International Journal for Rapid Research in Engineering Technology & Applied Science
Vol VIII Issue XI November 2022
ISSN Online:- (2455-4723)

Plant Disease Detection System

Amrita Bharani*,Apoorva Tripathi**, AishlyManglani***, Anupama Sahu****
Pir Mohammad*****
Department of computer science ,Acropolis Institute of technology and research, Indore,Madhya Pradesh , India*
bharaniamrita2001@gmail.com,tripathiapoorva24@gmail.com**,aishlymanglanics19@acropolis.in***, anupamasahucs19@acropolis.in****,pirmohammadcs19@acropolis.in*****

Abstract—One of the most important and tedious tasks in agriculture practices is detection of diseases on crops.It is a noteworthy risk to the growth and quality of the crop and affects the yield of crops. Early detection of these diseases still proves to be troublesome.The recent expansion of deep learning methods has found its application in plant disease detection, offering a robust tool with highly accurate results. This is an efficient step towards sustaining the crop and increasing the yield and thereby giving good profit to the farmers.The main aim of the proposed work is to find a solution to the problem of 38 different classes of plant diseases detection using the simplest approach while making use of minimal computing resources to achieve better results compared to the traditional models. VGG16 training model is deployed for detection and classification of plant diseases. Neural network models employ automatic feature extraction to aid in the classification of the input image into respective disease classes.

Keywords—Deep learning,Convulational neural network

1. INTRODUCTION

In India the major population relies on agriculture as their source of income.India ranks second globally in terms of farm yields. It was reported in the year 2018 that agriculture opened the doors of employment for more than 50% of the employees, hence contributing to 18–20% to country's GDP. India has thus proven to be one of the leading nations in terms of agricultural yield and productivity.It becomes very crucial to recognize the problems faced in this sector. There are several challenges faced by the farmers which act as a barrier to their income. The major one is the losses in yield caused by crop diseases. It's hard to observe and take care of plant diseases manually and needs a good amount of effort and a good expertise in plant diseases which is a barrier to farmers as most of them are illiterate and don't have the adequate knowledge for those diseases.

These diseases go unnoticed which damages the crop and farmers have to bear the loss of this crop failure.A solution which can provide early detection of these diseases can help in maintaining the crop and saving the crop from damage. Image processing with machine learning models is an effective solution to detect plant diseases.

Various Laboratory based approaches such as polymerase chain reaction, gas chromatography, mass spectrometry, thermography and hyper spectral techniques have been employed for disease identification. However, these techniques are not cost effective and are time consuming.These techniques are done by visual inspection or some chemical processes by experts. For doing so, a large team of experts as well as continuous observation of plants is needed, which costs high when we do with large farms.

Automatic detection of the diseases by simply seeing the symptoms on the plant leaves makes it easier as well as cheaper. The proposed solution for plant disease detection is computationally less expensive and requires less time for prediction.

CNN can sustain applications in agriculture, such as identifying diseases and quantifying areas affected. Diseases are usually identified by a specialist by mere visual observation. This method requires a lot of time on large farms and land. The use of convolutional neural networks for early stage plant disease detection and detection is effective in improving product quality.

Developing such an accurate image classifier for the diagnosis of plant diseases requires a large processed and validated dataset containing a variety of images of diseased and healthy plants. Is required. Project "Plant Village" has collected thousands of plant images and made them available for free use. The dataset has already been prepared and is available in three versions of his color, grayscale and segmented.

The proposed system of recognition of plant leaf diseases focuses on 14 varieties of plants which include apple, blueberry, cherry, corn, grape, orange, peach, pepper, potato, raspberry, soybean, squash, strawberry and tomato. This system is built on concepts of Deep Learning-Convolutional neural networks (CNN) is posted for formation of a statistical model that is executed on the input image and transforms the input to classify output tags.

2. LITERATURE REVIEW

The motivation of using machine learning in this project is explained below.

The manual process of identifying plant diseases (visual inspection) is often time consuming, labor intensive, expensive and subjective. This is one of the main reasons researchers are looking for alternative methods. Various ML approaches have been proposed to address this problem with a good accuracy score, reducing cost and subjectivity.

Early stage detection of plant diseases is one of the major concerns in the agricultural sector. Therefore, many of the researchers are working on the problem of detection and diagnosis of plant diseases. Some research results for detecting plant diseases are shown below.

Deep Learning for Image-Based Plant Detection" [1], the authors PrasannaMohanty et al.proposed an approach to detect plant diseases by training a convolutional neural network. A CNN model has been trained to identify healthy and diseased plants from 14 species. This model achieved 99.35% accuracy on the test set data. Using the model with images taken from trusted online sources, the model achieves an accuracy of 31.4%, which is better than a simple random-choice model, but with a more diverse set of training data. helps improve accuracy. Also, several other variations of model or neural network training can provide greater accuracy and pave the way for making plant disease detection more accessible to everyone.

In the publication,"Convolutional Neural Network-Based Starting v3 Models for Animal Classification" [2], Jyotsna Bankar et al. proposed using the Inception v3 model to classify animals into different species. Inception v3 can be used for object classification.This feature of Inception v3 makes it an effective tool for various image classifiers.

D.Tiwari et al. Al.[3]used transfer learning and various pretrained models on the potato leaf image dataset comparing 92% accuracy for neural network backpropagation and 95% accuracy for support vector machines and concluded that VGG 19 provides the best accuracy of 97.8%.

Wang-Su[4]'swork uses the original GoogleNet and a modified version of GoogleNet to classify 3767 images into 8 classes. These images were taken from the Flavia data set [8]. The modified version of GoogleNet contains twice as many Inception modules compared to the standard GoogleNet. When comparing two networks based on accuracy and performance. A modified version of GoogleNet performed slightly better than its counterpart.

[5] **According to S. Arivazhagan,** the disease detection cycle involves four basic stages: is detected, exposed bystanders are tracked by the fragmentation process, and valuable fragments are tracked. Surface findings are registered for retrieval. Finally, a classifier was used on the items classified to rank the diseases.

Prof. Sanjay, B. Dhaygude& et al… [6] The application of texture statistics to detect plant leaf disease is described. First, the color transform structure RGB is converted to HSV space, since HSV is a good color descriptor. Mask and remove green pixels with a precomputed threshold. Then, in the next step, segmentation is performed using a patch size of 32x32 to get useful segments. These segments are used for texture analysis with a color co-occurrence matrix. Finally, the texture parameters are compared with those of the normal leaf texture.

Bhumika S. Prajapati, Vipul K. Dabhi& et al… [7] In this study, the detection and classification of cotton leaf disease was performed using image processing and machine learning techniques. Background study removal and segmentation techniques were also described. Through this research, I came to the conclusion that a color space conversion from RGB to HSV is effective in removing the background .We also found that the threshold technique performed well compared to other background-removal techniques.Performed the color segmentation by masking green pixels in the image with no background and thresholding the resulting masked image to obtain binary images. This helps extract accurate disease characteristics. SVM was found to perform well in terms of disease classification accuracy. The proposed work has five main steps, three of which are implemented by:image acquisition, image preprocessing, and image segmentation

3. SOFTWARE REQUIREMENTS

3.1 SOFTWARE

- OS Version: Windows 10 64-bit
- IDE Used: Google Colaboratory
- Language: Python

4. METHODOLOGY

- Dataset Classification
- Building the CNN using transfer learning
- Training our network

- Testing

4.1 DATA CLASSIFICATION

Choosing an appropriate set of images to train a model is an important task.

To obtain the selected images, the centroid of each image is calculated. Centroids can be calculated using contours.

A contour is a curve that connects all points along the perimeter of a shape. Contours can be recognized very accurately on binary images. Therefore, each thresholded image must be converted to gray levels. You can use the Find Contours function for this purpose. The three arguments provided to this function are the source image, contour search mode, and contour approximation method. The output of the function contains images, contours and hierarchies. The output will contain all contours in the image. Each contour is an array of (x,y) coordinates of boundary points. A contour approximation method is used to specify the coordinates to be stored. CHAIN_APPROX_NONE saves all boundary points. However, not all edge points are required. You don't need all points to find a straight line contour, so just 2 points are enough. CHAIN_APPROX_SIMPLE provides this type of output by removing all redundant points and compressing the contours.

Image instants are calculated after the contour is found. Image moments are used to calculate the centroid or centroid of an object.The function cv2.Moments returns a dictionary of all moment values. From these moments, we can extract data such as centroid,area, etc.As we only need centroid of the image, it is given by the relations,

Cx = (M[―m10‖]/M[―m00‖]) and

Cy = (M["m01"] / M["m00"])

here, M is the dictionary of moments.

After calculating the centroid of each image using the method above, after ignoring all centroids, set a specific range of (x,y) coordinates. Images within the range are selected for further processing.

4.2 Building CNN using transfer learning

The advent of convolutional neural networks has made it possible to identify images.However, designing a CNN that identifies and divides objects into different classes is a complex task. Transfer learning makes it easy. For transfer learning, a 12GB TESLA k80 GPU was used to train models trained on the Plant Village dataset. Transfer learning also significantly reduces training time and provides much better performance for relatively small datasets.

Google publishes related models on the official TensorFlow website. 'Inception v3' is one of these models, trained on the ImageNet dataset and able to identify 1000 of classes such as TV, keyboard, car, and animal. This is one of the most widely used models for the classification of images [9]. The Inception v3 network is 48 layers deep and the input image size is 299 x 299. The network takes an image as input and gives labels as output. A feature of Inception v3 is factorization. The purpose of factoring the convolution is to reduce the parameters and connections while maintaining network efficiency.

The Inception Series of Convolutional Neural Networks is a series of neural networks that cannot be ignored in the history of convolutional neural networks. Before the advent of the Inception neural network, most neural networks only increased the depth of the network by increasing the number of convolutional layers to improve performance. Inception Neural Networks has changed this strategy. The Inception module proposed by Inception Neural Network uses different filter sizes and max pooling to reduce the dimensionality of the data. This has the advantage of obtaining much richer functionality with significantly reduced computation and fewer parameters.

a. Factor Convolutions

Factor Convolutions greatly reduce the number of pins and parameters without affecting system efficiency.factorization can be done with smaller convolutions.

b. Replace the 5 x 5 fold with two 3 x 3 folds. Or an asymmetric fold, such as a 3 x1 fold followed by a 1 x 3, replaces the 3 x 3 fold.BC Auxiliary Classifier Inception-v3 uses an auxiliary classifier as a normalizer. Batch normalization, introduced in Inception v2, is also used in helper classifiers.

c. Efficient Grid Size Reduction

Feature maps are usually reduced in size by max pooling. But this approach is either too greedy or too expensive. In Inception v3, max pooling gets 320 feature maps, which are concatenated to get 640 feature maps. Efficient grid size reduction in Inception v3 creates a cheap and efficient network.

4.3 MODEL TRAINING

A deep convolutional model can be used to classify labels specific to the task at hand. loads the Inception v3 model. A new class to detect is specified and the Inception v3 model is trained with different batches for the specified number of epochs, allowing to classify diseased plants using his Inception v3's image classification capabilities

Retraining scripts are the key vehicle for generating custom image classifiers in Inception v3. Train a new layer that performs the task of classifying custom classes. The script can be modified for parameters such as image_dir, middle_output_graphs_dir,output_graph,output_labels, distortion feature, number of training steps (epochs), learning rate. The folder is provided with the retrain script. Some of these images are kept for testing purposes.

The Inception v3 model is iteratively trained in different batches for the specified number of epochs. Each disease label is provided to the network along with the image

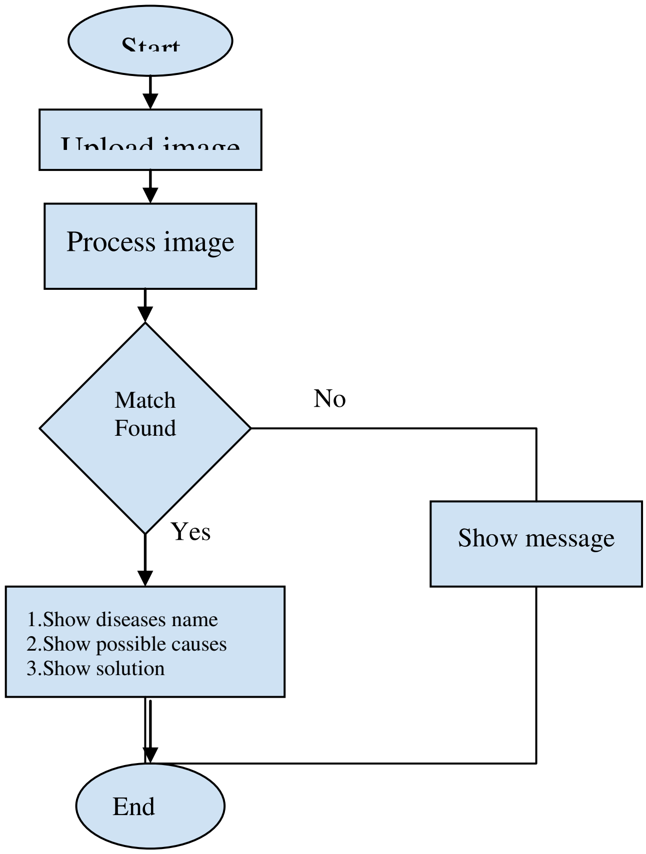
International Journal for Rapid Research in Engineering Technology & Applied Science
Vol VIII Issue XI November 2022
ISSN Online:- (2455-4723)

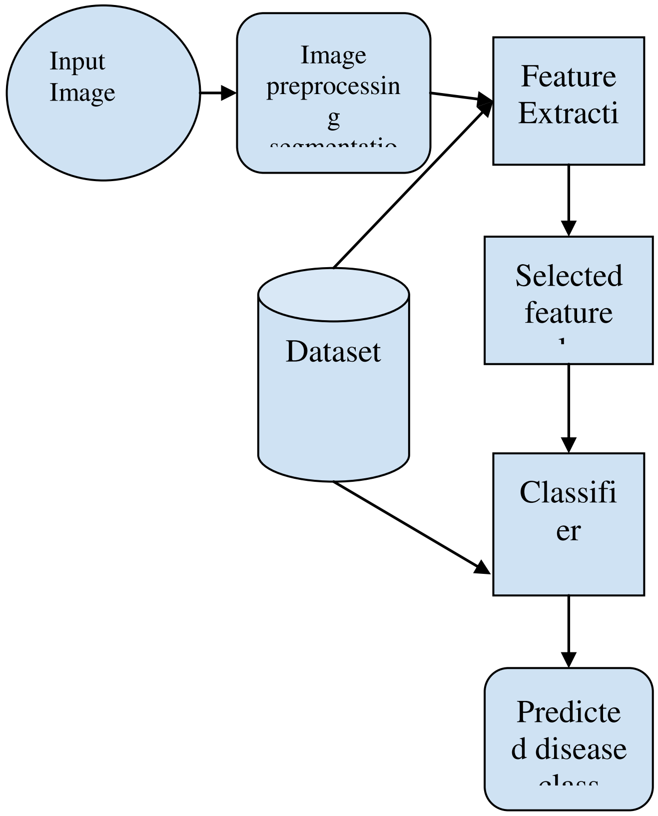
associated with the label. A separate set of images is provided for testing. Callback function is used to retrieve statistics during model training. Parameters such as loss, validation loss , and precision are retrieved in the callback. Use these values to monitor the model's performance over different epochs. Callbacks allow you to interact with model during model training. Callbacks can also be used to inhibit training after reaching a certain level of desired efficiency to prevent overfitting of the model.

4.4 TESTING

A trained model is tested on a set of images. A random image is introduced into his
network and the first label is compared to the original known label of the image. The parameters used for evaluation are F1 score, precision and recall. Accuracy is the percentage of true positive outcomes out of predicted positive outcomes. Recall indicates the proportion of true positive results that are correctly classified
F1 scores help maintain a balance between precision and recall.

4.5 FLOWCHART





4.5.5 OUTPUT

By using feature extraction Count vectorization, TF-IDF,our proposed model gets accuracy score of 0.93 with alpha set to 0.6

5.CONCLUSION

Plant diseases have been a significant concern in agriculture for years. Precision agriculture has enabled early disease detection and the minimization of losses through optimal decisions based on the results of DL methods.This paper proposes a CNN based method for plant disease classification using the leaves of diseased plants. Building such a neural network with high efficiency is a complex task. Transfer learning can be employed to achieve greater efficiency . Inception v3 is one of the models available that inherently have the capability to classify images and further can be trained to identify different classes.This project utilized to build 14 different plant leaf disease identification, detection and recognition systems.The neural network is trained with the Plant Village dataset. A Graphical User Interface is designed for this system. This GUI permits the user to choose the images from the dataset. Users can select any image from the dataset and the image gets loaded, following which the prediction of the disease will be shown on the User Interface. Convolutional neural network, trained for identifying and recognizing the plant leaf disease, could classify and predict the diseases correctly for almost all the images with few anomalies thus and obtained 94.8% accuracy. Optimal results were obtained by employing the methods specified in the paper. Thus, with implementation and use of these methods for plant disease classification losses in agriculture can be reduced.

6. FUTURE SCOPE

There is advancement in the field of agriculture which has made it possible to produce enough food to serve the demand of 7 billion people.However food security remains threatened due to several factors like climate change, the decline in pollinators, plant diseases, and others. Plant diseases are not only a threat to food security at the global scale, but can also have disastrous consequences for smallholder farmers whose livelihoods depend on healthy crops. In the developing world, more than 80 percent of the agricultural production is generated by smallholder farmers,and reports of yield loss of more than 50% due to pests and diseases are common.

There are many healthy ways to fight disease these days, but if the disease is caught in its early stages, it is far less damaging and easier to cure. Yields may be very low and additional costs for pesticides and fertilizers may be incurred. Proof of disease alone will not prevent farmers from all possible losses.So you can install another AI-controlled bot to perform the necessary steps to cure these diseased leaves.

7. REFERENCES

[1] Mohanty, S. P., Hughes, D. P., &Salathé, M. (2016). Using deep learning for image based plant disease detection. Frontiers in Plant Science,7(September),[1419].https://doi.org/10.3389/fpls.2016.01419

[2] Jyotsna Bankar, and Nitin R Gavai.Convolutional Neural Network based Inception v3 Model for Animal Classification. Global Journal of Advanced Research in Computer and correspondence Engineering, May 2018

[3] Tiwari, D.; Ashish, M.; Gangwar, N.; Sharma, A.; Patel, S.; Bhardwaj, S. Potato leaf disease detection using deep learning. Proceedings of the 2020 4th International Conference on Intelligent Computing and Control Systems (ICICCS), Madurai, India.

[4] Plant Leaf Recognition Using a Convolution Neural Network Wang-Su Jeon1 and Sang-Yong RheeDepartment of IT Convergence Engineering, Kyungnam University, Changwon, Korea Department of Computer Engineering, Kyungnam University, Changwon, K

[5] S.Arivazhagan, R. Newlin Shebiah, S.Ananthi, S.VishnuVarthini. 2013. Location of undesirable locale of plant leaves and grouping of plant leaf sicknesses utilizing surface elements. Agric Eng Int: CIGR Journal..

[6] Prof. Sanjay B. Dhaygude, Mr.NitinP.Kumbhar, "Agricultural plant Leaf Disease Detection Using Image Processing", International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering Vol. 2, Issue 1, January 2013.

[7] Bhumika S.Prajapati, Vipul K.DabhiHarshadkumar, B.Prajapati, "A Survey on Detection and Classification of Cotton Leaf Diseases", International Conference on Electrical, Electronics, and Optimization Techniques (ICEEOT) – 2016.