Bharath Prabhu N, 2021, 9:4
ISSN (Online): 2348-4098
ISSN (Print): 2395-4752

Leaf Disease Identification and Remedy Recommendation

Bharath Prabhu N, Girisha S, Varikuti Manush, Ravi Kiran S, Asst. Prof. Ms Savitri Kulkarni
Department of Computer Science and Engineering,
SJB Institute of Technology, Bengaluru, India.

Abstract- Detection of Diseases and identifying the pests can effectively reduce loses to the farmers. In this Paper, Convolution Neural Network is used in order to identify the crop that is been infected by Disease. At the same time, we provide a remedy to the user. The data that has been provided to the algorithm comes from publicly available datasets. The Accuracy of predicting the right disease was more than 88%. The result obtained shows us that the system is capable of predicting the disease and give corresponding guidance.

Keywords- Leaf Disease Identification, Convolution Neural Network.

I. INTRODUCTION

Every Person in India has at least one ancestor who has been Farmer for at least some time. The reason that person is not a farmer is due to the lack of knowledge which resulted in huge loses. There will be a lot of factors that will be affecting the growth of the crops like Climatic Conditions, no. of diseases, Soil Conditions etc.

Most of the people try to predict the disease from naked eye, where they can go wrong as the best of the best can also make mistakes due to number of new diseases every day. So the farmers started to make use of laboratories, many devices etc., which was found to be expensive.

So this resulted in inexperienced pesticide usage, which in turn reduced the crop to fight back against long term pathogens.

But plants play a huge role to every being on the earth. Without plants there exists no ecology. But in recent days, many types of plants are at risk. To protect this, we need a wily formed database and an efficient program. This paper aims at creating a classifier by making use of convolution neural network with one good camera.

II. LITERATURE SURVEY

Authors of [11] give a brief knowledge of using support vector machine. There aim was at proposing an automated system for diagnosing a paddy disease by making use of K-Means Clustering and feature extraction.

Whereas the authors of [12] tried to predict the plants that were affected by climatic conditions in Thailand. Accordingly, they state that system should have an application which can operate for specific disease diagnosis using rule-based model of data mining technique.

Authors of [13] focused on preparing a model that provided a method to identify the leaf disease by inspecting if an image which is subjected to examination is affected by a disease or not. They are generated by using different cluster sizes using image segmentation and thereby obtaining an optimized result.

In the paper [14] authors proposed a system where they were able to identify nitrogen deficiency in the plants and provide right quantity of fertilizers to the plants. However that digital image recognition of plant diseases is one of the thrust areas and hence came out with a model which comprises of back propagation networks and probabilistic neural networks.Thus huge amount of research is always going on in the domain of agriculture in order to yield better and satisfactory results.

International Journal of Science, Engineering and Technology

An Open Access Journal

III. MATERIALS AND METHODS

1. Dataset:

Image-based disease identification includes Training phase, Evaluation phase etc. it requires large amount of data. Hence the source of data is collected from Plant Village website. The image thus collated is labeled with four different categories-bacterial spot, yellow leaf curl virus and Late blight and healthy.

Correspondingly, a database compromising of more than 5000 images was used for training and around 1000 images were used to validate.

2. Process and Label of Images:

Several sets of images were collected from Plant Village website, which were spread across several formats having varying levels of resolutions and variations in quality. Thus for the purpose of acquiring reasonable feature extraction, the final images are provided as input to the classifier which are then pre-processed to achieve consistency.

IV. IMPLEMENTATION

This paper put forth a model which is used to train a CNN model in order to identify leaf disease. Tensor flow which is an open-source library is used for carrying out numerical computations in neural networks along with data flow graphs. Where nodes represent mathematical operations and graph edges represent multidimensional data arrays.

Convolution neural network in machine learning is a type of feed forward artificial neural network in which neurons are associated in a pattern that is simulated by the organization of animal visual cortex. Convolution neural networks may include global or local pooling layers which combines output of neuron clusters.

The first convolution layer filters the input image with 32 kernels of size 3*3. The input to the next convolution layer is the output of the first convolution layer. The last layer of convolution layer has 128 kernels of size 1*1 followed by a fully connected layer pf 512 neurons. The output of this layer is provided as input to the Soft Max function which produces a probability distribution of the four output classes. The below figure 2 depicts sample images from a database of types of leaves while.

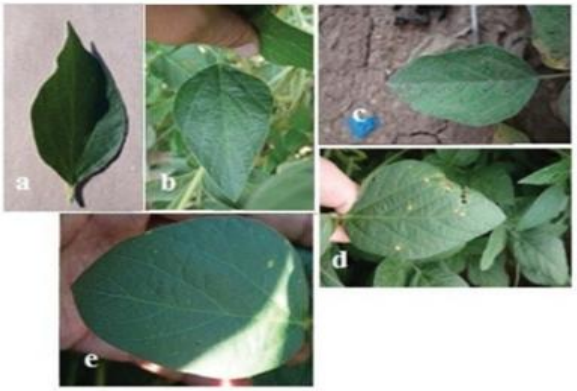

Fig 1. Sample images from the database are shown below.

1. Healthy image of a leaf is taken under a constant background
2. Image of a leaf taken under uncontrolled environment
3. Leaf images from a plant affected by

- Septorial leaf blight
- Frogeye leaf spot
- Downy Mildew

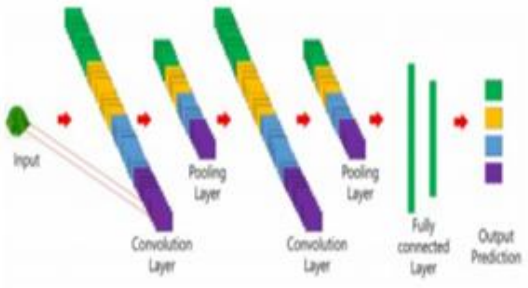

Fig 2. Illustrates working of CNN layers.

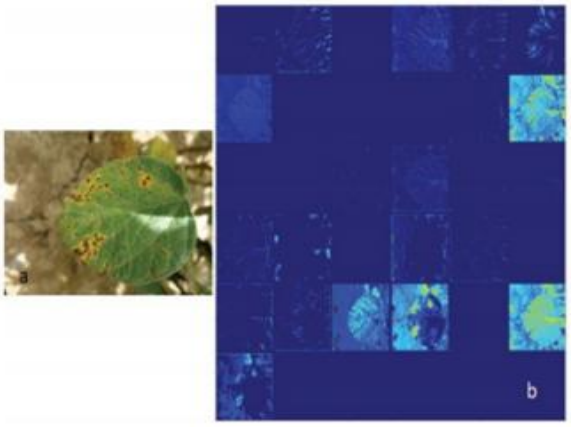

Fig 3. Visualizing feature extraction in the first activation layer.

- Sample image
- Feature maps of the first activation layer

V. RESULTS

The collected dataset is divided into 70% for the training, 10% for validation and 20% for testing. Different models with different architectures and learning are tested.

Trial and error methods are used to select the parameters of the network; the parameters in the architecture are kernel size, filter size, learning parameters etc. The below table depicts the classification results from different models using different architectures.

Table 1. Indicates graph of training accuracy versus actual accuracy.

	Architecture	Validation accuracy	Test accuracy
Grayscale	[3X3, 4X4]	77.60%	78.74%
	[5X5,5X5]	70.20%	70.07%
	[3X3, 4X4]	77.20%	78.67%
	[3X3, 2X2]	77.60%	77.87%
Color	[3X3,4X4, 1X1]	89.30%	88.20%
	[3X3,2X2,2X2]	89.50%	86.90%
	[3X3,4X4,3X3]	89.90%	88.00%
	[3X3,4X4]	88.00%	85.50%
	[3X3,2X2]	87.30%	85.30%
Segmented	[5X5, 3X3]	87.40%	86.00%
	[3X3,4X4]	87.60%	85.90%
	[3X3,2X2]	87.00%	85.50%

From the results, we found that colored images is better than gray scale and the segmented images. This shows the color feature is important to extract the features for classifying.

The model that provides good classification accuracy contains three convolutional layers followed by max pooling layer. The graphs of the training accuracy vs. validation accuracy of the model are shown. It can be seen from the graphs that the model is over fitting.

Over fitting happens when the model fits too well to the training set. It then becomes difficult for the model to generalize to new examples that were not in the training set.

Experiments were conducted to see how each model performs. Since the data is too small when compared to the total number of trainable parameters of the model, the first experiment was to increase the training data by rotating, flipping, rescaling of the images. Only the training data is augmented.

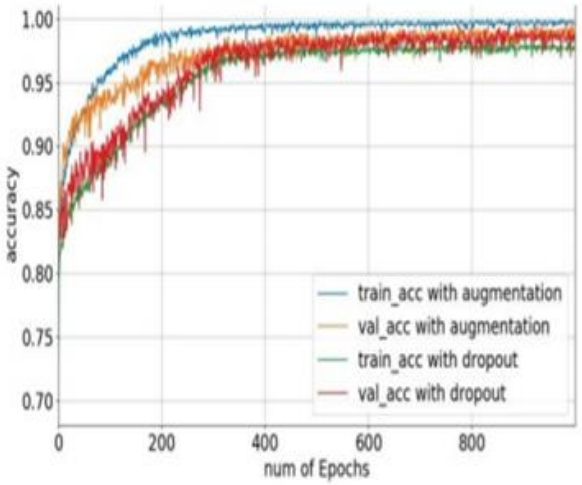

Fig 4. Training vs. Validation accuracy of the models.

Table 2. Illustrates the results of testing and validation accuracy.

Model	Validation accuracy	Test accuracy
Base model with aug. and dropout	99.21%	99.32%
Base model with aug. and L2 regularization	98.62%	98.73%

From the table 2, it is clear that the model provides the required efficiency

VI. CONCLUSION

Convolution Neural Network can be used to detect and classify plant diseases. The Network is trained using the images taken in the natural environment and achieved 99.32% classification ability. This shows the ability of Convolution Neural Network in extracting features of the environment which is required for plant disease classification.

Image classification, Image Categories, Feature Extraction, and Training data all are carried out using a python tool. The algorithm is implemented with training data and classification of given image dataset. The test image is given as input with trained data for detection and prediction analysis. From the results, it is clear that model provides reliable results.

International Journal of Science, Engineering and Technology

An Open Access Journal

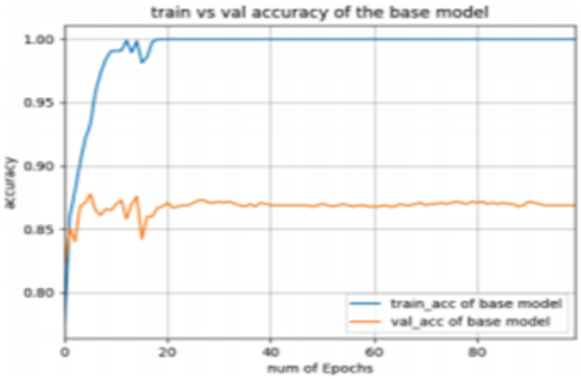

Fig 5. Training vs. Validation accuracy of the base models.

REFERENCES

[1] Tushar H. Jaware, Ravindra D Badgujar and Prashant G. Patil, Crop disease detection using image segmentation, National Conference on Advances in Communication and Computing, World Journal of Science and Technology, pp:190-194, Dhule, Maharashtra, India, 2012.

[2] Mokhled S. Al-Tarawneh, An Empirical Investigation of Olive Leave Spot Disease Using Auto-Cropping Segmentation and Fuzzy C-Means Classification, World Applied Sciences Journal, vol.23, no.9, pp: 1207-1211, 2013.

[3] Yan-Cheng Zhang, Han-Ping Mao, Bo Hu, Ming - Xi Li, Feature Selection of Cotton Disease leaves Image Based on Fuzzy feature Selection Techniques, Proceedings of the 2007 International Conference on Wavelet Analysis and Pattern Recognition, pp:124-129, Beijing, China, Nov.2007.

[4] Haiguang Wang, Guanlin Li, Zhanhong Ma, Xiaolong Li, Image Recognition of Plant Diseases Based on Back propagation Networks, 5th International Congress on Image and Signal Processing, pp-894-900, Chongqing, China, 2012.

[5] Simona E. Grigorescu, Nicolai Petkov, and Peter Kruizinga, Comparison of Texture Features Based on Gabor Filters, IEEE transaction on Image Processing, vol. 11, No. 10, pp-1160-1167, oct 2002.