IJNRD.ORG ISSN : 2456-4184
INTERNATIONAL JOURNAL OF NOVEL RESEARCH AND DEVELOPMENT (IJNRD) | IJNRD.ORG
An International Open Access, Peer-reviewed, Refereed Journal

Agrishield- Crop Disease Detection and Prediction using YOLOv8s

[1]Mr. Prateek Meshram, [2]Akshay Gade, [3]Divyanshi Rahate, [4] Ruhaan Hawaldar, [5]Durvesh Patil

[1]Professor , [2]Student, [3]Student, [4]Student, [5]Student
[1]Computer Engineering
[1]Dr. D. Y. Patil Institute of Engineering Management and Research, Akurdi, Pune, India

Abstract : In the case of plant disease detection, including leaf blotch, powdery mildew and rusts, conventional methods of machine learning such as feature selection and classification have been used. Most of them were successful but they could not identify the advanced and subtle signs, and even the early stages of disease development. For instance, approaches that relied purely on images were unable to capture complex features, so, deeper solutions were needed. In recent past, researchers have started using deep learning models, such as, convolutional neural networks, autoencoders for plant disease detection. These deep learning models are capable of learning features automatically from data, are suitable for high dimensional data and improve the accuracy of the output. For instance, YOLOV8S and other hybrid approaches are used in automatically detecting and classifying diseases in certain crops such as capsicum. However, there are still problems of robustness and generalization. The range of these systems can be expanded by incorporating more analytics to support plant disease detection. They can also grow to conduct multi-class detection which would mean they can detect several diseases at a time and give more information to farmers. Enhancements in model and system robustness will also enable the models to be generalized to a wider scope of applications to make the AI based disease classifiers to be more reliable.

Keywords: Machine learning; Convolutional Neural Network; Deep Learning; YOLOv8

INTRODUCTION

Plant diseases represent one of the highest threats to food security on a global scale, causing significant crop yield and quality losses. Early and accurate detection of plant diseases is indispensable in minimizing these losses, whereas the traditional disease management methods are often out-of-date. The majority of the traditional machine learning (ML) (IEEE, 2018) approaches are based on manual feature extraction and classification, which have been used to identify the normal plant diseases, which include leaf blotch, powdery mildew, and rust. However useful these techniques have proven, they can still be limited in some ways, mainly to the detection of early-stage symptoms or very subtle changes in the plant's health. At times manual image classification methods encounter great challenges in the extraction of complex high-dimensional features. There arises therefore a dire need for a complex approach capable of encapsulating the complexities inherent within an agricultural dataset. The recent advances in artificial intelligence, especially deep learning (DL) models, usher in a new generation of plant disease detection systems. Convolutional neural networks (CNNs) (O'Shea, 2015) and autoencoders are extensively being used to enhance the accuracy of disease detection due to their feature extraction capacity and ability to handle high-dimensional data. These models do not require manual feature engineering, as deep-learning-based models learn the features directly from the data and improve disease recognition at the wear level. Techniques include YOLO V8S (Varghese, 2024) and hybrids combining CNNs with other DL architectures, which have hit real promise with very high accuracy in classifying diseases on specific crops, including capsicum. The challenges most commonly faced yet some.

The future applications for plant disease detection systems are wide-ranging and hold tremendous potential to transform agricultural methodologies. A major area of innovation is the addition of multi-class classification functionality, which enables a system to detect multiple diseases at once. This would give farmers a general idea of plant health, leading to more informed decision-making. Also, as more comprehensive and varied data gathering becomes more prevalent in agriculture, introducing larger and more varied datasets into these models can train them and make them more useful for a broader set of diseases and crops. Further improvements in analytical techniques—such as ensemble models and hybrid deep learning models—could significantly enhance the flexibility and reliability of these systems. By improving how they generalize, plant disease models can become increasingly dependable in multiple agricultural settings, which will help drive their utilization in areas where complete access to sophisticated precision agriculture technology is impossible.

Integration of real-time feeds like weather conditions, soil moisture, and historical crop data can facilitate context-aware disease forecasting while potentially minimizing the need for intermittent human intervention. With increased research, these systems could further be developed to predict disease occurrences, warning farmers of outbreaks even before symptoms appear on the crops. By improving flexibility, scalability, and performance in various agricultural settings, next-generation plant disease detection systems will play a key role in global initiatives for the optimization of crop management practices. These advances have the ability to reduce pesticide dependence while encouraging sustainable agricultural practices. Therefore, as these technologies evolve further, they will offer solid support for the advancement of the agricultural industry in addressing existing challenges.

II. Literature Survey of the Papers

Year	Author	Technique	Observation
2023	M. Chithambarathanu and M. K. Jeyakumar	Random Forest, SVM, Decision Tree, CNN.	1. ML and DL techniques (SVM, CNN, LSTM) enhance crop pest detection and protection. 2. Automated pest monitoring reduces human error and increases efficiency.
2023	Harjeet Kauri, Deepak Prashar and Vipul Kumar	CNN, Random Forest, RestNet	1. DL models like AlexNet and ResNet are effective, but gaps remain. 2. Field images and model adaptation for disease severity are needed.
2023	Mr.Gopinath V, Ilakiya V, Nandhini R	CNN and LSTM	1. CNN and LSTM are used for plant disease prediction to enhance agricultural yield. 2. Image processing techniques aid in early detection of plant diseases for sustainable farming.
2023	Muhammad Shoaib, B. Shah, S. Sapagh, Akhtar Ali	CNN, SVM	1. ML and DL improve plant disease detection but require diverse data for robustness. 2. Collaboration between agriculture experts and DL models is essential for effective disease management
2023	Nishant Shelar , Suraj Shinde, Shubham Sawant	CNN	1. A Disease Recognition Model using CNN and leaf image classification detects plant diseases. 2. CNN processes pixel data for accurate plant disease detection through image recognition.
2023	Nauman Qadaeer, Thabit Sabbah and Muhammad attique khan	CNN	1. A lightweight modified CNN model with eight layers detects wheat diseases with 93% accuracy. 2. High-resolution images and a larger training set improve disease detection accuracy.
2023	Shengjie Leng, Yasenjiang Musha	YOLOv5	1. YOLOv5 model achieves 87.5% mAP@0.5, 5.4% higher than the original. 2. Uses CIPAM, FRAFM, and MobileBit for better detection.
2024	Dr. Amol Dhakne	CNN	1. CNN-based system enhances accuracy and versatility in medicinal plant classification. 2. Effectively identifies subtle botanical features for improved classification schemes.
2024	Rosemary Ngozi Ariwa, Caleb Markus, Nora Godwin Teneke	YOLOv8	1. Utilization of the YOLOv8 algorithm achieves 99.8% accuracy in maize leaf disease detection. 2. Highlights the broader applicability of deep learning techniques in real-world problem-solving.

Two researchers, M. Chithambarathanu and M. K. Jeyakumar, proposed a hybrid approach for the plant disease detection and classification. For this purpose, certain advanced image-processing techniques are applied to the analysis of textures of the plants. High-resolution images of healthy and diseased leaves from locally grown crops were acquired under varying environmental conditions and surroundings. The system has juxtaposed the started images with images in the sequence of acquisition, feature extraction, and classification, allowing the precise differentiation of healthy and diseased plants. This research was aimed at providing an applied and practical means of managing these diseases in agriculture, which would help the local farming community to monitor crop health efficiently.. (Jeyakumar2, 2023)

Harjeet Kaur, Deepak Prashar, and Vipul Kumar reviewed the application of deep learning models in plant disease detection, emphasizing its importance in improving crop quality and yield. The study focused on the use of CNNs for identification of various diseases in plants and presented a comparative analysis of the works done by different researchers. The paper emphasizes the effectiveness of deep learning techniques in agriculture and seems likely to revolutionize plant health monitoring and crop production.. (Harjeet Kaur, 2023)

An effective system for plant disease detection was developed by Mr. Gopinath V, Ilakiya V, and Nandhini R with the aim of reducing agricultural crop yield losses. Various image processing techniques along with CNN and LSTM models are employed for effective prediction of plant diseases. Their approach seeks to resolve existing challenges of manual inspection of diseases, allowing for accurate diagnostic and timely detection for improved sustainability in farming methods.. (Mr. Gopinath V, 2023)

Muhammad Shoaib, B. Shah, S. Sapagh, and Akhtar Ali explored the use of Machine Learning (ML) and Deep Learning (DL) techniques to detect plant diseases. The paper reviews advances from the years 2015 to 2022, proposing how ML and DL solved the accuracy-efficacy dilemma. It mentions some of the challenges, including limited availability of replicated data, low imaging quality, and features to distinguish healthy versus diseased plants. This insight in turn provides useful solutions to reduce these hindrances, thus helping build a more broader understanding for upcoming research in plant disease identification along with several recent advancements. (Muhammad Shoaib, 2023)

Nishant Shelar, Suraj Shinde, and Shubham Sawant constructed a Disease Recognition Model based on leaf image classification for plant disease detection. They used Convolutional Neural Networks (CNN) and a special form of neural networks meant for image processing that analyzes leaf images for efficient disease-identifying techniques. Such mechanisms of approach will lead to a far more accurate and lower-cost alternative than traditional ones, which are often time-consuming and imprecise.. (Nishant Shelar, 2023)

Nauman Qadaeer, Thabit Sabbah, and Muhammad Attique Khan developed a lightweight modified CNN model to detect wheat diseases, aiming to overcome the computational limitations of traditional CNNs. Their architecture features eight layers with convolutional filters of 16, 32, and 64. The model was tested on high-resolution images gathered from Azad Kashmir, using three dataset variations (S1, S2, and S3). Among them, the S3 dataset delivered the best performance, achieving an accuracy of 93%. The study highlights that training on larger, high-resolution datasets enhances disease detection accuracy, ultimately supporting better wheat production. (Nauman Qadaeer, 2023)

Shengjie Leng and Yasenjiang Musha developed a lightweight YOLOv5-based model for the quick detection of Northern corn leaf blight, a fungal disease that affects maize crops. Their model incorporates the Crucial Information Position Attention Mechanism (CIPAM) to preserve essential details, the Feature Restructuring and Fusion Module (FRAFM) to improve feature map integration, and the Mobile Bi-Level Transformer (MobileBit) for better scene comprehension. Experimental results indicate that the model achieves an 87.5% mAP@0.5 accuracy on the NLB dataset, surpassing the original YOLOv5 model by 5.4%. (Musha, 2023)

Dr. Amol Dhakne designed a Medicinal Plant Classification System using Convolutional Neural Networks (CNNs) to improve the identification and categorization of medicinal plant species. Traditional methods struggled to capture subtle and distinctive plant features, making accurate classification challenging. Their CNN-based approach enhances accuracy and adaptability by effectively recognizing complex patterns in plant images. With its flexible structure and hierarchical learning capabilities, the system offers a reliable solution for precise medicinal plant identification. (Dhakne, 2024)

Rosemary Ngozi Ariwa, Caleb Markus, and Nora Godwin Teneke created a plant disease forecasting system based on Convolutional Neural Networks and Long Short-Term Memory models. As plant diseases significantly influence agricultural productivity, early detection is crucial for sustainable agriculture. The conventional methods of identification are generally time-consuming and require much manual effort, so timely detection becomes difficult. The suggested method employs image processing methodologies to offer an efficient solution with improved crop yield and quality. Through the integration of CNN and LSTM, the system enhances speed and accuracy in disease prediction to foster improved agricultural practices. (Rosemary Ngozi Ariwa, 2024)

III. Background

3.1 Traditional Machine Learning Approaches

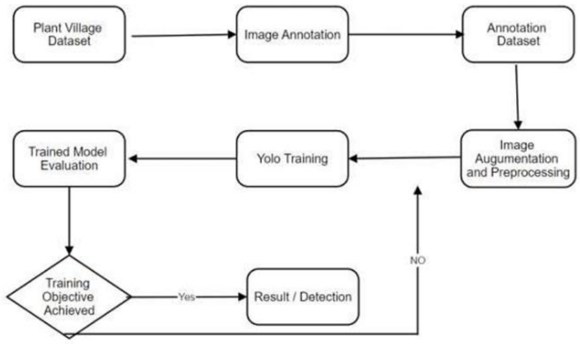
Conventional machine learning methods for detecting plant disease usually depend on feature extraction and classification. They process a preselected limited number of symptoms and need human intervention for image classifying. Though suitable for the identification of prevalent diseases, they tend to fail in identifying fine lines of infection at initial stages. The intricacy of plant diseases demands a sophisticated method with the ability to learn from big data and spot patterns that human eyes could miss.

3.2 The Role of Deep Learning

Deep learning, particularly convolutional neural networks , has transformed image recognition by automating the extraction of features. This removes the necessity for extensive manual labeling, enabling models to recognize complex patterns more effectively and accurately. Autoencoders further enrich deep learning by providing a way to carry out unsupervised learning and extract meaningful features from unlabeled data. However, despite all these advances, difficulties still arise in sustaining model resilience and increasing generalization under diverse agricultural environments

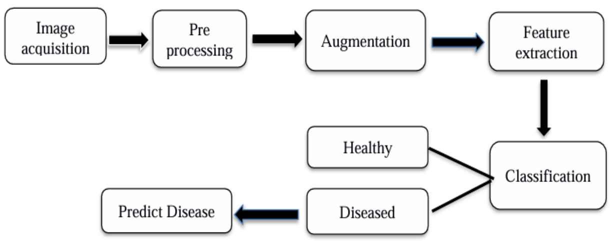
IV. Plant Village Dataset

Plant Village dataset is an open, public image dataset in wide use to assist in research for plant disease identification. It consists of over 54,000 labeled images of healthy and diseased leaves of 14 crops, which are staple foods such as tomatoes, peppers, and potatoes. The dataset has a large number of various plant diseases, including leaf spots, blights, molds, and bacterial spots, and as such is a suitable source for training machine learning models. By offering a large and diverse collection of images, the dataset enhances model performance and ensures that trained models have the ability to generalize well to real-world agricultural environments. This makes it a key element for researchers and developers constructing automated plant disease detection models such as those that are YOLO V8S-based (Varghese, 2024). To enable the Plant Village dataset to be used in training deep learning models to its full capacity, various preprocessing steps are performed before the images are loaded into YOLO V8S. The first thing is to resize all images into a consistent resolution that matches the input specifications of YOLO V8S. This supports uniformity within the dataset, reduces computation demands, and helps in making the model train faster. The second thing is to normalize images so that pixel intensities range between 0 and 1. This enhances stability in training and allows the model to learn more quickly and generalize better



Aside from these fundamental steps, data augmentation methods are also tasked with enhancing the variety of the dataset and preventing overfitting. Various various transformations such as random rotation, flipping, zooming, and cropping are employed to produce various views of the same picture. The augmentations mimic real-world variations in lighting conditions, directions, and environments to make the model more resilient against unseen data. Further, contrast enhancement and colour normalization are used to make the visual presentation consistent so that the images remain clear and informative despite changing lighting conditions during data collection. Through the implementation of these preprocessing steps, data is made suitable for deep learning operations so that the YOLO V8S model can better extract useful features. As a result, the model is more precise in the detection of plant diseases across different conditions and hence constitutes a useful means of improving crop health monitoring and farm yields. The combination of a dense dataset and advanced preprocessing techniques ensures that AI-driven disease detection technologies can provide real-time, accurate solutions to farmers and researchers who are fighting to protect crops against disease outbreaks.

V. Methodology



Development of an effective system for detecting and predicting crop diseases consists of a structured, multi-step process to enhance the accuracy and efficacy of disease detection. Drawing inspiration from Anjnaa (Anjnaa, 2019), this process consists of a number of critical stages including image acquisition, preprocessing, data augmentation, feature extraction, classification, and prediction of disease. Each of these stages is essential in facilitating the YOLO V8S model to effectively detect, classify, and predict plant diseases, rendering it a valuable tool for modern agricultural disease control. The method is explained below:

5.1 Image Acquisition

The initial and most critical process of developing an efficient disease detection system is to obtain high-quality images of crop leaves. These images serve as the foundation for model training, enabling the model to learn and detect disease symptoms efficiently. High-resolution images of healthy and diseased crop leaves are obtained from different agricultural fields and research datasets to provide a varied dataset. To improve the model's resilience to various conditions in the environment, photographs are captured in multiple lighting, orientations, and backgrounds. This strategy ensures that the system is not limited to one environment but is capable of accommodating real-world conditions. The database includes a variety of crops, such as tomatoes, potatoes, peppers, and wheat, and spans several diseases at different stages—from initial indications to full infections. By embracing such diversity, the model can better distinguish normal leaf texture variation from true disease symptoms and hence have better detection accuracy.

5.2 Preprocessing

Following image capture the images undergo a critical pre-processing technique to enhance their quality while removing background noise that could result in erroneous disease detection. The aim of this process is to get images ready so deep learning algorithms can provide accurate analyses from them. As an initial step in the preprocessing process researchers resize all images to conform to YOLO V8S specific input requirements. Keeping the dataset consistent in terms of image sizes using standardized images optimizes processing efficiency by retaining critical visual information. Images necessitate the need to do colour normalization since they are typically captured under varying lighting conditions. The standardization of colour tones across every image helps reduce environmental effects which would otherwise affect the model's learning outcome. The application of contrast enhancement serves to reveal delicate colour transitions which point to plant diseases. Plant infections show up through small colour changes which are hard to detect by eye so contrast adjustment highlights important features for better detection. The subsequent removal of backgrounds together with cropping ensures that only meaningful leaf sections are used for analysis. The model can concentrate exclusively on evaluating leaf diseases because unwanted background elements are removed during the process. A clean and homogeneous image dataset results from these preprocessing operations which leads to better disease detection and classification outcomes for the model.

5.3 Data Augmentation

To enhance the robustness and variability of the model, we applied data augmentation techniques that provide many diverse images to reduce probable overfitting and increase accuracy on unseen data. By transforming the leaves to simulate a different perspective and leaf orientation via operations like flipping, rotation, zooming, and cropping, the conditions of the images will begin to follow the randomness of their orientations by varying the amount of light. By this augmentation, we generate increased training data, improving the model's performance with different environmental complexities when applied in the real world.

5.4 Feature Extraction

In the current stage of this research, the YOLO V8S model is being utilized in feature extraction that is, it will find and extract essential aspects in the disease detection image. YOLO V8S an advanced object detection model and is known to detect multiple affected areas within the same image, thus, it is very efficient at plant disease identification. Apart from this, the model applies automatic feature engineering in which it retrieves patterns like color changes, spots, lesions, and texture differences straight away from the raw image data. Among the largest benefits of YOLO V8S use is its management of high-dimensional data, enabling it to identify intricate patterns signaling the occurrence and severity of the plant disease. Besides enhancing accuracy, this automated system significantly simplifies and accelerates the detection process. The model thus offers improved ability in generalizing from heterogeneous training data for disease diagnosis across various plant species and environmental conditions, thereby constituting a solid tool for agricultural use.

5.5 Classification: Healthy/Diseased

YOLO V8S then moves to classification after feature extraction, classifying the crop leaves as healthy or diseased. Through learning from training, it knows certain symptoms to label the diseased portions appropriately. In training, it learns from images that are labeled healthy and diseased leaves by calling its attention to pattern recognition that differentiates the two. Such specific categorizations enable farmers to act in time and focus on the area that should be treated instead of spraying chemicals on bowls. This creates localized intelligence to enhance disease control by minimizing chemical application and maximizing plant health, resulting in higher yield and sustainability.

5.6 Disease Prediction

In the disease prediction phase, the final step in the model extracts specific types of diseases affecting the crop. YOLO V8S is optimized for multi-class classification, making it possible to identify multiple diseases from a single image. During this analysis, the model identifies conditions such as leaf blotch, powdery mildew, or rust and provides a score of confidence for each prediction. This proactive ability to diagnose represents the key to early and accurate disease management. Instead of being dependent on the time-consuming and error-prone manual procedure for diagnosing the crop, farmers are given precise knowledge of the likely disease that may be exposing their crop. The system does not only identify the disease but also provides probability scores to assist farmers in making well-informed decisions concerning the applicable treatment. The timely intervention reduces crop losses and the indiscriminate use of pesticides and supports sustainable farming that eventually results in healthier yields and improved agricultural productivity.

Training and Evaluation of the Model

For this model, a very curated and augmented dataset is used to train it, thus the errors which contribute additional improvement are value-injected through backpropagation and optimization procedures. A part of the dataset is reserved as validation in an attempt to have the model generalize well, hence providing an evaluation of the performance on data the model did not see before-this helps minimize the chances of overfitting. Throughout the model's training, there is a constant tracking of essential performance metrics, such as accuracy, precision, recall, and F1-score to measure model efficiency in classifying the disease. Additionally, the model's disease identification and area localization effectiveness is assessed through IoU by comparing model predictions with actual affected parts. Such measures are tracked continuously, and hyperparameter fine-tuning for better performance to attain maximum accuracy and belief in disease detection in real situations is assured.

Deployment

In order to deploy our Crop Disease Detection System, A well thought through blend of heavyweight tools and techs that vibe across multiple platforms seamlessly. The backend is powered by Django, a high-performance and scalable Python web framework. Using Django helps us build a custom good server that talks to machine-learning model on one end and frontend another. It takes user requests, handles image uploads and returns disease predictions to consumers in a frictionless manner.

The model of machine-learning processes the PlantVillage dataset is available on Google Colab where there is free access to the high performance of cloud computing resources such as GPUs. Run the YOLO V8S model on Google Colab to check real-time predictions without using the users local machine. This cloud-opposed process allows large datasets and heavy computations to be run quickly which keeps your system available to users irrespectively of their device restrictions. We use Bootstrap for the frontend that makes sure web interface is responsive and usable on desktops (mvista, mobile etc.) by keeping it design-friendly and accessible. It permits users to upload images of crop leaves, get disease predictions with a single click and opens an insight for farmers. We set up a deployment resource that is both effective as well as scalable by using these technologies. Which creates a simplified, yet powerful platform for farmers to use with comparative ease at confirming disease quickly and accurately so they can then manage their crops intelligently, therefore increasing the overall production of agriculture.

VI. YOLOv8s and Its Benefits

6.1 Overview of YOLOv8s

YOLOv8s is a real-time object detection system renowned for its speed and accuracy. Unlike traditional object detection algorithms that propose regions of interest before classifying them, YOLOv8s employs a single neural network to predict bounding boxes and class probabilities simultaneously. This approach yields significant time savings, making it ideal for real-time applications in agriculture.

6.2 Advantages in Crop Disease Detection

Use of YOLOv8s for crop disease detection provides multiple advantages, including-Real time detection of diseases: Being able to process images in real time makes it easier and better for detecting crop diseases, hence predicting diseases. Higher Accuracy: YOLOv8s, with fine-tuning, delivers very high accuracy along with detection of multiple objects in a single frame, which is indeed beneficial for identifying several diseases in a single crop leaf. Scalability: The model can scale across different crops and a range of disease symptoms, to enable farmers to monitor diverse farming environments.

VII. Conclusion

YOLOv8s is an awesome technology for crop disease detection and prediction. The integration of YOLOv8s into AgriShield is one of the monumental leaps taken toward crop disease detection and prediction, offering farmers a real-time solution to arrive at accurate methods for combating agriculture challenges. YOLOv8s has the potential to work with high-dimensional data, laying a solid foundation for excellence in agricultural space. This innovation opens the way for all-round systems suitable for multi-class classification in order to pollinate action-oriented insights for farmers, thus fostering much more sustainable and efficient agricultural practices.

I. Acknowledgment

All authors declare that they have no conflicts of interest.
All participants provided informed consent before taking part in this study. Personal data was anonymized to ensure confidentiality.

References

Anjnaa, M. S., 2019. Hybrid System for Detection and Classification of Plant Disease.
Dhakne, M. P. M. a. D. A., 2024. Revier of medical plant classification system.
Harjeet Kaur, D. P. a. V. K., 2023. DEEP CNN MODELS IN PLANT DISEASE IDENTIFICATION.
IEEE, 2018. Machine Learning and Deep Learning Methods for Cybersecurity. p. 35381.
Jeyakumar2, M. C. ·. M. K., 2023. Survey on crop pest detection using deep learning and machine learning approaches.
Mr. Gopinath V, I. V. a. N. R., 2023. Plant disease prediction using hybrid model.
Muhammad Shoaib, B. S. S. S. a. A. A., 2023. An advanced deep learning models-based plant disease detection: A review of recent research.
Musha, S. L. a. Y., 2023. CEMLB-YOLO:Efficient Detection Model of Maize Leaf Blight in Complex Field Environments.
Nauman Qadaeer, T. S. a. M. A. K., 2023. A Convolutional Neural Network Model for Wheat Crop Disease Prediction.
Nishant Shelar, S. S. a. S. S., 2023. Plant Disease Detection Using Cnn.
O'Shea, K. T., 2015. An Introduction to Convolutional Neural Networks.
Rosemary Ngozi Ariwa, C. M. a. N. G. T., 2024. Plant diseasse detection using YOLO machine learning approach.
Varghese, R., 2024. YOLOv8: A Novel Object Detection Algorithm with Enhanced Performance and Robustness.