TEST Engineering & Management

May-June 2020
ISSN: 0193-4120 Page No. 6756 - 6761

Leaf Disease Detection Using Deep Learning Algorithm

Naadem Divya, Assistant Professor**,** IT Department Sreenidhi Institute of Science and Technology, Yamnampet, Hyderabad. (naademdivya@sreenidhi.edu.in)

Sunil Bhutada, Professor, IT Department Sreenidhi Institute of Science and Technology, Yamnampet, Hyderabad. (sunilb@sreenidhi.edu.in)

C.Venkata Sai Krishna, B. Tech IV year, IT Department Sreenidhi Institute of Science and Technology, Yamnampet, Hyderabad (saikittu987.sk@gmail.com)

Masapaka Bhanu Kumar, B. Tech IV year, IT Department Sreenidhi Institute of Science and Technology, Yamnampet, Hyderabad (mbhanukumar1004@gmail.com)

Article Info
Volume 83
Page Number: 6756 - 6761
Publication Issue:
May-June 2020

Article History
Article Received: *19 November 2019*
Revised: *27 January 2020*
Accepted: *24 February 2020*
Publication: *18 May 2020*

Abstract:
Predictive Modelling techniques are playing a significant role in the field of many areas like medicine, agriculture etc. The crops are affected by various factors such as soil conditions, humidity, climatic conditions and also by various diseases. Usually farmers detect the condition of crops/leaves by naked eyes which require a lot of labor, knowledge about leaf diseases and they also require expensive devices. The improper disease prediction by farmers may lead to the use of wrong pesticide which damages the crop. Many machine learning techniques like SVM, Decision Tree, etc. and Convolutional Neural Network (CNN) are applied to detect the diseases and a comparison study is done on accuracy and for various other factors.

Keywords: *Predictive Modelling, Crop, Disease Prediction, Convolutional Neural Network, Visualization, Accuracy, Remedial Actions*

1. INTRODUCTION:

One of the major threats that we are facing from many years in crops is disease identification where improper detection damages the crop growth and affects the quality which in turn reduces the economy of the nation. In this world, plant diseases cause tremendous loss to valuable food crops. More than 800 million people across the world are suffering with lack of adequate food due to at least 10% of crop losses [1]. Detecting the diseases accurately and precisely is very important and it's important to know when the disease will cross a certain threshold value which can causes tremendous loss to crops and the economy. If we can predict the diseases before it becomes a major outburst, then that will help the farmers to improve the crop production and management practices in an efficient, timely and responsive fashion. Plant disease identification can be done in many ways.

In olden days, plant disease identification relied on human visual examination [2]. Later on, many people came forward and proposed various machine learning models to detect the diseases of leaves where pre-processing of images is done by using histogram equalization, MATLAB, OpenCV etc. which consumes a lot of time for preprocessing for large data sets and for model training. For better

May-June 2020
ISSN: 0193-4120 Page No. 6756 - 6761

accuracy and automatic pre-processing of images, we have used convolutional neural network. Here each image is passed through many random filters like identity, sharpen, edgy detection, blur detection etc. and pooling is performed to reduce the dimensionality but retains the important features. Hence it removes the unwanted information while feeding to the model thus results in better accuracy.

Our proposed methodology is segregated into various contents like Section 2 literature survey which talks about the related work followed by Section 3 Problem Statement followed by Section 4 Proposed Methodology followed by Section 5 Results and Discussion gives the comparison between various algorithms with the proposed system followed by Section 6 Conclusion and Future Scope.

2. LITERATURE SURVEY:

Survey of Literature presents different machine learning algorithms that are used for the detection of leaf diseases. J.S. West *et al*. [3] introduced a new concept of an array of image-based diagnosis methods in the field of digital image processing. It requires an extra preprocessing step to extract the condensed information which is difficult for the computer to process it and manually done by experts [4]. Machine Learning detects the patterns in the given data set in a computational way. A general example is a machine detecting the handwritten digits [5]. The technique to identify the diseases through the plants by color co-occurrence method is used by the author Mrunalini [6] but this technique involved a lot of computations.

The study of neural networks started from long ago in the early 1940s. Convolutional Neural Networks (CNN) is mostly used in image recognition. Neocognitron was the first CNN which was introduced in the 1980s [7].

S. Raj Kumar *et al*. [8] they have used the technique consisting of hybrid features and geometric shape of a leaf by their methodology. A tool named Plant Village is available consisting of information regarding the crop health and crop diseases but the site is more useful to farmers rather than the professional plant pathologists and the information only about expertly identified leaves are present in the database.

H. Cartwright [9] used Artificial Neural Network (ANN) for disease identification where image processing is performed with respect to better feature extraction and this system gave better accuracy but did not provided any remedial solution. Steinwart *et al.* [10] used Support Vector Machines (SVM) for disease identification where strategies like vector control will give early information regarding the crop health but the method did not give good accuracy .The usage of sensors for health monitoring in agriculture should be rapid and cost-effective and the ground-based sensor system will monitor the health and diseases in plants and they have used various technologies like spectroscopic and imaging-based and volatile-based disease identification methods [11] but the maintenance cost is too high. A method named novel visualization has been introduced that detects the lesion caused by a plant disease where we extract the features responsible for making the decision and these features are learned in an unsupervised manner and compared with human identified symptoms describing how the model will make predictions [12].

Various machine learning algorithms are also used for identifying the plant diseases and we have used CNN in our proposed methodology and the comparison between them varies in the terms of accuracy, efficiency and pre-processing speed. The ultimate goal is to find out which algorithm works efficiently with huge data sets and gives good accuracy.

3. PROBLEM STATEMENT:

Usually diseases on the plants/leaves are identified by human inspection in the olden days but that

May-June 2020
ISSN: 0193-4120 Page No. 6756 - 6761

consumes a lot of time and most of the farmers will not be able to identify the diseases and if identified don't know the remedy to protect or does not get timely help to protect the crop. Identifying the diseases is not only important but also predicting the disease before it becomes a major outburst (becoming epidemic), as identifying the diseases in the later stage will destroy the entire crops and cause tremendous loss to the farmers and economy of the nation.

The objective of the proposed methodology is to determine the disease of leaves in the early stage and also provides the remedies. These remedies will provide the farmer to take the necessary steps to protect the crop and to halt the further spread of disease.

4. PROPOSED METHODOLOGY:

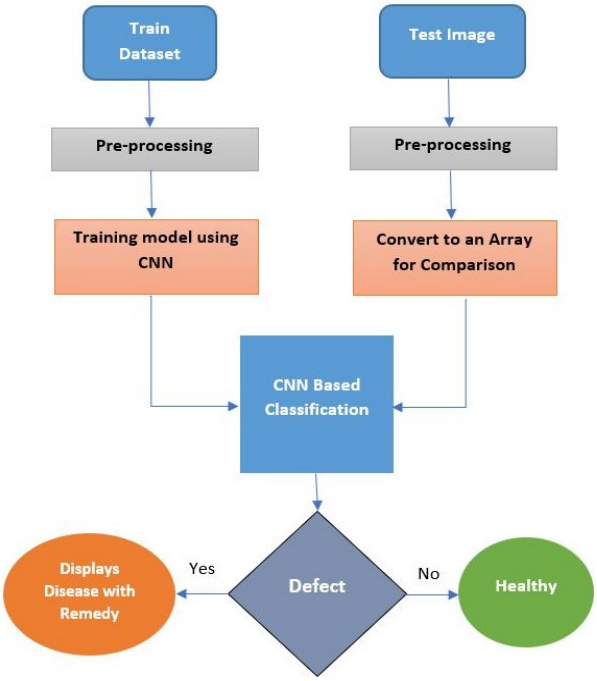
We can reduce the contamination of diseases by identifying the diseases in the early stage and suggesting the farmers with remedial actions. The images of the leaves are confined to a fixed size (50 x 50) and by using the python library open-cv, we are reading the images from the data set and storing them into the NumPy array files which are faster than the csv files for retrieval during the testing time for comparison.

The images are passed through many layers of Convolutional Neural Network and here the images are convoluted, passed through activation function and pooling is done to retain the important features. The model is trained with given data set and the test image is passed to the model where the model predicts whether the leaf is healthy or unhealthy. If it is unhealthy, proper remedies will be displayed.

4.1 Understanding Data:

We took the data set of tomato crop [13] with 6000 samples, where we considered 4000 (which are labeled) images of tomato in training which are divided into 4 categories such as Healthy, Late Blight, Virus and Bacteria and 200 images for testing (which are unlabeled).

4.2 System Architecture:



Convolutional Neural Network consists of many layers Conv2D, Activation, MaxPool2D, Drop out, Fully Connected and the mechanism of CNN can be depicted from the fig 4.a.

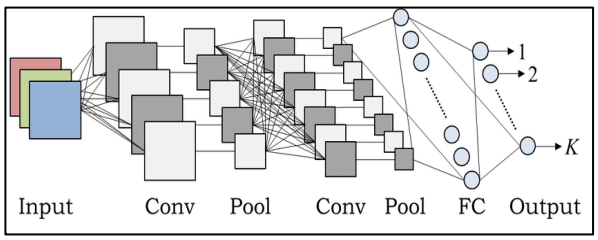


Fig 4.a

The leaf images are passed through the convolution layer where the aim is to extract features from the image by passing through random set of filters like Edgy Detection, Blur Detection, and Sharpen etc.

The Fig 4.b depicts the passage of image through convolutional layer by applying one filter and we have specified the stride as (3 X 3) in which the filter moves through 3 rows and 3 columns and hence the images are convoluted.

May-June 2020
ISSN: 0193-4120 Page No. 6756 - 6761

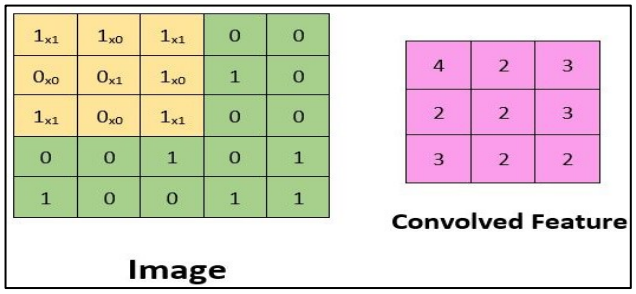


Fig 4.b

Later, the convoluted imaged are passed through activation layer (we have used relu as activation function) where the activation layer removes the unwanted information as you can observe in Fig 4.c

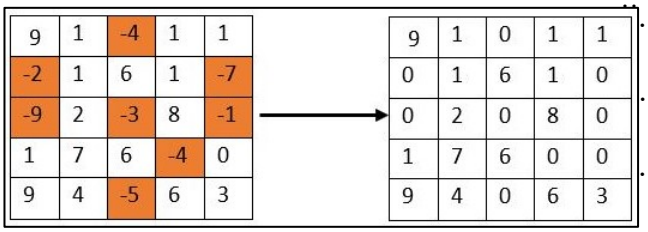


Fig 4.c

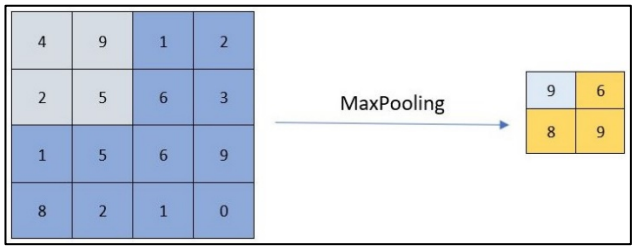


Fig 4.d

Later, the image is passed through pooling to reduce the dimensionality of each feature but retains the most important information. There are many types of pooling such as max pooling, average pooling, sum pooling. We have used max pooling (Fig 4.d) in our project.

We have passed the images through fives layers through different number of filters and repeating the same steps mentioned above.

At last, we will pass through the Fully connected layer to segregate the input image into various classes mentioned in the training data set.

In Regression layer, we will provide the optimizer and we have used Adam as our optimizer which gave good accuracy compared to other optimizers and we will build the model, we have used DNN as our model. In Convolutional Neural Network, images are fed into batch wise mode, the default size of a batch is 64 and we have used default size.

Pseudo code:

Input: Collect the Training and Testing data.

Convert the Training and Testing images into .npy files.

Pass the images through convolutional layers.

Model is trained.

Test the input image.

If the leaf is healthy, set the status as healthy.

If the leaf is unhealthy, set the status as unhealthy and provide the remedial actions.

Output: Status is displayed through user interface.

Experimental Steps:

a) Both the training and testing images are converted into .npy files for faster comparison in the later stages.

b) Images are passed through convolutional layer and the model is trained.

c) Input the test image, and comparison takes place between trained model and test image.

d) If the leaf is healthy, the proposed system prints the status as healthy.

e) If the leafy is unhealthy, then the proposed system will print the type of disease that the leaf is suffering and proper remedies will be provided.

f) The output can be observed through user interface.

May-June 2020
ISSN: 0193-4120 Page No. 6756 - 6761

5. RESULTS AND DISCUSSION:

As we have gone through papers and references, we have implemented the leaf disease identification using machine learning algorithms and compared with CNN.

Here the images are pre-processed using 2 methods.

Method 1: PIL (Python image Library) where the images are converted into grey scale and Image.Antialias is used to get rid of distortion in the images.

Method 2: Using Histogram Equalization, we have pre-processed the images.

	Using PIL Accuracy (in %)	Histogram Equalization Accuracy (in %)
KNN Classifier (K=5)	0.848	0.869
Decision Tree Classifier	0.831	0.858
Random Forest Classifier(n=400)	0.934	0.924
Logistic Regression	0.711	0.768
SVM classifier	0.7475	0.7475

In the proposed system as, we have used CNN and we have used different epochs and different optimizers to identify which one will provide better accuracy.

CNN Method

Optimizers	Epochs=6	Epochs=8	Epochs=10
Sgd	0.7148	0.7451	0.7972
Momentum	0.8715	0.8125	0.9252
rmsprop	0.9161	0.9537	0.9580
adam	0.9600	0.9753	0.9849

Among the various optimizers, Adam optimizer gave higher accuracy with 10 epochs.

From the above tables, we can clearly notice that Proposed System (CNN) gave more accuracy than the machine learning algorithms.

After the leaf is correctly predicted and if it is unhealthy proper remedial actions are taken.

For example: If the leaf is suffering with bacteria disease.

Remedy: Discard or destroy the affected plants. Do not compose them. Shift your plants yearly to prevent the infection. Copper fungicides should be used to control the disease.

Sample Output:

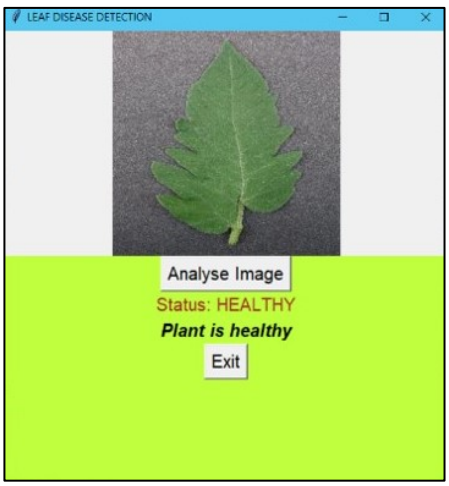


Fig 5.a

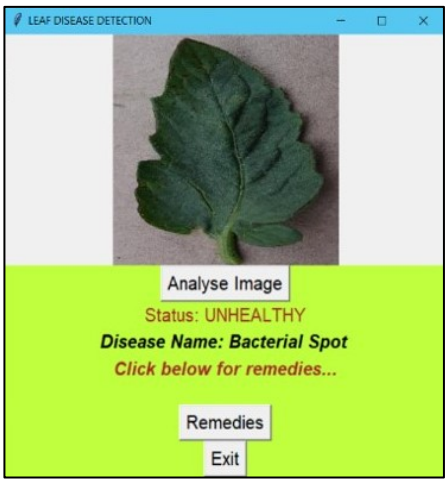


Fig 5.b

May-June 2020
ISSN: 0193-4120 Page No. 6756 - 6761

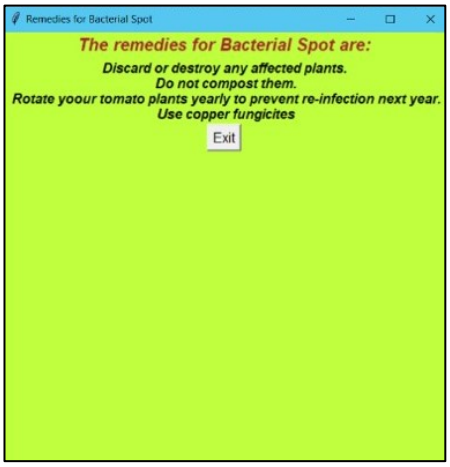


Fig 5.c

Fig 5.a: indicates that the leaf is healthy.

Fig 5.b: indicates that the leaf is unhealthy and displays the disease.

Fig 5.c: indicates the remedies to be followed.

6. CONCLUSION AND FUTURE SCOPE:

The proposed system was developed for the benefits of food growers (farmers) and in turn increases the productivity in agricultural sector. The proposed system can detect the diseases in the leaves and gives us proper remedies. In the proposed system, we have used Convolutional Neural Network which gave higher accuracy (98.4%) compared to machine learning algorithms. We can increase the speed of pre-processing the images by the use of Google GPU. In the future, with the help of drones we can perform the aerial surveillance of the crops and where the images are sent to the cloud and pre-processed and model will predict whether the plant is healthy or not. If it is unhealthy, proper remedies will be provided to the food growers through application.

REFERENCES

1. Erick D. De Wolf, Scott A. Isard, “Disease Cycle Approach to Plant Disease Prediction”, Annual Review of Phytopathology, vol. 45:203-220, 2007.
2. V. Singh, A. K. Misra, “Detection of Plant Leaf Diseases Using Image Segmentation and Soft Computing Techniques”, Information Processing in Agriculture, 2016.
3. J.S. West, Bravo, R. Oberti, D. Lemaire, D. Moshou, and H.A. McCartney, “The potential of optical canopy measurement for targeted control of field diseases”, Annual Review of Phytopathology, vol. 41, pp. 593-614, 2003.
4. A. Singh, B. Ganapathysubramanian, A. K. Singh, and S. Sarkar, “Machine Learning for high throughput stress phenotyping in plants”, Trends in Plant Science, vol. 21, no.2, pp. 110-124, 2016.
5. Li Deng, “The MNIST Database of Handwritten Digits images for Machine Learning Research”, MIT Press, November 2012.
6. Mrunalini R. et al., “An application of K-means clustering and artificial intelligence in pattern recognition for crop diseases”, 2011.
7. K. Fukushima, “Neocognitron: a self-organizing neural network model for a mechanism of visual pattern recognition,” Biological Cybernetics, vol. 36, no. 4, pp. 193-202, 1980.
8. S. Raj Kumar, S. Sowrirajan, “Automatic Leaf Disease Detection and Classification using Hybrid Features and Supervised Classifier”, International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering, vol. 5, Issue 6, 2016.
9. H. Cartwright, Ed., Artificial Neural Networks, Humana Press, 2015.
10. Steinwart and A. Christmann, Support Vector Machines, Springer Science & Business Media, New York, NY, USA, 2008.
11. S. Sankaran, A. Mishra, R. Ehsani, and C. Davis, “A review of advanced techniques for detecting plant diseases,” Computers and Electronics in Agriculture, vol. 72, no. 1, pp. 1-13, 2010.
12. S. Ghosal, D. Blystone, A. K. Singh, B. Ganapathysubramanian, A. Singh, and S. Sarkar, “An explainable deep machine vision framework for plant stress phenotyping, “Proceedings of the National Academy of Sciences of the United States of America, vol. 115, no. 18, pp.4613-4618, 2018.
13. https://www.kaggle.com/sohaibalam67/tomato-disease [Data set]