IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

RESEARCH CHALLENGES ON PLANT DISEASE DETECTION

[1] E.Gangadevi , [2] Dr. R.Shoba Rani

[1]Research Scholar, [2]Associate Professor,

[1,2]Department of Computer Science and Engineering, Dr. MGR Educational and Research Institute, Deemed to be University, Chennai, India-600095

Abstract: As agriculture is a very important field in the Economy and social life, there exist many challenges to overcome. In this aspect, plant disease detection is one of the major issues to farmers for early detection. Many research methods had come up with the lowest turnaround time and to produce an accurate result for early identification of plant disease through proposing many algorithms and techniques. Even though there are many pitfalls found then and there, each of the techniques has given its unique results by choosing appropriate datasets and algorithms. In this paper, we studied and investigated existing techniques for the detection of early plant diseases.

Index Terms: **Image processing, Machine learning, Deep learning, Plant disease**

1. INTRODUCTION

Diseases found in plants lead to a major loss in production as well as the economic growth of our country. Due to this reason, the quality and quantity of the agricultural products were decreased. Also, monitoring the large scale of farm fields is a challenging task for the farmers. Each plant disease's symptoms are different according to its behavior and its characteristics. As a human, having complete knowledge about each plant's diseases and their symptoms is a very tedious process. For detection and identification of plant diseases through the naked eye is followed in the traditional methods. Reducing the usage of pesticides is the main objective of the farmers which then relatively increases the quantity and quality of the food products. To overcome all the various challenges in the agricultural sector, an automatic disease detection system at the earlier stage was introduced at different levels with different emerging technologies.

In the earlier days, plant leaf disease classification was carried out through digital image processing techniques by applying image processing methods, analyzing various datasets, and through a complete understanding of them. The crops in the field must be continuously monitored from the sowing stage to the harvesting stage. The new technologies which are applied to detect plant diseases are very accurate, faster, inexpensive and time-consuming method when comparing with the traditional methods such as naked-eye observation. All these reasons made the researchers come up with different emerging techniques to solve the issues and help the farmers. This paper has taken a review on various techniques used by different researchers to detect plant diseases according to the different parameters like the nature of the plant, moisture in the soil, weather, type of irrigation, etc.

The image processing technique aims at recognizing the area or part of the leaf affected and measure those areas. Then, the image processing technique identifies the shape of the observed infected area from the leaf or from the stem of the plant which finds the color of that region. With all these findings, the technique will pre-process the images and will predict or detect whether the plant is affected by disease or not. In addition

to this observation, the infected region and the percentage of the affected region will also be measured. Smart farming has been introduced nowadays with image processing techniques and with machine learning and deep learning algorithms.

Even with the usage of image processing techniques, accurate, cost-effective, and time-consuming classification and detection of plant diseases are not possible. To overcome this situation, here comes Machine learning which is a subfield of Artificial Intelligence. The main objective of machine learning is to completely understand the infrastructure and different parameters of the data and fit the understandable data into models which can be in a good and consistent format ready for humans to make use of it. Machine learning algorithms make the computers train the inputs of data and apply statistical analysis and mathematical models to build models from sample data. This leads to automatic decision-making methods based on the data inputs.

The classification in machine learning depends upon the task assigned. Those classifications depend on the learning methods and learning the feedback to the given system is developed. The commonly used machine learning methods are supervised learning which trains algorithms that are completely based on the labeled inputs and output data managed by humans. The other machine learning method is unsupervised learning that uses unlabeled data to the algorithm which allows finding the structure of the input data. There are challenges in choosing the algorithms and the type of inputs we make use of for statistical and analytical purposes to come up with the desired solution. In this case, identifying plant diseases using its images from the dataset is a challenging task than using numerical data.

Deep learning which is a subset of machine learning applies a multi-layered structure of algorithms usually referred to as neural networks. In addition to machine learning strategies, deep learning helps us to cluster, classify and perform regression tasks. The property of feature extraction is an advantage in deep learning which leads to categorizing the data into several classes. Without the help of human intervention, identifying plant diseases can be carried out with advancements in technology.

2.RELATED WORK

In this related work, we have taken some papers related to the detection and prediction of plant diseases using various techniques. Various techniques with advanced algorithms to detect plant diseases have been developed by many researchers. When we detect the affected areas of the plant in the very earlier stage, the farmers can take appropriate action to prevent the farm which will avoid a huge loss. In this regard, a very fast, accurate, time-consuming, and less expensive system must be developed.

Guo et al. [1] proposed a plant disease identification method based on a deep learning algorithm applied for smart farming. Region Proposal Network (RPN) is used to identify and localize the leaves in their surroundings. Then segmentation is done for the images based on the RPN algorithm. These images are given as input into the transfer learning model which is trained by the dataset of affected leaves. Apart from this, other symptoms are also identified here. They concluded that this deep learning algorithm yields a better accuracy when comparing with the traditional methods.

Sagar et al [2] described a survey on different methods followed to detect plant diseases. They proposed that plant disease can be detected based on the type of family plants by using segmentation and classification. They have categorized the disease detection as monocot and Dicot family plant. A complete description of different types of plant families, various techniques of segmentation, feature extraction, and classification are studied here. They have concluded that the k-means clustering method is used for segmentation and the SVM classifier is used in the classification process by most of the researchers.

Rao et el. [3] describes a detection accuracy level method by forming a hybrid approach that helps them to overcome the issue of low illumination rate and noise issue. This method is implemented through the MATLAB simulation tool and the dataset they followed is from the Plant village database. The proposed model explains different steps for pre-processing, feature extraction, and classification methods. Fuzzy classifier method is also applied here.

Savita N. Ghaiwat et al. [4] proposed a survey in various classification techniques that is applied for finding plant diseases. From the given input, a k-nearest neighbor method is used for predicting the

diseases. In the case of non-linear separable training data, it is difficult to evaluate optimal parameters in SVM. This is the drawback found in this method.

Nanjesh et al. [5] detected unhealthy plant leaves using the RGB image acquisition method. This method converts RGB to the HIS format of an input image. By this method, mask, filter, and clearing green pixels are followed. With the help of a genetic algorithm from a soft computing technique, the classification of plant disease is done. The drawback found in this paper is we cannot give more inputs to compare the image on a large scale.

S.Ram and Anand [6] investigated the strategy of detecting and classifying apple fruit diseases which are divided into three steps as segmentation, feature extraction, and classification. For image segmentation, k-means clustering is followed. The different features are taken from the segmented image in which they are classified with the help of Multiclass Support Vector Machine. The method they suggested can be applied only to apple fruit plants and not to the other types of plants.

Yao et al. [7] presented a classification model for identifying the Rice disease spots. This approach uses medial filtering as noise reduction in the pre-processing phase. The feature extraction phase comprises extracting shape features and for efficient classification of disease, an SVM classifier is utilized.

M.Bhange et.al [8] A web-based tool has been developed to identify fruit diseases by uploading fruit images to the system. Features extraction has been done using parameters such as color, morphology, and CCV (color coherence vector). Clustering has been done using the k-means algorithm. SVM is used for classification as infected or non-infected. This work achieved an accuracy of 82% to identify pomegranate disease.

G. Saradhambal, et.al (2018) [9] proposed an approach to produce a system for automatic plant disease detection. Research was carried out to predict the infected area of the leaves by applying a k-means clustering algorithm and the Otsu's classifier. Both the shape and texture features were extracted in the proposed work. The shape-oriented features that were extracted in this work included area, color axis length, eccentricity, solidity, and perimeter, whereas the texture-oriented features were contrasted, correlation, energy, homogeneity and mean [10]. And lastly, classification in this research was done using a neural network-based classifier.

3.TABLE OF COMPARING EXISTING RESEARCH PAPERS

Paper	Authors /Reference	Year	Algorithm	Goals	Future directions
1	Muhammad Hammad Saleem et al [10]	2019	Deep Learning	Deep learning models are applied to visualize various plant diseases with performance metrics	DL models should be improved and a study should be made to understand types of plant diseases, classes, and size of datasets, learning rate, etc
2	Tete et al [11]	2017	K-means cluster	Applying digital image processing techniques like preprocessing, segmentation, feature extraction, and classification	To apply neural network concepts to get better accuracy
3	Surendar Kumar et al [12]	2015	Review on KNN, RBF, PNN, BPN, SVM	To make use of digital image processing techniques. Back propagation is also used to learn, train and transfer the functions	Because SVM is non-linearly separable, it is challenging to find optimal parameters. To overcome this, neural network concepts can be applied for detecting plant diseases.
4 99.35%	Mohanty et al [13]	2016	AlexNet and GoogleNet	Choice of deep learning architecture, training mechanisms, dataset type, and training set distributions are focused.	Image data extracted from a smartphone can be used with location and time data to improve accuracy.
5	Khirade [14]	2015	Basic Image processing Techniques	Types of plant diseases are classified and using the leaf image, digital processing steps are applied to detect plant diseases.	Can better apply advanced processing of image techniques.
6	Jiang [15]	2020	Transfer Learning Algorithm	Open-source data set is used to train the ResNet model to produce good versatility and efficient training.	The transfer Learning method seems to be very accurate, can follow the same in the future also.
7	Yan et al [16]	2020	Deep Learning Algorithm-RPN Algorithm	RPN is applied to localize the leaves. Reduces the domination of diseases on agricultural products.	Can apply a neural network to activate zero initial sets for different leaves to speed up the training process and the iteration over time.

8	Nagaraju et al [17]	2019	Deep learning techniques	With the help of the leaf images, a deep learning model is applied and received different results according to the algorithm	Need a systematic engineering design to develop extracting featured models
9	Sun et al [18]	2018	Image processing technology	To improve accuracy and intelligence.	Can modify the independent and dependent variables for greater potential

Table 1. Summary of existing research papers

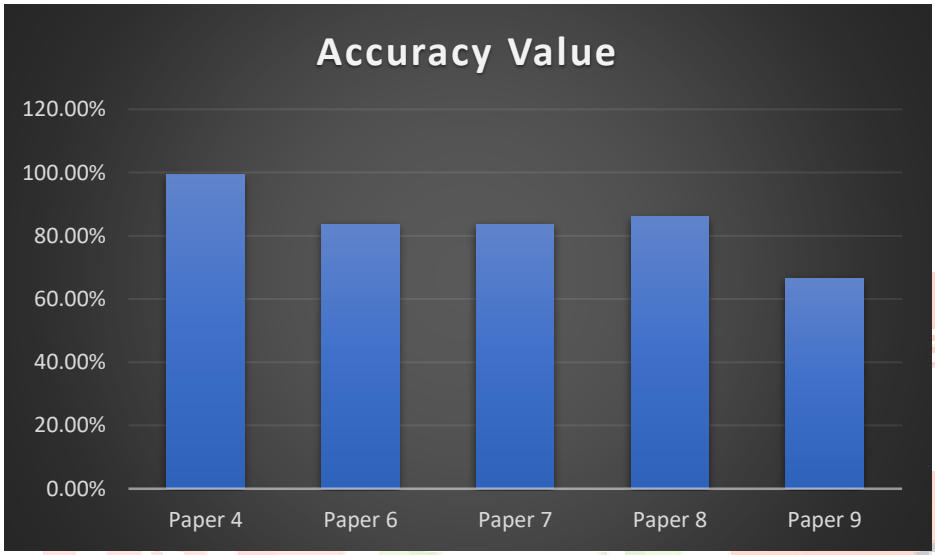

Fig1. Accuracy value for the reviewed papers

4. METHODOLOGIES FOR PLANT DISEASE DETECTION

There are many methods implemented to detect plant diseases. Some of the researchers concentrated on observing the type of the dataset. They intended to feed the quality data as input to train the model. Some of the researchers aimed to capture the images of the leaves through drones, a camera with sensors, and satellite images. This kind of research focused on monitoring the crop field on a continuous basis. On the other hand, researchers studied various types of diseases to analyze what kind of disease is attacked to the plants and how they can come out of it. Some researchers focused on implementing the datasets with different kinds of algorithms. They compared the result with different kinds of techniques to arrive at the maximum accuracy. In this aspect, we have concentrated on different methodologies followed by the researchers to detect plant diseases at the very early stage.The common methodologies applied by various researchers from the traditional method to the current methods are,

- Digital Image Processing Techniques
- Machine Learning Techniques
- Deep Learning Techniques

4.1 DIGITAL IMAGE PROCESSING TECHNIQUES

This technique is a traditional method of detecting plant diseases with the help of computers by analyzing digital images through different algorithms. During the processing of images, this technique can avoid noise and distortion by applying many algorithms to its input data. Digital image processing comprises of the following steps as shown in Fig 2.

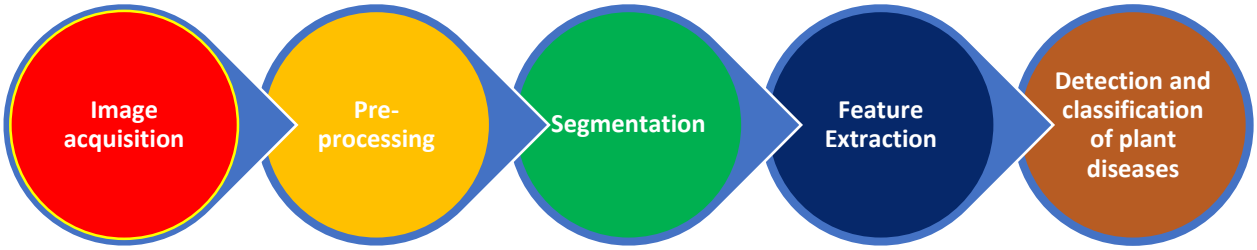

Fig 2. Sequential steps of Digital Image Processing

As a very initial step, image acquisition helps to confirm whether the image has been taken clearly from the datasets or been captured through the camera with proper quality. The format of an image can be electronic or in a physical way. This step assures us to get ready with the input data, especially images associated with our research. The second step is data pre-processing which is an important method followed in Data mining. The missing values are found and the noise found in image data is removed.

Segmentation helps us to categorize our work in an easier manner by dividing the areas into different segments according to the features presented in our datasets. To work with the huge volume of data, Feature extraction helps to minimize the number of resources required. Sufficient accuracy will also be maintained through feature extraction. As a result of all the above steps, through the appropriate algorithm, an easier way of detecting and classifying plant diseases can be achieved finally.

4.2 MACHINE LEARNING TECHNIQUES

The Machine Learning approach is widely used in many applications. Especially, in agriculture, the different Machine learning algorithms help the farmers to take the right decision to prevent and increase crop yields. Machine Learning algorithms such as SVM, CNN, k-means clustering, Minimum distance classifier, etc are applied. Training data will be inputted into the learning algorithm, which is completely studied to produce a hypothesis function. With this hypothesis function, any predicted result can be derived. This technique can be applied to predict the disease of plants beforehand. The model is shown in Fig 3.

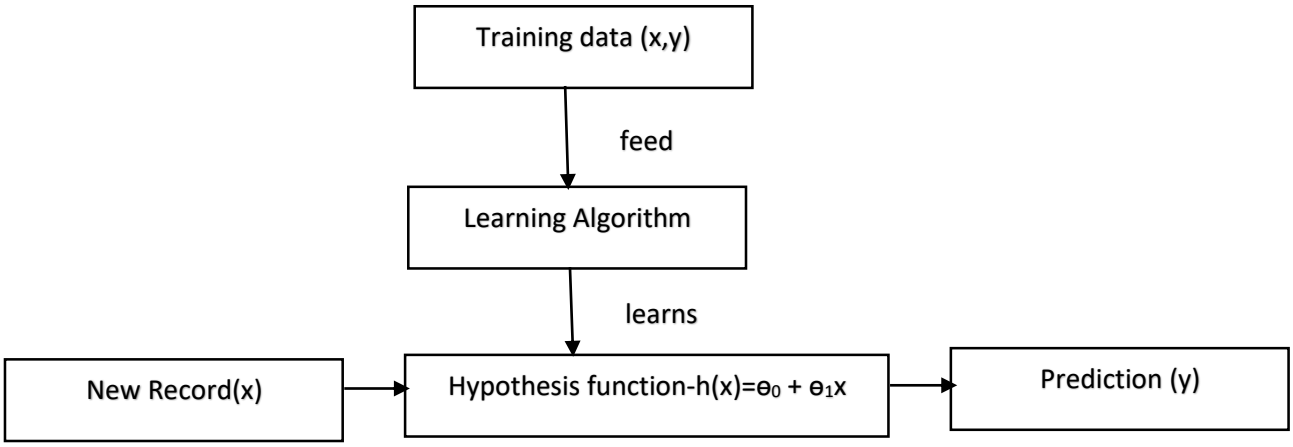

Fig 3. Hypothesis model

The applications in Data science follow the statistical and mathematical approaches before processing with machine learning algorithms. These Statistical and mathematical models help researchers to understand the data completely so as to apply those data for appropriate algorithms and to achieve better results. The statistical and mathematical model helps to build the data and to understand them clearly. The regression model in statistical analysis aims to understand the relationship between the dependent variable usually called an outcome variable and with many independent variables referred to as features. While predicting the results, there might be errors due to the noisy data which we fed or with any other reasons. Those error metrics can be applied to count the error rate for a forecasting model. The three different error metrics followed in the statistical approach are MSE, MAE, and RMSE. Linear regression error metrics and Bivariate analysis for a hypothesis are explained below.

Adjusted R-squared score=1- $\left[\frac{(1-R^2)\,(m-1)}{m-n-1}\right]$

As we used regression models to find the relationship between two categories, bivariate analysis is also a simple tool to measure the relationship among two sets of values. It also defines the categorical and quantitative values for both the input variable and independent variables as shown in Table 2. This kind of comparison helps researchers to produce a quality in the accuracy and in the result they display.

	Response/Dependent Variable		
Input/Independent Variable		Categorical	Quantitative
	Categorical	Chi-square test	Analysis of variance(ANOVA)
	Quantitative	Chi-square test	Correlation

Table 2. Bivariate analysis

Bivariate analysis: covariance

Covariance is a measure of how much two features vary together.

- It is a measure used to analyze the linear relationship between two variables.
- It answers the questions of how these two variables behave when paired.
- The positive value indicates the increase in a linear relationship.
- The negative value indicates the decrease in a linear relationship.

$$\text{Cov}(X,Y)=\frac{\sum_{i=1}^{n}(X_i-X)(Y_i-Y)}{n-1}$$

Covariance provides the direction. On the other hand, correlation provides direction and strength. Covariance result doesn't have any upper or lower bound and its size is dependent on the scale of the variables and it is not standardized. Correlation is always between -1 and +1. Its scale is independent of the scale of the variables and it is standardized. The model building process using Machine learning techniques are described in fig.4

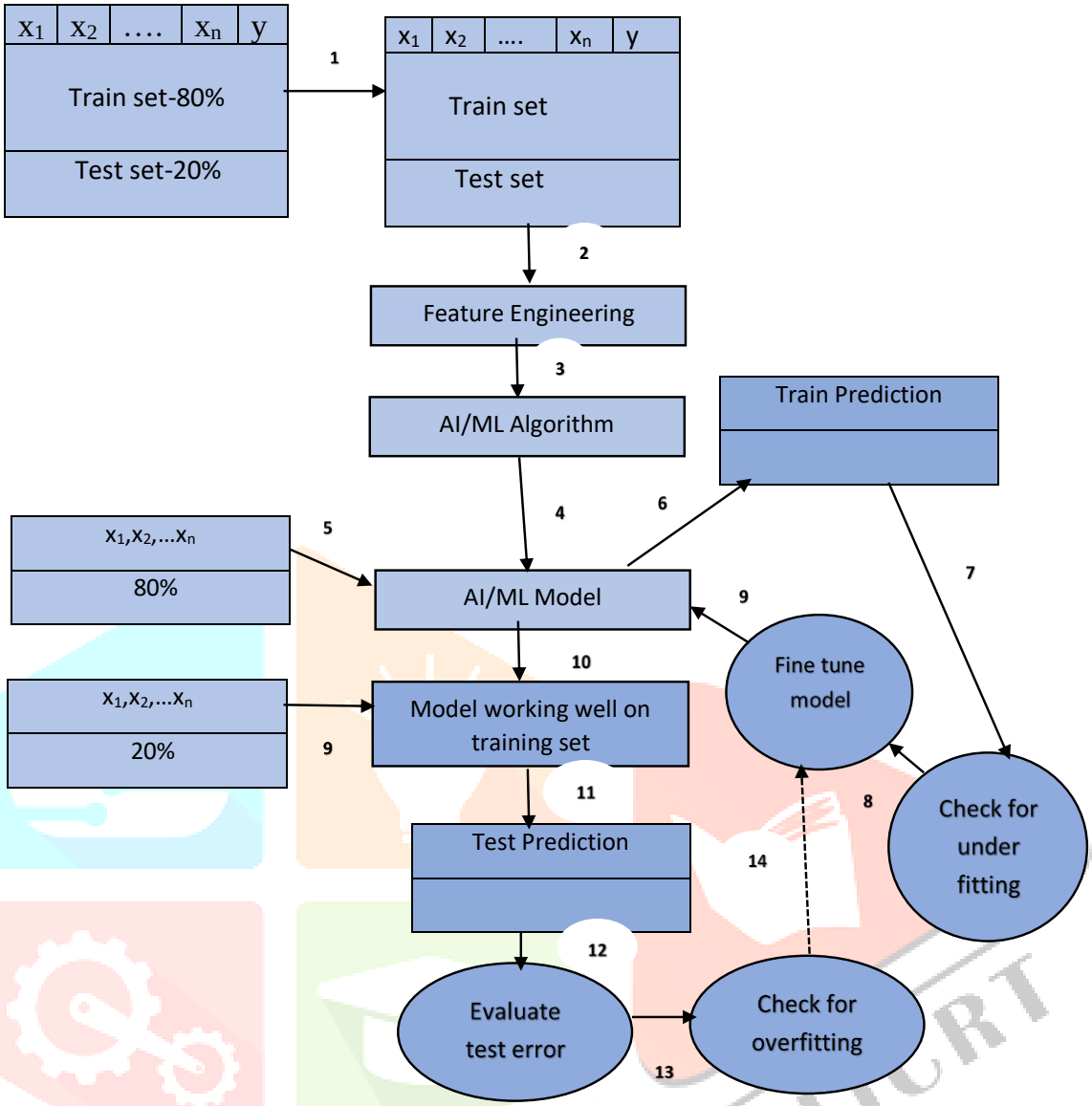

Fig 4. Model building process

4.3 DEEP LEARNING TECHNIQUES

With a large potential, Deep learning efficiently produces a flexible and intelligent method for image processing and data analysis which can yield an accurate result. Deep learning is a subset of machine learning, can be applied in various applications which are dependent on the images. So, in the case of detecting plant diseases, Deep learning plays a predominant role. The various techniques applied here are Convolutional neural networks, Region-based convolutional neural networks, faster R-CNN, Deep learning architectures, Visualization techniques, etc. In related works of this paper, we have summarized some research papers and their methodology followed with deep learning techniques. By observing the graph given, it is understood that deep learning models give higher accuracy of results when comparing with any other traditional models.

Abbreviations

KNN-K-Nearest Neighbour

RBF-Radial Basis Function

PNN-Probabilistic Neural Networks

BPN-Backpropagation Network

SVM-Support Vector Machine

CNN -Convolutional Neural Network

MSE-Mean Squared Error

RMSE- Root Mean Squared Error

MAE-Mean Absolute Error

5. CONCLUSION

This review explained the approaches using Image processing techniques, Machine learning models, and deep learning techniques for detecting plant leaf diseases. Also, the importance of using statistical methods and mathematical models is also explained. The main goal of us is to compare the data from a huge collection of datasets, analyze them, pre-process them and apply appropriate algorithms to come up with the desired solution. From all the research papers, it is also observed that most of them have used deep learning models. Because deep learning is suitable to work with images and to produce better accurate results. The identification of plant diseases at the earlier stage can be improved by applying computer vision concepts and natural language processing and with some advanced software tools.

6. REFERENCES

[1]- Yan Guo, Jin Zhang, Chengxin Yin, Xiaonan Hu, "Plant disease identification based on deep learning algorithm in smart farming", Hindawi,Discrete dynamics in nature and society,Volume 2020,Article id 2479172,11 pages.

[2] Sagar Patil, Anjali Chandavale," A survey on methods of plant disease detection", International Journal of Science and Research(IJSR), ISSN:2319-7064, Volume 4,Issue 2,Feb 2015.

[3]- Anusha Roa, S.B.Kulkarni ,"A Hybrid approach for plant leaf disease detection and classification using digital processing methods", International Journal of Electrical Engineering and Education.DOI: 10.1177/0020720920953126

[4]- Savita N. Ghaiwat, Parul Arora. Detection and Classification of Plant Leaf Diseases Using Image Processing Techniques: A Review. International Journal of Recent Advances in Engineering & Technology. ISSN (Online): 2347 – 2812, Volume-2, Issue-3,2014.

[5] - Tejoindhi M.R, Nanjesh B.R, JagadeeshGujanuru Math, AshwinGeetD'sa "Plant Disease Analysis Using Histogram Matching Based On Bhattacharya's Distance Calculation" International Conference on Electrical, Electroniocs and Optimization Techniques(ICEEOT)-2016 .

[6]- Shiv Ram Dubey, Anand Singh Jalal, "Detection and Classification of Apple Fruit Diseases using Complete Local Binary Patterns", International Conference on Computer and Communication Technology IEEE, 2012, pp. 346-351.

[7] - Yao Q, Guan Z, Zhou Y, Tang J, Hu Y, Yang B (2009) Application of support vector machine for detecting rice diseases using shape and color texture features. In: IEEE international conference on engineering computation ICEC, Hong Kong, pp 79–83.

[8] - Bhange, M., Hingoliwala, H.A., 'Smart Farming: Pomegranate Disease Detection Using Image Processing', Second International Symposium on Computer Vision and the Internet, Volume 58, pp. 280-288, 2015

[9] - Saradhambal, G., Dhivya, R., Latha, S., Rajesh, R., 'Plant Disease Detection and its Solution using Image Classification', International Journal of Pure and Applied Mathematics, Volume 119, Issue 14, pp. 879-884, 2018

[10] Muhammad Hammad Saleem,Johan Potgieter," Plant disease detection and classification by Deep learning", MDPI, 2019,8;doi:10.3390/plants8110468.

[11] Trimi Neha Tete,Sushma Kamlu,"Plant disease detection using different algorithms", Proceedings of 2nd international conference on research in intelligent and computing in Engineering, pp.103-106.DOI:10.15439/2017R24

[12] Surender Kumar, Rupinder Kaur,"Plant disease detection using image processing -A Review", International Journal of Computer Applications (00975-8887),Volume 124-No.16,August 2015.

[13] Sharada P.Mohanty,David P.Hudhes, Marcel Salathe," Using Deep Learning for Image Based Plant disease detection",Frontiers in plant science, doi:10.3389/fpls.2016.01419

[14] Sachin D.Khirade,A.B.Patil,"Plant disease detection using Image Processing",IEEE,International Conference on Computing Communication Control and Automation, doi: 10.1109/ICCUBEA.2015.153

[15]Han Jiang,"Research on plant leaf disease identification based on Transfer learning algorithm", Journal of Physics,doi: 10.1088/1742-6596/1576/1/012023

[16] Yan Guo, Jin Zhang, Chengxin Yin, Xiaonan Hu, "Plant disease identification based on deep learning algorithm in smart farming", Hindawi,Discrete dynamics in nature and society,Volume 2020,Article id 2479172,11 pages.

[17] M.nagaraju, Priyanka Chawla,"Systematic Review of Deep learning techniques in plant disease detection", Springer,doi:10.1007/s13198-020-00972-1

[18] Guiling Sun, Xinglong Jia, Tianyu Geng,"Plant diseases recognition based on image processing technology", Hindawi, Journal of electrical and computer engineering,Volume 2018,Article ID 6070129,7 pages.