks somewhat like this:
Input (image)->Convolution ->ReLU ->Convolution ->ReLU ->Pooling -> ReLU ->Convolution ->ReLU ->Pooling ->Fully Connected Layer.

Compared to other image classification algorithms, CNNs actually use very little preprocessing which means they are ideal for image classification.

CNNs have an input layer, and output layer, and hidden layers. The hidden layers usually consist of convolutional layers, ReLU layers, pooling layers, and fully connected layers.

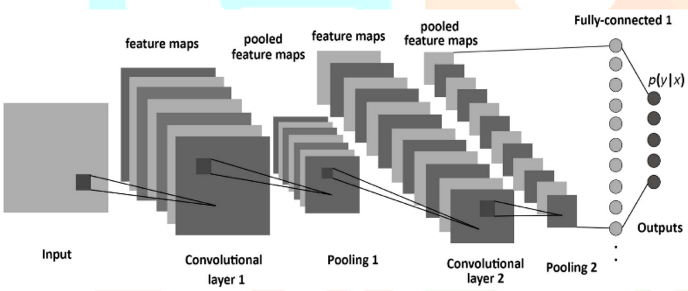


Fig 6. Basic Architecture of CNN

A CNN starts with an input image, applies many different filters to it to create a feature map then applies a ReLU function to increase non-linearity after that it applies a pooling layer to each feature map, flattens the pooled images into one long vector. Inputs the vector into a fully connected neural network. Processes the features through the network. The final fully connected layer provides the probability of the classes, and are trained through many epochs. This loops till we get a well-defined neural network with trained weights and feature detectors.

Convolution neural network or Convnets works with layers. Where one part involves extracting the features of the images and second part involves classifying the image.

We use convolutional neural network for the creation of model and specify filter size for the Conv layer and Pool layer and the shape on each layer

Where the shape is described in the form of (height, weight, channels). We feed the image in size of 64*64 pixel with 3 channels (RGB) red, green, blue.

Once this image is fed into the network first of all the **Conv2d** layer is applied to the image with 3 channels and the images are resized with output shape of (60,60,32). The output size of the image is given by the function know as Convolution Arithmetic: (W - F + 2P) / S + 1.
Where W = Input Size F = Filter Size P = Padding Size
S = Stride. The number of pixels by which we move the filter over the input matrix is called a stride.

The images are represented in the form of matrix and dot product of the image matrix and **kernel**[10] **(a filter of 3*3 matrix**) is done to identify the features of the image, filters are nothing but feature detectors.

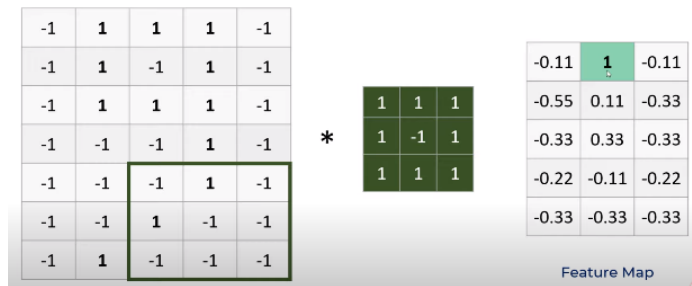


Fig 7. Conv2d

Relu Activation stands for rectified linear unit, it is the most commonly used function in CNN If you are unsure of which activation function to use on your network, ReLU is usually a good first choice. The ReLU function is simple and has no heavy calculations as there are no complicated mathematical calculaitons. The model, therefore, takes less time to train or run. One of the most important assets we consider as the benefit of using the ReLU activation function is sparsity. This function is applied after the conv2d layer which reduces the non-linearity in the images or in simple words it replaces the negative values in the matrix with 0 and for the positive values it returns them back as it is.

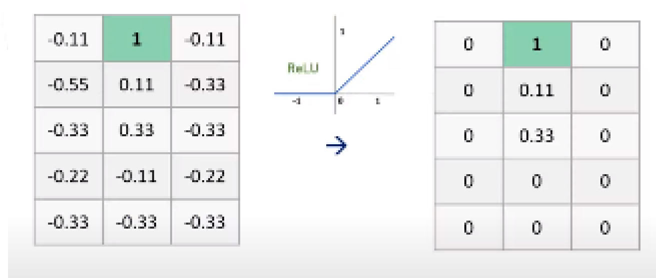


Fig 8. Relu Activation

MaxPooling till now we haven't talked about reducing the computation, therefore we use POOLING to reduce the dimensionality of the image for lesser computation. We use Maxpooling function here which takes the feature map and a window of 2*2 or 3*3 and pick out the largest number among it and forms a new feature map.

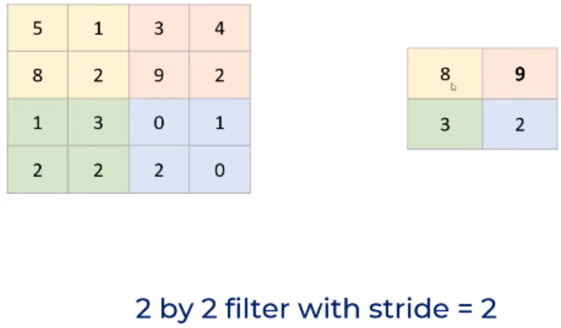
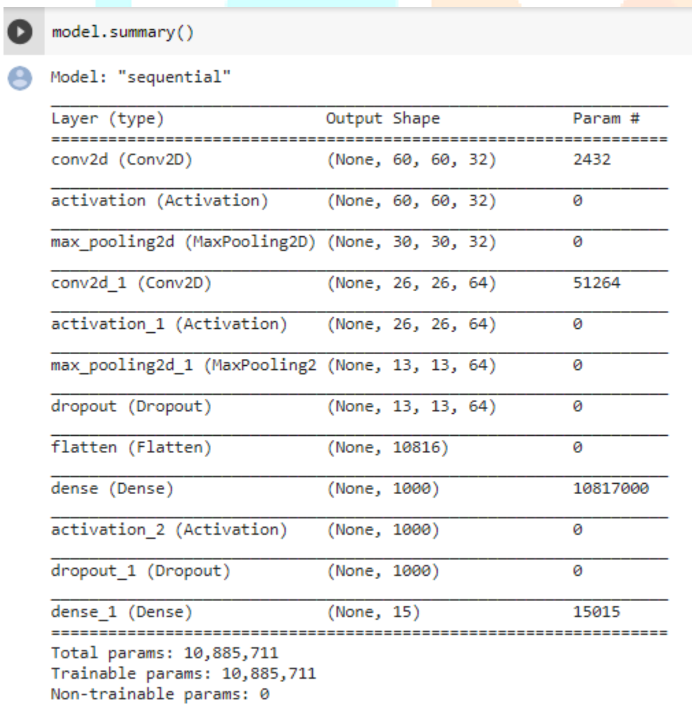


Fig 9. MaxPooling

Pooling[11] is a way to take big pictures and reduce them out while keeping important details in them/Maxpooling takes only the most relevant features from the images and filters out the noise in the image. **Conv2d, Relu** with **maxpooling** helps us in location invariant detection of features from an image it maintains a maximum value from each window, maintaining a good balance of each element within the window. This means that it does not really matter exactly where the feature fits as long as it fits somewhere inside the window. The result of this is that CNN can detect if a feature is in the image without worrying about where it is.

Dropout temporarily removes the use of randomly selected neurons to regularize the neural network. Once these layers are applied the feature map is then flattened in Single dimension and a fully connected neural network is formed.



```
model.summary()

Model: "sequential"
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
conv2d (Conv2D)              (None, 60, 60, 32)        2432
_________________________________________________________________
activation (Activation)      (None, 60, 60, 32)        0
_________________________________________________________________
max_pooling2d (MaxPooling2D) (None, 30, 30, 32)        0
_________________________________________________________________
conv2d_1 (Conv2D)            (None, 26, 26, 64)        51264
_________________________________________________________________
activation_1 (Activation)    (None, 26, 26, 64)        0
_________________________________________________________________
max_pooling2d_1 (MaxPooling2 (None, 13, 13, 64)        0
_________________________________________________________________
dropout (Dropout)            (None, 13, 13, 64)        0
_________________________________________________________________
flatten (Flatten)            (None, 10816)             0
_________________________________________________________________
dense (Dense)                (None, 1000)              10817000
_________________________________________________________________
activation_2 (Activation)    (None, 1000)              0
_________________________________________________________________
dropout_1 (Dropout)          (None, 1000)              0
_________________________________________________________________
dense_1 (Dense)              (None, 15)                15015
=================================================================
Total params: 10,885,711
Trainable params: 10,885,711
Non-trainable params: 0
_________________________________________________________________
```

Fig 10. Trained Model

B. Classification of the images(SOFTMAX Function)

Fully connected layers take high-quality filtered images and translate them into votes. Fully connected layers are the main basis for traditional neural networks. Instead of treating input as a two-dimensional list, they are treated as a single list

After maxpooling the slices are flattened into single dimension now we feed this layer as our final layer to **SOFTMAX** function and **predict 15 categories** i.e. the image dataset is divided into 15 classes out which 3 are images of healthy leaves and 12 are of diseased category.

So as model output we get a tensor of 1X15 size.

Now the SOFTMAX function calculates probability for each class form where the input belongs, the number of softmax units in the output layer should be equal to the no.of classes so that probability can be assigned to each unit which is also be termed as **Probability distribution.**

Formula to assign probability

$$f(x_i) = \exp(x_i) / \Sigma_j \exp(x_j) \ [12]$$

Exponential of a particular class over sum of exponential of each class.
For example:
A image of car is fed to a network and softmax assigns probability to each unit we find that the probability for the image being an car is highest which is are predicted output.

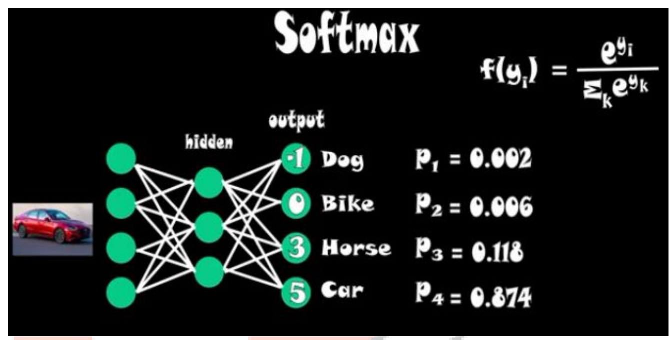


Fig 11.Example of Softmax Function

Therefore we come to know that the unit with the highest probability comes out to be our predicted output using the softmax function.

V. RESULTS

As CNN is basically designed for image classification it yields better accuracy. We find out that after training and testing the model to detect the crop disease, the attained accuracy are as follows:We get a training accuracy of 95.55% and testing accuracy of 95.69%, as we have performed only 70 EPOCHS (no of passes) to train this data, on increasing these number of epochs to 200-250 our accuracy of prediction will also rise up to 98-99% leading to correct prediction of the disease.

```
Epoch 00069: val_loss improved from 0.15648 to 0.14608, saving model to disease_selected_100.h5
Epoch 70/70
256/256 [==============================] - 90s 353ms/step - loss: 0.0578 - accuracy: 0.9799 - val_loss: 0.1691 - val_acc

Epoch 00070: val_loss did not improve from 0.14608
Traing finished.
Loading the best model...
Best Model loaded!
```

```
model.evaluate(X_test, Y_test)
```

```
64/64 [==============================] - 3s 42ms/step - loss: 0.1491 - accuracy: 0.9569
[0.1491481214761734, 0.9569261074066162]
```

Fig 12. Accuracy

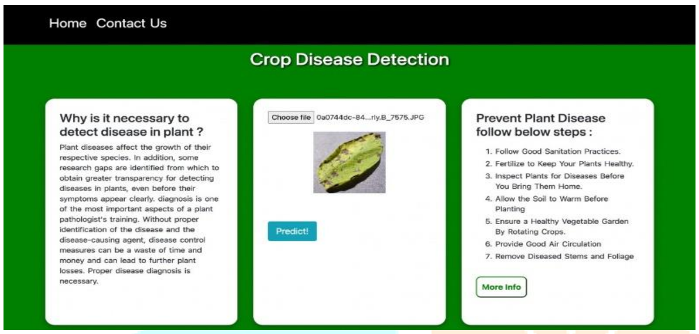


Fig 13. UI

Fig 14. Predicted Disease

VI. CONCLUSION

Image Processing proposes an innovative and automated system that is ideal to detect crop diseases in disease affected crops. Early detection through CNN, deep learning method helps the farmer to be alerted within time and appropriate measures can be taken to cure the disease. Images are fed into the network features are extracted from the image and then are classified as healthy or unhealthy. The accuracy can be increased by carrying out more number of epochs and better accuracy definitely leads to good prediction as well as accurate identification of the disease affecting the crop. This saves our farmers from great loss and boosting their agricultural growth. Therefore it can be concluded that CNN serves as the best approach to classify these images and detect the diseases out of all the available approaches.

ACKNOWLEDGEMENT

Firstly the student authors would like to extend their regards to Prof. Vineet Kumar associate professor Dept. of Computer Science and Engineering NIET, Gr.Noida for his kind permission to carry out this research work and also supervising us throughout the project. The authors would like to thank all the colleagues for their valuable inputs and the faculty of Dept. of Computer Science and Engineering for their support, encouragement and invaluable guidance for carrying out this project.

REFERENCE

[1] P. R. Rothe, R. V. Kshirsagar, "*Cotton Leaf Disease Identification using Pattern Recognition Techniques*", International Conference on Pervasive Computing (ICPC) IEEE, pp. 1-6, 2015
.

[2] Viraj A. Gulhane, Maheshkuma R. H. Kolekar, "*Diagnosis Of Diseases On Cotton Leaves Using Principal Component Analysis Classifier*", Annual IEEE India Conference (INDICON), 2014.

[3] Uma Shankar,Hafeez Ahmad,Suheel Ganai "*Seasonal Incidence and Management of Red Spider Mite, Tetranychus urticae Koch. Infesting Rose*"International Journal of Current Microbiology and Applied Sciences · September 2017.

[4] A. Miyatra, S. Solanki. 2014. "*Disease and Nutrient Deficiency Detection in Cotton Plant*," International Journal of Engineering Development and Research, vol. 2, no 2, pp. 2801-2804.

[5] Anuradha Badage, Aishwarya C, Ashwini N, Navitha K Singh, Neha Vijayananda "*Farmer Advisory: A Crop Disease Detection System*" International Research Journal of Engineering and Technology (IRJET) Volume: 06 Issue: 05 | May 2019

[6] V. Ramya, M. Anthuvan Lydia "Leaf Disease Detection and Classification using Neural Networks" International Journal of Advanced Research in Computer and Communication Engineering ISO 3297:2007 Certified Vol. 5, Issue 11, November 2016

[7] V. Ramya, M. Anthuvan Lydia "Leaf Disease Detection and Classification using Neural Networks" International Journal of Advanced Research in Computer and Communication Engineering ISO 3297:2007 Certified Vol. 5, Issue 11, November 2016

[8] Anuradha Badage "*Crop Disease Detection using Machine Learning: Indian Agriculture*" International Research Journal of Engineering and Technology (IRJET) Volume: 05 Issue: 09 | Sep 2018

[9] Anuradha Badage, Aishwarya C, Ashwini N, Navitha K Singh, Neha Vijayananda "*Farmer Advisory: A Crop Disease Detection System*" International Research Journal of Engineering and Technology (IRJET) Volume: 06 Issue: 05 | May 2019

[10] Vincent Dumoulin and Francesco Visin "*A guide to convolution arithmetic for deep learning*" arXiv:1603.07285v2 [stat.ML] 11 Jan 2018.

[11] Vincent Dumoulin and Francesco Visin "*A guide to convolution arithmetic for deep learning*" arXiv:1603.07285v2 [stat.ML] 11 Jan 2018.

[12] Chigozie Enyinna Nwankpa, Winifred Ijomah, Anthony Gachagan, and Stephen Marshall "*Activation Functions: Comparison of Trends in Practice and Research for Deep Learning*" arXiv:1811.03378v1 [cs.LG] 8 Nov 2018