Crop Disease Detection using Convolutional Neural Network

Aayush Singh
Department of Computer Engineering
Rajiv Gandhi Institute of Technology, Mumbai, India
aayushipsingh@gmail.com

Deepanshu Sonparote
Department of Computer Engineering
Rajiv Gandhi Institute of Technology, Mumbai, India
deepanshusonparote@gmail.com

Nitish Talekar
Department of Computer Engineering
Rajiv Gandhi Institute of Technology, Mumbai, India
nitishtalekar.nt503@gmail.com

Dilip S. Kale
Assistant Professor, Department of Computer Engineering
Rajiv Gandhi Institute of Technology, Mumbai, India
dilip.kale@mctrgit.ac.in

Abstract—**The crop agriculture industry faces the economic losses due to the pest infections, bacterial or viral contagions, the farmers lose nearly a fifth of the total profit on an average annually in India. This paper proposes a solution to the agricultural problem, which involves crop disease recognition by using machine learning and image processing techniques. In this paper, the study sets out to classify cotton crop images into classes, whether the crop is infected by a disease or not. Also, we endeavor applications that give the farmer readily available means to identify the diseases on their crop and take appropriate damage control actions. The dataset used to train the model was user created (mobile capture images with high-resolution camera) from various crop farms. Crops, of different varieties, containing three types of diseases, namely Rust, Mosaic Virus, Woolyaphids and Healthy plants are taken as classification ideals. The trained models have provided a performance reaching an high success rate in identifying the corresponding cotton plant disease. The model used in the study delivers significant accuracies of classification on the dataset used by employing Convolutional Neural Networks. The model is a very useful advisory or early warning tool for the farmers for identification of diseases in the early stage so that immediate action can be taken.**

Keywords— Crop; Cotton; Convolutional Neural Network; Image processing;

I. INTRODUCTION

Agriculture is the primary occupation in India. Crops are part and parcel of the agricultural industry. Nowadays, a tremendous loss in the quality and quantity of crop yield is observed, subject to various diseases in the plant. Crop Plant disease classification is a critical step, which can be useful in early detection of pest, insects, disease control, increase in productivity, among other examples. Farmers recognize disease manually with foregoing symptoms of plants, and with experts, whereas the actual diseases are hard to distinguish with naked eye, and it is time-consuming to predict whether the crop is healthy or not. Cotton is one of the major agricultural crops in India and it has a dominant impact on the overall Indian agriculture sector. Cotton plant leaf disease diagnosis is very difficult through simple observation. Due to the complexity, even experienced agronomists and plant pathologists often fail to successfully diagnose specific diseases and are consequently led to mistaken conclusions and treatments.

Our study helps to predict crop diseases in cotton plants by processing the images of the crop. For this, Image Processing techniques are used for fast, accurate and appropriate classification of diseases. Symptoms of diseases in cotton predominantly come out on leaves of plants. The existence of an automated system for the detection and diagnosis of plant diseases would offer a support system to the agronomist who is performing such diagnosis through observation of the leaves of infected plants.

The existing techniques for disease detection have utilized various image processing methods followed by varied classification techniques. The proposed system uses a convolutional neural network to classify the health of a cotton leaf plant. The flow diagram of the proposed system given in Fig. 1. consists of steps used to acquire the desired output.
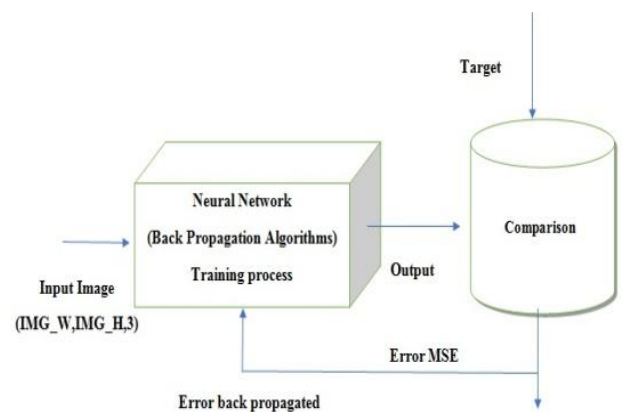

Fig. 1. Flow Diagram of Proposed System

II. Literature Review

Crop Yield Forecasting has been an area of interest for producers, agricultural-related organizations. Timely and accurate crop yield forecasts are essential for crop production. The existing techniques for disease detection have utilized various image processing methods followed by various classification techniques. However, some unconventional approaches have led to classification of diseases using unconventional factors. Chopda et al. [1] proposes a system which can predict the cotton crop diseases using decision tree with the help of the parameters like temperature, soil moisture, etc. based on the previous year data and through sensors. However, these data might not be fully dependable to predict or classify diseases.

Image classification and regression techniques play a very important role because they allow identifying, grouping, and organizing from a standardized system. We apply an algorithm for image segmentation technique on data for automatic detection and classification of plant leaf diseases. In [2], Kamble defines the application of texture analysis for detecting plant diseases with the help of different image processing technique. Further with the use of Decision-Making Module, the disease is classified. Singh and Misra [8] suggest different disease classification techniques that can be used for plant leaf disease detection and an algorithm for image segmentation, the advantage of using this method is that the plant diseases can be identified at an early stage or the initial stage.

Deep learning is a set of learning methods attempting to model data with complex architectures combining multiple non-linear transformations. The element of deep learning is the neural networks that are combined to form the deep neural networks. These techniques have enabled significant progress in the fields of image processing and image classification. Kulkarni [3] formulates an application of Deep Convolutional Neural Network to identify and classify crop disease on images, testing it on five classes of crops and three types of diseases for each class. Mique Jr, Eusebio L [4] proposed an application that will help farmers in detecting rice insect pests and diseases using Convolutional Neural Network (CNN) and image processing. The searching and comparison of captured images to a stack of rice pest images was implemented using a model based on CNN. Collected images were pre-processed and were used in training the model.

There exist several types of architectures for neural networks:

1) The multilayer perceptions, that are the oldest and simplest ones,
2) The Convolutional Neural Networks (CNN), particularly adapted for image processing
3) The recurrent neural networks and dense neural network (DNN) used for sequential data such as text or times series.

A different approach is taken by Petrellis [6] where mobile phone application for plant disease diagnosis is presented which is based on the detection of the disease signature that is expressed as a number of rules that concern the colour, the shape of the spots and historical weather data among other factors.

III. Diseases In Cotton Plant

A. Rust

Rust is a disease usually found in older leaves, where a yellow, orange-ish lesions are found on the upper layer of the leaves. These lesions grow with time and with time burst to release spores, which often coalesce and give rise to irregular dark brown spots.

Fig. 2.   Leaf infected with Rust disease

B. Mosaic Virus

Mosaic Virus is a viral infection in plants which can be transmitted via multiple carriers and does not get cured. The inability to cure this has made it necessary to conserve and defend against the contraction of virus. The virus affects the plant in multiple ways at multiple stages of a plant's lifetime. It can vary among Yellow/White spots, wrinkled leaves, stunted growth and development of 'warty' areas.

Fig. 3.   Leaf infected with Mosaic Virus

C. Woolly Aphids

Woolly Aphids are small parasite-like insects which utilize hosts such as cotton plants to lay eggs or feeding purposes. These generally feed on seeds, foliage or sometimes even roots of plants. The aphids are generally blue or green in colour, but appear white or cloudy due to the waxy material surrounding

their body. There are rare cases of severe attack by woolly aphids, therefore only a slight care of plants can ensure protection from these insects.

Fig. 4.   Woolly Aphids' attack on Stem of Plant

IV.   METHODOLOGY

A.   Dataset

The dataset contains a total of 8380 images. This is constituted by 2055 healthy plants images, 1800 having the Mosaic virus, 2155 plants being infected by Rust and 2370 samples of plants having Woolly aphids. Table 1 shows the composition of dataset of leaf images with class name and the number of plant leaf samples.

TABLE I.   AMOUNT OF DATA SAMPLES PER CLASS

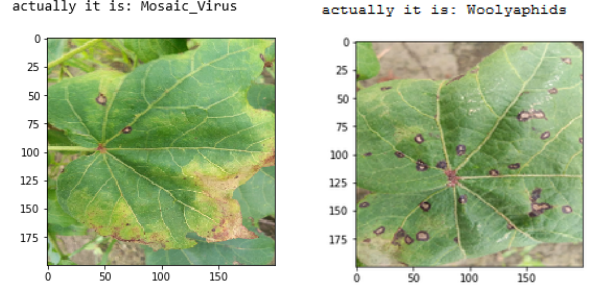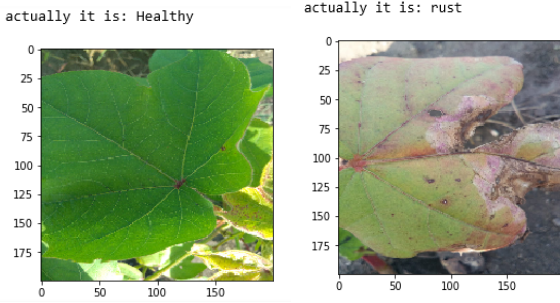| Image Class	Number of Samples
Healthy	2055
Mosaic virus	1800
Rust	2155
Woolly aphids	2370
Total	8380

Fig. 5.   Dataset Examples and their labels

B.   Preprocessing Data

For this initial process images of high resolution (4160 x 3120) are taken from datasets as input by setting IMG_W, IMG_H with 3 channels (RGB), for better visibility. The images are foremost pre-processed into a 4160 x 3120, RGB format with pixel values ranging from 0 to 255. Once normalized, the image is converted to a 224x224x3 input tensor and appended to the input array. The feature normalization used in the study is the min-max normalization. It is the ratio of the difference between the instance's feature value and the minimum value of a feature in the instance to the difference between the maximum and minimum values of features in the instance.

$$Norm(x) = \frac{x - xmin}{xmax - xmin}$$

C.   Classification Model

The convolution network model used for classification utilizes the architecture of a VGG16 Network. VGG16 model comprises of 13 Layers of convolution with increasing number on kernels after certain intervals, and 3 Dense layers for classification [10]. The VGG16 networks therefore has around 14M training parameters. To quicken up the training process and lower the time of execution, we have kept the number of kernels constant throughout the feature extraction process - 128. We have also added Batch Normalization layer after convolutions to drop unwanted nodes during the training process. These tweaks have proved to reduce the training parameter count to 9M and reduce training period significantly.
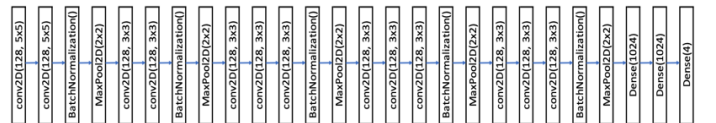

Fig. 6.   Layers of standard VGG16

Fig. 7.   Layers of modified VGG16

The activation function, ReLU is used for each of the hidden layers and the SoftMax function used in the last layer.

Adaptive gradient descent optimizer is used for optimization. The dataset is divided into a train test split of 70%-30%.

V. DISCUSSIONS AND RESULTS

The study was set out to classify if the crop is infected by a disease or is healthy. The training of model provided a significant quick learning process due to selected convolution layer structure and tweaks of layer addition. Adagrad is used as an optimization function to converge to the optimized classes quicker. The number of true values predicted correctly are represented as T, several false values predicted correctly are represented as F. The error was calculated using the error function mentioned.

$$Error = \left(\frac{F}{F + T}\right) * 100$$

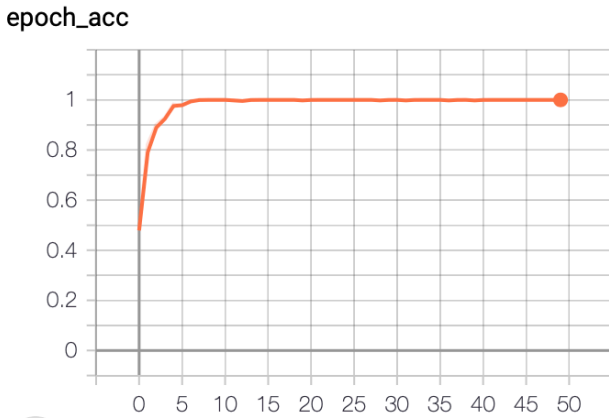The following graphs are mapping of training accuracy and loss over 50 epochs of training.

epoch_acc

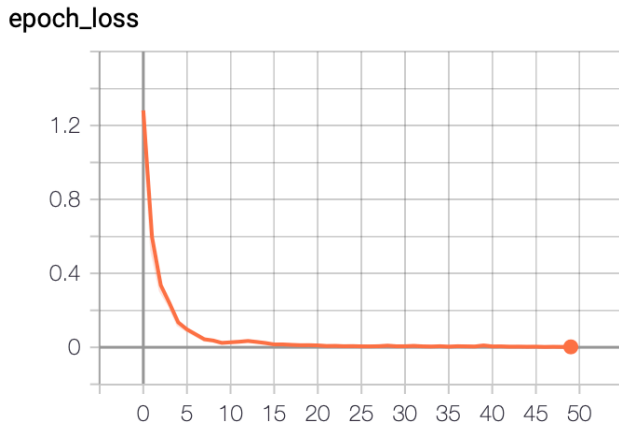Fig. 8.   Training Accuracy

epoch_loss

Fig. 9.   Training Loss

The training accuracy chart shows a significant consistency of training accuracy after 5 epochs of training, thus continuing to reduce loss over next 50 epochs. Consequently, the training loss can be observed to diminish exponentially over epoch 1 to 10, further decreasing as epoch progressing. This denotes the finer understanding of the pattern and regions in images, by the Network. Finally, on testing, we achieved a classification accuracy of 87.5%.

VI. CONCLUSIONS AND FUTURE SCOPE

In this paper, we have considered the cotton crop as it is the most important cash crop in India. Normally Rust, Mosaic virus and Woolly aphids are the hazardous diseases that the cotton crop suffers from in our country. Here, we consider a Convolutional Neural Network (VGG16) for crop disease recognition using leaf images for classification. There are several methods in computer vision for plant disease detection and classification process, but still, this research field is lacking. At present, there are no commercial solutions available in the market dealing with plant disease recognition based on the leaf images. We use a new approach of deep learning which automatically classifies and detect crop diseases from leaf images.

The data set of size 8380 images is collected from local farm with different diseases. The crop leaf images are considered for the training and testing purpose of the network. Initially, with the use of Gradient Descent and Back-Propagation algorithm classification are performed and it gives the prediction of diseases with 87.5% efficiency.

Furthermore, Convolutional Neural Network can be used for better classification accuracy. The main aim is to detect the crop leaf diseases from the database and train the images in such a way that the trained model gives the solution to farmers. The proposed model can recognize 11 different types of plant diseases. Here we consider plant stream and affected area by the disease boundaries, color variation, size and shape of plant leaves.

The future work of this project is to develop a complete system consisting of server-side components containing a trained model and an application for smart mobile devices with features such as displaying recognized diseases in plants, based on leaf images captured by the mobile phone camera. This application will serve as an aid to farmers, enabling fast and efficient recognition of plant diseases and facilitating the decision-making process when it comes to the use of chemical pesticides.

REFERENCES

[1]   Chopda, J., Raveshiya, H., Nakum, S. and Nakrani, V., 2018, January. Cotton Crop Disease Detection using Decision Tree Classifier. In 2018 International Conference on Smart City and Emerging Technology (ICSCET) (pp. 1-5). IEEE.

[2]   Kamble, J.K., 2018, February. Plant Disease Detector. In 2018 International Conference on Advances in Communication and Computing Technology (ICACCT) (pp. 97-101). IEEE.

[3]   Kulkarni, O., 2018, August. Crop Disease Detection Using Deep Learning. In 2018 Fourth International Conference on Computing Communication Control and Automation (ICCUBEA) (pp. 1-4). IEEE.

[4]   Mique Jr, E.L. and Palaoag, T.D., 2018, April. Rice Pest and Disease Detection Using Convolutional Neural Network. In Proceedings of the 2018 International Conference on Information Science and System (pp. 147-151). ACM.

[5]   Padol, P.B. and Yadav, A.A., 2016, June. SVM classifier based grape leaf disease detection. In 2016 Conference on advances in signal processing (CASP) (pp. 175-179). IEEE.

[6]   Petrellis, N., 2017, September. Mobile Application for Plant Disease Classification Based on Symptom Signatures. In Proceedings of the 21st Pan-Hellenic Conference on Informatics (p. 1). ACM.

[7]   Zeng, W., Li, M., Zhang, J., Chen, L., Fang, S. and Wang, J., 2018, October. High-Order Residual Convolutional Neural Network for Robust Crop Disease Recognition. In Proceedings of the 2nd International Conference on Computer Science and Application Engineering (p. 101). ACM.

[8] Singh, V. and Misra, A.K., 2015, March. Detection of unhealthy region of plant leaves using image processing and genetic algorithm. In 2015 International Conference on Advances in Computer Engineering and Applications (pp. 1028-1032). IEEE.

[9] Sardogan, M., Tuncer, A. and Ozen, Y., 2018, September. Plant leaf disease detection and classification based on CNN with LVQ algorithm. In 2018 3rd International Conference on Computer Science and Engineering (UBMK) (pp. 382-385). IEEE.

[10] Simonyan, K. and Zisserman, A., 2014. Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556.