Pest detection in Agricultural Plantation of Cotton crops using Convolutional Neural Network

Prof. Sandhya Potadar,Aakanksha Khare, Shalaka Buch, Aksha Khairmode
Department of electronics and telecommunication Engineering
MKSSS's Cummins College of Engineering for Women, Pune, Maharashtra

Abstract— **India is a diverse country with 70% of its people depending upon agriculture. With the ever-increasing need for food and shelter, there is a drastic increase in the need for clothing as well. For this, the production of cotton crops also needs to be increased. Pests and crop diseases play a vital role in reducing the yield. Identification and classification of crop pests are essential to aid farmers. Manual detection of pests and the use of economical friendly pesticides may result in a reduction of the destruction caused to crops. The extra use of pesticides may reduce yield and will pollute the air and the soil of that region. As Maharashtra is the Manchester of cotton production, the proposed system focuses on pest detection on cotton crops. To address this issue, a convolutional neural network architecture is being applied to the mixture of images collected from farmer's fields and open-source platforms. The data augmentation techniques are applied such as shear, rotation, brightening, translation, and zooming to avoid the network from overfitting. The CNN architecture learns and extracts high-level complex features in image classification applications instead of extracting handcrafted features by traditional methods. The results obtained from the proposed CNN model classifies the visuality of pests on cotton crops with a model accuracy of 91.54% and can be applied in the agriculture field.**

Keywords—Agriculture, Early pest detection, Image processing, CNN, Capsule Network.

I. INTRODUCTION

India is a diverse country with 70% of its people depending upon agriculture production. Production of crops in the agricultural sector is important for humans. With the ever-increasing need of food and shelter there is a drastic increase in need for clothing. For this production of crops should also be increased. The pests and crop diseases play a vital role in reducing the yield. Agriculture is the main source of livelihood and national income in developing countries and provides raw materials for industry. A nation's export depends on the agricultural sector. The biological parameters that affect agricultural production are the presence of pests in plants or plants. Pests are one of the leading causes of agricultural losses. Insects can be particularly harmful to plants as they feed on leaves, affect photosynthesis, and also act as vectors for several serious diseases. There are many chemical and biological methods. For pest control, but to achieve the maximum effectiveness of certain methods, careful monitoring of the entire property, which farmers can see with the naked eye, is generally recommended. They go about their daily activities. The main problem with this method is that by the time the infestation is detected, a lot of damage may already have been done. The early detection of pests is an important task for successful agriculture. Early pest detection requires a more systematic approach, especially on large farms and plantations. In contemporary global its miles dealt with as alternate. So, loss of manufacturing in the end impacts alternate in addition to the human society. In this evolving age of technology, use of era is completed everywhere and everywhere feasible. Agriculture is one of the sectors in which use of technology can be performed to grow productivity in addition to assist the farmers in gaining higher yields. To simplify this cycle of dwelling and contributing inside the discipline of agriculture and image processing we decided to enforce this venture. (machine vision machine -inspection of pest identity)

II. CONVOLUTIONAL NEURAL NETWORK

The algorithm used in this paper is Convolutional Neural Network.A Convolutional Neural Network (ConvNet/CNN) is a Deep Learning set of rules that can absorb an input picture, assign significance (learnable weights and biases) to numerous factors/items inside the image and is able to distinguish one from the other. It is a type of artificial neural community utilized in picture recognition and processing that is mainly designed to maintain pixel records and to carry out both generative and descriptive tasks. [1]The layers of CNN are as follows:

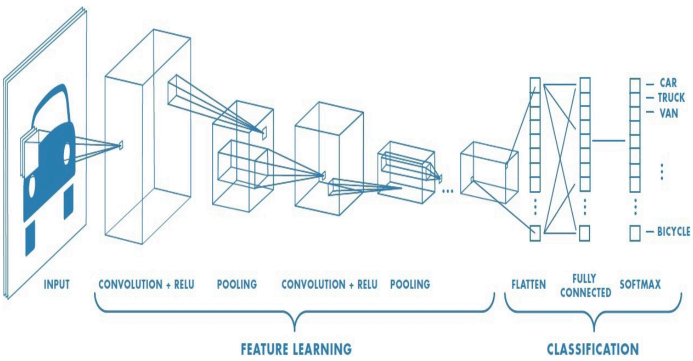

Fig 1: CNN Architecture

A. CONVOLUTIONAL LAYER

A “clear out”, sometimes called a “kernel”, is passed over the photograph, viewing some pixels at a time. The convolution operation is a dot product of the original pixel values with weights defined within the clear out. The results are summed up into one number that represents all of the pixels the filter observed. This layer applies fourteen 5x5 filters so as to be extracting 5x5-pixels from sub-areas. The first convolutional layer has sixty-four filters with the kernel length of 3 with the same padding. The identical padding has each the output tensor and input tensor have the identical width and peak of 256. TensorFlow will add zeros inside the rows and columns to ensure the same length.

B. ACTIVATION LAYER

The convolution layer generates a matrix, comparatively smaller in length than the original picture. This matrix is administered through an activation layer, which introduces non-linearity. The activation function is usually ReLu(Rectified linear unit).

C. POOLING LAYER

“Pooling” is the technique of down sampling and reducing the matrix size. A filter is exceeded over the consequences of the preceding layer and selects one range out of each institution of values (generally the maximum, this is referred to as max pooling). This lets in the network to train a whole lot faster, focusing on the most critical records in every photo feature. The pooling computation will reduce the extension of the statistics. We use the module max_pooling2d with a size of 2x2 and stride of two.

D. FLATTEN LAYER

In between the convolutional layer and the completely related layer, there's a ‘Flatten’ layer. Flattening transforms a -dimensional matrix of capabilities right into a vector that may be fed into a fully related neural community classifier.`Flatten` layer is a bit unusual in that the dimension of the matrix is larger than the dimension that may be used by a fully related classifier. Default fully related classifier is fully connected NxN with N neurons. Matrix has to have dimension at least equal to the number of neurons in the last layer of this network. Neurons in the last layer can connect to a lot of parameters because they already had to learn just a little bit about the data in order to classify it.

E. FULLY CONNECTED LAYER

A fully connected layer is a neural layer that combines the results of the convolutional layers to generate a prediction.In network based learning, the fully connected layer is an important part of neural networks because it allows us to apply back propagation. Back-propagation is a process of tuning the weights of multiple layers that are used to get input and predict output. By definition, fully connected means that all nodes in one layer connect to all nodes in the next layer.

III. LITERATURE REVIEW

In [1], the authors utilized a deep learning approach to determine the diseases and pests accurately on leaves and other parts of the crop. The proposed method comprises following steps: - the first step is image acquisition; second step is image preprocessing. After the dataset is ready it is given to the CNN model. In the next step, the feature extraction and classification are done using the transfer learning approach. The CNN with the transfer learning approach is used here, which is very effective in processing the large volume of data and gives higher accuracy.In [2], the authors focused on pest detection using Support Vector Machine. In this paper using image processing techniques an automatic pest identification system is proposed. Images of pests are accumulated and stored in a database. Based on the features extracted the Support Vector Machine is trained. Here color feature is used to train the Support Vector Machine to classify the leaf pixels and pest pixels. Morphological operations are used to remove the undesired elements in the classified image. The aim of this paper is to detect the pest and give the information about the number of pests found. The classification process here is done with the help of SVM.

In [3], the authors briefly sum up the progress achieved so far on the use of digital images and machine learning for effective pest monitoring and detection, thus providing a complete picture on the subject in a single source. It provides a detailed discussion on the research gaps and main weaknesses that still persist, with emphasis on technical aspects that obstruct practical adoption of the system. It proposes some possible directions for future research on this topic. In [4], they focused on early pest detection. Diseased plant images were acquired using scanners or cameras. Later the acquired image was processed to interpret the image contents by image processing methods. The segmentation algorithm alone cannot provide good quality output, it needed a pre-processing step. Preprocessing consisted of various steps like de-noising and image enhancement. Due to drawbacks and irregularities in pest images, it is mandatory to include the pre-processing step before segmentation process for quality and accurate output. The proposed system proved reliable for rapid detection of pests. The paper also introduced k-means clustering as an approach for segmentation.

In [5], the authors have proposed a technique for pest detection using the segmentation algorithm-means clustering in MATLAB. The prime focus was examination of small spots on plant leaves which are usually left undetermined by naked eye. Image acquisition and image preprocessing was carried out followed by K-means clustering which made the technique reliable in presence of extensive data. The proposed algorithm was successful in counting the pests on the leaves and displaying the infected area. In [6], an algorithm was proposed for automatic detection and estimation of Whitefly on cotton crops using image processing. The acquired leaf image was set through a set of procedures which included the steps color space conversion, background subtraction, thresholding followed by morphological operation. The technique proposed was successful in counting the number of whiteflies on the cotton leaf.

In [7], machine vision and image processing techniques were used in MATLAB for the detection of pests and disease on cotton crops. The color models RGB, HSI, YCbCr were implemented for extraction of damaged parts from the input cotton leaf image. The ratio of damage was chosen as the key aspect for determination of degree of damage caused by pests and diseases. Further the paper also discussed and compared the color models on the basis of the accuracy of the results obtained.

In [8], the authors explored the combination of different Image processing algorithms that can be used on a pre- processed image for plant disease detection. The paper discussed various segmentation techniques for partitioning of images along with the various feature extraction and classification techniques that can be implemented for extracting features from infected parts. Furthermore, the paper implored the usage of different ANN methods for classification of plant diseases.

IV. PROPOSED METHODOLOGY

The proposed workflow is as shown in Fig 2. First step is image acquisition followed by Image preprocessing, data augmentation and finally a traditional CNN model for classification. The steps of proposed work are mentioned below in detail:

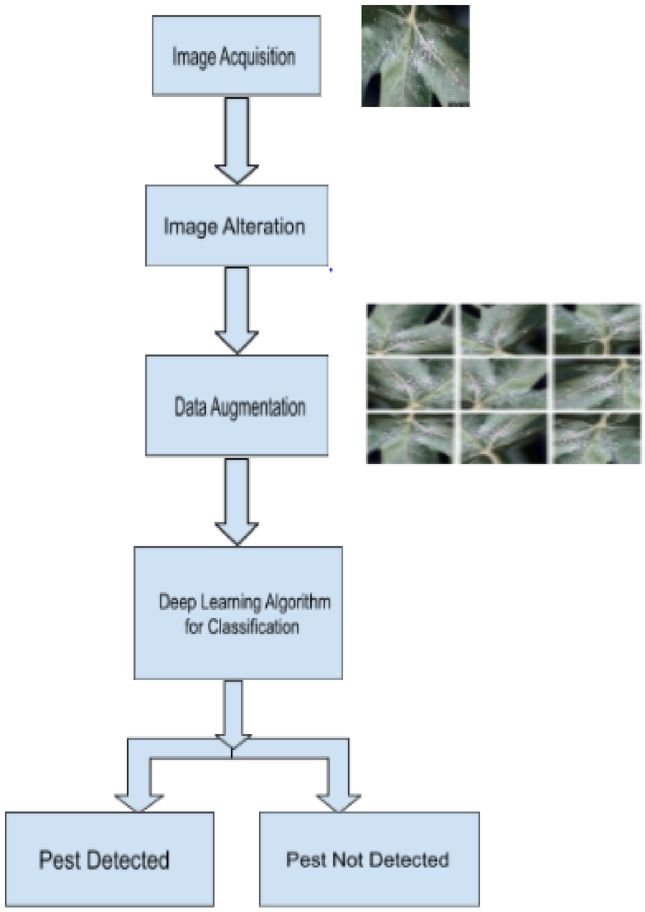

Fig 2: Basic flow of Pest detection

A. IMAGE ACQUISITION

Image acquisition or the gathering of images is done from many resources. The dataset so formed contains the real time images which were captured from a farm by a phone camera of 13 megapixels. Healthy plant images and diseased plant images both are required for training and testing. Keeping in consideration the need for varied pixel images, some images were taken from freely available open sources [9]. The collection of real time images collected from the field was a strenuous task as the cotton crops are seasonal.

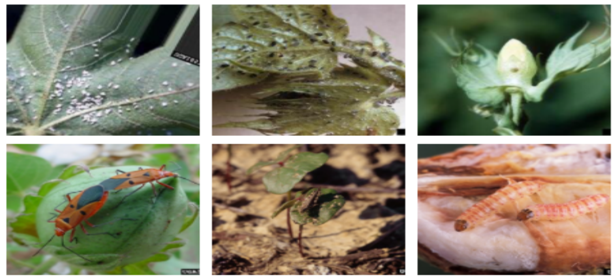

Fig.3: from top left a) Whitefly b) Aphid c) Boll-weevil second row from left d) Red cotton bug e) grasshopper f) Pink bollworm.

Whiteflies are sap- sucking bugs which typically lead to reduced plant vigor, upward curling of leaves, lint contamination with honey dew and related fungi. Cotton aphids, Aphis gossypii, are extraordinarily variable in length and shade, varying from mild yellow to darkish inexperienced or nearly black. The maximum common symptoms of damage consist of leaf crumpling and downward curling of leaves. Boll Weevil is a kind of a beetle that broadly speaking feeds on cotton buds, it's early degree assault signs may be diagnosed with the aid of small feeding punctures on the facet of the bud.
Red cotton bugs feed on developing and mature seeds, they generally stain the lint to ordinary yellow colour. Grasshoppers usually lead to partial or full defoliation of leaves while the crimson bollworm larvae burrow into cotton bolls to feed on the cotton seeds.

B. IMAGE ALTERATION

The images are collected from different sources, so the images have different size, format and resolution. For smooth operation of the CNN model, it needs to identify the pest images collected to be in the same format all over. To acquire more accurate and precise results the resolution and resizing of images is done. We resized all the images to 256 x 256 which is standard and always preferred.

C. DATA AUGMENTATION

The prediction accuracy of the CNN model hugely depends upon the diversity and the amount of dataset available during training. To attain this goal and ensure smooth running of CNN, data augmentation was performed. Execution of data augmentation led to an increase of dataset where initial images that were up to 1000 developed into a dataset of nearly 2000 images. The main operations performed on the images were, brightening, rotation, shearing, height shift, width shift, flipping, zooming and many more permutations which allowed to create a data set with good challenges.

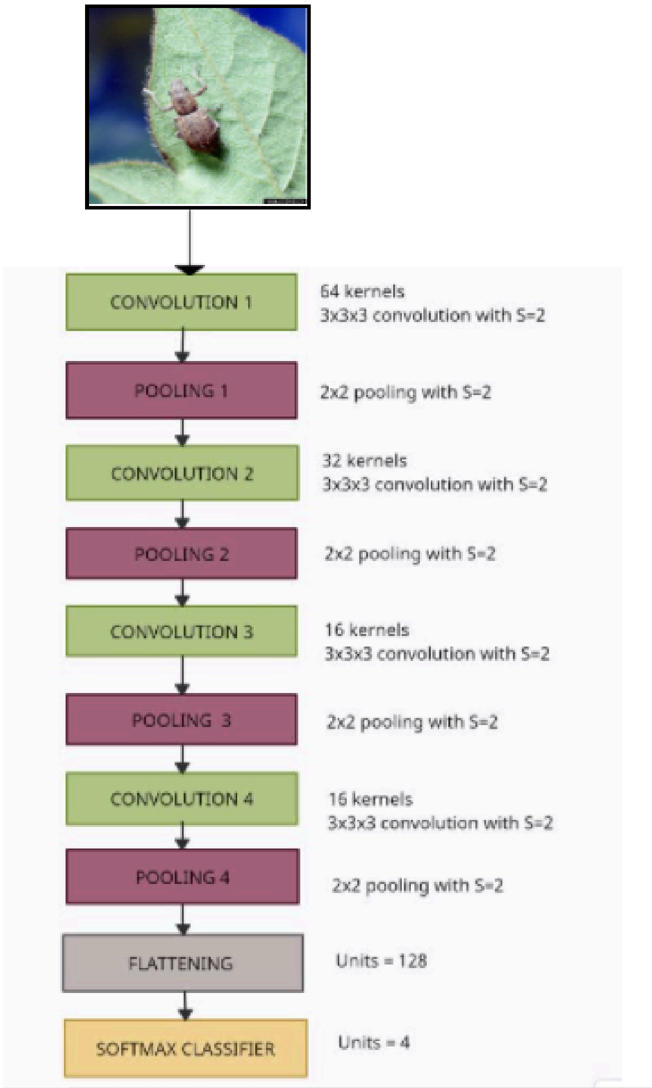

Fig. 4: Model Architecture

D. ALGORITHM FOR CLASSIFICATION

This stage starts with generating data from the location followed by target size, batching of images and categorizing it. Creating a CNN model from scratch with each line of code compiling convolution followed by pooling and then second convo-pooling which after takes us to flattening and finally the output layer of the model. The first convolutional layer has 64 filters with the kernel size of 3 with same padding. The same padding has both the output tensor and input tensor have the same width and height of 256. The activation used is 'ReLu'. The pooling computation will reduce the extension of the data. We can use the module max_pooling2d with a size of 2x2 and stride of 2. Flattening is done with units equal to 128. The model is trained at several different values of epochs. Epoch means training the model with training data for one cycle. The accuracy achieved after 20 such epochs is 91.54%. The accuracy increases with increase in the number of dataset images.

V. RESULTS

After compiling and running the code over the testing and validation data set, an accuracy of 91.5% was obtained after 20 epochs. The data set which was generated after augmentation created replicated images of some original images.

fig. 5(original)

fig.6(original)

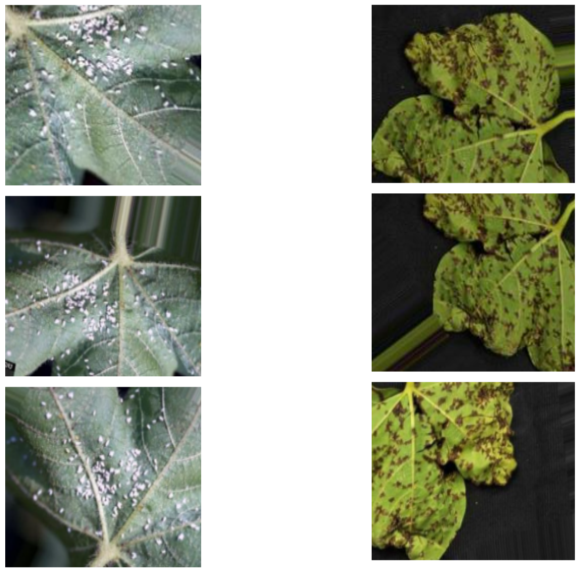

Fig. 7: Augmented output for both the original images

Metrics epoch	Accuracy	Loss
10	84%	0.30
20	91.54%	0.23

fig. 8 Training output

Here the accuracy calculations are done on testing data using the formula,

$$Accuracy = \frac{No\ of\ correctly\ predicted\ pest}{Total\ no\ of\ pests}$$

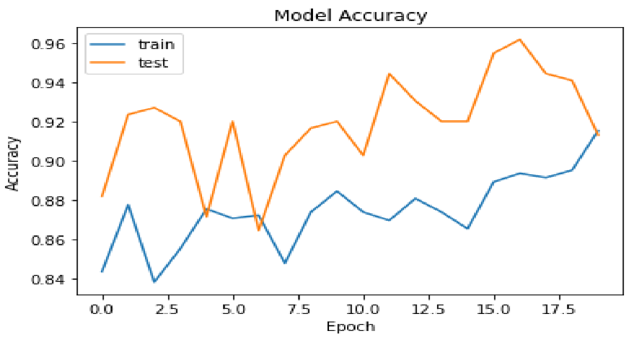

fig. 9 Model Accuracy Graph

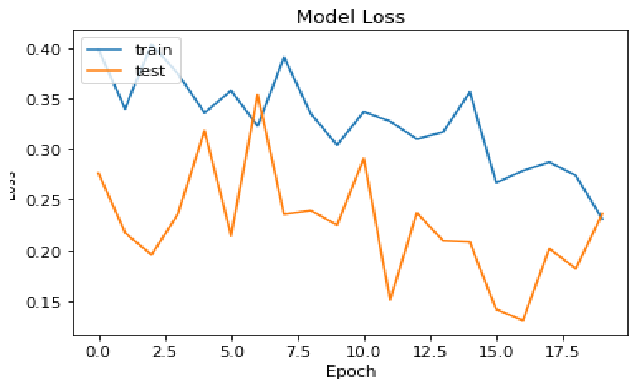

fig. 10 Model Loss Graph

VI. CONCLUSION

The proposed system is an advancement in the field of agriculture, which will be responsible for quick detection of the pest. Detection of pests at an early stage will help to prevent the crop damage. It will also aid in reducing the use of unwanted and not that needful pesticide and in return increase the natural value of soil. Implementation of Data augmentation results in a larger dataset in turn yielding a better accuracy after CNN implementation. The CNN model presented here gives an accuracy 91.54%. Capsule Network can be implemented instead of CNN to eliminate the need of a larger dataset, avoid loss due to max pooling and also address.

VII. REFERENCES

[1]. Pruthvi P. Patel, Dinesh Kumar B. Vaghela "Crop Diseases and Pests Detection Using Convolutional Neural Network",2019 IEEE ,978-1-5386-8158-9/19

[2]. Preetha Rajan, Radhakrishnan. B, Dr.L. Padma Suresh "Detection and Classification of Pests from Crop Images Using Support Vector Machine" ,2016 IEEE DOI: 10.1109/ICETT.2016.7873750

[3]. Jayme Garcia Arnal Barbedo,,"Detecting and Classifying Pests in Crops Using Proximal Images and Machine Learning: A Review" ,2020 Embrapa Agricultural Informatics, Campinas, SP 13083-886, Brazil ,AI 2020, 1, 312–328; doi:10.3390/ai1020021

[4]. Murali Krishnan,Jabert.G ," Pest Control in Agricultural Plantations Using Image Processing",2013 IOSR Journal of Electronics and Communication Engineering (IOSR-JECE),e-ISSN: 2278-2834,p- ISSN: 2278-8735.Volume 6, Issue 4(May. - Jun. 2013).

[5].K. Ashish Reddy, N. V. Megha Chandra Reddy and Sujatha. S , " Precision Method for Pest Detection in Plants using the Clustering Algorithm in Image Processing",International Conference on Communication and Signal Processing, July 28 - 30, 2020, India.

[6]. Monica N. Jige ,Varsha R. Ratnaparkhe, " Population Estimation of Whitefly for Cotton Plant Using Image Processing Approach", 2017 2nd IEEE International Conference On Recent Trends in Electronics Information & Communication Technology (RTEICT), May 19-20, 2017, India.

[7]. Qinghai He, Benxue Ma, Duanyang Qu, Qiang Zhang,Xinmin Hou,Jing Zhao , " Cotton Pests and Diseases Detection based on Image Processing" , TELKOMNIKA, Vol. 11, No. 6, June 2013, pp. 3445 ~ 3450 e-ISSN: 2087-278X

[8]Sachin D. Khirade,A. B. Patil, "Plant Disease Detection Using Image Processing" , 2015 International Conference on Computing Communication Control and Automation,DOI 10.1109/ICCUBEA.2015.153.

[9]. www.ipmimages.org