International Journal of Innovative Research in Science and Engineering
Volume 10, Issue 08, 2024
www.ijirse.com
ISSN: 2454-9665

Assessing the efficacy of Artificial Intelligence in crop disease detection and classification

Sumit Dey[1], Raj Kumar Goel[2]

[1,2]*Department of Computer Application, North-Eastern Hill University,*
Tura Campus, 794002 Meghalaya, India

ABSTRACT-
Agriculture worldwide is diverse, reflecting varying climates, cultures, and economic conditions. Plant diseases pose a significant threat to global food security, leading to substantial yield losses and economic repercussions. Traditional methods of disease detection and classification in plants often rely on visual inspection by experts, which can be time-consuming, labor-intensive, and subjective. However, the emergence of machine learning (ML) and deep learning (DL) approaches has revolutionized the field by offering automated, accurate, and efficient solutions for plant disease detection and classification. In this paper, we explore the various ML and DL techniques employed in this domain and discuss how they contribute to enhancing the performance and speed of plant disease detection and classification systems.
Keywords: Plant disease, Artificial intelligence, Machine Learning, Feature Extraction.

1. INTRODUCTION

Plant diseases cause a significant threat to the agriculture industry and have the potential to starve the entire human population. The failure of early detection is due to manual identification, which takes more time and resources, immediate control measures, and a lack of knowledge. Early identification of plant disease is crucial for crop protection which can mitigate threats to the quality and yield of a healthy plant. The infections in the leaves might affect the plant's survival, reducing the plant's life span to only 2-3 years[1, 2]. Plant illness can damage the plant's reproduction rate, resulting in inedible seeds. These seeds affect the soil, making it infertile for the plant's growth. The disease has also impacted fresh plants sown in the field, and the disease has been passed down through the generations in the soil, causing crop failure. During illness, the metabolism and transport of nutrients are disrupted. The reason of plant diseases could be the reason of diseases in different parts of a plant such leaf, root and stems; however, leave is one of the most important institutes to be observed to identify and detect infection. [3,4].It is a difficult task to monitor plant diseases manually due to its complex nature and time-consuming process. Artificial intelligence (AI) based computational models can detect leaf diseases in their early stages [5]. Traditional inspection across different plant fields is time-consuming and impractical for a wider plantation size, thus reducing crop production. Therefore, many smart agricultural practices are deployed to control plant diseases and pests. Most of these approaches, for example, use vision-based artificial intelligence (AI), machine learning (ML), or deep learning (DL) methods and models to provide disease detection solutions [6].

International Journal of Innovative Research in Science and Engineering
Volume 10, Issue 08, 2024
www.ijirse.com
ISSN: 2454-9665

Plant problems are caused by living organisms, such as fungi, bacteria, viruses, nematodes, insects, mites, and animals. A-biotic disorders are caused by non-living factors, such as drought stress, sunscald, freeze injury, wind injury, chemical injury, nutrient deficiency, or improper cultural practices, such as overwatering or planting conditions [7]. Unfortunately, the damage caused by these various living and non-living agents can appear very similar. Even with close observation of symptoms, accurate diagnosis can be difficult. For example, browning of leaves on an oak tree caused by drought stress may appear similar to leaf browning caused by oak wilt, a serious vascular disease, or the browning cause by anthracnose, a fairly minor leaf disease [8].

The first important step is to determine the identity of the plant and its requirements for healthy growth [9].A-biotic damage often occurs on many plant species. Drought stress or chemical injury will likely cause damage on several types of plants. In contrast, biotic disease problems are more limited to a certain species. The fungi that cause tomato leaf blight do not cause damage on sweet corn, for example.Biotic diseases sometimes show physical evidence (signs) of the pathogen, such as fungal growth, bacterial ooze, or nematode cysts, or the presence of mites or insects. Abiotic diseases do not show the presence of disease sign [10].

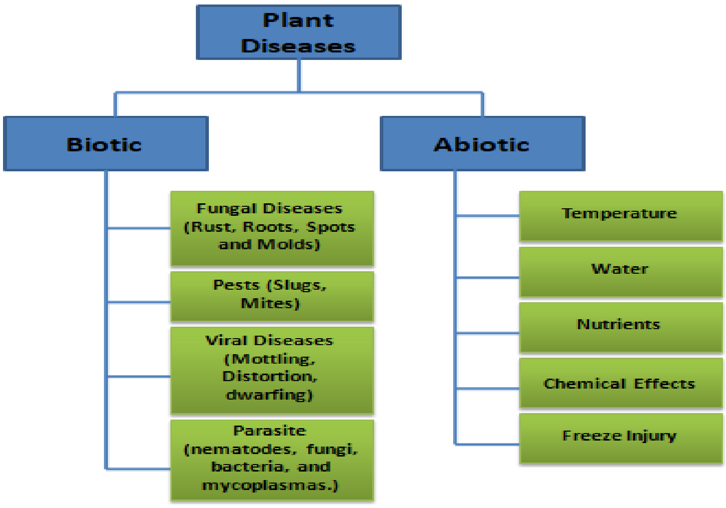

Figure1: Showing the types of plant disease and its causes.

2. LITERATURE REVIEW

The AI-based crop disease identification and classification literature have substantially improved over the years. Different algorithms have been applied such as CNNs and SVM, each having its own level of performance accuracy and efficiency. While the significance of large and accurately annotated data as crucial while training a robust model is widely acknowledged, variation in performance metrics among different studies expose gaps towards unification evaluation criteria. Many immediate advantages of AI in both early disease detection and yield improvement are rightly being discussed yet the challenges within developing scaleable solutions for resource limited environments implementing new-wave technologies (like AI) into traditional farming practices is persisting. The use of technology can significantly increase crop production all around the world. Previous research has determined the robustness of DL and ML techniques, k-means clustering (KMC), naive Bayes (NB), Digital soil prediction, feed-forward neural network (FFNN), support vector machine (SVM), k-nearest neighbor (KNN) classifier, fuzzy logic (FL), genetic algorithm (GA), artificial neural network (ANN), convolutional neural network (CNN), and so on [4,7].

A.F. Aji *et al.* [11] developed a mobile app to identify palm oil leaf diseases. Their app uses image processing and neural networks. To make the app fast enough for mobile use, they created a streamlined process. The app extracts six key features from palm leaf images and then uses a neural network to recognize disease patterns. P. Revathi and M. Hemalatha [12] developed a method to classify cotton leaf spot diseases using image processing. They used images taken with mobile phones to identify disease symptoms and categorize them with their proposed HPCCDD algorithm. This approach aims to improve farming practices by enabling early disease detection and targeted pesticide use.

Batule*et al.* [13] used RGB and HSV color analysis to detect leaf diseases. Their method involved image pre-processing, color thresholding, and feature extraction using K-means clustering. Khirade& Patil [14] focused on plant disease detection using image processing techniques. Their work involved image acquisition, pre-processing, segmentation, feature extraction, and classification. Zhang & Meng [15] developed an automatic citrus canker detection system for field-captured leaf images. Their method uses an improved AdaBoost algorithm to select key features for lesion segmentation. They propose a descriptor that combines color and texture features of canker lesions, and a two-level hierarchical detection structure to identify them.

S. D. Bauer *et al.* [16] investigated automatic classification methods for leaf diseases using multispectral images. They tested two pixel-wise approaches: k-nearest neighbor and adaptive Bayes classification with a Gaussian mixture model. The median classification rates were 91% for *Cercosporabeticola* and 86% for *Uromyces betae*. They also explored contextual classification using conditional random fields, which reduced common errors found in pixel-wise methods.S. Sladojevic*et al.* [17] developed a deep neural network for plant disease recognition using leaf images. They used feed-forward and feedback ANN architectures and created a dataset of 30,880 images for training and 2,589 for validation. The dataset, sourced from various languages and sources, includes fifteen classes and uses augmented images to improve feature learning.

U. Barman *et al.* [18] have performed a comparative assessment of pest damage identification of coconut plants using damage texture and color analysis. The Gray Level Co-occurrence Matrix (GLCM) and Gray Level Run Length Matrix (GLRLM) techniques are applied to extract the texture features of the damages, and two Artificial Neural Network (ANN) architectures are reported to classify the extracted data features of the damages into five different classes. H. Paul *et al.* [19] developed a DL-based method for detecting and classifying four potato leaf diseases (early-blight, septoria, late-blight, and black-leg) using explainable AI.

S. P. Mohanty *et al.* [20] proposed a DL model for disease detection using datasets of various crops. They employed Google-LeNet and AlexNet, achieving about 97.30% accuracy. However, the model's classification rate drops significantly with poor image background contrast. S. Ali *et al.* [21] developed a PCA-LDA classification system for potato leaf diseases using intelligent feature fusion. They extracted both handcrafted hybrid and deep features from RGB images, using TL-ResNet50 for deep feature extraction. The fused feature vector combines these features, with PCA selecting the most discriminant ones for LDA model development.

International Journal of Innovative Research in Science and Engineering
Volume 10, Issue 08, 2024
www.ijirse.com
ISSN: 2454-9665

3. METHODS OF PLANT DISEASE DETECTION

3.1 Traditional methods of plant disease detection

It relied on visual inspection by human experts, a process involving the physical examination of plants for symptoms indicative of disease. Trained agronomists or plant pathologists conduct these inspections in the field or laboratory, identifying characteristic signs such as lesions, discoloration, wilting, or abnormal growth patterns [22].

3.2 Laboratory-based methods

It represents another traditional approach to plant disease detection, involving the collection of plant samples followed by diagnostic tests to identify pathogens or disease-causing agents. These tests may include culturing, polymerase chain reaction (PCR), enzyme-linked immunosorbent assay (ELISA), or microscopy techniques [23]. While laboratory testing provides precise identification of pathogens, it is time-intensive and costly. Furthermore, sample collection can be invasive and may disrupt crop production, especially for perennial crops or sensitive plant species [24].

Table 1outlines several common leaf diseases, describing their symptoms, underlying causes, and the types of pathogens responsible. Downy Mildew (Plasmoparaviticola) is characterized by yellow to white patches on the leaves, which may develop into white to grayish growths. This disease, caused by a fungus, thrives in humid environments and is particularly damaging to grapevines. Shot Hole disease manifests as large holes in the leaves, resembling bullet holes. This condition is caused by a combination of viruses and water molds, with the holes forming when infected tissue falls away. Anthracnose presents with lesions on the leaves that become covered in pink, gelatinous masses of spores, indicating a fungal infection. Early Blight, another fungal disease, is often exacerbated by rain and can cause significant damage to leaves early in the growing season. Lastly, Leaf Curl is marked by areas of the leaf turning reddish, a symptom caused by a fungal infection that affects the leaf's appearance and health.

Table 1: Different symptoms of plant diseases.

Leaf diseases	**Leaf diseases symptoms**	**Description Pathogen group**
Anthracnose	Lesion becomes covered with pink, gelatinous masses of spores	Fungus
Downy Mildew (Plasmoparaviticola)	Yellow to white patches, and white to grayish	Fungus
Early blight	Early blight	Rain and fungus
Leaf curl	Turn reddish on leaf area	Fungus
Shot hole	Large holes appear in the leaf	Viruses and water mold

3.3 PLANT DISEASE DETECTION METHODS WITHCOMPUTATION

Plant disease detection methods encompass a variety of techniques, each possessing unique strengths and limitations. The utilization of machine learning algorithms stands as a prominent strategy for the classification of

International Journal of Innovative Research in Science and Engineering
Volume 10, Issue 08, 2024
www.ijirse.com
ISSN: 2454-9665

plant diseases utilizing image data [25]. These algorithms, which consist of supervised and unsupervised learning methods, are capable of extracting patterns from annotated datasets of both healthy and diseased plants to accurately categorize new samples [26]. Supervised learning approaches, exemplified by support vector machines (SVMs) and random forests, are specifically trained on labeled image data to effectively classify plants into different disease categories. Conversely, unsupervised learning methods, such as clustering algorithms, are designed to detect patterns and group similar plants together based on inherent similarities in their characteristics without the need for prior labelling [27].

Computer vision is essential for the automation of plant disease detection through the extraction of significant characteristics from images and the recognition of disease indications. Various image pre-processing methods, like normalization and enhancement, are employed to enhance the quality of input images, whereas feature extraction algorithms are utilized to capture pertinent data for disease categorization. Algorithms for object detection and segmentation are implemented to pinpoint and outline areas affected by disease in plant images, thus enabling precise identification and quantification of disease symptoms[28,29].

4. DATASETS AVAILABILITY FOR COMPUTATIONS

Among the publicly available datasets, the Plant Village Dataset stands out as one of the most comprehensive and widely used resources. It includes images of healthy and diseased plant leaves from 38 different classes, making it a popular choice for training DL models like VGG, ResNet, and DenseNet. This dataset is freely accessible on platforms like Kaggle, contributing significantly to advancements in plant disease detection.

Figure 2: Sample images from kaggle, PlantVillage dataset for 38 types of leaf diseases [30].

Another valuable public dataset is the Kaggle Plant Seedlings Classification, which offers images of around 960 unique plants belonging to 12 species at various growth stages. This dataset is particularly useful for training models such as AlexNet, GoogleNet, and InceptionV3. Additionally, the UCI Machine Learning Repository provides the Leaf Dataset, a smaller collection of leaf images from 30 different species. This dataset is well-suited for traditional ML algorithms like SVM, KNN, and Decision Trees. The Flavia Leaf Dataset, with images from 32 different plant species, is also widely used for feature extraction and classification tasks.

5. DISCUSSION

The integration of artificial intelligence (AI) technology into the detection and diagnosis of crop leaf diseases offers a sophisticated and efficient solution for precision agriculture. ML and DL techniques have significantly improved the automation of identifying and classifying crop leaf diseases, increasing the accuracy and speed of disease management. Advancements in ML and DL have enhanced the identification and management of infestations in crops. Image recognition techniques now make it possible to detect complex diseases and pests. However, much of the research has been confined to laboratory settings, relying heavily on collected images of plant diseases and pests, often focusing on specific features like disease spots, insect appearances, and leaf characteristics.

Traditional ML algorithms also play a crucial role. SVM (Support Vector Machine) is widely used for classification tasks, leveraging hyperplanes to separate classes. KNN (K-Nearest Neighbours) is an instance-based learning method that classifies based on the majority label among the nearest data points. Decision Tree models make decisions based on feature splits, and Naive Bayes applies probabilistic reasoning for classification tasks. Convolutional Neural Networks (CNNs) are essential in image recognition tasks. They automatically and adaptively learn spatial hierarchies of features through back-propagation by using multiple building blocks, such as convolution layers, pooling layers, and fully connected layers.

High-resolution images may still present challenges in identifying early-stage diseases and pests. Incorporating meteorological and plant health data, such as temperature and humidity, is essential for effective identification and prediction. However, this approach has been infrequently applied in early diagnostics. Plant growth is cyclical, seasonal, and region-specific. To build more robust and generalized models, it's important to gather sample images from various growth stages, seasons, and locations. The variability in plant diseases and pests throughout crop development, and differences in plant species across regions, means that many current research findings may not be universally applicable. Even if a model performs well in specific tests, its reliability across different times or locations isn't guaranteed. Most studies focus on images in the visible spectrum, but data beyond this range, such as near infrared and multispectral spectra, could enhance plant disease datasets.

VGG, including models like VGG19, is known for its deep, simple architectures consisting of small 3x3 filters that stack deeper to enhance the learning of complex features. DenseNet121 connects each layer to every other layer in a feed-forward fashion, which improves feature propagation and reduces the vanishing gradient problem. ResNet models, such as ResNet50 and ResNet101, utilize residual connections that allow gradients to flow through shortcut connections, enabling the training of much deeper networks. AlexNet revolutionized deep learning in computer vision with its use of ReLU activation and dropout layers to reduce overfitting.

GoogleNet, with its inception modules, and InceptionV3, which further optimizes these modules, are designed for computational efficiency while maintaining high performance. MobileNet is specifically optimized for mobile and embedded vision applications, focusing on reducing model size and latency. ResNetV2 are advanced variations that build upon the ResNet architecture to enhance performance and accuracy. DenseCNNs combine dense connectivity with CNN layers, ensuring efficient feature reuse and improved.

In the below table 2 showing that, the Kaggle PlantVillage and AGRONOMI-Net datasets are vital resources for developing ML models aimed at plant disease and pest recognition (PDPR). They include images of healthy and diseased plants, helping to train algorithms in detecting various agricultural issues. Plant Clef complements these by offering a broader range of plant species images, useful for both species' identification and biodiversity research.The Open Plant Disease Dataset and specialized datasets like Northern Leaf Blight (NLB) Lesions focus on specific plant diseases, such as NLB in maize, providing detailed data to train models for precise disease recognition. Additionally, datasets of insects from rice, maize, and soybean crops are used to develop models that can identify and classify common pests, contributing to better crop management and protection.Together, these datasets support advancements in precision agriculture, enabling automated systems to monitor plant health, identify diseases, and manage pests effectively.

Table 2: Showing Plant disease and pest detection from various datasets.

Dataset Name	**Types of Data**	**Disease/Pest Types Covered**	**Description**
Kaggle PlantVillage	RGB Images	38 crop species and 38 disease types	A publicly available dataset of over 54,000 images of diseased and healthy plant leaves, compiled from experts and citizen scientists
AGRONOMI-Net	RGB and thermal images	Multiple crop species and disease types	A dataset of over 3,000 images of various crops, compiled by the AGRONOMI-Net project for disease detection research
PDPR	RGB	Various crop species and diseases	A dataset of over 30,000 images of diseased and healthy plants collected from a field environment
Plant Clef	RGB Images	Multiple crop species and disease types	A dataset of over 9,000 images of plant leaves, used for the annual Plant CLEF benchmarking campaign
Open Plant Disease Dataset	RGB and infrared images	Multiple crop species and disease types	A dataset of over 8,000 images of plant leaves, compiled from various sources including university research and citizen scientists
NLB Lesions	RGB	NLB Disease	A dataset of images of corn plants affected by NLB collected from a field-environment
Insects from rice, maize, soybean	RGB	Rice Plant hoppers, Brown Plant hoppers and White flies.	A dataset of images of insects on rice, maize, & soybean plants collected from a field-environment.

6. CONCLUSION

Adopting advanced AI technologies such as ML and DL can help to overcome these challenges by enabling early identification of plant diseases. This study also addressed the challenges and limitations associated with using ML and DL for plant leaf disease identification, such as issues with data availability, imaging quality, and the differentiation between healthy and diseased plants. This advancement not only contributes to improved crop yield and quality, but also supports sustainable agricultural practices by enabling timely and targeted interventions. While AI technology is very promising for the detection and diagnosis of crop leaf diseases, practical implementation requires careful selection of appropriate models and methods to suit specific conditions and demands. This is particularly important in tropical regions, where the complexity of managing plant disease is increased.

Future research may focus on multi-dimensional fusion techniques to gather and recognize information on plant pests, as well as compiling a database of images covering a wide variety of wild plant pests and diseases. These innovations, along with complete and accurate datasets, could improve the overall performance of AI algorithms in crop disease identification and management.

REFERENCES

[1] Sudha, P. and Kumaran, P., 2023. Early detection and control of anthracnose disease in cashew leaves to improve crop yield using image processing and machine learning techniques. *Signal, Image and Video Processing*, *17*(7), pp.3323-3330.

[2] Nikith, B.V., Keerthan, N.K.S., Praneeth, M.S. and Amrita, T., 2023. Leaf disease detection and classification. *Procedia Computer Science*, *218*, pp.291-300.

[3] Vishnoi, S. and Goel, R.K., 2024. Climate smart agriculture for sustainable productivity and healthy landscapes. *Environmental Science & Policy*, *151*, p.103600.

[4] Kaur, P., & Gautam, V. (2021). Research patterns and trends in classification of biotic and abiotic stress in plant leaf. *Materials Today: Proceedings*, *45*, 4377-4382.

[5] Bania, R. K., Dey, S., & Buragohain, N. (2023). Hybrid ResNet50-PCA-based deep transfer learning approach for classification of tomato leaf diseases. *Journal of Biodiversity and Environmental Sciences, 22*(4), 42–48.

[6] Shafik, W., Tufail, A., Namoun, A., De Silva, L.C. and Apong, R.A.A.H.M., 2023. A systematic literature review on plant disease detection: Motivations, classification techniques, datasets, challenges, and future trends. *Ieee Access*, *11*, pp.59174-59203.

[7] Wilcoxson, R. D. (1996). Helen Hart, Remarkable Plant Pathologist (1900–1971). *Annual review of phytopathology*, *34*(1), 13-23.

[8] Costello, L. R. (2003). *Abiotic disorders of landscape plants: A diagnostic guide* (Vol. 3420). UCANR Publications.

[9] Gutschick, V. P. (1999). Biotic and abiotic consequences of differences in leaf structure. *The New Phytologist*, *143*(1), 3-18

[10] Kennelly, M., O'Mara, J., Rivard, C., Miller, G. L., & Smith, D. (2012). Introduction to abiotic disorders in plants. *The Plant Health Instructor*, *10*(1094), 10-20.

[11] A.F. Aji, Q. Munajat, A.P. Pratama, H. Kalamullah, J. Setiyawan, and A.M. Arymurthy. Detection of palm oil leaf disease with image processing and neural network classification on mobile device. International Journal of Computer Theory and Engineering, 5(3):528, 2013.

[12] P. Revathi and M. Hemalatha. Classification of cotton leaf spot diseases using image processing edge detection techniques. In proceedings of 2012 International Conference on Emerging Trends in Science, Engineering and Technology (INCOSET), pages 169–173. IEEE, 2012.

[13] V.B. Batule, G.U. Chavan, V.P. Sanap, and K.D. Wadkar. Leaf disease detection using image processing and support vector machine SVM. Journal for Research, 2(02), 2016.

[14] S.D. Khirade and A.B. Patil. Plant disease detection using image processing. In proceedings of 2015 International Conference on Computing Communication Control and Automation, pages 768–771. IEEE, 2015.

[15] M. Zhangand Q. Meng. Automatic citrus canker detection from leaf images captured in field. Pattern Recognition Letters, 32(15):2036–2046, 2011.

[16] S.D. Bauerand F. Korˇcand W. F¨orstner. The potential of automatic methods of classification to identify leaf diseases from multispectral images. Precision Agriculture, 12:361–377, 2011.

[17] S. Sladojevic, M. Arsenovic, A. Anderla, D. Culibrk, and D. Stefanovic. Deep neural networks based recognition of plant diseases by leaf image classification. Computational Intelligence and Neuroscience, 16 (1):11, 2016.

[18] U. Barman, C. Pathak, and N.K. Mazumder. Comparative assessment of pest damage identification of coconut plant using damage texture and color analysis. Multimedia Tools and Applications, 82:25083—-25105, 2023.

[19] H. Paul, S. Ghatak, S. Chakraborty, S.k.Pnadey, L.Dey, D. Show, and S. Maity. A study and comparison of deep learning based potato leaf disease detection and classification techniques using explainable ai. Multimedia Tools and Applications, 1–34, 2023.

[20] S.P. Mohanty, D.P. Hughes, and M. Salathe. Using deep learning for image-based plant disease detection. Frontiers in Plant Science, 7:1419, 2016.

[21] S. Ali, M. Hassan, J.Y. Kim, M.I. Farid, M. Sanaulla, and H. Mufti. FF-PCALDA: intelligent feature fusion based pca-lda classification system for plant leaf diseases. Applied Sciences, 12(7):3514, 2022.

[22] Baksh, H., Thottempudi, K., Khatal, M. M., GS, S., Das, N., Alam, M. A., ... & Priya, P. (2024). Advancing Agriculture through Artificial Intelligence, Plant Disease Detection Methods, Applications, and Limitations. *Journal of Advances in Biology & Biotechnology*, *27*(8), 730-739.

[23] Kushwah, N., Billore, V., Sharma, O. P., Singh, D., & Chauhan, A. P. S. (2024). Integrated Nutrient management for optimal plant health and crop yield. *Plant Science Archives*.

[24] Pethybridge SJ, Nelson SC. Classical and Bayesian approaches to the spatial analysis of disease incidence: The example of wheat stripe rust. Phytopathology. 2015;105(6):782-791

[25] Demilie, W. B. (2024). Plant disease detection and classification techniques: a comparative study of the performances. *Journal of Big Data*, *11*(1), 5

[26] Szegedy C, Liu W, Jia Y, Sermanet P, Reed S, Anguelov D, Rabinovich A. Going deeper with convolutions. In Proceedings of the IEEE conference on computer vision and pattern recognition. 2015;1-9

[27] Fatima S, Rasool A, Sajjad N, Bhat EA, Hanafiah MM, Mahboob M. Analysis and evaluation of penicillin production by using soil fungi. Biocatal. Agric. Biotechnol. 2019;21:101330

[28] Goel, R.K., Yadav, C.S., Vishnoi, S. and Rastogi, R., 2021. Smart agriculture–Urgent need of the day in developing countries. *Sustainable Computing: Informatics and Systems*, *30*, p.100512.

[29] Ritik Chawla, Kinjal Mondal, Pankaj, Meenakshi Sahu. Mechanisms of plant stress tolerance: Drought, salinity, and temperature extremes. Plant Science Archives. 2022;04-08.

[30] https://www.kaggle.com/datasets/mohitsingh1804/plantvillage