JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

RESEARCH ON PLANT LEAVES DISEASE DETECTION USING MACHINE LEARNING

Ms.SWETHA R , VAISHNAVI D, SOWBARNIGA A, SHRIDHARSHINI M
Assistant Professor,Principal, Student (UG),Student (UG), Student (UG)
Department of Computer Science and Engineering,
Vel Tech High Tech Dr. Rangarajan Dr. Sakunthala Engineering College, Chennai, India

Abstract: Considering the recent rate of population expansion, it is anticipated that by 2050, global crop productivity will need to increase by two fold. One of the biggest barriers to reaching this productivity goal is diseases and pests. Thus, it is crucial to create effective techniques for the automatic identification, detection, and forecasting of agricultural crops' pests and illnesses. Machine Learning (ML) strategies can be used to extract relationships and information from the data being worked on in order to execute such automation. This study reviews the literature on machine learning (ML) applications in the agricultural sector, with a particular emphasis on tomato crops. It focuses on the tasks of disease and pest classification, detection, and prediction. This research aims to promote the development of precision agriculture and smart farming by encouraging the creation of methods that will enable farmers to reduce their usage of chemical pesticides and herbicides while maintaining and raising the productivity and quality of their crops.

KEYWORDS: Agriculture; plant diseases; machine learning; detection and classification; dataset collection; smart farming

1 INTRODUCTION

Due to the very high infant death rate, the global human population grew very gradually until the year 1600. Around 1700 signified the achievement of the first billion, which was followed by a second billion in 1927 and the third billion in 1961. The global population reached its seventh billion in 2017. The primary cause of the recent decades' rapid population expansion has been improved medical care. The United Nations predicts that the world's population will reach 9.7 billion in 2050 and 10.9 billion by 2100.

Over the past few decades, rapid population development has led to increased demand for agricultural items, resulting in significant cultivation expansion. Crop yield production has to double by 2050 to keep up with the growing demand for food, biofuels, and animal products from increasing numbers of people. More importantly, diseases and insect pests continue to pose a threat to crop productivity. It is estimated that plant diseases and insect attacks cause 20% to 40% of the world's annual crop production to be lost, costing the economy $220 billion and $70 billion, respectively.

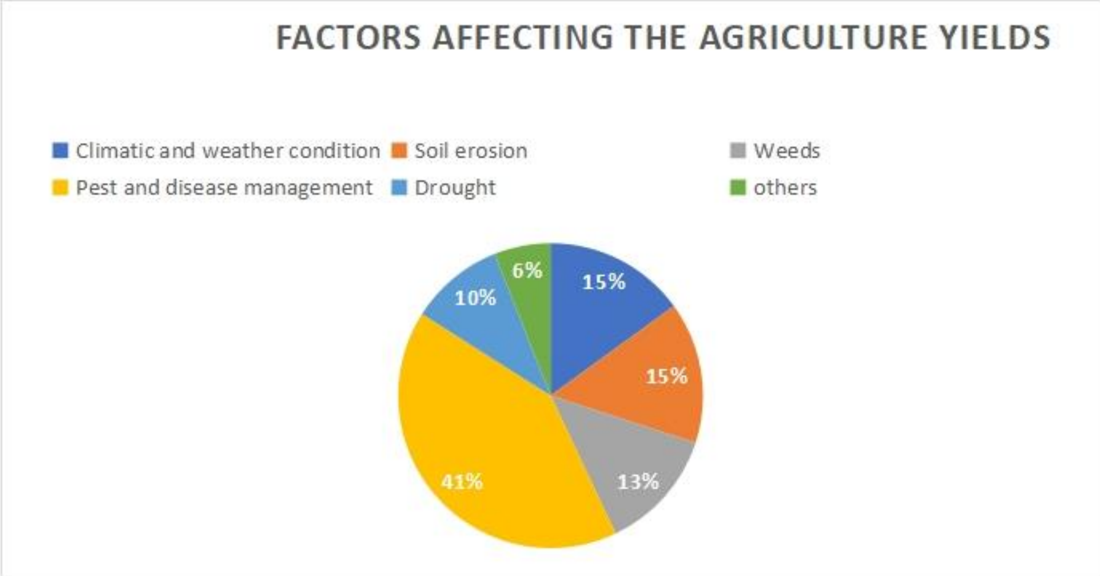
Farmers have traditionally used chemical treatments and pesticides to keep pests at away.But, it also have some negative effect such as human health and increased environmental damage to soil and groundwater. The traditional method of detecting the diseases such as visual inspection by farmers and experts. It is not suit for huge farms; it takes expertise and time. Therefore, automated techniques for crop monitoring and forecasting are needed to get beyond the drawbacks of manual detection.Machine learning (ML) has been employed to identify illnesses in medical imaging, classify enormous datasets, drive self-driving cars and conduct academic research in physics .ML-based agricultural applications are still in their early stages, but they are already showing prospective. Disease categorization from images can be done using Convolutional Neural Network (CNN) designs for various plants and diseases .Traditional machine learning approaches, such as feature extraction and classification, are widely employed in the field of plant disease diagnosis. These methods extract characteristics including color, texture, and shape to build a classifier that can distinguish between healthy and unhealthy plants.

ML-based techniques have aided the creation and growth of smart farming. This article also contribute projects progress, develop, and succeed.

.

FACTORS AFFECTING THE AGRICULTURE YIELDS:

This below pie chart shows the factors that affect the yields in agriculture. Among other factors, pest and disease management had the most impact, based on a review of numerous surveys.



LITERATURE REVIEW:

Wei, D., Juan, Z., Rujing, W., Chengjun, X., Tianjiao, C., and Wancai, L. (2019): The study concentrated on deep learning-based intelligent identification systems for disease and insect pests. In comparison to conventional techniques (Frontiers), they investigated the application of convolutional neural networks (CNNs) for image-based plant disease identification and found notable gains in processing speed and accuracy.

Yujian, L., Njuki, S., Too, E. C., and Yingchun, L. (2019): In order to identify plant diseases, this study compared various deep learning model fine-tuning techniques. They assessed different CNN topologies and emphasized how transfer learning can improve model performance on plant disease datasets (Frontiers).

Wang, B. (2022): Wang's research used cutting-edge deep learning algorithms to identify insect pests and crop illnesses. In order to detect diseases in crops like tomatoes and maize, the research focused on applying deep convolutional neural networks (DCNNs), which provide excellent accuracy rates and robust performance across a variety of environmental situations (Frontiers).

Potgieter, J., Saleem, M. H., and Arif, K. M. (2019): A thorough examination of deep learning models for the identification and categorization of plant diseases was given in this review. They talked about how different DL architectures, such as AlexNet, VGGNet, and ResNet, are implemented and how well they work to diagnose plant illnesses from picture data. In addition, the study suggested future strategies for improving model transparency and performance (MDPI) and highlighted a number of research needs.

Rumpf, T., Dehne, H.-W., Oerke, E.-C., Mahlein, A.-K., Steiner, U., and Plümer, L. (2010): This earlier study, which classified illnesses in wheat and sugar beet using support vector machines (SVM) and other machine learning algorithms, set the foundation for contemporary methods in plant disease detection. The potential of machine learning in agricultural applications was shown by their research (MDPI).

Dixit, P., Singh, N., Dubey, S. R., and Gupta, J. P. (2016): They looked into the use of CNNs in particular and deep learning techniques for the identification and categorization of apple leaf diseases. The study focused on how to increase disease identification accuracy (MDPI) by combining deep learning with image processing techniques.

Mahajan, G. R., Sahoo, R. N., and S. K. Behera (2018): The use of machine learning and remote sensing methods for the early diagnosis of plant diseases was the main focus of this work. They demonstrated the potential of merging various technologies for precise disease detection (MDPI) by using hyperspectral imaging and machine learning models to diagnose illnesses in rice crops.

METHODOLOGY:

The methodology for plant disease detection using machine learning entails a thorough approach that includes data collection, preprocessing, feature extraction, and model selection. CNN models function well with image databases for object recognition and classification. CNNs offer advantages, but they also have disadvantages, such as the requirement for large datasets and a protracted training phase.

DATASET:

The PlantVillage collection, which is available to the public , has 54,305 images representing 38 classifications of different plant illnesses. We separated the dataset into training, testing, and validation subsets in order to do our experimental study. More precisely, pre-trained models were trained on 80% of the PlantVillage dataset; the remaining 20% was set aside for testing and validation. There were 54,305 samples in all, of which 43,955 were set aside for training, 4,902 for validation, and 5,488 for testing purposes, covering all 38 classes of plant diseases. This is a thorough explanation of the dataset split.

PREPROCESSING:

Preprocessing is an essential initial step in leaf disease detection. In order to increase the machine learning model's accuracy in classifying diseases preprocessing is done before feature extraction.This step involves a number of procedures, including image resizing, noise removal, colour modification, morphological operations, and segmentation. Resizing the images to standard format 256 X 256 pixels due to the uneven size of input images.Noise should be removed before feature extraction in order to increase the accuracy.Noise is removed by various filtering techniques such as Wiener filter, median filter and Gaussian filter.

SEGMENTATION:

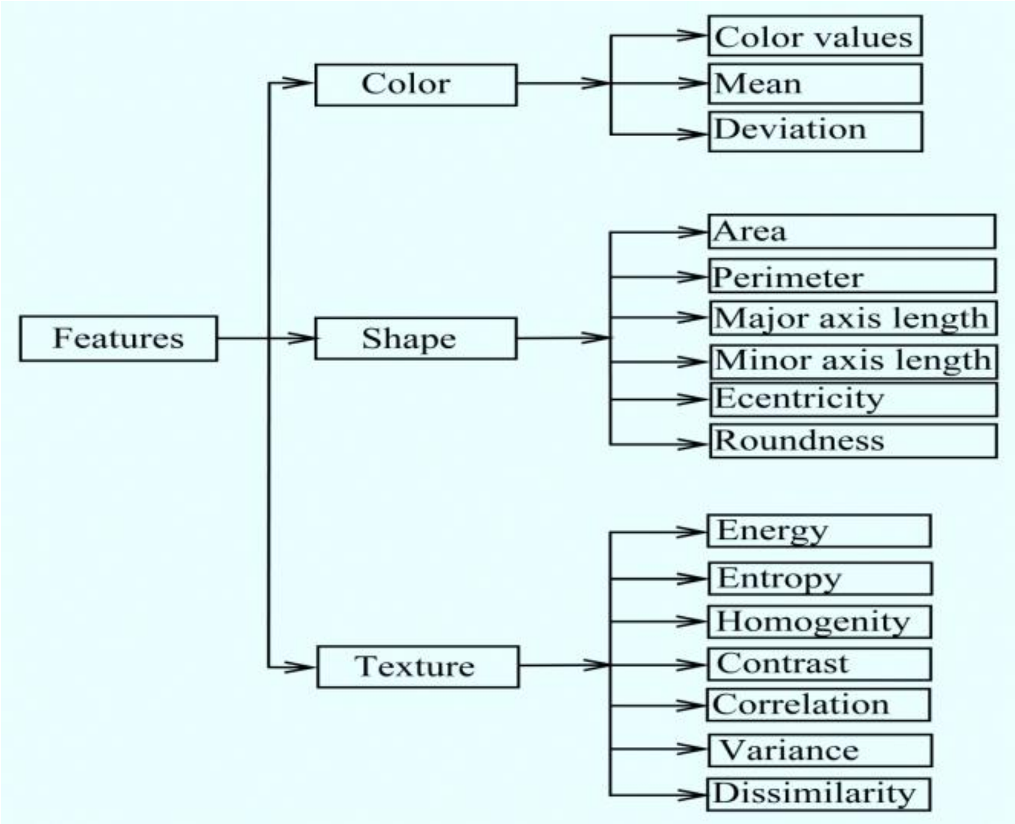
Segmentation is partitioning the images into discrete parts or segments according to the affected area. After completing this stage, a picture should be produced in which each pixel is a part of a distinct segment that represents a different area of the leaf. There are numerous segmentation techniques, including Otsu's segmentation and K-means clustering.

FEATURE EXTRACTION

One of the fundamental processes in machine learning is feature extraction. It is applied to explain significant data in mathematical form and in order to distinguish the classes by classification. There are two types of approaches in feature extraction: deep learning methods and handmade methods.

It is separated into shape features, color features, and texture features for handmade techniques. These techniques rely on manually extracting features from plant images. Color characteristics depend upon the various shades and intensities of color used to identify the diseased area, whereas shape features include minor/major axis length, area, circumference, deviation, etc. Several methods have been used for texture features, including local binary pattern (LBP), gray-level run-length method (GLRLM), etc.

While using deep learning techniques, all contextual and global data can be extracted to find the relevant features. These techniques are more robust and have better identification accuracy. Certain techniques used in the early research on plant disease identification rely on deep learning for feature extraction, including convolution neural networks (CNNs). Images are first entered into the CNN model, and these characteristics are subsequently loaded into a support vector machine, also known as an SVM or other machine learning classifier.



MACHINE LEARNING ALGORITHMS:

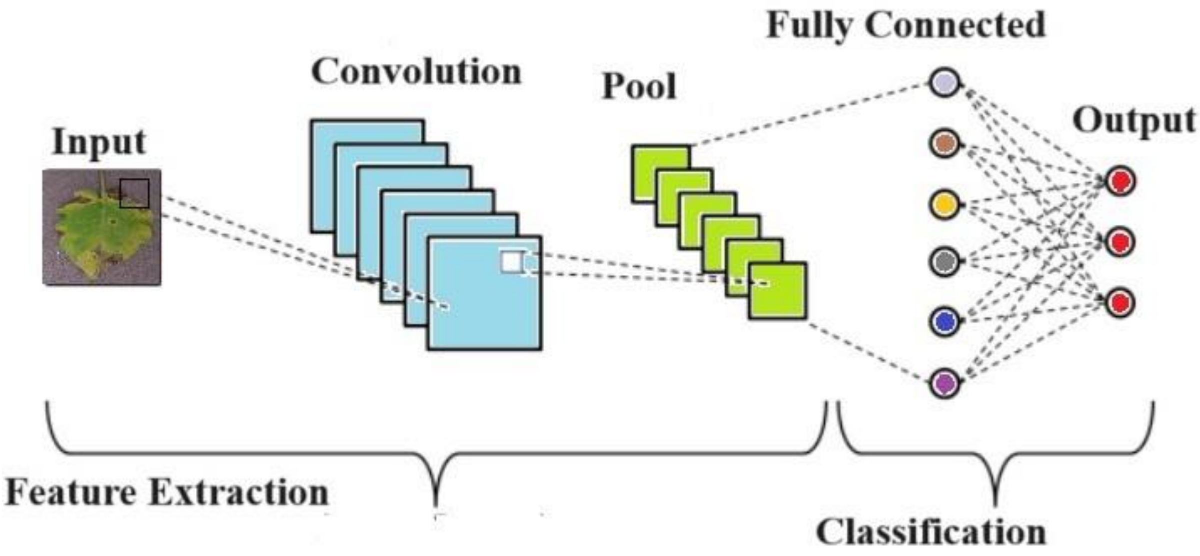
CONVOLUTIONAL NEURAL NETWORK(CNN):

CNN is a type of Deep Learning neural network architecture commonly used in Computer Vision. An artificial intelligence field called computer vision makes it possible for a computer to comprehend and analyze an image or other visual data. The utilization of so- many layers called convolutional layers in this kind of network allows for a hierarchical feature extraction process, wherein more specialized and complicated characteristics are extracted in deeper layers and simpler features, such edges, are extracted in the initial layers. Using pooling layers reduces the dimension of the input. Using these high-level features, fully connected neural networks function as classifiers and are typically positioned above the convolutional and pooling layers.
There are three types of layers in a normal Neural Network. They are

- Input layers
- Hidden layers
- Output layers

CNN ARCHITECTURE:

There are multiple layers in a convolutional neural network: input, pooling, convolutional, and fully connected.The convolutional layer processes the input image to extract features; the pooling layer minimizes computation by down sampling the image; and the fully connected layer produces the last result.To find the optimal filter, the network employs backpropagation and gradient descent.



Fig(1):Schematic representation of working of convolutional neural network

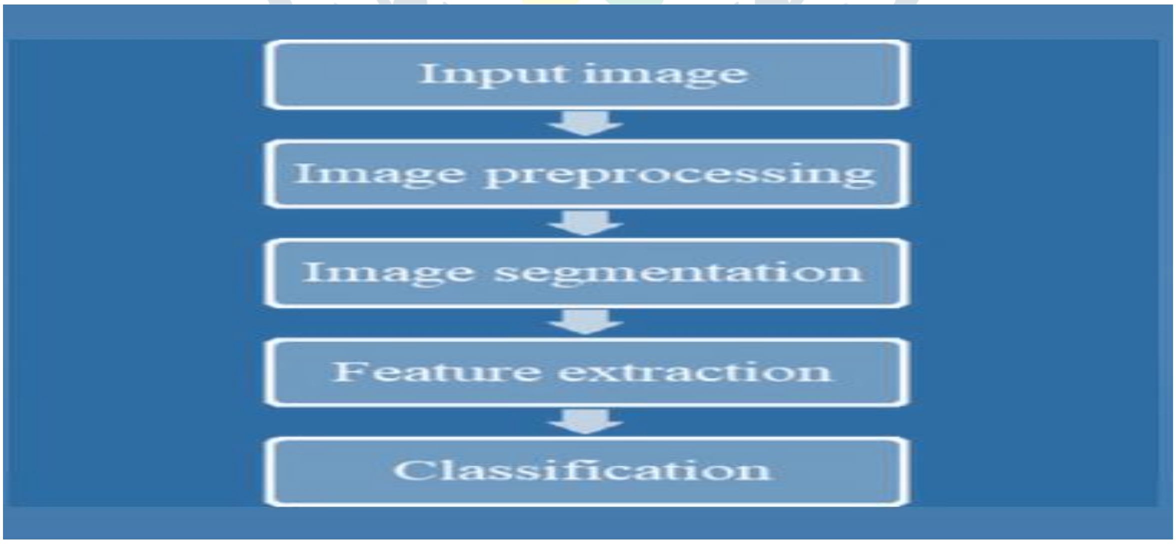
FEATURE SELECTION:

Feature selection is the crucial step in plant disease detection. The purpose of this particular phase is to prevent feature redundancy. This is accomplished by choosing the most discriminating traits and removing repetitious and unnecessary information. There are numerous feature selection techniques, including genetic algorithms (GA) and correlation-based feature selection (CFS). Correlation based feature selection is Selecting the features or based on the related images with the input image or target image.

CLASSIFICATION:

Classification is the process of classifying and categorizing plant diseases based on a trained dataset.There are two types of classification: supervised and unsupervised. K-nearest neighbor (KNN), support vector machine (SVM), random forest (RF), logistic regression (LR), decision tree (DT), naïve Bayes (NB), artificial neural network (ANN), and probabilistic neural network (PNN) are a few of the frequently used categorization approaches.

FLOWCHART:



Fig(2):representation of flow diagram

FUTURE SCOPE:

This survey provided an overview of current research on the use of machine learning (ML)-based methods for forecasting, detecting, and classifying illnesses and pests.

Early disease diagnosis would assist farmers in increasing crop productivity and addressing the costly domain expert issue. There are a number of gaps in the existing reservoir of literature, some of which are indicated as areas for further investigation in the quest to identify plant diseases. An important research challenge is the gathering of big datasets including a diverse range of photographs from various geographical regions. We also draw the conclusion from the survey that illness detection will be less reliable if the disease symptom varies greatly between phases of infection. Future research aims to create a dependable lightweight deep CNN.

CONCLUSION:

Plant diseases have a significant global influence on agriculture and the economy. As a result, a thorough analysis of earlier research on plant disease detection and classification utilizing AI-based methods is needed. The purpose of this study is to review the most recent work that has been done on applying deep learning and machine learning to the identification of plant diseases. As this study concludes, despite the introduction of a great deal of research, there are still some unanswered questions that need to be answered.

Future recommendations should address issues with this plant's disease identification system. It is advised to use a data augmentation strategy that creates a variety of plant photos and data sharing for the inadequate variety dataset.Other approaches should be used to improve accuracy while solving segmentation challenges. To improve the accuracy of the algorithm used to identify plant illnesses, additional plant photos should be gathered for diseases with comparable characteristics. The multiclassifier should be used to identify more than one illness in the image when there are multiple plant diseases. Lastly, effective techniques for improving image quality are advised for issues with plant leaf photos in order to increase classification accuracy.

REFERENCES:

1.Roser, M. Future population growth. In Our World in Data; University of Oxford: Oxford, UK, 2013.
2. Fróna, D.; Szenderák, J.; Harangi-Rákos, M. The challenge of feeding the world. Sustainability 2019, 11, 5816. [CrossRef]
3. Ray, D.K.; Mueller, N.D.; West, P.C.; Foley, J.A. Yield trends are insufficient to double global crop production by 2050. PLoS ONE 2013, 8, e66428. [CrossRef]
4. FAO. The Future of Food and Agriculture: Trends and Challenges; FAO: Rome, Italy, 2017.
5. Deutsch, C.A.; Tewksbury, J.J.; Tigchelaar, M.; Battisti, D.S.; Merrill, S.C.; Huey, R.B.; Naylor, R.L. Increase in crop losses to insect pests in a warming climate. Science 2018, 361, 916–919. [CrossRef] [PubMed]
6. Food and Agriculture Organization of the United Nations FAOSTAT Pesticides Use. Available online: https://www.fao.org/ faostat/en/#data/RP/visualize (accessed on 10 January 2022).
7. Kartikeyan, P.; Shrivastava, G. Review on emerging trends in detection of plant diseases using image processing with machine learning. Int. J. Comput. Appl. 2021, 975, 8887. [CrossRef]
8. Kamilaris, A.; Kartakoullis, A.; Prenafeta-Boldú, F.X. A review on the practice of big data analysis in agriculture. Comput. Electron. Agric. 2017, 143, 23–37. [CrossRef]
9. Rajpurkar, P.; Irvin, J.; Zhu, K.; Yang, B.; Mehta, H.; Duan, T.; Ding, D.; Bagul, A.; Langlotz, C.; Shpanskaya, K.; et al. Chexnet: Radiologist-level pneumonia detection on chest X-rays with deep learning. arXiv 2017, arXiv:1711.05225.
10. Wang, L.; Lee, C.Y.; Tu, Z.; Lazebnik, S. Training deeper convolutional networks with deep supervision. arXiv 2015, arXiv:1505.02496.
11. Soni, A.; Dharmacharya, D.; Pal, A.; Srivastava, V.K.; Shaw, R.N.; Ghosh, A. Design of a machine learning-based self-driving car. In Machine Learning for Robotics Applications; Springer: Berlin/Heidelberg, Germany, 2021; pp. 139–151.
12. Carleo, G.; Cirac, I.; Cranmer, K.; Daudet, L.; Schuld, M.; Tishby, N.; Vogt-Maranto, L.; Zdeborová, L. Machine learning and the physical sciences. Rev. Mod. Phys. 2019, 91, 045002. [CrossRef]
13. Mohanty, S.P.; Hughes, D.P.; Salathé, M. Using deep learning for image-based plant disease detection. Front. Plant Sci. 2016, 7, 1419. [CrossRef]
14. Xiao, Q.; Li, W.; Kai, Y.; Chen, P.; Zhang, J.; Wang, B. Occurrence prediction of pests and diseases in cotton on the basis of weather factors by long short term memory network. BMC Bioinform. 2019, 20, 688. [CrossRef]
15. Gutierrez, A.; Ansuategi, A.; Susperregi, L.; Tubío, C.; Ranki´c, I.; Lenža, L. A benchmarking of learning strategies for pest detection and identification on tomato plants for autonomous scouting robots using internal databases. J. Sens. 2019, 2019, 5219471. [CrossRef]
16. Plantix App. Available online: https://plantix.net/en/ (accessed on 11 January 2022).
17. Saillog. Available online: https://www.saillog.co/ (accessed on 11 January 2022).

18. CropDiagnosis. Available online: https://www.cropdiagnosis.com/portal/crops/en/home (accessed on 11 January 2022).
19. Gamaya. Available online: https://www.gamaya.com/ (accessed on 11 January 2022).
20. iFarmer. Available online: https://ifarmer.asia/ (accessed on 11 January 2022).