



# INTERNATIONAL RESEARCH JOURNAL ON ADVANCED SCIENCE HUB

e-ISSN : 2582 - 4376  
Open Access

## RSP SCIENCE HUB

(The Hub of Research Ideas)

Available online at www.rspsciencehub.com

Special Issue of First International Conference on Management, Science and Technology (ICMST 2021)

## Survey on Plant Diseases Prediction using Machine learning for better Crop Yield

Srilatha Toomula<sup>1</sup>, C. Jyothi Sree<sup>2</sup>

<sup>1</sup>Research Scholar, Department of Computer Science & Engineering, Osmania University, Hyderabad, India.

<sup>2</sup>Assistant Professor, Department of Computer Science, RBVRR Women's College, Hyderabad, India. toomula.srilatha@gmail.com<sup>1</sup>, jyothikjr13@gmail.com<sup>2</sup>

### Abstract

Agriculture is a process of growing crops, soil cultivating, it provides food and fabric and helps for growing country's economy. In India more than 50% of people directly and indirectly depend on the agriculture. For developing agriculture, main interceptions are weather hazards and the crop diseases. Weather hazards cannot be prevented, but the loss that occurs due to crop diseases can be reduced. This can be achieved by identifying the crop disease as early as possible and it is also important to identify the type of crop diseases for preventing the spreading of the disease. In India, we have 160 million hectares of arable land and it is second largest country after the United States. Identifying the crop diseases by human action manually is practically difficult and it is hard to identify the type of crop diseases. So, many researchers involved in to identify the crop diseases based on the image processing for helping the real-time gadgets which can be used to identify the crop disease and its types. This survey focuses on the investigation on the different surveys carried out and work related to the crop disease identification and detection based on the image processing. Computer vision and image processing-based work will help to detect of crop diseases along with many practical based applications like drones, IoT based devices etc. In recent studies most of the works are depending on the machine learning and deep learning-based image processing on various studies of predictions. After analyzing the related work on crop detection based on image processing, most of the works achieved better results based on deep learning algorithms compared to the machine learning algorithms.

**Keywords:** Deep Learning, Machine Learning, Crop Disease Detection, Crop Disease Identification.

### 1. Introduction

Farming had very important role in the economic growth of any Country. It is the field which highly affects the Gross Domestic Product (GDP) of the countries. Farming sector contributes main role GDP of India. There are different factors that affects the quality and quantity of crops cultivated. Due to some weather and local conditions these Farming plants are exposed to various diseases. And if these diseases remain undetected may cause some serious losses. In India itself around 20-

25percent of crops are lost due to diseases, pest, and weeds. In this paper, we are focusing on and doing Survey on Plant Diseases in section 2 in section -3 we are designing on overview given by different authors in there paper work section – 4 we are providing Conclusions on our Survey on Plant Diseases using Image Processing n final step, different machine learning and deep learning algorithms Gaussian Naïve Bayes, SVM, Logistic Regression, Linear Discriminate analysis, Random Forest.[1-5].

**2. Survey**

**2.1 Histogram of an Oriented Gradient (HOG)**

Shima et al. proposed concept of detection the plant diseases using Machine Learning algorithms by extracting Histogram of an Oriented Gradient features of the plant images. For Machine Learning algorithms require labelled data for classification and for predicting the class based on the features. In this survey they used three types of features taken from the images of leaves of plants for predicting the plant diseases. Means in HOG is combination the three types of elements, those are.

- Hu moments
- Haralick texture
- Color Histogram

Hu moments is initial stage for feature extraction, in this after conversion of the image RGB to Gray image with help of the pixels Hu moments will calculate. In simply Hu moments identify the shape of the leaves from images. Haralick texture is used to identify the texture of the image; generally the textures of the normal leaves and disease leaves are completely different. Based on the texture generate the texture value by using the pixels of the images store the values in the adjacency matrix. In this also conversion of the image RGB to Gray is required. Like Texture, Color of the leaves also different from the health plants and diseased plants. By collecting the color information diseases leaves can be predicted. For this image should convert to RGB to HSV color space is required. Based on this we can the pixels positions available in the given color. After collecting the HOG features, applied the classification algorithms for training and testing to find the accuracy. In this survey they applied SVM, Gaussian Naïve Bayes, Logistic Regression, Linear Discriminate analysis and Random forest algorithm. From these they found best accuracy for the Random Forest algorithm. The prediction process of detection of the plant disease is divided to 8 steps. [6-10].

- Step 1: Labelled Image Collection
- Step 2: Extract Fog Features
- Step 3: Add to Train Set
- Step 4: Train the RF Model
- Step 5: For each input image
- Step 6: Extract Fog Features
- Step 7: Predict the Results
- Step 8: Repeat Step 3.

In the following table we can see the accuracy achieved by ML algorithms for predicting the plant disease.

**Table.1 Accuracy Results of Machine Learning models**

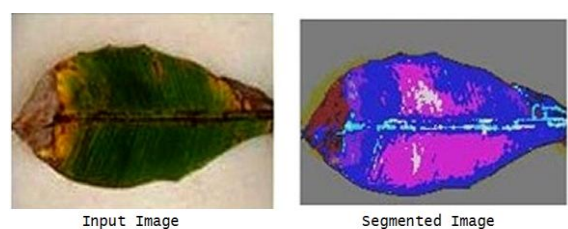
Sno	Algorithms	Accuracy
1	Gaussian Naïve Bayes	0.59
2	SVM	0.4
3	Logistic Regression	0.65
4	Linear Discriminate analysis	0.55
5	Random Forest	0.73

**2.2 Image Segmentation**

Vijai et. al. [7] surveyed on the for automatic prediction of leaf diseases. The proposed genetic algorithm is for image segmentation process on leaves to identify the region of the disease from the image and predict the type disease of the plant. This genetic algorithm is depends on the labelled data of collection of images which are identified by the type disease. This genetic algorithm is depending on two process, one is clustering and another one is classification. The proposed genetic algorithm steps are clearly depicted in the following.

Steps:

1. Image acquisition (Collection of images)
2. Identify the leaf image region from the image and increase the contrast of the image.
3. Green colored pixels are masked with red color to identify the segment part. After comparing intensity of the pixel color with threshold value then its turns to red color otherwise 0. In the following Figure 2 we can see the transformation of input image and segmented image.



**Fig.1. Leaf Segmentation Results**

4. Remove the masked, and then we can get the segment of the disease part. In figure 3 we can see the result of the leaf of lemon after removing the masked area.



**Fig. 2 Removing masked Region**

5. After of opting segments of the disease classify the leaf diseases.
6. These pixels are clustered according to the nearest centers of the clustered. For this used K-Means algorithm and used Euclidean distance between the pixels and the center of the clusters.
7. Based on the cluster values classify with SVM algorithm for the prediction of the type of diseases.

In this survey based on the type of plants the accuracy results showed 80 to 92 % of accuracy for detecting plant disease using k-m.

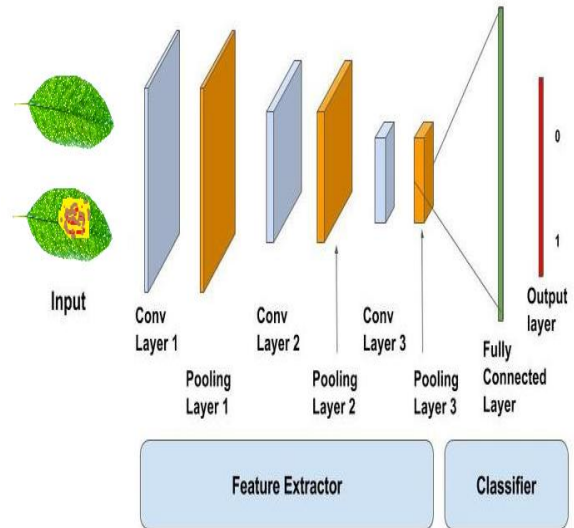
**2.3 Using Deep Learning**

Radanovich et al. [8] surveyed and compare the practical implementation results for the detection of plant disease by using machine learning algorithms and deep learning algorithm. In machine learning used SVM and K-NN algorithms and in Deep Learning classification used Convolution Neural Networks (CNN) algorithm. While implementing the machine learning algorithms using two faces training and testing.

**Table.2 Training and Testing**

Training	Testing
1. Collection of training images	1. Collection of training images
2. Region Segmentation	2. Region Segmentation
3. Feature extraction	3. Feature extraction
4. Feature normalization	4. Feature normalization
5. Classify with machine learning algorithms	using parameters from the training disease detection
6. Save model	5. Disease Prediction

In this survey demonstrated prediction concept using Deep Learning concepts. Deep Learning is the concept which includes three major steps. First one is same like Machine Learning, collection of the input data. But here we can also provide unstructured and unlabeled data like images. In second step we can combine the Feature Extraction and Classification algorithm. In the last phase prediction will same as Machine learning, prediction of the result according to the Classifications.[11-14].



**Fig 3. CNN Model**

Convolution Neural Networks or covnets are neural networks that share their parameters

Types of layers:

- InputLayer
- Convolution Layer
- Activation Function Layer
- Pool Layer
- Fully-Connected Layer

With CNN, there is no need for feature engineering, since CNN models are fed raw data and they are responsible for learning appropriate features. For image recognition problems, convolution networks (CNN) are mostly used. In the practical implementation used 80% data for training and 20% of data for testing.

**3. Overview**

In the following table we can see the accuracy achieved by ML algorithms and DL for predicting the plant disease.

**Table 3: Algorithms and its accuracy**

Sno	Algorithms	Accuracy
1	SVM	0.81
2	KNN	0.78
3	CNN	0.93

**Table 4: Accuracy Results of classification models**

Sno	Author	Goals	Limitations
1	Ramesh et al., 2018	<ul style="list-style-type: none"> <li>➤ Identification of features of the leaves images.</li> <li>➤ Classify with ML algorithms</li> </ul>	<ul style="list-style-type: none"> <li>➤ This concept is used to identify the healthy plant and diseased plant.</li> <li>➤ Cannot identify the type of disease.</li> </ul>
2	Vijai Singh et al., 2017	<ul style="list-style-type: none"> <li>➤ Based on Image segmentation can classify the diseased plants.</li> <li>➤ Low Computation cost</li> </ul>	<ul style="list-style-type: none"> <li>➤ Used very limited data for training and testing</li> </ul>
3	Radovanović et al. 2020	<ul style="list-style-type: none"> <li>➤ Compare the results between ML and DL algorithms.</li> <li>➤ Demonstrate CNN algorithm is best suitable algorithm for prediction plant disease.</li> </ul>	<ul style="list-style-type: none"> <li>➤ Heavy Computation cost required.</li> </ul>

**Conclusion**

Agriculture is basic and important requirement for civilization. For getting more production on crop in agriculture by identifying the crop disease as early as possible and it is also important to identify the type of crop diseases for preventing the spreading of the disease. Many researchers surveyed and proposed identify the crop diseases based on the image processing for helping the real-time gadgets. In this survey we investigate on the different works related to the crop disease identification and detection based on the image processing. In recent studies most of the works are depending on the machine learning and deep learning-based image processing on various

studies of predictions. After analysed the related work on crop detection based on image processing deep learning algorithms achieved best accuracy scores compared to the machine learning algorithms.

**References**

[1].Automatic estimation f crop disease severity levels based on vegetation index normalization  
 [2].H Zhao, C Yang, W Guo, L Zhang, D Zhang - Remote Sensing, 2020 - mdpi.com,  
 [3].Applications of remote sensing in pest monitoring and crop management  
 [4]. K. Ennouri, MA Triki, A Kallel - Bioeconomy for sustainable development, 2020 - Springer

- [5]. Review on Application of drones for Crop Health Monitoring and Spraying Pesticides and Fertilizer, G DEVI, N Sowmiya, K Yasoda, K Muthulakshmi... - J. Crit. Rev, 2020 - researchgate.net
- [6]. S. Ramesh et al., "Plant Disease Detection Using Machine Learning," 2018 International Conference on Design Innovations for 3Cs Compute Communicate Control (ICDI3C), Bangalore, 2018, pp. 41-45, doi: 10.1109/ICDI3C.2018.00017.
- [7]. Vijai Singh, A.K. Misra, Detection of plant leaf diseases using image segmentation and soft computing techniques, Information Processing in Agriculture, Volume 4, Issue 1, 2017, Pages 41-49, ISSN 2214-3173, <https://doi.org/10.1016/j.inpa.2016.10.005>.
- [8]. D. Radovanović and S. Đukanović, "Image-Based Plant Disease Detection: A Comparison of Deep Learning and Classical Machine Learning Algorithms," 2020 24th International Conference on Information Technology (IT), Zabljak, Montenegro, 2020, pp. 1-4, doi: 10.1109/IT48810.2020.9070664.
- [9]. United Nations, Department of Economic and Social Affairs, Population Division (2019). World Population Prospects 2019: Highlights (ST/ESA/SER.A/423).
- [10]. Savary, Serge, et al. "The global burden of pathogens and pests on major food crops." *Nature ecology & evolution* 3.3 (2019): 430.
- [11]. Mohanty, Sharada P., David P. Hughes, and Marcel Salathé. "Using deep learning for image-based plant disease detection." *Frontiers in plant science* 7 (2016): 1419.
- [12]. Fujita, E., et al. "A practical plant diagnosis system for field leaf images and feature visualization." *International Journal of Engineering & Technology* 7.4.11 (2018): 49-54.
- [13]. Haralick, Robert M., Karthikeyan Shanmugam, and Its' HakDinstein. "Textural features for image classification." *IEEE Transactions on systems, man, and cybernetics* 6 (1973): 610-621
- [14]. Wang P., Chen K., Yao L., Hu B., Wu X., Zhang J., et al. (2016)." Multimodal classification of mild cognitive impairment based on partial least squares".