The Scientific Temper (2023) Vol. 14 (1): 8-13
E-ISSN: 2231-6396, ISSN: 0976-8653
Doi: 10.58414/SCIENTIFICTEMPER.2023.14.1.02
https://scientifictemper.com/

RESEARCH ARTICLE

The effectiveness of machine learning and image processing in detecting plant leaf disease

Ashish Nagila*, Abhishek K. Mishra

Abstract

In our daily lives, the agricultural sector is crucial. Therefore, it is crucial to be clear about the steps taken to identify any diseases on agricultural plant leaves. Plant leaf disease is a significant issue or contributor to crop losses in an agricultural context. Some farmers are able to know every disease name and how to prevent them as it becomes increasingly crucial to recognize the sickness. Different plant leaf diseases appear during various seasons. This problem can be resolved using a deep learning-based approach by identifying the affected regions in plant leaf images, enabling farmers to better comprehend the disease. The primary goal of this research is to survey several image-processing methods for detecting plant diseases and to compare them. India is an agricultural nation, and the majority of its people depend on agriculture for a living. Focusing on farming with modern technology is essential to ensuring their comfort and ease of living. Crop productivity may be greatly increased by introducing new technologies. An autonomous plant disease detection method using image processing and a neural network methodology can be utilized to solve issues with plant and agricultural diseases. Plants can contract a wide range of illnesses. Different patterns are needed to detect various disorders.

Keywords: Data Augmentation, Feature extraction, Image acquisition, Plant leaf identification, Segmentation.

Introduction

Agriculture play an important role in our daily life activities. Agriculture is the backbone on the Indian economy, also most important part in our life. In India, over a billion people population is there out of which, over 70% of the population lives in rural areas, with 30% of country workforce. Internet of Things (IoT), AI, and unmanned aerial vehicles (UAV) improvements are combined with supporting agricultural areas in detecting plant leaf illnesses and reporting it properly to the appropriate people with the proper accuracy ranges. Due to the difficulties, farmers face everyday in modern society, no one is interested in farming or agriculture. So that all members of the youthful generation relocate to contemporary cities to live safely and avoid obstacles in the agricultural field. Plant disease can directly lead to stunted growth, causing bad effect on yields. Plants acts as an important resource for everyone in term of food. So, it is important to see that the plant is not affected by the disease. If disease occurs, then it is necessary to detect disease. Many diseases affect the plant leaf in different seasons according to it weather condition and climate where the leaf gets infected. In India, due to diversity in geographical topologies, there is variation in environmental conditions like temperature, humidity, and rainfall. Insect pest takes a huge toll on crops causing severe loss to the farming community.

Crop damage would result in a significant loss in productivity, which would have an impact on the economy. The most vulnerable component of plants, the leaves, are where disease symptoms first appear (Kiani E *et al.*, 2017). The crops must be inspected for illnesses from the beginning of their life cycle until they are ready to be harvested. Initially, specialists manually observed crop fields using the time-consuming method of traditional naked-eye surveillance to keep a check on the plants for illnesses (Surampalli Ashok *et al.*, 2020). A variety of methods have been used in recent years to create autonomous and semi-automatic plant disease detection systems. As of yet, these methods have proven to be quicker, less expensive, and more precise than farmers' manual observation, which has been the norm (Arun Pandian J *et al.*, 2019). This encourages

Department of CSE, IFTM University, Moradabad, Uttar Pradesh, India

***Corresponding Author:** Ashish Nagila, Department of CSE, IFTM University, Moradabad, Uttar Pradesh, India, E-Mail: ashishnagila01@gmail.com

How to cite this article: Nagila, A., Mishra, A.K. (2023). The effectiveness of machine learning and image processing in detecting plant leaf disease. The Scientific Temper, **14**(1): 8-13

Doi: 10.58414/SCIENTIFICTEMPER.2023.14.1.02

Source of support: Nil

Conflict of interest: None.

Received: 20/01/2023 **Accepted:** 12/02/2023 **Published** : 25/03/2023

scientists to develop technological systems that are more sophisticated and don't need human participation to detect plant diseases. This paper aims to explore several methods for identifying plant diseases and discuss them in terms of various parameters. The sections of the paper are as follows. The significance of plant disease detection is briefly discussed in the first section. The second section addresses methods employed and the latest work that has been done in the field. The basic methodology used to construct the disease detection system is included in Section 3. The fourth section concludes this essay and offers suggestions for the future.

Analysis

Using a fuzzy decision maker outside in a strawberry field, E. Kiani *et al.*. attempted to identify disease-infected leaves. Plant disease identification and segmentation had an overall accuracy of 97%, took 1.2 seconds to process, and was highly accurate (Kiani E *et al.*, 2017).

By using image processing methods like DTW and GLCM and segmentation to separate the leaf image into smaller parts of texture and determine the boundary of the tomato leaf and label on it, the authors of this paper (Surampalli Ashok *et al.*, 2020) have classified the tomato leaf disease using convolutional neural network algorithm. Convolutional neural networks are used for classification and detection to provide the results.

The author employed the pre-trained transfer learning model, i.e., VGG16, Res-Net, in this paper (Arun Pandian J et al., 2019) to describe the plant leaf disease dataset using image processing and PCA colour segmentation, GAN (Generative Adversarial Network), and neural style transfer technique. The Inception dataset, which combines deep learning with colour and position augmentation, has the best classification performance of any dataset. The data modification method using neural style transfer and PCA algorithms 89% of lessons learned through data manipulation were successfully transferred.

By H. Ali *et al.* research intends to employ the E colour difference technique to segregate the disease-affected area and uses colour histogram and textural cues to diagnose diseases, obtaining a 99.9% overall accuracy (Ali H *et al.*, 2017). Different classifiers have been employed, including fine KNN, Cubic SVM, boosted tree, and bagged tree classifiers. In comparison to the other classifiers, the bagged tree classifier outperforms them, scoring 99.5% 100, and 100% accuracy on the RGB, HSV, and LBP features, respectively. Classifiers using fine KNN, cubic SVM, and boosted trees fared well, reaching 88.9, 90.1, and 50.90% accuracy, respectively.

In this paper (Eftekhar Hossain Md. Farhad Hossain *et al.*, 2019), they used the KNN (K-nearest neighbors) classifier to perform machine learning algorithm approaches. The leaf disease image's texture is used to extract features for classification. They used a KNN classification approach to perform an algorithm classifier to categories the diseases Alternaria, anthracnose, bacterial blight, leaf spot, and canker of various plant species. This author has contrasted various segmentation methods and machine algorithms. And also performed different parameters like SSIM, MSE, DSC.The proposed approach can successfully detect and recognize the selected diseases.

The You Only Look Once (YOLO) technique was used in this paper (Sachin C *et al.,* 2020) to train the necessary network using a variety of different vegetable photos. Before training, the photographs were pulled from internet, and to get the output of labels from images, bounding boxes were performed around the vegetables. They compared many factors using various vegetable datasets, including cucumber, green apple, and green capsicum. When the network receives the inputs, the output will accurately label the recognized vegetable with the expected class categories and show bounding boxes around it.

This paper (Rajleen Kaur *et al.,* 2015) used the SVM method to implement on two datasets—one for training and the other for testing—after taking the image and preprocessing it. Then the image is processed to separate the portions of the black and background pixels. The third is used to segment a picture's healthy and unhealthy portions and detect disease. Only neural networks and principal component analysis (PCA) have been utilized in this, and they have been compared to both the standard SVM and the enhanced SVM with modifications.

In this paper (Ms. Kiran R *et al.*, 2019), image preprocessing techniques were used on unhealthy citrus leaf disease areas. Implementing image analysis and classification techniques for extracting and classifying citrus disease is the main task they have completed. The purpose model has been broken down into four sections: image preprocessing, which includes converting RGB to various color spaces, image enhancement; segmenting the region of interest using K-means clustering for statistical use in identifying the severity of plant leaf defects, feature extraction, and classification. The SVM algorithm was used to extract the texture feature using statistical GLCM and the color feature using any method of mean values.

K-means segmentation was chosen by D.A. Bashish *et al.*. to divide the leaf image into four clusters using squared Euclidean distances. For both color and texture features, the Color Co-occurrence technique is used for feature extraction (Bashish *et al.*, 2010). Classification is finally completed using a neural network detection strategy based on the Back Propagation method. The total system's accuracy in identifying and classifying diseases was determined to be 93%.

M. Bhange *et al.* By uploading fruit images to the system, a web-based tool for diagnosing fruit illnesses has been created (Bhange *et al.*, 2015). Features have been extracted

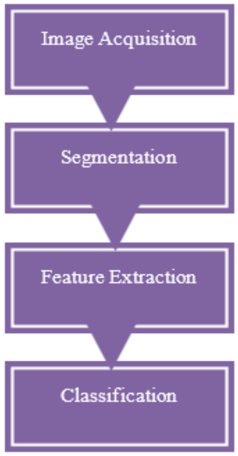

Figure 1: Plant disease detection system phases.

using variables including colour, morphology, and CCV (colour coherence vector). The k means algorithm has been used to perform clustering. SVM is used to determine whether something is contaminated or not. This research identified pomegranate illness with an accuracy rate of 82%.

Methods

Plant Disease Detection Process

As seen in Figure 1, there are fundamentally four stages to the plant disease detection procedure. Images are acquired during the first phase using a digital camera, a mobile device, or the internet. The image is divided into varying numbers of clusters in the second phase, each of which can be treated using a different technique. The following phase includes strategies for feature extraction, while the final step discusses the classification of disorders.

Image Acquisition

Plant leaf photos are collected during this phase using digital devices like cameras, smartphones, and other devices with the necessary resolution and size. Additionally, photos from the web can be used.

The application system developer is solely responsible for creating the picture database. The final phase of the detection system's classifier is more effective thanks to the image database (Omrani *et al.*, 2014).

Images Segmentation

This step seeks to make an image's representation more understandable and straightforward to examine (Gharge *et al.*, 2016). This stage is the essential method of image processing since it serves as the foundation for feature extraction. Images can be segmented using a variety of techniques, including thresholding, Otsu's algorithm, and k-means clustering. Based on a collection of features, the k-means clustering divides objects or pixels into K number of classes. By minimizing the sum of squares of distances between the items and their related clusters, classification is accomplished (Singh J *et al.*, 2018).

Table 1: Comparison table of different methods

S. No	*Title*	*Methodology*	*Datasets*	*Year*	*Accuracy in Percentage*
1.	Identification of plant disease infection using soft-computing: Application to modern botany. (Kiani E *et al.*, 2017)	Fuzzy Decision Maker	Strawberry leaf	2017	93
2.	Tomato Leaf Disease Detection UsingDeep Learning Techniques. (Surampalli Ashok *et al.*, 2020)	Convolutional neural network,segmentation	Tomato leaf	2020	98.05
3.	Data Augmentation on Plant Leaf Disease Image Dataset Using ImageManipulation and Deep Learning Techniques. (Arun Pandian J *et al.*, 2019)	GAN (Generative Adversarial Network) and Neural style Transfer and PCA	Mendeley datasets (healthy and diseasedleaf)	2019	89
4.	Symptom based automated detection of citrus diseases using color histogram and textural descriptors.	Classifiers using fine KNN, cubic SVM, boosted tree, and bagged tree classifiers	Lemon Leaf	2017	99.9
5.	A colour and texture-based approach fordetection and classification of plant leaf disease using KNN classifier. (Eftekhar Hossain Md. Farhad Hossain *et al.*, 2019)	KNN classifier, Feature extraction with GLCM	Alternata, Anthranoseand bacterial Blight	2019	96.76
6.	Vegetable Classification Using You Only Look Once Algorithm. (Sachin C *et al.*, 2020)	You only look Once (YOLO)	Cucumber, green apple, green capsicum	2020	61.6
7.	An Enhancement in Classifier SupportVector Machine to Improve Plant Disease Detection. (Rajleen Kaur *et al.*, 2015)	Enhanced SVM, PCA, neural network	Crop leaf image.	2015	84
8.	Unhealthy Region of citrus Leaf detection Using Image processingtechniques. (Ms. Kiran R *et al.*, 2019)	K-mean, SVM	Citrus leaf	2019	94
9.	A Framework for Detection and Classification of Plant Leaf and Stem Diseases	Network detection strategy based on the Back Propagation method	Wheat Leaf Image	2010	93
10.	Smart Farming: Pomegranate Disease Detection Using Image Processing	k-means algorithm for clustering and SVM.	Pomegranate Leaf	2015	82

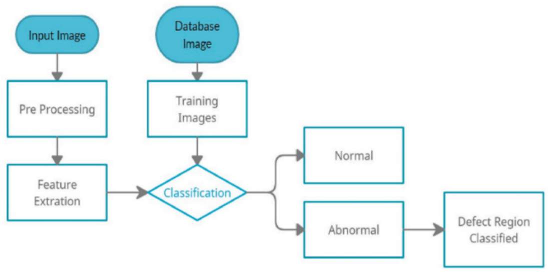

Figure 2: General steps for Leaf disease detection.

Feature Extraction

The area of interest is the outcome thus far after segmentation. Therefore, the characteristics from this area of interest must be extracted in this stage. To interpret a sample image, these features are required. Color, form, and texture are all possible bases for features (Dey *et al.*, 2016). The majority of academics now want to use textural traits to identify plant diseases. The system can be developed using a variety of feature extraction techniques, including histogram-based feature extraction, the gray-level co-occurrence matrix (GLCM), the color co-occurrence approach, and the spatial grey-level dependence matrix. The GLCM approach is a statistical technique for classifying textures.

Classification

Identifying whether the input image is healthy or unhealthy is the goal of the classification step. If it is determined that the image is diseased, some previously published papers have further categorized it into various disorders. A MATLAB software procedure, also known as a classifier, must be built in order to perform classification. Researchers have used a variety of classifiers recently, including Naive Bayes, support vector machines (SVM), artificial neural networks (ANN), back propagation neural networks (BPNN), and decision tree classifiers. The SVM classifier is found to be the most popular one. Although each classifier has benefits and drawbacks, SVM is a straightforward and reliable technique (Gavhale *et al.*, 2020).

Leaf Disease Detection using Machine Learning Techniques

Traditional farming methods put farmers' lives in danger, especially in places prone to drought. These methods involve manually collecting data, coping with bad weather, sprinkling pesticides on illnesses, and more. Given the state of conventional farming today, a pressing demand for predicted data in agriculture may help farmers recognize and address current issues. We would like to provide a strategy that uses a Decision Tree Classifier to forecast cotton crop illnesses based on temperature, soil moisture, and other variables to assist them in solving their issues.

The vast volumes of data that are essentially obtained nowadays from the crops must be estimated and used to the fullest. Figure 2 depicts the preprocessing process for the input image.

It focuses on the use of controlled machine learning techniques for maize plant disease identification, including Naive Bayes (NB), Decision Trees (DT), KNN, SVM, and Random Forest (RF). The categorization methodologies are examined and contrasted to choose the model that would forecast plant diseases with the maximum degree of accuracy.

- The system of classification is probabilistic.
- High independence presumption theorem.
- The value of one feature has no bearing on the value of another function.

The effectiveness of the classification approach to preprocessing benefits in the resulting portions and the good levels of prediction accuracy are both demonstrated in this paper.

NB Classifier

Gaussian Each attribute's continuous values are calculated by Naive Bayesian, and their distribution is determined by a Gaussian distribution, also known as the Normal Distribution. The bell-shaped bell curve that represents the regularity of the mean values and is calculated from the results of the Gaussian distribution is,

$$p\,(x_1/y) = \frac{1}{\sqrt{2\Pi\sigma_y^2}} \exp\left(-\frac{(x_i-\mu_y)^2}{2\sigma_y^2}\right) \rightarrow$$

KNN Classifier

KNN classifier is used for both classification and regression difficulties, but it typically aids in classification concerns. KNN is a distribution tree algorithm diagnosis. Non-parametric, or the structure derived from the attributes of the data collection, is used when there is no imagination for the distribution of the data. KNN is used for prediction when data sets don't follow fictitious mathematical theories. KNN doesn't require any training data to continue. The Slow Learning (SL) technique is referred called as such since all training data is helpful for testing data.

Decision Tree Classifier

DT classifier is the dominant and acceptable method for the classification and prediction process. DT has a structure of tree-based concept, where every internal node describes a feature test, every branch describes as a test result, and every terminal node holds a label of data. Decision trees can generate understanding rules quickly. A decision tree is a value-based method that is accessible and helpful because of its easily understandable flowchart estimation.

SVM Classifier

SVM analysis for classification and regression can be helpful. Within the two classes of data, SVM determines the hyperplane with the increased margin. The support vectors are the hyperplane vectors. Considering crucial circumstances, SVM may create a margin of hyperplane that divides the hyperplane vector into two classes that don't intersect. However, this is not always the case, so this classifier will look for support vector hyperplanes that have larger margins and fewer classification errors.

RF Classifier

RF is a supervised algorithm that can handle classification and regression techniques. It is mostly worried about problems with classification. A forest is a managed group of trees, and a strong forest has many trees. The RF also finds the dataset-based decision tree approach, just as the decision trees. It obtained the forecast results from every branch of the tree before selecting the better course of action. It is known as an ensemble approach, and it performs better than a single DT classifier since average performance reduces over-fitting.

MLP Classifier

The idea that is based on regression is known as an MLP. By using this technique, non-linear based learners are converted, changing the input dataset. Changes from the input data, which is a property that is linearly distinct. A hidden layer is created from the input data layer.

The MLP Classifier only employs one hidden layer, failing which it functions as an artificial neural network. Even the use of many concealed layers serves the objective of classification.

Result

The three main stages of plant disease detection are feature extraction, segmentation, and classification.

The study contrasted SVM and Naive Bayes, and many more classification methods for predicting disease. In terms of execution time and accuracy, the outputs of these two classifiers are contrasted.

Accuracy

The number of right points is what is used to define accuracy. divided by the sum of the points multiplied by 100

$$\text{Accuracy} = \frac{\text{Number of points correctly classified}}{\text{Total Number of points}} * 100$$

Execution Time

Execution time is defined as the interval between the start and stop times of an algorithm.

Execution time = End time of algorithm- start of the algorithm

Figure 3 compares the accuracy of different classifiers such as NB, KNN, DT, SVM, RF and MLP classifiers for detecting plant diseases.

As seen in Figure 4, the performance study compares the execution times of the different classifiers.

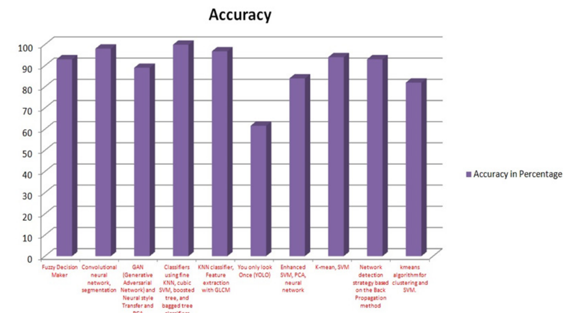

Figure 3: Accuracy comparisons.

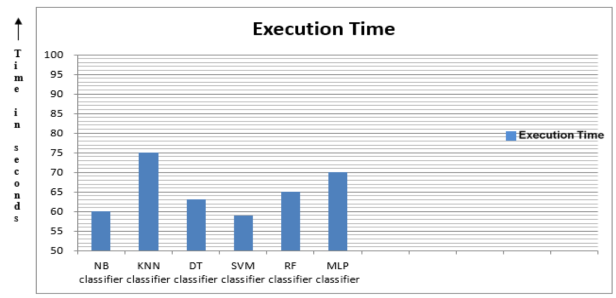

Figure 4: Execution time.

Discussion

In this study, many classifiers techniques for leaf disease are analyzed and compared as NB, KNN, DT, SVM, RF, and MLP. We compared and analyzed these techniques on the basis of two factor that are accuracy and execution time. When compared, SVM is found to be more accurate at detecting plant diseases.

Also, the SVM classifier executes faster. Additionally, a number of methods/mappings for identifying the disease symptoms were summarized. Here, the recent advancement of deep learning technologies for the diagnosis of plant leaf illnesses. We believe that this work will be a valuable resource for researchers attempting to identify plant diseases. Additionally, a comparison between deep learning and machine learning methods is done. Despite the fact that there has been a lot of notable development in recent years, there are still some research gaps that need to be filled in order to put into practice efficient strategies for plant disease identification.

Conclusion

The methods of plant disease detection utilizing image processing are reviewed and summarized in this work used by several scholars in recent years. BPNN, SVM, K-means clustering, Otsu's method, CCM, and SGDM were the main approaches used. These methods are employed to determine if the leaves are sick or healthy. The automation of the detection system employing intricate photos recorded in harsh environmental circumstances, such as outdoor lightning, presents a number of problems in this procedure. This paper comes to a conclusion that, despite having some drawbacks, these disease detection techniques demonstrate efficiency and accuracy sufficient to enable them to operate the system designed for the detection of leaf diseases.

Acknowledgment

The authors are highly thankful to the IFTM University. I would also like to express my deepest gratitude to Dr. Abhishek Kumar Mishra for helping in the research. This endeavor would not have been possible without his support. Last but not least I would also like to sincerely thank Mrs. Ritu Nagila for her technical support to complete this research.

Conflict of Interest

The authors declare that this research was carried out without any financial or commercial ties that might be viewed as a potential conflict of interest.

Reference

Ali, H., Lali, M.I., Nawaz, M.Z., Sharif, M., Saleem, B.A., 'Symptom based automated detection of citrus diseases using color histogram and textural descriptors', Computers and Electronics in Agriculture, Volume 138, pp. 92-104, 2017

Arun Pandian J, G Geetharamani, Annette B, Data Augmentation on Plant Leaf Disease Image Dataset Using Image Manipulation and Deep Learning Techniques, 9th International Conference on Advanced Computing (IACC), IEEE, (2019), pp. 199-204.

Bashish, D.A., Braik, M., Ahmad, S.B., 'A Fremework for Detection and Classification of Plant Leaf and Stem Diseases', International Conference on Signal and Image Processing, pp. 113-118, 2010.

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.

Dey, A.K., Sharma, M., Meshram, M.R., 'Image Processing Based Leaf Rot Disease, Detection of Betel Vine (Piper BetleL.)', International Conference on Computational Modeling and Security, Volume 85, pp. 748-754, 2016

Eftekhar Hossain Md. Farhad Hossain, Mohammad Anisur Rahaman, A Color and Texture Based Approach for the Detection and Classification of Plant Leaf Disease using KNN Classifier, International Conference on Electrical, Computer and Communication, (ECCE), IEEE, (2019), pp. 7-9

Gavhale, K.R., Gawande, U., 'An Overview of the Research on Plant Leaves Disease Detection using Image Processing Techniques', IOSR Journal of Computer Engineering, Volume 16, Issue 1, pp. 10-16, 2014 Jashraj Karnik, Anil Suthar (2020) "A survey of plant leaf disease identification using deep learning techniques" Dogo Rangsang Research journal, ISSN 2347-7180, Volume 10, Issue 12, Page No 91-95, UGCcare group I.

Gharge, S., Singh, P., 'Image Processing for Soybean Disease Classification and Severity Estimation', Emerging Research in Computing, Information, Communication and Applications, pp. 493- 500, 2016

Kiani, E., Mamedov, T., 'Identification of plant disease infection using soft-computing: Application to modern botany', 9th International Conference on Theory an d Application of Soft Computing, Computing with Words and Perception, Volume 120, pp. 893-900, 2017

Kiran R. Gavhale, Prof. Ujwalla Gawande, Mr. Kamal O. Hajari, Unhealthy Region of Citrus Leaf Detection Using Image Processing Techniques, International Conference for Convergence of Technology, IEEE, (2019), pp. 1-6.

Melike Sardogan, Adem Tuncer, Yunus Ozen, Plant Leaf Disease Detection and Classification based on CNN with LVQ Algorithm, 3rd International Conference on computer science and engineering, IEEE, (2018) pp. 382-385

Muammer Turkoglu, Davut Hanbay, Abdulkadir Sengur, Multi-model LSTM-based convolutional neural networks for detection of apple diseases and pests, Journal of Ambient Intelligence and Humanized Computing, (2019).

Omrani, E., Khoshnevisan, B., Shamshirband, S., Saboohi, H., Anuar, N.B., Nasir, M.H.N., 'Potential of radial basis function-based support vector regression for apple disease detection', Journal of Measurement, pp. 233-252, 2014

P. Ganesh, K. Volle, T. F. Burks, S. S. Mehta, Deep Orange: Mask R-CNN based Orange Detection and Segmentation, IFAC (International Federation of Automatic Control) Hosting by Elsevier, (2019), pp. 70- 75.

Rajleen Kaur, Dr. Sandeep Singh Kang, An Enhancement in Classifier Support Vector Machine to Improve Plant Disease Detection, 3rd International Conference on MOOCs, Innovation and Technology in Education (MITE), IEEE, (2015) pp. 135-140.

Sachin C, N Manasa, Vicky Sharma, Nippun Kumaar A. A, Vegetable Classification Using You Only Look Once Algorithm, International Conference on Cutting-edge Technologies in Engineering (ICon-CuTE) Shri Ramswaroop Memorial University, India, IEEE, (2020), pp. 101-107.

Singh, J., Kaur, H., 'A Review on: Various Techniques of Plant Leaf Disease Detection', Proceedings of the Second International Conference on Inventive Systems and Control, Volume 6, pp. 232-238, 2018

Surampalli Ashok, Gemini Kishore, Velpula Rajesh, S. Suchitra, S.G Gino, B. Pavithra, Tomato Leaf Disease Detection Using Deep Learning Techniques, Proceeding of the Fifth International Conference on Communication and Electronic Systems (ICCES), IEEE, (2020), pp. 979-983.

H. Ali, M.I. Lali, M.Z. Nawaz, M. Sharif, B.A. Saleem, Symptom based automated detection of citrus diseases using color histogram and textural descriptors, Computers and Electronics in Agriculture, Volume 138, 2017, Pages 92-104

D. Al Bashish, M. Braik and S. Bani-Ahmad, "A framework for detection and classification of plant leaf and stem diseases," *2010 International Conference on Signal and Image Processing*, Chennai, India, 2010, pp. 113-118, doi: 10.1109/ICSIP.2010.5697452.

Manisha Bhange, H.A. Hingoliwala, Smart Farming: Pomegranate Disease Detection Using Image Processing, Procedia Computer Science, Volume 58, 2015, Pages 280-288