ROLE OF ML AND DEEP LEARNING IN PLANT DISEASES IDENTIFICATION: A REVIEW

[1]Anshu Priya, [2]Mr. Ramesh Kumar
[1]M.Tech Scholar, [2]Head Of Department (H.O.D)
[1,2]Department of Computer Science & Engineering, K.K University, Nalanda (Bihar)

Abstract: It is defined as a condition where a plant's physiological function is abnormal due to continuous, prolonged 'irritation' caused by phytopathogenic organisms (infectious and biotic disease agents). The growth of plants can be adversely affected by diseases. To combat this, their early identification is essential. Recent advancements in Deep Learning (DL) are providing a strong foundation for improved accuracy when utilizing ML models for the detection and classification of plant diseases. Modified DL architectures and visualization techniques have been implemented to identify and classify signs of sickness in plants. A variety of performance metrics are also used to evaluate these architectures/technologies. Infected leaves are processed by applying pixel-based operations to improve the information. A feature extraction process follows, followed by an image segmentation process, and finally a classification process based on patterns extracted from the diseased leaves. This paper reviews the concept of Diseases detection in plants using ML and Deep Learning.

Keywords: *Plant Disease Detection, Machine Learning, Deep Learning.*

1. INTRODUCTION

Even with the development of modern agriculture techniques, plant diseases caused by pathogens (bacteria, viruses, fungi, etc.) still pose a serious threat to the agricultural industry today. Infections caused by plant pathogens are responsible for 20%-40% of global food loss. To minimize damage, one of the most essential approaches is to detect and identify plant diseases as early as possible, so that proper prevention or treatment can be implemented. [1]
The detection of plant diseases and pests is an important research topic in machine vision. This technology uses machine vision equipment to collect images of plants and to determine whether they contain diseases or pests [1]. Machine vision-based plant disease and pest detection equipment is currently used primarily in agriculture and has partly replaced the naked eye identification process. [1]

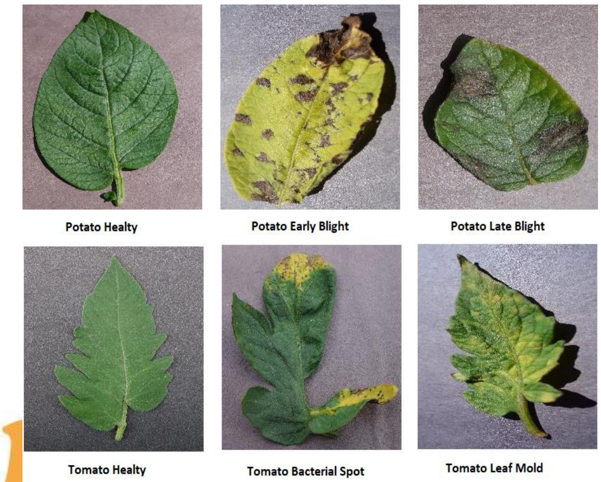

Fig 1. Plant Diseases Identification

It is common to use conventional image processing algorithms or manual feature design and classifiers as part of traditional machine vision-based plant diseases and pest detection methods [2]. In order to obtain uniform illumination, this type of imaging method selects the appropriate light source and shooting angle based on the different properties of plant diseases and pests. In spite of the fact that carefully constructed imaging schemes can greatly reduce the difficulty of classical algorithm design, they also add to the cost of application. On the other hand, under natural environments, it is unrealistic to expect classical algorithms to completely eliminate the impact of scene changes on recognition results.[2]

In real complex natural environment, plant diseases and pests detection is faced with many challenges, such as small differences between the lesion area and background, low contrast, large variations in the scale of the lesion area and various types, as well as a lot of noise in the lesion image. [3]

In addition, when collecting plant disease and pest images under natural light conditions, there are many disturbances, making traditional classical methods useless. [3]

Machine Learning in Plant Diseases Identification

Identifying and diagnosing life-threatening diseases in plants is the goal of plant health care. Plant mortality rates can be reduced if they are identified early on for any signs of disease. An early detection of such diseases can lower plant mortality rates. A type of artificial intelligence technology, machine learning (ML), allows researchers to enhance and develop without explicitly programming it.[4]

Some of popular machine learning algorithms are as follows,

Linear Regression: Using Linear Regression, you can visualize how random logs of wood would be arranged in increasing weight order to understand its working functionality. The catch, however, is that you cannot weigh each log. You have to determine its weight just by looking at its height and girth (visual analysis) and arranging them with a combination of these visible parameters. [5]

Decision Tree: One of the most popular algorithms in machine learning is the Decision Tree algorithm; it uses supervised learning to classify problems. The algorithm works well with both categorical and continuous dependent variables. Based on the most significant attributes and independent variables, it divides the population into two or more homogeneous sets. [5]

SVM (Support Vector Machine) Algorithm: As part of a classification algorithm, the SVM algorithm plots raw data points in an n-dimensional space (where n is the number of features you have). In order to classify the data, a particular coordinate is assigned to each feature, so lines called classifiers can be used to split the data and plot them on a graph. [6]

Naive Bayes Algorithm: The Naive Bayes classifier assumes that the presence of a particular feature is unrelated to the presence of any other feature. When calculating probability of a particular outcome, a Naive Bayes classifier would consider all of these features independently. Simple and effective, Naive Bayesian models outperform even highly sophisticated classification methods for massive datasets. [6]

The KNN algorithm (K-Nearest Neighbors): As far as the Data Science industry is concerned, this algorithm is used more often to solve classification problems than regression problems. It's a simple algorithm that stores all available cases and classifies any new cases by taking a majority vote from its k neighbors. Once this is determined, a distance function is used to determine which class the case belongs to. [6]

Machine learning's advantages in Plant Diseases Identification:

Trends and patterns are easily identified: It is possible to discover patterns and trends in large amounts of data using machine learning, which we cannot. For example, an e-commerce website like Amazon uses it to understand its users' browsing behavior and purchase histories in order to provide relevant products, deals, and reminders based on their preferences. As a result, relevant advertisements are displayed to them based on the results. [7]

No human intervention needed (automation): As ML allows machines to learn, it lets them make predictions as well as improve their algorithms on their own, so you don't have to babysit your project every step of the way. As an example, anti-virus software learns to filter new threats as they are recognized. ML is also good at detecting spam. [7]

Continual improvement: With experience, ML algorithms get better at making accurate and efficient predictions, which allows them to make better decisions. Consider making a weather forecast model. As the amount of data grows, your algorithms become more accurate. [7]

Data handling in multidimensional and multivariate environments : In dynamic or uncertain environments, Machine Learning algorithms are good at handling multidimensional and multivariate data. [7]

Deep Learning in Plant Diseases Identification

Plant diseases may cause substantial losses in food production and eradicate species diversity, making the agricultural sector a crucial part of economic growth and population growth. Early diagnosis of plant diseases using accurate or automatic detection techniques can enhance food production and minimize economic losses. Object detection and image classification systems have improved tremendously in recognition accuracy over the past few years thanks to deep learning. Since the number of diseases has increased rapidly, and farmers lack adequate knowledge, it has become increasingly difficult to identify and treat them. Computer vision combined with deep learning can solve this problem as leaves have similar textures and visual characteristics that make it easy to identify disease types. [8]

Some of Popular Algorithms in Plant Disease Detection are as follows,

Convolutional Neural Network (CNN): ConvNets are multilayer neural networks that are commonly used for image processing and object detection. In 1988, Yann LeCun developed the first CNN, which was known as LeNet, which was used to identify characters such as ZIP codes and digits. CNNs are commonly used to identify satellite images, process medical images, forecast time series, and detect anomalies. [9]

LSTMs (Long Short-Term Memory Networks): The LSTM is a type of Recurrent Neural Network (RNN) capable of learning and memorizing long-term dependencies. Time-series prediction is made possible by LSTMs because they remember previous inputs and have a chain-like structure where four interacting layers communicate uniquely. LSTMs are typically used for speech recognition, music composition, and pharmaceutical development in addition to time-series predictions. [9]

RNNs (Recurrent Neural Networks): As RNNs have directed cycles, they can feed the outputs from LSTMs as inputs to the current phase. RNNs are commonly used for image captioning, time-series analysis, natural-language processing, handwriting recognition, and machine translation due to their internal memories. [10]

GANs (Generative Adversarial Networks): There are two components to a GAN: the generator, which learns to generate fake data, and the discriminator, which learns from that fake data. GANs have gained popularity over the last few decades. They can be used to improve astronomical images and simulate gravitational lensing in dark matter research. Using image training, video game developers can upscale low-resolution, 2D textures in old video games to 4K or higher resolutions by using GANs.[10]

There are several advantages to deep learning over traditional machine learning methods, in Plant Diseases Identification:

- Feature learning algorithms can learn features automatically from data, which means they do not need to hand-engineer features. In tasks like image recognition, where the features are hard to define, this is especially useful.
- Deep learning algorithms are able to handle large and complex datasets that would be difficult for traditional machine learning algorithms to handle. [11]
- Deep learning algorithms achieve state-of-the-art performance on a wide range of problems, including image and speech recognition, natural language processing, and computer vision.
- Deep learning can uncover non-linear relationships in data that traditional methods would be unable to detect. [11]
- Deep learning algorithms can handle both structured and unstructured data, such as images, text, and audio.

2. CONCLUSION

Plant diseases and pests identification is a highly significant area of research in the realm of machine vision. Our solution takes advantage of specialized hardware to acquire images, then use machine learning techniques such as SVM, K-NN and CNN to make a judgment on whether there are any abnormalities present - such as burning or mould - on the foliage. This robust algorithm functions with high accuracy even when training samples are imperfect, achieving quick and reliable results.

REFERENCES

1. Yun Hwan Kima Seong Joon Yooa Yeong Hyeon Gua Jin Hee Limb Dongil Hana and Sung Wook Baik "Crop Pests Prediction Method using Regression and Machine Learning Technology: Survey" in IERI Procedia Elsevier vol. 6 2014.
2. G. Prem Rishi Kranth M. Hema Lalitha Laharika Basava and Anjali Mathur "Plant Disease Prediction using Machine Learning Algorithms" International Journal of Computer Applications (0975–8887) no. 25 November 2018.
3. Lujain Al Thunayan Nahed Al Sahdi and Liyakathunisa Syed "Comparative analysis of different classification algorithms for prediction of diabetes disease" Proceedings of the Second International Conference on Internet of things Data and Cloud Computing pp. 1-6 March 2017.
4. K. Thirunavukkarasu Ajay S. Singh MdIrfan and Abhishek Chowdhury "Prediction of Liver Disease using Classification Algorithms" 2018 4th International Conference on Computing Communication and Automation (ICCCA).
5. N. Gandhi L. J. Armstrong and O. Petkar "Predicting Rice crop yield using Bayesian networks" 2016 International Conference on Advances in Computing Communications and Informatics (ICACCI) pp. 795-799 2016.
6. Misha Denil David Matheson and Nando De Freitas "Narrowing thegap: Random forests in theory and in practice" International conference on machine learning (ICML) 2014.

 ISSN (Online): 2347 - 4718

7. S.G. Meshram M.J.S. Safari and K. Khosravi "Iterative classifier optimizer based pace regression and random forest hybrid models for suspended sediment load prediction" in Environ Sci Pollut Res Springer vol. 28 pp. 11637-11649 2021.
8. R. Beulah and M. Punithavalli "Prediction of sugarcane diseases using data mining techniques" Proc. IEEE International Conference on Advances in Computer Applications (ICACA) pp. 393-396 Oct. 2016.
9. N. Goel D. Jain and A. Sinha "Prediction model for automated leaf disease detection & analysis" Proc. IEEE 8th International Advance Computing Conference (IACC) pp. 360-365 Dec. 2018.
10. S. Kaur S. Pandey and S. Goel "Semi-automatic leaf disease detection and classification system for soybean culture" IET Image Processing vol. 12 no. 6 pp. 1038-1048 Feb. 2018.
11. L. Yashaswini H. Vani H. Sinchana and N. Kumar "Smart automated irrigation system with disease prediction" Proc. IEEE International Conference on Power Control Signals and Instrumentation Engineering (ICPCSI) pp. 422-427 Nov. 2017.

