International Journal of Engineering Inventions
e-ISSN: 2278-7461, p-ISSN: 2319-6491
Volume 11, Issue 3 [May.-June. 2022] PP: 92-97

Plant Leaf Disease Detection Using Deep Learning

B. Vineela Rani[1], K. Megha Nandi Vardhini[2], G. Bharathsai[3], G. Prudhvi[4], B. Manoj Kumar[5]
[12345 –] Department Of Computer Science And Engineering
RAGHU INSTITUTE OF TECHNOLOGY, Visakhapatnam, Andhra Pradesh, India.

ABSTRACT : Agricultural products are a basic necessary items in everyone's life. When crops get infected, it affects agricultural productivity and the country's economic resources. Timely detection of disease helps provide a means to control the spread of disease among plants. Leaf examination is one of the best methods for plant disease diagnosis. Machine Learning (ML) helps us to recognize and understand the information from the digital images. The dataset used for Plant Disease detection contains 9974 augmented images of diseased plant leaves. The raw image of a leaf is pre-processed, segmented and features like shape, colour, texture, vein etc. are extracted. Our proposed work detects the symptoms of plant diseases early and classifies plant diseases according to the symptoms using the Deep Learning (DL) process.

Date of Submission: 25-05-2022 Date of Acceptance: 05-06-2022

I. INTRODUCTION

Deep learning is a subset of machine learning and its is a part of it and which is based on artificial neural networks. Deep learning is also a kind of mimic of the human brain because the neural network can mimic the human brain [1]. It is a hype now it is used to be a lot of data and not enough processing power. Deep learning is a special kind of machine learning that learns to represent the world as nested layers and implement powerful and flexible implementations. Each concept is defined in relation to a simpler concept, and a more abstract representation is considered a lesser representation of the abstraction.

There are about 100 billion neurons in the human brain, and these neurons come together to form a picture of neurons connected through thousands of other neurons. The question here is how it recreates these neurons in the computer [2]. So we create artificial structures called artificial neural networks made up of nodes or neurons. There are multiple neurons for input and multiple neurons for outputs, between which there can be many neurons connected to each other in the hidden layer. The existance of pathogen is directly related to plant diseases. If you see powdery mildew on leaves, then the plant leaf have a fungal disease. Thick liquid exudates are mostly bacteria, and the ulcers themselves are made up of plant tissue and are a symptom but a sign of disease.

Symptoms may include significant changes in the color, shape, or function of plant in response to a pathogen. Leaf wilting is a classic symptom of Verticillium wilting caused by the plant pathogens Verticillium alboatrum and V. dahlia. Common symptoms of fire blight include brown, necrotic lesions surrounded by a bright yellow halo on the leaf edges or inside legume leaves [3].

II. LITERATURE SURVEY

Prof. Sanjay, B. Dhaygude : The application of texture statistics for detecting the plant leaf disease has been explained Firstly by color transformation structure RGB is converted into HSV space because HSV is a good color descriptor. Masking and removing of green pixels with pre-computed threshold level. Then in the A Literature Survey: Plant Leaf Diseases Detection Using Image Processing Techniques segmentation is performed using 32X32 patch size and obtained useful segments. These segments are used for texture analysis by color co-occurrence matrix.

Amandeep Singh, Maninder Lal Singh : The most significant challenge faced during the work was capturing the quality images with maximum detail of the leaf color. It is very typical task to get the image with all the details within a procesable memory. Such images are formed a through high resolution and thus are of 6-10MB of size. This was handled by using a Nikon made D5200 camera which served the task very well. Second challenge faced was to get rid of illumination conditions as from the start to the end of paddy crop season, illumination varies a lot even when the image acquiring time is fixed.

EXISTING SYSTEM

In the existing system, due to low contrast between the leaf and disease, the results were not accurate and our model should be designed for augmented or low contrast image dataset, but no paper has used the parameters that make our image data augmented.

PROPOSED SYSTEM

More noise is removed during the pre-processing step, due to which the disease on image is visible properly. Training image data augmented and train our model for low-resolution images. After cnn model compilation, another algorithm (k-fold cross validation) is verified for cross check of accuracy of our model.

III. DATASET

In our experiments we used three different formats for the Plant Village dataset which contains 6 different types of fruit leaves. We ran the experiment first with images of colored leaves and then with images of segmented leaves from the same dataset. Segmented images have a smoother background, which can provide more meaningful information that is easier to analyze. Finally, we used grayscale images from the same dataset to evaluate the performance of the implemented methods. All leaf images were divided into two sets, a training set and a test set.

IV. SYSTEM ARCHITECTURE

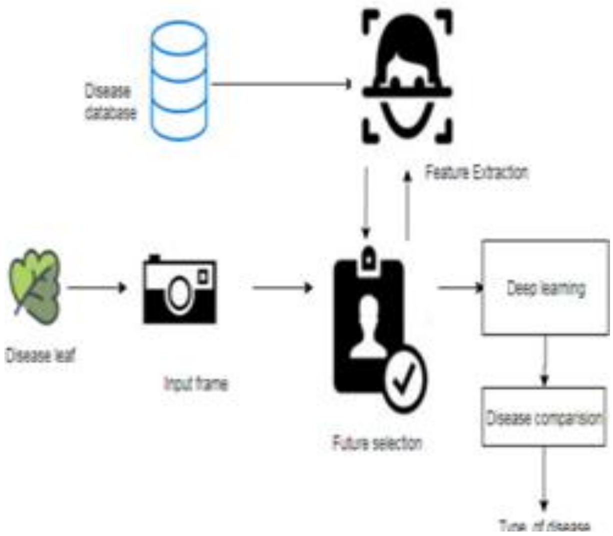


Fig : Architecture diagram

The architecture of the system has a database of all plant leaf diseases we considered. The module is trained number of times to achieve maximum accuracy. When a new image is passed to the module, its features are compared against already trained features in the database [7]. Then the correct result will be given.

V. LIBRARIES

Pandas: Panda is one of the most powerful and useful machine learning tools which helps in cleaning and analyzing data. It has functions that are used to test, clean, convert, and visualize data. Pandas library is an open source Python package built on top of Numpy by Wes McKinney.

Numpy: Used to perform various math and science tasks. It contains many advanced mathematical functions that work with multidimensional arrays and matrices, as well as mathematical concepts such as arrays and matrices.

Cv2: Can be used for tasks like face recognition, relative tracking, mark detection, etc. It provides standard software for computer vision applications and is designed to speed up the use of machine vision.

PIL: PIL is very important library and it stands for Python Imaging Library. The library supports a variety of image formats, including the popular JPEG and PNG formats.

Skimage: An open source Python package type for image preprocessing. It provides easyily usable features for reading, displaying, and saving images.

Scipy : Scientific Python, also known as Scipy, its is a Python numerical processing library. It is well known for some built in features for various mathematical problems frequently encountered in Machine Learning.

Sklearn.utils: This is a collection of various machine learning resources based on scikitlearn's fit / transform paradigm aimed at the common ecosystem and pipeline integration. The focus will be mostly on preliminary consideration and feature selection.

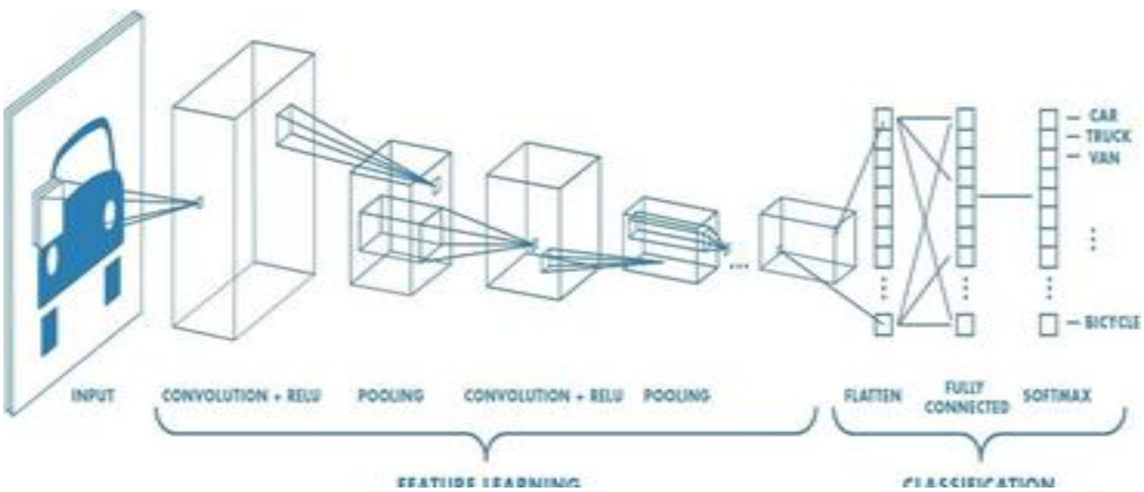
The sklearn library combines multiple powerful machine learning and mathematical modeling tools, which also includes division, deceleration, integration, and helps in reduction of size.

Keras: Keras is a powerful and easy-to-use Python library for developing and testing in-depth learning models. It encapsulates the mathematical libraries of Theano and TensorFlow and allows you to define and train neural network models with a few lines of code.

Tensorflow: TensorFlow is an open source platform for creating machine learning apps. It is a symbolic mathematical library that uses data flow and a diversified system to perform various tasks that focus on training and understanding deep neural networks.

Os: The OS module in Python provides functions for creating and extracting a directory (folder), downloading its contents, modifying and indexing the current index, etc. So, import it using the os import statement before using its functions.

VI. CNN MODEL STEPS



- **Conv2D:** A layer for compositing the images into multiple images activation is the activation function.

- **MaxPooling2D:** It is used to max pool the value from the given size matrix and the same is used for the next 2 layers.

- **Flatten:** It is used to flatten the dimensions of the image obtained after rotating it.

- **Dense:** It is used to make it fully connected model and hidden layer.

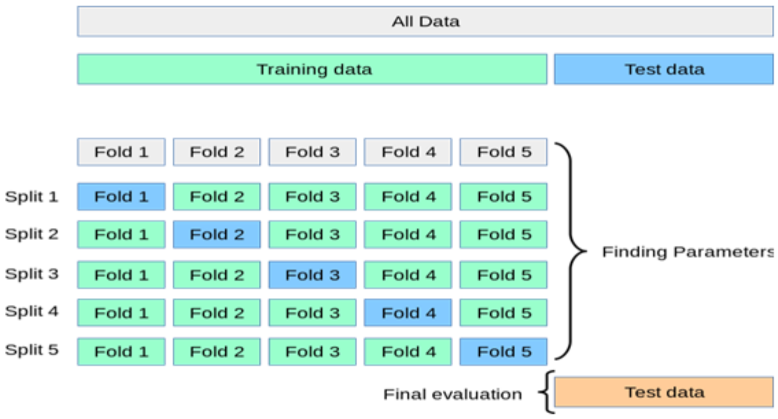
- **Dropout:** It is used to avoid overfitting on the dataset and is dense if the output layer contains only one neuron which decides to which category image belongs.

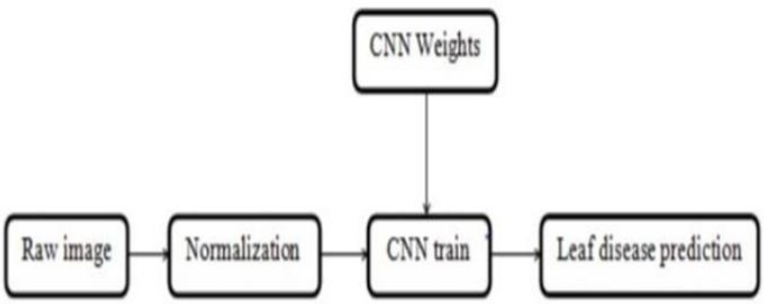
- **Image Data Generator**: It's miles that resizes the image, applies shear in some variety, zooms the image and does horizontal flipping with the image. This photograph data Generator consists of all possible orientation of the photo [6].

VII. K-CROSS VALIDATION

Epochs: It tells us the range of times a version can be educated in forward and backward skip.

Validation procedure: Validation data is used to feed the validation/test data into the version. Validation steps denotes the range of validation/check samples.



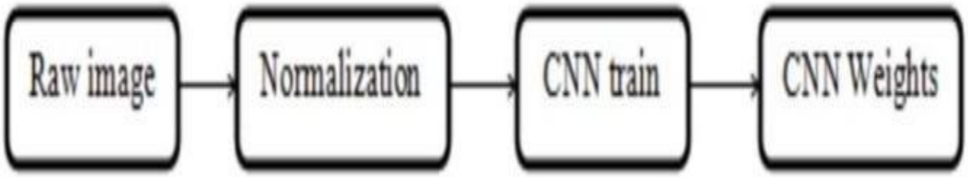


- Cross-validation is tresampling procedure used to evaluate machine learning models on a limited data sample.

- The system has a unmarried parameter referred to as ok that refers to the number of groups that a given information pattern is to be split into. As such, the system is often referred to as ok-fold move-validation. while a selected price for k is chosen, it is able to be utilized in area of okay within the reference to the model, such as k=10 becoming 10-fold cross-validation.

VIII. TRAINING AND TESTING

The dataset is preprocessed inclusive of photo reshaping, resizing, and tabular conversion. similar processing is likewise achieved on the test picture. A facts set including about 6 distinct types of foliar diseases is collected, from which any photograph can be used as a test photo for the software program.

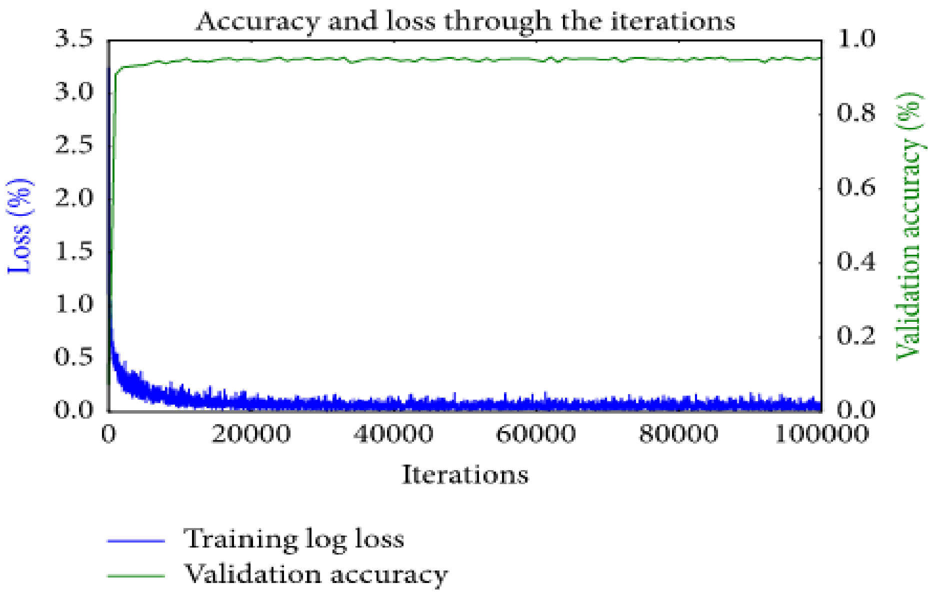


The training dataset is used to train the model (CNN) to identify test images and the diseases. Dense, Dropout, Activation, Flatten, Convolution2D, and MaxPooling2D these all are the layers present in CNN [8]. After successfully training the model, the software can identify diseases if plant species are included in the data set. After successful training and preprocessing, the test image is compared to the trained model to predict disease.

If a model finds any new image to predict its outcome, The model predicts any one of the classes from the 6 classes mentioned before. Here the test image we have given is potato leaf with septoria leaf spot [10].

IX. RESULTS

The results shown in this section refer to database wide training, including both the original and enhanced images. Convolutional networks are known to be able to learn features when trained on larger datasets, so the results obtained on the original image alone cannot be examined.
The algorithms training will give the 98.7 accuracy and the dataset is spilt and detect the leaf is healthy or unhealthy.



X. CONCLUSION

The focus was on how to predict plant disease models using CNN models using images from a given dataset (trained dataset) and past field datasets.

This provides the following insights into the prediction of plant leaf disease. Because the system processes the maximum number of plant leaf types, the farmer can learn about leaves that have not yet grown and list all possible plant leaves so that the farmer can decide which crop to grow.

The system also takes into account historical data so that farmers can get an idea of the demand and value of various plants in the market.

XI. FUTURE ENHANCEMENT

Agricultural departments want to automate crop yield determination (in real time) during the selection process. Automate this process by displaying forecast results in a web application or desktop application. Optimize work for implementation in an artificial intelligence environment.

In the future work, you can collect large datasets to improvise your system. In addition, it can be implemented to predict how the leaves will spread, and the system needs to provide the necessary pesticides and fertilizers for the leaves of infected plants.

REFERENCES

[1]. S. P. Mohanty, D. P. Hughes, and M. Salathé, ‘‘Using deep learning for image-based plant disease detection,’’ Frontiers Plant Sci., vol. 7, p. 1419, Sep. 2016.

[2]. J. D. Pujari, R. Yakkundi Math, and A. S. Byadgi, “Identification and classification of fungal disease affected on agriculture/horticulture crops using image processing techniques,” IEEE International Conference on the Computational Intelligence and Computing Research, 2014.

[3]. Balasubramanian Vijayalakshmi and Vasudev Mohan, “Kernel based PSO and FRVM: An automatic plant leaf type detection using texture, shape and colour features,”Computer and Electronics in Agriculture, vol. 125, pp. 99- 112, 2016.

[4]. X. Wang, M. Zhang, J. Zhu and S. Geng, “Spectral prediction of Phytophthora infestans infection on tomatoes using artificial neural network (ANN),” International Journal of Remote Sensing, pp. 1693– 1706, 2008.

[5]. Dong Pixia and Wang Xiangdong, “Recognition of Greenhouse Cucumber Disease Based on Image Processing Technology,” Open Journal of Applied Sciences, vol. 3, pp. 27-3, Mar. 2013.

[6]. M. Dyrmann, H. Karstoft, and H. S. Midtiby, ‘‘Plant species classification using convolutional neural network,’’ Biosystems Eng., vol. 151, pp. 72–80, Nov. 2016.

[7]. K. Simonyan and A. Zisserman, ‘‘Very deep convolutional networks for large-scale image recognition,’’ in Proc. Int. Conf. Learn. Repr., London, U.K., Apr. 2014, pp. 1–14.

[8]. S. Arivazhagan, R. Newlin Shebiah, S. Ananthi and S. Vishnu Varthini, “Detection of unhealthy region of plant leaves and classification of plant leaf diseases using texture features,” Commission Internationale Genie Rural(CIGR) journal, vol. 15, no. 1, pp. 211-217, 2013.

[9]. B. Liu, Y. Zhang, D.-J. He, and Y.-X. Li, ‘‘Identification of apple leaf diseases based on deep convolutional neural networks,’’ Symmetry, vol. 10,no. 1, pp. 11–19, Dec. 2018.

[10]. Huang, K. Y. (2007). Application of artificial neural network for detecting phalaenopsis seedling diseases using color and texture features. Comput.Electron.Agric.57,311.doi:10.1016/j.compag.2007.01.015.

[11]. T. Fushiki, “Estimation of prediction error by using K-fold cross-validation,” Stat. Comput., vol. 21, no. 2. pp. 137–146, 2011.

[12]. Nikita Jadhav, Himali Kasar, Sumita Chandak and Shivani Machha, "Crop Leaf Disease Diagnosis using Convolutional Neural Network", Published 2020 Biology International Journal of Trend in Scientific Research and Development.

[13]. J. G. A. Barbedo, "Factors influencing the use of deep learning for plant disease recognition", Biosyst. Eng., vol. 172, pp. 84-91, Aug. 2018.

[14]. C Szegedy and S Ioffe “Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift”, Arxiv.org, 2015.

[15]. X. Zhang, Y. Qiao, F. Meng, C. Fan and M. Zhang, "Identification of Maize Leaf Diseases Using Improved Deep Convolutional Neural Networks", IEEE Access, vol. 6, pp. 30370-30377, 2018.