

Paper Type: Original Article

Smart Agriculture for Beans: AI-Driven Disease Identification and Classification

Ahmed M. Ali ^{1,*}  and Hussam Elbehier ^{2,3} 

¹ Faculty of Computers and Informatics, Zagazig University, Zagazig 44519, Sharqiyah, Egypt; aabdelmounem@zu.edu.eg.

² Faculty of Information Systems and Computer Science, October 6th University 12585, Egypt; 212103294@o6u.edu.eg.

³ Vanridge University, USA; drhussam@vru-edu.net.

Received: 23 Dec 2024

Revised: 02 Mar 2025

Accepted: 28 Jun 2025

Published: 30 Jun 2025

Abstract

Introduce deep learning methods for automatic detection in this paper of plant disease using YOLOv8. The algorithm was developed using a unique dataset that included pictures of beans. also applied segmentation for the leaf to improve the localization of disease, the incorporated leaf segmentation techniques using color-based thresholding and morphological operations, which help isolate the affected regions from complex backgrounds. compared the results with previous work that depended on models like EfficientNet. The results show that the model achieved high accuracy and precision compared to previous work . The model reached an accuracy of 99.5%, Moreover the model provides real-time object detection and precise localization, making it suitable for practical agriculture applications more than the traditional ways.

Keywords: Bean Diseases; Recognition; Segmentation; YOLOV8.

1 | Introduction

The plants are important to humanity as they are one of the reasons for human survival on the Earth, and they represent food security for humanity, and agriculture is an essential factor for all living organisms [1].

So, the problem of plant disease and the insects that infect them greatly affects humanity and the prosperity of the nation and affects the production and quality of crops [2], and the traditional way to detect disease by the human Eye or by the Agricultural Experts and old, slow and not accurate so here that is the rule of Ai (deep learning) [3].

Detecting the plant disease at initial stages before the disease destroys all plants is critical for minimizing crop loss and ensuring agricultural productivity [4]. Traditional methods are manual, slow, and subject to human error, often resulting in delays or inaccurate. Moreover, the conventional machine learning models, such as CNNs and SVMs, focus on classification without providing the location of infected areas.

Contribution of the article:

- Develop an object detection model using YOLOv8 for identifying plant diseases.
- Segment the leaf and localize the affected areas on the plant leaves.



Corresponding Author: drhussam@vru-edu.net



<https://doi.org/10.61356/j.mawa.2025.7598>



Licensee **Multicriteria Algorithms with Applications**. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0>).

- Compare the proposed model's performance with traditional methods like CNN, SVM, and VGG.
- Achieve high accuracy and precision in detecting multiple plant disease types.

Contributions:

- Make an algorithm that takes a normal photo of a leaf and makes a ponder box for the disease to use the photo as a dataset, so the model can train on this photo with the ponder box.
- Custom YOLOv8 model trained for detecting disease in bean plants.
- Make a segmentation for the leaf to enhance detection accuracy.
- Comprehensive comparison with conventional deep learning models in terms of accuracy, precision, and localization.
- Achieving a detection accuracy of 99.5%, outperforming existing methods in classification and localization tasks.

2 | Literature Review

Several studies have explored some plant diseases and pest diseases by using deep learning and showing us the accuracy of each model that was used, and this is shown in Table 1.

Table 1. Comparison of Previous Studies on Plant Disease and Pests Using Deep Learning.

Name of paper	Model/methods	Authors	Year	Accuracy
Automation in Agriculture by Machine and Deep Learning Techniques [5]	FCN (Fully Convolutional Networks)	Muhammad Hammad Saleem1 · Johan Potgieter2 · Khalid Mahmood Arif1	2021	83.9%
Enhanced Pest and Disease Detection in Agriculture Using Deep [6]	CNN	Wenqi Li1, Xixi Han1*, Zhibo Lin1, Atta-ur Rahman2	2023	93.5%.
Automated Machine Learning Classification Framework to Predict Crop Yield and Detect Pest Patterns [7]	CNNs and deep learning	Gopi R1*, Tamil Selvi M2, Saranraj G3, Nagaraj P4, Parthiban K5 and Ranjith Kumar A6	2024	92.6%,
A Deep Learning and Social IoT approach for Plant Disease Prediction toward a Sustainable Agriculture [8]	DenseNet121, MobileNet, MobileNetV2, NASNetMobile	Giovanni Delnevo, Roberto Girau*, Chiara Ceccarini, and Catia Prandi	2015	94.58%
Revolutionizing Agriculture: Artificial Intelligence Assisted Plant Leaf Disease Detection using Deep Learning Principles [9]	CNN	V.S.Prakash, R. Vanitha, Sudam Vasant Nikam, Athappan V, R. Krishnamoorthy, S. Arun	2024	98.29%

This table summarizes key information from previous research on plant disease and pest detection, including methods applied and reported accuracy.

3 | Dataset

3.1 | Dataset Description

The data sets used for bean plants contain images, and they are divided into three classes representing several types of diseases. It consists of 3600 images divided into three classes, which represent diverse types of bean disease. Split into three sets (70%) for training, (15%) for validation, and (15%) for testing. As shown in Table 2, this table shows the distribution of the study's dataset, such as the names of the plants and the quantity of samples in the test, validation, and training sets.

Table 2. Distribution of the Dataset for plant disease and pest classes.

plant	Training set	Validation set	Evaluate set
Bean	2880	360	360

The dataset ensures a balanced representation of bean disease classes across all subsets, which contributes to robust model performance and reliable evaluations.

3.2 | Sample Images from the Dataset

As shown in Figure 1, the dataset includes images captured under different circumstances to guarantee the model's resilience in practical situations.

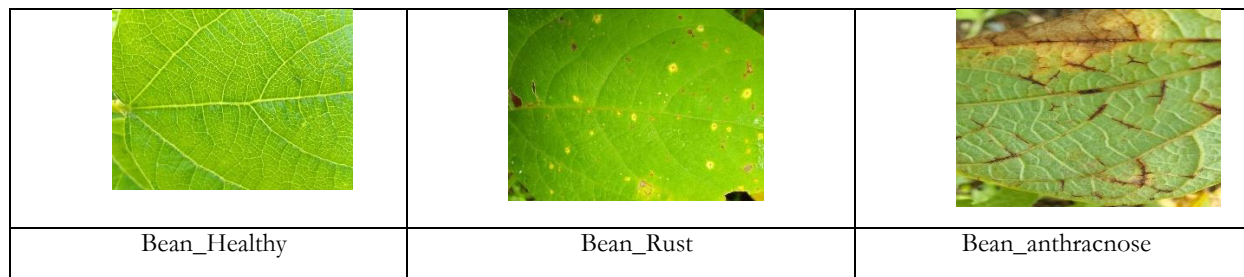


Figure 1. Sample images of bean disease are included in the dataset.

3.3 | Preprocessing Techniques

To ensure high-quality input, the following preprocessing steps were applied:

Annotation: the annotation process was performed by making an algorithm that uses image processing techniques, bounding boxes were generated around the infected areas or pests based on color segmentation and contour detection methods, and then, after making this, the annotation was saved in yaml file, which includes the class label and coordinates for each bounding box, as shown in Figure 2.

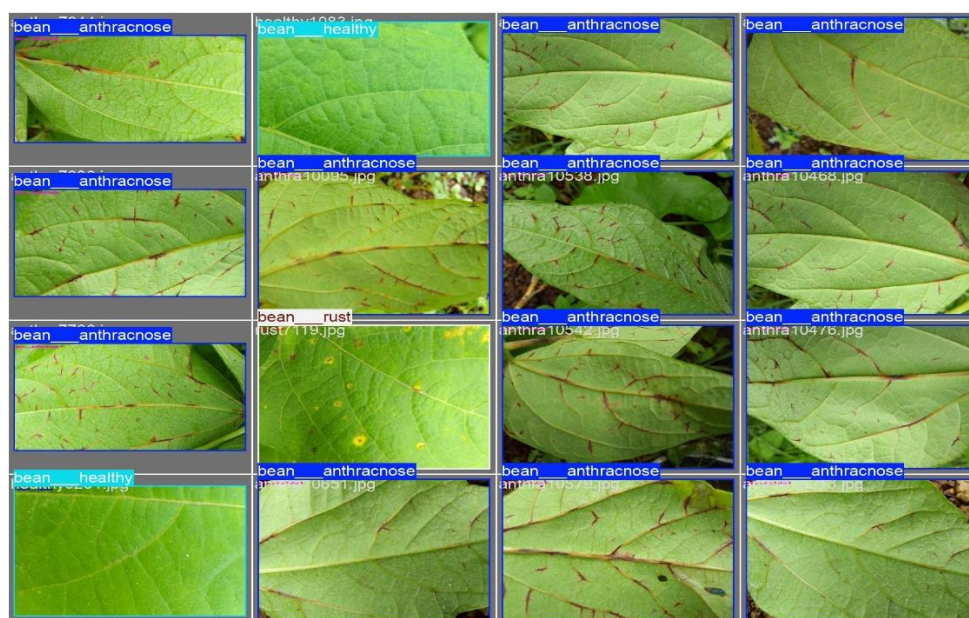


Figure 2. Sample of annotation images of plant disease and pests that are included in the dataset.

Resizing: all images were resized to a fixed resolution suitable for YOLOv8 input (640×640 pixels) as shown in Figure 3.

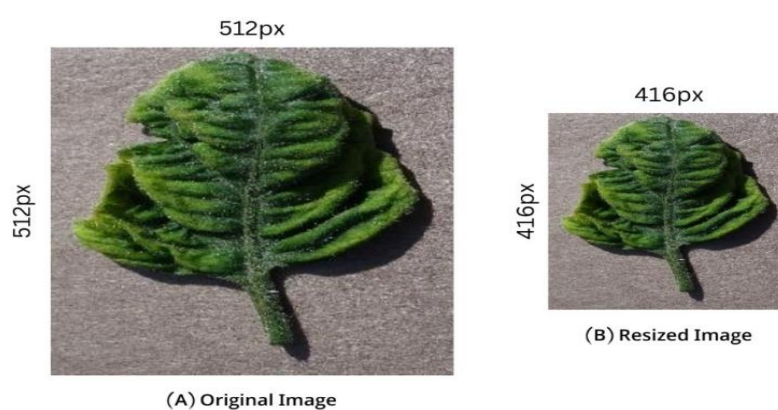


Figure 3. Sample of an image showing the effect of resizing.

Data cleaning was applied to remove the blurry or duplicate images.

Segmentation: applied to make a segmentation to remove the background from the image to make the model focus on the leaf to check if it is good or not, as shown in Figure 4.



Figure 4. Sample of the images before and after the segmentation.

The data was split the data into 70% for training 15% for testing, and 15% for validation.

4 | Methodology

4.1 | Model Architectures

Figure 5 presents the architecture of the YOLOv8n object detection model that was used this model to detect bean diseases [10].

YOLOv8 is a real-time object detection that is best from previous versions with several improvements in accuracy, speed, and flexibility [11].

The model architecture has three main components:

- Backbone:

A feature extractor that uses a modified version of CSPDarknet to extract rich spatial features from the input images.

- Neck:

A feature aggregation layer, based on PAN (Path Aggregation Network), that helps in combining low-level and high-level features for better object localization.

- Head:

In charge of final forecasts, which include object confidence scores, class probabilities, and bounding box coordinates.

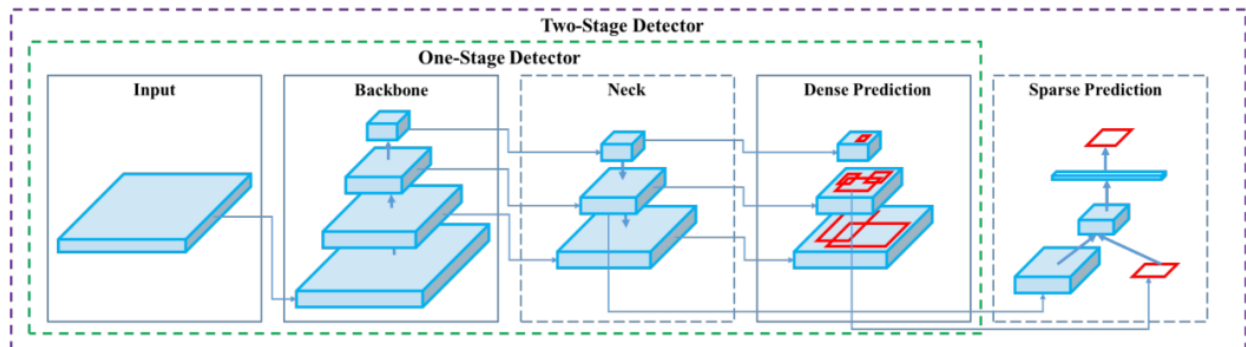


Figure 5. The architecture of YOLOv8n.

4.1.1 | YOLOv8n

It is the nano version of the YOLOv8 family, designed to offer a good balance between speed, model, size, and detection accuracy. It is especially suitable for real-time applications and devices where computational resources are limited.

The choice of YOLOv8n in this work was based on its lightweight structure and fast inference speed,

This gives us training and deployment on large datasets without needing high-end hardware or computational resources [12, 13].

In the bean dataset:

As shown in Figure 6 shows the connection between each dataset class's F1 score and the confidence threshold: Bean_anthraconse, bean_healthy, bean_rust.

Figure 7 illustrates how precision varies with different confidence thresholds for each class (Bean_anthraconse, bean_healthy, bean_rust).

And in Figures 8 and 9: show that the images of bean in validation phase before the model evaluate them self and after the model takes a photo and gets results.

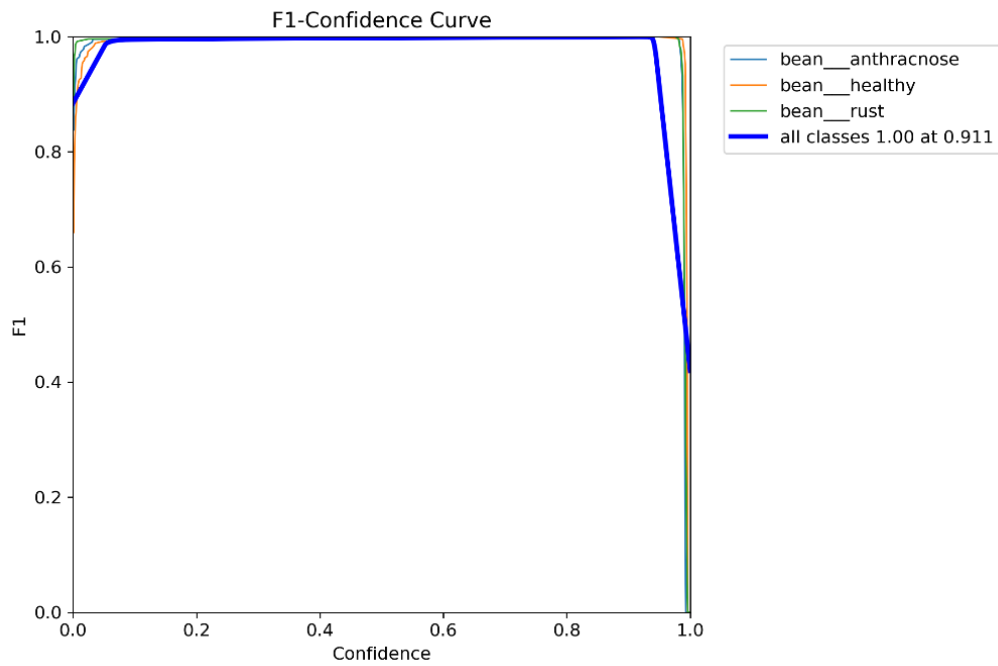


Figure 6. F1-Confidence Curve for Bean Disease classification.

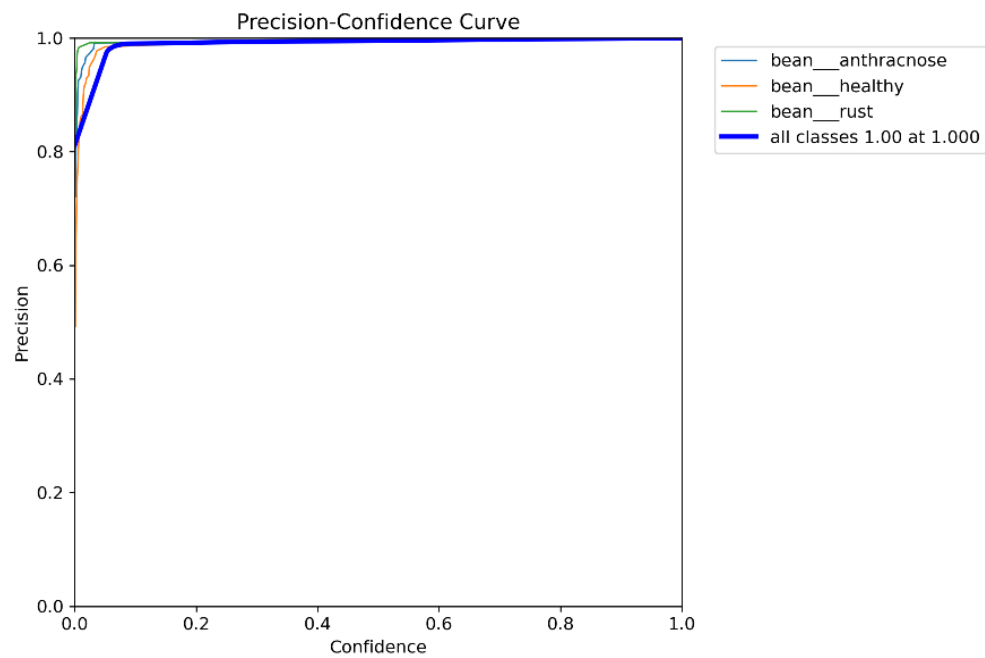


Figure 7. Precision-Confidence Curve for Bean Disease classification.

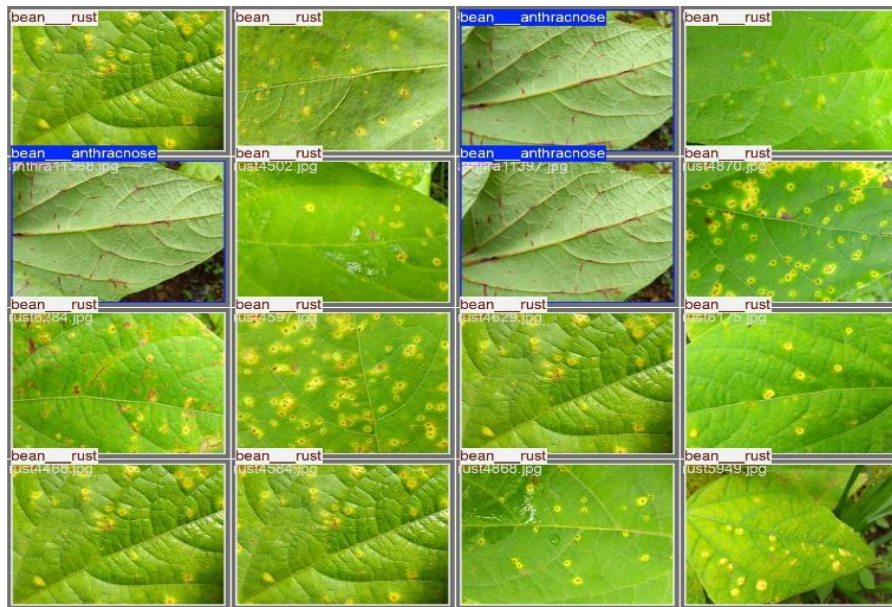


Figure 8. Images for bean in validation phase before model predict.

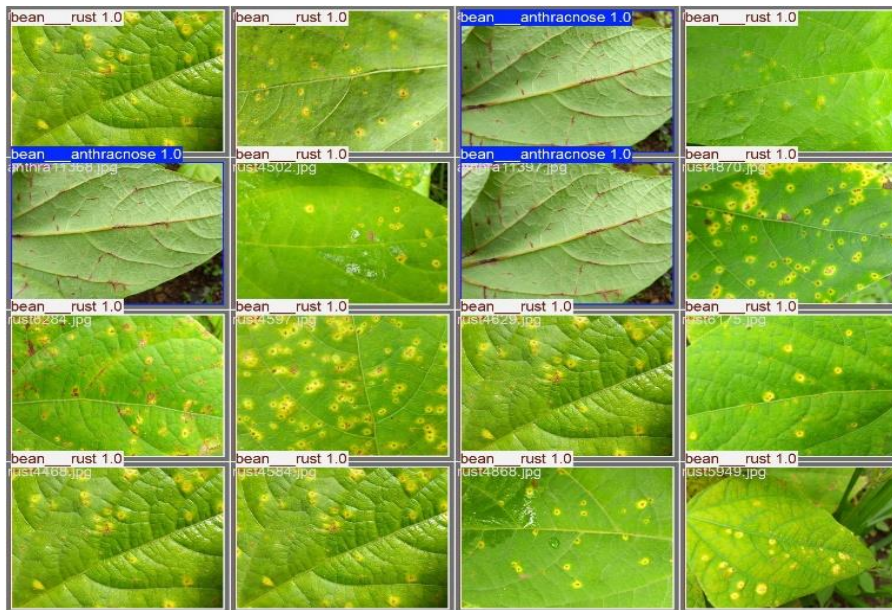


Figure 9. Images for bean in the validation phase after model prediction.

4.2 | Training Procedure

Training model using YOLOv8n is implemented using Ultralytics YOLOv8 library, training was performed in a Google Colab environment using a GPU to accelerate computing.

Optimizer:

Stochastic Gradient Descent (SGD) optimizer with a weight decay of 0.0005, momentum of 0.937, and learning rate of 0.01 by default. The learning rate is scheduled using a cosine decay strategy over the training epochs.

Loss Function: The model uses a composite loss function that includes:

- Box Loss (CIoU-based) for bounding box regression,
- Objectness Loss (Binary Cross-Entropy) to predict object presence,
- Classification Loss (Binary Cross-Entropy) for multi-class classification per anchor.

Batch Size: eight images in a batch.
Epochs: 50 epochs with early stopping based on validation loss.

4.3 | Evaluation Metrics

The following metrics were used to evaluate model performance:

- Accuracy: the percentage of photos that were correctly classified.
- Precision and Recall: To evaluate class-specific performance.
- F1-Score: precision and recall harmonic meaning.
- Inference Time: The amount of time needed to categorize one picture.

5 | Results and Discussion

5.1 | Model Performance

The performance of the evaluated architecture is summarized in Table 1. YOLOv8 emerged as the best-performing model, achieving an accuracy of 99.5% with a loss of 0.00170.
As it shows in Table 3, the table describes each plant’s models and pest model that I used and the accuracy, loss, precision, and recall.

Table 3. Performance metrics of the models.

Name/model	Accuracy	precision	Recall
Bean model	99.5%	99.3%	99.6%

This table displays accuracy, precision, and recall while summarizing the performance of the models under evaluation.

5.2 | Insights from Comparative Analysis

- Accuracy vs. Efficiency: YOLOv8n is appropriate for deployment because it strikes a compromise between accuracy and runtime efficiency.
- Impact of Preprocessing: Data segmentation significantly improved model generalization, with this segmentation of datasets showing a 5-10% increase in accuracy.
- The heatmap: in Figure 10, this gives a graphic summary of the relationships between the data set’s distinctive properties. Using color gradients, it highlights areas where certain features are strongly correlated, which can inform decisions about feature engineering and model optimization. Understanding these relationships can help in reducing multicollinearity and improving the accuracy of the predictive models.

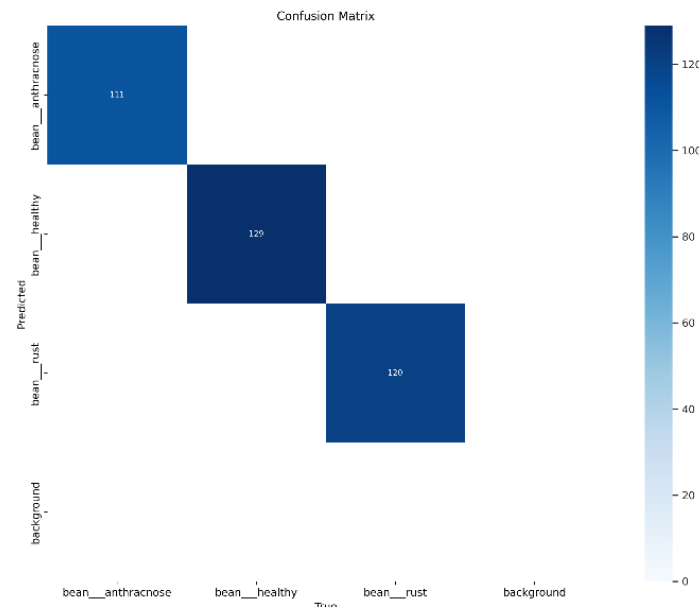


Figure 10. confusion matrix for bean diseases.

6 | Conclusion and Future Work

The goal of the TP, a basic optimization problem in operations research, is to find the least costly way of distributing goods, given supply and demand restrictions, from multiple sources to multiple destinations.

6.1 | Conclusion

In this work, we suggested a thorough deep learning method for identifying bean diseases. Using the YOLOv8 architecture. The model was developed and evaluated: disease detection in plant-like beans. The model achieves precision, recall, and accuracy values consistently of 99.5%.

Specifically, the models trained on datasets, ranging from 3500 to 3600 images were validated and evaluated on well-balanced sets.

Compared to recent studies, as shown in Table 1, the models achieved better results, especially in precision and recall. This shows that YOLOv8 is a strong and reliable choice for detecting plant diseases. Aim in the future to improve the system by evaluating it in more plant types and using it in real-time applications like drones.

6.2 | Future Work

And building on the insights gained, future work can explore the following directions:

- **Model Optimization:** Examine quantization and pruning strategies to enhance edge device inference speed and minimize model size.
- **Dataset Expansion:** include more plant classes and real-world images from varying conditions to enhance the robustness of the models and generalization.
- **Hybrid Approaches:** Experiment with combining YOLOv8 with Transformer-based models' mechanisms to capture more complex patterns in plant diseases and pest appearances.
- **Real-Time Field Deployment:** Build a real-time monitoring system that integrates trained models with devices such as drones, smartphones, or Raspberry Pi-based sensors to support instant detection.

Funding

This research has no funding source.

Data Availability

The datasets generated during and/or analyzed during the current study are not publicly available due to the privacy-preserving nature of the data, but are available from the corresponding author upon reasonable request.

Conflicts of Interest

The authors declare that there is no conflict of interest in the research.

Ethical Approval

This article does not contain any studies with human participants or animals performed by any of the authors.

References

- [1] Nurudeen, Habeebullahi. The Importance of Plants to the Environment - Green Habitat Initiative. 2 Aug. 2022, <https://www.greenhabitat.ng/the-importance-of-plants-to-the-environment/>.
- [2] Biosciences, National Research Council (US) Committee on. 'Plant Diseases and Insect Pests.' New Directions for Biosciences Research in Agriculture: High-Reward Opportunities, National Academies Press (US), 1985. [www.ncbi.nlm.nih.gov](https://www.ncbi.nlm.nih.gov/books/NBK216440/), <https://www.ncbi.nlm.nih.gov/books/NBK216440/>.
- [3] Khakimov, A., et al. Traditional and Current-Pro prospective Methods of Agricultural Plant Diseases Detection: A Review.' IOP Conference Series: Earth and Environmental Science, vol. 951, no. 1, Jan. 2022, p. 012002. DOI.org (Crossref), <https://doi.org/10.1088/1755-1315/951/1/012002>.
- [4] Demilie, Wubetu Barud. 'Plant Disease Detection and Classification Techniques: A Comparative Study of the Performances.' Journal of Big Data, vol. 11, no. 1, Jan. 2024, p. 5. BioMed Central, <https://doi.org/10.1186/s40537-023-00863-9>. CloseDeleteEdit
- [5] Saleem, Muhammad Hammad, et al. 'Automation in Agriculture by Machine and Deep Learning Techniques: A Review of Recent Developments.' Precision Agriculture, vol. 22, no. 6, Dec. 2021, pp. 2053–91. Springer Link, <https://doi.org/10.1007/s11119-021-09806-x>.
- [6] Li, Wenqi, et al. 'Enhanced Pest and Disease Detection in Agriculture Using Deep Learning-Enabled Drones.' Acadlore Transactions on AI and Machine Learning, vol. 3, no. 1, Jan. 2024, pp. 1–10. DOI.org (Crossref), <https://doi.org/10.56578/ataiml030101>.
- [7] Academia.Edu - Find Research Papers, Topics, Researchers. https://www.academia.edu/signup?a_id=120708753. Accessed 6 May 2025.
- [8] Delnevo, Giovanni, et al. 'A Deep Learning and Social IoT Approach for Plants Disease Prediction Toward a Sustainable Agriculture.' IEEE Internet of Things Journal, vol. 9, no. 10, May 2022, pp. 7243–50. IEEE Xplore, <https://doi.org/10.1109/JIOT.2021.3097379>.
- [9] Prakash, V. S., et al. Revolutionizing Agriculture: Artificial Intelligence Assisted Plant Leaf Disease Detection Using Deep Learning Principles.' 2024 10th International Conference on Communication and Signal Processing (ICCSP), 2024, pp. 586–91. IEEE Xplore, <https://doi.org/10.1109/ICCSP60870.2024.10543984>.
- [10] Boesch, Gaudenz. 'YOLOv8: A Complete Guide [2025 Update]'. Viso.Ai, 18 Dec. 2024, <https://viso.ai/deep-learning/yolov8-guide/>.
- [11] Reis, Dillon, et al. Real-Time Flying Object Detection with YOLOv8. arXiv:2305.09972, arXiv, 22 May 2024. arXiv.org, <https://doi.org/10.48550/arXiv.2305.09972>.
- [12] Ultralytics. YOLOv8. <https://docs.ultralytics.com/models/yolov8>. Accessed 6 May 2025
- [13] Katumba, A., Okello, W. S., Murindanyi, S., Nakatumba-Nabende, J., Bomera, M., Mugalu, B. W., & Acur, A. (2024). Leveraging edge computing and deep learning for the real-time identification of bean plant pathologies. Smart Agricultural Technology, 9, 100627.

Disclaimer/Publisher's Note: The perspectives, opinions, and data shared in all publications are the sole responsibility of the individual authors and contributors, and do not necessarily reflect the views of Sciences Force or the editorial team. Sciences Force and the editorial team disclaim any liability for potential harm to individuals or property resulting from the ideas, methods, instructions, or products referenced in the content.