ISSN No:-2456-2165

Crop Disease Detection using CNN

Ajay Rasave[1]
PG. Student,
ME Artificial Intelligence,
Department of Computer Engineering,
KJ Somaiya Institute of Engineering and Information Technology, Sion, Mumbai

Dr. Sunita R. Patil[2]
Professor, Department of Computer Engineering,
KJ Somaiya Institute of Engineering and Information Technology, Sion, Mumbai

Abstract:- Good health is a first priority of all human being. And to maintain health the first need is a healthy food, which we get from our agriculture resources. According to the United Nations' World Population Prospects (WPP), India is projected to surpass China as the world's most populous country in 2023. And to maintain the appetite and health of such huge population it's very important for agriculture field to grow the more and healthy crops. Now a days there are number of diseases which are causing crop underproduction. Delay in identifying the actual disease will result into more loss and cost of remedy. So, it's very important to have the knowledge of diseases for early detection. In today's date Artificial intelligence technologies such as Classification algorithms using Deep Learning are so efficient and are capable of predicting the disease based on the images of the crops. Hence, in this paper have reviewed the most efficient classification algorithms and proposed few most suitable models for image-based disease detection.

Keywords:- *CNN, Inception V3, VGG16, VGG19, ResNet18.*

I. INTRODUCTION

India is known as a Krishi Pradhan i.e.an agrarian country. Economy is majorly dependent on agriculture sector. Farmers are tending towards newer ways of farming. They are now trying to use the new technologies for better crop production. In recent days there are many new and old types of diseases which are affecting crop production. One of main reason is that wrong diagnosis of disease which results into wrong treatment and wastage of money. Many times, it becomes difficult to identify unknown diseases with traditional ways. So, it's very crucial to use new and smart ways of disease detection for early and perfect remedy.

Now a days Computer Vision technology are so efficient that they can play a vital role in identifying diseases on crop with just an image. Classification algorithms such as Convolution Neural Network (CNN), K-Means, Residual Neural Networks, etc. are famous for classifying different kinds of images, based on different parameters such as color, shape, etc.

Convolution Neural Networks have shown their performance and accuracy in image classification activities. So, in this paper we are going to consider different types of CNN algorithms.

A. Convolution Neural Networks:

CNN which also referred as Convolution Networks (ConvNet) is one of the Deep Learning algorithms. It isconsisting of different layers such as pooling Layers and convolution layers and Fully Connected layersas shown in Fig.1. The term Convolution in CNN stands for a mathematical operation [1]. Mathematically Convolution can be defined as an operation on two functions which gives an output as a new function which is inspired by the previous functions.

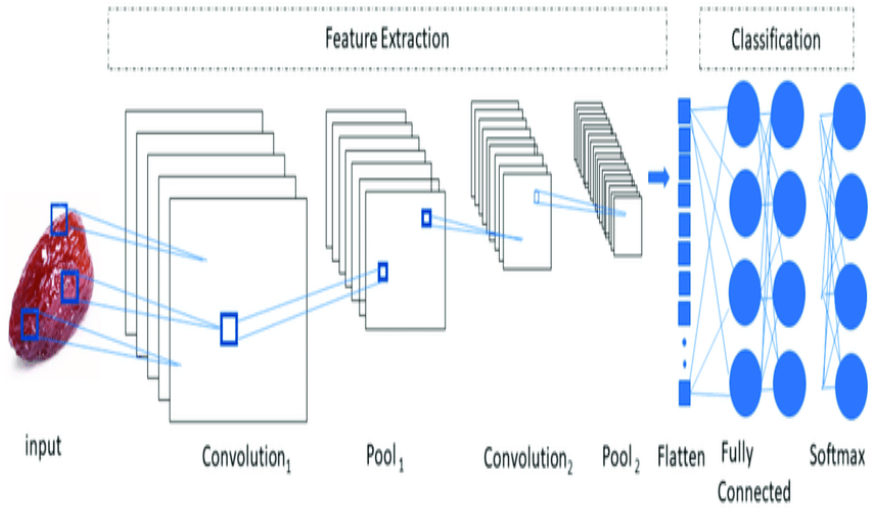

Fig. 1: Convolutional Neural Network [2]

ISSN No:-2456-2165

CNN has two main sections, first Feature Extraction model which is combined of Convolution Layers and Pooling Layers. And the second one is Classification model which mainly consist of Fully Connected Layers.

B. Feature Extraction Model

In this part the network tries to learn detecting the features from the input data. Here the convolution and pooling are sequenced as shown in Fig.1.

➢ *Convolution*

This process is mainly responsible for applying different learnable filters and extract the features from the given image. Filters are applied on the image pixels and output given is nothing but 2-D map i.e. a feature map. This convolution is be formulated as equation 1.

$$M_i = b_i + \sum_k W_{ik} \star X_k \quad (1)$$

Where,

M_i = feature map
$\star$ = convolution operator,
X_k = k th input channel,
W_{ik} = sub kernel of that channel,
bi = bias term

Now, the ReLu function is applied on each convolution to make it non-Linear.

➢ *Pooling:*

This is a Max pooling layer which responsible for reducing the feature map size. It reduces the convolution map size by sub sampling layer.

C. Classification model

The output given buy feature extraction layer, which is nothing, but a feature map is processed in this model in such a way that each neuron will be connected to all the features from this feature map i.e. in a fully connected networks manner. In this classification process SoftMax activation function is used. And it can be formulated as in equation 2, given below.

$$y_c = \varsigma(z)_c = \frac{e^{z_c}}{\sum_{d=1}^{C} e^{z_d}} \text{ for c = 1}\cdots\text{C} \quad (2)$$

Where,

ς is Soft Max function

input a C-dimensional vector z outputs a C-dimensional vector y of real values between 0 and 1.[3]

II. LITERATURE SURVEY

Due to huge scope in agriculture domain and advance commuting technologies there are lots of research done on crop diseases detection with deep learning. In this section we have done the literature survey on recent research done for crop diseases classification with different machine learning algorithms.

In the research [4], Sachin B. Jadhav, Vishwanath R. Udupi and Sanjay B. Patilproposed a crop disease detection mechanism using Convolutional Neural Network (CNN) classification algorithms namely AlexNet and GoogleNet. They trained and compared these two models on around 649 unhealthy images containing diseases like Bacterial blight , Brown spot and Frogeye leaf spot. and 550 healthy images of soyabean crop. They observed that AlexNet performed more accurately over GoogleNet. AlexNet gave accuracy around 98% whereas GoogleNet achieved accuracy around 96%.

R. Hu et al. in [5] proposed their own CNN model for corn leaf disease identification model based on the data augmentation using GoogleNet. And compared the results with other pre trained algorithms like ResNet18, VGG16, and VGG19. They achieved the accuracy of around 97.6% with their own model, which was more than compared with others. They used around 4354 corn leaf images from Plant Village datastore containing 4 kinds of diseases. They first trained their own model with proper fine tuning. Then for same dataset they used transfer learning models of VGG16, VGG19, ResNet18. And compared the accuracies, performances, precisions, Recall, F1 scores.

In the research [6] proposed by Selvaraj et al. they proposed a system to detect banana leaf disease and pets based on ResNet50 and Inception v2. They achieved the accuracy of around 90% which was comparatively better than MobileNetV1. They collected banana leaf images from Africa and south part of India having around 18 different kinds of diseases.

As per the research [7] proposed by Godliver Owomugisha et al., Linear SVC classifier gave the accuracy of around 99% for the ORB features. In this research they implemented a system to measure the severity of a disease on the leaf of cassava plants. They used three classifiers namely Linear SVC, KNN and Extra Trees to extract different features from the diseased images. They trained the modules on around 7386 images containing healthy and 4 different classes of diseases. And the severity was measured in a range from 1 to 5. Where severity 1 representing healthy image and severity 5 as unhealthy.

ISSN No:-2456-2165

III. THE PROPOSED MECHANISM

Considering the similar research works done in image classification, Convolution Neural Network i.e. CNN has proved its capability, performance and accuracy, which is comparatively far better than other algorithms. And hence our proposed system is based on the CNN classifiers. Proposed solution is as shown in Fig. 2. In order to make our system fast we use transfer learning mechanism i.e. Pretrained CNN modules will be used. That will save the time required to train initial common steps.

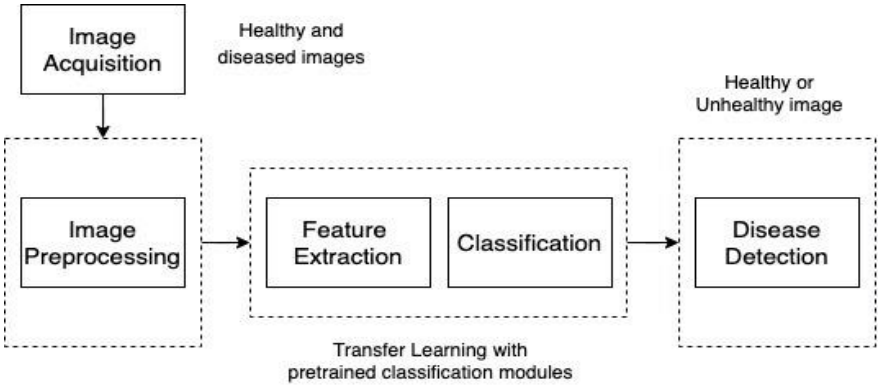

Fig. 2: Proposed System

Proposed system is composed of four commonly used steps in CNN modules. First step is image acquisition which is a image capturing layer. This step is responsible to avail image data ready for further processing. Second stage is Image Pre-processing, where images will be filtered, resized as per the requirement. Third one is Feature Extraction which find out the different features from images required for the classification. The fourth step is classification stage where images will be classified based on the features extracted in previous step. And a final fifth step which tells us about health of crop.

Finally, complete content and organizational editing before formatting. Please take note of the following items when proofreading spelling and grammar:

A. Image Acquisition

Image Acquisition [8] is the first step of the proposed system. The image data collected from any sources can be consist of images of different size, lightening, background. Or it can be impacted because of any corrupt format. So, this stage will make sure that image data provided to the next step is in appropriate format.

We are going to use image data from Plant Village data source. Our dataset is consisting of around 6354 images of Tomato Leaf containing different kinds of disease such as Bacterial Spot, Late Blight, Leaf Mold, Septoria Leaf Spot, Yellow Leaf Curl Virus, Mosaic Virus, etc.

B. Image Pre-processing:

In deep learning Image pre-processing [9] is one of the important steps. It also referred as the data cleaning. The aim of the image pre-processing is to modify data in such a way that it will be more efficient and suitable for specific application. Modification here means that reduce any distortion in images, adjust colour, brightness, or improve any specific feature which will help to improve the performance of the application. To do this we can different kinds of image pre-processing techniques such as Image Filtering and Segmentation, Pixel brightness corrections or transformations, Fourier transform and Image restauration, Geometric Transformations, etc.

Images from plant Village datastore are collected from many kinds of diseases. Images for same plant for same disease are captured in multiple ways like different angle/orientation [10]. It can consist of images of different resolutions, quality, brightness. So, it's crucial to pre-process images in order to make them in proper size, brightness, signal level, etc. In the proposed system Images will be resized to 256 × 256 pixels.

C. Feature Extraction:

The output of a deep learning-based prediction system is based on the number of features/factors it has. Having more numbers of features will make it even harder or complex to predict result [11]. Feature extraction layer is responsible to produce new features from existing one. These new features are reduced features, and they represent the same information from previous features.

As our proposed system uses the Transfer Learning mechanism, we will use pretrained models which will save the time required for few time-consuming steps. Here the proposed CNN module needs to be fine-tuned in order to get highest accuracy.

D. Classification:

This is the main step of the CNN module. This stage is responsible to predict out the most suitable class for input image by comparing the features captured by previous layer. Or it gives the probability that the input belongs to particular class [12].

E. Disease Detection:

This is the final stage of the proposed system. It takes output of classification as the input. This stage in general referred image post-processing layer. In this stage we can get further information regarding the health of the crop such as if leaf is diseased then how much portion is infected or which area is infected. Also this layer can further help us to suggest remedies for respective diseases.

ISSN No:-2456-2165

IV. PERFORMANCE EVALUATION

We are going to use accuracy, precision, recall, and F1-score as performance evaluation metrics [13].

A. Accuracy

Accuracy is one metric for evaluating classification models. Informally, accuracy in percentage is the percentage of correct predictions over the total number of predictions done by model. Formally, accuracy has the following definition:

$$Accuracy = \frac{Number\ of\ correct\ predictions}{Total\ numbers\ of\ predictions} \quad (3)$$

For binary classification, accuracy can also be calculated in terms of positives and negatives predictions as follows:

$$Accuracy = \frac{TP+TN}{TP+TN+FP+NP} \quad (4)$$

Where,

TN – True Positive
TN – True Negative
FP – False Positive
FN – False Negative

B. Precision:

Precision is nothing but the fraction of Positives over total positives i.e. False positives and True Positives. It can be formulated as,

$$\text{Precision} = \frac{[\text{True Positives}]}{[\text{True Positives} + \text{False Positives}]} \quad (5)$$

Recall:

Recall is portion of correctly identified positives. It is formulated as,

$$\text{Recall} = \frac{[\text{True Positives}]}{[\text{True Positives} + \text{False Negatives}]} \quad (6)$$

C. F1 Score:

F1 Score is useful to get measure of wrongly classified cases compare to Accuracy Metric. It can be formulated as below.

$$F1\ Score = 2 * \frac{\text{Recall} * \text{Precision}}{\text{Recall} + \text{Precision}} \quad (7)$$

V. CONCLUSION

Due to huge variety of diseases with huge variety of symptoms its usually difficult for farmers to identify the correct disease on the crop in order to take correct remedy action. Which further lead to more production loss and the wasted cost on wrong remedy because of wrong diagnosis. Hence it is very crucial to find out solution which can identify crop disease with fast and easier way. Artificial Intelligence with variety of technology can help to solve this problem with efficient ways. Deep Learning modules such as Inception V3, VGG, ResNet, etc. are very popular deep learning algorithms in image classification field. So, our proposed system will be implemented with these algorithms. In the proposed approach we will use transfer learning mechanism in order to improve performance of the system. The dataset is consist of 6354 tomato leaf images containing six types of diseases. These algorithms will be trained on same set of data. And the performance of each module will be compared with the other one. Performance will be evaluated based on Accuracy, Precision, Recall and F1 score etc.

REFERENCES

[1.] Ian Goodfellow, Yoshua Bengio and Aaron Courville. "Deep Learning", MIT Press, p. 326, 2016.

[2.] Pérez, Dalila & García-Vázquez, Juan & Salomon-Torres, Ricardo. (2021). Evaluation of Convolutional Neural Networks' Hyperparameters with Transfer Learning to Determine Sorting of Ripe Medjool Dates. Agriculture. 11. 1-12. 10.3390/agriculture11020115.

[3.] Jihen Amara, Bassem Bouaziz, and Alsayed Algergawy, "A Deep Learning-based Approach for Banana Leaf Diseases Classification", BTW 2017–Workshopband, pp. 83-85, 2017

[4.] Sachin B. Jadhav, Vishwanath R. Udupi, Sanjay B. Patil, "Identification of plant diseases using convolutional neural networks", International Journal of Information Technology, 2020.

[5.] Rongjie Hu, Shan Zhang, Peng Wang, Guoming Xu, Daoyong Wang and Yuqi Qian. 2020. The identification of corn leaf diseases based on transfer learning and data augmentation. In Proceedings of 2020 3rd International Conference on Computer Science and Software Engineering (CSSE'20). Beijing, China, 8 pages. https://doi.org/10.1145/3403746.3403905

[6.] M. G. Selvaraj, A. Vergara, H. Ruiz, N. Safari, S. Elayabalan, W. Ocimati, and G. Blomme, "AI-powered banana diseases and pest detection," *Plant Methods*, vol. 15, no. 1, pp. 1–11,2019

[7.] Godliver Owomugisha, Ernest Mwebaze, "Machine Learning for Plant Disease Incidence and Severity Measurements from Leaf Images," 2016 15th IEEE International Conference on Machine Learning and Applications, 2016 IEEE

[8.] Ivan Zhou. "Designing Image Acquisition for Machine Vision." LANDING AI. https://landing.ai/designing-image-acquisition-for-machine-vision/#open-overlay (accessed Mar. 30, 2023).

[9.] Great Learning Team. "Introduction to Image Pre-processing | What is Image Pre-processing?." Great Learning. https://www.mygreatlearning.com/blog/introduction-to-image-pre-processing/ (accessed Mar. 30, 2023).

[10.] Mohanty SP, Hughes DP and Salathé M (2016), "Using Deep Learning for Image-Based Plant Disease Detection,". Front. Plant Sci. 7:1419, Volume 7 – 2016

[11.] Dishaa Agarwal. "Guide For Feature Extraction Techniques." Analytics Vidhya. https://www.analyticsvidhya.com/blog/2021/04/guide

-for-feature-extraction-techniques/ (accessed Mar. 30, 2023).
[12.] Anne Bonner. “The Complete Beginner’s Guide to Deep Learning: Convolutional Neural Networks and Image Classification.” Towards Data Science. https://towardsdatascience.com/wtf-is-image-classification-8e78a8235acb (accessed Mar. 30, 2023).
[13.] Koo Ping Shung. “Accuracy, Precision, Recall or F1?.” https://towardsdatascience.com/accuracy-precision-recall-or-f1-331fb37c5cb9 (accessed Mar. 30, 2023).