JOURNAL OF ELECTRONIC VOLTAGE AND APPLICATION
e-ISSN: 2716-6074

JEVA

Vol. 6 No. 2 (2025) 126-139
https://publisher.uthm.edu.my/ojs/index.php/jeva

Performance Analysis of YOLOv8, YOLOv9, and YOLOv11 for Corn Leaf Disease Detection

Shafieka Nadia Sabarudin[1], Nik Shahidah Afifi Md Taujuddin[1*], Suhaila Sari[1], Ahmad Raqib Ab Ghani[2]

[1] *Department of Electronics Engineering, Faculty of Electrical and Electronic Engineering, Universiti Tun Hussein Onn Malaysia, Parit Raja, 86400 Batu Pahat, Johor, MALAYSIA*
[2] *Department of Civil Engineering, Faculty of Civil Engineering and Built Environment, Universiti Tun Hussein Onn Malaysia, Parit Raja, 86400 Batu Pahat, Johor, MALAYSIA*

*Corresponding Author: shahidah@uthm.edu.my
DOI: https://doi.org/10.30880/jeva.2025.06.02.012

Article Info

Received: 16 July 2025
Accepted: 5 November 2025
Available online: 9 December 2025

Keywords

YOLOv8, YOLOv9, YOLOv11, corn leaf disease, deep learning, image processing, precision agriculture, Convolution Neural Network (CNN).

Abstract

Corn is a crucial crop for agricultural yield and food security, yet it faces significant threats from various foliar diseases that can diminish both growth and quality. Conventional visual assessment techniques often require substantial labor and are susceptible to errors in diagnosis. This study introduces a Corn Leaf Disease Detection System that employs Convolutional Neural Networks (CNNs) and evaluates the performance of three models: YOLOv8, YOLOv9, and YOLOv11. The methodology involves capturing high-resolution images of corn leaves afflicted by diseases such as Northern Leaf Blight and Common Rust. These images undergo a pre-processing phase to enhance clarity and are standardized for input into the detection framework. CNN is employed for intricate classification tasks, while YOLOv11 is utilised for real-time disease detection. The dataset comprises 3,000 images, which are augmented to expand the training set to 6,000 samples. Among the evaluated models, YOLOv11 demonstrated superior performance, achieving an F1-score of 0.93, with precision at 0.94 and recall at 0.92 by epoch 100. These findings highlight the system's operational efficiency and robustness in effectively detecting corn leaf diseases.

1. Introduction

Crop diseases are becoming a great source of concern to farmers today because there are enormous losses in agricultural production, not only in quantity of agricultural produce but also in its quality [1]. Some of these are potato (Solanum tuberosum), corn (Zea mays) which are very susceptible to leaf diseases. Northern Leaf Blight, Common Rust, and Southern Leaf Blight, which affect the photosynthetic capability of the plant and may become an epidemic should they not be treated. These diseases can be so severe that the crops affected may have to be destroyed and in some extreme cases the farmer may end up abandoning the field for some time, causing monetary loss as well as leaving the food supply very unstable. There is no existing commercial diagnostic system to detect these diseases in their early stages on a widespread level, which means that most of the time these diseases can be only identified when symptoms are already high. One aspect that is emerging with good results is the use of deep learning and computer vision because with these researchers have been able to identify and classify the different diseases of the corn leaves with high precision resulting in the possibility of more efficient and automated agricultural surveillance [2]. Although they are severe, existing practices in detecting corn leaf diseases are still carried out visually by farmers who are not accurate most of the time due to the tedious nature of the practices [3].

2. Literature Review

2.1 Corn Leaf Disease

Corn (Zea mays) is one of the most important cereal crops in the world, only next to rice and wheat by their overall world production. But productivity is also severely affected by the numerous foliar diseases which endanger yield as well as quality. These diseases are of immense importance in sustaining agricultural production and food security since they can be accurately diagnosed and treated in time. Two of the main diseases of corn leaf determined by the latest works are Northern Corn Leaf Blight (NCLB) caused by Exserohilum turcicum that is the most common and the most harmful. Maydis Leaf Blight (MLB) aka Southern Corn Leaf Blight (SCLB), caused by Bipolaris Maydis is generally prevalent in humid climatic regions. Banded Leaf and Sheath Blight (BLSB) created by Rhizoctonia solani, that arises in South and Southeast Asia. Common Rust: is a disease caused by Puccinia sorghi which shows reddish-brown pustules and Gray Leaf Spot (GLS): is an illness brought about by Cercospora zeae-maydis or Cercospora zeina which evokes serious leaf flowering. And finally, it is Brown Spot which is caused by Physoderma maydis which normally occurs in localized regions like Nebraska. Such diseases have various symptoms that include lesions, necrotic spots, and discoloration, which upon observation, normally appear on the leaves and can be taken as major symptoms of diagnosis. Common ways of diagnosing diseases are usually resource consuming and slow. Therefore, the combination of image-based methods and deep learning algorithms, especially Convolutional Neural Networks (CNNs), provides an effective approach that can be used to identify diseases as fast as possible and allow farmers to intervene as early as possible [4].

2.2 You Only Look Once (YOLO) Algorithm

The YOLO (You Only Look Once) architecture is a popular real-time object detections framework that makes detections by processing the whole pictures in one pass via convolutional neural network (CNN) allowing to do detections fast and precisely. YOLOv8 is architecture, where performance is improved through the more efficient backbone networks, anchor-free detection, and work with such tasks as segmentation and classification. In this research work [5], the author proposes computer vision software to identify defects of mango fruits black spot, brown spot, and mango scab using YOLOv8. The number of augmented images in a controlled setting was 2,160 people and trained the model across 300 epochs. The trained Model has been put into a desktop application that was built using Tkinter to realize real-time defect detection based on a high-definition camera setup. It completed the evaluation with great performance and averaged 95% Average Precision (mAP) at IoU 0.50 and average recall of 93 per cent. Class-specific accuracy was found to be 100 percent on black spot, 80 percent on brown spot and 83.33 percent on mango scab respectively and an overall accuracy of 87.7 percent was found out in real-time. Hence, all these findings prove that YOLOv8 can be effective in automated defect inspection in fruit based on quality, but still some modifications are required to be done to develop variations in size and texture.

The YOLO (You Only Look Once) line of real-time object detecting algorithms grew through the implementation of the next version, YOLOv9, that induces the Generalized Efficient Layer Aggregation Network (GELAN) to enhance accuracy, scalability, and speed. The research paper of [6] is an evaluation of the YOLOv9 model of traffic accidents detection on 4, 019 labeled pictures. The test cases where the model was run were three-Yolov8 (100 epochs), Yolov9 (100 epochs), and Yolov9 (150 epochs). YOLOv9 compared to YOLOv8 achieved better results on all the important metrics, i.e., precision 94.2%, recall 85.6%, and mAP50-95 76.6, proving its resilience and efficiency of detecting traffic accidents in different conditions. Compared to SELAN that could lose more information, GELAN architecture boosts feature aggregation as well as limits information loss and precision in complicated situations is consequently higher. YOLOv9 is more stable and precise than YOLOv8 that reached 89.8 percent precision and 74.2 percent mAP50-95, becoming a potential candidate to implement real-time systems to detect accidents and other safety-related scenarios.

In the paper [7], the author demonstrates a real-time detection of strawberry ripeness via the YOLOv11 deep learning model to improve the automated harvesting. Through the labeled dataset of 3,100 strawberry images and training the different models of YOLOv11 (Nano, Small, Medium, Large), the researchers discovered that the most effective one was the Medium (YOLOv11-m), which offers the precision rate of 88.3%, Recall of 85.7%, and mAP@0.5 which is 92.8%. The system has shown good performance in difficult scenarios such as poor lighting conditions and fruit overlaps, and thus it can be deployed in real time embodied in agricultural robots. This paper points out the AI-based embedded system integration to automate agriculture, providing a low-cost and scalable capable of ripeness detection of fruits. It also covers the main problems of precision farming and illustrates how object detection models can help in the optimization of the perfect harvest time and save manual labor and increase the quality of the harvest.

J. of Electronic Voltage and Application Vol. 6 No. 2 (2025) p. 126-139

2.3 Previous Research About Corn Leaf Disease Detection

This recent research has added greatly to the development of corn leaf diseases through deep learning. Paper [8] suggested a model based on Convolutional Neural Network (CNN) since the model demonstrated a high accuracy of 98.78% in detecting three well known diseases namely Cercospora Leaf Spot, Common Rust, and Northern Leaf Blight. This model facilitates fast and precise detection quickly with minimal time of training due to its insistence on efficient training based on optimization of the architecture and preprocessing strategies like resizing images and normalization.

Based on the deep learning practice, paper [9] created a real-time system of classifying corn leaves as disease-affected and healthy using the YOLOv5 model with GoPro cameras. It has features of autonomous disease detection based on high-resolution images and unsupervised learning to target flexibility. Producers receive real-time alerts on their smartphones so that things can be mitigated in time. Diseases such as rust, leaf blight, eyespot and gray leaf spot can be efficiently identified using the model and the identification tool helps in sustaining farming since the model makes results available in a cost-effective manner that can be deployed in the field.

In addition to these methods, paper [10] have singled out the Corn Gray Leaf Spot (CGLS) disease, to introduce a solution using CNN based model that performs a multi-classification of the disease in five levels. The model was trained on a self-composed dataset that resulted in an accuracy value of 95.33 to identify high-severity cases of severity. The work demonstrates the relevance of the severity-level detections in the field of disease management and contributes to future research, which will elaborate on the current study with a more powerful real-world dataset and an opportunity to transform it into practice.

Northern Corn Leaf Blight (NLB) was explored under the field-based conditions where an optimized YOLOv3 model was applied in paper [11]. Such additions as Dense blocks and modules CBAM were introduced to enhance such accuracy without any loss in performance speed. The object detection was reannotated based on the dataset, and the given Dense-Attention model outperformed the base YOLOv3 in terms of its efficiency and accuracy in the setting of the complicated environment, achieving 0.821 of the AP0.5.

Based on its related objectives, another previous work used the YOLOv8 model to detect some of the prevalent corn diseases such as the gray leaf spot, and southern rust. It employs the latest functionalities like Squeeze-and-Excitation and Convolutional Block Attention Modules to have good metrics of 89.3 percent precision, 91.7 percent recall and a notable 0.940 mAP. These findings show that YOLOv8 can easily address complicated conditions in the field with high detection scale and short inferences distance thus highly appropriate in smart agriculture and real-time disease tracking [12] .

Lastly, researchers in the Philippine setting created a corn disease detector system with the help of the YOLOv5 module to detect corn rust, leaves blight, and gray leaf spot. Compared to their model, theirs resulted in a mean average precision (mAP) of 97.5 and detection accuracy of between 98.9 and 99.43. Running on a phenotyping device, the process showed live image identification, and the inference rate inside the system averaged 1.11 seconds. These results evidence the possibility of the use of the YOLO-based models in practice to automatically identify a disease within the agriculture industry [13].

3. Methodology

The flowchart (as in Fig. 1) illustrates the overall process of training a corn leaf disease detection model. The model is tested Google Colab. It begins with image data acquisition, followed by segmentation and augmentation to enhance dataset quality. Initially, the dataset consisted of approximately 3,000 images. After applying pre-processing and augmentation techniques, the dataset was expanded to around 6,000 images. The processed dataset was then split into training, validation, and testing sets using an 80:10:10 ratio. This ratio was chosen based on its proven effectiveness in achieving high accuracy and low loss, as supported by the referenced study on tomato leaf disease classification [14], where the 80:10:10 split achieved the highest validation accuracy of 92.42% with minimal loss when compared to other ratios like 70:15:15.

The model is trained iteratively until it achieves a loss of less than 1% and an accuracy above 80%. Once these performance benchmarks are met, the model proceeds to testing and evaluation to assess its ability to accurately classify and differentiate between various corn leaf diseases.

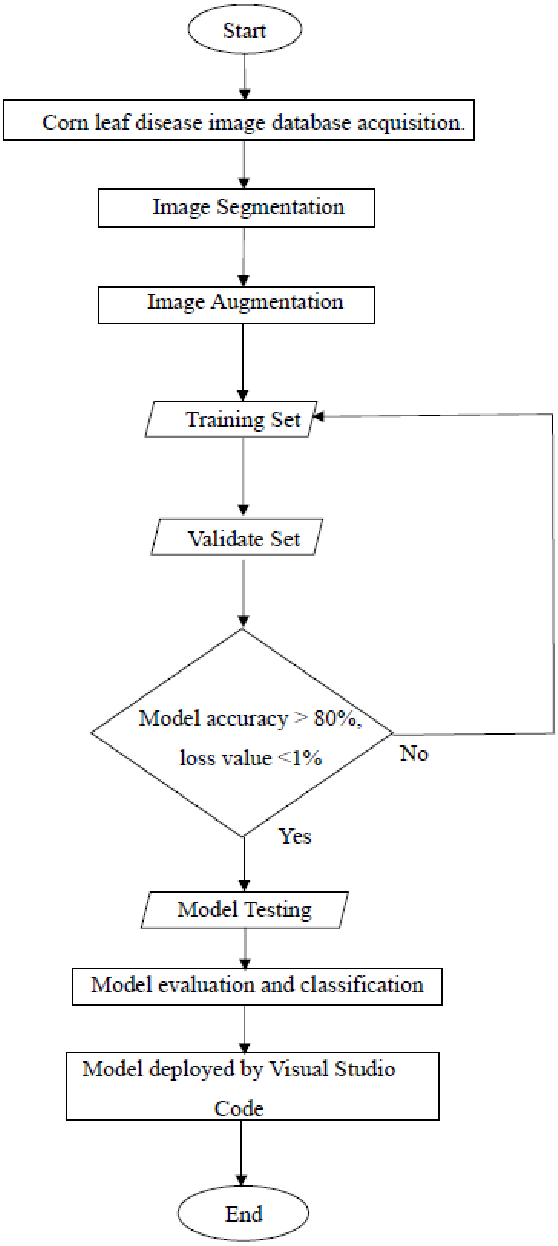


Fig. 1 *Flowchart of the project*

The initial phase in any image processing project is image acquisition, which provides the foundational data needed to run the program. In this study, the images used were of corn leaves categorized into three classes: healthy, infected with Northern Leaf Blight, and infected with Common Rust. These images were obtained from Kaggle, consisting of 3,000 original samples. All images were annotated using the Roboflow platform under the Corn Leaf Disease Project. To enhance model robustness, preprocessing steps such as auto-orientation and resizing were applied. This was followed by augmentation techniques including horizontal flipping and 90° rotations (both clockwise and counterclockwise), effectively expanding the dataset to 6,000 images. The final dataset was split using an 8:1:1 ratio into training, validation, and testing sets. Fig. 2 provides an overview of the dataset distribution, and Fig. 3 presents sample images for each class: Healthy, Northern Blight, and Common Rust.

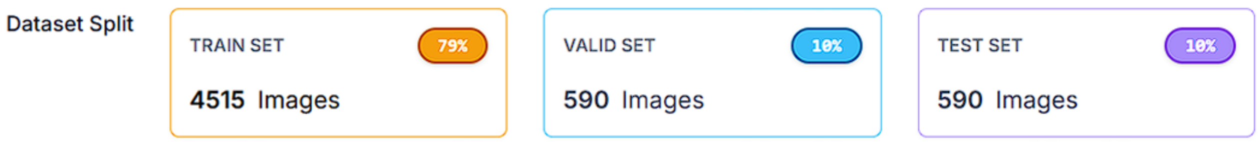


Fig. 2 *Dataset split*

(a) **(b)** **(c)**

Fig. 3 *Image sample of corn leaf with (a) Healthy; (b)Northern Blight; and (c) Common Rust*

Image annotation is a vital process that involves labeling images within a dataset to support the training of machine learning models, particularly in the realm of computer vision. This process is essential for supervised learning, where annotated data forms the foundation for teaching models to recognize specific features or objects. During annotation, various techniques such as bounding boxes, polygons, or segmentation masks are employed to assign labels to different elements within an image, clearly indicating their location and category. The precision and quality of these annotations significantly impact the model's performance, as they represent the ground truth from which the model learns to interpret visual information.

In this project dataset, Auto Orient and Resize were utilized. The Auto-Orient feature helps correct any inconsistencies in image orientation by removing EXIF data, ensuring that every image is displayed correctly as it is stored. Additionally, resizing the images to 640x640 pixels through stretching creates a uniform input size across the dataset, which can lead to better performance during model training and inference. These preprocessing steps are essential for achieving optimal results in image classification tasks, as they enhance both the uniformity and quality of the data being analyzed.

The data augmentation process for this project was carried out using Roboflow, a platform offering a suite of augmentation tools. Roboflow's Augmentations feature facilitated the application of selected techniques, including horizontal flipping and 90° rotations in both clockwise and counterclockwise directions. These augmentations were chosen to maximize the dataset, effectively tripling the number of outputs per original image while maintaining the dataset's integrity. By defining these specific augmentation strategies, the process ensured that the object detection model was exposed to a broader variety of image orientations, improving its robustness and adaptability. This targeted augmentation approach allowed for experimentation with the impact of these transformations on model performance, ensuring the enhanced dataset could contribute effectively to the training process.

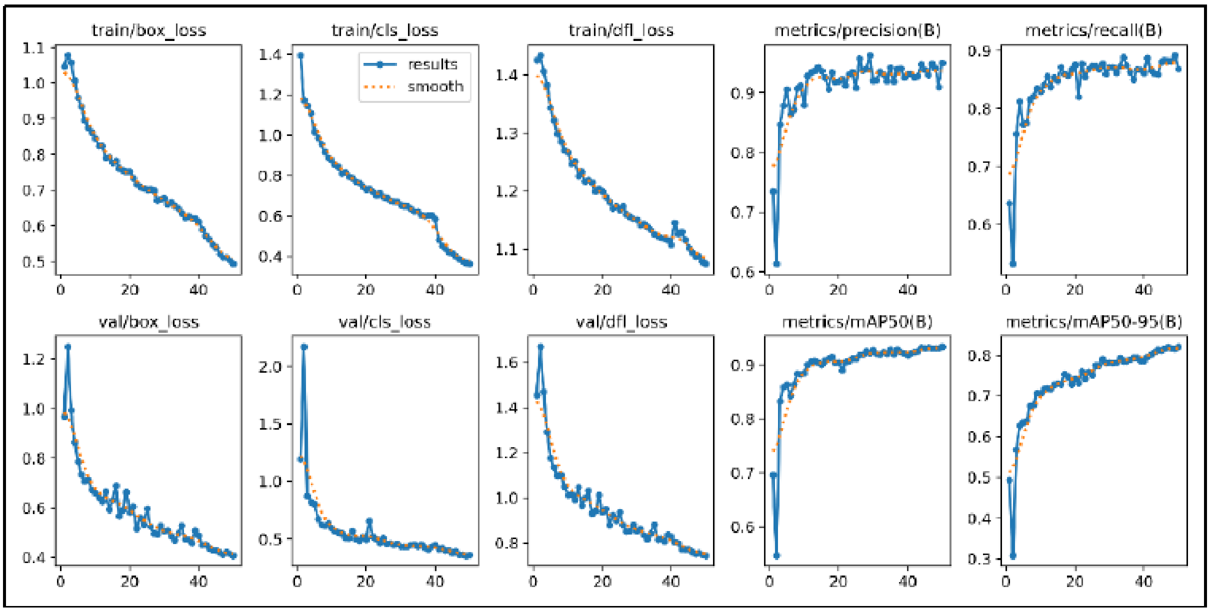
The Training and Testing process represents a critical phase of the image acquisition and object detection project, as it enables the construction and refinement of a model capable of effectively detecting and classifying objects within images. For this project, a comparative analysis will be conducted on several YOLO (You Only Look Once) model versions, specifically YOLOv8, YOLOv9, and YOLOv11. This analysis will be performed using Google Collab and will evaluate the models based on defined criteria, including accuracy, processing speed, and the complexity of the objects being detected. Each successive YOLO version introduces improvements in speed and accuracy, and this comparison will provide insights into the performance differences among these iterations.

4. Result

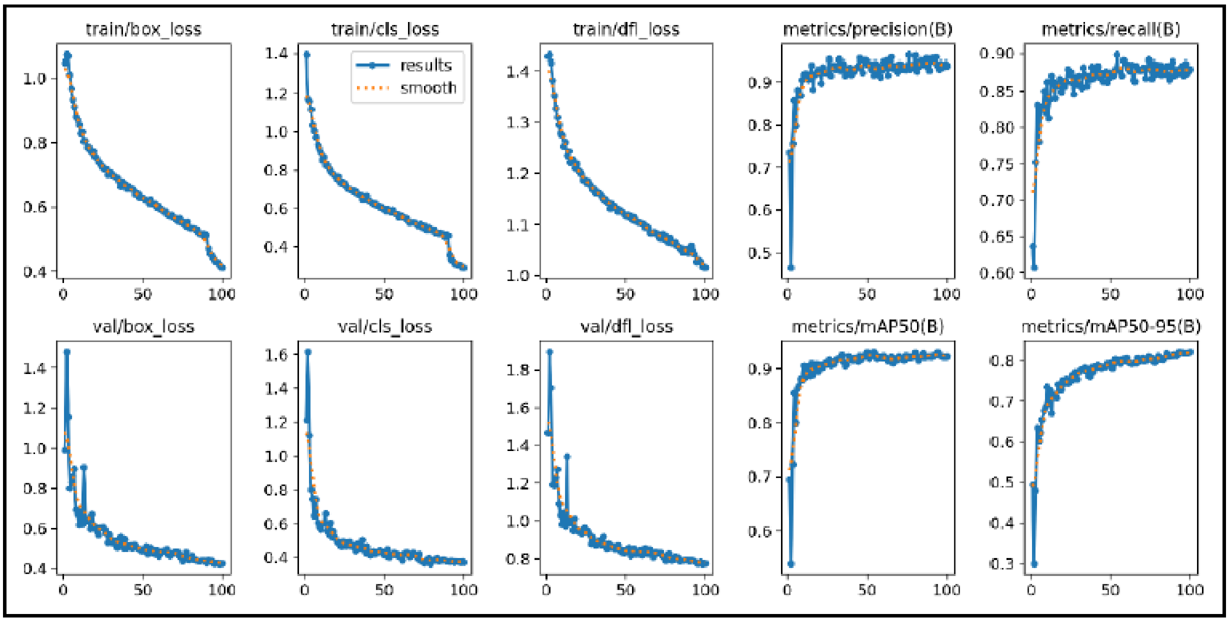
In this study, three different types of YOLO models were developed and compared based on their performance in disease recognition. The YOLO models that were developed are YOLOv8, YOLOv9 and YOLOV11. The dataset consisted of three different corn leaf conditions which are Healthy, Northern Blight and Common Rust. The YOLO models were trained using Google Colab. In the training phase of this study, the YOLO models were trained using 6000 datasets in 80:10:10 dataset ratio in terms of training, validation, and testing. All the training was done online using a personal laptop with AMD Ryzen 3 5300U with Radeon Graphics (2.60 GHz) and 8 GB RAM. The result is analyzed based on F1 Confidence (the maximum F1-score achieved at a specific confidence threshold), Precision Confidence (the precision value achieved at a particular confidence threshold), Precision-Recall Confidence (summarizing the overall performance of the model across all confidence thresholds) and Recall Confidence (recall value at a specific confidence threshold).

4.1 Performance Analysis of YOLOv8 Across Different Epoch Value

The training results for YOLOv8 at Fig. 4 and Table 1 give a quick understanding of how the model trained. After 50 epochs, the model gets a mAP@0.5 of about 0.93 and a mAP@0.5:0.95 around 0.82. Even though this offers a strong start, the lower scores at stricter maps imply that the localization of bounding boxes is not very accurate so far. The model's precision and recall are 0.87 and 0.86, which means it can find most of the objects but sometimes miss or include irrelevant ones. The completion of Epoch 100 shows major progress in all categories. mAP@0.5 goes up to 0.95, while the mAP@0.5:0.95 also increases to around 0.85, reflecting tighter and more reliable object localizations. Precision and recall are both over 0.9 which results in better detection and less identification of objects. In the last epoch, the performance of the model peaks and mAP@0.5 reaches a value of nearly 0.97, mAP@0.5:0.95 gets close to 0.89 and precision as well as recall are close to 0.92 and 0.94. Like the other scores, the F1 value increases consistently as time goes by. Eventually such training results in a model that is accurate and can be put into action.

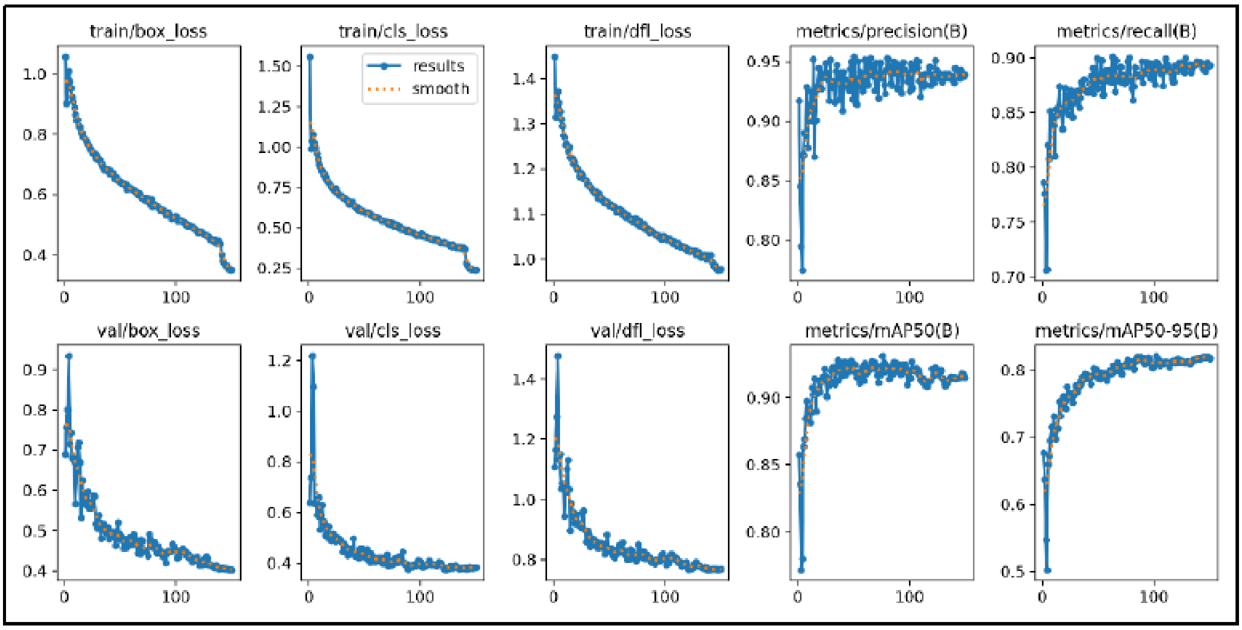


(a)



(b)

Fig. 4 *Training results for YOLOv8 (a) 50 Epoch; (b) 100 Epoch; (c) 150 Epoch*



(c)

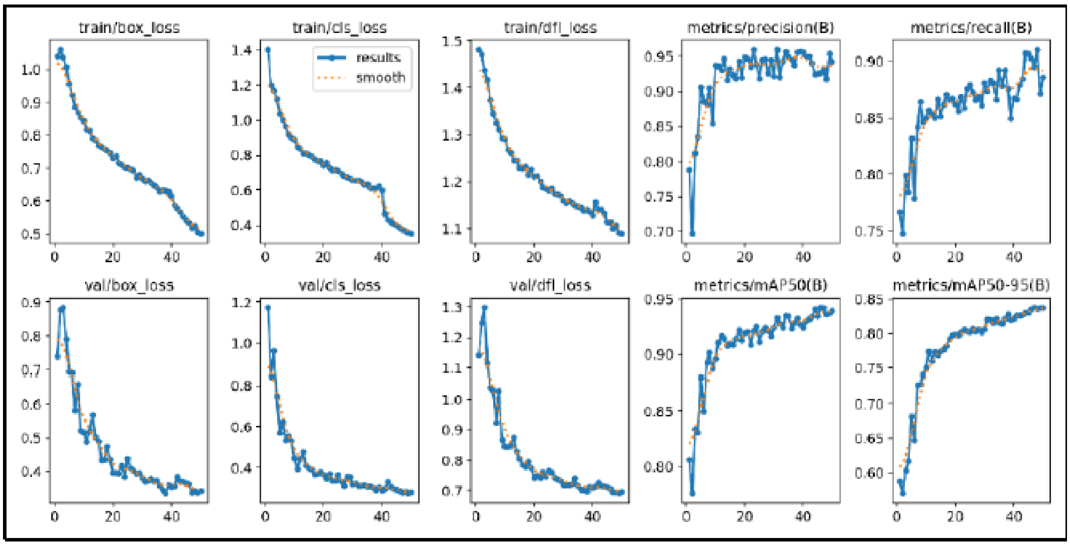
Fig. 4 *Training results for YOLOv8 (a) 50 Epoch; (b) 100 Epoch; (c) 150 Epoch (continued)*

Table 1 *Performance comparison of YOLOv8 model at different Epochs*

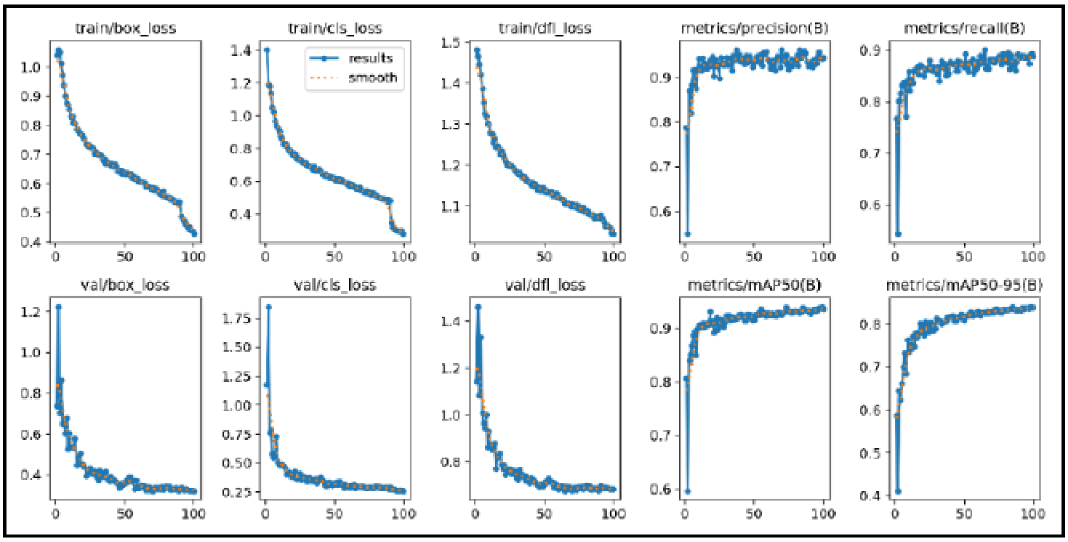
Epoch	Types	F1 Confidence	P Confidence	PR Confidence	R Confidence	Training Time
50	All	0.86	0.87	0.86	0.86	1.215 hrs
	Healthy	0.96	0.99	0.98	0.96	
	Northern	0.75	0.92	0.84	0.74	
	Rust	0.84	0.94	0.89	0.84	
100	All	0.9	0.91	0.9	0.89	2.560 hrs
	Healthy	0.97	0.99	0.99	0.97	
	Northern	0.81	0.94	0.88	0.8	
	Rust	0.89	0.96	0.93	0.88	
150	All	0.93	0.94	0.93	0.92	3.753 hrs
	Healthy	0.98	0.99	0.99	0.98	
	Northern	0.87	0.96	0.91	0.86	
	Rust	0.92	0.97	0.94	0.91	

4.2 Performance Analysis of YOLOv9 Across Different Epoch Values

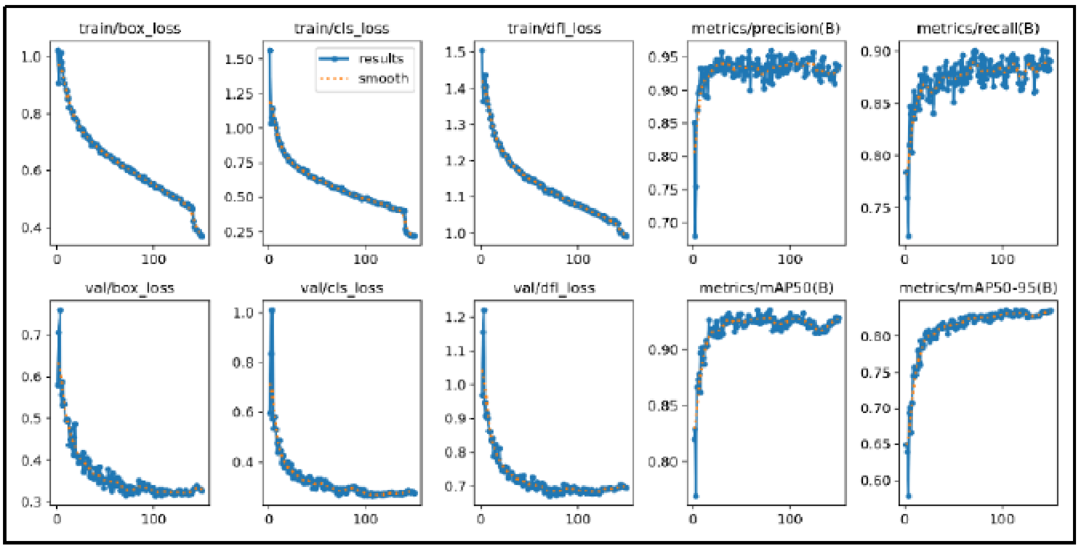
By referring to Training Results for YOLOv9 in Fig. 5 and Table 2, details are explained with respect to model loss, accuracy, and generalization. The difference is that when Epoch reaches 50, the training loss still stays relatively high, which indicates that the model is learning fundamental details and object formations. Generalization performance can be erratic, due to overfitting early patterns or too noisy representation in the training data. At Epoch 100, training is stabilized. The loss starts to decrease consistently and the accuracy levels off, implying efficient learning and better generalization. At Epoch 150, the model works best with minimum loss value, steady learning curve and high accuracy, which implies it is very adapted to the dataset. This stage verifies that the interior weights of YOLOv9 have certainly been fine-tuned, and its object detection can be efficiently supported by high confidence, accuracy, and speed.



(a)



(b)



(c)

Fig. 5 *Training results for YOLOv9 (a) 50 Epoch; (b) 100 Epoch; (c) 150 Epoch*

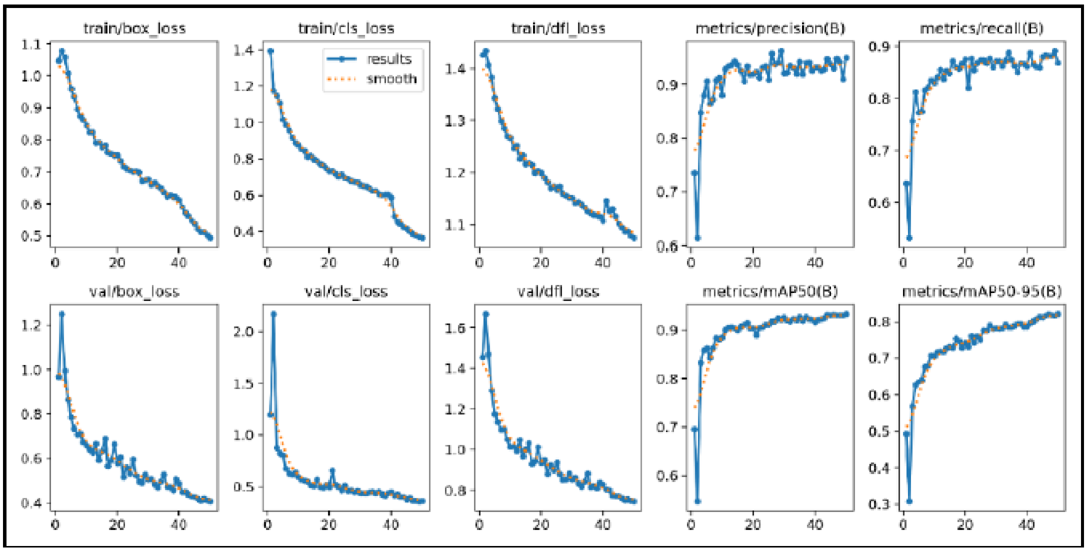
Table 2 *Performance comparison of YOLOv9 model at different Epochs*

Epoch	Types	F1 Confidence	P Confidence	PR Confidence	R Confidence	Training time
50	All	0.91	0.93	0.94	0.90	1.813 hours
	Rust	0.914	0.92	0.94	0.91	
	Northern	0.855	0.89	0.90	0.82	
	Healthy	0.97	0.98	0.98	0.96	
100	All	0.91	0.95	0.94	0.89	3.278 hours
	Rust	0.939	0.96	0.96	0.92	
	Northern	0.838	0.89	0.88	0.79	
	Healthy	0.97	0.98	0.98	0.96	
150	All	0.91	0.95	0.93	0.89	4.789 hours
	Rust	0.93	0.95	0.94	0.91	
	Northern	0.838	0.90	0.88	0.79	
	Healthy	0.97	0.99	0.99	0.96	

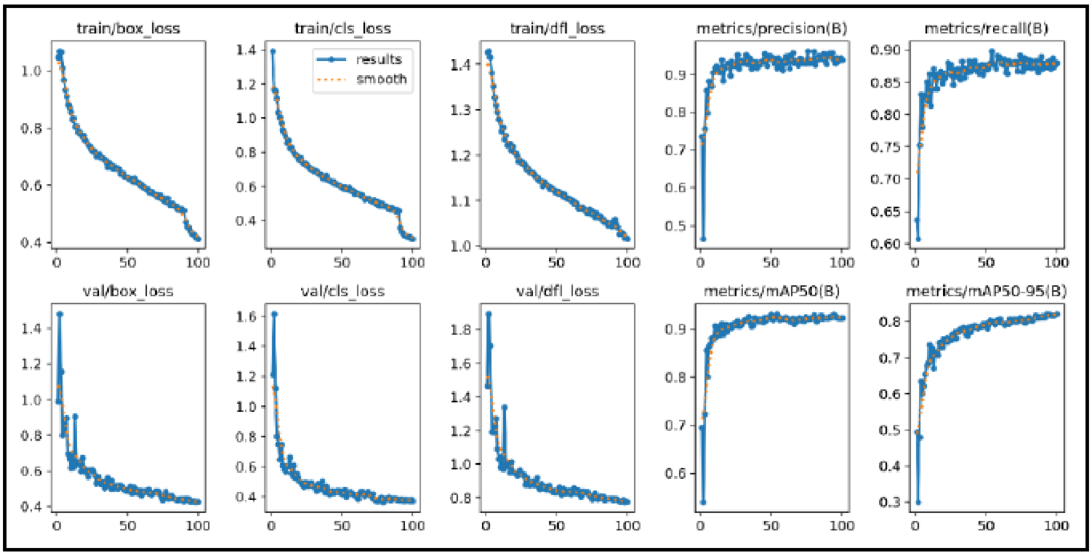
4.3 Performance Analysis of YOLOv11 Across Different Epoch Values

By referring to Fig. 6 and Table 3, the training results for YOLOv11 shown main performance measures which are training loss, validation loss, mean Average Precision (mAP), and when applicable, accuracy or Intersection over Union (IoU). The training loss at Epoch 50 might be high relatively and a significant difference between training loss and validation loss might advise possible overfitting or underfitting. The model will have improved greatly by Epoch 100, as both the loss values are predicted to have come down. At the same time, an increase in mAP values indicates an improvement in the detection quality, since the model learns to produce more accurate bounding boxes on different object classes, which is a required objective in object detection missions.

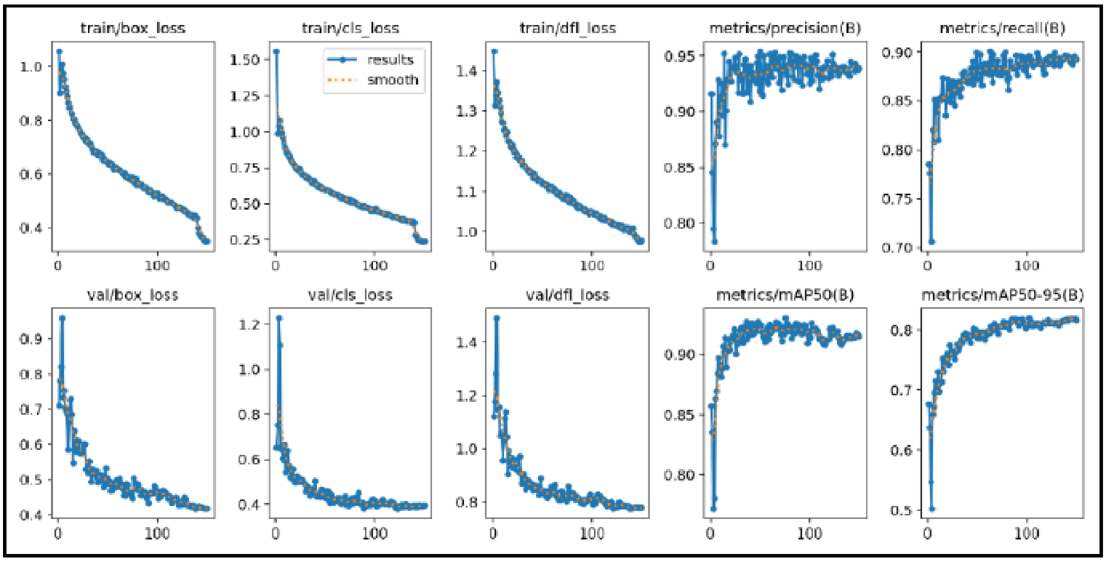
At Epoch 150, the loss curves have flattened, which is generally an indication of convergence but since there is still a gap between training and validation loss, it could be an indication of a mild instance of overfitting. The desirable result during this phase is a stable and consistent mAP curve that would indicate the model is approaching optimal performance and is well-calibrated. These evaluation criteria will play a critical role in deciding whether the model is good to be deployed to the real world or it needs additional work.



(a)



(b)



(c)

Fig. 6 *Training results for YOLOv11 (a) 50 Epoch; (b) 100 Epoch; (c) 150 Epoch*

Table 3 *Performance comparison of YOLOv9 model at different Epochs*

Epoch	Types	F1 Confidence	P Confidence	PR Confidence	R Confidence	Training time
50	All	0.908	0.950	0.950	0.869	1.237 hours
	Rust	0.923	0.938	0.938	0.908	
	Northern	0.821	0.923	0.923	0.739	
	Healthy	0.975	0.990	0.990	0.960	
100	All	0.914	0.956	0.956	0.875	2.483 hours
	Rust	0.939	0.973	0.973	0.908	
	Northern	0.833	0.909	0.909	0.768	
	Healthy	0.968	0.988	0.988	0.949	
150	All	0.920	0.945	0.945	0.896	3.687 hours
	Rust	0.947	0.963	0.963	0.931	
	Northern	0.845	0.896	0.896	0.800	
	Healthy	0.967	0.977	0.977	0.957	

J. of Electronic Voltage and Application Vol. 6 No. 2 (2025) p. 126-139

4.4 Comparison in Performance Between the YOLO Models

Table 4 shows a comparison of three YOLO model versions, YOLOv8, YOLOv9, and YOLOv11, for detecting different Corn Leaf Diseases, namely Rust, Northern Leaf Blight, and Healthy leaves. Metrics F1 Score has been used to evaluate each class, the overall F1 Score and maximum recall (across classes), training time in 150 epochs.

According to Table 4, YOLOv8 outperforms other methods and obtains the highest F1 Score (0.93) among all the disease categories, which is the most balanced in disease categories. It also tops the table for detecting Northern Leaf Blight with 0.87 and has the highest Max Recall (0.98), which therefore makes it the most sensitive for identifying Healthy leaves correctly. On the other hand, the YOLOv11 has a slightly lower overall F1 Score compared to YOLOv8 (-1.1 %) while it slightly increases the Rust Disease detection (0.95). Further, YOLOv11 has the lowest training time cost with only 3.69 hours, reducing about 1% when compared to the 3.75 hours for YOLOv8, a small but significant efficiency optimization.

Table 5 shows a comparison of three YOLO model versions, YOLOv8, YOLOv9, and YOLOv11, for detecting different Corn Leaf Diseases, namely Rust, Northern Leaf Blight, and Healthy leaves. Metrics F1 Score has been used to evaluate each class, the overall F1 Score and maximum recall (across classes), training time in 150 epochs.

According to table 4.5, YOLOv8 outperforms other methods and obtains the highest F1 Score (0.93) among all the disease categories, which is the most balanced in disease categories. It also tops the table for detecting Northern Leaf Blight with 0.87 and has the highest Max Recall (0.98), which therefore makes it the most sensitive for identifying Healthy leaves correctly. On the other hand, the YOLOv11 has a slightly lower overall F1 Score compared to YOLOv8 (-1.1 %) while it slightly increases the Rust Disease detection (0.95). Further, YOLOv11 has the lowest training time cost with only 3.69 hours, reducing about 1% when compared to the 3.75 hours for YOLOv8, a small but significant efficiency optimization.

As can be seen in Table 5 when viewing the mAP (overall), YOLOv9 performed worse. It reports the lowest F1 Score in all but one category (Rust, where it is 0.01 better than YOLOv8). But it is weaker in F1 Score for both Northern Leaf Blight (0.84) and Overall (0.91). To make things worse, YOLOv9, for a totally unknown reason, also gets the longest training period of 4.79 hours, which is significantly larger (+27.7%) than YOLOv11 or (+27.7%) higher than that of YOLOv8. The lower accuracy and longer training time of the YOLOv9 model implies that it may not be suitable to be the certain type of data or task data, and the performance of the model is decreased.

Finally, the YOLOv11 achieves the best trade-off between the accuracy, class balance, and the computational efficiency. Its ability to recognise all types of leaves and particularly the Rust Disease class that is hard to recognise, and its ability to learn perfectly and fast, is a solid reason for it to be implemented in a mobile or web-based leaf illness detection software.

Table 4 *Performance comparison of YOLOv8, YOLOv9 and YOLOv11 model at different Epochs*

Epoch	Model	Type	F1	Precision	PR	Recall	Time
50	YOLOv8	All	0.86	0.87	0.86	0.86	1.215 hours
		Healthy	0.96	0.99	0.98	0.96	
		Northern	0.75	0.92	0.84	0.74	
		Rust	0.84	0.94	0.89	0.84	
	YOLOv9	All	0.91	0.93	0.94	0.9	1.813 hours
		Healthy	0.97	0.98	0.98	0.96	
		Northern	0.855	0.89	0.9	0.82	
		Rust	0.914	0.92	0.94	0.91	
	YOLOv11	All	0.908	0.95	0.95	0.869	1.237 hours
		Healthy	0.975	0.99	0.99	0.96	
		Northern	0.821	0.923	0.923	0.739	
		Rust	0.923	0.938	0.938	0.908	
100	YOLOv8	All	0.93	0.94	0.93	0.92	3.753 hours
		Healthy	0.98	0.99	0.99	0.98	
		Northern	0.87	0.96	0.91	0.86	
		Rust	0.92	0.97	0.94	0.91	
	YOLOv9	All	0.91	0.95	0.93	0.89	4.789 hours
		Healthy	0.97	0.99	0.99	0.96	
		Northern	0.838	0.9	0.88	0.79	
		Rust	0.93	0.95	0.94	0.91	
	YOLOv11	All	0.92	0.945	0.945	0.896	3.687 hours
		Healthy	0.967	0.977	0.977	0.957	
		Northern	0.845	0.896	0.896	0.8	
		Rust	0.947	0.963	0.963	0.931	
150	YOLOv8	All	0.93	0.94	0.93	0.92	3.753 hours
		Healthy	0.98	0.99	0.99	0.98	
		Northern	0.87	0.96	0.91	0.86	
		Rust	0.92	0.97	0.94	0.91	
	YOLOv9	All	0.91	0.95	0.93	0.89	4.789 hours
		Healthy	0.97	0.99	0.99	0.96	
		Northern	0.838	0.9	0.88	0.79	
		Rust	0.93	0.95	0.94	0.91	
	YOLOv11	All	0.92	0.945	0.945	0.896	3.687 hours
		Healthy	0.967	0.977	0.977	0.957	
		Northern	0.845	0.896	0.896	0.8	
		Rust	0.947	0.963	0.963	0.931	

Table 5 *YOLO model comparison based on F1 score, recall, and training time*

Model	F1 Score (All)	F1 Score (Rust)	F1 Score (Northern)	F1 Score (Healthy)	Max Recall	Training Time (150 Epochs)
YOLOv8	0.93	0.92	0.87	0.98	0.98 (Healthy)	3.75 hrs
YOLOv9	0.91	0.93	0.84	0.97	0.96 (Healthy)	4.79 hrs
YOLOv11	0.92	0.95	0.85	0.97	0.96 (Healthy)	3.69 hrs

J. of Electronic Voltage and Application Vol. 6 No. 2 (2025) p. 126-139

5. Conclusion

To sum up, this project has conducted a comparative investigation of YOLOv8, YOLOv9 and YOLOv11 models used to identify disease in corn leaves, concerning their accuracy, precision, recall, F1 score and training time. The models had their training and evaluations done in the same conditions to have a fair comparison. Although YOLOv8 achieved the best overall F1 score of 0.93 and the maximum recall of 0.98, YOLOv11 posted the most stable overall results on all classes and measures. Markedly, out of all the models, YOLOv11 was most successful in identifying Common Rust with a precision of 0.963 and F1 score of 0.947, yet the model was the least time-exhaustive taking not even 24 hours, or to be more precise 3.687 hours to train.

Although reaching a high precision rate (0.945) and demonstrating satisfactory results in the process of Rust Disease detection, YOLOv9 was not as consistent when it came to overall class balance and needed much more time to train. As compared with YOLOv8 and YOLOv9, the technique provided by YOLOv11 was more efficient and stable, having a high detection rate at a less computational expense. That is why, in terms of acquiring the best trade-off model among the three in corn leaf disease detection, it is possible to regard YOLOv11 as the best trade-off model in the case when both the accuracy of detection and the efficiency of the training are needed.

Acknowledgement

This research was supported by Universiti Tun Hussein Onn Malaysia (UTHM) through Teir 1 (Q361).

Conflict of Interest

Authors declare that there is no conflict of interest regarding the publication of the paper.

Author Contribution

The author, Shafieka Nadia Binti Sabarudin, affirms only taking responsibility of the following: ***conception and design of the study, collection of the dataset, annotation of the images, training and evaluation of the models, development of the application, analysis and interpretation of the results, as well as preparation of the manuscript****. The supervisor, Dr. Nik Shahidah Afifi Md Taujuddin, Dr Suhaila Sari and Ahmad Raqib Ab Ghani have* ***offered academic advice and suggested appropriate literature*** *as well as useful comments to enhance the system and research direction.*

References

[1] X. Zhang, Y. Qiao, F. Meng, C. Fan, and M. Zhang, "Identification of maize leaf diseases using improved deep convolutional neural networks," *IEEE Access*, vol. 6, pp. 30370–30377, 2018, doi: 10.1109/ACCESS.2018.2844405.

[2] H. Amin, A. Darwish, A. E. Hassanien, and M. Soliman, "End-to-End Deep Learning Model for Corn Leaf Disease Classification," *IEEE Access*, vol. 10, pp. 31103–31115, 2022, doi: 10.1109/ACCESS.2022.3159678.

[3] D. A. Noola and B. Dayanand R, "Computer Aided Corn Leaf Disease Identification System," in *Proceedings - 2nd International Conference on Smart Electronics and Communication, ICOSEC 2021*, Institute of Electrical and Electronics Engineers Inc., 2021, pp. 930–937. doi: 10.1109/ICOSEC51865.2021.9591863.

[4] P. Bachhal, V. Kukreja, and S. Ahuja, "Maize Disease classification using Deep Learning Techniques: A Review," in *2023 International Conference on Advancement in Computation and Computer Technologies, InCACCT 2023*, Institute of Electrical and Electronics Engineers Inc., 2023, pp. 259–264. doi: 10.1109/InCACCT57535.2023.10141847.

[5] K. K. Villanueva, J. A. M. Galindo, A. J. R. Tamayo, J. E. C. Rosal, and D. I. E. Hisola, "Development of a Computer Vision Application for Mango (Mangifera Indica L.) Fruit Defect Detection Using YOLOv8 Architecture," in *2024 IEEE International Conference on Artificial Intelligence in Engineering and Technology (IICAIET)*, IEEE, Aug. 2024, pp. 483–487. doi: 10.1109/IICAIET62352.2024.10730444.

[6] S. H. Samba, I. Y. B. Agranata, L. Tsanaullailla, and F. Hamami, "Traffic Accident Detection Analysis Using YOLOv9 Algorithm," in *2024 9th International Conference on Informatics and Computing, ICIC 2024*, Institute of Electrical and Electronics Engineers Inc., 2024. doi: 10.1109/ICIC64337.2024.10957461.

[7] T. Soyak, N. B. Özcan, and G. Çinarer, "Strawberry Ripeness Detection: Optimizing Harvesting Efficiency with YOLOv11," in *ICHORA 2025 - 2025 7th International Congress on Human-Computer Interaction, Optimization and Robotic Applications, Proceedings*, Institute of Electrical and Electronics Engineers Inc., 2025. doi: 10.1109/ICHORA65333.2025.11017268.

[8] A. K. S. and H. Das Kshyanaprava Panda Panigrahi, *A CNN Approach for Corn Leaves Disease Detection to support Digital Agricultural System*. IEEE, 2020.

[9] S. Santhi, M. Mahima, R. Vengatesan, and M. Gautham Raja, "Corn Leaf Disease Detection using YOLOv5," in *7th International Conference on Inventive Computation Technologies, ICICT 2024*, Institute of Electrical and Electronics Engineers Inc., 2024, pp. 1335–1338. doi: 10.1109/ICICT60155.2024.10544783.

[10] A. Baliyan, V. Kukreja, V. Salonki, and K. S. Kaswan, "Detection of Corn Gray Leaf Spot Severity Levels using Deep Learning Approach," in *2021 9th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions), ICRITO 2021*, Institute of Electrical and Electronics Engineers Inc., 2021. doi: 10.1109/ICRITO51393.2021.9596540.

[11] B. Song and J. Lee, "Detection of Northern Corn Leaf Blight Disease in Real Environment Using Optimized YOLOv3," in *2022 IEEE 12th Annual Computing and Communication Workshop and Conference, CCWC 2022*, Institute of Electrical and Electronics Engineers Inc., 2022, pp. 475–480. doi: 10.1109/CCWC54503.2022.9720782.

[12] Z. Wu, Y. Jiang, X. Li, and K. L. Chung, "Enhancing Precision Agriculture: YOLOv8 for Accurate Corn Disease and Pest Detection," in *2024 IEEE 7th International Conference on Electronic Information and Communication Technology, ICEICT 2024*, Institute of Electrical and Electronics Engineers Inc., 2024, pp. 980–984. doi: 10.1109/ICEICT61637.2024.10671002.

[13] Y. C. Austria, M. C. A. Mirabueno, D. J. D. Lopez, D. J. L. Cuaresma, J. R. MacAlisang, and C. D. Casuat, "EZM-AI: A Yolov5 Machine Vision Inference Approach of the Philippine Corn Leaf Diseases Detection System," in *4th IEEE International Conference on Artificial Intelligence in Engineering and Technology, IICAIET 2022*, Institute of Electrical and Electronics Engineers Inc., 2022. doi: 10.1109/IICAIET55139.2022.9936848.

[14] O. Nathanael, M. Hutagalung, and Y. Gamaliel, "Development of Convolutional Neural Network (CNN) Method for Classification of Tomato Leaf Disease Based on Android," in *2023 1st IEEE International Conference on Smart Technology: Advances in Smart Technology for Sustainable Well-Being, ICE-SMARTec 2023*, Institute of Electrical and Electronics Engineers Inc., 2023, pp. 60–65. doi: 10.1109/ICE-SMARTECH59237.2023.10461962.