Vol. 8, No. 2, April 2026

Journal of Information Systems and Informatics

Published By Asosiasi Doktor Sistem Informasi Indonesia

Enhancing YOLOv12-Based Detection Quality for Rice Leaf Diseases through Systematic Evaluation

Ida Mulyadi[1], Fahrim Irhamna[2], Chyquitha Danuputri[3], Ridwang[4], Ridha Awalia[5]

[1,2,3,4,5]Informatics Department, Faculty of Engineer, Muhammadiyah University of Makassar, Makassar, Indonesia

Received:
December 26, 2025
Revised:
March 10, 2026
Accepted:
April 1, 2026
Published:
April 12, 2026

Corresponding Author:
Author Name*:
Ida Mulyadi
Email*:
idamulyadi@unismuh.ac.id

DOI:
10.63158/journalisi.v8i2.1580

Abstract. One of the most significant staple crops in the world is rice, and one of the main causes of the drop in agricultural yields is illnesses that affect rice leaves. To avoid large agricultural losses, early diagnosis of these illnesses is essential. The goal of this project is to use YOLOv12, the most recent deep learning-based object detection architecture, to create a rice leaf disease detection system. The model was trained using a dataset of 4,744 photos of rice leaves that included three disease classes: Leaf Blast, Brown Spot, and Bacterial Leaf Blight. Methods to boost variability and enhance detection performance, image preprocessing with data augmentation was used. Standard object detection criteria, such as mean Average Precision (mAP), precision, and recall, were used to assess the model. The YOLOv12 model was highly effective in detecting rice leaf illnesses. According to the experimental data, it achieved a mAP of 97%, a precision of 96%, and a recall of 96.5%. The use of YOLOv12's greater efficiency and quality in detecting small objects—which is essential for identifying illness symptoms on leaves—is what makes this study successful. These results lay the groundwork for upcoming precision agricultural real-time monitoring applications.

Keywords: Plant disease monitoring, YOLOv12, object detection, precision agriculture, rice leaf disease detection

ISI Journal of Information Systems and Informatics
Published By Asosiasi Doktor Sistem Informasi Indonesia

1. INTRODUCTION

Rice is one of the most important staple crops in the world and serves as the primary food source for more than half of the global population. Its contribution to food availability, household nutrition, and rural economic stability makes rice production a crucial component of global food security. In many developing countries, rice is not only consumed daily but also plays a central role in agricultural livelihoods and national economic resilience. As a result, maintaining both the yield and quality of rice has become a major priority in agricultural research and production systems. However, despite its global significance, rice cultivation continues to face serious challenges from various biotic stresses, particularly leaf diseases that can substantially reduce crop productivity and threaten sustainable food supply.

Among the most damaging diseases affecting rice plants are Brown Spot, Blast, and Bacterial Leaf Blight. These diseases attack the leaf surface and tissues, reducing photosynthetic efficiency, weakening plant growth, and ultimately decreasing grain yield and quality. Their impact can become even more severe when infection occurs at an early growth stage or spreads rapidly under favorable environmental conditions. Because disease symptoms often begin as small lesions or subtle discoloration, early detection is essential to prevent wider infection and minimize economic losses. In practice, timely identification allows farmers and agricultural practitioners to apply control measures more precisely, thereby reducing unnecessary pesticide use and improving field management efficiency.

Traditionally, rice leaf diseases are identified through direct visual inspection by farmers, agricultural officers, or plant pathology experts. Although this method remains widely used, it is often labor-intensive, time-consuming, and highly dependent on the observer's knowledge and field experience. Manual inspection also becomes less practical for large-scale agricultural areas, where frequent monitoring is required to detect disease outbreaks before they become severe. In addition, similar visual symptoms between diseases, nutrient deficiencies, and environmental stress can lead to misclassification and inconsistent diagnosis. To address these limitations, recent studies have increasingly explored deep learning methods, particularly convolutional neural networks (CNNs), for automatic plant disease detection because of their strong ability to extract visual features directly from image data with high accuracy [1], [2], [3].

The application of CNN-based approaches has shown promising results in agricultural image analysis, especially for recognizing disease patterns that are difficult to detect consistently through conventional observation. These models are capable of learning complex image characteristics such as color variation, texture irregularity, and lesion shape, which makes them highly effective for identifying diseased plant tissues under different conditions [4], [5], [6]. Building on these advances, object detection frameworks have gained attention because they not only classify disease categories but also localize infected regions within images. According to [7], architectures such as YOLO (You Only Look Once) have demonstrated effectiveness for real-time detection of multiple disease instances in a single image. Earlier YOLO variants, including YOLOv3, YOLOv4, YOLOv5, YOLOv7, and YOLOv8, have progressively improved detection accuracy and processing speed, making them increasingly attractive for agricultural monitoring applications [8]. However, these models still face important challenges, particularly in detecting very small objects such as early-stage disease spots on rice leaves, while often requiring considerable computational resources that limit their deployment in field environments and mobile platforms [9], [10].

Despite the progress made in deep learning-based plant disease detection, there remains a significant gap in optimizing object detection models for high-quality identification of small rice leaf disease objects without sacrificing efficiency for real-world implementation [11]. Previous studies have not fully explored the potential of the latest YOLOv12 architecture, which is expected to offer better efficiency and stronger small-object detection capability than earlier versions. In addition, systematic evaluation under different dataset partitioning scenarios has not been sufficiently investigated, even though such analysis is important for assessing model robustness and generalization. Therefore, this study proposes a YOLOv12-based rice leaf disease detection system and evaluates its performance under three dataset partitioning scenarios using a dataset consisting of 4,744 images. This study is expected to contribute by improving the detection quality of minor disease lesions, providing empirical evidence regarding the effectiveness of YOLOv12 compared with previous YOLO models, and supporting the development of scalable real-time disease monitoring systems for precision agriculture.

ISI Journal of Information Systems and Informatics
Published By Asosiasi Doktor Sistem Informasi Indonesia

2. METHODS

3.1. Dataset Description

This study employed a total of 4,744 images of rice leaves representing three disease classes, namely Bacterial Leaf Blight, Brown Spot, and Leaf Blast. The dataset was designed to reflect actual field conditions so that the detection model could learn from image characteristics that commonly appear in real agricultural environments. A portion of the dataset, consisting of 1,000 images, was collected directly from rice fields in Jangan-jangan Village, Pujananting District, Barru Regency, South Sulawesi Province, Indonesia, using a digital camera [12]. These field images captured a range of naturally occurring variations, including differences in lighting intensity, camera distance, leaf orientation, background complexity, and image perspective. Incorporating such variations was important because these factors strongly influence detection quality and model generalization when the system is used outside controlled laboratory settings.

All images were manually annotated using Roboflow, where bounding boxes were assigned to visible disease regions in each image. This annotation process produced labeled object detection data that could be used to train the YOLOv12 model to recognize both the location and category of disease symptoms. To evaluate the robustness of the model under different data allocation settings, the dataset was divided into three experimental scenarios [13]. In Scenario A, the dataset was split into 80% training, 10% validation, and 10% testing, corresponding to 3,795 training images, 474 validation images, and 475 test images. In Scenario B, the dataset was divided into 70% training, 10% validation, and 20% testing, corresponding to 3,320 training images, 474 validation images, and 950 test images. In Scenario C, the dataset was split into 60% training, 20% validation, and 20% testing, corresponding to 2,846 training images, 949 validation images, and 949 test images. These different partitioning strategies were used to analyze the effect of data distribution on model repeatability, generalization, and detection performance. The test sets in each scenario were specifically used to calculate true positives (TP), false positives (FP), and false negatives (FN) during performance evaluation.

3.2. Preprocessing and Augmentation

To improve data quality and increase model robustness against real-world image variations, a series of preprocessing and augmentation techniques were applied before

ISI Journal of Information Systems and Informatics
Published By Asosiasi Doktor Sistem Informasi Indonesia

training [14], [15], [16]. The overall image enhancement pipeline is illustrated in Figure 1. In the preprocessing stage, all images were first resized to a uniform input resolution of 640 × 640 pixels in order to standardize spatial dimensions and ensure compatibility with the YOLOv12 architecture. After resizing, pixel values were normalized to reduce intensity variation and stabilize the training process. This standardization step is important because raw field images are often captured under inconsistent conditions, resulting in variations in brightness, scale, and visual contrast that may negatively affect model learning.

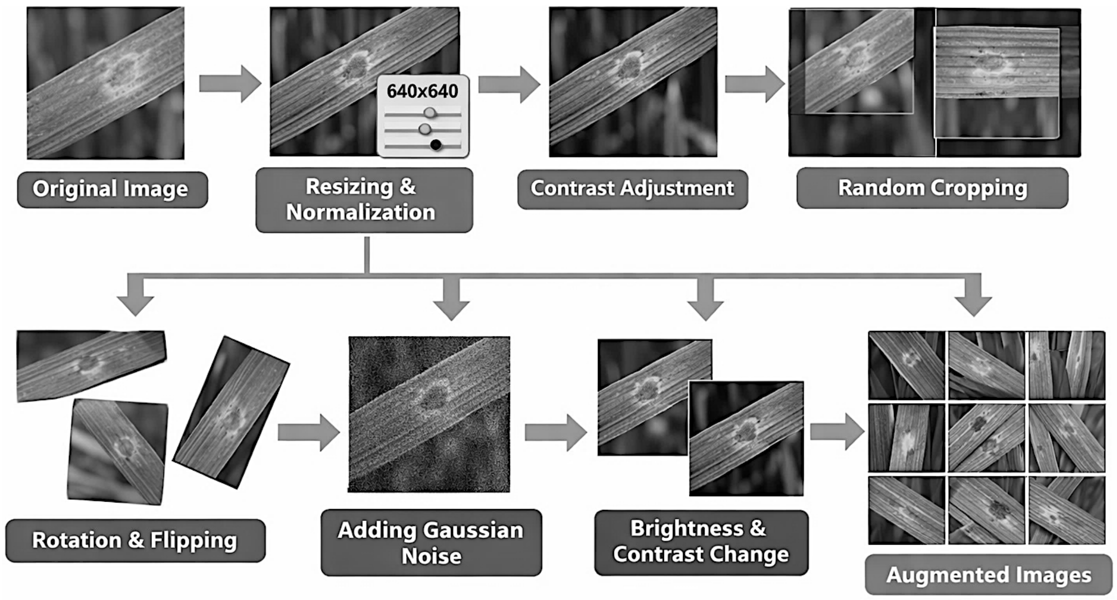

Figure 1. Preprocessing and Augmentation

Following preprocessing, several augmentation techniques were applied to increase the diversity of the training data and improve the model's ability to generalize to unseen field conditions. Random cropping was used to create variation in object framing and to simulate situations where leaves are only partially visible in the camera view. Contrast adjustment was applied to enhance the distinction between healthy and diseased tissue, especially when lesions exhibited weak visual contrast. Additional augmentation techniques included brightness and contrast changes to simulate different illumination conditions, rotation and flipping to introduce orientation diversity, and Gaussian noise injection to improve tolerance to image disturbance and sensor-related artifacts. These transformations were chosen because images captured in agricultural environments are frequently affected by changing sunlight intensity, camera movement, background

ISI Journal of Information Systems and Informatics
Published By Asosiasi Doktor Sistem Informasi Indonesia

clutter, and varying distances between the camera and the target leaf. By exposing the model to these variations during training, the risk of overfitting was reduced, and the resulting detector became more reliable for practical field deployment.

3.3. Training of YOLOv12 Model

The rice leaf disease detection model was developed using the YOLOv12 object detection architecture. The training process followed a transfer learning strategy by initializing the network with pre-trained YOLOv12 weights, which helped accelerate convergence and improve feature extraction performance [17], [18], [19]. This approach is particularly useful in agricultural image analysis, where disease symptoms may appear as small, irregular, and visually complex lesions. By starting from pre-trained weights, the model could leverage previously learned low-level and mid-level visual representations, such as edges, textures, and object contours, before adapting them to the specific task of detecting rice leaf diseases.

The training workflow consisted of four main stages. First, the annotated dataset was organized into training, validation, and test subsets according to the three experimental scenarios described in Section 2.1. Second, model initialization was performed using pre-trained YOLOv12 weights. Third, iterative learning was carried out using the training set, where the model learned to classify disease categories and localize their corresponding bounding boxes simultaneously. During this stage, key hyperparameters such as epoch number, batch size, input image size, and patience were configured to achieve stable optimization and to avoid overfitting. The validation set was used throughout training to monitor performance and identify the best model checkpoint. Finally, after the training process was completed, the best-performing model was exported as the final optimized object detection model for rice leaf disease recognition. This training design ensured that the resulting model was not only accurate on the training data but also sufficiently robust when evaluated on unseen images acquired under real field conditions.

3.4. Evaluation Metrics

The performance of the proposed model was assessed using standard metrics commonly employed in object detection studies [20], [21], [22], [23], [24]. The primary evaluation metrics used in this study were Precision, Recall, and mean Average Precision (mAP). These metrics were selected because they provide complementary information about

ISI Journal of Information Systems and Informatics
Published By Asosiasi Doktor Sistem Informasi Indonesia

the model's ability to produce correct detections, identify all relevant disease objects, and maintain strong localization quality across classes.

1) Precision measures how many predicted detections are actually correct and is defined as shown in Equation 1.

$$Precision = TP / (TP + FP) \quad (1)$$

where TP denotes true positives and FP denotes false positives. A high precision value indicates that the model produces relatively few incorrect detections.

2) Recall measures the ability of the model to detect all relevant disease objects present in the dataset and is defined as shown in Equation 2.

$$Recall = TP / (TP + FN) \quad (2)$$

where FN denotes false negatives. A high recall value indicates that the model misses only a small number of actual disease instances.

3) To provide a more comprehensive performance summary, mean Average Precision (mAP) was also used. This metric evaluates detection performance by considering the average precision for each class and then computing the mean value across all classes. It is expressed as shown in Equation 3.

$$mAP = (1 / C) \sum_{i=1}^{c} AP_i \quad (3)$$

where C is the total number of classes and AP_i is the average precision for class *i*. In object detection, mAP is particularly important because it reflects not only the correctness of class predictions but also the quality of object localization under Intersection over Union (IoU) criteria. In this study, mAP was used as one of the main indicators for comparing YOLOv12 performance across the three dataset partitioning scenarios.

ISI Journal of Information Systems and Informatics
Published By Asosiasi Doktor Sistem Informasi Indonesia

3.5. Comparative Evaluation Across Dataset Partitioning Scenarios

To investigate how data allocation influences model learning and generalization, a comparative evaluation was conducted across the three dataset partitioning scenarios. In this experiment, the same annotation strategy, preprocessing pipeline, augmentation settings, and YOLOv12 training configuration were maintained for all scenarios so that the comparison would remain fair and controlled. The only variable intentionally changed was the proportion of data assigned to the training, validation, and test sets. This design allowed the study to examine whether increasing the amount of training data would improve the model's ability to learn disease features more effectively, especially for small lesion objects that are often difficult to detect.

The comparative analysis focused on the resulting precision, recall, mAP, and the distribution of TP, FP, and FN values in each scenario. By using this approach, the study was able to evaluate not only overall detection accuracy but also the model's repeatability and robustness under different data availability conditions. A larger training set was expected to improve the model's representation learning and reduce prediction errors, while larger validation and test sets provided a more extensive basis for assessing generalization. This scenario-based evaluation therefore served as an important component of the experimental design, helping determine the most effective data split for rice leaf disease detection using YOLOv12. The detailed numerical results of this comparison are more appropriately presented in the Results and Discussion section.

3.6. Application Design

The workflow of the rice leaf disease detection application is illustrated in Figure 2. The process begins when the user launches the application on the target device. Once the application is active, the user captures an image of a rice leaf directly using the camera. This image serves as the main input for the detection system. After the image is captured, the trained YOLOv12 model performs real-time inference, allowing the application to analyze visual patterns associated with rice leaf diseases without requiring complex manual processing. This real-time capability is important for field deployment because it enables users to obtain rapid diagnostic feedback while inspecting crops directly in agricultural environments.

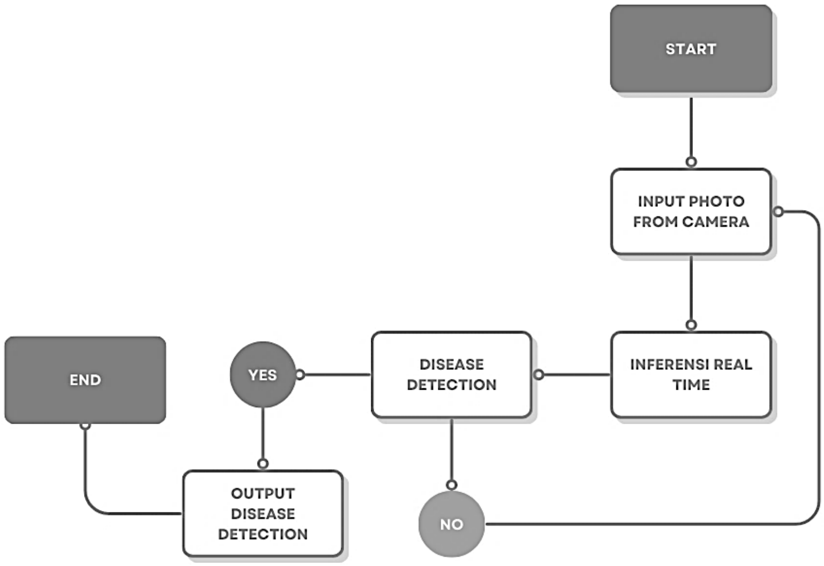

Figure 2. Application Design

After inference, the system evaluates the prediction results to determine whether disease symptoms have been successfully detected. If the model identifies a disease object with sufficient confidence, the application displays the output to the user, including the predicted disease class and the confidence score associated with the detection. The output may also include visual bounding boxes that highlight the infected area on the leaf image. If no disease is detected, or if the captured image is too unclear to support reliable analysis, the system returns a negative or inconclusive result and prompts the user to capture another image. The process ends after the user reviews the displayed result. This application design was developed to support practical and accessible disease diagnosis in the field, enabling farmers and agricultural practitioners to perform fast preliminary disease identification using a camera-based detection interface.

3. RESULTS AND DISCUSSION

3.1 Dataset

The 4,744 rice plant photos from the village of Jangan-jangan, Pujananting District, Barru Regency, make up the dataset used in this study. This dataset is then divided into three parts: training data, validation data, and testing data. After that, the pictures are uploaded to Roboflow for annotation. Three classifications are used in this study: brown spot,

ISI Journal of Information Systems and Informatics
Published By Asosiasi Doktor Sistem Informasi Indonesia

bacterial leaf blight, and leaf blast. The following are the findings from data gathering in the village of Jangan-jangan, Pujananting District, Barru Regency. The picture data of the rice plant leaves used in this investigation are shown in Figure 3. Bacterial leaf blight, brown spot, and leaf blast are the three types of leaf conditions.

Bacterial leaf blight | Brown spot | Leaf blast

Figure 3. Rice Plant Diseases

3.2 Preprocessing and Augmentation

High-quality, well-annotated leaf images are crucial for improving model accuracy in detecting Bacterial Leaf Blight. Figure 4 shows preprocessing and augmentation outcomes, with each leaf marked by distinct, non-overlapping bounding boxes indicating disease. These annotations reduce noise in training by capturing variations in leaf size and position. Accurate, consistent annotations and varied data enable YOLOv12 training to reliably identify Bacterial Leaf Blight under diverse conditions.

Figure 4. Annotation of Rice Leaves Infected with Bacterial Leaf Blight

The findings of the YOLOv12 model, which was trained on the annotated rice leaf dataset and is capable of accurately identifying brown spots, are displayed in Figure 5. The model shows that it can detect several lesions on a single leaf at once, as well as lesions of

ISI Journal of Information Systems and Informatics
Published By Asosiasi Doktor Sistem Informasi Indonesia

different sizes, locations, and orientations. These findings demonstrate that the model may be applied to automatic disease monitoring in the field, enabling quicker and more effective control actions, and that it can generalize to fresh leaves.

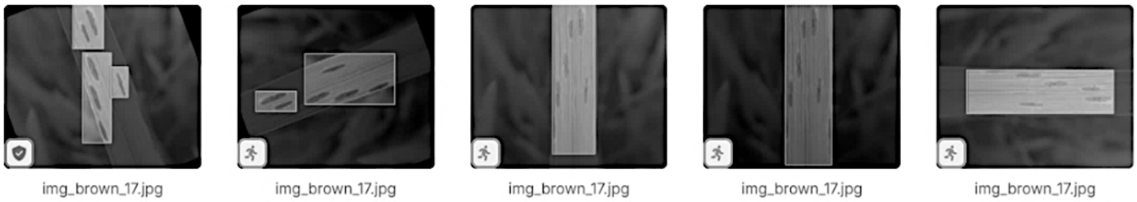

Figure 5. Annotation of Rice Leaves Infected with Brown Spot

The YOLOv12 model's detection results on augmented rice leaf pictures are displayed in Figure 6, indicating a strong capacity to detect Leaf Blast lesions. The model has a strong capacity for generalization, can identify several lesions on a single leaf, and maintains accuracy even when the leaves are positioned and oriented differently. This demonstrates that the model is more resilient to visual changes when data augmentation is used, making it suitable for autonomous disease monitoring in the field.

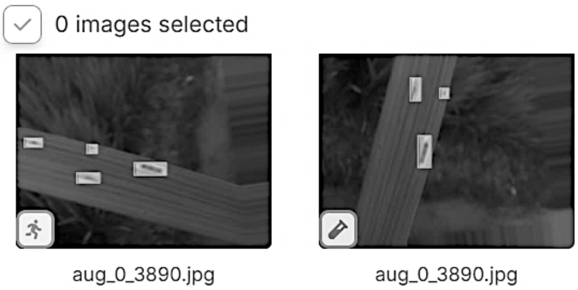

Figure 6. Annotation of Rice Leaves Infected with Leaf Blast

3.3 Training of YOLOv12 Model

The YOLOv12 model has been set up and is prepared for training with the rice leaf dataset, utilizing pretrained weights for rapid learning, as seen in Figure 7. The number of epochs, picture size, batch size, and early stopping are examples of hyperparameter settings that guarantee optimal training performance. The model may be able to correctly identify lesions on rice leaves because the convolutional layer's initial output shows that the network is prepared for the object detection optimization phase.

ISI Journal of Information Systems and Informatics
Published By Asosiasi Doktor Sistem Informasi Indonesia

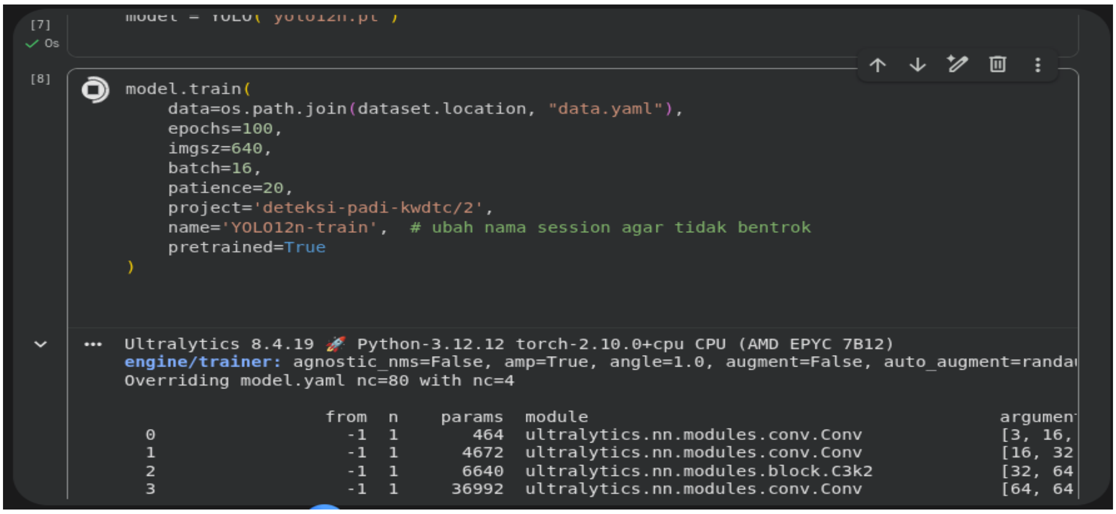

Figure 7. YOLOv12 Training

3.4 Evaluation Metrics

The evaluation results of the YOLOv12 model, which has been trained to identify items on rice leaves (dataset "Train 80%"), are displayed in Figure 8a. Model Performance: Recall: 96.5%, Precision: 96%, and mAP@50: 97%. The model's ability to correctly predict bounding boxes on the validation dataset is demonstrated by the high mAP value. High precision means that there are few false positives and the model's predictions are accurate. A high recall means that there are few false negatives because the model can identify most of the lesions that are currently present. The evaluation's findings verify that a higher percentage of training data aids YOLOv12 in identifying lesions on rice leaves with different sizes, shapes, and orientations. The model is prepared for automatic rice leaf disease monitoring. A thorough picture of the balance between the model's capacity to identify current lesions and prevent incorrect predictions is provided by the assessment of mAP, precision, and recall.

Figure 8a. Metric Evaluation Scenario 1

ISI Journal of Information Systems and Informatics
Published By Asosiasi Doktor Sistem Informasi Indonesia

The evaluation findings employing a dataset containing 70% training data ("Train 70") are displayed in Figure 8b. Model Performance: 95% mAP@50, 93.1% precision, and 88% recall. Although marginally lower than the model with 80% training data, the model's nevertheless high mAP value shows that it can predict bounding boxes reliably. A modest rise in False Positives (inaccurate predictions) is indicated by the accuracy decline relative to the 80% scenario. A rise in False Negatives is indicated by the decreased recall.

The model's capacity to identify all lesions is impacted by using 70% training data, which gives it less information to learn object properties than the 80% scenario. This evaluation result demonstrates that the YOLOv12 model's ability to detect lesions on different rice leaves is greatly impacted by a higher percentage of training data. Although the model's performance is marginally worse than in the 80% training data scenario, it can still detect lesions with accuracy and be utilized for automatic monitoring. This outcome demonstrates that, particularly for field applications, adequate training data is required to enhance accuracy (mAP), precision, and recall.

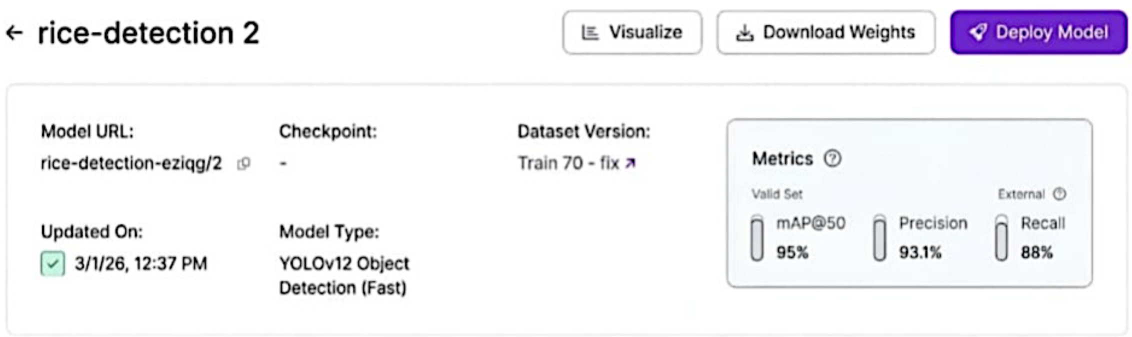

Figure 8b. Metric Evaluation Scenario 2

The evaluation findings employing a dataset containing 60% training data ("Train 60") are displayed in Figure 8c. Model Performance: mAP@50:91%, Recall: 91.5%, Precision: 91%. The model's inability to effectively anticipate bounding boxes is indicated by the lower mAP score when compared to the 70% and 80% situations. In comparison to models with more training data, precision equal to recall shows a drop in detection accuracy but a balance between True Positives and False Positives. A recall of 91.5% means that some lesions were found, although the probability of false negatives is slightly higher. In comparison to the 70% and 80% situations, using 60% training data gives the model less data variety to learn from, which lowers mAP and precision performance. This suggests

ISI Journal of Information Systems and Informatics
Published By Asosiasi Doktor Sistem Informasi Indonesia

that YOLOv12's capacity to identify lesions on rice leaves of different sizes, shapes, and orientations can be optimized with an adequate volume of training data. Although this model can be utilized for automatic monitoring, it is better to use a dataset with a greater training proportion (e.g., 80%) for maximum accuracy.

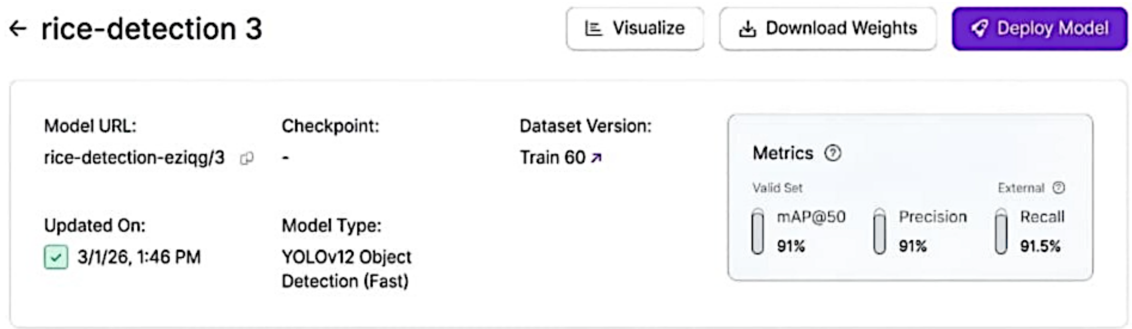

Figure 8c. Metric Evaluation Scenario 3

3.5 Comparison of Research Findings with Previous Rice Disease Detection Studies

YOLOv8n was used in earlier research with a dataset of 5,044. 4,032 photos were used for training, 506 images were used for validation, and 506 images were used for testing in the 80/10/10 scenario. The model's accuracy, recall, and mAP were 0.977, 0.990, and 0.995, respectively [25]. In contrast, this study only achieved a Recall of 96.5%, a precision of 96%, and mAP@50 utilizing a dataset of 4,744 pictures with Training: 3,795 images (80%), Validation: 474 images (10%), and Test: 475 images (10%). These findings show that while the model used in this study has strong detection skills, it is still not as good as the comparative model. The quantity of datasets employed in the study has a significant impact on this discrepancy.

A model's true robustness may not always be adequately shown by high performance metrics like high recall, precision, and mAP values. The dataset's features, including visual complexity, class distribution, similarity across disease classes, annotation quality, differences in lighting and background, and the clarity of symptoms in the photographs, can potentially have an impact on performance. Even with a few obstacles, the model can earn high scores if the dataset is reasonably simple and the classes are easily distinct. On the other hand, somewhat lower metric values may not always signify a poor model in more complicated and varied datasets.

ISI Journal of Information Systems and Informatics

Published By Asosiasi Doktor Sistem Informasi Indonesia

3.6 Mobile Application for Detecting Rice Leaf Diseases

This study has developed a mobile application for real-time disease detection on rice plant leaves, with a confidence level of 97%. Figure 9 displays the disease detection findings on several objects for each of the three types of diseases. The YOLOv12 Model can identify three different kinds of rice leaf illnesses in real time on mobile devices with a high confidence level (0.96-0.97), as shown in Figure 9a. These findings highlight the model's potential for field-based real-time disease monitoring applications that allow for quick lesion diagnosis on rice leaves. To verify the model's accuracy and stability in real-world agricultural activities, a number of tests were carried out using different kinds of leaves and in actual field settings.

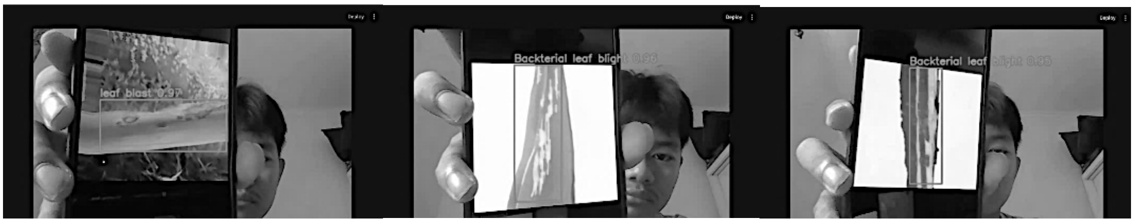

Figure 9a. Results of Leaf Blast disease detection

The YOLOv12 model can properly and consistently detect Brown Spot lesions on rice leaves, as shown by the various variations in the size and placement of the spots in Figure 9b. High confidence scores, which indicate accurate predictions, validate the model's promise for real-time monitoring of rice leaf diseases. More extensive field testing under different leaf situations is still required to guarantee the model's stability and dependability in real-world scenarios.

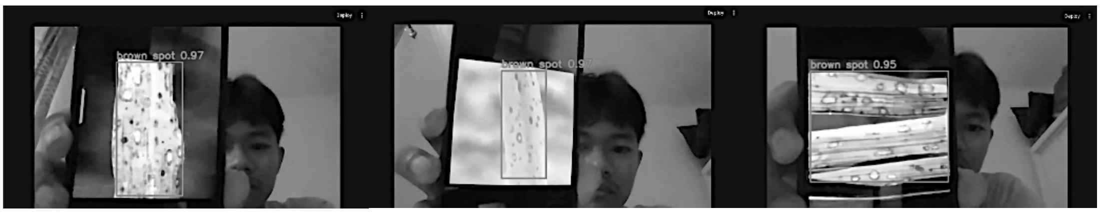

Figure 9b. Results of Brown Spot disease detection

Despite the lesions' varied locations, orientations, and sizes, the YOLOv12 model can reliably and precisely identify Leaf Blast lesions on rice leaves in picture 9c. The model's promise for real-time rice leaf disease monitoring is confirmed by a high confidence score, which denotes a trustworthy forecast.

ISI Journal of Information Systems and Informatics
Published By Asosiasi Doktor Sistem Informasi Indonesia

Figure 9c. Results of Leaf Blast disease detection

3.7 Discussion

The findings of this study demonstrate that the proposed YOLOv12-based rice leaf disease detection system is capable of identifying three important rice leaf diseases, namely Bacterial Leaf Blight, Brown Spot, and Leaf Blast, with high detection performance under field-oriented conditions. The dataset used in this study consisted of 4,744 images collected from rice fields in Jangan-jangan Village, Pujananting District, Barru Regency, which provided natural variations in leaf appearance, background complexity, shooting distance, and illumination conditions. This is an important aspect of the study because models trained on highly controlled laboratory images often perform well during testing but show reduced robustness when applied in real agricultural environments. By using field-based images and detailed object annotation through Roboflow, the model was exposed to more realistic disease patterns and environmental variability. As shown in Figure 3, the three disease classes exhibit different visual characteristics, but some symptoms may still overlap in terms of lesion shape, size, and color intensity, making automated detection a meaningful and challenging task.

The results of preprocessing and augmentation indicate that data preparation played a major role in improving the quality of model learning. The annotated images shown in Figure 4, Figure 5, and Figure 6 confirm that disease objects were localized using distinct bounding boxes that clearly separated lesion regions from the surrounding leaf area. This is particularly important in object detection tasks because the model does not only classify the disease category, but must also accurately determine the spatial position of the symptoms on the leaf surface. The augmentation strategies applied in this study, including resizing, normalization, random cropping, contrast adjustment, rotation, flipping, and Gaussian noise, appear to have strengthened the model's ability to handle different visual conditions. The model was able to detect lesions of varying sizes, positions, and orientations, and it also showed the capacity to identify multiple lesions on a single leaf. These findings suggest that augmentation improved the generalization ability of

ISI Journal of Information Systems and Informatics
Published By Asosiasi Doktor Sistem Informasi Indonesia

YOLOv12, especially for disease symptoms that appear small, irregular, or partially obscured in field images.

The training process also supports the suitability of YOLOv12 for this application. As illustrated in Figure 7, the model was initialized using pretrained weights, allowing it to begin training with previously learned visual features rather than starting from random parameter values. This transfer learning strategy likely contributed to faster convergence and more stable optimization, particularly because rice leaf disease symptoms can appear as subtle lesion patterns that require strong feature extraction capability. The selected training configuration, including the use of image size, batch size, epoch control, and early stopping, appears to have enabled the model to learn meaningful disease representations without excessive overfitting. The qualitative training outcome indicates that the convolutional layers were capable of extracting relevant visual patterns from the dataset and transforming them into reliable disease detection predictions. In the context of agricultural computer vision, this is significant because practical disease detection systems must balance accuracy, speed, and computational efficiency if they are to be deployed in real-world monitoring scenarios.

The comparison across the three dataset partitioning scenarios provides one of the most important findings of this study. Based on the evaluation results in Figure 8a, Figure 8b, and Figure 8c, the model performed best in Scenario 1, where 80% of the dataset was allocated for training. In this setting, the model achieved 96% precision, 96.5% recall, and 97% mAP@50, indicating a strong balance between accurate predictions and the ability to detect most actual disease lesions. These values show that the model produced relatively few false positives while also maintaining a low number of false negatives. When the proportion of training data was reduced to 70%, performance declined to 93.1% precision, 88% recall, and 95% mAP@50. The reduction became more apparent in the 60% training scenario, where the model achieved 91% precision, 91.5% recall, and 91% mAP@50. This pattern clearly indicates that the amount of training data strongly influences the learning capacity of YOLOv12, especially in tasks involving small and visually diverse disease objects. A larger training set provides a broader representation of symptom appearance, leaf texture variation, lesion orientation, and environmental disturbance, allowing the model to build a more robust internal representation of the target classes.

A closer interpretation of these metrics shows that the effect of training data is not limited to overall accuracy, but also directly affects the trade-off between false positives and false negatives. In Scenario 1, the high precision indicates that the model rarely assigned disease labels to irrelevant or healthy regions, while the high recall suggests that most lesions were successfully captured. In Scenario 2, the drop in recall to 88% is particularly meaningful because it implies that a greater number of actual disease lesions were missed. In agricultural practice, missed lesions can be more problematic than a small number of false alarms, especially during early disease development when timely intervention is critical to prevent further spread. Scenario 3 showed a more balanced precision and recall relationship, but the lower mAP still suggests weaker localization quality and reduced consistency in bounding box prediction. Taken together, these findings reinforce the idea that a higher proportion of training data improves both classification confidence and localization reliability, making the 80/10/10 split the most effective configuration among the three scenarios tested in this study.

When compared with previous research, the performance of the proposed model can be considered strong, although it still remains slightly below the benchmark reported in earlier work. A previous study using YOLOv8n and a dataset of 5,044 images under an 80/10/10 split reported 0.977 precision, 0.990 recall, and 0.995 mAP [25]. In contrast, the present study, using 4,744 images under the same data proportion scenario, achieved 96% precision, 96.5% recall, and 97% mAP@50. At first glance, this may suggest that the earlier model outperformed YOLOv12. However, this comparison should be interpreted carefully. Detection performance is influenced not only by model architecture, but also by dataset size, annotation consistency, class balance, lesion visibility, symptom similarity among classes, and the degree of visual complexity in the images. A dataset collected under more variable field conditions may produce slightly lower metric values while still representing a more realistic and challenging detection problem. Therefore, the difference in results does not necessarily indicate that the current model is weak; rather, it may reflect the greater heterogeneity and practical difficulty of the dataset used in this study.

This point is important because very high detection metrics do not always guarantee stronger real-world robustness. In plant disease detection, a model can obtain excellent precision, recall, and mAP when the images are visually clean, symptoms are large and

ISI Journal of Information Systems and Informatics
Published By Asosiasi Doktor Sistem Informasi Indonesia

clearly distinguishable, and the background is relatively simple. In contrast, field-acquired datasets often contain overlapping leaves, motion blur, nonuniform lighting, partial occlusion, and confusing background elements that make the task much more demanding. In such cases, slightly lower performance may still indicate that the model is functioning effectively under realistic deployment conditions. The present study therefore contributes valuable evidence that YOLOv12 can maintain high detection capability even when trained and evaluated on naturally variable rice field imagery. This strengthens the practical relevance of the model for precision agriculture applications, where the ability to operate reliably outside laboratory conditions is often more important than achieving ideal benchmark scores.

The mobile application results further confirm the practical potential of the proposed system. As shown in Figure 9, the trained YOLOv12 model was successfully integrated into a mobile-based real-time detection framework and was able to identify rice leaf diseases with confidence values in the range of 0.96–0.97. This is a promising outcome because it demonstrates that the model is not limited to offline experimental analysis, but can also be deployed in an operational interface that is useful for field users. The detection examples indicate that the model can recognize lesions under different leaf positions, scales, and image compositions, which is essential for mobile-based diagnosis where users may capture images from varying angles and distances. The practical implication is significant: farmers or agricultural officers can potentially use the application for rapid preliminary disease diagnosis, enabling faster decision-making regarding crop monitoring and treatment measures. Such a system could reduce reliance on manual visual inspection alone and support more responsive plant health management.

Even so, several limitations should be considered. The study focused on three disease classes only, while rice plants in real cultivation systems may also be affected by nutrient deficiencies, pest damage, mixed infections, or abiotic stress symptoms that resemble disease lesions. This means that future work should expand the dataset to include more symptom categories and more diverse field conditions in order to improve the model's discriminative ability. Another limitation concerns geographic scope, since the dataset was collected from one main location. Although the captured images already include realistic variability, broader multi-location sampling would strengthen the generalization capacity of the model across cultivars, climates, and agronomic practices. It would also be useful for future studies to compare YOLOv12 directly with other recent object detection architectures using the same dataset and experimental settings, so that architectural advantages can be evaluated more fairly. Additional validation on larger external datasets and longer-term field trials would further confirm the reliability of the mobile application in real agricultural operations.

4. CONCLUSION

This study demonstrates that the YOLOv12 model can be trained to identify three different types of leaf diseases: Bacterial Leaf Blight, Brown Spot, and Leaf Blast. The rice leaf dataset, which consists of 4,744 images from Jangan-jangan Village, Pujananting District, Barru Regency, was carefully annotated and varied through preprocessing and augmentation. The model demonstrated its capacity to properly identify lesions under controlled dataset settings by achieving mAP@50:97%, precision: 96%, and recall: 96.5% when trained with 80% of the training data using pretrained weights and optimal hyperparameters. These findings highlight how crucial it is to have a significant amount of training data, high-quality annotations, and sufficient image variety in order to support model performance. The model has been used in a mobile application demonstration for real-time disease detection, but further testing is required to guarantee the model's stability, generalization, and reliability before practical implementation in precision agriculture. This study is still restricted to a particular dataset and has not been thoroughly tested in real field conditions or across other datasets.

ACKNOWLEDGMENT

The author would like to sincerely thank everyone who has supported and contributed to this study. We would especially want to thank our staff for their advice and direction throughout the research process. We also thank Muhammadiyah University Makassar for providing the resources and facilities that made this research go smoothly. The author also thanks the Risetmu Team for their support, which allowed us to carry out this research. The author also expresses gratitude to friends and family for their encouragement and moral support during this research.

REFERENCES

[1] C. G. Simhadri, H. K. Kondaveeti, V. K. Vatsavayi, A. Mitra, and P. Ananthachari, "Deep learning for rice leaf disease detection: A systematic literature review on emerging trends, methodologies and techniques," *Inf. Process. Agric.*, vol. 12, no. 2, pp. 151–168, 2025, doi: 10.1016/j.inpa.2024.04.006.

[2] U. I. Ismail, H. N. Chua, R. Nordin, and M. K. Ahmed, "A comprehensive review of deep learning approaches for rice disease detection: Datasets, methodologies, and future directions," *Smart Agric. Technol.*, vol. 11, 2025, doi: 10.1016/j.atech.2025.100976.

[3] P. Pai *et al.*, "Deep learning-based automatic diagnosis of rice leaf diseases using ensemble CNN models," *Sci. Rep.*, vol. 15, no. 1, 2025, doi: 10.1038/s41598-025-13079-z.

[4] M. H. Bijoy *et al.*, "Towards Sustainable Agriculture: A Novel Approach for Rice Leaf Disease Detection Using dCNN and Enhanced Dataset," *IEEE Access*, vol. 12, pp. 34174–34191, 2024, doi: 10.1109/ACCESS.2024.3371511.

[5] J. Wang *et al.*, "Improved Lightweight YOLOv8 Model for Rice Disease Detection in Multi-Scale Scenarios," *Agronomy*, vol. 15, no. 2, 2025, doi: 10.3390/agronomy15020445.

[6] T. D. Bui and T. M. Do Le, "Ghost-Attention-YOLOv8: Enhancing Rice Leaf Disease Detection with Lightweight Feature Extraction and Advanced Attention Mechanisms," *AgriEngineering*, vol. 7, no. 4, 2025, doi: 10.3390/agriengineering7040093.

[7] L. E. S. C. M. 1 and odrigo O. C. C. 1, João D.S. Almeida1, Alexandre C. AraújoR1, "Detecting pancreatic masses on CT scans using YOLO architectures," doi: https://doi.org/10.1016/j.procs.2026.03.077.

[8] W. Gao, C. Zong, M. Wang, H. Zhang, and Y. Fang, "Intelligent identification of rice leaf disease based on YOLO V5-EFFICIENT," *Crop Prot.*, vol. 183, 2024, doi: 10.1016/j.cropro.2024.106758.

[9] Z. Xiuling, W. Huijuan, S. Yu, C. Gang, Z. Suhua, and Y. Quanbo, "Starting from the structure: A review of small object detection based on deep learning," *Image Vis. Comput.*, vol. 146, 2024, doi: 10.1016/j.imavis.2024.105054.

[10] H. Deng, S. Zhang, X. Wang, T. Han, and Y. Ye, "USD-YOLO: An Enhanced YOLO Algorithm for Small Object Detection in Unmanned Systems Perception," *Appl. Sci.*, vol. 15, no. 7, 2025, doi: 10.3390/app15073795.

[11] M. Nikouei *et al.*, "Small object detection: A comprehensive survey on challenges, techniques and real-world applications," *Intell. Syst. with Appl.*, vol. 27, 2025, doi: 10.1016/j.iswa.2025.200561.

[12] M. Xu, J. E. Park, J. Lee, J. Yang, and S. Yoon, "Plant disease recognition datasets in the age of deep learning: challenges and opportunities," *Front. Plant Sci.*, vol. 15, 2024, doi: 10.3389/fpls.2024.1452551.

[13] V. Singh, M. Pencina, A. J. Einstein, J. X. Liang, D. S. Berman, and P. Slomka, "Impact of train/test sample regimen on performance estimate stability of machine learning in cardiovascular imaging," *Sci. Rep.*, vol. 11, no. 1, pp. 1–8, 2021, doi: 10.1038/s41598-021-93651-5.

[14] B. Min, T. Kim, D. Shin, and D. Shin, "Data Augmentation Method for Plant Leaf Disease Recognition," *Appl. Sci.*, vol. 13, no. 3, 2023, doi: 10.3390/app13031465.

[15] A. Muhammad, Z. Salman, K. Lee, and D. Han, "Harnessing the power of diffusion models for plant disease image augmentation," *Front. Plant Sci.*, vol. 14, 2023, doi: 10.3389/fpls.2023.1280496.

[16] K. Antwi, K. E. Bennin, D. K. Pobi Asiedu, and B. Tekinerdogan, "On the application of image augmentation for plant disease detection: A systematic literature review," *Smart Agric. Technol.*, vol. 9, 2024, doi: 10.1016/j.atech.2024.100590.

[17] A. R. Muhammad, H. P. Utomo, P. Hidayatullah, and N. Syakrani, "Early Stopping Effectiveness for YOLOv4," *J. Inf. Syst. Eng. Bus. Intell.*, vol. 8, no. 1, pp. 11–20, 2022, doi: 10.20473/jisebi.8.1.11-20.

[18] M. L. Ali and Z. Zhang, "The YOLO Framework: A Comprehensive Review of Evolution, Applications, and Benchmarks in Object Detection," *Computers*, vol. 13, no. 12, 2024, doi: 10.3390/computers13120336.

[19] J. Leong, J. Zhao, B. Xue, W. Gibson, and M. Zhang, "Deep learning-based seabird detection in fisheries for seabird protection," *J. R. Soc. New Zeal.*, vol. 55, no. 6, pp. 2082–2102, 2025, doi: 10.1080/03036758.2025.2500998.

[20] Y. Tian *et al.*, "Development and evolution of YOLO in object detection: A survey," *Neurocomputing*, vol. 669, 2026, doi: 10.1016/j.neucom.2025.132436.

[21] M. A. M. Alhassan and E. Yılmaz, "Evaluating YOLOv4 and YOLOv5 for Enhanced Object Detection in UAV-Based Surveillance," *Processes*, vol. 13, no. 1, 2025, doi: 10.3390/pr13010254.

Journal of Information Systems and Informatics
Published By Asosiasi Doktor Sistem Informasi Indonesia

[22] N. Ma Muriyah, J. H. Sim, and A. Yulianto, "Evaluating YOLOv5 and YOLOv8: Advancements in Human Detection," *J. Inf. Syst. Informatics*, vol. 6, no. 4, pp. 2999–3015, 2024, doi: 10.51519/journalisi.v6i4.944.

[23] E. P. Silmina, T. Hardiani, and S. L. Mahfida, "The Effect of the Number of Classes on the Values Resulting from Evaluation Metrics in the YOLOv5 Model," *Int. J. Adv. Sci. Eng. Inf. Technol.*, vol. 15, no. 2, pp. 419–425, 2025, doi: 10.18517/ijaseit.15.2.20495.

[24] N. B. Džakula, R. Heriansyah, and F. Fadly, "Performance Evaluation of YOLOv10 and YOLOv11 on Blood Cell Object Detection Dataset," *Int. J. Adv. Artif. Intell. Mach. Learn.*, vol. 2, no. 2, pp. 95–103, 2025, doi: 10.58723/ijaaiml.v2i2.434.

[25] H. L. Gope, H. Fukai, F. M. Ruhad, and S. Barman, "Comparative analysis of YOLO models for green coffee bean detection and defect classification," *Sci. Rep.*, vol. 14, no. 1, 2024, doi: 10.1038/s41598-024-78598-7.