Feed Forward Neural Network for Plant Leaf Disease Detection and Classification

M. Ganthimathi[1], T. Nanthithasri[2], M. Sakthipriya[3], R. Shanmugapriya[4]

[1]Associate Professor, Dept. of CSE, Muthayammal Engineering College (Autonomous), Rasipuram, Tamilnadu, India.
Email: gandhimanimay02@gmail.com[1]

[2,3,4]UG Student, Dept. of CSE, Muthayammal Engineering College (Autonomous), Rasipuram, Tamilnadu, India.
Email: psakthi1210@gmail.com[3]

Abstract – Crop diseases are a major threat to food security. Their rapid identification remains difficult in many parts of the world due to the lack of the infrastructure. Plants are important as they are the source of energy supply for mankind. Plant diseases can affect leaves at any time between sowing and harvesting. This leads to huge loss on the production of crop and economical value of market. The leaf disease detection plays a vital role in agricultural field. However, it requires huge manpower, more processing time and extensive knowledge about plant diseases. Here, machine learning is applied to detect diseases in plant leaves as it analyzes the data from different aspects and classifies it into one of the predefined set of classes. The features like color, intensity and dimensions of the plant leaves are taken into consideration for classification.

Keywords -

I. Introduction

A. Machine Learning Techniques

Machine learning (ML) is the scientific study of algorithms and statistical models that computer systems use to perform a specific task without using explicit instructions, relying on patterns and inference instead. Machine learning algorithms build a mathematical model based on training data which is also known as sample data. This is used in order to make predictions or decisions without being explicitly programmed to perform the task. Machine learning algorithms are used in a wide variety of applications, where it is difficult or infeasible to develop a conventional algorithm for effectively performing the task. Machine learning techniques such as PCA, Feed Forward Neural Networks, Random Forest, and other feature extractions are used. Machine learning includes data collection, dataset preparation, feature extraction, preprocessing, feature selection, choosing and applying machine learning algorithms and performance evaluation. In this system machine learning methods are mainly applied to predict molecular biology and agriculture related to plant diseases in tomato leaves.

II. Leaf Diseases

Most leaf spots are caused by fungi, but some are caused by bacteria. Some insects also cause damage that appears like a leaf spot disease. Leaf spots on trees are very common and generally do not require spraying and may results some in defoliation of plant

A. Early Blight

Tomato leaf disease fungus, *Alternaria*, also causes leaf spot or early blight shown in Fig.1. Lower leaves show brown or black spots with dark edges, almost like a target. Stem ends of fruits may be attacked, showing large, sunken black areas with concentric rings. This tomato plant disease fungus usually strikes after plants set fruit.

B. Late Blight

The tomato plant disease late blight, caused by the fungus *Phytophthorainfestans*, occurs during periods of cool, rainy weather that may come at the end of a growing season. It looks almost like frost damage on leaves, causing irregular green-black splotches. Fruits may have large, irregular-shaped brown blotches that quickly become rotten as shown in Fig2. This tomato plant disease fungus also affects potatoes and can be transferred from them. Use the same controls as for septoria leaf spot.

C. Septoria Leaf Spot

Septoria leaf spot is one of the most common tomato plant leaf diseases. You can first detect this fungus as it creates a small, circular spot with a grayish-white center and dark edges. Small black spots may show up in the center. Affected tomato plant leaves turn yellow, wither, and fall off as shown in Fig. 3. Long periods of warm, wet weather contribute to

© IJRAD.

this tomato plant disease, and splashing water spreads spores to other leaves. Control leaf spot by not crowding your tomatoes. Leave enough space so air circulates and dries leaves. Avoid overhead watering. When watering tomatoes, water at the base of the plant. Also, water in the morning so wet leaves have time to dry before evening. A fungicide formulated for tomatoes can be used to treat affected plants.

D. Tomato Bacterial Diseases

Tomatoes can fall prey to several tomato plant bacterial diseases, including bacterial spot, bacterial speck, and bacterial canker. They are all slightly different but appear as spots on leaves and fruits. Use the same controls as for septoria leaf spot. Grow disease-resistant plants. Avoid rotating the same ground with peppers, which can host the same diseases. Avoid pruning and tying plants because the bacteria can enter any openings made during these procedures. Fixed copper sprays may reduce the spread if applied as soon as symptoms begin.

E. Mosaic Virus

Mosaic virus attacks many kinds of plants and is common in tomatoes. While mosaic virus does not kill the plant, it reduces number and quality of the fruits. Leaves may also grow in misshapen forms, resembling ferns shown in Fig. 4. Because the virus must enter through a cut in the plant, avoid handling the plant. Anyone who uses tobacco can easily transmit the di script tuned to perform well on our dataset

Fig. 1 Early blight

Fig. 2 Late Blight

Fig. 3 Spectoria spot

Fig. 4 Mosaic virus

III. PROPOSED SYSTEM

Across all our experiments, we use three different versions of the whole Plant Village dataset. We start with the Plant Village dataset as it is, in color. Then we experiment with a gray-scaled version of the Plant Village dataset, and finally we run all the experiments on a version of the dataset. Here where the leaves were segmented. Hence removing the extra background information which might have the potential to introduce some underlying factors in the dataset due to the regularized process of data collection in case of Plant Village dataset. Segmentation was automated by the means of a tuned script to perform well on our dataset.

IV. MODULE DESCRIPTION

A. Learning Approach

We analyze images of plant leaves, which have a spread of many class labels assigned to them. Each class label is a crop-disease pair, and we tried to predict the crop-disease pair given just the image of the plant leaf. We start with the PlantVillagedataset which consists of colored images. Then we experiment with a gray-scaled version of the images in the PlantVillage dataset.

We run all the experiments on a version of the PlantVillage dataset where the leaves were segmented. By removing all the extra background information which might have the potential to introduce some inherent bias in the dataset due to the regularized process of data collection in case of PlantVillage dataset. Segmentation was automated by the means of a tuned script to perform well on our dataset. Here we chose a technique that works based on a set of masks generated by analysis of the color, lightness, and saturation components of different parts of the leaf images. One of the steps in processing allowed us to easily fix color casts, which happened to be very strong in some of the subsets of the dataset, thus removing another potential bias. The sample datasets are shown in Figure 5.

B. Principal Component Analysis

Principal component analysis is a procedure that uses a transformation orthoganally. This convert a set of observations of possibly correlated variables into a set of values of linearly uncorrelated variables which is called principal components. The transformation is defined in a way that the first principal component has the largest possible variance. Then each succeeding component in turn has the highest variance possible only under the constraint that it is orthogonal to the preceding components. The resulting vectors are an uncorrelated orthogonal basis set. It is sensitive to the scaling of the original variables. PCA is mostly used as a tool in exploratory data analysis and for making predictive models. This analysis is a technique used to gives the variation and bring out strong patterns in a dataset. It is generally used to make data easy to explore and easy to visualize.

© IJRAD.

PCA used to implement the image recognition of plant disease based on the extracted color features, shape features and texture features from disease images and their combined features. As the result, PCA could reduce the dimensions of the feature data extracted from the images of plant diseases could reduce the number of neurons in the input layer, also the running time of the neural networks is reduced and acceptable recognition results could be obtained. PCA mainly used to provide a optimal neural networks, that could be constructed for plant disease identification. The flow chart for the disease prediction is shown in Fig. 6

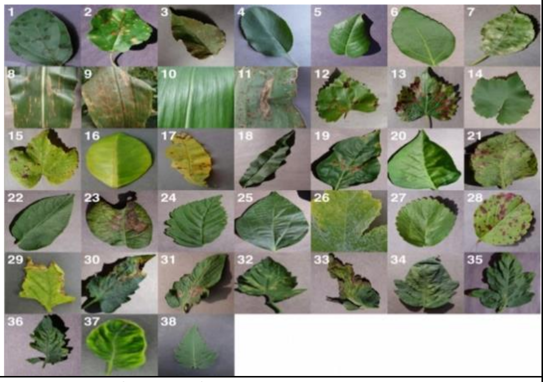

Fig. 5 Leaf images from the Plant Village dataset

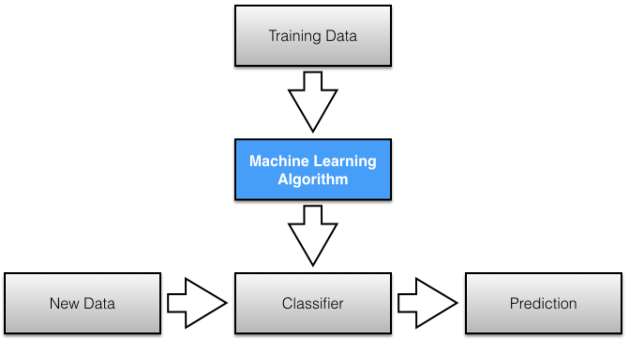

Fig. 6 Flow Chart

C. Feed Forward Neural Network

The feed forward neural network was the first and simplest type of artificial neural network. In this neural network, the information moves in only one direction, forward, from the input nodes, then goes through the hidden nodes and to the output nodes. There are no loops in the network. To do this, plant leaves images were taken as inputs and the inputs were changed into HIS color space representation. In this step, due emphasis was given to the H value and the green colored pixels are masked and removed to reduce the computation time because green color indicates healthy pixels. By taking only the infected portions of the leaves from the last step, segmentation was done using the K-medoid clustering. Color and texture feature were extracted using color co-occurrence method to train the model. Finally, the diseased leaves were classified using least square support vector machine taking the features as inputs.

D. Random Forest

Random forests are an ensemble learning method that is mainly used for classification, regression and other tasks. This operates by constructing a multitude of decision trees at training time. Then the outputting the class that is the mSode of the classes which is for classification or mean prediction which is for regression of the individual trees. Random decision forests correct for decision tree habit of over fitting to their training set. Random forest is a classifier that evolves from decision trees. It consists of many decision trees. To classify a new instance, each decision tree provides a classification for input data; random forest collects the classifications and chooses the most voted prediction as the result. The input of each tree is sampled data from the original dataset. In addition, a subset of features is randomly selected from the optional features to grow the tree at each node. Each tree is grown without pruning. Essentially, random forest enables many weak or weakly-correlated classifiers to form a strong classifier. In this algorithm there are two steps, in the first stage a random forest creation is produced and in the second step the values that are produced from step one are used and predictions are made. Here select m features from n feature. Random forest classifier creates a set of decision trees from randomly selected subset of training set. It then aggregates the votes from different decision trees to decide the final class of the test object.

E. Feature Extraction

Leaf disease detection plays a vital role in agricultural field. However, it requires huge manpower, more processing time and extensive knowledge about plant diseases. Therefore, machine learning is applied to detect diseases in plant leaves as it analyzes the data from different aspects and classifies it into one of the predefined set of classes. The morphological features and properties like color, intensity and dimensions of the plant leaves are taken into consideration. This paper presents an overview on various types of plant diseases and different classification techniques in machine learning that are used for identifying different kind of diseases in plant leaves.

V. CONCLUSION

The proposed ides deal with automatic disease detection of plant leaf using image processing techniques. It involves image acquisition, image preprocessing, image segmentation, feature extraction and classification. Development of automatic detection system using advanced computer technology such as image processing help to support the farmers in the identification of diseases at an early or initial stage and provide useful information for its control. To identify the disease is main purpose of the proposed approach. The experimental outcome shows the approach that it can identify the leaf diseases. The expansion of this work will concentrate on developing algorithms to increase the identification rate of classification process. The output of the process is shown in Fig. 7.

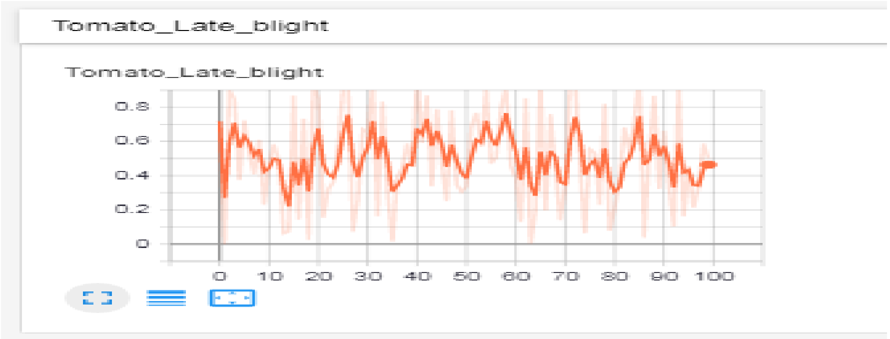

Fig. 7 Screen Shot

VI. FUTURE SCOPE

Future work can develop in hybrid algorithm such as other clustering method and PCA is ordered to improve the recognition rate of final classification process. Further needed to compute amount of disease present on leaf and will dedicate our future works on automatically estimating the severity of the detected disease. Would like to extend our work further on more plant disease detection.

REFERENCES

[1] Khirade SD, Patil AB, Plant, "Disease Detection Using Image Processing," Int. Conf on Comp Communication Control and Automation., IEEE, vol. 153, pp. 768-771, 2015.
[2] Surya Prabha D, Satheesh Kumar J, Study on Banana Leaf Disease Identification Using Image Processing Methods. Int Jour of Res in Comp Sc and Inf Tech IJRCSIT 2(A), pp. 89-94, 2014.
[3] Weizheng S, Yachun W, Zhanliang C, Hongda3 W, Grading Method of Leaf Spot Disease Based on Image Processing. Int Conf on Comp Sc and Soft Eng, IEEE, pp. 491-494, 2008.
[4] Phadikar S, Sil J Rice Disease Identification using Pattern Recognition Techniques, Proc. of 11th Int. Conf on Comp and Inf Tech, ICCIT, Khulna, Bangladesh, IEEE pp. 420-423, 2008.
[5] Sabah B, Sharma N Remote area plant disease detection using image processing. IOSR J Electron Commun. Eng., vol. 2, no. 6, pp. 31-40, 2012.
[6] Sanjay B, Patilet. Leaf Disease Severity Measurement Using Image Processing. Int Jour of Eng. and Tech., vol. 3, no. 5, pp. 297-301, 2011.
[7] Phadikar S, Jaya S Rice Disease Identification using Pattern Recognition Techniques. Proceed of 11th Int. Conf. on Comp and Inf Tech ICCIT, Khulna, Bangladesh, 2008.
[8] Al Bashish D, Braik M, Bani-Ahmad, A Framework for Detection and Classification of Plant Leaf and Stem Diseases. Int. Conf on Signal and Images Processing IEEE, 2010
[9] Bin Husin Z, Bin Abdul Aziz AH, Bin Md Shakaff AY, Mohamed Farook, RBS Feasibility Study on Plant Chilli Disease Detection Using Image Processing Techniques. Third Int. Conf on Intelligent Sys Modelling and Simulation IEEE, 2012.
[10] Wan MohdFadzil WMN, Shah Rizam MSB, Jilani R, Nooritawati MT Orchid Leaf Disease Detection using Border Segmentation Techniques. Conf on Systems Process and Control IEEE, Kuala Lumpur, Malaysia, 2014.

© IJRAD.